1) Д-ть тождество Вудбери:

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

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

▲ Покажем по определению, что

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

$$(A+UCV)(A^{-1} - A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}) =$$
$$= I - U(C^{-1}+VA^{-1}U)^{-1}VA^{-1} + UCVA^{-1} -$$
$$- UCVA^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1} =$$
$$= -(U+UCVA^{-1}U)(C^{-1}+VA^{-1}U)^{-1}VA^{-1} +$$
$$+ I + UCVA^{-1} = I + UCVA^{-1} - UC(C^{-1}+$$
$$+ VA^{-1}U)(C^{-1}+VA^{-1}U)^{-1}VA^{-1} =$$
$$= I + UCVA^{-1} - UCVA^{-1} = I$$ – ч.т.д. ■

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

$$\|uv^T - A\|_F^2 - \|A\|_F^2 = tr[(uv^T-A)^T(uv^T-A)] -$$
$$- tr(A^TA) = tr[(vu^T - A^T)(uv^T-A)] -$$
$$- tr(A^TA) = tr(vu^Tuv^T - A^Tuv^T - vu^TA +$$
$$+ A^TA - A^TA) = tr(u^Tuv^Tv) - tr(vu^TA +$$
$$+ vu^TA) = \|u\|^2\|v\|^2 - 2tr(vu^TA)$$

в) $a, u, v \in \mathbb{R}^d$

$$\operatorname{tr}\left[(2I_n + aa^T)^{-1}(uv^T + vu^T)\right] = \left\{ \begin{matrix} \text{тождество} \\ \text{Вудбери} \end{matrix} \right\} =$$

$$= \operatorname{tr}\left[\left(\frac{1}{2}I_n - \frac{1}{2}I_n a \left(I_1 + \frac{1}{2}a^T a\right)^{-1} a^T \cdot \frac{1}{2} I_n\right)\right.$$

$$\left.(uv^T + vu^T)\right] = \left\{ \; b = I_1 + \frac{1}{2}a^T a, \; b \in \mathbb{R} \right\} =$$

$$= \operatorname{tr}\left[\left(\frac{1}{2}I_n - \frac{1}{4b}aa^T\right)(uv^T + vu^T)\right] =$$

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

$$= \frac{1}{2}\operatorname{tr}(uv^T) + \frac{1}{2}\operatorname{tr}(uv^T) - \frac{1}{4b}\operatorname{tr}(vu^T aa^T) -$$

$$- \frac{1}{4b}\operatorname{tr}(vu^T aa^T) = \operatorname{tr}(uv^T) - \frac{1}{2b}\operatorname{tr}(vu^T aa^T) =$$

$$= \operatorname{tr}(uv^T) - \frac{1}{2b}u^T a \cdot a^T v = \operatorname{tr}(uv^T) -$$

$$- \frac{u^T a \cdot a^T v}{2\left(I_1 + \frac{1}{2}a^T a\right)} = \operatorname{tr}(uv^T) - \frac{u^T a \cdot a^T v}{2 + \|a\|^2}$$

$$= \operatorname{tr}(v^T u) - \frac{u^T a \cdot a^T v}{2 + \|a\|^2} = v^T u - \frac{u^T a \cdot a^T v}{2 + \|a\|^2}$$

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

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

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

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

$$= \operatorname{tr}(S^{-1} \cdot S) = \operatorname{tr}(I_d) = d$$

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

$f : E \longrightarrow \mathbb{R}$, $f(t) = \det(A - tI_n)$

$$df(t) = \det(A - tI_n) \cdot \operatorname{tr}[(A - tI_n)^{-1}(-I_n dt)] =$$

$$= -\det(A - tI_n) \operatorname{tr}[(A - tI_n)^{-1}] dt$$

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

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

$$= -\left[d(\det(A-tI_n))\,tr\left[(A-tI_n)^{-1}\right] + \right.$$

$$\left. + \det(A-tI_n)\,d\left(tr\left[(A-tI_n)^{-1}\right]\right)\right] dt_1 =$$

