Z. Wahrscheinlichkeitstheorie verw. Geb. 21, 207–214 (1972)

Approximation Theorems for Markov Operators*

CHOO-WHAN KIM

1. Introduction and Preliminaries

Let $(X, \mathscr{F}, \mu)$ be the unit interval, the Lebesgue measurable sets, and Lebesgue measure, respectively. All functions (maps) on X are $\mathscr{F}$-measurable real functions and will always be considered up to μ-equivalence. All sets that are referred to are elements of $\mathscr{F}$. We shall omit the phrase almost everywhere, it being understood where applicable.

By a Markov operator P on $L^\infty(X, \mathscr{F}, \mu)$ we shall mean a positive linear operator P from $L^\infty(X, \mathscr{F}, \mu)$ into $L^\infty(X, \mathscr{F}, \mu)$ such that $P1=1$ and $g_n \downarrow 0$ implies $Pg_n \downarrow 0$. A positive linear operator T from $L^1(X, \mathscr{F}, \mu)$ into $L^1(X, \mathscr{F}, \mu)$ such that $\int_X Tf\,d\mu = \int_X f\,d\mu$ for each $f \in L^1(X, \mathscr{F}, \mu)$ is called a Markov operator on $L^1(X, \mathscr{F}, \mu)$. It is well known that each Markov operator P on $L^\infty(X, \mathscr{F}, \mu)$ is the adjoint of a uniquely determined Markov operator T on $L^1(X, \mathscr{F}, \mu)$, that is, $P = T^*$. A Markov operator P on $L^\infty(X, \mathscr{F}, \mu)$ satisfying the condition that $\int_X Pg\,d\mu = \int_X g\,d\mu$ for each $g \in L^\infty(X, \mathscr{F}, \mu)$ is called a doubly stochastic operator. Every doubly stochastic operator P is uniquely extended to a positive linear operator from $L^p(X, \mathscr{F}, \mu)$ into itself with $\|P\|_p = 1$, $1 \leqq p < \infty$. Doob [3, p. 293] showed that each Markov operator P on $L^\infty(X, \mathscr{F}, \mu)$ induces a doubly stochastic operator P' in the sense of [3, p. 288]. We will briefly outline Doob's argument. Let P and T be, respectively, a Markov operator on $L^\infty(X, \mathscr{F}, \mu)$ and a Markov operator on $L^1(X, \mathscr{F}, \mu)$ such that $P = T^*$. Let $d\mu' = T1\,d\mu$ and let $Y = (x\colon T\,1(x) > 0)$. Then $\int_X Pg\,d\mu = \int_X g\,d\mu'$ for each $g \in L^\infty(X, \mathscr{F}, \mu)$. Note that $Pg = 0$ if $g = 0$ on Y. Define a positive linear operator P' from $L^\infty(X, \mathscr{F}, \mu')$ into $L^\infty(X, \mathscr{F}, \mu)$ by $P'g = Pg$. Define $T'f = Tf/T1$ on Y and $T'f = 0$ elsewhere, $f \in L^1(X, \mathscr{F}, \mu)$. Then T' is a positive linear operator from $L^1(X, \mathscr{F}, \mu)$ into $L^1(X, \mathscr{F}, \mu')$ such that $T'1 = 1$ and $T'^* = P'$. Both P' and T' are doubly stochastic in the sense of Doob. As corollaries of Brown's approximation theorems [1, Theorems 1 and 2] (see also [7, Theorem 2.2]) we have that

$$\text{(i)} \qquad \left|\int_X f(P' - T_{\phi_n})\,g\,d\mu\right| = \left|\int_X (T' - T_{\phi_n^{-1}})\,f \cdot g\,d\mu'\right| \to 0 \quad \text{as } n \to \infty$$

where $f \in L^1(X, \mathscr{F}, \mu)$ and $g \in L^\infty(X, \mathscr{F}', \mu')$. Here ϕ_n are bijective (invertible) measure preserving maps from $(X, \mathscr{F}, \mu)$ onto $(X, \mathscr{F}, \mu')$, and $T_{\phi_n} g(x) = g(\phi_n(x))$.

$$\text{(ii)} \qquad \int_X |P'g - Q'_n g|\,d\mu \to 0 \quad \text{and} \quad \int_X |T'f - S'_n f|\,d\mu' \to 0 \quad \text{as } n \to \infty$$

where $g \in L^1(X, \mathscr{F}, \mu')$ and $f \in L^1(X, \mathscr{F}, \mu)$. Q'_n and S'_n denote, respectively, convex combinations $\sum_t c_t T_{\phi_t}$ and $\sum_t c_t T_{\phi_t^{-1}}$ where ϕ_t are as in (i).

