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

Solving Non-linear SVM

非线性SVM求解

CHAPTER ONE

非线性分类器策略

Strategy of Non-linear Classifier

一、非线性分类器策略

▶ 构造非线性决策面

将判别函数$g(\boldsymbol{x})$从线性转化为非线性，此时决策面$g(\boldsymbol{x})=0$是特征空间中的一个超曲面

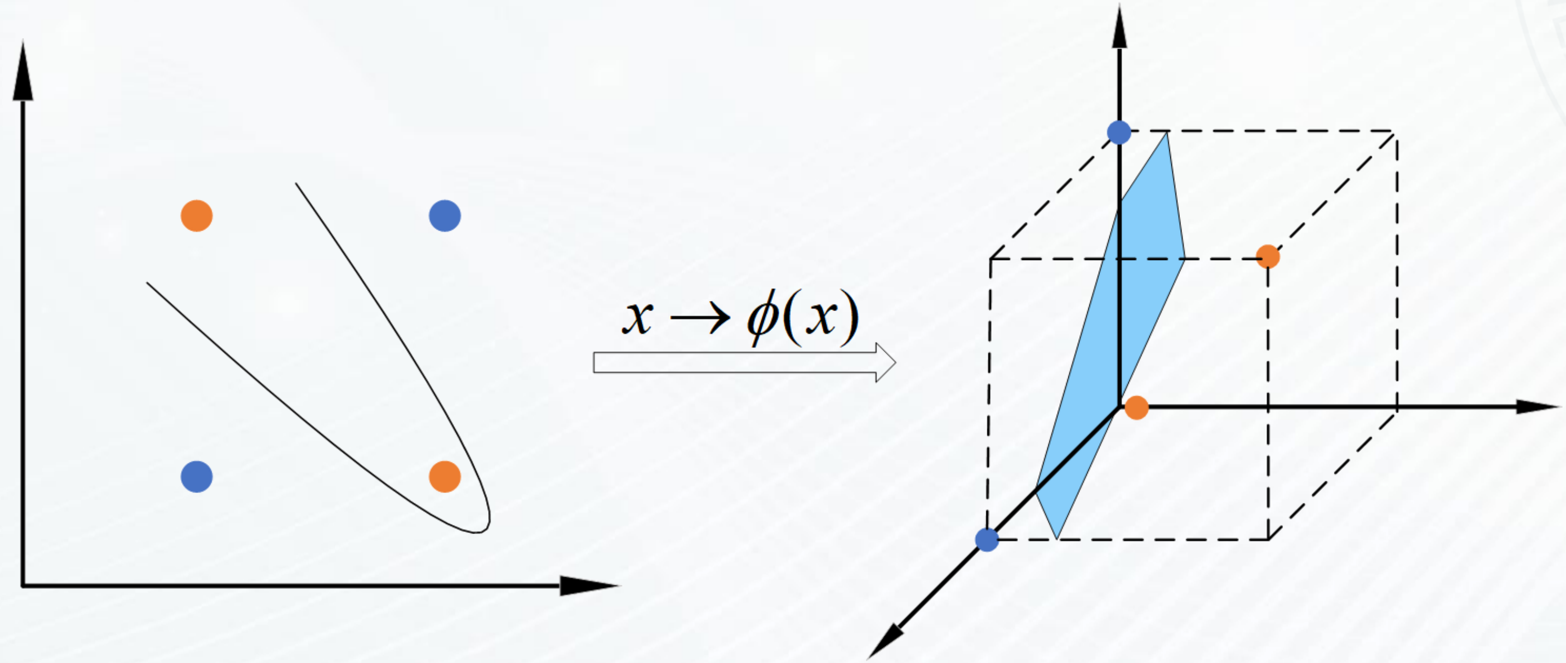
- **新问题**

  - 判别函数$g(\boldsymbol{x})$的非线性函数形式如何设定?
  - 判别函数$g(\boldsymbol{x})$的参数如何设计?
  - 非线性约束条件下的优化问题如何求解?

