Annales Academiæ Scientiarum Fennicæ
Series A. I. Mathematica
Volumen 17, 1992, 85–91

CONSISTENT STATISTICAL ESTIMATION IN SEMIMARTINGALE MODELS OF STOCHASTIC APPROXIMATION

A.V. Melnikov and A.E. Rodkina

Steklov Mathematical Institute, Vavilov Street 42, 117966 Moscow GSP-1, Russia

1. Consider, on a standard stochastic basis $\big(\Omega, F, (F_t)_{t\geq 0}, \mathbf{P}\big)$, the stochastic equation

$$X_t = X_0 + \int_0^t \gamma_s R(X_s)\, da_s + \int_0^t \gamma_s \sigma(s, X_s)\, dm_s, \tag{1}$$

in the space R^d, $d \geq 1$, where $a = (a_t, F_t)_{t\geq 0}$, $a_0 = 0$, is a predictable (continuous) increasing one-dimensional process; $m = (m_t, F_t)$, $m_0 = 0$, is a continuous local martingale with values in the space R^k, $k \geq 1$; $R(x) = \big(R_i(x)\big)_{i=1,\dots,d}$, and $\sigma(t,\omega,x) = \big(\sigma_{ij}(t,\omega,x)\big)_{i=1,\dots,d;j=1,\dots,k}$, are $(d\times 1)$ and, respectively, $(d\times k)$ matrix-valued predictable functions; $\gamma = (\gamma_t, F_t)_{t\geq 0}$ is a positive predictable process and $\mathbf{E}\,\|X_0\|^2 < \infty$ (here $\|x\|$ and (x,y) are the norm and scalar product of vectors x and y from a finite-dimensional Euclidean space).

The solution X_t, $t \geq 0$, of the equation (1) is an example of a continuous procedure to approximate stochastically the single root Θ of an equation of the form

$$R(\Theta) = 0. \tag{2}$$

The goal of this paper is to give general conditions of (a.s.) convergence of X_t to the root $\Theta = 0$ (for simplicity). The approach to stochastic approximation procedures, based on stochastic equations with respect to semimartingales, was suggested by Melnikov [1] in the case of a one-dimensional processes, a linear bounded real function $R(x)$ and, respectively, $\sigma \equiv 1$ (see also the corresponding references therein). Here we shall relax these conditions. In particular, our assumptions (see Theorem 1) on $\sigma(x) \sim x \ln^{1/2} x$ as $x \to \infty$ are typical unexplosing conditions for the strong solutions of stochastic equations. In addition, we do not make any particular growth assumptions for the one-dimensional model (1) (see Theorem 2). We also present two exampes which show that our results are nontrivial and new for classical diffusion models (see [2]).

doi:10.5186/aasfm.1992.1711

2. To formulate the basic result (Theorem 1) we need some notation.
By Δ we denote the class of functions $L(u)$, $u \geq 0$, defined by

$$L(u) = L_{\varepsilon,m}(u) = \text{const} \times \prod_{i=0}^{m} \ln_i(u + \varepsilon_m + \varepsilon), \qquad m = 0, 1, \ldots, \ \varepsilon > 0,$$

with
$e_m = \exp e_{m-1}$, $m \geq 1$, $e_0 = 0$;
$\ln_{i+1}(x) = \ln_i\big(\ln(x)\big)$, $\ln_0(x) = x$.

Denote by $C_t^m = \big(\langle m^i, m^j \rangle_t\big)_{i,j \leq k}$, the matrix of quadratic characteristics of the components m^i and m^j of the local martingale m (cf. (1)), and $\langle m \rangle_t = \operatorname{tr} C_t^m$.

We will consider the following assumption: there is a predictable increasing one-dimensional process V such that

(3) $\qquad C_t^m$ and a_t are absolutely continuous with respect to V_t,

and the corresponding densities are β_t and α_t, respectively.

All notions and notation not recalled here can be found in the books [3] or [4].

