CIS5200: Machine Learning Spring 2025

Homework 0

Release Date: January 16, 2025 *Due Date: January 28, 2025*

There are two parts of Homework 0: written and programming. The goal of written component is to give you an idea of the level of mathematical knowledge and maturity expected in this course. You should have seen all this material before; the goal of this homework is to encourage you to go back to some of the material and refresh your memory. The goal of the programming component is to help you get familiar with PyTorch and Google Colab that we will use for the rest of the semester.

- HW0 will count for 5% of the grade, and you will get full credit if you attempt all questions irrespective of whether the answer is correct, partially correct, or wrong.
- While we encourage collaboration on homework, in the specific case of HW0, **you are on your own**. It is a test of your level of readiness to take this course, so please work on it independently. Please use Ed Discussion only if you have clarifications about this homework.
- All written homework solutions are required to be formatted using LaTeX. Please use the template here. This is a good resource to get yourself started, if you have not previously used LaTeX.
- You will submit your solution for the written part of HW0 as a single PDF file via Gradescope. Instructions for how to submit the programming component of HW0 to Gradescope are included in the Colab notebook.
- The deadline is **11:59 PM ET**. Contact TAs on Ed if you face any issues.
- To get off the waitlist, complete HW0 and send your solutions to the Head TAs. Permission for remaining seats will be granted to those that complete HW0, as long as there is space availability.

Here is a list of resources to help brush up on the mathematical background:

- General Review - Mathematics for Machine Learning by Deisenroth, Marc Peter, A. Aldo Faisal, and Cheng Soon Ong, Cambridge University Press, 2020
- Linear Algebra Review
- Probability Review 1 and 2
- Additional Recommended Resources:
  - Matrix Calculus: The Matrix Calculus You Need For Deep Learning
  - Convex Optimization: Boyd and Vandenberghe, Chapter 2-3
  - Vector Calculus: Paul's Online Math Notes - Vector Calculus

Prerequisites: This course assumes familiarity with:

- Linear Algebra: matrix operations, eigenvalues/eigenvectors, vector spaces, matrix rank
- Multivariable Calculus: gradients, partial derivatives, chain rule
- Probability: basic probability rules, conditional probability, expectation, common distributions
- Optimization: convexity, critical points, gradient descent

Disclaimer: If you find HW0 to be very time consuming and extremely difficult, this course may not be right for you.

1 Written Questions

Q1 [*Linear Algebra*] Let A be a real-valued $n \times n$ matrix. Which of the following statements are true? Give a proof or counterexample.

1. If A is invertible, then $det(A^{-1}) = \frac{1}{det(A)}$
2. The sum of eigenvalues equals the trace of the matrix
3. If A has rank k, then exactly k eigenvalues are non-zero

Q2 [*Linear Algebra*] Let $A = \begin{bmatrix} 2 & -1 \\ 4 & -2 \end{bmatrix}$.

1. Find the nullspace of A.
2. Is the vector $[1, 1]^\top$ in the row space of A? Justify your answer.

Q3 [*Linear Algebra*] Consider the matrix $A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$.

1. What are the eigenvalues and corresponding eigenvectors of A?
2. Is A a PSD (positive semidefinite) matrix?
3. Since A is symmetric, its SVD can be written using eigenvalues and eigenvectors of A. Express the SVD of A using your answers from part (a). What do you notice about the relationship between singular values and eigenvalues in this case?

Q4 [*Calculus*] For column vector x ($n \times 1$ vector), answer the following questions:

1. Let $f(x) = \frac{1}{1+\exp(-w^\top x)}$ for column vector w, compute $\nabla_x f(x)$.
2. Let $f(x) = \|Ax - b\|_2^2$ for matrix $A \in \mathbb{R}^{n \times n}$ and n-dimensional vector b, compute $\nabla_x f(x)$.

Q5 [*Geometry*] Consider the hyperplane $w^\top x + b = 0$ for fixed vector $w \in \mathbb{R}^n$ and scalar $b \in \mathbb{R}$.

1. Under what conditions does the hyperplane pass through the origin?
2. What is the distance of any point x_0 from the hyperplane?

Q6 [*Vector Norms*] Let x be an n-dimensional vector.

1. If $\|x\|_\infty = 1$, what is the maximum possible value of $\|x\|_2$ in terms of n?
2. If $\|x\|_2 = 1$, what is the minimum possible value of $\|x\|_1$?

Q7 [*Convexity*] Consider the following functions:

1. Is $f(x) = x^3$ convex on $\mathbb{R}$? Justify your answer.
2. For what values of α is $f(x) = x^4 + \alpha x^2$ convex on $\mathbb{R}$?

Q8 [*Probability*] A spam detection system has the following properties:

- 80% of emails are legitimate (non-spam)
- For legitimate emails, the system has a 95% accuracy
- For spam emails, the system has a 90% accuracy

If the system flags an email as spam, what is the probability that it is actually spam?

Q9 [*Probability*] Let $X_1, X_2, ..., X_n$ be independent random variables where $X_i \sim N(\mu_i, \sigma_i^2)$.

1. Let $Y = \sum_{i=1}^n a_i X_i$ for fixed constants a_i. What is the distribution of Y? Express your answer in terms of μ_i, σ_i^2, and a_i.
2. If all $\mu_i = 0$ and $\sigma_i^2 = 1$, what is $P(\max_{1 \le i \le n} X_i > 2)$? Express your answer in terms of the standard normal CDF Φ.

Q10 [*Probability*] Consider rolling a fair six-sided die repeatedly.

1. What is the expected number of rolls needed to see a 6?
2. What is the expected number of rolls needed to see a 6 followed by a 6?

2 Programming Questions

Use the link here to access the Google Colaboratory (Colab) for the programming. Be sure to make a copy by going to "File", and "Save a copy in Drive". This assignment uses the PennGrader system for students to receive immediate feedback.