Module 6.5 : PCA : Interpretation 2

Given n orthogonal linearly independent vectors $P = p_1, p_2, \cdots, p_n$ we can represent x_i exactly as a linear combination of these vectors.

Given n orthogonal linearly independent vectors $P = p_1, p_2, \cdots, p_n$ we can represent x_i exactly as a linear combination of these vectors.

$$x_i = \sum_{j=1}^{n} \alpha_{ij} p_j \quad \text{[we know how to estimate } \alpha'_{ij}s \text{ but we will come back to that later]}$$

Given n orthogonal linearly independent vectors $P = p_1, p_2, \cdots, p_n$ we can represent x_i exactly as a linear combination of these vectors.

$$x_i = \sum_{j=1}^{n} \alpha_{ij} p_j \quad [\text{we know how to estimate } \alpha'_{ij}s \text{ but we will come back to that later}]$$

But we are interested only in the top-k dimensions (we want to get rid of noisy & redundant dimensions)

$$\hat{x}_i = \sum_{j=1}^{k} \alpha_{ik} p_k$$

Given n orthogonal linearly independent vectors $P = p_1, p_2, \cdots, p_n$ we can represent x_i exactly as a linear combination of these vectors.

$$x_i = \sum_{j=1}^{n} \alpha_{ij} p_j \quad \text{[we know how to estimate } \alpha'_{ij}s \text{ but we will come back to that later]}$$

But we are interested only in the top-k dimensions (we want to get rid of noisy & redundant dimensions)

$$\hat{x}_i = \sum_{j=1}^{k} \alpha_{ik} p_k$$

We want to select $p'_i s$ such that we minimise the reconstructed error

$$e = \sum_{i=1}^{m} (x_i - \hat{x}_i)^T (x_i - \hat{x}_i)$$

$$e = \sum_{i=1}^{m} (x_i - \hat{x}_i)^T (x_i - \hat{x}_i)$$

$$
\begin{aligned}
e &= \sum_{i=1}^{m} (x_i - \hat{x}_i)^T (x_i - \hat{x}_i) \\
&= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \alpha_{ij} p_j - \sum_{j=1}^{k} \alpha_{ij} p_j \right)^2
\end{aligned}
$$

$$
\begin{aligned}
e &= \sum_{i=1}^{m} (x_i - \hat{x}_i)^T (x_i - \hat{x}_i) \\
&= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \alpha_{ij} p_j - \sum_{j=1}^{k} \alpha_{ij} p_j \right)^2 \\
&= \sum_{i=1}^{m} \left(\sum_{j=k+1}^{n} \alpha_{ij} p_j \right)^2 = \sum_{i=1}^{m} \left(\sum_{j=k+1}^{n} \alpha_{ij} p_j \right)^T \left(\sum_{j=k+1}^{n} \alpha_{ij} p_j \right)
\end{aligned}
$$

$$
\begin{aligned}
e &= \sum_{i=1}^{m}(x_i - \hat{x}_i)^T(x_i - \hat{x}_i) \\
&= \sum_{i=1}^{m}\left(\sum_{j=1}^{n}\alpha_{ij}p_j - \sum_{j=1}^{k}\alpha_{ij}p_j\right)^2 \\
&= \sum_{i=1}^{m}\left(\sum_{j=k+1}^{n}\alpha_{ij}p_j\right)^2 = \sum_{i=1}^{m}\left(\sum_{j=k+1}^{n}\alpha_{ij}p_j\right)^T\left(\sum_{j=k+1}^{n}\alpha_{ij}p_j\right) \\
&= \sum_{i=1}^{m}\left(\alpha_{i,k+1}p_{k+1} + \alpha_{i,k+2}p_{k+2} + \ldots + \alpha_{i,n}p_n\right)^T\left(\alpha_{i,k+1}p_{k+1} + \alpha_{i,k+2}p_{k+2} + \ldots + \alpha_{i,n}p_n\right)
\end{aligned}
$$

$$
\begin{aligned}
e &= \sum_{i=1}^{m} (x_i - \hat{x}_i)^T (x_i - \hat{x}_i) \\
&= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \alpha_{ij} p_j - \sum_{j=1}^{k} \alpha_{ij} p_j \right)^2 \\
&= \sum_{i=1}^{m} \left(\sum_{j=k+1}^{n} \alpha_{ij} p_j \right)^2 = \sum_{i=1}^{m} \left(\sum_{j=k+1}^{n} \alpha_{ij} p_j \right)^T \left(\sum_{j=k+1}^{n} \alpha_{ij} p_j \right) \\
&= \sum_{i=1}^{m} \left(\alpha_{i,k+1} p_{k+1} + \alpha_{i,k+2} p_{k+2} + \ldots + \alpha_{i,n} p_n \right)^T \left(\alpha_{i,k+1} p_{k+1} + \alpha_{i,k+2} p_{k+2} + \ldots + \alpha_{i,n} p_n \right) \\
&= \sum_{i=1}^{m} \sum_{j=k+1}^{n} \alpha_{ij} p_j^T p_j \alpha_{ij} + \sum_{i=1}^{m} \sum_{j=k+1}^{n} \sum_{L=k+1, L \neq k}^{n} \alpha_{ij} p_j^T p_L \alpha_{iL}
\end{aligned}
$$

