Mathematical Problems in Engineering
Volume 7, pp. 87–95
Reprints available directly from the publisher
Photocopying permitted by license only

The Strong Law of Large Numbers for Dependent Vector Processes with Decreasing Correlation: "Double Averaging Concept"

ALEX S. POZNYAK*

CINVESTAV-IPN, Department of Automatic Control, A. P. 14-740, CP-07300, Mexico D.F., Mexico

(Received 10 October 2000)

A new form of the strong law of large numbers for dependent vector sequences using the "double averaged" correlation function is presented. The suggested theorem generalizes the well-known Cramer–Lidbetter's theorem and states more general conditions for fulfilling the strong law of large numbers within the class of vector random processes generated by a non stationary stable forming filters with an absolutely integrable impulse function.

Keywords: Law of large numbers; Correlation function; Forming filter; Dependent processes

AMS Classification Codes: 60Gxx, 93Exx, 94Bxx

1. INTRODUCTION

1.1. Problem Formulation

The strong law of large numbers tackles the problem

$$S_n := n^{-1}\sum_{t=1}^{n}\xi_t \overset{?}{\to} 0 \ (P - a.s.) \tag{1}$$

*e-mail: apoznyak@ctrl.cinvestav.mx

of the convergence to zero with probability one the time averaged process S_n generated by a discrete-time centered quadratically integrable random vector process $\xi_n \in R^n$ satisfying at each time $n = 1, 2, \ldots,$ the following conditions

$$E\{\xi_n\} = 0, E\{\xi_n \xi_n^{\top}\} = \Xi_n(\sigma_n^2 := tr\Xi_n < \infty) \tag{2}$$

In general, the random process $\{\xi_n\}$ may be dependent in time as well as may have increasing second moment.

This problem plays a key role in the analysis of the asymptotic behavior of recurrent algorithms analyzed in the identification [6, 9] and the stochastic adaptive control [2, 11] theories.

1.2. Motivating Example

This example deals with a simple identification problem. Estimating a vector parameter c^*, bazed on the observations

$$y_n = c^* + \xi_n$$

disturbed by the random (may be, dependent) noises ξ_n, and using the standart "averaging" estimate

$$\hat{c}_n := n^{-1} \sum_{t=1}^{n} y_t$$

it follows that the identification error at time n can be expressed as

$$\Delta_n := \hat{c}_n - c^* = n^{-1} \sum_{t=1}^{n} \xi_t = S_n$$

The analysis of the conditions, providing the convergence of the identification error to zero, directly leads to the problem (1) introduced above.

1.3. Preliminaries

A lot of fundamental results have been obtained for the independent random processes $\{\xi_n\}$ (see, for example, [5] and [8]). Later on, several elegant constructions, generalizing the strong law of large number to the class of dependent (*martingales* [7], weak and strong *mixing* [4] and *mixingales* [3]) processes have been propose. Unfortunately, most of the characteristics (such as "*mixing coefficients*"), participating in these constructions, are extremely complex for the direct calculation and turn out to be unapplicable in engineering practice. More clear and practically more useful results on the strong law of large numbers for dependent processes are contained in the publications operating with the *correlation function* as a main characteristic of a statistic dependence. The most advanced results in this direction have been obtained in [12] and [1] where the special decreasing conditions for the corresponding correlation function were introduced to guarantee the fulfilling of the strong law of large number.

1.4. Main Contribution

This note presents a new form of the strong law of large numbers using a special characteristic (the "double averaged" correlation function) of dependence which can be easily constructed based on correlation coefficients. This paper generalizes the earlier author's results [10] obtained for the scalar case. The well-known Cramer–Lidbetter's theorem is shown to be a partial case of the presented result as well as the intuitively used fact on the validity of the law of large number for the dependent processes generated by forming (may be, nonstationary) stable filters with white-noise type centered sequences in the input.

2. MAIN RESULT

Let all random sequences considered below be defined on the probability pace (Ω, F, P). For the given centered quadratic-integrable R^n-valued random process $\{\xi_n\}$, that is,

$$\begin{gathered}\xi_n \in R^n, E\{\xi_n\} = 0, \quad E\{\xi_n^\top \xi_n\} = \sigma_n^2 < \infty \\ n \in N^+ := \{1, 2, \ldots\}\end{gathered} \tag{3}$$

introduce the special characteristic, so-called, *the "double averaged" correlation function R_n defined by*

$$R_n := n^{-2} \sum_{t=1}^{n} \sum_{s=1}^{n} \rho_{t,s} = E\{S_n^\top S_n\} \tag{4}$$

where

$$\rho_{t,s} := E\{\xi_t^\top \xi_s\} \tag{5}$$

is the corresponding *correlation function.*

THEOREM 1 (The strong law of large numbers) *If for the vector process* $\{\xi_n\}$ (3) *the following series converges:*

$$\sum_{n \in N^+} \left(\frac{\sigma_n}{n} \sqrt{R_{n-1}} + \frac{1}{n^2} \sigma_n^2 \right) < \infty \tag{6}$$

*then "**the strong law of large numbers**" holds for this process, that is,*

$$S_n := \frac{1}{n} \sum_{t=1}^{n} \xi_t \overset{a.s.}{\to} 0$$

Remark 2 If the given process $\{\xi_t\}$ has a bounded variance, that is, $\sigma_n^2 \le \bar{\sigma}^2 < \infty$ and a "double averaged" correlation function R_n, decreasing as $R_n = O(n^{-\varepsilon})$ $(\varepsilon > 0)$, then the conditions of this theorem are fulfilled automatically.

