Пшеничников Глеб, 317

№1.

Док-ть тождество Вудбери:

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

$A \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{m\times m}$, $U \in \mathbb{R}^{n\times m}$, $V \in \mathbb{R}^{m\times n}$, $\det(A) \neq 0$, $\det(C) \neq 0$

▶ $$(A+UCV)\left(A^{-1} - A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}\right) =$$

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

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

$$= I + \underline{UCV}A^{-1} - \underline{UCV}A^{-1} = I$$ ■

N2 (a)

Упростить:

a) $\|uv^T - A\|_F^2 - \|A\|_F^2$, где $u \in \mathbb{R}^m$, $v \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$

$\|uv^T - A\|^2 = tr((uv^T - A)(uv^T - A)^T) =$

$= tr((uv^T - A)(vu^T - A^T)) =$

$= tr(uv^Tvu^T - uv^TA^T - Avu^T + AA^T) =$

$= tr(uv^Tvu^T) - tr(uv^TA^T) - tr(Avu^T) +$

$+ tr(AA^T) = \text{☆}$

1) $tr(uv^Tvu^T) = \{tr(AB) = tr(BA)\} = tr(v^Tvu^Tu) =$

$= \{v^Tv$ и u^Tu имеют размер $1 \times 1\} = tr(v^Tv) \cdot$

$\cdot tr(u^Tu) = \|v\|_F^2 \cdot \|u\|_F^2$

2) $tr(Avu^T) = tr(u^TAv) = \{$т.к. размер $1 \times 1\} = u^TAv$

3) $tr(uv^TA^T) = \{tr(B) = tr(B^T)\} = tr(Avu^T) = tr(u^TAv) = u^TAv$

4) $tr(AA^T) = \|A\|_F^2$

☆ $= \|v\|_F^2\|u\|_F^2 - 2u^TAv + \|A\|_F^2$

Тогда $\|uv^T - A\|_F^2 - \|A\|_F^2 = \boxed{\|v\|_F^2 \cdot \|u\|_F^2 - 2u^TAv}$

№2 (b).

Упростить:

b) $tr((2I_n + aa^T)^{-1}(uv^T + vu^T))$, $a, u, v \in \mathbb{R}^n$

Воспользуемся тождеством Вудбери:

$(2I_n + aa^T)^{-1} = (2I_n + aI_1a^T)^{-1} =$

$= \frac{1}{2}I_n - \frac{1}{2}I_n a(I_1 + a^T\frac{1}{2}I_n a)^{-1}a^T\frac{1}{2}I_n =$

$= \{I_1 + a^T\frac{1}{2}I_n a = 1 + \frac{1}{2}a^Ta = 1 + \frac{1}{2}\|a\|^2\} =$

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

$tr((2I_n + aa^T)^{-1}(uv^T + vu^T)) =$

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

$= tr(\frac{1}{2}I_n(uv^T + vu^T)) - tr(\frac{1}{4}\frac{aa^T}{1 + \frac{1}{2}\|a\|^2}(uv^T + vu^T)) = \circledast$

1) $tr(\frac{1}{2}I_n(uv^T + vu^T)) = tr(\frac{1}{2}uv^T + \frac{1}{2}vu^T) =$

$= \frac{1}{2}tr(uv^T) + \frac{1}{2}tr(vu^T) = \frac{1}{2}tr(vu^T) + \frac{1}{2}tr(vu^T) =$

$= tr(vu^T) = tr(u^Tv) = \{$т.к. размерн. $1\times1\} = u^Tv$

2) $tr(\frac{1}{4}\frac{aa^T}{1 + \frac{1}{2}\|a\|^2}(uv^T + vu^T)) =$

$= tr(\frac{1}{4}\frac{aa^T}{1 + \frac{1}{2}\|a\|^2}uv^T) + tr(\frac{1}{4}\frac{aa^T}{1 + \frac{1}{2}\|a\|^2}vu^T) =$

