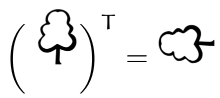
Yet Another Math for DS Course
Домашка №2

Матричное дифФфФфФференцирование (часть 2)
(продвинутая группа)



«Джек и бобовый стебель» (1890)

За эту домашку можно набрать 12 баллов. Два дополнительных балла бонусные и плюсуются к общему результату. Стоимость каждой задачи указана в скобочках. Баллы между пунктами внутри задачи распределяются равномерно, если около них не указано иного.

Решение работы нужно сдать в виде pdf-файла. Решения должны быть оформлены на листочке аккуратным почерком либо затеханы на компьютере. Если у вас плохой почерк, домашка должна быть затехана. Затехать домашку можно в overleaf, typora, colab, notion или другом любом удобном для вас сервисе.

Задача 1 (2 балла). Пусть $f(X) = \mathrm{tr}(A\,X\,B\,X^\top C)$, где $A, C \in \mathbb{R}^{p\times p}$, $B \in \mathbb{R}^{n\times n}$, $X \in \mathbb{R}^{p\times n}$. Найдите $\nabla_X f(X)$.

Задача 2 (2 балла). Пусть $f(X) = \mathrm{tr}\big(AXBX^{-1}\big)$, где $A, B, X \in \mathbb{R}^{n\times n}$. Найдите $\nabla_X f(X)$.

Задача 3 (2 балла). Пусть $f(X) = \det(X^T A X)$, где $A, X \in \mathbb{R}^{n\times n}$. Найдите $\nabla_X f(X)$.

Задача 4 (2 балла). Пусть $x \in \mathbb{R}^n \setminus \{0\}$ и $b \in \mathbb{R}^n$ фиксирован. Найдите градиент $\nabla_x f(x)$ функции

$$f(x) = \left(x^\top x\right)^{b^\top x}.$$

Задача 5 (2 балла). Предположим, что мы решаем задачу линейной регрессии, но при этом хотим, чтобы сумма всех обученных весов была равна единице. Пусть $X \in \mathbb{R}^{n\times d}$ полного ранга, $y \in \mathbb{R}^n$, а $\mathbb{1} \in \mathbb{R}^n$ – вектор из единиц. Найдите

$$\min_{w\in\mathbb{R}^d} (y - Xw)^T (y - Xw) \qquad \text{при условии} \qquad \mathbb{1}^\top w = 1.$$

Задача 6 (2 бонусных балла). Рассмотрим целевую функцию логистической регрессии

$$Q(w) = \frac{1}{n}\sum_{i=1}^{n} \log\left(1 + \exp(-y_i \langle w, x_i \rangle)\right), \quad y_i \in \{-1, +1\},\ x_i \in \mathbb{R}^d.$$

а) Найдите градиент ∇Q_w и упростите итоговое выражение таким образом, чтобы в нём участвовала сигмоидная функция

$$\sigma(z) = \frac{1}{1 + \exp(-z)}.$$

При решении данной задачи вам может понадобиться следующий факт (убедитесь, что он действительно выполняется):

$$\sigma'(z) = \sigma(z)(1 - \sigma(z)).$$

б) Выпишите, как будет выглядеть шаг градиентного спуска.

в) Найдите вторую производную целевой функции по w.

г) Разложите $Q(w)$ в ряд Тейлора до второго порядка в окрестности $w = 0$ и покажите, с какой квадратичной задачей совпадает минимизация этой аппроксимации.