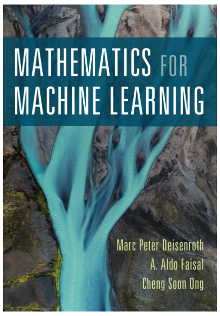
Mathematics for Machine Learning (Deisenroth, Faisal, Ong)

Chapter 4: Matrix Decompositions

Full solutions to exercises 4.1 - 4.12



CONTENTS (clickable links)

4.1) $A = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \\ 0 & 2 & 4 \end{pmatrix}$

<u>Laplace expansion on first row:</u>

$$\det(A) = 1\det\begin{pmatrix} 4 & 6 \\ 2 & 4 \end{pmatrix} - 3\det\begin{pmatrix} 2 & 6 \\ 0 & 4 \end{pmatrix} + 5\det\begin{pmatrix} 2 & 4 \\ 0 & 2 \end{pmatrix}$$

$$= 1(16-12) - 3(8) + 5(4) = 0$$

<u>Sarrus' rule</u>

$$\det(A) = 1(4)(4) + (2)(2)(5) + 0 - 0 - 2(3)(4) - 1(2)(6)$$

$$= 16 + 20 + 0 - 0 - 24 - 12 = 0.$$

4.2) Use fact that adding one row to another row doesn't change the determinant.

$$\begin{pmatrix} 2 & 0 & 1 & 2 & 0 \\ 2 & -1 & 0 & 1 & 1 \\ 0 & 1 & 2 & 1 & 2 \\ -2 & 0 & 2 & -1 & 2 \\ 2 & 0 & 0 & 1 & 1 \end{pmatrix} \xrightarrow[\substack{R3 \to R3 + R1 \\ R4 \to R4 - R1}]{R2 \to R2 - R1} \begin{pmatrix} 2 & 0 & 1 & 2 & 0 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & 1 & 2 & 1 & 2 \\ 0 & 0 & 3 & 1 & 2 \\ 0 & 0 & -1 & -1 & 1 \end{pmatrix}$$

$$\xrightarrow{R3 \to R3 + R2} \begin{pmatrix} 2 & 0 & 1 & 2 & 0 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 3 & 1 & 2 \\ 0 & 0 & -1 & -1 & 1 \end{pmatrix} \xrightarrow[R5 \to R5 + R3]{R4 \to R4 - 3R3} \begin{pmatrix} 2 & 0 & 1 & 2 & 0 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & -1 & 4 \end{pmatrix}$$

$$\xrightarrow{R5 \to R5 + R4} \begin{pmatrix} 2 & 0 & 1 & 2 & 0 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 0 & -3 \end{pmatrix}.$$

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

$$= 6.$$

4.3 a) $A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$

$$\det(A - \lambda I) = \det\begin{pmatrix} 1-\lambda & 0 \\ 1 & 1-\lambda \end{pmatrix} = (1-\lambda)^2.$$

Eigenvalues are $\lambda = 1$ (repeated).

Eigenvectors :

Need $v \in \mathbb{R}^2$ s.t $(A - I)\underline{v} = 0 \Leftrightarrow \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\underline{v} = \underline{0}$.

$\underline{v} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is valid.

Moreover, nullity $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = 2 - \text{rk}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = 2 - 1 = 1$ so $\underline{v} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ is the only valid choice, up to scaling.

So eigenspace associated to $\lambda = 1$ is $\langle e_2 \rangle = \langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \rangle$.

b) $B = \begin{pmatrix} -2 & 2 \\ 2 & 1 \end{pmatrix}$.

$$\det(B - \lambda I) = \det\begin{pmatrix} -2-\lambda & 2 \\ 2 & 1-\lambda \end{pmatrix} = (\lambda+2)(\lambda-1) - 4$$
$$= \lambda^2 + \lambda - 2 - 4 = (\lambda+3)(\lambda-2).$$

so eigenvalues are $\lambda_1 = -3$ & $\lambda_2 = 2$.

Eigenvectors :

For $\lambda_1 = -3$: Require $\underline{v} \in \mathbb{R}^2$ s.t $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}\underline{v} = \underline{0}$.
$\underline{v} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ works. As nullity $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = 1$ this is the only solution, up to scaling.