一、非线性分类器策略

▶ 另一种思路：非线性映射

- 将所有样本非线性映射到某个高维空间
- 使训练数据的不可分问题在该空间中变成线性可分问题。



CHAPTER TWO

核化SVM

Kernel SVM

二、核化SVM

核函数概念

- 对数据集进行非线性映射之后，原始的优化问题变为：

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

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

- 拉格朗日对偶问题为：

$$\max_{\boldsymbol{\alpha}}\left[\sum_{i=1}^{m}\alpha_i-\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_i\alpha_j y_i y_j \phi(\boldsymbol{x}_i)^T\phi(\boldsymbol{x}_j)\right]$$

$$s.t. \qquad \textstyle\sum_{i=1}^{m}\alpha_i y_i=0$$

$$0\leq\alpha_i\leq C, i=1,2,\dots,m$$

- 核函数：$K(\boldsymbol{x}_i,\boldsymbol{x}_j)$

$$K(\boldsymbol{x}_i,\boldsymbol{x}_j)=\langle\phi(\boldsymbol{x}_i),\phi(\boldsymbol{x}_j)\rangle=\phi(\boldsymbol{x}_i)^T\phi(\boldsymbol{x}_j)$$

二、核化SVM

核函数意义

- 原因：为什么将 $\phi(\boldsymbol{x}_i)^T\phi(\boldsymbol{x}_j)$ 定义为核函数?

- 挑战：$\phi(\boldsymbol{x}_i)$很难设计，为了保证非线性映射后的线性可分性，映射后的空间维数可能会很高，甚至是无穷大。

- 技巧：由于$\phi(\boldsymbol{x}_i)^T\phi(\boldsymbol{x}_j)$项不会被单独计算，因此我们并不需要知道$\phi(\boldsymbol{x}_i)$的确切值。

- 效果：使用核函数$K\left(\boldsymbol{x}_i,\boldsymbol{x}_j\right)=\phi(\boldsymbol{x}_i)^T\phi(\boldsymbol{x}_j)$能够解决问题。

二、核化SVM

▶ 重现非线性SVM的数学模型

- 利用核函数，拉格朗日对偶问题可重写为：

$$\max_{\boldsymbol{\alpha}}\left[\sum_{i=1}^{m}\alpha_i-\frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_i\alpha_j y_i y_j K\left(\boldsymbol{x}_i,\boldsymbol{x}_j\right)\right]$$

$$s.t.\qquad \sum_{i=1}^{m}\alpha_i y_i=0$$

$$0\le\alpha_i\le C, i=1,2,\dots,m$$

- 问题的解为：

$$h(\boldsymbol{x})=\boldsymbol{\omega}^T\phi(\boldsymbol{x})+\gamma$$

$$=\sum_{i=1}^{m}\alpha_i y_i\phi(\boldsymbol{x}_i)^T\phi(\boldsymbol{x})+\gamma$$

$$=\sum_{i=1}^{m}\alpha_i y_i K(\boldsymbol{x}_i,\boldsymbol{x})+\gamma$$

使用核化SVM分类时，
也只需要计算$K(\boldsymbol{x}_i,\boldsymbol{x})$，
不需要计算$\phi(\boldsymbol{x}_i)$

CHAPTER THREE

核函数

Kernel Function

三、核函数

▶ 核函数的判定与选择

- 什么样的函数才是核函数?

从核函数的定义来看，函数$K(\boldsymbol{x}_i,\boldsymbol{x}_j)$是核函数的充分必要条件是：

- $K(\boldsymbol{x}_i,\boldsymbol{x}_j)$可以写成映射函数$\phi(\boldsymbol{x}_i)$和$\phi(\boldsymbol{x}_j)$的内积；
- 映射函数$\phi(\boldsymbol{x})$能够将特征向量$\boldsymbol{x}$映射到一个新的希尔伯特空间。

定理 4.1（核函数判定）： 令$\mathcal{X}$为输入空间，函数K是定义在$\mathcal{X}\times\mathcal{X}$上的对称函数，当且仅当对于任意数据$D=\{\boldsymbol{x}_i\in\mathcal{X}|i=1,2,\ldots,m\}$，函数$K$的Gram矩阵$\mathcal{K}$均为半正定时，函数$K$是核函数。

➢ 这里函数K的Gram矩阵$\mathcal{K}$定义为:

$$\mathcal{K}=\begin{bmatrix} K(\boldsymbol{x}_1,\boldsymbol{x}_1) & \cdots & K(\boldsymbol{x}_1,\boldsymbol{x}_m) \\ \vdots & \ddots & \vdots \\ K(\boldsymbol{x}_m,\boldsymbol{x}_1) & \cdots & K(\boldsymbol{x}_m,\boldsymbol{x}_m) \end{bmatrix}$$

