

**2020年春季学期  
计算学部《机器学习》课程**

**Lab 1实验报告**

|  |  |
| --- | --- |
| 姓名 | 周牧云 |
| 学号 | 1180300315 |
| 班号 | 1803501 |
| 电子邮件 | zhou\_mu\_yun@163.com |
| 手机号码 | 13912263240 |

**目录**

[1 2](#_Toc51916153)

[1.1 2](#_Toc51916154)

[2 2](#_Toc51916155)

[3 2](#_Toc51916156)

建议写出：问题的描述，解决问题的思路，实验的做法，实验结果的分析，结论，自拟标题

# 实验目的

掌握最小二乘法求解（无惩罚项的损失函数）、掌握加惩罚项（2范数）的损失函数优化、梯度下降法、共轭梯度法、理解过拟合、克服过拟合的方法(如加惩罚项、增加样本)

# 实验要求

1. 生成数据，加入噪声；

2. 用高阶多项式函数拟合曲线；

3. 用解析解求解两种loss的最优解（无正则项和有正则项）

4. 优化方法求解最优解（梯度下降，共轭梯度）；

5. 用你得到的实验数据，解释过拟合。

6. 用不同数据量，不同超参数，不同的多项式阶数，比较实验效果。

7. 语言不限，可以用matlab，python。求解解析解时可以利用现成的矩阵求逆。梯度下降，共轭梯度要求自己求梯度，迭代优化自己写。不许用现成的平台，例如pytorch，tensorflow的自动微分工具。

# 三、实验内容

**1、算法原理**

本实验需要用多项式来拟合正弦函数。在m阶多项式中，有m+1个待定系数，m+1个系数（由低到高）组成的（列）向量记作w。要确定w，用最小二乘法。

设E(w) = 1/2 \* (Xw – Y)T(Xw – Y)，其中，X为多项式中各个未知项代入观测数据求得的矩阵，若记Xi为X的第i行的向量，则Xi[j]为第i个观测数据xi的j次方，记有n组观测数据，多项式最高次为m，易知X的维度为n \* (m+1)。Y为观测标签向量。即Y[j]为第j组观测数据的标签值（即y值）。从而问题转化为：求向量w，使得E(w)最小。

