

## Lista 3

CPE723 Otimização Natural

Olavo Sampaio

### Questão 1

Letra a)

Ver Figura 1, em anexo.

Letra b)

Para  $T = 1$ ,

$$P_{y|x} = \begin{bmatrix} 0.928 & 0.345 & 0.096 & 0.010 \\ 0.072 & 0.655 & 0.904 & 0.990 \end{bmatrix}$$

Letra c)

$$D = 12.036$$

Letra d)

$$y_{novo} = \begin{bmatrix} 1.482 \\ 6.470 \end{bmatrix}$$

Letra e)

Para  $T = 0.1$ ,

$$P_{y|x} = \begin{bmatrix} 1.0 & 0.002 & 0.0 & 0.0 \\ 0.0 & 0.998 & 1.0 & 1.0 \end{bmatrix}$$

$$D = 11.870$$

$$y_{novo} = \begin{bmatrix} 0.007 \\ 6.335 \end{bmatrix}$$

Letra f)

Para  $T = 50$ ,

$$P_{y|x} = \begin{bmatrix} 0.513 & 0.497 & 0.489 & 0.477 \\ 0.487 & 0.503 & 0.511 & 0.523 \end{bmatrix}$$

$$D = 13.088$$

$$y_{novo} = \begin{bmatrix} 4.663 \\ 4.834 \end{bmatrix}$$

#### Letra g)

Quanto menor a temperatura, mais determinístico será o pertencimento de um determinado elemento à uma dada classe, ou seja,  $P_{y|x}$  serão menos uniformes para cada  $x$ . Nesse caso, o sistema atinge valores da função custo mais próximos do mínimo global. Da mesma forma, para temperaturas mais altas, o algoritmo não tem muita informação e as probabilidades condicionais tendem a ser mais uniformes, o que leva a centroides próximos e custo distante do mínimo global.

Para  $T = 50$ , as distâncias entre cada ponto e os centroides tem menos influência nas probabilidades condicionais, que tendem a ser mais próximas e  $P_{y|x}$ , mais uniforme para cada  $x$ . Portanto, os centroides são mal posicionados e o custo tem valor mais elevado.

Para  $T = 1$ , as menores distâncias entre pontos e centroides terão maior influência sobre as probabilidades condicionais, o que leva o algoritmo a encontrar centroides e mínimo melhores.

Para  $T = 0.1$ , a temperatura é baixa o bastante para que as probabilidades sejam altamente influenciadas pelas distâncias e que cada  $x$ , quase garantidamente, pertença uma determinada classe, aproximando o problema do *hard clustering*. Isso permite que a solução atinja um mínimo menor que as anteriores, o menor custo do conjunto. O problema de usar temperaturas muito baixas é que o algoritmo pode não ficar preso em um mínimo local, sendo necessário elevar a temperatura para que ele escape.

#### Questão 2

O problema escolhido foi a clusterização de um conjunto de dados bidimensionais com 10 clusters. Os dados foram gerados a partir de distribuições normais com médias pré-determinadas e  $\sigma = 1$ . As médias foram escolhidas com uma distribuição uniforme sobre o intervalo  $(0, 20)$ , com semente do gerador aleatório fixa. A distribuição dos dados e centroides está na Figura 2.

Foi executado o algoritmo de DA, para  $T = 1$ , com critério de parada

$$custo\_novo - custo\_atual < \varepsilon,$$

com  $\varepsilon = 0.01$ .

No entanto, o algoritmo não convergiu. De fato, a função custo é sempre a mesma e os centroides não se aproximam de nenhum resultado razoável. Isso certamente se deve a um erro de implementação do DA.

