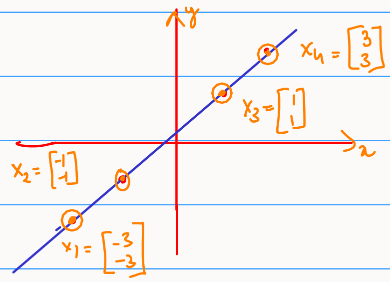
Quiz - 2 Solutions

(1)



The original data resides in $\mathbb{R}^2$ (two dim. space)

we want to reduce the dimensionality by 1.

$\Rightarrow \quad K = 1$

data matrix $X = \begin{bmatrix} -3 & -1 & 1 & 3 \\ -3 & -1 & 1 & 3 \end{bmatrix}$

$$\Rightarrow XX^T = \begin{bmatrix} -3 & -1 & 1 & 3 \\ -3 & -1 & 1 & 3 \end{bmatrix} \begin{bmatrix} -3 & -3 \\ -1 & -1 \\ 1 & 1 \\ 3 & 3 \end{bmatrix} = \begin{bmatrix} 20 & 20 \\ 20 & 20 \end{bmatrix}$$

Now, reduced dim. data is obtained as $\hat{x}_i = U^T x_i$

Here, $U \in \mathbb{R}^{2\times 1}$ and equal to the eigenvector of XX^T corresponding to the largest eigenvalue.

Now, find EVD of XX^T. Let $C = XX^T$

$$\Rightarrow \det(C - \lambda I) = 0 \Rightarrow \det\left(\begin{bmatrix} 20-\lambda & 20 \\ 20 & 20-\lambda \end{bmatrix}\right) = 0 \Rightarrow (20-\lambda)^2 - 20^2 = 0$$

$$\Rightarrow 20 - \lambda = \pm 20 \Rightarrow \lambda_1 = 0, \ \lambda_2 = 40$$

$$\Rightarrow \begin{bmatrix} 20-40 & 20 \\ 20 & 20-40 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = 0 \Rightarrow \begin{bmatrix} -20 & 20 \\ 20 & -20 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = 0 \Rightarrow x_1 = y_1$$

$\Rightarrow v_1 = \begin{bmatrix} x_1 \\ x_1 \end{bmatrix}$. Now, let us make sure that v_1 has unit norm

$$\Rightarrow v_1 \leftarrow \frac{v_1}{\|v_1\|_2} = \frac{1}{x_1\sqrt{2}} \begin{bmatrix} x_1 \\ x_1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\Rightarrow U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow W = U^T = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \end{bmatrix}$$

Now, the reduced dim. data can be found as follows:

$$\hat{x}_1 = W x_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} -3 \\ -3 \end{bmatrix} = \frac{-6}{\sqrt{2}} = -3\sqrt{2}$$

$$\hat{x}_2 = W x_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \end{bmatrix} = \frac{-2}{\sqrt{2}} = -\sqrt{2}$$

$$\hat{x}_3 = W x_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{2}{\sqrt{2}} = -\sqrt{2}$$

$$\hat{x}_4 = W x_4 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 3 \\ 3 \end{bmatrix} = \frac{6}{\sqrt{2}} = 3\sqrt{2}$$

$$\Rightarrow \hat{X} = \begin{bmatrix} -3\sqrt{2}, & -\sqrt{2}, & +\sqrt{2}, & 3\sqrt{2} \end{bmatrix}$$

Reconstruction:

$$\bar{x}_1 = U\hat{x}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}(-3\sqrt{2}) = \begin{bmatrix}-3\\-3\end{bmatrix}$$

$$\bar{x}_2 = U\hat{x}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}(-\sqrt{2}) = \begin{bmatrix}-1\\-1\end{bmatrix}$$

$$\bar{x}_3 = U\hat{x}_3 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}(\sqrt{2}) = \begin{bmatrix}1\\1\end{bmatrix}$$

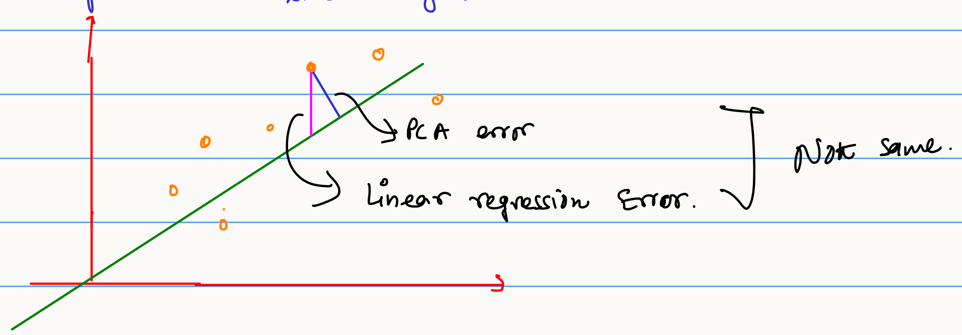
$$\bar{x}_4 = U\hat{x}_4 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}(3\sqrt{2}) = \begin{bmatrix}3\\3\end{bmatrix}$$

$\Rightarrow$ reconstruction Error $= \sum_{i=1}^{4} \|\bar{x}_i - x_i\|_2^2 = 0$

$\Rightarrow$ We can Recover the original data with zero error.

