ISSN:0128-7680
Pertanika J. Sci. & Technol. 4(1): 77-84 (1996)

The Rate of Mixing of Two-dimensional Markov Shifts

Mohd. Salmi Md. Noorani

Mathematics Department
Faculty of Mathematical Science
Universiti Kebangsaan Malaysia
43600 UKM Bangi, Selangor D.E., Malaysia

Received 28 May 1994

ABSTRAK

Takrif anjakan Markov matra dua berserta dengan sukatan Markov yang berpadanan diberi. Kemudian dipaparkan suatu syarat cukup untuk percampuran eksponen anjakan tersebut. Hasil ini mengitlakkan hasil yang diketahui umum dalam kasus matra satu.

ABSTRACT

The definitions of a two-dimensional Markov shift and the associated Markov measure are given. Then a sufficient condition for the exponential mixing of such shifts is provided. This generalizes the well-known result in the one-dimensional case.

Keywords: two-dimensional Markov shifts, Markov measure, mixing rate

INTRODUCTION

Unlike the theory of one-dimensional Markov shifts, the theory of higher dimensional analogue of these dynamical systems is filled with anomalies and difficulties (see Schmidt 1990 for a brief survey). In this paper, we make a small contribution towards a better understanding of these higher-dimensional Markov shifts by studying the rate of mixing of such shifts.

A standard result in the ergodic theory of one-dimensional Markov shifts is as follows: Let (T, X, C, μ) be a (one-dimensional) Markov shift where the Markov measure is given by some transition probability (p, P). Suppose A and B are arbitrary cylinder sets in C. Then the sequence $(\mu(A \cap T^{-n}B))_{n\geq 0}$ converges to $\mu(A)\mu(B)$ at an exponential rate as n tends to infinity, when the matrix P is irreducible and aperiodic (i.e., there exists some interger $N > 0$ such that all the entries of P^N are strictly positive). This result follows from the crucial matrix fact that when P is irreducible and aperiodic, then the sequence $(P^n(i,j))_{n\geq 0}$ converges exponentially fast to $p(j)$ as n tends to infinity, for all i,j. Note that an immediate corollary to the above result is that T is strong-mixing.

Our purpose in this short note is to generalize the aforementioned results to the case of a two-dimensional Markov shift. Observe that the dynamical system in questions consists of two commuting (invertible) measure-preserving transformations acting on the measurable space of functions from Z^2 to some fixed finite set together with the Markov measure. Here the Markov measure is defined by two commuting stochastic matrices P and Q such that they share a common stationary probability vector p (see below for details). Working analogously with the one-dimensional case, we need to look at the rate of convergence of the sequence $(P^m Q^n)(i,j)_{m,n\geq 0}$ to $p(j)$ as m, n tends to infinity, for all i,j. We show that if either P or Q is irreducible and aperiodic then the convergence rate of the aforementioned sequence is exponentially fast. This in turn implies the exponential convergence of measures on rectangle sets for the corresponding two-dimensional Markov shift (see Theorem 1). An immediate corollary to the above is that the two-dimensional Markov shift is strong-mixing.

DEFINITIONS AND RESULTS

Let Y be the finite set $\{1,2,\ldots,k\}$ equipped with the σ-algebra 2^Y. The measurable space (Y^{Z^2}, B) is defined as the space of all functions $x : Z^2 \to Y$ endowed with the product σ-algebra B. Recall that this means, B is the smallest σ-algebra such that the collection of all projection maps $\pi_F : Y^{Z^2} \to Y^F$ which is given by $\pi_F(x) = x_{|F}$ for each *finite* subset F of Z^2, is measurable. Of course, the set Y^F here is equipped with the product σ-algebra $\prod_{c \in F} 2^Y$. Given $x \in Y^{Z^2}$, then we shall write $x_{\mathbf{c}}$ for the value of the function x at $c \in Z^2$.

We shall be interested in the following subsets of Y^{Z^2}. First, let F be the set $\{\mathbf{c} = (c_1, c_2) \in Z^2 : a_t \leq c_t \leq a_t + u_t, t = 1,2\}$ for some given $a = (a_1, a_2) \in Z^2, \mathbf{u} = (u_1, u_2) \in (Z^+)^2$. Then, an (elementary) rectangle $R_{\mathbf{a},\mathbf{u}}$ is any subset of Y^{Z^2} which takes the form

$$R_{\mathbf{a},\mathbf{u}} = \{x \in Y^{Z^2} : x_{(c_1,c_2)} = i_{(c_1,c_2)}, \forall a_t \leq c_t \leq a_t + u_t, t = 1,2\}$$

for some fixed elements $i_{\mathbf{c}}$ of Y, for each $\mathbf{c} \in F$. It is clear that such subsets are measurable. Moreover, it is not difficult to see that the collection of such rectangles generates the product σ-algebra B.

