Convergence of random products of contractions in Hilbert space

By I. AMEMIYA and T. ANDO in Sapporo (Japan)

1. Introduction

Given two projections P_1 and P_2 in a Hilbert space, it is known that a product $T_n \cdots T_2 T_1$ converges strongly as $n \to \infty$ where $T_j = P_1$ or $T_j = P_2$ at random. The problem in this paper is to observe the case of a finite number of projections. The result is that weak convergence is always valid, while strong convergence is still unsettled. After several comments on the convergence of the iterates of a single contraction, the convergence problem of random products will be discussed for a wider class of contractions, including all non-negative definite contractions.

2. Iterates

Let T be a contraction in a Hilbert space, i. e., a linear operator with $\|T\| \leqq 1$, then the so-called mean ergodic theorem shows that the average $\frac{1}{n} \sum_{j=1}^{n} T^j$ converges strongly, as $n \to \infty$, to the projection onto the subspace of all vectors invariant under T, i. e., the null space of $I-T$, and that the orthogonal complement of the null space of $I-T$ coincides with the closure of its range (see [6], n° 143).

When do the iterates T^n themselves converge strongly or weakly? Since T operates as the identity on the null space of $I-T$, T^n converges if and only if $T^n f$ converges to 0 for all f in the range of $I-T$, so that T^n converges strongly or weakly according as $T^n(I-T)$ converges to 0 strongly or weakly.

Given f, $\|T^n f\|$ decreases monotonically with limit, say $\alpha \geqq 0$. If $\alpha = 0$, clearly $T^n(I-T)f \to 0$ strongly, and if $\alpha > 0$, with $g_n = \frac{T^n f}{\|T^n f\|}$, it follows $\|g_n\| = 1$, $T^n(I-T)f = \|T^n f\|(I-T)g_n$, and $\|Tg_n\| = \frac{\|T^{n+1}f\|}{\|T^n f\|} \to 1$. This observation leads to the following criterion.

T^n converges strongly or weakly, if T has the following property (S) *or* (W), *respectively:*

(S) $\|f_n\| \leqq 1$, $\|Tf_n\| \to 1$ *imply* $(I-T)f_n \to 0$ *strongly,*

(W) $\|f_n\| \leqq 1$, $\|Tf_n\| \to 1$ *imply* $(I-T)f_n \to 0$ *weakly.*

A non-negative definite contraction, in particular, a projection, has the property (S); in fact, if T is a non-negative definite contraction,

$$\|(I-T)f_n\|^2 = ((I-T)^2 f_n, f_n) \leqq$$
$$\leqq ((I-T)(I+T)f_n, f_n) = \|f_n\|^2 - \|Tf_n\|^2 \to 0$$

whenever $\|f_n\| \leqq 1$ and $\|Tf_n\| \to 1$.

The product of two (hence a finite number of) contractions, each of which has (S) or (W), also has the same property; in fact, if T_1, T_2 are contractions with the property, say (S), and if $\|f_n\| \leqq 1$ and $\|T_2 T_1 f_n\| \to 1$, then $\|T_1 f_n\| \leqq 1$ and $\|T_1 f_n\| \to 1$, so that

$$(I-T_2T_1)f_n = (I-T_1)f_n + (I-T_2)T_1 f_n \to 0$$

strongly. It should be mentioned that the statement about (S) was observed by HALPERIN [3] in proving the strong convergence of the iterates of a product of a finite number of projections.

The condition (W) has a simpler equivalent form (W′):

(W′) $\|Tf\| = \|f\|$ *implies* $Tf=f$.

Only the implication (W′)⇒(W) needs a proof. Since

$$\|f\|^2 - \|Tf\|^2 = ((I-T^*T)f, f)$$

and $I-T^*T$ is non-negative definite, $\|Tf\| = \|f\|$ is equivalent to $(I-T^*T)f=0$, so that (W′) implies that the null space of $I-T^*T$ is contained in that of $I-T$. By taking the orthogonal complements, it follows that the closure of the range of $I-T^*T$ contains the range of $I-T^*$. Now if $\|f_n\| \leqq 1$ and $\|Tf_n\| \to 1$, then

$$1 \geqq \|T^*Tf_n\| \geqq \|f_n\| \cdot \|T^*Tf_n\| \geqq (T^*Tf_n, f_n) = \|Tf_n\|^2 \to 1,$$

so that the property (S) for the non-negative definite contraction T^*T shows $(I-T^*T)f_n \to 0$ strongly, which, in turn, implies $(f_n, h) \to 0$ for all h in the closure of the range of $I-T^*T$. For an arbitrary g,

$$((I-T)f_n, g) = (f_n, (I-T^*)g) \to 0,$$

because $(I-T^*)g$ is in the closure of the range of $I-T^*T$. Thus (W′) implies (W).