Theorem 1. *Let X_t, $t \geq 0$, be a continuous solution of the equation* (1). *Suppose the following assumptions are fulfilled*

(1) $\int_0^\infty \gamma_s \, da_s = \infty$ a.s.,
(2) *there is a positive definite matrix C such that for all $x \in R^d \backslash \{0\}$, $\big(Cx, R(x)\big) < 0$,*
(3) *there are a function $L \in \Delta$ and a predictable positive process c_s such that* a.s. *for all $x \in R^d$*

$$\operatorname{tr} \sigma(t,x)\beta_t\sigma^*(t,x) \leq c_t L\big((Cx,x)\big) \quad \text{and} \quad \int_0^\infty c_t\gamma_t^2 \, dV_t < \infty.$$

Then $\mathbf{P}\big\{\lim_{t\to\infty} X_t = 0\big\} = 1$.

To prove the theorem we need the next lemmas. In what follows all the constants will be denoted by H or $H(\cdot)$.

Lemma 1.
(1) *Any function $L \in \Delta$ is well defined, positive, increasing and continuously differentiable.*
(2) *For any $\delta > 0$ there is ε such that for $x > \delta$ one has $L_{\varepsilon,m}(x) - 2xL'_{\varepsilon,m}(x) < 0$.*
(3) *For any $\varepsilon > 0$ and $m = 0, 1, \ldots,$ there is $H = H(\varepsilon, m)$ such that for $x > 0$ one has $xL'_{\varepsilon,m}(x) \leq HL_{\varepsilon,m}(x)$.*
In particular, for $\varepsilon = e_{m+1} - e_m$ one can choose $H(\varepsilon, m) = m + 1$.
(4) *For any $\lambda > 0$ there is $H = H(\lambda, \varepsilon)$ such that $L_{\varepsilon,m}(\lambda x) \leq H\big(L_{\varepsilon,m}(x) + 1\big)$.*

We shall not present a proof of the (analytic) Lemma 1. However, using the properties of $L = L_{\varepsilon,m} \in \Delta$ we note that the function

$$W(x) = W_{\varepsilon,m} = \int_0^{(Cx,x)} L_{\varepsilon,m}^{-1}(u)\,du$$

has two continuous partial derivatives and $W(x) \to \infty$ if and only if $\|x\| \to \infty$.

Lemma 2. *Let $Z_t = Z_0 + A_t + M_t$ be a positive continuous semimartingale with the martingale part M and $\mathbf{E}(Z_0) < \infty$, $A_t \le A_t^1 - A_t^2$, where A^1 and A^2 are continuous increasing processes. Then* (a.s.) $\{A_\infty^1 < \infty\} \subseteq \{Z \to\} \cap \{A_\infty^2 < \infty\}$.

Here $\{\omega : Z \to\}$ is the set of convergence of $Z_t(\omega)$ to a finite limit as $t \to \infty$.
Proof. Since $\mathbf{E}(Z_0) < \infty$, it follows from Theorem 7 [3; p. 115] that (a.s.)

$$\{A_\infty^1 < \infty\} \subseteq \{Z \to\} \cap \{(A^1 - A)_\infty < \infty\}.$$

Moreover, since $A_t \le A_t^1 - A_t^2$ we obtain (a.s.)

$$\{(A^1 - A)_\infty < \infty\} \subset \{A_\infty^2 < \infty\},$$

yielding the desired inclusion.

Proof of Theorem 1. Let us apply Ito's formula (see [4]) to $W_{\varepsilon,m}$, giving

$$\begin{aligned} W(X_t) = W(X_0) &+ 2\int_0^t \gamma_s L^{-1}(CX_s, X_s)(CX_s, R(X_s)\,da_s) \\ &+ \tfrac{1}{2}\operatorname{tr}\int_0^t \gamma_s^2 L_\varepsilon^{-2}((CX_s,X_s))\Big[L_\varepsilon((CX_s,X_s))C \\ &- (CX_s,X_s)'((CX_s,X_s)')^* \cdot L_\varepsilon'(CX_s,X_s)\Big]\sigma(s,X_s)\beta_s\sigma^*(s,X_s)\,dV_s \\ &+ 2\int_0^t \gamma_s L_\varepsilon^{-1}((CX_s,X_s))(CX_s, \sigma(s,X_s)\,dm_s). \end{aligned}$$

