RECURSIVE ESTIMATION OF DISTRIBUTIONAL FIX-POINTS

PAUL EMBRECHTS, * *ETH Zürich*

HARRO WALK, ** *Universität Stuttgart*

Abstract

In various stochastic models the random equation $S \stackrel{d}{=} \Psi \circ S$ of implicit renewal theory appears where the real random variable S and the stochastic process Ψ with index space and state space $\mathbb{R}$ are independent. By use of stochastic approximation the distribution function of S is recursively estimated on the basis of independent or ergodic copies of Ψ. Under integrability assumptions a.s. L_1-convergence is proved. The choice of gains in the recursion is discussed. Applications are given to insurance mathematics (perpetuities) and queueing theory (stationary waiting and queueing times).

RANDOM EQUATION; DISTRIBUTIONAL FIX-POINT; RECURSIVE ESTIMATION; STOCHASTIC APPROXIMATION; INSURANCE; STOCHASTIC DISCOUNTING; PERPETUITY; QUEUEING THEORY; G/G/1; STATIONARY WAITING TIME; STATIONARY QUEUEING TIME

AMS 1991 SUBJECT CLASSIFICATION: PRIMARY 62L20

SECONDARY 60H25; 62P05; 60K25

*Postal address: Department of Mathematics, ETH, CH-8092 Zürich, Switzerland

**Postal address: Mathematisches Institut A, Universität Stuttgart, Pfaffenwaldring 57, D-70569 Stuttgart, Germany

1. Introduction

Within so-called "implicit renewal theory", investigated by Goldie (1991), the following random equation is crucial:

$$S \stackrel{d}{=} \Psi \circ S\,, \tag{1}$$

where the real-valued random variable S is defined on a probability space $(\Omega, \mathcal{A}, P)$ and $\Psi : \mathbb{R} \times \Omega \to \mathbb{R}$ is $\mathcal{B}\otimes\mathcal{A}$-$\mathcal{B}$-measurable such that S and Ψ are independent. Here $\Psi = \{\Psi(t, \cdot);\ \ t \in \mathbb{R}\}$ can be considered as a stochastic process with state space $\mathbb{R}$ or as a random variable with values in $\mathcal{M}(\mathbb{R}, \mathbb{R})$, the space of $\mathcal{B}$-$\mathcal{B}$-measurable functions from $\mathbb{R}$ to $\mathbb{R}$; the right-hand side of (1) is given by

$$(\Psi \circ S)(\omega) = \Psi(S(\omega), \omega),\ \ \omega \in \Omega\,.$$

A distribution of S solving (1) is a fix-point of Ψ. Under suitable conditions a unique distributional fix-point exists (see the literature in Goldie (1991), especially Letac (1986)). Goldie (1991) investigated the asymptotic behaviour of the corresponding distribution function and gave numerous examples, together with other references, especially Kesten (1973). Further examples and references can be found in Aebi, Embrechts and Mikosch (1994) and in Embrechts and Goldie (1994). For an application of equations of the type (1) to stochastic volatility models in mathematical finance, see Embrechts, Klueppelberg and Mikosch (1997, Section 8.4).

In the present paper a stochastic approximation method for recursive estimation of the distribution function of S is proposed. It is assumed that an independent or a stationary ergodic sequence of copies $\Psi_1, \Psi_2, \ldots$ of Ψ can be observed or at least obtained by the observation of another sequence. Under regularity assumptions, mainly an integrability assumption of contraction type on Ψ, one obtains almost sure L_1-convergence of a recursively defined estimation sequence to the distribution function of S (Theorem 2). The proof is based on a general result on linear recursions in a Banach space for the ergodic case and for more general recursions in the independence case (Theorem 1 (Walk and Zsidó (1989)) together with Corollary 1). Applications of Theorem 2 concern so-called perpetuities. These are a.s. limits of stochastically discounted sums, encountered in insurance mathematics and finance (see Aebi, Embrechts and Mikosch (1994), Embrechts and Goldie (1994), Goldie and Grübel (1996)). Another application concerns the stationary waiting and queueing times for a G/G/1 queue (as to the context with (1) see section 5 in Goldie (1991)). The application in queueing theory allows to estimate the stationary waiting or queueing time distribution sequentially on the basis of observed independent interarrival times and service times under sharpened integrability assumptions. Also in the case that the distribution of Ψ is known or, on the basis of a finite sample observation, approximately known, one can estimate the distribution function of S by the rather simple recursion using an i.i.d. simulation sequence with the same or approximately the same distribution as Ψ. While this bootstrap method is based on (1), the bootstrap method of Aebi, Embrechts and Mikosch (1994) in the context of perpetuities is based on their definition by sums or series. The recursive procedure is of stochastic approximation type. When applying these results, one has a large freedom in the choice of the gain sequence (α_n) and the starting estimate. The choice $\alpha_n = c/n$ with c not too small yields an optimal order of L_2-convergence under slightly sharpened regularity assumptions. In order to obtain also an optimal limit covariance structure the averaging concept of Polyak (1990) and Ruppert (1991) with slower decreasing (α_n) and subsequent averaging of the estimates is used (Corollary 2).

2. Results on convergence

We first formulate a deterministic result on stochastic approximation procedures for linear problems in a Banach space (see Walk and Zsidó (1989), Theorem 1 and Corollary 1, and the literature cited there). It will be used later. In the procedure a so-called gain sequence (α_n) with $\alpha_n \in [0,1)$, $\alpha_n \to 0\ (n \to \infty)$, $\sum \alpha_n = \infty$ appears which can be freely chosen, e.g. $\alpha_n = 1/(n+1)\ (n \in \mathbb{N})$. Let

$$\beta_n := [(1-\alpha_1)\ldots(1-\alpha_n)]^{-1}, \ \ \gamma_n := \alpha_n \beta_n \ \ (n \in \mathbb{N}).$$

Thus $\beta_n = 1 + \gamma_1 + \ldots + \gamma_n \uparrow \infty$.

