

Косовцева Т.Р., Беляев В.В. Технологии обработки экономической информации. Адаптивные методы прогнозирования. Учебное пособие. – СПб: Университет ИТМО, 2016, - 31 с.

Учебное пособие предназначено для магистрантов, обучающихся по направлению 38.04.01 «Экономика», изучающих дисциплину «Компьютерные технологии в экономической науке и производстве».

В пособии изложены основные принципы построения моделей. Модели могут быть использованы для анализа и краткосрочного прогноза. Все примеры разобраны в среде MS Excel. Приведен пример решения задачи с помощью статистического пакета R.

Рекомендовано к печати на заседании Ученого совета факультета технологического менеджмента и инноваций, протокол № 8 от 31 марта 2016 г.



Университет ИТМО – ведущий вуз России в области информационных и фотонных технологий, один из немногих российских вузов, получивших в 2009 году статус национального исследовательского университета. С 2013 года Университет ИТМО – участник программы повышения конкурентоспособности российских университетов среди ведущих мировых научно-образовательных центров, известной как проект «5 в 100». Цель Университета ИТМО – становление исследовательского университета мирового уровня, предпринимательского по типу, ориентированного на интернационализацию всех направлений деятельности.

© Университет ИТМО, 2016
© Косовцева Т.Р., Беляев В.В.

СОДЕРЖАНИЕ

Введение.....	4
Экспоненциальное сглаживание.....	5
Выбор оптимального значения коэффициента адаптации.....	10
Адаптивные полиномиальные модели.....	12
Модель Хольта	16
Модель Хольта-Уинтерса.....	17
Модель Тейла-Вейджа	22
Применение статистического пакета R для построения модели Хольта-Уинтерса	22
Библиографический список.....	26
Приложение 1	27

ВВЕДЕНИЕ

Одно из перспективных направлений развития краткосрочного прогнозирования связано с адаптивными методами. Эти методы позволяют строить самокорректирующиеся модели, способные оперативно реагировать на изменение условий. Адаптивные методы учитывают различную информационную ценность уровней ряда, «старение» информации. Все это делает эффективным их применение для прогнозирования неустойчивых рядов с изменяющейся тенденцией. В адаптивных методах различную ценность уровней в зависимости от их «свежести» можно учесть с помощью системы весов, придаваемых этим уровням.

Оценивание коэффициентов адаптивной модели обычно осуществляется на основе рекуррентного метода, который формально отличается от метода наименьших квадратов, метода максимального правдоподобия и других методов тем, что не требует повторения всего объема вычислений при появлении новых данных.

Важнейшим достоинством адаптивных методов является построение самокорректирующихся моделей, способных учитывать результат прогноза, сделанного на предыдущем шаге.

Пусть модель находится в некотором состоянии, для которого определены текущие значения ее коэффициентов. На основе этой модели делается прогноз. При поступлении фактического значения оценивается ошибка прогноза (разница между этим значением и полученным по модели). Ошибка прогнозирования через обратную связь поступает в модель и учитывается в ней в соответствии с принятой процедурой перехода от одного состояния в другое.

В результате вырабатываются «компенсирующие» изменения, состоящие в корректировании параметров с целью большего согласования поведения модели с динамикой ряда. Затем рассчитывается прогнозная оценка на следующий момент времени, и весь процесс повторяется вновь.

Таким образом, адаптация осуществляется итеративно с получением каждой новой фактической точки ряда. Модель постоянно «впитывает» новую информацию, приспосабливается к ней и поэтому отражает тенденцию развития, существующую в данный момент.

Скорость (быстроту) реакции модели на изменения в динамике процесса характеризует так называемый *параметр адаптации*. Параметр адаптации должен быть выбран таким образом, чтобы обеспечивалось адекватное отображение тенденции при одновременной фильтрации случайных отклонений. Значение параметра адаптации может быть определено на основе эмпирических данных, выведено аналитическим способом или получено на основе метода проб.

В качестве критерия оптимальности при выборе параметра адаптации обычно принимают критерий минимума среднего квадрата ошибок прогнозирования.

На основе рассмотренных особенностей дадим определение группы методов прогнозирования, объединенных общим названием «адаптивные».

Адаптивными называются методы прогнозирования, позволяющие строить самокорректирующиеся (самонастраивающиеся) экономико-математические модели, которые способны оперативно реагировать на изменение условий путем учета результата прогноза, сделанного на предыдущем шаге, и учета различной информационной ценности уровней ряда.

Благодаря указанным свойствам адаптивные методы особенно удачно используются при краткосрочном прогнозировании (при прогнозировании на один или на несколько шагов вперед).

Указанное определение отражает основные характерные черты, присущие рассматриваемому подходу. В то же время деление на адаптивные и неадаптивные модели часто носит достаточно условный характер.

У истоков адаптивных методов лежит модель экспоненциального сглаживания.

Модификации и обобщения этой модели привели к появлению целого семейства адаптивных моделей с различными свойствами.

ЭКСПОНЕНЦИАЛЬНОЕ СГЛАЖИВАНИЕ

Метод экспоненциального сглаживания может быть использован как для сглаживания уровней временного ряда, так и для прогнозирования.

Особенность этого метода заключается в том, что в процедуре выравнивания каждого наблюдения используются только значения предыдущих уровней, взятых с определенным весом. Вес каждого наблюдения уменьшается по мере его удаления от момента, для которого определяется сглаживаемое значение.

Сглаженное значение наблюдения ряда S_t на момент времени t определяется по формуле:

$$S_t = \alpha \cdot y_t + (1 - \alpha) \cdot S_{t-1} \quad (1)$$

где α - сглаживающий параметр, характеризующий вес выравниваемого наблюдения, причем $0 < \alpha < 1$.

Обозначим $\beta = 1 - \alpha$. Тогда формулу (1) можно записать в виде (2).

$$S_t = \alpha \cdot y_t + \beta \cdot S_{t-1} \quad (2)$$

Величину S_{t-1} в формуле (2) можно представить в виде суммы фактического значения уровня y_{t-1} и сглаженного значения предшествующего ему наблюдения S_{t-2} , взятых с соответствующими

весами. Процесс такого разложения можно продолжить для членов S_{t-2} , S_{t-3} и т.д. В результате получится следующее выражение:

$$\begin{aligned}
 S_t &= \alpha \cdot y_t + \beta \cdot S_{t-1} = \alpha \cdot y_t + \beta \cdot \{\alpha \cdot y_{t-1} + \beta \cdot S_{t-2}\} = \\
 &= \alpha \cdot y_t + \beta \cdot \alpha \cdot y_{t-1} + \beta^2 \cdot \{\alpha \cdot y_{t-2} + \beta \cdot S_{t-3}\} = \\
 &= \alpha \cdot y_t + \alpha \cdot \beta \cdot y_{t-1} + \alpha \cdot \beta^2 \cdot y_{t-2} + \dots + \alpha \cdot \beta^k \cdot y_{t-k} + \dots \beta^t \cdot y_0
 \end{aligned} \tag{3}$$

Таким образом, сглаженное значение является взвешенной суммой всех предшествующих уровней ряда. Величина y_0 характеризует начало условия процесса.

Формулу (3) можно переписать короче через знак суммы:

$$S_t = \alpha \cdot \sum_{k=1}^{t-1} \beta^k \cdot y_{t-k} + \beta^t \cdot y_0 \tag{4}$$

где $t > 0$.

Проанализируем соотношение (4). Множитель β^k , стоящий перед y_{t-k} в каждом слагаемом, является относительным весом, который определяет величину вклада соответствующего уровня ряда в общую сумму. Поскольку $0 < \alpha < 1$, то и $0 < \beta < 1$, то с увеличением k значение β^k уменьшается.

Относительный вес каждого предшествующего уровня снижается по экспоненте по мере его удаления от момента, для которого вычисляется сглаженное значение, т.е. от давности наблюдения (отсюда произошло название этого метода сглаживания).

Пусть $\alpha = 0.25$, тогда $\beta = 0.75$. Тогда:

- вес слагаемого y_{t-1} будет равен $\alpha \cdot \beta = 0.25 \cdot 0.75 = 0.1875$;
- вес слагаемого y_{t-2} будет равен $\alpha \cdot \beta^2 = 0.25 \cdot 0.75^2 = 0.1406$;
- последующие значения приведены на рис.1.

На рис.2 показано экспоненциальное убывание веса соответствующего уровня ряда по мере удаления от текущего момента времени t .

