

机器学习课程实验报告

逻辑回归

学 号	1180301007
姓 名	赵锦涛
实验时间	2020 年 10 月

一、实验目的:

理解逻辑回归模型，掌握逻辑回归模型的参数估计算法。

二、实验要求:

实现两种损失函数的参数估计（1，无惩罚项；2. 加入对参数的惩罚），可以采用梯度下降、共轭梯度或者牛顿法等。

三、实验环境:

Python 3.8, Windows 10

四、实验原理:

逻辑回归是一种分类方法。对于二分类问题，给定输入 $X = \langle X_1, \dots, X_n \rangle$ ，输出 $Y \in \{0, 1\}$ ，有

$$P(Y = 1|X) = \frac{\exp(wX + b)}{1 + \exp(wX + b)}$$

$$P(Y = 0|X) = \frac{1}{1 + \exp(wX + b)}$$

扩充 X 向量为 $X = \{1, X_1, \dots, X_n\}$ ， w 向量为 $w = b, w_1, \dots, w_n$ ，则后验类概率可表示为

$$P(Y = 1|X) = \frac{\exp(wX)}{1 + \exp(wX)}$$

$$P(Y = 0|X) = \frac{1}{1 + \exp(wX)}$$

当训练得到参数 w 后，便可将样本代入求得相应的概率，样本分类类别取使概率较大的那类。

逻辑回归模型参数估计时可使用最大似然法估计方法参数，设

$$P(Y = 1|x) = p(x), P(Y = 0|x) = 1 - p(x)$$

似然函数为

$$\prod [p(x_i)]^{y_i} [1 - p(x_1)]^{1-y_i}$$

取对数有

$$\begin{aligned} L(w) &= \sum [y_i \ln p(x_i) + (1 - y_i) \ln(1 - p(x_i))] \\ &= \sum [y_i (wx_i) - \ln(1 + \exp(wx_i))] \end{aligned}$$

五、 实验结果与分析