Михеев Борис, 317.

№1 Док-ать: $(A+UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}$ (тождество Вудбери)
$A \in \mathbb{R}^{n\times n}, C \in \mathbb{R}^{m\times m}, U \in \mathbb{R}^{n\times m}, V \in \mathbb{R}^{m\times n}, \det A \neq 0, \det C \neq 0$

Докажем, что произведение $(A+UCV)$ и правой части тождества даёт I:

$(A+UCV)(A^{-1}-A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}) = I - U(C^{-1}+VA^{-1}U)^{-1}VA^{-1} +$
$+ UCVA^{-1} - UCVA^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1} = I + UCVA^{-1} -$
$-(U+UCVA^{-1}U)(C^{-1}+VA^{-1}U)^{-1}VA^{-1} = I + UCVA^{-1} -$
$-UC(C^{-1}+VA^{-1}U)(C^{-1}+VA^{-1}U)^{-1}VA^{-1} = I + UCVA^{-1} - UCVA^{-1} =$
$= I. \Rightarrow$ тождество верно ∎

№2

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

$\|uv^T - A\|_F^2 - \|A\|_F^2 = \langle uv^T - A, uv^T - A\rangle - \langle A, A\rangle = \langle uv^T, uv^T\rangle -$
$- \langle uv^T, A\rangle - \langle A, uv^T\rangle + \cancel{\langle A, A\rangle} - \cancel{\langle A, A\rangle} = \|uv^T\|_F^2 - 2\langle uv^T, A\rangle =$
$= \operatorname{tr} vu^Tuv^T - 2\operatorname{tr} vu^TA \underset{\text{цикличн. св-во следа}}{=} \operatorname{tr} \underbrace{v^Tvu^Tu}_{\in \mathbb{R}} - 2\operatorname{tr} \underbrace{u^TAv}_{\in \mathbb{R}} =$
$= v^Tvu^Tu - 2u^TAv = \boxed{\|v\|^2\|u\|^2 - 2\langle u, Av\rangle}$

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

Используем тождество Вудбери: $(A+UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}$,
где $A \in \mathbb{R}^{n\times n}, C \in \mathbb{R}^{m\times m}, U \in \mathbb{R}^{n\times m}, V \in \mathbb{R}^{m\times n}, \det A \neq 0, \det C \neq 0$.

$\operatorname{tr}((2I_n + aa^T)^{-1}(uv^T+vu^T)) = \left| \begin{array}{l} A = 2I_n \\ U = a \\ C = I_1 = 1 \\ V = a^T \end{array} \right| =$

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

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

$$= tr\left(\left(\frac{I_n}{2} - \frac{aa^T}{2(2+a^Ta)}\right)(uv^T + vu^T)\right) = tr\left(\frac{uv^T}{2} + \frac{vu^T}{2} - \frac{aa^Tuv^T}{2(2+a^Ta)} -\right.$$

$$\left. - \frac{aa^Tvu^T}{2(2+a^Ta)}\right) = tr\left(\frac{2uv^T + uv^Ta^Ta + 2vu^T + vu^Ta^Ta - aa^Tuv^T - aa^Tvu^T}{2(2+a^Ta)}\right)$$

$$= tr\,\frac{uv^T}{2} + tr\,\frac{vu^T}{2} - tr\,\frac{aa^Tuv^T}{2(2+a^Ta)} - tr\,\frac{aa^Tvu^T}{2(2+a^Ta)} =$$

линейность и цикличн. св-во следа

$$= \langle u, v\rangle - \frac{1}{2(2+a^Ta)}\left(tr\,aa^Tuv^T + tr\,aa^Tvu^T\right) =$$

$$= \langle u, v\rangle - \frac{1}{2(2+a^Ta)}\left(tr\,\underbrace{v^Taa^Tu}_{\in \mathbb{R}} + tr\,\underbrace{u^Taa^Tv}_{\in \mathbb{R}}\right) =$$

$$= u^Tv - \frac{1}{2(2+a^Ta)}\left(\underbrace{v^Taa^Tu}_{= (u^Taa^Tv)^T \in \mathbb{R}} + u^Taa^Tv\right) =$$

$$= u^Tv - \frac{u^Taa^Tv}{2+a^Ta} = \boxed{\langle u, v\rangle - \frac{\langle u, a\rangle \langle a, v\rangle}{2+\|a\|^2}}$$