Clearly (S) implies (W) and equivalently (W′). If a contraction T has (W′), its adjoint T^* has (W′) too. In fact, $\|T^*f\| = \|f\|$ implies $TT^*f=f$, so that

$$\|TT^*f\| = \|f\| = \|T^*f\|,$$

hence $TT^*f = T^*f$ by (W′) of T, and the assertion follows.

A contraction T is called *completely non-unitary* if $\|T^n f\| = \|T^{*n} f\| = \|f\|$ for all $n \geqq 0$ implies $f=0$. The decomposition theorem, proved idenpendently in [4] and [7], is quite useful in analysing an arbitrary contraction; it says that for a contraction T there is a uniquely determined closed linear subspace such that it reduces T and that T is unitary on it and is completely non-unitary on its orthogonal complement. Indeed, the subspace consists of all vectors f for which $\|T^n f\| = \|T^{*n} f\| = \|f\|$ for all $n \geqq 1$. Moreover SZ.-NAGY and FOIAŞ [7] proved that the spectral measure of the minimum unitary dilation of a completely non-unitary contraction is absolutely continuous with respect to the Lebesgue measure on the unit circle. This result can give the following improvement of the criterion (W′) for the weak convergence of the iterates:

T^n converges weakly, if $\|T^n f\| = \|T^{*n} f\| = \|f\|$ for all $n \geqq 1$ implies $Tf = f$.

Here is an alternative proof, not using spectral representation (cf. [2]). The hypothesis means that the unitary part of T in the decomposition mentioned above acts as the identity, so that there is no loss of generality in assuming the complete non-unitarity of T, which is, as in the proof of the implication (W′)⇒(W), equivalent to the statement that the intersection of all null spaces of A_n ($n = \pm 1, \pm 2, \dots$) consists of 0 only, where $A_n = I - T^{*n}T^n$ and $A_{-n} = I - T^n T^{*n}$ for $n \geqq 0$. By taking orthogonal complement, the linear span of the union of all the ranges of A_n ($n = \pm 1, \pm 2, \dots$) is dense, so that to prove the weak convergence, it suffices to show that for all non-zero integer n and vectors f, g $(T^j f, A_n g) \to 0$ as $j \to \infty$. Since $\|T^j f\|$ decreases as j increases, it results for $n \geqq 1$

$$(T^j f, A_n T^j f) = \|T^j f\|^2 - \|T^{n+j} f\|^2 \to 0 \quad \text{as} \quad j \to \infty,$$

so that the generalized Schwarz's inequality ([6], n° 104) for the scalar product induced by A_n yields

$$|(T^j f, A_n g)|^2 \leqq (T^j f, A_n T^j f) \cdot (g, A_n g) \leqq (T^j f, A_n T^j f) \cdot \|g\|^2 \to 0.$$

Since

$$A_{-n} T^j = T^n A_n T^{j-n} \quad \text{for} \quad j \geqq n \geqq 1,$$

the generalized Schwarz's inequality for the scalar product induced by A_{-n} yields

$$|(T^j f, A_{-n} g)|^4 \leqq (T^j f, A_{-n} T^j f)^2 \cdot (g, A_{-n} g)^2 \leqq$$

$$\leqq (T^j f, T^n A_n T^{j-n} f)^2 \cdot \|g\|^4 = (T^{*n} T^j f, A_n T^{j-n} f)^2 \|g\|^4 \leqq$$

$$\leqq (T^{j-n} f, A_n T^{j-n} f) \cdot (T^{*n} T^j f, A_n T^{*n} T^j f) \cdot \|g\|^4 \leqq$$

