Stationary Markov Chains and Independent Random Variables¹

M. ROSENBLATT

Introduction. The object of this paper is to obtain a necessary and sufficient condition for a stationary Markov chain $\{x_n\}$, $n = 0, \pm 1, \dots$, with an enumerable number of states to have a representation of the form

$$(1) \quad x_n = g(\alpha_n, \alpha_{n-1}, \dots),$$

where g is a Borel measurable function and $\{\alpha_n\}$, $n = 0, \pm 1, \dots$, is a process of independent uniformly distributed random variables $[0, 1]$. The condition is that $\{x_n\}$ be ergodic and have no periodic states. The proof makes use of ideas and techniques in [2].

Preliminary Remarks. Let $\{x_n\}$, $n = 0, \pm 1, \dots$, be a stationary Markov chain with an enumerable number of states which can be conveniently labeled by the integers $1, 2, \dots$. The probability structure of the chain is characterized by its one step transition probability matrix

$$P = (p_{ij}; i, j = 1, 2, \dots)$$

$$p_{ij} = P(x_{n+1} = j \mid x_n = i)$$

and its instantaneous stationary row vector of probabilities

$$p = (p_i; i = 1, 2, \dots)$$

$$p_i = P(x_n = i).$$

The instantaneous vector of probabilities is a left invariant vector of P (see [1]); that is to say,

$$pP = p.$$

Lemma 1 of [2] already indicates that a representation of the form (1) is possible only if the process $\{x_n\}$ is ergodic and has no periodic states. We have therefore to show that such a representation does exist if the process $\{x_n\}$ is ergodic and has no periodic states.

¹ Written in part while a guest of the Applied Mathematics Division, Brookhaven National Laboratory. Research supported by the Office of Naval Research.

An Auxiliary Process and a Family of Transformations. A continuous state Markov process $\{y_n\}$, $n = 0, \pm 1, \dots$, crucial in the proof of the desired result, is introduced. Let

$$(2) \quad y_n = 1 - x_n^{-1} + x_n^{-1}(1 + x_n)^{-1}u_n, \quad n = 0, \pm 1, \dots,$$

where the process $\{u_n\}$, $n = 0, \pm 1, \dots$, is a process of independent uniformly distributed random variables $[0, 1]$ and is independent of the process $\{x_n\}$. Notice that we have enlarged the probability space of the $\{x_n\}$ process by adjoining the $\{u_n\}$ process. The process $\{y_n\}$ is a stationary ergodic Markov process with transition probability density

$$(3) \quad p(y | y') = j(j+1)p_i \quad \text{if} \quad \begin{cases} 1 - \frac{1}{j} \leq y < 1 - \frac{1}{j+1} \\ 1 - \frac{1}{i} \leq y' < 1 - \frac{1}{i+1} \end{cases}$$

and instantaneous density function

$$p(y) = j(j+1)p_i \quad \text{if} \quad 1 - \frac{1}{j} \leq y < 1 - \frac{1}{j+1},$$

$i, j = 1, 2, \dots$. Note that $x_n = f(y_n)$ where

$$(4) \quad f(y) = j \quad \text{if} \quad 1 - \frac{1}{j} \leq y < 1 - \frac{1}{j+1}, \quad j = 1, 2, \dots$$

We now make a few comments about the Markov chain $\{x_n\}$ that will be important in constructing a family of functions associated with the process $\{y_n\}$. The terminology, construction and argument will be similar to that given in [2]. The set of states of $\{x_n\}$ is an irreducible closed set of persistent states [1]. Given a set of states S of $\{x_n\}$, let $N(S)$ denote the number of states in S . Call S a *contracting set of states* if there is a set of states $S' \subset S$ with $N(S') < N(S)$ such that for each $i \in S$ there is a $j \in S'$ with $p_{ij} > 0$. The set $S'' \subset S$ can be reached from $S' \subset S$ if for each $i \in S'$ there is a $j \in S''$ with $p_{ij} > 0$. S'' is *accessible from S'* if for some k there is a sequence of subsets (of S) $S_1 = S'$, $S_2, \dots$, $S_k = S''$ such that S_i can be reached from S_{i-1} , $j = 2, \dots, k$. We shall say that S is a *point contracting* if there is a set $S' \subset S$ consisting of one state that is accessible from S through subsets of S .

