Statistics and Data Analysis
Unit 05 – Lecture 05 Notes
Kernel PCA and t-SNE

Tofik Ali

February 17, 2026

Topic

Nonlinear dimensionality reduction: kernel PCA and t-SNE (visualization).

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

- Kernel PCA
- t-SNE
- Exercises
- Demo
- Summary

Learning Outcomes

- Explain why nonlinear methods are sometimes needed
- Describe kernel PCA idea (high level)
- Describe t-SNE purpose (visualization) and pitfalls
- Choose PCA vs t-SNE appropriately

Why these outcomes matter. **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. **t-SNE** is mainly for visualization. It preserves local neighborhood structure but can distort global distances. Do not treat t-SNE plots as proof of clusters; treat them as exploratory pictures that need validation.

Kernel PCA: Key Points

- Implicitly map to higher-dimensional space via kernel
- Apply PCA in that space
- Captures nonlinear structure

Explanation. **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.

t-SNE: Key Points

- Mainly for 2D/3D visualization
- Preserves local neighborhoods
- Global distances can be misleading

Explanation. **t-SNE** is mainly for visualization. It preserves local neighborhood structure but can distort global distances. Do not treat t-SNE plots as proof of clusters; treat them as exploratory pictures that need validation.

Exercises (with Solutions)

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

Exercise 1: Use case

Name one warning when interpreting t-SNE plots.

Solution

- Global distances and cluster sizes can be misleading.

Walkthrough. **t-SNE** is mainly for visualization. It preserves local neighborhood structure but can distort global distances. Do not treat t-SNE plots as proof of clusters; treat them as exploratory pictures that need validation.

Exercise 2: Randomness

What should you do if t-SNE changes across runs?

Solution

- Set seed and check stability.

Walkthrough. **t-SNE** is mainly for visualization. It preserves local neighborhood structure but can distort global distances. Do not treat t-SNE plots as proof of clusters; treat them as exploratory pictures that need validation.

Exercise 3: Kernel PCA benefit

Why kernel PCA can help on circular data?

Solution

- It can capture nonlinear manifold structure.

