Volume 381 15 July 2007 ISSN 0378-4371

ELSEVIER

PHYSICA A
STATISTICAL MECHANICS AND ITS APPLICATIONS

Recognized by the European Physical Society

Editors:
K.A. DAWSON
J.O. INDEKEU
H.E. STANLEY
C. TSALLIS

Complete Volume

Available online at
ScienceDirect
www.sciencedirect.com

http://www.elsevier.com/locate/physa

Available online at www.sciencedirect.com
ScienceDirect
Physica A 381 (2007) 71–81
PHYSICA A
www.elsevier.com/locate/physa

Parameter estimation for random dynamical systems using slice sampling

S.J. Hatjispyros[a,*], Theodoros Nicoleris[a], Stephen G. Walker[b]

[a]University of the Aegean, Greece
[b]University of Kent, UK

Received 14 October 2006; received in revised form 8 March 2007
Available online 18 March 2007

Abstract

We provide details on the full reconstruction of the dynamic equations from measured time series data, given the general class of the underlying physical process. Our results can be used by researchers in physical modelling and statistical mechanics interested in an efficient estimation of low dimensional models, incorporating dynamic as well as observational noise. Our approach is Bayesian, based on an auxiliary variables algorithm that is fast and accurate, and direct, in the sense that only uniform distributions need to be sampled. This method is simpler than other Bayesian approaches where one has to sample from non-standard–unknown distributions using MCMC methods.

Keywords: Slice sampler; Random dynamical systems

1. Introduction

Time varying phenomena in the physical sciences often exhibit irregular and complex behaviour. It has been known, since Poincare's pioneering work, that physical phenomena even in the macroscopic level that is purely deterministic, may present complicated and irregular behaviour. Later in the 1970s, Lorentz observed this type of behaviour in a simple deterministic dynamical model for meteorological fluid mechanics. It was termed chaos and initiated a new area in the mathematical and physical sciences. Such type of behaviour from simple and low dimensional dynamical systems was used in the modelling of a large variety of physical situations. For instance, as a simple model for turbulence [1], to understand the behaviour of plasmas [2] and the origin of kinetic theory in statistical mechanics [3,4], in modelling the behaviour of nonlinear electronic circuits [18], etc.

To compensate for a small number of degrees of freedom, such models can incorporate dynamic noise, perhaps as model error, resulting in what is known as random dynamical systems. In the case of laboratory or "real world" time series data we have often inaccurate measurements of the underlying dynamic process, making the states of the true system unobservable. Therefore, it is possible that observational noise is incorporated in such models.

*Corresponding author.
E-mail address: schatz@aegean.gr (S.J. Hatjispyros).

0378-4371/$ - see front matter
doi:10.1016/j.physa.2007.03.013

Suppose that system identification has led us to believe that the underlying physical process can be modelled as a low dimensional and nonlinear, time discretized, continuous state space parametric dynamical system that incorporates both dynamic and observational noise. To fully reconstruct the model of the dynamical system from measured time series data, provided that the general class of the underlying model is known, we need to estimate the state space parameters of the system comprised of the vector of parameters of the deterministic part and the characteristics of the dynamic noise components.

When dynamic and observational noise components are independent and identically distributed (iid) (typically normally distributed) the stochastic model is Markovian and techniques that make use of the likelihood function, that is the joint density of the observations given all the unknown state space parameters of the system, can be used to fully reconstruct the model. In the Bayesian statistical approach the unknown state space parameters are treated as random variables and any prior knowledge available to the researcher can be incorporated together with the likelihood function. Using Bayes theorem the posterior density, that is the conditional density of the state space parameters given the time series, can be obtained. The last and most crucial step is to sample from this posterior density. Then one recovers the probability that each of the state space parameters lies in a certain interval in the presence of the information given by the observed time series data.

In Refs. [5,6], computer intensive Markov chain Monte Carlo (MCMC) methods are used; specifically the Metropolis within Gibbs algorithm [7,8]. This is typically employed when one has to sample from non-standard distributions. Therefore, to sample from the posterior efficiently one has to employ an appropriate proposal (jumping) distribution. We believe that the auxiliary variables sampler (the slice sampler) is more appropriate for such a situation as the algorithm depends only on the proposed functional form of the underlying random dynamical system. Sampling then from the joint posterior is simpler and amenable to automatic construction of MCMC samplers [9,10] for such models.

We believe that the results of this work may be of interest to the physics community, especially those interested in physical modelling and statistical mechanics. For instance, our results may be used for the improved estimation of low dimensional models for observable data. Such low dimensional models are known by now to describe quite well certain systems e.g. fluid flows, electronic circuits, competing populations or even stochastic macroeconomic laws [11,17]. Our method may be used for accurate parameter estimation of such models and subsequently these may be used for prediction for control of the system. As an example, we may consider the problem of perturbing a chaotic trajectory to a periodic one (or vice versa). Chaos control was first successfully tested on the Henon map which since has become a canonical example of chaotic motion.

The paper is organized as follows. In Section 2 we provide a brief introduction on discrete time random dynamical systems. In Section 3 we describe in detail the auxiliary variables sampling algorithm for a k-dimensional dynamical system. In Section 4 we describe in detail a one-dimensional slice sampling algorithm. Namely we reconstruct the dynamic equation of the random logistic map with observational noise, which is the one-dimensional limit of the Henon map. In Section 5 we deal with a two-dimensional random map. We apply the slice sampling algorithm on the reconstruction of the dynamic equations of the Henon map with dynamic noise.

