SIT787 -Mathematics for AI

Trimester 2, 2021

Due: no later than the end of Week 7, Sunday 5 September 2021, 8:00pm AEST

Note:

- A proper way of presenting your solutions is part of the assessment. Please follow the order of questions in your submission.
- Your submission can be handwritten but it must be legible. Please write neatly. If I cannot read your solution, I cannot mark it.
- Provide the way you solve the questions. All steps (workings) to arrive at the answer must be clearly shown. I need to see your thoughts.
- For a final answer without a proper justification no score will be given.
- Only (scanned) electronic submission would be accepted via the unit site (Deakin Sync).
- Your submission must be in ONE pdf file. Multiple files and/or in different file format, e.g. .jpg, will NOT be accepted. If you need to change the format of your submission, it will be subject to the late submission penalty.

Question 1) Consider these vectors:

$$\boldsymbol{u} = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} 2 \\ 2 \\ -2 \\ 4 \end{bmatrix}, \text{ and } \boldsymbol{w} = \begin{bmatrix} 1 \\ 2 \\ 0 \\ -2 \end{bmatrix}$$

(i) Determine which two vectors are most similar to each other based on these norms:

(a) ℓ_2 norm:

$$\text{dist}(\boldsymbol{x}, \boldsymbol{y}) = ||\boldsymbol{x} - \boldsymbol{y}||_2 = ||\boldsymbol{y} - \boldsymbol{x}||_2 = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

(b) ℓ_1 norm:

$$\text{dist}(\boldsymbol{x}, \boldsymbol{y}) = ||\boldsymbol{x} - \boldsymbol{y}||_1 = ||\boldsymbol{y} - \boldsymbol{x}||_1 = \sum_{i=1}^{n} |x_i - y_i|$$

(c) ℓ_∞ norm:

$$\text{dist}(\boldsymbol{x}, \boldsymbol{y}) = ||\boldsymbol{x} - \boldsymbol{y}||_\infty = ||\boldsymbol{y} - \boldsymbol{x}||_\infty = \max_{i=1,\ldots,n} |x_i - y_i|$$

(ii) Determine which two vectors are most similar to each other based on the cosine similarity measure (you can use decimals here):

$$\cos(\theta) = \frac{\boldsymbol{x} \cdot \boldsymbol{y}}{||\boldsymbol{x}||||\boldsymbol{y}||}$$

(iii) Explain the reason behind the difference in result between (i) and (ii).

(iv) Can the difference be resolved? Give details of your suggestion, if you have any, and explain the outcome if your suggestions are applied. (you can use decimals here).

[2+2+3+8= 15 marks]

Question 2) Consider the following set of vectors:

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ h \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ k \end{bmatrix} \right\}$$

Based on different values of h and k, discuss the number of independent vectors in this set.

[10 marks]

Question 3) You are tasked with uncovering information about an incomplete matrix, some of whose entries are unknown and denoted as a, b, c, and d :

$$A = \begin{bmatrix} -1 & 0 & a \\ b & 4 & c \\ d & 0 & 0 \end{bmatrix}$$

Given that $\text{rank}(A) = 2$, how many distinct eigenvalues does A have? **[10 marks]**

Question 4) Consider the following matrix:

$$A = \begin{bmatrix} 3 & -3 & 0 \\ 3 & -1 & 2 \\ b & 0 & 2 \end{bmatrix}$$

Find the values for b (if possible) so that:

(i) The determinant of A is 4.

(ii) The rank of A is 2.

(iii) $\frac{1}{2}$ is an eigenvalue of A^{-1}.

(iv) The system $A\boldsymbol{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ has no solutions.

(v) The system $A\boldsymbol{x} = \begin{bmatrix} -3 \\ 1 \\ 2 \end{bmatrix}$ has infinitely many solutions.

[5 + 5 + 5 + 5+ 5 = 25 marks]

Now consider

$$A = \begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix}$$

for questions 5, 6, and 7.

Question 5) Find $S = A^T A$ matrix and

(a) Find the characteristic polynomial of S

(b) Find the eigenvalues of S, and call them λ_1 and λ_2 where $\lambda_1 \geq \lambda_2$.

(c) Find the eigenvectors S.

(d) Are the eigenvectors orthonormal? If they are not, convert them into orthonormal vectors using the Gram-Schmidt process. Call them $\boldsymbol{v_1}, \boldsymbol{v_2}$ after orthonormalisation.

(e) Using the eigenvalues of S, $\lambda_1 \geq \lambda_2$, set $\sigma_1 = \sqrt{\lambda_1}$ and $\sigma_2 = \sqrt{\lambda_2}$ and make this matrix

$$D = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix}$$

We will use this matrix later.
Make a matrix V using the orthonormal vectors you obtained from the previous part. The ordering of the columns of V should be the same as the ordering of the eigenvalues, that is $V = \begin{bmatrix} \boldsymbol{v_1} & \boldsymbol{v_2} \end{bmatrix}$. Show that V is an orthogonal matrix.

[3+3+3+3+3= 15 marks]

Question 6) Find $T = AA^T$ matrix and

(a) Find the characteristic polynomial of T

(b) Find the eigenvalues of T, and order them as $\lambda_1 \geq \lambda_2 \geq \lambda_3$

(c) Find the eigenvectors T.

(d) Are the eigenvectors orthonormal? If they are not, convert them into orthonormal vectors using the Gram-Schmidt process. Call them $\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}$ according to the order of their corresponding eigenvalues.

[3+3+3+3= 12 marks]

Question 7) Consider two orthonormal vectors you obtained in question 6.

(a) Using them we want to make three orthonormal vectors $\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}$ such that

$$\boldsymbol{u_1} = \frac{1}{\sigma_1} A\boldsymbol{v_1}$$

$$\boldsymbol{u_2} = \frac{1}{\sigma_2} A\boldsymbol{v_2}$$

Also, find the third vector $\boldsymbol{u_3}$,

$$\boldsymbol{u_3} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

such that $\boldsymbol{u_3}$ is a unit vector that is orthogonal to $\boldsymbol{u_1}$ and $\boldsymbol{u_2}$. In other words

$$||\boldsymbol{u_3}||_2 = 1 \text{ and } \boldsymbol{u_3} \perp \boldsymbol{u_1} \text{ and } \boldsymbol{u_3} \perp \boldsymbol{u_2}.$$

Put these three vectors as columns in a matrix $U = \begin{bmatrix} \boldsymbol{u_1} & \boldsymbol{u_2} & \boldsymbol{u_3} \end{bmatrix}$.

(b) Show that U is an orthogonal matrix.

(c) Compute UDV^T.

(d) Explain the relationship between $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}\}$ and $\{\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}\}$.

[5+2+3+3= 13 marks]

Question 1) Consider these vectors:

$$\boldsymbol{u} = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \end{bmatrix}, \boldsymbol{v} = \begin{bmatrix} 2 \\ 2 \\ -2 \\ 4 \end{bmatrix}, \text{ and } \boldsymbol{w} = \begin{bmatrix} 1 \\ 2 \\ 0 \\ -2 \end{bmatrix}$$

(i) Determine which two vectors are most similar to each other based on these norms:

(a) ℓ_2 norm:

$$\text{dist}(\boldsymbol{x}, \boldsymbol{y}) = ||\boldsymbol{x} - \boldsymbol{y}||_2 = ||\boldsymbol{y} - \boldsymbol{x}||_2 = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

(b) ℓ_1 norm:

$$\text{dist}(\boldsymbol{x}, \boldsymbol{y}) = ||\boldsymbol{x} - \boldsymbol{y}||_1 = ||\boldsymbol{y} - \boldsymbol{x}||_1 = \sum_{i=1}^{n} |x_i - y_i|$$

(c) ℓ_∞ norm:

$$\text{dist}(\boldsymbol{x}, \boldsymbol{y}) = ||\boldsymbol{x} - \boldsymbol{y}||_\infty = ||\boldsymbol{y} - \boldsymbol{x}||_\infty = \max_{i=1,\dots,n} |x_i - y_i|$$

(ii) Determine which two vectors are most similar to each other based on the cosine similarity measure (you can use decimals here):

$$\cos(\theta) = \frac{\boldsymbol{x} \cdot \boldsymbol{y}}{||\boldsymbol{x}||||\boldsymbol{y}||}$$

(iii) Explain the reason behind the difference in result between (i) and (ii).

