Available at WWW.MATHEMATICSWEB.ORG
POWERED BY SCIENCE DIRECT

Stochastic Processes and their Applications 103 (2003) 277–291

stochastic processes and their applications

www.elsevier.com/locate/spa

Asymptotic stability in distribution of stochastic differential equations with Markovian switching

Chenggui Yuan[1], Xuerong Mao[*,2]

Department of Statistics and Modelling Science, University of Strathclyde, Glasgow G1 1XH, UK

Received 19 June 2001; received in revised form 5 September 2002; accepted 6 September 2002

Abstract

Stability of stochastic differential equations with Markovian switching has recently been discussed by many authors, for example, Basak et al. (J. Math. Anal. Appl. 202 (1996) 604), Ji and Chizeck (IEEE Trans. Automat. Control 35 (1990) 777), Mariton (Jump Linear System in Automatic Control, Marcel Dekker, New York), Mao (Stochastic Process. Appl. 79 (1999) 45), Mao et al. (Bernoulli 6 (2000) 73) and Shaikhet (Theory Stochastic Process. 2 (1996) 180), to name a few. The aim of this paper is to study the asymptotic stability in distribution of nonlinear stochastic differential equations with Markovian switching.

Keywords: Generalized Itô's formula; Brownian motion; Markov chain; Asymptotic stability in distribution

1. Introduction

Stability of stochastic differential equations with Markovian switching has recently received a lot of attention. For example, Ji and Chizeck (1990) and Mariton (1990) studied the stability of a jump equation

$$dX(t) = A(r(t))X(t)\,dt, \tag{1.1}$$

where $r(t)$ is a Markov chain taking values in $S = \{1, 2, \dots, N\}$. Mao (1999) investigated the exponential stability for general nonlinear stochastic differential equations

* Corresponding author. Tel.: +44-141-548-3669; fax: +44-141-552-2079.
E-mail address: xuerong@stams.strath.ac.uk (X. Mao).
[1] Permanent address: Railway Campus, Central South University, Changsha 410075, Peoples Republic of China.
[2] Partially supported by the BBSRC/EPSRC and the Royal Society.

0304-4149/03/$ - see front matter © 2002 Elsevier Science B.V. All rights reserved.
PII: S0304-4149(02)00230-2

with Markovian switching

$$dX(t) = f(X(t), t, r(t))\,dt + g(X(t), t, r(t))\,dB(t). \tag{1.2}$$

Shaikhet (1996) took the time delay into account and considered the stability of a semi-linear stochastic differential delay equation with Markovian switching, while Mao et al. (2000) investigated the stability of a nonlinear stochastic differential delay equation with Markovian switching.

Most of these papers are concerned with asymptotic stability in probability or in mean square (i.e. the solution will tend to zero in probability or in mean square). However, this asymptotic stability is sometimes too strong and in this case it is useful to know whether or not the solution will converge in distribution (not necessary to converge to zero). This property is called asymptotic stability in distribution. Basak et al. (1996) discussed such stability for a semi-linear stochastic differential equation with Markovian switching of the form

$$dX(t) = A(r(t))X(t)\,dt + \sigma(X(t), r(t))\,dB(t). \tag{1.3}$$

Our aim is to establish much more general criteria on the asymptotic stability in distribution for a nonlinear stochastic differential equation with Markovian switching

$$dX(t) = f(X(t), r(t))\,dt + g(X(t), r(t))\,dB(t). \tag{1.4}$$

In Section 2, we shall give the formal definition of the asymptotic stability in distribution. In Section 3, a sufficient criterion on the asymptotic stability in distribution will be established under very general conditions, namely properties (P1) and (P2) (see the definitions below). Section 4 provides some sufficient conditions in terms of Lyapunov functions for properties (P1) and (P2) to hold and hence gives (indirectly) another criterion on the asymptotic stability in distribution in terms of Lyapunov functions. To make our theory more applicable, we establish a new criterion on the asymptotic stability in distribution in terms of M-matrices in Section 5. Let us emphasize that to apply this new criterion all we need to do is to verify the matrix $\mathscr{A}$ (see (5.5) below) formed by the coefficients of the equation is an M-matrix and this can be done very easily using the theory presented in Berman and Plemmons (1994). We also discuss an example to illustrate this new technique of M-matrices in the study of stochastic stability.

2. Stochastic differential equations with Markovian switching

Throughout this paper, unless otherwise specified, we let $(\Omega, \mathscr{F}, \mathscr{F}_t, P)$ be a complete probability space with a filtration $\mathscr{F}_t$ satisfying the usual conditions (i.e. it is right continuous and $\mathscr{F}_0$ contains all p-null sets). Let $B(t) = (B_t^1, \dots, B_t^m)^T$ be an m-dimensional Brownian motion defined on the probability space. Let $|\cdot|$ denote the Euclidean norm for vectors or the trace norm for matrices.

Let $r(t)$, $t \geqslant 0$, be a right-continuous Markov chain on the probability space taking values in a finite state space $S = \{1, 2, \ldots, N\}$ with generator $\Gamma = (\gamma_{ij})_{N \times N}$ given by

$$P\{r(t+\Delta) = j | r(t) = i\} = \begin{cases} \gamma_{ij}\Delta + o(\Delta) & \text{if } i \neq j, \\ 1 + \gamma_{ij}\Delta + o(\Delta) & \text{if } i = j, \end{cases}$$

where $\Delta > 0$. Here $\gamma_{ij} > 0$ is transition rate from i to j if $i \neq j$ while

$$\gamma_{ii} = -\sum_{i \neq j} \gamma_{ij}.$$

We assume that the Markov chain $r(\cdot)$ is independent of the Brownian motion $B(\cdot)$. It is well known that almost every sample path of $r(t)$ is a right-continuous step function and $r(t)$ is ergodic.

