Stochastic Processes and their Applications 50 (1994) 101–115
North-Holland

About Gaussian schemes in stochastic approximation

Alain Le Breton

L.M.C./I.M.A.G., Université Joseph Fourier, Grenoble, France

Received 25 February 1992
Revised 5 March 1993

A family of one-dimensional linear stochastic approximation procedures in continuous time where processes of errors are Gaussian martingales is considered. Under some general assumptions the asymptotic behaviour of these procedures is studied concerning strong consistency, rate of convergence and limiting law of involved estimates and costs. At first some asymptotic results for Gaussian martingales, associated quadratic functionals and functions with finite variation are discussed.

asymptotic normality * estimate * Gaussian martingale * rate of convergence * stochastic approximation * strong consistency * Wiener process

1. Introduction

Starting from the paper of Robbins and Monro (1951), many works have been devoted to the study of the asymptotic behaviour of stochastic approximation schemes similar to their procedure. The literature dealing with discrete time models is abundant; one can consult e.g. the articles by Lai and Robbins (1979, 1981), Kushner and Huang (1981), Wei (1987), Kushner and Yin (1987a, 1987b), Polyak (1990), Yin (1991) and also the books of Nevelson and Hasminskii (1973), Kushner and Clark (1978), Hall and Heyde (1980), Benveniste et al. (1987) and Duflo (1990). As far as we know the continuous time context has been less investigated; nevertheless Driml and Nedoma (1967), Melnikov (1989) and Melnikov and Rodkina (1992) provide some substantial material. In the present paper, we intend to develop a complete self-contained analysis of a family of scalar linear models in continuous time where the processes of errors are Gaussian martingales.

A basic probability space $(\Omega, \mathscr{F}, P)$ endowed with some filtration $(\mathscr{F}_t, t \geqslant 0)$ satisfying usual conditions is given on which all random variables and processes below are defined (for details on filtrations, semimartingales and stochastic integration we refer to Dellacherie and Meyer, 1980, or Jacod and Shiryaev, 1987). We consider the linear regression model

$$\mathrm{d}Y_t = \beta(\theta - X_t)\,\mathrm{d}V_t + \mathrm{d}m_t, \quad t \geqslant 0, \qquad Y_0 = 0. \tag{1}$$

Correspondence to: Dr. Alain Le Breton, L.M.C./I.M.A.G., Université Joseph Fourier, B.P. 53X, F-38041 Grenoble, France.

SSDI 0304-4149(93)E0024-9

Here the observed response process Y is real-valued, the process of errors m is a càdlàg real Gaussian $(\mathscr{F}_t)$-martingale $(m_0=0)$ and the scalar regressor process X is $(\mathscr{F}_t)$-predictable. We assume that V is a given càdlàg positive increasing deterministic function ($V_0=0$) such that $\langle m\rangle_t = Em_t^2 = \sigma^2 V_t$ with $\sigma>0$.

A simple example is when $m=\sigma W$ where W is a $(\mathscr{F}_t)$-standard Wiener process and $V_t=t$. Notice also that one can embed into model (1) the usual discrete time model

$$y_n = \beta(\theta - X_n) + \varepsilon_n\,, \quad n=1, 2, \dots,$$

where (ε_n) are i.i.d. Gaussian random variables with mean zero and variance σ^2. For that one sets $V_t=[t]$, $Y_t=\Sigma_{0<k\leqslant t} y_k$, $m_t=\Sigma_{0<k\leqslant t}\varepsilon_k$, $t\geqslant 0$, $\mathscr{F}_t=\sigma\{\varepsilon_k, 0<k\leqslant t\}$. The problem is to choose the levels X_s, $s\leqslant t$, and an estimate θ_t of the unknown parameter θ in view of the observation of Y up to time t in such a way that θ_t converges almost surely (a.s.) to θ as t tends to infinity while the cost $\int_0^t (X_s-\theta)^2\,\mathrm{d}V_s$ of the design X up to stage t is as small as possible. Notice that the parameter β in (1) may be also unknown. Of course we assume that $\beta\neq 0$. We follow the ideas of Lai and Robbins (1979) where a good medical motivation and a quite complete investigation were given for a discrete time general model.

Suppose that β is in fact *known*. To estimate θ at stage t, provided V_t is positive, it is natural to use the least squares estimator

$$\theta_t = V_t^{-1}\left[\int_0^t X_s\,\mathrm{d}V_s + \beta^{-1}Y_t\right]. \tag{2}$$

Then, due to (1), one gets

$$\theta_t - \theta = (\beta V_t)^{-1} m_t\,. \tag{3}$$

That equation shows that irrespective of how the levels X_t are fixed, whether preassigned or recursively chosen, the distribution of $V_t^{1/2}(\theta_t-\theta)$ is Gaussian with mean zero and variance of σ^2/β^2 and therefore, if $V_\infty=+\infty$, then θ_t is consistent in the mean square sense.

Notice that the estimate θ_t can be produced recursively. From (2), by using the rules of stochastic integration, for t_0 such that $V_{t_0}>0$ and $t>t_0$, we get

$$\mathrm{d}\theta_t = -[\mathrm{d}V_t/(V_tV_{t-})]\left[\int_0^{\cdot} X_s\,\mathrm{d}V_s + \beta^{-1}Y\right]_{t-} + V_t^{-1}[X_t\,\mathrm{d}V_t + \beta^{-1}\,\mathrm{d}Y_t]\,,$$

i.e.

$$\mathrm{d}\theta_t = V_t^{-1}[(X_t-\theta_{t-})\,\mathrm{d}V_t + \beta^{-1}\,\mathrm{d}Y_t]\,. \tag{4}$$

In order to have small cost, one is naturally led to choose the design X adaptively as $X=\theta_-$. Consequently, due to (4), we get $\mathrm{d}\theta_t = (\beta V_t)^{-1}\,\mathrm{d}Y_t$ and, taking (1) into account,

$$\mathrm{d}\theta_t = V_t^{-1}(\theta-\theta_{t-})\,\mathrm{d}V_t + (\beta V_t)^{-1}\,\mathrm{d}m_t\,. \tag{5}$$

A statement describing the asymptotic behaviour of this particular scheme has been given without proof in Le Breton (1992).

