Day 5: Mathematical Foundations of AI and Data science

Session By

Shankar J

Research Scholar, ECED

National Institute of Technology Calicut

Mathematical Foundations of AI and Data science

- PhD Scholar, NIT, Calicut
- Former academician- KTU
- Chair, Students Activities, IEEE Kerala Section 2021
- MoC, ATAL, NITI Aayog, Govt of India
- HAM / Amateur Radio – VU3FIY

Session By

Shankar J

Disclaimer

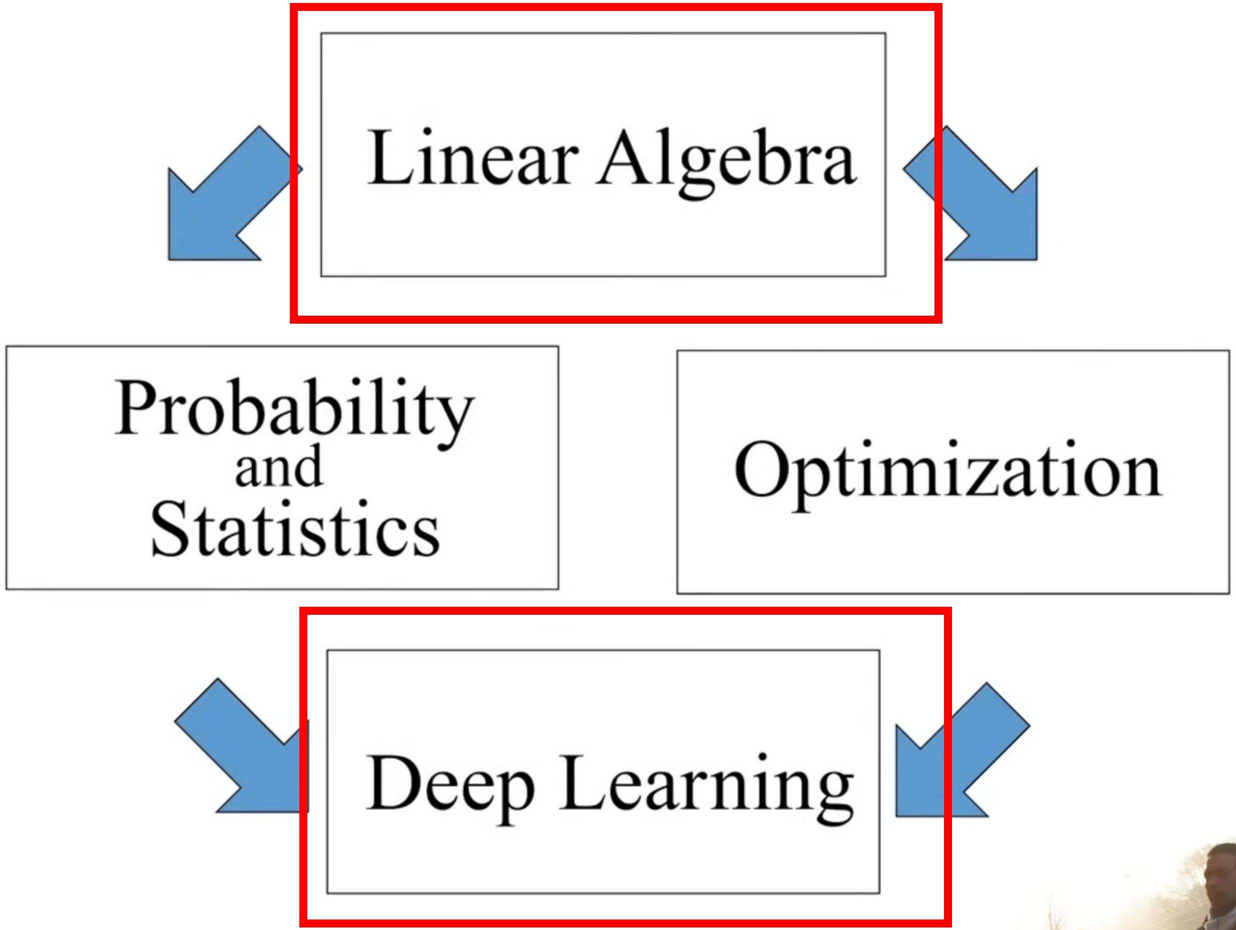
Shankar J

5 day - FDP: INTRODUCTION TO ARTIFICIAL INTELLIGENCE AND DATA SCIENCE (GNEST305)

2	**Mathematical Foundations of AI and Data science**: Role of linear algebra in Data representation and analysis – Matrix decomposition- Singular Value Decomposition (SVD)- Spectral decomposition- Dimensionality reduction technique-Principal Component Analysis (PCA). (Text-1)	11

CO2	Apply advanced mathematical concepts such as matrix operations, singular values, and principal component analysis to analyze and solve engineering problems.	K3

Apply



Symmetric Matrices
Orthogonal Matrices

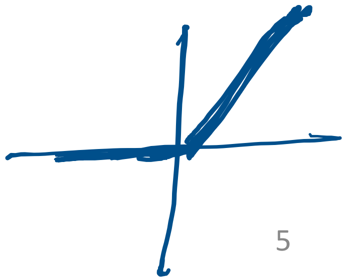
Factor a matrix

Symmetric * orthogonal mat
Orthogonal * diagonal * orthogonal mat (SVD)

If all linear algebra FAILS !!!