Theorem 1. *Let $\mathbb{B}$ be a real Banach space, $\mathcal{L}(\mathbb{B})$ the Banach space of bounded linear operators on $\mathbb{B}$ into $\mathbb{B}$ and $A \in \mathcal{L}(\mathbb{B})$ such that*

$$\text{spec } A \subset \{\lambda \in \mathbb{C};\ \ \text{re } \lambda > 0\}. \tag{2}$$

Let further $A_n \in \mathcal{L}(\mathbb{B})$ and $b, b_n \in \mathbb{B}\ \ (n \in \mathbb{N})$.

a) *Assume*

$$\overline{\lim}\ \frac{1}{n}\sum_{k=1}^{n} \|A_k\| < \infty, \tag{3}$$

$$\|\frac{1}{n}\sum_{k=1}^{n} A_k - A\| \longrightarrow 0, \tag{4}$$

$$\|\frac{1}{n}\sum_{k=1}^{n} b_k - b\| \longrightarrow 0, \tag{5}$$

and let $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{B}$ be defined by

$$x_{n+1} := x_n - \frac{\rho_n}{n}(A_n x_n - b_n) \tag{6}$$

with arbitrary $x_1 \in \mathbb{B}$ where $0 \le \rho_n/n \le 1$, $\rho_n \longrightarrow \rho \in (0,\infty)$, $\rho_n - \rho_{n+1} = O(1/n)$. Then

$$\|x_n - A^{-1}b\| \longrightarrow 0. \tag{7}$$

b) *Let* $\alpha_n \in [0,1)$ $(n \in \mathbb{N})$ *with* $\alpha_n \to 0$, $\sum \alpha_n = \infty$. *Assume*

$$\overline{\lim} \frac{1}{\beta_n} \sum_{k=1}^{n} \gamma_k \|A_k\| < \infty, \tag{3'}$$

$$\|\frac{1}{\beta_n} \sum_{k=1}^{n} \gamma_k A_k - A\| \longrightarrow 0, \tag{4'}$$

$$\|\frac{1}{\beta_n} \sum_{k=1}^{n} \gamma_k b_k - b\| \longrightarrow 0 \tag{5'}$$

with β_n, γ_n *as above, and let* (x_n) *be defined by*

$$x_{n+1} := x_n - \alpha_n(A_n x_n - b_n) \tag{6'}$$

with arbitrary $x_1 \in \mathbb{B}$. *Then* (7) *holds.*

The spectral condition, which guarantees existence of a unique solution of the equation $Ax - b = 0$ in $\mathbb{B}$, is fulfilled if $\|I - A\| < 1$ (I identity operator on $\mathbb{B}$). Because, by an index transformation, the condition $\rho_n/n \leq 1$ can be reduced to $\rho_n/n < 1$, the recursion (6′) is more general than (6).

We state a supplementary result.

Corollary 1. *Assume that* $\mathbb{B}$ *is a real separable Banach space and that* $(A_n)_{n\in\mathbb{N}}$ *is a sequence of independent identically distributed (i.i.d.)* $\mathcal{L}(\mathbb{B})$*-valued r.v.'s and* $(b_n)_{n\in\mathbb{N}}$ *is a sequence of i.i.d.* $\mathbb{B}$*-valued r.v.'s on a probability space* $(\Omega, \mathcal{A}, P)$ *with* $E\|A_n\|^2 < \infty$, $E\|b_n\|^2 < \infty$. *Let* $A := EA_n$ $(\in \mathcal{L}(\mathbb{B}))$, $b := Eb_n$ $(\in \mathbb{B})$ *and assume* (2). *Let* $(X_n)_{n\in\mathbb{N}}$ *be a sequence of* $\mathbb{B}$*-valued r.v.'s on* $(\Omega, \mathcal{A}, P)$ *with*

$$X_{n+1} = X_n - \alpha_n(A_n X_n - b_n), \ \ n \in \mathbb{N},$$

where $\alpha_n \in [0,1)$, $\sum \alpha_n^2 < \infty$, $\sum \alpha_n = \infty$. *Then*

$$\|X_n - A^{-1}b\| \longrightarrow 0 \quad \text{a.s.}$$

Proof. According to Theorem 1b it suffices to verify that for P-almost all $\omega \in \Omega$, relations (3′), (4′), (5′) hold. This is possible by the following Lemma applied to the case of Banach spaces $\mathbb{R}$, $\mathcal{L}(\mathbb{B})$, $\mathbb{B}$. □

Lemma 1. *Let $\mathbb{B}$ be a real separable Banach space, (α_n) a sequence in $[0,1)$ with $\sum \alpha_n^2 < \infty$, $\sum \alpha_n = \infty$, (β_n) and (γ_n) defined as above and (V_n) a sequence of i.i.d. $\mathbb{B}$-valued r.v.'s with $E\|V_n\|^2 < \infty$, $V := EV_n$. Then*

$$\|\frac{1}{\beta_n}\sum_{k=1}^{n} \gamma_k V_k - V\| \longrightarrow 0 \qquad \text{a.s.}$$

Proof. Without loss of generality let $V = 0$. Choose an arbitrary $\varepsilon > 0$. Below we use some arguments from Beck (1963) and Györfi and Masry (1990, Appendix A). For the identically distributed square integrable r.v.'s V_n one has a measurable function F on $\mathbb{B}$ with values in a finite subset of $\mathbb{B}$ such that, with $Y_n := F(V_n)$, $Z_n := V_n - (Y_n - EY_n)$, the relation

$$E\|Z_n\|^2 < \varepsilon, \;\; n \in \mathbb{N},$$

holds. Further one notes that

$$\|\frac{1}{\beta_n}\sum_{k=1}^{n} \gamma_k V_k\| \le |\; \frac{1}{\beta_n}\sum_{k=1}^{n} \gamma_k(\|Z_k\| - E\|Z_k\|)\;| + \sqrt{\varepsilon} + \|\frac{1}{\beta_n}\sum_{k=1}^{n} \gamma_k(Y_k - EY_k)\|,$$

$$E\|Y_n - EY_n\|^2 \le 2E\|Z_n\|^2 + 2E\|V_n\|^2 \le 2\varepsilon + 2E\|V_1\|^2.$$

Because of Kolmogorov's a.s. convergence theorem for series of independent square integrable real r.v.'s, which also applies to the case of random variables with values in a finite-dimensional Banach space, and since $\sum \alpha_n^2 E\|Z_n\|^2 < \infty$ and $\sum \alpha_n^2 E\|Y_n\|^2 < \infty$, we have a.s. convergence of $\sum \alpha_n(\|Z_n\| - E\|Z_n\|)$ and of $\sum \alpha_n(Y_n - EY_n)$. The Kronecker lemma now yields

$$\overline{\lim} \| \frac{1}{\beta_n} \sum_{k=1}^{n} \gamma_k V_k \| \leq \sqrt{\varepsilon} \quad \text{a.s.}$$

and thus the assertion. □

Theorem 1 and Corollary 1 allow to prove an almost sure convergence result (Theorem 2) for recursive estimation in the following situation of implicit renewal theory.