若不加入正则项，令损失函数导数为零，求w

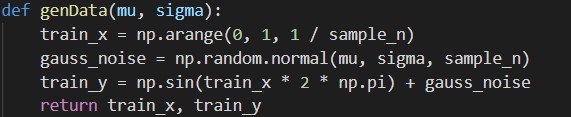
若加入正则项，令损失函数导数为零，求w

加入正则项，对损失函数用梯度下降，当损失函数收敛时，求w

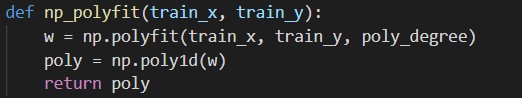
加入正则项，对损失函数用共轭梯度法，循环迭代m+1次，求w

**2、算法实现**

1）生成数据，加入噪声

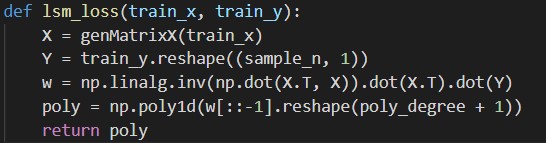


2）用高阶多项式函数拟合曲线；

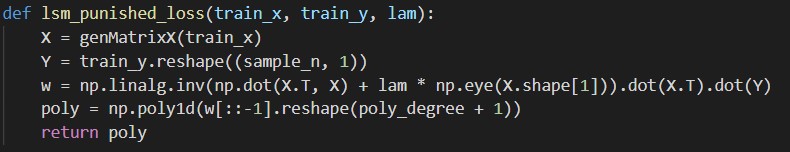


