**深 圳 大 学 实 验 报 告**

**实验课程名称： 最优化方法**

**实验项目名称： 最小二乘法实验**

**学院： 计算机与软件学院 专业： 软件工程（腾班）**

**报告人： 黄亮铭 学号： 202215502 班级： 腾班**

**同组人：**

**指导教师： 李炎然**

**实验时间： 2024年12月6日 - 2024年12月26日**

**实验报告提交时间： 2024年12月26日**

**教务处制**

**实验报告包含内容**

# 1实验目的与要求

1. 熟练最小二乘法优化模型的意义和求解手段；
2. 掌握最小二乘法的正规方程，能实现代码对其求解；
3. 掌握最速梯度下降法求解无约束最小二乘法问题。

# 2问题

读取附件“MatrixA\_b.mat”文件中的矩阵A和向量b。建立关于矩阵, 向量，未知向量最小二乘优化模型：



1）通过最小二乘法的正规方程，求出优化模型的准确解；

2）利用梯度下降法迭代求出模型“近似解”，通过设置迭代停止条件，分析“近似解”与“准确解”之间的误差。

# 3模型建立及求解

3.1最小二乘法求解

由于各种误差，在某些情况下难以求解出一组满足问题条件的解。而最小二乘法给出的解决方案就是寻找该问题的近似解，并尽可能地逼近原问题的解，使尽可能的小。

最小二乘法的数学模型如下所示：

3.1.2正规方程推导

目标函数为：

结合最小二乘法的原理，对上述目标函数进行求解，得到的应该满足如下的条件。

由于目标函数为可微函数，因此最优解满足梯度。

对求导得到：

即：

也即：

由此，我们通过最小二乘法得到了正规方程。如果为非奇异矩阵，则此时正规方程（原问题）有唯一解。

3.1.3正规方程求解

对于正规方程，我这里提供了两种求解方法。一种是使用高斯消元法直接对正规方程进行求解，另一种是使用QR分解进行求解。

这里重点讲解如何使用QR分解进行求解。

首先我们对进行QR分解，得到。然后代入正规方程，化简后得到。

3.1.4代码实践

在代码实现上，对于QR分解方面，我们可以选择使用实验3中自己编写的QR分解代码，也可以选择matlab库函数进行QR分解。为了结果的准确性，我选择使用matlab的库函数进行QR分解。

对于矩阵求逆方面，我们可以选择实验1中自己编写的高斯消元法求矩阵的逆的代码进行求解，也可以选择matlab的库函数进行求解。同样是为了准确性，我也选择了使用matlab的库函数进行矩阵求逆的操作。

通过QR分解和矩阵求逆的步骤之后，我们就可以很轻松地得到最优解的精确解（图1）。

|  |
| --- |
|  |

图1精确解

3.2梯度下降法求解

除了最小二乘法，梯度下降法也常用于优化问题最优解的逼近**。**梯度下降法的核心原理为：函数在某点的梯度方向，是函数值增长最快的方向。那么，梯度下降法就反其道而行之，沿着梯度的反方向去更新自变量，使得函数值逐步减小。