Now, when β is *unknown*, following Lai and Robbins (1979) approach, one can think of substituting for β some guess $b \neq 0$ of its value in the recursion defined by $X_t = \theta_{t-}$ and (2). Notice that, setting $a = b^{-1}\beta$, this corresponds to choose in (1) the design X adaptively as $X_t = X_t^a = \theta_{t-}^a$ with

$$\theta_t^a = V_t^{-1}\left[\int_0^t X_s^a \, dV_s + b^{-1} Y_t\right],$$

and therefore again $d\theta_t^a = (bV_t)^{-1} \, dY_t$ or

$$d\theta_t^a = aV_t^{-1}(\theta - \theta_{t-}^a) \, dV_t + (bV_t)^{-1} \, dm_t. \tag{6}$$

Of course (6) reduces to (5) when $a = 1$ i.e. b equals β. It appears as an analogue in continuous time of the linear version of the Robbins–Monro algorithm (see e.g. Hall and Heyde, 1980).

Our aim is to provide a complete asymptotic study of the approximation procedure defined by (6), when a is positive (i.e. the sign of b is the same as that of β), concerning strong consistency, rate of convergence and limiting law of the involved estimate θ^a and cost $C_t(a) = \int_0^t (\theta_{s-}^a - \theta)^2 \, dV_s$. At first, in Section 2, we discuss some preliminaries and auxiliary results about Gaussian martingales and functions with finite variation which we think are themselves of independent interest. Then the main theorems are stated and proved in Section 3.

2. Preliminaries and auxiliary results

In what follows $N = (N_t, t \geqslant 0)$ is a Gaussian martingale means that N is a càdlàg $(\mathscr{F}_t)$-adapted real-valued process such that $N_0 = 0$ and for all $0 \leqslant s \leqslant t$ the difference $N_t - N_s$ is a zero-mean Gaussian variable independent from $\mathscr{F}_s$ (i.e. N is a process with independent Gaussian increments on $(\Omega, \mathscr{F}, (\mathscr{F}_t), P)$ in the sense of Jacod and Shiryaev, 1987). Then the distribution of N is entirely characterized by the càdlàg positive increasing deterministic function of $\langle N \rangle$ defined as $\langle N \rangle_t = EN_t^2$, $t \geqslant 0$.

At first we state a law of iterated logarithm (LIL) (see Le Breton, 1992) which sharply defines the strong law of large numbers i.e. $\lim_{t\to+\infty} \langle N \rangle_t^{-1} N_t = 0$ a.s. if $\langle N \rangle_\infty = +\infty$.

Lemma 1. *Let N be a Gaussian martingale such that $\langle N \rangle_\infty = +\infty$. Then the following assertion holds*:

(LIL) $\quad \limsup_{t\to+\infty} [2\langle N \rangle_t \log\log \langle N \rangle_t]^{-1/2} |N_t| \leqslant 1 \quad$ *a.s.*

If moreover $\langle N \rangle_t^{-1} \Delta\langle N \rangle_t$ converges to zero as t goes to infinity, then equality holds in (LIL) *i.e.*

(LIL*) $\quad \limsup_{t\to+\infty} [2\langle N \rangle_t \log\log \langle N \rangle_t]^{-1/2} |N_t| = 1 \quad$ *a.s.* $\quad \square$

Later on, for any càdlàg process (possibly a deterministic function) $U = (U_t, t \geqslant 0)$,

$U_0=0$, for $t>0$ we set $U_{t-}=\lim_{s\uparrow t}U_s$ and $\Delta U_t=U_t-U_{t-}$. The set $J(u)=\{t>0: P(\Delta U_t\neq 0)>0\}$ is at most countable. Notice that for a Gaussian martingale N the equality $J(N)=J(\langle N\rangle)$ holds. For $K\subset]0,+\infty[$ we write $\sum_{s\in K}\Delta U_s$ instead of $\sum_{s\in K\cap J(U)}\Delta U_s$. The continuous part U^c of U is defined as $U^c=U-\sum_{s\leqslant t}\Delta U_s$. When U is a càdlàg positive increasing deterministic function such that $U_{t_0}>0$, (C_U) denotes the following condition

$$(\mathrm{C}_U)\qquad \sum_{t_0<s<+\infty}[U_s^{-1}\Delta U_s]^2<+\infty .$$

Now we investigate the a.s. asymptotic behaviour of some quadratic functional of a Gaussian martingale.

Proposition 1. *Let N be a Gaussian martingale. Assume that $\langle N\rangle_\infty=+\infty$ $(\langle N\rangle_{t_0}>0)$ and that condition $(\mathrm{C}_{\langle N\rangle})$ is fulfilled. Let φ be a measurable deterministic real-valued function defined on $]t_0,+\infty[$ such that $\langle N\rangle^2\varphi$ tends to one as t goes to infinity. Then the following assertion holds*:

$$(\mathrm{S1})\qquad \lim_{t\to+\infty}[\log\langle N\rangle_t]^{-1}\int_{t_0}^t\varphi_s N_{s-}^2\,\mathrm{d}\langle N\rangle_s=1\quad a.s.$$

Moreover in statement (S1) *one may substitute N_s^2 to N_{s-}^2 in the integral.*

Proof. At first let $\varphi=\varphi^*=[\langle N\rangle\langle N\rangle_-]^{-1}$. Then statement (S1) is proved in Le Breton (1992). Let us show that here one may substitute N_s^2 to N_{s-}^2 in the integral. In fact we shall prove that the difference

$$\begin{aligned}
&\int_{t_0}^t\varphi_s^*N_s^2\,\mathrm{d}\langle N\rangle_s-\int_{t_0}^t\varphi_s^*N_{s-}^2\,\mathrm{d}\langle N\rangle_s\\
&=\int_{t_0}^t 2\varphi_s^*N_{s-}\,\Delta N_s\,\mathrm{d}\langle N\rangle_s+\int_{t_0}^t\varphi_s^*(\Delta N_s)^2\,\mathrm{d}\langle N\rangle_s\\
&=2\sum_{t_0<s\leqslant t}\varphi_s^*N_{s-}\,\Delta\langle N\rangle_s\,\Delta N_s\\
&\quad+\sum_{t_0<s\leqslant t}\varphi_s^*\,\Delta\langle N\rangle_s[(\Delta N_s)^2-\Delta\langle N\rangle_s]+\sum_{t_0<s\leqslant t}\varphi_s^*(\Delta\langle N\rangle_s)^2,
\end{aligned}$$