$$= -\left[-\det(A-tI_n)\left(tr\left[(A-tI_n)^{-1}\right]\right)^2 dt_2 + \right.$$

$$+ \det(A-tI_n)\,tr\left[-(A-tI_n)^{-1}(-I_n dt_2)\cdot\right.$$

$$\left.\left.\cdot(A-tI_n)^{-1}\right]\right] dt_1 = \left[\det(A-tI_n)\cdot\right.$$

$$\left.\cdot\left(tr\left[(A-tI_n)^{-1}\right]\right)^2 - \det(A-tI_n)\cdot tr\left[(A-tI_n)^{-2}\right]\right]\cdot$$

$$\cdot dt_1\, dt_2$$

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

б) $f: \mathbb{R}_{++} \to \mathbb{R}$, $f(t) = \|(A+tI_n)^{-1}b\|$

$A \in S^n_+$, $b \in \mathbb{R}^n$

$$df(t) = d\langle (A+tI_n)^{-1}b, (A+tI_n)^{-1}b\rangle^{\frac{1}{2}} =$$

$$= \frac{1}{2}\|(A+tI_n)^{-1}b\|^{-1}\cdot 2\langle (A+tI_n)^{-1}b, d\left[(A+tI_n)^{-1}b\right]\rangle =$$

$$= \left\{ d\left[(A+tI_n)^{-1}b\right] = -(A+tI_n)^{-1} I_n\, dt\, (A+tI_n)^{-1}b = \right.$$

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

$$\cdot \langle (A+tI_n)^{-1}b, -(A+tI_n)^{-2}b\rangle dt =$$

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

$$= \{A \in S^n_+\} = -\|(A+tI_n)^{-1}b\|^{-1}\cdot tr[\underbrace{b^T(A+tI_n)^{-3}b}_{\in\mathbb{R}}]dt =$$

$$= -\|(A+tI_n)^{-1}b\|^{-1} \cdot b^T (A+tI_n)^{-3} b \, dt$$

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

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

$$= -[\, d(\|(A+tI_n)^{-1}b\|^{-1}) \cdot \underbrace{b^T (A+tI_n)^{-3} b}_{c} +$$

$$+ \|(A+tI_n)^{-1}b\|^{-1} d(b^T (A+tI_n)^{-3} b)] \, dt_1 =$$

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

$$\cdot (A+tI_n)^{-3} b \, c \, dt_2 + \|(A+tI_n)^{-1}b\|^{-1} b^T \cdot$$

$$\cdot (-3)(A+tI_n)^{-4} I_n \, dt_2 \cdot b] \, dt_1 =$$

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

$$- \|(A+tI_n)^{-1}b\|^{-3} (b^T (A+tI_n)^{-3} b)^2] \, dt_1 dt_2$$

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

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

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

$$df(x) = d[\tfrac{1}{2} \langle xx^T - A, xx^T - A \rangle] = \tfrac{1}{2} \cdot 2 \cdot$$

$$\cdot \langle xx^T - A, d(xx^T - A) \rangle = \langle xx^T - A, dx \cdot x^T + x \, d(x^T) \rangle =$$

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

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

$$= \langle xx^T x - Ax + xx^T x - A^T x, dx \rangle =$$

$= \{A \in S^n\} = \langle 2xx^Tx - 2Ax, dx\rangle =$

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

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

$d^2 f(x) = d\langle 2(xx^T - A)x, dx_1\rangle = \{A \in S^n\} =$

$= d\left[2x^T(xx^T - A)dx_1\right] = 2\left[(dx_2)^T(xx^T - A) +\right.$

$\left. + x^T\left(dx_2 \cdot x^T + x(dx_2)^T\right)\right] dx_1 =$

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

$\left. + x^T x (dx_2)^T dx_1\right] = \{(dx_2)^T(xx^T - A)dx_1 =$

$= (dx_1)^T(xx^T - A)dx_2$, т.к $(dx_2)^T(xx^T - A)dx_1 \in \mathbb{R}$;

