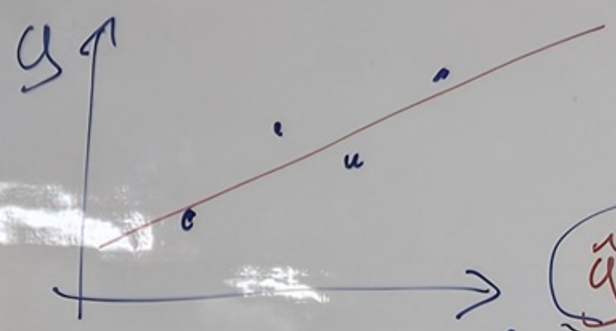
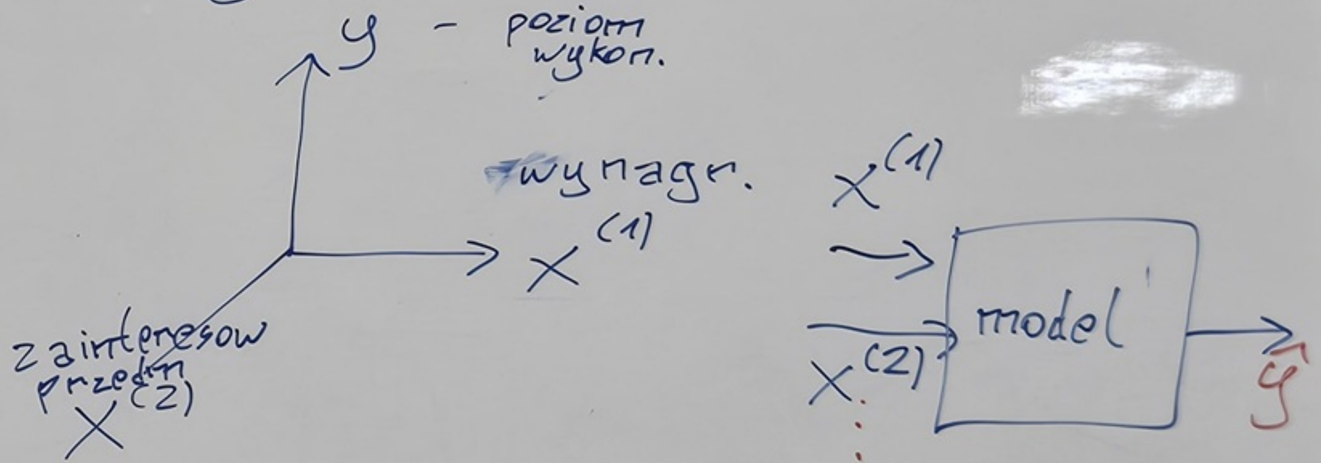
Dopasowanie modelu do danych



$\hat{y} = ax$

$\hat{y} = ax + b$

= uczenie maszynowe



$$\hat{y} = a_1 x^{(1)} + a_2 x^{(2)} + \dots + a_K x^{(K)} =$$

$$= \sum_{j=1}^{K} a_j x^{(j)} = \tilde{\theta}^T x = x^T \tilde{\theta}$$

$$\tilde{\theta} = \begin{bmatrix} a_1 \\ \vdots \\ a_K \end{bmatrix}$$

$x_i^{(j)}$ → nr. cechy

i - nr. obserw.

