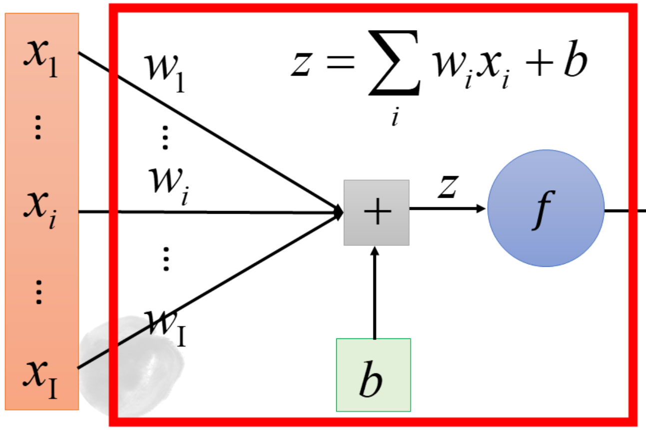
机器学习

苏州大学计算机科学与技术学院
自然语言处理实验室
主讲：周夏冰
邮箱：zhouxiabing@suda.edu.cn

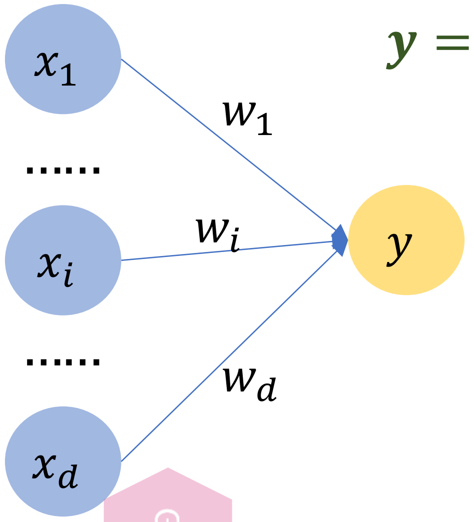
前情回顾



$$P_{w,b}(C_1|x) = \frac{1}{1+e^{-z}}$$

$$P_{w,b}(C_2|x) = 1 - P_{w,b}(C_1|x)$$

输入维度为d的x



$$y = \frac{1}{1+e^{-\left(\sum_{i=1}^{d} x_i w_i + b\right)}}$$

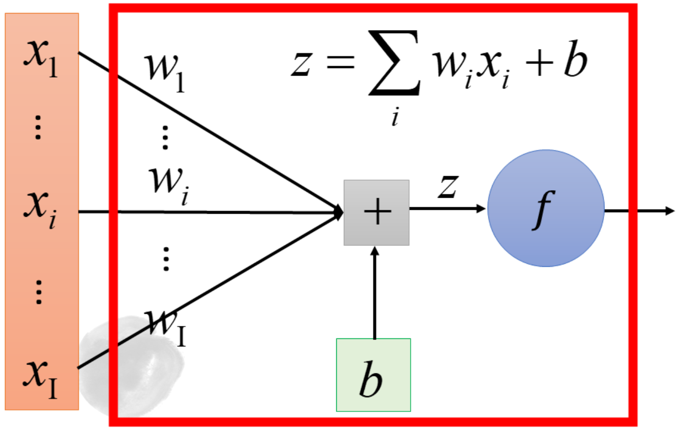
前情回顾

Step 1: $f_{w,b}(x) = \sigma(wx + b) = \frac{1}{1 + exp(-wx + b)}$ Output: between 0 and 1

Step 2: Cross entropy: $l(f(x^n), \hat{y}^n) = -\left[\hat{y}^n lnf(x^n) + (1 - \hat{y}^n)ln\left(1 - f(x^n)\right)\right]$

Step 3: Logistic regression: $w_i \leftarrow w_i - \eta \sum_n -\left(\hat{y}^n - f_{w,b}(x^n)\right) x_i^n$

前情回顾

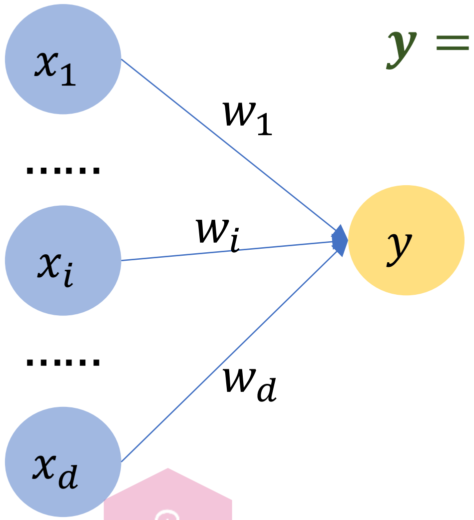


$$P_{w,b}(C_1|x) = \frac{1}{1+e^{-z}}$$

$$P_{w,b}(C_2|x) = 1 - P_{w,b}(C_1|x)$$

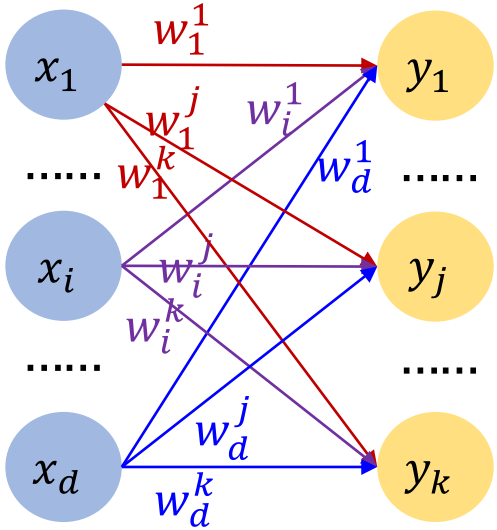
$$z_i = \sum_{j=1}^{d} x_j w_j^i + b_i$$

