Эконометрика-1

Филатов Александр Юрьевич
(Главный научный сотрудник, доцент ШЭМ ДВФУ)

alexander.filatov@gmail.com
http://vk.com/alexander.filatov, **http://vk.com/baikalreadings**

Лекции 6.1-6.2

Анализ временных рядов. Аналитические и алгоритмические тренды. Сезонность

Введение в анализ временных рядов

Матрица «объект-свойство»:

$$\begin{pmatrix} x_1^{(1)}(t) & \dots & x_1^{(p)}(t) \\ x_2^{(1)}(t) & \dots & x_2^{(p)}(t) \\ \dots & \dots & \dots \\ x_n^{(1)}(t) & \dots & x_n^{(p)}(t) \end{pmatrix}$$

$i = 1,\dots,n$ – объекты, наблюдения

$j = 1,\dots,p$ – свойства, переменные

$t = \{t_1,\dots,t_T\}$ – моменты времени

***n* = 1, *T* > 1 – временные ряды (time series).**

Если $p = 1$, то одномерный временной ряд, иначе – многомерный. Для многомерного можно учитывать взаимодействие величин по принципам прямой и обратной связи.

Будем рассматривать **дискретные одномерные временные ряды** с равноотстоящими наблюдениями $t_2 - t_1 = \ldots = t_T - t_{T-1} = \Delta t$: $\boldsymbol{y_1, \ldots, y_T}$.

Главная задача: кратко- и среднесрочный прогноз.

Отличия от пространственной выборки:

1. Элементы временного ряда не являются одинаково распределенными.
2. Элементы временного ряда не являются статистически независимыми.

Основные факторы, формирующие временной ряд

1. **Долговременные** – общая тенденция изменения признака, как правило, монотонная. Моделируется некоторой регрессионной функцией, чаще всего полиномом или экспонентой. T_t **– тренд**.
2. **Сезонные** – периодические колебания, происходящие в определенное время года. Моделируются с помощью дамми-переменных или отклонений от скользящего среднего. S_t **– сезонность**.
3. **Циклические** – долговременные циклы экономической, политической, демографической или иной природы. Часто имеют неопределенную, в т.ч. изменяющуюся длительность. Моделируются некоторыми периодическими функциями. ϕ_t **– цикл**.
4. **Случайные** – остатки, не поддающиеся объяснению. Разделяются на шоковые скачкообразные изменения и малые случайные отклонения, которые можно моделировать. ε_t **– случайные остатки**.

Основные задачи анализа временных рядов

1. Определить, какие факторы (долговременные, сезонные, циклические, случайные) присутствуют в модели.
2. Построить хорошие оценки для коэффициентов неслучайных функций T_t, S_t, ϕ_t.
3. Подобрать модель, адекватно описывающую поведение остатков ε_t и оценить ее параметры.

Аддитивная и мультипликативная модель:

$$y_t = \alpha_T T_t + \alpha_S S_t + \alpha_\varphi \varphi_t + \varepsilon_t$$

$$y_t = (T_t)^{\alpha_T} (S_t)^{\alpha_S} (\varphi_t)^{\alpha_\varphi} \varepsilon_t$$

$$\alpha_T, \alpha_S, \alpha_\varphi = \begin{cases} 1, \text{ если составляющая} \\ \quad \text{есть в модели} \\ 0, \text{ иначе} \end{cases}$$

Примеры:

Помесячные авиаперевозки – T, S, ϕ (?), ε.

Фондовые индексы – T, ϕ, ε.

Урожайность – T (?), ε.

Аккуратно с циклами: ## Кризис 1857, 1895, 1933, 1971, 2009 (38 лет!)

Неслучайная составляющая временного ряда

Необходимо выявить (желательно автоматически) факт наличия / отсутствия неслучайной (т.е. зависящей от t) составляющей.

H_0: $Ey_t = a = \text{const}$.

H_0: $Ey_t \neq \text{const}$ – возможны различные варианты конкретизации.

Проверка гипотезы о неизменности среднего:

1. Критерий серий, основанный на медиане – выявляет монотонные зависимости
2. Критерий восходящих и нисходящих серий – выявляет периодические зависимости.
3. Критерий Аббе.

Пример: курс доллара за 3 апреля – 16 мая 2018

57,29, 57,54, 57,76, 57,58, 57,83, 58,57, 62,37, 64,06, 62,07, 61,43, 62,28, 61,15, 61,55, 60,86, 61,32, 61,77, 61,66, 61,75, 62,60, 62,73, 62,00, 63,49, 63,20, 62,71, 63,01, 62,52, 61,74, 61,77, 61,92.

