Методы глубокого обучения, осень 2023

Домашнее задание. Матричное дифференцирование.

Срок сдачи (жёсткий, без возможности просрочки): 20 сентября, 23:59

Формат сдачи: pdf-файл, подготовленный с помощью $\TeX$.

Обозначения:

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

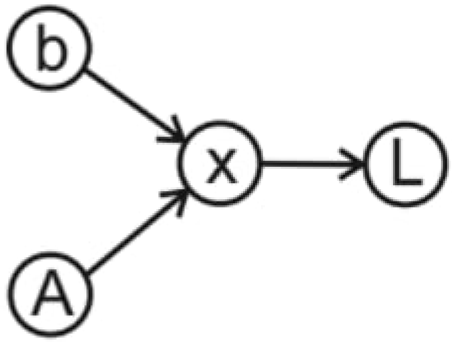
1. Найти градиент $\nabla_X f$ для функции $f(X) = \det(X^{-1} + A)$, где $X \in \mathbb{R}^{n\times n}$ – невырожденная матрица, $A \in \mathbb{R}^{n\times n}$ – произвольная матрица такая, что матрица $X^{-1} + A$ является невырожденной.

2. Для каждой из следующих функций найти градиент ∇f и гессиан $\nabla^2 f$:

   (a) $f : \mathbb{R}_{++} \to \mathbb{R}$, $f(t) = \|(A + tI_n)^{-1} b\|$, где $A \in \mathbb{S}^n_+$, $b \in \mathbb{R}^n$.

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

3. Имеется следующий граф вычислений:



   Здесь $A \in \mathbb{R}^{n\times n}$ — невырожденная матрица, $b \in \mathbb{R}^n$, а x – единственное решение СЛАУ $Ax = b$. L — функция потерь, которая каким-то образом зависит от x. Задача состоит в вычислении градиентов для прохода назад по этому графу, т.е. при известном $\nabla_x L$ найти $\nabla_b L$ и $\nabla_A L$.

4. Для каждой из следующих функций найти все точки стационарности и указать значения параметров, при которых они существуют:

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

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

   (c) $f : \mathbb{S}^n_{++} \to \mathbb{R}$, $f(X) = \langle X^{-1}, I_n\rangle - \langle A, X\rangle$, где $A \in \mathbb{S}^n$.

5. Рассмотрим симметричную матрицу $X \in \mathbb{R}^{n\times n}$ и её собственное разложение $X = Q^T \Lambda Q$, где Q – ортогональная матрица, состоящая из собственных векторов, а Λ – диагональная матрица с собственными значениями на диагонали. Дополнительно известно, что все собственные значения являются различными. Пусть f – некоторая скалярная функция потерь, которая каким-то образом зависит от собственных векторов и собственных значений матрицы X. Требуется по известным $\nabla_\Lambda f$ (диагональная матрица) и $\nabla_Q f$ найти $\nabla_X f$.