Pattern Recognition/Pattern Recognition

2016/2017

Exame Normal 19 June 2017 Duration: 2h00

Name: Tiago Silva **Number:** 20222176215 **Practical Class:**

AVISO

The Exam has a duration of 2h00m. The test is composed by five questions. The last question is a Matlab practical question. Each question must be answered in the framed box below it. Questions may be answered in Portuguese or English. This is a closed book test. You are allowed to use a calculator machine. As consultation you may use only 1 Page A4 with your own manuscript notes. Violation of the last rule ends up with exam cancellation, course failure and eventually you may be subject to disciplinary procedure. If you have any questions, you may ask. Good Luck!

Question	pts	Results	Graded by:
1)	20		
2)	20		
3)	20		
4)	10		
5)	30		

Graded by:

Name: **Number:** **Practical Class:**

Question 1 - Pre-processing

☐ **20pts**

Consider the pre-processing system in Figure 1. The system receives as input three features and has two main components: one that normalizes data by applying the z-score method; the other reduces data to one dimension by applying PCA. Consider that the system was trained resulting in the following system parameters:

- Normalizator
  - Mean vector: $\mu = \begin{bmatrix} 324.033 \\ 710.387 \\ 8.680 \end{bmatrix}$

    $3 \times 1 \times 1 \times 3 =$
  - Standard deviation vector: $\sigma = \begin{bmatrix} 201.353 \\ 361.216 \\ 2.277 \end{bmatrix}$
- PCA
  - Eigenvector matrix: $\mathbf{W} = \begin{bmatrix} -0.594 & -0.238 & -0.769 \\ -0.580 & -0.536 & 0.614 \\ -0.558 & 0.810 & 0.180 \end{bmatrix}$
  - Eigenvalues: $\lambda_1 = 2.768$, $\lambda_2 = 0.202$, and $\lambda_3 = 0.010$.

The eigenvectors are columns of $\mathbf{W}$ and are ordered by importance in decreasing order from the left to right.

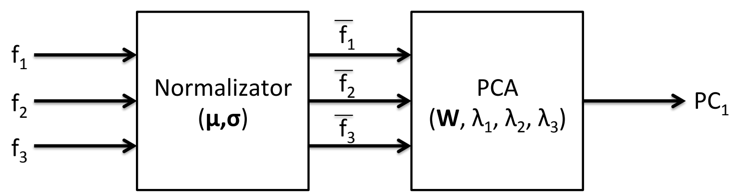


Figure 1: Pre-processing system

(a) What is the main relationship between eigenvectors? How would you prove this property?

(b) What is the percentage of variance preserved by the system?

(c) Consider that the feature vector $\mathbf{f} = \begin{bmatrix} 146.000 \\ 238.000 \\ 8.630 \end{bmatrix}$ is given at the system input. What will be the system output?

Name: **Number:** **Practical Class:**

Your answer to 1):

a) Eigenvectors are orthogonal. Just: do the dot product and it should be 0.

b) 99.5%

c)

After normalization:

$$\bar{f} = \begin{bmatrix} -0.8842 \\ -1.31 \\ -0.022 \end{bmatrix}$$

$$PC_1 = \bar{f}^T \cdot w_1 =$$

$$\underset{1\times 3}{} \qquad \underset{3\times 1}{}$$

$$= \begin{bmatrix} -0.8842 & -1.31 & -0.022 \end{bmatrix} \begin{bmatrix} -0.594 \\ -0.580 \\ -0.558 \end{bmatrix}$$

$$= 1,3$$

Name: **Number:** **Practical Class:**

Question 2 - Statistical Learning

☐ **20pts**

(a) Write the Bayes Classification Rule in the simplified form of the likelihood ratio $\Lambda(\mathbf{x})$ and explain its meaning.

