第 6 章　支持向量机

式 (6.9)

$$\boldsymbol{w}=\sum_{i=1}^{m}\alpha_i y_i \boldsymbol{x}_i$$

解析

式 (6.8) 可作如下展开:

$$\begin{aligned}L(\boldsymbol{w},b,\boldsymbol{\alpha})&=\frac{1}{2}\|\boldsymbol{w}\|^2+\sum_{i=1}^{m}\alpha_i\left(1-y_i\left(\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i+b\right)\right)\\&=\frac{1}{2}\|\boldsymbol{w}\|^2+\sum_{i=1}^{m}\left(\alpha_i-\alpha_i y_i\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i-\alpha_i y_i b\right)\\&=\frac{1}{2}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w}+\sum_{i=1}^{m}\alpha_i-\sum_{i=1}^{m}\alpha_i y_i\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}_i-\sum_{i=1}^{m}\alpha_i y_i b.\end{aligned}$$

对 $\boldsymbol{w}$ 和 b 分别求偏导数并分别令其等于 $\mathbf{0}$ 和 0，有

$$\frac{\partial L}{\partial \boldsymbol{w}}=\frac{1}{2}\times 2\times\boldsymbol{w}+\mathbf{0}-\sum_{i=1}^{m}\alpha_i y_i\boldsymbol{x}_i-\mathbf{0}=\mathbf{0}\Rightarrow\boldsymbol{w}=\sum_{i=1}^{m}\alpha_i y_i\boldsymbol{x}_i,$$

$$\frac{\partial L}{\partial b}=0+0-0-\sum_{i=1}^{m}\alpha_i y_i=0\Rightarrow\sum_{i=1}^{m}\alpha_i y_i=0.$$

值得一提的是，上述求解过程遵循的是“西瓜书”附录 B 中式 (B.7) 左侧的那段话:“在推导对偶问题时，常通过将拉格朗日函数 $L(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu})$ 对 $\boldsymbol{x}$ 求导并令导数为 0，来获得对偶函数的表达形式. ”那么这段话背后的缘由是什么呢？在这里我们认为有两种说法可以进行解释:

(1) 对于强对偶性成立的优化问题，其主问题的最优解 $\boldsymbol{x}^*$ 一定满足本章附注给出的 KKT 条件，而 KKT 条件中的条件 (1) 就要求最优解 $\boldsymbol{x}^*$ 能使得拉格朗日函数 $L(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu})$ 关于 $\boldsymbol{x}$ 的一阶导数等于 0;

证明参见参考文献 [1] 的 5.5 节