Two solutions for Exercise 2

Isao Sasano

April 14, 2015

This document gives Exercise 2 and two solutions for it. One is by substituting the values into the normal equations and the other is by substituting the values from the beginning.

Exercise 2 Fit a parabola (a square function) to the following four points so that (the half of) the sum of the squares of the distances of those points from the parabola is minimum, where the distance is measured in the vertical direction (the y-direction).

$$(-1, 0), (0, -1), (1, 0), (2, 1)$$

Solution 1 Let the function be $f(x) = ax^2 + bx + c$ and $(x_1, y_1) = (-1, 0), (x_2, y_2) = (0, -1), (x_3, y_3) = (1, 0), (x_4, y_4) = (2, 1)$. The half of the sum of the squares of the distances of these points from the parabola is given as follows.

$$J = \frac{1}{2}\sum_{i=1}^{4}(f(x_i) - y_i)^2 = \frac{1}{2}\sum_{i=1}^{4}(ax_i^2 + bx_i + c - y_i)^2$$

J takes the minimum value in the point where the partial derivatives of J with respect to a, b, and c are 0.

$$\frac{\partial J}{\partial a} = 0, \quad \frac{\partial J}{\partial b} = 0, \quad \frac{\partial J}{\partial c} = 0$$

Firstly the partial derivative of J with respect to a is calculated as follows.

$$\frac{\partial J}{\partial a} = \frac{\partial}{\partial a}\{\frac{1}{2}\sum_{i=1}^{4}(ax_i^2 + bx_i + c - y_i)^2\}$$

$$
\begin{aligned}
&= \frac{1}{2}\sum_{i=1}^{4}\frac{\partial}{\partial a}(ax_i^2+bx_i+c-y_i)^2 \\
&= \frac{1}{2}\sum_{i=1}^{4}2(ax_i^2+bx_i+c-y_i)\frac{\partial}{\partial a}(ax_i^2+bx_i+c-y_i) \\
&= \frac{1}{2}\sum_{i=1}^{4}2(ax_i^2+bx_i+c-y_i)x_i^2 \\
&= \sum_{i=1}^{4}(ax_i^2+bx_i+c-y_i)x_i^2 \\
&= \sum_{i=1}^{4}(ax_i^4+bx_i^3+cx_i^2-x_i^2y_i) \\
&= a\sum_{i=1}^{4}x_i^4+b\sum_{i=1}^{4}x_i^3+c\sum_{i=1}^{4}x_i^2-\sum_{i=1}^{4}x_i^2y_i
\end{aligned}
$$

Secondly the partial derivative of J with respect to b is calculated as follows.

$$
\begin{aligned}
\frac{\partial J}{\partial b} &= \frac{\partial}{\partial b}\{\frac{1}{2}\sum_{i=1}^{4}(ax_i^2+bx_i+c-y_i)^2\} \\
&= \frac{1}{2}\sum_{i=1}^{4}\frac{\partial}{\partial b}(ax_i^2+bx_i+c-y_i)^2 \\
&= \frac{1}{2}\sum_{i=1}^{4}2(ax_i^2+bx_i+c-y_i)\frac{\partial}{\partial b}(ax_i^2+bx_i+c-y_i) \\
&= \frac{1}{2}\sum_{i=1}^{4}2(ax_i^2+bx_i+c-y_i)x_i \\
&= \sum_{i=1}^{4}(ax_i^2+bx_i+c-y_i)x_i \\
&= \sum_{i=1}^{4}(ax_i^3+bx_i^2+cx_i-x_iy_i) \\
&= a\sum_{i=1}^{4}x_i^3+b\sum_{i=1}^{4}x_i^2+c\sum_{i=1}^{4}x_i-\sum_{i=1}^{4}x_iy_i
\end{aligned}
$$

Thirdly the partial derivative of J with respect to c is calculated as follows.

$$
\frac{\partial J}{\partial c} = \frac{\partial}{\partial c}\{\frac{1}{2}\sum_{i=1}^{4}(ax_i^2+bx_i+c-y_i)^2\}
$$

$$
\begin{aligned}
&= \frac{1}{2}\sum_{i=1}^{4}\frac{\partial}{\partial c}(ax_i^2+bx_i+c-y_i)^2 \\
&= \frac{1}{2}\sum_{i=1}^{4}2(ax_i^2+bx_i+c-y_i)\frac{\partial}{\partial c}(ax_i^2+bx_i+c-y_i) \\
&= \frac{1}{2}\sum_{i=1}^{4}2(ax_i^2+bx_i+c-y_i)1 \\
&= \sum_{i=1}^{4}(ax_i^2+bx_i+c-y_i) \\
&= a\sum_{i=1}^{4}x_i^2+b\sum_{i=1}^{4}x_i+c\sum_{i=1}^{4}1-\sum_{i=1}^{4}y_i
\end{aligned}
$$

Then we obtain the system of equations with respect to a, b, and c. The coefficients of the equations are computed as follows.

