Remarks on the theory of cycles in matrices

by

John S. Maybee

Purdue University

Department of Statistics

Division of Mathematical Sciences

Mimeograph Series no. 60

December 1965

Remarks on the Theory of Cycles in Matrices

John S. Maybee

Department of Mathematics

Purdue University

1. Introduction. This paper is concerned with the systematic development of the theory of cycles in matrices. We define a cycle as a product of certain distinct elements in a matrix; the precise definition is given below in section 2 (Definition 2.1). The concept is not new. The same idea has appeared both implicitly and explicitly in the literature ([3], [4]). Our contribution consists in exploiting the concept of matrix cycles and turning it into an apparently new technique for studying problems in matrix theory. The results have been somewhat surprizing.

This work has been primarily motivated by the desire to solve an old problem, namely that of finding an appropriate generalization of the well-known recurrance relation among the determinants of principal submatrices of a Jacobi matrix (see [5], [6], [2], [10]). The reason for desiring such a generalization, say to various types of so-called band matrices (or generalized Jacobi matrices), is based upon the fact that most of the theory of Jacobi matrices can be derived from the recurrance formula. As a result of more than a year's work on this problem I can say that the required generalization is quite subtle since there seem to be at least two fruitful ways of generalizing. Nevertheless the problem is solved for the most part in section 3 where, among other things, we give a formula analogues to the Jacobi recurrance formula for an arbitrary square matrix. This result is presented in theorem 3.1.

The concept of a matrix cycle and the determinant theorem just mentioned lead one in a natural way to introduce two new types of equality between (square) matrices. I have called these cyclic equality and principal equality, and the nature of cyclically and principally equal matrices is carefully investigat in sections 3 and 4. An unexpected dividend which accrues from these notions is a new method of characterizing the $n \times n$ matrices to which the entire Perron-Frobenius theory may be extended. (here the emphasis is on entire, since the basic result obviously extends to any matrix similar to an irreducible non-negative matrix.) This extension is given in section 5 where we derive in a very elementary fashion the Moroshima theorem [7] by the use of matrix cycles. We also show that a result first pointed out by Debreu and Herstein [8] is an elementary consequenc of the methods devised in sections 3 and 4.

The fundamental approach developed here can be expanded and extended to give results on other types of problems. One obvious extension would be to the theory of totally positive matrices and oscillation matrices ([5], [9], [10], [11], [12]). Another direction for further study lies in the investigation of stability of matrices, particularly to the theory of qualitative stability ([4], [13], [14]). We have already obtained interesting results in this area which will be reported upon in another place.

In order to keep the present paper to a reasonable length we have deliberately left out many concepts closely related to those presented here and, of course, many other possible applications of these concepts.

We restrict ourselves everywhere in what follows to the consideration of square matrices of order greater than or equal to three. This is mainly in order to avoid constant repetition of exceptions. Anyone interested in the 2×2 case can easily work out the results for himself.

Finally I wish to give credit to Professor J. Quirk (Economics) and John Rice (Mathematics) of Purdue University for patiently discussing many of the ideas contained herein on countless occasions. I also wish to thank the Statistics Department of Purdue University for inviting me to include this work in their mimeograph series.

2. Notation and fundamental concepts.

A square matrix of order n will usually be written in the notation of Gantmacher [1], $A = (a_{jk})_1^n$. Most of the remainder of our notation is identical to that of Marcus and Minc [2] or in the same spirit.

Let $s_n = (1,\ldots,n)$ be the set of the first n positive integers. A set of r indices satisfying the inequalities $1 \leq i_1 < i_2 < \ldots < i_r \leq n$ will be called an increasing multi-index of length r and written in the form $\mu = (i_1 \ldots i_r)$. $\Omega_{r,n}$ is the set of all such μ of length r in the set s_n, $0 \leq r \leq n$. (For $r = 0$ the set μ is taken to be the empty set.) The set $\Omega_{r,n}$ contains $\binom{n}{r}$ distinct elements. If $\mu \in \Omega_{r,n}$ is fixed, the set $\Omega_{p,\mu}$ denotes the set of increasing multi-indices of length $0 \leq p \leq r$ in μ. Clearly $\Omega_{p,\mu} \subset \Omega_{p,n}$.

For $\mu \in \Omega_{r,n}$, μ' denotes the complimentary set of indices.

Let $\mu \in \Omega_{r,n}$ be fixed, then $\mathcal{P}(\mu)$ will denote the set of permutations of μ and $\mathcal{C}(\mu)$ the subset of $\mathcal{P}(\mu)$ consisting of all cycles of length r. $\mathcal{C}(\mu)$ contains $(r-1)!$ distinct elements. Let τ denote an element of $\mathcal{C}(\mu)$.

We come now to the fundamental concept upon which the entire paper is built.

Definition 2.1: Let $\mu \in \Omega_{r,n}$ be fixed. The product

$$a(\mu,\tau) = a_{i_1\tau(i_1)}\, a_{i_2\tau(i_2)} \cdots a_{i_r\tau(i_r)}, \tau \in \mathcal{C}(\mu) \quad (2.1)$$

is called a cycle of length r in the matrix A.