(iv) Can the difference be resolved? Give details of your suggestion, if you have any, and explain the outcome if your suggestions are applied. (you can use decimals here).

[2+2+3+8= 15 marks]

To determine which two vectors are more similar, we need to find the distance between them.

$$\vec{u} - \vec{v} = \begin{bmatrix} 1-2 \\ 2-2 \\ 0-(-2) \\ 1-4 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 2 \\ -3 \end{bmatrix} \qquad \vec{u} - \vec{w} = \begin{bmatrix} 1-1 \\ 2-2 \\ 0-0 \\ 1-(-2) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 3 \end{bmatrix}$$

$$\vec{v} - \vec{w} = \begin{bmatrix} 2-1 \\ 2-2 \\ -2-0 \\ 4-(-2) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -2 \\ 6 \end{bmatrix}$$

a $\quad ||\vec{u} - \vec{v}||_2 = \sqrt{(-1)^2 + (0)^2 + (2)^2 + (-3)^2} = \sqrt{1+4+9} = \sqrt{14}$

$$||\vec{u} - \vec{w}||_2 = \sqrt{0^2 + 0^2 + 0^2 + 3^2} = \sqrt{9} = 3$$

$$||\vec{v} - \vec{w}||_2 = \sqrt{1^2 + 0^2 + (-2)^2 + 6^2} = \sqrt{1+4+36} = \sqrt{41}$$

as $3 \leq \sqrt{14} \leq \sqrt{41}$

$$\|\vec{u}-\vec{w}\|_2 \leq \|\vec{v}-\vec{w}\|_2 \leq \|\vec{u}-\vec{v}\|_2$$

based on l_2 norm, $\vec{u}$ and $\vec{w}$ are more similar.

(b) $\|\vec{u}-\vec{v}\|_1 = |-1| + |0| + |2| + |-3| = 1+2+3 = 6$

$\|\vec{u}-\vec{w}\|_1 = |0| + |0| + |0| + |3| = 3$

$\|\vec{v}-\vec{w}\|_1 = |1| + |0| + |-2| + |6| = 1+2+6 = 9$

as $3 \leq 6 \leq 9$

$$\|\vec{u}-\vec{w}\|_1 \leq \|\vec{u}-\vec{v}\|_1 \leq \|\vec{v}-\vec{w}\|_1$$

based on l_1 norm, $\vec{u}$ and $\vec{w}$ are more similar.

(c) $\|\vec{u}-\vec{v}\|_\infty = \max\{|-1|, |0|, |2|, |-3|\} = 3$

$\|\vec{u}-\vec{w}\|_\infty = \max\{|0|, |0|, |0|, |3|\} = 3$

$\|\vec{v}-\vec{w}\|_\infty = \max\{|1|, |0|, |-2|, |6|\} = 6$

as $3 \leq 3 \leq 6$

$$\|\vec{u}-\vec{w}\| \leq \|\vec{u}-\vec{v}\| \leq \|\vec{v}-\vec{w}\|$$

based on l_∞, $\vec{u}, \vec{v}$ are as similar as $\vec{u}$ and $\vec{w}$.

In summary

$$\begin{cases} l_2 \text{ norm: } & \vec{u} \text{ and } \vec{w} \text{ are more similar} \\ l_1 \text{ norm: } & \vec{u} \text{ and } \vec{w} \text{ are more similar} \\ l_\infty \text{ norm: } & \begin{cases} \vec{u} \text{ and } \vec{w} \text{ are more similar} \\ \vec{u} \text{ and } \vec{v} \text{ are more similar} \end{cases} \end{cases}$$

(ii) Cosine similarity measure:

$$\|\vec{u}\|_2 = \sqrt{1^2+2^2+0^2+1^2} = \sqrt{6}$$

$$\|\vec{v}\|_2 = \sqrt{2^2+2^2+(-2)^2+4^2} = \sqrt{4+4+4+16} = \sqrt{28}$$

$$\|\vec{w}\| = \sqrt{1^2+2^2+0^2+(-2)^2} = \sqrt{1+4+0+4} = \sqrt{9} = 3$$

$$\vec{u} \cdot \vec{v} = (1)(2)+(2)(2)+(0)(-2)+(1)(4) = 10$$

$$\vec{u} \cdot \vec{w} = (1)(1)+(2)(2)+(0)(0)+(1)(-2) = 3$$

$$\vec{v} \cdot \vec{w} = (2)(1)+(2)(2)+(-2)(0)+(4)(-2) = -2$$

$$\cos(\theta_{\vec{u},\vec{v}}) = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|\|\vec{v}\|} = \frac{10}{(\sqrt{6})(\sqrt{28})} = \frac{10}{(2.4)(5.3)} = \frac{10}{12.74} \simeq 0.79$$

$$\cos(\theta_{\vec{u},\vec{w}}) = \frac{\vec{u}\cdot\vec{w}}{\|\vec{u}\|\,\|\vec{w}\|} = \frac{3}{(\sqrt{6})(3)} = \frac{1}{2.4} \simeq 0.42$$

$$\cos(\theta_{\vec{v},\vec{w}}) = \frac{\vec{v}\cdot\vec{w}}{\|\vec{v}\|\,\|\vec{w}\|} = \frac{-2}{(\sqrt{28})(3)} = \frac{-2}{(5.3)(3)} = \frac{-2}{15.9} \simeq -0.12$$

$$0.79 \geq 0.42 \geq -0.12$$

based on cosine similarity measure

$\vec{u}$ and $\vec{v}$ are more similar.

In summary

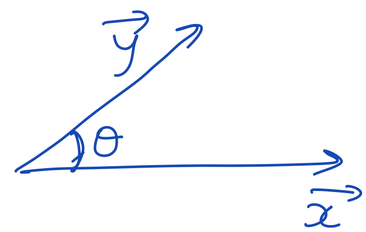
- l_2 norm: $\vec{u}$ and $\vec{w}$ are more similar
- l_1 norm: $\vec{u}$ and $\vec{w}$ are more similar
- l_∞ norm:
  - $\vec{u}$ and $\vec{w}$ are more similar
  - $\vec{u}$ and $\vec{v}$ are more similar
- cosine similarity: $\vec{u}$ and $\vec{v}$ are more similar.

(c) As we saw, these two similarity measure have different outcomes. Based on the Euclidean distance $\vec{u}$ and $\vec{w}$ are the most similar vectors, but the cosine similarity measure suggests $\vec{u}$ and $\vec{v}$ are the most similar ones.

Cosine similarity measure is related to Euclidean distance through this mathematical equation:

$$\|\vec{x}-\vec{y}\|_2^2 = (\vec{x}-\vec{y})^T(\vec{x}-\vec{y}) = \vec{x}\cdot\vec{x} - \vec{x}\cdot\vec{y} - \vec{y}\cdot\vec{x} + \vec{y}\cdot\vec{y}$$

$$= \|\vec{x}\|_2^2 + \|\vec{y}\|_2^2 - 2\,\vec{x}\cdot\vec{y}$$



$$\cos(\theta) = \frac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|_2\|\vec{y}\|_2}$$

$$\boxed{\|\vec{x}-\vec{y}\|_2^2} = \|\vec{x}\|_2^2 + \|\vec{y}\|_2^2 - 2\|\vec{x}\|_2\|\vec{y}\|_2 \boxed{\cos(\theta)}$$

If $\|\vec{x}\|_2$ and $\|\vec{y}\|_2$ can be any number, then it is hard to establish a meaningful relationship between the Euclidean distance and the cosine similarity measure.

If $\vec{x}$ and $\vec{y}$ are normalized; $||\vec{x}||_2 = 1$, $||\vec{y}||_2 = 1$

then $$||\vec{x} - \vec{y}||_2^2 = 1 + 1 - 2\cos(\theta)$$

$$||\vec{x} - \vec{y}||_2^2 = 2(1 - \cos(\theta))$$

In this case there is a linear relationship between the Euclidean distance and the cosie similarity measure.

If $||\vec{x} - \vec{y}||_2^2$ decreases, it makes $\cos(\theta)$ increases and get closer to 1, which means better similarity.