输入维度为d的x



$$y = \frac{1}{1+e^{-(\sum_{i=1}^{d} x_i w_i + b)}}$$

K个类别



$$y_1 = \frac{e^{z_1}}{\sum_{i=1}^{k} e^{z_i}}$$

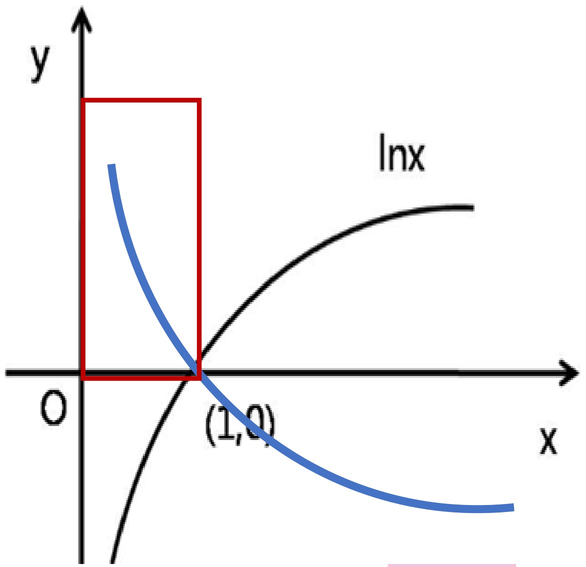
$$y_j = \frac{e^{z_j}}{\sum_{i=1}^{k} e^{z_i}}$$

$$y_k = \frac{e^{z_k}}{\sum_{i=1}^{k} e^{z_i}}$$

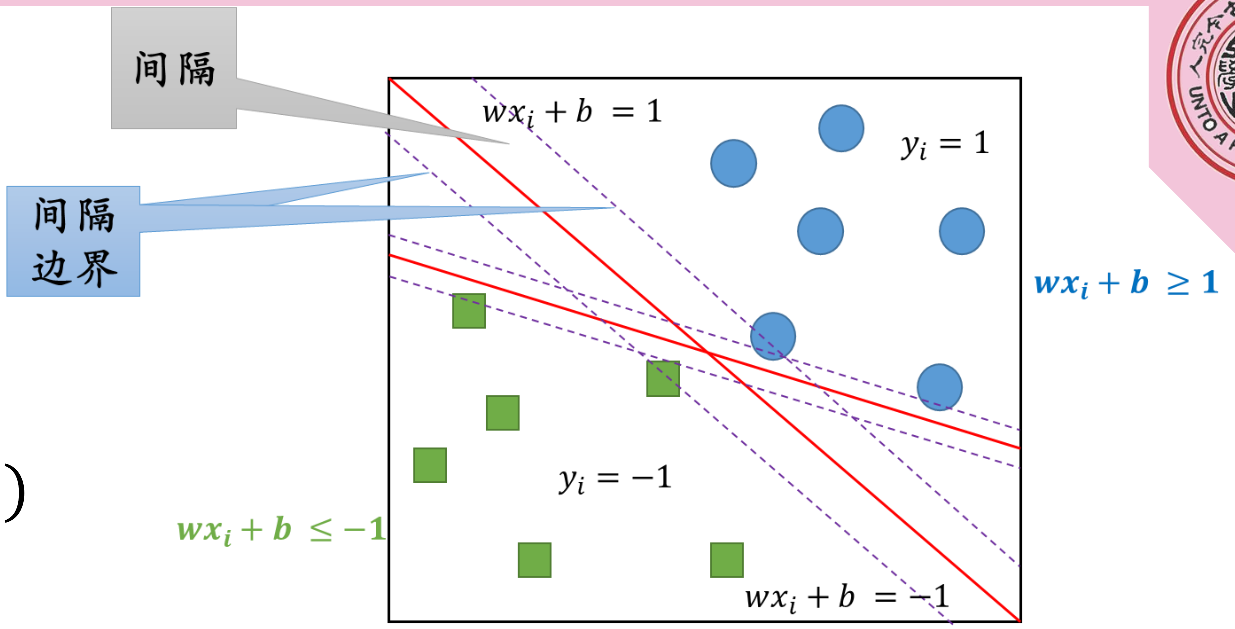
前情回顾

- 二分类
  - $l=-\sum_{i=1}^{N}[y_i \ln f(x_i)+(1-y_i)\ln(1-f(x_i))]$
- 多分类
  - $l=-\sum_{i=1}^{N}\sum_{k=1}^{K} y_i^k lnf(x_i)^k$

If x ∈ class 1	If x ∈ class 2	If x ∈ class 3
$\hat{y}=\begin{bmatrix}1\\0\\0\end{bmatrix}$	$\hat{y}=\begin{bmatrix}0\\1\\0\end{bmatrix}$	$\hat{y}=\begin{bmatrix}0\\0\\1\end{bmatrix}$
$-lny_1$	$-lny_2$	$-lny_3$



前情回顾



- 线性可分支持向量机
  - 决策函数 $f(x) = sign(w \cdot x + b)$
  - 目标：提高模型的泛化能力
  - 基本思想：寻找在特征空间上间隔最大的线性分类器

前情回顾

- 线性可分支持向量机
  - 决策函数 $f(x) = sign(w \cdot x + b)$
  - 目标：提高模型的泛化能力
  - 基本思想：寻找在特征空间上间隔最大的线性分类器
  - 目标函数
    - $\min\limits_{w,b} \frac{1}{2}\|w\|^2$
    - $s.t.\ \ y_i(wx_i + b) - 1 \geq 0, \quad i = 1, 2, \cdots, N$

