Суглобов Кирилл,
гр. 317

Домашнее задание
Матричные вычисления и матричное дифференцирование.

Обозначения:

- $\langle x, y \rangle$ — Евкл. скал. произв.
- $\|x\| \equiv \langle x, x \rangle^{\frac{1}{2}} = (x^T x)^{\frac{1}{2}}$ — Евкл. норма вектора.
- $\|A\|_F = \langle A, A \rangle^{\frac{1}{2}} = tr(A^T A)^{\frac{1}{2}}$ — Фробен. норма матрицы.
- I_n — единичная матрица $n \times n$.
- $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x_i \geq 0 \ \forall i\}$, $\mathbb{R}^n_{++} = \{x \in \mathbb{R}^n \mid x_i > 0 \ \forall i\}$
- $\mathbb{S}^n = \{A \in \mathbb{R}^{n \times n} \mid A = A^T\}$ — симметричность.
- $\mathbb{S}^n_+ = \{A \in \mathbb{S}^n \mid A$ — неотр. опред.$\}$,
  $\mathbb{S}^n_{++} = \{A \in \mathbb{S}^n \mid A$ — положительно опред.$\}$;

$\boxed{1}$ Док-ть тождество Вудбери:

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

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

═ Решение ═

1) Рассм. выр-е:

$$(C^{-1} + VA^{-1}U) = C^{-1}(I + CVA^{-1}U) = C^{-1}U^{-1}(U + UCVA^{-1}U) = C^{-1}U^{-1}(A + UCV)A^{-1}U;$$

Возьмём от него обратную матрицу $\Rightarrow (C^{-1} + VA^{-1}U)^{-1} = [C^{-1}U^{-1}(A + UCV)A^{-1}U]^{-1}$;

2) Рассм. выр-е: $A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}$;

Докажем, что оно равно $(A + UCV)^{-1}$ $\Longrightarrow$

$$\Rightarrow \left[I - (V^{-1}A)^{-1} \underbrace{(C^{-1} + VA^{-1}U)^{-1}}_{(C^{-1}U^{-1}(A+UCV)A^{-1}U)} (V^{-1})^{-1} \right] A^{-1} =$$

$$= \left[I - \left(\underbrace{V^{-1}C^{-1}U^{-1}}_{(UCV)^{-1}} (A+UCV) \underbrace{A^{-1} \underbrace{U U^{-1}}_{I} A}_{I} \right)^{-1} \right] A^{-1} =$$

$$\Rightarrow \left[I - \left((UCV)^{-1} \cdot (A+UCV) \right)^{-1} \right] A^{-1} =$$

$$= \left[I - (A+UCV)^{-1}(UCV) \right] A^{-1}; \quad I = (A+UCV)^{-1} \cdot (A+UCV)$$

$$\Rightarrow \left[(A+UCV)^{-1}(A+UCV) - (A+UCV)^{-1}(UCV) \right] A^{-1} =$$

$$= (A+UCV)^{-1} \underbrace{\left[A + \cancel{(UCV)} - \cancel{(UCV)} \right] A^{-1}}_{I} = (A+UCV)^{-1};$$

ч.т.д.

$\boxed{2}$ Упростить выр-е:

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

Решение

$\|B\|_F = tr(B^TB)^{\frac{1}{2}} \; \forall B \in \mathbb{R}^{m\times n} \Rightarrow \|uv^T - A\|_F^2 - \|A\|_F^2 =$

$$= \left[tr((uv^T - A)^T(uv^T - A))^{\frac{1}{2}} \right]^2 - \left[tr(A^TA)^{\frac{1}{2}} \right]^2 =$$

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

$$- vu^TA - A^Tuv^T + \cancel{A^TA} - \cancel{A^TA}) = tr(vu^Tuv^T) -$$

$$- tr(vu^TA) - tr(A^Tuv^T) = \{\text{циклич. св-во trace}\} \Rightarrow$$

$$\Rightarrow tr(\underbrace{\overbrace{v^Tv}^{=\|v\|^2}}_{\in \mathbb{R}}\underbrace{\overbrace{u^Tu}^{=\|u\|^2}}_{\in \mathbb{R}}) - tr(\underbrace{u^TAv}_{\in \mathbb{R}}) - tr(\underbrace{v^TA^Tu}_{\in \mathbb{R}}) = \quad (\Rightarrow = u^TAv)$$

$$= \{\text{trace от } x \in \mathbb{R} \text{ равен } x\} = \underline{\|v\|^2 \cdot \|u\|^2 - 2u^TAv};$$

$$\boxed{\text{Ответ: } v^Tvu^Tu - 2u^TAv \equiv (\|v\|\cdot\|u\|)^2 - 2u^TAv;}$$

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

Используем тождество Вудбери.

1) По тождеству Вудбери:

$$(2I_n + aa^T)^{-1} = \left\{\begin{matrix} A = 2I_n \in \mathbb{R}^{n\times n}, & C = 1 \in \mathbb{R}^{1\times 1}, \\ U = a \in \mathbb{R}^{n\times 1}, & V = a^T \in \mathbb{R}^{1\times n} \end{matrix}\right\} =$$

$$= \underset{A^{-1}}{(2I_n)^{-1}} - \underset{A^{-1}}{(2I_n)^{-1}} \cdot \underset{U}{a} \cdot (\underset{C^{-1}}{\tfrac{1}{1}} + \underset{V}{a^T} \cdot \underset{A^{-1}}{(2I_n)^{-1}} \cdot \underset{U}{a})^{-1} \cdot \underset{V}{a^T} \cdot \underset{A^{-1}}{(2I_n)^{-1}} =$$

$$= \underbrace{\tfrac{1}{2}I_n}_{\in \mathbb{R}^{n\times n}} - \underbrace{\tfrac{1}{2}I_n}_{\in \mathbb{R}^{n\times n}} \cdot \underbrace{a \cdot \overbrace{(1 + \tfrac{1}{2}a^Ta)^{-1}}^{\mathbb{R}} a^T}_{\mathbb{R}^{n\times n}} \cdot \underbrace{\tfrac{1}{2} \cdot I_n}_{\in \mathbb{R}^{n\times n}} = \{a^Ta = \|a\|^2\}$$

$$\Rightarrow \boxed{\tfrac{1}{2}I_n - \frac{1}{4(1 + \frac{1}{2}\|a\|^2)} \cdot aa^T \in \mathbb{R}^{n\times n}} \qquad \exists\, k = \frac{1}{4(1 + \frac{1}{2}\|a\|^2)}$$

