年 月 日 第 次作业

Logistic Regression 推导.

Guorui Shen, guorui233@outlook.com, Mar 26, 2020

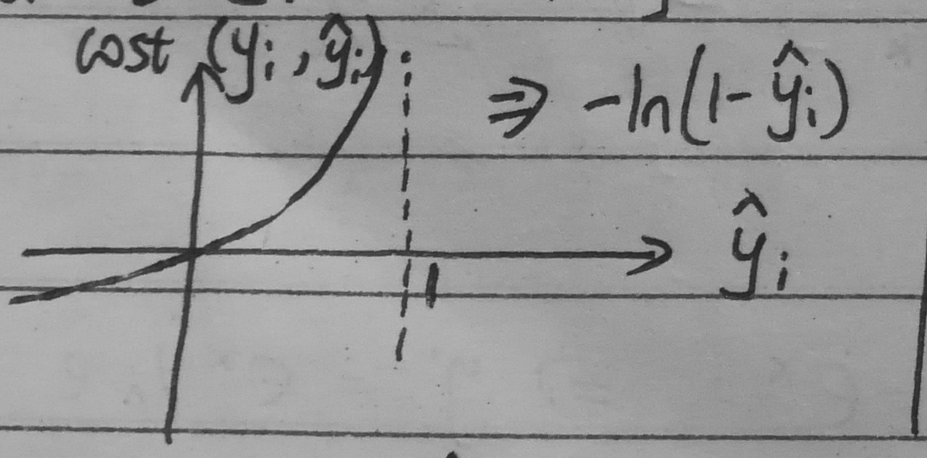
① 假设有 m 个数据点 $\{x_i, y_i\}_{i=1}^{m}$，其中 $x_i \in R^n$，$y_i = \begin{cases} 0 \\ 1 \end{cases}$ 表示二分类问题.

② 建立函数关系：$x_i \xrightarrow{f} y_i$，$y_i = f(\theta, x_i) = \frac{1}{1+e^{-\theta^T x_i}}$

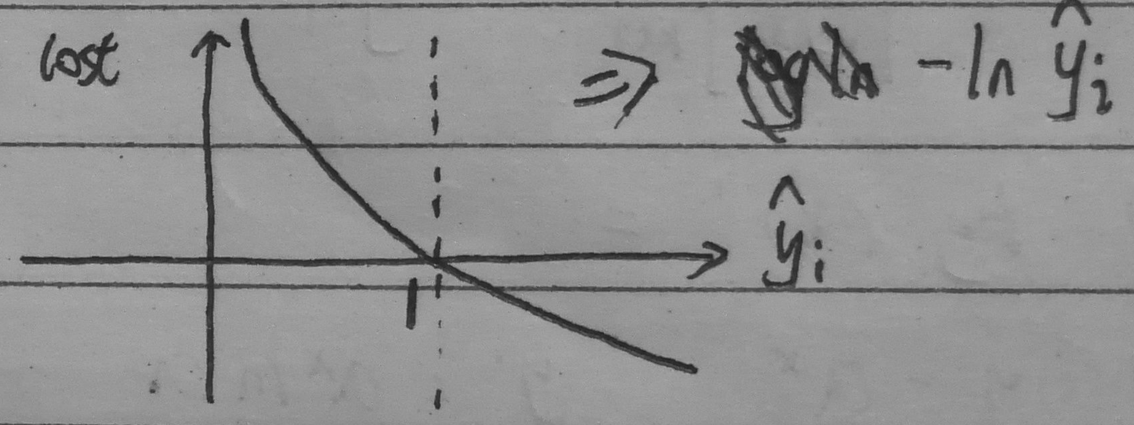
③ 建立损失函数 $cost(\hat{y}_i, y_i)$，其中 $\begin{cases} \hat{y}_i \text{ 是预测值} \\ y_i \text{ 是真实值.} \end{cases}$

原则：预测 $\begin{cases} \text{错误给予较大惩罚,} \\ \text{正确给予零惩罚.} \end{cases}$

a. 若 $y_i = 0$，则



b. 若 $y_i = 1$，则



所以 $cost(\hat{y}_i, y_i)$ 可以定义为：

$$cost(\hat{y}_i, y_i) = -y_i \ln \hat{y}_i - (1-y_i)\ln(1-\hat{y}_i)$$

$$= -y_i \ln f(\theta, x_i) - (1-y_i)\ln[1-f(\theta, x_i)]$$

所以，损失函数累加和为：

$$J(\theta) = \frac{1}{m}\sum_{i=1}^{m} cost(\hat{y}_i, y_i) = -\frac{1}{m}\sum_{i=1}^{m} y_i \ln f(\theta, x_i) + (1-y_i)\ln[1-f(\theta, x_i)]$$

年 月 日 第 次作业

㈣ Gradient descent, θ_j 更新 $\theta_j \leftarrow \theta_j - \alpha \frac{\partial J(\theta)}{\partial \theta_j}$

其中 $\frac{\partial J(\theta)}{\partial \theta_j} = -\frac{1}{m}\sum_{i=1}^{m} \frac{\partial}{\partial \theta_j}\left\{ y_i \ln \frac{1}{1+e^{-\theta^T x_i}} + (1-y_i)\ln\left[1 - \frac{1}{1+e^{-\theta^T x_i}}\right]\right\}$

$= -\frac{1}{m}\sum_{i=1}^{m} \frac{\partial}{\partial \theta_j}\left\{ -y_i \ln\left(1+e^{-\theta^T x_i}\right) - (1-y_i)\ln\left(1+e^{\theta^T x_i}\right)\right\}$

$= -\frac{1}{m}\sum_{i=1}^{m} \left\{ \frac{-y_i e^{-\theta^T x_i}\cdot(-x_{ij})}{1+e^{-\theta^T x_i}} - (1-y_i)\frac{e^{\theta^T x_i} x_{ij}}{1+e^{\theta^T x_i}}\right\}$

$= -\frac{1}{m}\sum_{i=1}^{m} \left\{ \frac{y_i x_{ij}}{1+e^{\theta^T x_i}} - (1-y_i)\frac{x_{ij} e^{\theta^T x_i}}{1+e^{\theta^T x_i}}\right\}$

$= -\frac{1}{m}\sum_{i=1}^{m} \frac{y_i x_{ij} - x_{ij} e^{\theta^T x_i} + y_i x_{ij} e^{\theta^T x_i}}{1+e^{\theta^T x_i}}$

$= -\frac{1}{m}\sum_{i=1}^{m} \frac{y_i(1+e^{\theta^T x_i}) - e^{\theta^T x_i}}{1+e^{\theta^T x_i}} \cdot x_{ij}$

$= -\frac{1}{m}\sum_{i=1}^{m} \left[y_i - \frac{e^{\theta^T x_i}}{1+e^{\theta^T x_i}}\right] x_{ij}$

$= -\frac{1}{m}\sum_{i=1}^{m} \left[y_i - \frac{1}{1+e^{-\theta^T x_i}}\right] x_{ij}$

$= \frac{1}{m}\sum_{i=1}^{m} \left[f(\theta, x_i) - y_i\right] x_{ij}$

所以: $\theta_j \leftarrow \theta_j - \alpha \frac{1}{m}\sum_{i=1}^{m} \left[f(\theta, x_i) - y_i\right] x_{ij}$

年　月　日 第　次作业

⑤ Logistic regression 是神经网络模型吗？ 是的.