	A	B	C	D	E
1				Слагаемое	
2	alpha	beta		№	Вес
3	0.25	0.75		1	0.1875
4				2	0.140625
5				3	0.105469
6				4	0.079102
7				5	0.059326
8				6	0.044495
9				7	0.033371
10				8	0.025028

Рис.1. Расчет весов уровня ряда

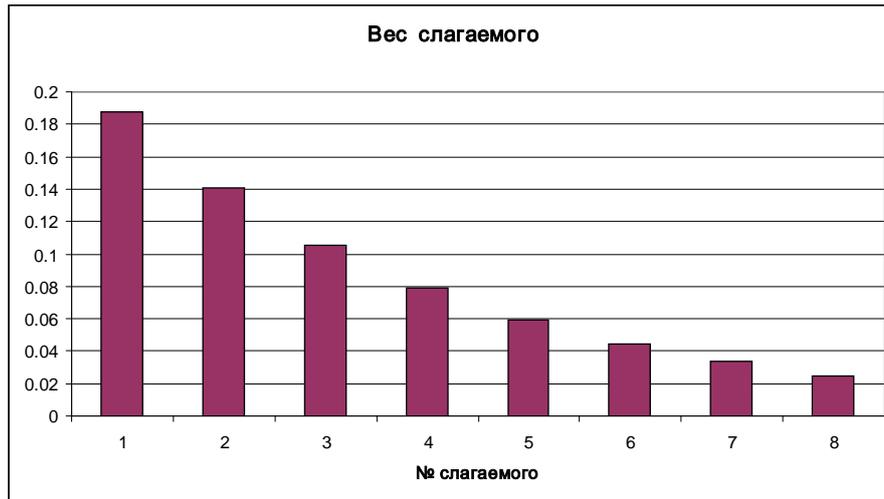


Рис. Изменение веса уровня ряда по мере его «старения».

Заметим, что слагаемое в (4) $\beta^t \cdot y_0$ при $t \rightarrow \infty$ также стремится к нулю, т.е. вклад начального условия также уменьшается по экспоненциальному закону. Вес, приписываемый этому значению, уменьшается по экспоненциальной зависимости по мере удаления от первого уровня. Поэтому для длинных временных рядов влияние неудачного выбора S_0 погашается.

При практическом использовании метода экспоненциального сглаживания возникают следующие затруднения: выбор сглаживающего параметра α и определение начального условия y_0 .

От численного значения параметра α зависит, насколько быстро будет уменьшаться вес предшествующих наблюдений и в соответствии с этим степень их влияния на сглаживаемый уровень. Чем больше значение параметра α , тем меньше сказывается влияние предшествующих уровней и соответственно меньшим оказывается сглаживающее воздействие экспоненциальной средней.

Задачу выбора параметра y_0 , определяющего начальные условия, предлагается решать следующим образом: если есть данные о развитии процесса в прошлом, то их среднее значение можно принять в качестве y_0 , если таких сведений нет, то в качестве y_0 используют исходное (первое) значение наблюдения временного ряда y_1 .

Английский математик Р. Браун показал, что математические ожидания временного ряда и экспоненциальной средней совпадут, но в то же время дисперсия экспоненциальной средней $D[S_t]$ меньше дисперсии временного ряда (σ_0^2):

$$D[S_t] = \frac{\alpha}{2 - \alpha} \cdot \sigma_0^2 \quad (5)$$

Из (5) видно, что при высоком значении α дисперсия экспоненциальной средней незначительно отличается от дисперсии ряда.

С уменьшением α дисперсия экспоненциальной средней сокращается, возрастает ее отличие от дисперсии ряда. Тем самым, экспоненциальная средняя начинает играть роль «фильтра», поглощающего колебания временного ряда.

Таким образом, с одной стороны, следует увеличивать вес более свежих наблюдений, что может быть достигнуто повышением α , согласно (2), с другой стороны, для сглаживания случайных отклонений величину α нужно уменьшить.

Эти два требования находятся в противоречии. Поиск компромиссного значения параметра сглаживания α составляет задачу оптимизации модели. Иногда поиск этого значения параметра осуществляется путем перебора. В этом случае в качестве оптимального выбирается то значение α , при котором получена наименьшая дисперсия ошибки.

При использовании экспоненциальной средней для краткосрочного прогнозирования предполагается, что модель ряда имеет вид:

$$y_t = a_{1,t} + \varepsilon_t \quad (6)$$

где $a_{1,t}$ - изменяемый во времени средний уровень ряда;

ε_t - случайные неавтокоррелированные отклонения с нулевым математическим ожиданием и дисперсией σ^2 .

Обозначим через $\hat{y}_\tau(t)$ прогноз, сделанный в момент t на τ - единиц времени (шагов) вперед;

Тогда для модели экспоненциального сглаживания.

$$\hat{y}_\tau(t) = \hat{a}_{1,t} \quad (7)$$

Вычисленная величина параметра $\hat{a}_{1,t}$ является оценкой существующей, но неизвестной величины $a_{1,t}$.

Параметр модели $\hat{a}_{1,t}$ определяется экспоненциальной средней

$$\hat{a}_{1,t} = S_t \quad (8)$$

$$S_0 = \hat{a}_{1,0} \quad (9)$$

Соотношение (1) после элементарного преобразования может быть записано в виде

$$S_t = S_{t-1} + \alpha \cdot (y_t - S_{t-1}) \quad (10)$$

Полагая, что величина $(y_t - S_{t-1})$ есть погрешность прогноза сделанного в момент времени $t-1$, на момент времени t , получаем величину нового прогноза (на момент времени $t+1$), как результат корректировки предыдущего прогноза с учетом его ошибки. В этом заключается адаптация модели.

В качестве ошибки модели будем использовать сумму квадратов ошибок (Sum Square Error - SSE), вычисляемый по формуле:

$$SSE = \sum_{t=1}^n \varepsilon_t^2 \quad (11)$$

Очевидно, что на ошибку модели SSE влияет величина параметра сглаживания α , которая может быть вычислена как решение соответствующей оптимизационной задачи. Соответствующая оптимизационная задача рассматривается в примерах, приведенных ниже.

Пример

Пусть нам известен набор данных, характеризующий количество пятен на Солнце за 25 месяцев с января 1981 года по январь 1983 года:

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1981	114.0	141.3	135.5	156.4	127.5	90.0	143.8	158.7	167.3	162.4	137.5	150.1
1982	111.2	163.6	153.8	122.0	82.2	110.4	106.1	107.6	118.8	94.7	98.1	127.0
1983	84.3											

Приняв коэффициент адаптации (сглаживающий параметр) α равным 0.5, сделать прогноз на февраль 1983 года используя модели экспоненциальной средней.

Решение

Всё решение приведено на рис.3 - 5.

Разместим исходные данные в таблице MS Excel в диапазоне A2:E28. Для построения модели скользящего среднего по формулам 1, 8, 9, необходимо вычислить S_0 - начальное значение S_t .

В качестве S_0 возьмем среднее арифметическое первых 5 значений временного ряда Y_t ,

$$S_0 = \frac{1}{5} \cdot \sum_{t=1}^5 y_t - \text{это значение равно } 134,940 \text{ и находится в ячейке F3.}$$

Далее в интервале F4: F28 вычисляем значения S_t для $t = 1, 2, \dots, 25$. Это и есть сглаженные значения уровней ряда.

Графики исходного временного ряда Y_t и сглаженных значений S_t приведен на рис.5.

Прогнозные значения $\hat{y}_t(t)$ получаются сдвигом на единицу значений S_t (интервал ячеек G4:G29). Искомое прогнозное значение $\hat{y}_1(25)$ т.е. количество пятен на Солнце в феврале 1983, оказывается в ячейке G29.

Таким образом, искомое прогнозное значение равно 99 (99.162 округлено до целых). Это значение при $t = 26$ также нанесено на графике (на рис.5).

Ошибка модели SSE равна 15008 и содержится в ячейке Н30.

	A	B	C	D	E	F	G	H
1	Год	Месяц		№№	yt	St	yt - прогн.	Ошибка^2
2			1	0		134.940		
3	1981	январь	1	1	114.0	124.470	134.940	438.484
4		февраль	1	2	141.3	132.885	124.470	283.249
5		март	1	3	135.5	134.193	132.885	6.838
6		апрель	1	4	156.4	145.296	134.193	493.173
7		май	1	5	127.5	136.398	145.296	316.707
24		октябрь	1	22	94.7	103.993	113.287	345.463
25		ноябрь	1	23	98.1	101.047	103.993	34.731
26		декабрь	1	24	127.0	114.023	101.047	673.576
27	1983	январь	1	25	84.3	99.162	114.023	883.476
28		февраль	1	26			99.162	
29							$SSE=$	15008.262
30								
31				alpha	0.5			
32				beta	0.5			

Рис.3. Режим отображения данных

	A	B	D	E	F	G	H
1	Год	Месяц	№№	yt	St	yt - прогн.	Ошибка^2
2			0		=СУММ(E3:E7)/5		
3	1981	январь	1	114	=\$E\$31*E3+(1-\$E\$31)*F2	=F2	=(E3-G3)^2
4		февраль	2	141.3	=\$E\$31*E4+(1-\$E\$31)*F3	=F3	=(E4-G4)^2
5		март	3	135.5	=\$E\$31*E5+(1-\$E\$31)*F4	=F4	=(E5-G5)^2
6		апрель	4	156.4	=\$E\$31*E6+(1-\$E\$31)*F5	=F5	=(E6-G6)^2
27	1983	январь	25	84.3	=\$E\$31*E27+(1-\$E\$31)*F26	=F26	=(E27-G27)^2
28		февраль	26			=F27	
29						$SSE=$	=СУММ(H3:H27)
30							
31			alpha	0.5			
32			beta	=1-E31			

Рис.4. Режим отображения формул

ВЫБОР ОПТИМАЛЬНОГО ЗНАЧЕНИЯ КОЭФФИЦИЕНТА АДАПТАЦИИ

Качество модели определяется величиной ошибки, чем она ниже тем качественнее построенная модель. Величина ошибки SSE зависит от величины коэффициента адаптации α , поэтому существует возможность подбора α таким образом, чтобы ошибка была минимальной.

