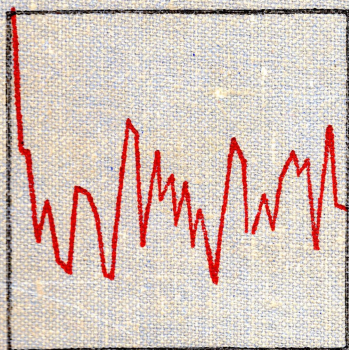


ДЖ. БОКС
Г. ДЖЕНКИНС

ВЫПУСК
1

АНАЛИЗ
ВРЕМЕННЫХ РЯДОВ
ПРОГНОЗ
И
УПРАВЛЕНИЕ





TIME SERIES ANALYSIS
Forecasting and control

by

GEORGE E. P. BOX
University of Wisconsin, U.S.A.

GWILYM M. JENKINS
University of Lancaster, U. K.

HOLDEN-DAY
San Francisco, Cambridge, London, Amsterdam
1970.

ДЖ. БОКС,
Г. ДЖЕНКИНС

АНАЛИЗ
ВРЕМЕННЫХ
РЯДОВ
ПРОГНОЗ
И
УПРАВЛЕНИЕ

Выпуск 1

Перевод с английского

А. Л. ЛЕВШИНА

Под редакцией

В. Ф. ПИСАРЕНКО

ИЗДАТЕЛЬСТВО «МИР»

МОСКВА 1974

В основу книги Бокса и Дженкинса положено использование данных о корреляционной функции (или функциях) одномерного и многомерного временных рядов. Особое внимание уделено нестационарным временным рядам, содержащим либо стационарные приращения, либо периодические нестационарности (что особенно важно для геофизических приложений). В первый выпуск вошли главы, содержащие основные сведения из корреляционной теории случайных процессов, выбор модели, оценивание ее параметров и проверку модели, а также модели для сезонных временных рядов.

Книга написана очень ясно и доступно; авторы, как правило, рассматривают конкретные примеры, доводимые до числовых результатов и позволяющие читателю научиться самостоятельно применять рекомендуемые методы. В конце книги приложены алгоритмы вычислений и таблицы используемых рядов. Книга будет весьма полезна специалистам по прикладной математике, геофизикам, физикам, астрономам, обработчикам данных наблюдений, экономистам, плановикам — всем лицам, встречающимся на практике с анализом и прогнозированием эмпирических величин, меняющихся со временем.

Редакция космических исследований, астрономии и геофизики

Предисловие к русскому изданию

Книга известных специалистов по статистике Дж. Бокса и Г. Дженкинса посвящена прикладным вопросам статистики случайных процессов. Один из авторов, Дженкинс, известен нашим читателям по недавно вышедшей книге [1], в которой подробно изложены вопросы статистического оценивания спектральной плотности стационарных процессов.

Настоящая книга посвящена применению параметрических моделей для различных задач, связанных с временными рядами, главным образом для задач прогнозирования и автоматического регулирования. Параметрические методы являются, вообще говоря, более эффективными по сравнению с непараметрическими (например, с обычным спектральным анализом), поскольку наблюдения при этом «расходятся» на оценивание меньшего числа параметров. Правда, параметрические модели описывают более узкий класс процессов, и их применение требует более полной априорной информации об изучаемых временных рядах.

Следует отметить, что если по спектральному анализу временных рядов с момента появления известной книги Блэкмана и Тьюки [2] вышло еще около десятка монографий, то по параметрическим методам для временных рядов книга Бокса и Дженкинса, по-видимому, первая. В силу этого она является, несомненно, оригинальной и представляет большой интерес для специалистов по случайным процессам.

Круг вопросов, затрагиваемых авторами, весьма широк. Сюда входят: выбор подходящей параметрической модели для изучаемого временного ряда и оценивание ее параметров; получение формул для прогноза ряда; оценивание передаточной функции линейного фильтра с помощью параметрической модели и изучение регулирующих схем.

Существенная особенность книги Бокса и Дженкинса состоит в том, что предлагаемый ими класс параметрических моделей содержит помимо традиционно рассматриваемых стационарных процессов и некоторые нестационарные процессы, а именно процессы со стационарными приращениями n -го порядка. Детальное математическое изучение этих процессов для случая, когда время изменяется непрерывно, было начато советскими математиками А. М. Ягломом и М. С. Пинскером [3, 4].

Однако практическое использование таких процессов до настоящего времени было не очень широким. Одна из причин — отсутствие литературы, в которой на доступном для прикладников уровне излагалась бы методика применения этих процессов к конкретным задачам практики. Книга Бокса и Дженкинса является первой и вполне успешной попыткой заполнить этот пробел.

Основной класс процессов, рассматриваемый в книге Бокса и Дженкинса, — это процессы z_t , удовлетворяющие следующей разностной схеме:

$$\begin{aligned} w_t &= \nabla^d z_t, & t &= 0, \pm 1, \pm 2, \dots, \\ \phi(B) w_t &= \theta(B) a_t, \end{aligned}$$

где ∇ — оператор взятия разности: $\nabla z_t = z_t - z_{t-1}$; B — оператор сдвига назад: $Bz_t = z_{t-1}$; $\phi(B)$, $\theta(B)$ — некоторые многочлены от B ; a_t — последовательность независимых одинаково распределенных случайных величин с конечной дисперсией (белый шум). Указанный класс процессов не охватывает всех процессов со стационарными n -ми приращениями, хотя и является достаточно широким. Он состоит из процессов, d -я разность которых является стационарным процессом с рациональной (относительно $e^{i\lambda}$, где λ — частота) спектральной плотностью. Отметим, что помимо указанного класса процессов авторы рассматривают еще и так называемые «сезонные процессы» (гл. 9), которые могут содержать периодические компоненты с известными периодами. При этом для перехода к стационарному процессу w_t с рациональной плотностью надо применить к процессу z_t вместо разностного оператора ∇^d некоторый другой разностный оператор, который уничтожит периодические компоненты.

Несмотря на то что обычные стационарные процессы с рациональной плотностью уже давно применяются в практических исследованиях (особенно в радиотехнике), эти применения, как правило, проводятся с помощью субъективных выводов без использования математической статистики. В книге Бокса и Дженкинса, напротив, видна постоянная тенденция к решению всех рассматриваемых задач на базе современной статистической теории оценивания и теории проверки статистических гипотез. Кроме того, как уже отмечалось, рассматривается более широкий класс процессов — процессы со стационарными n -ми приращениями (время дискретное).

Предлагаемые авторами практические методы выбора модели, оценивания коэффициентов этой модели и ее проверки

описаны подробно, с большим количеством иллюстраций и таблиц. При этом авторы стараются осветить и различные практические аспекты рассматриваемых задач. Из-за этого изложение порой может показаться несколько растянутым (особенно читателю с хорошей теоретической подготовкой). Однако легкость усвоения материала и его доступность от этого, несомненно, выигрывают.

В книге Бокса и Дженкинса (как и в упомянутой выше книге Дженкинса и Ваттса по спектральному анализу), как правило, нет строгих математических доказательств. Основной упор сделан на объяснение наглядного смысла предлагаемых методов и формул и на иллюстрацию их практического использования. Нередко вывод формул и утверждений основан на качественных, нестрогих рассуждениях. Это, по-видимому, неизбежно при рассмотрении практических ситуаций и возникающих при этом сложных, порой трудно формализуемых вопросов.

Хотя излагаемая авторами трактовка процессов со стационарными n -ми приращениями и не является математически строгой, получаемые ими результаты допускают строгое истолкование.

Так, например, процесс z_t можно рассматривать как обобщенный стационарный случайный процесс, определенный на подпространстве Φ_d основных функций $\varphi_d(t)$, у которых преобразование Фурье $\varphi_d(\lambda)$ имеет при $\lambda = 0$ нуль по крайней мере d -го порядка: $\varphi_d(0) = \varphi'_d(0) = \dots = \varphi_d^{(d-1)}(0) = 0$ (см., например, [4, 5]). Таким способом можно придать точный смысл представлению процесса z_t в проинтегрированной форме и в виде суммы бесконечного ряда независимых, одинаково распределенных случайных величин с коэффициентами, не стремящимися к нулю [см. гл. 4, формулы (4.1.8), (4.2.8) и далее]. При таком подходе распределения вероятностей могут быть определены не для значений процесса z_t , а лишь для линейных функционалов на Φ_d (например, конечных разностей порядка d).

Отметим, что хотя процесс z_t и является стационарным в обобщенном смысле, он, конечно, может и не быть стационарным в обычном смысле, и его реализации могут отражать эту «нестационарность» (например, значения реализации могут «уходить» от исходного уровня). В этом смысле оправдано использование авторами термина «нестационарный» по отношению к таким процессам. В то же время именно стационарность в обобщенном смысле позволяет использовать при прогнозировании методы, развитые для обычных стационарных процессов. В этой связи можно отметить, что проскальзывающая в книге мысль о том, что спектральные методы неприменимы к рассматриваемым процессам из-за их нестационарности, может несколько

дезориентировать читателя, поскольку все эти процессы имеют обобщенный спектр, обладающий при $\lambda = 0$ особенностью конечного степенного порядка. По-видимому, это недоразумение объясняется тем, что до сих пор нет практических методов измерения спектральных плотностей для процессов со стационарными n -ми приращениями.

Авторы с большим энтузиазмом пропагандируют предлагаемый ими параметрический подход для практических применений, поэтому вполне объясним их оптимизм в этом отношении. Однако читатель должен иметь в виду, что при изучении процессов, происходящих в природе и в обществе, вероятностные модели, естественно, не всегда применимы. Тем более, не всегда применима к встречающимся на практике временным рядам частная (хотя и достаточно широкая) модель, предложенная авторами. В связи с этим авторов можно было бы упрекнуть в отсутствии примеров, где предложенная ими модель не подходила бы. Думается, что такие отрицательные примеры были бы не менее поучительны для читателей, чем многочисленные примеры, иллюстрирующие успешное применение модели.

Книга Бокса и Дженкинса нашла положительный отклик за рубежом (см., например, [6]) и, несомненно, будет способствовать широкому практическому применению процессов с рациональным спектром в самых разных областях: геофизике, радиотехнике, экономике и т. д. Этому поможет и доступная, яркая форма изложения материала в книге. Отметим еще описание основных вычислительных алгоритмов, необходимых для использования предлагаемых методов. Эти алгоритмы могут служить основой для составления программ на ЭВМ, что также увеличивает практическую ценность книги Бокса и Дженкинса.

Для удобства читателей русское издание разделено на два выпуска. В первый выпуск вошли гл. 1—9, содержащие введение и краткую сводку излагаемых методов; основные сведения из корреляционной теории случайных процессов и процессов с рациональным спектром; теорию процессов со стационарными n -ми приращениями и с рациональным спектром, а также вывод формул для прогноза таких процессов; выбор модели, оценивание ее параметров и проверка модели; модели для сезонных временных рядов.

Второй выпуск будет содержать оценивание передаточных функций линейных фильтров, задачи автоматического управления в цепях с прямой и обратной связями, а также некоторые другие задачи теории регулирования и управления.

В. Ф. Писаренко

ЛИТЕРАТУРА

1. *Дженкинс Г., Баттс Д.*, Спектральный анализ и его приложения, вып. 1, изд-во «Мир», М., 1971, вып. 2, 1972.
2. *Blackman R. B., Tukey J. M.*, The measurement of power spectra from the point of view of communications engineering, Dover, New York, 1958.
3. *Яглом А. М., Пинскер М. С.*, Случайные процессы со стационарными приращениями n -го порядка, ДАН СССР, **90**, 385—388 (1953).
4. *Яглом А. М.*, Корреляционная теория процессов со случайными стационарными n -ми приращениями, Мат. сб., **37** (79), № 1, 141—196 (1956).
5. *Прохоров Ю. В., Розанов Ю. А.*, Теория вероятностей, Справочная матем. библиотека, изд-во «Наука», М., 1967.
6. *Hannan E. J.*, Math. Rev., **42**, № 5, 6019 (1971).

Предисловие

Большинство статистических методов имеет дело с моделями, в которых наблюдения предполагаются независимыми. Во многих приложениях зависимость между наблюдениями рассматривается как помеха; в планируемых экспериментах вводится *рандомизация* эксперимента. Это позволяет обосновать анализ, проводимый без учета зависимости наблюдений. Однако разнообразные данные в коммерции, экономике, технике и естественных науках поступают в форме *временных рядов*, в которых наблюдения зависимы и характер этой зависимости интересен сам по себе. Совокупность существующих методов анализа таких рядов зависимых наблюдений называется *анализом временных рядов*.

Спектральный анализ в частотной области образует один класс методов анализа временных рядов, но здесь мы почти не будем касаться этой важной темы. Основное содержание этой книги составляет построение стохастических (статистических) моделей для дискретных временных рядов во временной области и использование этих моделей в важнейших приложениях. Наша цель состоит в получении моделей, обладающих максимальной простотой и минимальным числом параметров и при этом адекватно описывающих наблюдения.

Получение таких моделей важно по следующим причинам:

1) они могут помочь нам понять природу системы, генерирующей временные ряды;

2) их можно использовать для *оптимального прогнозирования* будущих значений рядов;

3) если исследуются два или несколько связанных временных рядов, можно расширить модели так, чтобы они описывали динамические взаимосвязи между рядами и, следовательно, позволяли *оценивать передаточные функции*;

4) они могут быть использованы для выработки *стратегии оптимального управления*; эта стратегия указывает, каким образом надо изменять регулируемую переменную для того, чтобы минимизировать возмущение некоторой зависимой переменной.

Возможность оптимального прогнозирования, понимание динамических взаимосвязей и оптимальное управление имеют большое практическое значение. Например, прогнозирование оптимального сбыта нужно для планирования коммерческих опе-

раций; модели передаточных функций необходимы для улучшения проектирования и контроля заводских процессов; стратегии оптимального управления нужны для регулирования важных переменных этих процессов вручную или при помощи участвующих в процессе вычислительных машин.

Авторы более десяти лет имели дело с реальными данными, получаемыми в экономике и индустрии, и путем проб и ошибок, а также в результате продолжительных взаимодействий между теорией и практикой пытались выбирать, приспособливать и развивать практические методы решения таких задач. В результате этих работ и появилась эта книга.

Принятый здесь подход состоит, во-первых, в обсуждении класса моделей, обладающих достаточной гибкостью для описания практических ситуаций. В частности, временные ряды лучше всего описываются *нестационарными моделями*, в которых тренды и другие псевдоустойчивые характеристики, возможно меняющиеся во времени, рассматриваются скорее как статистические, а не детерминированные явления. Кроме того, временные ряды, связанные с экономикой или коммерцией, часто обладают заметными *сезонными*, или периодическими, компонентами; эти компоненты могут меняться во времени и должны описываться сезонными статистическими (возможно, нестационарными) моделями.

Далее рассматривается процесс *построения* моделей. Он заключается в установлении соответствия выбранного класса статистических моделей с имеющимися данными и существенно более сложен, чем простая подгонка модели. Поэтому вначале развиты методы *идентификации*, предназначенные для определения требуемого класса моделей; эти методы используют автокорреляционные и частные автокорреляционные функции. *Подгонка* идентифицированной модели к временному ряду при помощи функции правдоподобия может затем привести к оценкам максимального правдоподобия для параметров, или, если требуется, к байесовским апостериорным распределениям. Первоначально подогнанная модель не обязательно дает адекватное описание наблюдений. Поэтому разработан способ *диагностической проверки* для выявления неадекватности модели и выработки подходящих изменений; в случае необходимости может быть проведен еще один или несколько итеративных циклов идентификации, подгонки и диагностической проверки.

Если целью работы является прогноз, подогнанная статистическая модель используется для выработки оптимального прогнозирования при помощи простых рекуррентных вычислений. В частности, эта модель полностью определяет, будут ли прогнозируемые величины изменяться линейно, экспоненциально и т. д. Кроме того, подогнанная модель позволяет точно оценить,

как при прогнозе используются данные из прошлого, определить дисперсию ошибок прогноза и вычислить пределы, в которых с заданной вероятностью будут лежать будущие значения ряда.

Если модель включает в себя динамические соотношения, подобные итеративные циклы идентификации, подгонки и диагностической проверки позволяют получить подходящую стохастическую модель передаточной функции. В последней части книги стохастические модели передаточных функций, предложенные ранее, используются для построения регулирующих схем с прямой и обратной связями.

Приложения, данные в этой книге, ни в коей мере не исчерпывают всех возможностей, и мы надеемся, что приведенные примеры помогут читателю применить эти методы к решению собственных задач. В частности, разностные уравнения, используемые для представления передаточных функций и стохастических явлений, могут быть использованы как «кубики», которые, если их правильно собрать, годятся для описания разнообразных систем, встречающихся в технике, торговле и экономике. Более того, обсуждаемые и демонстрируемые в книге принципы построения моделей имеют весьма общие приложения.

ПЛАН КНИГИ

Эта книга состоит из следующих частей (звездочка у номера указывает разделы, которые при первом чтении можно опустить).

Введение и краткое содержание (гл. 1)

Эта глава содержит неформальное и крайне сжатое описание основных обсуждаемых проблем, определение и более полное объяснение которых даны позднее. Ее цель — дать общую картину областей, подлежащих последующему изучению, и изучающий, возможно, захочет возвращаться к ней по мере чтения последующих глав.

Часть I. Стохастические модели и основанное на них прогнозирование (гл. 2, 3, 4 и 5)

В главе 2 обсуждаются некоторые основные средства анализа временных рядов; затем в главах 3 и 4 вводится важный класс линейных стохастических моделей и обсуждаются их свойства. Сразу же после этого в главе 5 вводится прогнозирование. Такая последовательность удобна, поскольку форма оптимального прогнозирования вытекает из структуры стохастических моделей, обсуждаемой в главе 4.

Часть II. Построение стохастических моделей (гл. 6, 7, 8 и 9)

Эта часть книги описывает методику итеративного построения модели; здесь стохастические модели, введенные в части I, связываются с фактическими данными в форме временных рядов. В гл. 6—8 описаны соответственно процессы идентификации, оценивания и диагностической проверки моделей. Гл. 9 иллюстрирует весь процесс построения моделей; в ней показано, как объединить все эти идеи при построении сезонных моделей и как использовать эти модели для прогнозирования сезонных временных рядов.

Часть III. Построение модели передаточной функции (гл. 10 и 11)

В гл. 10 вводятся модели передаточных функций, связывающих выход системы с одним или несколькими ее входами.

В гл. 11 обсуждаются методы идентификации, оценивания и диагностической проверки модели передаточной функции в присутствии шума. Глава оканчивается описанием способов использования таких моделей в прогнозировании.

Часть IV. Проектирование дискретных схем регулирования (гл. 12 и 13)

В этих двух главах мы показываем, как ранее введенные стохастические модели и модели передаточных функций можно объединить при проектировании простых схем управления с прямой и обратной связями.

Первый вариант рукописи книги был завершен в 1965 г. и издан в 1966—1967 гг. как технические доклады № 72, 77, 79, 94, 95, 99, 103, 104, 116, 121 и 122 факультета статистики университета шт. Висконсин и № 1—4, 6—11 и 13 факультета системной техники университета в Ланкастере.

Июнь 1969 г.

Дж. Бокс, США

Г. М. Дженкинс, Великобритания

Введение и краткое содержание

В этой книге мы предложим методы построения, идентификации, подгонки и проверки моделей временных рядов и динамических систем. Эти методы будут удобны для дискретных систем с выборкой данных, т. е. таких систем, в которых возможность произвести наблюдение и предпринять регулирующие действия возникает через равные интервалы времени.

Мы продемонстрируем использование этих моделей временных рядов и динамических систем в трех важных прикладных областях.

1) *Прогнозирование* будущих значений временного ряда по его текущим и прошлым значениям.

2) Определение *передаточной* функции системы — определение динамической модели вход — выход; с помощью этой модели можно найти эффект на выходе инерционной системы по произвольно заданным рядам на входе.

3) Проектирование простых *регулирующих схем с прямой и обратной связью*, при помощи которых можно в максимально допустимых пределах компенсировать потенциальные отклонения выхода системы от желаемого номинала.

1.1. ТРИ ВАЖНЫЕ ПРАКТИЧЕСКИЕ ПРОБЛЕМЫ

1.1.1. Прогнозирование временных рядов

Использование доступных к моменту времени t наблюдений временного ряда для прогнозирования его значения в некоторый момент времени в будущем $t + l$ может явиться основой для а) планирования в экономике и торговле; б) планирования выпуска продукции; в) складского контроля и контроля выпуска; г) управления и оптимизации промышленных процессов. Как описано Холтом и др. [1], Брауном [2] и в монографии фирмы ICI по краткосрочному прогнозированию [3], существует необходимость в прогнозе вперед на интервал, называемый *временем упреждения* и зависящий от конкретной проблемы. Например, время упреждения в проблеме складского контроля определено

Гаррисоном [4] как период, начинающийся с момента передачи заказа на пополнение склада с завода и длящийся до тех пор, пока заказ не доставлен на склад.

Предполагается, что наблюдения доступны в *дискретные*, равноотстоящие моменты времени. Например, в проблеме прогнозирования сбыта z_t в текущем месяце t и сбыт z_{t-1} , z_{t-2} , ..., z_{t-3} в предыдущие месяцы могут быть использованы для прогноза сбыта с упреждением $l = 1, 2, 3, \dots, 12$ мес. Обозначим через $\hat{z}_t(l)$ сделанный в момент t прогноз сбыта z_{t+l} в некоторый момент $t+l$ в будущем, т. е. с *упреждением* l .

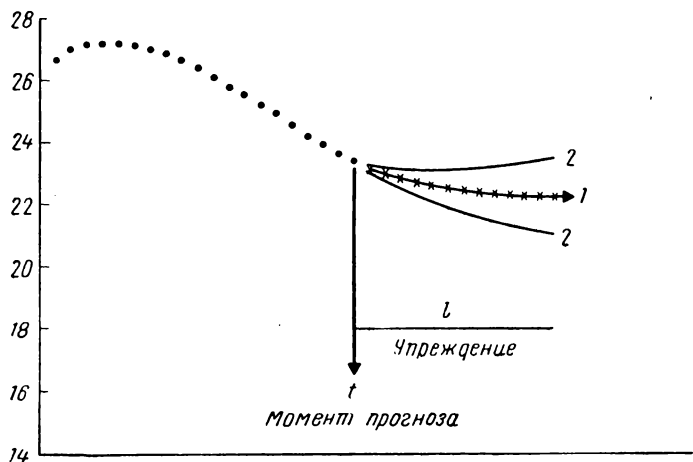


Рис. 1.1. Значения временного ряда (точки) вместе с прогнозирующей функцией (1) и 50%-ными вероятностными пределами (2).

Функция $\hat{z}_t(l)$, $l = 1, 2, \dots$, дающая в момент t прогнозы для всех будущих времен упреждения, будет называться *прогнозирующей функцией* в момент t . Наша цель — получить такую прогнозирующую функцию, у которой среднее значение квадрата отклонения $z_{t+l} - \hat{z}_t(l)$ истинного от прогнозируемого значения является наименьшим для каждого упреждения l .

В дополнение к вычислению наилучшего прогноза необходимо также указать его точность, чтобы, например, можно было оценить риск, связанный с решениями, основанными на прогнозировании. Точность прогноза может быть выражена *вероятностными пределами* по обе стороны от каждого прогнозируемого значения. Эти пределы можно вычислить для любого удобного набора вероятностей, например для 50 и 95%. Смысл этих пределов в том, что значение временного ряда, которое появится в соответствующее время, с указанной вероятностью будет лежать внутри этих пределов. Для иллюстрации на рис. 1.1 пока-

заны 20 последних значений временного ряда, обрывающегося на времени t . Там же показаны прогнозируемые величины на момент t для упреждений $l = 1, 2, \dots, 13$ вместе с 50%-ными вероятностными пределами.

Методы прогнозирования и оценки вероятностных пределов подробно обсуждаются в гл. 5.

1.1.2. Оценивание передаточных функций

Изучение динамики процессов представляет большой интерес для промышленности [5, 6]. Такие исследования проводятся, чтобы

- а) улучшить управление существующим производством,
- б) улучшить проектирование новых производств.

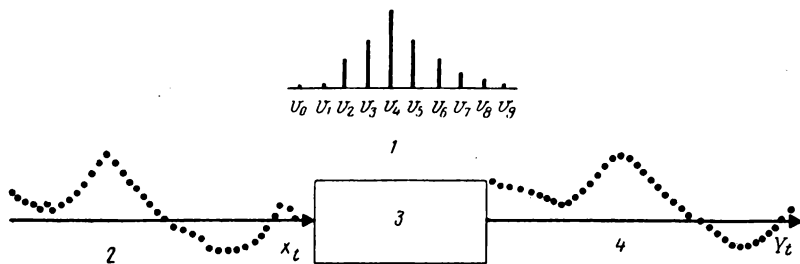


Рис. 1.2. Входной и выходной временные ряды в динамической системе. 1 — функция отклика, 2 — вход (скорость подачи воздуха), 3 — динамическая система (печь), 4 — выход (концентрация CO_2).

В частности, предложено несколько методов оценивания передаточных функций производственных единиц по записям процессов, состоящим из входных временных рядов X_t и выходных временных рядов Y_t . Участки таких записей показаны на рис. 1.2, где вход X_t — количество поступающего воздуха, а выход Y_t — концентрация двуокиси углерода, образующегося в печи. Наблюдения производились через 9-секундные интервалы. Гипотетическая функция отклика системы на единичный импульс, определяющая передаточную функцию системы, показана на рисунке вертикальными отрезками.

Классические методы оценивания модели передаточных функций, основанные на детерминированных возмущениях входа, таких, как дельта-импульс, ступенька или синусоида, не всегда дают хорошие результаты. Это связано с тем, что при разумных и допустимых по величине возмущениях реакция системы может быть замаскирована неконтролируемыми возмущениями, совокупность которых рассматривается как шум. Статистические методы оценивания моделей передаточных функций с учетом шума

описаны в главе 11. Оценка динамических реакций представляет значительный интерес в экономике, технике, биологии и многих других областях.

Проблема выбора формы статистического входа, обеспечивающего хорошую оценку передаточной функции при минимальном возмущении системы, кратко рассмотрена в главе 11. Во многих задачах хороший компромисс достигается использованием в качестве входного возмущения источника некоррелированных входных флуктуаций, например псевдослучайных двоичных последовательностей, используемых Хэммондом [7] и другими.

Другое важное приложение моделей передаточных функций— область прогнозирования. Если, например, можно определить динамическое соотношение между двумя временными рядами Y_t и X_t , то значения *обоих* рядов, принадлежащие прошлому, можно использовать для прогнозирования Y_t . В некоторых ситуациях этот подход может привести к значительному уменьшению ошибок в прогнозировании.

1.1.3. Проектирование дискретных регулирующих систем

В прошлом слово «control» (контроль, регулирование, управление) для статистиков означало обычно *методику контроля качества*, разработанную Шьюхартом [8] в США и Дадингом и Дженнетом [9] в Великобритании.

Недавно повысился интерес к применению принципов последовательного анализа для контроля качества, что привело к введению *диаграмм кумулятивных сумм* [10—12] и *диаграмм геометрического скользящего среднего* [13].

Слово «control» имеет другое значение для инженера по регулированию. Он подразумевает под ним контуры прямых и обратных регулирующих связей, вопросы динамики и устойчивости системы и конкретные устройства для выполнения регулирующих действий. Эти регулирующие устройства являются *автоматическими* в том смысле, что информация о процессе поступает к ним автоматически и информация на регулирование входов процесса автоматически выбирается из них.

В этой книге мы описываем статистический подход к прогнозированию временных рядов и к разработке регулирующих схем прямой и обратной связей, развитый в предшествующих статьях [14—20]. Обсуждаемые методы регулирования ближе к используемым инженерами по регулированию, чем к стандартным операциям по контролю качества, развитым статистиками. Это не означает, что мы считаем традиционные диаграммы контроля качества ненужными; просто они выполняют функции, которые мы не рассматриваем в этой работе. Важная функция стандарт-

ных диаграмм контроля качества — обеспечить способ непрерывного просмотра для обнаружения возможных причин отклонения. Удобное изображение данных о продукции позволяет лицам, управляющим процессом, быстро обнаружить изменения. Зная ответ на вопрос: «Когда произошли изменения этого типа?», мы можем спросить: «Почему они произошли?». Это создает непрерывно действующий стимул к улучшению процесса, часто ведущий к его переосмысливанию.

Напротив, регулирующие схемы, обсуждаемые в этой книге, предназначены для периодических, оптимальных регулировок управляющего переменного, влияние которого на некоторые характеристики качества выхода уже известно. Это позволяет уменьшить отклонения этой характеристики качества от некоторой желаемой величины.

Причина, по которой необходимо регулирование, заключается в неизбежном присутствии в процессе *возмущения* или *шума*. Когда мы можем измерять эти возмущения, возникает возможность провести необходимые компенсирующие изменения некоторых других переменных. Такой способ называется *регулированием с прямой связью*. Вместо этого или в дополнение к этому мы можем использовать отклонения от желаемого значения, или «сигнал ошибок» выходной характеристики, для вычисления компенсирующих изменений. Такой способ называется *регулированием с обратной связью*. В отличие от первого способа этот вид коррекции может быть использован даже тогда, когда источник возмущений не известен достаточно точно или величина возмущения не измерена. В общем плане может оказаться целесообразным использование регулировок как с прямой, так и с обратной связью; первая может быть использована для компенсации измеряемых возмущений, а вторая — для компенсации неизмеряемых.

Принятый здесь подход к регулированию заключается в выборе для описания возмущений соответствующего временного ряда, или *стохастической модели*, и для описания инерциальных характеристик системы — подходящей *модели передаточной функции*.

Тогда появляется возможность вывести *уравнение регулирования*. Это уравнение позволяет рассчитать необходимые действия в любой момент времени по заданным настоящему и предыдущему состояниям системы. Под «оптимальным» действием понимается такое, которое приводит к минимальной среднеквадратичной ошибке на выходе.

Выполнение регулирующего действия, указываемого уравнением регулирования, может быть достигнуто различными путями в зависимости от уровня технологии. В этой книге мы не будем указывать, какие средства должны быть использованы в

конкретных ситуациях. Однако стоит упомянуть, что на одном полюсе мы имеем регулирование, осуществляемое с применением ЭВМ; действие, указываемое уравнением регулирования, вычисляется на ЭВМ. Оно автоматически реализуется соответствующими датчиками, например, открывающими и закрывающими клапаны, регулирующие переменные процесса: температуру, давление, скорость потока. На промежуточном уровне существуют автоматические регуляторы, использующие для выполнения регулирующих действий различные пневматические и электрические устройства. На другом полюсе действие, предписываемое уравнением регулирования, определяется по соответствующей номограмме или диаграмме. Оператор периодически берет один или несколько отсчетов, наблюдений или определений, справляется по номограмме, какое действие необходимо предпринять, и затем вручную вводит необходимые изменения в процесс.

Даже в современной промышленности часто встречаются ситуации, где используется ручное регулирование, которое может быть улучшено. В гл. 12 кратко рассматривается подготовка диаграмм и номограмм, по которым в простых случаях можно определить оптимальные регулирующие действия. Например, верхняя диаграмма на рис. 1.3 показывает ряд измерений вязкости полимера, выполняемых с часовым интервалом в течение 42 ч. Вязкость должна поддерживаться вблизи значения 90 единиц. После каждого измерения вязкости оператор процесса использует номограмму, показанную в середине рисунка, для вычисления необходимой оптимальной регулировки управляемой переменной (подачи газа). На нижней диаграмме рис. 1.3 показана регулировка, сделанная в соответствии с номограммой.

На практике точная природа стохастических моделей и моделей передаточных функций, пригодных для любой конкретной ситуации и необходимых для разработки схем оптимального регулирования, обычно неизвестна. Далее, неизвестны соответствующие численные значения параметров этих моделей. Поэтому, как это почти всегда бывает при построении моделей, нам приходится использовать *итеративный* метод.

Любые доступные записи и дополнительная информация могут использоваться для разработок *пробной регулирующей схемы*, которая должна управлять процессом в течение некоторого периода. Данные рис. 1.3 были получены при реализации такой схемы. В гл. 12 показано, как по данным такой регулирующей схемы, состоящим из отклонений от требуемого значения характеристики качества и соответствующих регулировок, может быть проведен статистический анализ. В результате можно изменить модель и оценки параметров, что приведет к улучшению регулирующей схемы.

Хотя наши приложения часто имели отношения к ручным схемам регулирования, обсуждаемые методы могут быть с успехом использованы в процессах, управляемых вычислительной машиной, как описано, например, Офтоном [21]. При этом удается следить одновременно за большим числом регулируемых переменных, используя для каждой свой управляющий алгоритм. При помощи ЭВМ может быть также обеспечено взаимодействие

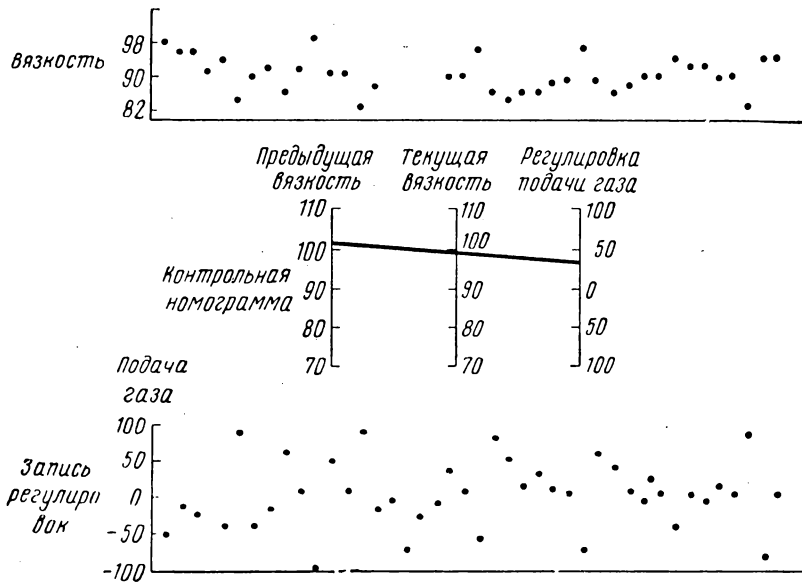


Рис. 1.3. Контроль вязкости. Запись наблюдаемой вязкости и регулировка подачи газа согласно номограмме.

между несколькими контурами управления. Многие современные схемы прямого цифрового управления (ПЦУ), использующие ЭВМ, основаны на замене стандартных одно-, двух- и трехступенчатых регуляторов соответствующими численными алгоритмами. Подход, который мы опишем здесь, может привести к более общим управляющим алгоритмам, чем те, которые применялись до сих пор, и более гибкому использованию цифровой вычислительной машины.

1.2. СТОХАСТИЧЕСКИЕ И ДЕТЕРМИНИРОВАННЫЕ ДИНАМИЧЕСКИЕ МАТЕМАТИЧЕСКИЕ МОДЕЛИ

Идея использования математических моделей для описания поведения физических объектов является общепризнанной. В частности, иногда удается получить модель, основанную на

физических законах, что дает возможность вычислить почти точное значение какой-либо зависящей от времени величины в любой момент времени. Например, мы можем вычислить траекторию ракеты, запущенной в известном направлении с известной скоростью. Если были бы возможны точные вычисления, такая модель была бы целиком *детерминированной*.

Вероятно, эта задача не является целиком детерминированной, потому что в ней может участвовать ряд неизвестных факторов, например переменная скорость ветра, которая может слегка отклонить ракету от курса. Во многих проблемах мы должны исследовать объекты, зависящие от времени, например, такие, как месячный сбыт газетной бумаги, в которых действует большое число неизвестных факторов. Для таких объектов нельзя предложить детерминированную модель, допускающую точное вычисление будущего поведения объекта. Тем не менее можно предложить модель, позволяющую вычислить *вероятность* того, что некоторое будущее значение будет лежать в определенном интервале. Такая модель называется вероятностной или *стохастической*. Модели временных рядов, необходимые для получения оптимального прогнозирования и регулирования, в действительности являются стохастическими. В дальнейшем необходимо различать вероятностную модель, или (как она иногда называется) *стохастический процесс*, и наблюдаемый временной ряд. Так, временной ряд z_1, z_2, \dots, z_N из N последовательных наблюдений рассматривается как выборочная реализация из бесконечной популяции таких временных рядов, генерируемых стохастическим процессом. Очень часто, говоря о стохастическом процессе, мы будем опускать слово «стохастический».

1.2.1. Стационарные и нестационарные стохастические модели для прогнозирования и регулирования

Важный класс стохастических моделей для описания временных рядов, который привлек большое внимание, — так называемые *стационарные* модели. Они основаны на предположении, что процесс остается в *равновесии* относительно *постоянного среднего уровня*. Однако в индустрии, коммерции и экономике, где прогнозирование имеет особо важное значение, многие временные ряды часто лучше описываются как *нестационарные* и, в частности, как не имеющие естественного среднего значения. Поэтому не удивительно, что методы экономического прогнозирования, предложенные Холтом [1, 22], Винтерсом [23], Брауном [2] и в монографии ICI [3], пользуются экспоненциально взвешенными скользящими средними; можно показать, что такой подход соответствует одному частному виду *нестационар-*

ного процесса. Хотя такие методы слишком частные, чтобы эффективно применяться ко всем временным рядам, тот факт, что они обеспечивают правильную прогнозирующую функцию, указывает на *характер нестационарной* модели, пригодной для этого класса проблем.

Стохастическая модель, для которой прогнозирование экспоненциально взвешенным скользящим средним является оптимальным, относится к классу *нестационарных* процессов, называемых процессами авторегрессии — проинтегрированного скользящего среднего (АРПСС), которые обсуждаются в гл. 4. Этот широкий класс процессов обеспечивает множество как стационарных, так и нестационарных моделей, которые адекватно описывают многие встречающиеся на практике временные ряды. Наш подход к прогнозированию заключается в отыскании вначале адекватной стохастической модели для исследуемого временного ряда. Как показано в гл. 5, когда подходящая модель ряда найдена, с ее помощью немедленно находится оптимальная прогнозирующая методика. Частным случаем таких методик является прогнозирование экспоненциально взвешенным скользящим средним.

Некоторые простые операторы. Мы будем широко пользоваться оператором сдвига назад B , определяемым как $Bz_t = z_{t-1}$; отсюда $B^m z_t = z_{t-m}$. Обратная операция выполняется оператором сдвига вперед $F = B^{-1}$, задаваемым как $Fz_t = z_{t+1}$; следовательно, $F^m z_t = z_{t+m}$. Еще один важный оператор — *разностный оператор со сдвигом назад* ∇ , который можно выразить через B как

$$\nabla z_t = z_t - z_{t-1} = (1 - B) z_t.$$

В свою очередь оператор, обратный ∇ , — это оператор суммирования S , выражаемый как

$$\begin{aligned} \nabla^{-1} z_t &= Sz_t = \sum_{j=0}^{\infty} z_{t-j} = \\ &= z_t + z_{t-1} + z_{t-2} + \dots = (1 + B + B^2 + \dots) z_t = \\ &= (1 - B)^{-1} z_t. \end{aligned}$$

Модель линейного фильтра. Применяемые нами стохастические модели используют тот факт [24], что временные ряды, в которых последовательные значения сильно зависимы, целесообразно рассматривать как генерируемые последовательностью *независимых* импульсов a_t . Эти импульсы — реализации *случайных* величин с фиксированным распределением, которое обычно предполагается нормальным с нулевым средним и дисперсией σ_a^2 . Такая последовательность случайных величин $a_t, a_{t-1}, a_{t-2}, \dots$ называется в технической литературе *белым шумом*.

Предполагается, что белый шум a_t можно трансформировать в процесс z_t при помощи *линейного* фильтра, показанного на рис. 1.4а. Операция линейной фильтрации заключается в вычислении взвешенной суммы предыдущих наблюдений, так что

$$z_t = \mu + a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = \mu + \psi(B) a_t. \quad (1.2.1)$$

В общем μ — параметр, определяющий «уровень» процесса, и

$$\psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \dots$$

— линейный оператор, преобразующий a_t в z_t и называемый *передаточной функцией* фильтра.

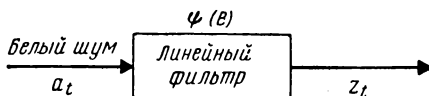


Рис. 1.4а. Представление временного ряда как выхода линейного фильтра.

Последовательность ψ_1, ψ_2, \dots , образованная весами, теоретически может быть конечной или бесконечной. Если эта последовательность (конечная или бесконечная) сходящаяся, фильтр называется *устойчивым*, а процесс z_t будет стационарным. Параметр μ будет тогда средним, вокруг которого процесс варьирует. В других случаях z_t нестационарен, и μ не имеет какого-либо особого смысла, кроме как некой точки отсчета уровня процесса.

Модели авторегрессии. Так называемая модель *авторегрессии* является исключительно полезной стохастической моделью для описания некоторых встречающихся на практике рядов. В этой модели текущее значение процесса выражается как конечная линейная совокупность *предыдущих значений процесса* и импульса a_t . Обозначим значения процесса в равноотстоящие моменты времени $t, t-1, t-2, \dots$ как $z_t, z_{t-1}, z_{t-2}, \dots$. Пусть $\tilde{z}_t, \tilde{z}_{t-1}, \tilde{z}_{t-2}, \dots$ будут отклонениями от μ , например $\tilde{z}_t = z_t - \mu$. Тогда

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \dots + \phi_p \tilde{z}_{t-p} + a_t \quad (1.2.2)$$

называется *процессом авторегрессии (АР) порядка p* . Такое название объясняется тем, что линейная модель

$$\tilde{z} = \phi_1 \tilde{x}_1 + \phi_2 \tilde{x}_2 + \dots + \phi_p \tilde{x}_p + a,$$

связывающая «зависимое» переменное z с множеством «независимых» переменных x_1, x_2, \dots, x_p , плюс член a , описывающий

ошибку, часто называется моделью *регрессии*; говорят, что z «регрессирует» на x_1, x_2, \dots, x_p . В (1.2.2) переменная z регрессирует на своих предшествующих значениях; поэтому модель *авторегрессирующая*. Если мы определим *оператор авторегрессии* порядка p как

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p,$$

то модель авторегрессии можно сжато описать как

$$\phi(B) \tilde{z}_t = a_t.$$

Эта модель содержит $p + 2$ неизвестных параметра: $\mu, \phi_1, \phi_2, \dots, \phi_p, \sigma_a^2$, которые на практике следует оценить по наблюдениям. Дополнительный параметр σ_a^2 — дисперсия белого шума a_t .

Нетрудно заметить, что модель авторегрессии является частным видом модели линейного фильтра (1.2.1). Например, \tilde{z}_{t-1} можно исключить из правой части (1.2.2) подстановкой

$$\tilde{z}_{t-1} = \phi_1 \tilde{z}_{t-2} + \phi_2 \tilde{z}_{t-3} + \dots + \phi_p \tilde{z}_{t-p-1} + a_{t-1}.$$

Аналогичным образом можно исключить \tilde{z}_{t-2} и т. д., получив в результате бесконечный ряд из a .

Символически это можно записать как

$$\phi(B) \tilde{z}_t = a_t,$$

что эквивалентно

$$\tilde{z}_t = \psi(B) a_t,$$

где

$$\psi(B) = \phi^{-1}(B).$$

Процессы авторегрессии могут быть стационарными или нестационарными. Чтобы процесс был стационарным, нужно выбрать ϕ так, чтобы веса ψ_1, ψ_2, \dots в $\psi(B) = \phi^{-1}(B)$ образовывали сходящийся ряд. Эти модели будут подробно обсуждаться в гл. 3 и 4.

Модели скользящего среднего. Модель авторегрессии (1.2.2) выражает отклонение \tilde{z}_t процесса в виде *конечной* взвешенной суммы p предыдущих отклонений процесса $\tilde{z}_{t-1}, \tilde{z}_{t-2}, \dots, \tilde{z}_{t-p}$ плюс случайный импульс a_t . Равным образом, как было только что показано, она выражает \tilde{z}_t как *бесконечную* взвешенную сумму a .

Другой тип моделей, имеющий большое значение в описании наблюдаемых временных рядов, — это так называемый процесс *скользящего среднего*. Пусть \tilde{z}_t линейно зависит от *конечного*

числа q предыдущих a . Такой процесс

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (1.2.3)$$

называется *процессом скользящего среднего (СС) порядка q* . Название «скользящее среднее» слегка вводит в заблуждение, так как веса $1, -\theta_1, -\theta_2, \dots, -\theta_q$, на которые умножаются a , не обязаны давать в сумме единицу или хотя бы быть положительными. Однако из-за общеупотребительности этого термина мы будем его придерживаться.

Если мы определим *оператор скользящего среднего* порядка q как

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q,$$

то модель скользящего среднего можно сжато записать как

$$\tilde{z}_t = \theta(B) a_t.$$

Она содержит $q + 2$ неизвестных параметра: $\mu, \theta_1, \dots, \theta_q, \sigma_a^2$, которые должны на практике оцениваться по наблюдениям.

Смешанные модели авторегрессии — скользящего среднего. Для достижения большей гибкости в подгонке моделей к наблюдаемым временным рядам иногда целесообразно объединить в одной модели и авторегрессию, и скользящее среднее. Это приводит к комбинированной модели авторегрессии — скользящего среднего

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \dots + \phi_p \tilde{z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}, \quad (1.2.4)$$

или

$$\phi(B) \tilde{z}_t = \theta(B) a_t,$$

в которой имеется $p + q + 2$ неизвестных параметра: $\mu; \phi_1, \dots, \phi_p; \theta_1, \dots, \theta_q; \sigma_a^2$, оцениваемых по наблюдениям.

На практике часто оказывается, что адекватное описание наблюдаемых временных рядов достигается при помощи моделей авторегрессии, скользящего среднего или комбинированной модели, в которых p и q не больше, а часто и меньше 2.

Нестационарные модели. Многие ряды, практически встречающиеся в промышленности или торговле (например, биржевые цены), обнаруживают нестационарный характер и, в частности, не колеблются относительно фиксированного среднего. Тем не менее их свойства могут быть в некотором смысле однородными. Например, хотя уровень, относительно которого происходят флуктуации, может быть разным в разные времена, поведение рядов с учетом различий в уровне оказывается во многом сходным. Мы покажем в гл. 4, что такой ряд может быть

представлен обобщенным оператором авторегрессии $\varphi(B)$, в котором один или несколько нулей полинома $\varphi(B)$ [т. е. один или несколько корней уравнения $\varphi(B) = 0$] равны единице. Тогда оператор $\varphi(B)$ можно записать

$$\varphi(B) = \phi(B)(1 - B)^d,$$

где $\phi(B)$ — стационарный оператор. При этом обобщенная модель, описывающая однородный нестационарный процесс, имеет вид

$$\varphi(B)z_t = \phi(B)(1 - B)^d z_t = \theta(B)a_t,$$

т. е.

$$\phi(B)\omega_t = \theta(B)a_t, \quad (1.2.5)$$

где

$$\omega_t = \nabla^d z_t. \quad (1.2.6)$$

Таким образом, однородный нестационарный процесс может быть описан моделью, которая требует, чтобы d -я разность процесса была стационарной. На практике d обычно равно 0, 1 или максимум 2.

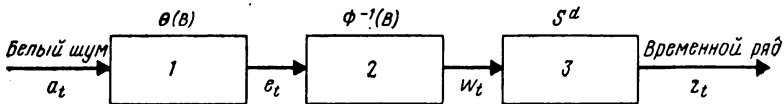


Рис. 1.46. Диаграмма модели авторегрессии — проинтегрированного скользящего среднего.

1 — фильтр скользящего среднего, 2 — стационарный фильтр авторегрессии, 3 — нестационарный фильтр суммирования.

Процесс, определенный (1.2.5) или (1.2.6), представляет собой эффективную модель для описания стационарных и нестационарных временных рядов и называется процессом *авторегрессии — проинтегрированного скользящего среднего* (АРПСС) порядка (p, d, q) . Процесс определен как

$$\omega_t = \phi_1 \omega_{t-1} + \dots + \phi_p \omega_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (1.2.7)$$

с $\omega_t = \nabla^d z_t$. Заметим, что если заменить ω_t на z_t — при $d = 0$, модель (1.2.7) будет содержать как частный случай и стационарную комбинированную модель (1.2.4), а также модели авторегрессии (1.2.2) и скользящего среднего (1.2.3).

Причина для включения в название АРПСС слова «проинтегрированный» (более точно, вероятно, было бы говорить «сум-

мированный») заключается в следующем. Соотношение, обратное (1.2.6), имеет вид

$$z_t = S^d w_t; \quad (1.2.8)$$

напомним, что S — оператор суммирования, определенный как

$$S w_t = \sum_{j=0}^{\infty} w_{t-j} = w_t + w_{t-1} + w_{t-2} + \dots$$

Следовательно, обобщенный процесс авторегрессии — проинтегрированного скользящего среднего (АРПСС) можно получить из белого шума a_t с помощью трех операций фильтрации, как показано на рис. 1.46. Первый фильтр имеет вход a_t , передаточную функцию $\theta(B)$ и выход e_t , где

$$e_t = a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} = \theta(B) a_t. \quad (1.2.9)$$

Второй фильтр имеет вход e_t , передаточную функцию $\phi^{-1}(B)$ и выход w_t , где

$$w_t = \phi_1 w_{t-1} + \dots + \phi_p w_{t-p} + e_t = \phi^{-1}(B) e_t. \quad (1.2.10)$$

Наконец, третий фильтр, согласно (1.2.8), имеет вход w_t и выход z_t с передаточной функцией S^d .

Как указано в гл. 9, для описания сезонных временных рядов можно использовать специальную форму модели (1.2.7).

1.2.2. Модели передаточных функций

Важный тип динамического соотношения между непрерывными входом и выходом, для которого можно найти много физических примеров, — это такой тип, у которого *отклонения* входа X и выхода Y от соответствующих средних значений связаны линейным дифференциальным уравнением вида

$$\begin{aligned} (1 + \Xi_1 D + \dots + \Xi_R D^R) Y(t) = \\ = (H_0 + H_1 D + \dots + H_S D^S) X(t - \tau), \end{aligned} \quad (1.2.11)$$

где D — дифференциальный оператор d/dt , Ξ и H — неизвестные параметры и τ — параметр, измеряющий *холостое время*, или *чистое запаздывание*, выхода относительно входа. Простейшим примером (1.2.11) является система, у которой скорость изменения выхода пропорциональна разности между входом и выходом, т. е.

$$\Xi \frac{dY}{dt} = X - Y,$$

и отсюда

$$(1 + \Xi D) Y = X.$$

Аналогично для дискретных данных в гл. 10 мы описываем систему, где вход X и выход Y , измеряемые через равные интервалы времени, связаны разностным уравнением

$$(1 + \xi_1 \nabla + \dots + \xi_r \nabla^r) Y_t = (\eta_0 + \eta_1 \nabla + \dots + \eta_s \nabla^s) X_{t-b}, \quad (1.2.12)$$

в котором дифференциальный оператор D заменен разностным оператором ∇ . Выражение вида (1.2.12), содержащее малое число параметров ($r \leq 2, s \leq 2$), часто можно использовать как аппроксимацию динамического соотношения более сложной природы.

Линейную модель (1.2.12) можно эквивалентным образом описать с помощью прошлых значений входа и выхода, подставив в (1.2.12) $\nabla = 1 - B$

$$\begin{aligned} (1 - \delta_1 B - \dots - \delta_r B^r) Y_t &= (\omega_0 - \omega_1 B - \dots - \omega_s B^s) X_{t-b} = \\ &= (\omega_0 B^b - \omega_1 B^{b+1} - \dots - \omega_s B^{b+s}) X_t, \end{aligned} \quad (1.2.13)$$

или

$$\delta(B) Y_t = \omega(B) B^b X_t = \Omega(B) X_t.$$

Другими словами, выход Y_t и вход X_t связаны линейным фильтром

$$Y_t = v_0 X_t + v_1 X_{t-1} + v_2 X_{t-2} + \dots = v(B) X_t, \quad (1.2.14)$$

передаточная функция которого

$$v(B) = v_0 + v_1 B + v_2 B^2 + \dots \quad (1.2.15)$$

может быть выражена как отношение двух полиномов

$$v(B) = \Omega(B) / \delta(B) = \delta^{-1}(B) \Omega(B).$$

Линейный фильтр (1.2.14) называют устойчивым, если ряд (1.2.15) сходится при $|B| \leq 1$. Ряд весов v_0, v_1, v_2, \dots , появляющихся в передаточной функции (1.2.15), называется *функцией отклика на единичный импульс*. Заметим, что для модели (1.2.12) первые b весов v_0, v_1, \dots, v_{b-1} равны нулю. Гипотетическая функция отклика для системы, показанной на рис. 1.2, изображена в центре этого рисунка.

Модель передаточной функции (1.2.13) позволяет иначе интерпретировать стохастические модели (1.2.4) и (1.2.5). Часто возмущения выхода z вызываются возмущениями некоторой переменной, с которой z динамически связано уравнением вида (1.2.12). Поэтому можно ожидать, что сложное стохастическое поведение случайной переменной z_t может быть выражено через другую случайную переменную a_t с более простыми свойствами соотношением

$$\delta(B) \tilde{z}_t = \Omega(B) a_t.$$

Если допустить возможность существования неустойчивого фильтра, у которого один или более корней уравнения $\delta(B) = 0$ равны единице, то, пользуясь ранее введенными обозначениями, можно получить

$$\varphi(B) \tilde{z}_t = \theta(B) a_t. \quad (1.2.16)$$

Стохастические модели, рассмотренные выше, как раз принадлежат к этому типу, причем a_t — источник *белого шума*. Поскольку (1.2.16) можно записать в виде

$$\tilde{z}_t = \varphi^{-1}(B) \theta(B) a_t,$$

то считается, что \tilde{z}_t можно получить пропусканием белого шума через линейный фильтр с передаточной функцией $\varphi^{-1}(B)\theta(B)$.

Выводы

1) Мы часто можем описать динамическое соотношение выхода Y и входа X с помощью линейного фильтра

$$Y_t = v_0 X_t + v_1 X_{t-1} + v_2 X_{t-2} + \dots = v(B) X_t,$$

где $v(B)$ — передаточная функция фильтра.

2) В свою очередь $v(B)$ часто можно компактно и достаточно точно представить в виде отношения двух полиномов малых степеней от B :

$$v(B) = \delta^{-1}(B) \Omega(B),$$

так что динамическое уравнение вход — выход можно записать как

$$\delta(B) Y_t = \Omega(B) X_t.$$

3) Мы будем считать, что ряд z_t с сильно зависимыми последовательными значениями может быть представлен как результат пропускания белого шума a_t через динамическую систему,

в которой отдельные корни уравнения $\delta(B) = 0$ могут быть единицами. Это позволяет получить модель процесса авторегрессии — проинтегрированного скользящего среднего

$$\varphi(B) z_t = \theta(B) a_t.$$

Модели с наложенным шумом. Мы видели, что задача оценивания подходящей модели, связывающей выход Y_t и вход X_t , эквивалентна оцениванию передаточной функции $v(B) = \delta^{-1}(B)\Omega(B)$. Однако эта задача на практике усложняется присутствием шума N_t , искажающего истинное соотношение между входом и выходом следующим образом:

$$Y_t = v(B) X_t + N_t,$$

где N_t и X_t независимы. Положим, как показано на рис. 1.5, что шум N_t может быть описан нестационарной стохастической

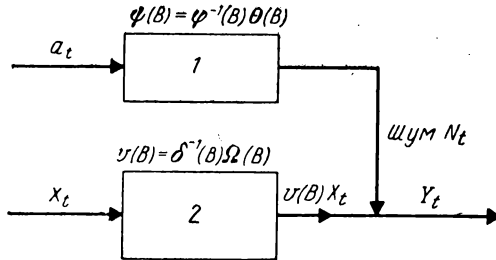


Рис. 1.5. Модель передаточной функции динамической системы с наложенным шумом.

1 — линейный фильтр, 2 — линейная динамическая система.

моделью вида (1.2.5) или (1.2.7), т. е.

$$N_t = \psi(B) a_t = \varphi^{-1}(B) \theta(B) a_t.$$

Тогда наблюдаемое соотношение между входом и выходом будет иметь вид

$$Y_t = v(B) X_t + \psi(B) a_t = \delta^{-1}(B) \Omega(B) X_t + \varphi^{-1}(B) \theta(B) a_t. \quad (1.2.17)$$

На практике необходимо оценить передаточную функцию $\psi(B) = \varphi^{-1}(B) \theta(B)$ линейного фильтра, описывающего шум, в дополнение к передаточной функции $v(B) = \delta^{-1}(B) \Omega(B)$, описывающей динамическое соотношение между входом и выходом. Методы получения таких оценок обсуждаются в гл. 11.

1.2.3. Модели дискретных систем регулирования

Как указывалось в разд. 1.1.3, регулирование — это попытка компенсировать возмущения, вносимые в систему. Некоторые из этих возмущений доступны для измерений, другие недоступны

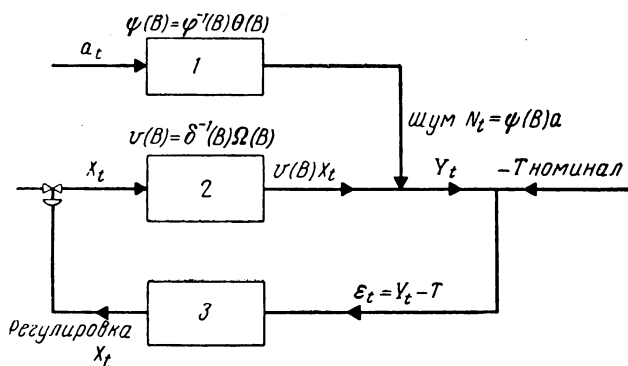


Рис. 1.6. Схема регулирующей обратной связи для компенсации неизмеряемых возмущений N_t .

1 — линейный фильтр, 2 — контролируемый процесс, 3 — регулятор.

и проявляют себя только как необъясненные отклонения от желаемого значения управляемой характеристики.

Для иллюстрации основных принципов регулирования рассмотрим один частный случай, когда на выход Y_t системы влияют недоступные для измерений возмущения. Предположим, что для поддержания выхода максимально близким к желаемому значению используется регулирование с обратной связью, осуществляющее коррекцию входного переменного X_t . Это иллюстрируется блок-схемой на рис. 1.6. Пусть N_t — влияние на выход различных неидентифицированных возмущений внутри системы; в отсутствие регулирования оно может вызывать отклонения выхода от желаемого значения, или *номинала*, T . Тогда, несмотря на коррекцию, вносимую в процесс, будет существовать ошибка

$$\epsilon_t = Y_t - T = v(B) X_t + N_t - T,$$

дающая отклонение выхода от номинала T . Цель заключается в таком выборе уравнения регулирования, чтобы ошибки ϵ были бы наименьшими в среднеквадратичном смысле. Уравнение ре-

гулирования описывает коррекцию $x_t = X_t - X_{t-1}$, предпринимаемую в момент t , как функцию текущего отклонения ε_t , предыдущих отклонений ε_{t-1} , ε_{t-2} , ... и предыдущих коррекций x_{t-1} , x_{t-2} Механизм (ручной, электрический, пневматический или электронный), выполняющий регулирующее действие, требуемое уравнением регулирования, называется *регулятором*.

Оказывается, что процедура проектирования оптимальной схемы эквивалентна прогнозированию отклонений от номинала, происходящих в *отсутствие регулирования*, и затем вычислению коррекции, необходимой для ликвидации отклонения. Следовательно, прогнозирование и проблемы регулирования тесно связаны. Для прогноза отклонений от номинала, происходящих в отсутствие регулирования, необходимо построить модель возмущения

$$N_t = \psi(B) a_t = \varphi^{-1}(B) \theta(B) a_t.$$

Вычисление коррекции x_t входного переменного в момент t , необходимой для ликвидации предсказанного изменения выхода, требует построения динамической модели с передаточной функцией

$$v(B) = \delta^{-1}(B) \Omega(B),$$

связывающей вход и выход. Результирующая коррекция x_t будет в общем состоять из линейной комбинации предыдущих коррекций, текущей и прошлых ошибок регулирования. Уравнение регулирования будет иметь вид

$$x_t = \zeta_1 x_{t-1} + \zeta_2 x_{t-2} + \dots + \chi_0 \varepsilon_t + \chi_1 \varepsilon_{t-1} + \chi_2 \varepsilon_{t-2} + \dots, \quad (1.2.18)$$

где $\zeta_1, \zeta_2, \dots, \chi_0, \chi_1, \chi_2, \dots$ — постоянные. Эти идеи обсуждаются в гл. 12.

1.3. ОСНОВНЫЕ ПОНЯТИЯ В ПОСТРОЕНИИ МОДЕЛЕЙ

1.3.1. Экономия

Мы видели, что математические модели, которые нам нужно использовать, содержат определенные константы или параметры, значения которых должны оцениваться по наблюдениям. На практике существенно, чтобы мы использовали *наименьшее возможное* число параметров для адекватного представления модели. Значение этого принципа *экономии* (parsimony) [25] в использовании параметров станет яснее в дальнейшем. В качестве предварительной иллюстрации мы рассмотрим следующий простой пример.

Положим, мы подогнали динамическую модель (1.2.12) вида

$$Y_t = (\eta_0 + \eta_1 \nabla + \eta_2 \nabla^2 + \dots + \eta_s \nabla^s) X_t, \quad (1.3.1)$$

работая с системой, адекватно описываемой следующим соотношением:

$$(1 + \xi \nabla) Y_t = X_t. \quad (1.3.2)$$

Модель (1.3.2) содержит только один параметр ξ , но для достаточно большого s может быть приближенно представлена моделью (1.3.1). Из-за присутствия экспериментальных ошибок мы можем не распознать правильное соотношение между коэффициентами подгоняемого уравнения. Это значит, что мы можем без нужды подогнать модель соотношением, содержащим s параметров, в то время как гораздо более простая форма (1.3.2), содержащая только один параметр, была бы вполне адекватной. Это может, например, привести к тому, что точность оценок выхода Y_t по данным значениям входа X_t, X_{t-1}, \dots будет далека от достижимой.

Наша цель состоит в получении не только адекватной, но и экономичной модели. Прогнозирование и регулирующие процедуры могут быть весьма неэффективными, если эти модели не адекватны или излишне расточительны в использовании параметров. Поэтому выбор модели требует осторожности и больших усилий. Процесс выбора по необходимости итеративен; иначе говоря, это процесс эволюции, адаптации, или проб и ошибок.

1.3.2. Итеративные этапы в выборе модели

Если физический механизм явления полностью ясен, можно вывести математическое выражение, точно описывающее явление. Тогда мы получаем *механистическую*, или *теоретическую*, модель. Хотя для получения адекватной механистической модели вначале может не хватать информации, применение адаптивных методов [26] иногда приводит к такой модели.

Во многих случаях для получения механистической модели нужны подробные сведения или значительные экспериментальные возможности, которых может не быть, и приходится прибегать к эмпирической модели. Конечно, точная механистическая и чисто эмпирическая модели представляют собой крайности. Обычно используемые модели занимают промежуточное положение. В частности, можно использовать неполные теоретические представления для указания подходящего класса матема-

тических функций, которые могут быть затем эмпирически подогнаны. Это означает, что число параметров модели и их численные значения оцениваются по экспериментальным данным.

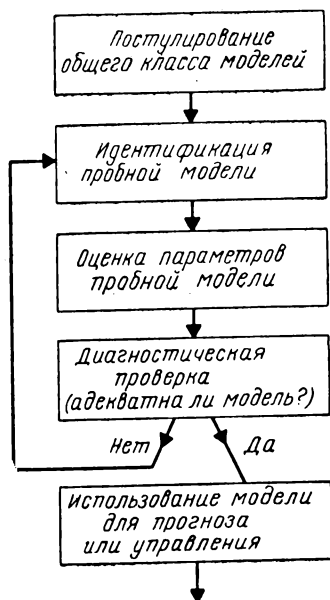


Рис. 1.7. Этапы итеративного подхода к построению моделей.

вид модели, но может дать и хорошие оценки численных значений ее параметров. Эти значения можно потом проверить, анализируя реальные данные.

На рис. 1.7 показан итеративный подход к построению моделей для прогнозирования и управления, используемый в этой книге.

1) Из взаимодействия теории и практики выбирается *полезный класс моделей* (для достижения поставленной цели).

2) Так как этот класс слишком обширен для непосредственной подгонки к данным, развиваются грубые методы *идентификации* подклассов этих моделей. Такие методы идентификации моделей используют имеющиеся данные и знание системы и дают для опробования соответствующие экономичные подклассы моделей. Кроме того, процесс идентификации может быть использован для получения грубых предварительных оценок параметров модели,

Такой подход принят в этой книге. Как мы указывали ранее, описываемые здесь стохастические и динамические модели могут быть обоснованы, по крайней мере частично, с теоретической точки зрения как имеющие правильные основные свойства.

Обычно полагают, что для анализа доступны последовательные значения членов изучаемого временного ряда или данных о входе — выходе. Следует провести не менее 50, а предпочтительнее 100 и более наблюдений. В случаях, где данные о 50 и более наблюдениях в прошлом недоступны, действуют, используя опыт и прошлую информацию для получения предварительной модели. Эта модель может время от времени корректироваться по мере поступления большего числа данных.

При подгонке динамических моделей теоретический анализ не только указывает на подходящий

3) Пробная модель *подгоняется* к данным; *оцениваются* ее параметры. Грубые оценки, получаемые на этапе идентификации, теперь можно использовать как начальные значения в более точных итеративных методах оценивания параметров.

4) *Диагностические проверки* позволяют выявить возможные дефекты подгонки и диагностировать их причины. Если такие дефекты не выявлены, модель готова к использованию. Если обнаружено какое-либо несоответствие, итеративные циклы идентификации, оценок и диагностической проверки повторяются до тех пор, пока не будет найдено подходящее представление модели.

Идентификация, оценивание и диагностическая проверка моделей временных рядов обсуждаются в гл. 6, 7, 8 и 9; те же процедуры для моделей передаточных функций — в гл. 11.

Часть I

Стохастические модели и основанное на них прогнозирование

В первой части этой книги, включающей гл. 2—5, описан важный класс стохастических моделей и их использование в прогнозировании.

Модель, описывающая вероятностную структуру последовательности наблюдений, называется *стохастическим процессом*. Временной ряд из N наблюдений $\mathbf{z}' = (z_1, z_2, \dots, z_N)$ рассматривается как выборочная реализация из бесконечной популяции выборок, которые могли бы генерироваться процессом. Главная цель статистического исследования — узнать свойства популяции по свойствам выборки. Например, сделать прогноз — это значит узнать вероятностное распределение *будущих наблюдений* по выборке \mathbf{z} значений из прошлого. Чтобы сделать это, нам необходимо уметь описывать стохастические процессы и временные ряды и знать классы стохастических моделей, пригодных для описания встречающихся на практике ситуаций.

Важный класс стохастических процессов, рассмотренный в гл. 2, — *стационарные процессы*. Предполагается, что они находятся в определенном статистическом равновесии и в частности колеблются относительно фиксированного среднего значения. Полезным инструментом для описания поведения стационарных процессов являются *автокорреляционная функция* и *спектр*.

Частным случаем стационарных стохастических процессов, особенно важным при моделировании временных рядов, являются процессы авторегрессии, скользящего среднего и смешанные процессы авторегрессии — скользящего среднего. Свойства этих процессов и в частности их корреляционная структура описаны в гл. 3.

Поскольку многие встречающиеся на практике временные ряды (например, биржевые цены и данные о сбыте) имеют нестационарные характеристики, стационарные модели, введенные в гл. 3, обобщаются в гл. 4 для получения полезного класса нестационарных моделей, называемых моделями авторегрессии — проинтегрированного скользящего среднего (АРПСС). Использование всех этих моделей в прогнозировании временных рядов обсуждается и иллюстрируется примерами в гл. 5.

Автокорреляционная функция и спектр

Главная особенность в разработке моделей временных рядов — предположение о некоторой форме *статистического равновесия*. Частное предположение такого рода (и, как будет показано позднее, излишне жесткое) — предположение о *стационарности*. Обычно стационарный временной ряд можно удобно описать его средним значением, дисперсией и *автокорреляционной функцией*, или, что эквивалентно, средним значением, дисперсией и *спектральной плотностью*. В этой главе мы рассмотрим свойства этих функций и в частности свойства автокорреляционной функции, которая многократно используется в последующих главах.

2.1. АВТОКОРРЕЛЯЦИОННЫЕ СВОЙСТВА СТАЦИОНАРНЫХ МОДЕЛЕЙ

2.1.1. Временные ряды и стохастические процессы

Временные ряды. Временной ряд — это множество наблюдений, генерируемых последовательно во времени. Если время непрерывно, временной ряд называется *непрерывным*. Если время изменяется дискретно, временной ряд *дискретен*. Наблюдения дискретного временного ряда, сделанные в моменты времени $\tau_1, \tau_2, \dots, \tau_t, \dots, \tau_N$, могут быть обозначены через $z(\tau_1), z(\tau_2), \dots, z(\tau_t), \dots, z(\tau_N)$. В этой книге рассматриваются только дискретные временные ряды, в которых наблюдения делаются через фиксированный интервал h . Когда имеется N *последовательных* значений такого ряда, доступных для анализа, мы пишем $z_1, z_2, \dots, z_t, \dots, z_N$, обозначая так наблюдения, сделанные в равноотстоящие моменты времени $\tau_0 + h, \tau_0 + 2h, \dots, \tau_0 + th, \dots, \tau_0 + Nh$. Во многих случаях значения τ_0 и h не важны, но если необходимо точно определить времена наблюдений, нужно указать эти два значения. Если мы принимаем τ_0 за начало и h за единицу времени, мы можем рассматривать z_t как наблюдение в *момент времени t* .

Дискретные временные ряды могут появляться двумя путями.

1) *Выборкой* из непрерывных временных рядов, например, в ситуации, показанной на рис. 1.2, где значения непрерывных входа и выхода газовой печи считаются с интервалом 9 с.

2) *Накоплением* переменной в течение некоторого периода времени; примерами могут служить дождевые осадки, которые обычно накапливают за такие периоды, как день или месяц, или выход партий продукта, накапливающегося за время цикла. Например, на рис. 2.1 показан временной ряд, состоящий из значений выхода 70 последовательных партий продукта химического процесса.

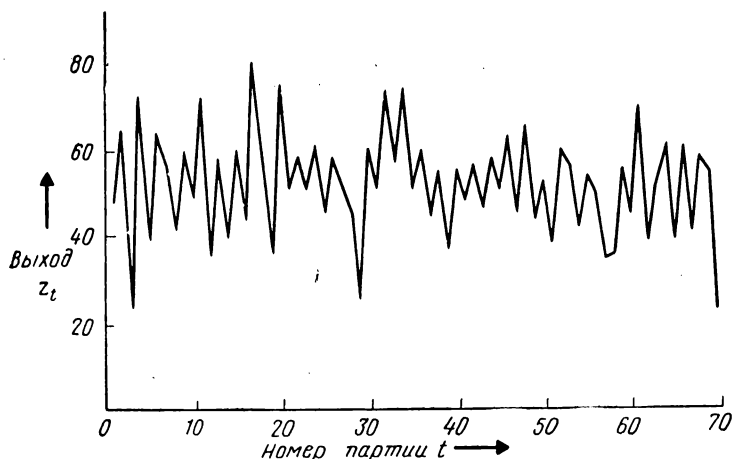


Рис. 2.1. Выход 70 последовательных партий продукта химического процесса.

Детерминированные и случайные временные ряды. Если будущие значения временного ряда точно определены какой-либо математической функцией, например, такой, как

$$z_t = \cos(2\pi ft),$$

временной ряд называют *детерминированным*. Если будущие значения могут быть описаны только с помощью распределения вероятностей, временной ряд называют *недетерминированным*, или просто *случайным*. Данные о партиях продукта на рис. 2.1 — это пример случайного временного ряда. Хотя в этом ряду имеется отчетливая тенденция к чередованию «вверх — вниз», невозможно точно предсказать выход следующей партии. В этой книге мы будем исследовать именно такие случайные временные ряды.

Стохастические процессы. Статистическое явление, развивающееся во времени согласно законам теории вероятности, на-

зывается *стохастическим процессом*. Мы часто будем называть его просто *процессом*, опуская слово «стохастический». Подлежащий анализу временной ряд может рассматриваться как одна

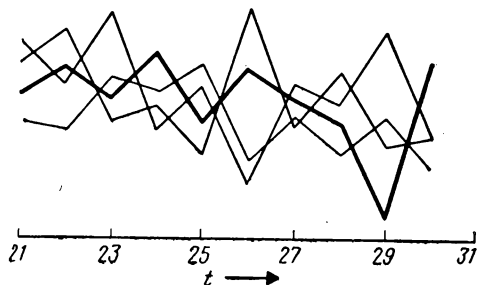


Рис. 2.2. Наблюдаемый временной ряд (жирная линия) и другие временные ряды, являющиеся реализациями одного и того же стохастического процесса.

частная *реализация* изучаемой системы, генерируемая скрытым вероятностным механизмом. Другими словами, анализируя *временной ряд*, мы рассматриваем его как *реализацию стохастического процесса*.

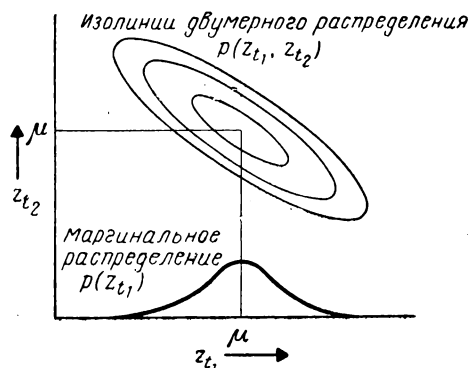


Рис. 2.3. Изолинии плотности двумерного распределения вероятности, описывающего стохастический процесс в моменты времени t_1 и t_2 ; там же маргинальное распределение в момент t_1 .

Например, анализируя данные о выходе партии продукта на рис. 2.1, мы можем представить себе другие множества наблюдений (другие реализации порождающего эти наблюдения стохастического процесса), которые могут быть генерированы той же самой химической системой, за те же $n = 70$ циклов. Так, например, на рис. 2.2 показаны выходы партий продукта с $t = 21$ по $t = 30$ (жирная линия) вместе с другими временными

рядами, которые *могли бы быть* получены из популяции временных рядов, определяемых тем же стохастическим процессом. Отсюда следует, что мы можем рассматривать наблюдение z_t в данное время t , скажем $t = 25$, как реализацию случайной величины z_t с плотностью вероятности $p(z_t)$.

Подобным образом наблюдения в любые два момента времени, $t_1 = 25$ и $t_2 = 27$, могут рассматриваться как реализации двух случайных величин z_{t_1} и z_{t_2} с совместной плотностью вероятности $p(z_{t_1}, z_{t_2})$. Например, на рис. 2.3 показаны изолинии плотности для такого совместного распределения вместе с маргинальным распределением в момент времени t_1 . В общем наблюдения, образующие временной ряд, могут быть описаны всевозможными N -мерными случайными величинами (z_1, z_2, \dots, z_N) с плотностью вероятности $p(z_1, z_2, \dots, z_N)$.

2.1.2. Стационарные стохастические процессы

Весьма специальный класс стохастических процессов, называемых стационарными процессами, основывается на предположении, что процесс находится в определенном *статистическом равновесии*. Стохастический процесс называется *строго стационарным*, если его свойства не зависят от изменения начала отсчета времени. Иными словами, если совместное распределение вероятностей m наблюдений $z_{t_1}, z_{t_2}, \dots, z_{t_m}$, сделанных в *любые* моменты времени t_1, t_2, \dots, t_m , такое же, что и для m наблюдений $z_{t_1+k}, z_{t_2+k}, \dots, z_{t_m+k}$, сделанных в моменты времени $t_1 + k, t_2 + k, \dots, t_m + k$. Поэтому, чтобы дискретный процесс был строго стационарным, взаимное распределение любой совокупности наблюдений не должно изменяться при сдвиге всех времен наблюдений вперед или назад на любое целое число k .

Среднее значение и дисперсия стационарного процесса. Когда $m = 1$, из предположения о стационарности следует, что распределение вероятности $p(z_t)$ одинаково для всех времен t и может быть записано как $p(z)$. Отсюда стохастический процесс имеет постоянное среднее значение

$$\mu = E[z_t] = \int_{-\infty}^{\infty} zp(z) dz, \quad (2.1.1)$$

определяющее уровень, относительно которого он флуктуирует, и постоянную дисперсию

$$\sigma_z^2 = E[(z_t - \mu)^2] = \int_{-\infty}^{\infty} (z - \mu)^2 p(z) dz, \quad (2.1.2)$$

измеряющую его *размах* относительно этого уровня. Поскольку распределение вероятности $p(z)$ одинаково для всех времен t , его форма может быть оценена по гистограмме наблюдений z_1, z_2, \dots, z_N временного ряда. Кроме того, среднее значение μ стохастического процесса можно оценить с помощью выборочного среднего временного ряда

$$\bar{z} = \frac{1}{N} \sum_{t=1}^N z_t, \quad (2.1.3)$$

а дисперсию σ_z^2 стохастического процесса — с помощью выборочной дисперсии

$$\hat{\sigma}_z^2 = \frac{1}{N} \sum_{t=1}^N (z_t - \bar{z})^2. \quad (2.1.4)$$

Автоковариация и коэффициенты автокорреляции. Из предположения о стационарности следует также, что совместное распределение вероятностей $p(z_{t_1}, z_{t_2})$ одинаково для всех времен t_1, t_2 , разделенных одним и тем же интервалом. Следовательно, природу этого совместного распределения можно оценить по диаграмме рассеяния, построенной по парам значений z_t, z_{t+k} временного ряда, разделенных постоянным интервалом, или *задержкой* k . Диаграммы рассеяния на рис. 2.4 построены по данным циклического процесса. На рис. 2.4, а показаны данные для задержки $k = 1$ (по одной оси отложено z_{t+1} , по другой z_t). На рис. 2.4, б показаны данные для задержки $k = 2$ (по одной оси отложено z_{t+2} , по другой z_t). Мы видим, что соседние значения временных рядов коррелированы; корреляция между z_t и z_{t+1} отрицательная, а между z_t и z_{t+2} положительная. Ковариация между значениями z_t и z_{t+k} , отделенными k интервалами времени, называется *автоковариацией* с задержкой k и определяется как

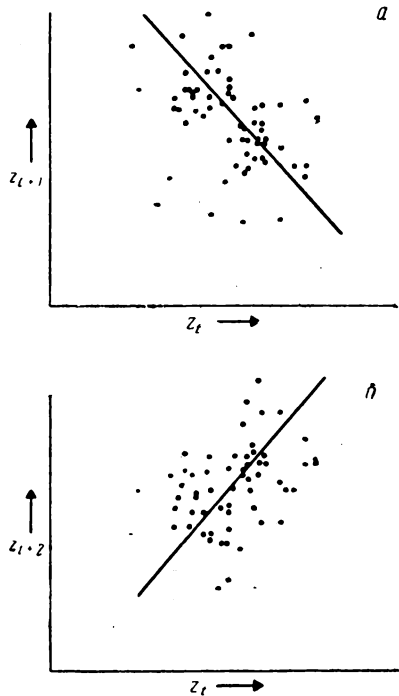


Рис. 2.4. Диаграммы разброса при задержках 1 и 2 для данных о выходе партий продукта на рис. 2.1.

$$\gamma_k = \text{cov}[z_t, z_{t+k}] = E[(z_t - \mu)(z_{t+k} - \mu)]. \quad (2.1.5)$$

Аналогично автокорреляция с задержкой k равна

$$\rho_k = \frac{E[(z_t - \mu)(z_{t+k} - \mu)]}{\sqrt{E[(z_t - \mu)^2]E[(z_{t+k} - \mu)^2]}} = \frac{E[(z_t - \mu)(z_{t+k} - \mu)]}{\sigma_z^2},$$

поскольку для стационарного процесса дисперсия $\sigma_z^2 = \gamma_0$ в момент времени $t+k$ та же, что и в момент времени t .

Таким образом, автокорреляция с задержкой k равна

$$\rho_k = \gamma_k / \gamma_0, \quad (2.1.6)$$

откуда вытекает, что $\rho_0 = 1$.

2.1.3. Положительная определенность и автоковариационная матрица

Ковариационная матрица, связанная со стационарным процессом, для наблюдений (z_1, z_2, \dots, z_n) , сделанных в n последовательных моментов времени, имеет вид

$$\begin{aligned} \Gamma_n &= \begin{bmatrix} \gamma_0 & \gamma_1 & \gamma_2 & \dots & \gamma_{n-1} \\ \gamma_1 & \gamma_0 & \gamma_1 & \dots & \gamma_{n-2} \\ \gamma_2 & \gamma_1 & \gamma_0 & \dots & \gamma_{n-3} \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \gamma_{n-1} & \gamma_{n-2} & \gamma_{n-3} & \dots & \gamma_0 \end{bmatrix} = \\ &= \sigma_z^2 \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{n-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{n-2} \\ \rho_2 & \rho_1 & 1 & \dots & \rho_{n-3} \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \dots & 1 \end{bmatrix} = \sigma_z^2 \mathbf{P}_n. \end{aligned} \quad (2.1.7)$$

Ковариационная матрица Γ_n такого вида (симметричная и с одними и теми же элементами на любой диагонали) называется автоковариационной матрицей; соответствующая корреляционная матрица \mathbf{P}_n называется автокорреляционной матрицей. Рассмотрим любую линейную функцию случайных переменных $z_t, z_{t-1}, \dots, z_{t-n+1}$

$$L_t = l_1 z_t + l_2 z_{t-1} + \dots + l_n z_{t-n+1}. \quad (2.1.8)$$

Так как для стационарного процесса $\text{cov}[z_i, z_j] = \gamma_{|j-i|}$, дисперсия L_t равна

$$\text{var}[L_t] = \sum_{i=1}^n \sum_{j=1}^n l_i l_j \gamma_{|j-i|}$$

и всегда больше нуля, если не все l_i равны нулю. Отсюда следует, что и автоковариационная, и автокорреляционная матрицы любого стационарного процесса положительно определенные *).

Условия, которым удовлетворяют автокорреляции стационарного процесса. Положительная определенность автокорреляционной матрицы (2.1.7) требует, чтобы определитель и все главные миноры этой матрицы были положительны. В частности, для $n = 2$

$$\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix} > 0,$$

так что $1 - \rho_1^2 > 0$

и, следовательно,

$$-1 < \rho_1 < 1.$$

Аналогично для $n = 3$ должно выполняться

$$\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix} > 0, \quad \begin{vmatrix} 1 & \rho_2 \\ \rho_2 & 1 \end{vmatrix} > 0,$$

$$\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix} > 0,$$

из чего следует

$$-1 < \rho_1 < 1,$$

$$-1 < \rho_2 < 1,$$

$$-1 < \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} < 1$$

и т. д. Так как \mathbf{P}_n должна быть положительно определенной для всех n , автокорреляции стационарного процесса должны удовлетворять очень большому числу условий. Как будет показано в разд. 2.2.3, все эти условия могут быть объединены в определении спектра.

Стационарность линейных функций. Из определения стационарности следует, что процесс L_t , полученный выполнением линейной операции (2.1.8) над стационарным процессом z_t , также стационарен. В частности, стационарны первая разность $\nabla z_t = z_t - z_{t-1}$ и высшие разности $\nabla^d z_t$. Этот результат особенно

*) Точнее, эти матрицы — положительно полуопределенные. — Прим. ред.

важен для исследования свойств нестационарных временных рядов, рассматриваемых в гл. 4.

Гауссовские процессы. Если распределение значений процесса для любого множества моментов времени — многомерное нормальное распределение, процесс называется *нормальным*, или *гауссовским*. Так как многомерное нормальное распределение полностью описывается его моментами первого и второго порядков, существование постоянного среднего значения μ и автоковариационной матрицы Γ_n вида (2.1.7) для всех n достаточно, чтобы обеспечить стационарность нормального процесса.

Слабая стационарность. Для того чтобы процесс был строго стационарным, вся его вероятностная структура должна зависеть только от разностей времен. Менее жесткое требование, называемое *слабой стационарностью* порядка f , заключается в том, чтобы моменты до некоторого порядка f зависели только от разностей времен. Например, существование среднего значения μ и автоковариационной матрицы Γ_n вида (2.1.7) достаточно для получения стационарности второго порядка. Тогда стационарность второго порядка плюс предположение о нормальности обеспечивают строгую стационарность.

2.1.4. Автоковариационные и автокорреляционные функции

В разд. 2.1.2 было показано, что автоковариационный коэффициент γ_k для задержки k измеряет ковариацию между двумя величинами z_t и z_{t+k} , разделенными интервалом k . Функция γ_k от задержки k называется *автоковариационной функцией* $\{\gamma_k\}$ стохастического процесса. Аналогично функция ρ_k от задержки k называется *автокорреляционной функцией* $\{\rho_k\}$ стохастического процесса. Заметим, что автокорреляционная функция безразмерна, т. е. независима от масштаба измерения временных рядов. Так как $\gamma_k = \rho_k \sigma_z^2$, знание автокорреляционной функции $\{\rho_k\}$ и дисперсии σ_z^2 эквивалентно знанию автоковариационной функции $\{\gamma_k\}$.

Автокорреляционная функция, представленная на рис. 2.5 как график диагональных элементов автокорреляционной матрицы, показывает, как изменяется корреляция между двумя любыми членами ряда по мере изменения расстояния между ними. Так как $\rho_k = \rho_{-k}$, автокорреляционная функция должна быть симметричной относительно нуля и на практике необходимо изображать только правую половину функции (для положительных k).

На рис. 2.6 показана правая половина автокорреляционной функции, приведенной на рис. 2.5. В дальнейшем, когда будем говорить об автокорреляционной функции, мы часто будем иметь в виду ее правую половину. В прошлом автокорреляционную функцию иногда называли *коррелограммой*.

Из ранее сказанного следует, что *нормальный* стационарный процесс z_t полностью описывается его средним значением μ и

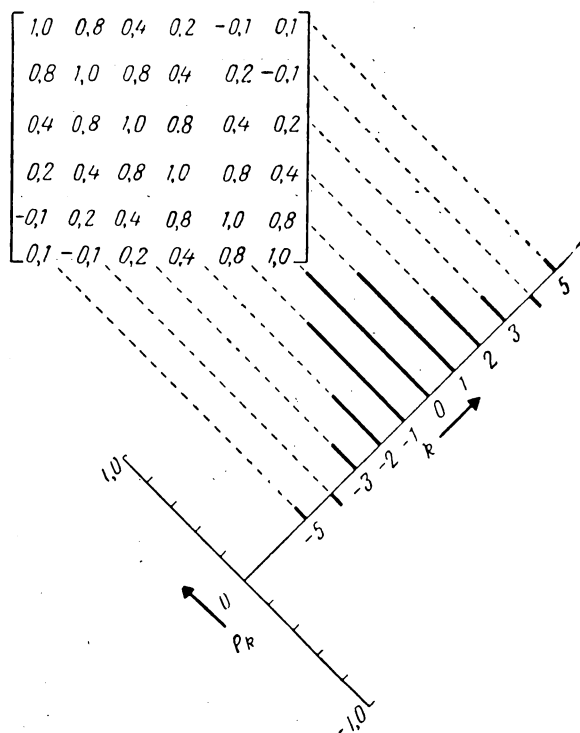


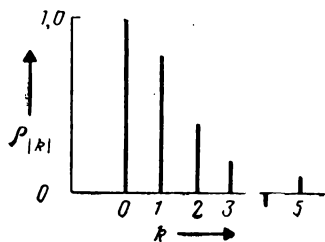
Рис. 2.5. Автокорреляционная матрица и соответствующая автокорреляционная функция.

его автоковариационной функцией $\{\gamma_k\}$, или, что эквивалентно, его средним значением μ , дисперсией σ_z^2 и автокорреляционной функцией $\{\rho_k\}$.

2.1.5. Оценивание автоковариационной и автокорреляционной функций

До сих пор мы рассматривали только теоретическую автокорреляционную функцию, которая описывает математически определенный стохастический процесс. На практике мы имеем конечный временной ряд z_1, z_2, \dots, z_N из N наблюдений, по которому мы можем найти только *выборочные оценки* автокорреляций.

В статистике имеется несколько оценок автокорреляционной функции, их свойства в деталях рассмотрены в [27]. Там сделан вывод, что наиболее удовлетворительной оценкой автокорреляции ρ_k при задержке k является



$$r_k = c_k / c_0, \quad (2.1.9)$$

где

$$c_k = \frac{1}{N} \sum_{t=1}^{N-k} (z_t - \bar{z})(z_{t+k} - \bar{z}),$$

$$k = 0, 1, 2, \dots, K \quad (2.1.10)$$

Рис. 2.6. Автокорреляционная функция с рис. 2.5 для неотрицательных задержек.

— выборочная оценка автоковариации γ_k , а \bar{z} — среднее значение временного ряда. Мы проиллюстрируем (2.1.10), вычислив r_1 по 10 первым значениям химического процесса на рис. 2.1, приведенным в табл. 2.1. Среднее значение первых десяти значений таблицы равно 51, и отклонения от среднего равны —4, 13, —28, 20, —13, 13, 4, —10, 8, —3.

Таблица 2.1

Ряд из 70 последовательных выходов партий
продукта химического процесса.
Этот ряд имеется также под названием
«ряд F» в сборнике временных рядов
в конце данного выпуска

1—15	16—30	31—45	46—60	61—70
47	44	50	62	68
64	80	71	44	38
23	55	56	64	50
71	37	74	43	60
38	74	50	52	39
64	51	58	38	59
55	57	45	59	40
41	50	54	55	57
59	60	36	41	54
48	45	54	53	23
71	57	48	49	
35	50	55	34	
57	45	45	35	
40	25	57	54	
58	59	50	45	

Тогда

$$\sum_{i=1}^9 (z_i - \bar{z})(z_{i+1} - \bar{z}) = \\ = (-4)(13) + (13)(-28) + \dots + (8)(-3) = -1497.$$

Отсюда

$$c_1 = \frac{-1497}{10} = -149,7.$$

Аналогично находим, что $c_0 = 189,6$. Отсюда

$$r_1 = \frac{c_1}{c_0} = \frac{-149,7}{189,6} = -0,79,$$

так как для практических целей достаточно округлять автокорреляцию до двух знаков после запятой. Вычисления, приведен-

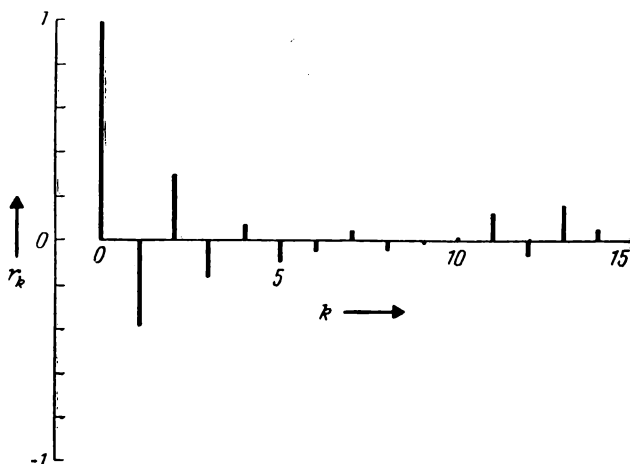


Рис. 2.7. Выборочная автокорреляционная функция данных о партиях продукта.

ные выше, чисто иллюстративны. На практике для получения полезной оценки автокорреляционной функции нам нужно по меньшей мере 50 наблюдений, и выборочные автокорреляции r_k должны быть вычислены для $k = 0, 1, \dots, K$, где K не больше чем примерно $N/4$.

Первые 15 значений r_k , полученные по всей серии из 70 наблюдений, приведены в табл. 2.2 и показаны на рис. 2.7. Выборочная автокорреляционная функция характеризуется знакопеременной корреляцией с тенденцией к затуханию по мере роста задержки. Такие автокорреляционные функции встречаются в данных о выпуске продукции и являются следствием эффектов «переноса». В этом частном примере высокий выход цикла

приводил к получению дегтярных остатков, которые не полностью устранялись из емкости и отрицательно влияли на выход следующего цикла.

Таблица 2.2

Выборочная автокорреляционная функция данных о партиях продукта

k	r_k	k	r_k	k	r_k
1	-0,39	6	-0,05	11	0,11
2	0,30	7	0,04	12	-0,07
3	-0,17	8	-0,04	13	0,15
4	0,07	9	-0,01	14	0,04
5	-0,10	10	0,01	15	-0,01

Программа для ЭВМ, позволяющая вычислить автоковариационную и автокорреляционную функции, описана под номером 1 в сборнике программ в конце книги.

2.1.6. Стандартные ошибки оценок автокорреляций

Для идентификации модели временного ряда методами, которые будут описаны в гл. 6, необходимо иметь грубый способ проверки того, является ли ρ_k практически нулем для задержек больше некоторой. Для этой цели можно использовать следующее выражение для дисперсии выборочного коэффициента автокорреляции стационарного нормального процесса, предложенное Бартлеттом [28]:

$$\text{var}[r_k] \approx \frac{1}{N} \sum_{v=-\infty}^{+\infty} (\rho_v^2 + \rho_{v+k}\rho_{v-k} - 4\rho_k\rho_v\rho_{v-k} + 2\rho_v^2\rho_k^2). \quad (2.1.11)$$

Например, если $\rho_k = \phi^{|k|}$ ($-1 < \phi < 1$), т. е. автокорреляционная функция затухает экспоненциально, (2.1.11) дает

$$\text{var}[r_k] \approx \frac{1}{N} \left[\frac{(1 + \phi^2)(1 - \phi^{2k})}{1 - \phi^2} - 2k\phi^{2k} \right], \quad (2.1.12)$$

и в частности

$$\text{var}[r_1] \approx \frac{1}{N} (1 - \phi^2).$$

Для любого процесса, у которого все автокорреляции ρ_v равны нулю при $v > q$, все члены, кроме первого, появляющиеся в правой части (2.1.11), равны нулю при $k > q$. Отсюда для дисперсии выборочных автокорреляций r_k с задержками k , большими, чем некоторое значение q , за которым теоретическую

автокорреляционную функцию можно полагать «затухшей», аппроксимация Бартлетта дает

$$\text{var}[r_k] \approx \frac{1}{N} \left\{ 1 + 2 \sum_{v=1}^q \rho_v^2 \right\}, \quad k > q. \quad (2.1.13)$$

Например, если k увеличивается и при условии, что ϕ не близко к 1, (2.1.12) быстро сходится к своему предельному значению

$$\lim_{k \rightarrow \infty} \text{var}[r_k] \approx \frac{1}{N} \left(\frac{1 + \phi^2}{1 - \phi^2} \right),$$

так же определяемому (2.1.13).

Чтобы практически использовать (2.1.13), нужно вместо теоретических автокорреляций ρ_h ($h = 1, 2, \dots, q$) подставлять выборочные автокорреляции r_h ; когда это сделано, квадратный корень из (2.1.13) мы будем называть *стандартной ошибкой при больших задержках*. В предположении, что теоретические автокорреляции ρ_h практически равны нулю при задержках, больших некоторой гипотетической задержки $k = q$, эта стандартная ошибка аппроксимирует стандартное отклонение r_k для больших задержек ($k > q$).

Подобные приближенные выражения для ковариации между выборочными корреляциями r_k и r_{k+s} при двух различных задержках k и $k + s$ были даны Бартлеттом [28]. В частности, аппроксимация для больших задержек сводится к

$$\text{cov}[r_k, r_{k+s}] \approx \frac{1}{N} \sum_{v=-\infty}^{\infty} \rho_v \rho_{v+s}. \quad (2.1.14)$$

Результат Бартлетта (2.1.14) указывает, что индивидуальные автокорреляции должны интерпретироваться с осторожностью, поскольку между соседними значениями может иметь место значительная ковариация. Этот эффект в некоторых случаях может искажать вид выборочной автокорреляционной функции, которая вопреки ожиданию может не затухать [27, 29].

Пример. Следующие выборочные автокорреляции были получены из временного ряда длиной $N = 200$ наблюдений, порождаемых стохастическим процессом, о котором было известно, что $\rho_1 = -0,4$ и $\rho_k = 0$ для $k \geq 2$:

k	1	2	3	4	5
r_k	-0,38	-0,08	0,11	-0,08	0,02
k	6	7	8	9	10
r_k	0,00	0,00	0,00	0,07	-0,08

В предположении, что ряд полностью случайный, имеем $q = 0$. Тогда для всех задержек (2.1.13) дает

$$\text{var}[r_k] \approx \frac{1}{N} = \frac{1}{200} = 0,005.$$

Соответствующая стандартная ошибка равна $0,07 = (0,005)^{1/2}$. Так как значение $r_1 = -0,38$ в 5 раз превосходит по модулю стандартную ошибку, можно заключить, что ρ_1 не равно нулю. Далее выборочные автокорреляции для задержек, больших 1, малы; следовательно, разумен вопрос, соответствует ли ряд принятой гипотезе (ее соответствие явлению мы обсудим позднее), что ρ_1 не равно нулю, а $\rho_k = 0 (k \geq 2)$. Используя (2.1.13) с $q = 1$ и подставляя r_1 вместо ρ_1 , находим, что при такой гипотезе выборочная дисперсия для больших задержек равна

$$\text{var}[r_k] \approx \frac{1}{200} [1 + 2(-0,38)^2] = 0,0064, \quad k > 1,$$

что дает стандартную ошибку 0,08. Поскольку выборочные автокорреляции для задержек, больших чем 1, малы по сравнению с этой стандартной ошибкой, нет причин сомневаться в адекватности модели $\rho_1 \neq 0, \rho_k = 0 (k \geq 2)$.

2.2. СПЕКТРАЛЬНЫЕ СВОЙСТВА СТАЦИОНАРНЫХ МОДЕЛЕЙ

2.2.1. Периодограмма

Другой способ анализа временного ряда основан на предположении, что он образован синусоидами и косинусоидами различных частот. Способ реализации этой идеи, предложенный в 1898 г. Шустером [30], основан на применении *периодограммы* (см. также [31]). Периодограмма вначале использовалась для обнаружения и оценок амплитуды синусоидальной компоненты, скрытой шумом. Мы будем пользоваться ею в дальнейшем для контроля случайности ряда (обычно ряда из остаточных ошибок после подгонки какой-либо модели), в котором периодические компоненты неизвестной частоты могут еще сохраниться.

Для иллюстрации вычислений периодограммы положим, что число наблюдений N нечетно: $N = 2q + 1$. Если мы подгоняем модель ряда Фурье

$$z_t = \alpha_0 + \sum_{i=1}^q (\alpha_i c_{it} + \beta_i s_{it}) + e_t, \quad (2.2.1)$$

где $c_{it} = \cos 2\pi f_i t$, $s_{it} = \sin 2\pi f_i t$, $f_i = i/N$ — i -я гармоника основной частоты $1/N$, то оценки (наименьших квадратов) коэф-

коэффициентов α_0 и (α_i, β_i) будут

$$a_0 = \bar{z}, \quad (2.2.2)$$

$$a_i = \frac{2}{N} \sum_{t=1}^N z_t c_{it}, \quad i = 1, 2, \dots, q, \quad (2.2.3)$$

$$b_i = \frac{2}{N} \sum_{t=1}^N z_t s_{it}. \quad (2.2.4)$$

Тогда периодограмма состоит из $q = (N - 1)/2$ значений

$$I(f_i) = \frac{N}{2} (a_i^2 + b_i^2), \quad i = 1, 2, \dots, q, \quad (2.2.5)$$

где $I(f_i)$ называется *интенсивностью* на частоте f_i .

Когда N четно, мы полагаем $N = 2q$, и (2.2.2) — (2.2.5) применимы для $i = 1, 2, \dots, q - 1$.

Для $i = q$ имеем

$$a_q = \frac{1}{N} \sum_{t=1}^N (-1)^t z_t,$$

$$b_q = 0$$

и

$$I(f_q) = I(0,5) = Na_q^2.$$

Заметим, что наиболее высокая частота составляет 0,5 цикла за временной интервал, поскольку наименьший период равен двум интервалам.

2.2.2. Дисперсионный анализ

Из таблицы дисперсионного анализа, относящейся к линейной регрессии (2.2.1) для нечетного N , устранив среднее значение, мы можем выделить $(N - 1)/2$ пар степеней свободы. Они связаны с парами коэффициентов (a_1, b_1) , (a_2, b_2) , \dots , (a_q, b_q) и, следовательно, с частотами $1/N, 2/N, \dots, q/N$. Очевидно, что значение периодограммы $I(f_i) = (N/2)(a_i^2 + b_i^2)$ есть просто «сумма квадратов» пар коэффициентов (a_i, b_i) , а следовательно, связано с частотой $f_i = i/N$. Таким образом,

$$\sum_{t=1}^N (z_t - \bar{z})^2 = \sum_{i=1}^q I(f_i). \quad (2.2.6)$$

Когда N четно, существует $(N - 2)/2$ пар степеней свободы и еще одна одиночная степень свободы, связанная с коэффициентом a_q .

Если ряд действительно случаен и не содержит регулярной синусоидальной компоненты, т. е.

$$z_t = \alpha_0 + e_t,$$

где α_0 — фиксированное среднее значение, а e — независимые, нормально распределенные случайные величины с нулевым средним значением и дисперсией σ^2 , то каждая компонента $I(f_i)$ имеет математическое ожидание $2\sigma^2$ и будет распределена как $\sigma^2\chi^2(2)$ независимо от всех других компонент *). Напротив, если ряд содержит регулярную случайную компоненту с частотой f_i , амплитудой A и фазой F , так что

$$z_t = \alpha_0 + \alpha \cos(2\pi f_i t) + \beta \sin(2\pi f_i t) + e_t,$$

где $\alpha = A \sin F$ и $\beta = A \cos F$, то сумма квадратов $I(f_i)$ будет иметь тенденцию к увеличению, так как ее математическое ожидание равно $2\sigma^2 + N(\alpha^2 + \beta^2)/2 = 2\sigma^2 + NA^2/2$.

Практически мало вероятно, чтобы частота f неизвестной регулярной синусоидальной компоненты точно совпадала с какой-либо из частот f_i , для которых могут быть вычислены интенсивности. В этом случае на периодограмме будет наблюдаться увеличение интенсивности в непосредственной близости от f .

Пример. При построении периодограммы обычно используется большое число наблюдений. Однако для иллюстрации процедуры построения мы воспользуемся набором из 12 среднемесячных температур (в градусах Цельсия) для центральной Англии в 1964 г., приведенным в табл. 2.3.

Таблица 2.3

Среднемесячные температуры в центральной Англии в 1964 г.

t	1	2	3	4	5	6	7	8	9	10	11	12
z_t	3,4	4,5	4,3	8,7	13,3	13,8	16,1	15,5	14,1	8,9	7,4	3,6
c_{1t}	0,87	0,50	0,00	-0,50	-0,87	-1,00	-0,87	-0,50	0,00	0,50	0,87	1,00

Третья строчка в табл. 2.3 содержит значения $c_{1t} = \cos(2\pi t/12)$, необходимые для вычисления a_1 по формуле

$$a_1 = \frac{1}{6}[(3,4)(0,87) + \dots + (3,6)(1,00)] = -5,30.$$

*) Подразумевается, что $\chi^2(m)$ — случайная переменная, имеющая χ^2 -распределение с m степенями свободы. Точное определение см., например, в приложении П7.1.

Значения a_i , b_i , $i = 1, 2, \dots, 6$, даны в табл. 2.4; им соответствует табл. 2.5 дисперсионного анализа. Как и ожидалось, главная компонента этих температурных данных имеет 12-месячный период с частотой $1/12$ цикла в месяц.

Таблица 2.4

Амплитуды синусоидальных и косинусоидальных компонент для температурных данных

i	a_i	b_i
1	-5,30	-3,82
2	0,05	0,17
3	0,10	0,50
4	0,52	-0,52
5	0,09	-0,58
6	-0,30	—

Таблица 2.5

Таблица дисперсионного анализа температурного ряда

i	Частота f_i	Период	Периодограмма $l(f_i)$	Число степеней свободы	Среднее квадратичное
1	$1/12$	12	254,96	2	127,48
2	$1/6$	6	0,19	2	0,10
3	$1/4$	4	1,56	2	0,78
4	$1/3$	3	3,22	2	1,61
5	$5/12$	$12/5$	2,09	2	1,05
6	$1/2$	2	1,08	1	1,08
Всего			263,10	11	23,92

2.2.3. Спектральная плотность и нормированный спектр

Для полноты мы кратко обсудим ниже понятия спектра и спектральной плотности. Применение этих важных понятий более полно описано в [27]. Мы не используем их для анализа временных рядов в этой книге, поэтому при первом чтении этот раздел можно опустить.

Выборочный спектр. При определении периодограммы (2.2.5) предполагается, что частоты $f_i = i/N$ являются гармониками

основной частоты $1/N$. Вводя спектр, мы ослабляем это предположение и позволяем частоте f изменяться непрерывно в диапазоне 0—0,5 Гц. Определение периодограммы может быть изменено следующим образом:

$$I(f) = \frac{2}{N} (a_f^2 + b_f^2), \quad 0 \leq f \leq \frac{1}{2}, \quad (2.2.7)$$

где $I(f)$ называется *выборочным спектром* [27]. Подобно периодограмме, он может быть использован для обнаружения и оценки амплитуд синусоидальной компоненты неизвестной частоты f , скрытой в шуме, и действительно это даже удобнее, если только не известно, что частота f связана гармонически с длиной ряда, т. е. $f = i/N$. Более того, он является отправным пунктом для теории спектрального анализа, использующей важное соотношение, приведенное в приложении П2.1. Это соотношение устанавливает связь выборочного спектра $I(f)$ и оценок c_k автоковариационной функции:

$$I(f) = 2 \left(c_0 + 2 \sum_{k=1}^{N-1} c_k \cos 2\pi f k \right), \quad 0 \leq f \leq \frac{1}{2}. \quad (2.2.8)$$

Таким образом, *выборочный спектр* — это косинус-преобразование Фурье выборочной автоковариационной функции.

Спектр. Периодограмма и выборочный спектр — удобные понятия анализа временных рядов, образованных смесью синусоид и косинусоид с *постоянными* частотами, скрытыми в шуме. Однако стационарные временные ряды такого типа, как описанные в разд. 2.1, характеризуются случайными изменениями частоты, амплитуды и фазы. Для таких рядов выборочный спектр $I(f)$ сильно флуктуирует и не допускает какой-либо разумной интерпретации [27].

Предположим, однако, что выборочный спектр был вычислен для временного ряда из N наблюдений, являющегося реализацией стационарного нормального процесса. Как уже говорилось выше, такой процесс не имеет никаких детерминированных синусоидальных или косинусоидальных компонент, но мы можем формально провести анализ Фурье и получить значения a_f , b_f для любой частоты f . Если повторные реализации N наблюдений порождены стохастическим процессом, мы можем собрать популяцию значений a_f , b_f и $I(f)$. Тогда мы можем найти среднее значение $I(f)$ по повторным реализациям*) длины N , а именно

$$E[I(f)] = 2 \left\{ E[c_0] + 2 \sum_{k=1}^{N-1} E[c_k] \cos 2\pi f k \right\}. \quad (2.2.9)$$

*) Т. е. среднее по ансамблю, или математическое ожидание. — Прим. ред.

Для больших N можно показать (см., например, [27]), что среднее значение автоковариации c_k в повторных реализациях стремится к теоретической автоковариации, т. е.

$$\lim_{N \rightarrow \infty} E[c_k] = \gamma_k.$$

Переходя к пределу в (2.2.9) для $N \rightarrow \infty$, определяем *спектр мощности* *) как

$$p(f) = \lim_{N \rightarrow \infty} E[I(f)] = 2 \left(\gamma_0 + 2 \sum_{k=1}^{\infty} \gamma_k \cos 2\pi f k \right), \quad 0 \leq f \leq 1/2. \quad (2.2.10)$$

Отметим, что так как

$$\begin{aligned} |p(f)| &\leq 2 \left(|\gamma_0| + 2 \sum_{k=1}^{\infty} |\gamma_k| |\cos 2\pi f k| \right) = \\ &= 2 \left(|\gamma_0| + 2 \sum_{k=1}^{\infty} |\gamma_k| \right), \end{aligned} \quad (2.2.11)$$

то для сходимости спектра γ_k должно убывать с ростом k настолько быстро, чтобы обеспечивать сходимость ряда (2.2.11). Так как *спектр мощности* — это *косинус-преобразование Фурье автоковариационной функции*, знание автоковариационной функции математически эквивалентно знанию спектра мощности и наоборот. Далее мы будем называть спектр мощности просто спектром.

Интегрируя (2.2.10) в пределах от 0 до $1/2$, найдем дисперсию процесса

$$\gamma_0 = \sigma_z^2 = \int_0^{1/2} p(f) df. \quad (2.2.12)$$

Следовательно, так же как периодограмма $I(f)$ показывает, каким образом дисперсия (2.2.6) ряда, состоящего из смеси синусоид и косинусоид, распределена между различными гармоническими компонентами, спектр $p(f)$ показывает, как дисперсия стохастического процесса распределена в непрерывном диапазоне частот. Можно интерпретировать $p(f)df$ как приближенное значение дисперсии процесса в частотном диапазоне от f до $f + df$.

Нормированный спектр **). Иногда более удобно определять спектр (2.2.10) при помощи автокорреляций ρ_k , а не автокова-

*) В математической литературе чаще используется термин «спектральная плотность». — *Прим. ред.*

**) В оригинале spectral density function (спектральная плотность). В нашей литературе спектральной плотностью называют функцию $p(f)$. Поэтому мы будем называть определяемую ниже функцию $g(f) = p(f)/\sigma_z^2$ *нормированным спектром*. — *Прим. ред.*

риаций γ_k . Результирующая функция

$$g(f) = p(f)/\sigma_z^2 = 2 \left\{ 1 + 2 \sum_{k=1}^{\infty} \rho_k \cos 2\pi f k \right\}, \quad 0 \leq f \leq 1/2 \quad (2.2.13)$$

называется *нормированным спектром*. Из (2.2.12) вытекает, что эта функция имеет свойство

$$\int_0^{1/2} g(f) df = 1.$$

Поскольку $g(f)$ положительна, она обладает теми же свойствами, что и обычные плотности вероятности. Аналогия распространяется и на оценивание этих двух функций, к обсуждению которого мы переходим.

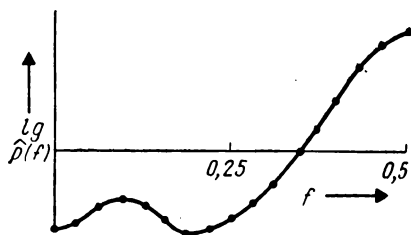


Рис. 2.8. Выборочный спектр мощности данных о партиях продукта.

Оценивание спектра. Казалось бы, что оценка спектра может быть получена из (2.2.10) заменой теоретических автоковариаций γ_k их выборочными оценками c_k . Согласно (2.2.8), это означает, что выборочный спектр принимается за оценку $p(f)$. Однако можно показать (см. [27]), что выборочный спектр стационарного временного ряда сильно флуктуирует вокруг теоретического спектра. Интуитивное объяснение этого факта заключается в том, что выборочный спектр соответствует использованию слишком узкого интервала в частотной области. Это аналогично использованию слишком узкого интервала группирования для гистограммы при оценке обычного распределения вероятностей. Используя модифицированную, или *сглаженную*, оценку

$$\hat{p}(f) = 2 \left\{ c_0 + 2 \sum_{k=1}^{N-1} \lambda_k c_k \cos 2\pi f k \right\}, \quad (2.2.14)$$

где λ_k — специально подобранные веса, называемые *корреляционным окном*, можно увеличить «ширину полосы» оценки и получить сглаженную оценку спектра.

На рис. 2.8 показана выборочная оценка спектра данных о партиях продукта. Видно, что дисперсия ряда сконцентрирована в основном на высоких частотах. Это вызвано быстрыми осцилляциями исходного ряда, показанного на рис. 2.1.

2.2.4. Простые примеры автокорреляционных функций и нормированных спектров

Для иллюстрации мы приведем ниже эквивалентные представления двух простых стохастических процессов при помощи

- 1) их теоретических моделей,
- 2) их теоретических автокорреляционных функций,
- 3) их теоретических спектров.

Рассмотрим два процесса

$$z_t = 10 + a_t + a_{t-1}, \quad z_t = 10 + a_t - a_{t-1},$$

где a_t, a_{t-1}, \dots — последовательность некоррелированных, случайных нормальных величин с нулевым средним значением и единичной дисперсией, т. е. белый шум. Используя определение (2.1.5)

$$\gamma_k = \text{cov}[z_t, z_{t+k}] = E[(z_t - \mu)(z_{t+k} - \mu)],$$

где $E[z_t] = E[z_{t+k}] = \mu = 10$, находим, что автоковариации этих двух статистических процессов равны

$$\gamma_k = \begin{cases} 2,0, & k=0 \\ 1,0, & k=1, \\ 0, & k \geq 2 \end{cases} \quad \gamma_k = \begin{cases} 2,0, & k=0 \\ -1,0, & k=1 \\ 0, & k \geq 2. \end{cases}$$

Тогда теоретические автокорреляционные функции равны

$$\rho_k = \begin{cases} 0,5, & k=1, \\ 0,0, & k \geq 2; \end{cases} \quad \rho_k = \begin{cases} -0,5, & k=1, \\ 0,0, & k \geq 2, \end{cases}$$

и, используя (2.2.13), получаем, что теоретические нормированные спектры имеют вид

$$g(f) = 2(1 + \cos 2\pi f), \quad g(f) = 2(1 - \cos 2\pi f).$$

Автокорреляционные функции и нормированные спектры показаны на рис. 2.9 вместе с выборками из временных рядов для каждого процесса.

1. Следует отметить, что для этих двух стационарных процессов знание либо автокорреляционной функции, либо нормированного спектра (со средним значением и дисперсией) эквивалентно знанию модели (при условии, что процесс нормальный).

2. Очевидно, что автокорреляционная функция объясняет некоторые свойства ряда. Сравнительная гладкость первого ряда объясняется положительной связью между последовательными значениями; знакопеременный характер второго ряда объясняется отрицательной связью между последовательными значениями.

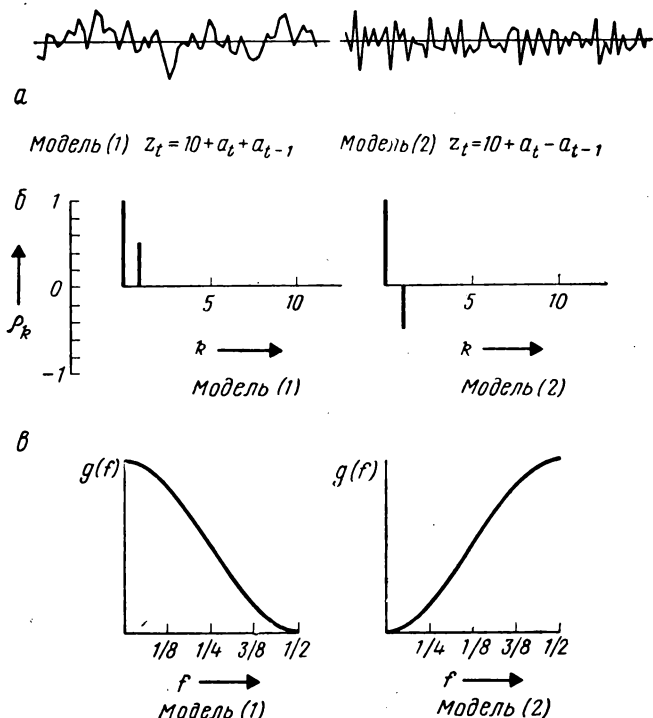


Рис. 2.9. Две простые стохастические модели (а), соответствующие им теоретические автокорреляционные функции (б) и спектральные плотности (в).

3. Спектральная плотность отражает другие, но эквивалентные характеристики ряда. Преобладание низких частот в первом ряде и высоких частот во втором проявляется в соответствующих спектрах.

2.2.5. Преимущества и недостатки автокорреляционных функций и нормированных спектров

Поскольку автокорреляционная функция и нормированный спектр являются взаимными трансформантами, они математически эквивалентны. Поэтому вопрос об их сравнительных преимуществах и недостатках не математический, а связан с тем,

что мы хотим нагляднее представить. Поскольку, как мы видели, каждая из этих функций характеризует разные свойства данных, их следует рассматривать не как соперников, а как союзников. Каждая из них способствует пониманию изучаемого стохастического процесса.

Получение выборочных оценок автокорреляционной функции и спектра — неструктурные *) подходы, аналогичные представлению эмпирической функции распределения гистограммой. Оба эти подхода позволяют данным из стационарных рядов «говорить самим за себя» и являются первым шагом в анализе временных рядов; точно так же гистограмма, являясь первым шагом в вероятностном анализе данных, указывает путь к отысканию подходящей параметрической модели для последующего анализа.

Параметрические модели временных рядов такого типа, как рассмотренные в разд. 2.2.4, не обязательно имеют простую автокорреляционную функцию или простой спектр. Используя любой из этих неструктурных подходов, мы, возможно, будем вынуждены оценивать значения корреляции при большом числе задержек или находить большое число ординат спектра, хотя данные могут быть описаны параметрической моделью всего с одним-двумя параметрами. Поскольку каждое значение корреляции или каждая ордината являются оцениваемыми параметрами, такие неструктурные методы очень расточительны в отношении параметров, в то время как подход, основанный на прямом отыскании модели, может быть более экономичным. С другой стороны, вначале мы, вероятно, не знаем, какова подходящая модель; поэтому использование того или другого неструктурного подхода оказывается необходимым для *идентификации* нужного типа модели. Аналогично, нарисовав гистограмму, легче выбрать подходящий тип распределения. Выбор между спектром и автокорреляционной функцией как средствами построения модели зависит от природы моделей, оказывающихся практически полезными. Те модели, которые мы нашли полезными и которые будут рассмотрены в последующих главах этой книги, проще описываются с помощью автокорреляционной функции. Именно этим инструментом мы пользуемся для идентификации рядов.

ПРИЛОЖЕНИЕ П2.1. СВЯЗЬ МЕЖДУ ВЫБОРОЧНЫМ СПЕКТРОМ И ОЦЕНКОЙ АВТОКОВАРИАЦИОННОЙ ФУНКЦИИ

Ниже будет дан вывод формулы (2.2.8):

$$I(f) = 2 \left\{ c_0 + 2 \sum_{k=1}^{N-1} c_k \cos 2\pi f k \right\}, \quad 0 \leq f \leq 1/2,$$

*) Т. е. непараметрические.

связывающей выборочный спектр $I(f)$ и оценки c_k автоковариационной функции. Пусть выборочные оценки a_f, b_f метода наименьших квадратов для синусоидальной и косинусоидальной компонент ряда на частоте f объединены в $d_f = a_f - ib_f$, где $i = \sqrt{-1}$. Тогда

$$I(f) = \frac{N}{2} (a_f - ib_f)(a_f + ib_f) = \frac{N}{2} d_f d_f^*, \quad (\text{П2.1.1})$$

где d_f^* комплексно сопряжено с d_f . Тогда, используя (2.2.3) и (2.2.4), получим

$$\begin{aligned} d_f &= \frac{2}{N} \sum_{t=1}^N z_t (\cos 2\pi f t - i \sin 2\pi f t) = \\ &= \frac{2}{N} \sum_{t=1}^N z_t e^{-i2\pi f t} = \frac{2}{N} \sum_{t=1}^N (z_t - \bar{z}) e^{-i2\pi f t}. \end{aligned} \quad (\text{П2.1.2})$$

Подставляя (П2.1.2) в (П2.1.1), получаем

$$I(f) = \frac{2}{N} \sum_{t=1}^N \sum_{t'=1}^N (z_t - \bar{z})(z_{t'} - \bar{z}) e^{-i2\pi f (t-t')}. \quad (\text{П2.1.3})$$

Так как

$$c_k = \frac{1}{N} \sum_{t=1}^{N-k} (z_t - \bar{z})(z_{t+k} - \bar{z}),$$

замена $k = t - t'$ преобразует (П2.1.3) в

$$I(f) = 2 \sum_{k=-N+1}^{N-1} c_k e^{-i2\pi f k} = 2 \left\{ c_0 + 2 \sum_{k=1}^{N-1} c_k \cos 2\pi f k \right\}, \quad 0 \leq f \leq 1/2,$$

что и требовалось доказать.

Линейные стационарные модели

Описана общая линейная стохастическая модель, основанная на предположении о том, что временной ряд генерируется совокупностью случайных импульсов. Для практических целей удобно пользоваться моделями, описываемыми малым числом параметров. Часто удается достигнуть экономии, представляя линейный процесс при помощи небольшого числа членов типа авторегрессии или скользящего среднего. Будут рассмотрены свойства результирующих моделей авторегрессии — скользящего среднего (АРСС), используемых в дальнейшем в процедурах построения моделей.

