

Exercício 2

A.P. Braga

April 1, 2017

Classificador Bayesiano - Gaussiana no \mathbb{R}^n

Nesta atividade os alunos irão utilizar o classificador bayesiano, com o modelo gaussiano multivariáveis, como uma ferramenta para o treinamento e classificação de amostras.

Para tal, será utilizado o conjunto de dados conhecido como Iris. Para carregar a base de dados Iris no R basta digitar `data(iris)`. Utilizando-se dessa base e da função de densidade de probabilidade normal fornecida, pede-se:

1. Dividir a base de dados iris em duas classes, uma correspondendo à espécie "setosa" (amostras 1 a 50) e a outra correspondendo às espécies "versicolor" e "virginica" (demais amostras).
2. Dividir cada uma das classes em dois conjuntos, o primeiro com 70% dos dados que será o conjunto de treinamento e o segundo com 30% dos dados que será usado para teste.
3. Utilizando o conjunto de treinamento, encontrar a média e a matriz de covariância para cada classe.
4. Utilizando-se da função de densidade de probabilidade normal (*normal probability density function (pdf)*) calcular a densidade de probabilidade de cada amostra do conjunto de treinamento para cada uma das classes.
5. Ainda para o conjunto de treinamento, estimar qual a classe de cada uma das amostras baseado nos valores de densidade de probabilidade encontrados no item anterior. Apresente a matriz de confusão.

Obs: Imprimir apenas a matriz de confusão. Use a função `table(y,ŷ)`.

6. Para o conjunto de teste, calcular a densidade de probabilidade de cada amostra para cada uma das classes, estimar a qual classe cada uma das amostras pertence baseado nos valores de densidade de probabilidade encontrados e apresentar a matriz de confusão.
7. Faça um loop para repetir o processo do item 1 ao 5 e calcule a média e desvio padrão dos acertos do grupo de teste após 30 repetições.

Código em R da função densidade de probabilidade normal para n variáveis:

```
pdfnvar<-function(x,m,k,n){(1/(sqrt((2*pi)^n*(det(k))))) *
exp(-0.5*(t(x-m)%*% (solve(k))%*(x-m)))}
```

Forma de Entrega: Relatório em .doc ou .pdf, descrevendo o que foi feito, mostrando as informações pedidas e explicando os resultados obtidos. O relatório deve ser colocado em um arquivo .zip junto com os códigos utilizados e enviado via Moodle.