Критерий серий, построенный на медиане

1. Переходим к вариационному ряду (сортируем в порядке возрастания): $y_1 < y_2 < \ldots < y_T$.
2. Определяем выборочную медиану $y_{med} = \begin{cases} y_{(T+1)/2}, \text{ если } T - \text{нечетное}, \\ \left(y_{T/2} + y_{T/2+1}\right)/2, \text{ если } T - \text{четное}. \end{cases}$
3. На основе исходного ряда записываем серии из «+» и «–»: «+», если $y_t > y_{med}$, «–», если $y_t < y_{med}$, $y_t = y_{med}$ не учитываются.
4. Вычисляются 2 характеристики:
   $\gamma\ (T)$ – общее число серий
   (последовательностей подряд идущих плюсов и минусов),
   $\tau\ (T)$ – протяженность самой длинной серии.
   Если ряд неслучайный, $\gamma\ (T)$ – достаточно мало, $\tau\ (T)$ – велико.
5. Если $\gamma(T) \leq \left[0{,}5\left(T + 2 - 1{,}96\sqrt{T-1}\right)\right]$ или $\tau(T) \geq \left[1{,}43\ln(T+1)\right]$, ряд неслучайный с вероятностью ошибки $\alpha = 0{,}05$.

Пример:

$x_{med} = 61{,}77$, $\gamma\ (29) = 8 < 10{,}31$, $\tau(29) = 8 > 4{,}86$, ряд неслучайный.

Критерий восходящих и нисходящих серий

1. На основе исходного ряда записываем серии из «+» и «–»:
   «+», если $y_t - y_{t-1} > 0$, «–», если $y_t - y_{t-1} < 0$,
   $y_t = y_{t-1}$ – не учитываются.
2. Вычисляются 2 характеристики:
   $\gamma(T)$ – общее число серий
   $\tau(T)$ – протяженность самой длинной серии.
   Если ряд неслучайный, $\gamma(T)$ – мало, $\tau(T)$ – велико.
3. Если $\gamma(T) \le \left[\frac{1}{3}(2T-1) - 1{,}96\sqrt{\frac{16T-29}{90}}\right]$ или $\tau(T) \ge \tau_0(T) = \begin{cases} 5, & T \le 26, \\ 6, & T \in [27; 153], \\ 7, & T \in [154; 1170], \\ 8, & T \ge 1171. \end{cases}$
   ряд неслучайный с вероятностью ошибки $\alpha = 0{,}05$.

Пример:

$\gamma(29) = 17 > 14{,}69$, $\tau(29) = 4 < 6$, ряд случайный (нет периодических колебаний, хотя может быть монотонная зависимость, выявленная предыдущим критерием).

Критерий Аббе (квадратов последовательных разностей)

1. Подсчитываем эмпирическое значение критерия

$$\gamma(T)=\frac{q^2}{s^2}=\frac{\dfrac{1}{2(T-1)}\sum\limits_{i=2}^{T}(y_t-y_{t-1})^2}{\dfrac{1}{T-1}\sum\limits_{i=1}^{T}(y_t-\overline{y})^2}.$$

2. Находим критическую точку

$$\gamma_\alpha(T)=1+\frac{u_\alpha}{\sqrt{T+0{,}5\left(1+u_\alpha^2\right)}},\ u_\alpha=НОРМСТОБР(\alpha).$$

3. Если $\gamma(T)\le\gamma_\alpha(T)$, ряд неслучайный с вероятностью ошибки α = 0,05.

Замечание:

Рекомендуется использовать критерий Аббе для больших выборок (n>60)

Пример:

γ (29) = 0,54/3,69 = 0,147 < 0,704, ряд неслучайный.

Полиномиальные тренды

Общая формула полиномиального тренда:

$$\hat{y}_t = \theta_0 + \theta_1 t + \theta_2 t^2 + \ldots + \theta_p t^p$$

$p = 1$ – линейный тренд, постоянный прирост;

$p = 2$ – квадратичный тренд, постоянное ускорение;

$p = 3$ – кубичный тренд, постоянное изменение ускорения (???)

Не рекомендуется использовать тренды высших степеней!

Темпы роста инфляции стали сокращаться.