前情回顾

- 线性可分支持向量机
  - 目标函数
    - $\min\limits_{w,b} \frac{1}{2}\|w\|^2$
    - $s.t.\ \ y_i(wx_i + b) - 1 \geq 0, \quad i = 1,2,\cdots,N$
  - 拉格朗日乘子法
    - 带有约束的优化问题
    - 约束条件函数与原函数联立
  - $L = \frac{1}{2}||w||^2 \boxed{+} \sum_{i=1}^{N} \alpha_i(\boxed{1 - y^i(w \cdot x^i + b)})$
  - $\min\limits_{b,w}(\max\limits_{\alpha_i \geq 0} L(w,b,\alpha))$

前情回顾

- 线性可分支持向量机
  - $L = \frac{1}{2}||\boldsymbol{w}||^2 + \sum_{i=1}^{N} \alpha_i(1 - y^i(\boldsymbol{w} \cdot \boldsymbol{x}^i + b))$
  - $\min_{b,w}(\max_{\alpha_i \geq 0} L(w, b, \alpha))$
  - 对偶算法
    - 增加模型可解释性
    - 自然引入核函数
    - $\max_{\alpha_i \geq 0}(\min_{b,w} L(w, b, \alpha))$

- 无论原始问题是否是凸的，拉格朗日对偶可以转化为凸优化问题；
- 对偶问题可以给出原始问题一个下界；

前情回顾

- 线性可分支持向量机
  - $L = \frac{1}{2}||\boldsymbol{w}||^2 + \sum_{i=1}^{N} \alpha_i(1 - y^i(\boldsymbol{w} \cdot x^i + b))$
  - $\min_{b,w}(\max_{\alpha_i \geq 0} L(w, b, \alpha))$
  - 对偶算法
    - 增加模型可解释性
    - 自然引入核函数
    - $\max_{\alpha_i \geq 0}(\min_{b,w} L(w, b, \alpha))$

原函数：$f(\theta) = \max_{\alpha,\beta} L(\theta, \alpha, \beta)$

对偶函数：$D(\alpha, \beta) = \min_{\theta} L(\theta, \alpha, \beta)$

$$D(\alpha^*, \beta^*) = \min_{\theta} f(\theta) + \sum \alpha^* g(\theta) + \sum \beta^* h(\theta)$$

$$\leq f(\theta^*) + \sum \alpha^* g(\theta^*) + \sum \beta^* h(\theta^*)$$

$$\leq f(\theta^*)$$

前情回顾

- 线性可分支持向量机
  - $L = \frac{1}{2}||\boldsymbol{w}||^2 + \sum_{i=1}^{N} \alpha_i(1 - y^i(\boldsymbol{w} \cdot \boldsymbol{x}^i + b))$
  - $\min_{b,w}(\max_{\alpha_i \geq 0} L(\boldsymbol{w}, b, \alpha))$
  - 对偶算法
    - 增加模型可解释性
    - 自然引入核函数
  - $\max_{\alpha_i \geq 0}\left(\boxed{\min_{b,w} L(\boldsymbol{w}, b, \alpha)}\right)$

$\theta = (\boldsymbol{w}, b)$

原函数：$f(\theta) = \max_{\alpha,\beta} L(\theta, \alpha, \beta)$

对偶函数：$D(\alpha, \beta) = \min_{\theta} L(\theta, \alpha, \beta)$

$$D(\alpha^*, \beta^*) = \min_{\theta} f(\theta) + \sum \alpha^* g(\theta) + \sum \beta^* h(\theta)$$

$$= f(\theta^*) + \sum \alpha^* g(\theta^*) + \sum \beta^* h(\theta^*)$$

$$\leq f(\theta^*)$$

前情回顾

- 线性可分支持向量机
  - $L = \frac{1}{2}||w||^2 + \sum_{i=1}^{N} \alpha_i (1 - y^i(w \cdot x^i + b))$
  - $\min_{b,w}(\max_{\alpha_i \geq 0} L(w, b, \alpha))$
  - 对偶算法
    - 增加模型可解释性
    - 自然引入核函数
    - $\max_{\alpha_i \geq 0}(\min_{b,w} L(w, b, \alpha))$

原函数：$f(\theta) = \max_{\alpha,\beta} L(\theta, \alpha, \beta)$

对偶函数：$D(\alpha, \beta) = \min_{\theta} L(\theta, \alpha, \beta)$

$$D(\alpha^*, \beta^*) = \min_{\theta} f(\theta) + \sum \alpha^* g(\theta) + \sum \beta^* h(\theta)$$

$$= f(\theta^*) + \sum \alpha^* g(\theta^*) + \sum \beta^* h(\theta^*)$$

$$= f(\theta^*)$$

$\alpha^* g(x^*) = 0$

前情回顾

- $\min\limits_{bw} L = \frac{1}{2}||w||^2 + \sum_{i=1}^{N} \alpha_i (1 - y^i (w \cdot x^i + b))$

$$\frac{\partial L}{\partial w_j} = w_j - \sum_{i=1}^{N} \alpha_i y^i x_j^i = 0 \quad w_j = \sum_{i=1}^{N} \alpha_i y^i x_j^i$$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^{N} \alpha_i y^i = 0$$

