Limit Theorems for Solutions of Stochastic Difference Equations in Banach Spaces with Applications

Valery Koval
Zhitomir Institute of Technology
Cherniakhovsky Str. 103
262005 Zhitomir, Ukraine

Rainer Schwabe
Technical University Darmstadt
Department of Mathematics
Schloßgartenstr. 7
D-64289 Darmstadt, Germany

Abstract

Exact rates of convergence which are of the law of the iterated logarithm type are investigated for recursive stochastic difference equations in Banach spaces. The results are applied to autoregressive processes and to an averaging method.

1. Introduction

Let $\mathbb{B}$ be a real separable Banach-space, let $||\cdot||$ be the corresponding norm and $L(\mathbb{B})$ the space of bounded linear operators on $\mathbb{B}$. We concider the recursive stochastic equation

$$X_n = A_n X_{n-1} + D_n V_n, \qquad n \geq 1, \tag{1.1}$$

where X_0 and V_n are $\mathbb{B}$ valued random variables and A_n and D_n are random operators in $L(\mathbb{B})$. We assume that the condition

$$\lim_{n\to\infty} ||A_n - A|| = 0 \qquad \text{almost surely (a. s.)} \tag{1.2}$$

is satisfied for some $A \in L(\mathbb{B})$ and that the spectral radius r of A is smaller than 1, i. e.

$$\lim_{n\to\infty} ||A^n||^{1/n} = r < 1. \tag{1.3}$$

We want to present conditions under which

$$\limsup_{n\to\infty} \; c_n \left\| \sum\nolimits_{i=1}^{n} X_i \right\| = \alpha \qquad \text{a. s.} \tag{1.4}$$

is satisfied for some nonincreasing sequence (c_n) of real numbers, $c_n \to 0$, and some non-random constant α, $0 < \alpha < \infty$.

Similar questions have been investigated, in particular, by Chen and Guo (1988) and Bosq (1993) under more restrictive conditions. For example, Chen and Guo (1988) assumed that $\mathbb{B} = \mathbb{R}^m$ and that (V_n) is a stationary ergodic martingale difference sequence, while Bosq (1993) assumed that $A_n = A$ and $D_n = I$ are time independent, I being the identity on $\mathbb{B}$, and that the sequence (V_n) is identically distributed with finite moments $E||V_n||^{2+\delta} < \infty$, for some $\delta > 0$. Moreover, both articles are confined to $c_n = (n \log\log n)^{-1/2}$. In case $D_n = I$ for (X_n) some log log-invariance principles were proved by Koval (1992) and Pechtl (1993). In the present note we offer some generalizations.

In Section 2 we present and discuss the main result and its consequences while the corresponding proofs are deferred to Section 3. The rate of convergence result exhibited in formula (1.4) has various applications as pointed out by Chen and Guo (1988) and Bosq (1993). Further applications are presented for autoregressive processes in Section 4 and for methods related to averaging procedures (see Polyak, 1990, and Polyak and Juditsky, 1992) in Section 5.

2. Results

We establish an almost sure representation of the partial sum process $(\sum_{i=1}^n X_i)$ by the error sum process $(\sum_{i=1}^n V_i)$. For the recursive stochastic equation scheme $X_n = A_n X_{n-1} + D_n V_n$ as defined in (1.1) let (c_n) be a non-increasing sequence of positive real numbers satisfying $c_n \to 0$.

Theroem 1. Let conditions (1.2), (1.3) and the following

$$\lim_{n\to\infty} ||D_n - D|| = 0 \quad \text{a. s.}, \tag{2.1}$$

for some $D \in L(\mathbb{B})$,

$$\lim_{n\to\infty} c_n \sum_{i=1}^n c_i^{-1} ||A_i - A_{i+1}|| = 0 \quad \text{a. s.}, \tag{2.2}$$

$$\lim_{n\to\infty} c_n \sum_{i=1}^n c_i^{-1} ||D_i - D_{i+1}|| = 0 \quad \text{a. s.}, \tag{2.3}$$

and

$$\lim_{n\to\infty} c_n ||V_n|| = 0 \quad \text{a. s.} \tag{2.4}$$

be satisfied. If, additionally,

$$\sup_n c_n \left\| \sum_{i=1}^n V_i \right\| < \infty \quad \text{a. s.}, \tag{2.5}$$

holds, then

$$\lim_{n\to\infty} c_n \left\| \sum_{i=1}^n X_i - (I-A)^{-1} D \sum_{i=1}^n V_i \right\| = 0 \quad \text{a. s.} \tag{2.6}$$

Remark 1. Condition (2.5) is satisfied if (V_n) obeys a law of the iterated logarithm type.

Remark 2. For the particular choice $c_n = (n \log\log n)^{-1/2}$ of the normalizing sequence condition (2.4) holds, if e. g. either

a) the V_n are independent identically distributed random variables with finite second moment $E||V_n||^2 < \infty$; or

b) (V_n) has bounded higher moments $\sup_n E||V_n||^{2+\delta} < \infty$, for some $\delta > 0$.

Moreover, if the sequence (V_n) is independent and $c_n = (s_n^2 \log\log s_n)^{-1/2}$ where $s_n = \sup_{||f||\le 1} \left(\sum_{i=1}^n Ef^2(V_i)\right)^{1/2}$, $f \in I\!B^*$, then condition (2.4) emerges from the growth behaviour of (V_n) in Kolmogorov's law of the iterated logarithm (see Ledoux and Talagrand, 1991, § 8).

Remark 3. If e. g. $A_n = A + a_n T$ for some random operator T and some real sequence (a_n) and if $c_n = (n \log\log n)^{-1/2}$, then condition (2.2) is satisfied in case $\sup_n n^\nu |a_n - a_{n+1}| < \infty$ for some $\nu > 1$, i. e. the sequence (a_n) is sufficiently smooth. In particular, the latter holds for $a_n = n^{-\gamma}$, $\gamma > 0$. Similar considerations can be made for condition (2.3).

Remark 4. By assumption (1.3) the existence of the inverse bounded linear operator $(I - A)^{-1}$ is guaranteed.

Next, we present some consequences of Theorem 1.

Corollary 1. Under the conditions of Theorem 1 the following representation holds

$$\limsup_{n\to\infty} \; c_n \left\|\sum_{i=1}^n X_i\right\| = \limsup_{n\to\infty} \; c_n \left\|(I - A)^{-1} D \sum_{i=1}^n V_i\right\| \qquad \text{a. s.}$$

Corollary 2. If the conditions of Theorem 1 are satisfied, if the inverse operator $D^{-1} \in L(I\!B)$ exists and if

$$\limsup_{n\to\infty} \; c_n \left\|\sum_{i=1}^n V_i\right\| = v \qquad \text{a. s.}$$

holds for some finite v, then

$$\frac{v}{||D^{-1}(I - A)||} \le \limsup_{n\to\infty} \; c_n \left\|\sum_{i=1}^n X_i\right\| \le ||(I - A)^{-1} D|| \; v \qquad \text{a. s.} \tag{2.7}$$

Corollary 3. If the conditions of Theorem 1 are satisfied, if D^{-1} exists and if

$$\lim_{n\to\infty} d\left(c_n \sum_{i=1}^n V_i \, , K\right) = 0 \qquad \text{a. s.,}$$

holds, where $d(x, K)$ is the distance of x from the set K of all limiting points of $(c_n \sum_{i=1}^n V_i)$ with probability 1, then $(I-A)^{-1}DK$ is the set of all limiting points of $(c_n \sum_{i=1}^n X_i)$ with probability 1 and

$$\lim_{n\to\infty} d\left(c_n \sum\nolimits_{i=1}^n X_i, (I-A)^{-1}DK\right) = 0 \quad \text{a. s.}$$

Remark 5. Corollaries 2 and 3 generalize the corresponding results by Chen and Guo (1988) and Bosq (1993). In particular, (2.6) implies a log log-invariance principle for the solutions (X_n) of the recursive stochastic equation (1.1) (see Koval, 1992, and Pechtl, 1993).