Пример «Динамика курса доллара с 3 апреля по 16 мая 2018»:

$$\underset{}{\hat{y}_t} = \underset{(0{,}55)}{58{,}87} + \underset{(0{,}032)}{0{,}160}\, t, \quad R^2 = 0{,}483.$$

$$\hat{y}_t = \underset{(0{,}67)}{56{,}67} + \underset{(0{,}103)}{0{,}585}\, t - \underset{(0{,}003)}{0{,}014}\, t^2, \quad R^2 = 0{,}695.$$

$$\hat{y}_t = \underset{(0{,}67)}{55{,}86} + \underset{(0{,}268)}{0{,}885}\, t - \underset{(0{,}021)}{0{,}039}\, t^2 + \underset{(0{,}0005)}{0{,}0005}\, t^3, \quad R^2 = 0{,}712.$$

$$\hat{y}_t = \underset{(0{,}31)}{57{,}70} + \underset{(0{,}023)}{0{,}019}\, t + \underset{(0{,}474)}{4{,}135}\, z, \quad R^2 = 0{,}868.$$

Экспоненциальный тренд – наиболее используемый в экономике

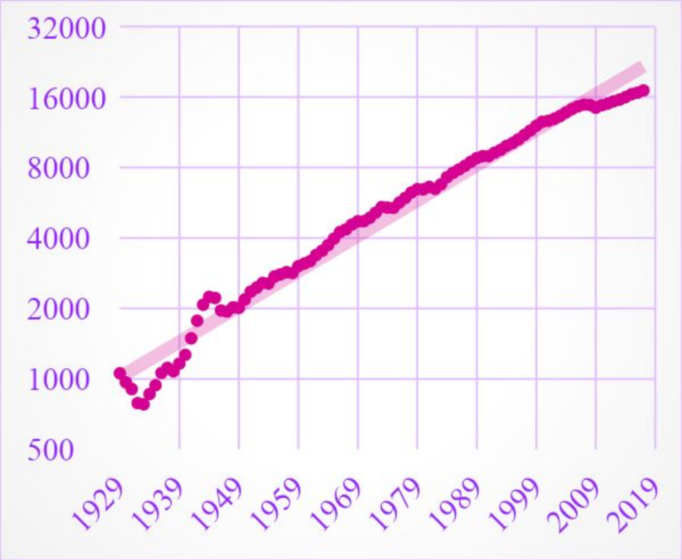
Единственный тренд, выявляющий **постоянный темп относительного прироста во времени**: экономический рост, уровень цен, выручка,…

5% / год = 132 раза /век,   10% / год = 13781 раз / век.

	y
1929	1 057
1930	967
1931	905
1932	788
1933	778
1934	862
1935	939
1936	1 061
1937	1 115
1938	1 078
…	…
2017	17 096

Каков темп экономического роста в США с 1929 г.? Происходит ли восстановление после кризиса с учетом роста экономики на 2,6%, 2,9%, 1,5% и 2,3% в 2014-2017?

$\hat{y}(t) = 1016^{0,035t}$, рост 3,5% в год.

