Assignment 4

Aakash
180260001

Q2) Given a set of N vectors $\mathcal{X} = \{x_1, x_2, \ldots, x_N\}$ each in $\mathbb{R}^d$, with $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$. Direction $\vec{e}$ such that $\sum_{i=1}^{N} \|(x_i-\bar{x}) - (e\cdot(x_i-\bar{x}))e\|^2$ is minimized is obtained by maximizing $e^t C e$, with C is the covariance matrix of vectors in $\mathcal{X}$, is the eigenvector $\underline{e}$ with the highest eigenvalue. Considering another direction f such that $f^t C f$ is maximised is, using the loss function (f is $\perp$ to e),

$$J(e,f) = \sum_{i=1}^{N} \|(x_i-\bar{x}) - (e^t(x_i-\bar{x}))e - (f^t(x_i-\bar{x}))f\|^2$$

$$= \|(x_i-\bar{x})\|^2 + \|(e^t(x_i-\bar{x}))e\|^2 + \|(f^t(x_i-\bar{x}))f\|^2$$
$$- 2\sum_{i=1}^{N}\{(e^t(x_i-\bar{x}))e^t(x_i-x) + (f^t(x_i-\bar{x}))f^t(x_i-x)\}$$

$$J(e,f) = -(e^tSe + f^tSf) + \sum_{i=1}^{N}\|x_i-\bar{x}\|^2$$

(Derivation similar to that done in class) $\left\{ 2(e^t(x_i-\bar{x}))e^t f(f^t(x_i-\bar{x})) \right.$ $\xrightarrow{} 0$ as e & f are perpendicular to each other.

Now, maximising

$\tilde{J}(e,f) = e^tSe + f^tSf$ with constraints $e^te = 1$ & $f^tf = 1$ and using Lagrange multipliers

$$\tilde{J}(e,f) = e^tSe + f^tSf - \lambda(e^te-1) - \mu(f^tf-1)$$

taking derivatives w.r.t $\vec{e}$ & $\vec{f}$ we get,

$$\frac{d\tilde{J}}{de} = 2Se - 2\lambda e = 0 \quad \& \quad \frac{d\tilde{J}}{df} = 2Sf - 2\mu f = 0$$

$$\Rightarrow \quad \underline{Se = \lambda e} \quad \& \quad \underline{Sf = \mu f}$$

$\Rightarrow$ $\underline{e^tSe = \lambda}$ & $\underline{f^tSf = \mu}$ as $S = (N-1)C$, it is a symmetric, positive semi-definite matrix i)

Thus using i) as $\underline{e \;\&\; f \text{ are orthonormal}}$ & $\lambda \neq \mu$. Thus to maximise $\tilde{J}(e,f)$, $\vec{e}$ corresponds to the eigenvalue which is the largest & $\vec{f}$ corresponds to the 2nd largest eigenvalue of C.

And due to $\underline{C \text{ being symmetric}}$ and $\underline{f \text{ being } \perp^{r}}$ to $e \Rightarrow \lambda \neq \mu$ & $\underline{\lambda > \mu}$.