INSTITUTE OF TECHNOLOGY OF CAMBODIA

Department: AMS

Assignment

MATHEMATICAL MODELING

Professor: LUEY SOKEA

Group8

1. SAO SAMARTH e20200084
2. TENG CHANSOPANHA e20201711
3. THONG CHHUNHER e20200711
4. THORNTHEA GECHHAI e20201321

Year: 2022-2023

Assignment

Deadline: 12 December, 2022

1. Derive the equations that minimize the sum of the squared deviations between a set of data points and the quadratic model $y = c_1 x^2 + c_2 x + c_3$ Use the equations to find estimates of c_1, c_2 and c_3 for the following set of data.

x	0.1	0.2	0.3	0.4	0.5
y	0.06	0.12	0.36	0.65	0.96

   Compute D and $\mathrm{d}_{\max}$ to bound $C_{\max}$ Compare the results with your solution to problem 3 in Section 3.2.

2. A general rule for computing a person's weight is as follows: For a female, multiply the height in inches by 3.5 and subtract 108; for a male, multiply the height in inches by 4.0 and subtract 128. If the person is small bone-structured, adjust this computation by subtracting 10%; for a large bone-structured person, add 10%. No adjustment is made for an average-size person. Gather data on the weight versus height of people of differing age, size, and gender. Using Equation (3.4), fit a straight line to your data for males and another straight line to your data for females. What are the slopes and intercepts of those lines? How do the results compare with the general rule?

3. Write a computer program that finds the least-squares estimates of the coefficients in the following models.
   a. $y = ax^2 + bx + c$
   b. $y = ax^n$

4. Write a computer program that computes the deviation from the data points and any model that the user enters. Assuming that the model was fitted using the least-squares criterion, compute D and $\mathrm{d}_{\max}$. Output each data point, the deviation from each data point, D; $\mathrm{d}_{\max}$; and the sum of the squared deviations.

1. For the following data, formulate the mathematical model that minimizes the largest deviation between the data and the model $y = c_1x^2 + c_2x + c3$. If a computer code is available, solve for the estimates of c_1, c_2, and c_3.

x	0.1	0.2	0.3	0.4	0.5
y	0.06	0.12	0.36	0.65	0.95

The objective is to find the estimates of c_1, c_2 and c_3, use $y = ax + b$ to obtain the derivation of equation to minimize the sum of squared deviations between set of data points and a quadratic model $y = c_1x^2 + c_2x + c_3$. Also calculate D and $d_{\max}$ to bound $c_{\max}$ along with comparison of obtained solutions with the solution of problem 2 of section 3.2.

The provided data is shown below as,

x	0.1	0.2	0.3	0.4	0.5
y	0.06	0.12	0.36	0.65	0.95

The normal equations are;

$$\begin{aligned} nc_1 + c_2\sum x + c_3\sum x^2 &= \sum y \\ c_1\sum x + c_2\sum x^2 + c_3\sum x^3 &= \sum xy \\ c_1\sum x^2 + c_2\sum x^3 + c_3\sum x^4 &= \sum x^2y \end{aligned}$$

Now compute the values and get,

$$\sum x = 1.5, \sum y = 2.14, \sum x^2 = 0.55, \sum xy = 0.873, \sum x^3 = 0.225, \sum x^2y = 0.3793, \sum x^4 = 0.305501$$

Substitute all these values in the normal equations to obtain,

$$\begin{aligned} 5c_1 + 1.5c_2 + 0.55c_3 &= 2.14 \\ 1.5c_1 + 0.55c_2 + 0.225c_3 &= 0.873 \\ 0.55c_1 + 0.225c_2 + 0.3055c_3 &= 0.3793 \end{aligned}$$

Solve the system of equations and get,

$$c_1 = 3.785,\ c_2 = 0.03857,\ c_3 = 0$$

Hence the second-degree equation namely the parabola is $y = 3.785x^2 + 0.03857x + 0$

calculate D ,

x	y	$y_i - y(x_i) = y_i - \left(3.785x_i^2 + 0.03857x_i + 0\right)$	$\left(y_i - y(x_i)\right)^2$
0.1	0.06	0.018293	0.0003346
0.2	0.12	−0.0391	0.00152881
0.3	0.36	0.007779	0.00006051
0.4	0.65	0.028972	0.00083937
0.5	0.95	−0.015535	0.00024133
			$\sum\left(y_i - y(x_i)\right)^2 = 0.00300462$

