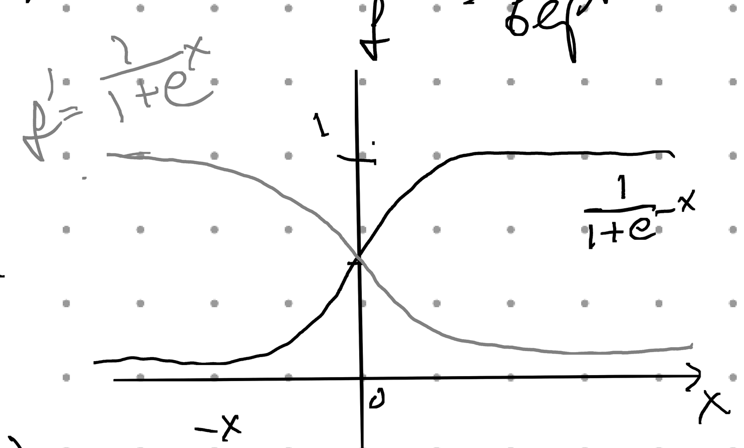
$\boxed{5.1}$ $f(x) = \log(x^4)\sin(x^3)$

$f'(x) = 4x^3 \cdot \frac{1}{x^4} \sin(x^3) + \log(x^4)\, 3x^2 \cos x^3 =$

$= \frac{4}{x}\sin(x^3) + \log(x)\, 12x^2 \cos x^3$

верно

$\boxed{5.2}$ $f(x) = \frac{1}{1+e^{-x}}$



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

Не верно

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

$= \left(\frac{e^x}{e^x} + e^x + \frac{1}{e^x} + 1\right)^{-1} = \frac{e^x}{e^{2x} + 2e^x + 1} = \frac{e^x}{(e^x+1)^2}$

верно

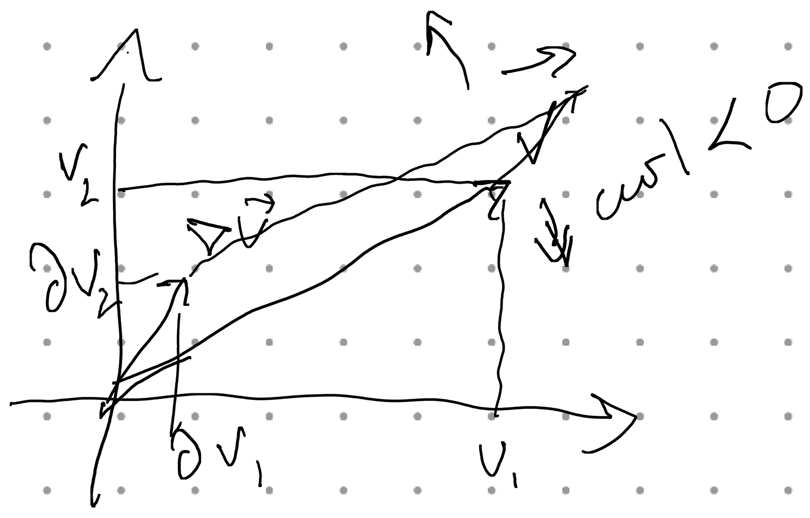
$\boxed{5.3}$ $f(x) = e^{-\frac{1}{2\sigma^2}(x-\mu)^2}$, $\mu, \sigma \in R$

$f' = \underbrace{e^{-\frac{1}{2\sigma^2}(x-\mu)^2}}_{f} \cdot -\frac{1}{\cancel{2}\sigma^2} \cdot \cancel{2}(x-\mu) = -\frac{x-\mu}{\sigma^2} f(x)$

