

# Machine Learning

Aula 09

Aprendizagem Supervisionada: Máquinas de Vetor de Suporte

Evandro J.R. Silva

Uninassau Teresina



# Sumário

1 Introdução

2 Hiperplano

3 Margem máxima

4 Vetores de suporte

5 Kernel trick

6 FIM



# 1 Introdução

## 2 Hiperplano

## 3 Margem máxima

## 4 Vetores de suporte

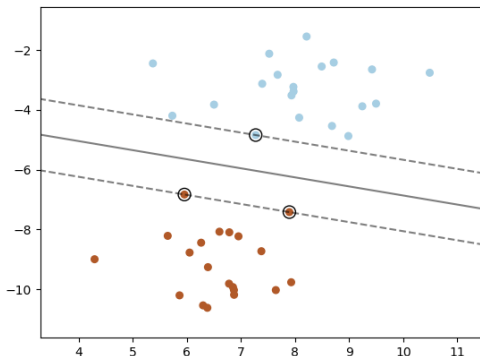
## 5 Kernel trick

## 6 FIM



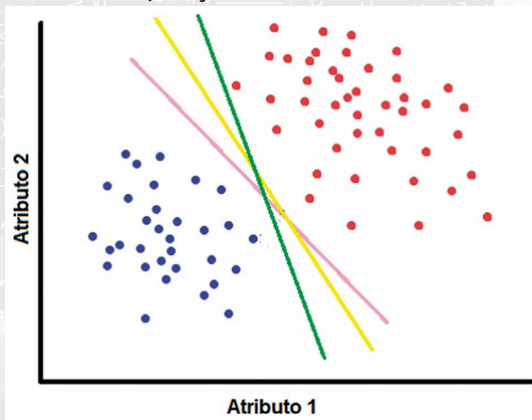
# Introdução

- **SVM (Support Vector Machine)**
  - É um algoritmo de classificação (pode ser usado também para regressão) que procura identificar o hiperplano que melhor separa duas classes.



# Introdução

- Na figura a seguir temos a borda/fronteira de decisão de 3 classificadores. Aparentemente os 3 estão classificando corretamente. Mas qual fez a melhor separação?



# Introdução

- Para conseguirmos entender o máximo possível, vamos seguir o seguinte roteiro:

- 1 O que é um **hiperplano**?
- 2 Hiperplano de **margem máxima** e o que é uma margem no contexto de classificação.
- 3 Vetores de suporte.
- 4 Kernel trick.

## 1 Introdução

## 2 Hiperplano

## 3 Margem máxima

## 4 Vetores de suporte

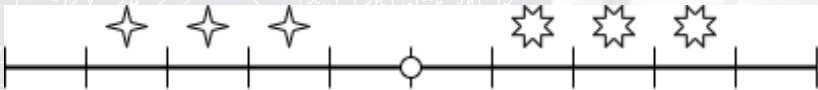
## 5 Kernel trick

## 6 FIM



# Hiperplano

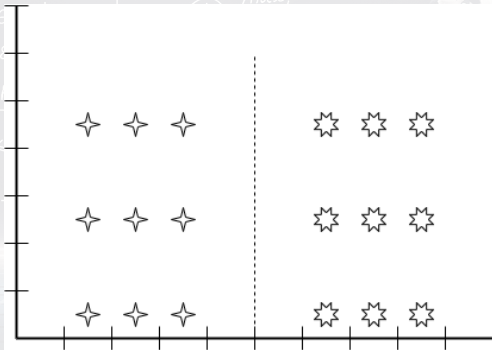
- Um **hiperplano** é um plano  $n$ -dimensional, onde  $n$  é indefinido.
- Para entendermos melhor, vamos ver como ele se apresenta em 1, 2 e 3 dimensões.
- Com  $n = 1$ , o hiperplano será apenas um ponto. Na imagem a seguir, o hiperplano é o ponto central.





