№1

a)

$$\mathrm{cov}(X,Y) = M[XY] - M[X]M[Y]$$
$$M[XY] = M[X]M[Y] + \mathrm{cov}(X,Y)$$

$$D[X] = M[X^2] - (M[X])^2$$

$$\Delta = \xi_2^2 - \xi_1\xi_3$$
$$M\Delta = M\xi_2^2 - M\xi_1\xi_3 =$$
$$= D\xi_2 + (M\xi_2)^2 - M\xi_1 \cdot M\xi_3 - \mathrm{cov}(\xi_1,\xi_3) =$$
$$= A_{22} + m_2^2 - m_1 m_3 - A_{13} = 6 + 4 - 3 - 1 = 6$$

b)

$$\eta_1 = \xi_1 + a\xi_3; \quad \eta_2 = \xi_2 + b\xi_3$$
$$a + b = 1$$

Для гаусових векторів відсутність кореляції між координатами еквівалентна їх незалежності.

$$\mathrm{cov}(\eta_1,\eta_2) = 0$$

$$\mathrm{cov}(\xi_1 + a\xi_3, \xi_2 + b\xi_3) =$$
$$= \mathrm{cov}(\xi_1,\xi_2) + b\,\mathrm{cov}(\xi_1,\xi_3) + a\,\mathrm{cov}(\xi_3,\xi_2) + ab\,\mathrm{cov}(\xi_3,\xi_3) =$$
$$= 2 + b - a + 12ab = 0$$

$$\begin{cases} 2 + b - a + 12ab = 0 \\ a + b = 1 \end{cases} \qquad \begin{cases} a = 1 - b \\ 2 + b - 1 + b + 12b - 12b^2 = 0 \end{cases}$$
$$12b^2 - 14b - 1 = 0$$
$$b = \frac{7 \pm \sqrt{5}1}{12}$$
$$a = \frac{5 \mp \sqrt{5}1}{12}$$

№2

Тут потрібно використати теорему про нормальну кореляцію.

a)

$$M(\xi_1|\xi_2) = M\xi_1 + \frac{\text{cov}(\xi_1, \xi_2)}{\text{cov}(\xi_2, \xi_2)}(\xi_2 - M\xi_2) = \\ = 1 + \frac{-1}{4}(\xi_2 - 0) = -\frac{1}{4}\xi_2 + 1$$

b)

$$D_{\xi_2,(\xi_1,\xi_3)} = \begin{pmatrix} A_{21} & A_{23} \end{pmatrix} = \begin{pmatrix} -1 & 1 \end{pmatrix}$$
$$D_{(\xi_1,\xi_3),(\xi_1,\xi_3)} = \begin{pmatrix} 1 & -2 \\ -2 & 9 \end{pmatrix}$$
$$D^{-1}_{(\xi_1,\xi_3),(\xi_1,\xi_3)} = \frac{1}{5}\begin{pmatrix} 9 & 2 \\ 2 & 1 \end{pmatrix}$$

$$M(\xi_2|(\xi_1, \xi_3)) = M\xi_2 + \frac{1}{5}\begin{pmatrix} -1 & 1 \end{pmatrix}\begin{pmatrix} 9 & 2 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} \xi_1 - M\xi_1 & \xi_3 - M\xi_3 \end{pmatrix} = \\ = 0 + \begin{pmatrix} -1.4 & -0.2 \end{pmatrix}\begin{pmatrix} \xi_1 - 1 \\ \xi_3 - 2 \end{pmatrix}$$

c)

$$D_{(\xi_1,\xi_2),\xi_3} = \begin{pmatrix} A_{13} \\ A_{23} \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$
$$D_{\xi_3,\xi_3} = 9$$

$$M((\xi_1, \xi_2)|\xi_3) = \begin{pmatrix} m_1 \\ m_2 \end{pmatrix} + \frac{1}{9}\begin{pmatrix} -2 \\ 1 \end{pmatrix}(\xi_3 - m_3) = \\ = \begin{pmatrix} 1 - \frac{2}{9}(\xi_3 - 2) \\ \frac{1}{9}(\xi_3 - 2) \end{pmatrix}$$

