

1. Выбор модели временного ряда по компромиссу «точность-надежность»

Анализ точности

Заранее выбран класс функций для оценки трендов временного ряда $X(t)$ -степенные полиномы. Задан диапазон $l=1-4$.

$$X_t = P_l(t) + \delta_{lt} = \sum_{i=0}^l a_i t^i + \delta_{lt}, \text{ где } l, a_i - \text{неизвестны.} \quad (1)$$

Полином порядка $l - P_l(t)$ выбирается как наиболее точный из допустимых по надежности. Основа процедуры подбора – сопоставление различных функций – полиномов $P_l(t)$ порядка l по степени их близости к ряду наблюдений. Критерий точности - $S^2(l)$ -минимум среднего квадрата невязок для каждого l , принадлежащего $[1, N-2]$ (в рамках заданного диапазона l этот показатель медленно убывающая функция от l). Но с ростом l , несмотря на увеличение точности, уменьшается статистическая надежность оценок полиномов, измеренная с помощью коэффициента вариации. Поэтому принятие решения о "наилучшем" полиноме порядка $l=k$ - разрешение компромисса "точность -надежность".

1. Решение задачи минимизации S^2 для $l=1,2,3, 4$.

Система нормальных уравнений имеет следующий вид:

$$\begin{array}{rclclclcl} 722666a_4 & + & 44100a_3 & + & 2870a_2 & + & 210a_1 & + & 20a_0 & = & 641,7 \\ 12333300a_3 & + & 722666a_3 & + & 44100a_2 & + & 2870a_1 & + & 210a_0 & = & 8499.9 \\ 216455810a_3 & + & 12333300a_3 & + & 722666a_2 & + & 44100a_1 & + & 2870a_0 & = & 130137.7 \\ 3877286700a_3 & + & 216455810a_3 & + & 12333300a_2 & + & 722666a_1 & + & 44100a_0 & = & 2136701.1 \\ 70540730666a_3 & + & 3877286700a_3 & + & 216455810a_2 & + & 12333300a_1 & + & 722666a_0 & = & 36584637.7 \end{array}$$

В результате ее решения получены следующие оценки коэффициентов a_i :

	$l=1$	$l=2$	$l=3$	$l=4$
a_0	4,263	8,871	7,967	10,875
a_1	2,650	1,393	1,854	-0,421
a_2		0,06	0,0062	0,467
a_3			0,0017	-0,032
a_4				$7,983 \cdot 10^{-4}$

Далее для каждого полинома $P_1(t)$, $P_2(t)$, $P_3(t)$ были получены показатели точности:

$$\begin{array}{ll} S^2(1) & = 3,950 \\ S^2(2) & = 0,756 \\ S^2(3) & = 0,696 \\ S^2(4) & = 0,595 \end{array}$$

$$\text{где } S^2(l) = \frac{1}{20} \min_{a_i} \sum_{t=1}^{20} [X(t) - P_l(t)]^2, \quad l=1,2,3,4$$

S^2 падает - точность растет по мере роста порядка полинома.

2. На втором этапе необходимо провести проверку на автокорреляцию случайных остатков δ_{it} уравнения (1). Модель временного ряда, построенная с помощью полинома порядка l признается допустимой, если случайные составляющие δ_{it} представляют собой белый шум. В этом случае, по теореме Гаусса-Маркова, оценки a_i , полученные из регрессионной модели (1) по методу МНК будут несмещенными и эффективными.