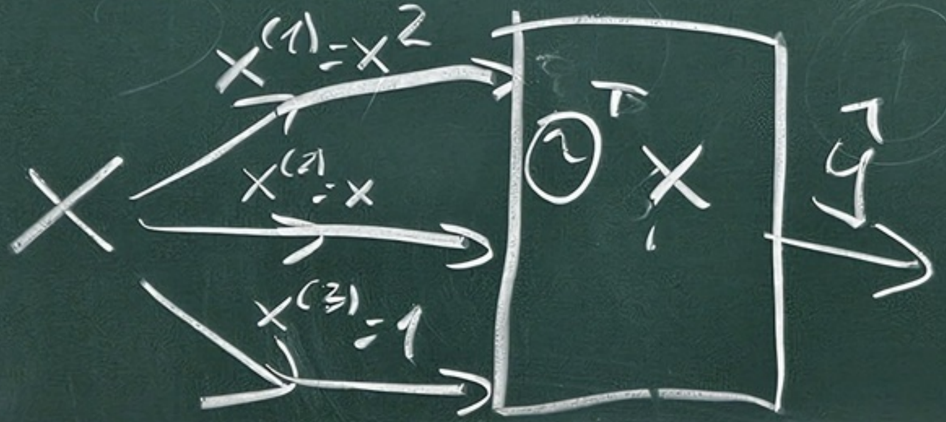
$$x = \begin{bmatrix} x^{(1)} \\ \vdots \\ x^{(K)} \end{bmatrix} \qquad x_i \in \mathbb{R}^K$$

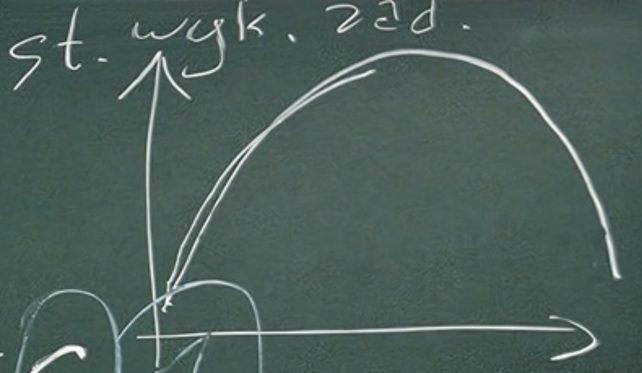
$$\hat{y} = a_1 x^{(1)} + a_2 x^{(2)}$$

$$\hat{y} = a x + b \cdot 1$$









$$\hat{y} = ax^2 + bx + c$$

$$\hat{y} = a_1 x^{(1)} + a_2 x^{(2)} + a_3 x^{(3)}$$

x	1	0	2	2
y	-1	2	0	1

$x^{(1)}$	1	0	2	2
$x^{(2)}$	1	1	1	1
y	-1	2	0	1

$$\hat{Y} = \Theta^T X = X^T \Theta$$

$$\begin{bmatrix} x_1^{(1)} & x_2^{(1)} & \cdots & x_N^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \cdots & x_N^{(2)} \\ \vdots & \vdots & & \\ x_1^{(k)} & x_2^{(k)} & \cdots & x_N^{(k)} \end{bmatrix} = X_{K \times N}$$

$$(A \pm B)^T = A^T \pm B^T$$

$$(A \cdot B)^T = B^T A^T$$

$N \times K \quad K \times L$

$K \times N \quad L \times K$

$$(A^T)^T = A$$

$$\begin{bmatrix} y_1 & y_2 & \cdots & y_N \end{bmatrix} = Y_{1 \times N}$$

$$\begin{bmatrix} \hat{y}_1 & \hat{y}_2 & \cdots & \hat{y}_N \end{bmatrix} = \hat{Y}_{1 \times N}$$

$$\begin{bmatrix} \Theta^T x_1 & \Theta^T x_2 & \cdots & \Theta^T x_N \end{bmatrix}$$

$$\Theta^T \begin{bmatrix} x_1 & x_2 & \cdots & x_N \end{bmatrix}$$

$$A A^{-1} = I$$

$$A^{-1} A = I$$

$$(A^{-1})^T = (A^T)^{-1}$$

$$(A \cdot B)^{-1} = B^{-1} A^{-1}$$

$$Y = \tilde{\Theta}^T X \quad / \cdot X^{-1}$$

$K \times N$

$$Y \cdot X^{-1} = \tilde{\Theta}^T$$

$$Y X^T = \tilde{\Theta}^T \underbrace{X X^T}_{K \times K} \quad / \cdot (X X^T)^{-1}$$

$$Y X^T (X X^T)^{-1} = \tilde{\Theta}^T$$

$$\boxed{\tilde{\Theta} = (X X^T)^{-1} X Y^T}$$