Example 1. Let $\mathbb{B}$ be a real separable Hilbert space and (V_n) a sequence of independent identically distributed $\mathbb{B}$-valued random variables with $EV_n = 0$, $E||V_n||^2 < \infty$ and covariance operator T. Then

$$\limsup_{n\to\infty} \frac{||(I-A)^{-1}D\sum_{i=1}^n V_i||}{(n\log\log n)^{1/2}} = \left(2||(I-A)^{-1}DTD^*(I-A^*)^{-1}||\right)^{1/2} \quad \text{a. s.,}$$

where A^* and D^* are the adjoint operators to A and D (see Ledoux and Talagrand, 1991, § 8). If conditions (1.2), (1.3) and (2.1) to (2.3) are satisfied then Corollary 1 yields

$$\limsup_{n\to\infty} \frac{||\sum_{i=1}^n X_i||}{(n\log\log n)^{1/2}} = \left(2||(I-A)^{-1}DTD^*(I-A^*)^{-1}||\right)^{1/2} \quad \text{a. s.}$$

3. Proofs

First, we prove Theorem 1. We rewrite the recursive stochastic equation (1.1) as

$$X_n = AX_{n-1} + DV_n + (D_n - D)V_n + (A_n - A)X_{n-1}\,.$$

By iterating we obtain

$$\begin{aligned} X_n \;&=\; \sum\nolimits_{i=1}^n A^{n-i}DV_i + A^{n-1}A_1X_0 + \sum\nolimits_{i=1}^n A^{n-i}(D_i - D)V_i \\ &\quad + \sum\nolimits_{i=1}^{n-1} A^{n-i-1}(A_{i+1} - A)X_i\,. \end{aligned}$$

Using the identity

$$\sum\nolimits_{k=1}^n \sum\nolimits_{i=1}^k A^{k-i}T_i = (I-A)^{-1}\sum\nolimits_{i=1}^n (I - A^{n+1-i})T_i\,,$$

where (T_n) stands for any appropriate sequence of $\mathbb{B}$-valued random variables, we take the sum of X_1 to X_n and rearrange terms to get

$$\begin{aligned}\sum_{k=1}^{n} X_k &= (I-A)^{-1}\Big\{D\sum_{i=1}^{n} V_i + (I-A^n)A_1X_0 - \sum_{i=1}^{n} A^{n+1-i}D_iV_i \\ &\quad + \sum_{i=1}^{n}(D_i-D)V_i + \sum_{i=1}^{n-1}\Big[A_i - A - A^{n-i}(A_{i+1}-A)\Big]X_i\Big\}.\end{aligned}$$

Now, by applying Abel's identity to the last two expressions we find

$$\begin{aligned}&\sum_{k=1}^{n} X_k = (I-A)^{-1}\Big\{D\sum_{i=1}^{n} V_i + (I-A^n)A_1X_0 - \sum_{i=1}^{n} A^{n+1-i}D_iV_i \\ &\quad + \sum_{i=1}^{n-1}(D_i-D_{i+1})\sum_{j=1}^{i} V_j + (D_n-D)\sum_{j=1}^{n} V_j \\ &\quad + \sum_{i=1}^{n-1}\Big[A_{i+1} - A_{i+2} - A^{n-i}(A_{i+1}-A) + A^{n-i-1}(A_{i+2}-A)\Big]\sum_{j=1}^{i} X_j\Big\}.\end{aligned}$$

Let

$$Y_n = c_n\sum_{k=1}^{n} X_k \qquad \text{and} \qquad Z_n = c_n\sum_{k=1}^{n} V_k\,. \tag{3.1}$$

With these notations the above equation becomes

$$\begin{aligned}Y_n &= (I-A)^{-1}\Big\{DZ_n + c_n(I-A^n)A_1X_0 - c_n\sum_{i=1}^{n} A^{n+1-i}D_iV_i \\ &\quad + c_n\sum_{i=1}^{n-1} c_i^{-1}(D_i-D_{i+1})Z_i + (D_n-D)Z_n \\ &\quad + c_n\sum_{i=1}^{n-1} c_i^{-1}\Big[A_{i+1} - A_{i+2} - A^{n-i}(A_{i+1}-A) + A^{n-i-1}(A_{i+2}-A)\Big]Y_i\Big\}.\end{aligned}$$

