Домашнее задание № 1. *Алеся Демешко, 797*

Задача 1. *Покажите, что для любых квадратных вещественных матриц выполнено $Tr(AB) = Tr(BA)$.*

Решение. Пусть $A = (a_{ij}), B = (b_{ij})$. Тогда $(AB)_{ii} = \sum_{k=1}^{n} a_{ik}b_{ki} \Rightarrow Tr(AB) = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}b_{kj}$. С другой стороны,

$$(BA)_{ii} = \sum_{l=1}^{n} a_{il}b_{li} \Rightarrow Tr(BA) = \sum_{m=1}^{n}\sum_{l=1}^{n} b_{ml}a_{lm}$$
$$= \sum_{l=1}^{n}\sum_{m=1}^{n} a_{lm}b_{ml} = \sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}b_{kj} \Longrightarrow Tr(AB) = Tr(BA)$$

□

Задача 2. *Доказать, что для любых матриц $A \in R^{m\times k}, B \in R^{k\times n}, C \in R^{m\times n}$ имеет место $\langle AB, C\rangle = \langle B, A^TC\rangle = \langle A, CB^T\rangle$*

Решение. По определению $\langle A, B\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ij}$.
$AB = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{l=1}^{k} a_{i,l}b_{l,j} \to \langle AB, C\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n}(AB)_{i,j}c_{i,j} = \sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{l=1}^{k} a_{i,l}b_{l,j}c_{i,j} = \sum_{l=1}^{k}\sum_{j=1}^{n} b_{l,j}\sum_{i=1}^{m} a_{i,l}c_{i,j} = \sum_{l=1}^{k}\sum_{j=1}^{n} b_{l,j}(A^T)_{l,i}c_{i,j} = \sum_{l=1}^{k}\sum_{j=1}^{n} b_{l,j}(A^TC)_{l,j} = \langle B, A^TC\rangle$
$\langle AB, C\rangle = \sum_{l=1}^{k}\sum_{j=1}^{n}\sum_{i=1}^{m} a_{i,l}b_{l,j}c_{i,j} = \sum_{i=1}^{m}\sum_{l=1}^{k} a_{i,l}\sum_{j=1}^{n} c_{i,j}b_{l,j} =$
$= \sum_{i=1}^{m}\sum_{l=1}^{k} a_{i,l}\sum_{j=1}^{n} c_{i,j}B^T_{j,l} = \sum_{i=1}^{m}\sum_{l=1}^{k} a_{i,l}(CB^T)_{i,l}$ □

Задача 3. *Пусть $x, y \in R^n$. Покажите, что $\langle xx^T, yy^T\rangle = \langle x, y\rangle^2$*

Доказательство. По определению $\langle x, y\rangle = \sum_{i=1}^{n} x_iy_i$, $\langle A, B\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} * b_{ij}$.
С одной стороны, что такое a_{ij}: $a_{ij} = (xx^T)_{ij} = x_ix_j$. Аналогично, $b_{ij} = (yy^T)_{ij} = y_iy_j$. То есть мы получаем, $\langle xx^T, yy^T\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} * b_{ij} = \sum_{i=1,j=1}^{n} x_ix_jy_iy_j$ (1).
С другой стороны, $\langle x, y\rangle^2 = (\sum_{i=1}^{n} x_iy_i)^2 = \sum_{i=1}^{n} x_iy_i\sum_{j=1}^{n} x_jy_j = \sum_{i=1,j=1}^{n} x_ix_jy_iy_j$ (2)
Получим в итоге (1) = (2). □

Задача 4. *Доказать неравенство Коши-Буняковского $|\langle x, y\rangle| \le \langle x, x\rangle^{1/2}\langle y, y\rangle^{1/2}$*

Доказательство. Рассмотрим скалярное произведение $\langle tx - y, tx - y\rangle = 0$.
$0 = \langle tx - y, tx - y\rangle = t^2\langle x, x\rangle - 2t\langle x, y\rangle + \langle y, y\rangle$ Так как у нас одно либо нет решений, то $D = 4\langle x, y\rangle^2 - 4\langle x, x\rangle\langle y, y\rangle \le 0 \Rightarrow \langle x, y\rangle^2 \le \langle x, x\rangle\langle y, y\rangle \Rightarrow \langle x, y\rangle \le \langle x, x\rangle^{1/2}\langle y, y\rangle^{1/2}$ □

Задача 5. *Пусть $x \in R^n$. Докажите следующие неравенства: $\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\|x\|_\infty$ $\frac{1}{\sqrt{n}}\|x\|_1 \le \|x\|_2 \le \|x\|_1$.*