$\underbrace{x^T dx_2}_{\in \mathbb{R}} \cdot \underbrace{x^T dx_1}_{\in \mathbb{R}} = (x^T dx_1)^T x^T dx_2 = (dx_1)^T x x^T dx_2$;

$\underbrace{x^T x}_{\in \mathbb{R}} \underbrace{(dx_2)^T dx_1}_{\in \mathbb{R}} = \cancel{tr(x^T x (dx_2)^T dx_1) =}$

$\cancel{= tr((dx_1)^T dx_2 x^T x) \; tr(dx_1 x^T x}$

$= \underbrace{(dx_2)^T (x^T x) dx_1}_{\in \mathbb{R}} = (dx_1)^T (x^T x) dx_2 \} =$

$= 2\left[(dx_1)^T(xx^T - A)dx_2 + (dx_1)^T xx^T dx_2 +\right.$

$\left. + (dx_1)^T(x^T x)dx_2\right] = 2(dx_1)^T\left[xx^T - A +\right.$

$\left. + xx^T + x^T x I_n\right] dx_2 = 2(dx_1)^T\left[2xx^T + x^T x I_n -\right.$

$\left. - A\right] dx_2$

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

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

$$df(x) = d\, e^{\langle x,x \rangle \ln \langle x,x \rangle} = e^{\langle x,x \rangle \ln \langle x,x \rangle} \cdot$$

$$\cdot \left[d(\langle x,x \rangle) \cdot \ln \langle x,x \rangle + \langle x,x \rangle \cdot d \ln \langle x,x \rangle \right] =$$

$$= \langle x,x \rangle^{\langle x,x \rangle} \left[2 \langle x, dx \rangle \ln \langle x,x \rangle + \right.$$

$$\left. + \langle x,x \rangle \frac{1}{\langle x,x \rangle} \cdot 2 \langle x, dx \rangle \right] = \langle x,x \rangle^{\langle x,x \rangle} \cdot$$

$$\cdot \langle 2x (\ln \langle x,x \rangle + 1), dx \rangle$$

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

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

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

$$\cdot 2\, d\left[(\ln \langle x,x \rangle + 1) x^T \right] dx_1 = \langle x,x \rangle^{\langle x,x \rangle} \cdot 4 \cdot$$

$$\cdot (\ln \langle x,x \rangle + 1)^2 \underbrace{x^T dx_2}_{\in \mathbb{R}} \cdot \underbrace{x^T dx_1}_{\in \mathbb{R}} + \langle x,x \rangle^{\langle x,x \rangle} \cdot$$

$$\cdot 2 \left[\frac{1}{\langle x,x \rangle} 2 \underbrace{x^T dx_2}_{\in \mathbb{R}} \underbrace{x^T dx_1}_{\in \mathbb{R}} + (\ln \langle x,x \rangle + 1) \cdot \right.$$

$$\left. \cdot (dx_2)^T dx_1 \right] = 2 \langle x,x \rangle^{\langle x,x \rangle} \left[2 (\ln \langle x,x \rangle + 1)^2 \cdot \right.$$

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

$$\left. + (\ln \langle x,x \rangle + 1) (dx_1)^T dx_2 \right] = 2 \langle x,x \rangle^{\langle x,x \rangle} (dx_1)^T \cdot$$

$$\cdot \left[2 (\ln \langle x,x \rangle + 1)^2 x x^T + \frac{2}{\langle x,x \rangle} x x^T + \right.$$

$$\left. + I_n (\ln \langle x,x \rangle + 1) \right] dx_2$$

$$\Rightarrow \nabla^2 f(x) = 2 \langle x,x \rangle^{\langle x,x \rangle} \left[2 (\ln \langle x,x \rangle + 1)^2 x x^T + \right.$$

$$\left. + \frac{2}{\langle x,x \rangle} x x^T + I_n (\ln \langle x,x \rangle + 1) \right]$$

c) $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $p \geq 2$

$f: \mathbb{R}^n \to \mathbb{R}$, $f(x) = \|Ax - b\|^p$

$$df(x) = d\langle Ax-b, Ax-b\rangle^{\frac{p}{2}} =$$

$$= \frac{p}{2}\langle Ax-b, Ax-b\rangle^{\frac{p}{2}-1} \cdot 2\langle Ax-b, A\,dx\rangle =$$

