Two solutions for Exercise 1

Isao Sasano

April 12, 2016

Exercise 1 (Least square method)

Fit a straight line (a linear function) to the following three points so that (the half of) the sum of the squares of the distances of those points from the straight line is minimum, where the distance is measured in the vertical direction (the y-direction).

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

In the following we show Solution 1 which assigns the values of the three points to the variables after obtaining the form of normal equation and Solution 2 which firstly assigns the values of the three points to the variables and then calculates the derivatives.

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

$$J = \frac{1}{2}\sum_{i=1}^{3}\{f(x_i) - y_i\}^2 = \frac{1}{2}\sum_{i=1}^{3}(ax_i + b - y_i)^2$$

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

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

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

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

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

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

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

Then we obtain the system of equations (called *normal equations*)

$$
\begin{aligned}
5a + 3b + 4 &= 0 \\
3a + 3b + 1 &= 0
\end{aligned}
$$

and $a = -\frac{3}{2}, b = \frac{7}{6}$ is the solution. Hence the function is obtained as follows.

$$
f(x) = -\frac{3}{2}x + \frac{7}{6}
$$

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

Solution 2 Let the line (the linear function) be $f(x) = ax + b$. The half of the sum of the squares of the distances of these points from the line is given as follows.

$$
\begin{aligned}
J &= \frac{1}{2}\Big\{\{f(0) - 1\}^2 + \{f(1) - 0\}^2 + \{f(2) - (-2)\}^2\Big\} \\
&= \frac{1}{2}\Big\{(b-1)^2 + (a+b)^2 + (2a+b+2)^2\Big\} \\
&= \frac{1}{2}\Big\{b^2 - 2b + 1 + a^2 + 2ab + b^2 + 4a^2 + b^2 + 4 + 4ab + 8a + 4b\Big\} \\
&= \frac{1}{2}\Big\{5a^2 + 3b^2 + 6ab + 8a + 2b + 5\Big\}
\end{aligned}
$$

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

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

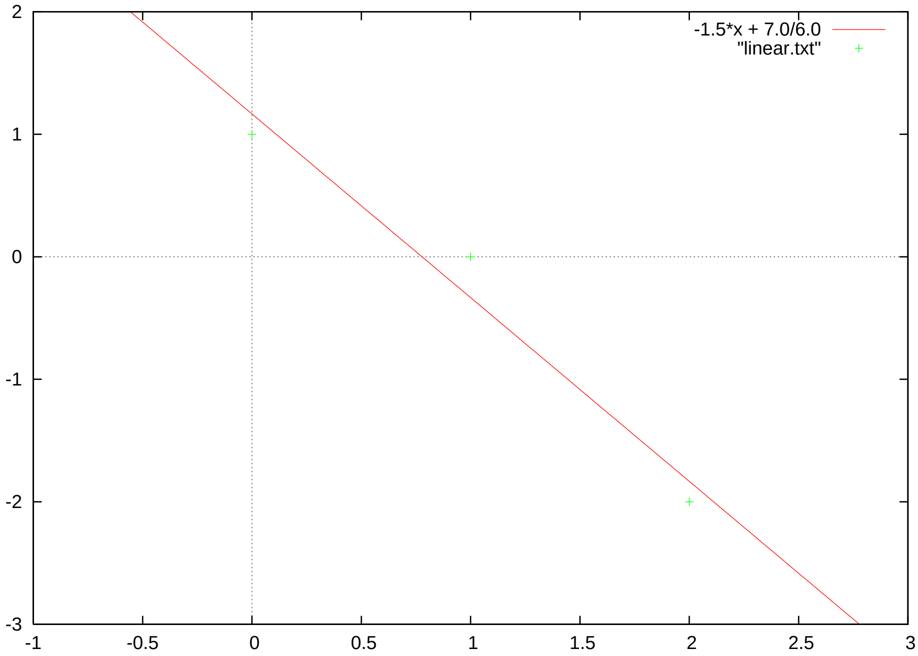

Figure 1: The straight line closest to the given three points

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

$$\begin{aligned}\frac{\partial J}{\partial a} &= \frac{1}{2}(10a + 6b + 8) \\ &= 5a + 3b + 4\end{aligned}$$

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

$$\begin{aligned}\frac{\partial J}{\partial b} &= \frac{1}{2}(6a + 6b + 2) \\ &= 3a + 3b + 1\end{aligned}$$

Then we obtain the system of equations (called *normal equations*)

$$\begin{aligned}5a + 3b + 4 &= 0 \\ 3a + 3b + 1 &= 0\end{aligned}$$

We omit the following since it is same as Solution 1.

Note 1 We may write the above normal equations by using matrices as follows.

$$\begin{pmatrix} 5 & 3 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -4 \\ -1 \end{pmatrix}$$

By solving this we obtain the following solution.

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -\frac{3}{2} \\ \frac{7}{6} \end{pmatrix}$$

Although we may use either of the two forms, the notation in matrices is appropriate for paper-and-pencil calculations in, say, Gaussian elimination when there are more than two variables.

Note 2 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 3 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 4 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 paper-and-pencil calculations. Iterative methods are to obtain approximate solutions and not to obtain the exact solutions.