



---

---

# Reconhecimento

---

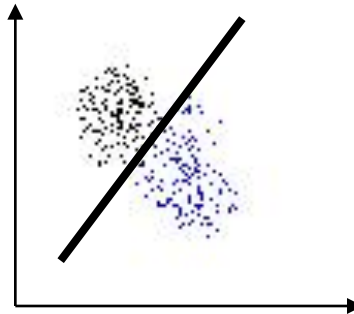
---



# Reconhecimento

## ■ Funções de decisão

- ▶ Determinar uma função que fundamente a decisão de atribuir um padrão a uma das classes do problema.



- Classe 1
- Classe 2

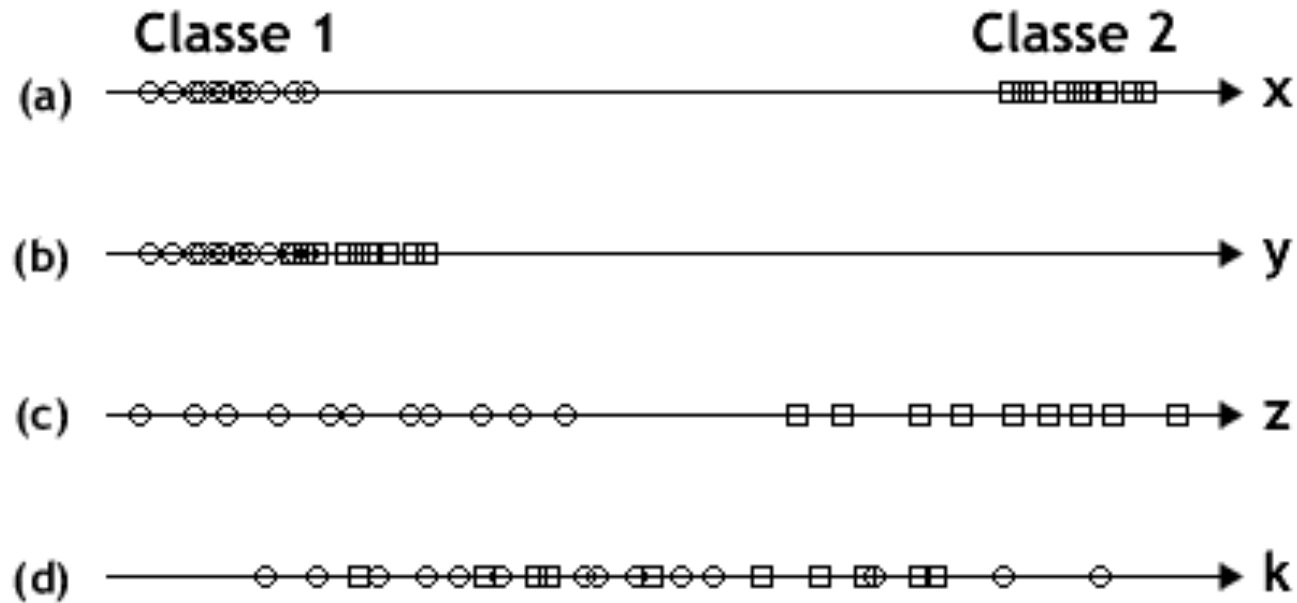
# Reconhecimento

## ■ Questões

- ▶ As características permitem boa separabilidade das classes?
- ▶ As classes são linearmente separáveis?
- ▶ A distribuição das classes é Normal?
- ▶ As distribuições têm covariâncias semelhantes?
- ▶ O tamanho da amostra é maior que o número de variáveis?
- ▶ As matrizes de covariância são singulares?