1. 用解析解求解两种loss的最优解（无正则项和有正则项）

无正则项：

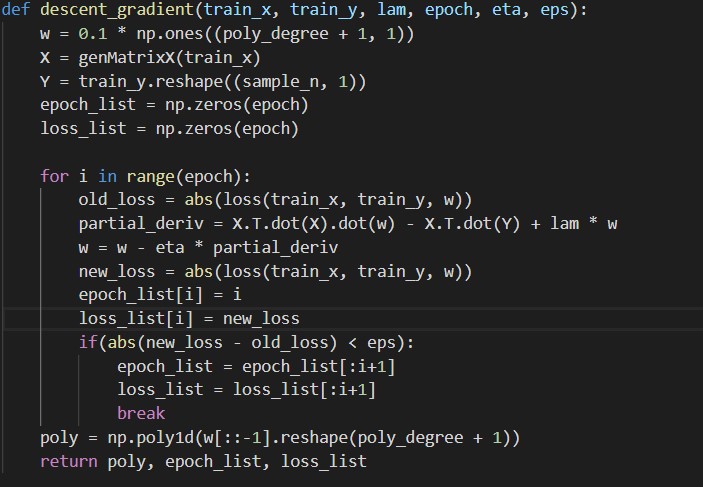


有正则项：

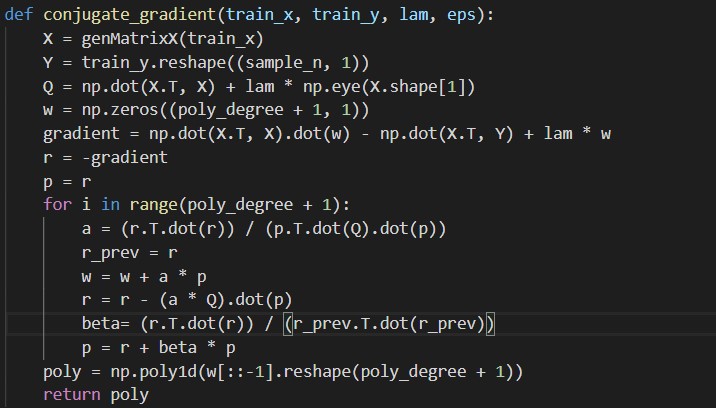


4）优化方法求解最优解（梯度下降，共轭梯度）

梯度下降法：

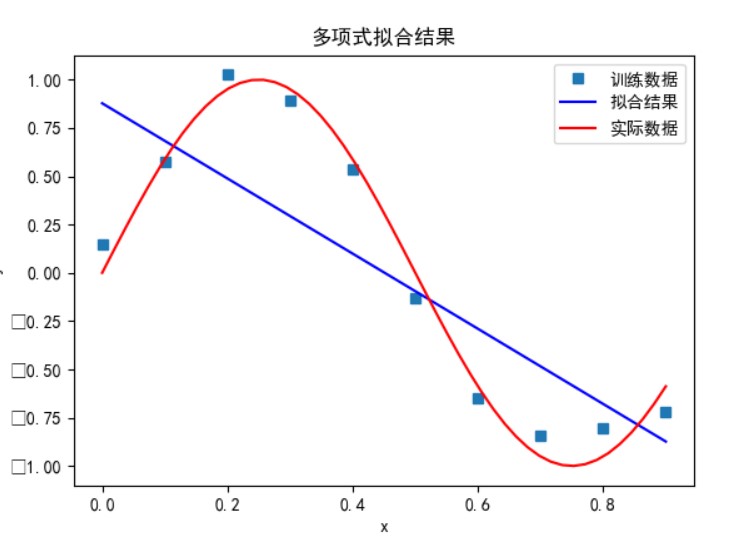


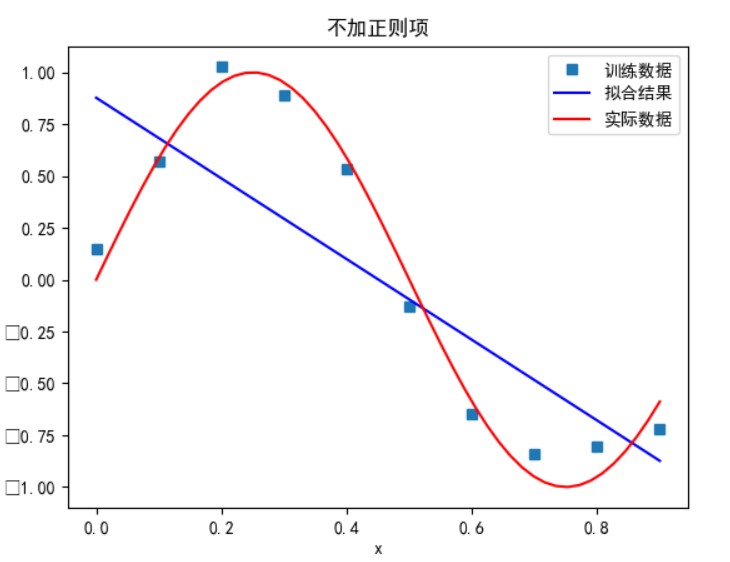
共轭梯度法：

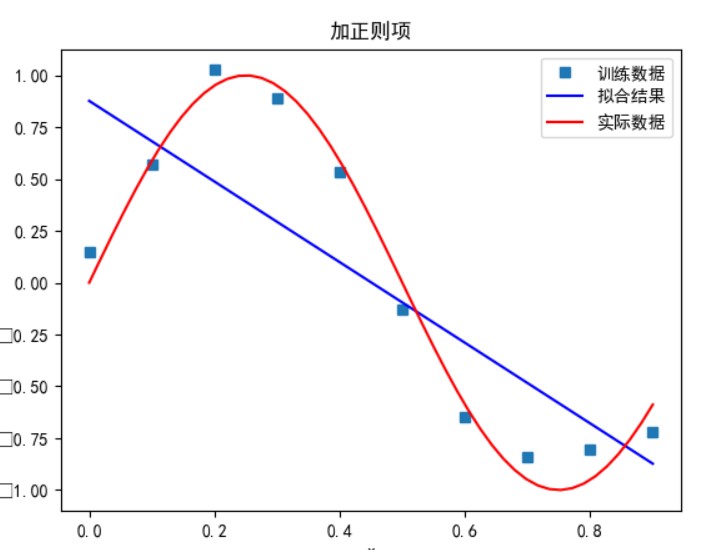


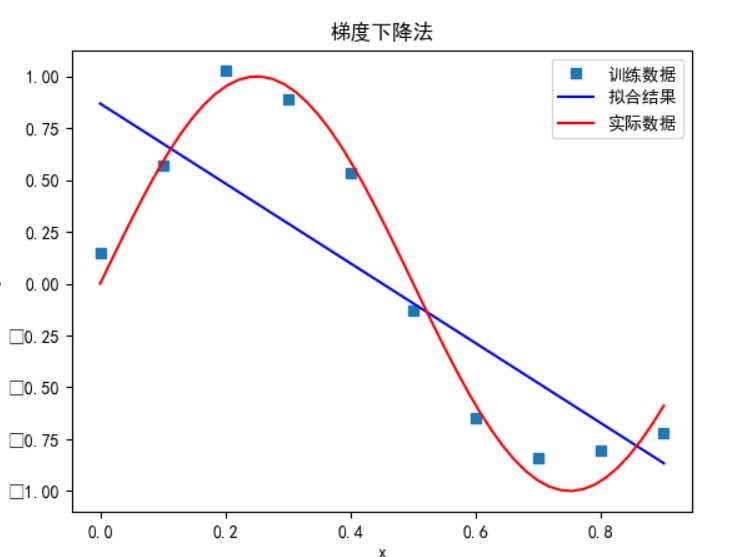
