Задание №1:

Д-ть: $(A+UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}$

▶ $(A+UCV)\cdot(A+UCV)^{-1} = I$.

=> Проверим

$(A+UCV)(A^{-1}-A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}) =$

$= I + UCVA^{-1} - (U + UCVA^{-1}U)(C^{-1}+VA^{-1}U)^{-1}VA^{-1} =$

$= I + UCVA^{-1} - UC(C^{-1}+VA^{-1}U)(C^{-1}+VA^{-1}U)^{-1}VA^{-1} =$

$= I + UCVA^{-1} - UCVA^{-1} = I$

=> Верно

■

Задание 2а

Упростить: $\|uv^T - A\|_F^2 - \|A\|_F^2 = f$

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

- $uv^T = \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} (v_1 \dots v_n) = \begin{pmatrix} u_1v_1 & \dots & u_1v_n \\ \vdots & \ddots & \vdots \\ u_mv_1 & \dots & u_mv_n \end{pmatrix}$

- $uv^T - A = \begin{pmatrix} u_1v_1 - a_{11} & u_1v_2 - a_{12} & \dots & u_1v_n - a_{1n} \\ u_2v_1 - a_{21} & u_2v_2 - a_{22} & \dots & u_2v_n - a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ u_mv_1 - a_{m1} & u_mv_2 - a_{m2} & \dots & u_nv_n - a_{mn} \end{pmatrix}$

- $\|X\|_F^2 = tr(X^TX)$ — сумма кв. столбцов матрицы

$\Rightarrow \|uv^T - A\|_F^2 = \sum_{i=1}^{n} \sum_{j=1}^{m} (u_jv_i - a_{ji})^2$

- $\|A_F^2\| = \sum_{i=1}^{n} \sum_{j=1}^{m} (a_{ji})^2$

$\Rightarrow f = \sum_{i=1}^{n} \sum_{j=1}^{m} (u_jv_i)^2 - 2u_jv_ia_{ji} + (a_{ji})^2 - (a_{ji})^2 =$

$= \sum_{i=1}^{n} \sum_{j=1}^{m} \left((u_jv_i)^2 - 2u_jv_ia_{ji}\right)$

Задача 2b $\quad f = tr((2I_n + aa^T)^{-1}(uv^T + vu^T))$

- $(2I_n + aa^T)^{-1} = \{$ по тожд. Вудберри, при $A = 2I_n,\ U = a,\ V = a^T,\ C = I_n \} =$

$= \frac{1}{2} I_n - \frac{1}{2} I_n \cdot a (I_n + \frac{1}{2} a^T a)^{-1} a^T \cdot \frac{1}{2} I_n =$

$= \frac{1}{2} I_n - \frac{1}{4} a (I_n + \frac{1}{2} \|a\|^2)^{-1} a^T = \{ (I_n + \frac{1}{2}\|a\|^2) = const$

$\Rightarrow \exists \quad c = I_n + \frac{1}{2}\|a\|^2 \} = \frac{1}{2} I_n - \frac{1}{4c} aa^T$

- $(\frac{1}{2} I_n - \frac{1}{4c} aa^T)(uv^T + vu^T) =$

$= \frac{1}{2}(uv^T + vu^T) - \frac{1}{4c}(aa^T(uv^T + vu^T))$ (*)

- $tr(\frac{1}{2}(uv^T + vu^T)) = \frac{1}{2}(tr(uv^T) + tr(vu^T)) = \{ tr(uv^T) = u_0 v_0 + u_1 v_1 + \ldots + u_{n-1} v_{n-1} = tr(vu^T) \} = tr(uv^T) = \langle u, v \rangle$

- $tr(aa^T(uv^T + vu^T)) = \{$ т.к. aa^T – симметрична, то

$tr(aa^T \cdot uv^T) = tr(A \cdot B) = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ji} b_{ij} = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ki} b_{ki} = [a_{ik} = a_{ki}] = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ki} b_{ki}$

$tr(aa^T \cdot vu^T) = tr(A \cdot B^T) = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ji} b_{ji} = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ki} b_{ki} \quad \} =$

$= 2\,tr(aa^T \cdot uv^T) = 2\,tr(v^T a a^T u) = 2 \langle v, a \rangle \cdot \langle a, u \rangle$

(*) $\Rightarrow f = \langle u, v \rangle - \frac{1}{2}(I_n + \frac{1}{2}\|a\|^2)^{-1} \cdot \langle v, a \rangle \cdot \langle a, u \rangle$

