Z. Wahrscheinlichkeitstheorie verw. Geb. 28, 173–177 (1974)

Sample Path Convergence of Stable Markov Processes

B. Jamison* and R. Sine

0.

In recent years several people [5, 6, 8] have sharpened the Krylov-Bogoliubov topological decomposition of a compact space relative to a Markov operator (due to Bebutov [1]) but at the cost of introducing the quite stringent condition of uniform mean stability. One of the authors [9] has shown that a portion of this theory holds under the weaker condition that the operator be stable, that is takes continuous functions into continuous functions. (For example, this condition guarantees that the union of the ergodic sets is closed.) In this paper we extend the results of [5] concerning the sample path behavior of the Markov processes by showing that the convergence of the paths to ergodic sets and ergodic kernels holds under the condition of stability. We also clarify the relationship between our Krylov-Bogoliubov type of ergodic decompositon and Foguel's recent topological analogue of Hopf's ergodic decomposition [4].

1.

Let S be a compact Hausdorff space, and $C(S)$ the Banach space of all real-valued continuous functions on S. A linear operator T on $C(S)$ is called a *Markov operator* on S if $T \geqq 0$ and $T(1)=1$. We denote by $\mathfrak{M}$ the collection of all $f \in C(S)$ with $Tf=f$; the members of $\mathfrak{M}$ are called *invariant*. A member μ of the dual space $C^*(S)$ is called a *probability* if $\mu \geqq 0$ and $\mu(1)=1$. Let $\mathscr{L}$ be the collection of all probabilities μ for which $T^*(\mu)=\mu$; the members of $\mathscr{L}$ are also called invariant. It is well known that $\mathscr{L}$ is not empty. An equivalence relation is defined on $S \times S$ as follows: $x \sim y$ iff $f(x)=f(y)$ for each $f \in \mathfrak{M}$. Let $\mathscr{D}$ be the decomposition of S induced by $\sim$. The member of $\mathscr{D}$ are obviously closed. A member E of $\mathscr{D}$ is called an *ergodic set* if there is a $\mu \in \mathscr{L}$ with $\mu(E)=1$. Let $\mathscr{E}$ be the class of all ergodic sets; its union $\Sigma\mathscr{E}$ is a non-empty closed set (see [9], Theorem 1.9). If $E \in \mathscr{E}$, we denote by K_E the closure of the union of the supports of all $\mu \in \mathscr{L}$ with $\mu(E)=1$. We call K_E the *ergodic kernel* (of E). For each $x \in S$ let δ_x be the unit mass concentrated at x, and let $P^n(x,.)=(T^*)^n \delta_x$. A non-empty closed set F is called *self-supporting* if $P(x,F)=1$ for all $x \in F$. Ergodic sets need not be self-supporting (see the second example on p. 162 of [9]), but their ergodic kernels are (this follows from Theorem 1.3 of [9] and the fact that the closure of the union of a collection of self-supporting sets is itself self-supporting, a consequence of Theorem 1.1 of [9]). For each $E \in \mathscr{E}$ let

$$E^1 = \bigcap_{n=1}^{\infty} \{x \colon x \in E,\ P^n(x,E)=1\}.$$

* Partially supported by National Science Foundation Grant SD GU 3171.

If $P^m(x, \{y: P^n(y, E)<1\})>0$, then $P^{m+n}(x, E)<1$, and it follows that E^1 is self-supporting; we call E^1 the *self-supporting part* of E. Clearly $K_E \subset E^1 \subset E$, and the inclusions may be proper.

Theorem 1. *For each $x \in S \setminus \Sigma\mathscr{E}$ there is an f on S with the following properties:*

(i) $f(x)>0$,

(ii) $0 \leqq f \leqq 1$,

(iii) f *is lower semicontinuous,*

(iv) $Tf \leqq f$,

(v) $\lim_n T^n f = 0$ *pointwise on* S.