- 带入

$$\max_{\alpha} L = -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y^i y^j (x^i \cdot x^j) + \sum_{i=1}^{N} \alpha_i$$

$$\sum_{i=1}^{N} \alpha_i y^i = 0 \qquad \alpha_i \geq 0$$

SVM对偶问题解

- 根据KKT条件：

$$\nabla_w L(w^*,b^*,\alpha^*) = w^* - \sum_{i=1}^{N} \alpha_i^* y_i x_i = 0 \quad \longrightarrow \quad w^* = \sum_{i=1}^{N} \alpha_i^* y_i x_i$$

$$\nabla_b L(w^*,b^*,\alpha^*) = -\sum_{i=1}^{N} \alpha_i^* y_i = 0 \qquad b^* = y_j - \sum_{i=1}^{N} \alpha_i^* y_i (x_i \cdot x_j)$$

$$\boxed{\alpha_i^*(y_i(w^* \cdot x_i + b^*) - 1) = 0}, \quad i = 1,2,\cdots,N$$

$$y_i(w^* \cdot x_i + b^*) - 1 \geqslant 0, \quad i = 1,2,\cdots,N$$

$$\alpha_i^* \geqslant 0, \quad i = 1,2,\cdots,N$$

对于正分量$\alpha_j > 0$

$y_j\left(w^* x_j + b^*\right) = 1$

前情回顾

- 线性可分支持向量机
  - $L = \frac{1}{2}||\boldsymbol{w}||^2 + \sum_{i=1}^{N} \alpha_i(1 - y^i(\boldsymbol{w} \cdot \boldsymbol{x}^i + b))$
  - $\min_{b,w}(\max_{\alpha_i \geq 0} L(\boldsymbol{w}, b, \boldsymbol{\alpha}))$
  - 对偶算法
    - 增加模型可解释性
    - 自然引入核函数
    - $\max_{\alpha_i \geq 0}(\min_{b,w} L(\boldsymbol{w}, b, \boldsymbol{\alpha}))$

kkT条件

$$\alpha_i^*\left(1 - y^i(\boldsymbol{w} \cdot \boldsymbol{x}^i + b)\right) = 0, i = 1, 2, \cdots, N$$

$$w_j = \sum_{i=1}^{N} \alpha_i y^i x_j^i$$

$$b = y^j - \sum_{i=1}^{N} \alpha_i y^i (\boldsymbol{x}^i \cdot \boldsymbol{x}^j)$$

$$\min_{\alpha} L = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i \alpha_j y^i y^j (\boldsymbol{x}^i \cdot \boldsymbol{x}^j) - \sum_{i=1}^{N} \alpha_i$$

$$\sum_{i=1}^{N} \alpha_i y^i = 0$$

$$\alpha_i \geq 0$$

SVM对偶问题

- **SVM目标函数：**
  - $\max\limits_{\alpha} L = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j\, y^i y^j(\boldsymbol{x}^i \cdot \boldsymbol{x}^j) + \sum_{i=1}^{N}\alpha_i$
  - $\min\limits_{\alpha} L = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j\, y^i y^j(\boldsymbol{x}^i \cdot \boldsymbol{x}^j) - \sum_{i=1}^{N}\alpha_i$
  - $\sum_{i=1}^{N}\alpha_i y^i = 0$
  - $\alpha_i \geq 0$

例7.2

$$\min_{\alpha} L = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y^i y^j(\boldsymbol{x}^i\cdot\boldsymbol{x}^j) - \sum_{i=1}^{N}\alpha_i$$
$$\sum_{i=1}^{N}\alpha_i y^i = 0$$
$$\alpha_i \geq 0$$

- **正实例$x_1 = (3，3)^T$, $x_2 = (4，3)^T$, 负实例$x_3 = (1，1)^T$**

$$\min_{\alpha}\quad \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j (x_i\cdot x_j) - \sum_{i=1}^{N}\alpha_i$$

$$= \alpha_1^2 y_1^2 (3,3)\cdot(3,3) + 2\alpha_1\alpha_2 y_1 y_2 (3,3)\cdot(4,3) + 2\alpha_1\alpha_3 y_1 y_3 (3,3)\cdot(1,1) + 2\alpha_2\alpha_3 y_2 y_3 (4,3)\cdot(1,1) + \alpha_2^2 y_2^2 (4,3)\cdot(4,3) + \alpha_3^2 y_3^2 (1,1)\cdot(1,1)$$

$$= \frac{1}{2}(18\alpha_1^2 + 25\alpha_2^2 + 2\alpha_3^2 + 42\alpha_1\alpha_2 - 12\alpha_1\alpha_3 - 14\alpha_2\alpha_3) - \alpha_1 - \alpha_2 - \alpha_3$$

$$\text{s.t.}\quad \alpha_1 + \alpha_2 - \alpha_3 = 0$$

$$\alpha_i \geqslant 0,\quad i = 1,2,3$$

$$\alpha_3 = \alpha_1 + \alpha_2 \longrightarrow s(\alpha_1,\alpha_2) = 4\alpha_1^2 + \frac{13}{2}\alpha_2^2 + 10\alpha_1\alpha_2 - 2\alpha_1 - 2\alpha_2$$

例7.2

$$\frac{\partial s}{\partial \alpha_1} = 8\alpha_1 + 10\alpha_2 - 2 = 0$$

$$\frac{\partial s}{\partial \alpha_2} = 13\alpha_2 + 10\alpha_1 - 2 = 0$$

- $s(\alpha_1, \alpha_2) = 4\alpha_1^2 + \frac{13}{2}\alpha_2^2 + 10\alpha_1\alpha_2 - 2\alpha_1 - 2\alpha_2$
- 1. 两个均不为0，可得：$\alpha_1 = \frac{3}{2}, \alpha_2 = -1$ ×
- 2. $\alpha_1 = 0, \alpha_2 = 0$ ×