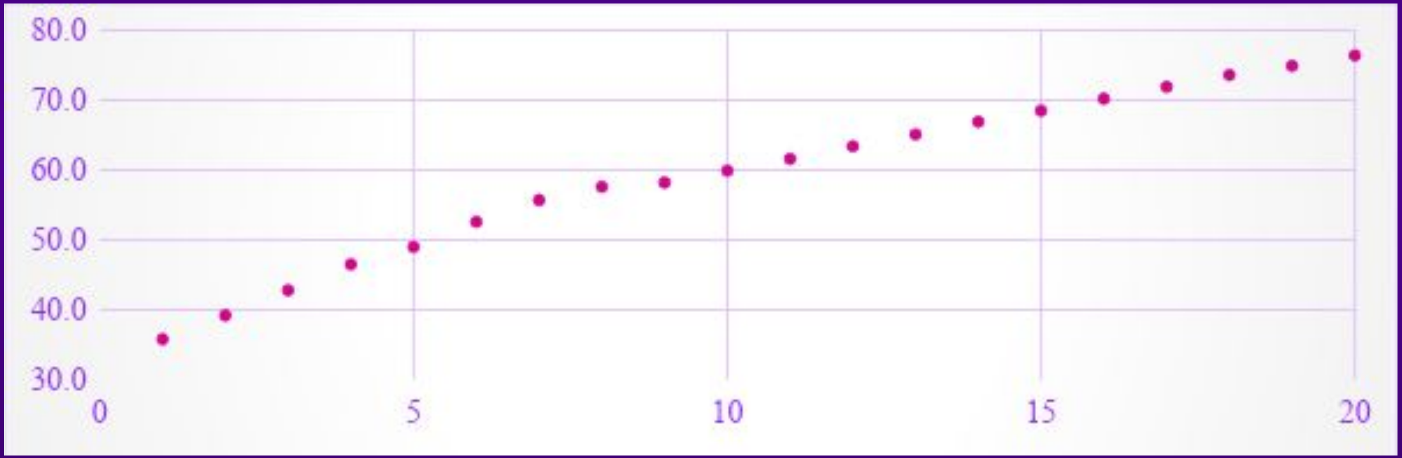


Аналитические тренды. Пример

	y	t
2013.1	35,8	1
2013.2	39,2	2
2013.3	42,8	3
2013.4	46,5	4
2014.1	49,0	5
2014.2	52,6	6
2014.3	55,7	7
2014.4	57,6	8
2015.1	58,2	9
2015.2	59,9	10
2015.3	61,6	11
2015.4	63,4	12
2016.1	65,1	13
2016.2	66,9	14
2016.3	68,5	15
2016.4	70,2	16
2017.1	71,9	17
2017.2	73,6	18
2017.3	74,9	19
2017.4	76,4	20

Задача: выявить долгосрочную тенденцию, построив аналитическую функцию от времени $T(t)$, позволяющую сделать долгосрочный прогноз.

Пример: имеется поквартальная динамика числа владельцев смартфонов в России за 2013-2017 гг.



Линейный тренд: постоянный абсолютный прирост.

$$\hat{y}(t) = 38{,}33 + 2{,}02t, \quad \hat{R}^2 = 0{,}973.$$

Квадратичный тренд: немонотонная зависимость.

$$\hat{y}(t) = 33{,}90 + 3{,}22t - 0{,}058t^2, \quad \hat{R}^2 = 0{,}994,$$

$$t_{\max} = 28\ (2019.4), \quad y_{\max} = 79{,}0.$$

Аналитические тренды. Пример

	y	t
2013.1	35,8	1
2013.2	39,2	2
2013.3	42,8	3
2013.4	46,5	4
2014.1	49,0	5
2014.2	52,6	6
2014.3	55,7	7
2014.4	57,6	8
2015.1	58,2	9
2015.2	59,9	10
2015.3	61,6	11
2015.4	63,4	12
2016.1	65,1	13
2016.2	66,9	14
2016.3	68,5	15
2016.4	70,2	16
2017.1	71,9	17
2017.2	73,6	18
2017.3	74,9	19
2017.4	76,4	20

Гиперболический тренд: насыщение, функция определена только при положительных значениях t.

$$\hat{y}(t) = 67{,}23 - 43{,}05/t, \quad \hat{R}^2 = 0{,}634.$$

Гиперболический тренд: модификация

$$\hat{y}(t) = 123 - 2004/(t+22), \quad \hat{R}^2 = 0{,}996.$$

Логарифмический тренд: наиболее медленный неограниченный рост, функция определена только для положительных t, не очень понятна интерпретация.

$$\hat{y}(t) = 28{,}80 + 14{,}50 \ln t, \quad \hat{R}^2 = 0{,}950.$$

Логарифмический тренд: модификация

$$\hat{y}(t) = -32{,}25 + 32{,}75 \ln(t-7), \quad \hat{R}^2 = 0{,}997.$$

Экспоненциальный тренд: постоянный относительный прирост.

$$\hat{y}(t) = 39{,}94 e^{0{,}036t}, \quad \hat{R}^2 = 0{,}929.$$

Степенной тренд: постоянная эластичность, не очень понятна интерпретация для временных рядов!

