

Parameter estimation in regulatory networks of biochemical reactions

Dirk Fey^{1,2,*}, Eric Bullinger^{1,2,*}

1. Hamilton Institute, National University Ireland, Maynooth, Ireland
 2. Industrial Control Centre, University of Strathclyde, Glasgow, Scotland, UK
- *E-mail: dirk.fey@nuim.ie; eric.bullinger@nuim.ie

Introduction

Dynamic models present a fundamental tool in systems biology, but rely on kinetic parameters, e.g. association and dissociation constants. Results arising from the model analysis depend crucially on these parameters.

The estimation of parameters presents a bottleneck in systems biology. Experimental studies on isolated reactions are usually expensive, time-consuming or even infeasible for large networks. As a consequence, many parameters must be inferred indirectly from measurements on a limited number of species concentrations and/or reaction rates. The rapid increase of availability and quality of biological data makes this a more and more feasible task [1–4].

Here we propose a two-step method that decouples the reconstruction of non-measured species and reactions from the estimation of parameters.

Method: Connecting identifiability and observability

Biological processes are often modeled with ordinary differential equations (ode's)

$$\frac{d}{dt}c = N \cdot r \quad (1)$$

where c is the vector of species concentrations, N the stoichiometric matrix and r the vector of reaction rates. In principle, the reaction rates may be any nonlinear function. Typical kinetics are mass action for signal transduction, Michaelis Menten for metabolic pathways and Hill for gene regulation. We consider all these kinetics, and products thereof:

$$r_i = r_{i,\text{nominal}} \cdot \frac{c_1^{\nu_{i,1}}}{K_{i,1}^{\eta_{i,1}} + c_1^{\eta_{i,1}}} \cdot \dots \cdot \frac{c_n^{\nu_{i,n}}}{K_{i,n}^{\eta_{i,n}} + c_n^{\eta_{i,n}}} \quad (2)$$

This allows to describe activation, inhibition and cooperativity in any combination.

We use the special form of the kinetics (2) to rewrite the system (through mathematical manipulation) such that the ode's do not depend on the parameters r_{nominal} and K anymore. Thereby, we introduce a new ode for each reaction rate and each denominator term in (2), the Hill variables $m_{i,j} = K_{i,j}^{\eta_{i,j}} + c_j^{\eta_{i,j}}$,

$$\frac{d}{dt}m_{i,j} = \eta_{i,j} \cdot c_j^{\eta_{i,j}-1} \cdot \dot{c}_j, \quad \frac{d}{dt}r_i = r_i \cdot \sum_{j=1}^n \left(\nu_{i,j} \cdot \frac{\dot{c}_j}{c_j} - \frac{\dot{m}_{i,j}}{m_{i,j}} \right) \quad (3)$$

In the extended system (1), (3) the parameters are hidden in the initial conditions.

The advantage of this model extension is the decoupling of estimation of states (species concentrations, reaction rates, ...) and parameters (r_{nominal} and K). This allows the following two-step approach:

1. Reconstruct all extended states using so-called observers, a well developed technique from systems theory [5];
2. Calculate the parameters based on the extended states by explicitly solving the sum of least squares criterion on the estimation error.

Interestingly, our approach establishes a systems theoretical link between parameter estimation and state reconstruction: The parameters r_{nominal} and K are identifiable if, and only if, the extended system (1), (3) is observable.

Proof of concept: Circadian rhythm in neurospora

The proposed method is applied to a simple model of the circadian rhythm in neurospora [6], composed of three species, three mass action and three Hill kinetic reactions. Thus, the extended model consists of twelve states: three species concentrations, six reaction rates and three Hill variables.

Now let us assume we measure the species concentrations and the Hill reaction rates. After initializing the non-measured states with 100% deviations, an observer based on [7] reconstructed the states successfully (Figure 1). The estimates of the parameters deviate only slightly from the true values (Table 1).

Concluding comments

A crucial task in quantitative modeling of systems biology is the choice of kinetic parameters. The presented approach presents a two-step strategy for parameter estimation that decouples the estimation of states and parameters. Importantly, it allows to identify suitable measurements and/or model structures. For the actual estimation using real world biological data, further research is necessary to incorporate discrete and noisy data.

References

- [1] Casey, F. P., Baird, D., Feng, Q., Gutenkunst, R. N., Waterfall, J. J., Myers, C. R., Brown, K. S., Cerione, R. A., and Sethna, J. P. *IET Sys. Biol.* **1**(3), 190–202 (2007).
- [2] Peifer, M. and Timmer, J. *IET Sys. Biol.* **1**(2), 78–88 (2007).
- [3] van Riel, N. and Sontag, E. *IEE Proc. Sys. Biol.* **153**(4), 263–274 (2006).
- [4] Anguelova, M., Cedersund, G., Johansson, M., Franzén, C., and Wennberg, B. *IET Sys. Biol.* **1**(4), 230–237 (2007).
- [5] Birk, J. and Zeitz, M. *Int. J. Contr.* **47**, 1823–1836 (1988).
- [6] Leloup, J. C., Gonze, D., and Goldbeter, A. *J Biol Rhythms* **14**(6), 433–448 (1999).

