

ROTATIONAL REPRESENTATIONS OF STOCHASTIC MATRICES

BY JOHN HAIGH¹

University of Guelph

Let (p_{ij}) be the matrix of a recurrent Markov chain with stationary vector $\pi > 0$, and let $\{J_i\}$ be a partition of the unit circle into sets J_1, \dots, J_n , $m(J_i) = \pi_i$, where m is Lebesgue measure. Suppose f_t defines rotation through distance t . The conditions under which p_{ij} can be written as $m(f_t(J_i) \cap J_j)/m(J_i)$ for all i and j , where each J_i is the union of at most $b(n)$ arcs, have recently been examined by Steve Alpern and Joel Cohen. Cohen conjectured that $b(n) = n - 1$, and proved $b(2) = 1$. Alpern proved that Cohen's conjecture was false for n sufficiently large, and gave bounds for $b(n)$. We give a construction that shows that $b(3) = 2$, and prove that $b(n)$ is nondecreasing.

1. Introduction. Let m denote Lebesgue measure on $X = [0, 1)$, and let $f_t(x) = x + t \pmod{1}$ be the shift transformation on X . Given any partition $J = \{J_i: i = 1, 2, \dots, n, m(J_i) > 0\}$ of X , Cohen (1981) pointed out that if

$$(1) \quad p_{ij} = m(f_t(J_i) \cap J_j)/m(J_i)$$

then $P = (p_{ij})$ is the matrix of a recurrent Markov chain on $\{1, 2, \dots, n\}$, with stationary vector $\pi = (\pi_1, \dots, \pi_n)$, $\pi_i = m(J_i)$. Conversely, given a stochastic matrix P , Cohen considered whether a representation of P in the form (1) exists, when each J_i is the union of at most $b(n)$ intervals. He conjectured that, if P is irreducible, such a representation always exists, and that $b(n) = n - 1$; proved this conjecture when $n = 2$; and showed that, for any n , there was some irreducible matrix P for which $b(n) \geq n - 1$.

Alpern (1983) showed that it was possible to extend the class of matrices from irreducible to recurrent, and used an elegant construction to show that, given any recurrent P , such a representation is possible, with $t = 1/t_0$, $t_0 = \text{l.c.m.}(1, 2, \dots, n)$, and $b(n) \leq c(n)$, where

$$(2) \quad c(n) = t_0(n^2 - n + 1) \leq \exp(\beta n)$$

for some constant β . In itself, this does not disprove Cohen's conjecture, that $b(n) = n - 1$, for some more parsimonious construction might exist. But Alpern also constructed particular stochastic matrices that imply that the conjecture is false, at least for $n = 6$ and $n \geq 8$.

We shall prove that, when P is recurrent, then $b(3) = 2$. We shall also show that, in contrast to the 2×2 case, essentially different representations of P

Received June 1984; revised September 1984.

¹ Partially supported by NSERC Grant A6187.

AMS 1980 subject classifications. Primary 15A51, secondary 28A65, 60J10.

Key words and phrases. Recurrent Markov chains, ergodic theory, measure-preserving transformations, mapping the unit interval.

might exist when $n \geq 3$. Finally, we show that $b(n) \geq b(n - 1)$, and use Alpern's construction to make an explicit conjecture about $b(n)$.

2. The case $n = 3$. Given any 3×3 recurrent matrix P with stationary vector π , we have

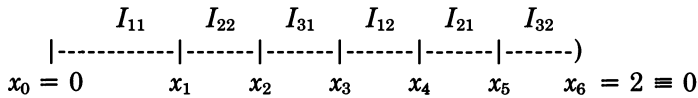
$$(3) \quad \pi_i = \pi_1 p_{1i} + \pi_2 p_{2i} + \pi_3 p_{3i} \quad i = 1, 2, 3$$

$$(4) \quad \pi_i = \pi_i p_{i1} + \pi_i p_{i2} + \pi_i p_{i3} \quad i = 1, 2, 3.$$

Equating the two expressions for π_i in (3) and (4), we have

$$(5) \quad \pi_2 p_{21} - \pi_1 p_{12} = \pi_1 p_{13} - \pi_3 p_{31} = \pi_3 p_{32} - \pi_2 p_{23} = \nu \quad (\text{say}).$$

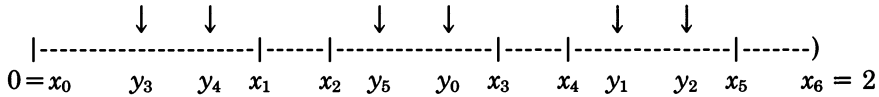
It turns out to be more convenient to define $f_t(x) = x + t \pmod{2}$, and work on the interval $[0, 2)$, so that $m(J_i) = 2\pi_i$. We now specify the division points x_1, x_2, \dots, x_5 , that partition $[0, 2)$ into the six intervals in the order shown, by fixing the lengths of these intervals.