converges a.s. to a finite limit. Notice that, due to $(\mathrm{C}_{\langle N\rangle})$, we have $\langle N\rangle_s\approx\langle N\rangle_{s-}$ for s tending to infinity and then also $\varphi_s^*\approx\langle N\rangle_s^{-2}$. Therefore, the series $\sum_{0<s\leqslant\cdot}\varphi_s^*(\Delta\langle N\rangle_s)^2$ converges. Finally, in order to show that (S1) holds, it suffices to prove that processes

$$\sum_{t_0<s\leqslant\cdot}\varphi_s^*N_{s-}\,\Delta\langle N\rangle_s\,\Delta N_s\quad\text{and}\quad\sum_{t_0<s\leqslant\cdot}\varphi_s^*\,\Delta\langle N\rangle_s[(\Delta N_s)^2-\Delta\langle N\rangle_s]\ ,$$

both converge a.s. to finite limits. But these processes are square integrable martingales with respective variances

$$\sum_{t_0<s\leqslant\cdot}(\varphi_s^*)^2\langle N\rangle_{s-}[\Delta\langle N\rangle_s]^3\quad\text{and}\quad 2\sum_{t_0<s\leqslant\cdot}(\varphi_s^*)^2[\Delta\langle N\rangle_s]^4\ .$$

Hence, since $(\varphi_s^*)^2\langle N\rangle_{s-}[\Delta\langle N\rangle_s]^3 \approx \langle N\rangle_s^{-3}[\Delta\langle N\rangle_s]^3$ and $(\varphi_s^*)^2[\Delta\langle N\rangle_s]^4 \approx \langle N\rangle_s^{-4}[\Delta\langle N\rangle_s]^4$, they converge a.s. to finite limits.

Now let φ be any measurable deterministic function φ such that $\langle N\rangle^2\varphi$ tends to one as t goes to infinity. Then $\varphi\varphi^*$ tends also to one. Therefore applying the Tœplitz Lemma we get that (S1) holds with either N_s^2 or N_{s-}^2 in the integral. $\square$

Concerning convergence in distribution we state:

Proposition 2. *Let N be a real Gaussian martingale. Assume that $\langle N\rangle_\infty = +\infty$ ($\langle N\rangle_{t_0} > 0$) and that condition ($C_{\langle N\rangle}$) is fulfilled. Let ϕ be a measurable deterministic real-valued function defined on $]t_0, +\infty[$ and tending to one as t goes to infinity. Then the following assertion holds*:

$$\text{(S2)} \qquad \langle N\rangle_t^{-2}\int_{t_0}^t \phi_s N_s^2\,\mathrm{d}\langle N\rangle_s \xrightarrow{\mathscr{L}} \int_0^1 W_s^2\,\mathrm{d}s \quad \textit{as } t \textit{ tends to infinity}$$

where $W=(W_t;\, t\in[0,1])$ is a standard Wiener process. Moreover in statement (S2) *one may substitute N_{s-}^2 to N_s^2 in the integral.*

Proof. At first let $\phi\equiv 1$. Then statement (S2) is proved in Le Breton (1992). Let us show that here one may substitute N_{s-}^2 to N_s^2 in the integral. It suffices to check that the quantity $\langle N\rangle_t^{-2}[\int_{t_0}^t N_s^2\,\mathrm{d}\langle N\rangle_s - \int_{t_0}^t N_{s-}^2\,\mathrm{d}\langle N\rangle_s]$ tends a.s. to zero as t tends to infinity. Due to the Kronecker Lemma, this holds provided that the difference $\int_{t_0}^t \langle N\rangle_s^{-2}N_s^2\,\mathrm{d}\langle N\rangle_s - \int_{t_0}^t\langle N\rangle_s^{-2}N_{s-}^2\,\mathrm{d}\langle N\rangle_s$ tends a.s. to a finite limit. But repeating the arguments of the proof of Proposition 1 with φ^* replaced by $\langle N\rangle^{-2}$ we get the result.

Now let ϕ be any measurable deterministic function tending to one as t goes to infinity. We shall show that the quantity

$$J(t) = E\left\{\langle N\rangle_t^{-2}\left|\int_{t_0}^t \phi_s N_s^2\,\mathrm{d}\langle N\rangle_s - \int_{t_0}^t N_s^2\,\mathrm{d}\langle N\rangle_s\right|\right\},$$

tends to zero. Clearly we have

$$J(t) \leqslant \langle N\rangle_t^{-2}\int_{t_0}^t |\phi_s - 1|\,\langle N\rangle_s\,\mathrm{d}\langle N\rangle_s\,.$$

Since $\mathrm{d}(\langle N\rangle_s^2) = (\langle N\rangle_s + \langle N\rangle_{s-})\,\mathrm{d}\langle N\rangle_s$ we get

$$J(t) \leqslant \tfrac{1}{2}\langle N\rangle_t^{-2}\int_{t_0}^t |\phi_s - 1|\,\mathrm{d}(\langle N\rangle_s^2) + \langle N\rangle_t^{-2}\sum_{t_0<s\leqslant t}|\phi_s - 1|\,(\Delta\langle N\rangle_s)^2\,.$$

The first term in the right hand side above converges a.s. to zero due to the Tœplitz Lemma.

The second term tends also to zero due to the Kronecker Lemma since the series $\sum_{t_0<s\leq \cdot}\langle N\rangle_s^{-2}|\phi_s-1|(\Delta\langle N\rangle_s)^2$ converges a.s. Finally $J(t)$ tends to zero so that the variable $\langle N\rangle_t^{-2}\int_{t_0}^t \phi_s N_s^2\,\mathrm{d}\langle N\rangle_s$ has the same limiting distribution as $\langle N\rangle_t^{-2}\int_{t_0}^t N_s^2\,\mathrm{d}\langle N\rangle_s$ and (S2) holds. It can be proved in a similar way that one may substitute N_{s-}^2 to N_s^2 in the integral. □

Now we investigate asymptotic properties of some functions with finite variation. Here V is a càdlàg positive increasing deterministic function ($V_0=0$, $V_{t_0}>0$). For any real number α we define the function L_α by

$$L_{\alpha,t}=\int_{t_0}^t V_s^{-\alpha}\,\mathrm{d}V_s\,,\quad t\geq t_0\,, \tag{7}$$

and we set for $v>0$,

$$F_\alpha(v)=\begin{cases}\log v & \text{if } \alpha=1\,,\\ v^{1-\alpha}/(1-\alpha) & \text{if } \alpha\neq 1\,.\end{cases} \tag{8}$$