3.1. ОБЩИЙ ЛИНЕЙНЫЙ ПРОЦЕСС

3.1.1. Две эквивалентные формы линейного процесса

В разд. 1.2.1 мы рассматривали представление стохастического процесса как выход линейного фильтра, на вход которого поступает белый шум a_t , т. е.

$$\tilde{z}_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = a_t + \sum_{j=1}^{\infty} \psi_j a_{t-j}, \quad (3.1.1)$$

где $\tilde{z}_t = z_t - \mu$ — отклонение процесса от некоторого начального уровня или, если процесс стационарен, от своего среднего значения. *Общий линейный процесс* (3.1.1) позволяет нам представить \tilde{z}_t как взвешенную сумму настоящего и прошлых значений другого процесса — белого шума*) a_t . Наиболее важная литература по теории линейных стохастических моделей — это [24, 28, 29, 32, 44, 92, 97—100, 102, 103]. Белый шум a_t можно рассматривать как *серию импульсов*, которые приводят в движение систему. Он состоит из последовательности некоррелированных

*) В математической литературе процесс (3.1.1) называется линейным, если a_t независимы, одинаково распределены, имеют конечную дисперсию,

кроме того, $\sum_{j=0}^{\infty} \psi_j^2 < \infty$. — Прим. ред.

случайных переменных с нулевым средним значением и постоянной дисперсией:

$$E[a_t] = 0, \quad \text{var}[a_t] = \sigma_a^2.$$

Так как случайные переменные некоррелированы, их автоковариационная функция должна иметь вид

$$\gamma_k = E[a_t a_{t+k}] = \begin{cases} \sigma_a^2, & k = 0, \\ 0, & k \neq 0. \end{cases} \quad (3.1.2)$$

Поэтому автокорреляционная функция белого шума имеет очень простую форму:

$$\rho_k = \begin{cases} 1, & k = 0, \\ 0, & k \neq 0. \end{cases} \quad (3.1.3)$$

Модель (3.1.1) может быть записана иначе, а именно, как взвешенная сумма прошлых значений \tilde{z} плюс добавочный импульс a_t :

$$\tilde{z}_t = \pi_1 \tilde{z}_{t-1} + \pi_2 \tilde{z}_{t-2} + \dots + a_t = \sum_{j=1}^{\infty} \pi_j \tilde{z}_{t-j} + a_t. \quad (3.1.4)$$

В форме (3.1.4) процесс можно объяснить как регрессию текущего отклонения \tilde{z}_t от уровня μ на прошлые отклонения процесса $\tilde{z}_{t-1}, \tilde{z}_{t-2}, \dots$.

Соотношения между весами ψ и π . Соотношения между весами ψ и π можно получить при помощи введенного ранее оператора сдвига назад

$$Bz_t = z_{t-1}, \quad B^j z_t = z_{t-j}.$$

Позднее нам понадобится также оператор сдвига вперед $F = B^{-1}$, такой, что

$$Fz_t = z_{t+1}, \quad F^j z_t = z_{t+j}.$$

Как пример использования оператора B рассмотрим модель

$$\tilde{z}_t = a_t - \theta a_{t-1} = (1 - \theta B) a_t,$$

в которой $\psi_1 = -\theta$, $\psi_j = 0$ при $j > 1$. Выражая a_t через \tilde{z} , получим

$$(1 - \theta B)^{-1} \tilde{z}_t = a_t.$$

Отсюда

$$a_t = (1 + \theta B + \theta^2 B^2 + \theta^3 B^3 + \dots) \tilde{z}_t,$$

и, выражая отклонение \tilde{z}_t через прошлые отклонения в виде (3.1.4), получаем

$$\tilde{z}_t = -\theta \tilde{z}_{t-1} - \theta^2 \tilde{z}_{t-2} - \theta^3 \tilde{z}_{t-3} - \dots + a_t,$$

так что для этой модели $\pi_j = -\theta^j$,

Вообще (3.1.1) можно записать в виде

$$\tilde{z}_t = \left(1 + \sum_{j=1}^{\infty} \psi_j B^j\right) a_t$$

или

$$\tilde{z}_t = \psi(B) a_t, \quad (3.1.5)$$

где

$$\psi(B) = 1 + \sum_{j=1}^{\infty} \psi_j B^j = \sum_{j=0}^{\infty} \psi_j B^j$$

с $\psi_0 = 1$. Как отмечалось в разд. 1.2.1, $\psi(B)$ называется *передаточной функцией* линейного фильтра, связывающего \tilde{z}_t с a_t . Она может рассматриваться также как *производящая функция* весов ψ , где B следует рассматривать как фиктивную переменную, чья j -я степень есть коэффициент при ψ_j .

Аналогично (3.1.4) можно записать

$$\left(1 - \sum_{j=1}^{\infty} \pi_j B^j\right) \tilde{z}_t = a_t$$

или

$$\pi(B) \tilde{z}_t = a_t. \quad (3.1.6)$$

Тогда

$$\pi(B) = 1 - \sum_{j=1}^{\infty} \pi_j B^j$$

— это производящая функция весов π . Применяя к обеим частям (3.1.6) оператор $\psi(B)$, получим

$$\psi(B) \pi(B) \tilde{z}_t = \psi(B) a_t = \tilde{z}_t,$$

Отсюда

$$\psi(B) \pi(B) = 1$$

и

$$\pi(B) = \psi^{-1}(B). \quad (3.1.7)$$

Соотношение (3.1.7) можно использовать для получения весов π по заданным весам ψ и наоборот.

3.1.2. Производящая функция автоковариаций линейного процесса

Основным инструментом анализа данных для идентификации моделей в гл. 6 будет автокорреляционная функция. Поэтому нам важно знать автокорреляционную функцию линейного процесса. В приложении П3.1 показано, что автоковариационная функция линейного процесса (3.1.1) имеет вид

$$\gamma_k = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k}. \quad (3.1.8)$$

В частности, принимая в (3.1.8) $k = 0$, мы находим, что дисперсия процесса равна

$$\gamma_0 = \sigma_z^2 = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j^2. \quad (3.1.9)$$

Отсюда следует, что если процесс имеет конечную дисперсию, веса ψ_j должны уменьшаться достаточно быстро, для того чтобы ряд в правой части (3.1.9) сходиллся.

Часто более удобным способом получения автоковариаций линейного процесса оказывается использование *производящей функции автоковариаций*

$$\gamma(B) = \sum_{k=-\infty}^{\infty} \gamma_k B^k, \quad (3.1.10)$$

в которой γ_0 (дисперсия процесса) — это коэффициент при $B^0 = 1$, в то время как γ_k (автоковариация для задержки k) — это коэффициент как при B^j , так и при $B^{-j} = F^j$. В приложении ПЗ.1 показано, что

$$\gamma(B) = \sigma_a^2 \psi(B) \psi(B^{-1}) = \sigma_a^2 \psi(B) \psi(F). \quad (3.1.11)$$

Например, положим

$$\tilde{z}_t = a_t - \theta a_{t-1} = (1 - \theta B) a_t,$$

так что $\psi(B) = 1 - \theta B$. Подставляя это в (3.1.11), получим

$$\gamma(B) = \sigma_a^2 (1 - \theta B)(1 - \theta B^{-1}) = \sigma_a^2 [-\theta B^{-1} + (1 + \theta^2) - \theta B].$$

Сравнивая с (3.1.11), находим, что автоковариации равны

$$\gamma_0 = (1 + \theta^2) \sigma_a^2,$$

$$\gamma_1 = -\theta \sigma_a^2,$$

$$\gamma_k = 0, \quad k \geq 2.$$

В дальнейшем B , рассматриваемая как фиктивная переменная в производящей функции, будет принимать и комплексные значения. В частности, иногда необходимо рассматривать отдельно случаи, когда $|B| < 1$, $|B| = 1$ или $|B| > 1$, т. е. когда комплексное число B лежит внутри, на или вне единичной окружности.

3.1.3. Условия стационарности и обратимости линейного процесса

Стационарность. Сходимость рядов (3.1.9) обеспечивает конечное значение дисперсии процесса. Мы видели также в разд. 2.1.3, что автоковариации и автокорреляции должны удовлетво-

рядом ряду условий, обеспечивающих стационарность. Для линейного процесса эти условия могут быть объединены в одно, а именно, что ряд $\psi(B)$ — производящая функция для весов ψ — должен сходиться при $|B| \leq 1$, т. е. для B , лежащих внутри или на единичной окружности. Этот результат обсуждается в приложении ПЗ.1.

Спектр линейного стационарного процесса. В приложении ПЗ.1 показано, что, если мы подставим B в виде $B = e^{-i2\pi f}$, где $i = \sqrt{-1}$, в производящую функцию автоковариаций (3.1.11), то получим уменьшенный в 2 раза спектр мощности. Отсюда спектр линейного процесса равен

$$p(f) = 2\sigma_a^2 \psi(e^{-i2\pi f}) \psi(e^{i2\pi f}) = 2\sigma_a^2 |\psi(e^{-i2\pi f})|^2 \quad 0 \leq f \leq 1/2. \quad (3.1.12)$$

В действительности (3.1.12) — хорошо известное выражение [27], связывающее спектр $p(f)$ выхода линейной системы с равномерным спектром $2\sigma_a^2$ входного белого шума, умноженным на квадрат коэффициента усиления системы $G^2(f) = |\psi(e^{-i2\pi f})|^2$.

Обратимость. Выше было показано, что веса ψ линейного процесса, если он стационарен, должны удовлетворять условию сходимости $\psi(B)$ внутри или на единичной окружности. Рассмотрим теперь ограничения, налагаемые на веса π , для получения свойства, называемого «обратимостью». Условие обратимости не зависит от условий стационарности и применимо также к нестационарным линейным моделям, которые мы введем в гл. 4.

Для иллюстрации смысла понятия обратимости рассмотрим снова модель

$$\tilde{z}_t = (1 - \theta B) a_t. \quad (3.1.13)$$

Выражая a через \tilde{z} , получим из (3.1.13)

$$a_t = (1 - \theta B)^{-1} \tilde{z}_t,$$

т. е.

$$a_t = \tilde{z}_t + \theta \tilde{z}_{t-1} + \theta^2 \tilde{z}_{t-2} + \dots \quad (3.1.14)$$

или

$$\tilde{z}_t = -\theta \tilde{z}_{t-1} - \theta^2 \tilde{z}_{t-2} - \dots + a_t, \quad (3.1.15)$$

и веса π модели в форме (3.1.15) равны $\pi_j = -\theta^j$. При любом значении θ (3.1.13) описывает обычный стационарный процесс. Однако, если $|\theta| \geq 1$, веса в разложении (3.1.14), выражающем a_t через текущее и *прошлые* значения \tilde{z} , образуют расходящийся ряд. Это означает, что текущее отклонение \tilde{z}_t в момент t в (3.1.15) зависит от $\tilde{z}_{t-1}, \tilde{z}_{t-2}, \dots, \tilde{z}_{t-j}, \dots$ с весами, растущими по мере роста j . Мы избегаем этой ситуации, требуя, чтобы веса π_j в «обращенном» разложении (3.1.14) образовывали сходящийся ряд, т. е. чтобы $|\theta| \leq 1$. В этом случае мы

будем называть ряд *обратимым*. Это условие будет, очевидно, выполняться, если ряд

$$\pi(B) = (1 - \theta B)^{-1} = \sum_{j=0}^{\infty} \theta^j B^j$$

сходится при $|B| \leq 1$, т. е. внутри или на единичной окружности.

В гл. 6, где мы рассмотрим вопрос о единственности этих моделей, мы увидим, что сходящееся разложение для a_t возможно и в случае, когда $|\theta| \geq 1$; при этом a_t выражается только через $z_t, z_{t+1}, z_{t+2}, \dots$, т. е. через настоящее и *будущее* значения процесса. Требование обратимости необходимо, если мы заинтересованы в разумной связи текущих событий с событиями в *прошлом*.

В общем линейный процесс

$$\pi(B) \tilde{z}_t = a_t$$

обратим, если веса π_j таковы, что ряд $\pi(B)$ сходится внутри или на единичной окружности.

Подводя итоги, отметим, что линейный процесс *стационарен*, если $\psi(B)$ сходится внутри или на единичной окружности, и *обратим*, если $\pi(B)$ сходится в той же области.

3.1.4. Процессы авторегрессии и скользящего среднего

Представления общего линейного процесса (3.1.1) и (3.1.4) не будут практически полезными, если они содержат неограниченное число параметров ψ_j и π_j . Мы рассмотрим теперь, как добиться экономии и все же сохранить достаточно представительные модели.

Процессы авторегрессии. Рассмотрим частный случай (3.1.4), когда только первые p весов ненулевые. Эта модель может быть записана как

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \dots + \phi_p \tilde{z}_{t-p} + a_t, \quad (3.1.16)$$

где мы используем символы $\phi_1, \phi_2, \dots, \phi_p$ для *конечного* набора весовых параметров. Процесс (3.1.16) называется процессом *авторегрессии* p -го порядка, или сокращенно процессом $AR(p)$. В частности, большое практическое значение имеют процессы авторегрессии первого ($p = 1$) и второго порядка ($p = 2$)

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + a_t,$$

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + a_t.$$

Теперь мы можем записать (3.1.16) в эквивалентной форме

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \tilde{z}_t = a_t$$

или

$$\phi(B)\tilde{z}_t = a_t. \quad (3.1.17)$$

Так как из (3.1.17) вытекает, что

$$\tilde{z}_t = \frac{1}{\phi(B)} a_t = \phi^{-1}(B) a_t,$$

то процесс авторегрессии можно трактовать как выход \tilde{z}_t линейного фильтра с передаточной функцией $\phi^{-1}(B)$, входом которого служит белый шум a_t .

Процессы скользящего среднего. Рассмотрим частный случай (3.1.1), когда только первые q из весов ψ ненулевые. Процесс имеет вид

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}, \quad (3.1.18)$$

где символы $-\theta_1, -\theta_2, \dots, -\theta_q$ используются для обозначения *конечного* набора весовых параметров. Процесс (3.1.18) называется процессом *скользящего среднего* *) порядка q , который иногда будет сокращенно обозначаться $CC(q)$. В частности, особенно важны для практики процессы первого ($q = 1$) и второго порядка ($q = 2$):

$$\tilde{z}_t = a_t - \theta_1 a_{t-1},$$

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}.$$

Мы можем также записать (3.1.18) в эквивалентной форме

$$\tilde{z}_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t,$$

или

$$\tilde{z}_t = \theta(B) a_t. \quad (3.1.19)$$

Отсюда следует, что процесс скользящего среднего можно трактовать как выход \tilde{z}_t линейного фильтра с передаточной функцией $\theta(B)$, на вход которого поступает белый шум a_t .

Смешанные процессы авторегрессии — скользящего среднего. Мы видели в разд. 3.1.1, что *конечный* процесс скользящего среднего

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} = (1 - \theta_1 B) a_t$$

может быть записан как *бесконечный* процесс авторегрессии

$$\tilde{z}_t = -\theta_1 \tilde{z}_{t-1} - \theta_1^2 \tilde{z}_{t-2} - \dots + a_t.$$

Поэтому, если процесс действительно типа $CC(1)$, его представление в виде процесса авторегрессии неэкономично. Аналогично процесс $AR(1)$ не может быть экономично представлен

*) Как отмечалось в гл. 1, термин «скользящее среднее» не вполне точен, поскольку сумма весов не равна 1. Однако этот термин общепринят, и мы будем им пользоваться.

с помощью процесса скользящего среднего. На практике для получения экономичной параметризации иногда бывает необходимо включить в модель как члены, описывающие авторегрессию, так и члены, моделирующие скользящее среднее.

Такой процесс

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \dots + \phi_p \tilde{z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

или

$$\phi(B) \tilde{z}_t = \theta(B) a_t \quad (3.1.20)$$

называется *смешанным процессом авторегрессии — скользящего среднего* порядка (p, q) и иногда сокращенно обозначается нами $APCC(p, q)$. Например, процесс $APCC(1, 1)$ имеет вид

$$\tilde{z}_t - \phi_1 \tilde{z}_{t-1} = a_t - \theta_1 a_{t-1}.$$

Так как (3.1.20) можно записать в виде

$$\tilde{z}_t = \phi^{-1}(B) \theta(B) a_t = \frac{\theta(B)}{\phi(B)} a_t = \frac{1 - \theta_1 B - \dots - \theta_q B^q}{1 - \phi_1 B - \dots - \phi_p B^p} a_t,$$

смешанный процесс авторегрессии — скользящего среднего можно интерпретировать как выход \tilde{z}_t линейного фильтра, передаточная функция которого есть отношение двух полиномов, на вход которого подается белый шум a_t .

В следующих разделах мы обсудим важные характеристики моделей авторегрессии, скользящего среднего и смешанного типа. Мы исследуем их дисперсии, автокорреляционные функции, спектры, а также условия стационарности и обратимости, накладываемые на их параметры.

3.2. ПРОЦЕССЫ АВТОРЕГРЕССИИ

3.2.1. Условия стационарности для процессов авторегрессии

Набор регулируемых параметров $\phi_1, \phi_2, \dots, \phi_p$ процесса $AR(p)$

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \dots + \phi_p \tilde{z}_{t-p} + a_t$$

или

$$(1 - \phi_1 B - \dots - \phi_p B^p) \tilde{z}_t = \phi(B) \tilde{z}_t = a_t$$

должен удовлетворять определенным условиям для того, чтобы процесс был стационарен.

Например, процесс авторегрессии первого порядка

$$(1 - \phi_1 B) \tilde{z}_t = a_t$$

может быть записан в виде

$$\tilde{z}_t = (1 - \phi_1 B)^{-1} a_t = \sum_{j=0}^{\infty} \phi_1^j a_{t-j}.$$

Отсюда

$$\psi(B) = (1 - \phi_1 B)^{-1} = \sum_{j=0}^{\infty} \phi_1^j B^j. \quad (3.2.1)$$

В разд. 3.1.3 указывалось, что стационарность будет выполняться, если $\psi(B)$ сходится при $|B| \leq 1$. Из (3.2.1) вытекает, что параметр ϕ_1 стационарного процесса $AP(1)$ должен подчиняться условию $|\phi_1| < 1$. Так как уравнение $1 - \phi_1 B = 0$ имеет корень $B = \phi_1^{-1}$, это эквивалентно условию, что корень уравнения $1 - \phi_1 B = 0$ должен лежать *вне* единичного круга.

Для произвольного процесса $AP(p)$ $\tilde{z}_t = \phi^{-1}(B) a_t$ мы получаем

$$\phi(B) = (1 - G_1 B)(1 - G_2 B) \dots (1 - G_p B)$$

и, разлагая на элементарные дроби, получим

$$\tilde{z}_t = \phi^{-1}(B) a_t = \sum_{i=1}^p \frac{K_i}{(1 - G_i B)} a_t.$$

Отсюда, для того чтобы ряд $\psi(B) = \phi^{-1}(B)$ был сходящимся при $|B| \leq 1$, необходимо $|G_i| < 1$, где $i = 1, 2, \dots, p$. Это эквивалентно тому, что корни уравнения $\phi(B) = 0$ должны лежать *вне* единичного круга. Корни уравнения $\phi(B) = 0$ можно рассматривать как нули полинома $\phi(B)$. Следовательно, условие стационарности заключается в том, что нули $\phi(B)$ должны лежать *вне* единичного круга. Аналогичные рассуждения применимы в случае, когда не все нули $\phi(B)$ различны. Уравнение $\phi(B) = 0$ называется *характеристическим уравнением* процесса.

Так как ряд

$$\pi(B) = \phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$$

конечен, обратимость процесса авторегрессии обеспечена без каких-либо дополнительных ограничений на его параметры.

3.2.2. Автокорреляционная функция и спектр процессов авторегрессии

Автокорреляционная функция. Важное рекуррентное соотношение для автокорреляционной функции стационарного процесса авторегрессии может быть найдено следующим образом. Умножим

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \dots + \phi_p \tilde{z}_{t-p} + a_t$$

на \tilde{z}_{t-k} и получим

$$\tilde{z}_{t-k} \tilde{z}_t = \phi_1 \tilde{z}_{t-k} \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-k} \tilde{z}_{t-2} + \dots + \phi_p \tilde{z}_{t-k} \tilde{z}_{t-p} + \tilde{z}_{t-k} a_t. \quad (3.2.2)$$

Переходя к математическим ожиданиям величин в (3.2.2), получаем разностное уравнение

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p}, \quad k > 0. \quad (3.2.3)$$

Отметим, что математическое ожидание $E[\tilde{z}_{t-k} a_t]$ равно нулю при $k > 0$, так как \tilde{z}_{t-k} может включать лишь импульсы a_j до момента $t-k$, не коррелированные с a_t . Поделив все члены (3.2.3) на γ_0 , находим, что автокорреляционная функция удовлетворяет аналогичному разностному уравнению

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p}, \quad k > 0. \quad (3.2.4)$$

Заметим, что это уравнение аналогично уравнению, которому удовлетворяет сам процесс \tilde{z}_t .

Запишем теперь (3.2.4) в виде

$$\phi(B) \rho_k = 0,$$

где $\phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$, и B действует на k , но не на t . Тогда, записав $\phi(B)$ в виде

$$\phi(B) = \prod_{i=1}^p (1 - G_i B),$$

получаем общее решение (3.2.4) в виде

$$\rho_k = A_1 G_1^k + A_2 G_2^k + \dots + A_p G_p^k, \quad (3.2.5)$$

где $G_1^{-1}, G_2^{-1}, \dots, G_p^{-1}$ — корни характеристического уравнения

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p = 0.$$

Стационарность требует, чтобы $|G_i| < 1$. Отсюда, если мы предположим, что все корни G_i не кратные, то на практике возможны два случая.

1) Корень G_i действителен, в силу чего член $A_i G_i^k$ в (3.2.5) убывает с ростом номера k как член геометрической прогрессии. Мы будем называть этот случай затухающей экспонентой.

2) Пара корней G_i, G_j комплексно сопряжена, в силу чего они образуют в ρ_k член

$$A d^k \sin(2\pi f k + F)$$

вида затухающей синусоиды.

В общем автокорреляционная функция стационарного процесса авторегрессии состоит из совокупности затухающих экспонент и затухающих синусоид.

Параметры авторегрессии, выраженные через автокорреляции: уравнения Юла — Уокера. Если мы подставим в (3.2.4) значения $k = 1, 2, \dots, p$, то получим систему линейных урав-

нений для $\phi_1, \phi_2, \dots, \phi_p$, со свободными членами $\rho_1, \rho_2, \dots, \rho_p$:

$$\begin{aligned}\rho_1 &= \phi_1 + \phi_2 \rho_1 + \dots + \phi_p \rho_{p-1}, \\ \rho_2 &= \phi_1 \rho_1 + \phi_2 + \dots + \phi_p \rho_{p-2}, \\ &\vdots \\ \rho_p &= \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \dots + \phi_p.\end{aligned}\quad (3.2.6)$$

Они обычно называются уравнениями Юла — Уокера [24, 32]. Оценки Юла — Уокера для параметров процесса получим, заменив теоретические значения автокорреляции ρ_k выборочными автокорреляциями r_k . Если мы перейдем к матричным обозначениям

$$\boldsymbol{\phi} = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{bmatrix}, \quad \boldsymbol{\rho}_p = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{bmatrix}, \quad \mathbf{P}_p = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{p-2} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \dots & 1 \end{bmatrix},$$

решение системы (3.2.6) — выражения для параметров $\boldsymbol{\phi}$ через автокорреляции — можно записать в виде

$$\boldsymbol{\phi} = \mathbf{P}_p^{-1} \boldsymbol{\rho}_p. \quad (3.2.7)$$

Дисперсия. Когда $k=0$, вклад члена $E[z_{t-k}a_t]$ в (3.2.2) (после перехода к математическим ожиданиям) равен $E[a_t^2] = \sigma_a^2$, так как единственный член в z_t , коррелированный с a_t , — это самый последний импульс a_t . Отсюда при $k=0$ $\gamma_0 = \phi_1 \gamma_{-1} + \phi_2 \gamma_{-2} + \dots + \phi_p \gamma_{-p} + \sigma_a^2$. Поделив все члены на $\gamma_0 = \sigma_z^2$ и заменив γ_{-k} на γ_k , получим выражение для дисперсии

$$\sigma_z^2 = \frac{\sigma_a^2}{1 - \rho_1 \phi_1 - \rho_2 \phi_2 - \dots - \rho_p \phi_p}. \quad (3.2.8)$$

Спектр. Для процесса АР(p)

$$\psi(B) = \phi^{-1}(B)$$

и

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p.$$

Отсюда, используя (3.1.12), получим выражение для спектра процесса авторегрессии

$$p(f) = \frac{2\sigma_a^2}{|1 - \phi_1 e^{-i2\pi f} - \phi_2 e^{-i4\pi f} - \dots - \phi_p e^{-i2\pi p f}|^2}, \quad 0 \leq f \leq 1/2. \quad (3.2.9)$$

Рассмотрим теперь два наиболее важных процесса авторегрессии, а именно процессы первого и второго порядка,

3.2.3. Процесс авторегрессии первого порядка (марковский процесс)

Процесс авторегрессии первого порядка имеет вид

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + a_t = a_t + \phi_1 a_{t-1} + \phi_1^2 a_{t-2} + \dots \quad (3.2.10)$$

Как было показано в разд. 3.2.1, для стационарности процесса необходимо, чтобы $-1 < \phi_1 < 1$.

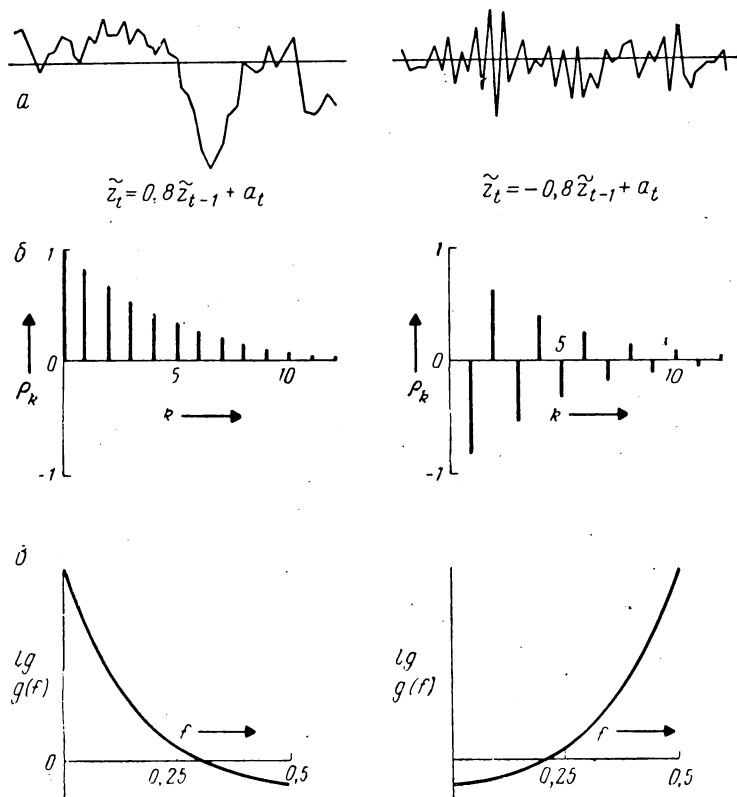


Рис. 3.1. Реализации процессов авторегрессии первого порядка (а), соответствующие им автокорреляционные функции (б) и спектральные плотности (в).

Автокорреляционная функция. Согласно (3.2.4), автокорреляционная функция процесса удовлетворяет разностному уравнению первого порядка

$$\rho_k = \phi_1 \rho_{k-1}, \quad k > 0, \quad (3.2.11)$$

которое при $\rho_0 = 1$ имеет решение

$$\rho_k = \phi_1^k, \quad k \geq 0. \quad (3.2.12)$$

Как показано на рис. 3.1, автокорреляционная функция экспоненциально затухает до нуля монотонно при положительном ϕ_1 , меняя знак, когда ϕ_1 отрицательно.

В частности, отметим, что

$$\rho_1 = \phi_1. \quad (3.2.13)$$

Дисперсия. Согласно (3.2.8), дисперсия процесса имеет вид

$$\sigma_z^2 = \frac{\sigma_a^2}{1 - \rho_1 \phi_1},$$

или, заменяя ρ_1 на ϕ_1 ,

$$\sigma_z^2 = \frac{\sigma_a^2}{1 - \phi_1^2}. \quad (3.2.14)$$

Спектр. Наконец, используя (3.2.9), найдем спектр процесса:

$$p(f) = \frac{2\sigma_a^2}{|1 - \phi_1 e^{-i2\pi f}|^2} = \frac{2\sigma_a^2}{1 + \phi_1^2 - 2\phi_1 \cos 2\pi f}, \quad 0 \leq f \leq \frac{1}{2}. \quad (3.2.15)$$

На рис. 3.1 показаны реализации процесса для случаев $\phi_1 = 0,8$, $\phi_1 = -0,8$, а также соответствующие автокорреляционные функции и спектры. Видно, что в случае большого положительного параметра $\phi_1 = 0,8$ соседние значения ряда близки и наблюдается заметный тренд. Это отражается на виде автокорреляционной функции, которая медленно спадает к нулю по экспоненциальному закону, и на виде спектра, в котором преобладают низкие частоты. В случае же, когда параметр ϕ_1 принимает большое по модулю отрицательное значение $-0,8$, ряд быстро осциллирует. Это находит отражение в поведении автокорреляционной функции, — она экспоненциально спадает до нуля, осциллируя с изменениями знака, — и в форме спектра — в нем преобладают высокие частоты.

3.2.4. Процесс авторегрессии второго порядка

Условие стационарности. Процесс авторегрессии второго порядка записывается в виде

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + a_t. \quad (3.2.16)$$

Для стационарности процесса необходимо, чтобы корни уравнения

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 = 0 \quad (3.2.17)$$

лежали вне единичного круга, т. е. чтобы параметры ϕ_1 и ϕ_2 находились в треугольной области

$$\begin{aligned} \phi_2 + \phi_1 &< 1, \\ \phi_2 - \phi_1 &< 1, \\ -1 < \phi_2 &< 1, \end{aligned} \quad (3.2.18)$$

показанной на рис. 3.2.

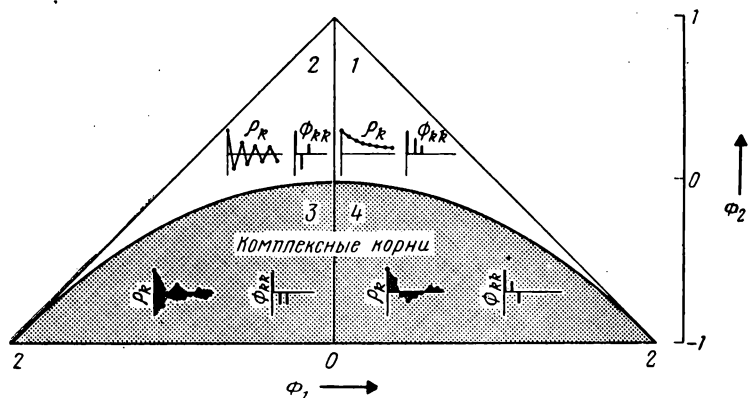


Рис. 3.2. Типичные автокорреляционные и частные автокорреляционные функции ρ_k и ϕ_{kk} для различных стационарных моделей АР(2).

Автокорреляционная функция. Из (3.2.4) получаем разностное уравнение второго порядка, которому удовлетворяет автокорреляционная функция рассматриваемого процесса:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2}, \quad k > 0. \quad (3.2.19)$$

Начальные значения $\rho_0 = 1$ и $\rho_1 = \phi_1 / (1 - \phi_2)$. Из (3.2.5) следует, что общее решение разностного уравнения (3.2.19) имеет вид

$$\rho_k = A_1 G_1^k + A_2 G_2^k = \frac{G_1(1 - G_2^2) G_1^k - G_2(1 - G_1^2) G_2^k}{(G_1 - G_2)(1 + G_1 G_2)}, \quad (3.2.20)$$

причем G_1^{-1} и G_2^{-1} — корни характеристического уравнения (3.2.17). Когда корни действительны, автокорреляционная функция состоит из совокупности затухающих экспонент. Это происходит, когда $\phi_1^2 + 4\phi_2 \geq 0$, и соответствует областям 1 и 2, лежащим выше параболической границы на рис. 3.2 (заимствованном из работы [33]). Конкретнее в области 1 автокорреляционная функция затухает, оставаясь положительной, что соответствует положительному доминирующему корню в (3.2.20). В области 2 затухающая функция автокорреляции знакопеременна, что соответствует отрицательному доминирующему корню.

Если корни G_1 и G_2 комплексные ($\phi_1^2 + 4\phi_2 < 0$), процесс авторегрессии второго порядка ведет себя как *псевдопериодический*. Это поведение отражается на функции автокорреляции, так как заменой $G_1 = de^{i2\pi f_0}$ и $G_2 = de^{-i2\pi f_0}$ в (3.2.20) получаем

$$\rho_k = \frac{\{\text{sign}(\phi_1)\}^k d^k \sin(2\pi f_0 k + F)}{\sin F}, \quad (3.2.21)$$

где $\text{sign}(\phi_1) = +1$, если $\phi_1 > 0$, и $\text{sign}(\phi_1) = -1$, если $\phi_1 < 0$. В любом случае мы называем (3.2.21) *затухающей синусоидой* с параметром затухания d , частотой f_0 и фазой F . Они связаны с параметрами процесса следующим образом:

$$d = \sqrt{-\phi_2} \quad (3.2.22)$$

и имеет тот же знак, что и ϕ_1 ,

$$\cos 2\pi f_0 = \frac{|\phi_1|}{2\sqrt{-\phi_2}}, \quad (3.2.23)$$

$$\text{tg } F = \frac{1 + d^2}{1 - d^2} \text{tg } 2\pi f_0. \quad (3.2.24)$$

Как показано на рис. 3.2, автокорреляционная функция в областях 3 и 4 — затухающая синусоида, причем фазовый угол F меньше 90° в области 4 и лежит между 90 и 180° в области 3. Это означает, что в области 4 автокорреляционная функция вначале (для нескольких первых задержек) положительна, а в области 3 всегда меняет знак при переходе от задержки 0 к задержке 1.

Уравнения Юла — Уокера. Подставляя $p = 2$ в (3.2.6), получим уравнения Юла — Уокера в виде

$$\begin{aligned} \rho_1 &= \phi_1 + \phi_2 \rho_1, \\ \rho_2 &= \phi_1 \rho_1 + \phi_2. \end{aligned} \quad (3.2.25)$$

Решив (3.2.25) относительно ϕ_1 и ϕ_2 , получим

$$\begin{aligned} \phi_1 &= \frac{\rho_1(1 - \rho_2)}{1 - \rho_1^2}, \\ \phi_2 &= \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}. \end{aligned} \quad (3.2.26)$$

Диаграмма В в сборнике таблиц и диаграмм в конце книги позволяет найти значения ϕ_1 и ϕ_2 для любых данных значений ρ_1 и ρ_2 . Эта диаграмма используется в гл. 6 и 7 для получения оценок $\hat{\phi}$ по значениям выборочных автокорреляций r_1 и r_2 .

Уравнения (3.2.25) могут быть решены также относительно ρ_1 и ρ_2 , что дает

$$\begin{aligned}\rho_1 &= \frac{\phi_1}{1 - \phi_2}, \\ \rho_2 &= \phi_2 + \frac{\phi_1^2}{1 - \phi_2}.\end{aligned}\tag{3.2.27}$$

Отсюда получаем начальные значения ρ в (3.2.19), указанные выше. Выражения для представления автокорреляционной функции в форме (3.2.20) или (3.2.21) удобны для объяснения

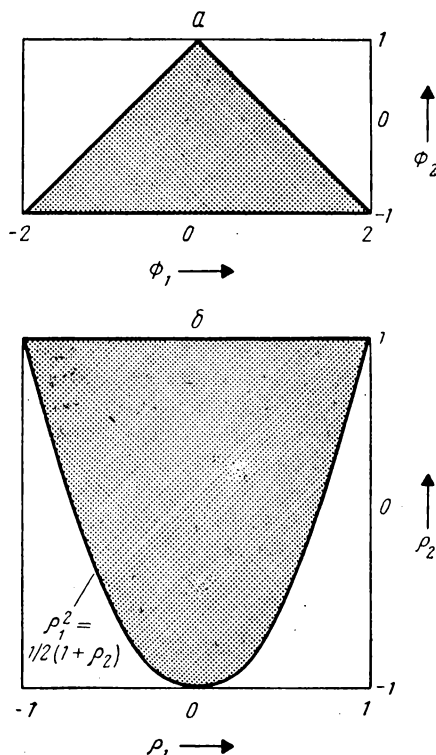


Рис. 3.3. Допустимые области значений a — ϕ_1, ϕ_2 ; b — ρ_1, ρ_2 для стационарного процесса $AP(2)$.

свойств различных типов практически встречающихся процессов. Однако для вычисления автокорреляций процесса $AP(2)$ при данных значениях ϕ_1 и ϕ_2 удобнее пользоваться непосредственно разностным уравнением (3.2.19).

Используя условия стационарности (3.2.18) и выражения (3.2.27) для ρ_1 и ρ_2 , можно показать, что допустимые значения ρ_1

и ρ_2 для стационарного процесса $AP(2)$ должны лежать в области

$$-1 < \rho_1 < 1,$$

$$-1 < \rho_2 < 1,$$

$$\rho_1^2 < \frac{1}{2}(\rho_2 + 1).$$

На рис. 3.3, а показана область допустимых значений параметров ϕ_1 и ϕ_2 , а на рис. 3.3, б — соответствующая область допустимых значений ρ_1 и ρ_2 .

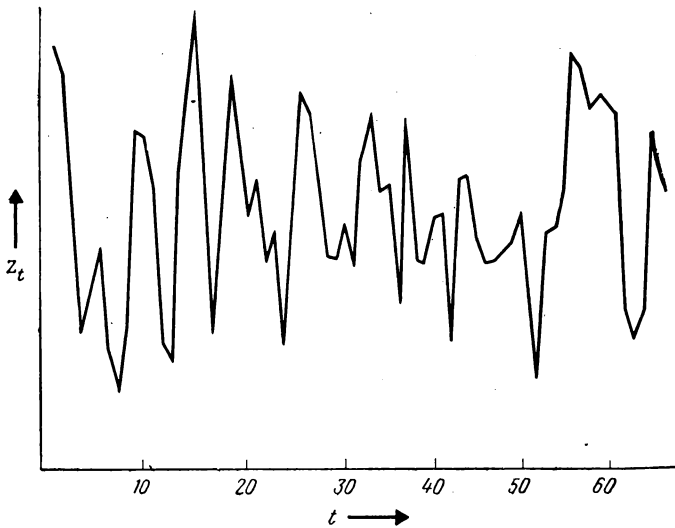


Рис. 3.4. Временной ряд, генерируемый процессом авторегрессии второго порядка.

Дисперсия. Из (3.2.8) следует, что дисперсия процесса равна

$$\sigma_z^2 = \frac{\sigma_a^2}{1 - \rho_1\phi_1 - \rho_2\phi_2} = \left(\frac{1 - \phi_2}{1 + \phi_2} \right) \frac{\sigma_a^2}{(1 - \phi_2)^2 - \phi_1^2}. \quad (3.2.28)$$

Спектр. Согласно (3.2.9), спектр выражается как

$$\begin{aligned} p(f) &= \frac{2\sigma_a^2}{|1 - \phi_1 e^{-i2\pi f} - \phi_2 e^{-i4\pi f}|^2} = \\ &= \frac{2\sigma_a^2}{1 + \phi_1^2 + \phi_2^2 - 2\phi_1(1 - \phi_2) \cos 2\pi f - 2\phi_2 \cos 4\pi f}, \quad 0 \leq f \leq \frac{1}{2}. \end{aligned} \quad (3.2.29)$$

Спектр также отражает псевдопериодическое поведение ряда в случае, когда корни характеристического уравнения комплексны. Для иллюстрации на рис. 3.4 показаны 70 членов ряда,

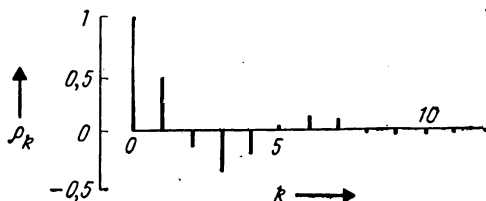


Рис. 3.5. Теоретическая автокорреляционная функция процесса авторегрессии второго порядка $\tilde{z}_t = 0,75\tilde{z}_{t-1} - 0,50\tilde{z}_{t-2} + a_t$.

генерируемого моделью авторегрессии второго порядка

$$\tilde{z}_t = 0,75\tilde{z}_{t-1} - 0,50\tilde{z}_{t-2} + a_t,$$

полученной подстановкой в (3.2.16) $\phi_1 = 0,75$ и $\phi_2 = -0,50$. На рис. 3.5 показана соответствующая теоретическая автокорреляционная функция, вычисленная по формуле (3.2.19) с начальными значениями $\rho_0 = 1$ и $\rho_1 = 0,75/[1 - (-0,5)] = 0,5$. Корни характеристического уравнения

$$1 - 0,75B + 0,5B^2 = 0$$

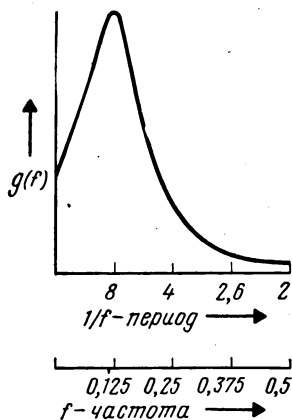
комплексны, так что наблюдаемую псевдопериодичность реализации ряда можно было предугадать: Это поведение четко проявляется в теоретической автокорреляционной функции на рис. 3.5; средний кажущийся период близок к 6.

Параметр затухания d и частота f_0 согласно (3.2.22) и (3.2.23) равны

$$d = \sqrt{0,50} = 0,71;$$

$$f_0 = \frac{\arccos 0,5303}{2\pi} = \frac{1}{6,2}.$$

Рис. 3.6. Теоретическая спектральная плотность процесса авторегрессии второго порядка $\tilde{z}_t = 0,75\tilde{z}_{t-1} - 0,50\tilde{z}_{t-2} + a_t$.



Таким образом, основной период функции автокорреляции равен 6,2.

Наконец, теоретический нормированный спектр, рассчитанный по формуле (3.2.29) и приведенный на рис. 3.6, показывает, что дисперсия ряда обусловлена в основном частотами, близкими к f_0 .

3.2.5. Частная автокорреляционная функция

Вначале мы можем не знать, каков порядок процесса авторегрессии, удовлетворительно описывающего наблюдаемый ряд. Эта задача аналогична задаче о числе независимых переменных, которые нужно учесть при множественной регрессии.

Использование частной автокорреляционной функции для анализа основано на том, что хотя процесс $AP(p)$ имеет бесконечно протяженную функцию автокорреляции, тем не менее он может быть описан при помощи p ненулевых функций от автокорреляций. Обозначим j -й коэффициент процесса авторегрессии порядка k через ϕ_{kj} , так что последний коэффициент будет равен ϕ_{kk} . Согласно (3.2.4), ϕ_{kj} удовлетворяет системе уравнений

$$\rho_j = \phi_{k1}\rho_{j-1} + \dots + \phi_{k(k-1)}\rho_{j-k+1} + \phi_{kk}\rho_{j-k}, \quad (3.2.30)$$

$$j = 1, 2, \dots, k,$$

что приводит к уравнениям Юла — Уокера (3.2.6) в виде

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-2} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \vdots \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \vdots \\ \rho_k \end{bmatrix} \quad (3.2.31)$$

или

$$\mathbf{P}_k \boldsymbol{\phi}_k = \boldsymbol{\rho}_k. \quad (3.2.32)$$

Решая эти уравнения последовательно для $k = 1, 2, 3$, получаем

$$\begin{aligned} \phi_{11} &= \rho_1, \\ \phi_{22} &= \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}, \\ \phi_{33} &= \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}}. \end{aligned} \quad (3.2.33)$$

В общем определитель в числителе состоит из тех же элементов, что и в знаменателе, но последний столбец заменен столбцом $\boldsymbol{\rho}_k$. Величина ϕ_{kk} , рассматриваемая как функция задержки k , называется функцией частной автокорреляции.

Для процесса авторегрессии порядка p частная автокорреляционная функция ϕ_{kk} будет ненулевой для $k \leq p$ и нулем для $k > p$. Другими словами, частная автокорреляционная функция

процесса авторегрессии p -го порядка *обрывается* на задержке, следующей за p . Частные автокорреляционные функции $\hat{\phi}_{kk}$ этого процесса показаны в каждой из четырех областей рис. 3.2.

3.2.6. Оценивание частной автокорреляционной функции

Частные автокорреляции можно оценивать последовательной подгонкой процессов авторегрессии порядка 1, 2, 3, ... методом наименьших квадратов (эта процедура будет описана в гл. 7) и нахождением выборочных оценок $\hat{\phi}_{11}$, $\hat{\phi}_{22}$, $\hat{\phi}_{33}$, ... последнего коэффициента, подгоняемого на каждом этапе. Возможен и другой путь — если значения параметров не слишком близки к границам нестационарности, можно использовать приближенные оценки Юла — Уокера. Выборочные частные автокорреляции можно получить, подставляя значения r_j вместо теоретических автокорреляций (в 3.2.30), что дает

$$r_j = \hat{\phi}_{k1}r_{j-1} + \hat{\phi}_{k2}r_{j-2} + \dots + \hat{\phi}_{k(k-1)}r_{j-k+1} + \hat{\phi}_{kk}r_{j-k}, \quad (3.2.34)$$

$$j = 1, 2, \dots, k,$$

и решая получающиеся уравнения для $k = 1, 2, \dots$. Простой рекуррентный метод такого расчета, предложенный Дарбином [34], приведен в приложении П3.2. Однако оценки, найденные по (3.2.34), очень чувствительны к ошибкам округления и не должны использоваться в случаях, когда значения параметров близки к границам нестационарности. Программы 1 и 3 в сборнике программ для ЭВМ в конце книги включают подпрограммы для вычисления выборочных автокорреляционной и частной автокорреляционной функций.

3.2.7. Стандартные ошибки оценок частной автокорреляции

Кенуй [35] показал (см. также [36] и [37]) в предположении, что порядок процесса авторегрессии равен p , что выборочные частные автокорреляции порядка $p+1$ и выше приближенно независимы и имеют дисперсию

$$\text{var}[\hat{\phi}_{kk}] \approx 1/n, \quad k \geq p+1.$$

Тогда стандартная ошибка (С.О.) выборочной частной автокорреляции $\hat{\phi}_{kk}$ равна

$$\text{С. О.}[\hat{\phi}_{kk}] = 1/\sqrt{n}, \quad k \geq p+1. \quad (3.2.35)$$

В табл. 3.1 приведены первые 15 значений выборочных частных автокорреляций для данных циклического процесса из

Таблица 3.1

Выборочные частные автокорреляционные функции
для процесса из табл. 2.1

k	$\hat{\phi}_{kk}$	k	$\hat{\phi}_{kk}$	k	$\hat{\phi}_{kk}$
1	-0,40	6	-0,15	11	0,18
2	0,19	7	0,05	12	-0,05
3	0,01	8	0,00	13	0,09
4	-0,07	9	-0,10	14	0,18
5	-0,07	10	0,05	15	0,01

табл. 2.1; они получены прямой подгонкой *) процессов авторегрессии возрастающего порядка. Эти частные автокорреляции

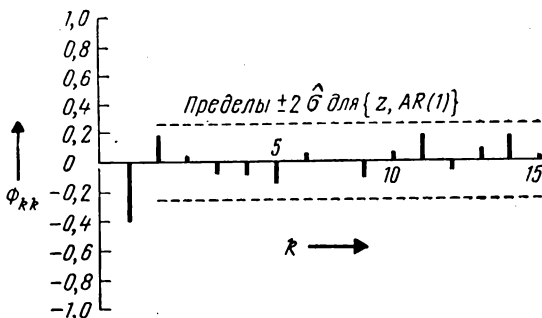


Рис. 3.7. Выборочная частная автокорреляционная функция данных о партиях продукта (рис. 2.1) и пределы, равные удвоенной стандартной ошибке, рассчитанные в предположении, что модель типа $AR(1)$.

показаны на рис. 3.7 и могут быть сопоставлены с автокорреляциями на рис. 2.7. Эти функции имеют примерно такое же поведение, как и для процесса $AR(1)$ при отрицательном ϕ_1 (см. рис. 3.1) или, возможно, для процесса $AR(2)$ с доминирующим отрицательным корнем (см. область 2 рис. 3.2). Пунктиром на рис. 3.7 показаны пределы в 2 С.О. для $\hat{\phi}_{22}$, $\hat{\phi}_{33}$, вычисленные по формулам (3.2.35) в предположении, что процесс есть $AR(1)$. Поскольку $\hat{\phi}_{22}$ — вторая по величине частная автокорреляция из всех рассмотренных, можно рассчитывать на то, что процесс является $AR(2)$. Использование частной автокорреляционной функции для идентификации моделей более детально рассмотрено в гл. 6.

*) Приближенные величины, согласующиеся до первого знака после запятой, могут быть получены из решения уравнения (3.2.34).

3.3. ПРОЦЕССЫ СКОЛЬЗЯЩЕГО СРЕДНЕГО

3.3.1. Условия обратимости для процессов скользящего среднего

Выведем условия, которым должны удовлетворять параметры $\theta_1, \theta_2, \dots, \theta_q$ процесса $CC(q)$

$$\begin{aligned}\tilde{z}_t &= a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} = \\ &= (1 - \theta_1 B - \dots - \theta_q B^q) a_t = \theta(B) a_t, \quad (3.3.1)\end{aligned}$$

чтобы этот процесс был обратимым.

В разд. 3.1.3 было показано, что процесс $CC(1)$

$$\tilde{z}_t = (1 - \theta_1 B) a_t$$

обратим, если $|\theta_1| < 1$, т. е. ряд

$$\pi(B) = (1 - \theta_1 B)^{-1} = \sum_{j=0}^{\infty} \theta_1^j B^j$$

сходится в единичном круге. Однако это эквивалентно утверждению, что корень $B = \theta_1^{-1}$ уравнения $(1 - \theta_1 B) = 0$ лежит *вне* единичного круга.

Условия обратимости для процессов CC более высокого порядка могут быть получены, если записать (3.3.1) как

$$a_t = \theta^{-1}(B) \tilde{z}_t.$$

Отсюда, если

$$\theta(B) = \prod_{j=1}^q (1 - H_j B),$$

то, разлагая на частные дроби $\pi(B) = \theta^{-1}(B)$, получим

$$\pi(B) = \theta^{-1}(B) = \sum_{j=1}^q \frac{M_j}{1 - H_j B}.$$

Этот ряд сходится, если $|H_j| < 1$ при $j = 1, 2, \dots, q$. Так как корни уравнения $\theta(B) = 0$ равны H_j^{-1} , отсюда следует, что условия обратимости процесса $CC(q)$ состоят в том, что корни характеристического уравнения

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q = 0 \quad (3.3.2)$$

лежат *вне* единичного круга. Отметим, что так как ряд

$$\psi(B) = \theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$$

конечен, процесс скользящего среднего является стационарным без каких-либо ограничений на параметры.

3.3.2. Автокорреляционная функция и спектр процесса скользящего среднего

Автокорреляционная функция. Как следует из (3.3.1), автоковариационная функция процесса $CC(q)$ равна

$$\gamma_k = E[(a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}) \times \\ \times (a_{t-k} - \theta_1 a_{t-k-1} - \dots - \theta_q a_{t-k-q})].$$

Следовательно, дисперсия процесса равна

$$\gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma_a^2 \quad (3.3.3)$$

и

$$\gamma_k = \begin{cases} (-\theta_k + \theta_1 \theta_{k+1} + \theta_2 \theta_{k+2} + \dots + \theta_{q-k} \theta_q) \sigma_a^2, & k = 1, 2, \dots, q, \\ 0, & k > q. \end{cases}$$

Отсюда автокорреляционная функция имеет вид

$$\rho_k = \begin{cases} \frac{-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \dots + \theta_q^2}, & k = 1, 2, \dots, q, \\ 0, & k > q. \end{cases} \quad (3.3.4)$$

Мы видим, что автокорреляционная функция процесса $CC(q)$ равна нулю для значений k , больших порядка процесса q . Другими словами, автокорреляционная функция процесса $CC(q)$ обрывается на задержке q .

Параметры скользящего среднего, выраженные через автокорреляции. Если $\rho_1, \rho_2, \dots, \rho_q$ известны, q уравнений (3.3.4) можно разрешить относительно параметров $\theta_1, \theta_2, \dots, \theta_q$. Однако в отличие от линейных уравнений Юла — Уокера (3.2.6) для процесса авторегрессии, уравнения (3.3.4) нелинейны. Поэтому, за исключением простого случая $q = 1$, который мы кратко обсудим здесь, эти уравнения решаются итеративно способом, рассмотренным в приложении П6.2. Подставляя в (3.3.4) вместо ρ_k их оценки r_k и решая получившиеся уравнения, можно найти начальные оценки параметров скользящего среднего. В отличие от соответствующих оценок авторегрессии, полученных при такой же замене ρ_k на r_k в уравнениях Юла — Уокера, результирующие оценки могут не обладать высокой статистической эффективностью. Тем не менее они могут оказаться полезными при грубых оценках параметров на этапе идентификации, рассматриваемом в гл. 6. Кроме того, они могут быть использованы как начальные приближения в итеративной процедуре (рассматриваемой в гл. 7), сходящейся к эффективным оценкам максимального правдоподобия.

Спектр. Для процесса $CC(q)$

$$\psi(B) = \theta(B)$$

и

$$\theta(B) = 1 - \theta_1 B - \theta_1 B^2 - \dots - \theta_q B^q.$$

Используя (3.1.12), получаем отсюда выражение для спектра процесса $CC(q)$

$$p(f) = 2\sigma_a^2 |1 - \theta_1 e^{-i2\pi f} - \theta_2 e^{-i4\pi f} - \dots - \theta_q e^{-i2\pi q f}|^2, \quad (3.3.5)$$

$$0 \leq f \leq \frac{1}{2}.$$

Обсудим теперь подробнее процессы скользящего среднего первого и второго порядков, имеющие большое практическое значение.

3.3.3. Процесс скользящего среднего первого порядка

Мы уже встречались с этим процессом, записанным в форме

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} = (1 - \theta_1 B) a_t.$$

В разд. 3.1.3 было показано, что для обратимости процесса необходимо, чтобы θ_1 находилось в диапазоне $-1 < \theta_1 < 1$. Процесс стационарен, конечно, для любых θ_1 .

Автокорреляционная функция. Используя (3.3.3), получим выражение для дисперсии процесса

$$\gamma_0 = (1 + \theta_1^2) \sigma_a^2$$

и из (3.3.4) — выражение для автокорреляционной функции

$$\rho_k = \begin{cases} \frac{-\theta_1}{1 + \theta_1^2}, & k = 1, \\ 0, & k \geq 2. \end{cases} \quad (3.3.6)$$

Из (3.3.6) для $k = 1$ находим

$$\theta_1^2 + \theta_1/\rho_1 + 1 = 0. \quad (3.3.7)$$

Так как произведение корней есть единица, видно, что если θ_1 — решение, то и θ_1^{-1} — тоже решение. Далее если θ_1 удовлетворяет условию обратимости $|\theta_1| < 1$, то другой корень θ_1^{-1} больше, чем единица, и этому условию не удовлетворяет. Например, если $\rho_1 = -0,4$, (3.3.7) имеет два решения: $\theta_1 = 0,5$ и $\theta_2 = 2,0$, но только решение $\theta_1 = 0,5$ обеспечивает обратимость процесса. Табл. А в сборнике таблиц и диаграмм в конце книги дает соответствующие обратимому процессу значения θ_1 во всем диапазоне возможных значений ρ_1 ($-0,5 < \rho_1 < 0,5$).

Спектр. Согласно (3.3.5), спектр процесса имеет вид

$$p(f) = 2\sigma_a^2 |1 - \theta_1 e^{-i2\pi f}|^2 = 2\sigma_a^2 (1 + \theta_1^2 - 2\theta_1 \cos 2\pi f), \quad 0 \leq f \leq \frac{1}{2}. \quad (3.3.8)$$

В общем когда θ_1 отрицательно, то ρ_1 положительно и в спектре доминируют низкие частоты. Напротив, при θ_1 положительном ρ_1 отрицательно, и в спектре доминируют высокие частоты.

Частная автокорреляционная функция. Из системы (3.2.31) при $\rho_1 = -\theta_1/(1 + \theta_1^2)$ и $\rho_k = 0$ для $k > 1$ после ряда алгебраических преобразований можно получить

$$\phi_{kk} = -\theta_1^k(1 - \theta_1^2)/[1 - \theta_1^{2(k+1)}].$$

Таким образом, $|\phi_{kk}| < \theta_1^k$, и в частной автокорреляционной функции преобладающим членом является затухающая экспонента. Если ρ_1 положительно и, следовательно, θ_1 отрицательно, частная автокорреляционная функция осциллирует с переменами знака. Если же ρ_1 отрицательно и, следовательно, θ_1 положительно, все значения частной автокорреляции отрицательны.

Отметим теперь взаимность процессов $AP(1)$ и $CC(1)$. В то время как автокорреляционная функция процесса $CC(1)$ обрывается после задержки 1, автокорреляционная функция процесса $AP(1)$ экспоненциально затухает с ростом задержки. Обратно, в то время как частная автокорреляционная функция процесса $CC(1)$ затухает и имеет доминирующим членом затухающую экспоненту, частная автокорреляционная функция процесса $AP(1)$ обрывается после задержки 1. Оказывается, что аналогичная приближенная взаимность имеет место и в общем случае.

3.3.4. Процесс скользящего среднего второго порядка

Условия обратимости. Процесс скользящего среднего второго порядка определен как

$$\tilde{z}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}$$

и стационарен для всех значений θ_1 и θ_2 . Однако он обратим только тогда, когда корни характеристического уравнения

$$1 - \theta_1 B - \theta_2 B^2 = 0 \quad (3.3.9)$$

лежат вне единичного круга, т. е.

$$\begin{aligned} \theta_2 + \theta_1 &< 1, \\ \theta_2 - \theta_1 &< 1, \\ -1 &< \theta_2 < 1. \end{aligned} \quad (3.3.10)$$

Эти условия аналогичны условиям (3.2.18) стационарности процесса $AP(2)$.

Автокорреляционная функция. Из (3.3.3) следует, что дисперсия процесса равна

$$\gamma_0 = \sigma_a^2 (1 + \theta_1^2 + \theta_2^2)$$

и из (3.3.4) — что автокорреляционная функция равна

$$\begin{aligned} \rho_1 &= \frac{-\theta_1(1 - \theta_2)}{1 + \theta_1^2 + \theta_2^2}, \\ \rho_2 &= \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2}, \\ \rho_k &= 0, \quad k \geq 3. \end{aligned} \quad (3.3.11)$$

Таким образом, автокорреляционная функция обрывается после задержки 2.

Из (3.3.10) и (3.3.11) вытекает, что первые две автокорреляции обратимого процесса СС(2) должны лежать внутри площади, ограниченной отрезками кривых

$$\begin{aligned} \rho_2 + \rho_1 &= -0,5, \\ \rho_2 - \rho_1 &= -0,5, \\ \rho_1^2 &= 4\rho_2(1 - 2\rho_2). \end{aligned} \quad (3.3.12)$$

Область обратимости (3.3.10) для параметров процесса показана на рис. 3.8, а, и соответствующая область (3.3.12) значений автокорреляций — на рис. 3.8, б. Последний рисунок позволяет оценить, согласуется ли данная пара значений ρ_1, ρ_2 с гипотезой, что модель — процесс СС(2). Если согласие имеется, значения θ_1 и θ_2 можно найти, решив нелинейные уравнения (3.3.11). Для облегчения такого расчета можно использовать диаграмму С в конце книги; она позволяет прямо находить значения θ_1 и θ_2 по данным ρ_1 и ρ_2 .

Спектр. Из (3.3.5) находим спектр

$$\begin{aligned} p(f) &= 2\sigma_a^2 |1 - \theta_1 e^{-i2\pi f} - \theta_2 e^{-i4\pi f}|^2 = \\ &= 2\sigma_a^2 [1 + \theta_1^2 + \theta_2^2 - 2\theta_1(1 - \theta_2) \cos 2\pi f - 2\theta_2 \cos 4\pi f], \quad (3.3.13) \\ 0 &\leq f \leq \frac{1}{2}. \end{aligned}$$

Заметим, что этот спектр с точностью до постоянного множителя $2\sigma_a^2$ обратен спектру (3.2.29) процесса авторегрессии второго порядка.

Частная автокорреляционная функция. Точное выражение для функции частной автокорреляции процесса СС(2) оказывается сложным, но главную роль в нем играет либо сумма двух экспоненциальных членов [если корни характеристического

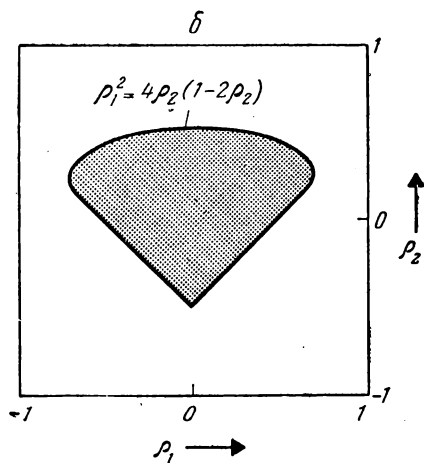
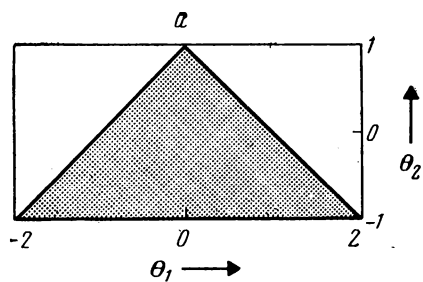


Рис. 3.8. Допустимые области значений a — θ_1, θ_2 ; b — ρ_1, ρ_2 для обратимого процесса $CC(2)$.

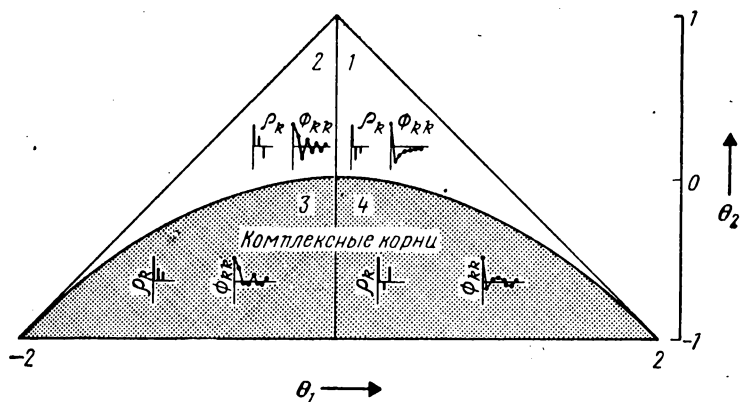


Рис. 3.9. Автокорреляционная и частная автокорреляционная функции ρ_k и ϕ_{kk} для различных моделей $CC(2)$.

уравнения (3.3.9) действительны], либо затухающая синусоида [если корни (3.3.9) комплексны]. Таким образом, эта функция ведет себя так же, как автокорреляция процесса $AP(2)$.

На рис. 3.9 (заимствованном из [33]) показаны автокорреляционные функции (кривые слева) и частные автокорреляционные функции (кривые справа) для различных значений параметров из области обратимости. Сравнение с рис. 3.2, на котором приведены соответствующие автокорреляции и частные автокорреляции, иллюстрирует взаимность процессов $CC(2)$ и $AP(2)$.

3.3.5. Взаимность процессов авторегрессии и скользящего среднего

Результаты предыдущего раздела выявили новые аспекты *взаимности* процессов авторегрессии и скользящего среднего. Как показано в табл. 3.2 в конце этой главы, из этой взаимности вытекает ряд следствий.

1) Для конечного процесса авторегрессии порядка p a_t может быть представлено как *конечная* взвешенная сумма предшествующих \tilde{z} , или \tilde{z}_t может быть представлено как бесконечная взвешенная сумма

$$\tilde{z}_t = \phi^{-1}(B) a_t$$

предшествующих a . Обратно, в конечном процессе скользящего среднего порядка q \tilde{z}_t может быть представлено как конечная взвешенная сумма предшествующих a или a_t — как бесконечная взвешенная сумма

$$\theta^{-1}(B) \tilde{z}_t = a_t$$

предшествующих \tilde{z} .

2) Конечный процесс CC имеет автокорреляционную функцию, обращающуюся в нуль после некоторой точки, но так как он эквивалентен бесконечному процессу AP , его частная автокорреляционная функция бесконечно протяженная. Главную роль в ней играют затухающие экспоненты и (или) затухающие синусоиды. Обратно, процесс AP имеет частную автокорреляционную функцию, обращающуюся в нуль после некоторой точки, но его автокорреляционная функция имеет бесконечную протяженность и состоит из совокупности затухающих экспонент и (или) затухающих синусоид.

3) Параметры процесса авторегрессии конечного порядка p не должны удовлетворять каким-нибудь условиям, для того чтобы этот процесс был обратимым. Однако, чтобы процесс был стационарен, корни $\phi(B) = 0$ должны лежать вне единичного круга. Обратно, параметры процесса CC не должны удовлетворять каким-нибудь условиям для того, чтобы процесс был ста-

ционарным. Однако, для того чтобы процесс СС был обратимым, корни $\theta(B) = 0$ должны лежать вне единичного круга.

4) Спектр процесса скользящего среднего обратен спектру соответствующего процесса авторегрессии.

3.4. СМЕШАННЫЕ ПРОЦЕССЫ АВТОРЕГРЕССИИ — СКОЛЬЗЯЩЕГО СРЕДНЕГО

3.4.1. Свойства стационарности и обратимости

Мы уже заметили в разд. 3.1.4, что для достижения экономичности может оказаться необходимым включить в модель как члены с авторегрессией, так и члены со скользящим средним. Таким образом, мы будем вынуждены использовать смешанную модель авторегрессии — скользящего среднего (АРСС)

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \dots + \phi_p \tilde{z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}, \quad (3.4.1)$$

т. е.

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \tilde{z}_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t$$

или

$$\phi(B) \tilde{z}_t = \theta(B) a_t,$$

где $\phi(B)$ и $\theta(B)$ — полиномы B степени p и q .

В дальнейшем мы будем пользоваться для этого процесса обозначением АРСС (p, q) . Его можно трактовать двумя способами, а именно:

а) как процесс авторегрессии порядка p

$$\phi(B) \tilde{z}_t = e_t,$$

где e_t подчиняется процессу скользящего среднего порядка q

$$e_t = \theta(B) a_t;$$

б) как процесс скользящего среднего порядка q

$$\tilde{z}_t = \theta(B) b_t,$$

где b_t подчиняется процессу авторегрессии порядка p

$$\phi(B) b_t = a_t,$$

так что

$$\phi(B) \tilde{z}_t = \theta(B) \phi(B) b_t = \theta(B) a_t.$$

Очевидно, что члены со скользящим средним в правой части (3.4.1) не повлияют на выводы разд. 3.2.1, определяющие условия стационарности процесса авторегрессии. Поэтому $\phi(B) \tilde{z}_t = \theta(B) a_t$ определяет стационарный процесс при условии, что

все корни характеристического уравнения $\phi(B) = 0$ лежат вне единичного круга. Аналогично, если процесс должен быть обратимым, корни уравнения $\theta(B) = 0$ должны лежать вне единичного круга.

3.4.2. Автокорреляционная функция и спектр смешанных процессов

Автокорреляционная функция. Выражение для автокорреляционной функции смешанного процесса может быть получено способом, аналогичным способу, использованному в разд. 3.2.2 для процессов авторегрессии. Умножив все члены в (3.4.1) на \tilde{z}_{t-k} и перейдя к математическим ожиданиям, мы найдем, что автоковариационная функция удовлетворяет разностному уравнению

$$\gamma_k = \phi_1 \gamma_{k-1} + \dots + \phi_p \gamma_{k-p} + \gamma_{za}(k) - \theta_1 \gamma_{za}(k-1) - \dots - \theta_q \gamma_{za}(k-q), \quad (3.4.2)$$

где $\gamma_{za}(k)$ — это взаимная ковариационная функция z и a , определяемая как $\gamma_{za}(k) = E[\tilde{z}_{t-k} a_t]$. Так как z_{t-k} зависит только от импульсов, которые произошли до момента $t-k$, очевидно, что

$$\gamma_{za}(k) = 0, \quad k > 0; \quad \gamma_{za}(k) \neq 0, \quad k \leq 0.$$

Из (3.4.2) следует

$$\gamma_k = \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + \dots + \phi_p \gamma_{k-p}, \quad k \geq q+1,$$

и

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_p \rho_{k-p}, \quad k \geq q+1, \quad (3.4.3)$$

или

$$\phi(B) \rho_k = 0, \quad k \geq q+1.$$

Это означает, что для процесса АРСС (p, q) существует q автокорреляций $\rho_q, \rho_{q-1}, \dots, \rho_1$, значения которых связаны зависимостью (3.4.2) с q параметрами скользящего среднего θ и p параметрами авторегрессии ϕ .

Далее, p значений $\rho_q, \rho_{q-1}, \dots, \rho_{q-p+1}$ необходимы как начальные значения для решения разностного уравнения $\phi(B) \rho_k = 0, k \geq q+1$, полностью определяющего автокорреляции при больших задержках. Если $q-p < 0$, вся автокорреляционная функция ρ_j для $j = 0, 1, 2, \dots$ будет состоять из совокупности затухающих экспонент и (или) затухающих синусоид, и ее свойства определяются полиномом $\phi(B)$ и начальными значениями. Если же $q-p \geq 0$, имеется $q-p+1$ начальных значений $\rho_0, \rho_1, \dots, \rho_{q-p}$, не укладывающихся в эту общую картину. Эти факты важны при идентификации смешанных рядов.

Дисперсия. При $k = 0$ получаем уравнение

$$\gamma_0 = \phi_1 \gamma_1 + \dots + \phi_p \gamma_p + \sigma_a^2 - \theta_1 \gamma_{za}(-1) - \dots - \theta_q \gamma_{za}(-q), \quad (3.4.4)$$

решая которое вместе с p уравнениями (3.4.2) для $k = 1, 2, \dots, p$, найдем $\gamma_0, \gamma_1, \dots, \gamma_p$.

Спектр. Из (3.1.12) найдем выражение для спектра смешанного процесса

$$\begin{aligned} p(f) &= 2\sigma_a^2 \frac{|\theta(e^{-i2\pi f})|^2}{|\phi(e^{-i2\pi f})|^2} = \\ &= 2\sigma_a^2 \frac{|1 - \theta_1 e^{-i2\pi f} - \dots - \theta_q e^{-i2\pi q f}|^2}{|1 - \phi_1 e^{-i2\pi f} - \dots - \phi_p e^{-i2\pi p f}|^2}, \quad 0 \leq f \leq \frac{1}{2}. \end{aligned} \quad (3.4.5)$$

Частная автокорреляционная функция. Процесс (3.4.1) можно записать в виде

$$a_t = \theta^{-1}(B) \phi(B) \tilde{z}_t,$$

где $\theta^{-1}(B)$ — бесконечный ряд из степеней B . Отсюда частная автокорреляционная функция бесконечна по протяженности. При больших задержках она ведет себя как частная автокорреляционная функция чистого процесса скользящего среднего; в ней преобладают члены типа затухающих экспонент и (или) затухающих синусоид, соотношение между которыми зависит от порядка скользящего среднего и значений параметров процесса.

3.4.3. Процесс авторегрессии первого порядка — скользящего среднего первого порядка

Важный для практики смешанный процесс авторегрессии первого порядка — скользящего среднего первого порядка АРСС(1,1) описывается формулой

$$\tilde{z}_t - \phi_1 \tilde{z}_{t-1} = a_t - \theta_1 a_{t-1}, \quad (3.4.6)$$

или, что равносильно,

$$(1 - \phi_1 B) \tilde{z}_t = (1 - \theta_1 B) a_t.$$

Выведем некоторые важнейшие свойства этого процесса.

Условия стационарности и обратимости. Прежде всего заметим, что процесс стационарен, если $-1 < \phi_1 < 1$, и обратим, если $-1 < \theta_1 < 1$. Обусловленная этим область допустимых значений параметров показана на рис. 3.10, а.

Автокорреляционная функция. Из (3.4.2) и (3.4.4) мы получаем

$$\gamma_0 = \phi_1 \gamma_1 + \sigma_a^2 - \theta_1 \gamma_{za}(-1),$$

$$\gamma_1 = \phi_1 \gamma_0 - \theta_1 \sigma_a^2,$$

$$\gamma_k = \phi_1 \gamma_{k-1}, \quad k \geq 2.$$

Умножив все члены (3.4.6) на α_{t-1} и переходя к математическим ожиданиям, получим

$$\gamma_{za}(-1) = (\phi_1 - \theta_1) \sigma_a^2.$$

Отсюда автоковариационная функция процесса равна

$$\begin{aligned} \gamma_0 &= \frac{1 + \theta_1^2 - 2\phi_1\theta_1}{1 - \phi_1^2} \sigma_a^2, \\ \gamma_1 &= \frac{(1 - \phi_1\theta_1)(\phi_1 - \theta_1)}{1 - \phi_1^2} \sigma_a^2, \\ \gamma_k &= \phi_1 \gamma_{k-1}, \quad k \geq 2. \end{aligned} \quad (3.4.7)$$

Отсюда следует, что автокорреляционная функция экспоненциально убывает от начального значения ρ_1 , зависящего от θ_1 и ϕ_1 *). Как показано на рис. 3.11, это затухание монотонное, если ϕ_1 положительно, и колебательное, если ϕ_1 отрицательно. Далее, знак ρ_1 определяется знаком $(\phi_1 - \theta_1)$ и указывает, происходит ли затухание в области положительных или отрицательных значений.

Пользуясь (3.4.7), можно выразить две первые автокорреляции через параметры процесса, а именно

$$\begin{aligned} \rho_1 &= \frac{(1 - \phi_1\theta_1)(\phi_1 - \theta_1)}{1 + \theta_1^2 - 2\phi_1\theta_1}, \\ \rho_2 &= \phi_1 \rho_1. \end{aligned} \quad (3.4.8)$$

Диаграмма D в конце книги построена так, что позволяет находить решения уравнений (3.4.8) для ϕ_1 и θ_1 по заданным ρ_1 и ρ_2 . Используя оценки r_1 и r_2 вместо ρ_1 и ρ_2 , можно получить таким образом исходные выборочные оценки параметров ϕ_1 и θ_1 .

Из (3.4.8) и условий стационарности и обратимости вытекает, что ρ_1 и ρ_2 должны лежать в области

$$\begin{aligned} |\rho_2| &< |\rho_1|, \\ \rho_2 &> \rho_1(2\rho_1 + 1), \quad \rho_1 < 0, \\ \rho_2 &> \rho_1(2\rho_1 - 1), \quad \rho_1 > 0. \end{aligned} \quad (3.4.9)$$

*) Автокорреляционная функция процесса AP(1) экспоненциально убывает от начального значения $\rho_0 \equiv 1$.

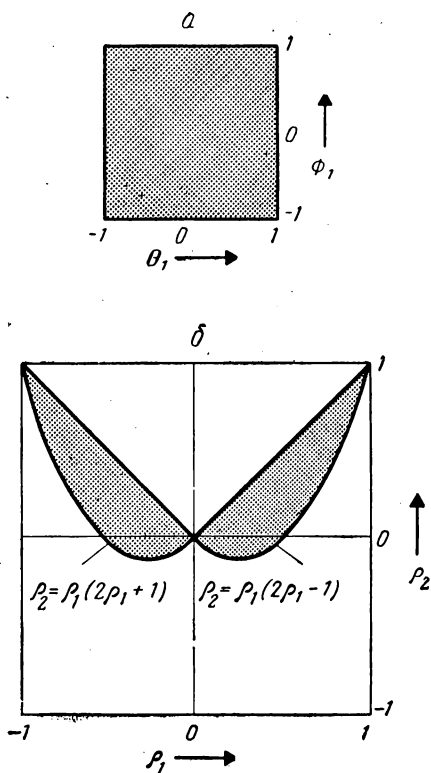


Рис. 3.10. Допустимые области значений a — θ_1, ϕ_1 ; b — ρ_1, ρ_2 для стационарного и обратимого процессов $ARCC(1,1)$.

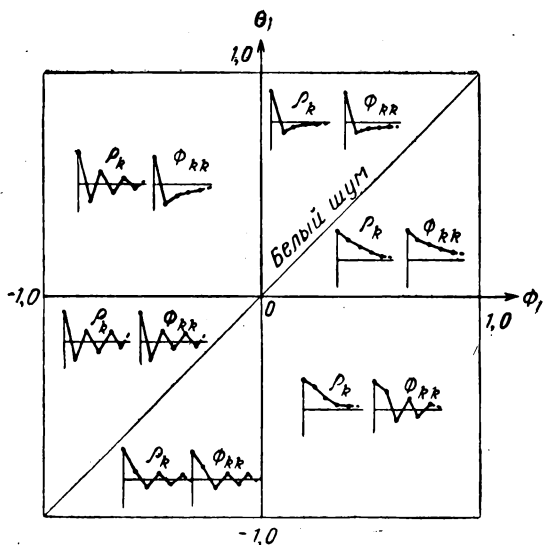


Рис. 3.11. Автокорреляционные и частные автокорреляционные функции ρ_k и ϕ_{kk} для различных моделей $ARCC(1,1)$.

Область допустимых значений ρ_1 и ρ_2 показана на рис. 3.10, б, она ограничивает диапазон допустимых комбинаций ρ_1 и ρ_2 для стационарного обратимого процесса АРСС(1,1).

Частная автокорреляционная функция. Частная автокорреляционная функция смешанного процесса АРСС(1,1), описываемого (3.4.6), состоит только из одного начального значения $\phi_{11} = \rho_1$. Поэтому она ведет себя подобно функции частной автокорреляции чистого процесса СС(1), и в ней преобладает затухающий экспоненциальный член. Как показано на рис. 3.11, при положительном θ_1 она монотонно затухает от значения ρ_1 , знак которого совпадает со знаком $(\phi_1 - \theta_1)$. При отрицательном θ_1 в ней преобладает экспоненциально затухающий осциллирующий член, амплитуда которого убывает от значения ρ_1 . Знак ρ_1 определяется знаком $(\phi_1 - \theta_1)$.

3.3.4. Резюме

На рис. 3.12 приведены допустимые области значений параметров и корреляций ρ_1 , ρ_2 для процессов АР(2), СС(2) и АРСС(1,1).

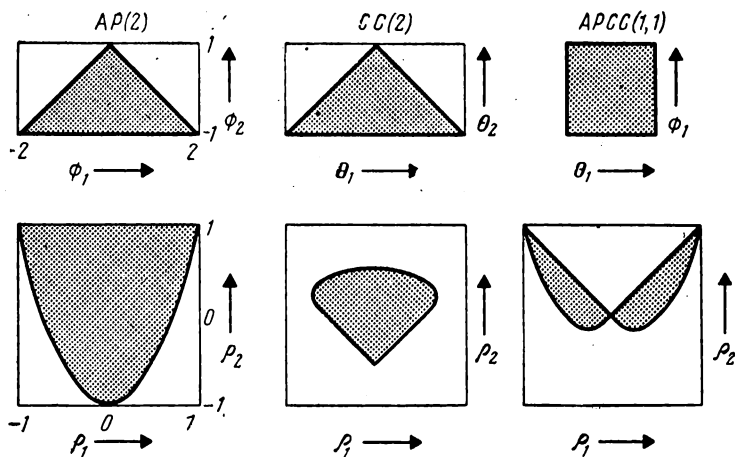


Рис. 3.12. Допустимые области значений параметров и величин ρ_1 , ρ_2 для стационарных обратимых процессов АР(2), СС(2) и АРСС(1,1).

АРСС(1,1) при условии, что они стационарны и обратимы. В табл. 3.2 дана сводная характеристика смешанных процессов авторегрессии — скользящего среднего и приведены также все важные результаты для процессов авторегрессии, скользящего среднего и смешанных процессов, которые будут нужны в гл. 6

Таблица 3.2

Сводка характеристик процессов авторегрессии, скользящего среднего и смешанных (АРСС) процессов

	Процессы авторегрессии	Процессы скользящего среднего	Смешанные процессы
Модель, выраженная через прошлые значения \tilde{z}	$\phi(B) \tilde{z}_t = a_t$	$\theta^{-1}(B) \tilde{z}_t = a_t$	$\theta^{-1}(B) \phi(B) \tilde{z}_t = a_t$
Модель, выраженная через прошлые значения a	$\tilde{z}_t = \phi^{-1}(B) a_t$	$\tilde{z}_t = \theta(B) a_t$	$\tilde{z}_t = \phi^{-1}(B) \theta(B) a_t$
π -веса	Конечный ряд	Бесконечный ряд	Бесконечный ряд
ϕ -веса	Бесконечный ряд	Конечный ряд	Бесконечный ряд
Условие стационарности	Корни $\phi(B) = 0$ лежат вне единичного круга	Всегда стационарен	Корни $\phi(B) = 0$ лежат вне единичного круга
Условие обратимости	Всегда обратима	Корни $\theta(B) = 0$ лежат вне единичного круга	Корни $\theta(B) = 0$ лежат вне единичного круга
Автокорреляционная функция	Бесконечна [затухающие экспоненты и (или) экспоненциально затухающие синусоиды]	Конечна (обрывается)	Бесконечна [после первых $q - p$ задержек затухающие экспоненты и (или) затухающие синусоиды]
Частная автокорреляционная функция	Конечна (обрывается)	Бесконечна [преобладают затухающие экспоненты и (или) экспоненциально затухающие синусоиды]	Бесконечна [после первых $p - q$ задержек преобладают затухающие экспоненты и (или) экспоненциально затухающие синусоиды]

для идентификации моделей по наблюдаемым временным рядам. В следующей главе мы обобщим модель АРСС и получим модели, которые могут описывать часто встречающийся на практике тип нестационарного поведения.

ПРИЛОЖЕНИЕ ПЗ.1. АВТОКОВАРИАЦИИ. ПРОИЗВОДЯЩАЯ ФУНКЦИЯ АВТОКОВАРИАЦИЙ И УСЛОВИЯ СТАЦИОНАРНОСТИ ОБЩЕГО ЛИНЕЙНОГО ПРОЦЕССА

Автоковариации. При задержке k автоковариация линейного процесса

$$\tilde{z}_t = \sum_{j=0}^{\infty} \psi_j a_{t-j}$$

с $\psi_0 = 1$, очевидно, равна

$$\begin{aligned} \gamma_k &= E[\tilde{z}_t \tilde{z}_{t+k}] = \\ &= E\left[\sum_{j=0}^{\infty} \sum_{h=0}^{\infty} \psi_j \psi_h a_{t-j} a_{t+k-h}\right] = \sigma_a^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+k}. \end{aligned} \quad (\text{ПЗ.1.1})$$

Здесь использовано свойство (3.1.2) автоковариационной функции белого шума.

Производящая функция автоковариации. Результат (ПЗ.1.1) может быть подставлен в выражение для производящей функции автоковариации

$$\gamma(B) = \sum_{k=-\infty}^{\infty} \gamma_k B^k, \quad (\text{ПЗ.1.2})$$

что приводит к

$$\gamma(B) = \sigma_a^2 \sum_{k=-\infty}^{\infty} \sum_{j=0}^{\infty} \psi_j \psi_{j+k} B^k = \sigma_a^2 \sum_{j=0}^{\infty} \sum_{k=-j}^{\infty} \psi_j \psi_{j+k} B^k,$$

так как $\psi_h = 0$ при $h < 0$. Обозначим $j+k=h$, так что $k=h-j$. Тогда

$$\gamma(B) = \sigma_a^2 \sum_{j=0}^{\infty} \sum_{h=0}^{\infty} \psi_j \psi_h B^{h-j} = \sigma_a^2 \sum_{h=0}^{\infty} \psi_h B^h \sum_{j=0}^{\infty} \psi_j B^{-j}$$

или

$$\gamma(B) = \sigma_a^2 \psi(B) \psi(B^{-1}) = \sigma_a^2 \psi(B) \psi(F), \quad (\text{ПЗ.1.3})$$

что совпадает с приведенной в тексте формулой (3.1.11).

Условия стационарности. Если мы сделаем замену $B = e^{-i2\pi f}$ и учтем, что $F = B^{-1} = e^{i2\pi f}$, то из выражения (ПЗ.1.2) для производящей функции автоковариации мы найдем уменьшенный в 2 раза спектр мощности. Отсюда спектр мощности линейного

процесса имеет вид

$$p(f) = 2\sigma_a^2 \psi(e^{-i2\pi f}) \psi(e^{i2\pi f}) = 2\sigma_a^2 |\psi(e^{-i2\pi f})|^2, \quad 0 \leq f \leq \frac{1}{2}. \quad (\text{ПЗ.1.4})$$

Следовательно, дисперсия процесса равна

$$\sigma_z^2 = \int_0^{1/2} p(f) df = 2\sigma_a^2 \int_0^{1/2} \psi(e^{-i2\pi f}) \psi(e^{i2\pi f}) df. \quad (\text{ПЗ.1.5})$$

Можно показать [98], что для сходимости интеграла (ПЗ.1.5) требуется сходимость бесконечного ряда $\psi(B)$ на единичном круге.

ПРИЛОЖЕНИЕ ПЗ.2. РЕКУРРЕНТНЫЙ МЕТОД ВЫЧИСЛЕНИЯ ОЦЕНОК ПАРАМЕТРОВ АВТОРЕГРЕССИИ

Покажем теперь, что оценки Юла — Уокера для параметров процесса $\text{AP}(p+1)$ можно получить, если известны оценки для процесса $\text{AP}(p)$, подогнанного к тому же наблюдаемому временному ряду. Как было описано в разд. 3.2.6, этот рекуррентный метод можно использовать для получения приближенной частной автокорреляционной функции.

Чтобы проиллюстрировать этот метод, рассмотрим уравнение (3.2.34). Оценки Юла — Уокера получены для случая $k = 2, 3$ из

$$\begin{aligned} r_2 &= \hat{\phi}_{21}r_1 + \hat{\phi}_{22}, \\ r_1 &= \hat{\phi}_{21} + \hat{\phi}_{22}r_1 \end{aligned} \quad (\text{ПЗ.2.1})$$

и

$$\begin{aligned} r_3 &= \hat{\phi}_{31}r_2 + \hat{\phi}_{32}r_1 + \hat{\phi}_{33}, \\ r_2 &= \hat{\phi}_{31}r_1 + \hat{\phi}_{32} + \hat{\phi}_{33}r_1, \\ r_1 &= \hat{\phi}_{31} + \hat{\phi}_{32}r_1 + \hat{\phi}_{33}r_2. \end{aligned} \quad (\text{ПЗ.2.2})$$

Коэффициенты $\hat{\phi}_{31}$ и $\hat{\phi}_{32}$ можно выразить через $\hat{\phi}_{33}$, используя последние два уравнения (ПЗ.2.2). Решение можно записать в матричном виде

$$\begin{pmatrix} \hat{\phi}_{31} \\ \hat{\phi}_{32} \end{pmatrix} = \mathbf{R}_2^{-1} \begin{pmatrix} r_2 - \hat{\phi}_{33}r_1 \\ r_1 - \hat{\phi}_{33}r_2 \end{pmatrix}, \quad (\text{ПЗ.2.3})$$

где

$$\mathbf{R}_2 = \begin{bmatrix} r_1 & 1 \\ 1 & r_1 \end{bmatrix}.$$

Теперь (ПЗ.2.3) можно записать иначе:

$$\begin{bmatrix} \hat{\phi}_{31} \\ \hat{\phi}_{32} \end{bmatrix} = \mathbf{R}_2^{-1} \begin{bmatrix} r_2 \\ r_1 \end{bmatrix} - \hat{\phi}_{33} \mathbf{R}_2^{-1} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix}. \quad (\text{ПЗ.2.4})$$

Используя то, что (ПЗ.2.1) теперь можно представить в виде

$$\begin{bmatrix} \hat{\phi}_{21} \\ \hat{\phi}_{22} \end{bmatrix} = \mathbf{R}_2^{-1} \begin{bmatrix} r_2 \\ r_1 \end{bmatrix},$$

получаем из (ПЗ.2.4)

$$\begin{bmatrix} \hat{\phi}_{31} \\ \hat{\phi}_{32} \end{bmatrix} = \begin{bmatrix} \hat{\phi}_{21} \\ \hat{\phi}_{22} \end{bmatrix} - \hat{\phi}_{33} \begin{bmatrix} \hat{\phi}_{22} \\ \hat{\phi}_{21} \end{bmatrix}$$

или

$$\begin{aligned} \hat{\phi}_{31} &= \hat{\phi}_{21} - \hat{\phi}_{33} \hat{\phi}_{22}, \\ \hat{\phi}_{32} &= \hat{\phi}_{22} - \hat{\phi}_{33} \hat{\phi}_{21}. \end{aligned} \quad (\text{ПЗ.2.5})$$

Чтобы завершить вычисление $\hat{\phi}_{31}$ и $\hat{\phi}_{32}$, нам нужно знать $\hat{\phi}_{33}$. Подставляя (ПЗ.2.5) в первое из уравнений (ПЗ.2.2), получим

$$\hat{\phi}_{33} = \frac{r_3 - \hat{\phi}_{21}r_2 - \hat{\phi}_{22}r_1}{1 - \hat{\phi}_{21}r_1 - \hat{\phi}_{22}r_2}. \quad (\text{ПЗ.2.6})$$

Таким образом, частная автокорреляция $\hat{\phi}_{33}$ вычисляется вначале по $\hat{\phi}_{21}$ и $\hat{\phi}_{22}$ с помощью (ПЗ.2.6), а затем два других коэффициента $\hat{\phi}_{31}$ и $\hat{\phi}_{32}$ могут быть найдены из (ПЗ.2.5).

В общем случае рекуррентные формулы, заимствованные у Дарбина [34], имеют вид

$$\hat{\phi}_{p+1,j} = \hat{\phi}_{pj} - \hat{\phi}_{p+1,p+1} \hat{\phi}_{p,p-j+1}, \quad j = 1, 2, \dots, p, \quad (\text{ПЗ.2.7})$$

$$\hat{\phi}_{p+1,p+1} = \frac{r_{p+1} - \sum_{j=1}^p \hat{\phi}_{pj} r_{p+1-j}}{1 - \sum_{j=1}^p \hat{\phi}_{pj} r_j}. \quad (\text{ПЗ.2.8})$$

Пример. В качестве примера рассмотрим вычисление оценок $\hat{\phi}_{31}$, $\hat{\phi}_{32}$ и $\hat{\phi}_{33}$ — параметров процесса АР(3), подогнанного к числам Вольфа для солнечных пятен. Выборочные автокорреляции с точностью до трех значащих цифр после запятой равны $r_1 =$

$= 0,806; r_2 = 0,428; r_3 = 0,070$. Тогда

$$\hat{\phi}_{21} = \frac{r_1(1-r_2)}{1-r_1^2} = 1,316,$$

$$\hat{\phi}_{22} = \frac{r_2-r_1^2}{1-r_1^2} = -0,632.$$

Из (ПЗ.2.6) получим

$$\hat{\phi}_{33} = \frac{0,070 - (1,316 \cdot 0,428 + 0,632 \cdot 0,806)}{1 - (1,316)(0,806) + (0,632)(0,428)} = 0,077.$$

Подставляя значения $\hat{\phi}_{21}, \hat{\phi}_{22}, \hat{\phi}_{33}$ в (ПЗ.2.5), получим

$$\hat{\phi}_{31} = 1,316 + (0,077)(0,632) = 1,365,$$

$$\hat{\phi}_{32} = -0,632 - (0,077)(1,316) = -0,733.$$

Напомним читателю, что эти оценки частных автокорреляции несколько отличаются от оценок максимального правдоподобия, полученных подгонкой процессов авторегрессии последовательно растущих порядков. Они очень чувствительны к ошибкам округления, в особенности когда процесс близок к нестационарному.

Линейные нестационарные модели

Многие эмпирические временные ряды (например, цены на бирже) ведут себя так, как будто они не имеют фиксированного среднего значения. Но даже при этом они выглядят однородными в том смысле, что если не учитывать локальный уровень или, возможно, локальный уровень и тренд, любая часть временного ряда по своему поведению во многом подобна любой другой. Модели, описывающие такое однородное нестационарное поведение, можно получить, предположив, что некая подходящая *разность* процесса стационарна. Рассмотрим теперь свойства важнейшего класса моделей, в которых d -я разность есть стационарный смешанный процесс авторегрессии — скользящего среднего. Эти модели называются процессами авторегрессии — проинтегрированного скользящего среднего (АРПСС).

4.1. ПРОЦЕССЫ АВТОРЕГРЕССИИ — ПРОИНТЕГРИРОВАННОГО СКОЛЬЗЯЩЕГО СРЕДНЕГО

4.1.1. Нестационарность процесса авторегрессии первого порядка

На рис. 4.1 показаны участки четырех временных рядов, полученных практически. Эти ряды наблюдались в задачах прогнозирования и контроля и уже по виду кажутся нестационарными. Ряды А, С и D представляют собой «нерегулируемые» выходные данные (концентрацию, температуру и вязкость) трех различных химических процессов. Эти ряды были выбраны, чтобы показать, как на выходные данные влияют такие нерегулируемые и неизмеряемые возмущения, как вариации расхода сырья и окружающей температуры. Температурный ряд С был получен временным отсоединением регуляторов на опытном заводе с записью последующих температурных флуктуаций. Оба ряда А и D получены в промышленных процессах, где необходимо было поддерживать некую выходную характеристику качества продукта как можно ближе к заданному уровню. Чтобы достигнуть такого регулирования, управляли другой переменной

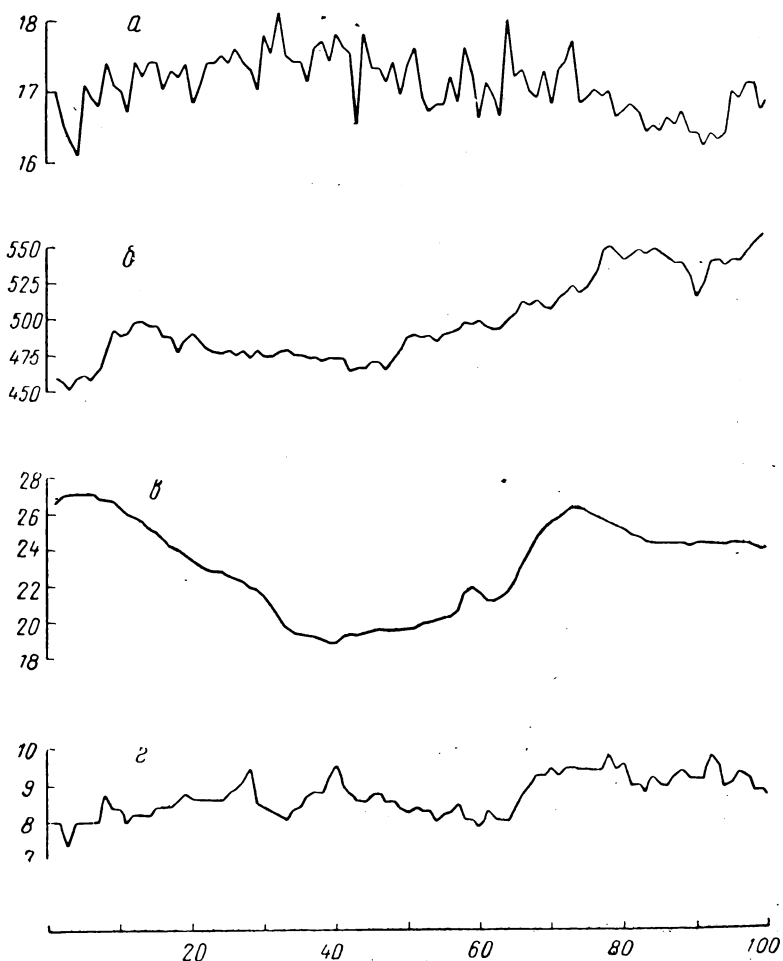


Рис. 4.1. Некоторые типичные временные ряды, встречающиеся в задачах прогноза и регулирования.

a — ряд А. «Нерегулируемая» концентрация, отсчет каждые два часа, химический процесс.

б — ряд В. Биржевые цены акций IBM, ежедневные данные.

в — ряд С. «Нерегулируемая» температура, отсчет каждую минуту, химический процесс.

г — ряд D. «Нерегулируемая» вязкость, отсчет каждый час, химический процесс.

процесса — это позволяло приближенно устранять вариации выхода. Однако эффект этих регулирующих действий на выход в каждом случае был точно известен, так что можно было ввести численную поправку за регулирующее воздействие. Это означает, что можно было с хорошей точностью рассчитать значения ряда, которые могли бы быть получены при отсутствии

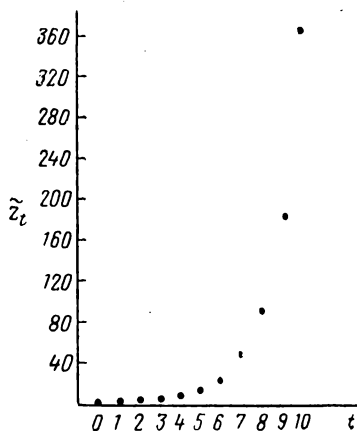


Рис. 4.2. Реализация нестационарного процесса авторегрессии первого порядка.

корректирующих действий. Здесь показаны именно эти компенсированные величины; они в дальнейшем называются «нерегулируемыми» рядами. Ряд В состоит из суточных цен акций компании IBM за период, начинающийся с мая 1961 г. Полные данные по каждому ряду приведены в сборнике временных рядов в конце этой книги. На рис. 4.1 показаны по 100 последовательных наблюдений каждого ряда; точки соединены прямыми линиями.

Существует неограниченное число различных проявлений нестационарности процесса. Однако те типы экономических и промышленных рядов, которые мы хотим анализировать, часто проявляют весьма специфическую однородную нестационарность. Такая нестационарность может быть представлена стохастической моделью, являющейся модифицированной формой модели АРСС. В гл. 3 была рассмотрена модель авторегрессии — скользящего среднего

$$\phi(B) \tilde{z}_t = \theta(B) a_t, \quad (4.1.1)$$

где $\phi(B)$ и $\theta(B)$ — полиномы от B степеней p и q . Для того чтобы процесс был стационарным, нужно, чтобы корни $\phi(B) = 0$ лежали вне единичного круга. Естественный путь получения нестационарного процесса заключается в ослаблении этого ограничения.

Для лучшего понимания возможностей этого подхода рассмотрим модель авторегрессии первого порядка

$$(1 - \phi B) \tilde{z}_t = a_t, \quad (4.1.2)$$

стационарную при $|\phi| < 1$. Изучим поведение этого процесса при значениях $\phi = 2$, лежащем вне диапазона стационарности.

В табл. 4.1 показано множество случайных величин нормального процесса с единичной дисперсией и соответствующие

значения ряда \tilde{z}_t , генерируемого моделью $\tilde{z}_t = 2\tilde{z}_{t-1} + a_t$ при $\tilde{z}_0 = 0,7$.

Таблица 4.1

Первые 11 значений нестационарного процесса авторегрессии
первого порядка

t	0	1	2	3	4	5	6	7	8	9	10
a_t		0,1	-1,1	0,2	-2,0	-0,2	-0,8	0,8	0,1	0,1	-0,9
\tilde{z}_t	0,7	1,5	1,9	4,0	6,0	11,8	22,8	46,4	92,9	185,9	370,9

Ряд представлен графиком на рис. 4.2. Видно, что после короткого вступительного периода ряд «скрывается» и, по существу, следует экспоненциальному закону; причем генерируемые значения a практически не играют никакой роли в дальнейшем ходе графика. Поведение рядов, генерируемых процессами высшего порядка, у которых нарушено условие стационарности, сходно. Далее, это поведение остается существенно тем же вне зависимости от того, включены ли в модель члены, описывающие процесс скользящего среднего.

4.1.2. Общая модель для нестационарного процесса, проявляющего однородность

Модель авторегрессии — проинтегрированного скользящего среднего. Хотя нестационарные модели рассмотренного выше типа могут представлять определенный интерес для описания взрывного или эволюционного поведения (такого, как размножение бактерий), ситуации, которыми мы занимаемся в этой книге, существенно иные. Мы видели, что процесс АРСС стационарен, если корни уравнения $\phi(B) = 0$ лежат *вне* единичного круга, и нестационарен, если корни лежат *внутри* единичного круга. Единственный нерассмотренный возможный случай — когда корни уравнения $\phi(B) = 0$ лежат *на* единичной окружности. Оказывается, что соответствующие модели очень важны, так как позволяют описывать однородные нестационарные временные ряды. В частности, несезонные ряды могут нередко хорошо описываться моделями, у которых один или несколько корней равны *единице*; эти модели и рассматриваются в настоящей главе *).

*) В гл. 9 мы рассмотрим модели, способные описывать сезонные колебания периода s , у которых характеристическое уравнение имеет корни на единичной окружности, равные корню s -й степени из единицы,

Рассмотрим модель

$$\varphi(B)\tilde{z}_t = \theta(B)a_t, \quad (4.1.3)$$

где $\varphi(B)$ — нестационарный оператор авторегрессии, такой, что d корней уравнения $\varphi(B) = 0$ равны единице, а остальные лежат вне единичного круга. Тогда модель (4.1.3) можно представить в виде

$$\varphi(B)\tilde{z}_t = \phi(B)(1-B)^d \tilde{z}_t = \theta(B)\dot{a}_t, \quad (4.1.4)$$

где $\phi(B)$ — стационарный оператор авторегрессии. Так как $\nabla^d \tilde{z}_t = \nabla^d z_t$ для $d \geq 1$, мы можем представить модель в виде

$$\phi(B)\nabla^d z_t = \theta(B)a_t. \quad (4.1.5)$$

Эквивалентное определение процесса можно дать двумя уравнениями

$$\phi(B)w_t = \theta(B)a_t \quad (4.1.6)$$

и

$$w_t = \nabla^d z_t. \quad (4.1.7)$$

Таким образом, мы видим, что модель соответствует предположению, что d -я разность ряда может быть представлена стационарным обратимым процессом АРСС. Другой способ трактовки этого процесса *) при $d \geq 1$ получим, обратив (4.1.7):

$$z_t = S^d w_t, \quad (4.1.8)$$

где S — бесконечный оператор суммирования, определенный как

$$Sx_t = \sum_{h=-\infty}^t x_h = (1 + B + B^2 + \dots)x_t = (1 - B)^{-1}x_t = \nabla^{-1}x_t.$$

Таким образом,

$$S = (1 - B)^{-1} = \nabla^{-1}.$$

Оператор $S^2 x_t$ определен аналогично:

$$S^2 x_t = Sx_t + Sx_{t-1} + Sx_{t-2} + \dots = \sum_{i=-\infty}^t \sum_{h=-\infty}^i x_h.$$

Далее

$$S^3 x_t = \sum_{j=-\infty}^t \sum_{i=-\infty}^j \sum_{h=-\infty}^i x_h \text{ и т. д.}$$

Уравнение (4.1.8) указывает, что процесс (4.1.5) можно получить суммированием (или интегрированием) процесса (4.1.6) d раз. Поэтому процесс (4.1.5) мы будем называть *процессом авторегрессии — проинтегрированного скользящего среднего*

*) См. предисловие.

(АРПСС): Модели АРПСС для нестационарных временных рядов, рассмотренные также Ягломом [38], имеют фундаментальное значение в проблемах прогнозирования и управления [14—20]. Для дальнейшей дискуссии о нестационарных процессах мы отсылаем к Заде и Рагаззини [39] и Калману [40, 41]. Еще ранее Тинтнером [91] была описана процедура анализа рядов, использующая разностный подход, получивший название «метода случайных разностей». Однако мотивы, методы и цели этой процедуры были весьма отличны от рассматриваемых здесь.

Как указывалось в гл. 1, модель (4.1.5) эквивалентна описанию процесса z_t как выхода линейного фильтра (если $d \neq 0$, это *неустойчивый* линейный фильтр), на входе которого белый шум a_t . Иначе мы можем рассматривать его как средство для *преобразования сильно зависимых и, возможно, нестационарных членов z_t процесса в последовательность некоррелированных случайных переменных a_t* , т. е. для преобразования процесса в белый шум.

Если в (4.1.5) оператор авторегрессии имеет порядок p , взята d -я разность и оператор скользящего среднего $\theta(B)$ имеет порядок q , мы говорим, что имеем модель АРПСС порядка (p, d, q) или просто процесс АРПСС (p, d, q) .

Две интерпретации модели АРПСС. Покажем теперь, что (4.1.5) — интуитивно оправданная модель для тех типов встречающихся в эксперименте временных рядов, которые мы хотим изучать. Прежде всего отметим основную особенность процесса авторегрессии первого порядка (4.1.2) для $|\phi| < 1$ и для $|\phi| > 1$. Она заключается в том, что локальное поведение ряда, генерируемого моделью, существенно зависит от *уровня \bar{z}_t* . Напротив, локальное поведение рядов такого типа, как показанный на рис. 4.1, кажется не зависящим от уровня.

Если нам нужны модели, у которых поведение процесса не зависит от его уровня, следует выбрать оператор авторегрессии $\varphi(B)$ таким, что

$$\varphi(B)(\bar{z}_t + c) = \varphi(B)\bar{z}_t,$$

где c — любая константа. Отсюда $\varphi(B)$ должно иметь вид

$$\varphi(B) = \phi_1(B)(1 - B) = \phi_1(B)\nabla.$$

Следовательно, класс процессов, имеющих желаемое свойство, должен иметь вид

$$\phi_1(B)\omega_t = \theta(B)a_t,$$

где $\omega_t = \nabla \bar{z}_t = \nabla z_t$. Далее, нельзя допустить, чтобы *разность ω_t* неконтролируемо росла. Это означает, что либо $\phi_1(B)$ должен быть стационарным оператором авторегрессии, либо $\phi_1(B) = \phi_2(B)(1 - B)$, где $\phi_2(B)$ — стационарный оператор

авторегрессии. В последнем случае то же рассуждение можно применить ко второй разности и т. д.

В результате мы приходим к заключению, что для представления нестационарных, но однородных временных рядов оператор в левой части (4.1.3) должен иметь вид $\phi(B)\nabla^d$, где $\phi(B)$ — стационарный оператор авторегрессии. Таким образом, мы приходим опять к модели (4.1.5). Смысл этой модели можно пояснить и с несколько другой точки зрения. Рассмотрим случай, когда в (4.1.4) $d = 0$, так что $\phi(B)\tilde{z}_t = \theta(B)a_t$. Требование,

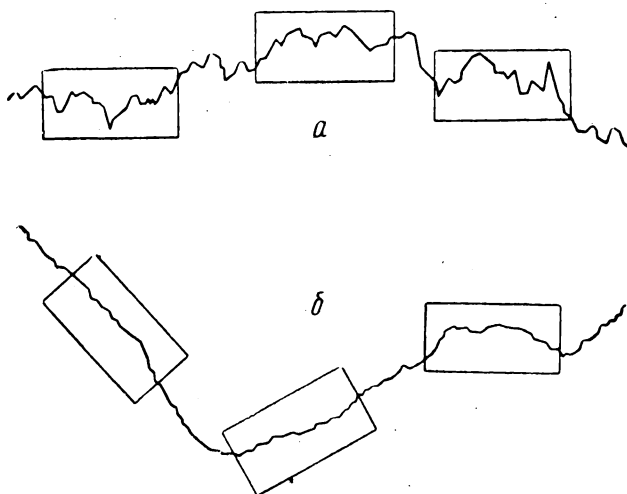


Рис. 4.3. Два типа однородного нестационарного поведения.

a — ряд с нестационарностью уровня, представимый моделью $\phi(B)\nabla z_t = \theta(B)a_t$; b — ряд с нестационарностью уровня и наклона, представимый моделью $\phi(B)\nabla^2 z_t = \theta(B)a_t$.

чтобы нули функции $\phi(B)$ лежали вне единичного круга, приводит к тому, что не только процесс \tilde{z}_t , но и процессы ∇z_t , $\nabla^2 z_t$, $\nabla^3 z_t$ и т. д. стационарны и имеют нулевые средние значения.

На рис. 4.3, a показан один тип нестационарных рядов, которые бы мы хотели описать моделью. Этот ряд однороден, если не принимать во внимание уровень; иными словами, отдельные участки ряда, за исключением смещения по вертикали, выглядят одинаковыми. Такое поведение можно описать, выдвинув требование, чтобы каждая разность была стационарна с нулевым средним значением, и допустив «свободные» изменения уровня *). Этого можно добиться, используя модель

$$\phi(B)\nabla z_t = \theta(B)a_t.$$

*) Приведенные рассуждения неприменимы на тех участках временного ряда на рис. 4.3, a , где уровень резко изменяется. — Прим. ред.

На рис. 4.3, б показан второй тип нестационарности, часто встречающийся на практике. Этот ряд не имеет ни фиксированного уровня, ни фиксированного наклона, но его поведение однородно, если мы допускаем различие в этих характеристиках. Такое поведение можно описать моделью

$$\phi(B) \nabla^2 z_t = \theta(B) a_t,$$

которая обеспечивает стационарность и нулевое среднее значение для всех разностей, кроме первой и второй, но допускает «свободные» изменения уровня и наклона *).

4.1.3. Общий вид процесса авторегрессии — проинтегрированного скользящего среднего

По причинам, которые будут рассмотрены ниже, иногда целесообразно рассматривать слегка обобщенную форму модели АРПСС (4.1.5), полученную добавлением постоянного члена θ_0 . Итак, довольно общий вид модели, которой мы будем пользоваться для описания временных рядов, это процесс авторегрессии — проинтегрированного скользящего среднего

$$\phi(B) z_t = \phi(B) \nabla^d z_t = \theta_0 + \theta(B) a_t, \quad (4.1.9)$$

где

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p,$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q.$$

В дальнейшем изложении

1) $\phi(B)$ будет называться *оператором авторегрессии*; предполагается, что этот оператор стационарен, т. е. корни $\phi(B) = 0$ лежат вне единичного круга.

2) $\phi(B) = \nabla^d \phi(B)$ будет называться *обобщенным оператором авторегрессии*; это нестационарный оператор, у которого d корней уравнения $\phi(B) = 0$ равны единице.

3) $\theta(B)$ будет называться *оператором скользящего среднего*; предполагается, что он обратим, т. е. что корни $\theta(B) = 0$ лежат вне единичного круга.

Когда $d = 0$, модель (4.1.9) описывает стационарный процесс. Требования стационарности и обратимости накладываются независимо, и в общем случае операторы $\phi(B)$ и $\theta(B)$ имеют разные порядки. Примеры областей стационарности для простых случаев $p = 1, 2$ и идентичных областей обратимости для $q = 1, 2$ были приведены в гл. 3.

Стохастические и детерминированные тренды. Мы видели в разд. 4.1.2, что когда постоянный член θ_0 опущен, модель

*) Приведенные рассуждения неприменимы на тех участках временного ряда на рис. 4.3, б, где наклон резко изменяется. — *Прим. ред.*

(4.1.9) может описывать ряды со *стохастическими* трендами, например случайным уровнем или наклоном ряда. В общем же случае мы можем включить в модель *детерминированную* функцию времени $f(t)$. В частности, автоматический учет детерминированного полиномиального тренда степени d обеспечивается условием, что θ_0 не равно 0. Например, когда $d = 1$, мы можем использовать модель с $\theta_0 \neq 0$ для оценок возможного детерминированного линейного тренда в присутствии нестационарного шума. Поскольку условие $\theta_0 \neq 0$ эквивалентно условию, что

$$E[w_t] = E[\nabla^d z_t] = \mu_w = \theta_0 / (1 - \phi_1 - \phi_2 - \dots - \phi_p)$$

не равно нулю, другой способ описания этой более общей модели (4.1.9) осуществляется в виде стационарного обратимого процесса АРСС, в котором $\tilde{w}_t = w_t - \mu_w$, т. е.

$$\phi(B) \tilde{w}_t = \theta(B) a_t. \quad (4.1.10)$$

В тех случаях, когда нет физических причин для существования детерминированной компоненты, среднее значение w может предполагаться нулевым, если только это не противоречит данным. Ниже, в случаях, когда $d > 0$, мы часто будем полагать, что $\mu_w = 0$, или, эквивалентно, $\theta_0 = 0$, если только сами данные или смысл задачи не будут указывать на то, что необходимо учитывать ненулевое среднее или в более общей форме детерминированную компоненту известного вида.

Некоторые важные специальные случаи модели АРСС. В гл. 3 мы познакомились с некоторыми важными специальными случаями модели (4.1.9), соответствующими стационарной ситуации $d = 0$. Следующие модели являются частными случаями нестационарной модели ($d \geq 1$), которая, по-видимому, часто встречается на практике:

1) Процесс (0, 1, 1)

$$\Delta z_t = a_t - \theta_1 a_{t-1} = (1 - \theta_1 B) a_t,$$

для которого $p = 0$, $d = 1$, $q = 1$, $\phi(B) = 1$, $\theta(B) = 1 - \theta_1 B$.

2) Процесс (0, 2, 2)

$$\nabla^2 z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} = (1 - \theta_1 B - \theta_2 B^2) a_t,$$

для которого $p = 0$, $d = 2$, $q = 2$, $\phi(B) = 1$, $\theta(B) = 1 - \theta_1 B - \theta_2 B^2$.

3) Процесс (1, 1, 1)

$$\nabla z_t - \phi_1 \nabla z_{t-1} = a_t - \theta_1 a_{t-1}$$

или

$$(1 - \phi_1 B) \nabla z_t = (1 - \theta_1 B) a_t,$$

для которого $p = 1$, $d = 1$, $q = 1$, $\phi(B) = 1 - \phi_1 B$, $\theta(B) = 1 - \theta_1 B$.

При описании несезонных временных рядов (сезонные модели рассмотрены в гл. 9) мы редко встречаемся с ситуацией, при которой p , d или q должны быть больше 2. Часто оказывается, что этим параметрам достаточно придать значения 0 или 1. Например, мы покажем, что ряды А, В, С, D, показанные на рис. 4.1, могут быть хорошо описаны *) простыми моделями, приведенными в табл. 4.2.

Таблица 4.2

Сводка простых нестационарных моделей, подогнанных к временным рядам на рис. 4.1

Ряд	Модель	Порядок модели
A	$\nabla z_t = (1-0,7B) a_t$	(0, 1, 1)
B	$\nabla z_t = (1+0,1B) a_t$	(0, 1, 1)
C	$(1-0,8B) \nabla z_t = a_t$	(1, 1, 0)
D	$\nabla z_t = (1-0,1B) a_t$	(0, 1, 1)

Нелинейное преобразование процесса z_t . Область полезных приложений модели (4.1.9) может быть значительно расширена, если можно заменить z_t в (4.1.9) на $z_t^{(\lambda)}$, где $z_t^{(\lambda)}$ — некоторое нелинейное преобразование z_t , включающее один или несколько параметров преобразования λ . Вид преобразования, которое следует применить, часто подсказывается ситуацией или может быть оценен из данных. Например, если мы интересуемся сбытом недавно введенного товара, мы можем выяснить, что объем сбыта быстро увеличивался и что нестационарную устойчивость проявляли *относительные*, а не абсолютные флуктуации. В таком случае, очевидно, целесообразно анализировать логарифм объема сбыта. Когда преобразование должно оцениваться по данным, один из способов такой оценки — использование подхода, описанного Боксом и Коксом [42].

4.2. ТРИ ФОРМЫ ПРЕДСТАВЛЕНИЯ МОДЕЛИ АВТОРЕГРЕССИИ — ПРОИНТЕГРИРОВАННОГО СКОЛЬЗЯЩЕГО СРЕДНЕГО

Рассмотрим теперь три «явных» представления общей модели (4.1.9). Каждое из них имеет свои достоинства.

*) Как будет подробнее рассмотрено дальше, в спорных случаях нестационарные модели часто удобнее использовать, чем стационарные. В частности, ни для одной из подгоняемых к этим рядам моделей не предполагалось, что z_t имеет фиксированное среднее. Однако в гл. 7 мы покажем, что в некоторых случаях стационарные модели дают несколько лучшее согласие.

Итак, текущее значение z_t процесса можно выразить

а) через предыдущие значения z и текущее и предшествующие значения a ; это достигается прямым использованием *разностного уравнения*;

б) только через *текущий и предшествующие импульсы* a_{t-j} ;

в) через взвешенную сумму *предшествующих значений* z_{t-j} процесса и текущий импульс a_t .

В этой главе мы в основном интересуемся *нестационарными* моделями, у которых $\nabla^d z_t$ — стационарный процесс и d больше нуля. Для таких моделей можно без потери общности опустить μ или, что то же самое, заменить \tilde{z}_t на z_t . Результаты этой главы будут, однако, приложимы и к стационарным моделям, у которых $d = 0$ при условии, что z_t интерпретируется как *отклонение* от среднего μ .

4.2.1. Представление модели в виде разностного уравнения

Прямое использование разностного уравнения позволяет нам выразить текущее значение z_t процесса через предшествующие значения z и текущее или предшествующие значения a . Так, если

$$\phi(B) = \phi(B)(1-B)^d = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_{p+d} B^{p+d},$$

общая модель (4.1.9) при $\theta_0 = 0$ может быть записана в виде

$$z_t = \phi_1 z_{t-1} + \dots + \phi_{p+d} z_{t-p-d} - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} + a_t. \quad (4.2.1)$$

В качестве примера рассмотрим процесс, описываемый моделью порядка (1,1,1):

$$(1 - \phi B)(1 - B)z_t = (1 - \theta B)a_t$$

(индексы 1 в ϕ_1 и θ_1 опущены для краткости). Тогда этот процесс можно записать в виде

$$[1 - (1 + \phi)B + \phi B^2]z_t = (1 - \theta B)a_t,$$

т. е.

$$z_t = (1 + \phi)z_{t-1} - \phi z_{t-2} + a_t - \theta a_{t-1}. \quad (4.2.2)$$

Для многих целей, и в частности для вычисления прогнозов в гл. 5, разностное уравнение (4.2.1) является наиболее удобным описанием модели.

4.2.2. Представление модели, использующее случайные импульсы

Представление модели через текущий и предшествующие импульсы. Мы видели в разд. 3.1.1, что линейная модель может быть представлена как выход линейного фильтра

$$z_t = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots = a_t + \sum_{j=1}^{\infty} \psi_j a_{t-j} = \psi(B)a_t, \quad (4.2.3)$$

входом которого является белый шум, или последовательность некоррелированных импульсов a_t . Иногда полезно представлять модель АРПСС в форме (4.2.3); в частности, веса ψ потребуются в гл. 5 для вычисления дисперсии прогнозов. Покажем теперь, что веса ψ процесса АРПСС можно получить прямо из представления модели в виде разностного уравнения.

Общее выражение для весов ψ . Если применить к обеим сторонам (4.2.3) обобщенный оператор авторегрессии $\phi(B)$, получим

$$\phi(B) z_t = \phi(B) \psi(B) a_t.$$

Однако, так как

$$\phi(B) z_t = \theta(B) a_t,$$

отсюда следует, что

$$\phi(B) \psi(B) = \theta(B). \quad (4.2.4)$$

Следовательно, веса ψ можно получить, приравнявая коэффициенты при одинаковых степенях B в уравнении

$$(1 - \phi_1 B - \dots - \phi_{p+d} B^{p+d})(1 + \psi_1 B + \psi_2 B^2 + \dots) = \\ = (1 - \theta_1 B - \dots - \theta_q B^q). \quad (4.2.5)$$

Заметим, что для j , больших, чем $p + d - 1$ и q , т. е. для таких, что

$$j > p + d - 1, \quad \text{если } p + d - 1 \geq q, \\ j > q, \quad \text{если } p + d - 1 < q,$$

веса ψ удовлетворяют разностному уравнению, определенному обобщенным оператором авторегрессии, т. е.

$$\phi(B) \psi_j = \phi(B) (1 - B)^d \psi_j = 0, \quad (4.2.6)$$

где B действует на индекс j . Тогда для достаточно больших j веса ψ_j представлены совокупностью полиномов, затухающих экспонент и затухающих синусоид от аргумента j .