$$\hat{y}(t) = 11{,}05 t^{0{,}59}, \quad \hat{R}^2 = 0{,}881.$$

Аналитические тренды. Прогноз

	y	*t*	лин	квад	гип	лог	эксп	степ	лин	квад	гип	лог	эксп	степ
2013.1	35,8	1	40,3	37,1	36,1	35,8	41,4	11,0	-4,5	-1,3	-0,3	0,0	-5,6	24,8
2013.2	39,2	2	42,4	40,1	39,7	39,7	42,9	16,7	-3,2	-0,9	-0,5	-0,5	-3,7	22,5
2013.3	42,8	3	44,4	43,1	43,0	43,1	44,5	21,2	-1,6	-0,3	-0,2	-0,3	-1,7	21,6
2013.4	46,5	4	46,4	45,9	46,1	46,3	46,1	25,1	0,1	0,6	0,4	0,2	0,4	21,4
2014.1	49,0	5	48,4	48,6	49,0	49,1	47,8	28,7	0,6	0,4	0,0	-0,1	1,2	20,3
2014.2	52,6	6	50,4	51,2	51,6	51,7	49,5	32,0	2,2	1,4	1,0	0,9	3,1	20,6
2014.3	55,7	7	52,4	53,6	54,1	54,2	51,3	35,0	3,3	2,1	1,6	1,5	4,4	20,7
2014.4	57,6	8	54,5	56,0	56,4	56,4	53,2	37,9	3,1	1,6	1,2	1,2	4,4	19,7
2015.1	58,2	9	56,5	58,3	58,6	58,5	55,2	40,7	1,7	-0,1	-0,4	-0,3	3,0	17,5
2015.2	59,9	10	58,5	60,4	60,6	60,5	57,2	43,3	1,4	-0,5	-0,7	-0,6	2,7	16,6
2015.3	61,6	11	60,5	62,4	62,5	62,4	59,3	45,8	1,1	-0,8	-0,9	-0,8	2,3	15,8
2015.4	63,4	12	62,5	64,3	64,3	64,2	61,4	48,2	0,9	-0,9	-0,9	-0,8	2,0	15,2
2016.1	65,1	13	64,5	66,1	66,0	65,8	63,7	50,6	0,6	-1,0	-0,9	-0,7	1,4	14,5
2016.2	66,9	14	66,5	67,8	67,5	67,4	66,0	52,9	0,4	-0,9	-0,6	-0,5	0,9	14,0
2016.3	68,5	15	68,6	69,3	69,0	69,0	68,4	55,1	-0,1	-0,8	-0,5	-0,5	0,1	13,4
2016.4	70,2	16	70,6	70,7	70,5	70,4	70,9	57,2	-0,4	-0,5	-0,3	-0,2	-0,7	13,0
2017.1	71,9	17	72,6	72,1	71,8	71,8	73,5	59,3	-0,7	-0,2	0,1	0,1	-1,6	12,6
2017.2	73,6	18	74,6	73,3	73,1	73,2	76,2	61,4	-1,0	0,3	0,5	0,4	-2,6	12,2
2017.3	74,9	19	76,6	74,4	74,3	74,4	79,0	63,4	-1,7	0,5	0,6	0,5	-4,1	11,5
2017.4	76,4	20	78,6	75,4	75,5	75,7	81,8	65,3	-2,2	1,0	0,9	0,7	-5,4	11,1
		21	80,6	76,2	76,6	76,9	84,8	67,2						
		28	94,8	79,0	83,1	84,2	109,1	79,7						
		40	118,9	70,8	90,9	93,8	167,7	98,5						
		120	280,1	-408	109,1	126,4	2959	189,1						

Алгоритмические тренды. Скользящее среднее

Задача: сглаживание значений временного ряда по *m* предыдущим, *m* последующим и текущей точке для устранения краткосрочных колебаний и выявления тенденции:

$$MA_t = \sum_{k=-m}^{m} w_k y_{t+k} = w_{-m} y_{t-m} + \ldots + w_{-1} y_{t-1} + w_0 y_t + w_1 y_{t+1} + \ldots + w_m y_{t+m}.$$

w_k – весовые коэффициенты, $\sum_{k=-m}^{m} w_k = 1.$

Весовые коэффициенты для скользящего среднего обычно симметричны ($w_k = w_{-k}$), однако традиционное для коэффициентов свойство неотрицательности ($w_k \geq 0$) для скользящего среднего выполняется не всегда.

На практике в качестве скользящего среднего часто используют **простое среднее арифметическое**, однако во многих случаях (когда предполагаемый тренд отличен от линейного вида) оптимальные весовые коэффициенты **не будут совпадать между собой**.

Наилучшие значения весовых коэффициентов w_k

Наилучшие (в смысле МНК) значения весовых коэффициентов выбираем в зависимости от ширины окна m и порядка аппроксимирующего полинома p.

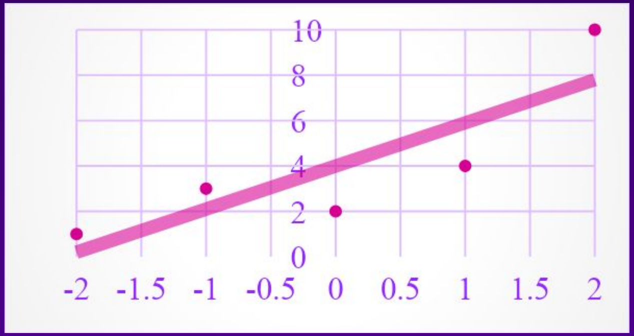
m	Pd	w_{-m}	w_{-m+1}	…	w_0
m_0	0 или 1	$\frac{1}{2m_0+1}$	$\frac{1}{2m_0+1}$	$\frac{1}{2m_0+1}$	$\frac{1}{2m_0+1}$
2	2 или 3	-3/35	12/35		17/35
3	2 или 3	-2/21	3/21	6/21	7/21
4	2 или 3	-21/231	14/231	39/231 54/231	59/231
3	4 или 5	5/231	-30/231	75/231	131/231
4	4 или 5	15/429	-55/429	30/429 135/429	179/429