5）用你得到的实验数据，解释过拟合。

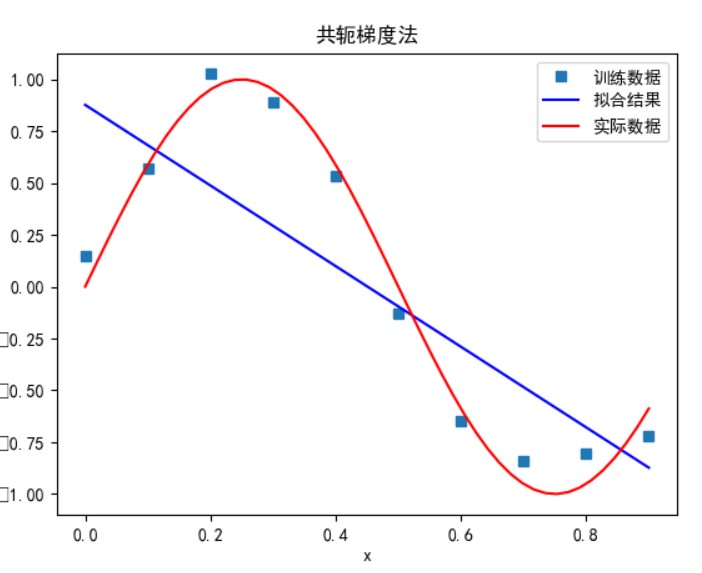
多项式次数为1时：



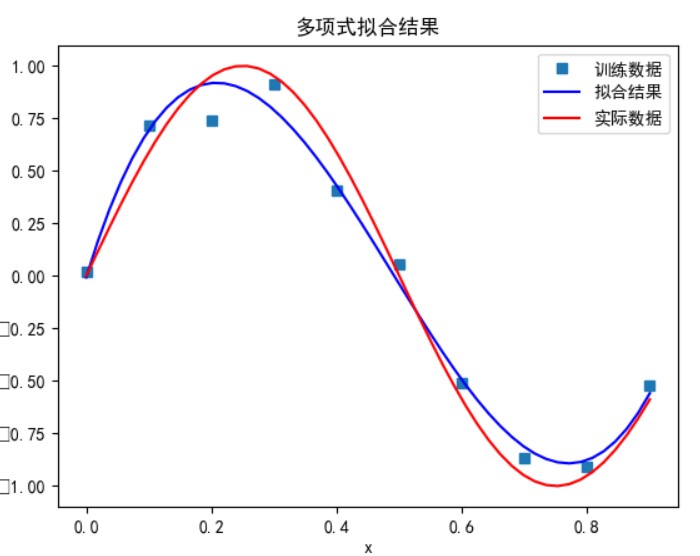


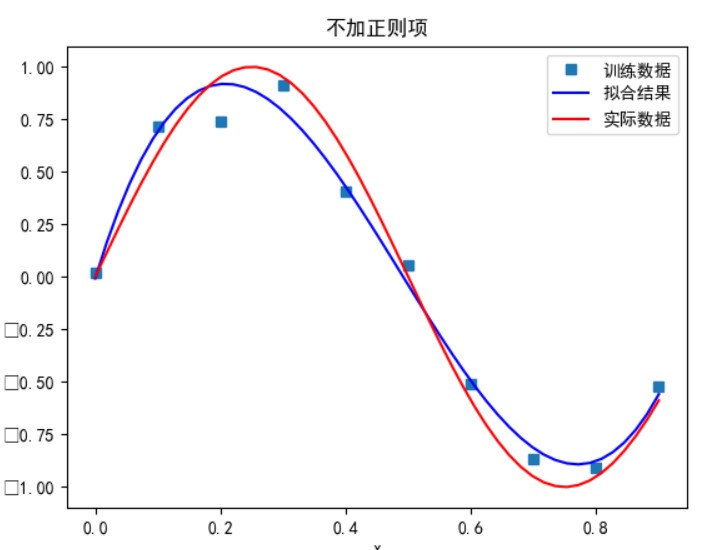


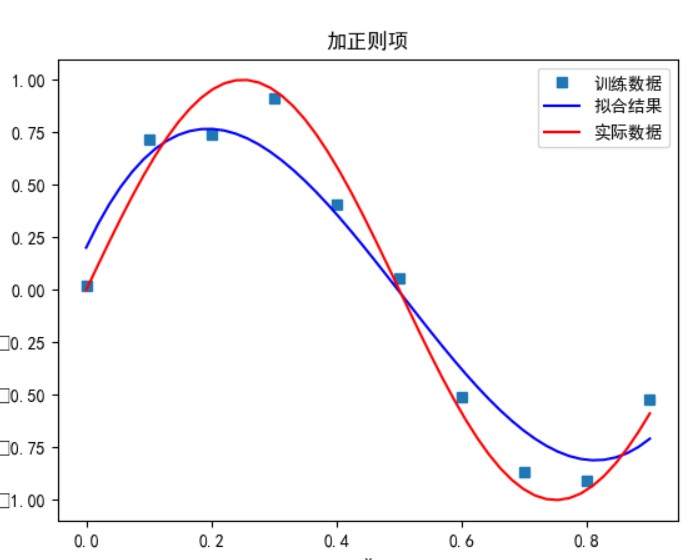


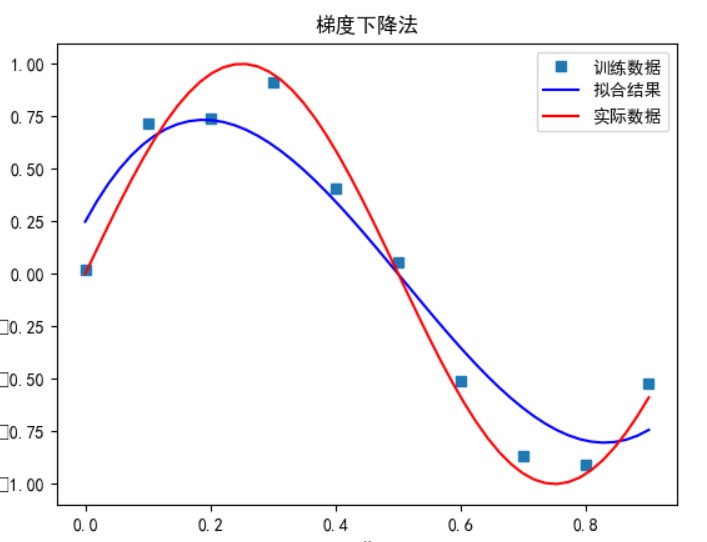


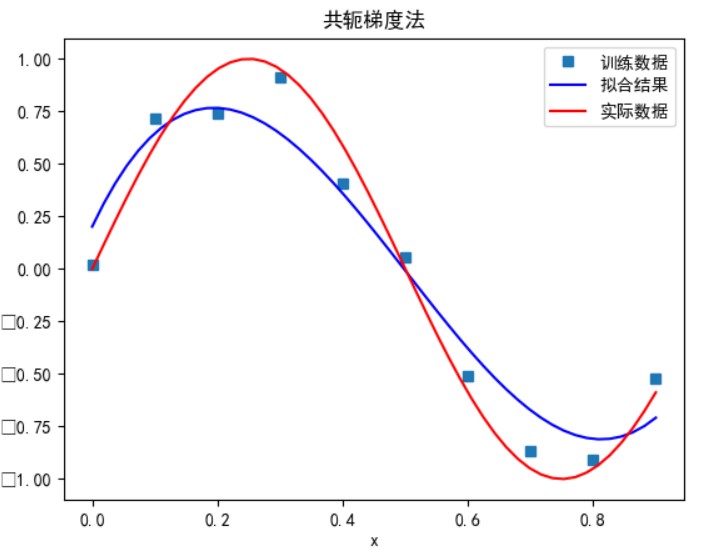
多项式次数为3时：



