

Statistics and Data Analysis

Unit 05 – Lecture 04 Notes

Tofik Ali

February 14, 2026

Topic

Factor analysis (latent factors) and LDA (supervised separation).

Learning Outcomes

- Explain factor analysis as latent-factor modeling (intuition)
- Differentiate PCA vs factor analysis (goal/assumptions)
- Explain LDA as supervised dimensionality reduction/classifier
- Interpret a 2D LDA projection

Detailed Notes

These notes are designed to be read alongside the slides. They expand each slide bullet into plain-language explanations, small worked examples, and common pitfalls. When a formula appears, emphasize (1) what each symbol means, (2) the assumptions needed to use it, and (3) how to interpret the final number in the problem context.

Factor Analysis

- Observed variables driven by a few latent factors
- Goal: explain correlations via factors
- Used for surveys/constructs

LDA

- Supervised: uses labels
- Finds projection maximizing class separation
- Can classify and visualize

Exercises (with Solutions)

Exercise 1: Supervised?

Is PCA supervised? Is LDA supervised?

Solution

- PCA is unsupervised; LDA is supervised.

Exercise 2: Goal

What does PCA optimize vs LDA (intuition)?

Solution

- PCA: variance captured; LDA: class separability.

Exercise 3: Use case

Labeled A/B/C data, want 2D plot separating classes. PCA or LDA?

Solution

- LDA (uses labels for separation).

Exit Question

Why can LDA separate classes better than PCA on labeled data?

Demo (Python)

Run from the lecture folder:

```
python demo/demo.py
```

Output files:

- `images/demo.png`
- `data/results.txt`

References

- Montgomery, D. C., & Runger, G. C. *Applied Statistics and Probability for Engineers*, Wiley.
- Devore, J. L. *Probability and Statistics for Engineering and the Sciences*, Cengage.
- McKinney, W. *Python for Data Analysis*, O'Reilly.