Mathematics 2L — Linear Modelling

Solutions 2

1 (i) Use the formulæ

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} + O(h^2),$$
$$f''(x) = \frac{f(x+h) - 2f(x) + f(x-h)}{h^2} + O(h^2),$$

at the midpoints. Here we have (i) $x = 0.2$, $h = 0.05$, so

$$f'(0.2) \simeq \frac{1.2840 - 1.1618}{2(0.05)} = 1.2220,$$

and

$$f''(0.2) \simeq \frac{1.2840 - 2(1.2214) + 1.1618}{(0.05)^2} = 1.2000,$$

to 4 decimal places.

1 (ii) $x = 1.2$, $h = 0.1$ so

$$f'(1.2) \simeq \frac{.2624 - .09531}{2(0.1)} = 0.8351, \quad [.8333]$$

and

$$f''(1.2) \simeq \frac{0.2624 - 2(0.1823) + 0.09531}{(0.1)^2} = -0.6820 \quad [-.6890],$$

2 (i) Here $n = \dfrac{b-a}{h} = 4$ and $x_j = \dfrac{j}{4}$ $(0 \le j \le 4)$
We have $y_0 = 0$, $y_4 = 1$ and need to determine $y_i = y(x_i)$ for $i = 1, 2, 3$.
The central difference formulæ yield

$$16\big(y_{j-1} - 2y_j + y_{j+1}\big) + x_j y_j = 1 \qquad (1 \le j \le 3)$$

so

$$\begin{array}{cccccc} 64y_0 & -128y_1 & +64y_2 & +y_1 & = & 4 \\ 64y_1 & -128y_2 & +64y_3 & +2y_2 & = & 4 \\ 64y_2 & -128y_3 & +64y_4 & +3y_3 & = & 4 \end{array}$$

and on putting in the boundary conditions $y_0 = 0$, $y_4 = 1$ we get

$$\begin{bmatrix} -127 & 64 & 0 \\ 64 & -126 & 64 \\ 0 & 64 & -125 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ -60 \end{bmatrix}$$

This leads to
$y_1 = \frac{1035190644441090}{5995576179081220} \simeq 0.1726590762$, $y_2 = \frac{1224027689451520}{3021392720166910} \simeq 0.4051203544$
and $y_3 = \frac{346169049088}{503576027136} \simeq 0.6874216214$

2(ii)
The central difference formulæ yield

$$16\big(y_{j-1} - 2y_j + y_{j+1}\big) + y_j = 0 \qquad (1 \le j \le 3)$$

The left end boundary condition is $y_0 = 1$. At the right end point we use $y_2 - 4y_3 + 3y_4 = 0$.
The system to solve is

$$\begin{bmatrix} -31 & 16 & 0 & 0 \\ 16 & -31 & 16 & 0 \\ 0 & 16 & -31 & 16 \\ 0 & 1 & -4 & 3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} -16 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

The exact solution is, as is easy to check,

$$y = \cos x + \tan 1 \sin x$$

To compare:

j	x_j	y_j	$y(x_j)$
1	0.25	$\frac{6334416}{4678179} \approx 1.3540345506$	1.3542212590
2	0.5	$\frac{244992}{150909} \approx 1.6234419418$	1.6242435991
3	0.75	$\frac{270336}{150909} \approx 1.7913842117$	1.7932783392
4	1	$\frac{8448}{4573} \approx 1.8473649683$	1.8508157177

3 (i)

Here $n = \dfrac{b-a}{h} = 3$ and $x_0 = 1$, $x_1 = \frac{4}{3}$, $x_2 = \frac{5}{3}$, $x_3 = 2$
We have $y_0 = 0$, $y_3 = 4$ and need to determine $y_i = y(x_i)$ for $i = 1, 2$.
The central difference formulæ yield

$$\frac{16}{9}\, 9\big(y_0 - 2y_1 + y_2\big) - \frac{4}{3}\,\frac{3}{2}\big(y_2 - y_0\big) + y_1 = \frac{16}{9}$$

and

$$\frac{25}{9}\, 9\big(y_1 - 2y_2 + y_3\big) - \frac{5}{3}\,\frac{3}{2}\big(y_3 - y_1\big) + y_2 = \frac{25}{9}$$

from which

$$\begin{bmatrix} -279 & 126 \\ 495 & -882 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 16 \\ -1570 \end{bmatrix}$$

