Бюрократия

$0.7 \cdot HW + 0.3 \cdot Test$ Филипп ТМ

$\uparrow$ 4 шт. $\uparrow$ 4 шт.

Небольшие воспоминания

$\underset{[m\times n]}{A}$ $\underset{[n\times k]}{B}$ $C = AB$ ~~BA~~

$A \in \mathbb{R}^{n\times m}$

$x \in \mathbb{R}^n$
вектор столбец
$x = \underset{[n\times 1]}{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}$

$AB \neq BA$

$(AB)^T = B^T A^T$

$tr(A) = tr\begin{pmatrix} a_{11} & & * \\ & a_{22} & \\ * & & \ddots \; a_{nn} \end{pmatrix} = \sum_{i=1}^{n} a_{ii}$

$\langle x, y\rangle = x^T \cdot y = (x_1, \dots, x_n) \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{i=1}^{n} x_i y_i$

$\|x\| = \sqrt{\langle x, x\rangle} = \sqrt{\sum x_i^2}$ $\|A\|_F = \sqrt{\sum_{i,j} a_{ij}^2} = \sqrt{\langle A, A\rangle_F} = \sqrt{tr(A^T A)}$

$\langle A, B\rangle_F = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij}\cdot b_{ij} = tr(A^T B)$

$$\underset{[2\times 3]}{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \qquad \underset{[2\times 3]}{B} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix}$$

$$tr\left(A^T B\right) = tr\left(\begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix} \right) =$$

$$= tr \begin{pmatrix} a_{11}b_{11} + a_{21}\cdot b_{21} & & * \\ & \bullet & \\ * & & \bullet \end{pmatrix} = \bullet + \bullet + \bullet$$

$$\sum_{ij} \left(x_{ij} - a_{ij}\right)^2 \to \min_A$$

$X \qquad A$

$\qquad\; A^T = A$

$$\begin{cases} tr\left(\left(X - A^T\right)(X - A)\right) \to \min\limits_A \\ A^T = A \end{cases}$$

Дифференциалы и производные

① $f(x): \mathbb{R} \to \mathbb{R}$

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

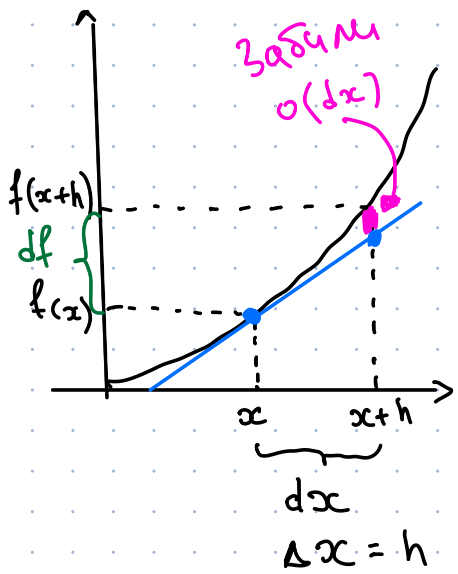
$f(x) = x^2$

$f'(x) = 2x$

<u>мысль:</u>

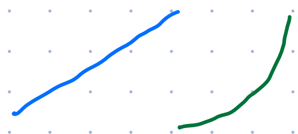
матан – искусство забивать

$$f(x+h) - f(x) = f'(x) \cdot h + o(h)$$



$$f(x+h) - f(x) = f'(x)\cdot h + \frac{1}{2}\cdot f''(x)\cdot h^2 + o(h^2)$$

$$df = f'(x)\cdot dx$$



② $f(x_1, \dots, x_n) : \mathbb{R}^n \to \mathbb{R}$

$$\nabla_x f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \dots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

$$df(x) = f(x_1 + dx_1, \dots, x_n + dx_n) - f(x_1, \dots, x_n)$$

f'_{x_1} $\qquad L[dx]$

$$df(x) = \underbrace{\frac{\partial f}{\partial x_1}}_{= f'_{x_1}} dx_1 + \frac{\partial f}{\partial x_2} dx_2 + \dots + \frac{\partial f}{\partial x_n} dx_n =$$

$$= \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\right)\cdot\begin{pmatrix} dx_1 \\ \vdots \\ dx_n \end{pmatrix} = \langle \nabla_x f, dx \rangle = \nabla_x^T f \cdot dx$$

$$f(x) = o(g(x))$$

$$\lim_{x\to 0} \frac{g(x)}{f(x)} = 0 \qquad \lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$$

