Lecture-9
Algorithmic Mathematics(CSC545)

Prepared by Asst. Prof. Bal Krishna Subedi

CDCSIT, TU

Exponential Regression

As we know, exponential equation be

$y=ae^{bx}$………(1)

Taking log on both sides

Then eq^n (1) became,

Log y= log a + bx …………(2)

Then comparing eq^n (2) with linear model and changing values (y by logy) for finding value of log a and b…Then

$$b = \frac{n\sum x\ \log y-(\sum x)(\sum \log y)}{n(\sum x2)-(\sum x)2}$$

$$\text{Log } a = \sum \log y/n - b \sum x/n$$

Exponential Regression

- To find value of a , We have to take anti log of log a, then place into equation of $y= a\ e^{bx}$.

However, this problem can be solved by using the algorithm given in the previous section in the following way: let us rewrite the equation using the conventional variables x and y as

$$y = a x^b$$

If we take logarithm on both the sides, we get

$$\ln y = \ln a + b \ln x \tag{10.11}$$

This equation is similar in form to the linear equation and, therefore, using the same procedure we can evaluate the parameters a and b.

$$b = \frac{n \sum \ln x_i \ln y_i - \sum \ln x_i \sum \ln y_i}{n \sum (\ln x_i)^2 - \left(\sum \ln x_i\right)^2} \tag{10.12}$$

$$\ln a = R = \frac{1}{n}\left(\sum \ln y_i - b \sum \ln x_i\right)$$

$$a = e^R \tag{10.13}$$

Similarly, we can linearise the exponential model shown in Eq. (10.9) by taking logarithm on both the sides. This would yield

$$\ln P = \ln P_0 + kt \ln e$$

Since, $\ln e = 1$,

we have $\ln P = \ln P_o + kt$

This is similar to the linear equation (10.14)

$$y = a + bx$$

where $y = \ln P$, $a = \ln P_0$, $b = k$, and $x = t$. We can now easily determine a and b and then P_0 and k.

There is a third form of nonlinear model known as *saturation-growth-rate* equation, as shown below:

$$p = \frac{k_1 t}{k_2 + t} \tag{10.15}$$

This can be linearised by taking inversion of the terms. That is

$$\frac{1}{p} = \left(\frac{k_2}{k_1}\right)\frac{1}{t} + \frac{1}{k_1} \tag{10.16}$$

This is again similar to the linear equation

$$y = a + bx$$

where

$$y = \frac{1}{p}, \qquad x = \frac{1}{t}$$

$$a = \frac{1}{k_1}, \qquad b = \frac{k_2}{k_1}$$

Once we obtain a and b, they could be transformed back into the original form for the purpose of analysis.

Example 10.2

Given the data table

x	1	2	3	4	5
y	0.5	2	4.5	8	12.5

fit a power-function model of the form

$$y = ax^b$$

Various quantities required in equation (10.12) are tabulated below:

x_i	y_i	$\ln x_i$	$\ln y_i$	$(\ln x_i)^2$	$(\ln x_i)(\ln y_i)$
1	0.5	0	−0.6931	0	0
2	2	0.6931	0.6931	0.4805	0.4804
3	4.5	1.0986	1.5041	1.2069	1.6524
4	8	1.3863	2.0794	1.9218	2.8827
5	12.5	1.6094	2.5257	2.5903	4.0649
Sum		4.7874	6.1092	6.1995	9.0804

Using Eq. (10.12),

$$b = \frac{(5)(9.0804) - (4.7874)(6.1092)}{(5)(6.1995) - (4.7874)^2}$$

$$= \frac{45.402 - 29.2472}{30.9975 - 22.9192}$$

$$= 1.9998$$

$$\ln a = \frac{6.1092 - (1.9998)(4.7847)}{5}$$

$$= -0.6929$$

$$a = 0.5001$$

Thus, we obtain the power-function equation as

$$y = 0.5001\, x^{1.9998}$$

Note that the data have been derived from the equation

$$y = \frac{x^2}{2}$$

The discrepancy in the computed coefficients is due to roundoff errors.

10.4 FITTING A POLYNOMIAL FUNCTION

When a given set of data does not appear to satisfy a linear equation, we can try a suitable polynomial as a regression curve to fit the data. The least squares technique can be readily used to fit the data to a polynomial.

