

Non-Linear Models

Mohammad Ehsanul Karim <wildscop@yahoo.com>
 Institute of Statistical Research and training;
 University of Dhaka, Dhaka – 1000, Bangladesh

Estimating Parameters from a Non-linear Regression Model

Illustrative Example:

Let the data¹ file is in “c:\ draper.txt”, which is as follows:

t	Y
1	0.80
4	0.45
16	0.04

Let that, we have to estimate the parameter q in the non-linear model $Y = e^{-qt} + e$ from the above observations.

Now for starting analysis, we need a “initial value”. But we have no prior information about what values q might take. There are a few suggestions for such situations: one of which is to make a grid search using Statistical Software. Let us demonstrate how we can do this. See updates on this at <http://www.angelfire.com/ab5/get5/nlfinal.PDF>

Using Statistical Software: SAS: Making a Grid Search

We write the following program in SAS (I used SAS 6.12) to get an estimate (note that we told SAS to do the search in Linearization technique by setting METHOD=NEWTON):

```
DATA DS;
INPUT T Y;
DATALINES;
1 0.80
4 0.45
16 0.04
RUN;
```

¹We are using this simple and small data set just for easy illustration purpose. This data set is taken from Draper, Smith (1998) “Applied Regression Analysis”, Third ed., Page 553, Exercise 24.A.

```

PROC NLIN DATA=DS METHOD=NEWTON;
PARAMETERS G=-100 TO 100 BY .1;
MODEL Y = EXP(-(G*T));
OUTPUT OUT=NLINOUT PREDICTED=PRED L95M=L95MEAN
U95M=U95MEAN L95=L95IND U95=U95IND;
RUN;

```

```

PROC PRINT DATA=NLINOUT;
RUN;

```

There will be an enormous search to find this estimate since we assumed our potential parameter (in the program, we wrote $q = G$) space to be -100 to 100 and told SAS to look within each 0.1 interval (so that we can get minimum residual SS or minima). However, after some iteration (when a convergent value will be found) SAS will bring the following output (edited to shorten the length – but the key parts of the output is unedited):

The SAS System 11:52 Sunday,

(Detailed Results on Non-Linear Least Squares Grid Search is omitted from here.)

Non-Linear Least Squares Iterative Phase

Dependent Variable Y		Method: Newton
Iter	G	Sum of Squares
0	0.200000	0.000352
1	0.203367	0.000302
2	0.203449	0.000302
3	0.203449	0.000302

NOTE: Convergence criterion met.

The SAS System 11:52 Sunday, April 6, 1997 50

Non-Linear Least Squares Summary Statistics			Dependent Variable Y
Source	DF	Sum of Squares	Mean Square
Regression	1	0.84379816630	0.84379816630
Residual	2	0.00030183370	0.00015091685
Uncorrected Total	3	0.84410000000	
(Corrected Total)	2	0.28940000000	

Parameter	Estimate	Asymptotic Std. Error	Asymptotic 95 % Confidence Interval Lower Upper
G	0.2034488738	0.00603799034	0.17746921275 0.22942853484

Asymptotic Correlation Matrix

Corr	G
ffffffffffffffffffff	
G	1

The SAS System 11:52 Sunday, April 6, 1997 51

OBS	T	Y	PRED	L95MEAN	U95MEAN	L95IND	U95IND
1	1	0.80	0.81591	0.79472	0.83711	0.75896	0.87286
2	4	0.45	0.44317	0.39712	0.48923	0.37307	0.51328
3	16	0.04	0.03857	0.02254	0.05461	-0.01666	0.09381

Therefore, from SAS, we get estimate of $q = 0.2034488738$. So we may suppose that the actual value may be around zero for now and check it in other ways.

Calculation By Hand: Non-linear Regression:

We will now go through some iterations just to see how the process works. Readers may think it as a slow motion of what happens to any Statistical Software when we run a non-linear regression.