Proof. Fix $x \in S \setminus \Sigma\mathscr{E}$. If $y \in \Sigma\mathscr{E}$, then $x \sim y$ is false, so there is a $g_y \in \mathfrak{M}$ with $g_y(x) \neq g_y(y)$; we may assume that $g_y(x)=0$ and $g_y(y)=1$. Let $O_y=\{z: g_y(z)>1/2\}$. The sets O_y, $y \in \Sigma\mathscr{E}$ are an open cover for the compact set $\Sigma\mathscr{E}$. Let $O_{y_1}, \ldots, O_{y_N}$ be a finite subcover, and $g_i = g_{x_i}$, $i=1, \ldots, N$. Let $g_0 = g_1 \vee \cdots \vee g_N \vee 0$. Observe that $g_0 \in C(S)$, $g_0 \geqq 0$, $g_0(x)=0$, $g_0 > \frac{1}{2}$ on $\Sigma\mathscr{E}$and $Tg_0 \geqq g_0$. We have $T^n g_0 \uparrow g \leqq \|g_0\|$, the convergence being pointwise on S. Clearly $Tg=g$, and g is lower semicontinuous. Let $f=g-g_0$. It is clear that, $0 \leqq f \leqq \|g_0\|$, f is lower semicontinuous, $Tf \leqq f$, and $T^n f \downarrow 0$. We now show that $f(x)>0$; since $g_0(x)=0$, this amounts to showing that $g(x)>0$. We have

$$g(x)=\lim_n T^n g_0(x)=\lim_n (1/n)(T+\cdots+T^n)\, g_0(x).$$

Tyĥonov's theorem assures us of the existence of a subset $\{n_t\}$ of the positive integers for which $(1/n_t)(T+\cdots+T^{n_t})\, h(x)$ converges for all $h \in C(S)$. The map $h \to \lim (1/n_t)(T+\cdots+T^{n_t})\, h(x)$ is a positive linear functional on $C(S)$ which sends 1 into 1, and thus corresponds to a probability measure μ on $C(S)$. Since the functional clearly assigns the same value to Th as it does to h, $\mu \in \mathscr{L}$. It follows that $\mu(\Sigma\mathscr{E})=1$. (This result, well known if S is metrizable, is an easy consequence of Theorem 1.11 of [9] and the Krein-Milman theorem.) Thus

$$\begin{aligned} g(x) &= \lim_n (1/n)(T+\cdots+T^n)\, g_0(x) = \lim (1/n_t)(T+\cdots+T^{n_t})\, g_0(x) \\ &= \int g_0 \, d\mu \geqq \tfrac{1}{2} > 0 \end{aligned}$$

since $g_0 \geqq \frac{1}{2}$ on $\Sigma\mathscr{E}$. If $\|f\|>1$, replace f by $f/\|f\|$. Now f has properties (i)–(v), and the theorem is proved.

Let Ω be the product space $\prod_{i=0}^{\infty} S_i$, where $S_i=S$, $i=0, 1, \ldots$, and let $\mathfrak{F}$ be the corresponding product σ-field $\prod_{i=0}^{\infty} \Sigma_i$, where for each $i=0, 1, \ldots, \Sigma_i=\Sigma$, the σ-field of Borel subsets of S. For each $i=0, 1, \ldots$, let X_i be the coordinate function on Ω; that is, $X_i(\omega)=\omega_i$, where $\omega=(\omega_0, \omega_1, \ldots)$ belongs to Ω. It is well known that for each probability measure μ on Σ there is a probability measure P_μ on $\mathfrak{F}$ for which $P_\mu(X_0 \in E)=\mu(E)$ for each $E \in \Sigma$ and $P_\mu(X_{n+k} \in E \mid X_0, \ldots, X_n)=P^k(X_n, E)$ P_μ-almost surely for each $n=0, 1, \ldots, k=1, 2, \ldots$ and $E \in \Sigma$. Then $X_0, X_1, \ldots$ is a Markov process; we call it the *process with initial distribution* μ. We denote the corresponding expectation operator by E_μ. If no misunderstanding can arise, we use P and E for P_μ and E_μ respectively. Let $A \in \Sigma$, and let $x_0, x_1, \ldots, x_n, \ldots$ be a sequence in S. We say that x_n *converges to* A, and write $x_n \to A$, if, any open neighborhood V of A, $x_n \in V$ for all but a finite number of values of n, and that x_n *converges to* A *in density if* $\lim_n (1/n) \sum_{k=1}^n 1_V(x_k) = 1$ for any such V. We write "$x_n \in A$ i.o." if $x_n \in A$ for infinitely many values of x, and "$x_n \in A$ ult." if $x_n \in A$ for all but a finite number of values of n.

