使用不同方法拟合线性模型报告

徐阳
xuyangx@buaa.edu.cn

Abstract

这是使用不同方法拟合线性模型的报告。本人使用了最小二乘法、梯度下降法、牛顿法在两组数据（x向量和y向量）之间建立了一次函数关系，并进行比较分析。但因为该组数据的线性程度较差，拟合结构不太理想。使用非线性模型进行拟合，发现效果能够提升。

Introduction

线性模型是学习深度模型的第一步。虽然我已经学习了很多更高层次的深度学习方法，但还没有系统地区分线性模型拟合的方法。

Methodology

下面是我进行线性拟合的方法，也包括非线性的方法

M1: 最小二乘法

通过最小化观测数据与模型预测值之间的残差平方和来拟合模型。目标是最小化损失函数

$$L(\beta) = \sum_{i=1}^{n}(y_i - X_i\beta)^2$$

通过求导并令导数为零，得到解析解

$$\beta = (X^T X)^{-1} X^T y$$

计算简单，适合小规模数据，但对大规模数据计算逆矩阵时效率低。

M2: 梯度下降法

通过迭代更新模型参数，沿着损失函数梯度的反方向逐步逼近最优解，损失函数与最小二乘法相同。参数更新公式：

$$\beta_{k+1} = \beta_k - \alpha \nabla L(\beta_k)\text{，其中 } \alpha \text{ 是学习率，} \nabla L(\beta_k) \text{ 是梯度}$$

适合大规模数据，内存需求低。但需调学习率，可能收敛慢或陷入局部最优

M3: 牛顿法

利用损失函数的二阶导数（Hessian矩阵）加速收敛，通过二阶泰勒展开逼近最优解损失函数与最小二乘法相同。参数更新公式：

$$\beta_{k+1} = \beta_k - H^{-1}\nabla L(\beta_k)\text{，其中 } H \text{ 是 Hessian 矩阵.}$$

