Задание 1.

а) $f: \mathbb{R}_{++} \to \mathbb{R}$ $A = A^t \succcurlyeq 0$

Под $\|\cdot\|$ будет пониматься $\|\cdot\|_2$

$f(t) = \|(A+tI)^{-1}b\|_2^2$

$g(x) = \|x\|^2$ $h(X) = X^{-1}$

$dg = \langle 2x, dx \rangle$ $dh = -X^{-1}(dX)X^{-1}$

$$df = \langle 2(A+tI)^{-1}b, d((A+tI)^{-1}b)\rangle = \langle 2(A+tI)^{-1}b, (d(A+tI)^{-1})b\rangle =$$
$$= \langle 2(A+tI)^{-1}b, -(A+tI)^{-1}d(A+tI)(A+tI)^{-1}b\rangle =$$
$$= \langle 2(A+tI)^{-1}b, -(A+tI)^{-1}dt(A+tI)^{-1}b\rangle =$$
$$= \langle 2(A+tI)^{-1}b, -(A+tI)^{-2}b\,dt\rangle =$$

т.к $dt \in \mathbb{R}$

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

еще заметим, что $(A+tI)^t = (A+tI)$
$(A+tI)^{-t} = (A+tI)^{-1}$

$$= -2b^t(A+tI)^{-3}b \cdot dt$$

по размерности важно

$$\Rightarrow \nabla f = -2b^t(A+tI)^{-3}b \in \mathbb{R}$$

Фиксируем $dt = dt_1$

$$d^2 f = d(-2b^t(A+tI)^{-3}b\,dt_1) = -2b^t\,d(A+tI)^{-3}b \cdot dt_1 =$$
$$= -2b^t(-3)(A+tI)^{-4} \cdot b \cdot dt_2\,dt_1 = 6b^t(A+tI)^{-4}b\,dt_2\,dt_1$$

т.к $\nabla^2 f \in \mathbb{R}$, то $dt_2 = dt_1 = dt$ и

$$\nabla^2 f = 6b^t(A+tI)^{-4}b$$

б) $f: \mathbb{R}^n \to \mathbb{R}$ $f(x) = \frac{1}{2}\|xx^t - A\|_F^2$ $A = A^t$

$$df = \frac{1}{2}d(\langle xx^t - A, xx^t - A\rangle_F) = \frac{1}{2}d(tr((xx^t-A)^t(xx^t-A))) =$$
$$= \frac{1}{2}\cdot d(tr(xx^txx^t - xx^tA - Axx^t + A^2)) =$$

$tr(xx^tA) = tr(Axx^t)$

$$= \frac{1}{2}d(tr(xx^txx^t - 2xx^tA + A^2)) =$$
$$= \frac{1}{2}tr(d(xx^txx^t) - 2d(xx^t)A) =$$
$$= \frac{1}{2}tr(d(xx^txx^t) - 2((dx)x^t + x(dx)^t)A) =$$
$$= \frac{1}{2}tr(x\,d(x^txx^t) + (dx)x^txx^t - 2((dx)x^t + x(dx)^t)A) =$$
$$= \frac{1}{2}tr(x((dx)^txx^t + x^t((dx)x^t + x(dx)^t)) + (dx)x^txx^t - 2((dx)x^t + x(dx)^t)A) =$$
$$= \frac{1}{2}tr(xx^tx(dx)^t + xx^t(dx)x^t + xx^tx(dx)^t + (dx)x^txx^t - 2(dx)x^tA - 2x(dx)^tA) =$$
$$= \frac{1}{2}[tr((dx)x^txx^t) + tr(x^txx^tdx) + tr((dx)x^txx^t) + tr(x^txx^tdx) - 2tr(x^tAdx) - 2tr(x^tAdx)] =$$
$$= \frac{1}{2}[4tr(x^txx^tdx) - 4tr(x^tAdx)] = 2tr((x^txx^t - x^tA)dx) =$$
$$= 2tr((\underbrace{xx^tx - Ax}_{\in\mathbb{R}})^t dx) = 2\langle \langle x,x\rangle x - Ax, dx\rangle$$

$$\Rightarrow \nabla f = 2(\langle x,x\rangle x - Ax) = 2(\langle x,x\rangle I - A)x$$

$h_i = (\sum_j x_j^2)x_i$

$h = \langle x,x\rangle x \in \mathbb{R}^n$

$$\frac{\partial h_i}{\partial x_j} = \begin{cases} 2x_i^2 + \sum_k x_k^2, & i = j \\ 2x_ix_j, & i \neq j \end{cases}$$

$\nabla h = \langle x,x\rangle I + 2xx^t$

$$\Rightarrow \nabla^2 f = 2(\langle x,x\rangle I + 2xx^t - A)$$

(c) $f: \mathbb{R}^n \setminus \{0\} \to \mathbb{R}$

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

Рассмотрим ф-ию

$g(x) = \langle x,x\rangle \ln(\langle x,x\rangle)$

Тогда $f(x) = e^{g(x)}$

$\nabla f(x) = e^{g(x)} \cdot \nabla g(x)$

$$dg = (d\langle x,x\rangle)\ln(\langle x,x\rangle) + \langle x,x\rangle\, d\ln(\langle x,x\rangle) =$$
$$= 2x^t(dx)\ln\langle x,x\rangle + \langle x,x\rangle\cdot\frac{1}{\langle x,x\rangle}\cdot 2x^t dx =$$
$$= (2\ln\langle x,x\rangle\, x + 2x)^t dx$$