Thus, $D = \left(\frac{0.00300462}{5}\right)^{\frac{1}{2}} = 0.02451$

So, minimum deviation is $D = 0.02451$.
The largest absolute deviation is 0.0391 when $x = 0.2$.
Then bound the $c_{\max}$ as,

$$D = 0.02451 \le c_{\max} \le 0.0391 = d_{\max}$$

Thus, $d_{mxz} = 0.0391$

On comparing the obtained solutions $y = 3.785x^2 + 0.03857x + 0$ and $D = 0.02451$ with the solutions of the problem 3 of section 3.2 which is $y = 3.7857x^2 + 0.03847x - 0.0050$ and $r_{\min} = 0.02833$,

it is clearly observed that the equation and value of estimates c_1, c_2 and c_3 both are different and the value of minimum deviation of the obtained solution is less than the minimum deviation of problem 3 of section 3.2.

Exercise2:

$$\left.\begin{aligned} a\sum_{i=1}^{m} x_i^2 + b\sum_{i=1}^{m} x_i &= \sum_{i=1}^{m} x_i y_i \\ a\sum_{i=1}^{m} x_i + mb &= \sum_{i=1}^{m} y_i \end{aligned}\right\}$$

Applying least square criterion to the model $y = ax + b$, it requires the minimization of

$$\begin{aligned} S &= \sum_{i=1}^{m} \left(y_i - f(x_i)\right)^2 \\ &= \sum_{i=1}^{m} [y_i - (ax_i + b)]^2 \end{aligned}$$

A necessary condition for optimality is that the two partial derivatives $\frac{\partial S}{\partial a}$ and $\frac{\partial S}{\partial b}$ equal zero.
Refer to page 122 of text book.
From the text book,

$$a = \frac{m\sum_{i=1}^{m} x_i y_i - \sum_{i=1}^{m} x_i \sum_{i=1}^{m} y_i}{m\sum_{i=1}^{m} (x_i)^2 - \left[\sum_{i=1}^{m} x_i\right]^2}$$

$$b = \frac{\sum_{i=1}^{m} (x_i)^2 \sum_{i=1}^{m} y_i - \sum_{i=1}^{m} (x_i y_i) \sum_{i=1}^{m} x_i}{m\sum_{i=1}^{m} (x_i)^2 - \left[\sum_{i=1}^{m} x_i\right]^2}$$

The following data was collected for the height and weight of normal men.

Height, in	60	62	64	66	68
Weight, kg	54	60	63	66	70

The above data was fitted to the straight line $\boldsymbol{y = ax + b}$, where slope is $\boldsymbol{a = 1.9}$ and intercept is $\boldsymbol{b = -59}$. The model is $\boldsymbol{y = 1.9x - 59}$

The following data was collected for the height and weight of normal women.

Height, in	60	62	64	66	68
Weight, kg	54	55	59	63	66

The above data was fitted to the straight line $\boldsymbol{y = ax + b}$, where slope is $\boldsymbol{a = 1.6}$ and intercept is $\boldsymbol{b = -43}$. The model is $\boldsymbol{y = 1.6x - 43}$.

It is assumed that the weights obtained in general rules are in pounds.
Using the general rule the weights of men are given below.

Height, in	60	62	64	66	68
Weight, kg	50.8	54.4	58	61.7	65.3

Using the general rule, the weights of women are given below.

Height, in	60	62	64	66	68
Weight, kg	46.3	49.4	52.6	55.8	59

From the above table, the general rule gives lesser values for the weight of both men and women.