2. Discrete random dynamical systems

We define the random recurrence relation $T : X \times Y \to X$ by

$$x_i = T(x_{i-1}, \eta_i), \quad i = 1, 2, \ldots, \tag{1}$$

where $x_i = (x_{i1}, \ldots, x_{ik})$ takes values in X, a Borel set of $\mathbb{R}^k$ and $\eta = \{\eta_i\}_{i \geqslant 1}$ is a p-dimensional random process having Y, some Borel set of $\mathbb{R}^p$, as its state space. We assume that for fixed x the function $T(x, y)$ is measurable in y, and for fixed y, $T(x, y)$ is continuous in x (see Ref. [12], and references therein). We could work with discontinuities in T but it is harder to do so and all our examples are based on continuous T so we assume this is true. Successive applications of the random map T generate realizations of the stochastic process x_i which depend solely on the initial value x_0 and on the particular realization of the process η. For a rigorous definition of a random dynamical system see Refs. [13–15]. When the random variables η_i are iid, the x_i process is Markovian with a stationary transition density.

Here we will only consider additive noise, $T(x,y) = f(x) + y$. The general case of multiplicative and additive noise can be treated in exactly the same way from a Bayesian perspective (by taking logarithms for example).

Denoting by $\theta \in \mathbb{R}^m$ any dependence of the deterministic map f on parameters, and setting $p = k$, relation (1) reads

$$x_i = f(x_{i-1}, \theta) + \eta_i, \quad \eta_i \overset{\text{iid}}{\sim} \pi(\cdot), \quad i = 1, 2, \ldots \quad (2)$$

We further assume that the random variables η_{ij} of the random vector $\eta_i = (\eta_{i1}, \ldots, \eta_{ik})$ are all independent of each other. Assuming that noisy measurements of the outputs x_i occur at all times for all $1 \leqslant i \leqslant n$, making the x_i sequence unobservable, we introduce the observable quantities $y_i = (y_{i1}, \ldots, y_{ik})$ such that

$$y_i = x_i + \varepsilon_i, \quad \varepsilon_i \overset{\text{iid}}{\sim} \nu(\cdot), \quad 1 \leqslant i \leqslant n, \quad (3)$$

again imposing the condition that the random variables ε_{ij} of the random vector $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{ik})$ are all independent of each other. In addition, let

$$\eta_{ij} \sim N(\cdot|0, 1/\xi_j), \quad \varepsilon_{ij} \sim N(\cdot|0, 1/\lambda_j), \quad 1 \leqslant j \leqslant k,$$

and denote by f_j the jth component of the vector function f. Then we obtain the following normal–normal model:

$$y_{ij}|x_{ij}, \lambda_j \sim N(\cdot|x_{ij}, 1/\lambda_j), \quad (4)$$

$$x_{ij}|x_{i-1}, \theta, \xi_j \sim N(\cdot|f_j(x_{i-1}, \theta), 1/\xi_j). \quad (5)$$

In what follows we let $y = (y_{ij})$ be the $n \times k$ array of observations, and $x = (x_{ij})$ be the corresponding $n \times k$ array of unobserved outputs generated by the recursion given in relation (2).

3. The slice sampler

Suppose that we are in a situation where we have observed y, and the functional form of the random dynamical system responsible for the output is known. The initial state of the system $x_0 = (x_{01}, \ldots, x_{0k})$, the parameters $\theta = (\theta_1, \ldots, \theta_m)$ and $\xi = (\xi_1, \ldots, \xi_k)$ as well as $\lambda = (\lambda_1, \ldots, \lambda_k)$, are all unknown quantities. We denote by $\vartheta = (x_0, \lambda, \xi, \theta)$ the $3k + m$ tuple of parameters we wish to estimate and by $\pi(\vartheta|y)$ the posterior density of ϑ given the data y. Since

$$\pi(\vartheta|y) = \int_{X^n} \pi(\vartheta, x|y) \, \mathrm{d}x,$$

we can sample ϑ from the joint posterior density $\pi(\vartheta, x|y)$, which is the joint posterior density of ϑ and x given the data y, and is given by

$$\pi(\vartheta, x|y) \propto \prod_{i=1}^{n} \prod_{j=1}^{k} \{N(y_{ij}|x_{ij}, 1/\lambda_j) \cdot N(x_{ij}|f_j(x_{i-1}, \theta), 1/\xi_j)\} \pi(\vartheta). \quad (6)$$

This follows from (4) and (5) and that the conditional density of x given the vector of parameters ϑ is given as

$$\begin{aligned} \pi(x|\vartheta) &= \pi(x_1, \ldots, x_n|x_0, \lambda, \xi, \theta) \\ &= \prod_{i=1}^{n} \pi(x_i|x_0, \ldots, x_{i-1}, \lambda, \xi, \theta) \\ &= \prod_{i=1}^{n} \pi(x_i|x_{i-1}, \xi, \theta). \end{aligned}$$

To estimate ϑ we must first compute the full conditionals of the joint posterior density, $\pi(\vartheta, x|y)$. But direct algorithms, such as the Gibbs sampler, are difficult to implement since it is required to sample from full conditionals looking like

$$\pi(x_i|x_{[-i]},y,\vartheta) \propto \prod_{j=1}^{k} \exp\left\{-\frac{\lambda_j}{2}(y_{ij}-x_{ij})^2\right\}$$
$$\times \prod_{j=1}^{k} \exp\left\{-\frac{\xi_j}{2}[(x_{ij}-f_j(x_{i-1},\theta))^2+(x_{i+1,j}-f_j(x_i,\theta))^2]\right\},$$

where $x_{[-i]}=(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n)$, which are of non-standard form for all $1\leqslant i<n$.