$= \frac{1}{4}\frac{1}{1 + \frac{1}{2}\|a\|^2}(tr(aa^Tuv^T) + tr(aa^Tvu^T)) =$

$= \begin{cases} tr(aa^Tuv^T) = tr(a^Tuv^Ta) = a^Tuv^Ta \\ tr(aa^Tvu^T) \underset{tr(B) = tr(B^T)}{=} tr(uv^Taa^T) = tr(a^Tuv^Ta) = a^Tuv^Ta \end{cases} =$

$= \frac{1}{4}\frac{1}{1 + \frac{1}{2}\|a\|^2} \cdot 2 \cdot a^Tuv^Ta = \frac{1}{2}\frac{a^Tuv^Ta}{1 + \frac{1}{2}\|a\|^2}$

$\circledast = u^Tv - \frac{1}{2}\frac{a^Tuv^Ta}{1 + \frac{1}{2}\|a\|^2} = \boxed{u^Tv - \frac{a^Tuv^Ta}{2 + \|a\|^2}}$

Упростить

c) $\sum_{i=1}^{n} \langle S^{-1} a_i, a_i \rangle$, где $a_1, \dots, a_n \in \mathbb{R}^d$, $S = \sum_{i=1}^{n} a_i a_i^T$, $\det(S) \neq 0$

$\langle S^{-1} a_i, a_i \rangle = \langle a_i, S^{-1} a_i \rangle = a_i^T S^{-1} a_i =$

$= \{$т.к. размерн $1 \times 1\} = tr(a_i^T S^{-1} a_i) =$

$= \{tr(AB) = tr(BA)\} = tr(S^{-1} a_i a_i^T)$

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

$= \{S = \sum_{i=1}^{n} a_i a_i^T\} = tr(S^{-1} \cdot S) = tr(I_d) = \boxed{d}$

№3 (a)

Найти 1-ю и 2-ю производ.

а) $f: E \to \mathbb{R}$, $f(t) = \det(A - tI_n)$, $A \in \mathbb{R}^{n\times n}$

$E = \{t \in \mathbb{R} : \det(A - tI_n) \neq 0\}$

$df = d(\det(A-tI_n)) = \{d\det X = \det X \langle X^{-T}, dX\rangle\} =$

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

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

$= \{X^{-1} = \frac{1}{\det X}\hat{X}$, где $\hat{X}$ – матрица алгебр. дополнения X_{ij} к x_{ij} матрицы $X\} =$

$= \det(A-tI_n)\langle \frac{1}{\det((A-tI_n)^T)} \widehat{(A-tI_n)^T}, -dtI_n\rangle =$

$= -\langle \widehat{(A-tI_n)}^T, I_n\rangle dt = -tr(\widehat{(A-tI_n)})\cdot dt$

$d^2 f = d(-tr(\widehat{(A-tI_n)})\, dt_1) =$

$= -d(\langle \widehat{(A-tI_n)}, I_n\rangle dt_1) = \{X^{-1} = \frac{1}{\det X}\hat{X}\} =$

$= -d(\langle (A-tI_n)^{-1}\det(A-tI_n), I_n\rangle dt_1) =$

$= \{d(X^{-1}) = -X^{-1}\cdot dX\cdot X^{-1};\ d(\det(A-tI_n)) = -tr(\widehat{(A-tI_n)})dt\} =$

$= -\langle \det(A-tI_n)\cdot(-1)\cdot(A-tI_n)^{-1}\cdot(-dt_2)\cdot(A-tI_n)^{-1} +$

$+ (A-tI_n)^{-1}\cdot(-1)\cdot tr(\widehat{A-tI_n})dt_2, I_n\rangle dt_1 =$

$= -\langle \widehat{(A-tI_n)}\cdot(A-tI_n)^{-1}dt_2 - (A-tI_n)^{-1}tr(\widehat{A-tI_n})dt_2, I_n\rangle\cdot$