Пример. Чтобы показать, как пользоваться (4.2.5), рассмотрим процесс (4.2.2) порядка $(1, 1, 1)$, для которого

$$\phi(B) = (1 - \phi B)(1 - B) = 1 - (1 + \phi)B + \phi B^2$$

и

$$\theta(B) = 1 - \theta B.$$

Подстановка в (4.2.5) приводит к

$$[1 - (1 + \phi)B + \phi B^2](1 + \psi_1 B + \psi_2 B^2 + \dots) = 1 - \theta B,$$

и отсюда

$$\begin{aligned}\psi_0 &= 1, \\ \psi_1 &= A_0 + A_1\phi, \\ \psi_2 &= A_0 + A_1\phi^2, \\ &\vdots \\ \psi_j &= A_0 + A_1\phi^j,\end{aligned}\tag{4.2.7}$$

где

$$A_0 = \frac{1-\theta}{1-\phi}, \quad A_1 = \frac{\theta-\phi}{1-\phi}$$

и $\psi_0 = A_0 + A_1 = 1$. Отсюда мы можем представить модель (4.2.2) эквивалентным выражением

$$z_t = \sum_{j=0}^{\infty} (A_0 + A_1\phi^j) a_{t-j}.\tag{4.2.8}$$

Так как $|\phi| < 1$, веса ψ_j при больших j стремятся к A_0 , так что импульсы a_{t-j} , которые внесли свой вклад в отдаленном прошлом, получают постоянный вес A_0 .

Усеченная форма представления модели через случайные импульсы. В некоторых задачах удобнее использовать выражение для модели, слегка отличающееся от (4.2.3). Пусть мы хотим выразить текущее значение процесса z_t через $t-k$ импульсов $a_t, a_{t-1}, \dots, a_{k+1}$, вошедших в систему после некоторого момента времени $k < t$. Этот момент времени k может, например, быть моментом первого наблюдения процесса.

Общая модель

$$\varphi(B) z_t = \theta(B) a_t\tag{4.2.9}$$

— это разностное уравнение с решением

$$z_t = C_k(t-k) + I_k(t-k).\tag{4.2.10}$$

Краткое рассмотрение линейных разностных уравнений проведено в приложении П4.1. Напомним читателю, что решение таких уравнений весьма сходно с решением линейных дифференциальных уравнений. *Функция $C_k(t-k)$* — это общее решение однородного разностного уравнения

$$\varphi(B) C_k(t-k) = 0.\tag{4.2.11}$$

В общем случае это решение состоит из *линейной* комбинации некоторых функций времени: степеней t^j , действительных экспонент G^t и комплексных экспонент $d^t \sin(2\pi f_0 t + F)$, где константы G, f_0, F — функции параметров модели (ϕ, θ) . *Частное решение $I_k(t-k)$* — это *любая* функция, удовлетворяющая

уравнению

$$\varphi(B)I_k(t-k) = \theta(B)a_t. \quad (4.2.12)$$

Не следует забывать, что в этом выражении B действует на t , а не на k . В приложении П4.1 показано, что этому уравнению при $t-k > q$ удовлетворяет функция

$$I_k(s-k) = \begin{cases} 0, & s \leq k, \\ \sum_{j=k+1}^s \psi_{s-j} a_j = a_s + \psi_1 a_{s-1} + \dots \\ \dots + \psi_{s-k-1} a_{k+1}, & s > k. \end{cases} \quad (4.2.13)$$

Рассмотрим в качестве иллюстрации рис. 4.4. Из вышесказанного вытекает, что любое наблюдение z_t можно связать с любым предшествующим временем k и представить в виде суммы

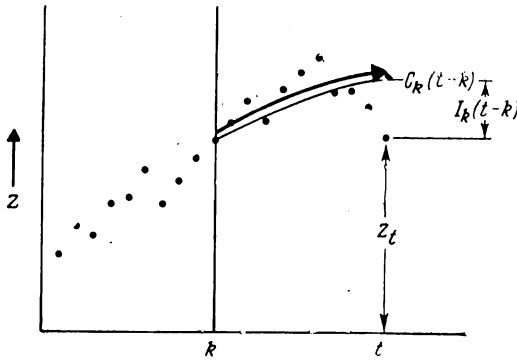


Рис. 4.4. Роль функции $C_k(t-k)$ и частного решения $I_k(t-k)$ в описании поведения временного ряда.

двух слагаемых. Первое слагаемое $C_k(t-k)$ — это компонента z_t , уже *определенная в момент времени k* . По нему можно оценить, как сказываются на значении ряда в момент t наблюдения, предшествующие моменту $k+1$. Иными словами, оно указывает, что было бы с процессом, если в момент времени k источник импульсов a_t был бы «выключен». Второе слагаемое $I_k(t-k)$ — это добавочная компонента, *не предсказуемая в момент времени k* , которая представляет результирующий эффект импульсов, образовавшихся после времени k .

Пример. Для иллюстрации рассмотрим опять пример

$$(1 - \phi B)(1 - B)z_t = (1 - \theta B)a_t.$$

Общее решение $C_k(t-k)$ соответствующего однородного разностного уравнения

$$(1 - \phi B)(1 - B)C_k(t-k) = 0$$

имеет вид

$$C_k(t-k) = b_0^{(k)} + b_1^{(k)} \phi^{t-k},$$

где $b_0^{(k)}$, $b_1^{(k)}$ — коэффициенты, зависящие от прошлого процесса и, что существенно, *изменяющиеся с номером k* .

Частное решение (4.2.13) при помощи формул (4.2.7) для весов ψ представимо в виде

$$I_k(t-k) = \sum_{j=k+1}^t (A_0 + A_1 \phi^{t-j}) a_j.$$

Таким образом, модель (4.2.8) можно записать в эквивалентном виде:

$$z_t = b_0^{(k)} + b_1^{(k)} \phi^{t-k} + \sum_{j=k+1}^t (A_0 + A_1 \phi^{t-j}) a_j. \quad (4.2.14)$$

Связь между усеченной и полной формой представления модели через случайные импульсы. Вернемся теперь к общему случаю. Мы всегда можем предполагать, что начало процесса находится в бесконечно далеком прошлом, так что функцию $C_{-\infty}(t)$ без потери общности можно считать равной нулю.

В этом случае

$$z_t = \sum_{j=-\infty}^t \psi_{t-j} a_j = \psi(B) a_t = I_{-\infty}(t) \quad (4.2.15)$$

— это полное выражение (4.2.3), описывающее модель.

Функция $C_k(t-k)$ также может быть выражена через веса ψ , так как, вычитая (4.2.10) из (4.2.15), мы получим для $t-k > q$

$$C_k(t-k) = I_{-\infty}(t) - I_k(t-k) = \sum_{j=-\infty}^k \psi_{t-j} a_j. \quad (4.2.16)$$

Можно показать, что функция $C_k(t-k)$ в форме (4.2.16) удовлетворяет (4.2.11), так как

$$\phi(B) [I_{-\infty}(t) - I_k(t-k)] = [\theta(B) - \theta(B)] a_t = 0.$$

Итак, для общей модели (4.2.9):

1) Значение процесса z_t может быть выражено как бесконечная взвешенная сумма текущего и предшествующих импульсов a_t :

$$z_t = \sum_{j=-\infty}^t \psi_{t-j} a_j = \sum_{j=0}^{\infty} \psi_j a_{t-j} = \psi(B) a_t.$$

2) Для $t-k > q$ значение z_t можно представить как взвешенную конечную сумму $t-k$ текущего и предшествующих импульсов, возникающих после момента времени k плюс функция $C_k(t-k)$ — общее решение однородного разностного уравне-

ния (4.2.1). Первое слагаемое состоит из первых $t-k$ членов бесконечной суммы $\sum_{j=0}^{\infty} \psi_j a_{t-j}$, так что

$$z_t = C_k(t-k) + \sum_{j=k+1}^t \psi_{t-j} a_j. \quad (4.2.17)$$

Наконец, функция $C_k(t-k)$ может быть принята равной усеченной бесконечной сумме, так что

$$C_k(t-k) = \sum_{j=-\infty}^k \psi_{t-j} a_j. \quad (4.2.18)$$

В качестве примера рассмотрим еще раз модель

$$(1 - \phi B)(1 - B)z_t = (1 - \theta B)a_t.$$

Мы можем выразить z_t либо как бесконечную взвешенную сумму a_j

$$z_t = \sum_{j=-\infty}^t (A_0 + A_1 \phi^{t-j}) a_j,$$

либо при помощи конечной взвешенной суммы

$$z_t = C_k(t-k) + \sum_{j=k+1}^t (A_0 + A_1 \phi^{t-j}) a_j.$$

Далее, функция $C_k(t-k)$ — это усеченная сумма

$$C_k(t-k) = \sum_{j=-\infty}^k (A_0 + A_1 \phi^{t-j}) a_j,$$

которую можно представить в виде

$$C_k(t-k) = b_0^{(k)} + b_1^{(k)} \phi^{t-k},$$

где

$$b_0^{(k)} = A_0 \sum_{j=-\infty}^k a_j = \frac{1-\theta}{1-\phi} \sum_{j=-\infty}^k a_j,$$

$$b_1^{(k)} = A_1 \sum_{j=-\infty}^k \phi^{k-j} a_j = \frac{\theta-\phi}{1-\phi} \sum_{j=-\infty}^k \phi^{k-j} a_j.$$

Общее решение однородного разностного уравнения как условное математическое ожидание. Одним из следствий соотношения (4.2.18) является (при $m > 0$)

$$\begin{aligned} C_k(t-k) &= C_{k-m}(t-k+m) + \psi_{t-k} a_k + \\ &+ \psi_{t-k+1} a_{k-1} + \dots + \psi_{t-k+m-1} a_{k-m+1}, \end{aligned} \quad (4.2.19)$$

что показывает, как изменяется функция $C_k(t - k)$ при изменении начала отсчета k . Обозначим теперь *условное математическое ожидание* z_t в момент времени k через $E_k[z_t]$.

Это есть математическое ожидание при условии, что известно все прошлое процесса вплоть до момента k , а то, что произойдет после момента k , неизвестно. При отыскании $E_k[z_t]$ следует учесть, что

$$E_k[a_j] = \begin{cases} 0, & j > k, \\ a_j, & j \leq k. \end{cases}$$

Это означает, что в момент времени k ожидаемые значения a в будущем равны нулю, и математические ожидания тех a , которые уже реализовались, равны этим реализовавшимся значениям.

Беря условные математические ожидания в момент k от обеих частей (4.2.17), получаем $E_k[z_t] = C_k(t - k)$. Отсюда для $(t - k) > q$ функция $C_k(t - k)$ выражает ожидаемое в момент времени k значение z_t процесса в будущем, основанное на знании прошлого. Частное решение $I_k(t - k)$ показывает, как это математическое ожидание изменится *последующими* событиями, представленными последовательностью импульсов $a_{k+1}, a_{k+2}, \dots, a_t$. В задаче прогнозирования, которую мы рассмотрим в гл. 5, окажется, что функция $C_k(t - k)$ — это прогноз z_t в момент времени k с минимальной среднеквадратичной ошибкой.

Уравнение (4.2.19) можно использовать для «подправления» этого прогноза.

4.2.3. Обращенное представление модели

Модель, выражающая z_t через a_t и предшествующие z . В разд. 3.1.1 было показано, что модель

$$z_t = \psi(B) a_t$$

можно записать также в обращенной форме:

$$\psi^{-1}(B) z_t = a_t,$$

или

$$\pi(B) z_t = \left(1 - \sum_{j=1}^{\infty} \pi_j B^j\right) z_t = a_t. \quad (4.2.20)$$

Здесь z_t равно бесконечной взвешенной сумме предыдущих значений z плюс случайный импульс, т. е.

$$z_t = \pi_1 z_{t-1} + \pi_2 z_{t-2} + \dots + a_t.$$

Согласно условию обратимости, веса π в (4.2.20) должны образовывать сходящийся ряд, т. е. $\pi(B)$ должно сходиться на единичной окружности или внутри нее.

Общее выражение для весов π . Чтобы вывести формулу для весов π общей модели АРСС, мы подставим (4.2.20) в уравнение

$$\varphi(B) z_t = \theta(B) a_t$$

и получим

$$\varphi(B) z_t = \theta(B) \pi(B) z_t.$$

Отсюда веса π могут быть явно выражены приравниванием коэффициентов при B в левой и правой частях уравнения

$$\varphi(B) = \theta(B) \pi(B), \quad (4.2.21)$$

или, в развернутом виде,

$$(1 - \varphi_1 B - \dots - \varphi_{p+d} B^{p+d}) = (1 - \theta_1 B - \dots - \theta_q B^q) \times \\ \times (1 - \pi_1 B - \pi_2 B^2 - \dots). \quad (4.2.22)$$

Нужно отметить, что для j , больших чем $p + d$ и q , т. е. для таких, что

$$j > p + d, \quad \text{если } p + d \geq q,$$

$$j > q, \quad \text{если } p + d < q,$$

веса π удовлетворяют разностному уравнению, определяемому оператором скользящего среднего

$$\theta(B) \pi_j = 0,$$

где B действует теперь на j . Отсюда для достаточно больших j веса π будут вести себя сходно с автокорреляционной функцией (3.2.5) процесса авторегрессии, т. е. состоять из суммы затухающих экспонент и затухающих синусоид.

Другой интересный факт заключается в том, что при $d \geq 1$ веса π в (4.2.20) в сумме равны 1. Это можно проверить, подставив в (4.2.21) значение $B=1$. Поскольку $\varphi(B) = \phi(B) (1-B)^d$ равно нулю при $B=1$, то $\theta(1) \neq 0$, так как корни уравнения $\theta(B)=0$ лежат вне единичного круга. Тогда из (4.2.21) следует, что $\pi(1) = 0$ и

$$\sum_{j=1}^{\infty} \pi_j = 1. \quad (4.2.23)$$

Поэтому при $d \geq 1$ процесс можно представить в форме

$$z_t = \bar{z}_{t-1}(\pi) + a_t, \quad (4.2.24)$$

где

$$\bar{z}_{t-1}(\pi) = \sum_{j=1}^{\infty} \pi_j z_{t-j}$$

— взвешенное среднее предыдущих значений процесса.

Пример. Рассмотрим опять в качестве примера процесс АРПСС(1, 1, 1)

$$(1 - \phi B)(1 - B)z_t = (1 - \theta B)a_t.$$

Тогда, пользуясь (4.2.21), получим

$$\pi(B) = \varphi(B)\theta^{-1}(B) = \{1 - (1 + \phi)B + \phi B^2\}(1 + \theta B + \theta^2 B^2 + \dots),$$

так что

$$\pi_1 = \phi + (1 - \theta), \quad \pi_2 = (\theta - \phi)(1 - \theta),$$

$$\pi_j = (\theta - \phi)(1 - \theta)\theta^{j-2}, \quad j \geq 3.$$

Первые семь весов π , соответствующие $\phi = -0,3$ и $\theta = 0,5$, приведены в табл. 4.3. Итак, t -е значение процесса может быть

Таблица 4.3

Первые семь весов процесса АРПСС (1, 1, 1) при
 $\phi = -0,3, \theta = 0,5$

j	1	2	3	4	5	6	7
π_j	0,2	0,4	0,2	0,1	0,05	0,025	0,0125

найдено как взвешенное среднее предшествующих значений плюс дополнительный импульс по формуле

$$z_t = (0,2z_{t-1} + 0,4z_{t-2} + 0,2z_{t-3} + 0,1z_{t-4} + \dots) + a_t.$$

Отметим, в частности, что веса π уменьшаются для удаляющихся значений z_{t-j} . Это свойство является следствием требования обратимости ряда, которое в данном случае равносильно требованию $-1 < \theta < 1$. Кстати, заметим, что для статистических моделей, описывающих практически встречающиеся временные ряды, сходящиеся веса обычно сравнительно быстро убывают до нуля. Поэтому, хотя z_t теоретически зависит от отдаленного прошлого, представление

$$z_t = \sum_{j=1}^{\infty} \pi_j z_{t-j} + a_t$$

обычно показывает, что z_t существенно зависит только от сравнительно недавних предшествующих значений временного ряда z_{t-j} . Это остается справедливым, несмотря на то что для нестационарных моделей с $d > 0$ веса ψ в представлении процесса в виде «взвешенных импульсов»

$$z_t = \sum_j \psi_j a_{t-j}$$

не сходятся. Это связано с тем, что вся информация z_t , поставляемая значениями давно возникших импульсов a_{t-j} , содержится в недавних значениях ряда z_{t-1}, z_{t-2}, \dots . В частности, условное математическое ожидание $E_h[z_t]$, теоретически определяемое через все значения процесса вплоть до момента k , может быть достаточно точно вычислено по нескольким *недавним* значениям временного ряда. Этот факт особенно важен при практическом прогнозировании.

4.3. ПРОЦЕССЫ ПРОИНТЕГРИРОВАННОГО СКОЛЬЗЯЩЕГО СРЕДНЕГО

Нестационарная модель, полезная для описания некоторых, часто встречающихся рядов, — это процесс $(0, 1, 1)$:

$$\nabla z_t = a_t - \theta a_{t-1}.$$

Модель содержит только два параметра θ и σ_a^2 . На рис. 4.5 показаны два временных ряда, генерированных этой моделью по

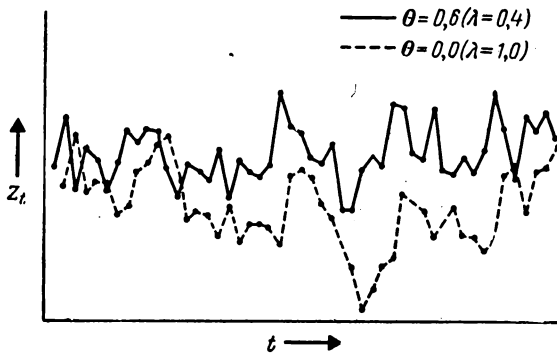


Рис. 4.5. Два временных ряда, генерированных процессом ПСС(0,1,1).

одной и той же последовательности случайных гауссовских величин a_t . Для первого ряда $\theta = 0,6$ и для второго $\theta = 0$. Модели этого типа часто оказываются полезными в задачах складского контроля (inventory control), в описании некоторых возмущений, возникающих в промышленных процессах, и в эконометрике. В гл. 7 будет показано, что этот простой процесс может при правильном подборе параметров хорошо аппроксимировать ряды А, В и D, показанные на рис. 4.1.

Другая весьма полезная модель — это процесс $(0, 2, 2)$

$$\nabla^2 z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2},$$

содержащий три параметра, а именно θ_1 , θ_2 и σ_a^2 . На рис. 4.6 показаны два ряда, генерируемые моделью, использующей одну и ту же последовательность гауссовских случайных чисел. Для первого ряда параметры $(\theta_1, \theta_2) = (0, 0)$, а для второго $(\theta_1, \theta_2) = (1, 5; -0, 8)$. Эти ряды в целом более гладкие, чем генерируемые процессом $(0, 1, 1)$. Модели $(0; 2, 2)$ полезны в описании

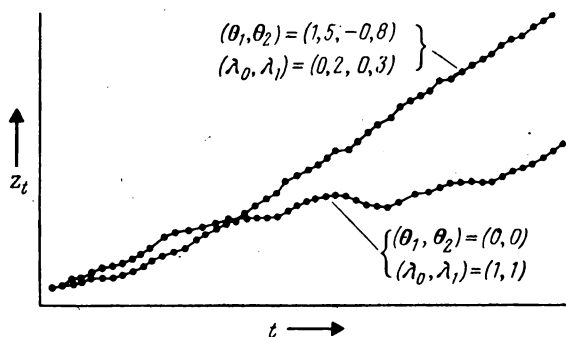


Рис. 4.6. Два временных ряда, генерированных процессом ПСС(0,2,2).

возмущенных систем с большой инерцией (таких, как ряд С). Обе модели $(0, 1, 1)$ и $(0, 2, 2)$ — частные случаи процессов класса

$$\nabla^d z_t = \theta(B) a_t. \quad (4.3.1)$$

Мы называем такие модели, как (4.3.1), процессами *проинтегрированного скользящего среднего* (ПСС) порядка $(0, d, q)$, их свойства будут рассмотрены в следующем разделе.

4.3.1. Процесс проинтегрированного скользящего среднего порядка $(0, 1, 1)$

Представление разностным уравнением. Процесс $(0, 1, 1)$

$$\nabla z_t = (1 - \theta B) a_t, \quad -1 < \theta < 1,$$

оказывается весьма полезным при описании экспериментальных рядов, и мы исследуем теперь его свойства более детально. Модель может быть описана через z и a в виде

$$z_t = z_{t-1} + a_t - \theta a_{t-1}. \quad (4.3.2)$$

Представление через случайные импульсы. Мы можем также выразить z_t только через a_t , суммируя обе стороны (4.3.2). Перед тем как это сделать, воспользуемся выражением для оператора в правой части через ∇ , а не через B . Тогда можно запи-

сать

$$1 - \theta B = (1 - \theta) B + (1 - B) = (1 - \theta) B + \nabla = \lambda B + \nabla,$$

где $\lambda = 1 - \theta$, и интервал обратимости для λ определен неравенствами $0 < \lambda < 2$. Отсюда

$$\nabla z_t = \lambda a_{t-1} + \nabla a_t,$$

и при суммировании получаем

$$z_t = \lambda S a_{t-1} + a_t, \quad (4.3.3)$$

откуда в представлении $z_t = \sum_{j=0}^{\infty} \psi_j a_{t-j}$ веса $\psi_0 = 1$, $\psi_j = \lambda$ для $j \geq 1$.

Если мы опишем модель с помощью импульсов a , появляющихся в системе после момента времени k , то получим, как в (4.2.17),

$$z_t = b_0^{(k)} + \lambda \sum_{j=k+1}^{t-1} a_j + a_t, \quad (4.3.4)$$

где функция $C_k(t-k) = b_0^{(k)}$ (своя константа b_0 для каждого k) — это общее решение однородного разностного уравнения $(1 - B)C_k(t-k) = 0$. Далее, так как

$$S a_{t-1} = \sum_{j=k+1}^{t-1} a_j + S a_k,$$

мы можем записать

$$b_0^{(k)} = \lambda S a_k = \lambda \sum_{j=-\infty}^k a_j.$$

Для этой модели, как мы видим, функция $C_k(t-k)$ — просто константа (т. е. полином от t нулевой степени), показывающая текущий «уровень» процесса и связанная с выбранной точкой отсчета k . Если точка отсчета переносится из $k-1$ в k , то b_0 подправляется согласно формуле

$$b_0^{(k)} = b_0^{(k-1)} + \lambda a_k.$$

Обращенное представление модели. В заключение мы можем рассмотреть модель в виде

$$\pi(B) z_t = a_t$$

или, что эквивалентно, в виде

$$z_t = \sum_{j=1}^{\infty} \pi_j z_{t-j} + a_t = \bar{z}_{t-1}(\pi) + a_t,$$

где $\bar{z}_{t-1}(\pi)$ — взвешенное скользящее среднее предыдущих значений процесса.

Используя (4.2.21), можно определить веса π для ПСС $(0, 1, 1)$ из уравнения

$$(1 - \theta B)\pi(B) = 1 - B,$$

т. е.

$$\pi(B) = \frac{1 - B}{1 - \theta B} = \frac{1 - \theta B - (1 - \theta)B}{1 - \theta B} =$$

$$= 1 - (1 - \theta)(B + \theta B^2 + \theta^2 B^3 + \dots),$$

откуда

$$\pi_j = (1 - \theta)\theta^{j-1} = \lambda(1 - \lambda)^{j-1}, \quad j \geq 1.$$

Таким образом, процесс можно записать в виде

$$z_t = \bar{z}_{t-1}(\lambda) + a_t. \quad (4.3.5)$$

Взвешенное скользящее среднее предыдущих значений процесса

$$\bar{z}_{t-1}(\lambda) = \lambda \sum_{j=1}^{\infty} (1 - \lambda)^{j-1} z_{t-j} \quad (4.3.6)$$

в этом случае — экспоненциально взвешенное скользящее среднее (ЭВСС).

Скользящее среднее (4.3.6) называется экспоненциально (или геометрически) взвешенным потому, что веса

$$\lambda \quad \lambda(1 - \lambda) \quad \lambda(1 - \lambda)^2 \quad \lambda(1 - \lambda)^3 \dots$$

убывают по степенному закону (т. е. как члены геометрической прогрессии). Весовая функция ПСС порядка $(0, 1, 1)$ с $\lambda = 0,4$ ($\theta = 0,6$) показана на рис. 4.7.

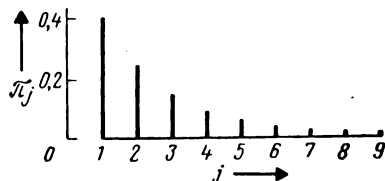


Рис. 4.7. Веса π процесса ПСС $(0,1,1)$ при $\lambda = 1 - \theta = 0,4$.

Хотя условие обратимости выполняется для $0 < \lambda < 2$, на практике наиболее часто сталкиваются со значениями λ в диапазоне между 0 и 1. Отметим, что если $\lambda = 1$, весовая функция содержит только одно ненулевое значение $\pi_1 = 1$. По мере того как λ приближается к нулю, экспоненциальные веса затухают все медленнее и ЭВСС захватывает все большую часть прошлого процесса. Наконец, при $\lambda = 0$ и $\theta = 1$ модель $(1 - B)z_t = (1 - B)a_t$ эквивалентна $z_t = \theta_0 + a_t$, где θ_0 — это среднее всех прошлых значений.

Из сравнения (4.3.3) и (4.3.5) становится очевидным, что

$$\bar{z}_{t-1}(\lambda) = \lambda S a_{t-1}. \quad (4.3.7)$$

Отсюда следует, в частности, что для этого процесса функция $b_0^{(k)} = C_k(t-k)$ в (4.3.4) равна

$$b_0^{(k)} = \bar{z}_k(\lambda),$$

т. е. экспоненциально взвешенному среднему значению процесса до начала отсчета k , так как (4.3.4) можно представить как

$$z_t = \bar{z}_k(\lambda) + \lambda \sum_{j=k+1}^{t-1} a_j + a_t.$$

Мы видим, что функция $C_k(t-k)$ как бы говорит нам, что известно о будущем значении процесса в момент t на базе данных о прошлом, *имеющихся к моменту времени k* . Для ПСС $(0, 1, 1)$ она содержит в себе информацию об «уровне» или положении процесса $b_0^{(k)} = \bar{z}_k(\lambda)$. В момент k наше знание будущего поведения процесса заключается в том, что он будет отклоняться от этого уровня «случайным блужданием», описываемым членом, математическое ожидание которого равно нулю, и поведение его мы не можем предсказать. Как только поступает новое наблюдение, т. е. как только начало отсчета переносится на момент времени $k+1$, уровень должен быть подправлен по формуле $b_0^{(k+1)} = \bar{z}_{k+1}(\lambda)$.

Важные свойства ПСС $(0, 1, 1)$. Так как этот процесс нестационарен, у него нет среднего значения. Однако экспоненциально взвешенное скользящее среднее $\bar{z}_t(\lambda)$ можно рассматривать как измерение положения или «уровня» процесса в момент времени t . Из определения (4.3.6) мы получаем хорошо известную рекуррентную формулу для ЭВСС

$$\bar{z}_t(\lambda) = \lambda z_t + (1 - \lambda) \bar{z}_{t-1}(\lambda). \quad (4.3.8)$$

Из этого выражения следует, что для модели ПСС $(0, 1, 1)$ каждый новый уровень достигается интерполяцией между новым наблюдением и предыдущим уровнем. Если $\lambda=1$, $\bar{z}_t(\lambda)=z_t$, что означает игнорирование всех указаний о положении процесса, полученных из предыдущих наблюдений. С другой стороны, если λ близко к нулю, $\bar{z}_t(\lambda)$ будет сильно зависеть от предыдущего значения $\bar{z}_{t-1}(\lambda)$, входящего с весом $1-\lambda$. Новому наблюдению будет дан малый вес λ .

Рассмотрим теперь два уравнения

$$\begin{aligned} z_t &= \bar{z}_{t-1}(\lambda) + a_t, \\ \bar{z}_t(\lambda) &= \bar{z}_{t-1}(\lambda) + \lambda a_t, \end{aligned} \quad (4.3.9)$$

второе из них получено подстановкой (4.3.5) в (4.3.8); его можно также вывести непосредственно из (4.3.7).

Мат [43] указал на то, что два уравнения (4.3.9) открывают удобный способ трактовки генерирования рассматриваемого процесса. Первое уравнение показывает, как при «уровне» системы $\bar{z}_{t-1}(\lambda)$ в момент t добавляется импульс a_t и образуется значение z_t . Однако второе уравнение показывает, что только λ -я доля импульса фактически «усваивается» уровнем и влияет на будущее; остающаяся доля $\theta = 1 - \lambda$ импульса рассеивается. Далее, после того как в результате «усвоения» импульса a_t установился новый уровень $\bar{z}_t(\lambda)$, в момент $t + 1$ в систему поступает новый импульс a_{t+1} .

Уравнения (4.3.9) с увеличенными на единицу индексами покажут тогда, как этот новый импульс вызывает появление z_{t+1} и как его λ -я доля «усваивается» системой при выработке нового уровня $\bar{z}_{t+1}(\lambda)$ и т. д.

Способ «подправления» текущего значения и уровня демонстрируется также записью уравнений (4.3.9) в виде бесконечных сумм импульсов a , а именно

$$\begin{aligned} z_t &= \lambda S a_{t-1} + a_t, \\ \lambda S a_t &= \lambda S a_{t-1} + \lambda a_t. \end{aligned} \quad (4.3.10)$$

Свойства ПСС $(0, 1, 1)$ с детерминированным смещением нуля

$$\nabla z_t = \theta_0 + (1 - \theta_1 B) a_t$$

обсуждаются в приложении П4.2.

4.3.2. Процесс проинтегрированного скользящего среднего порядка $(0, 2, 2)$

Представление разностным уравнением. Процесс

$$\nabla^2 z_t = (1 - \theta_1 B - \theta_2 B^2) a_t \quad (4.3.11)$$

весьма удобен для описания рядов, имеющих стохастический тренд (например, на рис. 4.6), и мы исследуем теперь его общие свойства внутри диапазона обратимости

$$-1 < \theta_2 < 1, \quad \theta_2 + \theta_1 < 1, \quad \theta_2 - \theta_1 < 1.$$

Как и ранее, можно явно выразить z_t через предшествующие z и a :

$$z_t = 2z_{t-1} - z_{t-2} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}.$$

Представление модели через случайные импульсы. Для того чтобы выразить z_t через a , мы должны представить оператор в правой части в разностном виде:

$$1 - \theta_1 B - \theta_2 B^2 = (\lambda_0 \nabla + \lambda_1) B + \nabla^2.$$

Приравнявая коэффициенты, найдем выражения для θ через λ и наоборот, а именно

$$\begin{aligned}\theta_1 &= 2 - \lambda_0 - \lambda_1, & \lambda_0 &= 1 + \theta_2, \\ \theta_2 &= \lambda_0 - 1, & \lambda_1 &= 1 - \theta_1 - \theta_2.\end{aligned}\quad (4.3.12)$$

Модель (4.3.11) можно тогда записать в виде

$$\nabla^2 z_t = (\lambda_0 \nabla + \lambda_1) a_{t-1} + \nabla^2 a_t. \quad (4.3.13)$$

Дважды суммируя (4.3.13), получаем

$$z_t = \lambda_0 S a_{t-1} + \lambda_1 S^2 a_{t-1} + a_t, \quad (4.3.14)$$

так что веса ψ для этого процесса оказываются равными

$$\psi_0 = 1, \quad \psi_1 = (\lambda_0 + \lambda_1), \quad \dots, \quad \psi_j = (\lambda_0 + j\lambda_1), \quad \dots$$

Важное преимущество использования для рассматриваемой модели выражений (4.3.13) или (4.3.14) по сравнению с (4.3.11) становится очевидным, если положить в (4.3.13) $\lambda_1 = 0$. Тогда

$$\nabla z_t = [1 - (1 - \lambda_0) B] a_t,$$

что соответствует процессу $(0, 1, 1)$ с $\theta = 1 - \lambda_0$. Однако, если мы положим в (4.3.11) $\theta_2 = 0$, получаем

$$\nabla^2 z_t = (1 - \theta_1 B) a_t.$$

Далее, в гл. 5, будет показано, что для рядов, генерируемых процессом $(0, 2, 2)$, оптимальные прогнозы лежат вдоль прямой линии, уровень и наклон которой непрерывно подправляются по мере поступления новых данных. Напротив, ряд, генерируемый процессом $(0, 1, 1)$, может давать информацию только для непрерывного подправления уровня, а не наклона. Во многих случаях весьма важно, можно ли предсказать и подправить линейный тренд и уровень. Если необходимо сделать выбор в пользу одной из этих моделей, то вопрос сводится к тому, равно или не равно нулю λ_1 в (4.3.14).

Область обратимости для ПСС $(0, 2, 2)$ та же, что и для СС (2) в гл. 3. Она может быть определена в координатах θ или

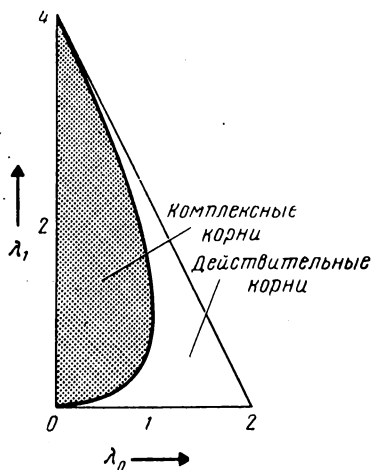


Рис. 4.8. Область обратимости для параметров λ_0 и λ_1 процесса ПСС $(0, 2, 2)$.

λ как

$$\begin{aligned} \theta_2 + \theta_1 &< 1, & 0 < 2\lambda_0 + \lambda_1 < 4, \\ \theta_2 - \theta_1 &< 1, & \lambda_1 > 0, \\ -1 < \theta_2 < 1, & \lambda_0 > 0. \end{aligned} \quad (4.3.15)$$

Треугольная область для θ была показана на рис. 3.8; соответствующая область для λ показана на рис. 4.8.

Усеченная форма представления модели через случайные импульсы. При помощи (4.3.14) усеченная формула для модели (4.2.17) может быть представлена в виде

$$z_t = C_k(t-k) + \lambda_0 \sum_{j=k+1}^{t-1} a_j + \lambda_1 \sum_{i=k+1}^{t-1} \sum_{j=k+1}^i a_j + a_t, \quad (4.3.16)$$

где функция $C_k(t-k)$ — общее решение однородного уравнения

$$(1-B)^2 C_k(t-k) = 0,$$

т. е.

$$C_k(t-k) = b_0^{(k)} + b_1^{(k)}(t-k),$$

и является полиномом от $(t-k)$ первой степени с коэффициентами, зависящими от положения начала отсчета k .

В приложении П4.3 показано, что эта функция может быть явно выражена через a :

$$C_k(t-k) = [(\lambda_0 - \lambda_1) S a_k + \lambda_1 S^2 a_k] + [\lambda_1 S a_k](t-k), \quad (4.3.17)$$

так что

$$\begin{aligned} b_0^{(k)} &= (\lambda_0 - \lambda_1) S a_k + \lambda_1 S^2 a_k, \\ b_1^{(k)} &= \lambda_1 S a_k. \end{aligned}$$

Из рассмотрения разностей $b_0^{(k)} - b_0^{(k-1)}$ и $b_1^{(k)} - b_1^{(k-1)}$ следует, что если начало отсчета перенесено из $k-1$ в k , то b_0 и b_1 корректируются согласно формулам

$$\begin{aligned} b_0^{(k)} &= b_0^{(k-1)} + b_1^{(k-1)} + \lambda_0 a_k, \\ b_1^{(k)} &= b_1^{(k-1)} + \lambda_1 a_k. \end{aligned} \quad (4.3.18)$$

Мы видим, что если эта модель подходит для описания процесса, математическое ожидание будущего поведения ряда по данным, доступным к моменту k , представляется прямой линией (4.3.17) с начальным положением $b_0^{(k)}$ и наклоном $b_1^{(k)}$. Фактически в момент времени t процесс отклонится от этой линии из-за влияния случайной компоненты

$$\lambda_0 \sum_{j=k+1}^{t-1} a_j + \lambda_1 \sum_{i=k+1}^{t-1} \sum_{j=k+1}^i a_j + a_t,$$

непредсказуемой в момент времени k . Далее, при переносе начала отсчета из $k-1$ в k данные о начальном положении и наклоне прямой должны быть скорректированы, согласно (4.3.18).

Обращенное представление модели. Наконец, рассмотрим модель в обращенном представлении

$$z_t = \sum_{j=1}^{\infty} \pi_j z_{t-j} + a_t = \bar{z}_{t-1}(\pi) + a_t.$$

Пользуясь (4.2.22) и приравнивая коэффициенты при одинаковых степенях B в

$$1 - 2B + B^2 = (1 - \theta_1 B - \theta_2 B^2)(1 - \pi_1 B - \pi_2 B^2 - \dots),$$

находим, что π -веса процесса ПСС(0, 2, 2) равны

$$\begin{aligned} \pi_1 &= 2 - \theta_1 = \lambda_0 + \lambda_1, \\ \pi_2 &= \theta_1(2 - \theta_1) - (1 + \theta_2) = \lambda_0 + 2\lambda_1 - (\lambda_0 + \lambda_1)^2, \quad (4.3.19) \\ (1 - \theta_1 B - \theta_2 B^2) \pi_j &= 0, \quad j \geq 3, \end{aligned}$$

где B действует на j .

Если корни характеристического уравнения $1 - \theta_1 B - \theta_2 B^2 = 0$ действительны, π -веса предшествующих z являются наложением двух затухающих экспонент. Если корни комплексные,

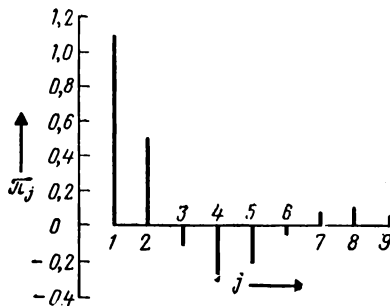


Рис. 4.9. Веса π процесса ПСС(0,2,2) с $\lambda_0 = 0,5$, $\lambda_1 = 0,6$.

веса ведут себя, как затухающая синусоида. На рис. 4.9 показаны веса для процесса при $\theta_1 = 0,9$ и $\theta_2 = -0,5$ (соответственно при $\lambda_0 = 0,5$ и $\lambda_1 = 0,6$).

Из рис. 3.9 и 4.8 видно, что при этих значениях коэффициентов характеристическое уравнение имеет комплексные корни (так как его дискриминант $\theta_1^2 + 4\theta_2 = -1,19 < 0$). Отсюда веса на рис. 4.9 должны вести себя как значения затухающей синусоиды, как это и происходит на самом деле.

4.3.3. Общий процесс проинтегрированного скользящего среднего порядка $(0, d, q)$

Представление разностным уравнением. Общий процесс проинтегрированного скользящего среднего порядка $(0, d, q)$ имеет вид

$$\nabla^d z_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t = \theta(B) a_t, \quad (4.3.20)$$

где нули функции $\theta(B)$ должны лежать вне единичного круга, если процесс обратим. Модель (4.3.20) может быть выражена в явном виде через прошлые значения z и a :

$$z_t = dz_{t-1} - \frac{1}{2}d(d-1)z_{t-2} + \dots + (-1)^{d+1}z_{t-d} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}.$$

Представление через случайные импульсы. Для того чтобы выразить z_t через a , запишем оператор в правой части (4.3.20) через $\nabla = 1 - B$. Получаем

$$(1 - \theta_1 B - \dots - \theta_q B^q) = (\lambda_{d-q} \nabla^{q-1} + \dots + \lambda_0 \nabla^{d-1} + \dots + \lambda_{d-1}) B + \nabla^d, \quad (4.3.21)$$

где λ , как и прежде, можно явно выразить через θ , приравняв коэффициенты при одинаковых степенях B .

Подставляя (4.3.21) в (4.3.20) и суммируя d раз, получаем

$$z_t = (\lambda_{d-q} \nabla^{q-d-1} + \dots + \lambda_0 S + \dots + \lambda_{d-1} S^d) a_{t-1} + a_t. \quad (4.3.22)$$

При $q > d$ из (4.3.22) следует, что в дополнение к d суммам мы получили $q - d$ дополнительных членов вида $\nabla^{q-d-1} a_{t-1}$, содержащих $a_{t-1}, a_{t-2}, \dots, a_{t+d-q}$.

Если мы представим это решение в виде конечных сумм импульсов a , поступающих в систему, начиная с некоторого момента k , мы получим уравнение в той же форме, но с дополнительным членом в виде функции, являющейся общим решением однородного уравнения

$$\nabla^d C_k(t-k) = 0,$$

которое представляет собой полином

$$C_k(t-k) = b_0^{(k)} + b_1^{(k)}(t-k) + b_2^{(k)}(t-k)^2 + \dots + b_{d-1}^{(k)}(t-k)^{d-1}.$$

Как и ранее, функция $C_k(t-k)$ описывает ту составляющую процесса, которая предсказуема в момент времени k . Аналогично коэффициенты $b_j^{(k)}$ могут быть представлены бесконечными суммами до момента времени k , т. е. $Sa_k, S^2a_k, \dots, S^d a_k$. Мы опять-таки можем найти, как изменятся эти коэффициенты при сдвиге начала отсчета из $k-1$ в k .

Обращенное представление модели. Наконец, модель может быть представлена обращенной формулой

$$\pi(B)z_t = a_t$$

или

$$z_t = \bar{z}_{t-1}(\pi) + a_t.$$

Веса π могут быть найдены приравниванием коэффициентов при одинаковых степенях B в (4.2.22), а именно

$$(1 - B)^d = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)(1 - \pi_1 B - \pi_2 B^2 - \dots). \quad (4.3.23)$$

Для данной модели их удобнее находить подстановкой численных значений в (4.3.23), нежели с помощью общей формулы.

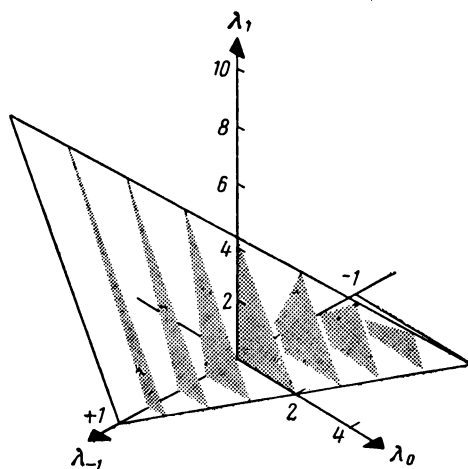


Рис. 4.10. Область обратимости для параметров λ_{-1} , λ_0 и λ_1 процесса ПСС(0,2,3).

Заметим, что из (4.3.23) вытекает, что для j , больших чем d и q , веса π удовлетворяют разностному уравнению

$$\theta(B)\pi_j = 0,$$

определяемому оператором скользящего среднего. Отсюда следует, что для достаточно больших j веса π_j ведут себя как значения суммы затухающих экспонент и синусоид.

Процесс ПСС порядка (0, 2, 3). Наконец, рассмотрим еще один достаточно интересный пример — процесс ПСС порядка (0, 2, 3):

$$\nabla^2 z_t = (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3) a_t.$$

Тем же способом, что и ранее, можно записать процесс в проинтегрированной форме как

$$z_t = \lambda_{-1} a_{t-1} + \lambda_0 S a_{t-1} + \lambda_1 S^2 a_{t-1} + a_t,$$

где соотношения между λ и θ имеют вид

$$\begin{aligned} \theta_1 &= 2 - \lambda_{-1} - \lambda_0 - \lambda_1, & \lambda_{-1} &= -\theta_3, \\ \theta_2 &= \lambda_0 - 1 + 2\lambda_{-1}, & \lambda_0 &= 1 + \theta_2 + 2\theta_3, \\ \theta_3 &= -\lambda_{-1}, & \lambda_1 &= 1 - \theta_1 - \theta_2 - \theta_3. \end{aligned}$$

Можно также воспользоваться усеченным представлением

$$z_t = b_0^{(k)} + b_1^{(k)}(t - k) + \lambda_{-1} a_{t-1} + \lambda_0 \sum_{j=k+1}^{t-1} a_j + \lambda_1 \sum_{i=k+1}^{t-1} \sum_{j=k+1}^i a_j + a_t.$$

Наконец, область обратимости определена неравенствами

$$\begin{aligned} \theta_1 + \theta_2 + \theta_3 &< 1, & \lambda_1 &> 0, \\ -\theta_1 + \theta_2 - \theta_3 &< 1, & 2\lambda_0 + \lambda_1 &< 4(1 - \lambda_{-1}), \\ \theta_3(\theta_3 - \theta_1) - \theta_2 &< 1, & \lambda_0(1 + \lambda_{-1}) &> -\lambda_1\lambda_{-1}, \\ |\theta_3| &< 1, & |\lambda_{-1}| &< 1 \end{aligned}$$

и показана на рис. 4.10.

В гл. 5 мы покажем, как можно оптимально предсказывать будущие значения временного ряда для модели АРПСС. При изучении способов прогноза мы будем широко пользоваться различными формами представления моделей, рассмотренными в этой главе.

ПРИЛОЖЕНИЕ П4.1. ЛИНЕЙНЫЕ РАЗНОСТНЫЕ УРАВНЕНИЯ

В этой книге мы часто рассматриваем линейные разностные уравнения. В частности, модель АРПСС связывает выход z_t со входом a_t с помощью разностного уравнения

$$\begin{aligned} z_t - \Phi_1 z_{t-1} - \Phi_2 z_{t-2} - \dots - \Phi_{p'} z_{t-p'} &= \\ &= a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}, \end{aligned} \quad (\text{П4.1.1})$$

где $p' = p + d$.

Уравнение (П4.1.1) можно также записать в виде

$$\Phi(B) z_t = \Theta(B) a_t,$$

где

$$\begin{aligned} \Phi(B) &= 1 - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_{p'} B^{p'}, \\ \Theta(B) &= 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q. \end{aligned}$$

Ниже мы дадим вывод общего решения разностного уравнения (П4.1.1) с начальными условиями, заданными в момент времени $k > t$.

1) Вначале будет показано, что общее решение можно записать в виде

$$z_t = C_k(t - k) + I_k(t - k),$$

где $C_k(t - k)$ — общее решение соответствующего однородного уравнения и $I_k(t - k)$ — частное решение нашего уравнения.

2) Будет получена общая формула для функции $C_k(t - k)$.

3) Наконец, будет выведена общая формула для частного решения $I_k(t - k)$.

Общее решение. Наши рассуждения аналогичны применяемым при выводе решения линейных дифференциальных или линейных алгебраических уравнений. Положим, что z'_t — любое частное решение уравнения

$$\varphi(B) z_t = \theta(B) a_t, \quad (\text{П4.1.2})$$

т. е.

$$\varphi(B) z'_t = \theta(B) a_t. \quad (\text{П4.1.3})$$

Вычитая (П4.1.3) из (П4.1.2), получаем

$$\varphi(B)(z_t - z'_t) = 0.$$

Таким образом, $z''_t = z_t - z'_t$ удовлетворяет уравнению

$$\varphi(B) z''_t = 0. \quad (\text{П4.1.4})$$

Тогда

$$z_t = z'_t + z''_t,$$

и, следовательно, общее решение (П4.1.2) есть сумма z''_t общего решения однородного разностного уравнения (П4.1.4) и z'_t — произвольного частного решения (П4.1.2). В дальнейшем мы обозначаем z''_t как $C_k(t - k)$, а z'_t как $I_k(t - k)$.

Вычисление функции $C_k(t - k)$. Простые корни. Рассмотрим однородное разностное уравнение

$$\varphi(B) z_t = 0, \quad (\text{П4.1.5})$$

где

$$\varphi(B) = (1 - G_1 B)(1 - G_2 B) \dots (1 - G_{p'} B), \quad (\text{П4.1.6})$$

и предположим вначале, что $G_1, G_2, \dots, G_{p'}$ все отличны друг от друга. Тогда, как будет показано ниже, общее решение (П4.1.5) в момент t для ряда, начинающегося в момент k , имеет вид

$$z_t = A_1 G_1^{t-k} + A_2 G_2^{t-k} + \dots + A_{p'} G_{p'}^{t-k}, \quad (\text{П4.1.7})$$

где A_i — константы. Тогда действительный корень уравнения $\varphi(B) = 0$ создает затухающее экспоненциальное слагаемое G^{t-k} в (П4.1.7). Пара сопряженных комплексных корней создает затухающее синусоидальное слагаемое в (П4.1.7).

Чтобы убедиться, что (П4.1.7) действительно удовлетворяет (П4.1.5), подставим (П4.1.7) в (П4.1.5) и получим

$$\varphi(B)(A_1 G_1^{t-k} + A_2 G_2^{t-k} + \dots + A_{p'} G_{p'}^{t-k}) = 0. \quad (\text{П4.1.8})$$

Рассмотрим теперь

$$\begin{aligned} \varphi(B) G_i^{t-k} &= (1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_{p'} B^{p'}) G_i^{t-k} = \\ &= G_i^{t-k-p'} (G_i^{p'} - \varphi_1 G_i^{p'-1} - \dots - \varphi_{p'}). \end{aligned}$$

Видно, что $\varphi(B) G_i^{t-k}$ обращается в нуль для каждого i , если

$$G_i^{p'} - \varphi_1 G_i^{p'-1} - \dots - \varphi_{p'} = 0,$$

т. е. если $B = 1/G_i$ является корнем уравнения $\varphi(B) = 0$. Так как из (П4.1.6) следует, что корни $\varphi(B) = 0$ — это $B = 1/G_i$, то $\varphi(B) G_i^{t-k} = 0$ при всех i и, следовательно, справедливо (П4.1.8). Это подтверждает, что (П4.1.7) есть общее решение уравнения (П4.1.5)*).

Чтобы доказать (П4.1.7) непосредственно, рассмотрим частный случай уравнения второго порядка

$$(1 - G_1 B)(1 - G_2 B) z_t = 0,$$

которое можно записать как

$$(1 - G_1 B) y_t = 0, \quad (\text{П4.1.9})$$

где

$$y_t = (1 - G_2 B) z_t. \quad (\text{П4.1.10})$$

Из (П4.1.9) вытекает

$$y_t = G_1 y_{t-1} = G_1^2 y_{t-2} = \dots = G_1^{t-k} y_k$$

и, следовательно,

$$y_t = D_1 G_1^{t-k},$$

*) Нужно еще доказать, что функции $G_1^{t-k}, \dots, G_{p'}^{t-k}$ линейно независимы. Это следует из того, что определитель, составленный из p последовательных значений этих функций, является определителем Вандермонда, отличным от нуля из-за того, что $G_1, \dots, G_{p'}$ различны. — Прим. ред.

где $D_1 = y_k$ — константа, определяемая начальным значением y_k . Поэтому (П4.1.10) можно записать как

$$\begin{aligned}
 z_t &= G_2 z_{t-1} + D_1 G_1^{t-k} = \\
 &= G_2 (G_2 z_{t-2} + D_1 G_1^{t-k-1}) + D_1 G_1^{t-k} = \\
 &\quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\
 &= G_2^{t-k} z_k + D_1 \{G_1^{t-k} + G_2 G_1^{t-k-1} + \dots + G_2^{t-k-1} G_1\} = \\
 &= G_1^{t-k} z_k + \frac{D_1}{1 - G_1/G_2} (G_1^{t-k} - G_2^{t-k}) = \\
 &= A_1 G_1^{t-k} + A_2 G_2^{t-k}, \tag{П4.1.11}
 \end{aligned}$$

где A_1, A_2 — константы, определяемые начальными значениями ряда. Эти рассуждения можно продолжить и показать, что общее решение (П4.1.5) при простых корнях $\varphi(B) = 0$ определяется формулой (П4.1.7).

Кратные корни. Примем теперь, что $\varphi(B) = 0$ имеет d -кратный корень G_0^{-1} , так что $\varphi(B)$ содержит множитель $(1 - G_0 B)^d$. В частности, рассмотрим решение (П4.1.11) уравнения второго порядка, у которого $G_1 = G_2 = G_0$. Тогда (П4.1.11) перейдет в

$$z_t = G_0^{t-k} z_k + D_1 G_0^{t-k} (t - k),$$

или

$$z_t = [A_0 + A_1 (t - k)] G_0^{t-k}.$$

В общем случае, если имеется d -кратный корень G_0 , как можно проверить прямой подстановкой в (П4.1.5), общее решение имеет вид

$$z_t = [A_0 + A_1 (t - k) + A_2 (t - k)^2 + \dots + A_{d-1} (t - k)^{d-1}] G_0^{t-k}. \tag{П4.1.12}$$

В частности, когда $G_0 = 1$, как в процессе ПСС $(0, d, q)$, решение принимает вид

$$z_t = A_0 + A_1 (t - k) + \dots + A_{d-1} (t - k)^{d-1}, \tag{П4.1.13}$$

т. е. является полиномом по $t - k$ степени $d - 1$.

В общем случае, когда $\varphi(B)$ можно разложить на множители

$$(1 - G_1 B)(1 - G_2 B) \dots (1 - G_p B)(1 - G_0 B)^d,$$

общее решение однородного уравнения $C_k(t - k)$ имеет вид

$$C_k(t - k) = G_0^{t-k} \sum_{j=0}^{d-1} A_j (t - k)^j + \sum_{i=1}^p D_i G_i^{t-k}, \tag{П4.1.14}$$

т. е. состоит из суммы затухающих экспоненциальных членов G^{t-k} , полиномиальных членов $(t-k)^j$, затухающих синусоидальных членов $d^{t-k} \sin(2\pi f_0 t + F)$ и их комбинаций.

Нахождение частного решения. Покажем теперь, что частное решение $I_k(t-k)$, удовлетворяющее уравнению

$$\varphi(B) I_k(t-k) = \theta(B) a_t, \quad t-k > q, \quad (\text{П4.1.15})$$

есть

$$\begin{aligned} I_k(s-k) &= 0, & s &\leq k, \\ I_k(1) &= a_{k+1}, \\ I_k(2) &= a_{k+2} + \psi_1 a_{k+1}, \end{aligned} \quad (\text{П4.1.16})$$

$$\begin{array}{ccc} \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \end{array}$$

$$I_k(t-k) = a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots + \psi_{t-k-1} a_{k+1}, \quad t > k,$$

где веса ψ те же, что и в представлении модели (4.2.3). Они удовлетворяют уравнению

$$\varphi(B) \psi(B) a_t = \theta(B) a_t. \quad (\text{П4.1.17})$$

Левая часть уравнения (П4.1.17) может быть представлена в виде

$$\begin{array}{l|l} a_t + \psi_1 a_{t-1} + \psi_2 a_{t-2} + \dots + \psi_{t-k-1} a_{k+1} & + \psi_{t-k} a_k + \dots \\ -\varphi_1 (& a_{t-1} + \psi_1 a_{t-2} + \dots + \psi_{t-k-2} a_{k+1} & + \psi_{t-k-1} a_k + \dots \\ -\varphi_2 (& \dots \dots \dots & \dots \dots \dots \\ \vdots & \vdots \vdots \vdots \vdots & \vdots \vdots \vdots \\ -\varphi_{p'} (& a_{t-p'} + \dots + \psi_{t-k-p'-1} a_{k+1} & + \psi_{t-k-p'} a_k + \dots \end{array} \quad (\text{П4.1.18})$$

а правая часть (П4.1.17) имеет вид

$$a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}.$$

Отсюда следует, что первые $q+1$ столбцов в этой таблице дают при суммировании по столбцу суммы $a_t, -\theta_1 a_{t-1}, \dots, -\theta_q a_{t-q}$.

Далее, левая часть в (П4.1.15), где $I_k(s-k)$ определено равенствами (П4.1.16), равна сумме членов в первых $(t-k)$ столбцах (отделенных справа вертикальной линией). Следовательно, если $t-k > q$, т. е. вертикальная линия проведена за q -м столбцом, сумма всех членов до вертикальной линии равна $\theta(B) a_t$. Это и показывает, что (П4.1.16) — частное решение разностного уравнения.

Пример. Рассмотрим процесс ПСС(0, 1, 1)

$$z_t - z_{t-1} = a_t - \theta a_{t-1}, \quad (\text{П4.1.19})$$

у которого $\psi_j = 1 - \theta$ для $j \geq 1$. Имеем

$$\begin{aligned} I_k(0) &= 0, \\ I_k(1) &= a_{k+1}, \\ &\vdots \\ I_k(t-k) &= a_t + (1 - \theta) \sum_{j=k+1}^{t-1} a_j, \quad t - k > 2. \end{aligned} \quad (\text{П4.1.20})$$

Если $z_t = I_k(t-k)$ — решение (П4.1.19), то

$$I_k(t-k) - I_k(t-k-1) = a_t - \theta a_{t-1},$$

и, как легко проверить, эта функция удовлетворяет (П4.1.20) при $t-k > 1$, т. е. при $t-k > q$.

ПРИЛОЖЕНИЕ П4.2. ПРОЦЕСС ПСС (0, 1, 1) С ДЕТЕРМИНИРОВАННЫМ ДРЕЙФОМ НУЛЯ

Общая модель $\phi(B) \nabla^d z_t = \theta_0 + \theta(B) a_t$ может быть также представлена в виде

$$\phi(B) \nabla^d z_t = \theta(B) \varepsilon_t,$$

где импульсы ε_t имеют ненулевое среднее значение $\xi = \theta_0 / (1 - \theta_1 - \dots - \theta_q)$. Например, модель ПСС(0, 1, 1) будет иметь вид

$$\nabla z_t = (1 - \theta B) \varepsilon_t,$$

причем $E[\varepsilon_t] = \xi = \theta_0 / (1 - \theta)$. В таком виде z_t может представлять выходную температуру реактора, тепло к которому поступает от нагревательного элемента с постоянной скоростью. Если

$$\varepsilon_t = \xi + a_t, \quad (\text{П4.2.1})$$

где a_t — белый шум с нулевым средним значением, тогда по отношению к началу отсчета k проинтегрированная форма представления модели будет

$$z_t = b_0^{(k)} + \lambda \sum_{j=k+1}^{t-1} \varepsilon_j + \varepsilon_t \quad (\text{П4.2.2})$$

с $\lambda = 1 - \theta$. Подставляя (П4.2.1) в (П4.2.2), выразим модель через a :

$$z_t = b_0 + \lambda \xi (t - k - 1) + \xi + \lambda \sum_{j=k+1}^{t-1} a_j + a_t. \quad (\text{П4.2.3})$$

Видно, что z_t содержит детерминированный наклон, или дрейф, описываемый членом $\lambda \xi (t - k - 1)$. Далее, если обозна-

чить «уровень» процесса в момент $t - 1$ как l_{t-1} , где

$$z_t = l_{t-1} + a_t,$$

то при переходе от времени $t - 1$ к t уровень изменится согласно формуле

$$l_t = l_{t-1} + \lambda \xi + \lambda a_t.$$

Изменение уровня содержит, таким образом, детерминированную компоненту $\lambda \xi$, а также стохастическую компоненту λa_t .

ПРИЛОЖЕНИЕ П4.3. СВОЙСТВА КОНЕЧНОГО ОПЕРАТОРА СУММИРОВАНИЯ

Соотношение между конечной и бесконечной суммами. В ряде приведенных ниже простых случаев можно проверить непосредственно, что соотношение между бесконечной и конечной суммами, первый член которых имеет индекс $k + 1$, где $t \geq k + 1$, имеет вид

$$Sx_t = \sum_{h=k+1}^t x_h + Sx_k, \quad (П4.3.1)$$

$$S^2x_t = \sum_{i=k+1}^t \sum_{h=k+1}^i x_h + S^2x_k + (t - k) Sx_k.$$

Последнее равенство можно также записать в виде

$$S^2x_t = \sum_{i=k+1}^t \sum_{h=k+1}^i x_h + b_0^{(k)} + b_1^{(k)}(t - k).$$

В общем d -кратная бесконечная сумма будет равна d -кратной конечной сумме плюс полином от t степени $d - 1$

$$b_0^{(k)} + b_1^{(k)}(t - k) + b_2^{(k)} \binom{t - k + 1}{2} + \dots + b_{d-1}^{(k)} \binom{t - k + d - 2}{d - 1}.$$

Применение к процессу ПСС(0, 2, 2). Проинтегрированная форма представления (4.3.14) процесса ПСС(0, 2, 2) имеет вид

$$z_t = \lambda_0 S a_{t-1} + \lambda_1 S^2 a_{t-1} + a_t;$$

используя (П4.3.1), это равенство можно преобразовать к

$$\begin{aligned} z_t &= \lambda_0 \left(\sum_{h=k+1}^{t-1} a_h + S a_k \right) + \lambda_1 \left\{ \sum_{i=k+1}^{t-1} \sum_{h=k+1}^i a_h + \right. \\ &\quad \left. + S^2 a_k + (t - k - 1) S a_k \right\} + a_t = \\ &= [(\lambda_0 - \lambda_1) S a_k + \lambda_1 S^2 a_k] + \lambda_1 S a_k (t - k) + \\ &\quad + \lambda_0 \sum_{h=k+1}^{t-1} a_h + \lambda_1 \sum_{i=k+1}^{t-1} \sum_{h=k+1}^i a_h + a_t. \end{aligned} \quad (П4.3.2)$$

Усеченная форма представления модели через случайные импульсы имеет вид

$$z_t = C_k(t-k) + \lambda_0 \sum_{h=k+1}^{t-1} a_h + \lambda_1 \sum_{i=k+1}^{t-1} \sum_{h=k+1}^i a_h + a_t; \quad (\text{П4.3.3})$$

приравнивая (П4.3.2) и (П4.3.3), можно получить выражение для функции $C_k(t-k)$ в виде бесконечной суммы случайных импульсов

$$C_k(t-k) = [(\lambda_0 - \lambda_1) S a_k + \lambda_1 S^2 a_k] + [\lambda_1 S a_k](t-k),$$

что совпадает с формулой (4.3.17), приведенной в основном тексте.

ПРИЛОЖЕНИЕ П4.4. ПРОЦЕССЫ АРПСС С ДОБАВЛЕННЫМ ШУМОМ

В этом приложении мы рассмотрим эффект добавления к общему процессу АРПСС (p, d, q) шума (например, ошибок измерений).

П4.4.1. Сумма двух независимых процессов скользящего среднего

В качестве необходимого введения ниже нам понадобится следующий факт. Рассмотрим стохастический процесс w_t , являющийся суммой двух независимых процессов скользящего среднего порядков q_1 и q_2 соответственно:

$$w_t = \theta_1(B) a_t + \theta_2(B) b_t, \quad (\text{П4.4.1})$$

где $\theta_1(B)$ и $\theta_2(B)$ — полиномы от B порядка q_1 и q_2 , процессы a_t и b_t — взаимно независимые белые шумы с нулевыми средними значениями.

Обозначим $q = \max(q_1, q_2)$; ясно, что автоковариационная функция γ_j процесса w_t должна быть равна нулю при $j > q$. Отсюда следует, что существует представление w_t как единого процесса скользящего среднего порядка q

$$w_t = \theta_3(B) u_t, \quad (\text{П4.4.2})$$

где процесс u_t — белый шум с нулевым средним значением.

Итак, сумма двух независимых процессов скользящего среднего — это еще один процесс скользящего среднего, порядок которого совпадает с максимальным порядком слагаемых процессов.

П4.4.2. Эффект добавления шума к общей модели

Коррелированный шум. Рассмотрим общую нестационарную модель порядка (p, d, q)

$$\phi(B) \nabla^d z_t = \theta(B) a_t. \quad (\text{П4.4.3})$$

Предположим, что мы не можем наблюдать z_t в чистом виде, а только $Z_t = z_t + b_t$, где b_t представляет собой некоторый внешний шум (например, измерительную ошибку) и может быть коррелирован. Мы хотим определить природу наблюдаемого процесса Z_t .

В общем случае мы имеем

$$\phi(B) \nabla^d Z_t = \theta(B) a_t + \phi(B) \nabla^d b_t.$$

Если шум является стационарным процессом АРСС порядка $(p_1, 0, q_1)$

$$\phi_1(B) b_t = \theta_1(B) \alpha_t, \quad (\text{П4.4.4})$$

где процесс α_t — белый шум, независимый от процесса a_t , то

$$\frac{\phi_1(B) \phi(B) \nabla^d Z_t}{p_1 + p + d} = \frac{\phi_1(B) \theta(B) a_t}{p_1 + q} + \frac{\phi(B) \theta_1(B) \nabla^d \alpha_t}{p + q_1 + d}, \quad (\text{П4.4.5})$$

где выражения под скобками показывают степени различных полиномов от B . Правая часть (П4.4.5) имеет теперь форму (П4.4.1). Пусть $P = p_1 + p$ и Q равно большему из чисел $(p_1 + q)$ и $(p + q_1 + d)$. Тогда можно записать

$$\phi_2(B) \nabla^d Z_t = \theta_2(B) u_t,$$

где u_t — белый шум, Z_t — процесс порядка (P, d, Q) .

Белый шум. Если, как это может оказаться в некоторых приложениях, добавленный шум белый, тогда в (П4.4.4) $\phi_1(B) = \theta_1(B) = 1$ и мы получаем

$$\phi(B) \nabla^d Z_t = \theta_2(B) u_t \quad (\text{П4.4.6})$$

при

$$\theta_2(B) u_t = \theta(B) a_t + \phi(B) \nabla^d b_t,$$

имеющем порядок (p, d, Q) , где Q — наибольшее из q и $(p + d)$. Если $p + d \leq q$, порядок процесса с шумом тот же, что у первоначального процесса. Единственный эффект дополнительного белого шума — изменение значений θ (но не ϕ).

Эффект добавления белого шума к процессу проинтегрированного скользящего среднего. В частности, процесс ПСС порядка $(0, d, q)$ с добавленным белым шумом остается процессом того же порядка, если $d \leq q$; в противном случае он становится процессом ПСС порядка $(0, d, d)$. В любом случае параметры процесса изменяются при добавлении шума. Природа этих из-

менений может быть выяснена приравниванием автоковариаций процесса с добавленным шумом автоковариациям простого процесса ПСС. Пример такой процедуры приводится ниже.

П4.4.3. Пример процесса ПСС (0, 1, 1) с добавленным белым шумом

Рассмотрим свойства процесса $Z_t = z_t + b_t$, когда

$$z_t = \lambda \sum_{j=1}^{\infty} a_{t-j} + a_t \quad (\text{П4.4.7})$$

и b_t и a_t — взаимно независимые процессы типа белого шума. Процесс Z_t имеет первую разность $W_t = Z_t - Z_{t-1}$ вида

$$W_t = [1 - (1 - \lambda)B]a_t + (1 - B)b_t. \quad (\text{П4.4.8})$$

Автоковариации для первых разностей W_t имеют вид

$$\begin{aligned} \gamma_0 &= \sigma_a^2 [1 + (1 - \lambda)^2] + 2\sigma_b^2, \\ \gamma_1 &= -\sigma_a^2 (1 - \lambda) - \sigma_b^2, \\ \gamma_j &= 0, \quad j \geq 2. \end{aligned} \quad (\text{П4.4.9})$$

Тот факт, что все γ_j , кроме первого, равны нулю, подтверждает, что процесс с добавленным к нему шумом является, как и ожидалось, процессом ПСС порядка (0, 1, 1). Для того чтобы явно выписать параметры этого процесса, предположим, что он может быть представлен в виде

$$Z_t = \Lambda \sum_{j=1}^{\infty} u_{t-j} + u_t, \quad (\text{П4.4.10})$$

где процесс u_t — белый шум. Процесс (П4.4.10) имеет автоковариации

$$\begin{aligned} \gamma_0 &= \sigma_u^2 \{1 + (1 - \Lambda)^2\}, \\ \gamma_1 &= -\sigma_u^2 (1 - \Lambda), \\ \gamma_j &= 0, \quad j \geq 2. \end{aligned} \quad (\text{П4.4.11})$$

Приравнивая (П4.4.9) и (П4.4.11), мы можем найти явные выражения для Λ и σ_u^2 . Отсюда

$$\begin{aligned} \frac{\Lambda^2}{1 - \Lambda} &= \frac{\lambda^2}{1 - \lambda + \sigma_b^2 / \sigma_a^2}, \\ \sigma_u^2 &= \sigma_a^2 \frac{\lambda^2}{\Lambda^2}. \end{aligned} \quad (\text{П4.4.12})$$

Положим, например, что первоначальный ряд имеет $\lambda = 0,5$ и $\sigma_b^2 = \sigma_a^2$, тогда $\Lambda = 0,33$ и $\sigma_u^2 = 2,25\sigma_a^2$.

П4.4.4. Связь между процессом ПСС (0, 1, 1) и случайным блужданием

Процесс

$$z_t = \sum_{j=0}^{\infty} a_{t-j} = \sum_{j=1}^{\infty} a_{t-j} + a_t, \quad (\text{П4.4.13})$$

который является процессом ПСС(0, 1, 1) с $\lambda = 1$, иногда называется *случайным блужданием*. Если a_t — шаги, сделанные вперед или назад в момент времени t , тогда z_t представляет положение блуждающего объекта в момент t .

Любой процесс ПСС(0, 1, 1) можно трактовать как случайное блуждание, скрытое в белом шуме b_t , который не коррелирован с прошлыми импульсами a_t . Если обозначить «зашумленный» процесс $Z_t = z_t + b_t$, где z_t определено (П4.4.13), то, используя (П4.4.12), получим

$$Z_t = \Lambda \sum_{j=1}^{\infty} u_{t-j} + u_t,$$

где

$$\frac{\Lambda^2}{1 - \Lambda} = \frac{\sigma_a^2}{\sigma_b^2}, \quad \sigma_u^2 = \frac{\sigma_a^2}{\Lambda^2}. \quad (\text{П4.4.14})$$

П4.4.5. Автоковариационная функция общей модели с добавленным коррелированным шумом

Предположим, что основной процесс типа АРПСС порядка (p, d, q)

$$\phi(B) \nabla^d z_t = \theta(B) a_t$$

и что наблюдается процесс $Z_t = z_t + b_t$, где b_t — стационарный процесс с автоковариационной функцией $\gamma_j(b)$, независимый от процесса a_t и, следовательно, от z_t . Пусть $\gamma_i(w)$ — автоковариационная функция для $w_t = \nabla^d z_t = \phi^{-1}(B) \theta(B) a_t$, и пусть $W_t = \nabla^d Z_t$. Мы хотим найти автоковариационную функцию для W_t . Имеем

$$\nabla^d (Z_t - b_t) = \phi^{-1}(B) \theta(B) a_t,$$

$$W_t = w_t + v_t,$$

где

$$v_t = \nabla^d b_t = (1 - B)^d b_t.$$

Отсюда

$$\gamma_j(W) = \gamma_j(w) + \gamma_j(v),$$

$$\gamma_j(v) = (1 - B)^d (1 - F)^d \gamma_j(b) = (-1)^d (1 - B)^{2d} \gamma_{j+d}(b)$$

и, следовательно,

$$\gamma_j(W) = \gamma_j(w) + (-1)^d (1 - B)^{2d} \gamma_{j+d}(b). \quad (\text{П4.4.15})$$

В качестве примера рассмотрим случай, когда коррелированный шум b_t добавлен к процессу ПСС(0, 1, 1), определенному как $w_t = \nabla z_t = (1 - \theta B) a_t$. Тогда автоковариации первых разностей W_t «зашумленного процесса» будут

$$\begin{aligned} \gamma_0(W) &= \sigma_a^2 (1 + \theta^2) + 2[\gamma_0(b) - \gamma_1(b)], \\ \gamma_1(W) &= -\sigma_a^2 \theta + [2\gamma_1(b) - \gamma_0(b) - \gamma_2(b)], \\ \gamma_j(W) &= [2\gamma_j(b) - \gamma_{j-1}(b) - \gamma_{j+1}(b)], \quad j \geq 2. \end{aligned}$$

В частности, если b_t — процесс авторегрессии первого порядка, т. е.

$$b_t = \phi b_{t-1} + \alpha_t,$$

то

$$\begin{aligned} \gamma_0(W) &= \sigma_a^2 (1 + \theta^2) + 2\sigma_b^2 (1 - \phi), \\ \gamma_1(W) &= -\sigma_a^2 \theta - \sigma_b^2 (1 - \phi)^2, \\ \gamma_j(W) &= -\sigma_b^2 \phi^{j-1} (1 - \phi)^2, \quad j \geq 2. \end{aligned}$$

Согласно (П4.4.5), результирующий «зашумленный» процесс $Z_t = z_t + b_t$ в этом случае определен как

$$(1 - \phi B) \nabla Z_t = (1 - \phi B) (1 - \theta B) a_t + (1 - B) \alpha_t,$$

т. е. имеет порядок (1, 1, 2).

Прогнозирование

Рассмотрев в гл. 4 некоторые свойства моделей АРПСС, мы покажем теперь, как они могут быть использованы для прогноза будущих значений наблюдаемого временного ряда. Во второй части книги мы рассмотрим проблему подгонки модели к фактическим данным. Здесь же мы будем считать модель известной *точно*, полагая, что ошибки в оценках параметров не будут заметно влиять на прогноз, что выполняется, если только число данных, используемых для подгонки модели, не слишком мало.

В этой главе мы рассмотрим несезонные временные ряды. Прогнозирование сезонных временных рядов описано в гл. 9. Мы покажем, как прогнозы с минимальной среднеквадратичной ошибкой можно получить непосредственно из представления модели в виде *разностного уравнения*.

Дальнейший рекуррентный расчет позволяет оценить доверительные интервалы прогнозов. Нужно подчеркнуть, что для практических вычислений прогнозов подход, основанный на использовании представления модели разностным уравнением, — простейший и наиболее изящный. Однако для лучшего понимания природы прогнозов мы должны исследовать их и с других позиций.

5.1. ПРОГНОЗЫ С МИНИМАЛЬНОЙ СРЕДНЕКВАДРАТИЧНОЙ ОШИБКОЙ И ИХ СВОЙСТВА

В разд. 4.2 мы рассмотрели три способа представления общей модели АРПСС

$$\varphi(B) z_t = \theta(B) a_t, \quad (5.1.1)$$

где $\varphi(B) = \phi(B) \nabla^d$. Мы начнем с того, что напомним эти три представления, так как каждое освещает различные стороны проблемы прогнозирования.

Мы займемся прогнозированием значения z_{t+l} , $l \geq 1$, в текущий момент времени t . Про такой прогноз будем говорить, что он делается в *момент t с упреждением l* . Подытожим теперь результаты разд. 4.2, заменив t на $t+l$ и k на t .

Три явные формы представления модели. Наблюдение z_{t+l} , генерируемое процессом (5.1.1), можно выразить следующим образом:

1) Непосредственно при помощи разностного уравнения

$$z_{t+l} = \varphi_1 z_{t+l-1} + \dots + \varphi_{p+d} z_{t+l-p-d} - \theta_1 a_{t+l-1} - \dots - \theta_q a_{t+l-q} + a_{t+l}. \quad (5.1.2)$$

2) Как бесконечную взвешенную сумму текущего и предшествующих импульсов a_j

$$z_{t+l} = \sum_{j=-\infty}^{t+l} \psi_{t+l-j} a_j = \sum_{j=0}^{\infty} \psi_j a_{t+l-j}, \quad (5.1.3)$$

где $\psi_0 = 1$, и, как в (4.2.5), веса ψ можно найти приравнянием коэффициентов в

$$\varphi(B)(1 + \psi_1 B + \psi_2 B^2 + \dots) = \theta(B). \quad (5.1.4)$$

Для $l > q$ модель может быть эквивалентно представлена в усеченной форме

$$z_{t+l} = C_t(l) + a_{t+l} + \psi_1 a_{t+l-1} + \dots + \psi_{l-1} a_{t+1}, \quad (5.1.5)$$

где функция $C_t(l)$ равна усеченной бесконечной сумме

$$C_t(l) = \sum_{j=-\infty}^t \psi_{t+l-j} a_j = \sum_{j=0}^{\infty} \psi_{l+j} a_{t-j}. \quad (5.1.6)$$

3) Как бесконечную взвешенную сумму предыдущих наблюдений плюс случайный импульс

$$z_{t+l} = \sum_{j=1}^{\infty} \pi_j z_{t+l-j} + a_{t+l}. \quad (5.1.7)$$

Если $d \geq 1$, то

$$\bar{z}_{t+l-1}(\pi) = \sum_{j=1}^{\infty} \pi_j z_{t+l-j} \quad (5.1.8)$$

будет взвешенным средним, так как при этом $\sum_{j=1}^{\infty} \pi_j = 1$. Как и в (4.2.22), веса π можно получить из равенства

$$\varphi(B) = (1 - \pi_1 B - \pi_2 B^2 - \dots) \theta(B). \quad (5.1.9)$$

5.1.1. Вывод формулы для прогнозов с минимальной среднеквадратичной ошибкой

Положим теперь, что, находясь в начале отсчета t , мы должны сделать прогноз $\hat{z}_t(l)$ величины z_{t+l} , который был бы линейной функцией текущего и предшествующих наблюдений

$z_t, z_{t-1}, z_{t-2}, \dots$. Тогда он будет также линейной функцией текущего и предшествующих импульсов $a_t, a_{t-1}, a_{t-2}, \dots$.

Предположим, что наилучший прогноз — это

$$\hat{z}_t(l) = \psi_l^* a_t + \psi_{l+1}^* a_{t-1} + \psi_{l+2}^* a_{t-2} + \dots,$$

где веса $\psi_l^*, \psi_{l+1}^*, \dots$ должны быть определены. Тогда, учитывая (5.1.3), находим, что среднеквадратичная ошибка прогноза равна

$$E[z_{t+l} - \hat{z}_t(l)]^2 = (1 + \psi_1^2 + \dots + \psi_{l-1}^2) \sigma_a^2 + \sum_{j=0}^{\infty} (\psi_{l+j} - \psi_{l+j}^*)^2 \sigma_a^2. \quad (5.1.10)$$

Она может быть минимизирована приравниванием $\psi_{l+j}^* = \psi_{l+j}$. Этот вывод является частным случаем более общих результатов теории прогнозирования, развитой Волдом [44], Колмогоровым [45—47], Винером [48] и Уиттлом [49]. В результате имеем

$$z_{t+l} = (a_{t+l} + \psi_1 a_{t+l-1} + \dots + \psi_{l-1} a_{t+1}) + (\psi_l a_t + \psi_{l+1} a_{t-1} + \dots), \quad (5.1.11)$$

$$z_{t+l} = e_t(l) + \hat{z}_t(l), \quad (5.1.12)$$

где $e_t(l)$ — ошибка прогноза $\hat{z}_t(l)$ с упреждением l .

Отсюда вытекает ряд важных выводов. Как и ранее, обозначим условное математическое ожидание z_{t+l} при условии, что все z до момента t известны, через $E_t[z_{t+l}]$. Тогда

$$1) \quad \hat{z}_t(l) = \psi_l a_t + \psi_{l+1} a_{t-1} + \dots = E_t[z_{t+l}]. \quad (5.1.13)$$

Это значит, что прогноз с минимальной среднеквадратичной ошибкой в момент t с упреждением l есть условное математическое ожидание z_{t+l} в момент t . Когда $\hat{z}_t(l)$ рассматривается как функция l при фиксированном t , мы будем называть ее *прогнозирующей функцией* для момента t .

2) Ошибка прогноза для упреждения l равна

$$e_t(l) = a_{t+l} + \psi_1 a_{t+l-1} + \dots + \psi_{l-1} a_{t+1}. \quad (5.1.14)$$

Так как

$$E_t[e_t(l)] = 0, \quad (5.1.15)$$

прогноз будет несмещенный. Далее, дисперсия ошибки прогноза равна

$$V(l) = \text{var}[e_t(l)] = (1 + \psi_1^2 + \psi_2^2 + \dots + \psi_{l-1}^2) \sigma_a^2. \quad (5.1.16)$$

3) Легко показать, что не только $\hat{z}_t(l)$ — прогноз z_{t+l} с минимальной среднеквадратичной ошибкой, но и любая линейная функция $\sum_{i=1}^L w_i \hat{z}_t(l)$ прогнозов — это прогноз с минимальной среднеквадратичной ошибкой соответствующей линейной функции $\sum_{i=1}^L w_i z_{t+i}$ будущих наблюдений. Например, положим, что мы получили при помощи (5.1.13) по ежемесячным данным прогнозы с минимальной ошибкой $\hat{z}_t(1)$, $\hat{z}_t(2)$ и $\hat{z}_t(3)$ сбыта товара на один, два и три месяца вперед. Тогда справедливо, что $\hat{z}_t(1) + \hat{z}_t(2) + \hat{z}_t(3)$ — это прогноз с минимальной среднеквадратичной ошибкой сбыта товара $z_{t+1} + z_{t+2} + z_{t+3}$ в течение следующего квартала.

4) *Остаточные ошибки как ошибки прогноза на один шаг вперед.* Из (5.1.4) следует, что ошибка прогноза на один шаг вперед есть

$$e_t(1) = z_{t+1} - \hat{z}_t(1) = a_{t+1}. \quad (5.1.17)$$

Следовательно, остаточные ошибки, генерирующие процесс, которые до сих пор вводились как множество независимых случайных величин или импульсов, оказываются *ошибками прогноза на шаг вперед*.

Отсюда следует, что для прогноза с минимальной среднеквадратичной ошибкой ошибки прогноза на шаг вперед должны быть некоррелированы. Это весьма разумно, так как если бы ошибки прогноза на шаг вперед были коррелированы, тогда ошибка прогноза a_{t+1} могла бы до некоторой степени быть предсказана по известным ошибкам прогноза $a_t, a_{t-1}, a_{t-2}, \dots$. Если предсказанная таким путем величина была бы \hat{a}_{t+1} , то $\hat{z}_t(1) + \hat{a}_{t+1}$ было бы лучшим прогнозом z_{t+1} , чем $\hat{z}_t(1)$.

5) *Корреляция между ошибками прогноза.* Хотя оптимальные ошибки прогноза с упреждением 1 (на шаг вперед) некоррелированы, ошибки прогноза с большим упреждением будут, вообще говоря, коррелированы. В разд. П5.1.1 мы выведем общее выражение для корреляции между ошибками прогноза $e_t(l)$ и $e_{t-j}(l)$, сделанными при том же самом упреждении с различных моментов времени t и $t-j$.

Далее справедливо также, что ошибки прогноза $e_t(l)$ и $e_t(l+j)$, сделанные для разных упреждений с одного и того же момента времени t , тоже коррелированы. Одно из следствий этого состоит в том, что часто функции прогноза лежат либо целиком выше, либо целиком ниже фактических значений ряда. В разд. П5.1.2 мы приведем общее выражение для корреляции между ошибками прогноза $e_t(l)$, $e_t(l+j)$, сделанными с того же момента времени.

5.1.2. Три основных представления прогноза

Мы видели, что прогноз с минимальной среднеквадратичной ошибкой $\hat{z}_t(l)$ для упреждения l — это условное математическое ожидание $E_t[z_{t+l}]$ случайной величины z_{t+l} в момент t .

Используя этот факт, мы можем записать выражения для прогноза любым из трех различных способов, соответствующих трем способам представления модели, приведенным в начале этого раздела. Для простоты обозначений временно примем, что квадратные скобки означают взятие условного математического ожидания в момент времени t . Именно

$$[a_{t+l}] = E_t[a_{t+l}], \quad [z_{t+l}] = E_t[z_{t+l}].$$

Для $l > 0$ имеются три различных способа выражения прогнозов.

Прогнозы, полученные из разностного уравнения. Переходя в (5.1.2) к условным математическим ожиданиям в момент t , получаем

$$[z_{t+l}] = \hat{z}_t(l) = \varphi_1 [z_{t+l-1}] + \dots + \varphi_{p+d} [z_{t+l-p-d}] - \theta_1 [a_{t+l-1}] \dots - \theta_q [a_{t+l-q}] + [a_{t+l}]. \quad (5.1.18)$$

Прогноз в проинтегрированном виде. Пользуясь (5.1.3), получаем

$$[z_{t+l}] = \hat{z}_t(l) = \psi_1 [a_{t+l-1}] + \dots + \psi_{l-1} [a_{t+1}] + \psi_l [a_t] + \psi_{l+1} [a_{t-1}] + \dots + [a_{t+l}], \quad (5.1.19)$$

что совпадает по форме с выражением (5.1.13), с которым мы уже встречались. Можно также воспользоваться усеченной формой представления модели (5.1.5) и при положительных $l > q$ получить

$$[z_{t+l}] = \hat{z}_t(l) = C_t(l) + [a_{t+l}] + \psi_1 [a_{t+l-1}] + \dots + \psi_{l-1} [a_{t+1}], \quad (5.1.20)$$

где $C_t(l)$ — общее решение соответствующего однородного разностного уравнения в момент t .

Прогноз как взвешенное среднее предшествующих наблюдений и прогнозов, сделанных в тот же момент с меньшими упреждениями. Наконец, перейдя в (5.1.7) к условным математическим ожиданиям, получим

$$[z_{t+l}] = \hat{z}_t(l) = \sum_{j=1}^{\infty} \pi_j [z_{t+l-j}] + [a_{t+l}]. \quad (5.1.21)$$

Нужно отметить, что прогноз с минимальной среднеквадратичной ошибкой определен через условное математическое ожида-

ние

$$[z_{t+l}] = E_t[z_{t+l}] = E[z_{t+l} | z_t, z_{t-1}, \dots],$$

для которого теоретически требуется знание всех прошлых z до бесконечности. Однако условие обратимости, которое мы наложили на общую модель АРПСС, требует, чтобы веса π (5.1.21) образовывали сходящийся ряд. Отсюда для вычисления прогноза с заданной точностью, начиная с некоторого k , влиянием z_{t-j} , $j > k$, можно пренебречь. На практике веса π обычно затухают довольно быстро, так что, какое бы представление модели ни использовалось в вычислениях, только умеренная часть ряда $z_t, z_{t-1}, \dots, z_{t-k}$ необходима для вычисления прогноза с достаточной точностью.

При вычислении условных математических ожиданий, входящих в выражения (5.1.18)–(5.1.21), мы должны учесть, что, если j — неотрицательное целое, то

$$\begin{aligned} [z_{t-j}] &= E_t[z_{t-j}] = z_{t-j}, \quad j = 0, 1, 2, \dots, \\ [z_{t+j}] &= E_t[z_{t+j}] = \hat{z}_t(j), \quad j = 1, 2, \dots, \\ [a_{t-j}] &= E_t[a_{t-j}] = a_{t-j} = z_{t-j} - \hat{z}_{t-j-1}(1), \quad j = 0, 1, 2, \dots, \\ [a_{t+j}] &= E_t[a_{t+j}] = 0, \quad j = 1, 2, \dots. \end{aligned} \quad (5.1.22)$$

Следовательно, чтобы получить прогноз $\hat{z}_t(l)$, нужно выразить модель для z_{t+l} с помощью любого из трех рассмотренных выше явных представлений и действовать с членами в правой части согласно следующим правилам:

Члены z_{t-j} ($j = 0, 1, 2, \dots$), известные уже к моменту t , оставить без изменения.

z_{t+j} ($j = 1, 2, \dots$), еще не известные, заменить их прогнозами $\hat{z}_t(j)$ на момент t .

Члены a_{t-j} ($j = 0, 1, 2, \dots$), уже известные, определить по $z_{t-j} - \hat{z}_{t-j-1}(1)$.

Члены a_{t+j} ($j = 1, 2, \dots$), еще не известные, заменить нулями. Для массовых расчетов удобнее пользоваться непосредственно представлением модели разностным уравнением (5.1.18).

Пример: прогнозирование с использованием представления модели разностным уравнением. В гл. 7 будет показано, что ряд S хорошо описывается моделью

$$(1 - 0,8B)(1 - B)z_{t+l} = a_{t+l},$$

т. е.

$$(1 - 1,8B + 0,8B^2)z_{t+l} = a_{t+l},$$

или

$$z_{t+l} = 1,8z_{t+l-1} - 0,8z_{t+l-2} + a_{t+l}.$$

Прогнозы в момент t будут иметь вид

$$\begin{aligned}\hat{z}_t(1) &= 1,8z_t - 0,8z_{t-1}, \\ \hat{z}_t(2) &= 1,8\hat{z}_t(1) - 0,8z_t, \\ \hat{z}_t(l) &= 1,8\hat{z}_t(l-1) - 0,8\hat{z}_t(l-2), \quad l=3, 4, 5, \dots\end{aligned}\quad (5.1.23)$$

Видно, что прогнозы легко вычисляются рекуррентным способом, начиная с $\hat{z}_t(1)$, $\hat{z}_t(2)$, ...

Хотя в приведенном выше примере в модели не имелось членов скользящего среднего, такие члены не вносят дополнительных трудностей. Так, далее в этой главе мы рассмотрим ряд, возникающий в проблеме регулирования, модель которого в момент $t+l$ имеет вид

$$\nabla^2 z_{t+l} = (1 - 0,9B + 0,5B^2) a_{t+l}.$$

Тогда

$$\begin{aligned}z_{t+l} &= 2z_{t+l-1} - z_{t+l-2} + a_{t+l} - 0,9a_{t+l-1} + 0,5a_{t+l-2}, \\ \hat{z}_t(1) &= 2z_t - z_{t-1} - 0,9a_t + 0,5a_{t-1}, \\ \hat{z}_t(2) &= 2\hat{z}_t(1) - z_t + 0,5a_t, \\ \hat{z}_t(l) &= 2\hat{z}_t(l-1) - \hat{z}_t(l-2), \quad l=3, 4.\end{aligned}$$

Вспомним, что в этих выражениях $a_t = z_t - \hat{z}_{t-1}(1)$, $a_{t-1} = z_{t-1} - \hat{z}_{t-2}(1)$, и процесс прогнозирования может быть начат подстановкой вместо неизвестных a значений их безусловных математических ожиданий, равных нулю.

В общем, если оператор скользящего среднего $\theta(B)$ имеет q -й порядок, уравнения прогноза для $\hat{z}_t(1)$, $\hat{z}_t(2)$, ..., $\hat{z}_t(q)$ будут зависеть непосредственно от a , но для прогнозов с большим упреждением такой прямой зависимости уже не будет. Не следует, конечно, думать, что влияние этих a не сказывается на прогнозах с большим упреждением. В рассмотренном примере $\hat{z}_t(3)$ зависит от $\hat{z}_t(2)$ и $\hat{z}_t(1)$, а они в свою очередь зависят от a_t и a_{t-1} .

5.2. ВЫЧИСЛЕНИЕ И ПОДПРАВЛЕНИЕ ПРОГНОЗА

5.2.1. Удобная схема для вычисления прогнозов

Во многих случаях необходимо делать прогноз на 1, 2, 3 ... шагов вперед. Такие прогнозы легко получать один из другого при помощи вычислительной схемы, показанной в табл. 5.1, которая иллюстрирует расчет прогнозов для ряда C на момент $t=20$ с упреждениями $l=1, 2, 3, \dots, 14$.

Диагональное расположение в таблице позволяет каждому прогнозу появляться против того значения ряда, которое он предсказывает. Так, $\hat{z}_{20}(6) = 22,51$, и этот прогноз, сделанный в момент $t = 20$, является прогнозом z_{26} и появляется против этой величины. Истинные значения z_t при $t = 21, 22$ и т. д. показаны курсивом для того, чтобы напомнить, что они не были известны в момент, когда делался прогноз.

Вычисления легко осуществляются при помощи формулы (5.1.23). Например,

$$\hat{z}_{20}(1) = (1,8 \times 23,4) - (0,8 \times 23,7) = 23,16,$$

$$\hat{z}_{20}(2) = (1,8 \times 23,16) - (0,8 \times 23,4) = 22,97$$

и т. д. Как только становится доступным новый элемент данных z_{21} , мы можем немедленно вычислить новую совокупность прогнозов, которая заполнит диагональ непосредственно ниже показанной в табл. 5.1. Используя то, что $a_t = z_t - \hat{z}_{t-1}(1)$, при каждом появлении нового z_t можно вычислить и ввести в таблицу a_t как разность ближайшего к нему справа и слева элементов таблицы. Например, как только становится известным $z_{21} = 23,1$, мы можем записать значение $a_{21} = -0,06 = 23,10 - 23,16$.

5.2.2. Вычисление весов ψ

Положим, что необходимы прогнозы с упреждением 1, 2, 3, ..., L . Чтобы найти доверительные пределы для этих прогнозов и уметь вычислять новые прогнозы путем коррекции старых, нужно найти веса $\psi_1, \psi_2, \dots, \psi_{L-1}$. Это осуществляется при помощи (5.1.4), а именно из уравнения

$$\varphi(B)\psi(B) = \theta(B), \quad (5.2.1)$$

т. е. приравнивая коэффициенты при одинаковых степенях B в

$$(1 - \varphi_1 B - \dots - \varphi_{p+d} B^{p+d})(1 + \psi_1 B + \psi_2 B^2 + \dots) = \\ = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q). \quad (5.2.2)$$

Зная значения φ и θ , получаем в результате выражения для ψ :

$$\begin{aligned} \psi_1 &= \varphi_1 - \theta_1, \\ \psi_2 &= \varphi_1 \psi_1 + \varphi_2 - \theta_2, \\ &\vdots \\ \psi_j &= \varphi_1 \psi_{j-1} + \dots + \varphi_{p+d} \psi_{j-p-d} - \theta_j, \end{aligned} \quad (5.2.3)$$

где $\psi_0 = 1$; $\psi_j = 0$ при $j < 0$ и $\theta_j = 0$ при $j > q$. Если K — наибольшее из целых $p + d - 1$ и q , тогда при $j > K$ веса ψ удо-

влетворяют разностному уравнению

$$\psi_j = \varphi_1 \psi_{j-1} + \varphi_2 \psi_{j-2} + \dots + \varphi_{p+d} \psi_{j-p-d}. \quad (5.2.4)$$

Следовательно, ψ легко вычисляются рекуррентным способом. Например, для модели $(1 - 0,8B + 0,8B^2)z_t = a_t$, описывающей ряд C , имеем

$$(1 - 1,8B + 0,8B^2)(1 + \psi_1 B + \psi_2 B^2 + \dots) = 1.$$

Прямым приравниванием коэффициентов при B^j или же используя (5.2.3) и (5.2.4) при $\varphi_1 = 1,8$ и $\varphi_2 = -0,8$, получаем

$$\psi_0 = 1,$$

$$\psi_1 = 1,8,$$

$$\psi_j = 1,8\psi_{j-1} - 0,8\psi_{j-2}, \quad j = 2, 3, 4, \dots$$

Отсюда

$$\psi_2 = (1,8 \times 1,8) - (0,8 \times 1,0) = 2,44,$$

$$\psi_3 = (1,8 \times 2,44) - (0,8 \times 1,8) = 2,95$$

и т. д. ψ для этого примера приведены во второй строке табл. 5.1.

5.2.3. Использование весов ψ при подправлении прогнозов

Интересно рассмотреть еще один способ вычисления прогнозов. Используя (5.1.13), можно выразить прогнозы $\hat{z}_{t+1}(l)$ и $\hat{z}_t(l+1)$ будущего наблюдения z_{t+l+1} , сделанные в моменты $t+1$ и t , как

$$\hat{z}_{t+1}(l) = \psi_l a_{t+1} + \psi_{l+1} a_t + \psi_{l+2} a_{t-1} + \dots,$$

$$\hat{z}_t(l+1) = \psi_{l+1} a_t + \psi_{l+2} a_{t-1} + \dots$$

Вычитая второе уравнение из первого, получим

$$\hat{z}_{t+1}(l) = \hat{z}_t(l+1) + \psi_l a_{t+1}. \quad (5.2.5)$$

Отсюда видно, что если прогноз величины z_{t+l+1} , сделанный в момент t , подправить добавлением ошибки прогноза на шаг вперед a_{t+1} с коэффициентом ψ_l , то получится прогноз той же величины z_{t+l+1} , но в момент $t+1$.

Это ведет к очень интересному выводу. Положим, что мы делаем прогнозы в момент t с упреждениями $1, 2, \dots, L$. Затем, как только становится известным z_{t+1} , мы можем найти $a_{t+1} = z_{t+1} - \hat{z}_t(1)$ и пропорционально подправить прогнозы $\hat{z}_{t+1}(l) = \hat{z}_t(l+1) + \psi_l a_{t+1}$ в момент $t+1$ для упреждений $1, 2, \dots, L-1$. Новый прогноз $\hat{z}_{t+1}(L)$ для упреждения L этим способом найти нельзя, но он легко вычисляется по прогнозам с меньшими упреждениями из разностного уравнения.

Вернемся еще раз к прогнозированию ряда С. В табл. 5.2 к прогнозам, показанным в табл. 5.1, добавлены прогнозы для $t = 21$. Они могут быть получены либо непосредственно (как были получены прогнозы для $t = 20$), либо путем использования корректирующего уравнения (5.2.5). Чтобы сделать этот расчет удобным, значения ψ расположены во второй строке.

Конкретно, как только мы узнаем $z_{21} = 23,1$, можем вычислить $a_{21} = 23,1 - 23,16 = -0,06$. Тогда

$$\hat{z}_{21}(1) = 22,86 = 22,97 + (1,8)(-0,06),$$

$$\hat{z}_{21}(2) = 22,67 = 22,81 + (2,44)(-0,06),$$

$$\hat{z}_{21}(3) = 22,51 = 22,69 + (2,95)(-0,06)$$

и т. д.

5.2.4. Вычисление вероятностных пределов прогнозов при произвольном упреждении

Из выражения (5.1.16) следует, что в общем дисперсия ошибки прогноза на l шагов вперед для любого момента t , определяемая как математическое ожидание величины

$$e_t^2(l) = \{z_{t+l} - \hat{z}_t(l)\}^2,$$

равна

$$V(l) = \left\{ 1 + \sum_{j=1}^{l-1} \psi_j^2 \right\} \sigma_a^2.$$

Например, по весам из табл. 5.1 были рассчитаны значения функции $V(l)/\sigma_a^2$ для ряда С, приведенные в табл. 5.3.

Таблица 5.3

Дисперсии ошибок прогноза для ряда С

l	1	2	3	4	5	6	7	8	9	10
$V(l)/\sigma_a^2$	1,00	4,24	10,19	18,96	30,24	43,86	59,46	76,79	95,52	115,41

Из предположения, что a подчиняются нормальному закону, следует, что при известных значениях процесса до момента t условное распределение вероятности $p(z_{t+l}|z_t, z_{t-1}, \dots)$ будущего значения процесса z_{t+l} будет также нормальным со средним значением $\hat{z}_t(l)$ и стандартным отклонением $\left\{ 1 + \sum_{j=1}^{l-1} \psi_j^2 \right\}^{1/2} \sigma_a$.

На рис. 5.1 показаны условные распределения вероятности будущих значений z_{21} , z_{22} , z_{23} для ряда С при наличии информации о ряде до момента $t = 20$.

Мы покажем в гл. 7, как можно получить оценку s_a^2 дисперсии σ_a^2 из наблюдаемых временных рядов. Когда число наблюдений, на которых основывается такая оценка, будет, скажем, не меньше 50, s_a можно заменить на σ_a , и приближенные $(1 - \epsilon)\%$ -ные вероятностные пределы $z_{t+l}(-)$ и $z_{t+l}(+)$ для z_{t+l} будут иметь вид:

$$z_{t+l}(\pm) = \hat{z}_t(l) \pm u_{\epsilon/2} \left\{ 1 + \sum_{j=1}^{l-1} \psi_j^2 \right\}^{1/2} s_a, \quad (5.2.6)$$

где $u_{\epsilon/2}$ — это квантиль уровня $1 - \epsilon/2$ стандартного нормального распределения.

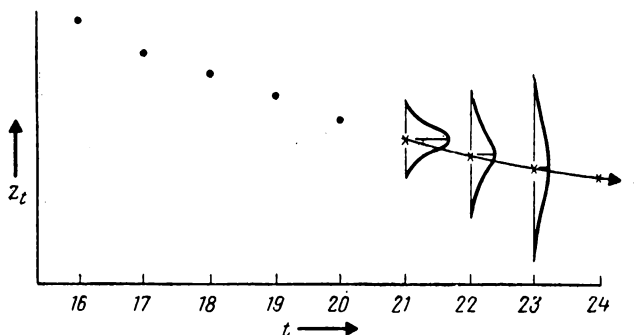


Рис. 5.1. Распределение условных вероятностей будущих значений z_{t+1} , z_{t+2} и z_{t+3} для ряда С при известной информации до момента $t = 20$.

В табл. 7.13 показано, что $s_a = 0,134$ для ряда С; следовательно, 50%-ные и 95%-ные пределы, например для $\hat{z}_t(2)$, равны:

50%-ные пределы: $\hat{z}_t(2) \pm (0,674)(1 + 1,8^2)^{1/2}(0,134) = \hat{z}_t(2) \pm 0,19$,
 95%-ные пределы: $\hat{z}_t(2) \pm 1,96(1 + 1,8^2)^{1/2}(0,134) = \hat{z}_t(2) \pm 0,55$.
 Величины, которые нужно прибавить и вычесть из прогноза, чтобы получить 50- и 95%-ные пределы, показаны в 4-й и 5-й строках головок табл. 5.1 и 5.2. Они относятся к значениям прогнозов, расположенным непосредственно ниже их.

На рис. 5.2 показан участок ряда С вместе с прогнозами на несколько шагов вперед (крестики) для моментов $t = 20$ и $t = 67$. Там же показаны 50- и 95%-ные вероятностные пределы для z_{20+l} при l , меняющемся от 1 до 14. Пределы $z_{t+l}(-)$ и $z_{t+l}(+)$ интерпретируются следующим образом. Если известна информация о временном ряде к моменту t , то с вероятностью $1 - \epsilon$ наблюдаемое значение z_{t+l} будет заключено в этих пределах, т. е.

$$\Pr \{z_{t+l}(-) < z_{t+l} < z_{t+l}(+)\} = 1 - \epsilon.$$

Необходимо также пояснить, что эти вероятности относятся только к *индивидуальным* прогнозам, а не к совокупности прогнозов с различными упреждениями. Например, справедливо, что с 95%-ной вероятностью значение z_{t+10} будет заключено

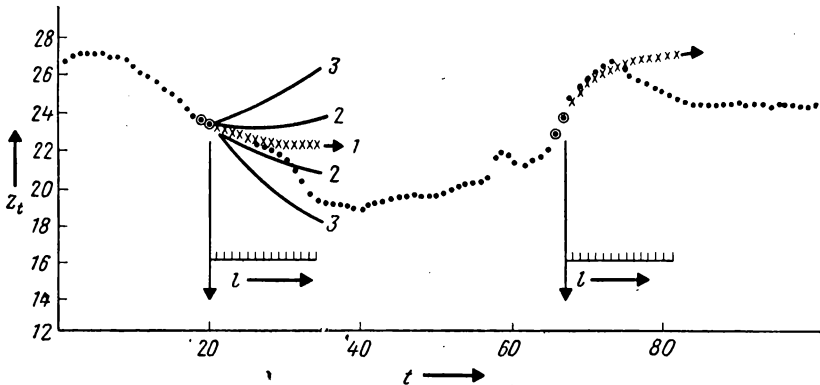


Рис. 5.2. Прогнозы и вероятностные пределы для ряда С.
1 — прогнозирующая функция, 2 — 50%-ные пределы, 3 — 95%-ные пределы.

в пределах для упреждения 10. Но нельзя ожидать, что ряд окажется одновременно внутри *всех* пределов с этой же вероятностью.

5.3. ПРОГНОЗИРУЮЩАЯ ФУНКЦИЯ И ВЕСА ПРОГНОЗА

Наиболее просто вычислять прогнозы только что описанным способом — прямым использованием разностного уравнения. С чисто *вычислительной* точки зрения другие представления модели менее удобны. Однако для изучения природы прогнозов полезно рассмотреть более подробно другие представления модели, приведенные в разд. 5.1.2, и в частности рассмотреть явное выражение для прогнозирующей функции.

5.3.1. Эвентуальная прогнозирующая функция, определенная оператором авторегрессии

В момент $t + l$ модель АРПСС может быть представлена в виде

$$\begin{aligned} z_{t+l} - \Phi_1 z_{t+l-1} - \dots - \Phi_{p+d} z_{t+l-p-d} = \\ = a_{t+l} - \theta_1 a_{t+l-1} - \dots - \theta_q a_{t+l-q}. \end{aligned} \quad (5.3.1)$$

Беря от обеих частей (5.3.1) условные математические ожидания при фиксированном до момента t прошлом, получим для $l > q$

$$\hat{z}_t(l) - \varphi_1 \hat{z}_t(l-1) - \dots - \varphi_{p+d} \hat{z}_t(l-p-d) = 0, \quad l > q, \quad (5.3.2)$$

где подразумевается, что $\hat{z}_t(-j) = z_{t-j}$ при $j \geq 0$. Разностное уравнение (5.3.2) имеет решение

$$\hat{z}_t(l) = b_0^{(t)} f_0(l) + b_1^{(t)} f_1(l) + \dots + b_{p+d-1}^{(t)} f_{p+d-1}(l) \quad (5.3.3)$$

для $l > q - p - d$. Отметим, что $\hat{z}_t(l)$ имеет ту же форму, что и функция C , введенная в гл. 4. В (5.3.3) $f_0(l), f_1(l), \dots, f_{p+d-1}(l)$ есть функции упреждения на время l . В общем случае они могут включать полиномы, экспоненты, синусы и косинусы и произведения этих функций. Для заданного момента t коэффициенты $b_j^{(t)}$ постоянны для всех упреждений l , но они изменяются от данного момента к следующему, подстраиваясь к соответствующей части рассматриваемого ряда. Всюду далее мы будем называть функцию, определенную (5.3.3), *эвентуальной функцией прогноза*; «эвентуальной» потому, что в случае, когда $q > p + d$, она дает прогнозы только для упреждений $l > q - p - d$.

Из (5.3.2) мы видим, что математическая форма функции прогноза определяется общим оператором авторегрессии $\varphi(B)$, т. е. свойствами f в (5.3.3). Конкретнее, он определяет, является ли функция прогноза полиномом, набором экспонент, набором синусов и косинусов или некоторой комбинацией этих функций.

5.3.2. Роль оператора скользящего среднего в определении начальных величин

В то время как оператор авторегрессии определяет природу эвентуальной прогнозирующей функции, оператор скользящего среднего отвечает за то, как эта функция должна быть «подогнана» к данным, и, следовательно, как надо вычислять и подправлять коэффициенты $b_0^{(t)}, b_1^{(t)}, \dots, b_{p+d-1}^{(t)}$. Например, рассмотрим процесс ПСС(0, 2, 3)

$$z_{t+l} - 2z_{t+l-1} + z_{t+l-2} = a_{t+l} - \theta_1 a_{t+l-1} - \theta_2 a_{t+l-2} - \theta_3 a_{t+l-3}.$$

Используя трактовку условного математического ожидания, данную в разд. 5.1.2, определим прогнозирующую функцию:

$$\hat{z}_t(1) = 2z_t - z_{t-1} - \theta_1 a_t - \theta_2 a_{t-1} - \theta_3 a_{t-2},$$

$$\hat{z}_t(2) = 2\hat{z}_t(1) - z_t - \theta_2 a_t - \theta_3 a_{t-1},$$

$$\hat{z}_t(3) = 2\hat{z}_t(2) - \hat{z}_t(1) - \theta_3 a_t,$$

$$\hat{z}_t(l) = 2\hat{z}_t(l-1) - \hat{z}_t(l-2), \quad l > 3.$$

Следовательно, эвентуальная прогнозирующая функция будет состоять из одной прямой линии

$$\hat{z}_t(l) = b_0^{(t)} + b_1^{(t)}l, \quad l > 1,$$

которая проходит через $\hat{z}_t(2)$ и $\hat{z}_t(3)$, как показано на рис. 5.3. Однако заметим, что если член θ_3 опущен, тогда $q - p - d = 0$

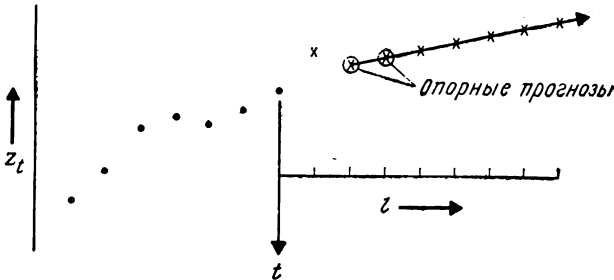


Рис. 5.3. Эвентуальная прогнозирующая функция процесса ПСС(0,2,3).

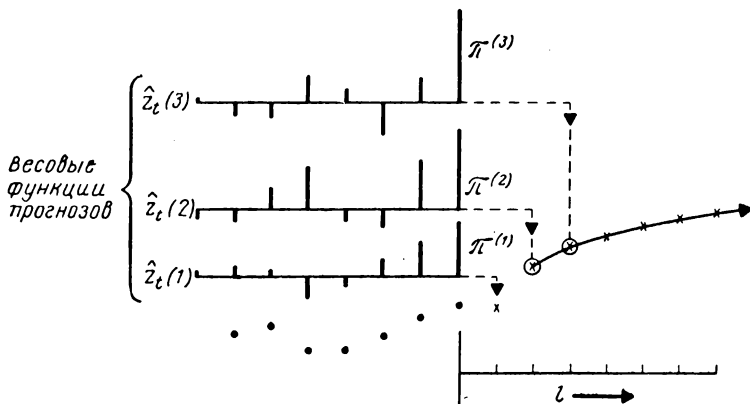


Рис. 5.4. Зависимость прогнозирующей функции от наблюдений для процесса (1,1,3) $(1 - \phi B) \nabla z_t = (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3) a_t$. Прогнозирующая функция проходит через $p + d = 2$ опорных значения $\hat{z}_t(q)$ и $\hat{z}_t(q-1)$, $q=3$.

и прогноз для *всех упреждений* будет задаваться прямой линией, проходящей через $\hat{z}_t(1)$ и $\hat{z}_t(2)$.

В общем поскольку только одна функция вида (5.3.3) может проходить через $p + d$ точки, то эвентуальная функция прогноза — это единственная кривая вида $\phi(B)$, которая проходит

через $p + d$ «опорные» величины $\hat{z}_t(q)$, $\hat{z}_t(q-1)$, ..., $\hat{z}_t(q-p-d+1)$, где $\hat{z}_t(-j) = z_{t-j}$, $j = 0, 1, 2$. В предельном случае, когда $q = 0$, т. е. модель есть чистая авторегрессия вида $\varphi(B)z_t = a_t$, кривая проходит через точки z_t , z_{t-1} , ..., $z_{t-p-d+1}$. Таким образом, опорные величины могут состоять из прогнозов или действительных значений ряда; они показаны на рисунках кружками.

Члены скользящего среднего, появляющиеся в представлении модели, помогают выбрать путь использования прошлого ряда для подгонки функции прогноза, определенной оператором авторегрессии $\varphi(B)$. Рис. 5.4 иллюстрирует ситуацию для модели порядка $(1, 1, 3)$, заданной как $(1 - \phi B)\nabla z_t = (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3)a_t$. Гипотетические весовые функции указывают линейную функциональную зависимость трех прогнозов $\hat{z}_t(1)$, $\hat{z}_t(2)$ и $\hat{z}_t(3)$ от $z_t, z_{t-1}, z_{t-2}, \dots$. Поскольку прогнозирующая функция содержит $p + d = 2$ коэффициентов, она однозначно определяется прогнозами $\hat{z}_t(3)$ и $\hat{z}_t(2)$, т. е. $\hat{z}_t(q)$ и $\hat{z}_t(q-1)$. Мы рассмотрим далее, как определяются упомянутые выше весовые функции прогноза.

5.3.3. Веса прогноза для упреждения l

Тот факт, что общая модель прогноза может быть также записана в обращенной форме

$$a_t = \pi(B)z_t = (1 - \pi_1 B - \pi_2 B^2 - \pi_3 B^3 - \dots)z_t, \quad (5.3.4)$$

позволяет нам записать прогноз в форме (5.1.21).

Подставляя в (5.1.21) условные математические ожидания, получаем

$$\hat{z}_t(l) = \sum_{j=1}^{\infty} \pi_j \hat{z}_t(l-j), \quad (5.3.5)$$

где, как и раньше, $\hat{z}_t(-h) = z_{t-h}$ для $h = 0, 1, 2, \dots$. Отсюда вообще

$$\hat{z}_t(l) = \pi_1 \hat{z}_t(l-1) + \dots + \pi_{l-1} \hat{z}_t(1) + \pi_l z_t + \pi_{l+1} z_{t-1} + \dots \quad (5.3.6)$$

и, в частности,

$$\hat{z}_t(1) = \pi_1 z_t + \pi_2 z_{t-1} + \pi_3 z_{t-2} + \dots$$

Прогнозы для больших упреждений могут быть также представлены непосредственно в виде линейных функций наблюдений z_t ,

z_{t-1}, z_{t-2}, \dots . Например, прогноз с упреждением 2 в момент t равен

$$\begin{aligned}\hat{z}_t(2) &= \pi_1 \hat{z}_t(1) + \pi_2 z_t + \pi_3 z_{t-1} + \dots \\ &= \pi_1 \sum_{j=1}^{\infty} \pi_j z_{t-j+1} + \sum_{j=1}^{\infty} \pi_{j+1} z_{t-j+1} = \sum_{j=1}^{\infty} \pi_j^{(2)} z_{t-j+1},\end{aligned}$$

где

$$\pi_j^{(2)} = \pi_1 \pi_j + \pi_{j+1}, \quad j = 1, 2, \dots \quad (5.3.7)$$

Продолжая таким же образом, нетрудно показать, что

$$\hat{z}_t(l) = \sum_{j=1}^{\infty} \pi_j^{(l)} z_{t-j+1}, \quad (5.3.8)$$

где

$$\pi_j^{(l)} = \pi_{j+l-1} + \sum_{h=1}^{l-1} \pi_h \pi_j^{(l-h)}, \quad j = 1, 2, \dots,$$

и $\pi_j^{(1)} = \pi_j$. Другие способы вычисления тех же весов приведены в приложении П5.2.

Как следует из (4.2.22) и (5.1.9), сами веса π_j могут быть получены в явном виде приравниванием коэффициентов в

Таблица 5.4

$$\begin{aligned}\theta(B)(1 - \pi_1 B - \pi_2 B^2 - \dots) &= \\ &= \varphi(B).\end{aligned}$$

Веса π для модели $\nabla^2 z_t = (1 - 0,9B + 0,5B^2) a_t$

l	$\pi_j = \pi_j^{(1)}$	$\pi_j^{(2)}$
1	1,100	1,700
2	0,490	0,430
3	-0,109	-0,463
4	-0,343	-0,632
5	-0,254	-0,336
6	-0,057	0,013
7	0,076	0,181
8	0,097	0,156
9	0,049	0,050
10	-0,004	-0,032
11	-0,028	-0,054
12	-0,023	-0,026

Когда эти величины известны, $\pi_j^{(l)}$ можно легко определить при помощи (5.3.9) или способами, указанными в приложении П5.2.

В качестве примера рассмотрим опять модель

$$\nabla^2 z_t = (1 - 0,9B + 0,5B^2) a_t,$$

подогнанную к ряду, часть которого показана на рис. 5.5. Приравнивая коэффициенты в

$$\begin{aligned}(1 - 0,9B + 0,5B^2) \times \\ \times (1 - \pi_1 B - \pi_2 B^2 \dots) &= \\ &= 1 - 2B + B^2,\end{aligned}$$

получаем веса $\pi_j = \pi_j^{(1)}$, по которым с помощью (5.3.7) можно вычислить веса $\pi_j^{(2)}$. Эти две последовательности весов для $j = 1, 2, \dots, 12$ приведены в табл. 5.4.

В этом примере прогнозы на один и два шага вперед, выраженные через наблюдаемые данные z_t, z_{t-1}, \dots , равны

$$\hat{z}_t(1) = 1,10z_t + 0,49z_{t-1} - 0,11z_{t-2} - 0,34z_{t-3} - 0,25z_{t-4} - \dots$$

и

$$\hat{z}_t(2) = 1,70z_t + 0,43z_{t-1} - 0,46z_{t-2} - 0,63z_{t-3} - 0,34z_{t-4} + \dots$$

Как показано на рис. 5.5, веса ведут себя как значения затухающей синусоиды.

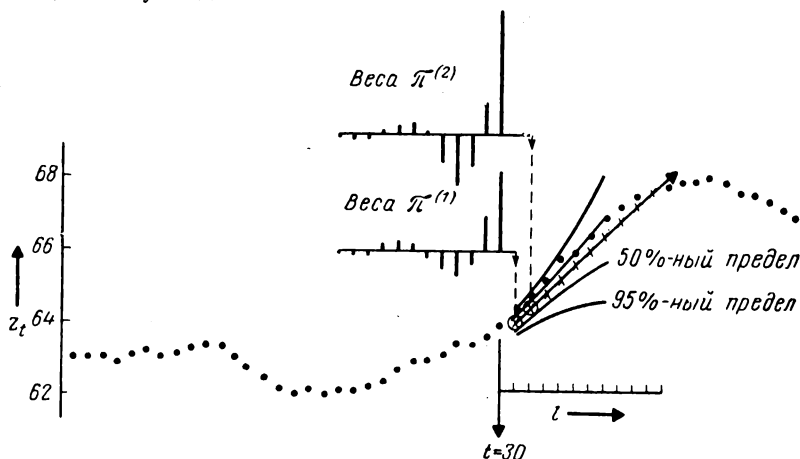


Рис. 5.5. Часть ряда, к которому подогнана модель $\nabla^2 z_t = (1 - 0,9B + 0,5B^2) a_t$, с прогнозирующей функцией для момента $t = 30$, весами прогноза и вероятностными пределами.

5.4. ПРИМЕРЫ ПРОГНОЗИРУЮЩИХ ФУНКЦИЙ И ИХ ПОДПРАВЛЕНИЯ

Рассмотрим теперь прогнозирующие функции для некоторых специальных случаев общей модели АРПСС. Мы представим их тремя различными способами, обсуждавшимися в разд. 5.1.2. Как уже отмечалось, легче всего вычислять прогнозы при помощи разностного уравнения. Другие формы представления полезны, так как помогают лучше понять в конкретных случаях природу прогнозирующей функции.

5.4.1. Прогнозирование процесса ПСС(0, 1, 1)

Модель этого процесса $\nabla z_t = (1 - \theta B) a_t$.

Способ разностного уравнения. Для момента $t + l$ модель можно представить в виде

$$z_{t+l} = z_{t+l-1} + a_{t+l} - \theta a_{t+l-1}.$$

Беря условные математические ожидания при известном до момента t прошлом, получаем

$$\begin{aligned}\hat{z}_t(1) &= z_t - \theta a_t, \\ \hat{z}_t(l) &= \hat{z}_t(l-1), \quad l \geq 2.\end{aligned}\quad (5.4.1)$$

Следовательно, для всех упреждений прогнозы в момент t будут следовать прямой линии, параллельной временной оси. Пользуясь тем, что $z_t = \hat{z}_{t-1}(1) + a_t$, можно записать (5.4.1) одним из двух удобных способов.

Первый из них

$$\hat{z}_t(l) = \hat{z}_{t-1}(l) + \lambda a_t, \quad (5.4.2)$$

где $\lambda = 1 - \theta$. Он говорит о том, что, если наш предыдущий прогноз $\hat{z}_{t-1}(l)$ отличается от фактического значения на a_t , мы должны подправить его на λa_t . Напомним, что, согласно разд. 4.1.3, λ указывает долю данного импульса a_t , входящую в «уровень» процесса. Поэтому разумно увеличить прогноз на часть λa_t от a_t , которая, как мы ожидаем, войдет в «уровень».

Второй способ записи (5.4.1) имеет вид

$$\hat{z}_t(l) = \lambda z_t + (1 - \lambda) \hat{z}_{t-1}(l). \quad (5.4.3)$$

Видно, что новый прогноз — это линейная интерполяция по аргументу λ между старым прогнозом и новым наблюдением. Выражение (5.4.3) ясно показывает, что если λ очень мало, мы в основном полагаемся на взвешенное среднее данных о прошлом и почти пренебрегаем новым наблюдением z_t . Напротив, если $\lambda = 1$, данные прошлого полностью игнорируются, $\hat{z}_t(1) = z_t$, и прогноз на все будущие времена равен текущему значению. При $\lambda > 1$ мы скорее вводим экстраполяцию, чем интерполяцию, между $\hat{z}_{t-1}(l)$ и z_t . Ошибка прогноза в (5.4.2) должна *увеличиться*, что показывает изменение в прогнозе.

Прогнозирующая функция в проинтегрированном виде. Эвентуальная прогнозирующая функция — это решение уравнения $(1 - B)\hat{z}_t(l) = 0$. Отсюда $\hat{z}_t(l) = b_0^{(t)}$, и, так как $q - p - d = 0$, этим обеспечивается прогноз для всех упреждений, т. е.

$$\hat{z}_t(l) = b_0^{(t)}, \quad l > 0. \quad (5.4.4)$$

Для *любого фиксированного момента* $b_0^{(t)}$ — константа, и прогнозы для всех упреждений будут следовать прямой линии, параллельной временной оси. Однако коэффициенты $b_0^{(t)}$ будут скорректированы, когда станут доступными новые наблюдения, и момент t , в который делается прогноз, сдвинется вперед. Поэтому прогнозирующую функцию можно трактовать как полином нулевой степени от упреждения l с коэффициентом, скорректированным в соответствии с положением исходной точки t .

Так как проинтегрированное представление модели имеет вид

$$z_t = \lambda S a_{t-1} + a_t,$$

то

$$\hat{z}_t(l) = b_0^{(l)} = \lambda S a_t.$$

Отметим также, что $\psi_j = \lambda$ ($j = 1, 2, \dots$), отсюда регулируемый коэффициент $b_0^{(l)}$ может быть скорректирован при переходе от момента t к $t + 1$, согласно формуле

$$b_0^{(t+1)} = b_0^{(t)} + \lambda a_{t+1}. \quad (5.4.5)$$

Прогнозы как взвешенное среднее предыдущих наблюдений. Поскольку для этого процесса веса $\pi_j^{(l)}$ в (5.3.8) являются также весами прогноза на один шаг вперед, мы можем записать

$$\hat{z}_t(l) = b_0^{(l)} = \lambda z_t + \lambda(1 - \lambda)z_{t-1} + \lambda(1 - \lambda)^2 z_{t-2} + \dots \quad (5.4.6)$$

Отсюда для модели ПСС(0, 1, 1) прогноз для всех времен в будущем — это экспоненциально взвешенное скользящее среднее настоящего и предшествующих значений z .

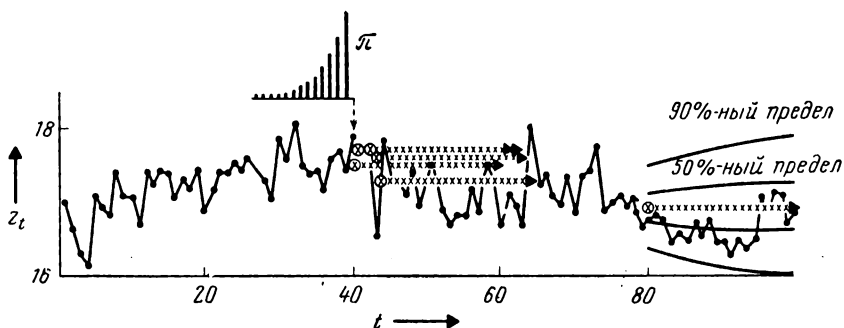


Рис. 5.6. Часть ряда А с прогнозами на моменты $t = 39, 40, 41, 42, 43$ и 79 .

Пример. Прогнозирование ряда А. В гл. 7 будет показано, что ряд А хорошо описывается моделью

$$(1 - B)z_t = (1 - 0,7B)a_t.$$

На рис. 5.6 показаны прогнозы в точках $t = 39, 40, 41, 42$ и 43 , а также в точке $t = 79$ для упреждений $1, 2, \dots, 20$. Веса π_j , которые для этой модели являются весами прогноза для любого упреждения, даны в табл. 5.5.

Эти веса для прогноза $\hat{z}_{39}(l)$ показаны на диаграмме в верхней части рис. 5.6.

Таблица 5.5

Веса прогноза, на которые умножаются предшествующие значения z для любого упреждения
(веса использованы при прогнозировании ряда А с помощью модели $\nabla z_t = (1 - 0,7B) a_t$)

l	π_l	l	π_l
1	0,300	7	0,035
2	0,210	8	0,025
3	0,147	9	0,017
4	0,103	10	0,012
5	0,072	11	0,008
6	0,050	12	0,006

Функции дисперсии. Поскольку для этой модели $\psi_j = \lambda$ ($j = 1, 2 \dots$), выражение (5.1.16) для дисперсии прогноза с упреждением l будет

$$V(l) = \sigma_a^2 \{1 + (l-1)\lambda^2\}. \quad (5.4.7)$$

Значение оценки $s_a^2 = 0,101$, соответствующей ряду А, было подставлено в формулу (5.4.7); с ее помощью были рассчитаны 50- и 95%-ные вероятностные пределы, показанные на рис. 5.6 (для момента $t = 79$).

5.4.2. Прогнозирование процесса ПСС(0, 2, 2)

Модель имеет вид процесса $\nabla^2 z_t = (1 - \theta_1 B - \theta_2 B^2) a_t$.

Способ разностного уравнения. В момент $t+l$ модель представлена в виде

$$z_{t+l} = 2z_{t+l-1} - z_{t+l-2} + a_{t+l} - \theta_1 a_{t+l-1} - \theta_2 a_{t+l-2}.$$

Беря условные математические ожидания при известном до момента t прошлом, получаем

$$\hat{z}_t(1) = 2z_t - z_{t-1} - \theta_1 a_t - \theta_2 a_{t-1},$$

$$\hat{z}_t(2) = 2\hat{z}_t(1) - z_t - \theta_2 a_t,$$

$$\hat{z}_t(l) = 2\hat{z}_t(l-1) - \hat{z}_t(l-2), \quad l \geq 3,$$

откуда можно вычислять прогнозы. Прогнозирование при помощи этого способа ряда на рис. 5.5 было продемонстрировано в разд. 5.1.2. Другой возможный способ генерирования первых $L-1$ прогнозов — подправление формулы (5.2.5)

$$\hat{z}_{t+1}(l) = \hat{z}_t(l+1) + \psi_l a_{t+1}. \quad (5.4.8)$$

Проинтегрированное представление модели имеет вид

$$z_t = \lambda_0 S a_{t-1} + \lambda_1 S^2 a_{t-1} + a_t, \quad \lambda_0 = 1 + \theta_2, \quad \lambda_1 = 1 - \theta_1 - \theta_2, \quad (5.4.9)$$

так что $\psi_j = \lambda_0 + j\lambda_1$ ($j = 1, 2, \dots$). Следовательно, скорректированная функция для этой модели равна

$$\hat{z}_{t+1}(l) = \hat{z}_t(l+1) + (\lambda_0 + l\lambda_1) a_{t+1}. \quad (5.4.10)$$

Прогноз при помощи проинтегрированного представления. Эвентуальная прогнозирующая функция — это решение уравнения $(1-B)^2 \hat{z}_t(l) = 0$, т. е. $\hat{z}_t(l) = b_0^{(t)} + b_1^{(t)}l$. Поскольку $q - p - d = 0$, эвентуальная прогнозирующая функция дает прогноз для любого упреждения, т. е.

$$\hat{z}_t(l) = b_0^{(t)} + b_1^{(t)}l, \quad l > 0. \quad (5.4.11)$$

Следовательно, прогнозирующая функция — это линейная функция упреждения l с коэффициентами, подстраиваемыми в соответствии с положением исходной точки t . Стохастическая модель в проинтегрированном представлении имеет вид

$$z_{t+l} = a_{t+l} + \lambda_0 S a_{t+l-1} + \lambda_1 S^2 a_{t+l-1},$$

и, переходя к условным математическим ожиданиям при известном к моменту t прошлому, получаем

$$\hat{z}_t(l) = \{\lambda_0 S a_t + \lambda_1 S^2 a_{t-1}\} + \{\lambda_1 S a_t\}l.$$

Подстраиваемые константы можно тогда определить как

$$\begin{aligned} b_0^{(t)} &= \lambda_0 S a_t + \lambda_1 S^2 a_{t-1}, \\ b_1^{(t)} &= \lambda_1 S a_t, \end{aligned} \quad (5.4.12)$$

а формулы, указывающие способ подстройки, будут

$$\begin{aligned} b_0^{(t)} &= b_0^{(t-1)} + b_1^{(t-1)} + \lambda_0 a_t, \\ b_1^{(t)} &= b_1^{(t-1)} + \lambda_1 a_t. \end{aligned} \quad (5.4.13)$$

Дополнительный член $b_1^{(t-1)}$, входящий в формулу для наклона $b_0^{(t)}$, представляет поправку, которую необходимо ввести в параметр положения b_0 , чтобы приспособить его к новой исходной точке.

Нужно также отметить, что λ_0 и λ_1 — доли импульса a_t , которые перейдут в параметры, описывающие положение и наклон соответственно.

Прогноз как взвешенное среднее предыдущих наблюдений. Для этой модели прогнозирующая функция — это прямая линия, проходящая через значения прогнозов $\hat{z}_t(1)$ и $\hat{z}_t(2)$. Иллюстрацией служит рис. 5.5, где показаны прогнозы, сделанные

в момент $t = 30$, и соответствующие весовые функции. Мы увидим, что зависимость всей прогнозирующей функции от предыдущих z отражает зависимость $\hat{z}_t(1)$ и $\hat{z}_t(2)$ от этих значений. Весовые функции для $\hat{z}_t(1)$ и $\hat{z}_t(2)$, показанные на этом рисунке, приведены в табл. 5.4.

Этот пример еще раз демонстрирует, что в то время как АР-оператор $\varphi(B)$ определяет вид используемой функции (в этом случае прямая линия), оператор СС существует при определении способа, которым эта функция «подгоняется» к данным о прошлом.

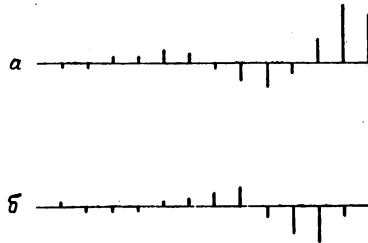


Рис. 5.7. Веса, приданные предыдущим z , определяющие положение и наклон ряда для модели $\nabla^2 z_t = (1 - 0,9 + 0,5B^2)a_t$.

a — веса для параметра положения $b_0^{(t)}$, b — веса для параметра наклона $b_1^{(t)}$.

Зависимость подстраивающихся коэффициентов прогнозирующей функции от предшествующих z . Поскольку для общей модели значения подстраивающихся коэффициентов прогнозирующей функции определяются величинами $\hat{z}_t(q)$, $\hat{z}_t(q-1)$, ..., $\hat{z}_t(q-p-d+1)$, которые в свою очередь могут быть выражены через наблюдаемые величины, ясно, что через наблюдения можно выразить и сами подстраивающиеся коэффициенты.

Например, в случае модели $\nabla^2 z_t = (1 - 0,9B + 0,5B^2)a_t$, изображенной на рис. 5.6, мы имеем

$$\hat{z}_t(1) = b_0^{(t)} + b_1^{(t)} = \sum_{j=1}^{\infty} \pi_j^{(1)} z_{t-j+1},$$

$$\hat{z}_t(2) = b_0^{(t)} + 2b_1^{(t)} = \sum_{j=1}^{\infty} \pi_j^{(2)} z_{t-j+1},$$

так что

$$b_0^{(t)} = 2\hat{z}_t(1) - \hat{z}_t(2) = \sum_{j=1}^{\infty} \{2\pi_j^{(1)} - \pi_j^{(2)}\} z_{t-j+1}$$

и

$$b_1^{(t)} = \hat{z}_t(2) - \hat{z}_t(1) = \sum_{j=1}^{\infty} \{\pi_j^{(2)} - \pi_j^{(1)}\} z_{t-j+1}.$$

Эти весовые функции приведены на рис. 5.7.

Дисперсия ошибки прогноза. Пользуясь (5.1.16) и тем, что $\psi_j = \lambda_0 + j\lambda_1$, получаем формулу для дисперсии прогноза с упреждением l :

$$V(l) = \sigma_a^2 \left\{ 1 + (l-1)\lambda_0^2 + \frac{1}{6}l(l-1)(2l-1)\lambda_1^2 + \lambda_0\lambda_1l(l-1) \right\}. \quad (5.4.14)$$

На рис. 5.5 показаны 50- и 95%-ные вероятностные пределы для прогноза при $t = 30$ (использовались значения $\lambda_0 = 0,5$ и $\lambda_1 = 0,6$ и значение выборочной оценки $s_a^2 = 0,032$).

5.4.3 Прогнозирование общего процесса ПСС(0, d, q)

В качестве примера рассмотрим процесс порядка (0, 1, 3)

$$(1 - B)z_{t+l} = (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3)a_{t+l}.$$

Беря условные математические ожидания при известном до момента t прошлом, мы получаем

$$\hat{z}_t(1) - z_t = -\theta_1 a_t - \theta_2 a_{t-1} - \theta_3 a_{t-2},$$

$$\hat{z}_t(2) - \hat{z}_t(1) = -\theta_2 a_t - \theta_3 a_{t-1},$$

$$\hat{z}_t(3) - \hat{z}_t(2) = -\theta_3 a_t,$$

$$\hat{z}_t(l) - \hat{z}_t(l-1) = 0, \quad l = 4, 5, 6, \dots$$

Отсюда $\hat{z}_t(l) = \hat{z}_t(3) = b_0^{(t)}$ для всех $l > 2$, как и следовало ожидать, так как $q - p - d = 2$. Как показано на рис. 5.8, про-

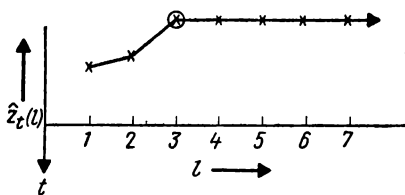


Рис. 5.8. Прогнозирующая функция для процесса ПСС(0,1,3).

гнозирующая функция делает два начальных «скачка», зависящих от предыдущих a , прежде чем выравняться до eventualной прогнозирующей функции.

Для процесса ПСС(0, d, q) eventualная прогнозирующая функция удовлетворяет уравнению $(1 - B)^d \hat{z}_t(l) = 0$; его решение — полином от l степени $d - 1$:

$$\hat{z}_t(l) = b_0^{(t)} + b_1^{(t)}l + b_2^{(t)}l^2 + \dots + b_{d-1}^{(t)}l^{d-1}.$$

Это выражение дает прогнозы $\hat{z}_t(l)$ для $l > q - d$. Коэффициенты $b_0^{(t)}, b_1^{(t)}, \dots, b_{d-1}^{(t)}$ должны последовательно корректироваться по

мере продвижения момента прогноза. Прогноз для момента t сделает $q-d$ начальных «скачков», зависящих от $a_t, a_{t-1}, \dots, a_{t-q+1}$, и затем перейдет в приведенный выше полином.

5.4.4. Прогнозирование процессов авторегрессии

Рассмотрим процесс порядка $(p, d, 0)$

$$\varphi(B) z_t = a_t.$$

Эвентуальная прогнозирующая функция в этом случае есть решение уравнения $\varphi(B) \hat{z}_t(l) = 0$. Она применима ко всем упреждениям и проходит через последние $p+d$ известных значений ряда. Например, модель ряда цен акций IBM (ряда B) очень близка к

$$(1 - B) z_t = a_t,$$

так что

$$\hat{z}_t(l) \approx z_t.$$

Лучший прогноз на будущее оказывается очень близким к текущему значению цены. Весовая функция для $\hat{z}_t(l)$ сосредоточена в точке t , и не происходит никакого осреднения по прошлому.

Стационарные модели авторегрессии. Процесс $\phi(B) \tilde{z}_t = a_t$ порядка $(p, 0, 0)$, где $\phi(B)$ — стационарный оператор и $\tilde{z}_t = z_t - \mu$ с $E[z_t] = \mu$, будет в общем случае иметь прогнозирующую функцию, являющуюся совокупностью экспонент и затухающих синусоид.

В частности, при $p = 1$ модель порядка $(1, 0, 0)$

$$(1 - \phi B) \tilde{z}_t = a_t, \quad -1 < \phi < 1$$

имеет прогнозирующую функцию, которая при любых $l > 0$ является решением уравнения $(1 - \phi B) \hat{z}_t(l) = 0$. Отсюда

$$\hat{z}_t(l) = b_0^{(l)} \phi^l, \quad l > 0. \quad (5.4.15)$$

Кроме того, $\hat{z}_t(1) = \phi \tilde{z}_t$, так что $b_0^{(1)} = \tilde{z}_t$

и

$$\hat{z}_t(l) = \tilde{z}_t \phi^l.$$

Следовательно, прогноз с минимальной среднеквадратичной ошибкой предсказывает, что текущее отклонение от среднего должно экспоненциально спадать до нуля. На рис. 5.9, а показан временной ряд, генерируемый процессом $(1 - 0,5B) \tilde{z}_t = a_t$, с его прогнозирующей функцией для момента $t = 14$. Поведение этой функции целиком определяется только отклонением \tilde{z}_{14} . Аналогично минимальная среднеквадратичная ошибка прогноза для процесса авторегрессии второго порядка такова, что текущее отклонение от среднего убывает до нуля как затухающая

синусоида или сумма двух экспонент. На рис. 5.9, б показан временной ряд, генерируемый процессом $(1 - 0,75B + 0,50B^2)\tilde{z}_t = a_t$, и его прогноз для момента $t = 14$. Поведение прогнозирующей функции для момента t целиком определяется последними двумя отклонениями \tilde{z}_{14} и \tilde{z}_{13} .

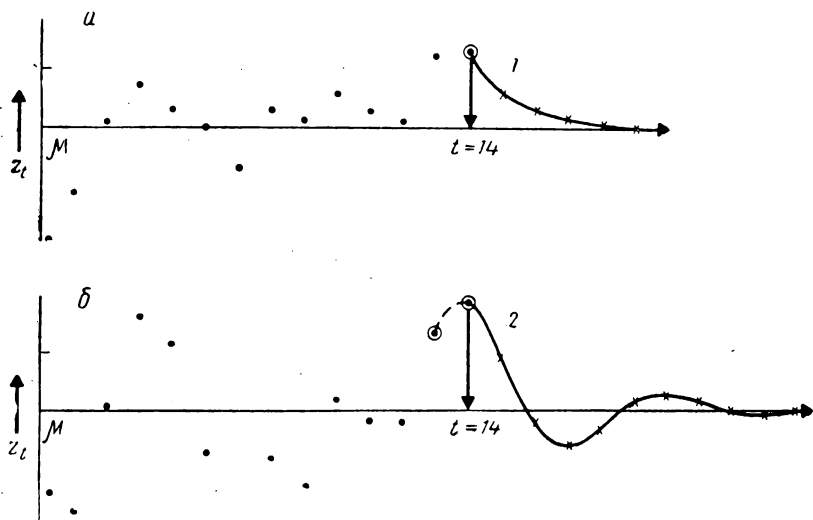


Рис. 5.9. Прогнозирующие функции процессов авторегрессии первого и второго порядка.

а — выборка из процесса авторегрессии первого порядка $(1 - 0,5B)\tilde{z}_t = a_t$ и прогнозирующая функция для момента $t = 14$.

б — выборка из процесса авторегрессии второго порядка $(1 - 0,75B + 0,50B^2)\tilde{z}_t = a_t$ и прогнозирующая функция на момент $t = 14$. 1 — ход прогноза, определяемый только по текущему отклонению, 2 — ход прогноза, определяемый по текущему и предпоследнему отклонениям.

Функция дисперсии для прогноза процесса $(1, 0, 0)$. Чтобы полнее проиллюстрировать использование (5.1.16), выведем функцию дисперсии для процесса авторегрессии первого порядка. Так как модель в момент $t + l$ может быть представлена в виде

$$\tilde{z}_{t+l} = a_{t+l} + \phi a_{t+l-1} + \dots + \phi^{l-1} a_{t+1} + \phi^l \tilde{z}_t,$$

то из (5.4.15) следует, что

$$e_t(l) = \tilde{z}_{t+l} - \hat{\tilde{z}}_t(l) = a_{t+l} + \phi a_{t+l-1} + \dots + \phi^{l-1} a_{t+1}.$$

Отсюда

$$V(l) = \text{var}[e_t(l)] = \sigma_a^2 (1 + \phi^2 + \dots + \phi^{2l-2}) = \frac{\sigma_a^2 (1 - \phi^{2l})}{1 - \phi^2}. \quad (5.4.16)$$

Мы видим, что для этого стационарного процесса при стремлении l к бесконечности дисперсия растет до постоянного значения $\sigma_a^2/(1-\phi^2)$, равного среднеквадратичному отклонению процесса относительно окончательного прогноза μ . Это поведение существенно иное, чем у нестационарных моделей, у которых функция дисперсии прогноза при больших упреждениях неограниченно растет.

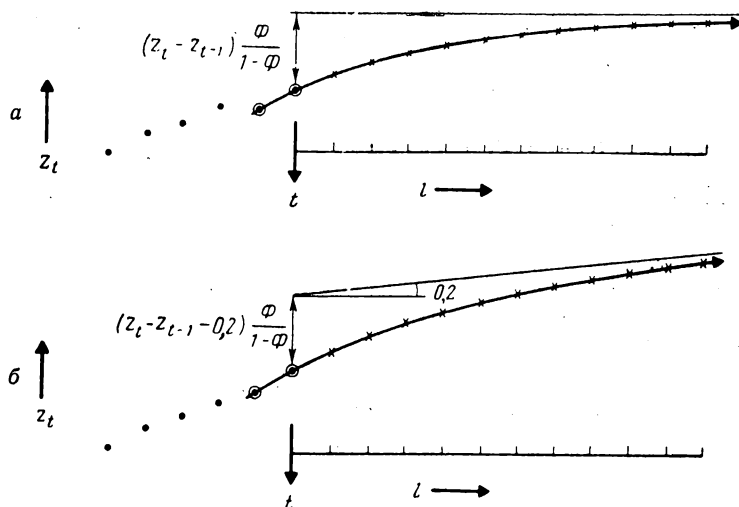


Рис. 5.10. Прогнозирующая функция для двух процессов $(1,1,0)$.
 a — прогнозы процесса $(1,1,0) \quad (1-0,8B) \nabla z_t = a_t$, b — прогнозы процесса $(1,1,0) \quad (1-0,8B) (\nabla z_t - 0,2) = a_t$.

Нестационарные модели авторегрессии порядка $(p, d, 0)$.
 Для модели

$$\phi(B) \nabla^d z_t = a_t$$

затухать до своего среднего значения при прогнозировании на несколько шагов вперед будет уже d -я разность процесса. Среднее $\nabla^d z_t$ будет, как правило, предполагаться нулевым, если только не появится каких-либо свидетельств в пользу противного. Если это необходимо, можно, как указывалось в гл. 4, ввести ненулевое среднее, заменив в модели $\nabla^d z_t$ на $\nabla^d z_t - \mu$. Например, рассмотрим модель

$$(1 - \phi B) (\nabla z_t - \mu) = a_t. \quad (5.4.17)$$

После замены t на $t + j$, переходя к условным математическим ожиданиям при известном на момент t прошлом, легко получим

(ср. с (5.4.15) и следующими за ним)

$$\hat{z}_t(j) - \hat{z}_t(j-1) - \mu = \phi^j (z_t - z_{t-1} - \mu).$$

Эта формула показывает, что прогнозируемая *разность* экспоненциально убывает от начального значения $z_t - z_{t-1}$ до своего среднего значения μ . Суммируя такие выражения для j , изменяющегося от 1 до l , получаем прогнозирующую функцию

$$\hat{z}_t(l) = z_t + \mu l + (z_t - z_{t-1} - \mu) \frac{\phi(1 - \phi^l)}{1 - \phi}, \quad l \geq 1,$$

которая асимптотически приближается к прямой линии

$$\hat{f}(l) = z_t + \mu l + (z_t - z_{t-1} - \mu) \frac{\phi}{1 - \phi}.$$

На рис. 5.10 показаны прогнозы для двух случаев $\phi = 0,8$, $\mu = 0$ и $\phi = 0,8$, $\mu = 0,2$. Мы увидим в гл. 7, что модель (5.4.17) с $\phi = 0,8$ и $\mu = 0$ хорошо описывает ряд С; прогнозы, основанные на этой модели, уже демонстрировались на рис. 5.1 и 5.2. Рассмотрим теперь прогнозирование некоторых важных смешанных моделей.

5.4.5. Прогнозирование процесса (1, 0, 1)

Способ разностного уравнения. Рассмотрим стационарную модель

$$(1 - \phi B)\tilde{z}_t = (1 - \theta B)a_t.$$

Прогнозы легко вычисляются из

$$\begin{aligned} \hat{\tilde{z}}_t(1) &= \phi \tilde{z}_t - \theta a_t, \\ \hat{\tilde{z}}_t(l) &= \phi \hat{\tilde{z}}_t(l-1), \quad l \geq 2. \end{aligned} \tag{5.4.18}$$

Прогнозы спадают к среднему, как члены геометрической прогрессии, аналогично процессу авторегрессии первого порядка, но прогноз на один шаг изменен введением множителя, зависящего от $a_t = z_t - \hat{z}_{t-1}(1)$. Веса ψ равны

$$\psi_j = (\phi - \theta) \phi^{j-1}, \quad j = 1, 2, \dots;$$

отсюда, используя (5.2.5), получаем правило для подправления прогнозов с упреждением 1, 2, ..., $L-1$ по значениям предыдущих прогнозов с упреждением 2, 3, ..., L :

$$\hat{\tilde{z}}_{t+1}(l) = \hat{\tilde{z}}_t(l+1) + (\phi - \theta) \phi^{l-1} a_{t+1}.$$

Проинтегрированное представление. Эвентуальная прогнозирующая функция для всех $l > 0$ — это решение уравнения

$$(1 - \phi B) \hat{z}_t(l) = 0, \text{ т. е.}$$

$$\hat{z}_t(l) = b_0^{(l)} \phi^l, \quad l > 0,$$

но

$$\hat{z}_t(1) = b_0^{(1)} \phi = \phi \tilde{z}_t - \theta a_t = \left\{ \left(1 - \frac{\theta}{\phi} \right) \tilde{z}_t + \frac{\theta}{\phi} \hat{z}_{t-1}(1) \right\} \phi,$$

и отсюда

$$\hat{z}_t(l) = \left\{ \left(1 - \frac{\theta}{\phi} \right) \tilde{z}_t + \frac{\theta}{\phi} \hat{z}_{t-1}(1) \right\} \phi^l. \quad (5.4.19)$$

Следовательно, прогноз отклонения с упреждением l экспоненциально затухает от своего начального значения, полученного линейной интерполяцией между предыдущим отклонением прогноза на шаг вперед и текущим отклонением. Когда $\phi = 1$, прогноз для всех упреждений принимает знакомый вид экспоненциально взвешенного среднего, и (5.4.19) становится равным (5.4.3).

Веса, приданные предыдущим наблюдениям. Веса π , а значит, и веса, приданные предыдущим наблюдениям для получения прогнозов на шаг вперед, равны

$$\pi_j = (\phi - \theta) \theta^{j-1}, \quad j = 1, 2, \dots$$

Заметим, что сумма весов для этого стационарного процесса равна $(\phi - \theta)/(1 - \theta)$, а не 1. Если бы ϕ было равно 1, процесс стал бы нестационарным процессом ПСС(0, 1, 1), сумма весов равнялась 1, а поведение генерируемого ряда не зависело от уровня z_t .

Например, к ряду А будет позднее подогнан процесс (1, 0, 1) с $\phi = 0,9$ и $\theta = 0,6$, и, следовательно, веса будут равны $\pi_1 = 0,30$, $\pi_2 = 0,18$, $\pi_3 = 0,11$, $\pi_4 = 0,07 \dots$ с суммой 0,75. Прогнозы (5.4.19) очень медленно спадают к среднему и для малых упреждений практически неотличимы от прогнозов, получаемых из процесса ПСС(0, 1, 1) для модели $\nabla z_t = a_t - 0,7a_{t-1}$, для которой веса равны $\pi_1 = 0,30$; $\pi_2 = 0,21$; $\pi_3 = 0,15$; $\pi_4 = 0,10$ с суммой 1. Последняя модель имеет то преимущество, что не привязывает процесса к фиксированному среднему значению.

Функция дисперсии. Так как веса ψ даны выражением

$$\psi_j = (\phi - \theta) \phi^{j-1}, \quad j = 1, 2, \dots,$$

отсюда вытекает, что функция дисперсии равна

$$V(l) = \sigma_a^2 \left\{ 1 + (\phi - \theta)^2 \left(\frac{1 - \phi^{2l-2}}{1 - \phi^2} \right) \right\} \quad (5.4.20)$$

и асимптотически растет до значения $\sigma_a^2(1 - 2\phi\theta + \theta^2)/(1 - \phi^2)$.

5.4.6. Прогнозирование процесса (1, 1, 1)

Другой важной смешанной моделью является нестационарный процесс (1, 1, 1)

$$(1 - \phi B) \nabla z_t = (1 - \theta B) a_t.$$

Способ разностного уравнения. В момент $t + l$ модель можно представить в виде

$$z_{t+l} = (1 + \phi) z_{t+l-1} - \phi z_{t+l-2} + a_{t+l} - \theta a_{t+l-1}.$$

Беря условные математические ожидания от обеих частей, получаем

$$\begin{aligned} \hat{z}_t(1) &= (1 + \phi) z_t - \phi z_{t-1} - \theta a_t, \\ \hat{z}_t(l) &= (1 + \phi) \hat{z}_t(l-1) - \phi \hat{z}_t(l-2), \quad l > 1. \end{aligned} \quad (5.4.21)$$

Представление в проинтегрированном виде. Так как $q < p + d$, эвентуальная прогнозирующая функция для всех $l > 0$ — это решение уравнения $(1 - \phi B)(1 - B)\hat{z}_t(l) = 0$, равное

$$\hat{z}_t(l) = b_0^{(t)} + b_1^{(t)} \phi^l.$$

Подставляя выражения для $\hat{z}_t(1)$ и $\hat{z}_t(2)$ в (5.4.21), получаем явные выражения

$$\begin{aligned} b_0^{(t)} &= z_t + \frac{\phi}{1 - \phi} (z_t - z_{t-1}) - \frac{\theta}{1 - \phi} a_t, \\ b_1^{(t)} &= \frac{\theta a_t - \phi (z_t - z_{t-1})}{1 - \phi}. \end{aligned}$$

Таким образом,

$$\hat{z}_t(l) = z_t + \phi \frac{1 - \phi^l}{1 - \phi} (z_t - z_{t-1}) - \theta \frac{1 - \phi^l}{1 - \phi} a_t. \quad (5.4.22)$$

Очевидно, что с ростом l прогноз стремится к $b_0^{(t)}$.

Веса, приданные предшествующим наблюдениям. Исключая a_t из (5.4.22), получаем другой вид выражения прогноза через предшествующие z :

$$\hat{z}_t(l) = \left\{ 1 - \frac{\theta - \phi}{1 - \phi} (1 - \phi^l) \right\} z_t + \left\{ \frac{\theta - \phi}{1 - \phi} (1 - \phi^l) \right\} \bar{z}_{t-1}(\theta), \quad (5.4.23)$$

где $\bar{z}_{t-1}(\theta)$ — это экспоненциально взвешенное скользящее среднее с параметром θ , т. е. $\bar{z}_{t-1}(\theta) = (1 - \theta) \sum_{j=1}^{\infty} \theta^{j-1} z_{t-j}$. Отсюда

веса π процесса состоят из «пики» в момент t и ЭВСС, начинающегося в момент $t - 1$. Если мы будем трактовать $(1 - \alpha)x + \alpha y$ как линейную интерполяцию между x и y для значения аргумента α , то прогноз (5.4.23) — это линейная интерполяция между z_t и $\bar{z}_{t-1}(\theta)$. Аргумент для единичного упрежде-

ния равен $\theta - \phi$, но по мере роста упреждения аргумент стремится к $(\theta - \phi)/(1 - \phi)$. Например, когда $\theta = 0,9$ и $\phi = 0,5$, прогноз на шаг вперед будет

$$\hat{z}_t(1) = 0,6z_t + 0,4\bar{z}_{t-1}(\theta),$$

а для больших упреждений

$$\hat{z}_t(\infty) = 0,2z_t + 0,8\bar{z}_{t-1}(\theta).$$

5.5. РЕЗЮМЕ

Результаты этой главы сводятся к следующему. Пусть \bar{z}_t будет отклонением наблюдаемого ряда от любой известной детерминированной функции $f(t)$. В частности, для стационарного ряда $f(t)$ может быть равно среднему значению ряда μ или нулю, так что \bar{z}_t образуют наблюдаемый ряд. Рассмотрим теперь общую модель АРПСС

$$\phi(B)\nabla^d \bar{z}_t = \theta(B)a_t$$

или

$$\phi(B)\bar{z}_t = \theta(B)a_t.$$

Прогноз с минимальной среднеквадратичной ошибкой. Пусть известны значения ряда до момента t . Тогда прогноз $\hat{\bar{z}}_t(l)$ ($l > 0$) с минимальной среднеквадратичной ошибкой — это условное математическое ожидание величины \bar{z}_{t+l} при заданных значениях \bar{z}_t, \bar{z}_{t-1}

$$\hat{\bar{z}}_t(l) = [\bar{z}_{t+l}] = E[\bar{z}_{t+l} | \bar{z}_t, \bar{z}_{t-1}, \dots].$$

Ошибки прогноза на шаг вперед. Отсюда следует, что ошибки прогноза с упреждением, равным единице (на шаг вперед), — это не коррелированные между собой импульсы, генерируемые моделью.

Вычисление прогнозов. На практике простейший способ вычисления прогнозов — непосредственное использование разностного уравнения, а именно

$$\begin{aligned} \hat{\bar{z}}_t(l) = & \varphi_1[\bar{z}_{t+l-1}] + \dots + \varphi_{p+d}[\bar{z}_{t+l-p-d}] + \\ & + [a_{t+l}] - \theta_1[a_{t+l-1}] - \dots - \theta_q[a_{t+l-q}]. \end{aligned} \quad (5.5.1)$$

Для вычисления условных математических ожиданий в (5.5.1) вместо \bar{z} , известных к этому моменту, подставляются их значения, вместо будущих значений z — их прогноз, вместо известных a — их значения и вместо будущих a — нули. Процесс прогнозирования может быть начат аппроксимацией неизвестных a нулями; практически удобные представления моделей и оценки параметров получают методами, описанными в гл. 6, 7 и 8.

Вероятностные пределы для прогнозов. Их можно получить а) вычислением вначале весов ψ по формулам

$$\begin{aligned}\psi_0 &= 1, \\ \psi_1 &= \varphi_1 - \theta_1, \\ \psi_2 &= \varphi_1\psi_1 + \varphi_2 - \theta_2 \\ &\vdots \\ \psi_j &= \varphi_1\psi_{j-1} + \dots + \varphi_{p+d}\psi_{j-p-d} - \theta_j,\end{aligned}\quad (5.5.2)$$

где $\theta_j = 0$ для $j > q$;

б) подстановкой весов ψ в

$$\tilde{z}_{t+l}(\pm) = \hat{z}_t(l) \pm u_{\varepsilon/2} \left(1 + \sum_{j=1}^{l-1} \psi_j^2\right)^{1/2} \sigma_a \quad (5.5.3)$$

для каждого нужного уровня вероятности ε и для каждого упреждения l ; здесь σ_a можно заменить практически его оценкой s_a — выборочным стандартным отклонением белого шума a_t , а $u_{\varepsilon/2}$ — квантиль уровня $1 - \varepsilon/2$ стандартного нормального распределения.

Подправка прогнозов. Когда становится известным новое отклонение \tilde{z}_{t+1} , прогноз для момента $t+1$ может быть скорректирован. Для этого нужно вычислить новую ошибку прогноза $a_{t+1} = \tilde{z}_{t+1} - \hat{z}_t(1)$ и использовать разностное уравнение (5.5.1), где t заменено на $t+1$. Другой возможный метод — использовать прогнозы $\hat{z}_t(1), \hat{z}_t(2), \dots, \hat{z}_t(L)$ на момент t , получить первые $L-1$ прогнозов $\hat{z}_{t+1}(1), \hat{z}_{t+1}(2), \dots, \hat{z}_{t+1}(L-1)$ на момент $t+1$ из формулы

$$\hat{z}_{t+1}(l) = \hat{z}_t(l+1) + \psi_l a_{t+1} \quad (5.5.4)$$

и затем найти последний прогноз $\hat{z}_{t+1}(L)$, используя разностное уравнение (5.5.1).

Другие способы представления прогнозов. Способы, описанные выше, охватывают все, что нужно для практического прогнозирования. Однако описанные ниже альтернативные представления позволяют углубить теоретическое понимание природы прогнозов, генерируемых разными моделями.

1) *Прогнозы в проинтегрированной форме.* Для $l > q - p - d$ прогнозы задаются одной кривой

$$\hat{z}_t(l) = b_0^{(t)} f_0(l) + b_1^{(t)} f_1(l) + \dots + b_{p+d-1}^{(t)} f_{p+d-1}(l), \quad (5.5.5)$$

определяемой «опорными» значениями $\hat{z}_t(q), \hat{z}_t(q-1), \dots, \hat{z}_t(q-p-d+1)$, где $\hat{z}_t(-j) = \tilde{z}_{t-j}$ ($j = 0, 1, 2, \dots$). Если $q > p + d$, первые $q - p - d$ прогнозов не лежат на этой кри-

вой. В общем стационарный оператор авторегрессии приводит к появлению в (5.5.5) затухающих экспоненциальных и синусоидальных членов, а нестационарный оператор ∇^d — полиномиальных членов.

Подстраивающиеся коэффициенты $b_j^{(t)}$ в (5.5.5) при переходе от момента t к моменту $t + 1$ могут быть скорректированы на величину, зависящую от последней ошибки прогноза на шаг вперед a_{t+1} , согласно общей формуле

$$\mathbf{b}^{(t+1)} = \mathbf{L}' \mathbf{b}^{(t)} + \mathbf{g} a_{t+1}, \quad (5.5.6)$$

выведенной в приложении П5.3.

2) *Прогноз как взвешенная сумма прошлых наблюдений.* С теоретической точки зрения полезно представить прогноз как взвешенную сумму прошлых наблюдений. Так, если модель представлена в обращенной форме

$$a_t = \pi(B) \tilde{z}_t = (1 - \pi_1 B - \pi_2 B^2 - \dots) \tilde{z}_t,$$

прогноз с упреждением 1 равен

$$\hat{\tilde{z}}_t(1) = \pi_1 \tilde{z}_t + \pi_2 \tilde{z}_{t-1} + \dots, \quad (5.5.7)$$

а прогнозы для больших упреждений можно получить из

$$\hat{\tilde{z}}_t(l) = \pi_1 [\tilde{z}_{t+l-1}] + \pi_2 [\tilde{z}_{t+l-2}] + \dots, \quad (5.5.8)$$

где условные математические ожидания известных \tilde{z} надо заменить их значениями, а для будущих \tilde{z} — их прогнозом.

Прогноз для любого упреждения можно записать и иначе, а именно как линейную функцию имеющихся наблюдений. Тогда

$$\hat{\tilde{z}}_t(l) = \sum_{j=1}^{\infty} \pi_j^{(l)} \tilde{z}_{t+j-1},$$

где $\pi_j^{(l)}$ — функции весов π .

ПРИЛОЖЕНИЕ П5.1. КОРРЕЛЯЦИЯ МЕЖДУ ОШИБКАМИ ПРОГНОЗА

П5.1.1. Автокорреляционная функция ошибок прогноза, сделанных в различные моменты времени

Хотя справедливо, что для оптимального прогноза ошибки прогноза с упреждением 1 некоррелированы, это, вообще говоря, не верно для прогнозов с большим упреждением. Рассмотрим прогнозы для упреждения l , сделанные в моменты t и $t - j$, где j — положительное целое число. Тогда, если $j = l$, $l + 1$, $l + 2$, ..., ошибки прогноза не будут содержать общих

компонент; однако для $j = 1, 2, \dots, l-1$ некоторые из a будут входить в ошибки обоих прогнозов. Конкретно

$$e_t(l) = z_{t+l} - \hat{z}_t(l) = a_{t+l} + \psi_1 a_{t+l-1} + \dots + \psi_{l-1} a_{t+1},$$

$$e_{t-j}(l) = z_{t+l-j} - \hat{z}_{t-j}(l) = a_{t-j+l} + \psi_1 a_{t-j+l-1} + \dots + \psi_{l-1} a_{t-j+1},$$

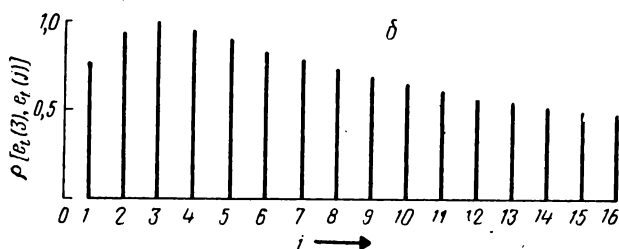
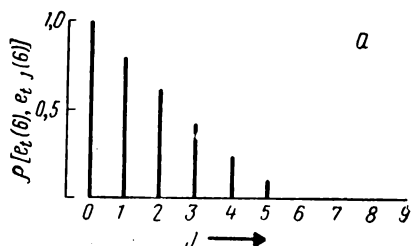


Рис. П5.1. Корреляция между различными ошибками прогноза для ряда С. а — автокорреляции ошибок прогноза ряда С для различных моментов с упреждением $l=6$. б — корреляция между ошибками прогноза ряда С для одного и того же момента t при упреждениях 3 и j .

и для $j < l$ автоковариация ошибок прогноза с упреждением l для задержки j равна

$$E[e_t(l) e_{t-j}(l)] = \sigma_a^2 \sum_{i=j}^{l-1} \psi_i \psi_{i-j}, \quad (\text{П5.1.1})$$

где $\psi_0 = 1$. Соответствующие автокорреляции равны

$$\rho[e_t(l), e_{t-j}(l)] = \begin{cases} \frac{\sum_{i=j}^{l-1} \psi_i \psi_{i-j}}{\sum_{i=0}^{l-1} \psi_i^2}, & 0 \leq j < l, \\ 0, & j \geq l. \end{cases} \quad (\text{П5.1.2})$$

В гл. 7 будет показано, что ряд С (рис. 4.1) хорошо описывается моделью $(1, 1, 0)$: $(1 - 0,8B) \nabla z_t = a_t$. Для иллюстрации фор-

мулы (П5.1.2) мы вычислили автокорреляционные функции ошибок прогноза для упреждения $l = 6$ в этой модели. В разд. 5.2.2 было показано, что веса $\psi_1, \psi_2, \dots, \psi_5$ для этой модели равны 1,80; 2,44; 2,95; 3,36; 3,69. Отсюда автоковариация для задержки 1 равна

$$E[e_t(6) e_{t-1}(6)] = \\ = \sigma_a^2 \{(1,80 \times 1,00) + (2,44 \times 1,80) + \dots + (3,69 \times 3,36)\} = 35,70 \sigma_a^2.$$

Делением на $E[e_t^2(6)] = 43,86 \sigma_a^2$ получаем $\rho[e_t(6), e_{t-1}(6)] = 0,81$. Первые 6 автокорреляций приведены в табл. П5.1 и на рис. П5.1, а. Как ожидалось, автокорреляции для задержек больше 5 равны нулю.

Таблица П5.1

Автокорреляции ошибок прогноза с упреждением 6 для ряда С

j	0	1	2	3	4	5	6
$\rho[e_t(6), e_{t-j}(6)]$	1,00	0,81	0,61	0,41	0,23	0,08	0,00

П5.1.2. Корреляция между ошибками прогноза в один и тот же момент времени при разных упреждениях

Положим, что мы сделали серию прогнозов для различных упреждений в *один и тот же* момент t . Тогда ошибки таких прогнозов будут коррелированы. Для $j = 1, 2, 3 \dots$ имеем

$$e_t(l) = z_{t+l} - \hat{z}_t(l) = a_{t+l} + \psi_1 a_{t+l-1} + \dots + \psi_{l-1} a_{t+1}, \\ e_t(l+j) = z_{t+l+j} - \hat{z}_t(l+j) = a_{t+l+j} + \psi_1 a_{t+l+j-1} + \dots \\ \dots + \psi_j a_{t+l} + \psi_{j+1} a_{t+l-1} + \dots + \psi_{l+j-1} a_{t+1},$$

так что ковариация между прогнозами на момент t с упреждениями l и $l+j$ равна $\sigma_a^2 \sum_{i=0}^{l-1} \psi_i \psi_{j+i}$, где $\psi_0 = 1$. Отсюда коэффициент корреляции между ошибками прогноза на момент t с упреждениями l и $l+j$ равен

$$\rho[e_t(l), e_t(l+j)] = \frac{\sum_{i=0}^{l-1} \psi_i \psi_{j+i}}{\left(\sum_{h=0}^{l-1} \psi_h^2 \sum_{g=0}^{l+j-1} \psi_g^2 \right)^{1/2}}. \quad (\text{П5.1.3})$$

Для того чтобы проиллюстрировать формулу (П5.1.3), мы вычислили для прогнозов ряда С, сделанных в один и тот же

момент, корреляцию между ошибкой прогноза с упреждением 3 и ошибками с упреждениями $j = 1, 2, 3, 4, \dots, 16$. Например, пользуясь (П5.1.3) и весами ψ из разд. 5.2.2, получаем

$$E[e_t(3)e_t(5)] = \\ = \sigma_a^2 \{(1,00 \times 2,44) + (1,80 \times 2,95) + (2,44 \times 3,36)\} = 15,94\sigma_a^2.$$

Корреляции для упреждений $j = 1, 2, \dots, 16$ приведены в табл. П5.2 и на рис. П5.1, б.

Таблица П5.2

Корреляция между ошибками прогноза с упреждением 3 и упреждениями j , сделанными в один и тот же момент времени (ряд С)

j	1	2	3	4	5	6	7	8
$\rho[e_t(3), e_t(j)]$	0,76	0,94	1,00	0,96	0,91	0,85	0,80	0,75
j	9	10	11	12	13	14	15	16
$\rho[e_t(3), e_t(j)]$	0,71	0,67	0,63	0,60	0,57	0,54	0,52	0,50

Как и ожидалось, прогнозы, сделанные на один и тот же момент времени с различными упреждениями, сильно коррелированы.

ПРИЛОЖЕНИЕ П5.2. ВЕСА ПРОГНОЗА ДЛЯ ПРОИЗВОЛЬНОГО УПРЕЖДЕНИЯ

В этом приложении мы рассмотрим альтернативную процедуру вычисления весов прогноза $\pi_l^{(i)}$, придаваемых предыдущим z , для любого упреждения l . Чтобы получить этот результат, мы используем тождество (3.1.7), а именно

$$(1 + \psi_1 B + \psi_2 B^2 + \dots)(1 - \pi_1 B - \pi_2 B^2 - \dots) = 1,$$

откуда можно выразить веса π через веса ψ и наоборот.

Приравнивая коэффициенты, получаем для $j \geq 1$

$$\psi_j = \sum_{i=1}^j \psi_{j-i} \pi_i \quad (\psi_0 = 1). \quad (\text{П5.2.1})$$

Так, например,

$$\psi_1 = \pi_1,$$

$$\pi_1 = \psi_1,$$

$$\psi_2 = \psi_1 \pi_1 + \pi_2,$$

$$\pi_2 = \psi_2 - \psi_1 \pi_1,$$

$$\psi_3 = \psi_2 \pi_1 + \psi_1 \pi_2 + \pi_3,$$

$$\pi_3 = \psi_3 - \psi_2 \pi_1 - \psi_1 \pi_2.$$

Из (5.3.6) получаем

$$\hat{z}_t(l) = \pi_1 \hat{z}_t(l-1) + \pi_2 \hat{z}_t(l-2) + \dots + \pi_{l-1} \hat{z}_t(1) + \pi_l z_t + \pi_{l+1} z_{t-1} + \dots \quad (\text{П5.2.2})$$

Так как каждый из прогнозов в (П5.2.2) сам является функцией наблюдений $z_t, z_{t-1}, z_{t-2}, \dots$, можно представить $\hat{z}_t(l)$ как

$$\hat{z}_t(l) = \pi_1^{(l)} z_t + \pi_2^{(l)} z_{t-1} + \pi_3^{(l)} z_{t-2} + \dots,$$

где веса прогноза с упреждением l могут быть вычислены из весов прогноза на шаг вперед $\pi_j^{(1)} = \pi_j$. Покажем теперь, что веса $\pi_j^{(l)}$ можно получить, используя тождество

$$\pi_j^{(l)} = \sum_{i=1}^l \psi_{l-i} \pi_{i+j-1}. \quad (\text{П5.2.3})$$

Например, веса для прогноза с упреждением 3 равны

$$\pi_1^{(3)} = \pi_3 + \psi_1 \pi_2 + \psi_2 \pi_1,$$

$$\pi_2^{(3)} = \pi_4 + \psi_1 \pi_3 + \psi_2 \pi_2,$$

$$\pi_3^{(3)} = \pi_5 + \psi_1 \pi_4 + \psi_2 \pi_3$$

и т. д. Чтобы вывести (П5.2.3), мы запишем

$$\begin{aligned} \hat{z}_t(l) &= \psi_l a_t + \psi_{l+1} a_{t-1} + \dots, \\ \hat{z}_{t+l-1}(1) &= \psi_1 a_{t+l-1} + \dots + \psi_l a_t + \psi_{l+1} a_{t-1} + \dots \end{aligned}$$

Вычитанием второго уравнения из первого получаем

$$\hat{z}_t(l) = \hat{z}_{t+l-1}(1) - \psi_1 a_{t+l-1} - \psi_2 a_{t+l-2} - \dots - \psi_{l-1} a_{t+1}.$$

Отсюда

$$\begin{aligned} \hat{z}_t(l) &= \pi_1 z_{t+l-1} + \pi_2 z_{t+l-2} + \dots + \pi_{l-1} z_{t+1} + \pi_l z_t + \pi_{l+1} z_{t-1} + \dots \\ &\dots + \psi_1 (-z_{t+l-1} + \pi_1 z_{t+l-2} + \dots + \pi_{l-2} z_{t+1} + \pi_{l-1} z_t + \pi_l z_{t-1} + \dots) + \\ &\quad + \psi_2 (-z_{t+l-2} + \dots + \pi_{l-3} z_{t+1} + \pi_{l-2} z_t + \pi_{l-1} z_{t-1} + \dots) + \dots \\ &\dots + \psi_{l-1} (-z_{t+1} + \pi_1 z_t + \pi_2 z_{t-1} + \dots). \end{aligned}$$

Из соотношений (П5.2.1) следует, что каждый из коэффициентов $z_{t+l-1}, \dots, z_{t+1}$ равен нулю, и после приведения подобных членов получаем доказываемую формулу (П5.2.3). Можно представить эту формулу и иначе, в рекуррентном виде

$$\pi_j^{(l)} = \pi_{j+1}^{(l-1)} + \psi_{l-1} \pi_j. \quad (\text{П5.2.4})$$

Пользуясь в качестве примера моделью $\nabla^2 z_t = (1 - 0,9B + +0,5B^2) a_t$, мы вычислили веса для упреждения 2. Уравнение

(П5.2.4) дает

$$\pi_j^{(2)} = \pi_{j+1} + \psi_1 \pi_j$$

и, пользуясь весами в табл. 5.4 при $\psi_1 = 1,1$, мы получаем, например,

$$\pi_1^{(2)} = \pi_2 + \psi_1 \pi_1 = 0,490 + (1,1)(1,1) = 1,700,$$

$$\pi_2^{(2)} = \pi_3 + \psi_1 \pi_2 = -0,109 + (1,1)(0,49) = 0,430$$

и т. д. Первые 12 весов приведены в табл. 5.4.

ПРИЛОЖЕНИЕ П5.3. ПРОГНОЗИРОВАНИЕ ПРИ ПОМОЩИ ОБЩЕГО ПРОИНТЕГРИРОВАННОГО ПРЕДСТАВЛЕНИЯ

П5.3.1. Общий метод получения проинтегрированного представления

Подчеркнем еще раз, что для практического вычисления прогнозов использование разностного уравнения — наиболее простой путь. Нижеследующее общее рассмотрение проинтегрированного представления нужно, чтобы совершенствовать уже полученные прогнозы. В этом рассмотрении вместо того, чтобы получить явное выражение для функции прогноза, как мы делали в примерах из разд. 5.4, будет удобнее выписать общее выражение для эвентуальной прогнозирующей функции, включающее $p + d$ подстраиваемых коэффициентов. Мы затем покажем, как нужно изменять эвентуальную прогнозирующую функцию, чтобы получать первые $q - p - d$ прогнозов при $q > p + d$. В завершение будет показано, как корректировать подстраиваемые коэффициенты при переходе от момента t к моменту $t + 1$.

Поскольку очевидно, что $\hat{z}_t(-j) = z_{t-j}$ для $j = 0, 1, 2, \dots$, то, используя рассуждения об условных математических ожиданиях в разд. 5.1.2, легко прийти к разностному уравнению для прогнозов

$$\begin{aligned} \hat{z}_t(1) - \varphi_1 \hat{z}_t(0) - \dots - \varphi_{p+d} \hat{z}_t(1-p-d) &= -\theta_1 a_t - \dots - \theta_q a_{t-q+1}, \\ \hat{z}_t(2) - \varphi_1 \hat{z}_t(1) - \dots - \varphi_{p+d} \hat{z}_t(2-p-d) &= -\theta_2 a_t - \dots - \theta_q a_{t-q+2}, \\ \vdots \quad \quad \quad \ddots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \dots & \quad \quad \quad (\text{П5.3.1}) \\ \hat{z}_t(q) - \varphi_1 \hat{z}_t(q-1) - \dots - \varphi_{p+d} \hat{z}_t(q-p-d) &= -\theta_q a_t, \\ \hat{z}_t(l) - \varphi_1 \hat{z}_t(l-1) - \dots - \varphi_{p+d} \hat{z}_t(l-p-d) &= 0, \quad l > q. \end{aligned}$$

Эвентуальная прогнозирующая функция — это решение последнего уравнения и может быть представлена как

$$\begin{aligned}\hat{z}_t(l) &= b_0^{(t)} f_0(l) + b_1^{(t)} f_1(l) + \dots + b_{p+d-1}^{(t)} f_{p+d-1}(l) = \\ &= \sum_{i=0}^{p+d-1} b_i^{(t)} f_i(l), \quad l > q - p - d. \quad (\text{П5.3.2})\end{aligned}$$

Когда $q \leq p + d$, эвентуальная прогнозирующая функция дает прогнозы $\hat{z}_t(1)$, $\hat{z}_t(2)$, $\hat{z}_t(3)$, ... для всех упреждений $l \geq 1$.

Как пример такой модели с $q \leq p + d$ рассмотрим

$$(1 - B)(1 - \sqrt{3}B + B^2)^2 z_t = (1 - 0,5B) a_t,$$

так что $p + d = 5$ и $q = 1$. Тогда

$$(1 - B)(1 - \sqrt{3}B + B^2)^2 \hat{z}_t(l) = 0, \quad l = 2, 3, 4, \dots,$$

где B действует теперь на l , а не на t . Решение этого разностного уравнения дает функцию прогноза

$$\begin{aligned}\hat{z}_t(l) &= b_0^{(t)} + b_1^{(t)} \cos \frac{2\pi l}{12} + b_2^{(t)} l \cos \frac{2\pi l}{12} + \\ &+ b_3^{(t)} \sin \frac{2\pi l}{12} + b_4^{(t)} l \sin \frac{2\pi l}{12}, \quad l = 1, 2, \dots\end{aligned}$$

Если $q > p + d$, то для упреждения $l \geq q - p - d$ прогнозирующая функция будет иметь дополнительные члены, содержащие a . Отсюда

$$\hat{z}_t(l) = \sum_{i=0}^{p+d-1} b_i^{(t)} f_i(l) + \sum_{i=0}^l d_{i+1} a_{t-i}, \quad l \leq q - p - d, \quad (\text{П5.3.3})$$

где $j = q - p - d - l$ и d можно выразить явно, подставив (П5.3.3) в (П5.3.1). Например, рассмотрим стохастическую модель

$$\nabla^2 z_t = (1 - 0,8B + 0,5B^2 - 0,4B^3 + 0,1B^4) a_t,$$

в которой $p + d = 2$, $q = 4$, $q - p - d = 2$ и $\varphi_1 = 2$, $\varphi_2 = -1$, $\theta_1 = 0,8$, $\theta_2 = -0,5$, $\theta_3 = 0,4$, $\theta_4 = -0,1$. Пользуясь рекуррентным соотношением (5.2.3), получаем $\psi_1 = 1,2$, $\psi_2 = 1,9$, $\psi_3 = 2,2$, $\psi_4 = 2,6$. Далее, из (П5.3.3) следует

$$\begin{aligned}\hat{z}_t(1) &= b_0^{(t)} + b_1^{(t)} + d_{10} a_t + d_{11} a_{t-1}, \\ \hat{z}_t(2) &= b_0^{(t)} + 2b_1^{(t)} + d_{20} a_t, \\ \hat{z}_t(l) &= b_0^{(t)} + b_1^{(t)} l, \quad l > 2.\end{aligned} \quad (\text{П5.3.4})$$

Кроме того, (П5.3.1) дает

$$\hat{z}_t(4) - 2\hat{z}_t(3) + \hat{z}_t(2) = 0,1a_t,$$

так что из (П5.3.4) имеем

$$d_{20}a_t = 0,1a_t,$$

и, следовательно, $d_{20} = 0,1$. Аналогично из (П5.3.1)

$$\hat{z}_t(3) - 2\hat{z}_t(2) + \hat{z}_t(1) = -0,4a_t + 0,1a_{t-1},$$

и отсюда, пользуясь (П5.3.4), имеем

$$-0,2a_t + d_{10}a_t + d_{11}a_{t-1} = -0,4a_t + 0,1a_{t-1},$$

что приводит к

$$d_{10} = -0,2, \quad d_{11} = 0,1.$$

Следовательно, прогнозирующая функция будет

$$\hat{z}_t(1) = b_0^{(t)} + b_1^{(t)} - 0,2a_t + 0,1a_{t-1},$$

$$\hat{z}_t(2) = b_0^{(t)} + 2b_1^{(t)} + 0,1a_t,$$

$$\hat{z}_t(l) = b_0^{(t)} + b_1^{(t)}l, \quad l > 2.$$

П5.3.2. Коррекция общего проинтегрированного представления

Коррекцию формул для коэффициентов можно осуществить, используя тождество (5.2.5) с $t+1$, замененным на t ,

$$\hat{z}_t(l) = \hat{z}_{t-1}(l+1) + \psi_l a_t.$$

Тогда для $l > q - p - d$

$$\sum_{i=0}^{p+d-1} b_i^{(t)} f_i(l) = \sum_{i=0}^{p+d-1} b_i^{(t-1)} f_i(l+1) + \psi_l a_t. \quad (\text{П5.3.5})$$

Решив $p+d$ таких уравнений для разных l , получим требуемую формулу для коррекции в виде

$$b_i^{(t)} = \sum_{j=0}^{p+d-1} g_{ij} b_j^{(t-1)} + v_i a_t.$$

Заметим, что коррекция каждого коэффициента функции прогноза зависит только от ошибки прогноза на один шаг $a_t = z_t - \hat{z}_{t-1}(1)$. В качестве примера рассмотрим опять стохастическую модель

$$(1-B)(1-\sqrt{3}B+B^2)z_t = (1-0,5B)a_t.$$

В разд. П5.3.1 было показано, что прогнозирующая функция для этого примера равна

$$\hat{z}_t(l) = b_0^{(t)} + (b_1^{(t)} + lb_2^{(t)}) \cos \frac{2\pi l}{12} + (b_3^{(t)} + lb_4^{(t)}) \sin \frac{2\pi l}{12}.$$

Удобно использовать тождество коррекции (П5.3.5) для частных значений $l = 6, 12, 24, 3, 9$. Пользуясь рекуррентным соотношением (5.2.3), получаем веса ψ для этих значений: $\psi_6 = 25 + 9\sqrt{3}$, $\psi_{12} = 317 + 184\sqrt{3}$, $\psi_{24} = 611 + 346\sqrt{3}$, $\psi_3 = 4 + 7\sqrt{3}$, $\psi_9 = 134 + 90\sqrt{3}$. Отсюда, используя (П5.3.5), получаем матричное уравнение

$$\begin{bmatrix} 1 & -1 & -6 & 0 & 0 \\ 1 & 1 & 12 & 0 & 0 \\ 1 & 1 & 24 & 0 & 0 \\ 1 & 0 & 0 & 1 & 3 \\ 1 & 0 & 0 & -1 & -9 \end{bmatrix} \begin{bmatrix} b_0^{(t)} \\ b_1^{(t)} \\ b_2^{(t)} \\ b_3^{(t)} \\ b_4^{(t)} \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & -\frac{\sqrt{3}}{2} & -\frac{7\sqrt{3}}{2} & -\frac{1}{2} & -\frac{7}{2} \\ 1 & \frac{\sqrt{3}}{2} & \frac{13\sqrt{3}}{2} & \frac{1}{2} & \frac{13}{2} \\ 1 & \frac{\sqrt{3}}{2} & \frac{25\sqrt{3}}{2} & \frac{1}{2} & \frac{25}{2} \\ 1 & -\frac{1}{2} & -2 & \frac{\sqrt{3}}{2} & 2\sqrt{3} \\ 1 & \frac{1}{2} & 5 & -\frac{\sqrt{3}}{2} & -5\sqrt{3} \end{bmatrix} \times$$

$$\begin{bmatrix} b_0^{(t-1)} \\ b_1^{(t-1)} \\ b_2^{(t-1)} \\ b_3^{(t-1)} \\ b_4^{(t-1)} \end{bmatrix} + \begin{bmatrix} 25 + 9\sqrt{3} \\ 317 + 184\sqrt{3} \\ 611 + 346\sqrt{3} \\ 4 + 7\sqrt{3} \\ 134 + 90\sqrt{3} \end{bmatrix} a_t.$$

Первые три уравнения легко решить относительно $b_0^{(t)}$, $b_1^{(t)}$, $b_2^{(t)}$, а четвертое и пятое — относительно $b_3^{(t)}$, $b_4^{(t)}$. Окончательное

решение будет

$$\begin{bmatrix} b_0^{(t)} \\ b_1^{(t)} \\ b_2^{(t)} \\ b_3^{(t)} \\ b_4^{(t)} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{\sqrt{3}}{2} & 0 & \frac{1}{2} \\ 0 & -\frac{1}{2} & -\frac{1}{2} & \frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} \\ 0 & 0 & -\frac{1}{2} & 0 & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} b_0^{(t-1)} \\ b_1^{(t-1)} \\ b_2^{(t-1)} \\ b_3^{(t-1)} \\ b_4^{(t-1)} \end{bmatrix} + \begin{bmatrix} \frac{195 + 112\sqrt{3}}{2} \\ \frac{-149 - 68\sqrt{3}}{2} \\ \frac{49 + 27\sqrt{3}}{2} \\ \frac{-244 - 113\sqrt{3}}{2} \\ \frac{19 + 5\sqrt{3}}{2} \end{bmatrix} a_t,$$

что дает требуемую формулу коррекции для каждого из 5 коэффициентов $b_0^{(t)}$, $b_1^{(t)}$, ..., $b_4^{(t)}$.

П5.3.3. Сравнение с методом взвешенных наименьших квадратов Брауна

Хотя проинтегрированное представление приводит к усложнению вычисления прогнозов, оно позволяет нам сравнить рассмотренный здесь прогноз с наименьшей среднеквадратичной ошибкой с другим достаточно известным типом прогноза.

Запишем

$$\mathbf{F}_l = \begin{bmatrix} f_0(l) & f_1(l) & \dots & f_{p+d-1}(l) \\ f_0(l+1) & f_1(l+1) & \dots & f_{p+d-1}(l+1) \\ \vdots & \vdots & & \vdots \\ f_0(l+p+d) & f_1(l+p+d) & \dots & f_{p+d-1}(l+p+d) \end{bmatrix},$$

$$\mathbf{b}^{(t)} = \begin{bmatrix} b_0^{(t)} \\ b_1^{(t)} \\ \vdots \\ b_{p+d-1}^{(t)} \end{bmatrix}, \quad \boldsymbol{\psi}_l = \begin{bmatrix} \psi_l \\ \psi_{l+1} \\ \vdots \\ \psi_{l+p+d} \end{bmatrix}.$$

Тогда, пользуясь (П5.3.5) для $l, l+1, \dots, l+p+d$, получим для $l > q - p - d$

$$\mathbf{F}_l \mathbf{b}^{(t)} = \mathbf{F}_{l+1} \mathbf{b}^{(t-1)} + \boldsymbol{\psi}_l a_t,$$

что дает

$$\mathbf{b}^{(t)} = (\mathbf{F}_l^{-1} \mathbf{F}_{l+1}) \mathbf{b}^{(t-1)} + (\mathbf{F}_l^{-1} \boldsymbol{\psi}_l) a_t$$

или

$$\mathbf{b}^{(t)} = \mathbf{L}' \mathbf{b}^{(t-1)} + \mathbf{g}_t a_t. \quad (\text{П5.3.6})$$

Это выражение имеет ту же алгебраическую форму, что и корректирующая функция, предлагаемая в методике «взвешенных наименьших квадратов» Брауна [2, 50]. Для сравнения, если мы обозначим ошибку прогноза, даваемую этим методом, e_t , то формулу коррекции Брауна можно записать в виде

$$\beta^{(t)} = L' \beta^{(t-1)} + h e_t, \quad (\text{П5.3.7})$$

где $\beta^{(t)}$ — вектор настраиваемых коэффициентов. Та же матрица L появляется в (П5.3.6) и (П5.3.7). Это неизбежно, так как первый множитель просто учитывает изменения в коэффициентах, возникающие при переходе к новому начальному моменту, и должен появиться в любой формуле такого типа. Рассмотрим, например, функцию прогноза в виде прямой линии

$$\hat{z}_{t-1}(l) = b_0^{(t-1)} + b_1^{(t-1)} l,$$

где $b_0^{(t-1)}$ — ордината в момент $t-1$ начала прогноза. Можно также записать

$$\hat{z}_{t-1}(l) = (b_0^{(t-1)} + b_1^{(t-1)}) + b_1^{(t-1)}(l-1),$$

где $(b_0^{(t-1)} + b_1^{(t-1)})$ — теперь ордината в момент t . Очевидно, если мы корректируем прогноз в момент t , коэффициент b_0 должен быть соответствующим образом подправлен, даже если бы функция прогноза осталась неизменной.

В общем матрица L не изменяет прогнозирующую функцию, а просто перемещает ее. Истинная корректировка производится вектором коэффициентов g или h . Далее мы увидим, что коэффициенты g , дающие прогнозы с минимальной среднеквадратичной ошибкой, и коэффициенты h Брауна в общем полностью отличны друг от друга.

Метод прогнозирования Брауна

1) Прогнозирующая функция выбирается из общего класса линейных комбинаций и произведений полиномов, экспонент, синусов и косинусов.

2) Выбранная прогнозирующая функция подгоняется к прошлым значениям при помощи методики «взвешенных наименьших квадратов». В этой методике коэффициенты оцениваются и корректируются так, что минимизируется сумма

$$S_\omega = \sum_{j=0}^{\infty} \omega_j \{z_{t-j} - \hat{z}_t(-j)\}^2 \quad (\text{П5.3.8})$$

взвешенных квадратов расхождения прошлых значений ряда и значений, даваемых для соответствующих времен в прошлом прогнозирующей функцией. Весовая функция ω_j выбирается произвольно, так чтобы она спадала по степенному закону $\omega_j = (1 - \alpha)^j$, где константа α , обычно называемая *константой сглаживания*, выбирается равной опять-таки произвольному значению в диапазоне 0,1—0,3.

Различие между прогнозами с минимальной среднеквадратичной ошибкой и прогнозами Брауна.

Чтобы проиллюстрировать эти замечания, рассмотрим прогнозирование цен акций IBM, обсуждавшееся Брауном [2, стр. 141]. В этом исследовании он использовал квадратичную модель, которая в обозначениях этой книги может быть представлена как

$$\hat{z}_t(l) = \beta_0^{(t)} + \beta_1^{(t)}l + \frac{1}{2} \beta_2^{(t)}l^2.$$

Он применил к этой модели свой метод «взвешенных наименьших квадратов», чтобы предсказать биржевые цены на три дня вперед. Результаты, полученные его методом, показаны для участка ряда IBM на рис. П5.2, где они сравниваются с прогнозами с минимальной среднеквадратичной ошибкой.

Метод «взвешенных наименьших квадратов» можно подвергнуть критике на следующих основаниях:

1) Вид прогнозирующей функции должен определяться оператором авторегрессии $\varphi(B)$ стохастической модели, а не выбираться произвольно. В частности, его нельзя уверенно выбирать в результате визуального просмотра ряда. Рассмотрим, например, биржевые цены акций IBM на рис. П5.2. Видно, что квадратичная функция может быть использована для *подгонок* коротких участков этого ряда к уже имеющимся значениям. Если бы такая подгонка имела отношение к прогнозированию, мы могли бы заключить, как и сделал Браун, что прогнозирующая функция должна быть полиномом второй степени.

Наиболее общий линейный процесс, для которого квадратичная функция давала бы прогнозы с *наименьшей среднеквадратичной ошибкой* при любом упреждении l , определяется моделью $(0, 3, 3)$

$$\nabla^3 z_t = (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3) a_t,$$

которая на основании тех же рассуждений, что и в разд. 4.3.3, может быть представлена в виде

$$\nabla^3 z_t = \nabla^3 a_t + \lambda_0 \nabla^2 a_{t-1} + \lambda_1 \nabla a_{t-1} + \lambda_2 a_{t-1}.$$

Однако в гл. 7 будет показано, что если правильно подогнать эту модель, оценки ее параметров, полученные методом наименьших квадратов, окажутся равными $\lambda_1 = \lambda_2 = 0$ и $\lambda_0 \approx 1, 0$. Отсюда $\nabla z_t = (1 - \theta B) a_t$ с $\theta = 1 - \lambda_0$, близкими к нулю, будет подходящей стохастической моделью, и соответствующий прогнозирующий полином будет $\hat{z}_t(l) = \beta_0^{(t)}$, т. е. нулевой, а не второй степени по l .

2) Выбор весовой функции ω_j в (П5.3.8) должен также определяться стохастической моделью, а не быть произвольным. Метод «взвешенных наименьших квадратов» будет давать про-

гнозы с минимальной среднеквадратичной ошибкой только в очень частном случае, когда

- а) процесс имел порядок $(0, 1, 1)$, так что $\nabla z_t = (1 - \theta B) a_t$;
- б) подгонялся полином нулевой степени;
- в) константа сглаживания α была принята равной нашему $\lambda = 1 - \theta$.

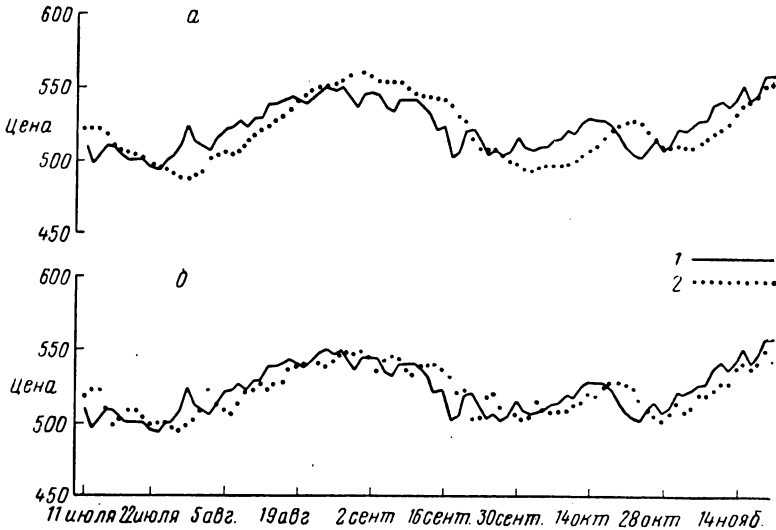


Рис. П5.2. Ряд биржевых цен акций IBM; сравнение с прогнозом на три шага вперед, основанным на лучшей модели ПСС(0,1,1), и квадратичным прогнозом Брауна для периода, начинающегося 11 июля 1960 г.

а — квадратичный прогноз Брауна, б — прогноз с минимальной среднеквадратичной ошибкой.

1 — данные; 2 — прогнозы на три шага вперед.

В данном примере, даже если бы была выбрана правильная степень полинома (нуль), значение $\alpha = \lambda = 0,1$, используемое Брауном, оказалось бы весьма не подходящим к случаю. Правильное значение λ для этого ряда близко к единице.

3) Методика экспоненциально взвешенных наименьших квадратов вынуждает все $p + d$ коэффициентов в корректирующем векторе \mathbf{h} быть функциями одного сглаживающего параметра α . Фактически же они должны быть функциями $p + q$ независимых параметров (ϕ, θ) .

Итак, различия между двумя методами нетривиальны, и интересно сравнить на данных IBM, как они действуют. Прогноз с минимальной среднеквадратичной ошибкой равен $\hat{z}_t(l) = b_0^{(t)}$ с корректировкой $b_0^{(t)} = b_0^{(t-1)} + \lambda a_t$, где $\lambda \approx 1,0$. Если взять λ

точно равным единице, это эквивалентно использованию

$$\hat{z}_t(l) = z_t,$$

т. е. лучший прогноз цен акций для всех будущих моментов времени есть текущая цена *). Идея, что биржевые цены ведут себя так, конечно, не нова и восходит к Бейгельеру [51]. При $\lambda = 1$ $z_t = Sa_t$, что свидетельствует, что z_t — случайное блуждание.

Для сопоставления прогноза с минимальной среднеквадратичной ошибкой (СКО) с квадратичными прогнозами Брауна было проведено прямое сравнение на базе ряда цен акций IBM за период с 11 июля 1960 г. по 10 февраля 1961 г. (150 наблюдений). Для этого участка ряда прогноз с минимальной СКО был получен для модели $Vz_t = a_t - \theta a_{t-1}$, $\theta = 0, 1$, или $\lambda = 1 - \theta = 0,9$. На рис. П5.2 показаны прогнозы с минимальной СКО для упреждения 3 и соответствующие значения квадратичного прогноза Брауна. Видно, что прогнозы с минимальной СКО, которые, по существу, эквивалентны использованию сегодняшней цены для предсказания цены на 3 дня вперед, значительно лучше, чем полученные с помощью существенно более сложной методики Брауна.

Среднеквадратичные ошибки прогноза при разных упреждениях, вычисленные прямым сравнением значений ряда и их прогнозов с упреждением l , показаны в табл. П5.3 для двух типов прогнозов. Видно, что квадратичные прогнозы Брауна имеют значительно большие среднеквадратичные ошибки, чем прогнозы, полученные методом минимальной среднеквадратичной ошибки.

Таблица П5.3

Сравнение среднеквадратичных ошибок прогнозов, полученных при различных упреждениях, по наилучшему из процессов ПСС(0, 1, 1) и по методике Брауна

Упреждение l	1	2	3	4	5	6	7	8	9	10
СКО (Браун)	102	158	218	256	363	452	554	669	799	944
СКО ($\lambda = 0,9$)	42	91	136	180	282	266	317	371	427	483

*) Этот результат приближенно справедлив, если предположить, что не имеется никакой полезной информации, кроме прошлых значений ряда, и что рассматривается прогнозирование на относительно короткий период времени. Для больших периодов становятся важными эффекты роста и инфляции.

Часть II

Построение стохастических моделей

Мы видели, что процесс АРПСС порядка (p, d, q) позволяет получить класс моделей, пригодных для описания временных рядов, которые не обязательно стационарны, но тем не менее в некотором смысле статистически однородны.

Процесс АРПСС определен уравнением

$$\phi(B)(1-B)^d z_t = \theta_0 + \theta(B) a_t,$$

где $\phi(B)$ и $\theta(B)$ — операторы от B степеней p и q соответственно, нули которых лежат вне единичного круга.

Мы отмечали, что эта модель весьма общая и объединяет модели авторегрессии, скользящего среднего, смешанные модели авторегрессии — скользящего среднего и проинтегрированные виды всех трех типов.

Итеративный подход к построению моделей. Увязка моделей этого типа с данными обычно лучше всего достигается трехступенчатой итеративной процедурой, основанной на идентификации, оценивании и диагностической проверке.

Под *идентификацией* подразумевается использование данных и любой информации о том, как был генерирован ряд, с целью отыскания подкласса экономичных моделей, заслуживающих опробования.

Под *оцениванием* подразумевается эффективное использование данных для получения выводов о параметрах, определяющих адекватность опробуемой модели.

Под *диагностической проверкой* подразумевается проверка согласования подогнанной модели с данными, чтобы вскрыть недостатки модели и улучшить ее.

В последующей гл. 6 мы рассмотрим идентификацию, в гл. 7 — оценивание и в гл. 8 — диагностическую проверку. Как работают все эти методы, будет показано в гл. 9 на примере моделирования сезонных временных рядов.

Идентификация моделей

В этой главе рассматриваются методы идентификации моделей несезонных временных рядов. Методы идентификации — это довольно грубые процедуры, применяемые к множеству данных для выявления типа моделей явления, которые имеет смысл использовать в дальнейших исследованиях. Конкретная цель состоит в том, чтобы получить некоторое указание на то, какие значения p , d и q нужны в общей линейной модели АРПСС, и выбрать некоторые начальные значения параметров. Полученная таким способом пробная модель является отправной точкой для применения более формальных и эффективных методов оценивания, описанных в гл. 7.

6.1. ЦЕЛИ ИДЕНТИФИКАЦИИ

6.1.1. Этапы процедуры идентификации

Прежде всего необходимо отметить, что идентификация и оценивание с неизбежностью перекрываются. Так, мы можем оценить параметры модели, которая оказалась более сложной, чем ожидалось, чтобы решить, что именно поддается упрощению. Здесь процедура оценивания выполняет роль идентификации. Нужно также пояснить, что идентификация неизбежно неточна. Она неточна потому, что вопрос о том, какие типы моделей встречаются на практике и в каких обстоятельствах — это черта поведения физического мира, и она не может быть решена чисто математическими рассуждениями. Далее, так как на этапе идентификации нельзя точно сформулировать задачу, приходится прибегать к статистически «неэффективным» методам. На этом этапе особенно полезны графические методы и должны высказываться суждения. Однако следует помнить, что предварительная идентификация не требует от нас ничего, кроме опробования класса моделей, который будет позднее эффективно подгоняться и проверяться.

Наша задача — идентифицировать подходящий подкласс моделей из общего семейства моделей АРПСС

$$\phi(B) \nabla^d z_t = \theta_0 + \theta(B) a_t, \quad (6.1.1)$$

который может быть использован для описания данного временного ряда. Наш подход состоит в следующем.

а) Мы будем брать конечную разность от z_t столько раз, сколько необходимо, чтобы обеспечить стационарность, в надежде свести изучаемый процесс к смешанному процессу авторегрессии — скользящего среднего

$$\phi(B) w_t = \theta_0 + \theta(B) a_t,$$

где

$$w_t = (1 - B)^d z_t = \nabla^d z_t.$$

б) Идентифицировать результирующий процесс АРПСС. Наш основной инструмент для реализации (а) и (б) — автокорреляционная и частная автокорреляционная функции. Они используются не только, чтобы облегчить угадывание вида модели, но также для получения приближенных оценок параметров. Такие приближения часто полезны на этапе оценивания, чтобы получить начальные значения для используемых на этом этапе итеративных процедур.

6.2. МЕТОДИКА ИДЕНТИФИКАЦИИ

6.2.1. Использование при идентификации автокорреляционной и частной автокорреляционной функций

Идентификация порядка разности. В разд. 3.4.2 было показано, что автокорреляционная функция стационарного смешанного процесса авторегрессии — скользящего среднего удовлетворяет разностному уравнению

$$\phi(B) \rho_k = 0, \quad k > q.$$

Кроме того, если $\phi(B) = \prod_{i=1}^p (1 - G_i B)$, решение этого разностного уравнения для k -й автокорреляции в предположении об отсутствии кратности корней имеет вид

$$\rho_k = A_1 G_1^k + A_2 G_2^k + \dots + A_p G_p^k, \quad k > q - p. \quad (6.2.1)$$

Условие стационарности, требующее, чтобы нули $\phi(B)$ лежали вне единичного круга, приводит к тому, что корни G_1, G_2, \dots, G_p лежат внутри единичного круга.

Из (6.2.1) следует, что для стационарной модели, у которой ни один из корней не лежит близко к границе единичного круга, автокорреляционная функция быстро затухает при средних и больших k . Положим теперь, что один действительный корень,

например G_1 , приближается к 1, так что

$$G_1 = 1 - \delta,$$

где δ малое положительное число. Тогда, поскольку для больших k

$$\rho_k \approx A_1(1 - k\delta),$$

автокорреляционная функция не будет быстро затухать, а будет спадать медленно и почти линейно. Подобные рассуждения можно провести и в случае, когда к единице приближается не один корень.

Следовательно, отсутствие у автокорреляционной функции тенденции к затуханию может рассматриваться как свидетельство того, что существует корень, близкий к 1. Выборочная автокорреляционная функция похожа на теоретическую. Отсюда отсутствие затухания выборочной автокорреляционной функции логично истолковать в том смысле, что процесс z_t ведет себя нестационарно, хотя возможно, его разность ∇z_t или какая-либо более высокая разность стационарна.

Оказалось, что нестационарность подсказывается отсутствием быстрого спада выборочной автокорреляционной функции. При этом не обязательно, чтобы выборочные корреляции при малых задержках были велики. Это иллюстрируется в приложении Пб.1, где вычислены ожидаемые значения выборочной автокорреляционной функции для нестационарного процесса $(0, 1, 1)$ $\nabla z_t = (1 - \theta B)a_t$. Как и следовало ожидать, величина $E[r_k]$ медленно убывает с k , ее начальное значение $E[r_1]$ зависит от значения θ и числа наблюдений в выборке и не должно быть близко к 1, если θ близко к 1. Мы покажем это также в разд. 6.3.4 для ряда А рис. 4.1.

По приведенным причинам предполагается, что необходимая для получения стационарности степень разности d достигнута, если автокорреляционная функция ряда $w_t = \nabla^d z_t$ быстро затухает. На практике d обычно равно 0, 1 или 2, и достаточно просмотреть примерно 20 первых значений автокорреляции исходного ряда, его первых и вторых разностей.

Идентификация результирующего стационарного процесса АРСС. Приняв предварительное решение о величине d , мы далее изучаем общий вид выборочных автокорреляционной и частной автокорреляционной функций соответствующего разностного ряда, чтобы найти указания к выбору порядков p и q операторов авторегрессии и скользящего среднего. При этом мы должны помнить характерные особенности поведения теоретической автокорреляционной функции и теоретической частной автокорреляционной функции для процессов авторегрессии, скользящего среднего и смешанного процесса, рассмотренные в гл. 3,

В то время как автокорреляционная функция процесса авторегрессии порядка p спадает плавно, ее частная автокорреляционная функция имеет обрыв после p -й задержки. Обратно, автокорреляционная функция процесса скользящего среднего порядка q обрывается после задержки q , в то время как ее частная автокорреляция плавно спадает с ростом задержки. Далее, автокорреляционная функция смешанного процесса, содержащая компоненту авторегрессии порядка p и компоненту скользящего среднего порядка q , после первых $q - p$ задержек представляется в виде суммы экспонент и затухающих синусоид. Обратно, частная автокорреляционная функция смешанного процесса приближенно представляется суммой экспонент и затухающих синусоид после $p - q$ задержек (см. табл. 3.2).

В общем поведение автокорреляционной функции процесса авторегрессии похоже на поведение частной автокорреляционной функции процесса скользящего среднего и наоборот. Например, автокорреляционная функция процесса авторегрессии первого порядка экспоненциально затухает, в то время как частная автокорреляционная функция обрывается после первой задержки. Соответственно для процесса скользящего среднего первого порядка автокорреляционная функция обрывается после первой задержки. Частная автокорреляционная функция — не точно экспоненциальная, но в ней преобладают экспоненциальные члены, определяющие ее общий вид.

Особенно важны процессы авторегрессии и скользящего среднего первого и второго порядков и простой смешанный процесс $(1, d, 1)$. Свойства теоретических автокорреляционной и частной автокорреляционной функций этих процессов сведены в табл. 6.1, которая требует внимательного изучения и является удобной справочной таблицей. Читатель должен также обращаться к рис. 3.2, 3.9 и 3.11, на которых показано типичное поведение автокорреляционной и частной автокорреляционной функций для процессов авторегрессии и скользящего среднего второго порядка и простого смешанного процесса $(1, d, 1)$.

Соотношение между выборочными и теоретическими автокорреляциями. Выборочные автокорреляции могут иметь довольно большие дисперсии и быть сильно коррелированы друг с другом. По этой причине, как указывал Кендалл [29], нельзя ожидать *детального* сходства выборочной автокорреляционной функции с теоретической. В частности, умеренно большие значения выборочной автокорреляции могут наблюдаться и после того, как теоретическая автокорреляционная функция затухла; в выборочной функции могут наблюдаться всплески и тренды, не имеющиеся в теоретической функции. При использовании выборочной автокорреляционной функции для идентификации обычно можно быть уверенным в ее главных характеристиках,

Таблица 6.1

**Поведение автокорреляционных функций для d -й разности
процесса АРСС (p, d, q)**

[Таблица А и диаграммы В, С и D, облегчающие вычисление приближенных оценок параметров процессов скользящего среднего первого и второго порядка, процесса авторегрессии второго порядка и смешанного процесса АРСС (1,1), приведены в конце данного выпуска.]

Порядок	(1, d , 0)	(0, d , 1)
Поведение ρ_k	Экспоненциально затухают	Только $\rho_1 \neq 0$
Поведение ϕ_{kk}	Только $\phi_{11} \neq 0$	Доминирует экспоненциальное затухание
Предварительные оценки	$\phi_1 = \rho_1$	$\rho_1 = \frac{-\theta_1}{1 + \theta_1^2}$
Допустимый диапазон	$-1 < \phi_1 < 1$	$-1 < \theta_1 < 1$
Порядок	(2, d , 0)	(0, d , 2)
Поведение ρ_k	Наложение экспонент и затухающих синусоид	Только $\rho_1 \neq 0, \rho_2 \neq 0$
Поведение ϕ_{kk}	Только $\phi_{11} \neq 0, \phi_{22} \neq 0$	Доминирует наложение экспонент и затухающих синусоид
Предварительные оценки	$\phi_1 = \frac{\rho_1(1 - \rho_2)}{1 - \rho_1^2}$ $\phi_2 = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$	$\rho_1 = \frac{-\theta_1(1 - \theta_2)}{1 + \theta_1^2 + \theta_2^2}$ $\rho_2 = \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2}$
Допустимый диапазон	$-1 < \phi_2 < 1$ $\phi_2 + \phi_1 < 1$ $\phi_2 - \phi_1 < 1$	$-1 < \theta_2 < 1$ $\theta_2 + \theta_1 < 1$ $\theta_2 - \theta_1 < 1$
Порядок	1, d , 1	
Поведение ρ_k	Экспоненциально затухают, начиная с первой задержки	
Поведение ϕ_{kk}	С первой задержки доминирует экспоненциальное затухание	
Предварительные оценки	$\rho_1 = \frac{(1 - \theta_1\phi_1)(\phi_1 - \theta_1)}{1 + \theta_1^2 - 2\phi_1\theta_1}$ $\rho_2 = \rho_1\phi_1$	
Допустимый диапазон	$-1 < \phi_1 < 1$ $-1 < \theta_1 < 1$	

а более тонкие черты этой функции могут и не отражать реальных эффектов. Поэтому может понадобиться подобрать и исследовать далее на этапах оценивания и диагностической проверки две или даже большее число возможных моделей процесса.

6.2.2. Стандартные ошибки выборочных автокорреляций и частных автокорреляций

Поскольку мы не знаем теоретических автокорреляций и выборочные значения, которые мы вычисляем, несколько отличны от соответствующих теоретических, важно иметь некоторые представления о том, каковы могут быть эти отличия. В частности, необходимо уметь оценить, становятся ли автокорреляции и частные автокорреляции практически нулевыми при задержках больше некоторого k . Для больших задержек мы можем вычислить стандартные ошибки оцениваемых автокорреляций, исходя из упрощенной формулы Бартлетта (2.1.13), в которой теоретические автокорреляции заменены выборочными оценками. Имеем

$$\hat{\sigma}[r_k] \approx \frac{1}{n^{1/2}} \{1 + 2(r_1^2 + r_2^2 + \dots + r_q^2)\}^{1/2}, \quad k > q. \quad (6.2.2)$$

Для частных автокорреляций мы воспользуемся результатом, выраженным формулой (3.2.35), а именно при гипотезе, что процесс — авторегрессия порядка p , стандартная ошибка частной автокорреляции порядка $p + 1$ и выше равна

$$\hat{\sigma}[\hat{\phi}_{kk}] \approx \frac{1}{n^{1/2}} \quad k > p. \quad (6.2.3)$$

Андерсон показал [52], что даже для не слишком больших n выборочный коэффициент автокорреляции распределен примерно нормально со средним значением нуль. Отсюда при гипотезе, что теоретическая автокорреляция равна нулю, оценка r_k , деленная на ее стандартную ошибку, также будет распределена примерно по нормальному закону с единичной дисперсией. То же справедливо и для частных автокорреляций. Эти факты можно использовать для получения нестрогих правил проверки, являются ли теоретические автокорреляции и частные автокорреляции при задержках, больших некоторой, практически нулевыми. Обычно достаточно вспомнить, что для нормального распределения отклонения, превышающие одну стандартную ошибку в любом направлении, имеют вероятность около $1/3$, в то время как отклонения, превышающие две стандартные ошибки, имеют вероятность около $1/20$.

6.2.3. Идентификация некоторых реальных временных рядов

В этом разделе описанные выше способы идентификации применяются к шести временным рядам, обозначенным А, В,, F. Ряды А — D показаны на рис. 4.1, ряд Е — на рис. 6.1 и ряд F — на рис. 2.1. Данные для всех этих рядов приведены в сборнике временных рядов в конце этой книги. Ряды А, В, С и D были описаны в гл. 4 и ряд F — в гл. 2. Ряд Е — ряд чисел Вольфа солнечных пятен, каждый его член — это среднегодовое

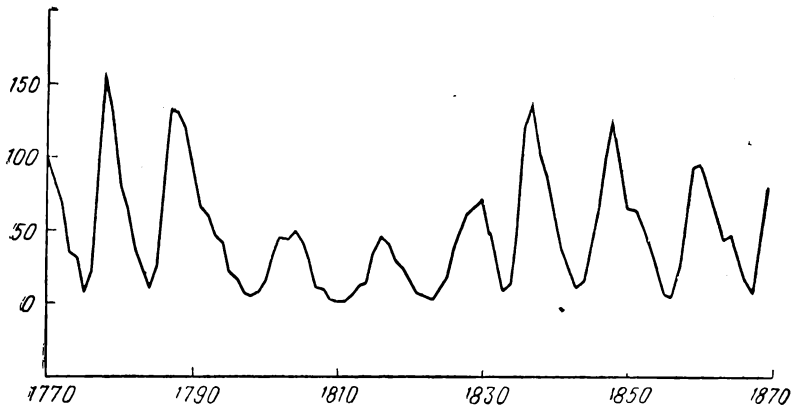


Рис. 6.1. Ряд Е — числа солнечных пятен по Вольфу (1770—1869).

число солнечных пятен за какой-либо год. Как отмечалось в гл. 4, следует ожидать, что ряды А, С и D обладают нестационарными свойствами, так как они представляют собой «нерегулируемое» поведение определенных выходов продукта.

Аналогично можно ожидать, что ряд биржевых цен В не имеет фиксированного уровня. С другой стороны, ряд F скорее всего стационарный, так как он представляет собой вариации выхода партий продукта, получаемых при равномерно регулируемых условиях. Подобным образом можно ожидать, что ряды солнечных пятен равновесны на больших интервалах.

Выборочные автокорреляции z , ∇z и $\nabla^2 z$ для рядов А — F приведены в табл. 6.2. В табл. 6.3 показаны соответствующие выборочные частные автокорреляции. Графическое представление корреляционных функций очень помогает пониманию их поведения; автокорреляции и частные автокорреляции ряда А показаны на рис. 6.2 и 6.3, ряда С — на рис. 6.4 и 6.5.

Для ряда А автокорреляции ∇z малы уже после первой задержки. Это указывает на то, что этот временной ряд может быть описан как процесс ПСС(0, 1, 1). Однако рассмотрение

автокорреляционной функции z показывает, что *при задержке больше 1* корреляции спадают довольно монотонно. Следовательно, имеется и другая возможность — что ряд описывается смешанным процессом АРСС(1, 0, 1). Частная автокорреляционная функция z ведет себя соответствующим этому предположению образом. Позднее мы увидим, что оба предположения приводят по существу к одной и той же модели.

Автокорреляции ряда S , показанные на рис. 6.4, свидетельствуют о необходимости по меньшей мере одного взятия разностей. Приближенно экспоненциальный спад корреляций первой

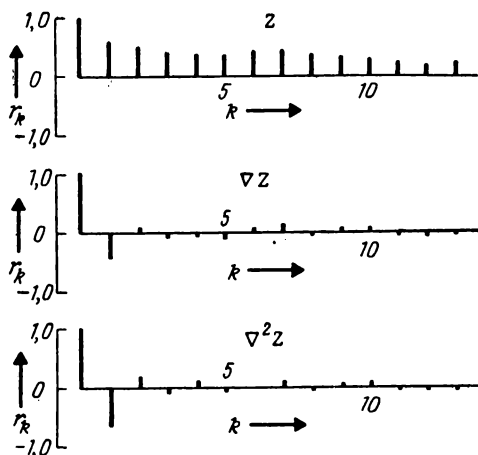


Рис. 6.2. Выборочные автокорреляции различных разностей ряда А.

разности указывает, что процесс имеет порядок (1, 1, 0) с параметром авторегрессии $\phi \approx 0,8$. С другой стороны, из малости автокорреляций $\nabla^2 z$ вытекает, что приемлем и процесс ПСС(0, 2, 0). К тем же выводам приводит рассмотрение частной автокорреляционной функции. Для отсчетов z приходим к обобщенному оператору авторегрессии второй степени; для ∇z — к оператору авторегрессии первого порядка, действующему на ∇z_t с $\phi \approx 0,8$; для $\nabla^2 z_t$ — к некоррелированному шуму. Отсюда следуют две возможности:

$$(1 - 0,8B)(1 - B)z_t = a_t,$$

$$(1 - B)^2 z_t = a_t.$$

Вторая модель очень похожа на первую, differing только выбором 0,8 вместо 1 для одного из коэффициентов авторегрессии.

При оценке свойств выборочных корреляционных функций весьма полезно проводить «контрольные» линии, параллельные

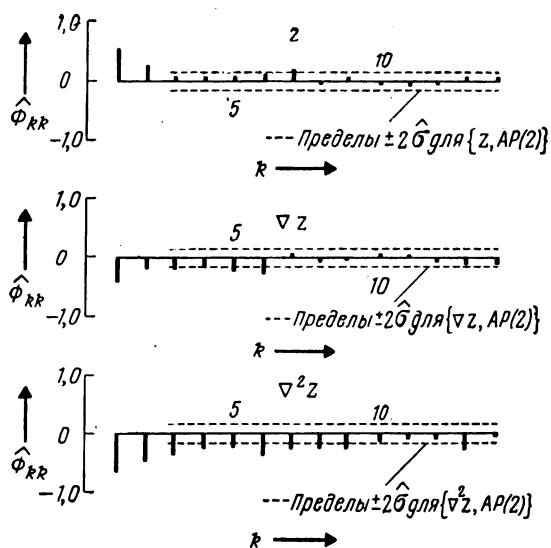


Рис. 6.3. Выборочные частные автокорреляции ряда A и его двух первых разностей.

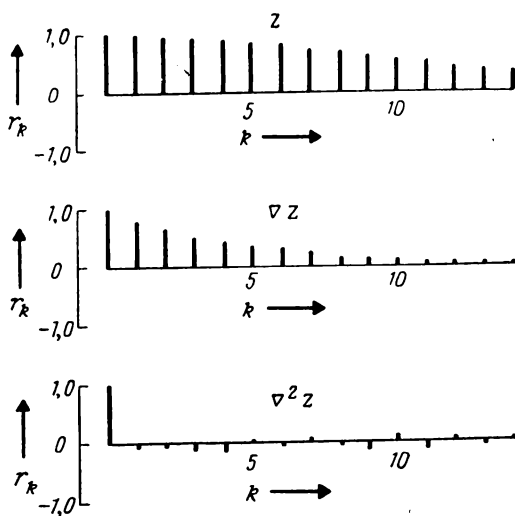


Рис. 6.4. Выборочные автокорреляции ряда C и его двух первых разностей.

Выборочные автокорреляции рядов А—F

Ряд А. Отсчеты концентраций в химическом процессе, каждые 2 часа

Автокорреляции

197 наблюдений

		1	2	3	4	5	6	7	8	9	10
z	Задержки	1—10	0,57	0,50	0,40	0,36	0,33	0,35	0,39	0,32	0,30
		11—20	0,19	0,16	0,20	0,24	0,14	0,18	0,20	0,20	0,18
Vz	Задержки	1—10	-0,41	0,02	-0,07	-0,01	-0,07	-0,02	0,15	-0,07	0,04
		11—20	-0,05	-0,06	-0,01	0,16	-0,17	0,03	0,01	0,08	-0,12
V^2z	Задержки	1—10	-0,65	0,18	-0,04	0,04	-0,04	-0,04	0,13	-0,11	0,04
		11—20	-0,02	-0,02	-0,04	0,18	-0,19	0,08	-0,03	0,09	-0,17
											0,26
											0,18
											0,02
											0,15
											0,02
											0,20

Ряд В. Цены акций IBM к закрытию биржи, ежедневно с 17 мая 1961 г. по 2 ноября 1962 г.

Автокорреляции

369 наблюдений

z	Задержки	1—10	0,99	0,99	0,98	0,97	0,96	0,95	0,94	0,93	0,92
		11—20	0,91	0,91	0,90	0,89	0,87	0,86	0,85	0,84	0,83
Vz	Задержки	1—10	0,09	0,09	-0,05	-0,04	0,12	0,07	0,04	-0,07	0,02
		11—20	0,08	0,05	-0,05	0,07	0,12	0,12	0,05	0,05	0,07
V^2z	Задержки	1—10	-0,45	-0,02	-0,04	0,00	-0,07	-0,01	0,04	-0,10	0,02
		11—20	0,04	0,04	-0,12	0,13	0,10	0,05	-0,04	-0,01	0,09

Ряд С. Отсчеты температур в химическом процессе, каждую минуту

Автокорреляции

226 наблюдений

z	Задержки	1—10	0,98	0,94	0,90	0,85	0,80	0,75	0,69	0,64	0,58	0,52
∇z	Задержки	11—20	0,47	0,41	0,36	0,30	0,25	0,20	0,15	0,10	0,05	0,00
	Задержки	1—10	0,80	0,65	0,53	0,44	0,38	0,32	0,26	0,19	0,14	0,14
	Задержки	11—20	0,10	0,09	0,07	0,07	0,07	0,07	0,09	0,05	0,04	0,04
$\nabla^2 z$	Задержки	1—10	—0,08	—0,07	—0,12	—0,06	0,01	—0,02	0,05	—0,05	—0,12	0,12
	Задержки	11—20	—0,12	0,07	—0,08	0,03	—0,01	—0,06	0,17	—0,10	—0,01	—0,02

Ряд D. Отсчеты вязкости химического процесса, каждый час

Автокорреляции

310 наблюдений

z	Задержки	1—10	0,86	0,74	0,62	0,53	0,46	0,41	0,35	0,31	0,27	0,24
∇z	Задержки	11—20	0,22	0,20	0,18	0,15	0,14	0,13	0,16	0,19	0,21	0,23
	Задержки	1—10	—0,05	—0,06	—0,07	—0,08	—0,06	0,00	—0,02	—0,02	—0,03	—0,06
	Задержки	11—20	—0,01	0,04	0,02	—0,07	—0,03	—0,09	—0,02	0,05	—0,01	0,06
$\nabla^2 z$	Задержки	1—10	—0,50	0,00	0,00	—0,01	—0,02	0,04	—0,01	0,00	0,01	—0,04
	Задержки	11—20	0,00	0,04	0,03	—0,06	0,04	—0,06	0,00	0,06	—0,06	0,06

Ряд Е. Числа солнечных пятен по Вольфу, каждый год

Автокорреляции

100 наблюдений

z	Задержки	1—10	0,81	0,43	0,07	-0,17	-0,27	-0,21	-0,04	0,16	0,33	0,41
∇z	Задержки	11—20	0,39	0,29	0,14	0,02	-0,06	-0,10	-0,14	-0,18	-0,17	-0,10
$\nabla^2 z$	Задержки	1—10	0,55	-0,02	-0,30	-0,40	-0,40	-0,33	-0,20	0,04	0,26	0,31
	Задержки	11—20	0,29	0,16	-0,03	-0,12	-0,10	-0,09	-0,09	-0,12	-0,14	-0,05
	Задержки	1—10	0,15	-0,31	-0,20	-0,11	-0,09	-0,02	-0,11	-0,04	0,19	0,05
	Задержки	11—20	0,13	0,09	-0,10	-0,11	0,04	0,01	0,00	-0,03	-0,10	-0,04

Ряд F. Выход продукта циклического химического процесса

Автокорреляции

70 наблюдений

z	Задержки	1—10	-0,39	0,30	-0,17	0,07	-0,10	-0,05	0,04	-0,04	-0,01	0,01
∇z	Задержки	11—20	0,11	-0,07	0,15	0,04	-0,01	0,17	-0,11	0,02	-0,05	0,02
$\nabla^2 z$	Задержки	1—10	-0,74	0,43	-0,27	0,16	-0,10	0,01	0,05	-0,05	0,04	-0,05
	Задержки	11—20	0,11	-0,16	0,12	-0,01	-0,08	0,16	-0,14	0,08	-0,07	0,03
	Задержки	1—10	-0,83	0,54	-0,33	0,21	-0,12	0,03	0,04	-0,06	0,06	-0,07
	Задержки	11—20	0,12	-0,16	0,11	0,00	-0,10	0,16	-0,15	0,10	-0,07	0,03

Выборочные частные автокорреляции*) рядов А — F

Ряд А. Отсчеты концентраций в химическом процессе, каждые 2 часа

Частные автокорреляции

197 наблюдений

z	Задержки										
		1—10	0,57	0,25	0,08	0,09	0,07	0,15	0,19	—0,03	0,01
Vz	Задержки	11—20	—0,09	—0,04	0,04	0,08	—0,15	0,06	0,13	0,09	—0,06
		1—10	—0,41	—0,19	—0,17	—0,14	—0,20	—0,23	0,01	—0,04	—0,01
V ² z	Задержки	11—20	0,02	—0,07	—0,10	0,13	—0,09	—0,15	—0,11	0,04	—0,08
		1—10	—0,66	—0,43	—0,33	—0,23	—0,20	—0,36	—0,23	—0,21	—0,23
		11—20	—0,07	—0,04	—0,25	0,00	0,04	—0,02	—0,16	—0,03	—0,22

Ряд В. Цены акций IBM к закрытию биржи, ежедневно с 17 мая 1961 г. по 2 ноября 1962 г.

Частные автокорреляции

369 наблюдений

z	Задержки										
		1—10	0,996	—0,09	0,01	0,05	0,02	0,02	—0,12	—0,05	—0,02
Vz	Задержки	11—20	—0,05	—0,09	—0,03	0,07	—0,08	0,06	—0,14	—0,10	—0,01
		1—10	0,09	—0,01	—0,05	—0,03	—0,02	0,13	0,05	0,02	—0,06
V ² z	Задержки	11—20	0,09	0,03	—0,08	0,08	—0,06	0,14	0,10	0,00	0,07
		1—10	—0,45	—0,28	—0,24	—0,20	—0,29	—0,17	—0,13	—0,03	—0,14
		11—20	—0,09	0,02	—0,13	0,01	—0,19	—0,13	—0,03	—0,10	—0,10

* Получены подгонкой процессов авторегрессии растущего порядка методом наименьших квадратов.

Ряд С. Отсчеты температур в химическом процессе, каждую минуту

Частные автокорреляции

226 наблюдений

z	Задержки	1—10	0,99	—0,81	—0,03	—0,02	—0,10	—0,07	—0,01	—0,03	0,04	—0,04
∇z	Задержки	11—20	—0,15	0,10	—0,14	0,01	—0,10	—0,02	—0,07	—0,11	0,11	—0,13
$\nabla^2 z$	Задержки	1—10	0,81	—0,01	—0,01	0,06	0,03	—0,03	—0,01	—0,08	0,00	0,10
		11—20	—0,14	0,10	—0,05	0,05	0,02	0,06	0,06	—0,17	0,09	0,00
	Задержки	1—10	—0,08	—0,08	—0,14	—0,10	—0,03	—0,05	0,02	—0,06	—0,16	0,09
		11—20	—0,14	0,01	—0,09	—0,02	—0,05	—0,09	0,13	—0,13	—0,03	—0,05

Ряд D. Отсчеты вязкости в химическом процессе, каждый час

Частные автокорреляции

310 наблюдений

z	Задержки	1—10	0,86	—0,02	0,00	0,01	0,03	0,03	—0,02	0,01	0,00	0,01
∇z	Задержки	11—20	0,05	0,01	—0,04	—0,03	0,07	0,04	0,10	0,06	0,00	0,06
$\nabla^2 z$	Задержки	1—10	—0,05	—0,06	—0,07	—0,09	—0,08	—0,03	—0,05	—0,05	—0,05	—0,09
		11—20	—0,05	0,01	—0,01	—0,10	—0,07	—0,13	—0,09	—0,02	—0,08	0,00
	Задержки	1—10	—0,50	—0,32	—0,24	—0,20	—0,22	—0,16	—0,14	—0,11	—0,07	—0,12
		11—20	—0,15	—0,12	—0,02	—0,06	—0,01	—0,07	—0,12	—0,06	—0,13	—0,07

Ряд Е. Числа солнечных пятен по Вольфу, каждый год

Частные автокорреляции

100 наблюдений

z	Задержки	1—10	0,81	-0,71	0,21	-0,15	0,10	0,10	0,18	0,23	0,01	0,00
		11—20	0,14	-0,16	0,12	0,03	-0,08	-0,14	-0,06	-0,12	0,00	0,05
Vz	Задержки	1—10	0,57	-0,48	-0,06	-0,27	-0,22	-0,26	-0,29	-0,05	-0,02	-0,16
		11—20	0,13	-0,15	-0,04	0,06	0,12	0,02	0,07	-0,06	-0,09	-0,06
V ² z	Задержки	1—20	0,15	-0,35	-0,10	-0,21	-0,16	-0,17	-0,36	-0,26	-0,09	-0,33
		11—20	-0,02	-0,13	-0,20	-0,21	-0,10	-0,13	0,00	0,03	-0,01	-0,08

Ряд F. Выход продукта циклического химического процесса

Частные автокорреляции

70 наблюдений

z	Задержки	1—10	-0,40	0,19	0,01	-0,07	-0,07	-0,15	0,05	0,00	-0,10	0,05
		11—20	0,18	-0,05	0,09	0,18	0,01	0,43	0,01	-0,14	0,11	0,18
Vz	Задержки	1—10	-0,76	-0,32	-0,16	-0,16	-0,09	-0,24	-0,15	-0,06	-0,18	-0,28
		11—20	-0,02	-0,16	-0,24	-0,06	-0,44	-0,02	0,12	-0,12	-0,17	-0,24
V ² z	Задержки	1—10	-0,83	-0,52	-0,38	-0,33	-0,15	-0,24	-0,26	-0,14	-0,09	-0,31
		11—20	-0,12	-0,09	-0,26	0,08	-0,38	-0,39	-0,07	-0,05	-0,03	-0,30

оси абсцисс на уровнях $\pm\sigma$ или $\pm 2\sigma$. Их должно сопровождать указание имеющейся гипотезы. Так, на рис. 6.5 надпись «пределы $\pm 2\hat{\sigma}$ для $\{z, AP(2)\}$ » на предельных линиях показывает,

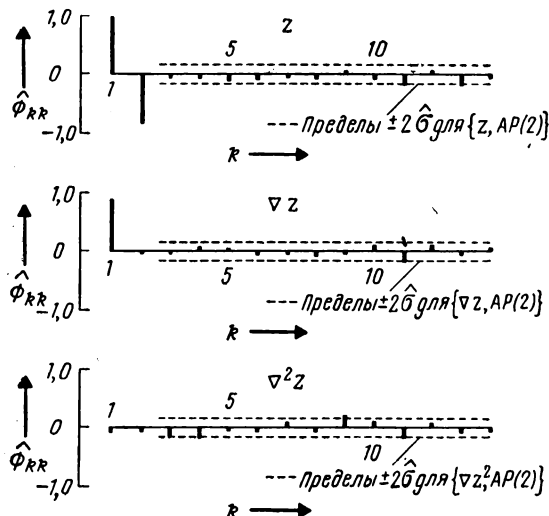


Рис. 6.5. Выборочные частные автокорреляции ряда S и его двух первых разностей.

что использованное значение $\hat{\sigma}$ приблизительно верно при гипотезе, что процесс — авторегрессия второго порядка. Пробная идентификация каждого из рядов $A - F$ дана в табл. 6.4.

Таблица 6.4

Пробная идентификация моделей рядов $A - F$

Ряд	Степень разности	Кажущаяся природа разности ряда	Идентификация для z_t
A	0	Смешанный АР первого порядка и СС первого порядка	(1, 0, 1)
	1	СС первого порядка	(0, 1, 1)
B	1	СС первого порядка	(0, 1, 1)
	1	АР первого порядка	(1, 1, 0)
C	2	Некоррелированный шум	(0, 2, 2) *
	0	АР первого порядка	(1, 0, 0)
D	1	Некоррелированный шум	(0, 1, 1) *
	0	АР второго порядка	(2, 0, 0)
E	0	АР третьего порядка	(3, 0, 0)
	0	АР второго порядка	(2, 0, 0)

*) Порядок оператора скользящего среднего, по-видимому, нулевой, но для дальнейшего обсуждения удобно сохранить более общие обозначения.

Нужно остановиться на трех других вопросах, связанных с процедурой такой идентификации.

1) Простое взятие разностей, которым мы пользуемся, не дает стационарности в рядах, содержащих сезонные компоненты. В гл. 9 мы обсудим подходящие модификации исследования таких рядов.

2) Как обсуждалось в гл. 4, ненулевое значение θ_0 в (6.1.1) указывает на существование систематического полиномиального тренда степени d . Для нестационарных моделей в табл. 6.4 значение $\theta_0 = 0$ вполне приемлемо для описания поведения ряда. Однако иногда могут существовать реальные физические явления, требующие введения ненулевого θ_0 . В других случаях может быть неясно, необходим ли такой член. Некоторые основанные на наблюдаемых данных указания о необходимости включения θ_0 в модель можно получить на этапе идентификации, сравнивая среднее значение \bar{w} ряда $w_t = \nabla^d z_t$ с его приближенной стандартной ошибкой, найденной по формуле $\sigma^2(\bar{w}) = = n^{-1} \sigma_w^2 [1 + 2r_1(w) + 2r_2(w) + \dots]$.

3) В разд. 3.4.2 отмечалось, что для любого процесса $APCC(p, q)$ с $p - q > 0$ вся положительная половина автокорреляционной функции будет суперпозицией затухающих синусоид и экспонент. Это, конечно, не мешает нам пробно идентифицировать q , так как а) частная автокорреляционная функция примет $p - q$ «аномальных» значений, прежде чем начнет вести себя как частная автокорреляция процесса $CC(q)$; б) q должно быть таким, что автокорреляционная функция могла бы, согласно общему правилу, принимать в качестве начальных значений величины от ρ_q до $\rho_{-(p-q-1)}$.

6.3. НАЧАЛЬНЫЕ ОЦЕНКИ ПАРАМЕТРОВ

6.3.1. Единственность оценок, получаемых по автоковариационной функции

Хотя справедливо, что данная модель $APCC$ имеет однозначно определенную автоковариационную функцию, обратное не справедливо. Сначала кажется, что это не позволяет использовать выборочные автоковариации как средства идентификации. Позднее, в разд. 6.4 будет показано, что выборочная автоковариационная функция может быть использована в этих целях. Причина в том, что, хотя существует множество моделей $APCC$, имеющих одну и ту же автоковариационную функцию, существует только одна модель, выражающая текущее значение $\nabla^d z_t = w_t$ через прошлое процесса в стационарной обратимой форме.

6.3.2. Начальные оценки для процессов скользящего среднего

Из формулы (3.3.4) следует, что первые q автокорреляций процесса $CC(q)$ не равны нулю и могут быть выражены через параметры модели

$$\rho_k = \frac{-\theta_k + \theta_1\theta_{k+1} + \theta_2\theta_{k+2} + \dots + \theta_{q-k}\theta_q}{(1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2)}, \quad k = 1, 2, \dots, q. \quad (6.3.1)$$

Выражение (6.3.1) для $\rho_1, \rho_2, \dots, \rho_q$ через $\theta_1, \theta_2, \dots, \theta_q$ дает q уравнений с q неизвестными. Предварительные оценки θ можно получить, подставив в (6.3.1) r_k вместо ρ_k и решив получающиеся нелинейные уравнения. Предварительную оценку σ_a^2 можно тогда получить из

$$\gamma_0 = \sigma_a^2(1 + \theta_1^2 + \dots + \theta_q^2),$$

заменяв θ их предварительными оценками и $\gamma_0 = \sigma^2$ его оценкой c_0 .

Предварительные оценки для процесса $(0, d, 1)$. Табл. А в конце этой книги связывает ρ_1 и θ_1 , и, заменив ρ_1 на $r_1(w)$, ее можно использовать для получения начальных оценок любого процесса $(0, d, 1)$: $w_t = (1 - \theta_1 B)a_t$, где $w_t = \nabla^d z_t$.

Предварительные оценки для процесса $(0, d, 2)$. Диаграмма С в конце книги связывает ρ_1 и ρ_2 с θ_1 и θ_2 и после подстановки $r_1(w)$ и $r_2(w)$ вместо ρ_1 и ρ_2 может быть использована для получения начальных оценок любого процесса $(0, d, 2)$.

При получении оценок этим способом необходимо помнить следующее.

1) Автоковариации являются вторыми моментами совместного распределения w . Отсюда, приравнивая выборочные моменты их теоретическим значениям, можно получать оценки параметров. Хорошо известно, что метод моментов не всегда эффективен, и можно продемонстрировать его малую эффективность в этих частных случаях. Однако грубые оценки этим способом могут оказаться полезными для получения более эффективных оценок, поскольку они дают приближенное представление о том, где в пространстве параметров «стоит искать» наиболее эффективные оценки.

2) В общем случае уравнения (6.3.1), полученные приравниванием моментов, имеют несколько решений. Например, для $q = 1$

$$\rho_1 = \frac{-\theta_1}{1 + \theta_1^2}, \quad (6.3.2)$$

и отсюда находим как

$$\theta_1 = -\frac{1}{2\rho_1} + \left\{ \frac{1}{(2\rho_1)^2} - 1 \right\}^{1/2},$$

так и

$$\theta_1 = -\frac{1}{2\rho_1} - \left\{ \frac{1}{(2\rho_1)^2} - 1 \right\}^{1/2} \quad (6.3.3)$$

как возможные решения. Далее, из табл. 6.2 следует, что автокорреляция с задержкой 1 для первой разности ряда А равна $\approx -0,4$. Подстановка этого значения в (6.3.3) дает пару решений $\theta_1 \approx 0,5$ и $\theta_1' \approx 2,0$. Однако выбранное значение $\theta_1 \approx 0,5$ — единственное, лежащее внутри интервала обратимости $-1 < \theta_1 < 1$. В разд. 6.4.1 показано, что только одно из возможных решений может удовлетворить условиям обратимости.

Примеры. Ряды А, В и D были идентифицированы в табл. 6.4 как возможные процессы ПСС(0, 1, 1). В разд. 4.3.1 было показано, что эта модель может быть описана различными способами

$$\begin{aligned} \nabla z_t &= (1 - \theta_1 B) a_t, \\ z_t &= \lambda_0 S a_{t-1} + a_t \quad (\lambda_0 = 1 - \theta_1), \\ z_t &= \lambda_0 \sum_{j=1}^{\infty} (1 - \lambda_0)^{j-1} z_{t-j} + a_t. \end{aligned}$$

Приближенные оценки параметров были получены при помощи табл. А в конце книги, они приведены в табл. 6.5.

Ряд С был пробно идентифицирован в табл. 6.4 как процесс ПСС(0, 2, 2)

$$\nabla^2 z_t = (1 - \theta_1 B - \theta_2 B^2) a_t,$$

или, что эквивалентно,

$$z_t = \lambda_0 S a_{t-1} + \lambda_1 S^2 a_{t-1} + a_t.$$

Так как первые две автокорреляции $\nabla^2 z_t$, данные в табл. 6.2,

Таблица 6.5

Начальные оценки параметров рядов А, В и D

Ряд	r_1	$\hat{\theta}_1$	$\hat{\lambda}_0 = 1 - \hat{\theta}_1$
А	-0,41	0,5	0,5
В	0,09	-0,1	1,1
D	-0,05	0,1	0,9

приближенно нулевые, то, используя диаграмму С в конце книги, находим $\hat{\theta}_1 = 0$, $\hat{\theta}_2 = 0$, $\hat{\lambda}_0 = 1 + \hat{\theta}_2 = 1$, $\hat{\lambda}_1 = 1 - \hat{\theta}_1 - \hat{\theta}_2 =$

= 1. Отсюда ряд можно представить как

$$\nabla^2 z_t = a_t$$

или

$$z_t = S a_{t-1} + S^2 a_{t-1} + a_t. \quad (6.3.4)$$

Это означает, что вторая разность $\nabla^2 z_t$ очень близка к чисто случайному ряду.

6.3.3. Начальные оценки для процессов авторегрессии

Если предположить, что исследуемый ряд — процесс авторегрессии первого или второго порядка, начальные оценки ϕ_1 и ϕ_2 можно получить, заменив теоретические автокорреляции ρ_j в формулах табл. 6.1 их выборочными оценками r_j , полученными из уравнений (3.2.6) Юла — Уокера. В частности, для процесса АР(1) $\hat{\phi}_{11} = r_1$ и для АР(2)

$$\begin{aligned} \hat{\phi}_{21} &= \frac{r_1(1-r_2)}{1-r_1^2}, \\ \hat{\phi}_{22} &= \frac{r_2-r_1^2}{1-r_1^2}, \end{aligned} \quad (6.3.5)$$

где ϕ_{pj} обозначает j -й авторегрессионный параметр процесса порядка p . Соответствующая формула, вытекающая из уравнений Юла — Уокера, для процессов высшего порядка может быть получена заменой ρ_j в (3.2.7) на r_j .

Отсюда

$$\hat{\phi} = R_p^{-1} r_p, \quad (6.3.6)$$

где R_p — выборочная корреляционная матрица размером $p \times p$, содержащая коэффициенты корреляции до порядка $p-1$, и r_p — вектор (r_1, r_2, \dots, r_p) . Например, если $p=3$, (6.3.6) имеет вид

$$\begin{bmatrix} \hat{\phi}_{31} \\ \hat{\phi}_{32} \\ \hat{\phi}_{33} \end{bmatrix} = \begin{bmatrix} 1 & r_1 & r_2 \\ r_1 & 1 & r_1 \\ r_2 & r_1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}. \quad (6.3.7)$$

Итеративный метод получения оценок для АР(p) по оценкам для АР($p-1$) был приведен в приложении ПЗ.2.

В гл. 7 будет показано, что в отличие от ситуации, возникающей с процессами СС, параметры авторегрессии, получаемые из (6.3.6), весьма близки к эффективным оценкам максимального правдоподобия.

Пример. Ряд E до взятия разностей *) похож на процесс авторегрессии второго порядка

$$(1 - \phi_1 B - \phi_2 B^2) \tilde{z}_t = a_t.$$

Подставляя выборочные оценки $r_1 = 0,81$ и $r_2 = 0,43$, полученные из табл. 6.2, из (6.3.5) находим $\hat{\phi}_1 = 1,34$ и $\hat{\phi}_2 = -0,66$. В качестве второго примера рассмотрим опять ряд C , идентифицированный либо как процесс $(1, 1, 0)$, либо как $(0, 2, 2)$. Первая модель дает

$$(1 - \phi_1 B) \nabla z_t = a_t$$

с $\hat{\phi}_1 \approx 0,8$, т. е. r_1 для ∇z равно 0,80.

Этот пример особенно интересен, потому что он ясно показывает, что две конкурирующие модели, идентифицированные для этого ряда, тесно связаны друг с другом. В предположении, что ряд — процесс порядка $(0, 2, 2)$, мы получили модель

$$(1 - B)(1 - B) z_t = a_t. \quad (6.3.8)$$

Альтернативная модель

$$(1 - 0,8B)(1 - B) z_t = a_t \quad (6.3.9)$$

очень похожа на (6.3.8).

6.3.4. Начальные оценки для смешанных процессов авторегрессии — скользящего среднего

В дальнейшем часто будет обнаруживаться, что либо сразу, либо после взятия нужного числа разностей ряд $w_t = \nabla^d z_t$ будет наиболее экономично описываться смешанным процессом АРСС

$$\phi(B) w_t = \theta(B) a_t.$$

В разд. 6.2.1 отмечалось, что указанием на смешанный процесс является тот факт, что автокорреляционная и частная автокорреляционная функции затухают. Другой факт, выражаемый формулами (3.4.3) и помогающий идентифицировать смешанный процесс, заключается в том, что после q задержек теоретические автокорреляции удовлетворяют разностному уравнению для чисто авторегрессионного процесса $\phi(B) w_t = a_t$.

