JOURNAL OF MULTIVARIATE ANALYSIS, 39, 236–245 (1991)

Recursive Approximate Maximum Likelihood Estimation for a Class of Counting Process Models

PETER SPREIJ

Free University, Amsterdam

Communicated by C. R. Rao

In this paper we present a recursive algorithm that produces estimators of an unknown parameter that occurs in the intensity of a counting process. The estimators can be considered as approximations of the maximum likelihood estimator. We prove consistency of the estimators and derive their asymptotic distribution by using Lyapunov functions and weak convergence for martingales. The conditions that we impose in order to prove our results are similar to those in papers on (quasi) least squares estimation. © 1991 Academic Press, Inc.

INTRODUCTION

We assume that we are given a complete probability space (Ω, F, P) together with a filtration $\{F_t\}_{t \geqslant 0}$, satisfying the usual conditions in the sense of [2]. All stochastic processes to be encountered below are assumed to be adapted with respect to the given filtration and have cadlag paths. Similarly the martingale property is also to be understood with respect to this filtration. Let $N: \Omega \times [0, \infty) \to N_0$ be a counting process, such that its Doob–Meyer decomposition takes the following form (in differential notation)

$$dN_t = \varphi_t^{\mathrm{T}} \theta \, dt + dm_t. \tag{1}$$

Here $\varphi: \Omega \times [0, \infty) \to [0, \infty)^d$ is a predictable process, $\theta \in (0, \infty)^d$ an unknown parameter and $m: \Omega \times [0, \infty) \to R$ a local martingale. Superscript T usually denotes transposition.

The purpose of this paper is to give a recursive scheme that generates estimators $\{\hat{\theta}\}$ of the unknown θ. This scheme is given below as the set of Equations (2a)–(2e) (see the appendix for the implementation of this scheme). In an earlier paper [6] we have presented a similar but slightly

Received October 10, 1989; revised January 27, 1991.

AMS classification numbers: 62F10, 62L12.

Key words and phrases: recursive estimation, counting processes.

0047-259X/91 $3.00

Copyright © 1991 by Academic Press, Inc.
All rights of reproduction in any form reserved.

different set of equations. For a heuristic derivation of these equations we refer to [6], where also an account for the terminology approximate maximum likelihood (AML) estimation can be found.

A recursive scheme is attractive in two ways. Suppose one has observations over an interval $[0, T]$, and using these observations one wants an estimator of θ. One way to obtain such an estimator is to use nonlinear optimization techniques applied to the likelihood ratio. The use of a recursive scheme like (2a)–(2e) may be considered as an alternative approach to get an estimator $\hat{\theta}_T$. Another potential application is in the field of adaptive control. In problems of this type one has to use at each time instant t of the observation interval a control law which depends on an estimate of the unknown parameter using the observation up to that ·time t and this naturally calls for an efficient way to update current estimates, when new observations come in. A recursive scheme provides such an efficient, easy to compute, way of updating estimators.

The conditions that we impose in Theorem 3 in order to prove a.s. convergence of the estimators are of the same form as those in, e.g., [1, 5], where (quasi) least squares estimation has been studied and considerably weaker than those in [6]. However, we do not need all the conditions of [1]. In the sequel θ_0 will denote the "true" parameter value, $\mathbf{1}$ is the vector in R^d whose entries are all equal to 1. After giving an assumption on the parameter space, we present our estimation algorithm and an analysis of its asymptotic properties.

The Results

Assumption 1. *θ_0 lies in a compact subset of $\mathbb{R}^d_+$. Hence there exists $\varepsilon > 0$ such that $\varepsilon < \theta_{0i} < 1/\varepsilon$, $\forall i = 1, ..., d$.*

AML Algorithm.