Проверка на автокорреляцию может быть проведена, например, с помощью теста Дарбина-Уотсона. В данной задаче для каждого из разбиений временного ряда на детерминированную (полином $P_l(t)$, $l=1,2,3,4$) и случайную ($\delta_l(t)$, $l=1,2,3,4$) составляющие (см. таблицу).

t	X_t	$P_1(t)$	$\delta_1(t)$	$P_2(t)$	$\delta_2(t)$	$P_3(t)$	$\delta_3(t)$	$P_4(t)$	$\delta_4(t)$
1	10.8	6.91286	3.88714	10.32422	0.47578	9.82934	0.97066	10.89080	-0.09080
2	12.1	9.56256	2.53744	11.89665	0.20335	11.71432	0.38568	11.65977	0.44023
3	12.7	12.21226	0.48774	13.58877	-0.88877	13.63218	-0.93218	13.01866	-0.31866
4	15	14.86195	0.13805	15.40059	-0.40059	15.59313	-0.59313	14.82336	0.17664
5	16.5	17.51165	-1.01165	17.33211	-0.83211	17.60738	-1.10738	16.94894	-0.44894
6	19.1	20.16135	-1.06135	19.38332	-0.28332	19.68515	-0.58515	19.28960	-0.18960
7	21.6	22.81105	-1.21105	21.55423	0.04577	21.83666	-0.23666	21.75872	-0.15872
8	24.5	25.46075	-0.96075	23.84484	0.65516	24.07211	0.42789	24.28884	0.21116
9	27.4	28.11045	-0.71045	26.25515	1.14485	26.40172	0.99828	26.83165	0.56835
10	30.8	30.76015	0.03985	28.78515	2.01485	28.83571	1.96429	29.35800	1.44200
11	31.9	33.40985	-1.50985	31.43485	0.46515	31.38429	0.51571	31.85791	0.04209
12	34	36.05955	-2.05955	34.20425	-0.20425	34.05767	-0.05767	34.34055	-0.34055
13	36.1	38.70925	-2.60925	37.09334	-0.99334	36.86607	-0.76607	36.83425	-0.73425
14	39.4	41.35895	-1.95895	40.10213	-0.70213	39.8197	-0.4197	39.38649	0.01351
15	42.7	44.00865	-1.30865	43.23062	-0.53062	42.92878	-0.22878	42.06394	0.63606
16	45.8	46.65835	-0.85835	46.4788	-0.6788	46.20353	-0.40353	44.95239	0.84761
17	49.4	49.30805	0.09195	49.84668	-0.44668	49.65414	-0.25414	48.15681	1.24319
18	53.6	51.95774	1.64226	53.33426	0.26574	53.29085	0.30915	51.80134	1.79866
19	55.8	54.60744	1.19256	56.94153	-1.14153	57.12386	-1.32386	56.02925	-0.22925
20	62.5	57.25714	5.24286	60.66851	1.83149	61.16339	1.33661	61.00300	1.49700

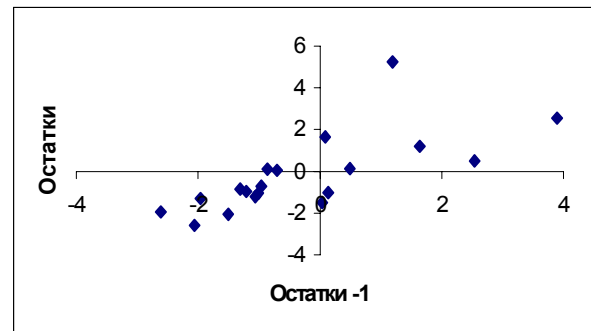
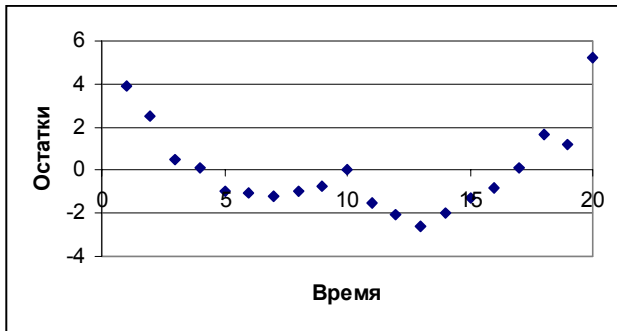
были рассчитаны коэффициенты автокорреляции:

$r_1(1)=0,717$, $r_1(2)=0,313$, $r_1(3)=0,289$, $r_1(4)=0,266$

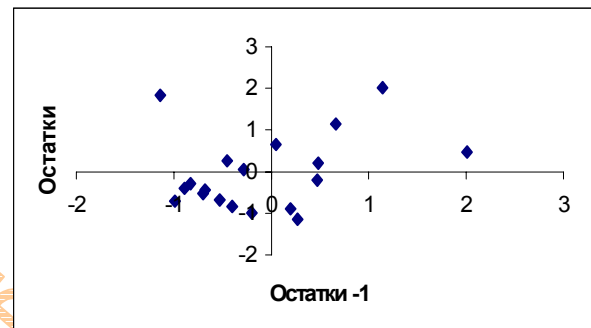
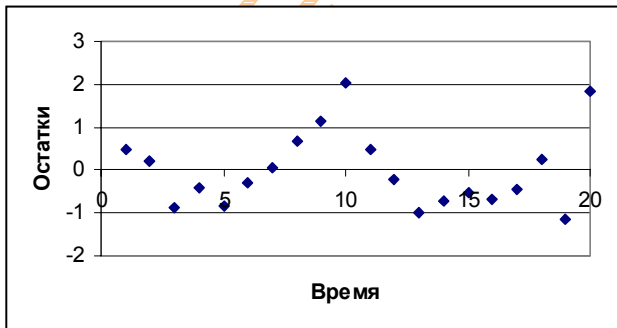
Близость r_1 к 1 свидетельствует о наличии линейного тренда в остатках. При $l=1$ коэффициент автокорреляции "плохой", но условно принято предположение о допустимости $P_1(t)$, $P_2(t)$, $P_3(t)$.

Графический анализ остатков

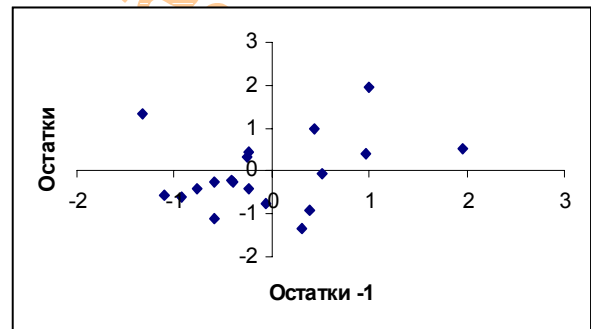
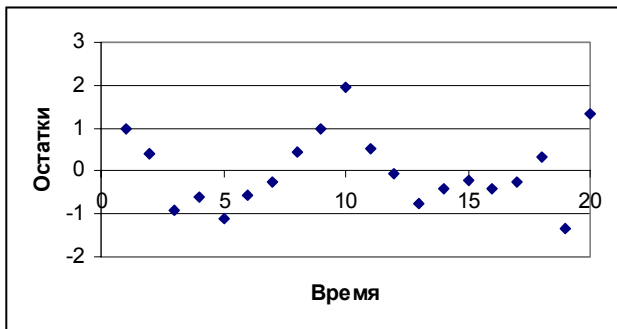
I = 1



