统计学习实验三：Logistics Regression

王嗣萱 2018110601014

1 实验原理

1.1 Logistic 分布

Logistic 分布是一种连续型的概率分布，其分布函数和密度函数分别为：

$$\begin{aligned} F(x) &= P(X \le x) = \frac{1}{1+e^{-(x-\mu)/\gamma}} \\ f(x) &= F^{'}(X \le x) = \frac{e^{-(x-\mu)/\gamma}}{\gamma(1+e^{-(x-\mu)/\gamma})^2} \end{aligned} \tag{1}$$

其中，μ 表示位置参数，γ 为形状参数。我们可以看下其图像特征：

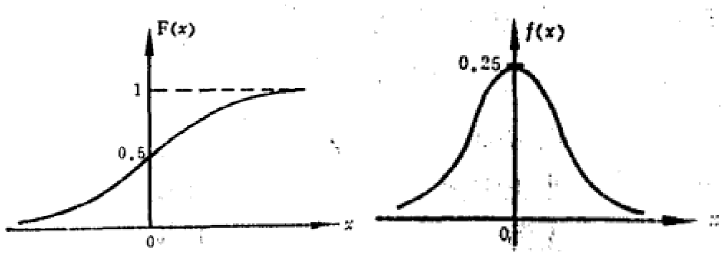

图 1: Logistic 分布

Logistic 分布是由其位置和尺度参数定义的连续分布。Logistic 分布的形状与正态分布的形状相似，但是 Logistic 分布的尾部更长，所以我们可以使用 Logistic 分布来建模比正态分布具有更长尾部和更高波峰的数据分布。在深度学习中常用到的 Sigmoid 函数就是 Logistic 的分布函数在 $\mu = 0, \gamma = 1$ 的特殊形式。

1.2 Logistic 回归

本次实验中考虑'0'，'1' 的二分类问题，即一个伯努利分布。

Sigmoid 函数，也称为逻辑函数（Logistic function）：

$$g(z) = \frac{1}{1+e^{-z}} \tag{2}$$

逻辑回归的假设函数形式如下：

$$\begin{aligned} & h_\theta(x) = g(\theta^T x), g(z) = \frac{1}{1+e^{-z}} \\ & P(y=1|x;\theta) = h_\theta(x) \\ & P(y=0|x;\theta) = 1 - h_\theta(x) \end{aligned} \tag{3}$$

可以将其合并为一个表达式：

$$P(y|x;\theta) = (h_\theta(x))^y (1-h_\theta(x))^{1-y} \tag{4}$$

logistic regression 的目标函数是根据最大似然思想求得的。似然函数为：

$$L(\theta) = \prod_{i=1}^{n} (h_\theta(x^i))^{y^i} (1-h_\theta(x^i))^{1-y^i} \tag{5}$$

对 $L(\theta)$ 求对数可以得到:

$$l(\theta) = -logL(\theta) = -\sum_{i=1}^{n} [y^i log(h_\theta(x^i)) + (1-y^i) log(1-h_\theta(x^i))] \tag{6}$$

使用 $J(\theta) = \frac{1}{m} l(\theta) = -\frac{1}{N} logL(w)$ 作为 logistic regression 的损失函数

1.3

2 Python 代码实现

3 结果及图形展示

4 总结体会