верно

$\boxed{5.4}$ Taylor

Curl 2d $(\vec{v})$



$$\begin{vmatrix} \frac{1}{\partial x} & \partial v_1 \\ \frac{1}{\partial y} & \partial v_2 \end{vmatrix} =$$

$$= \frac{\partial v_2}{\partial x} - \frac{\partial v_1}{\partial y} =$$

$$= ?$$

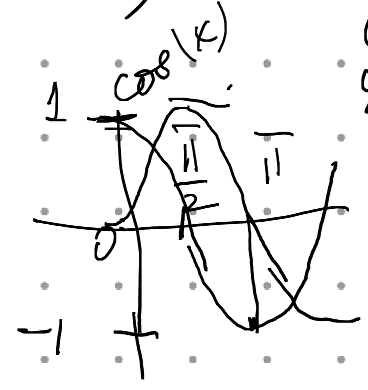
$\boxed{5.4}$ Taylor polynomials T_n

$f(x) = \sin(x) + \cos(x)$, $x_0 = 0$, $n = 0, \dots, 5$

$$T_0(x) = \frac{\cancel{\sin^{(0)}(x) + \cos^{(0)}(x)}\ \cancel{\sin(0) + \cos(0)}}{0!}(x-0)^0 = (I)$$

$\cos'(x) = -\sin(x)$

$\sin'(x) = \cos(x)$



$$(I) = \frac{0 + (1)}{1}(x)^0 = 1^0 = T_0(x)$$

$$T_1(x) = \frac{\sin'(0) + \cos'(0)}{1!}(x-0)^1 = T_0 + \frac{\cos(0) - \sin(0)}{1!}(x-1) = T_0 + x$$

$$T_2(x) = \frac{-\sin(0) - \cos(0)}{2!}(x)^2 = T_1 - \frac{x^2}{2}$$

$$T_3(x) = T_2 + \frac{-\cos(0) + \sin(0)}{3!}x^3 = T_2 - \frac{x^3}{3!}$$

$$T_4(x) = T_3 + \frac{\sin 0 + \cos 0}{4!}x^4 = T_3 + \frac{x^4}{4!}$$

$$T_5(x) = T_4 + \frac{\cos 0 - \sin 0}{5!}x^5 = T_4 + \frac{x^5}{5!}$$

$$T_6(x) = T_5 + \frac{-\sin 0 - \cos 0}{6!}x^6 = T_5 - \frac{x^6}{6!}$$

Верно

$\boxed{5.5}$

a) dim of $\dfrac{\partial f_i}{\partial \vec{x}}$

$f_1(\underset{\in \mathbb{R}^2}{\vec{x}}) = \sin(x_1)\cos(x_2) \in \mathbb{R}$

$$\frac{\partial f_1}{\partial \vec{x}} = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} \\ \dfrac{\partial f_1}{\partial x_2} \end{bmatrix}^T \in \mathbb{R}^{1\times 2} = \begin{bmatrix} \cos x_1 \cos x_2 \\ -\sin(x_1)\sin x_2 \end{bmatrix}^T$$

$$\frac{\partial f}{\partial x} = \frac{f(x + \partial x) - f(x)}{\partial x}$$

$f_2(\vec{x}, \vec{y}) = \overbrace{\underset{1\times n}{\vec{x}^T}\ \underset{n}{\vec{y}}}^{1} \in \mathbb{R}$, $\vec{x}, \vec{y} \in \mathbb{R}^n$

$$\mathbb{R}^n \ni \frac{\partial f_2}{\partial \vec{x}} = \left[\frac{\partial f_2(\vec{x}, \vec{y})}{\partial x_1} \ \dots \ \frac{\partial f_2(\vec{x}, \vec{y})}{\partial x_n}\right] = \cancel{\begin{bmatrix} y_1 + x_2y_2 + \dots + x_ny_n \\ x_1y_1 + y_2 + \dots + x_ny_n \\ \vdots \\ x_1y_1 + \dots + y_n \end{bmatrix}^T} = \vec{y}^T$$

$$\mathbb{R}^{n\times n} \ni f_3(\vec{x}) = \overbrace{\underset{n}{\vec{x}}\ \underset{1\times n}{\vec{x}^T}}^{n\times n}, \quad \vec{x} \in \mathbb{R}^n$$

Верно

$$\frac{\partial f_3}{\partial \vec{x}} = \begin{bmatrix} \dfrac{\partial(x_1x_1)}{\partial x_1} & \dots & \dfrac{\partial(x_1x_n)}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial(x_nx_1)}{\partial x_n} & \dots & \dfrac{\partial(x_nx_n)}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{n\times n} = \begin{bmatrix} 2x_1 & x_2 & \dots & x_n \\ x_1 & 2x_2 & \dots & x_n \\ x_1 & x_2 & \dots & x_n \\ \vdots & & & \\ x_1 & x_2 & \dots & 2x_n \end{bmatrix}$$