$$\leqq (T^{j-n} f, A_n T^{j-n} f) \cdot \|f\|^2 \cdot \|g\|^4 \to 0.$$

3. Random products

Let T_j ($j = 1, 2, \dots, N$) be a finite set of contractions. A mapping $r(\cdot)$ from the set of all positive integers to $\{1, \dots, N\}$ will be called a *(random) selection.* Given a random selection $r(\cdot)$, construct the sequence of contractions $\{S_n\}$ by setting $S_n = T_{r(n)} \cdots T_{r(2)} \cdot T_{r(1)}$, then what can be said about the convergence of S_n or of the average $\frac{1}{n} \sum_{j=1}^{n} S_j$? The random ergodic theorem (cf. [1]) shows that if each selection is considered as a point of the infinite product of the copies of the probability space $\{1, 2, \dots, N\}$ (on which each point has the same probability N^{-1}), then the average $\frac{1}{n} \sum_{j=1}^{n} S_j$ converges strongly for almost all selections. Without any further restriction on the T_j's this would be the best result.

Suppose now that each T_j ($j = 1, 2, \dots, N$) has (S). If a selection $r(\cdot)$ is *periodic*, i. e., $r(k+m) = r(k)$ for some m and all k, then $S_{mk} = (S_m)^k$, and since S_m has (S),

S_{mk} converges strongly, as $k \to \infty$, to the projection onto the null space of $I - S_m$. For an index n, take k such that $m(k-1) < n \leqq mk$, then

$$S_n = S_{n-m(k-1)} S_{m(k-1)}$$

where it is assumed that $S_0 = I$. If f is in the null space of $I - S_m$,

$$\|f\| = \|S_{m(k-1)}f\| \geqq \|S_n f\| \geqq \|S_{mk} f\| = \|f\|,$$

so that $S_n f = f$ because of (W′) for S_n, and if f is in the closure of the range of $I - S_m$,

$$\|S_n f\| \leqq \|S_{m(k-1)} f\| \to 0 \quad \text{as} \quad n \to \infty,$$

because $k \to \infty$ as $n \to \infty$. Thus S_n converges strongly.

When every T_j is a projection, PRAGER [5] derived the weak convergence of S_n from a *quasi-periodicity* assumption on the selection $r(\cdot)$; S_n converges weakly, if there is m such that every m consecutive $r(k), r(k+1), \ldots, r(k+m-1)$ contains each j at least once $(j = 1, 2, \ldots, N)$. On starting with an observation that in his proof the hypothesis that T_j is a projection is not essential, but only the property (W) is necessary, the goal of this paper is to derive the weak convergence from (W) without any periodicity assumption on a selection. It should be mentioned that when the Hilbert space is of finite dimension, PRAGER attained the same goal in the case of projections; he proved the strong convergence, but strong convergence is equivalent to the weak one in the finite dimensional case.

Theorem. *If T_j is a contraction with* (W) *or equivalently* (W′) $(j = 1, 2, \ldots, N)$, *then for any random selection $r(\cdot)$ the sequence*

$$S_n = T_{r(n)} \cdots T_{r(2)} T_{r(1)}$$

converges weakly as $n \to \infty$.

The proof will be divided into several steps.

(i) In what follows, by a weak neighborhood there is meant a convex symmetric neighborhood of 0 with respect to the weak topology. The condition (W) can be stated in the following form: for any weak neighborhood $\mathfrak{V}$ there is an $\varepsilon > 0$ such that $\|f\| \leqq 1$, $\|T_j f\| \geqq 1 - \varepsilon$ imply $(I - T_j) f \in \mathfrak{V}$.

(ii) The intersection of the null spaces of $I - T_k$ $(k = 1, 2, \ldots, j)$ coincides with the null space of $I - T_j \ldots T_2 T_1$. In fact, the former is obviously contained in the latter. If f is in the latter,

$$\|f\| = \|T_n \cdots T_2 T_1 f\| \leqq \|T_1 f\| \leqq \|f\|$$

so that $T_1 f = f$ by (W′) of T_1 and by induction $T_k f = f$ $(k = 1, 2, \ldots, j)$. Let Q_j stand for the projection onto the null space of $I - T_j \ldots T_2 T_1$, then $T_k Q_j = Q_j$ $(k = 1, 2, \ldots, j)$ so that $T_k^* Q_j = Q_j$ because a vector invariant under a contraction is also invariant under its adjoint (cf. [6], nº 143). Thus from $T_k Q_j = Q_j$ and $T_k^* Q_j = Q_j$, the commutativity of Q_j with T_k $(k = 1, 2, \ldots, j)$ follows.