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

$$\Rightarrow \nabla f(x) = p\,\|Ax-b\|^{p-2} A^T(Ax-b)$$

$$d^2 f(x) = d\left(p\|Ax-b\|^{p-2}(x^TA^T - b^T)A\,dx_1\right) =$$

$$= p\left[d\left(\|Ax-b\|^{p-2}\right)(x^TA^T - b^T)A\,dx_1 +\right.$$

$$\left. + \|Ax-b\|^{p-2}\,d\left[(x^TA^T-b^T)A\right]dx_1\right] =$$

$$= p\left[(p-2)\|Ax-b\|^{p-4}\underbrace{(x^TA^T-b^T)A\,dx_2}_{\in\mathbb{R}} \cdot\right.$$

$$\left.\cdot \underbrace{(x^TA^T-b^T)A\,dx_1}_{\in\mathbb{R}} + \|Ax-b\|^{p-2}\underbrace{(dx_2)^TA^TA\,dx_1}_{\in\mathbb{R}}\right] =$$

$$= p\left[(p-2)\|Ax-b\|^{p-4}(dx_1)^TA^T(Ax-b)(x^TA^T-b^T)A\cdot\right.$$

$$\left.\cdot\, dx_2 + \|Ax-b\|^{p-2}(dx_1)^TA^TA\,dx_2\right] =$$

$$= p\,(dx_1)^T\left[(p-2)\|Ax-b\|^{p-4}A^T(Ax-b)(x^TA^T-b^T)A +\right.$$

$$\left. + \|Ax-b\|^{p-2}A^TA\right]dx_2$$

$$\Rightarrow \nabla^2 f(x) = p\left[(p-2)\|Ax-b\|^{p-4}A^T(Ax-b)(x^TA^T-b^T)\cdot\right.$$

$$\left.\cdot A + \|Ax-b\|^{p-2}A^TA\right]$$

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

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

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

$= 2X^{-1}(-1)X^{-1}dX_2X^{-1} = -2X^{-2}dX_2X^{-1} \} =$

$= 2tr(X^{-2}dX_2X^{-1}dX_1)$

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

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

$= 2tr(X^{-1}dXX^{-\frac{1}{2}}X^{-\frac{1}{2}}dXX^{-1}) = \{X \in S^n_{++}, dX \in S^n\} = 2tr((X^{-\frac{1}{2}}dXX^{-1})^T X^{-\frac{1}{2}}dXX^{-1}) =$

$= 2\|X^{-\frac{1}{2}}dXX^{-1}\|^2_F > 0$

$\Rightarrow$ $d^2f(X)[dX, dX]$ имеет полож. знак

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

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

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

$d^2f(X) = \frac{1}{n^2}(\det X)^{\frac{1}{n}-1} \det X \langle X^{-T}, dX_2\rangle\langle X^{-1}, dX_1\rangle +$

$+ \frac{1}{n}(\det X)^{\frac{1}{n}}\langle -X^{-1}dX_2X^{-1}, dX_1\rangle =$

$= \frac{1}{n}(\det X)^{\frac{1}{n}}\left[\frac{1}{n}\langle X^{-1}, dX_2\rangle\langle X^{-1}, dX_1\rangle - \langle X^{-1}dX_2X^{-1}, dX_1\rangle\right]$