$$dX_t = \frac{Q_t \phi_t}{\phi_t^{\mathrm{T}} \hat{\theta}_{t-}} (dN_t - \phi_t^{\mathrm{T}} X_t dt), X_0 \tag{2a}$$

$$dQ_t = -\frac{Q_t \phi_t \phi_t Q_t}{\phi_t^T \hat{\theta}_t} dt, \qquad Q_0 > 0 \tag{2b}$$

$$\hat{\theta}_t = I_{1t} I_{2t} X_t + \varepsilon(1 - I_{1t})\,\mathbf{1} + \varepsilon^{-1}(1 - I_{2t})\,\mathbf{1} \tag{2c}$$

$$I_{1t} = 1_{\{\phi_t^T X_t \geqslant \varepsilon \phi_t^T \mathbf{1}\}} \tag{2d}$$

$$I_{2t} = 1_{\{\phi_t^T X_t \leqslant \varepsilon^{-1} \phi_t^T \mathbf{1}\}} \tag{2e}$$

Comment. Introducing the ε above is done to establish a.s. convergence of $\{\hat{\theta}_t\}$ to θ_0. If we compare (2) to the AML algorithm in [6] we see that we use the extra indicator process I_2. Clearly we require knowledge of ε to compute the $\hat{\theta}_t$. The proof of $\hat{\theta}_t \to \theta_0$ a.s. that we will give parallels to a certain extent the procedure in [1]. First we state an auxiliary result. Define $\bar{Q}_t^{-1} = \int_0^t \phi_s \phi_s^{\mathrm{T}} / \phi_s^{\mathrm{T}} \theta_0 \, ds$. Denote by $\underline{\lambda}_t$ the minimal eigenvalue of $\bar{Q}_t^{-1}$ and by $\bar{\lambda}_t$ its maximal eigenvalue.

LEMMA 2. *There exist constants $\underline{c}$ and $\bar{c}$ such that*

(i) $\bar{c} + \varepsilon^2 \bar{\lambda}_t \leqslant \lambda_{\max}(Q_t^{-1}) \leqslant \varepsilon^{-2} \bar{\lambda}_t + \bar{c}$

(ii) $\underline{c} + \varepsilon^2 \underline{\lambda}_t \leqslant \lambda_{\min}(Q_t^{-1}) \leqslant \varepsilon^{-2} \underline{\lambda}_t + \underline{c}$.

Proof. Define $\underline{c} = \inf_{|x|=1} x^{\mathrm{T}} Q_0^{-1} x$ and $\bar{c} = \sup_{|x|=1} x^{\mathrm{T}} Q_0^{-1} x$. Since $\varepsilon \phi^{\mathrm{T}} \mathbf{1} \leqslant \phi^{\mathrm{T}} \hat{\theta}_t \leqslant \varepsilon^{-1} \phi^{\mathrm{T}} \mathbf{1}$ we have for all $x \in \mathbb{R}^d$:

$$x^{\mathrm{T}} Q_0^{-1} x + \varepsilon^2 x^{\mathrm{T}} \bar{Q}_t^{-1} x \leqslant x^{\mathrm{T}} Q_t^{-1} x \leqslant x^{\mathrm{T}} Q_0^{-1} x + \varepsilon^{-2} x^{\mathrm{T}} \bar{Q}_t^{-1} x. \tag{3}$$

By taking infima in (3) in the right order we obtain (i). The second assertion follows by taking suprema.

THEOREM 3. *Consider the AML algorithm* (2). *Assume that $\underline{\lambda}_t \to \infty$ a.s. and that there exists a function $f\colon [0, \infty) \to [0, \infty)$ such that $\lim_{x \to \infty} f(x)/x = \infty$ and such that*

$$\sup_{t \geqslant 0} \frac{f(\log \bar{\lambda}_t)}{\underline{\lambda}_t} < \infty \qquad a.s.$$

Then $\hat{\theta}_t \to \theta_0$ a.s.

Remarks. 1. Observe that $\underline{\lambda}_t \to \infty$ a.s. implies that $N_t \to \infty$ a.s. because

$$\int_0^t \phi_s^{\mathrm{T}} \theta_0 \, ds = \theta_0^{\mathrm{T}} \bar{Q}_t^{-1} \theta_0 \geqslant \underline{\lambda}_t \theta_0^{\mathrm{T}} \theta_0.$$

2. A possible choice of f that can be found in the literature [1, 5] is $f(x) = x^{1+\alpha}$, with $\alpha > 0$.

The crucial step in the proof of Theorem 3 is Lemma 4 below. We will postpone the proof of this lemma and show first, after stating the lemma, how we use it in the proof of Theorem 3.

