ML

2 Теоретическое ДЗ

№1

① Док-ть, что $\exists a \in R^n, x \in R^n$, то $\frac{\partial(a^T x)}{\partial x} = a$

$a = (a_1, a_2, \ldots, a_n)^T$

$x = (x_1, x_2, \ldots, x_n)^T$

$a^T x = \sum_{i=1}^{n} a_i x_i$

$$\frac{\partial(a^T x)}{\partial x} = \left(\frac{\partial(a^T x)}{\partial x_1}, \frac{\partial(a^T x)}{\partial x_2}, \ldots, \frac{\partial(a^T x)}{\partial x_n} \right)^T$$

а т.к. $\frac{\partial(a^T x)}{\partial x_i} = a_i \Rightarrow \frac{\partial(a^T x)}{\partial x} = (a_1, a_2, \ldots, a_n)^T = a$

② Если $A \in R^{m \times n}$, $x \in R^n$, то $\frac{\partial(Ax)}{\partial x} = A$

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & & & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \quad x = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix}$$

$$g(x) = Ax = \underbrace{\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \cdot & \cdot & \cdot & - \\ a_{m1} & \dots & & a_{mn} \end{pmatrix}}_{m \times n} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}}_{n} = \underbrace{\begin{pmatrix} g_1 \\ g_2 \\ \dots \\ g_m \end{pmatrix}}_{m \times 1}$$

$$\frac{\partial (Ax)}{\partial x} = \frac{\partial (g(x))}{\partial x} = \left(\frac{\partial g}{\partial x_1}, \frac{\partial g}{\partial x_2}, \dots \frac{\partial g}{\partial x_n} \right) =$$

$$= \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_n} \\ \cdot & \cdot & - \\ \frac{\partial g_m}{\partial x_1} & \dots & \frac{\partial g_m}{\partial x_n} \end{pmatrix} \textcircled{=}$$

т.к. $g_j = \sum\limits_{i=1}^{n} a_{ji} x_i, \quad \forall j = \overline{1..m}$

$$\textcircled{=} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \cdot & - & \cdot & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = A$$

③ $A \in R^{n \times n}$, $x \in R^n$

док-ть, что $\dfrac{\partial (x^T A x)}{\partial x} = (A + A^T) x$

$$x^T A = (\underbrace{a_{11} x_1 + \ldots + a_{1n} x_n}_{g_1}, \ldots, \underbrace{a_{n1} x_1 + \ldots + a_{nn} x_n}_{g_n}) =$$

$= (g_1, g_2, \ldots, g_n)$, где $g_j = \sum_{i=1}^{n} a_{ji} x_i$

$$x^T A x = (g_1, g_2, \ldots, g_n) \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} = \sum_{k=1}^{n} g_k x_k$$

$$\frac{\partial (x^T A x)}{\partial x} = \left(\frac{\partial (x^T A x)}{\partial x_1}, \ldots, \frac{\partial (x^T A x)}{\partial x_n} \right)^T$$

$$\frac{\partial (x^T A x)}{\partial x_j} = \frac{\partial \sum_{k=1}^{n} \left(\sum_{i=1}^{n} a_{ki} x_i x_k \right)}{\partial x_j} =$$

$$= \underset{(\forall i = k = j)}{2 x_j a_{jj}} + \sum_{i = j \neq k}^{n} a_{kj} x_k + \sum_{k = j \neq i}^{n} a_{ji} x_i =$$

$$= 2 x_j a_{jj} + \sum_{i \neq j}^{n} x_i (a_{ij} + a_{ji}) = \sum_{i=1}^{n} x_i (a_{ij} + a_{ji})$$

ДАТА / DATE

ПН ВТ СР ЧТ ПТ СБ ВС
MON TUE WED THU FRI SAT SUN

Из этого следует, что если $A = A^T$, то есть $a_{ij} = a_{ji}$, то

$$\frac{\partial (x^T A x)}{\partial x} = 2Ax$$

(4) Если $x \in R^n$, то $\frac{\partial \|x\|^2}{\partial x} = 2x$

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} \qquad \|x\|^2 = x_1^2 + x_2^2 + \dots + x_n^2$$

$$\frac{\partial \|x\|^2}{\partial x_i} = 2x_i$$

Тогда: $$\frac{\partial \|x\|^2}{\partial x} = \begin{pmatrix} \frac{\partial \|x\|^2}{\partial x_1} \\ \dots \\ \frac{\partial \|x\|^2}{\partial x_n} \end{pmatrix} = \begin{pmatrix} 2x_1 \\ 2x_2 \\ \dots \\ 2x_n \end{pmatrix} = 2x$$

