#机器学习笔记

极大似然估计（Maximum Likelihood Estimation, MLE）是一种在统计学中用于从给定数据中估计模型参数的方法。它基于一个简单的原则：选择使观测数据出现概率（即"似然性"）最大的参数值。MLE是许多统计模型和机器学习算法中的核心技术。

基本原理：

- **模型假设**：首先，你需要有一个关于数据如何生成的概率模型，该模型包含了一些未知的参数。
- **似然函数**：接着，定义似然函数，它是关于这些参数的函数，表示在给定这些参数下观测到实际数据样本的概率。
- **最大化似然**：最后，通过找到使似然函数取最大值的参数，来估计这些未知参数。这些使似然函数最大的参数值被认为是给定数据下模型参数的最佳估计。

数学表示：

设 $X_1, X_2, \ldots, X_n$ 为随机样本，它们独立同分布于某个概率分布，该分布由参数 θ 决定。似然函数 $L(\theta)$ 定义为这组观测值的联合概率密度或质量函数，视为关于 θ 的函数：

$$L(\theta) = f(X_1, X_2, \ldots, X_n | \theta)$$

对于独立同分布的样本，似然函数通常是各个样本概率的乘积：

$$L(\theta) = \prod_{i=1}^{n} f(X_i | \theta)$$

求解过程：

1. **写出似然函数**：根据你的数据和模型，写出似然函数。
2. **对数似然**：实际操作中，通常取似然函数的对数（对数似然），因为对数转化乘积为求和，这样计算上更简便，不改变估计值。对数似然函数表示为：

$$\log L(\theta) = \sum_{i=1}^{n} \log f(X_i | \theta)$$

3. **求导数**：求对数似然函数关于参数 θ 的导数。

4. **设置导数为0求解**：找到使导数（或偏导数，如果有多个参数）为零的参数值，这通常涉及到求解方程或方程组。
5. **确定最大值**：验证所得解确实使似然函数最大化（而不是最小化），通常通过检查二阶导数或使用数值方法。

应用：

MLE被广泛应用于统计推断、参数估计和机器学习模型中，包括但不限于线性回归、逻辑回归、复杂的混合模型等。它是一种强大的方法，尤其是在模型正确指定，且大样本条件下，MLE估计的性质（如一致性和渐进正态性）使其成为实际应用中的可靠选择。