∎

Задание №2с

$f = \sum_{i=1}^{n} \langle S^{-1} a_i, a_i \rangle$, где

$a_1, \ldots, a_n \in \mathbb{R}^n$, $S = \sum_{i=1}^{n} a_i a_i^T$

$$\sum_{i=1}^{n} \langle S^{-1} a_i, a_i \rangle = \{ \langle A, B \rangle = tr(A^T B) \} =$$

$$= \sum_{i=1}^{n} tr(a_i^T S^{-T} a_i) = \{ tr(AB) = tr(BA),\ A = a_i^T S^{-T} \} =$$

$$= \sum_{i=1}^{n} tr(a_i a_i^T S^{-T}) = tr\left(\sum_{i=1}^{n} a_i a_i^T S^{-T}\right) = tr\left(\left(\sum_{i=1}^{n} (a_i a_i^T)\right) \cdot S^{-T}\right) =$$

$$= tr(S \cdot S^{-T}) \underset{\uparrow}{=} tr(I_n) = n$$

по построению S – симметричная м-ца $\Rightarrow S^{-1}$ – симметрична $\Rightarrow S^{-T} = S^{-1}$ $\Rightarrow S \cdot S^{-1} = I_n$

Ответ: $f = n$

$\blacksquare$

Задача №3а $f(t) = \det(A - tI_n)$

• Первая производная:

$\exists\ f(t) = f(g(t))\ ;\ y = g(t) = A - tI_n$

$df(t)[dt] = \det(y) \cdot \langle y^{-T}, dy \rangle \circeq$

$dy = d(A - tI_n) = -I_n dt$

$\circeq -\det(A - tI_n) \langle (A - tI_n)^{-T}, I_n dt \rangle =$

$= -\det(A - tI_n) \cdot tr((A - tI_n)^{-1}) dt$

• Вторая производная:

$d^2 f(t)[dt_1, dt_2] = -d(\det(A - tI_n) \cdot tr((A - tI_n)^{-1}) dt_1) =$

$= -(d(\det(A - tI_n)) \cdot tr((A - tI_n)^{-1}) + \det(A - tI_n) \cdot d(tr((A - tI_n)^{-1}))) dt_1 \circeq$

- $d(\det(A - tI_n)) = -\det(A - tI_n) tr((A - tI_n)^{-1}) dt_2$
- $d(tr((A - tI_n)^{-1})) = tr(d(A - tI_n)^{-1}) = -tr((A - tI_n)^{-1} \cdot (-dt_2 I_n)(A - tI_n)^{-1}) =$
  $= tr(((A - tI_n)^{-1})^2) dt_2$

$\circeq (\det(A - tI_n) \cdot tr^2((A - tI_n)^{-1}) - \det(A - tI_n) \cdot tr(((A - tI_n)^{-1})^2)) dt_1 dt_2 =$

$= \det(A - tI_n) \cdot (tr^2((A - tI_n)^{-1}) - tr(((A - tI_n)^{-1})^2)) dt_1 dt_2$

Ответ: ~~dt~~

$df = -\det(A - tI_n) \cdot tr((A - tI_n)^{-1}) dt_1$

$d^2 f = \det(A - tI_n) \cdot [tr^2((A - tI_n)^{-1}) - tr(((A - tI_n)^{-1})^2)] dt_1 dt_2$

$\Downarrow$

$f'(t) = -\det(A - tI_n) \cdot tr((A - tI_n)^{-1})$

$f''(t) = \det(A - tI_n) \cdot [tr^2((A - tI_n)^{-1}) - tr(((A - tI_n)^{-1})^2)]$

Задача 3b $f(t) = \|(A+tI_n)^{-1}B\| = \left(((A+tI_n)^{-1}B)^T((A+tI_n)^{-1}B)\right)^{1/2}$

• Первая производная: $\exists\ f(t) = f(g(h(t)))$

$$z = h(t) = (A+tI_n)^{-1}B$$

$$y = g(z) = g(h(t)) = z^Tz$$

$$f(y) = y^{1/2}$$

$$\Rightarrow df(t)[dt] = dy^{1/2} = \frac{1}{2y^{1/2}}dy = \frac{1}{2f(t)}dz^Tz =$$

$$= \frac{1}{2f(t)}\cdot 2z^Tdz = \frac{1}{f(t)}\cdot((A+tI_n)^{-1}B)^Td((A+tI_n)^{-1}B) = \{(A+tI_n)^{-T} = (A+tI_n)^{-1}\}$$

$$= \frac{-B^T(A+tI_n)^{-1}}{f(t)}\cdot(A+tI_n)^{-1}d(A+tI_n)(A+tI_n)^{-1}B =$$

$$= \frac{-1}{f(t)}B^T((A+tI_n)^{-1})^2\cdot I_n dt\cdot(A+tI_n)^{-1}B = \frac{-1}{f(t)}\cdot B^T((A+tI_n)^{-1})^3B\,dt$$