2) $\Rightarrow$ Исходное выр-е $= tr((\tfrac{1}{2}I_n - kaa^T)(uv^T + vu^T)) \Rightarrow$

$$\Rightarrow tr(\tfrac{1}{2}I_nuv^T + \tfrac{1}{2}I_nvu^T - kaa^Tuv^T - kaa^Tvu^T) \overset{\text{циклич. trace}}{=}$$

$$= \tfrac{1}{2}\underbrace{tr(I_nuv^T)}_{tr(\underbrace{v^Tu}_{\in \mathbb{R}})} + \tfrac{1}{2}\underbrace{tr(I_nvu^T)}_{tr(\underbrace{u^Tv}_{\in \mathbb{R}})} - k\underbrace{tr(aa^Tuv^T)}_{tr(\underbrace{a^Tu}_{\in \mathbb{R}}\underbrace{v^Ta}_{\in \mathbb{R}})} - k\underbrace{tr(aa^Tvu^T)}_{tr(\underbrace{a^Tv}_{\in \mathbb{R}}\underbrace{u^Ta}_{\in \mathbb{R}})} \Rightarrow$$

$\Rightarrow$ {т.к. $tr(x) = x$ для $x \in \mathbb{R}$} = {цикличность trace} =

$$= \frac{1}{2}(u^T v + \underbrace{v^T u}_{\in\mathbb{R} \Rightarrow = (v^T u)^T = u^T v}) - k\cdot(a^T u v^T a + tr(\underbrace{u^T a}_{\mathbb{R}} \underbrace{a^T v}_{\mathbb{R}})) \ominus$$

$tr(\underbrace{a^T u}_{\mathbb{R}} \underbrace{v^T a}_{\mathbb{R}}) = a^T u v^T a$

$\ominus\ u^T v - 2k a^T u v^T a \Rightarrow$

$$\Rightarrow \left\{k = \frac{1}{4(1 + \frac{1}{2}\|a\|^2)}\right\} =$$

$$= \underbrace{u^T v}_{\langle u, v\rangle} - \frac{1}{2 + \|a\|^2} \underbrace{a^T u}_{\mathbb{R} \Rightarrow \langle a, u\rangle = \langle u, a\rangle} \underbrace{v^T a}_{\mathbb{R} \Rightarrow \langle v, a\rangle}$$

$\boxed{\text{Ответ: } \langle u, v\rangle - \frac{\langle u, a\rangle\langle v, a\rangle}{2 + \|a\|^2}}$

(c) $\sum_{i=1}^{n} \langle S^{-1} a_i, a_i \rangle$, где $a_i \in \mathbb{R}^d$, $i = \overline{1, n}$;

$S = \sum_{i=1}^{n} a_i a_i^T$, $\det(S) \neq 0$;

Решение

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

$= \sum_{i=1}^{n} a_i a_i^T \equiv S \Rightarrow S$ — симметричная матрица

2) $\sum_{i=1}^{n} \langle S^{-1} a_i, a_i \rangle = \sum_{i=1}^{n} tr(a_i^T S^{-1} a_i) =$ {цикл. св-во tr.} =

$\langle S^{-1} a_i, a_i \rangle = ((S^{-1} a_i)^T a_i) = a_i^T \underbrace{(S^{-1})^T}_{\text{т.к. } S - \text{симм.} \Rightarrow S^{-1} - \text{симм.}} a_i$

Это число, поэтому можем взять от него trace, ничего не изменится

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

$$= tr\Big(\underbrace{\Big(\sum_{i=1}^{n} a_i a_i^T\Big)}_{S} S^{-1}\Big) =$$

$$= tr(I_d) = d$$

$\boxed{\text{Ответ: } d}$

3 Найти первую и вторую производные функций:

(a) $f: E \to \mathbb{R}$, $f(t) = \det(A - tI_n)$; где $A \in \mathbb{R}^{n\times n}$,
$E = \{t \in \mathbb{R} : \det(A - tI_n) \neq 0\}$;

Решение

1) $df[dt] = d(\det(A - tI_n)) = \det(A - tI_n)\,\mathrm{tr}((A - tI_n)^{-1} \cdot$

$\cdot \underbrace{d(A - tI_n)}_{dA - d(tI_n)}) = \det(A - tI_n)\,\mathrm{tr}((A - tI_n)^{-1}\underbrace{(-dt)}_{\in\mathbb{R}} I_n) =$

$dA = 0$, $d(tI_n) = \underbrace{(dt)}_{\in\mathbb{R}} I_n$

$= \underbrace{-\det(A - tI_n)\,\mathrm{tr}((A - tI_n)^{-1})}_{f'(t)}\,dt$

2) $d^2 f(t)[dt_1, dt_2] = d(df(t)[\underset{\text{фикс}}{dt_1}])(t)[dt_2] =$

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

$= -\underbrace{dt_1}_{= dt_1^T}\Big[-\det(A - tI_n)\,\mathrm{tr}((A - tI_n)^{-1})\underbrace{dt_2}_{\in\mathbb{R}}\,\mathrm{tr}((A - tI_n)^{-1}) +$

$+ \det(A - tI_n)\,\mathrm{tr}(\underbrace{d((A - tI_n)^{-1})}_{-(A - tI_n)^{-1} d(A - tI_n)(A - tI_n)^{-1}})\Big] =$

$d(A - tI_n) = \underbrace{(-dt_2)}_{\in\mathbb{R}} I_n$

$= dt_1^T \cdot \underbrace{\det(A - tI_n) \cdot \left[\mathrm{tr}^2((A - tI_n)^{-1}) - \mathrm{tr}(((A - tI_n)^{-1})^2)\right]}_{f''(t)} \cdot dt_2$

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

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

1) Рассм. $d(\|x\|)$, где $x \in \mathbb{R}^n \Rightarrow$

$$\Rightarrow d(\|x\|) = d((x^T x)^{\frac{1}{2}}) = \frac{1}{2}(x^T x)^{-\frac{1}{2}} \underbrace{d(x^T x)}_{} \ominus$$

$$\ominus \frac{1}{2}(x^T x)^{-\frac{1}{2}} \cdot 2x^T dx = \underline{\underline{(x^T x)^{-\frac{1}{2}} x^T dx}}, \quad x^T \underbrace{(I + I^T)}_{2I} dx$$

