Машинное обучение, ВМК МГУ, осень 2020

Теоретическое задание 1: матричные вычисления и матричное дифференцирование

Мягкий дедлайн: 30 ноября 2020, 01:00 (за каждый день просрочки снимается 1 балл)
Жёсткий дедлайн: 7 декабря 2020, 01:00
Формат сдачи: pdf-файл, подготовленный с помощью TEX, или скан рукописных листков (скан должен быть собран в один pdf-файл с порядком страниц, соответствующим задачам в задании).

Используемые обозначения:

- $\langle x, y \rangle$ — евклидово скалярное произведение;
- $\|x\| = \langle x, x \rangle^{1/2} = (x^T x)^{1/2}$ — евклидова норма вектора;
- $\|A\|_F = \langle A, A \rangle^{1/2} = \operatorname{tr}(A^T A)^{1/2}$ — матричная норма Фробениуса;
- $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n \mid x_i \geq 0 \ \forall i\}$, $\mathbb{R}^n_{++} = \{x \in \mathbb{R}^n \mid x_i > 0 \ \forall i\}$;
- I_n — единичная матрица размера $n \times n$;
- $\mathbb{S}^n = \{A \in \mathbb{R}^{n\times n} \mid A = A^T\}$;
- $\mathbb{S}^n_+ = \{A \in \mathbb{S}^n \mid A$ — неотр. определённая$\}$, $\mathbb{S}^n_{++} = \{A \in \mathbb{S}^n \mid A$ — полож. определённая$\}$.

Все задачи (каждый подпункт) оцениваются одинаково из общей суммы в 10 баллов.

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

2. Упростите каждое из следующих выражений:
   (a) $\|uv^T - A\|_F^2 - \|A\|_F^2$, где $u \in \mathbb{R}^m$, $v \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$.
   (b) $\operatorname{tr}((2I_n + aa^T)^{-1}(uv^T + vu^T))$, где $a, u, v \in \mathbb{R}^n$. *Подсказка: воспользуйтесь тождеством Вудбери.*
   (c) $\sum_{i=1}^n \langle S^{-1}a_i, a_i \rangle$, где $a_1, \dots, a_n \in \mathbb{R}^d$, $S = \sum_{i=1}^n a_i a_i^T$, $\det(S) \neq 0$.

3. Для каждой из следующих функций f найдите первую и вторую производную:
   (a) $f : E \to \mathbb{R}$, $f(t) = \det(A - tI_n)$, где $A \in \mathbb{R}^{n\times n}$, $E = \{t \in \mathbb{R} : \det(A - tI) \neq 0\}$.
   (b) $f : \mathbb{R}_{++} \to \mathbb{R}$, $f(t) = \|(A + tI_n)^{-1}b\|$, где $A \in \mathbb{S}^n_+$, $b \in \mathbb{R}^n$.

4. Для каждой из следующих функций f найдите градиент ∇f и гессиан $\nabla^2 f$:
   (a) $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = \frac{1}{2}\|xx^T - A\|_F^2$, где $A \in \mathbb{S}^n$;
   (b) $f : \mathbb{R}^n \backslash \{0\} \to \mathbb{R}$, $f(x) = \langle x, x \rangle^{\langle x, x \rangle}$;
   (c) $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = \|Ax - b\|^p$, где $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $p \geq 2$.

5. Пусть $f : \mathbb{S}^n_{++} \to \mathbb{R}$ — одна из следующих функций:
   (a) $f(X) = \operatorname{tr}(X^{-1})$;
   (b) $f(X) = (\det X)^{1/n}$.

Для каждого из указанных вариантов покажите, что вторая производная $d^2 f(X)[H, H]$ имеет постоянный знак для всех $X \in \mathbb{S}^n_{++}$ и приращений $H \in \mathbb{S}^n$. *Подсказка: в последнем случае воспользуйтесь неравенством Коши-Буняковского.*

6. Для каждой из следующих функций f найдите все точки стационарности и укажите, когда они существуют:

   (a) $f : \mathbb{R}^n \to \mathbb{R}$, $f(x) = \langle c, x \rangle + \frac{\sigma}{3}\|x\|^3$, где $c \in \mathbb{R}^n$, $c \neq 0$, $\sigma > 0$;

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

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

Бонус (2 балла) Пусть $X \in \mathbb{S}^n_{++}$. Вычислите следующее выражение:

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