For $\lambda_2 = 2$: Require $\underline{v} \in \mathbb{R}^2$ s.t $\begin{pmatrix} -4 & 2 \\ 2 & -1 \end{pmatrix}\underline{v} = 0$.
$\underline{v} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ works, and as nullity $\begin{pmatrix} -4 & 2 \\ 2 & -1 \end{pmatrix} = 1$, this is the only solution up to scaling.

Therefore, $E_{(\lambda_1 = -3)} = \langle \begin{pmatrix} 2 \\ -1 \end{pmatrix} \rangle$, $E_{(\lambda_2 = 2)} = \langle \begin{pmatrix} 1 \\ 2 \end{pmatrix} \rangle$.

4.4) $$A - \lambda I = \begin{pmatrix} -\lambda & -1 & 1 & 1 \\ -1 & 1-\lambda & -2 & 3 \\ 2 & -1 & -\lambda & 0 \\ 1 & -1 & 1 & -\lambda \end{pmatrix}$$

$$\xrightarrow{R1 \to R1 - R4} \begin{pmatrix} -\lambda-1 & 0 & 0 & 1+\lambda \\ -1 & 1-\lambda & -2 & 3 \\ 2 & -1 & -\lambda & 0 \\ 1 & -1 & 1 & -\lambda \end{pmatrix}$$

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

$R1 \to R1 + R3.$

$$= -(\lambda+1) \left[3 \det \begin{pmatrix} -1 & -\lambda \\ -1 & 1 \end{pmatrix} + \lambda \det \begin{pmatrix} 1-\lambda & -2 \\ -1 & -\lambda \end{pmatrix} + \det \begin{pmatrix} 0 & -\lambda & -1 \\ 2 & -1 & -\lambda \\ 1 & -1 & 1 \end{pmatrix} \right]$$

$$= -(\lambda+1) \left[3(-1-\lambda) + \lambda(-\lambda+\lambda^2-2) + \lambda(2+\lambda) - 1(-2+1) \right]$$

$$= -(\lambda+1) \left[-3 - 3\lambda - \lambda^2 + \lambda^3 - 2\lambda + 2\lambda + \lambda^2 + 2 - 1 \right]$$

$$= -(\lambda+1)(\lambda^3 - 3\lambda - 2) = -(\lambda+1)(\lambda-2)(\lambda+1)^2 = -(\lambda+1)^3(\lambda-2)$$

so eigenvalues are :

$\lambda_1 = 2$ (multiplicity 1)

$\lambda_2 = -1$ (multiplicity 3).

Eigenspace for $\lambda_1 = 2$.

Need to find $\ker(A-2I) = \ker \begin{pmatrix} -2 & -1 & 1 & 1 \\ -1 & -1 & -2 & 3 \\ 2 & -1 & -2 & 0 \\ 1 & -1 & 1 & -2 \end{pmatrix}$.

As $\det(A-2I) = 0$, $\dim(\ker(A-2I)) \geq 1$.

As algebraic multiplicity of $\lambda_1 = 2$ is 1, $\dim(E_2) = \dim(\ker(A-2I)) \leq 1$

$\Rightarrow \dim(E_2) = \dim(\ker(A-2I)) = 1$.

By looking at matrix, we can see that $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \in \ker(A-2I)$.

$\therefore$ eigenspace $E_2 = \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\rangle$.

Eigenspace for $\lambda_2 = -1$ (algebraic multiplicity $= 3$).

Need to find $\ker(A+I) = \ker \begin{pmatrix} 1 & -1 & 1 & 1 \\ -1 & 2 & -2 & 3 \\ 2 & -1 & 1 & 0 \\ 1 & -1 & 1 & 1 \end{pmatrix}$

$$A+I \xrightarrow[\substack{R3 \to R3 - 2R1 \\ R4 \to R4 - R1}]{R2 \to R2 + R1} \begin{pmatrix} 1 & -1 & 1 & 1 \\ 0 & 1 & -1 & 4 \\ 0 & +1 & -1 & -2 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[R3 \to R3 - R2]{R1 \to R1 + R2} \begin{pmatrix} 1 & 0 & 0 & 5 \\ 0 & 1 & -1 & 4 \\ 0 & 0 & 0 & -6 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Therefore $\text{null}(A+I) = 1$, and (for example) $(A+I)\begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} = \underline{0}$

$\therefore$ eigenspace $E_{-1} = \left\langle \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle$.

