Limit theorems for Banach-valued autoregressive processes Applications to real continuous time processes

Denis BOSQ

Abstract

We show that a large class of continuous time processes admits a Banach autoregressive representation. This fact allows us to obtain various limit theorems for continuous time processes. In particular we prove the law of iterated logarithm for processes which satisfy a stochastic differential equation.

Résumé

Nous montrons qu'une vaste classe de processus à temps continu possède une représentation autorégressive Banachique. Ceci nous permet d'obtenir des théorèmes limites pour les processus réels à temps continu. Par exemple nous établissons la loi du logarithme itéré pour des processus vérifiant une équation différentielle stochastique.

1 Introduction

The well known interpretation of a continuous time process as a random variable which takes values in a functional space is scarcely used in statistical inference except, for example, as an auxiliary tool in statistics for diffusion processes (see, e.g., LIPTSER and SHIRYAYEV (1977), KHASMINSKII and SKOROKHOD (1996)).

Received by the editors February 1996.

Communicated by M. Hallin.

1991 *Mathematics Subject Classification :* 60B, 60F, 60G, 62M.

Key words and phrases : Probability in Banach spaces, autoregressive processes, law of large numbers, central limit theorem, Ornstein-Uhlenbeck process, stochastic differential equations, law of iterated logarithm.

An outstanding exception is the book by GRENANDER (1981) where that technique is systematically used. However the statistical application of GRENANDER's results requires that independent copies of the observed process sample paths should be available. Other typical examples are ANTONIADIS-BEDER (1989), KUKUSH (1990), LECOUTRE (1990), DIPPON (1993),...

In the current paper we consider a model which allows us to use dependent copies of a process : a more realistic situation. An interesting example should be electricity consumption during n days in succession, which defines n dependent random variables $X_1, \ldots, X_n$ with values in the Banach space $C[0, \ell]$ where ℓ is the length of one day associated with some unit of measure.
These variables are clearly dependent (in particular $X_i(\ell) = X_{i+1}(0)$).

Our Banach autoregressive model takes the simple form

$$X_n = \rho(X_{n-1}) + \varepsilon_n \ , \ \ n \in \mathbb{Z} \ , \tag{1}$$

we obtain various limit theorems for this model and apply them to the asymptotic behaviour of functionals of real continuous time processes. As far as we know these Banach type results are new as are most of the applications. Some results about statistical prediction of these processes can be found in BOSQ (1991,a) and MOURID (1994).

For simulations and some applications see BOSQ (1991,b) and PUMO (1992). An application to prediction of electricity consumption appears in AIELLO et al. (1994). For an application to road traffic we refer to BESSE and CARDOT (1996).

Asymptotic results concerning dependent Banach valued random variables appear in DEHLING (1983), DENISEVSKII (1986), ZHURBENKO and ZUPAROV (1986) and PUMO (1992).

The rest of the paper is organized as follows : generalities on Banach autoregressive processes are given in section 2. In section 3 we provide some examples of continuous time processes which admit a Banach autoregressive representation. Section 4 is devoted to weak and strong laws of large numbers. We show that the strong law is satisfied at an exponential rate under suitable conditions. In section 5 we give necessary and sufficient conditions for the validity of the central limit theorem and of the law of iterated logarithm. Applications to real continuous time processes are considered in section 6. Proofs are postponed until section 7.

We wish to thank S. UTEV who provided us a technical lemma (see section 7) which allowed us to improve the statement of Theorem 4 and to simplify the proofs. We also thank an anonymous referee for useful comments and suggestions.

2 Banach valued autoregressive processes

Let us consider a separable Banach space $(B, \|\cdot\|)$ equipped with its Borel σ-Algebra $\mathcal{B}$. A ***B*-white noise** is a sequence of $\boldsymbol{B}$-valued independent random variables, defined on the same probability space $(\Omega, \mathcal{A}, P)$, having the same distribution and such that

$$0 < \sigma^2 = E \parallel \varepsilon_n \parallel^2 < \infty \ , \quad E\varepsilon_n = 0, \ n \in \mathbb{Z} \ .$$

Let ρ be a bounded linear map from B to B such that $\sum_{n \geq 0} \parallel \rho^n \parallel < \infty$ (with clear notations) and let $m \in B$. We set

$$X_n = m + \sum_{j=0}^{\infty} \rho^j(\varepsilon_{n-j}) \ , \ n \in \mathbb{Z} \ . \tag{2}$$

Then it is easy to prove that the series converges in $L^2_B(\Omega, \mathcal{A}, P)$ and almost surely.

Thus ε_i and $(X_j, j < i)$ are independent, (X_n) is strictly stationary and

$$X_n - m = \rho(X_{n-1} - m) + \varepsilon_n \ \ , \ n \in \mathbb{Z} \ . \tag{3}$$

We will say that $(X_n, \ n \in \mathbb{Z})$ is a *B-valued autoregressive process of order* 1 (ARB (1)). ρ is the correlation operator, (ε_n) the innovation, m the mean.

Note that if $x^* \in B^*$ (the topological dual of B) is an eigenvector of ρ^* (the adjoint of ρ), associated with an eigenvalue $\lambda \in]-1, +1[$ then $(x^*(X_n - m), n \in \mathbb{Z})$ is a real autoregressive process of order 1 (possibly degenerated).