Clearly if $\alpha=0$ then

$$L_{0,t}=V_t-V_{t_0}=F_0(V_t)-F_0(V_{t_0})\,.$$

We are going to study the asymptotic behaviour of the function L_α when $\alpha\neq 0$. At first let us consider the function δ_α defined as

$$\delta_\alpha=(\operatorname{sgn}\alpha)[F_\alpha(V)-F_\alpha(V_{t_0})-L_\alpha]\,. \tag{9}$$

We state:

Lemma 2. *Let V be a càdlàg positive increasing deterministic function ($V_0=0$, $V_{t_0}>0$) and α be an arbitrary real number. Let L_α, F_α and δ_α be defined respectively by (7), (8) and (9). Then the following assertions hold:*

(i) *The function δ_α is purely discontinuous positive and increasing; more precisely one can write $\delta_\alpha=\sum_{t_0<s\leq\cdot}\Delta\delta_{\alpha,s}$ where the jumps satisfy*

$$\begin{aligned}0\leq\Delta\delta_{\alpha,s}&=(\operatorname{sgn}\alpha)[F_\alpha(V_s)-F_\alpha(V_{s-})-V_s^{-\alpha}\Delta V_s]\\ &\leq(\operatorname{sgn}\alpha)(V_{s-}^{-\alpha}-V_s^{-\alpha})\Delta V_s\,.\end{aligned} \tag{10}$$

(ii) *If the condition (C_V) is fulfilled, then the series $\sum_{t_0<s<+\infty}V_s^{\alpha-1}\Delta\delta_{\alpha,s}$ converges.*

Proof. At first we prove assertion (i). Notice that $\dot F_\alpha(v)=v^{-\alpha}$ and $L_\alpha=\int_{t_0}^{\cdot}\dot F_\alpha(V_s)\,\mathrm{d}V_s$. Then applying the change of variable formula (see e.g. Dellacherie and Meyer, 1980, p. 171) we get

$$F_\alpha(V)-F_\alpha(V_{t_0})-L_\alpha=\sum_{t_0<s\leq\cdot}[F_\alpha(V_s)-F_\alpha(V_{s-})-V_s^{-\alpha}\Delta V_s]\,.$$

It is also easy to check that for every real number α the function $(\operatorname{sgn}\alpha)F_\alpha$ is concave. Therefore it comes that

$$(\operatorname{sgn}\alpha)\dot{F}_\alpha(V_s)\,\Delta V_s \leqslant (\operatorname{sgn}\alpha)[F_\alpha(V_s)-F_\alpha(V_s)] \leqslant (\operatorname{sgn}\alpha)\dot{F}_\alpha(V_{s-})\,\Delta V_s\,,$$

from which inequalities (10) can be obtained immediately. So, assertion (i) holds.

Now we show assertion (ii). Due to (10) we get

$$0 \leqslant V_s^{\alpha-1}\,\Delta\delta_{\alpha,s} \leqslant g(V_s^{-1}\,\Delta V_s)\,,$$

where $g(u)=(\operatorname{sgn}\alpha)u[(1-u)^{-\alpha}-1]$. But, when (C_V) is fulfilled, then $V_s^{-1}\,\Delta V_s$ tends to zero as s tends to infinity. Therefore the statement holds since, for u close to zero, $g(u)$ is equivalent to $|\alpha|u^2$. □

Remark 1. Lemma 2(i) means also that the continuous parts L_α^{c} and $[F_\alpha(V)-F_\alpha(V_{t_0})]^{\mathrm{c}}$ of L_α and $F_\alpha(V)-F_\alpha(V_{t_0})$ respectively coincide. In particular, for $\alpha=1$, we get

$$\begin{aligned} L_1^{\mathrm{c}} = L_1 - \sum_{t_0<s\leqslant\cdot} V_s^{-1}\,\Delta V_s &= \log(V/V_{t_0}) - \sum_{t_0<s\leqslant\cdot} \Delta\log V_s \\ &= [\log(V/V_{t_0})]^{\mathrm{c}}. \end{aligned} \tag{11}$$

Now we prove:

Proposition 3. *Let α be some real number and V, L_α and F_α be as in Lemma 2. Then the following assertions hold:*

(i) *If $\alpha>1$ then the function L_α converges to a finite limit as t tends to infinity.*

(ii) *If $\alpha\leqslant 1$ and conditions $V_\infty=+\infty$ and (C_V) are fulfilled, then*

$$\lim_{t\to+\infty} L_{\alpha,t}/F_\alpha(V_t)=1\,. \tag{12}$$

Proof. If $\alpha>1$ then, due to (9) and Lemma 2(i), we have

$$0\leqslant L_\alpha \leqslant F_\alpha(V)-F_\alpha(V_{t_0}) \leqslant V_{t_0}^{1-\alpha}/(\alpha-1)\,.$$

Therefore, since it is increasing, L_α converges to a finite limit and assertion (i) holds.

Now, let us assume that conditions $V_\infty=+\infty$ and (C_V) are fulfilled. Then, due to Lemma 2(ii), the series $\sum_{t_0<s<+\infty} V_s^{\alpha-1}\,\Delta\delta_{\alpha,s}$ converges.

For $\alpha=1$, taking Lemma 2(i) into account, this simply means that δ_1 converges to a finite limit and hence, due to (9), (12) holds.

For $\alpha<1$, applying the Kronecker Lemma, we get that $V_t^{\alpha-1}\delta_{\alpha,t}$ tends to zero as t tends to infinity and finally, due to (8) and (9), again (12) holds. □

Now for any fixed positive real number a we denote by $(\mathrm{C}_{V,a})$ the condition

$(\mathrm{C}_{V,a})$ $\qquad V_s^{-1}\,\Delta V_s < a^{-1}$ for every $s>t_0$.