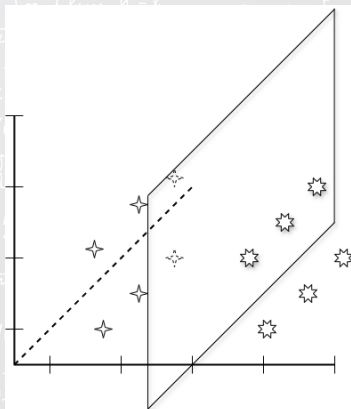
# Hiperplano

- Com  $n = 2$  o hiperplano passa a ser uma linha, ou seja, consiste nos infinitos pontos que compõem a linha.
- Na figura a seguir, o hiperplano é mostrado como uma linha tracejada.



# Hiperplano

- Com  $n = 3$  o hiperplano passa a ser um plano.
- Na figura a seguir, o hiperplano é mostrado como um plano.
- Com  $n \geq 4$  passa a ser difícil visualizar. Porém, em bases de dados reais é comum que o **espaço amostral** tenha várias dimensões. O SVM vai encontrar o melhor plano que passe entre as duas classes.



## 1 Introdução

## 2 Hiperplano

## 3 Margem máxima

## 4 Vetores de suporte

## 5 Kernel trick

## 6 FIM

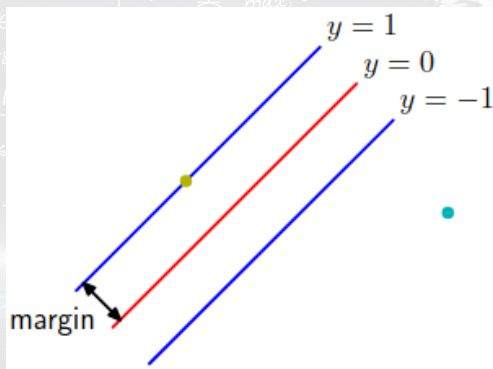


# Margem máxima

- No contexto de classificação, uma **margem** (ou **borda**) é a região do **espaço amostral** onde as instâncias de duas classes estão próximas umas das outras. Essa região também é chamada de **fronteira**.

# Margem máxima

- No caso específico do SVM, uma margem é definida como a distância perpendicular entre a fronteira de decisão e a instância mais próxima.



# Margem máxima

- O problema que o SVM tenta resolver abrange a maximização do tamanho dessa margem.
- Margens maiores possibilitam uma melhor generalização.

# Margem máxima

- O problema que o SVM tenta resolver abrange a maximização do tamanho dessa margem.
- Margens maiores possibilitam uma melhor generalização.
- O objetivo então é encontrar as maiores margens possíveis na fronteira entre duas classes.
- Como encontrar?

# Margem máxima

- O problema que o SVM tenta resolver abrange a maximização do tamanho dessa margem.
- Margens maiores possibilitam uma melhor generalização.
- Como encontrar? Com o uso de **vetores de suporte**.



## 1 Introdução

## 2 Hiperplano

## 3 Margem máxima

## 4 Vetores de suporte

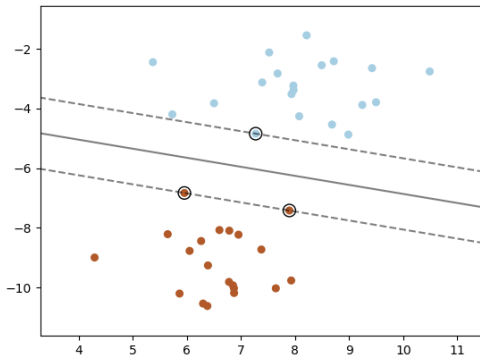
## 5 Kernel trick

## 6 FIM



# Vetores de suporte

- Os **vetores de suporte** são as instâncias das duas classes que são utilizadas para definir a margem de separação das classes.