(b) Suppose the following likelihoods for a two-class classification problem:

$$\begin{aligned} & \quad\quad\quad\; \mu \;\; \sigma^2 \\ p(\mathbf{x}|\omega_1) &= \mathcal{N}(0,1) \qquad \forall x \\ p(\mathbf{x}|\omega_2) &= \frac{1}{4} \qquad -2 < x < 2 \end{aligned}$$

Assume igual priors ($P(\omega_1) = P(\omega_2) = 0.5$.
Give the minimum error classification rule for this binary problem using formula in (a).
(you may want to draw a graph plot to help in your answer).

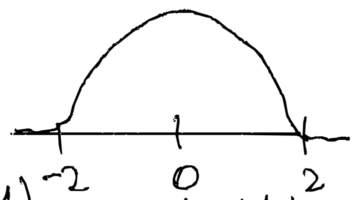
Your answer to 2):

a) $\Lambda(x) = \frac{P(\omega_2)}{P(\omega_1)} < \frac{P(x|\omega_1)}{P(x|\omega_2)}$ If true then ω_1 else ω_2

Compares the ratio of likelihoods against the ratio of priors. If the evidence for ω_1 outweighs the priors against it we decide ω_1

b) $\frac{P(x|\omega_1)}{P(x|\omega_2)} > 1 \Rightarrow P(x|\omega_1) > \frac{1}{4} \Rightarrow \omega_1$ else ω_2

$$\frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} > \frac{1}{4}$$



$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} > \frac{1}{4} \Rightarrow \frac{1}{\sqrt{2\pi}} \cdot \left(-\frac{1}{2}\right)x^2 > \ln\left(\frac{1}{4}\right)$$

$\Rightarrow$ solve to x

Name: Number: Practical Class:

Question 3 - SVM - Radial Basis Function Model

☐ 20pts

Consider the boundary decisions found by a Radial Basis Function SVM in Figure :

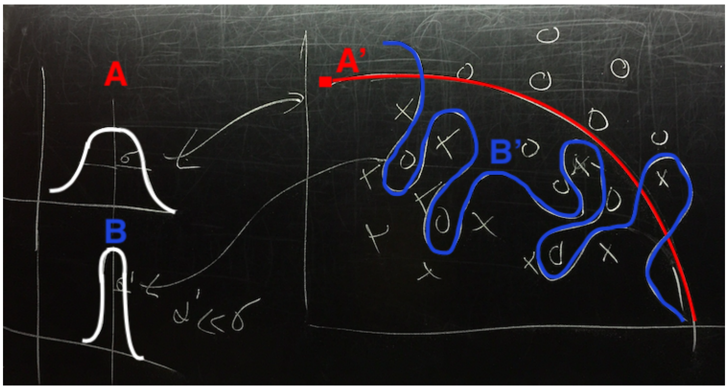


Figure 2: SVM Boundary Decision

(a) Use above SVM Radial Basis Function formula to give the rationale of the relation between A and B and the Boundary Decisions A' and B'.
(Please note that you should give a plenty justification including role of parameter σ; it is not enough to say that A corresponds to boundary decision A' and B corresponds to boundary decision B'.)

(b) According to above reasoning give your intuition which model would you choose A or B?

Your answer to 3):

A → A' because σ is large, the function yields localized decision boundaries hence overfitting. B → B' for the opposite reason. I would probably choose A because although underfitting it can still generalize better but depends on the problem.

Name: **Number:** **Practical Class:**

Question 4 - Classification Theory

☐ **10pts**

In the classes we learned that the total error rate of a classifier can be decomposed into two parts: (i)] Error rate due to the average performance of the classifier and (ii) Error rate due to the variation of training set.

(a) Identify each partial error above by their common names in pattern recognition

(b) In practical class you worked with cork stoppers dataset problem. Take the first 100 samples with classes ω_1 and ω_2. Suppose you have two linear classifiers, say, a linear Fisher Discriminant and a minimum distance classifier.

(b1.) Which of the above errors could you find?

(b2.) What would you do to determine the other error?

Your answer to 4):

a) Bias and variance respectively.

b) b1) Probably bias, they are simple models

b2) We could create diferent sets of samples and test the performance across them.