Exercise3: (a)
It is required to fit the quadratic model $y = ax^2 + bx + c$ with the given m data points. Applying least square criterion to the model $y = ax^2 + bx + c$, it requires the minimization of

$$\begin{aligned} S &= \sum_{i=1}^{m} \left(y_i - f(x_i)\right)^2 \\ &= \sum_{i=1}^{m} \left[y_i - \left(ax_i^2 + bx_i + c\right)\right]^2 \\ &= \sum_{i=1}^{m} y_i^2 - \sum_{i=1}^{m} 2y_i\left(ax_i^2 + bx_i + c\right) + \sum_{i=1}^{m} \left(ax_i^2 + bx_i + c\right)^2 \end{aligned}$$

The necessary conditions for optimality are $\frac{\partial S}{\partial a} = 0, \frac{\partial S}{\partial b} = 0$ and $\frac{\partial S}{\partial c} = 0$, that is there are 3 unknowns and 3 equations.
The three equations are formed as given below.
On solving, the first equation can be found as follows:

$$\frac{\partial}{\partial a}\left[\sum_{i=1}^{m} y_i^2 - \sum_{i=1}^{m} 2y_i\left(ax_i^2 + bx_i + c\right) + \sum_{i=1}^{m} \left(ax_i^2 + bx_i + c\right)^2\right] = 0$$

$$-\sum_{i=1}^{m} 2y_i x_i^2 + 2\sum_{i=1}^{m} \left(ax_i^2 + bx_i + c\right)x_i^2 = 0$$

$$-2\sum_{i=1}^{m} y_i x_i^2 + 2a\sum_{i=1}^{m} x_i^4 + 2b\sum_{i=1}^{m} x_i^3 + 2c\sum_{i=1}^{m} x_i^2 = 0$$

$$a\sum_{i=1}^{m} x_i^4 + b\sum_{i=1}^{m} x_i^3 + c\sum_{i=1}^{m} x_i^2 = \sum_{i=1}^{m} x_i^2 y_i \ldots\ldots (1)$$

On solving, the second equation can be found as follows:

$$\frac{\partial}{\partial b}\left[\sum_{i=1}^{m} y_i^2 - \sum_{i=1}^{m} 2y_i\left(ax_i^2 + bx_i + c\right) + \sum_{i=1}^{m} \left(ax_i^2 + bx_i + c\right)^2\right] = 0$$

$$-\sum_{i=1}^{m} 2y_i x_i + 2\sum_{i=1}^{m} \left(ax_i^2 + bx_i + c\right)x_i = 0$$

$$-2\sum_{i=1}^{m} y_i x_i + 2a\sum_{i=1}^{m} x_i^3 + 2b\sum_{i=1}^{m} x_i^2 + 2c\sum_{i=1}^{m} x_i = 0$$

$$a\sum_{i=1}^{m} x_i^3 + b\sum_{i=1}^{m} x_i^2 + c\sum_{i=1}^{m} x_i = \sum_{i=1}^{m} x_i y_i \ldots\ldots (2)$$

On solving, the third equation can be found as follows:

$$\frac{\partial}{\partial c}\left[\sum_{i=1}^{m} y_i^2 - \sum_{i=1}^{m} 2y_i\left(ax_i^2 + bx_i + c\right) + \sum_{i=1}^{m} \left(ax_i^2 + bx_i + c\right)^2\right] = 0$$

$$-\sum_{i=1}^{m} 2y_i + 2\sum_{i=1}^{m} \left(ax_i^2 + bx_i + c\right) = 0$$

$$-2\sum_{i=1}^{m} y_i + 2a\sum_{i=1}^{m} x_i^2 + 2b\sum_{i=1}^{m} x_i + 2cm = 0$$

$$a\sum_{i=1}^{m} x_i^2 + b\sum_{i=1}^{m} x_i + cm = \sum_{i=1}^{m} y_i \dots \dots (3)$$

Eliminate b and c from the above three equations and get the value of a.
Using equation (1) and (3),

$$m\left(a\sum_{i=1}^{m} x_i^4 + b\sum_{i=1}^{m} x_i^3 + c\sum_{i=1}^{m} x_i^2\right) - \sum_{i=1}^{m} x_i^2\left(a\sum_{i=1}^{m} x_i^2 + b\sum_{i=1}^{m} x_i + cm\right) = m\sum_{i=1}^{m} x_i^2 y_i - \sum_{i=1}^{m} x_i^2 \sum_{i=1}^{m} y_i$$

$$a\left[m\sum_{i=1}^{m} x_i^4 - \left(\sum_{i=1}^{m} x_i^2\right)^2\right] + b\left[m\sum_{i=1}^{m} x_i^3 - \sum_{i=1}^{m} x_i^2 \sum_{i=1}^{m} x_i\right] = m\sum_{i=1}^{m} x_i^2 y_i - \sum_{i=1}^{m} x_i^2 \sum_{i=1}^{m} y_i \dots \dots (4)$$

