ДЗ Deadline 17.11.2025 № 1–4

Мартынов А.Ю.
3825М1ПР1

№ 1

$a \in \mathbb{R}^n, x \in \mathbb{R}^n$

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

1) $a^T x = a_1 x_1 + \ldots + a_n x_n$; $f(x) = a^T x$

$\frac{\partial f(x)}{\partial x} = \nabla f(x)$ ← это вектор

$$\frac{\partial(a_1 x_1 + \ldots + a_n x_n)}{\partial x} = \begin{pmatrix} \frac{\partial(a_1 x_1 + \ldots + a_n x_n)}{\partial x_1} \\ \vdots \\ \frac{\partial(a_1 x_1 + \ldots + a_n x_n)}{\partial x_n} \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = a$$

2) $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$

$$Ax = \begin{bmatrix} a_{11} x_1 + \ldots + a_{1n} x_n \\ a_{21} x_1 + \ldots + a_{2n} x_n \\ \ldots \\ a_{m1} x_1 + \ldots + a_{mn} x_n \end{bmatrix} = \alpha \in \mathbb{R}^m \quad [\alpha_i - i\text{-ая строчка вектора } \alpha]$$

$$\frac{\partial(Ax)}{\partial x} = \overset{\text{якобиан}}{J(Ax)} = \begin{bmatrix} \frac{\partial \alpha_1}{\partial x_1} & \frac{\partial \alpha_1}{\partial x_2} & \ldots & \frac{\partial \alpha_1}{\partial x_n} \\ \frac{\partial \alpha_2}{\partial x_1} & \ldots & \ldots & \frac{\partial \alpha_2}{\partial x_n} \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial \alpha_m}{\partial x_1} & \ldots & \ldots & \frac{\partial \alpha_m}{\partial x_n} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ \vdots & a_{22} & & \vdots \\ a_{m1} & \ldots & \ldots & a_{mn} \end{bmatrix} = A$$

3) $A \in \mathbb{R}^{n \times n}$, $x \in \mathbb{R}^n$

$1 \times n \cdot n \times n = 1 \times n$ - строка

1. $x^T A = [x_1 \ldots x_n] \begin{bmatrix} a_{11} & a_{12} \\ a_{n1} & a_{nn} \end{bmatrix} = \left[\sum_{i=1}^n a_{i1} x_i \quad \sum_{i=1}^n a_{i2} x_i \quad \ldots \quad \sum_{i=1}^n a_{in} x_i \right]$

$1 \times n \cdot n \times 1 = [1]$ - скаляр; $x^T A x$ ~ скалярное умножение векторов в аффинном пр-ве

$$[\ldots] \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{j=1}^n x_j \left(\sum_{i=1}^n a_{ij} x_i \right)$$

$$\frac{\partial \sum_{j=1}^n x_j \left(\sum_{i=1}^n a_{ij} x_i \right)}{\partial x_k} = \sum_{j=1}^n \frac{\partial \left(x_j \left(\sum_{i=1}^n a_{ij} x_i \right) \right)}{\partial x_k} = \alpha$$

Теперь возможно несколько случаев.

1) $i \neq k, j \neq k$, тогда $\alpha = 0$.

2) $i = j = k$, тогда получаем диагональные элементы: $2a_{kk} x_k = (A^T_{kk} + A_{kk}) \cdot x_k$

3) $i = k, j \neq k$, тогда $\alpha = a_{kj} x_j = A^T_{jk} x_k$

4) $i \neq k, j = k$, тогда $\alpha = A_{ik} x_k$

После этого можем собрать итоговую матрицу A и получим $(A + A^T) x$

Если $A = A^T$, то тогда 3) и 4) эквивалентны, т.к. $A = A^T$, и опять же соберём матрицу A и получим $2Ax$ ЧТД

4) $\|x\|^2 = \sqrt{(x,x)}^2 = \left(\sqrt{x^T x}\right)^2$

$$\left[\frac{\partial(x_1^2 + \ldots + x_n^2)}{\partial x_i}\right]_{i=\overline{1,n}} \overset{\nabla \text{ градиент}}{=} \begin{bmatrix} \frac{\partial(x_1^2+\ldots+x_n^2)}{\partial x_1} \\ \vdots \\ \frac{\partial(x_1^2+\ldots+x_n^2)}{\partial x_n} \end{bmatrix} = \begin{bmatrix} 2x_1 \\ 2x_2 \\ \vdots \\ 2x_n \end{bmatrix} = 2x$$

