日期: /

一、求最大间隔分离超平面

1.1 求解方法

(1) 设已知训练集 $\mathbf{D}=\{(\boldsymbol{x}_1,y_1),(\boldsymbol{x}_2,y_2),\ldots,(\boldsymbol{x}_m,y_m)\}$, 其中 $\boldsymbol{x}_i\in\mathbb{R}^d$, $y_i\in\mathcal{Y}=\{1,-1\}$, $i=1,2,\ldots,m$

(2) 构造并求解凸二次规划

$$\min_{\boldsymbol{w},b}\|\boldsymbol{w}\|_2^2 \quad \text{s.t.} \quad y_i(\boldsymbol{w}^T\boldsymbol{x}_i+b)\geq 1,\ i=1,2,\ldots,m$$

求得解 $\boldsymbol{w}^*,b^*$

(3) 构造划分超平面 $(\boldsymbol{w}^*)^T\boldsymbol{x}+b^*=0$, 由此得到决策函数

$$f(\boldsymbol{x})=\operatorname{sgn}\left((\boldsymbol{w}^*)^T\boldsymbol{x}+b^*\right)$$

$$\operatorname{sgn}(a)=\begin{cases}1, & a\geq 0,\\ -1, & a<0.\end{cases}$$

1.2 例题

- 已知正例点是 $x_1=[1,2]^\top, x_2=[2,3]^\top, x_3=[3,3]^\top$， 负例点是 $x_4=[2,1]^\top, x_5=[3,2]^\top$，试求最大间隔分离超平面和分类决策函数，并在图上画出分离超平面、间隔边界及支持向量 （答案： $w_1=-1$, $w_2=2$, $b=-2$ ）

解：构造凸二次规划问题：

$$\min \|w\|_2^2=\min(w_1^2+w_2^2)$$

$$\text{s.t.}\begin{cases}1\cdot(w_1+2w_2+b)\geq 1\\ 1\cdot(2w_1+3w_2+b)\geq 1\\ 1\cdot(3w_1+3w_2+b)\geq 1\\ -1\cdot(2w_1+w_2+b)\geq 1\\ -1\cdot(3w_1+2w_2+b)\geq 1\end{cases}$$

求解得到 $w_1=-1$ $w_2=2$, $b=-2$

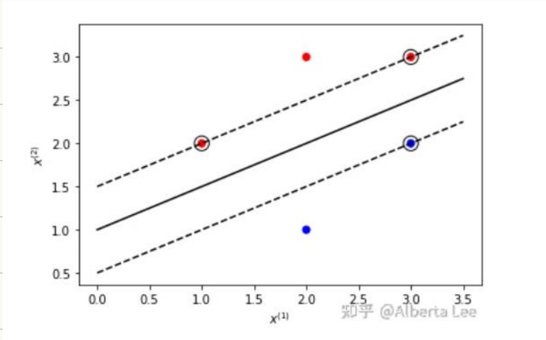
超平面为: $-x^{(1)}+2x^{(2)}-2=0$

分类决策函数: $f(x)=\mathrm{sgn}(-x^{(1)}+2x^{(2)}-2)$

$$\mathrm{sgn}(a)=\begin{cases}1, & a\geq 0\\ -1, & a<0\end{cases}$$

支持向量 $x_1=[1,2]^T$ $x_3=[3,3]^T$

$x_5=[3,2]^T$



二、求对偶问题

线性支持向量机还可以定义为以下形式:

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

$$\text{s.t.} \quad y_i(w\cdot x_i+b)\geqslant 1-\xi_i, i=1,2,\cdots,N \quad \xi_i\geqslant 0, i=1,2,\cdots,N$$

试求其对偶形式。

解答:
根据支持向量机的对偶算法，得到对偶形式，由于不能消去变量 ξ_i 的部分，所以拉格朗日因子也包含 β_i 。
拉格朗日函数为:

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

分别求 w,b,ξ 的偏导数:

$$\begin{cases}\nabla_w L=w-\sum\limits_{i=1}^N\alpha_i y_i x_i=0\\ \nabla_b L=-\sum\limits_{i=1}^N\alpha_i y_i=0\\ \nabla_\xi L=2C\xi_i-\alpha_i-\beta_i=0\end{cases}$$

化简可得:

$$\begin{cases}w=\sum\limits_{i=1}^N\alpha_i y_i x_i=0\\ \sum\limits_{i=1}^N\alpha_i y_i=0\\ 2C\xi_i-\alpha_i-\beta_i=0\end{cases}$$

可解得: $L=-\frac{1}{2}\sum_{i=1}^N\sum_{j=1}^N\alpha_i\alpha_j y_i y_j(x_i\cdot x_j)+\sum_{i=1}^N\alpha_i-\frac{1}{4C}\sum_{i=1}^N(\alpha_i+\beta_i)^2$