LEMMA 4. *Consider* (2). *Let $\tilde{X}_t = X_t - \theta_0$ and $P_t = \tilde{X}_t^{\mathrm{T}} Q_t^{-1} \tilde{X}_t$. Then $P_t = O(\log \bar{\lambda}_t)$ a.s. $(t \to \infty)$.*

Proof of Theorem 3.

$$\tilde{X}_t^{\mathrm{T}} \tilde{X}_t = \tilde{X}_t^{\mathrm{T}} Q_t^{-1/2} Q_t Q_t^{-1/2} \tilde{X}_t \leqslant \lambda_{\max}(Q_t) P_t$$

$$= \frac{P_t}{\lambda_{\min}(Q_t^{-1})} = \frac{f(\log \bar{\lambda}_t)}{\underline{\lambda}_t} \cdot \frac{\underline{\lambda}_t}{\lambda_{\min}(Q_t^{-1})} \cdot \frac{\log \bar{\lambda}_t}{f(\log \bar{\lambda}_t)} \cdot \frac{P_t}{\log \bar{\lambda}_t}. \tag{4}$$

Consider the right-hand side of (4). Its last factor is bounded in view of Lemma 4. The first factor is bounded because of the assumption in the theorem. The second factor is bounded because of Lemma 2 and the third factor tends to zero because of the assumption on f. We conclude that $\tilde{X}_t \to 0$ a.s. But now it is easy to show that $\hat{\theta}_t \to \theta_0$ a.s.:

$$\tilde{\theta}_t = \hat{\theta}_t - \theta_0 = \tilde{X}_t I_{1t} I_{2t} + (1 - I_{1t})(\varepsilon \mathbf{1} - \theta_0) + (1 - I_{2t})(\varepsilon^{-1}\mathbf{1} - \theta_0).$$

Since $\phi_t^{\mathrm{T}} \theta_0 > \phi_t^{\mathrm{T}} \mathbf{1}\varepsilon$ there is $\eta > 0$ such that $\phi_t^{\mathrm{T}} \theta_0 \geqslant \phi_t^{\mathrm{T}} \mathbf{1}(\varepsilon + \eta)$. Because $\tilde{X}_t \to 0$ we eventually have $|\tilde{X}_{it}| < \eta$, $\forall i$. But then

$$\phi_t^{\mathrm{T}} X_t = \phi_t^{\mathrm{T}} \tilde{X}_t + \phi_t^{\mathrm{T}} \theta_0 \geqslant -\phi_t^{\mathrm{T}} \mathbf{1}\eta + \phi_t^{\mathrm{T}} \mathbf{1}(\varepsilon + \eta) = \phi^{\mathrm{T}} \mathbf{1}\varepsilon.$$

Therefore $I_{1t} \to 1$. In a similar way one can prove that $I_{2t} \to 1$, which implies that $\tilde{\theta}_t \to 0$ a.s. ▮.

The proof of Lemma 4 involves a series of other lemmas.

LEMMA 5. *Let $P_0 > 0$, $P_0 \in \mathbb{R}^{k \times k}$, and let $P_t = P_0 + \int_0^t \xi(s)\,\xi(s)^{\mathrm{T}}\,ds$ for a left continuous function $\xi\colon [0, \infty) \to \mathbb{R}^k$. Then*

(i) $\int_0^t \xi(s)^{\mathrm{T}} P_s^{-1} \xi(s)\,ds = \log\det(P_t) - \log\det(P_0)$

(ii) $\int_0^t \xi(s)^{\mathrm{T}} P_s^{-1} \xi(s)\,ds = O(\log \lambda_{\max}(P_t))$.

Proof. [1].

LEMMA 6. *Let m be a quasi left-continuous locally square integrable martingale with $\langle m \rangle = A$. Let $f\colon [0, \infty) \to [0, \infty)$ be a differentiable increasing function with*

$$\lim_{x \to \infty} f(x) = \infty \qquad \textit{and} \qquad \int_0^\infty \frac{dx}{(1 + f(x))^2} < \infty.$$

Define $g_t = 1 + f(A_t)$. Then both $g_t^{-1} m_t$ and $g_t^{-2}[m, m]_t$ converge almost surely for $t \to \infty$. On $\{A_\infty = \infty\}$ both limits equal zero a.s.