5.6 a) $f(\vec{t}) = \sin(\log(\underbrace{\vec{t}^T \vec{t}}_{\in R}))$, $\vec{t} \in R^D$, $f \in R$

$$\frac{\partial f}{\partial \vec{t}} = \begin{bmatrix} \frac{\partial f}{\partial t_1} \\ \vdots \\ \frac{\partial f}{\partial t_D} \end{bmatrix}^T \qquad \frac{\partial f}{\partial t_1} = \sin(\log(t_1^2 + \dots + t_D^2)) = \cos(\log(\vec{t}^T \vec{t})) \cdot \frac{1}{\vec{t}^T \vec{t}} \cdot 2t_1$$

Верно

$$\frac{\partial f}{\partial \vec{t}} = 2\cos(\log(\vec{t}^T \vec{t})) \cdot \frac{1}{t^T t} \cdot t^T$$

b) $\underbrace{g(X)}_{\in R} = tr(\underbrace{A\cancel{X}B}_{\in R^{D \times D}})$, $A \in R^{D \times E}$, $X \in R^{E \times F}$, $B \in R^{F \times D}$, $tr \in R$

$$\frac{\partial g}{\partial X} = \begin{bmatrix} \frac{\partial g}{\partial X_{1,1}} & \frac{\partial g}{\partial X_{1,2}} & \dots & \frac{\partial g}{\partial X_{E,F}} \end{bmatrix} \quad E \cdot D$$

$$\frac{\partial g}{\partial X} = \frac{\partial tr}{\partial AXB} \cdot \frac{\partial AXB}{\partial X} = \frac{\partial tr}{\partial AXB} \cdot \frac{\partial AXB}{\partial XB} \cdot \underbrace{\frac{\partial XB}{\partial X}}_{E \times F}$$

$$\frac{\partial XB}{\partial X}$$

$$X_{e,:} B_{:,d} = \sum_f X_{e,f} \cdot B_{f,d}$$

$$\frac{\partial (X_{e,:} B_{:,d})}{\partial X_{e,f}} = B_{f,d}$$

$$\frac{\partial XB}{\partial X} \in R^{(E \times D) \times (E \times F)} =$$

$$g(X) = tr(\underset{D \times E}{A}\ \underset{E \times F}{X}\ \underset{F \times D}{B}) = \sum_i (AXB)_{ii}$$

$$A_{2\times2} X_{2\times2} = \begin{bmatrix} a_{11}x_{11} + a_{12}x_{21} & a_{11}x_{12} + a_{12}x_{22} \\ a_{21}x_{11} + a_{22}x_{21} & a_{21}x_{12} + a_{22}x_{22} \end{bmatrix}$$

