Assignment 1

Due Tuesday, October 5, 2021 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　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:

- $f(t\vec{x} + (1-t)\vec{y}) \leq tf(\vec{x}) + (1-t)f(\vec{y})$

- If f is differentiable: $f(\vec{y}) \geq f(\vec{x}) + (\nabla f(\vec{x}))^\top (\vec{y} - \vec{x})$

- If f is twice differentiable: $\mathrm{H}f(\vec{x}) \succeq 0$

a) Given $x \in \mathbb{R}$ and only using the definition of convex functions given above, prove that the rectified linear unit function, $\mathrm{ReLU}(x) := \max(x, 0)$, is convex.

b) Given a $A \in \mathbb{R}^{n \times n}$, $\vec{x} \in \mathbb{R}^n$, $\vec{b} \in \mathbb{R}^n$, and $\lambda \geq 0$, prove that $f(\vec{x}) = \left\| A\vec{x} + \vec{b} \right\|_2 + \lambda \left\| \vec{x} \right\|_1$ 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 $w_i \geq 0$
  - 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

c) Given $x \in \mathbb{R}$, prove that the logistic function $f(x) = \frac{1}{1+e^{-x}}$ is neither convex nor concave (i.e. $-f(x)$ is also not convex).

2　Taylor Expansion and Quadratic Form

a) Given $x \in \mathbb{R}$ and $y \in \mathbb{R}$, consider the following function

$$h(x, y) = -\cos(x^2) + e^{xy} - 2y^2$$

　(i)　Compute the **Gradient** and **Hessian matrix** of h

　(ii)　Find the **second order Taylor expansion** of h at the point $(x = x_0, y = y_0)$

　(iii)　Simplify the equation from (ii) at the point $(x_0 = 0, y_0 = 0)$

　(iv)　Use result from (iii), determine the definiteness for the **Hessian** of h around the point $(x_0 = 0, y_0 = 0)$.

b)　(i)　Let $\vec{x}, \vec{b} \in \mathbb{R}^n$ and $M \in \mathbb{R}^{n \times n}$ be a symmetric and invertible matrix, prove the following equation:

$$(\vec{x} - M^{-1}\vec{b})^\top M (\vec{x} - M^{-1}\vec{b}) = \vec{x}^\top M \vec{x} - 2\vec{b}^\top \vec{x} + \vec{b}^\top M^{-1} \vec{b}$$

(ii) Let $\vec{x}, \vec{\mu}, \vec{\theta} \in \mathbb{R}^n$, $A, B \in \mathbb{R}^{n \times n}$ be symmetric matrices, assume $A + B$ is invertible. Let $f(\vec{x})$ be the sum of two quadratic forms:

$$f(\vec{x}) = (\vec{x} - \vec{\mu})^\top A (\vec{x} - \vec{\mu}) + (\vec{x} - \vec{\theta})^\top B (\vec{x} - \vec{\theta})$$

Show that f can be written as a single quadratic form plus some constant term. (Hint: use (i))

3 SVD and Eigendecomposition

(In this question, you are allowed to use a computer to help you compute the eigenvalues/vectors.)

a) Given a matrix $A = \begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix}$, find an analytical expression of A^n. Show your work. (In the final expression, you should not need to multiply any matrix n times to obtain A^n)

b) Consider the following matrix:

$$A = \begin{pmatrix} \frac{1+4\sqrt{3}}{4\sqrt{2}} & \frac{4-\sqrt{3}}{4\sqrt{2}} \\ \frac{4\sqrt{3}-1}{4\sqrt{2}} & \frac{4+\sqrt{3}}{4\sqrt{2}} \end{pmatrix}$$

(i) Show that the following is a **Singular Value Decomposition** of A:

$$\begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}^\top$$

(ii) A matrix $R \in \mathbb{R}^{2 \times 2}$ is a **2D rotation matrix** if it has the following form:

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

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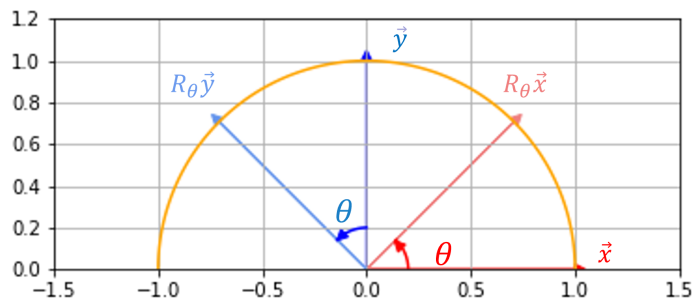
Figure 1: In this case, $\overrightarrow{x} = (1, 0)$ and $\overrightarrow{y} = (0, 1)$ are both rotated by $\theta = \frac{\pi}{4}$

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

(iii) Give an explanation for all geometric transformations performed by the SVD of A. In what order are the transformations performed?

(iv) The unit circle shown in Figure 2 has been transformed by a number of different 2D transformations. The transformation results are shown in Figure 3.
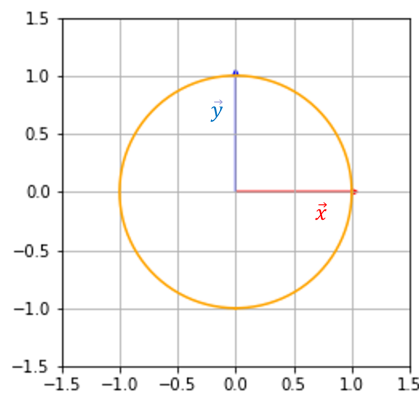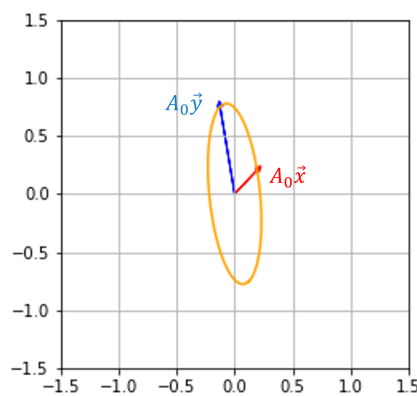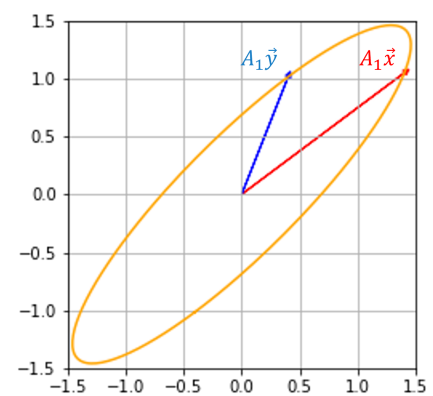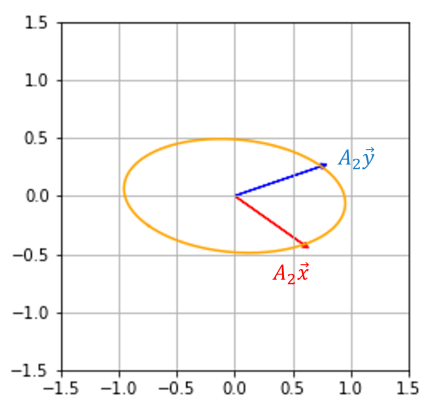

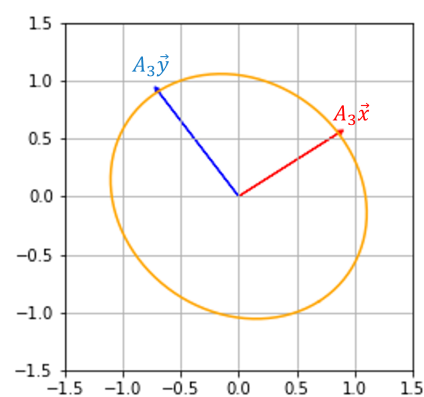Figure 2: unit circle and plane basis $\overrightarrow{x}, \overrightarrow{y}$

(a)

(b)

Figure 3: unit circle after different 2D transforms

Consider the geometric interpretation of the SVD for A. Which of the images from Figure 3 correctly shows the transformation of the unit circle by A? (i.e. Which of the matrices in the figure, A_0, A_1, A_2, A_3, is equal to A?) Please explain your choice.