Adding non linear fn: b/w matrices

$f(x) = x$ +ve
$f(x) = 0$ -ve



Source: Gilbert Strang MIT Lecture

Why Vectors and Matrices Matter

- ➢ Maps - each location : latitude and longitude

Vectors and matrices form the mathematical foundation

- ✓ Representing data points in feature spaces (vectors)
- ✓ Storing image pixels as numerical grids (matrices)
- ✓ Transforming data through rotations, scaling, and projections
- ✓ Solving systems of equations in optimisation problems

What is a Vector?

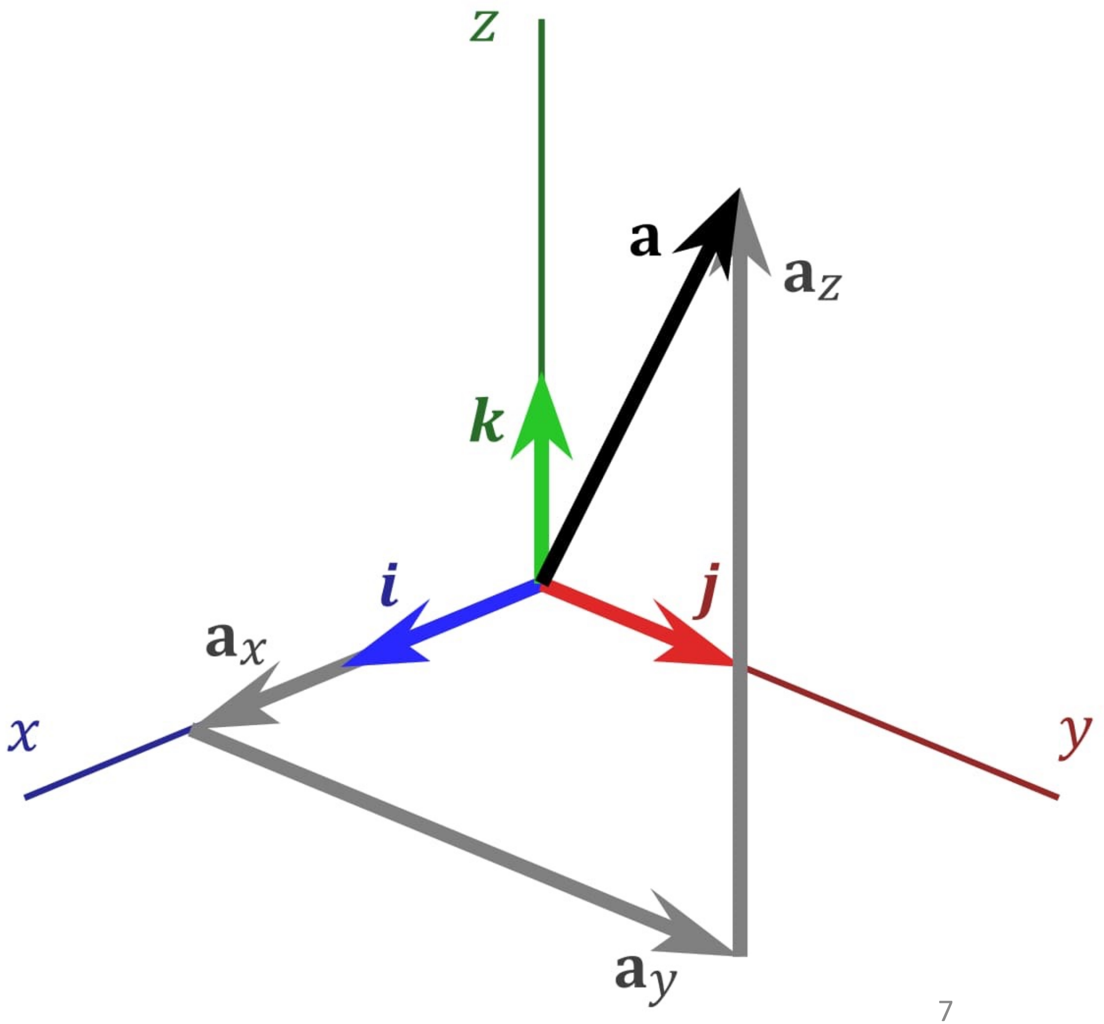
A vector is an ordered list of numbers that can represent a point or direction in space

$$\text{Column vector: } \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

$$\mathbb{R}^{\mathbf{n}}$$

$$\text{Row vector: } \mathbf{v} = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$$

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \quad \text{or sometimes written as} \quad [v_1, v_2, ..., v_n]^T$$