$$d^2 f(X)[dX, dX] = \frac{1}{n}(\det X)^{\frac{1}{n}}\left[\frac{1}{n}(X^{-1}, dX)^2 - \right.$$
$$\left. - (X^{-1}dX X^{-1}, dX)\right] = \frac{1}{n}(\det X)^{\frac{1}{n}}\left[\frac{1}{n}(tr(X^{-1}dX))^2 - \right.$$
$$\left. - tr(X^{-1}dX X^{-1}dX)\right] = \frac{1}{n}(\det X)^{\frac{1}{n}}\left[\frac{1}{n}\cdot\right.$$
$$\left.\cdot\left(tr(X^{-\frac{1}{2}}dX X^{-\frac{1}{2}})\right)^2 - tr(X^{-\frac{1}{2}}dX X^{-\frac{1}{2}}X^{-\frac{1}{2}}dX X^{-\frac{1}{2}})\right] =$$
$$= \frac{1}{n}(\det X)^{\frac{1}{n}}\left[\frac{1}{n}\left(tr(X^{-\frac{1}{2}}dX X^{-\frac{1}{2}})\right)^2 - \|X^{-\frac{1}{2}}dX X^{-\frac{1}{2}}\|_F^2\right] =$$
$$= \frac{1}{n}(\det X)^{\frac{1}{n}}\cdot A, \quad A \in \mathbb{R}$$
$$A = \frac{1}{n}\left(tr(X^{-\frac{1}{2}}dX X^{-\frac{1}{2}})\right)^2 - \|X^{-\frac{1}{2}}dX X^{-\frac{1}{2}}\|_F^2 =$$
$= \{$ Пусть b_{ij} - эл-ты матрицы $B = X^{-\frac{1}{2}}dX X^{-\frac{1}{2}}\} =$
$$= \frac{1}{n}(b_{11} + \dots + b_{nn})^2 - \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij}^2 \le$$
$$\le \frac{1}{n}(b_{11} + \dots + b_{nn})^2 - (b_{11}^2 + \dots + b_{nn}^2)$$
$$\frac{1}{n}(b_{11} + \dots + b_{nn})^2 \vee (b_{11}^2 + \dots + b_{nn}^2) \quad \Big| \sqrt{}$$
$$\frac{b_{11} + \dots + b_{nn}}{\sqrt{n}} \vee \sqrt{b_{11}^2 + \dots + b_{nn}^2} \quad \Big| \cdot \frac{1}{\sqrt{n}}$$
$$\frac{b_{11} + \dots + b_{nn}}{n} \le \sqrt{\frac{b_{11}^2 + \dots + b_{nn}^2}{n}}$$

нер-во о среднем квадратичном и среднем арифметическом

$\Rightarrow A \le 0$ и т.к. $X \in S_{++}^n$, то $\det X > 0$

$\Rightarrow d^2 f(X)[dX, dX] \le 0$

6) a) $c \in \mathbb{R}^n, \ c \neq 0, \ \sigma > 0$

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

$$df(x) = \langle c, dx \rangle + \frac{\sigma}{3} d\langle x, x \rangle^{\frac{3}{2}} = \langle c, dx \rangle +$$

$$+ \frac{\sigma}{2} \langle x, x \rangle^{\frac{1}{2}} \cdot 2\langle x, dx \rangle = \langle c + \sigma\|x\|x, dx \rangle$$

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

$$c + \sigma\|x\|x = 0$$

$$x = -\frac{1}{\sigma\|x\|}c, \ \exists\, x = ac, \ a \in \mathbb{R} \Rightarrow$$

$$ac = -\frac{1}{\sigma|a|\|c\|}c \Rightarrow c\left(a + \frac{1}{\sigma|a|\|c\|}\right) = 0$$

Т.к $c \neq 0$, то $a + \frac{1}{\sigma|a|\|c\|} = 0 \iff$

$\iff a|a| = -\frac{1}{\sigma\|c\|}$. Т.к $\sigma > 0, \|c\| > 0$, то $a < 0 \Rightarrow$

$$\Rightarrow -a^2 = -\frac{1}{\sigma\|c\|} \Rightarrow a = -\sqrt{\frac{1}{\sigma\|c\|}}$$

Отговор: $x = -\sqrt{\frac{1}{\sigma\|c\|}}\, c$

b) $a, b \in \mathbb{R}^n; \ a, b \neq 0; \ E = \{x \in \mathbb{R}^n \mid \langle b, x \rangle < 1\}$

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

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

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

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

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

$$\Leftrightarrow a(1-b^T x) = -b \Leftrightarrow a - ab^T x = -b \Leftrightarrow$$