The third term here is denoted by I. We note that $\operatorname{tr}(AB) \le \|A\| \operatorname{tr} B$ if the matrices A and B are symmetrical and $B \ge 0$. Using this fact and the properties of L_ε from Lemma 1 we obtain

$$\begin{aligned} I &\le H(C)\int_0^t \gamma_s^2 L_\varepsilon^{-2}\Big[L_\varepsilon((CX_s,X_s)) + (CX_s,X_s)L_\varepsilon'((CX_s,X_s))\Big]\operatorname{tr}\sigma_s\beta_s\sigma_s^*\,dV_s \\ &\le H\int_0^t \gamma_s^2 L_\varepsilon^{-1}(\cdot)\operatorname{tr}\sigma_s\beta_s\sigma_s^*\,dV_s. \end{aligned}$$

As a result we get

$$
\begin{aligned}
W(X_t) \le W(X_0) &+ \int_0^t \Big[2\gamma_s\alpha_s(CX_s, R(X_s)) + H\gamma_s^2 \operatorname{tr}\beta\sigma^*\Big] L^{-1} dV_s \\
&+ 2\int_0^t \gamma_s L^{-1}(CX_s, \sigma_s\, dm_s).
\end{aligned} \tag{4}
$$

Note that $\mathbf{E}\|X_0\|^2 < \infty$ implies $\mathbf{E}W(X_0) < \infty$. Thus, the inequality (4) gives us a possibility to apply Lemma 2 with

$$
Z_t = W(X_t), \qquad A_t^2 = 2\int_0^t \gamma_s\alpha_s L_\varepsilon^{-1}(CX_s, R(X_s, R(X_s)))\, dV_s,
$$

$$
A_t^1 = 2H\int_0^t \gamma_s^2 \operatorname{tr}\sigma\beta\sigma^*\, dV_s, \qquad M_t = 2\int_0^t \gamma_s L^{-1}(CX_s, \sigma_s\, dm_s).
$$

Therefore, we have (a.s.)

$$
\{A_\infty^1 < \infty\} \subseteq \{W(X_t) \to \,\} \cap \{A_\infty^2 < \infty\}.
$$

Using now the assumptions of the theorem we have (a.s.)

$$
A_\infty^1 \le H\int_0^\infty \gamma_s^2 L^{-1} c_s L\, dV_s = H\int_0^\infty \gamma_s^2 c_s\, dV_s < \infty,
$$

and therefore (a.s.)

$$
A_\infty^2 < \infty; \tag{5}
$$

$W(X_t)$ tends to a finite limit (a.s.) and

$$
\mathbf{P}\{\, \|X_t\|^2 \to \,\} = 1. \tag{6}
$$

Let us prove that $\|X_t\|^2 \to 0$ (a.s.) as $t \to \infty$. Assume that $\|X_t\|^2(\omega) \to b = b(\omega)$ and $b(\omega) \neq 0$ on a set Ω_1 with

$$
\mathbf{P}(\Omega_1) = \alpha_1 > 0. \tag{7}
$$

Lemma 3. *On the set* Ω_1

$$
y = \liminf_{t\to\infty} \big(-CX_t, R(X_t)\big) > 0. \tag{8}
$$

Proof. Suppose that $\liminf\big(-CX_t, R(X_t)\big) = 0$ on a subset $\Omega_2 \subseteq \Omega_1$ with $\mathbf{P}(\Omega_2) = \alpha_2 > 0$. Fix $\omega \in \Omega_2$ and construct $t_n = t_n(\omega) \to \infty$ $(n \to \infty)$ such that $\big(CX_{t_n}, R(X_{t_n})\big)(\omega) \to 0$ $(n \to \infty)$. The sequence $\big(X_{t_n}(\omega)\big)_{n\geq 1}$ is bounded because of (6), and we can take a subsequence $X_{t'_n}(\omega)$ which converges to $\xi(\omega)$. Using continuity of the scalar product and the function R we get, for $\omega \in \Omega_2$, that

$$\big(-CX_{t'_n}, R(X_{t'_n})\big)(\omega) \to \big(-C\xi, R(\xi)\big)(\omega) = 0.$$

This statement contradicts the second assumption of Theorem 1 because, in view of (7), $\xi \equiv 0$. Thus (8) is true, finishing the proof of Lemma 3.