$$
\begin{aligned}
e &= \sum_{i=1}^{m} (x_i - \hat{x}_i)^T (x_i - \hat{x}_i) \\
&= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \alpha_{ij} p_j - \sum_{j=1}^{k} \alpha_{ij} p_j \right)^2 \\
&= \sum_{i=1}^{m} \left(\sum_{j=k+1}^{n} \alpha_{ij} p_j \right)^2 = \sum_{i=1}^{m} \left(\sum_{j=k+1}^{n} \alpha_{ij} p_j \right)^T \left(\sum_{j=k+1}^{n} \alpha_{ij} p_j \right) \\
&= \sum_{i=1}^{m} (\alpha_{i,k+1} p_{k+1} + \alpha_{i,k+2} p_{k+2} + \ldots + \alpha_{i,n} p_n)^T (\alpha_{i,k+1} p_{k+1} + \alpha_{i,k+2} p_{k+2} + \ldots + \alpha_{i,n} p_n) \\
&= \sum_{i=1}^{m} \sum_{j=k+1}^{n} \alpha_{ij} p_j^T p_j \alpha_{ij} + \sum_{i=1}^{m} \sum_{j=k+1}^{n} \sum_{L=k+1, L \neq k}^{n} \alpha_{ij} p_j^T p_L \alpha_{iL} \\
&= \sum_{i=1}^{m} \sum_{j=k+1}^{n} \alpha_{ij}^2 \qquad\qquad (\because p_j^T p_j = 1, p_i^T p_j = 0 \quad \forall i \neq j)
\end{aligned}
$$

$$
\begin{aligned}
e &= \sum_{i=1}^{m} (x_i - \hat{x}_i)^T (x_i - \hat{x}_i) \\
&= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \alpha_{ij} p_j - \sum_{j=1}^{k} \alpha_{ij} p_j \right)^2 \\
&= \sum_{i=1}^{m} \left(\sum_{j=k+1}^{n} \alpha_{ij} p_j \right)^2 = \sum_{i=1}^{m} \left(\sum_{j=k+1}^{n} \alpha_{ij} p_j \right)^T \left(\sum_{j=k+1}^{n} \alpha_{ij} p_j \right) \\
&= \sum_{i=1}^{m} \left(\alpha_{i,k+1} p_{k+1} + \alpha_{i,k+2} p_{k+2} + \ldots + \alpha_{i,n} p_n \right)^T \left(\alpha_{i,k+1} p_{k+1} + \alpha_{i,k+2} p_{k+2} + \ldots + \alpha_{i,n} p_n \right) \\
&= \sum_{i=1}^{m} \sum_{j=k+1}^{n} \alpha_{ij} p_j^T p_j \alpha_{ij} + \sum_{i=1}^{m} \sum_{j=k+1}^{n} \sum_{L=k+1, L \neq k}^{n} \alpha_{ij} p_j^T p_L \alpha_{iL} \\
&= \sum_{i=1}^{m} \sum_{j=k+1}^{n} \alpha_{ij}^2 \qquad (\because p_j^T p_j = 1, p_i^T p_j = 0 \quad \forall i \neq j) \\
&= \sum_{i=1}^{m} \sum_{j=k+1}^{n} \left(x_i^T p_j \right)^2
\end{aligned}
$$

$$= \sum_{i=1}^{m} \sum_{j=k+1}^{n} \left(p_j^T x_i\right) \left(x_i^T p_j\right)$$

$$e = \sum_{i=1}^{m} (x_i - \hat{x}_i)^T (x_i - \hat{x}_i)$$

$$= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \alpha_{ij} p_j - \sum_{j=1}^{k} \alpha_{ij} p_j \right)^2$$

$$= \sum_{i=1}^{m} \left(\sum_{j=k+1}^{n} \alpha_{ij} p_j \right)^2 = \sum_{i=1}^{m} \left(\sum_{j=k+1}^{n} \alpha_{ij} p_j \right)^T \left(\sum_{j=k+1}^{n} \alpha_{ij} p_j \right)$$

