$\boxed{1}$

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

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

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

$$\Rightarrow 0 = A^{-1}UCV - A^{-1}U(C^{-1}+VA^{-1}U)^{-1}(V - VA^{-1}UCV) =$$

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

$$= A^{-1}UCV - A^{-1}UCV = 0$$

$\Rightarrow$ тождество верно

$\boxed{2}$

a) $\|uv^T - A\|_F^2 - \|A\|_F^2$

$\langle uv^T - A ; uv^T - A\rangle - \langle A ; A\rangle =$

$= \langle uv^T ; uv^T\rangle - \langle 2uv^T ; A\rangle + \langle A ; A\rangle - \langle A ; A\rangle$

$= \langle uv^T ; uv^T - 2A\rangle$

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

$= tr\left(\left[\frac{1}{2}I_n - \frac{1}{2}a\left(I_1 + a^T\frac{1}{2}a\right)^{-1}a^T\frac{1}{2}\right](uv^T + vu^T)\right)$

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

$= tr\left(\left[\frac{1}{2}I_n - \frac{aa^T}{4 + 2\|a\|^2}\right] \cdot 2\langle u, v\rangle\right) =$

$= tr\left(\left[I_n - \frac{aa^T}{2 + \|a\|^2}\right] \cdot \langle u, v\rangle\right) =$

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

$= \langle u, v\rangle\left(n - \frac{\|a\|^2}{2 + \|a\|^2}\right)$

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

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

т.к. S симметрич.

$= \sum_{i=1}^{n} tr \left(a_i^T \left(\sum_{j=1}^{n} a_j a_j^T\right)^{-1} a_i\right) =$ т.к. $tr AB = tr BA$

$= \sum_{i=1}^{n} tr \left(a_i a_i^T \left(\sum_{j=1}^{n} (a_j a_j)^T\right)^{-1}\right) = tr \left[\sum_{i=1}^{n} a_i a_i^T \cdot\right.$

$\left. \cdot \left(\sum_{j=1}^{n} a_j a_j^T\right)^{-1}\right] = tr\, I_d = d$

13

a) $f(t) = \det (A - t I_n)$

$d f(t) = d \det (A - t I_n) =$

$= \det (A - t I_n) \langle (A - t I_n)^{-T} ; d(-t I_n) \rangle =$

$= - \det (A - t I_n) \langle (A - t I_n)^{-T} ; I_n dt \rangle =$

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

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

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

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

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

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

b) $f(t) = \| (A + tI_n)^{-1} b \| =$ $\qquad A \in S_+^{n \times n}$

$$= \langle (A + tI_n)^{-1} b ; (A + tI_n)^{-1} b \rangle^{\frac{1}{2}}$$

$$df(t) = \frac{1}{2} \frac{1}{\| (A + tI_n)^{-1} b \|} \cdot 2 \langle (A + tI_n)^{-1} b ;$$

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

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

$$
= \underbrace{\frac{-1}{\|(A+tI_n)^{-1}b\|} \langle (A+tI_n)^{-1}b, (A+tI_n)^{-2}b\rangle}_{f'(t)} dt
$$

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

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

$$
+ \frac{1}{\|(A+tI_n)^{-1}b\|} \Big(\langle -(A+tI_n)^{-1} dt\, (A+tI_n)^{-1}b, (A+tI_n)^{-2}b\rangle
$$

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

$$
- (A+tI_n)^{-1}(A+tI_n)^{-1} dt\, (A+tI_n)^{-1}\rangle \Big) dt =
$$

$$
= \Bigg[\underbrace{\frac{-1}{\|(A+tI_n)^{-1}b\|^3} \langle (A+tI_n)^{-1}b, (A+tI_n)^{-2}b\rangle^2 + \frac{-1}{\|(A+tI_n)^{-1}b\|}\Big(\langle (A+tI_n)^{-2}b, (A+tI_n)^{-2}b\rangle + \langle (A+tI_n)^{-1}b, 2(A+tI_n)^{-3}b\rangle \Big)}_{f''(t)} \Bigg] dt^2
$$