三、核函数

▶ 常用核函数

一些常用的核函数

- 多项式核函数： $K(\boldsymbol{x}_i, \boldsymbol{x}_j) = (\boldsymbol{x}_i^T \boldsymbol{x}_j + 1)^q, q \geq 1$
- 高斯核函数： $K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp\left(-\frac{\|\boldsymbol{x}_i - \boldsymbol{x}_j\|^2}{2\sigma^2}\right), \sigma^2 > 0$
- Sigmoid核函数： $K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \tanh(\beta \boldsymbol{x}_i{}^T \boldsymbol{x}_j + \theta), \beta, \theta > 0$

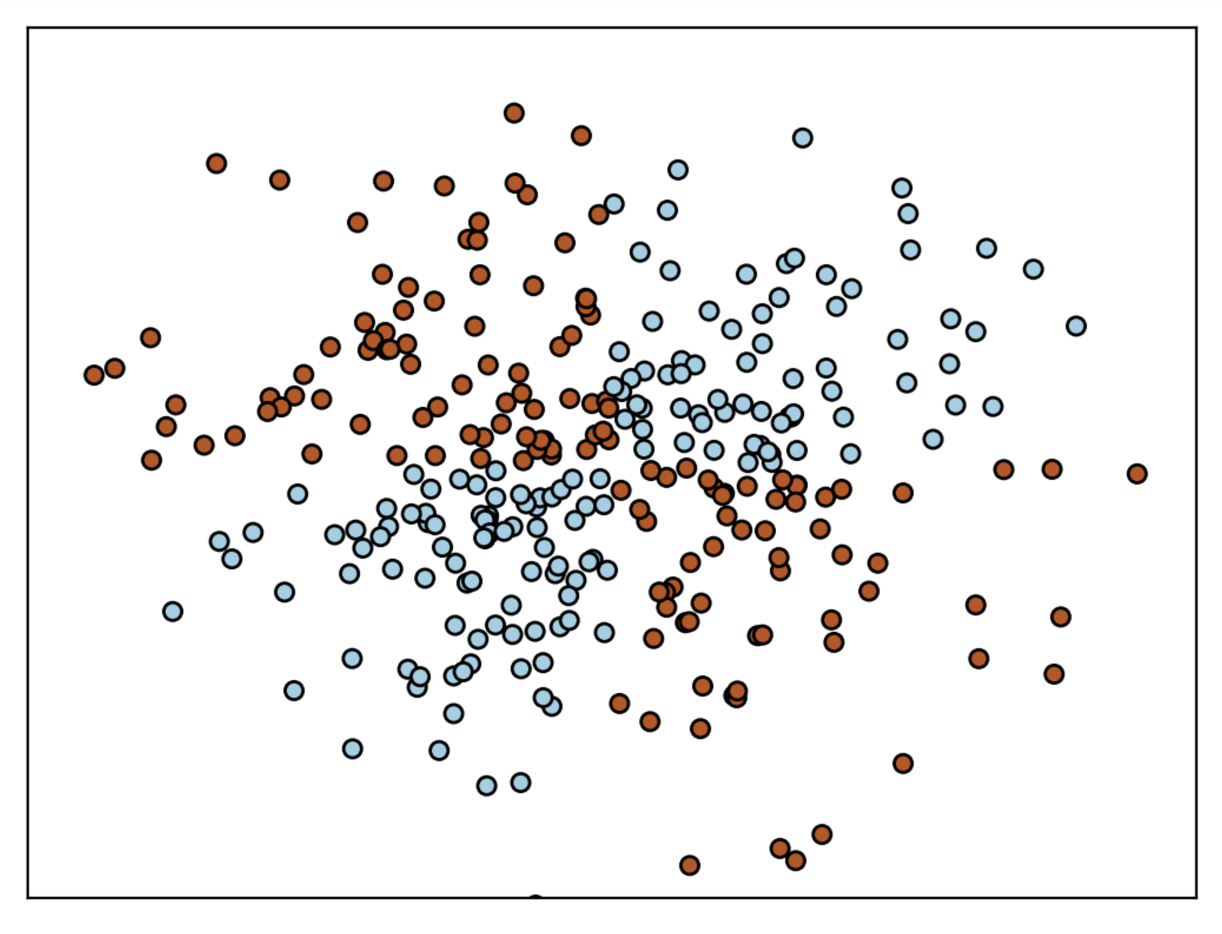
核函数的性质

- ➢ $K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \beta_1 K_1(\boldsymbol{x}_i, \boldsymbol{x}_j) + \beta_2 K_2(\boldsymbol{x}_i, \boldsymbol{x}_j)$