Iteration 0: We have the function $Y = f(\mathbf{x}, \mathbf{q}) + \mathbf{e} = e^{-qt} + \mathbf{e}$ in our hand with approximated initial value $\mathbf{q}_0 = 0$ (as our grid search indicates – note that, since we have to estimate only one value, this is just a scaler, but for multi-parameter case, we would have a vector).

$$f^0 = f(\mathbf{x}, \mathbf{q}_0) = e^{-0xt} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \text{ since } t \text{ is a } 3 \times 1 \text{ vector.}$$

$$\text{Now, } y_0 = \langle Y - f^0 \rangle = \begin{bmatrix} Y_1 - f_1^0 \\ Y_2 - f_2^0 \\ Y_3 - f_3^0 \end{bmatrix} = \begin{bmatrix} 0.80 - 1 \\ 0.45 - 1 \\ 0.04 - 1 \end{bmatrix} = \begin{bmatrix} -0.20 \\ -0.55 \\ -0.96 \end{bmatrix}$$

$$Z_0 = \left[\frac{\partial f(\mathbf{x}, \mathbf{q})}{\partial \mathbf{q}} \right]_{\mathbf{q}=\mathbf{q}_0} = (-t) \times e^{-0xt} = \begin{bmatrix} -1 \\ -4 \\ -16 \end{bmatrix}$$

$$Z_0'Z_0 = 1^2 + 4^2 + 16^2 = 273$$

$$(Z_0'Z_0)^{-1} = (273)^{-1} = 0.003663004$$

$$Z_0'(Y - f^0) = (-0.20) \times (-1) + (-0.55) \times (-4) + (-0.96) \times (-16) = 17.76$$

$$b_0 = (Z_0'Z_0)^{-1} Z_0'(Y - f^0) = 0.003663004 \times 17.76 = 0.06505495$$

$$\hat{\mathbf{q}}_1 = \hat{\mathbf{q}}_0 + b_0 = 0.06505495 + 0 = 0.06505495$$

Iteration 1: We have the function $Y = f(\mathbf{x}, \mathbf{q}) + \mathbf{e} = e^{-qt} + \mathbf{e}$ in our hand with approximated value $\hat{\mathbf{q}}_1 = 0.06505495$.

$$f^1 = f(\mathbf{x}, \mathbf{q}_1) = e^{-\hat{q}_1 t} = \begin{bmatrix} 0.9370160 \\ 0.7708821 \\ 0.3531441 \end{bmatrix} \text{ since } t \text{ is a } 3 \times 1 \text{ vector.}$$

$$\text{Now, } y_1 = \langle Y - f^1 \rangle = \begin{bmatrix} Y_1 - f_1^1 \\ Y_2 - f_2^1 \\ Y_3 - f_3^1 \end{bmatrix} = \begin{bmatrix} -0.1370160 \\ -0.3208821 \\ -0.3131441 \end{bmatrix}$$

$$Z_1 = \left[\frac{\partial f(\mathbf{x}, \mathbf{q})}{\partial \mathbf{q}} \right]_{\mathbf{q}=\hat{\mathbf{q}}_1} = (-t) \times e^{-\hat{q}_1 t} = \begin{bmatrix} -0.937016 \\ -3.083529 \\ -5.650305 \end{bmatrix}$$

$$Z_1'Z_1 = 42.3121$$

$$(Z_1'Z_1)^{-1} = (42.3121)^{-1}$$

$$Z_1'(Y - f^1) = 2.887195$$

$$b_1 = (Z_1'Z_1)^{-1}Z_1'(Y - f^1) = 0.06823569$$

$$\hat{\mathbf{q}}_2 = \hat{\mathbf{q}}_1 + b_1 = 0.1332906$$

Iteration 2: We have the function $Y = f(\mathbf{x}, \mathbf{q}) + \mathbf{e} = e^{-qt} + \mathbf{e}$ in our hand with approximated value $\hat{\mathbf{q}}_1 = 0.1332906$.