Notice that if $a\leqslant 1$ then condition $(\mathrm{C}_{V,a})$ is automatically fulfilled. Moreover when (C_V) is satisfied then, since $V_s^{-1}\,\Delta V_s$ tends to zero, for any $a>1$ one can choose t_0 large enough in order that $(\mathrm{C}_{V,a})$ is fulfilled. Provided $(\mathrm{C}_{V,a})$ is satisfied we can define (cf. Liptser and

Shiryaev, 1978, p. 256) the function ψ_a as the unique positive locally bounded solution of equation

$$\psi_a(t) = 1 - a \int_{t_0}^{t} \psi_a(s-) V_s^{-1} \, dV_s , \quad t \geq t_0 ,$$

i.e.

$$\psi_a(t) = \exp(-aL_{1,t}^{c}) \prod_{t_0 < s \leq t} (1 - a\Delta L_{1,s}) . \tag{13}$$

Here L_1^c is given by (11) and $\Delta L_{1,s} = V_s^{-1} \Delta V_s$. Notice that, due to (11) and (13) with $a = 1$, we have

$$\begin{aligned} \log \psi_1 &= -L_1^c + \sum_{t_0 < s \leq \cdot} \log(1 - V_s^{-1} \Delta V_s) \\ &= -[\log(V/V_{t_0})]^c - \sum_{t_0 < s \leq \cdot} \Delta \log V_s = -\log(V/V_{t_0}) . \end{aligned}$$

Therefore $\psi_1 = (V/V_{t_0})^{-1}$. We are going to study the asymptotic behaviour of the function ψ_a when $a > 0$, $a \neq 1$. At first let us consider the function γ_a defined as

$$\gamma_a = \operatorname{sgn}(1-a)[\log \psi_a + a \log(V/V_{t_0})] . \tag{14}$$

We state:

Lemma 3. *Let V be a càdlàg positive increasing deterministic function ($V_0 = 0$, $V_{t_0} > 0$) and a be a positive real number such that condition ($C_{V,a}$) is fulfilled. Let γ_a be defined by (14) where ψ_a is given by (13). Then the function γ_a is purely discontinuous positive and increasing; more precisely one can write $\gamma_a = \sum_{t_0 < s \leq \cdot} \Delta\gamma_{a,s}$ where the jumps satisfy*

$$0 \leq \Delta\gamma_{a,s} \leq a\{[V_s - (a \vee 1)\Delta V_s]^{-1} - V_s^{-1}\}\Delta V_s \tag{15}$$

with $a \vee 1 = \max(a, 1)$.

Proof. From (11) and (13) it is clear that

$$(\log \psi_a)^c = -aL_1^c = -a[\log(V/V_{t_0})]^c .$$

Therefore, due to (14), γ_a is purely discontinuous. Moreover since

$$(\Delta \log \psi_a)_s = \log(1 - aV_s^{-1}\Delta V_s) \; ; \; \Delta \log V_s = \log V_s - \log V_{s-} ,$$

it follows that

$$\Delta\gamma_{a,s} = \operatorname{sgn}(1-a)[\log(1 - aV_s^{-1}\Delta V_s) + a\Delta \log V_s] \tag{16}$$

or

$$\Delta\gamma_{a,s} = \operatorname{sgn}(1-a)[\log(aV_{s-} + (1-a)V_s) - a\Delta \log V_{s-} - (1-a) \log V_s] . \tag{17}$$

Now assume that $a<1$. Since $u\to\log u$ is concave, equation (17) shows that $\Delta\gamma_{a,s}\geqslant 0$. Moreover, applying (10) with $\alpha=1$ we get

$$0\leqslant \Delta\log V_s - V_s^{-1}\Delta V_s \leqslant (V_sV_{s-})^{-1}(\Delta V_s)^2\,. \tag{18}$$

Then, from (16) and (18) it follows that

$$0\leqslant \Delta\gamma_{a,s}\leqslant \log(1-aV_s^{-1}\Delta V_s)+aV_s^{-1}\Delta V_s+a(V_sV_{s-})^{-1}(\Delta V_s)^2\,.$$

Hence, because $\log(1-u)+u\leqslant 0$ for $u\in[0,1[$, we obtain

$$0\leqslant \Delta\gamma_{a,s}\leqslant a(V_sV_{s-})^{-1}(\Delta V_s)^2=a(V_{s-}^{-1}-V_s^{-1})\Delta V_s\,.$$

So the assertion in the lemma is proved for $a<1$ since here $V_s-(a\vee 1)\Delta V_s=V_{s-}$.

Now assume that $a>1$. Then rewriting (17) as

$$\begin{aligned}\Delta\gamma_{a,s}=-a\,\mathrm{sgn}(1-a)\{&\log V_{s-}-a^{-1}\log[aV_{s-}+(1-a)V_s]\\&-(1-a^{-1})\log V_s\}\,,\end{aligned} \tag{19}$$

and looking at V_{s-} as $V_{s-}=a^{-1}[aV_{s-}+(1-a)V_s]+(1-a^{-1})V_s$, again, due to the concavity of the logarithm and the fact that $-a\,\mathrm{sgn}(1-a)=a$, we get that $\Delta\gamma_{a,s}\geqslant 0$. Now let us rewrite (19) as

$$\begin{aligned}\Delta\gamma_{a,s}=a\{&\log(1-V_s^{-1}\Delta V_s)+V_s^{-1}\Delta V_s\}\\&+\{\log V_s-\log(V_s-a\Delta V_s)-aV_s^{-1}\Delta V_s\}\,.\end{aligned}$$

Then, because $\log(1-u)+u\leqslant 0$ for $u\in[0,1[$, we obtain

$$0\leqslant\Delta\gamma_{a,s}\leqslant\log V_s-\log(V_s-a\Delta V_s)-aV_s^{-1}\Delta V_s\,. \tag{20}$$

But, due to the concavity of the logarithm we have

$$aV_s^{-1}\Delta V_s\leqslant\log V_s-\log(V_s-a\Delta V_s)\leqslant a(V_s-a\Delta V_s)^{-1}\Delta V_s\,.$$

Therefore from (20) we get

$$0\leqslant\Delta\gamma_{a,s}\leqslant a\{[V_s-a\Delta V_s]^{-1}-V_s^{-1}\}\Delta V_s\,.$$

Finally the assertion in the lemma holds if $a>1$. □

Now we prove:

Proposition 4. *Let a be an arbitrary positive real number. Let V be a càdlàg positive increasing deterministic function ($V_0=0$) such that $V_{t_0}>0$ and condition (C_V) is fulfilled (and then also, changing t_0 if necessary, ($C_{V,a}$) is satisfied). Let ψ_a and γ_a be defined by (13) and (14) respectively. Then the following assertions hold*:

(i) *The function γ_a converges to a positive finite limit $\gamma_{a,\infty}$ as t tends to infinity.*

(ii) *If $V_\infty=+\infty$ then*

$$\lim_{t\to+\infty}(V_t/V_{t_0})^a\psi_a(t)=\kappa_a=\exp[\mathrm{sgn}(1-a)\gamma_{a,\infty}]\,.$$

Proof. Due to (15) we get

$$0 \leqslant \Delta\gamma_{a,s} \leqslant h(V_s^{-1}\,\Delta V_s)$$

where $h(u) = au\{[1-(a \vee 1)u]^{-1} - 1\}$. But when (C_V) is fulfilled then $V_s^{-1}\,\Delta V_s$ tends to zero as t tends to infinity. Therefore assertion (i) is true since, for u close to zero, $h(u)$ is equivalent to $a(a \vee 1)u^2$.

