



No exercício proposto, foram amostrados 400 dados ao todo, pertencentes a quatro classes perfeitamente balanceadas em relação a quantidade de amostras diferentes. As classes foram obtidas através da geração de dados aleatórios em torno dos pontos $(2, 2, \sigma = 0.6)$, $(4, 4, \sigma = 0.8)$, $(2, 4, \sigma = 0.2)$ e $(4, 2, \sigma = 1)$, pertencentes às classes **vermelha**, **verde**, **amarela** e **azul**, respectivamente. A figura 1 ilustra os dados amostrados no problema.

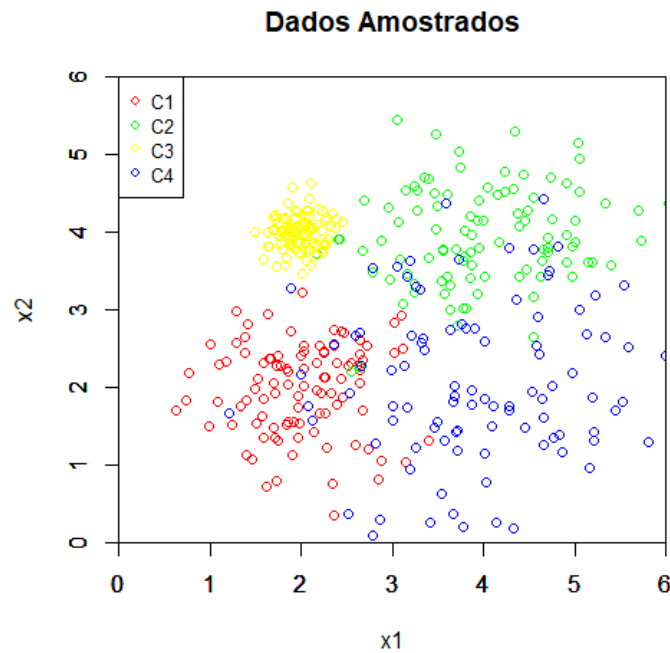


Figura 1: Dados amostrados para o exercício 1.

Utilizando um a biblioteca *plot3D* e *contour2D* do *R*, foram geradas as superfícies que mostram a densidade de probabilidade de cada função, mostradas nas figuras 2 e 3.

Por fim, foi gerado um grid para que a função classificasse amostras distribuídas uniformemente em um intervalo de $[0, 6]$ e, utilizando a matriz gerada com as classificações, foi plotado, utilizando a função *image2D*, a superfície de separação das variáveis, mostrada na figura 4.

Na figura 4 podemos perceber os pontos que seriam classificados incorretamente caso utilizássemos esta superfície de separação de classes como base para o classificador. É possível ver a predominância das classes de maior desvio padrão em áreas afastadas da média das 4 classes amostradas, como em torno do ponto $(0, 0)$ que encontra-se mais próximo visualmente da classe vermelha ou o ponto $(0, 6)$, que visualmente classificaria-se como de classe amarela, porém ambos foram classificados como azul. Isto mostra a alta sensibilidade do Classificador Bayesiano à variância dos dados amostrados.