Given the finite collection of states $1, 2, \dots, n$, our objective is to show that there is a finite set of states S containing $1, 2, \dots, n$ that is point contracting. Since the set of states of $\{x_n\}$ is an irreducible set of persistent states, there must be two paths from state 1 to state 1 with lengths that are relatively prime. Add all the states passed through in these paths to the collection $1, 2, \dots, n$. Further add all the states in paths of positive probability and minimal length from 1 to i and i to 1, $i = 2, \dots, n$. Call the collection of states included thus far S . Any two states i, j of S can be connected by a path of positive probability

containing only states of S . Notice that the matrix of transition probabilities of the process $\{x_n\}$ as restricted to the states of S is irreducible and aperiodic, that is, some sufficiently high power of the restricted matrix has all its elements positive.

Now if $S' \subset S$, $N(S') > 1$, there is a set of states $S'' \subset S$ with $N(S'') < N(S')$ that is accessible from S' . For suppose there were not such a set S'' . Every m -tuple of states of S is accessible from every m -tuple of states of S because of the irreducible and nonperiodic character of the transition matrix when restricted to the states of S . If there were no set of states S'' with the desired property, no m -tuple with $m = N(S')$ could be contracted. Take any two distinct states $i, j \in S$. Since they are in some m -tuple and no m -tuple is contracting, there is no state $k \in S$ such that $p_{ik}, p_{jk} > 0$. This means that for any $i \in S$ there is one and only $j = j(i)$ such that $p_{ij} > 0$. But this contradicts the fact that the transition probability matrix of $\{x_n\}$ as restricted to S is irreducible and aperiodic.

Starting with $S = S_0$ contract S_0 into a set $S_1 \subset S$ with $N(S_1) < N(S)$. Let $S_2 \subset S$ be a set with $N(S_2) < N(S_1)$ that is accessible from S_1 in a minimal number of steps. Continue this construction until we end up with a set $S_k \subset S$ with $N(S_k) = 1$. This will happen in a finite number of steps, at most $N^3(S)$ (see [2]). Thus S is point contracting.

Given the states $1, \dots, n$ we have seen that there is a set of states $S^{(n)}$ containing these that is point contracting. As before contract $S_0^{(n)} = S^{(n)}$ into a set $S_1^{(n)} \subset S^{(n)}$ with $N(S_1^{(n)}) < N(S^{(n)})$. Let $S_2^{(n)} \subset S^{(n)}$ be a set of states with $N(S_2^{(n)}) < N(S_1^{(n)})$ that is accessible from $S_1^{(n)}$ in a minimal number of steps. Continue this construction until a set $S_{k(n)}^{(n)}$ with one element is obtained. Let $M_1^{(n)}, \dots, M_{a(n)}^{(n)}$ be the sets of states with

$$M_1^{(n)} = S_0^{(n)}, \quad M_2^{(n)} = S_1^{(n)}, \dots, M_{2+r_1}^{(n)} = S_2^{(n)}, \dots, \\ M_{2+r_1+r_2}^{(n)} = S_3^{(n)}, \dots, M_{a(n)}^{(n)} = S_{k(n)}^{(n)}$$

where $a(n) = 2 + r_1 + \dots + r_k$ and $M_{i+1}^{(n)}$ can be reached from $M_i^{(n)}$. In other words the $M_i^{(n)}$ are the subsets of states of $S_0^{(n)}$ that are passed through in going from $S_0^{(n)}$ to $S_{k(n)}^{(n)}$. The total number of steps in going from $M_1^{(n)}$ to $M_{a(n)}^{(n)}$ can be at most $N^3(S^{(n)})$. Given any state $i \in M_i^{(n)}$ let $l = f_i^{(n)}(i)$ be the state of lowest index in $M_{i+1}^{(n)}$ for which $p_{il} > 0$.