习题7.4

日期: /

支持向量机-例题

$$\min_{w} \ \frac{1}{2}\| w \|^2$$

s.t. $y_i(w^T x_i + b) \geq 1, i = 1, \dots,$

正例 $\boldsymbol{x_1} = \binom{3}{3}, \boldsymbol{x_2} = \binom{4}{3}, \boldsymbol{y_1} = \boldsymbol{y_2} = \mathbf{1}$; 负例 $\boldsymbol{x_3} = \binom{1}{1}, \boldsymbol{y_3} = \mathbf{-1}$

$$\arg\min_{w,b} \max_{\lambda} L(w,b,\lambda) = \frac{1}{2}w^T w + \sum \lambda_i (1 - y_i(w^T x_i + b))] = \arg\max_{\lambda} \min_{w,b}$$

$$0 = \frac{\partial L}{\partial w} = w - \sum \lambda_i y_i x_i \Rightarrow w^* = \sum \lambda_i y_i x_i$$ 代入L得

$$0 = \frac{\partial L}{\partial b} = \sum \lambda_i y_i \Rightarrow \sum \lambda_i y_i = 0$$

$$\arg\max_{\lambda} L = \sum \lambda_i - \frac{1}{2}\sum_i \sum_j \lambda_i \lambda_j y_i y_j x_i x_j^T$$

$$= \arg\min_{\lambda} \frac{1}{2}\sum_i \sum_j \lambda_i \lambda_j y_i y_j x_i x_j^T - \sum \lambda_i .$$

KKT:

$\lambda_i \geq 0$

$1 - y_i(w^T x_i + b) \leq 0$

$\lambda_i[1 - y_i(w^T x_i + b)] = 0$

代入x_1, x_2, x_3得 (3,3) (4,3) (1,1)

$$\arg\min_{\lambda} \frac{1}{2}[18\lambda_1^2 + 21\lambda_1\lambda_2 - 6\lambda_1\lambda_3 + 21\lambda_2\lambda_1 + 25\lambda_2^2 - 7\lambda_2\lambda_3 - 6\lambda_3\lambda_1 - 7\lambda_3\lambda_2 + 2\lambda_3^2]$$

$$\xrightarrow{\lambda_3 = \lambda_1 + \lambda_2} \arg\min_{\lambda} 4\lambda_1^2 + \frac{13}{2}\lambda_2^2 + 10\lambda_1\lambda_2 - 2\lambda_1 - 2\lambda_2$$

$$0 = \frac{\partial L}{\partial \lambda_1} = 8\lambda_1 + 10\lambda_2 - 2$$
$$0 = \frac{\partial L}{\partial \lambda_2} = 13\lambda_2 + 10\lambda_1 - 2$$
$$\Rightarrow \begin{cases} \lambda_1 = 1.5 \\ \lambda_2 = -1 \ (\text{不满足KKT条件}) \end{cases}$$ 即 $\lambda_1 = 0$ 或 $\lambda_2 = 0$

$$\begin{cases} \lambda_1 = 0 \\ \lambda_2 = \frac{2}{13} \end{cases} \Rightarrow L = -0.1538, \quad \begin{cases} \lambda_2 = 0 \\ \lambda_1 = \frac{1}{4} \end{cases} \Rightarrow L = -0.25 \quad \therefore \text{取} \begin{cases} \lambda_1 = \frac{1}{4} \\ \lambda_2 = 0 \\ \lambda_3 = \frac{1}{4} \end{cases}$$

得 $w = \sum \lambda_i y_i x_i = \frac{1}{4}\times(3,3) + \frac{1}{4}\times(-1)\times(1,1) = (\frac{1}{2}, \frac{1}{2})$

代入点(3,3)得 $y(w^T x + b) = 1 \Rightarrow b = -2.$

即 $\frac{1}{2}x^{(1)} + \frac{1}{2}x^{(2)} - 2 = 0.$ 支持向量为 $x_1(3,3), x_3(1,1)$

日期: /

例 7.2 训练数据与例 7.1 相同。如图 7.4 所示，正例点是 $x_1=(3,3)^{\mathrm{T}}$，$x_2=(4,3)^{\mathrm{T}}$，负例点是 $x_3=(1,1)^{\mathrm{T}}$，试用算法 7.2 求线性可分支持向量机。

解 根据所给数据，对偶问题是

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

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

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

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

$$\begin{cases}13\alpha_2+10\alpha_1-2=0\\ 8\alpha_1+10\alpha_2-2=0\end{cases}$$

解这一最优化问题。将 $\alpha_3=\alpha_1+\alpha_2$ 代入目标函数并记为

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

对 α_1,α_2 求偏导数并令其为 0，易知 $s(\alpha_1,\alpha_2)$ 在点 $\left(\frac{3}{2},-1\right)^{\mathrm{T}}$ 取极值，但该点不满足约束条件 $\alpha_2\geqslant 0$，所以最小值应在边界上达到。

当 $\alpha_1=0$ 时，最小值 $s\left(0,\frac{2}{13}\right)=-\frac{2}{13}$；当 $\alpha_2=0$ 时，最小值 $s\left(\frac{1}{4},0\right)=-\frac{1}{4}$。于是 $s(\alpha_1,\alpha_2)$ 在 $\alpha_1=\frac{1}{4},\alpha_2=0$ 达到最小，此时 $\alpha_3=\alpha_1+\alpha_2=\frac{1}{4}$。

