

Тема: Адаптивные методы прогнозирования

Теоретический материал к теме.

Метод экспоненциального сглаживания может быть использован как для сглаживания уровней временного ряда, так и для прогнозирования. В нашем случае мы будем использовать его для прогнозирования.

Особенность экспоненциального сглаживания 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:

$$S_t = \alpha \cdot y_{t-1} + (1 - \alpha) \cdot S_{t-1} \quad (4.1)$$

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

y_{t-1} – фактический уровень ряда в момент времени $t - 1$;

S_{t-1} – сглаженное по экспоненте значение ряда в момент времени $t - 1$.

Если обозначить $\beta = 1 - \alpha$, то формула 4.1 преобразуется в следующий вид:

$$S_t = \alpha \cdot y_{t-1} + \beta \cdot S_{t-1} \quad (4.2)$$

В результате математических преобразований можно получить и другую формулу:

$$S_t = S_{t-1} + \alpha \cdot (y_{t-1} - S_{t-1}) \quad (4.3)$$

Величину $(y_{t-1} - S_{t-1})$ можно рассматривать как погрешность прогноза. Тогда новый прогноз S_t получается в результате корректировки предыдущего прогноза с учетом его ошибки. В этом и состоит адаптация модели.

Формулы 4.1 - 4.3 равнозначны и дают одинаковый результат. Можно пользоваться той, которая представляется более удобной. Формула 4.3 наглядно демонстрирует процесс обучения модели на своих же ошибках, допущенных на предыдущих шагах прогнозирования.

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).

Например, если $\alpha = 0,25$, то $\beta = 0,75$. Математические преобразования формулы 4.1 (которые мы не будем здесь рассматривать) показывают, что вес каждого уровня ряда y_t по мере удаления от момента прогнозирования t снижается и равен: $w_{y_{t-i}} = \alpha \cdot \beta^i$

Для нашей задачи 4.1 (при $\alpha = 0,25$, то $\beta = 0,75$) получим:

вес слагаемого y_{t-1} (безработица в октябре): $w_{y_{t-1}} = 0,25 \cdot 0,75^1 = 0,188$;

вес слагаемого y_{t-2} (безработица в сентябре): $w_{y_{t-2}} = 0,25 \cdot 0,75^2 = 0,141$;

вес слагаемого y_{t-3} (безработица в августе): $w_{y_{t-3}} = 0,25 \cdot 0,75^3 = 0,105$;

...

вес слагаемого y_{t-10} (безработица в январе): $w_{y_{t-10}} = 0,25 \cdot 0,75^{10} = 0,008$;

	<i>месяц</i>	<i>y</i>	<i>α</i>	<i>β</i>	<i>w</i>
<i>t-10</i>	янв	2,99	0,3	0,7	0,008
<i>t-9</i>	фев	2,66	0,25	0,75	0,019
<i>t-8</i>	мар	2,1	0,25	0,75	0,025
<i>t-7</i>	апр	2,2	0,25	0,75	0,033
<i>t-6</i>	май	2,4	0,25	0,75	0,044
<i>t-5</i>	июн	2,22	0,25	0,75	0,059
<i>t-4</i>	июл	1,97	0,25	0,75	0,079
<i>t-3</i>	авг	1,42	0,25	0,75	0,105
<i>t-2</i>	сен	2,22	0,25	0,75	0,141
<i>t-1</i>	окт	1,72	0,25	0,75	0,188
<i>t</i>	ноя				

На рисунке 4.1 наглядно демонстрируется снижение веса наблюдения по мере удаления от момента прогнозирования в прошлое:

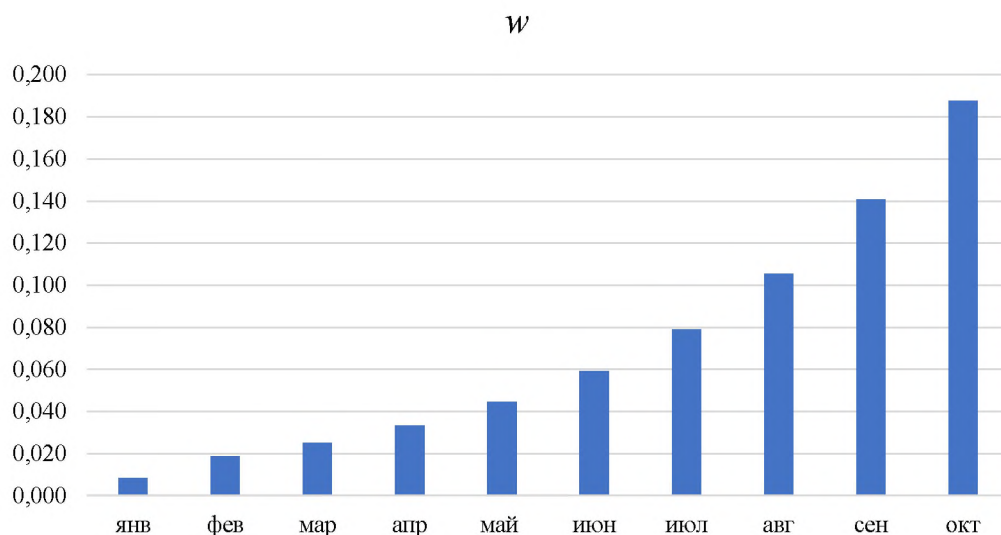


Рисунок 4.1

То есть при $\alpha = 0,25$, то $\beta = 0,75$ значение фактического уровня безработицы в октябре при прогнозировании будет иметь вес 18,8%; уровень безработицы в сентябре – вес 14,1% и т.д.

Изменение параметра сглаживания приведет к изменению весов уровней ряда при прогнозировании. Чем ближе α к единице, тем более высокий вес имеют последние уровни ряда. Чем ближе α к нулю, тем больший вес имеет сложившаяся тенденция за весь период наблюдений.

Задача 4.1. В таблице 4.1 представлены данные, характеризующие уровень безработицы в регионе. Определите прогнозное значение уровня безработицы на ноябрь с помощью метода экспоненциального сглаживания.

Таблица 4.1 – Уровень безработицы, %

Январь	Февраль	Март	Апрель	Май	Июнь	Июль	Август	Сентябрь	Октябрь
2,99	2,66	2,1	2,2	2,4	2,22	1,97	1,42	2,22	1,72

Решение. Для решения будем пользоваться одной из формул (4.1 – 4.3). Например, (4.3): $S_t = S_{t-1} + \alpha \cdot (y_{t-1} - S_{t-1})$.

Вначале необходимо задать параметр сглаживания α . Пусть $\alpha = 0,25$.

Построим вспомогательную таблицу:

i	месяц	y_t	S_t
0	янв	2,99	
1	фев	2,66	
2	мар	2,1	
3	апр	2,2	
4	май	2,4	
5	июн	2,22	
6	июл	1,97	
7	авг	1,42	
8	сен	2,22	
9	окт	1,72	
10	ноя		

Очевидно, что в экспоненту для января S_0 найти невозможно. Начальное значение экспоненты S_0 принимают равным фактическому уровню y_0 или среднему арифметическому из всех фактических уровней ряда: $\sum y / n$. В некоторых случаях используют среднее арифметическое из нескольких первоначальных уровней ряда. На практике выбор S_0 при достаточном количестве наблюдений (более 4-5) не влияет на конечный результат: прогноз будет одинаковым в любом случае.

Попробуем продемонстрировать это в нашем примере.

