

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ  
РОССИЙСКОЙ ФЕДЕРАЦИИ

МГТУ им Н.Э.Баумана

Факультет ФН

Кафедра вычислительной математики и математической физики

Соколов Арсений Андреевич

Лабораторная работа №2 по численным  
методам

3 курс, группа ФН11-53Б

Вариант 9

Преподаватель

\_\_\_\_\_ В. А. Кутыркин

«\_\_\_» \_\_\_\_\_ 2019 г.

Москва, 2019 г.

## Задание 2.1

**Задание.** Дана модель линейной регрессии:

$$Y = x_*^0 + z_1 x_*^1 + z_2 x_*^2 + z_3 x_*^3 + z_4 x_*^4 + z_5 x_*^5 + z_6 x_*^6 + \varepsilon \quad (1)$$

Для оценки неизвестных вектора тренда  $\succ x_* = [x_*^0, \dots, x_*^k] \in \succ \mathbb{E}^{k+1}$  и параметра  $\sigma$  случайной составляющей  $\varepsilon \sim N(0, \sigma)$  модели линейной регрессии 1 проводился эксперимент, в котором получены  $m = 20$  значений  $y^1, \dots, y^m \in \mathbb{R}$  регрессора модели 1 для  $m$  различных наборов  $\succ z^1 = [z_1^1, \dots, z_6^1], \succ z^m = [z_1^m, \dots, z_6^m]$  шести факторов модели 1.

Требуется получить оценки вектора тренда  $\succ x_* = [x_*^0, \dots, x_*^k] \in \succ \mathbb{E}^{k+1}$  и параметра  $\sigma$  случайной составляющей  $\varepsilon \sim N(0, \sigma)$  модели линейной регрессии 1. Если возможно, редуцировать модель регрессии 1 до приведённой модели. Результаты расчётов проиллюстрировать графически, сопроводив необходимыми комментариями.