Пример: скользящее среднее по 5 точкам при квадратичном тренде

$$MA_t = -\frac{3}{35}y_{t-2} + \frac{12}{35}y_{t-1} + \frac{17}{35}y_t + \frac{12}{35}y_{t+1} - \frac{3}{35}y_{t+2}.$$

Вывод весовых коэффициентов для квадратичного полинома

Решение задачи поиска наилучших весовых коэффициентов рассмотрим на примере $p = 2$, $m = 2$, т.е. для $T_t = \theta_0 + \theta_1 t + \theta_2 t^2$.



Критерий метода наименьших квадратов:

$$\sum \varepsilon_i = \left(y_{-2} - \theta_0 - \theta_1(-2) - \theta_2(-2)^2\right)^2 + \left(y_{-1} - \theta_0 - \theta_1(-1) - \theta_2(-1)^2\right)^2 +$$

$$+\left(y_0 - \theta_0 - \theta_1 \cdot 0 - \theta_2 \cdot 0^2\right)^2 + \left(y_1 - \theta_0 - \theta_1 \cdot 1 - \theta_2 \cdot 1^2\right)^2 + \left(y_2 - \theta_0 - \theta_1 \cdot 2 - \theta_2 \cdot 2^2\right)^2 =$$

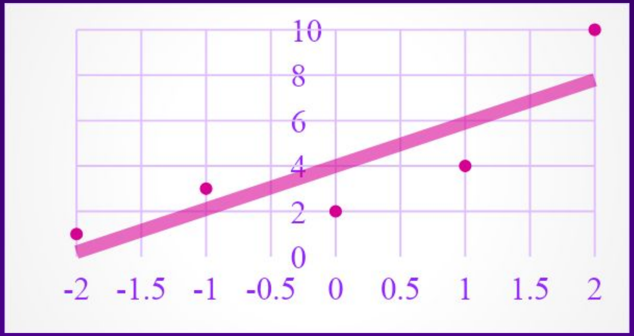
$$= \left(y_{-2} - \theta_0 + 2\theta_1 - 4\theta_2\right)^2 + \left(y_{-1} - \theta_0 + \theta_1 - \theta_2\right)^2 + \left(y_0 - \theta_0\right)^2 +$$

$$+\left(y_1 - \theta_0 - \theta_1 - \theta_2\right)^2 - \left(y_2 - \theta_0 - 2\theta_1 - 4\theta_2\right)^2 \to \min_{\theta_0, \theta_1, \theta_2}$$

Решение данной системы из 3 линейных уравнений:

$$\hat{\theta}_0 = -\frac{3}{35} y_{-2} + \frac{12}{35} y_{-1} + \frac{17}{35} y_0 + \frac{12}{35} y_1 - \frac{3}{35} y_2.$$

Определение скользящего среднего в краевых точках



По обычным формулам невозможно найти скользящее среднее в первых m и последних m точках. Однако **для коротких временных рядов** эти значения могут быть очень важны. В этом случае:

$$MA_t = \hat{\theta}_0 + \hat{\theta}_1 t + \ldots + \hat{\theta}_p t^p, \quad t = 1, \ldots, m,$$

$\hat{\theta}, \theta_1, \ldots, \hat{\theta}_p$ – коэффициенты полинома степени p, построенного по первым $(2m+1)$ точкам.

$$MA_t = \hat{\theta}_0 + \hat{\theta}_1 t + \ldots + \hat{\theta}_p t^p, \quad t = T - m + 1, \ldots, T,$$

$\hat{\theta}, \theta_1, \ldots, \hat{\theta}_p$ – коэффициенты полинома степени p, построенного по последним $(2m+1)$ точкам.

Для длинных временных рядов первые m и последние m значений скользящего среднего обычно не вычисляются.