* This work was supported in part by NSF Grant GP-5270 and in part by NRC Grant A-4844.

For brevity, we will denote $(X, \mathscr{F}, \mu)$ by X, and $L^p(X, \mathscr{F}, \mu)$ by $L^p(X)$, $1 \leqq p \leqq \infty$. Let $[L^p(X)]$ be the vector space of all bounded linear operators from $L^p(X)$ into itself. Throughout this paper, by a Markov operator we shall always mean a Markov operator on $L^\infty(X)$. Then the set $\mathscr{M}$ of all Markov operators is a convex subset of $[L^\infty(X)]$. Each non-singular map (μ-nonsingular, measurable point map) $\phi: X \to X$ defines a Markov operator T_ϕ by the formula $T_\phi g(x) = g(\phi(x))$. The set of such operators T_ϕ is denoted by Ψ. Let Ψ_i be the set of those $T_\phi \in \Psi$ such that ϕ is an injection. A μ-continuous Markov operator P is a Markov operator for which there is $0 < p(x, y) \in L^1(X \times X)$ such that $Pg(x) = \int_X p(x, y)\, g(y)\, d\mu(y)$ (see Moy [9]). We call $p(.,.)$ the kernel of P. A μ-continuous Markov operator P with kernel $p(.,.)$ is called a Hilbert-Schmidt type if $p(.,.) \in L^2(X \times X)$. The convex hull of a subset $\mathscr{E}$ of $\mathscr{M}$ is denoted by $\mathrm{ch}(\mathscr{E})$. Then $\mathrm{ch}(\mathscr{E}) \subset \mathscr{M}$. By the weak* operator topology in $\mathscr{M}$ we mean the topology inherited from the weak* operator topology in $[L^\infty(X)]$. Similarly the strong operator topology in $[L^\infty(X)]$ restricted to $\mathscr{M}$ is called the strong operator topology in $\mathscr{M}$. The uniform topology in $\mathscr{M}$ is the topology of $\mathscr{M}$ induced by the metric

$$\|P - Q\|_{\infty,1} = \sup\{\|(P-Q)\, g\|_1 : \|g\|_\infty \leqq 1\}.$$

The purpose of this paper is to prove that Ψ_i is dense in $\mathscr{M}$ in the weak* operator topology (Theorem 1), and $\mathrm{ch}(\Psi)$ is dense in $\mathscr{M}$ in the strong operator topology (Theorem 2). We will also prove the uniform approximation theorem (Theorem 3): for each μ-continuous Markov operator P, there is a sequence in $\mathrm{ch}(\Psi)$ converging to P in the uniform topology. If P is a Markov operator of Hilbert-Schmidt type, there is a sequence in $\mathrm{ch}(\Psi)$ converging to P in the $L^2(X)$-operator norm topology (Theorem 4). The results in this paper generalize approximation theorems of Brown [1, Theorems 1 and 2] and the author [7, Theorems 2.3 and 2.4].

We refer to Neveu [10] for the elementary concepts of probability theory and others related to Markov operators used in the text.

The author wishes to thank the referee for valuable suggestions which have led to the improved versions of Theorems 1 and 2 presented in this paper.