$$\frac{\partial s}{\partial \alpha_2} = 13\alpha_2 - 2 = 0$$

- 3. $\alpha_1 = 0, \alpha_2 \neq 0$，可得：$\alpha_2 = \frac{2}{13}$ $s = -\frac{2}{13}$
- 4. $\alpha_2 = 0, \alpha_1 \neq 0$，可得：$\alpha_1 = \frac{1}{4}$ $s = -\frac{1}{4}$ √ $\alpha_3 = \frac{1}{4}$

$$\frac{\partial s}{\partial \alpha_1} = 8\alpha_1 - 2 = 0$$

$$\alpha_i(1 - y_i(wx_i + b)) = 0$$

⇓

$$y_i(wx_i + b) = 1$$

⇓

α_1, α_3为支持向量

02

线性支持向量机

线性支持向量机

- 当数据存在噪声，数据往往线性不可分。
- 对于这种偏离正常位置很远的噪声点，称为特异点（outlier）
- 在原来的SVM模型里，outlier的存在有可能造成很大的影响，因为超平面本身就只有少数几个支持向量组成，如果这里面再有几个特异点，那影响就很大

线性支持向量机

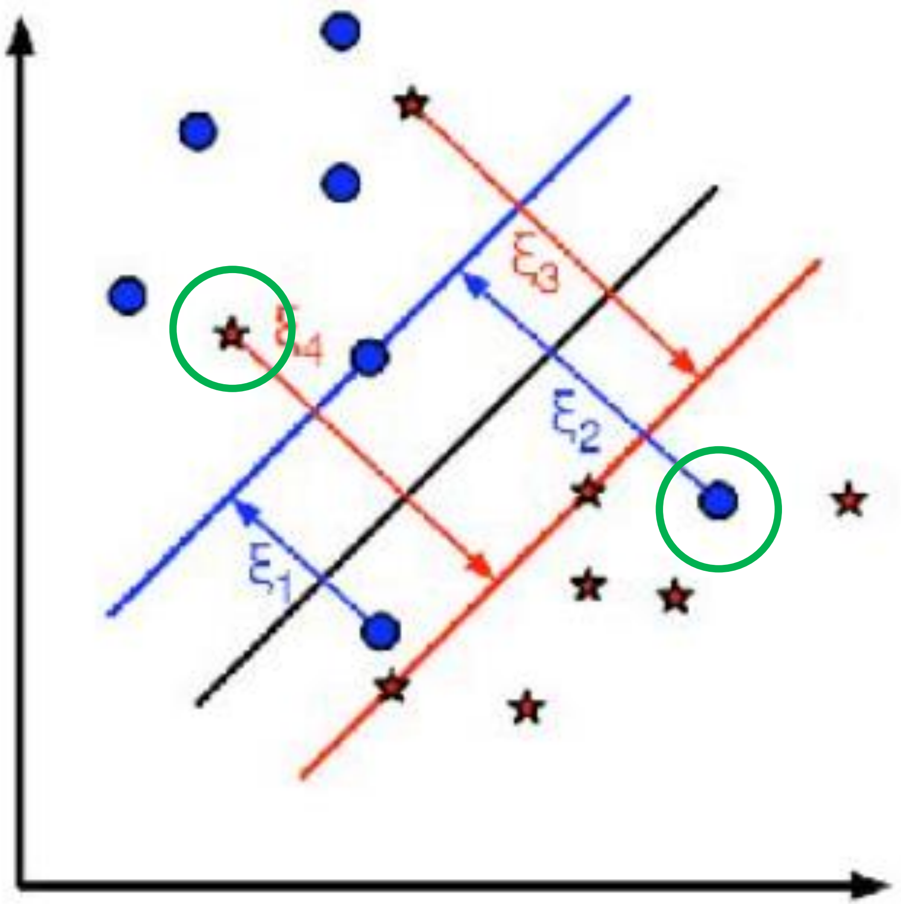
- $y_i(wx_i + b) \geq 1 - \xi_i$

- 引入变量，增加支付代价

$$\min_{w,b,\xi} \quad \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\xi_i$$

$$\text{s.t.} \quad y_i(w \cdot x_i + b) \geq 1 - \xi_i, \quad i = 1,2,\cdots,N$$

$$\xi_i \geq 0, \quad i = 1,2,\cdots,N$$



线性支持向量机

- 构建拉格朗日函数
  - $$L(w,b,\xi,\alpha,\mu)\equiv\frac{1}{2}\|w\|^2+C\sum_{i=1}^{N}\xi_i-\sum_{i=1}^{N}\alpha_i(y_i(w\cdot x_i+b)-1+\xi_i)-\sum_{i=1}^{N}\mu_i\xi_i$$
  - 先求极小值
    - 对 w,b,ξ求导

$$\nabla_w L(w,b,\xi,\alpha,\mu)=w-\sum_{i=1}^{N}\alpha_i y_i x_i=0$$

$$\nabla_b L(w,b,\xi,\alpha,\mu)=-\sum_{i=1}^{N}\alpha_i y_i=0$$

$$\nabla_{\xi_i} L(w,b,\xi,\alpha,\mu)=C-\alpha_i-\mu_i=0$$

线性支持向量机

- $$\min_{w,b,\xi} L(w,b,\xi,\alpha,\mu)=-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j(x_i\cdot x_j)+\sum_{i=1}^{N}\alpha_i$$

- 求极大值

$$\max_{\alpha}\quad -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j(x_i\cdot x_j)+\sum_{i=1}^{N}\alpha_i$$

$$\text{s.t.}\quad \sum_{i=1}^{N}\alpha_i y_i=0$$

$$C-\alpha_i-\mu_i=0$$

$$\alpha_i\geqslant 0$$

$$\mu_i\geqslant 0,\quad i=1,2,\cdots,N$$

$$\text{s.t.}\quad \sum_{i=1}^{N}\alpha_i y_i=0$$

$$0\leqslant\alpha_i\leqslant C,\quad i=1,2,\cdots,N$$

线性支持向量机

$$L(w,b,\xi,\alpha,\mu)\equiv\frac{1}{2}\|w\|^2+C\sum_{i=1}^{N}\xi_i-\sum_{i=1}^{N}\alpha_i(y_i(w\cdot x_i+b)-1+\xi_i)-\sum_{i=1}^{N}\mu_i\xi_i$$

