CS 15-759: Homework 4

Due: 3/14/2025, 11:59 PM on Canvas

Bonus points: If you find typos in my homework or lecture notes, please email me. You will earn +1 bonus points per typo found, and potentially more for especially egregious typos.

Hints: Hints are on the last page. It is recommended to think about the problem without hints for a while, and then look at the hints when stuck. The problems are meant to be difficult, so there is no shame in looking at the hints. If you make partial progress on problems (e.g., by following the hints) you will get partial points.

Problem 1: Leverage Scores are Sensitivities (10 Points)

Let $M \in \mathbb{R}^{n \times n}$ be a positive definite matrix, and $v \in \mathbb{R}^n$. Prove that

$$\max_{x \neq 0} \frac{\langle x, v\rangle^2}{x^\top M x} = v^\top M^{-1} v.$$

Why does this say that leverage scores are exactly the sensitivities?

Problem 2: Energy Increase Lemma (30 Points)

Prove the following statement. Let $g \in \mathbb{R}^n$ and $a_1, \dots, a_m \in \mathbb{R}^n$. For $r \in \mathbb{R}^m_{>0}$ define

$$\mathcal{E}(r) := \min_{g^\top x = -1} \sum_{i=1}^m r_i \langle a_i, x\rangle^2.$$

Prove that for $r, r' \in \mathbb{R}^m_{>0}$ such that $r'_i \geq r_i$ for all $i = 1, \dots, m$ that

$$\mathcal{E}(r') - \mathcal{E}(r) \geq \sum_{i=1}^m \frac{r'_i - r_i}{r'_i} r_i \langle a_i, \Delta_r\rangle^2,$$

where

$$\Delta_r := \mathrm{argmin}_{g^\top x = -1} \sum_{i=1}^m r_i \langle a_i, x\rangle^2.$$

Problem 3: Sparsification via Matrix Chernoff (30 Points)

For this problem you may use the following theorem, which is called *matrix Bernstein*.

Theorem 0.1. *Let $X_1, \ldots, X_m$ be independent symmetric random matrices satisfying $\mathbb{E}[X_i] = 0$, $\|X_i\| \le k$, and $\left\|\sum_{i=1}^m \mathbb{E}[X_i^2]\right\| \le \sigma^2$. Then for some constant $c > 0$, it holds that*

$$\Pr\left[\left\|\sum_{i=1}^m X_i\right\| \ge t\right] \le 2n \cdot \exp\left(-c \cdot \min\left\{\frac{t^2}{\sigma^2}, \frac{t}{k}\right\}\right).$$

Let $A^\top A = \sum_{i=1}^m a_i a_i^\top$, and define $\tau_i := a_i^\top (A^\top A)^{-1} a_i$ as the leverage scores. Set $p_i = \min\{1, C\varepsilon^{-2} \log n\}$ for sufficiently large C. Define independent random weights $w_1, \ldots, w_m$ as follows. $w_i = 1/p_i$ with probability p_i and 0 otherwise (with probability $1 - p_i$). Prove that for sufficiently large C,

$$\Pr\left[(1-\varepsilon)A^\top A \preceq \sum_{i=1}^m w_i a_i a_i^\top \preceq (1+\varepsilon)A^\top A\right] \ge 1 - n^{-10}.$$

Problem 4: Lewis Weights for the Huber Loss (30 Points)

Let $H : \mathbb{R} \to \mathbb{R}_{\ge 0}$ be the *Huber loss* function, defined as $H(x) = x^2/2$ for $|x| \le 1$, and $H(x) = |x| - 1/2$ for $|x| \ge 1$. Let $a_1, \ldots, a_m \in \mathbb{R}^n$ be full rank, and $\lambda > 0$. You can check that $H(x)$ is a convex function (though, this is not useful for the problem stated next).

Prove that there are unique weights $w_1, \ldots, w_m > 0$ such that for the matrix $M := \sum_{i=1}^m w_i a_i a_i^\top$ it holds that

$$\frac{H(\sqrt{a_i^\top M^{-1} a_i})}{w_i a_i^\top M^{-1} a_i} = \lambda$$

for all $i = 1, 2, \ldots, m$.

1 Hints

Problem 1: Use calculus. Or: apply the Cauchy-Schwarz inequality.

Problem 2: (Will flesh out later). Write $\mathcal{E}(r)$ as a dual maximization problem. Then use Δ_r as a certificate for this maximization problem.

Problem 3: You may assume that $A^\top A = I$, by changing $a_i \to (A^\top A)^{-1/2} a_i$.

Problem 4: Define a contractive map that has these weights as a fixed point.

References