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

Заметим, что S - симметричная матрица (как сумма симметричных матриц) $\Rightarrow$ S^{-1} - симметричная матрица $\Rightarrow$ $(S^{-1})^T = S^{-1}$

$$\sum_{i=1}^{n} \langle S^{-1}a_i, a_i\rangle = \sum_{i=1}^{n} tr\, a_i^TS^{-1}a_i \underset{\text{линейность } tr}{=} tr \sum_{i=1}^{n} a_i^TS^{-1}a_i \underset{\text{цикл. св-во, линейность}}{=} \sum_{i=1}^{n} tr\, a_ia_i^TS^{-1} =$$

$$= tr\,\underbrace{\sum_{i=1}^{n} a_ia_i^T}_{= S} S^{-1} = tr\, I_d = \boxed{d}$$

$A \in \mathbb{R}^{n \times n}$

№3

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

$df(t) = \det(A - tI_n)\, tr(A - tI_n)^{-1} d(A - tI_n) = -\det(A - tI_n)\, tr(A - tI_n)^{-1} dt \Rightarrow$

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

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

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

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

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

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

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

т.о. $\boxed{f''(t) = \det(A - tI_n)\left((tr(A - tI_n)^{-1})^2 - tr(((A - tI_n)^{-1})^2)\right)}$

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

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

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

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

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

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

$= \frac{-b^T ((A + tI_n)(A + tI_n)^T)^{-1} \underbrace{d((A + tI_n)(A + tI_n)^T)}_{= d((A + tI_n)^T (A + tI_n)),} ((A + tI_n)(A + tI_n)^T)^{-1} b}{2\|(A + tI_n)^{-1} b\|} =$

т.к. A - симм $\Rightarrow A + tI_n$ - симм.

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

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

$A+tI_n$ - симм.

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

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

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

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

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

$$d((A+tI_n)^{-1})^3 = d(A+tI_n)^{-1}(A+tI_n)^{-2} + (A+tI_n)^{-1}d(A+tI_n)^{-1}(A+tI_n)^{-1} +$$
$$+ (A+tI_n)^{-2}d(A+tI_n)^{-1} = -(A+tI_n)^{-1}dt(A+tI_n)^{-1}(A+tI_n)^{-2} +$$
$$+ (A+tI_n)^{-1}(-(A+tI_n)^{-1}dt(A+tI_n)^{-1})(A+tI_n)^{-1} +$$
$$+ (A+tI_n)^{-2}(-(A+tI_n)^{-1}dt(A+tI_n)^{-1}) = -3(A+tI_n)^{-4}dt$$

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

A-симм. $\Downarrow$ $A+tI_n$-симм.

$$= -2b^T(A+tI_n)^{-3}b\,dt.$$

Т.о, $d^2f(t) = \left(3b^T(A+tI_n)^{-4}b\|(A+tI_n)^{-1}b\|\,dt_1dt_2 - \frac{b^T(A+tI_n)^{-3}b\,dt_1\,b^T(A+tI_n)^{-3}b\,dt_2}{\|(A+tI_n)^{-1}b\|}\right)\cdot\frac{1}{\|(A+tI_n)^{-1}b\|^2} =$

$$= \frac{3b^T(A+tI_n)^{-4}b\,dt_1dt_2}{\|(A+tI_n)^{-1}b\|} - \frac{b^T(A+tI_n)^{-3}bb^T(A+tI_n)^{-3}b\,dt_1dt_2}{\|(A+tI_n)^{-1}b\|^3}$$

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

№4 a) $f: \mathbb{R}^n \to \mathbb{R}$; $f(x) = \frac{1}{2}\|xx^T - A\|_F^2$; $A \in \mathbb{S}^n$

$f(x) = \frac{1}{2}\|xx^T - A\|_F^2 = \frac{1}{2}tr((xx^T-A)^T(xx^T-A)) = \frac{1}{2}tr((xx^T-A^T)(xx^T-A)) \underset{A\in\mathbb{S}^n}{=} \frac{1}{2}tr(xx^T-A)^2$

$f(x) = \frac{1}{2}\langle xx^T-A, xx^T-A\rangle = \frac{1}{2}\langle xx^T, xx^T\rangle - \frac{2}{2}\langle xx^T, A\rangle + \frac{1}{2}\langle A, A\rangle =$