Moreover $EX_0 = m$ and $D = \rho C$ where C is the covariance operator of X_0 and D the cross covariance operator of (X_0, X_1) defined by

$$C(x^*) = E(x^*(X_0)X_1) \ \ , \ x^* \in B^* \ ,$$

and

$$D(x^*) = E(x^*(X_0)X_1) \ \ , \ x^* \in B^* \ .$$

Finally (X_n) is a Markov process and

$$E(X_n \mid X_j, \ j \leq n-1) = \rho(X_{n-1}) \ .$$

3 ARB representation of a real continuous time process

3.1 Representation of the ORNSTEIN-UHLENBECK process

Let us consider the real stationary Gaussian process

$$\xi_t = \int_{-\infty}^{t} e^{-\lambda(t-u)} dw(u) \ \ , t \in \mathbb{R}$$

where w is a bilateral standard Wiener process and λ a positive constant. We choose a version of (ξ_t) such that every sample path is continuous and, in order to obtain an ARB representation we take $B = C[0,1]$ and set

$$X_n(t) = \xi_{n+t} \ , \ t \in [0,1] \ , n \in \mathbb{Z} \ .$$

Now, taking into account that

$$E(\xi_{n+t} \mid \xi_s, s \leq n) = e^{-\lambda t}\xi_n \ , \ t \in [0,1]$$

we may set

$$\rho(x)(t) = e^{-\lambda t}x(1),\ t \in [0,1],\ x \in C[0,1]\ . \tag{4}$$

Then we have

$$\sum_{n\geq 0} \parallel \rho^n \parallel = \sum_{n \geq 0} e^{-(n-1)\lambda} = (1-e^{-\lambda})^{-1}$$

and

$$X_n = \rho(X_{n-1}) + \varepsilon_n \ \ , \ n \in \mathbb{Z}$$

where

$$\varepsilon_n(t) = \int_0^t e^{-\lambda(t-v)}dw(n+v),\ t \in [0,1], n \in \mathbb{Z}\ .$$

Hence, (ξ_t) has an ARB representation.

Now the adjoint ρ^* of ρ is defined by

$$\rho^*(\mu)(x) = x(1)\int_0^1 e^{-\lambda t}d\mu(t),\ \mu \in C^*[0,1]\ .$$

Thus the Dirac measure $\delta_{(1)}$ is the only eigenvector of ρ^* and the corresponding eigenvalue is $e^{-\lambda}$. Consequently $(X_n(1),\ n \in \mathbb{Z})$ is an $AR(1)$ which satisfies

$$X_n(1) = e^{-\lambda}X_{n-1}(1) + e^{-\lambda}\int_0^1 e^{\lambda v}dw(n+v),\ n \in \mathbb{Z}\ .$$

By stationarity $(X_n(t),\ n \in \mathbb{Z})$ is also an $AR(1)$ for every $t \in [0,1]$.

The above facts remain valid with a slight adaptation if we choose $B = C[a,b]$, $a < b$.

3.2 Representation of a stationary Gaussian process solution of a stochastic differential equation

Consider the stochastic differential equation of order k $(k \geq 2)$:

$$\sum_{\ell=0}^{k} a_\ell d\xi^{(\ell)}(t) = dw(t) \tag{5}$$

where $a_0, \ldots, a_k$ are constant, $a_k \neq 0$, and where $w(t)$ is a bilateral Wiener process. In (5) differentiation up to the order $k-1$ is ordinary when the order k derivative is defined in the ITO sense (cf. ASH and GARDNER (1975)).

We suppose that the roots $-\lambda_1, \ldots, -\lambda_k$ of the equations $\sum_{\ell=0}^{k} a_\ell \lambda^\ell = 0$ are real and such that $-\lambda_k < \ldots < -\lambda_1 < 0$.

Then by using theorem 2.8.2 in ASH and GARDNER (1975) we may assert that the only stationary solution of (5), is the Gaussian process

$$\xi_t = \int_{-\infty}^{t} g(t-u)dw(u) \ , \ t \in \mathbb{R} \tag{6}$$

where g is the GREEN's function of (5), that is, $g(t) = 0$ for $t < 0$ and, for $t \geq 0$, $g(t)$ is the unique solution of the problem

$$\begin{cases} \displaystyle\sum_{\ell=0}^{k} a_\ell x^{(\ell)}(t) = 0 \quad , & (7) \\ x(0) = \ldots = x^{(k-2)}(0) = 0 \quad , \quad x^{(k-1)}(0) = a_k^{-1} \quad . & (8) \end{cases}$$

We choose a version of (ξ_t) such that *every* sample path of (ξ_t) has $k-1$ continuous derivatives.

Now by using again ASH and GARDNER (1975, pp. 110-111) we obtain

$$E(\xi_{n+t} \mid \xi_s \leq n) = \sum_{j=0}^{k-1} \xi^{(j)}(n)\varphi_j(t), \ n \in \mathbb{Z}, \ t \in [0,1] \ , \tag{9}$$

where φ_j is the unique solution of (5) which satisfies

$$\varphi_j^{(\ell)}(0) = \delta_{j\ell} \ ; \ \ell = 0, \ldots, k-1 \ .$$

Then we define ρ on $B = C_{k-1}[0,1]$ by

$$\rho(x)(t) = \sum_{j=0}^{k-1} x^{(j)}(1)\varphi_j(t), \ t \in [0,1], \ x \in C_{k-1}[0,1] \ . \tag{10}$$

It is easy to prove that the only eigenelements of ρ are $\left(e^{-\lambda_r t}, e^{-\lambda_r}\right)$, $r = 1, \ldots, k$ and that

$$\| \rho^p \| = O\left(e^{-(p-1)\lambda_1}\right) \ .$$

Finally

$$X_n = \rho(X_{n-1}) + \varepsilon_n$$

where

$$X_n(t) = \xi_{n+t} \ , \ t \in [0,1], \ n \in \mathbb{Z}$$

and

$$\varepsilon_n(t) = \int_{n}^{n+t} g(n+t-u)dw(u), \ t \in [0,1], \ n \in \mathbb{Z} \ .$$

Note that (ξ_t) is not Markovian whereas (X_n) is a Markov process.

3.3 Process with seasonality

Let $(\eta_t, \ t \in \mathbb{R})$ be a real process such that

$$\eta_t = m(t) + \xi_t \quad , \ t \in \mathbb{R}$$

where (ξ_t) is a zero mean process admitting an ARB representation in B, a separable Banach space of real functions defined over $[0, \tau]$ and where $m(\cdot)$ is a non constant deterministic function with period τ and such that $t \mapsto m(t)$, $0 \le t \le 1$ belongs to B. Then (η_t) has clearly an ARB representation.

Note that (η_t) is not stationary when (X_n) is strictly stationary.