2. Basic Lemmas

A finite collection $\pi = (A, \ldots, A_n)$ of pairwise disjoint subsets of X such that $X = \bigcup_i A_i$ and $\mu(A_i) > 0$ for all i is called a finite partition of X. In what follows, by a partition π we shall mean a finite partition of X in the above sense of the term. If we denote by Π the family of all partitions π, then Π is a directed set, when ordered by the relation $\pi \leqq \pi'$ iff π' refines π. The conditional expectation operator U_π relative to a partition $\pi = (A_1, \ldots, A_n)$ is defined by

$$U_\pi g = \sum_i \frac{1}{\mu(A_i)} \Big(\int_{A_i} g\, d\mu \Big) 1_{A_i}, \qquad g \in L^\infty(X).$$

Here 1_{A_i} denotes the indicator function of the set A_i. It is easily seen that U_π is a doubly stochastic operator such that $U_\pi U_\pi = U_\pi$, $U_\pi^* = U_\pi$ and $U_\pi U_{\pi'} = U_{\pi'} U_\pi = U_\pi$

for $\pi \leqq \pi'$. We shall denote the net $(U_\pi: \pi \in \Pi)$ briefly by (U_π). It is well known [4, p. 500] that the net (U_π) converges to the identity operator I in the strong operator topology for $[L^\infty(X)]$. By the mean differentiation theorem of Helms [6, p. 446], the net (U_π) converges also to I in the strong operator topology for $[L^p(X)]$, $1 \leqq p < \infty$. We define also the following sequence of conditional expectation operators. Let Λ be the subfamily of Π consisting of all dyadic partitions

$$\Delta_n = \{[(k-1)\,2^{-n}, k2^{-n});\ k=1, \ldots, 2^n\}, \qquad n=1,2,\ldots.$$

Then Λ with relation $\leqq$ is a linearly ordered set and hence a directed set. Define the sequence (U_n) by $U_n = U_{\Delta_n}$, $n=1,2,\ldots$. Note that the sequence (U_n) is not a subnet of the net (U_π). It follows essentially from the martingale convergence theorem of Doob [2, p. 319] that the sequence (U_n) converges to I in the strong operator topology for $[L^p(X)]$, $1 \leqq p < \infty$. Let $\langle f, g\rangle = \int_X f \cdot g \, d\mu$ where $f \in L^1(X)$ and $g \in L^\infty(X)$.

We shall state and prove two lemmas which generalize Lemmas 2.1 and 2.2 of [7], respectively.

Lemma 1. *For each* $P \in \mathcal{M}$ *and* $\pi \in \Pi$, *there is* $T_\phi \in \Psi_i$ *such that*

$$U_\pi P U_\pi = U_\pi T_\phi U_\pi.$$

Proof. Let $\pi = (A_1, \ldots, A_n)$ and let $a_{ij} = \langle 1_{A_i}, P 1_{A_j}\rangle$ $(i, j=1, \ldots, n)$. Let $(B_{i1}, \ldots, B_{in})$ be a partition of A_i into disjoint sets such that $\mu(B_{ij}) = a_{ij}$ $(i, j=1, \ldots, n)$. Let $(C_{1j}, \ldots, C_{nj})$ be a partition of A_j into disjoint sets such that

$$\mu(C_{ij}) > 0 \qquad (i, j=1, \ldots, n).$$

Using the isomorphism theorems of Halmos and von Neumann [5, Theorems 1 and 2] we can show that there is a point map ϕ on X such that ϕ: $B_{ij} \to C_{ij}$ is a non-singular bijection whenever $\mu(B_{ij}) > 0$ $(i, j=1, \ldots, n)$. Clearly $\phi \in \Psi_i$. Note that ϕ is not necessarily a bijection on X. We now have

$$\langle 1_{A_i}, T_\phi 1_{A_j}\rangle = \sum_k \mu(A_i \cap \phi^{-1}(C_{kj})) = \sum_k \mu(A_i \cap B_{kj}) = \mu(B_{ij}) = \langle 1_{A_i}, P 1_{A_j}\rangle$$

for $i, j=1, \ldots, n$ and hence $U_\pi T_\phi U_\pi = U_\pi P U_\pi$.

In particular, if P is doubly stochastic, we choose C_{ij} such that $\mu(C_{ij}) = a_{ij}$ $(i, j=1, \ldots, n)$. In this case we can choose a bijective (invertible) measure preserving map ϕ satisfying $U_\pi P U_\pi = U_\pi T_\phi U_\pi$. This completes the proof.