$= \frac{1}{2}tr\, xx^Txx^T - \langle Ax, x\rangle + \frac{1}{2}\langle A, A\rangle \underset{\text{цикл св-во}}{=} \frac{1}{2}tr\,\underbrace{x^Tx}_{\in\mathbb{R}}x^Tx - \langle Ax, x\rangle + \frac{1}{2}\langle A,A\rangle =$

$= \frac{1}{2}\|x\|^4 - \langle Ax, x\rangle + \frac{1}{2}\langle A,A\rangle = \frac{1}{2}\langle x,x\rangle^2 - \langle Ax, x\rangle + \frac{1}{2}\langle A, A\rangle$

$df(x) = \frac{2\langle x,x\rangle}{2}d(x^Tx) - 2\langle Ax, dx\rangle = \frac{2x^Tx}{2}\cdot 2x^Tdx - 2\langle Ax, dx\rangle =$

$= 2x^Txx^Tdx - 2x^TAdx = 2x^T(xx^T-A)dx \Rightarrow$

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

$d^2f(x) = 2\langle d(xx^T\cancel{x} - Ax), dx_1\rangle = 2\langle dx_2\cdot x^Tx + x\cdot d(x^Tx) - Adx_2, dx_1\rangle =$

$= 2\langle dx_2\cdot \underbrace{x^Tx}_{\in\mathbb{R}} + 2x\cdot x^Tdx_2 - Adx_2, dx_1\rangle = 2\langle x^Tx\,dx_2 + 2xx^Tdx_2 - Adx_2, dx_1\rangle = 2\langle \underbrace{(x^TxI + 2xx^T - A)}_{\text{симм. матр}}dx_2, dx_1\rangle \underset{\text{св-во скал. произв}}{=}$

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

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

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

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

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

$=\left(2x^Tdx\underbrace{\ln x^Tx}_{\in\mathbb{R}}+\frac{x^Tx\cdot 2x^Tdx}{x^Tx}\right)e^{x^Tx\ln x^Tx}=(2\ln x^Tx\cdot x^Tdx+2x^Tdx)\,e^{x^Tx\ln x^Tx}=$

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

$\boxed{=2\langle x,x\rangle^{\langle x,x\rangle}(\ln\langle x,x\rangle+1)\,x.}$

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

$=2\langle(e^{x^Tx\ln x^Tx}(2x^Tdx_2\cdot\ln x^Tx+\frac{x^Tx}{x^Tx}2x^Tdx_2)(\ln x^Tx+1)x+$

$+e^{x^Tx\ln x^Tx}\frac{2x^Tdx_2}{x^Tx}\cdot x+e^{x^Tx\ln x^Tx}(\ln x^Tx+1)dx_2),dx_1\rangle=$

$=2e^{x^Tx\ln x^Tx}\langle(2\ln x^Tx\cdot x^Tdx_2+2x^Tdx_2)(\ln x^Tx+1)x+\frac{2x^Tdx_2x}{x^Tx}+(\ln x^Tx+1)dx_2,dx_1\rangle=$

$=2e^{x^Tx\ln x^Tx}\langle(2\ln^2x^Tx\cdot x^Tdx_2+2\ln x^Tx\cdot x^Tdx_2+2\ln x^Tx\cdot x^Tdx_2+2x^Tdx_2)x+$

$+\frac{2x^Tdx_2x}{x^Tx}+(\ln x^Tx+1)dx_2,dx_1\rangle=2e^{x^Tx\ln x^Tx}\langle 2\ln^2x^Tx\cdot x^Tdx_2x+$

$+4\ln x^Tx\cdot x^Tdx_2x+2x^Tdx_2x+\frac{2x^Tdx_2x}{x^Tx}+(\ln x^Tx+1)dx_2,dx_1\rangle=$

$=2e^{x^Tx\ln x^Tx}\langle 2(\ln x^Tx+1)^2x^Tdx_2x+\frac{2x^Tdx_2x}{x^Tx}+(\ln x^Tx+1)dx_2,dx_1\rangle=$

$=2e^{x^Tx\ln x^Tx}\langle 2\langle(\ln x^Tx+1)^2x,dx_2\rangle x+\frac{2x^Tdx_2x}{x^Tx}+(\ln x^Tx+1)dx_2,dx_1\rangle=$

$=2e^{x^Tx\ln x^Tx}\langle 2\langle(\ln x^Tx+1)^2x,dx_2\rangle x+\langle\frac{2x}{x^Tx},dx_2\rangle x+(\ln x^Tx+1)dx_2,dx_1\rangle=$