$\nabla g = 2(\ln\langle x,x\rangle + 1)x$

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

$$\nabla^2 f(x) = (e^{g(x)}\nabla g(x))(\nabla g(x))^t + e^{g(x)}\cdot\nabla^2 g(x) =$$
$$= e^{g(x)}[\nabla g(x)(\nabla g(x))^t + \nabla^2 g(x)]$$

$$d^2 g = 2\, d(\langle(\ln\langle x,x\rangle+1)x, dx_1\rangle) = 2\langle d((\ln\langle x,x\rangle+1)x), dx_1\rangle =$$
$$= 2\langle \frac{1}{\langle x,x\rangle}2x^t dx_2\cdot x + (\ln\langle x,x\rangle+1)dx_2, dx_1\rangle =$$
$$= 2\left(\left(\frac{2x^t dx_2\, x}{\langle x,x\rangle}\right)^t dx_1 + (\ln\langle x,x\rangle+1)(dx_2)^t dx_1\right) =$$
$$= 2\left(\frac{2}{\langle x,x\rangle} x^t(dx_2)^t x\, dx_1 + (\ln\langle x,x\rangle+1)(dx_2)^t dx_1\right) =$$
$$= 2\left(\frac{2}{\langle x,x\rangle}(dx_2)^t x x^t dx_1 + (\ln\langle x,x\rangle+1)(dx_2)^t dx_1\right) =$$
$$= \langle\left(\frac{4}{\langle x,x\rangle} xx^t + 2(\ln\langle x,x\rangle+1)I\right)dx_2, dx_1\rangle$$

$$\nabla^2 g = \frac{4}{\langle x,x\rangle}xx^t + 2(\ln\langle x,x\rangle+1)I$$

$$\Rightarrow \nabla^2 f = e^{g(x)}\left[4(\ln\langle x,x\rangle+1)^2 xx^t + \frac{4}{\langle x,x\rangle}xx^t + 2(\ln\langle x,x\rangle+1)I\right]$$

Задание 2.

a) $f: E\to\mathbb{R}$ $\quad E = \{x\in\mathbb{R}^n \mid Q(x)<0\}$ $\quad A = A^t \geq 0$

$f(x) = \ln(-Q(x))$ $\quad Q(x) = \frac12\langle Ax,x\rangle + \langle b,x\rangle + c$

$$df = \frac{1}{-Q(x)}\cdot -\left(\frac12 2\langle Ax,dx\rangle + \langle b,dx\rangle\right) = \langle\frac{Ax+b}{Q(x)}, dx\rangle$$

$$d^2 f = d\left(\langle\frac{Ax+b}{Q(x)}, dx_1\rangle\right) \underset{\uparrow}{=} \langle d\left(\frac{Ax+b}{Q(x)}\right), dx_1\rangle = \langle\frac{A(dx_2)Q(x) - (Ax+b)\langle Ax+b, dx_2\rangle}{Q^2(x)}, dx_1\rangle =$$

линейность дифферец.

$$= \langle\frac{A\cdot Q(x) - (Ax+b)(Ax+b)^t}{Q^2(x)}dx_2, dx_1\rangle$$

$$\nabla^2 f(x) = \frac{A\cdot Q(x) - (Ax+b)(Ax+b)^t}{Q^2(x)}$$

Хотим показать, что $\forall q\in E$, $\langle\nabla^2 f(x)\cdot q, q\rangle \overset{?}{\vee} 0$

т.к. $Q(x)<0$ $\forall x\in E$, то $Q^2(x)>0$ $\forall x\in E$ и не влияет на знак

Знаем, что $A\geq 0 \Rightarrow \forall q\in E$, $q^t A\cdot \underbrace{Q(x)}_{<0}\cdot q \leq 0$

Обозначим за $y = Ax+b$

yy^t – м-ца ранга 1

все ее миноры порядка ≥ 2 равны 0 (в том числе главные миноры, к-ые нужны для кр. Сильвестра для неотрицательной кв. ф-ии)

миноры порядка 1 это $y_i^2 \geq 0$

$y = (y_1 \dots y_n)^t \in\mathbb{R}^n$

$\Rightarrow$ по кр. Сильвестра $(Ax+b)(Ax+b)^t \geq 0$ – неотриц. опр. м-ца

а $-yy^t \leq 0$

$$\Rightarrow \frac{A\cdot Q(x) - (Ax+b)(Ax+b)^t}{Q^2(x)} \leq 0$$

(b) $f : \mathbb{R}^n_{++} \to \mathbb{R} \qquad f(x) = \left(\sum_j x_j^p\right)^{\frac{1}{p}} \quad p < 1$

$$\frac{\partial f}{\partial x_i} = \frac{1}{p}\left(\sum_k x_k^p\right)^{\frac{1}{p}-1} \cdot p\, x_i^{p-1} = \left(\sum_k x_k^p\right)^{\frac{1}{p}-1} \cdot x_i^{p-1}$$

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \left(\frac{1}{p} - 1\right)\left(\sum_k x_k^p\right)^{\frac{1}{p}-2} p\, x_j^{p-1}\, x_i^{p-1} \qquad i \neq j$$

$$\frac{\partial^2 f}{\partial x_i^2} = \left(\frac{1}{p} - 1\right)\left(\sum_k x_k^p\right)^{\frac{1}{p}-2} p\, x_i^{p-1} \cdot x_i^{p-1} + \left(\sum_k x_k^p\right)^{\frac{1}{p}-1}(p-1)\, x_i^{p-2}$$

$$\Rightarrow d^2 f(x) = \underset{\vee\, 0}{(1-p)} \underbrace{\left(\sum_k x_k^p\right)^{\frac{1}{p}-2}}_{\vee\, 0} \left(x^{p-1}(x^{p-1})^t - \left(\sum_k x_k^p\right) \cdot \mathrm{diag}(x^{p-2}) \right)$$

