Assignment 4 - Q5

Rishabh Shah, Shriram SB, Anmol Mishra
CS663

October 15, 2018

Exercise 5. Consider a set of N vectors $X = \{x_1, x_2, ..., x_N\}$ each in R^d with average vector $\bar{x}$. We have seen in class that the direction e such that $\sum_{i=1}^{N} ||x_i - \bar{x} - (e.(x_i - \bar{x}))e||^2$ is minimized, is obtained by maximizing $e^T Ce$, where C is the covariance matrix of the vectors in X. This vector e is the eigenvector of matrix C with the highest eigenvalue. Prove that the direction f perpendicular to e for which $f^T Cf$ is maximized, is the eigenvector of C with the second highest eigenvalue. For simplicity, assume that all non-zero eigenvalues of C are distinct and that $rank(C) > 2$.

Proof. We try to minimize the error of approximation J:

$$
\begin{aligned}
J\left(e_2\right) &= \sum_{i=1}^{N} ||(x_i - \bar{x}) - (e_1.(x_i - \bar{x}))e_1 - (e_2.(x_i - \bar{x}))e_2||^2 \\
&= \sum_{i=1}^{N} ||(x_i - \bar{x})||^2 + \sum_{i=1}^{N} ||a_i e_1||^2 + \sum_{i=1}^{N} ||b_i e_2||^2 - 2\sum_{i=1}^{N} a_i e_1^T (x_i - \bar{x}) - 2\sum_{i=1}^{N} b_i e_2^T (x_i - \bar{x}) \\
&= \sum_{i=1}^{N} ||(x_i - \bar{x})||^2 + \sum_{i=1}^{N} a_i^2 + \sum_{i=1}^{N} b_i^2 - 2\sum_{i=1}^{N} a_i^2 - 2\sum_{i=1}^{N} b_i^2 \\
&= \sum_{i=1}^{N} ||(x_i - \bar{x})||^2 - \sum_{i=1}^{N} a_i^2 - \sum_{i=1}^{N} b_i^2 \\
&= -\sum_{i=1}^{N} (e_2(x_i - \bar{x}))^2 + \sum_{i=1}^{N} ||(x_i - \bar{x})||^2 - \sum_{i=1}^{N} a_i^2 \\
&= -\sum_{i=1}^{N} (e_2^T (x_i - \bar{x})^T (x_i - \bar{x}) e_2) + \sum_{i=1}^{N} ||(x_i - \bar{x})||^2 - \sum_{i=1}^{N} a_i^2 \\
&= -e_2^T S e_2 + \sum_{i=1}^{N} ||(x_i - \bar{x})||^2 - \sum_{i=1}^{N} a_i^2
\end{aligned}
$$

Minimizing J with respect to e_2 is equivalent to maximizing $e_2^T S e_2$. We use the method of Lagrange Multipliers with the constraints, $e_2^T e_2 = 1$ and $e_2^T e_1 = 0$.

$\tilde{J}(e_2) = e_2^T S e_2 - \lambda(e_2^T e_2 - 1) - \gamma(e_2^T e_1)$

Taking derivative of $\tilde{J}$ w.r.t e_2 and setting it to 0, we get
$Se_2 - \lambda e_2 - \gamma e_1 = 0$

Since e_2 is a vector in the space perpendicular to e_1, it can be expressed as $e_2 = \sum_{i=2}^{N} \mu_i v_i$ where v_i are the eigenvectors of S.

$\sum_{i=2}^{N} \mu_i S v_i - \lambda e_2 - \gamma e_1 = 0$
$\sum_{i=2}^{N} \mu_i \lambda_i v_i - \lambda e_2 - \gamma e_1 = 0$

Multiplying with e_1^T, we get $\gamma e_1^T e_1 = 0$.(Since $v_i^T v_j = 0 \; \forall i \neq j$)
Since e_1 is not zero vector, γ is zero.
Thus, $Se_2 = \lambda e_2$, so e_2 is an eigenvector of S.
As, $e_2^T S e_2 = \lambda$ and we wish to maximize $e_2^T S e_2$ under the constraint $e_2^T e_1 = 0$, we would choose e_2 to be the eigenvector corresponding to $\mu_2 = 1$ and $\mu_i = 0 \forall i \neq 2$. Thus, e_2 is the eigenvector corresponding to the second largest eigenvalue of S.
Hence proved.

$\square$