Identification of genetic networks by a quadratic systems approach

Amato F.(1), Bansal M.(2), Cosentino C.(1), Curatola W.(1), Di Bernardo D.(2)

(1) Department of Experimental and Clinical Medicine, Università degli Studi di Catanzaro Magna Graecia, Catanzaro (2) Telethon Institute of Genetics and Medicine, Naples

Motivation

The aim of the present work is to identify a connection network describing the expression profiles of the genes involved in the cellular cycle of the yeast. In order to derive such connections from the microarray data, a mathematical model is needed, whose states represent the expression levels of the genes. However, since the cellular dynamics exhibit a limit cycle behavior [1] (i.e. roughly speaking, a periodic response, which is robust to small perturbations), it is not possible to describe the whole phenomenon by means of only linear models. Therefore, we investigate the suitability of quadratic systems [2], for the description of the genetic network dynamics. A quadratic system is made up of a set of first order differential equations, with linear terms plus quadratic terms (i.e. multiplications between two state variables or between a state and an input variable).

Methods

A common problem concerning the use of nonlinear models consists of the lack of fast and reliable identification techniques [3]. In the case of quadratic systems this is not an issue because we can exploit the same regression methods developed in the linear context. In order to verify the quality of these identification techniques in the quadratic case we have generated several random systems, having up to ten states, and a loose connection pattern (scale-free networks). The identification technique is based on successive refinement steps: in the first step a simple linear model is used to identify the genes that influence the expression of each gene. Then, the quadratic terms are introduced and a new identification is performed to identify the coefficients of these terms. If the identified coefficient is below a certain threshold, the related term is discarded in the next identification step, and so on.

Results

The technique presented here has been already tested in the linear framework on real data and yielded promising results [4]. As for the quadratic case, a huge percentage of the nonlinear terms are correctly identified, given a proper choice of the coefficients threshold. The foregoing works are focused on the identification from noisy data, and the combination of the technique with smoothing and clustering algorithms, in order to render it suitable for application on real-world data (in particular for the yeast) where high order systems must be considered.

Supplementary Information: References [1] H. K. Khalil, Nonlinear Systems, Prentice Hall, 2001. [2] M. Frayman, On the relationship between bilinear and quadratic systems, IEEE Transactions on Automatic Control, 567—568, 1975. [3] O. Nelles, Nonlinear System Identification: From Classical Approaches to Neural Networks and Fuzzy Models, Springer-Verlag, 2000. [4] T. S.

Gardner, D. Di Bernardo, D. Lorenz, J. J. Collins, Inferring genetic networks and identifying compound mode of action via expression profiling, Science, 102—104, 2003.