Nonlinear Regression Using the "Big Three" Spreadsheets

Roy W. Clark

Middle Tennessee State University, Murfreesboro, TN 37132, royclark@bellsouth.net

Abstract: Nonlinear regression using the spreadsheet Excel has appeared in the literature in several places. This paper extends the use of this technique to two other popular spreadsheets, Lotus 123 and Quattro Pro. Associated files, which can be downloaded, are included. These are the spreadsheet templates and ScreenCam movies that illustrate the use of equation solvers for minimization in these spreadsheets.

In a concise and excellent paper, Daniel C. Harris [1] has shown how a modern spreadsheet, Excel [2], can easily be made to do nonlinear regression. Weighting of the data and determination of the standard deviation in the parameters are covered as well. It is the purpose of this paper to extend Harris's method to the other two spreadsheets in wide use, Lotus 123 [3] and Quattro Pro [4], and to provide templates and movies for the less-than-expert spreadsheet user to succeed at nonlinear regression. Using nonlinear regression, data are fitted not to a best straight line but to some more complex nonlinear equation.

Examples in chemistry include exponential equations, quadratic and cubic equations, enzyme kinetics equations, and (Harris' example) the van Deemter equation used in chromatography. The spreadsheets provided here use Beer's law exponential decay, but the templates can be modified without difficulty by substituting most any nonlinear equation. Whereas linear regression adjusts two parameters, the slope and intercept of the straight line, using the fitting program, nonlinear regression may optimize three or more parameters to obtain the best fit of the data. In the examples here there are only two adjustable parameters. Needless to say, the fewer parameters the program is expected to adjust, the easier the task.

Beer's law data are usually plotted in the form of a straight line where *A*, the absorbance, is a function of *C*, the concentration of the solutions. Actually, this is a transformed version of Beer's law, $I = I_0 \times 10^{-\varepsilon bC}$. *I* is the intensity of the light exiting the cell; I_0 is the light entering the cell; *C* is the concentration of the solution; *b* is the path length; finally, ε (or sometimes *a*) is the decay constant, often called the molar absorptivity. Instruments measure the ratio of *I* to I_0, called the transmittance, *T*. Thus, Beer's law is really $T = 10^{-\varepsilon bC}$. Fitting with only one adjustable parameter, ε, would be equivalent to plotting the equation in the linear form and forcing it to go through zero. This is not done for linear Beer's law plots because the nonzero intercept is interpreted as an indication of the detection limits of the method; so, in the nonlinear form $T = T_0 \times 10^{-\varepsilon bC}$, where the symbol T_0 is another adjustable parameter. This symbol is chosen because there is no widely recognized symbol agreed upon for this, and this pre-exponential factor equals *T* when *C* is zero.

Taking the base-10 logarithm of the above equation results in

$$\log T = \log T_0 - \varepsilon bC$$

Multiplying through by –1 gives

$$-\log T = -\log T_0 + \varepsilon bC$$

A, the absorbance, is defined as –log *T* and so

$$A = -\log T_0 + \varepsilon bC$$

The intercept in straight-line Beer's law plots is –log T_0. T_0 is ideally 1, and the straight-line intercept is ideally zero. The path length is usually 1 cm; so, if we ignore *b*, the relevant equation for nonlinear regression in our spreadsheets is $T = T_0 \times 10^{-1\times k\times C}$.

T_0 and *k* are the adjustable parameters. We now only need *T* versus *C* data on which to perform the regression. The templates included here contain a set of such data. You can easily substitute your own. The movie files are ScreenCam [5] files showing the mouse actions to perform for each of the three spreadsheets. You may download the movie that is appropriate for your spreadsheet program. When viewing the movies, the spacebar can be used to pause. Tapping the spacebar again resumes the movie.

The following files are available for downloading: NONLNREG.123, the template for Lotus 123-97; NONLNREG.wb3, the template for Quattro Pro 8; NONLNREG.xl5, the template for Excel 97 (43rc1897.zip); NLREG123.exe, the movie for Lotus 123; NLREGQTR.exe, the movie for Quattro Pro; and NLREGEXL.EXE, the Movie for Excel 97 (43rc2897.zip).

Some Notes and Precautions

1) In the example spreadsheets the two parameters to be adjusted were both initially set to an arbitrary value of 1. For equations more complex than this one, the initial parameters may be more critical. As a general rule, set the initial parameters to values that approximate the expected values.
2) Some hand-held graphing calculators, for example the Texas Instruments TI-83 [6], seem to do exponential data fits, but, in reality, they do linear fits to the logarithmic form of the function. This does not yield the same parameters as true nonlinear fits. This technique is not recommended.
3) When using Lotus 123-97, there is another, seemingly simpler method that can be used to fit an exponential equation. Just enter the data, draw a scatter plot, select the

series (*y* column), and choose "series trend." The choices are linear, exponential, logarithmic, and power. Choose exponential, and the fit will be performed (to the base e). This, however, is another case in which a linear fit is performed and reported as exponential. As in the case for the TI-83 above, this is not a true nonlinear regression.

References

1. Harris, D. C. *J Chem. Educ.* **1998**, *75*, 119.
2. *Excel*™ is a software program produced by Microsoft Corp.
3. *Lotus -123*™ is a software program produced by Lotus Corp.
4. *Quattro Pro*™ is a software program produced by Corel Corp.
5. *Screen Cam*™ is a software program produced by Lotus Corp.
6. *TI-83* is a graphing calculator produced by Texas Instruments Corp.