$$f^2 = f(\mathbf{x}, \mathbf{q}_2) = e^{-\hat{q}_2 t} = \begin{bmatrix} 0.8752107 \\ 0.5867464 \\ 0.1185228 \end{bmatrix} \text{ since } t \text{ is a } 3 \times 1 \text{ vector.}$$

$$\text{Now, } y_2 = \langle Y - f^2 \rangle = \begin{bmatrix} Y_1 - f_1^2 \\ Y_2 - f_2^2 \\ Y_3 - f_3^2 \end{bmatrix} = \begin{bmatrix} -0.07521069 \\ -0.13674643 \\ -0.07852278 \end{bmatrix}$$

$$Z_2 = \left[\frac{\partial f(\mathbf{x}, \mathbf{q})}{\partial \mathbf{q}} \right]_{\mathbf{q}=\hat{\mathbf{q}}} = (-t) \times e^{-\hat{q}_2 t} = \begin{bmatrix} -0.8752107 \\ -2.3469857 \\ -1.8963644 \end{bmatrix}$$

$$Z_2'Z_2 = 9.870534$$

$$(Z_2'Z_2)^{-1} = (9.870534)^{-1}$$

$$Z_2'(Y - f^2) = 0.5356749$$

$$b_2 = (Z_2'Z_2)^{-1}Z_2'(Y - f^2) = 0.05427011$$

$$\hat{\mathbf{q}}_3 = \hat{\mathbf{q}}_2 + b_2 = 0.1875607$$

Iteration 3: We have the function $Y = f(\mathbf{x}, \mathbf{q}) + \mathbf{e} = e^{-qt} + \mathbf{e}$ in our hand with approximated value $\hat{\mathbf{q}}_3 = 0.1875607$.

$$f^3 = f(\mathbf{x}, \mathbf{q}_3) = e^{-\hat{q}_3 t} = \begin{bmatrix} 0.82897877 \\ 0.47225180 \\ 0.04973871 \end{bmatrix} \text{ since } t \text{ is a } 3 \times 1 \text{ vector.}$$

$$\text{Now, } y_3 = \langle Y - f^3 \rangle = \begin{bmatrix} Y_1 - f_1^3 \\ Y_2 - f_2^3 \\ Y_3 - f_3^3 \end{bmatrix} = \begin{bmatrix} -0.028978765 \\ -0.022251802 \\ -0.009738708 \end{bmatrix}$$

$$Z_3 = \left[\frac{\partial f(\mathbf{x}, \mathbf{q})}{\partial \mathbf{q}} \right]_{\mathbf{q}=\hat{\mathbf{q}}} = (-t) \times e^{-\hat{q}_3 t} = \begin{bmatrix} -0.8289788 \\ -1.8890072 \\ -0.7958193 \end{bmatrix}$$

$$Z_3'Z_3 = 4.888882$$

$$(Z_3'Z_3)^{-1} = (4.888882)^{-1}$$

$$Z_3'(Y - f^3) = 0.07380705$$

$$b_3 = (Z_3'Z_3)^{-1}Z_3'(Y - f^3) = 0.01509691$$

$$\hat{\mathbf{q}}_4 = \hat{\mathbf{q}}_3 + b_3 = 0.2026576$$

Iteration 4: We have the function $Y = f(\mathbf{x}, \mathbf{q}) + \mathbf{e} = e^{-qt} + \mathbf{e}$ in our hand with approximated value $\hat{\mathbf{q}}_4 = 0.2026576$

$$f^4 = f(\mathbf{x}, \mathbf{q}_4) = e^{-\hat{\mathbf{q}}_4 t} = \begin{bmatrix} 0.81655777 \\ 0.44457770 \\ 0.03906526 \end{bmatrix} \text{ since } t \text{ is a } 3 \times 1 \text{ vector.}$$

$$\text{Now, } y_4 = \langle Y - f^4 \rangle = \begin{bmatrix} Y_1 - f_1^4 \\ Y_2 - f_2^4 \\ Y_3 - f_3^4 \end{bmatrix} = \begin{bmatrix} -0.0165577742 \\ 0.0054223029 \\ 0.0009347429 \end{bmatrix}$$