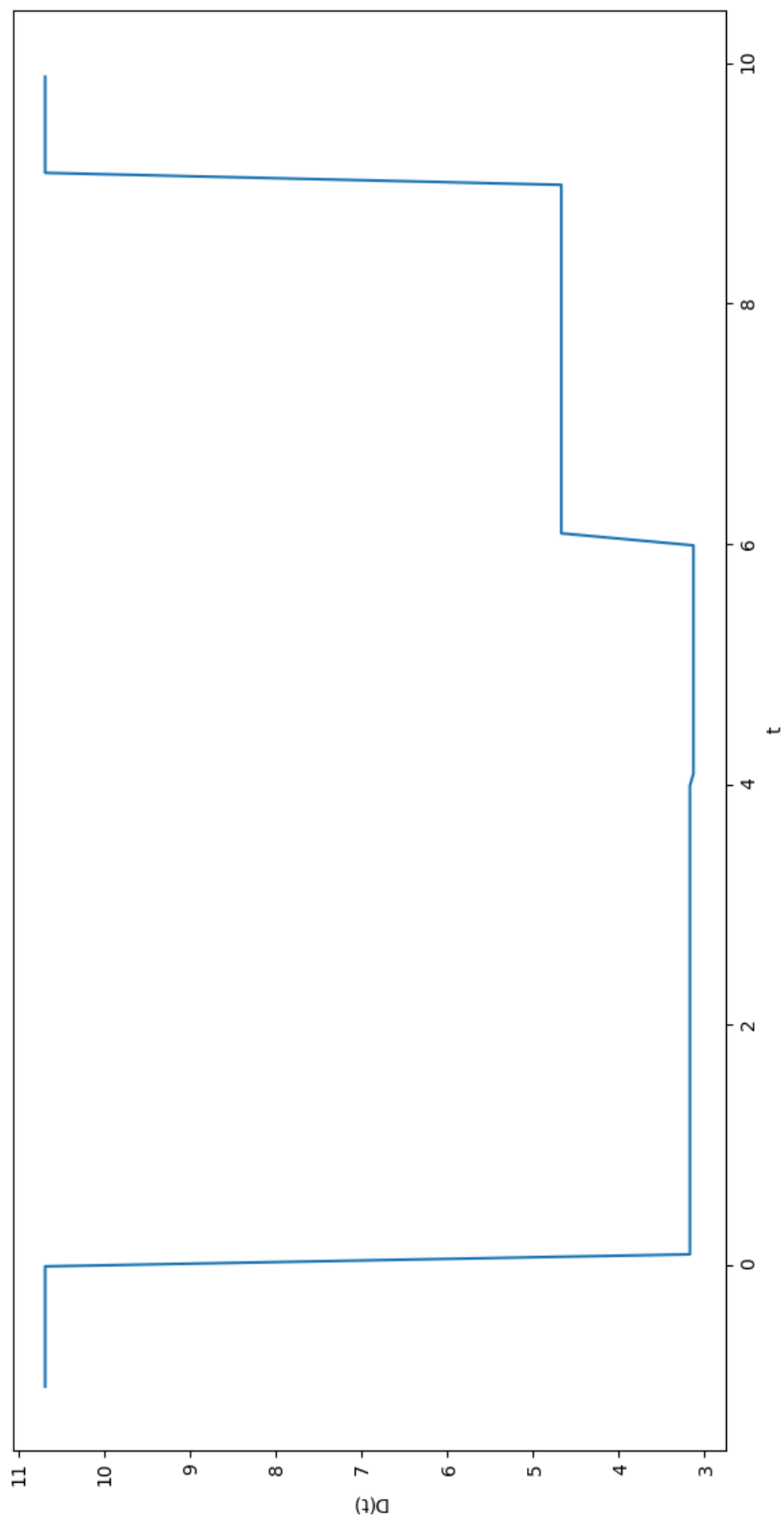


Figura 1: Questão 1ª.  $D(t)$  ao longo de -1, 10