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

Мы получили СЛАУ $Ax = d$, где $A = ab^T$, $d = a+b$. Причём, если $a = (a_1, \ldots, a_n)^T$, то

$$A = \begin{bmatrix} a_1 \cdot b^T \\ \vdots \\ a_n \cdot b^T \end{bmatrix} \Rightarrow \det(A) = 0 \text{ и } \operatorname{rang} A = 1.$$

По Th. Кронекера-Капелли СЛАУ совместна $\Leftrightarrow \operatorname{rang}(A|d) = \operatorname{rang} A = 1$, где $A|d$ – расширенная матрица, полученная из матрицы A приписыванием столбца d.

Ответ: x - стац. точка $\Leftrightarrow$ x - решение СЛАУ $ab^T x = a+b$, причём $\operatorname{rang}(ab^T | a+b)$ должен равняться 1.

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

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

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

$$\cdot \exp(-\langle Ax, x\rangle) \cdot (-1) \cdot 2\langle Ax, dx\rangle =$$

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

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

$$\underbrace{\exp(-\langle Ax, x\rangle)}_{>0}(c - 2\langle c, x\rangle Ax) = 0 \Rightarrow$$

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

$$Ax = \frac{1}{2\langle c, x\rangle} c$$

$$x = \frac{1}{2\langle c, x\rangle} A^{-1} c \quad , \quad] \quad x = a A^{-1} c, \; a \in \mathbb{R}$$

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

$$a^2 = \frac{1}{2\langle c, A^{-1}c\rangle} \Rightarrow a = \pm\sqrt{\frac{1}{2\langle c, A^{-1}c\rangle}}$$

Отметим, что если $A \in S^n_{++}$, то $\exists A^{-1}$ и $A^{-1} \in S^n_{++} \Rightarrow c^T A^{-1} c > 0 \quad \forall c \in \mathbb{R}^n, c \neq 0$,

т.е. выражение под корнем всегда полож.

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

Ответ: $x = \pm\sqrt{\frac{1}{2\langle c, A^{-1}c\rangle}} A^{-1}c$.

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

$$\lim_{k\to+\infty} tr\left(X^{-k} - (X^k + X^{2k})^{-1}\right) = \lim_{k\to+\infty} tr A_k$$

$$A_k = X^{-k} - (X^k + X^{2k})^{-1} = \cancel{\{\text{тождество Вудбери:}}$$

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

$= \{$тождество Вудбери: $A = I, U = I, V = I,$
$C = X^k$, $\det A \neq 0$, $\det C \neq 0\} =$

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

$$= X^{-k} - (I - (X^{-k}+I)^{-1})X^{-k} = X^{-k} - X^{-k} + (X^{-k}+I)^{-1} \cdot$$

$$\cdot X^{-k} = (X^k(X^{-k}+I))^{-1} = (I+X^k)^{-1} =$$

$= \{$т.к. матрица X симметрическая, то справедливо след. разложение: $X = QDQ^T$, где Q – орт. матрица, столбцы которой содержат ОНБ из с.в., а D – диаг. матрица с с.зн. матр A на диагонали, т.е $D = diag(\lambda_1, \dots, \lambda_n)$, где λ_i – с.з. $A\} =$

$$= (QQ^T + QD^kQ^T)^{-1} = \{Q^{-1} = Q^T, (Q^T)^{-1} = Q\} =$$

$= Q(I+D^k)^{-1}Q^T$. Тогда получаем:

$$\lim_{k\to+\infty} tr\, Q(I+D^k)^{-1}Q^T = \lim_{k\to+\infty} tr(I+D^k)^{-1} =$$

$$= \lim_{k\to+\infty} tr\, diag\left(\frac{1}{1+\lambda_1^k}, \dots, \frac{1}{1+\lambda_n^k}\right) = \lim_{k\to+\infty} \sum_{m=1}^n \frac{1}{1+\lambda_m^k} =$$

$$= \sum_{m=1}^n \lim_{k\to+\infty} \frac{1}{1+\lambda_m^k} \overset{\text{т.к. все } \lambda_m > 0 \ (X \in S^n_{++})}{=} r + \frac{p}{2},$$ где r – кол-во $\lambda_j < 1$, p – кол-во $\lambda_j = 1$