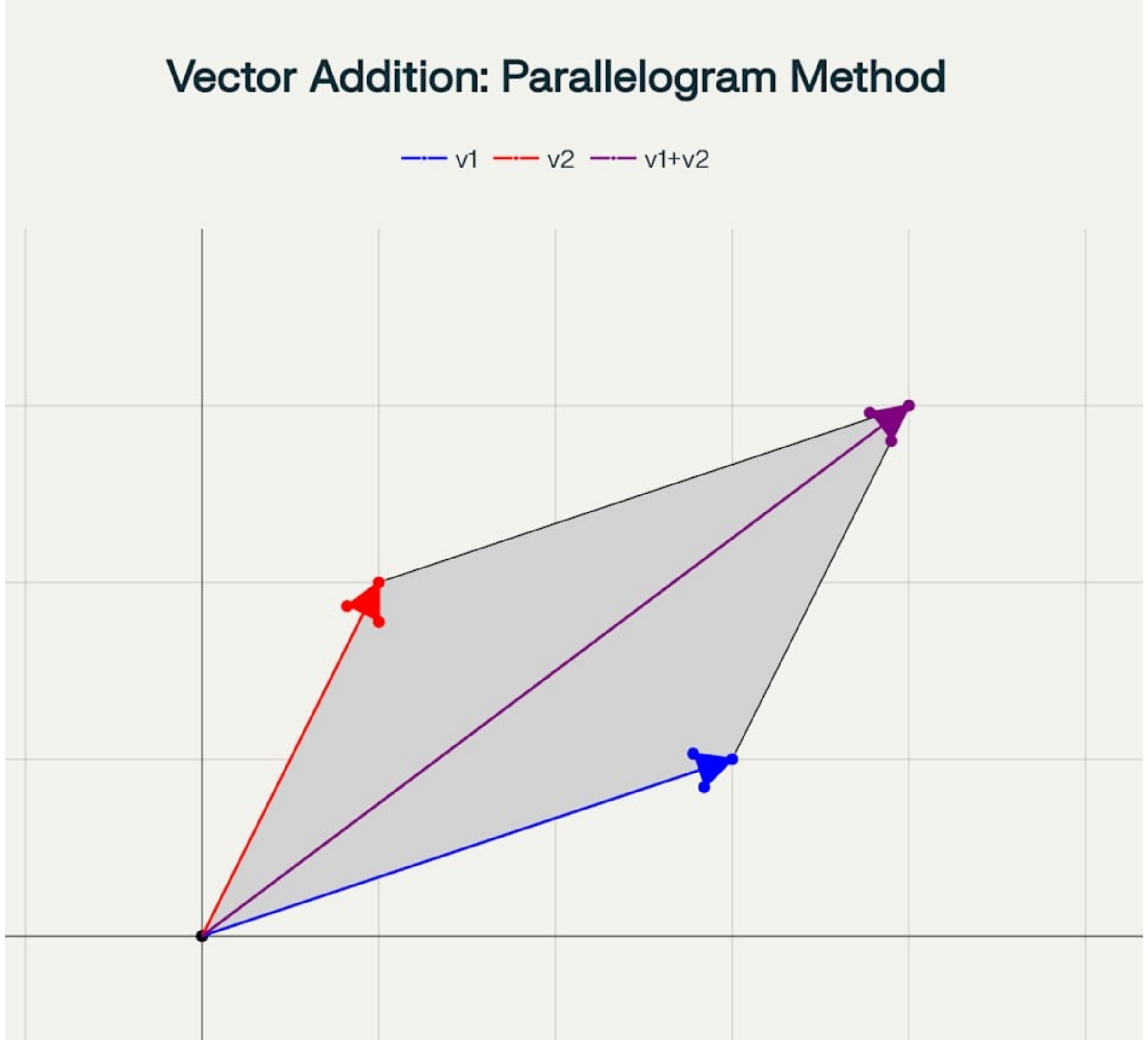


Basic Vector Operations

a) Addition

If $\mathbf{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$,

$$\mathbf{a} + \mathbf{b} = \begin{bmatrix} 1+3 \\ 2+4 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$



b) Scalar Multiplication

If $k = 2$ and $\mathbf{a} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$,

$$k\mathbf{a} = 2 \times \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

stretches or shrinks a vector

Vector Norms

Vector norms measure the "size" or "length" of vectors and are fundamental to understanding distance metrics in AI and data science

A vector norm is a function that maps a vector to a non-negative real number representing its "length".

L1 Norm (Manhattan Distance): Sum of the absolute values of vector elemer $\|x\|_1 = |x_1| + |x_2| + ... + |x_n|$

L2 Norm (Euclidean Distance): Square root of the sum of squared eleme $\|x\|_2 = \sqrt{x_1^2 + x_2^2 + ... + x_n^2}$

