Tools for simulating large biological systems - the need for optimization

Pedro Mendes

Virginia Bioinformatics Institute, Virginia Polytechnic and State University, 1750 Kraft Drive, suite 1100, Blacksburg, VA 24061, USA

mendes@vt.edu

Key Word: Optimization, Fitting, Simulation, Genomics

A much positive outcome of the emerging "omic" disciplines in biology (genomics, proteomics, etc.) is the trend away from extreme reductionism to systems descriptions and analyses. Interpretation of whole-genome level experiments requires building comprehensive models of biological systems where usually a large set of biochemical reactions is represented explicitly at some level of detail. Such models are challenging because, by necessity, they contain very large sets of parameters, which are difficult to estimate. To be successful it is important to design experiments to provide sufficient information to estimate parameters. Once the right data is available then it is necessary to process them with algorithms that are capable of handling models with large numbers of parameters.

Model fitting to experimental data is done through optimization algorithms, most times using sum of squares functions. These least-squares problems have specific properties that are used by some algorithms to accelerate conversion. These, such as the popular Levenberg-Marquardt method, can only find local minima, and therefore require a starting point in the neighborhood of the solution as input. This is not a problem if the error function contains only one minimum and is not too flat. In high-dimensional problems, such as large biological systems, the probability of occurrence of several minima and large flat hyper-surfaces in the error function increases. In those cases, local-optimization methods will not be sufficient for parameter estimation. There are a series of global optimization methods available that can be used for such purpose. The application and performance of these methods in estimating parameters in large biological systems will be discussed. Illustrations will be provided with examples produced with the program Gepasi that can carry out simulation and parameter estimation.