ML_hw Колесников 381803-2

I 1. $a \in R^n, x \in R^n$

док-во

$\Downarrow$

$$\frac{\partial(a^T x)}{\partial x} = a$$

$$a^T x = a_1 x_1 + \quad + a_n x_n$$

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

$$= (a_1 \quad a_n) = a^T \quad ?$$

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

$$Ax = \begin{pmatrix} \sum a_{1i} x_i \\ \\ \sum a_{mi} x_i \end{pmatrix} \quad \frac{\partial(Ax)}{\partial x_1} = \begin{pmatrix} a_{11} \\ \\ a_{m1} \end{pmatrix}$$

$$\frac{\partial(Ax)}{\partial x_n} = \begin{pmatrix} a_{1n} \\ \\ a_{mn} \end{pmatrix}$$

$$\frac{\partial(Ax)}{\partial x} = \begin{pmatrix} a_{11} & & a_{1n} \\ & & \\ a_{n1} & & a_{nn} \end{pmatrix} = A$$

3. $$\frac{\partial(x^T A x)}{\partial x} = x^T(A^T + A)$$

$$x^T A x = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij} x_i x_j$$

$$\frac{\partial(x^T A x)}{\partial x_1} = \sum_{i=2}^{n} a_{i1} x_i + 2a_{11} x_1 + \sum_{j=2}^{n} a_{1j} x_j$$

$$\frac{\partial(x^T A x)}{\partial x_k} = \sum_{i \neq k} a_{ik} x_i + a_{kk} x_k + a_{kk} x_k + \sum_{j \neq k} a_{kj} x_j = \sum_{i=1}^{n} a_{ik} x_i + \sum_{j=1}^{n} a_{1j} x_j$$

$$\frac{\partial (x^T A x)}{\partial x} = \begin{pmatrix} \sum a_{i1} x_i + \sum a_{1j} x_j \\ \\ \sum a_{in} x_i + \sum a_{nj} x_j \end{pmatrix}^T$$

$$x^T (A^T + A) = x^T A^T + x^T A$$

$$x^T A^T = (x_1 \quad x_n) \begin{pmatrix} a_{11} & \\ & a_{nn} \end{pmatrix} =$$

$$= \left(\sum x_i a_{1i}, \quad \sum a_{ni} x_i \right)$$

$$x^T A = (x_1 \quad x_n) \begin{pmatrix} a_{11} & a_{1n} \\ a_{21} & \\ & a_{nn} \end{pmatrix} =$$

$$= \left(\sum x_i a_{i1} \quad \sum x_i a_{in} \right)$$

$x^T(A^T + A) = (\Sigma + \Sigma, \quad , \Sigma + \Sigma)$

$= \dfrac{\partial(x^T A x)}{\partial x}$

—

если $A = A^T$:

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

4. $$\frac{\partial\|x\|^2}{\partial x} = 2x$$

$x = (x_1 \quad x_n)^T$

$\|x\|^2 = (x, x) = x_1^2 + \quad + x_n^2$

$$\frac{\partial\|x\|^2}{\partial x_i} = 2x_i \qquad \frac{\partial\|x\|^2}{\partial x} = (2x_1 \quad 2x_n)$$

$= 2(x_1 \quad x_n) = 2x^T$

5. $g(X) = \begin{pmatrix} g(x_1) \\ \vdots \\ g(x_n) \end{pmatrix}$

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

$$= \begin{pmatrix} g'(x_1) & 0 & \\ 0 & g'(x_2) & \\ 0 & & g'(x_n) \end{pmatrix} = \mathrm{diag}(g(x$$

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

$h = \begin{pmatrix} h_1 \\ \\ h_m \end{pmatrix}(x) \quad g = \begin{pmatrix} g_1 \\ \\ g_n \end{pmatrix}(h)$

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

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

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

$$\frac{\partial g(h(x))}{\partial h} \cdot \frac{\partial h(x)}{\partial x} = \begin{pmatrix} \sum_{i=1}^{n} \frac{\partial g_1}{\partial h_i} \frac{\partial h_i}{\partial x_1} & \sum_{i=1}^{m} \frac{\partial g_1}{\partial x_i} \frac{\partial h_i}{\partial x_n} \\ \sum_{i=1}^{n} \frac{\partial g_l}{\partial h_i} \frac{\partial h_i}{\partial x} & \sum \frac{\partial g}{\partial h_i} \frac{\partial h_i}{\partial x} \end{pmatrix}$$

1) $g(\beta) = \|X\beta - y\|^2$

$$\frac{\partial g(\beta)}{\partial \beta} = ? \qquad \frac{\partial^2 g(\beta)}{\partial \beta^T \partial \beta} = ?$$

$$\frac{\partial g(\beta)}{\partial \beta} = 2(X\beta - y)^T \frac{\partial (X\beta - y)}{\partial \beta} = \underline{2(X\beta - y)^T X}$$

$$\frac{\partial^2 g(\beta)}{\partial \beta^T \partial \beta} = \frac{\partial \left(\frac{\partial g\beta}{\partial \beta}\right)}{\partial \beta^T} = \frac{\partial (2(X\beta - y)^T X)}{\partial \beta^T}$$

$$= \frac{\partial (2\beta^T X^T X - y^T X)}{\partial \beta^T} = 2X^T X$$

$$\frac{\partial g(\beta)}{\partial \beta} = 0 \qquad 2(X\beta - y)^T X = 0$$

$$2X^T(X\beta - y) = 0$$

$X^T X\beta = X^T y$ z n.l.l. g

$\hat{\beta} = \arg\min \|X\beta - y\|^2$?