M4: 使用非线性模型

我通过观察样本数据发现，待拟合函数有非常强的周期性。我想利用利用人类的先验知识进行初步拟合，再利用梯度下降的方法进行进一步拟合，目标函数设置为：

$$y = a\sin bx + cx + d$$

其中，[a,b,c,d]是参数组，使用梯度下降的方法进行拟合。

Experimental Studies

下面是不同方法得到的实验图

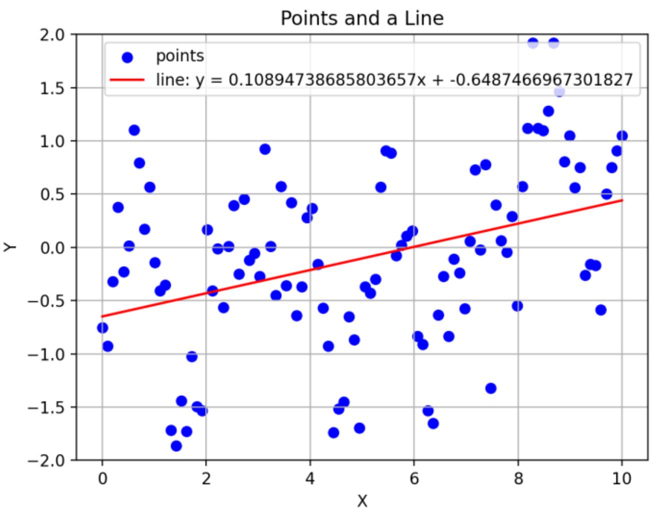

Figure 1: 最小二乘法结果

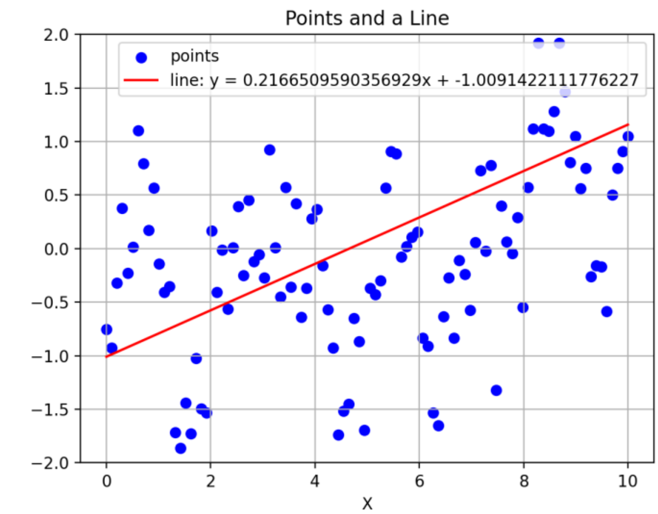

Figure 2: 梯度下降法结果

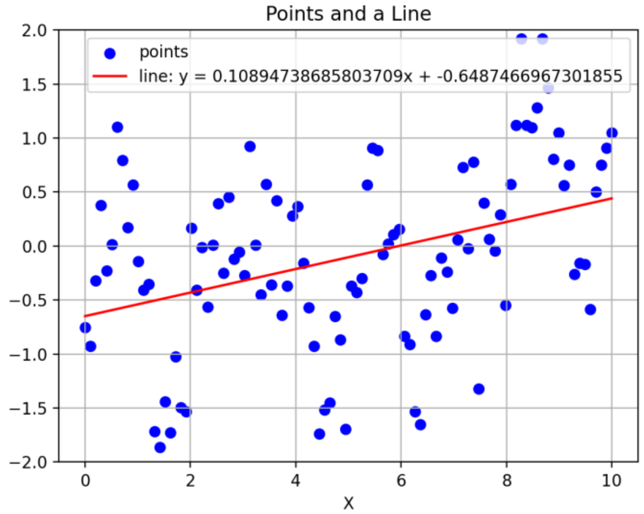

Figure 3： 牛顿法结果

Table 1: 使用均方误差衡量y=ax+b模型的准确性

New Table	a	b	Test Error	Train Error
M1	0.109	-0.649	0.298	0.306
M2	0.217	-1.009	0.349	0.372
M3	0.109	-0.649	0.298	0.306

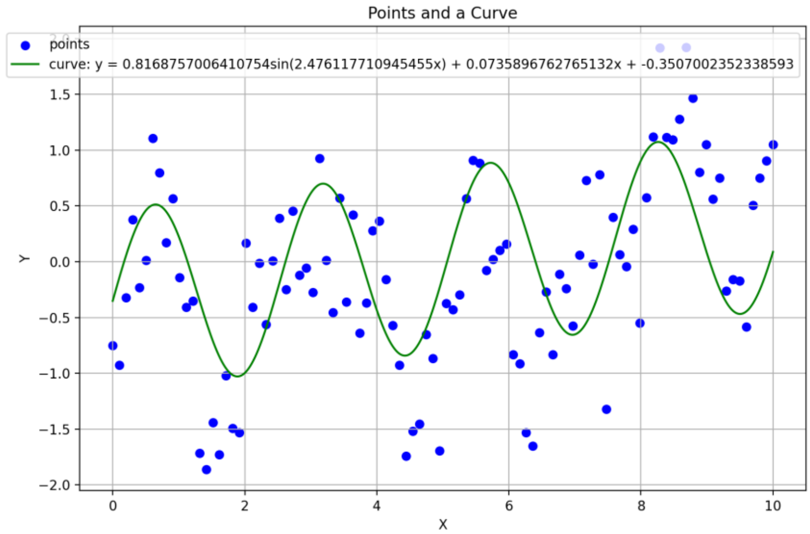

Figure 4： 非线性拟合

使用均方误差评估拟合效果，最后得到的训练误差是0.275，测试误差是0.255

Conclusions

发现三种方法中，最小二乘法和牛顿法拟合效果较好。而梯度下降法由于收敛较慢、容易陷入局部最优的特性，没有找到最优解。但这三种方法拟合出的模型都不如简单设置的非线性模型。