$a_{11}\ a_{12}$ $x_{11}\ x_{12}$

$a_{21}\ a_{22}$ $x_{21}\ x_{22}$

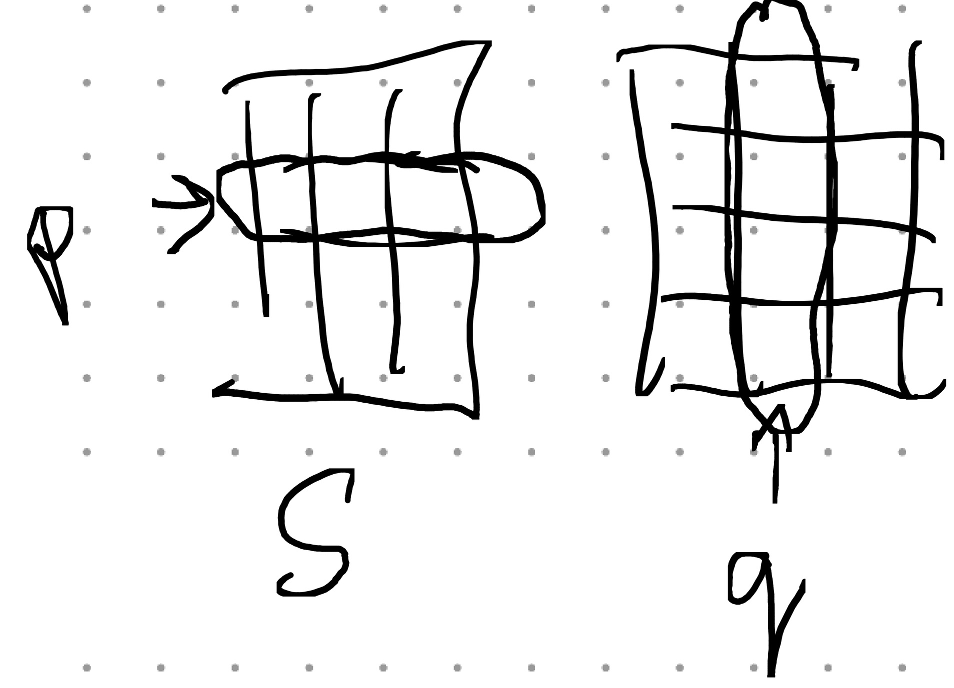
$$\frac{\partial AX}{\partial X} \in R^{(2\times2)\times(2\times2)} = \overbrace{\begin{bmatrix} \frac{\partial (AX)_{11}}{\partial X_{11}} & \dots & \frac{\partial (AX)_{11}}{\partial X_{22}} \\ \vdots & & \vdots \\ \frac{\partial (AX)_{22}}{\partial X_{11}} & \dots & \frac{\partial (AX)_{22}}{\partial X_{22}} \end{bmatrix}}^{2\times2} = \begin{bmatrix} a_{11} & 0 & a_{12} & 0 \\ 0 & a_{11} & 0 & a_{12} \\ a_{21} & 0 & a_{22} & 0 \\ 0 & a_{21} & 0 & a_{22} \end{bmatrix} =$$

2×2

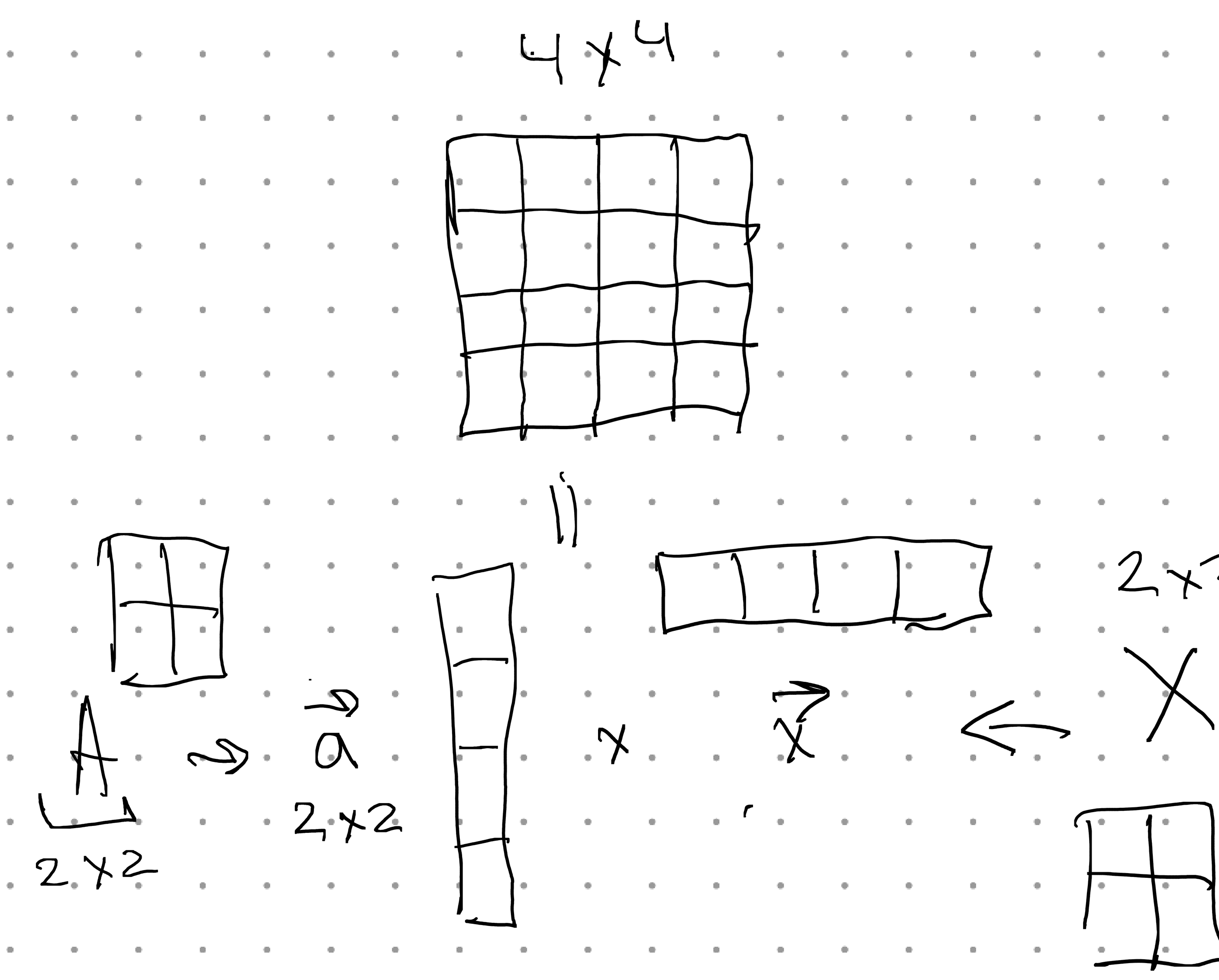
$$\frac{\partial \vec{x}^T \vec{a}}{\partial \vec{x}} = \begin{bmatrix} \frac{\partial x^T a}{\partial x_1} & \dots & \frac{\partial x^T a}{\partial x_n} \end{bmatrix} = [a_1 \dots a_n] = a^T$$