Consider a stochastic differential equation with Markovian switching of the form

$$\mathrm{d}X(t) = f(X(t), r(t))\,\mathrm{d}t + g(X(t), r(t))\,\mathrm{d}B(t) \tag{2.1}$$

on $t \geqslant 0$ with initial value $X(0) = x \in \mathbb{R}^n$, where

$$f : \mathbb{R}^n \times S \to \mathbb{R}^n \quad \text{and} \quad g : \mathbb{R}^n \times S \to \mathbb{R}^{n \times m}.$$

For the existence and uniqueness of the solution we shall impose a hypothesis:

(H) Both f and g satisfy the local Lipschitz condition and the linear growth condition. That is, for each $k = 1, 2, \ldots,$ there is an $h_k > 0$ such that

$$|f(x,i) - f(y,i)| + |g(x,i) - g(y,i)| \leqslant h_k |x - y|$$

for all $i \in S$ and those $x, y \in \mathbb{R}^n$ with $|x| \vee |y| \leqslant k$; and there is, moreover, an $h > 0$ such that

$$|f(x,i)| + |g(x,i)| \leqslant h(1 + |x|)$$

for all $x \in \mathbb{R}^n$ and $i \in S$.

It is known (cf. Mao, 1999) that under hypothesis (H), Eq. (2.1) has a unique continuous solution $X(t)$ on $t \geqslant 0$. Let $C^2(\mathbb{R}^n \times S; \mathbb{R}_+)$ denote the all nonnegative functions $V(x,i)$ on $\mathbb{R}^n \times S$ which are continuously twice differentiable in x. If $V \in C^2(\mathbb{R}^n \times S; \mathbb{R}_+)$, define an operator LV from $\mathbb{R}^n \times S$ to $\mathbb{R}$ by

$$LV(x,i) = \sum_{j=1}^{N} \gamma_{ij} V(x,j) + V_x(x,i) f(x,i) + \frac{1}{2} \operatorname{trace}[g^{\mathrm{T}}(x,i) V_{xx}(x,i) g(x,i)], \tag{2.2}$$

where

$$V_x(x,i) = \left(\frac{\partial V(x,i)}{\partial x_1}, \ldots, \frac{\partial V(x,i)}{\partial x_n} \right), \quad V_{xx}(x,i) = \left(\frac{\partial^2 V(x,i)}{\partial x_i \partial x_j} \right)_{n \times n}.$$

For the convenience of the reader we cite the generalized Itô formula established by Skorohod (1989, Lemma 3, p. 104) as a lemma.

Lemma 2.1. *Let $V \in C^2(\mathbb{R}^n \times S; \mathbb{R}_+)$ and τ_1, τ_2 be bounded stopping times such that $0 \leqslant \tau_1 \leqslant \tau_2$ a.s. If $V(X(t), r(t))$ and $LV(X(t), r(t))$ are bounded on $t \in [\tau_1, \tau_2]$ with probability* 1, *then*

$$EV(X(\tau_2), r(\tau_2)) = EV(X(\tau_1), r(\tau_1)) + E\int_{\tau_1}^{\tau_2} LV(X(s), r(s))\, ds. \tag{2.3}$$

In this paper, whenever we apply this formula we will define the bounded stopping times τ_1 and τ_2 such that $\{X(t)\colon \tau_1 \leqslant t \leqslant \tau_2\}$ is bounded in $\mathbb{R}^n$ with probability 1 and hence $V(X(t), r(t))$, etc. become bounded on $t \in [\tau_1, \tau_2]$.

We conclude this section by defining the asymptotic stability in distribution for Eq. (2.1). Let $y(t)$ denote the $\mathbb{R}^n \times S$-valued process $(X(t), r(t))$. Then $y(t)$ is a time homogeneous Markov process. Let $p(t, x, i, dy \times \{j\})$ denote the transition probability of the process $y(t)$. Let $P(t, x, i, A \times B)$ denote the probability of event $\{y(t) \in A \times B\}$ given initial condition $y(0) = (x, i)$, i.e.

$$P(t, x, i, A \times B) = \sum_{j \in B} \int_A p(t, x, i, dy \times \{j\}).$$

Definition 2.1. The process $y(t)$ is said to be asymptotically stable in distribution if there exists a probability measure $\pi(\cdot \times \cdot)$ on $\mathbb{R}^n \times S$ such that the transition probability $p(t, x, i, dy \times \{j\})$ of $y(t)$ converges weakly to $\pi(dy \times \{j\})$ as $t \to \infty$ for every $(x, i) \in \mathbb{R}^n \times S$. Eq. (2.1) is said to be asymptotically stable in distribution if $y(t)$ is asymptotically stable in distribution.

Obviously the asymptotic stability in distribution of $y(t)$ implies the existence of a unique invariant probability measure for $y(t)$.

3. Asymptotic stability in distribution

In this section, we will establish some sufficient criteria on the asymptotic stability in distribution for the solution process $y(t) = (X(t), r(t))$ of Eq. (2.1). To highlight the initial values, we let $r_i(t)$ be the Markov chain starting from state $i \in S$ at $t = 0$ and denote by $X^{x,i}(t)$ the solution of Eq. (2.1) with initial conditions $X(0) = x \in \mathbb{R}^n$ and $r(0) = i$.