Let $(\Omega, \mathcal{A}, P)$ be a probability space and let $\Psi : \mathbb{R} \times \Omega \to \mathbb{R}$ be $\mathcal{B} \otimes \mathcal{A}$-$\mathcal{B}$-measurable. Assume that the paths $\Psi(\cdot, \omega)$, $\omega \in \Omega$, are nondecreasing and Lipschitz continuous. Further assume that $E \mid \Psi(0, \cdot) \mid < \infty$ and $EL < 1$, where $L(\omega)$ is the minimal Lipschitz constant of $\Psi(\cdot, \omega)$. In the proof of Theorem 2, the application of Theorem 1a and Corollary 1 yields that boundedness of the operators A_n and the contraction condition $\|I - A\| < 1$ are implied by Lipschitz continuity of $\Psi(\cdot, \omega)$ and $EL < 1$, respectively. The argument following (15) in the proof of Theorem 2 yields existence of a unique distribution function F_S on $\mathbb{R}$ satisfying (1), where the real random variable S with distribution function F_S and the $\mathcal{M}(\mathbb{R}, \mathbb{R})$-valued random variable Ψ on the suitably enlarged probability space $(\Omega, \mathcal{A}, P)$ are independent. In this case $E \mid S \mid < \infty$. The contraction condition $EL < 1$ is also important for almost sure convergence of the stochastic approximation algorithm in Theorem 2. Goldie (1991, p. 127 and Corollary 2.4) considers random functions Ψ with $\Psi(t)$, for $\mid t \mid$ large, approximately equal to Mt. Here the random variable M satisfies $E\log|M| < 0$ and the "Cramér-condition" $E \mid M \mid^{\kappa} = 1$, for some $\kappa > 0$. The latter condition is important for the tail behavior of the distribution function F_S of S in (1) (Goldie (1991, Corollary 2.4)). In both applications below, see (13) and (14), the first condition of Goldie is sharpened to the contraction condition $EM < 1$ for nonnegative M, i.e. here $EL < 1$, where the second condition is not needed.

Let $F_0(t) = 1$ for $t \geq 0$, $F_0(t) = 0$ for $t < 0$, and let Γ_1 be the set of distribution functions F on $\mathbb{R}$ satisfying

$$\int \mid t \mid dF(t) \; (= d_1(F_0, F)) < \infty ,$$

with metric d_1 given by

$$d_1(F,G) = \int_{\mathbb{R}} \mid F(t) - G(t) \mid dt, \;\; F, G \in \Gamma_1 .$$

Note that d_1 is also the L_1-minimal metric in Γ_1 defined by

$$d_1(F,G) := \inf\{E \mid X - Y \mid ; \; F_X = F, F_Y = G\}, \;\; F, G \in \Gamma_1 ,$$

where F_X is the distribution function of the real random variable X (see Aebi, Embrechts and Mikosch (1994)).

Theorem 2. *Let on a probability space $(\Omega, \mathcal{A}, P)$ a stationary ergodic sequence $(\Psi_n)_{n\in\mathbb{N}}$ of copies of Ψ be given. Choose an arbitrary $R_1 \in \Gamma_1$ (e.g. $R_1 = F_0$) and $\rho_n \in \mathbb{R}_+$ with $\rho_n/n \leq 1$ $(n \in \mathbb{N})$, $\rho_n \longrightarrow \rho \in (0, \infty)$, $\rho_n - \rho_{n+1} = O(1/n)$ (e.g. $\rho_n = 1$). Let the sequence $(R_n)_{n\in\mathbb{N}}$ of Γ_1-valued random variables recursively be defined by*

$$R_{n+1}(\cdot, \omega) = (1 - \frac{\rho_n}{n}) R_n(\cdot, \omega) + \frac{\rho_n}{n} R_n(\Psi_n^{-1}(\cdot, \omega), \omega), \;\; \omega \in \Omega, \tag{10}$$

where

$$\Psi_n^{-1}(t, \cdot) := \sup\{s \in \mathbb{R}; \Psi_n(s, \cdot) \leq t\}, \;\; t \in \mathbb{R},$$

$$\sup \;\; \emptyset := -\infty, \;\; R_n(\infty, \cdot) := 1, \;\; R_n(-\infty, \cdot) := 0.$$

Then

$$d_1(R_n, F_S) \longrightarrow 0, \quad \text{i.e.} \int \mid R_n(t) - F_S(t) \mid dt \longrightarrow 0 \text{ a.s.} \tag{11}$$

If additionally the sequence (Ψ_n) *of copies of* Ψ *is independent,* $E \mid \Psi(0) \mid^2 < \infty$ *and* $EL^2 < \infty$, *then (11) also holds for* (R_n) *defined by the general recursion*

$$R_{n+1}(\cdot,\omega) = (1-\alpha_n)R_n(\cdot,\omega) + \alpha_n R_n(\Psi_n^{-1}(\cdot,\omega),\omega), \;\; \omega \in \Omega, \tag{12}$$

with a gain sequence (α_n) *chosen arbitrarily in* $[0,1)$ *with* $\sum \alpha_n^2 < \infty$, $\sum \alpha_n = \infty$.

Before we give the proof of this result (Section 3) we discuss two applications, one in insurance mathematics and one in queueing theory.

First application. Let the real random variables M, U be independent, where $E \log |M| < 0$, $E \log^+ |U| < \infty$. Let (M_1, U_1), $(M_2, U_2), \ldots$ be an independent sequence of copies of (M, U). Then the real-valued limit

$$S = \sum_{i=1}^{\infty} M_1 \ldots M_i U_i = \lim_{l \to \infty} \sum_{i=1}^{l} M_1 \ldots M_i U_i$$

of stochastically discounted sums, where the M_i's and U_i's are considered as observable discount factors and payments, respectively, exists a.s. (see Vervaat (1979)) and is known as a "perpetuity" in life insurance and finance. See also Embrechts, Klueppelberg and Mikosch (1997, Section 8.4). Set

$$\Psi(t) = M(U+t), \;\; t \in \mathbb{R}\,. \tag{13}$$

Then (1) is fulfilled for independent (M, U), S with uniqueness of the distribution of S (see Goldie (1991), Section 4, Aebi, Embrechts and Mikosch (1994), Embrechts and Goldie (1994), and the literature cited there). Set

