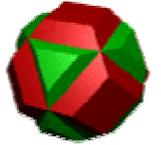




Statistical Model Fitting



Alan Welsh, Ralph Mansson, Lefteris Danos, Jeremy Frey, Jon Essex
Combechem Project, University of Southampton

Nonlinear Models

There are situations where we have a model that is considered appropriate for our data set, e.g. in interfacial second harmonic generation (SHG) experiments which investigate the structure and dynamics of interfaces, even in the presence of overlying bulk phases.

Nonlinear models require a numerical procedure to estimate the model parameters, e.g. a Gauss-Newton algorithm minimises the weighted sum of squares based on sensible starting values. The method also provides estimates of errors for the parameters.

Standard diagnostics methods and profile plots can be used to assess the quality of fit for a nonlinear model.

Linear Models

Regression analysis uses the relationship between a response variable and one or more explanatory variables to make predictions. When there is no prior model assumed for the data we consider a range of different models to identify the important variables – we let the data show the important subsets of the variables.

There is the important issue of model selection which can be approached by a combination of automatic selection procedures and information based criterion to reduce the dimensionality of the problem.

In some problems linear terms are not sufficient and greater flexibility can be introduced by using spline functions to describe the relationship to the explanatory variables.

Statistical Analysis using the R system

R is a language and environment for statistical computing and graphics. R provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, ...) and graphical techniques, and is highly extensible. R is designed around a true computer language, and it allows users to add additional functionality by defining new functions as required.

One of R's strengths is the ease with which well-designed publication-quality plots can be produced, including mathematical symbols and formulae where needed.

This software has been successfully used for linear and nonlinear modelling on projects as part of Combechem.

