IISEE Lecture Note

Data Processing

By

T. HARA

International Institute of

Seismology and Earthquake Engineering,

Building Research Institute

Contents

Due to copyright and references, only a part of chapters 5 and 9 are shown in this open version.

5. Polynomial Curve Fitting

In this chapter, using the results of Chapter 3 and 4, we fit a polynomial function to a set of given data. Suppose that we have N pairs of data, $(x_i, y_i)(i = 1, N)$. The observational equation is as follows:

$$
\begin{aligned}
y_1 &= a_1 + a_2 x_1 + a_3 x_1^2 + \ldots + a_M x_1^{M-1} + \varepsilon_1 \\
y_2 &= a_1 + a_2 x_2 + a_3 x_2^2 + \ldots + a_M x_2^{M-1} + \varepsilon_2 \\
&\cdot \\
&\cdot \\
y_N &= a_1 + a_2 x_N + a_3 x_N^2 + \ldots + a_M x_N^{M-1} + \varepsilon_N
\end{aligned} \tag{5-0-1}
$$

where ε_i is the error of i - th data.

EXERCISE 5

Let us develop a program to fit a polynomial curve to a set of given data. The structure of the program will be the following:

(1) Determine the order of the polynomial.

(2) Read the data. $\Rightarrow$ Display them.

(3) Construct the matrix $\mathbf{G}$ and the vector $\mathbf{d}$. $\Rightarrow$ Display them.

(4) Calculate the matrix $\mathbf{A}\ (= \mathbf{G}^T\mathbf{G})$ and $\mathbf{b}\ (= \mathbf{G}^T\mathbf{d})$. $\Rightarrow$ Display them.

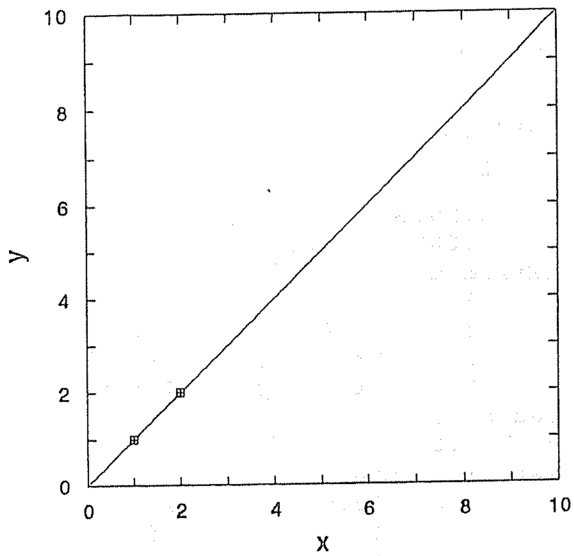
(5) Solve the normal equation, $\mathbf{Ax} = \mathbf{b}$. $\Rightarrow$ Display the solution.

(6) Plot the determined polynomial together with the data.

EXERCISE 5-1

Fit a straight line for the two data: (1.0, 1.0) and (2.0, 2.0). The result will be the following:

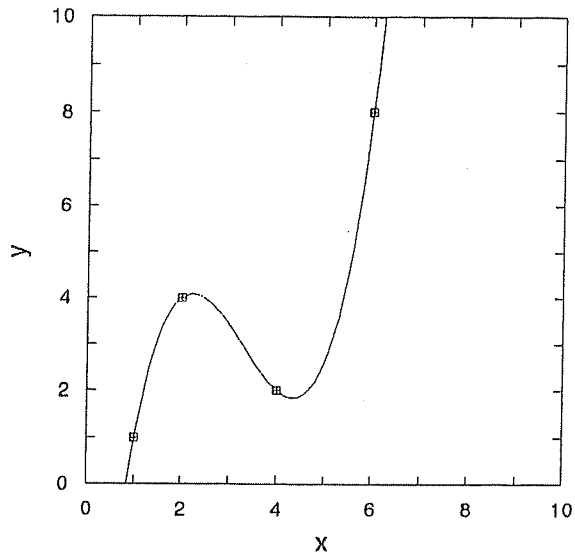
```
Input order:
1
Number of data:   2
Data (x,y):
    1.00000      1.00000
    2.00000      2.00000
Matrix G:
     1.0      1.0
     1.0      2.0
Vector d:
     1.0      2.0
Matrix A(=GtG):
     2.0      3.0
     3.0      5.0
Vector b:
     3.0      5.0
A inverse  :
     5.0     -3.0
    -3.0      2.0
Solution:
     0.0      1.0
```



EXERCISE 5-2

Fit the 3-rd order polynomial to the four data: (1.0, 1.0), (2.0, 4.0), (4.0, 2.0), (6.0, 8.0).