$$\Psi_n(t) = M_n(U_n + t), \;\; t \in \mathbb{R}\,,$$

with $n \in \mathbb{N}$ and choose $\rho_n/n \in [0,1]$ with $\rho_n \longrightarrow \rho \in (0,\infty)$, $\rho_n - \rho_{n+1} = O(1/n)$. Under the sharpened assumptions $M \geq 0$, $EM < 1$, $E \mid U \mid < \infty$ the conditions of Theorem 2 (first part) with recursion

$$R_{n+1}(t) = (1 - \frac{\rho_n}{n}) R_n(t) + \frac{\rho_n}{n} R_n(\frac{t}{M_n} - U_n), \;\; t \in \mathbb{R}, \;\; n \in \mathbb{N},$$

where

$$R_n(\frac{t}{M_n} - U_n) = F_0(t) \text{ for } M_n = 0,$$

are fulfilled. Thus

$$\int \mid R_n(t) - P[S \leq t] \mid dt \longrightarrow 0 \text{ a.s.}$$

This also holds for the general recursion with ρ_n/n replaced by $\alpha_n \in [0,1)$ satisfying $\sum \alpha_n^2 < \infty$, $\sum \alpha_n = \infty$, if additionally $EM^2 < \infty$, $EU^2 < \infty$ hold.

Second application. Let the integrable nonnegative-real random variables A, B be independent with $E(B - A) < 0$. Let (B_0, A_1) $(B_1, A_2), \ldots$ be independent copies of (B, A), where the A_i's and B_i's are observable interarrival times and service times, respectively, in a G/G/1 queueing system. As is well-known, the stationary waiting time W satisfies

$$W \stackrel{d}{=} \max_{m \geq 1} \sum_{i=1}^{m} (B_{i-1} - A_i)$$

(the right-hand side exists a.s.) and

$$W \stackrel{d}{=} (B - A + W)_+$$

for independent $(A, B), W$. Following Goldie (1991) one sets

$$S := e^W, \;\; M := e^{B-A}, \;\; U := 1$$

and obtains that for independent (A, B), S equation (1) is fulfilled with

$$\Psi(t) = \max\{U, Mt\}, \;\; t \in \mathbb{R}\,,$$

and uniqueness of the distribution of S. Now set

$$\Psi_n(t) = \max\{U_n, M_n t\}, \ \ t \in \mathbb{R}\,, \tag{14}$$

with

$$U_n = 1, \ \ M_n = e^{B_{n-1} - A_n} \ \ (n \in \mathbb{N})\,,$$

and choose $\rho_n/n \in [0,1]$ with $\rho_n \to \rho \in (0,\infty)$, $\rho_n - \rho_{n+1} = O(1/n)$. Under the sharpened assumptions

$$Ee^{B-A} < 1, \quad \text{i.e. } EM < 1,$$

the conditions of Theorem 2 with recursion

$$R_{n+1}(t) = (1 - \frac{\rho_n}{n})R_n(t) + \frac{\rho_n}{n} R_n(\frac{t}{e^{B_{n-1}-A_n}}), \ \ t \geq 1, \ \ n \in \mathbb{N},$$

are fulfilled. Thus, setting $t = e^u$, one obtains

$$\int_0^\infty \mid R_n(e^u) - P[W \leq u] \mid e^u du \longrightarrow 0 \quad \text{ a.s. .}$$

Analogously the stationary queueing time Q (waiting time plus service time) can be treated. One has

$$Q \stackrel{d}{=} B + (Q - A)_+\,.$$

In the above consideration, again following Goldie (1991), one replaces W by Q and sets

$$U := e^B\,.$$

Now, under the sharpened assumptions

$$Ee^B < \infty, \ \ Ee^{B-A} < 1$$

one uses the recursion

$$R_{n+1}(t) = \begin{cases} (1 - \frac{\rho_n}{n})R_n(t) + \frac{\rho_n}{n} R_n(\frac{t}{e^{B_{n-1}-A_n}}) & , \ t \geq e^{B_n}, \\ (1 - \frac{\rho_n}{n})R_n(t) & , \ 1 \leq t < e^{B_n}, \end{cases}$$

and obtains

$$\int_0^\infty \mid R_n(e^u) - P[Q \le u] \mid e^u du \longrightarrow 0 \quad \text{a.s. .}$$

This result also follows for the general recursion with ρ_n/n replaced by α_n as before, if additionally $Ee^{2(B-A)} < \infty$ holds.

3. Proof of Theorem 2.

Let $\mathbb{B} := L_1(\mathbb{R})$ and $K := \{F_0 - r;\ r \in \Gamma_1\}$. Then $K \subset \mathbb{B}$, because $\int_{\mathbb{R}} \mid F_0(s) - r(s) \mid ds = \int_{\mathbb{R}} \mid s \mid r(ds) < \infty$ for $r \in \Gamma_1$. If $v \in \mathbb{B}$ and if $\psi = \Psi(\cdot, \omega)$, $l = L(\omega)$ are realizations of Ψ and L, respectively, then $v \circ \psi^{-1} \in \mathbb{B}$, because

$$\int \mid v(\psi^{-1}(s)) \mid ds = \int \mid v(u) \mid \frac{d\psi(u)}{du} du \le l\|v\| < \infty.$$

If $v \in K$, then trivially $F_0 - v \in \Gamma_1$. Also $(F_0 - v) \circ \psi^{-1} \in \Gamma_1$, because with r.v.'s $\widetilde{U}, \widetilde{V}$ having distribution functions $F_0 - v$ $(\in \Gamma_1)$ and $(F_0 - v) \circ \psi^{-1}$, respectively, one has that

$$\begin{aligned} & \int \mid F_0(s) - (F_0 - v)(\psi^{-1}(s)) \mid ds = E \mid \widetilde{V} \mid \\ = \quad & E \mid \psi \circ \widetilde{U} \mid \le \mid \psi(0) \mid + E \mid \psi \circ \widetilde{U} - \psi(0) \mid \\ \le \quad & \mid \psi(0) \mid + lE \mid \widetilde{U} \mid = \mid \psi(0) \mid + l\|v\| < \infty. \end{aligned}$$

