
ЧАСТЬ 4

ВРЕМЕННЫЕ РЯДЫ: ДОПОЛНИТЕЛЬНЫЕ ГЛАВЫ.

МОДЕЛЬ СТОХАСТИЧЕСКОЙ ГРАНИЦЫ

Р а з д е л 5

СГЛАЖИВАНИЕ И ПРОГНОЗИРОВАНИЕ ВРЕМЕННЫХ РЯДОВ

Тема 5.1

АДАПТИВНЫЕ МЕТОДЫ, МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

Во второй части учебника уже говорилось о том, что под временным рядом (*time series*) в экономике понимается ряд значений некоторой переменной, измеренных в последовательные моменты времени. Для многих рядов измерения производятся через равные промежутки времени (годовые, квартальные, недельные, дневные данные). Если принять длину такого промежутка за единицу времени (год, квартал, неделя, день), то можно считать, что последовательные наблюдения x_1, \dots, x_T переменной x произведены в моменты $t = 1, \dots, T$. Для некоторых экономических и финансовых показателей производить измерения через равные промежутки времени не удастся. Например, значения биржевых индексов на момент закрытия фиксируются только в те дни, когда биржа работает. В последнем случае наблюдения x_1, \dots, x_T соответствуют T последовательным рабочим дням биржи.

В динамике уровней ряда могут наблюдаться «систематические эффекты»:

- некоторая **основная тенденция** (тренд — *trend*);
- **циклические колебания** (*cyclicity*);
- **сезонная изменчивость** (*seasonality*).

После учета этих систематических эффектов остается некоторая флуктуация, которая носит нерегулярный характер (*irregularity*). В связи с этим уровень ряда x_t часто представляется в виде:

- суммы тренда, циклической, сезонной и нерегулярной компонент (**аддитивная модель**):

$$x_t = T_t + C_t + S_t + I_t;$$

- произведения этих компонент (**мультипликативная модель**):

$$x_t = T_t C_t S_t I_t.$$

В рамках такого подхода исследование временного ряда состоит в выявлении и придании количественного выражения каждой из указанных компонент. При этом конечной целью исследования может являться использование полученных результатов для прогнозирования будущих значений ряда.

Нерегулярная составляющая ряда может быть уменьшена или удалена путем применения различных методов сглаживания ряда, так что при этом более отчетливо выявляются трендовые, сезонные и циклические составляющие ряда. В рамках темы 5.1 не делается никаких предположений о возможной вероятностной модели порождения ряда и рассматриваются методы сглаживания, не требующие явного описания такой модели. При изложении этих методов предполагается, что наблюдается некоторый временной ряд x_1, x_2, \dots, x_T (T — количество наблюдений). Для реализации рассматриваемых ниже методов используется пакет программ эконометрического анализа *Econometric Views* (версия 6). Приведены результаты применения этих методов к различным экономическим временным рядам.

Скользящие средние

Простое скользящее среднее (*SMA* — *simple moving average*) с периодом усреднения K (*K-point moving average*) формируется путем усреднения K последовательных значений ряда x_1, x_2, \dots, x_T . Если в качестве K берется нечетное число, $K = 2m + 1$, то значение x_i исходного ряда заменяется значением

$$sma_i = \frac{x_{i-m} + x_{i-m+1} + \dots + x_{i+m}}{K}.$$

Разумеется, значения sma_i определяются только для $i = m + 1, \dots, T - m$. В частном случае $K = 3$ имеем:

$$sma_i = \frac{x_{i-1} + x_i + x_{i+1}}{3}, \quad i = 2, \dots, T - 1;$$

если $K = 5$, то

$$sma_i = \frac{x_{i-2} + x_{i-1} + x_i + x_{i+1} + x_{i+2}}{5}, \quad i = 3, \dots, T - 2.$$

ПРИМЕР 5.1.1

Рассмотрим динамику значений индекса ММББ (Московской межбанковской валютной биржи) на периоде с 18 сентября 2007 г. по 14 февраля 2008 г. (ряд *IND* — рис. 5.1).

Сгладим этот ряд, используя трехпериодное скользящее среднее. Результат сглаживания приведен на рис. 5.2. Для наглядности на рис. 5.3 приведены графики исходного (*IND*) и сглаженного (*SMA_3*) рядов.

На рис. 5.4 приведен график разности $DELTA_3 = IND - SMA_3$, динамика этой разности не имеет систематического тренда.