Here we will describe how after the introduction of strategic latent variables, $u=(u_{ij})$ and $v=(v_{ij})$, $1\leqslant i\leqslant n$, $1\leqslant j\leqslant k$, all the full conditional distributions required for the Gibbs sampler become standard. See Ref. [9] for further details. In the following we use $\mathrm{Ga}(a,b)$ to denote a gamma distribution with mean a/b; $\mathscr{U}(a,b)$ to denote a uniform distribution on the interval (a,b); $\mathrm{Exp}(a)$ to denote an exponential distribution with mean $1/a$; and I to represent the indicator function.

For $1\leqslant i\leqslant n$ and $1\leqslant j\leqslant k$ consider the random variables

$$\begin{aligned}&u_{ij}|\xi_j \sim \mathrm{Ga}(\cdot|3/2,\xi_j/2),\\&x_{ij}|x_{i-1},u_{ij},\theta \sim \mathscr{U}(\cdot|f_j(x_{i-1},\theta)-\sqrt{u_{ij}},f_j(x_{i-1},\theta)+\sqrt{u_{ij}}),\end{aligned}\tag{7}$$

and

$$\begin{aligned}&v_{ij}|\lambda_j \sim \mathrm{Ga}(\cdot|3/2,\lambda_j/2),\\&y_{ij}|x_{ij},v_{ij} \sim \mathscr{U}(\cdot|x_{ij}-\sqrt{v_{ij}},x_{ij}+\sqrt{v_{ij}}).\end{aligned}\tag{8}$$

We also have

$$\pi(u_i|\xi)=\prod_{j=1}^{k}\pi(u_{ij}|\xi_j), \quad \pi(u|\xi)=\prod_{i=1}^{n}\pi(u_i|\xi),\tag{9}$$

$$\pi(v_i|\lambda)=\prod_{j=1}^{k}\pi(v_{ij}|\lambda_j), \quad \pi(v|\lambda)=\prod_{i=1}^{n}\pi(v_i|\lambda),\tag{10}$$

along with

$$\begin{aligned}&\pi(x_i|x_{i-1},\theta,u_i)=\prod_{j=1}^{k}\pi(x_{ij}|x_{i-1},\theta,u_{ij}),\\&\pi(x|\theta,u)=\prod_{i=1}^{n}\pi(x_i|x_{i-1},\theta,u_i)\end{aligned}\tag{11}$$

and

$$\begin{aligned}&\pi(y_i|x_i,v_i)=\prod_{j=1}^{k}\pi(y_{ij}|x_{ij},v_{ij}),\\&\pi(y|x,v)=\prod_{i=1}^{n}\pi(y_i|x_i,v_i).\end{aligned}\tag{12}$$

We now calculate the augmented likelihood

$$\pi(x,y,u,v|\vartheta)=\pi(v|\vartheta)\pi(u|v,\vartheta)\pi(x|u,v,\vartheta)\pi(y|x,u,v,\vartheta).$$

Using (9)–(12) one obtains

$$\pi(x,y,u,v|\vartheta)=\prod_{i=1}^{n}\prod_{j=1}^{k}\pi(u_{ij}|\xi_j)\pi(x_{ij}|x_{i-1},\theta,u_{ij})\times\prod_{i=1}^{n}\prod_{j=1}^{k}\pi(v_{ij}|\lambda_j)\pi(y_{ij}|x_{ij},v_{ij}).$$

From (7) and (8) the above-written equation becomes

$$\pi(x, y, u, v|\vartheta) \propto \prod_{i=1}^{n}\prod_{j=1}^{k} \xi_j^{3/2} \exp\{-0.5\xi_j u_{ij}\} I(u_{ij} > (x_{ij} - f_j(x_{i-1}, \theta))^2)$$
$$\times \prod_{i=1}^{n}\prod_{j=1}^{k} \lambda_j^{3/2} \exp\{-0.5\lambda_j v_{ij}\} I(v_{ij} > (y_{ij} - x_{ij}))^2). \tag{13}$$

The marginal of the joint posterior $\pi(\vartheta, x, u, v, |y)$ with respect to the variables u, v is given by (6). For the appropriate choice of independent priors for the parameters x_0, λ, ξ and θ, the corresponding Gibbs sampler has only standard full conditionals, as follows: observing that

$$\int_{\mathbb{R}^+} \lambda_j^{3/2} \mathrm{e}^{-v_{ij}\lambda_j/2} I(v_{ij} > (y_{ij} - x_{ij})^2)\,\mathrm{d}v_{ij} \propto N(y_{ij}|x_{ij}, 1/\lambda_j)$$

and

$$\int_{\mathbb{R}^+} \xi_j^{3/2} \mathrm{e}^{-u_{ij}\xi_j/2} I(u_{ij} > (x_{ij} - f_j(x_{i-1}, \theta))^2)\,\mathrm{d}u_{ij} \propto N(x_{ij}|f_j(x_{i-1}, \theta), 1/\xi_j),$$

the joint density just described gives the required posterior in (6) once we have integrated out the latent variables u and v.

