Page No.
Date

Why Hyperplanes matter in Datasets

- Each student's performance is a point in a high-dimensional hyperplane
- Decision boundaries in LDA are hyperplanes separating "Above Average" and "Below Average"
- PCA: Project data from a high-dimensional hyperplane to a lower-dimensional plane (2D).

Correlations

- A line models performance in one subject
- A plane models performance in two subjects
- A hyperplane models real-world students performance across many subjects
- Increasing dimensionality provides richer information but requires linear algebra techniques to analyze effectively

Interpretation:

Human visualization is no longer possible, but mathematical tools such as PCA, LDA, SVD help analyze patterns.

Dimensionality growth:
2D → 3D → nD

No. of Subjects	Dimentions	Geometric form
1	1D	Line
2	2D	Plane
3	3D	3D Space
n > 3	nD	Hyperplane

As the number of increases:

- Information captured increases
- Visualization becomes difficult.
- Dimensionality reduction becomes ~~es~~ essential

- If we select three subjects
- Each student is represented as:

$$[\mathbf{x} = [x_1, x_2, x_3]]$$

Student now lie in a 3D coordinate space

Interpretion:

Patterns in student performance become harder to visualise but capture more information

4 Hyperplane (Higher Dimention)

A hyperplane is a generalization of a line and plane to n dimensions
In our dataset

- With n subjects, each student is a point in n-dimensional space:

$$[\mathbf{x} = [x_1, x_2, \dots, x_n]]$$

All student collectively form a cloud of point in an n-dimention hyperplane

2 Plane (2D)

A plane exists in two-dimensional space

In our dataset:

- If we select two subjects
- Each student is represented as:

$[\mathbf{x} = [x_1, x_2]]$

Each student becomes a point on a 2D plane.

Interpretation:

We can visually compare students using a scatter plot, observing correlation between two subjects

3 3D Space (3D)

A three-dimensional space is defined by three independent axes

In our dataset:

Partb 5

Page No.
Date

LINE, Plane, and Hyperplane in the context of Student Performance Dataset

In this dataset, each student is represented by vectors of subject scores.

If a student has scores in n subjects, then that student corresponds to a point in n-dimensional space

1 LINE (1D)

A line exists in one-dimensional space

In our dataset:

- If we consider only one subject
- Each student has a single value

$$[\mathbf{x} = [x]]$$

All students lie along a single line (number line)

Interpretation:

Performance is compared using only one criterion