

Gaussian Process Modelling of Transcription Factor Networks using Markov Chain Monte Carlo

Michalis K. Titsias, Neil Lawrence and Magnus Rattray
School of Computer Science, University of Manchester
Manchester M13 9PL, UK

Ordinary differential equations (ODEs) can provide an useful framework for modelling the dynamics of biological networks. In this study, we focus on a small biological sub-system where a set of target genes are regulated by one transcription factor protein. The concentration of the protein and the gene specific kinetic parameters such as basal rates, decay rates and sensitivities are typically unknown. The objective of modelling is to estimate these quantities by making use of a set of observed gene expression levels.

We consider a Bayesian framework for modelling the system of ODEs that is based on Gaussian processes. The Gaussian process is used as the prior for the transcription factor protein and allows us to infer the concentration of the protein in a time continuous manner. We present a Markov chain Monte Carlo algorithm for a full Bayesian statistical inference. The essential property of our MCMC algorithm is that we efficiently infer the protein concentration by applying a novel sampling algorithm for Gaussian process models. We apply our technique to linear and non-linear models. Some preliminary results for the data used by Barenco *et al.* are given in the figure below.

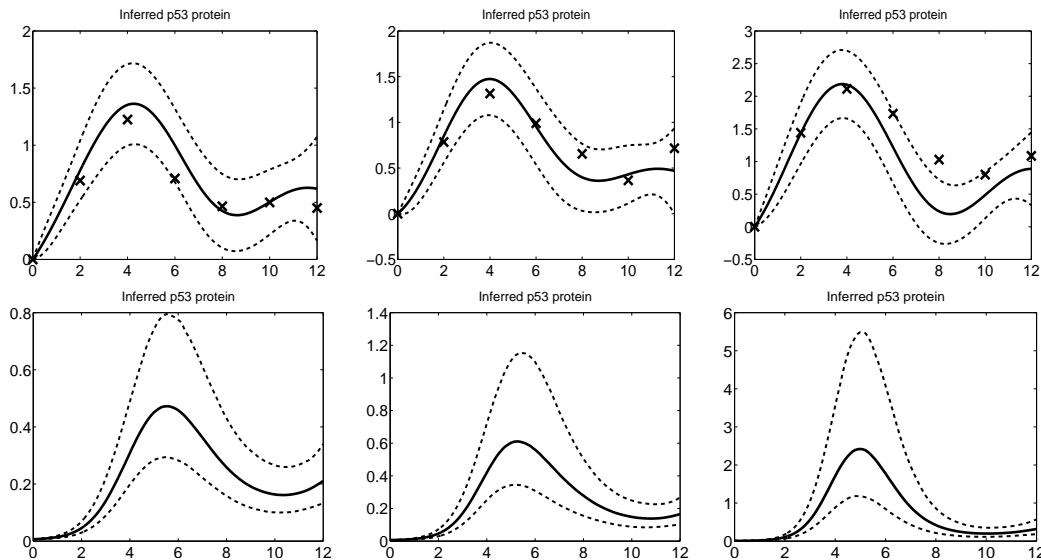


Figure 1: Inference of p53 activity using 5 target genes and three replicas; see Barenco *et al.*. The first row shows the inferred protein concentration for the three replicas using a linear model. The prediction of Barenco *et al.* is also shown as crosses. The second row shows the corresponding protein activities when a positivity (exponential) constraint is used together with the Michaelis-Menten kinetic equation.

References

M. Barenco, D. Tomescu, D. Brewer, R. Gallard, J. Stark, and M. Hubank. *Genome Biology*, 7(3):R25, 2006