Extra note: since $\dim(E_{-1}) = 1 < 3 =$ alg. mult of $\lambda_2 = -1$, the matrix A is <u>not</u> diagonalisable.

4.5)

Matrix	Invertible?	Diagonalisable?
$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	Yes, $\det = 1$.	Yes, it's already diagonal.
$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	No, $\det = 0$.	Yes, it's already diagonal.
$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$	Yes, $\det = 1$.	No. Concretely: - eigenvalue $\lambda = 1$ has multiplicity 2 $\dim(E_1) = \operatorname{null}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 1$. $(1 \neq 2)$ Theoretically: - Jordan Canonical form is unique & matrix is already in J.C.F. which is not diagonal. Aside: Go-to example for failure of Maschke's theorem for infinite groups $\left(\rho : \mathbb{Z} \to GL_2(\mathbb{C})\,;\ 1 \mapsto \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\right)$.
$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$	No, $\det = 0$.	No. - eigenvalue $\lambda = 0$ has multiplicity 2. $\dim(E_0) = \operatorname{null}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 1$ $(1 \neq 2)$ - Alternatively, it's not diagonalisable as it has a non-diagonal Jordan Canonical Form.

4.6 a) $A-\lambda I = \begin{pmatrix} 2-\lambda & 3 & 0 \\ 1 & 4-\lambda & 3 \\ 0 & 0 & 1-\lambda \end{pmatrix}$

$$\det(A-\lambda I) = (1-\lambda)\det\begin{pmatrix} 2-\lambda & 3 \\ 1 & 4-\lambda \end{pmatrix}$$

$$= (1-\lambda)[(2-\lambda)(4-\lambda)-3]$$

$$= (1-\lambda)(\lambda^2-6\lambda+5) = -(\lambda-1)(\lambda-5)(\lambda-1)$$

$$= -(\lambda-1)^2(\lambda-5).$$

So eigenvalues are :

$\lambda_1 = 5$ (multiplicity 1)

$\lambda_2 = 1$ (multiplicity 2).

Eigenspaces :

- For $\lambda_1 = 5$; $(1 \leqslant \dim(E_5) \leqslant 1 \iff \dim(E_5) = 1)$.

$$E_5 = \ker(A-5I) = \ker\begin{pmatrix} -3 & 3 & 0 \\ 1 & -1 & 3 \\ 0 & 0 & -4 \end{pmatrix}$$

It's clear to see that $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \in E_5$.

Since $\dim(E_5) = 1$, we have $E_5 = \left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle$.

- For $\lambda_2 = 1$; $\quad (1 \leq \dim(E_1) \leq 2 = \text{alg. multiplicity of } \lambda_2 = 1)$

$$E_1 = \ker(A - I) = \begin{pmatrix} 1 & 3 & 0 \\ 1 & 3 & 3 \\ 0 & 0 & 0 \end{pmatrix}$$

We can see that $rk(A - I) = 2 \Rightarrow null(A - I) = 3 - 2 = 1.$

clearly $\begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix} \in E_1.$

Since $\dim(E_1) = 1$, we have $E_1 = \left\langle \begin{pmatrix} 3 \\ -1 \\ 0 \end{pmatrix} \right\rangle.$

Finally, since $\dim(E_1) = 1 < 2 = \text{alg. multiplicity of } \lambda_2 = 1$

the matrix A is <u>not</u> diagonalisable.

b) $$A - \lambda I = \begin{pmatrix} 1-\lambda & 1 & & \\ & -\lambda & & 0 \\ 0 & & -\lambda & \\ & & & -\lambda \end{pmatrix}$$

$$\det(A - \lambda I) = (\lambda - 1)\lambda^3.$$

so eigenvalues are:

$\lambda_1 = 1$ (multiplicity 1).

$\lambda_2 = 0$ (multiplicity 3).

Eigenspaces:

- For $\lambda_1 = 1$ $\quad (1 \leq \dim(E_1) \leq 1 \Rightarrow \dim(E_1) = 1).$

$$E_1 = \ker(A - I) = \ker \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

We have $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \in E_1$, and since $\dim(E_1) = 1$; $E_1 = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\rangle$.