3.4 Hilbert-valued autoregressive processes

Let $(Z_t,\ t \in \mathbb{R})$ be a locally square integrable real zero mean process with independent increments. Then

$$\varepsilon_n(t) = Z_{n+t} - Z_n \ ,\ t \in [0,1],\ n \in \mathbb{Z}$$

defines an B-white noise, where $B = L^2[0,1]$.

Now let ρ be a linear operator defined by

$$(\rho x)(t) = \int_0^1 K(s,t)x(s)ds,\ t \in [0,1],\ x \in H$$

where $\int_{[0,1]^2} K^2(s,t)dsdt < 1$. Then $\| \rho \| < 1$ and $X_n(t) = \sum_{j \ge 0} \rho^j(\varepsilon_{n-j})(t)$, $0 \le t \le 1$, $n \in \mathbb{Z}$ is an ARB process.

4 Laws of large numbers

In the sequel we set $S_n = \sum_{i=1}^{n} X_i$, $\overline{\varepsilon}_n = \frac{1}{n}\sum_{i=1}^{n} \varepsilon_i$, $R = \sum_{n \ge 0} \| \rho^n \|$.

Theorem 1 *(Strong law)*
Let (X_n) be an ARB with mean m, then

(11) $$\overline{X}_n \to m \quad a.s. \ .$$

Furthermore if B is of type 2 and if (ε_n) satisfies

(12) $$E \| \varepsilon_n \|^k \le \frac{k!}{2} c^{k-2} \sigma^2 \ ,\ k \ge 2$$

where c is constant, then, for each $\eta > 0$

(13) $$P(\| \overline{X}_n - m \| > \eta) \le \alpha \exp(-\beta n)$$

where α and β are positive constants.

In order to state a weak law of large numbers we recall that a linear operator $\ell : B \mapsto B$ is said to be coercive with constant r if

$$\| \ell(x) \| \ge r \| x \| \ ,\ x \in B \ .$$

Now we have the following

Theorem 2 *(Weak law)*
Let (X_n) be an ARB with mean m

a) If B is of type p where $1 < p \leq 2$ with type constant c_1 then

$$E \parallel \overline{X}_n - m \parallel^p \leq \frac{(2c_1 R)^p}{n^{p-1}} E(\parallel \varepsilon_0 \parallel^p + \parallel \rho(X_0 - m) \parallel^p) \ . \tag{14}$$

b) If B is of cotype q where $2 \leq q < \infty$ with cotype constant c_2 and if the operators $\sum_{j=0}^{n} \rho^j$, $n \geq 0$ are coercive with the same constant r, then

$$E \parallel \overline{X}_n - m \parallel^q \geq \left(\frac{r}{2c_2}\right)^q \frac{1}{n^{q-1}} E(\parallel \varepsilon_0 \parallel^q) \ . \tag{15}$$

c) In particular if B is an Hilbert space and if ρ is symmetric compact then $\parallel \rho \parallel < 1$ and

$$\begin{aligned} \frac{1 - \parallel \rho \parallel}{1 + \parallel \rho \parallel} \frac{1}{n} E \parallel \varepsilon_0 \parallel^2 &\leq E \parallel \overline{X}_n - m \parallel^2 \\ &\leq \frac{1}{(1 - \parallel \rho \parallel)^2} \frac{1}{n} E \parallel X_0 - m \parallel^2 \ , \end{aligned} \tag{16}$$

furthermore the inequalities in (16) are equalities if and only if ρ vanishes.

For the definition of types and cotypes of a Banach space we refer to HOFFMAN and JORGENSEN (1973), MAUREY (1973) and LEDOUX and TALAGRAND (1991).

5 Central limit theorem and law of iterated logarithm

The central limit theorem for an ARB process has the following form :

Theorem 3 *(CLT)*
An ARB process (X_n) with mean m satisfies the CLT if and only if the innovation (ε_n) satisfies it and in that case

$$\sqrt{n}\left(\overline{X}_n - m\right) \xrightarrow{\mathcal{L}} N \sim \mathcal{N}(0, (I - \rho)^{-1} C_\varepsilon (I - \rho)^{*-1} \tag{17}$$

where C_ε denotes the covariance operator of ε_0.

Note that each of the following conditions implies the CLT for (ε_n) and thus for (X_n) :

a) B is of type 2, in particular B is a Hilbert space.

b) B is of cotype 2 and ε_0 is pregaussian.

c) $B = C[0, 1]$ and ε_0 is subgaussian.

d) $B = C[0,1]$ and there exists a square integrable r.v. M such that

$$|\varepsilon_0(t,\omega) - \varepsilon_0(s,\omega)| \leq M(\omega)|t-s| \ , \ \omega \in \Omega; \ s,t \in [0,1] \ .$$

Details about these results may be found in LEDOUX and TALAGRAND (1991).

In order to state the law of iterated logarithm we set

$$\begin{aligned} \operatorname{Log} n &= \max(1, \ell n\ n), \\ d(x,A) &= \inf\{\| x-y \|, \ y \in A\} \ , \ x \in B \ , \ A \subset B \ , \end{aligned}$$

and we denote by $C(x_n)$ the set of limit points of the sequence (x_n).

Theorem 4 *(LIL)*
Let (X_n) be a zero mean ARB. Then (X_n) satisfies the compact LIL if and only if (ε_n) satisfies it. In this case we have

$$\lim_{n\to\infty} d\left(\frac{S_n}{\sqrt{2n\log\log n}} \ , \ (I-\rho)^{-1}K\right) = 0 \quad a.s. \tag{18}$$

and

$$C\left(\frac{S_n}{\sqrt{2n\log\log n}}\right) = (I-\rho)^{-1}K \quad a.s. \tag{19}$$

where K is the closed unit ball of the reproducing Hilbert space associated with C_ε :

$$K = \{x \in B \ , \ x = E(\xi\varepsilon_0), \ \xi \in L^2(\Omega,\mathcal{A},P), \ E\xi = 0, \ E\xi^2 \leq 1\} \tag{20}$$

The compact LIL for Banach valued i.i.d. r.v.'s is due to KUELBS (1977). Note also that each condition (a), (b), (c), (d) above implies the LIL for (ε_n). Details and other references appear in LEDOUX and TALAGRAND (1991).