Note that the cycles of length 1 are just the elements on the main diagonal of A. On the other hand, a cycle of length greater than 1 never contains an element on the main diagonal. $A = (a_{jk})_1^n$ contains $(r-1)!\binom{n}{r}$ distinct cycles of length r. Every cycle of length r can also be written in the form

$$a_{j_1j_2}\, a_{j_2j_3} \cdots a_{j_{r-1}j_r}\, a_{j_rj_1}. \quad (2.2)$$

Definition 2.2: $A = (a_{jk})_1^n$ and $B = (b_jk)_1^n$ are said to be cyclically equal, written $A \underset{\mathcal{C}}{=} B$, if for each $1 \leq r \leq n$, all $\mu \in \Omega_{r,n}$, and all $\tau \in \mathcal{C}(\mu)$, we have

$$a(\mu,\tau) = b(\mu,\tau) .$$

Clearly $A = B$ implies $A \underset{\mathcal{C}}{=} B$. The converse is not true.

In fact, we shall see below that we can have $A \underset{C}{=} B$ with the elements of A and B different everywhere except on the main diagonal.

To see more clearly into the meaning of cyclic equality it is convenient to introduce the principal submatrices of the matrix A. For each $0 \leq r \leq n$ and each $\mu \in \Omega_{r,n}$ we denote by $A[\mu]$ the principal submatrix of A in rows and columns μ. For $r = 0$ this array is empty and for $r = n$ it is A itself. Let $d(A[\mu])$ denote the determinant of $A[\mu]$: we write $d(A)$ for $r = n$ and set $d(A[\mu]) = 1$ if $r = 0$. Let $\sigma(A[\mu])$ denote the spectrum of $A[\mu]$ and put

$$\hat{\sigma}(A) = \bigcup_{r=1}^{n} \bigcup_{\mu \in \Omega_{r,n}} \sigma(A[\mu]). \qquad (2.3)$$

We shall call $\hat{\sigma}(A)$ the total spectrum of A.

Lemma 2.1: If $A \underset{C}{=} B$ then for every $0 \leq r \leq n$ and every $\mu \in \Omega_{r,n}$, $d(A[\mu]) = d(B[\mu])$.

Proof: The proof consists in the observation that every permutation is the product of cycles.

It should be remarked that simple examples can be found for which $A \underset{C}{=} B$ and corresponding non-principal minors are not equal.

In view of the fact that A has $\binom{n}{r}$ principal minors or order r and $(r-1)!\ \binom{n}{r}$ cycles of length r, it seems unlikely that the converse of lemma 2.1 is true. We shall prove later that it is in fact false and establish the precise relationships between cycles and principal minors. Since this is so, there is room for another type of equality.

Definition 2.3: $A = (a_{jk})_1^n$ and $B = (b_{jk})_1^n$ are said to be principally equal, written $A \underset{\mathcal{P}}{=} B$, if for each $0 \leq r \leq n$ and all $\mu \in \Omega_{r,n}$,

$$d(A[\mu]) = d(B[\mu]).$$

For future reference we find it convenient to record here

Lemma 2.2: If $A \underset{\wp}{\equiv} B$ or if $A \underset{\mathcal{P}}{=} B$, $\hat{\sigma}(A) = \hat{\sigma}(B)$.

It will also be useful to recall that the matrix $A = (a_{jk})_1^n$ is irreducible if for each pair of indices j, (j,k) there is a chain

$$a_{jj_1}, a_{j_1j_2}, \ldots, a_{j_{m-1}j_m}, a_{j_mk}$$

with the property that

$$a_{jj_1}a_{j_1j_2} \cdots a_{j_mk} \neq 0.$$

3. Some combinatorial results.

In most of our work we shall assume that the elements of matrix A belong to the real or complex field. The concept, of a cycle in a matrix is essentially combinatorial in character, however, and we exploit this fact in the present section. Accordingly, the elements of A may belong to any field, unless otherwise indicated.

Our first result is an apparently new formula for the expansion of the determinant of a matrix of order n. I call this formula the expansion of a determinant by principal minors. It is a natural generalization of the well-known recurrance formula for the expansion of the determinant of a Jacobi matrix.

Theorem 3.1: Let $A = (a_{jk})_1^n$ and let $\mu \in \Omega_{n-1,n}$ be fixed. Then

$$d(A) = a_{\mu',\mu'}\, d(A[\mu]) + \sum_{r=0}^{n-2} (-1)^{n+1-r} \sum_{\nu \in \Omega_{r,\mu}} d(A[\nu]) \sum_{\tau \in \mathcal{C}(\nu')} a(\nu', \tau). \qquad (3.1)$$

In this formula ν' is computed relative to s_n rather than relative to μ.

Proof: The proof of formula (3.1) is accomplished in three steps. First we observe that each summand is a term appearing in the definition of the determinant. Secondly the formula contains the correct number of terms. In fact, let N be the number of terms in (3.1); then

$$N = (n-1)! + \sum_{r=0}^{n-2} \binom{n-1}{r} (n-r-1)!\, r!$$

$$= (n-1)! + \sum_{r=0}^{n-2} \frac{(n-1)!}{r!\,(n-r-1)!} (n-r-1)!\, r! = n!$$

Thirdly, it is only necessary to point out that no two summands are equal. This proves the theorem.

With the aid of (3.1) we can now study more closely the relationship between principal and cyclic equality of matrices. For this purpose it is convenient to define the sums

$$A_{(\mu)} = \sum_{\tau \in \mathcal{C}(\mu)} a(\mu, \tau) \qquad (3.2)$$

for each $1 \leq r \leq n$ and each $\mu \in \Omega_{r,n}$. Thus $A_{(\mu)}$ is the sum of the cycles of length r in $A[\mu]$. There are $(r-1)!$ terms in

$A_{(\mu)}$.