(d) This insight gives us a way to overcome the difference mentioned in part c.

Let's normalise these vectors

$$\hat{u} = \frac{1}{\|\vec{u}\|}\vec{u} = \frac{1}{\sqrt{6}}\begin{bmatrix}1\\2\\0\\1\end{bmatrix} = \begin{bmatrix}1/\sqrt{6}\\2/\sqrt{6}\\0\\1/\sqrt{6}\end{bmatrix} \simeq \begin{bmatrix}0.4\\0.8\\0\\0.4\end{bmatrix}$$

$\sqrt{6} \simeq 2.4$

$\sqrt{28} \simeq 5.3$

$$\hat{v} = \frac{1}{\|\vec{v}\|}\vec{v} = \frac{1}{\sqrt{28}}\begin{bmatrix}2/\sqrt{28}\\2/\sqrt{28}\\-2/\sqrt{28}\\4/\sqrt{28}\end{bmatrix} \simeq \begin{bmatrix}0.4\\0.4\\-0.4\\0.8\end{bmatrix}$$

$$\hat{w} = \frac{1}{\|\vec{w}\|}\vec{w} = \frac{1}{3}\begin{bmatrix}1\\2\\0\\-2\end{bmatrix} = \begin{bmatrix}1/3\\2/3\\0\\-2/3\end{bmatrix} \simeq \begin{bmatrix}0.3\\0.6\\0\\-0.6\end{bmatrix}$$

$$\hat{u} - \hat{v} = \begin{bmatrix}0.4 - 0.4\\0.8 - 0.4\\0 - (-0.4)\\0.4 - 0.8\end{bmatrix} = \begin{bmatrix}0\\0.4\\0.4\\-0.4\end{bmatrix}$$

$$\hat{u} - \hat{w} = \begin{bmatrix}0.4 - 0.3\\0.8 - 0.6\\0 - 0\\0.4 - (-0.6)\end{bmatrix} = \begin{bmatrix}0.1\\0.2\\0\\1\end{bmatrix}$$

$$\hat{v} - \hat{w} = \begin{bmatrix}0.4 - 0.3\\0.4 - 0.6\\-0.4 - 0\\0.8 - (-0.6)\end{bmatrix} = \begin{bmatrix}0.1\\0.2\\-0.4\\1.4\end{bmatrix}$$

$$\left.\begin{aligned} \|\hat{u},\hat{v}\|_2 &= 0.67 \\ \|\hat{u},\hat{w}\|_2 &= 1.1 \\ \|\hat{v},\hat{w}\|_2 &= 1.5 \end{aligned}\right\} \Rightarrow \vec{u},\vec{v} \text{ are the the most similar.}$$

$$\left.\begin{aligned} \cos(\theta_{\hat{u},\hat{v}}) &= \frac{\hat{u}\cdot\hat{v}}{(1)(1)} \cong 0.77 \\ \cos(\theta_{\hat{u},\hat{w}}) &= \frac{\hat{u}\cdot\hat{w}}{(1)(1)} \cong 0.41 \\ \cos(\theta_{\hat{v},\hat{w}}) &= \frac{\hat{v}\cdot\hat{w}}{(1)(1)} \cong -0.12 \end{aligned}\right\} \Rightarrow \vec{u},\vec{v} \text{ are the most similar.}$$

By this action, both similarities are in agreement.

Also, considering other norms, we have:

$$\|\hat{u}-\hat{v}\|_1 = |0| + |0.4| + |0.4| + |-0.4| = 1.2$$

$$\|\hat{u}-\hat{w}\|_1 = |0.1| + |0.2| + |0| + |1| = 1.3$$

$$\|\hat{v}-\hat{w}\|_1 = |0.1| + |0.2| + |-0.4| + |1.4| = 2.1$$

$\Longrightarrow \vec{u}$ and $\vec{v}$ are the most similar

$\|\hat{u}-\hat{v}\|_\infty = \max\{|0|, |0.4|, |0.4|, |-0.4|\} = 0.4$

$\|\hat{u}-\hat{w}\|_\infty = \max\{|0.1|, |0.2|, |0|, |1|\} = 1$

$\|\hat{v}-\hat{w}\|_\infty = \max\{|0.1|, |0.2|, |-0.4|, |1.4|\} = 1.4$

$\Longrightarrow \vec{u}$ and $\vec{v}$ are the most similar

There is an interesting relationship between vector norms $l_1, l_2,$ and l_∞, which is called equivalent norms. If $\vec{x} \in \mathbb{R}^n$, then we have

$$\|\vec{x}\|_\infty \leq \|\vec{x}\|_2 \leq \|\vec{x}\|_1 \leq \sqrt{n}\,\|\vec{x}\|_2 \leq n\,\|x\|_\infty$$

in other words, these norms change proportionally. However, in this problem, we mainly showed the relationship between l_2 and the cosine similarity measure.

Question 2) Consider the following set of vectors:

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ h \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ k \end{bmatrix} \right\}$$

Based on different values of h and k, discuss the number of independent vectors in this set.

[10 marks]

Let's make a matix using these vectors

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 2 & h & 1 \\ 1 & 1 & K \end{bmatrix}$$

I use Gaussiam elimination to convert this matrix into a row echelon form (upper triangular here as A is a square matrix).

$$\begin{bmatrix} 1 & 1 & 0 \\ 2 & h & 1 \\ 1 & 1 & K \end{bmatrix} \xrightarrow[R_3 \leftarrow R_3 - R_1]{R_2 \leftarrow R_2 - 2R_1} \begin{bmatrix} 1 & 1 & 0 \\ 0 & h-2 & 1 \\ 0 & 0 & K \end{bmatrix}$$

This matrix is upper triangular matrix.

if $K=0 \Rightarrow \text{rank}(A) = 2$ there are two independent columns $\Rightarrow$ there two independent vectors.

$$h \neq 2 \longrightarrow \begin{cases} K \neq 0 & \text{rank}(A) = 3 \\ K = 0 & \text{rank}(A) = 2 \end{cases}$$

$$h = 2 \longrightarrow \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & K \end{bmatrix} \xrightarrow{R_3 \leftarrow R_3 - KR_2} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\longrightarrow \text{rank}(A) = 2$$

Conclusion:

If $h = 2 \Longrightarrow \text{rank}(A) = 2$

If $h \neq 2$ and $K \neq 0 \Longrightarrow \text{rank}(A) = 3$

If $h \neq 2$ and $K = 0 \Longrightarrow \text{rank}(A) = 2$

If $K = 0 \longrightarrow \text{rank}(A) = 2$

Therefore, if $h \neq 2$ and $K \neq 0$, there are three independent vectors. Otherwise, there are two independent vectors.

Question 3) You are tasked with uncovering information about an incomplete matrix, some of whose entries are unknown and denoted as $a, b, c,$ and d :

$$A = \begin{bmatrix} -1 & 0 & a \\ b & 4 & c \\ d & 0 & 0 \end{bmatrix}$$

Given that $\text{rank}(A) = 2$, how many distinct eigenvalues does A have? [**10 marks**]

Because $\text{rank}(A)=2$, then $\det(A)=0$.

Let's find $\det(A)$ by expanding using the 3^{rd} row

$$\det(A) = (-1)^{3+1}(d)\det\begin{pmatrix} 0 & a \\ 4 & c \end{pmatrix} + (-1)^{3+2}(0)\det\begin{pmatrix} -1 & a \\ b & c \end{pmatrix}$$

$$+(-1)^{3+3}(0)\det\begin{pmatrix} -1 & 0 \\ b & 4 \end{pmatrix}$$

$$= d\left((0)(c)-4a\right) = -4ad = 0 \begin{cases} d=0 \\ \text{or} \\ a=0 \end{cases}$$

If $d=0$, $\det(A-\lambda I) = \det\left(\begin{bmatrix} -1-\lambda & 0 & a \\ b & 4-\lambda & c \\ 0 & 0 & -\lambda \end{bmatrix}\right)$

$$= (-1)^{3+3}(-\lambda)\det\begin{pmatrix} -1-\lambda & 0 \\ b & 4-\lambda \end{pmatrix}$$

$$= -\lambda\left((-1-\lambda)(4-\lambda) - 0\right)$$

$$= -\lambda\left(\lambda^2 - 4\lambda + \lambda - 4\right)$$

$$= -\lambda(\lambda^2 - 3\lambda - 4) = 0 \rightarrow \begin{cases} \lambda = 0 \\ \lambda = 4 \\ \lambda = -1 \end{cases}$$

There are three distinct eigenvalues.

