Neuroscience Research 104 (2016) 52–55

Contents lists available at ScienceDirect

Neuroscience Research

journal homepage: www.elsevier.com/locate/neures

Review article

Slow dynamics perspectives on the Embodied-Brain Systems Science

Shiro Yano [a,*], Takaki Maeda [b], Toshiyuki Kondo [a]

[a] Department of Computer and Information Sciences, Tokyo University of Agriculture and Technology, 2-24-16 Nakachou, Koganei-shi, Tokyo 184-8588, Japan

[b] Department of Neuropsychiatry, Keio University School of Medicine, 35 Shinanomachi, Shinjuku-ku, Tokyo 160-8582, Japan

ARTICLE INFO

Article history:
Received 2 September 2015
Received in revised form 5 November 2015
Accepted 10 November 2015
Available online 28 November 2015

Keywords:
Dynamical system
Fast-slow systems
Bayesian statistics
Dimension reduction
Statistical learning theory
Embodied-Brain Systems Science

ABSTRACT

Recent researches point out the importance of the fast-slow cognitive process and learning process of self-body. Bayesian perspectives on the cognitive system also attract research attentions. The view of fast-slow dynamical system has long attracted wide range of attentions from physics to the neurobiology. In many research fields, there is a vast well-organized and coherent behavior in the multi degrees-of-freedom. This behavior matches the mathematical fact that fast-slow system is essentially described with a few variables. In this paper, we review the mathematical basis for understanding the fast-slow dynamical systems. Additionally, we review the basis of Bayesian statistics and provide a fast-slow perspective on the Bayesian inference.

Contents

1. Introduction

The fast-slow dynamical system is a set of interacting objects such that at least one object varies much slower than the other objects. This viewpoint has attracted wide range of research attentions including physics, chemistry, sociology and neurobiology (Haken, 2004; Scheffer et al., 2012). One of the reasons for this is the mathematical fact that such system is essentially described only by the slow variables. Fast variables are enslaved to these slow variables. A number of experiences that the multi degrees-of-freedom (DOF) systems often show the DOF reduction encouraged the researchers to model them with a fast-slow system.

In the cognitive sciences, Bayesian perspectives on the cognitive systems attract lots of research attentions (Griffiths et al., 2008). Additionally, researchers point out the existence and importance of the fast-slow cognitive process about the self-body (Hagura and Haggard, 2015).

In this paper, we review the mathematical basis and related results of the fast-slow dynamical systems. Moreover we review the basis of Bayesian statistics and propose a fast-slow perspective on the Bayesian statistics. These review and perspectives would be helpful for advancing the Bayesian perspectives in cognitive science.

* Corresponding author. Tel.: +81 09047846062.
E-mail addresses: syano@cc.tuat.ac.jp (S. Yano), takaki@xa3.so-net.ne.jp (T. Maeda), t_kondo@cc.tuat.ac.jp (T. Kondo).

http://dx.doi.org/10.1016/j.neures.2015.11.002

2. Dynamical systems

2.1. Fast-slow dynamical system

We consider the following ordinary differential equation (ODE)

Definition 1 *(Fast-Slow ODE).*

$$\frac{dx}{dt} = f(x, y, \varepsilon) \tag{1}$$

$$\frac{dy}{dt} = \varepsilon g(x, y, \varepsilon), \tag{2}$$

with $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, $\varepsilon \in \mathbb{R}$, f: $\mathbb{R}^{n+m} \to \mathbb{R}^n$ and g: $\mathbb{R}^{n+m} \to \mathbb{R}^m$. A parameter ε is subjected to $0 \le \varepsilon \ll 1$.

The variable y is called slow variable due to $dy/dt \simeq 0$. Contrastingly the variable x is called fast variable (Jones, 1995).

We assume the functions f, g are the C^∞-differentiable functions and that f is hyperbolic at the equilibrium of Eq. (1). In other words, all of the eigenvalues of the Jacobian $\partial f/\partial x(x^*, y^*, \varepsilon)$ at any points $(x^*, y^*) \in \{(x, y)|f(x, y, \varepsilon) = 0\}$ have non-zero real parts. This assumption is important for applying the implicit function theorem on $f(x, y, \varepsilon) = 0$. As an especially important case $\varepsilon = 0$, we put down the function $x = h_0(y)$ as a solution of $0 = f(x, y, 0)$.

By changing the timescales t to $\tau = \varepsilon t$, Eqs. (1) and (2) become

$$\varepsilon \frac{dx}{d\tau} = f(x, y, \varepsilon) \tag{3}$$

$$\frac{dy}{d\tau} = g(x, y, \varepsilon). \tag{4}$$

The time scale τ is a slower unit of measurement than t. It is for this reason that system (1) and (2) is called the fast system and system (3) and (4) is called the slow system.

2.2. Dimensionality reduction and Synergetics

Roughly speaking, in the limit of $\varepsilon \to 0$, there exists a $\Delta > 0$ such that the trajectory of (3) and (4) starts from $(x_0, y_0) \in \mathbb{R}^{n+m}$, gets closer to the trajectory of Eqs. (5) and (6) during $-\Delta < \tau < \Delta$.

$$x = h_0(y) \tag{5}$$

$$\frac{dy}{d\tau} = g(h_0(y), y, 0) \tag{6}$$

It means that the variable x is enslaved to satisfy $x = h_0(y)$. For a following precise explanation about this reduction, we denote a manifold $\mathcal{M}_0 = \{(x, y)|x = h_0(y), y \in K\}$, where K is a compact domain in $\mathbb{R}^m$.

