# PCA Explained Purely as Matrix Multiplication (Step-by-Step)

## 1 Dataset

We consider the following realistic $6 \times 2$ dataset:

$$X = \begin{bmatrix} 10 & 25 \\ 12 & 27 \\ 13 & 28 \\ 15 & 31 \\ 16 & 33 \\ 18 & 35 \end{bmatrix} \qquad (6 \times 2)$$

Column meanings:

$$X_1 = \text{Ad Spend (in \$1000s)}, \qquad X_2 = \text{New Customer Signups}.$$

No statistics interpretation is needed. PCA will be presented purely as matrix multiplications.

## 2 Step 1: Center the Data

Compute the column means:

$$\mu = \begin{bmatrix} 14 \\ 29.83 \end{bmatrix}.$$

We create a $6 \times 1$ vector of ones:

$$\mathbf{1} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix},$$

and form the replicated mean matrix:

$$\mathbf{1}\mu^\top = \begin{bmatrix} 14 & 29.83 \\ 14 & 29.83 \\ 14 & 29.83 \\ 14 & 29.83 \\ 14 & 29.83 \\ 14 & 29.83 \end{bmatrix}.$$

Now subtract to obtain the centered matrix:

$$X_c = X - \mathbf{1}\mu^\top.$$

$$X_c = \begin{bmatrix} -4 & -4.83 \\ -2 & -2.83 \\ -1 & -1.83 \\ 1 & 1.17 \\ 2 & 3.17 \\ 4 & 5.17 \end{bmatrix} \qquad (6 \times 2)$$

# 3 Step 2: Covariance Matrix via One Dot Product

PCA uses the covariance matrix

$$\Sigma = \frac{1}{n-1} X_c^\top X_c \quad \text{with } n = 6.$$

Shapes:

$$X_c : \ 6 \times 2, \qquad X_c^\top : \ 2 \times 6, \qquad X_c^\top X_c : \ 2 \times 2.$$

Compute the matrix product:

$$X_c^\top X_c = \begin{bmatrix} 42 & 55 \\ 55 & 72.83 \end{bmatrix}.$$

Divide by $n - 1 = 5$:

$$\Sigma = \begin{bmatrix} 8.4 & 11.0 \\ 11.0 & 14.57 \end{bmatrix}.$$

This $2 \times 2$ matrix will be diagonalized by PCA.

# 4 Step 3: Eigenvalues (FULL Detailed Derivation)

We solve the eigenvalue equation:

$$\Sigma v = \lambda v,$$

which is equivalent to:

$$\det(\Sigma - \lambda I) = 0.$$

Write:

$$\Sigma - \lambda I = \begin{bmatrix} 8.4 - \lambda & 11.0 \\ 11.0 & 14.57 - \lambda \end{bmatrix}.$$

The determinant is:

$$(8.4 - \lambda)(14.57 - \lambda) - (11.0)^2 = 0.$$

Expand:

$$(8.4 - \lambda)(14.57 - \lambda) = 8.4(14.57) - 8.4\lambda - 14.57\lambda + \lambda^2.$$

Compute:

$$8.4 \times 14.57 \approx 122.36.$$

Thus the characteristic polynomial becomes:

$$\lambda^2 - 22.97\lambda + 122.36 - 121 = 0,$$

$$\lambda^2 - 22.97\lambda + 1.36 = 0.$$

Using the quadratic formula:

$$\lambda = \frac{22.97 \pm \sqrt{22.97^2 - 4(1.36)}}{2}.$$

Compute the discriminant:

$$\Delta = 22.97^2 - 5.44 \approx 522.17, \qquad \sqrt{\Delta} \approx 22.85.$$

Therefore:

$$\lambda_1 = \frac{22.97 + 22.85}{2} \approx 22.91, \qquad \lambda_2 = \frac{22.97 - 22.85}{2} \approx 0.06.$$

# 5 Step 4: Eigenvectors (FULL Derivation)

### Eigenvector for $\lambda_1 \approx 22.91$

Solve

$$(\Sigma - \lambda_1 I)v_1 = 0.$$

$$\Sigma - \lambda_1 I \approx \begin{bmatrix} -14.51 & 11.0 \\ 11.0 & -8.34 \end{bmatrix}.$$

Let $v_1 = \begin{bmatrix} x \\ y \end{bmatrix}$.
From first row:

$$-14.51x + 11y = 0 \quad \Rightarrow \quad y = 1.319x.$$

Choose $x = 1$, so eigenvector direction:

$$v_1 \propto \begin{bmatrix} 1 \\ 1.319 \end{bmatrix}.$$

Normalize:

$$\|v_1\| = \sqrt{1^2 + (1.319)^2} \approx 1.655.$$

$$v_1 = \begin{bmatrix} 0.604 \\ 0.797 \end{bmatrix}.$$

### Eigenvector for $\lambda_2 \approx 0.06$

Solve

$$(\Sigma - \lambda_2 I)v_2 = 0.$$

$$\Sigma - \lambda_2 I \approx \begin{bmatrix} 8.34 & 11 \\ 11 & 14.51 \end{bmatrix}.$$

Equation:

$$8.34x + 11y = 0 \quad \Rightarrow \quad y = -0.758x.$$

Normalize:

$$v_2 = \begin{bmatrix} 0.797 \\ -0.604 \end{bmatrix}.$$

### Eigenvector Matrix

$$V = \begin{bmatrix} 0.604 & 0.797 \\ 0.797 & -0.604 \end{bmatrix} \qquad (2 \times 2)$$

Columns = principal components.

## 6 Step 5: PCA Projection via Dot Product

The PCA transformation is simply the matrix product:

$$Z = X_c V.$$

Shapes:

$$X_c: \ 6 \times 2, \qquad V: \ 2 \times 2, \qquad Z: \ 6 \times 2.$$

Compute:

$$Z = \begin{bmatrix} -4 & -4.83 \\ -2 & -2.83 \\ -1 & -1.83 \\ 1 & 1.17 \\ 2 & 3.17 \\ 4 & 5.17 \end{bmatrix} \begin{bmatrix} 0.604 & 0.797 \\ 0.797 & -0.604 \end{bmatrix} = \begin{bmatrix} 6.27 & 0.27 \\ 3.47 & -0.12 \\ 2.07 & -0.31 \\ -1.53 & -0.09 \\ -3.73 & 0.32 \\ -6.53 & -0.07 \end{bmatrix}.$$

Column 1 = PC1 scores, the major direction of variation. Column 2 = PC2 scores, very small because the original variables were almost perfectly correlated.

## 7 Summary

PCA is nothing more than the following chain of matrix multiplications:

$$(1)\ \text{Centering:} \quad X_c = X - \mathbf{1}\mu^\top,$$

$$(2)\ \text{Covariance:} \quad \Sigma = \frac{1}{5} X_c^\top X_c,$$

$$(3)\ \text{Eigen-decomposition:} \quad \Sigma = V \Lambda V^\top,$$

$$(4)\ \text{Projection:} \quad Z = X_c V.$$

No statistical interpretation required. Just clean linear algebra.