Пусть $S_0 = y_0 = 2,99$. Тогда:

$$S_1 = S_0 + \alpha \cdot (y_0 - S_0) = 2,99 + 0,25 \cdot (2,99 - 2,99) = 2,99;$$

Внесем результаты в таблицу:

<i>i</i>	<i>месяц</i>	<i>y_t</i>	<i>S_t</i>
0	янв	2,99	2,99
1	фев	2,66	2,99
2	мар	2,1	
3	апр	2,2	
4	май	2,4	
5	июн	2,22	
6	июл	1,97	
7	авг	1,42	
8	сен	2,22	
9	окт	1,72	
10	ноя		

$$S_2 = S_1 + \alpha \cdot (y_1 - S_1) = 2,99 + 0,25 \cdot (2,66 - 2,99) = 2,91;$$

$$S_3 = S_2 + \alpha \cdot (y_2 - S_2) = 2,91 + 0,25 \cdot (2,1 - 2,91) = 2,71;$$

<i>i</i>	<i>месяц</i>	<i>y_t</i>	<i>S_t</i>
0	янв	2,99	2,99
1	фев	2,66	2,99
2	мар	2,1	2,91
3	апр	2,2	2,71
4	май	2,4	
5	июн	2,22	
6	июл	1,97	
7	авг	1,42	
8	сен	2,22	
9	окт	1,72	
10	ноя		

Аналогично проводим расчеты до S_{10} и заносим результаты в таблицу.

$$S_4 = S_3 + \alpha \cdot (y_3 - S_3) = 2,71 + 0,25 \cdot (2,2 - 2,71) = 2,58$$

...

$$S_9 = S_8 + \alpha \cdot (y_8 - S_8) = 2,11 + 0,25 \cdot (2,22 - 2,11) = 2,13$$

В результате получилась таблица:

<i>i</i>	<i>месяц</i>	<i>y_t</i>	<i>S_t</i>
0	янв	2,99	2,99
1	фев	2,66	2,99
2	мар	2,1	2,91
3	апр	2,2	2,71
4	май	2,4	2,58
5	июн	2,22	2,53
6	июл	1,97	2,46
7	авг	1,42	2,33
8	сен	2,22	2,11
9	окт	1,72	2,13
10	ноя		

Теперь можно построить прогноз безработицы на ноябрь:

$$S_{10} = S_9 + \alpha \cdot (y_9 - S_9) = 2,13 + 0,25 \cdot (1,72 - 2,13) = 2,03.$$

<i>i</i>	<i>месяц</i>	<i>y_t</i>	<i>S_t</i>
0	янв	2,99	2,99
1	фев	2,66	2,99
2	мар	2,1	2,91
3	апр	2,2	2,71
4	май	2,4	2,58
5	июн	2,22	2,53
6	июл	1,97	2,46
7	авг	1,42	2,33
8	сен	2,22	2,11
9	окт	1,72	2,13
10	ноя		2,03

На рисунке 4.2 наглядно представлены результаты прогнозирования.