This leads to $y_1 = \frac{25370993340}{25370993340} = 1$, and $y_2 = \frac{26825960700}{11457867960} \simeq 2.3412698413$.
The exact solution is

$$y = (\log_2 x - 1)\, x + x^2$$

whose values at x_1 and x_2 are [to 10dp] 0.9978277768 and 2.3393871014.

3 (ii) Start with forward differences at $x_0 = 1$ for y'. Then use the central differences at x_1 and x_2.

We get 3 equations for the unknown function values y_0, y_1 and y_2:

$$\begin{bmatrix} -3 & 4 & -1 \\ 162 & -279 & 126 \\ 0 & 495 & -882 \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 16 \\ -355 \end{bmatrix}$$

with solutions $y_0 \simeq 0.5243902439$, $y_1 \simeq 0.5745257453$ and $y_2 \simeq 0.7249322493$.
The exact solution is

$$y = \big(A + B \ln x\big)x + x^2$$

with

$$A = \frac{-1}{2(1 - \ln 2)}, \qquad B = \frac{4\ln 2 - 3}{2(1 - \ln 2)}.$$

This takes the values

$$0.6294456766 \qquad 0.6586889344 \qquad 0.7729162319$$

at the mesh points. We do not get such good accuracy this time.

4 The loop equations are

$$\begin{array}{lllcl} i_1 & +2(i_1 - i_2) & & = 21 & \textit{loop } 1 \\ 2(i_2 - i_1) & +i_2 & +2(i_2 - i_3) & = 0 & \textit{loop } 2 \\ 2(i_3 - i_2) & +i_3 & & = 0 & \textit{loop } 3 \end{array}$$

In matrix form:

$$\begin{bmatrix} 3 & -2 & 0 \\ -2 & 5 & -2 \\ 0 & -2 & 3 \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \\ i_3 \end{bmatrix} = \begin{bmatrix} 21 \\ 0 \\ 0 \end{bmatrix}$$

Now

$$\begin{aligned} \begin{bmatrix} 3 & -2 & 0 & 21 \\ -2 & 5 & -2 & 0 \\ 0 & -2 & 3 & 0 \end{bmatrix} &\sim \begin{bmatrix} 3 & -2 & 0 & 21 \\ 0 & 11 & -6 & 42 \\ 0 & -2 & 3 & 0 \end{bmatrix} \\ &\sim \begin{bmatrix} 3 & -2 & 0 & 21 \\ 0 & 11 & -6 & 42 \\ 0 & 0 & 21 & 84 \end{bmatrix} \end{aligned}$$

So $i_3 = 4$, $i_2 = 6$ and $i_1 = 11$.

5 Take

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 2 \end{bmatrix}, \quad Z = \begin{bmatrix} m \\ c \end{bmatrix}.$$

Then the least squares best fit line $y = mx + c$ is found by solving $A^T AZ = A^T B$, *ie* the equation

$$\begin{bmatrix} 30 & 10 \\ 10 & 4 \end{bmatrix} \begin{bmatrix} m \\ c \end{bmatrix} = \begin{bmatrix} 13 \\ 4 \end{bmatrix}$$

which has solution $c = -\frac{1}{2}$ and $m = \frac{3}{5}$. So the best fit line is $y = \dfrac{3}{5}x - \dfrac{1}{2}$.

6 As in **Q 5** we find

$$\begin{bmatrix} 21 & 7 \\ 7 & 4 \end{bmatrix} \begin{bmatrix} m \\ c \end{bmatrix} = \begin{bmatrix} 22 \\ 6 \end{bmatrix}$$

giving the best fit line $y = \frac{46}{35}x - \frac{4}{5}$.

