Chair for Theoretical Information Technology | RWTH AACHEN UNIVERSITY

Prof. Dr. Rudolf Mathar, Dr. Gholamreza Alirezaei, Dr. Arash Behboodi

Exercise 7

Friday, December 23, 2016

Problem 1. *(Fisher's Linear Discriminant Function for two classes)* If Fisher's linear discriminant function is used for classification into two classes C_1 and C_2, prove that an observation $\mathbf{x}$ is allocated to C_1 if $\mathbf{a}^T(\mathbf{x} - \frac{1}{2}(\overline{\mathbf{x}}_\mathbf{1} + \overline{\mathbf{x}}_\mathbf{2})) > 0$ with $\mathbf{a} = \mathbf{W}^{-1}(\overline{\mathbf{x}}_\mathbf{1} - \overline{\mathbf{x}}_\mathbf{2})$.

Problem 2. *(ML Discriminant Rule for two classes)* Suppose that ML discriminant rule is used for classification into two classes C_1 and C_2. The class distributions are Gaussian and known as $N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ and $N_p(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$ with $\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}$. The densities are:

$$f_l(\mathbf{u}) = \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{u} - \boldsymbol{\mu}_l)^T\boldsymbol{\Sigma}^{-1}(\mathbf{u} - \boldsymbol{\mu}_l)\right\}, \mathbf{u} \in \mathbb{R}^p . l = 1, 2.$$

Prove that the ML rule allocates $\mathbf{x}$ to the class C_1 if

$$\boldsymbol{\alpha}^T(\mathbf{x} - \boldsymbol{\mu}) > 0,$$

where $\boldsymbol{\alpha} = \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$ and $\boldsymbol{\mu} = \frac{1}{2}(\boldsymbol{\mu}_1 + \boldsymbol{\mu}_2)$.

Problem 3. *(Eigenvalues in Fisher's Linear Discriminant Analysis)* Let $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n]^T \in \mathbb{R}^{n\times p}$ be samples and $\mathbf{W}$ and $\mathbf{B}$ are matrices corresponding to within-group and between-group sum of squares. Define $\mathbf{S} = \mathbf{X}^T\mathbf{E}_n\mathbf{X}$. Suppose that $\mathbf{W}$ has rank p. Show that the following three eigenvectors are the same:

a) the eigenvector correpsonding to the largest eigenvalue of $\mathbf{W}^{-1}\mathbf{B}$

b) the eigenvector correpsonding to the largest eigenvalue of $\mathbf{W}^{-1}\mathbf{S}$

c) the eigenvector correpsonding to the smallest eigenvalue of $\mathbf{S}^{-1}\mathbf{W}$