We choose

$$(6) \quad m(I_{i1}) = 2\pi_i - \pi_i p_{ii}, \quad m(I_{i2}) = \pi_i p_{ii}$$

and we define $t = 1 - \nu$ (ν as in (5)). Then, writing $f_t(x_i) = y_i$, we can easily calculate the values of y_0, \dots, y_5 , and verify that these points are juxtaposed with x_0, \dots, x_6 as shown



(i.e., that $x_0 \leq y_3, y_4 \leq x_1$ etc.) and that

$$(7) \quad \begin{array}{ll} x_3 - y_0 = 2\pi_1 p_{13} & y_1 - x_4 = 2\pi_1 p_{12} \\ x_1 - y_4 = 2\pi_2 p_{21} & y_5 - x_2 = 2\pi_2 p_{23} \\ x_5 - y_2 = 2\pi_3 p_{32} & y_3 - x_0 = 2\pi_3 p_{31}. \end{array}$$

Thus, if $J_i = I_{i1} \cup I_{i2}$, clearly $m(J_i) = 2\pi_i$, and relations (7) show that (1) holds for $\{p_{ij}: i \neq j\}$; but, since the sets $\{J_i\}$ partition $[0, 2)$, (1) also holds for $\{p_{ii}\}$. Hence any 3×3 recurrent matrix has a representation of the form (1), with each J_i being the union of at most 2 intervals.

Cohen also asked whether such a representation, using the minimum number of intervals, was unique up to cyclic permutations of the sets $\{I_{ij}\}$. The answer is that it may not be unique. Suppose $\pi_1 p_{11} \geq |\nu|$ and $\pi_2 p_{22} \geq |\nu|$ (this is clearly

Alpern's result implies that $b(n) \geq H(n - 1)$. We offer the conjecture that $b(n) = H(n - 1)$. Some values are

n	2	3	4	5	6	7	8	9	10	11	12	13
$H(n - 1)$	1	2	3	4	6	6	12	15	20	30	30	60

The permutation matrix (p_{ij}) used to prove $b(n) \geq H(n - 1)$ is recurrent, but reducible. From it, define $Q = (q_{ij})$, where $q_{ij} = \epsilon$ when $p_{ij} = 0$, and $q_{ij} = 1 - \epsilon(n - 1)$ when $p_{ij} = 1$, where ϵ is strictly positive, but very small. Only minor changes in the argument used for (p_{ij}) are needed to show that $b(n) \geq H(n - 1)$ for the irreducible matrix Q .

We now prove that $b(n) \geq b(n - 1)$. Note that, if our conjecture that $b(n) = H(n - 1)$ is true, the table of values of $H(n - 1)$ shows that we do not have the strict inequality $b(n) > b(n - 1)$, for all values of n .

PROOF. Let P_0 be some $(n - 1) \times (n - 1)$ recurrent matrix in which any rotational representation requires some J_i to contain $b(n - 1)$ intervals. Let P_1 be the $n \times n$ matrix whose principal submatrix is P_0 , and $p_{nn} = 1$. Suppose $m(J_n) = a$, and that P_1 has a representation in which every J_i is the union of at most r intervals. Since $f_i(J_n) = J_n$, we see that the intervals in J_n can be split into families, each family consisting of equally sized intervals whose left endpoints are a multiple of $t = p/k$ apart.

Remove this J_n from $[0, 1)$, coalesce the remaining intervals, define $t = p(1 - a)/k$; this gives a representation of P_0 on $[0, 1 - a)$, using at most r intervals, so $r \geq b(n - 1)$; but $b(n) \geq r$, so $b(n) \geq b(n - 1)$. \square

Acknowledgement. I am very grateful to Murray Clayton for bringing this problem to my attention, and for allowing me to bounce my attempts to solve it against him. I also thank a referee for pointing out an oversight in a previous version of this paper.

REFERENCES

ALPERN, S. (1983). Rotational representations of stochastic matrices. *Ann. Probab.* **11** 789-794.
 COHEN, J. E. (1981). A geometric representation of a stochastic matrix: theorem and conjecture. *Ann. Probab.* **9** 899-901.

SCHOOL OF MATHEMATICAL AND PHYSICAL SCIENCES
 THE UNIVERSITY OF SUSSEX
 MATHEMATICS AND PHYSICS BUILDING
 FALMER, BRIGHTON BN1 9QH
 UNITED KINGDOM