We shall now move on to the notion of a Markov measure on (Y^{Z^2}, B). For this, assume that we are given two $k \times k$-matrices P and Q satisfying the following three properties:

1. P, Q are stochastic matrices such that $PQ = QP$.
2. There exists a probability vector $p = (p(1), \ldots, p(k))$ such that $pP = p$ and $pQ = p$.
3. If P^0, Q^0 denotes the 0–1 matrices which are compatible with P and Q respectively, then we require $P^0 Q^0 = Q^0 P^0$ and $P^0 Q^0$ is also a 0–1 matrix.

Let

$$R_{\mathbf{a},\mathbf{u}} = \{x \in Y^{Z^2}: x_{(c_1,c_2)} = i_{(c_1,c_2)}, \ \forall a_t \le c_t \le a_t + u_t, t = 1,2\}$$

be a rectangle, for some $\mathbf{a} = (a_1, a_2) \in Z^2$ and $\mathbf{u} = (u_1, u_2) \in (Z^+)^2$. We shall call $R_{\mathbf{a},\mathbf{u}}$ an allowable rectangle if, in addition,

$$P^0(x_{(c_1,c_2)}, x_{(c_1+1,c_2)}) = Q^0(x_{(c_1,c_2)}, x_{(c_1,c_2+1)}) = 1$$

for all $x \in R_{\mathbf{a},\mathbf{u}}$ and $a_t \le c_t \le a_t + u_t - 1, t = 1, 2$. We are now ready to define the Markov measure m on (Y^{Z^2}, B) associated with the matrices P and Q. Let $R_{\mathbf{a},\mathbf{u}} = \{x \in Y^{Z^2}: x_{(c_1,c_2)} = i_{(c_1,c_2)}, \forall a_t \le c_t \le a_t + u_t, t = 1,2\}$ be an allowable rectangle. Then the measure of $R_{\mathbf{a},\mathbf{u}}$ is taken to be

$$m(R_{\mathbf{a},\mathbf{u}}) = p(i_{(a_1,a_2)}) \prod_{e=0}^{u_1-1} P(i_{(a_1+e,a_2)}, i_{(a_1+e+1,a_2)}) \times$$
$$\prod_{f=0}^{u_2-1} Q(i_{(a_1+u_1,a_2+f)}, i_{(a_1+u_1,a_2+f+1)})$$

For non-allowable rectangles R, we take $m(R)$ to be zero. By using the Kolmogorov consistency theorem (see, e.g., Parthasarathy 1967), m extends uniquely to a probability measure on the product σ-algebra B. In analogy with the one-dimensional case, we shall call this measure m the Markov measure defined by the matrices P and Q.

We shall define the horizontal shift σ: $Y^{Z^2} \to Y^{Z^2}$ and the vertical shift τ: $Y^{Z^2} \to Y^{Z^2}$ by

$$(\sigma x)_{(c_1,c_2)} = x_{(c_1+1,c_2)} \text{ and } (\tau x)_{(c_1,c_2)} = x_{(c_1,c_2+1)}$$

for all $x \in Y^{Z^2}$ and $(c_1, c_2) \in Z^2$. Then, it is clear that σ and τ commute. Moreover, since each σ and τ preserve the measure m on the algebra A of finite disjoint union of rectangles then they are measure-preserving on the smallest σ-algebra containing A, which is, of course, the product σ-algebra B. Thus σ and τ are two commuting measure-preserving automorphisms acting on (Y^{Z^2}, B, m). We shall call the resulting (invertible) measure-preserving dynamical system $(Y^{Z^2}, B, m, \sigma, \tau)$ a (two-dimensional) Markov shift with transition probability (p, P, Q).

DISCUSSION

By working on rectangle sets, it can be shown that the assumption that P^0Q^0 is also a 0–1 matrix is needed to check consistency of the Markov measure (c.f. Kolmogorovs theorem).