5) $\frac{\partial g(x_i)}{\partial x_j}$; $\frac{\partial g(x_i)}{\partial x_j} = \begin{cases} g'(x_i), & i = j \\ 0, & j \neq i \end{cases}$ (*)

$$\frac{\partial \begin{pmatrix} g(x_1) \\ \vdots \\ g(x_n) \end{pmatrix}}{\partial x} = \begin{bmatrix} \frac{\partial g(x_1)}{\partial x_1} & \frac{\partial g(x_1)}{\partial x_2} & \cdots & \frac{\partial g(x_1)}{\partial x_n} \\ \frac{\partial g(x_2)}{\partial x_1} & \frac{\partial g(x_2)}{\partial x_2} & \cdots & \frac{\partial g(x_2)}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial g(x_n)}{\partial x_1} & \cdots & \cdots & \frac{\partial g(x_n)}{\partial x_n} \end{bmatrix} \overset{\text{пользуясь (*) получаем}}{=} \cancel{\begin{bmatrix} \frac{\partial g(x_1)}{\partial x_1} & 0 & \cdots & 0 \\ 0 & \partial_\theta & & \\ & & \ddots & \end{bmatrix}}$$

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

6) ~~Как~~ $h: \mathbb{R}^n \to \mathbb{R}^m$, $g: \mathbb{R}^m \to \mathbb{R}^p$, $x \in \mathbb{R}^n$, то $\frac{\partial g(h(x))}{\partial x} = \frac{\partial g(h(x))}{\partial h} \cdot \frac{\partial h(x)}{\partial x}$

$h(x) = \begin{bmatrix} h_1(x) \\ \vdots \\ h_m(x) \end{bmatrix}$; $\frac{\partial g_1(h(x))}{\partial x_i} = \frac{\partial g_1}{\partial h_1}\frac{\partial h_1}{\partial x_i} + \frac{\partial g_1}{\partial h_2}\frac{\partial h_2}{\partial x_i} + \ldots$

$$\frac{\partial g(h(x))}{\partial x} = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & & \\ \frac{\partial g_p}{\partial x_1} & \cdots & \frac{\partial g_p}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial g_1}{\partial h_1} & \cdots & \frac{\partial g_1}{\partial h_m} \\ \frac{\partial g_p}{\partial h_1} & \cdots & \frac{\partial g_p}{\partial h_m} \end{bmatrix} \cdot \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \cdots & \frac{\partial h_1}{\partial x_n} \\ \vdots & \cdots & \vdots \\ \frac{\partial h_m}{\partial x_1} & \cdots & \frac{\partial h_m}{\partial x_n} \end{bmatrix} =$$

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

№2

$g(\beta) = \|X\beta - y\|^2$ Описать реш.; найти $\nabla g(\beta)$ $\nabla^2_{\beta\beta^T} g(\beta)$

(P.S) вспоминаем ∇, Y....

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

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

$= \beta^T X^T X\beta - \underbrace{\beta^T X^T y}_{\text{скаляр}} - \underbrace{y^T X\beta}_{\text{скаляр}} + y^T y \Rightarrow$ $\cancel{\beta^T X^T X\beta}$.

скаляр скаляр, $\exists \alpha \in \mathbb{R}^{1\times 1}$: тогда $\alpha = \alpha^T$ - потребуется дальше.

$\beta^T X^T X\beta = \|X\beta\|^2$ (*)

$\beta^T X^T y = 2^T X\beta$, как скаляры $= (-2y^T X\beta)$

(через несколько шагов)

$\nabla g(\beta) = \frac{\partial \beta^T X^T X\beta}{\partial \beta} - \frac{\partial(\beta^T X^T y)}{\partial \beta} - \frac{\partial (y^T X\beta)}{\partial \beta} + \frac{\partial(y^T y)}{\partial \beta} \xrightarrow{0}$ восп. (*) $=$

$= \frac{\partial \|X\beta\|^2}{\partial \beta} - \left(\frac{\partial(\beta^T X^T y)}{\partial \beta}\right)^{T^T} - y^T X \overset{(\text{св-во } 1^{\circ})}{=} \frac{\partial \|X\beta\|^2}{\partial \beta} - \frac{\partial (y^T X\beta)^T}{\partial \beta} - y^T X =$

