Задание номер 4

Н. К. Животовский
nikita.zhivotovskiy@phystech.edu

8 мая 2017 г.

Задание принимается до 2.00 утра 22 мая по адресу slt.fupm.2017@gmail.com. Не забываем, что в начале текста задания *обязательно* указывается:

- С кем вы делали это задание.
- Какие источники (кроме материалов лекций) вы использовали.

Задание оформляется в формате pdf (текст набирается в latex/Word) и в таком виде, чтобы ваши коллеги могли разобрать текст решения. Задания, оформленные не в соответствии с указанными правилами, не принимаются. Желательно оставлять зазоры между задачами для пометок.

Упражнение 1. Покажите, что для модели гауссовской последовательности оценки $\hat{\theta}^{\mathrm{hard}}$ и $\hat{\theta}^{\mathrm{hard}}$ выражаются соответственно

1. $\hat{\theta}^{\mathrm{hard}} = \arg\min_{\theta \in \mathbb{R}^d} \left(\|Y - \theta\|_2^2 + 4\tau^2 \|\theta\|_0\right).$
2. $\hat{\theta}^{\mathrm{soft}} = \arg\min_{\theta \in \mathbb{R}^d} \left(\|Y - \theta\|_2^2 + 4\tau \|\theta\|_1\right).$

Указание. Напоминаем, что $\hat{\theta}_i^{\mathrm{hard}} = y_i \mathbf{I}[|y_i| > 2\tau]$ и $\hat{\theta}_i^{\mathrm{soft}} = y_i \left(1 - \frac{2\tau}{|y_i|}\right)_+$.

■

Упражнение 2. [Ridge regression] В модели линейной регрессии с фиксированным дизайном рассмотрим оценку

$$\hat{\theta}^{\mathrm{ridge}} = \arg\min_{\theta \in \mathbb{R}^d} \left(\frac{1}{n}\|Y - X\theta\|_2^2 + \tau \|\theta\|_2^2\right)$$

1. Докажите, что $\hat{\theta}^{\mathrm{ridge}}$ единственна и выпишите явную формулу для нее.
2. Рассчитайте смещение оценки $\hat{\theta}^{\mathrm{ridge}}$ и докажите, что его абсолютная величина ограничена $\|\theta^*\|_2$.

Упражнение 3. [Концентрация для χ_k^2] Докажите, что существуют абсолютные константы $c_1, c_2 > 0$ такие что для любого $\delta \in (0, c_1)$ выполнено

$$P(Z \leq (1-\delta)\mathbb{E}Z) \leq \exp(-c_2\delta^2 k)$$

и

$$P(Z \geq (1+\delta)\mathbb{E}Z) \leq \exp(-c_2\delta^2 k),$$

где $Z \sim \chi_k^2$.

Упражнение 4. [Операторная норма случайной матрицы] Пусть A — квадратная матрица $n \times n$ элементы которой есть независимые центрированные субгауссовские случайные величины с параметром σ.

1. Пусть $\mathcal{N}_\varepsilon$ — минимальное ε-покрытие единичной сферы в $\mathbb{R}^n$. Докажите, что

$$\sup_{x \in \mathcal{N}_\varepsilon} \|Ax\|_2 \leq \|A\|_{op} \leq \frac{1}{1-\varepsilon} \sup_{x \in \mathcal{N}_\varepsilon} \|Ax\|_2.$$

2. Докажите, что для числа покрытия единичной сферы выполнено

$$\left(\frac{1}{\varepsilon}\right)^n \leq \mathcal{N}_\varepsilon \leq \left(\frac{c_1}{\varepsilon}\right)^n,$$

для некоторой $c_1 > 0$. Расстояние считать евклидовым.

3. Используя неравенства для максимумов субгауссовских случайных величин покажите что для некоторой $c_2 > 0$

$$\mathbb{E}\|A\|_{op} \leq c_2 \sigma \sqrt{n}.$$

Задача 1. Приведите пример распределения P_X, для которого линейные решающие правила дают значение коэффициента рассогласования $\tau(\varepsilon)$ порядка $\frac{1}{\varepsilon}$. Что можно сказать о порядках сходимости активного обучения в этом случае?

Задача 2. [Оценивание ковариационной матрицы] Пусть дана i.i.d. выборка из n случайных векторов $X_1, \ldots, X_n$, таких что $X \in \mathbb{R}^d$ и $\mathbb{E}[XX^T] = \Sigma$ для положительно-определенной ковариационной матрицы Σ. Рассмотрим оценку $\hat{\Sigma} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i X_i^T$.

1. Предполагая, что для любого единичного вектора t случайная величина $X^T t$ является субгауссовской с параметром ограниченным σ, оцените порядок

$$\mathbb{E}\|\hat{\Sigma} - \Sigma\|_{op}.$$

2. Согласуются ли полученные порядки с одномерным случаем?