Consider a polynomial of degree $m - 1$

$$y = a_1 + a_2 x + a_3 x^2 + \dots + a_m x^{m-1} \quad (10.17)$$

$$= f(x)$$

If the data contains n sets of x and y values, then the sum of squares of the errors is given by

$$Q = \sum_{i=1}^{n} [y_i - f(x_i)]^2 \quad (10.18)$$

Since $f(x)$ is a polynomial and contains coefficients a_1, a_2, a_3, etc., we have to estimate all the m coefficients. As before, we have the following m equations that can be solved for these coefficients.

$$\frac{\partial Q}{\partial a_1} = 0$$

$$\frac{\partial Q}{\partial a_2} = 0$$

$$\cdots$$

$$\cdots$$

$$\cdots$$

$$\frac{\partial Q}{\partial a_m} = 0$$

Consider a general term,

$$\frac{\partial Q}{\partial a_j} = -2\sum_{i=1}^{n}[y_i - f(x_i)]\frac{\partial f(x_i)}{\partial a_j} = 0$$

$$\frac{\partial f(x_i)}{\partial a_j} = x_i^{j-1}$$

Thus, we have

$$\sum_{i=1}^{n}[y_i - f(x_i)]x_i^{j-1} = 0 \qquad j = 1, 2, \ldots, m$$

$$\sum\left[y_i x_i^{j-1} - x_i^{j-1} f(x_i)\right] = 0$$

Substituting for $f(x_i)$

$$\sum_{i=1}^{n} x_i^{j-1}\left(a_1 + a_2 x_i + a_3 x_i^2 + \ldots + a_m x_i^{m-1}\right) = \sum_{i=1}^{n} y_i x_i^{j-1}$$

These are m equations ($j = 1, 2 \ldots m$) and each summation is for $i = 1$ to n.

$$\begin{aligned}
&a_1 n + a_2 \sum x_i + a_3 \sum x_i^2 + \ldots + a_m \sum x_i^{m-1} = \sum y_i \\
&a_1 \sum x_i + a_2 \sum x_i^2 + a_3 \sum x_i^3 + \ldots + a_m \sum x_i^m = \sum y_i x_i \qquad (10.19) \\
&\vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
&a_1 \sum x_i^{m-1} + a_2 \sum x_i^m + a_3 \sum x_i^{m+1} + \ldots + a_m \sum x_i^{2m-2} = \sum y_i x_i^{m-1}
\end{aligned}$$

The set of m equations can be represented in matrix notation as follows

$$\mathbf{CA} = \mathbf{B}$$

where

$$\mathbf{C} = \begin{bmatrix} n & \sum x_i & \sum x_i^2 & \cdots & \sum x_i^{m-1} \\ \sum x_i & \sum x_i^2 & \sum x_i^3 & \cdots & \sum x_i^{m} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \sum x_i^{m-1} & \sum x_i^{m} & \cdots & \cdots & \sum x_i^{2m-2} \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_m \end{bmatrix} \qquad \mathbf{B} = \begin{bmatrix} \sum y_i \\ \sum y_i x_i \\ \sum y_i x_2 \\ \vdots \\ \sum y_i x_i^{m-1} \end{bmatrix}$$

Example 10.4

Fit a second order polynomial to the data in the table below:

x	1.0	2.0	3.0	4.0
y	6.0	11.0	18.0	27.0

The order of polynomial is 2 and therefore we will have 3 simultaneous equations as shown below:

$$a_1 n + a_2 \sum x_i + a_3 \sum x_i^2 = \sum y_i$$

$$a_1 \sum x_i + a_2 \sum x_i^2 + a_3 \sum x_i^3 = \sum y_i x_i$$

$$a_1 \sum x_i^2 + a_2 \sum a_i^3 + a_3 \sum x_i^4 = \sum y_i x_i^2$$

The sums of powers and products can be evaluated in a tabular form as shown below:

	x	y	x^2	x^3	x^4	yx	yx^2
	1	6	1	1	1	6	6
	2	11	4	8	16	22	44
	3	18	9	27	81	54	162
	4	27	16	64	256	108	432
Σ	10	62	30	100	354	190	644