Using equation (2) and (3),

$$m\left(a\sum_{i=1}^{m} x_i^3 + b\sum_{i=1}^{m} x_i^2 + c\sum_{i=1}^{m} x_i\right) - \sum_{i=1}^{m} x_i\left(a\sum_{i=1}^{m} x_i^2 + b\sum_{i=1}^{m} x_i + cm\right) = m\sum_{i=1}^{m} x_i y_i - \sum_{i=1}^{m} x_i \sum_{i=1}^{m} y_i$$

$$a\left(m\sum_{i=1}^{m} x_i^3 - \sum_{i=1}^{m} x_i \sum_{i=1}^{m} x_i^2\right) + b\left[m\sum_{i=1}^{m} x_i^2 - \left(\sum_{i=1}^{m} x_i\right)^2\right] = m\sum_{i=1}^{m} x_i y_i - \sum_{i=1}^{m} x_i \sum_{i=1}^{m} y_i \dots \dots (5)$$

Using equation (4) and (5),

$$\left\{a\left[m\sum_{i=1}^{m} x_i^4 - \left(\sum_{i=1}^{m} x_i^2\right)^2\right] + b\left[m\sum_{i=1}^{m} x_i^3 - \sum_{i=1}^{m} x_i^2 \sum_{i=1}^{m} x_i\right]\right\}\left[m\sum_{i=1}^{m} x_i^2 - \left(\sum_{i=1}^{m} x_i\right)^2\right]$$
$$-\left[a\left(m\sum_{i=1}^{m} x_i^3 - \sum_{i=1}^{m} x_i \sum_{i=1}^{m} x_i^2\right) + b\left[m\sum_{i=1}^{m} x_i^2 - \left(\sum_{i=1}^{m} x_i\right)^2\right]\right]\left(m\sum_{i=1}^{m} x_i^3 - \sum_{i=1}^{m} x_i^2 \sum_{i=1}^{m} x_i\right)$$
$$= \left(m\sum_{i=1}^{m} x_i^2 y_i - \sum_{i=1}^{m} x_i^2 \sum_{i=1}^{m} y_i\right)\left[m\sum_{i=1}^{m} x_i^2 - \left(\sum_{i=1}^{m} x_i\right)^2\right] - \left(m\sum_{i=1}^{m} x_i y_i - \sum_{i=1}^{m} x_i \sum_{i=1}^{m} y_i\right)\left(m\sum_{i=1}^{m} x_i^3 - \sum_{i=1}^{m} x_i^2 \sum_{i=1}^{m} x_i\right)$$