*) Ряд солнечных пятен неоднократно исследовался (см., например, работы Шустера [53], Юла [24] и Морана [54]). Этот ряд почти наверняка не может быть адекватно описан процессом авторегрессии второго порядка. Модель, исходящая из реального физического механизма явления, была бы, конечно, наиболее подходящей. Недавно опубликованные работы указывают на то, что эмпирически модель авторегрессии второго порядка можно было бы лучше подогнать к ряду, если бы z подверглись сначала некоторому преобразованию.

В частности, если автокорреляционная функция d -й разности спадает экспоненциально (если не считать искажения в ρ_1), следует предположить, что процесс имеет порядок $(1, d, 1)$, т. е.

$$(1 - \phi_1 B) \omega_t = (1 - \theta_1 B) a_t, \quad (6.3.10)$$

где $\omega_t = \nabla^d z_t$.

Приближенные значения параметров процесса (6.3.10) получены подстановкой выборочных оценок $r_1(\omega)$ и $r_2(\omega)$ вместо ρ_1 и ρ_2 в выражения (3.4.8) и приведены в табл. 6.1. Здесь

$$r_1 = \frac{(1 - \hat{\theta}_1 \hat{\phi}_1)(\hat{\phi}_1 - \hat{\theta}_1)}{1 + \hat{\theta}_1^2 - 2\hat{\phi}_1 \hat{\theta}_1},$$

$$r_2 = r_1 \hat{\phi}_1.$$

Диаграмма D в конце этой книги связывает ρ_1 , ρ_2 и ϕ_1 , θ_1 и может быть использована для получения начальных оценок параметров любого процесса $(1, d, 1)$.

Например, рассмотрим опять ряд A, который был идентифицирован как процесс порядка $(0, 1, 1)$ с $\theta_1 = 0,5$. Рассматривая автокорреляционную функцию от z_t (а не от $\omega_t = \nabla z_t$, что менее удобно), мы видим, что начиная с r_2 автокорреляции затухают примерно экспоненциально, хотя и довольно медленно. Отсюда следует другая возможная идентификация ряда A как стационарного процесса порядка $(1, 0, 1)$. Выборочные автокорреляции и соответствующие начальные оценки параметров равны при этом:

$$r_1 = 0,57, \quad r_2 = 0,50, \quad \hat{\phi}_1 \approx 0,87, \quad \hat{\theta}_1 \approx 0,48.$$

Эта идентификация дает приближенную модель порядка $(1, 0, 1)$:

$$(1 - 0,9B) \tilde{z}_t = (1 - 0,5B) a_t,$$

в то время как ранее идентифицированная модель порядка $(0, 1, 1)$, приведенная в табл. 6.5, имела вид

$$(1 - B) z_t = (1 - 0,5B) a_t.$$

Мы видим, что опять альтернативные модели очень близки.

Более общий метод получения начальных оценок параметров для смешанного процесса авторегрессии — скользящего среднего приведен в приложении П6.2.

Взаимная компенсация операторов авторегрессии и скользящего среднего. Альтернативные модели, идентифицированные выше, на самом деле еще более похожи, чем это кажется. Это вызвано тем, что малые изменения оператора авторегрессии смешанной модели могут быть почти точно скомпенсированы соответствующими изменениями оператора скользящего среднего. В частности, если мы имеем модель

$$\{1 - (1 - \delta) B\} \tilde{z}_t = (1 - \theta B) a_t,$$

где δ мало и положительно, мы можем получить

$$\begin{aligned}(1 - B) \tilde{z}_t &= \{1 - (1 - \delta) B\}^{-1} (1 - B) (1 - \theta B) a_t = \\ &= \{1 - \delta B [1 + (1 - \delta) B + (1 - \delta)^2 B^2 + \dots]\} (1 - \theta B) a_t = \\ &= \{1 - (\theta + \delta) B\} a_t + \text{члены с } a_{t-2}, a_{t-3}, \dots \text{ порядка } \delta.\end{aligned}$$

6.3.5. Выбор между стационарными и нестационарными моделями в спорных случаях

Произвол в идентификации рядов A, C, D в табл. 6.4 скорее кажущийся, чем фактический. Он возникает во всех случаях, когда корни $\phi(B) = 0$ приближаются к единице. В этих ситуациях не имеет существенного значения, включен ли корень, близкий к единице, в $\phi(B)$ или добавлена еще одна разность, соответствующая единичному корню. Более точная оценка проводится при помощи оценочных процедур, описанных в гл. 7; однако необходимо помнить следующее.

1) По временным рядам неизбежно ограниченной длины никогда не удастся *доказать*, что корень оператора авторегрессии точно равен единице.

2) Не существует, конечно, резкого перехода от стационарного поведения ряда к нестационарному. Это можно понять, рассматривая поведение простой смешанной модели

$$(1 - \phi_1 B)(z_t - \mu) = (1 - \theta_1 B) a_t.$$

Ряды, генерируемые такими моделями, ведут себя все более нестационарно по мере увеличения ϕ_1 до единицы. Например, ряд с $\phi_1 = 0,99$ может отходить от своего среднего μ и не возвращаться к нему на очень длительные периоды времени. Это выглядит так, как будто «притяжение», действующее на члены ряда со стороны среднего значения, становится по мере приближения ϕ_1 к 1 все меньше и меньше, и, наконец, когда $\phi_1 = 1$, μ полностью перестает влиять на поведение ряда.

В спорных случаях некоторыми преимуществами может обладать использование нестационарных моделей (например, если оценка ϕ_1 близка к единице, принятие ϕ_1 *равным* единице). Это, в частности, верно в задачах прогнозирования и регулирования. Когда ϕ_1 близко к единице, мы фактически не знаем, имеет ли среднее значение ряда какой-либо смысл или нет. Следовательно, более удобно использовать нестационарную модель, которая не содержит среднего значения μ . Если мы будем использовать такую модель, прогнозы будущего поведения не будут ни в какой степени зависеть от выборочного среднего, найденного по предыдущему периоду, которое может не иметь никакого отношения к будущему уровню ряда.

Для сравнения с более эффективными методами оценивания, описанными в гл. 7, интересно понять, какая дополнительная информация о модели может быть извлечена на этапе идентификации. Мы уже показали, как находить начальные оценки параметров $(\hat{\phi}, \hat{\theta})$ в модели АРСС, идентифицированной для соответствующего разностного ряда $w_t = \nabla^d z_t$. Чтобы завершить описание, мы покажем теперь, как получить предварительные оценки остаточной дисперсии σ_a^2 и, если процесс стационарен, предварительную оценку постоянного члена в модели.

6.3.6. Начальные оценки остаточной дисперсии

Начальная оценка остаточной дисперсии может быть получена подстановкой выборочной оценки c_0 вместо γ_0 в выражения из гл. 3. Так, подстановкой в (3.2.8) можно получить начальную оценку σ_a^2 для процесса АР:

$$\hat{\sigma}_a^2 = c_0 (1 - \hat{\phi}_1 r_1 - \hat{\phi}_2 r_2 - \dots - \hat{\phi}_p r_p). \quad (6.3.11)$$

Аналогично из (3.3.3) получаем начальную оценку для процесса СС

$$\hat{\sigma}_a^2 = \frac{c_0}{1 + \hat{\theta}_1^2 + \dots + \hat{\theta}_q^2}. \quad (6.3.12)$$

Выражение для оценки смешанного процесса более сложно, наиболее простой способ ее получения описан в приложении Пб.2. Для практически важного процесса АРСС(1, 1) она принимает вид (см. 3.4.7)

$$\hat{\sigma}_a^2 = \frac{1 - \hat{\phi}_1^2}{1 + \hat{\theta}_1^2 - 2\hat{\phi}_1\hat{\theta}_1} c_0. \quad (6.3.13)$$

Например, рассмотрим модель (1, 0, 1), идентифицированную для ряда А. Используя (6.3.13) с $\hat{\phi}_1 = 0,87$; $\hat{\theta}_1 = 0,48$ и $c_0 = 0,1586$, получаем выборочную оценку

$$\hat{\sigma}_a^2 = \frac{1 - (0,87)^2}{1 + (0,48)^2 - 2(0,87)(0,48)} 0,1586 = 0,098.$$

6.3.7. Приближенная стандартная ошибка для \bar{w}

Общий процесс АРПСС, для которого среднее значение μ_w ряда $w_t = \nabla^d z_t$ не обязательно равняется нулю, может быть записан любым из трех способов:

$$\phi(B)(w_t - \mu_w) = \theta(B)a_t, \quad (6.3.14)$$

$$\phi(B)w_t = \theta_0 + \theta(B)a_t, \quad (6.3.15)$$

$$\phi(B)w_t = \theta(B)(a_t + \xi), \quad (6.3.16)$$

где

$$\mu_w = \frac{\theta_0}{1 - \phi_1 - \phi_2 - \dots - \phi_p} = \frac{(1 - \theta_1 - \theta_2 - \dots - \theta_q) \xi}{1 - \phi_1 - \phi_2 - \dots - \phi_p}.$$

Отсюда, если $1 - \phi_1 - \phi_2 - \dots - \phi_p \neq 0$ и $1 - \theta_1 - \theta_2 - \dots - \theta_q \neq 0$, из $\mu_w = 0$ следует, что $\theta_0 = 0$ и что $\xi = 0$. В общем случае, когда $d=0$, μ_w может быть и не равным нулю. Рассмотрим эвентуальную функцию прогноза, связанную с общей моделью (6.3.14), в случае, когда $d > 0$. При $\mu_w = 0$ эта функция прогноза уже содержит настраиваемую полиномиальную компоненту степени $d-1$. Когда допускается ненулевое μ_w , в эту функцию вводится *фиксированный* полиномиальный член степени d . Например, если $d=2$ и $\mu \neq 0$, функция прогноза $\hat{z}_t(l)$ включает квадратичную компоненту от l , коэффициент которой фиксирован и не подстраивается к ряду. Поскольку такие модели при $d > 0$ часто неприемлемы, гипотеза, что $\mu_w = 0$, во многих случаях не будет противоречить данным. Поэтому мы предполагаем обычно $\mu_w = 0$, если не имеется каких-либо указаний на обратное.

На этом идентификационном этапе построения модели указание, нужно ли принять μ_w равным нулю или нет, можно получить из сравнения $\bar{w} = \sum w/n$ с его приближенной стандартной ошибкой. Если имеется $(n-d)$ разностей, то

$$\sigma^2(\bar{w}) = n^{-1} \gamma_0 \sum_{-\infty}^{+\infty} \rho_l = n^{-1} \sum_{-\infty}^{+\infty} \gamma_l,$$

т. е.

$$\sigma^2(\bar{w}) = n^{-1} \gamma(1), \quad (6.3.17)$$

где $\gamma(B)$ — это производящая функция автоковариации, определенная в (3.1.10), и $\gamma(1)$ — ее значение при $B = B^{-1} = 1$.

В качестве примера рассмотрим процесс $(1, d, 0)$:

$$(1 - \phi B)(w_t - \mu_w) = a_t$$

с $w_t = \nabla^d z_t$. Из (3.1.11) получаем

$$\gamma(B) = \frac{\sigma_a^2}{(1 - \phi B)(1 - \phi F)},$$

так что

$$\sigma^2(\bar{w}) = n^{-1} (1 - \phi)^{-2} \sigma_a^2.$$

Но $\sigma_a^2 = \sigma_w^2 (1 - \phi^2)$, так что

$$\sigma^2(\bar{w}) = \frac{\sigma_w^2}{n} \frac{1 - \phi^2}{(1 - \phi)^2} = \frac{\sigma_w^2}{n} \frac{1 + \phi}{1 - \phi}$$

и

$$\sigma(\bar{w}) = \sigma_w \left\{ \frac{1 + \phi}{n(1 - \phi)} \right\}^{1/2}.$$

Величины σ_w^2 и ϕ оцениваются соответственно по c_0 и r_1 , как определено формулами (2.1.9) и (2.1.10). Отсюда для процесса $(1, d, 0)$ искомая стандартная ошибка выражается как

$$\hat{\sigma}(\bar{w}) = \left\{ \frac{c_0(1+r_1)}{n(1-r_1)} \right\}^{1/2}.$$

Этим способом можно получить выражения для $\sigma(\bar{w})$, приведенные в табл. 6.6.

Таблица 6.6

Приближенная стандартная ошибка \bar{w} , где $w_t = \nabla^d z_t$ и z_t — процесс АРПСС (p, d, q)

$(1, d, 0)$ $\left\{ \frac{c_0(1+r_1)}{n(1-r_1)} \right\}^{1/2}$	$(0, d, 1)$ $\left\{ \frac{c_0(1+2r_1)}{n} \right\}^{1/2}$
$(2, d, 0)$ $\left\{ \frac{c_0(1+r_1)(1-2r_1^2+r_2)}{n(1-r_1)(1-r_2)} \right\}^{1/2}$	$(0, d, 2)$ $\left\{ \frac{c_0(1+2r_1+2r_2)}{n} \right\}^{1/2}$
$(1, d, 1)$ $\left\{ \frac{c_0}{n} \left(1 + \frac{2r_1^2}{r_1-r_2} \right) \right\}^{1/2}$	

Пробная идентификация моделей А — F. В табл. 6.7 дана сводка моделей, пробно идентифицированных по рядам А — F, вместе с предварительными оценками параметров. Эти модели используются как начальные приближения в более эффективных методах оценивания, описываемых в гл. 7.

6.4. МНОГОЗНАЧНОСТЬ МОДЕЛЕЙ

6.4.1. Многозначность моделей авторегрессии — скользящего среднего

В предположении о нормальности процесса знание первого и второго моментов распределения вероятностей дает полное знание распределения. В частности, если известны средние значения процесса $w_t = \nabla^d z_t$ и его автоковариационная функция, вероятностная структура w_t полностью определена. Покажем теперь, что хотя эта единственная вероятностная структура может быть представлена множеством линейных моделей, тем не менее един-

ственность модели *достигается* при введении соответствующих ограничений, обеспечивающих стационарность и обратимость.

Положим, что w_t имеет производящую функцию ковариаций $\gamma(B)$ и представимо линейной моделью

$$\phi(B) w_t = \theta(B) a_t, \quad (6.4.1)$$

где нули $\phi(B)$ и $\theta(B)$ лежат вне единичного круга. Тогда эта линейная модель может также быть представлена в виде

$$\prod_{i=1}^p (1 - G_i B) w_t = \prod_{j=1}^q (1 - H_j B) a_t, \quad (6.4.2)$$

где G_i^{-1} — корни $\phi(B) = 0$ и H_j^{-1} — корни $\theta(B) = 0$; G_i, H_j лежат вне единичного круга. Пользуясь (3.1.11), получим выражение для производящей функции ковариации w

$$\gamma(B) = \prod_{i=1}^p (1 - G_i B)^{-1} (1 - G_i F)^{-1} \prod_{j=1}^q (1 - H_j B) (1 - H_j F) \sigma_a^2.$$

Многозначность в выборе параметров скользящего среднего. Так как

$$(1 - H_j B)(1 - H_j F) = H_j^2 (1 - H_j^{-1} B)(1 - H_j^{-1} F),$$

отсюда следует, что любая из стохастических моделей

$$\prod_{i=1}^p (1 - G_i B) w_t = \prod_{j=1}^q (1 - H_j^{\pm 1} B) k a_t$$

может иметь ту же производящую функцию ковариации при надлежащем выборе константы k . В приведенных выше рассуждениях предполагалось, что для комплексных корней будут взяты обратные значения каждого из членов сопряженной пары. Однако, если действительный корень H лежит внутри единичного круга, H^{-1} будет лежать вне его, а если комплексная пара H_1 и H_2 лежит внутри, то пара H_1^{-1} и H_2^{-1} лежит вне единичного круга. Отсюда следует, что существует только *одна стационарная обратимая* модель вида (6.4.2), имеющая данную автоковариационную функцию.

Возвратное представление. Заметим, что $\gamma(B)$ не изменится, если в (6.4.2) заменить $(1 - G_i B)$ на $(1 - G_i F)$ или $(1 - H_j B)$ на $(1 - H_j F)$. Отсюда все стохастические модели

$$\prod_{i=1}^p (1 - G_i B^{\pm 1}) w_t = \prod_{j=1}^q (1 - H_j B^{\pm 1}) a_t$$

имеют идентичную ковариационную структуру. Однако представления моделей, содержащие оператор $B^{-1} = F$, обращаются

Таблица 6.7
Сводка данных о моделях, идентифицированных для рядов А—F, включающая начальные оценки параметров

Ряд	Степень разности	$\bar{w} \pm \hat{\sigma}(\bar{w})$	$\hat{\sigma}_w^2 = c_0^*$	Идентифицированная модель	$\hat{\sigma}_a^2$
А	0	$17,06 \pm 0,10$ $0,002 \pm 0,011$	0,1586 0,1364	$z_t - 0,87z_{t-1} = 2,45 + a_t - 0,48a_{t-1}$ $\nabla z_t = a_t - 0,53a_{t-1}$	0,098 0,107
	1				
В	1	$-0,28 \pm 0,41$	52,54	$\nabla z_t = a_t - 0,09a_{t-1}$	52,2
С	1	$-0,035 \pm 0,047$ $-0,003 \pm 0,008$	0,0532 0,0198	$\nabla z_t - 0,81\nabla z_{t-1} = a_t$ $\nabla^2 z_t = a_t - 0,09a_{t-1} - 0,07a_{t-2}$	0,019 0,020
	2				
D	0	$9,13 \pm 0,04$ $0,004 \pm 0,017$	0,3620 0,0965	$z_t - 0,86z_{t-1} = 1,32 + a_t$ $\nabla z_t = a_t - 0,05a_{t-1}$	0,093 0,096
	1				
E	0	$46,9 \pm 5,4$ $46,9 \pm 5,4$	1382,2 1382,2	$z_t - 1,32z_{t-1} + 0,63z_{t-2} = 14,9 + a_t$ $z_t - 1,37z_{t-1} + 0,74z_{t-2} - 0,08z_{t-3} = 13,7 + a_t$	289 287
	0				
F	0	$51,1 \pm 1,1$	139,8	$z_t + 0,32z_{t-1} - 0,18z_{t-2} = 58,3 + a_t$	115

*) Когда $d=0$, вместо w читай z .

к будущим w и (или) к будущим a . Следовательно, хотя существуют стационарные и обратимые представления, в которых w_t выражено через будущие w и a , имеется только одно представление (6.4.2), связывающее w_t целиком только с *прошлым* ряда.

Практически важную форму представления модели получим (хотя это и кажется несколько странным), заменив в (6.4.1) все B на F , т. е.

$$\phi(F) w_t = \theta(F) e_t,$$

где e_t — последовательность независимо распределенных случайных переменных, имеющих нулевое среднее значение и дисперсию $\sigma_e^2 = \sigma_a^2$. Это означает, что существует стационарное обратимое представление, в котором w_t выражено *целиком* через будущие w и e . Мы будем называть его *возвратным* представлением или просто *возвратным процессом*.

Уравнение (6.4.2) не является наиболее общей формой стационарной обратимой линейной модели, имеющей производящую функцию ковариации $\gamma(B)$. Например, обе части (6.4.2) могут быть умножены на любой множитель $(1 - QB)$. Отсюда процесс

$$(1 - QB) \prod_{i=1}^p (1 - G_i B) w_t = (1 - QB) \prod_{j=1}^q (1 - H_j B) a_t$$

имеет ту же ковариационную структуру, что и (6.4.2). Этот факт не создает особых трудностей при идентификации, так как мы, естественно, будем стремиться выбрать простейшее представление. Но в гл. 7 мы увидим, что нужно остерегаться возможности появления одинаковых множителей при подгонке процессов.

Итак, мы приходим к заключению, что стационарная обратимая модель, в которой текущее значение w_t выражено только через *прошлое*, и не содержащая одинаковых множителей, определяется своей ковариационной структурой единственным образом.

Четкое понимание многозначности модели важно по ряду причин.

1) Как следует из проведенных рассуждений, ковариационная функция может быть логически обоснованно использована для идентификации линейной стационарной обратимой модели, выражающей w_t через прошлое процесса.

2) Выяснена природа многозначности решений для параметров скользящего среднего, полученных приравниванием моментов.

3) Возвратный процесс

$$\phi(F) w_t = \theta(F) e_t,$$

полученный заменой B на F в линейной модели, полезен при оценке значения ряда в моменты времени, предшествующие первому наблюдению.

Рассмотрим теперь (2) и (3) более подробно.

6.4.2. Многозначные решения при оценке параметров скользящего среднего методом моментов

При оценке в разд. 3.3 q параметров $\theta_1, \theta_2, \dots, \theta_q$ в модели СС приравниванием ковариаций мы нашли, что решение неоднозначно. Каждой комбинации корней соответствует линейное представление, но только одна такая комбинация дает обратимое представление через прошлое процесса.

Например, рассмотрим процесс СС (1) для w_t

$$w_t = (1 - \theta_1 B) a_t$$

и положим, что $\gamma_0(w)$ и $\gamma_1(w)$ известны, и мы хотим найти θ_1 и σ_a^2 .

Так как

$$\gamma_0 = (1 + \theta_1^2) \sigma_a^2, \quad \gamma_1 = -\theta_1 \sigma_a^2, \quad \gamma_k = 0, \quad k > 1, \quad (6.4.3)$$

то

$$-\gamma_0/\gamma_1 = \theta_1^{-1} + \theta_1,$$

и если $(\theta_1 = 0, \sigma_a^2 = \sigma^2)$ — решение для данных γ_0 и γ_1 , то имеется решение $(\theta_1 = \theta^{-1}, \sigma_a^2 = \theta^2 \sigma^2)$. Тогда очевидно, что для данных γ_0 и γ_1 существует пара возможных процессов

$$w_t = (1 - \theta B) a_t$$

и

$$w_t = (1 - \theta^{-1} B) \alpha_t \quad (6.4.4)$$

с $\sigma_a^2 = \sigma^2 \theta^2$. Если $-1 < \theta < 1$, то (6.4.4) не является обратимым представлением.

Однако модель можно представить как

$$w_t = \{(1 - \theta^{-1} B)(-\theta F)\}(-\theta^{-1} B \alpha_t).$$

Тогда после подстановки

$$e_t = -\alpha_{t-1}/\theta$$

модель принимает вид

$$w_t = (1 - \theta F) e_t, \quad (6.4.5)$$

где e_t имеет ту же дисперсию, что и a_t . Поэтому (6.4.5) — это возвратный процесс, взаимный прямому процессу

$$w_t = (1 - \theta B) a_t. \quad (6.4.6)$$

Точно так же, как импульс a_t в (6.4.6) представим в виде сходящейся суммы текущего и *предыдущих* значений w

$$a_t = w_t + \theta w_{t-1} + \theta^2 w_{t-2} + \dots,$$

импульс e_t в (6.4.5) представим как сходящиеся суммы текущего и будущих значений w

$$e_t = w_t + \theta w_{t+1} + \theta^2 w_{t+2} + \dots$$

Следовательно, корень θ^{-1} *действительно создает* обратимый процесс, но только если допустимо представление импульса e_t через будущие значения w_t .

Область обратимости, показанная в табл. 6.1, ограничивает допустимые значения параметров *при условии*, что мы выражаем импульсы в терминах *прошлого*.

6.4.3. Использование возвратного процесса для определения начальных значений

Положим, что нам известен временной ряд w_1, w_2, \dots, w_n , представляющий процесс

$$\phi(B)w_t = \theta(B)a_t. \quad (6.4.7)$$

В гл. 7 мы встретимся с задачами, где понадобятся оценки значений w_0, w_{-1}, w_{-2} и т. д. ряда в моменты времени, *предшествовавшие* первому наблюдению. Эта необходимость обусловлена тем, что для определенных рекуррентных расчетов при оценках параметров модели нужны начальные значения. Пусть мы хотим оценить w_{-l} по данным w_1, \dots, w_n . Обсуждение в разд. 6.4.1 показало, что вероятностная структура w_1, \dots, w_n может быть объяснена как прямой моделью (6.4.7), так и возвратной моделью:

$$\phi(F)w_t = \theta(F)e_t. \quad (6.4.8)$$

Значение w_{-l} имеет точно такие же вероятностные соотношения с последовательностью w_1, w_2, \dots, w_n , как значение w_{n+l+1} с последовательностью $w_n, w_{n-1}, w_{n-2}, \dots, w_1$. Поэтому, чтобы оценить значение в момент времени, на $(l+1)$ опережающий начало наблюдений, мы можем сначала рассмотреть, какова будет оптимальная оценка или прогноз на $l+1$ шагов после окончания ряда, и затем применить эту процедуру к *обращенному ряду*. Другими словами, мы «прогнозируем» обращенные ряды. Мы называем это «прогнозированием назад».

ПРИЛОЖЕНИЕ П6.1. СРЕДНЕЕ ЗНАЧЕНИЕ ВЫБОРОЧНОЙ АВТОКОРРЕЛЯЦИОННОЙ ФУНКЦИИ НЕСТАЦИОНАРНОГО ПРОЦЕССА

Пусть ряд из N наблюдений z_1, z_2, \dots, z_N генерируется нестационарным процессом $(0, 1, 1)$

$$\nabla z_t = (1 - \theta B) a_t,$$

и вычисляются выборочные автокорреляции r_k , где

$$r_k = \frac{c_k}{c_0} = \frac{\sum_{t=1}^{N-k} (z_t - \bar{z})(z_{t+k} - \bar{z})}{\sum_{t=1}^N (z_t - \bar{z})^2}.$$

В предположении, что c распределены нормально, $r_k = c_k/c_0$ распределено независимо от c_0 , т. е. *)

$$E[c_k] = E[c_0 r_k] = E[c_0] E[r_k].$$

Следовательно,

$$E[r_k] = E[c_k]/E[c_0] = \frac{\sum_{t=1}^{N-k} E[(z_t - \bar{z})(z_{t+k} - \bar{z})]}{\sum_{t=1}^N E[(z_t - \bar{z})^2]}.$$

После очевидных, но громоздких алгебраических преобразований находим

$$E[r_k] = \frac{(N-k) \{ (1-\theta)^2 (N^2 - 1 + 2k^2 - 4kN) - 6\theta \}}{N(N-1) \{ (N+1)(1-\theta)^2 + 6\theta \}}. \quad (\text{П6.1.1})$$

Для θ , близких к нулю, $E[r_k]$ близко к единице, но для больших значений θ оно может быть значительно меньше единицы даже при малых k . Рис. П6.1 иллюстрирует этот факт; на нем показано $E[r_k]$ для $\theta = 0,8$ и $N = 100$ и 200 . Хотя, как и полагается для нестационарного процесса, математические ожидания не затухают достаточно быстро, видно, что они не приближаются к единице даже при очень малых задержках.

Подобный эффект может быть замечен всегда, когда параметры приближаются к значениям, при которых сокращение одинаковых множителей в обеих частях уравнения, описывающего модель, приводит к стационарному процессу. Так, в примере, приведенном выше, мы можем представить модель как

$$(1 - B) z_t = \{(1 - B) + \delta B\} a_t,$$

*) Это предположение является приближенным. В результате формулу (П6.1.1) и последующие выводы нельзя считать вполне обоснованными. — *Прим. ред.*

где $\delta = 0,2$. Когда δ стремится к нулю, поведение процесса будет все больше походить на поведение белого шума $z_t = a_t$, для которого автокорреляционная функция равна нулю для задержек $k > 0$.

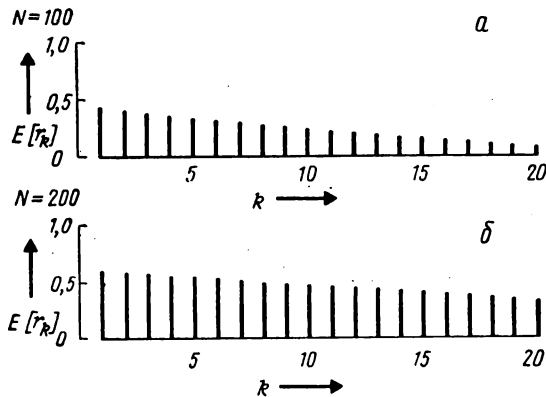


Рис. П6.1. Математические ожидания величин r_k для ряда, генерированного моделью $\nabla z_t = (1 - 0,8B)a_t$.

ПРИЛОЖЕНИЕ П6.2. ОБЩИЙ МЕТОД ПОЛУЧЕНИЯ НАЧАЛЬНЫХ ОЦЕНОК ПАРАМЕТРОВ СМЕШАННОГО ПРОЦЕССА АВТОРЕГРЕССИИ — СКОЛЬЗЯЩЕГО СРЕДНЕГО

В разд. 6.3 было показано, как получить начальные оценки для параметров простых моделей АРСС. В частности, в конце книги в сборнике таблиц и диаграмм приведены диаграммы В, С и D, позволяющие быстро получить начальные оценки для процессов АР(2), СС(2) и АРСС(1, 1). В настоящем приложении будет описан пригодный для программирования общий метод получения начальных оценок процесса АРСС(p, q): $\phi(B)w_t = \theta(B)a_t$. Соответствующая программа для ЭВМ описана под заголовком «Программа 2» в сборнике программ в конце этой книги.

В общем случае вычисление начальных оценок процесса АРСС(p, q) основано на первых $p + q + 1$ автоковариациях c_j [$j = 0, 1, \dots, (p + q)$] от $w_t = \nabla^d z_t$ и проводится в 3 этапа.

1) Параметры авторегрессии $\phi_1, \phi_2, \dots, \phi_p$ оцениваются по автоковариациям $c_{q-p+1}, \dots, c_{q+1}, c_{q+2}, \dots, c_{q+p}$.

2) На базе оценок $\hat{\phi}$, найденных в (1), вычисляются первые $q + 1$ автоковариации c'_j ($j = 0, 1, \dots, q$) полученного ряда

$$w'_t = w_t - \hat{\phi}_1 w_{t-1} - \dots - \hat{\phi}_p w_{t-p}.$$

3) Наконец, автоковариации c'_0, c'_1, \dots, c'_q используются при итеративном расчете начальных оценок параметров скользящего среднего $\theta_1, \theta_2, \dots, \theta_q$ и остаточной дисперсии σ_a^2 .

Начальные оценки параметров авторегрессии. Пользуясь результатом (3.4.3), можно получить начальные оценки параметров авторегрессии, решив p линейных уравнений

$$\begin{aligned} c_{q+1} &= \hat{\phi}_1 c_q + \hat{\phi}_2 c_{q-1} + \dots + \hat{\phi}_p c_{q-p+1}, \\ c_{q+2} &= \hat{\phi}_1 c_{q+1} + \hat{\phi}_2 c_q + \dots + \hat{\phi}_p c_{q-p+2}, \\ c_{q+p} &= \hat{\phi}_1 c_{q+p-1} + \hat{\phi}_2 c_{q+p-2} + \dots + \hat{\phi}_p c_q. \end{aligned} \quad (\text{П6.2.1})$$

Автоковариации найденного процесса скользящего среднего. Обозначим теперь $w'_t = \phi(B)w_t$ и будем анализировать этот процесс как процесс скользящего среднего

$$w'_t = \theta(B) a_t. \quad (\text{П6.2.2})$$

Прежде всего необходимо выразить автоковариации c'_j процесса w'_t через автоковариации c_j процесса w_t . Можно показать, что

$$c'_j = \sum_{i=0}^p \phi_i^2 c_j + \sum_{i=1}^p (\phi_0 \phi_i + \phi_1 \phi_{i+1} + \dots + \phi_{p-i} \phi_p) d_j, \quad (\text{П6.2.3})$$

где

$$\begin{aligned} j &= 0, 1, \dots, q, \\ d_j &= c_{j+1} + c_{j-1}, \\ \phi_0 &= -1. \end{aligned}$$

Начальные оценки параметров скользящего среднего. Пользуясь оценками автоковариаций c'_j , можно получить начальные оценки параметров скользящего среднего в найденном процессе (П6.2.2) при помощи какого-либо одного из двух итеративных процессов.

1. *Линейно-сходящийся процесс.* Из выражений

$$\begin{aligned} \gamma_0 &= (1 + \theta_1^2 + \dots + \theta_q^2) \sigma_a^2, \\ \gamma_k &= (-\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q) \sigma_a^2, \quad k \geq 1, \end{aligned}$$

для автоковариационной функции процесса $CC(q)$, приведенных в разд. 3.3.2, можно найти оценки параметров $\sigma_a^2, \theta_q, \theta_{q-1}, \dots, \theta_1$ в том точно порядке, как здесь указано, при помощи итераций

$$\begin{aligned} \sigma_a^2 &= \frac{c'_0}{1 + \theta_1^2 + \dots + \theta_q^2}, \\ \theta_j &= - \left(\frac{c'_j}{\sigma_a^2} - \theta_j \theta_{j+1} - \theta_2 \theta_{j+2} - \dots - \theta_{q-j} \theta_q \right) \end{aligned} \quad (\text{П6.2.4})$$

с условием, что $\theta_0 = 0$. Параметры $\theta_1, \theta_2, \dots, \theta_q$ приравниваются нулю в начале итеративной процедуры; значения θ_j и σ_a^2 , используемые в любом цикле вычисления, — это последние из доступных оценок этих величин. Например, в случае $q = 2$ уравнения (П6.2.4) имеют вид

$$\begin{aligned}\sigma_a^2 &= \frac{c'_0}{1 + \theta_1^2 + \theta_2^2}, \\ \theta_2 &= -\frac{c'_2}{\sigma_a^2}, \\ \theta_1 &= -\left(\frac{c'_1}{\sigma_a^2} - \theta_1\theta_2\right).\end{aligned}$$

2. *Квадратически сходящийся процесс.* Алгоритм Ньютона — Рафсона, обладающий более быстрой сходимостью, чем метод (1), был предложен Вилсоном [55]. Обозначим $\tau = (\tau_0, \tau_1, \dots, \tau_q)$, где

$$\tau_0^2 = \sigma_a^2, \quad \theta_j = -\tau_j/\tau_0, \quad j = 1, 2, \dots, q. \quad (\text{П6.2.5})$$

Тогда, если τ^i — оценка τ , полученная в результате i -й итерации, новые значения в результате $(i+1)$ -й итерации будут получены из формулы

$$\tau^{i+1} = \tau^i - (\mathbf{T}^i)^{-1} \mathbf{f}_i, \quad (\text{П6.2.6})$$

где

$$\mathbf{f}' = (f_0, f_1, \dots, f_q), \quad f_j = \sum_{i=0}^{q-j} \tau_i \tau_{i+j} - c'_j$$

и

$$\mathbf{T} = \begin{bmatrix} \tau_0 & \tau_1 & \dots & \tau_{q-2} & \tau_{q-1} & \tau_q \\ \tau_1 & \tau_2 & \dots & \tau_{q-1} & \tau_q & 0 \\ \tau_2 & \tau_3 & \dots & \tau_q & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ \tau_q & 0 & \dots & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} \tau_0 & \tau_1 & \tau_2 & \dots & \tau_q \\ 0 & \tau_0 & \tau_1 & \dots & \tau_{q-1} \\ 0 & 0 & \tau_0 & \dots & \tau_{q-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{bmatrix}.$$

Имея значения τ для каждой итерации, можно получить значения параметров из (П6.2.5).

Пример. Рассмотрим оценивание ϕ и θ в модели АРСС

$$(1 - \phi B) \tilde{z}_t = (1 - \theta B) a_t,$$

используя значения $c_0 = 1,25$, $c_1 = 0,50$, $c_2 = 0,40$, соответствующие процессу с $\phi = 0,8$, $\theta = 0,5$ и $\sigma_a^2 = 1,000$. Оценка ϕ полу-

Таблица П6.1.

Сходимость предварительных оценок σ_a^2 и θ_1
для процесса СС (1)

Итерация	Метод (1)		Метод (2)	
	σ_a^2	θ	σ_a^2	θ
0	—	0,000	1,250	0,0000
1	1,250	0,400	2,250	0,667
2	1,077	0,464	1,210	0,545
3	1,029	0,486	1,012	0,503
4	1,011	0,494	1,000	0,500
5	1,004	0,498	—	—
6	1,002	0,499	—	—
7	1,001	0,500	—	—
8	1,000	0,500	—	—

чена подстановкой $p = q = 1$ в (П6.2.1), в результате которой

$$c_2 = \phi c_1,$$

так что $\hat{\phi} = 0,8$. Отсюда, пользуясь (П6.2.3), находим, что две первые ковариации ряда

$$w'_t = \tilde{z}_t - \phi \tilde{z}_{t-1}$$

равны

$$c'_0 = (1 + \phi^2) c_0 - 2\phi c_1 = 1,25,$$

$$c'_1 = (1 + \phi^2) c_1 - \phi (c_2 + c_0) = -0,50.$$

Подставляя эти значения в (П6.2.4), находим формулы, на которых основан итеративный процесс (1):

$$\sigma_a^2 = \frac{1,25}{1 + \theta^2},$$

$$\theta = 0,5/\sigma_a^2.$$

Аналогично, подставляя их в (П6.2.6), получаем формулы, на которых основан итеративный процесс (2):

$$\begin{bmatrix} \tau_0^{i+1} \\ \tau_1^{i+1} \end{bmatrix} = \begin{bmatrix} \tau_0^i \\ \tau_1^i \end{bmatrix} - (\mathbf{T}^i)^{-1} \begin{bmatrix} f_0^i \\ f_1^i \end{bmatrix},$$

где

$$\begin{bmatrix} f_0^i \\ f_1^i \end{bmatrix} = \begin{bmatrix} (\tau_0^i)^2 + (\tau_1^i)^2 - c'_0 \\ \tau_0^i \tau_1^i - c'_1 \end{bmatrix}$$

и

$$\mathbf{T}^i = \begin{bmatrix} \tau_0^i & \tau_1^i \\ \tau_1^i & 0 \end{bmatrix} + \begin{bmatrix} \tau_0^i & \tau_1^i \\ 0 & \tau_0^i \end{bmatrix} = \begin{bmatrix} 2\tau_0^i & 2\tau_1^i \\ \tau_1^i & \tau_0^i \end{bmatrix}.$$

Дисперсия σ_a^2 и θ могут быть теперь вычислены по формулам $\sigma_a^2 = \tau_0^i$, $\theta = -\tau_1/\tau_0$.

Табл. П6.1 показывает, как сходятся итерации методами (1) и (2). Программа 2 в сборнике программ в конце книги может быть использована для вычисления начальных оценок параметров любого процесса АРПСС(p, d, q).

ПРИЛОЖЕНИЕ П6.3. ПРЯМОЙ И ВОЗВРАТНЫЙ ПРОЦЕССЫ ПСС ПОРЯДКА (0, 1, 1)

В разд. 6.4 мы встретились с прямой и возвратной моделями стационарных процессов. Интересно также рассмотреть соответствующие двойственные модели нестационарных процессов. В качестве примера рассмотрим процесс ПСС(0, 1, 1). Пусть a_t и e_t — последовательности случайного шума с дисперсией $\sigma_a^2 = \sigma_e^2$. Процессы скользящего среднего

$$w_t = (1 - \theta B) a_t, \quad w_t = (1 - \theta F) e_t$$

имеют одинаковые автоковариации. После введения обозначения $\varepsilon_t = -e_{t-1}$ можно представить возвратный процесс как

$$-w_t = (1 - \theta F) e_{t-1}.$$

Положим, что фактически делаются наблюдения z_t , первые разности которых образуют w_t , так что

$$w_t = (1 - B) z_t = -(1 - F) z_{t-1}.$$

Тогда можно выразить модель через z в одной из двух форм следующим образом:

$$\begin{aligned} (1 - B) z_t &= (1 - \theta B) a_t, & (1 - F) z_t &= (1 - \theta F) e_t, \\ a_t &= \frac{1 - B}{1 - \theta B} z_t = & e_t &= \frac{1 - F}{1 - \theta F} z_t = \\ &= \left\{ 1 - \frac{(1 - \theta) B}{1 - \theta B} \right\} z_t, & &= \left\{ 1 - \frac{(1 - \theta) F}{1 - \theta F} \right\} z_t, \\ a_t &= z_t - (1 - \theta)(z_{t-1} + & e_t &= z_t - (1 - \theta)(z_{t+1} + \\ & + \theta z_{t-2} + \theta^2 z_{t-3} + \dots) = & + \theta z_{t+2} + \theta^2 z_{t+3} + \dots) = \\ &\quad \quad \quad > & < \\ &= z_t - z_{t-1}, & &= z_t - z_{t+1}. \end{aligned}$$

где $\overset{>}{z}_{t-1}$ — возвратное экспоненциально взвешенное среднее, использующее значения z до времени t . Следовательно,

$$\overset{>}{z}_{t-1} - \overset{>}{z}_{t-2} = (1 - \theta) a_{t-1},$$

$$\overset{>}{z}_{t-1} = (1 - \theta) \overset{>}{z}_{t-1} + \theta \overset{>}{z}_{t-2}.$$

где $\overset{<}{z}_{t+1}$ — прямое экспоненциально взвешенное среднее, использующее значения z после времени t . Следовательно,

$$\overset{<}{z}_{t+1} - \overset{<}{z}_{t+2} = (1 - \theta) e_{t+1},$$

$$\overset{<}{z}_{t+1} = (1 - \theta) \overset{<}{z}_{t+1} + \theta \overset{<}{z}_{t+2}.$$

Положим, что имеются значения z в обе стороны от текущего момента t , ковариации разностей которых удовлетворяют (6.4.3),

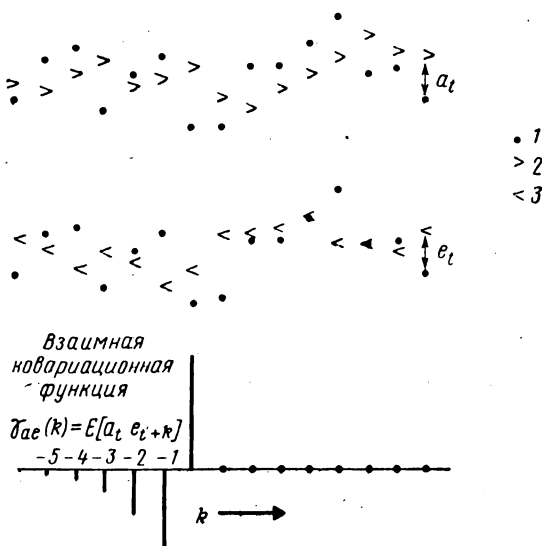


Рис. П6.2. Прямые и обратные экспоненциально взвешенные скользящие средние (ЭВСС).

1 — ряд, генерированный процессом (0,1,1), 2 — возвратные ЭВСС, 3 — прямые ЭВСС.

Из этого ряда можно построить два множества случайных величин

$\dots a_{t-2}, a_{t-1}, a_t, a_{t+1}, a_{t+2}, \dots$ и $\dots e_{t-2}, e_{t-1}, e_t, e_{t+1}, e_{t+2}, \dots$,

где a_t , например, — разность между z_t и $\overset{>}{z}_{t-1}$, возвратным ЭВСС, вычисленным по предыдущим значениям $z_{t-1}, z_{t-2}, z_{t-3}, \dots$, в то время как e_t — разность между z_t и $\overset{<}{z}_{t+1}$, прямым ЭВСС, вычисленным по значениям $z_{t+1}, z_{t+2}, z_{t+3}, \dots$.

Связь между a и e . Для получения связи между a и e можно написать

$$a_t = \frac{(\theta - B)}{1 - \theta B} e_t = \theta e_t - (1 + \theta) \frac{(1 - \theta) B}{1 - \theta B} e_t = \\ = \theta e_t - (1 + \theta) e_{t-1}^>. \quad (\text{П6.3.1})$$

Наоборот,

$$e_t = \frac{(\theta - F)}{1 - \theta F} a_t = \theta a_t - (1 + \theta) a_{t+1}^<, \quad (\text{П6.3.2})$$

где $e_t^>$ и $a_t^<$ — экспоненциально взвешенные скользящие средние (ЭВСС), определенные ранее.

Легко показать, например, что если a — это последовательность независимых случайных величин с нулевым средним значением и дисперсией σ^2 , то e , генерируемые по ним согласно (П6.3.2), обладают теми же свойствами.

Хотя между собой e независимы, a и e взаимно коррелированы. Используя (П6.3.1), находим

$$\gamma_{ae}(k) = E[a_t e_{t+k}] = \begin{cases} -(1 - \theta^2) \theta^{-(k+1)} \sigma^2, & k < 0, \\ \theta \sigma^2, & k = 0, \\ 0, & k > 0. \end{cases}$$

Возникающая ситуация иллюстрируется рис. П6.2. Вверху рисунка показана часть наблюдений процесса $(0, 1, 1)$ с $\theta = 0,5$, соответствующие возвратные ЭВСС и получающиеся a . Ниже показаны тот же ряд с прямыми ЭВСС и получающиеся e . Внизу показана взаимная ковариационная функция $\gamma_{ae}(k)$ процессов a_t и e_{t+k} .

Оценивание модели

После того как процесс идентификации привел к пробному варианту модели, нам необходимо получить эффективные оценки параметров. После этого подогнанная модель будет подвергнута диагностической проверке и тестам на качество подгонки. Как указывал Фишер, для того чтобы критерии согласия были подходящими, необходимо эффективно использовать данные в процессе подгонки. Если это не было сделано, неадекватность найденной модели может быть вызвана неэффективностью процедуры подгонки, а не тем, что неадекватен вид модели. Эта глава содержит общее рассмотрение методов правдоподобия и Байеса для оценок параметров стохастической модели. Всюду в этой главе полужирные буквы обозначают векторы и матрицы. Так, $\mathbf{X} = \{x_{ij}\}$ — матрица с элементом x_{ij} в i -й строке и j -м столбце и \mathbf{X}' — транспонированная матрица.

7.1. ИССЛЕДОВАНИЕ ФУНКЦИЙ ПРАВДОПОДОБИЯ И СУММЫ КВАДРАТОВ

7.1.1. Функция правдоподобия

Пусть мы имеем выборку из N наблюдений \mathbf{z} , которую мы считаем реализацией N -мерной случайной величины, для которой известно вероятностное распределение $p(\mathbf{z}|\xi)$ зависит от неизвестного параметра ξ . Вектор ξ мы используем для обозначения общего множества параметров, и, в частности, он может относиться к $p + q + 1$ параметрам (ϕ, θ, σ) модели АРПСС.

До того как появились данные, $p(\mathbf{z}|\xi)$ связывает плотность вероятности с каждым конкретным выходом \mathbf{z} эксперимента для фиксированного ξ . После того как данные появились, наступает момент для рассмотрения различных возможных значений ξ , которые могли бы привести к заданному множеству \mathbf{z} фактически полученных наблюдений. Для этих целей подходит *функция правдоподобия* $L(\xi|\mathbf{z})$, которая имеет ту же форму, что и $p(\mathbf{z}|\xi)$, но в которой теперь \mathbf{z} зафиксировано, а ξ переменные. Обычно важны только относительные значения $L(\xi|\mathbf{z})$, и поэтому функцию правдоподобия часто считают содержащей произвольную мультипликативную константу.

Часто удобнее работать с логарифмической функцией правдоподобия $\ln L(\xi|z) = l(\xi|z)$, содержащей произвольную аддитивную константу. Одна из причин, по которым функция правдоподобия имеет фундаментальное значение в теории оценивания, связана с «принципом правдоподобия», выдвинутым с разных позиций Фишером [56], Бернардом [57] и Бирнбаумом [58]. Этот принцип гласит (при условии правильности предполагаемой модели): все, что данные могут сказать о параметрах модели, содержится в функции правдоподобия, а все другие аспекты данных не имеют отношения к делу. При байесовском подходе функция правдоподобия также важна, так как она является той компонентой апостериорного распределения параметров, которая зависит от данных.

Для полного понимания ситуации с оцениванием необходимо проделать подробное аналитическое и графическое изучение функции правдоподобия; в байесовском подходе мы должны изучить апостериорное распределение параметров, которое в рассматриваемых ситуациях определяется в основном правдоподобием. Во многих примерах с выборками средних и больших размеров логарифмическая функция правдоподобия унимодальна и в достаточно большой окрестности максимума может быть аппроксимирована квадратичной функцией. Значения параметров, максимизирующие функцию правдоподобия или, что эквивалентно, логарифмическую функцию правдоподобия, называются *оценками максимального правдоподобия* (МП).

Вторые производные логарифмической функции правдоподобия дают меру «растянутости» функции правдоподобия и могут использоваться для вычисления приближенных стандартных ошибок оценок. Предельные свойства оценок максимального правдоподобия обычно доказываются для независимых наблюдений [59]. Но, как показано Уиттлом [87], они могут быть обобщены на стационарные временные ряды.

В последующем изложении мы будем исходить из того, что читатель знаком с некоторыми фундаментальными понятиями теории оценивания. Приложения П7.1 и П7.2 содержат сводки наиболее существенных результатов теории нормального распределения и линейного метода наименьших квадратов, необходимых в этой главе. Некоторые из важных предшествующих работ по оценке параметров моделей временных рядов можно найти в [34, 78, 88, 89, 93—95, 97—101].

7.1.2. Условное правдоподобие для процесса АРПСС

Пусть $N = n + d$ исходных наблюдений z образуют временной ряд, который мы обозначим $z_{-d+1}, \dots, z_0, z_1, z_2, \dots, z_n$. Предположим, что этот ряд генерирован моделью АРПСС

порядка (p, d, q) . По этим наблюдениям можно генерировать ряд \mathbf{w} из $n = N - d$ разностей $\omega_1, \omega_2, \dots, \omega_n$, где $\omega_t = \nabla^d z_t$. Отсюда общая задача подгонки параметров ϕ и θ модели АРПСС (6.1.1) эквивалентна задаче подгонки к ω стационарной смешанной модели АРСС (p, q) , которую можно представить в виде

$$a_t = \tilde{\omega}_t - \phi_1 \tilde{\omega}_{t-1} - \phi_2 \tilde{\omega}_{t-2} - \dots - \phi_p \tilde{\omega}_{t-p} + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q}, \quad (7.1.1)$$

где $\omega_t = \nabla^d z_t$ и $\tilde{\omega}_t = \omega_t - \mu$ с $E[\omega_t] = \mu$.

Для $d > 0$ часто разумно предположить, что $\mu = 0$ (см. обсуждение этого вопроса в разд. 4.1.3, 6.2.3, 6.3.5 и 6.3.7). Когда это не оправдано, мы полагаем, что μ можно заменить на

$\bar{\omega} = \sum_{i=1}^n \omega_i / n$. Для выборок такого объема, которые обычно рассматриваются в анализе временных рядов, эта аппроксимация адекватна. Однако, если это нужно, можно включить μ в число оцениваемых параметров. Описываемые здесь методики позволяют оценивать μ одновременно с другими параметрами.

Значения ω не могут быть сразу подставлены в (7.1.1) для вычисления a из-за трудностей с начальными условиями разностного уравнения. Однако если мы предположим, что p значений \mathbf{w}_* ряда ω и q значений \mathbf{a}_* ряда a даны до начала наблюдения ряда ω , то значения a_1, a_2, \dots, a_n , обусловленные этим выбором, можно вычислить по (7.1.1).

Следовательно, для любого данного набора параметров (ϕ, θ) и начальных значений $(\mathbf{w}_*, \mathbf{a}_*)$ мы можем последовательно вычислить множество значений $a_t(\phi, \theta | \mathbf{w}_*, \mathbf{a}_*, \mathbf{w})$ для $t = 1, 2, \dots, n$. Далее, в предположении, что a распределены по нормальному закону, получаем

$$p(a_1, a_2, \dots, a_n) \approx \sigma_a^{-n} \exp \left\{ - \left(\sum_{i=1}^n a_i^2 / 2\sigma_a^2 \right) \right\}.$$

Если дано множество величин \mathbf{w} , то условная логарифмическая функция правдоподобия для параметров (ϕ, θ, σ_a) при заданных $(\mathbf{w}_*, \mathbf{a}_*)$ будет равна

$$l_*(\phi, \theta, \sigma_a) = -n \ln \sigma_a - \frac{S_*(\phi, \theta)}{2\sigma_a^2}, \quad (7.1.2)$$

куда, в согласии со сказанным ранее, не включена аддитивная константа и где

$$S_*(\phi, \theta) = \sum_{i=1}^n a_i^2(\phi, \theta | \mathbf{w}_*, \mathbf{a}_*, \mathbf{w}). \quad (7.1.3)$$

Выше мы использовали звездочки как нижние индексы в функциях правдоподобия и суммы квадратов, чтобы подчеркнуть, что

они зависят от выбора начальных значений. Заметим, что в условное правдоподобие l_* данные входят только через условную сумму квадратов. Очевидно, что изолинии функции l_* для любого заданного значения σ_a в пространстве (ϕ, θ, σ_a) совпадают с изолиниями S_* , т. е. оценки максимального правдоподобия те же, что и оценки наименьших квадратов. Вообще говоря, мы можем в предположении о нормальности изучать поведение условного правдоподобия, исследуя условную сумму квадратов. В частности, для любого фиксированного σ_a l_* — линейная функция S_* .

7.1.3. Выбор начальных значений для вычисления условного правдоподобия

Рассмотрим вкратце способ расчета безусловного правдоподобия, которое в сущности и необходимо для оценивания параметров.

В ряде случаев, когда n не мало, достаточно хорошим приближением к безусловной функции правдоподобия является условная функция правдоподобия, в которой вместо элементов w_* и a_* в (7.1.3) подставлены подходящие численные значения. Один из способов — приравнять элементы w_* и a_* их безусловным математическим ожиданиям. Безусловные математические ожидания элементов a_* — нули, и если модель не содержит детерминированной части (в частности, при $\mu = 0$), безусловные математические ожидания w_* также будут равны нулю*). Однако это приближение оказывается плохим, если некоторые корни $\phi(B) = 0$ близки к единичной окружности, т. е. процесс приближается к нестационарному. В этом случае начальное значение данных w_1 может значительно отклоняться от его безусловного математического ожидания, и введение начальных значений этого типа создаст переходный процесс большой амплитуды, который будет медленно затухать. При подгонке модели порядка (p, d, q) мы будем иногда пользоваться более надежной процедурой аппроксимации. Она сводится к вычислению a , начиная с a_{p+1} , и приравниванию предварительно предыдущих a нулю. Таким образом, фактически имеющиеся значения полностью использованы для вычисления w .

Этим способом мы можем вычислить сумму квадратов только для $n - p = N - p - d$ значений ряда a_t , но для длинного ряда небольшая потеря информации не существенна. В случаях когда авторегрессионных членов не имеется, две описанные методики эквивалентны. Для сезонных рядов, рассмотренных в гл. 9,

) Если более приемлемо предположение $E[w_t] = \mu \neq 0$, мы можем подставить \bar{w} вместо каждого из элементов w_ .

условная аппроксимация не дает удовлетворительных результатов, и вычисление безусловного правдоподобия особенно необходимо.

Проиллюстрируем теперь на простом примере рекуррентный расчет условных сумм квадратов S_* .

Вычисление условной суммы квадратов для процесса $(0, 1, 1)$. Ряд В был пробно идентифицирован в табл. 6.4 как процесс ПСС $(0, 1, 1)$

$$\nabla z_t = (1 - \theta B) a_t \quad -1 < \theta < 1, \quad (7.1.4)$$

т. е.

$$a_t = w_t + \theta a_{t-1},$$

где $w_t = \nabla z_t$ и $E[w_t] = 0$. Нужно напомнить, что в гл. 6 была получена предварительная оценка методом моментов, из которой следовало, что для этих данных θ близко к нулю.

Вычисление нескольких первых a было проведено в табл. 7.1 для частного значения параметра $\theta = 0,5$. Элементы a вычислялись рекуррентным образом по формуле $a_t = w_t + 0,5a_{t-1}$ с точностью до одного знака после запятой. В соответствии со сказанным выше, чтобы начать процесс вычислений, a_0 приравнялось нулю. Это значение показано курсивом. Продолжая вычисления указанным путем, мы нашли, что

$$S_*(0,5) = \sum_{i=1}^{368} a_i^2 (\theta = 0,5 | a_0 = 0) = 27\,694.$$

Таблица 7.1

Рекуррентный расчет a для первых 10 значений ряда В при $\theta = 0,5$

t	z_t	$w_t = \nabla z_t$	$a_t = w_t + 0,5a_{t-1}$
0	460		<i>0</i>
1	457	-3	-3,0
2	452	-5	-6,5
3	459	7	3,8
4	462	3	4,9
5	459	-3	-0,6
6	463	4	3,7
7	479	16	17,8
8	493	14	22,9
9	490	-3	8,4

Рекуррентный расчет особенно удобен для реализации на ЭВМ. Перебирая значения θ от $-0,5$ до $+0,5$ с шагом $0,1$, мы вычисляли значения условной суммы квадратов $S_*(\theta)$ (при условии $a_0 = 0$); результаты показаны в 3-й строке табл. 7.2.

Таблица 7.2

Сумма квадратов для модели $\nabla z_t = (1 - \theta B)a_t$, подгоняемой к ряду В

θ	-0,5	-0,4	-0,3	-0,2	-0,1	0,0
$\lambda = (1 - \theta)$	1,5	1,4	1,3	1,2	1,1	1,0
$S_*(\theta)$	23 929	21 595	20 222	19 483	19 220	19 363
$S(\theta)$	23 929	21 594	20 222	19 483	19 220	19 363

θ	0,1	0,2	0,3	0,4	0,5	
$\lambda = (1 - \theta)$	0,9	0,8	0,7	0,6	0,5	
$S_*(\theta)$	19 896	20 851	22 315	24 471	27 694	
$S(\theta)$	19 896	20 849	22 315	24 478	27 691	

7.1.4. Безусловное правдоподобие; сумма квадратов; оценки наименьших квадратов

В приложении П7.4 показано, что для $N = n + d$ наблюдений, генерированных предполагаемой моделью АРПСС, безусловная логарифмическая функция правдоподобия имеет вид

$$l(\phi, \theta, \sigma_a) = f(\phi, \theta) - n \ln \sigma_a - \frac{S(\phi, \theta)}{2\sigma_a^2}, \quad (7.1.5)$$

где $f(\phi, \theta)$ — функция ϕ и θ . Безусловная сумма квадратов выражается как

$$S(\phi, \theta) = \sum_{t=-\infty}^n [a_t | \phi, \theta, w]^2, \quad (7.1.6)$$

где $[a_t | \phi, \theta, w] = E[a_t | \phi, \theta, w]$ обозначает условное математическое ожидание a_t при фиксированных ϕ , θ и w . В дальнейшем, когда из контекста будет ясно, о чем идет речь, мы будем сокращенно обозначать это условное ожидание $[a_t]$.

Обычно $f(\phi, \theta)$ значимо только при малых n . Для средних и больших n в (7.1.5) доминирует $S(\phi, \theta)/2\sigma_a^2$, и поэтому изолинии безусловной суммы квадратов в пространстве параметров (ϕ, θ) практически совпадают с изолиниями функции правдоподобия и логарифмической функции правдоподобия. Отсюда, в частности, следует, что оценки параметров, полученные минимизацией суммы квадратов (7.1.6), которые мы называем *оценками наименьших квадратов*, будут, как правило, очень близкими к оценкам

максимального правдоподобия. При байесовском подходе в предположениях, рассмотренных в разд. 7.4, для всех процессов $AP(p)$ и $CC(q)$ апостериорная плотность распределения по существу есть функция только $S(\phi, \theta)$. Отсюда оценки наименьших квадратов очень близки к оценкам, дающим максимум апостериорной плотности. Далее в этом разделе и в разд. 7.1.5 мы главным образом будем заниматься расчетом, анализом и использованием безусловной суммы квадратов $S(\phi, \theta)$, определенной в (7.1.6), и вычислением оценок наименьших квадратов.

При вычислении безусловной суммы квадратов $[a]$ будем находить по рекуррентной формуле, полученной из (7.1.1) взятием условных математических ожиданий от обеих частей. Предварительный расчет в обратном направлении дает значения $[w_{-j}]$, $j = 0, 1, 2, \dots$ (т. е. прогнозы назад), необходимые для начала рекуррентных вычислений в прямом направлении.

Вычисление безусловной суммы квадратов для процесса скользящего среднего. В качестве иллюстрации рассмотрим пример биржевых цен акций IBM, используя опять 10 первых *) значений ряда, приведенного в табл. 7.1. Для $w_t = Vz_t$ в разд. 6.4.3 было показано, что подгоняемая модель порядка $(0, 1, 1)$ может быть записана в прямой или в возвратной форме:

$$w_t = (1 - \theta B) a_t, \quad w_t = (1 - \theta F) e_t,$$

где $\mu = E[w_t]$ опять предполагается равным нулю. Отсюда можно записать

$$[e_t] = [w_t] + \theta [e_{t+1}], \quad (7.1.7)$$

$$[a_t] = [w_t] + \theta [a_{t-1}], \quad (7.1.8)$$

где $[w_t] = w_t$ для $t = 1, 2, \dots, n$ и является прогнозом w_t назад для $t \leq 0$. Эти два уравнения мы и используем в расчетах. Удобная схема вычислений показана в табл. 7.3.

Мы начинаем с внесения в таблицу известных величин. К ним относятся

а) значения данных z_0, z_1, \dots, z_9 , по которым мы можем вычислить первые разности w_1, w_2, \dots, w_9 ;

б) значения $[e_0], [e_{-1}], \dots$, равные нулю, так как e_0, e_{-1}, \dots распределены независимо от w ;

в) значения $[a_{-1}], [a_{-2}], \dots$, равные нулю, потому что для любого процесса $CC(q)$ a_{-q}, a_{-q-1}, \dots распределены независимо от w . Однако следует помнить, что в общем случае $[a_0], [a_{-1}], \dots, [a_{-q+1}]$ не равны нулю и должны быть получены прогнозированием назад. В этом примере таким путем получено $[a_0]$.

*) На практике, конечно, нельзя получить полезных оценок параметров по столь малому числу наблюдений, как десять. Мы используем это подмножество данных просто для иллюстрации вычисления.

Таблица 7.3

Вычисление $[a]$ по первым десяти значениям ряда B при $\theta = 0,5$

t	z_t	$[a_t]$	$0,5 [a_{t-1}]$	$[w_t]$	$0,5 [e_{t+1}]$	$[e_t]$
-1	[458,4]	0	0	0	0	0
0	460	1,6	0	1,6	-1,6	0
1	457	-2,2	0,8	-3,0	-0,1	-3,1
2	452	-6,1	-1,1	-5,0	4,8	-0,2
3	459	4,0	-3,0	7,0	2,6	9,6
4	462	5,0	2,0	3,0	2,3	5,3
5	459	-0,5	2,5	-3,0	7,6	4,6
6	463	3,8	-0,2	4,0	11,1	15,1
7	479	17,9	1,9	16,0	6,2	22,2
8	493	23,0	9,0	14,0	-1,5	12,5
9	490	8,5	11,5	-3,0	0	-3,0

Начиная с конца ряда, используем (7.1.7) для вычисления $[e_t]$ при $t = 9, 8, 7, \dots 1$. Этот возвратный процесс начинается при помощи той же аппроксимации, что описана ранее при вычислении условной суммы квадратов. В данном случае это приводит к тому, что $[e_{10}] = 0$. В общем эффект этой аппроксимации состоит во внесении в систему переходного процесса. В силу стационарности операторов $\phi(B)$ и $\theta(B)$ этот процесс для ряда умеренной длины будет почти всегда пренебрежимо мал к моменту начала ряда и поэтому не повлияет на расчет a . Как мы увидим далее, если необходимо, можно проконтролировать эффект этой аппроксимации в любом конкретном случае, проведя второй итеративный цикл.

Далее, чтобы начать рекуррентное заполнение табл. 7.3, в строку, соответствующую $t = 9$, мы вносим нуль (показан курсивом) в столбец 6 вместо неизвестной величины $0,5 [e_{10}]$. Затем, пользуясь (7.1.7), получаем

$$[e_9] = [w_9] + 0,5 [e_{10}] = w_9 + 0 = -3$$

и проставляем $0,5 [e_9] = -1,5$ в строке $t = 8$. Это позволяет нам вычислить $[e_8]$ и т. д. В завершение получаем

$$[e_0] = [w_0] + \theta [e_1],$$

т. е.

$$0 = [w_0] - 1,6,$$

что дает $[w_0] = 1,6$ и, следовательно, $[w_{-h}] = 0$, $h = 1, 2, 3 \dots$

Пользуясь (7.1.8) при $t = 0$, получаем

$$[a_0] = [w_0] + \theta [a_{-1}] = 1,6 + (0,5)(0) = 1,6$$

и можем продолжить вычисление вперед остающихся $[a]$. Сравнение значений a_t в табл. 7.1 и 7.3 показывает, что в этом частном примере переходный процесс, введенный изменением начальной величины, не сказывается при $t > 5$. Вычисляя таким образом члены всего ряда, находим безусловную сумму квадратов

$$S(0,5) = \sum_{t=0}^{368} [a_t | 0,5, w]^2 = 27\,691,$$

что для данного примера очень близко к значению условной суммы $S_*(0,5) = 27\,694$.

Безусловные суммы квадратов $S(\theta)$ для значений θ между $-0,5$ и $+0,5$ приведены в нижней строке табл. 7.2 и для этого частного примера очень близки к значениям условных сумм $S_*(\theta)$.

7.1.5. Общий способ вычисления безусловной суммы квадратов

В примере, приведенном выше, w_t был процессом скользящего среднего первого порядка с нулевым средним значением. Оказалось, что все прогнозы для упреждений больше 1 равны нулю, и, следовательно, только одно предварительное значение (прогноз назад $[w_0] = 1,6$) было необходимо для начала рекуррентных вычислений. Для процесса скользящего среднего q -го порядка были бы необходимы q ненулевых предварительных значений $[w_0], [w_{-1}], \dots, [w_{1-q}]$. Для оценки параметров авторегрессии имеются специальные способы, которые мы рассмотрим позже, в разд. 7.3.1. Однако в приложении П7.4 будет показано, что описываемая здесь процедура позволяет вычислять безусловные суммы квадратов для любой модели АРПСС с любой требуемой точностью аппроксимации.

Конкретнее, положим, что w_t генерируются стационарной прямой моделью

$$\phi(B) \tilde{w}_t = \theta(B) a_t, \quad (7.1.9)$$

где $\nabla^d z_t = w_t$ и $\tilde{w}_t = w_t - \mu$. Тогда они с равным успехом могли бы генерироваться возвратной моделью

$$\phi(F) \tilde{w}_t = \theta(F) e_t. \quad (7.1.10)$$

Как и ранее, мы сначала используем (7.1.10) для получения прогнозов назад $[\tilde{w}_{-j} | \phi, \theta, w]$. Теоретически присутствие оператора авторегрессии обеспечивает бесконечно протяженный ряд таких оценок. Однако из-за стационарного характера этого оператора на практике оценки $[\tilde{w}_t]$ в некоторый момент $t = -Q$ (где Q невелико) и для $t < -Q$ практически равны нулю.

Тогда с достаточной точностью можно записать

$$\tilde{w}_t = \phi^{-1}(B) \theta(B) a_t = \sum_{j=0}^{\infty} \psi_j a_{t-j} \approx \sum_{j=0}^Q \psi_j a_{t-j}.$$

Это означает, что первоначальный смешанный процесс может быть заменен процессом скользящего среднего порядка Q , и для этого можно использовать способ, описанный в разд. 7.1.4. В общем двойной набор уравнений для вычисления условных ожиданий $[a_t | \phi, \theta, w]$ получен взятием условных ожиданий от обеих частей (7.1.10) и (7.1.9), т. е. из уравнений

$$\phi(F)[\tilde{w}_t] = \theta(F)[e_t] \quad (7.1.11)$$

находим прогнозы назад и затем из уравнений

$$\phi(B)[\tilde{w}_t] = \theta(B)[a_t] \quad (7.1.12)$$

вычисляем $[a_t]$. Если мы найдем, что прогнозы для времен упреждений, больших некоторого Q , становятся пренебрежимо малыми, можно начинать рекуррентный расчет вперед по формулам

$$\begin{aligned} [e_{-j} | \phi, \theta, w] &= 0, & j &= 0, 1, 2, \dots, \\ [a_{-j} | \phi, \theta, w] &= 0, & j &> Q - 1. \end{aligned} \quad (7.1.13)$$

Вычисление безусловной суммы квадратов для смешанного процесса авторегрессии — скользящего среднего. В качестве примера рассмотрим следующие $n = 12$ последовательных значений $w_t = \nabla^d z_t$:

t	1	2	3	4	5	6	7	8	9	10	11	12
w_t	2,0	0,8	-0,3	-0,3	-1,9	0,3	3,2	1,6	-0,7	3,0	4,3	1,1

Положим, что мы хотим вычислить безусловную сумму квадратов $S(\phi, \theta)$, относящуюся к процессу АРПСС $(1, d, 1)$,

$$(1 - \phi B) w_t = (1 - \theta B) a_t,$$

где $\nabla^d z_t = w_t$ или, что то же самое, к возвратному процессу

$$(1 - \phi F) w_t = (1 - \theta F) e_t$$

в предположении, что среднее значение w равно нулю.

Конечно, оценки, базирующиеся на 12 наблюдениях, не имеют почти никакой ценности, но тем не менее этот короткий ряд послужит для объяснения природы вычислений. Проиллюстрируем это для значений параметров $\phi = 0,3$, $\theta = 0,7$. Тогда (7.1.11) и

(7.1.12) можно записать в виде

$$[e_t] = [w_t] - 0,3[w_{t+1}] + 0,7[e_{t+1}], \quad (7.1.14)$$

$$[a_t] = [w_t] - 0,3[w_{t-1}] + 0,7[a_{t-1}], \quad (7.1.15)$$

где $[w_t] = w_t$ ($t = 1, 2, \dots, n$).

Схема вычислений показана в табл. 7.4. Сначала в центральный столбец вносятся данные, затем в последний столбец вносятся нулевые значения для $[e_0]$, $[e_{-1}]$, $[e_{-2}]$, \dots . Возвратное уравнение (7.1.14) начинает использоваться точно таким же образом, как описано в разд. 7.1.3 для расчета условных величин. Из-за того что в модели имеется оператор авторегрессии *первого* порядка, мы начинаем в *одном* шаге от конца ряда и подставляем 0 в строчку $t = 11$ вместо неизвестного значения 0,7 $[e_{12}]$. Рекуррентный расчет ведется вперед при помощи (7.1.14). В табл. 7.4 и 7.5 данные приведены с точностью до одного знака после запятой; расчеты велись с точностью до двух знаков. Прогнозы назад $[w_{-j}]$ ($j = 0, 1, 2, \dots$) быстро затухают и при нашей точности вычислений равны нулю для $j > 4$. На основании (7.1.13) оценки $[a_{-j}]$ принимаются равными нулю для $j > 4$. Рекуррентный расчет вперед ведется по формуле (7.1.15); вычисляются $[a_t]$.

Безусловная сумма квадратов $S(\phi, \theta)$ находится суммированием квадратов всех вычисленных $[a_t]$. Отсюда

$$S(0,3; 0,7) = \sum_{t=-4}^{12} [a_t]^2 = 89,2.$$

На практике второй итеративный цикл не нужен почти никогда. Однако, чтобы показать быструю сходимость этого процесса даже для нереально короткого ряда из 12 наблюдений, в табл. 7.5 был проведен второй цикл итераций. В нем значение $a_{12} = 3,99$, найденное в предыдущей итерации, использовалось для начала новой итерации при вычислении прогнозов w вперед по (7.1.15). Полученные величины подставлялись в возвратные уравнения (7.1.14) для получения новых прогнозов назад $[w_0]$, $[w_{-1}]$, \dots . В общем для второго цикла вычислений мы используем соотношения

$$\begin{aligned} [a_{n+i} | \phi, \theta, w] &= 0, & i = 1, 2, \dots, \\ [e_{n+i} | \phi, \theta, w] &= 0, & i > Q', \end{aligned} \quad (7.1.16)$$

где Q' выбрано так, что $[w_{n+i}]$ пренебрежимо мало при $i > Q'$. В табл. 7.5 показано, что второй цикл вычисления приводит к $[w_0] = -1,04$, что совпадает со значением, полученным в первом цикле. Поэтому дальнейшее вычисление a будет давать те же результаты, что и ранее.

Таблица 7.4

Вычисление $[a]$ и $S(0,3; 0,7)$ по 12 значениям ряда, предположительно генерируемого процессом $(1-0,3B)w_t = (1-0,7B)a_t$

t	$[a_t]$	$0,7[a_{t-1}]$	$-0,3[w_{t-1}]$	$[w_t]$	$-0,3[w_{t+1}]$	$0,7[e_{t+1}]$	$[e_t]$
-4	-0,01	0,00	0,00	-0,01	0,01	0	0
-3	-0,04	-0,01	0,00	-0,03	0,03	0	0
-2	-0,11	-0,03	0,01	-0,09	0,09	0	0
-1	-0,36	-0,08	0,03	-0,31	0,31	0	0
0	-1,20	-0,25	0,09	-1,04	-0,60	1,64	0
1	1,47	-0,84	0,31	2,0	-0,24	0,58	2,34
2	1,23	1,03	-0,60	0,8	0,09	-0,06	0,83
3	0,32	0,86	-0,24	-0,3	0,09	0,13	-0,08
4	0,02	0,23	0,09	-0,3	0,57	-0,09	0,18
5	-1,80	0,01	0,09	-1,9	-0,09	1,86	-0,13
6	-0,39	-1,26	0,57	0,3	-0,96	3,32	2,66
7	2,84	-0,27	-0,09	3,2	-0,48	2,02	4,74
8	2,63	1,99	-0,96	1,6	0,21	1,08	2,89
9	0,66	1,84	-0,48	-0,7	-0,90	3,14	1,54
10	3,67	0,46	0,21	3,0	-1,29	2,78	4,49
11	5,97	2,57	-0,90	4,3	-0,33	0	3,97
12	3,99	4,18	-1,29	1,1			

$$S(0,3; 0,7) = \sum_{t=-4}^{12} [a_t]^2 = 89,2$$

Таблица 7.5

Вторая итерация вычислений, начатых в табл. 7.4

t	$[a_t]$	$0,7[a_{t-1}]$	$-0,3[w_{t-1}]$	$[w_t]$	$-0,3[w_{t+1}]$	$0,7[e_{t+1}]$	$[e_t]$
0				-1,04	-0,60	1,64	0
1				2,0	-0,24	0,58	2,34
2				0,8	0,09	-0,06	0,83
3				-0,3	0,09	0,12	-0,09
4				-0,3	0,57	-0,10	0,17
5				-1,9	-0,09	1,85	-0,14
6				0,3	-0,96	3,30	2,64

Продолжение

t	$[a_t]$	$0,7 [a_{t-1}]$	$-0,3 [w_{t-1}]$	$[w_t]$	$-0,3 [w_{t+1}]$	$0,7 [e_{t+1}]$	$[e_t]$
7				3,2	-0,48	2,00	4,72
8				1,6	0,21	1,05	2,86
9				-0,7	-0,90	3,09	1,49
10				3,0	-1,29	2,71	4,42
11				4,3	-0,33	-0,10	3,87
12	3,99			1,1	0,74	-1,98	-0,14
13	0	2,79	-0,33	-2,46	0,22	-0,60	-2,84
14	0	0	0,74	-0,74	0,07	-0,18	-0,85
15	0	0	0,22	-0,22	0,02	-0,06	-0,26
16	0	0	0,07	-0,07	0,01	-0,02	-0,08
17	0	0	0,02	-0,02	0,00	-0,01	-0,03
18	0	0	0,01	-0,01	0,00		-0,01

$$S(0,3; 0,7) = \sum_{t=1}^{18} [e_t]^2 = 89,3$$

Далее, как следует из приложения П7.4, $S(\phi, \theta | w)$ также быть может вычислено по сумме квадратов $[e_t]$. Пользуясь этим, из табл. 7.5 находим, что

$$S(0,3; 0,7) = \sum_{t=1}^{18} [a_t]^2 = 89,3,$$

что близко согласуется с величиной 89,2, найденной по a_t .

Мы видели, что при подгонке к ряду В стоимости акций IBM процесса порядка (0, 1, 1) *условные* суммы квадратов давали очень точное приближение к безусловным суммам. Что это не всегда верно, можно увидеть из данного примера.

Ранее, в разд. 7.1.3, упоминалось о двух условных суммах квадратов, которые можно использовать как приближения к безусловным суммам. Они были получены

1) началом рекуррентного расчета с первого доступного наблюдения, приравниванием всех неизвестных a или e нулю и всех w их безусловным математическим ожиданиям;

2) началом рекуррентного счета с p -го наблюдения при использовании только наблюденных значений для w и нуля вместо неизвестных a или e .

Две такие условные суммы квадратов и две безусловные суммы квадратов,

- 3) полученные из $[a_t]$,
- 4) полученные из $[e_t]$, сравниваются ниже.
- 3) Неизвестные w и e приняты равными нулю:

$$\sum_{t=1}^{12} (e_t | 0,3; 0,7, w_{13} = 0, e_{13} = 0, w)^2 = 101,0.$$

- 2) Неизвестные e приняты равными нулю:

$$\sum_{t=2}^{12} (e_t | 0,3; 0,7, w_{12} = 1,1; e_{12} = 0, w)^2 = 82,4.$$

- 3) Расчет безусловных сумм, первая итерация:

$$\sum_{t=-4}^{12} [a_t | 0,3; 0,7; w]^2 = 89,2.$$

- 4) Расчет безусловных сумм, вторая итерация:

$$\sum_{t=1}^{18} [e_t | 0,3; 0,7, w]^2 = 89,3.$$

Сумма квадратов в (1) — очень плохое приближение к (3) и (4), хотя расхождение, превышающее для этого ряда из 12 чисел 10%, уменьшилось бы при большей длине ряда. Это объясняется тем, что переходный эффект, вызванный выбором начальных значений, практически исчезает после $t = 12$. С другой стороны, когда учтено, что условная сумма в (2) основана на 11, а не на 12 квадратах величин, (2) дает лучшее приближение, чем (1). По уже обсуждавшимся причинам, если нужно пользоваться условным приближением, лучше брать его в форме (2), а не в форме (1). Однако, как будет показано в гл. 9, для сезонных рядов условная аппроксимация вообще неудовлетворительна, и следует, как правило, вычислять безусловную сумму квадратов.

7.1.6. Графическое исследование суммы квадратов

Сумма квадратов $S(\theta)$ для данных IBM, приведенных в табл. 7.2, показана на рис. 7.1. Минимальная сумма квадратов получается при $\theta = -0,09$ ($\lambda = 1,09$); это дает оценку наименьших квадратов, являющуюся хорошим приближением (в предположении о нормальности) к оценке максимального правдоподобия параметра θ .

Графический анализ сумм квадратов легко обобщается на случай двух параметров путем вычисления сумм квадратов по подходящей сетке значений параметров и вычерчивания изолиний. Как мы обсуждали в разд. 7.1.4, в предположении

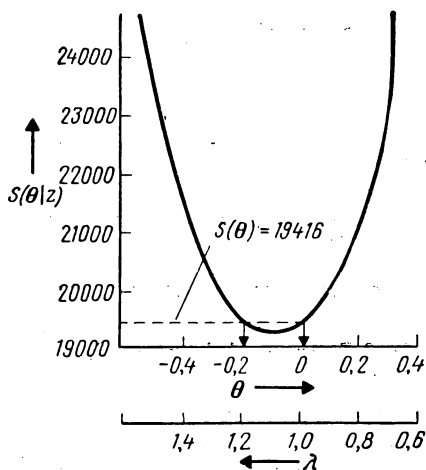


Рис. 7.1. График $S(\theta)$ для ряда В.

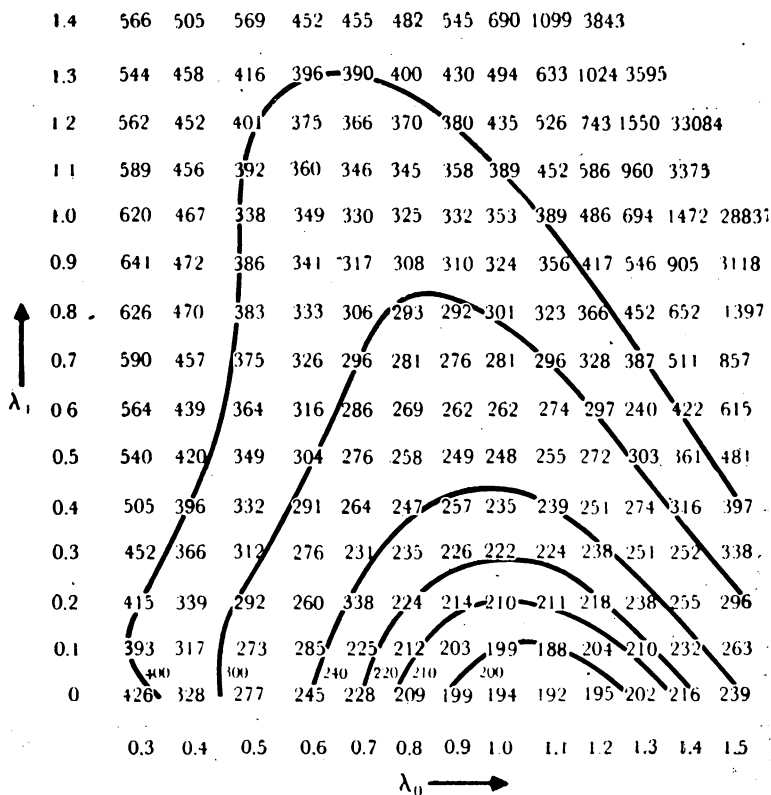


Рис. 7.2. Значения $S(\lambda_0, \lambda_1) \times 10^{-2}$ для ряда В в узлах сетки λ_0, λ_1 и приближенные изолинии.

о нормальности поведение этих изолиний почти такое же, как у функции максимального правдоподобия. Для большинства практических задач вполне пригодны приближенные карты изолиний, построенные на глаз по выходной таблице ЭВМ. На рис. 7.2 показана таблица значений $S(\lambda_0, \lambda_1)$ для ряда B , к которому был подогнан процесс ПСС(0, 2, 2)

$$\nabla^2 z_t = (1 - \theta_1 B - \theta_2 B^2) a_t = \{1 - (2 - \lambda_0 - \lambda_1) B - (\lambda_0 - 1) B^2\} a_t \quad (7.1.17)$$

или, в просуммированном виде,

$$z_t = \lambda_0 S a_{t-1} + \lambda_1 S^2 a_{t-1} + a_t.$$

Минимальное значение суммы квадратов на рис. 7.2 достигается при $\hat{\lambda}_0 \approx 1,09$, $\hat{\lambda}_1 \approx 0,0$. График подтверждает, таким образом, что лучшей моделью в этом случае является процесс

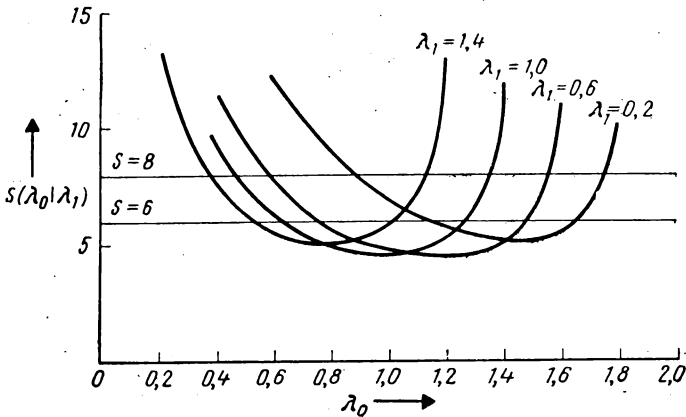


Рис. 7.3. Графики $S(\lambda_0, \lambda_1)$ как функции λ_0 для фиксированных значений λ_1 , используемые при построении изолиний для ряда S .

ПСС(0, 1, 1). Проиллюстрированный здесь способ подгонки несколько более сложными моделями, чем нужно, может давать полезное подтверждение правильности первоначальной идентификации. Усовершенствования модели должны, конечно, делаться в тех направлениях, где ожидается их необходимость.

Проведение изолиний. Устройства для автоматического построения графиков и карт по выходным таблицам ЭВМ становятся все более распространенными. С их помощью легко получать карты изолиний по значениям функции на сетке. Если необходимо построить изолинии аккуратно от руки, можно получить серию графиков $S(\lambda_0, \lambda_1)$ с аргументом λ_0 при различных значениях λ_1 . Проведя горизонтальные линии на выбранных

уровнях S , для которых будут строиться изолинии, перенесем точки пересечения на плоскость карты. Если необходимо, можно уточнить карту, повторив все построения для S как функции λ_1

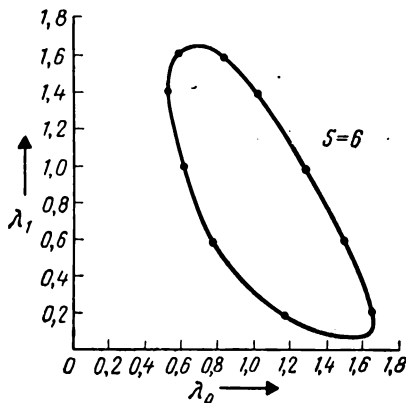


Рис. 7.4. Построение изолинии $S = 6$ для ряда С.

при различных фиксированных значениях λ_0 . Методика построения иллюстрируется табл. 7.6 и рис. 7.3 и 7.4 для процесса ПСС(0, 2, 2), подгоняемого к данным ряда С. В табл. 7.6 показана сетка λ_0, λ_1 со значениями суммы квадратов, а на

Таблица 7.6

Значения суммы квадратов в узлах сетки для ряда С
в предположении, что модель порядка (0, 2, 2)

λ_1	λ_0										
	0,0	0,2	0,4	0,6	0,8	1,0	1,2	1,4	1,6	1,8	2,0
2,0	381	11,5	7,55	6,60	9,00						
1,8	217	11,1	7,08	5,87	6,51	14,0					
1,6	571	12,3	6,92	5,52	5,40	7,45	881				
1,4	277	13,8	7,24	5,41	4,89	5,56	12,1				
1,2	949	15,9	7,94	5,56	4,72	4,75	6,55	649			
1,0	180	20,3	9,02	5,99	4,79	4,44	5,03	11,0			
0,8	761	19,7	10,1	6,61	5,09	4,44	4,47	6,05	497		
0,6	474	23,3	11,4	7,48	5,65	4,72	4,40	4,87	10,3		
0,4	443	26,8	13,6	8,97	6,67	5,39	4,73	4,66	6,00	394	
0,2	887	41,3	19,7	12,5	8,91	6,90	5,74	5,18	5,40	10,2	
0,0	3526	160	55,5	28,6	17,7	12,2	9,19	7,42	6,58	7,30	321

рис. 7.3 показаны графики $S(\lambda_0, \lambda_1)$ как функции λ_0 при различных значениях λ_1 . Абсциссы пересечения с линией $S(\lambda_0, \lambda_1) = 6$ используются для построения изолинии на рис. 7.4.

Три параметра. Если мы хотим изучить возможности совместной оценки трех параметров, можно построить набор двумерных диаграмм для различных фиксированных значений третьего параметра. В качестве примера на рис. 7.5 показаны некоторые диаграммы из соответствующих наборов для рядов А, С и D. В каждом случае подгонялась достаточно «сложная» модель

$$\nabla^2 z_t = (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3) a_t = \\ = \{1 - (2 - \lambda_{-1} - \lambda_0 - \lambda_1) B - (\lambda_0 + 2\lambda_{-1} - 1) B^2 + \lambda_{-1} B^3\} a_t$$

или

$$z_t = \lambda_{-1} a_{t-1} + \lambda_0 S a_{t-1} + \lambda_1 S^2 a_{t-1} + a_t.$$

Результирующие лучшие модели этого типа показаны в табл. 7.7 *).

Таблица 7.7

Модели ПСС, подогнанные к рядам А, С, D

Ряд	$\hat{\lambda}_{-1}$	$\hat{\lambda}_0$	$\hat{\lambda}_1$	Подогнанный ряд
А	0	0,3	0	$z_t = 0,3 S a_{t-1} + a_t$
С	0	1,1	0,8	$z_t = 1,1 S a_{t-1} + 0,8 S^2 a_{t-1} + a_t$
D	0	0,9	0	$z_t = 0,9 S a_{t-1} + a_t$

Включение дополнительного параметра (в частности, λ_{-1}) в процесс подгонки не являлось строго необходимым; это было сделано для иллюстрации эффекта «переусложнения» модели и того, как хорошо оправдывается наша идентификация этих рядов.

7.1.7. Описание «благоприятных» оценочных ситуаций; доверительные области

Функция правдоподобия строится, конечно, не только для того, чтобы показать максимальные значения правдоподобия. Эта функция в целом содержит весь объем информации, приходящей с данными. В некоторых проблемах возникает ситуация, когда функция правдоподобия имеет два и более максимума

*) Позднее, в разд. 7.2.5, будет показано, что несколько лучшее согласие можно в ряде случаев получить для очень близких к рассматриваемым моделей, содержащих члены с авторегрессией.

Необходимость соблюдения осторожности при интерпретации функции правдоподобия. Необходимо быть осторожным при интерпретации функции правдоподобия. Например, результаты, рассмотренные ниже и основанные на предположении, что логарифмическая функция правдоподобия вблизи ее максимума близка к квадратичной функции, явно не применимы к ситуациям оценивания трех параметров, показанным на рис. 7.5. Однако эти примеры не типичны, поскольку мы умышленно *перусложнили* модель. Если справедлива более простая модель, мы должны *ожидать*, что изолинии функций правдоподобия будут

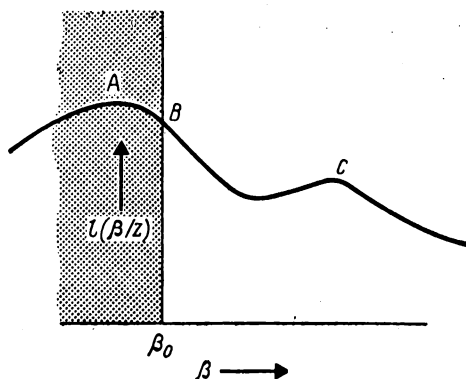


Рис. 7.6. Гипотетическая функция правдоподобия с ограничением $\beta > \beta_0$.

усечены вблизи максимума некоторой границей в пространстве параметров более высокой размерности. Квадратичная модель *может* быть использована, если подгонялась более простая *идентифицированная*, а не излишне усложненная модель.

Особая осторожность необходима, когда максимум функции правдоподобия может лежать на границе или около нее. Рассмотрим ситуацию, показанную на рис. 7.6, и положим известным априори, что параметр $\beta > \beta_0$. Максимальное правдоподобие внутри допустимой области значения β достигается в точке B, где $\beta = \beta_0$, а не в A или C. Отметим, что первая производная функции правдоподобия в этом случае *не равна нулю* в максимуме и квадратичная аппроксимация определенно не является адекватным представлением функции правдоподобия.

Использование метода правдоподобия в прошлом часто было не слишком умелым, и неопытность потребителя иногда ошибочно истолковывалась как слабость метода. Это использование обычно включало

1) дифференцирование логарифмической функции правдоподобия и приравнивание первых производных нулю для получения оценок максимального правдоподобия (МП);

2) получение приближенных дисперсий и ковариаций этих оценок по вторым производным логарифмической функции правдоподобия или их математическим ожиданиям.

Формальное применение этих приемов может привести к бессмысленным результатам. Это, во-первых, связано с тем элементарным фактом, что приравнивание производных нулю может не приводить нас к максимуму, и, во-вторых, с тем, что информация, содержащаяся в функции правдоподобия, выражается полностью оценками МП и вторыми производными этой функции, только если в интересующей нас области применимо квадратичное приближение. Узнать, выполняется ли это условие для новой задачи оценивания, можно обычно только путем тщательного аналитического и графического исследования.

Когда исследуется некоторый новый класс задач оценивания (например, возникающих при оценивании параметров модели АРСС), следует широко пользоваться графическими представлениями функции правдоподобия. После того как поведение данного класса моделей достаточно хорошо понято и представления о ситуации указывают на допустимость такого подхода, мы можем действовать более прямыми методами, к рассмотрению которых мы переходим. Эти результаты подробно описаны в приложениях П7.4 и П7.5. Мы начнем с рассмотрения выражений для дисперсий и ковариаций оценок максимального правдоподобия, пригодных в случае, когда логарифмическая функция правдоподобия приближенно квадратичная и размер выборки сравнительно велик.

В последующем изложении удобно ввести вектор β , $k = p + q$ элементов которого — параметры авторегрессии и скользящего среднего ϕ и θ . Тогда полный набор $p + q + 1 = k + 1$ параметров процесса АРСС может быть представлен как ϕ , θ , σ_a , или как β , σ_a , или просто как ξ .

Дисперсии и ковариации оценок МП. Для правильно параметризованной модели АРСС часто оказывается, что в пределах значимой области пространства параметров *) логарифмическая функция правдоподобия будет примерно квадратичной формой от компонент β (т. е. ϕ и θ), так что

$$l(\xi) = l(\beta, \sigma_a) \approx l(\hat{\beta}, \sigma_a) + \frac{1}{2} \sum_{i=1}^k \sum_{j=1}^k l_{ij} (\beta_i - \hat{\beta}_i) (\beta_j - \hat{\beta}_j), \quad (7.1.18)$$

*) Скажем, в пределах 95%-ной доверительной области.

где в рассматриваемом приближении производные

$$l_{ij} = \frac{\partial^2 l(\beta, \sigma_a)}{\partial \beta_i \partial \beta_j} \quad (7.1.19)$$

постоянны.

При больших n влиянием члена $f(\phi, \theta)$ или, что то же самое, $f(\beta)$ в (7.1.5) в большинстве случаев можно пренебречь. Отсюда, если $S(\beta)$ близко к квадратичной форме, и $l(\beta, \sigma_a)$ также будет близко к ней. Иначе говоря, $l(\beta, \sigma_a)$ будет близко к квадратичной форме от β , если условные математические ожидания $[a_i | \beta, w]$ в (7.1.6) локально близки к линейным функциям компонент вектора β .

Для средних и больших выборок при условии, что допустимо локальное квадратичное приближение (7.1.18), можно получить полезные приближенные выражения для дисперсий и ковариаций оценок и для доверительных интервалов.

Информационная матрица для параметров β . Матрица $E[l_{ij}] = I(\beta)$ размером $(k \times k)$ названа в [56, 87] *информационной матрицей* для параметров β ; здесь математическое ожидание взято по распределению w . Для данного значения σ_a матрица ковариаций $V(\hat{\beta})$ МП-оценок $\hat{\beta}$ при больших размерах выборки является обратной к информационной матрице

$$V(\hat{\beta}) \approx \{-E[l_{ij}]\}^{-1}. \quad (7.1.20)$$

Например, если $k = 2$, матрица ковариаций для большой выборки имеет вид

$$V(\hat{\beta}) = \begin{bmatrix} V(\hat{\beta}_1) & \text{cov}(\hat{\beta}_1, \hat{\beta}_2) \\ \text{cov}(\hat{\beta}_1, \hat{\beta}_2) & V(\hat{\beta}_2) \end{bmatrix} \approx - \begin{bmatrix} E[l_{11}] & E[l_{12}] \\ E[l_{12}] & E[l_{22}] \end{bmatrix}^{-1}.$$

Далее, используя (7.1.5), получаем

$$l_{ij} \approx - \frac{S_{ij}}{2\sigma_a^2}, \quad (7.1.21)$$

где

$$S_{ij} = \frac{\partial^2 S(\beta | w)}{\partial \beta_i \partial \beta_j}.$$

Кроме того, если для больших выборок мы аппроксимируем математические ожидания величин l_{ij} или S_{ij} фактически наблюдаемыми значениями, то, используя (7.1.20), получим

$$V(\hat{\beta}) \approx \{-E[l_{ij}]\}^{-1} \approx 2\sigma_a^2 \{E[S_{ij}]\}^{-1} \approx 2\sigma_a^2 \{S_{ij}\}^{-1}. \quad (7.1.22)$$

Отсюда для $k = 2$

$$V(\hat{\beta}) \approx 2\sigma_a^2 \begin{bmatrix} \frac{\partial^2 S(\beta)}{\partial \beta_1^2} & \frac{\partial^2 S(\beta)}{\partial \beta_1 \partial \beta_2} \\ \frac{\partial^2 S(\beta)}{\partial \beta_1 \partial \beta_2} & \frac{\partial^2 S(\beta)}{\partial \beta_2^2} \end{bmatrix}^{-1}.$$

Если $S(\beta)$ была бы точно квадратичной функцией β в рассматриваемой области пространства параметров, то все производные S_{ij} были бы постоянны во всей этой области. На практике S_{ij} будут несколько варьировать, и мы будем предполагать, что эти производные определяются в точке $\hat{\beta}$ либо вблизи нее. В приложениях П7.4 и П7.5 показано, что оценка *) σ_a^2 имеет вид

$$\hat{\sigma}_a^2 = S(\hat{\beta})/n \quad (7.1.23)$$

и что для больших выборок $\hat{\sigma}_a^2$ и $\hat{\beta}$ не коррелированы. Наконец, элементы (7.1.22) можно оценить, пользуясь тем, что

$$\text{cov}(\hat{\beta}_i, \hat{\beta}_j) \approx 2\hat{\sigma}_a^2 S^{ij}, \quad (7.1.24)$$

где матрица $\{S^{ij}\}$ равна

$$\{S^{ij}\} = \{S_{ij}\}^{-1}.$$

Выражение (7.1.24) при $j = i$ определяет дисперсию $V(\hat{\beta}_i)$.

Приближенные доверительные области для параметров. Эти результаты позволяют, в частности, получить приближенные значения дисперсии наших оценок. Извлекая корень из дисперсий, мы найдем стандартные отклонения, которые обычно называют *стандартными ошибками* оценок. Стандартную ошибку оценки $\hat{\beta}_i$ мы будем обозначать $\text{CO}[\hat{\beta}_i]$. Когда нужно рассматривать несколько параметров одновременно, необходимо иметь способ суждения о *совместной* точности оценок. Один из способов осуществления такого подхода — определение *доверительной области*. Можно показать (см., например, [61]), что $(1 - \varepsilon)$ -я доверительная область имеет следующее свойство: если делать повторные выборки размера n из той же популяции и строить доверительную область для каждой такой выборки, то $(1 - \varepsilon)$ -я часть этих областей будет включать точку с истинным значением параметра.

Если для данного σ_a $l(\beta, \sigma_a)$ — приближенно квадратичная функция β вблизи $\hat{\beta}$, то, пользуясь (7.1.20) (см. также приложение П7.1), можно получить приближенную $(1 - \varepsilon)$ -ю доверительную область в виде

$$-\sum_{i,j} E[l_{ij}](\beta_i - \hat{\beta}_i)(\beta_j - \hat{\beta}_j) < \chi_{\varepsilon}^2(k), \quad (7.1.25)$$

где $\chi_{\varepsilon}^2(k)$ — это квантиль уровня $(1 - \varepsilon)$ распределения χ^2 с k степенями свободы. Можно получить и другое приближенное

*) Можно высказать ряд доводов в пользу того, чтобы в знаменателе (7.1.23) вместо n стояло бы $n - k = n - p - q$, но для не слишком малых размеров выборки эта замена практически не влияет на результат.

выражение для доверительной области, если воспользоваться (7.1.22) и подставить туда оценку (7.1.23) вместо σ_a^2 *):

$$\sum_{i,j} S_{ij} (\beta_i - \hat{\beta}_i) (\beta_j - \hat{\beta}_j) < 2\hat{\sigma}_a^2 \chi_e^2(k). \quad (7.1.26)$$

Однако если поверхность $S(\beta)$ квадратичная, то

$$S(\beta) - S(\hat{\beta}) = \frac{1}{2} \sum_{i,j} S_{ij} (\beta_i - \hat{\beta}_i) (\beta_j - \hat{\beta}_j). \quad (7.1.27)$$

Тогда, пользуясь (7.1.23) и (7.1.26), находим, что приближенная $(1 - \epsilon)$ -я доверительная область ограничена изолинией суммы квадратов, для которой

$$S(\beta) = S(\hat{\beta}) \left\{ 1 + \frac{\chi_e^2(k)}{n} \right\}. \quad (7.1.28)$$

Примеры вычисления приближенных доверительных интервалов и областей.

1) Ряд В. Значения $S(\lambda)$ и ее разностей для ряда В приведены в табл. 7.8. Вторая разность $S(\lambda)$ не постоянна, и поэтому $S(\lambda)$ не является строго квадратичной функцией. Однако на интервале от $\lambda = 0,85$ до $\lambda = 1,35$ $\nabla^2(S)$ изменяется

Таблица 7.8

**$S(\lambda)$ и ее первые и вторые разности
для различных значений λ , ряд В**

$\lambda = 1 - \theta$	$S(\lambda)$	$\nabla(S)$	$\nabla^2(S)$
1,5	23 929	2 334	961
1,4	21 595	1 373	634
1,3	20 222	739	476
1,2	19 483	263	406
1,1	19 220	—143	390
1,0	19 363	—533	422
0,9	19 896	—955	509
0,8	20 851	—1464	692
0,7	22 315	—2156	1067
0,6	24 471	—3223	
0,5	27 694		

*) Можно вывести несколько более точное приближение, основанное на F -распределении Фишера и учитывающее выборочное распределение $\hat{\sigma}_a^2$. Однако это уточнение практически несущественно, если размер выборки не слишком мал.

незначительно, так что можно считать (7.1.28) достаточно хорошим приближением. Так как минимальное значение $S(\hat{\lambda}) = 19\,216$, то критическое значение $S(\lambda)$, определяющее приближенно 95 %-ный доверительный интервал, равно

$$S(\lambda) = 19\,216 \left\{ 1 + \frac{3,84}{369} \right\} = 19\,416.$$

Считывая значение λ , соответствующее $S(\lambda) = 19\,416$ на рис. 7.1, получаем приближенный доверительный интервал $0,98 < \lambda < 1,19$.

Можно поступить и по-другому, используя (7.1.26). Пользуясь второй разностью из табл. 7.8 при $\lambda = 1,1$, чтобы оценить производную, получаем

$$S_{11} = \frac{\partial^2 S}{\partial \lambda^2} \approx \frac{390}{(0,1)^2}.$$

Пользуясь (7.1.23), находим, что $\hat{\sigma}_a^2 = 19\,216/369 = 52,3$. Отсюда 95 %-ный доверительный интервал, согласно (7.1.26), будет

$$\frac{390}{(0,1)^2} (\lambda - 1,09)^2 < 2 \times 52,3 \times 3,84,$$

т. е.

$$|\lambda - 1,09| < 0,10.$$

Отсюда интервал равен $0,99 < \lambda < 1,19$, что хорошо согласуется с предыдущими вычислениями.

В этом примере, где имеется только один параметр λ , использование (7.1.25) и (7.1.26) эквивалентно использованию интервала

$$\hat{\lambda} \pm u_{\varepsilon/2} \hat{\sigma}(\hat{\lambda}),$$

где $u_{\varepsilon/2}$ — квантиль уровня $(1 - \varepsilon/2)$ стандартного нормального распределения. Приближенная стандартная ошибка $\hat{\sigma}(\hat{\lambda}) = \sqrt{2\hat{\sigma}_a^2 S_{11}^{-1}}$ находится из (7.1.24). В этом примере

$$V(\hat{\lambda}) = 2\hat{\sigma}_a^2 S_{11}^{-1} = \frac{2 \times 52,3 \times 0,1^2}{390} = 0,00268,$$

и приближенная стандартная ошибка $\hat{\lambda}$ равна

$$\hat{\sigma}(\hat{\lambda}) = \sqrt{\text{var}(\hat{\lambda})} = 0,052.$$

Отсюда приближенный доверительный интервал равен, как и ранее,

$$\hat{\lambda} \pm 1,96\hat{\sigma}(\hat{\lambda}) = 1,09 \pm 0,10.$$

Наконец, в разд. 7.2 мы покажем, что для больших выборок процесса СС(1) можно оценить (7.1.20) аналитически, что дает

$$V(\hat{\lambda}) \approx \frac{\lambda(2-\lambda)}{n}.$$

В рассматриваемом примере, заменяя λ на $\hat{\lambda} = 1,09$, находим

$$V(\hat{\lambda}) \approx 0,00269,$$

что хорошо согласуется с предыдущими оценками, дает ту же стандартную ошибку 0,052 и тот же доверительный интервал.

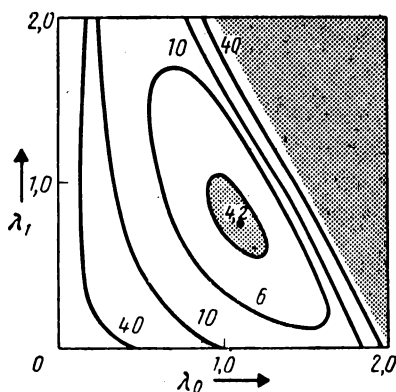


Рис. 7.7. Изолинии суммы квадратов для ряда С в предположении модели порядка (0,2,2); в центре заштрихована 95%-ная доверительная область.

2) Ряд С. При идентификации ряда С в качестве одной из опробованных моделей использовался процесс (0, 2, 2). Для иллюстрации использования формулы (7.1.28) в случае, когда число параметров больше 1, на рис. 7.7 заштрихована 95 %-ная доверительная область значения параметров λ_0 и λ_1 . В этом примере $S(\hat{\lambda}) = 4,20$; $n = 226$, $\chi^2_{0,05}(2) = 5,99$, так что 95 %-ная доверительная область ограничена изолинией

$$S(\lambda_0, \lambda_1) = 4,20 \left\{ 1 + \frac{5,99}{226} \right\} = 4,31.$$

7.2. НЕЛИНЕЙНОЕ ОЦЕНИВАНИЕ

7.2.1. Общий подход

Графическое представление суммы квадратов особенно важно при исследовании новых задач оценивания, так как позволяет выявить любые специфические особенности ситуации. Когда появляется уверенность, что какие-нибудь аномальные

явления маловероятны, можно воспользоваться другими методами.

Мы видели, что в большинстве случаев оценки максимального правдоподобия хорошо приближаются оценками наименьших квадратов, которые минимизируют

$$S(\phi, \theta) = \sum_{t=-\infty}^n [a_t | \phi, \theta, w]^2 = \sum_{t=-\infty}^n [a_t]^2.$$

На практике бесконечную сумму можно заменить более удобной конечной суммой $\sum_{t=1-Q}^n [a_t]^2$.

В общем значительные упрощения при минимизации по β суммы квадратов $\sum_{t=1}^n \{f_t(\beta)\}^2$ происходят, если каждое $f_t(\beta)$ ($t = 1, 2, \dots, n$) — линейная функция параметров β . Покажем теперь, что линейность $[a_t]$ имеет несколько различных смыслов по отношению к параметрам авторегрессии ϕ и параметрам скользящего среднего θ .

Для чистого процесса авторегрессии $[a_t] = \phi(B)[\tilde{w}_t]$ и

$$\frac{\partial [a_t]}{\partial \phi_i} = -[\tilde{w}_{t-i}] + \phi(B) \frac{\partial [\tilde{w}_t]}{\partial \phi_i}.$$

Для $u > 0$ $[\tilde{w}_u] = \tilde{w}_u$ и $\partial[\tilde{w}_u]/\partial \phi_i = 0$, в то же время для $u \leq 0$ как $[\tilde{w}_u]$, так и $\partial[\tilde{w}_u]/\partial \phi_i$ являются функциями ϕ . Поэтому, если пренебречь эффектом «начальных значений», $[a_t]$ линейны относительно ϕ . Напротив, для чистого процесса скользящего среднего

$$[a_t] = \theta^{-1}(B)[\tilde{w}_t], \quad \frac{\partial [a_t]}{\partial \theta_j} = -\theta^{-2}(B)[\tilde{w}_{t-j}] + \theta^{-1}(B) \frac{\partial [\tilde{w}_t]}{\partial \theta_j},$$

так что $[a_t]$ — всегда нелинейные функции параметров.

В разд. 7.3 мы увидим, что для процесса авторегрессии оценки наименьших квадратов и максимального правдоподобия существенно упрощаются. В данном разделе будет показано, как путем итеративного применения линейного метода наименьших квадратов можно найти оценки для любой модели АРСС.

Линеаризация модели. В дальнейшем символ β будет по-прежнему использоваться как общий символ для $k = p + q$ параметров (ϕ, θ) . Наша задача — минимизировать

$$\sum_{t=1-Q}^n [a_t | \tilde{w}, \beta]^2 = \sum_{t=1-Q}^n [a_t]^2.$$

Разлагая $[a_t]$ в ряд Тейлора в окрестности точки, соответствующей предполагаемым значениям параметров $\beta'_0 = (\beta_{1,0}, \beta_{2,0}, \dots$

..., $\beta_{k,0}$), получаем приближенное выражение

$$[a_t] = [a_{t,0}] - \sum_{i=1}^k (\beta_i - \beta_{i,0}) x_{i,t}, \quad (7.2.1)$$

где

$$[a_{t,0}] = [a_t | \mathbf{w}, \beta_0]$$

и

$$x_{i,t} = - \left. \frac{\partial [a_t]}{\partial \beta_i} \right|_{\beta=\beta_0}.$$

Если \mathbf{X} — матрица $\{x_{i,t}\}$ размера $(n+Q) \times k$, то $n+Q$ уравнений (7.2.1) можно представить в виде

$$[\mathbf{a}_0] = \mathbf{X}(\beta - \beta_0) + [\mathbf{a}],$$

где $[\mathbf{a}_0]$ и $[\mathbf{a}]$ — векторы-столбцы с $n+Q$ элементами. Поправка $\beta - \beta_0$, минимизирующая $S(\beta) = S(\phi, \theta) = [\mathbf{a}]' [\mathbf{a}]$, находится теперь линейным методом наименьших квадратов, т. е. регрессией $[\mathbf{a}_0]$ на \mathbf{x} . Поскольку $[a_t]$ не точно линейная функция параметров β , однократная поправка не обеспечит получение оценок наименьших квадратов. Поэтому полученные после первой поправки значения используются как новые предполагаемые значения, и процедура повторяется до получения сходящихся результатов. Сходимость оказывается более быстрой, если используются достаточно удачные начальные приближения, которые можно получить на этапе идентификации. Если начальные приближения неудачны, процесс вычисления параметров может оказаться расходящимся.

7.2.2. Численный метод нахождения производных

Как будет показано позднее, производные $x_{i,t}$ могут быть получены непосредственно. Однако для машинных расчетов оказался весьма эффективным общий нелинейный метод наименьших квадратов, в котором производные оцениваются численно. Это делается поочередным возмущением одного из параметров. Затем для данной модели вычисляются рекуррентным образом значения $[a_t | \mathbf{w}, \beta_{1,0}, \beta_{2,0}, \dots, \beta_{k,0}]$ для $t = 1 - Q, \dots, n$; при этом используется необходимое число значений «прогнозов назад». Затем вычисления повторяются для $[a_t | \mathbf{w}, \beta_{1,0} + \delta_1, \beta_{2,0}, \dots, \beta_{k,0}]$, затем для $[a_t | \mathbf{w}, \beta_{1,0}, \beta_{2,0} + \delta_2, \dots, \beta_{k,0}]$ и т. д. Величина, обратная по знаку требуемой производной, находится затем достаточно точно по формуле

$$x_{i,t} = \{[a_t | \mathbf{w}, \beta_{1,0}, \dots, \beta_{i,0}, \dots, \beta_{k,0}] - [a_t | \mathbf{w}, \beta_{1,0}, \dots, \beta_{i,0} + \delta_i, \dots, \beta_{k,0}]\} / \delta_i. \quad (7.2.2)$$

Описанный численный метод получения производных универсален и требует только наличия программы для вычисления $[a_t]$. Общие стандартные программы нахождения нелинейных оценок, нуждающиеся только в дополнительной подпрограмме для вычисления $[a_t]$, сейчас уже широко распространены [62]. В некоторых вариантах программы δ необходимо указать заранее; в других программа сама выполняет необходимые итерации для выбора подходящего δ . Некоторые программы включают специальные логические схемы, позволяющие избежать проскоков области минимума и ускорить сходимость [63].

При условии, что решение наименьших квадратов не лежит вблизи границы или на границе области, значения $\mathbf{X} = \mathbf{X}_{\hat{\beta}}$ из последней итерации можно использовать для вычисления приближенных дисперсий, ковариаций и доверительных интервалов. Конкретнее, $(\mathbf{X}_{\hat{\beta}}'\mathbf{X}_{\hat{\beta}})^{-1}\sigma_a^2$ аппроксимирует матрицу ковариаций $\hat{\beta}$, а σ_a^2 оценивается по $S(\hat{\beta})/n$.

Применение к процессу (0, 1, 1). В качестве элементарного примера рассмотрим подгонку к ряду A процесса (0, 1, 1)

$$w_t = (1 - \theta B) a_t$$

с $\mu = E[w_t] = 0$. Начало вычислений для пробного значения $\theta_0 = 0,5$ показано в табл. 7.9. Значения прогноза назад для $[a_0]$ были получены приравниванием $[e_7] = 0$ и применением обратной рекуррентной формулы $[e_t] = [w_t] + \theta[e_{t+1}]$. Большая точность могла бы быть достигнута применением этой рекуррентной формулы к более далеким значениям ряда. Значения $[a_t]$, найденные последовательно из уравнения $[a_t] = \theta[a_{t-1}] + [w_t]$ для $\theta = 0,50$ и $\theta = 0,51$, приведены в четвертом и пятом столбцах вместе со значениями отрицательных производных x_t , найденных по формуле (7.2.2). Для получения первой поправки для θ вычислим

$$\theta - \theta_0 = \sum_{t=0}^n [a_{t,0}] x_t / \sum_{t=0}^n x_t^2.$$

В этом примере, используя весь ряд длиной в 197 наблюдений, мы получили сходящийся результат после четырех итераций. Расчет шел следующим образом:

Итерация	θ
0	0,50
1	0,63
2	0,68
3	0,69
4	0,70
5	0,70

Таблица 7.9

Пример расчета производных по данным ряда А

t	z_t	w_t	$[a_t, 0] = [a_t 0,50]$	$[a_t 0,51]$	$x_t = 100 \{[a_t 0,50] - [a_t 0,51]\}$
0	17,0		0,2453	0,2496	—0,43
1	16,6	—0,40	—0,2773	—0,2727	—0,46
2	16,3	—0,30	—0,4387	—0,4391	0,04
3	16,1	—0,20	—0,4193	—0,4239	0,46
4	17,1	1,00	0,7903	0,7838	0,65
5	16,9	—0,20	0,1952	0,1997	—0,45
6	16,8	—0,10	—0,0024	0,0019	—0,43
7	17,4	0,60	0,5988	0,6010	—0,22
8	17,1	—0,30	—0,0006	0,0065	—0,71
9	17,0	—0,10	—0,1003	—0,0967	—0,36
10	16,7	—0,30	—0,3502	—0,3493	—0,09

В общем значения ϕ и θ , минимизирующие $S(\phi, \theta)$, могут быть найдены этим методом с любой необходимой степенью точности. Этот метод особенно удобен потому, что нам не нужно специально программировать расчет производных; не требуется никаких дополнительных процедур, кроме как для вычисления $[a_t]$.

Программа 3, описанная в сборнике машинных программ в конце этой книги, позволяет численно оценивать производные и включает возможность прогнозирования назад. Она позволяет найти оценки наименьших квадратов для параметров любого (p, d, q) процесса.

Теперь мы покажем также, что производные можно получить и непосредственно, но для этого нужны дополнительные рекуррентные расчеты.

7.2.3. Прямой метод нахождения производных

Для того чтобы проиллюстрировать методику, достаточно рассмотреть процесс $APCC(1, 1)$, который можно представить как

$$e_t = w_t - \phi w_{t+1} + \theta e_{t+1}$$

либо как

$$a_t = w_t - \phi w_{t-1} + \theta a_{t-1}.$$

В разд. 7.1.4 мы видели, что можно пользоваться обоими представлениями модели попеременно — одно из них дает начальные значения, которые используются для рекуррентного счета

по другому представлению. Предположим, что первое вычисление уже было сделано и мы получили значения $[e_t]$, $[a_t]$, $[w_0]$, $[w_{-1}]$, ..., $[w_{1-Q}]$, как в разд. 7.1.5; значениями $[w_{-Q}]$, $[w_{-Q-1}]$, ..., а значит, и $[a_{-Q}]$, $[a_{-Q-1}]$, ..., можно пренебречь. Покажем, как можно использовать такой двойственный подход для вычисления производных.

Обозначая частную производную $\partial[a_t]/\partial\phi$ как $a_t^{(\phi)}$, получим

$$e_t^{(\phi)} = w_t^{(\phi)} - \phi w_{t+1}^{(\phi)} + \theta e_{t+1}^{(\phi)} - [w_{t+1}], \quad (7.2.3)$$

$$a_t^{(\phi)} = w_t^{(\phi)} - \phi w_{t-1}^{(\phi)} + \theta a_{t-1}^{(\phi)} - [w_{t-1}], \quad (7.2.4)$$

$$e_t^{(\theta)} = w_t^{(\theta)} - \phi w_{t+1}^{(\theta)} + \theta e_{t+1}^{(\theta)} + [e_{t+1}], \quad (7.2.5)$$

$$a_t^{(\theta)} = w_t^{(\theta)} - \phi w_{t-1}^{(\theta)} + \theta a_{t-1}^{(\theta)} + [a_{t-1}]. \quad (7.2.6)$$

Имеем

$$\begin{aligned} [w_t] &= w_t; \\ w_t^{(\phi)} &= w_t^{(\theta)} = 0, \end{aligned} \quad t = 1, 2, \dots, n, \quad (7.2.7)$$

и

$$[e_{-j}] = 0, \quad j = 0, 1, \dots, n. \quad (7.2.8)$$

Рассмотрим (7.2.3) и (7.2.4). Приняв в (7.2.3) $e_{n+1}^{(\phi)} = 0$, можем начать рекуррентный расчет назад, который при помощи (7.2.7) и (7.2.8) позволит нам в результате найти $w_{-j}^{(\phi)}$ для $j = 0, 1, \dots, Q - 1$. Так как $a_{-Q}^{(\phi)}$, $a_{-Q-1}^{(\phi)}$, ... можно принять равными нулю, мы можем теперь использовать (7.2.4) для рекуррентного вычисления требуемых производных $a_t^{(\phi)}$. Аналогичным образом можно использовать (7.2.5) и (7.2.6) для вычисления производных $a_t^{(\theta)}$.

Рассмотрим опять в качестве примера расчет производных $x_t = -d[a_t]/d\theta$ для первой части ряда А, выполненный полностью численным путем в табл. 7.9. В табл. 7.10 приведен соответствующий расчет по формулам

$$\begin{aligned} e_t^{(\theta)} &= w_t^{(\theta)} + \theta e_{t+1}^{(\theta)} + [e_{t+1}], \\ -x_t &= a_t^{(\theta)} = w_t^{(\theta)} + \theta a_{t-1}^{(\theta)} + [a_{t-1}]. \end{aligned}$$

Значения $[a_t]$ и $[e_t]$, которые уже были вычислены, введены в таблицу, и вычисление $e_t^{(\theta)}$ начинается в этом примере приравнованием $e_6^{(\theta)} = 0$. Видно, что значения x_t очень близки к полученным в табл. 7.9 только численным способом.

7.2.4. Общий алгоритм наименьших квадратов для условной модели

При анализе длинных рядов мы нередко используем приближение, заключающееся в приравнивании начальных значений a , а следовательно, и x их безусловным математическим ожиданиям, т. е. нулю, и затем в непосредственном использовании прямой рекуррентной формулы. Так, в предыдущем примере можно использовать уравнения

$$a_t = \theta a_{t-1} + w_t, \quad x_t = -a_t^{(\theta)} = \theta x_{t-1} - a_{t-1}.$$

В результате в ряды a_t и x_t вносится переходный процесс, для ряда x_t этот процесс затухает несколько медленнее, так как x зависит от a . В качестве иллюстрации этим способом были вычислены значения a и x для рассмотренного примера. Оказалось, что, хотя в начале имеются расхождения, начиная с $t = 4$ для a_t и с $t = 8$ для x_t два знака после запятой совпадают. В некоторых примерах, где имеется большой объем данных (скажем, 200 и более наблюдений) за счет некоторой потери информации (отбрасывания примерно 20 первых вычисленных значений) можно добиться очень высокой точности.