Using the sums $A_{(\mu)}$ our determinant formula becomes

$$d(A) = a_{\mu'\mu'}\, d(A[\mu]) + \sum_{r=0}^{n-2} (-1)^{n+1-r} \sum_{\nu \in \Omega_{r,\mu}} d(A[\nu])A_{(\nu')}. \qquad (3.3$$

Lemma 3.2: $A \underset{\mathcal{P}}{=} B$ implies that, for every $1 \leq r \leq n$ and each $\mu \in \Omega_{r,n}$, $A_{(\mu)} = B_{(\mu)}$.

Proof. The proof is by indirection on r. For $r = 1$ the statement is obviously true. Suppose the result true for $r \leq p - 1$ and all $\mu \in \Omega_{r,n}$. Let $\mu_o \in \Omega_{p,n}$. By (3.3) we have for any fixed $\mu_1 \in \Omega_{p-1,\mu_o}$.

$$d(A[\mu_o]) = a_{\mu_1'\mu_1'}\, d(A[\mu_1]) + \sum_{q=1}^{p-2} (-1)^{p+1-q} \sum_{\nu \in \Omega_{q,\mu_1}} d(A[\nu]) \sum_{\tau \in \mathcal{G}(\nu')} a(\nu',\tau) + (-1)^{p+1} A_{(\mu_o)},$$

and

$$d(B[\mu_o]) = b_{\mu_1'\mu_1'}\, d(B[\mu_1]) \sum_{q=1}^{p-2} (-1)^{p-q+1} \sum_{\nu \in \Omega_{q,\mu_1}} d(B[\nu]) \sum_{\tau \in \mathcal{G}(\nu')} b(\nu',\tau) + (-1)^{p+1} B_{(\mu_o)}.$$

In these formulas the complements ν' are taken relative to μ_o so that $\sum_{\tau \in \mathcal{G}(\nu')} a(\nu',\tau) = A[\mu_o]_{(\nu')}$ and

$\sum_{\tau \in \mathcal{C}(\nu')} b(\nu',\tau) = B[\mu_o]_{(\nu')}$. By the inductive hypotheses these numbers are equal for A and B. Since $d(A[\mu_o]) = d(B[\mu_o])$, it follows that $A_{(\mu_o)} = B_{(\mu_o)}$, as was to be shown.

Definition 3.1: We say $A = (a_{jk})_1^n$ is of cyclic order p ($< n$) if every cycle of A of length greater than p is zero.

In another place [15] we have denoted the class of matrices of cyclic order two by the symbol $\mathcal{T}_2$. The best known subclass of $\mathcal{T}_2$ is the class of Jacobi (tridiagonal) matrices. The class $\mathcal{T}_2$ has many special properties. Lemma 3.2 makes it clear why this is the case. We state the result formally in

Lemma 3.3: $A \underset{\mathcal{C}}{=} B$ is equivalent to $A \underset{\mathcal{P}}{=} B$ if A and B are of cyclic order 2.

We proceed now to study other combinatorial results based upon the concept of cyclic order.

Lemma 3.4: If A is of cyclic order $p \geq 1$ there are at least $\frac{1}{2}[n(n-1) - p(p-1)]$ zeroes in A off the main diagonal.

Proof: Each element a_{jk} with $j \neq k$ appears in precisely $(n-2)!$ n-cycles of A. It follows that we must set at least n-1 distinct elements of A off the main diagonal equal to zero in order to make all n-cycles vanish. Without loss of generality we may set $a_{12},\ldots,a_{1n}$ equal to zero. Obviously all of the nonzero (n-1)-cycles of A are then (n-1)-cycles of $A[2,\ldots,n]$ and we have reduced the problem to the case of a matrix of order n-1. The process may be continued until we arrive at $A[n-p+1,\ldots,n]$ which contains cycles of length $\leq p$. Hence the least number of zeroes off the main diagonal of A is

$$p + (p + 1) + \dots + (n-1) = \frac{1}{2}n(n-1) - \frac{1}{2}p(p-1),$$

as was to be shown.

Note that our proof also shows how to construct a matrix of cyclic order p having no more than the least number of zeroes off the main diagonal.

Matrices having ''most'' off diagonal elements equal to zero are of considerable interest in a variety of problems in computation. Of particular interest are certain types of so called band matrices. In many applications these matrices enjoy a kind of symmetry.

Definition 3.2: The matrix $a = (a_{jk})_1^n$ is called combinatorially symmetric if $a_{jk} \neq 0$ implies $a_{kj} \neq 0$.

This concept appears in a paper of Drazin and Haynesworth [3] where such matrices are called special. These authors also introduce the concept of a cycle - not by any name - but they make only limited use of cycles.

The usual definition of a band matrix (also called a generalized Jacobi matrix) requires the nonzero elements to be on a prescribed number of diagonals adjacent to the main diagonal and usually implies combinatorial symmetry. Thus one defines $A = (a_{jk})_1^n$ to be a band matrix of order $p \geq 0$ if $a_{jk} = 0$ for $|j - k| > p$. The band width of A is defined to be the number $2p + 1$. A would then have at most $pn - p\frac{(p+1)}{2}$ nonzero elements above the main diagonal and at most the same number below. Moreover, if $r \geq p + 1$, $\mu \in \Omega_{r,n}$, the submatrix $A_{[\mu]}$ would have at most $pr - p\frac{(p+1)}{2}$ nonzero elements above the main diagonal and a like number below. By reversing the point of view just outlined we arrive naturally at

Definition 3.3: The matrix $A = (a_{jk})_1^n$ will be called a generalized band matrix of order p $(< n)$ if for each $p + 1 \leq r \leq n$, and all $\mu \in \Omega_{r,n}$, the submatrix $A[\mu]$ has at most $pr - p\frac{(p+1)}{2}$ nonzero elements above the main diagonal and at most $pr - \frac{p(p+1)}{2}$ nonzero elements below the main diagonal.

