Uczenie maszynowe

uczenie nadzorowane

dane + etykiety

obrazki: $\boxed{0}$ $\boxed{1}$ $\boxed{2}$ 28x28

etykiety: 0 1 2

$f: \mathbb{R}^{28\times 28} \to \{0, 1, \dots, 9\}$

regresja:

wzrost matki, wzrost ojca, <u>wzrost syna</u>

$f: \mathbb{R}^2 \to \mathbb{R}$

cel: znaleźć

$f: X \to Y$

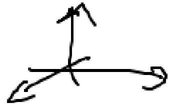
która będzie modelować jedną ze zmiennych

uczenie nienadzorowane

· klasteryzacja

$\mathbb{R}^{20}$

· redukcja wymiarów



(reinforcement learning)

uczenie przez wzmacnianie

np. szachy, go

MCTS

Regresja liniowa (nie będziemy zagłębiać się w teorię statystyczną ...)

$f : \mathbb{R}^m \to \mathbb{R}$

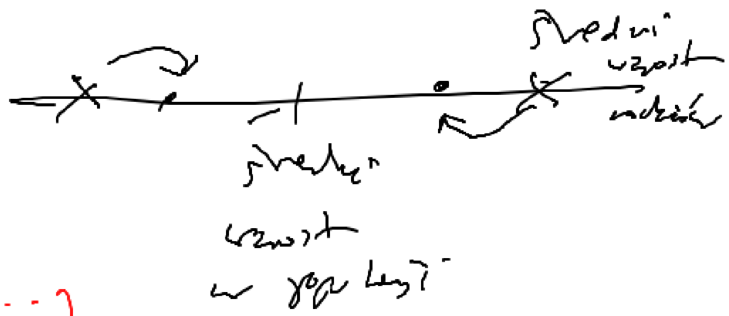
model: $f_\vartheta(x_i) = [1, x_i] \cdot \vartheta = \vartheta_0 + \sum_{j=1}^{m} x_{ij} \vartheta_j$, $i = 1, \dots, n$, $x_i \in \mathbb{R}^m$

np.

$$X = \begin{bmatrix} 163 & 175 \\ 167 & 190 \\ 173 & 172 \\ 150 & 161 \\ 176 & 181 \end{bmatrix} \qquad Y = \begin{bmatrix} 177 \\ 180 \\ 170 \\ 158 \\ 183 \end{bmatrix}$$

wzrost matki, wzrost ojca, wzrost syna

sir Galton



$$\tilde{X} = \begin{bmatrix} 1 & 163 & 175 \\ 1 & 167 & \vdots \\ 1 & \vdots & \vdots \\ 1 & \vdots & \vdots \\ 1 & \vdots & 181 \end{bmatrix}$$

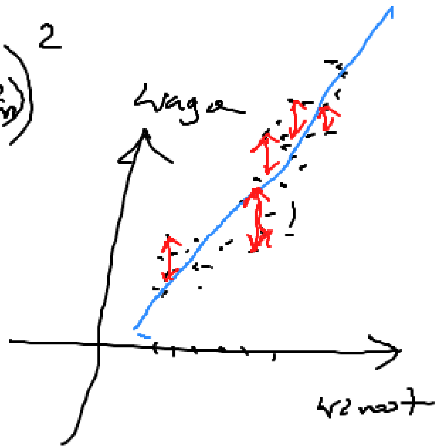
$$\tilde{X}^T \cdot \tilde{X} = \begin{bmatrix} \cdots \end{bmatrix} \cdot \begin{bmatrix} \vdots \end{bmatrix} = \begin{bmatrix} \cdots \end{bmatrix}$$

~~Staramy się z~~ Minimalizujemy następującą funkcję kosztu:

$$\underset{\substack{\uparrow \\ \text{prawdziwa} \\ \text{wartość}}}{y},\quad \underset{\substack{\uparrow \\ \text{modelowana} \\ \text{wartość}}}{\hat{y}}$$

