Homework1

Hanbyul Joo
4190.408 001, Artificial Intelligence

September 19, 2024

1 Instruction

In this homework, you are encouraged to study basic tools of linear algebra, probability and optimization. Check the following concepts and address it's meaning, algorithms, applications and so on. You may search for internet, but you should not copy and paste it whole sentences. Please make your answers as concise as possible, with only essential parts (each answer should be within a few lines).

- **Make sure to submit HW1**. Students who do not submit HW1 is considered not taking the course, and will be dropped.
- Homeworks need to be submitted electronically on ETL. **Only PDF generated from LaTeX is accepted.**
- File names must be in the following format: **20XX_1XXXX_YOUR-NAME_HW1.pdf**
- Collaborations on solving the homework is allowed. Discussions are encouraged but you should think about the problems on your own.
- If you do collaborate with someone or use a book or website, you are expected to write up your solution independently. That is, close the book and all of your notes before starting to write up your solution.
- Make sure you cite your work/collaborators at the end of the homework.
- **Honor Code**: This document is exclusively for Fall 2024, 4190.408 students with Professor Hanbyul Joo at Seoul National University. Any student who references this document outside of that course during that semester (including any student who retakes the course in a different semester), or who shares this document with another student who is not in that course during that semester, or who in any way makes copies of this document (digital or physical) without consent of Professor Hanbyul Joo is guilty of cheating, and therefore subject to penalty according to the Seoul National University Honor Code.

2 Linear algebra

1.1 1. Prove each of below statements

- For matrix $A, B \in \mathbb{R}^{n\times n}$, $||AB|| \leq ||A||||B||$
- Function $||x||_p = (\sum_{i=1}^{n} |x_i|^p)^{\frac{1}{p}}$ (known as p-norm) is norm (hint : hölder inequality)

2. (Linear independent) Show that the number of elements of a linearly independent set of n-vectors is at most n

3. (Matrix norm) Show that for matrix $A \in \mathbb{R}^{n\times n}$ and its maximum singular value σ_{max}, below holds

$$||A||^2 = \sigma_{max}^2$$

4. (Ax = 0) Solve below problem with SVD

$$\begin{aligned} &\underset{x}{\operatorname{argmin}} \;\; \|Ax\|_2^2 \\ &s.t. \quad \|x\| = 1 \end{aligned}$$

5. (Ax=b) Derive pseudo-inverse of A with A,b, and SVD components

3 Probability

1. Prove below statements for random variable A,B.

- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ (Use only Probability axiom!)
- If A,B are independent $P(A) = P(A|B^c)$

2. (Bayes' theorem) There is a factory with machines M_1, M_2, M_3 accounting for ratio of 0.2, 0.3, 0.5 of entire production each. The failure rate of each machine is known to be 0.03, 0.02, 0.01 each. If one randomly chosen product is found to be a failure, find the probability for each machine for manufacturing the product.

3. (Gaussian distributions) For random variable $X \sim \mathcal{N}(0, 1)$ and $Y \sim \mathcal{N}(0, 1)$ show that $Z = X + Y$ is gaussian and state its variance.

4. (Gaussian random vector) Show that for gaussian random vector X with covariance C, there always exist a matrix A that satisfies $C = AA^T$

5. (Conditional Expectation) There is a piece of wooden stick with a length of L. If we break the stick at random point 2 times, what is the average length of the remaining stick?

4 Optimization

1. Show that if function f is differentiable, $D_v f(x) = \nabla_x f(x)^T v$ holds

2. (Constrained least squares) Show that for n-vector x, a, p-vector d and matrix $C \in \mathbb{R}^{p \times n}$ with linearly independent row vectors, the solution of the problem

$$\begin{array}{ll} \text{minimize} & ||x - a||^2 \\ \text{subject to} & Cx = d \end{array} \tag{1}$$

is $\hat{x} = a - C^\dagger(Ca - d)$. *Hint.* Use directly below optimality condition(known as KKT condition)

$$\begin{bmatrix} 2I & C^T \\ C & 0 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} = \begin{bmatrix} 2a \\ d \end{bmatrix} \tag{2}$$

where p-vector z is lagrangian multiplier of formulation

$$L(\mathbf{x}, \mathbf{z}) = ||\mathbf{x} - a||^2 + z_1(\mathbf{c}_1^T\mathbf{x} - d_1) + \cdots + z_p(\mathbf{c}_p^T\mathbf{x} - d_p)$$

3. (Minimizing norms) Formulate the following problem as LP. Explain in detail the relation between the optimal solution of each problem and the solution of its equivalent LP. (LP form is something like below)

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x \leq b_i \;\; i = 1, \ldots, m \end{array} \tag{3}$$

- Minimize $||Ax - b||_\infty$
- Minimize $||Ax - b||_1$