Least Squares

1 In this problem we'll find the "best fit" line $y = mx + b$ to the points $(x_1, y_1) = (1, 2)$, $(x_1, y_1) = (3, 3)$, and $(x_3, y_3) = (5, 5)$, by directly minimizing the least squared error.

(a) Write down the squared error function $E(m, b) = e_1^2 + e_2^2 + e_3^2$, where (for example) $e_1 = mx_1 + b - y_1$.

(b) Take the two partial derivatives E_m and E_b, and write down the two equations $E_m = 0$ and $E_b = 0$.

(c) Solve the two equations you found in part (b) for m and b. What is the best fit line $y = mx + b$?

2 It turns out that the best fit line to the points (x_1, y_1) through (x_n, y_n) has slope m and y-intercept b, where m and b satisfy the linear system

$$\left(\sum_{i=1}^{n} x_i^2\right) m + \left(\sum_{i=1}^{n} x_i\right) b = \sum_{i=1}^{n} x_i y_i$$

$$\left(\sum_{i=1}^{n} x_i\right) m + nb = \sum_{i=1}^{n} y_i.$$

This can be written as

$$\begin{bmatrix} \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & n \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} x_i y_i \\ \sum_{i=1}^{n} y_i \end{bmatrix}.$$

This is sometimes called the *normal system.*

Find the normal system for the points from the previous problem by completing the following table:

i	x_i	y_i	x_i^2	$x_i y_i$
1	1	2		
2	3	3		
3	5	5		
Sum				

Is this the same linear system you found in problem 1(b)?

3 Find the least squares line given the data

$$(x_1, y_1) = (1, 1), \qquad (x_2, y_2) = (2, 4), \qquad (x_3, y_3) = (3, 6), \qquad (x_4, y_4) = (4, 9)$$

as follows:

(a) Complete the following table:

i	x_i	y_i	x_i^2	$x_i y_i$
1	1	1		
2	2	4		
3	3	6		
4	4	9		
Sum				

(b) Write down the normal system.

(c) Solve the normal system, either by row reduction (elimination) or by finding an inverse matrix. This will give you m and b.

(d) Write down the least squares line $y = mx + b$ using the slope m and intercept b that you found in part (c).

4 Repeat the previous problem with the following data. (That is, for each set of data, find the least squares line using the techniques of the previous problem.)

(a) $(1, 2)$, $(2, 4)$, $(3, 5)$, $(4, 7)$

(b) $(0, -2)$, $(1, 2)$, $(2, 0)$, $(3, 3)$, $(4, 5)$

For your convenience, here are two blank tables:

i	x_i	y_i	x_i^2	$x_i y_i$

i	x_i	y_i	x_i^2	$x_i y_i$

Answers:
1 $y = \frac{3}{4}x + \frac{13}{12}$ **3** $y = 2.6x - 1.5$
4 (a) $y = 1.6x + 0.5$ (b) $y = 1.5x - 1.4$