**Исходные данные.**

$$N = 9$$

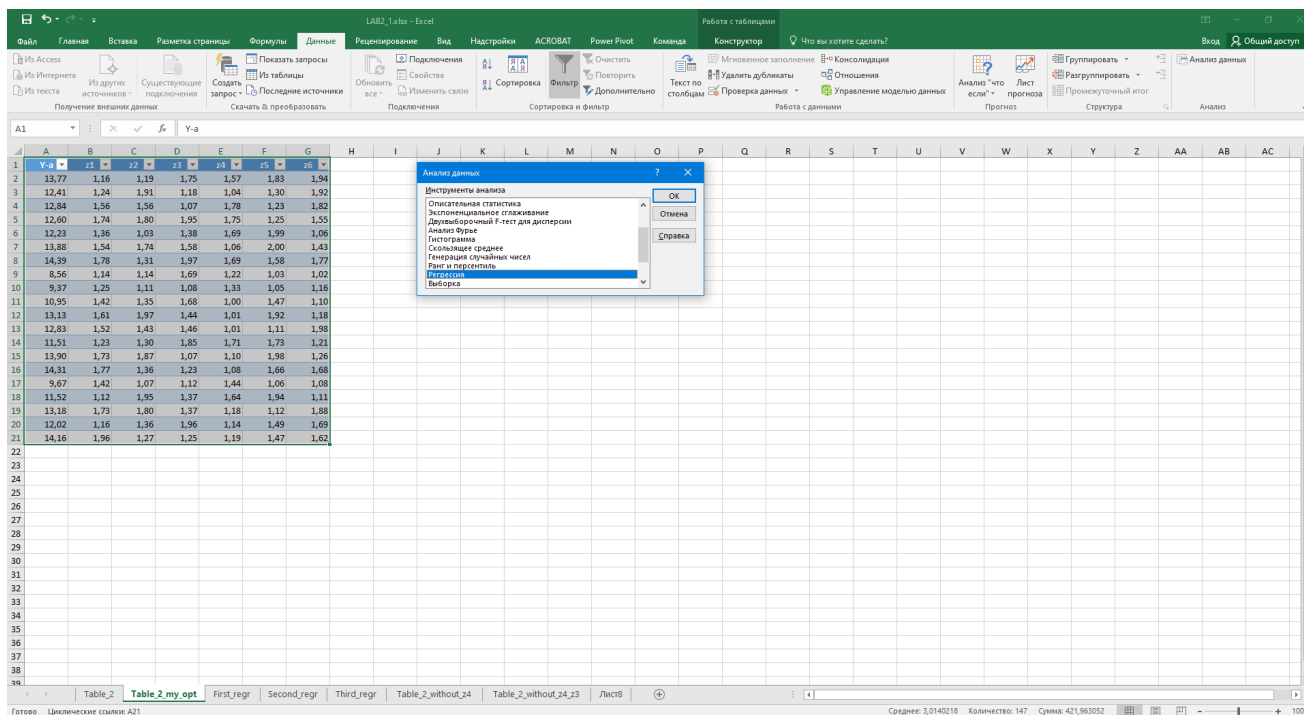
$$\alpha = 1$$

Таблица 1

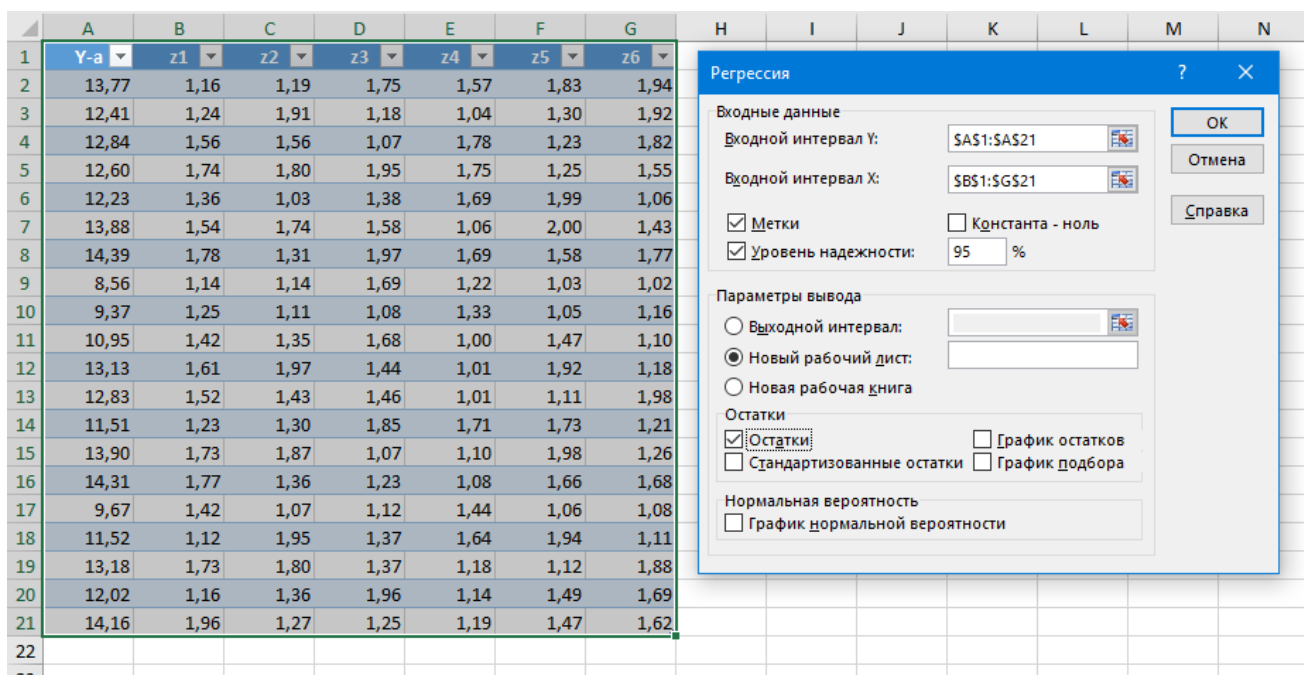
Y-a	z1	z2	z3	z4	z5	z6
13,77	1,16	1,19	1,75	1,57	1,83	1,94
12,41	1,24	1,91	1,18	1,04	1,30	1,92
12,84	1,56	1,56	1,07	1,78	1,23	1,82
12,60	1,74	1,80	1,95	1,75	1,25	1,55
12,23	1,36	1,03	1,38	1,69	1,99	1,06
13,88	1,54	1,74	1,58	1,06	2,00	1,43
14,39	1,78	1,31	1,97	1,69	1,58	1,77
8,56	1,14	1,14	1,69	1,22	1,03	1,02
9,37	1,25	1,11	1,08	1,33	1,05	1,16
10,95	1,42	1,35	1,68	1,00	1,47	1,10
13,13	1,61	1,97	1,44	1,01	1,92	1,18
12,83	1,52	1,43	1,46	1,01	1,11	1,98
11,51	1,23	1,30	1,85	1,71	1,73	1,21
13,90	1,73	1,87	1,07	1,10	1,98	1,26
14,31	1,77	1,36	1,23	1,08	1,66	1,68
9,67	1,42	1,07	1,12	1,44	1,06	1,08
11,52	1,12	1,95	1,37	1,64	1,94	1,11
13,18	1,73	1,80	1,37	1,18	1,12	1,88
12,02	1,16	1,36	1,96	1,14	1,49	1,69
14,16	1,96	1,27	1,25	1,19	1,47	1,62

