CS 663 Assignment 4

Sarthak Consul - 16D100012
Parthasarathi Khirwadkar - 16D070001
Bhishma Dedhia - 16D170005

QUESTION 5

$$J(\boldsymbol{e}) = \sum_{i=1}^{N} ||\boldsymbol{x_i} - \bar{\boldsymbol{x}} - \boldsymbol{e}^\intercal(\boldsymbol{x_i} - \bar{\boldsymbol{x}})\boldsymbol{e}||^2$$
$$= \sum_{i=1}^{N} ||\boldsymbol{x_i} - \bar{\boldsymbol{x}}||^2 - \boldsymbol{e}^\intercal \boldsymbol{S}\boldsymbol{e} \quad (\boldsymbol{S} = (N-1)\boldsymbol{C})$$

Thus the direction, $\boldsymbol{e}$, for which $J(\boldsymbol{e})$ is minimized is that which maximizes $\boldsymbol{e}^\intercal\boldsymbol{C}\boldsymbol{e}$.
Let the eigenvalues of S be $\lambda_1 > \lambda_1 > \cdots > \lambda_{n-1}$
To find the $\boldsymbol{f}$ that maximizes $J(\boldsymbol{f})$ given $\boldsymbol{f} \perp \boldsymbol{e}$ and $||\boldsymbol{f}|| = 1$, applying the Lagrangian Multiplier Test,

$$\nabla_f \Big[\boldsymbol{f}^\intercal\boldsymbol{S}\boldsymbol{f} - \lambda(\boldsymbol{f}^\intercal\boldsymbol{f} - 1) - \beta(\boldsymbol{e}^\intercal\boldsymbol{f})\Big] = 0$$
$$2\boldsymbol{S}\boldsymbol{f} - 2\lambda\boldsymbol{f} - \beta\boldsymbol{e} = 0$$

Pre-multipying both sides with $\boldsymbol{e}^\intercal$,

$$2(\boldsymbol{S}^\intercal\boldsymbol{e})^\intercal\boldsymbol{f} - 2\lambda\boldsymbol{e}^\intercal\boldsymbol{f} - \beta\boldsymbol{e}^\intercal\boldsymbol{e} = 0$$

As $\boldsymbol{f} \perp \boldsymbol{e}$, $\boldsymbol{e}^\intercal\boldsymbol{f} = 0$. Also $\boldsymbol{S}$ is symmetric and so $\boldsymbol{S}^\intercal\boldsymbol{e} = \boldsymbol{S}\boldsymbol{e} = \lambda_1\boldsymbol{e}$ and finally $\boldsymbol{e}^\intercal\boldsymbol{e} = 1$

$$\implies 2\lambda_1\boldsymbol{e}^\intercal\boldsymbol{f} - \beta = 0$$
$$\implies \beta = 0$$
$$\text{Thus } \boldsymbol{S}\boldsymbol{f} = \lambda\boldsymbol{f} \ \big(\text{i.e. } \boldsymbol{f} \text{ is an eigenvector of } \boldsymbol{S}\big)$$

Let the eigenvalue of $\boldsymbol{S}$ corresponding to $\boldsymbol{f}$ be μ

$$J(\boldsymbol{f}) = \mu^2 \ \text{ (As } \boldsymbol{f} \text{ is of unit norm)}$$

Maximizing $J(\boldsymbol{f})$ would mean that (f) corresponds to the largest eigenvalue, given that it should be orthogonal to $\boldsymbol{e}$.
$\mu \neq \lambda_1$ as it is given that $\boldsymbol{C}$ and so $\boldsymbol{S}$ has distinct eigenvalues and if $\mu = \lambda_1$, that would mean $\boldsymbol{f} = \boldsymbol{e}$ (and so $\boldsymbol{f} \not\perp \boldsymbol{e}$) $\Rightarrow\!\Leftarrow$
The next best choice of μ is λ_2 and since eigenvectors of symmetric matrices are orthogonal, no contradictions arise.
Thus $\boldsymbol{f}$ has to be the eignevector corresponding to the second largest eigenvalue