Koneru Lakshmaiah Education Foundation
(Category -1, Deemed to be University estd. u/s. 3 of the UGC Act, 1956)
❖ Approved by AICTE ❖ ISO 21001:2018 Certified
Campus: Green Fields, Vaddeswaram - 522 302, Guntur District, Andhra Pradesh, INDIA.
Phone No. +91 8645 - 350 200; www.klef.ac.in; www.klef.edu.in; www.kluniversity.in
Admin Off: 29-36-38, Museum Road, Governorpet, Vijayawada - 520 002. Ph: +91 - 866 - 3500122, 2576129

Course Name: Mathematics for AI

Session- 3 Linearly Independent & Dependent Vectors:

Definition: A set of vectors is **linearly independent** if the only way to write

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n = \vec{0}$$

is when

$$c_1 = c_2 = \cdots = c_n = 0.$$

This means **no vector in the set can be written as a combination of the others**.

Example:

$$\vec{v}_1 = (1,0), \vec{v}_2 = (0,1)$$

The equation $c_1(1,0) + c_2(0,1) = (0,0)$, gives c1=0,c2=0 .

So, these vectors are **linearly independent**.

Definition: A set of vectors is **linearly dependent** if there exist constants, **not all zero**, such that

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n = \vec{0}.$$

This means **at least one vector can be written as a combination of the others**.

Example:$\vec{v}_1 = (1,2), \vec{v}_2 = (2,4)$

Notice that

$$\vec{v}_2 = 2\vec{v}_1$$

So

$$2\vec{v}_1 - \vec{v}_2 = (0,0)$$

Since the coefficients are not all zero, these vectors are **linearly dependent**.

Example 1 (2D): $\vec{v}_1 = (2,3), \vec{v}_2 = (1,-1)$

There is **no scalar** k such that $(2,3) = k(1,-1)$, so neither vector is a multiple of the other.

Example 2 (3D): $\vec{v}_1 = (1,0,0), \vec{v}_2 = (0,1,0), \vec{v}_3 = (0,0,1)$

These are the standard basis vectors in $\mathbb{R}^3$.

The only solution to $c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 = (0,0,0)$ is $c_1 = c_2 = c_3 = 0$.

Example 1 (2D): $\vec{v}_1 = (3,6), \vec{v}_2 = (-1,-2)$

Notice that $\vec{v}_1 = -3\vec{v}_2$, So one vector is a scalar multiple of the other.

Example 2 (3D): $\vec{v}_1 = (1,2,3), \vec{v}_2 = (2,4,6), \vec{v}_3 = (1,0,1)$,

Since $\vec{v}_2 = 2\vec{v}_1$

one vector depends on another, regardless of $\vec{v}_3$.

Rank of a Matrix
Definition: The **rank of a matrix** is the **maximum number of linearly independent rows or columns** of the matrix.
Equivalently,

- Rank = dimension of the column space
- Rank = dimension of the row space

It represents the **amount of non-redundant information** in the matrix.
Example 1 (Solved)

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

Solution

- Second row = 2 × first row
- Rows are linearly dependent

$$\boxed{\mathrm{rank}(A) = 1}$$

Interpretation
Only **one independent direction** exists → redundancy present.
Example 2 (Solved)

$$B = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

Solution

- Two non-zero independent rows
- Third row is zero

$$\boxed{\mathrm{rank}(B) = 2}$$

Interpretation
Matrix contains **two independent features**.

- **Basis of a Vector Space**

Definition: A **basis** of a vector space is a set of vectors that:

1. Are **linearly independent**
2. **Span** the vector space

A basis is the **minimum set of vectors needed to represent all vectors** in the space.
Example 1: Consider vectors in $\mathbb{R}^2$:

$$v_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Solution

- v_1, v_2 are linearly independent
- Any vector (x, y) can be written as

$$xv_1 + yv_2$$

$$\boxed{\{v_1, v_2\} \text{ is a basis of } \mathbb{R}^2}$$

Example 2 :

$$v_1 = \begin{bmatrix}1\\2\\3\end{bmatrix}, v_2 = \begin{bmatrix}2\\4\\6\end{bmatrix}, v_3 = \begin{bmatrix}1\\0\\1\end{bmatrix}$$

Solution

- $v_2 = 2v_1 \rightarrow$ dependent
- v_1 and v_3 are independent

$$\boxed{\{v_1, v_3\} \text{ is a basis of the space spanned}}$$

Orthogonal Vectors

Definition: Two vectors are said to be **orthogonal** if their **dot product is zero**.

$$v \cdot w = 0$$

Orthogonal vectors represent **independent, non-overlapping information**.