If $a = 0 \Rightarrow \det(A - \lambda I) = \det\left(\begin{bmatrix} -1-\lambda & 0 & 0 \\ b & 4-\lambda & c \\ d & 0 & -\lambda \end{bmatrix}\right)$

$$= (-1)^{1+1}(-1-\lambda)\det\begin{pmatrix} 4-\lambda & c \\ 0 & -\lambda \end{pmatrix}$$

$$= (-1-\lambda)(-4\lambda + \lambda^2) = 0 \Rightarrow \begin{cases} \lambda = -1 \\ \lambda = 0 \\ \lambda = +4 \end{cases}$$

There are three distinct eigenvalues

If $a = 0$ and $d = 0 \Rightarrow$

$$\det(A - \lambda I) = \det\left(\begin{bmatrix} -1-\lambda & 0 & 0 \\ b & 4-\lambda & c \\ 0 & 0 & -\lambda \end{bmatrix}\right)$$

$$= (-1-\lambda)(-4\lambda + \lambda^2) = 0 \rightarrow \begin{cases} \lambda = -1 \\ \lambda = 0 \\ \lambda = 4 \end{cases}$$

again there are 3 distinct eigenvalues.

Question 4) Consider the following matrix:

$$A = \begin{bmatrix} 3 & -3 & 0 \\ 3 & -1 & 2 \\ b & 0 & 2 \end{bmatrix}$$

Find the values for b (if possible) so that:

(i) The determinant of A is 4.

(ii) The rank of A is 2.

(iii) $\frac{1}{2}$ is an eigenvalue of A^{-1}.

(iv) The system $A\boldsymbol{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ has no solutions.

(v) The system $A\boldsymbol{x} = \begin{bmatrix} -3 \\ 1 \\ 2 \end{bmatrix}$ has infinitely many solutions.

(i)

Let's convert A into row echelon form:

$$\begin{bmatrix} 3 & -3 & 0 \\ 3 & -1 & 2 \\ b & 0 & 2 \end{bmatrix} \xrightarrow[R_3 \leftarrow R_3 - \frac{b}{3}R_1]{R_2 \leftarrow R_2 - R_1} \begin{bmatrix} 3 & -3 & 0 \\ 0 & 2 & 2 \\ 0 & b & 2 \end{bmatrix} \xrightarrow{R_3 \leftarrow R_3 - \frac{b}{2}R_2}$$

$$\begin{bmatrix} 3 & -3 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 2-b \end{bmatrix}$$

$$\det(A) = (3)(2)(2-b) = 12-6b = 4 \Rightarrow 6b = 8 \rightarrow b = \frac{8}{6} = \frac{4}{3}$$

Or, you can use the definition of determinants:

$$\det(A) = 3\det\begin{pmatrix} -1 & 2 \\ 0 & 2 \end{pmatrix} + 3\det\begin{pmatrix} 3 & 2 \\ b & 2 \end{pmatrix}$$

$$= 3(-2) + 3(6-2b) = -6 + 18 - 6b$$

$$= 12 - 6b = 4 \Rightarrow \boxed{b = 4/3.}$$

(ii) We can use the row echelon form of A we obtained in part (i) to answer this question.

$$\begin{bmatrix} 3 & -3 & 0 \\ 0 & 2 & 2 \\ 0 & 0 & 2-b \end{bmatrix}$$

When $2-b=0$ or $\boxed{b=2}$, the rank of A is 2.

(iii) If λ is an eigenvalue of A, then $\frac{1}{\lambda}$ is an eigenvalue of A^{-1}. If we want to have $\frac{1}{2}$ as an eigenvalue of A^{-1}, it is equivalent to have 2 as an eigenvalue of A. Please be aware that we suppose A^{-1} exists, so all eigenvalues are nonzero. Therefore, we need to check two items:

(I) $\det(A) \neq 0$

(II) 2 is an eigenvalue of A.

We have seen that $\det(A) = 12 - 6b$.

We need $12 - 6b \neq 0$ or $\boxed{b \neq 2}$

$$A - \lambda I = \begin{bmatrix} 3-\lambda & -3 & 0 \\ 3 & -1-\lambda & 2 \\ b & 0 & 2-\lambda \end{bmatrix}$$

$$\det(A - \lambda I) = (3-\lambda)\left((-1-\lambda)(2-\lambda) - (2)(0)\right) - (-3)\left(3(2-\lambda) - 2b\right)$$

$$= (3-\lambda)(-1-\lambda)(2-\lambda) \quad +3(6-3\lambda-2b)$$

$$= (3-\lambda)(-1-\lambda)(2-\lambda) + 18 - 9\lambda - 6b$$

if 2 is an eigenvalue, it should make this determinant zero. so,

$$(3-2)(-1-2)(2-2) + 18 - 9(2) - 6b = 0$$

$$-6b = 0$$

$$\boxed{b = 0}$$

So, if $b = 0$, A^{-1} exists, and $\frac{1}{2}$ is an eigenvalue of it.

(iv) $A\vec{x} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ has no solution.

The augmented matrix

$$\left[\begin{array}{ccc|c} 3 & -3 & 0 & 1 \\ 3 & -1 & 2 & 0 \\ b & 0 & 2 & 0 \end{array}\right] \xrightarrow[R_3 \leftarrow R_3 - \frac{b}{3}R_1]{R_2 \leftarrow R_2 - R_1} \left[\begin{array}{ccc|c} 3 & -3 & 0 & 1 \\ 0 & 2 & 2 & -1 \\ 0 & b & 2 & -\frac{b}{3} \end{array}\right]$$

$$\xrightarrow{R_3 \leftarrow R_3 - \frac{b}{2}R_2} \left[\begin{array}{ccc|c} 3 & -3 & 0 & 1 \\ 0 & 2 & 2 & -1 \\ 0 & 0 & 2-b & -\frac{b}{3} + \frac{b}{2} \end{array}\right] = \left[\begin{array}{ccc|c} 3 & -3 & 0 & 1 \\ 0 & 2 & 2 & -1 \\ 0 & 0 & 2-b & \frac{b}{6} \end{array}\right]$$

To have no solution, we need $2-b=0$ and $\frac{b}{6} \neq 0$, which implies $\boxed{b=2.}$

(V) $Ax = \begin{bmatrix} -3 \\ 1 \\ 2 \end{bmatrix}$ has many solutions

The augmented matrix

$$\left[\begin{array}{ccc|c} 3 & -3 & 0 & -3 \\ 3 & -1 & 2 & 1 \\ b & 0 & 2 & 2 \end{array}\right] \xrightarrow[R_3 \leftarrow R_3 - \frac{b}{3}R_1]{R_2 \leftarrow R_2 - R_1} \left[\begin{array}{ccc|c} 3 & -3 & 0 & -3 \\ 0 & 2 & 2 & 4 \\ 0 & b & 2 & 2+b \end{array}\right]$$

$$\xrightarrow{R_3 \leftarrow R_3 - \frac{b}{2}R_2} \left[\begin{array}{ccc|c} 3 & -3 & 0 & -3 \\ 0 & 2 & 2 & 4 \\ 0 & 0 & 2-b & 2-b \end{array}\right]$$

When $\boxed{2-b=0}$, we have many solutions.

Now consider

$$A = \begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix}$$

for questions 5, 6, and 7.

Question 5) Find $S = A^T A$ matrix and

(a) Find the characteristic polynomial of S

(b) Find the eigenvalues of S, and call them λ_1 and λ_2 where $\lambda_1 \geq \lambda_2$.

(c) Find the eigenvectors S.

(d) Are the eigenvectors orthonormal? If they are not, convert them into orthonormal vectors using the Gram-Schmidt process. Call them $\boldsymbol{v_1}, \boldsymbol{v_2}$ after orthonormalisation.