The full conditionals for the Gibbs sampler, for u_{ij} and v_{ij}, $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant k$ are

$$\pi\big(v_{ij}|\cdot\big) \propto \exp(-\lambda_j v_{ij}/2) I(v_{ij} > (y_{ij} - x_{ij})^2), \tag{14}$$

$$\pi\big(u_{ij}|\cdot\big) \propto \exp(-\xi_j u_{ij}/2) I(u_{ij} > (x_{ij} - f_j(x_{i-1}, \theta))^2). \tag{15}$$

For the choice of priors $\pi(\lambda_j) \propto \lambda_j^{-1}$ and $\pi(\xi_j) \propto \xi_j^{-1}$ (diffuse gamma priors), sampling from the full conditionals for λ_j and ξ_j is also straightforward and involves only gamma distributions. This is a standard procedure for applying non-informative prior distributions and provides proper posterior distributions even though the priors are improper. Then

$$\pi\big(\lambda_j|\cdot\big) \propto \lambda_j^{-1} \prod_{i=1}^{n} \lambda_j^{3/2} \mathrm{e}^{-u_{ij}\lambda_j/2} \propto \mathrm{Ga}(\lambda_j|3n/2, V_{nj}/2), \tag{16}$$

$$\pi(\xi_j|\cdot) \propto \xi_j^{-1} \prod_{i=1}^{n} \xi_j^{3/2} \mathrm{e}^{-u_{ij}\xi_j/2} \propto \mathrm{Ga}(\xi_j|3n/2, U_{nj}/2), \tag{17}$$

where $V_{nj} = \sum_{i=1}^{n} v_{ij}$, $U_{nj} = \sum_{i=1}^{n} u_{ij}$.

The prior for the jth coordinate of the initial condition x_0 will be uniform on the interval (x_{0j}^-, x_{0j}^+) that is

$$\pi(x_{0j}) \propto I(x_{0j}^- < x_{0j} < x_{0j}^+)$$

and the full conditional for x_{0j} is given by

$$\pi(x_{0j}|\cdot) \propto I(x_{0j}^- < x_{0j} < x_{0j}^+) \prod_{r=1}^{k} I(u_{1r} > (x_{1r} - f_r(x_0, \theta))^2), \tag{18}$$

which is uniform in the intersection of the interval (x_{0j}^-, x_{0j}^+) and the set

$$\bigcap_{r=1}^{k} \{x_{0j} : u_{1r} > (x_{1r} - f_r(x_0, \theta))^2\}.$$

Another conditional density is that of x_{ij}, conditional on all the other parameters. For $1 \leqslant i \leqslant n$ we have

$$\pi(x_{ij}|\cdot) \propto I(v_{ij} > (y_{ij} - x_{ij})^2) I(u_{ij} > (x_{ij} - f_j(x_{i-1}, \theta))^2) \times \left\{ \prod_{r=1}^{k} I(u_{i+1,r} > (x_{i+1,r} - f_r(x_i, \theta))^2) \right\}^{1-\delta_{in}}, \tag{19}$$

where δ_{in} is Kronecker's delta. Clearly for $i<n$ and $1\leqslant j\leqslant k$ the full conditional of x_{ij} is uniform with support given by the intersection of the following sets:

$$S_{ij}^1=\{x_{ij}:y_{ij}-\sqrt{v_{ij}}<x_{ij}<y_{ij}+\sqrt{v_{ij}}\},$$
$$S_{ij}^2=\{x_{ij}:f_j(x_{i-1},\theta)-\sqrt{u_{ij}}<x_{ij}<f_j(x_{i-1},\theta)+\sqrt{u_{ij}}\},$$
$$S_{ij}^3=\bigcap_{r=1}^{k}\left\{x_{ij}:x_{i+1,r}-\sqrt{u_{i+1,r}}<f_r(x_i,\theta)<x_{i+1,r}+\sqrt{u_{i+1,r}}\right\}.$$

In the case $i=n$, the full conditional for x_{nj}, for $1\leqslant j\leqslant k$ is again uniform with support $S_{nj}^1\cap S_{nj}^2$.

Notice that at each iteration of the Gibbs sampler there is a least an x_{ij} which belongs to the intersection of the sets $S_{ij}^1\cap S_{ij}^2\cap S_{ij}^3$ because of the way the u_{ij} and v_{ij} were sampled in (14) and (15), respectively. Moreover, from the assumption of the continuity of the f_r for $1\leqslant r\leqslant k$ their intersection is an interval.

The final conditional density to sample is that for θ_l, $1\leqslant l\leqslant m$. The prior for θ_l will be taken to be uniform on the interval (θ_l^-,θ_l^+). Then its full conditional becomes

$$\pi(\theta_l|\cdot)\propto I(\theta_l^-<\theta<\theta_l^+)\prod_{i=1}^{n}\prod_{j=1}^{k}I\{u_{ij}>(x_{ij}-f(x_i,\theta)^2)\},\tag{20}$$

which is uniform with support given by the intersection of the sets (θ_l^-,θ_l^+) and

$$\bigcap_{i=1}^{n}\bigcap_{j=1}^{k}\{\theta_l:x_{ij}-\sqrt{u_{ij}}<f_j(x_i,\theta)<x_{ij}+\sqrt{u_{ij}}\}.$$

This completes the description of the full conditional distributions.

4. A slice sampler for the random logistic map

Here $k=1$ and $m=1$ with $f_1=f$ and $f(x,\theta)=1-\theta x^2$. Then the model in relations (4) and (5) becomes

$$x_i|x_{i-1},\theta,\xi\sim N(\cdot|f(x_{i-1},\theta),1/\xi),$$
$$y_i|x_i,\lambda\sim N(\cdot|x_i,1/\lambda).\tag{21}$$

A parameter of interest is θ and in a Bayesian context it is assigned a prior distribution, say $\pi(\theta)$. As well, we assign non-informative gamma prior distributions to both λ and ξ and a uniform prior on $(0,1)$ to the initial condition x_0.

The joint posterior $\pi(\vartheta,u,v,x|y)$ given by (13) becomes

$$\pi(\vartheta,u,v,x|y)\propto\pi(\vartheta)\prod_{i=1}^{n}\xi^{3/2}\mathrm{e}^{-u_i\xi/2}I(u_i>(x_i-f(x_{i-1},\theta))^2)\times\prod_{i=1}^{n}\lambda^{3/2}\mathrm{e}^{-v_i\lambda/2}I(v_i>(y_i-x_i)^2),$$

where $\vartheta=(x_0,\lambda,\xi,\theta)$ is the 4 tuple of parameters we wish to estimate.