It follows from (8) that for $\omega \in \Omega_1$ there is $T = T(\omega) > 0$ such that

$$-L^{-1}\big((CX_t, X_t)\big)\big(CX_t, R(X_t)\big) > \operatorname{const} y(\omega) > 0 \qquad \text{for } t \geq T(\omega).$$

Therefore, using the first condition of the theorem we have on Ω_1

$$\begin{aligned} -\int_{T(\omega)}^{\infty} \gamma_s \alpha_s L^{-1}(CX_s, X_s)\big(CX_s, R(X_s)\big)\, dV_s &= -\int_{T(\omega)}^{\infty} \gamma_s L^{-1}\big(CX_s, R(X_s)\big)\, da_s \\ &\geq \operatorname{const} y(\omega) \int_{T(\omega)}^{\infty} \gamma_s\, da_s = \infty. \end{aligned}$$

3. We consider now the one-dimensional $(d = k = 1)$ model (1) in order to obtain a result on a.s. convergence of X_t as $t \to \infty$ under wider assumptions.

Theorem 2. *Let the first and the second* $(C \equiv 1)$ *conditions of Theorem* 1 *be fulfilled. Moreover, suppose there exist a real number* $\delta > 0$ *and a predictable function* Δ_t *such that for all* $|x| \leq \delta$ *and* $t > 0$

$$(9) \qquad 2\gamma_t x R(x)\alpha_t + \gamma_t^2 \beta_t^2 \sigma^2(t, x) \leq \Delta_t, \qquad \text{where} \int_0^{\infty} \Delta_s\, dV_s < \infty.$$

Then the continuous strong solution X_t, $t \geq 0$, *of the equation* (1) *satisfies* $\mathbf{P}\{X_t \to 0\} = 1$ *as* $t \to \infty$.

Proof. Choose $\varepsilon = \varepsilon(\delta) > 0$, construct the functions L_ε and W_ε, and apply

Ito's formula. We then have

$$\begin{aligned}
W_\varepsilon(X_t) &= W_\varepsilon(X_0) + \int_0^t 2\gamma_s L_\varepsilon^{-1}(X_s^2) X_s R(X_s) I_{\{|X_s|\le\delta\}}\, da_s \\
&\quad + \int_0^t 2\gamma_s L_\varepsilon^{-1}(X_s^2) X_s R(X_s) I_{\{|X_s|>\delta\}}\, da_s \\
&\quad + \int_0^t \gamma_s^2 L_\varepsilon^{-2}(X_s^2)\big[L_\varepsilon(X_s^2) - 2X_s^2 L_\varepsilon'(X_s^2)\big] I_{\{|X_s|>\delta\}} \sigma^2(s, X_s)\, d\langle m\rangle_s \\
&\quad + \int_0^t \gamma_s^2 L_\varepsilon^{-2}(X_s^2)\big[-2X_s^2 L_\varepsilon'(X_s^2)\big] I_{\{|X_s|\le\delta\}} \sigma(s, X_s^2)\, d\langle m\rangle_s \\
&\quad + \int_0^t \gamma_s^2 L_\varepsilon^{-1}(X_s^2) I_{\{|X_s|\le\delta\}} \sigma^2(s, X_s)\, d\langle m\rangle_s \\
&\quad + \int_0^t \gamma_s L_\varepsilon^{-1}(X_s^2) 2X_s \sigma(s, X_s)\, dm \\
&= W_\varepsilon(X_0) - A_t^2 + A_t^1 + M_t,
\end{aligned}$$

where $-A^2$ is the sum of the 3rd, 4th and 5th terms, A^1 is the sum of the 2nd and 6th terms and, respectively, M is the last term.