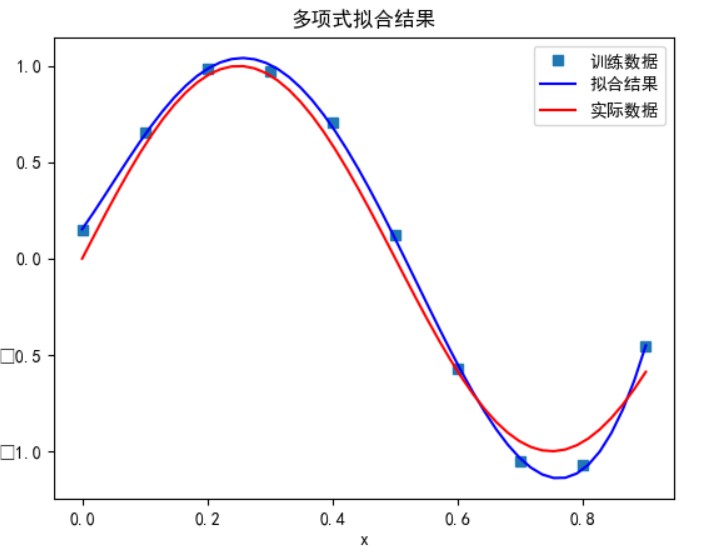


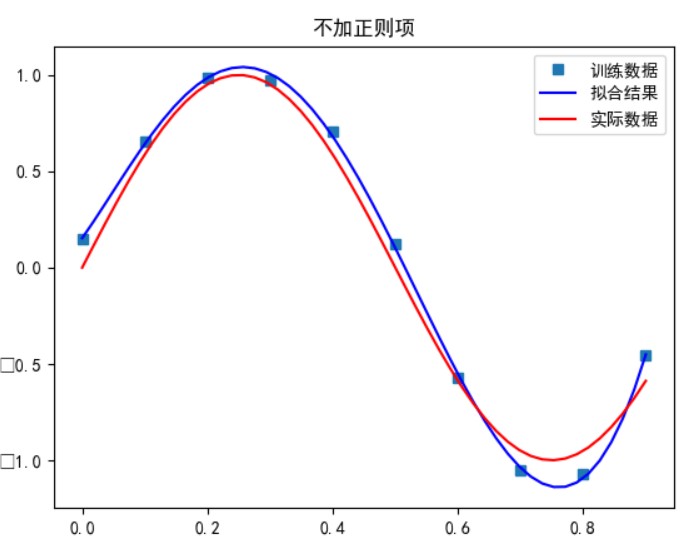


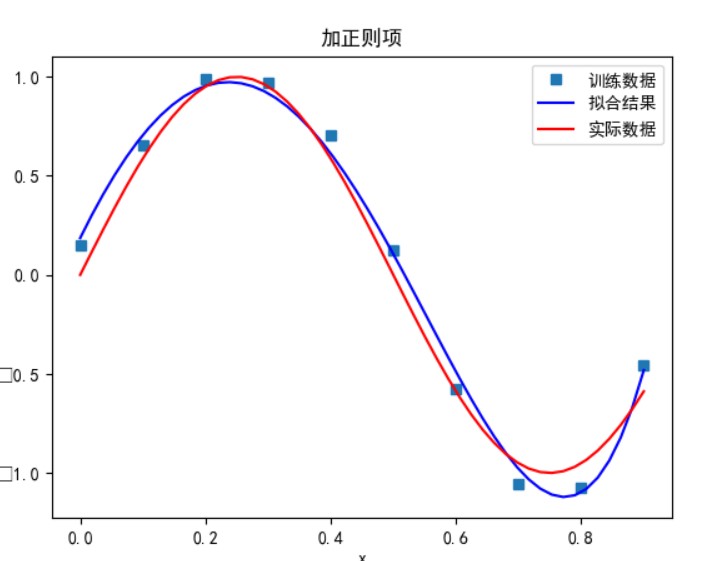


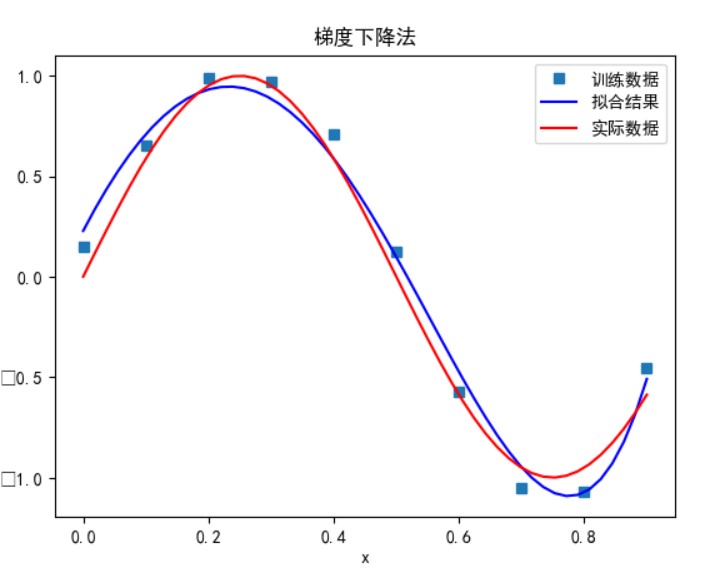


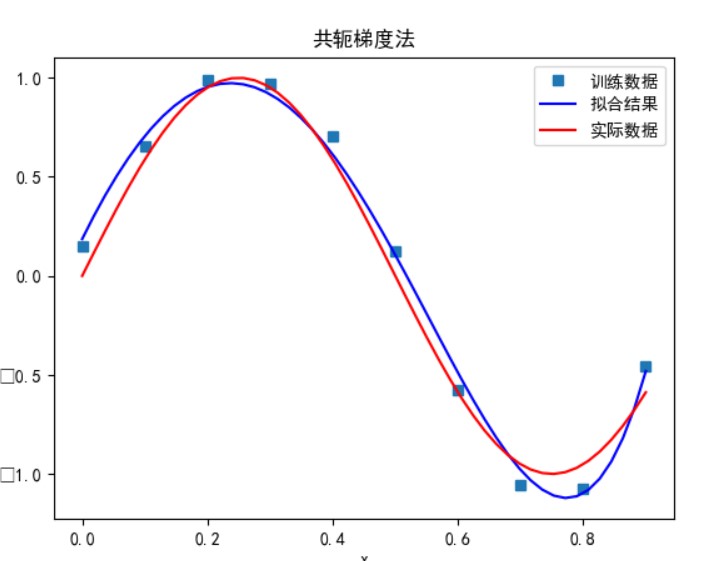
当多项式次数为5时：



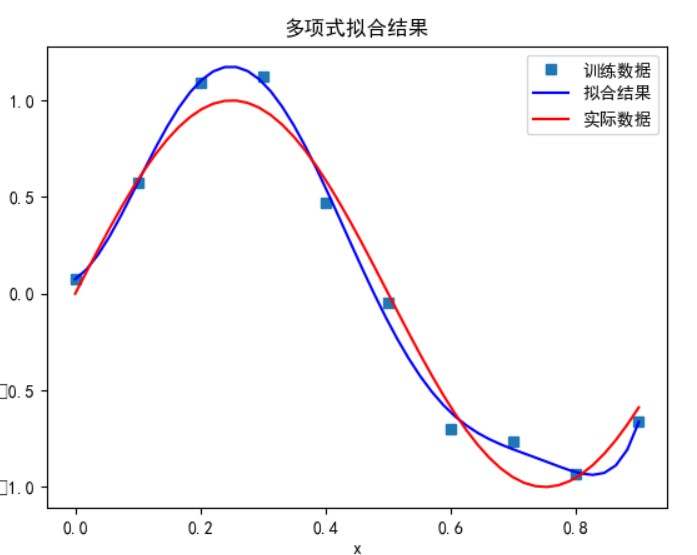


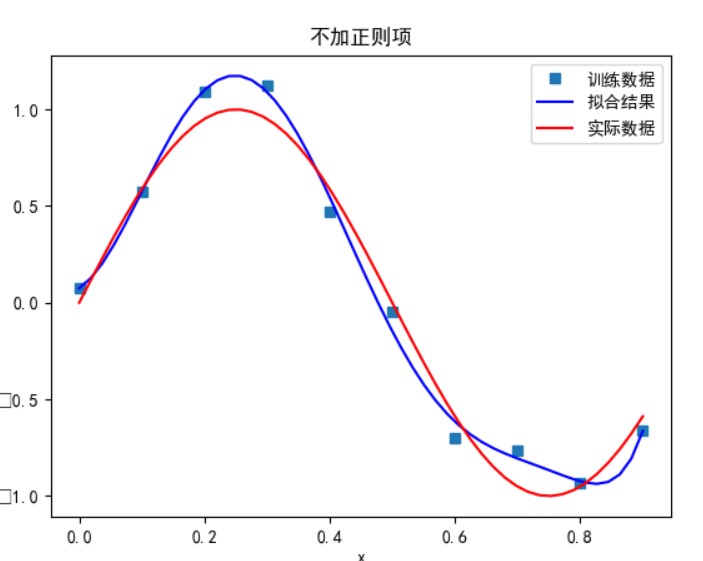