Для определения оптимальной величины α составим задачу нелинейного программирования.

В качестве целевой функции принимаем соотношение $SSE(\alpha) \rightarrow \min$, в качестве ограничений $0 \leq \alpha \leq 1$. Оптимизацию осуществим с помощью надстройки «Поиск решения». Решение приведено на рис.6 - 7.



Рис.5. Графики исходного временного ряда Y_t и сглаженных значений S_t

При α равном 0.306, величина ошибки модели SSE равна 14 394 (что на $(15\ 008 - 14\ 394) / 14\ 394 \times 100\% = 4.2\%$ меньше, чем ошибка прогноза при α равном 0.5.

Значение прогноза на февраль 1983 года равно 103 (103,47 округлено до целых).

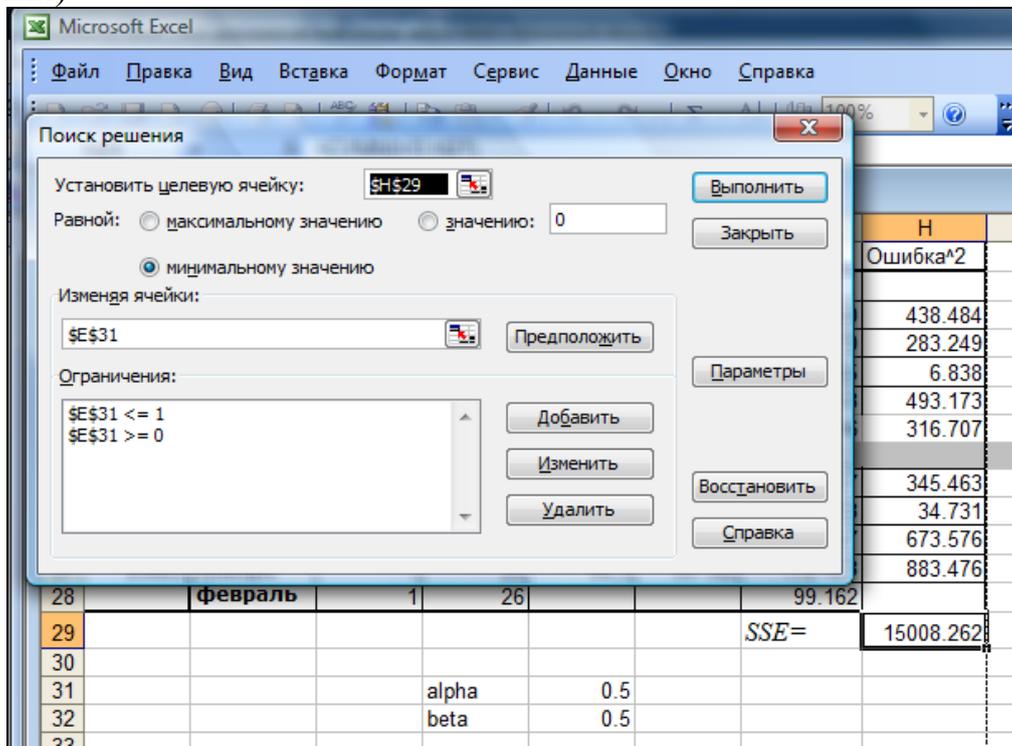


Рис.6. Решение с помощью надстройки «Поиск решения». Ввод условий задачи. Режим отображения данных

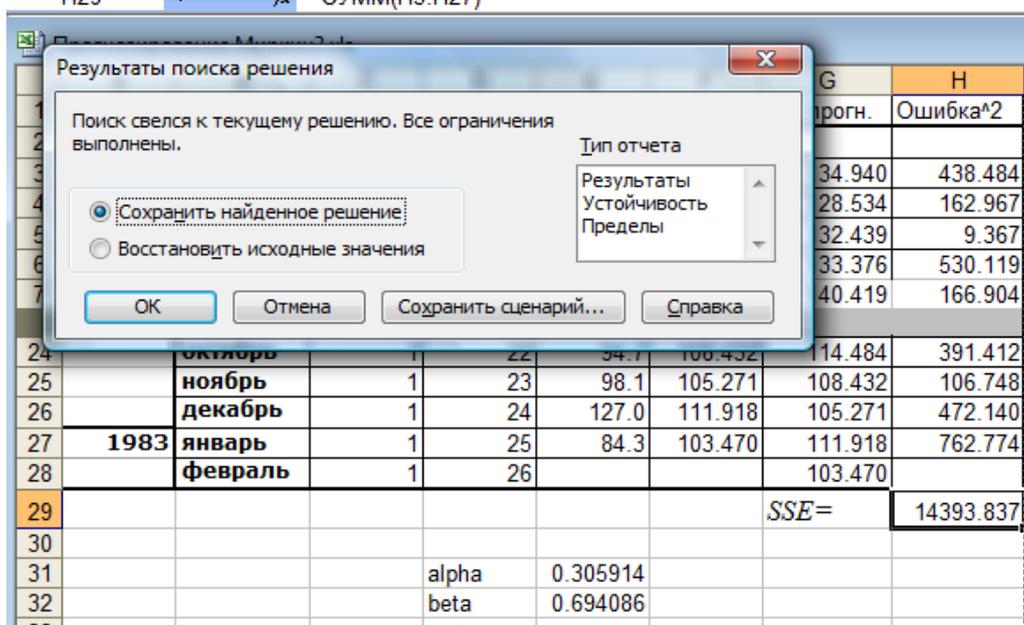


Рис.7. Решение с помощью надстройки «Поиск решения». Вывод результата.

АДАПТИВНЫЕ ПОЛИНОМИАЛЬНЫЕ МОДЕЛИ

Понятие экспоненциальной средней, которое рассматривалось ранее, можно обобщить в случае экспоненциальных средних более высоких порядков.

После того, как были вычислены экспоненциальные средние S_t по формуле (1), они сами образуют временной ряд. Обозначим его члены через $S_t^{(1)}$. Верхний индекс показывает, что это результат первого сглаживания. Этот ряд можно сгладить еще раз, воспользовавшись формулой, аналогичной формуле (1):

$$S_t^{(2)} = \alpha \cdot S_t^{(1)} + \beta \cdot S_{t-1}^{(2)} \quad (12)$$

Полученный сглаженный ряд $S_t^{(2)}$ по отношению к исходному является дважды сглаженным исходным, поэтому его называют экспоненциальной средней второго порядка.

Экспоненциальная средняя $S_t^{(k)}$ порядка k определяется соотношением:

$$S_t^{(k)} = \alpha \cdot S_t^{(k-1)} + \beta \cdot S_{t-1}^{(k)} \quad (13)$$

Если тренд исследуемого процесса описывается полиномом степени p , то прогноз на τ - шагов вперед осуществляется по формуле (14).

$$\hat{y}_\tau(t) = \hat{a}_1 + \hat{a}_2 \cdot \tau + \hat{a}_2 \cdot \tau^2 + \hat{a}_3 \cdot \tau^3 + \dots + \hat{a}_{p+1} \cdot \tau^p \quad (14)$$

где $\hat{a}_1, \hat{a}_2, \hat{a}_2, \hat{a}_3, \dots, \hat{a}_{p+1}$ - оценки параметров.

Р. Брауном и Р. Майером доказана фундаментальная теорема метода экспоненциального сглаживания и прогнозирования, которая утверждает, что $(p + 1)$ неизвестных коэффициентов полинома p -го порядка

$\hat{a}_1, \hat{a}_2, \hat{a}_2, \hat{a}_3, \dots, \hat{a}_{p+1}$ могут быть выражены через линейные комбинации экспоненциальных средних $S_t^{(k)}$, где $k = 1, 2, \dots, p + 1$.

Следовательно, для проведения прогноза с помощью этого метода необходимо вычислить экспоненциальные средние $S_t^{(k)}$, порядок которых изменяется от 1 до $(p + 1)$, а затем через их линейные комбинации — k определить коэффициенты полинома. Далее, пользуясь этим полиномом, произвести прогноз.

На практике обычно используются полиномы не выше второго порядка.

В таблице 2 приведены расчетные формулы.

Процедура прогнозирования временных рядов на основе адаптивных полиномиальных моделей **состоит из следующих этапов**.