№3

$$\begin{cases} \eta_1 = 3\xi_1 + 2\xi_2 \\ \eta_2 = 2\xi_1 - 2\xi_3 \end{cases}$$

Відомо, що якщо $\vec{\xi}$ - гаусовий вектор, то і $B\vec{\xi} + \vec{a}$ є таким та має параметри $\mathcal{N}(B\vec{m} + \vec{a}; BAB^T)$.

$$B = \begin{pmatrix} 3 & 2 & 0 \\ 2 & 0 & -2 \end{pmatrix}$$

$$\eta = B\xi$$

$$\vec{m}' = (6 + 2, 4 - 2) = (8, 2)$$

$$A' = BAB^T = \begin{pmatrix} 172 & 110 \\ 110 & 108 \end{pmatrix}$$

$$\varphi_{B\vec{\xi}}(\vec{t}) = \exp\left\{ i\vec{t} \cdot \vec{m}' - \frac{1}{2}\vec{t} \cdot A'\vec{t} \right\}$$

$$\varphi_\eta(\vec{t}) = \exp\left\{ i(8t_1 + 2t_2) - \frac{1}{2}(172t_1^2 + 220t_1t_2 + 108t_2^2) \right\}$$

η_1 та η_2 не є незалежними, бо $\mathrm{cov}(\eta_1, \eta_2) = 110 \neq 0$

№4

Формула оцінки параметрів моделі за матрицею плану Z та значеннями залежної величини X:

$$\hat{\beta} = \left(Z^T Z\right)^{-1} Z^T X$$

Тут беремо перший стовпець Z одиницями, другий - значення x. Матриця X співпадає із y.

[9]:
```
X = np.array([
    83, 62, 76, 77, 89, 74, 48, 78, 76, 51, 73, 89
])
Z = np.array([
    [1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1,  1 ],
    [71, 49, 80, 73, 93, 85, 58, 82, 64, 32, 87, 80]
])
n = X.shape[0]
```

[14]:
```
A = Z @ Z.T
A
```

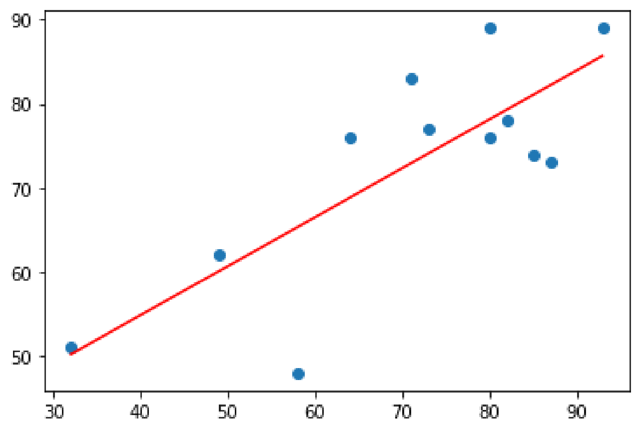
[14]:
```
array([[   12,   854],
       [  854, 64222]])
```

[15]:
```
w = np.linalg.inv(A) @ Z @ X
w
```

[15]:
```
array([31.60946116,  0.58160008])
```

[23]:
```
s = Z[1].argsort()
xs = Z[1][s]
ys = X[s]
plt.scatter(xs, ys)
plt.plot([xs[0], xs[-1]], [
    w[0] + xs[0]*w[1],
    w[0] + xs[-1]*w[1]
], c='r')
```

[23]:
```
[<matplotlib.lines.Line2D at 0x2aa90f660e0>]
```



Отримали рівняння регресії

$$\hat{y} = 31.6 + 0.58x$$

$$\beta_0 = 31.6; \quad \beta_1 = 0.58$$
$$\hat{y} = 31.6 + 0.58x$$

Спрогнозуємо бал підсумкової контрольної студенту, що отримав 84 бали за модульну:

$$y' = 31.6 + 0.58 \cdot 84 = 80.46$$

$$\hat{y} = 80.46$$

Помітимо, що значення 84 всередині навчального діапазону, тобто маємо справу з інтерполяцією. Шукаємо 0.95-довірчий інтервал для прогнозованого значення інтерполяції.

