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

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

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

Let $x, y \in \mathbb{R}$ and $t \in [0,1]$

$$f(x) = \max(x, 0)$$

$$\begin{aligned} f(tx + (1-t)y) &= \max(tx + (1-t)y, 0) \\ &\leq \max(tx, 0) + \max((1-t)y, 0) \\ &= t\max(x, 0) + (1-t)\max(y, 0) \\ &= tf(x) + (1-t)f(y) \end{aligned}$$

Therefore f is convex on $\mathbb{R}$

b) Given a $A \in \mathbb{R}^{n\times n}$, $\overrightarrow{x} \in \mathbb{R}^n$, $\overrightarrow{b} \in \mathbb{R}^n$, and $\lambda \geq 0$, prove that $f(\overrightarrow{x}) = \left\|A\overrightarrow{x} + \overrightarrow{b}\right\|_2 + \lambda \left\|\overrightarrow{x}\right\|_1$ is convex. For this part, you may use the following properties of convex functions:

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

Let $\overrightarrow{x_0}, \overrightarrow{x_1} \in \mathbb{R}^n$ and $t \in [0,1]$
Using triangle inequality, we have $\|\overrightarrow{x_0} + \overrightarrow{x_1}\| \leq \|\overrightarrow{x_0}\| + \|\overrightarrow{x_1}\|$ for any two vectors $\overrightarrow{x_0}, \overrightarrow{x_1}$.

$$\begin{aligned} \|t\overrightarrow{x_0} + (1-t)\overrightarrow{x_1}\| &\leq \|t\overrightarrow{x_0}\| + \|(1-t)\overrightarrow{x_1}\| \\ &= t\,\|\overrightarrow{x_0}\| + (1-t)\,\|\overrightarrow{x_1}\| \end{aligned}$$

So $g_2(\overrightarrow{x}) = \|\overrightarrow{x}\|_2$ and $g_1(\overrightarrow{x}) = \|\overrightarrow{x}\|_1$ are both convex.
By the second given property, $g_2(A\overrightarrow{x} + \overrightarrow{b})$ is also convex.
By the first given property, $f(\overrightarrow{x}) = \left\|A\overrightarrow{x} + \overrightarrow{b}\right\|_2 + \lambda \left\|\overrightarrow{x}\right\|_1 = 1 \times g_2(A\overrightarrow{x} + \overrightarrow{b}) + \lambda g_1(\overrightarrow{x})$ is convex.

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).

$$f(x) = \frac{1}{1+e^{-x}}$$
$$\frac{\partial f}{\partial x} = \frac{e^{-x}}{(1+e^{-x})^2}$$

$$\frac{\partial^2 f}{\partial^2 x} = -\frac{e^{-x}(e^{-x}-1)}{(1+e^{-x})^3} > 0 \text{ if } x < 0$$
$$< 0 \text{ if } x > 0$$

Therefore f is not convex nor concave

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

$$\nabla h(x,y) = \begin{pmatrix} \frac{\partial h}{\partial x} \\ \frac{\partial h}{\partial y} \end{pmatrix} = \begin{pmatrix} 2x\sin(x^2) + ye^{xy} \\ xe^{xy} - 4y \end{pmatrix}$$
$$\mathrm{H}h(x,y) = \begin{pmatrix} \frac{\partial^2 h}{\partial^2 x} & \frac{\partial^2 h}{\partial x \partial y} \\ \frac{\partial^2 h}{\partial y \partial x} & \frac{\partial^2 h}{\partial^2 y} \end{pmatrix} = \begin{pmatrix} 2\sin(x^2) + 4x^2\cos(x^2) + y^2e^{xy} & xye^{xy} + e^{xy} \\ xye^{xy} + e^{xy} & x^2e^{xy} - 4 \end{pmatrix}$$

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

The 2nd order Taylor expansion at (x_0, y_0) is:

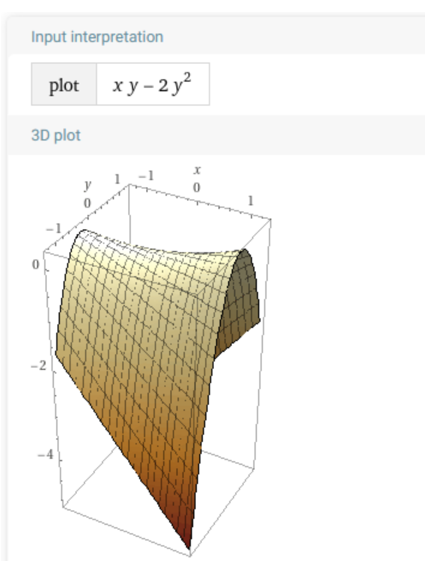
$$h(x_0, y_0) + \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}^\top \nabla h(x,y) + \frac{1}{2}\begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}^\top Hh(x,y)\begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}$$
$$= (-\cos(x_0^2) + e^{x_0y_0} - 2y_0^2) + \begin{pmatrix} x - x_0 & y - y_0 \end{pmatrix}\begin{pmatrix} 2x_0\sin(x_0^2) + y_0e^{x_0y_0} \\ x_0e^{x_0y_0} - 4y_0 \end{pmatrix} +$$
$$\begin{pmatrix} x - x_0 & y - y_0 \end{pmatrix}\frac{1}{2}\begin{pmatrix} 2\sin(x_0^2) + 4x_0^2\cos(x_0^2) + y_0^2e^{x_0y_0} & x_0y_0e^{x_0y_0} + e^{x_0y_0} \\ x_0y_0e^{x_0y_0} + e^{x_0y_0} & x_0^2e^{x_0y_0} - 4 \end{pmatrix}\begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}$$

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

$$\begin{aligned}
Let\ Q(x,y) &= h(0,0) + \begin{pmatrix} x-0 \\ y-0 \end{pmatrix}^T \nabla h(x,y) + \frac{1}{2}\begin{pmatrix} x-0 \\ y-0 \end{pmatrix}^T Hh(x,y)\begin{pmatrix} x-0 \\ y-0 \end{pmatrix} \\
&= 0 + \begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} 0 \\ 0 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & -4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \\
&= \begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} 0 & 1/2 \\ 1/2 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}
\end{aligned}$$

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

Let $\overrightarrow{p_0} = \begin{pmatrix} 0.1 \\ 0.1 \end{pmatrix}$, $\overrightarrow{p_1} = \begin{pmatrix} -0.15 \\ -0.05 \end{pmatrix}$
$Q(0.1, 0.1) = -0.01$
$Q(-0.15, -0.05) = 0.0025$
There exists two points $\overrightarrow{p_0}$ and $\overrightarrow{p_1}$ near (0,0) where $\overrightarrow{p_0}^\top \begin{pmatrix} 0 & 1/2 \\ 1/2 & -2 \end{pmatrix} \overrightarrow{p_0} < 0$ and $\overrightarrow{p_1}^\top \begin{pmatrix} 0 & 1/2 \\ 1/2 & -2 \end{pmatrix} \overrightarrow{p_1} > 0$. Hence, the Hessian of h is indefinite around (0,0).



For further illustration, the plot of the quadratic form Q(x,y) shows a saddle point at (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:

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

Recall that ${M^{-1}}^\top = {M^\top}^{-1}$ and $M = M^\top$

$$\begin{aligned}
(\vec{x} - M^{-1}\vec{b})^\top M(\vec{x} - M^{-1}\vec{b}) &= (\vec{x}^\top - (M^{-1}\vec{b})^\top)M(\vec{x} - M^{-1}\vec{b}) \\
&= (\vec{x}^\top - \vec{b}^\top(M^{-1})^\top)M(\vec{x} - M^{-1}\vec{b}) \\
&= (\vec{x}^\top - \vec{b}^\top M^{-1})M(\vec{x} - M^{-1}\vec{b}) \\
&= (\vec{x}^\top M - \vec{b}^\top M^{-1}M)(\vec{x} - M^{-1}\vec{b}) \\
&= (\vec{x}^\top M - \vec{b}^\top)(\vec{x} - M^{-1}\vec{b}) \\
&= \vec{x}^\top M\vec{x} - \vec{x}^\top MM^{-1}\vec{b} - \vec{b}^\top\vec{x} + \vec{b}^\top M^{-1}\vec{b} \\
&= \vec{x}^\top M\vec{x} - \vec{x}^\top\vec{b} - \vec{b}^\top\vec{x} + \vec{b}^\top M^{-1}\vec{b} \\
&= \vec{x}^\top M\vec{x} - 2\vec{b}^\top\vec{x} + \vec{b}^\top M^{-1}\vec{b}
\end{aligned}$$

(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))

$$\begin{aligned}
(\vec{x} - \vec{\mu})^\top A(\vec{x} - \vec{\mu}) &= \vec{x}^\top A\vec{x} - \vec{\mu}^\top A\vec{x} - \vec{x}^\top A\vec{\mu} + \vec{\mu}^\top A\vec{\mu} \\
&= \vec{x}^\top A\vec{x} - 2\vec{\mu}^\top A\vec{x} + \vec{\mu}^\top A\vec{\mu} \\
(\vec{x} - \vec{\theta})^\top B(\vec{x} - \vec{\theta}) &= \vec{x}^\top B\vec{x} - \vec{\theta}^\top B\vec{x} - \vec{x}^\top B\vec{\theta} + \vec{\theta}^\top B\vec{\theta}
\end{aligned}$$

Equation 1:

$$\begin{aligned}
f(\vec{x}) &= (\vec{x}^\top A\vec{x} - 2\vec{\mu}^\top A\vec{x} + \vec{\mu}^\top A\vec{\mu}) + (\vec{x}^\top B\vec{x} - 2\vec{\theta}^\top B\vec{x} + \vec{\theta}^\top B\vec{\theta}) \\
&= \vec{x}^\top(A + B)\vec{x} - 2(\vec{\mu}^\top A + \vec{\theta}^\top B)\vec{x} + (\vec{\mu}^\top A\vec{\mu} + \vec{\theta}^\top B\vec{\theta})
\end{aligned}$$

Let $M = A + B$, $\vec{b}^\top = (\vec{\mu}^\top A + \vec{\theta}^\top B)$ and $R = \vec{\mu}^\top A\vec{\mu} + \vec{\theta}^\top B\vec{\theta}$
We get the following:

$$\begin{aligned}
f(\vec{x}) &= \vec{x}^\top(A + B)\vec{x} - 2(\vec{\mu}^\top A + \vec{\theta}^\top B)\vec{x} + (\vec{\mu}^\top A\vec{\mu} + \vec{\theta}^\top B\vec{\theta}) \\
&= \vec{x}^\top M\vec{x} - 2b^\top\vec{x} + R
\end{aligned}$$

Since $M = A + B$ is symmetric and invertible, we can use equation (i):

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

Re-writting this equation, we get:

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

Use this in the expression of f, we now get:

$$\begin{aligned} f(\overrightarrow{x}) &= (\overrightarrow{x} - M^{-1}\overrightarrow{b})^\top M(\overrightarrow{x} - M^{-1}\overrightarrow{b}) - \overrightarrow{b}^\top M^{-1}\overrightarrow{b} + R \\ &= (\overrightarrow{x} - \overrightarrow{\psi})^\top M(\overrightarrow{x} - \overrightarrow{\psi}) + C \end{aligned}$$

Where $\overrightarrow{\psi} = M^{-1}\overrightarrow{b} = (A+B)^{-1}\overrightarrow{b}$ and
$C = R - \overrightarrow{b}^\top M^{-1}\overrightarrow{b} = \overrightarrow{\mu}^\top A\overrightarrow{\mu} + \overrightarrow{\theta}^\top B\overrightarrow{\theta} - \overrightarrow{b}^\top M^{-1}\overrightarrow{b}$, which is a constant term.
So f can indeed be written as a quadratic form, plus a constant term that is independent from $\overrightarrow{x}$.