- For $\lambda_2 = 0$ $\quad (1 \leq \dim(E_0) \leq 3)$.

$$E_0 = \ker(A) = \ker \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

We have $\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \in E_0$.

Moreover, these 3 vectors are linearly independent.

This along with $\dim(E_0) = 3 \Rightarrow E_0 = \left\langle \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle$.

As geometric multiplicities (i.e dimensions of eigenspaces) are equal to algebraic multiplicities, the matrix is diagonalisable.

We have :

$$\begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

$$A = P \quad D \quad P^{-1}$$

(our original matrix) $\quad$ (these are the eigenvectors) $\quad$ (diagonal matrix of eigenvalues) $\quad$ (note: $P^{-1} = P$)

4.7a) $A = \begin{pmatrix} 0 & 1 \\ -8 & 4 \end{pmatrix}$.

$$\det(A-\lambda I) = \det\begin{pmatrix} -\lambda & 1 \\ -8 & 4-\lambda \end{pmatrix} = \lambda(\lambda-4)+8 = \lambda^2-4\lambda+8.$$

The discriminant, $D = (-4)^2 - 4(1)(8) = -16$, is <0. Therefore there are no real eigenvalues, and A is not diagonalisable.

b) $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

The matrix is symmetric, so it is diagonalisable.

Observe that $\operatorname{im}(A) = \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle$.

Therefore, $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ must be an eigenvector.

Indeed, $A\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$. Thus 3 is an eigenvalue.

Moreover, let $A = PDP^{-1}$, then $\operatorname{rk}(A) = \operatorname{rk}(D) = \underline{1}$

(This is clear, as one can take the point of view of A & D representating the same linear map with different bases.)

Therefore 3 is the only non-zeo eigenvalue.

For eigenvalue 0, we require $E_0 = \ker\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$.

$E_0 = \left\langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\rangle$ is clear. (other choices exist.)

Therefore,
$$A = \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}}_{\text{"}P\text{"}} \underbrace{\begin{pmatrix} 3 & & \\ & 0 & \\ & & 0 \end{pmatrix}}_{\text{"}D\text{"}} \underbrace{\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}^{-1}}_{\text{"}P^{-1}\text{"}}$$

c) $A = \begin{pmatrix} 5 & 4 & 2 & 1 \\ 0 & 1 & -1 & -1 \\ -1 & -1 & 3 & 0 \\ 1 & 1 & -1 & 2 \end{pmatrix}$

$$A-\lambda I = \begin{pmatrix} 5-\lambda & 4 & 2 & 1 \\ 0 & 1-\lambda & -1 & -1 \\ -1 & -1 & 3-\lambda & 0 \\ 1 & 1 & -1 & 2-\lambda \end{pmatrix} \xrightarrow{R3 \to R3+R4} \begin{pmatrix} 5-\lambda & 4 & 2 & 1 \\ 0 & 1-\lambda & -1 & -1 \\ 0 & 0 & 2-\lambda & 2-\lambda \\ 1 & 1 & -1 & 2-\lambda \end{pmatrix}$$

As $\det(A) = \det(A^T)$, we can add a column to another column and not change the determinant

$$\xrightarrow{C3 \to C3-C4} \begin{pmatrix} 5-\lambda & 4 & 1 & 1 \\ 0 & 1-\lambda & 0 & -1 \\ 0 & 0 & 0 & 2-\lambda \\ 1 & 1 & \lambda-3 & 2-\lambda \end{pmatrix}$$

Thus, $\det(A-\lambda I) = -(2-\lambda)\det\begin{pmatrix} 5-\lambda & 4 & 1 \\ 0 & 1-\lambda & 0 \\ 1 & 1 & \lambda-3 \end{pmatrix}$

$$= -(2-\lambda)(1-\lambda)\det\begin{pmatrix} 5-\lambda & 1 \\ 1 & \lambda-3 \end{pmatrix}$$

$$= -(2-\lambda)(1-\lambda)\left[(5-\lambda)(\lambda-3)-1\right]$$

$$= +(2-\lambda)(1-\lambda)(\lambda^2-8\lambda+16)$$

$$= (\lambda-2)(\lambda-1)(\lambda-4)^2.$$

We have $\dim(E_2) = 1$, $\dim(E_1) = 1$, $1 \leq \dim(E_4) \leq 2$.

Therefore A is diagonalisable $\Leftrightarrow \dim(E_4) = 2. \Leftrightarrow \text{null}(A-4I) = 2.$

$$A-4I = \begin{pmatrix} 1 & 4 & 2 & 1 \\ 0 & -3 & -1 & -1 \\ -1 & -1 & -1 & 0 \\ 1 & 1 & -1 & -2 \end{pmatrix} \xrightarrow[R4 \to R4-R1]{R3 \to R3+R1} \begin{pmatrix} 1 & 4 & 2 & 1 \\ 0 & -3 & -1 & -1 \\ 0 & 3 & 1 & 1 \\ 0 & -3 & -3 & -3 \end{pmatrix}$$

We can see that $rk(A-4I) = 3$ (look at rows 1, 2, 4 in RHS matrix)
thus $\dim(E_4) = \text{null}(A-4I) = 1$ so A is not diagonalisable.

d) $A = \begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 3 & -6 & -4 \end{pmatrix}$

$$A - \lambda I = \begin{pmatrix} 5-\lambda & -6 & -6 \\ -1 & 4-\lambda & 2 \\ 3 & -6 & -4-\lambda \end{pmatrix} \xrightarrow{R3 \to R3 + 3R2} \begin{pmatrix} 5-\lambda & -6 & -6 \\ -1 & 4-\lambda & 2 \\ 0 & 6-3\lambda & 2-\lambda \end{pmatrix}$$

$$\therefore \det(A - \lambda I) = (5-\lambda) \det \begin{pmatrix} 4-\lambda & 2 \\ 6-3\lambda & 2-\lambda \end{pmatrix} + 1 \det \begin{pmatrix} -6 & -6 \\ 6-3\lambda & 2-\lambda \end{pmatrix}$$

$$= (5-\lambda)\left[(4-\lambda)(2-\lambda) - 2(6-3\lambda)\right] + (-6)(2-\lambda) + 6(6-3\lambda)$$

$$= (5-\lambda)(\lambda^2 - 4) - 12\lambda + 24$$

$$= (5-\lambda)(\lambda-2)(\lambda+2) - 12(\lambda-2)$$

$$= (\lambda-2)\left[(5-\lambda)(\lambda+2) - 12\right] = -(\lambda-2)(\lambda^2 - 3\lambda + 2)$$

$$= -(\lambda-2)^2(\lambda-1).$$

We have A diagonalisable $\Leftrightarrow \dim(E_2) = 2$.

- Eigenspace for $\lambda = 2$:

$$E_2 = \ker(A - 2I) = \ker \begin{pmatrix} 3 & -6 & -6 \\ -1 & 2 & 2 \\ 3 & -6 & -6 \end{pmatrix} \overset{\substack{R3 \to R3 - R1 \\ \text{then} \\ R1 \to \frac{1}{3}R1 \\ R2 \to R2 + R1}}{=\!=\!=} \ker \begin{pmatrix} 1 & -2 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

so $\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \in E_2$. As these vectors are linearly independent and $\dim(E_2) \leq 2$, we have $E_2 = \left\langle \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \right\rangle$.

Thus A is diagonalisable! We need to find basis of E_1.

$$E_1 = \ker(A - I) = \ker \begin{pmatrix} 4 & -6 & -6 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{pmatrix} \underset{\substack{R1 \rightleftarrows R2 \\ \text{then} \\ R1 \to -R1 \\ R2 \to R2 - 4R1 \\ R3 \to R3 - 3R1}}{=\!=\!=} \ker \begin{pmatrix} 1 & -3 & -2 \\ 0 & 6 & 2 \\ 0 & 3 & 1 \end{pmatrix}$$

$$\xrightarrow[\substack{R2 \rightleftarrows R3 \\ R1 \to R1+R2 \\ R3 \to R3 - 2R2.}]{} \quad \ker \begin{pmatrix} 1 & 0 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

So $$E_1 = \left\langle \begin{pmatrix} -1 \\ \frac{1}{3} \\ -1 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 3 \\ -1 \\ 3 \end{pmatrix} \right\rangle$$

So A is diagonalisable and we have

$$A = \begin{pmatrix} 3 & 2 & 2 \\ -1 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & & \\ & 2 & \\ & & 2 \end{pmatrix} \begin{pmatrix} 3 & 2 & 2 \\ -1 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}^{-1}$$

