

Classificador de Bayes - Exercício 07

Matheus Bitarães de Novaes

15 de Agosto de 2021

1 Introdução

Neste exercício, deve-se implementar um Classificador Bayesiano para separar duas classes de uma base de dados "duas espirais". Cada classe será modelada como uma mistura de gaussianas de duas dimensões, utilizando o algoritmo *K-means*.

2 Geração dos dados

Os dados das duas espirais foram gerados pela função `mlbench.spirals` do pacote `mlbench`, em R. São 3000 amostras de duas espirais com desvio padrão de 0.05, conforme figura 1.

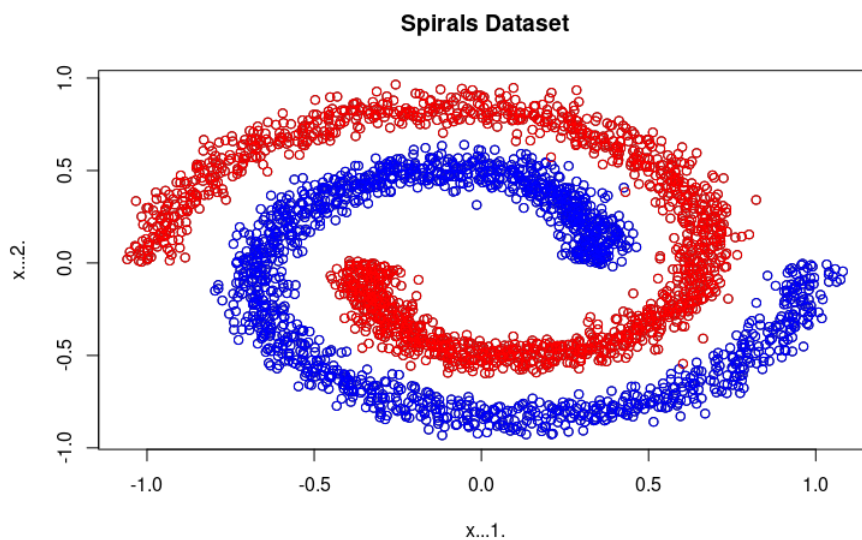


Figure 1: Amostras geradas para o problema

3 Clusterização

A clusterização foi feita através do algoritmo *K-Means*. Foi utilizada a função *kmeans*, nativa da linguagem R. A clusterização resultante pode ser vista na figura 2, onde pode-se ver os centros dos 30 *clusters* encontrados.

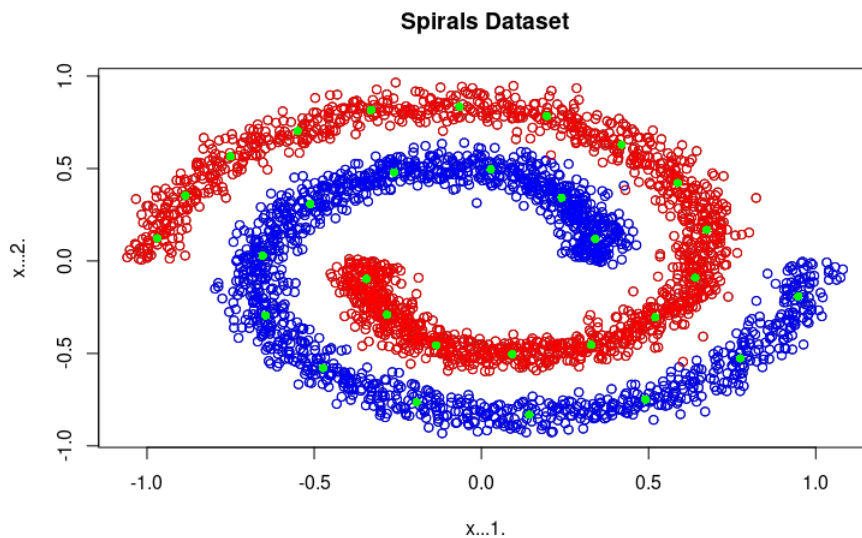


Figure 2: 30 *clusters* gerados através do algoritmo *K-means*

4 Geração das Gaussianas

Foi gerada uma gaussiana para cada cluster, onde o centro da gaussiana é o centro do cluster. Na figura 3 pode-se ver a região gerada pela soma das gaussianas, onde as gaussianas positivas são de uma classe e as gaussianas negativas da outra classe.

A partir desta região de gaussianas, pode-se avaliar qualquer ponto no espaço e checar seu sinal para classificá-lo em uma das duas classes do problema.

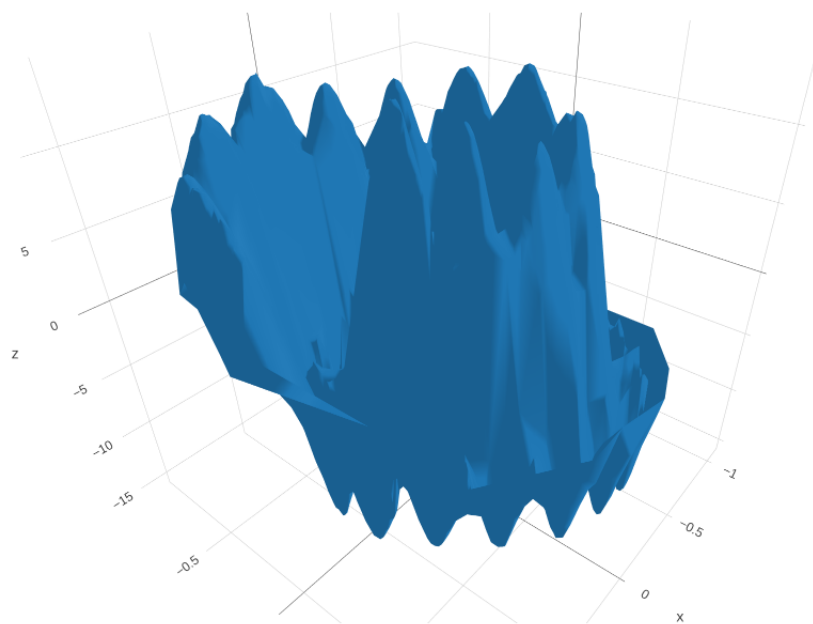


Figure 3: Região composta pela soma das gaussianas