$\cdot dt_1 = -tr[\widehat{(A-tI_n)}\cdot(A-tI_n)^{-1}]dt_1 dt_2 +$

$+ tr((A-tI_n)^{-1})\cdot tr(\widehat{A-tI_n})dt_1 dt_2$

$f'(t) = -tr(\widehat{A-tI_n})$

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

$+ tr((A-tI_n)^{-1})\cdot tr(\widehat{A-tI_n})$

N3(b)

b) $f: \mathbb{R}_{++} \to \mathbb{R}$; $f(t) = \|(A + tI_n)^{-1} b\|$, $A \in \mathbb{S}^n_+$, $b \in \mathbb{R}^n$

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

$$= \left[b^T(A+tI_n)^{-T}(A+tI_n)^{-1}b\right]^{\frac{1}{2}} = \{A \in \mathbb{S}^n \Rightarrow A + tI_n \in \mathbb{S}^n \Rightarrow$$

$$\Rightarrow (A+tI_n)^{-T} = (A+tI_n)^{-1}\} = \left[b^T(A+tI_n)^{-2}b\right]^{\frac{1}{2}}$$

$$f'(t) = \frac{1}{2}\left[b^T(A+tI_n)^{-2}b\right]^{-\frac{1}{2}} \frac{d}{dt}\left(b^T(A+tI_n)^{-2}b\right) =$$

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

$$= \frac{1}{2}\|(A+tI_n)^{-1}b\|^{-1}\left(b^T \cdot \frac{d}{dt}(A+tI_n)^{-2} \cdot b\right) =$$

$$= \boxed{-\frac{b^T(A+tI_n)^{-3}b}{\|(A+tI_n)^{-1}b\|}}$$

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

$$= \frac{3b^T(A+tI_n)^{-4}b \cdot \|(A+tI_n)^{-1}b\| - b^T(A+tI_n)^{-3}b \cdot \frac{-b^T(A+tI_n)^{-3}b}{\|(A+tI_n)^{-1}b\|}}{\|(A+tI_n)^{-1}b\|^2} =$$

$$= \boxed{\frac{3b^T(A+tI_n)^{-4}b}{\|(A+tI_n)^{-1}b\|} + \frac{\left[b^T(A+tI_n)^{-3}b\right]^2}{\|(A+I_n)^{-1}b\|^3}}$$

N4(a)

a) $f: \mathbb{R}^n \to \mathbb{R}$, $f(x) = \frac{1}{2}\|xx^T - A\|_F^2$, $A \in \mathbb{S}^n$

$df = \langle \nabla f, dx \rangle = tr(dx^T \nabla f(x)) = tr(\nabla f(x)\, dx^T)$

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

$= \frac{1}{2} tr((xx^T - A)(xx^T - A)) = \frac{1}{2} tr(xx^Txx^T - xx^TA - Axx^T + A^2)$

$d(xx^T) = dx \cdot x^T + x\, dx^T$

$df = \frac{1}{2} tr(d(xx^Txx^T - xx^TA - Axx^T + A^2)) =$

$= \frac{1}{2} tr(d(xx^T) \cdot xx^T + xx^T d(xx^T) - d(xx^T) \cdot A - A d(xx^T)) =$

$= \frac{1}{2} tr[dx \cdot x^Txx^T + x\, dx^T \cdot x \cdot x^T + xx^T dx x^T + xx^Tx\, dx^T -$

$- dx\, x^TA - x\, dx^TA - A\, dx\, x^T - Ax\, dx^T] =$

$= \{ tr(B) \overset{①}{=} tr(B^T), tr(AB) \overset{②}{=} tr(BA) \} =$

$= \frac{1}{2} tr(xx^T \overset{①}{x}\, dx^T + xx^T \overset{②}{x}\, dx^T + xx^T \overset{①②}{x}\, dx^T + xx^Tx\, dx^T -$

$- A^T \overset{①}{x}\, dx^T - A \overset{②}{x}\, dx^T - A^T \overset{①②}{x}\, dx^T - Ax\, dx^T) = \{A = A^T\} =$

