Statistics & Probability Letters 5 (1987) 71–73
North-Holland

January 1987

A STRONG LAW OF LARGE NUMBERS FOR VECTOR GAUSSIAN MARTINGALES AND A STATISTICAL APPLICATION IN LINEAR REGRESSION

A. Le BRETON

Laboratory TIM 3, University of Grenoble I, France

M. MUSIELA

Department of Statistics, University of New South Wales, Australia

Received February 1986
Revised April 1986

Abstract: For a d-dimensional gaussian martingale M with tensor increasing process we prove that $\langle M\rangle_t^+ M_t$ converges in $\mathbb{R}^d$ with probability 1 as $t \to \infty$ and the limit is zero a.s. iff tr $\langle M\rangle_t^+$ tends to zero. We apply this result to study the strong consistency of estimates in a linear regression model.

Keywords: Strong law of large numbers, gaussian martingales, linear regressin, strong consistency.

1. A strong law of large numbers

The classical strong law of large numbers for real martingales is an important tool in proving consistency for estimates in statistical problems concerned with one dimensional parameters. It seems that for multidimensional martingales no analogous general result is available. Here we present a version of the strong law of large numbers for vector gaussian martingales. We apply this result to prove the strong consistency of estimates of multidimensional parameters in a normal linear multiple regression model.

Let $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\geqslant 0}, P)$ be a probability space satisfying usual conditions on which a d-dimensional gaussian martingale M ($M_0 = 0$) is given. Let $\langle M\rangle$ be the tensor increasing process of M and $\langle M\rangle_t^+$ denote the Moore–Penrose generalized inverse of the matrix $\langle M\rangle_t$.

Theorem. *The process $\langle M\rangle^+ M$ converges in $\mathbb{R}^d$ with probability* 1 *as $t \to \infty$. Moreover*

$$\lim_{t\to\infty} \langle M\rangle_t^+ M_t = 0 \ a.s. \text{ if and only if}$$

$$\lim_{t\to\infty} \operatorname{tr} \langle M\rangle_t^+ = 0.$$

Proof. To prove the first statement we shall show that, for all $a \in \mathbb{R}^d$, $a^*\langle M\rangle_t^+ M_t$ converges in $\mathbb{R}$ with probability 1 as $t \to \infty$.

Note that, for all $t \in \mathbb{R}^+$, $\langle M\rangle_t = EM_tM_t^*$ is a symmetric positive ($\geqslant 0$) matrix and that $\langle M\rangle_t \geqslant \langle M\rangle_s$ for $s \leqslant t$. Then for $s \leqslant t$

$$\begin{aligned}\operatorname{Im}(\langle M\rangle_s) &= \ker^{\perp}(\langle M\rangle_s) \subset \ker^{\perp}(\langle M\rangle_s) \\ &\qquad \cup \ker^{\perp}(\langle M\rangle_t - \langle M\rangle_s) \\ &\qquad \subset (\ker(\langle M\rangle_s) \cap \ker(\langle M\rangle_t - \langle M\rangle_s))^{\perp} \\ &\qquad = \ker^{\perp}(\langle M\rangle_t) = \operatorname{Im}(\langle M\rangle_t).\end{aligned}$$

Set $L = \cup_{t\geqslant 0} \operatorname{Im}(\langle M\rangle_t)$. From the above considerations if follows that there exists some $t_0 \in \mathbb{R}^+$ such that $L = \operatorname{Im}(\langle M\rangle_t)$ for all $t \geqslant t_0$. Therefore if $a \in L^{\perp} = \ker(\langle M\rangle_t) = \ker(\langle M\rangle_t^+)$, then

$$a^*\langle M\rangle_t^+ M_t = 0 \quad \text{for } t \geqslant t_0.$$

On the other hand if $a \in L$ and for some $t \geqslant t_0$, $a^*\langle M\rangle_t a = 0$, then

$$\begin{aligned}a &\in \ker(\langle M\rangle_t^+) \cap \operatorname{Im}(\langle M\rangle_t) \\ &= \ker(\langle M\rangle_t) \cap \operatorname{Im}(\langle M\rangle_t) = \{0\}.\end{aligned}$$

Thus for $a \in L\setminus\{0\}$ and $t \geqslant t_0$, $a^*\langle M\rangle_t a > 0$. To achieve the proof we shall need the following lemma which is of independent interest.

