



*Tutorials and worked examples for simulation,
curve fitting, statistical analysis, and plotting.*

<https://simfit.uk>

<https://simfit.org.uk>

<https://simfit.silverfrost.com>

To simulate theoretical models or to fit deterministic models to experimental data it is frequently necessary to solve an ordinary nonlinear differential equation, or a system of ordinary nonlinear differential equations. This usually cannot be done by formal integration but is often possible using numerical methods under the following conditions.

1. The independent variable, say x , is nonnegative so that the system is parameterized for $x \geq 0$.
2. The model can be expressed naturally as a system of n nonlinear equations. Alternatively, there is one equation of order n where the highest order derivative can be expressed as a function of the independent variable, the dependent variable, and all the other derivatives, so that this can be transformed into a system of n equations.
3. The initial conditions at $x = 0$ are known, i.e., it is an initial value problem (IVP).
4. The range of integration and the number of intermediate independent variable points are defined.
5. An appropriate method is selected.

Definitions

The system of n equations must be expressed in the following form

$$y'_1 = f_1(x, y_1, y_2, \dots, y_n)$$

$$y'_2 = f_2(x, y_1, y_2, \dots, y_n)$$

...

$$y'_n = f_n(x, y_1, y_2, \dots, y_n)$$

where $y'_i = dy_i/dx$ for components y_1, y_2, \dots, y_n .

For example, consider the Lotka-Volterra predator-prey differential equations where y_1 are numbers of prey, say rabbits, while y_2 are numbers of predators, say foxes, so that

$$\frac{dy_1}{dx} = p_1 y_1 - p_2 y_1 y_2$$

$$\frac{dy_2}{dx} = -p_3 y_2 + p_4 y_1 y_2.$$

Here the model is expressed, using four fixed or variable parameters p_1, p_2, p_3 , and p_4 , as a system of two autonomous differential equations, i.e., x does not appear on the right hand sides of the defining equations.

However, the Van der Pol oscillator is the second order differential equation

$$\frac{d^2 z}{dx^2} + \mu(z^2 - 1) \frac{dz}{dx} + z = 0$$

which can be expressed as the following autonomous set of two equations using $y_1 = z$, $y_2 = dy_1/dx = dz/dx$, and $dy_2/dx = d^2z/dx^2$

$$\begin{aligned}\frac{dy_1}{dx} &= y_2 \\ \frac{dy_2}{dx} &= -y_1 - \mu(y_1^2 - 1)y_2.\end{aligned}$$

Of course the differential equations must be supplied along with the range of independent variable, initial conditions, and tolerance settings, but also the numerical method must be chosen subject to these considerations.

- For non-stiff problems where all components have similar relaxation times, the standard methods due to Runge-Kutta, or Adams etc. should be used.
- For stiff problems, where components have widely differing relaxation times, i.e., the condition number of the Jacobian matrix is large, it will be necessary to use the backward differentiation method (BDF) due to Gear. In this case the Jacobian $J_{ij} = \partial f_i / \partial y_j$ is used and there are two possibilities.
 1. The Jacobian matrix can be supplied along with the differential equations. This is the best technique to use, but has the great disadvantage that if the Jacobian matrix is not coded accurately the convergence may be slow or lead to an erroneous solution.
 2. To avoid this happening the Jacobian can be approximated at run time by finite differences. This is not so good as supplying an accurate explicit Jacobian, but will avoid the problem of having to supply an accurate explicit Jacobian.
- The recommended procedure is to experiment with the methods and tolerances interactively until a satisfactory outcome has been achieved.
- Usually a system of n equations will involve $k \geq 0$ parameters p_i that are to be fixed at known values or estimated by curve fitting a given data set. The rule used by SIMFIT is that there must be $k + n$ parameters, as the last n parameters i.e., $p_{k+1}, p_{k+2}, \dots, p_{k+n}$ are the initial conditions.

Simulation

A single differential equation can be simulated using SIMFIT program **makdat**, which is dedicated to generating data intended to be used by program **adderr** in order to add random error to mimic experimental data.

Systems of differential equations can be simulated by program **deqsol** which offers several options as follows.

- Use equations from a supplied library or user-defined model file.
- Change the method and tolerance factor in order to identify optimum conditions.
- Plot or tabulate selected components.
- Store orbits to be collected together for plotting.
- Plot phase diagrams for autonomous systems.
- Change the equations, parameter values, initial conditions, or range of integration interactively.

Simulation would normally be used to explore the outcome of changing parameters or initial conditions, especially to estimate starting parameters for subsequent curve-fitting.

Curve fitting differential equations

The SIMFIT program **qnfit** can be used to fit a single differential equation to data, and it is preferred for this purpose as it offers very extensive options for goodness of fit, parameter reliability, and model discrimination.

However, program **deqsol** has additional options making it more suitable for fitting systems of equations. This program provides options for overlaying multiple trajectories on data points before and after fitting in order to visualize the progress of optimization to a solution point.

It should be emphasized that curve fitting differential equations to experimental data is an extremely difficult subject requiring considerable expertise, and the following facts should be considered.

- The data should be supplied using the library file approach.
- The data points must be supplied in nondecreasing order with respect to the independent variable. This is to facilitate the ability to integrate methodically between fixed levels of the independent variable, but also to ensure that only the first integral for the model is calculated for groups of replicates.
- The system should be simulated before fitting in order to ensure that starting estimates and parameter limits are reasonable.
- There are two options available concerning initial conditions, i.e., the last n parameters.
 1. The initial conditions can be set at known fixed values.
 2. Some or all of the initial conditions can be estimated.

Wherever possible the first of these methods should be used because many serious problems can be encountered when attempting to estimate initial conditions along with other model parameters.

To summarize: fitting differential equations to data will only be satisfactory when these conditions are met.

1. The model proposed can simulate the data and makes sense theoretically.
2. The data are accurate with a high signal to noise ratio.
3. The range of independent variable is sufficiently extensive to expose the influence of all the parameters to be estimated.
4. The starting estimates are close to the parameters expected at the eventual solution point, and parameter limits prevent excessive movement of parameters away from those expected at the solution point.

Note that programs **qnfit** and **deqsol** are designed to terminate with fixed values for system components as a warning if these conditions are not met.