$=2e^{x^Tx\ln x^Tx}\langle\langle 2(\ln x^Tx+1)^2x+\frac{2x}{x^Tx},dx_2\rangle x+(\ln x^Tx+1)dx_2,dx_1\rangle=$

$=2e^{x^Tx\ln x^Tx}\langle\langle 2(\ln x^Tx+1)^2x+\frac{2x}{x^Tx},dx_2\rangle x,dx_1\rangle+$

$+2e^{x^Tx\ln x^Tx}\langle\underbrace{(\ln x^Tx+1)}_{\in\mathbb{R}}dx_2,dx_1\rangle=2e^{x^Tx\ln x^Tx}\langle 2(\ln x^Tx+1)^2x+\frac{2x}{x^Tx},dx_2\rangle x^Tdx_1+$

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

$\cdot\langle x,dx_1\rangle+2e^{x^Tx\ln x^Tx}\langle(\ln x^Tx+1)dx_1,dx_2\rangle=2e^{x^Tx\ln x^Tx}(\langle\langle x,dx_1\rangle\cdot$

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

$=2e^{x^Tx\ln x^Tx}(\langle\langle x,dx_1\rangle(2(\ln x^Tx+1)^2x+\frac{2x}{x^Tx})+(\ln x^Tx+1)dx_1,dx_2\rangle)=$

$$= 2e^{x^Tx\ln x^Tx} < \left(2(\ln x^Tx+1)^2 + \frac{2}{x^Tx}\right) xx^T dx_1 + (\ln x^Tx+1)dx_1, dx_2 > =$$

$$= 2e^{x^Tx\ln x^Tx} < \left(2(\ln x^Tx+1)^2 + \frac{2}{x^Tx}\right) xx^T \cancel{dx_1} + (\ln x^Tx+1) I) dx_1, dx_2 > \Rightarrow$$

$$\Rightarrow \boxed{\begin{aligned}\nabla^2 f(x) &= e^{x^Tx\ln x^Tx}\left(4(\ln x^Tx+1)^2 xx^T + \frac{4xx^T}{x^Tx} + 2(\ln x^Tx+1)I\right) = \\ &= \langle x,x\rangle^{\langle x,x\rangle}\left(4(\ln\langle x,x\rangle+1)^2 xx^T + \frac{4xx^T}{\langle x,x\rangle} + 2(\ln\langle x,x\rangle+1)I\right)\end{aligned}}$$

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

$$f(x) = \|Ax-b\|^p = \langle Ax-b, Ax-b\rangle^{\frac{p}{2}} = ((Ax-b)^T(Ax-b))^{\frac{p}{2}}$$

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

$$\cdot\, d((x^TA^T-b^T)(Ax-b)) = \frac{p}{2}((Ax-b)^T(Ax-b))^{\frac{p}{2}-1} d(x^TA^TAx - \underbrace{x^TA^Tb}_{\in\mathbb{R}} - b^TAx +$$

$$+ b^Tb) = \frac{p}{2}((Ax-b)^T(Ax-b))^{\frac{p}{2}-1}(2x^TA^TA\,dx - 2b^TA\,dx) =$$

$$= p((Ax-b)^T(Ax-b))^{\frac{p}{2}-1}(x^TA^TA - b^TA)dx \Rightarrow$$

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

$$d^2 f(x) = d < p((Ax-b)^T(Ax-b))^{\frac{p-2}{2}} A^T(Ax-b), dx_1 > =$$

$$= < d(p((Ax-b)^T(Ax-b))^{\frac{p-2}{2}} A^T(Ax-b)), dx_1 > = < \frac{p(p-2)}{2}((Ax-b)^T(Ax-b))^{\frac{p-4}{2}} \cdot$$

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

$$= < \frac{p(p-2)}{2}\|Ax-b\|^{p-4} \cdot 2(Ax-b)^TA\,dx_2\, A^T(Ax-b) + p\|Ax-b\|^{p-2}A^TA\,dx_2, dx_1 > =$$

$$= < p(p-2)\|Ax-b\|^{p-4} \underbrace{(Ax-b)^TA\,dx_2}_{\in\mathbb{R}} A^T(Ax-b) + p\|Ax-b\|^{p-2}A^TA\,dx_2, dx_1 > =$$

[illegible]

$$= < p(p-2)\|Ax-b\|^{p-4} A^T(Ax-b)(Ax-b)^TA\,dx_2 + p\|Ax-b\|^{p-2}A^TA\,dx_2, dx_1 > =$$

$$= < (p(p-2)\|Ax-b\|^{p-4}A^T(Ax-b)(Ax-b)^TA + p\|Ax-b\|^{p-2}A^TA)dx_2, dx_1 > \Rightarrow$$

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

№5 (a) $f:\mathbb{S}^n_{++}\to\mathbb{R}$, $f(X)=\operatorname{tr}(X^{-1})$.

