Principal Component Analysis Made Simple: Eigenvalues and Eigenvectors

1. Introduction

This document demonstrates Principal Component Analysis (PCA), a key method in machine learning for dimensionality reduction, using linear algebra concepts of eigenvalues and eigenvectors. It focuses on the mathematical process of finding principal components to project high-dimensional data into a lower-dimensional space while preserving variance.

Using an example of predicting a student's grade based on hours studied and practice problems solved, we will show the PCA process with calculations, eigenvalue/eigenvector computations, and visualizations. The content aligns with the topic 'Linear Algebra in Machine Learning'.

2. PCA Overview

PCA is a technique to reduce the dimensionality of data while retaining most of its variance. It transforms the data into a new coordinate system defined by principal components, which are directions (eigenvectors) of maximum variance, scaled by their importance (eigenvalues).

For our example, we have data with two features:

- x_1: hours studied
- x_2: practice problems solved

Our goal is to reduce this 2D data to 1D by finding the principal component using eigenvalues and eigenvectors of the data's covariance matrix.

3. Variance and Covariance

Variance measures how much a feature varies, and covariance measures how two features vary together. PCA finds directions (eigenvectors) where the data has the most variance, indicated by eigenvalues.

The covariance matrix for two features x_1 and x_2 is:

$$\mathbf{C} = \begin{bmatrix} \mathrm{Var}(x_1) & \mathrm{Cov}(x_1, x_2) \\ \mathrm{Cov}(x_2, x_1) & \mathrm{Var}(x_2) \end{bmatrix}$$

where:

- $\mathrm{Var}(x_i) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x}_i)^2$
- $\mathrm{Cov}(x_1, x_2) = \frac{1}{n} \sum_{i=1}^{n} (x_{1i} - \bar{x}_1)(x_{2i} - \bar{x}_2)$

4. Eigenvalues and Eigenvectors

The principal components are the eigenvectors of the covariance matrix, and their corresponding eigenvalues indicate the amount of variance they capture. For a matrix $\mathbf{C}$, we solve:

$$\mathbf{C}\mathbf{v} = \lambda\mathbf{v}$$

where:

- $\mathbf{v}$: eigenvector
- λ: eigenvalue

The eigenvector with the largest eigenvalue is the first principal component, capturing the most variance.

5. Example Dataset

Here is our data (hours studied, practice problems solved, and grade for reference):

Hours Studied (x_1)	Practice Problems (x_2)	Grade (y)
1	2	2.1
2	3	2.9
3	5	4.2

We'll use only x_1 and x_2 for PCA to reduce from 2D to 1D. The grade (y) is included for context but not used in PCA calculations.

6. Step-by-Step PCA Calculation

Step 1: Center the Data

Subtract the mean of each feature to center the data around the origin:

$$\bar{x}_1 = \frac{1+2+3}{3} = 2, \quad \bar{x}_2 = \frac{2+3+5}{3} = \frac{10}{3} \approx 3.333$$

Centered data:

$$\begin{gathered}(1-2, 2-3.333) = (-1, -1.333)\\(2-2, 3-3.333) = (0, -0.333)\\(3-2, 5-3.333) = (1, 1.667)\end{gathered}$$

Step 2: Compute the Covariance Matrix

Calculate variances and covariance:

$$\text{Var}(x_1) = \frac{(-1)^2 + 0^2 + 1^2}{3} = \frac{1+0+1}{3} = \frac{2}{3} \approx 0.667$$

$$\text{Var}(x_2) = \frac{(-1.333)^2 + (-0.333)^2 + (1.667)^2}{3} = \frac{1.777 + 0.111 + 2.778}{3} = \frac{4.666}{3} \approx 1.555$$

$$\mathrm{Cov}(x_1, x_2) = \frac{(-1)(-1.333) + (0)(-0.333) + (1)(1.667)}{3} = \frac{1.333 + 0 + 1.667}{3} = \frac{3}{3} = 1$$

Covariance matrix:

$$\mathbf{C} = \begin{bmatrix} 0.667 & 1 \\ 1 & 1.555 \end{bmatrix}$$

Step 3: Find Eigenvalues

Solve the characteristic equation $\det(\mathbf{C} - \lambda\mathbf{I}) = 0$:

$$\mathbf{C} - \lambda\mathbf{I} = \begin{bmatrix} 0.667 - \lambda & 1 \\ 1 & 1.555 - \lambda \end{bmatrix}$$

$$\begin{aligned} \det \begin{bmatrix} 0.667 - \lambda & 1 \\ 1 & 1.555 - \lambda \end{bmatrix} &= (0.667 - \lambda)(1.555 - \lambda) - (1)(1) \\ &= \lambda^2 - (0.667 + 1.555)\lambda + (0.667 \cdot 1.555 - 1) \\ &= \lambda^2 - 2.222\lambda + (1.037 - 1) = \lambda^2 - 2.222\lambda + 0.037 \end{aligned}$$

Solve the quadratic equation:

$$\lambda = \frac{2.222 \pm \sqrt{2.222^2 - 4 \cdot 1 \cdot 0.037}}{2} = \frac{2.222 \pm \sqrt{4.937 - 0.148}}{2} = \frac{2.222 \pm \sqrt{4.789}}{2}$$

$$\sqrt{4.789} \approx 2.188, \quad \lambda_1 \approx \frac{2.222 + 2.188}{2} = 2.205, \quad \lambda_2 \approx \frac{2.222 - 2.188}{2} = 0.017$$

Eigenvalues: $\lambda_1 \approx 2.205$, $\lambda_2 \approx 0.017$.

