Page No.

Date / /

Exercise - 1 The Geometry of matrices.

matrix and vector:

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 0.5 \end{bmatrix}, \quad \vec{v} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

(1) Because A is diagonal with positive entries on the diagonal, it represents a scaling transformation.
It scales the x-axis by a factor of 2 and the y-axis by a factor of 0.5. There is no rotation or shearing involved.

(2)

$$A \times \vec{v} = \begin{bmatrix} 2 & 0 \\ 0 & 0.5 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \cdot 1 + 0 \cdot 4 \\ 0 \cdot 1 + 0.5 \cdot 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

(3)

$$\det(A) = 2 \times 0.5 - 0 \times 0 = 1.0$$

$\because$ $\det(A) = 1$, the transformation neither expands nor contracts the area of space. It preserves area exactly.

Date / /

Exercise-9:- The stable Axes of Data covariance matrix:-

$$S = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$$

(1) charecteristic equation

$$\det(S - \lambda I) = 0 \Rightarrow \det \begin{bmatrix} 4-\lambda & 2 \\ 2 & 1-\lambda \end{bmatrix} = 0$$

(2) finding Eigenvalues:

Expanding the determinant

$$(4-\lambda)(1-\lambda) - (2)(2) = 0$$

$$4 - 4\lambda - \lambda + \lambda^2 - 4 = 0$$

$$\lambda^2 - 5\lambda = 0$$

$$\lambda(\lambda - 5) = 0$$

$$\boxed{\lambda_1 = 5, \ \lambda_2 = 0}$$

(3) finding eigenvector for $\lambda_1 = 5$

solving $(S - 5I)\vec{w} = 0$:

$$\begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix} \vec{w} = 0$$

Exercise-3: Building PCA from scratch

① Why standardize.

→ PCA finds directions of maximum variance.
→ If features are on vastly different scales · feature A in the range $[0, 0.001]$ and feature B in $[0, 1000]$ - then the covariance matrix will be dominated by feature B simply because its numerical values are large, not because it is more informative.

② PCA Objective function

→ we seek a unit vector $\vec{w}$ (with $\|\vec{w}\|=1$) that maximises the variance of the projected data:

$$\max_{\vec{w}} \vec{w}^T \Sigma \vec{w} \quad \text{subject to } \vec{w}^T\vec{w} = 1$$

③ Dimensions of the projection Matrix W

→ with $p=10$ original features and a target of $k=2$ dimensions, the projection matrix W must have dimensions 10×2.
→ Each of its 2 columns is one of the top eigenvectors.
→ The projected dataset $Z = XW$ then has shape

$$(n \times 10) \cdot (10 \times 2) = (n \times 2).$$

Page No.
Date / /

From row 1: $-w_1 + 2w_2 = 0 \Rightarrow w_1 = 2w_2$.
Choosing $w_2 = 1$:

$$\vec{w} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} ;\ \text{normalized} : \vec{w} = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

(4) variance explained by PC1

$$\frac{\lambda_1}{\lambda_1 + \lambda_2} = \frac{5}{5+0} = \frac{5}{5} = 100\%.$$

→ The first principle component captures 100% of the variance.
→ This makes sense because $\det(S) = 0$,
→ meaning the data is perfectly collinear - it lies on a 1D line.