This definition includes the diagonal matrices, $p = 0$, the Jacobi matrices, $p = 1$, and the class $\mathcal{J}_1$ of [15], $p = 1$.

Before examining generalized band matrices more closely, let us first note some ideas connected with definition 3.2.

If $a(\mu,\tau) = a_{i,\tau(i_1)}\ a_{i_2\tau(i_2)} \cdots a_{i_r\tau(i_r)}$ is a cycle of length r of A, then

$$a'(\mu,\tau) = a_{\tau(i_1,i_1}\ a_{\tau(i_2)i_2} \cdots a_{\tau(i_r)i_r}$$

is also a cycle of length r of A. We shall call $a'(\mu,\tau)$ the transpose of $a(\mu,\tau)$. Note that

$$a(\mu,\tau)\ a'(\mu,\tau) = a_{i_1\tau(i_1)}\ a_{\tau(i_1)i_1} \cdots a_{i_r\tau(i_r)}\ a_{\tau(i_r)i_r}.$$

Thus the product of transposed r-cycles in A can be written as the product of r 2-cycles of A. The following result is therefore immediate.

Lemma 3.5: $A = (a_{jk})_1^n$ is combinatorially symmetric if and only if $a(\mu,\tau) \neq 0$ implies $a'(\mu,\tau) \neq 0$ for all $2 \leq r \leq n$ and all $\mu \in \Omega_{r,n}$.

For our next result it is useful to have at hand the notion of the j-th cross of the matrix $A = (a_{jk})_1^n$. For fixed j this consists of the elements in the jth row and column of A excluding the element a_{jj}.

Lemma 3.6: Let $A = (a_{jk})_1^n$, be combinatorially symmetric and have $pn - p\frac{(p+1)}{2}$ nonzero elements above the main diagonal where p is an even integer satisfying $2 < p < n$. Then A is of cyclic order p if and only if all nonzero off diagonal elements of A are contained in $p/2$ crosses of A.

Proof: If all nonzero off diagonal elements of A are contained in $p/2$ crosses of A it is easy to see that A can have no nonzero cycles of length greater than p. This result does not require that A have a fall quota of nonzero off diagonal elements. To prove the converse suppose the assertion false, i.e., suppose all nonzero elements of A above the main diagonal lie in $p/2 + 1$ crosses. Each cycle of A of length greater than 1 intersects a given cross of A twice or not at all. Among the cycles of A intersecting the $p/2 + 1$ crosses of A containing nonzero elements there must be one different from zero, because A has a full quota of off diagonal elements.

Actually it is not essential to assume A has enough nonzero off diagonal elements to fill $p/2$ crosses in order to obtain such a result. We have not attempted here to formulate the most precise relation between even cyclic order and the location of nonzero elements. It is known, for example, that if A is of cyclic order 2 and irreducible then A is necessarily combinatorially symmetric.

In order to better illustrate the connection between the ideas we have been developing thus far, consider the following two examples.

$$A = \begin{pmatrix} x & x & x & & 0 & \\ x & x & x & x & & \\ x & x & x & x & x & \\ & x & x & x & x & x \\ & & x & x & x & x \\ 0 & & & x & x & x \end{pmatrix},$$

$$A_2 = \begin{pmatrix} x & x & x & x & x & x \\ x & x & x & x & x & x \\ x & x & x & & & \\ x & x & & x & 0 & \\ x & x & & & x & \\ x & x & 0 & & & x \end{pmatrix}.$$

The x's in these examples are intended to denote the location of nonzero elements of A. Both A_1 and A_2 are combinatorially symmetric and irreducible and both have the same number of nonzero off diagonal elements. Both are generalized band matrices of order $p = 2$. A_2 has no nonzero cycles of length greater than 4 while A_1 has nonzero cycles of all lengths as is easily seen by constructing $G(A_1)$, the directed graph of A_1. The key to the similarity between A_1 and A_2 lies in the fact that both have 35 nonzero cycles. This motivates the following definition.

Definition 3.4: By the combinatorial type of A we mean the set of integers (n_0, n_1, n_2) where n_0 is the number of nonzero elements on the main diagonal of A, n_1 is the number of nonzero elements of A above the main diagonal, and n_2 is the number of nonzero elements of A below the main diagonal.

Theorem 3.2: If A and B are irreducible and have the same combinatorial type, then A and B have the same number of non-zero cycles.

Proof. I have not been able to construct a proof of this theorem which does not make use of graph theoretic arguments. Since A and B are irreducible, their directed graphs, $G(A)$, $G(B)$, are strongly connected (see [17], [18] for the application of graph theory to matrices). Hence each path leaving a vertex of $G(A)$ or $G(B)$ can be regarded as an arc of a cycle containing the vertex in question. The result now follows readily.

Incidentally, the hypothesis that A and B be irreducible is essential as the following example shows:

$$A = \begin{pmatrix} x & 0 & X \\ 0 & x & 0 \\ x & 0 & x \end{pmatrix}, \quad B = \begin{pmatrix} x & x & 0 \\ 0 & x & 0 \\ x & 0 & x \end{pmatrix}.$$

From theorem 3.2 we obtain immediately an estimate of the number of nonzero cycles of a generalized band matrix.