$x \in$
$x^2 \in$ $o(\ln x)$
$e^x \in$

мн-во

Обобщим всё это на матрицы.

$$f: \underset{\Delta}{\underbrace{\mathbb{R}^{n\times k}}} \to \underbrace{\mathbb{R}} \qquad \lim_{\|H\|\to 0} \frac{f(X+H)-f(X)}{H}$$

Делим на матрицу?

$$f(X+H)-f(X) = \underbrace{L[dX]} + o(\|dX\|)$$

$$X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$$

$$\det X = x_{11}x_{22} - x_{21}x_{12}$$

чёт линейное

$$d(\det X) = L[dx_{11}, dx_{12}, dx_{21}, dx_{22}] + o(\|dx\|)$$

$$\frac{\partial f}{\partial x_{11}}, \dots, \frac{\partial f}{\partial x_{22}}$$

$$x_{22} \cdot dx_{11} - x_{21} \cdot dx_{12} - x_{12}\, dx_{21} + x_{11}\, dx_{22}$$

Таблица матричных производных

x \ $f(x)$	скаляр	вектор	матрица
скаляр	① $\underset{[1\times 1]}{f'(x)}\, dx$	④ $\underset{[m\times 1]}{J} \cdot \underset{[1\times 1]}{dx}$	$[n\times k]$
вектор	② $\underset{[1\times n]}{\nabla_x^T f(x)}\, \underset{[n\times 1]}{dx}$	③ $\underset{[m\times n]}{J} \cdot \underset{[n\times 1]}{dx}$	$m\times n\times k$
матрица	⑤ $tr(\underset{[k\times n]}{\nabla_X^T f} \cdot \underset{[n\times k]}{dX})$	$[m\times n\times k]$	$m\times \ell\times n\times k$

$$dA = 0$$ Все кроме α матрицы

$$d(\alpha \cdot X) = \alpha \cdot dX$$

$$d(X \cdot Y) = dX \cdot Y + X \cdot dY$$

$$\boxed{dX \cdot Y \neq Y \cdot dX}$$

$$d(X^T) = (dX)^T$$

$$(const)' = 0$$

$$(\alpha \cdot f)' = \alpha f'$$

$$(f \cdot g)' = f' g + f g'$$

Цыганские фокусы:

$S_{1\times1}$ $\quad S^T = S$

скаляр $\quad tr(S) = S$

Упражнение

$$f(x) = a^T x \qquad a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \qquad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

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

$$\nabla_x f(x)$$

$$f(x) = (a_1, \dots, a_n) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = a_1 x_1 + a_2 x_2 + \dots + a_n x_n$$

$$\frac{\partial f(x)}{\partial x_j} = a_j \qquad \nabla_x f = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = a$$

$$df(x) = d(a^T x) = a^T dx$$

$$f: \mathbb{R}^n \to \mathbb{R} \qquad \Rightarrow \nabla_x f = a$$

$$df = \nabla_x^T f \cdot dx$$

Упражнение

$$f(x) = \underset{[1\times n]}{x^T}\,\underset{[n\times n]}{A}\,\underset{[n\times 1]}{x} \quad [1\times 1] \qquad x \in \mathbb{R}^n \qquad A \in \mathbb{R}^{n\times n}$$