A second implication of the 0–1 assumption on the matrix P^0Q^0 is that for allowable rectangles $R_{\mathbf{a},\mathbf{u}}, m(R_{\mathbf{a},\mathbf{u}})$ is also given by

$$m(R_{\mathbf{a},\mathbf{u}}) = p(i_{(a_1,a_2)}) \prod_{e'=0}^{u_2-1} \dot{P}(i_{(a_1,a_2+e')}, i_{(a_1,a_2+e'+1)}) \times \prod_{f'=0}^{u_1-1} Q(i_{(a_1+f'_1,a_2+u_2)}, i_{(a_1+f'+1,a_2+u_2)})$$

Suppose we give the set Y the discrete topology. Then the Markov measure m is supported by the subshift of finite type

$$X = \{x \in Y^{Z^2}: P^0(x_{(e,f)}, x_{(e+1,f)}) = Q^0(x_{(e,f)}, x_{(e,f+1)}) = 1, \forall e, f \in Z\}$$

when we assume that the stationary probability vector p is a strictly positive vector. Note that the fact that X is non-empty follows from the commuting assumption on P^0 and Q^0.

Using well-known methods from the theory of one-dimensional Markov shifts, we prove:

Lemma 1

Suppose P and Q are two commuting $k \times k$-stochastic matrices such that there exists some probability vector $p = (p(1), \ldots, p(k))$ satisfying $pP = p = pQ$. If either P or Q is irreducible and aperiodic then the sequence $(P^m Q^n(i, j))_{m,n \geq 0}$ converges to $p(j)$ at an exponential rate as m, n tends to infinity, for all $i, j = 1, \ldots, k$, i.e., there exists constants $C > 0$, $0 \leq \alpha, \beta \leq 1$ with $\alpha\beta < 1$ such that

$$|P^m Q^n(i, j) - p(j)| \leq C\alpha^m \beta^n \qquad \textit{for all } m, n \geq 0$$

and for all $i, j = 1, \ldots, k$.

Proof

Without loss on generality, we shall assume that P is irreducible and aperiodic. Thus by the Perron-Frobenius theorem (see, for instance, Seneta 1981), the dominant eigenvalue 1 is a simple eigenvalue for P. Now let V be the subspace $\{v \in C^k : \langle p, v \rangle = 0\}$. (Here $\langle , \rangle$ denotes the standard inner-product in C^k). Then, it is easy to see that P leaves V invariant, i.e., $PV \subset V$. Moreover, if we denote the vector $(1,1,\ldots,1)$ by $\mathbf{1}$, then each $w \in C^k$ can be uniquely written as

$$w = w - \langle p, w \rangle 1 + \langle p, w \rangle 1$$

such that $w - \langle p, w \rangle 1 \in V$ and $\langle p, w \rangle 1 \in U$, where U is the one-dimensional subspace generated by the vector $\mathbf{1}$. Thus we have

$$C^k = V \oplus U.$$

Hence, by virtue of the simplicity of the eigenvalue 1 of P, we deduce that the spectral radius of $P_{|V}$ (the restriction of P to the invariant subspace V) is strictly less than 1. The spectral radius formula then implies there exists some $0 \leq \alpha < 1$ such that

$$|| P^m_{|V} || \leq C_1 \alpha^m$$

for all $m \geq 0$ and some constant $C_1 \geq 0$. (Here the matrix norm is the usual operator norm.) Now, if Q also has 1 as a simple eigenvalue then using the same argument as above, it gives us

$$|| Q^n_{|V} || \leq C_2 \beta^n$$

for all $n \geq 0$ and some constants $0 \leq \beta < 1, C_2 > 0$. Thus, in this case, we have

$$|| P^m_{|V} Q^n_{|V} || \leq C \alpha^m \beta^n$$

for all $m, n \geq 0$ and some constant $C > 0$.

On the other hand, when the eigenvalue 1 of Q is no longer simple the spectral radius of $Q_{|V}$ may equal 1. In this case, it suffices to note that $Q^n_{|V}$ is bounded so that

$$|| Q^n_{|V} || \leq C_3$$

for all $n \geq 0$ and for some constant $C_3 > 0$. This in turn implies that

$$|| P^m_{|V} Q^n_{|V} || \leq C' \alpha^m$$

for all $m, n \geq 0$, and some constant $C' > 0$. Hence, in either case, we have

$$|| P^m_{|V} Q^n_{|V} || \leq C \alpha^m \beta^n \qquad \forall m, n \geq 0$$

and for some constants $C > 0, 0 \leq \alpha, \beta \leq 1$ with $\alpha\beta < 1$. From this last inequality we deduce that

$$|| P^m Q^n v || \leq C \alpha^m \beta^n ||v|| \qquad \forall m, n \geq 0, \forall v \in V.$$

