Tutorial on Nonlinear Regression

<u>Introduction:</u>

In any simple nonlinear regression problem, one has to define the model in a form that contains only one dependent variable (y) as a function of a specified number of independent input variables (x_1, x_2, …, x_N) and fitting parameters (a_1, a_2, …, a_M). Then, the following regression procedure is carried out for the optimization of the fitting parameters that best describe y versus the independent variables with a well-more than (M+1)-point experimental data set.

1. Provide reasonable initial guesses for all of the fitting parameters.
2. Calculate a new variable, $y_{\text{calc},i}$, that applies the assumed fitting parameters to every data point (i) of the experimental data set.
3. Calculate the squared error (SE_i) objective function for every data point as shown below. Attention should be given to alternative forms of the objective function which may result in better correlations as shown in the following examples.

$$SE_i = \left(f\left(y_{calc,i}\right) - f\left(y_{\exp,i}\right)\right)^2 \tag{1}$$

   where f represents the best form for the objective function in terms of y.

4. Add the objective functions to obtain the sum-squared error (SSE):

$$SSE = \sum_{i=1}^{N_p} SE_i \tag{2}$$

   where N_p is the number of experimental data points.

5. Apply an optimization subroutine (e.g., the Microsoft Excel Solver) for minimizing the SSE by changing the fitting parameters. The final fitting parameters that result in the least SSE (LSSE) are the final results for the linear regression.
6. Calculate the average relative error (ARE) for the obtained correlation:

$$ARE(\%) = \frac{100}{N_p} \sum_{i=1}^{N_p} \frac{\left|y_{\exp,i} - y_{calc,i}\right|}{y_{\exp,i}} \tag{3}$$

Using the Microsoft Excel Solver for Nonlinear Optimization

Before using the Microsoft Excel Solver, make sure that the solver add-in is installed with your copy of Microsoft Excel. The following procedure is performed to install the Solver add-in:

1. On the Tools menu, click Add-Ins. If Solver is not listed in the Add-Ins dialog box, click Browse and locate the drive, folder, and file name for the "Solver.xla" add-in usually located in the "Library\Solver\" (e.g., "C:\win32app\MSOFFICE\Office\Library\Solver\") folder or run the Setup program if you can't locate the file.
2. In the Add-Ins dialog box, select the Solver Add-in check box. Note: Add-ins you select in the Add-Ins dialog box remain active until you remove them.

To Run the Solver for optimizing the problem, go to the Tools menu and select the solver. Then define the following parameters:

- Set Target Cell: Choose the cell that contains the objective function (in this case, the LSSE)
- Equal to: choose one of the optimization operations (Maximum, <u>Minimum</u>, Equal to: value).
- By changing cells: define the spreadsheet cell(s) that contain the fitting parameters.
- Constraints: add any constraints you may need.

You can click the "Options" button to select different options (e.g., search methods, positive fitting parameters, … etc.), and make sure to read through the help files to get more acquainted with different features of the Solver.

Examples

In each of the following examples, a dependent variable is defined and four forms of objective functions are minimized for obtaining the fitting parameters. These four forms are:

- LSSE in the dependent variable:

$$SSE1 = \sum_{i=1}^{N_p} \left(y_{calc,i} - y_{\exp,i} \right)^2 \qquad (4)$$

- Minimum ARE as defined in Equation 3.
- Minimum sum square error relative to experimental dependent variables:

$$SSE3 = \sum_{i=1}^{N_p} \frac{\left(y_{calc,i} - y_{\exp,i}\right)^2}{y_{\exp,i}} \tag{5}$$

- Minimum Average SE:

$$SSE4 = \frac{1}{N_p} SSE1 \tag{6}$$

A brief comparison between the results obtained from each objective function is shown in the accompanying Microsoft Excel workbook. The cells in this workbook are color-coded to indicate the objective function in each case (**bold red texts**) and the fitting parameters (yellow cells).

Example 1. Fitting Isothermal Data with the Isothermal Langmuir Model:

The Langmuir model is given by the following formula for correlating the amount adsorbed with the equilibrium pressure:

$$n = m\theta = \frac{mbP}{1 + bP} \tag{7}$$

The spreadsheet Example1 shows the application of this method with the four forms of objective functions (Labeled Langmuir1 through Langmuir4, respectively) and n as the dependent variable and P as the independent variable. Although the least ARE method results in a lower ARE, the back-predicted curve deviates considerably from the experimental data at high loadings with a considerable increase in the SSE. Minimizing SSE3 gives almost the same result as the minimum ARE, but to a lesser extent, and the results from minimizing SSE4 is almost identical to SSE1.

Example 2. Fitting Isothermal Data with the Virial Adsorption Isotherm Model:

The virial-type adsorption isotherm is given by the following formula for correlating the equilibrium pressure with the equilibrium amount adsorbed:

$$P = n \exp\left(A + Bn + Cn^2 + \cdots\right) \tag{8}$$

The spreadsheet Example2 shows the use of the four forms of the objective functions (Labeled Virial1 through Virial4, respectively) with ln(P) as the dependent variable and n as the independent variable. All forms give excellent correlations and differences between different methods are minimal.