Before explaining the theorem about the above mentioned dimensionality reduction, we introduce a term:

Definition 2 *(Locally invariant manifold (Chow et al., 2000)).* A submanifold $\mathcal{M} \subset \mathbb{R}^{n+m}$ with boundary $\partial\mathcal{M}$ is called locally invariant under (1) and (2), if, for any point $p \in \mathcal{M}/\partial\mathcal{M}$, there exists a $\Delta > 0$ such that $(x, y)_{t,p} \in \mathcal{M}$ for $t \in (-\Delta, \Delta)$, where $(x, y)_{t,p}$ is the solution of (1) and (2) with $(x, y)_{0,p} = p$.

Following theorem holds under a few appropriate assumptions (Jones, 1995)

Theorem 3 (Fenichel's theorem). *If $\varepsilon > 0$ is sufficiently small, there exists the locally invariant manifold under Eqs. (1) and (2) that $\mathcal{M}_\varepsilon = \{(x, y)|x = h_\varepsilon(y), y \in K\}$. Moreover h_ε is C^r for any $r < +\infty$ jointly in y and ε. $\mathcal{M}_\varepsilon$ is diffeomorphic to $\mathcal{M}_0$.*

Manifold M_ε is called slow manifold. Fenichel's theorem is known as the generalization of Tikhonov–Levinson theory (O'Malley, 2014). Tikhonov–Levinson theory assumes the stability of Eq. (1). Fenichel generalized this theory to be applicable for hyperbolic f at the equilibrium. Further historical review and extensions are reviewed in O'Malley (2014). Thus we get the reduced system Eq. (6).

We rewrite Eq. (6) to Eq. (7) without loss of generality.

$$\frac{dy}{d\tau} = g_0(y, \eta). \tag{7}$$

The vector field is parametrized by η. As noted before, the variable x is enslaved to this dynamics of the slow variable y.

Hermann Haken has investigated the mechanisms of the spontaneous emergence of new quantities and structures in the large degree of freedom system (Haken, 2004). He named this research field Synergetics. The fast-slow system is enslaved to the reduced system (7). Moreover once bifurcation occurs in this reduced system, the whole system spontaneously changes. A bifurcation of a dynamical system is a qualitative change on the system which is caused by parameters such as η (Crawford, 1991). A review of bifurcation theory is outside the scope of this paper. Readers are recommended to refer Crawford (1991) and Kuznetsov (2004). For this reason, the fast-slow phenomena have been one of the research subjects in Synergetics.

3. Bayesian statistics

3.1. Basis of Bayesian inference

At first, we introduce notations in this section. We represent a set of observed n samples as $X^n = (X_1, X_2, \ldots, X_n)$ which are independently taken from the true distribution $q(x)$, $x \in \mathbb{R}^n$. In general, true distribution $q(x)$ is unknown. Bayesian inference is a kind of the statistical inference which aims to construct a model of $q(x)$. It is based on the Bayes' theorem:

$$\varphi(w|X^n) = \frac{\varphi_0(w)\prod_{i=1}^n p(X_i|w)}{\int \varphi_0(w)\prod_{i=1}^n p(X_i|w)dw} \tag{8}$$

which consists of a conditional probability distribution $p(x|w)$, given a parameter $w \in \mathbb{R}^d$, prior distribution $\varphi_0(w)$ and samples X^n. The Bayes' theorem (8) is recursively derived from another form of the Bayes' theorem:

$$\varphi(w|X^n) = \frac{p(X_n|w)\varphi(w|X^{n-1})}{\int p(X_n|w)\varphi(w|X^{n-1})dw}, \tag{9}$$

where $\varphi_0(w) = \varphi(w|X^0)$. Bayesian inference is the updating process of prior distribution to the posterior distribution based on the Bayes' theorem. The denominator $Z = \int \varphi_0(w)\prod_{i=1}^n p(X_i|w)dw$ is called the marginal likelihood and the negative logarithm $-\ln Z$ is called the Bayes free energy (Watanabe, 2001a).

We are interested in the asymptotic agreement between true distribution $q(x)$ and predictive distribution $p(x|X^n)$ in the limit of $n \to \infty$, where predictive distribution is defined as

$$p(x|X^n) = \int p(x|w)\varphi(w|X^n)dw. \tag{10}$$

Watanabe (Watanabe, 2001b) showed that the generalization error $G(n) = E_{X^n}[d(q(\cdot), p(\cdot|X^n))]$ behaves $G(n) \to 0$ with $n \to \infty$ when $d(q(\cdot), p(\cdot|X^n))$ is Kull-back Leibler divergence (KL divergence):

$$d(q(\cdot), p(\cdot|X^n)) = \int q(x)\ln\frac{q(x)}{p(x|X^n)}dx. \tag{11}$$

KL divergence is a type of divergence function.

Definition 4 *(Divergence function (Gneiting and Raftery, 2007)).*

$$d(P, Q) = S(Q, Q) - S(P, Q) \quad P, Q \in \mathcal{P} \tag{12}$$