## Решение.

Для получения оценки вектора тренда  $\hat{x}_* = [x_*^0, \dots, x_*^k] \in \mathbb{R}^{k+1}$  и параметра  $\sigma$  случайной составляющей  $\varepsilon \sim N(0, \sigma)$  модели линейной регрессии 1 воспользуемся программой «Анализ данных», входящей в состав Excel. Переходя на вкладку «Данные», выбираем программу «анализ данных», в которой выбираем пункт «регрессия».



В открывшемся окне заполняем поля, согласно следующему скриншоту, нажимаем «ОК»:



На новом листе видим результат первой регрессии. Обратим внимание на строки 16 : 23. Для исключения незначимых параметров, найдём строки, в которых  $|t\text{-статистика}| < 2$ . Выберем минимальное значение. В данном случае – это строка 21, соответствующая  $z_4$ .

	A	B	C	D	E	F	G	I
1	ВЫВОД ИТОГОВ							
2								
3	<i>Регрессионная статистика</i>							
4	Множественный R	0,99999824						
5	R-квадрат	0,999996479						
6	Нормированный R-квадрат	0,999994854						
7	Стандартная ошибка	0,003816318						
8	Наблюдения	20						
9								
10	<i>Дисперсионный анализ</i>							
11		<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>		
12	Регрессия	6	53,77486566	8,962477611	615373,7628	1,13942E-34		
13	Остаток	13	0,000189336	1,45643E-05				
14	Итого	19	53,775055					
15								
16		<i>Коэффициенты</i>	<i>Стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>Нижние 95%</i>	<i>Верхние 95%</i>	
17	Y-пересечение	-1,007979953	0,009492771	-106,1839531	1,71971E-20	-1,028487838	-0,987472068	
18	$z_1$	2,998635451	0,003654101	820,6218926	4,93566E-32	2,990741245	3,006529657	
19	$z_2$	0,006189067	0,003088125	2,004150255	0,066345309	-0,000482422	0,012860557	
20	$z_3$	-0,001028465	0,002993614	-0,343552884	0,736675263	-0,007495775	0,005438846	
21	$z_4$	0,000461738	0,003182317	0,145094937	0,886861373	-0,00641324	0,007336716	
22	$z_5$	2,999563826	0,002748812	1091,222033	1,21439E-33	2,99362538	3,005502273	
23	$z_6$	3,00008523	0,002824029	1062,342043	1,72103E-33	2,993984285	3,006186175	
24								
25								
26		Минимальное по модулю значение, меньше 2						
27	ВЫВОД ОСТАТКА							
28								
29	<i>Наблюдение</i>	<i>Предсказанное Y-a</i>	<i>Остатки</i>					
30	1	13,77602522	-0,006025216					
31	2	12,40238099	0,007619008					
32	3	12,84272605	-0,002726051					
33	4	12,59846337	0,00153663					

Исключая из исходной таблицы столбец, соответствующий  $z_4$  получим новую таблицу:

Таблица 2

Y-a	z1	z2	z3	z5	z6
13,77	1,16	1,19	1,75	1,83	1,94
12,41	1,24	1,91	1,18	1,30	1,92
12,84	1,56	1,56	1,07	1,23	1,82
12,60	1,74	1,80	1,95	1,25	1,55
12,23	1,36	1,03	1,38	1,99	1,06
13,88	1,54	1,74	1,58	2,00	1,43
14,39	1,78	1,31	1,97	1,58	1,77
8,56	1,14	1,14	1,69	1,03	1,02
9,37	1,25	1,11	1,08	1,05	1,16
10,95	1,42	1,35	1,68	1,47	1,10
13,13	1,61	1,97	1,44	1,92	1,18
12,83	1,52	1,43	1,46	1,11	1,98
11,51	1,23	1,30	1,85	1,73	1,21
13,90	1,73	1,87	1,07	1,98	1,26
14,31	1,77	1,36	1,23	1,66	1,68
9,67	1,42	1,07	1,12	1,06	1,08
11,52	1,12	1,95	1,37	1,94	1,11
13,18	1,73	1,80	1,37	1,12	1,88
12,02	1,16	1,36	1,96	1,49	1,69
14,16	1,96	1,27	1,25	1,47	1,62