4.8) $$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

(If $A = U\Sigma V^T$ then $A^TA = (V\Sigma U^T)(U\Sigma V^T) = V\Sigma^2 V^T$.)

$$A^TA = \begin{pmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix} = \begin{pmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{pmatrix}.$$

Eigenvalues:

$$A^TA - \lambda I = \begin{pmatrix} 13-\lambda & 12 & 2 \\ 12 & 13-\lambda & -2 \\ 2 & -2 & 8-\lambda \end{pmatrix} \xrightarrow{R2 \to R2 - 6R3} \begin{pmatrix} 13-\lambda & 12 & 2 \\ 0 & 25-\lambda & -50+6\lambda \\ 2 & -2 & 8-\lambda \end{pmatrix}$$

$$\begin{aligned} \det(A^TA - \lambda I) &= (13-\lambda)\left[(25-\lambda)(8-\lambda) + 2(-50+6\lambda)\right] + 2\left[12(-50+6\lambda) - 2(25-\lambda)\right] \\ &= (13-\lambda)(\lambda^2 - 33\lambda + 200 - 100 + 12\lambda) + 2(72\lambda - 600 - 50 + 2\lambda) \\ &= -\lambda^3 + 34\lambda^2 - 225\lambda + 0 \\ &= -\lambda(\lambda^2 - 34\lambda + 225) \\ &= -\lambda(\lambda-9)(\lambda-25). \end{aligned}$$

$(-25)(-9) = 225.$
$-25 - 9 = -34$

Eigenspaces:

- $E_0 = \ker\begin{pmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{pmatrix}$ (Observe, $R1 = R2 + \frac{1}{2}R3$. Use this to remove $R1$. Then apply $R2 \to R2 - 6R3$ to clear entry in position $(2,1)$)

$$\xrightarrow[R2 \to R2 - 6R3]{R1 \to R1 - R2 - \frac{1}{2}R3} \ker\begin{pmatrix} 0 & 0 & 0 \\ 0 & 25 & -50 \\ 2 & -2 & 8 \end{pmatrix} \xrightarrow[\substack{R2 \to \frac{1}{25}R2 \\ R1 \to R1 + 2R2}]{R1 \rightleftarrows R3} \begin{pmatrix} 2 & 0 & 4 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

so $E_0 = \left\langle \begin{pmatrix} 2 \\ -2 \\ -1 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} 2/3 \\ -2/3 \\ -1/3 \end{pmatrix} \right\rangle$ ← normalised i.e. length = 1.

- $E_9 = \ker\begin{pmatrix} 4 & 12 & 2 \\ 12 & 4 & -2 \\ 2 & -2 & -1 \end{pmatrix} \xrightarrow[R2 \to R2 - 6R3]{R1 \to R1 - 2R3} \ker\begin{pmatrix} 0 & 16 & 4 \\ 0 & 16 & 4 \\ 2 & -2 & -1 \end{pmatrix}$

$$\xrightarrow[\substack{R3 \to R3 - R2 \\ R2 \to \frac{1}{16}R2}]{\substack{R1 \rightleftarrows R3 \\ R1 \to \frac{1}{2}R1}} \ker\begin{pmatrix} 1 & -1 & -\frac{1}{2} \\ 0 & 1 & \frac{1}{4} \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{R1 \to R1 + R2} \ker\begin{pmatrix} 1 & 0 & -\frac{1}{4} \\ 0 & 1 & \frac{1}{4} \\ 0 & 0 & 0 \end{pmatrix}$$

so $E_9 = \left\langle \begin{pmatrix} +1 \\ -1 \\ +4 \end{pmatrix} \right\rangle = \left\langle \frac{1}{\sqrt{18}}\begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix} \right\rangle = \left\langle \frac{1}{3\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 4 \end{pmatrix} \right\rangle$

- $E_{25} = \ker\begin{pmatrix} -12 & 12 & 2 \\ 12 & -12 & -2 \\ 2 & -2 & -17 \end{pmatrix}$. Col 1 = − Col 2 & $\dim(E_{25}) = 1$. So:

$E_{25} = \left\langle \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle = \left\langle \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\rangle$.