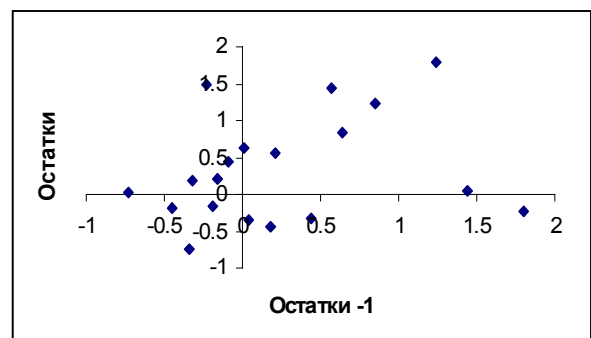
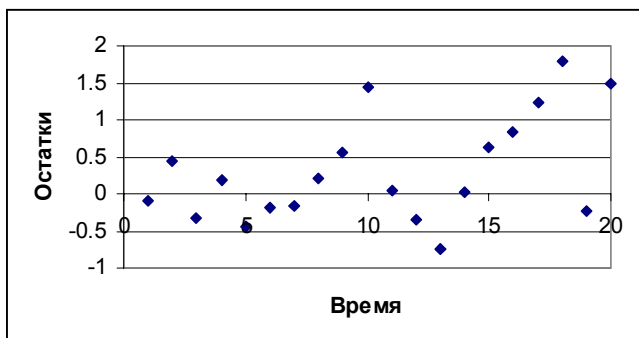
I = 2



I = 3



I = 4



Анализ надежности

Для анализа надежности полиномов были вычислены оценки дисперсии $\sigma_{\delta}^2(l)$ случайных составляющих δ_{it} .

$$\sigma_{\delta}^2(l) = \frac{S^2(l)20}{20-l-1}, l=1,2,3,4$$

$$\sigma_{\delta}^2(1) = 4,388, \sigma_{\delta}^2(2) = 0,894, \sigma_{\delta}^2(3) = 0,870, \sigma_{\delta}^2(4) = 0,793$$

Дисперсии оценок параметров $\hat{a}_{ii}$ полиномов $P_1(t)$, $P_2(t)$, $P_3(t)$, $P_4(t)$

$$\sigma_i^2(l) = a^{ii}(l) \cdot \sigma_{\delta}^2(l), i = 0,1,\dots,l, l=4,$$

где $a^{ii}(l)$ - i ый диагональный элемент обратной матрицы соответствующей системы нормальных уравнений, который может быть вычислен независимо от значений $X(t)$ в силу свойств регрессий в виде степенных функций от t .

	$\sigma_0(l)$	$\sigma_1(l)$	$\sigma_2(l)$	$\sigma_3(l)$	$\sigma_4(l)$
$l=1$	$2,166 \cdot 10^{-1}$	$8,123 \cdot 10^{-2}$			
$l=2$	$7,034 \cdot 10^{-1}$	$1,543 \cdot 10^{-1}$	$7,136 \cdot 10^{-3}$		
$l=3$	$1,019 \cdot 10^0$	$4,096 \cdot 10^{-1}$	$4,478 \cdot 10^{-2}$	$1,404 \cdot 10^{-3}$	
$l=4$	$1,387 \cdot 10^0$	$8,669 \cdot 10^{-1}$	$1,623 \cdot 10^{-1}$	$1,148 \cdot 10^{-2}$	$2,695 \cdot 10^{-4}$

Из таблицы СКО видно, что дисперсии оценок убывают по мере роста их номера i , таким образом, абсолютная надежность увеличивается, но более важен показатель относительной надежности старшего коэффициента полинома - коэффициент вариации:

$$B_i(l) = \frac{\sigma_i(l)}{|\hat{a}_{ii}|} < b = 0,3 - \text{(ограничение в соответствии с личными предпочтениями)}$$

Таблица $B_i(l)$

	$B_0(l)$	$B_1(l)$	$B_2(l)$	$B_3(l)$	$B_4(l)$
$l=1$	0,0508	0,0306			
$l=2$	0,0793	0,1108	0,1189		
$l=3$	0,1279	0,2209	7,2226	0,8259	
$l=4$	0,1275	2,0591	0,3475	0,3588	0,3376

Данному критерию соответствуют полиномы первого и второго порядков, они будут допустимыми с точки зрения надежности, однако из них полином $P_2(t)$ является более точным. Следовательно, $k=2$.

Прогноз на 23 период

$$P_2(23) = 8,871 + 1,393 \cdot 23 + 0,06 \cdot 23^2 = 72,65$$