Step 4: Find Eigenvectors

For $\lambda_1 = 2.205$:

$$\mathbf{C} - 2.205\mathbf{I} = \begin{bmatrix} 0.667 - 2.205 & 1 \\ 1 & 1.555 - 2.205 \end{bmatrix} = \begin{bmatrix} -1.538 & 1 \\ 1 & -0.650 \end{bmatrix}$$

Solve $(\mathbf{C} - 2.205\mathbf{I})\mathbf{v} = 0$:

$$\begin{bmatrix} -1.538 & 1 \\ 1 & -0.650 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$-1.538v_1 + v_2 = 0 \implies v_2 = 1.538v_1$$

Choose $v_1 = 1$, then $v_2 = 1.538$. Normalize:

$$\|\mathbf{v}\|^2 = 1^2 + 1.538^2 = 1 + 2.365 = 3.365, \quad \|\mathbf{v}\| \approx 1.834$$

$$\mathbf{v}_1 \approx \begin{bmatrix} \frac{1}{1.834} \\ \frac{1.538}{1.834} \end{bmatrix} \approx \begin{bmatrix} 0.545 \\ 0.838 \end{bmatrix}$$

For $\lambda_2 = 0.017$, repeat similarly (omitted for brevity, as we focus on the first principal component).

Step 5: Project Data

Project the centered data onto the first principal component $\mathbf{v}_1 \approx [0.545, 0.838]$:

$$\text{Projection} = (x_1 - \bar{x}_1) \cdot 0.545 + (x_2 - \bar{x}_2) \cdot 0.838$$

For each point:

$$\begin{aligned}
(-1, -1.333): &\quad (-1) \cdot 0.545 + (-1.333) \cdot 0.838 \approx -0.545 - 1.117 = -1.662 \\
(0, -0.333): &\quad 0 \cdot 0.545 + (-0.333) \cdot 0.838 \approx -0.279 \\
(1, 1.667): &\quad 1 \cdot 0.545 + 1.667 \cdot 0.838 \approx 0.545 + 1.397 = 1.942
\end{aligned}$$

The 1D projections are approximately -1.662, -0.279, and 1.942.

7. Variance Explained

The proportion of variance explained by the first principal component:

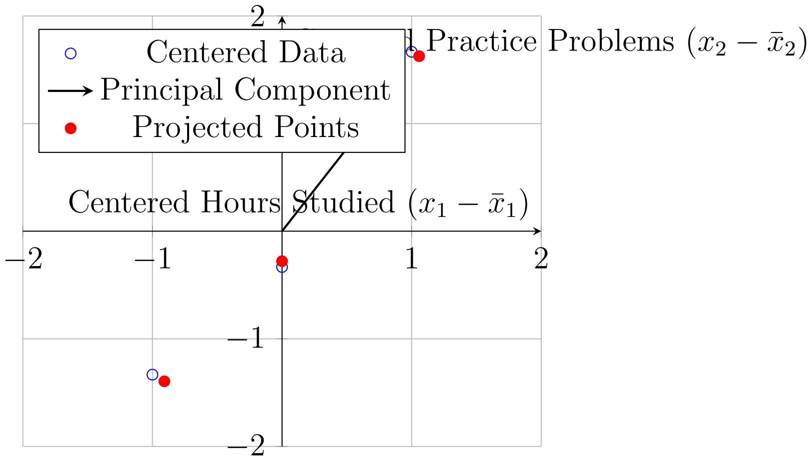
$$\frac{\lambda_1}{\lambda_1 + \lambda_2} = \frac{2.205}{2.205 + 0.017} \approx \frac{2.205}{2.222} \approx 0.992$$

This means 99.2% of the variance is captured by the first principal component, indicating effective dimensionality reduction.

8. Visualizations

This plot shows:

- Blue circles: centered data points.
- Black line: direction of the first principal component.
- Red dots: projected points on the principal component.



9. Final Results

Original (x_1, x_2)	Centered $(x_1 - \bar{x}_1, x_2 - \bar{x}_2)$	Projected Value
(1, 2)	(-1, -1.333)	-1.662
(2, 3)	(0, -0.333)	-0.279
(3, 5)	(1, 1.667)	1.942

- The first principal component is $\mathbf{v}_1 \approx [0.545, 0.838]$ with eigenvalue $\lambda_1 \approx 2.205$.
- The projected values represent the data in 1D, capturing 99.2% of the variance.
- This 1D representation can be used for further analysis, such as regression, with reduced complexity.

10. Summary

- PCA reduces dimensionality by finding eigenvectors (principal components) of the covariance matrix.
- Eigenvalues indicate the variance captured by each component.
- The first principal component maximizes variance, allowing effective data compression.
- The process involves centering data, computing the covariance matrix, finding eigenvalues/eigenvectors, and projecting data.