$$
\begin{array}{cccc}
\sum_{i=1}^{4}x_i^4 = 18, & \sum_{i=1}^{4}x_i^3 = 8, & \sum_{i=1}^{4}x_i^2 = 6, & \sum_{i=1}^{4}x_i = 2 \\
\sum_{i=1}^{4}1 = 4, & \sum_{i=1}^{4}x_i^2y_i = 4, & \sum_{i=1}^{4}x_iy_i = 2, & \sum_{i=1}^{4}y_i = 0
\end{array}
$$

Hence the system of equations is obtained as follows.

$$
\begin{aligned}
18a+8b+6c-4 &= 0 \quad \cdots(1) \\
8a+6b+2c-2 &= 0 \quad \cdots(2) \\
6a+2b+4c &= 0 \quad \cdots(3)
\end{aligned}
$$

By solving this, we obtain the solution.

$$
a=\frac{1}{2},\quad b=-\frac{1}{10},\quad c=-\frac{7}{10}
$$

Hence the function is obtained as follows.

$$
f(x)=\frac{1}{2}x^2-\frac{1}{10}x-\frac{7}{10}
$$

The function is depicted with the four points in Fig. 1. In Fig. 1 the green symbols $\times$ are the given points and the red curve is the square function.

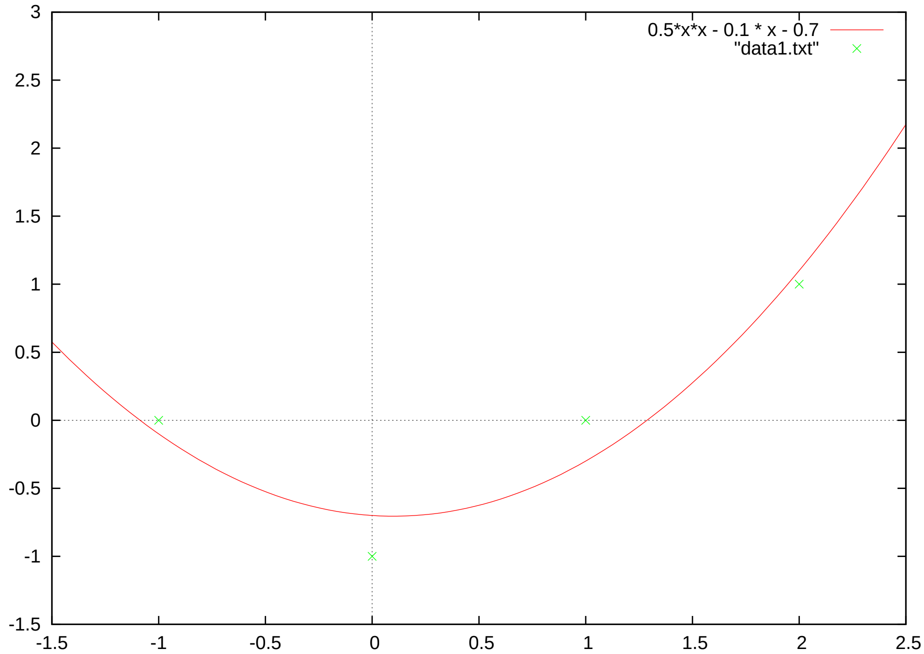

Figure 1: The parabola closest to the given four points

Solution 2 Let the function be $f(x) = ax^2 + bx + c$ and $(x_1, y_1) = (-1, 0), (x_2, y_2) = (0, -1), (x_3, y_3) = (1, 0), (x_4, y_4) = (2, 1)$. The half of the sum of the squares of the distances of these points from the parabola is calculated as follows.

$$\begin{aligned} J &= \frac{1}{2}\Big\{\{f(-1) - 0\}^2 + \{f(0) - (-1)\}^2 + \{f(1) - 0\}^2 + \{f(2) - 1\}^2\Big\} \\ &= \frac{1}{2}\Big\{(a - b + c)^2 + (c + 1)^2 + (a + b + c)^2 + (4a + 2b + c - 1)^2\Big\} \end{aligned}$$

J takes the minimum value in the point where the partial derivatives of J with respect to a, b, and c are 0.

$$\frac{\partial J}{\partial a} = 0, \quad \frac{\partial J}{\partial b} = 0, \quad \frac{\partial J}{\partial c} = 0$$

Firstly the partial derivative of J with respect to a is calculated as follows.

$$\begin{aligned} \frac{\partial J}{\partial a} &= \frac{\partial}{\partial a}\left\{\frac{1}{2}\Big\{(a - b + c)^2 + (c + 1)^2 + (a + b + c)^2 + (4a + 2b + c - 1)^2\Big\}\right\} \\ &= \frac{1}{2}\left\{\frac{\partial}{\partial a}(a - b + c)^2 + \frac{\partial}{\partial a}(c + 1)^2 + \frac{\partial}{\partial a}(a + b + c)^2 + \frac{\partial}{\partial a}(4a + 2b + c - 1)^2\right\} \\ &= \frac{1}{2}\Big\{2(a - b + c) + 2(a + b + c) + 2(4a + 2b + c - 1) \cdot 4\Big\} \\ &= (a - b + c) + (a + b + c) + (4a + 2b + c - 1) \cdot 4 \\ &= (a - b + c) + (a + b + c) + (16a + 8b + 4c - 4) \\ &= 18a + 8b + 6c - 4 \end{aligned}$$