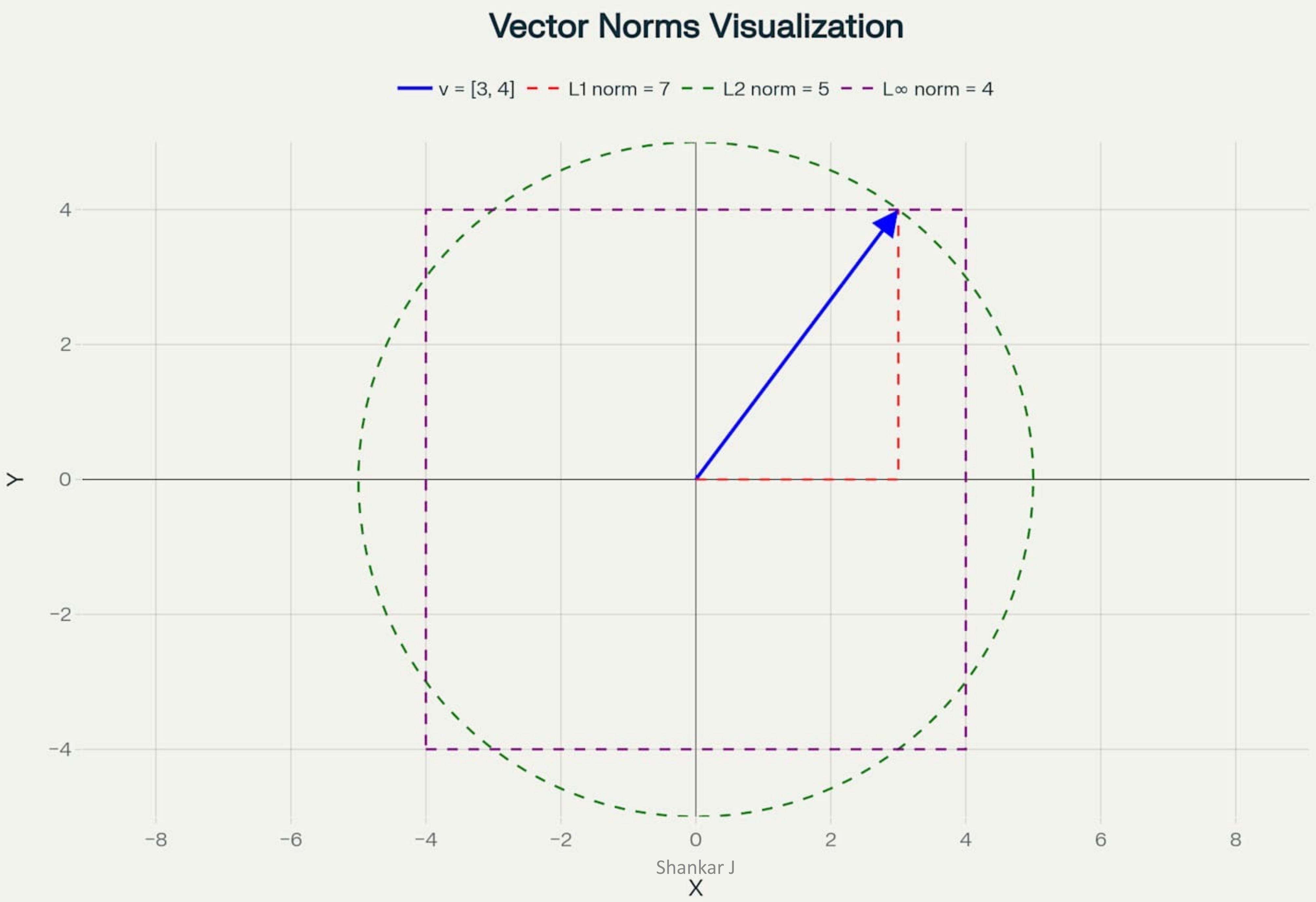
L∞ Norm (Maximum Norm): Maximum absolute value among all eleme $\|x\|_\infty = \max(|x_1|, |x_2|, ..., |x_n|)$

Applications in ML and Data Science

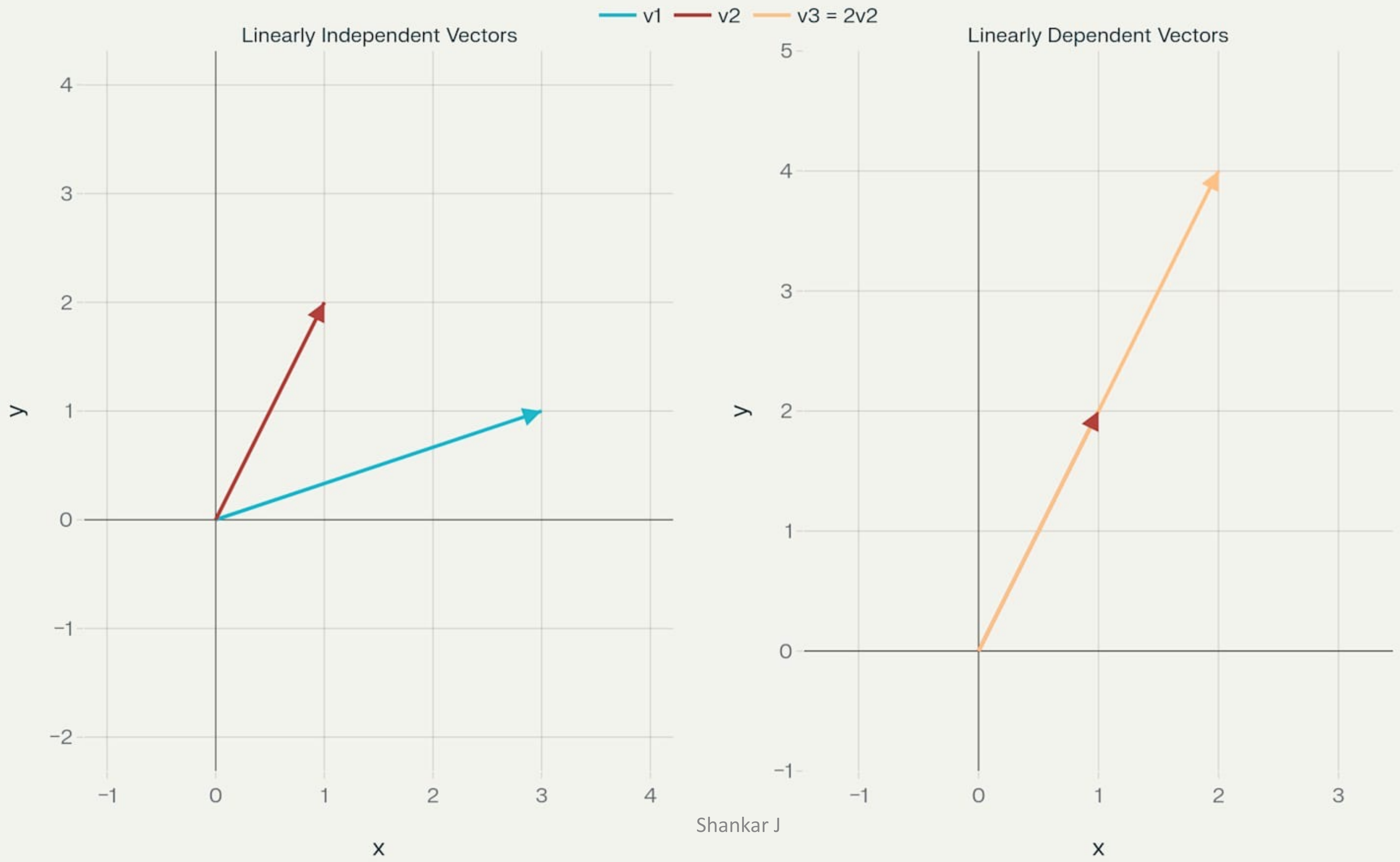
Norms are widely used in machine learning for:

- Distance metrics in k-Nearest Neighbours algorithms
- Loss functions and regularisation (L1/L2 regularisation)
- Measuring model convergence
- Feature normalisation and scaling

If x = [3, −4, 2],
L1 norm is 9
L2 norm is $\sqrt{29} \approx 5.39$
L ∞ norm is 4

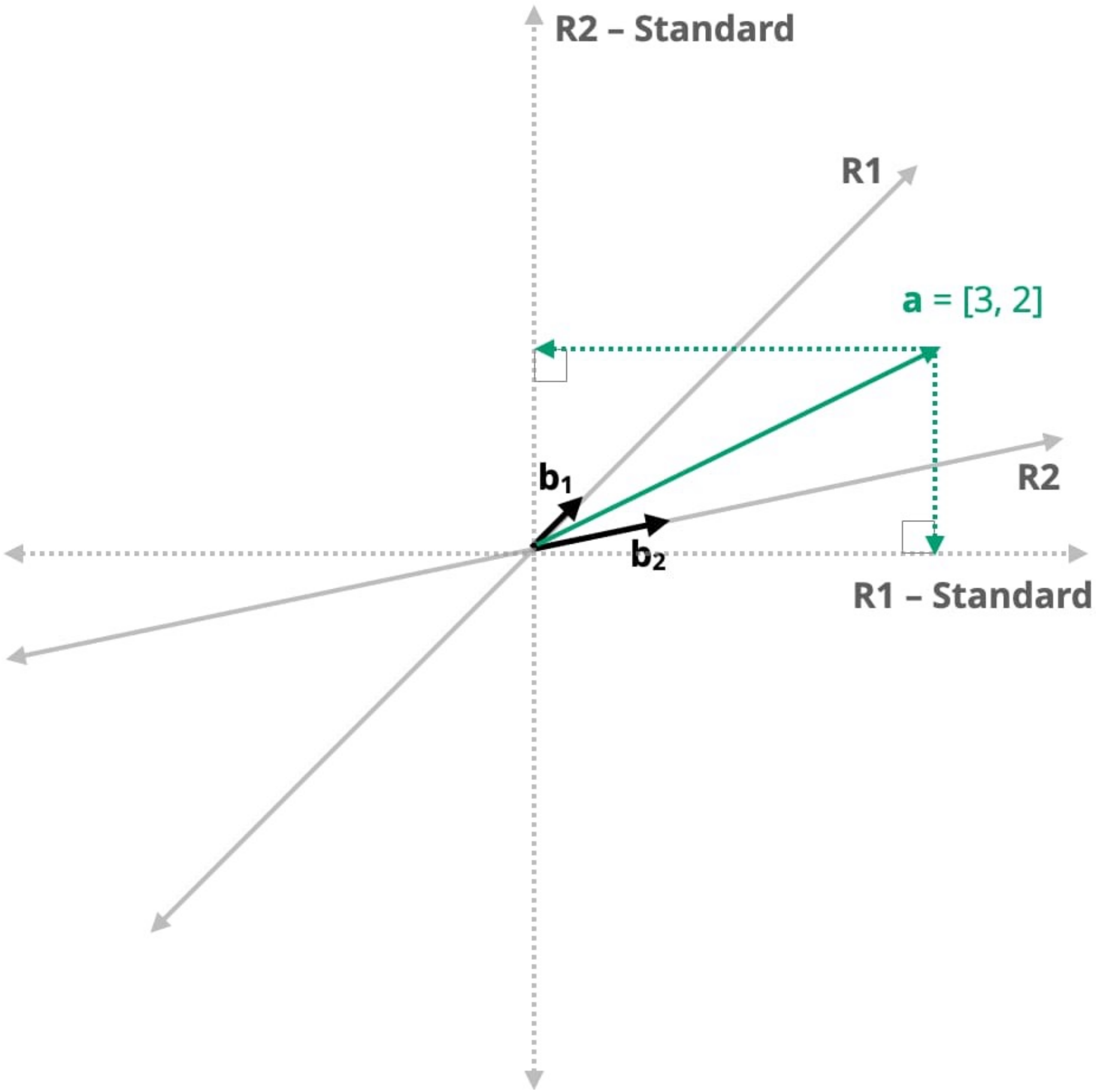

Vector Norms Visualization
v = [3, 4]
L1 norm = 7
L2 norm = 5
L∞ norm = 4
Y
X
−8
−6
−4
−2
0
2
4
6
8

Linear Independence vs. Dependence



A basis is a set of linearly independent vectors that span a vector space, providing a coordinate system for that space