$$t_0 = \frac{(\hat{y}_0 - y_0)\sqrt{n-2}}{\sqrt{\frac{1}{n}S(\hat{\beta}) + \frac{(\bar{x}-x_0)^2}{S_{xx}}}}$$

де n - к-сть спостережень, y_0 - істинне значення значення, $\hat{\beta}$ - обчислена оцінка параметрів,
$S(\beta) = \|X - Z\beta\|^2, \quad S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2$.

$$S(\hat{\beta}) = 27.14$$
$$S_{xx} = 1902.0$$

Інтервал:

$$P(g_1 < t_0 < g_2) = 0.95$$
$$t_0 \sim t(n-2)$$

Центральний інтервал матиме такі межі:

$$g_1 = t_{\frac{1-\gamma}{2};n-2}; g_2 = t_{\frac{1+\gamma}{2};n-2}$$

$$g_1 < \frac{(\hat{y}_0 - y_0)\sqrt{n-2}}{\sqrt{\frac{1}{n}S(\hat{\beta}) + \frac{(\bar{x}-x_0)^2}{S_{xx}}}} < g_2$$
$$\sqrt{\frac{1}{n}S(\hat{\beta}) + \frac{(\bar{x}-x_0)^2}{S_{xx}}} \cdot \frac{g_1}{\sqrt{n-2}} < \hat{y}_0 - y_0 < \sqrt{\frac{1}{n}S(\hat{\beta}) + \frac{(\bar{x}-x_0)^2}{S_{xx}}} \cdot \frac{g_2}{\sqrt{n-2}}$$
$$\hat{y}_0 - \sqrt{\frac{1}{n}S(\hat{\beta}) + \frac{(\bar{x}-x_0)^2}{S_{xx}}} \cdot \frac{g_2}{\sqrt{n-2}} < y_0 < \hat{y}_0 - \sqrt{\frac{1}{n}S(\hat{\beta}) + \frac{(\bar{x}-x_0)^2}{S_{xx}}} \cdot \frac{g_1}{\sqrt{n-2}}$$

Обчислюємо і отримуємо:

$$P(80.46 - 1.07 < y_0 < 80.46 + 1.07) = 0.95$$

$$P(79.39 < y_0 < 81.53) = 0.95$$

№5

a)

```
[2]: X = np.array([
         31, 58, 65, 73, 37, 44, 60, 91, 21, 84
     ])
     Z = np.array([
         [1,  1,  1,  1,  1,  1,  1,  1,  1,  1],
         [4,  9, 10, 14,  4,  7, 12, 22,  1, 17]
     ])
     n = X.shape[0]
```

```
[3]: A = Z @ Z.T
     A
```

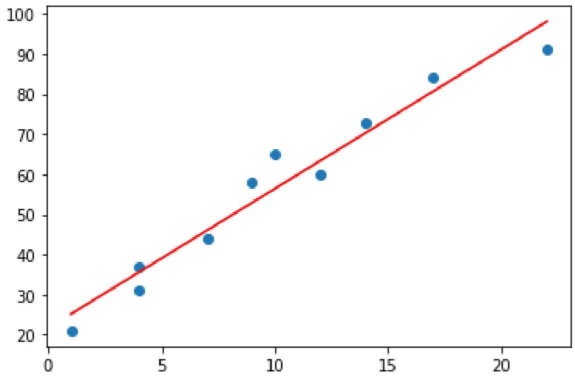
```
[3]: array([[  10,  100],
            [ 100, 1376]])
```

```
[4]: w = np.linalg.inv(A) @ Z @ X
     w
```

```
[4]: array([21.69255319,  3.47074468])
```

```
[7]: s = Z[1].argsort()
     xs = Z[1][s]
     ys = X[s]
     plt.scatter(xs, ys)
     plt.plot([xs[0], xs[-1]], [
         w[0] + xs[0]*w[1],
         w[0] + xs[-1]*w[1]
     ], c='r')
```

```
[7]: [<matplotlib.lines.Line2D at 0x2506d35bf10>]
```



