编 程 学 习 笔 记

Logistic Regression（LR）

Code

作 者 姓 名： 颜佳

目 录

[1 第一章 1](#_Toc76209995)

[1.1 第一节 1](#_Toc76209996)

[1.1.1 第一段 1](#_Toc76209997)

[1.2 1](#_Toc76209998)

[2 1](#_Toc76209999)

# 二分类问题



Sigmoid函数：



假设：



给定，通过该模型预测和的概率分别为：





即：



对于含有个样本的训练集，极大似然函数为：



两边同时取对数，得到极大对数似然函数：



要求极大似然函数的最大值，即求下列函数的极小值



对于一个样本的情况



当时，时没有损失，越趋近损失越大；

当时，时没有损失，越趋近损失越大；

梯度



对数似然函数的本质就是衡量在某个参数下，整体的估计和真实情况一样的概率，越大代表越相近。而损失函数的本质就是衡量预测值和真实值之间的差距，越小代表越相近。在二分类的情况下，对数似然函数的最大化的本质就是交叉熵的最小化。

当所求的模型时概率模型的条件概率分布，损失函数是对数损失函数时，经验风险最小化就等价于极大似然估计。

对数损失函数，即对数似然损失(Log-Likelihood Loss)，也称逻辑斯谛回归损失(Logistic Loss)或交叉熵损失(Cross-Entropy Loss)，是在概率估计上定义的。它常用于逻辑斯谛回归和神经网络，以及一些期望极大算法的变体，可用于评估分类器的概率输出。

对数损失通过惩罚错误的分类，实现对分类器的准确度(Accuracy)的量化。最小化对数损失基本等价于最大化分类器的准确度。为了计算对数损失，分类器必须提供对输入的所属的每个类别的概率值，不只是最可能的类别。对数损失函数的计算公式如下:

对数损失函数（概率密度估计问题）



式中：

——样本容量；

——分类的个数；

——一个二值指标，表示类别是否是输入实例的真实类别；

——模型或分类器预测输入实例属于类别的概率；

对于二分类问题，上式变为：

