Page No.
Date

Q1) a) Covariance matrix $(C) = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})(x_i - \bar{x})^T$

$$C^T = \frac{1}{N-1} \sum_{i=1}^{N} \left[(x_i-\bar{x})(x_i-\bar{x})^T\right]^T$$

$$C^T = \frac{1}{N-1} \sum_{i=1}^{N} (x_i-\bar{x})(x_i-\bar{x})^T = C$$

As $(AB)^T = B^T A^T$, $((x_i-\bar{x})^T)^T = x_i-\bar{x}$

$\Rightarrow C^T = C$ which means C is symmetric

Let $\overline{x_i} = x_i - \bar{x}$

$$\Rightarrow C = \frac{1}{N-1} \sum_{i=1}^{N} \overline{x_i}\,(\overline{x_i})^T$$

$\Rightarrow C = \frac{1}{N-1} X X^T$ where $X = [\bar{x}_1 | \bar{x}_2 | \ldots | \bar{x}_N]$

$\Rightarrow$ Let V is a column matrix $d \times 1$ (where x_i is $d \times 1$ matrix).

C is positive semidefinite if $V^T C V \geq 0 \quad \forall V$

now $V^T C V = \frac{1}{N-1} V^T X X^T V = \frac{1}{N-1} (X^T V)^T (X^T V)$

$\rightarrow$ Let $Z = X^T V$ $(\dim(Z) = d \times 1)$

$\Rightarrow V^T C V = \frac{1}{N-1} Z^T Z = \frac{1}{N-1} |Z|^2$ (Sum of squares of elements in Z)

$\Rightarrow V^T C V \geq 0 \quad \forall V$

$\Rightarrow C$ is symmetric, positive semidefinite.

b) Let C be our symmetric matrix (assume $d \times d$ matrix)

→ Let $(V_1, V_2, \dots V_d)$ be eigen vectors, $(\lambda_1, \lambda_2, \dots \lambda_d)$ be corresponding eigen values.

We know $C V_i = \lambda_i V_i$

→ Consider $V_i^T C V_j$

$$V_i^T C V_j = V_i^T (C V_j) = V_i^T \lambda_j V_j = \lambda_j V_i^T V_j$$

$$V_i^T C V_j = V_i^T C^T V_j \quad (C \text{ is symmetric})$$

$$V_i^T C^T V_j = (C V_i)^T V_j = (\lambda_i V_i)^T V_j = \lambda_i V_i^T V_j$$

$$\Rightarrow \lambda_i V_i^T V_j = \lambda_j V_i^T V_j$$

→ For eigen vectors corresponding to different eigen values (i.e $\lambda_i \neq \lambda_j$)

$\lambda_i \neq \lambda_j \Rightarrow V_i^T V_j = 0$ which shows V_i $\perp^{lar}$ to V_j

→ If $\lambda_i = \lambda_j$ then solutions form a vector space ($CV = \lambda V$) then we can consider any orthonormal basis of this space and consider them to be eigen vectors.

→ Now we can normalise these vectors $V_1, V_2, \dots V_d$ making them unit vectors (If V satisfy $CV = \lambda V$ then γV also satisfy)

→ Now since all these eigen vectors are pairwise $\perp^{lar}$, unit magnitude we can say eigen vectors of symmetric matrix are orthonormal

c) We know that $x_i = \bar{x} + \sum_{l=1}^{d} \alpha_{il} V_l$ (From definition of α_{il})

Given $\tilde{x}_i = \bar{x} + \sum_{l=1}^{K} \alpha_{il} V_l$ (Assume V_l sorted in eigen value order)

$$(N-1)C = \sum_{i=1}^{N} (x_i - \bar{x})(x_i - \bar{x})^T = \sum_{i=1}^{N} \left(\sum_{l=1}^{d} \alpha_{il} V_l\right)\left(\sum_{l=1}^{d} \alpha_{il} V_l^T\right)$$

$$\Rightarrow (N-1)CV_K = \sum_{i=1}^{N} \left(\sum_{l=1}^{d} \alpha_{il} V_l\right)\left(\sum_{l=1}^{d} \alpha_{il} V_l^T\right) V_K$$

$$\Rightarrow (N-1)\lambda_K V_K = \sum_{i=1}^{N} \left(\sum_{l=1}^{d} \alpha_{il} V_l\right) \alpha_{iK}$$

(As $CV_K = \lambda_K V_K$, $V_l^T V_K = 0$ for $l \neq K$, $V_K^T V_K = 1$ (normal))

$$\Rightarrow (N-1)\lambda_K V_K = \sum_{l=1}^{d} \left(\sum_{i=1}^{N} \alpha_{il} \alpha_{iK}\right) V_l$$

$$\Rightarrow (N-1)\lambda_K V_K = \sum_{l=1,\, l\neq K}^{d} \left(\sum_{i=1}^{N} \alpha_{il} \alpha_{iK}\right) V_l + \left(\sum_{i=1}^{N} (\alpha_{iK})^2\right) V_K$$

$$\Rightarrow \sum_{l=1,\, l\neq K}^{d} \left(\sum_{i=1}^{N} \alpha_{il} \alpha_{iK}\right) V_l + \left(\left(\sum_{i=1}^{N} (\alpha_{iK})^2\right) - (N-1)\lambda_K\right) V_K = 0$$

