

数值分析实验二

赵浩宇,2016012390, 计科 60

2018 年 3 月 25 日

1 实验要求

从随机的数据中找出其规律性, 给出其近似表达式的问题, 在生产实践和学实验中大量存在, 通常利用数据的最小二乘法求得拟合曲线。

用最小二乘法求一个多项式, 使得它与下列数据相拟合

x	-1.0,	-0.5,	0.0,	0.5,	1.0,	1.5,	2.0
y	-4.467,	-0.452,	0.551,	0.048,	-0.447,	0.549,	4.552

表 1: 实验数据

分别实现一次函数, 二次函数以及三次函数拟合, 计算拟合函数与实际值的误差, 并且绘制出曲线拟合图以及散点图。

2 算法描述

下面简述最小二乘法的算法。

1. 首先使用 `vector<vector<double>>` 与 `vector<double>` 抽象出矩阵与向量, 并且实现矩阵乘法, 矩阵与向量的乘法, 矩阵转置等基本矩阵操作。
2. 通过之前封装的矩阵与向量, 实现高斯消元解线性方程组的功能。即输入 A, b , 输出 x 使得 $Ax = b$ 。
3. 之后实现最小二乘法拟合的函数。若需要拟合 n 次多项式, 首先写出矩阵 A , A 的大小为 $m \times (n+1)$, 其中 m 是数据点的个数, 且 $A_{ij} = x[i]^j$, 其中 i, j 都从 0 开始计数。之后通过高斯消元解方程 $A^T A \theta = A^T b$, 其中 b 是数据点的函数值组成的向量。解出的 θ 即为拟合的多项式的系数。
4. 计算误差, 以及通过 Mathematica 进行画图。

3 程序清单

程序清单见表 2

文件	描述
code_lab2.cc	通过最小二乘法进行多项式拟合的源代码
plot.nb	通过 Mathematica Notebook 进行函数作图的源代码

表 2: 程序清单

4 运行结果

首先给出通过最小二乘法进行一次，二次，三次的曲线拟合的误差。其中 x, f, val, err 分别表示数据点，那个数据点的值，得到的多项式在那个数据点的值，以及误差。一次的误差见表 3，二次的误差见表 4，三次的误差见表 5。

x	f	val	err
-1.00	-4.4670	-2.958821	1.508179
-0.50	-0.4520	-1.956643	1.504643
0.00	0.5510	-0.954464	1.505464
0.50	0.0480	0.047714	0.000286
1.00	-0.4470	1.049893	1.496893
1.50	0.5490	2.052071	1.503071
2.00	4.5520	3.054250	1.497750

表 3: 一次多项式拟合

x	f	val	err
-1.00	-4.4670	-2.963524	1.503476
-0.50	-0.4520	-1.956643	1.504643
0.00	0.5510	-0.951643	1.502643
0.50	0.0480	0.051476	0.003476
1.00	-0.4470	1.052714	1.499714
1.50	0.5490	2.052071	1.503071
2.00	4.5520	3.049548	1.502452

表 4: 二次多项式拟合

之后给出一次，二次，三次多项式拟合的图像，以及散点图。通过观察一次多项式与二次多项式的系数可以知道，一次多项式与二次多项式及其相似，所以在 $[-5, 5]$ 这个比较小的范围，两个函数图像基本重合。函数图像与散点图见图 1。

通过误差表格以及函数图像，可以直观的感受三次多项式拟合情况最好，而一次二次函数拟合效果并不是那么好。这是因为原来的数据点基本符合一个三次函数。详细的一次，二次，三次多项式的

x	f	val	err
-1.00	-4.4670	-4.466190	0.000810
-0.50	-0.4520	-0.453976	0.001976
0.00	0.5510	0.551024	0.000024
0.50	0.0480	0.051476	0.003476
1.00	-0.4470	-0.449952	0.002952
1.50	0.5490	0.549405	0.000405
2.00	4.5520	4.552214	0.000214

表 5: 三次多项式拟合

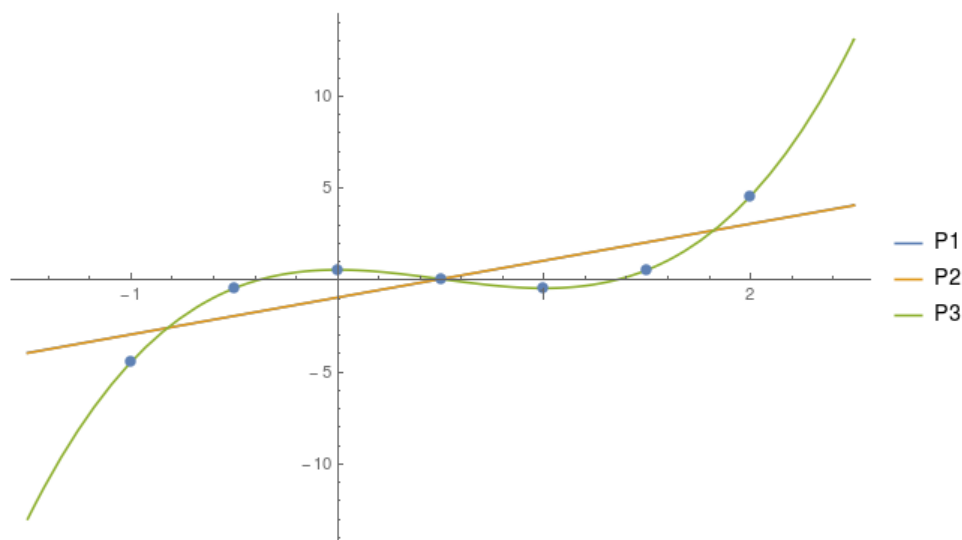


图 1: 一次，二次，三次多项式拟合的图像

表达式以及误差的数据请见 ls.out。

5 体会与展望

通过最小二乘法拟合曲线的实验，更加深入地理解了最小二乘法的工作原理，并且实现了一些相应的模块，实现了高斯消元法等基础算法。

其实最小二乘法拟合曲线依然会出现类似于 Lagrange 插值的病态性质，尤其是当需要拟合的函数次数较高时。希望以后可以多做一些有关过拟合与欠拟合之间 tradeoff 的实验。