UML2012

Exercise set 3

Solutions to be presented in the 30.3.2012 session

Exercise 1:

Denote by $U = (\mathbf{u}_1, \ldots, \mathbf{u}_m)$ the first $m \leq p$ principal component directions (weights) of the zero mean random variable $\mathbf{x} \in \mathbb{R}^p$. Assume you have n observations of $\mathbf{x}$, organized in the matrix $X \in \mathbb{R}^{p \times n}$.

1.1 Explain why the sample covariance can be expressed as $\frac{1}{n} X X^T$.

1.2 Let $\mathbf{z} = U^T \mathbf{x}$ and $Z = U^T X$. Give the horizontal rows of Z an interpretation in terms of PCA.

1.3 Show that the rows of Z are orthogonal to each other, and give an interpretation for this.

Exercise 2:

2.1 Assume the covariance matrix C of a random vector $\mathbf{x} \in \mathbb{R}^2$ has the form

$$C = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \tag{1}$$

Calculate the eigenvalues as function of ρ (by hand). What is the effect of correlation between the random variables on the eigenvalues?

2.2 Let $x_2 = a x_1 + \varepsilon$ so that

$$\mathbf{x} = \begin{pmatrix} x_1 \\ a x_1 + \varepsilon \end{pmatrix} \tag{2}$$

where a is some deterministic constant, and ε is a random variable which is uncorrelated with x_1. Also assume that $E(\mathbf{x}) = 0$. How do you have to choose the factor a and the noise ε so that $\mathbf{x}$ has covariance matrix C?

2.3 Calculate the variance of ε (the noise variance) and make a scatter plot of $\mathbf{x}$ for $\rho \in \{-1, -0.25, 0, 0.5, 1\}$ (either sketch by hand or make the plots with matlab/R)

2.4 Suppose $X \in \mathbb{R}^{2 \times n}$ is a sample of n observations of $\mathbf{x}$. What happens to the horizontal vectors X_{1*} and X_{2*} when $|\rho| \to 1$?

Exercise 3:

The objective function for a single factor is

$$J_{ls}(\mathbf{a}) = \|C - \mathbf{a}\mathbf{a}^T\|^2, \tag{3}$$

where C is the covariance matrix, and the squared norm $\|A\|^2$ of a matrix A is defined as $\|A\|^2 = \sum_{ij} A_{ij}^2$.

3.1 Show that $\sum_{ij} A_{ij}B_{ij} = \mathrm{Tr}(A^T B) = \mathrm{Tr}(AB^T)$.

3.2 Write down the details of how to get from Equation (3) to

$$J_{ls} = \|\mathbf{a}\|^4 - 2\mathbf{a}^T C\mathbf{a} + \mathrm{Tr}(CC^T) \tag{4}$$

3.3 Calculate the gradient of J_{ls}, as defined in Equation (4), with respect to $\mathbf{a}$.

3.4 Show that if the gradient is zero for some vector $\mathbf{v}$, then $\mathbf{v}$ is an eigenvector. Let $\mathbf{e}$ be an eigenvector of C with unit norm. Find all the scalars α such that the vector $\mathbf{a}^* = \alpha\mathbf{e}$ makes the gradient of J_{ls} zero.

3.5 Calculate the value of J_{ls} for $\mathbf{a}^*$. Conclude that the eigenvector $\mathbf{e}$ that has the eigenvalue with the largest norm minimizes J_{ls}.

Exercise 4:

The optimization problem for quartimax is to maximize

$$J(U) = \sum_{ij} G((AU)_{ij}) \tag{5}$$

under the constraint of orthogonality of U (see Section 5.5).

4.1 Derive an iterative update rule for the matrix U which solves the optimization problem for $G(y) = y^4$.

4.2 Show that $J(U)$ is constant for all orthogonal U if $G(y) = y^2$.