# Homework1

Hanbyul Joo
4190.408 001, Artificial Intelligence

September 15, 2025

## 1 Instruction

In this homework, you are encouraged to study basic tools of linear algebra, probability and optimization. Please make your answer **concisely but rigorously**.

- Homeworks need to be submitted electronically on ETL. PDF generated from LaTeX is preferred, but for HW1 only, scanned versions of hand-written answers will also be accepted. **However, if the handwriting is illegible, graders may not be able to award points.** Students who choose to submit hand-written answers should keep this in mind.
- Collaborations on solving the homework is allowed. Discussions are encouraged but you should think about the problems on your own.
- File names must be in the following format: **20XX_1XXXX_YOUR-NAME_HW1.pdf**

## 2 Linear algebra

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

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

($\|A\|$ means spectral norm of $A$)
2. (Ax = 0) For some matrix $A \in \mathbb{R}^{m\times n}$ and vector $x \in \mathbb{R}^n$ solve below problem with SVD

$$\begin{aligned} &\underset{x\in\mathbb{R}^n}{\text{argmin}} \quad \|Ax\|_2^2 \\ &\text{subject to } \|x\| = 1 \end{aligned}$$

3. ($Ax = b$) Derive the pseudo-inverse of $A \in \mathbb{R}^{m\times n}$ composed of $A = U\Sigma V^T$ (note that $\Sigma$ does not necessarily needs to be invertible).

## 3 Probability

1. (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.

2. (Gaussian distributions) For two independent random variables $X \sim \mathcal{N}(0, 1)$ and $Y \sim \mathcal{N}(0, 1)$, show that $Z = X + Y$ is gaussian.

3. (KL Divergence) Show that for any probability mass function $p, q \in \mathbb{R}^n$ $D_{KL}(p\|\|q) \geq 0$ holds.

*Hint.* For convex function $f : \mathbb{R}^n \to \mathbb{R}$ and random variable X, Jensen's inequality states

$$f(E(X)) \leq E(f(X))$$

## 4 Optimization

1. Let $f : \mathbb{R}^n \to \mathbb{R}$ be differentiable. (a) Show that for any $v \in \mathbb{R}^n$, $D_v f(x) = \nabla_x f(x)^T v$. (b) Among all unit vectors $u$ with $||u||_2 = 1$, show that the direction that yields the largest instantaneous *decrease* of $f$ at $x$ is $u^\star = -\nabla f(x)/||\nabla f(x)||_2$.
*Hint.* Use Cauchy–Schwarz inequality for (b).

2. Consider a convex function $f : \mathbb{R}^n \to \mathbb{R}$ and let $x^* \in \mathbb{R}^n$ be a local minimizer of $f$. Show that $x^*$ is also a global minimizer of $f$.