Write $A = U \Sigma V^T$, then

$$\Sigma = \begin{pmatrix} \sqrt{25} & 0 & 0 \\ 0 & \sqrt{9} & 0 \end{pmatrix} = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \quad \& \quad V^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{3\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{4}{3\sqrt{2}} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \end{pmatrix}$$

Finally; $A = U\Sigma V^T \Rightarrow U\Sigma = AV$ (note: $V^{-1} = V^T$)

$$\Leftrightarrow \begin{pmatrix} | & | & | \\ 5u_1 & 3u_1 & 0 \\ | & | & | \end{pmatrix} = AV \in M_{2\times 3}(\mathbb{R}).$$

$$AV = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & 2/3 \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & -2/3 \\ 0 & \frac{4}{3\sqrt{2}} & -1/3 \end{pmatrix} = \begin{pmatrix} 5/\sqrt{2} & \frac{3}{\sqrt{2}} & 0 \\ 5/\sqrt{2} & -3/\sqrt{2} & 0 \end{pmatrix}$$

so $U = \begin{pmatrix} \frac{1}{\sqrt{2}} & 1/\sqrt{2} \\ \frac{1}{\sqrt{2}} & -1/\sqrt{2} \end{pmatrix}$.

Altogether; $A = U\Sigma V^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{3\sqrt{2}} & -\frac{1}{3\sqrt{2}} & \frac{4}{3\sqrt{2}} \\ 2/3 & -2/3 & -1/3 \end{pmatrix}$

Check: Calculating U directly.

$AA^T = (U\Sigma V^T)(V\Sigma U^T) = U$ "Σ^2" U^T.

$$AA^T = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{pmatrix} = \begin{pmatrix} 17 & 8 \\ 8 & 17 \end{pmatrix}$$

$$\det(AA^T - \lambda I) = (17-\lambda)^2 - 64 = \lambda^2 - 34\lambda + 225 = (\lambda - 25)(\lambda - 9).$$

Eigenspaces:

$E_9 = \ker \begin{pmatrix} 8 & 8 \\ 8 & 8 \end{pmatrix}$

so $E_9 = \langle \begin{pmatrix} 1 \\ -1 \end{pmatrix} \rangle = \langle \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \rangle$.

$E_{25} = \ker \begin{pmatrix} -8 & 8 \\ 8 & -8 \end{pmatrix}$

so $E_{25} = \langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle = \langle \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle$

so "Σ^2" $= \begin{pmatrix} 25 & \\ & 9 \end{pmatrix} \Rightarrow \Sigma = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}$ & $U = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$

4.9) $A = \begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix}$ (If $A = U\Sigma V^T$ then $A^TA = V\Sigma^2V^T$).

$$A^TA = \begin{pmatrix} 2 & -1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$$

$$A^TA - \lambda I = \begin{pmatrix} 5-\lambda & 3 \\ 3 & 5-\lambda \end{pmatrix}.$$

We see that to obtain colinearity between $\begin{pmatrix} 5-\lambda \\ 3 \end{pmatrix}$ & $\begin{pmatrix} 3 \\ 5-\lambda \end{pmatrix}$ we can set $\lambda = 2$ or $\lambda = 8$.

- If $\lambda = 2$, $E_2 = \ker \begin{pmatrix} 3 & 3 \\ 3 & 3 \end{pmatrix} = \langle \begin{pmatrix} 1 \\ -1 \end{pmatrix} \rangle = \langle \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \rangle.$
- If $\lambda = 8$, $E_2 = \ker \begin{pmatrix} -3 & 3 \\ 3 & -3 \end{pmatrix} = \langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle = \langle \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle.$

Therefore $\Sigma = \begin{pmatrix} \sqrt{8} & 0 \\ 0 & \sqrt{2} \end{pmatrix} = \begin{pmatrix} 2\sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix}$

& $V = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \Leftrightarrow V^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$

We have $A = U\Sigma V^T \Leftrightarrow U\Sigma = AV \Leftrightarrow \begin{pmatrix} | & | \\ \sqrt{8}u_1 & \sqrt{2}u_2 \\ | & | \end{pmatrix} = AV.$

$$AV = \frac{1}{\sqrt{2}}\begin{pmatrix} 2 & 2 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 4 & 0 \\ 0 & -2 \end{pmatrix} = \begin{pmatrix} \sqrt{8} & 0 \\ 0 & -\sqrt{2} \end{pmatrix}$$