# Reconhecimento

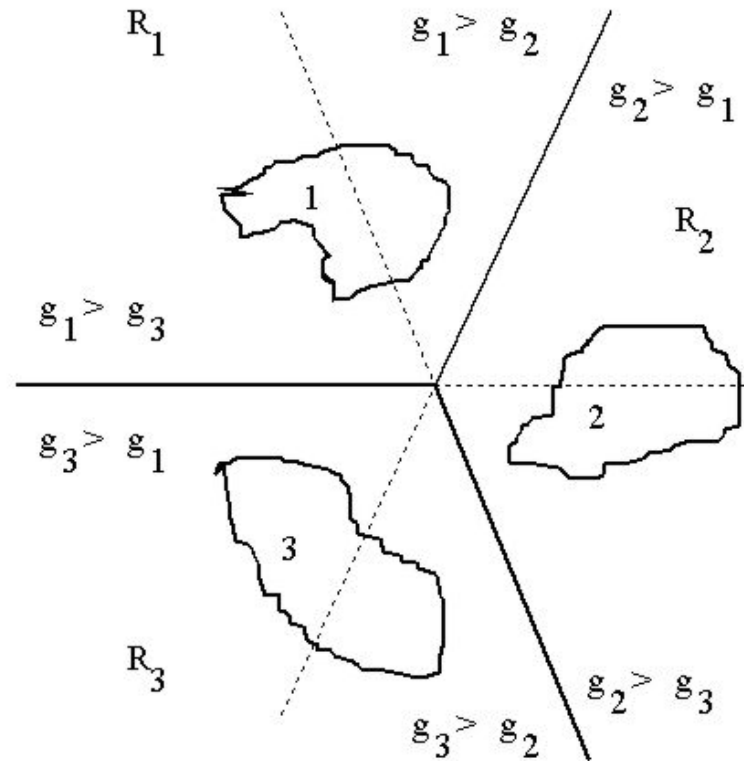
## Relação entre os descritores



característica

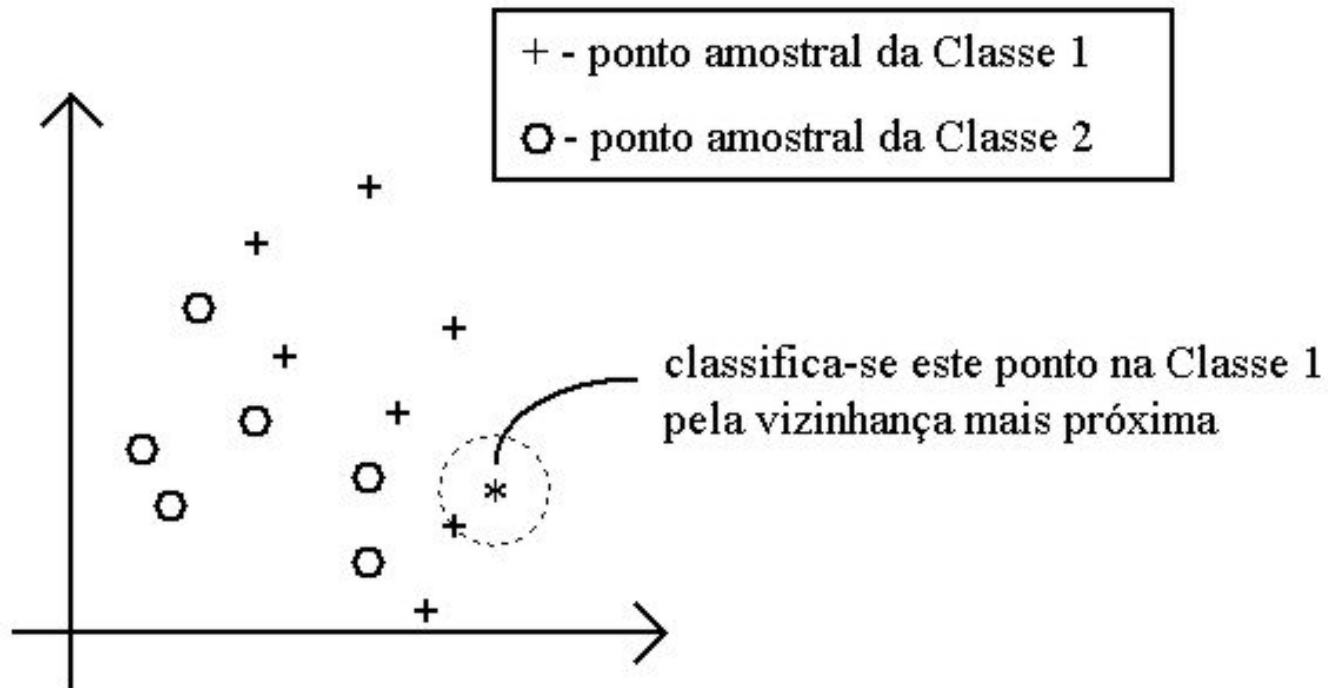
Ideal: Médias afastadas com desvio padrão pequeno.

## Funções discriminantes



# Reconhecimento

## Vizinho mais próximo



# Reconhecimento

## Classificação Supervisionada

Etapas:

- Escolher um conjunto de treinamento. Cada elemento do conjunto tem sua classe determinada.
- Escolher características discriminantes.
- Escolher um método/função de decisão.
- Determinar os parâmetros a partir do conjunto de treino.
- Testar com objetos fora do conjunto de treinamento.