$\boxed{4}$

a) $f(x) = \frac{1}{2}\|xx^T - A\|_F^2 =$

$= \frac{1}{2} \langle xx^T - A; xx^T - A \rangle =$

$df(x) = \frac{1}{2} \cdot 2 \langle xx^T - A; d(xx^T) \rangle =$

$\langle xx^T - A; dx \cdot x^T \rangle + \langle xx^T - A; x\, dx^T \rangle$

$= \operatorname{tr}(xx^T - A)^T dx \cdot x^T + \operatorname{tr}(xx^T - A)^T x\, dx^T =$

$= \operatorname{tr} x^T (xx^T - A)^T dx + \operatorname{tr} dx^T (xx^T - A)^T x =$

$= \operatorname{tr} x^T (xx^T - A)^T dx + \operatorname{tr} x^T (xx^T - A) dx =$

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

$\Rightarrow \nabla f = 2xx^T x - 2Ax$

$$d^2 f = d\langle 2(xx^T - A)x \,; dx_1\rangle =$$

$$= 2\langle d[\, x\|x\|^2 - A\ x \,] \,; dx_1\rangle =$$

$$= 2\langle \|x\|^2 dx_2 + x\cdot 2\|x\| \langle \|x\|^{-1} x ; dx_2\rangle - A\, dx_2 ;$$

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

$$dx_1\rangle$$

Приведем к каноннич. виду

$$\langle A dx_2 ; dx_1\rangle = \operatorname{tr}((dx_2)^T A^T dx_1) =$$

$$= \operatorname{tr}((dx_1)^T A\, dx_2) = \langle A^T dx_1 ; dx_2\rangle$$

$$\Rightarrow \langle [2 I_n \|x\|^2 + 4xx^T - 2A] dx_1 ; dx_2\rangle$$

$$\Rightarrow \nabla^2 f = 2[I_n \|x\| + 4xx^T - 2A]$$

b) $f(x) = \langle x, x\rangle^{\langle x, x\rangle} = \|x\|^{2\|x\|^2} =$

$$= e^{\|x\|^2 \cdot \ln \|x\|^2}$$

$$df = e^{\|x\|^2 \cdot \ln \|x\|^2} \cdot d[\|x\|^2 \cdot \ln \|x\|^2] =$$

$$= e^{\|x\|^2 \cdot \ln \|x\|^2} \cdot \left(2\|x\| \cdot d\|x\| \cdot \ln \|x\|^2 + \|x\|^2 \cdot 2 \cdot \frac{1}{\|x\|} \cdot d\|x\| \right) =$$

$$= \|x\|^{2\|x\|^2} \left(4\|x\| \cdot \ln\|x\| \cdot \|x\|^{-1}\langle x, dx\rangle + 2\|x\| \cdot \|x\|^{-1} \langle x, dx\rangle \right) =$$

$$= \|x\|^{2\|x\|^2}\left(4\ln\|x\|+2\right)\langle x, dx\rangle$$

$$\Rightarrow \nabla f = \|x\|^{2\|x\|^2}\left(4\ln\|x\|+2\right)x$$

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

$$= \|x\|^{2\|x\|^2}\left(4\ln\|x\|+2\right)^2\langle x, dx_2\rangle x$$

$$= 4\frac{1}{\|x\|}\cdot d\|x\| x + \left(4\ln\|x\|+2\right)dx_2 =$$

$$= \left[4\frac{1}{\|x\|^2}\cdot xx^T + \left(4\ln\|x\|+2\right)I_n\right]dx_2$$

$$\Rightarrow d^2 f = \|x\|^{2\|x\|^2} \langle (4\ln\|x\|+2)^2\, x \langle x, dx_2\rangle +$$

$$+ \left(4\|x\|^{-2} \cdot xx^T + (4\ln\|x\|+2) I_n\right) dx_2 ,\ dx_1 \rangle$$

$$\langle x\langle x, dx_2\rangle, dx_1\rangle = \langle xx^T dx_2 , dx_1\rangle$$

$$\Rightarrow d^2 f = \|x\|^{2\|x\|^2} \left\langle \left[\left((4\ln\|x\|+2)^2 + \frac{4}{\|x\|^2}\right) \cdot\right.\right.$$