8) $\{x_i\}_{i=1}^N$, $x_i \in \mathbb{R}^D$, $P \in \mathbb{R}^{D\times d}$

$$F(P) = \sum_{i=1}^N \|x_i - P(P^TP)^{-1}P^T x_i\|^2 = N\operatorname{tr}[(I-P(P^TP)^{-1}\cdot$$

$$\cdot P^T)^2 S], \quad S = \frac{1}{N}\sum_{i=1}^N x_i x_i^T$$

а) $\nabla_P F(P)$ – ? при условии, что $P^TP = I$.

Отметим, что $S^T = S$

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

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

$$d[P(P^TP)^{-1}P^T] = dP\cdot(P^TP)^{-1}P^T + Pd[(P^TP)^{-1}P^T] \boxed{=}$$

$$d[(P^TP)^{-1}P^T] = -(P^TP)^{-1}d(P^TP)(P^TP)^{-1}\cdot P^T +$$

$$+ (P^TP)^{-1}(dP)^T = -(P^TP)^{-1}[(dP)^TP + P^TdP](P^TP)^{-1}P^T +$$

$$+ (P^TP)^{-1}(dP)^T = -[(dP)^TP + P^TdP]P^T + (dP)^T =$$

$$= -(dP)^TPP^T - P^TdPP^T + (dP)^T$$

$$\boxed{=} dP\cdot P^T + P[-(dP)^TPP^T - P^TdPP^T + (dP)^T] =$$

$$= dP\cdot P^T - P(dP)^TPP^T - PP^TdPP^T + P(dP)^T$$

$\boxed{=} -2N\,tr[(I-PP^T)(dP\,P^T - P(dP)^T PP^T -$

$- PP^T dP\,P^T + P(dP)^T)S] = -2N\,tr[(dP\,P^T -$

$- P(dP)^T PP^T - PP^T dP\,P^T + P(dP)^T - PP^T dP\,P^T +$

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

$= -2N\,tr[P^T S\,dP - (dP)^T PP^T SP - P^T SPP^T dP +$

$+ (dP)^T SP - P^T SPP^T dP + (dP)^T PP^T SPP^T P +$

$+ P^T SPP^T PP^T dP - (dP)^T SPP^T P] =$

$= -2N\,tr[P^T S\,dP - P^T SPP^T dP - P^T SPP^T dP +$

$+ P^T S\,dP - P^T SPP^T dP + P^T PP^T SPP^T dP +$

$+ P^T SPP^T PP^T dP - P^T PP^T S\,dP] =$

$= -2N\,tr[P^T S\,dP - P^T SPP^T dP]$

$\Rightarrow \nabla_P F(P) = 2N(PP^T SP - SP)$

б) $S = Q\Lambda Q^T$, $Q = [q_1|q_2|\ldots|q_D] \in \mathbb{R}^{D\times D}$,

где q_i – с.в. матрицы S

I. Д-ем, что $\nabla_P F(P) = 0$ при

$P = [q_{u_1}|q_{u_2}|\ldots|q_{u_d}]$

$\nabla_P F(P) = 2N(PP^T SP - SP) = 0 \Rightarrow$

дост. показать, что $PP^T SP = SP$

$PP^T SP = PP^T Q\Lambda Q^T P$, $SP = Q\Lambda Q^T P$

Матрицы $Q^T P$ и $P^T Q$ имеют простой вид элементов. Найдём их:

$$(Q^T P)_{ij} = \sum_m Q^T_{im} P_{mj} = \sum_m Q_{mi} (q_{u_j})_m =$$

$$= \sum_m (q_i)_m (q_{u_j})_m = \{Q - \text{орт. мат.}\} =$$

$$= \delta_{i u_j} = \begin{cases} 1, & i = u_j \\ 0, & i \neq u_j \end{cases}$$

Поскольку $P^T Q = (Q^T P)^T$, то $(P^T Q)_{ij} = \delta_{j u_i}$

$$\xcancel{(P\underbrace{P^T Q}_{} \Lambda \underbrace{Q^T P}_{})_{ij} = \sum_m \sum_n \sum_l P_{im} \delta_{m u_n} \Lambda_{nl} \cdot}$$