Example 1:

$$v = \begin{bmatrix}1\\2\end{bmatrix}, w = \begin{bmatrix}2\\-1\end{bmatrix}$$

Solution

$$v \cdot w = (1)(2) + (2)(-1) = 2 - 2 = 0$$

$$\boxed{v \text{ and } w \text{ are orthogonal}}$$

Example 2:

$$v = \begin{bmatrix}1\\0\\0\end{bmatrix}, w = \begin{bmatrix}0\\3\\0\end{bmatrix}$$

Solution

$$v \cdot w = 0$$

$$\boxed{v \text{ and } w \text{ are orthogonal}}$$

Case Study: **Detecting Redundant Sensor Readings Using Matrix Rank**

An industrial system uses **four sensors** S_1, S_2, S_3, S_4 to monitor a physical process.
For a certain time interval, the sensor readings (after normalization) are represented by the columns of the matrix

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

Each column corresponds to the readings of one sensor across three-time instants. Identify which sensor readings are **redundant**. Also, determine the **minimum number of sensors** required to represent all sensor information without loss.

Basically, redundant means which is not necessary
In data science, a **redundant feature** is a variable that provides no additional information to a model beyond what is already provided by other variables.
Calling a blank sheet as empty
Temperature of a room in K and Celsius.

The Real-Life Example: The "Weather Station"

Imagine you are building a machine learning model to predict **Ice Cream Sales**. You collect data from a weather station that gives you three features:

1. **Temperature in Celsius** (x_1)
2. **Temperature in Fahrenheit** (x_2)
3. **Wind Speed** (x_3)

Why is there redundancy?

The **Temperature in Fahrenheit** is a redundant feature. Why? Because it doesn't tell the model anything new that the Celsius column hasn't already said. They are just the same physical reality expressed in different scales.

Redundant: Temperature in Fahrenheit (it's just a math transformation of Celsius).

Useful/Unique: Wind Speed (it measures something entirely different from temperature).

Methods to identify Redundancy:

The most common way to identify and quantify redundancy is through mathematics, specifically **Linear Dependence** and **Correlation Coefficients**.

Linear Dependence (Linear Algebra)

In a matrix of data, features are represented as **vectors**. If one feature vector is a scalar multiple of another (e.g., $V_2 = 2\ V_1$), they are **linearly dependent**. In mathematical terms, the "Rank" of your data matrix doesn't increase when you add a redundant feature.

Solution:

Concept: Matrix Rank, LI and LD.
How? We will see
First let us discuss the required concepts.

Definition: Matrix Rank

The **rank of a matrix** is defined as the **maximum number of linearly independent rows or columns** of the matrix.

Equivalently,

- Rank = dimension of the row space
- Rank = dimension of the column space

Key Points (Exam-Friendly)

- Row rank = Column rank
- Rank indicates the **true dimensionality** of the data
- A matrix has **full rank** if its rank equals the smaller of the number of rows or columns

Example 1: Matrix with Full Rank

Consider the matrix

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

Rows are linearly independent and Columns are linearly independent, hence $\text{rank}(A) = 2$

Example 2: Matrix with Dependent Rows and Columns

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

Second row = 2 × first row
Second column = 2 × first column

$$\text{rank}(B) = 1$$

Example 3: Rectangular Matrix

$$C = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix}$$

- Second row is a multiple of the first
- All columns depend on one vector

$$\text{rank}(C) = 1$$

Interpretation

- **High rank** → more independent information
- **Low rank** → redundancy present

Consider the vectors in $\mathbb{R}^2$:

$$v_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, v_2 = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

Step 1: Form a Linear Combination

Check whether there exist scalars c_1, c_2, not both zero, such that

$$c_1 v_1 + c_2 v_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

That is,

$$c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 3 \\ 6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Step 2: Solve the System

$$\begin{aligned} c_1 + 3c_2 &= 0 \\ 2c_1 + 6c_2 &= 0 \end{aligned}$$

The second equation is just **twice** the first, so there are infinitely many non-zero solutions, for example:

$$c_1 = -3, c_2 = 1$$

Case Study: Detecting Redundant Sensor Readings Using Matrix Rank

An industrial system uses four sensors $S_1, S_2, S_3, S_4$to monitor a physical process. For a certain time interval, the sensor readings (after normalization) are represented by the columns of the matrix

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

Each column corresponds to the readings of one sensor across three-time instants. Identify which sensor readings are redundant. Also, determine the minimum number of sensors required to represent all sensor information without loss.