Маємо рівняння прямої

$$\hat{y} = 21.69 + 3.47 \cdot x$$

b)

$$\hat{y} = 21.69 + 3.47 \cdot 10 = 56.4$$

"10 годин" - всередині навчального інтервалу. інтерполяція. Інтервал:

$$\hat{y}_0 - \sqrt{\frac{1}{n}S(\hat{\beta}) + \frac{(\bar{x} - x_0)^2}{S_{xx}}} \cdot \frac{g_2}{\sqrt{n-2}} < y_0 < \hat{y}_0 - \sqrt{\frac{1}{n}S(\hat{\beta}) + \frac{(\bar{x} - x_0)^2}{S_{xx}}} \cdot \frac{g_1}{\sqrt{n-2}}$$

Обчислюємо і маємо:

$$P(55.26 < y_0 < 57.54) = 0.95$$

c)

Гіпотеза

$$y \sim \mathcal{N}(a + bx, \sigma^2)$$

$$H_0 : b = 3$$

$$H_1 : b > 3$$

$$\alpha = 0.01$$

Статистика критерію значущості окремого параметра має вигляд

$$t_0 = \frac{\hat{b} - b}{\sqrt{S(\hat{a}, \hat{b}) \cdot (A^{-1})_{22}}} \sqrt{n-k} \sim t(n-k)$$

(Зауваження: $k = 2$)

$$\hat{a} = 21.69; \quad \hat{b} = 3.47$$

$$S(\hat{a}, \hat{b}) = 14.94; \quad (A^{-1})_{11} = 0.00266$$

Обчислюємо і маємо

$$|t_0| = 1.73 < t_{crit} = t_{1-\frac{0.01}{2};n-2} = 3.35$$

Отже, нульову гіпотезу не можна відхиляти. Можна зробити висновок, що в середньому кожна додаткова година підготовки до контрольної роботи підвищує бал на менше ніж три пункти.

Тобто, різниця $\hat{b} - b = 0.47$ недостатня для статистичної значимості результатів.

обчислення:

```
[33]: S = np.square(X - w @ Z).sum()
      S
```

```
[33]: 223.07819148936179
```

```
[34]: xbar = X.mean()
      xbar
```

```
[34]: 56.4
```

```
[36]: Ainv_ii = np.linalg.inv(A)[1,1]
      Ainv_ii
```

```
[36]: 0.002659574468085106
```

```
[37]: t_0 = np.sqrt(n-2) * (w[1] - 3)/np.sqrt(S * Ainv_ii)
      t_0
```

```
[37]: 1.728605059862597
```

```
[39]: t_crit = sp.stats.t(n-2).ppf(1-0.01/2)
      t_crit
```

```
[39]: 3.3553873313333953
```

d)

Будуємо 0.95-довірчий інтервал для b в термінах попереднього пункту.

Маємо таку центральну статистику для b:

$$t_0 = \frac{\hat{b} - b}{\sqrt{S(\hat{a}, \hat{b}) \cdot (A^{-1})_{11}}} \sqrt{n-k-1} \sim t(n-k-1)$$

$$P(g_1 < \frac{\hat{b} - b}{\sqrt{S(\hat{a}, \hat{b}) \cdot (A^{-1})_{11}}} \sqrt{n-k-1} < g_2) = \gamma$$

$$\frac{\sqrt{S(\hat{a}, \hat{b}) \cdot (A^{-1})_{11}}}{\sqrt{n-k-1}} g_1 < (\hat{b} - b) < \frac{\sqrt{S(\hat{a}, \hat{b}) \cdot (A^{-1})_{11}}}{\sqrt{n-k-1}} g_2$$

$$\hat{b} - \frac{\sqrt{S(\hat{a}, \hat{b}) \cdot (A^{-1})_{11}}}{\sqrt{n-k-1}} \cdot t_{\frac{1+\gamma}{2};n-2-1} < b < \hat{b} + \frac{\sqrt{S(\hat{a}, \hat{b}) \cdot (A^{-1})_{11}}}{\sqrt{n-k-1}} \cdot t_{\frac{1+\gamma}{2};n-2-1}$$