• Вторая производная:

$$d^2f(t)[dt_1,dt_2] = -d\left(\frac{B^T((A+tI_n)^{-1})^3B}{f(t)}\right)dt_1 =$$

$$= -\frac{d(B^T((A+tI_n)^{-1})^3B)f(t) - (B^T((A+tI_n)^{-1})^3B)\,df(t)}{f^2(t)}dt_1 \ (=)$$

$$\bullet\ d(B^T((A+tI_n)^{-1})^3B) = B^Td((A+tI_n)^{-1})^3B = 3B^T((A+tI_n)^{-1})^2d(A+tI_n)^{-1}B =$$

$$= -3B^T((A+tI_n)^{-1})^3d(A+tI_n)\cdot(A+tI_n)^{-1}B = -3B^T((A+tI_n)^{-1})^4B\,dt_2$$

$$\bullet\ df(t) = \frac{-1}{f(t)}B^T((A+tI_n)^{-1})^3B\,dt_2$$

$$(=)\ \frac{3B^T((A+tI_n)^{-1})^4B\,dt_1dt_2}{f(t)} - \frac{(B^T((A+tI_n)^{-1})^3B)^2dt_1dt_2}{f^3(t)}$$

Ответ: $df = \frac{-1}{\|(A+tI_n)^{-1}B\|}\cdot B^T((A+tI_n)^{-1})^3B\,dt_1$

$$d^2f = \left(\frac{3B^T((A+tI_n)^{-1})^4B}{\|(A+tI_n)^{-1}B\|} - \frac{(B^T((A+tI_n)^{-1})^3B)^2}{\|(A+tI_n)^{-1}B\|^3}\right)dt_1dt_2$$

$$f'(t) = \frac{-1}{\|(A+tI_n)^{-1}B\|}\cdot B^T((A+tI_n)^{-1})^3B$$

$$f''(t) = \frac{3B^T((A+tI_n)^{-1})^4B}{\|(A+tI_n)^{-1}B\|} - \frac{(B^T((A+tI_n)^{-1})^3B)^2}{\|(A+tI_n)^{-1}B\|^3}$$

Задача №4а $\quad f(x) = \frac{1}{2}\|xx^T - A\|_F^2, \quad A \in S^n$

$$f(x) = \frac{1}{2}\|xx^T - A\|_F^2 = \frac{1}{2}\mathrm{tr}\left((xx^T - A)^T(xx^T - A)\right) = \{(xx^T - A)^T = (xx^T - A), \text{ т.к. } xx^T \in S^n\} =$$

$$= \frac{1}{2}\mathrm{tr}\left((xx^T - A)^2\right)$$

• Первая производная:

$$df(x)[dx] = \frac{1}{2}d\,\mathrm{tr}\left((xx^T - A)^2\right) = \frac{1}{2}\mathrm{tr}\left(d(xx^T - A)^2\right) = \mathrm{tr}\left((xx^T - A)\,d\,xx^T\right) =$$

$$= 2\mathrm{tr}\left((xx^T - A)x\,dx^T\right) = 2\mathrm{tr}\left(((xx^T - A)\cdot x)dx^T\right) \overset{\text{трансп. обе части и сдвигаем}}{=} 2\mathrm{tr}\left(x^T(xx^T - A)^T dx\right) =$$

$$= 2\langle (xx^T - A)x, dx\rangle$$

$$\Rightarrow \nabla f = 2(xx^T - A)x$$

• Вторая производная:

$$d^2f(x)[dx_1, dx_2] = 2d\left(\mathrm{tr}\left(x^T(xx^T - A)^T dx_1\right)\right) = 2\mathrm{tr}\left(d(x^T(xx^T - A)^T)dx_1\right) =$$

$$= 2\mathrm{tr}\left((dx_2^T\cdot(xx^T - A)^T + x^T\,dxx^T)dx_1\right) \cancel{= 2\mathrm{tr}\left((dx_2^T\cdot(xx^T - A) + x^T\cdot 2(dx_2\cdot x^T\right.}$$

$$= 2\mathrm{tr}\left((dx_2^T\cdot(xx^T - A) + x^T(dx_2\,x^T + x\,dx_2^T))dx_1\right) =$$

$$= 2\mathrm{tr}\left(dx_2^T(xx^T - A)dx_1\right) + 2\mathrm{tr}\left(x^T dx_2 x^T dx_1\right) + 2\mathrm{tr}\left(x^T x\, dx_2^T dx_1\right) =$$

$$\{\text{трансп}\} = 2\mathrm{tr}\left(dx_2 dx_1^T(xx^T - A)\right) + 2\mathrm{tr}\left(dx_1^T x\cdot x^T dx_2\right) + 2\mathrm{tr}\left(dx_1^T dx_2 I x^T x\right) =$$