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

$$f(x) = (x_1, \ldots, x_n) \cdot \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

$$\frac{\partial f}{\partial x_i} = \ldots$$

дифф

критем

местб

$$df = \nabla_x^T f \cdot dx$$

$$df = d(x^T A x) = dx^T \cdot A \cdot x + x^T d(Ax) =$$

$$= \underset{[1\times n]}{dx^T} \cdot \underset{[n\times n]}{A} \cdot \underset{[n\times 1]}{x} + x^T \cdot \cancel{dA}\, x + x^T A\, dx =$$

0

$[1\times 1]$

скаляр

$$= x^T A^T dx + x^T A\, dx =$$

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

$[1\times 1] \qquad [1\times 1]$

$$= x^T (A^T + A)\, dx \qquad \nabla_x^T f$$

$$\nabla_x f = (A + A^T)\, x$$

$$[x^T (A^T + A)]^T = (A^T + A)^T x = (A + A^T)\, x$$

$$\nabla_x^2 f(x) = H = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \ldots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \ldots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

матрица Гессе

$$\underset{[n\times 1]}{g(x)} = \underset{[n\times n]}{(A + A^T)}\, \underset{[n\times 1]}{x}$$

вектор   вектор

$$dg = d((A + A^T)x) = \underbrace{(A + A^T)}_{\nabla_x^2 f}\, dx$$

$$\nabla_x^2 f = (A + A^T)$$

③ $g(x): \mathbb{R}^n \to \mathbb{R}^m$ $\quad dg = \mathcal{Y} \cdot dx$ матрица Якоби

$$\begin{cases} g_1(x_1, \ldots, x_n) \\ g_2(x_1, \ldots, x_n) \\ \ldots \\ g_m(x_1, \ldots, x_n) \end{cases}$$

$$\underset{[m \times 1]}{\begin{pmatrix} dg_1 \\ dg_2 \\ \vdots \\ dg_m \end{pmatrix}} = \underset{[m \times n]}{\begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \ldots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \ldots & \frac{\partial g_2}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial x_1} & \ldots & \frac{\partial g_m}{\partial x_n} \end{pmatrix}} \underset{[n \times 1]}{\begin{pmatrix} dx_1 \\ dx_2 \\ \vdots \\ dx_n \end{pmatrix}}$$

m - число ф.
n - число арг.

④ $f: \mathbb{R} \to \mathbb{R}^m$

$$x \xrightarrow{f} \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix} \qquad \frac{\partial f}{\partial x} = \begin{pmatrix} \frac{\partial f_1(x)}{\partial x} \\ \ldots \\ \frac{\partial f_m(x)}{\partial x} \end{pmatrix}$$

$$df = \begin{pmatrix} df_1(x) \\ \vdots \\ df_m(x) \end{pmatrix} = \underset{[m \times 1]}{\begin{pmatrix} \frac{\partial f_1(x)}{\partial x} \\ \ldots \\ \frac{\partial f_m(x)}{\partial x} \end{pmatrix}} \cdot \underset{[1 \times 1]}{(dx)} = \mathcal{Y} \cdot dx$$

Упражнение

$$y = \underset{[n \times 1]}{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}} \qquad X = \underset{[n \times k]}{\begin{bmatrix} x_{11} & x_{12} & \ldots & x_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \ldots & x_{nk} \end{bmatrix}} \qquad w = \underset{[k \times 1]}{\begin{bmatrix} w_1 \\ \vdots \\ w_k \end{bmatrix}}$$

n - кол-во набл.
k - кол-во факторов

$x_i = \begin{pmatrix} x_{i1} \\ \vdots \\ x_{ik} \end{pmatrix}$ - строка наблюдение (1 квартира)

кол-во комнат / площадь / этаж и т.д.

$$\hat{y}_i = w_1 x_{i1} + w_2 x_{i2} + \ldots + w_k x_{ik}$$

прогноз

$$L(x_i, y_i, w) = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \longrightarrow \min_w$$