Обчислюємо і маємо

$$P(2.84 < b < 4.1) = 0.95$$

```
[47]: d = np.sqrt( S * Ainv_ii / (n-2) ) * sp.stats.t(n-2).ppf((1+0.95)/2)
      d
```

```
[47]: 0.6279856549036167
```

```
[48]: w[1] - d, w[1] + d
```

```
[48]: (2.842759025947448, 4.098730335754681)
```

№6

$$y = ax^b$$

$$\ln y = \ln a + b \ln x$$

$$y' = a' + b'x'$$

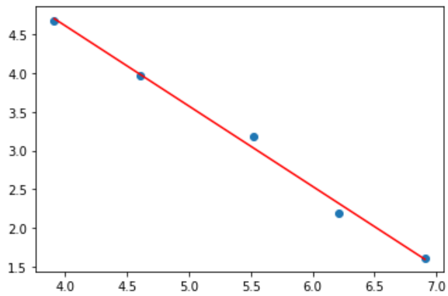
Маємо модель лінійної регресії. Обчислюємо та отримуємо

$$\hat{a}' = 8.777; \quad \hat{b}' = -1.04$$

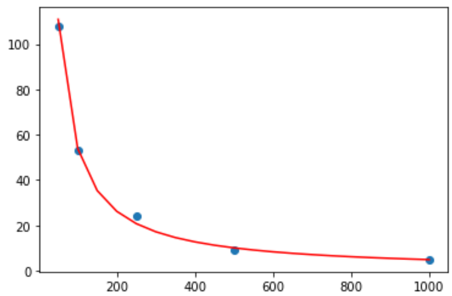
$$\hat{a} = 6481; \quad \hat{b} = -1.04$$

$$\hat{y} = 6481 \cdot x^{-1.04}$$

```
[21]: w = w1.copy()
      w[0] = np.exp(w[0])
      w

[21]: array([ 6.48077044e+03, -1.03994020e+00])

...

[23]: [<matplotlib.lines.Line2D at 0x21292c9afe0>]
```




```
...

[24]: [<matplotlib.lines.Line2D at 0x21292d02bc0>]
```



Оцінка вартості виробництва партії з 300 деталей:

$$y_{300} = 6481 \cdot 300^{-1.04} = 17.2$$

№7

Будуємо регресію із $n = 6, k = 3$.

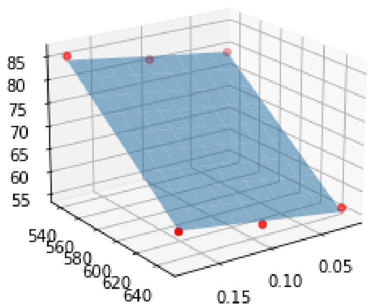
```
[7]: X = np.array([
         [0.02, 0.02, 0.10, 0.10, 0.18, 0.18],
         [538, 649, 538, 649, 538, 649]
     ])
     Y = np.array([
         78.9, 56.2, 80.9, 57.4, 85.3, 60.7
     ])
     n = Y.shape[0]
```

```
[8]: Z = np.vstack([np.ones(n), X])
     A = Z @ Z.T
     B = np.linalg.inv(A)
```

```
[9]: w = B @ Z @ Y
     w
```

```
[9]: array([192.67933559,  34.0625    ,  -0.21261261])
```

```
[25]: <mpl_toolkits.mplot3d.art3d.Poly3DCollection at 0x204e837bd60>
```



$$\hat{y} = 192.68 + 34.06x_1 - 0.2126x_2$$

$$\hat{y}_0 = 192.68 + 34.06 \cdot 0.14 - 0.2126 \cdot 593 = 71$$

№8

Доповнюємо Z стовпцями (x_i^2) та обчислюємо регресію із $k = 3$.

```
[2]: X = np.array(
         [3, 5, 7, 9, 10]
     )
     Y = np.array([
         14.5, 35.7, 62.8, 100, 121
     ])
     n = Y.shape[0]
```