Обознач.: $x^{p-1} = (x_1^{p-1}, \ldots, x_n^{p-1})$

$x_j \in \mathbb{R}^n_{++}$

$$d^2 f(x)[h,h] = \overset{>0}{C} \left(h^t x^{p-1} \cdot (x^{p-1})^t h - h^t \left(\sum_k x_k^p\right) \cdot \mathrm{diag}(x^{p-2})\, h \right) =$$

$$= C \left(\langle x^{p-1}, h \rangle^2 - \left(\sum_k x_k^p\right)\left(\sum_k x_k^{p-2} \cdot h_k^2\right)\right) \underset{\text{К-Б}}{\leq} C\left(\|x^{\frac{p}{2}}\|^2 \cdot \|x^{\frac{p}{2}-1} \circ h\|^2 - \|x^{\frac{p}{2}}\|^2 \cdot \|x^{\frac{p}{2}-1} \circ h\|^2 \right) = 0$$

$\langle x^{p-1}, h \rangle^2 = \langle x^{\frac{p}{2}}, x^{\frac{p}{2}-1} \circ h \rangle^2 \leq \|x^{\frac{p}{2}}\|^2 \cdot \|x^{\frac{p}{2}-1} \circ h\|^2 \ ; \quad \sum_k x_k^{p-2} \cdot h_k^2 = \|x^{\frac{p}{2}-1} \circ h\|^2$

$\circ$ - адамарово произв.

$\sum_k x_k^p = \|x^{\frac{p}{2}}\|^2$

$$\Rightarrow d^2 f(x) \leq 0$$

(c) $f : S^n_{++} \to \mathbb{R}$

$$f(X) = \langle X^{-1}, A \rangle_F \qquad A \in S^n_+$$

$$df = d\langle A, X^{-1} \rangle = \langle A, dX^{-1} \rangle =$$

$$= \langle A, -X^{-1}(dX)X^{-1} \rangle = \langle -X^{-t} A X^{-t}, dX \rangle =$$

$$= \langle -X^{-1} A X^{-1}, dX \rangle$$

$(X^{-1})^t = (X^t)^{-1} = X^{-1}$

Фикс $dX = dX_1$

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

$$+ X^{-1} A(-X^{-1} dX_2 X^{-1}), dX_1 \rangle = \langle X^{-1} dX_2 X^{-1} A X^{-1} +$$

$$+ X^{-1} A X^{-1} dX_2 X^{-1}, dX_1 \rangle$$

это какая-то билин. ф-ма

$$d^2 f[H,H] = \langle X^{-1} H X^{-1} A X^{-1} + X^{-1} A X^{-1} H X^{-1}, H \rangle \ominus$$

X - пол. опр. $\qquad \exists\, U : U^{-1} = U^t \quad X = U \Lambda U^t$

$X = X^t \Rightarrow$

$\Lambda = \mathrm{diag}(\lambda_1, \ldots \lambda_n)$

$\lambda_j > 0$ - с.ч.

Тогда $X^{-1} = (U \Lambda U^t)^{-1} = U \Lambda^{-1} U^t$ - собств. разлож.

у X^{-1} с.ч. $\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_n} > 0$

$\Rightarrow X^{-1} > 0 \Rightarrow \exists!\, (X^{-1})^{\frac{1}{2}}$ - корень $\quad X^{-\frac{1}{2}} = U \Lambda^{-\frac{1}{2}} U^t > 0$

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

$$\ominus \langle X^{-\frac{1}{2}} \underline{X^{-\frac{1}{2}} H X^{-\frac{1}{2}}}\, \underline{X^{-\frac{1}{2}} A X^{-\frac{1}{2}}} X^{-\frac{1}{2}} + X^{-\frac{1}{2}} \underline{X^{-\frac{1}{2}} A X^{-\frac{1}{2}}}\, \underline{X^{-\frac{1}{2}} H X^{-\frac{1}{2}}} X^{-\frac{1}{2}}, H \rangle =$$

$$= \langle X^{-\frac{1}{2}} H X^{-\frac{1}{2}} X^{-\frac{1}{2}} A X^{-\frac{1}{2}}, X^{-\frac{1}{2}} H X^{-\frac{1}{2}} \rangle + \langle X^{-\frac{1}{2}} A X^{-\frac{1}{2}} X^{-\frac{1}{2}} H X^{-\frac{1}{2}}, X^{-\frac{1}{2}} H X^{-\frac{1}{2}} \rangle =$$

$$= \langle \underbrace{X^{-\frac{1}{2}} H X^{-\frac{1}{2}}}_{B}\, \underbrace{X^{-\frac{1}{2}} A X^{-\frac{1}{2}}}_{C}, \ldots$$

$B = B^t, \ C = C^t$

$$= \langle BC, B \rangle + \langle CB, B \rangle = 2\langle C, BB \rangle = 2\langle X^{-\frac{1}{2}} A X^{-\frac{1}{2}}, X^{-\frac{1}{2}} H X^{-\frac{1}{2}} X^{-\frac{1}{2}} H X^{-\frac{1}{2}} \rangle =$$

