机器学习复习3

2022年7月6日 星期三 22:20

支持向量机 SVM（一）

- Logistic & SVM

$h_\theta(x) = g(\theta^T x) = \frac{1}{1+e^{-\theta^T x}}$

$g(z) = \frac{1}{1+e^{-z}}$

$P(y=1|x;\theta) = h_\theta(x)$

$P(y=0|x;\theta) = 1-h_\theta(x)$

Show in plot

- function margin & geometric margin（函数间隔与几何间隔）

给定训练样本，其中 x 为 features & labels

the distance: $\gamma^{(i)} = y^{(i)}\left(\left(\frac{w}{||w||}\right)^T x^{(i)} + \frac{b}{||w||}\right)$

- 最优间隔 classfier.

$$\begin{cases} \max_{\gamma,w,b} \frac{\hat{\gamma}}{||w||} & s.t.\ y^{(i)}(w^T x^{(i)}+b) \geq \hat{\gamma} \\ \min_{\gamma,w,b} \frac{1}{2}||w||^2 & s.t.\ y^{(i)}(w^T x^{(i)}+b) \geq 1,\ i=1,\cdots,m. \end{cases}$$

- 拉格朗日对偶（Lagrange duality）

$\mathcal{L}(w,\beta) = f(w) + \sum_{i=1}^{l} \beta_i h_i(w)$

eg: $\min_w f(w)$

s.t. $g_i(w) \leq 0$

$h_i(w) = 0$

化为 $\mathcal{L}(w,\alpha,\beta) = f(w) + \sum_{i=1}^{k} \alpha_i g_i(w) + \sum_{i=1}^{l} \beta_i h_i(w)$

Def 一个新函数: $\theta_P(w) = \max_{\alpha,\beta:\alpha_i \geq 0} \mathcal{L}(w,\alpha,\beta)$

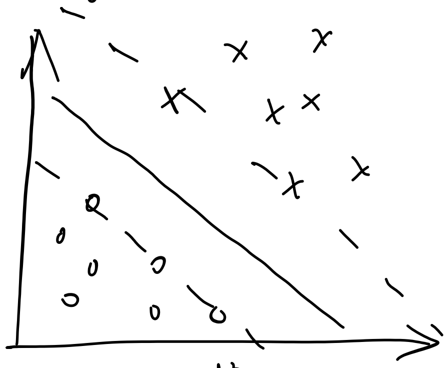
$$\Rightarrow \theta_P(w) = \begin{cases} f(w) & \text{if } w \text{ satisfies primal constraints} \\ \infty & \text{otherwise.} \end{cases}$$

...

- 最优间隔分类器

$\min_{\gamma,w,b} \frac{1}{2}||w||^2$

s.t. $y^{(i)}(w^T x^{(i)} + b) \geq 1,\ i=1,\cdots,m$



构造 Largrange 函数:

$\mathcal{L}(w,b,\alpha) = \frac{1}{2}||w||^2 - \sum_{i=1}^{m} \alpha_i [y^{(i)}(w^T x^{(i)}+b)-1]$

$d^* = \max_{\alpha,\beta:\alpha_i \geq 0} \min \mathcal{L}(w,\alpha,\beta)$

结果化简为:

$\mathcal{L}(w,b,\alpha) = \sum_{i=1}^{m} \alpha_i - \frac{1}{2}\sum_{i,j=1}^{m} y^{(i)} y^{(j)} \alpha_i \alpha_j (x^{(i)})^T x^{(j)}$

$b^* = -\frac{\max_{i:y^{(i)}=-1} w^{*T} x^{(i)} + \min_{i:y^{(i)}=1} w^{*T} x^{(i)}}{2}$

总结：先根据 w 和 b 做一次运算，判断正负

再用内积的方式求 w 和 b！