$$\left.\cdot\, xx^T + (4\ln\|x\|+2) I_n\right] dx_2 ,\ dx_1 \rangle =$$

$$= \|x\|^{2\|x\|^2} \left\langle \left[\left((4\ln\|x\|+2)^2 + \frac{4}{\|x\|^2}\right) xx^T +\right.\right.$$

$$\left.+ (4\ln\|x\|+2) I_n\right] dx_1 ,\ dx_2 \rangle$$

$$\nabla^2 f = \|x\|^{2\|x\|^2} \left((4\ln\|x\|+2)^2 + \frac{4}{\|x\|^2}\right) xx^T +$$

$$+ (4\ln\|x\|+2) I_n$$

c) $f(x) = \|Ax-b\|^p$

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

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

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

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

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

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

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

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

$$\cdot \langle d A^T(Ax-b), dx_1\rangle$$

$$\underline{\langle A^T(Ax-b), dx_2\rangle \cdot \langle A^T(Ax-b), dx_1\rangle} =$$

$$= (Ax-b)^T A\, dx_2 \cdot (Ax-b)^T A\, dx_1 =$$

$$= dx_2^T A^T(Ax-b)(Ax-b)^T A\, dx_1 =$$

$$= \langle A^T(Ax-b)(Ax-b)^T A\, dx_2 ; dx_1\rangle =$$

$$= \langle A^T(Ax-b)(Ax-b)^T A\, dx_1 ; dx_2 \rangle =$$

$$\Rightarrow d^2 f = \langle \left[p(p-2) \|Ax-b\|^{p-4} \cdot A^T(Ax-b)(Ax-b)^T A + p\|Ax-b\|^{p-2} A^T A \right] dx_1 ; dx_2 \rangle$$

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

$\boxed{5}$

a) $f(x) = tr(X^{-1})$

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

$$d^2 f = tr(X^{-1} dx_2 X^{-1} dx_1 X^{-1}) +$$

$$+ tr(X^{-1} dx_1 X^{-1} dx_2 X^{-1}) =$$

$$= tr(X^{-T} dx_1^T X^{-T} d(x_2^T X^{-T}) + tr(X^{-1} dx_1 X^{-1} dx_2 X^{-1}) =$$

$$= 2\, tr(X^{-1} dx_1 X^{-1} dx_2 X^{-1}) \ (=)$$

Y мас $dx_1 = dx_2 = dx$

$\circledcirc\ 2\,tr\left(X^{-1}dX\,X^{-\frac{1}{2}}X^{-\frac{1}{2}}dX\,X^{-1}\right) =$

$= 2\,\langle X^{-\frac{1}{2}T}dX^{T}X^{-T},\ X^{-\frac{1}{2}}dX\,X^{-1}\rangle =$

$= 2\,\| X^{-\frac{1}{2}}dX\,X^{-1}\|^2$ – имеет положит. знак

б) $f(X) = (\det X)^{\frac{1}{n}}$

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

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

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

$dX_1 = dX_2 = dX$

$\circledcirc\ \frac{1-n}{n^2}(\det X)^{\frac{1-2n}{n}}\ \langle X^{-1}, dX\rangle^2 -$

$-\frac{1}{n}(\det X)^{\frac{1-n}{n}}\ \| X^{-\frac{1}{2}}dX\,X^{-\frac{1}{2}}\|^2$

т.к. X – положит. определенная $(\det X) > 0$

$\Rightarrow$ получаем, что знак отрицательный

$\boxed{6}$

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

$$df = \langle c, dx \rangle + \frac{\sigma}{3} \cdot 3 \|x\|^2 \cdot \|x\|^{-1} \langle x, dx \rangle$$