Corollary. *Let D be the dissipative part of S in the sense of Foguel. Then $S \setminus \Sigma\mathscr{E} \subset D$. Let K be a compact subset of $S \setminus \Sigma\mathscr{E}$. Then, for any probability measure μ on S,*

$$P_\mu(X_n \in K \text{ i.o.}) = 0.$$

Proof. According to Foguel's definition [4], the dissipative set D consists of all points x for which there is an f satisfying conditions (i)–(v) of Theorem 1, so the first assertion of the corollary is immediate. Let K be a compact subset of $S \setminus \Sigma\mathscr{E}$. We deduce from the theorem via a simple compactness argument the existence of functions $f_1, \ldots, f_N$ satisfying conditions (ii)–(v) of the theorem and for which

$$\inf\{f_i(x): x \in K\} \geqq \delta > 0, \qquad i = 1, \ldots, N.$$

Let

$$K_i = \{x: x \in K, f_i(x) \geqq \delta\}, \qquad i = 1, \ldots, N.$$

Fix i. Since $Tf_i \leqq f_i$, $\{f_i(X_n), n \geqq 0\}$ is a bounded supermartingale, hence converges. But

$$E(\lim_n f_i(X_n)) \leqq \liminf_n E f_i(X_n) = \liminf E(T^n f_i(X_0)) = 0$$

by virtue of (v). Since $f_i = \delta$ on K_i, this shows that $P(X_n \in K_i \text{ i.o.}) = 0$. But $K = K_1 \cup \cdots \cup K_N$, so $P(X_n \in K \text{ i.o.}) = 0$.

The dissipative part of D of S is open, so its complement C, called the conservative set, is closed. Although $C \subset \Sigma\mathscr{E}$, as we have just observed, it is not true in general that $C = \Sigma\mathscr{E}$. We give an example to show this. Let Z be the set of positive integers with the discrete topology, and let $S = Z \cup \{\infty\}$ be the one-point compactification of Z. For each real-valued function f on S let $Tf(n) = f(n+1)$ for $n \in Z$ and $Tf(\infty) = f(\infty)$. Clearly T is a Markov operator on $C(S)$. Let $h(n) = 1/n$, $n \in Z$, and $h(\infty) = 0$. Then h satisfies conditions (ii)–(v), which shows that $Z \subset D$. Since $T^n f(\infty) = f(\infty)$ for all n, conditions (i) and (v) are incompatible if $x = \infty$, so in fact $Z = D$, and $C = \{\infty\}$. But the only members of $\mathscr{E}$ are the constant functions, so S is the only ergodic set, and $\Sigma\mathscr{E} = S$.

If S is a compact metric space, then any compact set, $\Sigma\mathscr{E}$ in particular, is the intersection of a countable number of its open neighborhoods. It then follows from the corollary that $X_n \to \Sigma\mathscr{E}$ P_μ-almost surely for any initial distribution μ. We now sharpen this result.

Theorem 2. *Let T be a Markov operator on a compact metric space S. Let μ be any probability measure on S, and $\{X_n, n \geqq 0\}$ the process with initial distribution μ. Then there is a set $N \in \mathfrak{F}$ with $P_\mu(N) = 0$ such that for all $\omega \notin N$ there is an $E \in \mathscr{E}$ such that $X_n(\omega)$ converges to the escape-proof part E^1 of E and converges in density to the ergodic kernel K_E contained in E.*

Proof. Assume the hypothesis of the theorem. Let d be a metric for S. Let $f \in C(S)$. It was observed by Breiman ([2], see also [5]) that as a consequence of the "stability theorem" on p. 387 of [7] we have:

$$(1/n)\sum_{k=1}^{n}\{Tf(X_k)-f(X_{k+1})\}\to 0 \qquad P_\mu\text{-almost surely.}$$

Since $C(S)$ is separable relative to the uniform norm, there is an $N_1\in\mathfrak{F}$ with $P(N_1)=0$ such that if $\omega\notin N_1$, then

$$(1/n)\sum_{k=1}^{n}\{Tf(X_k(\omega))-f(X_{k+1}(\omega))\}\to 0$$

for each $f\in C(S)$. As observed in the proof of Theorem 3.2 of [5], this implies that for each $f\in C(S)$ and each $m=1,2,\dots$ we have

$$\lim \frac{1}{n}\sum_{k=1}^{n}\left\{\left(\frac{1}{m}\sum_{j=1}^{m}(T^j f)(X_k(\omega))\right)-f(X_{k+1}(\omega))\right\}=0 \tag{1}$$

provided $\omega\notin N_1$.

