Biometric Person Authentication – Seminar
Summer Semester 2023
Quality and Usability Lab, TU Berlin

Assignment 1 – Due on 05 June 2023 at 12:00 CEST

Number of questions = 2, Total marks = 8

Question 1 - Bivariate Statistics (5 marks)

Preliminaries:

- Download the file DavisData.mat from ISIS to your computer and load the file into your Matlab Workspace using the Matlab function `load()`.
- The Workspace will now contain two matrices, xF of size 112x2 and xM of size 88x2. They contain 112 and 88 samples, respectively, of observations for female and male students in a university course.
- Each sample comprises the student's body mass in kg and body height in cm.

For each class, determine the *bivariate normal distribution* of mass and height values.

a) Determine the mean vectors $\boldsymbol{\mu}_F$ and $\boldsymbol{\mu}_M$ using the Matlab function `mean()` **[1 mark]**

b) Determine the covariance matrices $\boldsymbol{\Sigma}_F$ and $\boldsymbol{\Sigma}_M$ using the Matlab function `cov()` **[1 mark]**
Note: `cov()` offers 2 options for calculating the covariance. Choose the unbiased option.

c) Determine the standard deviations $\boldsymbol{\sigma}_{Fm}$, $\boldsymbol{\sigma}_{Fh}$, $\boldsymbol{\sigma}_{Mm}$ and $\boldsymbol{\sigma}_{Mh}$ of mass and height from the diagonal elements of the covariances, and the correlation coefficients r_{Fmh} and r_{Mmh} between mass and height from the (identical!) off-diagonal elements of the covariance
Note: The correlation coefficient is defined as $r_{12} = \sigma_{12}/\sqrt{\sigma_{11}\sigma_{22}}$ where σ_{11} and σ_{22} are the diagonal elements of $\boldsymbol{\Sigma}$ and $\sigma_{12} = \sigma_{21}$ is the off-diagonal element of $\boldsymbol{\Sigma}$. **[1 mark]**

d) Explain in plain English the meaning of all the statistics you have calculated in a) to c) and what those statistics tell you about this population of 200 university students **[2 marks]**

Question 2 - Pattern recognition (3 marks)

Given the 2 *bivariate normal models* $\mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}_F, \boldsymbol{\Sigma}_F)$ and $\mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}_M, \boldsymbol{\Sigma}_M)$ determined in Question 1 and the 4 test vectors $\boldsymbol{x}_1 = (62, 168)$, $\boldsymbol{x}_2 = (64, 170)$, $\boldsymbol{x}_3 = (66, 172)$ and $\boldsymbol{x}_4 = (68, 174)$, use the function `mvnpdf()` to calculate $p(\boldsymbol{x}_i|\boldsymbol{\mu}_F, \boldsymbol{\Sigma}_F)$ and $p(\boldsymbol{x}_i|\boldsymbol{\mu}_M, \boldsymbol{\Sigma}_M)$ for $i = 1, \cdots, 4$.

a) Classify the test vectors with equal prior probabilities and cost matrix $\lambda_{ij} = \delta_{ij}$ where δ_{ij} is the *Kronecker delta*, i.e. the decision cost is 0 for $i = j$ and 1 for $i \neq j$. **[1 mark]**

b) Classify the test vectors with prior probabilities of $P(\omega_F) = 0.4$ and $P(\omega_M) = 0.6$, and $\lambda_{ij} = \delta_{ij}$. **[1 mark]**

c) Classify the test samples with the prior probabilities of b) and cost matrix $\Lambda = (\lambda_{ij}) = \begin{pmatrix} 0 & 1.5 \\ 0.75 & 0 \end{pmatrix}$, i.e., the penalty for misclassifying a real male as female is $\lambda_{12} = 1.5$ and the penalty for misclassifying a real female as male is $\lambda_{21} = 0.75$. **[1 mark]**

--- End of Assignment 1 ---