Extracting the Measurement Information Core for Model Selection of Dynamical Systems

Alberto G. Busetto[1,2], Cheng Soon Ong[1] and Joachim M. Buhmann[1,2]

[1]Departement Informatik, ETH Zurich, Zurich, Switzerland
[2]Competence Center for Systems Physiology and Metabolic Diseases, Zurich, Switzerland
{busettoa,chengsoon.ong,jbuhmann}@inf.ethz.ch

Time-dependent processes are often mathematically modeled as systems of differential or difference equations. When the knowledge about the process is partial, conflicting hypotheses are represented as alternative models, which are selected according to observational evidence [1]. Designing maximally informative experiments is desirable, but it is also challenging for nonlinear and partially observable systems [2]. Whereas the standard approaches are based on the linear-Gaussian assumption, in this work we propose a strategy based on the extraction of a measurement information core, for the purpose of active model selection. The core consists of a subset of state variables of the dynamical system Σ, whose property of high informativeness is invariant under a range of initial conditions and parameters. The information core is identified according to the expected "surprise" experienced by the observer from the result of the designed measurement [3].

In a n-dimensional state space, let the process model be described by the possibly stochastic differential equation

$$dx(t) = f(x(t), \theta, t)dt + \sigma_w(x(t))dW(t),$$

where $W(t)$ is a Wiener process, whose infinitesimal variance is σ_w. In this formulation, the deterministic component f and its parameters θ are unknown. A set of alternative models $\{\mathcal{M}_i\}_{i=1}^q$ is hypothesized, each one associated with its respective deterministic function f_i, whose parameters are uncertain. The time-discrete measurement process can be formalized as $y(t_i) = Hx(t_i) + v_i$, where y is an m-dimensional measurement vector, $H \in \{0,1\}^{m\times n}$ is the Boolean measurement matrix and v_i are random variables that describe the measurement noise.

The goal of optimal experimental design is the specification of H which maximizes the information gain, subject to an upper bound on the dimension of the observation vector and to additional feasibility constraints. The information gain is quantified by the expected dissimilarity between the prior and the posterior, that is by the mutual information between the models and the data

$$\mathrm{I}(\mathcal{M}, \mathcal{D}(H)) = \mathbb{E}_{\mathcal{D}}\left[\mathrm{KL}(p(\mathcal{M}|\mathcal{D}(H))||p(\mathcal{M}))\right],$$

where $\mathcal{D}(H)$ denotes the dataset measured according to the selected H. The initial conditions are unknown, the parameters are uncertain and, moreover, their marginalization is computationally infeasible. Therefore, the identification of the maximally informative subset of variables must be robust against perturbations of the system configuration $C = (x(t_0), \theta)$.

The maximum amount of information is obtained when the complete state space is measurable, that is when H is a permutation matrix. This defines an upper bound on the experimentally available information, given by $\mathrm{I}^{\max} = \mathrm{I}(\mathcal{M}, \mathcal{D}(\mathrm{I}_n))$. For a given $\alpha \in [0, 1]$, we denote the solution of the optimal experimental design as

$$H_\alpha^*(C) = \arg \min_{H \in \mathcal{S}_\alpha(C)} \sum_{i,j} H_{i,j},$$

where

$$\mathcal{S}_\alpha(C) = \{H | \mathrm{I}_C(\mathcal{M}|\mathcal{D}(H)) \geq \alpha \mathrm{I}_C^{\max}\}.$$

The function $\mathrm{m}(H)$ defines the subset of state variables x_i that are measured when the experiment specified by H is performed. Finally, the measurement information core of Σ is obtained as

$$\mathrm{core}(\Sigma) = \bigcap_{C \in \mathcal{C}} \mathrm{m}(H_\alpha^*(C)),$$

for a set of system configurations $\mathcal{C}$.

This research builds on the observation that, despite nonlinearity and uncertain system condition, very few variables are necessary to discriminate between conflicting hypotheses. In the case of the 19 proposed models and parameters of the biochemical TOR pathway [4], we computed that approximately half of the experimentally available information comes from a subset of variables whose cardinality is one order of magnitude smaller than the dimension of the state space.

The extraction of the measurement information core facilitates the identification of key mechanisms in the studied system, which are the ones playing a central role in the model selection. We expect our strategy to highlight specific dynamical behaviors, providing a persistent source of information for the design of critical experiments.

References

[1] Vyshemirsky, V. and Girolami, M.A., "Bayesian ranking of biochemical system models", *Bioinformatics*, 24(6), 833–839 (2007).

[2] Banga, J.R. and Balsa-Canto, E., "Parameter estimation and optimal experimental design", *Essays in Biochemistry*, 45, 195–109 (2008).

[3] Baldi, P. and Itti, L., "Attention: bits versus wows", *Proceedings of the IEEE International Conference on Neural Networks and Brain*, Beijing, China, PL56–PL61 (2005).

[4] Kuepfer, L., Peter, M., Sauer, U. and Stelling, J., "Ensemble modeling for analysis of cell signaling dynamics", *Nature Biotechnology*, 25, 1001–1006 (2007).

Topic: Learning Algorithms.

Preference: Oral.