(e) Using the eigenvalues of S, $\lambda_1 \geq \lambda_2$, set $\sigma_1 = \sqrt{\lambda_1}$ and $\sigma_2 = \sqrt{\lambda_2}$ and make this matrix

$$D = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix}$$

We will use this matrix later.
Make a matrix V using the orthonormal vectors you obtained from the previous part. The ordering of the columns of V should be the same as the ordering of the eigenvalues, that is $V = \begin{bmatrix} \boldsymbol{v_1} & \boldsymbol{v_2} \end{bmatrix}$. Show that V is an orthogonal matrix.

[3 + 3 + 3 + 3 + 3 = 15 marks]

$$S = A^T A = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 11 & 1 \\ 1 & 11 \end{bmatrix}$$

(a)

$$\det(S - \lambda I) = \det\begin{pmatrix} 11-\lambda & 1 \\ 1 & 11-\lambda \end{pmatrix} = (11-\lambda)^2 - 1 = \lambda^2 - 22\lambda + 120$$

(b) To find eigenvalues, put $\det(S - \lambda I) = 0$

$$\Rightarrow \lambda^2 - 22\lambda + 120 \rightarrow \lambda = \frac{22 \pm \sqrt{(-22)^2 - 4(1)(120)}}{2}$$

$$= \frac{22 \pm 2}{2} \Rightarrow \begin{cases} \lambda_1 = 12 \\ \lambda_2 = 10 \end{cases}$$

(c) To find eigenvectors, for each eigenvalue λ, we need to solve the system $(S - \lambda I)\vec{v} = \vec{0}$.

$\lambda_1 = 12 \longrightarrow S - 12I = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$

$(S - 12I)\vec{x_1} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \longrightarrow \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

We use Gaussian elimination to solve this.

$\left[\begin{array}{cc:c} -1 & 1 & 0 \\ 1 & -1 & 0 \end{array}\right] \xrightarrow{R_2 \leftarrow R_2 + R_1} \left[\begin{array}{cc:c} -1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right]$

$\Rightarrow x_2$ is a free variable. $\boxed{x_2 = t \in \mathbb{R}}$

using the first row, $-x_1 + x_2 = 0$

$-x_1 = -t \rightarrow \boxed{x_1 = t}$

$\vec{x_1} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} t \\ t \end{bmatrix} = t\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad t \in \mathbb{R}$

set $t = 1$. $\vec{x_1} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

$\lambda_1 = 10 \longrightarrow S - 12I = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

$(S - 12I)\vec{x}_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

We use Gaussian elimination to solve this.

$$\left[\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & 1 & 0 \end{array}\right] \xrightarrow{R_2 \leftarrow R_2 - R_1} \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

$\Rightarrow x_2$ is a free variable. $\boxed{x_2 = t \in \mathbb{R}}$

using the first row, $x_1 + x_2 = 0$

$x_1 = -t \longrightarrow \boxed{x_1 = -t}$

$$\vec{x}_2 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -t \\ t \end{bmatrix} = t\begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad t \in \mathbb{R}$$

set $t = 1$. $\vec{x}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

The eigenvectors of S are

$\lambda_1 = 12 \longrightarrow \vec{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$\lambda_2 = 10 \longrightarrow \vec{x}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$

(d) $\vec{x_1} \cdot \vec{x_2} = (1)(-1) + (1)(1) = 0$

$\Rightarrow \quad \vec{x_1} \perp \vec{x_2}$.

So, we do not need to orthognalize them. only we need to normalize them.

$$\hat{x}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \vec{v_1}$$

$$\hat{x}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \vec{v_2}$$

(e) $\sigma_1 = \sqrt{\lambda_1} = \sqrt{12} \qquad \sigma_2 = \sqrt{\lambda_2} = \sqrt{10}$

$$D = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \sqrt{12} & 0 \\ 0 & \sqrt{10} \\ 0 & 0 \end{bmatrix}$$

To show that V is an orthogonal matrix, we need to show that $VV^T = V^TV = I$.

$$V = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \qquad V^T = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -1/\sqrt{2} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$VV^T = \begin{bmatrix} \frac{1}{2} + \frac{1}{2} & \frac{1}{2} - 1/2 \\ \frac{1}{2} - 1/2 & 1/2 + 1/2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

Also, we need to show $V^T V = I$.

$$V^T V = \begin{bmatrix} 1/\sqrt{2} & \frac{1}{\sqrt{2}} \\ -1/\sqrt{2} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}+\frac{1}{2} & \frac{1}{2}-\frac{1}{2} \\ \frac{1}{2}-\frac{1}{2} & \frac{1}{2}+\frac{1}{2} \end{bmatrix} = I$$

$\Rightarrow V$ is an orthogonal matrix.

Question 6) Find $T = AA^T$ matrix and

(a) Find the characteristic polynomial of T

(b) Find the eigenvalues of T, and order them as $\lambda_1 \geq \lambda_2 \geq \lambda_3$

(c) Find the eigenvectors T.

(d) Are the eigenvectors orthonormal? If they are not, convert them into orthonormal vectors using the Gram-Schmidt process. Call them $\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}$ according to the order of their corresponding eigenvalues.

[3+3+3+3 = 12 marks]

$$T = \begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 0 & 2 \\ 0 & 10 & 4 \\ 2 & 4 & 2 \end{bmatrix}$$

ⓐ characteristic polynomial $= \det(T - \lambda I)$

$$T - \lambda I = \begin{bmatrix} 10-\lambda & 0 & 2 \\ 0 & 10-\lambda & 4 \\ 2 & 4 & 2-\lambda \end{bmatrix}$$

$$\det(T - \lambda I) = (10-\lambda)\det\begin{pmatrix} 10-\lambda & 4 \\ 4 & 2-\lambda \end{pmatrix} + 2\det\begin{pmatrix} 0 & 10-\lambda \\ 2 & 4 \end{pmatrix}$$

$$= (10-\lambda)\left((10-\lambda)(2-\lambda) - 16\right) + 2\left(-2(10-\lambda)\right)$$

$$= (10-\lambda)\left(20 - 10\lambda - 2\lambda + \lambda^2 - 16\right) - 4(10-\lambda)$$

$$= (10-\lambda)\left[\lambda^2 - 12\lambda + 4 - 4\right]$$

$$= (10-\lambda)\left[\lambda^2 - 12\lambda\right]$$

(b) To find eigenvalues, put $\det(T-\lambda I)=0$

$$\Rightarrow (10-\lambda)(\lambda)(\lambda-12)=0 \Rightarrow \begin{cases} \lambda_1 = 12 \\ \lambda_2 = 10 \\ \lambda_3 = 0 \end{cases}$$

(c) To find eigenvectors, we solve the system $(T-\lambda I)\vec{v}_1 = \vec{0}$ for each eigenvalue.

$$\lambda_1 = 12 \Rightarrow T-12I = \begin{bmatrix} -2 & 0 & 2 \\ 0 & -2 & 4 \\ 2 & 4 & -10 \end{bmatrix}$$

$$(T-\lambda I)\vec{v}_1 = 0 \Rightarrow \begin{bmatrix} -2 & 0 & 2 \\ 0 & -2 & 4 \\ 2 & 4 & -10 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\rightarrow \left[\begin{array}{ccc:c} -2 & 0 & 2 & 0 \\ 0 & -2 & 4 & 0 \\ 2 & 4 & -10 & 0 \end{array}\right] \xrightarrow{R_3 \leftarrow R_3 + R_1} \left[\begin{array}{ccc:c} -2 & 0 & 2 & 0 \\ 0 & -2 & 4 & 0 \\ 0 & 4 & -8 & 0 \end{array}\right]$$

$$\xrightarrow{R_3 \leftarrow R_3 + 2R_2} \left[\begin{array}{ccc:c} -2 & 0 & 2 & 0 \\ 0 & -2 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

v_3 is a free variable $\Rightarrow \boxed{v_3 = t \in \mathbb{R}}$

using 2^{nd} row: $-2v_2 + 4v_3 = 0 \Rightarrow -2v_2 = -4t$
$\rightarrow \boxed{v_2 = 2t}$

using 1^{st} row: $-2v_1 + 2v_3 = 0 \rightarrow -2v_1 = -2t$

$\rightarrow \boxed{v_1 = t}$

$$\vec{v_1} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} t \\ 2t \\ t \end{bmatrix} = t \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \quad t \in \mathbb{R}.$$

put $t=1 \Rightarrow \vec{v_1} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$.