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)

The matrix A has eigenvalues $\lambda_1 = 5$ and $\lambda_2 = 3$

Its eigenvectors are $v_0 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$, $v_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$

Since A is symmetrical, eigendecomposition can be performed:

$$A = Q\Lambda Q^\top = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

Note that Q is orthogonal so $Q^\top Q = I$

Therefore:

$$
\begin{aligned}
A^n &= (Q\Lambda Q^\top)(Q\Lambda Q^\top)(Q\Lambda Q^\top)...(Q\Lambda Q^\top) \\
&= Q\Lambda(Q^\top Q)\Lambda(Q^\top Q)...(Q^\top Q)\Lambda Q^\top \\
&= Q\Lambda\Lambda...\Lambda Q^\top \\
&= Q\Lambda^n Q^\top \\
&= \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 5^n & 0 \\ 0 & 3^n \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \\
&= \frac{1}{2}\begin{bmatrix} 5^n+3^n & 5^n-3^n \\ 5^n-3^n & 5^n+3^n \end{bmatrix}
\end{aligned}
$$

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:

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

(1). $U = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$, we need to prove that U is orthogonal, i.e. $UU^\top = I$.

$$
\begin{aligned}
UU^\top &= \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix} \\
&= \begin{pmatrix} \frac{\sqrt{2}}{2}\times\frac{\sqrt{2}}{2} + (-\frac{\sqrt{2}}{2})(-\frac{\sqrt{2}}{2}) & \frac{\sqrt{2}}{2}\times\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}(-\frac{\sqrt{2}}{2}) \\ \frac{\sqrt{2}}{2}\times\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}(-\frac{\sqrt{2}}{2}) & \frac{\sqrt{2}}{2}\times\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}\times\frac{\sqrt{2}}{2} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I
\end{aligned}
$$

(2). Σ is diagonal with real non-negative entries by inspection.

(3). $V = \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}$, we need to prove that V is orthogonal, i.e. $VV^\top = I$.

$$
\begin{aligned}
VV^\top &= \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \\
&= \begin{pmatrix} \frac{\sqrt{3}}{2}\times\frac{\sqrt{3}}{2} + \frac{1}{2}\times\frac{1}{2} & \frac{\sqrt{3}}{2}(-\frac{1}{2}) + -\frac{1}{2}\times\frac{\sqrt{3}}{2} \\ (-\frac{1}{2})\frac{\sqrt{3}}{2} + \frac{\sqrt{3}}{2}\times\frac{1}{2} & (-\frac{1}{2})(-\frac{1}{2}) + \frac{\sqrt{3}}{2}\times\frac{\sqrt{3}}{2} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I
\end{aligned}
$$

(4). We need to verify that $U\Sigma V^\top = A$ by multiply all the matrices.

$$U\Sigma = \begin{pmatrix} \frac{\sqrt{2}}{2} * 2 + 0 & 0 - \frac{\sqrt{2}}{2} \times \frac{1}{2} \\ 2\frac{\sqrt{2}}{2} + 0 & 0 + \frac{\sqrt{2}}{2} \times \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \sqrt{2} & -\frac{1}{2\sqrt{2}} \\ \sqrt{2} & \frac{1}{2\sqrt{2}} \end{pmatrix}$$

$$U\Sigma V^\top = \begin{pmatrix} \sqrt{2} & -\frac{1}{2\sqrt{2}} \\ \sqrt{2} & \frac{1}{2\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}\sqrt{3}}{2} + \frac{1}{4\sqrt{2}} & \frac{\sqrt{2}}{2} - \frac{\sqrt{3}}{4\sqrt{2}} \\ \frac{\sqrt{2}\sqrt{3}}{2} - \frac{1}{4\sqrt{2}} & \frac{\sqrt{2}}{2} + \frac{\sqrt{3}}{4\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{1+4\sqrt{3}}{4\sqrt{2}} & \frac{4-\sqrt{3}}{4\sqrt{2}} \\ \frac{4\sqrt{3}-1}{4\sqrt{2}} & \frac{4+\sqrt{3}}{4\sqrt{2}} \end{pmatrix} = A$$