$$= \sum_{i=1}^{m} \left(\alpha_{i,k+1} p_{k+1} + \alpha_{i,k+2} p_{k+2} + \ldots + \alpha_{i,n} p_n\right)^T \left(\alpha_{i,k+1} p_{k+1} + \alpha_{i,k+2} p_{k+2} + \ldots + \alpha_{i,n} p_n\right)$$

$$= \sum_{i=1}^{m} \sum_{j=k+1}^{n} \alpha_{ij} p_j^T p_j \alpha_{ij} + \sum_{i=1}^{m} \sum_{j=k+1}^{n} \sum_{L=k+1, L \neq k}^{n} \alpha_{ij} p_j^T p_L \alpha_{iL}$$

$$= \sum_{i=1}^{m} \sum_{j=k+1}^{n} \alpha_{ij}^2 \qquad (\because p_j^T p_j = 1, p_i^T p_j = 0 \quad \forall i \neq j)$$

$$= \sum_{i=1}^{m} \sum_{j=k+1}^{n} \left(x_i^T p_j\right)^2$$

$$
\begin{aligned}
e &= \sum_{i=1}^{m} (x_i - \hat{x}_i)^T (x_i - \hat{x}_i) \\
&= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \alpha_{ij} p_j - \sum_{j=1}^{k} \alpha_{ij} p_j \right)^2 \\
&= \sum_{i=1}^{m} \left(\sum_{j=k+1}^{n} \alpha_{ij} p_j \right)^2 = \sum_{i=1}^{m} \left(\sum_{j=k+1}^{n} \alpha_{ij} p_j \right)^T \left(\sum_{j=k+1}^{n} \alpha_{ij} p_j \right) \\
&= \sum_{i=1}^{m} \left(\alpha_{i,k+1} p_{k+1} + \alpha_{i,k+2} p_{k+2} + \ldots + \alpha_{i,n} p_n \right)^T \left(\alpha_{i,k+1} p_{k+1} + \alpha_{i,k+2} p_{k+2} + \ldots + \alpha_{i,n} p_n \right) \\
&= \sum_{i=1}^{m} \sum_{j=k+1}^{n} \alpha_{ij} p_j^T p_j \alpha_{ij} + \sum_{i=1}^{m} \sum_{j=k+1}^{n} \sum_{L=k+1, L \neq k}^{n} \alpha_{ij} p_j^T p_L \alpha_{iL} \\
&= \sum_{i=1}^{m} \sum_{j=k+1}^{n} \alpha_{ij}^2 \qquad (\because p_j^T p_j = 1, p_i^T p_j = 0 \quad \forall i \neq j) \\
&= \sum_{i=1}^{m} \sum_{j=k+1}^{n} \left(x_i^T p_j \right)^2
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{i=1}^{m} \sum_{j=k+1}^{n} \left(p_j^T x_i \right) \left(x_i^T p_j \right) \\
&= \sum_{j=k+1}^{n} p_j^T \left(\sum_{i=1}^{m} x_i x_i^T \right) p_j
\end{aligned}
$$

$$e = \sum_{i=1}^{m} (x_i - \hat{x}_i)^T (x_i - \hat{x}_i)$$

$$= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \alpha_{ij} p_j - \sum_{j=1}^{k} \alpha_{ij} p_j \right)^2$$

$$= \sum_{i=1}^{m} \left(\sum_{j=k+1}^{n} \alpha_{ij} p_j \right)^2 = \sum_{i=1}^{m} \left(\sum_{j=k+1}^{n} \alpha_{ij} p_j \right)^T \left(\sum_{j=k+1}^{n} \alpha_{ij} p_j \right)$$

$$= \sum_{i=1}^{m} (\alpha_{i,k+1} p_{k+1} + \alpha_{i,k+2} p_{k+2} + \ldots + \alpha_{i,n} p_n)^T (\alpha_{i,k+1} p_{k+1} + \alpha_{i,k+2} p_{k+2} + \ldots + \alpha_{i,n} p_n)$$

$$= \sum_{i=1}^{m} \sum_{j=k+1}^{n} \alpha_{ij} p_j^T p_j \alpha_{ij} + \sum_{i=1}^{m} \sum_{j=k+1}^{n} \sum_{L=k+1, L \neq k}^{n} \alpha_{ij} p_j^T p_L \alpha_{iL}$$

$$= \sum_{i=1}^{m} \sum_{j=k+1}^{n} \alpha_{ij}^2 \qquad (\because p_j^T p_j = 1, p_i^T p_j = 0 \quad \forall i \neq j)$$

$$= \sum_{i=1}^{m} \sum_{j=k+1}^{n} \left(x_i^T p_j \right)^2$$

$$= \sum_{i=1}^{m} \sum_{j=k+1}^{n} \left(p_j^T x_i \right) \left(x_i^T p_j \right)$$

$$= \sum_{j=k+1}^{n} p_j^T \left(\sum_{i=1}^{m} x_i x_i^T \right) p_j$$

$$= \sum_{j=k+1}^{n} p_j^T m C p_j \qquad \left[\because \frac{1}{m} \sum_{i=1}^{m} x_i x_i^T = \frac{X^T X}{m} = C \right]$$