Т.е. $\Rightarrow \underline{\underline{\frac{1}{\|x\|} x^T dx}}$

$$df = df(t)[dt] = d(\|(A+tI_n)^{-1} \cdot b\|) = \frac{1}{\|(A+tI_n)^{-1}b\|} \cdot$$

$$\cdot \underbrace{((A+tI_n)^{-1}b)^T}_{\left(\begin{array}{cc} b^T(A^T + I_n^T t^T)^{-1} & \\ \parallel & \parallel \\ A,\ \text{т.к.}\ \in S^n_+ & t,\ \text{т.к.}\ \in \mathbb{R} \end{array}\right)} \underbrace{d((A+tI_n)^{-1}b)}_{\underbrace{d((A+tI_n)^{-1})}_{-(A+tI_n)^{-1}\underbrace{d(A+tI_n)}_{dt}(A+tI_n)^{-1}}b + (A+tI_n)^{-1}\underbrace{db}_{\to 0}} \ominus$$

$$\ominus \frac{-1}{\|(A+tI_n)^{-1}b\|} \cdot b^T(A+tI_n)^{-1}(A+tI_n)^{-1}\underbrace{dt}_{\in \mathbb{R}}(A+tI_n)^{-1}b$$

$$\Rightarrow \underbrace{-\frac{b^T((A+tI_n)^{-1})^3 b}{\|(A+tI_n)^{-1}b\|}}_{f'(t)} dt \Rightarrow df = \underbrace{-\frac{b^T(A+tI_n)^{-3}b}{f}}_{f'(t)} dt$$

2) $d^2 f(t)[dt_1, dt_2] = d(df(t)[\underset{\text{фикс}}{dt_1}])(t)[dt_2] =$

$$= d\left(-\frac{b^T((A+tI_n)^{-1})^3 b}{\|(A+tI_n)^{-1}b\|} \underbrace{dt_1}_{\substack{\in \mathbb{R} \\ \text{и } = dt_1^T}}\right) = dt_1^T\left[-\frac{d(b^T(A+tI_n)^{-3}b) \cdot f - b^T(A+tI_n)^{-3}b \cdot df}{\underbrace{\|(A+tI_n)^{-1}b\|^2}_{f^2}}\right] \ominus \left\{ d(b^T(A+tI_n)^{-3}b) = \underbrace{(db^T)}_{0} \cdot (A+tI_n)^{-3}b + \right.$$

$$+\ b^T \cdot \underbrace{d\left((A+tI_n)^{-3} b\right)}_{\underbrace{d(A+tI_n)^{-3}}_{-3(A+tI_n)^{-4}\underbrace{d(A+tI_n)}_{\underbrace{dt_2}_{\in \mathbb{R}}}} \cdot b \,+\, (A+tI_n)^{-3} \cdot \underbrace{db}_{0}} = -3 b^T (A+tI_n)^{-4} b\, dt_2 \Big\}$$

$$\equiv dt_1^T \cdot \frac{1}{f^2} \cdot \Big[3 b^T (A+tI_n)^{-4} b f - b^T (A+tI_n)^{-3} b \cdot b^T (A+tI_n)^{-3} \cdot$$

$$\cdot b \cdot \frac{1}{f} \Big] dt_2 \Rightarrow f''(t) = \frac{1}{f} \cdot 3 b^T (A+tI_n)^{-4} b \, -$$

$$- \underbrace{b^T (A+tI_n)^{-3} b}_{\in \mathbb{R}} \cdot \underbrace{b^T (A+tI_n)^{-3} b}_{\in \mathbb{R}} \cdot \frac{1}{f^3} =$$

$$= \frac{3 \operatorname{tr}\left(b^T (A+tI_n)^{-4} b\right)}{f} - \frac{\operatorname{tr}^2\left(b^T (A+tI_n)^{-3} b\right)}{f^3};$$

Ответ: $f'(t) = -\dfrac{b^T (A+tI_n)^{-3} b}{f}$

$$f''(t) = \frac{3 \operatorname{tr}\left(b^T (A+tI_n)^{-4} b\right)}{f} - \frac{\operatorname{tr}^2\left(b^T (A+tI_n)^{-3} b\right)}{f^3}$$

$\boxed{4}$ Найти ∇f и $\nabla^2 f$ для функций:

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

1) $df(x)[dx] = d\left(\frac{1}{2} \| x x^T - A \|_F^2\right) \Rightarrow \{ x x^T = (x x^T)^T \}$

~~[illegible]~~

$\Rightarrow$

$$\Rightarrow \frac{1}{2} d(\langle xx^T - A, xx^T - A\rangle) = \frac{1}{2} d(\overbrace{\langle xx^T, xx^T\rangle}^{tr(xx^Txx^T)} -$$

$$- 2\langle xx^T, A\rangle - \langle A, A\rangle) = \frac{1}{2}\Big[tr(d(xx^T)xx^T +$$

$$+ xx^T d(xx^T)) - 2d\, tr(xx^TA) - \underbrace{d\langle A, A\rangle}_{0}\Big] =$$

$$= \frac{1}{2}\Big[tr(d(xx^T)xx^T) + tr(xx^Td(xx^T)) - 2tr(\underbrace{d(x^TAx)}_{\substack{x^T(A+A^T)dx \\ \| \\ A,\text{ т.к. } \in S^n}})\Big] =$$

$$= tr(xx^T \underbrace{d(xx^T)}_{dx\cdot x^T + x\cdot dx^T}) - tr(2x^TA\,dx) \ominus$$

$$\ominus \underbrace{tr(xx^Tdx\cdot x^T)}_{tr(x^Txx^Tdx)} + \underbrace{tr(xx^Txdx^T)}_{tr(dx^Txx^Tx)} - tr(2x^TA\,dx) \ominus$$

$$\underbrace{dx^Tx}_{\in\mathbb{R}} \Rightarrow = (dx^Tx)^T = x^Tdx$$

$$\Rightarrow tr(x^Tdx\,x^Tx) = tr(x^Txx^Tdx)$$

$$\ominus 2tr(\underbrace{x^Tx}_{\mathbb{R}}\underbrace{x^Tdx}_{\mathbb{R}}) - 2tr(\underbrace{x^TAdx}_{\mathbb{R}}) = \{\text{снимаем trace с чисел}\} =$$