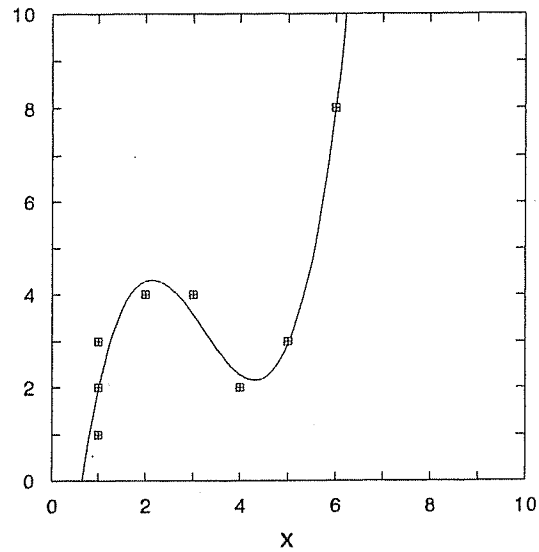
```
Input order:
3
Number of data:  4
Data (x,y):
    1.00000     1.00000
    2.00000     4.00000
    4.00000     2.00000
    6.00000     8.00000
Matrix G:
     1.0     1.0     1.0     1.0
     1.0     2.0     4.0     8.0
     1.0     4.0    16.0    64.0
     1.0     6.0    36.0   216.0
Vector d:
     1.0     4.0     2.0     8.0
Matrix A(=GtG):
     4.0    13.0    57.0   289.0
    13.0    57.0   289.0  1569.0
    57.0   289.0  1569.0  8833.0
   289.0  1569.0  8833.050817.0
Vector b:
    15.0    65.0   337.0  1889.0
A inverse :
    20.3   -23.9     7.5    -0.7
   -23.9    29.6    -9.5     0.9
     7.5    -9.5     3.1    -0.3
    -0.7     0.9    -0.3     0.0
Solution:
    -8.4    13.5    -4.6     0.5
```



EXERCISE 5-3

Fit the 3-rd order polynomial to the eight data:

x	1.0	1.0	1.0	2.0	3.0	4.0	5.0	6.0
y	1.0	2.0	3.0	4.0	4.0	2.0	3.0	8.0




```
Input order:
3
Number of data:   8
Data (x,y):
     1.00000      1.00000
     1.00000      2.00000
     1.00000      3.00000
     2.00000      4.00000
     3.00000      4.00000
     4.00000      2.00000
     5.00000      3.00000
     6.00000      8.00000
Matrix G:
      1.0      1.0      1.0      1.0
      1.0      1.0      1.0      1.0
      1.0      1.0      1.0      1.0
      1.0      2.0      4.0      8.0
      1.0      3.0      9.0     27.0
      1.0      4.0     16.0     64.0
      1.0      5.0     25.0    125.0
      1.0      6.0     36.0    216.0
Vector d:
      1.0      2.0      3.0      4.0      4.0      2.0      3.0      8.0
Matrix A(=GtG):
      8.0     23.0     93.0    443.0
     23.0     93.0    443.0   2277.0
     93.0    443.0   2277.012203.0
    443.0 2277.012203.067173.0
Vector b:
     27.0     97.0    453.0   2377.0
A inverse  :
      8.1    -10.2      3.3     -0.3
    -10.2     13.5     -4.4      0.4
      3.3     -4.4      1.5     -0.1
     -0.3      0.4     -0.1      0.0
Solution:
     -5.9     11.4     -4.0      0.4
```

EXERCISE 5-4

Try to fit the 3-rd order polynomial to the three data, (1.0, 1.0), (2.0, 1.0), (4.0, 3.0) and what happens?

