Zad.2 $X = [x_1, \dots, x_m]^T \quad Y = [y_1, \dots, y_m]^T \quad W = \begin{bmatrix} w^1 & & 0 \\ & w^2 & \\ 0 & & w^m \end{bmatrix} \quad \theta = [\theta_1 \dots \theta_m]^T$

$$J(\theta) = \frac{1}{2} \sum_i w^i \cdot (x_i \theta - y_i)^2 = Tr\left(W (X\theta - Y)^T (X\theta - Y)\right)$$

$$\frac{\partial J}{\partial \theta} = \frac{\partial}{\partial \theta} \left(\frac{1}{2} Tr \left(W (X\theta - Y)^T (X\theta - Y) \right) \right) = \frac{\partial}{\partial \theta} \left(\frac{1}{2} (X\theta - Y)^T W (X\theta - Y) \right) =$$

$$= \frac{\partial}{\partial \theta} \left(\frac{1}{2} (\theta^T X^T - Y^T) W (X\theta - Y) \right) = \frac{\partial}{\partial \theta} \frac{1}{2} \left(\theta^T X^T W X \theta - \theta^T X^T W Y - Y^T W X \theta + Y^T W Y \right) =$$

$$= \frac{1}{2} \left(2\theta X^T W X - 2 \cancel{\theta^T} X^T W Y \right) \cancel{= 0}$$

$$\theta_* = (X^T W Y)(X^T W X)^{-1}$$

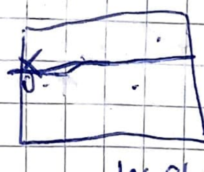
α - smooth parameter

$\frac{\text{ALL PTS}}{\text{LOCAL PTS}}$

1. Get subsets of data $\rightarrow$ KNN
2. Local points $\rightarrow$ fit by polynomial 0, 1, 2 degree
3. Weight: $W(x) = (1 - |d|^3)^3$ closer the point to the focal $\rightarrow$ more weight / influence it has

d - odległość ovy punktu od przewidywanego w danej

$\leftarrow$ rodzaj wdcit jedy



ALBO PARABOLA

W ODWIIE NAYWIĘKSZE WAGI MAJĄ NAJBLIŻSZE FOCU

Zad.3

3 ważenia

Podejrzani: 9 10 11 12

1. ⌊1 2 3 4⌋ ⌊5 6 7 8⌋ = → ⌊9 10⌋ ⌊12⌋ ....

Równowagi ✓

~~BSU 5. cięższa~~

~~2. ⌊1 2 3 4⌋ ⌊9 10 11 12⌋ ← po tym ważeniu wiemy w której czwórce jest zła kulka i czy jest cięższa czy lżejsza~~

Jednym ważeniem z wiedzą

~~Np.~~

2. ① ⑤ ⑨ ② ③ ⑥

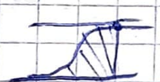
→ to samo nierówność podejrzani ① ⑥

→ Równość ④ ⑦ ⑧ podejrzane ④ ⑦ ⑧

→ INNA nierówność ② ③ < ⑤ ④⑦

↓

ważymy dwie z tej samej strony

Zad. 4 

Dystrybuanta: $F_{Min(x,y)}(z) = P(Min(x,y) \le z) = 1 - P(x > z \wedge y > z) =$

$= 1 - P(x > z)P(y > z) = 1 - (1-z)(1-z) = 2z - z^2$

F. Gęstości

$F(z) = \int_0^z f(t)\,dt \qquad f(z) = \frac{dF}{dz}(z) = 2 - 2z$

$EZ = \int_0^1 z f(z)\,dz = \int_0^1 (2+2z) z\,dz = 2\left[\frac{z^2}{2} - \frac{z^3}{2}\right]_0^1 = \frac{1}{3}$

