

Построение и оценка моделей машинного обучения

Теоретическое домашнее задание №1

Задача 1. Найдите производную по матрице $A \in \mathbb{R}^{n \times n}$

$$\frac{\partial}{\partial A} \log \det A = ?$$

Задача 2. Найдите производную по вектору $a \in \mathbb{R}^n$

$$\frac{\partial}{\partial a} (a^T \exp(aa^T)a) = ?,$$

где $\exp(B)$ — **матричная экспонента**, $B \in \mathbb{R}^{n \times n}$. Матричной экспонентой обозначают ряд

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

Задача 3. В методе t-SNE, который широко используется для визуализации данных, задача построения низкоразмерных представлений объектов сводится к минимизации функционала

$$\begin{aligned} C &= \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}} \rightarrow \min_{\{y_i\}} \\ q_{ij} &= \frac{(1 + \|y_i - y_j\|^2)^{-1}}{\sum_{k \neq l} (1 + \|y_k - y_l\|^2)^{-1}} \\ p_{i|j} &= \frac{\exp(-\|x_i - x_j\|^2 / 2\sigma_i^2)}{\sum_{k \neq i} \exp(-\|x_i - x_k\|^2 / 2\sigma_i^2)}, \quad p_{i|i} = 0 \\ p_{ij} &= \frac{p_{i|j} + p_{j|i}}{2N} \end{aligned}$$

Найдите производную, которая необходима для решения задачи градиентным спуском:

$$\frac{\partial C}{\partial y_i} = ?$$

Задача 4. Рассмотрим задачу обучения линейной регрессии

$$Q(w) = (y - Xw)^T(y - Xw) \rightarrow \min_w$$

Будем решать её с помощью градиентного спуска. Допустим, мы находимся на некоторой итерации t , и хотим выполнить очередной шаг

$$w^{(t)} = w^{(t-1)} - \alpha \nabla_w Q(w^{(t-1)})$$

При известных $y, X, w^{(t-1)}$, найдите длину шага α , при которой уменьшение значения функционала будет наибольшим:

$$Q(w^{(t-1)} - \alpha \nabla_w Q(w^{(t-1)})) \rightarrow \min_{\alpha}$$