$A \in S^n_+ \Rightarrow \exists\, A^{\frac{1}{2}}$

$$= 2\langle \underline{X^{-\frac{1}{2}} A^{\frac{1}{2}}} A^{\frac{1}{2}} X^{-\frac{1}{2}}, X^{-\frac{1}{2}} H X^{-\frac{1}{2}} X^{-\frac{1}{2}} H X^{-\frac{1}{2}} \rangle =$$

$$= 2\langle A^{\frac{1}{2}} X^{-\frac{1}{2}}, A^{\frac{1}{2}} X^{-\frac{1}{2}} X^{-\frac{1}{2}} H X^{-\frac{1}{2}} \underline{X^{-\frac{1}{2}} H X^{-\frac{1}{2}}} \rangle =$$

$$= 2\langle A^{\frac{1}{2}} X^{-\frac{1}{2}}, A^{\frac{1}{2}} X^{-\frac{1}{2}} X^{-\frac{1}{2}} H X^{-\frac{1}{2}} X^{-\frac{1}{2}} H X^{-\frac{1}{2}} \rangle =$$

$$= 2\|A^{\frac{1}{2}} X^{-1} H X^{-\frac{1}{2}}\|_F^2 \geq 0$$

$$\Rightarrow d^2 f \geq 0$$

(d) $f: S^n_{++} \to \mathbb{R}$ $\quad f(X) = (\det X)^{\frac{1}{n}} = e^{g(X)}$

$g(X) = \frac{1}{n}\ln(\det X)$

Знаем, что $d(\det X) = \det X \langle X^{-t}, dX \rangle = \det X \langle X^{-1}, dX \rangle$ т.к $X \in S^n_{++}$

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

$$d^2 f = d\left(e^{g(X)} \cdot \frac{1}{n} \langle X^{-1}, dX_1 \rangle\right) = \frac{1}{n}\left(e^{g(X)} \cdot \frac{1}{n} \langle X^{-1}, dX_2 \rangle \langle X^{-1}, dX_1 \rangle + e^{g(X)} \langle -X^{-1} dX_2 X^{-1}, dX_1 \rangle\right) = \frac{(\det X)^{\frac{1}{n}}}{n}\left[\frac{1}{n}\langle X^{-1}, dX_2\rangle\langle X^{-1}, dX_1\rangle - \langle X^{-1} dX_2 X^{-1}, dX_1 \rangle\right]$$

Посчитаем значение кв. формы на $dX_1 = dX_2 = H$

$\frac{\det X^{\frac{1}{n}}}{n} > 0$ т.к $X \in S^n_{++}$, можем не учитывать

$$d^2 f[H,H] = \underbrace{\frac{(\det X)^{\frac{1}{n}}}{n}}_{C>0=}\left[\frac{1}{n}\langle X^{-1}, H\rangle^2 - \langle X^{-1} H X^{-1}, H \rangle\right] = C\left(\frac{1}{n}\langle I, X^{-\frac{1}{2}} H X^{-\frac{1}{2}}\rangle^2 - \langle X^{-\frac{1}{2}} H X^{-\frac{1}{2}}, X^{-\frac{1}{2}} H X^{-\frac{1}{2}} \rangle\right) \le$$

$\exists X^{-\frac{1}{2}}$, т.к $X \in S^n_{++}$

$$\underset{\text{К-Б}}{\le} C\left(\frac{1}{n} \underbrace{\|I\|_F^2}_{n} \cdot \|X^{-\frac{1}{2}} H X^{-\frac{1}{2}}\|_F^2 - \|X^{-\frac{1}{2}} H X^{-\frac{1}{2}}\|_F^2\right) = 0$$

$\Rightarrow d^2 f \le 0$

Задание 3

(а) • $r_k = (0.99)^{2^k}$

$\frac{r_{k+1}}{r_k} = \frac{0.99^{2^{k+1}}}{0.99^{2^k}} = 0.99^{2^k} \xrightarrow[k \to +\infty]{} 0$ $\quad$ верх. пред = нижн. предел = предел.

$\Rightarrow r_k$ имеет суперлинейн. скор. сход.

$r_k^{\frac{1}{2^k}} \to 0.99$ $\Rightarrow$ имеет место квадр. сходимость по тесту корней для сверх. лин. сход порядка 2.

• $r_k = \frac{1}{k!}$ $\quad \frac{r_{k+1}}{r_k} = \frac{1}{k+1} \xrightarrow[k \to +\infty]{} 0 \Rightarrow$ суперлиней. ск. сход. $\quad \frac{r_{k+1}}{r_k^2} = \frac{1}{(k+1)!} \cdot (k!)^2 = \frac{k!}{k+1} \to +\infty \Rightarrow$ нет кв. сходимости по тесту отношений

б) • $r_k = \frac{1}{k^k}$ $\quad \overline{\lim_{k \to +\infty}} \left(\frac{1}{k^k}\right)^{\frac{1}{k}} = \overline{\lim_{k \to +\infty}} \frac{1}{k} = 0$ – суперлиней. $\quad r_k^{\frac{1}{2^k}} = (k^{-k})^{\frac{1}{2^k}} = k^{-\frac{k}{2^k}} = e^{-\frac{k}{2^k}\ln k} \to 1 \Rightarrow$ нет кв. сход по тесту корней