1. Выбирается вид модели экспоненциального сглаживания, задается значение параметра сглаживания α . При выборе порядка адаптивной полиномиальной модели могут использоваться различные подходы, например, графический анализ, метод последовательных разностей и др.

2. Определяются начальные условия. Например, для полиномиальной модели первого порядка необходимо определить $\hat{a}_{1,0}$ и $\hat{a}_{2,0}$. Чаще всего в качестве этих оценок берут коэффициенты соответствующих полиномов, полученные методом наименьших квадратов. Начальные условия для модели нулевого порядка обычно получают усреднением нескольких первых уравнений ряда. Зная эти оценки, с помощью указанных в таблице формул, находят начальные значения экспоненциальных средних. Например, для полиномиальной модели первого порядка $S_0^{(1)}$ и $S_0^{(2)}$

3. Производится расчет значений соответствующих экспоненциальных средних $S_t^{(1)}$, $S_t^{(2)}$ и т.д. до $S_t^{(p)}$.

4. Находятся оценки коэффициентов модели $\hat{a}_1, \hat{a}_2, \hat{a}_2, \hat{a}_3, \dots, \hat{a}_{p+1}$.

5. Осуществляется прогноз на одну точку вперед, находится отклонение фактического значения временного ряда от прогнозируемого. Шаги с 3 по 5 данной процедуры повторяются для всех для $t = 1, 2, \dots, n$, где n - длина ряда.

6. Окончательная прогнозная модель формируется на последнем шаге в момент $t = n$. Прогноз получается на базе выражения (14) путем подстановки в него последних значений коэффициентов и времени упреждения τ .

Положительной особенностью рассмотренных моделей следует считать то, что при поступлении новой, свежей информации все расчеты повторять нет необходимости. Достаточно принять в качестве начальных

условий последние значения функций сглаживания $S_t^{(1)}$, $S_t^{(2)}$ и т.д. до $S_t^{(p)}$ и повторить расчеты в п.п. 3-6 один раз.

Вычисление по этим моделям можно улучшить за счет выбора параметра сглаживания α . Этот выбор можно сделать или методом сплошного перебора или путем решения соответствующей оптимизационной задачи, целью которой является уменьшение ошибки модели E .

Таблица 2

Основные формулы для прогнозирования по адаптивным полиномиальным моделям

Степень модели	Начальные условия	Экспоненциальные средние
n=0	$S_0^{(1)} = \hat{a}_{1,0}$	$S_t^{(1)} = \alpha \cdot y_t + \beta \cdot S_{t-1}^{(1)}$
n=1	$S_0^{(1)} = \hat{a}_{1,0} - \frac{\beta}{\alpha} \cdot \hat{a}_{2,0}$ $S_0^{(2)} = \hat{a}_{1,0} - \frac{2\beta}{\alpha} \cdot \hat{a}_{2,0}$	$S_t^{(1)} = \alpha \cdot y_t + \beta \cdot S_{t-1}^{(1)}$ $S_t^{(2)} = \alpha \cdot S_t^{(1)} + \beta \cdot S_{t-1}^{(2)}$
n=2	$S_0^{(1)} = \hat{a}_{1,0} - \frac{\beta}{\alpha} \cdot \hat{a}_{2,0} + \frac{\beta \cdot (2 - \alpha)}{2\alpha^2} \cdot \hat{a}_{3,0}$ $S_0^{(2)} = \hat{a}_{1,0} - \frac{2\beta}{\alpha} \cdot \hat{a}_{2,0} + \frac{\beta \cdot (3 - 2\alpha)}{\alpha^2} \cdot \hat{a}_{3,0}$ $S_0^{(3)} = \hat{a}_{1,0} - \frac{3\beta}{\alpha} \cdot \hat{a}_{2,0} + \frac{3\beta \cdot (4 - 3\alpha)}{2\alpha^2} \cdot \hat{a}_{3,0}$	$S_t^{(1)} = \alpha \cdot y_t + \beta \cdot S_{t-1}^{(1)}$ $S_t^{(2)} = \alpha \cdot S_t^{(1)} + \beta \cdot S_{t-1}^{(2)}$ $S_t^{(3)} = \alpha \cdot S_t^{(2)} + \beta \cdot S_{t-1}^{(3)}$
Степень модели	Оценка коэффициентов	Модель прогноза
n=0	$\hat{a}_{1,t} = S_t^{(1)}$	$\hat{y}_t(t) = \hat{a}_{1,t}$
n=1	$\hat{a}_{1,t} = 2 \cdot S_t^{(1)} - S_t^{(2)}$ $\hat{a}_{1,t} = \frac{\alpha}{\beta} \cdot [S_t^{(1)} - S_t^{(2)}]$	$\hat{y}_t(t) = \hat{a}_{1,t} + \tau \cdot \hat{a}_{2,t}$
n=2	$\hat{a}_{1,t} = 3 \cdot (S_t^{(1)} - S_t^{(2)}) + S_t^{(3)}$ $\hat{a}_{2,t} = \frac{\alpha}{2\beta^2} \cdot [(6 - 5\alpha) \cdot S_t^{(1)} - 2(5 - 4\alpha) \cdot S_t^{(2)} + (4 - 3\alpha) \cdot S_t^{(3)}]$ $\hat{a}_{3,t} = \frac{\alpha^2}{\beta^2} \cdot [S_t^{(1)} - 2 \cdot S_t^{(2)} + S_t^{(3)}]$	$\hat{y}_t(t) = \hat{a}_{1,t} + \tau \cdot \hat{a}_{2,t} + \frac{1}{2} \tau^2 \cdot \hat{a}_{3,t}$

Модель Брауна имеют массу достоинств:

- логичная и ясная идея построения модели;
- простота компьютерной реализации метода;
- возможность оптимизации модели путем подбора параметра адаптации.

Недостатком этой модели является использование единственного параметра адаптации α . Это снижает диапазон применения, поскольку

невозможно описать все многообразие рядов с помощью одной изменяемой величины.

Общим недостатком всех адапционных моделей является невозможность использования другой информации об исследуемом явлении, кроме самих значений временного ряда.

Пример (продолжение)

Рассмотрим применение модели Брауна первого порядка ($p=1$) на том же примере. Решение приведено на рис.8-9, которое осуществлено по предложенному выше плану. Оптимальное значение параметра адаптации находится в ячейке E37. В качестве результата прогноза на февраль 1983 года получено значение 95,7, которое находится в ячейке J35.

	A	B	C	D	E	F	G	H	I	J	K
1	Количество солнечных пятен, единиц в месяц										
2			ЛИНЕЙН()		4.21	122.31					
3			по первым 5 уровням		5.221401	17.31743					
4					0.178108	16.51152					
5					0.650115	3					
6					177.241	817.891					
7											
8	Год	Месяц	№№	yt	St	St[2]	tau	at[1]	at[2]	yt - прогн.	Ошибка^2
9				0	106.624	90.938		122.310	4.210		
10	1981	январь		1	114.0	108.185	1	121.782	3.649	126.520	156.750
11		февраль		2	141.3	115.192	1	131.437	4.360	125.432	251.808
12		март		3	135.5	119.489	1	135.684	4.347	135.796	0.088
30		сентябрь		21	118.8	117.942	1	109.859	-2.170	104.415	206.929
31		октябрь		22	94.7	113.024	1	102.774	-2.751	107.689	168.714
32		ноябрь		23	98.1	109.866	1	99.295	-2.837	100.022	3.696
33		декабрь		24	127.0	113.492	1	108.016	-1.470	96.458	932.835
34	1983	январь		25	84.3	107.315	1	98.127	-2.466	106.546	494.878
35		февраль		26			1			95.662	
36										SSE=	15668.683
37				alpha	0.211599					MSE=	626.747
38				beta	0.788401						

Рис.8. Лист Excel в режиме отображения данных

	A	B	C	D	E	F
2					=ЛИНЕЙН(D10:D14;C10:C14;1;1)	=ЛИНЕЙН(D10:D14;C10:C14;1;1)
3					=ЛИНЕЙН(D10:D14;C10:C14;1;1)	=ЛИНЕЙН(D10:D14;C10:C14;1;1)
4					=ЛИНЕЙН(D10:D14;C10:C14;1;1)	=ЛИНЕЙН(D10:D14;C10:C14;1;1)
5					=ЛИНЕЙН(D10:D14;C10:C14;1;1)	=ЛИНЕЙН(D10:D14;C10:C14;1;1)
6					=ЛИНЕЙН(D10:D14;C10:C14;1;1)	=ЛИНЕЙН(D10:D14;C10:C14;1;1)
7						
8	Год	Месяц	№№	yt	St[1]	St[2]
9			0		=F2-\$E\$38/\$E\$37*E2	=F2-2*\$E\$38/\$E\$37*E2
10	1981	январь	1	114	=\$E\$37*D10+(1-\$E\$37)*E9	=\$E\$37*E10+(1-\$E\$37)*F9
11		февраль	2	141	=\$E\$37*D11+(1-\$E\$37)*E10	=\$E\$37*E11+(1-\$E\$37)*F10
33		декабрь	24	127	=\$E\$37*D33+(1-\$E\$37)*E32	=\$E\$37*E33+(1-\$E\$37)*F32
34	1983	январь	25	84.3	=\$E\$37*D34+(1-\$E\$37)*E33	=\$E\$37*E34+(1-\$E\$37)*F33
35		февраль	26			
36						
37				alpha	0.5	
38				beta	=1-E37	

