Machine Learning and Pattern Recognition

东北大学"智能+X"新工科课程系列

机器学习与模式识别

东北大学 信息科学与工程学院
人工智能系、智能感知与机器人研究所
陈东岳

Mathematic Modelling of SVM problem

SVM问题的数学模型

CHAPTER ONE

问题原型

Prototype

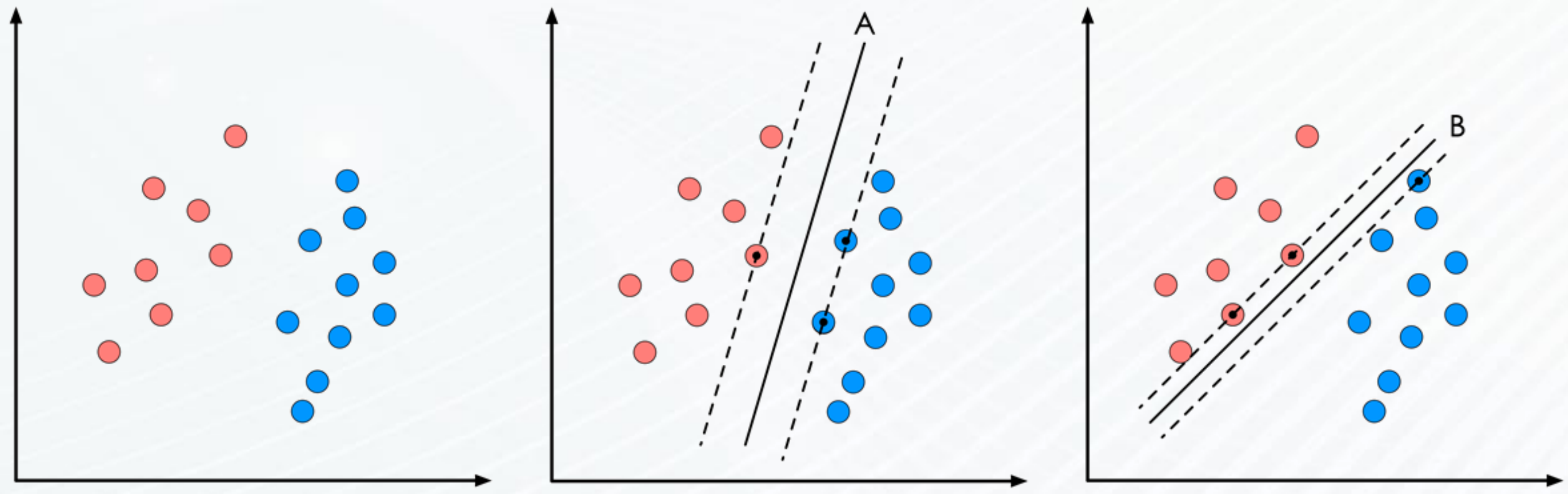
一、问题原型

什么是线性可分问题?

- 如果平面中至少存在一条线，且所有蓝点在该线的一侧，而所有红点在另一侧，则两组是线性可分离的。

线性分类器的性能比较：

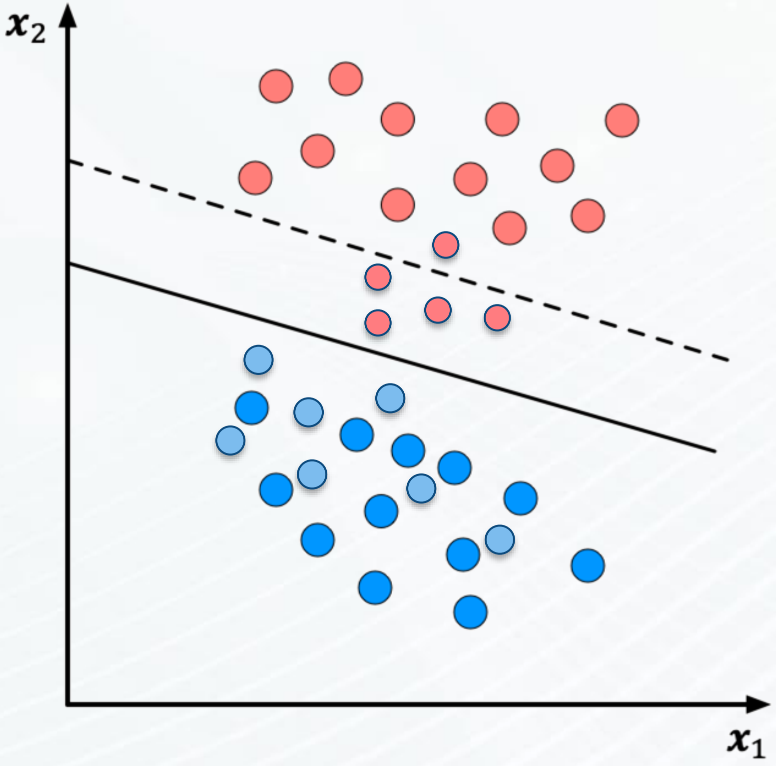
- 哪个分类器的性能更好?



