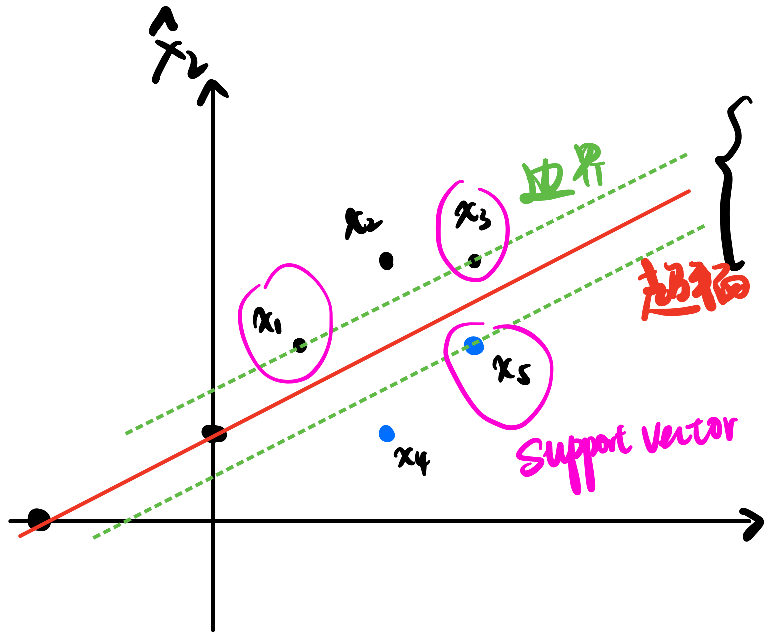
Homwork 9.



1. $\begin{cases} \text{设超平面为: } w^Tx+b=0 \\ \text{在约束条件 } y_i(w^Tx_i+b_i)\geq 1 \text{ 下,} \end{cases}$

max: $\frac{1}{\|w\|}$ (1)

$\Updownarrow$

Lagrange 对偶问题:

$\begin{cases} \text{在 } \alpha_i\geq 0, \text{且 } \sum_{i=1}^{n}\alpha_i y_i=0 \text{ 的约束} \\ \max\limits_{\alpha} \sum_{i=1}^{n}\alpha_i-\frac{1}{2}\sum_{i,j}\alpha_i\alpha_j y_i y_j (x_i^T x_j) \quad (2) \end{cases}$

将约束: $\alpha_1+\alpha_2+\alpha_3-\alpha_4-\alpha_5=0$ 代入 (2).

$$\max f(\alpha_1,\alpha_2,\alpha_3,\alpha_4)=2(\alpha_1+\alpha_2+\alpha_3)-\frac{1}{4}(4\alpha_1^2+2\alpha_2^2+\alpha_3^2+2\alpha_4^2+4\alpha_1\alpha_2+6\alpha_1\alpha_4+2\alpha_2\alpha_3+2\alpha_2\alpha_4)$$

$$\begin{cases} \frac{\partial f}{\partial \alpha_1}=0 \\ \frac{\partial f}{\partial \alpha_2}=0 \\ \frac{\partial f}{\partial \alpha_3}=0 \\ \frac{\partial f}{\partial \alpha_4}=0 \end{cases} \Rightarrow \begin{cases} \alpha_1=-0.4<0 \\ \alpha_2=1.2 \\ \alpha_3=0 \\ \alpha_4=0 \end{cases}$$ 不满足约束条件 $\alpha_i\geq 0$

那么最大值只能在边界处取到: $\begin{cases} \alpha_1,\alpha_2,\alpha_3,\alpha_4=0 \\ \alpha_1+\alpha_2+\alpha_3-\alpha_4=0 \end{cases}$

由于从 f 方程形式可以看到 $\frac{\partial f}{\partial \alpha_4}=-\alpha_4-\frac{3}{2}\alpha_1-\frac{1}{2}\alpha_2<0$

故只能取 $\alpha_4=0$.

以下尝试解

$$\begin{cases} \text{若 } \alpha_1=0, & \frac{\partial f}{\partial \alpha_2}=0 \ \frac{\partial f}{\partial \alpha_3}=0 \Rightarrow \alpha_2=0, \alpha_3=2, \ f=2. \\ \text{若 } \alpha_2=0, & \frac{\partial f}{\partial \alpha_1}=0, \frac{\partial f}{\partial \alpha_3}=0 \Rightarrow \alpha_1=0.5, \alpha_3=2, \ f=2.5. \\ \text{若 } \alpha_3=0, & \frac{\partial f}{\partial \alpha_1}=0, \frac{\partial f}{\partial \alpha_2}=0 \Rightarrow \alpha_1=0, \alpha_2=1 \ \ f=1. \end{cases}$$

从上: 得: $\alpha_1=\frac{1}{2}, \alpha_2=0, \alpha_3=2, \alpha_4=0, \alpha_5=\frac{5}{2}$

超平面为: $w=\sum\alpha_i y_i x_i$ $b=y_i-w^Tx_i$ $i=1,3,5$ 其中之一

即 $-x_1+2x_2-2=0$

决策函数 $\mathrm{sign}(-x_1+2x_2-3)$.

2. $L_{CE} = -\left[y \log \sigma(w \cdot x + b) + (1-y) \log(1 - \sigma(w^T x + b)) \right]$

首先，令 $z = w^T x + b$

$$\frac{\partial \log \sigma(w^T x + b)}{\partial w_j} = \frac{\partial \log \sigma(z)}{\partial z} \cdot \frac{\partial z}{\partial w_j} = \frac{\sigma'(z)}{\sigma(z)} \cdot x_j = (1 - \sigma(z)) x_j .$$

$$\frac{\partial \log(1 - \sigma(w^T x + b))}{\partial w_j} = \frac{\partial \log(1 - \sigma(z))}{\partial z} \frac{\partial z}{\partial w_j} = \frac{-\sigma'(z)}{1 - \sigma(z)} \cdot x_j = -\sigma(z) x_j .$$

所以：
$$\frac{\partial L_{CE}(w, b)}{\partial w_j} = -y \left[1 - \sigma(w^T x + b) \right] x_j + (1-y) \sigma(w^T x + b) x_j .$$
$$= -y x_j + \sigma(w^T x + b) x_j$$