$$= 2\langle (xx^T - A)dx_1, dx_2\rangle + 2\langle xx^T dx_1, dx_2\rangle + 2\langle x^T x\cdot I_n dx_1, dx_2\rangle =$$

$$= 2\langle ((xx^T - A) + xx^T + x^T x\cdot I_n)dx_1, dx_2\rangle$$

$$\Rightarrow \nabla^2 f = 2\left((xx^T - A) + xx^T + x^T x\cdot I_n\right)$$

Ответ: $\quad \nabla f = 2(xx^T - A)x$

$\qquad\quad \nabla^2 f = 2\left((xx^T - A) + xx^T + x^T x\cdot I_n\right)$

Задача №46 $f(x) = \langle x,x\rangle^{\langle x,x\rangle}$

$$f(x) = x^Tx^{x^Tx} = \left(e^{\ln x^Tx}\right)^{x^Tx} = e^{x^Tx\ln(x^Tx)}$$

• Первая производная:

$$df(x) = de^{x^Tx\ln(x^Tx)} = e^{x^Tx\ln(x^Tx)} d(x^Tx\ln(x^Tx)) =$$

$$= e^{x^Tx\ln(x^Tx)} \cdot (d(x^Tx)\cdot\ln(x^Tx) + x^Tx\cdot d\ln(x^Tx)) =$$

$$= e^{x^Tx\ln(x^Tx)}\cdot\left(2x^Tdx\cdot\ln(x^Tx) + x^Tx\cdot\frac{2x^Tdx}{x^Tx}\right) = e^{x^Tx\ln(x^Tx)}\cdot(2x^Tdx\,(\ln(x^Tx)+1)) =$$

$$= 2(\ln(x^Tx+1)\, e^{x^Tx\ln(x^Tx)}\, x^Tdx$$

$$\Rightarrow \nabla f = 2\langle x,x\rangle^{\langle x,x\rangle}(\ln(x^Tx)+1)\,x$$

• Вторая производная:

$$d^2f[dx_1,dx_2] = 2d\left(\langle x,x\rangle^{\langle x,x\rangle}(\ln(x^Tx)+1)\,x^Tdx_1\right) =$$

$$= 2\left(d\langle x,x\rangle^{\langle x,x\rangle}\cdot(\ln(x^Tx)+1)\,x^T + \langle x,x\rangle^{\langle x,x\rangle} d((\ln(x^Tx)+1)\,x^T)\right)dx_1 \ominus$$

$$\cdot\, d\langle x,x\rangle^{\langle x,x\rangle} = 2\langle x,x\rangle^{\langle x,x\rangle}(\ln(x^Tx)+1)\,x^Tdx_2$$

$$\cdot\, d((\ln(x^Tx)+1)\,x^T) = d(\ln(x^Tx)+1)\cdot x^T + (\ln(x^Tx)+1)\,dx_2^T =$$

$$= \frac{2x^Tdx}{x^Tx}\cdot x^T + \cancel{2}\,(\ln(x^Tx)+1)\,dx_2^T$$

$$\ominus\; 2\langle x,x\rangle^{\langle x,x\rangle}\left(2(\ln(x^Tx)+1)^2 x^Tdx_2\, x^Tdx_1 + 2\frac{x^Tdx_2}{x^Tx}x^T + (\ln(x^Tx)+1)\,dx_2^T\right)dx_1 =$$

= {транспонируем выражение, т.к. оно из $\mathbb{R}$} =

$$= dx_1^T\langle x,x\rangle^{\langle x,x\rangle}\left(4(\ln(x^Tx)+1)^2 x(dx_2^Tx) + 4\frac{1}{x^Tx}\cdot x(dx_2^Tx) + 2(\ln(x^Tx)+1)\,dx\cdot I_n\cdot dx_1\right) =$$

тракт. то, что в каждой скобках

$$= dx_1^T\langle x,x\rangle^{\langle x,x\rangle}\left(4(\ln(x^Tx)+1)^2 xx^Tdx_2 + 4\frac{1}{x^Tx}xx^Tdx_2 + 2(\ln(x^Tx)+1)\cdot I_n\cdot dx_1\right) =$$

$$= dx_1^T\langle x,x\rangle^{\langle x,x\rangle}\left(4(\ln(x^Tx)+1)^2 xx^T + 4\frac{1}{x^Tx}xx^T + 2(\ln(x^Tx)+1)\cdot I_n\right)dx_2$$

$$\nabla^2 f = \langle x,x\rangle^{\langle x,x\rangle}\left(4(\ln(x^Tx)+1)^2xx^T + 4\frac{1}{x^Tx}xx^T + 2(\ln(x^Tx)+1)\,I_n\right)$$