Define

$$\begin{aligned}R_n &= (I-A)^{-1}\Big\{c_n(I-A^n)A_1X_0 - c_n\sum_{i=1}^{n} c_i^{-1}A^{n+1-i}D_i(c_iV_i) \\ &\quad + c_n\sum_{i=1}^{n-1} c_i^{-1}(D_i-D_{i+1})Z_i + (D_n-D)Z_n\Big\}, \\ G_{n,i} &= (I-A)^{-1}c_nc_i^{-1}\Big[A_{i+1} - A_{i+2} - A^{n-i}(A_{i+1}-A) + A^{n-i-1}(A_{i+2}-A)\Big].\end{aligned}$$

Then Y_n can, finally, be written as

$$Y_n = (I-A)^{-1}DZ_n + R_n + \sum_{i=1}^{n-1} G_{n,i}Y_i\,. \tag{3.2}$$

Next, we show that

$$\lim_{n\to\infty} ||R_n|| = 0 \ a.s. \tag{3.3}$$

In the following, let M denote a positive random variable which is finite with probability one which may differ from formula to formula. Conditions (1.3), (2.5) and the monotonicity of the sequence (c_n) yield

$$||R_n|| \le M\left(c_n + \sum_{i=1}^{n} q^{n-i}c_i||V_i|| + c_n\sum_{i=1}^{n-1} c_i^{-1}||D_i - D_{i+1}|| + ||D_n - D||\right),$$

for some q, $0 < q < 1$. Hence the convergence (3.3) of $||R_n||$ follows from conditions (2.1), (2.3), (2.4), $c_n \to 0$ and the Toeplitz Lemma. Next, we will prove that

$$\lim_{n\to\infty} \sum_{i=1}^{n-1} ||G_{n,i}|| = 0 \quad \text{a. s.} \tag{3.4}$$

This is achieved by using the bound

$$\sum_{i=1}^{n-1} ||G_{n,i}|| \leq M \left(c_n \sum_{i=1}^{n-1} c_i^{-1} ||A_{i+1} - A_{i+2}|| + \sum_{i=1}^{n} q^{n-i} ||A_i - A|| \right),$$

which, together with (1.2) and (2.2), implies the convergence result (3.4) by means of the Toeplitz Lemma. Now, by (2.5) and (3.3) we have

$$\sup_n ||(I - A)^{-1} D Z_n + R_n|| < \infty \quad \text{a. s.}$$

and by condition (3.4) and Lemma 2 in Walk and Zsidó (1989) it follows that

$$M = \sup_n ||Y_n|| < \infty \quad \text{a. s.}$$

Consequently, according to (3.2),

$$||Y_n - (I - A)^{-1} D Z_n|| \leq ||R_n|| + M \sum_{i=1}^{n-1} ||G_{n,i}||,$$

which, together with (3.3) and (3.4), implies the assertion (2.6) of the theorem. □

Remark 6. A closer look at the proof of Theorem 1 shows that condition (1.2) may be replaced by the alternative condition $\sum_{n=1}^{\infty} ||A^n|| < \infty$ used by Bosq (1993) without changing the assertion of Theorem 1.

Next, we proceed to the corollaries. Corollary 1 is an immediate consequence of Theorem 1 and the standard inequality

$$\left| \limsup_{n\to\infty} ||Y_n|| - \limsup_{n\to\infty} ||(I - A)^{-1} D Z_n|| \right| \leq \limsup_{n\to\infty} ||Y_n - (I - A)^{-1} D Z_n||,$$

where we make use of the notations defined in (3.1).

For Corollary 2 the upper bound in (2.7) is straightforward from Corollary 1. For the corresponding lower bound we use the relation (3.2). By the existence of D^{-1} we have

$$D^{-1}(I - A) Y_n = Z_n + D^{-1}(I - A) \left(R_n + \sum_{i=1}^{n-1} G_{n,i} Y_i \right)$$

and, hence,

$$v = \limsup_{n\to\infty} ||Z_n|| = \limsup_{n\to\infty} ||D^{-1}(I - A) Y_n|| \leq ||D^{-1}(I - A)|| \limsup_{n\to\infty} ||Y_n||.$$

with probability 1, which implies the lower bound in (2.7).

Corollary 3 follows from Theorem 1 by standard arguments.