Theorem 3.3. Let $A = (a_{jk})_1^n$ be a generalized band matrix of order $p (\leq [\frac{n}{2}])$, then A has at most

$$N = \sum_{r=1}^{2p} (r-1)! \binom{n}{r} \tag{3.4}$$

nonzero cycles.

4. Elementwise characterization of principal and cyclic equality.

We assume henceforth that all matrices are over the real or the complex field. Our first aim is to study cyclic equality more closely. It turns out that the set of matrices B such that $B \underset{\mathcal{C}}{=} A$ for a given matrix $A = (a_{jk})_1^n$ can be nicely characterized if A is irreducible.

Theorem 4.1. Let $A = (a_{jk})_1^n$ be irreducible and let $B = (b_{jk})_1^n$. Then $B \underset{\mathcal{C}}{=} A$ if and only if there exists a diagonal matrix D such that $d(D) \neq 0$ and

$$B = D^{-1}AD \,. \tag{4.1}$$

Proof. We suppose first that $a_{jk} \neq 0$ for all j and k. Clearly $B \underset{\mathcal{C}}{=} A$ implies $b_{jj} = a_{jj}$, $j = 1,\ldots,n$. Since

$$b_{kj}\, b_{jk} = a_{jk}\, a_{kj} \,, \quad j < k \,,$$

we may put

$$\epsilon_{jk} = \frac{b_{jk}}{a_{jk}} = \frac{a_{kj}}{b_{kj}} \,.$$

It follows that B may be written in the form

$$B = \begin{pmatrix} a_{11} & \epsilon_{12}a_{12} & \cdots & \epsilon_{1n}a_{1n} \\ \epsilon_{12}^{-1}a_{21} & a_{22} & \cdots & \epsilon_{2n}a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \epsilon_{1n}^{-1}a_{n1} & \epsilon_{2n}^{-1}a_{n2} & \cdots & a_{nn} \end{pmatrix} . \tag{4.2}$$

From this representation it is clear that the matrix E defined by

$$E = \begin{pmatrix} 1 & \epsilon_{12} & \epsilon_{13} & \cdots & \epsilon_{1n} \\ \epsilon_{12}^{-1} & 1 & \epsilon_{23} & \cdots & \epsilon_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \epsilon_{1n}^{-1} & \epsilon_{2n}^{-1} & \cdots & \cdots & 1 \end{pmatrix}$$

must be such that every cycle equals 1. E is combinatorially symmetric so that transposed cycles will yield the same relations among the ϵ_{jk}. If $i < j < k$, we have

$$\epsilon_{ij}\,\epsilon_{jk}\,\frac{1}{\epsilon_{ik}} = 1.$$

hence

$$\epsilon_{ik} = \epsilon_{ij}\,\epsilon_{jk}.$$

In particular $\epsilon_{i,i+2} = \epsilon_{i,i+1}\,\epsilon_{i+1,i+2}$ and an easy inductive argument shows that

$$\epsilon_{jk} = \epsilon_{j,j+1}\,\epsilon_{j+1,j+2} \cdots \epsilon_{k-1,k} \quad j < k-1.$$

If we now put $\epsilon_{jk} = \sigma_j/\sigma_k$ all these ralations are satisfied and it follows that

$$B = D^{-1}AD$$

where $D = \text{diag}\,(\sigma_1,\ldots,\sigma_n)$, $d(D) \neq 0$.

Now it is an easy matter to verify that every matrix B of the form (4.1) is such that $B \underset{\mathcal{C}}{=} A$ regardless of the presence of zeroes in A. Hence it only remains to prove that $B \underset{\mathcal{C}}{=} A$ implies (4.1) if the hypothesis $a_{jk} \neq 0$ is described.

Suppose, in fact, that $a_{jk} = 0$ for some $j < k$. If $a_{kj} \neq 0$ and there exists a nonzero cycle of length ≥ 3 containing a_{kj}, we must set $b_{jk} = 0$. If $a_{kj} \neq 0$ and every cycle containing a_{kj} contains a zero element of A, then A is reducible. Hence, if $a_{jk} = 0$ and $a_{kj} \neq 0$ we put $b_{jk} = 0$ and $b_{kj} = \epsilon_{jk}^{-1}$. If $a_{kj} = 0$ we must set $b_{kj} = 0$ as well as $b_{jk} = 0$ by a similar argument. If $a_{jk} = 0$ for some $j > k$ the situation is similar except that we put $b_{kj} = \epsilon_{kj}$ or zero. Hence in any case B still has the form (4.2) except that now some entries are zero.

Let $i_1, \ldots, i_m$ be any set of distinct indices such that $a_{i_1 i_2}, a_{i_2 i_3}, \ldots, a_{i_{m-1} i_m}$ is a chain of nonzero elements of A. If any cycle containing these elements is different from zero we obtain as before the consistancy relationships

$$\epsilon_{i, i_m} = \epsilon_{i_1 i_2} \cdots \epsilon_{i_{m-1} i_m} .$$

Every such relationship is satisfied by setting $\epsilon_{jk} = \sigma_j / \sigma_k$. Since A is irreducible, at least $n - 1$ ϵ_{jk} differ from zero, and the matrix D is determined up to the condition $d(D) \neq 0$. This completes the proof of the theorem.