The span of a set of vectors is the set of all possible linear combinations of those vectors



$(3,2) \in \mathbb{R}^2$

$b_1 \ (1,0)$

$b_2 \ (0,1)$

$(3,2) = 3b_1 + 2b_2$

Orthogonality and Orthonormal Sets

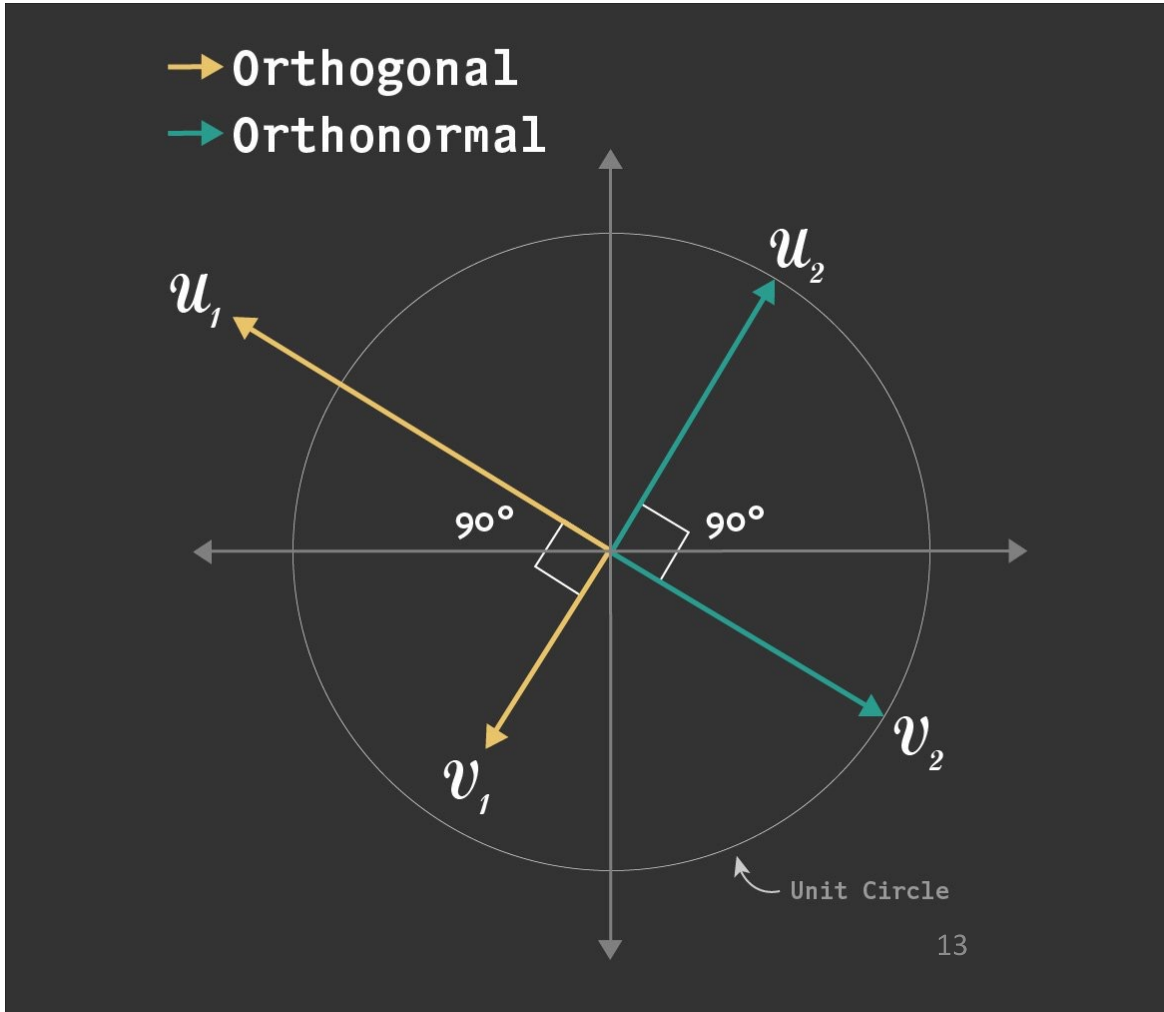
Two vectors u and v are orthogonal if their dot product is zero: u·v=0

An orthogonal set is a collection of vectors where each pair of distinct vectors is orthogonal

If all vectors in an orthogonal set have unit length (norm = 1), then it's an orthonormal set

The Gram-Schmidt process converts a linearly independent set of vectors into an orthogonal or orthonormal set.