We want to minimize e

$$\min_{p_{k+1},p_{k+2},\cdots,p_n} \sum_{j=k+1}^{n} p_j^T mCp_j \qquad s.t. \quad p_j^T p_j = 1 \quad \forall j = k+1, k+2, \cdots, n$$

We want to minimize e

$$\min_{p_{k+1}, p_{k+2}, \cdots, p_n} \sum_{j=k+1}^{n} p_j^T mC p_j \qquad s.t. \quad p_j^T p_j = 1 \quad \forall j = k+1, k+2, \cdots, n$$

The solution to the above problem is given by the eigen vectors corresponding to the smallest eigen values of C (**Proof : refer Slide 26**).

We want to minimize e

$$\min_{p_{k+1}, p_{k+2}, \cdots, p_n} \sum_{j=k+1}^{n} p_j^T m C p_j \qquad s.t. \quad p_j^T p_j = 1 \quad \forall j = k+1, k+2, \cdots, n$$

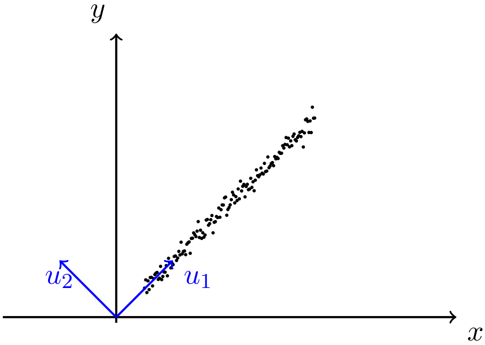
The solution to the above problem is given by the eigen vectors corresponding to the smallest eigen values of C (**Proof : refer Slide 26**).

Thus we select $P = p_1, p_2, \cdots, p_n$ as eigen vectors of C and retain only top-k eigen vectors to express the data [or discard the eigen vectors $k+1, \cdots, n$]

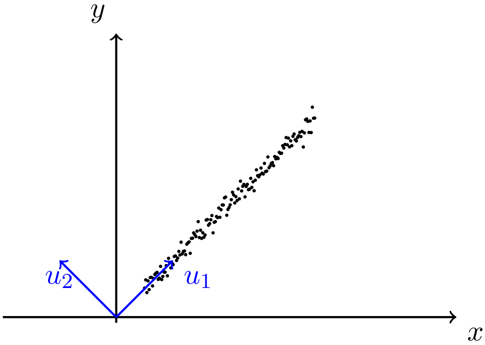
Key Idea

Minimize the error in reconstructing x_i after projecting the data on to a new basis.

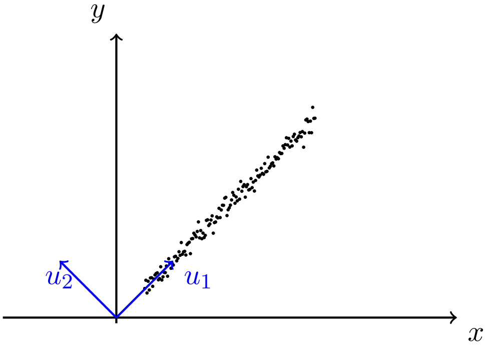
Let's look at the ***'Reconstruction Error'*** *in the context of our toy example*



- $u_1 = [1, 1]$ and $u_2 = [-1, 1]$ are the new basis vectors

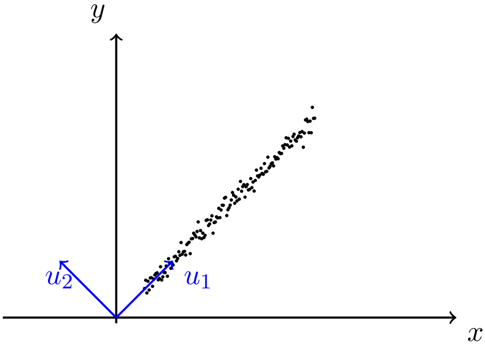


- $u_1 = [1, 1]$ and $u_2 = [-1, 1]$ are the new basis vectors
- Let us convert them to unit vectors $u_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ & $u_2 = \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$