$\lambda = 10 \rightarrow \quad T - 10I = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 4 \\ 2 & 4 & -8 \end{bmatrix}$

we need to solve $(T-10I)\vec{v_2} = \vec{0}$

$$\Rightarrow \begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 4 \\ 2 & 4 & -8 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

using Gaussian elimination:

$$\left[\begin{array}{ccc|c} 0 & 0 & 2 & 0 \\ 0 & 0 & 4 & 0 \\ 2 & 4 & -8 & 0 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_3} \left[\begin{array}{ccc|c} 2 & 4 & -8 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 2 & 0 \end{array}\right]$$

$$\xrightarrow{R_3 \leftarrow R_3 - \frac{1}{2}R_2} \left[\begin{array}{ccc|c} 2 & 4 & -8 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

v_2 is a free variable $\Rightarrow v_2 = t \quad t \in \mathbb{R}$

using 2nd row, $4v_3 = 0 \longrightarrow v_3 = 0$

using 1^{st} row, $2v_1 + 4v_2 - 8v_3 = 0$

$2v_1 + 4t = 0 \rightarrow 2v_1 = -4t$

$\rightarrow v_1 = -2t$

$$\vec{v_2} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} -2t \\ t \\ 0 \end{bmatrix} = t\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \quad t \in \mathbb{R}$$

put $t=1 \Rightarrow \vec{v_2} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$

$\lambda_3 = 0 \Rightarrow T - (0)I = T$

$$(T-(0)I)\vec{v_3} = \vec{0} \rightarrow \begin{bmatrix} 10 & 0 & 2 \\ 0 & 10 & 4 \\ 2 & 4 & 2 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

using Gaussian elimination

$$\left[\begin{array}{ccc|c} 10 & 0 & 2 & 0 \\ 0 & 10 & 4 & 0 \\ 2 & 4 & 2 & 0 \end{array}\right] \xrightarrow{R_3 \leftarrow R_3 - \frac{1}{5}R_1} \left[\begin{array}{ccc|c} 10 & 0 & 2 & 0 \\ 0 & 10 & 4 & 0 \\ 0 & 4 & \frac{8}{5} & 0 \end{array}\right]$$

$$\xrightarrow{R_3 \leftarrow R_3 - \frac{4}{10}R_2} \left[\begin{array}{ccc|c} 10 & 0 & 2 & 0 \\ 0 & 10 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

v_3 is a free variable $\Rightarrow v_3 = t \quad t \in \mathbb{R}$

using 2^{nd} row $\rightarrow$ $10v_2 + 4v_3 = 0$

$10v_2 = -4t \rightarrow v_2 = -\frac{2}{5}t$

using 1^{st} row $\rightarrow$ $10v_1 + 2v_3 = 0$

$10v_1 = -2t \rightarrow v_1 = -\frac{1}{5}t$

$$\vec{v_3} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} -\frac{1}{5}t \\ -\frac{2}{5}t \\ t \end{bmatrix} = t\begin{bmatrix} -\frac{1}{5} \\ -\frac{2}{5} \\ 1 \end{bmatrix} \quad t \in \mathbb{R}$$

put $t = -5$ $\rightarrow$ $\vec{v_3} = \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix}$

$\lambda_1 = 12 \longrightarrow \vec{v_1} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$

$\lambda_2 = 10 \longrightarrow \vec{v_2} = \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$

$\lambda_3 = 0 \longrightarrow \vec{v_3} = \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix}$

(d) $\vec{v_1} \cdot \vec{v_2} = (1)(-2) + (2)(1) + (1)(0) = 0 \Rightarrow \vec{v_1} \perp \vec{v_2}$

$\vec{v_1} \cdot \vec{v_3} = (1)(1) + (2)(2) + (1)(-5) = 0 \longrightarrow \vec{v_1} \perp \vec{v_3}$

$\vec{v_2} \cdot \vec{v_3} = (-2)(1) + (1)(2) + (0)(-5) = 0 \Rightarrow \vec{v_2} \perp \vec{v_3}$

These vectors are mutually orthogonal. So, we don't need to orthognalize them. We only need to normalize them.

$$\hat{v}_1 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix} = \vec{x_1}$$

$$\vec{v_2} = \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{bmatrix} = \vec{x_2}$$

$$\vec{v_3} = \frac{1}{\sqrt{30}} \begin{bmatrix} 1 \\ 2 \\ -5 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{30} \\ 2/\sqrt{30} \\ -5/\sqrt{30} \end{bmatrix} = \vec{x_3}$$

Question 7) Consider two orthonormal vectors you obtained in question 6.

(a) Using them we want to make three orthonormal vectors $\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}$ such that

$$\boldsymbol{u_1} = \frac{1}{\sigma_1} A\boldsymbol{v_1}$$

$$\boldsymbol{u_2} = \frac{1}{\sigma_2} A\boldsymbol{v_2}$$

Also, find the third vector $\boldsymbol{u_3}$,

$$\boldsymbol{u_3} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

such that $\boldsymbol{u_3}$ is a unit vector that is orthogonal to $\boldsymbol{u_1}$ and $\boldsymbol{u_2}$. In other words

$$||\boldsymbol{u_3}||_2 = 1 \text{ and } \boldsymbol{u_3} \perp \boldsymbol{u_1} \text{ and } \boldsymbol{u_3} \perp \boldsymbol{u_2}.$$

Put these three vectors as columns in a matrix $U = \begin{bmatrix} \boldsymbol{u_1} & \boldsymbol{u_2} & \boldsymbol{u_3} \end{bmatrix}$.

(b) Show that U is an orthogonal matrix.

(c) Compute UDV^T.

(d) Explain the relationship between $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}\}$ and $\{\boldsymbol{x_1}, \boldsymbol{x_2}, \boldsymbol{x_3}\}$.

[5+2+3+3= 13 marks]

$$\vec{v_1} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \qquad \sigma_1 = \sqrt{\lambda_1} = \sqrt{12} \qquad \sigma_2 = \sqrt{\lambda_2} = \sqrt{10}$$

$$\vec{v_2} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

$$\vec{u_1} = \frac{1}{\sqrt{12}} \begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \frac{1}{\sqrt{12}} \begin{bmatrix} \frac{3}{\sqrt{2}} - \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} + \frac{3}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$= \frac{1}{\sqrt{12}} \begin{bmatrix} 2/\sqrt{2} \\ 4/\sqrt{2} \\ 2/\sqrt{2} \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{24}} \\ 4/\sqrt{24} \\ 2/\sqrt{24} \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} 2/\sqrt{4} \\ 4/\sqrt{4} \\ 2/\sqrt{4} \end{bmatrix}$$

$$= \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

$$\vec{u}_2 = \frac{1}{\sqrt{10}} \begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \frac{1}{\sqrt{10}} \begin{bmatrix} -\frac{3}{\sqrt{2}} - \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} + \frac{3}{\sqrt{2}} \\ -1/\sqrt{2} + \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$= \frac{1}{\sqrt{10}} \begin{bmatrix} -4/\sqrt{2} \\ 2/\sqrt{2} \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{4}{\sqrt{20}} \\ 2/\sqrt{20} \\ 0 \end{bmatrix} = \frac{1}{\sqrt{5}} \begin{bmatrix} -\frac{4}{\sqrt{4}} \\ 2/\sqrt{4} \\ 0 \end{bmatrix}$$

$$= \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$$

$$\vec{u}_3 = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \qquad \vec{u}_3 \perp \vec{u}_1 \Rightarrow \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0$$

$$\Rightarrow \boxed{a + 2b + c = 0 \quad \text{I}}$$

$$\vec{u}_3 \perp \vec{u}_2 \Rightarrow \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = 0$$

$$\Rightarrow \boxed{-2a + b = 0 \quad \text{II}}$$

$$\|\vec{u}_3\|_2 = 1 \Rightarrow \sqrt{a^2 + b^2 + c^2} = 1$$

$$\Rightarrow \boxed{a^2 + b^2 + c^2 = 1 \quad \text{III}}$$

To find $\|\vec{u}_3\|$, we need to solve this nonlinear system:

$$\begin{cases} a+2b+c=0 \\ -2a+b=0 \\ a^2+b^2+c^2=1 \end{cases}$$

$-2a+b=0 \rightarrow \boxed{b=2a}$ *

plug $b=2a$ in the 1^{st} equation:

$$a+4a+c=0 \longrightarrow \boxed{c=-5a} \quad **$$

plug * and ** in the 3^{rd} equation:

$$a^2+4a^2+25a^2=1 \Rightarrow 30a^2=1$$

$$\rightarrow a^2=\frac{1}{30} \Rightarrow a=\pm\frac{1}{\sqrt{30}}$$

when $a=\frac{1}{\sqrt{30}} \rightarrow b=\frac{2}{\sqrt{30}}$ and $c=\frac{-5}{\sqrt{30}}$

$$\Rightarrow \vec{u}_3=\begin{bmatrix} 1/\sqrt{30} \\ 2/\sqrt{30} \\ -5/\sqrt{30} \end{bmatrix}$$

when $a=-\frac{1}{\sqrt{30}} \rightarrow b=-\frac{2}{\sqrt{30}}$, $c=\frac{+5}{\sqrt{30}} \rightarrow \vec{u}_3=\begin{bmatrix} -1/\sqrt{30} \\ \frac{-2}{\sqrt{30}} \\ 5/\sqrt{30} \end{bmatrix}$

We have two options for $\vec{u}_3$, and they are parallel vectors. Let's cosider only the first one. Choosing the other one will not change the results.

$$\vec{u}_1 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

$$\vec{u}_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}$$

$$\vec{u}_3 = \begin{bmatrix} 1/\sqrt{30} \\ 2/\sqrt{30} \\ -5/\sqrt{30} \end{bmatrix}$$

*please compare $\vec{u}_1$, $\vec{u}_2$, and $\vec{u}_3$ with $\vec{x}_1$, $\vec{x}_2$, and $\vec{x}_3$. We obtained them through completly different procedures.

$$U = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{30}} \\ \frac{1}{\sqrt{6}} & 0 & \frac{-5}{\sqrt{30}} \end{bmatrix} \qquad U^T = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{30}} & -\frac{5}{\sqrt{30}} \end{bmatrix}$$

$$UU^T = \begin{bmatrix} \frac{1}{6}+\frac{4}{5}+\frac{1}{30} & \frac{2}{6}-\frac{2}{5}+\frac{2}{30} & \frac{1}{6}-\frac{5}{30} \\ \frac{2}{6}-\frac{2}{5}+\frac{2}{30} & \frac{4}{6}+\frac{1}{5}+\frac{4}{30} & \frac{2}{6}-\frac{10}{30} \\ \frac{1}{6}-\frac{5}{30} & \frac{2}{6}-\frac{10}{30} & \frac{1}{6}+\frac{25}{30} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I$$

$$U^TU = \begin{bmatrix} \frac{1}{6}+\frac{4}{5}+\frac{1}{30} & \frac{2}{6}-\frac{2}{5}+\frac{2}{30} & \frac{1}{6}-\frac{5}{30} \\ \frac{2}{6}-\frac{2}{5}+\frac{2}{30} & \frac{4}{6}+\frac{1}{5}+\frac{4}{30} & \frac{2}{6}-\frac{10}{30} \\ \frac{1}{6}-\frac{5}{30} & \frac{2}{6}-\frac{10}{30} & \frac{1}{6}+\frac{25}{30} \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I$$

U is an orthogonal matrix, as $UU^T = U^TU = I$.

(c)

$$U = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{30}} \\ \frac{1}{\sqrt{6}} & 0 & \frac{-5}{\sqrt{30}} \end{bmatrix} \quad V = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

$$D = \begin{bmatrix} \sqrt{12} & 0 \\ 0 & \sqrt{10} \\ 0 & 0 \end{bmatrix}$$

$$UDV^T = \underset{3\times 3}{\begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{30}} \\ \frac{1}{\sqrt{6}} & 0 & \frac{-5}{\sqrt{30}} \end{bmatrix}} \left(\underset{3\times 2}{\begin{bmatrix} \sqrt{12} & 0 \\ 0 & \sqrt{10} \\ 0 & 0 \end{bmatrix}} \underset{2\times 2}{\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}} \right)$$

$$= \underset{3\times 3}{\begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{30}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{30}} \\ \frac{1}{\sqrt{6}} & 0 & \frac{-5}{\sqrt{30}} \end{bmatrix}} \underset{3\times 2}{\begin{bmatrix} \sqrt{6} & \sqrt{6} \\ -\sqrt{5} & \sqrt{5} \\ 0 & 0 \end{bmatrix}}$$

$$= \begin{bmatrix} 1+2+0 & 1-2+0 \\ 2-1+0 & 2+1+0 \\ 1+0+0 & 1+0+0 \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix} = A$$

(d)

What do we have so far?

$\{\vec{v_1}, \vec{v_2}\}$ eigenvectors of A^TA.

$$A^TA\vec{v_1} = \lambda_1 \vec{v_1} \qquad A^TA\vec{v_2} = \lambda_2 \vec{v_2}$$

$\{\vec{x_1}, \vec{x_2}, \vec{x_3}\}$ eigenvectors of AA^T.

$$AA^T\vec{x_1} = \lambda_1\vec{x_1} \qquad AA^T\vec{x_2} = \lambda_2 \vec{x_2} \qquad AA^T\vec{x_3} = \lambda_3 \vec{x_3}$$

Also, about $\{\vec{u_1}, \vec{u_2}, \vec{u_3}\}$

$$\vec{u_1} = \frac{1}{\sigma_1} A\vec{v_1} \quad \text{and} \quad \vec{u_2} = \frac{1}{\sigma_2} A\vec{v_2}.$$

$$\vec{u_1} = \frac{1}{\sigma_1} A\vec{v_1} \Longrightarrow A^T\vec{u_1} = \frac{1}{\sigma_1}\left(A^TA\vec{v_1}\right)$$

$$= \frac{1}{\sigma_1}\left(\lambda_1 \vec{v_1}\right) = \frac{\lambda_1}{\sigma_1}\vec{v_1}$$

$$A^T\vec{u_1} = \frac{\lambda_1}{\sigma_1}\vec{v_1} \Rightarrow AA^T\vec{u_1} = \frac{\lambda_1}{\sigma_1}(A\vec{v_1})$$

$\Rightarrow AA^T\vec{u_1} = \lambda_1 \vec{u_1}$

$\Rightarrow \vec{u_1}$ is an eigenvalue of the matrix AA^T corresponding to λ_1. Therefore, it is equal or parallel to $\vec{x}_1$.

Also similarly, we can show that $\vec{u}_2$ is an eigenvalue of AA^T corresponding to λ_2, therefore, it is equal or parallel to $\vec{x}_2$.

$\vec{u}_3 \perp \vec{u}_1$, $\vec{u}_3 \perp \vec{u}_2$, and at same time $\vec{x}_3 \perp \vec{x}_1$, $\vec{x}_3 \perp \vec{x}_1$, therefore, as it is a 3-dimensional space, $\vec{x}_3$ and $\vec{u}_3$ should be equal or parallel to each other.

SIT787: Mathematics for AI
Assessment 2 - Further Learning
Singular Value Decomposition (SVD)

Asef Nazari

There is a famous saying that "The SVD is the Swiss Army knife of matrix decompositions". In this short tutorial, we will see what it is and how we can use it.

