

Homework2

付宇辉 无 05 2020010768

i. 记

$$\mathbf{X} = \begin{bmatrix} \widehat{x}_1 \\ \widehat{x}_2 \\ \dots \\ \widehat{x}_N \end{bmatrix}, \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_N \end{bmatrix}$$

所以损失函数可以表示为:

$$L(\widehat{\mathbf{w}}; \mathbf{X}, \mathbf{y}) = \frac{1}{N} \|\mathbf{y} - \mathbf{X}\widehat{\mathbf{w}}\|^2$$

$$\widehat{\mathbf{w}}^* = \arg \min_{\widehat{\mathbf{w}}} \frac{1}{N} \|\mathbf{y} - \mathbf{X}\widehat{\mathbf{w}}\|^2$$

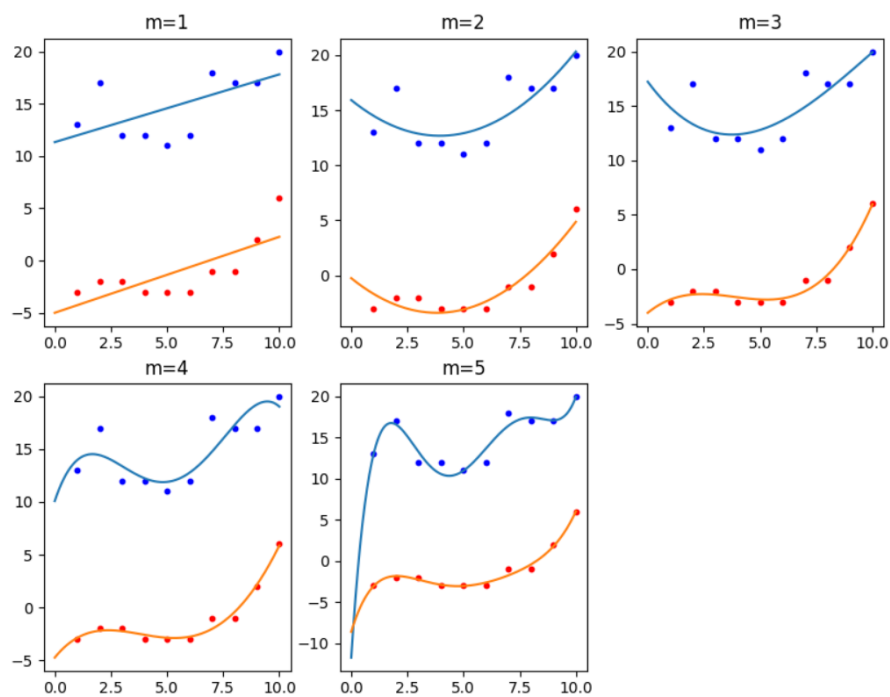
对损失函数求 $\widehat{\mathbf{w}}$ 求偏导得到:

$$\frac{\partial L}{\partial \widehat{\mathbf{w}}} = -\frac{2}{N} \mathbf{X}^T (\mathbf{y} - \mathbf{X}\widehat{\mathbf{w}})$$

令 $\frac{\partial L}{\partial \widehat{\mathbf{w}}} = 0$ 得到:

$$\widehat{\mathbf{w}}^* = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$$

ii.



代码在附件中

w^* 的值为

```
[[11.33333333]
 [ 0.64848485]]

[[10.08333333]
 [ 6.35567211]
 [-2.89262821]
 [ 0.44298757]
 [-0.02083333]]

[[15.91666667]
 [-1.64318182]
 [ 0.20833333]]

[[4.75000000e+00]
 [ 2.63986014e+00]
 [-8.57080420e-01]
 [ 9.17832168e-02]
 [-2.18531469e-03]]

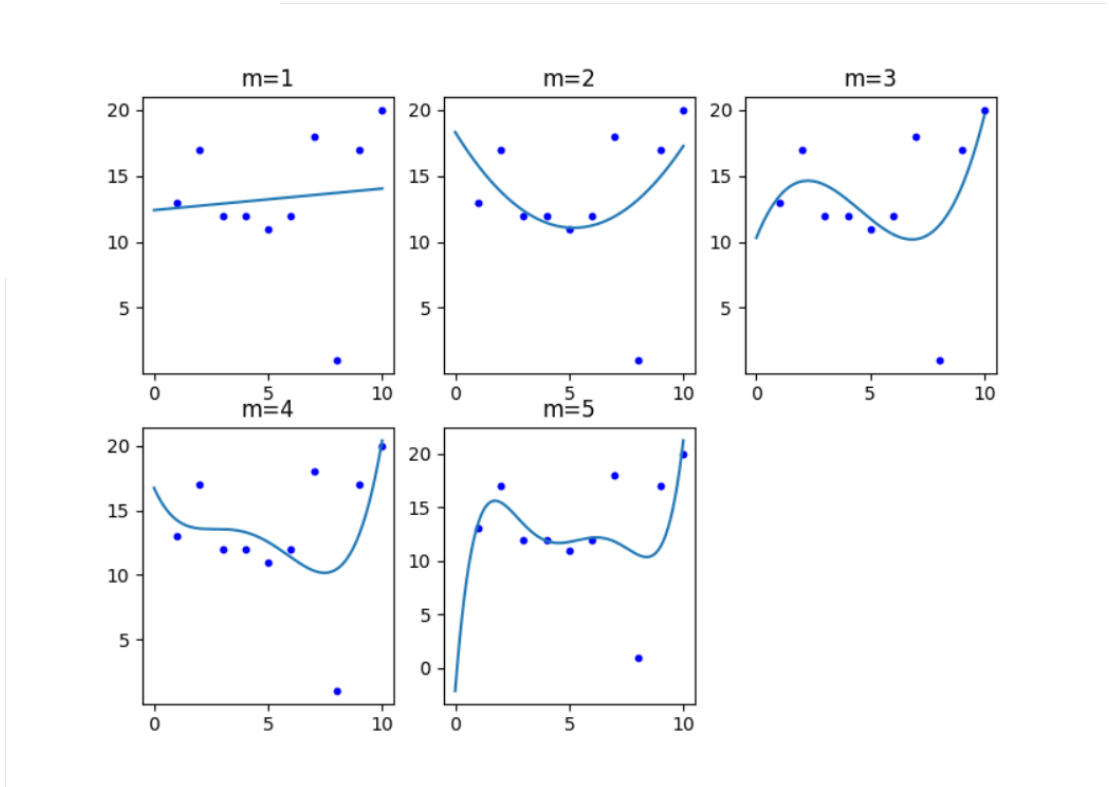
[[1.72333333e+01]
 [-2.81099456e+00]
 [ 4.61538462e-01]
 [-1.53457653e-02]]

[[1.17333333e+01]
 [ 4.23201165e+01]
 [-2.17724359e+01]
 [ 4.66392774e+00]
 [-4.40384615e-01]
 [ 1.52564103e-02]]

[[8.60000000e+00]
 [ 8.98652681e+00]
 [-4.18881119e+00]
 [ 8.36655012e-01]
 [-7.62237762e-02]
 [ 2.69230769e-03]]

[[1.67832168]
 [-0.50524476]
 [ 0.04370629]]
```

iii.



降低异常点对结果影响可能的方法：

1. 使用比较简单的模型进行拟合，可以降低异常点对结果的影响；
2. 先进行一步筛选，剔除样本中明显不合理的点；
3. 增加采样点（数据）可以减少异常点的影响。