(a) (b) (c)

一、问题原型

- **什么是模型的泛化能力?**
  - 分类器对未知数据的处理能力
- **如何根据泛化性能考虑模型设计?**



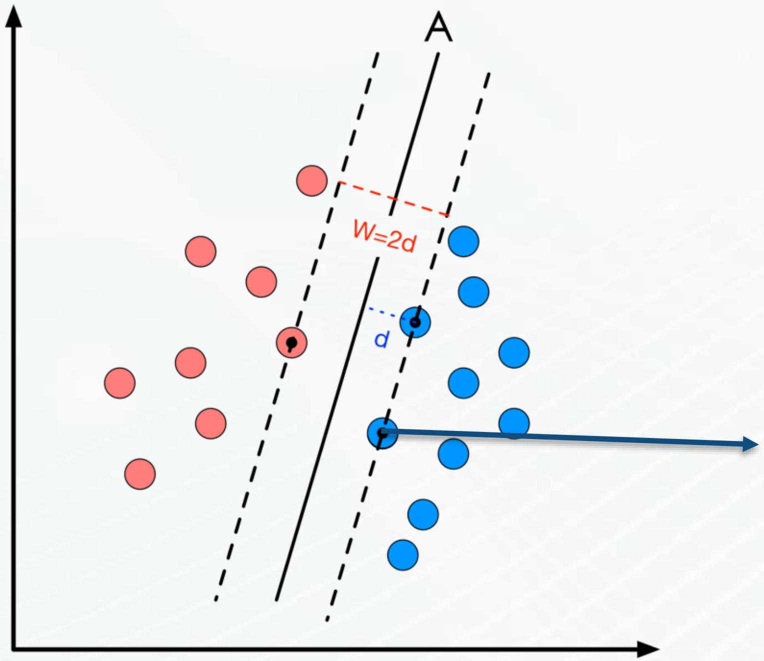
一、问题原型

▶ SVM问题原型

- 决策面由方向参数ω和空间位置参数γ共同决定

间隔:

$W = 2d$



目标

找到能产生**最大间隔**的决策面

支持向量

每类中距离决策面最近的样本

- 决策面必须与两类的支撑向量具有相同的距离

CHAPTER ONE

数学模型

Mathematic Model

二、数学模型

SVM约束优化问题 —— “三要素”

- 优化变量：决策面方程

如何定义？

- 目标函数：分类间隔（最大化）

如何计算？

- 约束条件：所有训练样本被正确分类

如何用数学形式表达？

二、数学模型

▶ 要素1：优化变量——超平面方程

- 线性决策超平面方程

$$g(\boldsymbol{x}) = \boldsymbol{\omega}^T\boldsymbol{x} + \gamma = 0 \Longrightarrow \text{参数：} \boldsymbol{\omega}\text{和}\gamma$$