$$\frac{\partial a^T X b}{\partial X} = \frac{\partial a^T X b}{\partial a^T X} \cdot \frac{\partial a^T \cancel{X}}{\partial \cancel{X}} = \frac{\partial a^T X b}{\partial a^T}$$

$$\boxed{(ST)_{pq} = \sum_i S_{pi} T_{iq}}$$



$$\frac{\partial (ST)_{pq}}{\partial S_p} = T_q^T$$

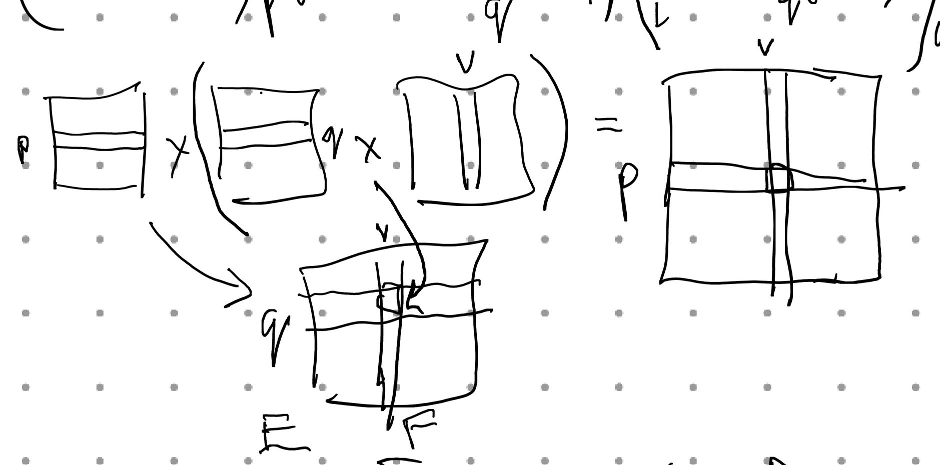


(continued)