(Here the vector norm is the usual l_1 norm.) Recall that given $w \in C^k$, then $w- < p, w > 1 \in V$. Thus, since $P^m Q^n$ is stochastic for all $m, n \geq 0$, we have

$$|| P^m Q^n w- < p, w > 1 || \leq C \alpha^m \beta^n || w - < p, w > 1 ||$$

for all $m, n \geq 0$ and all $w \in C^k$. Furthermore, by taking $w = (0, \ldots, 0, 1, 0, \ldots, 0)$, the j-th unit vector, it is easy to see that

$$|P^m Q^n(i, j) - p(j)| \leq || P^m Q^n w - < p, w > 1 ||$$

for all $m, n \geq 0$ and $i, j = 1, 2, \ldots, k$. Thus

$$|P^m Q^n(i, j) - p(j)| \leq C' \alpha^m \beta^n \qquad \forall m, n \geq 0,$$

and for all $i, j = 1, 2, \ldots, k$ where $C' = C \max\{||w - < p, w > 1\,||: w = j$–th unit vector, $j = 1, 2, \ldots, k\}$. This gives us the required result.

Observe that the essential ingredient in the above proof is the fact that either P or Q has a simple dominant eigenvalue 1 and the rest of the spectrum has a modulus strictly less than 1. Of course, a similar observation also holds in the case of a one-dimensional Markov shift. The following theorem is the main result of this note.

Theorem 1
Let $(Y^{Z^2}, B, m, \sigma, \tau)$ be a Markov shift with transition probability (p, P, Q). Suppose either P or Q is irreducible and aperiodic. Then given any rectangle $A, B \in B$, there exists an interger $N > 0$ and constants $C > 0$, $0 \leq \alpha, \beta \leq 1$ with $\alpha\beta < 1$ such that

$$|m(A \cap \sigma^{-m}\tau^{-n}B) - m(A)m(B)| \leq C\alpha^m\beta^n$$

for all intergers $m, n \geq N$.

Proof
Let $A, B \in B$ be two arbitrary rectangles. Then, by definition, there exists $\mathbf{a} = (a_1, a_2), \mathbf{b} = (b_1, b_2) \in Z^2, \mathbf{u} = (u_1, u_2), \mathbf{v} = (v_1, v_2) \in (Z^+)^2$ such that $A = R_{\mathbf{a},\mathbf{u}}$ and $B = R_{\mathbf{b},\mathbf{v}}$ where

$$R_{\mathbf{a},\mathbf{u}} = \{x \in Y^{Z^2}: x_{(c_1,c_2)} = i_{(c_1,c_2)}, \forall a_t \leq c_t \leq a_t + u_t, t = 1, 2\}$$

and

$$R_{\mathbf{b},\mathbf{v}} = x \in Y^{Z^2}: x_{(d_1,d_2)} = i'_{(d_1,d_2)}, \forall b_t \leq d_t \leq b_t + v_t, t = 1, 2\}$$

Hence

$$A \cap \sigma^{-m}\tau^{-n}B = \{x \in Y^{Z^2}: x_{(c_1,c_2)} = i_{(c_1,c_2)}, \forall a_t \leq c_t \leq a_t + u_t$$
$$\text{and } x_{(d_1+m,d_2+n)} = i'_{(d_1,d_2)}, \forall b_t \leq d_t \leq b_t + v_t, t = 1, 2\}$$

Now, let $m > a_1 + u_1 - b_1$ and $n > a_2 + u_2 - b_2$. Then, in particular, $A \cap \sigma^{-m}\tau^{-n}B$ is a finite disjoint union of elementary rectangles $R_1, R_2, \ldots, R_k$, say. If either A or B is non-allowable, then each of the R_i's are non-allowable. Thus, in this case, we have $m(A \cap \sigma^{-m}\tau^{-n}B) = \Sigma_{i=1}^{k} m(R_i) = 0$. Moreover, since $m(A)m(B)$ is also zero, then the required result holds trivially in this case. We are now left with the case when the elementary rectangles A and B are both allowable. Then, by the assumption P^0Q^0 is also a 0–1 matrix (see above), it is straightforward (but tedious) to check that

