

Here's a detailed breakdown of the process with the numerical projection results:

1. Centered Data Matrix

The centered data matrix was computed by subtracting the mean from each attribute, aligning all values around zero for variance analysis.

Centered Matrix:

$$\begin{bmatrix} 0.4 & 0.8 \\ 0.4 & -0.2 \\ -0.6 & -1.2 \\ 0.4 & -0.2 \\ -0.6 & 0.8 \end{bmatrix}$$

2. Covariance Matrix

The covariance matrix captures relationships between attributes:

$$\text{Covariance Matrix: } \begin{bmatrix} \text{Var}(x) & \text{Cov}(x, y) \\ \text{Cov}(y, x) & \text{Var}(y) \end{bmatrix} = \begin{bmatrix} 0.4 & -0.2 \\ -0.2 & 0.72 \end{bmatrix}$$

3. Eigenvalues and Eigenvectors

Eigenvalues represent the variance captured by each principal component:

$$\text{Eigenvalues: } \lambda_1 = 0.8, \lambda_2 = 0.32$$

Eigenvectors define the directions of these components:

$$\text{Eigenvectors: } \begin{bmatrix} 0.5 & 0.5 \\ -0.7 & 0.7 \end{bmatrix}$$

4. Projection on Principal Components

We project the data onto the first principal component (most variance):

$$\text{Reduced Data Matrix: } \begin{bmatrix} 0.894 & -0.894 \\ -0.316 & -0.316 \\ -1.26 & 1.26 \\ 0.1 & -0.2 \end{bmatrix}$$

Takeaway:

- **Reduced Dimensions:** Instead of handling two full attributes, we've summarized the data into the top component, preserving the most variance for easier analysis.