Problem Set 4

自硕21 崔晏菲 2021210976

1.解：当样本线性不可分时，

假设有 l 个二分类样本 $(y_1, x_1), \cdots (y_l, x_l)$，$y \in \{-1, 1\}$

那么原问题的形式为

$$\min \ \Phi(w,\xi) = \frac{1}{2}(w \cdot w) + C\sum_{i=1}^{l}\xi_i$$

$$s.t. \quad y_i(w \cdot x_i + b) \geq 1 - \xi_i, \quad i = 1, 2, \cdots, l$$

拉格朗日函数为 $L(w, b, \xi, \lambda) = \frac{1}{2}(w \cdot w) + C\sum_{i=1}^{l}\xi_i - \sum_{i=1}^{l}\lambda_i[(wx_i + b)y_i - 1 + \xi_i] - \sum_{i=1}^{l}\gamma_i\xi_i$

有

$$\begin{cases} \dfrac{\partial L}{\partial w} = w - \sum_{i=1}^{l}\lambda_i y_i x_i = 0 \\ \dfrac{\partial L}{\partial b} = -\sum_{i=1}^{l}\lambda_i y_i = 0 \\ \left.\begin{array}{l} \lambda_i[(wx_i + b)y_i - 1 + \xi_i] = 0, \quad i = 1, 2, \cdots, l \\ \gamma_i\xi_i = (C - \lambda_i)\xi_i = 0, \quad i = 1, 2, \cdots, l \end{array}\right\} \text{KKT条件} \end{cases}$$

$$\Rightarrow \begin{cases} w = \sum_{i=1}^{l}\lambda_i y_i x_i \\ \sum_{i=1}^{l}\lambda_i y_i = 0 \end{cases}$$

得 $W(\lambda) = \sum_{i=1}^{l}\lambda_i - \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}\lambda_i\lambda_j y_i y_j (x_i \cdot x_j)$

即对偶问题为

$$\max \ W(\lambda) = \sum_{i=1}^{l}\lambda_i - \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}\lambda_i\lambda_j y_i y_j (x_i \cdot x_j)$$

$$s.t. \quad \sum_{i=1}^{l}\lambda_i y_i = 0$$

$$\text{且 } 0 \leq \lambda_i \leq C, \quad i = 1, \cdots, l$$

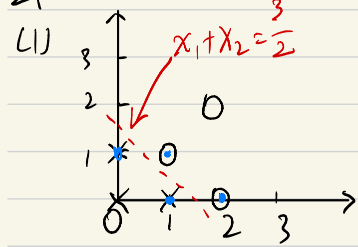
我们有 $\|w_0\|^2 = w_0^T \cdot w_0 = \sum_{i=1}^{l}\sum_{j=1}^{l}\lambda_i\lambda_j y_i y_j (x_i \cdot x_j) = \sum_{i=1}^{l}\lambda_i^0$

根据KKT条件，当样本被错分时，因为 $(C - \lambda_i)\xi_i = 0$，而错分时 $\xi_i > 0$

故 $\lambda_i = C$. 证毕

2、

(1)



W_1: O , $y_i=1$

W_2: X , $y_i=-1$

最优超平面如右图红色所示

最优Margin为 $\frac{\sqrt{2}}{4}$

支持向量为蓝色

只有 $\alpha_2=\alpha_4=0$

(2) 在对偶问题中

$$\max \quad W(\alpha)=\sum_{i=1}^{6}\alpha_i-\frac{1}{2}\sum_{i=1}^{6}\sum_{j=1}^{6}\alpha_i\alpha_j y_i y_j (x_i\cdot x_j)$$

$$=\alpha_1+\alpha_3+\alpha_5+\alpha_6-\frac{1}{2}\left(2\alpha_1^2+4\alpha_3^2+\alpha_5^2+\alpha_6^2+4\alpha_1\alpha_3-2\alpha_1\alpha_5-2\alpha_1\alpha_6-4\alpha_3\alpha_5\right)$$

$\alpha_1,\alpha_3,\alpha_5,\alpha_6>0$

$$\begin{cases}\frac{\partial W}{\partial \alpha_1}=1-(2\alpha_1+2\alpha_3-\alpha_5-\alpha_6)=0\\ \frac{\partial W}{\partial \alpha_3}=1-(4\alpha_3+2\alpha_1-2\alpha_5)=0\\ \frac{\partial W}{\partial \alpha_5}=1-(\alpha_5-\alpha_1-2\alpha_3)=0\\ \frac{\partial W}{\partial \alpha_6}=1-(\alpha_6-\alpha_1)=0\end{cases}$$

解得

$w=\sum_{i=1}^{6}\alpha_i y_i x_i=\begin{pmatrix}1\\1\end{pmatrix}$

$b=\frac{3}{2}$

和第(1)问一致