Если мы примем описанное выше приближение, то придем к интересному общему алгоритму для этой условной модели. Общая модель может быть представлена в виде

$$a_t = \theta^{-1}(B) \phi(B) \tilde{w}_t,$$

Таблица 7.10

Пример рекуррентного расчета производных для данных ряда А

t	$[a_{t-1}]$	$\theta a_{t-1}^{(\theta)}$	$x_t = -a_t^{(\theta)}$	$e_t^{(\theta)}$	$\theta e_{t+1}^{(\theta)}$	$[e_{t+1}]$
0			-0,43	$(-w_0^{(\theta)} = -0,43)$	0,06	-0,49
1	0,25	0,22	-0,47	0,12	0,30	-0,18
2	-0,28	0,24	0,04	0,60	0,36	0,24
3	-0,44	-0,02	0,46	0,73	-0,15	0,88
4	-0,42	-0,23	0,65	-0,30	-0,05	-0,25
5	0,79	-0,33	-0,46	-0,10	0	-0,10
6	0,20	0,23	-0,43	0		
7	0,00	0,22	-0,22			
8	0,60	0,11	-0,71			
9	0,00	0,35	-0,35			
10	-0,10	-0,18	-0,08			

где $w_t = \nabla^d z_t$, $\tilde{w}_t = w_t - \mu$ и

$$\begin{aligned}\theta(B) &= 1 - \theta_1 B - \dots - \theta_l B^l - \dots - \theta_q B^q, \\ \phi(B) &= 1 - \phi_1 B - \dots - \phi_l B^l - \dots - \phi_p B^p.\end{aligned}$$

Если начальные приближения для параметров $\beta = (\phi, \theta)$ равны $\beta_0 = (\phi_0, \theta_0)$, то

$$a_{t,0} = \theta_0^{-1}(B) \phi_0(B) \tilde{w}_t$$

и

$$-\left. \frac{\partial a_t}{\partial \phi_j} \right|_{\beta_0} = u_{j,t} = u_{t-j}, \quad -\left. \frac{\partial a_t}{\partial \theta_i} \right|_{\beta_0} = x_{i,t} = x_{t-i},$$

где

$$u_t = \theta_0^{-1}(B) \tilde{w}_t = \phi_0^{-1}(B) a_{t,0}, \quad (7.2.9)$$

$$x_t = -\theta_0^{-2}(B) \phi_0(B) \tilde{w}_t = -\theta_0^{-1}(B) a_{t,0}. \quad (7.2.10)$$

Значения a , x и u можно вычислять рекуррентным образом, положив начальные значения a , x и u равными нулю, а именно

$$a_{t,0} = \tilde{w}_t - \phi_{1,0} \tilde{w}_{t-1} - \dots - \phi_{p,0} \tilde{w}_{t-p} + \theta_{1,0} a_{t-1,0} + \dots + \theta_{q,0} a_{t-q,0}, \quad (7.2.11)$$

$$u_t = \theta_{1,0} u_{t-1} + \dots + \theta_{q,0} u_{t-q} + \tilde{w}_t = \quad (7.2.12)$$

$$= \phi_{1,0} u_{t-1} + \dots + \phi_{p,0} u_{t-p} + a_{t,0}, \quad (7.2.13)$$

$$x_t = \theta_{1,0} x_{t-1} + \dots + \theta_{q,0} x_{t-q} - a_{t,0}. \quad (7.2.14)$$

В соответствии с (7.2.1) приближенное уравнение линейной регрессии принимает вид

$$a_{t,0} = \sum_{j=1}^p (\phi_j - \phi_{j,0}) u_{t-j} + \sum_{i=1}^q (\theta_i - \theta_{i,0}) x_{t-i} + a_t. \quad (7.2.15)$$

Получаемые отсюда поправки — это коэффициенты регрессии $a_{t,0}$ на u_{t-j} и x_{t-i} . Прибавляя поправки к приближениям (ϕ_0, θ_0) , получаем «вторые приближения» и заменяем ими первые во второй итерации, в которой вновь вычисляются значения $a_{t,0}$, x_t и u_t до тех пор, пока результаты не начнут сходиться.

Другой вариант алгоритма. Выражение (7.2.15) можно также представить в виде

$$\begin{aligned}a_{t,0} &= \sum_{j=1}^p (\phi_j - \phi_{j,0}) B^j \phi_0^{-1}(B) a_{t,0} - \sum_{i=1}^q (\theta_i - \theta_{i,0}) B^i \theta_0^{-1}(B) a_{t,0} + a_t \\ &= -\{\phi(B) - \phi_0(B)\} \phi_0^{-1}(B) a_{t,0} + \{\theta(B) - \theta_0(B)\} \theta_0^{-1}(B) a_{t,0} + a_t,\end{aligned}$$

т. е.

$$a_{t,0} = -\phi(B) \{\phi_0^{-1}(B) a_{t,0}\} + \theta(B) \{\theta_0^{-1}(B) a_{t,0}\} + a_t, \quad (7.2.16)$$

что является другой трактовкой этого алгоритма.

Приложение к процессу ПСС(0, 2, 2). Чтобы проиллюстрировать расчеты для приближения с помощью условной модели, рассмотрим оценки наименьших квадратов для $\hat{\theta}_1$, $\hat{\theta}_2$ ряда С, основанные на модели (0, 2, 2):

$$w_t = (1 - \theta_1 B - \theta_2 B^2) a_t$$

С

$$w_t = \nabla^2 z_t,$$

$$a_{t,0} = w_t + \theta_{1,0} a_{t-1,0} + \theta_{2,0} a_{t-2,0},$$

$$x_t = -a_{t,0} + \theta_{1,0} x_{t-1} + \theta_{2,0} x_{t-2}.$$

Расчет для начальных значений $\theta_{1,0} = 0,1$ и $\theta_{2,0} = 0,1$ приведен в табл. 7.11.

Таблица 7.11

Нелинейная оценка θ_1 и θ_2 для процесса ПСС (0, 2, 2)

t	z_t	∇z_t	$\nabla^2 z_t = w_t$	$a_{t,0}$	x_{t-1}	x_{t-2}
1	26,6	—	—	0	0	0
2	27,0	0,4	—	0	0	0
3	27,1	0,1	-0,3	-0,300	0	0
4	27,2	0,1	0,0	-0,030	0,300	0
5	27,3	0,1	0,0	-0,033	0,060	0,300
6	26,9	-0,4	-0,5	-0,533	0,069	0,060
7	26,4	-0,5	-0,1	-0,156	0,546	0,069
8	26,0	-0,4	0,1	0,039	0,218	0,546
9	25,8	-0,2	0,2	0,189	0,038	0,218

Первые поправки $\theta_{1,0}$ и $\theta_{2,0}$ были найдены «регрессией» $a_{t,0}$ на x_{t-1} и x_{t-2} , затем процедура повторялась до тех пор, пока не была достигнута сходимость. Результаты итераций показаны в табл. 7.12; использовались начальные значения $\theta_{1,0} = 0,1$ и $\theta_{2,0} = 0,1$.

7.2.5. Сводка моделей, подогнанных к рядам А — F

В табл. 7.13 дана сводка моделей, подогнанных к рядам А—F итеративным методом наименьших квадратов, описанным в разд. 7.2.1 и 7.2.2. Подгоняемые модели были идентифицированы в гл. 6, и их свойства были описаны в итоговой табл. 6.4. Из табл. 7.13 видно, что для рядов А, С и D были идентифицированы и подогнаны две возможные модели. Для рядов А и D одна из конкурирующих моделей включает стационарный оператор авторегрессии $(1 - \phi B)$ вместо нестационарного оператора $(1 - B)$ в другой. Рассмотрение табл. 7.13 показывает, что

Таблица 7.12
Сходимость итераций для θ_1 и θ_2

Итерация	θ_1	θ_2
0	0,1000	0,1000
1	0,1247	0,1055
2	0,1266	0,1126
3	0,1286	0,1141
4	0,1290	0,1149
5	0,1292	0,1151
6	0,1293	0,1152
7	0,1293	0,1153
8	0,1293	0,1153

Таблица 7.13

Сводка моделей, подогнанных к рядам А — F (значения под каждой оценкой — стандартные ошибки этих оценок)

Ряд	Число наблюдений	Подгоняемая модель	Остаточная дисперсия *)
A	197	$z_t - 0,92z_{t-1} = 1,45 + a_t - 0,58a_{t-1}$ ($\pm 0,04$) ($\pm 0,08$) $\nabla z_t = a_t - 0,70a_{t-1}$ ($\pm 0,05$)	0,097 0,101
B	369	$\nabla z_t = a_t + 0,09a_{t-1}$ ($\pm 0,05$)	52,2
C	226	$\nabla z_t - 0,82\nabla z_{t-1} = a_t$ ($\pm 0,04$) $\nabla^2 z_t = a_t - 0,13a_{t-1} - 0,12a_{t-2}$ ($\pm 0,07$) ($\pm 0,07$)	0,018 0,019
D	310	$z_t - 0,87z_{t-1} = 1,17 + a_t$ ($\pm 0,03$) $\nabla z_t = a_t - 0,06a_{t-1}$ ($\pm 0,06$)	0,090 0,096
E	100	$z_t = 14,35 + 1,42z_{t-1} - 0,73z_{t-2} + a_t$ ($\pm 0,07$) ($\pm 0,07$) $z_t = 11,31 + 1,57z_{t-1} - 1,02z_{t-2} + 0,21z_{t-3} + a_t$ ($\pm 0,10$) ($\pm 0,15$) ($\pm 0,10$)	228 218
F	70	$z_t = 58,87 - 0,34z_{t-1} + 0,19z_{t-2} + a_t$ ($\pm 0,12$) ($\pm 0,12$)	113

*) Получена из $S(\hat{\phi}, \hat{\theta})/n$.

в обоих случаях модель с авторегрессией дает несколько меньшую остаточную дисперсию, хотя, как уже отмечалось, модели очень близки. Хотя стационарная модель подгоняется несколько лучше, модель ПСС(0, 1, 1) предпочтительней в этих случаях, поскольку в отличие от стационарной модели она свободна от предположения, что ряд имеет фиксированное среднее значение. Это особенно важно для прогноза будущих членов ряда. Если уровень в действительности изменится, модель с $d > 0$ отреагирует на его изменение, в то время как модель с $d = 0$ будет тяготеть к устаревшему среднему уровню.

Пределы, указанные под коэффициентами в табл. 7.13, — это стандартные ошибки выборочных оценок, полученные из ковариационной матрицы $(\mathbf{X}'_{\hat{\beta}} \mathbf{X}_{\hat{\beta}})^{-1} \hat{\sigma}_a^2$, как описано в разд. 7.2.1. Отметим, что оценка $\hat{\phi}_3$ процесса АР(3), подогнанного к ряду Е солнечных пятен, в 2,1 раза больше своей стандартной ошибки; это свидетельствует о том, что улучшение, достигаемое подгонкой процесса авторегрессии третьего порядка по сравнению с процессом авторегрессии второго порядка, — на пределе значимости. Это согласуется с выводами Морана [54].

7.2.6*. Информационные матрицы при больших выборках и оценки ковариаций

Обозначим через $[\mathbf{U} : \mathbf{X}]$ матрицу размером $n \times (p + q)$ из значений u и x , определенных в (7.2.13) и (7.2.14), в случае, когда элементы β_0 — истинные значения параметров; пусть размер выборки n достаточно велик, так что можно пренебречь концевыми эффектами. Тогда информационная матрица для смешанной модели АРСС будет

$$\mathbf{I}(\phi, \theta) = E \left[\begin{array}{c|c} \mathbf{U}'\mathbf{U} & \mathbf{U}'\mathbf{X} \\ \hline \mathbf{X}'\mathbf{U} & \mathbf{X}'\mathbf{X} \end{array} \right] \sigma_a^{-2} = \quad (7.2.17)$$

$$= n\sigma_a^{-2} \left[\begin{array}{ccc|ccc} \gamma_{uu}(0) & \gamma_{uu}(1) & \dots \gamma_{uu}(p-1) & \gamma_{ux}(0) & \gamma_{ux}(-1) & \dots \gamma_{ux}(1-q) \\ \gamma_{uu}(1) & \gamma_{uu}(0) & \dots \gamma_{uu}(p-2) & \gamma_{ux}(1) & \gamma_{ux}(0) & \dots \gamma_{ux}(2-q) \\ & \vdots & & & & \\ \gamma_{uu}(p-1) & \gamma_{uu}(p-2) & \dots \gamma_{uu}(0) & \gamma_{ux}(p-1) & \gamma_{ux}(p-2) & \dots \gamma_{ux}(p-q) \\ \gamma_{ux}(0) & \gamma_{ux}(1) & \dots \gamma_{ux}(p-1) & \gamma_{xx}(0) & \gamma_{xx}(1) & \dots \gamma_{xx}(q-1) \\ \gamma_{ux}(-1) & \gamma_{ux}(0) & \dots \gamma_{ux}(p-2) & \gamma_{xx}(1) & \gamma_{xx}(0) & \dots \gamma_{xx}(q-2) \\ & \vdots & & & \vdots & \\ \gamma_{ux}(1-q) & \gamma_{ux}(2-q) & \dots \gamma_{ux}(p-q) & \gamma_{xx}(q-1) & \gamma_{xx}(q-2) & \dots \gamma_{xx}(0) \end{array} \right], \quad (7.2.18)$$

где $\gamma_{uu}(k)$ и $\gamma_{xx}(k)$ — автоковариации u и x и $\gamma_{ux}(k)$ — взаимные ковариации, определенные как

$$\gamma_{ux}(k) = \gamma_{xu}(-k) = E[u_t x_{t+k}] = E[x_t u_{t-k}].$$

Выборочная ковариационная матрица для оценок максимального правдоподобия в случае больших выборок может быть получена из соотношения

$$V(\hat{\phi}, \hat{\theta}) = I^{-1}(\phi, \theta).$$

Оценки $I(\phi, \theta)$ и, следовательно, $V(\hat{\phi}, \hat{\theta})$ можно получить, вычисляя u и x при $\beta_0 = \hat{\beta}$ и опуская знак математического ожидания в (7.2.17) или подстановкой стандартных выборочных оценок автокорреляций и взаимных корреляций в (7.2.18). Теоретические выражения для больших выборок можно получить, используя тот факт, что когда элементы β_0 равны истинным значениям параметров, из уравнений (7.2.13) и (7.2.14) следует, что производные u_t и x_t являются процессами авторегрессии:

$$\phi(B)u_t = a_t, \quad \theta(B)x_t = -a_t.$$

Отсюда автоковариации, присутствующие в (7.2.18), соответствуют чистым процессам авторегрессии, а взаимные ковариации обратны по знаку взаимным ковариациям между этими процессами, генерируемыми теми же a .

Проиллюстрируем этот результат на нескольких примерах.

Ковариационная матрица оценок параметров для процессов $AR(p)$ и $CC(q)$. Пусть $\Gamma_p(\phi)$ — автоковариационная матрица размера $p \times p$ для p последовательных наблюдений процесса $AR(p)$ с параметрами $\phi' = (\phi_1, \phi_2, \dots, \phi_p)$. Пользуясь (7.2.18), находим ковариационную матрицу оценок $\hat{\phi}$ размера $p \times p$

$$V(\hat{\phi}) = n^{-1} \sigma_a^2 \Gamma_p^{-1}(\phi). \quad (7.2.19)$$

Пусть $\Gamma_q(\theta)$ — автоковариационная матрица размера $q \times q$ для последовательных наблюдений процесса $CC(q)$ с параметрами $\theta' = (\theta_1, \theta_2, \dots, \theta_q)$. Тогда, пользуясь (7.2.18), получаем ковариационную матрицу оценок $\hat{\theta}$ размера $q \times q$

$$V(\hat{\theta}) = n^{-1} \sigma_a^2 \Gamma_q^{-1}(\theta). \quad (7.2.20)$$

Иногда полезно параметризовать процесс $ARCC$ через корни многочленов $\phi(B)$ и $\theta(B)$. В этом случае ковариационная матрица оценок параметров имеет особенно простую форму.

Ковариация корней многочленов, соответствующих процессу $ARCC$. Рассмотрим процесс $ARCC(p, q)$, параметризованный че-

рез корни многочленов $\phi(B)$, $\theta(B)$ (они предполагаются действительными), так что

$$\prod_{i=1}^p (1 - G_i B) \tilde{w}_t = \prod_{j=1}^q (1 - H_j B) a_t$$

или

$$a_t = \prod_{i=1}^p (1 - G_i B) \prod_{j=1}^q (1 - H_j B)^{-1} \tilde{w}_t.$$

Производные a тогда имеют вид

$$u_{i,t} = -\frac{\partial a_t}{\partial G_i} = (1 - G_i B)^{-1} a_{t-1},$$

$$x_{j,t} = -\frac{\partial a_t}{\partial H_j} = -(1 - H_j B)^{-1} a_{t-1}.$$

Отсюда, пользуясь (7.2.18), получаем для больших выборок информационную матрицу для корней $n^{-1} \mathbf{I}(\mathbf{G}, \mathbf{H}) =$

$$= \begin{bmatrix} (1-G_1^2)^{-1} & (1-G_1 G_2)^{-1} & \dots & (1-G_1 G_p)^{-1} & -(1-G_1 H_1)^{-1} & \dots & -(1-G_1 H_q)^{-1} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ (1-G_1 G_p)^{-1} & (1-G_2 G_p)^{-1} & \dots & (1-G_p^2)^{-1} & -(1-G_p H_1)^{-1} & \dots & -(1-G_p H_q)^{-1} \\ \hline -(1-G_1 H_1)^{-1} & -(1-G_2 H_1)^{-1} & \dots & -(1-G_p H_1)^{-1} & (1-H_1^2)^{-1} & \dots & (1-H_1 H_q)^{-1} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ -(1-G_1 H_q)^{-1} & -(1-G_2 H_q)^{-1} & \dots & -(1-G_p H_q)^{-1} & (1-H_1 H_q)^{-1} & \dots & (1-H_q^2)^{-1} \end{bmatrix}. \quad (7.2.21)$$

AP(2). В частности, для процесса авторегрессии второго порядка

$$(1 - G_1 B)(1 - G_2 B) \tilde{w}_t = a_t,$$

$$\begin{aligned} \mathbf{V}(\hat{G}_1, \hat{G}_2) &\approx n^{-1} \begin{bmatrix} (1 - G_1^2)^{-1} & (1 - G_1 G_2)^{-1} \\ (1 - G_1 G_2)^{-1} & (1 - G_2^2)^{-1} \end{bmatrix}^{-1} = \\ &= \frac{1}{n} \frac{1 - G_1 G_2}{(G_1 - G_2)^2} \begin{bmatrix} (1 - G_1^2)(1 - G_1 G_2) & -(1 - G_1^2)(1 - G_2^2) \\ -(1 - G_1^2)(1 - G_2^2) & (1 - G_2^2)(1 - G_1 G_2) \end{bmatrix}. \end{aligned} \quad (7.2.22)$$

В точности аналогичные формулы можно получить для процесса скользящего среднего второго порядка.

APCC(1, 1). Таким же образом для процесса APCC(1, 1)

$$(1 - \phi B) \tilde{w}_t = (1 - \theta B) a_t,$$

и, подставив $\phi = G_1$ и $\theta = H_1$ в (7.2.21), получаем

$$\begin{aligned} V(\hat{\phi}, \hat{\theta}) &\approx n^{-1} \begin{bmatrix} (1 - \phi^2)^{-1} & -(1 - \phi\theta)^{-1} \\ -(1 - \phi\theta)^{-1} & (1 - \theta^2)^{-1} \end{bmatrix}^{-1} = \\ &= \frac{1}{n} \frac{1 - \phi\theta}{(\phi - \theta)^2} \begin{bmatrix} (1 - \phi^2)(1 - \phi\theta) & (1 - \phi^2)(1 - \theta^2) \\ (1 - \phi^2)(1 - \theta^2) & (1 - \theta^2)(1 - \phi\theta) \end{bmatrix}. \end{aligned} \quad (7.2.23)$$

Результаты для этих двух процессов иллюстрируют их двойственность, рассмотренную в приложении П7.5.

7.3*. РЕЗУЛЬТАТЫ ОЦЕНИВАНИЯ ДЛЯ НЕКОТОРЫХ ЧАСТНЫХ МОДЕЛЕЙ

В приложениях П7.4 и П7.5 дан вывод некоторых результатов оценивания для ряда частных случаев. Ниже эти результаты, а также полученные ранее в этой главе собраны вместе для удобства просмотра.

7.3.1*. Процессы авторегрессии

Оценки параметров чистого процесса авторегрессии можно получить, решив некоторые *линейные уравнения*. В приложении П7.5 показано,

- 1) как можно получать точные оценки наименьших квадратов, решая систему линейных уравнений (см. также разд. 7.4.3);
- 2) как, слегка изменив коэффициенты этих уравнений, можно получить хорошее приближение к точным уравнениям максимального правдоподобия;
- 3) как, используя выборочные автокорреляции в качестве коэффициентов линейных уравнений Юла — Уокера, можно найти приближенные оценки наименьших квадратов и максимального правдоподобия.

Оценки, полученные в (1), конечно, идентичны тем, что находятся прямой минимизацией $S(\phi)$, описанной в общем виде в разд. 7.2. Оценки (3) — это хорошо известные приближения Юла и Уокера. Они полезны как начальные оценки на этапе идентификации, но в некоторых ситуациях могут заметно отличаться от оценок (1) и (2). Чтобы проиллюстрировать это, мы сравним оценки (3), полученные из уравнений Юла — Уокера, с оценками (1) наименьших квадратов, которые были сведены в табл. 7.13.

Оценки Юла — Уокера. Оценки Юла — Уокера (6.3.6) имеют вид

$$\hat{\phi} = R^{-1}r,$$

где

$$\mathbf{R} = \begin{bmatrix} 1 & r_1 & \dots & r_{p-1} \\ r_1 & 1 & \dots & r_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p-1} & r_{p-2} & \dots & 1 \end{bmatrix}, \quad \mathbf{r} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_p \end{bmatrix}. \quad (7.3.1)$$

В частности, оценки для процессов авторегрессии первого и второго порядков равны

$$\begin{aligned} \text{AP (1):} \quad \hat{\phi}_1 &= r_1, \\ \text{AP (2):} \quad \hat{\phi}_1 &= \frac{r_1(1-r_2)}{1-r_1^2}, \quad \hat{\phi}_2 = \frac{r_2-r_1^2}{1-r_1^2}. \end{aligned} \quad (7.3.2)$$

В приложении П7.5 показано, что приближенное значение $S(\hat{\phi})$ дает формула

$$S(\hat{\phi}) = \sum_{t=1}^n \tilde{\omega}_t^2 \{1 - \mathbf{r}'\hat{\phi}\}, \quad (7.3.3)$$

и отсюда

$$\hat{\sigma}_a^2 = S(\hat{\phi})/n = c_0(1 - \mathbf{r}'\hat{\phi}), \quad (7.3.4)$$

где c_0 — дисперсия ω . Выражение того же вида связывает σ_a^2 и теоретическую дисперсию ω [см. (3.2.8)] γ_0 , а именно

$$\sigma_a^2 = \gamma_0(1 - \rho'\phi),$$

где элементы ρ и ϕ — теоретические значения. Отсюда из (7.2.19) вытекает, что ковариационная матрица для оценок $\hat{\phi}$ имеет вид

$$\mathbf{V}(\hat{\phi}) = n^{-1} \sigma_a^2 \mathbf{\Gamma}^{-1} = n^{-1} (1 - \rho'\phi) \mathbf{P}^{-1}, \quad (7.3.5)$$

где $\mathbf{\Gamma}$ и \mathbf{P} — автоковариационная и автокорреляционная матрицы p последовательных значений процесса $\text{AP}(p)$, определенные (2.1.7).

В частности, для процессов авторегрессии первого и второго порядков находим

$$\text{AP (1):} \quad \mathbf{V}(\hat{\phi}) = n^{-1} (1 - \phi^2), \quad (7.3.6)$$

$$\text{AP (2):} \quad \mathbf{V}(\hat{\phi}_1, \hat{\phi}_2) = n^{-1} \begin{bmatrix} 1 - \phi_2^2 & -\phi_1(1 + \phi_2) \\ -\phi_1(1 + \phi_2) & 1 - \phi_2^2 \end{bmatrix}. \quad (7.3.7)$$

Оценки дисперсий и ковариаций получаем подстановкой в (7.3.5) оценок параметров. Отсюда

$$\hat{\mathbf{V}}(\hat{\phi}) = n^{-1} (1 - \mathbf{r}'\hat{\phi}) \mathbf{R}^{-1}. \quad (7.3.8)$$

Примеры. Автокорреляции для задержек 1 и 2 для рядов С, D, E и F, идентифицированных как возможные процессы авторегрессии (в случае ряда С это относится к первым разностям), показаны ниже.

Ряды	Пробная идентификация	Степень разности	Выборочные автокорреляции	n
С	(1, 1, 0)	1	$r_1 = 0,805$	226
D	(1, 0, 0)	0	$r_1 = 0,861$	310
E	(2, 0, 0)	0	$r_1 = 0,806, r_2 = 0,428$	100
F	(2, 0, 0)	0	$r_1 = -0,390, r_2 = 0,304$	70

Пользуясь (7.3.2), получаем оценки Юла — Уокера, приведенные в табл. 7.14, вместе со стандартными ошибками, вычисленными по (7.3.6) и (7.3.7).

Таблица 7.14

Оценки Юла — Уокера для рядов С—F

Ряд	Оценки и их стандартные ошибки	Коэффициенты корреляции между оценками
С	$\hat{\phi}_1 = 0,81 \quad \pm 0,04$	
D	$\hat{\phi}_1 = 0,86 \quad \pm 0,03$	
E	$\left\{ \begin{array}{l} \hat{\phi}_1 = 1,32 \\ \hat{\phi}_2 = -0,63 \end{array} \right. \quad \begin{array}{l} \pm 0,08 \\ \pm 0,08 \end{array}$	$-0,81$
F	$\left\{ \begin{array}{l} \hat{\phi}_1 = -0,32 \\ \hat{\phi}_2 = 0,18 \end{array} \right. \quad \begin{array}{l} \pm 0,12 \\ \pm 0,12 \end{array}$	$0,39$

Из (7.3.7) получаем выражение для коэффициента корреляции между оценками двух параметров процесса авторегрессии второго порядка:

$$\rho(\hat{\phi}_1, \hat{\phi}_2) = \frac{\text{cov}(\hat{\phi}_1, \hat{\phi}_2)}{\sqrt{V(\hat{\phi}_1) V(\hat{\phi}_2)}} = \frac{-\phi_1}{1 - \phi_2} = -\rho_1.$$

Заметим, что для ряда E существует большая отрицательная корреляция между оценками. Это означает, что доверительная область для $\hat{\phi}_1$ и $\hat{\phi}_2$ будет вытягиваться вдоль диагонали, идущей из верхнего левого угла на плоскости (ϕ_1, ϕ_2) . Отсюда следует, что эти оценки нестабильны; в этом заключается причина относительно больших расхождений между оценками наименьших квадратов из табл. 7.13 и оценками Юла — Уокера из табл. 7.14 для этого конкретного ряда.

7.3.2.* Процессы скользящего среднего

Оценки максимального правдоподобия $\hat{\theta}$ для процессов скользящего среднего можно в простых случаях получить графически, как это было показано в разд. 7.1.6, а в более общей ситуации — итеративным расчетом, описанным в разд. 7.2.1.

Из (7.2.20) следует, что для средних и больших выборок ковариационная матрица оценок параметров процесса скользящего среднего q -го порядка имеет ту же форму, что и соответствующая матрица для процесса авторегрессии того же порядка.

Отсюда для процессов скользящего среднего первого и второго порядков, согласно (7.3.6) и (7.3.7), получаем

$$\text{СС (1): } V(\hat{\theta}) = n^{-1} (1 - \theta^2), \quad (7.3.9)$$

$$\text{СС (2): } V(\hat{\theta}_1, \hat{\theta}_2) = n^{-1} \begin{bmatrix} 1 - \theta_2^2 & -\theta_1(1 + \theta_2) \\ -\theta_1(1 + \theta_2) & 1 - \theta_2^2 \end{bmatrix}. \quad (7.3.10)$$

7.3.3.* Смешанные процессы

Оценки максимального правдоподобия $(\hat{\phi}, \hat{\theta})$ для смешанных процессов, как и для процессов скользящего среднего, в простых случаях можно получить графически, а в более сложных — итеративным расчетом. Для выборок средних и больших размеров ковариационную матрицу можно найти, оценив и обратив информационную матрицу (7.2.18). В важном частном случае процесса АРСС(1, 1)

$$(1 - \phi B) \tilde{w}_t = (1 - \theta B) a_t$$

получаем, как и в (7.2.23),

$$V(\hat{\phi}, \hat{\theta}) = n^{-1} \frac{1 - \phi\theta}{(\phi - \theta)^2} \begin{bmatrix} (1 - \phi^2)(1 - \phi\theta) & (1 - \phi^2)(1 - \theta^2) \\ (1 - \phi^2)(1 - \theta^2) & (1 - \theta^2)(1 - \phi\theta) \end{bmatrix}. \quad (7.3.11)$$

Заметим, что при $\phi = \theta$ дисперсии $\hat{\phi}$ и $\hat{\theta}$ бесконечны. Этого следовало ожидать, так как в этих случаях множители $(1 - \phi B) = (1 - \theta B)$ в обеих частях уравнения модели сокращаются и оно приобретает вид

$$\tilde{w}_t = a_t.$$

Это — частный случай избыточности параметров, которую мы более подробно обсудим в разд. 7.3.5.

7.3.4.* Разделение линейных и нелинейных компонент при оценивании

Иногда интересно провести анализ, в котором оценки параметров смешанной модели разделены на линейные и нелинейные части. Рассмотрим общую смешанную модель, которую мы

представили как

$$a_t = \phi(B) \theta^{-1}(B) \tilde{w}_t$$

или

$$a_t = \phi(B)(e_t | \theta), \quad (7.3.12)$$

где

$$(e_t | \theta) = \theta^{-1}(B) \tilde{w}_t,$$

т. е.

$$\tilde{w}_t = \theta(B)(e_t | \theta). \quad (7.3.13)$$

Для любого данного набора θ величину e можно вычислить рекуррентно при помощи формулы (7.3.13), которую можно записать в виде

$$e_t = \tilde{w}_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \dots + \theta_q e_{t-q}.$$

Расчет можно начать, положив неизвестные e равными нулю. Когда e вычислены, легко получить условные оценки $\hat{\phi}_\theta$. Это — выборочные параметры авторегрессии в линейной модели (7.3.12), которую можно представить в виде

$$a_t = e_t - \phi_1 e_{t-1} - \phi_2 e_{t-2} - \dots - \phi_p e_{t-p}. \quad (7.3.14)$$

Как уже объяснялось в разд. 7.3.1, оценки наименьших квадратов для параметров авторегрессии можно найти прямым решением простой системы линейных уравнений и приближенно при помощи уравнений Юла — Уокера. В простых случаях мы можем изучить поведение $S(\hat{\phi}_\theta, \theta)$ и найти абсолютный минимум суммы квадратов вычислением этой суммы на сетке значений θ и построением изолиний.

Пример, относящийся к ряду C. В соответствии с одним из вариантов пробной идентификации ряда C, приведенным в табл. 6.4, этот ряд, возможно, генерируется моделью (1, 1, 0)

$$(1 - \phi B) w_t = a_t$$

с $w_t = \nabla z_t$ и $E[w_t] = 0$. Мы хотели изучить ситуацию оценивания по этим данным в случае более сложной модели

$$(1 - \phi B) w_t = (1 - \theta_1 B - \theta_2 B^2) a_t.$$

Следуя проведенным выше рассуждениям, можно трактовать этот процесс как комбинацию нелинейной модели

$$e_t = w_t + \theta_1 e_{t-1} + \theta_2 e_{t-2}$$

и линейной модели

$$a_t = e_t - \phi e_{t-1}.$$

Последовательность значений e вычислялась рекуррентным образом для каждого выбранного значения нелинейного параметра $\theta = (\theta_1, \theta_2)$ внутри области обратимости. Пользуясь прибли-

жением Юла — Уокера, можно найти оценку $\hat{\phi}_0 = r_1(\varepsilon)$, а также

$$S(\hat{\phi}_0, \theta) \approx \sum_{t=1}^n \varepsilon_t^2 \{1 - r_1^2(\varepsilon)\}.$$

Эта сумма квадратов была вычислена для сетки значений θ_1 и θ_2 ; ее рельеф показан на рис. 7.8. Видно, что минимум расположен близко к точке $\theta_1 = \theta_2 = 0$, в которой $r_1(\varepsilon) = 0,805$. Таким

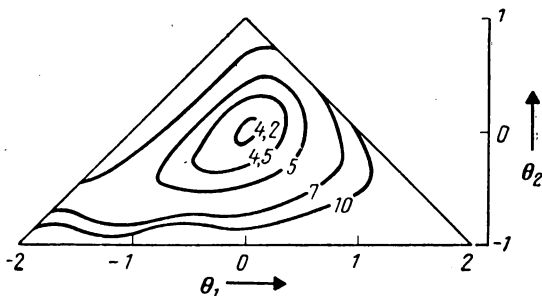


Рис. 7.8. Изолинии суммы $S(\hat{\phi}_0, \theta)$, построенной над областью допустимых значений параметров θ .

образом, подтверждено, что внутри всего класса моделей порядка $(1, 1, 2)$ и ниже простейшая модель $(1, 1, 0)$

$$(1 - 0,8B) \nabla z_t = a_t$$

обеспечивает хорошее описание наблюдаемого ряда.

7.3.5. Избыточность параметров

Модель

$$\phi(B) \tilde{w}_t = \theta(B) a_t$$

идентична модели

$$(1 - \alpha B) \phi(B) \tilde{w}_t = (1 - \alpha B) \theta(B) a_t,$$

в которой операторы авторегрессии и скользящего среднего умножены на один и тот же множитель $(1 - \alpha B)$. Значительные трудности в процедуре оценивания могут быть вызваны тем фактом, что подгоняется модель с избыточным (лишним) множителем. Поэтому требуется позаботиться о том, чтобы в модели не появлялись избыточные или почти избыточные множители. Например, общий множитель в модели АРСС(2, 1)

$$(1 - 1,3B + 0,4B^2) \tilde{w}_t = (1 - 0,5B) a_t$$

можно увидеть только после разложения левой части на множители

$$(1 - 0,5B)(1 - 0,8B) \tilde{w}_t = (1 - 0,5B) a_t,$$

что приводит к

$$(1 - 0,8B) \tilde{w}_t = a_t.$$

На практике трудности вызывает не только наличие одинаковых множителей в обеих частях уравнения, но и ситуация, когда имеются почти одинаковые множители. Например, пусть истинная модель была

$$(1 - 0,4B)(1 - 0,8B) \tilde{w}_t = (1 - 0,5B) a_t. \quad (7.3.15)$$

Если делается попытка подогнать эту модель, можно ожидать крайнюю нестабильность в оценках параметров из-за близости множителей $(1 - 0,4B)$ и $(1 - 0,5B)$ по разные стороны (7.3.15). В этой ситуации можно найти комбинации значений параметров, дающих близкие $[a]$ и, следовательно, близкие значения правдоподобия; изменение значений параметра слева почти компенсируется соответствующим изменением параметра справа. Поверхность суммы квадратов в трехмерном пространстве параметров будет сходна с косоцилиндрической, и вместо четко определенного минимума будет найдена линия «почти наименьших квадратов».

С несколько иной точки зрения мы можем представить модель (7.3.15) в виде бесконечного оператора авторегрессии. Применяя необходимые разложения, находим

$$(1 - 0,700B - 0,030B^2 - 0,015B^3 - 0,008B^4 - \dots) \tilde{w}_t = a_t.$$

Итак, модель весьма аккуратно описывается как

$$(1 - 0,7B) \tilde{w}_t = a_t. \quad (7.3.16)$$

Нестабильность оценок, полученных при попытке подогнать модель АРСС(2, 1), возникает потому, что мы пытались подогнать три параметра в ситуации, где почти достаточно одного.

Предварительная идентификация как средство избежать избыточности параметров. Главная цель применения идентификационной процедуры перед подгонкой модели — избежать трудностей, возникающих из-за избыточности параметров, или, в более позитивном смысле, достичь *экономичности* параметризации. Так, в только что рассмотренном примере для временного ряда в несколько сотен наблюдений выборочная автокорреляционная функция по данным, генерируемым моделью (7.3.15), была бы неотличима от генерируемой простым процессом авторегрессии (7.3.16).

Это привело бы нас к подгонке процесса АР(1), в нормальной ситуации полностью адекватного наблюдениям. Только для временного ряда в несколько тысяч наблюдений может потребоваться более сложная модель, и имеющейся информации окажется достаточно для получения надежных оценок дополнительных параметров.

Избыточность процесса $APCC(1, 1)$. Простейший процесс, в котором возникает возможность прямого сокращения множителя, — это процесс $APCC(1, 1)$.

$$(1 - \phi B) \tilde{w}_t = (1 - \theta B) a_t.$$

В частности, если $\phi = \theta$, то, каковы бы они ни были,

$$\tilde{w}_t = a_t,$$

и это означает, что \tilde{w} генерируется белым шумом. Данные не содержат информации об общих параметрах, и в соответствии с (7.3.11) $\hat{\phi}$ и $\hat{\theta}$ имеют бесконечную дисперсию. Далее, при любых значениях ϕ и θ $S(\phi, \theta)$ — константа на прямой $\phi = \theta$, как

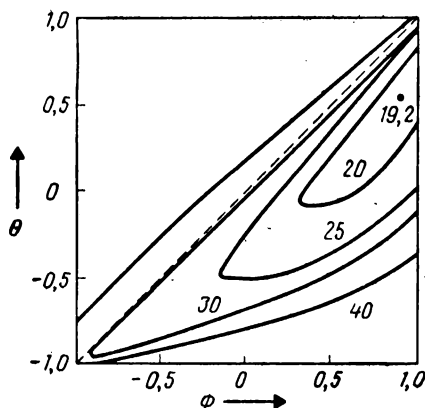


Рис. 7.9. Изолинии суммы квадратов для ряда А.

это видно на карте рельефа суммы квадратов для ряда А, показанной на рис. 7.9. Для этих данных значения оценок наименьших квадратов $\hat{\phi} = 0,92$, $\hat{\theta} = 0,58$ соответствуют точке, которая не слишком близка к прямой $\phi = \theta$, и в оценке этих параметров трудностей не возникает.

На практике, если принять рекомендуемую нами методику идентификации, трудностей, связанных с избыточностью, можно избежать. Процесс $APCC(1, 1)$, у которого ϕ почти равно θ , в нормальной ситуации будет идентифицирован как белый шум или, если различие заметно, как процесс $AP(1)$ или $CC(1)$ с одним небольшим коэффициентом. В итоге

1) следует избегать смешанных процессов, содержащих почти одинаковые множители, и помнить о трудностях, которые могут возникнуть, если такие множители имеются;

2) такие процессы будут исключаться автоматически, если разумно пользоваться процедурами идентификации и оценивания.

7.4. ОЦЕНИВАНИЕ ПРИ ПОМОЩИ ТЕОРЕМЫ БАЙЕСА

7.4.1.* Теорема Байеса

В этом разделе мы опять будем пользоваться обозначением ξ для общего вектора параметров. Теорема Байеса гласит, что если $p(\xi)$ — априорное распределение вероятности для ξ , то апостериорное распределение вероятности $p(\xi|z)$ для ξ после получения данных z можно следующим образом выразить через априорную вероятность и функцию правдоподобия:

$$p(\xi|z) = \frac{p(\xi) L(\xi|z)}{\int p(\xi) L(\xi|z) d\xi}. \quad (7.4.1)$$

Знаменатель просто обеспечивает нормировку: интеграл от $p(\xi|z)$ равен единице. Существенную роль в выражении играет числитель — видно, что апостериорная вероятность пропорциональна априорной вероятности, умноженной на функцию правдоподобия. Сэведж [64] показал, что априорная и апостериорная вероятности могут интерпретироваться как субъективные вероятности. В частности, во многих случаях до того, как появились данные, мы очень мало знаем о ξ и готовы согласиться, что в интересующей нас области а priori любые значения ξ равновероятны. В этом случае $p(\xi)$ можно считать в этой области равномерным, и, следовательно, $p(\xi|z)$ будет пропорционально функции правдоподобия.

Следует заметить, что при этом не требуется, чтобы распределение ξ было равномерным во всем диапазоне, который для некоторых параметров может быть бесконечным. Говоря о *локальной* равномерности, мы подразумеваем, что распределение приблизительно постоянно в области, где функция правдоподобия существенна, и не принимает очень больших значений вне этой области.

Так, если ξ — вес стула, то мы определенно можем сказать а priori, что он весит больше унции (≈ 30 г) и меньше тонны. Также вероятно, что, когда мы сделаем наблюдение z , взвесив стул на весах со стандартным отклонением σ , мы можем честно признаться, что нас равно удовлетворили бы любые априорные значения ξ в пределах $z \pm 3\sigma$. Исключение составил бы случай, когда вес тяжелого по внешнему виду стула оказался бы, скажем, равным 10 унциям. В этом случае априорное распределение и функция правдоподобия оказались бы несопоставимыми, и мы, конечно, не стали бы применять теорему Байеса, а проверили бы весы, и если бы они оказались правильными, исследовали стул более внимательно.

В этой идее есть, конечно, некий произвол. Пусть мы предположили, что априорное распределение ξ локально равномерно. Отсюда следует, что распределение любой линейной функции ξ также локально равномерно. Однако априорное распределение некоего нелинейного преобразования $\alpha = \alpha(\xi)$ (например, $\alpha = \lg \xi$) может и *не быть* локально равномерным. Этот произвол обычно не играет существенной роли, если мы можем получить довольно точные оценки ξ . Мы будем тогда рассматривать ξ только в весьма узком диапазоне, в котором преобразование ξ (скажем, $\lg \xi$) часто очень близко к линейному.

Джеффрис [65] указывал, что лучше выбрать преобразование $\alpha(\xi)$ так, чтобы мера информации Фишера $I_\alpha = -E[\partial^2 l / \partial \alpha^2]$ была бы независимой от α и, следовательно, от ξ . Это эквивалентно выбору $\alpha(\xi)$ таким образом, чтобы предельная дисперсия ее оценки максимального правдоподобия была независима от ξ , и достигается выбором априорного распределения ξ , пропорциональным $\sqrt{I_\xi}$.

Джеффрис оправдывал этот выбор, исходя из его инвариантности по отношению к используемой параметризации. Тот же результат может быть получен [66] из следующих рассуждений. Если для больших выборок функция правдоподобия для $\alpha(\xi)$ приближается к нормальной кривой, то среднее значение и дисперсия этой кривой заключает в себе всю информацию, которую можно рассчитывать извлечь из данных. Положим теперь, что можно найти преобразование $\alpha(\xi)$, для которого аппроксимирующая нормальная кривая имеет почти постоянную дисперсию, какими бы ни были истинные значения параметров. Тогда при этой параметризации *единственная* информация, на которую можно рассчитывать, выражается в параметре, определяющем *положение* функции правдоподобия. Не знать ничего существенного а priori об этой ожидаемой информации означает, что любые возможные *положения* α мы а priori считаем равновероятными. Иначе говоря, α следует считать локально равномерным.

Обобщение правила Джеффриса на случай нескольких параметров заключается в том, что совместное априорное распределение параметров ξ считается пропорциональным

$$|I_\xi|^{1/2} = \left| -E \left\{ \frac{\partial^2 l}{\partial \xi_i \partial \xi_j} \right\} \right|^{1/2}. \quad (7.4.2)$$

В [67] подчеркивалось, что саму функцию правдоподобия лучше рассматривать и графически представлять в такой метрике α , для которой I_α независимо от α . Если это сделано, то функция правдоподобия и апостериорная функция плотности распределения с равномерным априорным распределением пропорциональны,

7.4.2.* Байесовское оценивание параметров

Рассмотрим теперь оценивание параметров в модели АРПСС с байесовской точки зрения.

В приложении П7.4 показано, что точное выражение для функции правдоподобия временного ряда \mathbf{z} длиной $N = n + d$ для процесса АРПСС (p, d, q) имеет вид

$$L(\boldsymbol{\phi}, \theta | \mathbf{z}) = \sigma_a^{-n} f(\boldsymbol{\phi}, \theta) \exp \left\{ -\frac{S(\boldsymbol{\phi}, \theta)}{2\sigma_a^2} \right\}, \quad (7.4.3)$$

где

$$S(\boldsymbol{\phi}, \theta) = \sum_{t=-\infty}^n [a_t | \mathbf{z}, \boldsymbol{\phi}, \theta]^2 = \sum_{t=-\infty}^n [a_t | \mathbf{w}, \boldsymbol{\phi}, \theta]^2. \quad (7.4.4)$$

Если априорная информация о σ_a , $\boldsymbol{\phi}$ и θ отсутствует, то, поскольку информация о σ_a не дает нам информации о $\boldsymbol{\phi}$ и θ , рационально, следуя Джеффрису, принять априорное распределение $\boldsymbol{\phi}$, θ и σ_a в виде

$$p(\boldsymbol{\phi}, \theta, \sigma_a) \sim |I(\boldsymbol{\phi}, \theta)|^{1/2} \sigma_a^{-1}.$$

Отсюда следует, что апостериорное распределение имеет вид

$$p(\boldsymbol{\phi}, \theta, \sigma_a | \mathbf{z}) \sim \sigma_a^{-(n+1)} |I(\boldsymbol{\phi}, \theta)|^{1/2} f(\boldsymbol{\phi}, \theta) \exp \left\{ -\frac{S(\boldsymbol{\phi}, \theta)}{2\sigma_a^2} \right\}. \quad (7.4.5)$$

Если теперь проинтегрировать (7.4.5) от нуля до бесконечности по σ_a , то мы получим точное совместное апостериорное распределение параметров $\boldsymbol{\phi}$ и θ :

$$p(\boldsymbol{\phi}, \theta | \mathbf{z}) \sim |I(\boldsymbol{\phi}, \theta)|^{1/2} f(\boldsymbol{\phi}, \theta) \{S(\boldsymbol{\phi}, \theta)\}^{-n/2}. \quad (7.4.6)$$

7.4.3.* Процессы авторегрессии

Если z_t есть процесс $(p, d, 0)$, то $w_t = \nabla^d z_t$ будет чистым процессом авторегрессии порядка p . В приложении П7.5 показано, что для такого процесса множители $|I(\boldsymbol{\phi})|^{1/2}$ и $f(\boldsymbol{\phi})$, которые во всех случаях являются существенно менее важными, чем множитель $S(\boldsymbol{\phi})$, практически взаимно сокращаются. Это приводит к очень простому результату, что при данных предположениях параметры $\boldsymbol{\phi}$ процесса АР(p) для \mathbf{w} имеют апостериорное распределение

$$p(\boldsymbol{\phi} | \mathbf{z}) \sim \{S(\boldsymbol{\phi})\}^{-n/2}. \quad (7.4.7)$$

В этом случае рельеф суммы квадратов, приближенно совпадающий с рельефом функции правдоподобия при отсутствии априорной информации, совпадает также с рельефом апостериорной вероятности,

Совместное распределение параметров авторегрессии. В приложении П7.5 показано, что для чистого процесса АР оценки наименьших квадратов для ϕ , минимизирующие $S(\phi) = \phi_u' D \phi_u$, имеют вид

$$\hat{\phi} = D_p^{-1} d, \quad (7.4.8)$$

где

$$\phi_u = \begin{bmatrix} 1 \\ \phi_1 \\ \vdots \\ \phi_p \end{bmatrix}, \quad d = \begin{bmatrix} D_{12} \\ D_{13} \\ \vdots \\ D_{1, p+1} \end{bmatrix},$$

$$D_p = \begin{bmatrix} D_{22} & D_{23} & \dots & D_{2, p+1} \\ D_{23} & D_{33} & \dots & D_{3, p+1} \\ \vdots & \vdots & \dots & \vdots \\ D_{2, p+1} & D_{3, p+1} & \dots & D_{p+1, p+1} \end{bmatrix}, \quad D = \left[\begin{array}{c|c} D_{11} & -d' \\ \hline -d & D_p \end{array} \right] \quad (7.4.9)$$

и

$$D_{ij} = D_{ji} = \tilde{w}_i \tilde{w}_j + \tilde{w}_{i+1} \tilde{w}_{j+1} + \dots + \tilde{w}_{n+1-i} \tilde{w}_{n+1-j}. \quad (7.4.10)$$

Отсюда следует, что

$$S(\phi) = v s_a^2 + (\phi - \hat{\phi})' D_p (\phi - \hat{\phi}), \quad (7.4.11)$$

где

$$s_a^2 = S(\hat{\phi})/v, \quad v = n - p, \quad (7.4.12)$$

и

$$S(\hat{\phi}) = \hat{\phi}_u' D \hat{\phi}_u = D_{11} - \hat{\phi}' D_p \hat{\phi} = D_{11} - d' D_p^{-1} d. \quad (7.4.13)$$

Тогда мы можем записать

$$p(\phi | z) \sim \left\{ 1 + \frac{(\phi - \hat{\phi})' D_p (\phi - \hat{\phi})}{v s_a^2} \right\}^{-n/2}. \quad (7.4.14)$$

Это можно записать и иначе:

$$p(\phi | z) \sim \left\{ 1 + \frac{\frac{1}{2} \sum_i \sum_j S_{ij} (\phi_i - \hat{\phi}_i) (\phi_j - \hat{\phi}_j)}{v s_a^2} \right\}^{-n/2}, \quad (7.4.15)$$

где

$$S_{ij} = \partial^2 S(\phi) / \partial \phi_i \partial \phi_j = 2D_{i+1, j+1}.$$

Отсюда следует, что параметры процесса авторегрессии имеют многомерное t -распределение (П7.1.13) с $v = n - p$ степенями свободы.

На практике для частного случая $p = 1$ распределение $(\phi - \hat{\phi})/s_{\hat{\phi}}$ точно совпадает с t -распределением Стьюдента с $n - 1$ степенями свободы, причем из общих результатов, полученных выше, следует, что

$$\hat{\phi} = D_{12}/D_{22}, \quad s_{\hat{\phi}} = \left\{ \frac{1}{(n-1)} \frac{D_{11}}{D_{22}} \left[1 - \frac{D_{12}^2}{D_{11}D_{22}} \right] \right\}^{1/2}. \quad (7.4.16)$$

Величина $s_{\hat{\phi}}$ при больших выборках стремится к $[(1 - \phi^2)/n]^{1/2}$ и в рамках выборочной теории идентична со стандартной ошибкой для $\hat{\phi}$ при больших выборках. Однако, используя это и подобные выражения в рамках байесовского подхода, следует помнить, что случайными переменными являются именно параметры (в этом случае ϕ). Такие величины, как $\hat{\phi}$ и $s_{\hat{\phi}}$, являющиеся функциями уже полученных данных, рассматриваются как фиксированные.

Приближение нормальным законом. Для выборок размером $n > 50$, которыми мы обычно интересуемся, t -распределение может с достаточной точностью аппроксимироваться нормальным распределением. Это означает, что распределение ϕ очень близко к p -мерному совместному нормальному распределению $N\{\hat{\phi}, D_p^{-1}s_a^2\}$ со средним значением $\hat{\phi}$ и матрицей ковариаций $D_p^{-1}s_a^2$.

Байесовские области наивысшей плотности вероятности. Суммируя то, что можно узнать о вероятности различных значений ϕ по апостериорному распределению, полезно указать область *наивысшей плотности вероятности*, кратко называемую областью НПВ [104]. Байесовская $(1 - \varepsilon)$ -я область НПВ имеет следующие свойства:

1) любое значение параметра внутри этой области имеет более высокое значение вероятности, чем вне этой области;

2) апостериорная вероятность этой области равна $1 - \varepsilon$.

Так как ϕ имеет многомерное t -распределение, из (П7.1.4) следует, что

$$P\{(\phi - \hat{\phi})' D_p (\phi - \hat{\phi}) < p s_a^2 F_\varepsilon(p, v)\} = 1 - \varepsilon \quad (7.4.17)$$

определяет точную $(1 - \varepsilon)$ -ю область НПВ для ϕ . Для $v = n - p > 100$

$$p F_\varepsilon(p, v) \approx \chi_\varepsilon^2(p).$$

Кроме того,

$$(\phi - \hat{\phi})' D_p (\phi - \hat{\phi}) = \frac{1}{2} \sum_{i,l} S_{il} (\phi_i - \hat{\phi}_i) (\phi_l - \hat{\phi}_l).$$

Отсюда приближенно область НПВ, определенная (7.4.17), имеет вид

$$\sum_{i,j} S_{ij}(\phi_i - \hat{\phi}_i)(\phi_j - \hat{\phi}_j) < 2s_a^2 \chi_e^2(p) \quad (7.4.18)$$

и, если положить $\hat{\sigma}_a^2 = s_a^2$, идентична с доверительной областью, определенной (7.1.26).

Хотя эти приближенные области и идентичны, нужно помнить, что их интерпретация различна. С точки зрения выборочной теории говорят, что если доверительная область вычислена согласно (7.1.26), то для каждой серии повторных выборок $(1 - \epsilon)$ -я часть этих областей будет включать точку с истинным значением параметра. С байесовской точки зрения рассматривается результат единственной фактически наблюдавшейся выборки \mathbf{z} . В предположении, что априорное распределение вероятности локально равномерно, область НПВ включает $(1 - \epsilon)$ -ю часть результирующего вероятностного распределения $\boldsymbol{\phi}$ при данном \mathbf{z} , обладающую наивысшей плотностью. Другими словами, вероятность того, что значение $\boldsymbol{\phi}$, вызывавшее данные \mathbf{z} , лежит внутри НПВ, равна $1 - \epsilon$.

Пользуясь (7.4.11), (7.4.12) и (7.4.18) для больших выборок, находим, что приближенная байесовская область НПВ ограничена изолинией, для которой

$$S(\boldsymbol{\phi}) = S(\hat{\boldsymbol{\phi}}) \left\{ 1 + \frac{\chi_e^2(p)}{n} \right\}, \quad (7.4.19)$$

т. е. точно соответствует доверительной области, определенной (7.1.28).

7.4.4.* Процессы скользящего среднего

Если \mathbf{z} есть процесс проинтегрированного скользящего среднего порядка $(0, d, q)$, то $\omega = \nabla^d \mathbf{z}$ является чистым процессом скользящего среднего порядка q . В приложении П7.5 показано, что в этом случае множители $|\mathbf{I}(\boldsymbol{\theta})|^{1/2}$ и $f(\boldsymbol{\theta})$ в (7.4.6), которые во всех случаях существенно менее важны, чем множитель с $S(\boldsymbol{\theta})$, для больших выборок взаимно сокращаются. Тогда, согласно (7.4.7), параметры $\boldsymbol{\theta}$ процесса $CC(q)$ для ω имеют апостериорное распределение

$$p(\boldsymbol{\theta} | \mathbf{z}) \sim \{S(\boldsymbol{\theta})\}^{-n/2}, \quad (7.4.20)$$

и в этом случае рельеф суммы квадратов для не слишком малых выборок практически *совпадает* с рельефом апостериорной плотности вероятности. Так как $[a_i]$ не является линейной функцией $\boldsymbol{\theta}$, $S(\boldsymbol{\theta})$ не является точно квадратичной функцией $\boldsymbol{\theta}$, но для больших выборок внутри изучаемой области обычно это

приближенно верно. В таком случае имеем приближенно

$$S(\theta) = v s_a^2 + \frac{1}{2} \sum_{i,j} S_{ij}(\theta_i - \hat{\theta}_i)(\theta_j - \hat{\theta}_j),$$

где $v s_a^2 = S(\hat{\theta})$ и $v = n - q$. Отсюда, подставляя это выраже-

Таблица 7.15

Расчет приближенной апостериорной плотности $p(\lambda | z)$ для ряда В

λ	$10^{57} \times [S(\lambda)] - 184,5$	$p(\lambda z)$
1,300	4	0,001
1,275	33	0,006
1,250	212	0,036
1,225	1 007	0,171
1,200	3 597	0,609
1,175	9 956	1,685
1,150	21 159	3,582
1,125	34 762	5,884
1,100	44 473	7,528
1,075	44 988	7,615
1,050	35 835	6,066
1,025	22 563	3,819
1,000	11 277	1,908
0,975	4 540	0,769
0,950	1 457	0,247
0,925	372	0,063
0,900	76	0,012
0,900	76	0,012
0,875	12	0,002
0,850	2	0,000
Всего 236 325		40,000

ние для $S(\theta)$ в (7.4.20) и используя экспоненциальное приближение, получаем

1) для больших выборок θ приближенно распределено по многомерному нормальному закону $N\{\hat{\theta}, 2(S_{ij})^{-1} s_a^2\}$;

2) приближенная область НПВ определена формулой (7.4.18) или (7.4.19), если в них заменить p на q и ϕ на θ .

Пример: апостериорное распределение $\lambda = 1 - \theta$ для процесса ПСС(0, 1, 1). В качестве примера в табл. 7.15 приведены вычисления приближенного апостериорного распределения плотности вероятности $p(\lambda | z)$ по данным ряда В. Во втором столбце таблицы приведены вычисленные ординаты $\{S(\lambda)\}^{-n/2} \times 10^{57}$ для значений λ с шагом $h = 0,025$. Их сумма Σ равна 236 325. Поделив ординаты на $h\Sigma$, получаем функцию апостериорной плотности,

площадь под графиком которой с достаточной точностью равна 1 *).

*) Приближенное численное интегрирование для получения «нормированной» кривой с единичной площадью, которое при желании можно выполнить и более точно, позволяет получить истинные плотности вероятности. Почти для всех практических приложений этого делать не нужно — достаточно вычислить и построить в удобном масштабе $\{S(\lambda)\}^{-n/2}$, пропорциональную истинной плотности.

Распределение показано на рис. 7.10. Видно, что оно приближенно нормальное с модой $\hat{\lambda} = 1,09$ и стандартным отклонением около 0,05. 95%-ный байесовский интервал покрывает тот же диапазон $0,98 < \lambda < 1,19$, что и доверительный интервал.

7.4.5.* Смешанные процессы

Если z есть процесс АРПСС (p, d, q) , то $\omega = \nabla^d z$ является процессом АРСС (p, q) $\phi(B)\tilde{\omega}_t = \theta(B)a_t$. В приложении П7.5 показано, что для такого процесса множители $|I(\phi, \theta)|^{1/2}$ и $f(\phi, \theta)$ в (7.4.5) не сокращаются точно. Вместо этого мы показываем, что

$$\begin{aligned} |I(\phi, \theta)|^{1/2} f(\phi, \theta) &= \\ &= J(\phi^* | \phi, \theta). \end{aligned} \quad (7.4.21)$$

В (7.4.21) ϕ^* — это $p+q$ параметров, полученных перемножением операторов авторегрессии и скользящего среднего

$$\begin{aligned} (1 - \phi_1^* B - \phi_2^* B^2 - \dots - \phi_{p+q}^* B^{p+q}) &= \\ &= (1 - \phi_1 B - \dots - \phi_p B^p)(1 - \theta_1 B - \dots - \theta_q B^q), \end{aligned}$$

и J — якобиан преобразования ϕ^* в (ϕ, θ) , т. е.

$$p(\phi, \theta | z) \sim J(\phi^* | \phi, \theta) \{S(\phi, \theta)\}^{-n/2}. \quad (7.4.22)$$

В частности, для процесса АРСС (1,1) $\phi_1^* = \phi + \theta$, $\phi_2^* = -\phi\theta$, $J = |\phi - \theta|$ и

$$p(\phi, \theta | z) \sim |\phi - \theta| \{S(\phi, \theta)\}^{-n/2}. \quad (7.4.23)$$

В этом случае видно, что якобиан будет играть определяющую роль в окрестности прямой $\phi = \theta$ и будет давать нулевую плотность на этой прямой. Это разумно, так как сумма квадратов $S(\phi, \theta)$ будет принимать одинаковое конечное значение $\sum_{t=1}^n \tilde{\omega}_t^2$ для любого $\phi = \theta$ и соответствовать исследованию возможности, что $\tilde{\omega}_t$ — белый шум. Однако в нашем выводе мы не накладывали ограничений на диапазон параметров.

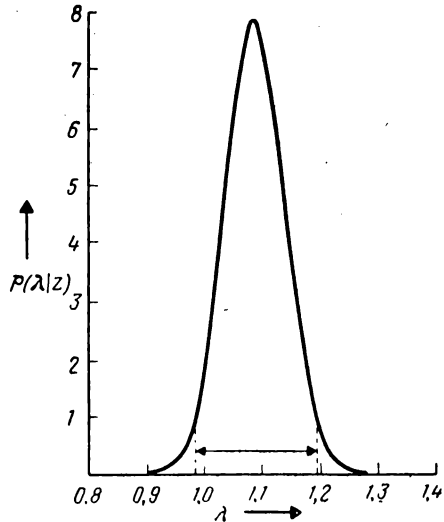


Рис. 7.10. Апостериорная плотность $p(\lambda | z)$ для ряда В.

Возможность $\phi = \theta$ связана, таким образом, с неограниченным диапазоном для (равных) параметров. Эффект ограничения пространства параметров введением, например, условий стационарности и обратимости ($-1 < \phi < 1$, $-1 < \theta < 1$) даст малые положительные значения плотности, но это уточнение вряд ли практически существенно.

Байесовский анализ еще раз подтверждает вывод разд. 7.3.5, что при оценивании смешанных моделей могут встретиться трудности; в частности, они могут встретиться при итеративном поиске решений, когда близка избыточность параметров. Мы уже видели, что использование предварительной идентификации позволяет избежать этих трудностей.

ПРИЛОЖЕНИЕ П7.1. ОБЗОР ТЕОРИИ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ

П7.1.1. Разбиение положительно определенной квадратичной формы

Рассмотрим положительно определенную квадратичную форму $Q_p = \mathbf{x}' \Sigma^{-1} \mathbf{x}$. Пусть вектор \mathbf{x} размера $p \times 1$ разбит на две части после p_1 -го элемента, так что $\mathbf{x}' = (\mathbf{x}_1' : \mathbf{x}_2') = (x_1, x_2, \dots, x_{p_1}, x_{p_1+1}, \dots, x_p)$, и пусть матрица Σ размера $p \times p$ также разбита на части после p_1 -й строки и p_1 -го столбца, так что

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}' & \Sigma_{22} \end{bmatrix}.$$

Тогда, так как

$$\begin{aligned} \mathbf{x}' \Sigma^{-1} \mathbf{x} &= (\mathbf{x}_1' : \mathbf{x}_2') \begin{bmatrix} \mathbf{I} & -\Sigma_{11}^{-1} \Sigma_{12} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \Sigma_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & (\Sigma_{22} - \Sigma_{12}' \Sigma_{11}^{-1} \Sigma_{12})^{-1} \end{bmatrix} \times \\ &\quad \times \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ -\Sigma_{12}' \Sigma_{11}^{-1} & \mathbf{I} \end{bmatrix} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix}. \end{aligned}$$

$Q_p = \mathbf{x}' \Sigma^{-1} \mathbf{x}$ всегда может быть представлена как сумма двух квадратичных форм Q_{p_1} и Q_{p_2} , содержащих соответственно p_1 и p_2 элементов, где

$$Q_p = Q_{p_1} + Q_{p_2},$$

$$Q_{p_1} = \mathbf{x}_1' \Sigma_{11}^{-1} \mathbf{x}_1, \quad (\text{П7.1.1})$$

$$Q_{p_2} = (\mathbf{x}_2 - \Sigma_{12}' \Sigma_{11}^{-1} \mathbf{x}_1)' (\Sigma_{22} - \Sigma_{12}' \Sigma_{11}^{-1} \Sigma_{12})^{-1} (\mathbf{x}_2 - \Sigma_{12}' \Sigma_{11}^{-1} \mathbf{x}_1).$$

Определитель Σ можно представить в виде

$$|\Sigma| = |\Sigma_{11}| |\Sigma_{22} - \Sigma_{12}' \Sigma_{11}^{-1} \Sigma_{12}|. \quad (\text{П7.1.2})$$

П7.1.2. Два полезных интеграла

Пусть $\mathbf{z}'\mathbf{C}\mathbf{z}$ — положительно определенная квадратичная форма от \mathbf{z} из q элементов, так что $\mathbf{z}' = (z_1, z_2, \dots, z_q)$, где $-\infty < z_i < \infty$, $i = 1, 2, \dots, q$, и пусть a , b и m — положительные действительные числа. Тогда можно показать, что

$$\int_R \left\{ a + \frac{\mathbf{z}'\mathbf{C}\mathbf{z}}{b} \right\}^{-(m+q)/2} d\mathbf{z} = \frac{(b\pi)^{q/2} \Gamma(m/2)}{a^{m/2} |\mathbf{C}|^{1/2} \Gamma((m+q)/2)}, \quad (\text{П7.1.3})$$

где q -кратный интеграл берется по всему пространству R возможных \mathbf{z} и

$$\frac{\int_{\mathbf{z}'\mathbf{C}\mathbf{z} > qF_0} \left\{ 1 + \frac{\mathbf{z}'\mathbf{C}\mathbf{z}}{m} \right\}^{-(m+q)/2} d\mathbf{z}}{\int_R \left\{ 1 + \frac{\mathbf{z}'\mathbf{C}\mathbf{z}}{m} \right\}^{-(m+q)/2} d\mathbf{z}} = \int_{F_0}^{\infty} p(F|q, m) dF, \quad (\text{П7.1.4})$$

где функция $p(F|q, m)$ называется F -распределением с q и m степенями свободы и определена как

$$p(F|q, m) = \frac{(q/m)^{q/2} \Gamma((m+q)/2)}{\Gamma(q/2) \Gamma(m/2)} F^{(q-2)/2} \left(1 + \frac{q}{m} F \right)^{-(m+q)/2}. \quad (\text{П7.1.5})$$

Если $m \rightarrow \infty$, то

$$\left\{ 1 + \frac{\mathbf{z}'\mathbf{C}\mathbf{z}}{m} \right\}^{-(m+q)/2} \rightarrow e^{-(\mathbf{z}'\mathbf{C}\mathbf{z})/2},$$

и, обозначив $qF = \chi^2$, получаем из (П7.1.4)

$$\frac{\int_{\mathbf{z}'\mathbf{C}\mathbf{z} > \chi_0^2} e^{-(\mathbf{z}'\mathbf{C}\mathbf{z})/2} d\mathbf{z}}{\int_R e^{-(\mathbf{z}'\mathbf{C}\mathbf{z})/2} d\mathbf{z}} = \int_{\chi_0^2}^{\infty} p(\chi^2|q) d\chi^2, \quad (\text{П7.1.6})$$

где функция $p(\chi^2|q)$ называется χ^2 -распределением с q степенями свободы и определена как

$$p(\chi^2|q) = \frac{1}{2^{q/2} \Gamma(q/2)} (\chi^2)^{(q-2)/2} e^{-\chi^2/2}. \quad (\text{П7.1.7})$$

Здесь и далее $p(x)$ используется как *общее* обозначение для функции плотности вероятности случайной величины x .

П7.1.3. Нормальное распределение

Говорят, что случайная величина x распределена нормально со средним μ и стандартным отклонением σ или имеет распределение $N(\mu, \sigma^2)$, если ее плотность вероятности равна

$$p(x) = (2\pi)^{-1/2} (\sigma^2)^{-1/2} e^{-(x-\mu)^2/2\sigma^2}. \quad (\text{П7.1.8})$$

Отсюда нормированная величина $u = (x - \mu)/\sigma$ имеет распределение $N(0, 1)$. Табл. Е в конце книги дает ординаты $p(u = u_\epsilon)$ и значения u_ϵ , для которых $P\{u > u_\epsilon\} = \epsilon$ для заданного ϵ .

Многомерное нормальное распределение. Говорят, что векторная случайная величина $x' = (x_1, x_2, \dots, x_p)$ имеет совместное p -мерное нормальное распределение $N\{\mu, \Sigma\}$, если плотность вероятности равна

$$p(x) = (2\pi)^{-p/2} |\Sigma|^{-1/2} e^{-(x-\mu)' \Sigma^{-1} (x-\mu)/2}. \quad (\text{П7.1.9})$$

Изоповерхности плотности вероятности — эллипсоиды, определяемые уравнениями $(x - \mu)' \Sigma^{-1} (x - \mu) = \text{const}$. В качестве

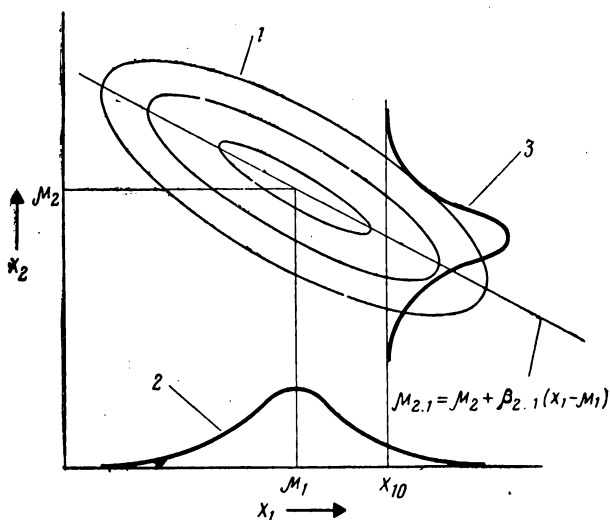


Рис. П7.1. Изолинии двумерного нормального распределения (1); там же показаны маргинальное распределение $p_1(x_1)$ (2) и условное распределение $p(x_2|x_{10})$ при $x_1 = x_{10}$ (3).

иллюстрации на рис. П7.1 показаны эллиптические изолинии для двумерного нормального распределения.

В точке $x = \mu$ многомерное распределение имеет максимальную плотность вероятности

$$\max p(x) = p(\mu) = (2\pi)^{-p/2} |\Sigma|^{-1/2}.$$

Распределение χ^2 как вероятность непопадания в область, ограниченную изоповерхностью многомерного нормального распределения. Для p -мерного нормального распределения (П7.1.9) вероятность непопадания в область, ограниченную поверхностью, заданной уравнением

$$(x - \mu)' \Sigma^{-1} (x - \mu) = \chi_0^2,$$

равна χ^2 -интегралу с p степенями свободы

$$\int_{\chi_0^2}^{\infty} p(\chi^2 | p) d\chi^2,$$

где плотность χ^2 -распределения определена формулой (П7.1.7). В табл. F в конце книги приведены значения $\chi_e^2(p)$, для которых при заданном ϵ $P\{\chi^2 > \chi_e^2(p)\} = \epsilon$.

Маргинальные и условные распределения для многомерного нормального распределения. Положим, что вектор из $p = p_1 + p_2$ случайных величин разбит на две части после p_1 -го элемента, так что $\mathbf{x}' = (\mathbf{x}'_1; \mathbf{x}'_2) = (x_1, x_2, \dots, x_{p_1}; x_{p_1+1}, \dots, x_{p_1+p_2})$ и матрица ковариаций имеет вид

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma'_{12} & \Sigma_{22} \end{bmatrix}.$$

Тогда, пользуясь (П7.1.1) и (П7.1.2), можно записать многомерное нормальное распределение для $p = p_1 + p_2$ величин как *маргинальное* распределение \mathbf{x}_1 , умноженное на *условное* распределение \mathbf{x}_2 при данном \mathbf{x}_1 , т. е.

$$\begin{aligned} p(\mathbf{x}) &= p(\mathbf{x}_1, \mathbf{x}_2) = p(\mathbf{x}_1) p(\mathbf{x}_2 | \mathbf{x}_1) = \\ &= (2\pi)^{-p_1/2} |\Sigma_{11}|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{x}_1 - \mu_1)' \Sigma_{11}^{-1}(\mathbf{x}_1 - \mu_1)\right\} \times \\ &\times (2\pi)^{-p_2/2} |\Sigma_{22.11}|^{-1/2} \exp\left\{-\frac{1}{2}(\mathbf{x}_2 - \mu_{2.1})' \Sigma_{22.11}^{-1}(\mathbf{x}_2 - \mu_{2.1})\right\}, \quad (\text{П7.1.10}) \end{aligned}$$

где

$$\Sigma_{22.11} = \Sigma_{22} - \Sigma'_{12} \Sigma_{11}^{-1} \Sigma_{12} \quad (\text{П7.1.11})$$

и $\mu_{2.1} = \mu_2 + \beta_{2.1}(\mathbf{x}_1 - \mu_1) = E(\mathbf{x}_2 | \mathbf{x}_1)$ определяет гиперплоскость регрессии в $p_1 + p_2$ -мерном пространстве, которая прослеживает точку среднего значения p_2 элементов \mathbf{x}_2 , в то время как p_1 элементов \mathbf{x}_1 изменяются. Матрица коэффициентов регрессии размером $p_2 \times p_1$ определяется формулой $\beta_{2.1} = \Sigma'_{12} \Sigma_{11}^{-1}$.

Как маргинальное, так и условное распределения для многомерного нормального закона сами являются многомерными нормальными распределениями. Видно, что для *многомерного нормального распределения* условное распределение $p(\mathbf{x}_2 | \mathbf{x}_1)$ с точностью до сдвига сохраняется при любом \mathbf{x}_1 .

Одномерные маргинальные плотности. В частности, маргинальная плотность для одного элемента x_i ($i = 1, 2, \dots, p$) равна $N(\mu_i, \sigma_i^2)$ — одномерной нормальной плотности со средним значением, равным i -му элементу μ , и дисперсией, равной i -му диагональному элементу Σ .

Двумерное нормальное распределение. В качестве примера на рис. П7.1 показаны маргинальное и условное распределения для двумерного нормального распределения. В этом случае маргинальное распределение x_1 есть $N(\mu_1, \sigma_1^2)$, а условное распределение x_2 при данном x_1 равно

$$N\left\{\mu_2 + \rho \frac{\sigma_2}{\sigma_1}(x_1 - \mu_1), \sigma_2^2(1 - \rho^2)\right\},$$

где $\rho = (\sigma_1/\sigma_2)\beta_{2.1}$ — коэффициент корреляции между x_1 и x_2 .

П7.1.4. Распределение Стьюдента

Говорят, что случайная величина x имеет нормированное t -распределение Стьюдента $t(\mu, s^2, \nu)$ со средним значением μ , нормирующим параметром s и ν степенями свободы, если

$$p(x) = (2\pi)^{-1/2} (s^2)^{-1/2} \left(\frac{\nu}{2}\right)^{-1/2} \Gamma\left(\frac{\nu+1}{2}\right) \Gamma^{-1}\left(\frac{\nu}{2}\right) \left\{1 + \frac{(x-\mu)^2}{\nu s^2}\right\}^{-(\nu+1)/2}. \quad (\text{П7.1.12})$$

Отсюда стандартное t -отклонение $t = (x - \mu)/s$ имеет распределение $t(0, 1, \nu)$. В табл. Г в конце книги приведены значения t_e , для которых при заданном ϵ $P(t > t_e) = \epsilon$.

Переход к нормальному распределению. Для больших ν произведение

$$\left(\frac{\nu}{2}\right)^{-1/2} \Gamma\left(\frac{\nu+1}{2}\right) \Gamma^{-1}\left(\frac{\nu}{2}\right)$$

стремится к единице, в то время как крайний справа множитель в (П7.1.12) стремится к $e^{-(1/2s^2)(x-\mu)^2}$. Поэтому если мы для больших ν примем $s^2 = \sigma^2$, то t -распределение стремится к нормальному распределению (П7.1.8).

Многомерное t -распределение. Пусть μ' будет вектором $(\mu_1, \mu_2, \dots, \mu_p)$ размером $(p \times 1)$ и S — положительно определенной матрицей размером $(p \times p)$. Говорят, что векторная случайная величина x [68, 69] имеет нормированное t -распределение $t(\mu, S, \nu)$ с вектором средних значений μ , нормирующей матрицей S и ν степенями свободы, если

$$p(x) = (2\pi)^{-p/2} |S|^{-1/2} \left(\frac{\nu}{2}\right)^{-p/2} \Gamma\left(\frac{\nu+p}{2}\right) \Gamma^{-1}\left(\frac{\nu}{2}\right) \times \\ \times \left\{1 + \frac{(x-\mu)' S^{-1} (x-\mu)}{\nu}\right\}^{-(\nu+p)/2}. \quad (\text{П7.1.13})$$

Изоповерхности плотности вероятности многомерного t -распределения — эллипсоиды, определяемые уравнением

$$(x - \mu)' S^{-1} (x - \mu) = \text{const.}$$

Переход к многомерному нормальному распределению. Для больших ν произведение

$$\left(\frac{\nu}{2}\right)^{-p/2} \Gamma\left(\frac{\nu+p}{2}\right) \Gamma^{-1}\left(\frac{\nu}{2}\right)$$

стремится к единице; крайняя правая скобка в (П7.1.13) стремится к $e^{-(\mathbf{x}-\boldsymbol{\mu})' \mathbf{S}^{-1}(\mathbf{x}-\boldsymbol{\mu})/2}$. Отсюда, если для больших ν примем $\mathbf{S} = \boldsymbol{\Sigma}$, многомерное t -распределение стремится к многомерному нормальному распределению (П7.1.9).

F-распределение как вероятность непадания в область, ограниченную изоповерхностью многомерного t -распределения. Пользуясь (П7.1.4), можно выразить вероятность непадания в область, ограниченную изоповерхностью p -мерного t -распределения $t(\boldsymbol{\mu}, \mathbf{S}, \nu)$, заданную уравнением

$$(\mathbf{x} - \boldsymbol{\mu})' \mathbf{S}^{-1}(\mathbf{x} - \boldsymbol{\mu}) = pF_0,$$

как F -интеграл с p и ν степенями свободы

$$\int_{F_0}^{\infty} p(F | p, \nu) dF,$$

где функция плотности для F определена формулой (П7.1.5). Для больших ν $P\{F > F_e(p, \nu)\} = P\{\chi^2 > \chi_e^2(p)\}$, где $pF = \chi^2$. Отсюда, как и следовало ожидать, вероятность непадания в область, ограниченную заданной изоповерхностью t -распределения, для больших ν равна аналогичной вероятности для многомерного нормального распределения, к которому стремится многомерное t -распределение.

Маргинальное t -распределение. Если $p = p_1 + p_2$ -мерный вектор \mathbf{x} , распределенный, согласно уравнению (П7.1.13), разбить на две части после p_1 -го элемента, так что $\mathbf{x}' = (\mathbf{x}'_1; \mathbf{x}'_2)$, и разбить \mathbf{S} подобным образом, получим

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}'_{12} & \mathbf{S}_{22} \end{bmatrix}.$$

Обозначив

$$\mathbf{S}_{22.11} = \mathbf{S}_{22} - \mathbf{S}'_{12} \mathbf{S}_{11}^{-1} \mathbf{S}_{12},$$

$$\boldsymbol{\mu}_{2.1} = \boldsymbol{\mu}_2 + \boldsymbol{\beta}_{2.1}(\mathbf{x}_1 - \boldsymbol{\mu}_1), \quad \boldsymbol{\beta}_{2.1} = \mathbf{S}'_{12} \mathbf{S}_{11}^{-1},$$

получим

$$\begin{aligned} p(\mathbf{x}_1, \mathbf{x}_2) &= (2\pi)^{-(p_1+p_2)/2} |\mathbf{S}_{11}|^{-1/2} |\mathbf{S}_{22.11}|^{-1/2} \left(\frac{\nu}{2}\right)^{-(p_1+p_2)/2} \times \\ &\quad \times \Gamma\left(\frac{\nu+p_1+p_2}{2}\right) \Gamma^{-1}\left(\frac{\nu}{2}\right) \times \\ &\quad \times \left\{ 1 + \frac{(\mathbf{x}_1 - \boldsymbol{\mu}_1)' \mathbf{S}_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1)}{\nu} + \frac{(\mathbf{x}_2 - \boldsymbol{\mu}_{2.1})' \mathbf{S}_{22.11}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_{2.1})}{\nu} \right\}^{-(\nu+p_1+p_2)/2}. \end{aligned}$$

Используя теперь предварительный результат (П7.1.3) с

$$a = 1 + \frac{(\mathbf{x}_1 - \boldsymbol{\mu}_1)' \mathbf{S}_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1)}{v},$$

$b = v$, $m = v + p_1$, $q = p_2$, $\mathbf{C} = \mathbf{S}_{22.11}^{-1}$, получим

$$p(\mathbf{x}_1) = (2\pi)^{-p_1/2} |\mathbf{S}_{11}|^{-1/2} \left(\frac{v}{2}\right)^{-p_1/2} \Gamma\left(\frac{v + p_1}{2}\right) \Gamma^{-1}\left(\frac{v}{2}\right) \times \\ \times \left\{ 1 + \frac{(\mathbf{x}_1 - \boldsymbol{\mu}_1)' \mathbf{S}_{11}^{-1} (\mathbf{x}_1 - \boldsymbol{\mu}_1)}{v} \right\}^{-(v+p_1)/2}. \quad (\text{П7.1.14})$$

Итак, если p -мерный вектор \mathbf{x} имеет многомерное t -распределение (П7.1.13), маргинальное распределение любых p_1 переменных, образующих вектор \mathbf{x}_1 , — это p_1 -мерное t -распределение $t(\boldsymbol{\mu}_1, \mathbf{S}_{11}, v)$.

Одномерные маргинальные плотности. В частности, маргинальное распределение для одного элемента x_i ($i = 1, 2, \dots, p$) будет $t(\mu_i, s_{ii}, v)$ — одномерное t -распределение со средним значением, равным i -му элементу $\boldsymbol{\mu}$, нормирующим множителем, равным положительному значению квадратного корня из i -го диагонального элемента \mathbf{S} , и v степенями свободы.

Условное t -распределение. Условное распределение $p(\mathbf{x}_2 | \mathbf{x}_1)$ можно получить как отношение совместного распределения $p(\mathbf{x}_1, \mathbf{x}_2)$ и маргинального распределения $p(\mathbf{x}_1)$:

$$p(\mathbf{x}_2 | \mathbf{x}_{10}) = \frac{p(\mathbf{x}_{10}, \mathbf{x}_2)}{p(\mathbf{x}_{10})}.$$

Поделив, получаем

$$p(\mathbf{x}_2 | \mathbf{x}_{10}) = \text{const} \left\{ 1 + \frac{(\mathbf{x}_2 - \boldsymbol{\mu}_{2.1})' \{c(\mathbf{x}_{10}) \mathbf{S}_{22.11}^{-1}\} (\mathbf{x}_2 - \boldsymbol{\mu}_{2.1})}{v + p_1} \right\}^{-(v+p_1+p_2)/2}, \quad (\text{П7.1.15})$$

где $c(\mathbf{x}_{10})$ (для данного \mathbf{x}_{10} — это фиксированная константа) равно

$$c(\mathbf{x}_{10}) = \frac{v + p_1}{v + (\mathbf{x}_{10} - \boldsymbol{\mu}_1)' \mathbf{S}_{11}^{-1} (\mathbf{x}_{10} - \boldsymbol{\mu}_1)}. \quad (\text{П7.1.16})$$

Тогда распределение \mathbf{x}_2 при данном $\mathbf{x}_1 = \mathbf{x}_{10}$ — многомерное t -распределение:

$$t\{\boldsymbol{\mu}_2 + \boldsymbol{\beta}_{2.1}(\mathbf{x}_{10} - \boldsymbol{\mu}_1), c^{-1}(\mathbf{x}_{10}) \mathbf{S}_{22.11}, v + p_1\},$$

где $\boldsymbol{\beta}_{2.1} = \mathbf{S}'_{12} \mathbf{S}_{11}^{-1}$.

Двумерное t -распределение. Как и раньше, некоторое представление об общей многомерной ситуации можно получить, изучив двумерное распределение $t(\boldsymbol{\mu}, \mathbf{S}, v)$. Диаграммное представление, аналогичное показанному на рис. П7.1, на первый

взгляд подобно двумерному нормальному распределению. Однако маргинальные распределения — это одномерные t -распределения, имеющие v степеней свободы, в то время как условные распределения — это t -распределения с $v + 1$ степенями свободы. Далее нормирующий множитель для условного распределения $p(x_2|x_{10})$, например, зависит от x_{10} . Это — явное отличие от условного распределения для нормального случая, где дисперсия не зависит от x_{10} .

ПРИЛОЖЕНИЕ П7.2. ОБЗОР ЛИНЕЙНОЙ ТЕОРИИ НАИМЕНЬШИХ КВАДРАТОВ

П7.2.1. Нормальные уравнения

Предполагается, что модель имеет вид

$$w_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + e_i, \quad (\text{П7.2.1})$$

где w_i ($i = 1, 2, \dots, n$) — наблюдения, полученные из эксперимента, в котором независимые переменные $x_{i1}, x_{i2}, \dots, x_{ik}$ принимали известные *фиксированные* значения; β_i — неизвестные коэффициенты, которые нужно оценить из данных, и e_i — некоррелированные ошибки с нулевыми средними значениями и одинаковой дисперсией σ^2 .

Уравнения (П7.2.1) можно представить в матричной форме

$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1k} \\ x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

или

$$\mathbf{w} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}, \quad (\text{П7.2.2})$$

где предполагается, что \mathbf{X} имеет полный ранг k . Теорему наименьших квадратов Гаусса можно сформулировать так [70]: оценки $\hat{\boldsymbol{\beta}}' = (\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k)$ параметров $\boldsymbol{\beta}$, линейные относительно наблюдений и минимизирующие среднеквадратичную ошибку любой линейной функции параметров $\lambda_1\beta_1 + \lambda_2\beta_2 + \dots + \lambda_k\beta_k$, можно получить, минимизируя сумму квадратов

$$S(\boldsymbol{\beta}) = \mathbf{e}'\mathbf{e} = (\mathbf{w} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{w} - \mathbf{X}\boldsymbol{\beta}). \quad (\text{П7.2.3})$$

Для определения минимума $S(\boldsymbol{\beta})$ разложим вектор $\mathbf{w} - \mathbf{X}\boldsymbol{\beta}$ на два вектора $\mathbf{w} - \mathbf{X}\hat{\boldsymbol{\beta}}$ и $\mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})$:

$$\mathbf{w} - \mathbf{X}\boldsymbol{\beta} = \mathbf{w} - \mathbf{X}\hat{\boldsymbol{\beta}} - \mathbf{X}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}). \quad (\text{П7.2.4})$$

Если мы выберем $\hat{\beta}$ так, что

$$(\mathbf{X}'\mathbf{X})\hat{\beta} = \mathbf{X}'\mathbf{w}, \quad (\text{П7.2.5})$$

то

$$S(\beta) = S(\hat{\beta}) + (\beta - \hat{\beta})' \mathbf{X}'\mathbf{X} (\beta - \hat{\beta}) \quad (\text{П7.2.6})$$

и векторы $\mathbf{w} - \mathbf{X}\hat{\beta}$ и $\mathbf{X}(\beta - \hat{\beta})$ ортогональны. Так как второй член в правой части (П7.2.6) — положительно определенная квадратичная форма, отсюда следует, что минимум $S(\beta)$ достигается при $\beta = \hat{\beta}$, где $\hat{\beta}$ определяется *нормальными уравнениями* (П7.2.5).

П7.2.2. Оценка остаточной дисперсии

Согласно (П7.2.3) и (П7.2.5), сумма квадратов в точке минимума равна

$$S(\hat{\beta}) = \mathbf{w}'\mathbf{w} - \hat{\beta}'\mathbf{X}'\mathbf{X}\hat{\beta}. \quad (\text{П7.2.7})$$

Далее, если мы введем

$$s^2 = S(\hat{\beta})/(n - k), \quad (\text{П7.2.8})$$

можно показать [71], что $E[s^2] = \sigma^2$ и, следовательно, s^2 — несмещенная оценка σ^2 .

П7.2.3. Ковариационная матрица оценок

Такая матрица определяется как

$$\begin{aligned} \mathbf{V}(\hat{\beta}) &= \text{cov}[\hat{\beta}, \hat{\beta}'] = \\ &= \text{cov}[(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{w}, \mathbf{w}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}] = (\mathbf{X}'\mathbf{X})^{-1} \sigma^2, \end{aligned} \quad (\text{П7.2.9})$$

так как $\text{cov}(\mathbf{w}, \mathbf{w}') = \mathbf{I}\sigma^2$.

П7.2.4. Доверительные области

В предположении о нормальности [71] квадратичные формы $S(\hat{\beta})$ и $(\beta - \hat{\beta})' \mathbf{X}'\mathbf{X} (\beta - \hat{\beta})$ в (П7.2.6) независимо распределены как χ^2 соответственно с $n - k$ и k степенями свободы. Отсюда

$$\frac{(\beta - \hat{\beta})' \mathbf{X}'\mathbf{X} (\beta - \hat{\beta})}{S(\hat{\beta})} \frac{n - k}{k}$$

распределено как $F(k, n - k)$. Из (П7.2.8) следует, что

$$(\beta - \hat{\beta})' (\mathbf{X}'\mathbf{X}) (\beta - \hat{\beta}) \leq k s^2 F_\varepsilon(k, n - k) \quad (\text{П7.2.10})$$

определяет $(1 - \varepsilon)$ -ю доверительную область для β .

П7.2.5. Коррелированные ошибки

Пусть ошибки имеют известную ковариационную матрицу V , где $V^{-1} = PP'/\sigma^2$. Тогда (П7.2.2) можно записать иначе:

$$P'w = P'X\beta + P'e$$

или

$$w^* = X^*\beta + e^*. \quad (\text{П7.2.11})$$

Ковариационная матрица $e^* = P'e$ равна

$$\text{cov}[P'e, e'P] = P'VP = I\sigma^2.$$

Следовательно, можно применять обычную теорию наименьших квадратов с $V = I\sigma^2$ к трансформированной модели (П7.2.11), в которой w заменено на $w^* = P'w$ и X на $X^* = P'X$.

ПРИЛОЖЕНИЕ П7.3. ПРИМЕРЫ ВЛИЯНИЯ ОШИБОК ОЦЕНИВАНИЯ ПАРАМЕТРОВ НА ВЕРОЯТНОСТНЫЕ ПРЕДЕЛЫ ПРОГНОЗОВ

Дисперсии и вероятностные пределы для прогнозов, указанные в разд. 5.2.4, основывались на предположении, что параметры (ϕ, θ) модели АРПСС известны точно. На практике их необходимо заменять выборочными оценками $(\hat{\phi}, \hat{\theta})$. Для того чтобы получить некоторое представление о влиянии ошибок оценивания на дисперсию ошибок прогноза, мы рассмотрим частный случай нестационарного процесса ПСС(0, 1, 1) и стационарного процесса АР(1). Будет показано, что для этих процессов в случае, когда оценки параметров основаны на рядах не слишком малой длины, эффект ошибок оценивания мал.

Процессы ПСС(0, 1, 1). Записав выражения $\nabla z_t = a_t - \theta a_{t-1}$ для $t+l, t+l-1, \dots, t+1$ и сложив их, получим

$$z_{t+l} - z_t = a_{t+l} + (1 - \theta)(a_{t+l-1} + \dots + a_{t+1}) - \theta a_t.$$

Обозначим через $\hat{z}_t(l|\theta)$ прогноз с упреждением l в случае, когда параметр θ известен точно. Взяв условные математические ожидания в момент t для $l = 1, 2, \dots$, получим

$$\hat{z}_t(1|\theta) = z_t - \theta a_t, \quad \hat{z}_t(l|\theta) = \hat{z}_t(1|\theta), \quad l \geq 2.$$

Отсюда ошибка прогноза с упреждением l равна

$$\begin{aligned} e_t(l|\theta) &= z_{t+l} - \hat{z}_t(l|\theta) = \\ &= a_{t+l} + (1 - \theta)(a_{t+l-1} + \dots + a_{t+1}), \end{aligned}$$

и дисперсия ошибки этого прогноза имеет вид

$$V(l) = E_t[e_t^2(l|\theta)] = \sigma_a^2 \{1 + (l-1)\lambda^2\}, \quad (\text{П7.3.1})$$

где $\lambda = 1 - \theta$.

Однако если заменить θ ее выборочной оценкой $\hat{\theta}$, полученной по ряду из n наблюдений $w_t = \nabla z_t$, то

$$\hat{z}_t(1|\hat{\theta}) = z_t - \hat{\theta} \hat{a}_t, \quad \hat{z}_t(l|\hat{\theta}) = \hat{z}_t(1|\hat{\theta}), \quad l \geq 2,$$

где $\hat{a}_t = z_t - \hat{z}_{t-1}(1|\hat{\theta})$. Отсюда ошибка прогноза с упреждением l , использующего $\hat{\theta}$, будет

$$\begin{aligned} e_t(l|\hat{\theta}) &= z_{t+l} - \hat{z}_t(l|\hat{\theta}) = \\ &= z_{t+l} - z_t + \hat{\theta} \hat{a}_t = \\ &= e_t(l|\theta) - (\theta a_t - \hat{\theta} \hat{a}_t). \end{aligned} \quad (\text{П7.3.2})$$

Так как $\nabla z_t = (1 - \theta B) a_t = (1 - \hat{\theta} B) \hat{a}_t$, отсюда следует, что

$$\hat{a}_t = \left(\frac{1 - \theta B}{1 - \hat{\theta} B} \right) a_t,$$

и, исключая из (П7.3.2) \hat{a}_t , получим

$$\hat{e}_t(l|\hat{\theta}) = e_t(l|\theta) - \frac{(\theta - \hat{\theta})}{1 - \hat{\theta} B} a_t.$$

Далее,

$$\begin{aligned} \frac{\theta - \hat{\theta}}{1 - \hat{\theta} B} a_t &= \frac{\theta - \hat{\theta}}{1 - \theta B} \left\{ 1 + \frac{(\theta - \hat{\theta}) B}{1 - \theta B} \right\}^{-1} a_t = \\ &= \frac{\theta - \hat{\theta}}{1 - \theta B} \left\{ 1 - \frac{(\theta - \hat{\theta}) B}{1 - \theta B} \right\} a_t = \\ &= (\theta - \hat{\theta})(a_t + \theta a_{t-1} + \theta^2 a_{t-2} + \dots) - \\ &- (\theta - \hat{\theta})^2(a_{t-1} + 2\theta a_{t-2} + 3\theta^2 a_{t-3} + \dots). \end{aligned} \quad (\text{П7.3.3})$$

В предположении, что прогноз и оценка $\hat{\theta}$ по существу основаны на непересекающихся данных, $\hat{\theta}$ и a_t, a_{t-1}, \dots будут независимы. Кроме того, для не слишком малых выборок $\hat{\theta}$ приближенно нормально распределено относительно среднего значения θ с дисперсией $(1 - \theta^2)/n$. При этом можно показать, что дисперсия выражения (П7.3.3) имеет вид

$$\frac{\sigma_a^2}{n} \left(1 + \frac{3}{n} \frac{1 + \theta^2}{1 - \theta^2} \right).$$

При условии, что θ не слишком близко к единице,

$$\text{var} [e_t(l|\hat{\theta})] \approx \sigma_a^2 \{ 1 + (l-1) \lambda^2 \} + \frac{\sigma_a^2}{n}. \quad (\text{П7.3.4})$$

Ясно, что пропорциональное изменение дисперсии будет наибольшим для $l = 1$, когда точная ошибка прогноза уменьшается

до σ_a^2 . В этом случае для выборочных оценок параметров, основанных на не слишком коротких рядах, вероятностные пределы увеличиваются в $(n+1)/n$ раз.

Процессы авторегрессии первого порядка. Записав выражения для модели $\tilde{z}_t = \phi \tilde{z}_{t-1} + a_t$ для момента $t+l$ и перейдя к условным математическим ожиданиям в момент t , получаем прогноз с упреждением l , основанный на точном значении параметра,

$$\hat{\tilde{z}}_t(l|\phi) = \phi \hat{\tilde{z}}_t(l-1|\phi) = \phi^l \tilde{z}_t.$$

Аналогично

$$\hat{\tilde{z}}_t(l|\hat{\phi}) = \hat{\phi} \hat{\tilde{z}}_t(l-1|\hat{\phi}) = \hat{\phi}^l \tilde{z}_t,$$

и, следовательно,

$$e_t(l|\hat{\phi}) = e_t(l|\phi) + (\phi^l - \hat{\phi}^l) \tilde{z}_t.$$

Отсюда следует, что

$$E_t[e_t^2(l|\hat{\phi})] = E_t[e_t^2(l|\phi)] + \tilde{z}_t^2 E_t[(\phi^l - \hat{\phi}^l)^2],$$

так что в среднем с учетом (5.4.16)

$$\text{var}[e_t(l|\hat{\phi})] \approx \sigma_a^2 \frac{(1 - \phi^{2l})}{(1 - \phi^2)} + \sigma_a^2 \frac{E[(\phi^l - \hat{\phi}^l)^2]}{1 - \phi^2}. \quad (\text{П7.3.5})$$

Когда $l=1$,

$$\text{var}[e_t(1|\hat{\phi})] \approx \sigma_a^2 + \left(\frac{\sigma_a^2}{1 - \phi^2} \right) \left(\frac{1 - \phi^2}{n} \right) = \sigma_a^2 \left(1 + \frac{1}{n} \right). \quad (\text{П7.3.6})$$

Для $l > 1$ имеем

$$\begin{aligned} \phi^l - \hat{\phi}^l &= \phi^l - \{\phi - (\phi - \hat{\phi})\}^l \approx \phi^l - \phi^l \left\{ 1 - \frac{\phi - \hat{\phi}}{\phi} \right\}^l \approx \\ &\approx l \phi^{l-1} (\phi - \hat{\phi}). \end{aligned}$$

Итак, в среднем

$$\text{var}[e_t(l|\hat{\phi})] \approx \text{var}[e_t(l|\phi)] + \frac{l^2 \phi^{2(l-1)}}{n} \sigma_a^2,$$

и различие опять порядка n^{-1} .

ПРИЛОЖЕНИЕ П7.4. ТОЧНАЯ ФУНКЦИЯ ПРАВДОПОДОБИЯ ДЛЯ ПРОЦЕССА СКОЛЬЗЯЩЕГО СРЕДНЕГО

Чтобы получить функцию правдоподобия в этом случае, необходимо найти выражение для функции плотности вероятности ряда $\mathbf{w}'_n = (w_1, w_2, \dots, w_n)$ в предположении, что он генерируется стационарной моделью скользящего среднего порядка q

$$\tilde{w}_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}, \quad (\text{П7.4.1})$$

где $\tilde{w}_t = w_t - \mu$, $\mu = E[w_t]$. В предположении, что a и, следовательно, \tilde{w} распределены нормально, совместную плотность вероятности можно записать в виде

$$p(\mathbf{w}_n | \theta, \sigma_a^2) = (2\pi\sigma_a^2)^{-n/2} |\mathbf{M}_n^{(0, q)}|^{1/2} \exp\left(-\frac{\tilde{\mathbf{w}}_n' \mathbf{M}_n^{(0, q)} \tilde{\mathbf{w}}_n}{2\sigma_a^2}\right), \quad (\text{П7.4.2})$$

где $(\mathbf{M}_n^{(0, q)})^{-1} \sigma_a^2$ — ковариационная матрица \mathbf{w} размером $n \times n$. Во многих приложениях, где $w_t = \nabla^d z_t$, $d > 0$, разумно предполагать, что $\mu = 0$ и, следовательно, $\tilde{w}_t = w_t$. Когда это не предполагается и, следовательно, $E[w_t]$ не известно, μ можно рассматривать просто как дополнительный оцениваемый параметр, входящий в вектор θ . Рассмотрим теперь удобный способ вычисления $\tilde{\mathbf{w}}_n' \mathbf{M}_n^{(0, q)} \tilde{\mathbf{w}}_n$ и для простоты положим, что $\mu = 0$ и, следовательно, $w_t = \tilde{w}_t$.

Пользуясь моделью (П7.4.1), можно написать $n + q$ уравнений

$$\begin{aligned} a_{1-q} &= a_{1-q}, \\ &\vdots \\ a_{-1} &= a_{-1}, \\ a_0 &= a_0, \\ a_1 &= w_1 + \theta_1 a_0 + \theta_2 a_{-1} + \dots + \theta_q a_{1-q}, \\ a_2 &= w_2 + \theta_1 a_1 + \theta_2 a_0 + \dots + \theta_q a_{2-q}, \\ &\vdots \\ a_n &= w_n + \theta_1 a_{n-1} + \theta_2 a_{n-2} + \dots + \theta_q a_{n-q}. \end{aligned}$$

После подстановки выражения для a_1 в выражение для a_2 , выражений для a_1 и a_2 — в a_3 и т. д. мы можем записать $n + q$ -мерный вектор $\mathbf{a}' = (a_{1-q}, a_{2-q}, \dots, a_n)$ через n -мерный вектор $\mathbf{w}'_n = (w_1, w_2, \dots, w_n)$ и через q -мерный вектор предварительных значений $\mathbf{a}'_* = (a_{1-q}, a_{2-q}, \dots, a_0)$. Имеем

$$\mathbf{a} = \mathbf{L} \mathbf{w}_n + \mathbf{X} \mathbf{a}_*,$$

где \mathbf{L} — матрица размером $(n + q) \times n$ и \mathbf{X} — матрица размером $(n + q) \times q$, элементы которой — функции элементов θ , временно считающихся зафиксированными.

Совместное распределение $n + q$ величин, являющихся элементами \mathbf{a} , равно

$$p(\mathbf{a} | \sigma_a^2) = (2\pi\sigma_a^2)^{-(n+q)/2} \exp\left(-\frac{\mathbf{a}' \mathbf{a}}{2\sigma_a^2}\right).$$

Замечая, что наше преобразование имеет единичный якобиан, находим совместное распределение \mathbf{w}_n и \mathbf{a}_* :

$$p(\mathbf{w}_n, \mathbf{a}_* | \theta, \sigma_a) = (2\pi\sigma_a^2)^{-(n+q)/2} \exp\left\{-\frac{S(\theta, \mathbf{a}_*)}{2\sigma_a^2}\right\},$$

где

$$S(\theta, \mathbf{a}_*) = (\mathbf{L}\mathbf{w}_n + \mathbf{X}\mathbf{a}_*)'(\mathbf{L}\mathbf{w}_n + \mathbf{X}\mathbf{a}_*). \quad (\text{П7.4.3})$$

Пусть $\hat{\mathbf{a}}_*$ — вектор значений, минимизирующих $S(\theta, \mathbf{a}_*)$. Тогда, пользуясь (П7.2.6), получаем

$$S(\theta, \mathbf{a}_*) = S(\theta) + (\mathbf{a}_* - \hat{\mathbf{a}}_*)' \mathbf{X}'\mathbf{X}(\mathbf{a}_* - \hat{\mathbf{a}}_*),$$

где

$$S(\theta) = (\mathbf{L}\mathbf{w}_n + \mathbf{X}\hat{\mathbf{a}}_*)'(\mathbf{L}\mathbf{w}_n + \mathbf{X}\hat{\mathbf{a}}_*) \quad (\text{П7.4.4})$$

— функция только наблюдений \mathbf{w}_n , но не предварительных значений \mathbf{a}_* . Итак,

$$\begin{aligned} p(\mathbf{w}_n, \mathbf{a}_* | \theta, \sigma_a) &= \\ &= (2\pi\sigma_a^2)^{-(n+q)/2} \exp\left\{-\frac{1}{2\sigma_a^2}[S(\theta) + (\mathbf{a}_* - \hat{\mathbf{a}}_*)' \mathbf{X}'\mathbf{X}(\mathbf{a}_* - \hat{\mathbf{a}}_*)]\right\}. \end{aligned}$$

Однако, так как

$$p(\mathbf{w}_n, \mathbf{a}_* | \theta, \sigma_a) = p(\mathbf{w}_n | \theta, \sigma_a) p(\mathbf{a}_* | \mathbf{w}_n, \theta, \sigma_a),$$

то отсюда следует, что

$$\begin{aligned} p(\mathbf{a}_* | \mathbf{w}_n, \theta, \sigma_a) &= \\ &= (2\pi\sigma_a^2)^{-q/2} |\mathbf{X}'\mathbf{X}|^{1/2} \exp\left[-\frac{(\mathbf{a}_* - \hat{\mathbf{a}}_*)' \mathbf{X}'\mathbf{X}(\mathbf{a}_* - \hat{\mathbf{a}}_*)}{2\sigma_a^2}\right], \quad (\text{П7.4.5}) \end{aligned}$$

$$p(\mathbf{w}_n | \theta, \sigma_a) = (2\pi\sigma_a^2)^{-n/2} |\mathbf{X}'\mathbf{X}|^{-1/2} \exp\left\{-\frac{S(\theta)}{2\sigma_a^2}\right\}. \quad (\text{П7.4.6})$$

Мы приходим к следующим выводам.

1) Из (П7.4.5) видно, что $\hat{\mathbf{a}}_*$ — это условное ожидание \mathbf{a}_* при данных \mathbf{w}_n и θ . Пользуясь обозначениями, введенными в разд. 7.1.4, получаем

$$\hat{\mathbf{a}}_* = [\mathbf{a}_* | \mathbf{w}_n, \theta] = [\mathbf{a}_*],$$

откуда $[\mathbf{a}] = \mathbf{L}\mathbf{w}_n + \mathbf{X}[\mathbf{a}_*]$, и, пользуясь (П7.4.4), находим

$$S(\theta) = \sum_{t=1-q}^n [a_t]^2. \quad (\text{П7.4.7})$$

Хотя $\hat{\mathbf{a}}_*$ можно получить непосредственно методом наименьших квадратов, на практике их легче вычислять, используя тот факт, что $\hat{\mathbf{a}}_* = [\mathbf{a}_*]$, и получая $[\mathbf{a}_*]$ по методике «прогнозирования назад», описанной в разд. 7.1.4 и 7.1.5.

2) Сравнивая (П7.4.6) и (П7.4.2), получаем

$$|\mathbf{X}'\mathbf{X}|^{-1} = |\mathbf{M}_n^{(0, q)}|$$

и

$$S(\theta) = \mathbf{w}_n' \mathbf{M}_n^{(0, q)} \mathbf{w}_n.$$

3) Чтобы вычислить

$$S(\theta) = \sum_{t=1-q}^n [a_t]^2,$$

можно получить величины $[a_t] = [a_t | \mathbf{w}_n, \theta]$, пользуясь оценками $[\mathbf{a}_*]' = ([a_{1-q}], [a_{2-q}], \dots, [a_0])$ для предварительных значений, найденных прогнозированием назад, и вычислив элементы $[a_1], [a_2], \dots, [a_n]$ рекуррентным путем по формулам

$$[a_t] = [w_t] + \theta_1 [a_{t-1}] + \theta_2 [a_{t-2}] + \dots + \theta_q [a_{t-q}],$$

где $[w_t] = w_t$ ($t = 1, 2, \dots, n$).

4) Наконец, пользуясь (П7.4.6) и (П7.4.7), получаем точное выражение для безусловной функции правдоподобия

$$L(\theta, \sigma_a | \mathbf{w}_n) = (\sigma_a^2)^{-n/2} |\mathbf{X}'\mathbf{X}|^{-1/2} \exp \left\{ -\frac{1}{2\sigma_a^2} \sum_{t=1-q}^n [a_t]^2 \right\}. \quad (\text{П7.4.8})$$

Так, если $q = 1$, легко показать, что \mathbf{X} состоит из $(n+1)$ -мерного вектора-столбца с элементами $1, \theta, \theta^2, \dots, \theta^n$. Тогда $\mathbf{X}'\mathbf{X} = [1 - \theta^{2(n+1)}] / (1 - \theta^2)$. Для $n > 50$ членом $\theta^{2(n+1)}$ обычно можно пренебречь, и мы получаем достаточно точный результат:

$$L(\theta, \sigma_a | \mathbf{w}_n) = (\sigma_a^2)^{-n/2} (1 - \theta^2)^{1/2} \exp \left\{ -\frac{1}{2\sigma_a^2} \sum_{t=0}^n [a_t]^2 \right\}. \quad (\text{П7.4.9})$$

Обобщение на процессы авторегрессии и смешанные процессы. Описанный выше метод получения безусловной функции правдоподобия легко обобщить на общую смешанную модель

$$\phi(B) \tilde{w}_t = \theta(B) a_t, \quad (\text{П7.4.10})$$

которая при $w_t = \nabla^d z_t$ определяет общий процесс АРПСС. Заметим сначала, что эта модель может быть записана при помощи бесконечного оператора скользящего среднего

$$\tilde{w}_t = (1 + \psi_1 B + \psi_2 B^2 + \dots) a_t. \quad (\text{П7.4.11})$$

Для процессов, представляющих интерес во многих задачах, веса ψ_1, ψ_2, \dots быстро уменьшаются, так что (П7.4.11) с любой желаемой степенью точности можно аппроксимировать *конечным* процессом скользящего среднего некоторого порядка Q :

$$\tilde{w}_t = (1 + \psi_1 B + \psi_2 B^2 + \dots + \psi_Q B^Q) a_t.$$

Подходящее значение Q можно выбрать, исходя из факта, что для конечного скользящего среднего порядка Q наблюдения через $Q + 1$ интервал некоррелированы, и, следовательно, прогнозы $[\tilde{w}]$ для упреждений, больших чем Q , будут равны нулю. В процессе вычислений, описанных в разд. 7.1.5, находятся прогнозы для времен $0, -1, -2, \dots, -Q + 1$, и Q выбирается как точка, за которой прогнозируемые значения $[\tilde{w}]$ становятся практически нулями. Поэтому не возникает необходимости в вычислении ψ , так как рекуррентный расчет осуществляется прямо по выражению для общей модели (П7.4.10). Пример расчета дан в разд. 7.1.5.

Следовательно, в общем случае функция правдоподобия для ряда z из $n + d$ значений, генерируемого любым процессом АРПСС, имеет вид

$$L(\phi, \theta, \sigma_a | z) = (2\pi\sigma_a^2)^{-n/2} |M_n^{(p, q)}|^{1/2} \exp \left\{ -\frac{S(\phi, \theta)}{2\sigma_a^2} \right\}, \quad (\text{П7.4.12})$$

где

$$S(\phi, \theta) = \sum_{t=-\infty}^n [a_t | z, \phi, \theta]^2 \quad (\text{П7.4.13})$$

и значения $[a_t | z, \phi, \theta] = [a_t | w_n, \phi, \theta]$ практически можно вычислить рекуррентным способом с суммированием, начинающимся от некоторой точки $t = 1 - Q$, за которой $[a_t]$ пренебрежимо малы.

В качестве примера рассмотрим процесс авторегрессии первого порядка для w_t :

$$w_t - \phi w_{t-1} = a_t, \quad (\text{П7.4.14})$$

где w_t может быть d -й разностью $\nabla^d z_t$ истинных наблюдений, и мы располагаем рядом z из $n + d$ наблюдений. Для того чтобы вычислить функцию правдоподобия (П7.4.12), нам нужно знать

$$S(\phi) = \sum_{t=-\infty}^n [a_t | w_n, \phi]^2 = \sum_{t=-\infty}^1 [a_t | w_n, \phi]^2 + \sum_{t=2}^n (w_t - \phi w_{t-1})^2.$$

Искомые условные ожидания (прогнозы назад) равны

$$[w_j | w_n, \phi] = \phi^{1-j} w_1, \quad j = 0, -1, -2, \dots,$$

поэтому

$$[a_t | w_n, \phi] = \phi^{-t+1} (1 - \phi^2) w_1, \quad t = 1, 0, -1, \dots$$

Отсюда

$$S(\phi) = (1 - \phi^2) w_1^2 + \sum_{t=2}^n (w_t - \phi w_{t-1})^2 \quad (\text{П7.4.15})$$

— результат, который может быть получен и более явными способами, рассмотренными в приложении П7.5.

ПРИЛОЖЕНИЕ П7.5. ТОЧНАЯ ФУНКЦИЯ ПРАВДОПОДОБИЯ ДЛЯ ПРОЦЕССА АВТОРЕГРЕССИИ

Положим, что данный ряд $\mathbf{w}'_n = (w_1, w_2, \dots, w_n)$ генерируется стационарной моделью авторегрессии p -го порядка

$$w_t - \phi_1 w_{t-1} - \phi_2 w_{t-2} - \dots - \phi_p w_{t-p} = a_t,$$

причем временно предполагается, что w имеют среднее значение $\mu = 0$, но, как и прежде, допускается обобщение на случай $\mu \neq 0$. Предполагая, что a , а следовательно, и w подчиняются нормальному закону, получаем функцию совместной плотности вероятности:

$$p(\mathbf{w}_n | \boldsymbol{\phi}, \sigma_a) = (2\pi\sigma_a^2)^{-n/2} |\mathbf{M}_n^{(p,0)}|^{1/2} \exp \left\{ -\frac{\mathbf{w}'_n \mathbf{M}_n^{(p,0)} \mathbf{w}_n}{2\sigma_a^2} \right\}; \quad (\text{П7.5.1})$$

из-за обратимости общего процесса матрица $\mathbf{M}_n^{(p,0)}$ размером $n \times n$ симметрична относительно *обеих* главных диагоналей. Можно сказать, что такая матрица *дважды* симметрична. Имеем

$$p(\mathbf{w}_n | \boldsymbol{\phi}, \sigma_a) = p(w_{p+1}, w_{p+2}, \dots, w_n | \mathbf{w}_p, \boldsymbol{\phi}, \sigma_a) p(\mathbf{w}_p | \boldsymbol{\phi}, \sigma_a),$$

где $\mathbf{w}'_p = (w_1, w_2, \dots, w_p)$. Первый множитель в правой части можно получить, используя распределение

$$p(a_{p+1}, \dots, a_n) = (2\pi\sigma_a^2)^{-(n-p)/2} \exp \left\{ -\frac{1}{2\sigma_a^2} \sum_{t=p+1}^n a_t^2 \right\}. \quad (\text{П7.5.2})$$

Для фиксированных \mathbf{w}_p (a_{p+1}, \dots, a_n) и (w_{p+1}, \dots, w_n) связаны преобразованием

$$\begin{aligned} a_{p+1} &= w_{p+1} - \phi_1 w_p - \dots - \phi_p w_1, \\ &\vdots \\ a_n &= w_n - \phi_1 w_{n-1} - \dots - \phi_p w_{n-p} \end{aligned}$$

с единичным якобианом. Отсюда имеем

$$\begin{aligned} p(w_{p+1}, \dots, w_n | \mathbf{w}_p, \boldsymbol{\phi}, \sigma_a) &= \\ &= (2\pi\sigma_a^2)^{-(n-p)/2} \exp \left\{ -\frac{1}{2\sigma_a^2} \sum_{t=p+1}^n (w_t - \phi_1 w_{t-1} - \dots - \phi_p w_{t-p})^2 \right\}. \end{aligned}$$

Кроме того,

$$p(\mathbf{w}_p | \boldsymbol{\phi}, \sigma_a) = (2\pi\sigma_a^2)^{-p/2} |\mathbf{M}_p^{(p,0)}|^{1/2} \exp \left\{ -\frac{1}{2\sigma_a^2} \mathbf{w}'_p \mathbf{M}_p^{(p,0)} \mathbf{w}_p \right\}.$$

Тогда

$$p(\mathbf{w}_n | \boldsymbol{\phi}, \sigma_a) = (2\pi\sigma_a^2)^{-n/2} |\mathbf{M}_p^{(p,0)}|^{1/2} \exp\left\{-\frac{S(\boldsymbol{\phi})}{2\sigma_a^2}\right\}, \quad (\text{П7.5.3})$$

где

$$S(\boldsymbol{\phi}) = \sum_{i=1}^p \sum_{j=1}^p m_{ij}^{(p)} w_i w_j + \sum_{t=p+1}^n (w_t - \phi_1 w_{t-1} - \dots - \phi_p w_{t-p})^2. \quad (\text{П7.5.4})$$

Кроме того,

$$\begin{aligned} \mathbf{M}_p^{(p,0)} = \{m_{ij}^{(p)}\} &= \{\gamma_{|t-j|}\}^{-1} \sigma_a^2 = \\ &= \begin{bmatrix} \gamma_0 & \gamma_1 & \dots & \gamma_{p-1} \\ \gamma_1 & \gamma_0 & \dots & \gamma_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{p-1} & \gamma_{p-2} & \dots & \gamma_0 \end{bmatrix}^{-1} \sigma_a^2, \end{aligned} \quad (\text{П7.5.5})$$

где $\gamma_0, \gamma_1, \dots, \gamma_{p-1}$ — теоретические автоковариации процесса, а $|\mathbf{M}_p^{(p,0)}| = |\mathbf{M}_n^{(p,0)}|$.

Пусть $n = p + 1$, так что

$$\begin{aligned} \mathbf{w}_{p+1}' \mathbf{M}_{p+1}^{(p,0)} \mathbf{w}_{p+1} &= \sum_{i=1}^p \sum_{j=1}^p m_{ij}^{(p)} w_i w_j + \\ &+ (w_{p+1} - \phi_1 w_p - \phi_2 w_{p-1} - \dots - \phi_p w_1)^2. \end{aligned}$$

Тогда

$$\mathbf{M}_{p+1}^{(p)} = \begin{bmatrix} \mathbf{M}_p^{(p)} & \begin{smallmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{smallmatrix} \\ \begin{smallmatrix} 0 & 0 & \dots & 0 \end{smallmatrix} & 0 \end{bmatrix} + \begin{bmatrix} \phi_p^2 & \phi_p \phi_{p-1} & \dots & -\phi_p \\ \phi_p \phi_{p-1} & \phi_{p-1}^2 & \dots & -\phi_{p-1} \\ \vdots & \vdots & \ddots & \vdots \\ -\phi_p & -\phi_{p-1} & \dots & 1 \end{bmatrix},$$

и элементы $\mathbf{M}_p^{(p)} = \mathbf{M}_p^{(p,0)}$ можно найти, пользуясь соображением о двойной симметричности как $\mathbf{M}_p^{(p)}$, так и $\mathbf{M}_{p+1}^{(p)}$. Так, например,

$$M_2^{(1)} = \begin{bmatrix} m_{11}^{(1)} + \phi_1^2 & -\phi_1 \\ -\phi_1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -\phi_1 \\ -\phi_1 & m_{11}^{(1)} + \phi_1^2 \end{bmatrix},$$

и после приравнивания элементов двух матриц имеем

$$M_1^{(1)} = m_{11}^{(1)} = 1 - \phi_1^2.$$

Продолжая действовать таким же образом, для процессов порядка 1 и 2 находим

$$\begin{aligned} \mathbf{M}_1^{(1)} &= 1 - \phi_1^2, \quad |\mathbf{M}_1^{(1)}| = 1 - \phi_1^2, \\ \mathbf{M}_2^{(2)} &= \begin{bmatrix} 1 - \phi_2^2 & -\phi_1(1 + \phi_2) \\ -\phi_1(1 + \phi_2) & 1 - \phi_2^2 \end{bmatrix}, \\ |\mathbf{M}_2^{(2)}| &= (1 + \phi_2)^2 \{(1 - \phi_2)^2 - \phi_1^2\}. \end{aligned}$$

Например, для $p = 1$

$$p(\mathbf{w}_n | \phi, \sigma_a) = (2\pi\sigma_a^2)^{-n/2} (1 - \phi^2)^{1/2} \exp \left\{ -\frac{1}{2\sigma_a^2} \left[(1 - \phi^2) \omega_1^2 + \sum_{t=2}^n (\omega_t - \phi \omega_{t-1})^2 \right] \right\},$$

что совпадает с результатом (П7.4.15). Эта процедура должна привести к матрицам $\mathbf{M}_p^{(p)}$, элементы которых — квадратичные функции ϕ .

Из (П7.5.4) очевидно, что $S(\phi) = \mathbf{w}_n' \mathbf{M}_n^{(p)} \mathbf{w}_n$ — квадратичная форма не только от ω , но и от параметров ϕ . Если представить ϕ'_u в виде $\phi'_u = (1, \phi_1, \phi_2, \dots, \phi_p)$, становится очевидным, что для некоторой матрицы \mathbf{D} размером $(p+1) \times (p+1)$, элементы которой — квадратичные функции, справедливо равенство

$$\mathbf{w}_n' \mathbf{M}_n^{(p)} \mathbf{w}_n = \phi'_u \mathbf{D} \phi_u.$$

Запишем теперь

$$\mathbf{D} = \begin{bmatrix} D_{11} & -D_{12} & -D_{13} & \dots & -D_{1,p+1} \\ -D_{12} & D_{22} & D_{23} & \dots & D_{2,p+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -D_{1,p+1} & D_{2,p+1} & D_{3,p+1} & \dots & D_{p+1,p+1} \end{bmatrix}. \quad (\text{П7.5.6})$$

Рассмотрение (П7.5.4) показывает, что элементы D_{ij} — симметричные суммы квадратов и произведения членов с задержками, определяемые как

$$D_{ij} = D_{ji} = \omega_i \omega_j + \omega_{i+1} \omega_{j+1} + \dots + \omega_{n+1-i} \omega_{n+1-j}, \quad (\text{П7.5.7})$$

где сумма D_{ij} содержит $n - (i-1) - (j-1)$ слагаемых.