For $v \in \mathbb{B}$ let $v \circ \Psi^{-1}$ denote the $\mathbb{B}$-valued r.v. $\omega \to v(\Psi^{-1}(\cdot, \omega))$, for $v \in K$ let $(F_0 - v) \circ \Psi^{-1}$ denote the Γ_1-valued r.v. $\omega \to (F_0 - v)(\Psi^{-1}(\cdot, \omega))$. The $\mathbb{B}$-valued r.v.'s $v \circ \Psi^{-1}$, $v \in \mathbb{B}$, and the K-valued r.v.'s $F_0 - (F_0 - v) \circ \Psi^{-1}$, $v \in K$, are integrable. This follows because

$$E\|v \circ \Psi^{-1}\| \le EL \cdot \|v\| < \infty \text{ for } v \in \mathbb{B},$$

$$E\|F_0 - (F_0 - v) \circ \Psi^{-1}\| \le E \mid \Psi(0, \cdot) \mid + EL \cdot \|v\| < \infty \text{ for } v \in K.$$

One defines the operator $A \in \mathcal{L}(\mathbb{B})$ by

$$Av := v - Ev \circ \Psi^{-1},\ v \in \mathbb{B},$$

further

$$b := E(F_0 - F_0 \circ \Psi^{-1}).$$

One then has

$$\|I - A\| \leq EL < 1,$$

thus A satisfies condition (2) in Theorem 1. The unique solution $A^{-1}b$ of the equation $Ax - b = 0$ in $\mathbb{B}$ lies in the closed convex subset K of $\mathbb{B}$. In fact, for $v \in K$ one has

$$v - (Av - b) = E[F_0 - (F_0 - v) \circ \Psi^{-1}] \in K;$$

therefore the sequence (z_n) defined by $z_{n+1} := (I - A)z_n + b$ with $z_1 \in K$ lies in K and converges to $A^{-1}b$ according to Banach's fixed point theorem. The equation $Ax - b = 0$ in K with unique solution $A^{-1}b$ is equivalent to the equation

$$w(t) = Ew(\Psi^{-1}(t, \cdot)), \;\; t \in \mathbb{R} \tag{15}$$

for $w \in \Gamma_1$, whose unique solution is $F_0 - A^{-1}b$. Let S be a real-valued r.v. with distribution function F_S such that S and Ψ are independent. Because of

$$P[\Psi \circ S \leq t] = P[S \leq \Psi^{-1}(t, \cdot)] = EF_S(\Psi^{-1}(t, \cdot)), \;\; t \in \mathbb{R},$$

equation (1) is equivalent to (15), and its unique solution has the distribution function $F_S = F_0 - A^{-1}b$.

First recursion (10) is considered. One sets

$$X_n := F_0 - R_n, \;\; n \in \mathbb{N},$$

and from (10) one obtains

$$X_{n+1} = X_n - \frac{\rho_n}{n}(A_n X_n - b_n), \;\; n \in \mathbb{N},$$

where A_n is a random bounded linear operator on $\mathbb{B}$ into $\mathbb{B}$ with

$$A_n v = v - v \circ \Psi_n^{-1}, \ \ v \in \mathbb{B},$$

and where

$$b_n = F_0 - F_0 \circ \Psi_n^{-1}$$

is a $\mathbb{B}$-valued r.v. This recursion corresponds to recursion (6) in Theorem 1. One verifies conditions (3), (4), (5) of Theorem 1a for P-almost all realizations. With $L_n(\omega)$ as minimal Lipschitz constant of $\Psi_n(\cdot, \omega)$, $\omega \in \Omega$, one obtains as above that

$$\|v - A_n v\| \leq L_n \|v\|, \ \ v \in \mathbb{B}.$$

Thus

$$\|A_n\| \leq 1 + L_n, \ \ E\|A_n\| \leq 1 + EL < \infty,$$

further

$$\|b_n\| = \mid \Psi_n(0) \mid, \ \ E\|b_n\| = E \mid \Psi(0) \mid < \infty.$$

The a.s. ergodic theorem for $\mathbb{R}$ yields (3) because of the stationarity assumption. Because of the stationarity and ergodicity assumptions and $EA_n v = Av$ for all $v \in \mathbb{B}$, $EA_n = A$, $Eb_n = b$, an a.s. ergodic theorem for separable Banach spaces (Tempel'man (1972), Landers and Rogge (1978), Györfi and Masry (1990); compare Krengel (1985), § 4.2), here for $\mathcal{L}(\mathbb{B})$ and for $\mathbb{B}$, yields (4) and (5). Theorem 1a now yields

$$\|X_n - A^{-1} b\| \longrightarrow 0 \text{ a.s.},$$

and thus (11).

Finally recursion (12) is considered. With

$$X_n := F_0 - R_n, \ \ n \in \mathbb{N},$$

from (12) one obtains

$$X_{n+1} = X_n - \alpha_n(A_n X_n - b_n), \ \ n \in \mathbb{N},$$

which corresponds to (6′), with A_n, b_n as before. By the additional assumptions of independence of (Ψ_n) and of square integrability of $\Psi(0)$ and L one obtains (11) via Corollary 1. □

4. Choice of the gain sequence

Below, the assumptions in the second part of Theorem 2 are used, namely independence of the sequence (Ψ_n) of copies of Ψ together with $E|\Psi(0)|^2 < \infty$ and $EL^2 < \infty$.

The simplest choice of the gains ρ_n/n in Theorem 2 is c/n with $0 < c \leq 1$, especially $1/n$. Let in Γ_1 a further metric d be given by

$$d(F, G) := \Big(\int_{\mathbb{R}} (F(t) - G(t))^2 dt \Big)^{1/2}.$$

Apparently convergence with respect to the metric d_1 yields convergence with respect to the metric d. Remark b) below yields convergence of $nEd(R_n, F_S)^2$ for c so large that $EL < 1 - 1/(2c)$. The limit would be minimized if one would replace the factor c by a certain linear transformation depending on the unknown stochastic behaviour of Ψ. In fact, in one-dimensional and finite-dimensional problems of stochastic approximation this concept is used via consistent estimation of (a matrix of) derivatives (see Venter (1967), Fabian (1968), Polyak and Tsypkin (1979), Wei (1987), Remarks 5.10 and 5.16 in part I of Ljung, Pflug and Walk (1992) with further references). In finite-dimensional linear estimation, the choice of the coefficients optimal in an L_2-sense can be done recursively by use of matrix transformations (for linear regression by a recursive least squares method, see e.g. Ljung and Söderström (1983); compare also Polyak (1990), equation (11); for non-linear time series models see Thavaneswaran and Abraham (1988)). The averaging concept in stochastic approximation due to Polyak (1990) and Ruppert (1991) avoids matrix transformations and leads to the same optimal limit. Corresponding to this, in our infinite-dimensional context for the recursion sequence (R_n) we take a slowly decreasing gain sequence (α_n) with

