Parameter estimation of ODE's using support vector regression and qualititative constraints.

Nicolas Brunel, 23/11/2007.

Dynamical systems used for the modeling of biological networks (such as gene regulatory networks or metabolic networks) are generally based on Ordinary Differential Equations (ODE's). Although nonlinear ODE's (of the form $\dot{x}(t) = f(t, x(t), \theta)$) are commonly used in Systems Biology, their estimation from real data remains a diffcult task because of the high number of parameters to estimate compared with the relatively low number of observations. Moreover, statistical inference gives rise to complex optimization problems due to the implicit link between observations and the parameters θ (in particular due to the nonlinearity of f).

When concentration profiles (time series) are discretely observed, it was shown that a preliminary nonparametric estimation $\hat{x}$ of the solution of the ODE enables to define a rather simple criterion to minimize for the determination of f, contrary to more classical methods such as Least Mean Squares. We show the versatility of this method, by using spline regression and Support Vector Regression (SVR) for the computation of $\hat{x}$. We are also interested in showing that the method based on SVR can be easily modified in order to introduce some qualitative knowledge on the behavior of the solution of the ODE. Indeed, a slight modification of SVR, the so-called semiparametric SVR proposed by Smola et al., enables to impose a pre-defined shape to the estimated solution. This information about the shape can help to convert a qualitative information into a useful mathematical knowledge for statistical learning which helps for a better exploration of the parameter space, especially when few data are available. For instance, a key property of dynamical systems that we want to control can be the oscillating behavior of solutions, which is a relevant feature for numerous biological systems.