MSE

$$(y - Xw) = \begin{pmatrix} y_1 - x_1^T w \\ y_2 - x_2^T w \\ \ldots \\ y_n - x_n^T w \end{pmatrix}$$

$\langle x_i, w \rangle$

$$\frac{1}{n}(y - Xw)^T (y - Xw) = \frac{1}{n}\sum_{i=1}^{n} (y_i - x_i^T w)^2$$

$$h^T h = (h_1, \ldots, h_n) \cdot \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = \sum_i h_i^2 = \langle h, h \rangle$$

$$L(w) = \frac{1}{n}(y - Xw)^T (y - Xw) \longrightarrow \min_w$$

$$\underset{w}{L : \mathbb{R}^k \to \mathbb{R}}$$

w – переменная
X, y – константы

$0 - Xdw$
$dy - dXw$

$$dL(w) = \frac{1}{n} d(y - Xw)^T \cdot (y - Xw) + \frac{1}{n}(y - Xw)^T \cdot d(y - Xw) =$$

$$= -\frac{1}{n}\Big[\underbrace{\overset{[1\times k][k\times n]\ [n\times 1]}{dw^T X^T (y - Xw)}}_{[1\times 1]} + (y - Xw)^T X\, dw \Big] = *$$

$$[d(y - Xw)]^T = (-Xdw)^T = -dw^T X^T$$

$$[dw^T X^T (y - Xw)]^T = (y - Xw)^T X\, dw$$

$$* = -\frac{2}{n}(y - Xw)^T X\, dw$$

$$\nabla_w L = -\frac{2}{n} X^T (y - Xw)$$

$$-\frac{2}{n}X^T(y-Xw)=0$$

$$X^Ty - X^TXw = 0$$

$$X^TXw = X^Ty \quad \Big| \cdot (X^TX)^{-1} \text{ слева}$$

$$w = (X^TX)^{-1}X^Ty$$ МНК оценка

А это точно точка минимума?

$\nabla^2_w L$ $\qquad g(w) = -\frac{2}{n}\underset{[k\times n]}{X^T}\underset{[n\times 1]}{(y-Xw)}$

$g: \mathbb{R}^k \to \mathbb{R}^k \qquad dg = J\,dx$

$$dg = -\frac{2}{n}X^T d(y-Xw) = \underbrace{-\frac{2}{n}X^T(-X)}_{J}\,dw$$

$$\nabla^2_w L = \frac{2X^TX}{n}$$
$$\parallel$$
$$H$$

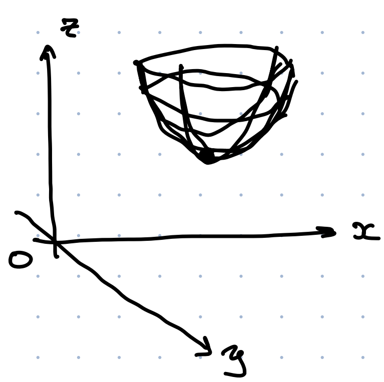
Критерий Сильвестра:

H – полож. опр $\Rightarrow$ минимум

H – отр. опр $\Rightarrow$ максимум

Опр.

$A \succ 0$ если $\forall v \in \mathbb{R}^n \quad v^TAv > 0$



$$v^T \cdot \frac{2X^TX}{n} \cdot v = \frac{2}{n}\underbrace{(Xv)^T}_{z}\underbrace{Xv}_{z} = \frac{2}{n}z^Tz = \frac{2}{n}\sum_{i=1}^{k} z_i^2 > 0$$

$\Rightarrow$ мы в точке минимума

$f' = 0 \qquad w_1 \quad w_2$

x^0

f''