- **满足KKT**

$$\nabla_w L(w^*,b^*,\xi^*,\alpha^*,\mu^*)=w^*-\sum_{i=1}^{N}\alpha_i^* y_i x_i=0$$

$$\nabla_b L(w^*,b^*,\xi^*,\alpha^*,\mu^*)=-\sum_{i=1}^{N}\alpha_i^* y_i=0$$

$$\nabla_\xi L(w^*,b^*,\xi^*,\alpha^*,\mu^*)=C-\alpha^*-\mu^*=0$$

$$\alpha_i^*(y_i(w^*\cdot x_i+b^*)-1+\xi_i^*)=0$$

$$\mu_i^*\xi_i^*=0$$

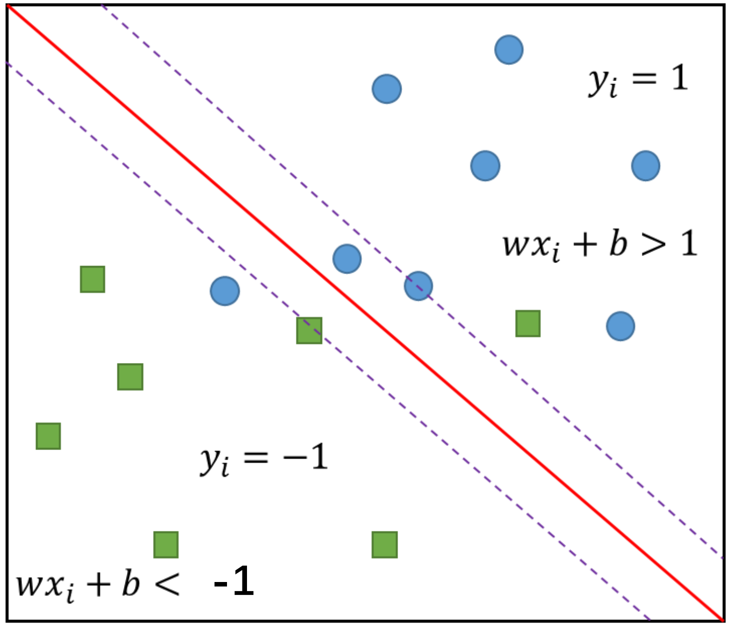
$$w^*=\sum_{i=1}^{N}\alpha_i^* y_i x_i$$

对于分量 $0<\alpha_j<C$

$$y_j\left(w^*x_j+b^*\right)=1$$

$$b^*=y_j-\sum_{i=1}^{N}y_i\alpha_i^*(x_i\cdot x_j)$$

支持向量



$\alpha_i^* < C$
$C - \alpha_i - \mu_i = 0 \rightarrow \mu_i \neq 0$
$\mu_i \xi_i = 0 \rightarrow \xi_i = 0$
$\alpha_i(y_i(wx_i + b) - 1 + \xi_i) = 0$
$y_j(w^* x_j + b^*) = 1$

$\alpha_i^* = C$
$\alpha_i(y_i(wx_i + b) - 1 + \xi_i) = 0$
$y_i(wx_i + b) = 1 - \xi_i$

$0 < \xi_i < 1$
$y_i(wx_i + b)$是正数，分类正确
x_i在间隔边界和超平面之间

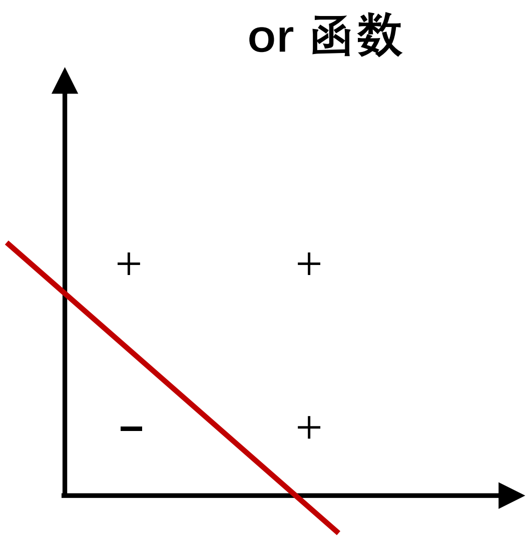
$\xi_i = 1$
$y_i(wx_i + b) = 0$
x_i在超平面上

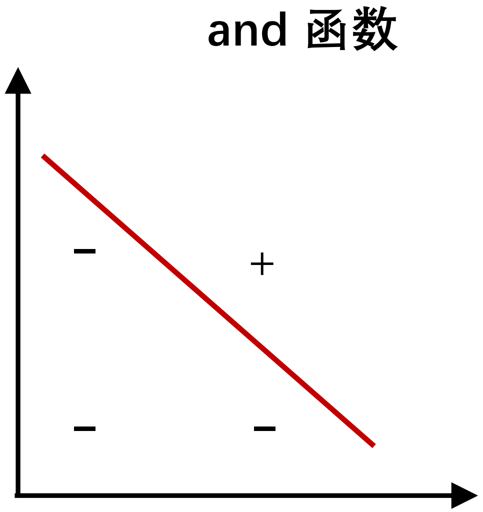
$\xi_i > 1$
$y_i(wx_i + b)$是负数，分类错误
x_i在超平面另外一侧