This is essential for QR decomposition



What is a Matrix?

A matrix is a rectangular array of numbers arranged in rows and columns

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Each **column of A** is a **vector in $\mathbb{R}$**

Each **row of A** is a **vector in $\mathbb{R}^n$**

Special Types of Matrices

Identity matrix:

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

Diagonal matrix: All non-diagonal elements are zero

Symmetric matrix: $A=A^T$

Basic Matrix Operations

a) Addition/Subtraction

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$$

$$A + B = \begin{bmatrix} 1+5 & 2+6 \\ 3+7 & 4+8 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}$$

b) Scalar Multiplication

$$2A = 2 \times \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}$$

c) Transpose

$$A^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$$

d) Matrix Multiplication

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$$

$$A\mathbf{x} = \begin{bmatrix} 1*5+2*6 \\ 3*5+4*6 \end{bmatrix} = \begin{bmatrix} 5+12 \\ 15+24 \end{bmatrix} = \begin{bmatrix} 17 \\ 39 \end{bmatrix}$$

Row Matrix	Only one row ($1 \times n$)	$[8\ 9\ 7\ 6]$ (order 1×4)
Column Matrix	Only one column ($m \times 1$)	$\begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}$ (order 3×1)
Zero/Null Matrix	All elements are zero	$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$
Singleton Matrix	Only one element (1×1)	$[4]$
Square Matrix	Rows = Columns ($n \times n$)	$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$
Rectangular Matrix	Rows ≠ Columns ($m \times n, m \neq n$)	$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$
Diagonal Matrix	Square; non-diagonal elements are zero	$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$

Scalar Matrix	Diagonal with all diagonal elements equal	$\begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$
Identity (Unit) Matrix	Diagonal elements are 1, rest are 0	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
Upper Triangular Matrix	Square; all elements below main diagonal are zero	$\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$
Lower Triangular Matrix	Square; all elements above main diagonal are zero	$\begin{bmatrix} 1 & 0 \\ 4 & 3 \end{bmatrix}$
Symmetric Matrix	Square; $A = A^T$	$\begin{bmatrix} 2 & 3 \\ 3 & 4 \end{bmatrix}$
Skew-Symmetric Matrix	Square; $A = -A^T$	$\begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix}$

Hermitian Matrix	Square; $A = A^\dagger$ (conjugate transpose)	$\begin{bmatrix} 2 & i \\ -i & 3 \end{bmatrix}$
Skew-Hermitian Matrix	Square; $A = -A^\dagger$	$\begin{bmatrix} 0 & 2+i \\ -2+i & 0 \end{bmatrix}$
Orthogonal Matrix	Square; $AA^T = I$	$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$

$$\begin{bmatrix}2&1&3\\3&1&4\\5&7&12\end{bmatrix}\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} = x_1\begin{bmatrix}2\\3\\5\end{bmatrix}+x_2\begin{bmatrix}1\\1\\7\end{bmatrix}+x_3\begin{bmatrix}3\\4\\12\end{bmatrix}$$

A

Two independent column

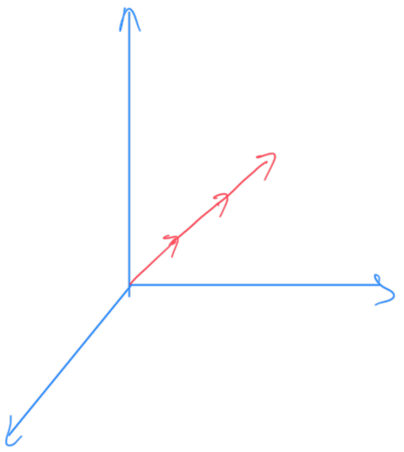
Rank $(A) = 2$

all Ax = column space $C(A)$

$$A = \begin{bmatrix}1&3&8\\1&3&8\\1&3&8\end{bmatrix} \qquad C(A) = \text{line} = \begin{bmatrix}1\\1\\1\end{bmatrix}\begin{bmatrix}1&3&8\end{bmatrix}$$

rank $(A) = 1$

$$uv^T$$



Basis

Row Reduced Echelon Form

Basis for column spaces

$$\underset{A}{\begin{bmatrix} 2 & 1 & 3 \\ 3 & 1 & 4 \\ 5 & 7 & 12 \end{bmatrix}} = \underset{C_{3\times 2}}{\begin{bmatrix} 2 & 1 \\ 3 & 1 \\ 5 & 7 \end{bmatrix}} \underset{R_{2\times 3}}{\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}}$$

$\rightarrow$ basis for row space

Column space of $A = C(A) \rightarrow$ Range

Column rank = 2

= row rank

Row space of A = Column space of $A^T = C(A^T)$

Matrix Multiplication AB

$$\begin{bmatrix} \cdots\cdots & \\ \text{row of } A & \\ & \end{bmatrix} \begin{bmatrix} \text{col} & \vdots \\ \text{of} & \vdots \\ B & \vdots \end{bmatrix} = \left.\begin{matrix} \text{dot product} \\ r \cdot c \end{matrix}\right\} \text{ Basic way}$$

Need deeper way !!!