We choose $\varepsilon = \varepsilon(\delta)$ such that (see Lemma 1)

$$L_\varepsilon^{-2}(x^2)\big[L_\varepsilon(x^2) - 2x^2 L_\varepsilon'(x^2)\big] I_{\{|x|>\delta\}} < 0.$$

Therefore A^2 is an increasing positive process. Applying Lemma 2 to $Z_t = W_\varepsilon(X_t)$ we get (a.s.)

$$\{A_\infty^1 < \infty\} \subseteq \{W_\varepsilon(X_t) \to\ \} \cap \{A_\infty^2 < \infty\}. \tag{10}$$

Note that using (9) and Lemma 1 one can show that (a.s.)

$$\begin{aligned}
A_\infty^1 &= \int_0^\infty \big[2\gamma_s X_s R(X_s) + \gamma_s^2 \beta_s \sigma^2(X_s)\big] L_\varepsilon^{-1}(X_s^2) I_{\{|X_s|\le\delta\}}\, dV_s \\
&\le L_\varepsilon^{-1}(0) \int_0^\infty \Delta_s\, dV_s < \infty.
\end{aligned}$$

Therefore, it follows from (10) that $W_\varepsilon(X_t)$ converges (a.s.) to a finite limit as $t \to \infty$. Now, the same method as used at the end of the proof of Theorem 1 can be applied to finish the proof of Theorem 2.

4. Example 1. The example shows that the second condition of Theorem 1 is weaker than the "usual" condition $\big(x, R(x)\big) < 0$.

Let us take the following positive definite matrix

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 8 \end{pmatrix}, \qquad \text{and } R(x) = \begin{pmatrix} -x_1 + 8x_2 \\ -x_2 - x_1 \end{pmatrix}, \qquad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Then for $x_1 = x_2 = 1$ we have $\big(x, R(x)\big) = 5 > 0$. But $\big(Cx, R(x)\big) = -x_1^2 - 8x_2^2 < 0$ for all $x \neq 0$.

Example 2. Consider a one-dimensional model (1) with $a_t \equiv t$, $m_t \equiv W_t$ (Wiener process), $\gamma_t > 0$, $\sigma(t,x) = \gamma_t^{-1/2}(-2xR(x))^{1/2}$; $t > 0$, $x \in R^1$. Let $\int_0^\infty \gamma_s \, ds = \infty$ and $\int_0^\infty \gamma_s^2 \, ds = \infty$ (for example, $\gamma_t = const!$). In this case we have the trivial condition (9):

$$2\gamma_t x R(x) + \gamma_t^2 \sigma(t,x) \equiv 0.$$

Using Theorem 2 we get $x_t \to 0$ (a.s.) as $t \to \infty$. As a result we have the a.s. convergence of X_t without the classical condition $\int_0^\infty \gamma_s^2 \, ds < \infty$ (see [2]).

Remark 1. The method of this paper has been used to investigate Kiefer–Wolfowitz procedures for semimartingales (see [1]–[2]). This paper deals with the continuous model (1) only. However, the discontinuous case can be treated too, and we are going to carry it out in another paper.

Remark 2. Our conditions 2)–3) of the theorems are quite similar to the so-called "monotony conditions" from [5], where the corresponding existence and uniqueness results are proved for strong solutions of stochastic equations with respect to continuous semimartingales.

References

[1] MELNIKOV, A.V.: Stochastic approximation procedures for semimartingales. - In Statistics and Control of Stochastic Processes, edited by A.N. Shiryaev, Nauka, Moscow, 1989, 147–156 (Russian).

[2] NEVELSON, M.B., and R.Z. KHASMINSKII: Stochastic approximation and recurrent estimation. - Nauka, Moscow, 1972 (Russian).

[3] LIPTSER, R.SH., and A.N. SHIRYAEV: Theory of martingales. - Nauka, Moscow, 1986 (Russian).

[4] GIHMAN, I.I., and A.V. SKOROHOD: Stochastic differential equations and their applications. - Naukova Dumka, Kiev, 1982 (Russian).

[5] KRYLOV, N.V., and B.L. ROZOVSKII: On evolution stochastic equations. - In Modern Problems in Mathematics, VINITI, Moscow, 1979, 71–146 (Russian).