Lemma 2. *For each* $P \in \mathcal{M}$ *and* $\pi \in \Pi$, *there is* $S \in \mathrm{ch}(\Psi)$ *such that*

$$U_\pi P U_\pi = U_\pi S U_\pi = S U_\pi.$$

Proof. Given a partition $\pi = (A_1, \ldots, A_n)$, let $a_{ij} = \langle 1_{A_i}, P 1_{A_j}\rangle$ and $p_{ij} = a_{ij}/\mu(A_i)$ $(i, j=1, \ldots, n)$. Set $U = U_\pi$. Then we have

$$UPU 1_{A_j} = UP 1_{A_j} = \sum p_{ij} 1_{A_i} \qquad (j=1, \ldots, n).$$

Note that the $n \times n$ matrix $\rho=(p_{ij})$ is a row-stochastic matrix, that is, $p_{ij} \geqq 0$ $(i, j=1, \ldots, n)$ and $\sum_j p_{ij}=1$ for each i. Hence the matrix ρ is a convex combination $\rho=\sum_t c_t v_t$ where v_t are $n \times n$ row-stochastic matrices having exactly one entry 1 in each row (see [8, p. 133]). Given such a matrix v_t, there is a unique map Φ_t from the set $(1, \ldots, n)$ into itself such that

$$(v_t)_{ij}=\delta_{\Phi_t(i), j} \quad (i, j=1, \ldots, n)$$

where $(v_t)_{ij}$ are entries of v_t and $\delta_{k,j}$ is the Kroneker delta. Then we have

$$p_{ij}=\sum_t c_t \delta_{\Phi_t(i), j} \quad (i, j=1, \ldots, n).$$

Let ϕ_t be a point map on X such that ϕ_t: $A_i \rightarrow A_{\Phi_t(i)}$ is a non-singular bijection for each i (i$=1, \ldots, n$). Such a map exists by an argument given in the proof of Lemma 1. Clearly ϕ_t is a non-singular map on X. Since Φ_t is not necessarily an injection, the same is true for ϕ_t. Define the operator S by

$$S=\sum_t c_t T_{\phi_t}.$$

Then $S \in \mathrm{ch}(\Psi)$. Thus it remains to show that S satisfies $UPU=SU=USU$. Note that for each t and j,

$$T_{\phi_t} 1_{A_j}=\sum_i \delta_{\Phi_t(i), j} 1_{A_i}.$$

It follows that for each j $(j=1, \ldots, n)$,

$$S 1_{A_j}=\sum_t c_t \sum_i \delta_{\Phi_t(i), j} 1_{A_i}=\sum_i p_{ij} 1_{A_i}=UPU 1_{A_j}.$$

We have at once the assertion.

The following version of Lemma 2 will be needed later. Let Φ be the set of those $T_\phi \in \Psi$ such that ϕ is a measure preserving map. Let Φ_b denote the set of those $T_\phi \in \Phi$ such that ϕ is a bijective (invertible) measure preserving map.

Lemma 3. *Let n be a fixed positive integer and let $\pi=(A_1, \ldots, A_n)$ be such that $\mu(A_i)=1/n$, $i=1, \ldots, n$. Then for each $P \in \mathscr{M}$, there is $S \in \mathrm{ch}(\Psi)$ with the following properties:*

(i) *$UPU=USU=SU$ where $U=U_\pi$.*

(ii) *S extends uniquely to a positive element of $[L^p(X)]$ with*

$$0<\|S\|_p \leqq \|S\|_1^{1/p} \leqq n^{1/p} \quad (1 \leqq p<\infty).$$

In particular, if P is doubly stochastic, $S \in \mathrm{ch}(\Phi_b)$.

