Shaan Ul Haque(180070053)
Samarth Singh (180050090)
Niraj Mahajan (180050069)

1 Answer

We need to prove that after the eigen-vector corresponding to largest eigen-value is chosen the next best choice is to choose the eigen-vector with second largest eigen value. Let $\mathbf{e}$, be the eigen-vector with largest eigen-value. Let $\mathbf{f}$ be a unit vector perpendicular to $\mathbf{e}$. Then according to question we want to minimize the cost function:

$$J(\mathbf{f}) = \sum_{i=1}^{N} ||b_i\mathbf{f} - (\mathbf{x_i} - \mathbf{x} - a_i\mathbf{e})||^2$$

where, $a_i = \mathbf{e^T}(\mathbf{x_i}\text{-}\mathbf{x})$; $b_i = \mathbf{f^T}(\mathbf{x_i}\text{-}\mathbf{x})$; $\mathbf{x}$=average value of $\mathbf{x_i}$

$$J(\mathbf{f}) = \sum_{i=1}^{N} ||b_i\mathbf{f}||^2 + \sum_{i=1}^{N} ||\mathbf{x_i} - \mathbf{x} - a_i\mathbf{e}||^2 - 2\sum_{i=1}^{N} b_i\mathbf{f}^T(\mathbf{x_i} - \mathbf{x} - a_i\mathbf{e})$$

$$J(\mathbf{f}) = \sum_{i=1}^{N} b_i^2 + \sum_{i=1}^{N} ||\mathbf{x_i} - \mathbf{x} - a_i\mathbf{e}||^2 - 2\sum_{i=1}^{N} ||b_i\mathbf{f}||^2$$

$$\implies J(\mathbf{f}) = -\sum_{i=1}^{N} b_i^2 + \sum_{i=1}^{N} ||\mathbf{x_i} - \mathbf{x} - a_i\mathbf{e}||^2$$

(Since, $\mathbf{f}$ is a unit vector, $\mathbf{f}^T\mathbf{e} = 0,\ \mathbf{f}^T(\mathbf{x_i} - \mathbf{x}) = b_i$)
We want to minimize J($\mathbf{f}$) and to minimize it we should maximize the first term in the latter equation.

$$\sum_{i=1}^{N} b_i^2 = \sum_{i=1}^{N} \mathbf{f}^T(\mathbf{x_i} - \mathbf{x})(\mathbf{x_i} - \mathbf{x})^T\mathbf{f}$$

We define $\sum_{i=1}^{N}(\mathbf{x_i} - \mathbf{x})(\mathbf{x_i} - \mathbf{x})^T = S$, where S = (N-1)C and C is the covariance matrix
Thus the above equation becomes,

$$\sum_{i=1}^{N} b_i^2 = \mathbf{f}^T\mathbf{S}\mathbf{f}$$

Thus, we need to maximize the above equation under the constraint, $\mathbf{f}^T\mathbf{f} = 1$. Using Langrange's multiplier method we get,

$$G(\mathbf{f}) = \mathbf{f}^T\mathbf{S}\mathbf{f} - \lambda(\mathbf{f}^T\mathbf{f} - 1)$$
$$G'(\mathbf{f}) = \mathbf{S}\mathbf{f} - \lambda\mathbf{f} = 0$$
$$\implies \mathbf{S}\mathbf{f} = \lambda\mathbf{f}$$

Thus, $\mathbf{f}$ needs to be an eigenvector of $\mathbf{S}$. Rearranging the equation gives,

$$\mathbf{f}^T\mathbf{S}\mathbf{f} = \lambda$$

Thus, to maximize $\mathbf{f}^T\mathbf{S}\mathbf{f}$ we need maximum eigen-value. But we have already chosen the largest eigenvalue and its corresponding vector, thus the next best option is the second largest eigenvalue. Hence, proved that $\mathbf{f}$, which is perpendicular to $\mathbf{e}$, is the eigenvector corresponding to the second largest eigenvalue.

1.1 Another approach

We wanted the maximize $\mathbf{f}^T\mathbf{S}\mathbf{f}$. A more intuitive solution which also gives a physical interpretation is as follows:

$$\mathbf{f}^T\mathbf{S}\mathbf{f} = \mathbf{f}^T(\mathbf{S}\mathbf{f})$$

Let $\mathbf{f}^T\mathbf{S} = \mathbf{c}$, and the above equation becomes $\mathbf{f}^T\mathbf{c}$. The dot product of two vectors is maximized (under the constraint of constant magnitude of vectors) when angle between the two is $0°$ and thus $\mathbf{c}$ should be $\alpha\mathbf{f}$, where α is some constant. Written mathematically,

$$\mathbf{S}\mathbf{f} = \alpha\mathbf{f}$$

Thus, $\mathbf{f}$ is an eigenvector of $\mathbf{S}$ matrix.