

# Assignment 7

## Pattern Recognition - André E. Lazzaretti

1. Gere um conjunto de dados tridimensional com 900 exemplos, os quais correspondem a duas classes – sendo os 100 primeiros com média nula e matriz de covariância  $\mathbf{S}_1 = \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 0.01 \end{bmatrix}$ . Os demais estão organizados em 8 grupos de 100 exemplos. Cada grupo segue uma distribuição Gaussiana com a seguinte matriz de covariância:  $\mathbf{S}_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0.01 \end{bmatrix}$ . Sendo as médias:

- $[a \ 0 \ 0]$ ;
- $[a/2 \ a/2 \ 0]$ ;
- $[0 \ a \ 0]$ ;
- $[-a/2 \ a/2 \ 0]$ ;
- $[-a \ 0 \ 0]$ ;
- $[-a/2 \ -a/2 \ 0]$ ;
- $[0 \ -a \ 0]$ ;
- $[a/2 \ -a/2 \ 0]$ .

Considere  $a=6$ .

- a) Plote os dados em 3-D e visualize-os sob vários ângulos para perceber a organização dos dados no espaço.
  - b) Aplique o LDA nos dados. Projete os dados em um subespaço gerado pelos autovetores que correspondem aos autovalores não-negativos do produto das matrizes  $\mathbf{S}_w^{-1}\mathbf{S}_b$ . O que isso significa? Comente os resultados.
- 
2. Gere uma espiral de Arquimedes tridimensional como um pacote de 11 espirais de Arquimedes bidimensionais idênticas, uma após a outra. Uma espiral bidimensional é descrita em coordenadas polares pela equação  $r=a\theta$ , onde  $a$  é um parâmetro definido pelo usuário. No caso, os pontos em uma espiral 3-D são gerados como segue. Para cada  $\theta$ , utilize os valores de  $\theta_{\text{init}}$  até  $\theta_{\text{fin}}$  com passo  $\theta_{\text{step}}$  e calcule:
    - $R = a\theta$
    - $x = r\cos\theta$
    - $y = r\sin\theta$Os 11 pontos são na forma  $(x, y, z)$ , onde  $z = -1, -0.8, -0.6, \dots, 0.8, 1$  são pontos da espiral. Utilize  $a=0.1$ ,  $\theta_{\text{init}} = 0.5$ ,  $\theta_{\text{fin}} = 2.05*\pi$ ,  $\theta_{\text{step}} = 0.2$ . Plote a espiral 3-D de tal forma que todas as espirais 2-D sejam plotadas com o mesmo símbolo e todos os grupos de 11 pontos na forma  $(x, y, z)$ , onde  $x$  e  $y$  são fixados e  $z$  assume os valores  $-1, -0.8, -0.6, \dots, 0.8, 1$ , são plotados com a mesma cor.
  - a) Aplique kernel PCA para uma dimensão  $m=2$  para diferentes parâmetros do kernel Gaussiano e plote os resultados.
  - b) Repita o processo com PCA para os dois componentes principais.
  - c) Compare e comente os resultados.

Exemplo da espiral 2-D e 3-D:

