Applied Mathematics, 2012, 3, 2123-2132
http://dx.doi.org/10.4236/am.2012.312A293 Published Online December 2012 (http://www.SciRP.org/journal/am)

Stochastic Approximation Method for Fixed Point Problems

Ya. I. Alber[1], C. E. Chidume[2], Jinlu Li[3]
[1]Department of Mathematics, The Technion-Israel Institute of Technology, Haifa, Israel
[2]The Abdus Salam International Centre for Theoretical Physics, Trieste, Italy
[3]Department of Mathematical Sciences, Shawnee State University, Portsmouth, USA
Email: alberya@yahoo.com, alberya@gmail.com

Received August 28, 2012; revised November 28, 2012; accepted December 5, 2012

ABSTRACT

We study iterative processes of stochastic approximation for finding fixed points of weakly contractive and nonexpansive operators in Hilbert spaces under the condition that operators are given with random errors. We prove mean square convergence and convergence almost sure (a.s.) of iterative approximations and establish both asymptotic and nonasymptotic estimates of the convergence rate in degenerate and non-degenerate cases. Previously the stochastic approximation algorithms were studied mainly for optimization problems.

Keywords: Hilbert Spaces; Stochastic Approximation Algorithm; Weakly Contractive Operators; Nonexpansive Operators; Fixed Points; Convergence in Mean Square; Convergence Almost Sure (a.s.); Nonasymptotic Estimates of Convergence Rate

1. Introduction

In this paper the following problem is solved: To find a fixed point x^* of the operator $T:G\to H$, in other words, to find a solution $x^*\in G$ of the equation

$$x=Tx, \tag{1.1}$$

where T is a Lipschitz continuous mapping, H is a Hilbert space, $G\subseteq H$ is a closed convex subset. We suppose that x^* exists, *i.e.*, the fixed point set N of T is nonempty. Note in different particular cases of the Equation (1.1), for example, when $T:G\to G$, the solution existence and solution uniqueness can be proved under some additional assumptions.

We separately consider two classes of mappings *T*: the class of weakly contractive maps and more general class of nonexpansive ones. Let us recall their definitions.

Definition 1.1. *A mapping* $T:G\to H$ *is said to be weakly contractive of class* $C_{\phi(t)}$ *on a closed convex subset* $G\subset H$ *if there exists a continuous and increasing function* $\phi(t)$ *defined on* IR^+ *such that* ϕ *is positive on* $IR^+/\{0\},\phi(0)=0,\lim_{t\to+\infty}\phi(t)=+\infty$, *and for all*

$$x,y\in G,\|Tx-Ty\|\le\|x-y\|-\phi(\|x-y\|). \tag{1.2}$$

Remark 1.2. *It follows from* (1.2) *that* $\psi(t)\le t$ *and in real problems an argument* t *of the function* $\psi(t)$ *doesn't necessary approaches to* ∞ *obeying the condition* $T:G\to H$ (*see the example in Remark* 3.4).

Definition 1.3. *A mapping* $T:G\to H$ *is said to be nonexpansive on the closed convex subset* $G\subset H$ *if for all* $x,y\in G$

$$\|Tx-Ty\|\le\|x-y\|.$$

It is obvious that the class of weakly contractive maps is contained in the class of nonexpansive maps because the right-hand side of (1.2) is estimated as

$$0\le\|x-y\|-\phi(\|x-y\|)\le\|x-y\|, \tag{1.3}$$

and it contains the class of strongly contractive maps because $\phi(t)=(1-q)t$ with $0<q<1$ gives us

$$\|Tx-Ty\|\le q\|x-y\|. \tag{1.4}$$

We study the following algorithm of stochastic approximation:

$$x_{n+1}=Pr_G(x_n-\alpha_nS_nx_n),n=1,2,\cdots,x_1\in G, \tag{1.5}$$

where Pr_G is the metric projection operator from H onto *G* and deterministic step-parameters α_n satisfy the standard conditions:

$$\sum_{n=1}^{\infty}\alpha_n=\infty \text{ and } \sum_{n=1}^{\infty}\alpha_n^2<\infty. \tag{1.6}$$

The factor S_nx_n in (1.5) is an infinite-dimensional vector of random observations of the clearance operator $F=I-T$ at random points $x_n\in G$ given for all $n\ge1$ on the same probability space (Ω,A,P). We set

$$S_n x_n = \zeta_n + \xi_n, \quad (1.7)$$

where $\zeta_n = x_n - Tx_n$ and ξ_n is a sequence of independent random vectors with the conditions

$$E[\xi_n] = 0 \text{ and } E\left[\|\xi_n\|^2\right] \le C_1, 0 < C_1 < \infty. \quad (1.8)$$

Here E is a symbol of the mathematical expectation. In order to calculate conditional mathematical expectations of different random variables we define the σ-subalgebra $A_n := \sigma(x_1, x_2, \cdots, x_n)$ on (Ω, A, P). And then $E[\xi_n / A_n]$ means A_n-measurable function with the following property: for any $B \in A_n$

$$\int_B \xi_n \mathrm{d}P(\omega) = \int_B E[\xi_n / A_n] \mathrm{d}P(\omega).$$

We also assume in the sequel that ζ_n is A_n-measurable for all $n \ge 1$.

Let us recall the mean square convergence and almost sure (a.s.) convergence.

We say that the sequence $\{\xi_n\}$ of random variables $\xi_n(\omega)$ converges in mean square to ξ if ξ exists and

$$\lim_{n\to\infty} E\left[\|\xi_n - \xi\|^2\right] = 0.$$

The sequence ξ_n converges to ξ almost surely or with probability 1 if

$$P\left(\lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)\right) = 1.$$

Almost sure convergence and convergence in mean square imply convergence in the sense of probability: The sequence $\{\xi_n\}$ of random variables $\xi_n(\omega)$ converges in the sense of probability to $\xi(\omega)$ if for all $\varepsilon > 0$

$$\lim_{n\to\infty} P\left(\|\xi_n(\omega) - \xi(\omega)\| \ge \varepsilon\right) = 0.$$

So, we consider iterative processes of stochastic approximation in the form (1.5) for finding fixed points of weakly contractive (Definition 1.1) and nonexpansive (Definition 1.3) mappings in Hilbert spaces under the conditions (1.8). We prove mean square convergence and convergence almost sure of iterative approximations and establish both asymptotic and nonasymptotic estimates of the convergence rate. Perhaps, we present here the first results of this sort for fixed point problems. Formerly the stochastic approximation methods were studied mainly to find minimal and maximal points in optimization problems (see, for example, [1-6] and references within).

2. Auxiliary Recurrent Inequalities

Lemma 2.1. [3,4] *Let* $\{\mu_k\}, \{\rho_k\}$ and $\{\gamma_k\}$ *be sequences of nonnegative real numbers satisfying the recurrent inequality.*

$$\mu_{k+1} \le (1+\rho_k)\mu_k + \gamma_k, k = 1, 2, \cdots \quad (2.1)$$

Assume that $\sum_{k=1}^{\infty} \rho_k < \infty$ and $\sum_{k=1}^{\infty} \gamma_k < \infty$. Then $\{\mu_k\}$ is bounded and converges to some limit.

Lemma 2.2. [3,4] *Let* $\{\mu_k\}, \{\zeta_k\}, \{\alpha_k\}$ and $\{\gamma_k\}$ *be sequences of nonnegative real numbers satisfying the recurrent inequality.*

$$\mu_{k+1} \le (1+\rho_k)\mu_k - \alpha_k \psi(\mu_k) + \gamma_k, k = 1, 2, \cdots, \quad (2.2)$$

where $\sum_{k=1}^{\infty} \alpha_k = \infty, \sum_{k=1}^{\infty} \rho_k < \infty$ and either $\sum_{k=1}^{\infty} \gamma_k < \infty$ or

$$\lim_{k\to\infty} \frac{\gamma_k}{\alpha_k} = 0. \quad (2.3)$$

Assume that $\psi(t)$ is continuous and increasing function defined on IR^+ such that ψ is positive on $IR^+/\{0\}$, $\psi(0) = 0$. Then $\lim_{k\to\infty} \mu_k = 0$. There exists an infinite subsequence $\{k_l\}, l = 1, 2, \cdots,$ such that

$$\mu_{k_l} \le C_0 \psi^{-1}\left(\frac{1}{\sum_{k=1}^{k_l} \alpha_k} + \frac{\gamma_{k_l}}{\alpha_{k_l}}\right),$$

where $C_0 \ge \prod_{k=1}^{\infty}(1+\rho_k)$.

In the following two lemmas we want to present nonasymptotic estimates for the whole sequence $\mu_k, k \ge 1$. For this the stronger requirements are made of parameters α_k and function $\psi(t)$ in the recurrent inequality.

Suppose that $\alpha(t)$ such that $\alpha(k) = \alpha_k$, $F(t)$ and $\Phi(t)$ are antiderivatives from $\alpha(t)$ and $\frac{1}{\psi(t)}$, respectively, with arbitrary constants C (without loss of generality, one can put $C = 0$), *i.e.*

$$F(t) = \int \alpha(t) \mathrm{d}t, \Phi(t) = \int \frac{\mathrm{d}t}{\psi(t)}.$$

Observe that $F(t)$ has the following properties:

i) $F'(t) = \alpha(t)$;

ii) $F(t)$ is strictly increasing on $[1, \infty)$ and $F(t) \to \infty$ as $t \to \infty$;

iii) The function $g(t) = \frac{\alpha(t)}{F(t)} = \frac{F'(t)}{F(t)}$ is decreasing;

iv) $G(t) = \ln F(t) \to \infty$ as $t \to \infty$.

Introduce the following denotations:

1) $\psi^{-1}(z)$ and $\Phi^{-1}(z)$ are the inverse functions to $\psi(t)$ and $\Phi(t)$, respectively;

2) $v(s) = \psi^{-1}\left[\frac{c_0 \gamma(s-1)}{\alpha(s-1)}\right] + \gamma(s-1), s \ge 2, c_0 > 1$ is a

fixed control parameter;

3) $u(s,C)=\Phi^{-1}\left[\Phi(C)-a\left(F(s)-F(2)\right)\right]$, $C\ge 0, a=\dfrac{c_0-1}{c_0}<1$;

4) $w(s,v(r))=\Phi^{-1}\left[\Phi(v(r))-a\left(F(s)-F(2)\right)\right]$, $2\le r<\infty$;

5) $Q(\bar{c})=\Phi^{-1}\left[\Phi(\bar{c})-a\left(F(2)-F(1)\right)\right]$, where $\bar{c}>0$ is an arbitrary fixed number;

We present now the based condition (**P**): The graphs of the scalar functions $v(s)$ and $w(s,v(r))$ with any fixed $r\in[2,\infty)$ are intersected on the interval $[2,\infty)$ not more than at two points s_1 and s_2 (we do not consider contact points as intersection ones excepting $s=2$ if any).

For example, the graphs of the functions $v(s)$ and $w(s,v(r))$ calculated for $\alpha(s)=\dfrac{b}{s^{\kappa}}, b>0, 0<\kappa\le 1,$ and $\psi(t)=t^{\nu}, \nu\ge 1,$ satisfy the condition (**P**).

Lemma 2.3. [3,4] *Assume that 1) the property* (**P**) *is carried out for the function* $w(s,v(2))=u(s,v(2))$ *and* $v(s)$; 2) $u(s,v(2))\ge v(s)$ *as* $s\to\infty$; 3) *the control parameter* c_0 *is chosen such that*

$$u(s,v(2))\ge v(s) \text{ as } s\to 2. \tag{2.4}$$

Then for the sequence $\{\mu_k\}$ generated by the inequality

$$\mu_{k+1}\le\mu_k-\alpha_k\psi(\mu_k)+\gamma_k,\ k=1,2,\cdots, \tag{2.5}$$

it follows: $\lim_{k\to\infty}\mu_k=0$ and for all $k\ge 1$

$$\mu_k\le u(k,C), C=\max\{Q(\mu_1),v(2)\}. \tag{2.6}$$

Lemma 2.4. [3,4] *Assume that* 1) *the property* (**P**) *is carried out for all the function* $w(s,v(2r))$ *and* $v(s)$; 2) $u(s,v(2))\le v(s)$ *as* $s\to\infty$. *Then for the sequence* $\{\mu_k\}$ *generated by the inequality* (2.5) $\lim_{k\to\infty}\mu_k=0$. *In addition*,

a) if $Q(\mu_1)\le v(2)$ and the control parameter c_0 is chosen such that $u(s,v(2))\le v(s)$ as $s\to 2$, then for all $k\ge 2$

$$\mu_k\le v(k); \tag{2.7}$$

b) in all remaining cases

$$\mu_k\le u(k,C), C=\max\{Q(\mu_1),v(2)\},\quad 1\le k\le\bar{s}, \tag{2.8}$$

$$\mu_k\le v(k), k>\bar{s}, \tag{2.9}$$

where $\bar{s}$ is a unique root of the equation

$$u(s,C)=v(s) \tag{2.10}$$

on the interval $[2,\infty)$.

The following lemmas deal with another sort of recurrent inequalities:

Lemma 2.5. [7,8] *Let* $\{\mu_k\},\{\alpha_k\},\{\beta_k\}$ and $\{\gamma_k\}$ *be sequences of non-negative real numbers satisfying the recurrence inequality.*

$$\mu_{k+1}\le\mu_k-\alpha_k\beta_k+\gamma_k,\ k=1,2,\cdots. \tag{2.11}$$

Assume that

$$\sum_{k=1}^{\infty}\alpha_k=\infty \text{ and } \sum_{k=1}^{\infty}\gamma_k<\infty.$$

Then:

i) There exists an infinite subsequence $\{\beta_{\ell_k}\}\subset\{\beta_k\}$ such that

$$\beta_{\ell_k}\le\frac{1}{\sum_{j=1}^{\ell_k}\alpha_j}, \tag{2.12}$$

and, consequently, $\lim_{k\to\infty}\beta_{\ell_k}=0$;

ii) if $\lim_{k\to\infty}\alpha_k=0$ and there exists $\kappa>0$ such that

$$|\beta_{k+1}-\beta_k|\le\kappa\alpha_k \tag{2.13}$$

for all $k\ge 1$, then $\lim_{k\to\infty}\beta_k=0$.

Lemma 2.6. [7,8] *Let* $\{\mu_k\},\{\alpha_k\},\{\beta_k\}$ and $\{\gamma_k\}$ *be sequences of non-negative real numbers satisfying the recurrence inequality* (2.11). *Assume that* $\sum_{k=1}^{\infty}\alpha_k=\infty$ *and* (2.3) *is satisfied. Then there exists an infinite subsequence* $\{\beta_{\ell_k}\}\subset\{\beta_k\}$ *such that* $\lim_{k\to\infty}\beta_{\ell_k}=0$.

3. Mean Square Convergence of Stochastic Approximations

Theorem 3.1. *Assume that* $T:G\to H$ *is a weakly contractive mapping of the class* $C_{\phi(\tau)}$, $\psi_1(\tau)=\tau\phi(\tau)$ *is a convex function with respect to* $t=\tau^2$ *and* $E\left[\|x_1-x^*\|^2\right]<\infty$. *Then the sequence* $\{x_n\}$ *generated by* (1.5)-(1.7) *converges in mean square to a unique fixed point* x^* *of* T. *There exists an infinite subsequence* $\{n_l\}, l=1,2,\cdots,$ *such that*

$$E\left[\|x_{n_l}-x^*\|^2\right]\le C_0\psi^{-1}\left(\frac{1}{2\sum_{n=1}^{n_l}\alpha_n}+C_1\alpha_{n_l}\right), \tag{3.1}$$

where $\psi(\tau^2)=\psi_1(\tau)$ and some positive constant C_0 satisfies the inequality

$$\prod_{n=1}^{\infty}\left(1+8\alpha_n^2\right)\le C_0. \tag{3.2}$$

Remark 3.2. $1 < C_0 < \infty$ *exists in view of the second condition in* (1.6).

Proof. First of all, we note that the method (1.5)-(1.7) guarantees inclusion $x_n \in G$ for all $n \geq 1$. Since the metric projection operator Pr_G is nonexpansive in a Hilbert space and $x^* \in G$ exists, we can write

$$\begin{aligned}
&\left\|x_{n+1} - x^*\right\|^2 \\
&= \left\|Pr_G\left(x_n - \alpha_n\left(x_n - Tx_n + \xi_n\right)\right) - Pr_G x^*\right\|^2 \\
&\leq \left\|x_n - x^*\right\|^2 - 2\alpha_n\left(\zeta_n + \xi_n, x_{n+1} - x^*\right) \\
&= \left\|x_n - x^*\right\|^2 - 2\alpha_n\left(\zeta_n, x_n - x^*\right) \\
&-2\alpha_n\left(\xi_n, x_n - x^*\right) + \left\|x_{n+1} - x_n\right\|^2.
\end{aligned} \tag{3.3}$$

Let us evaluate the first scalar product in (3.3). We have

$$\begin{aligned}
&\left(x_n - Tx_n, x_n - x^*\right) \\
&= \left(x_n - x^* - Tx_n + Tx^*, x_n - x^*\right) \\
&= \left\|x_n - x^*\right\|^2 - \left(Tx_n - Tx^*, x_n - x^*\right) \\
&\geq \left\|x_n - x^*\right\|^2 - \left\|Tx_n - Tx^*\right\|\left\|x_n - x^*\right\| \\
&\geq \left\|x_n - x^*\right\|^2 - \left(\left\|x_n - x^*\right\| - \phi\left(\left\|x_n - x^*\right\|\right)\right)\left\|x_n - x^*\right\| \\
&\geq \phi\left(\left\|x_n - x^*\right\|\right)\left\|x_n - x^*\right\| = \psi_1\left(\left\|x_n - x^*\right\|\right)
\end{aligned} \tag{3.4}$$

We remember that $\psi\left(\tau^2\right) = \psi_1(\tau)$. Then the inequalities (3.3) and (3.4) yield

$$\begin{aligned}
&\left\|x_{n+1} - x^*\right\|^2 \leq \left\|x_n - x^*\right\|^2 - 2\alpha_n\psi\left(\left\|x_n - x^*\right\|^2\right) \\
&-2\alpha_n\left(\xi_n, x_n - x^*\right) + \alpha_n^2\left\|\zeta_n + \xi_n\right\|^2.
\end{aligned} \tag{3.5}$$

Applying the conditional expectation with respect to A_n to the both sides of (3.5) we obtain

$$\begin{aligned}
&E\left[\left\|x_{n+1} - x^*\right\|^2 / A_n\right] \\
&\leq \left\|x_n - x^*\right\|^2 - 2\alpha_n\psi\left(\left\|x_n - x^*\right\|^2\right) \\
&-2\alpha_n E\left[\left(\xi_n, x_n - x^*\right)/A_n\right] + \alpha_n^2 E\left[\left\|\zeta_n + \xi_n\right\|^2 / A_n\right].
\end{aligned} \tag{3.6}$$

It is easy to see that

$$\begin{aligned}
&E\left[\left\|\zeta_n + \xi_n^2\right\|/A_n\right] \\
&\leq 2E\left[\left\|\zeta_n\right\|^2/A_n\right] + 2E\left[\left\|\xi_n\right\|^2/A_n\right] \\
&= 2\left\|\zeta_n\right\|^2 + 2E\left[\left\|\xi_n\right\|^2/A_n\right].
\end{aligned} \tag{3.7}$$

Since T is weakly contractive and therefore nonexpansive, one gets

$$\left\|x_n - Tx_n\right\|^2 \leq 4\left\|x_n - x^*\right\|^2.$$

Taking into account (3.7), the inequality (3.9) is estimated as follows:

$$\begin{aligned}
&E\left[\left\|x_{n+1} - x^*\right\|^2 / A_n\right] \\
&\leq \left(1 + 8\alpha_n^2\right)\left\|x_n - x^*\right\|^2 - 2\alpha_n\psi\left(\left\|x_n - x^*\right\|^2\right) \\
&-2\alpha_n E\left[\left(\xi_n, x_n - x^*\right)/A_n\right] + 2\alpha_n^2 E\left[\left\|\xi_n\right\|^2/A_n\right].
\end{aligned} \tag{3.8}$$

Now the unconditional expectation implies

$$\begin{aligned}
&E\left[\left\|x_{n+1} - x^*\right\|^2\right] \\
&\leq \left(1 + 8\alpha_n^2\right)E\left[\left\|x_n - x^*\right\|^2\right] - 2\alpha_n E\left[\psi\left(\left\|x_n - x^*\right\|^2\right)\right] \\
&-2\alpha_n E\left[\left(\xi_n, x_n - x^*\right)\right] + 2\alpha_n^2 E\left[\left\|\xi_n\right\|^2\right]
\end{aligned} \tag{3.9}$$

Next we need the Jensen inequality for a convex function $\psi\left(\left\|x_n - x^*\right\|^2\right)$:

$$E\left[\psi\left(\left\|x_n - x^*\right\|^2\right)\right] \leq \psi\left(E\left[\left\|x_n - x^*\right\|^2\right]\right)$$

(see [9,10]). This allows us to rewrite (3.9) in the form

$$\begin{aligned}
&E\left[\left\|x_{n+1} - x^*\right\|^2\right] \\
&\leq \left(1 + 8\alpha_n^2\right)E\left[\left\|x_n - x^*\right\|^2\right] \\
&-2\alpha_n\psi\left(E\left[\left\|x_n - x^*\right\|^2\right]\right) + 2\alpha_n^2 E\left[\left\|\xi_n\right\|^2\right]
\end{aligned} \tag{3.10}$$

because of

$$E\left[\left(\xi_n, x_n - x^*\right)\right] = 0.$$

Denoting $\lambda_n = E\left[\left\|x_n - x^*\right\|^2\right]$ we have

$$\lambda_{n+1} \leq \left(1 + 8\alpha_n^2\right)\lambda_n - 2\alpha_n\psi\left(\lambda_n\right) + 2C_1\alpha_n^2, \tag{3.11}$$

where in view of Definition 1.1, $\psi(\lambda)$ is a continuous and increasing function with $\psi(0) = 0$. Due to (6), from Lemma 2.2 it follows

$$\lim_{n\to\infty} E\left[\left\|x_n - x^*\right\|^2\right] = 0$$

and the estimate (3.1) holds too. The theorem is proved. □

Remark 3.3. *If a fixed point of weakly contractive mapping* $T: G \to H$ *exists, then it is unique* [11].

Remark 3.4. *The following example was presented in* [11]: *Let* $Tx = \sin x$, $G = [0,1]$ *and* $T: G \to G$. *It has been shown in* [11] *that*

$$\left|\sin x-\sin y\right|\le\left|x-y\right|-\frac{1}{8}\left|x-y\right|^{3}$$

for all $0\le x\le y\le 1$. Then

$$\phi(\tau)=\frac{1}{8}\tau^{3},\psi_{1}(\tau)=\frac{1}{8}\tau^{4}\text{ and }\psi(t)=\frac{1}{8}t^{2}.$$

Definition 3.5. *Let a nonexpansive mapping $T:G\to H$ have a unique fixed point x^{*}. T is said to be weakly sub-contractive on the closed convex subset $G\subset H$ if there exists continuous and increasing function $\psi(t)$ defined on IR^{+} such that ψ is positive on $IR^{+}/\{0\}$, $\psi(0)=0$, $\lim_{t\to+\infty}\phi(t)=+\infty$, such that for all $x\in G$.*

$$\left\langle x-Tx,x-x^{*}\right\rangle\ge\psi\left(\left\|x-x^{*}\right\|^{2}\right).\qquad(3.12)$$

Theorem 3.6. *Assume that a mapping $T:G\to H$ is weakly sub-contractive and the function $\psi(t)$ in* (3.12) *is convex on IR^{+} Then the results of Theorem* 3.1 *holds for the sequence $\{x_{n}\}$ generated by* (1.5)-(1.7).

The second inequality in (1.6) can be omitted if we assume not less than linear growth of $\psi(\lambda)$ "on infinity" and put $\alpha_{n}\to 0$ as $n\to\infty$

Theorem 3.7. *Assume that a mapping $T:G\to H$ is weakly sub-contractive and the function $\psi(t)$ in* (3.12) *is convex on IR^{+} Suppose that instead of* (1.6) *the conditions*

$$\lim_{n\to\infty}\alpha_{n}=0\text{ and }\sum_{n=1}^{\infty}\alpha_{n}=\infty.\qquad(3.13)$$

hold. In addition, let $\alpha_{n}\le\bar{\alpha}<0.5$ and

$$\lim_{\lambda\to\infty}\frac{\psi(\lambda)}{\lambda}\ge 4\bar{\alpha}.\qquad(3.14)$$

Then the sequence $\{x_{n}\}$ generated by (1.5), (1.7) and (3.13) converges in mean square to x^{*}. There exists an infinite subsequence $\{n_{l}\},l=1,2,\cdots,$ such that

$$E\left[\left\|x_{n_{l}}-x^{*}\right\|^{2}\right]\le\psi^{-1}\left(\frac{1}{2\sum_{n=1}^{n_{l}}\alpha_{n}}+\frac{C_{3}\alpha_{n_{l}}}{2}\right).$$

where

$$C_{3}=2C_{1}+8\max\left\{\left\|x^{*}-x_{1}\right\|^{2},C_{2}\right\},$$
$$C_{2}=\lambda_{*}+2C_{1}\bar{\alpha}^{2}+8\bar{\alpha}^{2}\lambda_{*},\qquad(3.15)$$
$$\lambda_{*}=\psi^{-1}\left(C_{1}\bar{\alpha}+4\bar{\alpha}\lambda_{*}\right).$$

Proof. Consider the inequality (11) in the form

$$\lambda_{n+1}\le\lambda_{n}-2\alpha_{n}\psi(\lambda_{n})+2C_{1}\alpha_{n}^{2}+8\alpha_{n}^{2}\lambda_{n},\qquad(3.16)$$

where $\lambda_{n}=E\left[\left\|x_{n}-x^{*}\right\|^{2}\right]$. Observe that it is derived by making use of (3.4) and the nonexpansivity property of T. We shall show that λ_{n} are bounded for all $n\in I=\{1,2,\cdots\}$. Indeed, since $\psi(\lambda)$ is a convex increasing continuous function, we conclude that $\psi_{1}(\lambda)=\frac{\psi(\lambda)}{\lambda}$ is nondecreasing and since (3.14) holds, the inequality $\psi(\lambda)\le C_{1}\bar{\alpha}+4\bar{\alpha}\lambda$ has a solution $\lambda\le\lambda_{*}$, where λ_{*} is the unique root of the scalar equ0 ation $\psi(\lambda)=C_{1}\bar{\alpha}+4\bar{\alpha}\lambda$. Together with this, (3.4) and (3.14) are co-ordinated by the parameter $\bar{\alpha}<0.5$.

Only one alternative can happen for each $n\in I$: either

$$H_{1}:=-2\alpha_{n}\psi(\lambda_{n})+2C_{1}\alpha_{n}^{2}+8\alpha_{n}^{2}\lambda_{n}\ge 0$$

or

$$H_{2}:=-2\alpha_{n}\psi(\lambda_{n})+2C_{1}\alpha_{n}^{2}+8\alpha_{n}^{2}\lambda_{n}<0.$$

Denote $I_{1}=\{n\in I|H_{1}\text{ is true}\}$ and $I_{2}=\{n\in I|H_{2}\text{ is true}\}$. It is clear that $I_{1}\cup I_{2}=I$. From the hypothesis H_{1}, it arises

$$\psi(\lambda_{n})\le C_{1}\bar{\alpha}+4\bar{\alpha}\lambda_{n},$$

and then $\lambda_{n}\le\lambda_{*}$ for all $n\in I_{1}$. From the hypothesis H_{2}, we have: $\lambda_{n+1}<\lambda_{n}$ for all $n\in I_{2}$. Consider all the possible cases:

1) $I_{2}=\varnothing$. Then $\lambda_{n}\le\lambda_{*}$ for all $n\in I$.

2) $I_{1}=\varnothing$. Then $\lambda_{n}\le\lambda_{1}$ for all $n\in I$.

3) Let $I_{1}=\{1,2,\cdots,N_{0}\}$ and $I_{2}=\{N_{0}+1,N_{0}+2,\cdots\}$. Then $\lambda_{n}\le\lambda_{*}$ for $n=1,2,\cdots,N_{0}$. By (3.16), $\lambda_{N_{0}+1}\le C_{2}$. It is obvious that $\lambda_{n}\le C_{2}$ for $N_{0}+2,N_{0}+3,\cdots$. Therefore, $\lambda_{n}\le C_{2}$ for all $n\in I$.

4) Let $I_{2}=\{1,2,\cdots,N_{0}\}$ and $I_{1}=\{N_{0}+1,N_{0}+2,\cdots\}$. Then $\lambda_{n}\le\lambda_{1}$ for $n=1,2,\cdots,N_{0}$. and $\lambda_{n}\le\lambda_{*}$ for $N_{0}+1,N_{0}+2,\cdots$ Thus, $\lambda_{n}\le\lambda_{*}$ for all $n\in I$.

5) Let I_{1} and I_{2} be unbounded sets. Consider an arbitrary interval

$$I_{s}=\left[n_{s}+1,n_{s+1}-1\right]\subset I_{2},$$

where $n_{s},n_{s+1}\in I_{1},s=1,3,5,\cdots$. It is easy to be sure that $\lambda_{n_{s}}\le\lambda_{*}$ and $\lambda_{n}\le C_{2}$ for all $n\in I_{s}$.

6) The other situations of bounded and unbounded sets I_{1} and I_{2} are covered by the items 1)-5). Consequently, we have the final result: $\lambda_{n}\le\max\{\lambda_{1},C_{2}\}$ for all $n\in I$.

Thus, we obtain the inequality

$$\lambda_{n+1}\le\lambda_{n}-2\alpha_{n}\psi(\lambda_{n})+C_{3}\alpha_{n}^{2},\qquad(3.17)$$

where C_{3} is defined by (3.15). Now Lemma 2.2 with the condition (2.3) implies the result. □

Remark 3.8. *For a linear function $\psi(\tau)=c\tau$ which* *is convex and concave at the same time we suppose* $2 \geq c > 4\bar{\alpha}$.

Remark 3.9. *If* G *is bounded or more generally* $\{x_n\}$ *is bounded, then the inequality* (3.17) (*with some different constant* C_3) *immediately follows from* (3.16).

4. Estimates of the Mean Square Convergence Rate

Using Lemmas 2.3 and 2.4 we are able to give two general theorems on the nonasymptotic estimates of the mean square convergence rate for sequence $\{x_n\}$ generated by the stochastic approximation algorithm (1.5)-(1.7).

Again we introduce denotations 1)-5) from Section 2 induced now by the recurrent inequality (3.11):

1) $\psi^{-1}(z)$ and $\Phi^{-1}(z)$ are the inverse functions to $\psi(t)$ and $\Phi(t)$, respectively;

2) $v(s) = \psi^{-1}\left[C_1 c_0 \alpha(s-1)\right] + 2C_1\alpha^2(s-1), c_0 > 1$ is a fixed control parameter;

3) $u(s,C_2) = \Phi^{-1}\left[\Phi(C_2) - a\left(F(s) - F(2)\right)\right], C_2 > 0,$

$a = \dfrac{c_0 - 1}{c_0} < 1$

4) $w(s,v(r)) = \Phi^{-1}\left[\Phi(v(r)) - a\left(F(s) - F(2)\right)\right],$

$2 \leq s, r < \infty;$

5) $Q = \Phi^{-1}\left[\Phi\left(\left\|x_1 - x^*\right\|^2\right) - a\left(F(2) - F(1)\right)\right].$

Introduce also the basic condition (**P**).

Theorem 4.1. *Assume that all the conditions of Theorem 3.1 are fulfiled and*

i) the condition (**P**) holds for the functions $w(s,v(2)) = u(s,v(2))$ and $v(s)$;

ii) $u(s,v(2)) \geq v(s)$ as $s \to \infty$;

iii) $c_0 > 1$ is chosen such that $u(s,v(2)) \geq v(s)$ as $s \to 2$.

Then the sequence $\{x_n\}$ generated by (1.5)-(1.7) converges in average to a unique fixed point x^* of T and for all $n \geq 1$

$$E\left[\left\|x_n - x^*\right\|^2\right] \leq C_0 u(n, C_2), \quad C_2 = \max\{Q, v(2)\}. \tag{4.1}$$

Theorem 4.2. *Assume that all the conditions of Theorem* 3.1 *are fulfiled and*

i) the condition (**P**) holds for the functions $v(s)$ and $w(s,v(z))$ with any fixed $z \in [2,\infty)$;

ii) $u(s,v(2)) \leq v(s)$ as $s \to \infty$;

iii) If $Q \leq v(2)$ and $c_0 > 1$ is chosen such that $u(s,v(2)) \leq v(s)$ as $s \to 2$, then the sequence $\{x_n\}$ generated by (1.5)-(1.7) converges in average to a unique fixed point x^* of T and for all $n \geq 1$

$$E\left[\left\|x_n - x^*\right\|^2\right] \leq C_0 v(n); \tag{4.2}$$

iv) In all the remaining cases, (4.1) holds for $1 \leq n \leq \bar{s}$ and (4.2) for $n \geq \bar{s}$, where $\bar{s}$ is a unique root of the equation $u(s,C_2) = v(s)$ on the interval $[2,\infty)$.

Let us provide the examples of functions $\psi(\tau)$ and $\Phi(\lambda)$ suitable for Theorems 4.1 and 4.2 (see [12,13]).

1) Below in Corollaries 4.3-4.6 we use the functions $\psi(\tau) = \tau^\nu$ with $\nu \geq 1$. For them

$$\Phi(\lambda) = \begin{cases} \ln\lambda, \text{ if } \nu = 1, \\ \dfrac{\lambda^{1-\nu}}{1-\nu}, \text{ if } \nu \neq 1, \end{cases} \tag{4.3}$$

and

$$\Phi^{-1}(z) = \begin{cases} \exp z, \text{ if } \nu = 1, \\ \left[(1-\nu)z\right]^{\frac{1}{1-\nu}}, \text{ if } \nu \neq 1, \end{cases} \tag{4.4}$$

2) If $\psi(\tau) = \exp\tau - 1, \tau > 0,$ then

$$\Phi(\tau) = \ln(1 - \exp\tau) \text{ and } \Phi^{-1}(z) = -\ln(1 - \exp z).$$

3) If $\psi(\tau) = \dfrac{\tau}{\tau+1}, \tau > 0,$ then

$$\Phi(\tau) = \ln\tau + \tau \text{ and } \Phi^{-1}(z) \leq \ln(1 + \exp z).$$

4) If $\psi(\tau) = \dfrac{\tau^2}{\tau+1}, \tau > 0,$ then

$$\Phi(\tau) = \ln\tau - \frac{1}{\tau}.$$

In this example we are unable to define $\Phi^{-1}(z)$ in analitical form, therefore suggest to calculate it numerically by computer.

We next present very important corollaries from Theorems 4.1 and 4.2, where their assumptions automatically guarantee accomplishment of the condition (**P**) (see [4]). The functions $\psi(\tau)$ coincide with the point 1) above.

Corollary 4.3. *Assume that* $T: G \to H$ *is a strongly contractive mapping, that is*, (1.4) *is satisfied with* $0 < q < 1$. *Let in* (1.5) $\alpha_n = \dfrac{b}{n}, b > 0.$ *Then*

$$F(t) = b\ln t, \Phi(\tau) = \frac{\ln\tau}{q_1}, \Phi^{-1}(z) = \mathrm{e}^{q_1 z}, q_1 = 1 - q,$$

$$v(s) = \left(\frac{2c_0 C_1}{b} + 2C_1\right)\frac{1}{s-1}, u(s,C) = C\left(\frac{2}{s}\right)^{abq_1},$$

$$Q = \left\|x_1 - x^*\right\|^2 \left(\frac{1}{2}\right)^{abq_1}.$$

I. Suppose that $b > \frac{1}{q_1}$ and $c_0 \ge \frac{bq_1}{bq_1 - 1}$. Then $\lim_{n\to\infty} E\left[\left\|x_n - x^*\right\|^2\right] = 0$ and

1) If $Q \le v(2)$ and $c_0 \ge \frac{bq_1}{bq_1 - 2}$, we have for all $n > 1$

$$E\left[\left\|x_n - x^*\right\|^2\right] \le v(n); \tag{4.5}$$

2) In all the remain cases

$$E\left[\left\|x_n - x^*\right\|^2\right] \le u(n,C), \quad C = \max\{Q, v(2)\}, 1 < n \le \bar{s}, \tag{4.6}$$

and

$$E\left[\left\|x_n - x^*\right\|^2\right] \le v(s), n \ge \bar{s}, \tag{4.7}$$

where $\bar{s}$ is a unique root of the equation $u(s,C) = v(s)$ on the interval $[2,\infty)$.

II. Suppose that $\frac{1}{2q_1} < b < \frac{1}{q_1}$ and $c_0 \ge \frac{bq_1}{1 - bq_1} > 1$. Then $\lim_{n\to\infty} E\left[\left\|x_n - x^*\right\|^2\right] = 0$ and the estimate (4.6) holds for all $n \ge 1$.

Corollary 4.4. *Assume that* $T : G \to H$ *is a strongly contractive mapping, that is, in* (1.4) *is satisfied with* $0 < q < 1$. *Let in* (1.5) $\alpha_n = \frac{b}{n^\kappa}, b > 0, 0 < \kappa < 1$. *Then*

$$F(t) = \frac{b}{1-\kappa} t^{1-\kappa}, \Phi(\tau) = \frac{\ln \tau}{q_1},$$

$$\Phi^{-1}(z) = \left(\frac{1-\kappa}{b} z\right)^{\frac{1}{1-\kappa}}, q_1 = 1 - q,$$

$$v(s) = \left(\frac{2c_0C_1}{b} + 2C_1\right)\frac{1}{s-1},$$

$$u(s,C) = C \exp\left\{-\frac{abq_1}{1-\kappa}\left(s^{1-\kappa} - 2^{1-\kappa}\right)\right\},$$

$$Q = \left\|x_1 - x^*\right\|^2 \exp\left\{-\frac{abq_1}{1-\kappa}\left(2^{1-\kappa} - 1\right)\right\}.$$

Suppose that $c_0 \ge \frac{bq}{bq - 1}$. Then $\lim_{n\to\infty} E\left[\left\|x_n - x^*\right\|^2\right] = 0$ and

1) If $Q \le v(2)$ and $c_0 \ge \frac{bq_1}{bq_1 - 2}$, we have for all $n > 1$

$$E\left[\left\|x_n - x^*\right\|^2\right] \le v(n); \tag{4.8}$$

2) In all the remain cases the estimates (4.6) and (4.7) hold.

Corollary 4.5. *Assume that* $T : G \to H$ *is a weakly contractive mapping of the class* $C_{\phi(t)} = t^{2\nu-1}, \nu > 1$, *that is, in Theorem* 3.1 $\psi(\tau) = \tau^\nu$. *Let in* (1.5) $\alpha_n = \frac{b}{n}, b > 0$. *Then*

$$F(t) = b \ln t, \Phi(\tau) = \frac{\tau^{1-\nu}}{1-\nu}, \Phi^{-1}(z) = \left[(1-\nu)z\right]^{\frac{1}{1-\nu}},$$

$$u(s,C) = C\left[1 + (\nu-1)C^{\nu-1}ab\ln\frac{s}{2}\right]^{-\frac{1}{\nu-1}},$$

$$Q = \left\|x_1 - x^*\right\|^2\left[1 + (\nu-1)\left\|x_1 - x^*\right\|^{2(\nu-1)}ab\ln 2\right]^{-\frac{1}{\nu-1}}.$$

If $c_0 > 1$ is chosen from the condition

$$\left(\frac{2c_0C_1}{b} + 2C_1\right)^{\nu-1} ab \le \nu^{-1},$$

then $\lim_{n\to\infty} E\left[\left\|x_n - x^*\right\|^2\right] = 0$ and for all $n > 1$

$$E\left[\left\|x_n - x^*\right\|^2\right] \le u(n,C), \quad C = \max\left\{Q, \left(\frac{2C_1c_0}{b}\right)^{\frac{1}{\nu}} + 2C_1\right\}. \tag{4.9}$$

Corollary 4.6. *Assume that* $T : G \to H$ *is a weakly contractive mapping of the class* $C_{\phi(t)} = t^{2\nu-1}, \nu > 1$, *that is, in Theorem* 3.1 $\psi(\tau) = \tau^\nu$. *Let in* (1.5) $\alpha_n = \frac{b}{n^\kappa}, b > 0, 0 < \kappa < 1$. *Then*

$$F(t) = \frac{b}{1-\kappa}t^{1-\kappa}, \Phi(\tau) = \frac{\tau^{1-\nu}}{1-\nu}, \Phi^{-1}(z) = \left[(1-\nu)z\right]^{\frac{1}{1-\nu}},$$

$$v(s) = \left[\left(\frac{2c_0C_1}{b}\right)^{\frac{1}{\nu}} + 2C_1\right]\left(\frac{1}{s-1}\right)^{\frac{\kappa}{\nu}},$$

$$u(s,C) = C\left[1 + \frac{\nu-1}{1-\kappa}C^{\nu-1}ab\left(s^{1-\kappa} - 2^{1-\kappa}\right)\right]^{-\frac{1}{\nu-1}},$$

$$Q = \left\|x_1 - x^*\right\|^2\left[1 + \frac{\nu-1}{1-\kappa}\left\|x_1 - x^*\right\|^{2(\nu-1)}ab\left(2^{1-\kappa} - 1\right)\right]^{-\frac{1}{\nu-1}}.$$

I. Suppose that

$$\frac{\kappa}{\nu} < \frac{1-\kappa}{\nu-1}.$$

1) If $Q \le \nu(2)$ and $c_0 > 1$ is chosen from the condition

$$\left[\left(\frac{2c_0C_1}{b}\right)^{\frac{1}{\nu}} + 2C_1\right]^{\nu-1} ab \ge \frac{2^{\kappa}\kappa}{\nu},$$

then $\lim_{n\to\infty} E\left[\left\|x_n - x^*\right\|^2\right] = 0$ and for all $n > 1$

$$E\left[\left\|x_n - x^*\right\|^2\right] \le \nu(n). \tag{4.10}$$

2) In all the remain cases the estimates (4.6) and (4.7) hold.

II. Suppose that

$$\frac{\kappa}{\nu} \ge \frac{1-\kappa}{\nu-1}.$$

If $c_0 > 1$ is chosen from the condition

$$\left[\left(\frac{2c_0C_1}{b}\right)^{\frac{1}{\nu}} + 2C_1\right]^{\nu-1} ab \le \frac{\kappa}{\nu},$$

then $\lim_{n\to\infty} E\left[\left\|x_n - x^*\right\|^2\right] = 0$ and for all $n \ge 1$

$$E\left[\left\|x_n - x^*\right\|^2\right] \le u(n, C), C = \max\{Q, \nu(2)\}. \tag{4.11}$$

In addition to the examples presented in this section, we produce the functions $\psi(\tau)$ and $\Phi(\lambda)$ which have $\tau = 0$ as a tangency point of the infinite degree multiplicity and given logarithmic estimates of the convergence rate.

We define the function $\psi(\tau)$ by the following way:

$$\psi(\tau) = -\exp(-f(\tau))\left[f'(\tau)\right]^{-1}, \tau > 0,$$

where $f(\tau)$ is differentiable and decreasing function, $\lim_{\tau\to 0} f(\tau) = \infty$, and

$$\psi^{(l)}(0) = \lim_{\tau\to 0+} \psi^{(l)}(\tau), l = 0,1,2,\cdots,$$

where (l) denote the derivative degrees of the function $\psi(\tau), \psi^{(0)}(\tau) = \psi(\tau)$. It is easy to see that

$$\Phi(\tau) = -\exp f(\tau)$$

and

$$\Phi^{-1}(z) = f^{-1}(\ln(-z)).$$

In particular,

i) $f(\tau) = \frac{1}{\tau}, \psi(\tau) = \tau^2 \exp\left(-\frac{1}{\tau}\right)$. We have $\Phi(\tau) = -\exp\frac{1}{\tau}$ and $\Phi^{-1}(z) = \frac{1}{\ln(-z)}$. We have to verify that $\psi(\tau)$ is convex. In fact, it is true because

$$\frac{d^2\psi(\tau)}{d\tau} = \left(2 + \frac{2\tau+1}{\tau^2}\right)\exp\left(-\frac{1}{\tau}\right) > 0, \forall \tau > 0.$$

Beside this, it is easy to see that $\psi(\tau) \le \tau$, at least, on the interval $[0,1]$. In the next examples we leave to readers to check these properties.

ii) $f(\tau) = \frac{1}{\tau^s},\ s \ge 1,\ \psi(\tau) = \exp\left(-\frac{1}{\tau^s}\right)\tau^{s+1}$.

We have $\Phi(\tau) = -\exp\frac{1}{\tau^s}$ and $\Phi^{-1}(z) = (\ln(-z))^{-\frac{1}{s}}$.

iii) $f(\tau) = \exp\frac{1}{\tau^s}, \psi(\tau) = \exp\left(-\left[\frac{1}{\tau^s} + \exp\left(\frac{1}{\tau^s}\right)\right]\right)\tau^{s+1}$

We have

$$\Phi(\tau) = -\exp\left(\exp\frac{1}{\tau^s}\right) \text{ and } \Phi^{-1}(z) = (\ln\ln(-z))^{-\frac{1}{s}}.$$

5. Almost Sure Convergence of Stochastic Approximations for Nonexpansive Mappings

Consider next the almost surely convergence of stochastic approximations. First of all, we need the stochastic analogy of Lemma 2.5:

Lemma 5.1. *Let* $\{\alpha_k\}$ *be sequences of non-negative real numbers and* $\{\beta_k\}$ *be sequence of random* A_k*-measurable variables, a.s. nonnegative for all* $k \ge 1$. *Assume that*

$$\sum_{k=1}^{\infty} \alpha_k = \infty \text{ and } \sum_{k=1}^{\infty} \alpha_k\beta_k < \infty.$$

If $\lim_{k\to\infty} \alpha_k = 0$ and there exists $c > 0$ such that for all $k \ge 1$

$$\left|\beta_{k+1} - E\left[\beta_k / A_n\right]\right| \le c\alpha_k \text{ a.s.}, \tag{5.1}$$

then $\lim_{k\to\infty} \beta_k = 0$ a.s.

The proof can be provided by the scheme of nonstochastic case (see Proposition 2 in [8]) or as it was done in [5].

We need also the following lemma from [14] as applied to our case of Hilbert spaces (the concepts of modulus of convexity $\delta_B(\varepsilon)$ of Banach spaces B or Hilbert spaces H can be found in [15] and [16]).

Lemma 5.2. *If* $F = I - T$ *with a nonexpansive mapping* $T: G \to H$, *then for all* $x, y \in D(T)$,

$$(Fx - Fy, x - y) \ge R_1^2 \delta_H\left(\frac{\|Fx - Fy\|}{2R_1}\right),$$

where

$$R_1 = \sqrt{2^{-1}\left(\|x-y\|^2 + \|Tx-Ty\|^2\right)} \le \|x-y\|.$$

If $\|x\| \le R$ and $\|y\| \le R$ with $x, y \in D(T)$, then $R_1 \le 2R$ and

$$(Fx - Fy, x - y) \ge L^{-1}R^2\delta_H\left(\frac{\|Fx-Fy\|}{4R}\right).$$

Theorem 5.3. *Assume that a mapping* $T: G \to H$ *is nonexpansive and its fixed point set* N *is nonempty. If* (1.8) *holds and* $E\left[\|\xi_n\|/A_n\right] \le C_0$, *then the sequence* $\{x_n\}$ *generated by* (1.5)-(1.7) *weakly almost surely converges to some* $\tilde{x} \in N$.

Proof. Let $x^* \in N$. We next use Lemma 5.2 and the estimate (see [17], p. 49)

$$\delta_H(\varepsilon) \ge \frac{\varepsilon^2}{8}$$

to get

$$\left(x - Tx, x - x^*\right) \ge R_1^2\delta_H\left(\frac{\|x-Tx\|}{2R_1}\right) \ge \frac{\|x-Tx\|^2}{32}.$$

In this case the inequality (3.3) implies

$$\begin{aligned}&\left\|x_{n+1} - x^*\right\|^2 \\ &\le \left(1+8\alpha_n^2\right)\left\|x_n - x^*\right\|^2 - (32)^{-1}\|\alpha_n x_n - Tx_n\|^2 \quad (5.2)\\ &-2\alpha_n\left(\xi_n, x_n - x^*\right) + 2\alpha_n^2\|\xi_n\|^2.\end{aligned}$$

Similarly to (3.10), we have

$$\begin{aligned}&E\left[\left\|x_{n+1} - x^*\right\|^2 / A_n\right] \\ &\le \left(1+8\alpha_n^2\right)\left\|x_n - x^*\right\|^2 - 2\alpha_n E\left[\|x_n - Tx_n\|^2 / A_n\right] \quad (5.3)\\ &- 2\alpha_n E\left[\left(\xi_n, x_n - x^*\right)/A_n\right] + 2\alpha_n^2 E\left[\|\xi_n\|^2/A_n\right].\end{aligned}$$

Denote $\lambda_n = E\left[\left\|x_n - x^*\right\|^2\right]$ and $\beta_n = E\left[\|x_n - Tx_n\|^2\right]$ and apply the unconditional expectation to both sides of (5.3). Then

$$\lambda_{n+1} \le \left(1+8\alpha_n^2\right)\lambda_n - 2\alpha_n\beta_n + 2C_1\alpha_n^2. \quad (5.4)$$

It follows from this that

$$\lambda_{n+1} \le \left(1+8\alpha_n^2\right)\lambda_n + 2C_1\alpha_n^2.$$

Since $\sum_1^\infty \alpha_n^2 < \infty$ and due to Lemma 2.1, we conclude that $\{\lambda_n\}$ is bounded. Consequently, $\{x_n\}$ is bounded a.s. that follows from the theory of convergent quasimartingales (see [5,18]).

We now need Lemma 5.1. It is not difficult to see that

$$\sum_1^\infty \alpha_n\beta_n = \sum_1^\infty \alpha_n E\left[\|x_n - Tx_n\|^2\right] < \infty. \quad (5.5)$$

The last gives us

$$\sum_1^\infty \alpha_n \|x_n - Tx_n\|^2 < \infty \text{ a.s.}$$

Next we evaluate the following difference:

$$\begin{aligned}&\|x_{n+1} - Tx_{n+1}\|^2 - \|x_n - Tx_n\|^2 \\ &\le \left(\|x_{n+1} - Tx_{n+1}\| + \|x_n - Tx_n\|\right) \\ &\times\left(\|x_{n+1} - Tx_{n+1}\| - \|x_n - Tx_n\|\right)\end{aligned}$$

It is easy to see that $\|x_n - Tx_n\|$ is bounded a.s. Indeed, since $\|x_n\| \le C_1$ a.s., there exists a constant $C_2 > 0$ such that

$$\|x_n - Tx_n\| = 2\left\|x_n - x^*\right\| \le C_2 \text{ a.s.}$$

Therefore

$$\|x_{n+1} - Tx_{n+1}\| + \|x_n - Tx_n\| \le 2C_2 \text{ a.s.}$$

It is obviously that

$$\begin{aligned}&\|x_{n+1} - Tx_{n+1}\| - \|x_n - Tx_n\| \\ &\le \left\|\left(x_{n+1} - x_n\right) + \left(Tx_n - Tx_{n+1}\right)\right\| \le 2\|x_{n+1} - x_n\| \\ &= 2\left\|\Pr_G x_n - \alpha_n\left(x_n - Tx_n + \xi_n\right) - \Pr_G x_n\right\| \\ &\le 2\alpha_n\|x_n - Tx_n\| + 2\alpha_n\|\xi_n\|.\end{aligned}$$

Thus,

$$\begin{aligned}&E\left[\|x_{n+1} - Tx_{n+1}\|^2 / A_n\right] - \|x_n - Tx_n\|^2 \\ &\le 4\alpha_n C_2^2 + 4\alpha_n C_2 E\left[\|\xi_n\|/A_n\right] \\ &\le 4C_2\left(C_2 + C_0\right)\alpha_n \text{ a.s..}\end{aligned}$$

By Lemma 5.1, $\|x_n - Tx_n\| \to 0$ a.s. as $n \to \infty$.

Since $\{x_n\}$ is bounded a.s., there is a subsequence $\{x_{n_k}\}$ weakly convergent to some point $\tilde{x}$. Since G is convex and closed, consequently, weakly closed, we assert that $\tilde{x} \in G$. It is known that a nonexpansive mapping T is weakly demiclosed, therefore, $\tilde{x} \in N$ a.s. Weak almost surely convergence of whole sequence $\{x_n\}$ is shown by the standard way [8]. □

Corollary 5.4. *Assume that* $T: G \to H$ *is a weakly contractive mapping of the class* $C_{\phi(t)}$. *If* $\sum_1^\infty \alpha_n = \infty$, $\sum_1^\infty \alpha_n^2 < \infty$ *and* $E\left[\|\xi_n\|/A_n\right] \le C_0$, *then the sequence* $\{x_n\}$ *generated by* (1.5)-(1.7) *strongly almost surely converges to unique fixed point* x^* *of* T.

Proof. We have from (3.4)

$$\|x_n - Tx_n\|\left\|x_n - x^*\right\| \ge \psi\left(\left\|x_n - x^*\right\|^2\right). \quad (5.6)$$

Since $\{x_n\}$ is bounded a.s. and $\|x_n - Tx_n\| \to 0$ a.s. as $n \to \infty$, we conclude that $\psi\left(\|x_n - x^*\|^2\right) \to 0$ a.s. The proof follows due to the properties of the function ψ. □

Remark 5.5. *It is clear that all the results remain still valid for self-mappings* $T: G \to G$. *However, in this case, unlike any deterministic situation, the algorithm* (1.5)-(1.7) *must use the projection operator* Pr_G *because the vector* $v_n = x_n - \alpha_n S_n x_n$ *not always belongs to* G. *If* $T_n x_n = Tx_n + \xi_n \in G$ *for all* $n \ge 1$ *and* $0 < \alpha_n < 1$, *then* (1.5) *can be replaced by*

$$x_{n+1} = (1-\alpha_n) x_n + \alpha_n T_n x_n, n = 1, 2, \cdots, x_1 \in G.$$

REFERENCES

[1] M. T. Wasan, "Stochastic Approximation," Cambridge University Press, Cambridge, 1969.

[2] M. B. Nevelson and R. Z. Hasminsky, "Stochastic Approximation and Recursive Estimation," AMS Providence, Rhode Island, 1973.

[3] Ya. Alber and S. Shilman, "Nonasymptotic Estimates of the Convergence Rate of Stochastic Iterative Algorithms," *Automation and Remote Control*, Vol. 42, 1981, pp. 32-41.

[4] Ya. Alber and S. Shilman, "General Nonasymptotic Estimates of the Convergence Rate of Iterative Stochastic Algorithms," *USSR Computational Mathematics and Mathematical Physics*, Vol. 25, No. 2, 1985, pp. 13-20. doi:10.1016/0041-5553(85)90099-0

[5] K. Barty, J.-S. Roy and C. Strugarek, "Hilbert-Valued Perturbed Subgradient Algorithms," *Mathematics of Operation Research*, Vol. 32, No. 3, 2007, pp. 551-562. doi:10.1287/moor.1070.0253

[6] X. Chen and H. White, "Asymptotic Properties of Some Projection-Based Robbins-Monro Procedures in a Hilbert Space," *Studies in Nonlinear Dynamics & Econometrics*, Vol. 6, No. 1, 2002, pp. 1-53. doi:10.2202/1558-3708.1000

[7] Ya. I. Alber and A. N. Iusem, "Extension of Subgradient Techniques for Nonsmooth Optimization in Banach Spaces," *Set-Valued Analysis*, Vol. 9, No. 4, 2001, pp. 315-335. doi:10.1023/A:1012665832688

[8] Ya. Alber, A. Iusem and M. Solodov, "On the Projection Subgradient Method for Nonsmooth Convex Optimization in a Hilbert Space," *Mathematical Programming*, Vol. 81, No. 1, 1998, pp. 23-35. doi:10.1007/BF01584842

[9] A. P. Dempster, N. M. Laird and D. B. Rubin, "Maximal Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society*, Vol. 39, No. 1, 1977, pp. 185-197.

[10] W. Rudin, "Real and Complex Analysis," McGraw-Hill, New York, 1978.

[11] Ya. I. Alber and S. Guerre-Delabriere, "Principle of Weakly Contractive Maps in Hilbert Spaces," *Operator Theory, Advances and Applications*, Vol. 98, 1997, pp. 7-22.

[12] Ya. I. Alber, "Recurrence Relations and Variational Inequalities," *Soviet Mathematics Doklady*, Vol. 27, 1983, pp. 511-517.

[13] Ya. Alber, S. Guerre-Delabriere and L. Zelenko, "The Principle of Weakly Contractive Maps in Metric Spaces," *Communications on Applied Nonlinear Analysis*, Vol. 5, No. 1, 1998, pp. 45-68.

[14] Ya. I. Alber, "New Results in Fixed Point Theory," Preprint, Haifa Technion, 2000.

[15] J. Diestel, "The Geometry of Banach Spaces," *Lecture Notes in Mathematics*, No. 485, Springer, Berlin, 1975.

[16] T. Figiel, "On the Moduli of Convexity and Smoothness," *Studia Mathematica*, Vol. 56, No. 2, 1976, pp. 121-155.

[17] Ya. Alber and I. Ryazantseva, "Nonlinear Ill-Posed Problems of Monotone Type," Springer, Berlin, 2006.

[18] M. Métivier, "Semimartingales," De Gruyter, Berlin, 1982. doi:10.1515/9783110845563