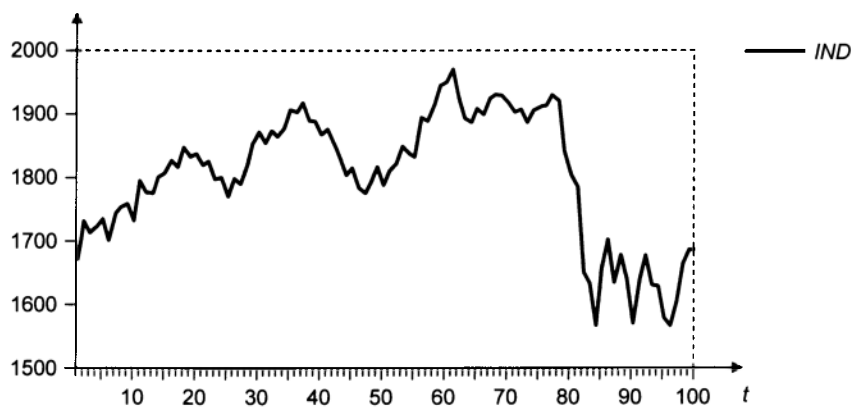


Рис. 5.1

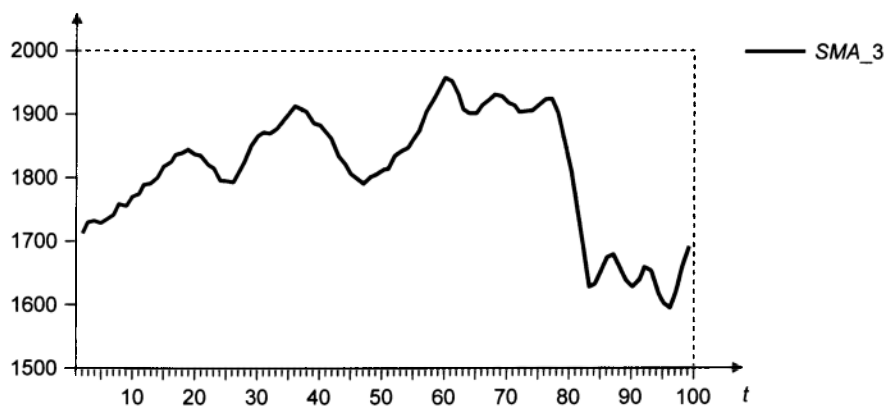


Рис. 5.2

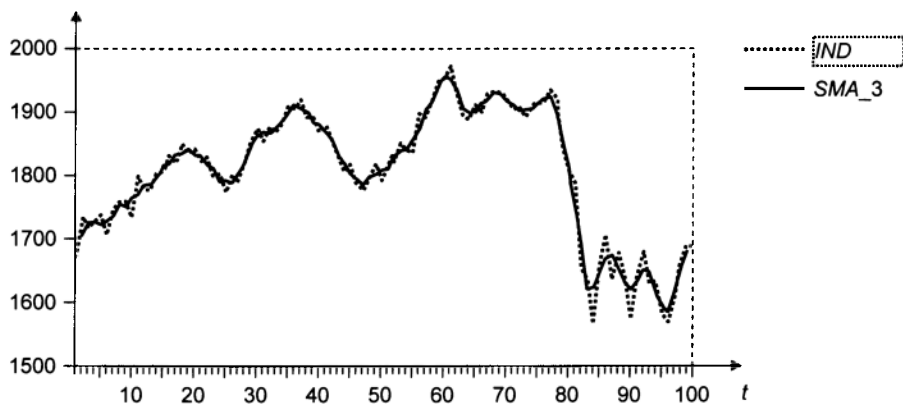


Рис. 5.3

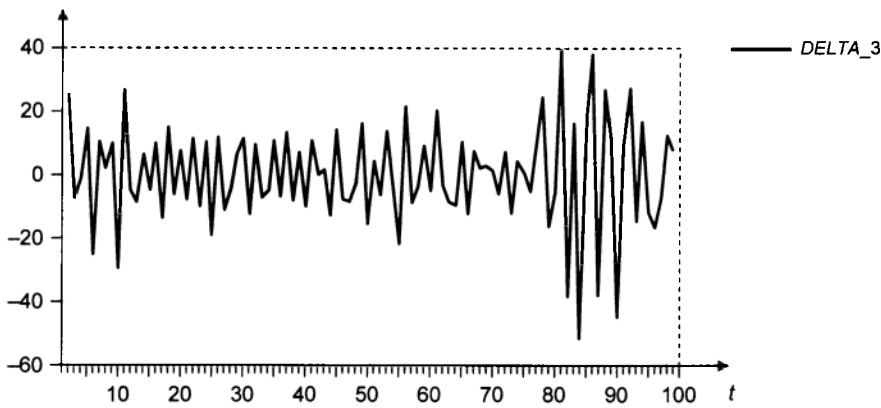


Рис. 5.4

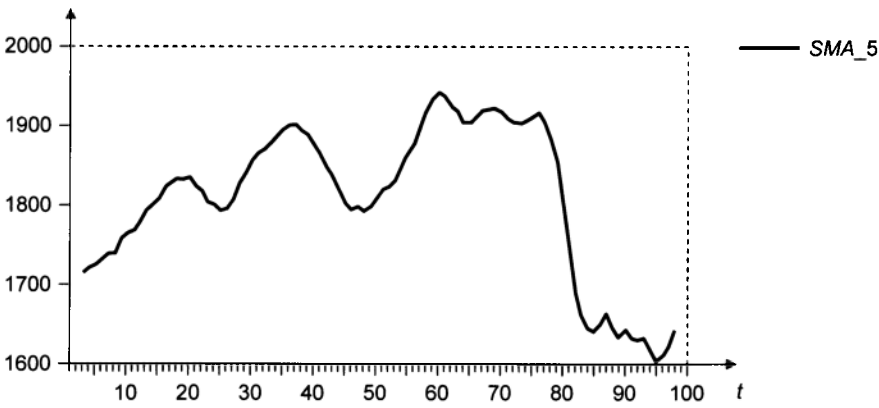


Рис. 5.5

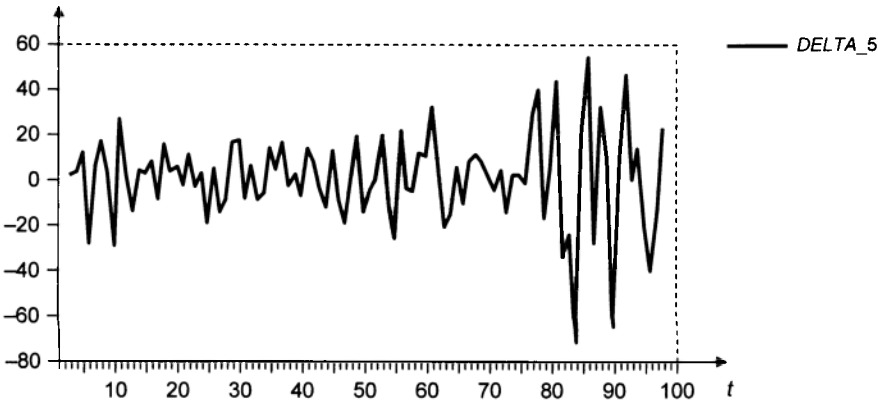


Рис. 5.6

Используем теперь 5-периодное скользящее среднее. График сглаженного ряда приведен на рис. 5.5, а график разности $DELTA_5 = IND - SMA_5$ — на рис. 5.6.■

При рассмотрении месячных и квартальных данных в динамике рядов часто проявляются сезонные колебания, в таких случаях естественно использовать простые скользящие средние с периодами усреднения $K = 12$ и $K = 4$, соответственно. Но возникает следующий вопрос. Когда используется усреднение по нечетному количеству точек, например, по $K = 3$ точкам, то усредненное значение

значение $\frac{x_{i-1} + x_i + x_{i+1}}{3}$ сопоставляется i -му периоду времени, т.е. центральному из трех наблюдений. Если же взять, например, $K = 4$, какому периоду

времени следует сопоставить усредненное значение $\frac{x_{i-1} + x_i + x_{i+1} + x_{i+2}}{4}$?

По аналогии со случаем нечетного K это должен быть период $i + 0.5$, и это порождало бы значения $sma_{2.5}, sma_{3.5}, \dots, sma_{T-1.5}$. Чтобы обойти это затруднение, производят дополнительное усреднение полученного ряда простых скользящих средних, используя малое четное значение периода усреднения K (например, $K = 2$). В последнем случае получаем в итоге последовательность значений $sma_3^*, sma_4^*, \dots, sma_{T-2}^*$, где

$$sma_i^* = \frac{sma_{i-0.5} + sma_{i+0.5}}{2},$$

которую называют **центрированным скользящим средним** (*centered moving average*). Заметим, что при этом

$$\begin{aligned} sma_i^* &= \frac{1}{2} \left(\frac{x_{i-2} + x_{i-1} + x_i + x_{i+1}}{4} + \frac{x_{i-1} + x_i + x_{i+1} + x_{i+2}}{4} \right) = \\ &= \frac{1}{4} (0.5x_{i-2} + (x_{i-1} + x_i + x_{i+1}) + 0.5x_{i+2}). \end{aligned}$$

ПРИМЕР 5.1.2

Рассмотрим временной ряд $AGRO$, представляющий квартальные данные об индексе реального объема сельскохозяйственного производства (1993:01 = 100) за 12-летний период — с I квартала 1994 г. по IV квартал 2005 г. Динамика этого ряда имеет ярко выраженный сезонный характер (рис. 5.7).

Сгладим этот ряд, используя центрированное среднее с периодом 4. Получим сглаженный ряд $SMA_4_CENTERED$ (рис. 5.8), не определенный для первых двух и для последних двух наблюдений исходного ряда.

На графике разности $DELTA = AGRO - SMA_4_CENTERED$ (рис. 5.9) видно, что динамика этой разности не имеет систематического тренда, который взял на себя сглаженный ряд.■