$$= 2(x^Txx^T - x^TA)dx = \underbrace{2x^T(xx^T - A)}_{\nabla f^T(x)}dx$$

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

$$2)\ d^2f(x)[dx_1, dx_2] = d(df(x)[\underset{\text{фикс}}{dx_1}])(x)[dx_2] =$$

$$\Rightarrow d(2\cdot x^T(xx^T - A)dx_1) = 2d(\underbrace{x^T(xx^T - A)dx_1}_{\in\mathbb{R} \Rightarrow = (x^T(xx^T-A)dx_1)^T}) =$$

$$= 2d(dx_1^T\underbrace{(xx^T - A)^T}_{(xx^T - A)}x) \ominus$$

$$\ominus dx_1^T\cdot\Big[2\cdot d((xx^T - A)x)\Big]; \text{ Рассм. } d((xx^T - A)x) \Rightarrow$$

$$\Rightarrow \underbrace{d(xx^T - A)}\cdot x + (xx^T - A)dx_2 \ominus$$

$$dxx^T \Rightarrow dx_2 \cdot x^T + \cancel{x}dx_2^T$$

$$\ominus\; dx_2 \underbrace{x^T x}_{\substack{\cap \\ \mathbb{R}}} + x\underbrace{dx_2^T x}_{\in \mathbb{R} \Rightarrow \text{Транспонируем}} + (xx^T - A)dx_2 \Rightarrow$$

$$\mathbb{R} \Rightarrow dx_2 \cdot (x^T x) = (x^T \cdot x)dx_2$$

$$\Rightarrow (x^T x)dx_2 + xx^T dx_2 + xx^T dx_2 - A dx_2 =$$

$$= [\underbrace{x^T x}_{\|x\|^2} I + 2xx^T - A]dx_2$$

$$\Rightarrow d^2 f(x)[dx_1, dx_2] = dx_1^T \cdot \underbrace{\left[2\cdot\left(\|x\|^2 I + xx^T + (xx^T - A)\right)\right]}_{\nabla^2 f(x)} dx_2$$

$\Rightarrow$ **Ответ:** 1) $\nabla f(x) = 2(xx^T - A)x = 2(x\underbrace{x^T x}_{\|x\|} - Ax) =$

$$= 2(\|x\|^2 \cdot x - Ax) \Longrightarrow \underline{\underline{2(\|x\|^2 I - A)x}}$$

2) $\nabla^2 f(x) = \underline{\underline{2(\|x\|^2 I + 2xx^T - A)}}$

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

1) $df(x)[dx] = d\left(\langle x,x\rangle^{\langle x,x\rangle}\right) = d\left(e^{\langle x,x\rangle \ln\langle x,x\rangle}\right) =$

$$= e^{\langle x,x\rangle \ln\langle x,x\rangle} d(\underbrace{\langle x,x\rangle}_{x^T x} \ln \underbrace{\langle x,x\rangle}_{x^T x}) \ominus \Big\{ d(\|y\|) = d\left((y^T y)^{\frac{1}{2}}\right) =$$

$$= \frac{1}{2}\cdot(y^T y)^{-\frac{1}{2}} d(y^T y) = \frac{1}{2\|y\|}[\underbrace{dy^T \cdot y}_{\in \mathbb{R} \Rightarrow \text{Транспонируем}} + y^T dy] = \frac{1}{\|y\|} y^T dy \text{ — на всякий случай}\Big\}$$

$$= \langle x,x\rangle^{\langle x,x\rangle}\Big(\underbrace{d(x^T x)}_{\in\mathbb{R}}\ln(x^T x) + \cancel{x^T x}\cdot\frac{1}{\cancel{x^T x}}\underbrace{d(x^T x)}_{\in\mathbb{R}}\Big) =$$

$$= \langle x,x\rangle^{\langle x,x\rangle}(\ln\langle x,x\rangle + 1)\underbrace{d(x^T x)}_{\underbrace{dx^T\cdot x}_{\in\mathbb{R} \Rightarrow \text{транспонируем}} + x^T dx = 2x^T dx} \Rightarrow$$

$$\Rightarrow df(x)[dx] = \underbrace{\underbrace{2\langle x,x\rangle^{\langle x,x\rangle}(\ln\langle x,x\rangle + 1)}_{\in\mathbb{R}} x^T}_{\nabla^T f(x)} dx \Rightarrow$$

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

2) $$d^2 f(x)[dx_1, dx_2] = d\Big(\underbrace{2\langle x,x\rangle^{\langle x,x\rangle}(\ln\langle x,x\rangle + 1)}_{\in\mathbb{R}}\underbrace{x^T dx_1}_{\in\mathbb{R}}\Big)$$

Транспонирование

$$\Rightarrow dx_1^T\big(2\cdot d[x e^{\langle x,x\rangle \ln\langle x,x\rangle}(\ln\langle x,x\rangle+1)]\big) \Rightarrow$$

$\Rightarrow$ Рассм. $\Rightarrow$ $$dx_2\cdot \underbrace{e^{\langle x,x\rangle\ln\langle x,x\rangle}(\ln\langle x,x\rangle+1)}_{\in\mathbb{R}} + x\cdot d\big(e^{\langle x,x\rangle\ln\langle x,x\rangle}\cdot$$

$$\cdot(\ln\langle x,x\rangle+1)\big);$$ Рассм. $$d(e^{\langle x,x\rangle\ln\langle x,x\rangle})\overbrace{(\ln\langle x,x\rangle+1)}^{\in\mathbb{R}} +$$

$$+ e^{\langle x,x\rangle\ln\langle x,x\rangle}\underbrace{d(\ln\langle x,x\rangle+1)}_{\frac{1}{\langle x,x\rangle}\cdot \underbrace{dx^T x}_{2x^T dx_2}} = 2\langle x,x\rangle^{\langle x,x\rangle}(\ln\langle x,x\rangle+1)^2 x^T dx_2 +$$

$$+ \langle x,x\rangle^{\langle x,x\rangle}\frac{1}{\langle x,x\rangle}\cdot 2x^T dx_2 = 2\langle x,x\rangle^{\langle x,x\rangle-1}\big((\ln\langle x,x\rangle+1)^2\cdot$$