$= \frac{1}{2} tr(4xx^Tx\, dx^T - 4Ax\, dx^T) = tr(2(xx^Tx - Ax)\, dx^T)$

$\Rightarrow \boxed{\nabla f(x) = 2(xx^Tx - Ax)}$

$d^2 f = d(tr(2(xx^Tx - Ax)\, dx_1^T)) = d\langle 2(xx^Tx - Ax), dx_1 \rangle =$

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

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

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

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

$\nabla^2 f = 2(x^Tx I_n + 2xx^T - A) = \boxed{2(\|x\|^2 I_n + 2xx^T - A)}$

N4(b)

b) $f: \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $f(x) = \langle x, x \rangle^{\langle x, x \rangle}$

$d(a^b) \overset{?}{=} a^b d(\ln a^b)$

$a^b = e^{\ln(a^b)} = e^{b \ln a}$

$d(a^b) = d(e^{b\ln a}) = e^{b\ln a} d(b \ln a) = a^b d(\ln a^b)$

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

$= (x^Tx)^{x^Tx} [\ln(x^Tx) d(x^Tx) + x^Tx \, d(\ln(x^Tx))] =$

$= \{ d(x^Tx) = d\langle x, x\rangle = 2\langle x, dx\rangle;\ d\ln B = \frac{1}{B} dB \} =$

$= (x^Tx)^{x^Tx} [\ln(x^Tx) \cdot 2\langle x, dx\rangle + x^Tx \cdot \frac{1}{x^Tx} \cdot 2\langle x, dx\rangle] =$

$= 2(x^Tx)^{x^Tx} [\ln(x^Tx) + 1] \cdot \langle x, dx\rangle$

$\nabla f = 2(x^Tx)^{x^Tx} [\ln(x^Tx) + 1] x =$

$= \boxed{2[\ln(x^Tx) + 1] (x^Tx)^{(x^Tx)} \cdot x}$

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

$= 2\langle d((x^Tx)^{x^Tx}[\ln(x^Tx)+1] \cdot x), dx_1\rangle =$

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

$= 2\langle [2(x^Tx)^{x^Tx}(\ln(x^Tx)+1)x\langle x, dx_2\rangle + (x^Tx)^{x^Tx} dx_2](\ln(x^Tx)+1) +$

$+ \frac{d(x^Tx)}{x^Tx} \cdot (x^Tx)^{x^Tx} \cdot x, dx_1\rangle = \{\langle x, dx_2\rangle = x^T dx_2\}$

$= 2\langle 2(x^Tx)^{x^Tx}(\ln(x^Tx)+1)^2 x x^T dx_2 + (x^Tx)^{x^Tx}(\ln(x^Tx)+1) dx_2 +$

$+ (x^Tx)^{x^Tx} \cdot \frac{1}{x^Tx} 2 x x^T dx_2, dx_1\rangle =$

$= \langle (x^Tx)^{x^Tx} [(4(\ln(x^Tx)+1)^2 + \frac{4}{x^Tx}) x x^T + 2(\ln(x^Tx)+1) I_n] dx_2, dx_1\rangle$

$\boxed{\nabla f^2 = (x^Tx)^{x^Tx} [4(\ln(x^Tx)+1)^2 \cdot x x^T + \frac{4 x x^T}{x^Tx} + 2(\ln(x^Tx)+1) I_n]}$

№4 (c)

c) $f: \mathbb{R}^n \to \mathbb{R}$, $f(x) = \|Ax - b\|^p$, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $p \geq 2$

$$df = p \cdot \|Ax-b\|^{p-1} \cdot d(\|Ax-b\|) = p\|Ax-b\|^{p-1} d(\sqrt{\|Ax-b\|^2}) =$$

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

$$= p\|Ax-b\|^{p-2} \cdot \frac{1}{2} d(\langle Ax-b, Ax-b\rangle) =$$

$$= p\|Ax-b\|^{p-2} \langle Ax-b, d(Ax-b)\rangle = p\|Ax-b\|^{p-2} \langle Ax-b, Adx\rangle =$$

$$= p\|Ax-b\|^{p-2} \langle A^T(Ax-b), dx\rangle$$

$$\boxed{\nabla f = p\|Ax-b\|^{p-2} \cdot A^T(Ax-b)}$$

$$d^2 f = d\langle p\|Ax-b\|^{p-2} \cdot A^T(Ax-b), dx_1\rangle =$$

$$= p\langle d(\|Ax-b\|^{p-2}) \cdot A^T(Ax-b) + \|Ax-b\|^{p-2} \cdot d(A^T(Ax-b)), dx_1\rangle =$$