Доказательство. 1) $\|x\|_\infty = max_{1\le i\le n}|x| \le |x_{max}| + (\sum_{i=1,i!=i_{max}}^{n} x_i^2)^{1/2} = ((x_{max})^2)^{1/2}2 + (\sum_{i=1,i!=i_{max}}^{n} x_i^2)^{1/2} = (\sum_{i=1}^{n} x_i^2)^{1/2} = \|x\|_2$
$\|x\|_2 = (\sum_{i=1}^{n} x_i^2)^{1/2} \le (\sum_{i=1}^{n} x_{max}^2)^{1/2} = (nx_{max}^2)^{1/2} = \sqrt{n}\|x\|_\infty$
2) Докажем, что $\frac{1}{\sqrt{n}}(|x_1| + ... + |x_n|) \le \sqrt{x_1^2 + ... + x_n^2}$. Возведём в квадрат: $\frac{1}{n}(|x_1| + ... + |x_n|)^2 \le x_1^2 + ... + x_n^2 \Leftrightarrow (|x_1| + ... + |x_n|)^2 \le n(x_1^2 + ... + x_n^2) \Leftrightarrow (|x_1| + ... + |x_n|)^2 \le (1 + ... + 1)(x_1^2 + ... + x_n^2)$, где $(1 + ... + 1) = n$, неравенство верно так как это неравенство Коши-Буняковского.
$\sqrt{x_1^2 + ... + x_n^2} \le |x_1| + ... + |x_n|$. Возведём в квадрат: $x_1^2 + ... + x_n^2 \le (|x_1| + ... + |x_n|)^2$, что очевидно верно. □

Задача 6. *Упростите каждое из следующих выражений:*

(a) $\det(AXB(C^{-T}X^TC)^{-T})$*, где* $A, B, C, D \in R^{n\times n}, \det(C) \neq 0, \det(C^{-T}X^TC) \neq 0$*;*

(b) $\|uv^T - A\|_F^2 - \|A\|_F^2$*, где* $u \in R^m, v \in R^n, A \in R^{m\times n}$*;*

(c) $Tr((2I_n + aa^T)^{-1}(uv^T + vu^T))$*, где* $a, u, v \in R^n$*;*

(d) $\sum_{i=1}^n \langle S^{-1}a_i, a_i\rangle$*, где* $a_1, ..., a_n \in R^n, S := \sum_{i=1}^n a_i a_i^T, \det(S) \neq 0$*;*

Доказательство. (a) Из линейной алгебры $(AB)^T = B^TA^T$ и $(AB)^{-1} = B^{-1}A^{-1}$. Тогда $(C^{-T}X^TC)^{-T} = (C^{-T}X^{-T}C)^T = CX^{-1}C^{-T}$. В итоге получим

$$\begin{aligned}
&\det(AXBCX^{-1}C^{-1}) = \det(A)\det(X)\det(B)\det(C)\det(X^{-1})\det(C^{-T}) = \\
&= \det(A)\det(B)\det(X)\det(X^{-1})\det(C)\det(C^{-T}) = \\
&= \det(A)\det(B)\det(C)\det(C^{-T})
\end{aligned}$$

(b) $\|uv^T - A\|_F^2 - \|A\|_F^2 = \sum_{i=1}^m\sum_{j=1}^n(u_iv_j - A_{ij})^2 - = \sum_{i=1}^m\sum_{j=1}^n A_{ij}^2 = \sum_{i=1}^m\sum_{j=1}^n(u_iv_j - 2A_{ij})(u_iv_j) = \sum_{i=1}^m\sum_{j=1}^n u_i^2v_j^2 - \sum_{i=1}^m\sum_{j=1}^n 2u_iA_{ij}v_j = \langle u, u\rangle\langle v, v\rangle - 2\langle vu^T, A\rangle$

(c)

(d)

□

Задача 7. *Пусть f это одна из следующих функций. Нужно вычислить первую и вторую производные f.*

(a) $f : E \to R$ *- функция* $f(t) := \det(A - tI_n)$*, где* $A \in R^{n\times n}, E := \{t \in R : \det(A - tI_n \neq 0\}$

(b) $f : R_{++} \to R$ *- функция* $f(t) := |(A + tI_n)^{-1}b|^2$*, где* $A \in S_+^n, B \in R^n$

Доказательство. (a) Первая производная:

$$\begin{aligned}
&d(\det(A - tI_n)) = \det(A - tI_n)\langle(A - tI_n)^{-T}, d(A - tI_n)\rangle = \\
&= \det(A - tI_n)\langle(A - tI_n)^{-T}, -I_ndt_1\rangle = -\det(A - tI_n)\langle(A - tI_n)^{-T}, I_ndt_1\rangle
\end{aligned}$$