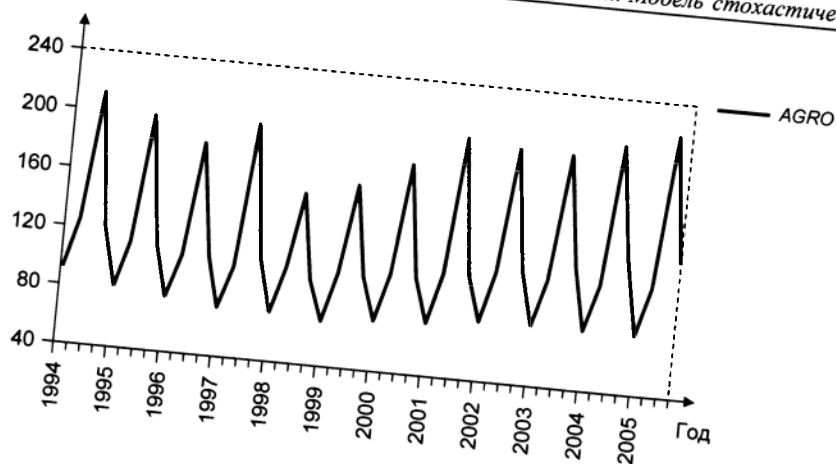


Рис. 5.7

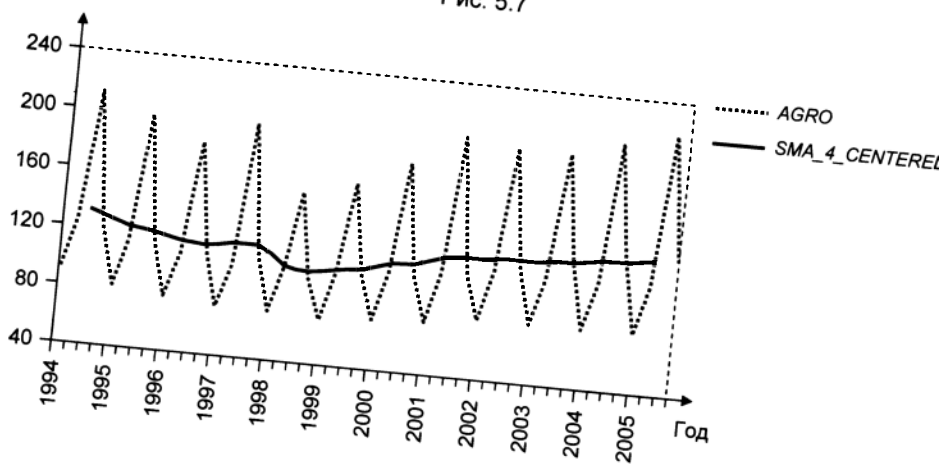


Рис. 5.8

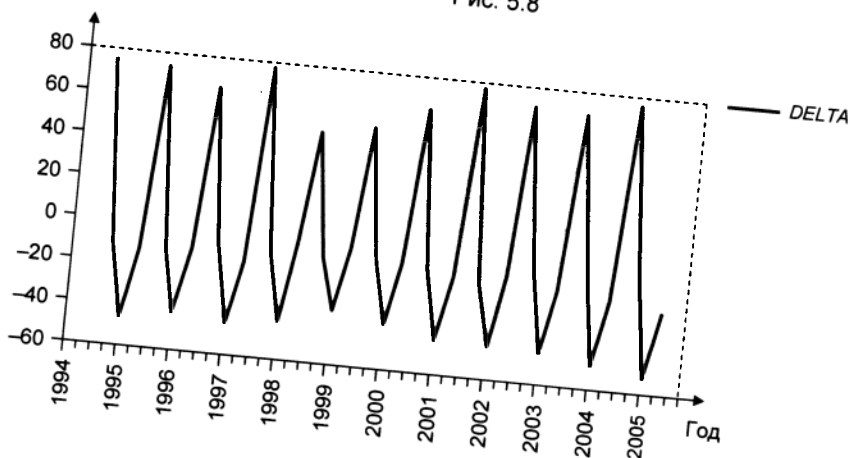


Рис. 5.9

ПРИМЕР 5.1.3

Рассмотрим временной ряд *IMP*, представляющий квартальные данные об импорте товаров и услуг в Российскую Федерацию (в млрд долл.) за 12-летний период — с I квартала 1994 г. по IV квартал 2005 г. Динамика этого ряда показана на рис 5.10.

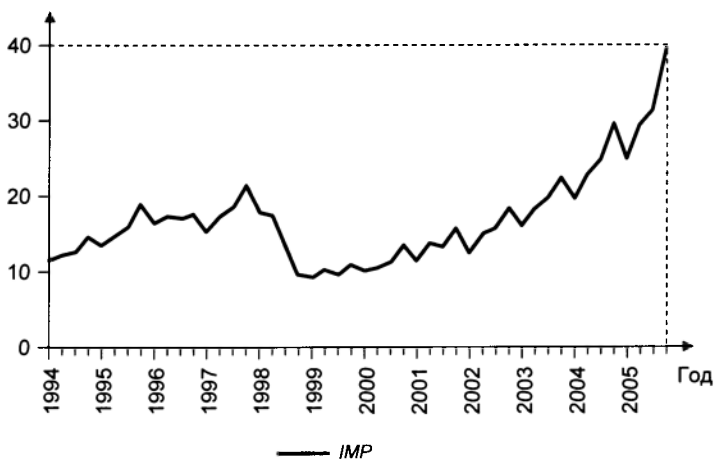


Рис. 5.10

Сгладим этот ряд, опять используя центрированное среднее с периодом 4. Для наглядности поместим исходный и сглаженный ряды на одном графике (рис. 5.11).

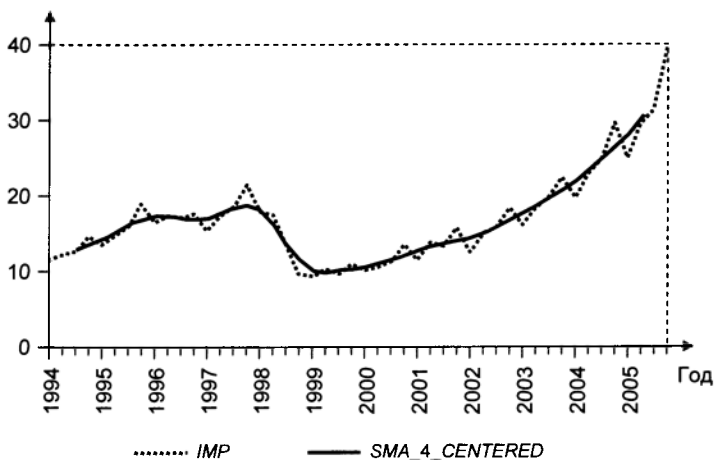


Рис. 5.11

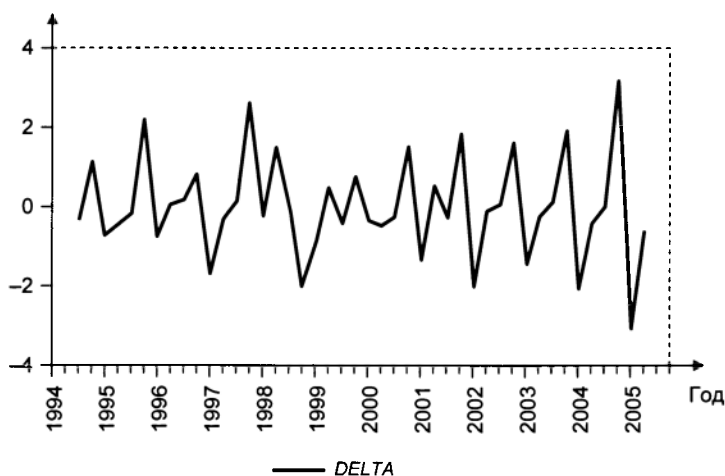


Рис. 5.12

Рассмотрим график разности $DELTA = IMP - SMA_4_CENTERED$ (рис. 5.12). Динамика этой разности не имеет систематического тренда. ■

Фильтр Ходрика — Прескотта

Этот инструмент используется для получения сглаженной оценки долговременного тренда ряда x_t . Соответствующий алгоритм подбирает сглаженный ряд s_t , для которого минимизируется сумма

$$\sum_{t=1}^T (x_t - s_t)^2 + \lambda \sum_{t=1}^T ((s_{t+1} - s_t) - (s_t - s_{t-1}))^2.$$

Параметр $\lambda > 0$ управляет гладкостью получаемого ряда. Чем больше λ , тем более гладким является ряд s_t . При $\lambda \rightarrow \infty$ ряд s_t приближается к линейному тренду. В пакете *EViews* по умолчанию используются следующие значения параметра λ :

- $\lambda = 100$ для годовых данных;
- $\lambda = 1600$ для квартальных данных;
- $\lambda = 14\,400$ для месячных данных.

ПРИМЕР 5.1.4

Применяя фильтр Ходрика — Прескотта с $\lambda = 1600$ к ряду $AGRO$, получаем ряд $HPTREND$, интерпретируемый как «долговременный тренд». График этого ряда вместе с графиком ряда $AGRO$ представлен на рис. 5.13. ■

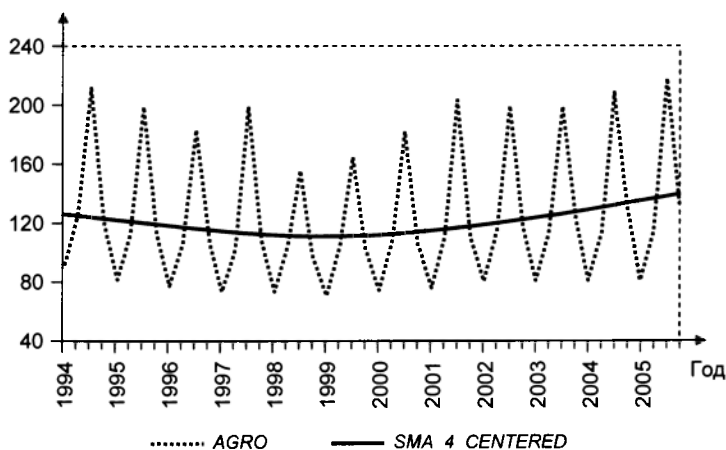


Рис. 5.13

ПРИМЕР 5.1.5

Применяя фильтр Ходрика — Прескотта с $\lambda = 1600$ к ряду *IMP*, выделяем из этого ряда «долговременный тренд» *HPTREND*. Графики рядов *IMP* и *HPTREND* представлены на рис. 5.14. ■