Ответ: $\nabla f = 2\langle x,x\rangle^{\langle x,x\rangle}(\ln(x^Tx)+1)\,x$

$$\nabla^2 f = \langle x,x\rangle^{\langle x,x\rangle}\left(4(\ln(x^Tx)+1)^2xx^T + 4\frac{1}{x^Tx}xx^T + 2(\ln(x^Tx)+1)\,I_n\right)$$

Задача №4с $f(x) = \|Ax - b\|^p$, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $p \geq 2$

$$f(x) = ((Ax-b)^T(Ax-b))^{p/2}$$

• Первая производная:

$$df(x)[dx] = d((Ax-b)^T(Ax-b))^{\frac{p}{2}} = \frac{p}{2}((Ax-b)^T(Ax-b))^{\frac{p}{2}-1} d((Ax-b)^T(Ax-b)) =$$

$$= \frac{p}{2}((Ax-b)^T(Ax-b))^{\frac{p}{2}-1} \cdot 2(Ax-b)^T d(Ax-b) =$$

$$= \frac{p}{2}((Ax-b)^T(Ax-b))^{\frac{p}{2}-1} \cdot 2(Ax-b)^T A\,dx = p((Ax-b)^T(Ax-b))^{\frac{p}{2}-1}(Ax-b)^T A\,dx$$

$$\Rightarrow \nabla f = p((Ax-b)^T(Ax-b))^{\frac{p}{2}-1} \cdot A^T(Ax-b)$$

• Вторая производная:

$$d^2f(x)[dx_1, dx_2] = dp((Ax-b)^T(Ax-b))^{\frac{p}{2}-1}(Ax-b)^T A\,dx_1 =$$

$$= p\left(d((Ax-b)^T(Ax-b))^{\frac{p}{2}-1} \cdot (Ax-b)A + ((Ax-b)^T(Ax-b))^{\frac{p}{2}-1} d(Ax-b)^T \cdot A\right) dx_1 =$$

$$\cdot\ d((Ax-b)^T(Ax-b))^{\frac{p}{2}-1} = \left(\frac{p}{2}-1\right) \cdot ((Ax-b)^T(Ax-b))^{\frac{p}{2}-2} \cdot 2(Ax-b)^T A\,dx_2$$

$$d(Ax-b)^T = (d(Ax-b))^T = (A\,dx)^T = dx_2^T A^T$$

$$= p\cdot\left(\frac{p}{2}-1\right) \cdot ((Ax-b)^T(Ax-b))^{\frac{p}{2}-2} \cdot 2 \cdot (Ax-b)^T A\,dx_2 \cdot (Ax-b)^T A\,dx_1 + p((Ax-b)^T(Ax-b))^{\frac{p}{2}-1} dx_2^T A^T A\,dx_1$$

$$= \{1 \text{ слагаемое: } (Ax-b)^T A\,dx_1 = dx_1^T A^T(Ax-b);\ dx_2^T \cdot A^T A\,dx_1 = dx_1^T A^T A\,dx_2\} =$$

$$= dx_1^T\left((p-2)p \cdot A^T(Ax-b) \cdot [(Ax-b)^T(Ax-b)]^{\frac{p}{2}-2}(Ax-b)^T A + pA^T A((Ax-b)^T(Ax-b))^{\frac{p}{2}-1}\right) dx_2$$

$$\Rightarrow \nabla^2 f = (p-2)pA^T(Ax-b) \cdot [(Ax-b)^T(Ax-b)]^{\frac{p}{2}-2}(Ax-b)^T A + pA^T A[(Ax-b)^T(Ax-b)]^{\frac{p}{2}-1}$$

Ответ: $\nabla f = p[(Ax-b)^T(Ax-b)]^{\frac{p}{2}-1} \cdot A^T(Ax-b)$

$$\nabla^2 f = (p-2)pA^T(Ax-b)[(Ax-b)^T(Ax-b)]^{\frac{p}{2}-2}(Ax-b)^T A + pA^T A[(Ax-b)^T(Ax-b)]^{\frac{p}{2}-1}$$

Задача №5а $f(X) = tr(X^{-1})$

• Первая производная:

$$df(X)[dX] = d(tr(X^{-1})) = tr(d(X^{-1})) = tr(-X^{-1}dX X^{-1}) =$$

$$= \cancel{tr} - tr((X^{-1})^2 dX)$$

• Вторая производная: $dX = dX_1 = const$

$$d^2f(X)[dX, dX] = -d(tr((X^{-1})^2 dX_1)) = -tr(d(X^{-1})^2 dX_1) =$$

$$= -tr(2X^{-1}d(X^{-1})dX_1) = -tr(-2X^{-1} \cdot X^{-1}dX_2 \cdot X^{-1}dX_1) =$$

$$= 2tr((X^{-1})^2 dX_2 \cdot X^{-1} \cdot dX_1) = 2tr(X^{-1}dX_2 X^{-1}dX_1 X^{-1})$$