Step 1: Observe Linear Dependence

- Column 2 is **2 times** column 1
- Column 3 is **3 times** column 1
- Column 4 is **equal** to column 1

Thus, all columns are scalar multiples of the first column.

Step 2: Determine the Rank

Since all columns depend on one column,

$$\text{rank}(A) = 1$$

Step 3: Identify Redundant Sensors

- Only one sensor provides independent information
- Sensors S_2, S_3,and $S_4$are **redundant**

Step 4: Minimum Sensors Required

$$\boxed{1 \text{ sensor}}$$

is sufficient to represent all readings without information loss.

Step 5: Interpretation

Matrix rank reveals the **true dimensionality** of sensor data.
A rank smaller than the number of sensors indicates **redundancy**.

Final Answer

Using matrix rank, redundant sensor readings are identified and removed, leading to efficient data representation and reduced computational cost.

Case Study 2 : Removing Redundant Features in Smart City Air-Quality Monitoring

A **smart city** installs multiple sensors at each monitoring station to assess **air quality**.
For one station, the following **six features** are recorded every hour:

1. f_1: Concentration of Carbon Monoxide (CO)
2. f_2: Concentration of Nitrogen Oxides (NO)
3. f_3: Air Quality Index (AQI)
4. f_4: Weighted pollution score
5. f_5: Temperature
6. f_6: Heat index

The city wants to **reduce storage and computation costs** by removing **redundant features**, without losing information.

Mathematical Representation

Each feature is represented as a **column vector** of sensor readings over time:

$$X = [f_1 \ f_2 \ f_3 \ f_4 \ f_5 \ f_6]$$

All vectors belong to the vector space $\mathbb{R}^n$, where $n$is the number of time samples.

Given Relationships (from environmental science)

- AQI is computed using CO and NO
- Weighted pollution score is a linear combination of AQI and pollutant concentrations
- Heat index is computed from temperature

Mathematically:

$$\begin{aligned} f_3 &= af_1 + bf_2 \\ f_4 &= cf_1 + df_2 + ef_3 \\ f_6 &= kf_5 \end{aligned}$$

Step-by-Step Solution Using Mathematical Concepts

Linear Independence (Detecting Redundancy)

A set of vectors is **linearly independent** if none can be written as a linear combination of others.
From the given relations:

- $f_3$depends on f_1, f_2

- f_4 depends on f_1, f_2, f_3
- f_6 depends on f_5

Hence,

$$\{f_1, f_2, f_3, f_4, f_5, f_6\}$$

is a **linearly dependent set**.

How this helps?

Linear dependence **mathematically proves** that some features do not add new information.

Rank (Finding True Dimensionality)

Although there are **6 features**, only these are independent:

$$\{f_1, f_2, f_5\}$$

Therefore,

$$\text{rank}(X) = 3$$

How this helps

Rank tells us the **true number of informative features**, independent of how many are recorded.

Basis (Selecting Essential Features)

A **basis** is a minimal linearly independent set that spans the same feature space.

Choose the basis:

$$\mathcal{B} = \{f_1 \text{ (CO)}, f_2 \text{ (NO)}, f_5 \text{ (Temperature)}\}$$

All other features can be generated from this basis.

How this helps

Basis selection gives a **systematic rule** for which features to keep and which to remove.

Orthogonality (Avoiding Overlap of Information)

The original features are **not orthogonal**, because:

- AQI overlaps with CO and NO
- Heat index overlaps with temperature

Non-orthogonal features have **high correlation**.

By keeping only, the basis features (or orthogonalizing them if needed), each retained feature contributes **distinct information**.

How this helps

Orthogonality ensures that no retained feature duplicates information from another.

Final Result: Redundant Feature Removal

Removed Features

- AQI
- Weighted pollution score
- Heat index

Retained Features

- CO concentration
- NO concentration
- Temperature

Dimensionality Reduction: $6 \rightarrow 3$

Final Conclusion :

In this smart-city air-quality case study, redundant features arise because several sensor variables are linear combinations of others. Linear independence identifies redundancy, rank reveals true dimensionality, basis selection determines essential features, and orthogonality ensures non-overlapping information, enabling effective removal of redundant features without loss of data.

Case study 3: A smart building uses **four sensors** to monitor environmental conditions.
Each sensor records readings at **three-time instants**.
The sensor data (after normalization) is represented by the matrix

$$A = \begin{bmatrix} 1 & 2 & 3 & 1 \\ 2 & 4 & 6 & 2 \\ 3 & 6 & 9 & 3 \end{bmatrix}$$

Each **column** of the matrix corresponds to a sensor:

- S_1, S_2, S_3, S_4

The objective is to **identify and remove redundant sensors** without losing information.