Therefore, $U\Sigma V^\top$ is a valid SVD of A

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

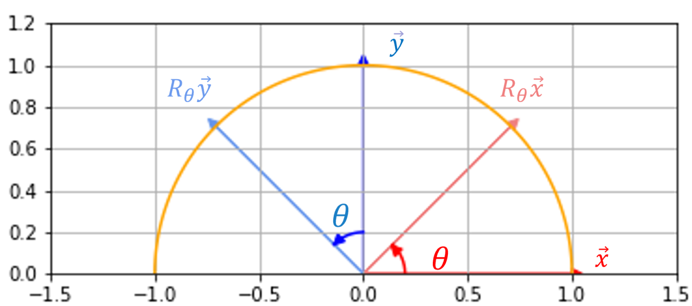


Figure 1: In this case, $\overrightarrow{x} = (1, 0)$ and $\overrightarrow{y} = (0, 1)$ are both rotated by $\theta = \frac{\pi}{4}$

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 $\theta_{V^\top}$.

$$U = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$$

can be written as

$$U = \begin{pmatrix} \cos(\frac{\pi}{4}) & -\sin(\frac{\pi}{4}) \\ \sin(\frac{\pi}{4}) & \cos(\frac{\pi}{4}) \end{pmatrix}$$

$\Rightarrow \theta_U = \frac{\pi}{4}$, U is a rotation matrix

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

can be written as

$$V^\top = \begin{pmatrix} \cos(-\frac{\pi}{6}) & -\sin(-\frac{\pi}{6}) \\ \sin(-\frac{\pi}{6}) & \cos(-\frac{\pi}{6}) \end{pmatrix}$$

$\Rightarrow \theta_{V^\top} = -\frac{\pi}{6}$, $V^\top$ is a rotation matrix

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

1. $V^\top$ rotates everything clockwise by $\frac{\pi}{6}$
2. Σ scales the circle by a factor of 2 in the x-direction, and by a factor of $1/2$ in the y-direction
3. U rotates everything counter-clockwise by $\frac{\pi}{4}$

(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.

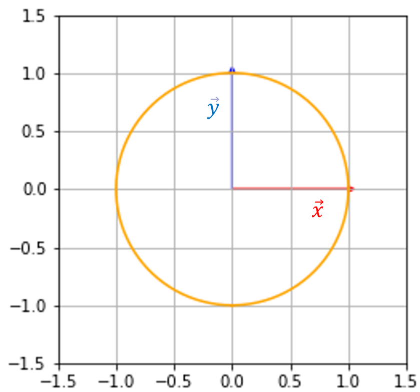
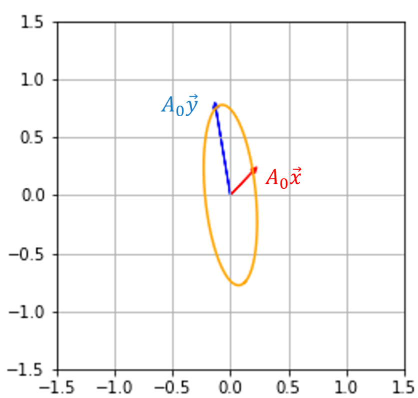
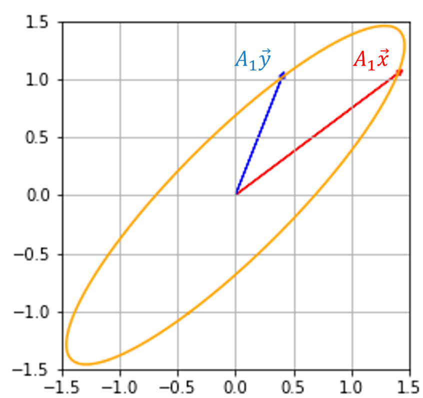


