

6.4 软间隔与正则化.

1. 硬间隔: 所有样本都必须划分正确

2. 软间隔: 允许某些样本不满足. $y_i(w^T x_i + b) \geq 1$

最大化间隔同时, 不满足约束样本在边界附近.

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \ell_{\eta_i}(y_i(w^T x_i + b) - 1) \rightarrow$$

$$\ell_{\eta_i}(z) = \begin{cases} 1 & z \leq 0 \\ 0 & \text{else} \end{cases}$$

6.3

3. 软间隔支持向量机

$$\begin{cases} y_i(w^T x_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0, \quad i=1, 2, \dots, m \\ \min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \xi_i \end{cases}$$

$$L(w, b, \alpha, \xi, \mu) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^m \xi_i$$

$$+ \sum_{i=1}^m \alpha_i (1 - \xi_i - y_i(w^T x_i + b)) - \sum_{i=1}^m \mu_i \xi_i$$

4. 通式

软间隔SVM的目标函数形式同硬间隔类似, 唯一不同的是 α_i 的范围.

$$0 \leq \alpha_i \leq C$$

$$\min_f \Omega(f) + C \sum_{i=1}^m \lambda(f(x_i), y_i)$$

↑
结构风险

描述模型的某些性质

↑
正则化程度

描述模型与训练数据的契合程度