Definition 3.1. Eq. (2.1) is said to have property (P1) if for any $(x, i) \in \mathbb{R}^n \times S$ and any $\varepsilon > 0$, there exists a constant $R > 0$ such that

$$P\{|X^{x,i}(t)| \geqslant R\} < \varepsilon \quad \forall t \geqslant 0. \tag{3.1}$$

Eq. (2.1) is said to have property (P2) if for any $\varepsilon > 0$ and any compact subset K of $\mathbb{R}^n$, there exists a $T = T(\varepsilon, K) > 0$ such that

$$P\{|X^{x,i}(t) - X^{y,i}(t)| < \varepsilon\} \geqslant 1 - \varepsilon \quad \forall t \geqslant T \tag{3.2}$$

whenever $(x, y, i) \in K \times K \times S$.

We observe that property (P1) guarantees that for any $(x,i)\in\mathbb{R}^n\times S$, the family of transition probabilities $\{p(t,x,i,\mathrm{d}y\times\{j\})\colon t\geqslant 0\}$ is tight. That is, for any $\varepsilon>0$ there is a compact subset $K=K(\varepsilon,x,i)$ of $\mathbb{R}^n$ such that

$$P(t,x,i,K\times S)\geqslant 1-\varepsilon \quad \forall t\geqslant 0. \tag{3.3}$$

We can now state our main result.

Theorem 3.1. *Let* (H) *hold. If Eq.* (2.1) *has properties* (P1) *and* (P2), *then Eq.* (2.1) *is asymptotically stable in distribution.*

To prove this theorem we need to introduce more notations. Let $\mathscr{P}(\mathbb{R}^n\times S)$ denote all probability measures on $\mathbb{R}^n\times S$. For $P_1,P_2\in\mathscr{P}(\mathbb{R}^n\times S)$ define metric $d_{\mathbb{L}}$ as follows:

$$d_{\mathbb{L}}(P_1,P_2)=\sup_{f\in L}\left|\sum_{i=1}^{N}\int_{\mathbb{R}^n} f(x,i)P_1(\mathrm{d}x,i)-\sum_{i=1}^{N}\int_{\mathbb{R}^n} f(x,i)P_2(\mathrm{d}x,i)\right| \tag{3.4}$$

and

$$\mathbb{L}=\{f:\mathbb{R}^n\times S\to\mathbb{R}: |f(x,i)-f(y,j)|\leqslant |x-y|+|i-j| \text{ and } |f(\cdot,\cdot)|\leqslant 1\}. \tag{3.5}$$

Let us now present three lemmas.

Lemma 3.1. *Under* (H), *for every* $p>0$ *and any compact subset* K *of* $\mathbb{R}^n$,

$$\sup_{(x,i)\in K\times S} E\left[\sup_{0\leqslant s\leqslant t}|X^{x,i}(s)|^p\right]<\infty \quad \forall t\geqslant 0. \tag{3.6}$$

For the proof of this lemma please see Mao (1999).

Lemma 3.2. *Let* (H) *hold and Eq.* (2.1) *have property* (P2). *Then, for any compact subset* K *of* $\mathbb{R}^n$,

$$\lim_{t\to\infty} d_{\mathbb{L}}(p(t,x,i,\cdot\times\cdot),p(t,y,j,\cdot\times\cdot))=0 \tag{3.7}$$

uniformly in $x,y\in K$ *and* $i,j\in S$.

Proof. For any pair of $i,j\in S$, define the stopping time

$$\beta_{ij}=\inf\{t\geqslant 0\colon r_i(t)=r_j(t)\}. \tag{3.8}$$

Recall that $r_i(t)$ is the Markov chain starting from state $i\in S$ at $t=0$ and due to the ergodicity of the Markov chain, $\beta_{ij}<\infty$ a.s. (cf. Anderson, 1991). So, for any $\varepsilon>0$, there exists a positive number T such that

$$P\{\beta_{ij}\leqslant T\}>1-\frac{\varepsilon}{8} \quad \forall i,j\in S. \tag{3.9}$$

For such T, by Lemma 3.1, there is a sufficiently large $R>0$ for

$$P(\Omega_{x,i}) > 1-\frac{\varepsilon}{16} \quad \forall (x,i)\in K\times S, \tag{3.10}$$

where $\Omega_{x,i}=\{|X^{x,i}(t)|\leqslant R \ \forall t\in[0,T]\}$.

Now, fix any $x,y\in K$ and $i,j\in S$. Let I_G denote the indicator function for set G and set $\Omega_1=\Omega_{x,i}\cap\Omega_{j,y}$. For any $f\in\mathbb{L}$ and $t\geqslant T$, compute

$$\begin{aligned}
&|Ef(X^{x,i}(t),r_i(t))-Ef(X^{y,j}(t),r_j(t))|\\
&\quad\leqslant 2P\{\beta_{ij}>T\}+E(I_{\{\beta_{ij}\leqslant T\}}|f(X^{x,i}(t),r_i(t))-f(X^{y,j}(t),r_j(t))|)\\
&\quad\leqslant \frac{\varepsilon}{4}+E[I_{\{\beta_{ij}\leqslant T\}}E(|f(X^{x,i}(t),r_i(t))-f(X^{y,j}(t),r_j(t))| \mid \mathscr{F}_{\beta_{ij}})]\\
&\quad\leqslant \frac{\varepsilon}{4}+E[I_{\{\beta_{ij}\leqslant T\}}E|f(X^{u,k}(t-\beta_{ij}),r_k(t-\beta_{ij}))-f(X^{v,k}(t-\beta_{ij}),r_k(t-\beta_{ij}))|]\\
&\quad\leqslant \frac{\varepsilon}{4}+E[I_{\{\beta_{ij}\leqslant T\}}E(2\wedge|X^{u,k}(t-\beta_{ij})-X^{v,k}(t-\beta_{ij})|)]\\
&\quad\leqslant \frac{\varepsilon}{4}+2P(\Omega-\Omega_1)+E[I_{\Omega_1\cap\{\beta_{ij}\leqslant T\}}E(2\wedge|X^{u,k}(t-\beta_{ij})-X^{v,k}(t-\beta_{ij})|)],
\end{aligned} \tag{3.11}$$

where $u=X^{x,i}(\beta_{ij})$, $v=X^{y,j}(\beta_{ij})$ and $k=r_i(\beta_{ij})=r_j(\beta_{ij})$. Note that given $\omega\in\Omega_1\cap\{\beta_{ij}\leqslant T\}$, $|u|\vee|v|\leqslant R$. So, by property (P2), there exists a constant T_1 such that

$$E(2\wedge|X^{u,k}(t-\beta_{ij})-X^{v,k}(t-\beta_{ij})|)<\frac{\varepsilon}{2} \quad \forall t\geqslant T+T_1. \tag{3.12}$$