Proof. Notation is as in the proof of Lemma 2. Assertion (i) is obvious. To prove Assertion (ii), the following maps ϕ_t will be used in the definition of S. For a fixed t, we choose a map ϕ_t such that ϕ_t: $A_i \rightarrow A_{\Phi_t(i)}$ is a measure preserving bijection for each i ($i=1,\dots,n$). Then we have for each set E and j,

$$\|T_{\phi_t} 1_{E\cap A_j}\|_1 = \Big\|\sum_i \delta_{\Phi_t(i),j}\, 1_{\phi_t^{-1}(E\cap A_j)\cap A_i}\Big\|_1 = \sum_i \delta_{\Phi_t(i),j}\, \mu(E\cap A_j),$$

$$\|S 1_{E\cap A_j}\|_1 \leqq \sum_t c_t \sum_i \delta_{\Phi_t(i),j}\, \mu(E\cap A_j) = \sum_i p_{ij}\, \mu(E\cap A_j)$$

and thus

$$\|S 1_E\|_1 \leqq \beta\, \mu(E) \qquad \text{where } \beta = \max_j \Big(\sum_i p_{ij}\Big).$$

It is straightforward to show that S is uniquely extended to a positive linear operator on $L^1(X)$, denoted also by S, with

$$0 < \|S\|_1 = \beta \leqq n.$$

For example, if P is defined by $Pg = n\langle 1_{A_1}, g\rangle$ for $g\in L^\infty(X)$, then $\beta = n$. By the Riesz convexity theorem [4, p. 525] and $\|S\|_\infty = 1$, we have for $1<p<\infty$,

$$0 < \|S\|_p \leqq \|S\|_1^{1/p}\, \|S\|_\infty^{1-1/p} \leqq \|S\|_1^{1/p} \leqq n^{1/p}.$$

Suppose P is doubly stochastic. Then the matrix ρ is doubly stochastic so that it is a convex combination of permutation matrices v_t. In this case the maps ϕ_t defined above are measure preserving bijections and hence $S\in \mathrm{ch}(\Phi_b)$.

The following is a reformulation of [7, Lemma 2.5].

Lemma 4. *Let U_n be the conditional expectation operator relative to a dyadic partition $\Delta_n = (A_1, \dots, A_{2^n})$. Then there exist measure preserving maps θ_1 and θ_2 on X such that for each $j=1,2$, $\theta_j(A_i)\subset A_i$ for all A_i in Δ_n and if $V=\frac{1}{2}(T_{\theta_1}+T_{\theta_2})$,*

$$\|V^{2k} - U_n\|_2 \leqq 2^{-k}, \qquad k=1,2,\dots.$$

3. Approximations

We begin by proving

Theorem 1. *For each Markov operator P, there is a net $(P_\pi\colon \pi\in\Pi)$ in Ψ_i converging to P in the weak* operator topology.*

Proof. It suffices to prove that for each Markov operator P, there is a net $(P_\pi\colon\ \pi\in\Pi)$ *in* Ψ_i *such that*

$$\lim_\pi \langle f, P_\pi g\rangle = \langle f, Pg\rangle$$

where $f\in L^1(X)$ and $g\in L^\infty(X)$. We may assume without loss of generality that $\|f\|_1\leqq 1$ and $\|g\|_\infty \leqq 1$. Since the net (U_π) converges to I in the strong operator topology for $[L^p(X)]$, $1\leqq p\leqq\infty$, there is $\pi_0\in\Pi$ such that

$$\|f - U_\pi f\|_1 < \varepsilon/4 \quad \text{and} \quad \|g - U_\pi g\|_\infty < \varepsilon/4$$

for all $\pi \geqq \pi_0$. Here $\varepsilon>0$ is arbitrary. Let $\phi=\phi_\pi$ and $U=U_\pi$ be as in Lemma 1 where $\pi\geqq\pi_0$. It follows that

$$\begin{aligned}|\langle f, (P-T_\phi)g\rangle| &\leqq |\langle f, (P-UPU)g\rangle| + |\langle f, (UT_\phi U - T_\phi)g\rangle|\\ &\leqq |\langle f-Uf, Pg\rangle| + |\langle Uf, P(g-Ug)\rangle| + |\langle Uf, T_\phi(Ug-g)\rangle| + |\langle Uf-f, T_\phi g\rangle|\\ &\leqq 2\|f-Uf\|_1 + 2\|g-Ug\|_\infty < \varepsilon.\end{aligned}$$

By setting $P_\pi = T_{\phi_\pi}$ we complete the proof.