$\cancel{\partial\|X\beta\|}$ $\frac{\partial \|X\beta\|^2}{\partial \beta} = \frac{\partial \|X\beta\|^2}{\partial (X\beta)} \cdot \frac{\partial (X\beta)}{\partial \beta} = 2X\beta \cdot \frac{\partial X\beta}{\partial \beta}$

св-в

$\nabla g(\beta) = 2(X\beta)\frac{\partial (X\beta)}{\partial \beta} - 2y\frac{\partial (X\beta)}{\partial \beta} = 2(X\beta - y)X \Rightarrow 0$

Найдем реш. $2(X\beta - y)X \Rightarrow 0$

$\nabla^2_{\beta\beta^T}$

$X \neq 0$, тогд. $X\beta - y \Rightarrow 0$.

$X\beta = y$ – ~~...~~ единств. уже было.

$\nabla^2 g(\beta) = \frac{\partial\, 2(X\beta - y)X}{\partial \beta^T} \overset{\frac{\partial yx}{\partial \beta^T} = 0}{=} 2\frac{\partial X\beta X}{2\partial \beta^T} = 2\frac{\partial X\beta X}{\partial \beta X}$.

но провер. на min/max

$\left(\frac{\partial \beta X^T}{\partial \beta^T}\right)^T = 2X^2 > 0$

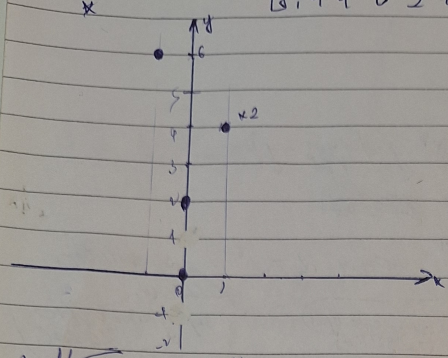
$\Rightarrow$ Точка min, тогд. $\exists \hat\beta$: $X\hat\beta = y$ и ~~...~~ $\hat\beta = \operatorname{argmin} g(\beta)$

№3.

стр 1

1) Изобразить точки

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



2) ~~Не~~ Применяя МНК к квадр. модели вида

$f(x) = \beta_0 + \beta_1 x + \beta_2 x^2$ и постр. график

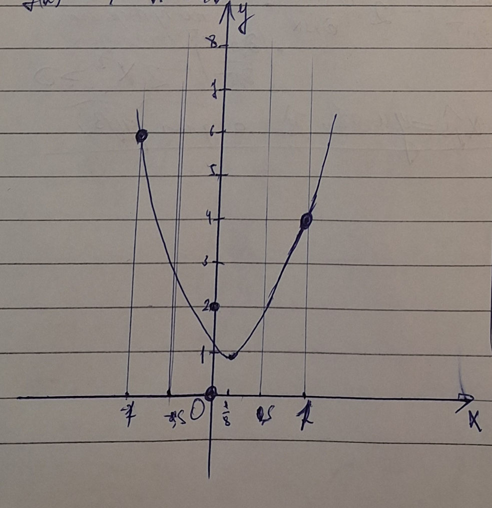
$$y = \begin{bmatrix} 4 \\ 4 \\ 0 \\ 2 \\ 6 \end{bmatrix} \quad X = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{bmatrix}$$

$$X^T X = \begin{bmatrix} 5 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 3 \end{bmatrix}, \quad X^T y = \begin{bmatrix} 16 \\ 2 \\ 14 \end{bmatrix}$$

$X^T X \beta = X^T y$ – решая отн-но β

$$\begin{bmatrix} 5 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 3 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 16 \\ 2 \\ 14 \end{bmatrix} \rightarrow \beta = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$$

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



№ 3 стр 2

3) Квадр. модель рег. с $\lambda = 1$

$$X^T X + \lambda I = \begin{bmatrix} 5 & 1 & 3 \\ 1 & 3 & 1 \\ 3 & 1 & 3 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 1 & 3 \\ 1 & 4 & 1 \\ 3 & 1 & 4 \end{bmatrix}$$

Решаем систему

$$\begin{bmatrix} 6 & 1 & 3 \\ 1 & 4 & 1 \\ 3 & 1 & 4 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 16 \\ 2 \\ 14 \end{bmatrix} \rightarrow \beta_0 = \frac{3}{2}, \quad \beta_1 = -\frac{1}{2}, \quad \beta_2 = \frac{5}{2}$$

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