The full conditionals for the Gibbs sampler, for u_i and v_i are from relations (14) and (15)

$$\pi(v_i|\cdot)\propto\exp(-\lambda v_i/2)I(v_i>(y_i-x_i)^2),$$
$$\pi(u_i|\cdot)\propto\exp(-\xi u_i/2)I(u_i>(x_i-f(x_{i-1},\theta))^2),\quad 1\leqslant i\leqslant n.$$

It is worth pointing out that letting $\pi(w)=\mathscr{U}(0,1)$ and $u=-2\lambda^{-1}\log(w)+K$ one obtains

$$\pi(u)\propto\mathrm{e}^{-\lambda u/2}I(u>K).$$

Therefore the conditional densities for all the latent variables which are truncated exponential random variables, can be sampled as

$$u_i=-2\lambda^{-1}\log w_i+(y_i-x_i)^2$$

and

$$v_i = -2\xi^{-1}\log w_i^* + (x_i - f(x_{i-1},\theta))^2,$$

where the w_i and w_i^* are iid uniform on the unit interval.

Using diffuse gamma priors for λ and ξ, one obtains from relations (16) and (17) the full conditionals

$$\pi(\lambda|\cdot) \propto \mathrm{Ga}(\lambda|3n/2, V_n/2), \quad \pi(\xi|\cdot) \propto \mathrm{Ga}(\xi|3n/2, U_n/2),$$

with $U_n = \sum_{i=1}^n u_i$ and $V_n = \sum_{i=1}^n v_i$.

The prior for x_0 will be taken to be uniform on $(0,1)$, more commonly known as a non-informative flat prior. For the full conditional of x_0 relation (18) gives

$$\pi(x_0|\cdot) \propto I(0<x_0<1)I(u_1>(x_1 - f(x_0,\theta))^2),$$

which is uniform with support given by the intersection of $(0,1)$ and

$$\{x_0 : -\sqrt{u_1} + x_1 < f(x_0,\theta) < \sqrt{u_1} + x_1\}.$$

From (19) one obtains the density of x_i conditional on all the other parameters,

$$\pi(x_i|\cdot) \propto I(v_i>(y_i - x_i)^2)I(u_i>(x_i - f(x_{i-1},\theta))^2) \times (I(u_{i+1}>(x_{i+1} - f(x_i,\theta))^2)^{1-\delta_{in}}.$$

Clearly for $i<n$ the full conditional of x_i is uniform with support given by the intersection of the three sets

$$\begin{aligned} S_i^1 &= \{x_i : y_i - \sqrt{v_i} < x_i < y_i + \sqrt{v_i}\}, \\ S_i^2 &= \{x_i : f(x_{i-1},\theta) - \sqrt{u_i} < x_i < f(x_{i-1},\theta) + \sqrt{u_i}\}, \\ S_i^3 &= \{x_i : x_{i+1} - \sqrt{u_{i+1}} < f(x_i,\theta) < x_{i+1} + \sqrt{u_{i+1}}\}. \end{aligned}$$

From the discussion in the previous section it is to be understood that the intersection of the three sets is the interval (α_i, β_i) with

$$\alpha_i = \max\left\{ y_i - \sqrt{v_i}, 1 - \theta x_{i-1}^2 - \sqrt{u_i}, \sqrt{\frac{1 - x_{i+1} - \sqrt{u_{i+1}}}{\theta}} \right\}$$

and

$$\beta_i = \min\left\{ y_i + \sqrt{v_i}, 1 - \theta x_{i-1}^2 + \sqrt{u_i}, \sqrt{\frac{1 - x_{i+1} + \sqrt{u_{i+1}}}{\theta}} \right\}.$$

For $i = n$, the full conditional for x_n is again uniform on the interval $S_n^1 \cap S_n^2$.

The final conditional density to sample is that for θ, which is given by relation (20). It can be verified that the map $f(\cdot,\theta) : X \to X$, where $X = [-\tau(\theta), \tau(\theta)]$ and $\tau(\theta) = (1 + \sqrt{1+4\theta})/2\theta$ for $0<\theta\leqslant 2$ is compact. Setting the prior for θ to be uniform on the interval $(0,2)$, its full conditional becomes

$$\pi(\theta|\cdot) \propto I(0<\theta<2)\prod_{i=1}^n I(u_i>(x_i - f(x_{i-1},\theta))^2),$$

which is uniform with interval given by the intersection of the interval $(0,2)$ and

$$\bigcap_{i=1}^n \{\theta : x_i - \sqrt{u_i} < f(x_{i-1},\theta) < x_i + \sqrt{u_i}\}.$$

In the case of the logistic map, all of these intervals are easy to find.