It therefore follows from (3.10)–(3.12) that

$$|Ef(X^{x,i}(t),r_i(t))-Ef(X^{y,j}(t),r_j(t))|\leqslant \frac{\varepsilon}{4}+\frac{\varepsilon}{4}+\frac{\varepsilon}{2}=\varepsilon \quad \forall t\geqslant T+T_1.$$

Since f, etc. are arbitrary, we must have that

$$\sup_{f\in\mathbb{L}}|Ef(X^{x,i}(t),r_i(t))-Ef(X^{y,j}(t),r_j(t))|\leqslant\varepsilon \quad \forall t\geqslant T+T_1,$$

namely,

$$d_{\mathbb{L}}(p(t,x,i,\cdot\times\cdot),p(t,y,j,\cdot\times\cdot))\leqslant\varepsilon \quad \forall t\geqslant T+T_1$$

for all $x,y\in K$ and $i,j\in S$. The proof is complete. □

Lemma 3.3. *Let* (H) *hold. If Eq.* (2.1) *has properties* (P1) *and* (P2), *then for any* $(x,i)\in\mathbb{R}^n\times S$, $\{p(t,x,i,\cdot\times\cdot)\colon\ t\geqslant 0\}$ *is Cauchy in the space* $\mathscr{P}(\mathbb{R}^n\times S)$ *with metric* $d_{\mathbb{L}}$.

Proof. Fix any $(x,i)\in\mathbb{R}^n\times S$. We need to show that for any $\varepsilon>0$, there is a $T>0$ such that

$$d_{\mathbb{L}}(p(t+s,x,i,\cdot\times\cdot),p(t,x,i,\cdot\times\cdot))\leqslant\varepsilon \quad \forall t\geqslant T,\ s>0.$$

This is equivalent to

$$\sup_{f\in\mathbb{L}} |Ef(X^{x,i}(t+s), r_i(t+s)) - Ef(X^{x,i}(t), r_i(t))| \leqslant \varepsilon \quad \forall t \geqslant T,\ \ s > 0. \tag{3.13}$$

For any $f \in \mathbb{L}$ and $t, s > 0$, compute

$$\begin{aligned}
&|Ef(X^{x,i}(t+s), r_i(t+s)) - Ef(X^{x,i}(t), r_i(t))| \\
&\quad = |E[E(f(X^{x,i}(t+s), r_i(t+s))|\mathscr{F}_s)] - Ef(X^{x,i}(t), r_i(t))| \\
&\quad = \left| \sum_{l=1}^{N} \int_{\mathbb{R}^n} Ef(X^{z,l}(t), r_l(t))\, p(s,x,i,\mathrm{d}z \times \{l\}) - Ef(X^{y,j}(t), r_i(t)) \right| \\
&\quad \leqslant \sum_{l=1}^{N} \int_{\mathbb{R}^n} |Ef(X^{z,l}(t), r_l(t)) - Ef(X^{x,i}(t), r_i(t))|\ p(s,x,i,\mathrm{d}z \times \{l\}) \\
&\quad \leqslant 2P(s,x,i,\bar{B}_R \times S) \\
&\qquad + \sum_{l=1}^{N} \int_{B_R} |Ef(X^{z,l}(t), r_l(t)) - Ef(X^{x,i}(t), r_i(t))|\ p(s,x,i,dz \times \{l\}).
\end{aligned} \tag{3.14}$$

where $B_R = \{x \in \mathbb{R}^n\colon\ |x| \leqslant R\}$ and $\bar{B}_R = \mathbb{R}^n - B_R$. By property (P1) (or (3.3)), there is a positive number R sufficiently large for

$$P(s,x,i,\bar{B}_R \times S) < \frac{\varepsilon}{4} \quad \forall s \geqslant 0. \tag{3.15}$$

On the other hand, by Lemma 3.2, there is a $T > 0$ such that

$$\sup_{f\in\mathbb{L}} |Ef(X^{z,l}(t), r_l(t)) - Ef(X^{x,i}(t), r_i(t))| < \frac{\varepsilon}{2} \quad \forall t \geqslant T \tag{3.16}$$

whenever $(z,l) \in B_R \times S$. Substituting (3.15) and (3.16) into (3.14) yields

$$|Ef(X^{x,i}(t+s), r_i(t+s)) - Ef(X^{x,i}(t), r_i(t))| < \varepsilon \quad \forall t \geqslant T,\ \ s > 0.$$

Since f is arbitrary, the desired inequality (3.13) must hold. $\square$

We can now easily prove our main result Theorem 3.1.