• $r_k = \begin{cases} (0.99)^{2^k}, & k = 2q \\ \frac{r_{k-1}}{k}, & k = 2q-1 \end{cases}$

k-нечет $\frac{r_k}{r_{k-1}} = \frac{r_{k-1}}{k(r_{k-1})} = \frac{1}{k} \xrightarrow{k \to +\infty} 0$

k-чет. $\frac{r_k}{r_{k-1}} = \frac{(0.99)^{2^k}}{\frac{r_{k-2}}{k-1}} = \frac{(0.99)^{2^k}}{0.99^{2^{k-2}}}(k-1) = (0.99)^{\frac{3}{4} \cdot 2^k}(k-1) \xrightarrow{k \to +\infty} 0$

$\underline{\lim_{k \to +\infty}} \frac{r_k}{r_{k-1}} = 0 \quad \overline{\lim_{k \to +\infty}} \frac{r_k}{r_{k-1}} = 0$

$\Rightarrow$ суперлин.

k-чет $(r_k)^{\frac{1}{2^k}} \to 0.99$

k-нечет $(r_k)^{\frac{1}{2^k}} = \left(\frac{r_{k-1}}{k}\right)^{\frac{1}{2^k}} = \left(\frac{0.99^{2^{k-1}}}{k}\right)^{\frac{1}{2^k}} = \frac{0.99^{\frac{1}{2}}}{k^{\frac{1}{2^k}}} \to 0.99^{\frac{1}{2}}$ $\quad k^{\frac{1}{2^k}} \to 1$

$\Rightarrow \overline{\lim} r_k^{\frac{1}{2^k}} = 0.99$

$\Rightarrow$ кв. скор. сход.

Задание 4

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

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

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

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

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

x – т. стационарности $\forall c \in \mathbb{R}^n \setminus \{0\}$, $\sigma > 0$

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

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

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

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

$ab^t x = a + b$ $\quad Ax = a + b$ – СЛУ

$A = ab^t$ $\quad rk A = 1$

по т-ме Крон-Кап. совместна $\Leftrightarrow rk(A|a+b) = rk A$

$\Rightarrow a = (a_1, \dots, a_n)^t$, $b = (b_1, \dots, b_n)^t$

$$\left(\begin{array}{ccc|c} a_1 b_1 & \dots & a_1 b_n & a_1 + b_1 \\ a_2 b_1 & \dots & a_2 b_n & \alpha_2(a_1 + b_1) \\ \dots & & & \dots \\ a_n b_1 & \dots & a_n b_n & \alpha_n(a_1 + b_1) \end{array}\right) \to (a_1 b_1 \dots a_1 b_n \mid a_1 + b_1)$$

$$X = \begin{pmatrix} \frac{a_1 + b_1 - c_2 \cdot a_1 b_2 - \dots - c_n a_1 b_n}{a_1 b_1} \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

Чтобы система была совместной $\alpha_j = \frac{a_j}{a_1}$

$c_2, \dots, c_n \in \mathbb{R}$ и такие, что $x \in E$ $\quad a + b = \begin{pmatrix} a_1 + b_1 \\ \frac{a_2}{a_1}(a_1 + b_1) \\ \dots \\ \frac{a_n}{a_1}(a_1 + b_1) \end{pmatrix}$

Это при условии, что $a_1 \ne 0$, $b_1 \ne 0$

Но $a \ne 0$, $b \ne 0$ $\Rightarrow \exists i, j : a_i \ne 0, b_j \ne 0$

Тогда берем самую верх. левую такую позицию

и эл. преобр $(0 \dots 0 \; a_i b_j \dots a_i b_n \mid a_i + b_i)$

$$X = \begin{pmatrix} c \\ 0 \\ \dots \\ \frac{a_i + b_i - c_{j+1} a_i b_{j+1} - \dots - c_n a_i b_n}{a_i b_j} \\ c_{j+1} \\ \dots \\ c_n \end{pmatrix}$$

Также если какой-то $a_k = 0$, то $b_k = 0$

В общем x – решение системы $ab^t x = a + b$

берем a и b такие, что она совместна.

(c) $f: \mathbb{R}^n \to \mathbb{R}$ $\qquad c \in \mathbb{R}^n \setminus \{0\}, \ A \in S^n_{++}$

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

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

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

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

$$c - 2\langle c, x\rangle Ax = 0 \qquad \text{т.к } A \in S^n_{++} \Rightarrow \exists A^{-1}$$

$$A^{-1}c - 2\langle c, x\rangle x = 0$$

$\Rightarrow$ векторы $A^{-1}c$ и x - лин. зависимые

т.е $x = \gamma \cdot A^{-1}c \quad \gamma \in \mathbb{R}$

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

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

$\langle c, A^{-1}c\rangle > 0$ т.к $A^{-1} \in S^n_{++}$

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

$\forall c \in \mathbb{R}^n \setminus \{0\}$
$\forall A \in S^n_{++}$

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

(4d) $f: S^n_{++} \to \mathbb{R}$ $\qquad A = A^t$

$$f(X) = \langle X^{-1}, I\rangle - \langle A, X\rangle$$

$$df = \langle I, -X^{-1}(dX)X^{-1}\rangle - \langle A, dX\rangle =$$

$$= \langle -X^{-1} \cdot X^{-1} - A, dX\rangle$$

$$-X^{-2} = A$$

$$X^{-2} = -A$$

$$\exists A \leq 0 \Rightarrow -A \geq 0$$

$$\Rightarrow \exists B: B^2 = -A$$

$$\Rightarrow X = B^{-1}$$

при $A \leq 0$

Задание 5

$\{x_i\}_{i=1}^N \quad x_i \in \mathbb{R}^D \qquad d \le D$

$P \in \mathbb{R}^{D\times d} \qquad P(P^tP)^{-1}P^t x$ – проекция

$$F(P) = N\cdot tr((I - P(P^tP)^{-1}P^t)^2 S) \to \min_P$$

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

$P(P^tP)^{-1}P^t$ – ортопроектор $\Rightarrow I - P(P^tP)^{-1}P^t$ – ортопроектор