$\boxed{5.6\,b}$ continued

$$\frac{\partial S_{p,:} T_{:,q}}{\partial S_p} = T_{:,q}^T$$

$$(A X B)_{pv} = \sum_{q}^{E} A_{pq} \left(\sum_{i}^{F} (X_{qi} B_{iv}) \right)_{qv} =$$



$$= \sum_{q}^{E} \sum_{i}^{F} A_{pq} X_{qi} B_{iv} \qquad p, v = 1..D$$

reshaped:

$$\frac{\partial (A X B)_{pv}}{\underset{E \times F}{\partial X}} = \begin{bmatrix} A_{p1} B_{1v} & \dots & A_{pE} B_{1v} \\ \vdots & & \vdots \\ A_{p1} B_{Fv} & \dots & A_{pE} B_{Fv} \end{bmatrix}$$

$$tr(A X B) = \sum_{p=1}^{D} \left(\sum_{q}^{E} \sum_{i}^{F} A_{pq} X_{qi} B_{ip} \right) \in R$$

diagonal

$$\frac{\partial\, tr(A X B)}{\partial X_{q,i}} = \sum_{p}^{D} A_{pq} B_{ip}$$

$$\frac{\partial\, tr(A X B)}{\underset{E \times F}{\partial X}} = \underset{E \times D}{A^T} \; \underset{D \times F}{B^T}$$

Верно

5.7

a) $f(z) = \overbrace{\log(1+z)}^{\in R}$, $\overset{\in R}{z = x^T x}$, $x \in R^D$

$$\frac{df}{dx} = \frac{\partial \log(1+z)}{\partial z} \cdot \frac{\partial z}{\partial x}$$

$$\frac{\partial(1+x^T x)}{\partial x} = 2x^T \in R^{1\times D}$$

$$\frac{\partial \log(1+z)}{1+z} = \frac{1}{1+x^T x} \in R$$

$$\boxed{\frac{\partial f}{\partial x} = \frac{2x^T}{1+x^T x} \in R^{1\times D}}$$ — Ответ

Верно

b) $f(z) = \sin(z)$, $\underset{\in R^E}{z = A\vec{x} + \vec{b}}$, $A \in R^{E\times D}$, $x \in R^D$, $b \in R^E$

$$\underbrace{\frac{\partial \sin(z)}{\partial z}}_{E\times E} \cdot \underbrace{\frac{\partial z}{\partial x}}_{E\times D} \qquad z_e = \sum_d (A_{ed} x_d + b_e)$$

$$\frac{\partial A\vec{x}+\vec{b}}{\partial \vec{x}} = A \Leftarrow \frac{\partial z_e}{\partial x_d} = A_{ed}$$

$$\frac{\partial \sin z}{\partial z} = \operatorname{diag}(\cos(z)) \in R^{E\times E}$$

$$\frac{df}{dx} = \operatorname{diag}(\cos(Ax+b)) \cdot A \in R^{E\times D}$$

Ответ.

Верно

5.8

a) $\overset{\in R^1}{f(z) = e^{-\frac{1}{2}z}}$

$z = g(\vec{y}) = \underset{1\times D}{\vec{y}^T} \underset{D\times D}{S^{-1}} \underset{D\times 1}{\vec{y}} \in R$ $\quad g(\vec{x}) = (x-\mu)^T S^{-1}(x-\mu)$

$y = h(\vec{x}) = \vec{x} - \vec{\mu} \in R^D$

$\vec{x}, \vec{\mu} \in R^D$, $S \in R^{D\times D}$

$$\underbrace{\frac{df}{dx}}_{1\times D} = \underbrace{\frac{\partial f}{\partial z}}_{1\times 1} \underbrace{\frac{\partial z}{\partial y}}_{1\times D} \underbrace{\frac{\partial y}{\partial x}}_{D\times D}$$

$$\frac{\partial y}{\partial x} = I_D \qquad \left(\frac{\partial y}{\partial x}\right)_{ij} = \frac{\partial(x_i - \mu_i)}{\partial x_j} = \begin{cases} 1, & i=j \\ 0, & i\neq j \end{cases}$$

$$\left(\frac{\partial z}{\partial y}\right)_i = \frac{\partial y_k \sum_k S^{-1}_{kp} y_p}{\partial y_i}$$

$$\frac{\partial z}{\partial y} = \underset{1\times D}{y^T}\underset{D\times D}{(S^{-1} + S^{-1^T})} = 2y^T S^{-1} \text{ ?}$$

Верно?

$$\frac{\partial f}{\partial z} = e^{-\frac{1}{2}z} \cdot \left(-\frac{1}{2}\right) =$$

$$\frac{\partial f}{\partial x} = e^{-\frac{1}{2}z} \cdot \left(-\frac{1}{2}\right) \cdot y^T \underbrace{(S^{-1} + S^{-1^T})}$$

Верно?