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

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

$$\sigma \|x\| x = -c$$

$$\Rightarrow x = \lambda c$$

$$\sigma |\lambda| \|c\| \lambda c = -c$$

$$\sigma |\lambda| \|c\| \lambda = -1$$

$$\Rightarrow -\lambda |\lambda| = \frac{1}{\|c\| \sigma}$$

$$-\lambda^2 = \frac{1}{\|c\| \sigma}$$

$$\lambda = -\sqrt{\frac{1}{\|c\| \sigma}}$$

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

При $\|c\|, \sigma \neq 0$

$\Rightarrow$ при $c, \sigma \neq 0$

и $\sigma > 0$

б)

$$f(x) = \langle a, x\rangle - \ln(1 - \langle b, x\rangle)$$

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

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

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

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

$a + b = a\langle b, x\rangle \Rightarrow a$ и b - линейно зависимы

$$\Rightarrow b = \frac{\|b\|}{\|a\|} a$$

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

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

Возьмем $x_0 = b\left(\frac{1}{\|b\|^2} + \frac{1}{\|b\|\|a\|}\right)$

Тогда $\left(\frac{1}{\|b\|^2} + \frac{1}{\|b\|\|a\|}\right) \cdot \|b\|^2 = 1 + \frac{\|b\|}{\|a\|}$

$\Rightarrow x_0$ - решение.

$$x = x_0 + x' \Rightarrow \langle b, x'\rangle = 0$$

(*) $\sum_{i=1}^{n} b_i x_i^1 = 0$, т.к. $\exists j : b_j \neq 0$ ($b \neq 0$ по условию) будем считать без нарушения общности, что $i=1$

Тогда решение (*):

$$x^1 = \lambda_2 \begin{pmatrix} y_2 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} y_3 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \ldots + \lambda_n \begin{pmatrix} y_n \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

$$b_1 y_i + b_i = 0 \Rightarrow y_i = -\frac{b_i}{b_1} \quad i = 2 \ldots n$$

$\lambda_i \in \mathbb{R}$ $i = 2 \ldots n$

$$\Rightarrow \quad x = b \left(\frac{1}{\|b\|^2} + \frac{1}{\|b\| \|a\|} \right) + x^1$$

a, b - линейно зависимы - условие на параметры

c) $f(x) = \langle c, x \rangle \cdot e^{-\langle Ax, x \rangle}$

$df = \langle c, dx \rangle \cdot e^{-\langle Ax, x \rangle} -$

$- \langle c, x \rangle e^{-\langle Ax, x \rangle} \cdot (\langle A dx, x \rangle + \langle Ax, dx \rangle)$

$$\langle A\,dx, x\rangle = tr(dx^T A^T x) = tr(x^T A\,dx) =$$

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

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

$$\nabla f = e^{-\langle Ax, x\rangle}(c - \langle c, x\rangle \cdot 2Ax) = 0$$

$e^{-\langle Ax, x\rangle} \neq 0$, где $\forall x \in \mathbb{R}^n$

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

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

A^{-1} существует, т.к. $\det(A) \neq 0$, $A \in S^n_{++}$

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

$$x = \lambda A^{-1} c$$

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

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

$\langle c, A^{-1} c\rangle$ должно быть больше 0

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

$$X = \sqrt{\frac{1}{2\langle C, A^{-1}C\rangle}} \cdot A^{-1}C$$

$\boxed{7}$

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

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

Используем тождество Вудбери

$A = X^k$; $U = V = X^k$; $C = I_n$

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

$$\circeq \lim_{k\to+\infty} \operatorname{tr}(I_n + X^k)^{-1} = \lim_{k\to+\infty} \sum_{i=1}^{n} \frac{1}{1+\lambda_i^k} =$$

$$= \sum_{i=1}^{n} \lim_{k\to+\infty} \frac{1}{1+\lambda_i^k} \circeq$$

λ_i - собств. знач. X

Тогда собств. знач. $(I_n + X^k)^{-1}$ будут $\frac{1}{1+\lambda_i^k}$ и $1+\lambda_i^k \neq 0$ $\forall k \geq 0$, т.к. $\lambda_i > 0$ (X - положит. опр.)