$$\alpha_n \to 0, \ \sum \alpha_n = \infty, \ \alpha_n/\alpha_{n+1} = 1 + o(\alpha_n), \tag{16}$$

e.g. $\alpha_n = (n+1)^{-\alpha}$ with $0 < \alpha < 1$ and then use

$$\overline{R}_n := \frac{1}{n} \sum_{k=1}^{n} R_k$$

as an estimate of F_S. Considering $F_0 - R_n =: X_n$, $F_0 - F_S =: \vartheta$, $F_S - R_n = X_n - \vartheta$, $F_S - \overline{R}_n = \overline{X}_n - \vartheta$ as random elements in the Hilbert space $L_2(\mathbb{R})$ one obtains $d(\overline{R}_n, F_S) \to 0$ a.s. by Theorem 2 and convergence (in the operator norm, even in a trace norm) of the operators $nE\big((\overline{X}_n - \vartheta) \otimes (\overline{X}_n - \vartheta)\big)$ to a covariance operator (S-operator) with minimal trace, which is the limit of $Ed(R_n, F_S)^2$. This follows directly from Corollary 2 below, which is formulated in a general Hilbert space setting.

Let $\mathbb{H}$ denote a separable real Hilbert space (with scalar product $<,>$ and norm $\| \, \|$) and $\mathcal{L}(\mathbb{H})$ the Banach space of bounded linear operators on $\mathbb{H}$ into $\mathbb{H}$. $I : \mathbb{H} \to \mathbb{H}$ denotes the identity operator. As to the following notations, see Reed and Simon (1980). For $A \in \mathcal{L}(\mathbb{H})$ let tr A denote its trace, spec A its spectrum, A^* the adjoint operator, further $\sqrt{A}$ its square root if A is symmetric positive-semidefinite. Under the norm $\|A\|_1 := \operatorname{tr} \sqrt{AA^*}$ $(\geq \|A\|)$ the space $\mathcal{L}_1(\mathbb{H}) \subset \mathcal{L}(\mathbb{H})$ of nuclear operators on $\mathbb{H}$ is a separable Banach space. It contains the space of all symmetric positive-semidefinite nuclear operators (S-operators).

$x \otimes y := < x, \cdot > y$ denotes the tensor product of x and y ($x, y \in \mathbb{H}$), which is considered as an element of $\mathcal{L}(\mathbb{H})$. The covariance operator of an $\mathbb{H}$-valued random variable X with $E\|X\|^2 < \infty$ and $EX = 0$ is an S-operator defined by $S := E(X \otimes X)$ with $\|S\|_1 = tr\, S$ $= E\|X\|^2$.

The following corollary, under an independence assumption, extends Theorem 3 of Polyak (on parameter estimation) to a more general linear recursion and to the Hilbert space case.

Corollary 2. *Let $((A_n, b_n))_{n\in\mathbb{N}}$ be an independent sequence of identically distributed $\mathcal{L}(\mathbb{H}) \times \mathbb{H}$-valued random variables with $E\|A_n\|^2 < \infty$, $E\|b_n\|^2 < \infty$. For $A := EA_n$ assume* (2). *Let $\vartheta = A^{-1}b$ denote the unique solution of the equation $Ax - b = 0$ in $\mathbb{H}$, where $b := Eb_n$. Choose α_n in* $(0,1)$ *such that* (16) *is fulfilled and let X_1 be an $\mathbb{H}$-valued random variable with $E\|X_1\|^2 < \infty$. Let the sequences (X_n), $(\overline{X}_n)$ of (square integrable) $\mathbb{H}$-valued random variables be defined by*

$$X_{n+1} := X_n - \alpha_n(A_n X_n - b_n), \tag{17}$$

$$\overline{X}_n := \frac{1}{n}\sum_{k=1}^{n} X_k \quad (n \in \mathbb{N}).$$

Then

$$\|nE((\overline{X}_n - \vartheta) \otimes (\overline{X}_n - \vartheta)) - K\|_1 \to 0$$

with S-operator $K = A^{-1}S(A^{-1})^*$ where $S := E((b_n - A_n\vartheta) \otimes (b_n - A_n\vartheta))$, thus

$$nE\|X_n - \vartheta\|^2 \to \ \mathrm{tr}\ K.$$

Remarks concerning Corollary 2.a) By Corollary 1, almost sure convergence $\|X_n - \vartheta\| \to 0$, $\|\overline{X}_n - \vartheta\| \to 0$ hold without the assumption $\alpha_n/\alpha_{n+1} = 1 + o(\alpha_n)$.

b) If one replaces in (17) the factor α_n by the operator $\frac{1}{n}\Gamma$ (special case $\frac{c}{n}I$ with $c > 0$) and assumes

$$\text{spec }(\Gamma A - \frac{I}{2}) \subset \{\lambda \in \mathbb{C};\ \text{re } \lambda > 0\},$$

then

$$\|nE((X_n - \vartheta) \otimes (X_n - \vartheta)) - \widetilde{K}\|_1 \to 0,$$

where the S-operator $\widetilde{K}$ is uniquely determined by

$$(\Gamma A - \frac{I}{2})\widetilde{K} + \widetilde{K}((\Gamma A)^* - \frac{I}{2}) = \Gamma S \Gamma^*.$$

tr $\widetilde{K}$ is minimized by $\Gamma = A^{-1}$; in this case $\widetilde{K}$ equals K in Corollary 2. The convergence result can be obtained from Lemma 2b below (compare also Walk (1988), where functional central limit theorems are derived).

The proof of Corollary 2 is based on Lemma 2 which will be proved first. Both proofs follow the lines in Polyak (1990), but also use functional analytic tools.

