Exercise 2 - PCA

$D = \{ (4,5)^T, (-2,3)^T, (1,1)^T \}$

1) Compute the mean to center the data

$$\bar{x} = \frac{1}{3}\left[\begin{pmatrix}4\\5\end{pmatrix} + \begin{pmatrix}-2\\3\end{pmatrix} + \begin{pmatrix}1\\1\end{pmatrix}\right] = \begin{pmatrix}1\\3\end{pmatrix}$$

$$x_1^* = x_1 - \bar{x} = \begin{pmatrix}4\\5\end{pmatrix} - \begin{pmatrix}1\\3\end{pmatrix} = \begin{pmatrix}3\\2\end{pmatrix}$$

$$x_2^* = x_2 - \bar{x} = \begin{pmatrix}-2\\3\end{pmatrix} - \begin{pmatrix}1\\3\end{pmatrix} = \begin{pmatrix}-3\\0\end{pmatrix}$$

$$x_3^* = x_3 - \bar{x} = \begin{pmatrix}1\\1\end{pmatrix} - \begin{pmatrix}1\\3\end{pmatrix} = \begin{pmatrix}0\\-2\end{pmatrix}$$

2) Form the matrix $X = (x_1 | \dots | x_n)^T$ and compute $A = X^T X$

$$X = \begin{bmatrix}3 & 2\\-3 & 0\\0 & -2\end{bmatrix}$$

$$A = X^T X = \begin{bmatrix}3 & -3 & 0\\2 & 0 & -2\end{bmatrix}\begin{bmatrix}3 & 2\\-3 & 0\\0 & -2\end{bmatrix} = \begin{bmatrix}18 & 6\\6 & 8\end{bmatrix}$$

3) Solve for the eigenvectors of A

$\det(A - \lambda I) = 0$

$$\begin{vmatrix}18-\lambda & 6\\6 & 8-\lambda\end{vmatrix} = 0$$

$$(18-\lambda)(8-\lambda)-36=0$$

$$144-26\lambda+\lambda^2-36=0$$

$$\lambda^2-26\lambda+108=0$$

Eigenvalues: $\lambda_1 = 13+\sqrt{61}$ $\quad \lambda_2 = 13-\sqrt{61}$

①

$$(A-\lambda_1 I)\vec{v_1} = \vec{0}$$

$$\begin{bmatrix} 18-13-\sqrt{61} & 6 \\ 6 & 8-13-\sqrt{61} \end{bmatrix} \begin{pmatrix} v_1^{(1)} \\ v_1^{(2)} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{cases} (5-\sqrt{61})\, v_1^{(1)} + 6\, v_1^{(2)} = 0 \\ 6\, v_1^{(1)} + (-5-\sqrt{61})\, v_1^{(2)} = 0 \end{cases} \iff \begin{cases} v_1^{(1)} = 5+\sqrt{61} \\ v_1^{(2)} = 6 \end{cases}$$

$$\vec{v_1} = \begin{pmatrix} 5+\sqrt{61} \\ 6 \end{pmatrix} \quad \text{Normalized} \quad \hat{v}_1 = \frac{\vec{v_1}}{\|\vec{v_1}\|_2} \approx \begin{pmatrix} 0.9056 \\ 0.4242 \end{pmatrix}$$

②

$$(A-\lambda_2 I)\vec{v_2} = \vec{0}$$

$$\begin{bmatrix} 18-13+\sqrt{61} & 6 \\ 6 & 8-13+\sqrt{61} \end{bmatrix} \begin{pmatrix} v_2^{(1)} \\ v_2^{(2)} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{cases} (5+\sqrt{61})\, v_2^{(1)} + 6\, v_2^{(2)} = 0 \\ 6\, v_2^{(1)} + (-5+\sqrt{61})\, v_2^{(2)} = 0 \end{cases} \iff \begin{cases} v_2^{(1)} = 5-\sqrt{61} \\ v_2^{(2)} = 6 \end{cases}$$

$$\vec{v_2} = \begin{pmatrix} 5-\sqrt{61} \\ 6 \end{pmatrix} \quad \text{Normalized} \quad \hat{v}_2 = \frac{\vec{v_2}}{\|\vec{v_2}\|_2} \approx \begin{pmatrix} -0.4242 \\ 0.9056 \end{pmatrix}$$

4) The principal components are the normalized eigenvectors ordered by decreasing size of eigenvalues.

$$V = (\hat{v}_1 \mid \hat{v}_2) = \begin{pmatrix} 0.9056 & -0.4242 \\ 0.4242 & 0.9056 \end{pmatrix}$$