$$= p\langle (p-2) \cdot \|Ax-b\|^{p-4} A^T(Ax-b) \cdot (Ax-b)^T A dx_2 +$$

$$+ \|Ax-b\|^{p-2} \cdot A^T \cdot A dx_2, dx_1\rangle =$$

$$= p\langle [(p-2)\|Ax-b\|^{p-4} A^T(Ax-b)(Ax-b)^T A +$$

$$+ \|Ax-b\|^{p-2} A^T A] \cdot dx_2, dx_1\rangle$$

$$\boxed{\nabla^2 f = p(p-2)\|Ax-b\|^{p-4} A^T(Ax-b)(Ax-b)^T A + p\|Ax-b\|^{p-2} A^T A}$$

N5(a)

L10

a) $f: S^n_{++} \to \mathbb{R},\ f(x) = tr(x^{-1})$

$df = d(tr(x^{-1})) = tr(d(x^{-1}))$

$dx^{-1} \overset{?}{=} -x^{-1}\, dx\, x^{-1}$

$xx^{-1} = I$

$d(xx^{-1}) = 0$

$dx\, x^{-1} + x\, d(x^{-1}) = 0 \Rightarrow d(x^{-1}) = -x^{-1} dx\, x^{-1}$

$df = tr(-x^{-1} dx\, x^{-1}) = \left\{ \begin{array}{l} tr(AB) = \langle A, B^T \rangle \\ tr(AB) = \sum_{ij} A_{ij} B_{ji} \\ \langle A, B^T \rangle = \sum_{ij} A_{ij} B^T_{ij} = \sum_{ij} A_{ij} B_{ji} \end{array} \right\} =$

$= -\langle x^{-1}, (dx\, x^{-1})^T \rangle = -\langle x^{-1}, x^{-1} dx \rangle$ (т.к. $x = x^T$).

$d^2 f = d(-\langle x^{-1}, x^{-1} dx \rangle) = -\langle d(x^{-1}), x^{-1} dx \rangle - \langle x^{-1}, d(x^{-1}) dx \rangle =$

$= \langle x^{-1} dx\, x^{-1}, x^{-1} dx \rangle + \langle x^{-1}, x^{-1} dx\, x^{-1} dx \rangle =$

$= tr(x^{-1} dx\, x^{-1} dx\, x^{-1}) + tr(x^{-1} dx\, x^{-1} dx\, x^{-1}) =$

$= 2\, tr(x^{-1} dx\, x^{-1} dx\, x^{-1})$

$\exists\ A = x^{-1} dx\, x^{-\frac{1}{2}}$

$x^{-1} dx\, x^{-1} dx\, x^{-1} = (x^{-1} dx\, x^{-\frac{1}{2}})(x^{-\frac{1}{2}} dx\, x^{-1}) =$

$= \{x^{-1} = x^{-T},\ x^{-\frac{1}{2}} = (x^{-\frac{1}{2}})^T\} = AA^T$

$2\, tr(x^{-1} dx\, x^{-1} dx\, x^{-1}) = 2\, tr(AA^T) = 2 \sum_{ij} A_{ij} A^T_{ji} =$

$= 2 \sum A_{ij}^2 > 0$

$\Rightarrow \boxed{d^2 f > 0}$

№5(б)

