

МИНОБРНАУКИ РОССИИ
САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
«ЛЭТИ» ИМ. В.И. УЛЬЯНОВА (ЛЕНИНА)
Кафедра АМ

ОТЧЕТ
по лабораторной работе №5
по дисциплине «Статистический анализ»
Тема: Линейная регрессия.

Студент гр. 8374

Пихтовников К.С.

Преподаватель

Чирин А.В.

Санкт-Петербург

2020

Цель работы:

Сформировать вектор из n иксов, сгенерировать гауссовский случайный шум Z , сформировать вектор из игреков и построить оценки параметров, а также оценить остаточную регрессию.

Вариант 15:

$$x_{\min} = -1.8$$

$$x_{\max} = 2.6$$

$$n = 40$$

$$a_2 = 0.3$$

$$a_1 = -1.7$$

$$a_0 = -3.3$$

$$c_1 = 7.2$$

$$c_0 = -1.4$$

$$\sigma = 1.2$$

Выполнение работы:

1. Линейная регрессия

1. Сформировать вектор n иксов (от x_{\min} до x_{\max} с равномерным шагом).

```
step = (xmax-xmin)/(n-1);  
x(n) = 0;  
x(1) = xmin;  
for i = 1:(n-1)  
    x(i+1) = xmin + step*i;  
end
```

2. Сгенерировать гауссовский случайный шум Z – выборку из n независимых случайных величин с мат.ожиданием 0 и заданной дисперсией σ^2 .

```
z = sigma*randn(n, 1).';
```

3. Сформировать вектор икрегов: $y_k = c_1 \cdot x_k + c_0 + Z_k$.

```
y = c1*x+c0+z;
```

4. Построить оценки параметров c_1 и c_0 по методу наименьших квадратов.

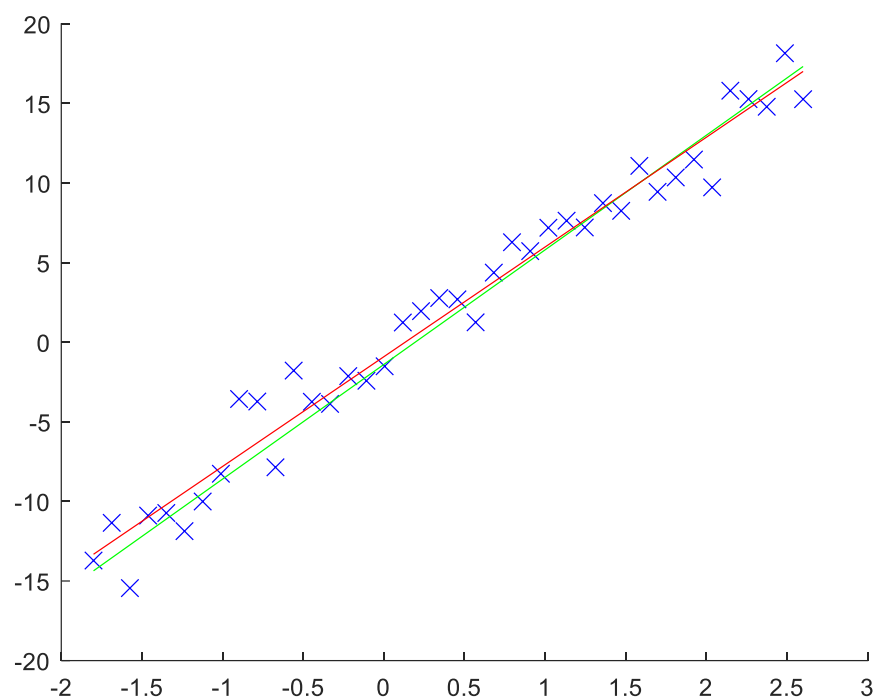
```
c = polyfit(x, y, 1);  
c1_apr = c(1);  
c0_apr = c(2);
```

Результат:

c1_apr		c0_apr		c1_apr		c0_apr	
1x1 double		1x1 double		1x1 double		1x1 double	
	1	2		1	2		2
1	6.8929			-0.9129			

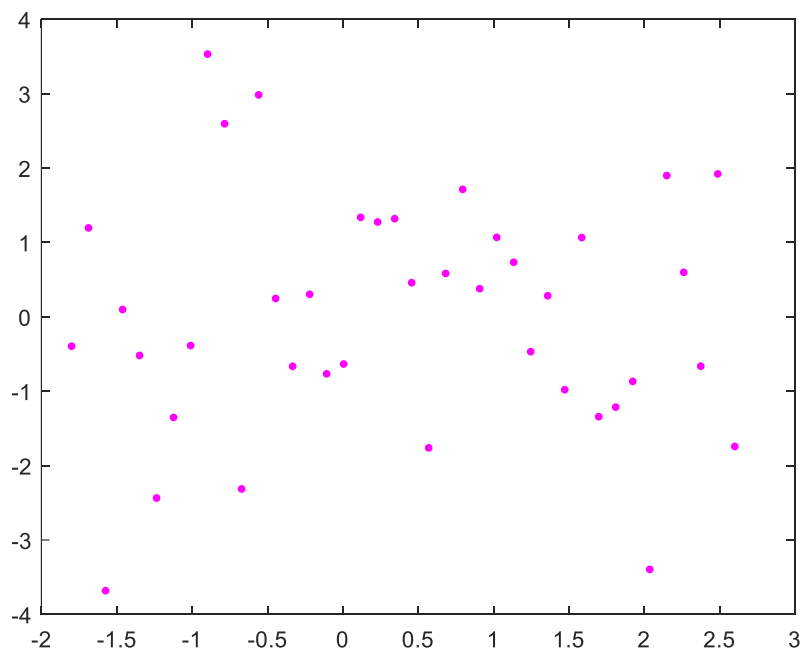
5. Графики: облака точек (x_k, y_k) , истинной зависимости $y = c_1 \cdot x + c_0$ и линию регрессии $\hat{y}_k = \hat{c}_1 \cdot x_k + \hat{c}_0$.