Lemma. *Let $a \in L = \cup_{t\geqslant 0} \operatorname{Im}(\langle M\rangle_t)$, $a \neq 0$, and let, for $t \geqslant t_0 = \inf(t: L = \operatorname{Im}(\langle M\rangle_t)]$,*

$$N_t = \left(a^*\langle M\rangle_t^+ a\right)^{-1} a^*\langle M\rangle_t^+ M_t.$$

and $\mathscr{F}_t^N$ denote the σ-algebra generated by $\{N_u: t_0 \leqslant u \leqslant t\}$.
Then the process (N_t) is a $(\mathscr{F}_t^N)$ martingale on $[t_0,\infty[$ with $((a^*\langle M\rangle_t^+ a)^{-1})$ as increasing process.

Proof. Note that for $t \geqslant s \geqslant t_0$ and $a \in L\setminus\{0\}$ we have

$$\begin{aligned} EN_tN_s &= \left(a^*\langle M\rangle_t^+ a\right)^{-1} a^*\langle M\rangle_t^+\langle M\rangle_s\langle M\rangle_s^+ \\ &\qquad \times a\left(a^*\langle M\rangle_s^+ a\right)^{-1} \\ &= \left(a^*\langle M\rangle_s^+ a\right)^{-1} = EN_s^2. \end{aligned}$$

This proves that for all $u \leqslant s$ the gaussian random variables $N_t - N_s$ and N_u are independent. Consequently, for $s \leqslant t$, $N_t - N_s$ and $\mathscr{F}_s^n = \sigma(N_u: t_0 \leqslant u \leqslant s)$ are independent. Therefore $E(N_t - N_s/\mathscr{F}_s^N) = E(N_t - N_s) = 0$ and (N_t) is a $(\mathscr{F}_t^N)$ gaussian martingale. Moreover, its increasing process $(\langle N\rangle_t)$ is given by

$$\langle N\rangle_t = EN_t^2 = \left(a^*\langle M\rangle_t^+ a\right)^{-1}.$$

To complete the proof of the first statement of the theorem note that for $a \in L\setminus\{0\}$ and $t \geqslant t_0$ we have

$$a^*\langle M\rangle_t^+ M_t = \langle N\rangle_t^{-1} N_t.$$

Therefore, by the strong law of large numbers for real martingales (see Liptser (1980) for example), $a^*\langle M\rangle_t^+ M_t$ converges in $\mathbb{R}$ with probability 1 as $t \to \infty$.

The second statement follows from the first since by Fatou's lemma

$$E \lim_{t\to\infty} M_t^*(\langle M\rangle_t^+)^2 M_t \leqslant \lim_{t\to\infty} \operatorname{tr}\langle M\rangle_t^+. \quad \square$$

2. Application in linear regression

Consider the continuous time multiple normal linear regression model

$$Y_t = \int_0^t (\mathrm{d}\langle M\rangle_s) X_s\theta + \sigma M_t,$$

where the measurable deterministic matrix valued function $X: \mathbb{R}^+ \to \mathbb{R}^d \otimes \mathbb{R}^p$ is such that for all $T \in \mathbb{R}^+$

$$\int_0^T \operatorname{tr} X_t X_t^* \,\mathrm{d}\langle M\rangle_t < \infty$$

and the parameter (θ, σ^2) belongs to $\Theta \times \mathbb{R}_+^*$ with

$$\Theta = \bigcup_{T\geqslant 0} \operatorname{Im}\left(\int_0^T X_t^* \,\mathrm{d}\langle M\rangle_t X_t\right).$$

Let us choose the estimate θ_T of θ in view of $(Y_t;\ t \leqslant T)$ which is given by

$$\theta_T = \left(\int_0^T X_t^* \,\mathrm{d}\langle M\rangle_t X_t\right)^+ \int_0^T X_t^* \,\mathrm{d}Y_t.$$

Note that θ_T is the maximum likelihood estimate of θ provided $\Theta \subset \operatorname{Im}\left(\int_0^T X_t^* \,\mathrm{d}\langle M\rangle_t X_t\right)$. Since for T large enough one can write

$$\theta_T - \theta = \sigma\left(\int_0^T X_t^* \,\mathrm{d}\langle M\rangle_t X_t\right)^+ \int_0^T X_t^* \,\mathrm{d}M_t;$$

using our theorem we obtain

Corollary. *The estimate θ_T is strongly consistent if and only if*

$$\lim_{T\to\infty} \operatorname{tr}\left(\int_0^T X_t^* \,\mathrm{d}\langle M\rangle_t X_t\right)^+ = 0.$$

