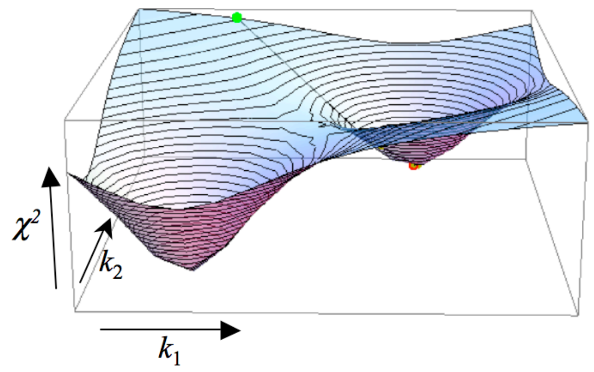
Considerations during numerical-methods-based Curve Fitting

1. Local versus Global Minima

Nonlinear least-squares fitting is an iterative process that attempts to find a minimum on a complicated multidimensional surface. There always is the chance of getting trapped in a local minimum. For this reason, try different initial guesses to attempt to locate either other local minima or of course preferably the global minimum.



2. Parameter Correlation

A problem frequently encountered in curve fitting is the occurrence of strongly correlated parameters. Two or more parameters may be adjusted agonistically or antagonistically with no loss in fitting quality. In the case of a model that depends linearly on its parameters, the correlation can be systematically described by a correlation matrix that derives from the variance-covariance matrix (multi-parameter analog of the variance σ^2). Each coefficient of the correlation matrix indicates the correlation between two fit coefficients as a number between *-1* and *1*. The correlation between two coefficients is perfect if the corresponding element is *1*, it is a perfect inverse correlation if the element is *-1*, and there is no correlation if it is *0*. Fits in which an element of the correlation matrix is very close to *1* or *-1* signals "identifiability" problems: the absolute value of strongly correlated parameters cannot be determined. An example of this is the correlation between k_2 and k_{-2} in figure 1.

	k_1	k_2	k_{-2}	k_3	k_4	k_{-4}
k_1	1.00	-0.76	-0.74	0.11	-0.87	-0.63
k_2	-0.76	1.00	0.99	-0.41	0.58	0.41
k_{-2}	-0.74	0.99	1.00	-0.37	0.55	0.41
k_3	0.11	-0.41	-0.37	1.00	-0.20	-0.15
k_4	-0.87	0.58	0.55	-0.20	1.00	0.82
k_{-4}	-0.63	0.41	0.41	-0.15	0.82	1.00

Figure 1. Correlation Matrix of a six-parameter system

Correlation matrices can be good indications of correlated variables but are only accurate for linear systems. If the fitting function is nonlinear in the fit coefficients, then the error estimates are an approximation. The quality of the approximation will therefore depend on the nature of the nonlinearity at the particular location in parameter space. One may manually check for parameter correlations by supplying different initial guesses. This should naturally be assisted by one's chemical intuition since correlations are rooted in the system under investigation.

A more systematic approach to determining correlation coefficients for nonlinear cases is to compute the χ^2-surface by sampling parameter space. Plotting χ^2 with respect to two other parameters may give a plot like the one shown in figure 2.

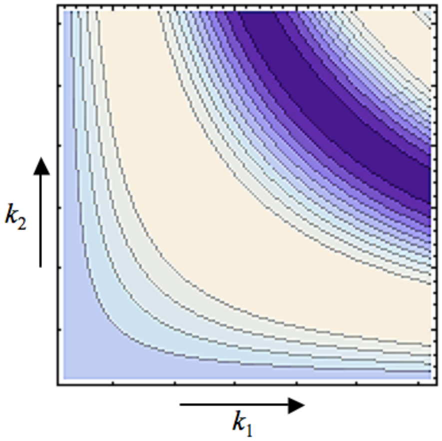


Figure 2. Contour Plot of χ^2-surface.

The example shows an inverse correlation of the two parameters. Decreasing k_1 while increasing k_2 yields an equally low value in χ^2 and therefore an equally good fit as indicated by the deep blue color.

The next three sections are devoted to technical challenges of numerical methods during curve fitting.

3. The fudge factor 'epsilon' in Levenberg-Marquardt fitting

Epsilon (also referred to as lambda λ in other sources) is a fudge factor used in the *Levenberg-Marquardt* algorithm of nonlinear least-squares fitting. Epsilon values are used to calculate partial derivatives with respect to each fit coefficient. The partial derivatives in turn are used to determine the search direction for the coefficients that give the smallest chi-square. If the function depends only weakly on a coefficient the derivatives may appear to be zero. The solution to that problem is to set an epsilon value large enough to give a nonzero difference in the function output. Alternatively, if the function depends strongly on a coefficient epsilon should be set to a lower value. A good choice for epsilon will greatly improve the quality and efficiency of the fit. Choosing epsilon values is an empirical process and can be thought of as a separate adjustable parameter for curve fitting. As a rule of thumb, adjust epsilon in powers of ten (1, 0.1, 0.01 etc.).

4. Numerical Integration Methods

One of the primary hurdles of numerical integration is the presence of stiff differential equations. This occurs frequently in chemical kinetics where concentrations abruptly change curvature. An example is shown in figure 3.

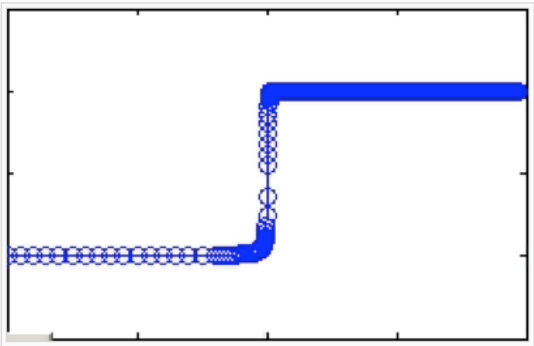


Figure 3a. non-Stiff Solver

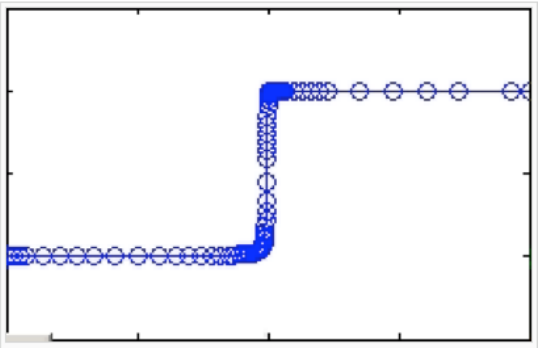


Figure 3b. Stiff Solver

A non-Stiff solver spends much effort in integrating the later parts of the curve because it needed to dramatically reduce its step-size at the point of large curvature. The stiff solver is able to adapt to the changes in curvature better and leads to increased performance. *Bulirsch-Stoer* and *Runge-Kutta* methods are adequate solvers for smooth (nonstiff) problems with the latter being more commonly used. Trial runs indicate that *Adams-Moulton* may be fastest for nonstiff problems. *BDF* is definitely the preferred algorithm for stiff problems. *Runge-Kutta* is a robust method that may work on problems that fail with other methods.

5. Computer Precision during Numerical Integration

Computers have a limit in the number of digits they can handle. One may reach this limit if values differ by many orders of magnitude. To avoid such computer precision problems, it is suggested to scale parameters and concentrations to a value of around unity. This involves modifying the model by introducing pseudo-rate constants and pseudo-concentrations. An example is shown below:

$d[\mathrm{A}]/dt = k_1[\mathrm{A}][\mathrm{B}] - k_{-1}\,[\mathrm{C}]$
$[\mathrm{A}]_0 = 0.1$ (decreases to 0.0 over time)
$[\mathrm{B}]_0 = 2.5\ 10^{-5}$ (remains low over time)
$[\mathrm{C}]_0 = 0.0$ (increases to 0.1 over time)

$k_1 = 0.25$
$k_{-1} = 1.51\cdot 10^{-7}$

Rewrite the differential equation above as:

$d[\mathrm{A}]/dt = k_1[\mathrm{A}][\mathrm{B'}] - k_{-1'}[\mathrm{C}]$

where $k_{-1} = 10^{-7}\cdot k_{-1'} = 1.51\cdot 10^{-7}$
$[\mathrm{B}]_0 = 10^{-5}\cdot [\mathrm{B'}]_0 = 2.5\ 10^{-5}$

k_{-1} and $[\mathrm{B}]_0$ were scaled to values closer to unity. The numerical integration will now proceed more smoothly and more accurately.