⑤ $f(X): \mathbb{R}^{n\times k} \to \mathbb{R}$

$$\nabla_X f(X) = \begin{pmatrix} \frac{\partial f}{\partial x_{11}} & \frac{\partial f}{\partial x_{12}} & \cdots & \frac{\partial f}{\partial x_{1k}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f}{\partial x_{n1}} & \frac{\partial f}{\partial x_{n2}} & \cdots & \frac{\partial f}{\partial x_{nk}} \end{pmatrix} \qquad dX = \begin{pmatrix} dx_{11} & dx_{12} & \cdots & dx_{1k} \\ \vdots & \vdots & \ddots & \vdots \\ dx_{n1} & dx_{n2} & \cdots & dx_{nk} \end{pmatrix}$$

$\underset{[n\times k]}{X}$ $\qquad [n\times k]$

$$df = \frac{\partial f}{\partial x_{11}} dx_{11} + \frac{\partial f}{\partial x_{12}} dx_{12} + \ldots + \frac{\partial f}{\partial x_{nk}} dx_{nk}$$

$$df(X) = tr\left(\nabla_X^T f(X) \cdot dX\right)$$

Упражнение

1×1

$$f(X) = \underset{[1\times n]}{a^T} \underbrace{XAX}_{[n\times n]} \underset{[n\times 1]}{a}$$

$a \in \mathbb{R}^n \quad X \in \mathbb{R}^{n\times n}$

$A \in \mathbb{R}^{n\times n}$

$tr(A+C) = tr(A) + tr(C)$

$tr(ABC) = tr(CAB) = tr(BCA)$

$\nabla_X f$ – ?

$$df = d(a^T X A X a) = \underset{1\times 1}{a^T \cdot dX \cdot A X a} + \underset{1\times 1}{a^T X A \cdot dX \cdot a} =$$

$$= tr(\underset{1\times n}{a^T \cdot dX} A X \underset{n\times 1}{a}) + tr(\underset{1\times n}{a^T} X A \cdot dX \underset{n\times 1}{a}) =$$

$$= tr\left(AXaa^T dX + aa^T XA\, dX\right) =$$

$$= tr\left(\left[AXaa^T + aa^T XA\right] dX\right)$$

$$tr\left(\nabla_X^T f(X) \cdot dX\right) \qquad (AB)^T = B^T A^T$$

$$\nabla_X f(X) = \left[AXaa^T + aa^T XA\right]^T =$$

$$= (AXaa^T)^T + (aa^T XA)^T = \underbrace{aa^T}_{[n\times n]}\underbrace{X^T A^T}_{[n\times n]} + \underbrace{A^T}_{[n\times n]}\underbrace{X^T aa^T}_{[n\times n]}$$

$$\underset{[n\times n]}{X}$$

Упражнение

$$f(x) = xx^T x \qquad x \in \mathbb{R}^n \qquad \nabla_x f(x)$$

$$f(x): \mathbb{R}^n \to \mathbb{R}^n$$

$$\underset{[1\times 1]}{(dx^T \cdot x)^T} = \underset{[1\times 1]}{x^T dx}$$

$$df = J dx$$

$$df = d(xx^Tx) = \underbrace{dx}_{n\times 1} \cdot \underbrace{x^T x}_{\substack{1\times 1 \\ \text{scalar}}} + \underbrace{x}_{n\times 1} \cdot \underbrace{dx^T}_{1\times n} \cdot \underbrace{x}_{n\times 1} + \underbrace{xx^T}_{n\times n} \cdot \underbrace{dx}_{n\times 1} =$$

$$= \underbrace{\underset{[1\times 1]}{x^T x} \cdot \overset{I_n}{} \underset{[n\times 1]}{dx}}_{[n\times 1]} + \underbrace{\underset{[n\times n]}{x \cdot x^T}\, \underset{[n\times 1]}{dx}}_{[n\times 1]} + \underbrace{\underset{[n\times n]}{xx^T}\, \underset{[n\times 1]}{dx}}_{[n\times 1]} =$$

$$= \left[x^T x \cdot \underset{\cancel{1\times 1}\ n\times n}{I_n} + \underset{n\times n}{2xx^T}\right] dx \qquad J = \left[x^T x \cdot I_n + 2xx^T\right]$$