Substituting these values, we get

$$4a_1 + 10a_2 + 30a_3 = 62$$
$$10a_1 + 30a_2 + 100a_3 = 190$$
$$30a_1 + 100a_2 + 354a_3 = 644$$

Solving these equations gives

$$a_1 = 3$$
$$a_2 = 2$$
$$a_3 = 1$$

Therefore, the least squares quadratic polynomial is

$$y = 3 + 2x + x^2 \qquad \text{(verify using table data)}$$

10.5 MULTIPLE LINEAR REGRESSION

There are a number of situations where the dependent variable is a function of two or more variables. For example, the salary of a salesperson may be expressed as

$$y = 500 + 5x_1 + 8x_2$$

where x_1 and x_2 are the number of units sold of products 1 and 2, respectively. We shall discuss here an approach to fit the experimental data where the variable under consideration is a linear function of two independent variables.

Let us consider a two-variable linear function as follows:

$$y = a_1 + a_2x + a_3z \qquad (10.20)$$

The sum of the squares of errors is given by

$$Q = \sum_{i=1}^{n}(y_i - a_1 - a_2x_i - a_3z_i)^2$$

Differentiating with respect to a_1, a_2 and a_3, we get,

$$\frac{\partial Q}{\partial a_1} = -2\sum (y_i - a_1 - a_2 x_i - a_3 z_i)$$

$$\frac{\partial Q}{\partial a_2} = -2\sum (y_i - a_1 - a_2 x_i - a_3 z_i)\, x_i$$

$$\frac{\partial Q}{\partial a_3} = -2\sum (y_i - a_1 - a_2 x_i - a_3 z_i)\, y_i$$

Setting these partial derivatives equal to zero results in

$$n a_1 + \left(\sum x_i\right) a_2 + \left(\sum z_i\right) a_3 = \sum y_i$$

$$\left(\sum x_i\right) a_1 + \left(\sum x_i^2\right) a_2 + \left(\sum x_i z_i\right) a_3 = \sum y_i x_i$$

$$\left(\sum z_i\right) a_1 + \left(\sum x_i z_i\right) a_2 + \left(\sum z_i^2\right) a_3 = \sum y_i z_i$$

These are three simultaneous equations with three unknowns and, therefore, can be expressed in matrix form as

$$\begin{bmatrix} n & \sum x_i & \sum z_i \\ \sum x_i & \sum x_i^2 & \sum x_i z_i \\ \sum z_i & \sum x_i z_i & \sum z_i^2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i x_i \\ \sum y_i z_i \end{bmatrix} \quad (10.21)$$

This equation can be solved using any standard method. This is a two-dimensional case and, therefore, we obtain a regression "plane" rather than "line".

We can easily extend Eq. (10.21) to the more general case

$$y = a_1 + a_2 x_1 + a_3 x_2 + \ldots + a_{m+1} x_m$$

Example 10.6

Given the table of data

x	1	2	3	4
z	0	1	2	3
y	12	18	24	30

Obtain a regression plane to fit the data.

The various sums of powers and products required for evaluation of coefficients are tabulated below:

	x	z	y	x^2	z^2	xz	yx	yz
	1	0	12	1	0	0	12	0
	2	1	18	4	1	2	36	18
	3	2	24	9	4	6	72	48
	4	3	30	16	9	12	120	90
Σ	10	6	84	30	14	20	240	156

On substitution of these values in Eq. (10.21) we get

$$4a_1 + 10a_2 + 5a_3 = 84$$

$$10a_1 + 30a_2 + 20a_3 = 240$$

$$6a_1 + 20a_3 + 14a_3 = 156$$

Solution of these equations results in

$$a_1 = 10$$

$$a_2 = 2$$

$$a_3 = 4$$

Thus, the regression plane is

$$y = 10 + 2x + 4z$$

Assignment# 9a

Use the exponential model

$$y = a\,e^{bx}$$

to fit the data

x	0.4	0.8	1.2	1.6	2.0	2.4
y	75	100	140	200	270	375

Assignment# 9b

- Prepare a report on "B-splines and Approximation of function in curve fitting"

Thank You

Any Query??