

Computer experiment 1 (deadline: **3/6**)

1.

- 5.2 a.** (i) Generate four sets, each one consisting of 100 two-dimensional vectors, from the normal distributions with mean values $[-10, -10]^T$, $[-10, 10]^T$, $[10, -10]^T$, $[10, 10]^T$ and covariance matrices equal to $0.2 * I$. These sets constitute the data set for a four-class two-dimensional classification problem (each set corresponds to a class).
- a. (ii) Compute the S_w , S_b , and S_m scatter matrices.
- a. (iii) Compute the value for the criterion J_3 .
- b. Repeat (a) when the mean vectors of the normal distributions that generate the data are $[-1, -1]^T$, $[-1, 1]^T$, $[1, -1]^T$, $[1, 1]^T$.
- c. Repeat (a) when the covariance matrices of the normal distributions that generate the data are equal to $3 * I$.

2.

- 5.3** Generate two sets, each one consisting of 100 five-dimensional vectors, from the normal distributions with mean values $[0, 0, 0, 0, 0]^T$ and $[0, 2, 2, 3, 3]^T$ and covariance matrices equal to

$$\begin{bmatrix} 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1.5 \end{bmatrix}.$$

Their composition forms the data set for a two-class five-dimensional classification problem (each set corresponds to a class). Using the J_3 criterion find the best combination of features if:

- a. they are considered individually.
- b. they are considered in pairs.
- c. they are considered in triples.
- d. Justify the results.

3.

- 5.4 a.** (i) Generate two sets, each one consisting of 100 two-dimensional vectors, from the normal distributions with mean values $[2, 4]^T$ and $[2.5, 10]^T$ and covariance matrices equal to the 2×2 identity matrix I . Their composition forms the data set for a two class two dimensional classification problem (each set corresponds to a class).
- a. (ii) Compute the value of the FDR index for both features.
- b. Repeat (a) when the covariance matrices of the normal distributions that generate the data are both equal to $0.25 * I$.
- c. Discuss the results.

作業繳交規定

如果報告中只有答案或數據將會扣分：

手寫部分請寫上**計算過程**，

程式 Report 部分請**描述原理**並針對結果進行**討論、比較**

程式 Code 部分請上傳到**助教提供**之 **FTP**，勿上傳至 ilms

程式 **readme.txt** 部分請說明如何執行你的程式 (如何 **compile** 主程式...等)

程式語言沒有限制，可用 Matlab、C++、.....

程式作業評分

程式完成度:50 分 (有產生結果、結果正確性)

Report 內容:50 分 (描述原理 20 分+結果比較與討論 30 分)

抄襲程式經確認後該題零分。

繳交期限

請均於規定日期當天 **23:59** 之前繳交，本課堂作業遲交第一天分數會打**八折**(所得分數 = 原始分數 * **0.8**)，第二天以後視同缺交均為**零分**，請同學注意。

程式作業上傳

上傳內容:

Report 報告 (裡面不需要附 Code)、**Code 程式** (請加基本註解)、**readme.txt** (如何 compile 程式)

上傳方式:

使用 FTP 軟體，如 FileZilla，連線至

主機 140.114.85.181

使用者 PR2017

密碼 pr2017

登入後到當次作業資料夾，ex. Lab1

在資料夾內上傳 105062XXX.zip 檔，(記得要壓縮!)