Q2) Let $V_1, V_2 \cdots V_d$ are eigen vectors and $\lambda_1, \lambda_2 \cdots \lambda_d$ are corresponding eigen values of C. (which are sorted in order)

As $V_1, V_2, \cdots V_d$ are $\perp^{lar}$ they are linearly independent, no. of vector = dimension

$\Rightarrow V_1, V_2, \cdots V_d$ form orthonormal basis of $\mathbb{R}^d$ (spectral theorem)
(Proof in Q_1 and using this previous statement justify basis)

$$\Rightarrow \quad f = \sum_{i=1}^{d} a_i V_i$$

$\rightarrow$ As f is direction vector it should be unit magnitude $\Rightarrow f^T f = 1$

$$\Rightarrow f^T f = \left(\sum_{i=1}^{d} a_i V_i^T\right)\left(\sum_{i=1}^{d} a_i V_i\right) = \sum_{i=1}^{d} a_i^2 = 1$$

$\rightarrow$ We know $C = VDV^T$ where $V = [V_1 | V_2 | \cdots | V_d]$ and D is diagonal matrix $\begin{bmatrix} \lambda_1 & 0 & \cdots & \vdots \\ 0 & \lambda_2 & \ddots & \\ 0 & \cdots & \cdots & \lambda_d \end{bmatrix}$

$$f^T C f = f^T V D V^T f = (V^T f)^T D (V^T f)$$

$$V^T f = [V_1 | V_2 | \cdots | V_d]^T \sum_{i=1}^{d} a_i V_i = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_d \end{bmatrix}$$

$$\Rightarrow f^T C f = [a_1 \ a_2 \ \cdots \ a_d] \begin{bmatrix} \lambda_1 & \cdots & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ 0 & 0 & \lambda_3 & 0 \\ 0 & 0 & 0 & \lambda_d \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_d \end{bmatrix}$$

Page No.	
Date	

$$\Rightarrow f^T C f = \sum_{i=1}^{d} a_i^2 \lambda_i$$

→ Given $f \perp^{lar}$ to V_1

$$\Rightarrow V_1^T f = 0 \Rightarrow a_1 = 0$$

$$\Rightarrow \sum_{i=2}^{d} a_i^2 = 1 \quad , \quad \sum_{i=2}^{d} a_i^2 \lambda_i = f^T C f$$

$$\Rightarrow f^T C f = \lambda_2 + \sum_{i=3}^{d} a_i^2 (\lambda_i - \lambda_2)$$

Since $\lambda_2 \geq \lambda_i \ \forall i \geq 3$ so under these conditions $f^T C f$ is maximized when $a_i^2 = 0 \ \forall i \geq 3$

$$\Rightarrow a_2^2 = 1 \quad \Rightarrow \quad a_2 = \pm 1$$ (Both give same direction)

$$\Rightarrow f = \sum_{i=1}^{d} a_i V_i = V_2$$

→ Now extending this to $g \perp^{lar}$ to e, f $(e = V_1, f = V_2)$

$$g = \sum_{i=1}^{d} a_i V_i$$ (As $V_1, V_2 \ldots V_d$ form orthonormal basis)

$V_1^T g = 0 = a_1$ ($V_1 \perp^{lar}$ to g)

$V_2^T g = 0 = a_2$ ($V_2 \perp^{lar}$ to g)

$$\Rightarrow g = \sum_{i=3}^{d} a_i V_i$$

→ As $g^T g = 1 \Rightarrow \sum_{i=1}^{d} a_i^2 = 1 \Rightarrow \sum_{i=3}^{d} a_i^2 = 1$

$$g^T C g = \sum_{i=1}^{d} a_i^2 \lambda_i = \sum_{i=3}^{d} a_i^2 \lambda_i$$

$$\Rightarrow g^T C g = \lambda_3 + \sum_{i=4}^{d} a_i^2 (\lambda_i - \lambda_3)$$

As $\lambda_3 \geq \lambda_i \ \forall i \geq 4$ so $g^T C g$ is maximised when $a_i^2 = 0 \ \forall i \geq 4$

$\Rightarrow a_3^2 = 1 \Rightarrow a_3 = \pm 1$ (gives same direction)

$$\Rightarrow g = \sum_{i=1}^{d} a_i v_i = V_3$$

$\Rightarrow$ We can extend this to K^{th} being eigen vector corresponding to K^{th} largest eigen value which is an important step in PCA.