• Подставим приращения:

$$d^2f[H,H] = 2tr(X^{-1}HX^{-1}HX^{-1})$$

1) Если $X \in S^n_{++}$, $\Rightarrow X^{-1} \in S^n_{++}$

2) т.к. $X \in S^n$, то $X^{-1} \in S^n$, $\Rightarrow (X^{-1})^T = X^{-1}$

3) По критерию положит. опред. матрицы:

$\exists B : det(B) \neq 0$, и $X^{-1} = B^T B$

$\Rightarrow d^2f[H,H] = 2tr(X^{-T}HB^T \cdot BHX^{-1}) =$

$\exists D = X^{-T}HB^T$

$= 2tr(DD^T) = 2\|D\|_F^2$

$\Rightarrow$ по определению нормы

$d^2f[H,H] \geq 0$, $\forall D$ ($=0$ при $D=0$)

$\Rightarrow d^2f[H,H]$ сохраняет пост. знак

Задача №5б $f(X) = (\det X)^{1/n}$

• Первая производная: $\exists\ f(X) = f(g(X))$, $g(X) = y = \det X$

$$df(y)[dy] = dy^{1/n} = \frac{1}{n\cdot y^{\frac{n-1}{n}}}\,dy = \frac{d(\det X)}{n\cdot(\det X)^{\frac{n-1}{n}}} = \frac{\det X \cdot tr(X^{-1}dX)}{n\,(\det X)^{\frac{n-1}{n}}} =$$

$$= \frac{1}{n}(\det X)^{\frac{1}{n}} \cdot tr(X^{-1}dX)$$

• Вторая производная $\exists\ dX = dX_1$,

$$d^2f(X)[dX_1, dX_2] = \frac{1}{n}d\left((\det X)^{\frac{1}{n}} tr(X^{-1}dX_1)\right) = \frac{1}{n}\left(d(\det X)^{1/n} \cdot tr(X^{-1}dX_1) +\right.$$

$$\left.+ (\det X)^{1/n} \cdot d\,tr(X^{-1}dX_1)\right) = \frac{1}{n^2}(\det X)^{\frac{1}{n}} tr(X^{-1}dX_2)\,tr(X^{-1}dX_1) + \cancel{(\det X)^{\frac{1}{n}}}$$

$$+ \frac{1}{n}(\det X)^{1/n} tr(dX^{-1}dX_1) = \frac{1}{n^2}(\det X)^{\frac{1}{n}} tr(X^{-1}dX_2)\,tr(X^{-1}dX_1)$$

$$+ \frac{1}{n}(\det X)^{1/n} \cdot tr(-X^{-1}dX_2X^{-1}dX_1) =$$

$$= \frac{1}{n^2}(\det X)^{1/n} tr(X^{-1}dX_2)\,tr(X^{-1}dX_1) - \frac{1}{n}(\det X)^{1/n} tr(X^{-1}dX_2X^{-1}dX_1) =$$

$$= \frac{1}{n^2}(\det X)^{1/n}\left(tr(X^{-1}dX_2)\cdot tr(X^{-1}dX_1) - n\,tr(X^{-1}dX_2X^{-1}dX_1)\right)$$

• Подставим приращения:

$$d^2f(X)[H,H] = \frac{1}{n^2}(\det X)^{1/n}\left(tr^2(X^{-1}H) - n\,tr(X^{-1}HX^{-1}H)\right)$$

1) т.к. $X \in S^n_{++} \Rightarrow \det X > 0 \Rightarrow \frac{1}{n^2}(\det X)^{1/n} > 0$

2) Рассм. $tr^2(X^{-1}H) - n\,tr(X^{-1}H\,X^{-1}H) =$

$= \cancel{\ldots}$ 1) $X^{-T} = X^{-1}$, $H^T = H$ – в силу симметр.

2) т.к. и X^{-1}, и H – симм. $\Rightarrow X^{-1}H = HX^{-1}$

$= tr^2(X^{-1}H\,I_n) - n\,tr(X^{-1}H\,X^{-1}H) = \langle X^{-1}H, I\rangle^2 - n\langle X^{-1}H, X^{-1}H\rangle$

~~3) Сравним эти значения:~~ $\cancel{\exists\ A = X^{-1}H}$

~~(в силу симм. X^{-1} и H, будем менять порядок индексов на удоб.)~~

$$\cancel{A = \begin{pmatrix} \sum_{i=1}^n x_{1i}h_{1i} & \sum_{i=1}^n x_{2i}h_{2i} & - & \sum_{i=1}^n x_{1i}h_{ni} \\ \sum_{i=1}^n x_{2i}h_{1i} & \sum_{i=1}^n x_{2i}h_{2i} & - & \sum_{i=1}^n x_{2i}h_{ni} \\ & \ldots & & \end{pmatrix} = (a_{ij})}$$

$$\cancel{\Rightarrow (tr(A))^2 = \left(\sum_{i=1}^n a_{ii}\right)^2 = (a_{11} + \ldots + a_{nn})^2.}$$