Finally assertion (ii) follows immediately from (14). □

3. Main results

Recall that we intend to study the asymptotic behaviour of θ^a satisfying (6) and that of the corresponding cost. In what follows we suppose that some t_0 is chosen such that $V_{t_0} > 0$ and condition $(\mathrm{C}_{V,a})$ holds for $a = b^{-1}\beta$ (which is assumed to be positive). Then, for any given real value θ_0, equation (6) has a unique solution starting at time t_0 from the initial condition θ_0. More precisely, due to (6), we can write $Z = \theta^a - \theta$ as the solution of the stochastic differential equation

$$\mathrm{d}Z_t = -aV_t^{-1}Z_{t-}\,\mathrm{d}V_t + (bV_t)^{-1}\,\mathrm{d}m_t\,, \quad t \geqslant t_0\,, \qquad Z_{t_0} = \theta_0 - \theta\,,$$

i.e. Z is the Gaussian process given by

$$Z_t = \psi_a(t)\left[Z_{t_0} + \int_{t_0}^{t} \psi_a^{-1}(s)(bV_s)^{-1}\,\mathrm{d}m_s\right], \quad t \geqslant t_0\,, \tag{21}$$

where ψ_a is defined by (13).

Our first statement gives the precise asymptotic behaviour of θ^a.

Theorem 1. *Let $a > 0$ and θ_0 an arbitrary real number. Assume that $V_{t_0} > 0$, $V_\infty = +\infty$ and condition (C_V) is fulfilled (and then also, changing t_0 if necessary, $(\mathrm{C}_{V,a})$ is satisfied). Let θ^a be the solution process of equation (6) starting from θ_0 at time t_0. Then the following assertions hold:*

(i) *If $a > \frac{1}{2}$ then*

$$\limsup_{t\to+\infty}\,[V_t/2\log\log V_t]^{1/2}|\theta_t^a - \theta| = (\sigma/|\beta|)g^{1/2}(a) \quad a.s.\,,$$

and

$$V_t^{1/2}(\theta_t^a - \theta) \xrightarrow{\mathscr{L}} \mathrm{N}(0, (\sigma^2/\beta^2)g(a))\,,$$

where $g(a) = a^2/(2a-1)$.

(ii) *If $a = \frac{1}{2}$ then*

$$\limsup_{t\to+\infty}\,[V_t/(2\log V_t)(\log\log\log V_t)]^{1/2}|\theta_t^a - \theta| = \sigma/(2|\beta|) \quad a.s.\,,$$

and

$$(V_t/\log V_t)^{1/2}(\theta_t^a-\theta) \xrightarrow{\mathscr{L}} \mathrm{N}(0, \sigma^2/(4\beta^2)) .$$

(iii) *If* $a<\frac{1}{2}$ *then there exists a finite (Gaussian) random variable* χ *such that*

$$\lim_{t\to+\infty} V_t^a(\theta_t^a-\theta)=\chi \quad a.s.$$

Proof. At first we consider the case $a>\frac{1}{2}$. Due to Proposition 4(ii) we have

$$\psi_a(t)\approx\kappa_a V_{t_0}^a V_t^{-a} . \tag{22}$$

So, in view of (21), we can write

$$V_t^{1/2}(\theta_t^a-\theta)=b^{-1}V_t^{1/2}\psi_a(t)N_t+\mathrm{o}(1) \quad \text{a.s.} \tag{23}$$

where $N=(N_t;\, t\geqslant t_0)$ is the Gaussian martingale defined by

$$N_t=\int_{t_0}^t \psi_a^{-1}(s)V_s^{-1}\,\mathrm{d}m_s\,, \quad t\geqslant t_0 . \tag{24}$$

Notice that

$$\langle N\rangle_t=EN_t^2=\sigma^2\int_{t_0}^t \psi_a^{-2}(s)V_s^{-2}\,\mathrm{d}V_s\,, \quad t\geqslant t_0 . \tag{25}$$

Now, in view of (22) we have $V_s^{-2}\psi_a^{-2}(s)\approx\kappa_a^{-2}V_{t_0}^{-2a}V_s^{-2(1-a)}$ where $2(1-a)<1$. Then, from (25), applying Proposition 3(ii) and the Tœplitz Lemma, we get that

$$\langle N\rangle_t\approx\sigma^2(2a-1)^{-1}\kappa_a^{-2}V_{t_0}^{-2a}V_t^{2a-1}\,, \quad \log\log\langle N\rangle_t\approx\log\log V_t . \tag{26}$$

Therefore we also have

$$|b|^{-1}V_t^{1/2}\psi_a(t)\approx(\sigma/|\beta|)g^{1/2}(a)\langle N\rangle_t^{-1/2} . \tag{27}$$

Using (23), applying (LIL*) (cf. Lemma 1) to N and taking (26) and (27) into account we get that the first part of assertion (i) holds.

Since the distribution of the random variable $b^{-1}V_t^{1/2}\psi_a(t)N_t$ is Gaussian with mean zero and variance $b^{-2}V_t\psi_a^2(t)\langle N\rangle_t$, due to (23) and (27), it follows that the second part of assertion (i) also holds.

Now we assume that $a=\frac{1}{2}$. Here $V_t^{1/2}\psi_{1/2}(t)\approx\kappa_{1/2}V_{t_0}^{1/2}$ and, by use of arguments similar to those above, one can see that (23) and (24) still hold with

$$\langle N\rangle_t\approx\sigma^2\kappa_{1/2}^{-2}V_{t_0}^{-1}\log V_t\,, \qquad \log\log\langle N\rangle_t\approx\log\log\log V_t\,, \tag{26'}$$

and then achieve the proof of assertion (ii).