1. $\boldsymbol{\omega} = [\omega_1, \omega_2]^T$

   - 控制决策面方向

2. 截距$-\gamma/\omega_1$和$-\gamma/\omega_2$

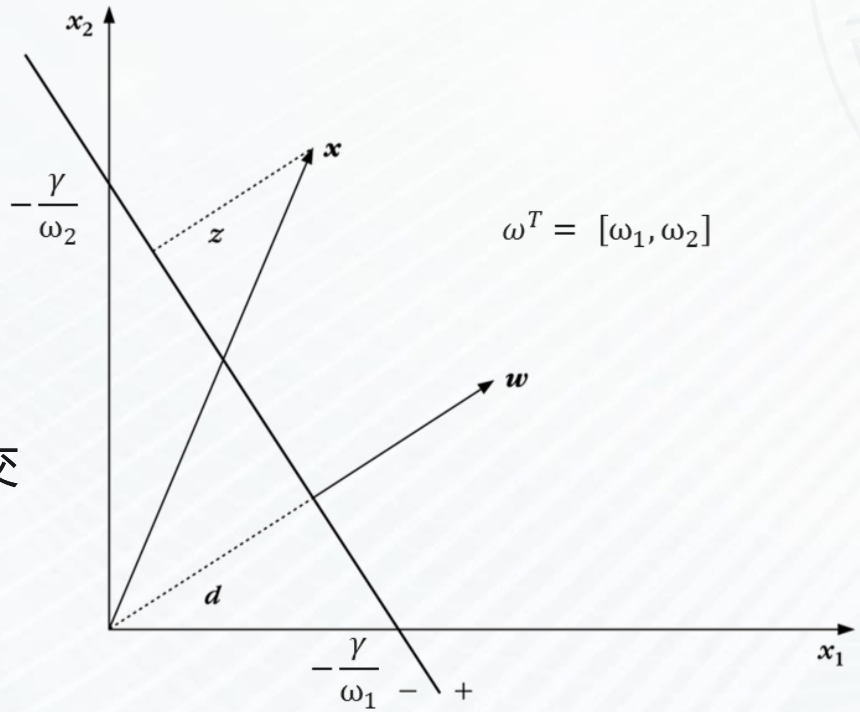
   - 决定决策面位置

3. $\boldsymbol{\omega} \perp g(\boldsymbol{x}) = 0$

   - 向量$\boldsymbol{\omega}$与决策面$g(\boldsymbol{x}) = 0$正交

4. 点$\boldsymbol{x}$到决策面的距离

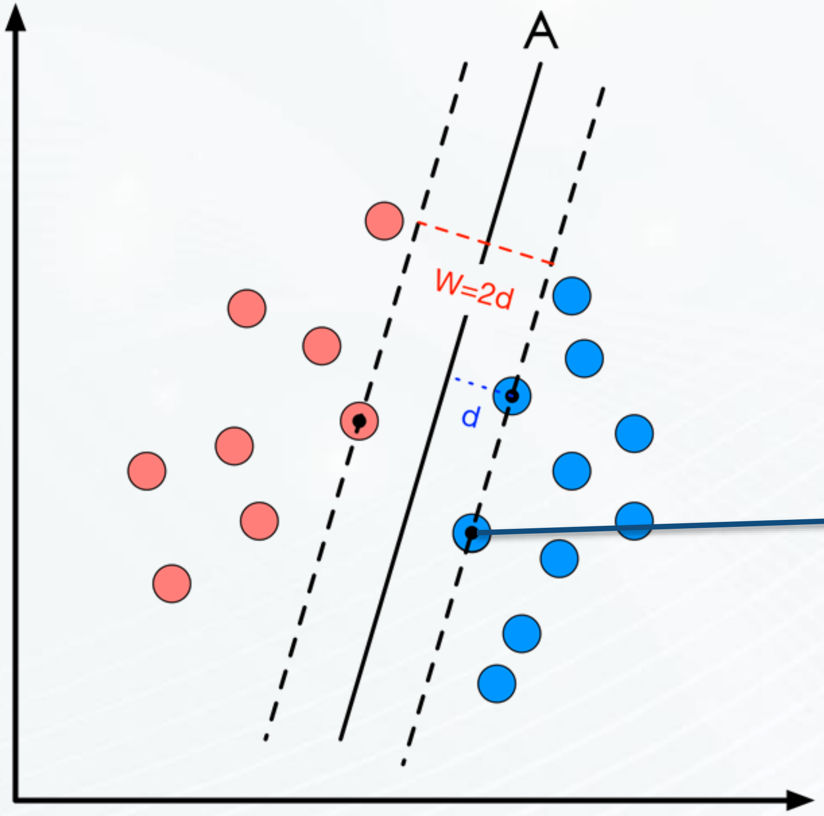
   - $d = |g(\boldsymbol{x})|/\|\boldsymbol{\omega}\|$



二、数学模型

▶ 要素2：目标函数——间隔的计算

- 分类间隔：支持向量到决策面距离的2倍



$$W = 2d$$

$$d = \frac{|\boldsymbol{\omega}^T \boldsymbol{x}_s + \gamma|}{\|\boldsymbol{\omega}\|}$$

支持向量: $\boldsymbol{x}_s$

二、数学模型

▶ 要素3：约束条件——基本描述

- 对方向参数向量$\boldsymbol{\omega}$的约束

  - 并非所有方向都能找到使训练样本被完全正确分类的决策面

- 对位置参数γ的约束

  - 决策超平面的位置必须是间隔区域的中心

- 方程$d = \frac{|\boldsymbol{\omega}^T x_S + \gamma|}{\|\boldsymbol{\omega}\|}$对支撑向量$\boldsymbol{x}_S$的约束

  - 支持向量到中心轴的距离为d

  - 支撑向量的选择决定了决策超平面的方向和位置

二、数学模型

▶ 要素3：约束条件——约束条件的整合

- 整合三个约束条件：步骤1

  - 假设每个样本有标签y_i

$$y_i = \begin{cases} +1 & x_i \in A\text{类} \\ -1 & x_i \in B\text{类} \end{cases}$$