Results for random logistic: We ran the slice/Gibbs sampler for the logistic map using $\theta = 0.5$, $\xi = 100$, $\lambda = 500$ and $x_0 = 0.5$. Fifty data points were generated and the slice sampler was run for 20,000 iterations. The posterior distributions for the parameters of interest, namely θ and x_0, are presented in Fig. 1 as histograms of the samples

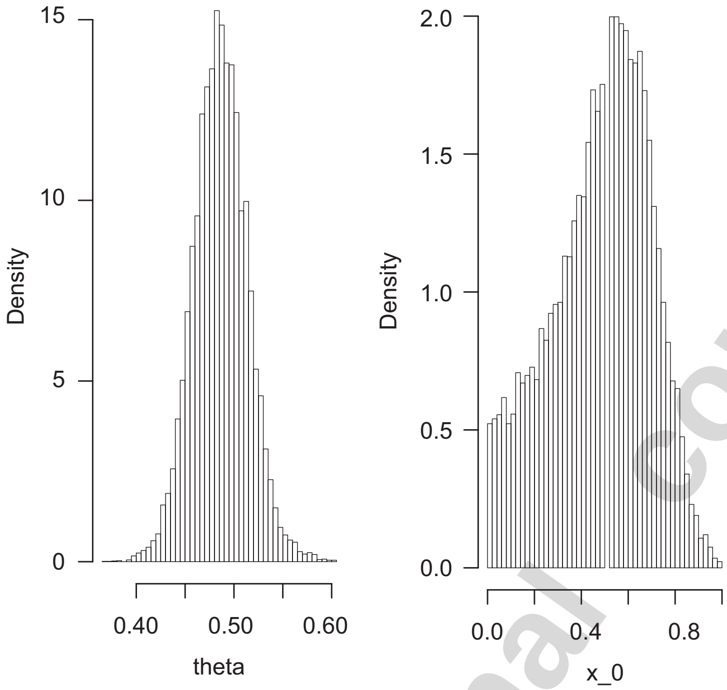

Fig. 1. Histogram representation of the posterior distributions for θ and x_0.

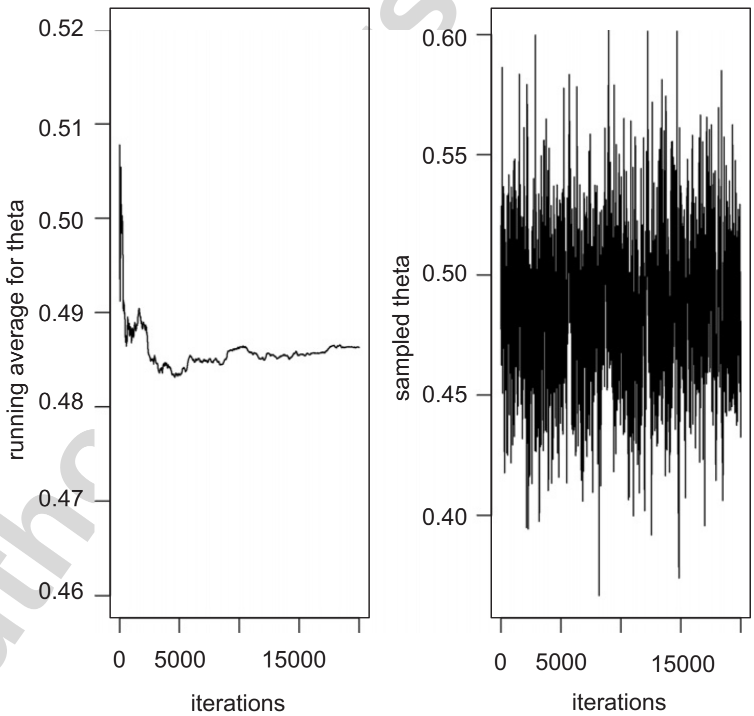

Fig. 2. Running average for θ and sampled values of θ over iterations.

from the chain. The mean values from the samples are 0.49 and 0.48 for θ and x_0, respectively. As a check that the chain has converged we plot the running average of the θ chain and to see that the chain is mixing well we also plot the θ chain. These are presented in Fig. 2. Fig. 3 provides the autocorrelation functions for the θ series. These pictures clearly show that the chain has no problems converging and is mixing well.

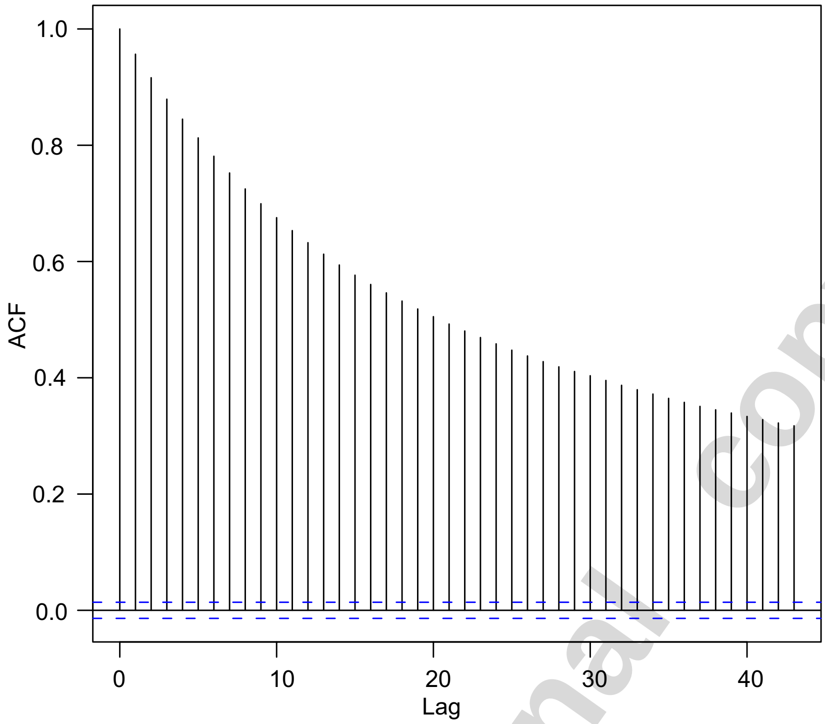

Fig. 3. Autocorrelation function for θ series.

