Homework 1

1 习题一

Q: 验证感知机为什么不能表示异或。

A: 假设存在感知机能表示异或，则存在 $w=(w_1,w_2),b$，使得

$$\begin{cases} w\cdot(1,1)+b=w_1+w_2+b<0 \\ w\cdot(0,0)+b=b<0 \\ w\cdot(0,1)+b=w_2+b>0 \\ w\cdot(1,0)+b=w_1+b>0 \end{cases} \tag{1}$$

前两式相加，后两式相加，得到矛盾。所以不存在感知机表示异或。

2 习题二

Q: 采用 $\sum_{i=1}^N \xi_i^2$ 来定义软间隔最大化问题如下：

$$\min_{w,b,\xi} \quad \frac{1}{2}\|w\|^2 + C\sum_{i=1}^N \xi_i^2$$

$$\begin{aligned} s.t. \quad & y_i(w\cdot x_i+b)\geq 1-\xi_i, \quad i=1,2,\cdots,N. \\ & \xi_i \geq 0, \quad i=1,2,\cdots,N. \end{aligned}$$

给出其对偶形式。

A: 该问题的拉格朗日函数为

$$L(w,b,\xi,\alpha,\mu)=\frac{1}{2}\|w\|^2+C\sum_{i=1}^N\xi_i^2-\sum_{i=1}^N\alpha_i(y_i(wx_i+b)-1+\xi_i)-\sum_{i=1}^N\mu_i\xi_i$$

$$\begin{cases} \nabla_w L = w-\sum_{i=1}^N\alpha_i y_i x_i=0 \\ \nabla_b L = -\sum_{i=1}^N \alpha_i y_i = 0 \\ \nabla_{\xi_i} L = 2C\xi_i-\alpha_i-\mu_i=0 \end{cases} \tag{2}$$

整理得 $L=-\frac{1}{2}\|w\|^2+\sum_{i=1}^N\alpha_i-C\sum_{i=1}^N\xi_i^2$。将梯度条件带入整理得

$$\min_{w,b,\xi} L = -\frac{1}{2}\sum_{i=1}^N\sum_{j=1}^N\alpha_i\alpha_j y_i y_j(x_i\cdot x_j)+\sum_{i=1}^N\alpha_i-\frac{1}{4C}\sum_{i=1}^N(\alpha_i+\mu_i)^2$$

对偶问题为对上式关于 α_i,μ_i 求极大。显然 μ_i 应取值为 0，故对偶问题为

$$\min_{\alpha} \quad \frac{1}{2}\sum_{i=1}^N\sum_{j=1}^N\alpha_i\alpha_j y_i y_j(x_i\cdot x_j)-\sum_{i=1}^N\alpha_i+\frac{1}{4C}\sum_{i=1}^N\alpha_i^2$$

$$\begin{aligned} s.t. \quad & \sum_{i=1}^N \alpha_i y_i = 0 \\ & \alpha_i \geq 0, \quad i=1,2,\cdots,N. \end{aligned}$$