б) $f: S^n_{++} \to \mathbb{R},\ f(X) = (\det(X))^{\frac{1}{n}}$

$$df = \frac{1}{n}\det(X)^{\frac{1}{n}-1} d(\det(X)) = \frac{1}{n}\det(X)^{\frac{1}{n}-1}\cdot \det X \cdot \langle X^{-T}, dX\rangle =$$

$$= \frac{1}{n}\det(X)^{\frac{1}{n}}\cdot\langle X^{-T}, dX\rangle = \{X^{-T} = X^{-1}\} = \frac{1}{n}\det(X)^{\frac{1}{n}}\langle X^{-1}, dX\rangle$$

$$d^2 f = \frac{1}{n} d\left(\det(X)^{\frac{1}{n}}\langle X^{-1}, dX_1\rangle\right) =$$

$$= \frac{1}{n}\left[d(\det(X)^{\frac{1}{n}})\langle X^{-1}, dX_1\rangle + \det(X)^{\frac{1}{n}} d(\langle X^{-1}, dX_1\rangle)\right] =$$

$$= \frac{1}{n}\left[\frac{1}{n}\det(X)^{\frac{1}{n}}\langle X^{-1}, dX_2\rangle\langle X^{-1}, dX_1\rangle - \right.$$

$$\left. - \det(X)^{\frac{1}{n}}\langle X^{-1}dX_2X^{-1}, dX_1\rangle\right] =$$

$$= \frac{1}{n}\det(X)^{\frac{1}{n}}\left[\frac{1}{n}\langle X^{-1}, dX\rangle^2 - \langle X^{-1}dXX^{-1}, dX\rangle\right] =$$

$$= \frac{1}{n}\det(X)^{\frac{1}{n}}\left[\frac{1}{n}\mathrm{tr}(X^{-1}dX)^2 - \mathrm{tr}(X^{-1}dXX^{-1}dX)\right]$$

По неравенству Коши-Буняковского

$$|\mathrm{tr}(AB)|^2 \le \mathrm{tr}(AA^T)\,\mathrm{tr}(BB^T)$$

$A = X^{-1}dX \Rightarrow A^T = X^{-1}dX$

$B = I \Rightarrow B^T = I$ $\Big| \Rightarrow$

$$\Rightarrow |\mathrm{tr}(X^{-1}dX)|^2 \le \mathrm{tr}(X^{-1}dXX^{-1}dX)\cdot n \Rightarrow$$

$$\Rightarrow \boxed{d^2 f < 0}$$

№6(a)

a) $f: \mathbb{R}^n \to \mathbb{R}$, $f(x) = \langle c, x\rangle + \frac{\sigma}{3}\|x\|^3$, $c \in \mathbb{R}^n$, $c \neq 0$, $\sigma > 0$

$df = d(\langle c, x\rangle) + \frac{\sigma}{3} d(\|x\|^3) = \left\{ d(\frac{1}{3}\|x\|^3) = \|x\| \langle x, dx\rangle \right\} =$

$= \langle c, dx\rangle + \sigma \|x\| \langle x, dx\rangle = \langle c + \sigma \|x\| x, dx\rangle$

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

$c + \sigma \|x\| x = \vec{0}$

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

$x = -\frac{c}{\sigma} \cdot \frac{1}{\|x\|}$

Возьмем норму:

$\|x\| = +\frac{\|c\|}{\sigma} \cdot \frac{1}{\|x\|}$

$\|x\|^2 = \frac{\|c\|}{\sigma}$

$\|x\| = \pm\sqrt{\frac{\|c\|}{\sigma}}$, т.к. $\|x\| > 0$, то $\|x\| = \sqrt{\frac{\|c\|}{\sigma}}$

Тогда: $x = -\frac{c}{\sigma} \cdot \sqrt{\frac{\sigma}{\|c\|}} = \boxed{-\frac{c}{\sqrt{\sigma \|c\|}}\text{, где } c \neq 0,\ \sigma > 0}$

точки стационарн.

№6(b)

$f: E \to \mathbb{R}$, $f(x) = \langle a, x\rangle - \ln(1 - \langle b, x\rangle)$, $a, b \in \mathbb{R}^n$, $a, b \neq 0$

$E - \{x \in \mathbb{R}^n / \langle b, x\rangle < 1\}$

$$df = d(\langle a, x\rangle) - d(\ln(1 - \langle b, x\rangle) = \langle a, dx\rangle +$$