Вторая производная:

$$\begin{aligned}
&d^2(\det(A - tI_n)) = d(-\det(A - tI_n)\langle(A - tI_n)^{-T}, I_ndt_1\rangle) = \\
&= -d(\det(A - tI_n))\langle(A - tI_n)^{-T}, I_ndt_1\rangle - \det(A - tI_n)d(\langle(A - tI_n)^{-T}, I_ndt_1\rangle) = \\
&= -\det(A - tI_n)\langle(A - tI_n)^{-T}, I_ndt_2\rangle\langle(A - tI_n)^{-T}, I_ndt_1\rangle - \det(A - tI_n)\langle d(A - tI_n)^{-T}, I_ndt_1\rangle = \\
&= -\det(A - tI_n)\langle(A - tI_n)^{-T}, I_ndt_2\rangle\langle(A - tI_n)^{-T}, I_ndt_1\rangle - \det(A - tI_n)\langle(d(A - tI_n)^{-1})^T, I_ndt_1\rangle = \\
&= -\det(A - tI_n)\langle(A - tI_n)^{-T}, I_ndt_2\rangle\langle(A - tI_n)^{-T}, I_ndt_1\rangle - \det(A - tI_n)* \\
&* \langle(-(A - tI_n)^{-1})(-I_ndt_2)(A - tI_n)^{-1})^T, I_ndt_1\rangle = -\det(A - tI_n)\langle(A - tI_n)^{-T}, I_ndt_2\rangle* \\
&* \langle(A - tI_n)^{-T}, I_ndt_1\rangle - \det(A - tI_n)\langle(A - tI_n)^{-1}(I_ndt_2)(A - tI_n)^{-1})^T, I_ndt_1\rangle
\end{aligned}$$

(b) Первая производная:

$$\begin{aligned}
&d(|(A + tI_n)^{-1}b|^2) = 2|(A + tI_n)^{-1}b|d(|(A + tI_n)^{-1}b|) = 2|(A + tI_n)^{-1}b|d((A + tI_n)^{-1}b) = \\
&= 2(-(A + tI_n)^{-1}d(A + tI_n)(A + tI_n)^{-1}b)(A + tI_n)^{-1}b = 2(-(A + tI_n)^{-1}I_ndt_1(A + tI_n)^{-1}b)* \\
&* (A + tI_n)^{-1}b = 2((A + tI_n)^{-1})^3bdt_1
\end{aligned}$$

Вторая производная:

$$d(2((A+tI_n)^{-1})^3 bdt_1) = 2*3((A+tI_n)^{-1})^2 bdt_1 d(((A+tI_n)^{-1}) =$$
$$= 6((A+tI_n)^{-1})^2 bdt_1(-A+tI_n)^{-1} d(A+tI_n)(A+tI_n)^{-1}) = 6((A+tI_n)^{-1})^4 bdt_1 dt_2$$

□

Задача 8. *Пусть f одна из следующих функций. Для каждой вычислите градиент и матрицу Гессе.*
По определению, матрица Гессе:
$$\begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ . & . & . \\ \frac{\partial^2 f}{\partial x_1 x_n} & \dots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

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

(b) *$f : R^n \backslash \{0\} \to R$ - функция $f(x) := \frac{\langle Ax,x \rangle}{|x|^2}$, где $A \in S^n$*

(c) *$f : R^n \backslash \{0\} \to R$ - функция $f(x) := \langle x,x \rangle^{\langle x,x \rangle}$*

Доказательство. (a) $f = \frac{1}{2}\|xx^T - A\|_F^2 = \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n x_i x_j - a_{i,j}^2$;
$\nabla f = \frac{1}{2}(\sum_{i\neq 1}^n 2x_1 + 2x_i, ..., \sum_{i\neq n}^n 2x_n + 2x_i) = (\sum_{i\neq 1}^n x_1 + x_i, ..., \sum_{i\neq n}^n x_n + x_i)$
Найдём теперь $\nabla^2 f$. Каждая строка матрицы будет иметь вид: $(1, ..., 1)$ т. к. в каждой компоненте градиента встречается один раз x_i, и при взятии производной он становится 1, а все остальные x_j обнуляются. В итоге получим единичную матрицу I_n

(b)

$$f = \frac{< Ax, x >}{|x|^2} = \frac{\sum_{i=1}^n (Ax)_i x_i}{\sum_{i=1}^n x_i^2} = \frac{\sum_{i=1}^n x_i \sum_{j=1}^n a_{i,j} x_j}{\sum_{i=1}^n x_i^2}$$

$$\frac{\partial f}{\partial x_k} = \frac{(\sum_{i=1}^n \sum_{j=1}^n a_{i,j} x_i x_j)' * (\sum_{i=1}^n x_i^2) + (\sum_{i=1}^n \sum_{j=1}^n a_{i,j} x_i x_j) * (\sum_{i=1}^n x_i^2)'}{(\sum_{i=1}^n x_i^2)^2} =$$
$$= \frac{(\sum_{i\neq k}^n 2x_k a_{k,k} + x_i a_{k,i} + x_i a_{i,k}) * (\sum_{i=1}^n x_i^2) + (\sum_{i=1}^n \sum_{j=1}^n a_{i,j} x_i x_j) * (2x_k)}{(\sum_{i=1}^n x_i^2)^2}$$