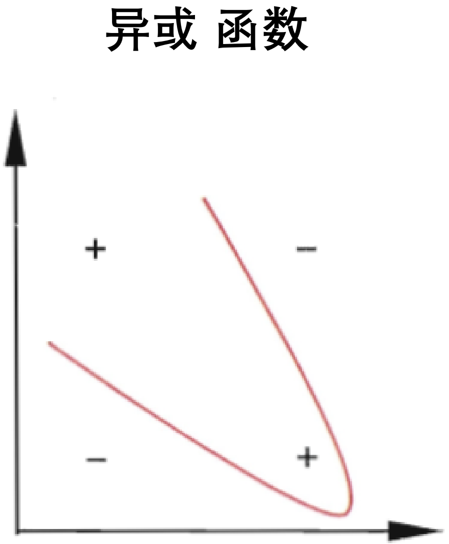
03

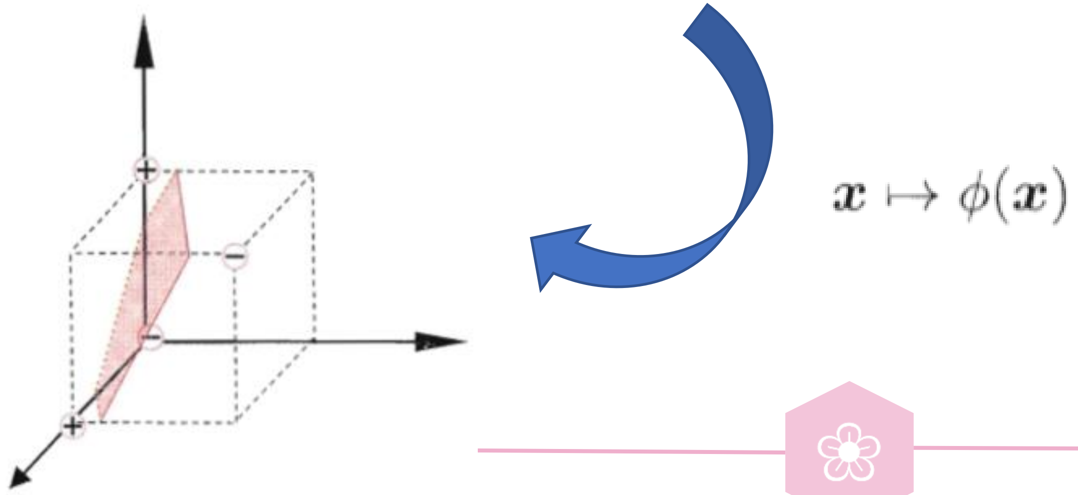
非线性支持向量机

核函数





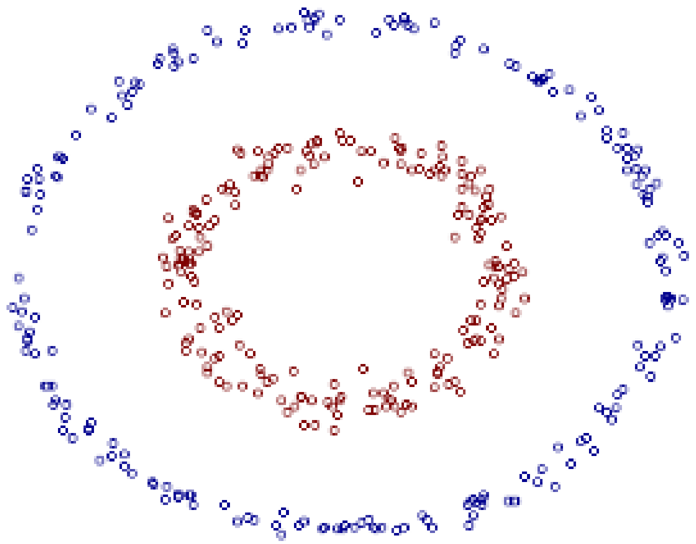


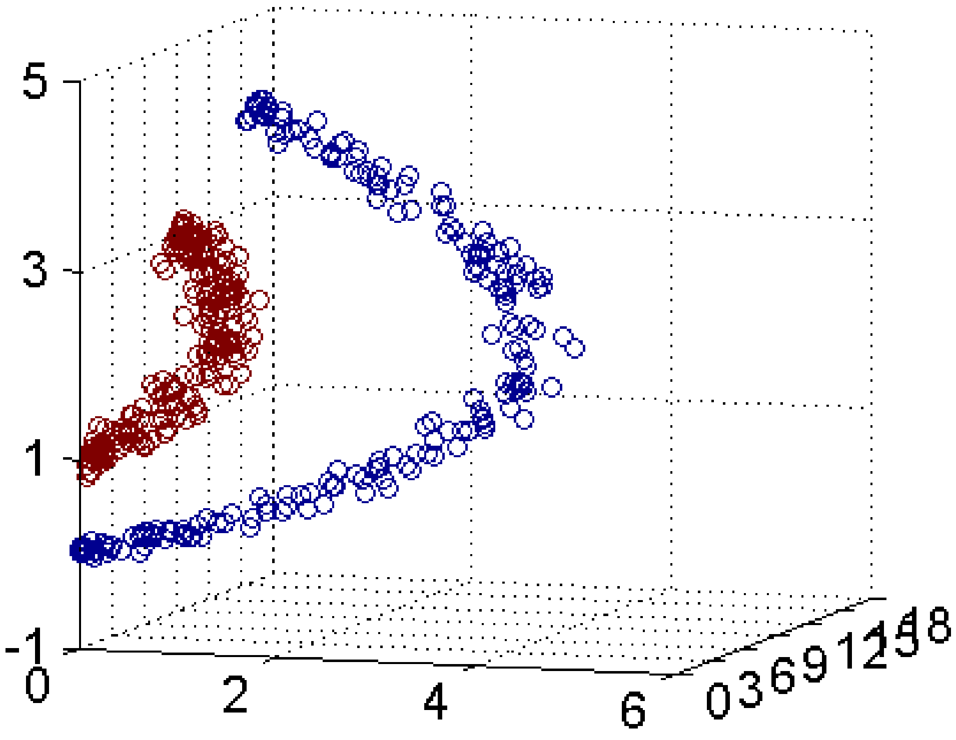


如果原始空间是有限维，即属性有限

那么一定存在一个高位特征空间使样本可分

核函数





核函数

- 定义：设χ是输入空间，又设$\mathcal{H}$为特征空间（希尔伯特空间），如果存在一个χ到$\mathcal{H}$的映射：

$$\phi(x): \chi \rightarrow \mathcal{H}$$

使得对所有$x, z \epsilon \chi$，函数$K(x, z)$满足条件

$$K(x, z) = \phi(x) \cdot \phi(z)$$

则称$K(x, z)$为核函数，$\phi(x)$为映射函数，$\cdot$为内积

核函数

- 例1：假设输入空间是$\mathcal{R}^2$，核函数是$K(x,z)=(x\cdot z)^2$，试找出其相关的特征空间$\mathcal{H}$和映射$\phi(x)\colon \mathcal{R}^2\rightarrow\mathcal{H}$
  - $x=(x^1,x^2)$ $\quad$ $\mathrm{z}=(z^1,z^2)$
  - $K(x,z)=(x\cdot z)^2=\left(x^1z^1+x^2z^2\right)^2=\left(x^1z^1\right)^2+2\mathrm{x}^1\mathrm{z}^1\mathrm{x}^2\mathrm{z}^2+\left(\mathrm{x}^2\mathrm{z}^2\right)^2$
  - 三维：$\phi(x)=\left(\left(x^1\right)^2,\sqrt{2}x^1x^2,\left(x^2\right)^2\right)^T$ $\quad$ $\boxed{\phi(x)\cdot\phi(z)}$

$$\phi(x)=\frac{1}{\sqrt{2}}\left(\left(x^1\right)^2-\left(x^2\right)^2,2x^1x^2,\left(x^1\right)^2+\left(x^2\right)^2\right)^{\mathrm{T}}\quad\boxed{\phi(x)\cdot\phi(z)}$$

核函数应用

- 对偶形式

$$\max_{\boldsymbol{\alpha}} \sum_{i=1}^{m}\alpha_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_i\alpha_j y_i y_j \boxed{\boldsymbol{x}_i^{\mathrm{T}}\boldsymbol{x}_j}$$

- 转换

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

- 重写

$$\max_{\boldsymbol{\alpha}} \sum_{i=1}^{m}\alpha_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_i\alpha_j y_i y_j \boxed{\kappa(\boldsymbol{x}_i,\boldsymbol{x}_j)}$$

核技巧：不显示地定义映射函数，只定义核函数

核函数应用

- 目标函数：$W(\alpha)=\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j K(x_i,x_j)-\sum_{i=1}^{N}\alpha_i$

