

# 机器学习基础作业 2

2025 年 4 月 15 日

**问题 1.** 证明讲义中的如下结论: SMO 算法中未经剪辑时  $\alpha_2$  的最优解为

$$\alpha_2 = \alpha_2^* + y_2 \frac{(y_2 - f(x_2)) - (y_1 - f(x_1))}{\eta}$$

其中  $\eta = K_{11} + K_{22} - 2K_{12}$ .

证明. 由求导推得的结论,

$$\alpha_2 = \frac{s(K_{11} - K_{12})\gamma + y_2(v_1 - v_2) - s + 1}{\eta}$$

又

$$f(x) = \sum_{i=1}^N \alpha_i^* y_i x_i \cdot x + b^*, \quad v_i = \sum_{j=3}^N \alpha_j^* y_j x_j \cdot x_i + b^*, \quad s = y_1 y_2, \quad \gamma = \alpha_1 + s \alpha_2$$

代入  $f(x_1)$  和  $f(x_2)$ , 并相减得到

$$\begin{aligned} f(x_1) - f(x_2) &= \sum_{i=1}^N \alpha_i^* y_i x_i (x_1 - x_2) \\ &= \alpha_1^* y_1 x_1 \cdot (x_1 - x_2) + \alpha_2^* y_2 x_2 \cdot (x_1 - x_2) + \sum_{j=3}^N \alpha_j^* y_j x_j \cdot (x_1 - x_2) \\ &= \alpha_1^* y_1 (K_{11} - K_{12}) + \alpha_2^* y_2 (K_{12} - K_{22}) + (v_1 - v_2) \\ &= (\gamma - s \alpha_2^*) s y_2 (K_{11} - K_{12}) + \alpha_2^* y_2 (K_{12} - K_{22}) + (v_1 - v_2) \\ &= s y_2 \gamma (K_{11} - K_{12}) - \alpha_2^* y_2 \eta + (v_1 - v_2) \end{aligned}$$

把  $v_1 - v_2$  代回  $\alpha_2$  的表达式得到

$$\begin{aligned} \alpha_2 &= \frac{s(K_{11} - K_{12})\gamma + y_2(f(x_1) - f(x_2)) - s y_2 \gamma (K_{11} - K_{12}) + \alpha_2^* y_2 \eta - s + 1}{\eta} \\ &= \alpha_2^* + y_2 \frac{(y_2 - f(x_2)) - (y_1 - f(x_1))}{\eta} \end{aligned}$$

□

**问题 2.** 考虑样本的重要性, 给定训练数据集  $D = \{(x_i, y_i, p_i)\}_{i=1}^N$ , 其中  $x_i \in \mathcal{X} \subseteq \mathbb{R}^n$ ,  $y_i \in \mathcal{Y} = \{-1, 1\}$ ,  $p_i \in (0, 1)$  为样本  $x_i$  的权重,  $i = 1, 2, \dots, N$ . 在考虑基于  $D$  构建线性支持向量机模型时, 在惩罚项中需要以  $\sum_{i=1}^N p_i \xi_i$  代替  $\sum_{i=1}^N \xi_i$ .

1. 基于  $D$  给出线性支持向量机的原始最优化问题.
2. 利用拉格朗日乘子法给出其对偶问题.

证明. 1. 直接按要求替换, 得到

$$\begin{aligned} \min_{w, b, \xi} \quad & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N p_i \xi_i \\ \text{s.t.} \quad & y_i (w \cdot x_i + b) \geq 1 - \xi_i, \quad i = 1, 2, \dots, N \end{aligned}$$

2. 引入 Lagrange 乘子  $\alpha_i, \beta_i \geq 0$ , 得到

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

对  $w, b, \xi$  偏导为 0, 得到

$$\begin{aligned} w &= \sum_{i=1}^n \alpha_i y_i x_i \\ 0 &= \sum_{i=1}^n \alpha_i y_i \\ \beta_i &= C p_i - \alpha_i \end{aligned}$$

代入  $L(w, b, \xi, \alpha, \beta)$ , 计算得

$$\begin{aligned} L(w, b, \xi, \alpha, \beta) &= \frac{1}{2} \|w\|^2 + (C p_i - \alpha_i - \beta_i) \xi_i + \sum_{i=1}^n \alpha_i - \left( \sum_{i=1}^n \alpha_i y_i x_i \right) \cdot w - b \sum_{i=1}^n \alpha_i y_i \\ &= -\frac{1}{2} \|w\|^2 + \sum_{i=1}^n \alpha_i = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j x_i \cdot x_j \end{aligned}$$

因而对偶问题为

$$\begin{aligned} \max_{\alpha} \quad & \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j x_i \cdot x_j \\ \text{s.t.} \quad & \sum_{i=1}^n \alpha_i y_i = 0, \quad 0 \leq \alpha_i \leq C p_i, \quad i = 1, 2, \dots, N \end{aligned}$$

□

**问题 3.** 证明异或问题非线性可分.

证明. 设异或问题是线性可分的, 则存在  $w \in \mathbb{R}^2, b \in \mathbb{R}$  使得超平面  $w \cdot x + b = 0$  满足

$$\begin{cases} w \cdot (0, 0) + b > 0 \\ w \cdot (0, 1) + b < 0 \\ w \cdot (1, 0) + b < 0 \\ w \cdot (1, 1) + b > 0 \end{cases}$$

将式 1 和式 4 相加, 将式 2 和式 3 相加, 得到  $0 > w \cdot (1, 1) + b > 0$ , 矛盾. 因此异或问题非线性可分. □