MS-E2112 - Multivariate statistical analysis – Home exercise 3

Srikanth Gadicherla, 546195

January 29, 2016

Home Exercise 3: Maximizing Variance

Let x denote a p variate random vector with a finite mean vector μ and a finite covariance matrix Σ and let y_k denote the k^{th} principal component of x. Let $b \in \mathbb{R}^p$, $b^T b = 1$. Assume that $b^T x$ is uncorrelated with first $k-1$ components of x. Show that $var(y_k) \geq var(b^T x)$

Proof Let $y = \Gamma^T (x - \mu)$ where $\Gamma \in \mathbb{R}^{p \times p}$ is Orthogonal.

where $\Gamma^T \Sigma \Gamma = \Lambda = diag(\lambda_1, \dots \lambda_p)$ is diagnol &

$$\lambda_1 \geq \dots \geq \lambda_p.$$

Let ϑ_i denote i^{th} column of Γ. Assume that $b^T b = 1$.

Since the set $\{\vartheta_1, \dots \vartheta_p\}$ is an orthonormal basis of $\mathbb{R}^p$, the vector b can be written as

$$b = c_1 \vartheta_k + c_2 \vartheta_{k+1} + \dots + c_p \vartheta_p ..$$

Now, since first $k-1$ components are uncorrelated for b, so $b^T b = 0$ for those. $\vartheta_i^T \vartheta_i = 0$ for first '$k-1$' components.

$$\therefore var[b^T x] = b^T \Sigma b = \sum_{j=k}^{p} c_j \vartheta_j^T \left(\sum_{i=k}^{p} \lambda_i \vartheta_i \vartheta_i^T \right) \sum_{\hat{k}=k}^{p} c_{\hat{k}} \vartheta_{\hat{k}}$$

$$= \sum_{i=k}^{p} \lambda_i c_i^2$$

Since $b^T b = 1$, we have that $\sum_{i=k}^{p} c_i^2 = 1$.

Thus λ_k is the largest eigenvalue, the variance of $b^T x$ i.e., $var[b^T x]$ is maximized when $c_k = 1$ and $c_i = 0$ $i \neq k$ & $i > k$, and consequently $b = \vartheta_k$.

$$\therefore \quad Var[y_k] \geq Var[b^T x]$$