High Dimensional Statistics | Prof. Dr. Podolskij Mark | Homework 4

Anton Zaitsev | 0230981826 | anton.zaitsev.001@student.uni.lu | University of Luxembourg

May 7, 2024

Exercise 2

Let X be a k-dimensional vector with $\mathbb{E}[X] = 0 \in \mathbb{R}^k$ and $\mathbb{E}[XX^T] = \Sigma \in \mathbb{R}^{k \times k}$. The first principal component $\beta_1 \in \mathbb{R}^k$ is the eigenvector of Σ that corresponds to the largest eigenvalue λ_1 of Σ. The goal is to find vector $\beta \in \mathbb{R}^k$ s.t. $||\beta||^2 = 1, < \beta_1, \beta >= 0$ which maximizes $var(\beta^T X)$ and to deduce that β is the eigenvector of Σ that corresponds to the second largest eigenvalue λ_2 of Σ, i.e. $\beta = \beta_2$.

Notice that

$$var(\beta^T X) = \beta^T \mathbb{E}[XX^T]\beta = \beta^T \Sigma \beta$$

Since Σ is the covariance matrix, is it symmetric and positive definite. We can perform eigendecomposition of Σ:

$$\Sigma = O \Lambda O^T$$

Here, $O \in \mathbb{R}^{k \times k}$ - matrix of orthonormal (unitary in length and orthogonal to each other) eigenvectors of $\Sigma : O = [\beta_1 | \beta_2 | \dots | \beta_k]$
Matrix $\Lambda \in \mathbb{R}^{k \times k}$ - diagonal matrix, where each element on the diagonal is the eigenvalue of Σ s.t. $\Lambda_{11} = \lambda_1, \Lambda_{22} = \lambda_2, \dots, \Lambda_{kk} = \lambda_k, \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k$. Note that each column O_i is the eigenvector of λ_i. Then, we have the following problem:

$$\underset{||\beta||^2=1, <\beta,\beta_1>=0}{\text{argmax}} \{\beta^T O \Lambda O^T \beta\}$$

Note that since we seek for a vector β that is orthogonal to β_1 and unitary in length, we then essentially seek for a vector in the set $\{\beta_2, \dots, \beta_k\}$. This is due to the fact that eigenvectors in O are orthonormal and form a basis for the vector space $\mathbb{R}^{k \times k}$. The problem is then can be rewritten in the following way:

$$\underset{||\beta||^2=1, <\beta,\beta_1>=0}{\text{argmax}} \{\beta^T O \Lambda O^T \beta\} =$$

$$= \underset{<\beta_i,\beta_1>=0, i \in \{2,\dots,k\}}{\text{argmax}} \left\{ \begin{bmatrix} \beta_{i1} \beta_{i2} \dots \beta_{ik} \end{bmatrix}_{1 \times k} \begin{bmatrix} \beta_{11} \dots \beta_{d1} \\ \beta_{12} \dots \beta_{d2} \\ \vdots \\ \beta_{1d} \dots \beta_{dd} \end{bmatrix}_{k \times k} \begin{bmatrix} \lambda_1 0 \dots 0 \\ 0 \lambda_2 \dots 0 \\ \vdots \\ 00 \dots \lambda_d \end{bmatrix}_{k \times k} \begin{bmatrix} \beta_{11} \dots \beta_{1d} \\ \beta_{21} \dots \beta_{2d} \\ \vdots \\ \beta_{d1} \dots \beta_{dd} \end{bmatrix}_{k \times k} \begin{bmatrix} \beta_{i1} \\ \beta_{i2} \\ \vdots \\ \beta_{ik} \end{bmatrix}_{k \times 1} \right\}$$

When we multiply vector β^T with O, we get a row vector of size $1 \times k$ with 1 at position i and zeros on other positions: $\beta^T O = \begin{bmatrix} 0_1 \dots 1_i \dots 0_k \end{bmatrix}_{1 \times k}$. Then we scale this vector by λ_i (remember that λ_i's are ordered in the decreasing order in Λ): $\begin{bmatrix} 0_1 \dots \lambda_i \dots 0_k \end{bmatrix}_{1 \times k}$. For $O^T \beta$, we get a column vector of size $k \times 1$ with 1 at position i and zeros on other positions: $O^T \beta = \begin{bmatrix} 0_1 \\ \vdots \\ 1_i \\ \vdots \\ 0_k \end{bmatrix}_{k \times 1}$. Finally, we take the product of the row vector and column vector and get λ_i as a result. The largest eigenvalue for $i \in \{2, \dots, k\}$ is λ_2. Thus, vector β, which is unitary in length and orthogonal to β_1 and that maximizes $var(\beta^T X)$ is β_2 - eigenvector of Σ that corresponds to the second largest eigenvalue of Σ.