Remark. Let $\mathcal{M}_1$ be the convex set of all Markov operators on $L^1(X)$. Let Θ be the set of those $T\in\mathcal{M}_1$ such that $T^*\in\Psi_i$. An immediate corollary to Theorem 1 is that Θ is dense in $\mathcal{M}_1$ in the weak operator topology for $[L^1(X)]$. Since every convex set in $[L^1(X)]$ has the same closure in the weak operator and the strong operator topologies [4, p. 477], the convex hull of Θ is dense in $\mathcal{M}_1$ in the strong operator topology. A result [7, Proposition 1.1] may be used to show that the mapping $T\to T^*$ of $\mathcal{M}_1$ onto $\mathcal{M}$ is not continuous when $\mathcal{M}_1$ and $\mathcal{M}$, respectively, endowed with the strong operator topology for $[L^1(X)]$ and the strong operator topology for $[L^\infty(X)]$. Hence the convex hull of Ψ_i is not dense in $\mathcal{M}$ in the strong operator topology.

We have the following strong approximation theorem for $\mathcal{M}$.

Theorem 2. *For each Markov operator P, there is a net $(S_\pi\colon \Pi)$ in* $\mathrm{ch}(\Psi)$ *converging to P in the strong operator topology.*

Proof. A subbase at a Markov operator P for the strong operator topology is given by set of the form

$$N(P,g,\varepsilon)=\{Q\colon \|(P-Q)g\|_\infty<\varepsilon\}$$

where $g\in L^\infty(X)$ and $\varepsilon>0$. Choose $\pi_0\in\Pi$ such that

$$\|Pg-U_\pi Pg\|_\infty<\varepsilon/3 \quad\text{and}\quad \|g-U_\pi g\|_\infty<\varepsilon/3$$

for all $\pi\geqq\pi_0$. Let $S=S_\pi$ and $U=U_\pi$ be as in Lemma 2 where $\pi\geqq\pi_0$. We have that $S\in\mathrm{ch}(\Psi)$ and

$$\begin{aligned}\|(P-S)g\|_\infty &\leqq \|(P-UPU)g\|_\infty+\|(SU-S)g\|_\infty\\ &\leqq \|Pg-UPg\|_\infty+\|UP(g-Ug)\|_\infty+\|S(Ug-g)\|_\infty<\varepsilon.\end{aligned}$$

Thus the net $(S_\pi g)$ is eventually in $N(P,g,\varepsilon)$. This completes the proof.

For μ-continuous Markov operator P, we prove the following uniform approximation theorem.

Theorem 3. *For each μ-continuous Markov operator P, there is a sequence (Q_k) in* $\mathrm{ch}(\Psi)$ *such that*

$$\|P-Q_k\|_{\infty,1}\to 0 \quad \textit{as } k\to\infty.$$

Proof. Recall that U_n denotes the conditional expectation operator relative to a dyadic partition Δ_n. Let p be the kernel for P and let $(U_n\otimes U_n)p$ be the kernel for U_nPU_n. Since

$$\|P-U_nPU_n\|_{\infty,1}\leqq\|p-(U_n\otimes U_n)p\|_{L^1(X\times X)}\to 0 \quad \text{as } n\to\infty,$$

we can choose a positive integer n such that

$$\|P-U_nPU_n\|_{\infty,1}<\varepsilon/2,$$

where $\varepsilon>0$ is arbitrary. Choose a positive integer k_0 such that $2^{n+1}<2^{k_0}\varepsilon$. Let V be the operator as in Lemma 4. We have then

$$\|V^{2k}-U_n\|_2<\varepsilon/2^{n+1} \quad \text{for all } k\geqq k_0.$$

If S is the Markov operator as in Lemma 3 where $\pi=\Delta_n$, then $U_n P U_n = S U_n$ and $\|S\|_1 \leqq 2^n$. It follows that for each $g \in L^\infty(X)$ and $k(\geqq k_0)$,