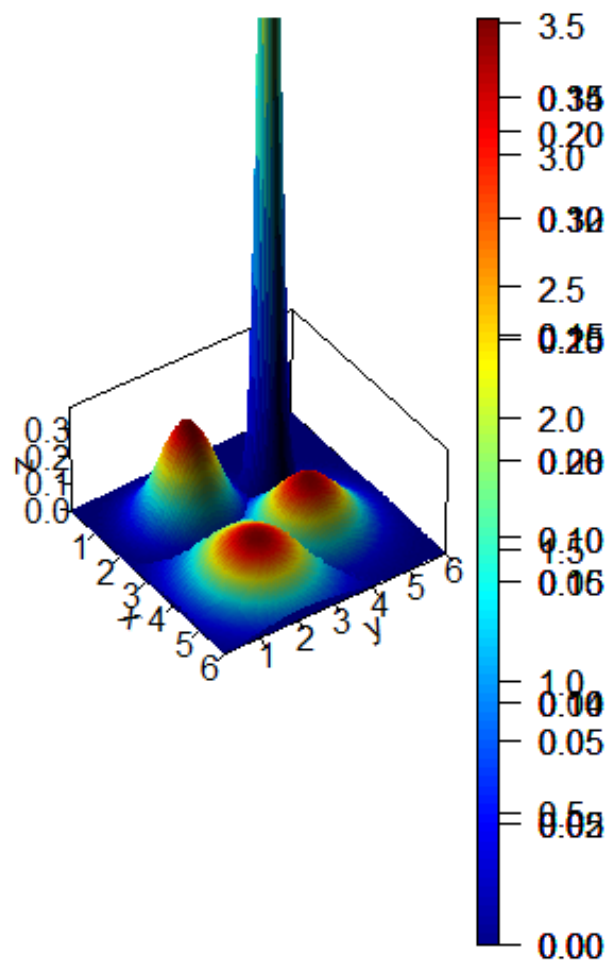


Figura 2: Curva de densidade de probabilidade vista em 3 dimensões.

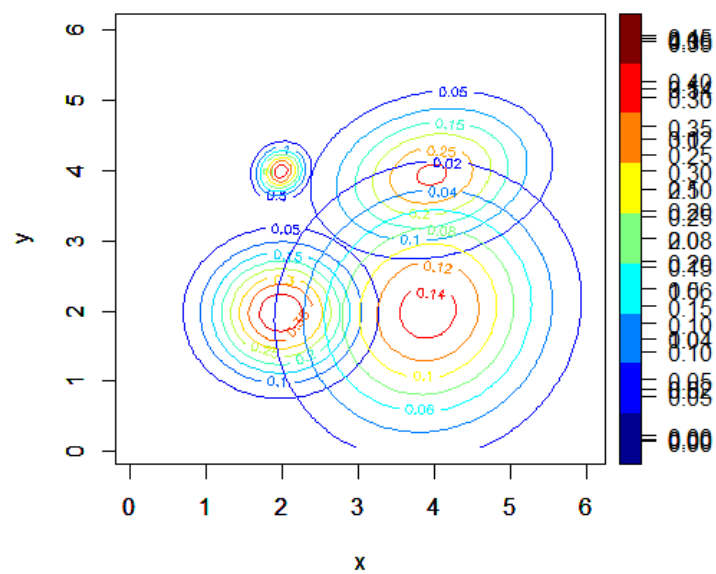


Figura 3: Curva de densidade de probabilidade em 2 dimensões.

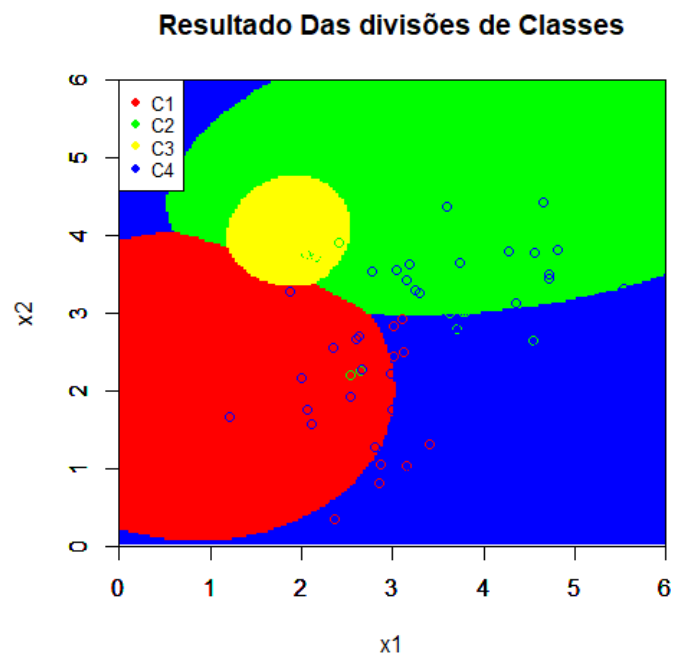


Figura 4: Resultado das superfícies de separação das classes.