

# Машинное обучение, ФКН ВШЭ

## Теоретическое домашнее задание №2

### Матрично-векторное дифференцирование и градиентный спуск

**Задача 1.** Пусть  $f(x) = \ln \langle Ax, x \rangle$  где  $A \in \mathbb{R}^{n \times n}$  — положительно определенная матрица. Найдите производную  $\nabla_x f(x)$ .

**Задача 2.** Пусть  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$ ,  $x \in \mathbb{R}^n$ . Найдите производную  $\nabla_x f(x)$  функции

$$f(x) = \sin \|Ax + b\|_2$$

**Задача 3.** Рассмотрим следующий важный функционал, который вы встретите уже достаточно скоро:

$$Q(w) = \frac{1}{\ell} \sum_{i=1}^{\ell} \ln(1 + e^{-y_i \langle w, x_i \rangle})$$

Здесь  $y_i \in \{+1, -1\}$ , а  $w, x \in \mathbb{R}^d$ .

Найдите производную  $\nabla_w Q(w)$ . Выпишите, как выглядела бы итерация градиентного спуска в данном случае.

**Задача 4.** В случае одномерной Ridge-регрессии минимизируется функция со штрафом:

$$Q(w) = \|y - Xw\|^2 + \lambda w^2,$$

где  $\lambda$  — положительный параметр, штрафующий функцию за слишком большие значения  $w$ .

1. Найдите производную  $\nabla_w Q(w)$ , выведите формулу для оптимального  $w$ .
2. Найдите вторую производную  $\nabla_w^2 Q(w)$ . Убедитесь, что мы оказались в точке минимума.
3. Выпишите шаг градиентного спуска в матричном виде.
4. Сделайте то же самое, но для многомерного случая:

$$Q(w) = \frac{1}{\ell} \|y - Xw\|^2 + \lambda \|w\|^2 = \frac{1}{\ell} \sum_{i=1}^{\ell} (y_i - \langle x_i, w \rangle)^2 + \lambda \|w\|^2$$

где  $X \in \mathbb{R}^{\ell \times d}$ ,  $w \in \mathbb{R}^d$ ,  $x, y \in \mathbb{R}^{\ell}$ .

**Задача 5.** Пусть  $f(x) = x^T \exp(xx^T)x$ , где  $x \in \mathbb{R}^n$ , а  $\exp(B)$  — матричная экспонента,  $B \in \mathbb{R}^{n \times n}$ . Матричной экспонентой обозначают ряд

$$I_n + \frac{B}{1!} + \frac{B^2}{2!} + \frac{B^3}{3!} + \frac{B^4}{4!} + \dots = \sum_{k=0}^{\infty} \frac{B^k}{k!}.$$

Найдите производную  $\nabla_x f(x)$ .

**Задача 6.** Рассмотрим симметричную матрицу  $A \in \mathbb{R}^{n \times n}$  и ее спектральное разложение  $A = Q\Lambda Q^T$ . Пусть  $\lambda \in \mathbb{R}^n$  — это диагональ матрицы  $\Lambda$  (то есть вектор, составленный из собственных значений  $A$ ). Найдите производные:

1.  $\nabla_{\lambda} \operatorname{tr}(A)$

2.  $\nabla_Q \operatorname{tr}(A)$

**Задача 7.** Рассмотрим задачу обучения линейной регрессии с функцией потерь Log-Cosh:

$$Q(w) = \frac{1}{\ell} \sum_{i=1}^{\ell} \ln(\cosh(w^T x_i - y_i))$$

Выпишите формулу для градиента  $\nabla_w Q(w)$ . Запишите ее в матричном виде, используя матрицу объектов-признаки  $X$  и вектор целевых переменных  $y$ .

**Задача 8.** Найдите симметричную матрицу  $X$ , наиболее близкую к матрице  $A$  по норме Фробениуса  $(\sum_{i,j} (x_{ij} - a_{ij})^2)$ . Иными словами, решите задачу условной матричной минимизации

$$\begin{cases} \|X - A\|_F^2 \rightarrow \min_X \\ X^T = X \end{cases}$$

**Hint:** Надо будет выписать лагранжиан. А ещё пригодится тот факт, что  $\sum_{i,j} (x_{ij} - a_{ij})^2 = \|X - A\|_F^2 = \operatorname{tr}((X - A)^T(X - A))$ .