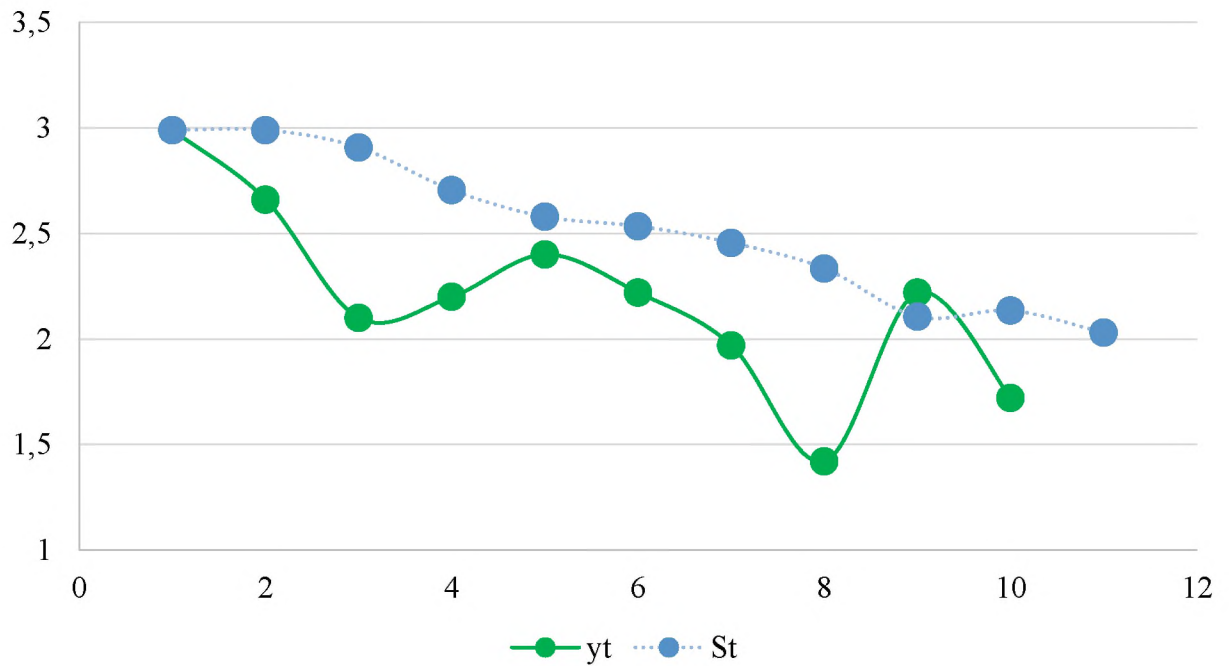


Рисунок 4.2

На данном этапе задача решена.

Дополнение.

Сравним результаты прогнозирования при следующих условиях:

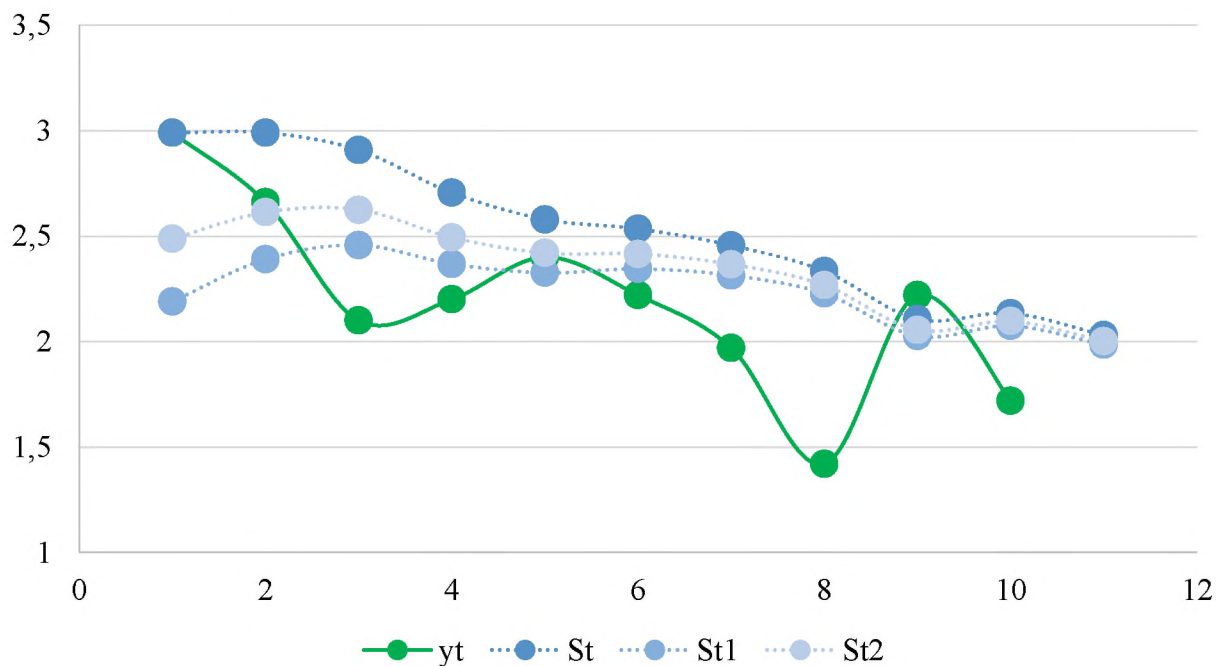
1) $S_0 = \frac{\sum_{i=0}^9 y}{10} = 2,19$ - среднее значение из всех десяти месяцев (от 0 до 9).