Let $f\in\mathfrak{M}$. Then $\{f(X_n), n\geqq 0\}$ is a bounded martingale, so converges P_μ-almost surely. But a subset of a separable metric space is separable so $\mathfrak{M}$ is separable relative to the uniform norm. It follows that there is an $N_2\in\mathfrak{F}$ with $P_\mu(N_2)=0$ such that $\{f(X_n(\omega))\}$ converges for each $f\in\mathfrak{M}$ provided that $\omega\notin N_2$. We have already established the existence of an $N_3\in\mathfrak{F}$ with $P(N_3)=0$ such that $X_n(\omega)\to\Sigma\mathscr{E}$ provided $\omega\notin N_3$. Suppose $\omega\notin N_2\cup N_3$. Then the sequence $\{X_n(\omega)\}$ has all its cluster points in $\Sigma\mathscr{E}$. Suppose one of the cluster points is in the ergodic set E. Then all the cluster points must be in E. For suppose the sequence also has cluster points in another ergodic set F. There is an $f\in\mathfrak{M}$ which assumes different constant values on E and F. But then $\{f(X_n(\omega))\}$ cannot possibly converge, and this contradicts $\omega\notin N_3$. Since S is compact, the fact that all the cluster points of $\{X_n(\omega)\}$ belong to E implies that $X_n(\omega)\to E$.

Now let $\mathscr{C}$ be a countable base for the topology of S. Let $\mathscr{U}$ be the class of all finite unions of members of $\mathscr{C}$; $\mathscr{U}$ is countable. For each $\varepsilon>0$ and $A\in\Sigma$ let $A_\varepsilon=\bigcup_{n=1}^{\infty}\{x\colon P^n(x,A)\geqq\varepsilon\}$. A sequence $x_0,x_1,\dots$ in S is said to have property M_A if for each $\varepsilon>0$ $x_n\in A_\varepsilon$ i.o. $\Rightarrow x_n\in A$ i.o. It follows from a theorem of Doeblin (see [3], Prop. 7) that for each $A\in\Sigma$, $\{X_n(\omega)\}$ has property M_A for P_μ-almost all $\omega\in\Omega$. Thus there is an $N_4\in\mathfrak{F}$ with $P_\mu(N_4)=0$ such that, if $\omega\notin N_4$, then $\{X_n(\omega)\}$ has property M_U for each $U\in\mathscr{U}$. Now suppose $\omega\notin N_2\cup N_3\cup N_4$. All cluster points of the sequence $\{X_n(\omega)\}$ then belong to some ergodic set E. We claim that in fact $X_n(\omega)\to E^1$. For suppose not. Then $\{X_n(\omega)\}$ clusters at some $y\in E\setminus E^1$. For some value of m we have $P^m(y,E)<1$, so there are $\varepsilon>0$ and $\delta>0$ such that $P^m(y,E^\delta)>2\varepsilon$, where $E^\delta=\{z\colon d(z,E)>\delta\}$. Since E^δ is open, a routine argument using the regularity of $P(y,.)$ demonstrates that there is a $U\in\mathscr{U}$ with $U\subset E^\delta$ and $P^m(y,U)>\gamma>0$. Let $V=\{z\colon P^m(z,U)>\gamma\}$. Then V is an open neighborhood of y, so $X_n(\omega)\in V$ i.o. But $V\subset U_\gamma$; since $\omega\notin N_4$, we have $X_n(\omega)\in U$ i.o. Since $U\subset E^\delta$, this in turn implies that $d(X_n(\omega),E)>\delta$ for infinitely many values of n, which contradicts $X_n(\omega)\to E$. So all the cluster points of $X_n(\omega)$ belong to E^1, and $X_n(\omega)\to E^1$.

