PCA by implying SVD

Prepared by: Sujal Sinha
School of Physical Sciences, NISER, Bhubaneswar

November 26,2025

Abstract

This report explains Principal Component Analysis (PCA) using singular value decomposition (SVD). It shows the precise mathematics, a small numeric worked example (2×2), and clear visual diagrams drawn for visual interpretation.

1 Introduction (What PCA solves)

Principal Component Analysis (PCA) identifies the directions of greatest variability in a dataset and orders them. It reduces dimensionality while preserving the most important structure.

Given data with n observations and p features arranged in matrix $X \in \mathbb{R}^{n\times p}$, PCA finds orthogonal directions $v_1, \dots, v_p$ (principal components) such that projections onto those directions maximize variance.

Two common computational routes:

- Eigen-decomposition of the covariance matrix $C = \frac{1}{n}X^\top X$.
- Singular value decomposition (SVD) of the centered data matrix X_c.

SVD is numerically stable and is the practical method in practice.

2 Centering and Covariance

2.1 Center the data

Given data rows $x_i \in \mathbb{R}^p$, compute column means

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \in \mathbb{R}^p.$$

Center each row:

$$X_c = X - \mathbf{1}\bar{x}^\top,$$

so each column of X_c has mean zero.

2.2 Covariance matrix

The (unbiased or biased up to scale) covariance matrix is

$$C = \frac{1}{n}X_c^\top X_c \quad (\text{or } \frac{1}{n-1} \text{ depending on convention}).$$

PCA finds eigenpairs (λ_j, v_j) such that

$$Cv_j = \lambda_j v_j, \quad \lambda_1 \geq \lambda_2 \geq \cdots \geq 0.$$

Eigenvector v_j is the j-th principal direction and λ_j is the variance along that direction.

3 SVD use in PCA

3.1 SVD definition

For any $X_c \in \mathbb{R}^{n\times p}$ the singular value decomposition is

$$X_c = U\Sigma V^\top,$$

where

- $U \in \mathbb{R}^{n\times n}$ (or $n \times r$) has orthonormal columns (left singular vectors),
- $\Sigma \in \mathbb{R}^{n\times p}$ is diagonal (nonnegative) with singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq 0$,
- $V \in \mathbb{R}^{p\times p}$ is orthogonal (right singular vectors).

3.2 Relation to covariance

Compute

$$X_c^\top X_c = V\Sigma^\top U^\top U\Sigma V^\top = V\Sigma^\top \Sigma V^\top = V\Lambda V^\top,$$

where $\Lambda = \Sigma^\top\Sigma$ is diagonal with entries σ_j^2. Thus:

$$\text{eigenvalues}(X_c^\top X_c) = \{\sigma_j^2\}, \qquad \text{eigenvectors} = \text{columns of } V.$$

So PCA directions = columns of V; the variance explained by component j equals σ_j^2/n (if using $1/n$ covariance).

3.3 Why SVD is used:

- It works for rectangular matrices.
- It avoids forming $X_c^\top X_c$ explicitly (which can magnify numerical errors).
- We get both directions (V) and coordinates (UΣ).

4 Jacobi rotation (short derivation for 2×2)

The Jacobi rotation eliminates an off-diagonal element in a symmetric matrix by a plane rotation.

Let

$$A = \begin{pmatrix} a & b \\ b & d \end{pmatrix}, \qquad b \neq 0.$$

We look for an orthogonal rotation matrix

$$R(\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

such that $A' = R^\top AR$ has zero in the off-diagonal entry $A'_{12} = 0$. Working out A'_{12} gives

$$A'_{12} = (a-d)\sin\theta\cos\theta + b(\cos^2\theta - \sin^2\theta).$$

Using double-angle identities ($\sin 2\theta = 2\sin\theta\cos\theta,\ \cos 2\theta = \cos^2\theta - \sin^2\theta$), set $A'_{12} = 0$ to get

$$\frac{1}{2}(a-d)\sin 2\theta + b\cos 2\theta = 0 \quad\Longrightarrow\quad \tan 2\theta = \frac{2b}{a-d}.$$

One convenient choice for the angle is

$$\theta = \frac{1}{2} \arctan\left(\frac{2b}{a-d}\right),$$

which is what Jacobi implementation uses.

Iterating pairwise eliminations across the matrix diagonal(picking maximum off diagonal term for least runtime) produces an approximate diagonal matrix whose diagonal entries are the eigenvalues.

5 Worked 2×2 numerical example

Let

$$A = \begin{pmatrix} 4 & 3 \\ 3 & 2 \end{pmatrix}.$$

Compute $\theta = \frac{1}{2} \arctan \frac{2\cdot 3}{4-2} = \frac{1}{2} \arctan(3) \approx 0.6367$ rad. Then $c = \cos\theta \approx 0.804$ and $s = \sin\theta \approx 0.594$. Form R and compute $A' = R^\top A R$. This yields a diagonalized matrix approximately:

$$A' \approx \begin{pmatrix} 5.854 & 0 \\ 0 & 0.146 \end{pmatrix},$$

so eigenvalues $\lambda_1 \approx 5.854,\ \lambda_2 \approx 0.146$.

If this symmetric matrix were $X^\top X$, singular values of X would be $\sigma_j = \sqrt{\lambda_j}$.

6 How the algorithm works

The code computes $A^\top A$, finds eigenpairs (via Jacobi), takes square roots of eigenvalues to get singular values S, and uses the eigenvectors as V. PCA uses V and S^2 for directions and variances respectively. It then center the data and apply SVD on the centered data X_c.

7 Visual Intuitions

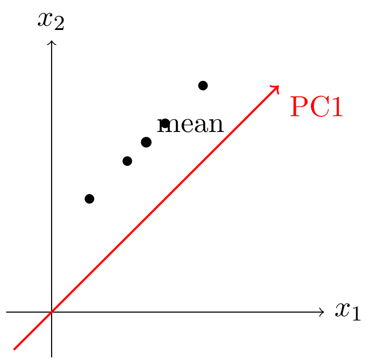

Figure 1: Data points roughly on a line; the red arrow shows the first principal component.

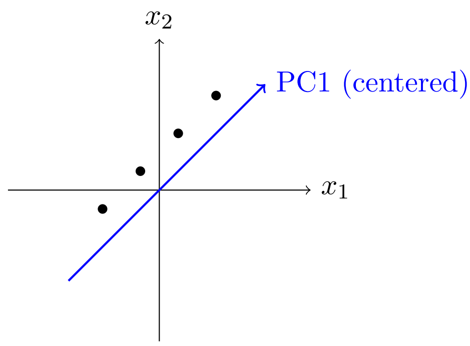

Figure 2: After subtracting the mean, the data centers at the origin; PCA finds its main axis.

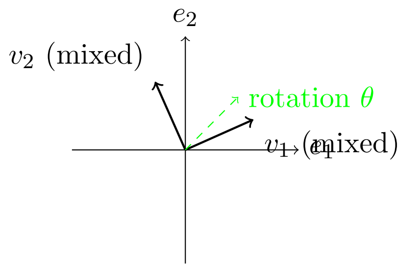

Figure 3: Jacobi rotation aligns coordinate axes to reduce off-diagonal entries in a symmetric 2×2 block.

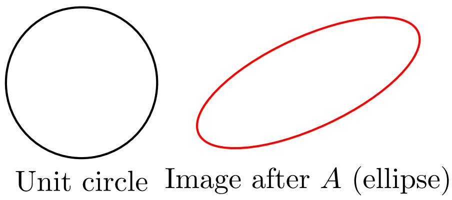

Figure 4: SVD finds the principal axes of the ellipse (directions and magnitudes of stretch).

Appendix: Key formulas

$$
\begin{aligned}
\text{Centering: } & X_c = X - \mathbf{1}\bar{x}^\top, \\
\text{Covariance: } & C = \frac{1}{n} X_c^\top X_c, \\
\text{Eigenproblem: } & Cv = \lambda v, \\
\text{SVD: } & X_c = U \Sigma V^\top, \\
\text{Relation: } & X_c^\top X_c = V \Sigma^\top \Sigma V^\top.
\end{aligned}
$$