$$\frac{\partial^2 f}{\partial x_k x_t} = \frac{(\sum_{i\neq k}^n 2x_k a_{k,k} + x_i a_{k,i} + x_i a_{i,k})' * (\sum_{i=1}^n x_i^2) + (\sum_{i\neq k}^n 2x_k a_{k,k} + x_i a_{k,i} + x_i a_{i,k}) * (\sum_{i=1}^n x_i^2)'}{(\sum_{i=1}^n x_i^2)^4} +$$
$$+ \frac{(\sum_{i=1}^n \sum_{j=1}^n a_{i,j} x_i x_j)' * (2x_k) + (\sum_{i=1}^n \sum_{j=1}^n a_{i,j} x_i x_j) * (2x_k)'}{(\sum_{i=1}^n x_i^2)^4} = \frac{(a_{k,t} + a_{t,k}) * (\sum_{i=1}^n x_i^2)}{(\sum_{i=1}^n x_i^2)^4} +$$

$$+ \frac{(\sum_{i\neq k}^n 2x_k a_{k,k} + x_i a_{k,i} + x_i a_{i,k}) * (2x_t) + (\sum_{i\neq t}^n 2x_t a_{t,t} + x_i a_{t,i} + x_i a_{i,t}) * (2x_k)}{(\sum_{i=1}^n x_i^2)^4}$$

(c) $\langle x,x\rangle^{\langle x,x\rangle} = \exp^{\langle x,x\rangle \ln \langle x,x\rangle}$

$$\frac{\partial f}{\partial x_k} = \exp^{\langle x,x\rangle \ln \langle x,x\rangle}(2x_k \ln\langle x,x\rangle + \langle x,x\rangle \frac{1}{\langle x,x\rangle} 2x_k) = \exp^{\langle x,x\rangle \ln \langle x,x\rangle}(2x_k \ln\langle x,x\rangle + 2x_k)$$

$$\frac{\partial^2 f}{\partial x_k \partial x_t} = \exp^{\langle x,x\rangle \ln \langle x,x\rangle}(2x_t \ln\langle x,x\rangle + 2x_t)(2x_k \ln\langle x,x\rangle + 2x_k) + \exp^{\langle x,x\rangle \ln \langle x,x\rangle}\frac{2x_k 2x_t}{\langle x,x\rangle}$$

□

Задача 9. *Пусть $f : S^n_{++} \to R$ - одна из следующих функций:*

(a) $f(X) := Tr(X^{-1})$

(b) $f(X) := \langle X^{-1}v, v\rangle$, *где* $v \in R^n$

(c) $f(X) := (\det(X))^{1/n}$

Для каждого из указанных вариантов покажите, что вторая производная имеет постоянный знак

Доказательство. (a)

(b) Найдём первую производную: $d(\langle X^{-1}v, v\rangle) = \langle dX^{-1}v, v\rangle = \langle -X^{-1}(dX)X^{-1}v, v\rangle$. Аналогично, вторая производная будет представлять из себя скалярное произведение, которое всегда положительно.

(c) Найдём первую производную:

$$d((\det(X))^{1/n}) = \frac{1}{n}((\det(X)^{-\frac{n-1}{n}})d(\det(X)) = \frac{1}{n}((\det(X)^{-\frac{n-1}{n}})\det(X)\langle X^{-T}, dx_1\rangle =$$
$$= \frac{1}{n}((\det(X)^{\frac{1}{n}})\langle X^{-T}, dx_1\rangle$$

Найдём вторую производную:

$$d(\frac{1}{n}((\det(X)^{\frac{1}{n}})\langle X^{-T}, dx_1\rangle) = \frac{1}{n}(d((\det(X)^{\frac{1}{n}}))\langle X^{-T}, dx_1\rangle + (\det(X)^{\frac{1}{n}})d(\langle X^{-T}, dx_1\rangle)) =$$
$$= \frac{1}{n}(\frac{1}{n}((\det(X)^{\frac{1}{n}})\langle X^{-T}, dx_2\rangle\langle X^{-T}, dx_1\rangle + (\det(X)^{\frac{1}{n}})(\langle dX^{-T}, dx_1\rangle + \langle X^{-T}, ddx_1\rangle))$$

Заметим, что знак второй производной зависит только от знака детерминанта, так как все скалярные произведения положительны. Так как у нас матрица положительно определённая, то следовательно у неё все главные миноры положительны, а значит и самый большой главный минор то есть детерминант. В итоге, получили, что вторая производная всегда положительна.

□