6 Some applications to real continuous time processes

The above results allow us to obtain limit theorems for continuous time process. In this section we provide some examples.

Let $(\xi_t, \ t \in \mathbb{R})$ be a measurable process with admit an ARB representation with $B = C[0,1]$. For every signed bounded measure μ on $[0,1[$ we define

$$\mu_n(A) = \sum_{j=0}^{n-1} \mu(A(j)), \ A \in \mathcal{B}_{[0,1[}$$

with $A_{(j)} = \{x : x+j \in A \cap [j,j+1[\}$.

Then we have strong laws for (ξ_t) :

Corollary 1 *For every signed bounded measure μ on* $]0,1]$

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \xi_t d\mu_{[T]}(t) = \int_0^T E\xi_t d\mu_{[T]}(t) \ \ a.s. \tag{21}$$

In order to state a CLT we consider the eigenelements (e_j,λ_j), $j\geq 1$ of the kernel $E(\xi_s\xi_t)$ defined by

$$\lambda_j e_j(t) = \int_0^1 E(\xi_s\xi_t)e_j(s)ds,\ t\in[0,1],\ j\geq 1$$

with $\int_0^1 e_j^2(t)dt = 1,\ j\geq 1.$

Corollary 2 *If (ξ_t) is zero mean, if $\lambda_1,\ldots,\lambda_k$ are distinct and if there exists a square integrable r.v. L such that*

$$|\xi(t,\omega)-\xi(s,\omega)| \leq L(\omega)|t-s|\ ;\ s,t\in[0,1],\ \omega\in\Omega \tag{22}$$

then

$$\frac{1}{\sqrt{n}}\left(\int_0^1\left(\sum_{j=0}^{n-1}\xi_{j+t}\right)e_j(t)dt,\ 1\leq j\leq k\right) \xrightarrow{\mathcal{L}} N_k \tag{23}$$

where N_k is a zero mean k dimensional Gaussian vector with covariance matrix

$$\begin{pmatrix} \lambda_1 & O \\ & \\ O & \lambda_k. \end{pmatrix}$$

Note that condition (22) could be replaced by any assumption which ensures that (X_n) satisfies the CLT.

Corollary 3 *If (ξ_t) is Gaussian then*

$$\lim_{T\to\infty} d\left(\frac{1}{\sqrt{2T\log\log T}}\int_0^T \xi_t d\mu_{[T]}(t), K'\right) = 0 \ \ a.s. \tag{24}$$

and

$$C\left(\frac{1}{\sqrt{2T\log\log T}}\int_0^T \xi_t d\mu_{[T]}(t)\right) = K' \quad a.s. \tag{25}$$

where

$$K' = \left\{\int_0^1 y(t)d\mu(t),\ y\in(I-\rho)^{-1}K\right\}. \tag{26}$$

In particular if (ξ_t) is the ORNSTEIN-UHLENBECK process and if μ is the Lebesgue measure on $[0,1]$ then

$$K' = \left[-\frac{1}{c}\ ,\ +\frac{1}{c}\right]\ .$$

For related results concerning the ORNSTEIN-UHLENBECK process we refer to STOICA (1992).

More generally corollary 3 is valid if (ξ_t) satisfies (5).

7 PROOFS

We first state and prove UTEV's technical lemma.

Lemma 1 *(S. UTEV)*
Let (Y, Y_k) be an equidistributed sequence of nonnegative r.v.'s with finite second moment and (a_k) be a summable positive sequence.
Then

$$I_n[(Y_k)] := \frac{1}{\sqrt{n}} \sum_{j=1}^{n} a_{n-j} Y_j \longrightarrow_{n\to\infty} 0 \quad a.s. \tag{27}$$

Proof :
We observe that there exists a function $0 \le g(x) \uparrow +\infty$ as $x \to \infty$ such that

$$E[Y^2 g(Y)] < \infty \ ,$$

whence there exists a nonnegative sequence $0 \le t_k \downarrow 0$ such that

$$\sum_{k=1}^{\infty} P(Y \ge t_k \sqrt{k}) < \infty \ .$$

Let $X_k = Y_k 1_{(Y_k < t_k \sqrt{k})}$. By the Borel-Cantelli lemma $P(X_k \ne Y_k \text{ i.o.}) = 0$. Thus it remains to prove (27) with (Y_k) replaced by (X_k) i.e. for $I_n[(X_k)]$. By construction

$$I_n[(X_k)] \le \frac{1}{\sqrt{n}} \sum_{j=1}^{n} (a_{n-j} t_j \sqrt{j})$$

$$= \frac{1}{\sqrt{n}} \left[\sum_{j=1}^{[\sqrt{n}]} (a_{n-j} t_j \sqrt{j}) \right] + \frac{1}{\sqrt{n}} \left[\sum_{j=[\sqrt{n}]+1}^{n} (a_{n-j} t_j \sqrt{j}) \right]$$

$$\le \frac{1}{\sqrt{n}} (1 + n^{1/4}) \left(\sum_{j=0}^{\infty} a_j \right) t_1 + \frac{1}{\sqrt{n}} (\sqrt{n}) \left(\sum_{j=0}^{\infty} a_j \right) t_{[\sqrt{n}]}$$

$\to 0$ as $n \to \infty$. ■

In the sequel we may and do suppose that $m = 0$.

Proof of theorem 1

1) First suppose that $\| \rho \| < 1$. Then (1) and $(I-\rho)^{-1} = \sum_{j \ge 0} \rho^j$ yield

$$\overline{X}_n = (I-\rho)^{-1} \overline{\varepsilon}_n \quad - \tfrac{1}{n} (I-\rho)^{-1} (\rho^n \varepsilon_1 + \ldots + \rho \varepsilon_n) + \tfrac{1}{n} (\rho + \ldots + \rho^n)(X_0). \tag{28}$$

We treat each term separately.