9. Application to hypocenter determination

Here, we apply the theory derived in Chapter 8 to hypocenter determination in homogeneous half-space. The observational equation is given by:

(9-1) $$t_i = t_0 + \frac{L_i}{V} + \varepsilon_i$$

where t_i is the arrival time observed at the i-th station, t_0 is the origin time, V is the velocity of the seismic wave, ε_i is the error for the i-th station, and L_i is the distance between the event and the i-th station defined by:

(9-2) $$L_i = \left(\left(x_e - x_s^i\right)^2 + \left(y_e - y_s^i\right)^2 + \left(z_e - z_s^i\right)^2 \right)^{1/2}$$

where $x_e, y_e, z_e, x_s^i, y_s^i, z_s^i$ are the latitude, longitude and depth of the event and the i-th station in the Cartesian coordinate, respectively (we consider the case in which we can neglect the curvature of the earth's surface). The partial derivatives of the arrival time of the i-th station with respect to hypocentral parameters are given by:

(9-3) $$\begin{cases} \dfrac{\partial t_i}{\partial t_0} = 1 \\ \dfrac{\partial t_i}{\partial x_e} = \dfrac{1}{V}\dfrac{1}{2}\left(\left(x_e - x_s^i\right)^2 + \left(y_e - y_s^i\right)^2 + \left(z_e - z_s^i\right)^2 \right)^{-1/2} 2\left(x_e - x_s^i\right) = \dfrac{1}{V}\dfrac{1}{L_i}\left(x_e - x_s^i\right) \\ \dfrac{\partial t_i}{\partial y_e} = \dfrac{1}{V}\dfrac{1}{L_i}\left(y_e - y_s^i\right) \\ \dfrac{\partial t_i}{\partial z_e} = \dfrac{1}{V}\dfrac{1}{L_i}\left(z_e - z_s^i\right) \end{cases}$$

Note that the partial derivatives with respect to spatial parameters depend on the guess. The matrix $\mathbf{G}_k$ in eq. (8-1-5) is given by:

(9-4) $$\begin{bmatrix} \dfrac{\partial t_1}{\partial t_0} & \dfrac{\partial t_1}{\partial x_e} & \dfrac{\partial t_1}{\partial y_e} & \dfrac{\partial t_1}{\partial z_e} \\ \dfrac{\partial t_2}{\partial t_0} & \dfrac{\partial t_2}{\partial x_e} & \dfrac{\partial t_2}{\partial y_e} & \dfrac{\partial t_2}{\partial z_e} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \dfrac{\partial t_N}{\partial t_0} & \dfrac{\partial t_N}{\partial x_e} & \dfrac{\partial t_N}{\partial y_e} & \dfrac{\partial t_N}{\partial z_e} \end{bmatrix}$$

We assume the following:

(9-5) $$\begin{cases} \mathbf{C}_{\mathbf{d}} = \sigma^2 \mathbf{I} \\ \mathbf{C}_{\mathbf{m}} = \gamma^2 \mathbf{I} \end{cases}$$

The second equation in eqs. (8-1-6) reduces as:

(9-6) $$\left(\frac{1}{\sigma^2} \mathbf{G}_k{}^{\mathrm{T}} \mathbf{G}_k + \frac{1}{\gamma^2} \mathbf{I} \right) (\mathbf{m} - \mathbf{m}_k) = \frac{1}{\sigma^2} \mathbf{G}_k{}^{\mathrm{T}} (\mathbf{d} - \mathbf{f}(\mathbf{m}_k)) + \frac{1}{\gamma^2} (\mathbf{m}_{\mathbf{0}} - \mathbf{m}_k)$$
$$\therefore\ \mathbf{m}_{k+1} = \mathbf{m}_k + \left(\mathbf{G}_k{}^{\mathrm{T}} \mathbf{G}_k + \varepsilon^2 \mathbf{I} \right)^{-1} \left(\mathbf{G}_k{}^{\mathrm{T}} (\mathbf{d} - \mathbf{f}(\mathbf{m}_k)) + \varepsilon^2 (\mathbf{m}_{\mathbf{0}} - \mathbf{m}_k) \right)$$

where $\varepsilon^2 = \sigma^2 / \gamma^2$.