Matrix Analysis Homework

Yuxun Qu

October 2019

9.Using least squares techniques, fit the following data

x	-5	-4	-3	-2	-1	0	1	2	3	4	5
y	2	7	9	12	13	14	14	13	10	8	4

with a line $y = \alpha_0 + \alpha_1 x$ and then fit the data with a quadratic $y = \alpha_0 + \alpha_1 x + \alpha_2 x^2$.Determine which of these two curves best fits the data by computing the sum of the squares of the errors in each case.

Solution:

对于线性拟合而言，令

$$\mathbf{X} = [\alpha_0, \alpha_1]^T, \mathbf{A} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ ... & ... \\ 1 & x_{11} \end{bmatrix} = \begin{bmatrix} 1 & -5 \\ 1 & -4 \\ ... & ... \\ 1 & 5 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} y_1 \\ y_2 \\ ... \\ y_{11} \end{bmatrix} = \begin{bmatrix} 2 \\ 7 \\ ... \\ 4 \end{bmatrix}$$

最小二乘法满足

$$\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$$

$$\begin{bmatrix} 11 & 0 \\ 0 & 110 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 106 \\ 20 \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} \frac{106}{11} \\ \frac{2}{11} \end{bmatrix}$$

所以线性拟合曲线为

$$y = \frac{106}{11} + \frac{2}{11}x$$

均方误差为 e_1 满足

$$e_1 = \sum_{k=1}^{n} (y - a - bx)^2 = \frac{1792}{11}$$

对于二次拟合而言，令

$$\mathbf{X} = [\alpha_0, \alpha_1, \alpha_2]^T, \mathbf{A} = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ ... & ... & ... \\ 1 & x_{11} & x_{11}^2 \end{bmatrix} = \begin{bmatrix} 1 & -5 & 25 \\ 1 & -4 & 16 \\ ... & ... & ... \\ 1 & 5 & 25 \end{bmatrix}, \mathbf{b} = \begin{bmatrix} y_1 \\ y_2 \\ ... \\ y_{11} \end{bmatrix} = \begin{bmatrix} 2 \\ 7 \\ ... \\ 4 \end{bmatrix}$$

最小二乘法满足

$$\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b}$$

$$\begin{bmatrix} 11 & 0 & 110 \\ 0 & 110 & 0 \\ 110 & 0 & 1958 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 106 \\ 20 \\ 688 \end{bmatrix}$$

$$\mathbf{x} = \begin{bmatrix} \frac{1998}{143} \\ \frac{2}{11} \\ \frac{-62}{143} \end{bmatrix}$$

所以二次拟合曲线为

$$y = \frac{1998}{143} + \frac{2}{11}x - \frac{62}{143}x^2$$

均方误差为 e_1 满足

$$e_1 = \sum_{k=1}^{n}(y - a - bx - cx^2)^2 = \frac{232}{143}$$

比较两者可知二次函数的拟合效果更好.