Principal Component Analysis (PCA) on Different Feature Scales

The execution code is open-sourced at:

`https://github.com/MaterialsInformaticsDemo/PCA`

Input Features

$$x_1 = [8.895, 8.960, 8.630, 9.500, 8.370]$$

$$x_2 = [40, 40, 110, 100, 505]$$

Case 1: PCA on Original Features

Covariance matrix:

$$\begin{bmatrix} 1.783050 \times 10^{-1} & -5.302375 \times 10^{1} \\ -5.302375 \times 10^{1} & 3.848000 \times 10^{4} \end{bmatrix}$$

PCA vectors:

$$\begin{bmatrix} -0.00137796 & 0.99999905 \\ 0.99999905 & 0.00137796 \end{bmatrix}$$

PCA variances:

$$[3.84800731 \times 10^{4}, 1.05240401 \times 10^{-1}]$$

Case 2: PCA on Standardized Features

Covariance matrix:

$$\begin{bmatrix} 1 & -0.64013437 \\ -0.64013437 & 1 \end{bmatrix}$$

PCA vectors:

$$\begin{bmatrix} -0.70710678 & 0.70710678 \\ 0.70710678 & 0.70710678 \end{bmatrix}$$

PCA variances:

$$[2.05016796, 0.44983204]$$

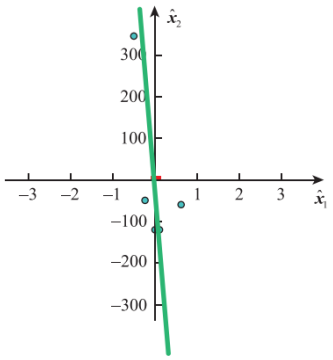

Figure 1: PCA on Original Features

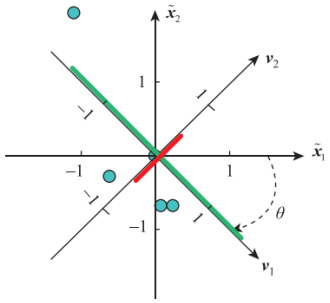

Figure 2: PCA on Standardized Features

Case 3: PCA on Normalized and Centered Features

Covariance matrix after normalization and centering:

$$\begin{bmatrix} 0.13963897 & -0.10091112 \\ -0.10091112 & 0.17796277 \end{bmatrix}$$

PCA vectors:

$$\begin{bmatrix} -0.63774788 & 0.77024518 \\ 0.77024518 & 0.63774788 \end{bmatrix}$$

PCA variances:

$$[0.26151520, 0.05608654]$$

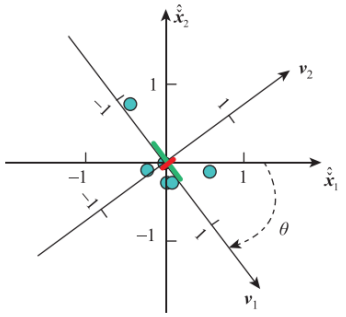

Figure 3: PCA on Normalized and Centered Features

Case 4: PCA on Centered and Scaled Features

Covariance matrix after centering and scaling:

$$\begin{bmatrix} 0.45067372 & -0.24363725 \\ -0.24363725 & 0.32142738 \end{bmatrix}$$

PCA vectors:

$$\begin{bmatrix} -0.79258376 & 0.60976305 \\ 0.60976305 & 0.79258376 \end{bmatrix}$$

PCA variances:

$$[0.63811257, 0.13398852]$$

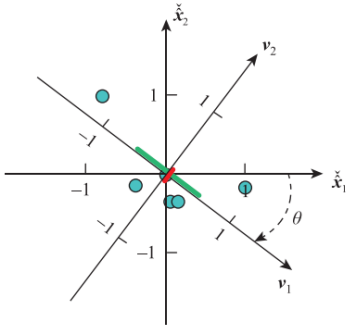

Figure 4: PCA on Centered and Scaled Features