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
```

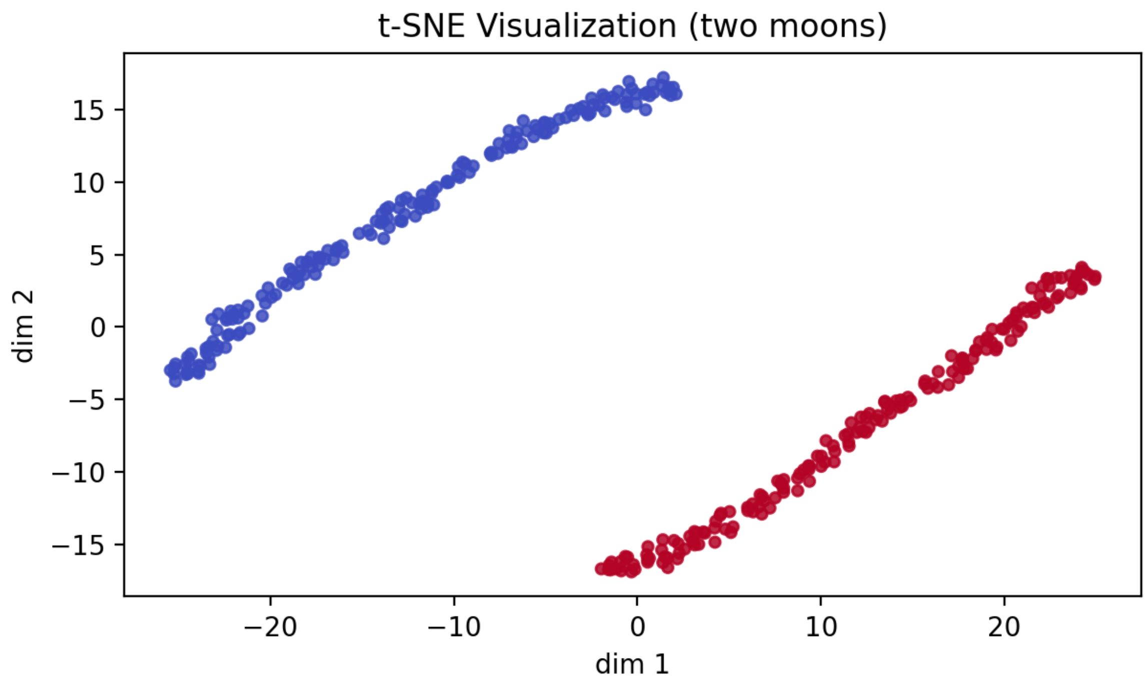
Output files:

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

What to show and say.

- Creates a nonlinear dataset and compares PCA-like vs nonlinear embeddings.
- Produces a 2D visualization to discuss neighborhood preservation.
- Use it to warn that t-SNE is mainly for visualization, not modeling features.

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 should we avoid using t-SNE coordinates directly as model features (usually)?

Suggested answer (for revision). t-SNE coordinates are not stable/global-metric features and can distort distances; it is primarily a visualization tool, not a feature generator.

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

Quick Links

Overview Kernel PCA t-SNE Exercises Demo Summary

Agenda

- Overview
- Kernel PCA
- t-SNE
- Exercises
- Demo
- Summary

Learning Outcomes

- Explain why nonlinear methods are sometimes needed
- Describe kernel PCA idea (high level)
- Describe t-SNE purpose (visualization) and pitfalls
- Choose PCA vs t-SNE appropriately

Kernel PCA: Key Points

- Implicitly map to higher-dimensional space via kernel
- Apply PCA in that space
- Captures nonlinear structure

t-SNE: Key Points

- Mainly for 2D/3D visualization
- Preserves local neighborhoods
- Global distances can be misleading

Exercise 1: Use case

Name one warning when interpreting t-SNE plots.

Solution 1

- Global distances and cluster sizes can be misleading.

Exercise 2: Randomness

What should you do if t-SNE changes across runs?

Solution 2

- Set seed and check stability.

Exercise 3: Kernel PCA benefit

Why kernel PCA can help on circular data?

Solution 3

- It can capture nonlinear manifold structure.

Mini Demo (Python)

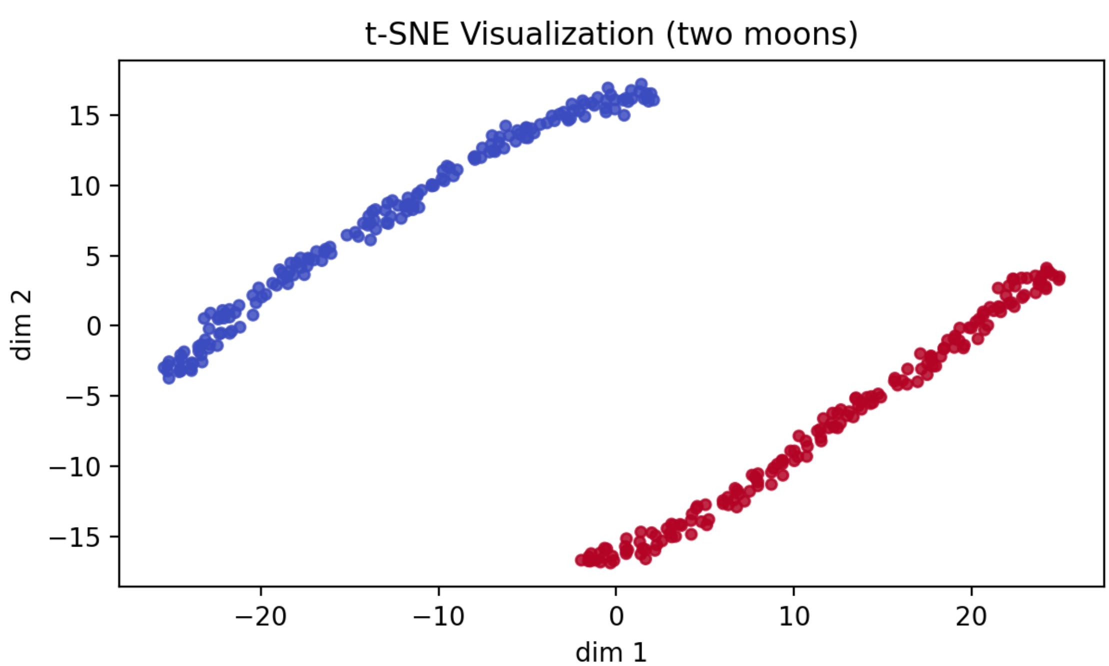
Run from the lecture folder:

```
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 should we avoid using t-SNE coordinates directly as model features (usually)?