Statistics and Data Analysis
Unit 05 – Lecture 04 Notes
Factor Analysis and Discriminant Analysis (LDA)

Tofik Ali

February 17, 2026

Topic

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

How to Use These Notes

These notes are written for students who are seeing the topic for the first time. They follow the slide order, but add the missing 'why', interpretation, and common mistakes. If you get stuck, look at the worked exercises and then run the Python demo.

Course repository (slides, demos, datasets): `https://github.com/tali7c/Statistics-and-Data-Analysis`

Time Plan (55 minutes)

- 0–10 min: Attendance + recap of previous lecture
- 10–35 min: Core concepts (this lecture's sections)
- 35–45 min: Exercises (solve 1–2 in class, rest as practice)
- 45–50 min: Mini demo + interpretation of output
- 50–55 min: Buffer / wrap-up (leave 5 minutes early)

Slide-by-slide Notes

Title Slide

State the lecture title clearly and connect it to what students already know. Tell students what they will be able to do by the end (not just what you will cover).

Quick Links / Agenda

Explain the structure of the lecture and where the exercises and demo appear.

- Overview

- Factor Analysis
- LDA
- Exercises
- Demo
- Summary

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

Why these outcomes matter. Always state assumptions clearly. Common assumptions in classical tests include independence of observations, roughly normal errors (or a large-sample justification), and similar variances across groups. Violations do not automatically invalidate a result, but they change how much you should trust the p-value and confidence interval. **PCA** finds new axes (principal components) that capture maximum variance. It is a rotation of the feature space. Because PCA is variance-based, it is sensitive to scaling: standardize features first unless all features are already comparable.

Factor Analysis: Key Points

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

Explanation. Correlation measures the strength of a linear association between two variables. It is symmetric (no X/Y direction) and does not imply causation. Outliers can inflate or hide correlation, so always look at the scatter plot.

LDA: Key Points

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

Exercises (with Solutions)

Attempt the exercise first, then compare with the solution. Focus on interpretation, not only arithmetic.

Exercise 1: Supervised?

Is PCA supervised? Is LDA supervised?

Solution

- PCA is unsupervised; LDA is supervised.

Walkthrough. **PCA** finds new axes (principal components) that capture maximum variance. It is a rotation of the feature space. Because PCA is variance-based, it is sensitive to scaling: standardize features first unless all features are already comparable.

Exercise 2: Goal

What does PCA optimize vs LDA (intuition)?

Solution

- PCA: variance captured; LDA: class separability.

Walkthrough. **PCA** finds new axes (principal components) that capture maximum variance. It is a rotation of the feature space. Because PCA is variance-based, it is sensitive to scaling: standardize features first unless all features are already comparable.

Exercise 3: Use case

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

Solution

- LDA (uses labels for separation).

Walkthrough. **PCA** finds new axes (principal components) that capture maximum variance. It is a rotation of the feature space. Because PCA is variance-based, it is sensitive to scaling: standardize features first unless all features are already comparable.

Mini Demo (Python)

Run from the lecture folder:

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

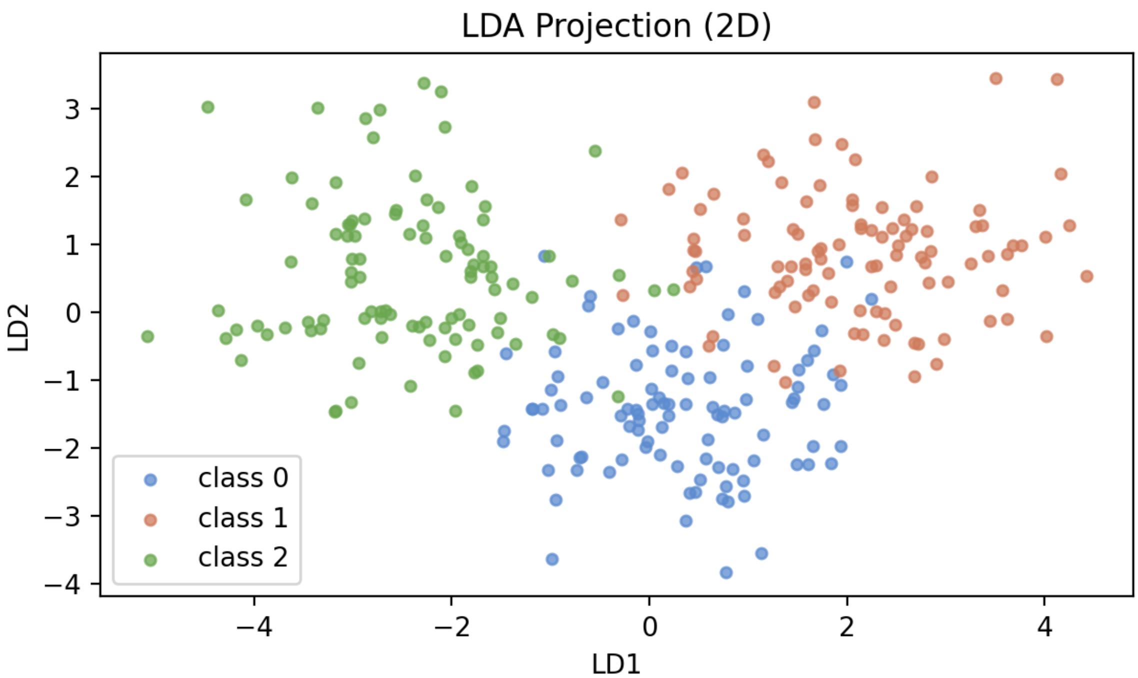
Output files:

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

What to show and say.

- Illustrates factor-like structure and an LDA-style supervised projection.
- Use it to compare 'variance capture' (PCA) vs 'class separation' (LDA).
- Discuss when interpretability vs prediction is the goal.

Demo Output (Example)



Summary

- Key definitions and the main formula.
- How to interpret results in context.
- How the demo connects to the theory.

Exit Question

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

Suggested answer (for revision). LDA uses labels to maximize class separation, while PCA ignores labels and only maximizes variance, so LDA can separate classes better.

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.

Appendix: Slide Deck Content (Reference)

The material below is a reference copy of the slide deck content. Exercise solutions are explained in the main notes where applicable.

Title Slide

https://github.com/tali7c/Statistics-and-Data-Analysis

Quick Links

Overview Factor Analysis LDA Exercises Demo Summary

Agenda

- Overview
- Factor Analysis
- LDA
- Exercises
- Demo
- Summary

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

Factor Analysis: Key Points

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

LDA: Key Points

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

Exercise 1: Supervised?

Is PCA supervised? Is LDA supervised?

Solution 1

- PCA is unsupervised; LDA is supervised.

Exercise 2: Goal

What does PCA optimize vs LDA (intuition)?

Solution 2

- PCA: variance captured; LDA: class separability.

Exercise 3: Use case

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

Solution 3

- LDA (uses labels for separation).

Mini Demo (Python)

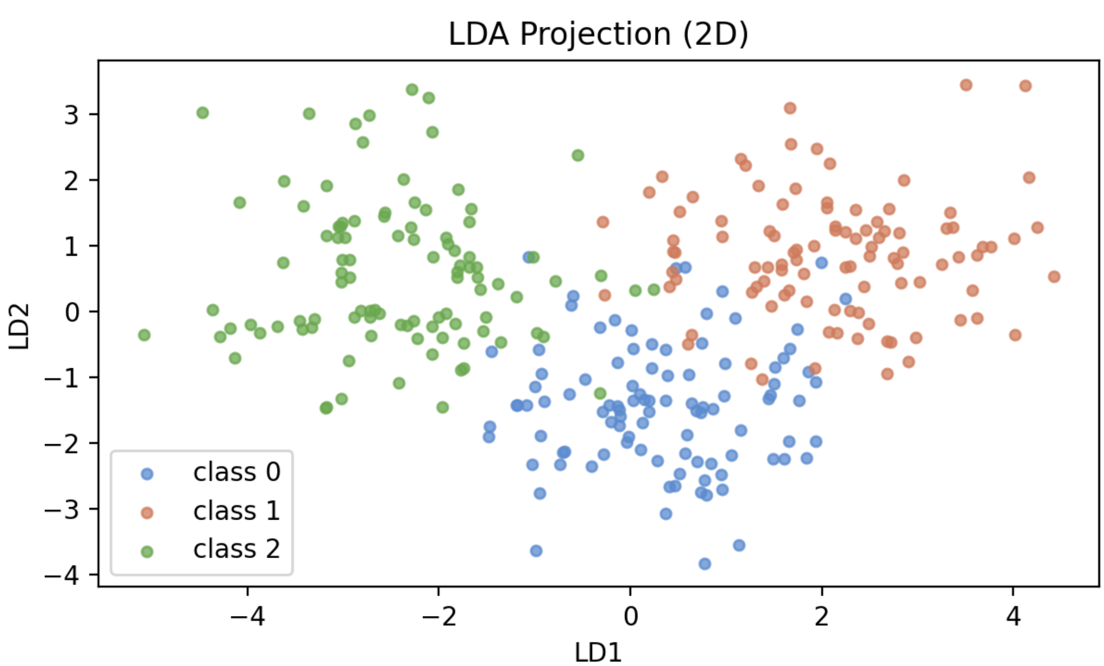
Run from the lecture folder:

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

Outputs:

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

Demo Output (Example)



Summary

- Key definitions and the main formula.
- How to interpret results in context.
- How the demo connects to the theory.

Exit Question

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