$$m(A \cap \sigma^{-m}\tau^{-n}B) = p(i_{(a_1,a_2)}) \prod_{e=0}^{u_1-1} P(i_{(a_1+e,a_2)}, i_{(a_1+e+1,a_2)}) \times$$
$$\prod_{f=0}^{u_2-1} Q(i_{(a_1+u_1,a_2+f)}, i_{(a_1+u_1,a_2+f+1)}) P^{m'} Q^{n'} (i_{(a_1+u_1,a_2+u_2)}, i'_{(b_1,b_2)}) \times$$
$$\prod_{e=0}^{v_1-1} P(i'_{(b_1+e,b_2)}, i'_{(b_1+e+1,b_2)}) \prod_{f=0}^{v_2-1} Q(i'_{(b_1+v_1,b_2+f)}, i'_{(b_1+v_1,b_2+f+1)})$$

where $m' = b_1 + m - (a_1 + u_1) > 0, n' = b_2 + n - (a_2 + u_2) > 0$. Observe that since one of P or Q is aperiodic, then the stationary probability vector p is strictly positive. Hence

$$m(A \cap \sigma^{-m}\tau^{-n}B) = \frac{m(A)m(B)}{p(i'_{(b_1,b_2)})} P^{m'} Q^{n'} (i_{(a_1+u_1,a_2+u_2)}, i'_{(b_1,b_3)}).$$

So that

$$\left| \frac{m(A \cap \sigma^{-m}\tau^{-n}B)}{m(A)m(B)} - 1 \right| = \frac{1}{p(i'_{(b_1,b_2)})} \left| P^{m'} Q^{n'} (i_{(a_1+u_1,a_2+u_2)}, i'_{(b_1,b_2)}) - p(i'_{(b_1,b_2)}) \right|$$

Thus, by combining the previous line and Lemma 1, we gather that there exists $0 \leq \alpha, \beta \leq 1$ with $\alpha\beta < 1$ such that

$$|m(A \cap \sigma^{-m}\tau^{-n}B) - m(A)m(B)| \leq C \alpha^{m'} \beta^{n'}$$

for all $m', n' > 0$ and for some constant $C > 0$. Finally, by taking $N = \max((a_1 + u_1) - b_1, (a_2 + u_2) - b_2, 1)$, we deduce that

$$|m(A \cap \sigma^{-m}\tau^{-n}B) - m(A)m(B)| \leq C' \alpha^m \beta^n$$

for all $m, n \geq N$ and some constant $C' > 0$. This then gives us the required result. □

Let T_1, T_2 be two commuting measure-preserving transformations acting on the probability space (Z, D, ν). Then the resulting dynamical system is said to be strong-mixing if

$$\lim_{m,n \to \infty} m(A \cap T_1^{-m} T_2^{-n} B) = m(A)m(B)$$

for all $A, B \in D$.

Recall that for Markov shifts, disjoint unions of rectangles form an algebra that generates the product σ-algebra. Hence, by using a standard approximation theorem (see, e.g., Walters 1981), we have an immediate corollary to Theorem 1.

Corollary 1
Let $(Y^{Z^2}, B, m, \sigma, \tau)$ *be a Markov shift with transition probability* (p, P, Q). *If either* P *or* Q *is irreducible and aperiodic then the Markov shift is strong-mixing* .

Observe that if P, say, is irreducible and aperiodic and Q is the identity matrix then we can identify the two-dimensional Markov shift with the one-dimensional Markov shift with transition probability (p, P). Thus we can retrieve the well-known mixing result for one-dimensional Markov shifts from the above corollary. As a final remark, note that all the above results can be generalized for Markov shifts of arbitrary dimensions.

ACKNOWLEDGEMENTS

We would like to express our gratitude to Prof. W. Parry for helpful discussions through-out the course of this work. We also would like thank the referee for the helpful suggestions to improve the presentation of this paper.

REFERENCES

PARTHASARATHY, K. L. 1967. *Probability Measures on Metric Spaces*. New York: Academic Press.

SCHMIDT, K. 1990. *Algebraic Ideas in Ergodic Theory*. CBMS Regional Conf. Ser. in Math., No. 76. Providence RI: Amer. Math. Soc.

SENETA, E. 1981. *Non-negative Matrices and Markov Chains*, 2nd edn. New York: Springer-Verlag.

WALTERS, P. 1981. *An Introduction to Ergodic Theory*, G.T.M. 79. Berlin: Springer-Verlag.