Proof Since for any $n = 1, 2, \ldots$

$$\|S_n\|^2 = \left(1 - \frac{1}{n}\right)^2 \|S_{n-1}\|^2 + v_n \le \left(1 - \frac{1}{n}\right) S_{n-1}^2 + v_n$$

where

$$v_n := 2\frac{1}{n}\left(1 - \frac{1}{n}\right) S_{n-1}^\top \xi_n + \frac{1}{n^2} \|\xi_n\|^2$$

then the back iterations imply

$$S_n^2 \le \pi_n S_1^2 + \pi_n \sum_{t=2}^{n} \pi_t^{-1} v_t$$

with

$$\pi_n := \prod_{t=2}^{n} (1 - t^{-1})$$

By the Kronecker's lemma (see, for example, Appendix in [11]) S_n tends to zero if with probability 1 the following sequence

$$r_n^{(1)} := \sum_{t=1}^{n} v_t$$

converges. To fulfill this, it is sufficient to show that under the conditions of this theorem the series

$$r_n^{(2)} := \sum_{t=1}^{n} \frac{1}{t^2} \|\xi_t\|^2, \quad r_n^{(3)} := \sum_{t=1}^{n} \frac{1}{t} |S_{t-1}^{\mathrm{T}} \xi_t|$$

converge with probability one that is true if

$$\sum_{t=1}^{\infty} \frac{1}{t^2} \sigma_n^2 < \infty, \sum_{t=1}^{\infty} \frac{1}{t} \{|S_{t-1}^{\mathrm{T}} \xi_t|\} < \infty$$

By the Cauchy-Bouniakovski-Shwartz inequality, it follows

$$\sum_{t=1}^{\infty} \frac{1}{t} E\{|S_{t-1}^{\mathrm{T}} \xi_t|\} \leq \sum_{t=1}^{\infty} \frac{1}{t} \sqrt{E\{\|S_{t-1}\|^2\} \sigma_n^2}$$

that together with the identity

$$E\{\|S_{t-1}\|^2\} = R_{t-1}$$

directly leads to the result of this theorem. The theorem is proven. ■

3. IMPORTANT PARTIAL CASES

In this section, two partial cases, most important for the identification and adaptive control applications, are considered in detail.

3.1. The Cramer–Lidbetter Condition

COROLLARY 3 *Assume that the correlation coefficients* $\rho_{t,s}$ (5) *of the given random process* (3) *satisfy the Cramer–Lidbetter's condition* [1], *that is,*

$$|\rho_{t,s}| \leq K \frac{t^{\alpha} + s^{\alpha}}{1 + |t - s|^{\beta}}$$

where K, α, β-nonnegative constants verifying

$$2\alpha < min\{1, \beta\}$$

Then the strong law of large numbers holds, that is,

$$S_n := \frac{1}{n}\sum_{t=1}^{n} \xi_t \overset{a.s.}{\to} 0$$

Proof Since

$$E\{\|\xi_t\|^2\} = \sigma_n^2 = \rho_{t,t} \leq 2Kt^{\alpha}$$

then

$$\begin{aligned} R_n &\leq \frac{K}{n^2}\sum_{t=1}^{n}\sum_{s=1}^{n}\frac{t^{\alpha}+s^{\alpha}}{1+|t-s|^{\beta}} \\ &= \frac{2K}{n^2}\left(\sum_{t=1}^{n} t^{\alpha} + \sum_{t=1}^{n}\sum_{s<t}\frac{t^{\alpha}+s^{\alpha}}{1+(t-s)^{\beta}}\right) \\ &\leq \frac{2K}{n^{1-\alpha}} + 2KI_n \end{aligned}$$

where

$$I_n := \frac{1}{n^2}\sum_{t=1}^{n}\sum_{s<t}\frac{t^{\alpha}+s^{\alpha}}{1+(t-s)^{\beta}} = I_n' + I_n''$$

$$\begin{aligned} I_n' &:= \frac{1}{n^2}\sum_{t=1}^{n} t^{\alpha}\sum_{s<t}\frac{1}{1+(t-s)^{\beta}} \\ &\leq \frac{1}{n^2}\int_0^n t^{\alpha}\left\{\begin{array}{ll} (t^{\beta-1}-1/1-\beta), & \beta \neq 1 \\ \ln t, & \beta = 1 \end{array}\right\} dt \\ &\leq Const\left\{\begin{array}{ll} n^{\alpha-\beta}, & \beta < 1 \\ n^{\alpha+\varepsilon-1}, & \beta = 1 \\ n^{\alpha-1}, & \beta > 1 \end{array}\right\}, \varepsilon > 0 \end{aligned}$$

$$\begin{aligned} I_n'' &:= \frac{1}{n^2}\sum_{t=1}^{n}\sum_{s<t}\frac{s^{\alpha}}{1+(t-s)^{\beta}} \\ &\leq \frac{1}{n^2}\sum_{t=1}^{n} t^{\alpha}\sum_{s<t}\frac{1}{1+(t-s)^{\beta}} = I_n' \end{aligned}$$