→ As $V_1, V_2 \cdots V_d$ are linearly independent so $\lambda_1 V_1 + \lambda_2 V_2 \cdots \lambda_d V_d = 0$

$$\Rightarrow \lambda_1 = \lambda_2 = \cdots \lambda_d = 0$$

$$\Rightarrow \sum_{i=1}^{N} \alpha_{il} \alpha_{iK} = 0 \ (\text{if } l \neq K), \quad \sum_{i=1}^{N} \alpha_{iK}^2 = (N-1)\lambda_K$$

→ Now let us use these to find required answer

$$\frac{1}{N}\sum_{i=1}^{N}\|\tilde{x}_i - x_i\|^2 = \frac{1}{N}\sum_{i=1}^{N}(\tilde{x}_i - x_i)^T(\tilde{x}_i - x_i)$$

$$= \frac{1}{N}\sum_{i=1}^{N}\left(\sum_{l=k+1}^{d}\alpha_{il}V_l\right)^T\left(\sum_{l=k+1}^{d}\alpha_{il}V_l\right)$$

$$= \frac{1}{N}\sum_{i=1}^{N}\left(\sum_{l=k+1}^{d}\alpha_{il}V_l^T\right)\left(\sum_{l=k+1}^{d}\alpha_{il}V_l\right)$$

$$= \frac{1}{N}\sum_{i=1}^{N}\sum_{l=k+1}^{d}\alpha_{il}^2 = \frac{1}{N}\sum_{i=1}^{N}\sum_{l=k+1}^{d}(\alpha_{il})^2$$

$$= \frac{1}{N}\sum_{l=k+1}^{d}\sum_{i=1}^{N}(\alpha_{il})^2$$

$$= \frac{1}{N}\sum_{l=k+1}^{d}(N-1)\lambda_l = \underline{\frac{N-1}{N}\sum_{l=k+1}^{d}\lambda_l}$$

→ Now it is given all eigen values are very small (as all eigen value above k+1 are very small)

$\Rightarrow \sum_{l=k+1}^{d}\lambda_l$ is also small $\Rightarrow \frac{N-1}{N}\sum_{l=k+1}^{d}\lambda_l$ is also small

$\Rightarrow$ The error $\underline{\frac{1}{N}\sum_{i=1}^{N}\|\tilde{x}_i - x_i\|^2}$ is very small

d) Find principal components of (X_1, X_2)

→ We should find eigen vectors of covariance matrix of (X_1, X_2)

→ Covariance matrix of $(X_1, X_2) = \begin{bmatrix} Var(X_1) & CoVar(X_1, X_2) \\ CoVar(X_1, X_2) & Var(X_2) \end{bmatrix}$

a) In first part $Var(X_1) = 100$, $Var(X_2) = 1$, $CoVar(X_1, X_2) = 0$ (uncorrelated)

$\Rightarrow$ Covariance matrix $C = \begin{bmatrix} 100 & 0 \\ 0 & 1 \end{bmatrix}$

$CV = \lambda V$

$\Rightarrow (C - \lambda I)V = 0$

As $V \neq 0 \Rightarrow |C - \lambda I| = 0 \Rightarrow \begin{vmatrix} 100-\lambda & 0 \\ 0 & 1-\lambda \end{vmatrix} = 0$

$\Rightarrow \lambda = 1$ or $\lambda = 100$ are eigen values

1) Eigen vector with eigen value = 100

$CV = 100V \Rightarrow \begin{bmatrix} 100 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 100a \\ 100b \end{bmatrix}$

$\begin{bmatrix} 100a \\ b \end{bmatrix} = \begin{bmatrix} 100a \\ 100b \end{bmatrix}$

As $b = 100b \Rightarrow b = 0 \Rightarrow V = a\begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$\Rightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is of unit magnitude so eigen vector corresponding to $\lambda = 100$

2) Eigen vector with eigen value = 1

$CV = V \Rightarrow \begin{bmatrix} 100 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$

$\begin{bmatrix} 100a \\ b \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$

Page No.
Date

As $100a = a \Rightarrow a = 0$

$$\Rightarrow V = \begin{bmatrix} 0 \\ b \end{bmatrix} = b \begin{bmatrix} 0 \\ 1 \end{bmatrix} \rightarrow \text{unit vector}$$

$\Rightarrow \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is eigen vector corresponding to $\lambda = 1$

$\Rightarrow$ The principal components are $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

b) Given $Var(X_1) = Var(X_2)$ (Assume μ)

$$\Rightarrow C = \begin{bmatrix} \mu & 0 \\ 0 & \mu \end{bmatrix}$$

$CV = \lambda V$

$\Rightarrow (C - \lambda I)V = 0$ As $V \neq 0 \Rightarrow |C - \lambda I| = 0$

$$\Rightarrow \begin{vmatrix} \mu - \lambda & 0 \\ 0 & \mu - \lambda \end{vmatrix} = 0$$

$\rightarrow \lambda = \mu, \mu$ are eigen values

$\rightarrow$ Eigen vector corresponding to $\lambda = \mu$

$$\Rightarrow CV = \mu V \Rightarrow \begin{bmatrix} \mu & 0 \\ 0 & \mu \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \mu a \\ \mu b \end{bmatrix}$$

$$\downarrow$$

$$\begin{bmatrix} \mu a \\ \mu b \end{bmatrix} = \begin{bmatrix} \mu a \\ \mu b \end{bmatrix}$$

$\Rightarrow$ All vectors satisfy $CV = \mu V$

So this forms a vector space (As discussed in b) so any orthonormal basis can be taken as eigen vectors

i.e any two $\perp^{lar}$ vectors of unit magnitude works in this case and are principal components