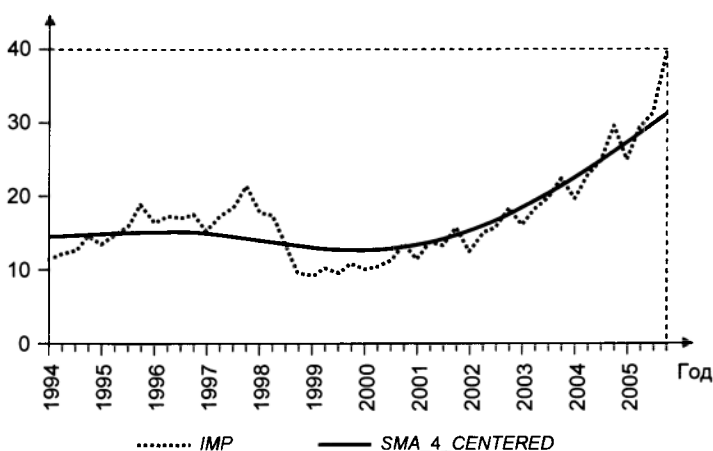


Рис. 5.14

ПРИМЕР 5.1.6

Применим фильтр Ходрика — Прескотта к ряду *IND*. Здесь данные дневные, с пропусками на выходные дни. Попробуем использовать различные значения параметра сглаживания: 100 (как для годовых данных), 1600 (как для квартальных данных), 14 400 (как для месячных данных) и 100 000. Получим результаты, отраженные на графиках, приведенных на рис. 5.15 (а, б, в, г). ■

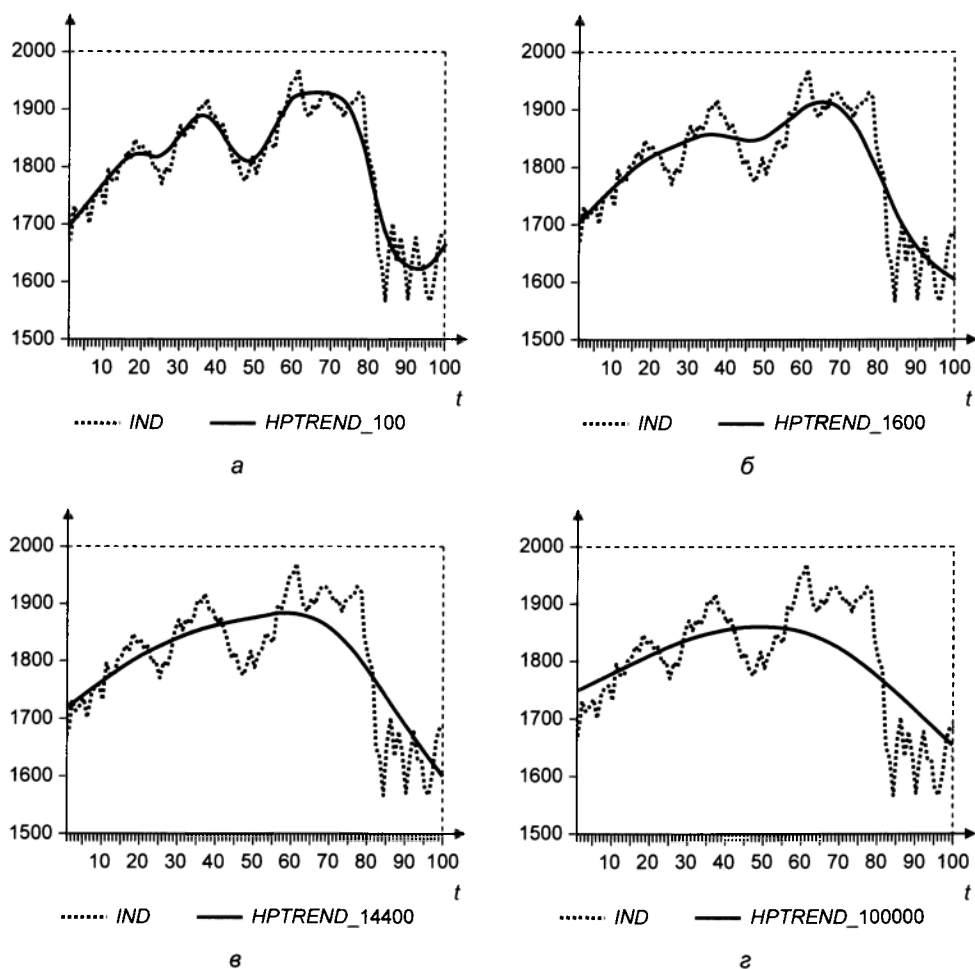


Рис. 5.15

Сглаживание ряда и прогнозирование будущих значений ряда

Прогнозирование будущих значений ряда с использованием простого скользящего среднего опирается на понятие **локального «текущего уровня»** (*current level*), вокруг которого происходят случайные (нерегулярные) флуктуации ряда. Этот уровень может изменяться с течением времени, и для периода i ($i = K, K + 1, \dots, T$) он определяется как

$$L_i = \frac{x_{i-K+1} + x_{i-K+2} + \dots + x_i}{K}$$

(т.е. среднее арифметическое K последних значений ряда).

Прогнозные значения ряда на любое количество h шагов вперед приравняются последнему текущему уровню L_T , который можно получить на основании имеющихся наблюдений x_1, x_2, \dots, x_T :

$$\hat{x}_{T+h} \equiv L_T = \frac{x_{T-K+1} + x_{T-K+2} + \dots + x_T}{K}.$$

Такой прогноз не принесет большой пользы, если ряд содержит сезонную компоненту и/или имеет выраженный возрастающий или убывающий тренд.

Взвешенное скользящее среднее (*WMA* — *Weighted Moving Average*) — метод сглаживания, в котором наблюдения усредняются с различными весами, так что наибольшие веса приписываются последним наблюдениям.

При использовании усреднения ряда x_1, x_2, \dots, x_T по трем периодам ($K = 3$)

$$wma_i = w_1 x_{i-1} + w_2 x_i + w_3 x_{i+1},$$

где w_1, w_2, w_3 — веса, приписываемые усредняемым наблюдениям: $w_3 \geq w_2 \geq w_1 \geq 0$ и $w_1 + w_2 + w_3 = 1$.

Например, можно положить

$$w_1 = 1/6, \quad w_2 = 2/6, \quad w_3 = 3/6,$$

и при таком выборе весов

$$wma_i = \frac{x_{i-1} + 2w_2 x_i + 3w_3 x_{i+1}}{6}.$$

При усреднении ряда x_1, x_2, \dots, x_T по K периодам прогнозы на h шагов вперед приравняются последнему «текущему уровню» L_T , вычисляемому по формуле:

$$L_T = w_1 x_{T-K+1} + w_2 x_{T-K+2} + \dots + w_K x_T,$$

где $w_K \geq w_{K-1} \geq \dots \geq w_1 \geq 0$, $w_1 + w_2 + \dots + w_K = 1$.

Под **экспоненциальным сглаживанием** (*Exponential Smoothing*) понимают методику, при которой наблюдения усредняются с разными весами, и при этом значения весов, приписываемых прошлым наблюдениям, убывают экспоненциальным образом по мере «старения» наблюдений.

Простое экспоненциальное сглаживание (*SES* — *Single Exponential Smoothing*) — метод, подходящий для рядов, случайным образом флуктуирующих вокруг постоянного среднего (уровня), в которых не проявляются тренды и сезонные колебания. В этом методе «текущий уровень» ряда в i -м периоде определяется рекуррентным образом:

$$L_i = \alpha x_i + (1 - \alpha)L_{i-1},$$

где α — некоторая постоянная, $0 < \alpha < 1$, так что «текущий уровень» ряда в i -м периоде есть взвешенное среднее текущего значения ряда и «текущего уровня» ряда в $(i - 1)$ -м периоде.

Из указанной формулы получаем последовательно:

$$\begin{aligned} L_i &= \alpha x_i + (1-\alpha)L_{i-1} = \alpha x_i + (1-\alpha)(\alpha x_{i-1} + (1-\alpha)L_{i-2}) = \\ &= \alpha x_i + \alpha(1-\alpha)x_{i-1} + (1-\alpha)^2 L_{i-2} = \dots = \\ &= \alpha x_i + \alpha(1-\alpha)x_{i-1} + \alpha(1-\alpha)^2 x_{i-2} + \dots + \alpha(1-\alpha)^{i-1} x_1 + (1-\alpha)^i L_0, \end{aligned}$$

т.е. в формировании значения L_i участвуют все значения x_i, x_{i-1}, \dots, x_1 , но в соответствующую сумму они входят с экспоненциально убывающими коэффициентами.