⑤ g - скалярная, $g(x)$ - применение g к каждой компоненте $x \in \mathbb{R}^n$, то

$$\frac{\partial g(x)}{\partial x} = \mathrm{diag}(g'(x))$$

Док-во:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} \qquad g(x) = \begin{pmatrix} g(x_1) \\ g(x_2) \\ \dots \\ g(x_n) \end{pmatrix}$$

$$\frac{\partial g(x)}{\partial x} = \begin{pmatrix} \frac{\partial g(x_1)}{\partial x_1} & \frac{\partial g(x_1)}{\partial x_2} & \dots & \frac{\partial g(x_1)}{\partial x_n} \\ \dots & \dots & \dots & \dots \\ \frac{\partial g(x_n)}{\partial x_1} & \frac{\partial (g(x_n))}{\partial x_2} & \dots & \frac{\partial (g(x_n))}{\partial x_n} \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{\partial g(x_1)}{\partial (x_1)} & 0 & 0 & 0 \dots 0 \\ 0 & \ddots & & \\ 0 & & \ddots & \\ 0 & & & \frac{\partial g(x_n)}{\partial x_n} \end{pmatrix} = \mathrm{diag}(g'(x)) =$$

ДАТА / DATE

ПН ВТ СР ЧТ ПТ СБ ВС
MON TUE WED THU FRI SAT SUN

6) если $h: R^n \to R^m$, $g: R^m \to R^p$ $x \in R^n$, то

$$\frac{\partial g(h(x))}{\partial x} = \frac{\partial g(h(x))}{\partial h} \frac{\partial h(x)}{\partial x}$$

Док-во:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{pmatrix} \quad h(x) = \begin{pmatrix} h_1(x) \\ h_2(x) \\ \cdots \\ h_m(x) \end{pmatrix} \quad g(x) = \begin{pmatrix} g_1(x) \\ g_2(x) \\ \cdots \\ g_p(x) \end{pmatrix}$$

$$\frac{\partial g(h(x))}{\partial x} = \begin{pmatrix} \frac{\partial g_1(h(x))}{\partial x_1} & \frac{\partial g_1(h(x))}{\partial x_2} & \cdots & \frac{\partial g_1(h(x))}{\partial x_n} \\ \frac{\partial g_2(h(x))}{\partial x_1} & \frac{\partial g_2(h(x))}{\partial x_2} & \cdots & \frac{\partial g_2(h(x))}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial g_p(h(x))}{\partial x_1} & \frac{\partial g_p(h(x))}{\partial x_2} & \cdots & \frac{\partial g_p(h(x))}{\partial x_n} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{i=1}^{m} \frac{\partial g_1(h_i(x))}{\partial h_i(x)} \frac{\partial h_i(x)}{\partial x_1} & \cdots & \sum_{i=1}^{m} \frac{\partial g_1(h(x))}{\partial h_i(x)} \frac{\partial h_i(x)}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \sum_{i=1}^{m} \frac{\partial g_p(h_i(x))}{\partial h_i(x)} \frac{\partial h_i(x)}{\partial x_1} & \cdots & \sum_{i=1}^{m} \frac{\partial g_p(h(x))}{\partial h_i(x)} \frac{\partial h_i(x)}{\partial x_n} \end{pmatrix} =$$

ДАТА / DATE
ПН ВТ СР ЧТ ПТ СБ ВС
MON TUE WED THU FRI SAT SUN

$$= \begin{pmatrix} \frac{\partial g_1(h(x))}{\partial h_1} & \frac{\partial g_1(h(x))}{\partial h_2} & \cdots & \frac{\partial g_1(h(x))}{\partial h_m} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial g_p(h(x))}{\partial h_1} & \frac{\partial g_p(h(x))}{\partial h_2} & \cdots & \frac{\partial g_p(h(x))}{\partial h_m} \end{pmatrix} \times$$

$$\times \begin{pmatrix} \frac{\partial h_1(x)}{\partial x_1} & \frac{\partial h_1(x)}{\partial x_2} & \cdots & \frac{\partial h_1(x)}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial h_m(x)}{\partial x_1} & \frac{\partial h_m(x)}{\partial x_2} & \cdots & \frac{\partial h_m(x)}{\partial x_n} \end{pmatrix} =$$

$$= \frac{\partial g(h(x))}{\partial h} \cdot \frac{\partial h(x)}{\partial x}$$

№3

① дана выборка

x	1	1	0	0	-1
y	4	4	0	2	6

Построить модель $f(x) = \beta_0 + \beta_1 x + \beta_2 x^2$

$$\exists\ X = \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ 1 & x_4 & x_4^2 \\ 1 & x_5 & x_5^2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix} \quad y = \begin{pmatrix} 4 \\ 4 \\ 0 \\ 2 \\ 6 \end{pmatrix}$$