Proof of Theorem 3.1. By definition, we need to show that there exists a probability measure $\pi(\cdot \times \cdot) \in \mathscr{P}(\mathbb{R}^n \times S)$ such that for any $(x,i) \in \mathbb{R}^n \times S$, the transition probabilities $\{p(t,x,i,\cdot \times \cdot)\colon\ t \geqslant 0\}$ converge weakly to $\pi(\cdot \times \cdot)$. Recalling the well-known fact that the weak convergence of probability measures is a metric concept (cf. Ikeda and Watanabe, 1981, Proposition 2.5), we therefore need to show that for any $(x,i) \in \mathbb{R}^n \times S$,

$$\lim_{t\to\infty} d_{\mathbb{L}}(p(t,x,i,\cdot \times \cdot), \pi(\cdot \times \cdot)) = 0. \tag{3.17}$$

By Lemma 3.3, $\{p(t,0,1,\cdot \times \cdot)\colon\ t \geqslant 0\}$ is Cauchy in the space $\mathscr{P}(\mathbb{R}^n \times S)$ with metric $d_{\mathbb{L}}$. So there is a unique $\pi(\cdot \times \cdot) \in \mathscr{P}(\mathbb{R}^n \times S)$ such that

$$\lim_{t\to\infty} d_{\mathbb{L}}(p(t,0,1,\cdot \times \cdot), \pi(\cdot \times \cdot)) = 0.$$

Now, for any $(x,i)\in\mathbb{R}^n\times S$, by Lemma 3.2,

$$\lim_{t\to\infty} d_{\mathbb{L}}(p(t,x,i,\cdot\times\cdot),\pi(\cdot\times\cdot))$$
$$\leqslant \lim_{t\to\infty}[d_{\mathbb{L}}(p(t,0,1,\cdot\times\cdot),\pi(\cdot\times\cdot))+d_{\mathbb{L}}(p(t,x,i,\cdot\times\cdot),p(t,0,1,\cdot\times\cdot))$$
$$=0$$

as required. □

4. Sufficient criteria for properties (P1) and (P2)

Theorem 3.1 depends on properties (P1) and (P2). It is therefore necessary to establish sufficient criteria for these properties so that Theorem 3.1 is applicable. On the other hand, property (P1) is concerned with boundedness while property (P2) is associated with uniformly asymptotic stability. The study on both of them has its own right. The importance of this section is therefore clear.

We shall need two more notations. Let $\mathscr{K}$ denote the family of nondecreasing functions $\mu:\mathbb{R}_+\to\mathbb{R}_+$ such that $\mu(0)=0$ while $\mathscr{K}_\infty$ denote the family of functions $\mu\in\mathscr{K}$ such that $\mu(u)\to\infty$ as $u\to\infty$.

The following lemma gives a criterion for property (P1).

Lemma 4.1. *Assume that there exist functions $V\in C^2(\mathbb{R}^n\times S;\ \mathbb{R}_+)$, $\mu\in\mathscr{K}_\infty$ and positive numbers β and λ_1 such that*

$$\mu(|x|)\leqslant V(x,i) \tag{4.1}$$

and

$$LV(x,i)\leqslant -\lambda_1 V(x,i)+\beta \tag{4.2}$$

for all $(x,i)\in\mathbb{R}^n\times S$. Then Eq. (2.1) has property (P1).

Proof. Fix any $(x,i)\in\mathbb{R}^n\times S$ and write $X^{x,i}(t)=X(t)$. Let k be a positive integer. Define the stopping time

$$\rho_k=\inf\{t>0\colon\ |X(t)|\geqslant k\}.$$

Clearly, $\rho_k\to\infty$ almost surely as $k\to\infty$. Let $t_k=\rho_k\wedge t$ for any $t\geqslant 0$. The generalized Itô formula (i.e. Lemma 2.1) shows that

$$E[e^{\lambda_1 t_k}V(X(t_k),r_i(t_k))]=V(x,i)+E\int_0^{t_k} e^{\lambda_1 s}LV(X(s),r_i(s))\,ds$$
$$+\lambda_1 E\int_0^{t_k} e^{\lambda_1 s}V(X(s),r_i(s))\,ds.$$

By conditions (4.1) and (4.2),

$$E[e^{\lambda_1 t_k}V(X(t_k),r_i(t_k))]\leqslant V(x,i)+\beta\int_0^t e^{\lambda_1 s}\,ds=V(x,i)+\frac{\beta}{\lambda_1}[e^{\lambda_1 t}-1].$$

Letting $k \to \infty$ gives

$$EV(X(t), r_i(t)) \leqslant \frac{1}{c_1}\left(\frac{\beta}{\lambda_1} + V(x,i)\right). \tag{4.3}$$

This, together with (4.1), yields

$$E\mu(|X(t)|) \leqslant C \quad \forall t \geqslant 0,$$

where C denotes the right-hand side term of (4.3). Therefore

$$P\{|X(t)| \geqslant R\} \leqslant \frac{E\mu(|X(t)|)}{R} \leqslant \frac{C}{R} \quad \forall t \geqslant 0.$$

Now for any $\varepsilon > 0$, choosing R sufficiently large for $C/R < \varepsilon$, we get the result. □

In what follows we shall establish a criterion for property (P2). Clearly, we need to consider the difference between two solutions of Eq. (2.1) starting from different initial values, namely

$$X^{x,i}(t) - X^{y,i}(t) = x - y + \int_0^t [f(X^{x,i}(s), r_i(s)) - f(X^{y,i}, r_i(s))]\,\mathrm{d}s$$
$$+ \int_0^t [g(X^{x,i}(s), r_i(s)) - g(X^{y,i}, r_i(s))]\,\mathrm{d}B(s). \tag{4.4}$$

For a given function $U \in C^2(\mathbb{R}^n \times S;\ \mathbb{R}_+)$, we define an operator $\mathscr{L}U : \mathbb{R}^n \times \mathbb{R}^n \times S \to \mathbb{R}$ associated with Eq. (4.4) by

$$\mathscr{L}U(x,y,i) = \sum_{j=1}^{N} \gamma_{ij} U(x-y,j) + U_x(x-y,i)[f(x,i) - f(y,i)]$$
$$+ \frac{1}{2}\operatorname{trace}([g(x,i) - g(y,i)]^{\mathrm{T}} U_{xx}(x-y,i)[g(x,i) - g(y,i)]). \tag{4.5}$$

Please note the difference between this operator and the other one LV defined by (2.2).