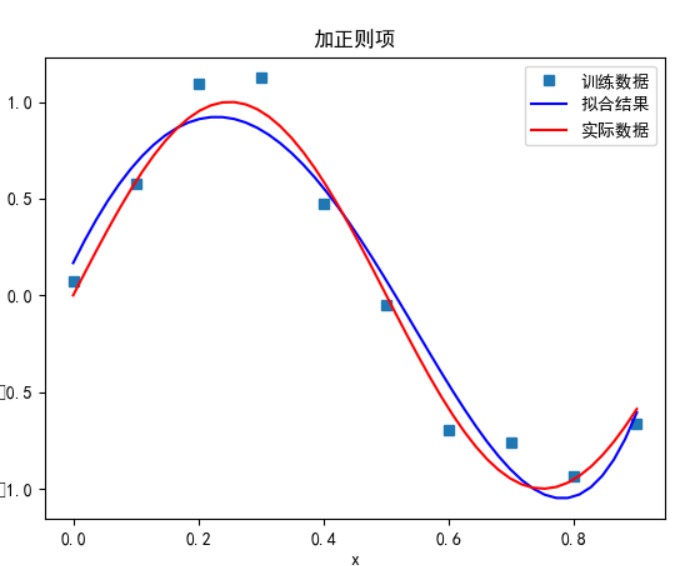


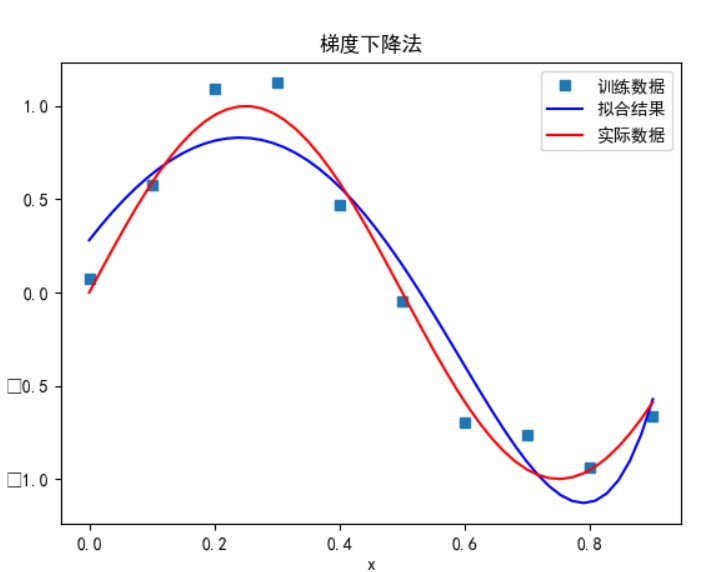


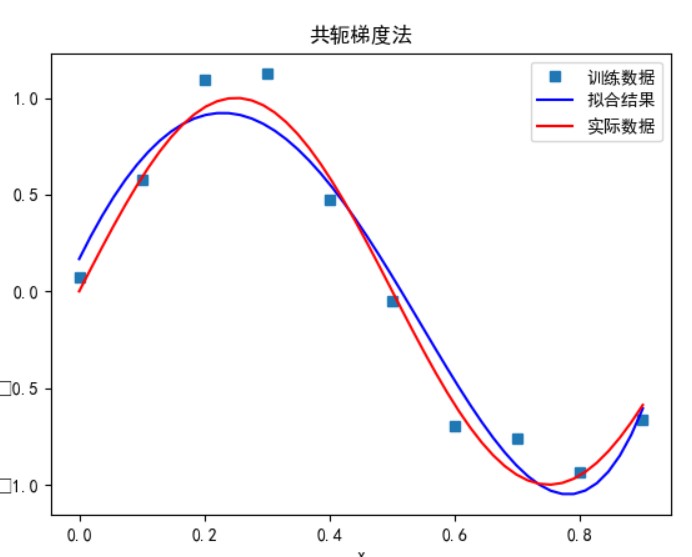
当多项式次数为7时：



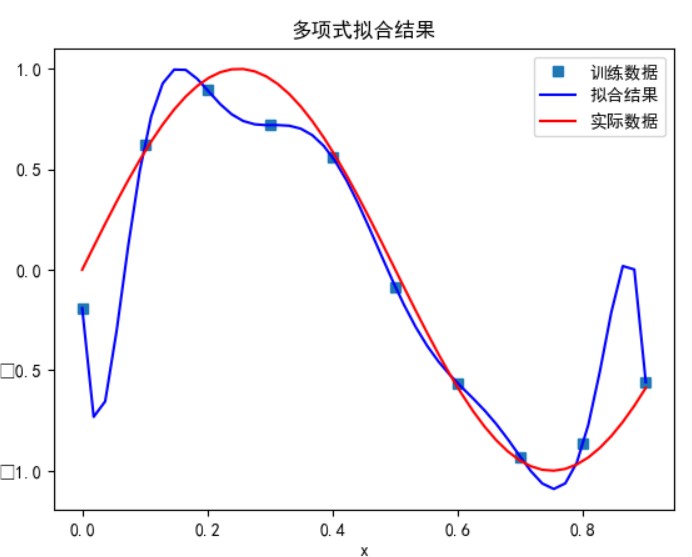


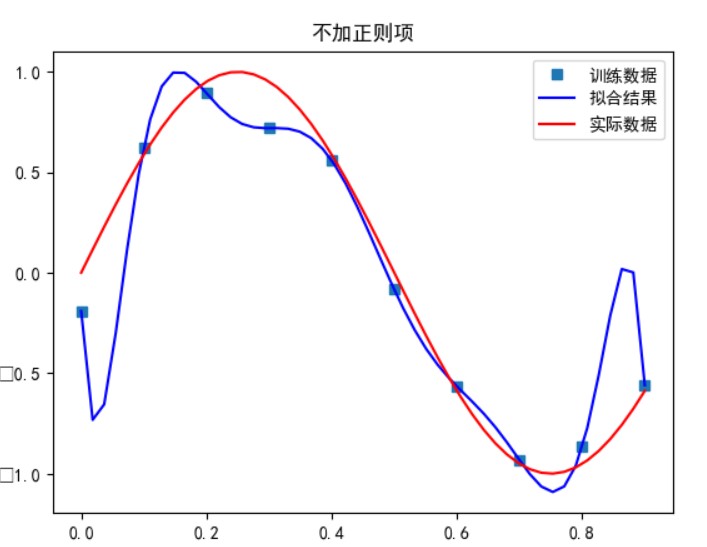


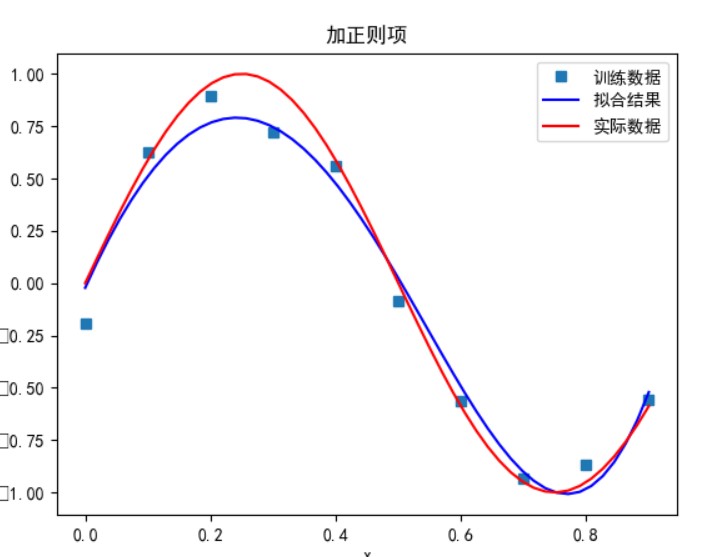


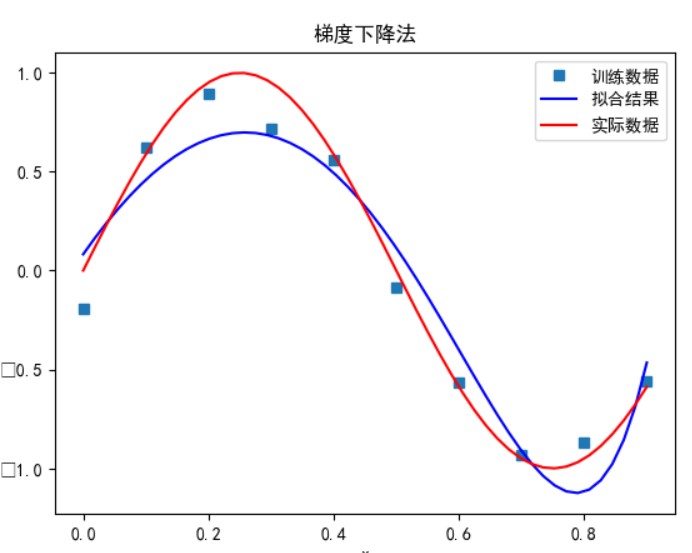