By using the strong law of large numbers for i.i.d. r.v.'s (cf. LEDOUX and TALAGRAND 1991, p. 189) and the continuity of $(I-\rho)^{-1}$ we obtain

$$(I-\rho)^{-1} \overline{\varepsilon}_n \to 0 \quad \text{a.s.}$$

On the other hand, from UTEV's lemma we infer that

$$\begin{aligned} &\| \frac{1}{n}(I-\rho)^{-1}(\rho^n\varepsilon_1 + \ldots + \rho\varepsilon_n) \| \\ &\quad \leq \| (I-\rho)^{-1} \| \frac{1}{n}(\| \rho^n \| \| \varepsilon_1 \| + \ldots + \| \rho \| \| \varepsilon_n \|) \to 0 \ \text{ a.s} \quad \text{as } \ n \to \infty \,. \end{aligned}$$

Finally

$$\frac{1}{n} \| (\rho + \ldots + \rho^n) X_0 \| \leq \frac{R \| X_0 \|}{n} \to 0 \,.$$

Collecting the above results we get $\overline{X}_n \to 0$ a.s. .

1. If $\| \rho \| \geq 1$ the condition $\sum \| \rho^n \| < \infty$ implies the existence of an integer n_0 such that

$$\| \rho^n \| < 1 \quad \text{for} \quad n \geq n_0 \,.$$

Now we have

$$X_{kn_0+\ell} = \sum_{j=0}^{n_0-1} \rho^j (\varepsilon_{kn_0+\ell-j}) + \rho^{n_0}(X_{(k-1)n_0+\ell}) \,,$$

$k \in \mathbb{Z}$; $\ell = 0, 1, \ldots, n_0 - 1$.

Thus to each ℓ we can associate an ARB process

$$\left(Y_n^\ell, \ n \in \mathbb{Z}\right) := (X_{nn_0+\ell}, \ n \in \mathbb{Z})$$

with autocorrelation operator ρ^{n_0} an innovation $\left(\sum_{j=0}^{n_0-1} \rho^j \left(\varepsilon_n n_0 + \ell - j \right), \ n \in \mathbb{Z} \right)$.

These ARB processes are connected with (X_n) by the identity

$$\overline{X}_n = \frac{q_n}{n_0 q_n + r_n} \left(\sum_{r=0}^{n_0-1} \frac{1}{q_n} \sum_{j=0}^{q_n'} Y_j^r \right) \,,$$

where $n = n_0 q_n + r_n$, $0 \leq r_n \leq n_0 - 1$ and $q_n' = q_n$ or $q_n + 1$.

As $q_n \to \infty$ the first part of the proof implies $\frac{1}{q_n} \sum_{j=0}^{q_n'} Y_j^\ell \to 0$ a.s. so that $\overline{X}_n \to 0$ a.s. and the strong law of large numbers is thus established.

3) We now turn to the proof of (13). For this purpose we again use decomposition (28).

Under conditions (12) YURINSKI (1976)'s inequality applies to (ε_n) : if $\overline{\eta} = \eta - \frac{a_n}{b_n} > 0$ where $a_n = E \| \varepsilon_1 + \ldots + \varepsilon_n \|$ and $b_n = \left(\sum_{j=1}^n E \| \varepsilon_j \|^2 \right)^{1/2} = \sigma\sqrt{n}$ then

$$P\left(\left\| \sum_{j=1}^n \varepsilon_j \right\| \geq \eta\sigma\sqrt{n} \right) \leq \exp\left(-\frac{\overline{\eta}^2}{8} \frac{1}{1 + \overline{\eta}c/2\sigma\sqrt{n}} \right) \,.$$

Since B is of type 2 (say, with constant c_1) we have

$$a_n \leq \left(E \left\| \sum_{j=1}^{n} \varepsilon_j \right\|^2 \right)^{1/2} \leq 2c_1 \sigma \sqrt{n} \ ,$$

thus the inequality is valid if $\eta > 2c_1$. Now for any $\gamma > 0$ we can choose $\eta = \dfrac{\gamma}{\sigma}\sqrt{n}$ and for n large enough we obtain

$$P(\| \overline{\varepsilon}_n \| \geq \gamma) \leq \exp(-\beta_1 n)$$

where β_1 is a strictly positive constant. It follows that, for any $\gamma > 0$,

$$P(\| (I-\rho)^{-1}\overline{\varepsilon}_n \| \geq \gamma) \leq \exp(-\beta_2 n)$$

where $\beta_2 > 0$.

Similarly it is easy to check that

$$P(\| (I-\rho)^{-1}\overline{\varepsilon}_n \| \geq \gamma) \leq \exp(-\beta_3 n) \ \ , \ \beta_3 > 0 \ .$$

Now if there exists $\lambda > 0$ such that $E(e^{\lambda \| X_0 \|}) < \infty$ we obtain by standard manipulations

$$P\left(\left\| \frac{1}{n}(\rho + \ldots + \rho^n) X_0 \right\| \geq \gamma \right) \leq e^{-\frac{\lambda\gamma}{R} n} E\left(e^{\lambda \| X_0 \|} \right) \ .$$

It remains to show the existence of λ. For this aim it suffices to write $\| X_0 \| \leq \sum_{j=0}^{\infty} \| \rho^j \| \| \varepsilon_j \|$ and to choose $\lambda < \left(c \max_{j\geq 0} \| \rho^j \| \right)^{-1}$, then using the exponential power series we get

$$E\left(e^{\lambda \| \rho^j \| \| \varepsilon_j \|} \right) \leq \exp\left(\lambda \| \rho^j \| + \frac{\lambda^2 \| \rho^j \|^2 \sigma^2}{1-\lambda c M} \right)$$

where $M = \max_{j\geq 0} \| \rho^j \|$. Therefore

$$E\left(e^{\lambda \| X_0 \|} \right) \leq \exp\left(\lambda \sum_{j\geq 0} \| \rho^j \| + \frac{\lambda^2 \sigma^2}{1-\lambda c M} \sum_{j\geq 0} \| \rho^j \|^2 \right)$$

and the proof is complete. ■

Proof of theorem 2

By iterating (1) we get

$$X_k = \sum_{j=0}^{k-1} \rho^j(\varepsilon_{k-j}) + \rho^k(X_0) \ , \ \ k \geq 1 \tag{29}$$

thus

$$\overline{X}_n = \frac{1}{n} \sum_{j=0}^{n-1} (I - \rho + \ldots + \rho^j)\varepsilon_{n-j} + \frac{1}{n}(\rho + \ldots + \rho^n)(X_0) \ .$$