Рис.9. Лист Excel в режиме отображения формул (начало)

	G	H	I	J	K
7					
8	tau	at[1]	at[2]	yt - прогн.	Ошибка^2
9		=2*E9-F9	=\$E\$37/\$E\$38*(E9-F9)		
10	1	=2*E10-F10	=\$E\$37/\$E\$38*(E10-F10)	=H9+I9*G10	=(D10-J10)^2
11	1	=2*E11-F11	=\$E\$37/\$E\$38*(E11-F11)	=H10+I10*G11	=(D11-J11)^2
33	1	=2*E33-F33	=\$E\$37/\$E\$38*(E33-F33)	=H32+I32*G33	=(D33-J33)^2
34	1	=2*E34-F34	=\$E\$37/\$E\$38*(E34-F34)	=H33+I33*G34	=(D34-J34)^2
35	1			=H34+I34*G35	
36				SSE=	=СУММ(K10:K34)
37				MSE=	=CPЗНАЧ(K10:K34)

Рис.10. Лист Excel в режиме отображения формул (окончание)

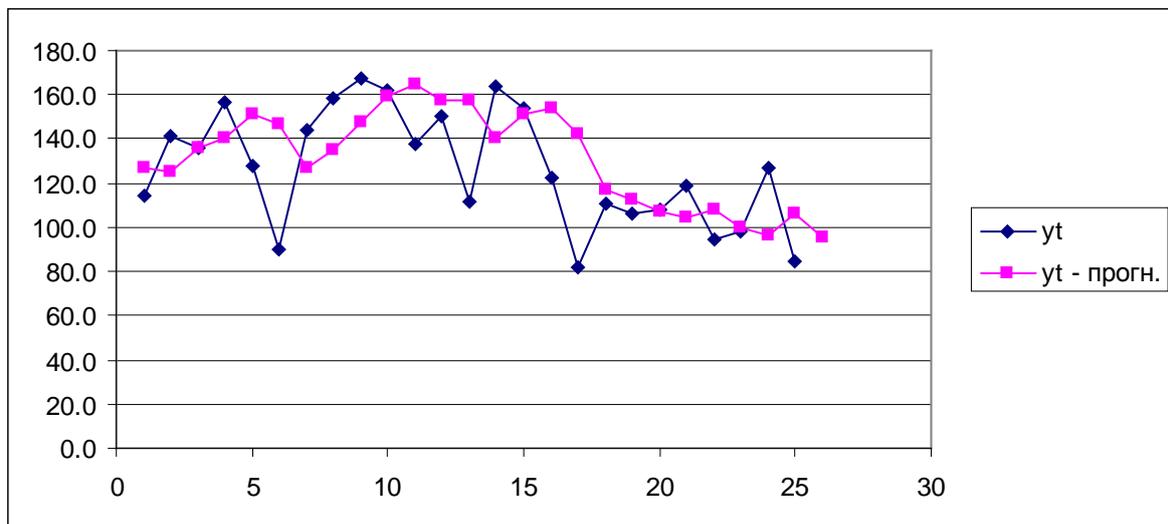


Рис.11.

МОДЕЛЬ ХОЛЬТА

Двухпараметрическая модель Хольта является развитием модели Брауна для полинома первой степени.

Прогноз по модели Хольта на τ шагов вперед определяется выражением:

$$\hat{y}_\tau(t) = (\hat{a}_{1,t} + \tau \hat{a}_{2,t}) \quad (15)$$

Обновление коэффициентов осуществляется следующим образом:

$$\hat{a}_{1,\tau} = \alpha_1 y_t + (1 - \alpha_1)(\hat{a}_{1,t-1} + \hat{a}_{2,t-1})$$

$$\hat{a}_{2,\tau} = \alpha_2 (\hat{a}_{1,t} - \hat{a}_{1,t-1}) + (1 - \alpha_2)\hat{a}_{2,t-1},$$

где α_1, α_2 - параметры адаптации, $0 < \alpha_1, \alpha_2 < 1$;

$\hat{a}_{1,t}$ и $\hat{a}_{2,t}$ - характеристики тенденции развития;

Каждый параметр принадлежит интервалу $[0; 1]$, чем ближе параметр к единице, тем больший вес приписывается последним наблюдениям.

Очевидно, что в этой модели прогнозные оценки зависят от прошлых и текущих уровней временного ряда, начальных значений коэффициентов $\hat{a}_{1,0}$ и $\hat{a}_{2,0}$, и параметров адаптации α_1, α_2 .

МОДЕЛЬ ХОЛЬТА-УИНТЕРСА

Метод Хольта-Уинтерса является модификацией метода экспоненциального сглаживания для сезонных рядов. В качестве модели ряда используется его представление в виде комбинации линейного тренда с сезонной составляющей, наложенной мультипликативно.

Прогноз по модели Хольта-Уинтерса на τ шагов вперед определяется выражением:

$$\hat{y}_\tau(t) = (\hat{a}_{1,t} + \tau \hat{a}_{2,t}) \hat{f}_{t-l+\tau} \quad (16)$$

Обновление коэффициентов осуществляется следующим образом:

$$\hat{a}_{1,t} = \alpha_1 \frac{y_t}{\hat{f}_{t-l}} + (1 - \alpha_1)(\hat{a}_{1,t-1} + \hat{a}_{2,t-1})$$

$$\hat{f}_t = \alpha_2 \frac{y_t}{\hat{a}_{1,t}} + (1 - \alpha_2) \hat{f}_{t-l}$$

$$\hat{a}_{2,t} = \alpha_3 (\hat{a}_{1,t} - \hat{a}_{1,t-1}) + (1 - \alpha_3) \hat{a}_{2,t-1},$$

где $0 < \alpha_1, \alpha_2, \alpha_3 < 1$

$\hat{a}_{1,t}$ и $\hat{a}_{2,t}$ - характеристики тенденции развития;

$\hat{f}_1, \hat{f}_{t-1}, \dots, \hat{f}_{t-l+1}$ - характеристики мультипликативного сезонного фактора;

l - число фаз в полном сезонном цикле (для ежемесячных наблюдений $l = 12$, для квартальных - $l = 4$);

$\alpha_1, \alpha_2, \alpha_3$ - параметры адаптации. Каждый параметр принадлежит интервалу $[0; 1]$, чем ближе параметр к единице, тем больший вес приписывается последним наблюдениям.

Очевидно, что в этой модели прогнозные оценки зависят от прошлых и текущих уровней временного ряда, начальных значений коэффициентов $\hat{a}_{1,0}$ и $\hat{a}_{2,0}$, значений сезонного фактора $(\hat{f}_1, \hat{f}_{t-1}, \dots, \hat{f}_{t-l+1})$ для каждой фазы цикла и параметров адаптации $\alpha_1, \alpha_2, \alpha_3$.

Пример

В качестве примера применения предложенной методики можно рассмотреть модель для прогнозирования общего потребления топлива в Великобритании (в пересчете на млн.т нефти) по данным ежеквартального потребления топлива до 2010 года [3].

	A	B	C	D
1	TABLE 1.2 Inland energy consumption:			
2	primary fuel input basis			
3	QUARTERLY DATA			
4				
5			t	$Y(t)$
6			1	2
7			0	
8	2007	Quarter 1	1	64.7
9	2007	Quarter 2	2	51.2
10	2007	Quarter 3	3	48.9
11	2007	Quarter 4	4	62.7
12	2008	Quarter 1	5	65.7
13	2008	Quarter 2	6	52.0
14	2008	Quarter 3	7	46.9
15	2008	Quarter 4	8	60.9
16	2009	Quarter 1	9	63.2
17	2009	Quarter 2	10	46.9
18	2009	Quarter 3	11	44.1
19	2009	Quarter 4	12	57.1
20	2010	Quarter 1	13	65.2
21	2010	Quarter 2	14	47.5
22	2010	Quarter 3	15	44.2
23	2010	Quarter 4	16	61.7

Рис.12. Исходные данные.