Есть 3 случая:

1) $\lambda_i = 1 \Rightarrow \lim\limits_{k\to+\infty} \frac{1}{1+\lambda_i^k} = \frac{1}{2}$

2) $0 < \lambda_i < 1 \Rightarrow \lim\limits_{k\to+\infty} \frac{1}{1+\lambda_i^k} = 1$

3) $1 < \lambda_i \Rightarrow \lim\limits_{k\to+\infty} \frac{1}{1+\lambda_i^k} = 0$

$$\circledast \begin{cases} 0, & \text{если } \forall i \quad 1 < \lambda_i \\ m \cdot \frac{1}{2} + \ell, & \text{где } m - \text{кол-во } \lambda_i = 1 \\ & \phantom{\text{где }} \ell - \text{кол-во } \lambda_i < 1 \end{cases}$$

8 a) $F(P) = N \operatorname{tr}\left(\left(I - P(P^T P)^{-1} P^T \right)^2 S \right)$

$dQ = d\left[I - P(P^T P)^{-1} P^T \right] = -d\left[P(P^T P)^{-1} P^T \right] =$

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

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

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

$- P(P^T P)^{-1} dP^T = \{ \text{св-во } P^T P = I \}$

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

$$dQ\,[I - P(P^TP)^{-1}P^T] = dQ\,[I - PP^T]$$

$$= \cancel{-dP\,P^T} + P(\cancel{dP^T P + P^T dP})P^T - P\,dP^T +$$

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

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

$$[I - P(P^TP)^{-1}P^T]\,dQ = [I - PP^T]\,dQ =$$

$$= -dP\,P^T + P(d\cancel{P^T P + P^T dP})P^T - \cancel{P\,dP^T} +$$

$$+ PP^T dP\,P^T - PP^T\cancel{P(dP^T P + P^T dP)}P^T + P\cancel{P^T P\,dP^T} =$$

$$= (PP^T - I_D)\,dP\,P^T$$

$$dF(P) = N\,\mathrm{tr}\Big(\big(dQ\,[I - P(P^TP)^{-1}P^T] +$$

$$+ [I - P(P^TP)^{-1}P^T]\,dQ\big)S\Big) =$$

$$= N\,tr\left(\left(p\,dp^T(pp^T - I_D) + (pp^T - I_D)dp\,p^T\right)S\right)$$

$$= N\,tr\left(S^T(pp^T - I_D)dp\,p^T + (pp^T - I_D)dp\,p^T S\right) =$$

$$= N\,tr\left(p^T S(pp^T - I_D)dp + p^T S(pp^T - I_D)dp\right)$$

$S = S^T$, т.к. $S = \frac{1}{N}\sum_{i=1}^{N} x_i x_i^T$

$$\Rightarrow \nabla_p F(p) = 2N(pp^T - I_D)\cdot Sp$$

б) $S = Q\Lambda Q^T = Q\Lambda Q^{-1}$

$\nabla_p F(p) = 0$, если p состоит из любых d векторов

$Q^{-1} = Q^T \Rightarrow Q^T Q = QQ^T = I$

Пусть матрица p состоит из столбцов $1\ 2 \ldots d$, можем так считать без нарушения общности

Тогда $p \cdot p^T \cdot S \cdot p = p \cdot p^T \cdot Q\Lambda Q^T p$

$$\underset{d\times D}{P^T Q} = \begin{pmatrix} \underbrace{\begin{matrix} 1 & & \emptyset \\ & \ddots & \\ \emptyset & & 1 \end{matrix}}_{d} & \underbrace{\emptyset}_{D-d} \end{pmatrix}$$ $\Rightarrow$ на диагонали

стоят 1, в остальных местах 0.

т.к. $q_i^T \cdot q_j = \begin{cases} 1, & i=j \\ 0, & i \neq j \end{cases}$

$$P^T Q \cdot \Lambda = \begin{pmatrix} \underbrace{\begin{matrix} \lambda_1 & & \emptyset \\ & \lambda_2 \ddots & \\ \emptyset & & \lambda_d \end{matrix}}_{d} & \underbrace{\emptyset}_{D-d} \end{pmatrix}$$

$$P^T \cdot Q \cdot \Lambda \cdot Q^T = \begin{pmatrix} \underbrace{\begin{matrix} \lambda_1 & & \emptyset \\ & \lambda_2 \ddots & \\ \emptyset & & \lambda_d \end{matrix}}_{d} & \underbrace{\emptyset}_{D-d} \end{pmatrix} \begin{pmatrix} q_1 \\ \vdots \\ q_D \end{pmatrix} = Y$$