~~Распишем лишь по~~ $\cancel{\frac{1}{k^{k-1}}}$ ~~сл-ду:~~ $\cancel{a_{11}^2 + 2a_{11}a_{21} + \ldots}$ ~~(все попар. попр.)~~

в силу н-ва Коши-Буняковск. $\langle X^{-1}H, I_n\rangle^2 \leq \|X^{-1}H\|^2 = \langle X^{-1}H, X^{-1}H\rangle$

$\Rightarrow \langle X^{-1}H, I_n\rangle^2 - n\langle X^{-1}H, X^{-1}H\rangle < 0$ при $n > 0$,

и м.б. $= 0$ при $n = 1$

Ответ: сохраняет знак „–" при $n > 1$,
0 при $n = 1$

Задача №6а $f(x) = \langle c, x\rangle + \frac{\sigma}{3}\|x\|^3$, $c \in \mathbb{R}^n$, $c \neq 0$, $\sigma > 0$

$$f(x) = \langle c, x\rangle + \frac{\sigma}{3}\|x\|^3 = c^T x + \frac{\sigma}{3}\|x\|^3$$

• Первая производная:

$$df(x) = d\left(c^T x + \frac{\sigma}{3}\|x\|^3\right) = d(c^T x) + \frac{\sigma}{3} d\|x\|^3 =$$

$$= c^T dx + \frac{\sigma}{3} d\|x\|^3 \ominus$$

$$d\|x\|^3 = d(x^T x)^{3/2} = \frac{3}{2}(x^T x)^{1/2} d x^T x = \frac{3}{2}\|x\| \cdot 2x^T dx = 3\|x\| x^T dx$$

$$\ominus\; c^T dx + \sigma\|x\| x^T dx = (c^T + \sigma\|x\| x^T) dx$$

$$\Rightarrow \nabla f = (c + \sigma\|x\| x)$$

• Точки стационарности: $c + \sigma\|x\| x = 0$

$$x = \frac{-c}{\sigma\|x\|}$$

$\Rightarrow \exists$ при $\|x\| \neq 0$

Задача №6 $f(x) = \langle a,x \rangle - \ln(1-\langle b,x \rangle)$, $a,b \in \mathbb{R}^n \neq 0$, $\langle b,x \rangle < 1$

$$df(x)[dx] = \langle a,dx \rangle - \frac{d(1-\langle b,x \rangle)}{1-\langle b,x \rangle} = \langle a,dx \rangle + \frac{\langle b,dx \rangle}{1-\langle b,x \rangle} =$$

$$= \langle a + \frac{b}{1-\langle b,x \rangle}, dx \rangle$$

$$\Rightarrow \nabla f = a + \frac{b}{1-\langle b,x \rangle}$$

$$\Rightarrow a + \frac{b}{1-\langle b,x \rangle} = \vec{0} \qquad | \cdot (1-\langle b,x \rangle)$$

$$a(1-\langle b,x \rangle) + b = \vec{0}$$

$$a + b = a\langle b,x \rangle$$

$$ab^T x = a + b$$

• По теореме кронекера-капели, решение $x\ \exists$

$\Leftrightarrow$ $rk(ab^T) = rk(ab^T | (a+b))$

$\Rightarrow$ Ответом ~~р~~ явл-ся решение $ab^T \cdot x = a + b$

– СЛАУ

$\boxed{\text{Задача №6c}}$ $f(x) = \langle c, x\rangle \exp(-\langle Ax, x\rangle)$, $c \in \mathbb{R}^n \neq 0$, $A \in S^n_{++}$

• Первая производная:

$$df(x)[dx] = \langle c, dx\rangle \exp(-\langle Ax,x\rangle) + \langle c,x\rangle \exp(-\langle Ax,x\rangle)\cdot d(-\langle Ax,x\rangle) =$$

$$= \exp(-\langle Ax,x\rangle)\cdot(\langle c,dx\rangle - 2\langle c,x\rangle\cdot\langle Ax,dx\rangle) =$$

$$= \exp(-\langle Ax,x\rangle)\cdot(\langle c,dx\rangle - \langle 2\langle c,x\rangle Ax, dx\rangle) =$$

$$= \exp(-\langle Ax,x\rangle)\cdot\langle c - 2\langle c,x\rangle Ax, dx\rangle = \langle \exp(-\langle Ax,x\rangle)\cdot(c - 2\langle c,x\rangle Ax), dx\rangle$$

$$\Rightarrow \nabla f = \exp(-\langle Ax,x\rangle)\cdot(c - 2\langle c,x\rangle Ax)$$