Почему так?

$$\alpha \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \alpha a_1 \\ \alpha a_2 \end{pmatrix}$$

$$\begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \alpha a_1 + \cancel{0 \cdot a_2} \\ \cancel{0 \cdot a_1} + \alpha a_2 \end{pmatrix}$$

Упражнение

Найдите симметричную матрицу $\underset{[n \times n]}{X}$ наиболее близкую к $\underset{[n \times n]}{A}$ по $\|\cdot\|_F^2$

$$\begin{cases} \|X - A\|_F^2 \to \min\limits_X \\ X^T = X \end{cases}$$

$$\|X - A\|_F^2 = tr\left((X-A)^T(X-A)\right)$$

$$\| \quad \sum_{i,j} (x_{ij} - a_{ij})^2$$

$$\mathcal{L} = tr\left((X-A)^T(X-A)\right) + \sum_{i,j} (x_{ij} - x_{ji}) \cdot \lambda_{ij}$$

$$x_{12} = x_{21} \qquad tr\left(\underset{[n \times n]}{\Lambda^T} \cdot (X - X^T)\right)$$

$$x_{ij} = x_{ji}$$

$$\mathcal{L} = tr\left((X-A)^T(X-A) + \Lambda^T \cdot (X - X^T)\right)$$

$$\frac{\partial \mathcal{L}}{\partial X} = ?$$

$$d\mathcal{L} = tr\left[d(X-A)^T(X-A) + \Lambda^T d(X - X^T)\right] =$$

$$= tr\left(dX^T\cdot(X-A) + (X-A)^T dX + \Delta^T dX - \Delta^T dX^T\right) =$$

$(X-A)^T dX$ $\qquad$ $dX\Delta$ $\quad ()^T$

ΔdX

$$= tr\left(\left(2(X-A)^T + \Delta^T - \Delta\right) dX\right)$$

$$\frac{\partial \mathcal{L}}{\partial X} = 2X - 2A + \Delta - \Delta^T = 0 \qquad ()^T$$

$$2X^T - 2A^T + \Delta^T - \Delta = 0 \quad +$$

$$X + X^T - A - A^T = 0 \qquad X = X^T$$

из условия задачи

$$2X - A - A^T = 0$$

$$X = \frac{1}{2}(A + A^T)$$

Производная сложной функции

$$(f(g(x)))' = g'(x)\cdot f'(g(x))$$

$$h = f\circ g : \mathbb{R}^n \to \mathbb{R}$$

$g : \mathbb{R}^n \to \mathbb{R}^m$ $\qquad J \quad [n\times m]$

$f : \mathbb{R}^m \to \mathbb{R}$ $\qquad \nabla_x f \quad [m\times 1]$

$$\nabla_x[f(g(x))] = \underset{[n\times m]}{J}\cdot \underset{[m\times 1]}{\nabla_x f}$$

Упражнения

$f(x) = x^T A x$ $\qquad \nabla_x f = (A + A^T) x \qquad \mathbb{R}^n \to \mathbb{R}$

$g(z) = \ln z$ $\qquad g'_z = \frac{1}{z} \qquad \mathbb{R} \to \mathbb{R}$

$$h(x) = g(f(x)) = \ln(x^T A x)$$

$$\nabla_x h = \underbrace{\frac{1}{x^T A x}}_{1 \times 1} \cdot (A + A^T) x$$

Ещё несколько производных

Упражнение

$\underset{[n \times n]}{f(X)} = X^{-1}$ $\qquad \underset{[n \times n]}{X} \in \mathbb{R}^{n \times n}$ $\qquad \nabla_X f$ – не смогу выписать в двумерном виде

$$df = dX^{-1} = ?$$

а) $X X^{-1} = I_n$

$$d(X X^{-1}) = dI_n = 0$$

$$dX \cdot X^{-1} + X \cdot dX^{-1} = 0$$

$$dX^{-1} = -X^{-1} dX \, X^{-1}$$

б) можно искать производные по определению.