$df(X)=d\operatorname{tr}(X^{-1})=\operatorname{tr}dX^{-1}=\operatorname{tr}(-X^{-1}dX\,X^{-1})=-\operatorname{tr}(X^{-1}dX\,X^{-1})\underset{\text{цикл св-во след}}{=}$

$-\operatorname{tr}(X^{-2}dX)$

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

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

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

$=2\operatorname{tr}X^{-2}dX_2X^{-1}dX_1.$

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

$X\in\mathbb{S}^n_{++}\Rightarrow X^{-1}\in\mathbb{S}^n_{++}\Rightarrow\exists X^{-\frac12}\in\mathbb{S}^n_{++}:(X^{-\frac12})^2=X^{-1}\Rightarrow$

$\Rightarrow d^2f(X)[dX,dX]=2\operatorname{tr}X^{-1}dX\,X^{-\frac12}X^{-\frac12}dX\,X^{-1}=2\operatorname{tr}(X^{-1}dX\,X^{-\frac12})(X^{-1}dX\,X^{-\frac12})^T=$

$=2\operatorname{tr}(X^{-1}dX\,X^{-\frac12})^T(X^{-1}dX\,X^{-\frac12})=2\operatorname{tr}\|X^{-1}dX\,X^{-\frac12}\|^2>0\Rightarrow$

(т.к. $X\neq0$, $dX\neq0$)

$\Rightarrow\boxed{d^2f(X)[dX,dX]>0.}$

(b) $f:\mathbb{S}^n_{++}\to\mathbb{R}:f(X)=(\det X)^{\frac1n}$

$df(X)=\frac1n(\det X)^{\frac1n-1}\cdot d\det X=\frac1n(\det X)^{\frac1n-1}\det X\cdot\operatorname{tr}X^{-1}dX=$

$=\frac1n(\det X)^{\frac1n}\operatorname{tr}X^{-1}dX$

$d^2f(X)=\frac{1}{n^2}(\det X)^{\frac1n-1}(\det X)\operatorname{tr}X^{-1}dX_2\operatorname{tr}X^{-1}dX_1+\frac1n(\det X)^{\frac1n}\operatorname{tr}dX^{-1}dX_1=$

$=\frac{1}{n^2}(\det X)^{\frac1n}\operatorname{tr}X^{-1}dX_2\operatorname{tr}X^{-1}dX_1+\frac1n(\det X)^{\frac1n}\operatorname{tr}(-X^{-1}dX_2X^{-1}dX_1)=$

$=\frac{1}{n^2}(\det X)^{\frac1n}\operatorname{tr}X^{-1}dX_2\operatorname{tr}X^{-1}dX_1-\frac1n(\det X)^{\frac1n}\operatorname{tr}X^{-1}dX_2X^{-1}dX_1=$

$=\frac1n(\det X)^{\frac1n}\left(\frac1n\operatorname{tr}X^{-1}dX_2\operatorname{tr}X^{-1}dX_1-\operatorname{tr}X^{-1}dX_2X^{-1}dX_1\right)$

$X\in\mathbb{S}^n_{++}\Rightarrow\det X>0.$

$d^2f(X)[dX,dX]=\frac1n(\det X)^{\frac1n}\left(\frac1n\operatorname{tr}X^{-1}dX\operatorname{tr}X^{-1}dX-\operatorname{tr}X^{-1}dX\,X^{-1}dX\right)=$

$=\frac1n(\det X)^{\frac1n}\left(\frac1n\operatorname{tr}^2X^{-1}dX-\operatorname{tr}X^{-1}dX\,X^{-1}dX\right)=$

$=\frac1n(\det X)^{\frac1n}\left(\frac1n\operatorname{tr}^2X^{-1}dX-\operatorname{tr}(X^{-1}dX)^2\right)=\frac1n(\det X)^{\frac1n}\left(\frac1n\langle X^{-1},dX\rangle^2-\right.$

