PACIFIC JOURNAL OF MATHEMATICS
Vol. 21, No. 2, 1967

CONCERNING NONNEGATIVE MATRICES AND DOUBLY STOCHASTIC MATRICES

RICHARD SINKHORN AND PAUL KNOPP

This paper is concerned with the condition for the convergence to a doubly stochastic limit of a sequence of matrices obtained from a nonnegative matrix A by alternately scaling the rows and columns of A and with the condition for the existence of diagonal matrices D_1 and D_2 with positive main diagonals such that D_1AD_2 is doubly stochastic.

The result is the following. The sequence of matrices converges to a doubly stochastic limit if and only if the matrix A contains at least one positive diagonal. A necessary and sufficient condition that there exist diagonal matrices D_1 and D_2 with positive main diagonals such that D_1AD_2 is both doubly stochastic and the limit of the iteration is that $A \neq 0$ and each positive entry of A is contained in a positive diagonal. The form D_1AD_2 is unique, and D_1 and D_2 are unique up to a positive scalar multiple if and only if A is fully indecomposable.

Sinkhorn [6] has shown that corresponding to each positive square matrix A there is a unique doubly stochastic matrix of the form D_1AD_2 where D_1 and D_2 are diagonal matrices with positive main diagonals. The matrices D_1 and D_2 are themselves unique up to a scalar factor. The matrix D_1AD_2 can be obtained as a limit of the sequence of matrices generated by alternately normalizing the rows and columns of A. But it was shown by example that for nonnegative matrices the iteration does not always converge, and even when it does, the D_1 and D_2 do not always exist.

Marcus and Newman [4] and Maxfield and Minc [5] gave some consideration to this problem for symmetric matrices.

In a recent communication with H. Schneider, the authors learned that Brualdi, Parter and Schneider [2] have independently obtained some of the results of this paper by employing different techniques.

DEFINITIONS. If A is an $N \times N$ matrix and σ is a permutation of $\{1, \cdots, N\}$, then the sequence of elements $a_{1,\sigma(1)}, \cdots, a_{N,\sigma(N)}$ is called the diagonal of A corresponding to σ. If σ is the identity, the diagonal is called the main diagonal.

If A is a nonnegative square matrix, A is said to have total support if $A \neq 0$ and if every positive element of A lies on a positive diagonal. A nonnegative matrix that contains a positive diagonal is said to have support.

The notation $A[\mu \mid \nu]$, $A(\mu \mid \nu]$, etc. is that of [3, pp. 10–11].