$x^2 \qquad 2x \qquad \lim\limits_{h \to 0} \frac{(x+h)^2 - x^2}{h} = \lim\limits_{h \to 0} \frac{x^2 + 2xh + h^2 - x^2}{h} =$

$$= \lim_{h\to 0} \frac{2x+h}{1} = 2x \qquad df = f(X+H) - f(X) = \ldots$$

dX^{-1} – ?

$$dX^{-1} = \underbrace{(X+H)}_{X}{}^{-1} - X^{-1} = \left[(I_n + HX^{-1})X\right]^{-1} - X^{-1} =$$

$$= X^{-1}(I_n + HX^{-1})^{-1} - X^{-1} = *$$

геометрическая прогрессия

$(-HX^{-1})$

ряд Неймана: $(I_n - H)^{-1} = \sum_{k=0}^{\infty} H^k$ $\quad \max_i |\lambda_i| < 1$ для H

У нас $\forall i,j \; H_{ij} \to 0 \quad \|H\| \to 0 \quad \lambda_i \to 0$

$$H = U \Sigma U^T \qquad \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

$$(I_n + HX^{-1})^{-1} = I_n - HX^{-1} + \underbrace{(HX^{-1})^2 - \ldots}_{\text{Давайте забьем } H^2}$$

$o(\|H\|)$

$$* = X^{-1}(I_n - HX^{-1}) - X^{-1} = -X^{-1}HX^{-1}$$

$$dX^{-1} = -X^{-1}\,dX\,X^{-1}$$

Упражнение

$X \in \mathbb{R}^{n\times n} \qquad f(X) = tr(X) \qquad \nabla_X f = I_n$

$$d\ tr\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22}\end{pmatrix} = d(a_{11}+a_{22}) = da_{11} + da_{22}$$

$$tr\begin{pmatrix} da_{11} & da_{12} \\ da_{21} & da_{22}\end{pmatrix} = da_{11} + da_{22} \quad //$$

$\nabla_X^T f$

$$df = d\,tr(X) = tr(dX) = tr(I_n dX)$$

$A = I_n \cdot A \qquad 7 = 1 \cdot 7$

Упражнение

$X \in \mathbb{R}^{n\times n} \qquad f(X) = det(X)$

$f : \mathbb{R}^{n\times n} \to \mathbb{R}$

Нафига брать производную определителя?

Нормальное р.-е.

$$det\begin{pmatrix} \Delta & x & * \\ \Delta & x & * \\ \Delta & x & * \end{pmatrix} = \overbrace{\Delta}^{(-1)^{1+1}} \cdot \underbrace{det\begin{pmatrix} x & * \\ x & * \end{pmatrix}} + \overbrace{x}^{(-1)^{1+2}} \cdot \underbrace{det\begin{pmatrix} \Delta & * \\ \Delta & * \end{pmatrix}} + \underbrace{*}_{(-1)^{1+3}} \cdot det\begin{pmatrix} \Delta & x \\ \Delta & x \end{pmatrix}$$

Разложение по строке

$$\frac{\partial}{\partial x_{ij}}(det\, X) = \frac{\partial}{\partial x_{ij}}\left(\sum_k (-1)^{i+k} \cdot x_{ik} \cdot M_{ik}\right) = (-1)^{i+j} \cdot \underline{M_{ij}}$$

$$(A^{-1})_{ij} = \frac{1}{\det A} \underbrace{(-1)^{i+j} \cdot M_{ji}}_{\frac{\partial}{\partial x_{ji}}(\det X)}$$

$$\nabla_X \det X = \det X \cdot X^{-T}$$

$$d(\det X) = \operatorname{tr}(\underbrace{\det X \cdot X^{-1}}_{\nabla_X^T f} \cdot dX)$$