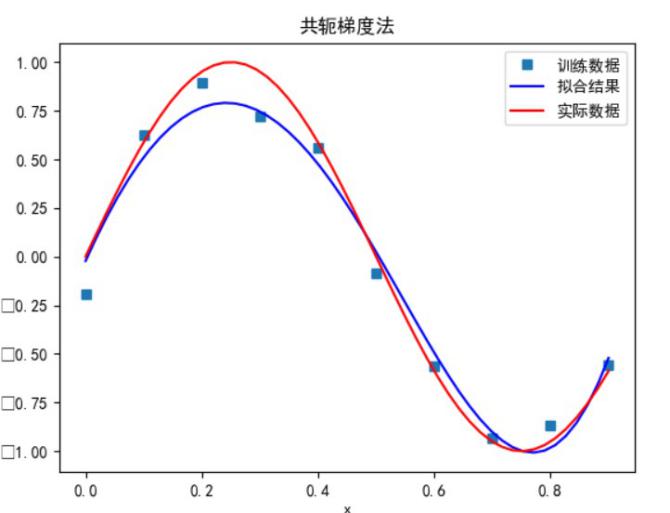
当多项式次数为9时：









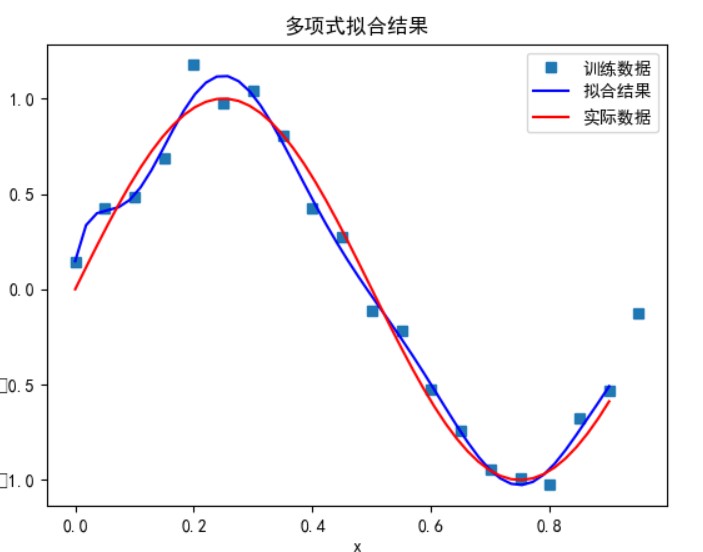


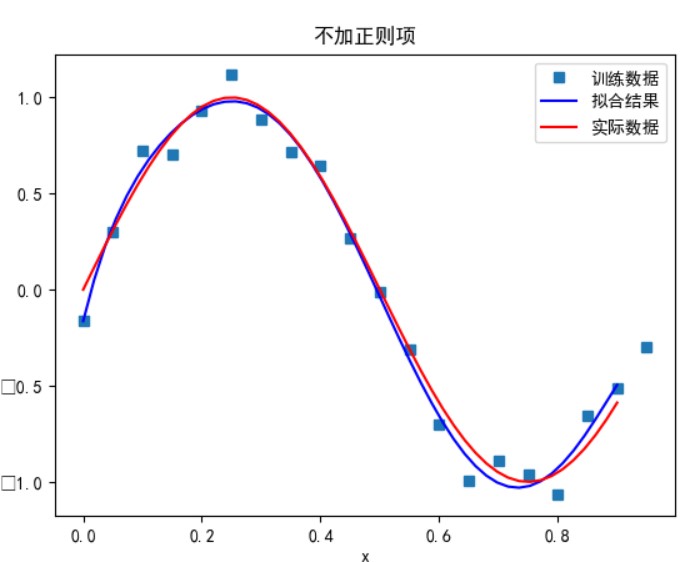
可以看到，多项式次数并不是越高拟合的效果就越好。

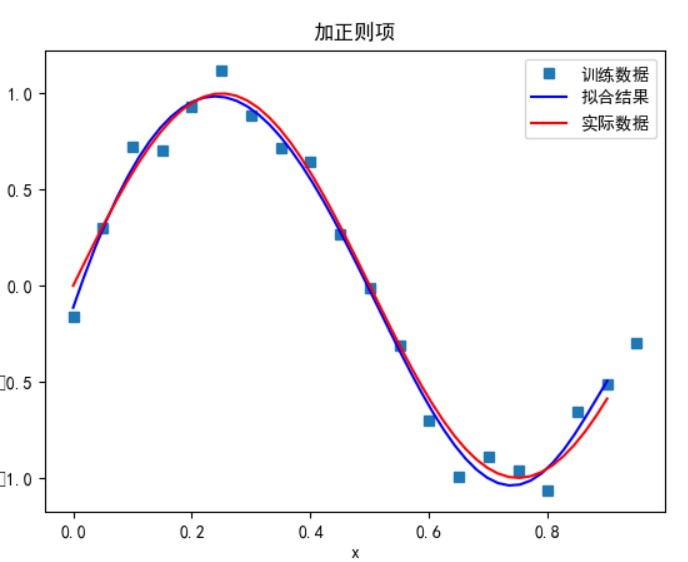
次数太小时无法正确拟合，次数在5时拟合的效果比较好，但再之后多项式次数过大时拟合的效果反而会变差，也就是出现了过拟合的情形。过拟合是由于样本数量过少，模型能力过强导致的。此时如果没有采用加入正则项等优化处理时拟合效果会非常差。