# Reconhecimento

## Classificação Não-Supervisionada

Etapas:

- Escolher um conjunto de treinamento. As classes dos elementos do conjunto não são conhecidas.
- Escolher características discriminantes e o número de classes.
- Escolher um método para o agrupamento de elementos.
- Determinar os parâmetros a partir das classes encontradas.
- Testar com objetos fora do conjunto de treinamento.



---

# Reconhecimento

## K-Means

1. Escolher  $K$  pontos no espaço de características
2. Agrupar os objetos mais próximos a cada ponto.
3. Mover os  $K$  pontos para a posição média dos grupos correspondentes
4. Repetir os passos 2 e 3 até que as médias não variem mais

Veja: <http://www.kovan.ceng.metu.edu.tr/~maya/kmeans/index.html>

# Reconhecimento

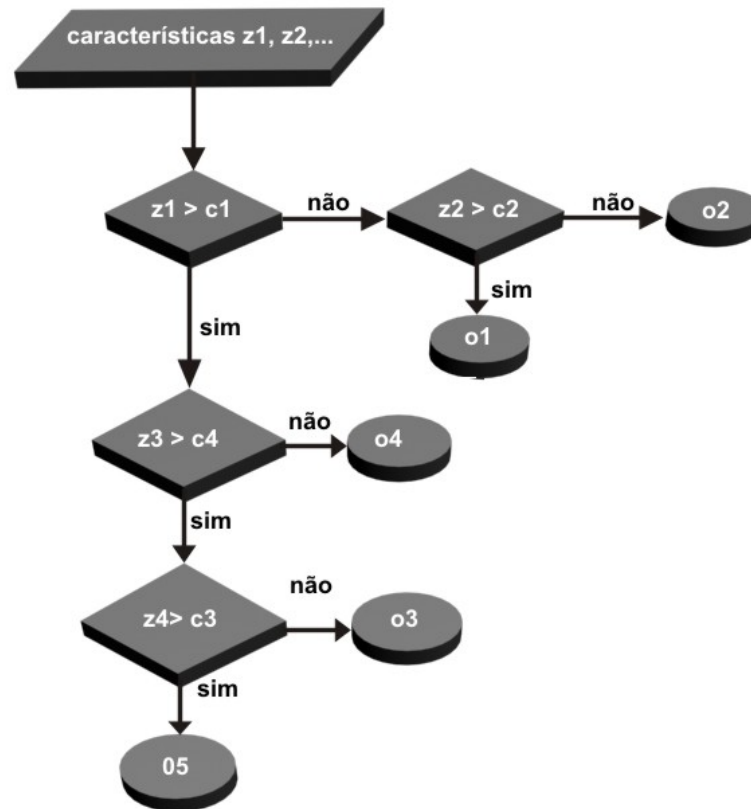
## ■ Avaliação de resultados

- ▶ Falso positivo: Sistema que detecta quadrados diz que uma imagem de uma bola é um quadrado.
- ▶ Falso negativo: Sistema que detecta quadrados diz que uma imagem de um quadrado não é um quadrado.
- ▶ Matriz de confusão

Obj \ Classe	Círculo	Quadrado	Triângulo
Círculo	10	2	1
Quadrado	5	20	2
Triângulo	0	0	10

# Reconhecimento

## Árvores de decisão



# Reconhecimento

## Distância Euclideana

Dados dois vetores P e Q

$$P = (p_1, p_2, \dots, p_n)$$

$$Q = (q_1, q_2, \dots, q_n)$$

a distância euclidiana entre eles é de

$$\sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} = \sqrt{\sum_{i=1}^n (p_i - q_i)^2}.$$

## Distância Euclideana

Função de decisão:

O padrão  $\mathbf{x}$  pertence a classe  $c_i$  com média  $\mathbf{m}_i$  se

$$\sqrt{(\mathbf{x} - \mathbf{m}_i)^T (\mathbf{x} - \mathbf{m}_i)} < \sqrt{(\mathbf{x} - \mathbf{m}_j)^T (\mathbf{x} - \mathbf{m}_j)}, \forall i \neq j$$

# Reconhecimento

Exemplo: suponha um sistema que tenha resultado na seguinte tabela.

Classe azul	obj	Caract.A	Caract B	Caract.C	Caract.D	Caract.E
	1	4	10	8	50	3
	2	5	12	7	55	1
	3	4	9	8	48	2
Classe amarela	4	6	10	9	54	3
	5	5	20	7	21	35
	6	3	21	8	19	30
	7	4	22	9	25	30
	8	5	21	6	15	25
	9	4	20	8	20	30

# Reconhecimento

Baseado nesses valores, trace e reflita:

- a) Gráfico de característica A x C
- b) Gráfico de característica B x D
- c) Gráfico de característica B x E
- d) Verifique em que classe um objeto, detectado por esse sistema, está, sendo seus valores: (7,12,7, 60, 3)

*Obs. Faça uma nova coluna com os resultados de todas as comparações para melhor observar.*

- e) Alguma das características, sozinha, poderia reconhecer um objeto como classe azul ou classe amarela?