Полученные графики, где зеленым цветом – истинная зависимость, красным – регрессионная:



6. Остатки $\varepsilon_k = y_k - \hat{y}_k$ (в зависимости от x).

Полученный график:



7. Оценить остаточную дисперсию:

$$\hat{\sigma}^2 = \frac{\sum_{k=1}^n \varepsilon_k^2}{n-2}$$

Остаточная дисперсия =

2.6581

2. Квадратичная регрессия

1. Использовать вектор иксов из первой части.
2. Сгенерировать гауссовский случайный шум Z – выборку из n независимых случайных величин с мат.ожиданием 0 и заданной дисперсией σ^2 . (из 1 части)
3. Сформировать вектор икрегов: $y_k = a_2 \cdot x_k^2 + a_1 \cdot x_k + a_0 + Z_k$.

$$y = a_2 \cdot x.^2 + a_1 \cdot x + a_0 + z;$$

4. Построить X-матрицу плана, а также $A = X^T X$ и A^{-1} .

```
X = [x'.^2 x' ones(n, 1)];
A = X'*X;
C = inv(A);
```

```
A=
    273.0990    83.9713    74.2427
    83.9713    74.2427    16.0000
    74.2427    16.0000    40.0000

C=
    0.0109   -0.0087   -0.0167
   -0.0087    0.0217    0.0075
   -0.0167    0.0075    0.0530
```

5. Построить оценки параметров a_2 , a_1 и a_0 по методу наименьших квадратов.

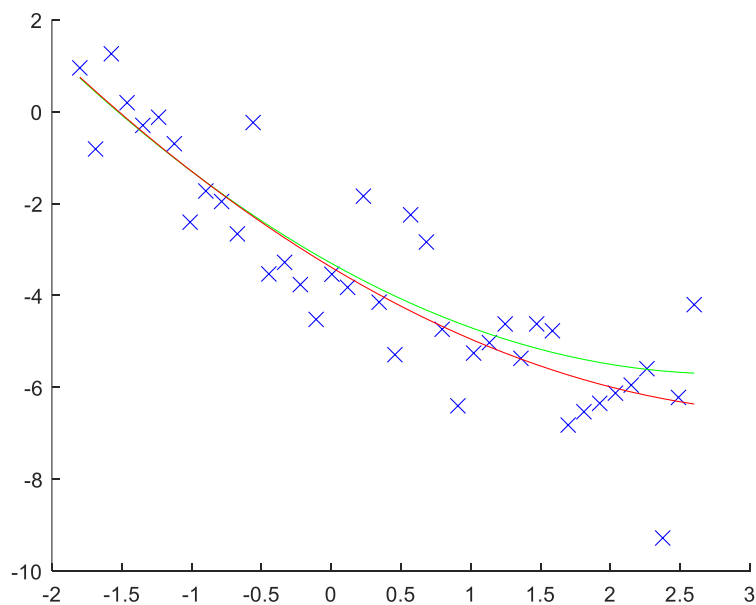
```
a = polyfit(x, y, 2);
a2Apr = a(1);
a1Apr = a(2);
a0Apr = a(3);
```

Результат:

a2Apr		a1Apr		a0Apr		
1x1 double		1x1 double		1x1 double		
	1	1	2	1	2	3
1	0.2621	-1.8269		1	-3.3866	

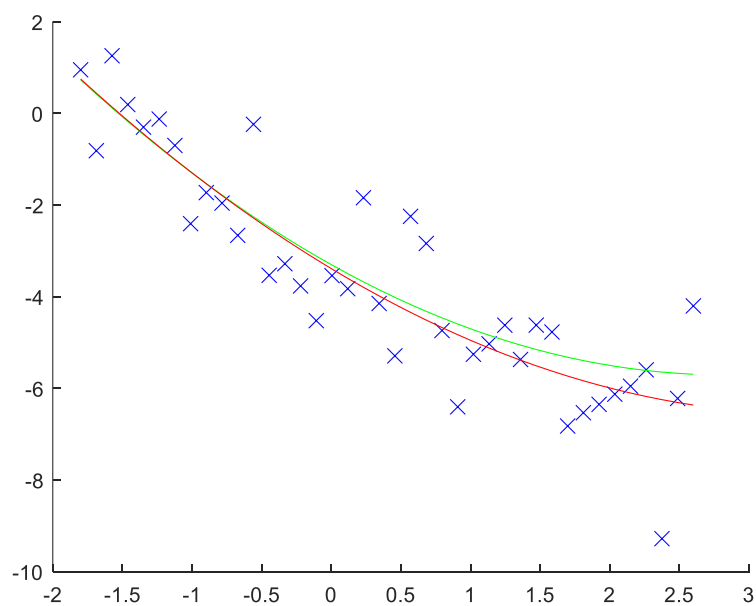
6. Графики: облака точек (x_k, y_k) , истинной зависимости $y = a_2 \cdot x^2 + a_1 \cdot x + a_0$ и линию регрессии $\hat{y}_k = \hat{a}_2 \cdot x_k^2 + \hat{a}_1 \cdot x_k + \hat{a}_0$.

Полученные графики, где зеленым цветом – истинная зависимость, красным – регрессионная:



7. Остатки $\varepsilon_k = y_k - \hat{y}_k$ (в зависимости от x).

Полученный график:



8. Оценить остаточную дисперсию:

$$\hat{\sigma}^2 = \frac{\sum_{k=1}^n \varepsilon_k^2}{n-3}$$

Остаточная дисперсия=

1.2456