5. A slice sampler for the random Henon map

The Henon map is a second order nonlinear map $x_i = \theta_2 - \theta_1 x_{i-1}^2 + \theta_3 x_{i-2}$. Letting $x_{i1} = x_i$ and $x_{i2} = \theta_3 x_{i-1}$ we obtain the deterministic law

$$\begin{aligned} x_{i1} &= f_1(x_{i-1}, \theta) = \theta_2 - \theta_1 x_{i-1,1}^2 + x_{i-1,2}, \\ x_{i2} &= f_2(x_{i-1}, \theta) = \theta_3 x_{i-1,1}. \end{aligned}$$

The model for the associated perturbed map with no observational noise is

$$x_{ij} | x_{i-1}, \theta, \xi_j \sim N(\cdot | f_j(x_{i-1}, \theta), 1/\xi_j), \quad j = 1, 2$$

with $\theta = (\theta_1, \theta_2, \theta_3)$, $\xi = (\xi_1, \xi_2)$, and target vector $\vartheta = (\theta, \xi, x_0)$.

Here our data consist of n consecutive observations $x = (x_1, \ldots, x_n)$ and the posterior density we want to sample from is $\pi(\vartheta | x)$. Now

$$\pi(\vartheta | x) \propto \pi(\vartheta) \prod_{i=1}^{n} \prod_{j=1}^{2} N(x_{ij} | f_j(x_{i-1}, \theta), 1/\xi_j) \tag{22}$$

and the extended joint posterior including the $2n$ latent variables u_{ij} is proportional to

$$\pi(\vartheta, u | x) \propto \pi(\vartheta) \prod_{i=1}^{n} \prod_{j=1}^{2} \xi_j^{3/2} \mathrm{e}^{-u_{ij} \xi_j / 2} I(u_{ij} > (x_{ij} - f_j(x_{i-1}, \theta))^2).$$

The full conditionals for u_{ij} and ξ_j for $1 \leqslant i \leqslant n$ and $j = 1, 2$ are the same as in relations (14) and (17), respectively. The priors for the initial conditions x_{01} and x_{02} will be taken to be uniform on the intervals (x_{01}^-, x_{01}^+) and (x_{02}^-, x_{02}^+), respectively. For the full conditionals of x_{01} and x_{02} one has

$$\begin{aligned} \pi(x_{01} | \cdot) &\propto I(x_{01}^- < x_0 < x_{01}^+) I(u_{11} > (x_{11} - f_1(x_0, \theta))^2) \times I(u_{12} > (x_{12} - f_2(x_0, \theta))^2), \\ \pi(x_{02} | \cdot) &\propto I(x_{02}^- < x_1 < x_{02}^+) I(u_{11} > (x_{11} - f_1(x_0, \theta))^2). \end{aligned}$$