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

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

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

$a(1 - \langle b, x\rangle) + b = 0$ | умножим скалярно на b

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

$$\langle a, b\rangle(1 - \langle b, x\rangle) + \|b\|^2 = 0$$

$$1 - \langle b, x\rangle = -\frac{\|b\|^2}{\langle a, b\rangle}$$

$$\langle b, x\rangle = 1 + \frac{\|b\|^2}{\langle a, b\rangle}$$

Заметим условия существ.:

$$\langle b, x\rangle < 1 \Rightarrow 1 + \frac{\|b\|^2}{\langle a, b\rangle} < 1$$

$$\frac{\|b\|^2}{\langle a, b\rangle} < 0 \Rightarrow \boxed{\langle a, b\rangle < 0}$$

$$\nabla f = a + \frac{b}{1 - \langle b, x\rangle} = 0$$

$$a - a\langle b, x\rangle + b = 0$$

$$a - ab^T x + b = 0$$

$$\boxed{x = \frac{a + b}{ab^T}, \text{ где } \langle a, b\rangle < 0}$$

↑

точки стационарн.

$\sim 6(c)$

c) $f: \mathbb{R}^n \to \mathbb{R}$, $f(x) = \langle c, x\rangle e^{-\langle Ax, x\rangle}$, $c \in \mathbb{R}^n$, $c \neq 0$, $A \in S^n_{++}$

$$df = \langle c, dx\rangle e^{-\langle Ax,x\rangle} + \langle c, x\rangle e^{-\langle Ax, x\rangle} \cdot 2 \cdot (-1) \cdot \langle Ax, dx\rangle \Rightarrow$$

$$\Rightarrow \nabla f = e^{-\langle Ax, x\rangle}\left[c - 2\langle c, x\rangle \cdot Ax\right]$$

$$e^{-\langle Ax, x\rangle}\left[c - 2\langle c, x\rangle Ax\right] = 0$$

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

$$Ax = \frac{c}{2\langle c, x\rangle}$$

Т.к. $A \in S^n_{++}$, то $\exists A^{-1}$

(1) $x = \dfrac{A^{-1}c}{2\langle c, x\rangle}$ | умножим скал на c

$$\langle c, x\rangle = \frac{\langle c, A^{-1}c\rangle}{2\langle c, x\rangle}$$

$$2\langle c, x\rangle^2 = \langle c, A^{-1}c\rangle$$

$$\langle c, x\rangle = \pm\sqrt{\tfrac{1}{2}\langle c, A^{-1}c\rangle}$$

в (1): $x = \pm\dfrac{A^{-1}c}{2\sqrt{\frac{1}{2}\langle c, A^{-1}c\rangle}} = \boxed{\pm\dfrac{A^{-1}c}{\sqrt{2c^T \cdot A^{-1}c}}, \text{ где } c \neq 0}$

точки стационарн

N7

$X \in S^n_{++}$

$$\lim_{k \to +\infty} tr(X^{-k} - (X^k + X^{2k})^{-1}) - ?$$

$$tr(X^{-k} - (X^k + X^{2k})^{-1}) = tr(X^{-k}) - tr((X^k + X^{2k})^{-1}) =$$

$$= tr(X^{-k}) - tr(X^{-k}(I_n + X^k)^{-1}) =$$

$$= tr\left(X^{-k}(I_n - (I_n + X^k)^{-1})\right) =$$

$$= tr\left[X^{-k}\left(I_n - \frac{I_n}{1+X^k}\right)\right] =$$

$$= tr\left[X^{-k} \cdot \frac{I_n + X^k - I_n}{1+X^k}\right] =$$

$$= tr\left[X^{-k} \cdot X^k \cdot \frac{1}{1+X^k}\right] = tr\left[\frac{I}{1+X^k}\right]$$

Т.к. $\det(X) > 0$, то $X^k \to +\infty \Rightarrow$

$$\Rightarrow \frac{I}{1+X^k} \to 0$$

$$\boxed{\lim_{k \to +\infty} tr(X^{-k} - (X^k + X^{2k})^{-1}) = 0}$$

№8(a)