Lemma 2 (compare lemmas 1, 5, 3, 4 in Polyak (1990)). *Let* $A \in \mathcal{L}(\mathbb{H})$ *satisfy* (2). *Let* $0 < \alpha_n < 1$ *with* $\alpha_n \to 0$ $(n \to \infty)$, $\sum \alpha_n = \infty$ *and let* Z_n, M_n, Q_n, J_n, L_n, D_n, F_n, G_n, $H_n \in \mathcal{L}_1(\mathbb{H})$ $(n \in \mathbb{N})$.

a) *If*

$$Z_{n+1} = Z_n - \alpha_n A Z_n - \alpha_n Z_n A^* + \alpha_n M_n + \alpha_n \|Z_n\|_1 Q_n \tag{18}$$

or

$$Z_{n+1} = Z_n - \alpha_n A Z_n + \alpha_n M_n + \alpha_n \|Z_n\|_1 Q_n, \tag{19}$$

with $\|M_n\|_1 \to 0$, $\|Q_n\|_1 \to 0$, *then*

$$\|Z_n\|_1 \to 0.$$

Let further $S \in \mathcal{L}_1(\mathbb{H})$ and then $K \in \mathcal{L}_1(\mathbb{H})$ be the unique solution of

$$AK + KA^* = S. \tag{20}$$

b) *If*

$$Z_{n+1} = Z_n - \frac{1}{n}(A + \frac{I}{2})Z_n - \frac{1}{n}Z_n(A + \frac{I}{2})^* + \frac{1}{n^2}S + \frac{1}{n^2}J_n + \frac{1}{n}\|Z_n\|_1 L_n$$

with $\|J_n\|_1 \to 0$, $\|L_n\|_1 \to 0$, *then*

$$\|nZ_n - K\|_1 \to 0.$$

Let additionally $\alpha_n/\alpha_{n+1} = 1 + o(\alpha_n)$ *(thus* (α_n) *satisfies* (16)*).*

c) *If*

$$Z_{n+1} = Z_n - \alpha_n A Z_n - \alpha_n Z_n A^* + \alpha_n^2 S + \alpha_n^2 D_n + \alpha_n^2 \|Z_n\|_1 F_n$$

with $\|D_n\|_1 \to 0$, $\|F_n\|_1 \to 0$, *then*

$$\|\alpha_n^{-1} Z_n - K\|_1 \to 0.$$

d) *If*

$$Z_{n+1} = Z_n - \alpha_n A Z_n + \frac{\alpha_n}{n+1}U + \frac{\alpha_n}{n+1}G_n + \frac{\alpha_n}{n+1}\|Z_n\|_1 H_n$$

with $\|G_n\|_1 \to 0$, $\|H_n\|_1 \to 0$, *then*

$$\|(n+1)Z_n - A^{-1}K\|_1 \to 0.$$

Proof. a) Let B_1, B_2, B be bounded linear operators on $\mathcal{L}_1(\mathbb{H})$ defined by

$$B_1 Z = AZ, \;\; B_2 Z = ZA^* \text{ for } Z \in \mathcal{L}_1(\mathbb{H}), \;\; B = B_1 + B_2.$$

Relation (2), i.e. $\|e^{-uA}\| \to 0 \ (u \to \infty)$, implies

$$\|e^{-uB_1}\| = \|e^{-uB_2}\| \to 0 \ (u \to \infty)$$

and, by commutativity of B_1 and B_2,

$$\|e^{-uB}\| = \|e^{-uB_1}e^{-uB_2}\| \leq \|e^{-uB_1}\|^2 \to 0 \ (u \to \infty)\,.$$

Thus

$$\text{spec } B \subset \{\lambda \in C; \ \text{re } \lambda > 0\}. \tag{21}$$

In the following let I denote the identity operator on $\mathcal{L}_1(\mathbb{H})$ and notice that for a bounded linear operator D on $\mathcal{L}_1(\mathbb{H})$ into $\mathcal{L}_1(\mathbb{H})$ and $E \in \mathcal{L}_1(\mathbb{H})$ also $DE \in \mathcal{L}_1(\mathbb{H})$ with $\|DE\|_1 \leq \|D\|\|E\|_1$ (see Reed and Simon (1980), Ch. VI, Problem 28(a)). (18) can be written in the form

$$Z_{n+1} = (I - \alpha_n B)Z_n + \alpha_n M_n + \alpha_n \|Z_n\|_1 Q_n$$

and yields

$$Z_{n+1} = \sum_{k=1}^{n} \alpha_k B_{n,k}(M_k + \|Z_k\|_1 Q_k) + B_{n,1}Z_1$$

with

$$\begin{aligned} B_{n,k} &= (I - \alpha_n B)\ldots(I - \alpha_k B) \text{ for } k = 1, \ldots, n, \\ B_{n,n+1} &= I \end{aligned}$$

according to Walk and Zsidó (1989) (first part of Lemma 1a). Thus

$$\|Z_{n+1}\|_1 \leq \sum_{k=1}^{n} \alpha_k \|B_{n,k}\|(\|M_k\|_1 + \|Z_k\|_1\|Q_k\|_1) + \|B_{n,1}\|\|Z_1\|_1.$$

Note that $\|M_n\|_1 \to 0$, $\|Q_n\|_1 \to 0$, further

$$\|B_{n,k}\| \leq c(\beta_k/\beta_n)^{\varepsilon} \quad (k = 1, \ldots, n; \ n \in \mathbb{N})$$

for some constants $c < \infty$, $\varepsilon > 0$ by (21) (see Walk and Zsidó (1989), Lemma 3b). One hence obtains an inequality of the form

$$\|Z_{n+1}\|_1 \leq \sum_{k=1}^{n} t_{n,k}\|Z_k\|_1 + r_n$$

with $t_{n,k} \in \mathbb{R}_+$, $r_n \in \mathbb{R}_+$ satisfying $r_n \to 0$, $t_{n,k} \to 0 \ \ (n \to \infty)$ for each $k \in \mathbb{N}$ and

$$\sum_{k=1}^{n} t_{n,k} \to 0 \quad (n \to \infty).$$

Thus $\|Z_n\|_1 \to 0$ (compare Walk and Zsidó (1989), Lemma 2b). Recursion (17) can be treated in the same way.

b,c,d). Because of (21), (20) has a unique solution. One establishes recursions for $nZ_n - K$, $\alpha_n^{-1}Z_n - K$, $(n+1)Z_n - A^{-1}K$ (here noticing $1/n = o(\alpha_n)$ according to Polyak (1990)) and uses each time part a). □

