Excel Solver Reference

Tanvi Karwal

Refer to chapter 8 of Taylor for details on all of the following.

Formula for a χ^2

A χ^2 determines how well your data fits your model, or vice versa. It quantifies this through the formula:

$$\chi^2 = \sum_{i=1}^{N} (y_i^{data} - y_i^{theory})^2 \tag{1}$$

where you have N measurements of observations y_i^{data} and theoretically expected values y_i^{theory}. It is essentially the sum of squared differences between your theory and your data. The smaller your χ^2, the better your model and data agree.

In eq. (1), we have defined quite a simple χ^2. It is completely independent of the errors on the data and the theory. A more sophisticated fit includes a weight for each data point to account for the associated error. We would want our χ^2 to be more sensitive to changes in an observation that has a small error and more insensitive to changes in an observation with a large error. We would do this by altering eq. (1) as follows:

$$\boxed{\chi^2 = \sum_{i=1}^{N} \frac{(y_i^{data} - y_i^{theory})^2}{\sigma_i^2}} \tag{2}$$

The error σ_i here should include the errors on your data as well as your theory. Using the quadrature rule (chapter 3),

$$\boxed{\sigma_i^2 = \sigma_{i,data}^2 + \sigma_{i,theory}^2} \tag{3}$$

The above works not just for a linear model (or linearised data), but for any model. That is, y^{theory} can be any complicated function of your parameters and independent variables:

$$\begin{gathered} y = ax + b \\ y = ax^2 + bx + c \\ y = \sin(ax + b) - \cos(cz) \\ y = a\exp(-x/b) \end{gathered} \tag{4}$$

In these examples, your **parameters** are a, b, c and your **independent variables** are x, z.

For a linear model

In a linear model, the above formulae simplify to a great extent. You would calculate y^{theory} as:

$$\boxed{y_i^{theory} = mx_i + c} \tag{5}$$

where m is your slope and c, your y-intercept. The error σ_{theory} is:

$$\boxed{\sigma_{i,theory} = m\delta x_i} \tag{6}$$

where δx_i is the error on the ith x-value.

Your total error then becomes:

$$\boxed{\sigma_i^2 = (\delta y_i)^2 + (m\delta x_i)^2} \tag{7}$$

where δy_i is the error on the measured y-value.

Finally, your total χ^2 becomes:

$$\boxed{\chi^2 = \sum_{i=1}^{N} \frac{(y_i^{data} - mx_i - c)^2}{(\delta y_i)^2 + (m\delta x_i)^2}} \tag{8}$$

Programming with Solver

In order to constrain the parameters of your model based on data, you can do the following in Excel:

1) Linearise your data such that your y is a linear function of your x. Also determine the errors δy and δx. Refer to Taylor chapter 3 for more on how to propagate errors.

2) Postulate, based on theory, what the slope m and y-intercept c of your model should be. Record these values in spreadsheet cells.

3) Use your linear model to calculate theoretical values y^{theory} and their errors σ_{theory} based on your x-values, m and c. Use cells as variables to calculate these.

4) Calculate the χ^2 (as shown in eq. (8)) in a spreadsheet cell.

5) Use Solver to find a fit by minimising the χ^2-cell, while varying the parameters m and c. Usually, you should also ensure m and c are permitted to be negative.

The above will most often give you a solution. Don't panic if it doesn't. Caveats include a reasonable model for the data, and other local minima being in close proximity.

If you attempt to fit obviously non-linear data with a linear model, Solver might have trouble finding you a good solution. Therefore having a reasonable model is important, or recognising what subset of your data to fit.

Solver finds these solutions by taking a random step away from your postulated solution and determining whether the χ^2 improves in that direction. Therefore, it is possible for Solver to get stuck in local minima, or locally optimal solutions. One solution to this problem is running Solver a few times, starting with a different postulated solution each time. Then, your globally optimal solution is the one for which the value of χ^2 is smallest.