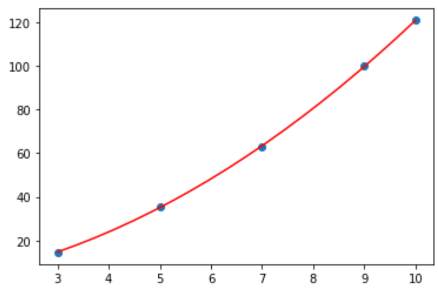
```
[3]: Z = np.vstack([np.ones(n), X, X*X])
     A = Z @ Z.T
     B = np.linalg.inv(A)
```

```
[4]: w = B @ Z @ Y
     w
```

```
[4]: array([-0.79666931,  2.16582078,  1.00130849])
```

```
[9]: s = X.argsort()
     plt.scatter(X, Y)

     xs = np.linspace(3,10,20)
     ys = w[0] + w[1]*xs + w[2]*xs*xs
     plt.plot(xs,ys, c='r')
```

```
[9]: [<matplotlib.lines.Line2D at 0x2580d2f6ef0>]
```



$$\hat{y} = -0.8 + 2.17x + 1 \cdot x^2$$

Тобто прискорення $a = 1$.

Вже вивели такий вигляд довірчого інтервалу:

$$\hat{a} - \frac{\sqrt{S(\hat{\beta}) \cdot (A^{-1})_{33}}}{\sqrt{n-k}} \cdot t_{\frac{1+\gamma}{2};n-3} < a < \hat{a} + \frac{\sqrt{S(\hat{\beta}) \cdot (A^{-1})_{33}}}{\sqrt{n-k}} \cdot t_{\frac{1+\gamma}{2};n-3}$$

Обчислюємо і маємо

$$P(0.744 < a < 1.259) = 0.95$$

Також маємо наступну статистику для σ^2:

$$\frac{S(\hat{\beta})}{\sigma^2} \sim \chi^2(n-k)$$

$$g_1 = \chi^2_{\frac{1-\gamma}{2};n-k}; \quad g_2 = \chi^2_{\frac{1+\gamma}{2};n-k}$$

$$P\left(g_1 < \frac{S(\hat{\beta})}{\sigma^2} < g_2\right) = \gamma$$

$$P\left(\frac{S(\hat{\beta})}{g_2} < \sigma^2 < \frac{S(\hat{\beta})}{g_1}\right) = \gamma$$

Обчислюємо і маємо

$$P(0.12 < \sigma^2 < 17.4) = 0.95$$

Обчислення:

```
[10]: S = np.square(Y - w @ Z).sum()
      S

[10]: 0.8822283901665476

[12]: Ainv_ii = np.linalg.inv(A)[2,2]
      Ainv_ii

[12]: 0.008128469468675568

[13]: t_crit = sp.stats.t(n-3).ppf(0.5 + 0.95/2)
      t_crit

[13]: 4.302652729911275

[14]: d = np.sqrt( S * Ainv_ii / (n-3) ) * t_crit
      d

[14]: 0.25764172693783616

[15]: w[2] - d, w[2] + d

[15]: (0.7436667583913081, 1.2589502122669805)

[16]: g1 = sp.stats.chi2(n-3).ppf(0.5-0.95/2)
      g2 = sp.stats.chi2(n-3).ppf(0.5+0.95/2)
      S/g2, S/g1

[16]: (0.11957945510833423, 17.423080045357487)
```

