

作业一：曲线拟合

实验目的：实现多项式曲线拟合

实验过程：

实验题目中要求对 $y=\sin(x)$ 曲线离散取点，并加入高斯噪音，在实验中通过 `numpy` 库的 `random` 函数实现：

```
y2 = y1 + 0.4*np.random.randn(len(x1))
```

对于离散分布的点可以用一条多项式曲线拟合：

$$y(x, w) = w_0 + w_1x + w_2x^2 + \dots + w_Mx^M$$

在实验中根据 `degree` 的不同，分别采用不同的函数构造，代码如下：

```
def func_d3(x,p):
    w0, w1, w2, w3 = p
    return w0+w1*x+w2*x**2+w3*x**3
def func_d9(x,p):
    w0, w1, w2, w3, w4, w5, w6, w7, w8, w9= p
    return w0+w1*x+w2*x**2+w3*x**3+w4*x**4+w5*x**5
    +w6*x**6+w7*x**7+w8*x**8+w9*x**9
```

在具体的实现过程中使用最小二乘法来拟合曲线，

```
plsq_d3 = leastsq(residuals_d3, p0, args=(y2, x1))
```

其中，`p0` 为拟合参数的初始值，`args` 为需要拟合的实验数据在拟合过程中，使得误差函数

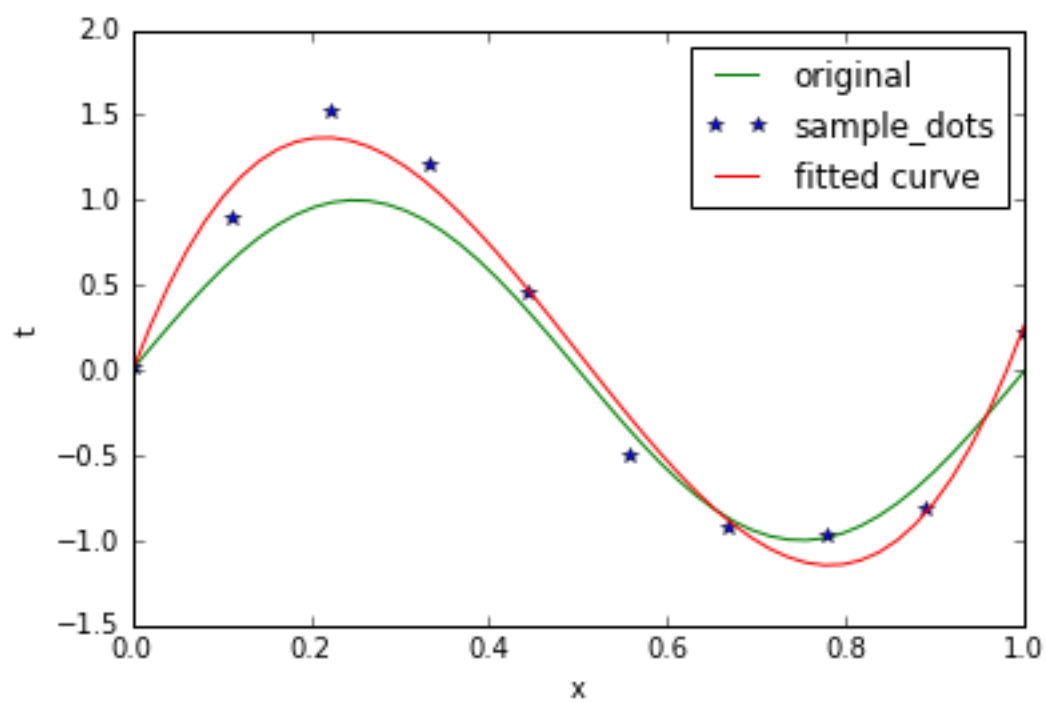
$$E(w) = \sum_{i=1}^n (y(x_i, w) - t(x_i))^2$$