$$\left\{a\left[m\sum_{i=1}^{m} x_i^4-\left(\sum_{i=1}^{m} x_i^2\right)^2\right]\right\}\left[m\sum_{i=1}^{m} x_i^2-\left(\sum_{i=1}^{m} x_i\right)^2\right]$$
$$-\left[a\left(m\sum_{i=1}^{m} x_i^3-\sum_{i=1}^{m} x_i\sum_{i=1}^{m} x_i^2\right)\right]\left(m\sum_{i=1}^{m} x_i^3-\sum_{i=1}^{m} x_i^2\sum_{i=1}^{m} x_i\right)$$
$$=\left(m\sum_{i=1}^{m} x_i^2y_i-\sum_{i=1}^{m} x_i^2\sum_{i=1}^{m} y_i\right)\left[m\sum_{i=1}^{m} x_i^2-\left(\sum_{i=1}^{m} x_i\right)^2\right]-\left(m\sum_{i=1}^{m} x_iy_i-\sum_{i=1}^{m} x_i\sum_{i=1}^{m} y_i\right)\left(m\sum_{i=1}^{m} x_i^3-\sum_{i=1}^{m} x_i^2\sum_{i=1}^{m} x_i\right)$$
$$a=\frac{\left(m\sum_{i=1}^{m} x_i^2y_i-\sum_{i=1}^{m} x_i^2\sum_{i=1}^{m} y_i\right)\left[m\sum_{i=1}^{m} x_i^2-\left(\sum_{i=1}^{m} x_i\right)^2\right]-\left(m\sum_{i=1}^{m} x_iy_i-\sum_{i=1}^{m} x_i\sum_{i=1}^{m} y_i\right)\left(m\sum_{i=1}^{m} x_i^3-\sum_{i=1}^{m} x_i^2\sum_{i=1}^{m} x_i\right)}{\left[m\sum_{i=1}^{m} x_i^4-\left(\sum_{i=1}^{m} x_i^2\right)^2\right]\left[m\sum_{i=1}^{m} x_i^2-\left(\sum_{i=1}^{m} x_i\right)^2\right]-\left(m\sum_{i=1}^{m} x_i^3-\sum_{i=1}^{m} x_i\sum_{i=1}^{m} x_i^2\right)\left(m\sum_{i=1}^{m} x_i^3-\sum_{i=1}^{m} x_i^2\sum_{i=1}^{m} x_i\right)}$$
$$a\left(m\sum_{i=1}^{m} x_i^3-\sum_{i=1}^{m} x_i\sum_{i=1}^{m} x_i^2\right)+b\left[m\sum_{i=1}^{m} x_i^2-\left(\sum_{i=1}^{m} x_i\right)^2\right]=m\sum_{i=1}^{m} x_iy_i-\sum_{i=1}^{m} x_i\sum_{i=1}^{m} y_i$$
$$b\left[m\sum_{i=1}^{m} x_i^2-\left(\sum_{i=1}^{m} x_i\right)^2\right]=\left(m\sum_{i=1}^{m} x_iy_i-\sum_{i=1}^{m} x_i\sum_{i=1}^{m} y_i\right)-a\left(m\sum_{i=1}^{m} x_i^3-\sum_{i=1}^{m} x_i\sum_{i=1}^{m} x_i^2\right)$$
$$b=\frac{\left(m\sum_{i=1}^{m} x_iy_i-\sum_{i=1}^{m} x_i\sum_{i=1}^{m} y_i\right)-a\left(m\sum_{i=1}^{m} x_i^3-\sum_{i=1}^{m} x_i\sum_{i=1}^{m} x_i^2\right)}{\left[m\sum_{i=1}^{m} x_i^2-\left(\sum_{i=1}^{m} x_i\right)^2\right]}$$

From equation (3),

$$\begin{aligned} a\sum_{i=1}^{m} x_i^2+b\sum_{i=1}^{m} x_i+cm &= \sum_{i=1}^{m} y_i \\ c &= \frac{\sum_{i=1}^{m} y_i-\left(a\sum_{i=1}^{m} x_i^2+b\sum_{i=1}^{m} x_i\right)}{m} \end{aligned}$$

The program for finding the least-squares estimates of the coefficients in the model $y = ax^2 + bx + c$ was written using Maple.

In the Maple file enter the values in the array x and y also enter m, the number of data points. In the Maple file now you have dummy values.

(b)

Applying least square criterion to the model $y = ax^n$, it required the minimization of

$$\begin{aligned} S &= \sum_{i=1}^{m} \left(y_i-f(x_i)\right)^2 \\ &= \sum_{i=1}^{m} [y_i-ax_i^n]^2 \\ &= \sum_{i=1}^{m} y_i^2-\sum_{i=1}^{m} 2y_i(ax_i^n)+\sum_{i=1}^{m} a^2x_i^{2n} \end{aligned}$$

The necessary condition for optimality is $\frac{\partial S}{\partial a} = 0$, that is

$$
\begin{aligned}
\frac{\partial}{\partial a}\left[\sum_{i=1}^{m} y_i^2 - \sum_{i=1}^{m} 2y_i(ax_i^n) + \sum_{i=1}^{m} a^2 x_i^{2n}\right] &= 0 \\
-\sum_{i=1}^{m} 2y_i x_i^n + \sum_{i=1}^{m} 2a x_i^{2n} &= 0 \\
-\sum_{i=1}^{m} y_i x_i^n + a\sum_{i=1}^{m} x_i^{2n} &= 0 \\
a &= \frac{\sum_{i=1}^{m} y_i x_i^n}{\sum_{i=1}^{m} x_i^{2n}}
\end{aligned}
$$

The program for finding the least-squares estimates of the coefficients in the model

$y = ax^n$ was written using Maple.
In the Maple file enter the values in the array x and y also enter m and n. In the Maple file now you have dummy values.