$\left.-\langle dX^TX^{-1},X^{-1}dX\rangle\right)=\frac1n(\det X)^{\frac1n}\left(\frac1n\langle dX^TX^{-1},I\rangle^2-\langle dX^TX^{-1},X^{-1}dX\rangle\right)\le$

$\underset{\text{КБШ}}{\le}\frac1n(\det X)^{\frac1n}\left(\frac1n\|dX^TX^{-1}\|^2\underbrace{\|I\|^2}_{n}-\|dX^TX^{-1}\|^2\right)=0\Rightarrow$

$\Rightarrow\boxed{d^2f(X)[dX,dX]\le0}$

№6 a) $f: \mathbb{R}^n \to \mathbb{R}$; $c \in \mathbb{R}^n$; $c \neq 0$

$f(x) = \langle c, x \rangle + \frac{\sigma}{3}\|x\|^3$; $\sigma > 0$

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

$$df(x) = c^T dx + \frac{\sigma}{3}\cdot\frac{3}{2}(x^T x)^{\frac{1}{2}}\cdot 2x^T dx = c^T dx + \sigma\|x\|\, x^T dx =$$

$$= (c^T + \sigma\|x\|\, x^T)\, dx \Rightarrow \nabla f(x) = c + \sigma\|x\|\, x = 0.$$

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

$\Rightarrow \|x\| = \sqrt{\frac{\|c\|}{\sigma}} \Rightarrow$ подставим в исходное ур-е: $\Rightarrow x\sqrt{\frac{\|c\|}{\sigma}} = -\frac{c}{\sigma}$

$x = -\frac{c\sqrt{\sigma}}{\sigma\sqrt{\|c\|}} = -\frac{c}{\sqrt{\|c\|\sigma}}$. Т.о, $\exists$-ет ! точка стационарности:

$\boxed{x = -\frac{c}{\sqrt{\|c\|\sigma}}}$, $\exists$ при $\forall c, \sigma$, соответствующих условию задачи.

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

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

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

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

$\Rightarrow \nabla f(x) = a + \frac{b}{1 - b^T x} = 0 \Rightarrow a$ и b – л.з. $\Rightarrow \exists c \in \mathbb{R}$: $a = cb$ $\Rightarrow$, $c \neq 0$

$\Rightarrow cb + \frac{b}{1 - b^T x} = 0 \underset{b \neq 0}{\Rightarrow} c + \frac{1}{1 - b^T x} = 0$; $\frac{1}{1 - b^T x} = -c$; $1 = c(b^T x - 1)$

$b^T x - 1 = \frac{1}{c}$; $b^T x < 1$ по условию $\Rightarrow \frac{1}{c} < 0 \Rightarrow c < 0$

Т.о, $b^T x = \frac{1}{c} + 1$, $c < 0$ – ~~множество искомых точек~~ уравнение задающее множество искомых точек стац-ти. (c задается из усл $\Rightarrow$ л3-ти векторов a и b) При этом вектора a, b – л.з. : $\exists c \in \mathbb{R}$: $a = cb$.
$c \neq 0$
$c < 0$

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

$$f(x) = c^T x \cdot e^{-x^T A x}$$

$$df(x) = c^T dx \cdot e^{-x^T A x} + c^T x \cdot e^{-x^T A x}(-2x^T A dx) = e^{-x^T A x}(c^T - 2c^T x \cdot x^T A)dx$$

$$\nabla f(x) = e^{-x^T A x}(c - 2Axx^T c) = 0$$

$$c - 2Axx^T c = 0; \quad c - 2\langle x, c\rangle Ax = 0; \quad Ax = \frac{c}{2\langle x, c\rangle} \Rightarrow x = \frac{A^{-1} c}{2\langle x, c\rangle}$$

$$x\langle x, c\rangle = \frac{A^{-1}c}{2}; \quad xx^T c = \frac{A^{-1}c}{2} \Rightarrow c^T x x^T c = \frac{c^T A^{-1} c}{2}$$

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

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

— 2 т. стац-ти, ∃-ют при ∀ знач.ях c, A, соответствующих условию.