$$\cdot\langle x,x\rangle + 1\big)x^T dx_2 \Big| \Rightarrow d^2 f(x) = dx_1^T\cdot\Big[2\cdot\big(\langle x,x\rangle^{\langle x,x\rangle}\cdot$$

$$\cdot(\ln\langle x,x\rangle+1)\cdot I + 2\langle x,x\rangle^{\langle x,x\rangle}(\ln\langle x,x\rangle+1)^2\cdot xx^T +$$

$$+ 2\langle x,x\rangle^{\langle x,x\rangle-1}\cdot xx^T)\Big]dx_2 \Rightarrow$$

$$\Rightarrow \nabla^2 f(x) = 2\langle x,x\rangle^{\langle x,x\rangle}\Big(2(\ln\langle x,x\rangle+1)^2 xx^T + (\ln\langle x,x\rangle+1)\cdot$$

$$\cdot I + \frac{2}{\langle x,x\rangle} xx^T\Big)$$

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

2) $\nabla^2 f(x) = 2\langle x,x\rangle^{\langle x,x\rangle}\Big(2(\ln\langle x,x\rangle+1)^2 xx^T +$

$+(\ln\langle x,x\rangle+1)\cdot I_n + \frac{2}{\langle x,x\rangle}\cdot xx^T\Big)$

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

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

$-b\|) = \{d(\|c\|) = \frac{1}{\|c\|}c^T dc\} = p\|Ax-b\|^{p-2}(x^TA^T-b^T)\cdot$

$\cdot \underbrace{d(Ax-b)}_{Adx} \Rightarrow$ $\nabla f(x) = p\|Ax-b\|^{p-2}A^T(Ax-b)$

2) $d^2f(x)[dx_1, dx_2] = d\Big(p\|Ax-b\|^{p-2}\cdot\underbrace{(x^TA^T-b^T)Adx_1}_{\in\mathbb{R} \Rightarrow \text{транспонируем}}\Big) =$

$\Rightarrow dx_1^T\cdot\Big[p\cdot\underbrace{d\Big(\underbrace{\|Ax-b\|^{p-2}}_{\in\mathbb{R}}A^T(Ax-b)\Big)}_{\text{рассм. выр-е}}\Big];$

$\Rightarrow d(\|Ax-b\|^{p-2})A^T(Ax-b) + \|Ax-b\|^{p-2}d(A^T(Ax-b)) \Rightarrow$

$$\Rightarrow \left\{ d(\|Ax-b\|^{p-2}) = (p-2)\|Ax-b\|^{p-3} \cdot \frac{1}{\|Ax-b\|} \cdot (Ax-b)^T \cdot A dx_2 \right\}$$

$$\Rightarrow \left\{ d(A^T(Ax-b)) = \underbrace{dA^T}_{0} \cdot (Ax-b) + A^T d(Ax-b) = A^T A dx_2 \right\}$$

$$\Rightarrow (p-2)\|Ax-b\|^{p-4} \underbrace{(Ax-b)^T A dx_2}_{\in \mathbb{R}} \underbrace{A^T(Ax-b)}_{\in \mathbb{R}^{n\times 1}} + \|Ax-b\|^{p-2} \cdot \underbrace{A^T A}_{\in \mathbb{R}^{n\times n}} \underbrace{dx_2}_{\in \mathbb{R}^{n\times 1}}$$

$$\Longrightarrow \Big[(p-2)\|Ax-b\|^{p-4} A^T (Ax-b)(Ax-b)^T A +$$
$$+ \|Ax-b\|^{p-2} A^T A \Big] dx_2;$$

$$\Rightarrow d^2 f(x)[dx_1, dx_2] = dx_1^T \Big[p\|Ax-b\|^{p-2} A^T \big((p-2)\|Ax-b\|^{p-2} \cdot (Ax-b) \cdot$$
$$\cdot (Ax-b)^T + I \big) A \Big] dx_2;$$

$$\Rightarrow \nabla^2 f(x) = p\|Ax-b\|^{p-2} A^T \Big[(p-2)\|Ax-b\|^{p-2} \cdot (Ax-b)(Ax-b)^T + I_m \Big] A$$

Ответ: 1) $\nabla f(x) = p\|Ax-b\|^{p-2} A^T (Ax-b)$

2) $\nabla^2 f(x) = p\|Ax-b\|^{p-2} A^T \Big[(p-2)\|Ax-b\|^{p-2} \cdot (Ax-b) \cdot$
$\cdot (Ax-b)^T + I_m \Big] A$

$\boxed{5}$ Для ф-ций показать, что $d^2 f$ — знакоопределённый (имеет пост. знак) и установить этот знак:

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

1) $\Rightarrow df(x)[dx] = d(\operatorname{tr}(X^{-1})) = \operatorname{tr}(dX^{-1}) = -\operatorname{tr}(X^{-1} \cdot$
$\cdot dX \cdot X^{-1}) = -\operatorname{tr}(X^{-2} dX) \Rightarrow \nabla f = -X^{-2T} =$

$= \{X \in S^n_{++}\} = -X^{-2};$

2) $d^2 f(X)[dX_1, dX_2] = d(-tr(X^{-2} dX_1)) =$

$= -tr(\underbrace{d(X^{-2} dX_1)}) =$

$\left(d(X^{-1} \cdot X^{-1} \cdot dX_1) = d(X^{-1}) X^{-1} dX_1 + X^{-1} d(X^{-1} dX_1) \right.$

$\left. \Rightarrow -X^{-1} dX_2 X^{-1} X^{-1} dX_1 + X^{-1} [dX^{-1} dX_1 + X^{-1} \underbrace{d(dX_1)}_{0}] \right)$

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

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

[illegible]

Рассм. $tr(X^{-2} dX_2 X^{-1} dX_1) \Rightarrow$

$\overset{\shortparallel}{X^{-2T}}$ $\overset{\shortparallel}{dX_2^T}$ $\overset{\shortparallel}{X^{-T}}$ $\overset{\shortparallel}{dX_1^T}$, т.к. все эти матрицы $\in S^n_{++}$

$\Rightarrow$ это равно $tr((dX_1 X^{-1} dX_2 X^{-2})^T) \equiv tr(dX_1 X^{-1} dX_2 X^{-2})$