As may be expected, the characterization of the set of matrices B such that $B \underset{\mathcal{P}}{=} A$, A a given matrix is not quite so simple. Let us assume first, as in the proof of theorem 4.1, that $a_{jk} \neq 0$. One finds again that B must have the form given in (4.2). Now let $i < j < k$ define a principal submatrix of order three. We must then have

$$\epsilon_{ij} a_{ij} \epsilon_{jk} a_{jk} \epsilon_{ik}^{-1} a_{ki} + \epsilon_{ij}^{-1} a_{ji} \epsilon_{jk}^{-1} a_{kj} \epsilon_{ik} a_{ik}$$

$$= a_{ij} a_{jk} a_{ki} + a_{ji} a_{kj} a_{ik}.$$

This yields the relation

$$(\epsilon_{ij}\,\epsilon_{jk} - \epsilon_{ik})[\epsilon_{ik}^{-1}\,a_{ij}\,a_{jk}\,a_{kj} - \epsilon_{ij}^{-1}\,\epsilon_{jk}^{-1}\,a_{ji}\,a_{kj}\,a_{ik}] = 0 \quad (4.3)$$

If the transposed three-cycles appearing in (4.3) are equal, then necessarily

$$\epsilon_{ij}\,\epsilon_{jk} = \epsilon_{ik} \quad (4.4)$$

In the contrary case either (4.4) holds or else

$$\epsilon_{ik} = \epsilon_{ij}\,\epsilon_{jk}(a_{ij}\,a_{jk}\,a_{ki})(a_{ji}\,a_{kj}\,a_{ik})^{-1} .$$

It will be most convenient to introduce the choice function

$$\{(a_{ij}\,a_{jk}\,a_{ki})(a_{ji}\,a_{ki}\,a_{ik})^{-1}\} = \begin{cases} 1 \\ \text{or } (a_{ij}\,a_{jk}\,a_{ki})(a_{ji}\,a_{kj}\,a_{ik})^{-1}. \end{cases} \quad (4.5)$$

It is to be understood that the value 1 is to be assigned if either three-cycle is zero or if both are zero, i.e., there is no choice in this case.

One may prove by an inductive argument that the elements b_{jk}, $j < k-1$, of B have the form

$$b_{jk} = \epsilon_{jj+1} \cdots \epsilon_{k-1,k}\{(a_{jj+1}\,a_{j+1j+2}\,a_{j+2j})(a_{j+1j}\,a_{j+2j+1}\,a_{j+1j+2})^{-1}\}\times$$

$$\times \cdots \{(a_{jk-1}\,a_{k-1k}\,a_{kj})(a_{k-1j}\,a_{kk-1}\,a_{jk})^{-1}\}a_{jk},$$

with b_{kj} defined in terms of the reciprocals. In the case $a_{jk} \neq 0$ for all j and k, the number of choices for b_{jk} is 2^{k-j-1} if $j < k$ and each pair of transposed three-cycles is distinct.

These considerations lead us to formulate

Definition 4.1: Let $A = (a_{jk})_1^n$ be irreducible. We define the set of matrices $\{A\}$ by the requirement that $\tilde{A} \in \{A\}$, $\tilde{A} = (a_{jk})_1^n$ satisfy

(1) $\tilde{a}_{jk} = a_{jk}$, if $|j - k| \leq 1$,

(2) $\tilde{a}_{jk} = \{(a_{jj+1}\, a_{j+1j+2}\, a_{j+2j})(a_{j+1j}\, a_{j+2j+1}\, a_{jj+2})^{-1}\}$ x

x... $\{(a_{jk-1}\, a_{k-1k}\, a_{kj})(a_{k-1j}\, a_{kk-1}\, a_{jk})^{-1}\}\, a_{jk}$,

if $j < k-1$, and

(3) $\tilde{a}_{kj} = \dfrac{1}{\tilde{a}_{jk}}\, a_{jk}\, a_{kj}$.

Extension of the arguments used to prove thoerem 4.1 now permit us to prove

Theorem 4.2: Let $A = (a_{jk})_1^n$ be irreducible and let $B = (b_{jk})_1^n$. Then $B \underset{\mathcal{P}}{=} A$ if and only if there exists a diagonal matrix D and a matrix $\tilde{A}$ in $\{A\}$ such that $d(D) \neq 0$ and

$$B = D^{-1}\tilde{A}D \qquad (4.6)$$

We note that the set $\{A\}$ is not empty since it contains A. If $a_{ij}\, a_{jk}\, a_{ki} = a_{ji}\, a_{kj}\, a_{ik}$ for every triple $i < j < k$, or if A is of cyclic order 2. Then $\{A\}$ consists of the single matrix A.

Here is a consequence of theorem 4.2 which generalizes the concept of a symmetric matrix.

Theorem 4.3: Let $A = (a_{jk})_1^n$ be real and irreducible and suppose

(1') $\operatorname{sgn} a_{jk} = \operatorname{sgn} a_{kj}$, $j \neq k$,

(2') $a_{ij}\, a_{jk}\, a_{ki} = a_{ji}\, a_{kj}\, a_{ik}$ for every triple $i < j < k$.

Then $\hat{\sigma}(A)$ is a subset of the real line.

Proof: By (4.6) and the remark above, every matrix having the form $B = D^{-1}AD$ has the same total spectrum as A. It will suffice to establish the theorem in the case where all $a_{jk} \neq 0$. for $j \neq k$.

A continuity argument will then yield the result in the general case.