Example 1. Let (ε_n) be a sequence of independent normally distributed random vectors in $\mathbb{R}^d$ with $E\varepsilon_n = 0$ and variance $E\varepsilon_n\varepsilon_n^* = V_n$. Moreover, let X_n, $n = 1,2,\ldots$, be $d \times p$ known matrices, $X_0 = 0$. Consider the discrete time linear regression model

$$Z_n = V_n X_n \theta + \sigma\varepsilon_n, \quad n = 1,2,\ldots,$$

where the parameter (θ, σ^2) belongs to $\Theta \times \mathbb{R}_+^*$ with

$$\Theta = \bigcup_{n=1}^{\infty} \operatorname{Im}\left(\sum_{k=1}^{n} X_k^* V_k X_k\right).$$

Set $M_t = \sum_{n=0}^{[t]} \varepsilon_n$, $t \geqslant 0$, $\varepsilon_0 = 0$, and $\mathscr{F}_t = \sigma(\varepsilon_0, \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{[t]})$, $t \geqslant 0$. It is clear that $(M_t, \mathscr{F}_t)$ is a gaussian martingale with tensor increasing process

$$\langle M \rangle_t = \sum_{n=0}^{[t]} V_n.$$

Define $(Y_n;\ n = 0,1,\dots)$ by $Y_0 = 0$, $Y_{n+1} = Y_n + Z_n$ and $Y_t = Y_n$, $X_t = X_n$ for $0 \leqslant n \leqslant t < n+1$. Then we can write our model in the form

$$Y_t = \int_0^t (\mathrm{d}\langle M \rangle_s) X_s \theta + \sigma M_t$$

and in view of the corollary we obtain that the estimate

$$\theta_n = \left(\sum_{k=1}^{n} X_k^* V_k X_k \right)^+ \sum_{k=1}^{n} X_k^* Z_k$$

is strongly consistent if and only if

$$\lim_{n \to \infty} \operatorname{tr} \left(\sum_{k=1}^{n} X_k^* V_k X_k \right)^+ = 0.$$

This result was proved by Anderson and Taylor (1976) in the case $d = 1$, $V_n = 1$ and $\Theta = \mathbb{R}^p$.

Example 2. Let (M_t) be a d-dimensional gaussian martingale with $\langle M \rangle_t = tI$, where I is the identity matrix. It is well known (see Jacod (1979) for example) that (M_t) is a d-dimensional standard Brownian motion. Consider the continuous time multiple regression model

$$Y_t = \int_0^t X_s \theta \, \mathrm{d}s + \sigma M_t, \quad t \geqslant 0,$$

where $X: \mathbb{R}_+ \to \mathbb{R}^d \otimes \mathbb{R}^p$ is a measurable deterministic function such that $\int_0^T \operatorname{tr} X_t^* X_t \, \mathrm{d}t < \infty$ for all $T \in \mathbb{R}^+$ and the parameter (θ, σ^2) belongs to $\Theta \times \mathbb{R}_+^*$ with

$$\Theta = \bigcup_{T \geqslant 0} \operatorname{Im}\left(\int_0^T X_t^* X_t \, \mathrm{d}t \right).$$

Then the estimate $\theta_T = (\int_0^T X_t^* X_t \, \mathrm{d}t)^+ \int_0^T X_t^* \, \mathrm{d}Y_t$ is strongly consistent if and only if

$\lim_{t \to \infty} \operatorname{tr}\left(\int_0^T X_t^* X_t \, \mathrm{d}t \right)^+ = 0.$

References

Anderson, T.W. and J.B. Taylor (1976), Strong consistency of least squares estimates in normal linear regression, *Ann. Statist.* 4, 788–790.

Jacod, J. (1979), *Calcul Stochastique et Problèmes de Martingales* (Springer-Verlag, Berlin, Heidelberg, New York).

Lipster, R.S. (1980), A strong law of large numbers for local martingales, *Stochastics* 3, 217–228.