Lemma 4.2. *If there exist functions $U \in C^2(\mathbb{R}^n \times S;\ \mathbb{R}_+)$, $\mu_1 \in \mathscr{K}_\infty$ and $\mu_2 \in \mathscr{K}$ such that*

$$U(0,i) = 0 \quad \forall i \in S, \tag{4.6}$$

$$\mu_1(|x|) \leqslant U(x,i) \quad \forall (x,i) \in \mathbb{R}^n \times S, \tag{4.7}$$

$$\mathscr{L}U(x,y,i) \leqslant -\mu_2(|x-y|) \quad \forall (x,y,i) \in \mathbb{R}^n \times \mathbb{R}^n \times S, \tag{4.8}$$

then Eq. (2.1) *has property* (P2).

Proof. For any $\varepsilon \in (0,1)$, by the continuity of U and (4.6), we can choose $\alpha \in (0,\varepsilon)$ sufficiently small for

$$\frac{\sup_{|x| \leqslant \alpha, i \in S} U(x,i)}{\mu_1(\varepsilon)} < \frac{\varepsilon}{2}. \tag{4.9}$$

Let K be any compact subset of $\mathbb{R}^n$ and fix any $x, y \in K$ and $i \in S$. Define the stopping times

$$\tau_\alpha = \inf\{t \geqslant 0: |X^{x,i}(t) - X^{y,i}(t)| \leqslant \alpha\}$$

and

$$\tau_\beta = \inf\{t \geqslant 0: |X^{x,i}(t) - X^{y,i}(t)| \geqslant \beta\},$$

where $\beta > \alpha$. Let $t_\beta = \tau_\beta \wedge t$. By the generalized Itô formula and (4.7), we can derive that for any $t > 0$,

$$\begin{aligned} \mu_1(\beta) P\{\tau_\beta \leqslant t\} &\leqslant EU(X^{x,i}(t_\beta) - X^{y,i}(t_\beta), r_i(t_k)) \\ &= U(x-y,i) + E\int_0^{t_\beta} \mathscr{L}U(X^{x,i}(s), X^{y,i}(s), r_i(s))\,\mathrm{d}s \\ &\leqslant U(x-y,i). \end{aligned}$$

Consequently,

$$P\{\tau_\beta \leqslant t\} \leqslant \frac{U(x-y,i)}{\mu_1(\beta)}.$$

Noting that $U(x-y,i)$ is bounded when $(x,y,i) \in K \times K \times S$, this implies that there exists a $\beta = \beta(K,\varepsilon) > 0$ such that

$$P\{\tau_\beta < \infty\} \leqslant \frac{\varepsilon}{4}. \tag{4.10}$$

Fix the β and let $t_\alpha = \tau_\alpha \wedge \tau_\beta \wedge t$. By the generalized Itô formula and (4.8), we can derive that for any $t > 0$,

$$\begin{aligned} 0 &\leqslant EU(X^{x,i}(t_\alpha) - X^{y,i}(t_\alpha), r_i(t_\alpha)) \\ &= U(x-y,i) + E\int_0^{t_\alpha} \mathscr{L}U(X^{x,i}(s), X^{y,i}(s), r_i(s))\,\mathrm{d}s \\ &\leqslant U(x-y,i) - \mu_2(\alpha) E(\tau_\alpha \wedge T_\beta \wedge t). \end{aligned}$$

Consequently,

$$tP\{\tau_\alpha \wedge \tau_\beta \geqslant t\} \leqslant E(\tau_\alpha \wedge \tau_\beta \wedge t) \leqslant \frac{U(x-y,i)}{\mu_2(\alpha)}.$$

Therefore, there exists a constant $T = T(K,\varepsilon) > 0$ such that

$$P\{\tau_\alpha \wedge \tau_\beta \leqslant T\} > 1 - \frac{\varepsilon}{4}.$$

By (4.10), we have

$$1-\frac{\varepsilon}{4}<P\{\tau_\alpha\wedge\tau_\beta\leqslant T\}\leqslant P\{\tau_\alpha\leqslant T\}+P\{\tau_\beta<\infty\}\leqslant P\{\tau_\alpha\leqslant T\}+\frac{\varepsilon}{4},$$

which yields

$$P\{\tau_\alpha\leqslant T\}\geqslant 1-\frac{\varepsilon}{2}. \tag{4.11}$$

Now, define stopping times

$$\sigma=\inf\{t\geqslant\tau_\alpha\wedge T:\ |X^{x,i}(t)-X^{y,i}(t)|\geqslant\varepsilon\}.$$

Let $t>T$ and compute

$$\begin{aligned}
&P(\{\tau_\alpha\leqslant T\}\cap\{\sigma\leqslant t\})\mu_1(\varepsilon)\\
&\quad\leqslant EI_{\{\tau_\alpha\leqslant T,\sigma\leqslant t\}}U(X^{x,i}(\sigma\wedge t)-X^{y,i}(\sigma\wedge t),r_i(\sigma\wedge t))\\
&\quad\leqslant EI_{\{\tau_\alpha\leqslant T\}}U(X^{x,i}(\tau_\alpha\wedge t)-X^{y,i}(\tau_\alpha\wedge t),r_i(\tau_\alpha\wedge t))\\
&\quad\leqslant EI_{\{\tau_\alpha\leqslant T\}}U(X^{x,i}(\tau_\alpha)-X^{y,i}(\tau_\alpha),r_i(\tau_\alpha))\\
&\quad\leqslant P\{\tau_\alpha\leqslant T\}\sup_{|x|\leqslant\alpha,i\in S}U(\alpha,i).
\end{aligned}$$