Определение скользящего среднего по четному числу точек

Скользящее среднее иногда применяется для устранения сезонных и иных циклических колебаний, и нам потребуется проводить усреднение по четному числу точек:

1) Помесячные данные, усреднение за год – 12 точек;
2) Поквартальные/посезонные данные, усреднение за год – 4 точки;
3) Почасовые данные, усреднение за сутки – 24 точки.

Простое усреднение по периоду, не равному циклу, дает смещенные оценки. Например, при летнем пике

$MA_{\text{лето}} = \left(y_{\text{весна}} + y_{\text{лето}} + y_{\text{осень}}\right)/3$ завышает результаты;

$MA_{\text{лето}} = \left(y_{\text{зима}} + y_{\text{весна}} + y_{\text{лето}} + y_{\text{осень}} + y_{\text{зима}}\right)/5$ занижает результаты.

Решение: крайние значения берем с вдвое меньшим весом.

Например, в случае линейного тренда

для усреднения по сезонам $w_{-2} = w_2 = 1/8$, остальные $w_k = 1/4$,

для усреднения по месяцам $w_{-6} = w_6 = 1/24$, остальные $w_k = 1/12$,

для усреднения по часам $w_{-12} = w_{12} = 1/48$, остальные $w_k = 1/24$.

Экспоненциально взвешенное скользящее среднее

Скользящее среднее применялось для **интерполяции данных** (усреднения внутри диапазона временного ряда). Если мы хотим осуществить прогноз, потребуются другие методы.

При экстраполяции важно дисконтирование наблюдений (последние более важны, чем более старые).

Для стационарного временного ряда (с неизменным средним и дисперсией) наилучший прогноз EMA_t является решением задачи

$$\left(y_t - EMA_t\right)^2 + \lambda\left(y_{t-1} - EMA_t\right)^2 + \ldots + \lambda^{m-1}\left(y_{t-m+1} - EMA_t\right)^2 \to \min,$$

$$EMA_t = \frac{y_t + \lambda y_{t-1} + \lambda^2 y_{t-2} + \ldots + \lambda^{m-1} y_{t-m+1}}{1 + \lambda + \lambda^2 + \ldots + \lambda^{m-1}} =$$

$$= \frac{1-\lambda}{1-\lambda^m}\left(y_t + \lambda y_{t-1} + \lambda^2 y_{t-2} + \ldots + \lambda^{m-1} y_{t-m+1}\right).$$

Для длинных временных рядов $EMA_t \approx \lambda \cdot EMA_{t-1} + (1-\lambda) \cdot y_t$.

Сезонность и ее устранение с помощью скользящего среднего

Сезонность: $\hat{y}_t = T_t + S_t$ – аддитивная форма,

$\hat{y}_t = T_t \cdot S_t$ – мультипликативная форма.

Способы устранения сезонной компоненты:

1. Метод дамми-переменных.
2. Использование скользящего среднего.

Алгоритм:

1. Выравнивание ряда с помощью скользящего среднего по 4 сезонам, 12 месяцам и т.д.
2. Расчет сезонной компоненты S_t и ее корректировка (для аддитивной формы сезонность должна быть в среднем нулевой, для мультипликативной – единичной).
3. Устранение сезонной компоненты $y_t - S_t$ или y_t / S_t.
4. Построение тренда T_t.
5. Получение прогнозных значений $T_t + S_t$ или $T_t \cdot S_t$.
6. Расчет ошибок ε_t, вычисление коэффициента детерминации.

Численный пример. Аддитивная сезонность

	y	MA	y-MA	S	y – S	T	T+S	ε
весна13	1,5			-0,280	1,780	1,741	1,462	0,038
лето13	2,6			0,780	1,820	1,803	2,582	0,018
осень13	1,7	1,663	0,037	0,183	1,517	1,864	2,047	-0,347
зима13	0,9	1,700	-0,800	-0,683	1,583	1,926	1,243	-0,343
весна14	1,4	1,888	-0,488	-0,280	1,680	1,987	1,707	-0,307
лето14	3	2,113	0,888	0,780	2,220	2,048	2,828	0,172
осень14	2,8	2,263	0,538	0,183	2,617	2,110	2,293	0,507
зима14	1,6	2,350	-0,750	-0,683	2,283	2,171	1,489	0,111
весна15	1,9	2,363	-0,463	-0,280	2,180	2,233	1,953	-0,053
лето15	3,2	2,400	0,800	0,780	2,420	2,294	3,074	0,126
осень15	2,7	2,488	0,213	0,183	2,517	2,356	2,539	0,161
зима15	2	2,550	-0,550	-0,683	2,683	2,417	1,734	0,266
весна16	2,2	2,563	-0,363	-0,280	2,480	2,479	2,199	0,001
лето16	3,4	2,563	0,838	0,780	2,620	2,540	3,320	0,080
осень16	2,6	2,663	-0,062	0,183	2,417	2,602	2,784	-0,184
зима16	2,1	2,738	-0,638	-0,683	2,783	2,663	1,980	0,120
весна17	2,9	2,713	0,188	-0,280	3,180	2,724	2,445	0,455
лето17	3,3	2,713	0,588	0,780	2,520	2,786	3,566	-0,266
осень17	2,5			0,183	2,317	2,847	3,030	-0,530
зима17	2,2			-0,683	2,883	2,909	2,226	-0,026