$L = \|X\beta - y\|^2$ для такой ф-ции потерь решением будет $X^T X \beta = X^T y$

$$X^T X = \begin{pmatrix} 5 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 3 \end{pmatrix}$$

$$X^T y = \begin{pmatrix} 16 \\ 2 \\ 14 \end{pmatrix}$$

$$\begin{cases} 5\beta_0 + \beta_1 + 3\beta_2 = 16 \\ \beta_0 + 3\beta_1 + \beta_2 = 2 \\ 3\beta_0 + \beta_1 + 3\beta_2 = 14 \end{cases} \Rightarrow \hat{\beta} = \begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix}$$

$$f(x) = 1 - x + 4x^2$$

DATA / DATE
ПН ВТ СР ЧТ ПТ СБ ВС
MON TUE WED THU FRI SAT SUN

№3

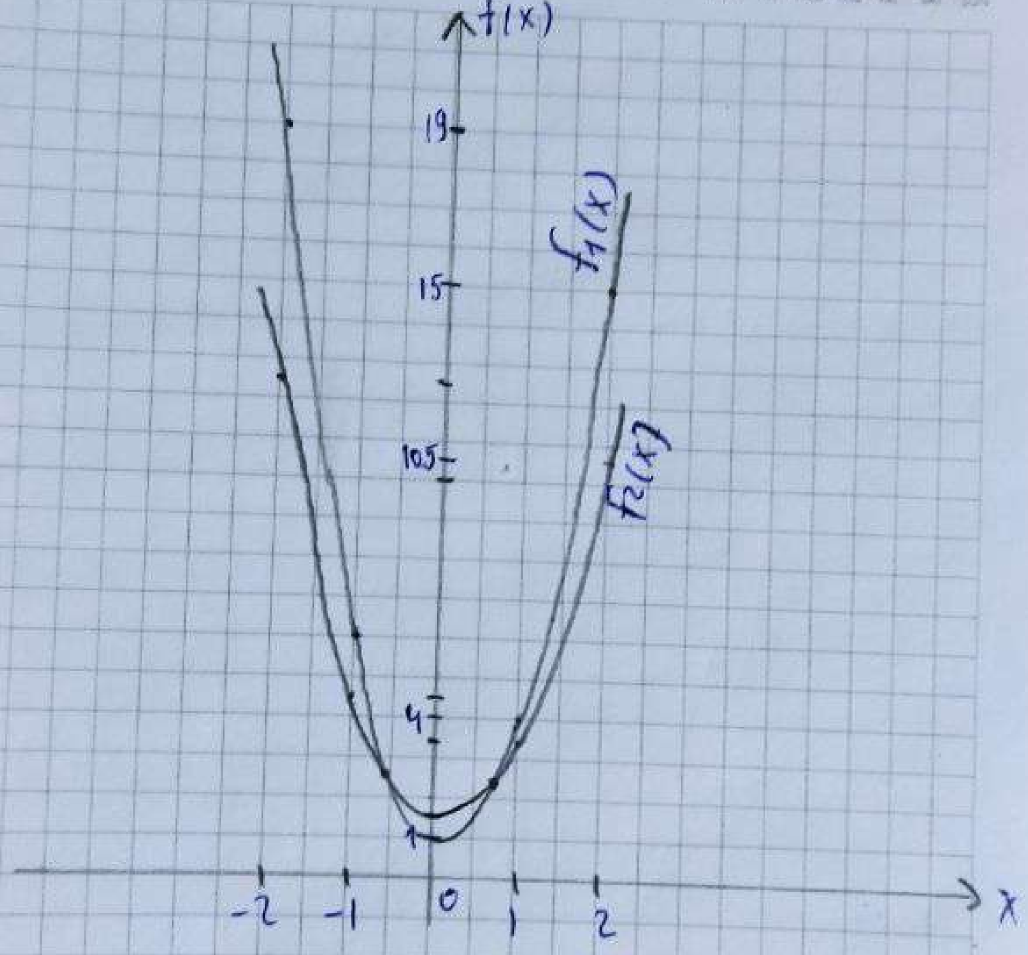
3) $f(x) = \beta_0 + \beta_1 x + \beta_2 x^2$, $\lambda = 1$

регуляриз. система: $(X^T X + \lambda I)\beta = X^T y$

$$X^T X + \lambda I = \begin{pmatrix} 6 & 1 & 3 \\ 1 & 4 & 1 \\ 3 & 1 & 4 \end{pmatrix}$$

$$\begin{cases} 6\beta_0 + \beta_1 + 3\beta_2 = 16 \\ \beta_0 + 4\beta_1 + \beta_2 = 2 \\ 3\beta_0 + \beta_1 + 4\beta_2 = 14 \end{cases} \Rightarrow \hat{\beta} = \begin{pmatrix} 1.5 \\ -0.5 \\ 2.5 \end{pmatrix}$$

$$f(x) = \frac{1}{2}(3 - x + 5x^2)$$



$f_1(x) = 4x^2 - x + 1$

$f_2(x) = \frac{1}{2}(5x^2 - x + 3)$