№9

Доповнюємо Z стовпцями (x_i^2) та обчислюємо регресію із $k = 3$.

```
[3]: X = np.array(
         [20, 16, 10, 11, 14]
     )
     Y = np.array([
         22, 41, 120, 89, 56
     ])
     n = Y.shape[0]
```

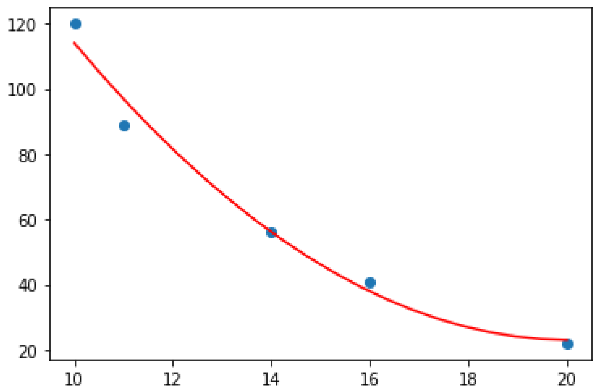
```
[5]: Z = np.vstack([np.ones(n), X, X*X])
     A = Z @ Z.T
     B = np.linalg.inv(A)
```

```
[6]: w = B @ Z @ Y
     w
```

```
[6]: array([384.39253883, -35.99751719,   0.89642221])
```

```
[8]: plt.scatter(X, Y)

     xs = np.linspace(10,20,20)
     ys = w[0] + w[1] * xs + w[2] * xs * xs
     plt.plot(xs, ys, c='r')
```

```
[8]: [<matplotlib.lines.Line2D at 0x20b000bab60>]
```



$$\hat{y} = 384.4 - 36x + 0.895x^2$$

Перевіримо гіпотезу ($\alpha = 0.05$)

$$H_0 : \beta_2 = 0$$

$$H_1 : \beta_2 \neq 0$$

Запишемо вже знайому статистику

$$t_0 = \frac{\hat{\beta}_2 - 0}{\sqrt{S(\beta) \cdot (A^{-1})_{33}}} \sqrt{n-k} \sim t(n-k); \quad t_{crit} = t_{1-\frac{\alpha}{2};n-k}$$

Обчислюємо

$$t_0 = 2.94 < t_{crit} = 4.3$$

Отже, не можна відхилити нульову гіпотезу. Залежність, скоріш за все, не має сенсу описувати параболічно.

Обчислення:

```
[7]: S = np.square(Y - w @ Z).sum()
     S
```

```
[7]: 108.03934300993143
```

```
[8]: Ainv_ii = np.linalg.inv(A)[2,2]
     Ainv_ii
```

```
[8]: 0.0017188693659281181
```

```
[11]: t_crit = sp.stats.t(n-3).ppf(1 - 0.05/2)
      t_crit
```

```
[11]: 4.302652729911275
```

```
[10]: t_0 = np.sqrt(n-3) * w[2] / np.sqrt(Ainv_ii * S)
      t_0
```

```
[10]: 2.941813933823414
```

№10

Оцінкою максимальної правдоподібності для коефіцієнта кореляції ρ є

$$r = \frac{\sum_{i=1}^{n} y_i(x_i - \bar{x})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_1 - \bar{y})^2}} = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}}$$

Обчислюємо і отримуємо

$$r = 0.936$$

Значення близьке до 1, а значить регресія є значущою - є сенс вважати, що існує лінійна залежність. А $r^2 = 0.876$ вказує на те, що лінійна модель пояснює 87% варіативності.

Статистикою критерію гіпотези $(H_0 : \rho = 0)$ проти $(H_1 : \rho \neq 0)$ є

$$z = \frac{\sqrt{n-3}}{2} \ln \frac{1+r}{1-r}$$

$$z_{crit} : \quad |z| > u_{1-\frac{\alpha}{2}}$$

$$z = 4.5 > z_{crit} = 2.576$$

Отже, можна відхилити нульову гіпотезу про некорельованість цих наборів даних.
Тобто, можна прийняти, що дані корельовані між собою.