$\Rightarrow (I - P(P^tP)^{-1}P^t)^2 = I - P(P^tP)^{-1}P^t$

$$F(P) = N\cdot tr((I - P(P^tP)^{-1}P^t) S)$$

$$dF = -N\,tr(d(P(P^tP)^{-1}P^t)\,S) = -N\,tr(dP\,(P^tP)^{-1}P^tS - P(P^tP)^{-1}(P^t dP + (dP)^tP)(P^tP)^{-1}P^tS + P(P^tP)^{-1}dP^tS) \;(=)$$

- $d(P^tP)^{-1} = -(P^tP)^{-1}d(P^tP)(P^tP)^{-1} = -(P^tP)^{-1}(P^tdP + (dP^t)P)(P^tP)^{-1}$
- $d(P^tP) = P^tdP + (dP^t)P$
- $d((P^tP)^{-1}P^t) = -(P^tP)^{-1}(P^tdP + (dP)^tP)(P^tP)^{-1}P^t + (P^tP)^{-1}dP^t$

$(=) -N(\langle dP, SP(P^tP)^{-1}\rangle - \langle P(P^tP)^{-1}P^tdP, SP(P^tP)^{-1}\rangle - \langle P(P^tP)^{-1}(dP)^tP, SP(P^tP)^{-1}\rangle + \langle P(P^tP)^{-1}, S\,dP\rangle) = -N(\langle SP(P^tP)^{-1}, dP\rangle - \langle P(P^tP)^{-1}P^tSP(P^tP)^{-1}, dP\rangle - \langle P(P^tP)^{-1}, SP(P^tP)^{-1}P^tdP\rangle + \langle SP(P^tP)^{-1}, dP\rangle) =$

$= -N\langle SP(P^tP)^{-1} - P(P^tP)^{-1}P^tSP(P^tP)^{-1} - P(P^tP)^{-1}P^tSP(P^tP)^{-1} + SP(P^tP)^{-1}, dP\rangle = -N\langle 2SP(P^tP)^{-1} - 2P(P^tP)^{-1}P^tSP(P^tP)^{-1}, dP\rangle$

(*) Теперь пользуемся тем, что $P^tP = I$

$\overset{(*)}{=} -2N\langle SP - PP^tSP, dP\rangle$

$$\nabla f(P) = -2N(SP - PP^tSP)$$

б) $S = Q\Lambda Q^t$

$Q = (q_1 \dots q_D), \; Q^tQ = I$

Рассмотрим матрицу перестановки $C \in \mathbb{R}^{D\times d}$ (ровно одна „1" в каждом столбце, строки могут быть нулевые) такую, что $C^tC = I$. С помощью нее выбираются столбцы Q.

$\exists\, P = Q\cdot C$. Тогда $\nabla f(P) = -2N(SP - PP^tSP) = -2N(Q\Lambda \underbrace{Q^tQ}_{I} C - QC\cdot C^t\underbrace{Q^tQ}_{I}\Lambda \underbrace{Q^tQ}_{I} C) = -2N(Q\Lambda C - QCC^t\Lambda C) = -2N[Q(\Lambda C - CC^t\Lambda C)] \;(=)$

$= -2N(Q\Lambda C - QCC^t\Lambda C)$

Пусть C: выбрали $q_{i_1}, \dots, q_{i_d} \Rightarrow \Lambda C = [\lambda_{i_1} e_{i_1}, \dots, \lambda_{i_d} e_{i_d}]$, $e \in \mathbb{R}^D$

λ_{i_k} – с.ч. для q_{i_k}, $k = 1\dots d$; e_{i_k} – единичный орт

$CC^t = diag([q_1 \in B], \dots, [q_D \in B]) \Rightarrow CC^t\Lambda C = diag([q_1 \in B], \dots, [q_D \in B])\cdot[\lambda_{i_1}e_{i_1}, \dots, \lambda_{i_d}e_{i_d}] = ([q_{i_1} \in B]\lambda_{i_1}e_{i_1}, \dots, [q_{i_d} \in B]\lambda_{i_d}e_{i_d}) = \Lambda C$

B – мн-во столбцов q_i из Q, кот-ые выбрали

$[q_i \in B]$ – индикатор, что взяли i-ый вектор

$(=) -2N(Q(\Lambda C - \Lambda C)) = 0$

На самом деле P – любая м-ца полного ранга

$\exists\, P = Q\cdot R$, $P \in \mathbb{R}^{D\times d}$, $R \in \mathbb{R}^{D\times d}$; R – полного ранга, чтобы $\exists (R^tR)^{-1}$

$\nabla f(P) = -2N(SP(P^tP)^{-1} - P(P^tP)^{-1}P^tSP(P^tP)^{-1}) =$

$= -2N(Q\Lambda Q^tQR(R^tR)^{-1} - QR(R^tR)^{-1}R^tQ^tQ\Lambda Q^tQR(R^tR)^{-1}) =$

$= -2N(Q\Lambda R(R^tR)^{-1} - Q\underbrace{R(R^tR)^{-1}R^t}_{\text{ортопроектор}}\Lambda R(R^tR)^{-1}) =$

$= -2N(Q(\Lambda - R(R^tR)^{-1}R^t\Lambda)R(R^tR)^{-1})$

Если $\Lambda \in Im(R)$, т.е. $\Lambda = R\cdot U^t$ – скелетное разложение, $U \in \mathbb{R}^{D\times d}$

$\Rightarrow R(R^tR)^{-1}R^t\underbrace{RU^t}_{\Lambda} = RU^t = \Lambda$

$\Rightarrow$ для таких R и U $\nabla f(P) = 0$

Док-ем, что минимум достигается на собств. векторах, соотв. максим. с.ч.

$F(P) = N\,tr((I - PP^t)S) = N\,tr(Q\Lambda Q^t - QCC^tQ^tQ\Lambda Q^t) =$

$P = Q\cdot C$

$= N\,tr(\Lambda - CC^t\Lambda) \to \min_C$

$\Rightarrow tr(CC^t\Lambda) \to \max_C$

$CC^t\Lambda = diag([q_1 \in B]\lambda_1, \dots, [q_D \in B]\lambda_D)$

$\Rightarrow$ нужно q_j такие, что они соответ. максимальным с.ч.

Почему нельзя лучше?

По условию мы ищем минимум $F(P)$ на пространстве ортогон. матриц $P: P^tP = I$

$\Rightarrow F(P) = N\,tr((I - PP^t)S) = N\,tr(Q\Lambda Q^t - PP^tQ\Lambda Q^t) =$

$= N\,tr(\Lambda - PP^tQ\Lambda Q^t) \;(=)$

P – ортогональная, можем представить ее как $P = Q\cdot L$, $Q \in \mathbb{R}^{D\times d}$, $L \in \mathbb{R}^{D\times d}$, $L^tL = I$

$\Rightarrow P^tP = L^tQ^tQL = I$

$(=) N\,tr(\Lambda - QLL^tQ^tQ\Lambda Q^t) =$

$= N\,tr(\Lambda - LL^t\Lambda) = (*)$

$$tr(LL^t\Lambda) = \sum_{i=1}^D (LL^t\Lambda)_{ii} = \sum_{i=1}^D\sum_{k=1}^D (LL^t)_{ik}\Lambda_{ik} = \sum_{i=1}^D (LL^t)_{ii}\lambda_i =$$

$$= \sum_{i=1}^D\left(\sum_{k=1}^d L_{ik}^2\right)\lambda_i = \sum_{i=1}^D \lambda_i\cdot\|l_i\|_2^2 \le \sum_{i=1}^D \lambda_i [q_i \in B]$$

$L = \begin{bmatrix} l_1 \\ \vdots \\ l_D \end{bmatrix}$, $l_j \in \mathbb{R}^d$ Для м-цы с ортогонал. столбцами норма строк $\|l_j\| \le 1$

Также $\lambda_i \ge 0$, $i = 1\dots D$

т.к м-ца ковариации неотриц. определ.

Чтобы максимизировать эту сумму

т.к каждое слагаемое неотрицат.

хотим перед бОльшими λ_i ставить бОльший вес $\|l_i\|^2 \le 1$

$\Rightarrow$ нужно расставить d „1" в строках с номерами наибол. с.ч

(d „1" т.к норма столбцов равна 1, а их d штук)

$\Rightarrow L$ – это м-ца перестановки C, выбирающей q_j

$(*) \ge N\,tr(\Lambda - CC^t\Lambda) = F(P^*)$

$P^* = Q\cdot C$

$\{ f(x) = \frac{\langle Ax, x\rangle}{\|x\|^2}, \ \|x\|^2 = 1 \}$

$L(x,\lambda) = \langle Ax, x\rangle - \lambda x^t x$

$dL = \langle 2Ax, dx\rangle - 2\lambda \langle x, dx\rangle = 0$

$Ax = \lambda x \quad x$ – с.в

Задание 6.

$f: \mathbb{R}^n \setminus \{0\} \to \mathbb{R} \qquad A \in S^n$

$f(x) = \frac{\langle Ax, x\rangle}{\|x\|^2}$ – отношение Релея

$$df = \frac{d(\langle Ax, x\rangle)\cdot \|x\|^2 - \langle Ax, x\rangle \cdot d(\|x\|^2)}{\|x\|^4} =$$

$$= \frac{\langle 2Ax, dx\rangle \cdot \|x\|^2 - \langle Ax, x\rangle \langle 2x, dx\rangle}{\|x\|^4} =$$

$$= \left\langle \frac{2Ax}{\|x\|^2} - \frac{2\langle Ax, x\rangle x}{\|x\|^4}, dx \right\rangle = 2\left\langle \frac{Ax}{\|x\|^2} - \frac{f(x) x}{\|x\|^2}, dx\right\rangle = 2\left\langle \frac{(A - f(x)\cdot I)x}{\|x\|^2}, dx\right\rangle$$

$\Downarrow$

$A \frac{x}{\|x\|^2} = f(x) \frac{x}{\|x\|^2}$

$\Rightarrow \frac{x}{\|x\|^2}$ – с.в с с.з $f(x)$

Подставим $x = q_j$ – с.в A $\quad \|q\|_2 = 1$

(т.к $A = A^t \Rightarrow \exists Q, \Lambda : A = Q\Lambda Q^t$)