3.2.1迭代公式推导

设函数可微，根据泰勒公式，在的一阶公式为：

如果足够小，则有

根据柯西不等式，我们可以得到

即：

当时，等式成立。

由此我们可以推出迭代公式：

3.2.2求解过程

为了消去求导之后的常数，我们可以在目标函数的前面乘以一个常量，该常量不会影响目标函数的最优解，即我们将目标函数从乘以一个常量变为。

令，我们可以得到

通过迭代下降求解：

此外，我们还可以通过线性搜索估计：

即是最佳步长。令是关于的凸函数，则有

上述就是我们完整的求解过程。经过上述步骤，我们成功得到了迭代求解公式和的最佳步长。

3.2.3代码实践

代码编写思路如下：

1. 初始化。
2. 循环：①求解；②求解；③使用迭代公式进行更新。
3. 重复步骤2），直到满足迭代条件。

在迭代条件方面，我选择了双迭代条件判断循环是否结束。第一个迭代条件是迭代次数，第二个迭代条件则是相邻迭代解之间的“相对接近程度”。

为探究最优迭代次数，应观察分析不同迭代次数对于目标求解的影响。本次实验中，迭代次数与目标函数值之间的关系如图2所示。

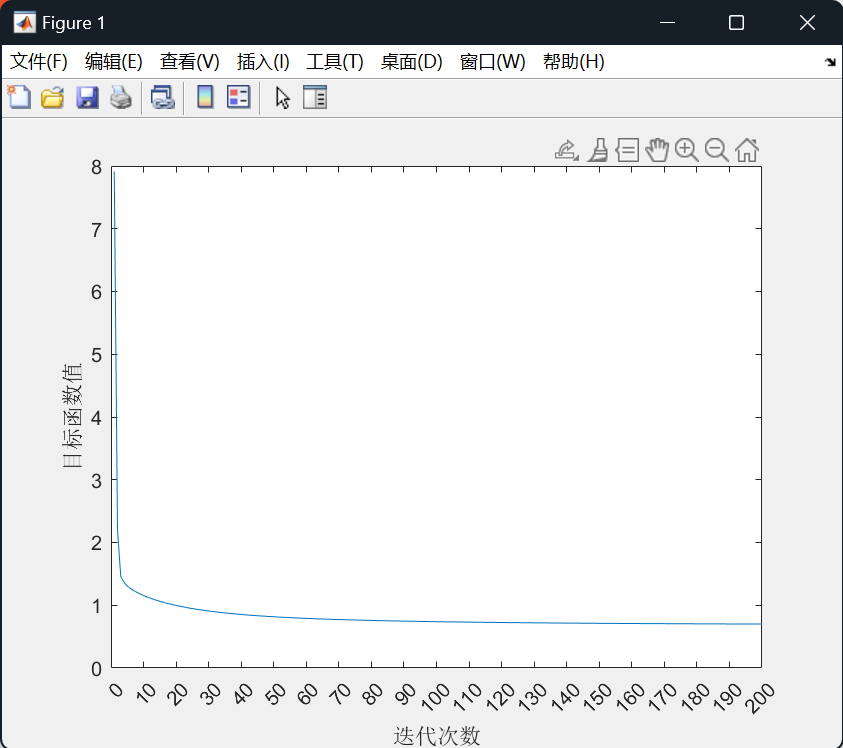


图2迭代次数与目标函数值

从图中我们可以看出，迭代次数在40之后，随迭代次数增大，目标函数趋于稳定。故在本实验中，选取迭代次数为40作为循环停止条件是一个比较好的选择**。**

同理，为探究最优阈值，应观察分析不同阈值对于目标求解的影响。本次实验中，迭代次数与相邻迭代解之间的“相对接近程度”之间的关系如图3所示。

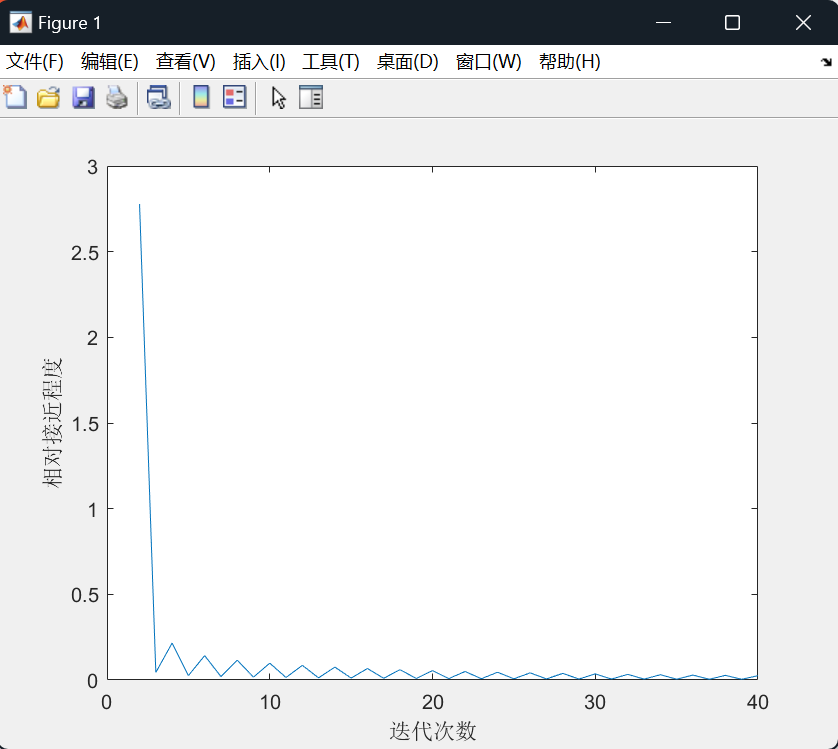


图3迭代次数与相邻迭代解之间的“相对接近程度”之间的关系

从图3中我们可以看出，随迭代次数增加，相邻迭代解之间的“相对接近程度”波动下降。经统计分析后，我选择相邻迭代解之间的“相对接近程度”的阈值为0.001,作为梯度下降法的循环停止条件。