№2 $X \in \mathbb{S}^n_{++}$.

$$\lim_{k\to+\infty} tr(X^{-k} + (X^k + X^{2k})^{-1}) \underset{\text{тожд. Вудбери (см. №1)}}{=} \left| \begin{matrix} A = X^k \\ U = X^k \\ V = X^k \\ C = I_n \end{matrix} \right| = \lim_{k\to+\infty} tr(X^{-k} - X^{-k} + X^{-k} X^k \circ$$

$$\circ (I + X^k X^{-k} X^k)^{-1} X^k X^{-k}) = \lim_{k\to+\infty} tr(\underbrace{I + X^k}_{\text{симм. матр}})^{-1} = \left| \begin{matrix} \text{След симметр матрицы} - \\ - \sum \text{ее с.з. } \lambda_j \\ \text{с учетом кратности} \end{matrix} \right| =$$

$$= \left| \begin{matrix} \text{с.з. } I + X^k \text{ имеют} \\ \text{вид } \lambda_i^k + 1, \text{ где} \\ \lambda_i - \text{с.з. матрицы } X \\ \text{(согласно свойствам} \\ \text{собств значений и} \\ \text{симм. матриц)} \Rightarrow \\ \Rightarrow \text{с.з. } (I + X^k)^{-1} \text{ имеют} \\ \text{вид } \frac{1}{\lambda_i^k + 1} \end{matrix} \right| = \lim_{k\to+\infty} \sum_{i=1}^{n} \frac{1}{\lambda_i^k + 1} = (\text{т.к. } k \to +\infty) =$$

$$= \lim_{k\to+\infty}\Big(\sum_{\substack{i=1 \\ 0<\lambda_i<1}}^{n} \frac{1}{\lambda_i^k + 1} + \sum_{\substack{i=1 \\ \lambda_i=1}}^{n} \frac{1}{\lambda_i^k + 1} + \sum_{\substack{i=1 \\ \lambda_i>1}}^{n} \frac{1}{\lambda_i^k + 1} \Big) =$$

$$\xrightarrow{k\to\infty} 1, \quad \xrightarrow{k\to\infty} \frac{1}{2}, \quad \xrightarrow{k\to\infty} 0$$

$$= \boxed{\sum_{i=1}^{n}[0 < \lambda_i < 1] + \frac{1}{2}\sum_{i=1}^{n}[\lambda_i > 1]}$$

индикаторы (скобки Айверсона)

№3) $\{x_i\}_{i=1}^N$, $x_i \in \mathbb{R}^D$, $P \in \mathbb{R}^{D\times d}$, $\text{proj}\, x = P(P^TP)^{-1}P^Tx$.

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

$S = \frac{1}{N}\sum_{i=1}^N x_ix_i^T$ - выборочная матрица коварнации

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

$= N\,tr(I - P(P^TP)^{-1}P^T - P(P^TP)^{-1}P^T + P(P^TP)^{-1}\cancel{P^TP(P^TP)^{-1}}P^T)S =$

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

$dF(P) = -N\,tr\, dP(P^TP)^{-1}P^TS = -N\,tr(dP(P^TP)^{-1}P^T + P(-(P^TP)^{-1} \cdot 2P^TdP \cdot$

$\cdot (P^TP)^{-1})P^T + P(P^TP)^{-1}dP^T)S = -N\,tr(dP(P^TP)^{-1}P^T - 2P(P^TP)^{-1}P^TdP \cdot$

$\cdot (P^TP)^{-1}P^T + P(P^TP)^{-1}dP^T)S = -N(tr\,dP(P^TP)^{-1}P^TS - 2\,tr\,P(P^TP)^{-1}P^T \cdot$

$\cdot dP(P^TP)^{-1}P^TS + tr\,P(P^TP)^{-1}dP^TS) \underset{trA = trA^T}{=} -N(tr(P^TP)^{-1}P^TS\,dP -$

$-2\,tr(P^TP)^{-1}P^TS \cdot P(P^TP)^{-1}P^TdP + tr(P^TP)^{-1}P^TSdP) =$

$= -2N(tr(P^TP)^{-1}P^TSdP - tr(P^TP)^{-1}P^TSP(P^TP)^{-1}P^TdP) =$

$= -2N\,tr((P^TP)^{-1}P^TS - (P^TP)^{-1}P^TSP(P^TP)^{-1}P^T)dP =$

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

$= -2N \langle (I - P(P^TP)^{-1}P^T)SP(P^TP)^{-1}, dP \rangle \Rightarrow$

$\Rightarrow \nabla_P F(P) = 2N(P(P^TP)^{-1}P^T - I)SP(P^TP)^{-1}$

Используем св-во ортогональности столбцов P: $P^TP = I \Rightarrow$

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

b) $\nabla_P F(P) = 2N(PP^T - I_D)SP$. По усл. $S = Q\Lambda Q^T$, Λ - диаг матр с с. з на диаг, $Q = (q_1|\dots|q_D) \in \mathbb{R}^{D\times D}$ - ортог матр с с.в. q_i по столб.

