Math 126 Problem 19: Least Squares Regression

Spring 2027

Submission Instructions

Please upload your work to Gradescope by **11:59pm on Monday, April 19, 2027 (Pacific time).**

- You are encouraged to discuss ideas and work with your classmates. However, you *must acknowledge* your collaborators at the top of each solution on which you collaborated with others and you *must write* your solutions independently.
- Your solutions to theory questions must be written legibly, or typeset in LaTeX or markdown.
- I recommend that you write your solutions in a Python notebook using Google Colab.
- You should submit your solutions as a *single PDF* via the assignment on Gradescope.

Grading: The point of the problem set is for *you* to learn. To this end, I hope to disincentivize the use of LLMs by *not* grading your work for correctness. Instead, you will grade your own work by comparing it to my solutions. This self-grade is due the Friday *after* the problem set is due, also on Gradescope.
Original notebook: Google Colab.

Least Squares Regression

Let $n \geq d$. Consider a full-rank, feature matrix $\mathbf{A} \in \mathbb{R}^{n \times d}$ and a target vector $\mathbf{b} \in \mathbb{R}^n$. The regression problem is to find a minimizing vector

$$\mathbf{x}^* = \arg\min_{\mathbf{x} \in \mathbb{R}^d} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2.$$

In class, we showed how to quickly solve the regression problem by projecting our feature matrix and target vector into a lower dimensional space. In this problem, you will derive the exact solution for $\mathbf{x}^*$ and explore how long it takes to compute it.

Part A: Optimal Solution

Show that $\mathbf{x}^* = \left(\mathbf{A}^\top \mathbf{A}\right)^{-1} \mathbf{A}^\top \mathbf{b}$.

Hint: Compute the partial derivative with respect to x_k of the ℓ_2-norm and set it equal to 0.

Part B: Time Complexity

Explain why computing $\left(\mathbf{A}^\top \mathbf{A}\right)^{-1} \mathbf{A}^\top \mathbf{b}$ takes $O(nd^2)$ time. Then, load a regression dataset of your choice (e.g., from `scikit-learn`), build $\mathbf{A}$ and $\mathbf{b}$ and find the optimal solution $\mathbf{x}^*$.