达到最小，根据定义可知误差值即实验数据与拟合函数所得值之间的差。

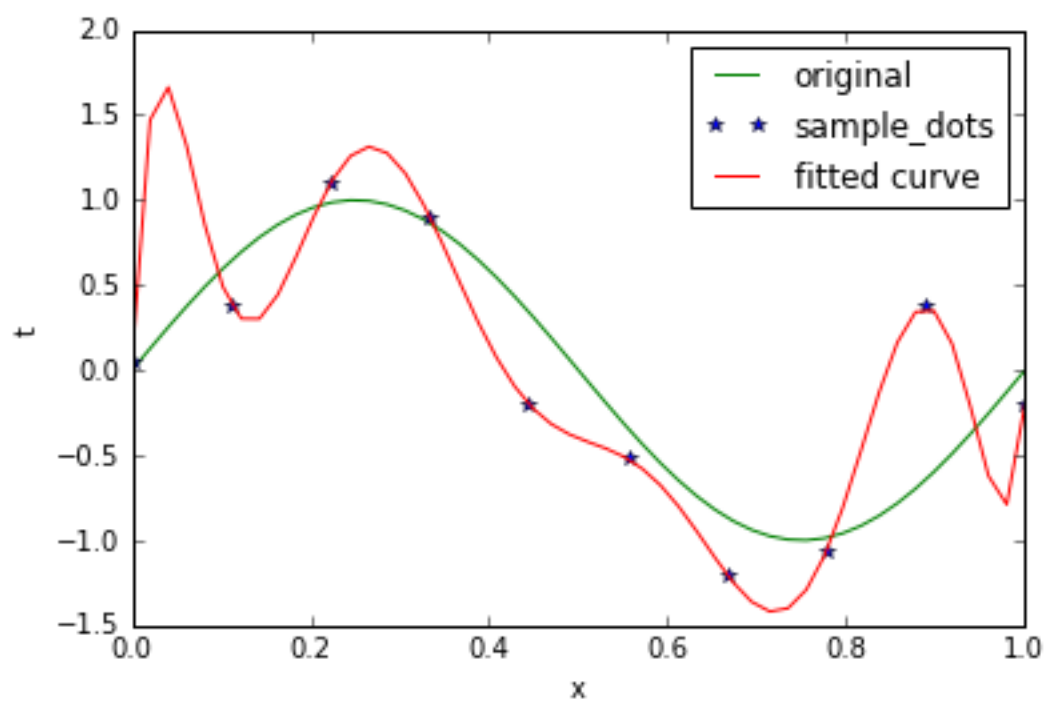
最后，实验要求中要求加入正则项再进行曲线拟合。

$$E(w) = \sum_{i=1}^n (y(x_i, w) - t(x_i))^2 + \lambda \|w\|^2$$

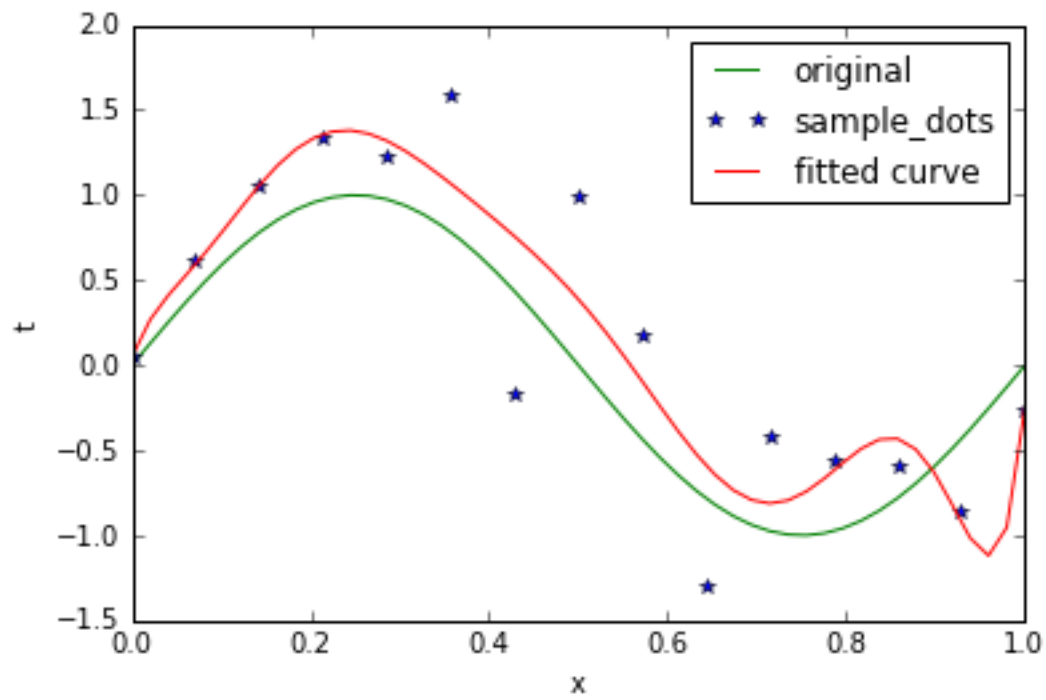
实验结果:



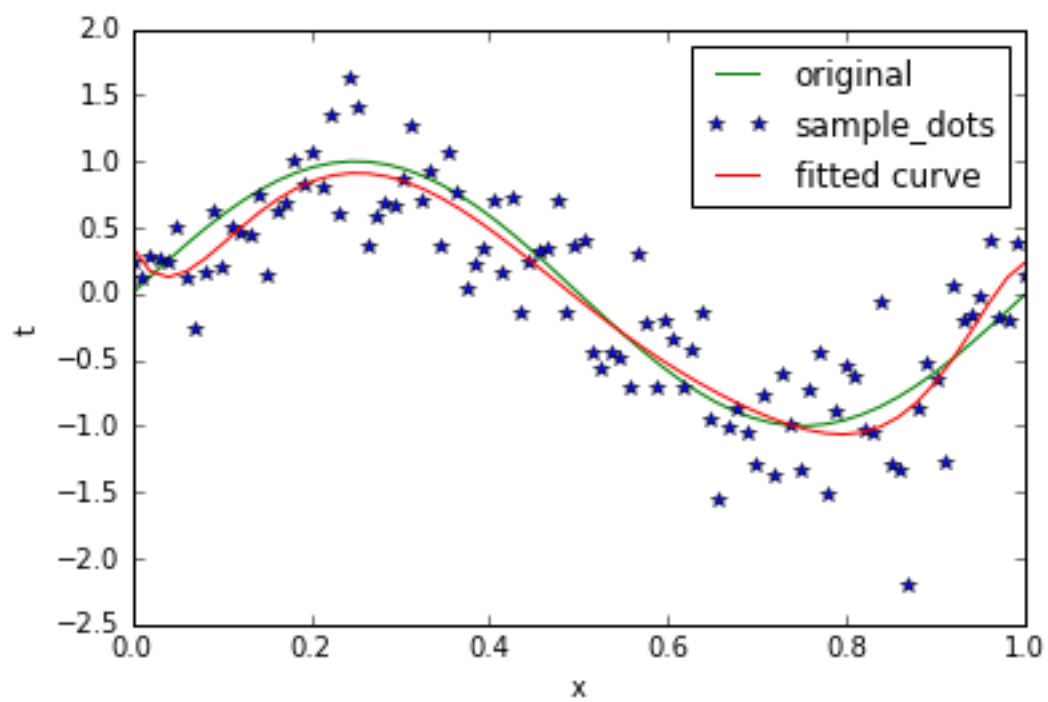
fit degree 3 curves in 10 samples



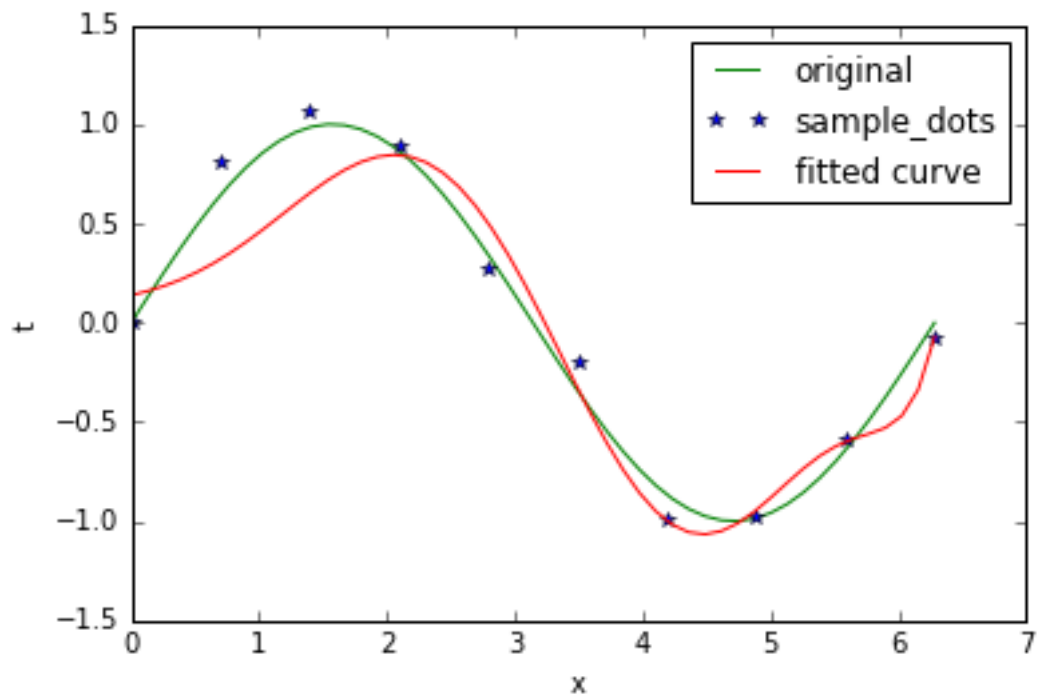
fit degree 9 curves in 10 samples



fit degree 9 curves in 15 samples



fit degree 9 curves in 100 samples



fit degree 9 curve in 10 samples but with regularization term

实验小结：从实验结果中可以看出，在相同 sample 数目的情况下，degree 越小拟合效果越好；在相同 degree 的情况下，sample 数目越多，拟合效果越好；在相同 sample 和 degree 的情况下，带有正则项时拟合效果更好。