(iii) Let $P_j = I - Q_j$, then for any weak neighborhood $\mathfrak{V}$, there is another $\mathfrak{W}$ such that $\|f\| \leqq 1$, $(I - T_k) f \in \mathfrak{W}$ $(k = 1, 2, \ldots, j)$ imply $P_j f \in \mathfrak{V}$. In fact, since $(I - T_k) f = 0$ $(k = 1, 2, \ldots, j)$ is equivalent to $P_j f = 0$ by (ii), the mapping which assigns to $P_j f$ the ordered j-tuple $\{(I - T_1) f, \ldots, (I - T_j) f\}$ is one-to-one. Since $(I - T_k) f = (I - T_k) P_j f$ $(k = 1, 2, \ldots, j)$ by (ii), it is continuous from the image of the unit ball under P_j into the product of j copies of the Hilbert space, when they are provided with their respective weak topologies. Since the domain is compact, the mapping is bi-continuous and the assertion is just the statement that the inverse mapping is continuous at the origin.

(iv) Let $\mathfrak{M}_j$ be the collection of contractions which are in a multiplicative semi-group with unit generated by j of the contractions $\{T_k\}_1^N$ $(j=1, 2, \ldots, N)$ and let $\mathfrak{M}_0=\{I\}$. Given a weak neighborhood $\mathfrak{V}$ and $S\in\mathfrak{M}_j$, there is a positive number $\varepsilon=\varepsilon(\mathfrak{V}, j)$ depending only on $\mathfrak{V}$ and j such that $\|f\|\leqq 1$, $\|Sf\|\geqq 1-\varepsilon$ implies $(I-S)f\in\mathfrak{V}$. Proof proceeds by induction on j as follows. The assertion for $j=0$ is trivial. Suppose that the assertion is true up to $j-1$. Only S in $\mathfrak{M}_j$ but not in $\mathfrak{M}_{j-1}$ needs consideration. There is no loss of generality in assuming that S is in the multiplicative semi-group generated by $T_1, T_2, \ldots, T_j$. For any index $1\leqq k\leqq j$, S can be written in the form

$$S=R_1T_kR_2=R_3T_kR_4$$

where R_1, R_4 are in $\mathfrak{M}_{j-1}$. Given a weak neighborhood $\mathfrak{V}$, take $\mathfrak{W}$ which is in relation of (iii) to $\mathfrak{V}$, and choose a weak neighborhood $\mathfrak{U}$ such that

$$4\mathfrak{U}+4T_i\mathfrak{U}\subseteqq\mathfrak{W}\qquad(i=1, \ldots, j),$$

which is possible because of the weak continuity of T_i. Now by the inductive assumption and (i) it is possible to take a positive number ε, independent of S, such that $\|g\|\leqq 1$, $\|Rg\|\geqq 1-\varepsilon$ with $R\in\mathfrak{M}_{j-1}$ or $R=T_k$ imply $(I-R)g\in\mathfrak{U}$. Now if $\|f\|\leqq 1$ and $\|Sf\|\geqq 1-\varepsilon$, obviously $1\geqq\|R_4f\|\geqq 1-\varepsilon$ and $\|T_kR_4f\|\geqq 1-\varepsilon$, hence $(I-R_4)f\in\mathfrak{U}$ and $(I-T_k)R_4f\in\mathfrak{U}$, so that

$$(I-T_k)f=(I-R_4)f+(I-T_k)R_4f-T_k(I-R_4)f\in 2\mathfrak{U}+T_k\mathfrak{U}\subseteqq\tfrac{1}{2}\mathfrak{W}$$

and quite similarly

$$(I-T_k)Sf=(R_1-I)T_kR_2f+T_k(I-R_1)T_kR_2f+T_k(I-T_k)R_2f\in\mathfrak{U}+2T_k\mathfrak{U}\subseteqq\tfrac{1}{2}\mathfrak{W}.$$