Proof. This is a simple application of Lemma A (see Appendix). Consider $g_t^{-1} m_t$. Define $X_t = g_t^{-2} m_t^2$. Then application of the stochastic calculus rule yields

$$\begin{aligned} dX_t &= -2g_t^{-3} f'(A_t)\, m_t^2\, dA_t + g_t^{-2}(2m_{t-}\, dm_t + d[m, m]_t) \\ &= -2g_t^{-1} f'(A_t)\, X_t\, dA_t + g_t^{-2}\, dA_t + g_t^{-2}(2m_{t-}\, dm_t + d([m, m]_t - A_t)). \end{aligned}$$

Notice that $f'(A_t) \geqslant 0$. Application of Lemma A immediately yields the desired result, since

$$\int_0^\infty g_t^{-2}\, dA_t = \int_0^{A_\infty} \frac{dx}{(1 + f(x))^2} < \infty.$$

On $\{A_\infty = \infty\}$ Lemma A also yields that $X_t \to 0$ because

$$\int_0^\infty g_t^{-1} f'(A_t)\, X_t\, dA_t = \int_0^\infty X_t\, d\log g_t.$$

The statement about $g_t^{-2}[m, m]_t$ can be proved similarly. ▮.

Remarks. 1. The statements of the lemma can be summarized as

$$m_t = o(g_t) + O(1) \qquad \text{and} \qquad [m, m]_t = o(g_t^2) + O(1).$$

2. Of course we may replace g_t in the lemma by $f(A_t)$ since we consider the behaviour for $t \to \infty$.

3. Convenient choices of f in applications are $f(x) = x^{(1+\alpha)/2}$, with $\alpha > 0$.

Proof of Lemma 4. For $\tilde{X}$ we have the following equation:

$$d\tilde{X}_t = \frac{Q_t \phi_t}{\phi^{\mathrm{T}} \hat{\theta}_{t-}} (dm_t - \phi_t^{\mathrm{T}} \tilde{X}_t\, dt).$$

Hence,

$$\begin{aligned} dP_t &= d(\tilde{X}_t^{\mathrm{T}} Q_t^{-1} \tilde{X}_t) \\ &= 2 \frac{\phi_t^{\mathrm{T}} \tilde{X}_{t-}}{\phi_t^{\mathrm{T}} \hat{\theta}_{t-}} (dm_t - \phi_t^{\mathrm{T}} \tilde{X}_t dt) + \frac{(\tilde{X}_t^{\mathrm{T}} \phi_t)^2}{\phi_t^{\mathrm{T}} \hat{\theta}_t} dt + \frac{\phi_t^{\mathrm{T}} Q_t \phi_t}{(\phi_t^{\mathrm{T}} \hat{\theta}_{t-})^2} dN_t \end{aligned}$$

or

$$\begin{aligned} P_t - P_0 &+ \int_0^t \frac{(\tilde{X}_s^{\mathrm{T}} \phi_s)^2}{\phi_s^{\mathrm{T}} \hat{\theta}_s} ds \\ &= 2 \int_0^t \phi_s^{\mathrm{T}} \hat{\theta}_{s-}\, dm_s + \int_0^t \frac{\phi_s^{\mathrm{T}} Q_s \phi_s}{(\phi_s^{\mathrm{T}} \hat{\theta}_s)^2} \phi_s^{\mathrm{T}} \theta_0\, ds + \int_0^t \frac{\phi_s^{\mathrm{T}} Q_s \phi_s}{(\phi_s^{\mathrm{T}} \hat{\theta}_{s-})^2} dm_s. \end{aligned} \tag{6}$$

Write (6) in obvious notation as

$$P_t - P_0 + L_t = 2M_{1t} + R_t + M_{2t}. \tag{7}$$

Compute

$$\langle M_1 \rangle_t = \int_0^t \frac{(\phi_s^{\mathrm{T}} \tilde{X}_s)^2}{\phi_s^{\mathrm{T}} \hat{\theta}_s} \cdot \frac{\phi_s^{\mathrm{T}} \theta_0}{\phi_s^{\mathrm{T}} \hat{\theta}_s} ds.$$

Observe that

$$\varepsilon^2 \leqslant \frac{\phi_s^{\mathrm{T}} \theta_0}{\phi_s^{\mathrm{T}} \hat{\theta}_s} \leqslant \varepsilon^{-2}.$$

Hence $\varepsilon^2 L_t \leqslant \langle M_1 \rangle_t \leqslant \varepsilon^{-2} L_t$. Hence $M_{1t} = o(L_t) + O(1)$ in view of Lemma 6 (take $f(x) = x$), and Remarks 1 and 2 that follow this lemma. Consider now R_t and notice that

$$\varepsilon^2 \int_0^t \frac{\phi_s^{\mathrm{T}} Q_s \phi_s}{\phi_s^{\mathrm{T}} \hat{\theta}_s} ds \leqslant R_t \leqslant \varepsilon^{-2} \int_0^t \frac{\phi_s^{\mathrm{T}} Q_s \phi_s}{\phi_s^{\mathrm{T}} \hat{\theta}_s} ds. \tag{8}$$

The integrals in the extreme sides of (8) are of the form encountered in Lemma 5. (Take $\xi(s) = \phi_s/(\phi_s^{\mathrm{T}} \hat{\theta}_s)^{1/2}$, $Q_t^{-1} = P_t$.) Therefore $R_t = O(\log \lambda_{\max}(Q_t^{-1}))$. The last term to analyze in (7) is M_{2t}:

$$\begin{aligned} \langle M_2 \rangle_t &= \int_0^t \frac{(\phi_s^{\mathrm{T}} Q_s \phi_s)^2}{(\phi_s^{\mathrm{T}} \hat{\theta}_s)^4} \phi_s^{\mathrm{T}} \theta_0 ds = \int_0^t - \frac{\phi_s^{\mathrm{T}} dQ_s \phi_s}{(\phi_s^{\mathrm{T}} \hat{\theta}_s)^2} \frac{\phi_s^{\mathrm{T}} \theta_0}{\phi_s^{\mathrm{T}} \hat{\theta}_s} ds \\ &\leqslant \int_0^t \frac{\phi_s^{\mathrm{T}} \phi_s}{\phi_s^{\mathrm{T}} \hat{\theta}_s)^2} \cdot \frac{\phi_s^{\mathrm{T}} \theta_0}{\phi_s^{\mathrm{T}} \hat{\theta}_s} d\,\mathrm{tr}(-Q_s) \leqslant \varepsilon^{-4} \int_0^t d\,\mathrm{tr}(-Q_s) \\ &\leqslant \varepsilon^{-4} \mathrm{tr}(Q_0) < \infty. \end{aligned}$$

From Lemma 6 we conclude that $M_2/\langle M_2 \rangle$ converges to a finite limit and, since $\langle M_2 \rangle_t \leqslant \varepsilon^{-4} \mathrm{tr}(Q_0)$, M_2 is a.s. bounded. Collecting the above results we obtain from (7)

$$P_t - P_0 + L_t = o(L_t) + O(1) + O(\log \lambda_{\max}(Q_t^{-1})) + O(1)$$

or

$$P_t - P_0 + L_t(1 + o(1)) = O(1) + O(\log \lambda_{\max}(Q_t^{-1})).$$

From Lemma 2 we obtain, after dividing by $\log \bar{\lambda}_t$,

$$\frac{P_t}{\log \bar{\lambda}_t} + (1 + o(1)) \frac{L_t}{\log \bar{\lambda}_t} = O(1).$$

Since both P_t and $(1 + o(1)) L_t$ are (eventually) nonnegative we obtain $P_t = O(\log \bar{\lambda}_t)$, as was to be proven. ∎.

We close this section by proving that the limit distribution of the AML estimators defined by (2) is asymptotically normal.