	A	B	C	D	E	F	G	H	I
40	Начальные значения параметров				Вычисляем по 8 наблюдениям.				
41	Линейн()				Сезонная составляющая				
42	по 8 на	b(0)	a(0)			1-год	2-год	Среднее знач.	
43		-0.4333	58.5733		1	1.11339382	1.166	1.139	F(-3)
44		1.2626	6.3757		2	0.88771576	0.929	0.908	F(-2)
45		0.0193	8.1824		3	0.8529693	0.844	0.849	F(-1)
46		0.1178	6		4	1.10306744	1.104	1.104	F(0)
47		7.8867	401.7095						
48									
49		alpha1=	alpha=	0.27295	1-alpha=	0.7270521			
50		alpha2=	beta=	0.1	1-beta=	0.9			
51		alpha3=	gamma=	0.28253	1-gamma=	0.7174712			
52									

Рис.13. Вычисление начальных приближений для параметров. Режим отображения данных

	B	C	D	E
28	t	№№ кварталов	Y(t)	Y(t) теорет
29	1	2	3	4
30	0	1		
31	1	1	=D8	=\$B\$43*B31+\$C\$43
32	2	2	=D9	=\$B\$43*B32+\$C\$43
33	3	3	=D10	=\$B\$43*B33+\$C\$43
34	4	4	=D11	=\$B\$43*B34+\$C\$43
35	5	1	=D12	=\$B\$43*B35+\$C\$43
36	6	2	=D13	=\$B\$43*B36+\$C\$43
37	7	3	=D14	=\$B\$43*B37+\$C\$43
38	8	4	=D15	=\$B\$43*B38+\$C\$43
39				
40				Вычисляем по 8 наб
41				Сезонная составляющ
42	b(0)	a(0)		
43	=ЛИНЕЙН(D31:D38;B31:B38;1;1)	=ЛИНЕЙН(D31:D38;B31:B38;1;1)		1
44	=ЛИНЕЙН(D31:D38;B31:B38;1;1)	=ЛИНЕЙН(D31:D38;B31:B38;1;1)		2
45	=ЛИНЕЙН(D31:D38;B31:B38;1;1)	=ЛИНЕЙН(D31:D38;B31:B38;1;1)		3
46	=ЛИНЕЙН(D31:D38;B31:B38;1;1)	=ЛИНЕЙН(D31:D38;B31:B38;1;1)		4
47	=ЛИНЕЙН(D31:D38;B31:B38;1;1)	=ЛИНЕЙН(D31:D38;B31:B38;1;1)		

Рис.14. Вычисление начальных приближений для параметров. Режим отображения данных (начало)

	E	F	G	H	I
41	Сезонная сс				
42		1-год	2-год	Среднее знач.	
43	1	=D31/E31	=D35/E35	=СРЗНАЧ(F43:G43)	F(-3)
44	2	=D32/E32	=D36/E36	=СРЗНАЧ(F44:G44)	F(-2)
45	3	=D33/E33	=D37/E37	=СРЗНАЧ(F45:G45)	F(-1)
46	4	=D34/E34	=D38/E38	=СРЗНАЧ(F46:G46)	F(0)

Рис.15. Вычисление начальных приближений для параметров. Режим отображения данных. (окончание.)

	B	C	D	E	F	G	H	I	J	K
53									Абс.погр.,	Отн.погр.,
54	t	№ квартал	Y(t)	tau	a1	a2	F	Y(t) теорет	E(t)	%
55	1	2	3		4	5	6	7		
56	-3	(-3)1					1.139			
57	-2	(-2)2					0.908			
58	-1	(-1)3					0.849			
59	0	(0)4		1	58.5733	-0.4333	1.104			
60	1	1	64.73263	1	57.776798	-0.46965	1.134078	66.2485	-1.5159123	0.023418056
61	2	2	51.22705	1	57.058685	-0.49449	0.905355	52.0539	-0.8268625	0.016141131
74	15	3	44.21527	1	52.398418	-0.38984	0.843749	44.2073	0.0079876	0.000180652
75	16	4	61.72703	1	53.122147	-0.27848	1.117893	57.2371	4.4899429	0.072738681
76										0.423007116

Рис.16. Построение значений модельных значений. Режим отображения данных

	B	C	D	E	F	G
53						
54	t	квар	Y(t)	tau	a1	a2
55	1	2	3	4		5
56	-3	(-3)1				
57	-2	(-2)2				
58	-1	(-1)3				
59	0	(0)4		1	=C43	=B43
60	1	1	=D8	1	=\$D\$49*D60/H56+\$F\$49*(F59+G59)	=\$D\$50*(F60-F59)+\$F\$50*G59
61	2	2	=D9	1	=\$D\$49*D61/H57+\$F\$49*(F60+G60)	=\$D\$50*(F61-F60)+\$F\$50*G60
74	15	3	=D22	1	=\$D\$49*D74/H70+\$F\$49*(F73+G73)	=\$D\$50*(F74-F73)+\$F\$50*G73
75	16	4	=D23	1	=\$D\$49*D75/H71+\$F\$49*(F74+G74)	=\$D\$50*(F75-F74)+\$F\$50*G74

Рис.17. Построение значений модельных значений. Режим отображения формул.

	H	I	J	K
53			Абс.погр.,	Отн.погр.,
54	F	Y(t) теорет	E(t)	%
55	6	7		
56	=H43			
57	=H44			
58	=H45			
59	=H46			
60	=\$D\$51*D60/F60+\$F\$51*H56	=(F59+G59*E60)*H56	=D60-I60	=ABS(J60)/D60
61	=\$D\$51*D61/F61+\$F\$51*H57	=(F60+G60*E61)*H57	=D61-I61	=ABS(J61)/D61
74	=\$D\$51*D74/F74+\$F\$51*H70	=(F73+G73*E74)*H70	=D74-I74	=ABS(J74)/D74
75	=\$D\$51*D75/F75+\$F\$51*H71	=(F74+G74*E75)*H71	=D75-I75	=ABS(J75)/D75
76				=СУММ(K60:K75)

Рис.18. Построение значений модельных значений. Режим отображения формул.

	A	B	C	D	E	F	G	H	I	J	K
77			Прогноз		tau						
78		13				53.155178	-0.39242	1.167417		факт	
79		14				52.785934	-0.3901	0.899217			
80		15				52.398418	-0.38984	0.843749			
81		16				53.122147	-0.27848	1.117893			
82	2011	Quarter 1			1				61.6906	60.7	1.63%
83	2011	Quarter 2			2				47.2675		
84	2011	Quarter 3			3				44.1168		
85	2011	Quarter 4			4				58.1396		

Рис.19. Построение значений прогнозных значений на 2011 год. Режим отображения формул.

	A	B	C	D	E	F	G
77			Прог		tau		
78		13				=F72	=G72
79		14				=F73	=G73
80		15				=F74	=G74
81		16				=F75	=G75
82	2011	Quar			1		
83	2011	Quar			2		
84	2011	Quar			3		
85	2011	Quar			4		

Рис.20. Построение значений прогнозных значений на 2011 год. Режим отображения формул.

	H	I	J	K
77				
78	=H72		факт	
79	=H73			
80	=H74			
81	=H75			
82		=(F\$81+\$G\$81*E82)*H	60.7	=ABS(I82-J82)/J82
83		=(F\$81+\$G\$81*E83)*H		
84		=(F\$81+\$G\$81*E84)*H		
85		=(F\$81+\$G\$81*E85)*H		

Рис.21. Построение значений прогнозных значений на 2011 год. Режим отображения формул.

При далеко не оптимальных значениях параметров адаптации $\alpha_1 = \alpha_2 = \alpha_3 = 0.5$ относительная ошибка прогноза на 1 квартал 2011 года составила 6.95%.

При оптимальных значениях параметров адаптации $\alpha_1 = 0.27$, $\alpha_2 = 0.09$, $\alpha_3 = 0.28$ относительная ошибка прогноза на 1 квартал 2011 года составила 1.63%.

Динамика фактических и прогнозных уровней потребления топлива при оптимальных значениях параметров приведена на рисунке 22.

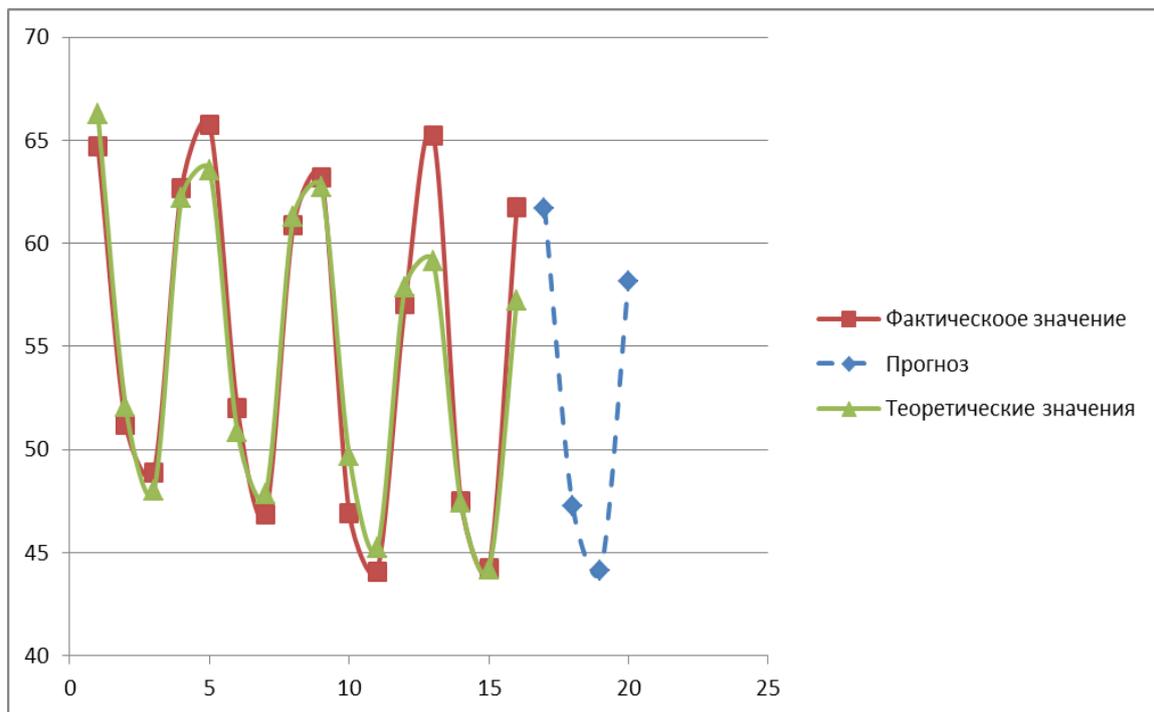


Рис. 22. График фактических и прогнозных уровней потребления топлива в Великобритании.