5. Funkcja złożona:

Korzystaj z reguły łańcucha

$$L_j = -\log(S(x)_j)$$

$$S(x)_j = \frac{e^{x_j}}{\sum_{k=0}^{N} e^{x_k}}$$

$$\frac{\partial L_j}{\partial x_i} = \frac{\partial L_j}{\partial S(x)_j} \cdot \frac{\partial S(x)_j}{\partial x_i}$$

$$\left(\frac{f(x)}{g(x)}\right)' = \frac{f(x)'g(x) - f(x)'g(x)}{(g(x))^2}$$

$$M = \sum_{k=0}^{N} e^{x_k}$$

$$\frac{\partial L_j}{\partial S(x)_j} = -\frac{1}{\partial S(x)_j}$$

$$\frac{\partial S(x)_j}{\partial x_i} = \frac{\partial}{\partial x_i}\left(\frac{e^{x_j}}{\sum_{k=0}^{N} e^{x_k}}\right) = \left(\frac{e^{x_i} M \cdot \delta_{i,j} - e^{x_i} e^{x_i}}{M^2}\right) =$$

$$= \begin{cases} S(x)_i(1 - S(x)_j) & i = j \\ -S(x)_i S(x)_j & i \neq j \end{cases}$$

$$\frac{\partial L_j}{\partial x_i} = -\frac{1}{S(x)_j}\left[S(x)_i(\overset{\delta_{i,j}}{1} - S(x)_j)\right] = \begin{cases} S(x)_i & i \neq j \\ \frac{S(x)_i(1 - S(x)_j)}{S(x)_j} & i = j \end{cases}$$

$$\nabla = [\, S(x)_0, \ldots, (1 - S(x)_j), \ldots, S(x)_n \,]$$

zad. 6

$$\varepsilon_m(x) = h_m(x) - y \qquad E_{AV} = \frac{1}{M}\sum_{m=1}^{M} E_x[\varepsilon_m(x)^2], \quad E_x[\varepsilon_m(x)] = 0, \quad \overset{**}{E_x[\varepsilon_m(x)\varepsilon_l(x)] = 0} \;\; m \neq l$$

$$h_m(x) = \frac{1}{M}\sum_{m=1}^{M} h_m(x)$$

$$E_x[e_m^2(x)] = E_x[(h_m(x) - y)^2] =$$

bo niezależne **

$$= E_x\left[\left(\frac{1}{M}\sum_{m=1}^{M} h_m(x) - y\right)^2\right] = E_x\left[\left(\frac{1}{M}\sum_{m=1}^{M} e_m(x)\right)^2\right] = E_x\left[\frac{1}{M^2}\sum e_m^2(x)\right] =$$

E[ ] liniowy

$$= \frac{1}{M^2}\sum_{m=1}^{M} E_x[e_m^2(x)] = \frac{1}{M} E_{AV}$$

zad.7

$$x_1 \cdot x_2 \cdots x_n$$
$$\Downarrow$$
$$\log(x_1 \cdot x_2 \cdots x_n)$$
$$\Downarrow$$
$$\log(x_1) + \log(x_2) + \ldots + \log(x_n)$$
$$\Downarrow$$

$$LSE(x_1, \ldots, x_n) = \log\left(\sum_{n=1}^{N} \exp(x_n)\right)$$

• Dla x_n dużych lub małych wartości e^{x_n} może nie być numerycznie poprawne

$$\log \sum_{i=1}^{N} x_i = \log \sum_{i=1}^{N} e^{\log x_i} = \log\left(e^{M} \cdot \sum_{i=1}^{N} e^{\log x_i - M}\right) = M + \underbrace{\sum_{i=1}^{N} e^{\log x_i - M}}_{[1, N]}$$

$$M = \max\{\log x_1, \ldots, \log x_n\}$$

$$0 < e^{\log x_i - M} \leq 1$$

Podobnie w softmax, problem może być gdy x_i duże:

$$M = \max\{x_1, x_2 \ldots x_n\}$$

$$\frac{\exp(x_i)}{\sum_{j=1}^{N} \exp(x_j)} \overset{*}{=} \frac{\exp(x_i - M)}{\sum_{j=1}^{N} \exp(x_j - M)}$$

$$* \ \frac{\exp(x+c)}{\sum_i \exp(x_i+c)} = \frac{e^x e^c}{\sum_i e^{x_i} e^c} = \frac{e^c}{e^c} \cdot \frac{e^x}{\sum_i e^{x_i}}$$