基于逻辑回归模型的研究

熊舒鹏

（昆明理工大学 机器人工程 昆明市）

1. **逻辑回归（Logical regression）**

## 逻辑回归定义

**(Grok3)**

逻辑回归（Logistic Regression）是机器学习中一种经典的分类算法，特别适用于二分类问题。它的核心思想是通过数学公式，将输入特征映射为一个概率值，用于预测样本属于某个类别的可能性。下面，我将详细推导逻辑回归模型的数学公式，包括其概率表示、损失函数以及参数估计的过程。

**逻辑回归的目标与基本假设**

逻辑回归的目标是预测一个样本属于正类（通常标记为1）的概率。在二分类问题中，类别标签为0（负类）或1（正类）。给定输入特征向量X，逻辑回归模型输出一个介于0和1之间的值，表示样本属于正类的。

逻辑回归的基础是从线性回归出发。线性回归的模型形式为：

*z=****w****T****x****+b*

其中，**w**是权重向量，**x**是特征向量，b是偏置项，z是线性组合的结果。

然而，线性回归的输出z是连续值，无法直接表示概率。为了将其转换为概率，我们引入 sigmoid函数。

**Sigmoid函数与概率表示**

Sigmoid函数定义为：

****

该函数将任意实数z映射到(0,1)区间，非常适合表示概率。因此，逻辑回归假设样本属于正类的概率为：



相应地，属于负类的概率为：



为了统一表达，我们可以将条件概率写为：

****

其中，是真实的类别标签。这样，当y=1时，概率为σ(z)；当y=0时，概率为1−σ(z)。

**参数估计：最大似然估计**

逻辑回归的参数w和b需要通过训练数据来估计。我们使用 最大似然估计（Maximum Likelihood Estimation, MLE）来找到最优参数。

假设m个独立的训练样本，其中。给定参数w和b，整个训练数据集的似然函数是所有样本概率的乘积：



其中，*z=****w****T****x****i+b*，是样本i的预测概率，记为pi。

由于连乘计算复杂且容易导致数值下溢，我们取对数，将似然函数转换为对数似然函数：



目标是最大化对数似然函数log L(***w***,b)。

**损失函数定义**

在实践中，我们通常将问题转化为最小化损失函数。逻辑回归的损失函数定义为负对数似然函数的平均值，称为**交叉熵损失函数**：



其中，

训练逻辑回归模型的任务就是通过优化算法最小化J(w,b)。

**梯度计算与参数更新**

为了最小化损失函数J(w,b)，我们使用梯度下降法。需要计算损失函数对参数w和b的偏导数。

(1) Sigmoid函数的导数

Sigmoid函数有一个重要性质：



这在梯度计算中非常有用。

#### (2) 对zi的偏导数

对于单个样本*i*的损失项：



其中pi=σ(zi)。

计算li对zi的偏导数：



由于：



代入得：



因此：



#### (3) 对w和b的梯度

根据链式法则：



#### (4) 梯度下降更新规则

使用学习率*α*，参数更新规则为：



通过迭代更新，直到损失函数收敛，得到最优的w和b。