Proof of Corollary 2. Without loss of generality $b = \vartheta = 0$ may be assumed. From (2), (17) and the independence assumption one obtains

$$\|EX_n\| \to 0. \tag{22}$$

Let $U_n := E(X_n \otimes X_n)$, $T_n := E(\overline{X}_n \otimes \overline{X}_n)$, $R_n := E(\overline{X}_n \otimes X_{n+1})$. By (17), (22) and the independence assumption one obtains

$$\begin{aligned} U_{n+1} &= U_n - \alpha_n A U_n - \alpha_n U_n A^* + \alpha_n^2 E(A_n U_n A_n^*) - \alpha_n^2 E((A_n(EX_n)) \otimes b_n) \\ &\quad -\alpha_n^2 E(b_n \otimes (A_n(EX_n))) + \alpha_n^2 S \\ &= U_n - \alpha_n A U_n - \alpha_n U_n A^* + \alpha_n^2 S + \alpha_n^2 T_n + \alpha_n^2 \|U_n\|_1 V_n \end{aligned}$$

with $\|T_n\|_1 \to 0$, $\|V_n\|_1 \to 0$, noticing

$$\|E((A_n(EX_n)) \otimes b_n)\|_1 \leq \|EX_n\|(E\|A_n\|^2)^{1/2}(E\|b_n\|^2)^{1/2} \to 0$$

and

$$\|A_n U_n A_n^*\|_1 \le \|A_n\|^2 \|U_n\|_1.$$

Thus, by Lemma 2c,

$$U_n = \alpha_n K + \alpha_n W_n, \tag{23}$$

with $\|W_n\|_1 \to 0$. One uses the recursion for X_{n+2}, the recursion

$$\overline{X}_{n+1} = \Big(1 - \frac{1}{n+1}\Big)\overline{X}_n + \frac{1}{n+1} X_{n+1} \tag{24}$$

and once more the independence assumption and obtains

$$R_{n+1} = \Big(1 - \frac{1}{n+1}\Big)(I - \alpha_n A) R_n + \frac{1}{n+1}(I - \alpha_n A) U_{n+1}\,.$$

Then, because of $n\alpha_n \to \infty$ and (23),

$$R_{n+1} = (I - \alpha_n A) R_n + \frac{\alpha_n}{n+1} U + \frac{\alpha_n}{n+1} G_n + \frac{\alpha_n}{n+1} \|R_n\|_1 H_n$$

with $\|G_n\|_1 \to 0$, $\|H_n\|_1 \to 0$ and thus, by Lemma 2d,

$$\|(n+1)R_n - A^{-1}K\|_1 \to 0.$$

This together with (24) and $\|U_n\|_1 \to 0$ yields

$$T_{n+1} = \Big(1 - \frac{2}{n}\Big) T_n + \frac{1}{n^2}(A^{-1}K + (A^{-1}K)^*) + \frac{1}{n^2} J_n + \frac{1}{n}\|T_n\|_1 L_n$$

with $\|J_n\|_1 \to 0$, $\|L_n\|_1 \to 0$, where $A^{-1}K + (A^{-1}K)^* = A^{-1}S(A^{-1})^*$ because of (20). The assertion then follows by Lemma 2b. $\square$

Acknowledgment. The authors thank the referee for suggestions which led to a considerable improvement of the paper.

References

AEBI, M., EMBRECHTS, P. AND MIKOSCH, T. (1994) Stochastic discounting, aggregate claims and the bootstrap. *Adv. Appl. Prob.* **26**, 183–206.

EMBRECHTS, P. AND GOLDIE, C.M. (1994) Perpetuities and random equations. In *Asymptotic Statistics: Proc. Fifth Prague Symp.*, September 1993. Ed. P. Mandl and M. Huskova. Physica-Verlag, Heidelberg, pp. 75–86.

EMBRECHTS, P., KLUEPPELBERG, C. AND MIKOSCH, T. (1997) *Modelling Extremal Events for Insurance and Finance.* Springer-Verlag, Berlin.

FABIAN, V. (1968) On asymptotic normality in stochastic approximation. *Ann. Math. Statist.* **39**, 1327–1332.

GOLDIE, C.M. (1991) Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Prob.* **1**, 126–166.

GOLDIE, C.M. and GRÜBEL, R. (1996) Perpetuities with thin tails. *Adv. Appl. Prob.* **28**, 463–480.

GYÖRFI, L. AND MASRY, E. (1990) The L_1 and L_2 strong consistency of recursive kernel density estimation from dependent samples. *IEEE Trans. Inf. Theory* - **IT 36**, 531–539.

KESTEN, H. (1973) Random difference equations and renewal theory for products of random matrices. *Acta Math.* **131**, 207–248.

KRENGEL, U. (1985) *Ergodic Theorems.* de Gruyter, Berlin.

LANDERS, D. AND ROGGE, L. (1978) An ergodic theorem for Fréchet-valued random variables. *Proc. Amer. Math. Soc.* **72**, 49–53.

LETAC, G. (1986) A contraction principle for certain Markov chains and its applications. *Random Matrices and their Applications. Proc. AMS-IMS-SIAM Joint Summer Research Conf. 1984.* Ed. J.E. Cohen, H. Kesten and C.M. Newman. *Contemp. Math.* **50**, 263–273, Amer. Math. Soc., Providence, R.I.

LJUNG, L., PFLUG, G. AND WALK, H. (1992) *Stochastic Approximation and Optimization of Random Systems.* Birkhäuser, Basel.

LJUNG, L. AND SÖDERSTRÖM, I. (1983) *Theory and Practice of Recursive Identification.* M.I.T. Press, Cambridge, Mass.

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

POLYAK, B.T. AND TSYPKIN, YA.Z. (1979) Adaptive estimation algorithms (convergence, optimality, stability). *Automat. Remote Control* **40**, 378–389.

REED, M. AND SIMON, B. (1980) *Methods of Modern Mathematical Physics, Vol. I, Functional Analysis.* Academic Press, New York.

RUPPERT, D. (1991) Stochastic approximation. *Handbook of Sequential Analysis.* Ed. B.K. Gosh and P.K. Sen. Marcel Dekker, New York, pp. 503–529.

TEMPEL'MAN, A.A. (1972) Ergodic theorems for general dynamical systems. *Trans. Moscow Math. Soc.* **26**, 94–132.

THAVANESWARAN, A. AND ABRAHAM, B. (1988) Estimation for nonlinear time series models using estimating equations. *J. Time Ser. Anal.* **9**, 99–108.

VENTER, J.H. (1967) An extension of the Robbins-Monro procedure. *Ann. Math. Statist.* **38**, 181–190.

VERVAAT, W. (1979) On a stochastic difference equation and representation of nonnegative infinitely divisible random variables. *Adv. Appl. Prob.* **11**, 750–783.

WALK, H. (1988) Limit behaviour of stochastic approximation processes. *Statistics & Decisions* **6**, 109–128.

WALK, H. AND ZSIDÓ, L. (1989) Convergence of the Robbins-Monro method for linear problems in a Banach space. *J. Math. Anal. Appl.* **139**, 152–177.

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