$$\Rightarrow Y_{ij} = \lambda_i \cdot q_{ij} \quad i = 1 \dots d \quad j = 1 \dots D$$

$$P \cdot Y = (q_1 \dots q_d) \cdot \begin{pmatrix} \lambda_1 q_1 \\ \lambda_2 q_2 \\ \dots \\ \lambda_d q_d \end{pmatrix} = (\lambda_1 q_1 \dots \lambda_d q_d) \cdot \begin{pmatrix} q_1 \\ \vdots \\ q_d \end{pmatrix}$$

$$P \cdot Y \cdot P = (\lambda_1 q_1 \dots \lambda_d q_d) \begin{pmatrix} q_1 \\ \vdots \\ q_d \end{pmatrix} (q_1 \dots q_d) =$$

$$= (\lambda_1 q_1 \dots \lambda_d q_d) \cdot I_d =$$

$$\Rightarrow P P^T Q \Lambda Q^T P = (\lambda_1 q_1 \dots \lambda_d q_d)$$

$$Q^T P = \begin{pmatrix} q_1 \\ \vdots \\ q_D \end{pmatrix} (q_1 \dots q_d) = \begin{array}{c} \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \\ & 0 & \end{pmatrix} \end{array} \begin{array}{l} \}\, d \\ \\ \}\, D \end{array}$$

$$\Lambda Q^T P = \begin{pmatrix} \lambda_1 & & 0 \\ 0 & \ddots & \lambda_d \\ & 0 & \end{pmatrix}$$

$$Q \Lambda Q^T P = (q_1 \dots q_D) \begin{pmatrix} \lambda_1 & & 0 \\ 0 & \ddots & \lambda_d \\ & 0 & \end{pmatrix} = (\lambda_1 q_1 \dots \lambda_d q_d)$$

$$\Rightarrow (P P^T - I) S P = 0 \Rightarrow \nabla f = 0$$

Вторая часть задачи 8b:

$$\frac{F(P)}{N} = tr\left(\left(I - P(P^T P)^{-1} P^T\right)^2 S\right) =$$

$$= tr\left(\left(I - 2P(P^T P)^{-1} P^T + P(P^T P)^{-1} P^T P (P^T P)^{-1} P^T\right) S\right)$$

$$= tr\left(S - P(P^T P)^{-1} P^T S\right) =$$

$$= tr\left(Q\Lambda Q^T - P(P^TP)^{-1}P^T Q\Lambda Q^T\right) =$$

$$= \underbrace{tr\left(Q^TQ\Lambda\right)}_{tr\,\Lambda = const} - tr\left(P(P^TP)^{-1}P^TQ\Lambda Q^T\right)$$

То есть $\arg\min\limits_P F(P) =$

$$= \arg\max_P tr\left(P(P^TP)^{-1}P^TQ\Lambda Q^T\right) =$$

Как мы выяснили в прошлой части 8b

$$P^TQ\Lambda Q^T = Y = \begin{pmatrix}\lambda_1 q_1\\ \vdots\\ \lambda_d q_d\end{pmatrix} = \begin{pmatrix} q_1\\ \vdots \\ q_d\end{pmatrix}\cdot \begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_d\end{pmatrix}$$

$$= \begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_d\end{pmatrix}\begin{pmatrix} q_1\\ \vdots \\ q_d\end{pmatrix}$$

$$= \arg\max_P tr\left(P(P^TP)^{-1}\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_d\end{pmatrix}P^T\right) =$$

$$\arg\max_P tr\left(P^TP(P^TP)^{-1}\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_d\end{pmatrix}\right) =$$

$= \arg\max\limits_P \sum\limits_{i=1}^{d} \lambda_i = P'$, где P' состоит из собств. векторов, отвечающих наибольшим собств. знач.

В решениях я использовала формулы, выведенные в файле sem05_vect_matrix_diff.pdf, который выложен на гитхаб.

В частности: $f(x) = \|x\|_2 \quad x \in \mathbb{R}^n \setminus \{0\}$

$$\Rightarrow df(x) = \langle \|x\|_2^{-1} x, dx \rangle$$