$\{x_i\}_{i=1}^N$ - выборка, $x_i \in \mathbb{R}^D$

$P \in \mathbb{R}^{D\times d}$

$P(P^TP)^{-1}P^Tx$ - проекция

$$F(P) = \sum_{i=1}^N \|x_i - P(P^TP)^{-1}P^Tx_i\|^2 = N\,tr((I-P(P^TP)^{-1}P^T)^2 S) \longrightarrow \min_P$$

$$S = \frac{1}{N}\sum_{i=1}^N x_i x_i^T \Rightarrow S = S^T$$

а) $dF(P) = N\cdot tr(d[(I-P(P^TP)^{-1}P^T)^2 S]) =$

$= N\,tr(2(I-P(P^TP)^{-1}P^T)\cdot d(I-P(P^TP)^{-1}P^T)S) =$

$= N\,tr(-2(I-P(P^TP)^{-1}P^T)[+dP(P^TP)^{-1}P^T +$

$+ P\,d(P^TP)^{-1}P^T + P(P^TP)^{-1}dP^T]S) =$

$= \{ d(P^TP)^{-1} = -(P^TP)^{-1}(dP^TP + P^TdP)(P^TP)^{-1}$ и $P^TP = I\} =$

$= N\,tr(-2(I-PP^T)[dPP^T - P(dP^TP + P^TdP)P^T +$

$+ P\,dP^T]S) =$

$= -2N\,tr(dPP^TS - PP^TdPP^TS - P(dP^TP + P^TdP)P^TS +$

$+ PP^TP(dP^TP + P^TdP)P^TS + PdP^TS - PP^TPdP^TS) =$

$= \{P^TP = I\} = -2N\,tr(dPP^TS - PP^TdPP^TS -$

$- PdP^TPP^TS - PP^TdPP^TS + PdP^TPP^TS + PP^TdPP^TS +$

$+ PdP^TS - PdP^TS) = -2N\,tr(dPP^TS - PP^TdPP^TS) =$

$= -2N\,tr(P^TS\,dP - P^TSPP^TdP) =$

$= -2N\,tr((P^TS - P^TSPP^T)dP) \Rightarrow$

$$\Rightarrow \boxed{\nabla F(P) = -2N(SP - PP^TSP)}$$

№8 (b)

17

$\nabla F = -2N(SP - PP^T SP)$

$S = Q\Lambda Q^T$

$\nabla F = -2N(Q\Lambda Q^T P - PP^T Q\Lambda Q^T P)$

1) $Q^T P$ — матрица $D \times d$, где в строках $i_1 \dots i_d$ стоят 1, остальные элем. нули

2) $\Lambda Q^T P$ — то же самое, но строки умножены на соответств. собств. знач.

3) $Q\Lambda Q^T P = P\,\mathrm{diag}(\lambda_{i_1}, \dots, \lambda_{i_d})$, т.к. каждый столбец умнож на его собств. знач

4) $PP^T x$ — проекция на подпр-во, натянутое на столбцы P

$P = [q_{i_1} | q_{i_2} | \dots | q_{i_d}]$, q_{i_j} — собств. вект. S

$\Rightarrow PP^T q_{i_j} = q_{i_j}$

$\Rightarrow PP^T Q\Lambda Q^T P = Q\Lambda Q^T P = P\,\mathrm{diag}(\lambda_{i_1} \dots \lambda_{i_d})$

$F(P) = N\,\mathrm{tr}((I - PP^T)^2 S) = N\,\mathrm{tr}((I - PP^T)^2 Q\Lambda Q^T) =$

$\cancel{= N\,\mathrm{tr}(Q\Lambda Q^T + 2PP^T Q\Lambda Q^T + PP^T Q\Lambda Q^T PP^T)}$

$\cancel{= N\,\mathrm{tr}((I - PP^T)}$

$= \mathrm{tr}(\Lambda Q^T (I - PP^T)^2 Q)$

$Q^T (I - PP^T)^2 Q$ — диагональная матр. на диагон. 1 для невыбр. собст. вект и 0 для выбранных

$\Rightarrow F(P) = N \sum_{j \notin \{i_1, \dots, i_d\}} \lambda_j \Rightarrow$ минимум при выборе d макс

$P = [q_{i_1} | \dots | q_{i_d}]$ с.з.