

Homework 8

PB17000297 罗晏宸

May 3 2020

1

试证明对于不含冲突数据（即特征向量完全相同但标记不同）的训练集，必存在一个与训练集一致（即训练误差为 0）的决策树。

解

证明. 决策树中叶节点生成于样本属于同一类或者样本在当前属性集上取值相同时，对于不含冲突数据的训练集，不会出现在当前属性集上取值相同但不同标记的样本，因此决策树的叶节点均为同类样本，即从根节点到每个叶节点的路径都代表一个向量，这个决策树与训练集是一致的。 \square

2

已知正例点 $\mathbf{x}_1 = (1, 2)^T$, $\mathbf{x}_2 = (2, 3)^T$, $\mathbf{x}_3 = (3, 3)^T$, 负例点 $\mathbf{x}_4 = (2, 1)^T$, $\mathbf{x}_5 = (3, 2)^T$, 试求 Hard Margin SVM 的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、间隔边界以及支持向量。

解 按照训练集数据构造约束最优化问题

$$\min_{\omega, b} \frac{1}{2}(\omega_1^2 + \omega_2^2) \quad (1.1)$$

$$\text{s. t.} \begin{cases} \omega_1 + 2\omega_2 + b \geq 1 \\ 2\omega_1 + 3\omega_2 + b \geq 1 \\ 3\omega_1 + 3\omega_2 + b \geq 1 \\ -2\omega_1 - \omega_2 - b \geq 1 \\ -3\omega_1 - 2\omega_2 - b \geq 1 \end{cases} \quad (1.2)$$

Lagrange 函数

$$L(\omega, b, \alpha) = \frac{1}{2}\|\omega\|^2 - \sum_{i=1}^5 \alpha_i y_i (\omega \cdot \mathbf{x}_i + b) + \sum_{i=1}^5 \alpha_i$$

其中, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_5)^T$ 为 Lagrange 乘子向量

根据 Lagrange 对称性, 原问题的对偶问题是极大极小问题

$$\max_{\alpha} \min_{\omega, b} L(\omega, b, \alpha)$$

(1) 求 $\min_{\omega, b} L(\omega, b, \alpha)$

将 Lagrange 函数 $L(\omega, b, \alpha)$ 分别对 ω 、 b 求偏导数并令其等于 0

$$\nabla_{\omega} L(\omega, b, \alpha) = \omega - \sum_{i=1}^5 \alpha_i y_i \mathbf{x}_i = 0$$

$$\nabla_b L(\omega, b, \alpha) = \sum_{i=1}^5 \alpha_i y_i = 0$$

得

$$\begin{aligned} \omega &= \sum_{i=1}^5 \alpha_i y_i \mathbf{x}_i = (\alpha_1 + 2\alpha_2 + 3\alpha_3 - 2\alpha_4 - 3\alpha_5, 2\alpha_1 + 3\alpha_2 + 3\alpha_3 - \alpha_4 - 2\alpha_5)^T \\ \sum_{i=1}^5 \alpha_i y_i &= \alpha_1 + \alpha_2 + \alpha_3 - \alpha_4 - \alpha_5 = 0 \end{aligned}$$

代入 Lagrange 函数, 得到

$$\begin{aligned}
& \min_{\omega, b} L(\omega, b, \alpha) \\
&= -\frac{1}{2} \sum_{i=1}^5 \sum_{j=1}^5 \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_j) + \sum_{i=1}^5 \alpha_i \\
&= \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 + \alpha_5 + \\
& \quad \frac{1}{2} (-5\alpha_1^2 - 16\alpha_2\alpha_1 - 18\alpha_3\alpha_1 + 8\alpha_4\alpha_1 + 14\alpha_5\alpha_1 \\
& \quad - 13\alpha_2^2 - 18\alpha_3^2 - 5\alpha_4^2 - 13\alpha_5^2 - 30\alpha_2\alpha_3 + 14\alpha_2\alpha_4 \\
& \quad + 18\alpha_3\alpha_4 + 24\alpha_2\alpha_5 + 30\alpha_3\alpha_5 - 16\alpha_4\alpha_5)
\end{aligned}$$