4. Applications to real valued difference equations of order m

We consider the homogeneous linear stochastic difference equation

$$x_n = a_1 x_{n-1} + a_2 x_{n-2} + \ldots + a_m x_{n-m} + v_n, \quad n \geq 1, \tag{4.1}$$

in $I\!R$ with non-random real coefficients $a_1, \ldots, a_m$, where $x_0, x_{-1}, \ldots, x_{-m+1}$ and v_n, $n \geq 1$, are real valued random variables. The characteristic equation for the autoregressive process (4.1) is given by

$$\lambda^m - (a_1 \lambda^{m-1} + a_2 \lambda^{m-2} + \ldots + a_m) = 0. \tag{4.2}$$

For further readings on autoregressive processes we refer to the monograph by Brockwell and Davis (1991).

Theorem 2. Let the roots of the characteristic equation (4.2) be contained in the complex unit circle and let (c_n) and (v_n) satisfy the conditions

$$\lim_{n\to\infty} c_n |v_n| = 0 \quad \text{a. s.} \tag{4.3}$$

and

$$\sup_n c_n \left| \sum_{i=1}^n v_i \right| < \infty \quad \text{a. s.,} \tag{4.4}$$

then

$$\limsup_{n\to\infty} \; c_n \left| \sum_{i=1}^n x_i \right| = \frac{1}{\left|1 - \sum_{k=1}^m a_k\right|} \limsup_{n\to\infty} \; c_n \left| \sum_{i=1}^n v_i \right| \quad \text{a. s.} \tag{4.5}$$

Corollary 4. Let the roots of the characteristic equation (4.2) be contained in the complex unit circle, and let the sequence (v_n) be independent and identically distributed with $Ev_n = 0$ and $Ev_n^2 = \sigma^2 < \infty$, then

$$\limsup_{n\to\infty} \frac{\left|\sum_{i=1}^n x_i\right|}{(n \log\log n)^{1/2}} = \frac{\sqrt{2}\sigma}{\left|1 - \sum_{k=1}^m a_k\right|} \quad \text{a. s.}$$

Remark 7. A strong law of large numbers has been investigated for example by Gaposhkin (1988) for solutions of the homogeneous linear stochastic difference equation (4.1).

Proof of Theorem 2. Let $||\cdot||_2$ denote the Euclidean norm on $\mathbb{R}^m$, and let

$$X_n = \begin{pmatrix} x_n \\ x_{n-1} \\ \vdots \\ x_{n-m+1} \end{pmatrix}, \quad V_n = \begin{pmatrix} v_n \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} a_1 & a_2 & \cdots & a_{m-1} & a_m \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}$$

We, thus, consider the recursive m-dimensional stochastic equation

$$X_n = AX_{n-1} + V_n, \quad n \geq 1,$$

in $\mathbb{R}^m$. The matrix A is a Frobenius matrix and its characteristic polynomial is given by

$$\det(\lambda I - A) = \lambda^m - (a_1\lambda^{m-1} + a_2\lambda^{m-2} + ... + a_m). \tag{4.6}$$

Hence, by the assumptions of the theorem the spectral radius of A is less than one. Thus, by (4.3), (4.4) and Theorem 1 we obtain

$$\lim_{n\to\infty} c_n \left\|\sum_{i=1}^n X_i - (I-A)^{-1}\sum_{i=1}^n V_i\right\|_2 = 0 \quad \text{a. s.} \tag{4.7}$$

In what follows we use a representation by coordinates. We start with $(I-A)^{-1}$. As $\sum_{i=1}^n V_i = (\sum_{i=1}^n v_i, 0, ..., 0)^\top$ it suffices to determine the first column of $(I-A)^{-1}$. It can be easily shown that all entries in the first column are equal to $\det(I-A)^{-1}$. Moreover, for $\lambda = 1$ it follows from (4.6) that $\det(I-A) = 1 - \sum_{k=1}^m a_k \neq 0$, and, hence, by (4.7)

$$\lim_{n\to\infty} c_n \left|\sum_{i=1}^n x_i - \left(1 - \sum_{k=1}^m a_k\right)^{-1} \sum_{i=1}^n v_i\right| = 0 \quad \text{a. s.}$$

which proves the theorem. □

Corollary 4 follows by (4.5), $c_n = (2\sigma^2 n \log\log n)^{-1/2}$ and the classical Hartman-Wintner law of the iterated logarithm.

Remark 8. Theorem 2 can be readily extended to difference equations of order m with varying coefficients $x_n = a_{1n}x_{n-1} + a_{2n}x_{n-2} + ... + a_{mn}x_{n-m} + b_n v_n$.

