机器学习第三周作业

樊泽羲 2200010816

2023 年 9 月 30 日

题目 1. 定义如下支持向量机：

$$
\begin{aligned}
&min_{w,b,\xi}\ \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}\xi_i^2 \\
&\quad s.t.\ y_i(w\cdot x_i + b) \geq 1-\xi_i, i=1,2,\cdots,N \\
&\qquad \xi_i \geq 0, i=1,2,\cdots,N
\end{aligned}
$$

试求其对偶形式

解答.

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

原问题化为无约束优化问题：$min_{w,b,\xi}max_{\alpha_i,\beta_i\geq 0}L(w,b,\xi,\alpha,\beta)$

对偶问题: $max_{\alpha_i,\beta_i\geq 0}min_{w,b,\xi}L(w,b,\xi,\alpha,\beta)$

对 L 的各个变量求导：

$$\frac{\partial L}{\partial w}=w-\sum_{i=1}^{N}\alpha_iy_ix_i=0\Longrightarrow w=\sum_{i=1}^{N}\alpha_iy_ix_i$$

$$\frac{\partial L}{\partial b}=-\sum_{i=1}^{N}\alpha_iy_i=0\Longrightarrow \sum_{i=1}^{N}\alpha_iy_i=0$$

$$\frac{\partial L}{\partial \xi_i}=2C\xi_i-\alpha_i-\beta_i=0\Longrightarrow \xi_i=\frac{\alpha_i+\beta_i}{2C}$$

代入整理可得：

$$min_{w,b,\xi}L=-\frac{1}{2}\sum_{i,j}\alpha_i\alpha_jx_ix_jy_iy_j-\frac{1}{4C}\sum_i(\alpha_i+\beta_i)^2+\sum_i\alpha_i$$

从而最终的对偶问题为：

$$max_{\alpha_i,\beta_i\geq 0}-\frac{1}{2}\sum_{i,j}\alpha_i\alpha_jx_ix_jy_iy_j-\frac{1}{4C}\sum_i(\alpha_i+\beta_i)^2+\sum_i\alpha_i$$

题目 2. 证明讲义中的公式 (35)

解答. 设当前选择的需要更新的拉格朗日乘子为 α_1 和 α_2, 其他拉格朗日乘子 $\alpha_i (3 \le i \le N)$ 在本轮参数更新中保持不变. 由松弛互补条件 $\sum_{i=1}^{N} \alpha_i y_i = 0$ 可知

$$\alpha_1 y_1 + \alpha_2 y_2 + \sum_{i=3}^{N} \alpha_i y_i = 0.$$

上式两边乘以 y_1 得到

$$\begin{aligned} &\alpha_1 + \alpha_2 y_1 y_2 + \textstyle\sum_{i=3}^{N} \alpha_i y_1 y_i = 0. \\ &\Longrightarrow \alpha_1 + s\alpha_2 = \gamma. \end{aligned}$$

其中$\gamma = -\sum_{i=3}^{N} \alpha_i y_1 y_i, s = y_1 y_2 \in \{-1, +1\}$

令 $K_{ij} = x_i \cdot x_j$ 且 $v_i = \sum_{j=3}^{N} \alpha_j y_j K_{ij} (i = 1, 2)$, 则将对偶问题转化为下面关于拉格朗日乘子 α_1 和 α_2 的优化问题:

$$\begin{aligned} \max\nolimits_{\alpha_1, \alpha_2} \quad & W_1(\alpha_1, \alpha_2), \\ \text{s.t.} \quad & 0 \le \alpha_1, \alpha_2 \le C, \\ & \alpha_1 + s\alpha_2 = \gamma. \end{aligned}$$

其中

$$\begin{aligned} W_1(\alpha_1, \alpha_2) = & \alpha_1 + \alpha_2 - \frac{1}{2} K_{11} \alpha_1^2 - \frac{1}{2} K_{22} \alpha_2^2 \\ & - s K_{12} \alpha_1 \alpha_2 - y_1 \alpha_1 v_1 - y_2 \alpha_2 v_2 \end{aligned}$$

由约束 $\alpha_1 + s\alpha_2 = \gamma$ 可得

$$\alpha_1 = \gamma - s\alpha_2$$

将其代入 $W_1(\alpha_1, \alpha_2)$, 并令

$$W_2(\alpha_2) = W_1(\gamma - s\alpha_2, \alpha_2).$$

令 $W_2(\alpha_2)$ 对 α_2 的导数为 0 , 可以得到

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

其中

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

又有：

$$\begin{aligned}
\alpha_2 &= \frac{s(K_{11} - K_{22})\gamma + y_2(v_1 - v_2) - s + 1}{\eta} \\
&= \frac{y_1y_2(K_{11} - K_{22})\gamma + y_2(f(x_1) - f(x_2)) + \alpha_2^\star\eta - y_1y_2\gamma(K_{11} - K_{22}) - s + 1}{\eta} \\
&= \frac{y_2f(x_1) - y_2f(x_2) - y_1y_2 + y_2^2}{\eta} + \alpha_2^\star \\
&= y_2 \cdot \frac{(f(x_1) - y_1) - (f(x_2) - y_2)}{\eta} + \alpha_2^\star \\
&= y_2 \cdot \frac{(y_2 - f(x_2)) - (y_1 - f(x_1))}{\eta} + \alpha_2^\star
\end{aligned}$$

这样就得到了未剪辑时 α_2 的更新公式