Как и при использовании скользящего среднего, прогнозные значения на h шагов вперед приравниваются последнему «текущему уровню» L_T , так что прогнозное значение ряда на период $(T+h)$, получаемое на основании наблюдений x_1, x_2, \dots, x_T , определяется по формуле:

$$\hat{x}_{T+h} = L_T \quad \text{для любого } h = 1, 2, \dots$$

Заметим еще, что

$$\hat{x}_{T+1} = \hat{x}_T + \alpha(x_T - \hat{x}_T),$$

где \hat{x}_T — прогножное значение ряда на период T , полученное на основании наблюдений x_1, x_2, \dots, x_{T-1} .

Иными словами, прогножное значение \hat{x}_{T+1} на период $T+1$ отличается от прогнозного значения \hat{x}_T , полученного на период T , на величину, пропорциональную ошибке прогноза на период T , так что этот механизм обеспечивает адаптивный характер процедуры.

Параметр сглаживания α определяет роль последнего наблюдения в формировании «текущего уровня» ряда. Если значение α близко к 1, то влияние на «текущий уровень» всех предыдущих наблюдений оказывается незначительным. При малых значениях α «текущий уровень» определяется в большей степени прошлой динамикой ряда.

Многие специалисты рекомендуют использовать значения α в пределах 0.1—0.5, при этом аналитики большинства фирм при обработке рядов используют свои традиционные значения α . Для динамично развивающихся фирм и рынков характерны более высокие значения α , чем для более консервативных компаний и стабильных рынков; для прогнозов используют более высокие значения α , чем для анализа предшествующих тенденций.

Для реального построения прогнозных значений необходимо определить также значение L_0 . Часто в качестве него берется само x_1 .

Параметр сглаживания можно подобрать по имеющимся данным, выбрав значение, минимизирующее сумму квадратов ошибок одношаговых прогнозов, получаемых последовательным добавлением наблюдений.

ПРИМЕР 5.1.7

Применим процедуру простого экспоненциального сглаживания к ряду $x_t = IND_t$, не задавая заранее значение параметра α , а предоставляя программе самой выбрать оптимальное значение этого параметра (результаты приведены в табл. 5.1). Это приводит к значению $\alpha = 0.914$.

Таблица 5.1

Простое экспоненциальное сглаживание ряда

Sample: 1 100; Included observations: 100;

Method: Single Exponential; Original Series: IND; Forecast Series: INDSM01

Parameters:	Alpha	0.914
	Sum of Squared Residuals	127940.1
	Root Mean Squared Error	35.76871
End of Period Levels:	Mean	1687.725

В табл. 5.1:

- *Sum of Squared Residuals* — сумма квадратов расхождений между значениями исходного и сглаженного рядов:

$$SSE = \sum_{i=1}^T (x_i - s_i)^2;$$

- *Root Mean Squared Error (RMSE)* — квадратный корень из суммы квадратов расхождений, деленной на число наблюдений:

$$RMSE = \sqrt{\frac{1}{T} \sum_{i=1}^T (x_i - s_i)^2};$$

- *Mean* — «текущий уровень» в конце периода наблюдений:

$$Mean = L_T.$$

На рис. 5.16 приведены графики исходного и сглаженного ряда.

Похоже, что график сглаженного ряда получается сдвигом вправо графика исходного ряда. Чтобы проверить это, совместим график сглаженного ряда INDSM01 и сдвинутый вправо на один период времени график исходного ряда, т.е. график ряда $IND(-1)$.

Получим практически одинаковые графики (рис. 5.17), и это означает, что прогнозные значения \tilde{x}_{t+1} практически равны текущим значениям x_t . Почему так, можно понять, построив вероятностную модель порождения ряда, что мы и сделаем в дальнейшем.

Вместе с тем посмотрим, что изменит некритический выбор одного из рекомендуемых значений 0.1—0.5, а именно $\alpha = 0.2$. В этом случае получим достаточно сглаженный ряд INDSM02 (рис. 5.18), который, однако,

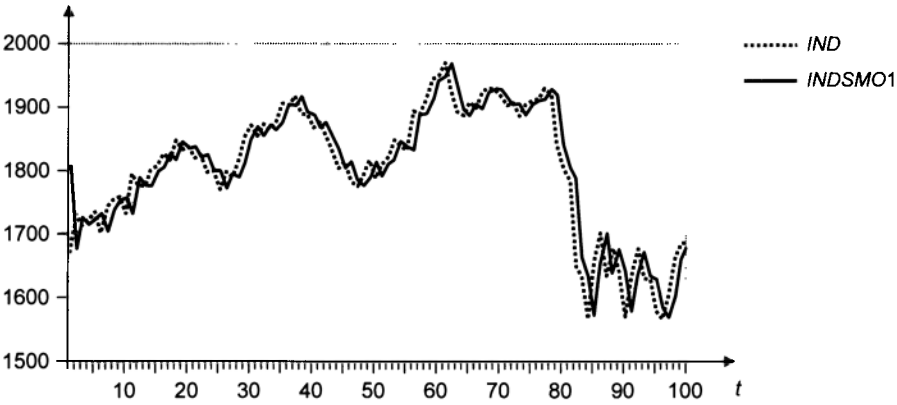


Рис. 5.16

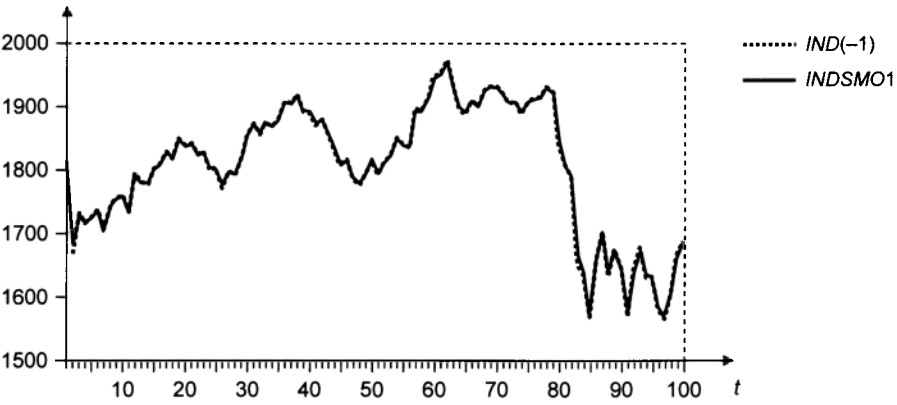


Рис. 5.17

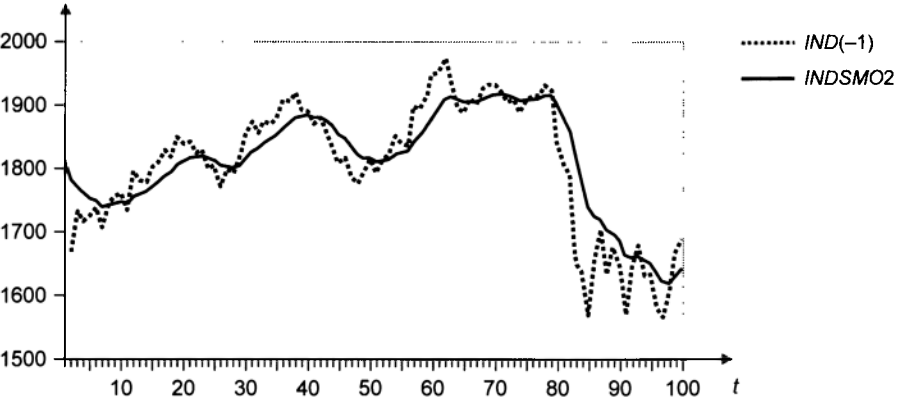


Рис. 5.18

Таблица 5.2

Простое сглаживание ряда с некритическим выбором значения α

Sample: 1 100; Included observations: 100;

Method: Single Exponential; Original Series: IND; Forecast Series: INDSM02

Parameters:	Alpha	0.2
	Sum of Squared Residuals	313564.0
	Root Mean Squared Error	55.99679
End of Period Levels:	Mean	1650.558

дает весьма плохие прогнозы (табл. 5.2): среднеквадратическая ошибка ($RMSE = 55.99679$) намного больше, чем при оцененном $\alpha = 0.914$, где $RMSE = 35.76871$. ■

Двойное экспоненциальное сглаживание, метод Брауна (*DES — Double Exponential Smoothing*) — метод, подходящий для рядов с выраженным линейным трендом, наличие которого учитывается при прогнозировании значений ряда на несколько периодов вперед.