Сезонность:

$$\tilde{S}_1 = \begin{pmatrix} -0{,}488 - 0{,}463 \\ -0{,}363 + 0{,}188 \end{pmatrix} / 4,$$

$$\tilde{S}_1 = -0{,}281, \;\; \tilde{S}_2 = 0{,}778,$$

$$\tilde{S}_3 = 0{,}181, \;\; \tilde{S}_4 = -0{,}684.$$

Корректировка:

$$\left(\tilde{S}_1 + \ldots + \tilde{S}_4\right)/4 = -0{,}015,$$

$$S_1 = -0{,}280, \;\; S_2 = 0{,}780,$$

$$S_3 = 0{,}183, \;\; S_4 = -0{,}683.$$

Тренд:

$$\hat{T}_t = 1{,}680 + 0{,}061t.$$

Точность модели:

$$\hat{K}_d = 1 - \frac{D\varepsilon}{Dy} = 0{,}850.$$

Численный пример. Мультипликативная сезонность

	y	MA	y/MA	S	y/S	T	$T \cdot S$	ε
весна13	1,5			0,876	1,712	1,675	1,468	0,032
лето13	2,6			1,336	1,946	1,745	2,332	0,268
осень13	1,7	1,663	1,023	1,090	1,559	1,815	1,979	-0,279
зима13	0,9	1,700	0,529	0,697	1,292	1,885	1,313	-0,413
весна14	1,4	1,888	0,742	0,876	1,597	1,955	1,713	-0,313
лето14	3	2,113	1,420	1,336	2,245	2,025	2,706	0,294
осень14	2,8	2,263	1,238	1,090	2,568	2,095	2,284	0,516
зима14	1,6	2,350	0,681	0,697	2,296	2,165	1,509	0,091
весна15	1,9	2,363	0,804	0,876	2,168	2,235	1,959	-0,059
лето15	3,2	2,400	1,333	1,336	2,395	2,305	3,081	0,119
осень15	2,7	2,488	1,085	1,090	2,476	2,375	2,590	0,110
зима15	2	2,550	0,784	0,697	2,870	2,445	1,704	0,296
весна16	2,2	2,563	0,859	0,876	2,510	2,515	2,204	-0,004
лето16	3,4	2,563	1,327	1,336	2,544	2,585	3,455	-0,055
осень16	2,6	2,663	0,977	1,090	2,384	2,655	2,896	-0,296
зима16	2,1	2,738	0,767	0,697	3,014	2,726	1,899	0,201
весна17	2,9	2,713	1,069	0,876	3,309	2,796	2,450	0,450
лето17	3,3	2,713	1,217	1,336	2,469	2,866	3,830	-0,530
осень17	2,5			1,090	2,293	2,936	3,201	-0,701
зима17	2,2			0,697	3,157	3,006	2,094	0,106

Сезонность:

$$S_1 = \begin{pmatrix} 0{,}742 + 0{,}804 \\ + 0{,}859 + 1{,}069 \end{pmatrix} / 4,$$

$$\tilde{S}_1 = 0{,}868, \quad \tilde{S}_2 = 1{,}324,$$

$$\tilde{S}_3 = 1{,}081, \quad \tilde{S}_4 = 0{,}690.$$

Корректировка:

$$\left(\tilde{S}_1 + \ldots + \tilde{S}_4\right)/4 = 0{,}991,$$

$$S_1 = 0{,}876, \quad S_2 = 1{,}336,$$

$$S_3 = 1{,}090, \quad S_4 = 0{,}697.$$

Тренд:

$$\hat{T}_t = 1{,}680 + 0{,}061t.$$

Точность модели:

$$\hat{K}_d = 1 - \frac{D\varepsilon}{Dy} = 0{,}778.$$

Спасибо за внимание!

alexander.filatov@gmail.com
http://vk.com/alexander.filatov, **http://vk.com/baikalreadings**