Note that $(c + 1)^2$ is a constant with respect to a and hence the following equation holds.

$$\frac{\partial}{\partial a}(c + 1)^2 = 0$$

Secondly the partial derivative of J with respect to b is calculated as follows.

$$\begin{aligned} \frac{\partial J}{\partial b} &= \frac{\partial}{\partial b}\left\{\frac{1}{2}\Big\{(a - b + c)^2 + (c + 1)^2 + (a + b + c)^2 + (4a + 2b + c - 1)^2\Big\}\right\} \\ &= \frac{1}{2}\left\{\frac{\partial}{\partial b}(a - b + c)^2 + \frac{\partial}{\partial b}(c + 1)^2 + \frac{\partial}{\partial b}(a + b + c)^2 + \frac{\partial}{\partial b}(4a + 2b + c - 1)^2\right\} \\ &= \frac{1}{2}\Big\{2(a - b + c) \cdot (-1) + 2(a + b + c) + 2(4a + 2b + c - 1) \cdot 2\Big\} \end{aligned}$$

$$\begin{aligned}
&= (a-b+c)\cdot(-1)+(a+b+c)+(4a+2b+c-1)\cdot 2\\
&= (-a+b-c)+(a+b+c)+(8a+4b+2c-2)\\
&= 8a+6b+2c-2
\end{aligned}$$

Note that $(c+1)^2$ is a constant with respect to b and hence the following equation holds.

$$\frac{\partial}{\partial b}(c+1)^2 = 0$$

Thirdly the partial derivative of J with respect to c is calculated as follows.

$$\begin{aligned}
\frac{\partial J}{\partial c} &= \frac{\partial}{\partial c}\left\{\frac{1}{2}\left\{(a-b+c)^2+(c+1)^2+(a+b+c)^2+(4a+2b+c-1)^2\right\}\right\}\\
&= \frac{1}{2}\left\{\frac{\partial}{\partial c}(a-b+c)^2+\frac{\partial}{\partial c}(c+1)^2+\frac{\partial}{\partial c}(a+b+c)^2+\frac{\partial}{\partial c}(4a+2b+c-1)^2\right\}\\
&= \frac{1}{2}\left\{2(a-b+c)+2(c+1)+2(a+b+c)+2(4a+2b+c-1)\right\}\\
&= (a-b+c)+(c+1)+(a+b+c)+(4a+2b+c-1)\\
&= 6a+2b+4c
\end{aligned}$$

Then we obtain the system of equations with respect to a, b, and c.

$$\begin{aligned}
18a+8b+6c-4 &= 0 \quad \cdots(1)\\
8a+6b+2c-2 &= 0 \quad \cdots(2)\\
6a+2b+4c &= 0 \quad \cdots(3)
\end{aligned}$$

We show the process of solving this system of equations. Here we solve it as high-school student may solve it.

Divide (1), (2), and (3) by two.

$$\begin{aligned}
9a+4b+3c-2 &= 0 \quad \cdots(4)\\
4a+3b+c-1 &= 0 \quad \cdots(5)\\
3a+b+2c &= 0 \quad \cdots(6)
\end{aligned}$$

Eliminate c from (5) and (6). Subtract (6) from the twice of (5) and obtain

$$5a+5b-2=0 \quad \cdots(7)$$

Secondly eliminate c from (4) and (5). Subtract (4) from the threefold of (5) and obtain

$$3a+5b-1=0 \quad \cdots(8)$$

(7) and (8) is a system of equations with respect to a and b. Subtract (8) from (7) and obtain

$$2a - 1 = 0$$

So we obtain

$$a = \frac{1}{2}$$

Substitute this into (8) and obtain

$$\frac{3}{2} + 5b - 1 = 0$$

and hence obtain

$$b = -\frac{1}{10}$$

Substitute these into (5) and we obtain

$$4 \cdot (-\frac{1}{2}) + 3 \cdot (-\frac{1}{10}) + c - 1 = 0$$

By solving this we obtain

$$c = -\frac{7}{10}$$

Hence the function is obtained as follows.

$$f(x) = \frac{1}{2}x^2 - \frac{1}{10}x - \frac{7}{10}$$

The function is depicted with the four points in Fig. 1.

(Note 1) Although we calculate the exact solution without considering significant figures, actual data obtained by experiments have errors. So it does not make sense for us to calculate the exact solution and we should calculate the solution with taking into account the significant figures. In this class we do not care about the errors and calculate the exact solution. Refer to Chapter 19, 20, and 21 for the numerical analysis, where we care about the errors.

(Note 2) In the reference book, the distances are written as $|y_i - f(x_i)|$. We take the squares of them, so the order of subtraction does not affect.

(Note 3) Methods for solving linear systems of equations are largely classified into two: direct and iterative methods. The Gauss elimination is a direct one and the Gauss-Seidel method is an iterative one.

Direct methods can be used for calculation on paper. Iterative methods are to obtain approximate solutions and not to obtain the exact solutions.