We now construct a family of Borel functions $T_\alpha(y')$ on $y' \in [0, 1]$, $0 \leq \alpha < 1$, with the following properties. Given any $y' \in [0, 1]$, $T_\alpha(y')$ is a one-one function of $\alpha \in [0, 1]$ onto the set of $y \in [0, 1]$ such that $p(y | y') > 0$. Let $a(y, y')$ be the function on the set $\{(y, y') | 0 \leq y' < 1, p(y | y') > 0\}$ defined by setting $a(y, y')$ equal to the value α , $0 \leq \alpha < 1$, such that $y = T_\alpha(y')$. The family T_α is to be such that $a(y, y')$ is Borel measurable and

$$(5) \quad P(a(y_{n+1}, y_n) \leq a | y_n = y') = \int_{a(y, y') \leq a} p(y | y') dy = a, \quad 0 \leq a < 1,$$

independent of y' . Notice that then $\alpha_n = a(y_n, y_{n-1})$ is independent of $y_{n-1}, y_{n-2}, \dots$ and therefore the random variables $\{\alpha_n\}$ are independent, uniformly distributed $[0, 1)$. Further $y_n = T_{\alpha_n}(y_{n-1})$ because of the one-one character of $T_\alpha(y')$ as a function of α for fixed y' . There are many such families of transformations. We shall be interested in a family convenient for our purposes. The transformations of the family $\{T_\alpha\}$ will map each interval

$$\left[1 - \frac{1}{i}, 1 - \frac{1}{i+1}\right), \quad i = 1, 2, \dots$$

into a single point. The point will, of course, generally depend on i, α . In fact we shall set up nonoverlapping intervals $(\gamma_i^{(n)}, \delta_i^{(n)})$, $0 < \gamma_i^{(n)} < \delta_i^{(n)} < 1$, with the following property. If $\gamma_i^{(n)} < \alpha < \delta_i^{(n)}$, then

$$(6) \quad T_\alpha \left[1 - \frac{1}{i}, 1 - \frac{1}{i+1}\right) \subset \left[1 - \frac{1}{f_i^{(n)}(i)}, 1 - \frac{1}{f_i^{(n)}(i)+1}\right)$$

for all $i \in M_i^{(n)}$. Let the minimal positive transition probability p_{ij} , $i, j \in S^{(n)}$, be $\Delta^{(n)}$. We shall take all the differences

$$(7) \quad \delta_i^{(n)} - \gamma_i^{(n)} < \frac{\Delta^{(n)}}{N^3(S^{(n)})2^{n+2}}.$$

It is clear that one can find a family of transformations $\{T_\alpha\}$ satisfying these conditions.

We want to show that

$$(8) \quad E | y_n - E(y_n | \alpha_n, \alpha_{n-1}, \dots, \alpha_{n-k}) | \rightarrow 0$$

as $k \rightarrow \infty$. An argument paralleling that of Theorem 2 of [2] indicates that (8) holds if and only if

$$(9) \quad w_n(y) - w_n(y') \rightarrow 0$$

in $\alpha_1, \alpha_2, \dots$ measure (the α 's independent uniformly distributed) for all y, y' in a set of $p(\cdot)$ measure one, where

$$(10) \quad w_r(y) = T_{\alpha_r}(T_{\alpha_{r-1}}(\dots(T_{\alpha_1}(y))\dots)).$$

Let $0 \leq y, y' < 1 - 1/(n+1)$. Set

$$A(m) = \{\gamma_j^{(n)} < \alpha_{m+j} < \delta_j^{(n)}; j = 1, \dots, a(n)\}.$$

By the construction of the transformations $\{T_\alpha\}$, if $(\alpha_1, \alpha_2, \dots) \in A(m)$ and $r > m + 2N^3(S^{(n)})$,

$$w_r(y) = w_r(y').$$

By the Borel-Cantelli lemma at least one of the independent events

$$A(4mN^3(S^{(n)})), \quad m = 1, 2, \dots,$$

occurs with $(\alpha_1, \alpha_2, \dots)$ measure one. Therefore

$$w_r(y) - w_r(y') \rightarrow 0$$

in $(\alpha_1, \alpha_2, \dots)$ measure for any $y, y', 0 \leq y, y' < 1 - 1/(n+1)$. Since this is true for any n , (9) holds and hence we can say that

$$y_n = g(\alpha_n, \alpha_{n-1}, \dots)$$

for some Borel measurable g . Since $x_n = f(y_n)$,

$$x_n = f(g(\alpha_n, \alpha_{n-1}, \dots)),$$

and our representation theorem is established.

REFERENCES

- [1] W. FELLER, *An Introduction to Probability Theory and Its Applications*, New York.
- [2] M. ROSENBLATT, Stationary processes as shifts of functions of independent random variables, *Journal of Mathematics and Mechanics*, 8 (1959), 665-681.

Brown University