• Стац. точки: $\exp(-\langle Ax,x\rangle)\cdot(c - 2\langle c,x\rangle Ax) = \vec{0}$

• сократим на $\exp(-\langle Ax,x\rangle) \neq 0$

$$c - 2\langle c,x\rangle Ax = 0$$

$$c = 2Ax\langle c,x\rangle$$

$$c = 2Ax\langle x,c\rangle$$

$$c = 2Axx^Tc \qquad | \cdot A^{-1} \text{ слева}$$

~~справа~~

$$A^{-1}c = 2xx^Tc$$

$$\cancel{\Rightarrow xx^T = \frac{1}{2}A^{-1}}$$

$$\Rightarrow \cancel{2x}\; xx^Tc = \frac{1}{2}A^{-1}c$$

Бонус | $X \in S^n_{++}$, Найти $\lim\limits_{k\to\infty} tr(X^{-k} - (X^k + X^{2k})^{-1})$

$\lim\limits_{k\to\infty} tr(X^{-k} - (X^{+k} + X^{2k})^{-1}) = \{$тожд. Вудберри: $A = X^k$, $U = V = X^k$, $C = I_n\} = \lim\limits_{k\to\infty} tr(X^{-k} - X^{-k} + X^{-k}X^{+k}(I_n + X_k)^{-1} \cdot X^k X^{-k}) =$

$= \lim\limits_{k\to\infty} tr((I_n + X^k_k)^{-1})$

• Преобразуем $tr((I_n + X^k_k)^{-1})$

для X^k воспользуемся теоремой Жордана:

$\exists\ C_k \in \mathbb{R}^{n\times n}, \det(C) \neq 0 : X^k = C_k J_k C_k^{-1}$, где

J_k - жорд. форма (верхнетреуг. м-ца с диаг. из с.з. X^k)

$\Rightarrow tr((I_n + X^k)^{-1}) = tr((I_n + C_k J_k C_k^{-1})^{-1}) = tr((C_k C_k^{-1} + C_k J_k C_k^{-1})^{-1}) =$

$= tr((C_k(I_n + J_k)C_k^{-1})^{-1}) = tr(C_k \cdot (C_k(I_n + J_k))^{-1}) =$

$= tr(C_k \cdot (I_n + J_k)^{-1} C_k^{-1}) = \{$след одинаков для подоб. матриц$\} = tr((I_n + J_k)^{-1})$

• Т.к. J_k - верхнетреугольная $\Rightarrow I_n + J_k$ - верхнетр.

$\Rightarrow$ обратная к ним так же будет верхнетреугольной, и диагональ обратной будет

• $\exists\ \lambda_1, \ldots, \lambda_n$ - с.з. X $\Rightarrow J_1 = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$

Найдём для $X^2 = X \cdot X = C J_1 C^{-1} \cdot C J_1 C^{-1} = C J_1 J_1 C^{-1}$

$\Rightarrow J_2 = \begin{pmatrix} \lambda_1^2 & & * \\ & \ddots & \\ 0 & & \lambda_n^2 \end{pmatrix}$

и далее для $\forall$ степени $\Rightarrow (I_n + J_k) = \begin{pmatrix} \lambda_1^k+1 & & * \\ & \ddots & \\ 0 & & \lambda_n^k+1 \end{pmatrix}$

$\Rightarrow (I_n + J_k)^{-1} = \begin{pmatrix} \frac{1}{\lambda_1^k+1} & & * \\ & \ddots & \\ 0 & & \frac{1}{\lambda_n^k+1} \end{pmatrix}$

$\Rightarrow tr((I_n + J_k)^{-1}) = \frac{1}{\lambda_1^k+1} + \ldots + \frac{1}{\lambda_n^k+1}$

• Слагаемое $\frac{1}{\lambda^k+1}$ при $k \to \infty$ м.б.

а) $= 1$, $\lambda < 1$ (оно больше 0 по неотр-ти)

б) $= \frac{1}{2}$, $\lambda = 1$

в) $= 0$, $\lambda > 1$

$\Rightarrow \lim\limits_{k \to \infty} tr\left((I_n + J_n^k)^{-1}\right) = m + \frac{p}{2}$, где

m - число с.з. X : с.з. < 1

p - число с.з. X : с.з. $= 1$

Ответ: $\lim\limits_{k \to \infty} tr\left(X^{-k} - (X^k + X^{2k})^{-1}\right) = m + \frac{p}{2}$

m - число с.з. λ_i X : $\lambda_i < 1$

p - число с.з. λ_i X : $\lambda_i = 1$