

第二次作业

1. 交叉熵损失函数:

$$L(w) = -y \log(\sigma(xw)) - (1-y) \log(1 - \sigma(xw))$$

$$\text{Sigmoid: } \sigma(z) = \frac{1}{1 + e^{-z}}$$

2. a. 交叉熵损失函数对参数 w 的一阶导数.

$$\frac{\partial L(w)}{\partial w} = \frac{1}{N} \sum_{i=1}^N [\sigma(x^{(i)}w) - y^{(i)}] x^{(i)T}$$

b. 参数更新公式:

$$w_{t+1} = w_t - r \frac{\partial L(w)}{\partial w}$$

3. a. 交叉熵损失函数 $L(w)$ 对参数 w 的二阶导数 (海森矩阵) H 的表达式:

$$H = \frac{\partial^2 L(w)}{\partial w \partial w^T} = \sum x x^T \sigma(xw) (1 - \sigma(xw))$$

b. 判断半正定性: $v^T H v = (\sum x^T v)^2 \sigma(xw) (1 - \sigma(xw)) \geq 0$

因为交叉熵函数的海森矩阵是半正定的, 所以该函数是凸函数,
且交叉熵损失函数是概率分布的凸组合, 表现仍为凸的