Выполняя новую регрессию с аналогичными параметрами, получаем:

	A	B	C	D	E	F	G
1	ВЫВОД ИТОГОВ						
2							
3	Регрессионная статистика						
4	Множественный R	0,999998237					
5	R-квадрат	0,999996473					
6	Нормированный R-квадрат	0,999995214					
7	Стандартная ошибка	0,003680473					
8	Наблюдения	20					
9							
10	Дисперсионный анализ						
11		df	SS	MS	F	Значимость F	
12	Регрессия	5	53,77486536	10,75497307	793966,473	1,20788E-37	
13	Остаток	14	0,000189642	1,35459E-05			
14	Итого	19	53,775055				
15							
16		Коэффициенты	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%
17	Y-пересечение	-1,007351577	0,008146614	-123,6527937	1,1235E-22	-1,024824327	-0,98987883
18	z1	2,998606903	0,003518918	852,1389684	2,0741E-34	2,991059575	3,006154231
19	z2	0,006098587	0,002916848	2,090814315	0,05526377	-0,000157429	0,012354603
20	z3	-0,000950537	0,00284021	-0,334671293	0,74283428	-0,00704218	0,005141107
21	z5	2,999597948	0,002641246	1135,675543	3,7191E-36	2,99393304	3,005262857
22	z6	3,000082501	0,002723445	1101,576299	5,6988E-36	2,994241292	3,005923709
23							
24	Минимальное по модулю значение, меньшее двух						
25							
26	ВЫВОД ОСТАТКА						
27							
28	Наблюдение	Предсказанное Y-a	Остатки				
29	1	13,77598231	-0,005982312				
30	2	12,40244993	0,007550069				
31	3	12,84246739	-0,002467392				
32	4	12,59826077	0,001739228				
33	5	12,22634445	0,003655551				

В данном случае, незначимым параметром оказался параметр  $z_3$ . Прделаем аналогичную процедуру: исключим  $z_3$  из Таблицы 2 и выполним для полученной таблицы регрессию.

Таблица 3

Y-a	z1	z2	z5	z6
13.77	1.16	1.19	1.83	1.94
12.41	1.24	1.91	1.30	1.92
12.84	1.56	1.56	1.23	1.82
12.60	1.74	1.80	1.25	1.55
12.23	1.36	1.03	1.99	1.06
13.88	1.54	1.74	2.00	1.43
14.39	1.78	1.31	1.58	1.77
8.56	1.14	1.14	1.03	1.02
9.37	1.25	1.11	1.05	1.16
10.95	1.42	1.35	1.47	1.10
13.13	1.61	1.97	1.92	1.18
12.83	1.52	1.43	1.11	1.98
11.51	1.23	1.30	1.73	1.21
13.90	1.73	1.87	1.98	1.26
14.31	1.77	1.36	1.66	1.68
9.67	1.42	1.07	1.06	1.08
11.52	1.12	1.95	1.94	1.11
13.18	1.73	1.80	1.12	1.88
12.02	1.16	1.36	1.49	1.69
14.16	1.96	1.27	1.47	1.62

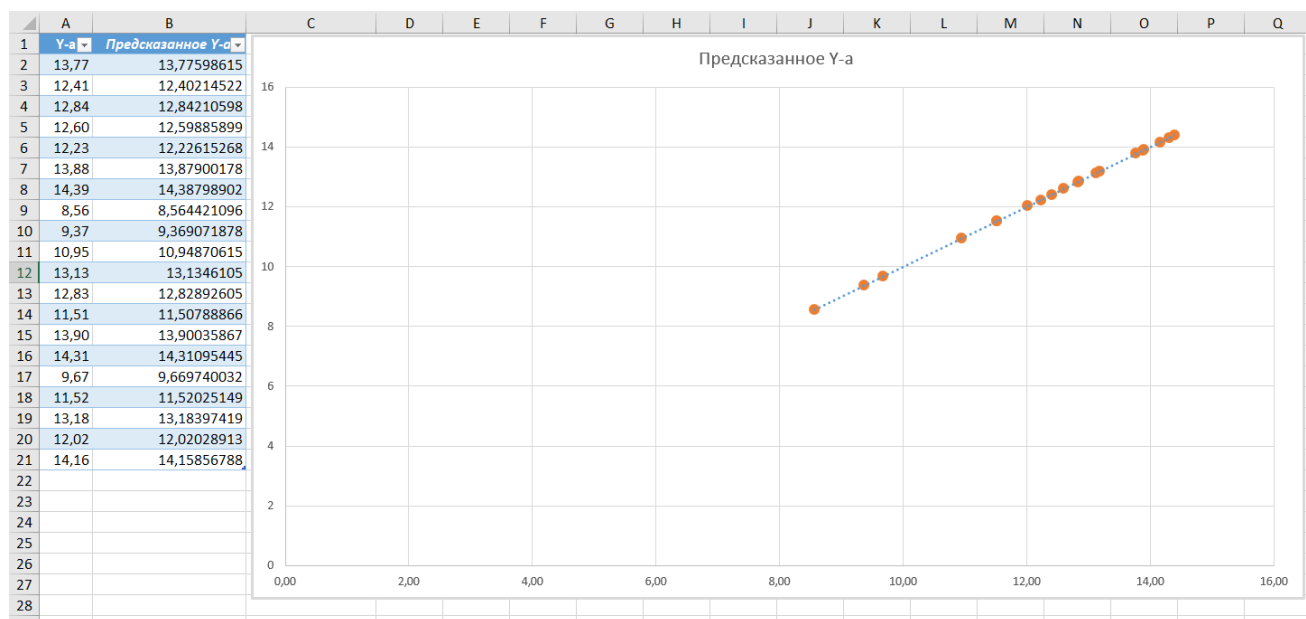
	A	B	C	D	E	F	G
1	ВЫВОД ИТОГОВ						
2							
3	Регрессионная статистика						
4	Множественный	0,999998223					
5	R-квадрат	0,999996445					
6	Нормированный	0,999995497					
7	Стандартная оши	0,003569869					
8	Наблюдения	20					
9							
10	Дисперсионный анализ						
11		df	SS	MS	F	Значимость F	
12	Регрессия	4	53,77486384	13,44371596	1054908,291	1,14956E-40	
13	Остаток	15	0,000191159	1,2744E-05			
14	Итого	19	53,775055				
15							
16		Коэффициенты	андартная ошиб	t-статистика	P-Значение	Нижние 95%	Верхние 95%
17	Y-пересечение	-1,00876898	0,006750058	-149,4459817	3,22125E-25	-1,023156387	-0,994381573
18	z1	2,998829292	0,003351761	894,7025519	7,11391E-37	2,991685183	3,005973402
19	z2	0,006253765	0,002793218	2,238910895	0,040746186	0,000300163	0,012207367
20	z5	2,999417674	0,00250803	1195,925939	9,15692E-39	2,994071936	3,004763413
21	z6	2,999901227	0,002588831	1158,785923	1,46982E-38	2,994383264	3,00541919

Мы видим, что на данном шаге уже нет значений, для которых минимум модуля  $t$ -статистики превышал бы 2. Значит полученные параметры являются значимыми. Оценка  $\sigma$  соответствует значению «Стандартная ошибка» из ячейки B7, то есть  $\hat{\sigma} = 0.003569869$ .

Кроме того, отталкиваясь от значений коэффициентов, находящихся в ячейках B17 : B21, можем выписать явное выражение:

$$Y = -1.00876898 + 2.998829292x_*^1 + 0.006253765x_*^2 + 2.999417674x_*^5 + 2.999901227x_*^6$$

На листе с последней регрессией спустимся ниже и обнаружим таблицу «Вывод остатка». Отталкиваясь от неё и от изначальных данных  $y - \alpha$  построим график:



## Задание 2.2

**Задание.** Дана модель полиномиальной (кубической) регрессии:

$$Y = x_*^0 + tx_*^1 + t^2x_*^2 + t^3x_*^3 + \varepsilon \quad (2)$$

Для оценки неизвестных вектора тренда  ${}^>x_* = [x_*^0, x_*^1, x_*^2, x_*^3] \in {}^>\mathbb{E}^4$  и параметра  $\sigma$  случайной составляющей  $\varepsilon \sim N(0, \sigma)$  модели линейной регрессии 2 проводился эксперимент, в котором получены  $m = 20$  значений  $y^1, \dots, y^m \in \mathbb{R}$  регрессора модели 2 для  $m$  попарно различных значений  $t_1, \dots, t_m \in \mathbb{R}$  единственного фактора модели 2.