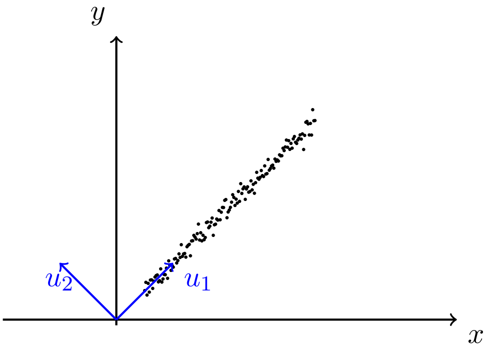
- $u_1 = [1, 1]$ and $u_2 = [-1, 1]$ are the new basis vectors
- Let us convert them to unit vectors $u_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ & $u_2 = \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$

- Consider the point $x = [3.3, 3]$ in the original data



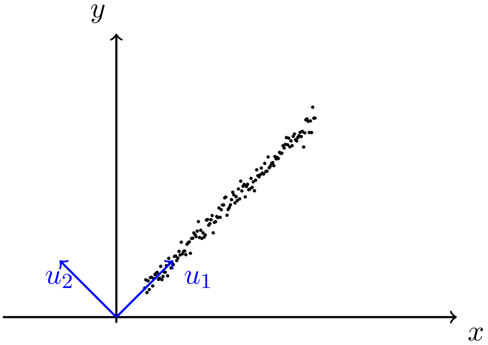
- $u_1 = [1, 1]$ and $u_2 = [-1, 1]$ are the new basis vectors
- Let us convert them to unit vectors $u_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ & $u_2 = \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$

- Consider the point $x = [3.3, 3]$ in the original data
- $\alpha_1 = x^T u_1 = 6.3/\sqrt{2}$
  $\alpha_2 = x^T u_2 = 0.3/\sqrt{2}$



- $u_1 = [1, 1]$ and $u_2 = [-1, 1]$ are the new basis vectors
- Let us convert them to unit vectors $u_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ & $u_2 = \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$
- Consider the point $x = [3.3, 3]$ in the original data
- $\alpha_1 = x^T u_1 = 6.3/\sqrt{2}$
  $\alpha_2 = x^T u_2 = 0.3/\sqrt{2}$
- the perfect reconstruction of x is given by (using $n = 2$ dimensions)

$$x = \alpha_1 u_1 + \alpha_2 u_2 = \begin{bmatrix} 3.3 & 3 \end{bmatrix}$$



- $u_1 = [1, 1]$ and $u_2 = [-1, 1]$ are the new basis vectors
- Let us convert them to unit vectors $u_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$ & $u_2 = \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$

- Consider the point $x = [3.3, 3]$ in the original data
- $\alpha_1 = x^T u_1 = 6.3/\sqrt{2}$
  $\alpha_2 = x^T u_2 = 0.3/\sqrt{2}$
- the perfect reconstruction of x is given by (using $n = 2$ dimensions)

$$x = \alpha_1 u_1 + \alpha_2 u_2 = \begin{bmatrix} 3.3 & 3 \end{bmatrix}$$

- But we are going to reconstruct it using fewer (only $k = 1 < n$ dimensions, ignoring the low variance u_2 dimension)

$$\hat{x} = \alpha_1 u_1 = \begin{bmatrix} 3.15 & 3.15 \end{bmatrix}$$

(reconstruction with minimum error)

Recap

- The eigen vectors of a matrix with distinct eigenvalues are linearly independent

Recap

- The eigen vectors of a matrix with distinct eigenvalues are linearly independent
- The eigen vectors of a square symmetric matrix are orthogonal

Recap

- The eigen vectors of a matrix with distinct eigenvalues are linearly independent
- The eigen vectors of a square symmetric matrix are orthogonal
- PCA exploits this fact by representing the data using a new basis comprising only the top-k eigen vectors

Recap

- The eigen vectors of a matrix with distinct eigenvalues are linearly independent
- The eigen vectors of a square symmetric matrix are orthogonal
- PCA exploits this fact by representing the data using a new basis comprising only the top-k eigen vectors
- The $n - k$ dimensions which contribute very little to the reconstruction error are discarded

Recap

- The eigen vectors of a matrix with distinct eigenvalues are linearly independent
- The eigen vectors of a square symmetric matrix are orthogonal
- PCA exploits this fact by representing the data using a new basis comprising only the top-k eigen vectors
- The $n - k$ dimensions which contribute very little to the reconstruction error are discarded
- **These are also the directions along which the variance is minimum**