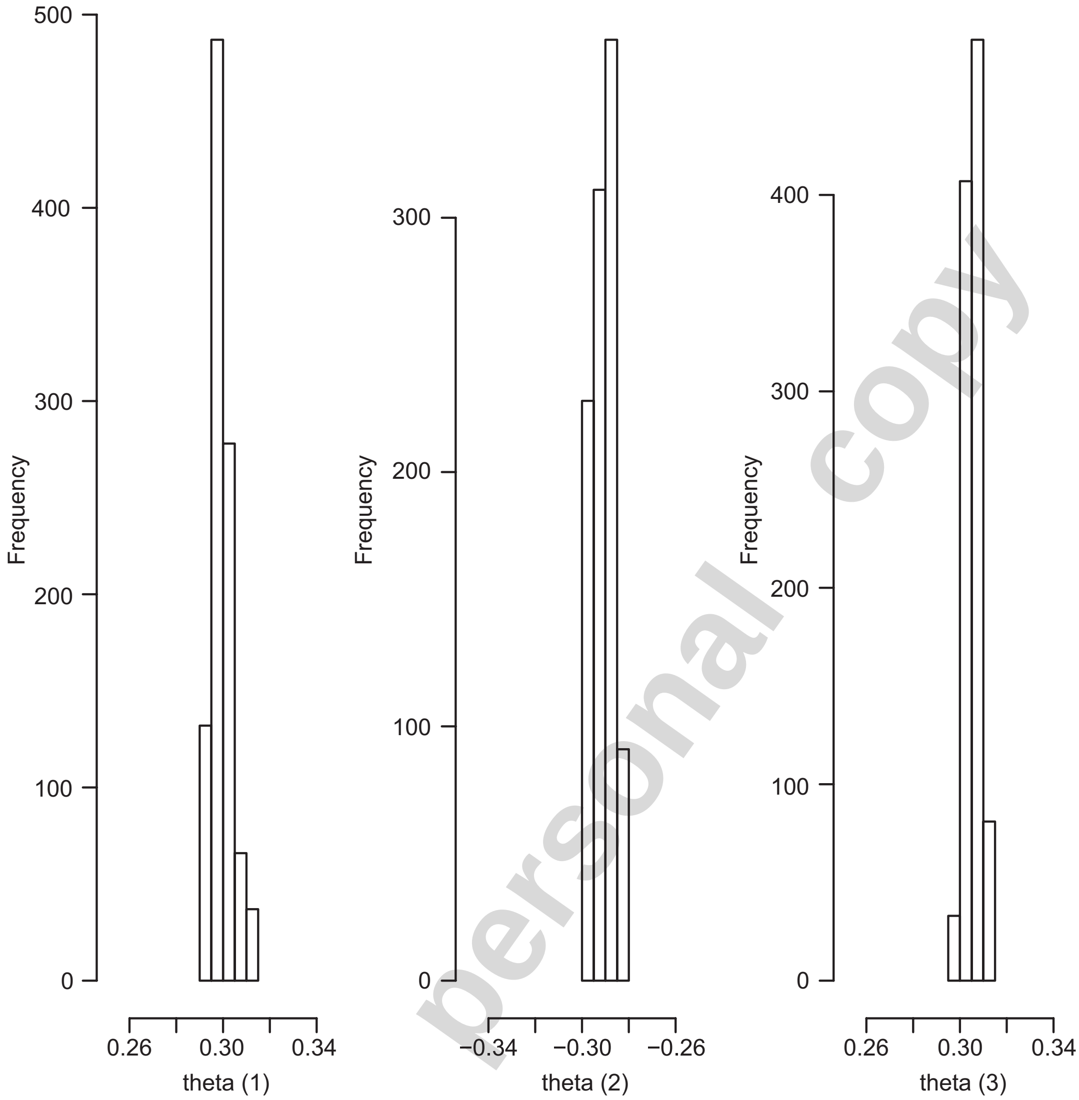

Fig. 4. Histogram representation of the posterior distributions for θ_1, θ_2 and θ_3.

The final conditional densities to sample is that for θ_1, θ_2 and θ_3. Uniform priors yield the following full conditionals:

$$\pi(\theta_s|\cdot) \propto I(\theta_s^- < \theta_s < \theta_s^+) \prod_{i=1}^{n} I(u_{i1} > (x_{i1} - f_1(x_{i-1}, \theta))^2), \quad s = 1, 2$$

and

$$\pi(\theta_3|\cdot) \propto I(\theta_3^- < \theta_3 < \theta_3^+) \prod_{i=1}^{n} I(u_{i2} > (x_{i2} - f_2(x_{i-1}, \theta))^2).$$

Numerical results for the Henon map: In this case we take 1000 data points and the parameters are $\theta_1 = 0.3$, $\theta_2 = -0.3$, $\theta_3 = 0.3$, $\xi = 50$, $\lambda = 100$ and $x_{01} = x_{02} = 0.1$. The chain was run for 1000 iterations, which due to the large sample size, is adequate for obtaining the posterior distributions. The histograms of the samples for the three θ parameters are presented in Fig. 4. As can be seen, these are very tight about the true values. This is purely due to the large sample size.

6. Discussion

Slice sampling algorithms are becoming popular throughout the statistical analysis of complicated models. For example, in Ref. [16] a slice sampling algorithm is being introduced for the ARCH model. Here we have introduced the idea for random dynamical systems. As we have demonstrated, the algorithms are fast and accurate, and are direct in that only standard distributions, that is uniform distributions, need to be sampled.

Acknowledgments

The authors are grateful to Professor A.N. Yannacopoulos for his constructive comments. The first author is financed by the Greek National Education Ministry jointly with European Union—research contract Pythagoras. The third author is financed by an EPSRC Advanced Research Fellowship. The paper was commenced during a visit of the third author to the University of the Aegean.

References

[1] D. Ruelle, F. Takens, On the nature of turbulence, Comm. Math. Phys. 20 (1971) 167–192;
D. Ruelle, F. Takens, On the nature of turbulence, Comm. Math. Phys. 23 (1971) 343–344.
[2] A. Lichtenberg, M. Lieberman, Regular and Stochastic Motion, Springer, New York, 1983.
[3] P. Gaspard, G. Nicolis, Transport properties, Lyapunov exponents and entropy per unit time, Phys. Rev. Lett. 65 (1990) 1693–1696.
[4] A.N. Yannacopoulos, G. Rowlands, Local transport coefficients for chaotic systems, J. Phys. A: Math. Gen. 30 (1997) 1503–1525.
[5] R. Meyer, N. Christensen, Bayesian reconstruction of chaotic dynamical systems, Phys. Rev. E 62 (2000) 3535–3542.
[6] R. Meyer, N. Christensen, Fast Bayesian reconstruction of chaotic dynamical systems via extended Kalman filtering, Phys. Rev. E 65 (2001) 016201–8.
[7] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, Equations of state calculations by fast computing machines, J. Chem. Phys. 21 (1953) 1087–1092.
[8] W.K. Hastings, Monte Carlo sampling methods using Markov chains with applications, Biometrika 57 (1970) 97–109.
[9] P. Damien, J. Wakefield, S.G. Walker, Gibbs sampling for Bayesian non-conjugate and hierarchical models by using auxiliary variables, J. R. Statist. Soc. B 61 (part 2) (1999) 331–344.
[10] R.M. Neal, Slice sampling, Ann. Statist. 31 (2003) 705–741.
[11] H. Voss, J. Timmer, J. Kurths, Nonlinear dynamical system identification from uncertain and indirect measurements, J. Bifur. Chaos 14 (2004) 1905–1933.
[12] A. Lasota, M.C. Mackey, Chaos, fractals, and noise, Applied Mathematical Sciences, vol. 97, Springer, New York, 1994.
[13] L. Arnold, Random dynamical systems, Springer Monographs in Mathematics, Springer, Berlin, 2003.
[14] K.R. Schenk-Hoppé, Bifurcation of the randomly perturbed logistic map, Discussion Paper 353, University of Bielefeld, 1997.
[15] S.J. Hatjispyros, A.N. Yannacopoulos, A random dynamical system model of a stylized equity market, Physica A 347 (2005) 583–612.
[16] S.G. Giakoumatos, P. Dellaportas, D. Politis, Bayesian analysis of the unobserved ARCH model, Statist. Comput. 15 (2005) 103–111.
[17] K.R. Schenk-Hoppé, Poverty traps and business cycles in a stochastic overlapping generations economy with S-shaped law of motion, J. Macroecon. 27 (2005) 275–288.
[18] R. Madan, Chua's circuit: a paradigm for chaos, World Scientific, 1993.