$$\frac{\partial^2 g(\beta)}{\partial\beta\partial\beta} = 2X^T X > 0 \Leftarrow$$ решение

Этого условия (β^*: $\frac{\partial g}{\partial \beta} = 0$) — лок мин

$$g(\beta) = g(\beta^*) + (g'(\beta^*), h) + \frac{1}{2}(g''(\beta^*)h, h) -$$

$$g(\beta) = g(\beta^*) + \underset{\varepsilon > 0}{\varepsilon} > g(\beta^*)$$

т.к. $g'' > 0$

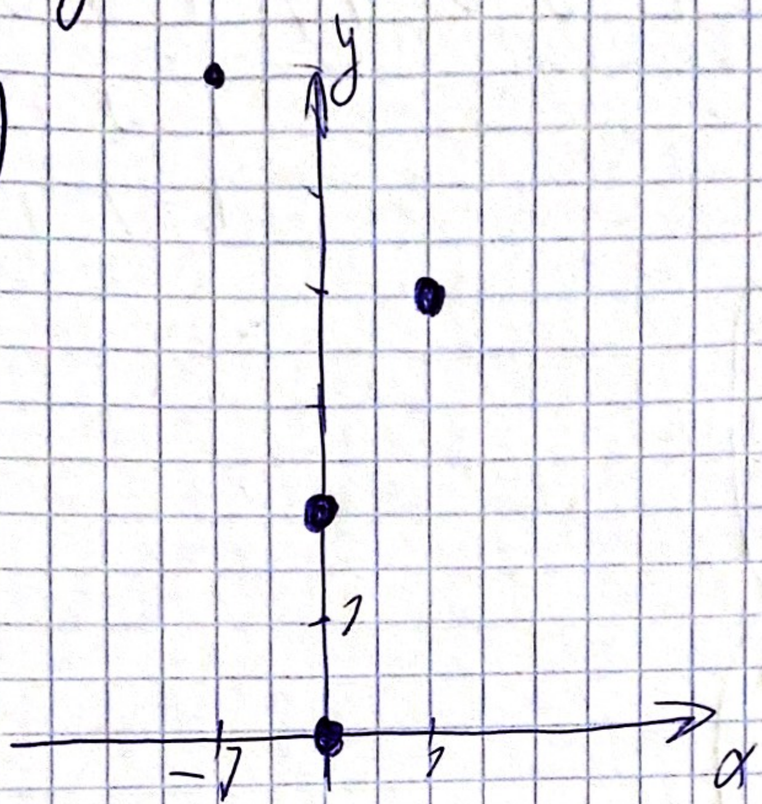
$\Downarrow$

β^* — мин мнк

III

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

1)



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

$$X = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & -1 \end{pmatrix} \quad \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}$$

$$y = (4 \;\; 4 \;\; 0 \;\; 2 \;\; 6)^T$$

$$X^T X = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & -1 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix} =$$

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

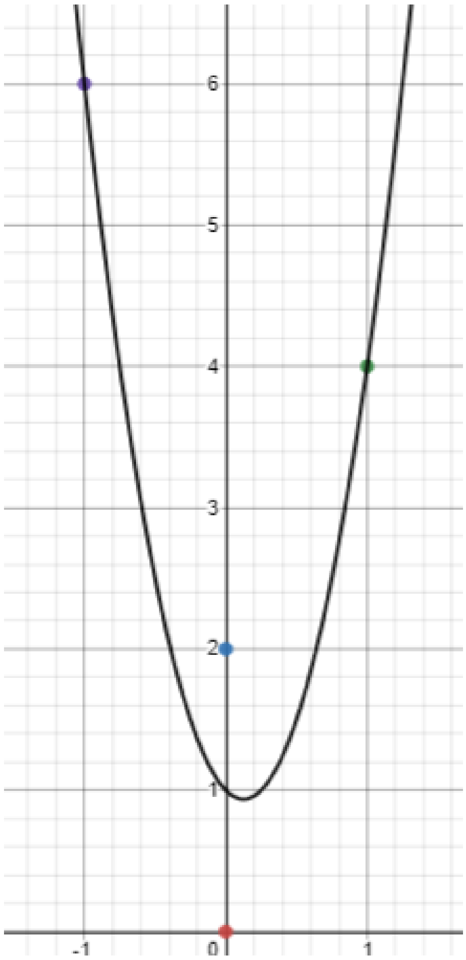
$$X^T y = \begin{pmatrix} 16 \\ 2 \\ 14 \end{pmatrix}$$

$$X X^T \beta = X^T y$$

$$\begin{pmatrix} 5 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 3 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} 16 \\ 2 \\ 14 \end{pmatrix}$$

$\beta_0 = 7 \quad \beta_1 = -1 \quad \beta_2 = 4$

$$6(E x) = 7 - x + 4x^2$$



3) $\lambda = 2$ $\quad (X^T X + \lambda I)\beta = X^T y$

$$\begin{pmatrix} 6 & 1 & 3 \\ 1 & 4 & 1 \\ 3 & 1 & 4 \end{pmatrix} \beta = \begin{pmatrix} 16 \\ 2 \\ 14 \end{pmatrix}$$

$\beta_0 = \frac{3}{2} \quad \beta_1 = -\frac{1}{2} \quad \beta_2 = \frac{5}{2}$

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