Let us put $\epsilon_{jj+1} = (a_{j+1j}/a_{jj+1})^{1/2}$, $j = 1,\ldots,n-1$. Then

$$\epsilon_{jk} = [(a_{j+1j} \cdots a_{kk-1})(a_{jj+1} \cdots a_{k-1k})^{-1}]^{1/2} .$$

and we shall have

$$b_{jk} = (\operatorname{sgn} a_{jk})[(a_{jj+1} \cdots a_{k-1k})^{-1}(a_{j+1j} \cdots a_{kk-1}\, a_{jk})]^{1/2}(a_{jk}a_{kj})^{1/2}$$

$$= (\operatorname{sgn} a_{jk})\,(a_{jk}\, a_{kj})^{1/2},$$

since the quotient under the first square root can be written as the product of quotients of transposed three-cycles. Similarly,

$$b_{kj} = (\operatorname{sgn} a_{kj})\,(a_{jk}\, a_{kj})^{1/2} .$$

Thus A is principally equal to the symmetric matrix B and the theorem is proved.

In order to extend this result to complex matrices we introduce the concept of a combinatorially hermitian matrix. We say $A = (a_{jk})_1^n$ is combinatorially hermitian if a_{jj} is real, $j = 1,\ldots,n$ and

$$a_{jk} = \alpha_{jk}\, z_{jk},\ j \neq k,$$

where α_{jk} is real and

$$z_{jk} = \bar{z}_{kj}.$$

Theorem 4.4: Let A be a complex, irreducible, combinatorially hermitian matrix and suppose

(1'') $$\operatorname{sgn} \alpha_{jk} = \operatorname{sgn} \alpha_{kj} ,$$

(2'') $$a_{ij}\, a_{jk}\, a_{ki} = \overline{a_{ji}\, a_{kj}\, a_{ik}} , \quad i < j < k.$$

Then $\sigma(A)$ is a subset of the real line.

5. Matrices of non-negative type.

The aim of the present section is to introduce a class of matrices enjoying the same properties as the non-negative matrices. The results of section 4 on cyclic and principal equality enable us to sort out all of the various kinds of matrices to which the Perron- Frobenius theory has been applied (all finite matrices). Moreover, the proofs are in every case considerably simplified. We shall not dwell upon the implication of our results here; we hope to return to this problem on another occasion.

The following definition is quite obviously motivated by our previous work.

Definition 5.1: $A = (a_{jk})_1^n$ will be called a Perron-Frobenius matrix if A is irreducible and $a_{jk} \geq 0$ for all j,k. A will be called a positive Perron-Frobenius matrix if A also satisfies the condition $a_{jk} \neq 0$ for $j \neq k$. A will be called a matrix of non-negative type if it is principally equal to a Perron-Frobenius matrix.

Lemma 5.1: Let A be a Perron-Frobenius matrix, then every matrix in $\{A\}$ is also a Perron-Frobenius matrix.

Proof: Every nonzero cycle of a Perron-Frobenius matrix is positiv It follows from the definition of the class $\{A\}$ that every element is a non-negative matrix. Since A is irreducible, every element of $\{A\}$ is also. This proves the lemma.

From lemma 5.1 we conclude immediately by formula (4.6) that if B is principally equal to a Perron-Frobenius matrix. it must also be cyclically equal to a Perron-Frobenius matrix. Therefore we have

Theorem 5.2: A is a matrix of non-negative type if and only if $A \underset{\rho}{=} B$ for some Perron-Frobenius matrix B.

The importance of the approach we are using rests upon two aspects of the classical Perron-Frobenius theory which we are able to preserve. These are the relation between the spectral radius of A and the spectral radius of any principal sub-matrix of A and the fact that A is invariant on a cone (actually a minehedral cone in the sense of Krein and Rutman [19]).

Let us turn now to the results of Moroshima [7].

Definition 5.2: A will be called a Moroshima matrix if A is real with $a_{jk} \neq 0$ if $j \neq k$, and

(1) $a_{jj} \geq 0$, $j = 1,\ldots,n$,

(2) $a_{jk}\, a_{kj} > 0$, $j \neq k$,

(3) $a_{ij}\, a_{jk}\, a_{ki} > 0$, $i \neq j \neq k \neq i$.

We shall say A is a matrix of positive type if A is cyclicall equal to a positive Perron-Frobenius matrix.

Theorem 5:3: A is a Moroshima matrix if and only if A is a real matrix of positive type.

Proof: If A is a real matrix of positive type it is clearly a Moroshima matrix. To prove the converse, assume the conditions (1)- (3) are satisfied. Clearly we must show every cycle of lengt greater than 1 of A is positive. Let $a(\mu,\tau)$ be a cycle of length r of A and write it in the form

$$a_{j_1j_2}\, a_{j_2j_3} \ldots a_{j_rj_1}\ , \ (j_1,\ldots,j_r) \ \text{distinct.}$$

Then

$$a(\mu,\tau) = \frac{(a_{j_1j_2}\, a_{j_2j_3}\, a_{j_3j_1})(a_{j_1j_3}\, a_{j_3j_4}\, a_{j_4j_1})\ldots(a_{j_1j_{r-1}}\, a_{j_{r-1}j_r}\, a_{j_rj_1})}{(a_{j_3j_1}\, a_{j_1j_3})(a_{j_4j_1}\, a_{j_1j_4})\ldots(a_{j_{r-1}j_1}\, a_{j_1j_{r-1}})}$$

It now follows from (2) and (3) that

$$a(u,\tau) > 0$$

and the theorem is proved.