  - 对于所有样本$\boldsymbol{x}_i$，正确的决策面应该满足：

$$\begin{cases} \boldsymbol{\omega}^T \boldsymbol{x}_i + \gamma > 0 & \forall\, y_i = 1 \\ \boldsymbol{\omega}^T \boldsymbol{x}_i + \gamma < 0 & \forall\, y_i = -1 \end{cases}$$

二、数学模型

▶ 要素3：约束条件——约束条件的整合

- 整合三个约束条件：步骤2
  - 决策面位于间隔区域的中心，所有支持向量到决策面距离为d；
  - 其他样本到决策面距离一定大于d，因此有：

$$\begin{cases} \dfrac{\boldsymbol{\omega}^T\boldsymbol{x}_i+\gamma}{\|\boldsymbol{\omega}\|} \geq d, & \forall y_i=1 \\ \dfrac{\boldsymbol{\omega}^T\boldsymbol{x}_i+\gamma}{\|\boldsymbol{\omega}\|} \leq -d, & \forall y_i=-1 \end{cases}$$

令： $$\boldsymbol{\omega}_d=\frac{\boldsymbol{\omega}}{\|\boldsymbol{\omega}\|d}, \qquad \gamma_d=\frac{\gamma}{\|\boldsymbol{\omega}\|d}$$

$$\begin{cases} \boldsymbol{\omega}_d^T\boldsymbol{x}_i+\gamma_d \geq 1, & \forall\ y_i=1 \\ \boldsymbol{\omega}_d^T\boldsymbol{x}_i+\gamma_d \leq -1, & \forall\ y_i=-1 \end{cases}$$

二、数学模型

▶ 要素3：约束条件——约束条件的整合

- 整合三个约束条件：步骤3

  - 比较两个线性方程：$\boldsymbol{\omega}_d^T x + \gamma_d = 0 \qquad (1)$

    $$\boldsymbol{\omega}^T x + \gamma = 0 \qquad (2)$$

  - 其中：$\boldsymbol{\omega}_d = \frac{\boldsymbol{\omega}}{\|\boldsymbol{\omega}\| d}$，$\gamma_d = \frac{\gamma}{\|\boldsymbol{\omega}\| d}$ 所以：(1) 和 (2) 是同一个超平面！

  - 使用$\boldsymbol{\omega}$和γ对变量$\boldsymbol{\omega}_d$和γ_d进行重命名，组合约束可以写为：

$$\begin{cases} \boldsymbol{\omega}^T \boldsymbol{x}_i + \gamma \geq 1, & \forall y_i = 1 \\ \boldsymbol{\omega}^T \boldsymbol{x}_i + \gamma \leq -1, & \forall y_i = -1 \end{cases}$$

  - 将约束条件精简为统一表述：$y_i\left(\boldsymbol{\omega}^T \boldsymbol{x}_i + \gamma\right) \geq 1, \ \forall \boldsymbol{x}_i$

    $\boldsymbol{x}_i$是支撑向量时，等号成立

二、数学模型

▶ 建立数学模型

- 间隔宽度W=2d可以被简化：

$$d = \frac{\left|\omega^T x_s + \gamma\right|}{\|\omega\|} = \frac{1}{\|\omega\|}$$

- 为什么可以简化？

  - 间隔宽度由决策面方向确定
  - 决策面方向须满足约束条件；

- 目标函数的新形式：

$$\max_{\omega,\gamma} 2d \Longrightarrow \max_{\omega,\gamma} \frac{2}{\|\omega\|} \Longrightarrow \min_{\omega,\gamma} \frac{1}{2}\|\omega\|^2$$

二、数学模型

▶ 建立数学模型

- 线性SVM优化问题的数学描述

$$\min_{\boldsymbol{\omega},\gamma} \frac{1}{2}\|\boldsymbol{\omega}\|^2$$

$$\text{s. t.} \quad y_i\left(\boldsymbol{\omega}^T \boldsymbol{x}_i + \gamma\right) \geq 1, \qquad i = 1,2, \ldots, N$$

- 数学模型特性

  - 目标函数为二次多项式函数，属于有约束凸优化问题

  - N个不等式约束条件，每个训练样本$\{\boldsymbol{x}_i, \ y_i\}$对应一个约束条件；

总 结

SVM数学模型

- 问题原型
  - 线性分类器性能比较
  - 泛化能力解释
  - 提出问题原型
    - 间隔
    - 支持向量
    - 提出问题
- 数学模型
  - SVM优化问题概述
  - 优化变量
    - 超平面方程
  - 目标函数
    - 间隔计算
  - 约束条件
    - 基本描述
    - 多条约束条件的整合
  - 建立数学模型
    - 规范化数学描述
    - 凸优化问题的提出

M L & P R

THANK YOU

感谢聆听