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

Рассм. $d^2 f(X)[H, H] = 2 tr(X^{-\frac{1}{2}} X^{-\frac{1}{2}} H X^{-1} X^{-1} H) =$

$= 2 tr(X^{-\frac{1}{2}} H X^{-1} X^{-1} H X^{-\frac{1}{2}}) = \underline{\underline{2 \|X^{-1} H X^{-\frac{1}{2}}\|_F^2 > 0}},$

$\Rightarrow (X^{-1} H X^{-\frac{1}{2}})^T$

(т.к. это $X^{-\frac{1}{2}T} H^T X^{-T} = X^{-\frac{1}{2}} H X^{-1}$)

т.к. $X^{-1} H X^{-\frac{1}{2}} \neq 0$,

ведь $|X^{-1} H X^{-\frac{1}{2}}| =$

$= \frac{1}{|X|^{3/2}} |H| \neq 0$, $\underset{0}{\cancel{=}}$

т.к. $\frac{1}{det(X)^{3/2}} \neq 0$

$\boxed{\underline{\text{Ответ}}: d^2 f(X) > 0}$

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

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

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

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

$= \frac{1}{n}\left[d\left((\det X)^{\frac{1}{n}}\right) tr(X^{-1}dX) - (\det X)^{\frac{1}{n}} tr(X^{-1}dX_2X^{-1}dX_1)\right.$

$\Rightarrow \frac{1}{n}\left[\frac{1}{n}(\det(X))^{\frac{1}{n}} tr(X^{-1}dX_2) tr(X^{-1}dX_1) - (\det(X))^{\frac{1}{n}} \cdot\right.$

$\left.\cdot\, tr(X^{-1}dX_2X^{-1}dX_1)\right] \Longrightarrow$ Рассм. $d^2f(X)[H, H] \Rightarrow$

$$\Rightarrow \frac{1}{n}(\det(X))^{\frac{1}{n}}\left(\frac{1}{n}tr^2(X^{-1}H) - \underbrace{tr(X^{-1}H)^2}_{tr(X^{-1}HX^{-1}H)}\right) = \frac{1}{n}(\det X)^{\frac{1}{n}} \cdot$$

$$\cdot \left[\frac{\langle X^{-1}, H\rangle^2}{n} - \langle HX^{-1}, X^{-1}H\rangle\right] \Rightarrow$$

$\Rightarrow$ Рассм. скобку $\Rightarrow \left(\frac{1}{n}\langle HX^{-1}, I_n\rangle^2 - \langle HX^{-1}, X^{-1}H\rangle\right) \le$

$$\underset{\text{(нер-во Коши-Буняковского)}}{\le} \frac{1}{n}\cdot \|HX^{-1}\|^2 \cdot \underbrace{\|I\|^2}_{n} - \|HX^{-1}\|^2 = 0$$

Ответ: $d^2f(X)[H,H] \le 0$, т.е. $\begin{cases} = 0, \text{ при } X = H \\ < 0, \text{ при } X \ne H \end{cases}$

$\boxed{6}$ Найти точки стационарности; знач. пар-ров в них;

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

$$f(x) = c^T x + \frac{\sigma}{3}\underbrace{\langle x, x\rangle}_{x^Tx}^{\frac{3}{2}}$$

$$df(x) = c^T dx + \frac{\sigma}{3}\cdot\frac{3}{2}\cdot\underbrace{(x^Tx)^{\frac{1}{2}}}_{\|x\|}\underbrace{d(x^Tx)}_{2x^Tdx} \Rightarrow$$

$$\Rightarrow df(x) = \underbrace{(c^T + \sigma\|x\|x^T)}_{\nabla^T f(x)} dx \Rightarrow \nabla f(x) = c + \sigma\|x\|x = 0$$

$$\Rightarrow x\|x\| = -\frac{c}{\sigma} \Rightarrow \left\{ \text{выразим } \|x\|: \underbrace{\|x\|x\|\|}_{\|x\|^2} = \frac{\|c\|}{\sigma} \Rightarrow \right.$$

$$\left.\Rightarrow \|x\| = \sqrt{\frac{\|c\|}{\sigma}}\right\} \Rightarrow x = -\frac{c\cdot\sqrt{\sigma}}{\sigma\cdot\sqrt{\|c\|}} = -\frac{c}{\sqrt{\sigma\|c\|}} \Rightarrow$$

$\Rightarrow$ **Ответ:**
$\exists!$ стац. т. $\Rightarrow x = -\dfrac{c}{\sqrt{\|c\|_B}}$
(при усл. задачи: $c \neq 0$, $B \neq 0$)

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

$\Downarrow$

$$f(x) = a^T x - \ln(1 - b^T x)$$

$$df = a^T dx - \frac{1}{1 - b^T x}(-b^T dx) =$$

$$= \underbrace{\left(a^T + \frac{b^T}{1 - b^T x}\right)}_{\nabla^T f} dx \Rightarrow \quad \nabla f(x) = a + \frac{1}{1 - b^T x} \cdot b = 0$$

~~$\Rightarrow$ {векторы $a, b \Rightarrow \parallel \Rightarrow \exists \alpha: a = \alpha b$, $\alpha \in \mathbb{R} \setminus \{0\}$} $\Rightarrow$~~

~~$\Rightarrow \alpha b + \frac{b}{1 - b^T x} = 0 \Rightarrow$ {$b \neq 0$} сокращаем $\Rightarrow 1 - b^T x = -\frac{1}{\alpha}$~~

~~$\Rightarrow$ {$b^T x < 1$ по усл.} $\Rightarrow -\frac{1}{\alpha} > 0 \Rightarrow \alpha < 0$~~

~~Ответ: мн-во точек стационарности задаётся ур-ем $\Rightarrow$~~

$$\Rightarrow -a + a b^T x = b \Rightarrow b = -a + \langle b, x\rangle a$$

$$\Rightarrow \langle b, x\rangle a = b + a \Rightarrow \langle b, x\rangle \underbrace{\langle a, a\rangle}_{\|a\|^2} = \langle b, a\rangle + \underbrace{\langle a, a\rangle}_{\|a\|^2}$$

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

$$\left(\frac{\langle b, b\rangle}{\|b\|^2} = \left\langle \frac{b}{\|b\|^2}, b\right\rangle\right) \Rightarrow$$

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

$\Rightarrow$ **Ответ:** $\boxed{x = \frac{a}{\|a\|^2} + \frac{b}{\|b\|^2}}$ $(a, b \neq 0$ по усл.)