Выбор оптимальных значений параметров адаптации позволил уменьшить относительную ошибку прогноза более чем в 4 раза.

МОДЕЛЬ ТЕЙЛА-ВЕЙДЖА

Эта трехпараметрическая трендсезонная модель, сочетающая линейный рост с аддитивной сезонностью.

$$y_{\tau}(t) = a_{1,t} + g_t + u_t \quad (17)$$

где $a_{1,t} = a_{1,t-1} + a_{2,t}$

$a_{1,t}$ - величина уровня ряда после удаления сезонных колебаний;

$a_{2,t}$ - аддитивный коэффициент роста, g_t - аддитивный коэффициент сезонности, u_t - белый шум.

Прогноз по этой модели на τ шагов вперед определяется выражением:

$$\hat{y}_{\tau}(t) = \hat{a}_{1,t} + \hat{a}_{2,t} \cdot \tau + \hat{g}_{t-l+\tau} \quad (18)$$

Обновление коэффициентов осуществляется следующим образом:

$$\hat{a}_{1,t} = \alpha_1 \cdot (y_t - \hat{g}_{t-l}) + (1 - \alpha_1) \cdot (\hat{a}_{1,t-1} + \hat{a}_{2,t-1})$$

$$\hat{g}_t = \alpha_2 \cdot (y_t - \hat{a}_{1,t}) + (1 - \alpha_2) \cdot \hat{g}_{t-l}$$

$$\hat{a}_{2,t} = \alpha_3 \cdot (\hat{a}_{1,t} - \hat{a}_{1,t-1}) + (1 - \alpha_3) \cdot \hat{a}_{2,t-1}$$

$$0 < \alpha_1, \alpha_2, \alpha_3 < 1$$

где

$\alpha_{1,t}$ - характеристика тенденции развития;

$g_1, g_{t-1}, \dots, g_{t-l+1}$ - аддитивный сезонный фактор;

l - число фаз в полном сезонном цикле (для ежемесячных наблюдений $l = 12$, для квартальных - $l = 4$);

ПРИМЕНЕНИЕ СТАТИСТИЧЕСКОГО ПАКЕТА R ДЛЯ ПОСТРОЕНИЯ МОДЕЛИ ХОЛЬТА-УИНТЕРСА

Использование электронных таблиц MS Excel для построения адаптивных моделей достаточно трудоемко и их применение оправданно только для учебных целей для пояснения алгоритмов реализации этих методов.

Рассмотрим применение статистического пакета R¹ на примере исследования потребления топлива в Великобритании с помощью модели Хольта-Уинтерса. Синтаксис использованных команд приведен в Приложении 1.

¹ Статистический пакет R – язык программирования для статистических расчетов и работы с графикой. Его особенностью - это свободная программная среда с открытым исходным кодом [1].

Будем предполагать что исходные данные находятся в файле "DATA_KURS.txt" в текущем каталоге. Решение приведено ниже. Консольная команда в R начинается с приглашения «>».

```
# чтение данных read.table в набор d
>d<-read.table("DATA_KURS.txt",h=T)
#формирование временного ряда y_ts из d
>y_ts=ts(d,frequency=4,start=c(1995,1))
#вывод временного ряда

> y_ts
      Qtr1 Qtr2 Qtr3 Qtr4
1995 63.6 49.6 44.6 61.8
1996 67.1 52.6 45.5 63.9
1997 64.6 51.4 46.2 61.9
1998 63.7 54.6 48.8 63.6
1999 67.0 52.4 47.5 64.1
2000 67.7 53.2 48.7 64.4
2001 70.5 53.2 50.4 63.1
2002 66.4 52.0 48.7 62.5
2003 67.1 51.8 47.6 64.1
2004 67.8 52.3 48.9 64.6
2005 68.2 54.5 48.6 64.1
2006 70.8 52.9 47.9 61.5
2007 64.7 51.2 48.9 62.7
2008 65.7 52.0 46.9 60.9
2009 63.2 46.9 44.1 57.1
2010 65.2 47.5 44.2 61.7

# Выделим некоторое подмножество из временного ряда
# Сделано для сопоставления расчетов в MS Excel
> y_tsw=window(y_ts,start=c(2007,1),end=c(2010,4))

> y_tsw
      Qtr1 Qtr2 Qtr3 Qtr4
2007 64.7 51.2 48.9 62.7
2008 65.7 52.0 46.9 60.9
2009 63.2 46.9 44.1 57.1
2010 65.2 47.5 44.2 61.7

# Построение модели HoltWinters. Результат находится
# в объекте mdHoltWinters
#
```

```

> mdHoltWinters<- HoltWinters(y_tsw,seasonal =
"multiplicative")

# Результат находится в объекте mdHoltWinters
# Вывод списка свойств объекта mdHoltWinters
#
> summary(mdHoltWinters)
              Length Class  Mode
fitted         48      mts  numeric
x              16       ts  numeric
alpha           1    -none- numeric
beta            1    -none- numeric
gamma           1    -none- numeric
coefficients    6    -none- numeric
seasonal        1    -none- character
SSE             1    -none- numeric
call           3    -none- call
>
# Вывод значений свойств объекта mdHoltWinters
# SSE сумма квадратов отклонений
> mdHoltWinters$SSE
[1] 57.42872

# Вывод значений свойств объекта mdHoltWinters
# alpha, beta и gamma параметры модели
> mdHoltWinters$alpha
      alpha
0.3741533
>
> mdHoltWinters$beta
      beta
      0
>
> mdHoltWinters$gamma
      gamma
      1
# Вывод значений свойств объекта mdHoltWinters
# fitted - теоретические значения уровней временного
# ряда
#
> mdHoltWinters$fitted
      xhat      level  trend      season
2008 Q1 65.43496 57.31250 -0.135 1.1444179

```

```

2008 Q2 52.18070 57.26415 -0.135 0.9133813
2008 Q3 48.54730 57.05513 -0.135 0.8529022
2008 Q4 61.06879 56.19749 -0.135 1.0892986
2009 Q1 64.09991 56.00451 -0.135 1.1473146
2009 Q2 50.52892 55.57604 -0.135 0.9113992
2009 Q3 44.91274 53.95127 -0.135 0.8345569
2009 Q4 57.97746 53.45190 -0.135 1.0874124
2010 Q1 60.13410 53.01499 -0.135 1.1371807
2010 Q2 47.30030 54.54676 -0.135 0.8693029
2010 Q3 44.85146 54.49771 -0.135 0.8250408
2010 Q4 58.08797 54.06728 -0.135 1.0770539

```

```

# Построение прогнозных значений
# (точечный и интервальный прогнозы)
# на 4 периода вперед
# с помощью модели mdHoltWinters
# Результат находится в объекте dw

```

```

> dw <- predict(mdHoltWinters, 4, prediction.interval
= TRUE)

```

```

# Вывод результат прогнозирования из объекта dw
# значения точечного прогноза: fit ;
# интервальный прогноз:
# lwr - нижняя граница
# upr - верхняя граница

```

```

> dw

```

	fit	upr	lwr
2011 Q1	65.80397	70.28206	61.32588
2011 Q2	47.86549	52.50725	43.22373
2011 Q3	44.78432	49.66653	39.90211
2011 Q4	61.09627	98.34779	23.84475

```

# Вывод графиков фактических, теоретических и
прогнозных
# значений
> plot(mdHoltWinters,dw,main="Потребление топлива в
Великобритании",ylab="Факт / Прогноз",xlab="Годы")