6）用不同数据量，不同超参数，不同的多项式阶数，比较实验效果。

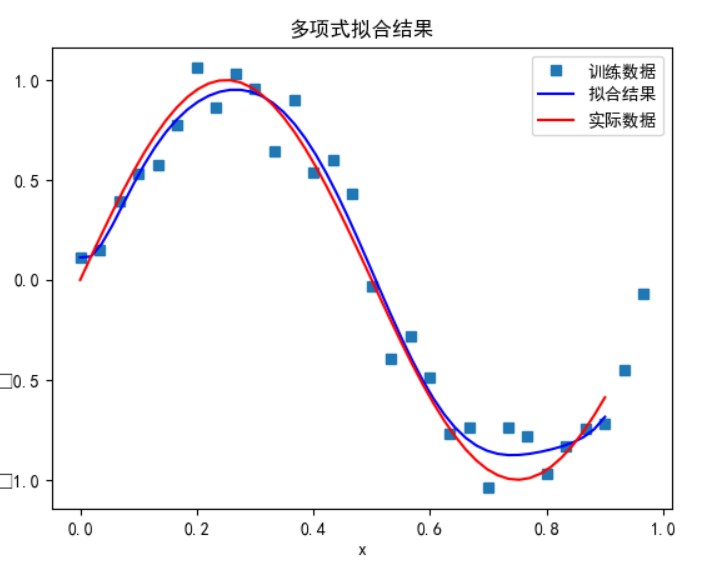
数据量为20，多项式次数为9，超参数为0.0001时：

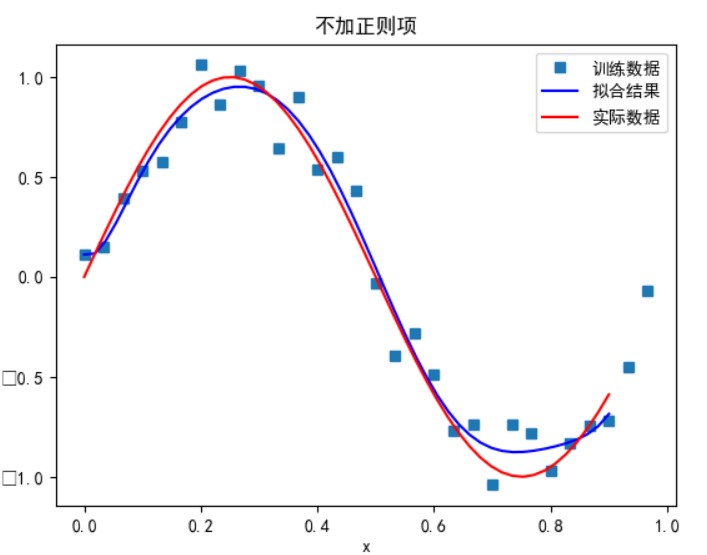


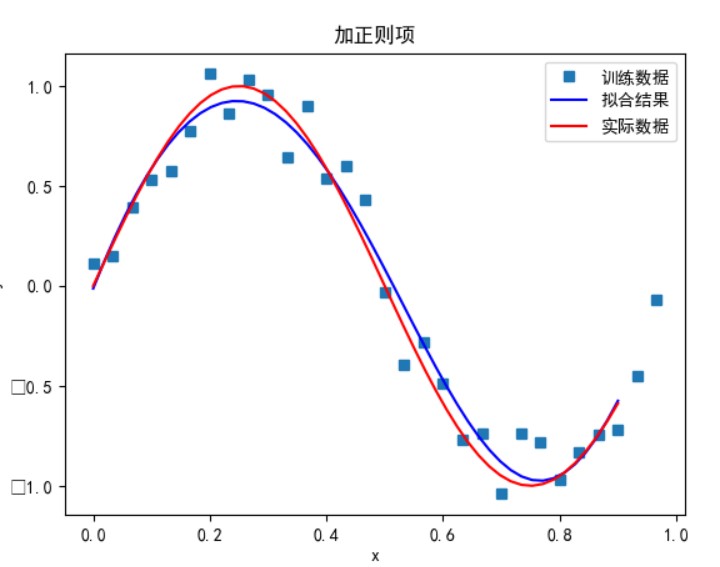




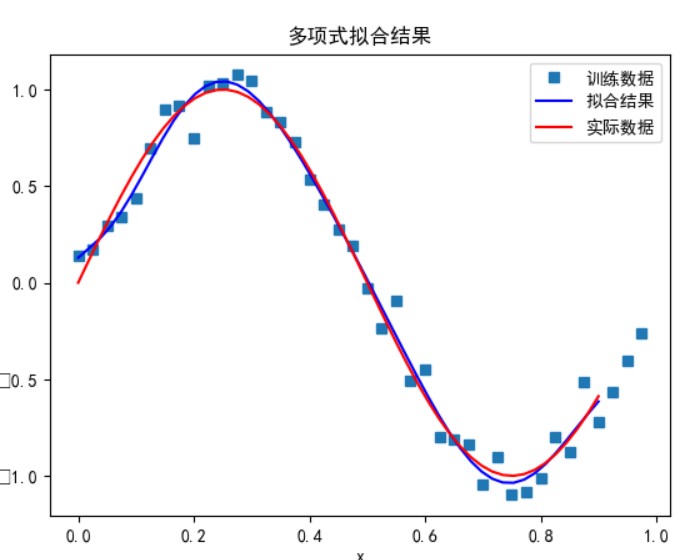
数据量为30，多项式次数为9，超参数为0.0001时：

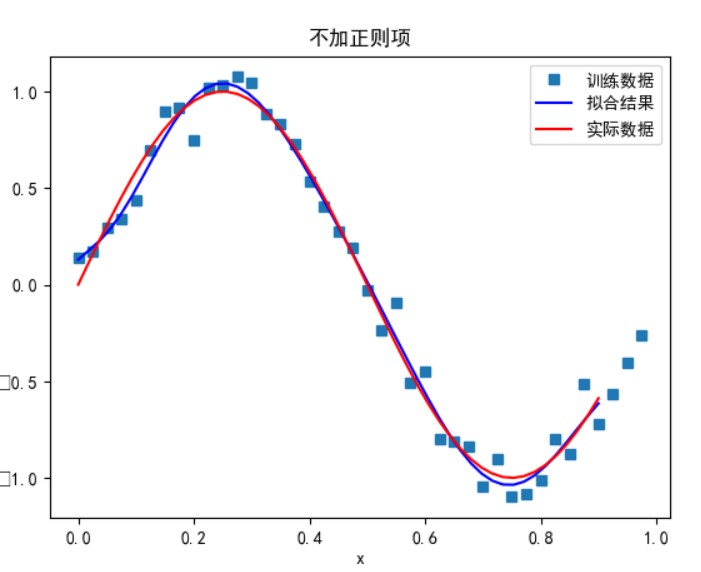


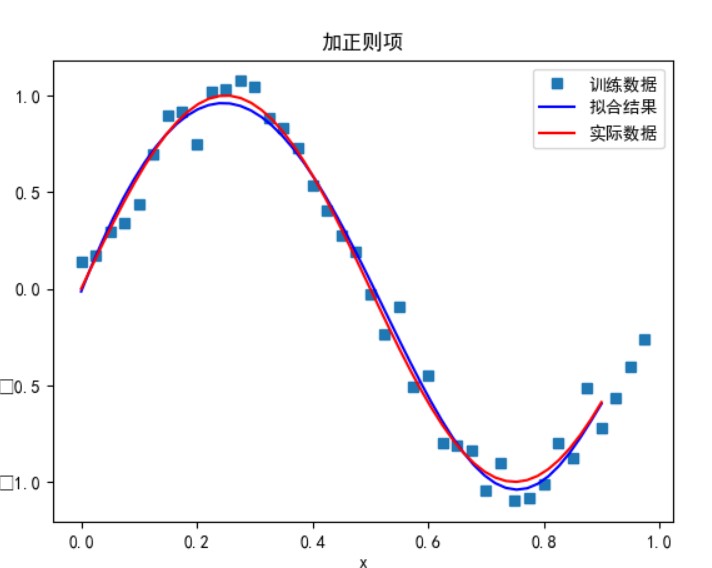




数据量为40，多项式次数为9，超参数为0.0001时：

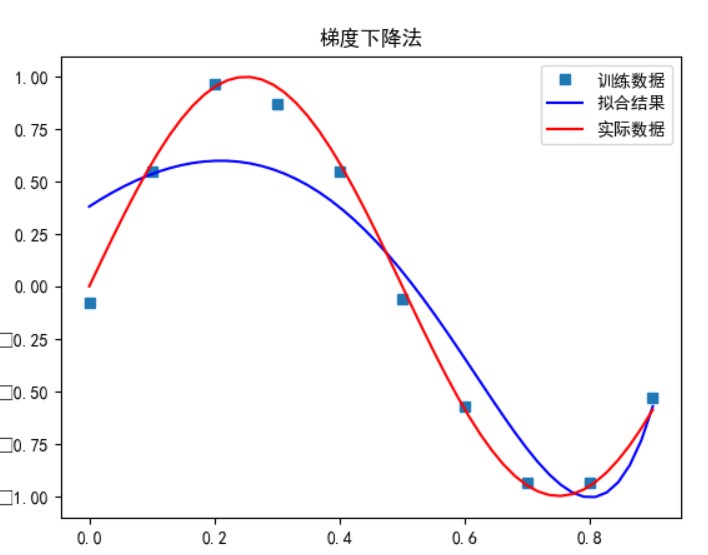




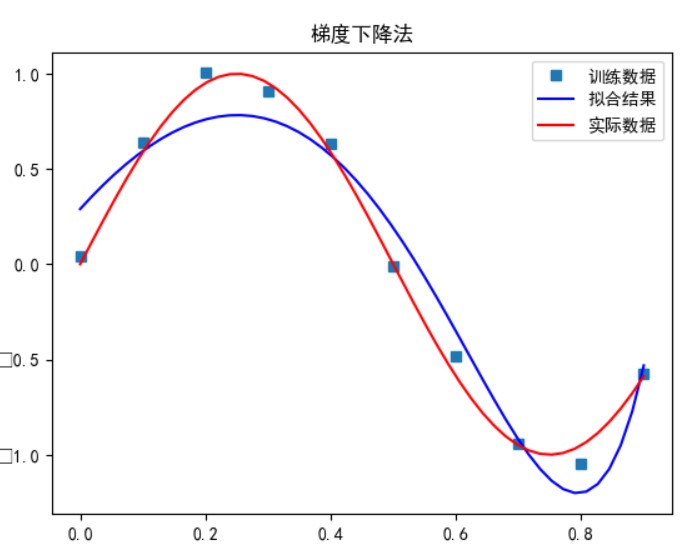


可以看到，相同条件下样本数据越多，拟合效果越好

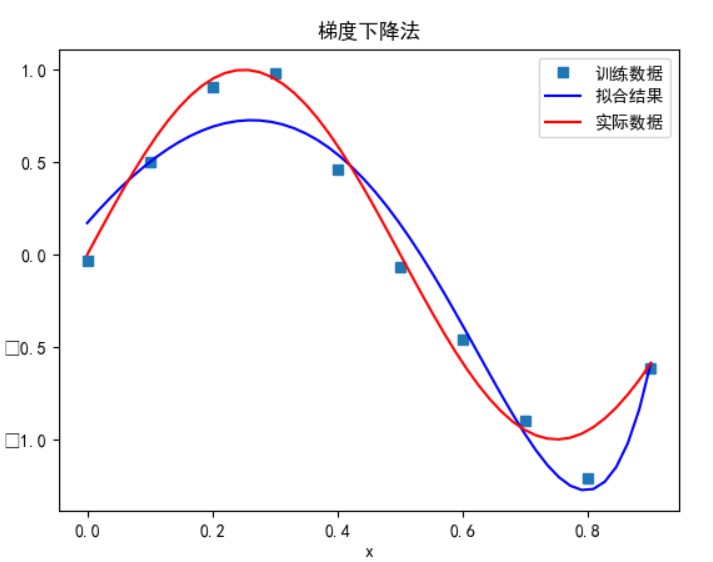
数据量为40，多项式次数为9，超参数为0.01时：



数据量为40，多项式次数为9，超参数为0.001时：



数据量为40，多项式次数为9，超参数为0.0001时：



可以看到超参数对拟合效果影响很有限