Finally let $a<\frac{1}{2}$. From (21) we can write

$$V_t^a(\theta_t^a-\theta)=V_t^a\psi_a(t)\{Z_{t_0}+b^{-1}N_t\}\,,$$

where N is again the martingale defined in (24). Here $V_t^a \psi_a(t) \approx \kappa_a V_{t_0}^a$ and N converges a.s. to a finite (Gaussian) random variable. Indeed, since $\psi_a^{-2}(s) V_s^{-2} \approx \kappa_a^{-2} V_{t_0}^{-2a} V_s^{-2(1-a)}$ with $2(1-a) > 1$, due to (25), applying Proposition 3(i) we get that $\langle N \rangle_\infty < \infty$. Therefore assertion (iii) holds. □

Now let us describe the behaviour of the cost associated to algorithm (6) up to time t i.e.

$$C_t(a) = \int_{t_0}^t (\theta_{s-}^a - \theta)^2 \, \mathrm{d}V_s = \int_{t_0}^t Z_{s-}^2 \, \mathrm{d}V_s \,, \tag{28}$$

where Z is given by (21).

Theorem 2. *Let notations and assumptions be the same as in Theorem 1. Then the following assertions hold*:

(i) *If* $a > \frac{1}{2}$ *then*

$$\lim_{t \to +\infty} [\log V_t]^{-1} C_t(a) = (\sigma^2 / \beta^2) g(a) \quad a.s.$$

(ii) *If* $a = \frac{1}{2}$ *then*

$$[\log V_t]^{-2} C_t(a) \xrightarrow{\mathscr{L}} (\sigma^2 / (4\beta^2)) \int_0^1 W_s^2 \, \mathrm{d}s \,,$$

where $W = (W_t; t \in [0, 1])$ *is a standard Wiener process.*

(iii) *If* $a < \frac{1}{2}$ *then*

$$\lim_{t \to +\infty} V_t^{2a-1} C_t(a) = \chi^2 / (1 - 2a) \quad a.s. \,,$$

where χ *is defined in Theorem* 1(iii).

Proof. Due to (21), (24), (25) and (28) we can write

$$C_t(a) = Z_{t_0}^2 K_1(t) + 2b^{-1} Z_{t_0} K_2(t) + b^{-2} \sigma^{-2} K_3(t) \,,$$

where

$$K_1(t) = \int_{t_0}^t \psi_a^2(s-) \, \mathrm{d}V_s \,, \qquad K_2(t) = \int_{t_0}^t \psi_a^2(s-) N_{s-} \, \mathrm{d}V_s \,,$$

$$K_3(t) = \int_{t_0}^t \psi_a^2(s-) \psi_a^2(s) V_s^2 N_{s-}^2 \, \mathrm{d}\langle N \rangle_s \,.$$

From (13) and (22) we observe that for s tending to infinity

$$\psi_a(s-) \approx \psi_a(s) \approx \kappa_a V_{t_0}^a V_s^{-a} \,. \tag{29}$$

In what follows we denote by $\lambda(t)$ the normalizing factor of $C_t(a)$ which is involved in the investigated statement. In the first two cases we shall show that when t goes to infinity, for $i = 1, 2$, $\lambda(t) K_i(t)$ tends to zero a.s. and $\lambda(t) K_3(t)$ tends to the right limit a.s.

At first we consider the case $a>\frac{1}{2}$. Here, due to (29), since $2a>1$, applying Proposition 3(i) we get that $K_1(\infty)<+\infty$ and then $\lambda(t)K_1(t)$ tends to zero. Concerning $K_2(t)$, due to (LIL*) (cf. Lemma 1), we can write a.s.

$$|K_2(t)| \leqslant c \int_{t_0}^{t} \psi_a^2(s-)\langle N\rangle_{s-}^{1/2}(\log\log\langle N\rangle_{s-})\,\mathrm{d}V_s\,, \tag{30}$$

for some positive (random) constant c. But from (25), (26) and (29), we can see that

$$\langle N\rangle_s^{-1}\,\Delta\langle N\rangle_s \approx (2a-1)V_s^{-1}\,\Delta V_s \tag{31}$$

and then in (26) one may replace $\langle N\rangle_t$ by $\langle N\rangle_{t-}$. Therefore, using (29) again, the function appearing in the integral of inequality (30) is, up to a multiplicative constant, equivalent to $V_s^{-(a+1/2)}(\log\log V_s)^{1/2}$. Consequently, since $a+\frac{1}{2}>1$, we have a.s.

$$|K_2(t)| \leqslant C \int_{t_0}^{t} V_s^{-\varepsilon}\,\mathrm{d}V_s\,,$$

for some positive (random) constant C and some $\varepsilon>1$. Again, applying Proposition 3(i), we get that $K_2(\infty)$ is finite a.s. and then $\lambda(t)K_2(t)$ tends to zero a.s. Now we consider $K_3(t)$. We notice that (31) means that condition $(\mathrm{C}_{\langle N\rangle})$ is satisfied since (C_V) is fulfilled and also (26) says that $\lambda^{-1}(t)=\log V_t\approx(2a-1)^{-1}\log\langle N\rangle_t$. Choosing $\varphi_s=\psi_a^2(s-)\psi_a^2(s)V_s^2$ we see that the product $\varphi_s\langle N\rangle_s^2$ tends to $\sigma^4/(2a-1)^2$. Then, applying Proposition 1, it is easy to get that $b^{-2}\sigma^{-2}\lambda(t)K_3(t)$ tends a.s. to $(\sigma^2/\beta^2)g(a)$. Finally assertion (i) holds.