So, finally, the following upper estimate for the "double averaged" correlation function R_n holds:

$$R_n \le Const \left\{ \begin{array}{ll} n^{\max\{\alpha-1,\alpha-\beta\}}, & \beta < 1 \\ n^{\alpha+\varepsilon-1}, & \beta = 1 \\ n^{\alpha-1}, & \beta > 1 \end{array} \right\}, \varepsilon > 0$$

The substitution of the right-hand side of the last inequality in (6) implies the result of this corollary. ■

3.2. Dependent Processes Generated by Stable Forming Filters

COROLLARY 4 *Consider a centered random independent vector process* $\{\xi_n\}$ *with finite variances* σ_n^2 *satisfying*

$$\sum_{n \in N^+} \frac{1}{n(n-1)} \sigma_n \sqrt{\sum_{r=0}^{n-1} \sigma_r^2} < \infty$$

and generating the random vector sequence $\{\zeta_n\}$ *according to the following expression*

$$\zeta_n = \sum_{t=0}^{n} h_{n,t} \xi_t$$

where the impulse response matrix function $h_{n,t}$ *for any* $t \le n$ *satisfies*

$$\|h_{n,t}\| \le \hat{h}(n-t)$$
$$H := \sum_{\tau=0}^{\infty} \hat{h}(\tau) < \infty$$

(*such impulse function corresponds to a stable, may be, nonstationary forming filter*). *Then for the random sequence* $\{\zeta_n\}$ *the strong law of large numbers holds, that is,*

$$\frac{1}{n} \sum_{t=1}^{n} \zeta_t \overset{a.s.}{\rightarrow} 0$$

Proof The inequality

$$R_n \le \frac{H^2}{n^2} \sum_{r=0}^{n} \sigma_r^2$$

implies

$$\frac{1}{n}\sqrt{R_{n-1}\sigma_n^2} \leq \frac{H}{n(n-1)}\sigma_n\sqrt{\sum_{r=0}^{n-1}\sigma_r^2}$$

that, together with the accepted assumptions, proves this corollary.

■

4. CONCLUSION

A version of the strong law of large numbers for dependent vector sequences is presented in this paper. The new characteristics of the dependence, called **the "double averaged" correlation function**, provides very simple proof of the main result which generalizes the well-known Cramer–Lidbetter's theorem and gives the simple explanation why the intuitively used fact (that the strong law holds for the class of random processes generated by a non stationary but stable forming filters with an absolutely integrable impulse function) turns out to be true.

References

[1] Cramer, H. and Lidbetter, M. (1969). *Stationary Random Processes*, Princeton: Princeton University Press.
[2] Duncan, T. E., Guo, L. and Pasik-Duncan, B. (1999). Adaptive Continuous-Time Linear Quadratic Gaussian Control, *IEEE Transaction on AC*, **44**(9), 1653–1662.
[3] Hall, P. and Heyde, C. C. (1980). *Martingale Limit Theory and Its Application*, Academic Press, NY.
[4] Iosifescu, M. and Theodorescu, R. (1969). *Random Processes and Learning*, Die Grundlehren der Mathematischen Wissenschaften, Berlin.
[5] Kolmogorov, A. N. (1956). *Foundations of the Theory of Probability*, Chelsea, NY.
[6] Ljung, L. (1987). *System Identification: Theory for the User*, Prentice Hall.
[7] Neveu, J. (1965). *Mathematical Foundations of the Calculus of Probability*, Holden-Day, San Francisco.
[8] Petrov, V. V. (1975). *Sums of Independent Random Variables*, Springer-Verlag, Berlin–NY.
[9] Poznyak, A. S. and Tzypkin, Ya. Z. (1989). The Generalized Instrumental Variables Method in Problems of Linear Plants Identification, *Reports of Academy of Sciences of USSR*, ser. Kibernetica i Teoriya Regulirovaniya, **306**(5), 1068–1072.

[10] Poznyak, A. S. (1992). New Form or Strong Large Numbers Law for Dependent Sequences and its Applications to the Analysis of Recurrent Identification Algorithms, *Reports of Russian Academy of Sciences*, ser. Kibernetica i Teoriya Regulirovaniya, **323**(4), 645–648.

[11] Poznyak, A., Najim, K. and Gomez, E. (2000). *Self-Learning Control of Finite Markov Chains*, Marcel-Dekker, NY.

[12] Stout, W. F. (1974). *About Sure Convergence*, Academic Press, NY.