THEOREM 7. *Assume that* $\{\hat{\theta}_t\}$ *given by* (2) *is a.s. convergent. Assume that there exist* $P\colon [0, \infty) \to R^{d \times d}$ *and* $h\colon [0, 1] \times [0, \infty) \to [0, \infty)$ *such that*

(i) *h is increasing in each of its arguments and for all t, T,*

$$h(t, T) \leqslant h(1, T) = T.$$

(ii) *all $P(t)$ are symmetric positive definite for $t>0$, P is increasing to infinity, continuous, and $R(t)=\lim_{T\to\infty} P(T)^{-1/2} P(h(t,T) P(T)^{-1/2}$ exists and is positive definite for all $t>0$.*

(iii) *$P(t)^{1/2}\bar{Q}^{1/2}\to I$ in probability. Then $\bar{Q}^{-1/2}\tilde{\theta}_t \xrightarrow{\mathscr{L}} N(O,I)$.*

Proof. Since both I_{1t} and I_{2t} tend to 1, eventually $X_t=\tilde{\theta}_t$. Therefore, it is sufficient to prove that $\bar{Q}_t^{-1/2}\tilde{X}_t \xrightarrow{\mathscr{L}} N(0,I)$.

From (2) we obtain $Q_t^{-1}\tilde{X}_t=Q_0^{-1}\tilde{X}_0+M_t$, where

$$M_t=\int_0^t \frac{\varphi}{\varphi^T\tilde{\theta}_-}\,dm.$$

Hence the asymptotic law of $\bar{Q}_t^{-1/2}\tilde{X}_t$ is the same as that of

$$\bar{Q}_t^{-1/2}Q_t\bar{Q}_t^{-1/2}\bar{Q}_t^{1/2}M_t. \tag{9}$$

As in [6] it is easy to prove that

$$\bar{Q}_t^{-1/2}Q_t\bar{Q}_t^{-1/2}\to I \qquad \text{a.s.} \tag{10}$$

Hence it suffices to establish that the asymptotic law of $\bar{Q}_t^{1/2}M_t$ is $N(0,I)$. Now

$$\langle M\rangle_t=\int_0^t \frac{\varphi_s\varphi_s^{\mathrm{T}}}{\varphi_s^{\mathrm{T}}\theta_0}\left(\frac{\varphi_s^{\mathrm{T}}\theta_0}{\varphi_s^{\mathrm{T}}\hat{\theta}_s}\right)^2 ds.$$

Similar to the proof for (10), it is possible to show that

$$\bar{Q}_t^{-1/2}\langle M\rangle_t\bar{Q}_t^{-1/2}\to I \qquad \text{a.s.}$$

Then using condition (iii) of the theorem, it follows that

$$P(t)^{-1/2}\langle M\rangle_t P(t)^{-1/2}\to I \qquad \text{in probability.} \tag{11}$$

Define now for each $T\in(0,\infty)$ and $\lambda\in R^d$ a new martingale (w.r.t. the filtration $\{F_{h(t,T)}\}_{t\in[0,1]}$) $Z^{T,\lambda}$ by $Z_t^{T,\lambda}=\lambda^{\mathrm{T}}P(T)^{-1/2}M_{h(t,T)}$.

Now let W be some continuous Gaussian martingale with quadratic variation $\langle W\rangle_t=\lambda^{\mathrm{T}}R(t)\lambda$. Such a W exists on a suitable filtered probability space, since $R(t)$ is continuously increasing. We claim

$$Z^{T,\lambda}\xrightarrow{\mathscr{L}} W. \tag{12}$$

We prove (12) by checking the conditions of Lemma B (see Appendix). Compute

$$\begin{aligned}\langle Z^{T,\lambda}\rangle_t&=\lambda^{\mathrm{T}}P(T)^{-1/2}\langle M\rangle_{h(t,T)}P(T)^{-1/2}\lambda\\&=\lambda^{\mathrm{T}}P(T)^{-1/2}P(h(t,T))^{1/2}(P(h(t,T))^{-1/2}\langle M\rangle_{h(t,T)}\\&\quad\times P(h(t,T))^{-1/2})P(h(t,T))^{1/2}P(T)^{-1/2}\lambda.\end{aligned}$$

From conditions (ii) and (11) we then obtain for $T \to \infty$,

$$\langle Z^{T,\lambda} \rangle_t \to \lambda^{\mathrm{T}} R(t)\, \lambda$$

which corresponds to condition (i) of Lemma B.