Since the relation is valid for $k=1, 2, \ldots, j$, (iii) guarantees $P_jf\in\frac{1}{2}\mathfrak{V}$ and $P_jSf\in\frac{1}{2}\mathfrak{V}$, consequently $P_j(I-S)f\in\mathfrak{V}$. As in (ii), $I-P_j$ is just the projection onto the null space of $I-S$, because S has T_k as a factor $(k=1, 2, \ldots, j)$, so that

$$(I-S)f=(I-S)P_jf=P_j(I-S)f\in\mathfrak{V}.$$

(v) S_nf converges weakly for all f. In fact, if $\|S_nf\|\to 0$, the assertion is trivial. If $\inf_n\|S_nf\|>0$, given a weak neighborhood $\mathfrak{V}$, take $\varepsilon=\varepsilon(\mathfrak{V}, N)$ in (iv), then for sufficiently large $n\geqq m$ we have $\|S_nf\|\geqq(1-\varepsilon)\|S_mf\|$. Since $S_n=S\cdot S_m$ for some $S\in\mathfrak{M}_N$ and

$$\|Sg\|=\frac{\|S_nf\|}{\|S_mf\|}\geqq 1-\varepsilon\quad\text{with}\quad g=\frac{S_mf}{\|S_mf\|},$$

(iv) guarantees $(I-S)g\in\mathfrak{V}$, so that

$$S_mf-S_nf=\|S_mf\|(I-S)g\in\|f\|\mathfrak{V}.$$

The weak convergence follows from the arbitrariness of $\mathfrak{V}$. This completes the proof.

*

When every index j appears infinitely many times in a selection $r(\cdot)$, the limit of the sequence S_n is the projection onto the subspace of vectors invariant under all T_j $(j=1, 2, \ldots, N)$. In fact, with the notations in the proof of the Theorem, for sufficiently large m $S_m Q_N = Q_N$, and if $f = P_N f$ and $\inf_n \|S_n f\| > 0$, for any weak neighborhood $\mathfrak{V}$ and sufficiently large m there is $n > m$ such that $r(m+1), \ldots, r(n)$ contains every j at least once $(j=1, 2, \ldots, N)$ and $\frac{\|S_n f\|}{\|S_m f\|} \geqq 1 - \varepsilon$ where $\varepsilon = \varepsilon(\mathfrak{V}, N)$, so that as in (iv) and (v) $P_N\left(\frac{S_m f}{\|S_m f\|}\right) \in \mathfrak{V}$, hence

$$S_m f = S_m P_N f = P_N S_m f \in \|f\| \mathfrak{V}.$$

The arbitrariness of $\mathfrak{V}$ implies the weak convergence of $S_n P_N$ to 0. Thus S_n converges weakly to Q_N.

*

Corollary. *If T_j is a contraction with* (W) *or equivalently* (W′) $(j=1, 2, \ldots, N)$, *then for any random selection $r(\cdot)$ the sequence*

$$S_n = T_{r(1)} T_{r(2)} \cdots T_{r(n)}$$

converges weakly as $n \to \infty$.

Proof. It is proved in § 2 that T_j^* has (W) $(j=1, 2, \ldots, N)$, so that by the Theorem, the product $T_{r(n)}^* \ldots T_{r(2)}^* T_{r(1)}^*$ converges weakly, hence $(S_n f, g) = (f, T_{r(n)}^* \ldots T_{r(2)}^* T_{r(1)}^* g)$ converges for all f, g.

This completes the proof.

References

[1] A. Beck and J. T. Schwartz, A vector valued random ergodic theorem, *Proc. Amer. Math. Soc.*, **8** (1957), 1049–1059.

[2] R. S. Foguel, Powers of a contraction in Hilbert space, *Pacific J. Math.*, **13** (1963), 551–562.

[3] I. Halperin, The product of projection operators, *Acta Sci. Math.*, **23** (1962), 96–99.

[4] H. Langer, Ein Zerspaltungssatz für Operatoren im Hilbertraum, *Acta Math. Acad. Hungar.*, **12** (1961), 441–445.

[5] M. Prager, On a principle of convergence in a Hilbert space, *Czech. Math. J.*, **10** (1960), 271–282.

[6] F. Riesz and B. Sz.-Nagy, *Leçons d'analyse fonctionnelle* (Budapest, 1952).

[7] B. Sz.-Nagy and C. Foiaş, Sur les contractions de l'espace de Hilbert. IV, *Acta Sci. Math.*, **21** (1960), 251–259.

(Received November 6, 1964)