(2) 求 $\min_{\omega, b} L(\omega, b, \alpha)$ 对 α 的极大

将目标函数转换为极小, 得到与之等价的对偶最优化问题

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^5 \sum_{j=1}^5 \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_j) - \sum_{i=1}^5 \alpha_i \quad (2.1)$$

$$\text{s. t. } \begin{cases} \alpha_1 + \alpha_2 + \alpha_3 - \alpha_4 - \alpha_5 = 0 \\ \alpha_i \geq 0, \quad i = 1, 2, 3, 4, 5 \end{cases} \quad (2.2)$$

解得 $\alpha^* = (-2, 5, -3, 0, 0)^T$ 。考虑原始最优化问题(1.1)~(1.2)和对偶最优化问题(2.1)~(2.2), 由 $\alpha_2^* = 5 > 0$ 原始最优化问题有解

$$\begin{aligned}
\omega^* &= \sum_{i=1}^5 \alpha_i^* y_i \mathbf{x}_i = (-1, 2)^T \\
b^* &= 1 - \sum_{i=1}^5 \alpha_i^* y_i (\mathbf{x}_i \cdot \mathbf{x}_2) = -2
\end{aligned}$$

于是最大间隔分离超平面为

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

其中 $\mathbf{x}_1 = (1, 2)^T$, $\mathbf{x}_3 = (3, 3)^T$ 与 $\mathbf{x}_5 = (3, 2)^T$ 为支持向量, 如图1所示。分类决策函数为

$$f(x) = \text{sign} \left\{ \sum_{i=1}^5 \alpha_i^* y_i (x \cdot \mathbf{x}_i) + b^* \right\} = \text{sign} \{-\mathbf{x}^{(1)} + 2\mathbf{x}^{(2)} - 2\}$$

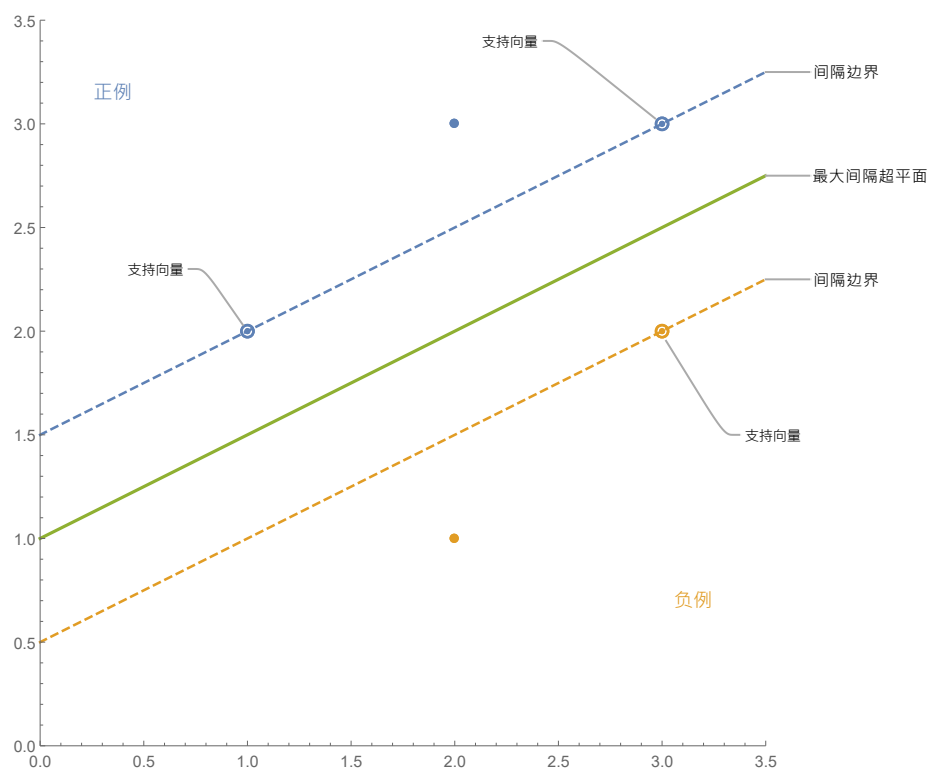


图 1: 最大间隔分离器、间隔边界与支持向量