Then if B is of type p with constant c_1 the independence of $X_0, \varepsilon_1, \ldots, \varepsilon_n$ together with proposition 9.11 in LEDOUX and TALAGRAND (1991, p. 248) imply

$$\begin{aligned} E \parallel \overline{X}_n \parallel^p &\leq \frac{(2c_1)^p}{n^p}\left[\left(\sum_{j=0}^{n-1} E \parallel (I+\ldots+\rho^j)\varepsilon_{n-j} \parallel^p\right) + E \parallel (\rho+\ldots+\rho^n)X_0 \parallel^p\right] \\ &\leq \frac{(2c_1R)^p}{n^p}(nE \parallel \varepsilon_0 \parallel^p + E \parallel \rho X_0 \parallel^p) \end{aligned}$$

which proves (14).

If B is of cotype q with constant c_2, we have similarly $E \parallel \overline{X}_n \parallel^q \geq \frac{(2c_2)^{-q}}{n^q}[\sum_{j=0}^{n-1} E(\parallel$ $(I+\ldots+\rho^j)\varepsilon_{n-j} \parallel^q)$

$$+E \parallel (I+\ldots+\rho^{n-1})(\rho X_0) \parallel^q]$$

and by uniform coercivity

$$\begin{aligned} E \parallel \overline{X}_n \parallel^q &\geq \frac{(2c_2)^{-q}r^q}{n^q}\left[\sum_{j=0}^{n-1} E \parallel \varepsilon_{n-j} \parallel^q + E \parallel \rho(X_0) \parallel^q\right] \\ &\geq \left(\frac{r}{2c_2}\right)^q \frac{1}{n^{q-1}} E \parallel \varepsilon_0 \parallel^q \end{aligned}$$

which is (15).

Now assume that B is a Hilbert space equipped with its scalar product $< \cdot\,,\, \cdot >$. If ρ is symmetric compact it has a spectral expansion (cf. AKHIEZER and GLAZMAN (1981)) :

$$\rho(x) = \sum_{j=0}^{\infty} \alpha_j < x,\ \psi_j > \psi_j \quad ,\ x \in B$$

with $\rho(\psi_j) = \alpha_j \psi_j\ ,\ j \geq 0$ and where (ψ_j) is an orthonormal basis of B. Furthermore $\parallel \rho \parallel = |\alpha_0| \geq |\alpha_1| \geq \ldots$

Since $\sum_{n\geq 0} \parallel \rho^n \parallel = \sum_{n\geq 0} |\alpha_0|^n < \infty$ we get $\parallel \rho \parallel < 1$.

Moreover

$$\begin{aligned} \left\| \left(\sum_{j=0}^{n} \rho^j\right)(x) \right\|^2 &= \sum_{j=0}^{\infty} \left(1+\alpha_j+\ldots+\alpha_j^n\right)^2 < x,\ \psi_j >^2 \\ &\geq \left(\frac{1- \parallel \rho \parallel}{1+ \parallel \rho \parallel}\right)^2 \parallel x \parallel^2\ ,\quad n \geq 0 \end{aligned}$$

which means that the operators $\sum_{j=0}^{n} \rho^j$, $n \geq 0$ are uniformly coercive with constant $r = \frac{1- \parallel \rho \parallel}{1+ \parallel \rho \parallel}$.

Since a Hilbert space is of type and cotype 2 we may apply (14) and (15) with $2c_1 = 2c_2 = 1$, and noting that $E \parallel X_0 \parallel^2 = E \parallel \varepsilon_0 \parallel^2 + E \parallel \rho(X_0) \parallel^2$ we obtain (16).

Finally if $\rho = 0$ then $X_0 = \varepsilon_0$ and inequalities (16) are equalities. Conversely equalities in (16) entail

$$1 \geq \frac{(1- \parallel \rho \parallel)^3}{1+ \parallel \rho \parallel} = \frac{E \parallel \rho(X_0) \parallel^2 + E \parallel \varepsilon_1 \parallel^2}{E \parallel \varepsilon_1 \parallel^2} \geq 1$$

which is possible only if $\rho = 0$. ■

Proof of Theorem 3

Consider the decomposition

$$\sqrt{n}\overline{X}_n = (I - \rho)^{-1}\sqrt{n}\overline{\varepsilon}_n + \Delta_n$$

where

$$\Delta_n = (I - \rho)^{-1}\frac{\rho^n\varepsilon_1 + \ldots + \rho\varepsilon_n}{\sqrt{n}} + \frac{(\rho + \ldots + \rho^n)(X_0)}{\sqrt{n}} \,. \tag{30}$$

Using UTEV's lemma we obtain

$$\Delta_n \longrightarrow 0 \quad \text{a.s.}$$

Then, if

$$\sqrt{n}\overline{\varepsilon}_n \xrightarrow{\mathcal{L}} \quad N_1 \sim \mathcal{N}(O, C_\varepsilon)$$

the continuity of $(I - \rho)^{-1}$ implies

$$(I - \rho)^{-1}\sqrt{n}\overline{\varepsilon}_n \xrightarrow{\mathcal{L}} \quad (I - \rho)^{-1}N_1$$

hence

$$\sqrt{n}\overline{X}_n \xrightarrow{\mathcal{L}} \quad (I - \rho)^{-1}N_1 \,.$$

(cf. BILLINGSLEY (1968), Theorem 4.4, p. 27).

Conversely if $\sqrt{n}\overline{X}_n$ satisfies the CLT, the asymptotic normality of $(I-\rho)^{-1}\sqrt{n}\overline{\varepsilon}_n$ follows by using the same method. Since $(I - \rho)$ is continuous we can conclude that $\sqrt{n}\overline{\varepsilon}_n$ satisfies the CLT. ■

Proof of Theorem 4

We set $u_n = \sqrt{2n \log\log n}$ and consider the decomposition

$$\frac{S_n}{u_n} = (I - \rho)^{-1}\frac{\sum_{j=1}^{n}\varepsilon_j}{u_n} + \frac{\sqrt{n}}{u_n}\Delta_n$$

where Δ_n is defined by (30). Thus from the proof of Theorem 3 it follows that $\dfrac{\sqrt{n}}{u_n}\Delta_n \to 0$ a.s..