$$Z_4 = \left[\frac{\partial f(\mathbf{x}, \mathbf{q})}{\partial \mathbf{q}} \right]_{\mathbf{q}=\hat{\mathbf{q}}} = (-t) \times e^{-\hat{\mathbf{q}}_4 t} = \begin{bmatrix} -0.8165578 \\ -1.7783108 \\ -0.6250441 \end{bmatrix}$$

$$Z_4' Z_4 = 4.219836$$

$$(Z_4' Z_4)^{-1} = (4.219836)^{-1}$$

$$Z_4'(Y - f^4) = 0.003293584$$

$$b_4 = (Z_4' Z_4)^{-1} Z_4'(Y - f^4) = 0.0007805005$$

$$\hat{\mathbf{q}}_5 = \hat{\mathbf{q}}_4 + b_4 = 0.2034381$$

Iteration 5: We have the function $Y = f(\mathbf{x}, \mathbf{q}) + \mathbf{e} = e^{-qt} + \mathbf{e}$ in our hand with approximated value $\hat{\mathbf{q}}_5 = 0.2034381$

$$f^5 = f(\mathbf{x}, \mathbf{q}_5) = e^{-\hat{\mathbf{q}}_5 t} = \begin{bmatrix} 0.81592070 \\ 0.44319189 \\ 0.03858044 \end{bmatrix} \text{ since } t \text{ is a } 3 \times 1 \text{ vector.}$$

$$\text{Now, } y_5 = \langle Y - f^5 \rangle = \begin{bmatrix} Y_1 - f_1^5 \\ Y_2 - f_2^5 \\ Y_3 - f_3^5 \end{bmatrix} = \begin{bmatrix} -0.015920699 \\ 0.006808111 \\ 0.001419557 \end{bmatrix}$$

$$Z_5 = \left[\frac{\partial f(\mathbf{x}, \mathbf{q})}{\partial \mathbf{q}} \right]_{\mathbf{q}=\hat{\mathbf{q}}} = (-t) \times e^{-\hat{\mathbf{q}}_5 t} = \begin{bmatrix} -0.8159207 \\ -1.7727676 \\ -0.6172871 \end{bmatrix}$$

$$Z_5' Z_5 = 4.189475$$

$$(Z_5' Z_5)^{-1} = (4.189475)^{-1}$$

$$Z_5'(Y - f^5) = 4.455581 \times 10^{-05}$$

$$b_5 = (Z_5' Z_5)^{-1} Z_5'(Y - f^5) = 1.063518 \times 10^{-05}$$

$$\hat{\mathbf{q}}_6 = \hat{\mathbf{q}}_5 + b_5 = 0.2034487$$

Iteration 6: We have the function $Y = f(\mathbf{x}, \mathbf{q}) + \mathbf{e} = e^{-qt} + \mathbf{e}$ in our hand with approximated value $\hat{\mathbf{q}}_6 = 0.2034487$

$$f^6 = f(\mathbf{x}, \mathbf{q}_6) = e^{-\hat{q}_6 t} = \begin{bmatrix} 0.81591202 \\ 0.44317304 \\ 0.03857388 \end{bmatrix} \text{ since } t \text{ is a } 3 \times 1 \text{ vector.}$$

$$\text{Now, } y_6 = \langle Y - f^6 \rangle = \begin{bmatrix} Y_1 - f_1^6 \\ Y_2 - f_2^6 \\ Y_3 - f_3^6 \end{bmatrix} = \begin{bmatrix} -0.015912022 \\ 0.006826964 \\ 0.001426121 \end{bmatrix}$$

$$Z_6 = \left[\frac{\partial f(\mathbf{x}, \mathbf{q})}{\partial \mathbf{q}} \right]_{\mathbf{q}=\hat{\mathbf{q}}} = (-t) \times e^{-\hat{q}_6 t} = \begin{bmatrix} -0.8159120 \\ -1.7726921 \\ -0.6171821 \end{bmatrix}$$