2) $S_0 = \frac{\sum_{i=0}^3 y}{4} = 2,49$ - среднее значение за первые 4 месяца (от 0 до 3).

Сравнительные результаты представлены в таблице:

<i>i</i>	месяц	y_t	S_t	S_{t1}	S_{t2}
0	янв	2,99	2,99	2,19	2,49
1	фев	2,66	2,99	2,39	2,61
2	мар	2,1	2,91	2,46	2,62
3	апр	2,2	2,71	2,37	2,49
4	май	2,4	2,58	2,33	2,42
5	июн	2,22	2,53	2,34	2,42
6	июл	1,97	2,46	2,31	2,37
7	авг	1,42	2,33	2,23	2,27
8	сен	2,22	2,11	2,03	2,06
9	окт	1,72	2,13	2,07	2,10
10	ноя		2,03	1,99	2,00

И на рисунке 4.3:



Можно видеть, что с увеличением количества уровней ряда выбранное в качестве точки отсчета значение экспоненты S_0 не влияет на результат прогнозирования.

После решения задачи 4.1:

Поэкспериментируйте с изменением параметра альфа: проследите, как меняется прогноз при изменении параметра от 0 до 1. (В приложенном файле «Данные для задач_тема 4.xlsx» есть привязанные формулы, прогноз будет меняться автоматически.)

Задачи для самостоятельного решения.

Задача 4.2. Имеются значения курса доллара:

Год	Месяц	Курс доллара, руб.
2015	Май 2015	51,7
	Июнь 2015	52,8
	Июль 2015	55,8
	Август 2015	60,3
	Сентябрь 2015	65,3
	Октябрь 2015	65,7
	Ноябрь 2015	63,8
	Декабрь 2015	66,7
2016	Январь 2016	72,9
	Февраль 2016	76,3
	Март 2016	75,9
	Апрель 2016	67,6
	Май 2016	66,2
	Июнь 2016	66,0
	Июль 2016	64,2
	Август 2016	65,9
	Сентябрь 2016	62,3
	Октябрь 2016	63,4
	Ноябрь 2016	63,9
	Декабрь 2016	61,5
2017	Январь 2017	60,3
	Февраль 2017	58,2
	Март 2017	57,5
	Апрель 2017	56,98
	Май 2017	56,51
	Июнь 2017	59,08
	Июль 2017	59,54
	Август 2017	58,73
	Сентябрь 2017	58,01
	Октябрь 2017	57,87
	Ноябрь 2017	58,33
	Декабрь 2017	57,6
2018	Январь 2018	56,29
	Февраль 2018	55,67
	Март 2018	57,26
	Апрель 2018	61,99
	Май 2018	62,59
	Июнь 2018	62,75
	Июль 2018	62,78
	Август 2018	68,08
	Сентябрь 2018	65,59
	Октябрь 2018	65,31
	Ноябрь 2018	

Сделайте прогноз на ноябрь 2018 года с помощью экспоненциального сглаживания. Параметр сглаживания выберите самостоятельно исходя из интуитивного представления или на основании экспериментального подбора. Для решения пользуйтесь файлом «Данные для задач_тема 4.xlsx».

Примечание: одного верного ответа для задачи 4.1 нет. Каждый может получить свой результат в зависимости от выбранного параметра сглаживания альфа. Правильным будет такой прогноз, который совпал с реальным значением курса валюты. При этом учитывайте, что в октябре 2018 года этот курс не был известен.

После решения задачи 4.2:

1) Подумайте, на основании какого критерия (количественного или качественного) можно оценить точность выбранного параметра сглаживания.

2) Найдите фактическое значение курса доллара на конец ноября 2018 года и сравните результат прогнозирования с фактическим уровнем. Проанализируйте, насколько верным оказался Ваш прогноз. Подумайте, как можно улучшить результат.

3) попытайтесь продолжить ряд фактических значений до апреля 2020 года; проанализируйте произошедшие изменения; попытайтесь сделать прогноз на май 2020 года.