这样，$\alpha_1^*=\alpha_3^*=\frac{1}{4}$ 对应的实例点 x_1,x_3 是支持向量。根据式 (7.25) 和式 (7.26) 计算得

$$w_1^*=w_2^*=\frac{1}{2}$$

$$b^*=-2$$

$w=\sum \alpha_i x_i y_i$

α	x		y
$\frac{1}{4}$	3	3	1
0	4	3	1
$\frac{1}{4}$	1	1	-1

分离超平面为

$$\frac{1}{2}x^{(1)}+\frac{1}{2}x^{(2)}-2=0$$

分类决策函数为

$$f(x)=\operatorname{sign}\left(\frac{1}{2}x^{(1)}+\frac{1}{2}x^{(2)}-2\right)$$ ■

$y_i(w^{\mathrm{T}}x+b)=1$

$1\times\left(\frac{3}{2}+\frac{3}{2}+b\right)=1$

$b=-2$

$\Leftarrow \begin{cases}w_1=\frac{3}{4}-\frac{1}{4}=\frac{1}{2}\\ w_2=\frac{3}{4}-\frac{1}{4}=\frac{1}{2}\end{cases}$

日期: /

例 7.1 数据与例 2.1 相同。已知一个如图 7.4 所示的训练数据集，其正例点是 $x_1=(3,3)^{\mathrm{T}}$，$x_2=(4,3)^{\mathrm{T}}$，负例点是 $x_3=(1,1)^{\mathrm{T}}$，试求最大间隔分离超平面。

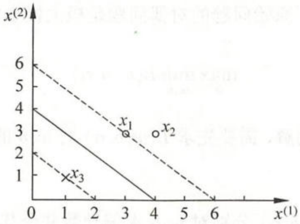


图 7.4 间隔最大分离超平面示例

解 按照算法 7.1，根据训练数据集构造约束最优化问题：

$$
\begin{aligned}
\min_{w,b} \quad & \frac{1}{2}(w_1^2+w_2^2) \\
\text{s.t.} \quad & 3w_1+3w_2+b \geqslant 1 \\
& 4w_1+3w_2+b \geqslant 1 \\
& -w_1-w_2-b \geqslant 1
\end{aligned}
$$

求得此最优化问题的解 $w_1=w_2=\dfrac{1}{2}$，$b=-2$。于是最大间隔分离超平面为

$$\frac{1}{2}x^{(1)}+\frac{1}{2}x^{(2)}-2=0$$

其中，$x_1=(3,3)^{\mathrm{T}}$ 与 $x_3=(1,1)^{\mathrm{T}}$ 为支持向量。

对偶问题: $\min\left(\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j x_i x_j y_i y_j-\sum_{i=1}^{n}\alpha_i\right)$

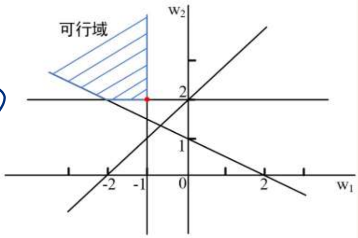
日期: /

- 已知正例点是 $x_1=[1,2]^\top, x_2=[2,3]^\top, x_3=[3,3]^\top$， 负例点是 $x_4=[2,1]^\top, x_5=[3,2]^\top$，利用线性可分支持向量机的原始形式求最大间隔分离超平面和分类决策函数，并在图上画出分离超平面、间隔边界及支持向量（答案： $\mathrm{w}_1=-1$，$\mathrm{w}_2=2$， $b=-2$ ）。这里 $\mathbf{w}=(\mathrm{w}_1,\mathrm{w}_2)^\mathrm{T}$，$\boldsymbol{x}=(x^{(1)},x^{(2)})^\mathrm{T}$。
- **解**：SVM原始问题： $\min\limits_{\mathbf{w},b}\|\mathbf{w}\|^2=\min\limits_{\mathbf{w},b}(\mathrm{w}_1^2+\mathrm{w}_2^2) \quad s.t.$

$$\begin{cases} w_1+2w_2+b\geq 1 & ① \\ 2w_1+3w_2+b\geq 1 & ② \\ 3w_1+3w_2+b\geq 1 & ③ \\ -2w_1-w_2-b\geq 1 & ④ \\ -3w_1-2w_2-b\geq 1 & ⑤ \end{cases}$$

上述不等式约束可进一步化简为

$$\begin{cases} -w_1+w_2\geq 2 \\ -2w_1\geq 2 \\ 2w_2\geq 2 \\ w_1+2w_2\geq 2 \\ w_2\geq 2 \end{cases}$$



不等式约束的解的可行域

由上以及结合右图可知为使 $\mathrm{w}_1^2+\mathrm{w}_2^2$ 最小，有 $(\mathrm{w}_1,\mathrm{w}_2)=(-1,2)$。将 $\mathrm{w}_1,\mathrm{w}_2$ 代入约束条件得 $b=-2$

$\begin{cases} ① \\ ② \\ ③ \end{cases} + \begin{cases} ④ \\ ⑤ \end{cases} \Rightarrow$ 画图

日期:　　　/

日期:　　　/