Наконец, мы можем записать точное выражение для плотности вероятности и, следовательно, для функции правдоподобия:

$$p(\mathbf{w}_n | \phi, \sigma_a) = L(\phi, \sigma_a | \mathbf{w}_n) = (2\pi\sigma_a^2)^{-n/2} |\mathbf{M}_p^{(p)}|^{1/2} \exp \left\{ \frac{-S(\phi)}{2\sigma_a^2} \right\}, \quad (\text{П7.5.8})$$

где

$$S(\boldsymbol{\phi}) = \mathbf{w}_p' \mathbf{M}_p^{(p)} \mathbf{w}_p + \sum_{t=p+1}^n (w_t - \phi_1 w_{t-1} - \dots - \phi_p w_{t-p})^2 = \boldsymbol{\phi}'_u \mathbf{D} \boldsymbol{\phi}_u, \quad (\text{П7.5.9})$$

и логарифмическая функция правдоподобия равна

$$l(\boldsymbol{\phi}, \sigma_a | \mathbf{w}_n) = -\frac{n}{2} \ln \sigma_a^2 + \frac{1}{2} \ln |\mathbf{M}_p^{(p)}| - \frac{S(\boldsymbol{\phi})}{2\sigma_a^2}. \quad (\text{П7.5.10})$$

Оценки максимального правдоподобия. Дифференцируя по σ_a и каждому из $\boldsymbol{\phi}$ в (П7.5.10), получаем

$$\frac{\partial l}{\partial \sigma_a} = -\frac{n}{\sigma_a} + \frac{S(\boldsymbol{\phi})}{\sigma_a^3}, \quad (\text{П7.5.11})$$

$$\frac{\partial l}{\partial \phi_j} = M_j + \sigma_a^{-2} (D_{1,j+1} - \phi_1 D_{2,j+1} - \dots - \phi_p D_{p+1,j+1}), \quad (\text{П7.5.12})$$

$$j = 1, 2, \dots, p,$$

где

$$M_j = \frac{\partial \left(\frac{1}{2} \ln |\mathbf{M}_p^{(p)}| \right)}{\partial \phi_j}.$$

Отсюда можно получить оценки максимального правдоподобия — для этого надо приравнять все производные нулю и решить получающиеся уравнения.

Из (П7.5.11) сразу получаем

$$\hat{\sigma}_a^2 = S(\hat{\boldsymbol{\phi}})/n. \quad (\text{П7.5.13})$$

Оценки $\hat{\boldsymbol{\phi}}$. При решении уравнения (П7.5.12) возникают трудности, так как в общем случае величины M_j ($j = 1, 2, \dots, p$) — сложные функции $\boldsymbol{\phi}$. Рассмотрим кратко три возможных приближенных способа решения.

1) Оценки наименьших квадратов.

В то время как математическое ожидание $S(\boldsymbol{\phi})$ пропорционально n , значение $|\mathbf{M}_p^{(p)}|$ от n не зависит; для средних и больших выборок в (П7.5.8) доминирует член, содержащий $S(\boldsymbol{\phi})$, а член с $|\mathbf{M}_p^{(p)}|$ сравнительно мал.

Если мы пренебрежем влиянием этого члена, то

$$l(\boldsymbol{\phi}, \sigma_a | \mathbf{w}_n) \approx -\frac{n}{2} \ln \sigma_a^2 - \frac{S(\boldsymbol{\phi})}{2\sigma_a^2}, \quad (\text{П7.5.14})$$

и оценки $\hat{\boldsymbol{\phi}}$ параметров $\boldsymbol{\phi}$, полученные путем максимизации (П7.5.14), совпадают с оценками наименьших квадратов, полученными минимизацией $S(\boldsymbol{\phi})$. Так, из (П7.5.9) следует, что

$S(\phi) = \phi'_u D \phi_u$, где D — матрица симметричных сумм квадратов и произведений размером $(p+1) \times (p+1)$, определенная в (П7.5.7). Дифференцируя, находим значения, дающие минимум сумм

$$\begin{aligned} D_{12} &= \hat{\phi}_1 D_{22} + \hat{\phi}_2 D_{23} + \dots + \hat{\phi}_p D_{2, p+1}, \\ D_{13} &= \hat{\phi}_1 D_{23} + \hat{\phi}_2 D_{33} + \dots + \hat{\phi}_p D_{3, p+1}, \\ D_{1, p+1} &= \hat{\phi}_1 D_{2, p+1} + \hat{\phi}_2 D_{3, p+1} + \dots + \hat{\phi}_p D_{p+1, p+1}, \end{aligned} \quad (\text{П7.5.15})$$

которые в очевидных матричных обозначениях можно записать как

$$d = D_p \hat{\phi},$$

так что

$$\hat{\phi} = D_p^{-1} d.$$

Эти оценки наименьших квадратов максимизируют апостериорную плотность вероятности (7.4.15).

2) Приближенные оценки максимального правдоподобия.

Вспомним полученный ранее результат (3.2.3), который можно записать в виде

$$\gamma_j - \phi_1 \gamma_{j-1} - \phi_2 \gamma_{j-2} - \dots - \phi_p \gamma_{j-p} = 0, \quad j > 0. \quad (\text{П7.5.16})$$

Беря математические ожидания от обеих частей (П7.5.12) и пользуясь тем, что $E[\partial l / \partial \phi_j] = 0$, получаем

$$\begin{aligned} M_j \sigma_a^2 + (n-j) \gamma_j - (n-j-1) \phi_1 \gamma_{j-1} - \\ - (n-j-2) \phi_2 \gamma_{j-2} - \dots - (n-j-p) \phi_p \gamma_{j-p} = 0. \end{aligned} \quad (\text{П7.5.17})$$

Умножив (П7.5.16) на n и вычтя результат из (П7.5.17), получим

$$M_j \sigma_a^2 = j \gamma_j - (j+1) \phi_1 \gamma_{j-1} - \dots - (j+p) \phi_p \gamma_{j-p}.$$

Далее, используя $D_{i+1, j+1} / (n-j-i)$ как оценку $\gamma_{|j-i|}$, получаем естественную оценку $M_j \sigma_a^2$:

$$j \frac{D_{1, j+1}}{n-j} - (j+1) \phi_1 \frac{D_{2, j+1}}{n-j-1} - \dots - (j+p) \phi_p \frac{D_{p+1, j+1}}{n-j-p}.$$

Подставляя эту оценку в (П7.5.12), находим

$$\begin{aligned} \frac{\partial l}{\partial \phi_j} \approx n \sigma_a^{-2} \left\{ \frac{D_{1, j+1}}{n-j} - \phi_1 \frac{D_{2, j+1}}{n-j-1} - \dots - \phi_p \frac{D_{p+1, j+1}}{n-j-p} \right\}, \\ (j=1, 2, \dots, p), \end{aligned} \quad (\text{П7.5.18})$$

что ведет к системе линейных уравнений в форме (П7.5.15), в которых вместо D_{ij} стоит

$$D_{ij}^* = nD_{ij}/\{n - (i - 1) - (j - 1)\}.$$

3) Оценки Юла — Уокера.

Наконец, если n не мало, мы можем приближенно заменить симметричные суммы квадратов и произведений в (П7.5.15) соответствующими оценками автоковариаций, умноженными на n .

Например, D_{ij} , где $|i - j| = k$, нужно заменить на $nc_k = \sum_{t=1}^{n-k} \tilde{w}_t \tilde{w}_{t+k}$.

Разделив все члены полученных уравнений на nc_0 , получим следующие соотношения, содержащие выборочные оценки автокорреляций $r_k = c_k/c_0$:

$$\begin{aligned} r_1 &= \hat{\phi}_1 + \hat{\phi}_2 r_1 + \dots + \hat{\phi}_p r_{p-1}, \\ r_2 &= \hat{\phi}_1 r_1 + \hat{\phi}_2 + \dots + \hat{\phi}_p r_{p-2}, \\ &\vdots \\ r_p &= \hat{\phi}_1 r_{p-1} + \hat{\phi}_2 r_{p-2} + \dots + \hat{\phi}_p. \end{aligned}$$

Это хорошо известные уравнения Юла — Уокера.

В матричных обозначениях (7.3.1) их можно записать как $\mathbf{r} = \mathbf{R}\hat{\boldsymbol{\phi}}$, и отсюда

$$\hat{\boldsymbol{\phi}} = \mathbf{R}^{-1}\mathbf{r}, \quad (\text{П7.5.19})$$

что соответствует уравнению (3.2.7), в котором \mathbf{p}_p заменено на \mathbf{r} , а \mathbf{P}_p на \mathbf{R} .

Чтобы продемонстрировать различия трех оценок, рассмотрим случай $p = 1$. Тогда $M_1\sigma_a^2 = -\gamma_1$, и в соответствии с (П7.5.12) точная оценка максимального правдоподобия для ϕ будет решением уравнения

$$-\gamma_1 + \sum_{t=2}^n w_t w_{t-1} - \phi \sum_{t=2}^{n-1} w_t^2 = 0.$$

Приближение (1) полностью игнорирует слагаемое γ_1 , что дает

$$\hat{\phi} = \frac{\sum_{t=2}^n w_t w_{t-1}}{\sum_{t=2}^{n-1} w_t^2} = \frac{D_{12}}{D_{22}}.$$

Приближение (2) соответствует подстановке вместо γ_1 его выборочной оценки $\sum_{t=2}^n w_t w_{t-1} / (n-1)$, что приводит к

$$\hat{\phi} = \frac{\sum_{t=2}^n w_t w_{t-1} / (n-1)}{\sum_{t=2}^n w_t^2 / (n-2)} = \frac{n-2}{n-1} \frac{D_{12}}{D_{22}}.$$

Приближение (3) заменяет числитель и знаменатель этой дроби стандартными выборочными оценками автоковариаций (2.1.10), что дает

$$\hat{\phi} = \frac{\sum_{t=2}^n w_t w_{t-1}}{\sum_{t=1}^n w_t^2} = \frac{c_1}{c_0} = r_1.$$

Обычно, как и в этом примере, для не слишком малых выборок различия между оценками, получаемыми путем разных приближений, малы. Как правило, мы пользуемся оценками наименьших квадратов из приближения (1). Эти оценки можно, конечно, вычислить и прямо по (П7.5.15). Однако, предполагая возможность расчетов на ЭВМ, вряд ли стоит рассматривать отдельно процессы авторегрессии; мы нашли, что более просто, даже в случае подгонки процессов авторегрессии, пользоваться общим итеративным алгоритмом, описанным в разд. 7.2.1, позволяющим находить оценки наименьших квадратов для любого процесса АРСС.

Оценки σ_a^2 . Пользуясь приближением (3), а также (П7.5.9) и (П7.5.13), находим

$$\hat{\sigma}_a^2 = \frac{S(\hat{\phi})}{n} = c_0 (1 : \hat{\phi}') \left[\begin{array}{c|c} 1 & -\mathbf{r}' \\ \hline -\mathbf{r} & \mathbf{R} \end{array} \right] \left(\begin{array}{c} 1 \\ \hline \hat{\phi} \end{array} \right).$$

Выполнив умножение в правой части и учтя, что $\mathbf{r} - \mathbf{R}\hat{\phi} = 0$, найдем

$$\hat{\sigma}_a^2 = c_0 (1 - \mathbf{r}'\hat{\phi}) = c_0 (1 - \mathbf{r}'\mathbf{R}^{-1}\mathbf{r}) = c_0 (1 - \hat{\phi}'\mathbf{R}\hat{\phi}). \quad (\text{П7.5.20a})$$

Легко показать, что σ_a^2 можно аналогичным образом выразить через теоретические корреляции

$$\sigma_a^2 = \gamma_0 (1 - \rho'\hat{\phi}) = \gamma_0 (1 - \rho'\mathbf{P}_p^{-1}\rho) = \gamma_0 (1 - \hat{\phi}'\mathbf{P}_p\hat{\phi}), \quad (\text{П7.5.20б})$$

что согласуется с результатом (3.2.8). Сходные выражения для $\hat{\sigma}_a^2$ можно получить при помощи приближений (1) и (2).

Информационная матрица. Дифференцируя (П7.5.11) и (П7.5.18) второй раз, получим

$$-\frac{\partial^2 l}{\partial \sigma_a^2} = -\frac{n}{\sigma_a^2} + \frac{3S(\boldsymbol{\phi})}{\sigma_a^4}, \quad (\text{П7.5.21a})$$

$$\frac{\partial^2 l}{\partial \sigma_a \partial \phi_j} \approx -\sigma_a^{-1} \frac{\partial l}{\partial \phi_j}, \quad (\text{П7.5.21б})$$

$$-\frac{\partial^2 l}{\partial \phi_i \partial \phi_j} \approx \frac{n}{\sigma_a^2} \left(\frac{D_{i+1, j+1}}{n-i-j} \right). \quad (\text{П7.5.21в})$$

Далее, так как

$$E \left[\frac{\partial l}{\partial \phi_j} \right] = 0,$$

отсюда следует, что для средних и больших выборок

$$E \left[-\frac{\partial^2 l}{\partial \sigma_a \partial \phi_j} \right] \approx 0$$

и

$$|I(\boldsymbol{\phi}, \sigma_a)| \approx |I(\boldsymbol{\phi})| I(\sigma_a),$$

где

$$I(\sigma_a) = E \left[-\frac{\partial^2 l}{\partial \sigma_a^2} \right] = \frac{2n}{\sigma_a^2}.$$

Далее, пользуясь (П7.5.21в), получим

$$I(\boldsymbol{\phi}) = -E \left[\frac{\partial^2 l}{\partial \phi_i \partial \phi_j} \right] = \frac{n}{\sigma_a^2} \boldsymbol{\Gamma}_p = \frac{n\gamma_0}{\sigma_a^2} \mathbf{P}_p = n(\mathbf{M}_p^{(p)})^{-1}. \quad (\text{П7.5.22})$$

Отсюда

$$|I(\boldsymbol{\phi}, \sigma_a)| = r \frac{2n^{p+1}}{\sigma_a^2} |\mathbf{M}_p^{(p, 0)}|^{-1}.$$

Дисперсии и ковариации выборочных оценок параметров авторегрессии. При условиях, подробно рассмотренных в [87], матрица, обратная информационной, является асимптотической матрицей ковариаций оценок максимального правдоподобия (МП). Более того, если логарифмическая функция правдоподобия приближенно квадратична и ее максимум не близок к границе, даже для выборки средних размеров элементы этой матрицы хорошо аппроксимируют дисперсии и ковариации оценок. Пользуясь (П7.5.22) и (П7.5.20б), находим

$$\begin{aligned} \mathbf{V}(\hat{\boldsymbol{\phi}}) &= \mathbf{I}^{-1}(\hat{\boldsymbol{\phi}}) \approx n^{-1} \mathbf{M}_p^{(p)} = n^{-1} \sigma_a^2 \boldsymbol{\Gamma}_p^{-1} = \\ &= n^{-1} (1 - \boldsymbol{\rho}' \mathbf{P}_p^{-1} \boldsymbol{\rho}) \mathbf{P}_p^{-1} = \\ &= n^{-1} (1 - \hat{\boldsymbol{\phi}}' \mathbf{P}_p \hat{\boldsymbol{\phi}}) \mathbf{P}_p^{-1} = n^{-1} (1 - \boldsymbol{\rho}' \hat{\boldsymbol{\phi}}) \mathbf{P}_p^{-1}. \end{aligned} \quad (\text{П7.5.23})$$

В частности, для процессов авторегрессии первого и второго порядков

$$V(\hat{\phi}) \approx n^{-1}(1 - \phi^2),$$

$$V(\hat{\phi}_1, \hat{\phi}_2) \approx n^{-1} \begin{bmatrix} 1 - \phi_2^2 & -\phi_1(1 + \phi_2) \\ -\phi_1(1 + \phi_2) & 1 - \phi_2^2 \end{bmatrix}. \quad (\text{П7.5.24})$$

Оценки дисперсий и ковариаций можно получить, заменив в (П7.5.24) параметры их выборочными оценками. Например, подставив r вместо ρ и $\hat{\phi}$ вместо ϕ в (П7.5.23), получим

$$\hat{V}(\hat{\phi}) = n^{-1}(1 - r'\hat{\phi})R^{-1}. \quad (\text{П7.5.25})$$

Распространение результатов для авторегрессий на более общие процессы. Имеется интересное соотношение, позволяющее нам распространить результаты оценивания для моделей авторегрессии на более общие модели. Пусть наблюдения генерируются моделью АРСС(p, q)

$$\phi(B)w_t = \theta(B)a_t, \quad (\text{П7.5.26})$$

где $w_t = \nabla^d z_t$. Модель можно представить в виде

$$\prod_{i=1}^p (1 - G_i B) w_t = \prod_{j=1}^q (1 - H_j B) a_t. \quad (\text{П7.5.27})$$

Пусть $H_1^0, H_2^0, \dots, H_q^0$ — значения параметров скользящего среднего, несколько отличающиеся от H_1, H_2, \dots, H_q , и положим, что делается преобразование

$$\prod_{j=1}^q (1 - H_j^0 B)^2 x_t = w_t. \quad (\text{П7.5.28})$$

Заменив неизвестные предварительные значения нулями, мы можем использовать это выражение для генерирования соответствующего ряда x для каждого временного ряда w из n наблюдений, порождаемого моделью (П7.5.27). Обозначим это преобразование через $T^{-1}x = w$, так что $x = Tw$, и так как T^{-1} — треугольная матрица с диагональными элементами, равными 1, $|T| = 1$. Примем $H_1 = H_1^0 + \delta_1, H_2 = H_2^0 + \delta_2, \dots, H_q = H_q^0 + \delta_q$. Тогда функция правдоподобия для параметров (G, H) , т. е. для параметров $(G, H^0 + \delta)$, будет

$$L(G, H^0 + \delta, \sigma_a^2 | w, (p, q)).$$

Такой выбор параметров показывает, что w_t генерируются моделью АРСС(p, q)

$$\prod_{i=1}^p (1 - G_i B) w_t = \prod_{j=1}^q \{1 - (H_j^0 + \delta_j) B\} a_t,$$

и, следовательно, x_t генерируются моделью

$$\prod_{i=1}^p (1 - G_i B) \prod_{j=1}^q (1 - H_j^0 B)^2 x_t = \prod_{j=1}^q \{1 - (H_j^0 + \delta_j) B\} a_t,$$

или с хорошей степенью приближения при малых δ моделью $AP(p + q)$

$$\prod_{i=1}^p (1 - G_i B) \prod_{j=1}^q \{1 - (H_j^0 - \delta_j) B\} x_t = a_t.$$

Соответствующая функция правдоподобия имеет вид

$$L\{\mathbf{G}, \mathbf{H}^0 - \delta, \sigma_a^2 | \mathbf{x}, (p + q, 0)\}.$$

Для средних и больших выборок приближения на концах ряда не будут оказывать существенного влияния; функция правдоподобия мала для больших δ , так что существенны только малые отклонения. Отсюда следует, что

$$L\{\mathbf{G}, \mathbf{H}^0 + \delta, \sigma_a^2 | \mathbf{w}, (p, q)\} \approx L\{\mathbf{G}, \mathbf{H}^0 - \delta, \sigma_a^2 | \mathbf{x}, (p + q, 0)\}. \quad (\text{П7.5.29})$$

Немедленно получаем следующие результаты.

1) Положим, что информационная матрица для параметров (\mathbf{G}, \mathbf{H}) , связанная с моделью $APCC(p, q)$

$$\prod_{i=1}^p (1 - G_i B) w_t = \prod_{j=1}^q (1 - H_j B) a_t,$$

есть $\mathbf{I}\{\mathbf{G}, \mathbf{H} | (p, q)\}$. Положим, что информационная матрица для параметров (\mathbf{G}, \mathbf{H}) чистой модели $AP(p + q)$

$$\prod_{i=1}^p (1 - G_i B) \prod_{j=1}^q (1 - H_j B) w_t = a_t$$

равна

$$\mathbf{I}\{\mathbf{G}, \mathbf{H} | (p + q, 0)\} = \left[\begin{array}{c|c} \mathbf{I}_{GG} & \mathbf{I}_{GH} \\ \hline \mathbf{I}'_{GH} & \mathbf{I}_{HH} \end{array} \right],$$

где матрица разбита на части после p -х строки и столбца. Тогда для средних и больших выборок

$$\mathbf{I}\{\mathbf{G}, \mathbf{H} | (p, q)\} \approx \mathbf{I}\{\mathbf{G}, -\mathbf{H} | (p + q, 0)\} = \left[\begin{array}{c|c} \mathbf{I}_{GG} & -\mathbf{I}_{GH} \\ \hline -\mathbf{I}'_{GH} & \mathbf{I}_{HH} \end{array} \right]. \quad (\text{П7.5.30})$$

2) На рассмотренном уровне приближения отсюда следует, что определители информационных матриц для процессов $APCC(p, q)$ и $AP(p + q)$ идентичны.

3) Так как для средних и больших выборок матрица, обратная информационной, является хорошим приближением

к ковариационной матрице параметров $V(G, H)$, получаем

$$V\{G, H | (p, q)\} \approx V\{G, -H | (p+q, 0)\}. \quad (\text{П7.5.31})$$

4) Запишем теперь

$$\prod_{i=1}^p (1 - G_i B) \prod_{j=1}^q (1 - H_j B) = 1 - \phi_1^* B - \phi_2^* B^2 - \dots - \phi_{p+q}^* B^{p+q}.$$

Тогда, пользуясь (П7.5.22), находим

$$|I(\phi^*)| \approx n^{p+q} |M_{p+q}^{(p+q, 0)}|^{-1} = n^{p+q} |M_n^{(p, q)}|^{-1}. \quad (\text{П7.5.32})$$

Кроме того,

$$|I(G, H)| = |I(\phi^*)| J^2(\phi^* | G, H), \quad (\text{П7.5.33})$$

где

$$J(\phi^* | G, H) = \left| \frac{\partial \phi^*}{\partial (G, H)} \right| \quad (\text{П7.5.34})$$

— якобиан преобразования.

Процессы скользящего среднего; дисперсии и ковариации оценок МП. Легко показать, что для больших выборок ковариационная матрица оценок МП чистого процесса скользящего среднего точно такая же, как для чистого процесса авторегрессии того же порядка. Тогда, согласно (П7.5.24), для процессов скользящего среднего первого и второго порядков имеем

$$V(\hat{\theta}) \approx n^{-1} (1 - \theta^2), \quad V(\hat{\theta}_1, \hat{\theta}_2) \approx n^{-1} \begin{bmatrix} 1 - \theta_2^2 & -\theta_1(1 + \theta_2) \\ -\theta_1(1 + \theta_2) & 1 - \theta_2^2 \end{bmatrix}.$$

Смешанные процессы авторегрессии — скользящего среднего; дисперсии и ковариации оценок МП. Для того чтобы проиллюстрировать, как используется (П7.5.31), рассмотрим процесс АРСС(1, 1)

$$(1 - \phi B) w_t = (1 - \theta B) a_t,$$

который мы сопоставим с процессом АР(2)

$$(1 - \phi B)(1 - \theta B) x_t = a_t. \quad (\text{П7.5.35})$$

Матрица ковариаций для $\hat{\phi}$ и $\hat{\theta}$ в модели авторегрессии (П7.5.35), согласно (7.2.22), имеет вид

$$\begin{aligned} V\{\hat{\phi}, \hat{\theta} | (2, 0)\} &\approx \\ &\approx n^{-1} \frac{1 - \phi\theta}{(\phi - \theta)^2} \begin{bmatrix} (1 - \phi^2)(1 - \phi\theta) & -(1 - \phi^2)(1 - \theta^2) \\ -(1 - \phi^2)(1 - \theta^2) & (1 - \theta^2)(1 - \phi\theta) \end{bmatrix}. \end{aligned}$$

Отсюда соответствующие дисперсии и ковариации для процесса АРСС(1, 1) равны

$$\begin{aligned} V\{\hat{\phi}, \hat{\theta} | (1, 1)\} &\approx \\ &\approx n^{-1} \frac{1 - \phi\theta}{(\phi - \theta)^2} \begin{bmatrix} (1 - \phi^2)(1 - \phi\theta) & (1 - \phi^2)(1 - \theta^2) \\ (1 - \phi^2)(1 - \theta^2) & (1 - \theta^2)(1 - \phi\theta) \end{bmatrix}. \end{aligned}$$

Диагностическая проверка модели

После того как модель идентифицирована и параметры оценены, подгоняемая модель подвергается *диагностической проверке*. Один из полезных методов проверки модели состоит в использовании *избыточного числа параметров*, т. е. в оценивании параметров для несколько более общей модели, чем ожидаемая. Этот метод исходит из того, что мы можем угадать неадекватные свойства модели. Следовательно, полезно дополнить этот подход более общими способами проверки, основанными на остаточных ошибках после подгонки модели. Они позволяют найти в данных указание, какие изменения модели необходимы. Мы опишем две такие проверки, которые используют 1) автокорреляционную функцию остаточных ошибок; 2) кумулятивную периодограмму остаточных ошибок.

8.1. ПРОВЕРКА СТОХАСТИЧЕСКИХ МОДЕЛЕЙ

8.1.1. Общие принципы

Предположим, что при помощи методов, описанных в гл. 6 и 7, по конкретному временному ряду была идентифицирована модель и оценены ее параметры. Остается решить вопрос, адекватна ли эта модель. Если будут обнаружены свидетельства серьезной неадекватности, возникнет необходимость узнать, как нужно изменить модель на следующем итеративном цикле. То, что мы делаем, только отчасти описывается словами «критерий согласия». Нам нужно обнаружить, *что именно* неадекватно в модели, чтобы узнать, как ее изменить. Например, среди известных процедур, не относящихся к анализу временных рядов, изучение остаточных ошибок в дисперсионном анализе, описанное Анскомбом и Тьюки [72, 73], диаграммы на нормальной вероятностной бумаге и другие методы, предложенные Даниелем [74] на основе критики факторного анализа, могут быть названы диагностическими проверками.

Никакая модель не может быть абсолютно правильной. Отсюда следует, что при достаточном объеме данных статистическое испытание может отвергнуть модель, которая тем не менее

вполне адекватна решаемой задаче. Напротив, испытания могут не выявить серьезных отклонений от сделанных предположений, потому что эти испытания нечувствительны к некоторым видам отклонений. Лучший путь — разработать как можно более чувствительные статистические процедуры, но быть готовым, если к тому имеются достаточные основания, использовать модели, в которых число параметров слегка занижено.

Ясно, что диагностические проверки должны ставить модель «под угрозу», т. е. быть чувствительными к вероятным отклонениям. Никакая система диагностической проверки не может быть всеобъемлющей, поскольку всегда можно просмотреть некоторые необычные свойства данных. Однако, если тщательно разработанные диагностические проверки были применены к модели, подогнанной к достаточно большому объему данных, и не выявили серьезных отклонений, мы имеем большие основания пользоваться этой моделью.

8.1.2. Введение избыточных параметров

Один из методов, который может быть использован для диагностической проверки, — это введение в модель *избыточных* параметров. Идентифицировав модель и считая ее правильной, мы затем подгоняем более сложную модель. Это ставит под угрозу идентифицированную модель, потому что более сложная модель содержит дополнительные параметры, с помощью которых можно ликвидировать возможные отклонения. Нужно тщательно продумать вопрос о том, как следует усложнить модель. В частности, в согласии с дискуссией об избыточности модели в разд. 7.3.5, было бы глупо добавлять множители в обе части уравнения, описывающего модель АРСС, *одновременно*. Если анализ не выявил, какие добавления нужны, это, конечно, не означает, что наша модель верна. Модель может лишь выдержать испытание. Все, что мы можем делать, — это «испытывать все, крепко держась за хорошее» (послание св. Павла к фессалонянам).

Пример введения избыточности. В качестве примера рассмотрим опять данные о стоимости акций IBM. Для этого анализа использовались данные, приведенные как ряд B' в сборнике временных рядов в конце книги. Этот ряд состоит из биржевых цен акций IBM за период *) с 29 июня 1959 г. по 30 июня 1960 г. Модель $(0, 1, 1)$

$$\nabla z_t = (1 - \theta B) a_t$$

*) Данные об акциях IBM, рассмотренные ранее и названные рядом B , покрывают другой период, а именно с 17 мая 1961 г. по 2 ноября 1962 г.

с $\hat{\lambda}_0 = 1 - \hat{\theta} = 0,90$ была идентифицирована и подогнана к 255 имевшимся наблюдениям.

Модель $(0, 1, 1)$ можно также представить в проинтегрированной форме (4.3.3):

$$z_t = \lambda_0 S a_{t-1} + a_t.$$

Более общая модель, рассмотренная в процедуре введения избыточных параметров, — это процесс $(0, 3, 3)$

$$\nabla^3 z_t = (1 - \theta_1 B - \theta_2 B^2 - \theta_3 B^3) a_t,$$

или в проинтегрированной форме

$$z_t = \lambda_0 S a_{t-1} + \lambda_1 S^2 a_{t-1} + \lambda_2 S^3 a_{t-1} + a_t.$$

Непосредственной причиной такого усложнения модели было желание проверить предположение, высказанное Брауном [2], что ряд следует прогнозировать при помощи подстраиваемой *квадратичной* прогнозирующей функции. В гл. 5 было показано, что для процесса ПСС $(0, q, q)$ оптимальная прогнозирующая функция — это подстраиваемый полином степени $q = 1$. Тогда для указанной выше избыточной модели $(0, 3, 3)$ оптимальная прогнозирующая функция с упреждением l — это квадратный трехчлен от l :

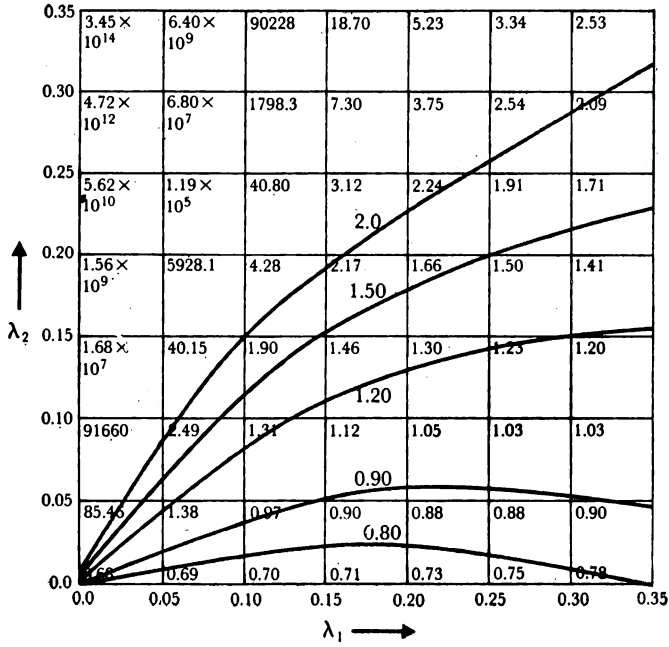
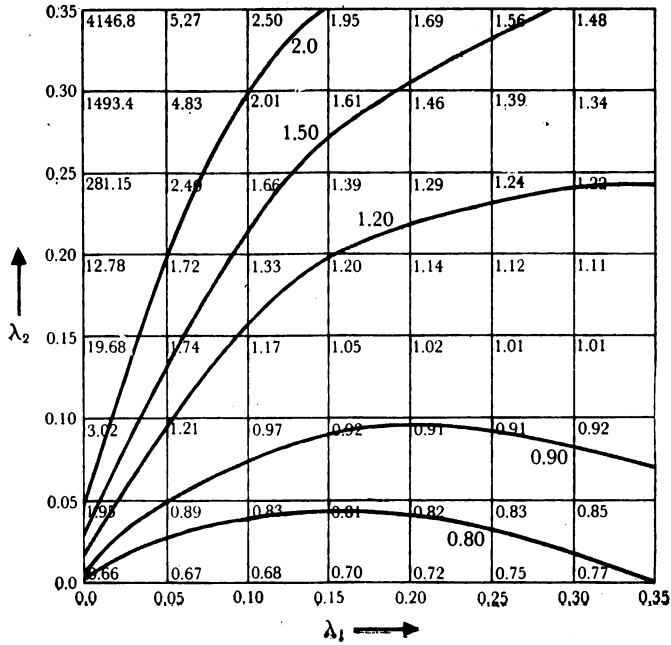
$$\hat{z}_t(l) = b_0^{(t)} + b_1^{(t)} l + b_2^{(t)} l^2,$$

где коэффициенты $b_0^{(t)}$, $b_1^{(t)}$, $b_2^{(t)}$ подстраиваются по мере поступления новых данных.

Однако модель, идентифицированная нами как процесс ПСС $(0, 1, 1)$, имеет прогнозирующую функцию

$$\hat{z}_t(l) = b_0^{(t)}. \quad (8.1.1)$$

Это «полином от l » нулевой степени. Следовательно, модель указывает, что прогноз $\hat{z}_t(l)$ не зависит от l , т. е. в любое время t прогноз на шаг вперед, на два шага вперед и т. д. одинаков. Другими словами, ряд содержит информацию только о своем будущем *уровне*, но не о наклоне или кривизне. На первый взгляд это странно, поскольку, оглядываясь в прошлое ряда, можно увидеть вполне явные линейный и криволинейные тренды. Следовательно, целесообразно проверить, действительно ли существуют ненулевые значения λ_1 и λ_2 , которые могли бы вызвать тренды. Сетка с величинами суммы квадратов $S(\lambda_1, \lambda_2 | \lambda_0)$ для значений $\lambda_0 = 0,7; 0,9; 1,1$ показана на рис. 8.1; можно видеть, что минимум близок к $\hat{\lambda}_0 = 0,9$, $\hat{\lambda}_1 = 0$, $\hat{\lambda}_2 = 0$. Также ясно, что значения $\lambda_1 > 0$ и $\lambda_2 > 0$ приводят к большим суммам квадратов и, следовательно, отклонения от идентифицированной модели ПСС $(0, 1, 1)$ в этом направлении

(a) $\lambda_0 = 0.7$ (b) $\lambda_0 = 0.9$ 

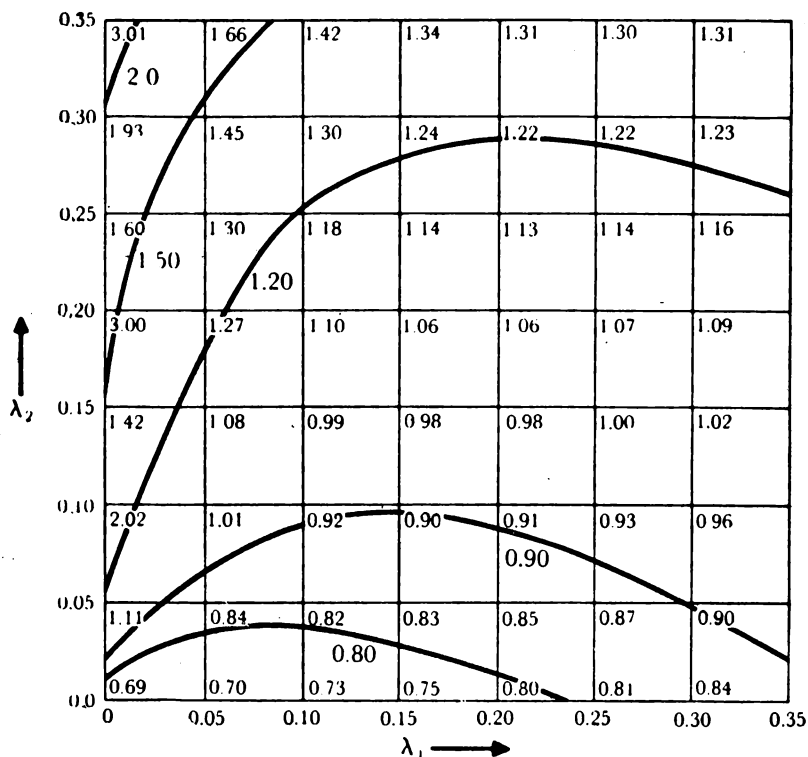
(с) $\lambda_0 = 1.1$ 

Рис. 8.1. Сетка и изолинии суммы квадратов для ряда В' и расширенной модели порядка (0,3,3).

противопоказаны. В частности, это свидетельствует о том, что квадратичная прогнозирующая функция будет ухудшать, а не улучшать прогноз по сравнению с (8.1.1), что и было подтверждено в разд. П5.3.3.

8.2. ДИАГНОСТИЧЕСКИЕ ПРОВЕРКИ, ПРИМЕНЯЕМЫЕ К ОСТАТОЧНЫМ ОШИБКАМ

Методы введения избыточных параметров путем обобщения модели в некотором направлении хороши в том случае, если мы знаем, какого рода отклонений следует бояться. Методики, менее зависящие от таких представлений, основываются на анализе *остаточных ошибок*. Хотя некоторые представления о том,

чего искать, полезны и в этом случае, эти методы создают больше возможностей находить способ изменения моделей непосредственно из данных.

8.2.1. Проверка при помощи автокорреляций

Пусть подгонялась модель

$$\phi(B)\tilde{w}_t = \theta(B)a_t$$

с $w_t = \nabla^d z_t$, и были получены оценки МП параметров $(\hat{\phi}, \hat{\theta})$. Тогда величины

$$\hat{a}_t = \hat{\theta}^{-1}(B)\hat{\phi}(B)\tilde{w}_t \quad (8.2.1)$$

мы будем называть *остаточными ошибками*. Можно доказать, что для адекватной модели

$$\hat{a}_t = a_t + O\left(\frac{1}{\sqrt{n}}\right).$$

По мере увеличения длины ряда \hat{a}_t становится все ближе к белому шуму a_t . Следовательно, можно ожидать, что изучение \hat{a}_t может выявить и указать природу неадекватности модели. В частности, некоторые характерные черты выборочной автокорреляционной функции \hat{a}_t могут указать на необходимость определенных изменений модели. Этот вопрос рассмотрен подробнее в разд. 8.3.

Предположим, что модель верна, и мы *знали* точные значения параметров ϕ и θ . Тогда, пользуясь (2.1.13) и результатом Андерсона [52], получаем, что выборочные автокорреляции $r_k(a)$ ряда a должны быть некоррелированы и распределены приблизительно нормально относительно нулевого среднего значения с дисперсией n^{-1} и, следовательно, со стандартной ошибкой $n^{-1/2}$. Эти факты можно использовать для оценки статистической значимости кажущихся отклонений этих автокорреляций от нуля.

На практике мы не знаем *истинных* значений параметров. Мы располагаем только выборочными оценками $(\hat{\phi}, \hat{\theta})$, по которым, пользуясь (8.2.1), мы можем вычислить не a , а \hat{a} . Автокорреляции $r_k(\hat{a})$ ряда \hat{a} могут дать ценную информацию о недостаточно хорошей подгонке и возможной природе неадекватности модели. Однако, как указывал Дарбин [75], может оказаться рискованным придавать статистическую значимость кажущимся отклонениям этих автокорреляций $r_k(\hat{a})$ от их теоретических нулевых значений исходя из стандартной ошибки $n^{-1/2}$, соответствующей $r_k(a)$. Дарбин показал, например, что для процесса $AR(1)$ с параметром ϕ дисперсия $r_1(\hat{a})$ равна $\phi^2 n^{-1}$, что может оказаться существенно *меньше*, чем n^{-1} . Дисперсии и ковариации всех автокорреляций \hat{a} для больших выборок любого

процесса АРСС были выведены Боксом и Пирсом [77]. Эти авторы показали, что хотя во всех случаях для малых задержек возможно уменьшение дисперсии и сильная корреляция $r_k(\hat{a})$, при больших задержках эти эффекты быстро исчезают. Поэтому использование $n^{-1/2}$ в качестве стандартной ошибки для $r_k(\hat{a})$ будет приводить к недооценке статистической значимости

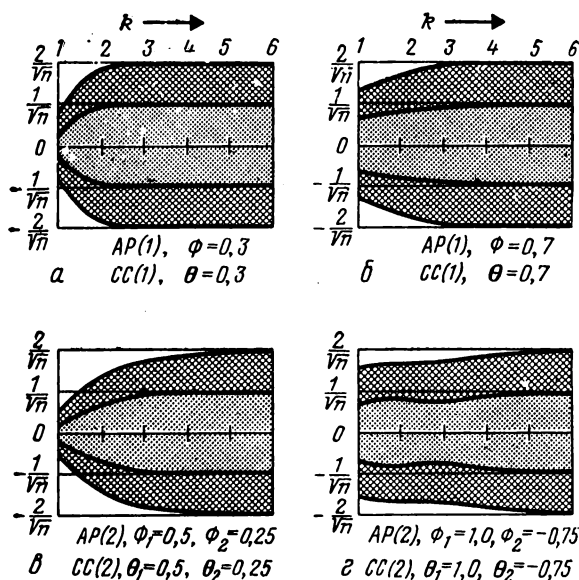


Рис. 8.2. Пределы стандартных ошибок для остаточных автокорреляций $r_k(\hat{a})$.

кажущегося отклонения от нуля для автокорреляций при малых задержках, но обычно вполне оправдано для средних и больших задержек.

В качестве примера на рис. 8.2 показаны пределы в одну и две стандартные ошибки $r_k(\hat{a})$ для большой выборки в случае двух процессов авторегрессии первого порядка и двух процессов авторегрессии второго порядка. Эти графики пригодны также для оценки стандартных ошибок процессов скользящего среднего с теми же параметрами (как указано на рисунке).

Можно сделать вывод, что при не слишком больших задержках можно рассматривать $n^{-1/2}$ как верхнюю границу стандартных ошибок $r_k(\hat{a})$, а не как сами стандартные ошибки. Если мы пользуемся $n^{-1/2}$ как стандартной ошибкой $r_k(\hat{a})$ при малых задержках, то можем серьезно недооценить значимость кажущихся расхождений.

8.2.2. Совокупный критерий согласия

Во многих случаях предпочтительнее не рассматривать отдельные $r_k(\hat{a})$, а оценить неадекватность модели, например, по первым 20 автокорреляциям, рассматриваемым как *единое целое*. Положим, что мы располагаем первыми K автокорреляциями *) любого процесса АРСС(p, d, q). Можно показать [77], что если подгоняемая модель удовлетворительна, то

$$Q = n \sum_{k=1}^K r_k^2(\hat{a}) \quad (8.2.2)$$

распределено приближенно как $\chi^2(K - p - q)$, где $n = N - d$ — число значений w , используемых при подгонке модели. С другой стороны, если модель не соответствует данным, среднее значение Q поднимется. Следовательно, общий, или «совокупный», критерий проверки гипотезы об адекватности модели, предложенный с учетом описанных выше трудностей, можно провести, сопоставив наблюдаемое значение Q с таблицей χ^2 -распределения (такой, как таблица F в конце книги).

Модель (0, 2, 2), подогнанная к ряду С. Для иллюстрации применения совокупного критерия (8.2.2) в табл. 8.1 приведены первые 25 автокорреляций $r_k(\hat{a})$ остаточных ошибок процесса ППС(0, 2, 2) $\nabla^2 z_t = (1 - 0,13B - 0,12B^2)a_t$ — одной из моделей, подгонявшихся в гл. 7 к ряду С.

Таблица 8.1

Автокорреляции $r_k(\hat{a})$ остаточных ошибок модели $\nabla^2 z_t = (1 - 0,13B - 0,12B^2)a_t$, подогнанной к ряду С

k	$r_k(\hat{a})$	k	$r_k(\hat{a})$	k	$r_k(\hat{a})$	k	$r_k(\hat{a})$	k	$r_k(\hat{a})$
1	0,020	6	-0,033	11	-0,129	16	-0,050	21	0,007
2	0,032	7	0,022	12	0,063	17	0,153	22	0,132
3	-0,125	8	-0,056	13	-0,084	18	-0,092	23	0,012
4	-0,078	9	-0,130	14	0,022	19	-0,005	24	-0,012
5	-0,011	10	0,093	15	-0,006	20	-0,015	25	-0,127

Так как число наблюдений $n = 224$, приблизительная верхняя граница стандартной ошибки автокорреляции равна

*) Здесь предполагается, что K достаточно велико, так что веса ψ_j в модели, представленной в виде

$$\hat{\psi}_t = \phi^{-1}(B) \theta(B) a_t = \psi(B) a_t,$$

будут пренебрежимо малы для $j > K$.

$1/\sqrt{224} \approx 0,07$. По сравнению с этой границей стандартной ошибки все величины $r_3(\hat{a}) = -0,125$, $r_9(\hat{a}) = -0,130$, $r_{11}(\hat{a}) = -0,129$, $r_{17}(\hat{a}) = 0,153$, $r_{22}(\hat{a}) = 0,132$, $r_{25}(\hat{a}) = -0,127$ слишком велики. Конечно, редкие большие отклонения могут встречаться даже в случайном ряде, но при рассмотрении этих результатов в целом должно появиться подозрение о недостаточной подгонке модели.

Чтобы сделать более строгое заключение, сравним величину

$$Q = 224[(0,020)^2 + (0,032)^2 + \dots + (-0,127)^2] = 33,7$$

с таблицей χ^2 с 23 степенями свободы. 90- и 95 %-ные квантили для χ^2 с 23 степенями свободы равны соответственно 32,0 и 35,2. Следовательно, возникает некоторое сомнение в адекватности этой модели.

Модель (1, 1, 0), подгоняемая к ряду С. Первые 25 автокорреляций остаточных ошибок для модели $(1 - 0,82B) \nabla z_t = a_t$, которая в гл. 7 была признана наиболее подходящей для описания ряда С, приведены в табл. 8.2. Для этой модели $Q = 225 \sum_{k=1}^{25} r_k^2(\hat{a}) = 28,9$. Сравнение с таблицей χ^2 для 24 степеней свободы показывает, что нет оснований для сомнения в адекватности этой модели.

Таблица 8.2

Автокорреляции $r_k(\hat{a})$ остаточных ошибок для модели
 $(1 - 0,82B) \nabla z_t = a_t$, подгоняемой к ряду С

k	$r_k(\hat{a})$	k	$r_k(\hat{a})$	k	$r_k(\hat{a})$	k	$r_k(\hat{a})$	k	$r_k(\hat{a})$
1	-0,007	6	0,019	11	-0,098	16	-0,039	21	0,001
2	-0,002	7	0,073	12	0,074	17	0,165	22	0,129
3	-0,061	8	-0,030	13	-0,054	18	-0,033	23	0,014
4	-0,014	9	-0,097	14	0,034	19	-0,004	24	-0,017
5	0,047	10	0,133	15	0,002	20	-0,009	25	-0,129

В табл. 8.3 подытожены значения статистики Q , основанной на 25 остаточных автокорреляциях для моделей из табл. 7.13, подгоняемых к рядам А—F.

Рассмотрение табл. 8.3 показывает, что имеются только два подозрительно больших значения Q . Одно из них — это значение 33,7, полученное после подгонки модели $(0, 2, 2)$ к ряду С, которое мы уже обсуждали. Другое значение $Q = 37,1$ получено после подгонки модели $(0, 1, 1)$ к ряду В. Оно свидетельствует о возможной неадекватности модели, так как 95- и

Таблица 8.3

Сводка результатов совокупного критерия, примененного к остаточным ошибкам различных моделей, подгоняемых к рядам А — F

Ряд	$n = N - d$	Подгоняемая модель	Q	Число степеней свободы
A	197	$z_t - 0,92z_{t-1} = 1,45 + a_t - 0,58a_{t-1}$	26,5	23
	196	$\nabla z_t = a_t - 0,70a_{t-1}$	29,9	24
B	368	$\nabla z_t = a_t + 0,09a_{t-1}$	37,1	24
C	225	$\nabla z_t - 0,82\nabla z_{t-1} = a_t$	28,9	24
	224	$\nabla^2 z_t = a_t - 0,13a_{t-1} - 0,12a_{t-2}$	33,7	23
D	310	$z_t - 0,87z_{t-1} = 1,17 + a_t$	10,8	24
	309	$\nabla z_t = a_t - 0,06a_{t-1}$	18,0	24
E	100	$z_t - 1,42z_{t-1} + 0,73z_{t-2} = 14,35 + a_t$	23,7	23
	100	$z_t - 1,57z_{t-1} + 1,02z_{t-2} - 0,21z_{t-3} = 11,31 + a_t$	23,7	22
F	70	$z_t + 0,34z_{t-1} - 0,19z_{t-2} = 58,87 + a_t$	11,3	23

97,5%-ные квантили для χ^2 с 24 степенями свободы равны соответственно 36,4 и 39,3. Возможную природу этой неадекватности мы рассмотрим в следующем разделе.

8.2.3. Неадекватность модели, возникающая из-за изменения значений параметров

Одна из возможных интересных форм неадекватности модели может быть связана с тем, что в течение продолжительного интервала времени форма модели остается той же, а параметры изменяются. Имеются основания считать, что такие явления могут объяснить неадекватность модели (0, 1, 1), подгоняемой к ряду данных IBM.

Таблица 8.4

Сравнение моделей ПСС (0, 1, 1), подогнанных к первой и второй половинам ряда B

	n	$\hat{\theta}$	$\hat{\lambda} = 1 - \hat{\theta}$	$\hat{\sigma}(\hat{\lambda}) = \left\{ \frac{\hat{\lambda}(1-\hat{\lambda})}{n} \right\}^{1/2}$	Q	Число степеней свободы
Первая половина	184	-0,29	1,29	$\pm 0,070$	22,6	24
Вторая половина	183	-0,03	1,03	$\pm 0,074$	33,9	24
Полный ряд	368	-0,09	1,09	$\pm 0,052$	37,1	24

Табл. 8.4 содержит результаты подгонки процесса $(0, 1, 1)$ отдельно к первой и второй половинам ряда B , а также ко всему ряду.

Обозначив оценки, полученные для двух половин $\hat{\lambda}^{(1)}$ и $\hat{\lambda}^{(2)}$, находим, что стандартная ошибка $\hat{\lambda}^{(1)} - \hat{\lambda}^{(2)}$ равна

$$\sqrt{(0,070)^2 + (0,074)^2} = 0,102.$$

Поскольку разность $\hat{\lambda}^{(1)} - \hat{\lambda}^{(2)} = 0,26$ в 2,6 раза больше ее стандартной ошибки, весьма вероятно, что действительно произошло изменение λ . Рассмотрение величин Q дает основание предположить, что модель $(0, 1, 1)$ с параметрами, изменяемыми для различных интервалов времени, может более точно объяснить наблюдения.

8.2.4. Проверка при помощи кумулятивной периодограммы

В некоторых ситуациях, особенно при подгонке сезонных временных рядов, о которых мы будем говорить в гл. 9, можно опасаться, что мы неадекватно учли *периодический* характер ряда. Следовательно, мы должны исследовать периодичность остаточных ошибок. Автокорреляционная функция не является чувствительным индикатором таких отклонений от случайности, поскольку эффект периодичности будет, как правило, распыляться на несколько значений автокорреляций. С другой стороны, периодограммы разработаны специально для обнаружения периодичности на фоне белого шума.

Периодограмма временного ряда a_t , $t = 1, 2, \dots, n$, согласно разд. 2.2.1, — это

$$I(f_i) = \frac{2}{n} \left[\left(\sum_{t=1}^n a_t \cos 2\pi f_i t \right)^2 + \left(\sum_{t=1}^n a_t \sin 2\pi f_i t \right)^2 \right], \quad (8.2.3)$$

где $f_i = i/n$ — частота. Таким образом, это результат корреляции a_t с синусоидами и косинусоидами различных частот. Периодичность с данной частотой f_i , имеющаяся в остаточных ошибках, подчеркивается при корреляции с синусоидой или косинусоидой той же частоты и дает поэтому большое значение $I(f_i)$.

Кумулятивная периодограмма. Бартлетт [78] (см. также [27]) показал, что *кумулятивная периодограмма* является эффективным средством обнаружения периодического отклонения от случайности.

Спектр мощности $p(f)$ белого шума имеет постоянное значение $2\sigma_a^2$ в частотном диапазоне 0—0,5 Гц. Следовательно,

график кумулятивного спектра белого шума как функции f

$$P(f) = \int_0^f p(g) dg \quad (8.2.4)$$

имеет вид прямой линии, идущей от точки $(0; 0)$ к точке $(0,5; \sigma_a^2)$, и, следовательно, график $P(f)/\sigma_a^2$ — прямая линия, идущая от $(0; 0)$ к $(0,5; 1)$.

Как упоминалось в разд. 2.2.3, $I(f)$ дает оценку спектра на частоте f . Действительно, для белого шума $E[I(f)] = 2\sigma_a^2$, и $I(f)$ — несмещенная оценка. Отсюда следует, что $\frac{1}{n} \sum_{i=1}^l I(f_i)$ — несмещенная оценка проинтегрированного спектра $P(f_j)$, и

$$C(f_j) = \frac{\sum_{i=1}^l I(f_i)}{ns^2} \quad (8.2.5)$$

— оценка $P(f_j)/\sigma_a^2$, где s^2 — оценка σ_a^2 . Мы будем называть $C(f_i)$ *нормированной кумулятивной периодограммой*.

Если бы модель была адекватна, а параметры ее *точно* известны, a , вычисляемые из данных, обладали бы свойствами белого шума. Для белого шума график $C(f_i)$ имел бы разброс относительно прямой, соединяющей точки $(0; 0)$ и $(0,5; 1)$. С другой стороны, из-за неадекватности модели ряд a становится неслучайным, и его кумулятивная периодограмма должна систематически отклоняться от этой линии. В частности, периодичности a приводили бы к появлению больших соседних значений $I(f_j)$. Эти большие ординаты усиливали бы друг друга в $C(f_j)$ и образовывали бы заметное отклонение от ожидаемой прямой линии.

На практике мы не знаем точно значений параметров и предполагаем только их выборочными оценками. Мы не имеем значений a , а только выборочные остаточные ошибки \hat{a} . Однако для больших выборок периодограмма ряда \hat{a} имеет свойства, сходные со свойствами периодограммы белого шума a . Поэтому внимательное изучение периодограммы ряда \hat{a} может служить полезной дополнительной диагностической проверкой, особенно для выявления неучтенных периодичностей.

Пример: ряд С. Мы видим, что ряд C хорошо описывается моделью $(1, 1, 0)$

$$(1 - 0,82B) \nabla z_t = a_t$$

и несколько хуже — моделью ПСС $(0, 2, 2)$

$$\nabla^2 z_t = (1 - 0,13B - 0,12B^2) a_t,$$

почти эквивалентной первой. Мы проиллюстрируем критерий кумулятивной периодограммы на примере анализа остаточных ошибок \hat{a} после подгонки к ряду неадекватной модели ПСС(0, 1, 1)

$$Vz_t = (1 - \theta B)a_t.$$

Для этой модели оценка наименьших квадратов для θ равна $-0,65$. График нормированной кумулятивной периодограммы остаточных ошибок этой модели показан на рис. 8.3, а. Рассмотрение этого рисунка показывает, что имеются существенные отклонения кумулятивной периодограммы от линейности. Эти отклонения очень заметны на низких частотах и похожи на то, что должно было бы быть, например, при недостаточном взятии разностей. На рис. 8.3, б показан соответствующий график для наилучшей модели ПСС(0, 2, 2). Точки кумулятивной периодограммы теперь весьма близки к ожидаемой прямой, хотя, как мы видели в табл. 8.3, имеются свидетельства неадекватности и этой модели.

Разумно иметь на графике как шкалу периодов, так и шкалу частот. Это облегчает идентификацию отходов от прямой в случаях, когда остаточные ошибки имеют периодичность. Например, в месячных ценах отклонения вблизи периодов 12, 24, 36 и т. д. месяцев могут свидетельствовать о неадекватном учете сезонных явлений.

Вероятностное соотношение между кумулятивной периодограммой и проинтегрированным спектром — точно такое же, как между эмпирической и теоретической функциями распределения. По этой причине мы можем оценить отклонения периодограммы от ожидаемой для белого шума при помощи критерия Колмогорова [79]. Пользуясь этим критерием, мы можем провести около теоретической прямой предельные линии. Они имеют следующий смысл: если ряд \hat{a}_t — белый шум, кумулятивная периодограмма будет отклоняться от прямой настолько, чтобы пересечь эти предельные линии только с указанной вероятностью. Однако, поскольку \hat{a} — это подогнанные значения, а не точные a , они будут не точно следовать процессу белого шума даже для правильной модели. Поэтому использование в качестве теста неадекватности модели пределов Колмогорова позволяет только приближенно оценивать вероятности. Однако полезно указывать эти пределы на кумулятивной периодограмме для грубых оценок, какие отклонения считать сомнительными, а какие исследовать подробнее.

Эти предельные линии обладают тем свойством, что для чисто случайных рядов кумулятивная периодограмма выходила бы за них в ϵ -й части всех случаев. Они проводятся на расстояниях $\pm K_\epsilon / \sqrt{q}$ выше и ниже теоретической прямой; здесь

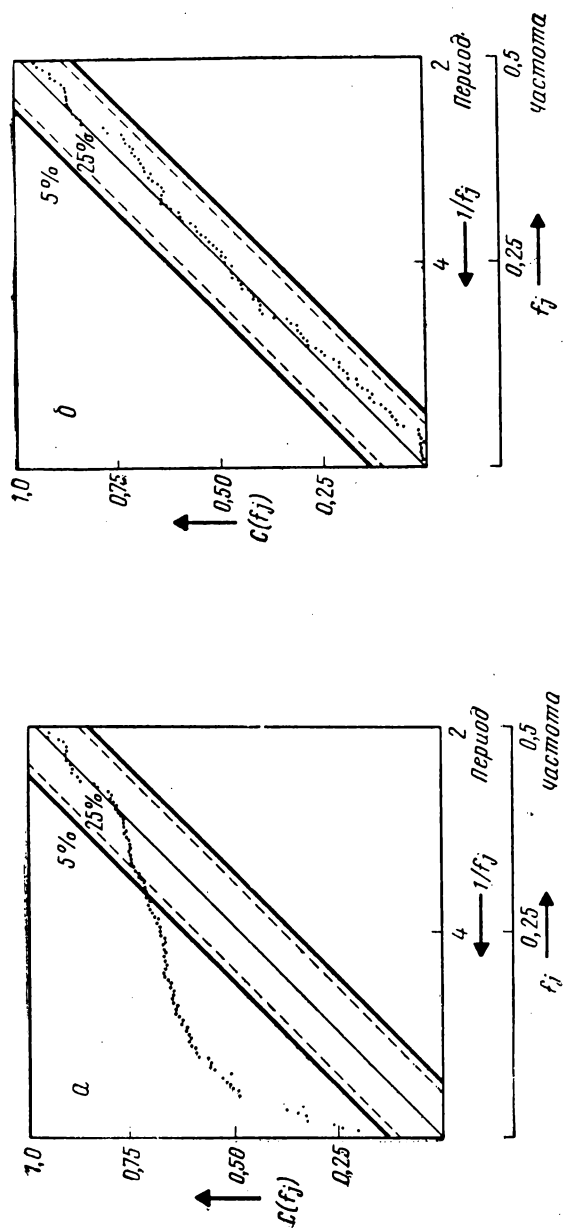


Рис. 8.3. Ряд C — кумулятивные периодограммы остаточных ошибок для наилучшей подогнанной модели.
 a — порядка (0,1,1), $б$ — порядка (0,2,2).

$q = (n - 2)/2$ для четных n и $(n - 1)/2$ для нечетных. Приближенные значения K_ε приведены в табл. 8.5.

Таблица 8.5

Коэффициенты для вычисления приближенных вероятностных пределов в критерии кумулятивной периодограммы

ε	0,01	0,05	0,10	0,25
K_ε	1,63	1,36	1,22	1,02

Для ряда C $q = (224 - 2)/2 = 111$, и 5%-ные предельные линии, проведенные на рис. 8.3, отстоят от теоретической прямой по ординате на $\pm 1,36/\sqrt{111} = \pm 0,13$. 25%-ные предельные линии отстоят от теоретической на $\pm 1,02/\sqrt{111} = \pm 0,10$.

Выводы. Каждая из описанных выше проверочных процедур имеет существенные преимущества и недостатки. Проверки, основанные на изучении выборочных автокорреляционных функций и кумулятивных периодограмм, могут указывать на особенности ряда, *существование которых не подозревалось*; однако они не слишком чувствительны. Тесты на конкретный вид отклонений, проводимые путем введения в модель избыточных параметров, весьма чувствительны, но могут не выявить те несоответствия, на которые они не рассчитаны.

8.3. ИСПОЛЬЗОВАНИЕ ОСТАТОЧНЫХ ОШИБОК ДЛЯ ИЗМЕНЕНИЯ МОДЕЛИ

8.3.1. Природа корреляций остаточных ошибок при использовании неверной модели

Если автокорреляционная функция остаточных ошибок относительно некоторой подогнанной модели указывает на неадекватность модели, необходимо рассмотреть, каким образом следует изменить модель. В разд. 8.3.2 мы покажем, как использовать автокорреляции остаточных ошибок для определения способа изменения. Рассмотрим вкратце эффект подгонок неверной модели при помощи автокорреляционной функции остаточных ошибок.

Пусть верная модель имеет вид

$$\phi(B) \tilde{w}_t = \theta(B) a_t,$$

а используется неверная модель

$$\phi_0(B) \tilde{w}_t = \theta_0(B) b_t.$$

Тогда остаточные ошибки b_t относительно неверной модели будут коррелированы, и, так как

$$b_t = \theta_0^{-1}(B) \theta(B) \phi_0(B) \phi^{-1}(B) a_t, \quad (8.3.1)$$

производящая функция автоковариаций b будет равна

$$\sigma_a^2 \{ \theta_0^{-1}(B) \theta_0^{-1}(F) \theta(B) \theta(F) \phi_0(B) \phi_0(F) \phi^{-1}(B) \phi^{-1}(F) \}. \quad (8.3.2)$$

Рассмотрим в качестве примера процесс ППС(0, 1, 1), в котором вместо верного значения θ мы пользуемся неким другим значением θ_0 . Тогда остаточные ошибки b_t будут следовать смешанному процессу (1, 0, 1)

$$(1 - \theta_0 B) b_t = (1 - \theta B) a_t,$$

и, пользуясь (3.4.8), получаем

$$\rho_1 = \frac{(1 - \theta\theta_0)(\theta_0 - \theta)}{1 + \theta^2 - 2\theta\theta_0},$$

$$\rho_j = \rho_1 \theta_0^{j-1}, \quad j = 2, 3, \dots$$

Предположим, что в процессе ПСС(0, 1, 1)

$$\nabla z_t = (1 - \theta B) a_t$$

мы используем значение $\theta_0 = 0,8$ вместо истинного значения $\theta = 0$. Тогда

$$\begin{aligned} \theta_0 &= 0,8, & \theta &= 0,0, \\ \rho_1 &= 0,8, & \rho_j &= 0,8^j. \end{aligned}$$

Отсюда b сильно коррелированы и подчиняются процессу авто-регрессии

$$(1 - 0,8B) b_t = a_t.$$

8.3.2. Использование остаточных ошибок для изменения модели

Предположим, что остаточные ошибки b_t для модели

$$\phi_0(B) \nabla^{d_0} z_t = \theta_0(B) b_t \quad (8.3.3)$$

выглядят неслучайными. Используя автокорреляционную функцию b_t , можно методами гл. 6 идентифицировать модель

$$\bar{\phi}(B) \nabla^{\bar{d}} b_t = \bar{\theta}(B) a_t \quad (8.3.4)$$

для ряда b_t . Исключая b_t из (8.3.3) и (8.3.4), получаем новую модель

$$\phi_0(B) \bar{\phi}(B) \nabla^{d_0} \nabla^{\bar{d}} z_t = \theta_0(B) \bar{\theta}(B) a_t, \quad (8.3.5)$$

которая может быть подогнана и диагностически проверена.

Например, предположим, что ряд был неверно идентифицирован как процесс ПСС(0, 1, 1) и после подгонки привел к модели

$$\nabla z_t = (1 + 0,6B) b_t. \quad (8.3.6)$$

Пусть также по ряду остаточных ошибок этой модели была идентифицирована модель

$$\nabla b_t = (1 - 0,8B) a_t. \quad (8.3.7)$$

Тогда, исключив b_t из (8.3.6) и (8.3.7), получим новую модель

$$\nabla^2 z_t = (1 - 0,2B - 0,48B^2) a_t,$$

что указывает на необходимость опробования в качестве модели процесса ПСС(0, 2, 2).

Модели сезонных рядов

В гл. 4—8 мы рассмотрели свойства класса линейных стохастических моделей, полезных для описания стационарных и нестационарных временных рядов, и показали, как эти модели можно использовать для прогнозирования. Мы рассмотрели далее практические проблемы идентификации, подгонки и диагностической проверки, возникающие при сопоставлении этих моделей с фактическими данными. В этой главе рассмотрено применение описанных методов к прогнозированию сезонных временных рядов; в них продемонстрировано также совместное использование принципов, изложенных в предыдущих главах.

9.1. ЭКОНОМИЧЕСКИЕ МОДЕЛИ СЕЗОННЫХ ВРЕМЕННЫХ РЯДОВ

На рис. 9.1 показан график суммарного количества пассажиров международных авиалиний США за 1952, 1953 и 1954 гг. Это часть более длинного ряда (данные за 12 лет), описанного Брауном [2] и приведенного в сборнике временных рядов в конце этой книги как ряд G. Ряд обладает ярко выраженной сезонной периодичностью, поскольку наибольшее количество путешествий приходится на поздние месяцы лета; отмечается также весенний пик. Подобные сезонные свойства имеют многие другие ряды, например данные о сбыте.

В общем мы говорим, что ряд имеет периодичность с периодом s , когда сходные особенности ряда повторяются после s опорных временных интервалов. В приведенном выше примере опорный временной интервал равен 1 мес., а период $s = 12$ мес. Известны, однако, примеры, когда s принимает другие значения. Так, для квартальных данных $s = 4$ означает сезонную периодичность в пределах года *). Иногда имеется несколько периодичностей. Так, поскольку имеется тенденция оплачивать счета ежемесячно, естественно ожидать, что еженедельные банковские сводки имеют месячную периодичность $s \approx 4$, в то время как месячные сводки имеют годовую периодичность $s = 12$.

*) Здесь опорный временной интервал равен одному кварталу. — *Прим. ред.*

9.1.1. Сопоставление подгонки и прогнозирования

Одной из причин неудач в анализе временных рядов в прошлом было смешение процедур *подгонки* и *прогнозирования* ряда. Пусть, например, ряд имеет тенденцию к увеличению значений в течение некоторого периода и, кроме того, проявляет сезонные особенности. Общепринятый метод анализа — произвольно разделить ряд на три компоненты: «тренд», «сезонную компоненту» и «случайную компоненту». Тренд можно подогнать полиномом, а сезонную компоненту — рядом Фурье. Прогноз производится путем экстраполяции этих подогнанных функций в будущее.

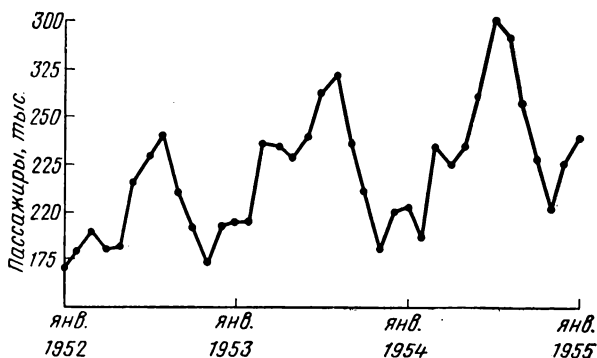


Рис. 9.1. Суммарное число пассажиров международных авиалиний в тысячах (часть ряда G).

Такие методы могут приводить к серьезным заблуждениям. Например, мы видели ранее, что поведение биржевых цен акций IBM (ряд B) хорошо аппроксимируется моделью случайного блуждания

$$z_t = \sum_{j=-\infty}^t a_j, \quad (9.1.1)$$

из которой вытекает, что $\hat{z}_t(l) = z_t$. Другими словами, лучший прогноз будущих значений стоимости акций — это сегодняшняя цена. Правда, короткие участки ряда B выглядят так, что их можно аппроксимировать кривыми второго порядка. Это просто отражает тот факт, что сумма случайных отклонений может иногда иметь подобный вид. Однако нет основания для использования квадратичной прогнозирующей функции, поскольку она дает очень плохие прогнозы. Конечно, истинные систематические эффекты, имеющие естественное объяснение, следует учитывать включением в модель соответствующей детерминированной компоненты. Например, если известно, что в систему регулярно

поступает тепло, было бы разумным объяснить результирующее увеличение температуры при помощи подходящей детерминированной функции времени, добавленной к стохастической компоненте.

9.1.2. Сезонные модели, включающие подстраивающиеся синусоиды и косинусоиды

Общая линейная модель

$$\tilde{z}_t = \sum_{j=1}^{\infty} \pi_j \tilde{z}_{t-j} + a_t = \sum_{j=1}^{\infty} \psi_j a_{t-j} + a_t \quad (9.1.2)$$

с подходящими значениями коэффициентов π_j и ψ_j вполне адекватна для описания многих временных рядов. Проблема заключается в выборе *экономичной параметризации* для такой модели. Как уже отмечалось выше, это не математическая проблема, а вопрос о том, как узнать тенденции поведения всевозможных явлений. Решать эту проблему можно, опробуя различные идеи на наблюдаемых временных рядах и развивая концепции, дающие полезные результаты.

Мы видели, что для несезонных рядов обычно удается получить полезное и экономичное представление вида

$$\varphi(B) \tilde{z}_t = \theta(B) a_t. \quad (9.1.3)$$

Более того, обобщенный оператор авторегрессии $\varphi(B)$ определяет эвентуальную прогнозирующую функцию, являющуюся решением разностного уравнения

$$\varphi(B) \hat{z}_t(l) = 0,$$

где подразумевается, что B действует на l . При описании сезонных явлений прогнозирующая функция должна обладать способностью к предсказанию периодичности. На первый взгляд кажется, что $\varphi(B)$ должно дать прогнозирующую функцию в виде суммы синусоид и косинусоид, возможно, с добавлением полиномиальных членов; такая функция позволит предсказать изменения уровня и периодические (сезонные) изменения. Такая прогнозирующая функция может быть естественным образом получена в рамках общей модели (9.1.3). Например, для ежемесячных данных прогнозирующая функция в виде синусоиды с 12-месячным периодом, подстраивающейся фазой и амплитудой удовлетворяет разностному уравнению

$$(1 - \sqrt{3}B + B^2) \hat{z}_t(l) = 0,$$

где B действует на l . Однако неверно, что *экономичный* способ описания периодического поведения — это представление его

суммой синусоид и косинусоид. Например, потребуется много гармонических компонент, чтобы описать данные о сбыте, на которые влияют рождественские и пасхальные праздники и другие сезонные распродажи. В качестве экзотического примера можно указать, что распродажа изделий для фейерверка в Великобритании происходит в основном в недели, предшествующие празднованию раскрытия «порохового заговора» Гая Фокса 5 ноября. Попытка описать отдельный «выброс» распродажи изделий для фейерверка непосредственно при помощи синусоид и косинусоид была бы крайне неэкономичной. Ясно, что требуется рассмотреть эту проблему более внимательно.

В нашем предыдущем анализе мы не всегда оценивали *все* компоненты $\varphi(B)$. Когда для получения стационарности нужно было брать разности d раз, мы писали $\varphi(B) = \phi(B)(1-B)^d$, что эквивалентно приравниванию d корней уравнения $\varphi(B) = 0$ единице. Если такое представление оказывалось адекватным, мы могли продолжать более простой анализ ряда $w_t = \nabla^d z_t$. Таким образом, мы использовали $\nabla = 1 - B$ как упрощающий оператор. В других задачах могут оказаться полезными другие типы упрощающих операторов. Например, потребление нефти как топлива зависит от внешней температуры, которая из-за вращения Земли вокруг Солнца меняется приблизительно как синусоида с периодом 12 мес. При анализе сбыта нефти может оказаться целесообразным введение упрощающего оператора $1 - \sqrt{3}B + B^2$ в качестве одной из компонент обобщенного оператора авторегрессии $\varphi(B)$. Если такое представление оказалось полезным, мы можем продолжать дальнейший, более простой анализ ряда $w_t = (1 - \sqrt{3}B + B^2)z_t$. Этот оператор, как можно заметить, принадлежит к классу однородных нестационарных операторов и имеет нули $e^{\pm i(2\pi/12)}$ на единичной окружности.

9.1.3. Общая мультипликативная модель сезонного ряда

Упрощающий оператор $1 - B^s$. Фундаментальным фактом, относящимся к сезонным временным рядам, является сходство наблюдений, разделенных интервалом s . Следовательно, можно ожидать, что операция $B^s z_t = z_{t-s}$ будет играть особо важную роль в анализе сезонных рядов; далее, так как в ряде $z_t, z_{t-s}, z_{t-2s}, \dots$ можно ожидать нестационарности, может оказаться полезным упрощающий оператор $\nabla_s z_t = (1 - B^s)z_t = z_t - z_{t-s}$. Устойчивый нестационарный оператор $(1 - B^s)$ имеет s нулей $e^{i(2\pi k/s)}$ ($k = 0, 1, \dots, s-1$), равномерно распределенных на единичной окружности. Далее, эвентуальная прогнозирующая функция удовлетворяет уравнению $(1 - B^s)\hat{z}_t(l) = 0$ и поэтому

может (но не обязана) быть представлена в виде набора синусоид и косинусоид

$$\hat{z}_t(l) = b_0^{(t)} + \sum_{j=1}^{[s/2]} [b_{1j}^{(t)} \cos(2\pi jl/s) + b_{2j}^{(t)} \sin(2\pi jl/s)],$$

где b — подстраиваемые коэффициенты и $[s/2] = \frac{1}{2}s$ для четных s , $[s/2] = \frac{1}{2}(s-1)$ для нечетных s .

Мультипликативная модель. Когда мы имеем дело с рядом, проявляющим сезонные особенности с известным периодом s , полезно представить данные в виде таблицы, состоящей из s столбцов, как, например, табл. 9.1, содержащая логарифмы данных об авиаперевозках. (Перед анализом данных о распродажах и других временных рядах этого типа часто переходят к логарифмам, поскольку сопоставимыми при разных объемах распродажи могут быть процентные флуктуации.)

Таблица 9.1

**Натуральные логарифмы месячных количеств (в тыс.)
пассажирских перевозок на международных авиалиниях (ряд G)**

Годы	Янв.	Февр.	Март	Апр.	Май	Июнь	Июль	Авг.	Сент.	Окт.	Нояб.	Дек.
1949	4,718	4,771	4,883	4,860	4,796	4,905	4,997	4,997	4,913	4,779	4,644	4,771
1950	4,745	4,836	4,949	4,905	4,828	5,004	5,136	5,136	5,063	4,890	4,736	4,942
1951	4,977	5,011	5,182	5,094	5,147	5,182	5,293	5,293	5,215	5,088	4,984	5,112
1952	5,142	5,193	5,263	5,199	5,209	5,334	5,438	5,489	5,342	5,252	5,147	5,268
1953	5,278	5,278	5,464	5,460	5,434	5,493	5,576	5,606	5,468	5,352	5,193	5,303
1954	5,318	5,236	5,460	5,425	5,455	5,576	5,710	5,680	5,557	5,434	5,313	5,434
1955	5,489	5,451	5,587	5,595	5,598	5,753	5,897	5,849	5,743	5,613	5,468	5,628
1956	5,649	5,624	5,759	5,746	5,762	5,924	6,023	6,004	5,872	5,724	5,602	5,724
1957	5,753	5,707	5,875	5,852	5,872	6,045	6,146	6,146	6,001	5,849	5,720	5,817
1958	5,829	5,762	5,892	5,852	5,894	6,075	6,196	6,225	6,001	5,883	5,737	5,820
1959	5,886	5,835	6,006	5,981	6,040	6,157	6,306	6,326	6,138	6,009	5,892	6,004
1960	6,033	5,969	6,038	6,133	6,157	6,282	6,433	6,407	6,230	6,133	5,966	6,068

Структура табл. 9.1 подчеркивает, что в периодических данных важен не один, а два временных интервала. В этом примере эти интервалы соответствуют месяцу и году. Конкретнее, мы ожидаем, что существуют определенные соотношения между а) наблюдениями за последовательные месяцы данного года, б) наблюдениями того же месяца в последовательные годы. Ситуация несколько сходна с встречаемой при изучении таблиц с двумя входами в дисперсионном анализе, где можно ожидать

сходства между наблюдениями в одной и той же строке или в одном и том же столбце.

Если рассмотреть с этой точки зрения табл. 9.1, то сезонный эффект должен проявиться в ней так: наблюдения за какой-либо месяц (скажем, апрель) некоторого года должны быть связаны с наблюдениями за тот же месяц предыдущего года. Пусть t -е наблюдение z_t относится к апрелю. Тогда мы можем связать это наблюдение с наблюдением в предыдущем апреле моделью вида

$$\Phi(B^s) \nabla_s^D z_t = \Theta(B^s) \alpha_t, \quad (9.1.4)$$

где $s = 12$, $\nabla_s = 1 - B^s$ и $\Phi(B^s)$, $\Theta(B^s)$ — полиномы B^s степеней P и Q соответственно, удовлетворяющие условиям стационарности и обратимости. Подобным образом модель

$$\Phi(B^s) \nabla_s^D z_{t-1} = \Theta(B^s) \alpha_{t-1} \quad (9.1.5)$$

может быть использована для связи наблюдений за март этого и предшествующего года и т. д. для любого из 12 мес. Кроме того, обычно оказывается разумным предположение, что параметры Φ и Θ , содержащиеся в этих ежемесячных моделях, примерно одинаковы для всех месяцев.

Ошибки α_t , α_{t-1} , ... этих моделей не обязательно должны быть некоррелированы. Например, количество авиапассажиров в апреле 1960 г., будучи связанным с количеством пассажиров в апреле предыдущего года, связано также с количеством пассажиров в марте, феврале, январе 1960 г. и т. д. Поэтому можно ожидать, что α_t в (9.1.4) связано с α_{t-1} в (9.1.5), с α_{t-2} и т. д. Следовательно, чтобы учесть эти связи, мы вводим вторую модель

$$\phi(B) \nabla^d \alpha_t = \theta(B) a_t, \quad (9.1.6)$$

где a_t — белый шум, а $\phi(B)$ и $\theta(B)$ — полиномы B степеней p и q соответственно, удовлетворяющие условиям стационарности и обратимости; $\nabla = \nabla_1 = 1 - B$.

Подставляя (9.1.6) в (9.1.4), получаем окончательную общую мультипликативную модель

$$\phi(B) \Phi_p(B^s) \nabla^d \nabla_s^D z_t = \theta_q(B) \Theta_Q(B^s) a_t, \quad (9.1.7)$$

где в этом частном примере $s = 12$. Индексы p , P , q , Q в (9.1.7) были введены, чтобы напомнить читателю о порядках различных операторов. Говорят, что результирующий мультипликативный процесс имеет порядок $(p, d, q) \times (P, D, Q)_s$. Аналогичные рассуждения можно использовать для получения моделей с тремя и более периодическими компонентами, учитывающими многообразие сезонных явлений.

9.2. ПРЕДСТАВЛЕНИЕ ДАННЫХ ОБ АВИАПЕРЕВОЗКАХ МУЛЬТИПЛИКАТИВНОЙ МОДЕЛЮ $(0, 1, 1) \times (0, 1, 1)_{12}$

В дальнейшем в этой главе мы рассмотрим основные виды моделей сезонных рядов введенного нами типа и их потенциальные возможности при прогнозировании. Мы рассмотрим также проблемы идентификации, оценивания и диагностической проверки, возникающие при привязке таких моделей к данным. Для этого мы не нуждаемся в каких-либо новых принципах, а будем просто применять методики и идеи, которые подробно обсуждались в гл. 6—8.

В этом разделе мы детально рассмотрим частный пример. В разд. 9.3 мы рассмотрим аспекты общей модели сезонных рядов, требующие специального внимания.

Детальный пример, рассматриваемый в этом разделе, заключается в применении модели $(0, 1, 1) \times (0, 1, 1)_{12}$ к данным об авиаперевозках (ряду G). В разд. 9.2.1 будет рассмотрена сама модель, в разд. 9.2.2 — основанное на ней прогнозирование, в разд. 9.2.3 — идентификация, в разд. 9.2.4 — подгонка и, наконец, в разд. 9.2.5 — ее диагностическая проверка.

9.2.1. Мультипликативная модель $(0, 1, 1) \times (0, 1, 1)_{12}$

Мы уже видели, что простая и широко применимая стохастическая модель для нестационарных временных рядов, не содержащая сезонных компонент, — это процесс ПСС $(0, 1, 1)$. Положим, что в соответствии с рассуждениями разд. 9.1.3 мы пришли к модели

$$\nabla_{12} z_t = (1 - \Theta B^{12}) \alpha_t,$$

связывающей z , разделенные интервалом времени в 1 год. Положим далее, что мы пришли к модели

$$\nabla \alpha_t = (1 - \theta B) \alpha_t,$$

связывающей α , отстоящие на один месяц, где θ и Θ в общем случае — разные величины. Комбинируя эти выражения, мы получаем мультипликативную модель сезонного ряда

$$\nabla \nabla_{12} z_t = (1 - \theta B)(1 - \Theta B^{12}) \alpha_t \quad (9.2.1)$$

порядка $(0, 1, 1) \times (0, 1, 1)_{12}$. В явном виде эту модель можно представить так:

$$z_t - z_{t-1} - z_{t-12} + z_{t-13} = \alpha_t - \theta \alpha_{t-1} - \Theta \alpha_{t-12} + \theta \Theta \alpha_{t-13}. \quad (9.2.2)$$

Интервал обратимости для этой модели, получаемый из условия, что корни уравнения $(1 - \theta B)(1 - \Theta B^{12}) = 0$ лежат вне

единичного круга, определен неравенствами

$$-1 < \theta < 1 \quad \text{и} \quad -1 < \Theta < 1.$$

Заметим, что оператор скользящего среднего

$$(1 - \theta B)(1 - \Theta B^{12}) = 1 - \theta B - \Theta B^{12} + \theta \Theta B^{13}$$

в правой части (9.2.1) порядка $q + sQ = 1 + 12(1) = 13$.

В разд. 9.2.3—9.2.5 мы покажем, что прологарифмированные данные об авиаперевозках хорошо описываются моделью вида (9.2.1), где с достаточной точностью $\hat{\theta} = 0,4$, $\hat{\Theta} = 0,6$, $\hat{\sigma}_a^2 = 1,34 \cdot 10^{-3}$. Удобно вначале рассмотреть, как, пользуясь этой моделью с этими значениями параметров, можно предсказать будущие значения ряда.

9.2.2. Прогнозирование

В гл. 4 мы видели, что имеются три разных способа рассмотрения общей модели, каждый из которых привел в гл. 5 к различным взглядам на прогноз. Мы рассмотрим теперь, как эти три подхода можно применить к прогнозированию с помощью модели сезонного ряда (9.2.1).

Способ разностного уравнения. Прогнозы удобнее всего *вычислять* прямо из самого разностного уравнения.

Поскольку

$$z_{t+l} = z_{t+l-1} + z_{t+l-12} - z_{t+l-13} + a_{t+l} - \theta a_{t+l-1} - \Theta a_{t+l-12} + \theta \Theta a_{t+l-13}, \quad (9.2.3)$$

после подстановки $\theta = 0,4$, $\Theta = 0,6$ немедленно получаем прогноз с минимальной среднеквадратичной ошибкой на момент t с упреждением l :

$$\hat{z}_t(l) = [z_{t+l-1} + z_{t+l-12} - z_{t+l-13} + a_{t+l} - 0,4a_{t+l-1} - 0,6a_{t+l-12} + 0,24a_{t+l-13}]. \quad (9.2.4)$$

Как и в гл. 5, мы называем

$$[z_{t+l}] = E[z_{t+l} | \theta, \Theta, z_t, z_{t-1}, \dots]$$

условным математическим ожиданием z_{t+l} в момент t . В этом выражении параметры считаются известными точно, и ряд z_t, z_{t-1}, \dots предполагается известным достаточно далеко в прошлое.

Практическое применение зависит от ряда фактов:

а) обратимые модели после подгонок к реальным рядам дают прогнозы, существенно зависящие только от сравнительно недавних значений ряда;

б) прогнозы нечувствительны к малым изменениям значений параметров, вносимым, например, ошибками оценивания.

Имеем

$$[z_{t+j}] = \begin{cases} z_{t+j} & j \leq 0 \\ \hat{z}_t(j) & j > 0 \end{cases}, \quad (9.2.5)$$

$$[a_{t+j}] = \begin{cases} a_{t+j} & j \leq 0 \\ 0 & j > 0 \end{cases}. \quad (9.2.6)$$

Отсюда, чтобы получать прогнозы, как в гл. 5, мы просто заменяем неизвестные z прогнозами, а неизвестные a — нулями. Известные a — это, конечно, уже вычисленные ошибки прогноза на шаг вперед, т. е. $a_t = z_t - \hat{z}_{t-1}(1)$.

Например, для получения прогноза на три месяца вперед имеем

$$z_{t+3} = z_{t+2} + z_{t-9} - z_{t-10} + a_{t+3} - 0,4a_{t+2} - 0,6a_{t-9} + 0,24a_{t-10}.$$

Беря условные математические ожидания в момент t , получаем

$$\hat{z}_t(3) = \hat{z}_t(2) + z_{t-9} - z_{t-10} - 0,6a_{t-9} + 0,24a_{t-10},$$

т. е.

$$\begin{aligned} \hat{z}_t(3) = \hat{z}_t(2) + z_{t-9} - z_{t-10} - 0,6[z_{t-9} - \hat{z}_{t-10}(1)] + \\ + 0,24[z_{t-10} - \hat{z}_{t-11}(1)]. \end{aligned}$$

Следовательно,

$$\begin{aligned} \hat{z}_t(3) = \hat{z}_t(2) + 0,4z_{t-9} - 0,76z_{t-10} + \\ + 0,6\hat{z}_{t-10}(1) - 0,24\hat{z}_{t-11}(1). \end{aligned} \quad (9.2.7)$$

Здесь прогноз выражается через предыдущие z и предыдущие прогнозы \hat{z} . Хотя нетрудно записать отдельные выражения для любого упреждения, вычисления прогнозов лучше вести, непосредственно используя единственное выражение (9.2.4), элементы в правой части которого определены в (9.2.5) и (9.2.6).

На рис. 9.2 показаны прогнозы для упреждений до 36 месяцев, сделанные на один и тот же произвольный момент — июль 1957 г. Видно, что простая модель, содержащая только два параметра, хорошо воспроизводит сезонные особенности и дает превосходные прогнозы. Следует помнить, конечно, что, как и все предсказания, получаемые из общей линейной стохастической модели, прогнозирующая функция подстраивается к данным. Когда происходят изменения в сезонных явлениях, они соответствующим образом отражаются в прогнозе. Нужно заметить, что если прогноз на месяц вперед дает завышенное значение, у всех более отдаленных прогнозов на тот же момент времени существует тенденция к завышению. Этого следовало ожидать, потому что, как указано в приложении П5.1, ошибки

прогноза на один и тот же момент времени с различными упреждениями сильно коррелированы. Конечно, прогноз далеко вперед, скажем на 36 мес., неизбежно может содержать значительную ошибку. Однако на практике первоначально отдаленный прогноз будет непрерывно корректироваться, и по мере уменьшения упреждения будет достигаться все большая точность.

Описанная процедура прогнозирования устойчива к умеренным изменениям значений параметров. Так, если мы используем

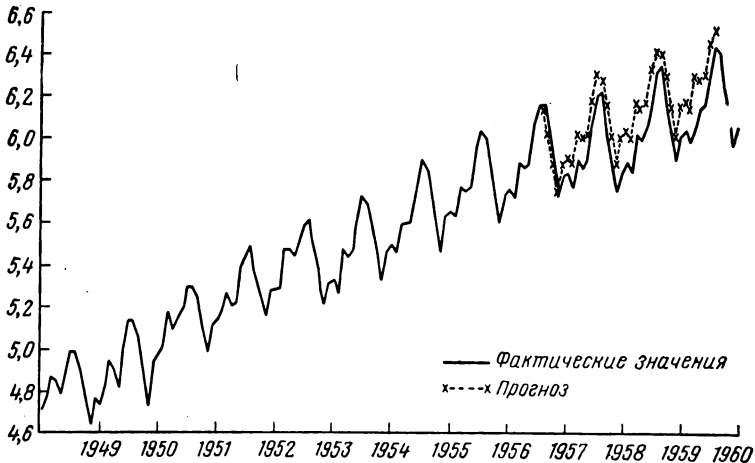


Рис. 9.2. Ряд Г с прогнозами на 1, 2, 3, ..., 36 месяцев вперед, сделанными в один и тот же произвольный момент, июль 1957 г. По оси ординат — натуральные логарифмы суммарного числа пассажиров за месяц (в тысячах).

вместо $\theta = 0,4$ и $\Theta = 0,6$ значения $\theta = 0,5$ и $\Theta = 0,5$, прогнозы не сильно изменятся. Это верно даже для прогнозов на несколько шагов вперед, например на 12 мес.

Прогнозирующая функция, ее коррекция и дисперсия ошибки прогноза. Как говорилось в гл. 5, способ разностного уравнения наиболее прост и удобен для фактического вычисления и коррекции прогнозов. Однако само разностное уравнение не объясняет в достаточной степени природу вычисляемых с его помощью прогнозов и характер коррекции. Чтобы оценить эти аспекты (а не предложить еще одну вычислительную процедуру), рассмотрим теперь прогнозирование с других точек зрения.

Прогнозирующая функция. Используя (5.1.12), получаем

$$z_{t+l} = \hat{z}_t(l) + e_t(l), \quad (9.2.8)$$

где

$$e_t(l) = a_{t+l} + \psi_1 a_{t+l-1} + \dots + \psi_{l-1} a_{t+1}.$$

Оператор скользящего среднего в правой части (9.2.1) имеет порядок 13. Отсюда, согласно (5.3.2), для $l > 13$ прогнозы удовлетворяют разностному уравнению

$$(1 - B)(1 - B^{12}) \hat{z}_t(l) = 0, \quad l > 13, \quad (9.2.9)$$

где B действует на l .

Пусть теперь $r = 0, 1, 2, \dots, m = 1, 2, \dots, 12$ и $l = (r, m)$ — упреждение в r лет и m месяцев, так что, например, $l = 15 =$

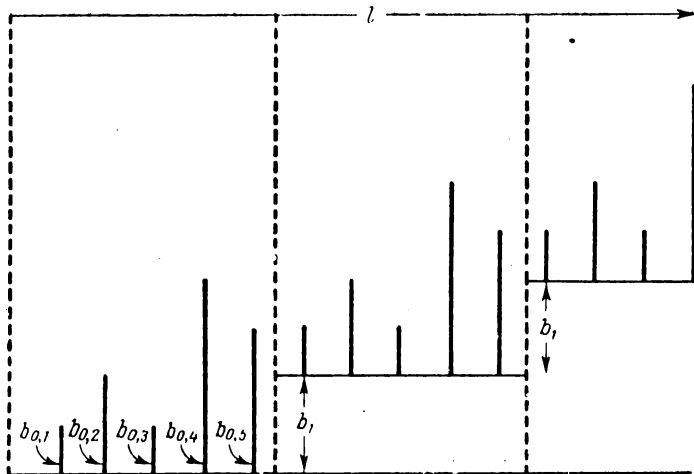


Рис. 9.3. Сезонная прогнозирующая функция, полученная по модели $\nabla \nabla_s z_t = (1 - \theta B) \times (1 - \theta B^s) a_t$ при $s = 5$.

$= (1, 3)$. Тогда прогнозирующая функция, являющаяся решением (9.2.9) с начальными условиями, определяемыми первыми 13 прогнозами, имеет вид

$$\hat{z}_t(l) = \hat{z}_t(r, m) = b_{0,m}^{(t)} + r b_1^{(t)}, \quad l > 0. \quad (9.2.10)$$

Эта прогнозирующая функция содержит 13 регулируемых коэффициентов $b_{0,1}^{(t)}, b_{0,2}^{(t)}, \dots, b_{0,12}^{(t)}, b_1^{(t)}$. Они представляют эффект 12 мес. и один годовой эффект и определяются первыми 13 прогнозами. Природу этой функции лучше понять из рис. 9.3, на котором показана прогнозирующая функция этого типа, но с периодом $s = 5$, так что имеется 6 регулируемых коэффициентов $b_{0,1}^{(t)}, b_{0,2}^{(t)}, \dots, b_{0,5}^{(t)}, b_1^{(t)}$.

Веса ψ . Чтобы определить формулы коррекции и получить дисперсию ошибки прогноза $e_t(l)$, нам нужны веса ψ в выражении для модели вида $z_t = \sum_{j=0}^{\infty} \psi_j a_{t-j}$. Оператор скользящего

среднего в (9.2.1) можно записать

$$(1 - \theta B)(1 - \Theta B^{12}) = (\nabla + \lambda B)(\nabla_{12} + \Lambda B^{12}),$$

где $\lambda = 1 - \theta$, $\Lambda = 1 - \Theta$, $\nabla_{12} = 1 - B^{12}$.

Тогда модель (9.2.1) можно представить как

$$\nabla \nabla_{12} z_t = (\nabla + \lambda B)(\nabla_{12} + \Lambda B^{12}) a_t.$$

Суммируя, получим

$$z_t = \lambda S a_{t-1} + \Lambda S_{12} a_{t-12} + \lambda \Lambda S S_{12} a_{t-13} + a_t,$$

где $S_{12} = 1 + B^{12} + B^{24} + \dots$. Таким образом, веса ψ этого процесса равны

$$\psi_0 = 1,$$

$$\psi_1 = \psi_2 = \dots = \psi_{11} = \lambda,$$

$$\psi_{12} = \lambda + \Lambda,$$

$$\psi_{13} = \psi_{14} = \dots = \psi_{23} = \lambda(1 + \Lambda),$$

$$\psi_{24} = \lambda(1 + \Lambda) + \Lambda,$$

$$\psi_{25} = \psi_{26} = \dots = \psi_{35} = \lambda(1 + 2\Lambda),$$

$$\psi_{36} = \lambda(1 + 2\Lambda) + \Lambda$$

и т. д. Представив ψ_j как $\psi_{r,m}$, где $r = 0, 1, 2 \dots$ и $m = 1, 2, 3, \dots$, 12 относятся соответственно к годам и месяцам, получаем

$$\psi_{r,m} = \lambda(1 + r\Lambda) + \delta\Lambda, \quad (9.2.11)$$

где

$$\delta = \begin{cases} 1, & \text{когда } m = 12, \\ 0, & \text{когда } m \neq 12. \end{cases}$$

Коррекция. Общая формула для коррекции (5.2.5) имеет вид

$$\hat{z}_{t+1}(l) = \hat{z}_t(l+1) + \psi_l a_{t+1}.$$

Тогда, если $m \neq s = 12$, то

$$b_{0,m}^{(t+1)} + r b_{1,m}^{(t+1)} = b_{0,m+1}^{(t)} + r b_{1,m+1}^{(t)} + (\lambda + r\Lambda) a_{t+1},$$

и, приравнявая коэффициенты при одинаковых степенях r , получаем

$$\begin{aligned} b_{0,m}^{(t+1)} &= b_{0,m+1}^{(t)} + \lambda a_{t+1}, \\ b_{1,m}^{(t+1)} &= b_{1,m+1}^{(t)} + \Lambda a_{t+1}. \end{aligned} \quad (9.2.12)$$

В случае же если $m = s = 12$, то

$$b_{0,12}^{(t+1)} + r b_{1,12}^{(t+1)} = b_{0,1}^{(t)} + (r+1) b_{1,1}^{(t)} + (\lambda + \Lambda + r\Lambda) a_{t+1},$$

и в этом случае

$$\begin{aligned} b_{0,12}^{(t+1)} &= b_{0,1}^{(t)} + b_{1,1}^{(t)} + (\lambda + \Lambda) a_{t+1}, \\ b_{1,12}^{(t+1)} &= b_{1,1}^{(t)} + \Lambda a_{t+1}. \end{aligned} \quad (9.2.13)$$

При анализе этих соотношений следует помнить, что $b_{0,m}^{(t+1)}$ — откорректированное $b_{0,m+1}^{(t)}$. Поэтому, если момент t , в который делается прогноз, был январь какого-либо года, тогда $b_{0,2}^{(t)}$ — это коэффициенты для марта. После того, как прошел месяц, нужно передвинуть момент прогноза на февраль, и откорректированный вариант коэффициента для марта теперь будет $b_{0,1}^{(t+1)}$.

Дисперсия ошибок прогноза. Зная веса ψ , мы можем вычислить дисперсию ошибок прогноза для любого упреждения l . Воспользуемся для этого результатом (5.1.16), а именно

$$V(l) = [1 + \psi_1^2 + \dots + \psi_{l-1}^2] \sigma_a^2. \quad (9.2.14)$$

Принимая $\lambda = 0,6$, $\Lambda = 0,4$, $\sigma_a^2 = 1,34 \cdot 10^{-3}$ в (9.2.11) и (9.2.14), находим выборочные стандартные отклонения $\hat{\sigma}(l)$ логарифмических ошибок прогноза для данных об авиаперевозках; результаты для l от 1 до 36 приведены в табл. 9.2.

Таблица 9.2

**Выборочные стандартные отклонения ошибок прогноза
с различным упреждением для логарифмов ряда авиаперевозок**

Упреждения $\hat{\sigma}(l) \cdot 10^{-2}$	1 3,7	2 4,3	3 4,8	4 5,3	5 5,8	6 6,2	7 6,6	8 6,9	9 7,2	10 7,6	11 8,0	12 8,2
Упреждения $\hat{\sigma}(l) \cdot 10^{-2}$	13 9,0	14 9,5	15 10,0	16 10,5	17 10,9	18 11,4	19 11,7	20 12,1	21 12,6	22 13,0	23 13,3	24 13,6
Упреждения $\hat{\sigma}(l) \cdot 10^{-2}$	25 14,4	26 15,0	27 15,5	28 16,0	29 16,4	30 17,0	31 17,4	32 17,8	33 18,3	34 18,7	35 19,2	36 19,6

Прогнозы как взвешенные средние предыдущих наблюдений. Если представить модель в виде

$$z_t = \sum_{j=1}^{\infty} \pi_j z_{t-j} + a_t,$$

прогноз на шаг вперед будет

$$\hat{z}_t(1) = \sum_{j=1}^{\infty} \pi_j z_{t+1-j}.$$

Веса π можно получить, приравнявая коэффициенты при одинаковых степенях B в уравнении

$$(1 - B)(1 - B^{12}) = (1 - \theta B)(1 - \theta B^{12})(1 - \pi_1 B - \pi_2 B^2 - \dots).$$

Отсюда

$$\begin{aligned}\pi_j &= \theta^{j-1}(1-\theta), \quad j=1, 2, \dots, 11, \\ \pi_{12} &= \theta^{11}(1-\theta) + (1-\theta), \\ \pi_{13} &= \theta^{12}(1-\theta) - (1-\theta)(1-\theta), \\ (1-\theta B - \theta B^{12} + \theta \theta B^{13})\pi_j &= 0, \quad j \geq 14.\end{aligned}\quad (9.2.15)$$

На рис. 9.4 показаны эти веса для значений параметров $\theta = 0,4$ и $\Theta = 0,6$. Почему весовая функция имеет такой вид, как

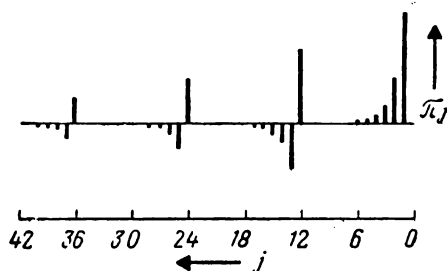


Рис. 9.4. Веса π для процесса $(0,1,1) \times (0,1,1)_{12}$, подогнанного к ряду G ($\theta = 0,4$, $\Theta = 0,6$).

на рис. 9.4, можно понять из следующего. Процесс (9.2.1) можно описать как

$$a_{t+1} = \left\{ 1 - \frac{\lambda B}{1 - \theta B} \right\} \left\{ 1 - \frac{\Lambda B^{12}}{1 - \Theta B^{12}} \right\} z_{t+1}. \quad (9.2.16)$$

Используем теперь символ $\text{ЭВСС}_\lambda(z_t)$ для обозначения экспоненциально взвешенного скользящего среднего с параметром λ от значений $z_t, z_{t-1}, z_{t-2}, \dots$, так что

$$\text{ЭВСС}_\lambda(z_t) = \frac{\lambda}{1 - \theta B} z_t = \lambda z_t + \lambda \theta z_{t-1} + \lambda \theta^2 z_{t-2} + \dots$$

Аналогично пусть символ $\text{ЭВСС}_\Lambda(z_t)$ обозначает экспоненциально взвешенное среднее с параметром Λ от значений $z_t, z_{t-12}, z_{t-24}, \dots$, так что

$$\text{ЭВСС}_\Lambda(z_t) = \frac{\Lambda}{1 - \Theta B^{12}} z_t = \Lambda z_t + \Lambda \Theta z_{t-12} + \Lambda \Theta^2 z_{t-24} + \dots$$

Подставляя в (9.2.16) $\hat{z}_t(1) = z_{t+1} - a_{t+1}$, получаем

$$\hat{z}_t(1) = \text{ЭВСС}_\lambda(z_t) + \text{ЭВСС}_\Lambda(z_{t-11}) - \text{ЭВСС}_\lambda(z_{t-12}). \quad (9.2.17)$$

Таким образом, прогноз — это ЭВСС, взятое по предыдущим месяцам, измененное вторым ЭВСС от расхождений между такими ежемесячными ЭВСС и фактически наблюдаемыми в предыдущие годы.

Например, пусть мы пытались предсказать сбыт универсального магазина в декабре. Большая доля распродажи приходится на рождественские подарки. Первый член в правой части (9.2.17) будет ЭВСС за предыдущие месяцы по ноябрь включительно. Однако мы знаем, что эта величина существенно ниже реальной, и мы поправили ее, прибавив к ней второе ЭВСС за предыдущие годы, взятое по *расхождениям* между фактической распродажей в декабре и соответствующими ежемесячными ЭВСС, взятыми по предыдущим месяцам в эти годы.

Прогнозы для упреждений $l > 1$ можно генерировать непосредственно по весам π , подставляя прогнозы для меньших упреждений вместо неизвестных величин. Так,

$$\hat{z}_t(2) = \pi_1 \hat{z}_t(1) + \pi_2 z_t + \pi_3 z_{t-1} + \dots$$

Можно поступить и иначе — вычислить явные значения весов, придаваемых непосредственно $z_t, z_{t-1}, z_{t-2}, \dots$, по формуле (5.3.9) или (П5.2.3).

9.2.3. Идентификация

Идентификация несезонного процесса ПСС(0, 1, 1) опирается на тот факт, что после взятия первых разностей автокорреляции для всех задержек, больших единицы, равны нулю. Для мультипликативного процесса $(0, 1, 1) \times (0, 1, 1)_{12}$ (9.2.1) ненулевые автокорреляции $\nabla \nabla_{12} z_t$ соответствуют только задержкам 1, 11, 12 и 13.

Действительно, эти автоковариации равны

$$\begin{aligned} \gamma_0 &= (1 + \theta^2)(1 + \Theta^2) \sigma_a^2, \\ \gamma_1 &= -\theta(1 + \Theta^2) \sigma_a^2, \\ \gamma_{11} &= \theta \Theta \sigma_a^2, \\ \gamma_{12} &= -\Theta(1 + \theta^2) \sigma_a^2, \\ \gamma_{13} &= \theta \Theta \sigma_a^2. \end{aligned} \tag{9.2.18}$$

В табл. 9.3 приведены выборочные автокорреляции логарифмированных данных авиаперевозок для

- исходного прологарифмированного ряда z_t ,
- ряда ∇z_t , полученного из (а) взятием разностей с шагом в месяц,
- ряда $\nabla_{12} z_t$, полученного из (а) взятием разностей с шагом в год,
- ряда $\nabla \nabla_{12} z_t$, полученного из (а) взятием разностей через месяц и через год.

Таблица 9.3

Выборочные автокорреляции различных разностей логарифмированных данных по авианерезовкам

Задержки		Автокорреляции											
(a) z	1—12	0,95	0,90	0,85	0,81	0,78	0,76	0,74	0,73	0,73	0,74	0,76	0,76
	13—24	0,72	0,66	0,62	0,58	0,54	0,52	0,50	0,49	0,50	0,50	0,52	0,52
	25—36	0,48	0,44	0,40	0,36	0,34	0,31	0,30	0,29	0,30	0,30	0,31	0,32
	37—48	0,29	0,24	0,21	0,17	0,15	0,12	0,11	0,10	0,10	0,11	0,12	0,13
(b) ∇z	1—12	0,20	-0,12	-0,15	-0,32	-0,08	0,03	-0,11	-0,34	-0,12	-0,11	0,21	0,84
	13—24	0,22	-0,14	-0,12	-0,28	-0,05	0,01	-0,11	-0,34	-0,11	-0,08	0,20	0,74
	25—36	0,20	-0,12	-0,10	-0,21	-0,06	0,02	-0,12	-0,29	-0,13	-0,04	0,15	0,66
	37—48	0,19	-0,13	-0,06	-0,16	-0,06	0,01	-0,11	-0,28	-0,11	-0,03	0,12	0,59
(c) $\nabla_{12} z$	1—12	0,71	0,62	0,48	0,44	0,39	0,32	0,24	0,19	0,15	-0,01	-0,12	-0,24
	13—24	-0,14	-0,14	-0,10	-0,15	-0,10	-0,11	-0,14	-0,16	-0,11	-0,08	0,00	-0,05
	25—36	-0,10	-0,09	-0,13	-0,15	-0,19	-0,20	-0,19	-0,15	-0,22	-0,23	-0,27	-0,22
	37—48	-0,18	-0,16	-0,14	-0,10	-0,05	0,02	0,04	0,10	0,15	0,22	0,29	0,30
(d) $\nabla \nabla_{12} z$	1—12	-0,34	0,11	-0,20	0,02	0,06	0,03	-0,06	0,00	0,18	-0,08	0,06	-0,39
	13—24	0,15	-0,06	0,15	-0,14	0,07	0,02	-0,01	-0,12	0,04	-0,09	0,22	-0,02
	25—36	-0,10	0,05	-0,03	0,05	-0,02	-0,05	-0,05	0,20	-0,12	0,08	-0,15	-0,01
	37—48	0,05	0,03	-0,02	-0,03	-0,07	0,10	-0,09	0,03	-0,04	-0,04	0,11	-0,05

Автокорреляции z велики и не затухают при больших задержках. В то время как взятие простых разностей в общем уменьшает корреляцию, сохраняется очень сильная периодическая компонента. Об этом свидетельствуют очень большие корреляции при задержках 12, 24, 36 и 48. Взятие разностей с периодом 12 приводит к корреляциям, стабильно положительным, а затем стабильно отрицательным. Наконец, взятие комбинированных разностей $\nabla\nabla_{12}$ заметно уменьшает корреляцию всюду.

В предположении, что модель имеет вид (9.2.1), дисперсии выборочных автокорреляций для больших задержек аппроксимировались формулой Бартлетта (2.1.13), которая в этом случае имела вид

$$\text{var}[r_k] \approx \frac{1 + 2(\rho_1^2 + \rho_{11}^2 + \rho_{12}^2 + \rho_{13}^2)}{n}, \quad k > 13. \quad (9.2.19)$$

Заменяя ρ в (9.2.19) выборочными корреляциями и полагая $n = 144 - 13 = 131$, где $n = 131$ — число разностей $\nabla\nabla_{12}z_t$, получаем стандартную ошибку $\hat{\sigma}(r) \approx 0,11$.

В табл. 9.4 наблюдаемые частоты значений 35 автокорреляций r_k , $k > 13$, сравниваются с частотами нормального распределения с нулевым средним значением и стандартным отклонением 0,11. Эта грубая проверка показывает, что модель заслуживает дальнейших исследований.

Таблица 9.4

Сравнение наблюдаемых и ожидаемых частот автокорреляций $\nabla\nabla_{12}z_t$ для задержек, больших 13

	Ожидаемые для нормального распределения с нулевым средним значением и станд. откл. 0,11	Наблюдаемые *)
$0 < r < 0,11$	23,9	27,5
$0,11 < r < 0,22$	9,5	7,0
$0,22 < r $	1,6	0,5
	35,0	35,0

* Наблюдения на границе двух ячеек гистограммы приписываются как $\frac{1}{2}$ каждой из соседних ячеек.

Предварительные оценки. Как и для несезонных моделей, приравнивая наблюдаемые корреляции их математическим ожиданиям, можно получить приближенные оценки для параметров θ и Θ . Подставляя выборочные оценки $r_1 = -0,34$ и $r_{12} = -0,39$ в выражения

$$\rho_1 = \frac{-\theta}{1 + \theta^2}, \quad \rho_{12} = \frac{-\theta}{1 + \theta^2},$$

получаем грубые оценки $\hat{\theta} \approx 0,39$ и $\hat{\Theta} \approx 0,48$. В приложении П9.1 дана таблица, суммирующая поведение автокорреляционной функции для некоторых видов моделей сезонных рядов; эта таблица полезна для идентификации и получения предварительных оценок параметров.

9.2.4. Оценивание

На рис. 9.5 показаны изолинии суммы квадратов $S(\theta, \Theta)$ для данных об авиаперевозках, к которым подгонялась модель

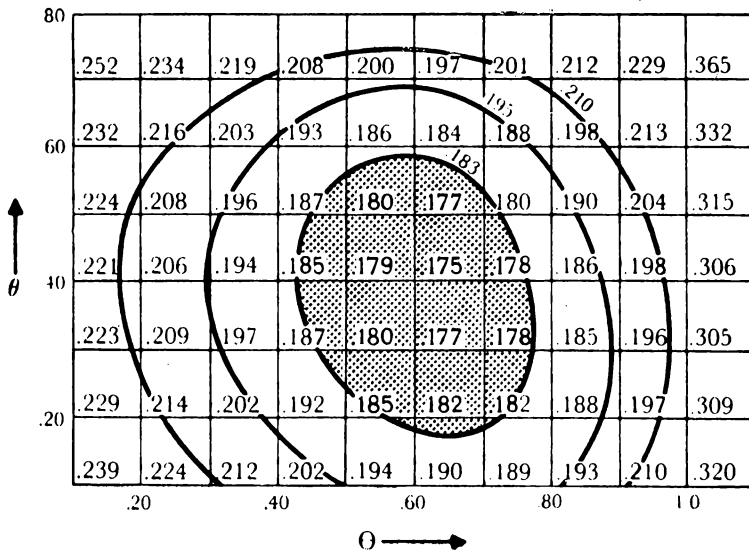


Рис. 9.5. Подгонка модели $\nabla \nabla_{12} z_t = (1 - \theta B)(1 - \Theta B^{12})a_t$ к ряду G: изолинии $S(\theta, \Theta)$; заштрихована 95%-ная доверительная область.

(9.2.1); там же показана соответствующая 95%-ная доверительная область.

Оценки наименьших квадратов очень близки к $\hat{\theta} = 0,4$ и $\hat{\Theta} = 0,6$. Значения $S(\theta, \Theta)$ в узлах сетки были вычислены по методике, описанной в гл. 7. Там было показано, что если дано n наблюдений w линейного процесса, описываемого формулой

$$\phi(B)w_t = \theta(B)a_t$$

или эквивалентной формулой

$$\phi(F)w_t = \theta(F)e_t,$$

квадратичная форма $\mathbf{w}'\mathbf{M}_n\mathbf{w}$, стоящая в показателе функции правдоподобия, всегда может быть выражена как сумма квадратов условных математических ожиданий a или e , т. е.

$$\mathbf{w}'\mathbf{M}_n\mathbf{w} = S(\boldsymbol{\phi}, \theta) = \sum_{t=-\infty}^n [a_t]^2 = \sum_{t=1}^{\infty} [e_t]^2,$$

где

$$[a_t] = [a_t | \mathbf{w}, \boldsymbol{\phi}, \theta],$$

$$[e_t] = [e_t | \mathbf{w}, \boldsymbol{\phi}, \theta].$$

Далее там было показано, что $S(\boldsymbol{\phi}, \theta)$ играет главную роль в оценивании параметров как в теории выборочных распределений, так и в методе правдоподобия или в байесовском подходе.

Вычисления для моделей сезонных рядов осуществляются точно так же, как описано в разд. 7.1.5 для несезонных моделей. Мы покажем, как проводятся вычисления оценок для данных об авиаперевозках по модели

$$\nabla\nabla_{12}z_t = w_t = (1 - \theta B)(1 - \theta B^{12})a_t.$$

Схема вычислений $[a]$ показана в табл. 9.5. Если имеется N наблюдений z , то в общем случае с разностным оператором $\nabla^d \nabla_s^D$ можно вычислить $n = N - d - sD$ значений w . Поэтому удобно использовать такие обозначения, чтобы первое наблюдение в

Таблица 9.5

Данные по авиаперевозкам.
Вычислительная таблица для $[a]$ и, следовательно, для $S(\theta, \theta)$

z_t	t	$[a_t]$	$[w_t]$	$[e_t]$
		0	0	0
z_{-12}	-12	$[a_{-12}]$	$[w_{-12}]$	0
z_{-11}	-11	$[a_{-11}]$	$[w_{-11}]$	0
.
.
z_0	0	$[a_0]$	$[w_0]$	0
z_1	1	$[a_1]$	w_1	$[e_1]$
z_2	2	$[a_2]$	w_2	$[e_2]$
.
.
z_{131}	131	$[a_{131}]$	w_{131}	$[e_{131}]$

ряду z имело индекс $1 - d - sD$. Первое наблюдение в ряду w будет тогда иметь индекс 1, а последнее n . В ряду данных об авиаперевозках имеется $N = 144$ наблюдений. В соответствии с этим в табл. 9.5 они обозначены как $z_{-12}, z_{-11}, \dots, z_{131}$. Величины w , полученные взятием разностей, образуют ряд w_1, w_2, \dots, w_n , где $n = 131$. Эти значения приведены в центре таблицы.

Фундаментальные рекуррентные формулы для расчета в прямом и обратном направлениях можно получить, перейдя к условным математическим ожиданиям в соответствующих представлениях модели. В этом примере получаем

$$[e_t] = [w_t] + \theta[e_{t+1}] + \Theta[e_{t+12}] - \theta\Theta[e_{t+13}], \quad (9.2.20)$$

$$[a_t] = [w_t] + \theta[a_{t-1}] + \Theta[a_{t-12}] - \theta\Theta[a_{t-13}]. \quad (9.2.21)$$

В общем случае модель сезонных рядов может включать стационарный оператор авторегрессии $\phi(B)\Phi(B^s)$ степени $(p + sP)$. Если мы хотим, чтобы расчет в обратном направлении, описанный в разд. 7.1.5, уходил в прошлое как можно дальше, нужно начинать рекурсию с вычисления приближенного значения $[e_{n-sP}]$, получаемого приравниванием всех неизвестных $[e]$ нулю. В данном примере $p = P = 0$, и отсюда, пользуясь (9.2.20), можно начинать с

$$[e_{131}] = w_{131} + \{\theta \times 0\} + \{\Theta \times 0\} - \{\theta\Theta \times 0\},$$

$$[e_{130}] = w_{130} + \{\theta \times [e_{131}]\} + \{\Theta \times 0\} - \{\theta\Theta \times 0\}$$

и т. д. до получения $[e_1]$. Поскольку $[e_{-j}] = 0$ при $j \geq 0$, мы можем использовать (9.2.20) для вычисления прогнозов назад: $[w_0], [w_{-1}], \dots, [w_{-12}]$. Далее, все значения $[w_{-j}]$ для $j > 12$ равны нулю, и, поскольку каждое $[a_t]$ — это функция ранее наблюдавшихся $[w]$, отсюда следует (это вытекает и непосредственно из представления модели), что $[a_{-j}] = 0$ при $j > 12$. Поэтому (9.2.21) теперь может быть использовано непосредственно для вычисления $[a]$ и, следовательно, для вычисления

$$S(\theta, \Theta) = \sum_{t=-12}^{131} [a_t]^2. \text{ Почти во всех интересных случаях переход-}$$

ные процессы, введенные аппроксимацией в начале обратной рекурсии, будут пренебрежимо слабо влиять на вычисление предыдущих $[w]$. Поэтому вычисленное этим способом $S(\theta, \Theta)$ будет весьма точным. Однако, как указывалось в разд. 7.1.5, можно еще больше повысить точность, повторив итерации «вперед и назад». Следующая итерация включает повторное вычисление $[e]$ и начинается с использования прогнозов $[w_{n+1}], [w_{n+2}], \dots, [w_{n+13}]$, полученных по уже вычисленным $[a]$.

Итеративное вычисление оценок наименьших квадратов $\hat{\theta}$, $\hat{\Theta}$. Как уже обсуждалось в разд. 7.2, в каждой новой ситуации или при появлении каких-либо трудностей полезно рисовать изолинии сумм квадратов, но в простых случаях можно прямо воспользоваться методикой итеративной линеаризации для получения оценок наименьших квадратов и их приближенных стандартных ошибок. Процедура вычислений была описана в разд. 7.2.1, и при оценке параметров сезонных моделей не возникает новых трудностей.

В настоящем примере можно приближенно записать

$$a_{t,0} = (\theta - \theta_0) x_{1,t} + (\Theta - \Theta_0) x_{2,t} + a_t,$$

где

$$x_{1,t} = - \left. \frac{\partial a_t}{\partial \theta} \right|_{\theta_0, \Theta_0}, \quad x_{2,t} = - \left. \frac{\partial a_t}{\partial \Theta} \right|_{\theta_0, \Theta_0}$$

и где θ_0 и Θ_0 — предполагаемые значения, а $a_{t,0} = [a_t | \theta_0, \Theta_0]$. Как было показано в разд. 7.2.2, производные легче всего рассчитывать численно. Действуя таким образом и пользуясь в качестве начальных значений предварительными оценками $\hat{\theta} = 0,39$, $\hat{\Theta} = 0,48$, полученными в разд. 9.2.3 по выборочным автокорреляциям, продолжаем итерации (см. табл. 9.6).

Таблица 9.6

Итеративное оценивание θ и Θ по логарифмированным данным авиаперевозок

Итерация	θ	Θ
Начальные значения	0,390	0,480
1	0,404	0,640
2	0,395	0,612
3	0,396	0,614
4	0,396	0,614

В результате трех итераций можно получить значения параметров с точностью до двух знаков после запятой, удовлетворяющей все практические требования. Выборочная дисперсия остаточных ошибок равна $\hat{\sigma}_a^2 = 1,34 \cdot 10^{-3}$. Матрица, обратная матрице сумм квадратов и произведений x , полученная на последней итерации, используется для вычисления стандартных ошибок оценок. Оценки наименьших квадратов и соответствую-

щие стандартные ошибки равны

$$\hat{\theta} = 0,40 \pm 0,08,$$

$$\hat{\Theta} = 0,61 \pm 0,07,$$

что хорошо согласуется со значениями, полученными на рис. 9.5.

Процедура оценивания может быть кратко описана следующим образом. Имея достаточно хорошие оценки параметров, полученные по автокорреляциям, можно, как и ранее, вычислить прогнозы ω назад. Полученные начальные значения используются на всех итерациях, пока процесс не сойдется. В этот момент еще одна итерация, использующая подправленные значения прогнозов назад, обычно обеспечивает окончательную сходимость. Однако, располагая быстродействующей ЭВМ, вряд ли целесообразно пользоваться такими ухищрениями. Действуя другим способом (как в разд. 7.2.3.), путем рекуррентного счета можно получить производные с любой желаемой точностью.

Дисперсии и ковариации оценок, сделанных по большим выборкам. Как и в разд. 7.2.6, можно получить формулы для дисперсий и ковариаций оценок параметров по большим выборкам. В этом случае

$$x_{1,t} \approx -(1 - \theta B)^{-1} a_{t-1} = - \sum_{j=0}^{\infty} \theta^j B^j a_{t-1},$$

$$x_{2,t} \approx -(1 - \Theta B^{12})^{-1} a_{t-12} = - \sum_{i=0}^{\infty} \Theta^i B^{12i} a_{t-12}.$$

Отсюда для больших выборок информационная матрица имеет вид

$$I(\theta, \Theta) = n \begin{bmatrix} (1 - \theta^2)^{-1} & \theta^{11} (1 - \theta^{12}\Theta)^{-1} \\ \theta^{11} (1 - \theta^{12}\Theta)^{-1} & (1 - \Theta^2)^{-1} \end{bmatrix}.$$

Если $|\theta|$ не близко к единице, элементы, не лежащие на диагонали, малы, и приближенные значения дисперсий и ковариаций $\hat{\theta}$ и $\hat{\Theta}$ равны

$$\begin{aligned} V(\hat{\theta}) &\approx n^{-1}(1 - \theta^2), & V(\hat{\Theta}) &\approx n^{-1}(1 - \Theta^2), \\ \text{cov}(\hat{\theta}, \hat{\Theta}) &\approx 0. \end{aligned} \quad (9.2.22)$$

В данном примере, подставляя в (9.2.22) значения $\hat{\theta} = 0,40$, $\hat{\Theta} = 0,61$ и $n = 131$, получаем величины

$$\begin{aligned} V(\hat{\theta}) &\approx 0,0064, & V(\hat{\Theta}) &\approx 0,0048, \\ \sigma(\hat{\theta}) &\approx 0,08, & \sigma(\hat{\Theta}) &\approx 0,07, \end{aligned}$$

которые при данной точности совпадают со значениями, полученными непосредственно при итерациях. Интересно отметить также, что оценки параметров $\hat{\theta}$ и $\hat{\Theta}$, связанных соответственно с месяцами и годами, практически не коррелированы.