Observe that for the jumps of $Z^{T,\lambda}$ we have

$$\begin{aligned} |\Delta Z_t^{T,\lambda}|^2 &= \left| \lambda^{\mathrm{T}} P(T)^{-1/2} \frac{\varphi_t}{\varphi_t^{\mathrm{T}} \hat{\theta}_{t-}} \right|^2 \\ &\leqslant \lambda^{\mathrm{T}} P(T)^{-1}\, \lambda \frac{\varphi_t^{\mathrm{T}} \varphi_t}{(\varphi_t^{\mathrm{T}} \hat{\theta}_{t-})^2} \leqslant \varepsilon^{-2} \lambda^{\mathrm{T}} P(T)^{-1}\, \lambda. \end{aligned}$$

Hence the jumps of $Z^{T,\lambda}$ are bounded by a deterministic quantity that tends to zero. Hence also the second condition of Lemma B is satisfied and (12) follows. In particular, $\lambda^{\mathrm{T}} P(T)^{-1/2} M_T = Z_1^{T,\lambda} \xrightarrow{\mathscr{L}} N(0, \lambda^{\mathrm{T}} \lambda)$. Finally by noticing that $\bar{Q}_t^{1/2} M_t = \bar{Q}_t^{1/2} P(t)^{-1/2} P(t)^{1/2} M_t$ and by using condition (iii) again, we have finished the proof.

As a final remark we mention that the behaviour of these AML algorithms, in general, will be superior to a least squares algorithm like in [1]. The easiest way, although it does not give a complete account, to see this is to assume that the process φ in (1) is deterministic. Then the Fisher information matrix at time t becomes $\bar{Q}_t^{-1}$. Hence from Theorem 7 we see that our estimators have an asymptotic variance that equals the Cramer–Rao bound. It is also this observation that led us to considering the algorithm (2).

Appendix

The next lemma generalizes a result in [7].

Lemma A. *Let X be a nonnegative stochastic process such that $X_t = X_0 + A_t - B_t + M_t$. Here A and B are predictable increasing processes with $A_0 = B_0 = 0$ and M is a local martingale. Assume that* $\lim_{t \to \infty} A_t < \infty$ *a.s. Then both* $\lim_{t \to \infty} X_t$ *and* $\lim_{t \to \infty} B_t$ *exist and are finite.*

Proof. Without loss of generality we assume that $X_0 = 0$ a.s. Let $\{T_n\}$ be a fundamental sequence for M [3]. Let $\{S_n\}$ be stopping times defined by $S_n = \inf\{t > 0 : A_t > n\}$. Each S_n is then predictable and hence there exist for each n another sequence of stopping times $\{S'_{n,k}\}_{k \geqslant 0}$ announcing S_n. Define $R_n = \sup\{S'_{n,k} : k \leqslant n\}$. Then $R_n < S_n$ and $A_{R_n} \leqslant n$. Furthermore, $\{R_n = \infty\} \uparrow \Omega$.

Now for all k, n $\{M_{t \wedge T_n \wedge R_k}\}_{t \geqslant 0}$ is a uniformly integrable martingale, and

$$M_{t \wedge T_n \wedge R_k} \geqslant X_{t \wedge T_n \wedge R_k} - A_{t \wedge T_n \wedge R_k} \geqslant -k.$$

In particular, $\{M^-_{t \wedge T_n \wedge R_k}\}_{n \geqslant 0}$ is uniformly integrable ($M^-_s = \max(0, -M_s)$). Hence,

$$\begin{aligned} E[M_{t \wedge R_k} | F_s] &= E[\lim_{n \to \infty} M_{t \wedge R_k \wedge T_n} | F_s] \\ &\leqslant \liminf_{n \to \infty} E[M_{t \wedge R_k \wedge T_n} | F_s] \\ &= \liminf_{n \to \infty} M_{s \wedge R_k \wedge T_n} = M_{s \wedge R_k}. \end{aligned}$$

Here the inequality follows from Fatou's lemma. So we see that $\{M_{t \wedge R_k}\}_{t \geqslant 0}$ is a supermartingale with $M^-_{t \wedge R_k} \leqslant k$. Hence the convergence theorem [3] for supermartingales tells us that $\lim_{t \to \infty} M_{t \wedge R_k}$ exists and is finite a.s. But then also $\lim_{t \to \infty} (X_{t \wedge R_k} + B_{t \wedge R_k})$ exists and is finite. Since both $X_{t \wedge R_k}$ and $B_{t \wedge R_k}$ are nonnegative and B is increasing, we obtain that both $\lim_{t \to \infty} X_{t \wedge R_k}$ and $\lim_{t \to \infty} B_{t \wedge R_k}$ exist and are finite.