$\exists!$ точка стационарности

(C) $f: \mathbb{R}^n \to \mathbb{R}$, $f(x) = \langle c, x\rangle \exp(-\langle Ax, x\rangle)$,

где $c \in \mathbb{R}^n$, $c \neq 0$, $A \in S^n_{++}$;

$df = d(c^T x \cdot e^{-x^TAx}) = c^T dx \cdot e^{-x^TAx} + c^T x \cdot e^{-x^TAx} \cdot$

$\cdot \underbrace{d(-x^TAx)}_{-2x^TAdx} \Rightarrow \underbrace{(c^T e^{-x^TAx} - 2c^Txe^{-x^TAx}x^TA)}_{\nabla^T f(x)} dx$

$\Rightarrow \nabla f(x) = \underbrace{e^{-x^TAx}}_{>0} \cdot \underbrace{(c - 2Axx^Tc)}_{} = 0$

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

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

$\Rightarrow c^Txx^Tc = c^TA^{-1}\cdot\frac{1}{2}\cdot c \Rightarrow \langle x^Tc, x^Tc\rangle = \langle x, c\rangle^2 =$

$= \frac{1}{2}\cdot c^TA^{-1}c \Rightarrow x^Tc = \pm\sqrt{\frac{c^TA^{-1}c}{2}}$

$\Rightarrow$ возвращаемся к (*):

$\pm x\sqrt{\frac{c^TA^{-1}c}{2}} = \frac{A^{-1}c}{2} \Rightarrow$

$\Rightarrow x = \pm\sqrt{\frac{2}{c^TA^{-1}c}}\frac{A^{-1}c}{2} \Rightarrow$ **Ответ:** $x = \pm\frac{1}{\sqrt{2c^TA^{-1}c}}A^{-1}c$

$\exists$ 2 стац. точки $\forall c, A$, удовл. усл-ю

$\boxed{7}$ $X \in S^n_{++}$; $\lim\limits_{k\to+\infty}\left[tr\left(X^{-k} - (X^k + X^{2k})^{-1}\right)\right] = ?$

= Решение =

1) $\Rightarrow \lim\limits_{k\to+\infty}\left[tr\left(X^{-k} - X^{-k}(I_n + X^k)^{-1}\right)\right] =$

$= \lim\limits_{k\to+\infty}\left[tr\left(X^{-k}\cdot\left(I_n - (I_n + X^k)^{-1}\right)\right)\right] =$

$= \left\{ \text{тожд-во Вудбери: } A = I_n;\ U, V = I_n;\ C = X^k \Rightarrow \Rightarrow (I_n + X^k)^{-1} = \underline{I_n - I_n(X^{-k} + I_n)^{-1}} \right\} =$

$= \lim\limits_{k\to+\infty}\left[tr\left(X^{-k}\cdot\left(\cancel{I_n} - \cancel{I_n} + I_n(X^{-k} + I_n)^{-1}\right)\right)\right] =$

$= \lim\limits_{k\to+\infty} tr\left(\underbrace{X^{-k}(X^{-k} + I_n)^{-1}}_{(I_n + X^k)^{-1}}\right) \Rightarrow \underline{\underline{\lim\limits_{k\to+\infty}\left[tr\left((I_n + X^k)^{-1}\right)\right]}}$

2) ~~[illegible]~~

$tr(I_n + X^k)^{-1} = \left\{ X^k = QD^kQ^T, \text{ где } D - \text{диаг.};\ QQ^T = Q^TQ = I \right\} =$

$= tr(QQ^T + QD^kQ^T)^{-1} =$

$= tr\left((Q\cdot(I_n + D^k)Q^T)^{-1}\right) = tr\left(Q^{-T}(I_n + D^k)Q^{-1}\right) =$

$= tr\left((I_n + D^k)^{-1}\underbrace{(Q^TQ)^{-1}}_{I_n}\right) = tr\left((I_n + D^k)^{-1}\right);$

$\underline{\underline{X \in S^n_{++}}} \Rightarrow \underline{\underline{\text{с.з. у } D^k > 0}}$

$\Rightarrow$ с.з. у $I_n + D^k > 1$;

$\Rightarrow \lim\limits_{k\to+\infty} tr\left(\underbrace{(I_n + D^k)^{-1}}_{diag\{\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}\}}\right)$

$\boxed{\begin{array}{l} \text{] } I_n + D^k = diag\{\lambda_1, \dots, \lambda_n\} \\ (\lambda_i = 1 + \sigma_i^k;\ \sigma_i - \text{с.з } X) \end{array}}$

$\Rightarrow$ Искомый предел $\Rightarrow \lim\limits_{k\to+\infty} \sum\limits_{i=1}^{n} \frac{1}{1+\sigma_i^k}$, т.е. ($\sigma_i$ — с.з. X)

$\Rightarrow$ Если 1) σ_i (какое-то из) < 1, то $\frac{1}{1+\sigma_i^k} \xrightarrow[k\to+\infty]{} 1$

2) σ_i (какое-то из) $= 1$, то $\frac{1}{1+\sigma_i^k} \xrightarrow[k\to+\infty]{} \frac{1}{2}$

3) σ_i (какое-то из) > 1, то $\frac{1}{1+\sigma_i^k} \xrightarrow[k\to+\infty]{} 0$

$\Rightarrow$ **Ответ:** Предел $= \#(\sigma_i < 1) + \frac{1}{2}\#(\sigma_i = 1)$,

где σ_i — с.з. матрицы X;

8 Задача оптимизации:

$$F(P) = \sum_{i=1}^{N} \|x_i - P(P^T P)^{-1} P^T x_i\|^2 =$$

$$= N\,tr((I - P(P^T P)^{-1} P^T)^2 S) \to \min_P$$

$\left(S = \frac{1}{N}\sum\limits_{i=1}^{N} x_i x_i^T\right.$ — выб. матр. ков. для норм. выборки$\left.\right)$

(a) Найти $\nabla_P F(P)$, если $P^T P = I$;