We learnt that for a square matrix $A_{n\times n}$ we can define eigenvalues and eigrnvectors. A nonzero vector $\mathbf{x}$ is called an eigenvector of A if $A\mathbf{x} = \lambda\mathbf{x}$. However, it is impossible to have eigenvalues and eigenvectors for rectangular matrices. Even for square matrices, eigenvalues and eigenvectors may have some issues:

- The eigenvalues may be complex
- The eigenvectors may not be orthogonal
- Even there are **deficit matrices** that do not have a complete set of eigenvectors

We considered the symmetric matrices that are the best well-behaved matrices. If $S = S^T$, then all n eigenvalues are real numbers and n eigenvectors $\mathbf{q}$ can be chosen orthogonal. In addition, we can decompose or diagonalise $S = Q\Lambda Q^T$ where

$$Q = \begin{bmatrix}\mathbf{q}_1, \dots, \mathbf{q}_n\end{bmatrix} \text{ and } \Lambda = \text{diag}\{\lambda_1, \dots, \lambda_n\}$$

However, this is not the case for all matrices. So, we need to construct a setting that it works for all matrices, in particular, rectangular matrices. Singular value decomposition (SVD) fills this gap. So, let's consider a recatngular matrix $A_{m\times n}$. We make two other matrices using this matrix: A^TA and AA^T. Now, we check what properties these two matrices have:

- A^TA is a $n\times n$ and AA^T is a $m \times m$ matrices. Therefore, trhey are square matrices.
- They are symmetric matrices. $(AA^T)^T = (A^T)^TA^T = AA^T$ and $(A^TA)^T = A^T(A^T)^T = A^TA$. You know, and I know how important it is for a matrix to be symmetric!
- A^TA and AA^T are positive semi-definite. A symmetric square matrix S is called positive semidefnite if
  - S has all positive eigenvalues, or
  - for all nonzero vectors $\mathbf{x}$, $\mathbf{x}^TS\mathbf{x} \geq 0$, or
  - All the leading determinants are non-negative.
- A^TA has the same positive eigenvalues as AA^T

AA^T is a symmetric $m \times m$ matrix. It has m real eigenvalues, and we can choose the m eignvectors to be orthonormal (orthogonal and unit vector). Let's call the eigenvectors $\mathbf{u}_i$.

$$AA^T\mathbf{u}_i = \lambda_i\mathbf{u}_i, \text{ for } i = 1, \dots, m$$

A^TA is a symmetric $n \times n$ matrix. It has n real eigenvalues, and we can choose the n eignvectors to be orthonormal (orthogonal and unit vector). Let's call the eigenvectors $\mathbf{v}_i$.

$$A^TA\mathbf{v}_i = \lambda_i \mathbf{v}_i, \text{ for } i = 1, \dots, n$$

Also, only r of the λ_i's are postive, $r = \text{rank}(A)$, and $r \leq \min\{m, n\}$. Let's define $\sigma_i = \sqrt{\lambda_i}$. Therefore,

$$AA^T\mathbf{u}_i = \sigma_i^2\mathbf{u}_i, \text{ and } A^TA\mathbf{v}_i = \sigma_i^2\mathbf{v}_i$$

The $\mathbf{u}$'s are called left singular vectors of A and the $\mathbf{v}$'s are the right singular vectors of A. The σ's are singular values. The connection between $\mathbf{u}$'s and $\mathbf{v}$'s is

$$A\mathbf{v}_i = \sigma\mathbf{u_i}$$

Now we are ready to build our matrices U with columns of normalised eigenvectors of AA^T, V with columns of normalised eigenvectors of A^TA, and Σ consisiting of ordered singular values

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$$

$\mathbf{u}_i$ and $\mathbf{v}_i$ are corresponding to the singukar value σ_i, and they should appear in their matrices following the order of σ_i's.

$$U_{m\times m} = \begin{bmatrix}\mathbf{u}_1 & \ldots & \mathbf{u}_m\end{bmatrix}, V_{n\times n} = \begin{bmatrix}\mathbf{v}_1 & \ldots & \mathbf{v}_n\end{bmatrix} \text{ and } \Sigma_{m\times n} = \begin{bmatrix}\sigma_1 & 0 & 0 & \ldots & 0\\ 0 & \sigma_2 & 0 & \ldots & 0\\ \vdots & \vdots & \vdots & \ldots & \vdots\\ 0 & 0 & \sigma_r & \ldots & 0\\ 0 & 0 & 0 & \ldots & 0\end{bmatrix}$$

So we have the SVD decomposition

$$A = U\Sigma V^T \text{ or } AV = U\Sigma \text{ or } A^TU = V\Sigma^T$$

This decomposition helps us writing $A\mathbf{v}_i = \sigma\mathbf{u_i}$ and $A^T\mathbf{u}_i = \sigma\mathbf{v_i}$

Example: Find the SVD of the matrix

$$A = \begin{bmatrix}1 & 0 & 1 & 0\\ 0 & 1 & 0 & 1\end{bmatrix}$$

Solution: A is a 2×4 matrix with rank $r = 2$.

$$AA^T = \begin{bmatrix}2 & 0\\ 0 & 2\end{bmatrix} \text{ and } A^TA = \begin{bmatrix}1 & 0 & 1 & 0\\ 0 & 1 & 0 & 1\\ 1 & 0 & 1 & 0\\ 0 & 1 & 0 & 1\end{bmatrix}$$

Eigenvalues of AA^T is immidiately seen. They are $\lambda_1 = \lambda_2 = 2$. The eigenvectors of AA^T are

$$\mathbf{u}_1 = \begin{bmatrix}0\\ 1\end{bmatrix} \text{ and } \mathbf{u}_1 = \begin{bmatrix}1\\ 0\end{bmatrix}$$

They are already unit vectors.

Finding eigenvalues of A^TA may need some work! Its characteristic polynomial is $\det(A^TA - \lambda I) = \lambda^4 - 4\lambda^3 + 4\lambda^2$ with eigenvalues $2, 2, 0, 0$! Pay attention that technically you don't need to calculate the eigenvalues of A^TA. The nonzero eigenvalues of A^TA is the same as eigenvalues of AA^T. The eigenvectors of A^TA are

$$\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \end{bmatrix} \text{ and } \mathbf{v}_4 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

However, they are not unit vectors. We need to normalise them.

$$\mathbf{v}_1 = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} \text{ and } \mathbf{v}_4 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

Our singular values are $\sqrt{2}$ and $\sqrt{2}$.

To check $A\mathbf{v}_i = \sigma\mathbf{u_i}$ we have

$$\frac{1}{\sigma_1}A\mathbf{v}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \mathbf{u}_1$$

$$\frac{1}{\sigma_2}A\mathbf{v}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \mathbf{u}_2$$

Also, see that $A^T\mathbf{u}_i = \sigma\mathbf{v_i}$

$$\frac{1}{\sigma_1}A^T\mathbf{u}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \mathbf{v}_1$$

$$\frac{1}{\sigma_1}A^T\mathbf{u}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} = \mathbf{v}_2$$

Now we are ready to build our matrices U with columns of normalised eigenvectors of AA^T, V with columns of normalised eigenvectors of A^TA, and Σ consisiting of singular values:

$$U_{2\times 2} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \Sigma_{2\times 4} = \begin{bmatrix} \sqrt{2} & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \end{bmatrix} \text{ and } V_{4\times 4} = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}$$

Now, we can check $A = U\Sigma V^T$.

$$U\Sigma V^T = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} \sqrt{2} & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} = A$$

So, to find $\mathbf{u}_i$'s and $\mathbf{v}_i$'s, consider which one of A^TA and AA^T has smaller dimensions. Find its eigenvalues and eigenvectors. Uisng λ_i's compute the singular vatlues σ_i's. Then using either $A\mathbf{v}_i = \sigma\mathbf{u_i}$ or $A^T\mathbf{u}_i = \sigma\mathbf{v_i}$ find singular vectors.

Theorem: For every $A \in \mathbb{R}^{m\times n}$ there exists $m \times m$ othogonal U matrix and $n \times n$ orthogonal V matrix such that U^TAV ia an $m \times n$ diagonal matrix Σ that has values $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$ in its diagonal. This theorem means

- Every A has decomposition $A = U\Sigma V^T$ whihc is called the singular value decomposition
- The values σ_i are the singular values
- Columns of U are the left singular vectors of A and columns of V are the right singular vectors of A.

Every matrix $A = U\Sigma V^T$ has a pseudoinverse $A^\perp = V\Sigma^\perp U^T$. For the matrix Σ, to obtain the pseudoinverse $\Sigma^\perp$ put $\frac{1}{\sigma_i}$ instead of each nonzero singular value σ_i and then transpose the matrix.

$$A^\perp = \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$$

You can check that $AA^\perp A = A$ and $A^\perp AA^\perp = A^\perp$

Problem: Given $A \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$, find $\mathbf{x} \in \mathbb{R}^\ltimes$ minimising $||A\mathbf{x} - \mathbf{b}||$. The optimal solution for this problem is $x = A^\perp\mathbf{b}$.