Now we assume that $a=\frac{1}{2}$. Here due to (29), applying Proposition 3(i) with $\alpha=1$, we get that $K_1(t)\approx\kappa_{1/2}^2 V_{t_0}\log V_t$ and then $\lambda(t)K_1(t)$ tends to zero. Concerning $K_2(t)$ we can still write (30) a.s. But from (25), (26′) and (29), we can see that

$$\langle N\rangle_s^{-1}\,\Delta\langle N\rangle_s \approx (V_s\log V_s)^{-1}\Delta V_s \tag{31′}$$

and then in (26′) one may replace $\langle N\rangle_t$ by $\langle N\rangle_{t-}$. Therefore, using (29) again, the function appearing in the integral of inequality (30) is, up to a multiplicative constant, equivalent to $V_s^{-1}(\log V_s)^{1/2}(\log\log V_s)^{1/2}$. Now in order to prove that $\lambda(t)K_2(t)$ tends to zero a.s., due to the Kronecker Lemma, it suffices to show that

$$J=\int_{t_0}^{+\infty} V_s^{-1}(\log V_s)^{-3/2}(\log\log V_{s-})^{1/2}\,\mathrm{d}V_s<+\infty\,.$$

But, defining $a(s)=\inf\{t\colon V_t>s\}$ and using the Lebesgue Lemma on the transformation of Stieltjes integrals (see e.g. Dellacherie and Meyer, 1980), we get that

$$J=\int_{V_{t_0}}^{+\infty} V_{a(s)}^{-1}(\log V_{a(s)})^{-3/2}(\log\log(V_-)_{a(s)})^{1/2}\,\mathrm{d}s<+\infty\,.$$

Then, since $V_{a(s)}\geqslant s$ and $(V_-)_{a(s)}\leqslant s$, we have $J<+\infty$. Now we consider $K_3(t)$. We notice that (31′) means that condition ($C_{\langle N\rangle}$) is satisfied since (C_V) is fulfilled and also (26′) says that $\lambda^{-1}(t)=(\log V_t)^2 \approx \sigma^{-4}\kappa_{1/2}^4 V_{t0}^2 \langle N\rangle_t^{-2}$. Choosing $\phi_s=\psi_a^2(s-)\psi_a^2(s)V_s^2$ we see that the product ϕ_s tends to $\kappa_{1/2}^4 V_{t0}^2$. Then, applying Proposition 2, it is easy to get that $b^{-2}\sigma^{-2}\lambda(t)K_3(t)$ tends in distribution to $(\sigma^2/(4\beta^2))\int_0^1 W_s^2\,\mathrm{d}s$. Hence assertion (ii) holds.

Finally let $a<\frac{1}{2}$. Here we write

$$V_t^{2a-1}C_t(a)=V_t^{2a-1}\int_{t0}^{t}[V_s^{2a}Z_{s-}^2]V_s^{-2a}\,\mathrm{d}V_s,$$

and we notice that since $2a<1$, due to Proposition 3(ii), $\int_{t0}^t V_s^{-2a}\,\mathrm{d}V_s \approx (1-2a)^{-1}V_t^{1-2a}$. Then using assertion (iii) of Theorem 1 and the Tœplitz Lemma we get that assertion (iii) holds. □

Remark 2. (a) The best rate of convergence in Theorem 1 above is clearly obtained for $a>\frac{1}{2}$. Moreover the factor $g(a)$ has its minimum value 1 for $a=1$ i.e. $b=\beta$.

(b) Of course in the discrete time case where errors are i.i.d. Gaussian random variables (cf. Section 1) the assumptions of Theorems 1 and 2 are fulfilled. Then our assertions reduce to corresponding ones in Theorem 2 of Lai and Robbins (1979). Notice that our assertion (ii) in Theorem 1 is more precise than their result in case of $a=\frac{1}{2}$.

(c) In fact in Theorem 2 one may replace $C_t(a)$ by $\int_{t0}^t(\theta_s^a-\theta)^2\,\mathrm{d}V_s$.

Acknowledgement

The author wishes to thank a referee for valuable suggestions on the first version of the paper.

References

A. Benveniste, M. Métivier and P. Priouret, Algorithmes adaptatifs et approximations stochastiques (Masson, Paris, 1987).

C. Dellacherie and P.A. Meyer, Probabilités et potentiel. Théorie des martingales (Hermann, Paris, 1980) Chap. V à VIII.

M. Driml and J. Nedoma, Stochastic approximation for continuous random processes, in: Trans. 2nd Prague Conf. Inform. Theory (Academia, Prague, 1967) pp. 145–158.

M. Duflo, Méthodes récursives aléatoires (Masson, Paris, 1990).

P. Hall and C.C. Heyde, Martingale Limit Theory and its Applications (Academic Press, New York, 1980).

J. Jacod and A.N. Shiryaev, Limit Theorems for Stochastic Processes (Springer, Berlin, 1987).

H.J. Kushner and D.S. Clark, Stochastic Approximation for Constrained and Unconstrained Systems (Springer, Berlin, 1978).

H.J. Kushner and G. Yin, Asymptotic properties of distributed and communicating stochastic approximation algorithms, SIAM J. Control Optim. 25 (1987a) 1266–1290.

H.J. Kushner and G. Yin, Stochastic approximation algorithms for parallel and distributed processing, Stochastics 22 (1987b) 219–250.

T.L. Lai and H. Robbins, Adaptive design and stochastic approximation, Ann. Statist. 7(6) (1979) 1196–1221.

T.L. Lai and H. Robbins, Consistency and asymptotic efficiency of slope estimates in stochastic approximation schemes, Z. Wahrsch. Verw. Gebiete 56 (1981) 329–360.

A. Le Breton, Propriétés asymptotiques des martingales gaussiennes et application, C.R. Acad. Sci. Paris Sér. I 315 (1992) 839–844.

R.S. Liptser and A.N. Shiryaev, Statistics of Random Processes II, Applications (Springer, Berlin, 1978).

A.V. Melnikov, Stochastic approximation for semimartingales, in: A.N. Shiryaev, ed., Statistics and Control of Stochastic Processes (Nauka, Moscow, 1989) pp. 147–156.

A.V. Melnikov and A.E. Rodkina, Consistent statistical estimation in semimartingale model of stochastic approximation, Ann. Acad. Sci. Fenn. Ser. A I Math. 17 (1992) 85–91.

M.B. Nevelson and R.Z. Khasminskii, Stochastic Approximation and Recursive Estimation (Amer. Math. Soc., Providence, RI, 1973).

B.T. Polyak, New method of stochastic approximation type, Automat. Remote Control 51 (1990) 937–946.

H. Robbins and S. Monro, A stochastic approximation method, Ann. Math. Statist. 22 (1951) 400–407.

C.Z. Wei, Multivariate adaptive stochastic approximation, Ann. Statist. 15 (1987) 1115–1130.

G. Yin, On extensions of Polyak's averaging approach to stochastic approximation, Stochastics 36 (1991) 245–264.