Assignment 1

Due October 06, 2025 at 11:59pm

This assignment is to be done individually.

Important Note: The university policy on academic dishonesty (cheating) will be taken very seriously in this course. You may not provide or use any solution, in whole or in part, to or by another student.

You are encouraged to discuss the concepts involved in the questions with other students. If you are in doubt as to what constitutes acceptable discussion, please ask! Further, please take advantage of office hours offered by the instructor and the TA if you are having difficulties with this assignment.

DO NOT:

- Give/receive code or proofs to/from other students
- Use Google to find solutions for assignment

DO:

- Meet with other students to discuss assignment (it is best not to take any notes during such meetings, and to re-work assignment on your own)
- Use online resources (e.g. Wikipedia) to understand the concepts needed to solve the assignment.

Submitting Your Assignment

The assignment must be submitted online on Canvas. You must submit a report in **PDF format**. You may typeset your assignment in LaTeX or Word, or submit neatly handwritten and scanned solutions. We will not be able to give credit to solutions that are not legible.

1 Linear Algebra

1.1 Taylor Expansions

Given $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, consider a nonlinear function $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$ as follows:

$$f(\vec{x}) = 2x_1^2 + x_2^2 + x_3^2 + 2x_2x_3 \tag{1}$$

a) Compute the **Gradient** and **Hessian matrix** of f.

b) Find the **second order Taylor Expansion** at the point $(x_1 = 0, x_2 = 0, x_3 = 0)$.

c) State whether f is **Convex** or **Strictly Convex**. Prove your claim.

1.2 Matrix Rank and Inverse

Prove that if $A \in \mathbb{R}^{m\times n}$, with $m \geq n$, is full rank, then $A^\top A$ is invertible via the following steps:

a) Prove that $A\vec{x} = \vec{0}$ if and only if $\vec{x} = \vec{0}$.

b) Prove that $A^\top A$ is positive definite.

c) Prove that any symmetric positive definite matrix is always invertible by using Eigendecomposition to construct the inverse.

2 SVD and Eigendecomposition

Consider the matrix $A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 1 \end{bmatrix}$.

a) Define the matrix $B = AA^\top$, compute the matrix B.

b) Show that the following is an Eigendecomposition of the matrix B:

$$\begin{bmatrix} \frac{-2}{\sqrt{30}} & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{6}} \\ \frac{-5}{\sqrt{30}} & 0 & \frac{-1}{\sqrt{6}} \\ \frac{-1}{\sqrt{30}} & \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 6 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{-2}{\sqrt{30}} & \frac{-5}{\sqrt{30}} & \frac{-1}{\sqrt{30}} \\ \frac{1}{\sqrt{5}} & 0 & \frac{-2}{\sqrt{5}} \\ \frac{2}{\sqrt{6}} & \frac{-1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

c) Find the singular value matrix for A. That is, if $A = U\Sigma V^\top$ is the SVD of A, find Σ.

d) Consider a SVD of a matrix D as follows:

$$D = \begin{bmatrix} \frac{1}{2} & \frac{-\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 1/2 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{-\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}^\top$$

A matrix $R_\theta \in \mathbb{R}^{2\times 2}$ is a 2D rotation matrix if it has the following form:

$$R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

where $\theta \in \mathbb{R}$. Geometrically speaking, $R_\theta \vec{v}$ rotates $\vec{v}$ counterclockwise by angle θ, for any $\vec{v} \in \mathbb{R}^2$, as shown in Figure 1.

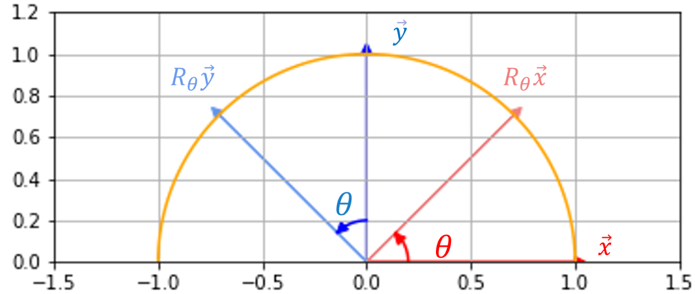

Figure 1: In this case, $\vec{x} = (1, 0)$ and $\vec{y} = (0, 1)$ are both rotated by $\theta = \frac{\pi}{4}$.

Show that $U = \begin{bmatrix} \frac{1}{2} & \frac{-\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix}$ and $V^\top = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{-\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}^\top$ are both rotation matrices and find their corresponding rotational angles θ_U and $\theta_{V^\top}$.

e) Explain all geometric transformations performed by the SVD of the matrix D. In what order are the transformations performed?

3 Convexity

For any $\vec{x}, \vec{y} \in \mathbb{R}^n$ and any $t \in [0, 1]$, a function f is said to be convex if it satisfies any of these conditions:

- $f(t\vec{x} + (1-t)\vec{y}) \le tf(\vec{x}) + (1-t)f(\vec{y})$
- If f is differentiable: $f(\vec{y}) \ge f(\vec{x}) + (\nabla f(\vec{x}))^\top (\vec{y} - \vec{x})$
- If f is twice differentiable: $\mathrm{H}f(\vec{x}) \succeq 0$

a) Given $x \in \mathbb{R}$ and *only* using the definition(s) of convex functions above, prove that the *Huber loss* with parameter $\delta > 0$,

$$\text{Huber}_\delta(x) \;:=\; \begin{cases} \frac{1}{2}x^2, & |x| \le \delta, \\ \delta|x| - \frac{1}{2}\delta^2, & |x| > \delta, \end{cases}$$

is convex.

b) Given $A \in \mathbb{R}^{n\times n}$, $\vec{x} \in \mathbb{R}^n$, $\vec{b} \in \mathbb{R}^n$, and $\lambda \ge 0$, prove that

$$f(\vec{x}) \;=\; \left\| A\vec{x} + \vec{b} \right\|_2 \;+\; \lambda \left\| \vec{x} \right\|_\infty$$

is convex. For this part, you may use the following properties of convex functions:

- $\sum_i w_i f_i(\vec{x})$ is convex if f_i are convex and $w_i \ge 0$.
- For any $A \in \mathbb{R}^{n\times n}$ and $\vec{b} \in \mathbb{R}^n$, $f(A\vec{x} + \vec{b})$ is convex if f is convex.
- $g(f(\vec{x}))$ is convex if f is convex and g is convex and non-decreasing.

You may also use the triangle inequality and positive homogeneity for norms:

$$\|\vec{u} + \vec{v}\| \le \|\vec{u}\| + \|\vec{v}\|, \qquad \|\alpha\,\vec{u}\| = |\alpha|\,\|\vec{u}\| \;\; (\alpha \in \mathbb{R}).$$

Hint: First show that any norm is convex using these two properties; in particular, both the 2-norm and the ∞-norm are convex. Then combine with the properties above.

c) Given $x \in \mathbb{R}$, consider the *Swish* activation

$$f(x) \;=\; x\,\sigma(x) \;=\; \frac{x}{1+e^{-x}},$$

widely used in deep learning. Prove that f is neither convex nor concave on $\mathbb{R}$ (i.e., both f and $-f$ fail to be convex).

Hint: You may use that $\sigma'(x) = \sigma(x)\big(1 - \sigma(x)\big)$ and verify that $f''(x)$ changes sign on $\mathbb{R}$.