This, together with (4.9), yields

$$P(\{\tau_\alpha\leqslant T\}\cap\{\sigma\leqslant t\})<\frac{\varepsilon}{2}. \tag{4.12}$$

By (4.11) and (4.12), we obtain

$$P\{\sigma\leqslant t\}\leqslant P(\{\tau_\alpha\leqslant T\}\cap\{\sigma\leqslant t\})+P\{\tau_\alpha>T\}<\varepsilon.$$

Letting $t\to\infty$ we have

$$P\{\sigma<\infty\}\leqslant\varepsilon. \tag{4.13}$$

This means that for any $(x,y,i)\in K\times K\times S$, we must have

$$P\{|X^{x,i}(t)-X^{y,i}(t)|<\varepsilon\}\geqslant 1-\varepsilon\quad\forall t\geqslant T$$

as required. The proof is therefore complete. □

5. Criterion in terms of M-matrices

To make our theory more applicable, let us now use the results obtained previously to establish a new criterion in terms of M-matrices, which can be verified easily in applications. For the convenience of the reader, let us cite some useful results on M-matrices. For more detailed information please see Berman and Plemmons (1994). We will need a few more notations. If B is a vector or matrix, by $B\gg 0$ we mean all elements of B are positive. If B_1 and B_2 are vectors or matrices with same dimensions we write $B_1\gg B_2$ if and only if $B_1-B_2\gg 0$. Moreover, we also adopt here the traditional notation by letting

$$Z^{N\times N} = \{A = (a_{ij})_{N\times N}\colon\ a_{ij} \leqslant 0,\ i \neq j\}.$$

Definition 5.1. A square matrix $A = (a_{ij})_{N\times N}$ is called a nonsingular M-matrix if A can be expressed in the form $A = sI - B$ with $s > \rho(B)$ while all the elements of B are nonnegative, where I is the identity matrix and $\rho(B)$ the spectral radius of B.

It is easy to see that a nonsingular M-matrix A has nonpositive off-diagonal and positive diagonal entries, that is

$$a_{ii} > 0 \text{ while } a_{ij} \leqslant 0 \quad i \neq j.$$

In particular, $A \in Z^{N\times N}$. There are many conditions which are equivalent to the statement that A is a nonsingular M-matrix and we now cite some of them for the use of this paper.

Lemma 5.1. *If $A \in Z^{N\times N}$, then the following statements are equivalent*:

(1) *A is a nonsingular M-matrix.*
(2) *A is semipositive*; *that is, there exists $x \gg 0$ in R^N such that $Ax \gg 0$.*
(3) *A^{-1} exists and its elements are all nonnegative.*
(4) *All the leading principal minors of A are positive*; *that is*

$$\begin{vmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{vmatrix} > 0 \quad \textit{for every } k = 1, 2, \ldots, N.$$

The following result gives a new criterion on asymptotic stability in distribution where the conditions are described in terms of an M-matrix.

Theorem 5.1. *Let* (H) *hold. Assume that for all $x, y \in \mathbb{R}^n$ and $i \in S$,*

$$x^{\mathrm{T}} f(x,i) \leqslant \beta_i |x|^2 + \alpha, \tag{5.1}$$

$$(x-y)^{\mathrm{T}}(f(x,i) - f(y,i)) \leqslant \beta_i |x-y|^2, \tag{5.2}$$

$$|g(x,i)|^2 \leqslant \delta_i |x|^2 + \alpha, \tag{5.3}$$

$$|g(x,i) - g(y,i)|^2 \leqslant \delta_i |x-y|^2, \tag{5.4}$$

where α, β_i and δ_i are constants. If

$$\mathscr{A} := -\mathrm{diag}(2\beta_1 + \delta_1, \ldots, 2\beta_N + \delta_N) - \Gamma \tag{5.5}$$

is an M-matrix, *then Eq.* (2.1) *is asymptotically stable in distribution.*

Proof. By Lemma 5.1 there is a vector $\vec{q} = (q_1, \dots, q_N)^T \gg 0$ such that

$$\vec{\lambda} = (\lambda_1, \dots, \lambda_N)^T := \mathscr{A}\vec{q} \gg 0.$$

Set $\lambda = \min_{1 \leqslant i \leqslant N} \lambda_i > 0$ and $q = \max_{1 \leqslant i \leqslant N} q_i > 0$. Define functions $V, U : \mathbb{R}^n \times S \to \mathbb{R}_+$ by

$$V(x,i) = U(x,i) = q_i |x|^2.$$

By (5.1)–(5.4) we compute the operator LV from $\mathbb{R}^n \times S$ to $\mathbb{R}$ as follows:

$$\begin{aligned} LV(x,i) &= 2q_i x^T f(x,i) + q_i |g(x,i)|^2 + \sum_{j=1}^N \gamma_{ij} q_j |x|^2 \\ &\leqslant \left(2\beta_i q_i + \delta_i q_i + \sum_{i=1}^N \gamma_{ij} q_j \right) |x|^2 + 3 q_i \alpha \\ &= -\lambda_i |x|^2 + 3 q_i \alpha \\ &\leqslant -\lambda |x|^2 + 3 q \alpha. \end{aligned}$$