Само название метода говорит о том, что процедура экспоненциального сглаживания здесь применяется дважды (с одним и тем же значением α): сначала — к исходному ряду, а затем — к сглаженному ряду:

$$\tilde{x}_i = \alpha x_i + (1 - \alpha) \tilde{x}_{i-1}, \quad \tilde{\tilde{x}}_i = \alpha \tilde{x}_i + (1 - \alpha) \tilde{\tilde{x}}_{i-1}.$$

При этом значение \hat{x}_{T+h} прогноза на h периодов времени вперед вычисляется по формуле:

$$\hat{x}_{T+h} = \left(2 + \frac{\alpha h}{1 - \alpha}\right) \tilde{x}_T - \left(1 + \frac{\alpha h}{1 - \alpha}\right) \tilde{\tilde{x}}_T = (2\tilde{x}_T - \tilde{\tilde{x}}_T) + \frac{\alpha}{1 - \alpha} (\tilde{x}_T - \tilde{\tilde{x}}_T) h,$$

или

$$\hat{x}_{T+h} = a + bh, \quad h = 1, 2, \dots,$$

$$\text{где } a = 2\tilde{x}_T - \tilde{\tilde{x}}_T, \quad b = \frac{\alpha}{1 - \alpha} (\tilde{x}_T - \tilde{\tilde{x}}_T).$$

Значения a и b суть параметры локального прямолинейного тренда ряда. Прогнозные значения лежат на прямой, определяемой этими параметрами.

ПРИМЕР 5.1.8

На периоде с I квартала 1994 г. по IV квартал 1997 г. количество безработных в Российской Федерации (в млн. человек) (ряд *UNEMP*) изменялось следующим образом (рис. 5.19).

Ряд имеет выраженный линейный тренд. Применим двойное экспоненциальное сглаживание для получения прогнозных значений этого ряда на 1997 г. по наблюдениям этого ряда за 1994—1996 гг. Сначала положим $\alpha = 0.2$. Результаты отражены на рис. 5.20 и в табл. 5.3.

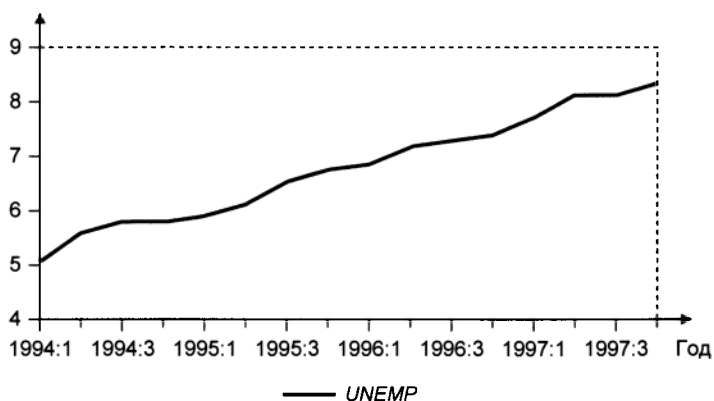


Рис. 5.19

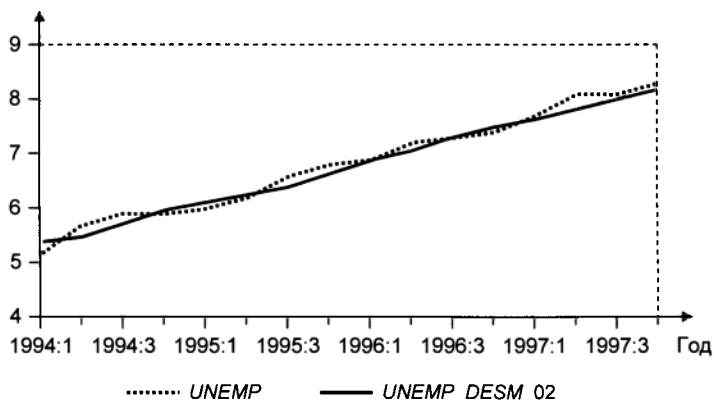


Рис. 5.20

Таблица 5.3

Двойное экспоненциальное сглаживание ряда

Sample: 1994:1 1996:4; Included observations: 12;

Method: Double Exponential; Original Series: UNEMP; Forecast Series: UNEMP_DESM_02

Parameters:	Alpha	0.2
	Sum of Squared Residuals	0.224375
	Root Mean Squared Error	0.136740
End of Period Levels:	Mean	7.257440
	Trend	0.183511

Если оставить программе подбор оптимального значения параметра α , то это приводит к значению $\alpha = 0.162$ и практически к тем же результатам (рис. 5.21 и табл. 5.4).

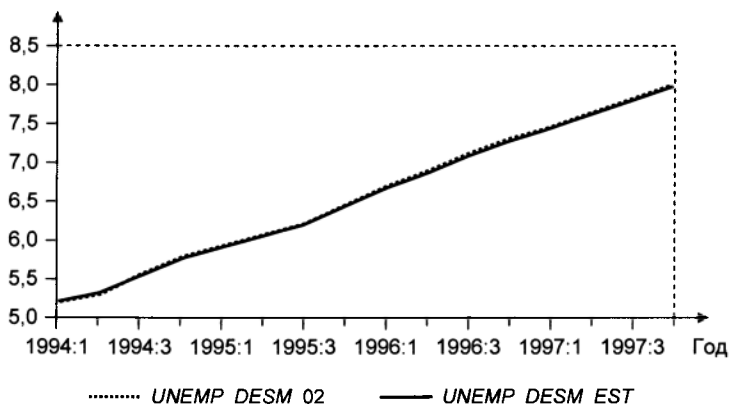


Рис. 5.21

Таблица 5.4

Двойное сглаживание ряда с подбором оптимального значения α

Sample: 1994:1 1996:4; Included observations: 12;

Method: Double Exponential; Original Series: UNEMP; Forecast Series: UNEMP_DESM_EST

Parameters:	Alpha	0.162
	Sum of Squared Residuals	0.221352
	Root Mean Squared Error	0.135816
End of Period Levels:	Mean	7.247008
	Trend	0.181239

Заметим, что если для сглаживания и прогнозирования этого ряда использовать простое экспоненциальное сглаживание с оцениваемым значением α , то это значение оценивается как $\alpha = 0.999$, и получим результаты, приведенные в табл. 5.5 и на рис. 5.22.

Таблица 5.5

Простое экспоненциальное сглаживание ряда с оцениваемым α

Sample: 1994:1 1996:4; Included observations: 12;

Method: Single Exponential; Original Series: UNEMP; Forecast Series: UNEMPMSM

Parameters:	Alpha	0.999
	Sum of Squared Residuals	1.040402
	Root Mean Squared Error	0.294449
End of Period Levels:	Mean	7.199900

Как видно на рис. 5.23, здесь прогнозы на один шаг вперед практически совпадают с текущими значениями ряда; прогнозы на несколько шагов вперед не отличаются от прогнозов на один шаг вперед. ■