7 Using the table

x	-1	0	1	2
x^2	1	0	1	4
y	3	1	2	5

take

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 4 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 3 \\ 1 \\ 2 \\ 5 \end{bmatrix}, \quad Z = \begin{bmatrix} a \\ b \end{bmatrix},$$

and solve $A^TAZ = A^TB$. Now

$$A^TA = \begin{bmatrix} 18 & 6 \\ 6 & 4 \end{bmatrix} \qquad A^TB = \begin{bmatrix} 25 \\ 11 \end{bmatrix}$$

and the solutions are $b = \frac{4}{3}$, $a = \frac{17}{18}$, giving best fit quadratic $y = \frac{17}{18}x^2 + \frac{4}{3}$.

8 Taking

$$A = \begin{bmatrix} -1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 2 & 1 \\ 2 & 3 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} -0.4 \\ 0.3 \\ 1.6 \\ 2.5 \end{bmatrix}, \quad Z = \begin{bmatrix} a \\ b \\ c \end{bmatrix},$$

and forming the equation $A^T AZ = A^T B$ we obtain the system

$$\begin{aligned} 6a &+ 9b + 2c = 7.0 \\ 9a &+ 14b + 4c = 11.1 \\ 2a &+ 4b + 4c = 4.0 \end{aligned}$$

which has the solution $a = 0.30,\ b = 0.50,\ c = 0.35$. Thus the best fit plane is $z = 0.30x + 0.50y + 0.35$.

9 After taking logarithms, the two laws have the forms $\ln y = n \ln x + \ln a$ and $\ln y = kx + \ln c$. The corresponding table of data is:

x	1	2	3	4	5
$\ln x$	0	0.693	1.099	1.386	1.609
y	1	3	7	15	30
$\ln y$	0	1.099	1.946	2.708	3.401

.

For the first law, let $\alpha = \ln a$ and apply the least squares method to find n and α. Then taking

$$A = \begin{bmatrix} 0 & 1 \\ 0.693 & 1 \\ 1.098 & 1 \\ 1.386 & 1 \\ 1.609 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1.099 \\ 1.946 \\ 2.708 \\ 3.401 \end{bmatrix}, \quad Z = \begin{bmatrix} n \\ \alpha \end{bmatrix},$$

and forming the equation $A^T AZ = A^T B$, we obtain the system

$$\begin{array}{rcll} 6.1958n & + & 4.786\alpha & = 12.124 \\ 4.786n & + & 5\alpha & = 9.154 \end{array}$$

which has solution $n \simeq 2.082$, $\alpha \simeq -0.162$ or $a \simeq 0.850$. So the best fit law of this type is $y = f(x) = 0.850x^{2.082}$.
For the second law, let $\gamma = \ln c$ and apply the least squares method to find k and γ. Then taking

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \\ 5 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1.099 \\ 1.946 \\ 2.708 \\ 3.401 \end{bmatrix}, \quad Z = \begin{bmatrix} k \\ \gamma \end{bmatrix},$$

and forming the equation $A^T AZ = A^T B$, we obtain the system

$$\begin{array}{rcll} 55k & + & 15\gamma & = 35.873 \\ 15k & + & 5\gamma & = 9.1538 \end{array}$$

which has solution $k \simeq 0.841$, $\gamma \simeq -0.693$ or $c \simeq 0.500$. So the best fit law of this type is $y = g(x) = 0.500e^{0.841x}$.
To compare:

x	1	2	3	4	5
$f(x) = 0.85x^{2.082}$	0.85	3.599	8.371	15.237	24.248
$g(x) = 0.5e^{0.841x}$	1.159	2.688	6.233	14.452	33.510

Now we can calculate the sums of squares of the errors for each curve: $\sum_{i=1}^{5}[y_i - f(x_i)]^2 \simeq 35.4$, and $\sum_{i=1}^{5}[y_i - g(x_i)]^2 \simeq 13.3$. So g seems to be better.