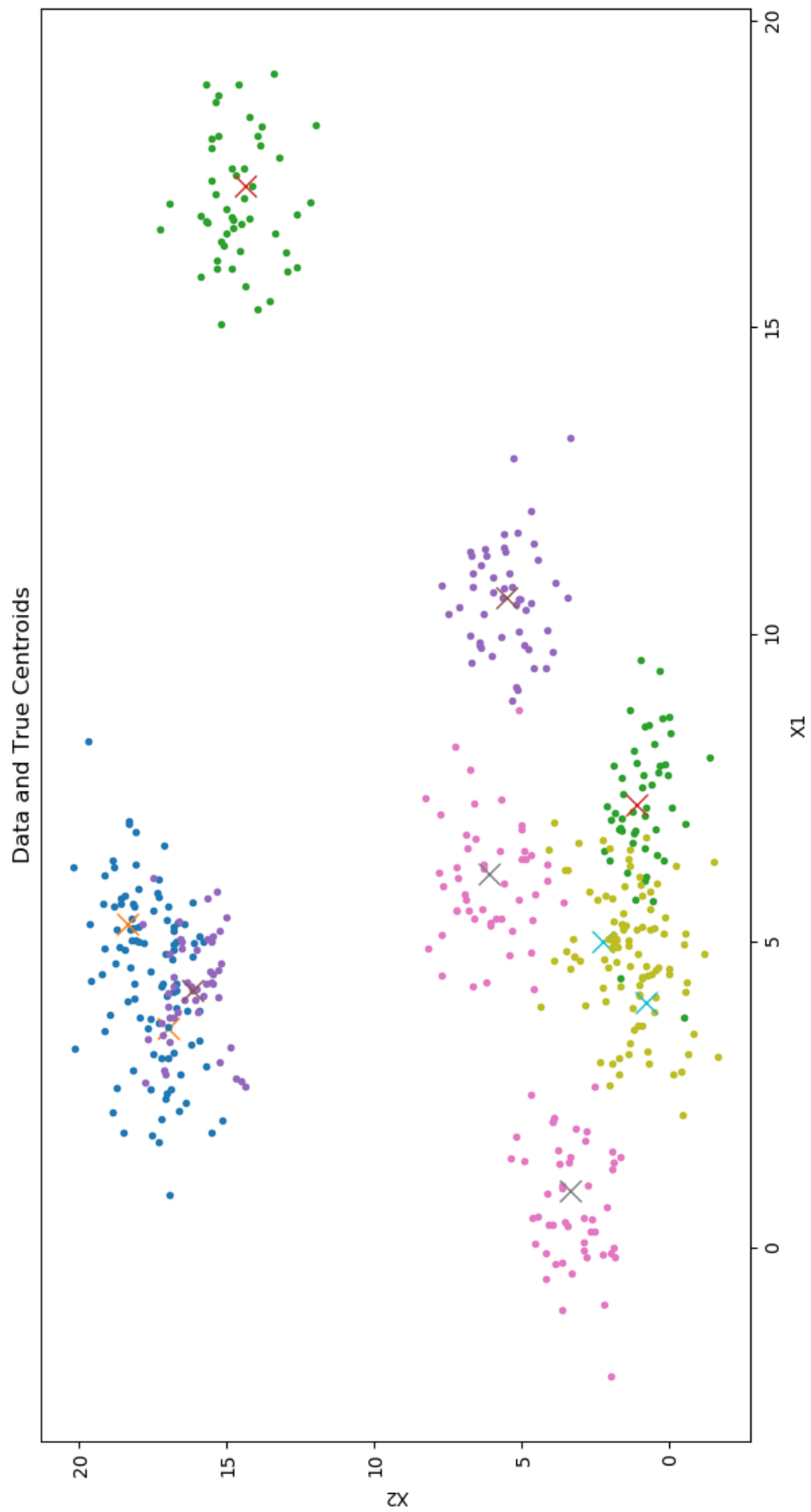


Figura 2: Questão 2. Dados e centroides verdadeiros