$$Z_6' Z_6 = 4.189064$$

$$(Z_6' Z_6)^{-1} = (4.189064)^{-1}$$

$$Z_6' (Y - f^6) = 5.276473 \times 10^{-07}$$

$$b_6 = (Z_6' Z_6)^{-1} Z_6' (Y - f^6) = 1.259583 \times 10^{-07}$$

$$\hat{\mathbf{q}}_7 = \hat{\mathbf{q}}_6 + b_6 = 0.2034489$$

Iteration 7: We have the function $Y = f(\mathbf{x}, \mathbf{q}) + \mathbf{e} = e^{-qt} + \mathbf{e}$ in our hand with approximated value $\hat{\mathbf{q}}_7 = 0.2034489$

$$f^7 = f(\mathbf{x}, \mathbf{q}_7) = e^{-\hat{q}_7 t} = \begin{bmatrix} 0.8159119 \\ 0.4431728 \\ 0.0385738 \end{bmatrix} \text{ since } t \text{ is a } 3 \times 1 \text{ vector.}$$

$$\text{Now, } y_7 = \langle Y - f^7 \rangle = \begin{bmatrix} Y_1 - f_1^7 \\ Y_2 - f_2^7 \\ Y_3 - f_3^7 \end{bmatrix} = \begin{bmatrix} -0.015911919 \\ 0.006827188 \\ 0.001426199 \end{bmatrix}$$

$$Z_7 = \left[\frac{\partial f(\mathbf{x}, \mathbf{q})}{\partial \mathbf{q}} \right]_{\mathbf{q}=\hat{\mathbf{q}}} = (-t) \times e^{-\hat{q}_7 t} = \begin{bmatrix} -0.8159119 \\ -1.7726912 \\ -0.6171808 \end{bmatrix}$$

$$Z_7' Z_7 = 4.189059$$

$$(Z_7' Z_7)^{-1} = (4.189059)^{-1}$$

$$Z_7' (Y - f^7) = 6.236455 \times 10^{-09}$$

$$b_7 = (Z_7' Z_7)^{-1} Z_7' (Y - f^7) = 1.488749 \times 10^{-09}$$

$$\hat{\mathbf{q}}_8 = \hat{\mathbf{q}}_7 + b_7 = 0.2034489$$

Since $\hat{\mathbf{q}}_8 = \hat{\mathbf{q}}_7$ we terminate the process and declare $\hat{\mathbf{q}}_8 = 0.2034489$ as our desired estimate.

Using Statistical Software: R: Non-linear Regression

Now, using Statistical Software like R (I used R 1.7.1) we can do the whole procedure just in a blink (and of course, without pain!) by writing the following commands in the R console:

```
> options(prompt=" R> " )
R> draper<-read.table("c:\\draper.txt",header=T)
R> draper.data<-data.frame(draper)
R> attach(draper.data)
R> nls(Y~exp(-(g*t)), data=draper.data, start = c(g =
0), trace = T)
1.2641 : 0
0.2197979 : 0.06505495
0.03052205 : 0.1332906
0.001429753 : 0.1875607
0.000304435 : 0.2026576
0.0003018342 : 0.2034381
0.0003018337 : 0.2034487
0.0003018337 : 0.2034489
Nonlinear regression model
  model: Y ~ exp(-(g * t))
  data: draper.data
      g
0.2034489
residual sum-of-squares: 0.0003018337

R> coef(nls(Y~exp(-(g*t)), data=draper.data, start = c(g
= 0), trace = T))
(---)
      g
0.2034489
```

Note that, the estimated value of $q = 0.2034489$ approximately in all our (slow motion hand calculation and R results within a blink) cases – which matches our SAS result too. Thus, we declare 0.2034489 as our estimate of the parameter q based on the given data (See <http://snipurl.com/ehsan> for the details discussions on it and more).