[7] Gauthier, J., Hammouri, H., and Othman, S. *IEEE T. Automat. Contr.* **37**(6), 875–880 (1992).

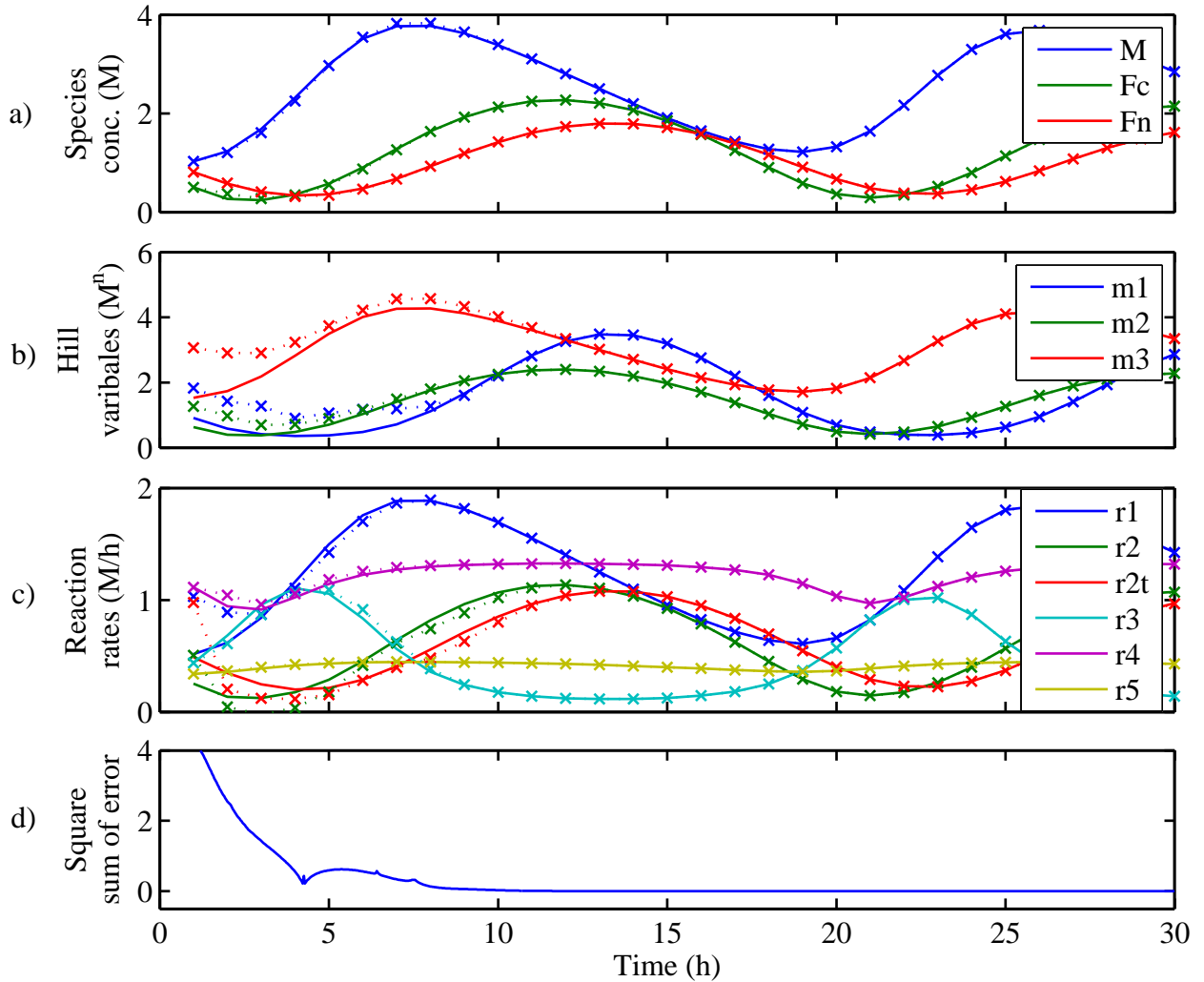


Figure 1: The true (solid lines) and estimated (x) values. The observer was initialized with all non measured values 100% above their true values. Solid lines: true values; dotted lines with x: observed values. Plot d) shows the sum of the error squares ($\sum_j (x_{j,est} - x_{j,true})^2$).

	Nominal rates						Hill constants		
	ks	k1	k2	vs	vd	vm	K1	Kd	Km
True	0.5	0.5	0.6	0.4	1.4	0.505	0.5	0.13	0.5
Estimated	0.5000	0.4986	0.5982	0.3987	1.3999	0.5050	0.5075	0.1287	0.4977

Table 1: True and estimated parameter values after the observer has converged ($t > 10$ h).