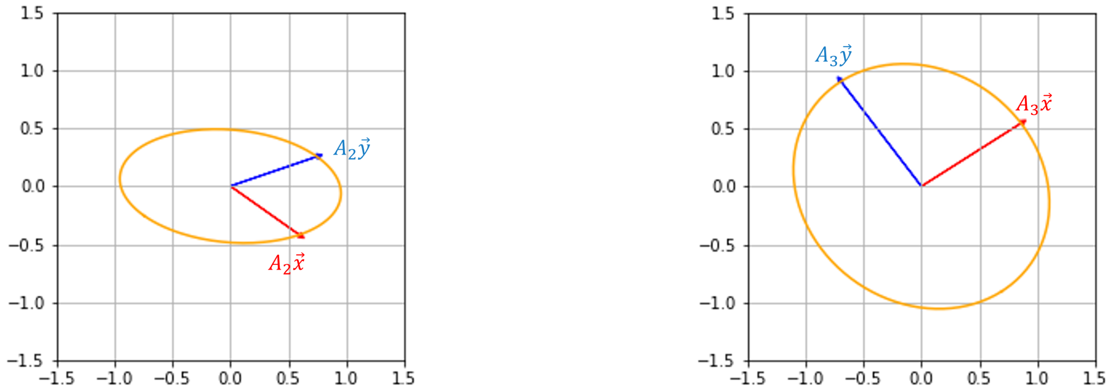
Figure 2: unit circle and plane basis $\overrightarrow{x}$, $\overrightarrow{y}$



(a)



(b)



(c) (d)

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.

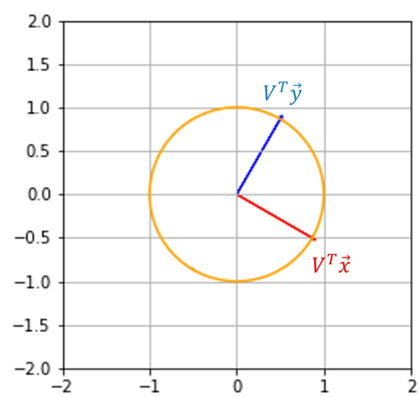


Figure 4: V^T rotates everything clockwise by $\frac{\pi}{6}$

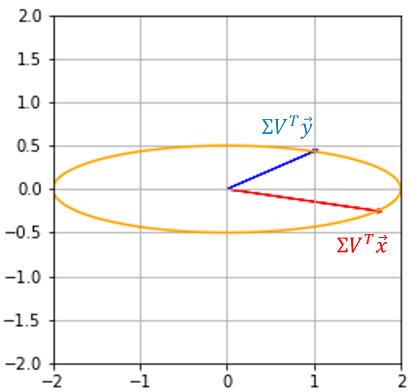


Figure 5: Σ scales the circle by a factor of 2 in the x-axis direction, and by a factor of $1/2$ in the y-axis direction

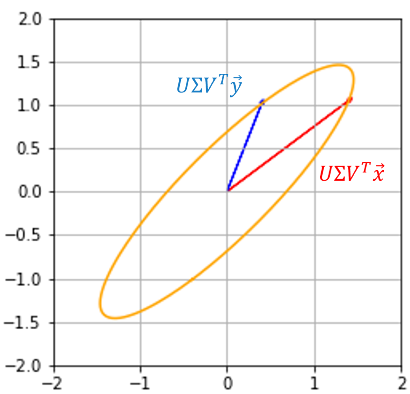


Figure 6: U rotates everything counter-clockwise by $\frac{\pi}{4}$

Therefore, the correct figure is (b).