Требуется получить оценки вектора тренда  ${}^>x_* = [x_*^0, x_*^1, x_*^2, x_*^3] \in {}^>\mathbb{E}^4$  и параметра  $\sigma$  случайной составляющей  $\varepsilon \sim N(0, \sigma)$  модели полиномиальной (кубической) регрессии 2. Результаты расчётов проиллюстрировать графически, сопроводив необходимыми комментариями.

**Исходные данные.**

$$N = 9$$

$$\alpha = 1$$

Таблица 4

$Y$	$t_1$	$t_1^2$	$t_1^3$
-5,75	0,05	0,0025	0,00013
-5,56	0,1	0,01	0,001
-5,34	0,15	0,0225	0,00338
-5,17	0,2	0,04	0,008
-4,98	0,25	0,0625	0,01563
-4,82	0,3	0,09	0,027
-4,65	0,35	0,1225	0,04288
-4,48	0,4	0,16	0,064
-4,31	0,45	0,2025	0,09113
-4,13	0,5	0,25	0,125
-3,93	0,55	0,3025	0,16638
-3,72	0,6	0,36	0,216
-3,49	0,65	0,4225	0,27463
-3,24	0,7	0,49	0,343
-2,95	0,75	0,5625	0,42188
-2,64	0,8	0,64	0,512
-2,29	0,85	0,7225	0,61413
-1,91	0,9	0,81	0,729
-1,48	0,95	0,9025	0,85738
-1	1	1	1



## Решение.

Для получения оценки вектора тренда  $\hat{x}_* = [x_*^0, x_*^1, x_*^2, x_*^3] \in \mathbb{E}^4$  и параметра  $\sigma$  случайной составляющей  $\varepsilon \sim N(0, \sigma)$  модели полиномиальной (кубической) регрессии 2 воспользуемся, аналогично первой части задания, программой «Анализ данных», входящей в состав Excel. Переходя на вкладку «Данные», выбираем программу «анализ данных», в которой выбираем пункт «регрессия».

Выставляя параметры регрессии, аналогичные предыдущей части задания, получим:

	A	B	C	D	E	F	G
1	ВЫВОД ИТОГОВ						
2							
3	Регрессионная статистика						
4	Множественный	0,999992029					
5	R-квадрат	0,999984058					
6	Нормированный	0,999981068					
7	Стандартная ош	0,00604793					
8	Наблюдения	20					
9							
10	Дисперсионный анализ						
11		df	SS	MS	F	Значимость F	
12	Регрессия	3	36,70913476	12,23637825	334533,3628	1,39304E-38	
13	Остаток	16	0,000585239	3,65775E-05			
14	Итого	19	36,70972				
15							
16		Коэффициенты	Стандартная ошибка	t-статистика	P-Значение	Нижние 95%	Верхние 95%
17	Y-пересечение	-5,995126935	0,006606346	-907,4799954	3,98659E-39	-6,009131763	-5,981122107
18	t1	4,944286994	0,053162981	93,00244124	2,65512E-23	4,83158651	5,056987479
19	t1^2	-4,880817482	0,116158538	-42,01858581	8,34435E-18	-5,127062581	-4,634572382
20	t1^3	4,928324842	0,072834696	67,66452102	4,25437E-21	4,773922183	5,082727501
21							
22							
23							
24	ВЫВОД ОСТАТКА						
25							
26	Наблюдение	Предсказанное Y	Остатки				
27	1	-5,759498588	0,009498588				
28	2	-5,544578086	-0,015421914				

Отметим, что все параметры полученной модели значимые.

Оценка  $\sigma$  соответствует значению «Стандартная ошибка» из ячейки B7, то есть  $\hat{\sigma} = 0.00604793$ .

Кроме того, отталкиваясь от значений коэффициентов, находящихся в ячейках B17 : B2, можем выписать явное выражение:

$$Y = -5.995126935 + 4.94428699432064t - 4.88081748197511t^2 + 4.92832484217597t^3$$

Отталкиваясь от полученных «Предсказанных  $Y$ » и изначальных данных график:

