Московский Государственный Университет

Матричные вычисления

Выполнил: Курцев Д.В.

Группа: 317

Факультет Вычислительной математики и кибернетики

Кафедра Математических методов прогнозирования

Ноябрь 2021

Задача 1

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

Воспользуемся Теоремой. Если A - квадратная матрица и $AB = I$, то $B = A^{-1}$.

Несложно увидеть, что $(A+UCV) \subset \mathbb{R}^{n\times n}$ Т.е. достаточно доказать, что $(A+UCV)(A^{-1} - A^{-1}U(C^{-1}+VA^{-1}U)^{-1}VA^{-1}) = I$.

Обозначим $M = U(C^{-1}+VA^{-1}U)^{-1}$

$(A+UCV)(A^{-1} - A^{-1}MVA^{-1}) = I - MVA^{-1} + UCVA^{-1} - UCVA^{-1}MVA^{-1}$

Таким образом, достаточно показать, что $UCVA^{-1} = MVA^{-1} + UCVA^{-1}MVA^{-1}$

Но $MVA^{-1} + UCVA^{-1}MVA^{-1} = (I + UCVA^{-1})MVA^{-1}$

То есть, достаточно показать, что $UC = (I + UCVA^{-1})M$

$(I+UCVA^{-1})M = (I+UCVA^{-1})(U(C^{-1}+VA^{-1}U)^{-1}) = (U + UCVA^{-1}U)(C^{-1}+VA^{-1}U)^{-1} = U(I+CVA^{-1}U)(C^{-1}+VA^{-1}U)^{-1} = UC(C^{-1}+VA^{-1}U)(C^{-1}+VA^{-1}U)^{-1} = UC$.

Ч.т.д.

Задача 2

$a)$ $\|uv^T - A\|_F^2 - \|A\|_F^2 = \langle uv^T - A, uv^T - A\rangle - \langle A, A\rangle = \langle uv^T, uv^T\rangle - 2\langle uv^T, A\rangle + \langle A, A\rangle - \langle A, A\rangle = tr((uv^T)^T uv^T) - 2tr(vu^TA) = tr(u^Tuv^Tv) - 2tr(u^TAv) = tr(\|u\|_F^2\, \|v\|_F^2) - 2u^TAv = \|u\|_F^2\, \|v\|_F^2 - 2u^TAv$.

Ответ: $\|u\|_F^2\, \|v\|_F^2 - 2u^TAv$

b)

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

По тождеству Вудбери, доказанному в задаче 1 получим:

$$(2I_n+aa^T)^{-1} = \frac{1}{2}I_n-\frac{1}{2}I_n a(1+a^T\frac{1}{2}I_n a)^{-1}a^T\frac{1}{2}I_n = \frac{1}{2}I_n-\frac{1}{4}a(1+\frac{1}{2}a^T a)^{-1}a^T =$$
$$\frac{1}{2}I_n - \frac{aa^T}{4(1+\frac{1}{2}a^T a)} = \frac{1}{2}I_n - \frac{aa^T}{2(2+\langle a,a\rangle)}$$

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

$$-\ tr(\frac{aa^T uv^T + aa^T vu^T}{2+\|a\|^2})\ ] = \langle u,v\rangle - \frac{1}{2(2+\|a\|^2)}tr(aa^T uv^T + aa^T vu^T) \quad (=)$$

Заметим, что:

$tr(aa^T uv^T) = tr(a^T uv^T a = tr(\langle a,u\rangle\langle a,v\rangle) = \langle a,u\rangle\langle a,v\rangle$

Аналогично получаем, что $tr(aa^T vu^T) = \langle a,u\rangle\langle a,v\rangle$

Тогда:

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

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

c)

$S = \sum\limits_{i=1}^{n} a_i a_i^T \Rightarrow S \subset \mathbb{S}^d \Rightarrow S^{-1} \subset \mathbb{S}^d \Rightarrow S^{-1} = S^{-T}$

$\sum\limits_{i=1}^{n}\langle S^{-1}a_i, a_i\rangle = \sum\limits_{i=1}^{n} tr(a_i^T S^{-1} a_i) = \sum\limits_{i=1}^{n} tr(S^{-1}a_i a_i^T) = tr[\sum\limits_{i=1}^{n}(S^{-1}a_i a_i^T)] =$

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

Ответ: d

Задача 3

a)$f(t) = det(A - tI_n)$

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

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

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

Первое слагаемое мы уже знаем, т.к. в нём стоит $df(t)$:

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

$d(tr(A - tI_n)^{-1}) = tr(d(A - tI_n)^{-1}) = tr(-(A - tI_n)^{-1}\ d(A - tI_n)\ (A - tI_n)^{-1}) = tr(-(A - tI_n)^{-1}\ (-I_n)dt\ (A - tI_n)^{-1}) = tr(A - tI_n)^{-2}\ dt$

Тогда получаем:

$(=)\ det(A - tI_n)\ [tr^2(A - tI_n)^{-1} - tr(A - tI_n)^{-2}]\ dt^2$

Ответ: $f'(t) = -det(A - tI_n)\ \ tr(A - tI_n)^{-1}$

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

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

$A \in \mathbb{S}^n \Rightarrow A + tI_n \in \mathbb{S}^n \Rightarrow (A + tI_n)^{-1} \in \mathbb{S}^n \Rightarrow (A + tI_n)^{-1} = (A + tI_n)^{-T}$

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

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

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

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

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

$$+\frac{b^T(A+tI_n)^{-3}b\ d(\|(A+tI_n)^{-1}b\|)}{\|(A+tI_n)^{-1}b\|^2}dt \quad (=)$$

$d(b^T(A+tI_n)^{-3}b) = b^T\ d(A+tI_n)^{-3}\ b = b^T\ 3(A+tI_n)^{-2}d(A+tI_n)^{-1}\ b =$ $-3b^T\ (A+tI_n)^{-4}\ b\ dt$, где $d(A+tI_n)^{-1}$ был посчитан в пункте а).

Тогда получаем:

$$(=)\ \frac{3b^T\ (A+tI_n)^{-4}\ b\|(A+tI_n)^{-1}b\| - \frac{(b^T(A+tI_n)^{-3}b)^2}{\|(A+tI_n)^{-1}b\|}}{\|(A+tI_n)^{-1}b\|^2}dt^2 =$$

$$\frac{3b^T\ (A+tI_n)^{-4}\ b\|(A+tI_n)^{-1}b\|^2 - (b^T(A+tI_n)^{-3}b)^2}{\|(A+tI_n)^{-1}b\|^3}dt^2$$

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

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

Задача 4

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

$df(x) = \frac{1}{2}\ d\langle xx^T - A, xx^T - A\rangle = \frac{1}{2}\,(\langle d(xx^T - A),\ xx^T - A\rangle + \langle xx^T - A,\ d(xx^T - A)\rangle) = \frac{1}{2}\ 2\ \langle xx^T - A, d(xx^T - A)\rangle = \langle xx^T - A, d(x)\,x^T - x\,(dx)^T)\rangle = \langle (xx^T - A)x, dx\rangle + tr[(xx^T - A)^T x(dx)^T] = \langle (xx^T - A)x, dx\rangle + tr[(dx)^T(xx^T - A)x] = \langle (xx^T - A)x, dx\rangle + \langle (xx^T - A)x, dx\rangle = 2\,\langle (xx^T - A)x, dx\rangle \Rightarrow \nabla f(x) = 2(xx^T - A)x$

$d^2f(x) = d\,2\,\langle (xx^T - A)x, dx_1\rangle = 2\,\langle d[(xx^T - A)x], dx_1\rangle \quad (=)$
$d(xx^T - A)x = d(xx^Tx - Ax) = d(x\langle x, x\rangle) - d(Ax) = \langle x, x\rangle\, dx_2 + x\, d\langle x, x\rangle - Adx_2 = \|x\|^2 I_n\, dx_2 + 2x\,\langle x, dx_2\rangle - Adx_2 = \|x\|^2 I_n\, dx_2 + 2xx^T dx_2 - Adx_2 = (\ \|x\|^2 I_n + 2xx^T - A)dx_2$
Тогда получаем:
$(=)\ \ 2\langle(\|x\|^2 I_n + 2xx^T - A)dx_2, dx_1\rangle = 2\langle(\|x\|^2 I_n + 2xx^T - A)^T dx_1, dx_2\rangle = 2\langle(\ \|x\|^2 I_n + 2xx^T - A)\, dx_1, dx_2\rangle \Rightarrow \nabla^2 f(x) = 2(\|x\|^2 I_n + 2xx^T - A)$

Ответ: $\nabla f(x) = 2(xx^T - A)x$

$\nabla^2 f(x) = 2(\|x\|^2 I_n + 2xx^T - A)$

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

Для начала найдём $dg(x)$, где $g(x) = x^x = e^{x\,lnx}$

$dg(x) = de^{x\,lnx} = e^{x\,lnx}\,d(x\,lnx) = x^x(lnx + 1)dx$

Тогда отсюда следует, что

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

$d^2f(x) = 2\,df(x)\;(ln\|x\|^2+1)\,\langle x, dx_1\rangle + 2f(x)\;d(ln\|x\|^2+1)\;\langle x, dx_1\rangle +$
$2f(x)\;(ln\|x\|^2+1)\;d\langle x, dx_1\rangle \quad (=)$

$$d(ln\langle x, x\rangle + 1) = \frac{2\langle x, dx_2\rangle}{\langle x, x\rangle}$$

$d\langle x, dx_1\rangle = \langle dx_1, dx_2\rangle$ Тогда получим:

$(=) \;\; 2\;[2f(x)(ln\|x\|^2+1)^2\;\langle x, dx_1\rangle\langle x, dx_2\rangle + \frac{2f(x)}{\|x\|^2}\langle x, dx_1\rangle\langle x, dx_2\rangle +$
$2f(x)(ln\|x\|^2+1)\langle dx_1, dx_2\rangle] = \{\langle x, dx_1\rangle\langle x, dx_2\rangle = x^T dx_1\;x^T dx_2 =$
$dx_1^T\,xx^T dx_2 = \langle xx^T\,dx_1, dx_2\rangle\} = 4f(x)(ln\|x\|^2+1)^2\;\langle xx^T\,dx_1, dx_2\rangle +$
$\frac{4f(x)}{\|x\|^2}\langle xx^T\,dx_1, dx_2\rangle + 2f(x)(ln\|x\|^2+1)I_n\langle dx_1, dx_2\rangle \Rightarrow$
$\nabla^2 f(x) = f(x)\;[4(ln\|x\|^2+1)^2\;xx^T + \frac{4}{\|x\|^2}xx^T + 2(ln\|x\|^2+1)I_n]$

Ответ: $\nabla f(x) = 2f(x)\;(ln\|x\|^2+1)x$

$\nabla^2 f(x) = f(x)[4(ln\|x\|^2+1)^2\;xx^T + \frac{4}{\|x\|^2}xx^T + 2(ln\|x\|^2+1)I_n]$

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

$df(x) = d\langle Ax-b, Ax-b\rangle^{\frac{p}{2}} = \frac{p}{2}\langle Ax-b, Ax-b\rangle^{\frac{p}{2}-1}\;d\langle Ax-b, Ax-b\rangle =$
$p\|Ax-b\|^{p-2}\;\langle Ax-b, d(Ax-b)\rangle = p\|Ax-b\|^{p-2}\;\langle Ax-b, Adx\rangle =$
$p\|Ax-b\|^{p-2}\;\langle A^T(Ax-b), dx\rangle \Rightarrow \nabla f(x) = p\|Ax-b\|^{p-2}\;A^T(Ax-b)$

$d^2f(x) = p\;d(\langle Ax-b, Ax-b\rangle^{\frac{p}{2}-1})\;\;\langle A^T(Ax-b), dx_1\rangle + p\|Ax-$
$b\|^{p-2}\;d\langle A^T(Ax-b), dx_1\rangle \quad (=)$
$d(\langle Ax-b, Ax-b\rangle^{\frac{p}{2}-1}) = (\frac{p}{2}-1)\;\langle Ax-b, Ax-b\rangle^{\frac{p}{2}-2}\;2\langle Ax-b, Adx_2\rangle =$
$(p-2)\;\|Ax-b\|^{p-4}\;\langle A^T(Ax-b), dx_2\rangle$
$d\langle A^T(Ax-b), dx_1\rangle = \langle A^TAdx_2, dx_1\rangle = \langle A^TAdx_1, dx_2\rangle$
$(=) \;\; p(p-2)\;\|Ax-b\|^{p-4}\;\langle A^T(Ax-b), dx_2\rangle\;\langle A^T(Ax-b), dx_1\rangle +$
$p\|Ax-b\|^{p-2}\;\langle A^TAdx_1, dx_2\rangle =${Аналогично пункту b) получаем}$=$
$p(p-2)\;\|Ax-b\|^{p-4}\;\langle A^T(Ax-b)(Ax-b)^TAdx1, dx_2\rangle + p\|Ax-$

$b\|^{p-2}\ \langle A^T A dx_1, dx_2\rangle \Rightarrow \nabla^2 f(x) = p(p-2)\ \|Ax-b\|^{p-4} A^T(Ax-b)(Ax-b)^T A + p\|Ax-b\|^{p-2} A^T A$

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

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

Задача 5

a) $f(X) = tr(X^{-1}),\ \ X \in \mathbb{S}^n_{++} \Rightarrow X^{-1} \in \mathbb{S}^n_{++}$

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

$d^2 f(X) = d(-tr(X^{-2}\, dX)) = -tr(dX^{-2}\, dX_1) = -tr(2X^{-1} d(X^{-1})\, dX_1) = -2\, tr[X^{-1}(-X^{-1}) dX_2\, X^{-1} dX_1] = 2\, tr(X^{-2} dX_2\, X^{-1}\, dX_1) =$

$2\ \ tr(X^{-1} dX_2\ \ X^{-1}\ \ dX_1\ \ X^{-1}) = 2\,\langle dX_2\ \ X^{-1},\ X^{-1}\, dX_1\, X^{-1}\rangle = \{X^{-1} \in \mathbb{S}^n_{++} \Rightarrow \exists L : X^{-1} = LL^T\} = 2\,\langle dX\ X^{-1},\ LL^T\, dX\, X^{-1}\rangle =$

$2\,\langle L^T dX\, X^{-1},\ L^T\, dX\, X^{-1}\rangle > 0,\ = 0 \Leftrightarrow X^{-1} \equiv 0$

Ч.т.д.

b) $f(X) = (det\, X)^{\frac{1}{n}},\ \ X \in \mathbb{S}^n_{++} \Rightarrow X^{-1} \in \mathbb{S}^n_{++}$

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

$d^2 f(x) = d(\frac{1}{n}(det\ X)^{\frac{1}{n}}\ \langle X^{-1},\ dX\rangle) = \frac{1}{n}(d(det\ X)^{\frac{1}{n}})\ \langle X^{-1},\ dX\rangle + \frac{1}{n}(det\, X)^{\frac{1}{n}}\, d\langle X^{-1},\ dX\rangle \quad (=)$

$\langle X^{-1},\ dX\rangle = tr(X^{-1} dX) = tr(X^{-\frac{1}{2}} X^{-\frac{1}{2}}\, dX) = tr(X^{-\frac{1}{2}}\, dX\, X^{-\frac{1}{2}}) = \{M = X^{-\frac{1}{2}}\, dX\, X^{-\frac{1}{2}}\} = trM = \langle I_n, M\rangle$

$d\langle X^{-1}, dX\rangle = \langle dX^{-1}, dX\rangle = \langle -X^{-1} dX\ X^{-1}, dX\rangle = tr(-X^{-1} dX\ X^{-1} dX)$

$= tr(-X^{-\frac{1}{2}}\, dX\, X^{-\frac{1}{2}}\, X^{-\frac{1}{2}}\, dX\, X^{-\frac{1}{2}}) = tr(M^T M) = \|M\|^2$

(=) $\frac{1}{n}(detX)^{\frac{1}{n}}[\frac{1}{n}\langle X^{-1}, dX\rangle^2 - \langle -X^{-1}dXX^{-1}, dX\rangle] = \frac{1}{n}(detX)^{\frac{1}{n}}(\frac{1}{n}\langle I_n, M\rangle^2 - \|M\|^2)$

По неравенству Коши-Буняквского получаем:

$\langle I_n, M\rangle^2 \leq \|I_n\|^2 \, \|M\|^2 = n\|M\|^2 \quad \Rightarrow$

$\Rightarrow \ \frac{1}{n}\langle I_n, M\rangle^2 \leq \|M\|^2, \quad X \in \mathbb{S}^n_{++} \Rightarrow det(X) > 0$

Получим, что $d^2 f(X) \leq 0$

Ч.т.д.

Задача 6

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

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

По задаче 4 c) при $A = I_n, b = 0$ получим, что $d(\|x\|^3) = 3\|x\|\langle x, dx\rangle$

Тогда $df(x) = \langle c + \frac{\sigma}{3}3\,\|x\|\,x, dx\rangle = (c^T + \sigma\|x\|\,x^T)dx$

Условие стационарности точки: $\nabla f(x) = 0$. То есть:

$c + \sigma\|x\|\,x = 0 \Rightarrow \|x\|\,x = -\frac{c}{\sigma} \Rightarrow \|x\|^2 = \|\frac{c}{\sigma}\|$. Тогда при условии, что $c \neq 0$ получим:

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

Ответ: $x = -\frac{c}{\sigma\sqrt{\|\frac{c}{\sigma}\|}}$, $c \neq 0$

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

$df(x) = d(\langle a, x\rangle) - d(ln(1 - \langle b, x\rangle)) = \langle a, dx\rangle - \frac{-\langle b,\, dx\rangle}{1-\langle b,\, x\rangle} = \langle a + \frac{b}{1-\langle b,\, x\rangle},\ dx\rangle \quad \Rightarrow \quad \nabla f(x) = a + \frac{b}{1-\langle b,\, x\rangle} = 0 \quad \Rightarrow \quad (1 - \langle b,\ x\rangle)a = -b \quad \Rightarrow a$ и b должны быть коллинеарны. Так же заметим, что $\langle b,\ x\rangle < 1 \quad \Rightarrow \quad 1 - \langle b,\ x\rangle > 0$. Следовательно a и b должны быть противоположно направленными. Пусть $b = \beta a,\ \ \beta \in \mathbb{R}, \beta < 0$.

Заметим, что $\beta = -\frac{\|b\|}{\|a\|}$. Тогда получим:
$(1-\langle \beta a, x\rangle)a = -\beta a \;\;\Rightarrow\;\; \{$Так как это верно для $\forall a\} \;\Rightarrow\; \beta\langle a, x\rangle = 1+\beta \;\;\Rightarrow\;\; a^T x = \frac{1+\beta}{\beta}$. То есть x - решение уравнения $\sum\limits_{i=1}^{n} a_i x_i = 1+\frac{1}{\beta}$
Ответ: $a^T x = 1 - \frac{\|a\|}{\|b\|}$, $\;a$ и b - противоположно направленные векторы.

c) $f(x) = \langle c,\, x\rangle\, e^{-\langle Ax,\, x\rangle}$
$df(x) = \langle c,\, dx\rangle\, e^{-\langle Ax,\, x\rangle} - \langle c,\, x\rangle\, e^{-\langle Ax,\, x\rangle}\, d\langle Ax,\, x\rangle = \langle c,\, dx\rangle\, e^{-\langle Ax,\, x\rangle} - 2\langle c,\, x\rangle\, e^{-\langle Ax,\, x\rangle}\,\langle Ax,\, dx\rangle \Rightarrow \nabla f(x) = e^{-\langle Ax,\, x\rangle}\, c - 2\langle c,\, x\rangle\, e^{-\langle Ax,\, x\rangle}\, Ax = e^{-\langle Ax,\, x\rangle}(c - 2\langle c,\, x\rangle\, Ax) = 0 \;\Rightarrow\; c - 2\langle c,\, x\rangle\, Ax = 0 \;\Rightarrow\; \langle c,\, x\rangle\, Ax = \frac{1}{2}c$.
Из того, что $A \in \mathbb{S}^n_{++} \;\Rightarrow\; detA > 0 \;\Rightarrow\; \exists A^{-1},\; A^{-1} \in \mathbb{S}^n_{++}$.
Тогда получим, что:
$\frac{1}{2}\, A^{-1}c = \langle c,\, x\rangle x = x\langle c,\, x\rangle = xx^T c \;\;\Rightarrow\;\; xx^T c = \frac{1}{2}\, A^{-1}c$.
Заметим так же, что:
$c^T xx^T c = x^T cx^T c = (x^T c)^2 = \frac{1}{2}\, c^T A^{-1} c = a \;\;\Rightarrow\;\; x^T c = \pm\sqrt{a}$. (Так как $A^{-1} \in \mathbb{S}^n_{++}$, то $a > 0$)
Тогда отсюда следует:
$\pm x\sqrt{a} = \frac{1}{2}\, A^{-1}c \;\Rightarrow\; x = \pm\frac{1}{2\sqrt{a}}\, A^{-1}c$