- ➢ $K(\boldsymbol{x}_i, \boldsymbol{x}_j) = K_1(\boldsymbol{x}_i, \boldsymbol{x}_j)\, K_2(\boldsymbol{x}_i, \boldsymbol{x}_j)$
- ➢ $K(\boldsymbol{x}_i, \boldsymbol{x}_j) = g(\boldsymbol{x}_i) K_1(\boldsymbol{x}_i, \boldsymbol{x}_j) g(\boldsymbol{x}_j), \ \forall g(\boldsymbol{x})$

示 例

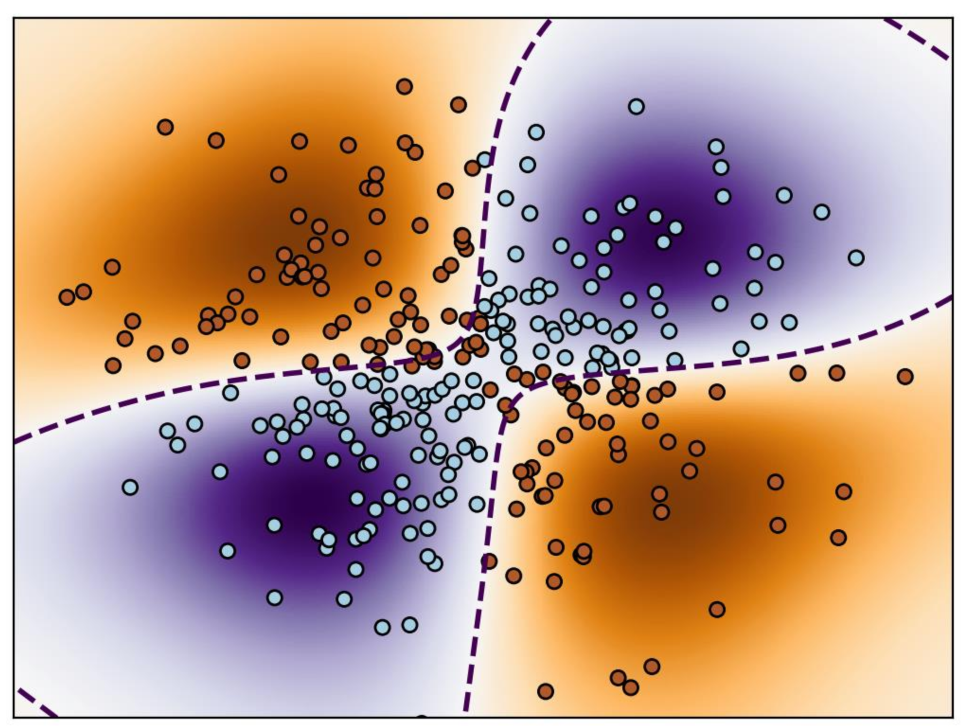
示例——非线性SVM的应用

- 线性不可分的两类样本：蓝色 vs 红色



示例——非线性SVM的应用

- 使用非线性SVM算法设计分类器。
- 调用 sklearn 库中的svm.NuSVC()函数;
- 函数参数使用默认设置即可（高斯核函数）。
- 分类效果如下:



总 结

- 非线性SVM求解
  - 非线性分类器策略
    - 构造非线性决策面
    - 非线性映射
  - 核化SVM
    - 引入核函数
    - 核函数的技巧
    - 非线性映射
  - 核函数
    - 核函数的判定
    - 常用核函数
      - 多项式核
      - 高斯核
      - Sigmoid核
  - 示例：非线性SVM的应用

S M A R T

THANK YOU

感谢聆听

MADE BY DONGYUE CHEN