в производную

$$\frac{Aq_j - \frac{\langle Q\Lambda Q^t q_j, q_j\rangle}{\|q_j\|^2} q_j}{\|q_j\|^2} = \lambda_j q_j - \left\langle Q\Lambda \begin{pmatrix} q_1^t \\ \vdots \\ q_n^t \end{pmatrix} q_j, q_j \right\rangle q_j =$$

$= \lambda_j q_j - \langle Q\Lambda e_j, q_j\rangle q_j = \lambda_j q_j - \langle \lambda_j e_j, e_j\rangle q_j =$

$= \lambda_j q_j - \lambda_j q_j = 0$

$x = Qy \qquad A \cdot Qy - \frac{\langle AQy, Qy\rangle}{\langle Qy, Qy\rangle} Qy = Q\Lambda y - \frac{\langle \Lambda y, y\rangle}{\langle y, y\rangle} Qy = 0$

$Q\left(\Lambda y - \frac{\langle \Lambda y, y\rangle}{\langle y, y\rangle} y\right) = 0 \quad Q$ – полного ранга $\Rightarrow \exists$ тривиал. нулевое решение системы

$\Lambda y = \frac{\langle \Lambda y, y\rangle}{\langle y, y\rangle} y \iff \lambda_j y_j = \frac{\sum_k \lambda_k y_k^2}{\|y\|^2} y_j \qquad \forall j = 1..n$

$\Lambda y = f_\Lambda(y) y \qquad (\Lambda - f_\Lambda(y)) y = 0$

$$d^2 f = 2\left\langle \frac{(A dx_2 - \cancel{\langle Ax, dx_2\rangle}^{0} - f(x) dx_2)\|x\|^2 - (A - f(x) I)x \cdot 2x^t dx_2}{\|x\|^4}, dx_1 \right\rangle =$$

$$= 2\left\langle \frac{\|x\|^2 (A - f(x)\cdot I) - 2(A - f(x) I) x x^t}{\|x\|^4} dx_2, dx_1 \right\rangle =$$

$Q\Lambda y = \frac{\langle \Lambda y, y\rangle}{\langle y, y\rangle} Qy$

т.к т. стационар.

(1) $\nabla f(x)|_{x=q_j} = 0$ в точке стац. q_j – с.в

$= 2\left\langle \frac{(A - f(x)\cdot I)(\|x\|^2 I - 2x x^t)}{\|x\|^4} dx_2, dx_1 \right\rangle$

$(A - \lambda_j \cdot I)(I - q_j q_j^t) =$

$= A - A q_j q_j^t - \lambda_j I + \lambda_j q_j \cdot q_j^t =$

$= A - \lambda_j q_j q_j^t - \lambda_j I + \lambda_j q_j \cdot q_j^t =$

$= A - \lambda_j \cdot I$ – надо определить пол./отр. опред.

(2) $x = Qy$ – т. стац. $\qquad \|x\| = \|y\|$

$\left(A - \frac{\langle \Lambda y, y\rangle}{\langle y, y\rangle} I\right)(\langle y, y\rangle \cdot I - 2Qyy^tQ^t) =$

$= \langle y, y\rangle A - 2Q\Lambda \cancel{Q^t}^{I} Q yy^t Q^t - \langle \Lambda y, y\rangle \cdot I + 2\frac{\langle \Lambda y, y\rangle}{\langle y, y\rangle} Qyy^tQ^t =$

$= \langle y, y\rangle A - 2Q\Lambda yy^tQ^t - \langle \Lambda y, y\rangle \cdot I + 2Q\Lambda yy^tQ^t =$

$= \langle y, y\rangle A - \langle \Lambda y, y\rangle \cdot I$

$$(1) \ u^t(A - \lambda_j I)u = v^t Q^t (Q\Lambda Q^t - \lambda_j I) Qv = v^t(\Lambda - \lambda_j \cdot I)v = \sum_{k=1}^n v_k^2 (\underbrace{\lambda_k - \lambda_j}_{\geq 0})$$

$\exists u \in \mathbb{R}^n \setminus \{0\}: u = Q \cdot v$

если x – собств. вект. для λ_{min} $\Rightarrow x$ – мин.

$\Rightarrow d^2 f \geq 0$

если x – с.в для λ_{max} $\Rightarrow x$ – макс.

$\Rightarrow d^2 f \leq 0$

иначе неопред.

} глобал. т.к отношение Релея ограничен. λ_{min} и λ_{max}

(2) Случай т. стац $x = Qy$

$u = Qv$

$v^t Q^t (\langle y, y\rangle Q\Lambda Q^t - \langle \Lambda y, y\rangle \cdot I) Qv = v^t(\langle y, y\rangle \Lambda - \langle \Lambda y, y\rangle \cdot I)v = \|y\|^2 \cdot v^t\left(\Lambda - \frac{\langle \Lambda y, y\rangle}{\langle y, y\rangle} \cdot I\right)v = \|y\|^2 \cdot v^t(\Lambda - f_\Lambda(y) \cdot I)v =$

$= \|y\|^2 \sum_{j=1}^n v_k^2 (\lambda_j - f_\Lambda(y))$

$\lambda_j = \frac{\langle \Lambda y, y\rangle}{\langle y, y\rangle}$ из условий в градиенте

для $\forall y_j \neq 0$

для других y_j знак может меняться

$\Rightarrow$ это седла, кроме y: $\langle \Lambda y, y\rangle = \lambda_{max}$

$\langle \Lambda y, y\rangle = \lambda_{min}$