9.2.5. Диагностическая проверка

Перед тем как перейти к дальнейшему, проверим адекватность модели путем анализа остаточных ошибок после подгонки модели.

Проверка по автокорреляциям. Выборочные автокорреляции остаточных ошибок

$$\hat{a}_t = \nabla \nabla_{12} z_t + 0,40 \hat{a}_{t-1} + 0,61 \hat{a}_{t-12} - 0,24 \hat{a}_{t-13}$$

приведены в табл. 9.7.

Таблица 9.7

Выборочные автокорреляции остаточных ошибок при подгонке модели $\nabla \nabla_{12} z_t = (1 - 0,40B)(1 - 0,61B^{12})a_t$ к данным авиаперевозок (ряд G)

Задержка k	Автокорреляции $r_k(d)$												Стандартная ошибка (верхняя граница)
1-12	0,02	0,02	-0,13	-0,14	0,05	0,06	-0,07	-0,04	0,10	-0,08	0,02	-0,01	0,09
13-24	0,03	0,04	0,05	-0,16	0,03	0,00	-0,11	-0,10	-0,03	-0,03	0,22	0,03	0,09
25-36	-0,02	0,06	-0,04	-0,06	-0,05	-0,08	-0,05	0,12	-0,13	0,00	-0,06	-0,02	0,09
37-48	0,11	0,07	-0,02	-0,05	-0,10	-0,02	-0,04	0,00	-0,08	0,03	0,04	0,06	0,09

$$\sum_{k=1}^{48} r_k^2(d) = 0,2726$$

Некоторые автокорреляции кажутся довольно большими по сравнению с верхней границей их стандартных ошибок, в особенности велико значение $r_{23} = 0,22$, в 2,5 раза превышающее эту верхнюю границу. Однако следует ожидать, что среди 48 отклонений могут встретиться и большие отклонения.

Общая проверка осуществляется при помощи статистики $Q = n \sum_{k=1}^{48} r_k^2(d)$, которая (см. разд. 8.2.2) приближенно распределена как χ^2 с 46 степенями свободы (подгоняются два параметра). Наблюдаемое значение $Q = 131 \cdot 0,2726 = 35,7$, и, если данная модель верна, значений Q , больших этого, следует ожидать в 86% случаев. Эта проверка не дает никаких свидетельств о неадекватности модели.

Проверка по периодограмме. Кумулятивная периодограмма остаточных ошибок показана на рис. 9.6. 5- и 25%-ные вероятностные пределы критерия Колмогорова, как указывалось

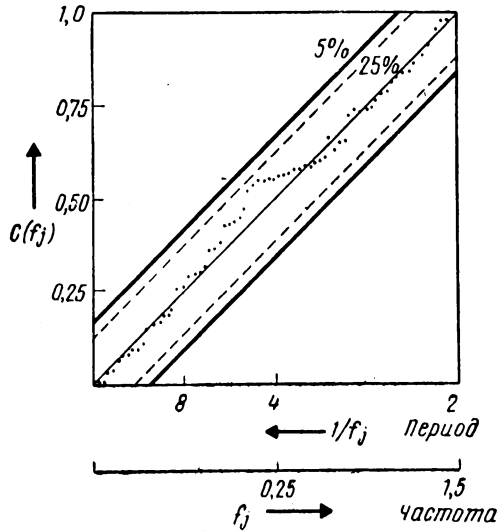


Рис. 9.6. Проверка остаточных ошибок модели $\nabla \nabla_{12} z_t = (1 - 0,40B)(1 - 0,61B^{12})a_t$, подогнанной к ряду G, при помощи кумулятивной периодограммы.

в разд. 8.2.4, дают лишь очень грубые указания значимости случайных отклонений; в этом примере они не выявляют значимых отклонений от предложенной модели.

9.3. НЕКОТОРЫЕ АСПЕКТЫ БОЛЕЕ ОБЩИХ МОДЕЛЕЙ СЕЗОННЫХ РЯДОВ

9.3.1. Мультипликативные и немумультипликативные модели

В предыдущих разделах мы обсудили методы анализа сезонных временных рядов, и в частности рассмотрели пример мультипликативной модели. Мы видели, что эта модель может хорошо описывать наблюдения с помощью довольно малого числа параметров. Нам остается рассмотреть некоторые другие модели сезонных рядов этого типа и при необходимости связанные с ними процедуры идентификации, оценивания, диагностической проверки и прогнозирования.

Предположим в общем случае, что имеется сезонный эффект с периодом s . Тогда общий класс мультипликативных моделей можно описать способом, продемонстрированным в табл. 9.8.

В мультипликативной модели предполагается, что «периодический» ход ряда описывается моделью

$$\Phi_P(B^s) \nabla_s^D z_{r,m} = \Theta_Q(B^s) a_{r,m},$$

в то же время «внутри цикла» α связаны соотношением

$$\phi_P(B) \nabla^d a_{r,m} = \theta_q(B) a_{r,m}.$$

Очевидно, можно изменить порядок рассмотрения моделей и в любом случае прийти к общей мультипликативной модели

$$\phi_P(B) \Phi_P(B^s) \nabla^d \nabla_s^D z_{r,m} = \theta_q(B) \Theta_Q(B^s) a_{r,m}, \quad (9.3.1)$$

где $a_{r,m}$ — процесс белого шума с нулевым средним значением. Практическая ценность модели (9.3.1) зависит от того, удастся ли с ее помощью экономично параметризовать фактические наблюдения. Имеется много примеров удачного использования таких моделей для описания разнообразных сезонных рядов, относящихся к совершенно различным проблемам [80].

Таблица 9.8

Таблица с двумя входами для мультипликативных моделей сезонных рядов

Внутри циклов \longrightarrow								
∇^d								
		Параметры скользящего среднего ($\theta_1, \theta_2, \dots, \theta_q$)						
		Параметры авторегрессии ($\phi_1, \phi_2, \dots, \phi_p$)						
Между циклами \downarrow	∇_s^D	$z_{1,1}$	$z_{1,2}$			$z_{1,m}$	\dots	$z_{1,s}$
	Параметры скользящего среднего ($\theta_1, \theta_2, \dots, \theta_q$)	$z_{2,1}$	$z_{2,2}$			$z_{2,m}$	\dots	$z_{2,s}$
	Параметры авторегрессии ($\phi_1, \phi_2, \dots, \phi_p$)	$z_{r,1}$	$z_{r,2}$			$z_{r,m}$	\dots	$z_{r,s}$

Все же получить полностью адекватное описание всех сезонных временных рядов при помощи мультипликативных моделей невозможно. Иногда оказывается полезной модификация, оставляющая смешанный оператор скользящего среднего немультимпликативным. Это означает, что мы заменяем оператор $\theta_q(B) \Theta_Q(B^s)$ в правой части (9.3.1) более общим оператором скользящего среднего $\theta_q^*(B)$. Вместо (или кроме) этого может

оказаться необходимым заменить оператор авторегрессии $\Phi_p(B)\Phi_p(B^s)$ слева более общим оператором авторегрессии $\Phi_p^*(B)$. Некоторые типы немультимпликативных моделей приведены в приложении П9.1. Им даны номера 4, 4а, 5 и 5а.

Но даже в тех случаях, когда оказалось необходимым применить немультимпликативную модель, из имеющегося опыта следует, что в качестве начального приближения полезна наилучшим образом подогнанная мультимпликативная модель; этот прием позволяет построить лучшую немультимпликативную модель. Подобная ситуация встречается при использовании таблиц дисперсионного анализа, имеющих два входа, где предположение об аддитивности строк и столбцов может быть верным или неоправданным, но во всех случаях полезно как начальное приближение.

Общая стратегия для привязки мультимпликативных или немультимпликативных моделей к данным та же, что описана и проиллюстрирована с некоторыми подробностями в разд. 9.2. Руководствуясь видом автокорреляционной функции, приходим к следующему:

1. Применяем к наблюдаемому ряду разностные операторы ∇ и ∇_s для достижения стационарности.
2. По виду автокорреляционной функции разностного ряда выбираем пробную модель.
3. По значениям соответствующих автокорреляций разностного ряда получаем предварительные оценки параметров. Их можно использовать как начальные значения при поиске оценок наименьших квадратов.
4. После подгонки диагностическая проверка остаточных ошибок может либо подтвердить правильность модели, либо указать пути ее улучшения, что приводит к новой подгонке и повторению диагностических проверок.

9.3.2. Идентификация

Полезные сведения для идентификации моделей содержатся в приложении П9.1, где показана структура ковариаций $w_t = \nabla^d \nabla_s^D z_t$ для нескольких простых моделей сезонных рядов. Этот перечень не претендует на исчерпывающий охват возможных моделей. Однако мы полагаем, что он включает некоторые наиболее употребительные модели, и читателю будет нетрудно определить характеристики других моделей, которые могут оказаться полезными. Нужно подчеркнуть, что довольно простые модели (такие, как модели 1 и 2 в приложении П9.1) адекватно описывают многие сезонные ряды.

Полезность получения автоковариаций мультимпликативного процесса объясняется тем, что в таком процессе производящая

функция автоковариаций (3.1.11) есть произведение производящих функций компонент. Поэтому, если компоненты в (9.3.1)

$$\phi_p(B) \nabla^d z_t = \theta_q(B) a_t, \quad \Phi_p(B^s) \nabla_s^D a_t = \Theta_q(B^s) a_t$$

имеют производящие функции автоковариаций $\gamma(B)$ и $\Gamma(B^s)$, то производящая функция автоковариаций z в модели (9.3.1) равна

$$\gamma(B) \Gamma(B^s).$$

Нужно также помнить, что может оказаться полезным параметризовать более общие модели как отклонения от подогнанных к рядам мультипликативных форм иллюстрируемым ниже способом.

Трехпараметрический немультимпликативный оператор

$$1 - \theta_1 B - \theta_{12} B^{12} - \theta_{13} B^{13}, \quad (9.3.2)$$

применяемый в моделях 4 и 5, можно записать как

$$(1 - \theta_1 B)(1 - \theta_{12} B^{12}) - \kappa B^{13}, \quad \kappa = \theta_1 \theta_{12} - (-\theta_{13}).$$

Если оценка κ велика по сравнению с ее стандартной ошибкой, это указывает на необходимость введения немультимпликативной модели, в которой θ_{13} никак не связано с θ_1 и θ_{12} . С другой стороны, если κ мало, то, приняв $\theta_1 = \theta$, $\theta_{12} = \Theta$, мы аппроксимируем этой моделью мультипликативную модель $(0, 1, 1) \times \times (0, 1, 1)_{12}$.

9.3.3. Оценивание

При оценивании параметров общих моделей сезонных рядов не возникает новых проблем. Безусловная сумма квадратов вычисляется теми же общими методами, описанными в разд. 7.1.5 и проиллюстрированными на примере в разд. 9.2.4. Как всегда, построение графиков может упростить трудные ситуации. В простых случаях итеративная процедура метода наименьших квадратов с численным расчетом производных обеспечивает быструю сходимость к оценкам наименьших квадратов и дает приближенные дисперсии и ковариации оценок. В случае необходимости всегда можно предложить рекуррентную процедуру для прямого вычисления производных.

Дисперсии и ковариации оценок по большим выборкам. Информационная матрица $I(\phi, \theta, \Phi, \Theta)$, относящаяся к большой выборке, может быть получена вычислением $E[X'X]$, где, как и в разд. 7.2.6, X — это матрица производных с обратным знаком размером $n \times (p + q + P + Q)$. Итак, для общей мультипликативной модели

$$a_t = \theta^{-1}(B) \Theta^{-1}(B^s) \phi(B) \Phi(B^s) w_t,$$

где

$$w_t = \nabla^d \nabla_s^D z_t,$$

нужные производные равны

$$\begin{aligned} \frac{\partial a_t}{\partial \theta_t} &= \theta^{-1}(B) B^t a_t, & \frac{\partial a_t}{\partial \theta_t} &= \theta^{-1}(B^s) B^{st} a_t, \\ \frac{\partial a_t}{\partial \phi_t} &= -\phi^{-1}(B) B^t a_t, & \frac{\partial a_t}{\partial \Phi_t} &= -\Phi^{-1}(B^s) B^{st} a_t. \end{aligned}$$

Приближенные значения дисперсий и ковариаций оценок получаем, как и ранее, обращением матрицы $I(\phi, \theta, \Phi, \Theta)$.

9.3.4. Эвентуальные прогнозирующие функции для различных моделей сезонных рядов

Рассмотрим теперь характеристики эвентуальных прогнозирующих функций для моделей сезонных рядов. Для модели сезонного ряда с одной периодичностью s эвентуальная прогнозирующая функция на момент t с упреждением l — это решение разностного уравнения

$$\phi(B) \Phi(B^s) \nabla^d \nabla_s^D \hat{z}_t(l) = 0.$$

В табл. 9.9 приведено это решение для различных разностных уравнений и указано также число начальных значений, от которых зависит прогнозирующая функция.

На рис. 9.7 показано поведение каждой прогнозирующей функции для $s = 4$. Удобно рассматривать время упреждения $l = rs + m$ как относящееся к прогнозу на r лет и m кварталов вперед. На рис. 9.7 необходимое число начальных значений (требуемых для начала прогноза), показанных крупными точками, было выбрано произвольно, и ход прогнозирующей функции был прослежен до конца четвертого периода. В случаях, когда разностное уравнение включало параметр авторегрессии, его значение было принято равным 0,5.

Константы $b_{0,m}$, b_1 и др., появляющиеся в табл. 9.9, строго говоря, должны обозначаться как $b_{0,m}^{(t)}$, $b_1^{(t)}$ и т. д., так как каждая из них зависит от момента t прогноза и перестраивается при каждом изменении t . Индекс t был временно опущен для упрощения записи. Оператор под номером (1) стационарен и содержит среднее значение μ . Он является оператором авторегрессии по отношению к сезонной компоненте, затухающей с каждым периодом и приближающейся к среднему значению.

Оператор (2) нестационарен по отношению к сезонной компоненте. Прогнозы для конкретного квартала разных лет связаны полиномиальной зависимостью нулевой степени. Поэтому

исходный прогноз сезонной компоненты точно повторяется в прогнозах на будущие годы.

Оператор (3) нестационарен по отношению к опорному интервалу, но стационарен по отношению к сезонной компоненте. На рис. 9.7 (3) показан общий уровень прогноза, асимптотически переходящий в новый уровень $b_0 + b_1/(1 - \Phi)$, где в то же

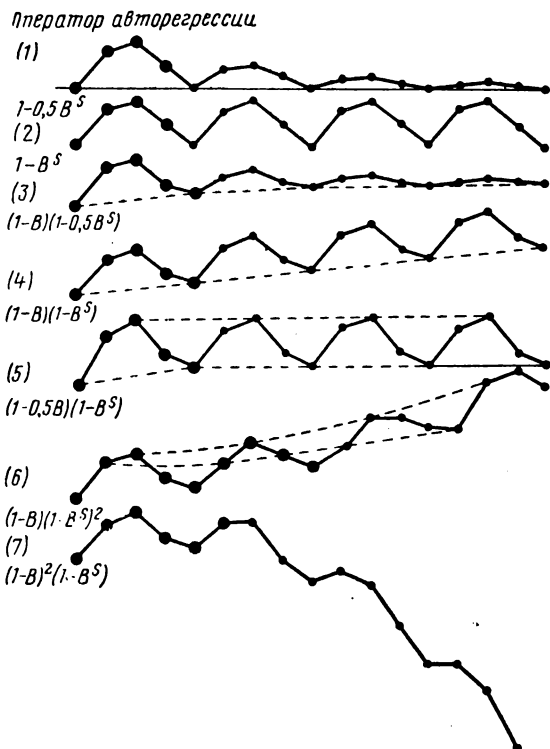


Рис. 9.7. Поведение сезонной прогнозирующей функции при различных видах общего сезонного оператора авторегрессии.

время накладывающаяся предсказуемая компонента сезонного эффекта экспоненциально затухает.

Оператор (4) — это предельный случай оператора (3) при стремлении Φ к 1. Оператор нестационарен как по отношению к опорному интервалу, так и к периодической компоненте. Исходный начальный прогноз воспроизводится в последующие годы, в то время как годовой уровень возрастает. Этот вид прогнозирующей функции дает модель мультипликативного процесса $(0, 1, 1) \times (0, 1, 1)_{12}$, подогнанная к данным по авиаперевозкам.

Таблица 9.9

Эвентуальные прогнозирующие функции для различных операторов авторегрессии.
Все коэффициенты b настраиваемые и зависят от момента прогноза t

Оператор авторегрессии	Эвентуальная прогнозирующая функция $z(r, m)$	Число начальных значений, от которых зависит прогнозирующая функция
(1) $1 - \Phi B^s$	$\mu + (b_{0,m} - \mu) \Phi^r$	s
(2) $1 - B^s$	$b_{0,m}$	s
(3) $(1 - B)(1 - \Phi B^s)$	$b_0 + (b_{0,m} - b_0) \Phi^r + b_1 \left(\frac{1 - \Phi^r}{1 - \Phi} \right)$	$s + 1$
(4) $(1 - B)(1 - B^s)$	$b_{0,m} + b_1 r$	$s + 1$
(5) $(1 - \phi B)(1 - B^s)$	$b_{0,m} + b_1 \phi^{m-1} \left(\frac{1 - \phi^{sr}}{1 - \phi^s} \right)$	$s + 1$
(6) $(1 - B)(1 - B^s)^2$	$b_{0,m} + b_{1,m} r + \frac{1}{2} b_{2r}(r - 1)$	$2s + 1$
(7) $(1 - B)^2(1 - B^s)$	$b_{0,m} + [b_1 + (m - 1)b_2]r + \frac{1}{2} b_{2sr}(r - 1)$	$s + 2$

Оператор (5) нестационарен по отношению к сезонной компоненте, но стационарен по опорному интервалу. Ход прогноза экспоненциально приближается к асимптотическому ходу прогноза по опорному интервалу

$$\hat{z}_t(\infty, m) = b_{0,m} + b_1 \phi^{m-1} / (1 - \phi^s).$$

Оператор (6) нестационарен по отношению к опорному интервалу и сезонной компоненте. С годами замечен общий квадратичный тренд, а сезонный ход испытывает мелкие изменения. Отдельные кварталы имеют не только свой собственный уровень $b_{0,m}$, но и собственную скорость изменения уровня $b_{1,m}$. Поэтому в случаях, когда этот вид прогнозирующей функции удобен, могут возникать интересные ситуации. Например, при увеличении упреждения может одновременно предсказываться величина разницы в сбыте товаров летом и зимой от года к году и уменьшение разницы сбыта товаров осенью и летом.

Оператор (7) опять-таки нестационарен как по опорному интервалу, так и по сезонной компоненте, и здесь опять имеется квадратичный тренд при переходе от года к году плюс линейные изменения прогнозов при переходе от квартала к кварталу. Однако в этом случае квартальные приращения имеют постоянную *скорость* изменения.

ПРИЛОЖЕНИЕ П9.1

АВТОКОВАРИАЦИИ ДЛЯ НЕКОТОРЫХ МОДЕЛЕЙ СЕЗОННЫХ РЯДОВ

Модель	(Автоковариации w_t)/ σ_a^2	Особые свойства
<p>(1) $w_t = (1 - \theta B)(1 - \theta B^s)a_t$ $w_t = a_t - \theta a_{t-1} - \theta a_{t-s} + \theta \theta a_{t-s-1}$ $s \geq 3$</p>	<p>$\gamma_0 = (1 + \theta^2)(1 + \theta^2)$ $\gamma_1 = -\theta(1 + \theta^2)$ $\gamma_{s-1} = \theta\theta$ $\gamma_s = -\theta(1 + \theta^2)$ $\gamma_{s+1} = \gamma_{s-1}$ Все остальные ковариации равны нулю</p>	<p>(a) $\gamma_{s-1} = \gamma_{s+1}$ (b) $\rho_{s-1} = \rho_{s+1} = \rho_1 \rho_s$</p>
<p>(2) $(1 - \Phi B^s)w_t = (1 - \theta B)(1 - \theta B^s)a_t$ $w_t - \Phi w_{t-s} = a_t - \theta a_{t-1} - \theta a_{t-s} + \theta \theta a_{t-s-1}$ $s \geq 3$</p>	<p>$\gamma_0 = (1 + \theta^2) \left[1 + \frac{(\theta - \Phi)^2}{1 - \Phi^2} \right]$ $\gamma_1 = -\theta \left[1 + \frac{(\theta - \Phi)^2}{1 - \Phi^2} \right]$ $\gamma_{s-1} = \theta \left[\theta - \Phi - \frac{\Phi(\theta - \Phi)^2}{1 - \Phi^2} \right]$ $\gamma_s = -(1 + \theta^2) \left[\theta - \Phi - \frac{\Phi(\theta - \Phi)^2}{1 - \Phi^2} \right]$ $\gamma_{s+1} = \gamma_{s-1}$ $\gamma_j = \Phi \gamma_{j-s} \quad j \geq s+2$ Для $s \geq 4$ $\gamma_2, \gamma_3, \dots, \gamma_{s-2}$ равны нулю</p>	<p>(a) $\gamma_{s-1} = \gamma_{s+1}$ (b) $\gamma_j = \Phi \gamma_{j-s} \quad j \geq s+2$</p>

Модель	(Автоковариации w_t)/ σ_a^2	Особые свойства
$(3) \quad w_t = (1 - \theta_1 B - \theta_2 B^2)(1 - \theta_1 B^s - \theta_2 B^{2s}) a_t$ $w_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta_1 a_{t-s} + \theta_1 \theta_1 a_{t-s-1} +$ $+ \theta_2 \theta_1 a_{t-s-2} - \theta_2 a_{t-2s} + \theta_1 \theta_2 a_{t-2s-1} + \theta_2 \theta_2 a_{t-2s-2}$ $s \geq 5$	$\gamma_0 = (1 + \theta_1^2 + \theta_2^2)(1 + \theta_1^2 + \theta_2^2)$ $\gamma_1 = -\theta_1(1 - \theta_2)(1 + \theta_1^2 + \theta_2^2)$ $\gamma_2 = -\theta_2(1 + \theta_1^2 + \theta_2^2)$ $\gamma_{s-2} = \theta_2 \theta_1(1 - \theta_2)$ $\gamma_{s-1} = \theta_1 \theta_1(1 - \theta_2)(1 - \theta_2)$ $\gamma_s = -\theta_1(1 + \theta_1^2 + \theta_2^2)(1 - \theta_2)$ $\gamma_{s+1} = \gamma_{s-1}$ $\gamma_{s+2} = \gamma_{s-2}$ $\gamma_{2s-2} = \theta_2 \theta_2$ $\gamma_{2s-1} = \theta_1 \theta_2(1 - \theta_2)$ $\gamma_{2s} = -\theta_2(1 + \theta_1^2 + \theta_2^2)$ $\gamma_{2s+1} = \gamma_{2s-1}$ $\gamma_{2s+2} = \gamma_{2s-2}$	<p>(a) $\gamma_{s-2} = \gamma_{s+2}$</p> <p>(b) $\gamma_{s-1} = \gamma_{s+1}$</p> <p>(c) $\gamma_{2s-2} = \gamma_{2s+2}$</p> <p>(d) $\gamma_{2s-1} = \gamma_{2s+1}$</p>
Все остальные автоковариации равны нулю		

Модель	(Автоковариация w_t)/ σ_a^2	Особые свойства
(3а) Частный случай модели (3) $w_t = (1 - \theta_1 B - \theta_2 B^2)(1 - \theta B^s) a_t$ $w_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \theta a_{t-s} +$ $+ \theta_1 \theta a_{t-s-1} + \theta_2 \theta a_{t-s-2}$ $s \geq 5$	$\gamma_0 = (1 + \theta_1^2 + \theta_2^2)(1 + \theta^2)$ $\gamma_1 = -\theta_1(1 - \theta_2)(1 + \theta^2)$ $\gamma_2 = -\theta_2(1 + \theta^2)$ $\gamma_{s-2} = \theta_2 \theta$ $\gamma_{s-1} = \theta_1 \theta(1 - \theta_2)$ $\gamma_s = -\theta(1 + \theta_1^2 + \theta_2^2)$ $\gamma_{s+1} = \gamma_{s-1}$ $\gamma_{s+2} = \gamma_{s-2}$ Все остальные автоковариации равны нулю	(а) $\gamma_{s-2} = \gamma_{s+2}$ (б) $\gamma_{s-1} = \gamma_{s+1}$
(3б) Частный случай модели (3) $w_t = (1 - \theta B)(1 - \theta_1 B^s - \theta_2 B^{2s}) a_t$ $w_t = a_t - \theta a_{t-1} - \theta_1 a_{t-s} + \theta \theta_1 a_{t-s-1} -$ $- \theta_2 a_{t-2s} + \theta \theta_2 a_{t-2s-1}$ $s \geq 3$	$\gamma_0 = (1 + \theta^2)(1 + \theta_1^2 + \theta_2^2)$ $\gamma_1 = -\theta(1 + \theta_1^2 + \theta_2^2)$ $\gamma_{s-1} = \theta \theta_1(1 - \theta_2)$ $\gamma_s = -\theta_1(1 + \theta^2)(1 - \theta_2)$ $\gamma_{s+1} = \gamma_{s-1}$ $\gamma_{2s-1} = \theta \theta_2$ $\gamma_{2s} = -\theta_2(1 + \theta^2)$ $\gamma_{2s+1} = \gamma_{2s-1}$ Все остальные автоковариации равны нулю	(а) $\gamma_{s-1} = \gamma_{s+1}$ (б) $\gamma_{2s-1} = \gamma_{2s+1}$

Модель	(Автоковариации w_t)/ σ_a^2	Особые свойства
<p>(4) $w_t = (1 - \theta_1 B - \theta_s B^s - \theta_{s+1} B^{s+1}) a_t$ $w_t = a_t - \theta_1 a_{t-1} - \theta_s a_{t-s} - \theta_{s+1} a_{t-s-1}$ $s \geq 3$</p>	<p>$\gamma_0 = 1 + \theta_1^2 + \theta_s^2 + \theta_{s+1}^2$ $\gamma_1 = -\theta_1 + \theta_s \theta_{s+1}$ $\gamma_{s-1} = \theta_1 \theta_s$ $\gamma_s = \theta_1 \theta_{s+1} - \theta_s$ $\gamma_{s+1} = -\theta_{s+1}$</p> <p>Все остальные автоковариации равны нулю</p>	<p>(а) В общем $\gamma_{s-1} \neq \gamma_{s+1}$ $\gamma_1 \gamma_s \neq \gamma_{s+1}$</p>
<p>(4а) Частный случай модели (4) $w_t = (1 - \theta_1 B - \theta_s B^s) a_t$ $w_t = a_t - \theta_1 a_{t-1} - \theta_s a_{t-s}$ $s \geq 3$</p>	<p>$\gamma_0 = 1 + \theta_1^2 + \theta_s^2$ $\gamma_1 = -\theta_1$ $\gamma_{s-1} = \theta_1 \theta_s$ $\gamma_s = -\theta_s$</p> <p>Все остальные автоковариации равны нулю</p>	<p>(а) В отличие от модели (4) $\gamma_{s+1} = 0$</p>

Модель	(Автоковариация w_t)/ σ_a^2	Особые свойства
(5) $(1 - \Phi B^s) w_t = (1 - \theta_1 B - \theta_s B^s - \theta_{s+1} B^{s+1}) a_t$ $w_t - \Phi w_{t-s} = a_t - \theta_1 a_{t-1} - \theta_s a_{t-s} - \theta_{s+1} a_{t-s-1}$ $s \geq 3$	$\gamma_0 = 1 + \theta_1^2 + \frac{(\theta_s - \Phi)^2}{1 - \Phi^2} + \frac{(\theta_{s+1} + \theta_1 \Phi)^2}{1 - \Phi^2}$ $\gamma_1 = -\theta_1 + \frac{(\theta_s - \Phi)(\theta_{s+1} + \theta_1 \Phi)}{1 - \Phi^2}$ $\gamma_{s-1} = (\theta_s - \Phi) \left[\theta_1 + \Phi \frac{(\theta_{s+1} + \theta_1 \Phi)}{1 - \Phi^2} \right]$ $\gamma_s = -(\theta_s - \Phi) \left[1 - \Phi \frac{(\theta_s - \Phi)}{1 - \Phi^2} \right] +$ $+ (\theta_{s+1} + \theta_1 \Phi) \left[\theta_1 + \Phi \frac{(\theta_{s+1} + \theta_1 \Phi)}{1 - \Phi^2} \right]$ $\gamma_{s+1} = -(\theta_{s+1} + \theta_1 \Phi) \left[1 - \Phi \frac{(\theta_s - \Phi)}{1 - \Phi^2} \right]$ $\gamma_i = \Phi \gamma_{i-s} \quad i \geq s+2$ <p>Для $s \geq 4$ $\gamma_2, \dots, \gamma_{s-2}$ все равны нулю</p>	(a) $\gamma_{s-1} \neq \gamma_{s+1}$ (b) $\gamma_i = \Phi \gamma_{i-s}$ $i \geq s+2$
(5a) Частный случай модели (5) $(1 - \Phi B^s) w_t = (1 - \theta_1 B - \theta_s B^s) a_t$ $w_t - \Phi w_{t-s} = a_t - \theta_1 a_{t-1} - \theta_s a_{t-s}$ $s \geq 3$	$\gamma_0 = 1 + \frac{\theta_1^2 + (\theta_s - \Phi)^2}{1 - \Phi^2}$ $\gamma_1 = -\theta_1 \left[1 - \Phi \frac{(\theta_s - \Phi)}{1 - \Phi^2} \right]$ $\gamma_{s-1} = \frac{\theta_1(\theta_s - \Phi)}{1 - \Phi^2}$ $\gamma_s = \frac{\Phi \theta_1^2 - (\theta_s - \Phi)(1 - \Phi \theta_s)}{1 - \Phi^2}$ $\gamma_i = \Phi \gamma_{i-s} \quad i \geq s+1$ <p>Для $s \geq 4$ $\gamma_2, \dots, \gamma_{s-2}$ все равны нулю</p>	(a) В отличие от модели (5) $\gamma_{s+1} = \Phi \gamma_1$

Вспомогательные материалы

Эта часть книги — сборник вспомогательных материалов, полезных при анализе временных рядов. Она содержит ряд программ для ЭВМ, необходимых для анализа временных рядов, таблицы и диаграммы для получения предварительных оценок моделей авторегрессии — скользящего среднего вместе со стандартными таблицами уровней значимости нормального, χ^2 - и t -распределений. Далее следует полная сводка всех временных рядов, анализировавшихся в этой книге.

Описание программ для ЭВМ

Программа 1. Идентификация стохастической модели.

Программа 2. Предварительное оценивание стохастической модели.

Программа 3. Оценивание стохастической модели.

Программа 4. Прогнозирование с помощью стохастической модели.

Использование ЭВМ при построении модели

ЭВМ обладает большим быстроедействием и высокой точностью, и по сравнению с ней человеческий ум медлен и неточен. Однако ум обладает способностью к критическим и индуктивным умозаключениям, практически отсутствующей у современных машин. Это позволяет внести удачное разделение труда, согласно которому мозг и машина делают то, к чему каждый из них лучше приспособлен. Модели, построенные вычислительной машиной чисто автоматически, вряд ли будут очень хороши, и даже могут быть очень плохими. Описанные ниже программы не нацелены на автоматическое построение моделей. Скорее, при правильном их использовании они дают исследователю большую свободу в экспериментировании с данными и могут породить идеи, нужные для плодотворного итерирования. Когда форма пробной модели выбрана, часто можно достичь исключительно быстрой сходимости в получении окончательной модели.

пользуясь нелинейной итеративной процедурой, даже в тех случаях, когда использовались довольно неудачные начальные значения. Это позволяет обойтись без этапа предварительного оценивания, реализуемого программой 2. В благоприятных случаях после пробной идентификации *формы* модели при помощи программы 1 исследователь может прямо переходить к оцениванию при помощи программы 3.

Программы, описанные ниже, были составлены на АЛГОЛе в Ланкастерском университете Дж. Хэмптоном, Г. Вилсоном, Э. Ходжкинсон и П. Блант. Они представлены здесь в виде вычислительных алгоритмов. Хотя необходимые входные и желаемые выходные данные полностью указаны, организация программ здесь не обсуждается, так как она будет сильно зависеть от конкретных условий применения и требований, предъявляемых потребителем.

ПРОГРАММА 1.

ИДЕНТИФИКАЦИЯ СТОХАСТИЧЕСКОЙ МОДЕЛИ

1.1. Общее описание

Программа вводит временной ряд, соответствующим образом трансформирует каждое наблюдение, производит разностные операции и затем вычисляет следующие величины и функции:

- 1) среднее значение \bar{w} и дисперсию s_w^2 ,
- 2) автоковариационную функцию $(acvf)c_k$,
- 3) автокорреляционную функцию $(acf)r_k$,
- 4) частную автокорреляционную функцию $(pacf)\hat{f}_n$.

1.2. Входные данные

Минимальной входной информацией, необходимой для вычислений, являются

значения временного ряда	$\{z_t\}, \quad t = 1, 2, \dots, N,$
число наблюдений	$N,$
порядок операции взятия несезонных разностей	$d \geq 0$
порядок операции взятия сезонных разностей	$D \geq 0,$
период сезонности	$s \geq 1, N - d - sD > 1,$
максимальная задержка $acvf, acf$	$K,$
максимальная задержка $pacf$	$L \leq K,$
параметры преобразования	$\lambda, m.$

1.3. Вычисления

Преобразование и взятие разностей. Для каждого набора значений (d, D, λ, m) ряд преобразуется следующим образом:

$$z'_t = \begin{cases} (z_t + m)^\lambda & \lambda \neq 0, \\ \ln(z_t + m) & \lambda = 0, \end{cases}$$

где, если $\lambda \neq 0$, параметр m выбран так, что $z_t + m$ положительно для всех t , и, если $\lambda = 1$, m приравняется нулю, так что z остаются неизменными. Затем находится разностный ряд, такой, что

$$w_t = \nabla^d \nabla_s^D z'_t, \quad t = 1, 2, \dots, n,$$

где

$$\nabla w_t = w_t - w_{t-1},$$

$$\nabla_s w_t = w_t - w_{t-s},$$

$$n = N - d - sD.$$

Далее вычисляются следующие величины.

Среднее значение и дисперсия

$$\bar{w} = \frac{1}{n} \sum_{t=1}^n w_t,$$

$$s_w^2 = c_0,$$

c_0 определено ниже.

Автоковариационная функция

$$c_k = \frac{1}{n} \sum_{t=1}^{n-k} (w_t - \bar{w})(w_{t+k} - \bar{w}),$$

где $k = 0, 1, \dots, K$.

Автокорреляционная функция

$$r_k = \frac{c_k}{c_0},$$

где $k = 0, 1, \dots, K$.

Частная автокорреляционная функция

$$\hat{\phi}_{ll} = \begin{cases} r_1, & l = 1 \\ \frac{r_l - \sum_{j=1}^{l-1} \hat{\phi}_{l-1,j} r_{l-j}}{1 - \sum_{j=1}^{l-1} \hat{\phi}_{l-1,j} r_j}, & l = 2, 3, \dots, L, \end{cases}$$

где

$$\hat{\phi}_{lj} = \hat{\phi}_{l-1,j} - \hat{\phi}_{ll} \hat{\phi}_{l-1,l-j}, \quad j = 1, 2, \dots, l-1.$$

Частная автокорреляционная функция может быть оценена более точно при помощи программы 3 путем вычисления оценки наименьших квадратов последнего из параметров авторегрессии $\hat{\phi}_l$ в последовательности процессов авторегрессии $AP(l)$, $l = 1, 2, \dots$, подгоняемых к данным.

1.4. Выход

Выходная информация должна включать все входные данные, а также

параметры преобразования	$\lambda, m,$
значения разностного преобразованного ряда	$\{w_t\}, t = 1, 2, \dots, n,$
число значений w_t	$n,$
среднее значение ряда w	$\bar{w},$
дисперсию ряда w	$s_w^2,$
$acvf$ ряда w	$c_k, k = 0, 1, \dots, K,$
acf ряда w	$r_k, k = 0, 1, \dots, K,$
$pacf$ ряда w	$\hat{\phi}_l, l = 1, 2, \dots, L.$

1.5. Дополнения

В целях обеспечения наибольшей общности следует ввести управляющие параметры, определяющие режим работы программы.

ПРОГРАММА 2. ПРЕДВАРИТЕЛЬНОЕ ОЦЕНИВАНИЕ СТОХАСТИЧЕСКОЙ МОДЕЛИ

2.1. Общее описание

Пользуясь полученными по программе 1 автоковариациями, эта программа вычисляет для несезонной модели

$$w_t = \phi_1 w_{t-1} + \dots + \phi_p w_{t-p} + \theta_0 + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q},$$

идентифицированной по надлежащим образом преобразованному разностному ряду, следующие величины:

- 1) начальные оценки $\hat{\phi}_0 = (\hat{\phi}_{10}, \hat{\phi}_{20}, \dots, \hat{\phi}_{p0})$ несезонных параметров авторегрессии,
- 2) начальные оценки $\hat{\theta}_0 = (\hat{\theta}_{10}, \hat{\theta}_{20}, \dots, \hat{\theta}_{q0})$ несезонных параметров скользящего среднего,
- 3) начальную оценку $\hat{\theta}_0$ общего постоянного члена,
- 4) начальную оценку $\hat{\sigma}_a^2$ дисперсии белого шума.

2.2. Входные данные

Минимальная входная информация включает

$$\begin{array}{ll} \text{число параметров авторегрессии} & p \geq 0, \\ \text{число параметров скользящего среднего} & q \geq 0, \quad p + q > 0, \\ \text{среднее значение ряда } \omega_t & \bar{\omega}, \\ \text{автоковариации ряда } \omega_t & c_k, \quad k = 0, 1, \dots, K, \\ & K \geq p + q \end{array}$$

где c_k и $\bar{\omega}$ вычислены по программе 1.

2.3. Вычисления

Оценки параметров авторегрессии $\hat{\phi}$. Если $p > 0$, решим систему из p линейных уравнений

$$A\phi_0 = x,$$

где

$$A_{ij} = c_{|q+i-j|},$$

$$x_i = c_{q+i},$$

$$i, j = 1, 2, \dots, p.$$

Оценки параметров скользящего среднего $\hat{\theta}$. 1) По известным автоковариациям c_k ряда ω_t вычисляется следующая модифицированная последовательность ковариаций c'_j :

$$c'_j = \begin{cases} \sum_{i=0}^p \sum_{k=0}^p \hat{\phi}_{i0} \hat{\phi}_{k0} c_{|j+i-k|}, & p > 0 \quad (\hat{\phi}_{00} = -1), \\ c_j, & p = 0, \end{cases}$$

где $j = 0, 1, \dots, q$.

2) При помощи алгоритма Ньютона — Рафсона

$$\tau^{i+1} = \tau^i - h,$$

где

$$T^i h = f^i,$$

вычисляется вектор τ^{i+1} на $(i+1)$ -й итерации по его значению τ^i на i -й итерации, где

$$\tau = (\tau_0, \tau_1, \dots, \tau_q),$$

$$f_j = \sum_{i=0}^{q-j} \tau_i \tau_{i+j} - c'_j,$$

$$f' = (f_0, f_1, \dots, f_q),$$

$$T = \begin{bmatrix} \tau_0 & \tau_1 & \dots & \dots & \dots & \tau_q \\ \tau_1 & \tau_2 & \dots & \dots & \dots & \tau_q \\ \vdots & \vdots & \dots & \dots & \dots & \vdots \\ \vdots & \vdots & \dots & \dots & 0 & \vdots \\ \vdots & \vdots & \dots & \dots & 0 & \vdots \\ \tau_q & \tau_q & \dots & \dots & \dots & \tau_q \end{bmatrix} + \begin{bmatrix} \tau_0 & \tau_1 & \dots & \dots & \dots & \tau_q \\ & \tau_0 & \tau_1 & \dots & \dots & \tau_{q-1} \\ & & \vdots & \dots & \dots & \vdots \\ & & & \vdots & \dots & \vdots \\ & & & & \vdots & \vdots \\ & & & & & \tau_0 \end{bmatrix}$$

с начальными значениями

$$\tau_0 = \sqrt{c_0}, \quad \tau_1 = \tau_2 = \dots = \tau_q = 0.$$

3) Когда $|f_j| < \epsilon$, $j = 0, 1, \dots, q$, для некоторого заданного значения ϵ , итеративная процедура считается завершенной, и по последнему значению τ находятся оценки параметров согласно формуле

$$\hat{\theta}_{j0} = -\tau_j / \tau_0, \quad j = 1, 2, \dots, q.$$

Оценка общей константы $\hat{\theta}_{00}$

$$\hat{\theta}_{00} = \begin{cases} \bar{w} \left(1 - \sum_{i=1}^p \hat{\phi}_{i0} \right), & p > 0, \\ \bar{w}, & p = 0. \end{cases}$$

Оценка дисперсии белого шума $\hat{\sigma}_a^2$

$$\hat{\sigma}_a^2 = \begin{cases} \tau_0^2, & q > 0, \\ c_0 - \sum_{i=1}^p \hat{\phi}_i c_i, & q = 0, \end{cases}$$

где τ_0 вычислено согласно пункту 2.

2.4. Выход

Выходная информация должна включать все входные данные, а также

начальные оценки несезонных

параметров авторегрессии

$$\hat{\phi}_0 = (\hat{\phi}_{10}, \hat{\phi}_{20}, \dots, \hat{\phi}_{p0}),$$

начальные оценки несезонных

параметров скользящего среднего

$$\hat{\theta}_0 = (\hat{\theta}_{10}, \hat{\theta}_{20}, \dots, \hat{\theta}_{q0}),$$

начальную оценку общей константы

$$\hat{\theta}_{00},$$

начальную оценку дисперсии

белого шума

$$\hat{\sigma}_a^2.$$

Вычисления и выдача данных могут быть проведены для нескольких порядков взятия разностей и для всех моделей (p, q) от $(1, 0)$ и $(0, 1)$ до (p, q) включительно.

2.5. Дополнения

В целях большей общности следует ввести управляющие параметры, определяющие режим работы программы. На практике вычислительные алгоритмы программ 1 и 2 могут быть объединены в единую программу, позволяющую, если необходимо, идентифицировать и предварительно оценивать параметры стохастической модели.

ПРОГРАММА 3.

ОЦЕНИВАНИЕ СТОХАСТИЧЕСКОЙ МОДЕЛИ

3.1. Общее описание

Программа вводит начальные оценки параметров. Они могут быть получены при помощи программы 2 для несезонных моделей и по методике, описанной в гл. 9, для сезонных моделей. Программа вычисляет

- 1) оценки наименьших квадратов μ , ϕ , Φ , θ , Θ и σ_a^2 в сезонной модели

$$\phi(B)\Phi(B^S)(w_t - \mu) = \theta(B)\Theta(B^S)a_t,$$

где $w_t = \nabla^d \nabla_S^p z'_t$,

z'_t — преобразованный временной ряд, определенный в программе 1,

μ — среднее значение ряда w ;

- 2) стандартные ошибки оценок и оценку корреляционной матрицы ошибок;
- 3) автокорреляционную функцию остаточных ошибок, соответствующих оценкам наименьших квадратов, и связанную с ними χ^2 -статистику.

3.2. Входные параметры

Минимальная информация, необходимая для расчетов, включает

$\{z_t\}$, $N, d, D, s, K, \lambda, m'$,	определены в программе 1,
p, q ,	определены в программе 2,
число сезонных пара-	

метров авторегрессии	$P \geq 0$,
число сезонных параметров скользящего среднего	$Q \geq 0, P + Q > 0$,
управляющий параметр для среднего значения μ	$M = 1$ (μ включено в модель) $M = 0$ (μ не включено в модель),
начальную оценку μ	\bar{w} , если $M = 1$, 0, если $M = 0$,
начальные оценки не-сезонных параметров авторегрессии	$\hat{\phi} = (\hat{\phi}_{10}, \hat{\phi}_{20}, \dots, \hat{\phi}_{p0})$,
начальные оценки сезонных параметров авторегрессии	$\hat{\Phi} = (\hat{\Phi}_{10}, \hat{\Phi}_{20}, \dots, \hat{\Phi}_{p0})$,
начальные оценки не-сезонных параметров скользящего среднего	$\hat{\theta} = (\hat{\theta}_{10}, \hat{\theta}_{20}, \dots, \hat{\theta}_{q0})$,
начальные оценки сезонных параметров скользящего среднего	$\hat{\Theta} = (\hat{\Theta}_{10}, \hat{\Theta}_{20}, \dots, \hat{\Theta}_{q0})$,
максимально допустимое число итераций	I.

3.3. Вычисления

Вычисление суммы квадратов остаточных ошибок. 1) При наличии начальных w , «спрогнозированных назад», вычисление остаточных ошибок для заданного набора параметров проводится в два этапа:

$$\alpha_t = (w_t - \mu) - \sum_{i=1}^p \phi_i (w_{t-i} - \mu) + \sum_{j=1}^q \theta_j \alpha_{t-j},$$

$$a_t = \alpha_t - \sum_{i=1}^P \Phi_i \alpha_{t-is} + \sum_{j=1}^Q \Theta_j a_{t-js},$$

где мы воспользовались сокращенным обозначением $a_t = [a_t | \mu, \phi, \Phi, \theta, \Theta, w]$,

μ присутствует, только если $M = 1$,

$t = Q', Q' + 1, \dots, n$,

Q — отрицательное значение момента времени t ; для времен меньше Q значения прогнозируемых назад величин пренебрежимо малы,

2) Для того чтобы начать рекуррентные вычисления в прямом направлении в (1), используется процедура из программы 4 для прогнозирования назад значений $\omega_0, \omega_{-1}, \dots$ до таких значений времени, при которых $\omega_t - \hat{\mu}$ пренебрежимо мало.

3) Для данных значений параметров $(\mu, \hat{\phi}, \hat{\Phi}, \hat{\theta}, \hat{\Theta})$ сумма квадратов остаточных ошибок вычисляется как

$$S(\mu, \hat{\phi}, \hat{\Phi}, \hat{\theta}, \hat{\Theta}) = \sum_{t=Q}^n a_t^2.$$

Вычисление оценок наименьших квадратов. Значения параметров, минимизирующих сумму квадратов остаточных ошибок, получены методом оптимизации с ограничениями, предложенным Марквардтом [63]. Этот метод описан в конце программы. Входной параметр I определяет максимально допустимое число итераций.

Стандартные ошибки и корреляционная матрица. Оценка остаточной дисперсии получается из оценки суммы квадратов на последней итерации по формуле

$$\hat{\sigma}_a^2 = \frac{1}{n - p - q - P - Q - M} S(\hat{\mu}, \hat{\phi}, \hat{\Phi}, \hat{\theta}, \hat{\Theta}),$$

ковариационная матрица оценок V находится по формуле

$$V = \{V_{ij}\} = (X'X)^{-1} \hat{\sigma}_a^2,$$

где X — регрессионная матрица в линеаризованной модели, вычисленная на последней итерации в процедуре Марквардта.

Стандартные ошибки равны

$$s_i = \sqrt{V_{ii}}, \quad i = 1, 2, \dots, p + q + P + Q + M,$$

а элементы R_{ij} корреляционной матрицы вычисляются как

$$R_{ij} = V_{ij} / \sqrt{V_{ii} V_{jj}}.$$

Наконец, оценка $\hat{\theta}_0$ общей константы равна

$$\hat{\theta}_0 = \hat{\mu}G,$$

где

$$G = \left(1 - \sum_{i=1}^p \hat{\phi}_i\right) \left(1 - \sum_{j=1}^P \hat{\Phi}_j\right).$$

Диагностические проверки. Остаточные автокорреляции вычисляются по остаточным ошибкам \hat{a}_t , соответствующим оценкам наименьших квадратов, согласно формуле

$$r_{aa}(k) = c_{aa}(k) / c_{aa}(0),$$

где

$$c_{aa}(k) = \frac{1}{n} \sum_{t=1}^{n-k} (\hat{a}_t - \bar{a})(\hat{a}_{t+k} - \bar{a}),$$

$$\bar{a} = \frac{1}{n} \sum_{t=1}^n \hat{a}_t$$

и

$$k = 0, 1, \dots, K.$$

Наконец, статистика χ^2 вычисляется как

$$\chi^2 = n \sum_{k=1}^K r_{aa}^2(k)$$

и сравнивается с χ^2 -распределением с

$$\nu = K - M - p - q - P - Q$$

степенями свободы.

3.4. Выход

Выходная информация должна включать все входные данные и

$$\hat{\mu}, \hat{\phi}, \hat{\Phi}, \hat{\theta}, \hat{\Theta}, S(\hat{\mu}, \hat{\phi}, \hat{\Phi}, \hat{\theta}, \hat{\Theta})$$

на каждой итерации, а также следующую информацию на последней итерации (т.е. когда процедура привела к сходящемуся результату, или поиск оказался неудачным, или число выполненных итераций достигло заданного максимального числа I, а сходимость не достигнута):

остаточные ошибки, соответствующие оценкам наименьших квадратов
оценку остаточной дисперсии (белого шума)
ковариационную матрицу оценок
корреляционную матрицу оценок
оценку общей константы
остаточные автокорреляции
 χ^2 -статистику

$$\{\hat{a}_t\}, \quad t = 1, 2, \dots, n,$$

$$\hat{\sigma}_a^2,$$

$$\mathbf{V},$$

$$\mathbf{s} = (s_1, s_2, \dots, s_{M+p+q+P+Q}),$$

$$\mathbf{R},$$

$$\hat{\theta}_0,$$

$$r_{aa}(k), \quad k = 1, 2, \dots, K,$$

$$\chi^2.$$

3.5. Дополнения

Для достижения большей общности следует ввести управляющие параметры, определяющие режим работы программы. Можно включить в программу табулирование $S(\hat{\mu}, \hat{\phi}, \hat{\Phi}, \hat{\theta}, \hat{\Theta})$ в окрестности минимума путем расчета этой функции по сетке с центром, соответствующим оценкам наименьших квадратов.

АЛГОРИТМ МАРКВАРДА ДЛЯ НЕЛИНЕЙНОГО МЕТОДА НАИМЕНЬШИХ КВАДРАТОВ *)

1. Задаваемые параметры

Обозначим через $\beta = (\beta_1, \beta_2, \dots, \beta_k)$ все параметры модели, т.е. $\hat{\beta} = (\mu, \phi, \Phi, \theta, \Theta)$; необходимо задать начальные значения β_0 , а также параметры π и F_2 , ограничивающие поиск, и параметр сходимости ϵ . Во время поиска необходимо оценивать значения $a_t = [a_t | \hat{\beta}, w]$ и производные

$$x_{i,t} = - \frac{\partial a_t}{\partial \beta_i}$$

на каждом шаге итерации.

2. Вычисление производных

Производные находятся по остаточным ошибкам, вычисленным в соответствии с разд. 3.3 программы 3, согласно формуле

$$x_{it} = \{a_t(\beta_{1,0}, \dots, \beta_{i,0}, \dots, \beta_{k,0}) - \\ - a_t(\beta_{1,0}, \dots, \beta_{i,0} + \delta_i, \dots, \beta_{k,0})\} / \delta_i.$$

3. Итерация

Этап 1. По известным $a_t, x_{i,t}$, найденным для текущих значений параметров, находятся следующие величины:

1) матрица размером $k \times k$

$$A = \{A_{ij}\},$$

где

$$A_{ij} = \sum_{t=Q'}^n x_{i,t} x_{j,t};$$

2) вектор g с элементами g_1, g_2, \dots, g_k , где

$$g_i = \sum_{t=Q'}^n x_{i,t} a_t;$$

*) Алгоритм слегка модифицирован Г. Вилсоном.

3) нормирующие величины

$$D_i = \sqrt{A_{ii}}.$$

Этап 2. Модифицированные (введением нормирующих множителей и наложением ограничений) линеаризованные уравнения

$$A^*h^* = g^*$$

конструируются согласно формулам

$$A_{ij}^* = A_{ij}/D_i D_j, \quad i \neq j,$$

$$A_{ii}^* = 1 + \pi,$$

$$g_i^* = g_i/D_i.$$

Уравнения решаются относительно h^* ; этот вектор денормируется для получения поправок параметров

$$h_i = h_i^*/D_i.$$

Вычисляются новые значения параметров

$$\beta = \beta_0 + h$$

и оценивается сумма квадратов остаточных ошибок $S(\beta)$.

Этап 3. 1) Если $S(\beta) < S(\beta_0)$, исследуются поправки параметров h . Если они все меньше ϵ , достигнута сходимость, и матрица A^{-1} размером $k \times k$ используется для вычисления ковариационной матрицы оценок согласно программе 3. В противном случае β_0 принимает значение β , π уменьшается в F_2 раз и продолжают вычисления на этапе 1.

2) Если $S(\beta) > S(\beta_0)$, ограничивающий параметр π увеличивается в F_2 раз и вычисления продолжают в соответствии с этапом 2. Почти во всех случаях удастся найти наименьшую сумму квадратов. Однако π имеет верхнее граничное значение, и если оно превышено, поиск прекращается. Когда поиск закончился успешно, либо в соответствии с критерием 1 этапа 3, либо (предположительно) после заданного числа итераций, по программе 3 вычисляются остаточная дисперсия и ковариационная матрица оценок.

ПРОГРАММА 4. ПРОГНОЗИРОВАНИЕ С ПОМОЩЬЮ СТОХАСТИЧЕСКОЙ МОДЕЛИ

4.1. Общее описание

Программа вводит оценки наименьших квадратов ($\hat{\theta}_0$, $\hat{\phi}$, $\hat{\Phi}$, $\hat{\theta}$, $\hat{\Theta}$) параметров общей модели, подогнанной при помощи программы 2, вместе со значениями временного ряда.

Программа вычисляет

1) прогнозирующую функцию $\hat{z}_{N-b}(l)$, $l = 1, 2, \dots, L$ для каждого момента времени, отстоящего на b временных единиц $b = 0, 1, \dots, B$ от конца ряда. Прогнозы вычисляются вначале для преобразованного временного ряда z'_t , а затем переводятся в прогнозы для исходного ряда z_t ;

2) верхний и нижний вероятностные пределы $z_t(+)$ и $z_t(-)$; вероятность для некоторого будущего значения z_t находиться между этими пределами $P\{z_t(-) < z_t < z_t(+)\}$ равна заданной величине α .

4.2. Входные параметры

Минимальная информация, необходимая для выполнения расчетов, включает следующие величины:

$\{z_t\}$, N , d , D , s , K	определены в программе 1,
p , P , q , Q , M	определены в программе 2,
управляющий параметр для постоянной составляющей	$M = 1$ (постоянная составляющая включена), $M = 0$ (постоянная составляющая не включена),
максимальное упреждение прогнозирующей функции	L ,
оценки наименьших квадратов постоянной составляющей	$\hat{\theta}_0$,
оценки наименьших квадратов несезонных параметров авторегрессии	$\hat{\phi}_1, \hat{\phi}_2, \dots, \hat{\phi}_p$,
оценки наименьших квадратов сезонных параметров авторегрессии	$\hat{\Phi}_1, \hat{\Phi}_2, \dots, \hat{\Phi}_p$,
оценки наименьших квадратов несезонных параметров скользящего среднего	$\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_q$,
оценки наименьших квадратов сезонных параметров скользящего среднего	$\hat{\Theta}_1, \hat{\Theta}_2, \dots, \hat{\Theta}_Q$,
оценка наименьших квадратов остаточной дисперсии	$\hat{\sigma}_a^2$.

4.3. Вычисления

Разъединение операторов. Для целей прогноза сезонная модель представляется в виде

$$z_t = \Phi_1^* z_{t-1} + \dots + \Phi_p^* z_{t-p} + \theta_0 + a_t - \Theta_1^* a_{t-1} - \dots - \Theta_Q^* a_{t-Q},$$

где неизвестные параметры заменены их оценками. Для получения Φ_i^* сначала находятся параметры $\Phi'_1, \dots, \Phi'_{p'}$ оператора

$$\Phi'(B) = \Phi(B^s) \phi(B) = 1 - \Phi'_1 B - \dots - \Phi'_{p'} B^{p'},$$

где $p' = p + sP$.

Для этого используется формула

$$\Phi'_k = - \sum_{i+s=P} \sum_j \Phi_i \phi_j,$$

двойная сумма находится суммированием по всем $j = 0, \dots, p_j$ и всем $i = 0, \dots, P$. Константы Φ_0 и ϕ_0 принимаются равными -1 . Далее, параметры $\Phi_1^*, \dots, \Phi_p^*$ оператора

$$\Phi^*(B) = \nabla^d \nabla_s^D \Phi'(B) = 1 - \Phi_1^* B - \dots - \Phi_{p^*}^* B^{p^*}$$

с $p^* = p' + d + sD$ получаются в результате следующей трехступенчатой процедуры:

1)

$$\Phi''_j = \begin{cases} \Phi'_j, & 0 \leq j < s, \\ \Phi'_j - \Phi'_{j-s}, & s \leq j \leq p', \\ -\Phi'_{j-s}, & p' < j \leq p' + s. \end{cases}$$

2) Этап 1 повторяется D раз с заменой Φ'_j на Φ''_j , p' на $p'' = p' + s$ на каждой итерации для получения коэффициентов $\nabla_s^D \Phi'(B)$.

3) Аналогично применяем d раз этап 1 с $s = 1$ к значениям, полученным на этапе 2, для нахождения окончательных оценок коэффициентов $\Phi_1^*, \dots, \Phi_{p^*}^*$.

Операторы скользящего среднего Θ_k^* находятся таким же способом, как и Φ' , по формуле

$$\Theta_k^* = - \sum_{i+s=P} \sum_j \Theta_i \theta_j,$$

где $i = 0, 1, \dots, Q$; $j = 0, 1, \dots, q$, $\Theta_0 = -1$, $\theta_0 = -1$.

Генерирование прогнозов. 1) Для трансформированного ряда z'_t прогнозы $z'_{N-b}(l)$ вычисляются по формуле

$$z'_{N-b}(l) = \theta_0 + \sum_{i=1}^{p+sP+d+sD} \Phi_i^* [z'_{N-b-l+i}] - \sum_{j=1}^{q+SQ} \Theta_j^* [a_{N-b-l+i}],$$

где

$$[z'_{N-b-l+i}] = \begin{cases} z'_{N-b}(l-i), & l > i, \\ z'_{N-b-i+l}, & l \leq i, \end{cases}$$

$$[a_{N-b-l+i}] = \begin{cases} 0, & l > j, \\ z'_{N-b-j+l} - z'_{N-b-j+l-1}(1), & l \leq j, \end{cases}$$

$l = 1, 2, \dots, L$ и z'_t определены в программе 1. Прогнозы находятся для каждого отодвигаемого назад момента времени $b = 0, 1, \dots, B$.

2) Прогнозы $\hat{z}_{N-b}(l)$ для исходного ряда находятся по формуле

$$\hat{z}_{N-b}(l) = \begin{cases} z'_{N-b}(l), & \lambda = 1, \\ \{z'_{N-b}(l)\}^{1/\lambda} - m, & \lambda \neq 0, \\ \exp \{z'_{N-b}(l)\} - m, & \lambda = 0. \end{cases}$$

Точность прогнозов. 1) Для прогнозов преобразованного ряда верхний и нижний вероятностные пределы равны

$$z'_{N-b+l}(\pm) = z'_{N-b}(l) \pm u \sqrt{V(l)},$$

где $u = 0,68; 1,65; 1,96$ или $2,58$ в зависимости от того, лежит ли будущее значение между этими пределами с вероятностью в интервале $0,50; 0,90; 0,95$ или $0,99$ соответственно. Функция дисперсий равна

$$V(l) = \hat{\sigma}_a^2 \sum_{j=0}^{l-1} \psi_j^2,$$

где

$$\psi_j = \begin{cases} 1, & j = 0, \\ \sum_{i=1}^j \Phi_i^* \psi_{j-i} - \Theta_j^*, & j \geq 1 \end{cases}$$

для $j = 1, 2, \dots, L$ и

$$\Phi_i^* = 0, \quad i > p + sP + d + sD,$$

$$\Theta_j^* = 0, \quad j > q + sQ.$$

2) Для прогнозов исходного ряда верхний и нижний пределы равны

$$z_{N-b+l}(\pm) = \begin{cases} z'_{N-b+l}(\pm), & \lambda = 1, \\ \{z'_{N-b+l}(\pm)\}^{1/\lambda} - m, & \lambda \neq 0, \\ \exp\{z'_{N-b+l}(\pm)\} - m, & \lambda = 0. \end{cases}$$

4.4. Выход

Выходная информация должна содержать все входные данные, а также

параметры обобщенного оператора авторегрессии

Φ^* ,

параметры обобщенного оператора скользящего среднего

Θ^* ,

веса ψ

$\psi_j, j = 1, 2, \dots, L,$

прогнозы и их вероятностные пределы для данного вероятностного уровня

$z_{N-b+l}(+),$

$\hat{z}_{N-b}(l),$

$z_{N-b+l}(-)$

для $l = 1, 2, \dots, L$ и выбранных начальных моментов, заданных значениями $b = 0, 1, \dots, B$.

4.5. Дополнения

Для целей большей общности полезно ввести управляющие параметры, определяющие режим работы программы. Практически удобно получать прогнозы в виде матрицы, в которой каждый элемент какого-либо столбца соответствует прогнозируемому значению z_t .

СБОРНИК ТАБЛИЦ И ДИАГРАММ

Таблица А. Таблица связи автокорреляции ρ_1 с параметром θ процесса скользящего среднего первого порядка.

Диаграмма В. Диаграмма связи автокорреляций ρ_1 и ρ_2 с параметрами ϕ_1 и ϕ_2 процесса авторегрессии второго порядка.

Диаграмма С. Диаграмма связи автокорреляций ρ_1 и ρ_2 с параметрами θ_1 и θ_2 процесса скользящего среднего второго порядка,

Диаграмма D. Диаграмма связи автокорреляций ρ_1 и ρ_2 с параметрами ϕ и θ смешанного процесса авторегрессии — скользящего среднего первого порядка.

Таблица E. Уровни значимости и ординаты стандартного нормального распределения.

Таблица F. Уровни значимости χ^2 -распределения.

Таблица G. Уровни значимости t -распределения.

Диаграммы B, C и D (в слегка измененном виде) приводятся с согласия авторов работы [33]. Таблицы E, F и G (в уплотненной форме) взяты из «Таблиц биометрики для статистиков», т. I, с разрешения правления журнала «Биометрика».

Таблица A

Таблица связи ρ_1 и θ для процесса скользящего среднего первого порядка

θ	ρ_1	θ	ρ_1
0,00	0,000	0,00	0,000
0,05	-0,050	-0,05	0,050
0,10	-0,099	-0,10	0,099
0,15	-0,147	-0,15	0,147
0,20	-0,192	-0,20	0,192
0,25	-0,235	-0,25	0,235
0,30	-0,275	-0,30	0,275
0,35	-0,315	-0,35	0,315
0,40	-0,349	-0,40	0,349
0,45	-0,374	-0,45	0,374
0,50	-0,400	-0,50	0,400
0,55	-0,422	-0,55	0,422
0,60	-0,441	-0,60	0,441
0,65	-0,457	-0,65	0,457
0,70	-0,468	-0,70	0,468
0,75	-0,480	-0,75	0,480
0,80	-0,488	-0,80	0,488
0,85	-0,493	-0,85	0,493
0,90	-0,497	-0,90	0,497
0,95	-0,499	-0,95	0,499
1,00	-0,500	-1,00	0,500

Таблица может быть использована для получения первичных оценок параметров $(0, d, 1)$ процесса $w_t = (1 - \theta B) a_t$, где $w_t = \nabla d z_t$, заменой ρ_1 на $r_1(w)$.

Диаграмма В

Диаграмма связи ρ_1 и ρ_2 с ϕ_1 и ϕ_2 для процесса авторегрессии второго порядка

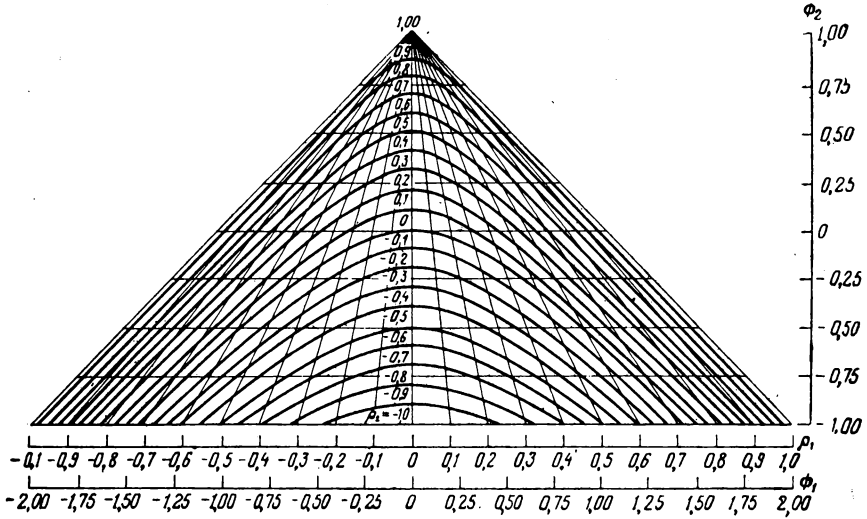


Диаграмма может использоваться для получения первичных оценок параметров процесса $(2, d, 0)$ $(1 - \phi_1 B - \phi_2 B^2) \omega_t = a_t$, где $\omega_t = \nabla^d z_t$, заменой ρ_1 и ρ_2 на $r_1(\omega)$, $r_2(\omega)$.

Диаграмма С

Диаграмма связи автокорреляций ρ_1 и ρ_2 с параметрами θ_1 и θ_2 процесса скользящего среднего второго порядка

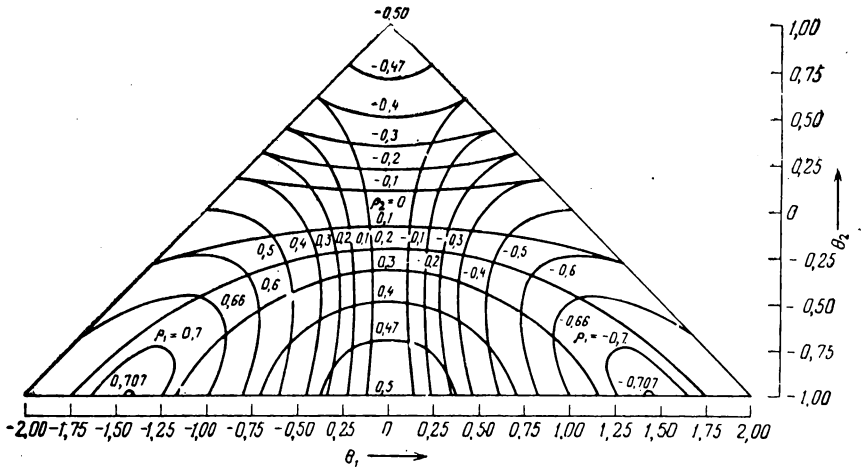


Диаграмма может использоваться для получения первичных оценок параметров процесса $(0, d, 2)$ $\omega_t = (1 - \theta_1 B - \theta_2 B^2) a_t$, где $\omega_t = \nabla^d z_t$, заменой ρ_1 и ρ_2 на $r_1(\omega)$ и $r_2(\omega)$.

Диаграмма D

Диаграмма связи автокорреляций ρ_1 и ρ_2 с параметрами ϕ и θ смешанного процесса авторегрессии — скользящего среднего первого порядка

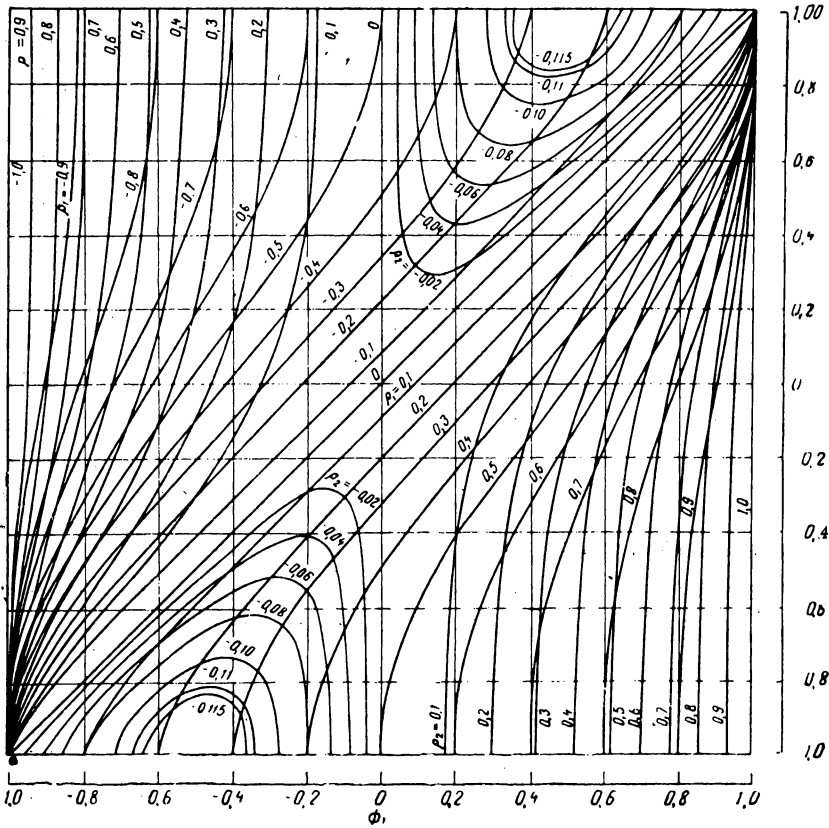


Диаграмма может быть использована для получения первичных оценок параметров процесса $(1, d, 1)$ $(1 - \phi B)w_t = (1 - \theta B)a_t$, где $w_t = \nabla^d z_t$, заменой ρ_1 и ρ_2 на $r_1(w)$ и $r_2(w)$.

Таблица Е
Уровни значимости и ординаты стандартного нормального распределения

u_e	ε	$p(u_e)$	u_e	ε	$q(u_e)$
0,0	0,500	0,3989	1,6	0,055	0,1109
0,1	0,460	0,3969	1,7	0,045	0,0940
0,2	0,421	0,3910	1,8	0,036	0,0790
0,3	0,382	0,3814	1,9	0,029	0,0656
0,4	0,345	0,3683	2,0	0,023	0,0540
0,5	0,309	0,3521	2,1	0,018	0,0369
0,6	0,274	0,3322	2,2	0,014	0,0355
0,7	0,242	0,3123	2,3	0,011	0,0283
0,8	0,212	0,2897	2,4	0,008	0,0224
0,9	0,184	0,2661	2,5	0,006	0,0175
1,0	0,159	0,2420	2,6	0,005	0,0136
1,1	0,136	0,2179	2,7	0,003	0,0104
1,2	0,115	0,1942	2,8	0,003	0,0079
1,3	0,097	0,1714	2,9	0,002	0,0059
1,4	0,081	0,1497	3,0	0,001	0,0044
1,5	0,067	0,1295			

* В таблице приведены значения u_e , для которых $P\{u > u_e\} = \varepsilon$, а также значения плотности вероятности.

Уровни значимости

p	ε						
	0,995	0,99	0,975	0,95	0,9	0,75	0,5
1	—	—	—	—	0,016	0,102	0,455
2	0,010	0,020	0,051	0,103	0,211	0,575	1,39
3	0,072	0,115	0,216	0,352	0,584	1,21	2,37
4	0,207	0,297	0,484	0,711	1,06	1,92	3,36
5	0,412	0,554	0,831	1,15	1,61	2,67	4,35
6	0,676	0,872	1,24	1,64	2,20	3,45	5,35
7	0,989	1,24	1,69	2,17	2,83	4,25	6,35
8	1,34	1,65	2,18	2,73	3,49	5,07	7,34
9	1,73	2,09	2,70	3,33	4,17	5,90	8,34
10	2,16	2,56	3,25	3,94	4,87	6,74	9,34
11	2,60	3,05	3,82	4,57	5,58	7,58	10,3
12	3,07	3,57	4,40	5,23	6,30	8,44	11,3
13	3,57	4,11	5,01	5,89	7,04	9,30	12,3
14	4,07	4,66	5,63	6,57	7,79	10,2	13,3
15	4,60	5,23	6,26	7,26	8,55	11,0	14,3
16	5,14	5,81	6,91	7,96	9,31	11,9	15,3
17	5,70	6,41	7,56	8,67	10,1	12,8	16,3
18	6,26	7,01	8,23	9,39	10,9	13,7	17,3
19	6,84	7,63	8,91	10,1	11,7	14,6	18,3
20	7,43	8,26	9,59	10,9	12,4	15,5	19,3
21	8,03	8,90	10,3	11,6	13,2	16,3	20,3
22	8,64	9,54	11,0	12,3	14,0	17,2	21,3
23	9,26	10,2	11,7	13,1	14,8	18,1	22,3
24	9,89	10,9	12,4	13,8	15,7	19,0	23,3
25	10,5	11,5	13,1	14,6	16,5	19,9	24,3
26	11,2	12,2	13,8	15,4	17,3	20,8	25,3
27	11,8	12,9	14,6	16,2	18,1	21,7	26,3
28	12,5	13,6	15,3	16,9	18,9	22,7	27,3
29	13,1	14,3	16,0	17,7	19,8	23,6	28,3
30	13,8	15,0	16,8	18,5	20,6	24,5	29,3

Таблица таких $\chi^2(p)$, что $P\{\chi^2(p) > \chi^2_\varepsilon(p)\} = \varepsilon$, где p — число степеней свободы,

χ^2 -распределения

ε							p
0,25	0,1	0,05	0,025	0,01	0,005	0,001	
1,32	2,71	3,84	5,02	6,63	7,88	10,8	1
2,77	4,61	5,99	7,38	9,21	10,6	13,8	2
4,11	6,25	7,81	9,35	11,3	12,8	16,3	3
5,39	7,78	9,49	11,1	13,3	14,9	18,5	4
6,63	9,24	11,1	12,8	15,1	16,7	20,5	5
7,84	10,6	12,6	14,4	16,8	18,5	22,5	6
9,04	12,0	14,1	16,0	18,5	20,3	24,3	7
10,2	13,4	15,5	17,5	20,1	22,0	26,1	8
11,4	14,7	16,9	19,0	21,7	23,6	27,9	9
12,5	16,0	18,3	20,5	23,2	25,2	29,6	10
13,7	17,3	19,7	21,9	24,7	26,8	31,3	11
14,8	18,5	21,0	23,3	26,2	28,3	32,9	12
16,0	19,8	22,4	24,7	27,7	29,8	34,5	13
17,1	21,1	23,7	26,1	29,1	31,3	36,1	14
18,2	22,3	25,0	27,5	30,6	32,8	37,7	15
19,4	23,5	26,3	28,8	32,0	34,3	39,3	16
20,5	24,8	27,6	30,2	33,4	35,7	40,8	17
21,6	26,0	28,9	31,5	34,8	37,2	42,3	18
22,7	27,2	30,1	32,9	36,2	38,6	43,8	19
23,8	28,4	31,4	34,2	37,6	40,0	45,3	20
24,9	29,6	32,7	35,5	38,9	41,4	46,8	21
26,0	30,8	33,9	36,8	40,3	42,8	48,3	22
27,1	32,0	35,2	38,1	41,6	44,2	49,7	23
28,2	33,2	36,4	39,4	43,0	45,6	51,2	24
29,3	34,4	37,7	40,6	44,3	46,9	52,6	25
30,4	35,6	38,9	41,9	45,6	48,3	54,1	26
31,5	36,7	40,1	43,2	47,0	49,6	55,5	27
32,6	37,9	41,3	44,5	48,3	51,0	56,9	28
33,7	39,1	42,6	45,7	49,6	52,3	58,3	29
34,8	40,3	43,8	47,0	50,9	53,7	59,7	30

Уровни значимости t-распределения

Таблица G

v	ε					
	0,25	0,10	0,05	0,025	0,01	0,005
1	1,00	3,08	6,31	12,71	31,82	63,66
2	0,82	1,89	2,92	4,30	6,96	9,92
3	0,76	1,64	2,35	3,18	4,54	5,84
4	0,74	1,53	2,13	2,78	3,75	4,60
5	0,73	1,48	2,02	2,57	3,36	4,03
6	0,72	1,44	1,94	2,45	3,14	3,71
7	0,71	1,42	1,90	2,36	3,00	3,50
8	0,71	1,40	1,86	2,31	2,90	3,36
9	0,70	1,38	1,83	2,26	2,82	3,25
10	0,70	1,37	1,81	2,23	2,76	3,17
11	0,70	1,36	1,80	2,20	2,72	3,11
12	0,70	1,36	1,78	2,18	2,68	3,06
13	0,69	1,35	1,77	2,16	2,65	3,01
14	0,69	1,34	1,76	2,14	2,62	3,00
15	0,69	1,34	1,75	2,13	2,60	2,95
16	0,69	1,34	1,75	2,12	2,58	2,92
17	0,69	1,33	1,74	2,11	2,57	2,90
18	0,69	1,33	1,73	2,10	2,55	2,88
19	0,69	1,33	1,73	2,09	2,54	2,86
20	0,69	1,33	1,72	2,09	2,53	2,84
30	0,68	1,31	1,70	2,04	2,46	2,75
40	0,68	1,30	1,68	2,02	2,42	2,70
60	0,68	1,30	1,67	2,00	2,39	2,66
120	0,68	1,29	1,66	1,98	2,36	2,62
∞	0,67	1,28	1,64	1,96	2,33	2,58

Таблица таких значений $t_{\varepsilon}(v)$, что $P\{t(v) > t_{\varepsilon}(v)\} = \varepsilon$, где v — число степеней свободы.

СБОРНИК ВРЕМЕННЫХ РЯДОВ, АНАЛИЗИРУЕМЫХ В КНИГЕ

- Ряд А. Отсчеты концентраций химического процесса, каждые два часа
 Ряд В. Биржевые цены акций IBM к закрытию биржи, ежедневно с 17 мая 1961 по 2 ноября 1962 г.
 Ряд В'. Биржевые цены акций IBM к закрытию биржи, ежедневно с 29 июня 1959 по 30 июня 1960 г.
 Ряд С. Отсчеты температур химического процесса, каждую минуту.
 Ряд D. Отсчеты вязкости химического процесса, каждый час
 Ряд Е. Числа солнечных пятен Вольфа, ежегодно
 Ряд F. Ряд 70 последовательных выходов партий продукта циклического химического процесса (дан в табл. 2.1 текста)
 Ряд G. Пассажирские перевозки на международных авиалиниях — месячные итоги (в тысячах пассажиров) с января 1949 по декабрь 1960.

Отсчеты концентраций химического процесса, каждые 2 часа *)

1	17,0	41	17,6	81	16,8	121	16,9	161	17,1
2	16,6	42	17,5	82	16,7	122	17,1	162	17,1
3	16,3	43	16,5	83	16,4	123	16,8	163	17,1
4	16,1	44	17,8	84	16,5	124	17,0	164	17,4
5	17,1	45	17,3	85	16,4	125	17,2	165	17,2
6	16,9	46	17,3	86	16,6	126	17,3	166	16,9
7	16,8	47	17,1	87	16,5	127	17,2	167	16,9
8	17,4	48	17,4	88	16,7	128	17,3	168	17,0
9	17,1	49	16,9	89	16,4	129	17,2	169	16,7
10	17,0	50	17,3	90	16,4	130	17,2	170	16,9
11	16,7	51	17,6	91	16,2	131	17,5	171	17,3
12	17,4	52	16,9	92	16,4	132	16,9	172	17,8
13	17,2	53	16,7	93	16,3	133	16,9	173	17,8
14	17,4	54	16,8	94	16,4	134	16,9	174	17,6
15	17,4	55	16,8	95	17,0	135	17,0	175	17,5
16	17,0	56	17,2	96	16,9	136	16,5	176	17,0
17	17,3	57	16,8	97	17,1	137	16,7	177	16,9
18	17,2	58	17,6	98	17,1	138	16,8	178	17,1
19	17,4	59	17,2	99	16,7	139	16,7	179	17,2
20	16,8	60	16,6	100	16,9	140	16,7	180	17,4
21	17,1	61	17,1	101	16,5	141	16,6	181	17,5
22	17,4	62	16,9	102	17,2	142	16,5	182	17,9
23	17,4	63	16,6	103	16,4	143	17,0	183	17,0
24	17,5	64	18,0	104	17,0	144	16,7	184	17,0
25	17,4	65	17,2	105	17,0	145	16,7	185	17,0
26	17,6	66	17,3	106	16,7	146	16,9	186	17,2
27	17,4	67	17,0	107	16,2	147	17,4	187	17,3
28	17,3	68	16,9	108	16,6	148	17,1	188	17,4
29	17,0	69	17,3	109	16,9	149	17,0	189	17,4
30	17,8	70	16,8	110	16,5	150	16,8	190	17,0
31	17,5	71	17,3	111	16,6	151	17,2	191	18,0
32	18,1	72	17,4	112	16,6	152	17,2	192	18,2
33	17,5	73	17,7	113	17,0	153	17,4	193	17,6
34	17,4	74	16,8	114	17,1	154	17,2	194	17,8
35	17,4	75	16,9	115	17,1	155	16,9	195	17,7
36	17,1	76	17,0	116	16,7	156	16,8	196	17,2
37	17,6	77	16,9	117	16,8	157	17,0	197	17,4
38	17,7	78	17,0	118	16,3	158	17,4		
39	17,4	79	16,6	119	16,6	159	17,2		
40	17,8	80	16,7	120	16,8	160	17,2		

**Ряд В. Биржевые цены акций IBM к закрытию биржи,
ежедневно с 17 мая 1961 г. по 2 ноября 1962 г. *)**

460	471	527	580	551	523	333	394	330
457	467	540	579	551	516	330	393	340
452	473	542	584	552	511	336	409	339
459	481	538	581	553	518	328	411	331
462	488	541	581	557	517	316	409	345
459	490	541	577	557	520	320	408	352
463	489	547	577	548	519	332	393	346
479	489	553	578	547	519	320	391	352
493	485	559	580	545	519	333	388	357
490	491	557	586	545	518	344	396	
492	492	557	583	539	513	339	387	
498	494	560	581	539	499	350	383	
499	499	571	576	535	485	351	388	
497	498	571	571	537	454	350	382	
496	500	569	575	535	462	345	384	
490	497	575	575	536	473	350	382	
489	494	580	573	537	482	359	383	
478	495	584	577	543	486	375	383	
487	500	585	582	548	475	379	388	
491	504	590	584	546	459	376	395	
487	513	599	579	547	451	382	392	
482	511	603	572	548	453	370	386	
479	514	599	577	549	446	365	383	
478	510	596	571	553	455	367	377	
479	509	585	560	553	452	372	364	
477	515	587	549	552	457	373	369	
479	519	585	556	551	449	363	355	
475	523	581	557	550	450	371	350	
479	519	583	563	553	435	369	353	
476	523	592	564	554	415	376	340	
476	531	592	567	551	398	387	350	
478	547	596	561	551	399	387	349	
479	551	596	559	545	361	376	358	
477	547	595	553	547	383	385	360	
476	541	598	553	547	393	385	360	
475	545	598	553	537	385	380	366	
475	549	595	547	239	360	373	359	
473	545	592	550	538	364	382	356	
474	549	592	544	533	365	377	355	
474	547	588	541	525	370	376	367	
474	543	582	532	513	374	379	357	

Продолжение ряда В

465	540	576	525	510	359	386	361
466	539	578	542	521	335	387	355
467	532	589	555	521	323	386	348
471	517	585	558	521	306	389	343
8888							

*) 369 наблюдений (читать сверху вниз)

Ряд В'. Биржевые цены акций IBM к закрытию биржи,
ежедневно с 29 июня 1959 г. по 30 июня 1960 г. *)

445	425	406	441	415	461
448	421	407	437	420	463
450	414	410	427	420	463
447	410	408	423	424	461
451	411	408	424	426	465
453	406	409	428	423	473
454	406	410	428	423	473
454	413	409	431	425	475
459	411	405	425	431	499
440	410	406	423	436	485
446	405	405	420	436	491
443	409	407	426	440	496
443	410	409	418	436	504
440	405	407	416	443	504
439	401	409	419	445	509
435	401	425	418	439	511
435	401	425	416	443	524
436	414	428	419	445	525
435	419	436	425	450	541
435	425	442	421	461	531
435	423	442	422	471	529
433	411	433	422	467	530
429	414	435	417	462	531
428	420	433	420	456	527
425	412	435	417	464	525
427	415	429	418	463	519
425	412	439	419	465	514
422	412	437	419	464	509
409	411	439	417	456	505
407	412	438	419	460	513
423	409	435	422	458	525

Продолжение ряда В'

422	407	433	423	453	519
417	408	437	422	453	519
421	415	437	421	449	522
424	413	444	421	447	522
414	413	441	419	453	
419	410	440	418	450	
429	405	441	421	459	
426	410	439	420	457	
425	412	439	413	453	
424	413	438	413	455	
425	411	437	408	453	
425	411	441	409	450	
424	409	442	415	456	

*) 255 наблюдений (читать сверху вниз)

Ряд С. Отсчеты температур химического процесса,
каждую минуту *)

26,6	19,6	24,4	21,1	24,4
27,0	19,6	24,4	20,9	24,2
27,1	19,6	24,4	20,8	24,2
27,1	19,6	24,4	20,8	24,1
27,1	19,6	24,5	20,8	24,1
27,1	19,7	24,5	20,8	24,0
26,9	19,9	24,4	20,9	24,0
26,8	20,0	24,3	20,8	24,0
26,7	20,1	24,2	20,8	23,9
26,4	20,2	24,2	20,7	23,8
26,0	20,3	24,0	20,7	23,8
25,8	20,6	23,9	20,8	23,7
25,6	21,6	23,7	20,9	23,7
25,2	21,9	23,6	21,2	23,6
25,0	21,7	23,5	21,4	23,7
24,6	21,3	23,5	21,7	23,6
24,2	21,2	23,5	21,8	23,6
24,0	21,4	23,5	21,9	23,6
23,7	21,7	23,5	22,2	23,5
23,4	22,2	23,7	22,5	23,5
23,1	23,0	23,8	22,8	23,4
22,9	23,8	23,8	23,1	23,3
22,8	24,6	23,9	23,4	23,3
22,7	25,1	23,9	23,8	23,3

Продолжение ряда С

22,6	25,6	23,8	24,1	23,4
22,4	25,8	23,7	24,6	23,4
22,2	26,1	23,6	24,9	23,3
22,0	26,3	23,4	24,9	23,2
21,8	26,3	23,2	25,1	23,3
21,4	26,2	23,0	25,0	23,3
20,9	26,0	22,8	25,0	23,2
20,3	25,8	22,6	25,0	23,1
19,7	25,6	22,4	25,0	22,9
19,4	25,4	22,0	24,9	22,8
19,3	25,2	21,6	24,8	22,6
19,2	24,9	21,3	24,7	22,4
19,1	24,7	21,2	24,6	22,2
19,0	24,5	21,2	24,5	21,8
18,9	24,4	21,1	24,5	21,3
18,9	24,4	21,0	24,5	20,8
19,2	24,4	20,9	24,5	20,2
19,3	24,4	21,0	24,5	19,7
19,3	24,4	21,0	24,5	19,3
19,4	24,3	21,1	24,5	19,1
19,5	24,4	21,2	24,4	19,0
				18,8

*) 226 наблюдений (читать сверху вниз)

Ряд D. Отсчеты вязкости химического процесса, каждый час*)

8,0	8,6	9,3	9,8	9,4	9,6	9,4
8,0	8,4	9,5	9,6	9,6	9,6	10,0
7,4	8,3	9,4	9,6	9,6	9,6	10,0
8,0	8,4	9,0	9,4	9,6	9,6	10,0
8,0	8,3	9,0	9,4	10,0	9,6	10,2
8,0	8,3	8,8	9,4	10,0	9,0	10,0
8,0	8,1	9,0	9,4	9,6	9,4	10,0
8,8	8,2	8,8	9,6	9,2	9,4	9,6
8,4	8,3	8,6	9,6	9,2	9,4	9,0
8,4	8,5	8,6	9,4	9,2	9,6	9,0
8,0	8,1	8,0	9,4	9,0	9,4	8,6
8,2	8,1	8,0	9,0	9,0	9,6	9,0
8,2	7,9	8,0	9,4	9,6	9,6	9,6
8,2	8,3	8,0	9,4	9,8	9,8	9,6

Продолжение ряда D

8,4	8,1	8,6	9,6	10,2	9,8	9,0
8,4	8,1	8,0	9,4	10,0	9,8	9,0
8,4	8,1	8,0	9,2	10,0	9,6	8,9
8,6	8,4	8,0	8,8	10,0	9,2	8,8
8,8	8,7	7,6	8,8	9,4	9,6	8,7
8,6	9,0	8,6	9,2	9,2	9,2	8,6
8,6	9,3	9,6	9,2	9,6	9,2	8,3
8,6	9,3	9,6	9,6	9,7	9,6	7,9
8,6	9,5	10,0	9,6	9,7	9,6	8,5
8,6	9,3	9,4	9,8	9,8	9,6	8,7
8,8	9,5	9,3	9,8	9,8	9,6	8,9
8,9	9,5	9,2	10,0	9,8	9,6	9,1
9,1	9,5	9,5	10,0	10,0	9,6	9,1
9,5	9,5	9,5	9,4	10,0	10,0	9,1
8,5	9,5	9,5	9,8	8,6	10,0	
8,4	9,5	9,9	8,8	9,0	10,4	
8,3	9,9	9,9	8,8	9,4	10,4	
8,2	9,5	9,5	8,8	9,4	9,8	
8,1	9,7	9,3	8,8	9,4	9,0	
8,3	9,1	9,5	9,6	9,4	9,6	
8,4	9,1	9,5	9,6	9,4	9,8	
8,7	8,9	9,1	9,6	9,6	9,6	
8,8	9,3	9,3	9,2	10,0	8,6	
8,8	9,1	9,5	9,2	10,0	8,0	
9,2	9,1	9,3	9,0	9,8	8,0	
9,6	9,3	9,1	9,0	9,0	8,0	
9,0	9,5	9,3	9,0	9,7	8,0	
8,8	9,3	9,1	9,4	9,6	8,4	
8,6	9,3	9,5	9,0	9,4	8,8	
8,6	9,3	9,4	9,0	9,2	8,4	
8,8	9,9	9,5	9,4	9,0	8,4	
8,8	9,7	9,6	9,4	9,4	9,0	
8,6	9,1	10,2	9,6	9,6	9,0	

*) 310 наблюдений (читать сверху вниз)

Ряд E. Числа солнечных пятен Вольфа, ежегодно *)

1770	101	1795	21	1820	16	1845	40
1771	82	1796	16	1821	7	1846	62
1772	66	1797	6	1822	4	1847	98
1773	35	1798	4	1823	2	1848	124
1774	31	1799	7	1824	8	1849	96
1775	7	1800	14	1825	17	1850	66

1776	20	1801	34	1826	36	1851	64
1777	92	1802	45	1827	50	1852	54
1778	154	1803	43	1828	62	1853	39
1779	125	1804	48	1829	67	1854	21
1780	85	1805	42	1830	71	1855	7
1781	68	1806	28	1831	48	1856	4
1782	38	1807	10	1832	28	1857	23
1783	23	1808	8	1833	8	1858	55
1784	10	1809	2	1834	13	1859	94
1785	24	1810	0	1835	57	1860	96
1786	83	1811	1	1836	122	1861	77
1787	132	1812	5	1837	138	1862	59
1788	131	1813	12	1838	103	1863	44
1789	118	1814	14	1839	86	1864	37
1790	90	1815	35	1840	63	1865	40
1791	67	1816	46	1841	37	1866	16
1792	60	1817	41	1842	24	1867	7
1793	47	1818	30	1843	11	1868	37
1794	41	1819	24	1844	15	1869	74

*) 100 наблюдений

**Ряд F. Ряд 70 последовательных
выходов партий продукта циклического
химического процесса *)**

47	44	50	62	68
64	80	71	44	38
23	55	56	64	50
71	37	74	43	60
38	74	50	52	39
64	51	58	38	59
55	57	45	59	40
41	50	54	55	57
59	60	36	41	54
48	45	54	53	23
71	57	48	49	
35	50	55	34	
57	45	45	35	
40	25	57	54	
58	59	50	45	

*) 70 наблюдений; этот ряд приведен также в табл. 2.1 текста.

**Ряд G. Пассажирские перевозки на международных авиалиниях
(месячные итоги в тысячах пассажиров) с января 1949 по декабрь 1960 г. *)**

Годы	январ.	февр.	март	апр.	май	июнь	июль	авг.	сент.	окт.	ноябрь	дек.
1949	112	118	132	129	121	135	148	148	136	119	104	118
1950	115	126	141	135	125	149	170	170	158	133	114	140
1951	145	150	178	163	172	178	199	199	184	162	146	166
1952	171	180	193	181	183	218	230	242	209	191	172	194
1953	196	196	236	235	229	243	264	272	237	211	180	201
1954	204	188	235	227	234	264	302	293	259	229	203	229
1955	242	233	267	269	270	315	364	347	312	274	237	278
1956	284	277	317	313	318	374	413	405	355	306	271	306
1957	315	301	356	348	355	422	465	467	404	347	305	336
1958	340	318	362	348	363	435	491	505	404	359	310	337
1959	360	342	406	396	420	472	548	559	463	407	362	405
1960	417	391	419	461	472	535	622	606	508	461	390	432

*) 144 наблюдения

ЛИТЕРАТУРА

1. *Holt C. C., Modigliani F., Muth J. F., Simon H. A.*, Planning Production, Inventories and Work Force, Prentice-Hall, New Jersey (1963).
2. *Brown R. G.*, Smoothing, Forecasting and Prediction of Discrete Time-Series, Prentice-Hall, New Jersey, 1962.
3. Short Term Forecasting, I. C. I. Monograph, № 2, Oliver and Boyd, Edinburgh, 1964.
4. *Harrison P. J.*, Short-term sales forecasting, *Applied Stat.*, **14**, 102 (1965).
5. *Aström K. J.*, Numerical identification of linear dynamic systems from normal operating records, *Theory of Self-adaptive Control Systems*, 96, Plenum Press, 1966.
6. *Hutchinson A. W., Shelton R. J.*, Measurement of dynamic characteristics of full-scale plant using random perturbing signals: an application to a refinery distillation column, *Trans. Inst. Chem. Eng.*, **45**, 334 (1967).
7. *Briggs P. A. N., Hammond P. H., Hughes M. T. G., Plumb G. O.*, Correlation analysis of process dynamics using pseudo-random binary test perturbations, *Inst. Mech. Eng., Advances in Automatic Control*, Paper 7, Nottingham, U.K., April 1965.
8. *Shewhart W. A.*, The Economic control of the Quality of Manufactured Product, Macmillan, New York, 1931.
9. *Dudding B. P., Jennet W. J.*, Quality control charts, British Standard 600R, 1942.
10. *Page E. S.*, On problems in which a change in a parameter occurs at an unknown point, *Biometrika*, **44**, 249 (1957).
11. *Page E. S.*, Cumulative sum charts, *Technometrics*, **3**, 1 (1961).
12. *Barnard G. A.*, Control charts and stochastic processes, *Jour. Royal Stat. Soc.*, **B21**, 239 (1959).
13. *Roberts S. W.*, Control chart tests based on geometric moving averages, *Technometrics*, **1**, 239 (1959).
14. *Box G. E. P., Jenkins G. M.*, Some statistical aspects of adaptive optimization and control, *J. Roy. Stat. Soc.*, **B24**, 297 (1962).
15. *Box G. E. P., Jenkins G. M.*, Further contributions to adaptive quality control: simultaneous estimation of dynamics: non-zero costs, *Bull. Intl. Stat. Inst.*, 34th Session, 943, Ottawa, Canada, 1963.
16. *Box G. E. P., Jenkins G. M.*, Mathematical models for adaptive control and optimization, *A. I. Ch. E. — I. Chem. E. Symp. Ser.*, **4**, 61 (1965).
17. *Box G. E. P., Jenkins G. M., Bacon D. W.*, Models for forecasting seasonal and non-seasonal time series, *Advanced Seminar on Spectral Analysis of Time Series*, ed. B. Harris, 271, John Wiley, New York, 1967.
18. *Box G. E. P., Jenkins G. M.*, Discrete models for feedback and feedforward control, *The Future of Statistics*, ed. D. G. Watts, 201, Academic Press, New York, 1968.
19. *Box G. E. P., Jenkins G. M.*, Some recent advances in forecasting and control, I, *Applied Stat.*, **17**, 91 (1968).
20. *Box G. E. P., Jenkins G. M.*, Discrete models for forecasting and control, *Encyclopedia of Linguistics, Information and Control*, 162, Pergamon Press, 1969.
21. *Oughton K. D.*, Digital computer controls paper machine, *Ind. Electron.*, **5**, 358 (1965).

22. *Holt C. C.*, Forecasting trends and seasonals by exponentially weighted moving averages, O. N. R. Memorandum, № 52, Carnegie Institute of Technology, 1957.
23. *Winters P. R.*, Forecasting Sales by exponentially weighted moving averages, *Management Sci.*, **6**, 324 (1960).
24. *Yule G. U.*, On a method of investigating periodicities in disturbed series, with special reference to Wölfer's sunspot numbers, *Phil. Trans.*, **A226**, 267 (1927).
25. *Tukey J. W.*, Discussion emphasizing the connection between analysis of variance and spectrum analysis, *Technometrics*, **3**, 191 (1961).
26. *Box G. E. P., Hunter W. G.*, The experimental study of physical mechanisms, *Technometrics*, **7**, 23 (1965).
27. *Jenkins G. M., Watts D. G.*, *Spectral Analysis and Its Applications*, Holden — Day, San Francisco, 1968. (Русский перевод: Г. Дженкинс, Д. Варрс, Спектральный анализ и его приложения, изд-во «Мир», вып. 1, 2, 1971, 1972.)
28. *Bartlett M. S.*, On the theoretical specification properties of sampling properties of autocorrelated time series, *J. Roy. Stat. Soc.*, **B8**, 27 (1946).
29. *Kendall M. G.*, On the analysis of oscillatory time series, *J. Roy. Stat. Soc.*, **108**, 93 (1945).
30. *Schuster A.*, On the investigation of hidden periodicities, *Terr. Mag.*, **3**, 13 (1898).
31. *Stokes G. C.*, Note on searching for periodicities, *Proc. Royal Soc.*, **29**, 122 (1879).
32. *Walker G.*, On periodicity in series of related terms, *Proc. Royal Soc.*, **A131**, 518 (1931).
33. *Stralkowski C. M.*, Lower order autoregressive-moving average stochastic models and their use for the characterization of abrasive cutting tools, Ph. D. Thesis, University of Wisconsin, 1968.
34. *Durbin J.*, The fitting of time series models, *Rev. Int. Inst. Stat.*, **28**, 233 (1960).
35. *Quenouille M. H.*, Approximate tests of correlation in time series, *J. Roy. Stat. Soc.*, **B11**, 68 (1949).
36. *Jenkins G. M.*, Tests of hypothesis in the linear autoregressive model, I: *Biometrika*, **41**, 405 (1954); II: *Biometrika*, **43**, 186 (1956).
37. *Daniels H. E.*, The approximate distribution of serial correlation coefficients, *Biometrika*, **43**, 169 (1956).
38. *Яглом А. М.*, Корреляционная теория процессов со случайными стационарными n -ми приращениями, *Матем. сб.*, **37** (79), 141, 1955.
39. *Zadeh L. A., Ragazzini J. R.*, An extension of Wiener's theory of prediction, *J. Appl. Phys.*, **21**, 645 (1950).
40. *Kalman R. E.*, A new approach to linear filtering and prediction problems, *J. of Basic Eng.*, Series D82, 35 (1960).
41. *Kalman R. E., Bucy R. S.*, New results in linear filtering and prediction theory, *J. of Basic Eng.*, Series D83, 5 (1961).
42. *Box G. E. P., Cox D. R.*, An analysis of transformations, *J. Roy. Stat. Soc.*, **B26**, 211 (1964).
43. *Muth J. F.*, Optimal properties of exponentially weighted forecasts of time series with permanent and transitory components, *J. Amer. Stat. Assoc.*, **55**, 299 (1960).
44. *Wold H. O.*, *A Study in The Analysis of Stationary Time Series*, Almqvist and Wicksell, Uppsala, 1932, 2nd ed. 1954.
45. *Kolmogoroff A.*, Sur l'interpolation et l'extrapolation des suites stationnaires, *Compt. Rend.*, **208**, 2043 (1939).
46. *Колмогоров А. Н.*, Стационарные последовательности в Гильбертовском пространстве, *Бюлл. МГУ*, **2**, № 6, 1—40, 1941.

47. Колмогоров А. Н., Интерполяция и экстраполяция стационарных случайных последовательностей, Изв. АН СССР, сер. матем., т. 5, № 3, 1941.
48. Wiener N., Extrapolation, Interpolation and Smoothing of Stationary Time Series, John Wiley, New York, 1949.
49. Whittle P., Prediction and Regulation by Linear Least — Squares Methods, English Universities Press, London, 1963.
50. Brown R. G., Meyer R. F., The fundamental theorem of exponential smoothing, Operations Res., 9, 673 (1961).
51. Bachelier L., Theorie de la speculation, Ann. Sci. Ec. norm. sup., Paris, Series 3, 17, 21 (1900).
52. Anderson R. L., Distribution of the serial correlation coefficient, Ann. Math. Stat., 13, 1 (1942).
53. Schuster A., On the periodicities of sunspots, Phil. Trans. Royal Soc., A206, 69 (1906).
54. Moran P. A. P., Some experiments in the prediction of sunspots numbers, J. Roy. Stat. Soc., B16, 112 (1954).
55. Wilson G. J., Factorization of the generating function of a pure moving average process, SIAM Jour. Num. Analysis, 6, 1 (1969).
56. Fisher R. A., Statistical Methods and Scientific Inference, Oliver and Boyd, Edinburgh, 1956. (Русский перевод: Р. Фишер, Статистические методы для исследователей, Госстатиздат, М., 1958.)
57. Barnard G. A., Statistical inference, J. Royal Stat. Soc., B11, 116 (1949).
58. Birnbaum A., On the foundations of statistical inference, J. Amer. Stat. Assoc., 57, 269 (1962).
59. Rao C. R., Linear Statistical Inference and Its Applications, John Wiley, New York, 1965. (Русский перевод: Р. С. Рав, Линейные статистические методы и их приложения, М., «Наука», 1968.)
60. Box G. E. P., Draper N. R., The Bayesian estimation of common parameters from several responses, Biometrika, 52, 355 (1965).
61. Wilks S. S., Mathematical Statistics, John Wiley, New York, 1962. (Русский перевод: С. С. Уилкс, Математическая статистика, М., «Наука», 1967.)
62. Both G. W., Peterson T. I., Non-linear estimation, IBM Share Programm Pa. No. 687 WL NLI, 1958.
63. Marquardt D. W., An algorithm for least squares estimation of non-linear parameters, J. Ind. Appl. Math., 11, 431 (1963).
64. Savage L. J., The Foundations of Statistical Inference, Methuen, London, 1962.
65. Jeffreys H., Theory of Probability, 3rd. ed., Clarendon Press, Oxford, 1961.
66. Box G. E. P., Taio G. C., Bayesian Inference, Addison — Wesley, в печати.
67. Jenkins G. M., Contribution to the discussion of the paper «Relationships between Bayesian and confidence limits for predictors» by A. R. Thatcher, J. Roy. Stat. Soc., B26, 176 (1964).
68. Cornish E. A., The multivariate t-distribution associated with a set of normal sample deviates, Austral J. Phys., 7, 531 (1954).
69. Dunnett C. W., Sobel M., A bivariate generalization of Student's *t*-distribution, with tables for special cases, Biometrika, 31, 153 (1954).
70. Barnard G. A., The logic of least squares, J. Royal Stat. Soc., B25, 124 (1963).
71. Plackett R. L., Principles of Regression Analysis, Clarendon Press, Oxford, 1960.
72. Anscombe F. J., Examination of residuals, Proc. 4th. Berkeley Symp., 1, 1 (1961).

73. *Anscombe F. J., Tukey J. W.*, The examination and analysis of residuals, *Technometrics*, 5, 141 (1963).
74. *Daniel C.*, Use of half normal plots in interpreting factorial experiments, *Technometrics*, 1, 311 (1959).
75. *Durbin J.*, Testing for serial correlation in least-squares regression when some of the regressors are lagged dependent variables, *Econometrica*, 37 (1969).
76. *Durbin J.*, An alternative to the bounds tests for testing serial correlation in least squares regression, *Economica*, 37, in press, 1969.
77. *Box G. E. P., Pierce D. A.*, Distribution of residual autocorrelations in autoregressive-integrated moving average time series models, *J. Amer. Stat. Ass.*, 64 (1970).
78. *Bartlett M. S.*, *Stochastic Processes*, Cambridge University Press, Cambridge, 1955.
79. *Hald A.*, *Statistical Theory with Engineering Applications*, John Wiley, New York, 1952. (Русский перевод: А. Хальд, Математическая статистика с техническими приложениями, М., ИЛ, 1956.)
80. *Bacon D. W.*, Seasonal time series, Ph. D. Thesis, University of Wisconsin, Madison, 1965.
81. *Young A. J.*, *An Introduction to Process Control Systems Design*, Longman Green, New York, 1955.
82. *Hougen J. O.*, Experience and Experiments with Process Dynamics, *Chemical Engineering Progress Monograph Series*, 60, № 4 (1964).
83. *Kotnour K. D., Box G. E. P., Altpeter R. J.*, A discrete predictor-controller applied to sinusoidal perturbation adaptive optimization, *Instr. Soc. Amer. Trans.*, 5, 225 (1966).
84. *Pierce D. A.*, Distribution of residual correlations in dynamic/stochastic time series models, University of Wisconsin Tech. Rep. 173, August 1968.
85. *Box G. E. P., Jenkins G. M., Wichern D. W.*, Least squares analysis with a dynamic model, University of Wisconsin Technical Report 105, 1967.
86. *Box G. E. P.*, Use and abuse of regression, *Technometrics*, 8, 625 (1966).
87. *Whittle P.*, Estimation and Information in Stationary time series, *Arkiv für Matematik*, 2, 423 (1953).
88. *Barnard G. A., Jenkins G. M., Winsten C. B.*, Likelihood inference and time series, *Jour. Royal Stat. Soc.*, A125, 321 (1962).
89. *Kendall M. G., Stuart A.*, *The Advanced Theory of Statistics*, Vol. 3, Griffin, London, 1966.
90. *Tintner G.*, *The Variate Difference Method*, Principia Press, Bloomington, Indiana, 1940.
91. *Tintner G., Rao J. N. K.*, On the variate difference method, *Australian Journal of Statistics*, 5, 106 (1963).
92. *Слуцкий Е. Е.*, Сложение случайных причин как источник циклических процессов, *Вопросы конъюнктуры*, т. 3, 1, 1927.
93. *Mann H. B., Wald A.*, On the statistical treatment of linear stochastic difference equations, *Econometrica*, 11, 173 (1943).
94. *Koopmans T. C.*, Serial correlation and quadratic forms in normal variables, *Ann. Math. Stat.*, 13, 14 (1942).
95. *Koopmans T. C.* (ed.), *Statistical Inference in Dynamic Economic Models*, John Wiley, New York, 1950.
96. *Cox D. P.*, Prediction by exponentially weighted moving averages and related methods, *J. Roy. Stat. Soc.*, B23, 414 (1961).
97. *Hannan E. J.*, *Time Series Analysis*, Methuen, London, 1960. (Русский перевод: Э. Хеннан, Анализ временных рядов, М., «Наука», 1964.)
98. *Grenander U., Rosenblatt M.*, *Statistical Analysis of Stationary Time Series*, John Wiley, New York, 1957.
99. *Quenouille M. H.*, *Analysis of Multiple Time Series*, Hafner, New York, 1957.

100. *Quenouille M. H.*, Associated Measurements, Butterworth, London, 1952.
101. *Whittle P.*, Hypothesis Testing in Time Series Analysis, University of Uppsala publication, 1951.
102. *Doob J. L.*, Stochastic Processes, John Wiley, New York, 1953. (Русский перевод: Дж. Дуб., Вероятностные процессы, М., ИЛ, 1956.)
103. *Robinson E. A.*, Multichannel Time Series Analysis, Holden — Day, San Francisco, 1967.
104. *Box G. E. P., Tiao G. C.*, Multiparameter problems from a Bayesian point of view, *Ann. Math. Statist.*, **36**, 1468 (1968).
105. *Reilly P. M.*, Personal Communication, 1967.
106. *Noble B.*, Methods Based on the Wiener — Hopf Technique for the Solution of Partial Differential Equations, Pergamon Press, New York, 1958.
107. *Wilson G. J.*, Optimal Control — A General Method of Obtaining the Feedback Scheme which Minimizes the Output Variance, Subject to a Constraint on the Variability of the Control Variable, Technical Report № 20, Dept. of Systems Engineering University of Lancaster.

Предметный указатель

- Автоковариации 361
Автоковариационная матрица 44
— функция 46, 142
Автокорреляционная матрица 44
— функция 37, 46, 59
— — линейного процесса 65
— — нестационарного процесса 224
— — ошибок прогноза 177
— — процесса авторегрессии 71
— — — АРПСС 197
— — — АРСС 92
— — — скользящего среднего 85
— — частная 81, 88, 93
- Безусловное правдоподобие 237
Белый шум 23, 29, 63
- Вероятностные пределы 155, 176
Веса прогноза 152, 160, 173
— — для произвольного упреждения 180
Возвратное представление 221, 229
Временные ряды 39
Время упреждения 15
Выборочный спектр 56
- Гауссовские процессы 46
- Диагностическая проверка 11, 35, 313, 352, 374
Дисперсии и ковариации оценок МП 252, 309, 351
Дисперсионный анализ 53
Доверительная область 254, 294, 347
- Идентификация моделей 11, 34, 61, 192, 344, 355, 367
— реальных временных рядов 199, 220
— — цели и методика 193
Избыточность параметров 275, 314
- Информационная матрица 253, 267, 309
Итеративный подход 34
- Ковариационная матрица оценок 268, 294
Кумулятивная периодограмма 323
- Линейные разностные уравнения 132
Линейный процесс 63
— фильтр 24
- Маргинальное распределение 289
Марквардта алгоритм 376
Марковский процесс 74
Метод Брауна 187
Многозначность моделей 218
Модели временных рядов 10, 22
— передаточных функций 28
Модель авторегрессии 24
— АРПСС 111
— мультипликативная 335, 353
- Наивысшая плотность вероятности (НПВ) 282
Начальные оценки параметров 209, 220, 225
Нормальное распределение 286
— уравнение 293
- Обратимость 67, 84
Оператор суммирования 138
Остаточные ошибки 317, 327, 373
Оценивание 347, 356, 369
— передаточных функций 17
— спектра 58
Оценки максимального правдоподобия (МП) 233, 245, 305
— Юла — Уокера 270, 307

- Периодограмма 52
 Подгонка моделей 11, 35
 Подправление прогнозов 153, 176
 Подстраивающиеся коэффициенты 166
 Примеры временных рядов:
 выходные данные химических процессов 102, 202, 389
 пассажирские перевозки 330, 348, 396
 температурный ряд 54
 цены акций компании IBM 104, 190, 205, 390
 число солнечных пятен 199, 207, 394
 Принцип экономии параметров 33
 Прогнозирование назад 223, 259
 — процессов ППС 162
 — — авторегрессии 169
 Прогнозирующая функция 16, 146, 162, 339, 357, 378
 — — эвентуальная 157, 166, 182
 Прогнозы с минимальной среднеквадратичной ошибкой 144, 175
 — — — подправление 153
 — — — способы представления 148, 176
 — — — сравнение с методом Брауна 186
 — — — схема вычисления 150
 Программы для ЭВМ 366
 Производящая функция 65
 — — автоковариаций 66, 98
 Прointегрированное представление 182
 Процессы авторегрессии (АР) 24, 68, 270, 280, 297
 — — проинтегрированного скользящего среднего (АРПСС) 23, 27, 102, 106
 — — — с добавленным шумом 139
 — — скользящего среднего (АРСС) 26, 69, 91, 213
 — проинтегрированного скользящего среднего (ПСС) 121
 — — — с детерминированным дрейфом нуля 137
 — скользящего среднего (СС) 25, 69, 84, 273, 283, 297, 312
 Распределение маргинальное 289
 — нормальное 286
 — Стьюдента 290
 Регулирование с прямой и обратной связями 19, 31
 Ряд солнечных пятен 100, 394
 Сборник временных рядов 388
 Сводка моделей для рядов А — F 265
 Сезонные процессы 6
 Случайное блуждание 142
 Смешанные процессы 273, 285, 300, 312
 Совокупный критерий согласия 320
 Спектр 37, 56
 — нормированный 57
 — процесса авторегрессии 71
 — — АРСС 92
 — — СС 85
 Стандартные ошибки оценок автокорреляций 50
 — — — — выборочных 198
 — — — — частных 82
 Стохастические модели 22, 37
 — процессы 40
 Сумма квадратов 235, 245
 — — безусловная 237
 — — вычисление на ЭВМ 373
 — — графическое представление 245, 275, 347
 — — условная 236
 Таблицы и диаграммы 381
 Теорема Байеса 278
 Тренды 109
 Уравнения регулирования 19, 32
 — Юла — Уокера 73, 81, 99
 Уровни значимости 385
 Условное правдоподобие 233
 Формула Бартлетта 51, 198
 Формы представления модели АРПСС 144
 Функция отклика на единичный импульс 29
 — правдоподобия 232, 251, 297
 Характеристическое уравнение 71
 Холостое время 28
 Экономичность параметризации 332
 Экспоненциально взвешенное скользящее среднее (ЭВСС) 124, 229

Оглавление

ПРЕДИСЛОВИЕ К РУССКОМУ ИЗДАНИЮ	5
ПРЕДИСЛОВИЕ	10
ПЛАН КНИГИ	13
Глава 1. ВВЕДЕНИЕ И КРАТКОЕ СОДЕРЖАНИЕ	15
1.1. Три важные практические проблемы	15
1.2. Стохастические и детерминированные динамические математические модели	21
1.3. Основные понятия в построении моделей	32
Часть I. СТОХАСТИЧЕСКИЕ МОДЕЛИ И ОСНОВАННОЕ НА НИХ ПРОГНОЗИРОВАНИЕ	
Глава 2. АВТОКОРРЕЛЯЦИОННАЯ ФУНКЦИЯ И СПЕКТР	39
2.1. Автокорреляционные свойства стационарных моделей	39
2.2. Спектральные свойства стационарных моделей	52
Приложение П2.1. Связь между выборочным спектром и оценкой автоковариационной функции	61
Глава 3. ЛИНЕЙНЫЕ СТАЦИОНАРНЫЕ МОДЕЛИ	63
3.1. Общий линейный процесс	63
3.2. Процессы авторегрессии	70
3.3. Процессы скользящего среднего	84
3.4. Смешанные процессы авторегрессии — скользящего среднего	91
Приложение П3.1. Автоковариации. Производящая функция автоковариаций и условия стационарности общего линейного процесса	98
Приложение П3.2. Рекуррентный метод вычисления оценок параметров авторегрессии	99
Глава 4. ЛИНЕЙНЫЕ НЕСТАЦИОНАРНЫЕ МОДЕЛИ	102
4.1. Процессы авторегрессии — проинтегрированного скользящего среднего	102
4.2. Три формы представления модели авторегрессии — проинтегрированного скользящего среднего	111
4.3. Процессы проинтегрированного скользящего среднего	121
Приложение П4.1. Линейные разностные уравнения	132
Приложение П4.2. Процесс ПСС (0, 1, 1) с детерминированным дрейфом нуля	137
Приложение П4.3. Свойства конечного оператора суммирования	138
Приложение П4.4. Процессы АРПСС с добавленным шумом , , . . .	139

Глава 5. ПРОГНОЗИРОВАНИЕ 144

5.1. Прогнозы с минимальной среднеквадратичной ошибкой и их свойства	144
5.2. Вычисление и подправление прогноза	150
5.3. Прогнозирующая функция и веса прогноза	157
5.4. Примеры прогнозирующих функций и их подправления	162
5.5. Резюме	175
Приложение П5.1. Корреляция между ошибками прогноза	177
Приложение П5.2. Веса прогноза для произвольного упреждения . .	180
Приложение П5.3. Прогнозирование при помощи общего проинтегрированного представления	182

Часть II. ПОСТРОЕНИЕ СТОХАСТИЧЕСКИХ МОДЕЛЕЙ

Глава 6. ИДЕНТИФИКАЦИЯ МОДЕЛЕЙ 193

6.1. Цели идентификации	193
6.2. Методика идентификации	194
6.3. Начальные оценки параметров	209
6.4. Многозначность моделей	218
Приложение П6.1. Среднее значение выборочной автокорреляционной функции нестационарного процесса	224
Приложение П6.2. Общий метод получения начальных оценок параметров смешанного процесса авторегрессии — скользящего среднего	225
Приложение П6.3. Прямой и возвратный процессы ПСС порядка $(0, 1, 1)$	229

Глава 7. ОЦЕНИВАНИЕ МОДЕЛИ. 232

7.1. Исследование функций правдоподобия и суммы квадратов . .	232
7.2. Нелинейное оценивание	257
7.3*. Результаты оценивания для некоторых частных моделей . .	270
7.4*. Оценивание при помощи теоремы Байеса	278
Приложение П7.1. Обзор теории нормального распределения . . .	286
Приложение П7.2. Обзор линейной теории наименьших квадратов .	293
Приложение П7.3. Примеры влияния ошибок оценивания параметров на вероятностные пределы прогнозов	295
Приложение П7.4. Точная функция правдоподобия для процесса скользящего среднего	297
Приложение П7.5. Точная функция правдоподобия для процесса авторегрессии	302

Глава 8. ДИАГНОСТИЧЕСКАЯ ПРОВЕРКА МОДЕЛИ 313

8.1. Проверка стохастических моделей	313
8.2. Диагностические проверки, применяемые к остаточным ошибкам	317
8.3. Использование остаточных ошибок для изменения модели . .	327

Глава 9. МОДЕЛИ СЕЗОННЫХ РЯДОВ 330

9.1. Экономические модели сезонных временных рядов	330
9.2. Представление данных об авиаперевозках мультипликативной моделью $(0, 1, 1) \times (0, 1, 1)_{12}$	336
9.3. Некоторые аспекты более общих моделей сезонных рядов . .	353
Приложение П9.1	361

ВСПОМОГАТЕЛЬНЫЕ МАТЕРИАЛЫ

ОПИСАНИЕ ПРОГРАММ ДЛЯ ЭВМ	366
Программа 1. Идентификация стохастической модели	367
Программа 2. Предварительное оценивание стохастической модели	369
Программа 3. Оценивание стохастической модели	372
Алгоритм Марквардта для нелинейного метода наименьших квадратов	376
Программа 4. Прогнозирование с помощью стохастической модели	377
СБОРНИК ТАБЛИЦ И ДИАГРАММ	381
СБОРНИК ВРЕМЕННЫХ РЯДОВ, АНАЛИЗИРУЕМЫХ В КНИГЕ	388
ЛИТЕРАТУРА	397
ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ	402

УВАЖАЕМЫЙ ЧИТАТЕЛЬ!

Ваши замечания о содержании книги, ее оформлении, качестве перевода и др. просим присылать по адресу: 129820, Москва, И-110, ГСП, 1-й Рижский пер., д. 2, изд-во «Мир»

Дж. Бокс, Г. Дженкинс

**АНАЛИЗ ВРЕМЕННЫХ РЯДОВ
ПРОГНОЗ И УПРАВЛЕНИЕ**

Редактор Э. А. Медушевская

Художник С. А. Бычков

Художественный редактор И. А. Шаврова

Технический редактор В. П. Сизова

Корректоры Т. П. Пашковская и В. И. Постнова

Сдано в набор 24/VIII 1973 г.

Подписано к печати 9/I 1974 г.

Вумага тип. № 2 60×90¹/₁₆ = 12,75 бум. л.

25,50 печ. л., Уч.-изд. л. 22,50. Изд. № 27/6871.

Цена 1 р. 72 к. Зак. 777

ИЗДАТЕЛЬСТВО «МИР»

Москва, 1-й Рижский пер., 2

Ордена Трудового Красного Знамени

Ленинградская типография № 2

имени Евгении Соколовой Союзполиграфпрома

при Государственном комитете Совета

Министров СССР по делам издательств,

полиграфии и книжной торговли

198052, Ленинград, Л-52,

Измайловский проспект, 29

1р. 72 к.