COMS 4995-1 S20 Homework 0 (due January 31, 2020)

Instructions

Create an account on Gradescope, and enroll in this course with the entry code 975V76. In Gradescope, set your "Student ID #" to your UNI, using only lowercase letters and numbers (e.g., abc1234). If your Student ID is not set to your UNI in this way, you are at risk of being unenrolled within Gradescope!

Submit your write-up on Gradescope as a neatly typeset PDF document by 11:00 PM of the due date. Please use TEX, LATEX, or a similar system.

On Gradescope, be sure to select the pages containing your answer for each problem. More details can be found on the Gradescope Student Workflow help page.

(If you don't select pages containing your answer to a problem, you'll receive a zero for that problem.)

Make sure **your name and your UNI** appear prominently on the first page of your write-up.

Comments

- It's possible that you've seen some (or all!) of these problems in a previous context. If this is the case, just please re-derive the solutions from basic principles without referring back to your old notes. Also, please indicate in the write-up if you had seen a problem before. (You won't lose any credit for this; it would just be helpful for us to know about this fact.)
- This is a "calibration" homework, meaning that you should use it to self-assess whether this course is suitable for you.
- This homework assignment *will* be graded, but it will count for less than half of each subsequent homework assignment in the final grade.

Problem 1 (15 points)

(a) Let $\boldsymbol{A} \in \mathbb{R}^{n\times d}$ and $\boldsymbol{b} \in \mathbb{R}^n$ be given. Consider the *least squares* optimization problem

$$\min_{\boldsymbol{w}\in\mathbb{R}^d} \|\boldsymbol{A}\boldsymbol{w} - \boldsymbol{b}\|_2^2.$$

Prove that a vector $\boldsymbol{w} \in \mathbb{R}^d$ is a minimizer of the least squares problem if and only if it solves the *normal equations*

$$\boldsymbol{A}^\top \boldsymbol{A}\boldsymbol{w} - \boldsymbol{A}^\top \boldsymbol{b}.$$

(b) Prove that a solution to the normal equations of minimum Euclidean norm must be (i) unique and (ii) contained in the row space of $\boldsymbol{A}$.

(c) For $p \geq 1$, the l^p norm for vectors in $\mathbb{R}^n$ is defined by

$$\|\boldsymbol{x}\|_p := \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \qquad (\forall \boldsymbol{x} = (x_1, \dots, x_n) \in \mathbb{R}^n).$$

Prove that $\|\boldsymbol{x}\|_1 \leq \sqrt{n}\|\boldsymbol{x}\|_2$ for every $\boldsymbol{x} \in \mathbb{R}^n$.

Problem 2 (10 points)

Let X be a non-negative integer-valued random variable.

(a) Prove that $\Pr(X \neq 0) \leq \mathbb{E}X$.

(b) Prove that

$$\Pr(X = 0) \leq \frac{\operatorname{var}(X)}{\mathbb{E}(X^2)}.$$

Hint: For (b), use the Cauchy-Schwarz inequality.

Problem 3 (10 points)

Let vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_d \in \mathbb{R}^n$ and a positive integer k be given. Prove that for any vector $\boldsymbol{u}$ in the convex hull of $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_d$, there exist non-negative integers $k_1, \ldots, k_d$ with $\sum_{i=1}^d k_i = k$ and

$$\left\| \boldsymbol{u} - \frac{1}{k} \sum_{i=1}^{d} k_i \boldsymbol{v}_i \right\|_2^2 \leq \frac{\max_i \|\boldsymbol{v}_i\|_2^2 - \|\boldsymbol{u}\|_2^2}{k}.$$

Hint: Recall that a vector $\boldsymbol{u}$ is in the convex hull of $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_d$ if there exist non-negative real numbers $\alpha_1, \ldots, \alpha_d \geq 0$ with $\sum_{i=1}^d \alpha_i = 1$ such that

$$\boldsymbol{u} = \sum_{i=1}^{d} \alpha_i \boldsymbol{v}_i.$$

This suggests a natural proof via the probabilistic method.

Problem 4 (10 points)

Let $X_1, \dots, X_n$ be mean-zero random variables taking values in the interval $[-1, +1]$, and let $S := \sum_{i=1}^{n} X_i$.

(a) Assume that $X_1, \dots, X_n$ are *uncorrelated* (i.e., $\mathbb{E}[X_i X_j] = 0$ for all $1 \leq i < j \leq n$). Use Chebyshev's inequality to fill in the blank ($\square$) in the following statment: for any $\delta \in (0, 1)$,

$$\Pr(|S| \geq \square) \leq \delta.$$

(Therefore, with probability at least $1 - \delta$, we have $|S| \leq \square$.) Try to make the quantity you fill-in as small as possible (up to constant factors), and prove that it is correct.

(b) Now assume that $X_1, \dots, X_n$ are *independent*. Use Hoeffding's inequality to fill in the blank ($\square$) in the same statement as in Part (a). Try to make the quantity you fill-in as small as possible (up to constant factors), and prove that it is correct.

(c) *Optional*: Is the inequality you derived in Part (a) tight?

Note: UML Appendix B more-or-less has the answers to this problem, except that the conditions in their statements of Chebyshev's and Hoeffding's inequalities are overly restrictive. Explain why the conclusions are still valid even with the less restrictive conditions that this problem requires.