$$\xcancel{\cdot \delta_{l u_j} = \{\text{т.к. } \Lambda \text{ диаг. матрица}\} = \sum_m \sum_n P_{im} \cdot}$$

$$\xcancel{\cdot \delta_{m u_n} \Lambda_{nn} \cdot \delta_{n u_j} =}$$

$$(P\underbrace{P^T Q}_{} \Lambda \underbrace{Q^T P}_{})_{ij} = \sum_m \sum_n \sum_l P_{im} \delta_{n u_m} \Lambda_{nl} \cdot$$

$$\cdot \delta_{l u_j} = \{\text{т.к. } \Lambda \text{ - диаг. матрица}\} = \sum_m \sum_n P_{im} \cdot$$

$$\cdot \delta_{n u_m} \Lambda_{nn} \cdot \delta_{n u_j} = \{\delta_{\alpha\beta} = 1 \Leftrightarrow \alpha = \beta\} =$$

$$= \{n = u_m, n = u_j \Rightarrow u_m = u_j \Rightarrow m = j\} =$$

$$= P_{ij} \Lambda_{u_j u_j} = (q_{u_j})_i \cdot \Lambda_{u_j u_j}$$

$$(Q\Lambda \underbrace{Q^T P})_{ij} = \sum_m \sum_n Q_{im}\Lambda_{mn}\delta_{nu_j} =$$

$$= \sum_m Q_{im}\Lambda_{mm}\delta_{mu_j} = Q_{iu_j}\Lambda_{u_j u_j} =$$

$$= (q_{u_j})_i \Lambda_{u_j u_j}$$

$\Rightarrow PP^T SP = SP \Rightarrow \nabla_P F(P) = 0$ – ч.т.д.

II. Д-ем, что $\min\limits_P F(P)$ достигается для матрицы $P = [q_{u_1} | \dots | q_{u_d}]$: с.в. q_{u_i} отвечают наибольшим с.зн.

Т.к. min – т. экстремума, то ~~он достигается~~ необх. усл., что $\nabla_P F(P) = 0$ – это показано в I.

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

$$= N\,tr((I - PP^T)S) = N\,tr(S - PP^TS) =$$

$$= \underbrace{N\,tr\,S}_{const} - \underbrace{N}_{const} tr(PP^TS) \to \min_P \Leftrightarrow tr(PP^TS) \to \max_P$$

$$tr(PP^TS) = tr(P\underbrace{P^TQ}\Lambda Q^T) = \cancel{\quad}$$

$$= \sum_{i=1}^{D} (P\underbrace{P^TQ}\Lambda Q^T)_{ii} = \sum_i \sum_m \sum_n \sum_l P_{im}\delta_{nu_m} \cdot \Lambda_{nl} \cdot Q^T_{li} = \sum_i \sum_m \sum_n P_{im}\delta_{nu_m}\Lambda_{nn} Q_{in} =$$

$$= \sum_i \sum_m P_{im} \Lambda_{u_m u_m} Q_{i u_m} = \sum_i \sum_m (q_{u_m})_i \cdot$$

$$\cdot \Lambda_{u_m u_m} \cdot (q_{u_m})_i = \sum_m \Lambda_{u_m u_m} \sum_i \left[(q_{u_m})_i \right]^2 =$$

$$= \sum_m \Lambda_{u_m u_m} \cdot \langle q_{u_m}, q_{u_m} \rangle = \sum_m \Lambda_{u_m u_m}$$

$\Rightarrow tr(PP^T S)$ достигает максимума, если в матрицу P взяли те с.в., которые соответствуют макс. с. зн. $\Rightarrow$

$\Rightarrow F(P)$ достигает минимума — ч.т.д.