- 决策函数：$f(x)=sign(\sum_{i=1}^{N}\alpha_i^* y_i \phi(x_i)\cdot\phi(x)+b^*)$

$$=sign(\sum_{i=1}^{N}\alpha_i^* y_i K(x_i,x)+b^*)$$

常用核函数

表 6.1 常用核函数

名称	表达式	参数
线性核	$\kappa(\boldsymbol{x}_i, \boldsymbol{x}_j) = \boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{x}_j$	
多项式核	$\kappa(\boldsymbol{x}_i, \boldsymbol{x}_j) = (\boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{x}_j)^d$	$d \geqslant 1$ 为多项式的次数
高斯核	$\kappa(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp\left(-\frac{\lVert \boldsymbol{x}_i - \boldsymbol{x}_j \rVert^2}{2\sigma^2}\right)$	$\sigma > 0$ 为高斯核的带宽(width)
拉普拉斯核	$\kappa(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp\left(-\frac{\lVert \boldsymbol{x}_i - \boldsymbol{x}_j \rVert}{\sigma}\right)$	$\sigma > 0$
Sigmoid 核	$\kappa(\boldsymbol{x}_i, \boldsymbol{x}_j) = \tanh(\beta \boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{x}_j + \theta)$	tanh 为双曲正切函数, $\beta > 0$, $\theta < 0$

非线性支持向量机

- $\min_{\alpha} \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j y_i y_j K(x_i, x_j) - \sum_{i=1}^{N}\alpha_i$
- s.t. $\sum_{i=1}^{N}\alpha_i y_i = 0$
- $0 \le \alpha_i \le C$

$$w_i = \sum_{i=1}^{N}\alpha_i y_i \phi(x_i) \qquad b^* = y_j - \sum_{i=1}^{N}\alpha_i^* y_i K(x_i, x_j)$$

实验作业

- sklearn是一个Python第三方提供的非常强力的机器学习库
- 基于数据集iris, 构建逻辑回归模型和支持向量机

Sepal.Leng	Sepal.Widt	Petal.Leng	Petal.Widtl	Species
5.1	3.5	1.4	0.2	setosa
4.9	3	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5	3.6	1.4	0.2	setosa

  - 150个样本，4个特征，3个类别
  - from sklearn.datasets import load_iris
- 提交实验报告pdf即可，包含实验设置（数据、参数等）、实验分析、结果展示、必要设计模块解释 （提交截止时间3.30，pdf上传到学习通对应题目位置即可）