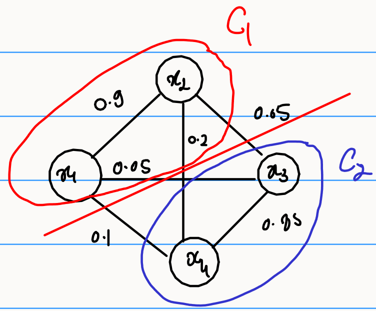
$\Rightarrow$ This is a special case as all the original data points lie on a line (1- dim. subspace).

PCA vs. Linear regression for line fitting problem



Quiz - 1 Solutions

Ans 3



$$\text{RatioCut}(C_1, C_2) = \sum_{i=1}^{2} \frac{1}{|C_i|} \sum_{r \in C_i} \sum_{s \notin C_i} w_{r,s}$$

$$|C_1| = 2, \quad |C_2| = 2$$

$$\text{RatioCut}(C_1, C_2) = \frac{1}{2} \sum_{r \in C_1} \sum_{s \notin C_1} w_{r,s} + \frac{1}{2} \sum_{r \in C_2} \sum_{s \notin C_2} w_{r,s}$$

$$\text{RatioCut}(C_1, C_2) = \frac{1}{2}\left(w_{1,3} + w_{1,4} + w_{2,3} + w_{2,4}\right) + \frac{1}{2}\left(w_{3,1} + w_{3,2} + w_{4,1} + w_{4,2}\right)$$

$$= \frac{1}{2}\left(0.05 + 0.1 + 0.2 + 0.05\right) + \frac{1}{2}\left(0.05 + 0.05 + 0.2 + 0.1\right)$$

$$= \frac{1}{2}(0.4) + \frac{1}{2}(0.4) = 0.4$$

①

$$\|x_1\|_2 = \sqrt{1^2 + 1^2} = \sqrt{2}, \quad \|x_2\|_2 = \sqrt{1^2 + 1^2} = \sqrt{2}, \quad \|x_3\|_2 = \sqrt{4+1} = \sqrt{5}$$

$$\|x_1 - x_2\|_2 = \|[1-1, 0-1, 1-0]^T\|_2 = \sqrt{1+1} = \sqrt{2}$$ ✓

$$\|x_2 - x_3\|_2 = \|[1-0, 1-2, 0-1]\|_2 = \sqrt{1+1+1} = \sqrt{3}$$

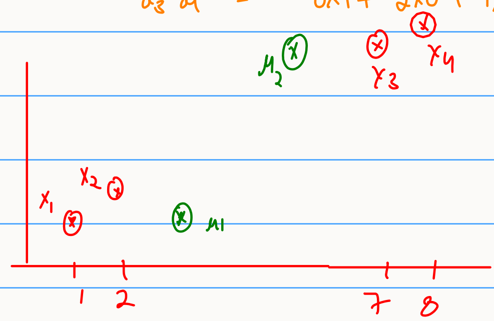
$$\|x_3 - x_1\|_2 = \|[0-1, 2-0, 1-1]^T\|_2 = \sqrt{4+1} = \sqrt{5}$$

$$x_1^T x_2 = 1 \times 1 + 0 \times 1 + 1 \times 0 = 1$$

$$x_2^T x_3 = 1 \times 0 + 1 \times 2 + 1 \times 0 = 2$$

$$x_3^T x_1 = 0 \times 1 + 2 \times 0 + 1 \times 1 = 2$$

②



given initial centers, we have to find clusters

$$d(x_1, \mu_1) < d(x_1, \mu_2) \Rightarrow x_1 \in C_1$$

$$d(x_2, \mu_1) < d(x_2, \mu_2) \Rightarrow x_2 \in C_1$$

$$d(x_3, \mu_2) < d(x_3, \mu_1) \Rightarrow x_3 \in C_2$$

$$d(x_4, \mu_2) < d(x_4, \mu_1) \Rightarrow x_4 \in C_2$$

$$\Rightarrow C_1 = \{x_1, x_2\}, \quad C_2 = \{x_3, x_4\}$$

Now, update the cluster centers

$$\Rightarrow \mu_1 = \frac{1}{2}(x_1 + x_2) = \begin{bmatrix} 1.5 \\ 1.5 \end{bmatrix}$$

$$\mu_2 = \frac{1}{2}(x_3 + x_4) = \begin{bmatrix} 7.5 \\ 7.5 \end{bmatrix}$$

Now, again find clustering using these updated cluster centers.

You will see that again $C_1 = \{x_1, x_2\}$, and $C_2 = \{x_3, x_4\}$

No change in clusters $\Rightarrow$ No change in cluster centers

$\Rightarrow$ Converged.