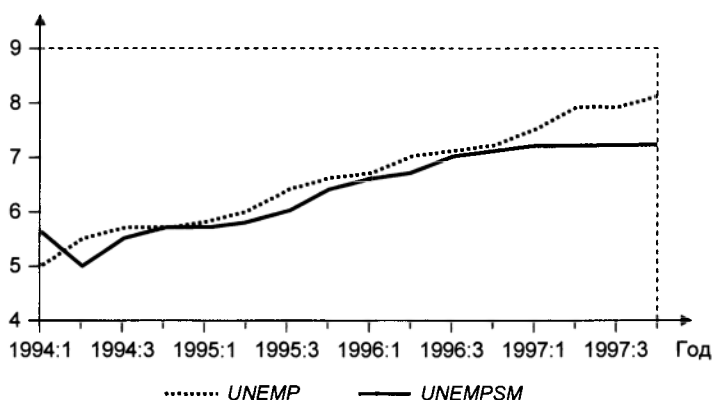


Рис. 5.22

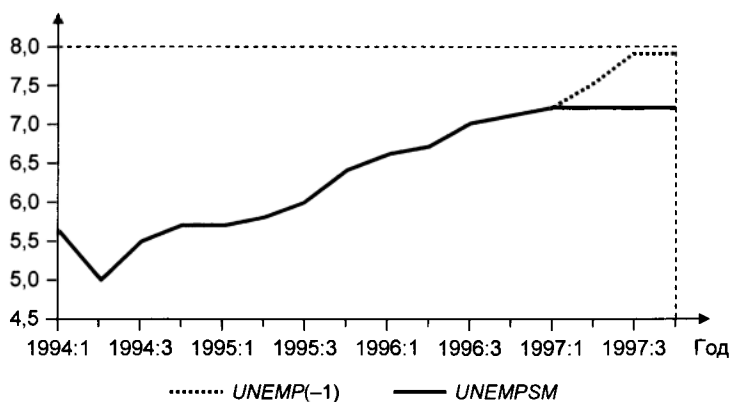


Рис. 5.23

Метод Хольта (*Holt's Linear Trend Algorithm*) используется для прогнозирования рядов с выраженным линейным трендом без сезонных составляющих. При этом прогноз осуществляется в направлении текущего локального линейного тренда, определяемого текущим уровнем ряда и текущим угловым коэффициентом локального тренда.

Пусть

L_i — текущий уровень ряда;

T_i — текущий угловой коэффициент локального тренда, соответствующие периоду i .

Эти величины вычисляются рекуррентным образом по следующим формулам:

$$L_i = \alpha x_i + (1 - \alpha)(L_{i-1} + T_{i-1}), \quad 0 < \alpha < 1,$$

$$T_i = \beta (L_i - L_{i-1}) + (1 - \beta)T_{i-1}, \quad 0 < \beta < 1.$$

Иначе говоря:

- «текущий уровень» для периода i есть взвешенное среднее значения ряда x_i в этом периоде и $(L_{i-1} + T_{i-1})$ — прогноза этого значения по локальному тренду, оцененному по предыдущим наблюдениям;
- «текущий угловой коэффициент» для периода i есть взвешенное среднее приращения «текущего уровня» $(L_i - L_{i-1})$ и значения «текущего углового коэффициента» для периода $(i - 1)$.

Соответственно значение \hat{x}_{i+h} прогноза на h периодов времени вперед, сделанного в период i , вычисляется по формуле:

$$\hat{x}_{i+h} = L_i + T_i h.$$

ПРИМЕР 5.1.9

Применим метод Хольта к ряду данных о безработице в России, предоставив компьютеру возможность выбора оптимальных значений параметров α и β (название сглаженного ряда *UNEMP_HW_EST*, предпрогнозный период 1994:1—1996:4). Результаты приведены в табл. 5.6. На рис. 5.24 изображены график ряда *UNEMP* в интервале 1994:1—1997:4, а также график сглаженного ряда *UNEMP_HW_EST* на предпрогножном периоде и сделанные по нему прогнозы на 4 квартала 1997 г.

Таблица 5.6

Метод Хольта с выбором оптимальных значений α и β
Sample: 1994:1 1996:4; Included observations: 12;
Method: Holt-Winters No Seasonal; Original Series: UNEMP;
Forecast Series: UNEMP_HW_EST

Parameters:	Alpha	1.0000
	Beta	0.0000
	Sum of Squared Residuals	0.232222
	Root Mean Squared Error	0.139111
End of Period Levels:	Mean	7.200000
	Trend	0.233333

Сравним прогнозы на 1997 г., полученные методами Брауна и Хольта (с оцениваемыми параметрами). Для этого вычислим сумму квадратов ошибок прогнозов и *RMSE* на прогнозном периоде. Непосредственный подсчет этих величин (в протоколе оценивания в пакете *EViews* они не приводятся) дает результаты, указанные в табл. 5.7 (для метода Брауна) и в табл. 5.8 (для метода Хольта). Алгоритм Хольта дает большее значение угловому коэффициенту локального тренда (рис. 5.25), при этом качество прогноза несколько лучше.■

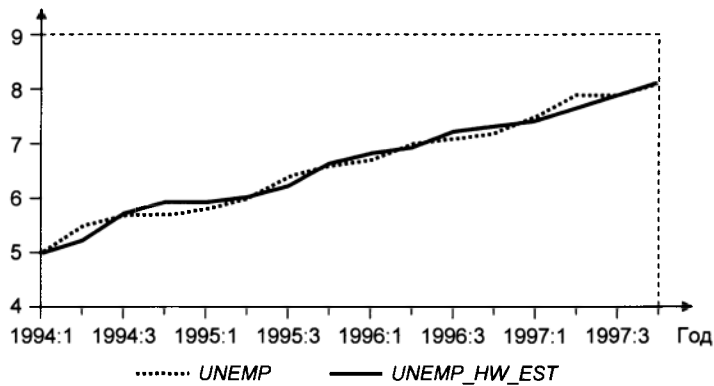


Рис. 5.24

Таблица 5.7

Прогноз, полученный с помощью метода Брауна

End of Period Levels:	Sum of Squared Residuals	0.117881
	Root Mean Squared Error	0.171667
	Mean	7.247008
	Trend	0.181239

Таблица 5.8

Прогноз, полученный с помощью метода Брауна

End of Period Levels:	Sum of Squared Residuals	0.060000
	Root Mean Squared Error	0.122474
	Mean	7.200000
	Trend	0.233333

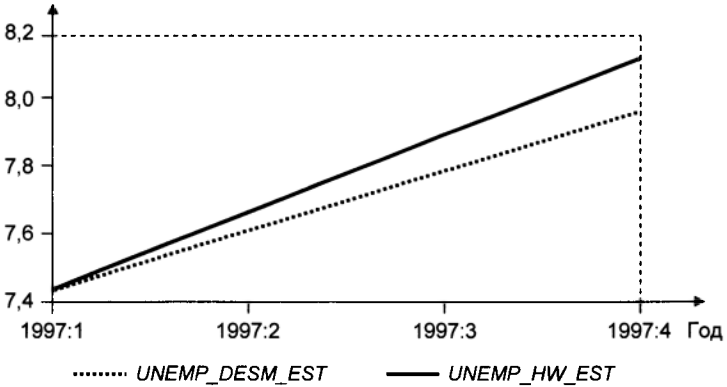


Рис. 5.25

Метод Хольта — Винтерса (Holt — Winters's Algorithm). Этот метод был предложен Винтерсом как обобщение метода Хольта, допускающее наличие сезонного фактора в динамике ряда. К паре уравнений Хольта, определяющих текущий уровень и текущий угловой коэффициент ряда, здесь добавляется третье уравнение, определяющее текущий сезонный фактор S_i . При этом сезонный фактор может быть как аддитивным, так и мультипликативным.

Если сезонный фактор **аддитивный**, то используется следующая система:

$$\begin{cases} L_i = \alpha(x_i - S_{i-s}) + (1-\alpha)(L_{i-1} + T_{i-1}), & 0 < \alpha < 1, \\ T_i = \beta(L_i - L_{i-1}) + (1-\beta)T_{i-1}, & 0 < \beta < 1, \\ S_i = \gamma(x_i - L_i) + (1-\gamma)S_{i-s}, & 0 < \gamma < 1, \end{cases}$$

где $s = 4$ для квартальных данных и $s = 12$ для месячных данных.

Значение \hat{x}_{i+h} прогноза на h периодов времени вперед, сделанного в период i , вычисляется по формуле:

$$\begin{aligned} \hat{x}_{i+h} &= L_i + T_i h + S_{i+h-s} \quad \text{для } h = 1, 2, \dots, s, \\ \hat{x}_{i+h} &= L_i + T_i h + S_{i+h-2s} \quad \text{для } h = s+1, s+2, \dots, 2s, \end{aligned}$$

и т.д. Прогноз представляет собой прямую линию, на которую накладываются аддитивные сезонные факторы.