3.3误差分析

首先我们需要比较最小二乘法和梯度下降法之间的区别。

1. **原理**：最小二乘法基于数学推导，通过对误差平方和函数求导，令导数为零得出正规方程，直接求解得到使误差平方和最小的解析解。梯度下降法是迭代式优化算法，依据目标函数当前的梯度方向，朝着函数值下降最快的反方向，按设定的学习率更新参数，持续逼近目标函数最小值。
2. **收敛性**：最小二乘法对于这类凸优化问题，只要相关矩阵（本次实验为）可逆，就能直接获取全局最优解；而梯度下降法收敛速度受学习率、数据特征以及初始化等因素影响，在非凸函数场景下可能陷入局部最优解，不过凸优化场景（本实验）理论上也能收敛到全局最优解。

由上述最小二乘法和梯度下降法之间的区别分析，我们可以知道“精确解”和“近似解”之间的误差主要来源于梯度下降法的初始化的初始解。这个解往往与“准确解”的距离较远，所以每一次迭代的步长的方向和长度都是尽量“减小”误差，但是由于步长的问题，最后得到的解可能还是会与“准确解”存在一定的误差。

图4是比较不同迭代次数下（忽略“相对接近程度”这一循环停止条件）的“近似解”和“精确解”之间的误差，误差的计算方法为。

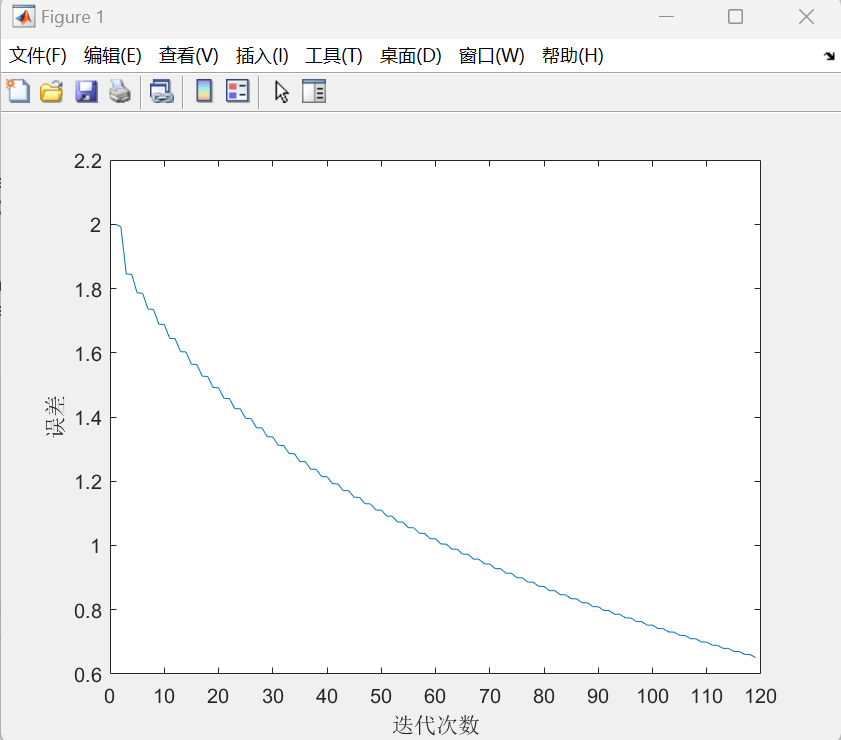


图4“近似解”和“精确解”之间的误差

由图4可以看出，即使迭代次数再多，它们之间仍然存在误差。迭代次数为120的误差近似为0.65139。

# 4小结（可含个人心得体会）

1. 经过本次实验，我熟悉了如何通过最小二乘法的正规方程求解最小二乘优化模型。
2. 通过本次实验，我熟悉了如何使用梯度下降法迭代求解问题。通过不断迭代，解可以不断接近最优解，但是仍然存在误差。
3. 在本实验中，“精确解”和“最优解”之间存在误差。这个误差可能是由于初始值和步长引起的，在其他非凸优化问题中，误差还可能是因为陷入了局部最优问题。一个解决办法为使用模拟退火算法，即算法有概率选择次优解而非每次都选择最优解。

|  |
| --- |
| 指导教师批阅意见：  成绩评定：  指导教师签字：  2024 年 12 月 日 |
| 备注： |

注：1、报告内的项目或内容设置，可根据实际情况加以调整和补充。

2、教师批改学生实验报告时间应在学生提交实验报告时间后10日内。