We need the following result. Let $E\in\mathscr{E}$. Let $f\in C(S)$ be into $[0,1]$ and equal to 0 on K_E. Then $(1/n)(Tf+\cdots+T^n f)$ goes uniformly to 0 on E^1. For suppose not. Then there is an $\varepsilon>0$ and $x_1,x_2,\dots,x_n,\dots$ in E^1 together with $n_1<n_2<\cdots<n_k<\cdots$ for which $(1/n_k)\sum_{j=1}^{n_k}T^j f(x_k)\geqq\varepsilon$ for each $k=1,2,\dots$. By taking a subsequence if necessary, we may assume that $(1/n_k)\sum_{j=1}^{n_k}T^{*j}\delta_{x_k}$ converges in the weak topology as $k \to \infty$ to a measure λ. Then $\lambda \in \mathscr{L}$; furthermore, since E^1 is self-supporting, $\lambda(E^1)=1$. Since $\lambda(E) \geqq \lambda(E^1)=1$, $\lambda(K_E)=1$ by definition of K_E. Since $f=0$ on K_E, we have $0=(g,\lambda)=\lim_k (f,(1/n_k)\sum_{j=1}^{n_k} T^{*j}\delta_{x_k}) \geqq \varepsilon$, which is a contradiction.

We now complete the proof of the theorem by showing that if $\omega \notin N_1 \cup N_2 \cup N_3 \cup N_4$, then $\{X_n(\omega)\}$, which converges to the escape-proof part E^1 of some $E \in \mathscr{E}$ by virtue of ω not belonging to $N_2 \cup N_3 \cup N_4$, converges in density to the ergodic kernel K_E of E. This amounts to showing that for each open set V containing K_E, $\lim_n (1/n)\sum_{k=1}^n 1_{V^c}(X_k(\omega))=0$ for all such ω. Let $\varepsilon>0$. Pick $f \in C(S)$, with $f \geqq 0$, $f=1$ on V^c, $f=0$ in K_E. By virtue of what we showed in the preceding paragraph, there is an m with $(1/m)\sum_{k=1}^m T^k f \leqq \varepsilon$ on E^1. Since $X_n(\omega) \to E^1$,

$$\limsup_n (1/m)\sum_{j=1}^m (T^j f)(X_n(\omega)) \leqq \varepsilon,$$

so

$$\limsup_n \frac{1}{n}\sum_{k=1}^n \left(\frac{1}{m}\sum_{j=1}^m (T^j f)(X_n(\omega))\right) \leqq \varepsilon. \tag{2}$$

Because $\omega \notin N_1$, (1) holds, but (1) and (2) imply that

$$\limsup_n (1/n)\sum_{k=1}^n f(X_{k+1}(\omega)) \leqq 2\varepsilon.$$

Since $1_{V^c} \leqq f$, $\limsup_n (1/n)\sum_{k=1}^n 1_{V^c}(X_k(\omega)) \leqq 2\varepsilon$. But ε is arbitrary, so this completes the proof of the theorem.

References

1. Beboutov, M.: Markov chains with a compact state space. Rec. Math. (Mat. Sb.) **10** (52), 213–238 (1942)
2. Breiman, L.: The strong law of large numbers for a class of Markov chains. Ann. Math. Statist. **31**, 801–803 (1960)
3. Chung, K. L.: The general theory of Markov processes according to Doeblin. Z. Wahrscheinlichkeitstheorie verw. Gebiete **2**, 230–254 (1964)
4. Foguel, S. R.: Ergodic decompositions of a topological space. Israel J. Math. **7**, 164–167 (1969)
5. Jamison, B.: Ergodic decomposition induced by certain Markov operators. Trans. Amer. Math. Soc. **117**, 451–468 (1965)
6. Lloyd, S. P.: On certain projections in spaces of continuous functions. Pacific J. Math. **13**, 171–175 (1963)
7. Loève, M.: Probability Theory. (3rd ed.) Princeton: Van Nostrand 1963
8. Rosenblatt, M.: Equicontinuous Markov operators. Teor. Verojatnost. i Primenen. **9**, 205–222 (1964)
9. Sine, R.: Geometric theory of a single Markov operator. Pacific J. Math. **27**, 155–166 (1968)

B. Jamison
State University of New York
at Albany
Department of Mathematics
1400 Washington Avenue
Albany, New York 12222/USA

R. Sine
University of Rhode Island
Kingston, Rhode Island 02881/USA

(Received January 30, 1972)