Now if (ε_n) satisfies the compact LIL (cf. LEDOUX and TALAGRAND (1991), p. 210) we have (18) and (19) since $(I - \rho)^{-1}$ is one-one.

Conversely if there exists a compact set K_1 such that $d\left(\frac{S_n}{u_n}, K_1\right) \longrightarrow 0$ a.s. and $C(\frac{S_n}{u_n}) = K_1$, then $d\left(\frac{\sum_{j=1}^n \varepsilon_j}{u_n}, (I-\rho)K_1\right) \to 0$ a.s. and $C\left(\frac{\sum_{j=1}^n \varepsilon_j}{u_n}\right) = (I-\rho)$. Furthermore $K := (I-\rho)K_1$ is compact and, by LEDOUX and TALAGRAND (1991)'s Theorem 8.5 (p. 210), K is the unit ball of the reproducing kernel Hilbert space associated with C_{ε_0}. ■

Proof of Corollary 1

First we have

$$\begin{aligned} \left|\tfrac{1}{T}\int_{[T]}^{T} \xi_t d\mu_{[T]}(t)\right| &\leq \parallel \mu \parallel \tfrac{X_{[T]}}{T} \\ &\leq \parallel \mu \parallel \tfrac{[T]}{T} \left\|\tfrac{S_{[T]}-S_{[T-1]}}{[T]}\right\| \end{aligned}$$

which tends to zero a.s. by Theorem 1.

$$\begin{aligned} \left|\tfrac{1}{T}\int_{0}^{[T]} \xi_t d\mu_{[T]}(t)\right| &= \left|\int_0^1 \left(\tfrac{1}{T}\sum_{j=0}^{[T-1]} \xi_{j+t}\right) d\mu(t)\right| \\ &\leq \parallel \mu \parallel \tfrac{[T]}{T} \left\|\tfrac{1}{[T]}\sum_{j=0}^{[T-1]} X_0\right\| \end{aligned}$$

which tends to zero a.s. again by Theorem 2. ■

Proof of Corollary 2

To begin we show that (X_n) satisfies the CLT. For this aim we write

$$\varepsilon_{n+t} - \varepsilon_{n+s} = (\xi_{n+t} - \xi_{n+s}) - E(\xi_{n+t} - \xi_{n+s} \mid \xi_u, u \leq n)$$

Using (22) and the monotonicity of conditional expectation we obtain

$$(|\xi_{n+t} - \xi_{n+s}| \ \mid \xi_u,\ u \leq n) \leq |t-s| E(|L| \mid \xi_u,\ u \leq n)$$

thus

$$|\varepsilon_{n+t} - \varepsilon_{n+s}| \leq [|L| + E(|L| \mid \xi_u, u \leq n)]|t-s|$$

which means that (ε_n) satisfies condition d in section 5. Then, by Theorem 3, (X_n) satisfies the CLT.

We now consider the bounded signed measures on $[0, 1[$ defined by

$$d\mu_i(t) := e_i(t)dt \ ; i = 1, \ldots, k \ ,$$

and we note that

$$\int_0^1 \left(\sum_{j=0}^{n-1} \xi_{j+t}\right) e_j(t)dt = \mu_i\left(\sum_{j=0}^{n-1} X_j\right) \ ; \ i = 1, \ldots, k \ ;$$

hence the result by a classical continuity argument. ■

Proof of Corollary 3

First note that

$$\left|\frac{1}{\sqrt{2T\log\log T}}\int_{[T]}^{T}\xi_t d\mu_{[T]}(t)\right| \leq \parallel \mu \parallel \frac{\parallel X_{[T]} \parallel}{\sqrt{2[T]\log\log[T]}}$$

then, by using UTEV's lemma with $a_1 = 1,\ a_j = 0,\ j > 1$ and $Y_j = \parallel X_j \parallel,\ j \geq 1$, we see that the bound tends to zero a.s.

Consequently we may and do suppose that T is an integer. Now we have

$$\frac{1}{u_T}\int_0^T \xi_t d\mu_T(t) = \int_0^1 \left(\frac{1}{u_T}\sum_{j=1}^{T-1} X_j(t)\right) d\mu(t)$$

where $u_T = (2T\log\log T)^{-1/2}$. Then, since $\mu \in C^*[0,1]$, (24) is a straightforward consequence of (18) in Theorem 4. It remains to prove (25).

For that purpose we consider $\Omega_0 \in \mathcal{A}$ such that $P(\Omega_0) = 1$ and $d(Z_n(\omega), K_1) \to 0$, $C(Z_n(\omega)) = K_1,\ \omega \in \Omega_0$ where $Z_n = \dfrac{1}{u_n}\displaystyle\sum_{i=0}^{n-1} X_i$ and $K_1 = (I-\rho)^{-1}K$.

First it is obvious that $C(\mu(Z_n(\omega)) \supset \mu(K_1),\ \omega \in \Omega_0$.

Conversely if y is a limit point of $\mu(Z_n(\omega))$ where $\omega \in \Omega_0$, then there exists a subsequence $Z_{n'}(\omega)$ such that $\mu(Z_{n'}(\omega))$ converges to y. Since $(Z_{n'}(\omega))$ is relatively compact we may extract a new subsequence $(Z_{n''}(\omega))$ which converges to $z \in K_1$.