$$\|(U_n P U_n - S V^{2k}) g\|_1 = \|S(U_n - V^{2k}) g\|_1 \leqq \|S\|_1 \|(U_n - V^{2k}) g\|_1$$
$$\leqq \|S\|_1 \|(U_n - V^{2k}) g\|_2 \leqq \|S\|_1 \|U_n - V^{2k}\|_2 \|g\|_2 < \frac{\varepsilon}{2} \|g\|_\infty$$

and hence

$$\|U_n P U_n - S V^{2k}\|_{\infty,1} \leqq \varepsilon/2.$$

Thus we have

$$\|P - S V^{2k}\|_{\infty,1} < \varepsilon \quad \text{for all } k \geqq k_0.$$

If we set $Q_k = S V^{2k}$, then $Q_k \in \mathrm{ch}(\Psi)$ because $V^{2k} \in \mathrm{ch}(\Phi) \subset \mathrm{ch}(\Psi)$ and $S \in \mathrm{ch}(\Psi)$. In particular, if P is doubly stochastic, then by Lemma 3(ii), $S \in \mathrm{ch}(\Phi_b)$ and hence $Q_k \in \mathrm{ch}(\Phi)$. This completes the proof.

We state the $L^2(X)$-operator norm approximation theorem for Markov operators of Hilbert-Schmidt type.

Theorem 4. *For each Markov operator P of Hilbert-Schmidt type, there is a sequence (Q_k) in $\mathrm{ch}(\Psi)$ such that*

$$\|P - Q_k\|_2 \to 0 \quad \text{as } k \to \infty.$$

Proof. This proof is similar to that of Theorem 3. If P has the kernel $p \in L^2(X \times X)$, then P is extended to a positive linear operator on $L^2(X)$ with $\|P\|_2 \leqq \|p\|_{L^2(X \times X)}$.

Note that

$$\|P - U_n P U_n\|_2 \leqq \|p - (U_n \otimes U_n) p\|_{L^2(X \times X)} \to 0 \quad \text{as } n \to \infty.$$

Thus we can choose n such that

$$\|P - U_n P U_n\|_2 < \varepsilon/2$$

where $\varepsilon > 0$ is arbitrary. Choose k_0 such that $2^{n+1} < 2^{k_0} \varepsilon$. Let S and V be as in the proof of Theorem 3. For each $k \geqq k_0$, we have from Lemma 3(ii) that

$$\|U_n P U_n - S V^{2k}\|_2 = \|S(U_n - V^{2k})\|_2 \leqq \|S\|_2 \|U_n - V^{2k}\|_2 < \varepsilon/2 \sqrt{2^n} < \varepsilon/2$$

and thus

$$\|P - S V^{2k}\|_2 < \varepsilon.$$

Since $S V^{2k} \in \mathrm{ch}(\Psi)$, the theorem follows.

References

1. Brown, J.R.: Approximation theorems for Markov operators. Pacific J. Math. **16**, 13–23 (1966).
2. Doob, J.L.: Stochastic processes. New York: Wiley 1953.
3. – A ratio operator limit theorem. Z. Wahrscheinlichkeitstheorie verw. Geb. **1**, 288–294 (1963).
4. Dunford, N., Schwartz, J.T.: Linear operators, part I. New York: Interscience 1967.
5. Halmos, P.R., Neumann, J. von: Operator methods in classical mechanics. Ann. of Math. II. Ser. **43**, 332–350 (1942).

6. Helms, L.L.: Mean convergence of martingales. Trans. Amer. math. Soc. **87**, 439–446 (1958).
7. Kim, C.W.: Uniform approximation of doubly stochastic operators. Pacific J. Math. **26**, 515–527 (1968).
8. Marcus, M., Minc, H.: A survey of matrix theory and matrix inequalities. Boston: Allyn and Bacon 1964.
9. Moy, S.-T.C.: λ-continuous Markov chains. Trans. Amer. math. Soc. **117**, 68–91 (1965).
10. Neveu, J.: Mathematical foundations of the calculus of probability. San Francisco: Holden-Day 1965.

Choo-whan Kim
Department of Mathematics
Simon Fraser University
Burnaby, B.C.
Canada

(Received December 1, 1970/April 14, 1971)