5.8 b

$$\underset{\in R}{f(x)} = tr(\underbrace{\vec{x}\vec{x}^T}_{D\times D} + \underbrace{\sigma^2 I}_{D\times D}), \quad x \in R^D$$

$$\begin{bmatrix} x_1 x_1 + \sigma^2 & & x_1 x_D \\ \vdots & & \vdots \\ x_D x_1 & \dots & x_D x_D + \sigma^2 \end{bmatrix}$$

$$\frac{\partial f}{\partial x_i} = \frac{\partial \sum_j (x_j x_j + \sigma^2)}{\partial x_i} =$$

$$= 0 + \dots + 0 + 2x_i + 0 + \dots + 0$$

$$\frac{\partial f}{\partial x} = 2x^T \in R^{1\times D}$$

Верно

5.8, c

$f = \tanh(z) \in R^M$

$\vec{z} = \underset{M\times N}{A} \underset{N}{\vec{x}} + \underset{M}{\vec{b}}$, $x \in R^N$, $A \in R^{M\times N}$, $b \in R^M$

$$\frac{\partial f}{\partial z} = \operatorname{diag}(1 - \tanh^2(z))$$

$$\left(\frac{\partial z}{\partial x}\right)_i = \frac{\partial \sum_n (A_{mn} \cdot x_n) + b_m}{\partial x_i} = A_{mj}$$

$$\frac{\partial z}{\partial x} = A$$

$$\frac{\partial f}{\partial x} = \operatorname{diag}(1 - \tanh^2(Ax+b)) \cdot A$$

Верно

$\boxed{5.9}$

$$\overset{D}{\vec{x}},\ \overset{E}{\vec{z}},\ \overset{F}{\vec{v}}$$

$$g(\vec{x},\vec{z},\vec{v}) := \log p(\vec{x},\vec{z}) - \log q(\vec{z},\vec{v}) \in R$$

$$\vec{z} := t(\overset{G}{\vec{\epsilon}},\overset{F}{\vec{v}}) \quad , p,q,t - \text{diff}, (\in R?)$$

$$\underline{\frac{dg}{d\vec{v}} = ?} \qquad \frac{dg}{d\vec{v}} \in R^{1\times F}$$

$$\frac{dg}{d\vec{v}} = \frac{1}{p(x,z)} \cdot \frac{dp}{dv} - \frac{1}{q(z,v)}\left(\frac{dq}{dv}\right)$$

$$\frac{dp}{dv} = \frac{dp}{dz}\underbrace{\frac{dz}{dt}\frac{dt}{dv}}$$

$$\frac{dq}{dv} = \begin{bmatrix} \frac{\partial q}{\partial z} & \frac{\partial q}{\partial v} \end{bmatrix} \begin{bmatrix} \frac{\partial z}{\partial v} \\ \frac{\partial v}{\partial v} \end{bmatrix} = \frac{\partial q}{\partial t}\underbrace{\frac{\partial t}{\partial v}} + \frac{\partial q}{\partial v}$$

$$\frac{dg}{dv} = \frac{1}{p}\frac{\partial p}{\partial t}\frac{\partial t}{\partial v} - \frac{1}{q}\left(\frac{\partial q}{\partial t}\frac{\partial t}{\partial v} + \frac{\partial q}{\partial v}\right) =$$

$$= \left(\frac{1}{p}\frac{\partial p}{\partial z} - \frac{1}{q}\frac{\partial q}{\partial z}\right)\frac{\partial t}{\partial v} + \frac{1}{\partial q}\frac{\partial q}{\partial v}$$

Вдерно