По опр-ю ортог-ти Q: $QQ^T = Q^TQ = I \Rightarrow \begin{pmatrix} q_1^T \\ \vdots \\ q_D^T \end{pmatrix}(q_1|\dots|q_D) =$ (ОНБ)

$= (q_1|\dots|q_D)\begin{pmatrix} q_1^T \\ \vdots \\ q_D^T \end{pmatrix} = I \Rightarrow q_i^Tq_j = \begin{cases} 1; & i=j \\ 0; & i\neq j \end{cases} \quad \forall i,j = \overline{1,D}$

$(q_1|\dots|q_D)\begin{pmatrix} q_1^T \\ \vdots \\ q_D^T \end{pmatrix} = \left\{\sum_{j,k=1}^D q_{ji}q_{kt}\right\}_{i,t=1}^D$

По опр-ю, если X - с.в. матрицы S, отвечающий с.з. λ $\Rightarrow$

$\Rightarrow SX = \lambda X$

$\exists\ P = (q_{k1} | \ldots | q_{kd}) \in \mathbb{R}^{D\times d}$, $1 \le k_1 < \ldots < k_d \le D$. $q_{ki} \sim \lambda_{ki}$
(d различных с.в.) с.в. с.з.

Тогда $\nabla_P F(P) = 2N(PP^T - I_D)SP = 2N(PP^TSP - SP) =$

$= 2N(PP^T(\lambda_{k1} q_{k1} | \ldots | \lambda_{kd} q_{kd}) - (\lambda_{k1} q_{k1} | \ldots | \lambda_{kd} q_{kd})) =$

$$= 2N\left(P\begin{pmatrix} \lambda_{k1} & & 0 \\ & \ddots & \\ 0 & & \lambda_{kd}\end{pmatrix} - (\lambda_{k1} q_{k1} | \ldots | \lambda_{kd} q_{kd})\right) =$$

$= 2N((\lambda_{k1} q_{k1} | \ldots | \lambda_{kd} q_{kd}) - (\lambda_{k1} q_{k1} | \ldots | \lambda_{kd} q_{kd})) = 0$

При этом т.к. $Q^TQ = I_D \Rightarrow P^TP = I_d$.

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

$\exists\ P = (q_{k1} | \ldots | q_{kd})$, $q_{ki} \sim \lambda_{ki}$, $1 \le k_1 < \ldots < k_d \le D$. Т.о. $P^TP = I_d$
(т.к. сост из с.в. q_{ki})

$$F(P) = N\,tr((I - P(\underbrace{P^TP}_{= I_d})^{-1}P^T)^2 S) = N\,tr((I - PP^T)^2 S) =$$

$$= N\,tr(I - 2PP^T + P\underbrace{P^TP}_{= I_d}P^T)S = N\,tr(I - PP^T)S = N\,tr(S - PP^TS) =$$

$= N(tr S - tr PP^TS) = N(tr S - tr P^TSP) = N(tr S -$

$$- tr\left[\begin{pmatrix} q_{k1}^T \\ \vdots \\ q_{kd}^T \end{pmatrix} (\lambda_{k1} q_{k1} | \ldots | \lambda_{kd} q_{kd})\right]) = N\left(\sum_{i=1}^{D} \lambda_i - \sum_{i=1}^{d} \lambda_{ki}\right).$$

Данное выр-е достигает своего min при max возможном значении $\sum\limits_{i=1}^{d} \lambda_{ki}$. Т.о, $F(P)$ достигает ~~своего~~ минимума, если матрица P состоит из с.в. q_i, отвечающих наибольшим с.з. матрицы S:

$P = (q_{k1} | \ldots | q_{kd})$, где ~~$\lambda_1 \ge \ldots \ge \lambda_d \ge \lambda_{d+1} \ge \ldots \ge \lambda_D$~~

$\lambda_{ki}, i = \overline{1, d}$ - d наибольших с.з. матрицы S