## Modelo Vetorial

Cálculo do ângulo entre 2 vetores:

$$\cos(\mathbf{u}, \mathbf{v}) = \frac{\mathbf{u} \bullet \mathbf{v}}{|\mathbf{u}| |\mathbf{v}|}$$

Função de decisão: O padrão  $\mathbf{x}$  pertence a classe  $c_i$  com média  $\mathbf{m}_i$  se

$$\cos(\mathbf{x}, \mathbf{m}_i) > \cos(\mathbf{x}, \mathbf{m}_j), \forall i \neq j$$



# Reconhecimento

## Distância de Mahalanobis

$$D_i(\mathbf{x}) = (\mathbf{x} - \mathbf{m}_i)^T \mathbf{C}_i^{-1} (\mathbf{x} - \mathbf{m}_i)$$

$$C_i[j, k] = \sum_{s=1}^N \frac{(\mathbf{x}_s[j] - \mathbf{m}_i[j])(\mathbf{x}_s[k] - \mathbf{m}_i[k])}{N - 1}$$

► Amostra com  $N$  objetos

Função de decisão: O padrão  $\mathbf{x}$  pertence a classe  $c_i$  com média  $\mathbf{m}_i$  e matrix de covariância  $\mathbf{C}_i$  se

$$D_i(\mathbf{x}) < D_j(\mathbf{x}), \forall i \neq j$$

# Distância de Mahalanobis

Distingue-se da distância euclidiana já que leva em conta as correlações entre as características e diferenças de unidades

Ex: classificar um peixe baseado no seu comprimento e altura

- Vetor  $(x_1, x_2)$ , sendo  $x_1$  comprimento e  $x_2$  altura
- Comprimento varia entre 50 e 100cm, altura entre 10 e 20cm

Caso 1: Calcular a distância euclidiana entre o peixe e a média de cada classe

- Ruim, pois dá a mesma importância para o comprimento e para altura – mas o comprimento varia muito mais que a altura

$$d_e(\vec{x}_1 \vec{x}_2) = \sqrt{(x_{11} - x_{12})^2 + (x_{21} - x_{22})^2}$$

# Distância de Mahalanobis

Caso 2: Incluir a media de dispersão de cada variável (comprimento e altura), e a variância para cada característica

$$d_2(\vec{x}_1 \vec{x}_2) = \sqrt{\left(\frac{(x_{11} - x_{12})}{\sigma_1}\right)^2 + \left(\frac{(x_{21} - x_{22})}{\sigma_2}\right)^2}$$

- Ruim, pois existe uma correlação entre altura e comprimento: um salmão mais comprido sofre também uma mudança na sua altura

Caso 3: Incluir a covariância entre  $x_1$  e  $x_2$ : covariância

$$d_m(\vec{x}_1 \vec{x}_2) = \sqrt{(\vec{x}_1 - \vec{x}_2)^T \Sigma^{-1} (\vec{x}_1 - \vec{x}_2)}$$

# Distância de Mahalanobis

Exemplo: suponha que há dois grupos de dados, cada um dos quais consiste em duas variáveis ( $x$ ,  $y$ ). Verifique a que classe pertence o vetor (6,4)

Amostras	x1	y1	x2	y2
1	5	3	9	1
2	1	5	7	3
3	4	6	8	2
4	5	7	7	1
5	2	5	8	1
6	3	5	7	3
7	4	4	9	2
8	5	5	8	1
9	3	4	7	2
10	1	3	8	1

---

# Distância de Mahalanobis

- Passos:
  - Calcular a média aritmética de cada grupo
  - Centrar os dados na média aritmética de cada grupo
  - Calcular as matrizes de covariância dos grupos
  - Calcular as matrizes inversas:  $A.A^{-1} = I$
  - Calcular as distâncias

# Reconhecimento

## Classificador de Bayes

$$d_i(\mathbf{x}) = \ln P(c_i) - \frac{1}{2} \ln |\mathbf{C}_i| - \frac{1}{2} (\mathbf{x} - \mathbf{m}_i)^T \mathbf{C}_i^{-1} (\mathbf{x} - \mathbf{m}_i)$$

- Inclui a probabilidade a priori de cada classe.

# Reconhecimento

## Redes Neurais