New way → columns × rows

$$AB = \begin{bmatrix} \text{col } k \end{bmatrix} \begin{bmatrix} \text{row } k \end{bmatrix} = \begin{pmatrix} \text{col } k \\ \text{of } A \end{pmatrix} \begin{pmatrix} \text{row } k \\ \text{of } B \end{pmatrix}$$

$$A = \begin{bmatrix} 1 & 3 & 8 \\ 1 & 3 & 8 \\ 1 & 3 & 8 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 8 \end{bmatrix}$$

Matrix decomposition (Factorisations)

$A = LU \longrightarrow$ elimination $\longrightarrow$ Solving linear system $\longrightarrow$ Lower Δ matrix $\times$ Upper Δ matrix

$A = QR \longrightarrow$ Gram - Schmidt

$$S = Q \Lambda Q^T = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_n^T \end{bmatrix}$$

$$A = X \Lambda X^{-1}$$

$A = U \Sigma V^T =$ (ortho) (diag) (ortho) $\longrightarrow$ Singular Value Decomposition (SVD)

Fundamental Subspaces

Column space $C(A)$ $\dim = r$

Row space $C(A^T)$ $\dim = r$

Null space $N(A)$ $\dim = n - r$

Null space $N(A^T)$ $\dim = m - r$

$$A = \begin{bmatrix} 1 & 2 & 4 \\ 2 & 4 & 8 \end{bmatrix} \quad \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 4 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$m = 2$ $n = 3$ $r = 1$

$\dim N(A) = n - r = 2$

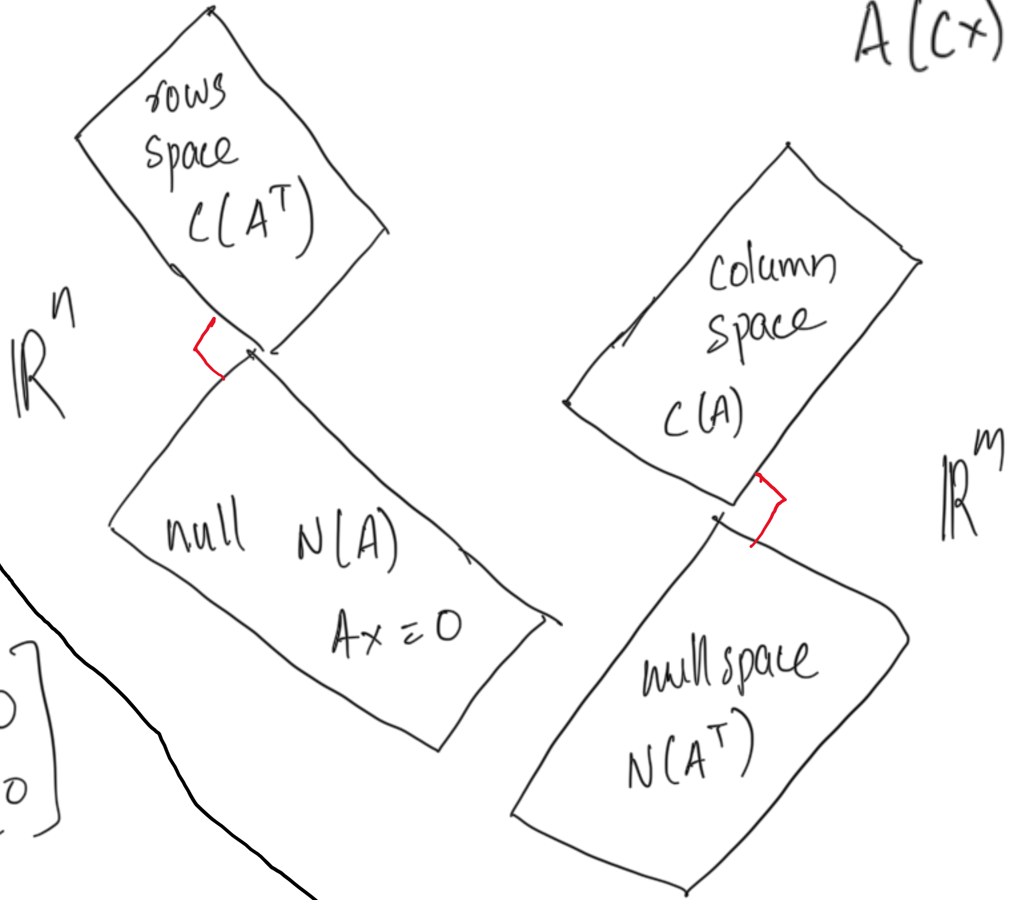
A $m \times n$ rank r

Nullspace = all solutions to $Ax = 0$

vector

$A(cx) = 0$ $Ay = 0$

$A(x+y) = 0$



$m \times n$

2×3

row → n components

columns → m components

GNEST305

LinkedIn

Resources

Shankar J
+91 9344775577
https://shankar-jayaraj.github.io/