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

Задача 7

$tr(X^{-k} - (X^k + X^{2k})^{-1}) = tr[X^{-k} - ((I_n + X^k)X^k)^{-1}] = tr[X^{-k}(I_n + X^k)^{-1}(I_n + X^k) - X^{-k}\,(I_n + X^k)^{-1}] = tr[X^{-k}(I_n + X^k)^{-1}((I_n +$

$X^k) - I_n)] = tr[X^{-k}(I_n + X^k)^{-1}\, X^k] = tr[X^k\, X^{-k}(I_n + X^k)^{-1}] =$ $tr[(I_n + X^k)^{-1}] \quad (=)$

Так как $X \in \mathbb{S}^n_{++} \quad \Rightarrow \quad \exists V, D$, где V - ортоганальная, а D - диагональная матрицы, такие что $X = VDV^{-1}$, где у матрицы D на диагонали стоят собственные значения матрицы X

$(=)\;\; tr[(V^kV^{-k}+V^kD^kV^{-k})^{-1}] = tr(V^k(I_n+D^k)V^{-k})^{-1} = tr(V^k(I_n+D^k)^{-1}V^{-k}) = tr(V^{-k}V^k(I_n + D^k)^{-1}) = tr(I_n + D^k)^{-1} = \sum\limits_{i=1}^{n} \frac{1}{1+\lambda_i^k}$, где λ_i - i-ое собственное значение X, $\lambda_i > 0,\; i = \overline{1, n}$, т.к. $X \in \mathbb{S}^n_{++}$

Тогда получим 3 разных случая:

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

2) $\lambda > 1 \;\Rightarrow\; \lim\limits_{k\to+\infty} \frac{1}{1+\lambda^k} = 0$

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

Таким образом, получим:

$$\lim_{k\to+\infty} tr(X^{-k} - (X^k + X^{2k})^{-1}) = \sum_{i=1}^{n}(\tfrac{1}{2}\mathbb{1}\{\lambda_i = 1\} + \mathbb{1}\{0 < \lambda_i < 1\})$$

Ответ: $\sum\limits_{i=1}^{n}(\frac{1}{2}\mathbb{1}\{\lambda_i = 1\} + \mathbb{1}\{0 < \lambda_i < 1\})$

Задача 8

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

$S = \frac{1}{N}\sum\limits_{i=1}^{N} x_ix_i^T \quad \Rightarrow \quad S \in \mathbb{S}^D$

а) $dF(P) = Ndtr((I-P(P^TP)^{-1}P^T)^2S) = Ntr(d[(I-P(P^TP)^{-1}P^T)^2]S) = 2N\, tr[(I - P(P^TP)^{-1}P^T)\, d(I - P(P^TP)^{-1}P^T)\, S] \quad (=)$

$d(I - P(P^TP)^{-1}P^T) = -dP\,(P^TP)^{-1}P^T - P(P^TP)^{-1}\,dP^T + P(P^TP)^{-1}(dP^T\,P + P^T\,dP)(P^TP)^{-1}P^T.$

Тогда получим, что:

$(I-P(P^TP)^{-1}P^T)d(I-P(P^TP)^{-1}P^T) = -dP(P^TP)^{-1}P^T - P(P^TP)^{-1}dP^T$
$+ P(P^TP)^{-1}\, dP^T\ P(P^TP)^{-1}P^T + P(P^TP)^{-1}P^T\ dP\ (P^TP)^{-1}P^T +$
$P(P^TP)^{-1}P^T\, dP\, (P^TP)^{-1}P^T + P(P^TP)^{-1}P^TP(P^TP)^{-1}\, dP^T$
$-\, P(P^TP)^{-1}P^TP(P^TP)^{-1}\, dP^T\ P(P^TP)^{-1}P^T -$
$P(P^TP)^{-1}P^TP(P^TP)^{-1}P^T\, dP\, (P^TP)^{-1}P^T =$
$P(P^TP)^{-1}P^T\, dP\, (P^TP)^{-1}P^T - dP\, (P^TP)^{-1}P^T \quad \Rightarrow$
$(=)\ \ 2N\, tr[P(P^TP)^{-1}P^T\, dP\, (P^TP)^{-1}P^TS - dP\, (P^TP)^{-1}P^TS] =$
$2N\, tr[(P^TP)^{-1}P^TSP(P^TP)^{-1}P^T\, dP - (P^TP)^{-1}P^TSdP] \quad \Rightarrow$
$\nabla F(P) = 2N[(P^TP)^{-1}P^TSP(P^TP)^{-1}P^T - (P^TP)^{-1}P^TS]^T = \{P^TP \in \mathbb{S}^d \Rightarrow (P^TP)^{-1} \in \mathbb{S}^d\} = 2N(P(P^TP)^{-1}P^TSP(P^TP)^{-1} - SP(P^TP)^{-1})$
Далее учтём, что $P^TP = I \ \Rightarrow\ \nabla F(P) = 2N\,(PP^TSP - SP)$

Ответ: $\nabla F(P) = 2N\,(PP^TSP - SP)$

b)

Пусть $S = Q\Lambda Q^T$, где Λ - диагональная матрица с собственными значениями, а $Q = [q_1|...|q_D]$ - матрица, состоящая из собственных векторов. Пусть матрица $P = [p_1|...|p_d]$, состоит из любых различных d столбцов матрицы Q. Тогда:
$SP = Q\Lambda Q^TP$. Обозначим $M = Q^TP \in \mathbb{R}^{D\times d}$. Произведя умножение, получим, что

$$M_{ij} = \begin{cases} 1 & q_i = p_j \\ 0 & \text{иначе} \end{cases}$$

Заметим, что в матрице M ровно d единиц, причём все они располагаются в разных строках и столбцах. Умножив Λ на M получим матрицу $B = \Lambda M$, у которой вместо единиц матрицы M стоят собственные значения, соответствующие собственным векторам p_j. Тогда $P' = QB$ - это матрица состоящая из собственных векторов p_j, умноженных на их собственные значения, т.е. $P' = [\lambda_1 p_1|...|\lambda_d p_d]$ $PP^TSP = PP^TQ\Lambda Q^TP = PM^T\Lambda M = PM^TB$. Умножив матрицы M^T и B, получим диагональную матрицу $D = M^TB$, у которой на главной диагонали стоят собственые значения, соответствующие собственным векторам p_j, идущие в том же порядке, что и p_j в P. Тогда получим, что $PD = P'$.

Слеловательно $\nabla F(P) = PP^TSP - SP = P' - P' = 0$

Покажем теперь, что минимум $F(P)$ достигается, если P состоит из столбцов, отвечающих наибольшим собственным значениям.

Обозначим $A = I - PP^T$, учитывая $P^TP = I, (P^TP)^{-1} = I$. Тогда $F(P) = N\,tr(A^2S) = N\,tr((I - PP^T)AS) = N\,tr(AS(I - PP^T)) = N\,tr(AS - ASPP^T) = N\,[tr(AS) - tr(ASPP^T)]$.

Но, ранее было показано, что $0 = \nabla F(P) = PP^TSP - SP = (PP^T - I)SP = -ASP \;\; \Rightarrow \;\; tr(ASPP^T) = 0$

Отсюда получим, что:

$F(P) = N\,trAS = tr((I - PP^T)S) = trS - trPP^TS \quad \Rightarrow$ минимум функции $F(P)$ достигается при максимуме $trPP^TS = trPP^TQ\Lambda Q^T = trPM^T\Lambda Q^T = trPB^TQ^T = trPP'^T = trP'^TP$. Но матрица P' стоит из тех же самых собственных векторов, что и P, только умноженных на их собственные значения. Поэтому произведение $P'^TP \in \mathbb{R}^{d\times d}$ будет диагональной матрицей, состоящей из собственных значений λ_i векторов p_j.

Тогда получим, что $trPP^TS = \sum\limits_{i=1}^{d} \lambda_i$. Отсюда следует, что минимум $F(P)$ достигается для матрицы P, состоящей из собственных векторов q_i, отвечающих наибольшим собственным значениям матрицы S.

Ч.т.д.