Если сезонный фактор **мультипликативный**, то используется следующая система:

$$\begin{cases} L_i = \frac{\alpha x_i}{S_{i-s}} + (1-\alpha)(L_{i-1} + T_{i-1}), & 0 < \alpha < 1, \\ T_i = \beta(L_i - L_{i-1}) + (1-\beta)T_{i-1}, & 0 < \beta < 1, \\ S_i = \frac{\gamma x_i}{L_i} + (1-\gamma)S_{i-s}, & 0 < \gamma < 1. \end{cases}$$

Значение \hat{x}_{i+h} прогноза на h периодов времени вперед, сделанного в период i , вычисляется по формуле:

$$\begin{aligned} \hat{x}_{i+h} &= (L_i + T_i h) S_{i+h-s} \quad \text{для } h = 1, 2, \dots, s, \\ \hat{x}_{i+h} &= (L_i + T_i h) S_{i+h-2s} \quad \text{для } h = s+1, s+2, \dots, 2s, \end{aligned}$$

и т.д.

Чаще применяют вариант с мультипликативной сезонностью, поскольку в аддитивном случае сглаженный ряд имеет постоянные сезонные колебания, величина которых не зависит от общего уровня значений ряда, а в мультипликативном случае величина сезонных колебаний может изменяться в зависимости от общего уровня значений ряда. Можно, однако, применить оба варианта метода Хольта — Винтерса и выбрать из них лучший по качеству ретроспективных прогнозов.

ПРИМЕР 5.1.10

Вернемся к ряду *AGRO*, имеющему ярко выраженный сезонный характер (рис. 5.26). Применим к нему алгоритм Хольта — Винтерса с оцениваемыми параметрами. Оценивание проведем по данным до IV квартала 2003 г. включительно, а прогнозные значения вычислим на период с I квартала 2004 г. по IV квартал 2005 г. В табл. 5.9 приведены результаты применения алгоритма с аддитивной сезонностью, а в табл. 5.10 — результаты применения алгоритма с мультипликативной сезонностью.

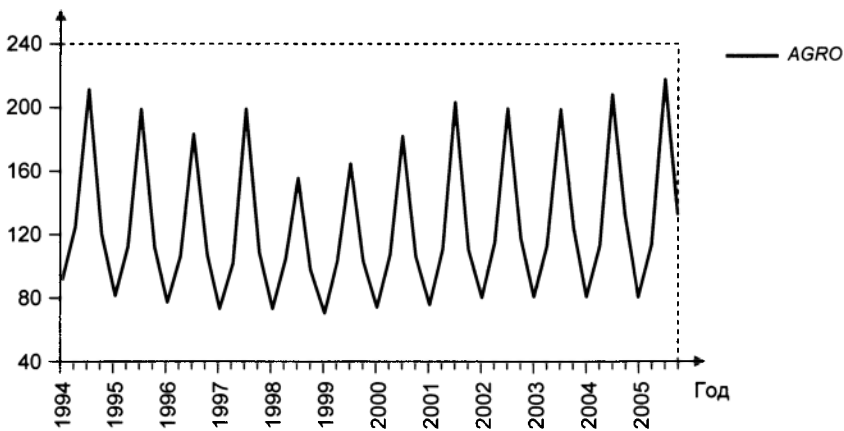


Рис. 5.26

Таблица 5.9

Применение метода Хольта — Винтерса с аддитивной сезонностью
Sample: 1994:1 2003:4; Included observations: 40;
Method: Holt-Winters Additive Seasonal; Original Series: AGRO;
Forecast Series: AGRO_HW_AD

Parameters:	Alpha		0.4600
	Beta		0.0000
	Gamma		0.0000
	Sum of Squared Residuals		2634.295
	Root Mean Squared Error		8.115255
End of Period Levels:	Mean		127.3298
	Trend		-0.201389
	Seasonals:	2003:1	-43.68208
		2003:2	-12.21069
		2003:3	66.45069
		2003:4	-10.55792

Таблица 5.10

Применение метода Хольта — Винтерса с мультипликативной сезонностью

Sample: 1994:1 2003:4; Included observations: 40;

Method: Holt-Winters Multiplicative Seasonal; Original Series: AGRO;

Forecast Series: AGRO_HW_MU

Parameters:	Alpha		0.6200
	Beta		0.0900
	Gamma		0.0000
	Sum of Squared Residuals		1808.309
	Root Mean Squared Error		6.723669
End of Period Levels:	Mean		129.8795
	Trend		1.073999
	Seasonals:	2003:1	0.627968
		2003:2	0.897011
		2003:3	1.565845
		2003:4	0.909176

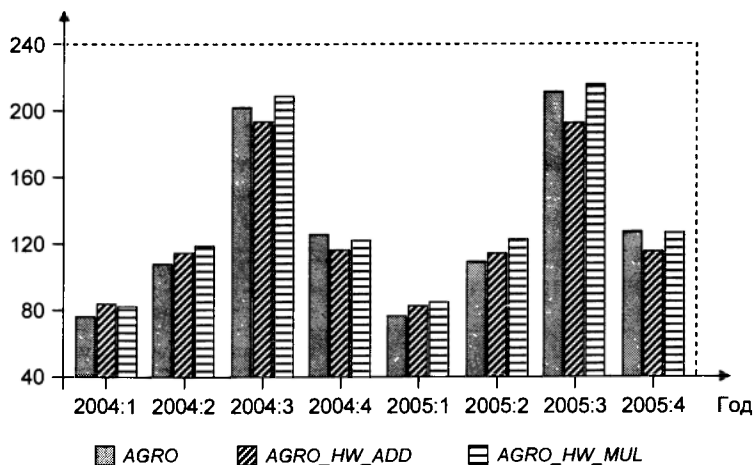


Рис. 5.27

Вариант с мультипликативной сезонностью дает лучшие результаты ретроспективных прогнозов. Диаграмма на рис. 5.27 дает возможность сравнить перспективные прогнозы по аддитивной и мультипликативной моделям. Сумма квадратов ошибок прогнозов на период с I квартала 2004 г. по IV квартал 2005 г. равна 813.75 для аддитивной модели и 489.98 для мультипликативной, так что мультипликативная модель оказалась лучшей по этому показателю и для перспективного прогнозирования. ■

Экстраполирование тренда, оцененного методом наименьших квадратов

Выше было показано, как можно получать прогнозы на основе выделения трендовых и сезонных составляющих методом сглаживания. Часто на основе визуального анализа графика ряда удается делать более определенные суждения о характере тренда. В динамике ряда на рассматриваемом периоде может явно проявляться линейный, полиномиальный или экспоненциальный характер тренда. В таких случаях прогнозирование может опираться на результаты оценивания параметров функции, описывающей тренд на всем периоде наблюдений, а само оценивание можно производить, используя метод наименьших квадратов.

ПРИМЕР 5.1.11

Ряд *UNEMP* — ряд данных о безработице в Российской Федерации на периоде с I квартала 1994 г. по IV квартал 1997 г. (квартальные данные) имеет выраженный линейный тренд (рис. 5.28).

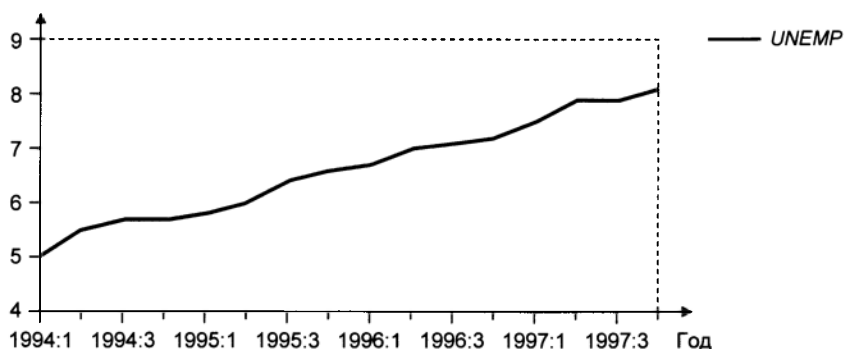


Рис. 5.28

Прогнозные значения этого ряда на 1997 г. получим по наблюдениям ряда за 1994—1996 гг., оценивая методом наименьших квадратов модель

$$UNEMP_t = \theta_1 + \theta_2 t + \varepsilon_t, \quad t = 1, \dots, 12.$$

Подобранная модель имеет вид:

$$UNEMP_t = 4.963 + 0.194t.$$

Прогнозные значения $UNEMP_t$ ряда на 1997 г. (табл. 5.11) получим по этой формуле, подставив в правую часть последовательно значения $t = 13, 14, 15, 16$. На рис. 5.29 сопоставлены наблюдаемые и прогнозные значения ряда на прогнозном периоде.

В табл. 5.12 приведены характеристики качества сделанного прогноза. ■