Reading Report:
A Tutorial on Principal Component Analysis

Boyuan Du (2024151470021)
School of Software, **Sichuan University**
Email:2024151470021@stu.scu.edu.cn

September 19, 2025

Course / Assignment: Reading report on PCA .
Primary Source: J. Shlens, "A Tutorial on Principal Component Analysis," 2014.

Abstract

One paragraph (150–200 words): objective of PCA, main insights from the tutorial, and your MNIST takeaway. Mention the ED/SVD equivalence briefly. PCA is a powerful and fundamental technique for extracting useful information from complex(high-dimensional) data by lowering dimension onto basic features of the data. The essence of PCA lies in its ability to identify the directions of maximum variance in the data, which are represented by the principal components. By projecting the data onto these components, PCA effectively reduces redundancy and highlights the most informative aspects of the dataset. The tutorial emphasizes two main approaches to PCA: eigen-decomposition (ED) of the covariance matrix and singular value decomposition (SVD) of the data matrix. Both methods yield equivalent results with the help of eigenvector. However they has a common weakness that they rely on the assumption that principal components are orthogonal which limits the two techniques to extract non-vertical principal component, but Independent Component Analysis can handle this problem efficiently. "about the experiment of MNIST..." **Keywords:** PCA, covariance, eigen-decomposition, SVD, variance.

Contents

1 Introduction

Briefly introduce PCA as a linear, non-parametric technique for variance-maximizing, decorrelating representation; follow the change-of-basis view in Shlens (2014).

2 Core Concepts and Questions (Q0–Q13)

Q0. Essence of a Matrix

Linear map / basis-change views; minimal example.

Q1. What Problems Does PCA Address?

Redundancy, meaningful axes of variation.

Q2. Assumptions and Limits

Linearity; variance→structure; orthogonality; failure cases.

Q3. Basis in Linear Algebra

Basis, orthonormality, projection.

Q4. Covariance and Redundancy

Cov, off-diagonal dependence, diagonal dominance.

Q5. SNR, Variance, Redundancy

Var, SNR; why diagonalize C_X.

Q6. Principal Component Meaning

Variance-maximizing directions under orthogonality.

Q7. PCA as Basis Transformation

$Y = PX$; geometric interpretation.

Q8. Re-expressing Inputs

Centering; dimensionality reduction by top-k rows of P.

Q9. PCA and Covariance

$C_Y = PC_XP^\top$; diagonalization.

Q10. PCA, ED, SVD

ED of C_X vs SVD route; equivalence and numerics.

Q11. Intuition: ED and SVD

Rotations + stretch interpretations.

Q12. Objective Function

$$\max_{\|w\|=1} w^\top C_X w \quad \Longleftrightarrow \quad \min_{\mathrm{rank}(\tilde{X})\le k} \|X - \tilde{X}\|_F^2 \text{ (Eckart–Young)} \tag{1}$$

Q13. Why Dimension Reduction Works

Order by eigenvalues; preserve most variance.

3 Experiments on MNIST (Q14)

3.1 Dataset and Setup

Use 2,000 train + 2,000 test images. Preprocess: flatten, center, optionally scale.

3.2 Method

PCA via ED or SVD. Report explained variance ratio for top-2/top-50 as sanity check.

3.3 Results: 2D Visualization

Figure 1: PCA to 2D on MNIST. Distinct colors/markers per digit class.

3.4 Analysis

Comment on class overlap; limits of linear PCA; when kernel PCA/ICA might help.

4 Discussion and Conclusion

Summarize SNR/covariance → diagonalization → ED/SVD routes; practical gains and limits.

Compliance Checklist

- Self-contained definitions (matrix, basis, projection, variance, covariance, SNR, redundancy, PC).
- High logic, concise writing.
- Q0–Q13 covered; Q14 plots and analysis present.
- Notation table completed in Section A.

A Notation (Alphabetical Summary)

B Derivations (Optional)

Sketch why eigenvectors of C_X diagonalize C_Y; Eckart–Young link.

C Reproducibility (Optional)

OS, Python/Matlab version, libs, seed, commands to regenerate figures.

Table 1: Symbols used throughout the report.

Notation	**Definition**	**Corresponding to**
$X \in \mathbb{R}^{m \times n}$	Zero-mean data (m dims, n samples)	feature matrix
$C_X = \frac{1}{n} X X^\top$	Covariance of X	covariance
$P = [p_1^\top \cdots p_m^\top]$	Orthonormal PC basis	principal axes
$Y = PX$	Data in PCA space	PCA coordinates
λ_i	Eigenvalue of C_X	explained variance
k	Target dimension	NA

References

[1] J. Shlens, *A Tutorial on Principal Component Analysis*, 2014.