CS-663 Assignment 4 Q2

Soham Naha (193079003)
Akshay Bajpai (193079002)
Mohit Agarwala (19307R004)

November 6, 2020

2

Consider a set of N vectors $\mathcal{X} = \{\boldsymbol{x_1}, \boldsymbol{x_2}, ..., \boldsymbol{x_N}\}$ each in $\mathbb{R}^d$, with average vector $\bar{\boldsymbol{x}}$. We have seen in class that the direction $\boldsymbol{e}$ such that $\sum_{i=1}^{N} \|\boldsymbol{x_i} - \bar{\boldsymbol{x}} - (\boldsymbol{e} \cdot (\boldsymbol{x_i} - \bar{\boldsymbol{x}}))\boldsymbol{e}\|^2$ is minimized, is obtained by maximizing $\boldsymbol{e}^t \boldsymbol{C} \boldsymbol{e}$, where $\boldsymbol{C}$ is the covariance matrix of the vectors in $\mathcal{X}$. This vector $\boldsymbol{e}$ is the eigenvector of matrix $\boldsymbol{C}$ with the highest eigenvalue. Prove that the direction $\boldsymbol{f}$ perpendicular to $\boldsymbol{e}$ for which $\boldsymbol{f}^t \boldsymbol{C} \boldsymbol{f}$ is maximized, is the eigenvector of $\boldsymbol{C}$ with the second highest eigenvalue. For simplicity, assume that all non-zero eigenvalues of $\boldsymbol{C}$ are distinct and that rank($\boldsymbol{C}$) > 2. [10 points]

Proof:

We have to maximise f^tCf given the constraints:

- f is perpendicular to e
- f is a unit vector
- $e^tCe = \lambda_1$.

We use the method of Lagrange multipliers to do so, as in the proof in class, with two constraints:

- $f^tf - 1 = 0$
- $f^te = 0$

So, the lagrangian equation to solve becomes :

$$J(f) = f^tCf - \lambda(f^tf - 1) - \delta(f^te)$$

Taking derivative of $J(f)$ with respect to f^t and setting it to zero, we get

$$2Cf - 2\lambda f - \delta e = 0$$

Left multiplying both sides by e we get

$$2e^tCf - 2\lambda e^tf - \delta e^te = 0$$

We have $Ce = \lambda_1 e$ or taking transpose both sides, $e^tC^t = \lambda_1 e^t$.
Now, as C is a co-variance matrix, hence it is symmetric as co-variance of x_i and x_j is the same as co-variance of x_j and x_i. Thus we have $e^tC^t = e^tC = \lambda_1 e^t$.
Right multiplying with f we get $e^tCf = \lambda_1 e^tf = 0$ As $e^tf = 0$ (i.e. f and e are perpendicular) and $e^tCf = 0$, hence $\delta = 0$ and so

$$Cf - \lambda f = 0$$
$$f^tCf = \lambda$$

To maximise f_tCf we have to choose the largest λ. As we have assumed that $rank(C) > 2$ and as e is the eigenvector corresponding to λ_1 hence the next largest value of λ can be λ_2 (the second largest eigenvalue).
Thus f is the eigenvector corresponding to λ_2, the second largest eigen value.