$\Downarrow$

$$\Rightarrow dF = N d(\langle (I - P(P^T P)^{-1} P^T)^2, S^T \rangle) =$$

$$= N \langle S^T, d(I - P(P^T P)^{-1} P^T) \cdot (I - P(P^T P)^{-1} P^T) + (I - P(P^T P)^{-1} \cdot$$

$$\cdot P^T) d(I - P(P^T P)^{-1} P^T) \rangle \ominus \{ d(P(P^T P)^{-1} P^T) =$$

$$= dP \cdot (P^T P)^{-1} P^T + P(d(P^T P)^{-1} \cdot P^T + (P^T P)^{-1} dP^T) =$$

$$= \{ P^T P = I, \text{ кроме } d(P^T P) \} \Rightarrow$$

$$\Rightarrow dP \cdot P^T + P dP^T - \cancel{P\, P^{-1}} P^{-T} \underbrace{(dP^T P)}_{(dP^T \cdot P + P^T dP)} P^{-1} \cancel{P^{-T}} P^T =$$

$$= (dP) P^T + P(dP^T) - (P^{-T} (dP^T) + P^T (dP) P^{-1}) \Rightarrow$$

$$\Rightarrow \underline{\underline{dP \cdot P^T + P \cdot dP^T - P(dP^T \cdot P + P^T \cdot dP) P^T}} \} =$$

$= \{$ Аналогично, в исх. выр-ии подставляем $P^T P = I \} \Rightarrow$

$$\ominus N \langle S^T, (dP - P(dP^T \cdot P + P^T \cdot dP) \underbrace{P^T (I - PP^T)}_{O_n} +$$

$$+ P \cdot dP^T \cdot (I - PP^T) + \underbrace{(I - PP^T) P}_{O_n} \cdot (dP^T - P(dP^T \cdot P +$$

$$+ P^T \cdot dP) P^T + (I - PP^T) \cdot dP \cdot P^T \rangle \Rightarrow$$

$$\Rightarrow -N \langle S^T, P dP^T (I - PP^T) + (I - PP^T) dP P^T \rangle =$$

$$= -N (\langle SP, (I - PP^T)^T dP \rangle + \langle dP, (I - PP^T)^T S^T P \rangle) =$$

$$= \{ S - \underline{\text{симм. матрица}} \} = -N \langle dP, (I - PP^T) \underbrace{(S^T + S)}_{2S} P \rangle \Big| \Rightarrow$$

$\Rightarrow dF = -2N\langle (I - PP^T)SP, dP\rangle \Rightarrow$

Ответ: $\Rightarrow \boxed{\nabla_P F(P) = -2N(I_D - PP^T)SP;}$

(b) 1) Д-ть, что $\nabla_P F(P) = 0$ для P, сост. из $\forall$ d разл. собст. век. q_i по столбцам, где Q: $S = Q\underset{\text{диаг с с.з.}}{\Lambda} Q^T$, $Q = [q_1|\dots|q_D] \in \mathbb{R}^{D\times D}$

2) И что минимум $F(P) \Rightarrow$ для P, сост. из q_i, соотв наиб. с.з.

= Решение =

1) $S = Q\Lambda Q^T$ — подставляем в $\nabla_P F(P)$

$\Rightarrow$ $\nabla_P F(P) = -2N(I - PP^T)Q\Lambda Q^T P$; $P = [q_{i_1}|\dots|q_{i_d}]$

$\cancel{\nabla_P F(P) = -2N[Q - PP^T Q]\Lambda Q^T P}$

$\exists\, q_{k_i} \sim \lambda_{k_i} \Rightarrow P = (q_{k_1}|\dots|q_{k_d}) \in \mathbb{R}^{D\times d}$; $1 \le k_1 < \dots < k_d \le D$

$\Rightarrow \nabla_P F(P) = 2N(PP^TSP - SP) = 2N(PP^T(\lambda_{k_1}q_{k_1}|\dots|\lambda_{k_d}q_{k_d}) - (\lambda_{k_1}q_{k_1}|\dots|\lambda_{k_d}q_{k_d}))$, т.к. P сост. из собств. век. Q. ($S = Q\Lambda Q^T$)

$\Rightarrow \{P^T(\lambda_{k_1}q_{k_1}|\dots|\lambda_{k_d}q_{k_d}) = \operatorname{diag}\{\lambda_{k_1},\dots,\lambda_{k_d}\}\} \Leftrightarrow$

$\Leftrightarrow 2N(\underbrace{P\cdot \operatorname{diag}\{\lambda_{k_1},\dots,\lambda_{k_d}\}}_{(\lambda_{k_1}q_{k_1}|\dots|\lambda_{k_d}q_{k_d})} - (\lambda_{k_1}q_{k_1}|\dots|\lambda_{k_d}q_{k_d})) = 0$

2) Теперь:

$$F(P) = N tr((I - P(P^T P)^{-1} P^T)^2 S),$$

$$P = [q_{k_1} | \dots | q_{k_d}]; \quad 1 \le k_1 < \dots < k_d \le D$$

q_{k_i} – с.в.
λ_{k_i} – с.з.
$q_{k_i} \sim \lambda_{k_i}$

$$F(P) = \{P^T P = I_d\} =$$

$$= N tr((I - PP^T)^2 S) =$$

$$\left(I - 2PP^T + P\underbrace{P^T P}_{I}P^T = I - PP^T \right)$$

$$= N tr(S - PP^T S) = N[tr S - tr(PP^T S)] =$$

$$= N\left[tr S - tr\left(\begin{bmatrix} q_{k_1}^T \\ \dots \\ q_{k_d}^T \end{bmatrix} \cdot [\lambda_{k_1} q_{k_1} | \dots | \lambda_{k_d} q_{k_d}]\right)\right] =$$

$$= N \cdot \left(\sum_{i=1}^{D} \lambda_i - \sum_{i=1}^{d} \lambda_i\right) \longrightarrow \min_{\lambda_i}$$

$\Rightarrow$ min при ~~[illegible]~~ max выр-е $\sum_{i=1}^{d} \lambda_i$, т.е. при max. λ_i, при max с.з.

$\Rightarrow$ Ответ: $F(P)$ достигает min при матрице P, кот. сост. из собств. вект.-столбцов q_i, кот. отвечают наиб. собст. знач. матр. S.

(a), (b) ч.т.д.