On the set $\{R_k = \infty\}$ these limits equal $\lim_{t \to \infty} X_t$ and $\lim_{t \to \infty} B_t$, respectively. But $\{R_k = \infty\} \uparrow \Omega$, which finishes the proof. ∎

The following lemma is a special case of a much more general result on weak convergence of locally square integrable martingales, that can be found in, for instance, the monograph by Jacod and Shiryaev [4].

LEMMA B (Central limit theorem). *Let M, M^n, $n \geqslant 0$ be real valued locally square integrable martingales defined on a suitable filtered probability space. Let M be a continuous Gaussian martingale with $C_t = \langle M \rangle_t = EM_t^2$. Assume that the following two conditions hold*:

(i) $\langle M^n \rangle_t \to C_t$ *in probability as $n \to \infty$ for all t.*

(ii) $\sup_t |\Delta M_t^n| \leqslant c_n$, *where $\{c_n\}$ is a deterministic sequence with $\lim_{n \to \infty} c_n = 0$. Then $\{M^n\}$ converges weakly to M for $n \to \infty$. Note: $M^n \xrightarrow{\mathscr{L}} M$.*

Using a suitable discretization of time, (2a)–(2e) can be implemented in the following way:

Let $0 = t_0 < t_1 \cdots < T_N = T$. Suppose that t_i is the actual time instant and that we have computed X_{t_i}, $\hat{\theta}_{t_i}$, Q_{t_i}. New observations at t_{i+1} are $N_{t_{i+1}}$ and $\varphi_{t_{i+1}}$. Then we compute from the discretized version of (2a)–(2e):

$$Q_{t_{i+1}} = Q_{t_i} - \frac{Q_{t_i}\varphi_{t_{i+1}}\varphi^{\mathrm{T}}_{t_{i+1}}Q_{t_i}}{\varphi^{\mathrm{T}}_{t_{i+1}}\hat{\theta}_{t_i}}(t_{i+1} - t_i)$$

$$X_{t_i+1} = X_{t_i} + \frac{Q_{t_{i+1}}\varphi_{t_{i+1}}}{\varphi^{\mathrm{T}}_{t_{i+1}}\hat{\theta}_{t_i}}(N_{t_{i+1}} - N_{t_i} - \varphi^{\mathrm{T}}_{t_{i+1}}X_{t_i}(t_{i+1} - t_i))$$

$$\hat{\theta}_{t_{i+1}} = I_{1t_{i1}}I_{2t_{i+1}}X_{t_{i+1}} + \varepsilon(1 - I_{1t_{i+1}})\,\mathbf{1} + \varepsilon^{-1}(1 - I_{2t_{i+1}})\,\mathbf{1}.$$

Finally, for $i = N - 1$ the result of the computation is an estimate $\hat{\theta}_T$.

References

[1] Christopeit, N. (1986). Quasi-least squares estimation in semi-martingale regression models. *Stochastics* **16** 255–278.

[2] Dellacherie, C., and Meyer, P. A. (1975). *Probabilités et Potentiel,* Chaps. I–IV. Hermann, Paris.

[3] Dellacherie, C., and Meyer, P. A. (1980). *Probabilités et Potentiel,* Chaps. V–VIII. Hermann, Paris.

[4] Jacod, J., and Shiryaev, A. N. (1987). *Limit Theorems for Stochastic Processes.* Springer-Verlag, New York/Berlin.

[5] Lai, T. L., and Wei, C. Z. (1982). Least squares estimates in stochastic regression models. *Ann. Statist.* **10** 154–166.

[6] Spreij, P. J. C. (1986). Recursive parameter estimation for counting processes with linear intensity. *Stochastics* **18** 277–312.

[7] Van Schuppen, J. H. (1983). Convergence results for continuous time adaptive stochastic filtering algorithms. *J. Math. Anal. Appl.* **96** 209–225.