年 月 日 第 次作业

Logistic Regression | Guorui Shen, guorui233@outlook.com. May 23, 2020.

① $\{x_i, y_i\}_{i=1}^{N}$, $x_i \in R^n$, $y_i = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ or $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ ，表示二分类问题.

$$y_i = \begin{bmatrix} y_{i1} \\ y_{i2} \end{bmatrix} = \begin{bmatrix} y_{i1} \\ 1-y_{i1} \end{bmatrix}$$

② 建立函数关系 $x_i \longrightarrow y_i$, $\hat{y}_i = \begin{bmatrix} \hat{y}_{i1} \\ \hat{y}_{i2} \end{bmatrix} = \begin{bmatrix} \frac{1}{1+e^{-\theta^T x_i}} \\ \frac{e^{-\theta^T x_i}}{1+e^{-\theta^T x_i}} \end{bmatrix} = \begin{bmatrix} \sigma(\theta^T x_i) \\ 1-\sigma(\theta^T x_i) \end{bmatrix}$

其中 $\sigma(t) = \frac{1}{1+e^{-t}}$ 是 Sigmoid 函数，满足 $\sigma'(t) = \sigma(t) \cdot [1-\sigma(t)]$

③ 损失函数 (cross entropy) :

$$loss(y_i, \hat{y}_i) = -y_{i1} \cdot \ln \sigma(\theta^T x_i) - (1-y_{i1}) \ln\left(1-\sigma(\theta^T x_i)\right)$$

④ 求导: $\frac{\partial loss}{\partial \theta} = -y_{i1} \frac{\sigma(\theta^T x_i) \cdot [1-\sigma(\theta^T x_i)]}{\sigma(\theta^T x_i)} \cdot x_i + (1-y_{i1}) \frac{\sigma(\theta^T x_i)[1-\sigma(\theta^T x_i)]}{1-\sigma(\theta^T x_i)} x_i$

$$= [\sigma(\theta^T x_i) - y_{i1}] \cdot x_i$$

⑤ 总的损失函数 $J(\theta) = \frac{1}{N} \sum_{i=1}^{N} loss(y_i, \hat{y}_i)$

所以 $\theta \leftarrow \theta - \lambda \cdot \frac{1}{N} \sum_{i=1}^{N} [\sigma(\theta^T x_i) - y_{i1}] \cdot x_i$.