```

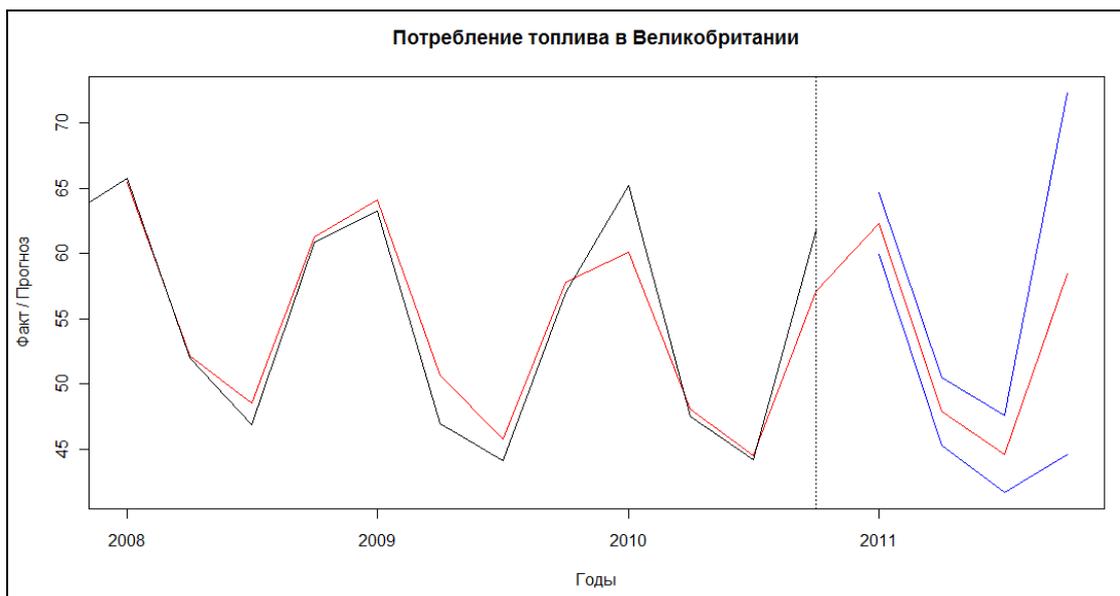


Рис. 24. График фактических и прогнозных уровней потребления топлива в Великобритании.

Полученные результаты немного отличаются от решения, полученного в MS Excel. Причиной этого является различный способ выбора стартовых значений коэффициентов, что и приводит к отличиям в оптимальных значениях параметров адаптации. Прогнозные значения при этом различаются несущественно.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Лукашин Ю.П. Адаптивные методы краткосрочного прогнозирования временных рядов. — М.: Финансы и статистика, 2003
2. Материалы сайта "Department of energy and climate change. Energy sector statistics»: http://www.decc.gov.uk/en/content/cms/statistics/energy_stats/source/total/total.aspx
3. Балдин Е.М. Введение в R [Электронный ресурс]. – Электрон. дан. – 2008 – Режим доступа: <http://www.inp.nsk.su/~baldin/DataAnalysis/R/R-01-intro.pdf>
4. Материалы сайта R FUNCTIONS FOR TIME SERIES ANALYSIS <https://cran.r-project.org/doc/contrib/Ricci-refcard-ts.pdf>

ПРИЛОЖЕНИЕ 1

Список функций R , используемых для анализа временных рядов, приведен в [4].

Рассмотрим функции R , которые используются в примере из раздела «Применение статистического пакета R для построения модели Хольта-Уинтерса» учебного пособия.

1. Функция `HoltWinters`

Описание

Вычисляет параметры модели Хольта-Уинтерса для временного ряда. Неизвестные параметры адаптации определяются как оптимальные для минимизации суммы квадратов отклонений фактических значений временного ряда от теоретических значений.

Синтаксис

```
HoltWinters(x, alpha = NULL, beta = NULL, gamma =
NULL,
          seasonal = c("additive",
"multiplicative"),
          start.periods = 2, l.start = NULL,
b.start = NULL,
          s.start = NULL,
          optim.start = c(alpha = 0.3, beta = 0.1,
gamma = 0.1),
          optim.control = list())
```

Аргументы

<code>x</code>	объект типа <code>ts</code> (временной ряд)
<code>alpha</code>	параметр модели
<code>beta</code>	параметр модели. Если равен <code>FALSE</code> , функция будет выполнять экспоненциальное сглаживание.
<code>gamma</code>	параметр используется для оценки сезонной компоненты. Если равен <code>FALSE</code> , то строится модель не учитывающая сезонность.
<code>seasonal</code>	строковый параметр. Определяет тип вхождения сезонной составляющей. Принимает значения <code>"additive"</code> (по умолчанию) or <code>"multiplicative"</code> .

Результаты работы

Результатом работы является объект класса "**HoltWinters**", со следующими компонентами:

<code>fitted</code>	теоретические уровни временного ряда
<code>x</code>	исходный временной ряд
<code>alpha</code>	параметр модели α_1
<code>beta</code>	параметр модели α_2
<code>gamma</code>	параметр модели α_3
<code>coefficients</code>	вектор, содержащий коэффициенты модели $a, b, s1, \dots, sr$, характеризующие тренд и сезонную составляющую.
<code>SSE</code>	сумма квадратов отклонений, полученная в результате оптимизации.

2. Функция `predict.HoltWinters`

Описание

Функция для построения точечного и интервального прогноза на основании модели Хольта-Уинтерса для временного ряда.

Синтаксис

```
predict(object, n.ahead = 1, prediction.interval =  
FALSE,  
        level = 0.95, ...)
```

Аргументы

<code>object</code>	объект класса " HoltWinters ",
<code>n.ahead</code>	горизонт планирования(количество периодов прогноза).
<code>prediction.interval</code>	интервальный прогноз, тип <code>logical</code> . Если <code>TRUE</code> , вычисляются верхняя и нижняя границы интервального прогноза.
<code>level</code>	доверительная вероятность для интервального прогноза.

Результат работы

Результатом работы является временной ряд с предсказанными значениями .

Если построен интервальный прогноз, то временной ряд выводится в виде таблицы с колонками *fit*, *lwr* и *upr* , содержащий прогнозные значения, нижнюю и верхнюю границы интервального прогноза, соответственно.

3. Функция `read.table`

Описание

Команда *read.table*, производит чтение данных из файлов различного типа

Синтаксис

```
read.table(file, head = FALSE, sep=";" ...)
```

Аргументы

`file` - название файла, из которого должны быть прочитаны данные.

`head` - (`head=TRUE`) указывает, содержит ли файл названия переменных в его первой строке .

`sep` - вид разделителя.

4. Функция `summary`

Описание

Вывод основных свойств объекта

Синтаксис

```
summary(object)
```

`object` - объект, свойства которого желательно увидеть

5. Функция `ts`

Описание

Команда *ts*, преобразует объекты (векторы, матрицы) в объекты типа *ts* (time series).

Синтаксис

```
ts(data = NA, start = 1, end=numeric(), frequency= 1)
```

Аргументы

<code>data</code>	вектор или матрица наблюдений.
<code>start</code>	время первого наблюдения
<code>end</code>	время последнего наблюдения.
<code>Frequency</code>	количество периодов времени на один цикл.

6. Функция `Window`

Описание

Функция для извлечения подмножества из временного ряда.

Синтаксис

```
window(x, start = NULL, end = NULL, frequency = NULL)
```

Аргументы

<code>x</code>	объект типа <code>ts</code> (time series)
<code>start</code>	время первого наблюдения .
<code>end</code>	время последнего наблюдения.
<code>Frequency</code>	количество периодов времени на один цикл

7. Функция `plot.ts`

Описание

Построение графиков временных рядов.

Синтаксис

```
plot.ts(x)  
x объект типа ts (time series).
```

8. Функция `plot`

Описание

Для изображения графиков используется функция **`plot(x, y)`**.

Здесь x — вектор значений абсцисс и y — вектор значений ординат.

Если указан только один аргумент — `plot(y)`, то предполагается, что $x=1:n$, где n — длина вектора y .

Кроме этих двух аргументов, функция принимает ещё много аргументов, отвечающих за оформление и некоторые другие аспекты графического вывода.

Графики появляются в отдельном окне (или в нескольких отдельных окнах), называемых графическими устройствами.

Функция **`plot()`** допускает множество дополнительных параметров. Вот некоторые из них:

- $\log = "x"$ или $\log = "y"$ или $\log = "xy"$ делает соответственно масштаб по оси абсцисс, ординат или по обеим осям логарифмическим;
- $type = \text{тип-графика}$ позволяет указать тип выводимого графика; некоторые значения параметра — "p" (точки, по умолчанию), "l" (линии), "b" (линии и точки), "o" (линии и точки перекрываются), "n" (ничего не рисуется);
- $xlab = \text{название-оси-абсцисс}$ и $ylab = \text{название-оси-ординат}$ позволяет указать подписи к оси абсцисс и ординат соответственно;
- $main = \text{основная-надпись}$ и $sub = \text{дополнительная-надпись}$ создают надписи сверху и снизу графика;
- $col = \text{цвет}$ задает цвет графика; возможны значения "blue", "red", "green", "cyan", "magenta", "yellow", "black" и другие.