Also, compute the operator $\mathscr{L}U$ from $\mathbb{R}^n \times \mathbb{R}^n \times S$ to $\mathbb{R}$:

$$\begin{aligned} \mathscr{L}U(x,y,i) &= 2q_i (x-y)^T (f(x,i) - f(y,i) + q_i |g(x,i) - g(y,i)|^2 \\ &\quad + \sum_{j=1}^N \gamma_{ij} q_j |x-y|^2 \\ &\leqslant \left(2\beta_i q_i + \delta_i q_i + \sum_{i=1}^N \gamma_{ij} q_j \right) |x-y|^2 \\ &\leqslant -\lambda |x-y|^2. \end{aligned}$$

By Lemmas 4.1 and 4.2, Eq. (2.1) has properties (P1) and (P2) so the conclusion follows from Theorem 3.1. □

Let us emphasize that to apply Theorem 5.1 all we need to do is to verify the matrix $\mathscr{A}$ defined by (5.5) is an M-matrix and this can be done very easily using the theory presented in Berman and Plemmons (1994), e.g. Lemma 5.1. We now discuss an example to illustrate this new technique of M-matrices in the study of stochastic stability.

Example 5.1. Let $B(t)$ be a scalar Brownian motion. Let α and σ be constants. Consider the Ornstein–Uhlenbeck process

$$\mathrm{d}X(t) = \alpha X(t)\,\mathrm{d}t + \sigma\,\mathrm{d}B(t), \qquad t \geqslant 0. \tag{5.6}$$

Given initial value $X(0) = x_0 \in \mathbb{R}^n$, it has the unique solution

$$x(t) = e^{\alpha t} x_0 + \sigma \int_0^t e^{\alpha(t-s)} \, dB(s). \tag{5.7}$$

It is easy to observe that when $\alpha < 0$, the distribution of the solution $X(t)$ converges to the normal distribution $N(0, \sigma^2/2|\alpha|)$ as $t \to \infty$ for arbitrary x_0, but when $\alpha \geqslant 0$, the distribution will not converge. In other words, Eq. (5.6) is asymptotically stable in distribution if $\alpha < 0$ but it is not if $\alpha \geqslant 0$.

Now let $r(t)$ be a right-continuous Markov chain taking values in $S = \{1, 2\}$ with generator

$$\Gamma = (\gamma_{ij})_{2\times 2} = \begin{pmatrix} -4 & 4 \\ \gamma & -\gamma \end{pmatrix},$$

where $\gamma > 0$. Assume that $B(t)$ and $r(t)$ are independent. Consider a one-dimensional stochastic differential equation with Markovian switching

$$dX(t) = \alpha(r(t))X(t) \, dt + \sigma \, dB(t) \tag{5.8}$$

on $t \geqslant 0$, where σ is a constant, $\alpha(1) = 1$ and $\alpha(2) = -\frac{1}{2}$. This system can be regarded as the result of two equations

$$dX(t) = X(t) \, dt + \sigma \, dB(t) \tag{5.9}$$

and

$$dX(t) = -\tfrac{1}{2} X(t) \, dt + \sigma \, dB(t) \tag{5.10}$$

switching from one to the other according to the law of the Markov chain. From the property of the Ornstein–Uhlenbeck process (5.6) we observe that Eq. (5.9) is not asymptotically stable in distribution through Eq. (5.10). However, we shall see that due to the Markovian switching the overall system (5.8) will be asymptotically stable in distribution. In fact, with obvious definitions of f and g, it is easy to see conditions (5.1)–(5.4) hold with

$$\beta_1 = 1, \quad \beta_2 = -\tfrac{1}{2}, \quad \delta_1 = \delta_2 = 0, \quad \alpha = \sigma^2.$$

So the matrix defined by (5.5) becomes

$$\mathscr{A} = -\mathrm{diag}(2, -1) - \Gamma = \begin{pmatrix} 2 & -4 \\ -\gamma & 1+\gamma \end{pmatrix}.$$

Since $\gamma > 0$, this is an M-matrix if and only if

$$2(1+\gamma) - 4\gamma > 0, \quad \text{namely } \gamma < 1.$$

By Theorem 5.1, we can therefore conclude that Eq. (5.8) is asymptotically stable in distribution if $\gamma \in (0, 1)$.

Acknowledgements

The authors would like to thank the referees and the associated editor for their useful comments and suggestions. The authors would also like to thank the EPSRC/BBSRC, the Royal Society and the University of Strathclyde for the financial support.

References

Anderson, W.J., 1991. Continuous-Time Markov Chains. Springer, Berlin.

Basak, G.K., Bisi, A., Ghosh, M.K., 1996. Stability of a random diffusion with linear drift. J. Math. Anal. Appl. 202, 604–622.

Berman, A., Plemmons, R.J., 1994. Nonnegative Matrices in the Mathematical Sciences. SIAM, Philadelphia, PA.

Ikeda, N., Watanabe, S., 1981. Stochastic Differential Equations and Diffusion Processes. North-Holland, Amsterdam.

Ji, Y., Chizeck, H.J., 1990. Controllability, stabilizability and continuous-time Markovian jump linear quadratic control. IEEE Trans. Automat. Control. 35, 777–788.

Mao, X., 1999. Stability of stochastic differential equations with Markovian switching. Stochastic Process. Appl. 79, 45–67.

Mao, X., Matasov, A., Piunovskiy, A.B., 2000. Stochastic differential delay equations with Markovian switching. Bernoulli 6, 73–90.

Mariton, M., 1990. Jump Linear Systems in Automatic Control. Marcel-Dekker, New York.

Shaikhet, L., 1996. Stability of stochastic hereditary systems with Markov switching. Theory Stochastic Process. 2, 180–184.

Skorohod, A.V., 1989. Asymptotic Methods in the Theory of Stochastic Differential Equations. American Mathematical Society, Providence, RI.