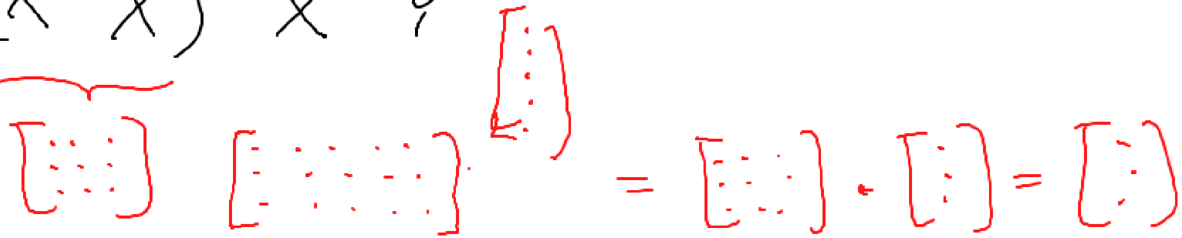
$$L(y, \hat{y}) = \| y - \hat{y} \|_2^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 =$$

$$= \sum_{i=1}^{n} \left(y_i - (\vartheta_0 + x_i(\vartheta_1, \dots, \vartheta_m))\right)^2$$



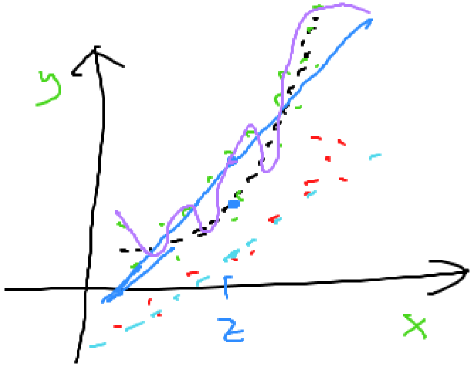
Przy założeniu, że $\operatorname{rank}\Big(\overbrace{\begin{bmatrix} 1 & \\ 1 & \\ \vdots & X \\ 1 & \end{bmatrix}}^{\tilde{X}}\Big) = m+1$, istnieje dokładnie jeden $\hat{\vartheta} \in \mathbb{R}^{m+1}$ minimalizujący L,

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



Numerycznie lepiej jest użyć innych metod.

Zakładając, że
$y_i = g(x_i) + e_i$, $e_i \sim N(0,\sigma^2)$, iid , $i = 1, \dots, n$



$\hat{f}_\vartheta(z) = \vartheta_0 + z \cdot (\vartheta_1, \dots, \vartheta_m)$ — model

błąd dla $\hat{f}_\vartheta(z) = \sigma^2 + Bias^2(\hat{f}_\vartheta(z)) + Var(\hat{f}_\vartheta(z))$

Decyzje: mniej skomplikowany model : duże małe

baza $1, x, x^2, \dots, x^{10}$
liniowa

bardziej -||- ——— małe duże

Regularyzacja

Ridge: regresja liniowa z funkcją kosztu

$$L(y,\hat{y}) = \|y-\hat{y}\|_2^2 + \alpha \|(\vartheta_1, \dots, \vartheta_m)\|_2^2$$

współczynnik ≥ 0 regularyzacji — ustalany z góry

$\alpha = 0$: zwykła regresja

Metody iteracyjne:

1) $\vartheta = $ (cokolwiek)

powtarzamy:

2) obliczamy $\left(\frac{\partial L}{\partial \vartheta}(y,\hat{y})\right) = \nabla_\vartheta L$

3) $\vartheta := \vartheta - \underbrace{c}_{\text{mała stała}} \cdot \nabla_\vartheta L$

numpy

array ⟵ (wielowymiarowa) macierz złożona z elementów tego samego typu

np. wektor $(1, 3, 5.5) \in \mathbb{R}^3$ shape = (3,)

$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \in \mathbb{R}^{2\times 4}$ shape = (2, 4)

$((\mathbb{R}^5)^4)^2$ shape = (2, 4, 5)

a = np.array([1, 2, 3, 4, 5])

a[1:3] (= [2, 3])

a[1:3] = [44, 47] ⟿ teraz a = [1, 44, 47, 4, 5] operacje arytmetyczne na obiektach array:

$$[2, 3] \cdot [4, 5] \rightarrow [8, 15]$$

broadcast: pozwala wykonywać operacje ~~[illegible]~~ w przypadku, gdy wymiary nie są takie same

a: (15, 7, 1)

b: (1, 14)

a: (4, 3)

b: (4, 1)

a*b