Thus, $U = \begin{pmatrix} | & | \\ u_1 & u_2 \\ | & | \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$

Altogether;

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \sqrt{8} & 0 \\ 0 & \sqrt{2} \end{pmatrix}\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

4.10) From Exercise 4.8) we have:

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{3\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{4}{3\sqrt{2}} \\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3} \end{pmatrix}$$

The rank-1 approximation is:

$$\sigma_1 A_1 = \sigma_1 u_1 v_1^T = 5 \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{pmatrix}$$

$$= 5 \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} = \frac{5}{2} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix}.$$

4.11) Suppose $\lambda \in \text{Spectrum}(A^TA) \setminus \{0\}$.

Then $A^TA\,\underline{v} = \lambda \underline{v}$ for some $\underline{v} \in \mathbb{R}^n$

Multiplying by A yields $A(A^TA)\underline{v} = \lambda A \underline{v}$

$$\Leftrightarrow (AA^T)(A\underline{v}) = \lambda(A\underline{v})$$

Thus, $\lambda \in \text{Spectrum}(AA^T)$.

The converse is completely symmetric. Let's just write it.

Conversely, suppose $\lambda \in \text{Spectrum}(AA^T) \setminus \{0\}$.

Then $AA^T\underline{v} = \lambda \underline{v}$ for some $\underline{v} \in \mathbb{R}^m$.

Multiplying by A^T yields $A^T(AA^T\underline{v}) = \lambda A^T\underline{v}$

$$\Leftrightarrow (A^TA)(A^T\underline{v}) = \lambda(A^T\underline{v})$$

so $\lambda \in \text{Spectrum}(A^TA)$.

4.12) $$\max_{\underline{x}\neq \underline{0}} \frac{\|A x\|_2}{\|x\|_2} \overset{\text{norm is homogeneous.}}{=} \max_{\underline{x}\neq \underline{0}} \left\| A \frac{x}{\|x\|_2} \right\|_2 .$$

Therefore, w.l.o.g, we may assume that $\|x\|_2 = 1$.

So $$\max_{\underline{x}\neq \underline{0}} \frac{\|A x\|_2}{\|x\|_2} = \max_{\substack{\underline{x} \text{ s.t} \\ \|x\|_2 = 1}} \|A \underline{x}\|_2 .$$

Write $A = U \Sigma V^T$, where U & V are orthogonal.

$$= \max_{\|x\|=1} \|U \Sigma V^T x\|_2 .$$

Two things:

1) U is orthogonal, therefore $\|Uy\|_2 = \|y\|_2$.
In other words, U doesn't change the length.

2) V^T is (~~invertible &~~) orthogonal, so the set $\{V^T x\}_{\|x\|=1}$ is equal to the set $\{y\}_{\|y\|=1}$. (Define $y := V^T \underline{x}$ for each $\underline{x}$. $\|y\| = \|x\|$ since V^T orthogonal).

Thus;

$$\max_{\|x\|=1} \|U \Sigma V^T x\|_2 = \max_{\|y\|=1} \|\Sigma y\|_2$$

Intuitively, you maximise this by taking the vector in the direction which is scaled most by $\Sigma = \begin{pmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \end{pmatrix}$

So $$\max_{\|y\|=1} \|\Sigma y\|_2 = \Sigma \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \sigma_1 .$$

Let's actually prove this...

Let $y = (y_1, \dots, y_n)$ s.t $\|y\|_2 = 1$.

Then $\|\Sigma y\|_2^2 = \left\| \begin{pmatrix} \sigma_1 y_1 \\ \vdots \\ \sigma_n y_n \end{pmatrix} \right\|_2^2 = \sigma_1^2 y_1^2 + \dots + \sigma_n^2 y_n^2$

$$\le \sigma_1^2 y_1^2 + \sigma_1^2 y_2^2 + \dots + \sigma_1^2 y_n^2$$

$$= \sigma_1^2 (y_1^2 + \dots + y_n^2) = \sigma_1^2 \|y\|_2^2 = \sigma_1^2$$

so indeed $\max\limits_{\|y\|=1} \|Ay\|_2 = \sqrt{\sigma_1^2} = \sigma_1$.