Therefore $\mu(Z_{n''}(\omega))$ converges to $\mu(z)$, thus $y = \mu(z)$ and finally $y \in \mu(K_1) = K'$, hence (25).
We now turn to the case of an ORNSTEIN-UHLENBECK process. Recall that here $\rho(x)(t) = e^{-\lambda t}x(1),\ 0 \leq t \leq 1$, hence

$$\begin{aligned}[(I-\rho)^{-1}(x)](t) &= \sum_{j=0}^{\infty}(\rho^j(x))(t)\\ &= x(t) + \tfrac{e^{-\lambda t}}{1-e^{-\lambda}}x(1)\ .\end{aligned}$$

Consequently

$$K' = \left\{\int_0^1 x(t)dt + \frac{x(1)}{1-e^{-\lambda}}\int_0^1 e^{-\lambda t}dt\ ,\ x \in K\right\}$$

where K is the unit closed ball of the reproducing kernel Hilbert space associated with $(\xi_t,\ 0 \leq t \leq 1)$. Now

$$\varepsilon_0(t) = \xi_t - e^{-\lambda t}\xi_0$$

and

$$\begin{gathered}K = \{x \in C[0,1],\ x(t) = E(Z\varepsilon_0(t)),\ 0 \leq t \leq 1\ ,\\ Z \in L^2(\Omega,\mathcal{A},P),\ EZ = 0,\ EZ^2 \leq 1\}\ .\end{gathered}$$

Then, K' may be written under the form

$$K' = \{E\left[Z\left(\int_0^1 \xi_t dt + \frac{\xi_1 - \xi_0}{\lambda}\right)\right], Z \in L^2(\Omega, \mathcal{A}, P),$$

$$EZ = 0,\ EZ^2 \leq 1\}.$$

Clearly the maximum is reached if

$$\begin{cases} Z = \alpha\left(\int_0^1 \xi_t dt + \dfrac{\xi_1 - \xi_0}{\lambda}\right) \\ EZ^2 = 1 \\ \alpha > 0 \end{cases}$$

hence $\sup K' = \left[E\left(\int_0^1 \xi_t dt + \frac{\xi_1 - \xi_0}{\lambda}\right)^2\right]^{1/2}$ and after elementary calculations we get $\sup K' = \frac{1}{\lambda}$. Obviously $\inf K' = -\sup K' = -\frac{1}{\lambda}$ and the proof is complete since K' is convex. ■

References

[1] AKHIEZER N.I. and GLAZMAN, I.M. (1981). *Theory of linear operators in Hilbert space.* I. Pitman, London.

[2] AIELLO, S. and CAVALLINI, A. and MONTANARI, G.C. (1994) preprint University of BOLOGNA.

[3] ANTONIADIS, A. and BEDER, S.H. (1989). Joint estimation of the mean and the covariance of Banach valued Gaussian vector. *Statistics*, 20, 1, 77-93.

[4] ASH, R.B. and GARDNER, M.F. (1975). *Topics in stochastic processes.* Academic Press, New York.

[5] BESSE, Ph. et CARDOT, H. (1996). Approximation spline de la prévision d'un processus fonctionnel autorégressif d'ordre 1. Preprint.

[6] BILLINGSLEY, P. (1968). *Convergence of probability measures.* Wiley, New York.

[7] BOSQ, D. (1991 a). Modelization, nonparametric estimation and prediction for continuous time processes, in *Nonparametric functional estimation and related topics*, G. ROUSSAS, Ed., Kluwer Acad., Dordrecht 509-529.

[8] BOSQ D. (1991 b). Prévision d'une fonction aléatoire. *Statistica* LI, 2, 153-164.

[9] CRAMÉR, H. and LEADBETTER, M.R. (1967). *Stationary and related stochastic processes.* Wiley, New York.

[10] DEHLING, H. (1983). Limit theorems for sums of weakly dependent Banach valued random variables. *Z. Wahrscheinlichktstheorie verw. Geb.* 63, 393-432.

[11] DENISEVSKII, N.A. (1986). The strong law of large numbers for a linear process in Banach space (in Russian) *Dotkl. Akad. Nauls. Ukrain SSR*, A, 12, 57-61.

[12] DIPPON, J. (1993). Asymptotic confidence spheres in Banach spaces via covariance operators. *J. of multiv. An.* 47, 48-58.

[13] GRENANDER, U. (1981). *Abstract inference*. Wiley.

[14] HOFFMAN JORGENSEN, J. (1973). Sums of independent Banach space valued random variables. *Aarhus Univ. preprint series* 1972/73, 15.

[15] KHASMINSKII, R. and SKOROKHOD, A. (1996). A paremeter estimation problem for indirect observations. *Prob. Inf. Trans.* (to appear).

[16] KUELBS, J. (1975). The law of the iterated logarithm and related strong convergence theorems for Banach space valued random variables. Ecole d'été de Probabilités de St Flour 1975. *Lecture notes in Mathematics*, 539, 225-314. Springer-Verlag.

[17] KUKUSH, A.G. (1990). On the probability of large deviation of an estimator of a nonlinear regression parameter in Hilbert space. *Th. Pr. Math. Stat.* 40, 51-58.

[18] LECOUTRE, J.P. (1990) Uniform consistency of a class of regression function estimators for Banach space valued random variables. *Stat. Prob. Letters*, 10, 145-149.

[19] LEDOUX, M. and TALAGRAND, M. (1991). *Probability in Banach space.* Springer-Verlag.

[20] LIPSTER, R.S. and SHIRYAEV, A.N. (1977). *Statistics of random processes I.* Springer-Verlag.

[21] MAUREY, B. (1973). Espaces de cotypes *p. Séminaire Maurey-Schwartz* 1972-73. Ecole Polytechnique.

[22] MOURID, T. (1995). PhD thesis (Univ. Paris VI).

[23] PUMO, B. (1991). PhD thesis (Univ. Paris VI).

[24] STOICA, G. (1992). Functional LIL's for the Ornstein-Uhlenbeck process. *C.R. Acad. Sci. Paris* 315, I, 1089-1093.

[25] UTEV, S. (1995). Personal communication.

[26] YURINSKI, V.V. (1976). Exponential inequalities for sums of random vectors. *Journal of mult. Analysis* 6, 473-499.

[27] ZHURBENKO, I.G. and ZUPAROV, T.M. (1986). The central limit theorem for stationary processes with mixing in a Hilbert space. *Soviet Math. Dokl.* 33, 1, 66-69.

Université Pierre et Marie Curie
74, rue Dunois
75013 Paris
France