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

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

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

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

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

=> Верно

■

Задание 2а

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

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

- $uv^T = \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} (v_1 \ldots v_n) = \begin{pmatrix} u_1 v_1 & \ldots & u_1 v_n \\ \vdots & \ddots & \vdots \\ u_m v_1 & \ldots & u_m v_n \end{pmatrix}$

- $uv^T - A = \begin{pmatrix} u_1 v_1 - a_{11} & u_1 v_2 - a_{12} & \ldots & u_1 v_n - a_{1n} \\ u_2 v_1 - a_{21} & u_2 v_2 - a_{22} & \ldots & u_2 v_n - a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ u_m v_1 - a_{m1} & u_m v_2 - a_{m2} & \ldots & u_m v_n - a_{mn} \end{pmatrix}$

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

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

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

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

$= \sum_{i=1}^{n} \sum_{j=1}^{m} \left((u_j v_i)^2 - 2u_j v_i a_{ji} \right)$

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

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

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

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

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

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

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

• $tr\left(\frac{1}{2}(uv^T + vu^T)\right) = \frac{1}{2}\left(tr(uv^T) + tr(vu^T)\right) = \{ tr(uv^T) = u_0v_0 + u_1v_1 + \ldots + u_{n-1}v_{n-1} = tr(vu^T) \} = tr(uv^T) = \langle u, v\rangle$

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

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

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

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

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

■

Задание №2c

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

■

Задача №3а $f(t) = \det(A - tI_n)$

• Первая производная:

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

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

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

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

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

• Вторая производная:

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

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

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

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

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

Ответ: ~~dt~~

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

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

$\Downarrow$

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

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

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

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

$$z = h(t) = (A+tI_n)^{-1}b$$
$$y = g(z) = g(h(t)) = z^Tz$$
$$f(y) = y^{1/2}$$

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

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

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

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

• Вторая производная:

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

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

$$\cdot\ d(b^T((A+tI_n)^{-1})^3 b) = b^T d((A+tI_n)^{-1})^3 b = 3b^T((A+tI_n)^{-1})^2 d(A+tI_n)^{-1} b =$$

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

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

$$\textcircled{=}\ \frac{3b^T((A+tI_n)^{-1})^4 b\,dt_1dt_2}{f(t)} - \frac{(b^T((A+tI_n)^{-1})^3 b)^2 dt_1 dt_2}{f^3(t)}$$

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

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

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

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

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

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

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

• Первая производная:

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

транспон. или 2 части и сдвигаем

$$= 2tr\left((xx^T-A)x\,dx^T\right) = 2tr\left(((xx^T-A)\cdot x)dx^T\right) = 2tr\left(x^T(xx^T-A)^T dx\right) =$$

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

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

• Вторая производная:

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

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

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

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

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

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

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

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

Ответ: $\nabla f = 2(xx^T-A)x$

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

Задача №4б $\quad f(x) = \langle x,x\rangle^{\langle x,x\rangle}$

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

• Первая производная:

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

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

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

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

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

• Вторая производная:

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

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

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

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

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

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

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

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

трансп. то, что в каждых скобках

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

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

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

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

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

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:

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

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

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

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

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

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

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

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

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

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

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: $dX = dX_1 = const$

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

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

$= 2tr((X^{-1})^2 dX_2\cdot X^{-1}\cdot dX_1) = 2tr(X^{-1}dX_2X^{-1}dX_1X^{-1})$

- 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 $X^{-1} = B^TB$

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

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

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

$\Rightarrow$ по определению нормы

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

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