1 Introdução

2 Hiperplano

3 Margem máxima

4 Vetores de suporte

5 Kernel trick

6 FIM



The background features a complex collage of mathematical concepts related to Support Vector Machines (SVM) and kernel methods. It includes:

- Equations and Limits:** Various mathematical expressions such as  $\lim_{n \rightarrow \infty} \frac{n^2 - x}{3}$ ,  $\lim_{n \rightarrow \infty} \sqrt[n]{A} = 1$ ,  $\lim_{n \rightarrow \infty} \sigma^n$ ,  $\lim_{n \rightarrow \infty} \frac{1}{n}$ , and  $\lim_{n \rightarrow \infty} \frac{1}{n^2}$ .
- Geometry:** A 3D diagram of a hyperplane in a coordinate system with axes  $x_1, x_2, x_3$  and a normal vector  $c_3 \cdot c_x$ .
- Calculus and Analysis:** Formulas for derivatives like  $\frac{d}{df} \left\{ \frac{x_n}{y_n} \right\}$  and  $\frac{d}{df} \left\{ \frac{x_n + y_n}{y_n} \right\}$ , and a graph showing a function  $f(x)$  with a margin  $\epsilon$ .
- Algebra and Set Theory:** Expressions like  $\{x_n\} \subset \mathbb{R}$ ,  $\{y_n\} \subset \mathbb{R}$ ,  $\{x_n\} \cdot \{y_n\} = \{x_n \cdot y_n\}$ , and  $\{x_n\} \cdot \{y_n\} = \{x_n + y_n\}$ .
- Optimization:** Formulas for maximizing the margin, such as  $\max_i \{x_n\} : x_n = \frac{1}{n}$  and  $\min \{x_n\} : x_n = \frac{1}{n}$ .

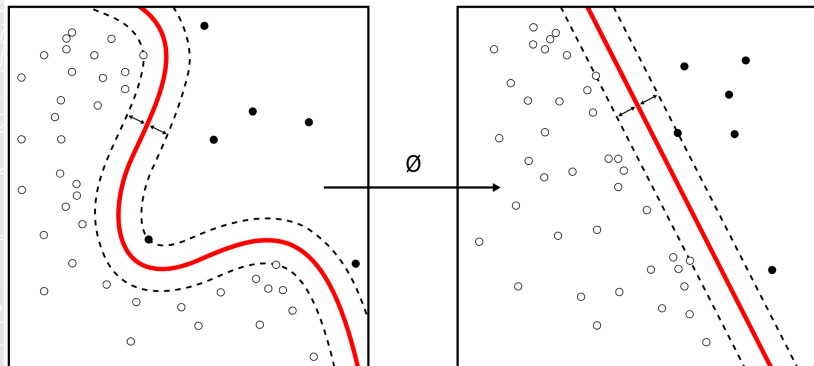
# Kernel trick

- O SVM é um **classificador linear**, ou seja, ele procura uma reta (a melhor, na verdade) que separa duas classes.
- Mas, e se as classes não forem linearmente separáveis?

# Kernel trick

- O SVM é um **classificador linear**, ou seja, ele procura uma reta (a melhor, na verdade) que separa duas classes.
- Mas, e se as classes não forem linearmente separáveis?
- Neste caso temos os *kernel tricks*
  - O **espaço amostral** é transformado em outro de maior dimensionalidade, de forma que as classes possam ser linearmente separáveis.

# Kernel trick



# Kernel trick

- Funções kernel:
  - **Linear** — a função padrão;
  - **Polinomial** — que funciona como uma regressão polinomial;
  - **RBF (Radial Basis Function)** — interessante para classes em formato circular.
  - **Sigmoide.**

# Kernel trick

- Funções kernel:
  - **Linear** — a função padrão;
  - **Polinomial** — que funciona como uma regressão polinomial;
  - **RBF (Radial Basis Function)** — interessante para classes em formato circular.
  - **Sigmoide.**
- E se a base de dados for multiclasse?



# Kernel trick

- Funções kernel:
  - **Linear** — a função padrão;
  - **Polinomial** — que funciona como uma regressão polinomial;
  - **RBF (Radial Basis Function)** — interessante para classes em formato circular.
  - **Sigmoide.**
- E se a base de dados for multiclasse?
  - SVM pode ser adaptado: OVO (*One vs. One*) ou OVR (*One vs. Rest*).
  - Ou seja, a aplicação do SVM em bases multiclasse nos leva aos **Comitês de Classificadores** [VEREMOS SOBRE ELES NA PRÓXIMA AULA].

## 1 Introdução

## 2 Hiperplano

## 3 Margem máxima

## 4 Vetores de suporte

## 5 Kernel trick

## 6 FIM