5. Applications to averaging

We consider the operator equation

$$x = Ax + b \tag{5.1}$$

in a separable Banach space $I\!B$, when the spectral radius of $A \in L(I\!B)$ is smaller than 1. The solution θ of (5.1) will be searched for by the method of stepwise approximation

$$x_n = Ax_{n-1} + b, \quad n \geq 1, \quad x_0 \in I\!B.$$

In case of underlying random disturbances, which can be used e.g. to model computational errors, we have

$$Y_n = AY_{n-1} + b + V_n, \quad n \geq 1, \quad Y_0 \in I\!B. \tag{5.2}$$

In such situations the search for an approximate solution $\widehat{\theta}_n$ of (5.1) can be realized by means of the averaging method

$$\widehat{\theta}_n = \frac{1}{n} \sum\nolimits_{i=1}^{n} Y_i .$$

(see Polyak, 1990, and Polyak and Juditsky, 1992).

We present a bound on the rate of convergence for $||\widehat{\theta}_n - \theta|| \to 0$ achieved by this method under the mild regularity condition

$$\limsup_{n\to\infty} \frac{\|\sum_{i=1}^n V_i\|}{(n \log\log n)^{1/2}} \leq v \qquad \text{a. s.}$$

on the random noise. Let $X_n = Y_n - \theta$. Then X_n satisfies the recursive equation $X_n = AX_{n-1} + V_n$, $n \geq 1$. By Corollary 2 we obtain

$$\limsup_{n\to\infty} \left(\frac{n}{\log\log n} \right)^{1/2} ||\widehat{\theta}_n - \theta|| = \limsup_{n\to\infty} \frac{\|\sum_{i=1}^n X_i\|}{(n \log\log n)^{1/2}} \leq ||(I-A)^{-1}||\, v \qquad \text{a. s.}$$

which is the same rate of convergence of order $(n^{-1} \log\log n)^{1/2}$, achieved by the Robbins-Monro procedure

$$Z_n = Z_{n-1} - \frac{1}{n}(FZ_{n-1} - f + V_n), \quad n \geq 1, \quad Z_0 \in I\!B,$$

for searching the root θ of the equation $F\theta = f$, where $f \in I\!B$, $F \in L(I\!B)$, in case $\inf\{Re\lambda : \lambda \in \mathrm{spec}(F) > \frac{1}{2}\}$, (see Walk, 1988).

Acknowledgement

The work of the first author was supported by a grant of the German Academic Exchange Service (DAAD), while the work of the second author was supported by the research grant Ku719/2-1 of the German Research Association (DFG) at the Free University, Berlin.

The present research was completed while the first author was visiting the Department of Mathematics at the Free University, Berlin.

References

[1] Bosq, D. (1993). Propriétés asymptotiques des processus autoregressifs Banachiques. Applications. *C. R. Acad. Sci. Sér. 1* **316**, 607–610.

[2] Brockwell, P. J. and Davis, R. A. (1991). *Time Series: Theory and Methods. 2nd ed.* Springer, New York.

[3] Chen, H. F. and Guo, L. (1988). Laws of iterated logarithm for state variables with applications. *Acta Math. Sin. New Ser.* **4**, 343–355.

[4] Gaposhkin, V. F. (1988). On the law of large numbers for linear autoregressive processes. *Izv. VUZ. Matematika* **4**, 12–21 (in Russian).

[5] Koval, V. A. (1992). Loglog-invariance principle for solutions of stochastic difference equation in a Banach space. In *Random processes and infinite-dimensional analysis.* (Skorohod, A. V., ed.) Inst. Math. Akad. Nauk Ukraine, Kiev, 57–62 (in Russian).

[6] Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces.* Springer, New York.

[7] Pechtl, A. (1993). Arithmetic means and invariance principles in stochastic approximation. *J. Theor. Probab.* **6**, 153–173.

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

[9] Polyak, B. T. and Juditsky, A. B. (1992). Acceleration of stochastic approximation by averaging. *SIAM J. Contr. Optimiz.* **30**, 838–855.

[10] Walk, H. (1988). Limit behaviour of stochastic approximation processes. *Statist. Decis.* **6**, 109–128.

[11] 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.