Actually Moroshima stated his theorem without the restriction $a_{jk} \neq 0$, $j \neq k$, but his proof (which is quite different and much more complicated than the above proof) does not extend to this case. Moroshima also gave the condition

$$(3') \qquad a_{ij}\, a_{jk}\, a_{ik} > 0$$

in place of our condition (3); but these are clearly equivalent conditions when we assume $a_{jk} \neq 0$ for $j \neq k$. One can construct a counterexample to the Moroshima theorem in the form given in [7]. Nevertheless the Moroshima result can be extended to the nonpositive case by the following considerations.

Assuming for the moment that $a_{jk} \neq 0$ for $j \neq k$ and replacing condition (3) by Moroshima's condition (3') , one can show that there exists a permutation matrix P such that

$$P'AP = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \tag{5.1}$$

where A_{11} and A_{22} are square Perron-Frobenius matrices and $-A_{21}$, $-A_{12}$ are (in general rectangular) Perron-Frobenius matrices. This fact was first pointed out by Debreu and Herstein [8]. Now discard the hypothesis that $c_{jk} \neq 0$ for $j \neq k$ and assume there exists a permutative matrix P such that A satisfies (5.1). We shall call such a matrix A a DHM matrix (Debreu-Herstein-Moroshima matrix). The following theorem follows easily from our previous results.

Theorem 5.4: Let A be a real irreducible matrix. Then A is a matrix of non-negative type if and only if A is either a Perron-Frobenius matrix on a DHM matrix.

Let us consider next the complex case. There seems to be nothing comparable to either theorem 5.3 or theorem 5.4 in the literature for the case of complex matrices. Yet it is obvious that theorem 5.3 and our proof of the theorem do not in any way rest upon the hypothesis that A is a real matrix. Hence we may state

Theorem 5.3': Let A be a complex matrix satisfying the condition $a_{jk} \neq 0$ if $j \neq k$. Then A is of non-negative type (actually, positive type) if and only if A satisfies the conditions (1), (2), (3) of definition 5.2.

The analogue of theorem 5.4 for complex matrices must be obtained by starting with a Perron-Frobenius matrix B and a complex diagonal matrix $D = \text{diag}\{z_1, \ldots, z_n\}$. Setting

$$A = D^{-1}BD$$

we obtain $a_{jk} = \bar{z}_j^{-1} b_{jk} z_k$. This shows that A is combinatorially hermetian. Beyond this fact it is not trivial to give an easily recognizable form for the complex matrices of non-negative type similar to the form (5.1) for a DHM matrix.

In fact, consider the Perron-Frobenius matrix

$$B = \begin{pmatrix} + & + & 0 & + \\ + & + & + & + \\ + & 0 & + & + \\ + & 0 & + & + \end{pmatrix},$$

and let A be obtained by the diagonal matrix

$D = (+, -, +, +) + i(-, -, +, -)$:

Here, of course, the magnitudes of the elements are irrelevant. D^{-1} has the pattern given by

$$D^{-1} = (+, -, +, +) - i(-, -, +, -)$$

from which one readily obtains

$$R(A) = \begin{pmatrix} + & - & 0 & + \\ - & + & - & - \\ + & 0 & + & + \\ + & 0 & + & + \end{pmatrix} + \begin{pmatrix} + & + & 0 & + \\ + & + & - & + \\ - & 0 & + & - \\ + & 0 & - & + \end{pmatrix},$$

$$I(A) = -\begin{pmatrix} 0 & + & - & - \\ - & 0 & - & - \\ + & - & 0 & + \\ - & + & - & 0 \end{pmatrix} + \begin{pmatrix} 0 & - & + & - \\ + & 0 & - & + \\ - & - & 0 & - \\ - & - & + & 0 \end{pmatrix},$$

Each summand in $R(A)$ is a DHM-matrix, but $R(A)$ itself is not a DHM-matrix. Neither summand in $I(A)$ is a DHM-matrix.

We conclude this section with an examples of some significance for the theory of finite difference approximations tc differential equations. Let $A = (a_{jk})_1^n$, where a_{jk} has the property that

$$(-1)^{j+k} |a_{jk}| = a_{jk}. \tag{5.2}$$

(Such matrices were called sign-regular by Gantmacher and Krein in their book [5] if the matrix $|A|$ with elements $|a_{jk}|$ is totally non-negative. See also Karlin [11], [12]. The diagonal matrix

$$D = \text{diag } (d_1,\ldots,d_n),\ d_j = (-1)^{j+1}$$

has the property that

$$D^{-1} AD = |A|$$

so that A is a matrix of non-negative type.

Now it is easy to see that the standard finite difference approximations to the operator

$$\frac{d^p}{dx^p} \left(a(x) \frac{d^p u}{dx^p}\right)$$

on a finite interval with sufficiently small mesh widths all lead to a matrix of the form of A. The same is true of the approximations known to the author for elliptic partial differential equations not involving mixed partial derivatives. It is probable that a wide class of such difference operators for boundary value problems of elliptic type yield matrices of this form. In another place [16] we have introduced methods for analizing the form of such matrices and we believe these methods could be applied to obtain information regarding the above conjecture.

Bibliography:

1. F.R. Gantmacher, Matrix theory (2 volumes), Chelsea Publishing Company, New York, N.Y. (1959).
2. M. Marcus and H. Menc, A survey of matrix theory and matrix inequalities, Allyn and Bacon, Inc., Boston (1964).
3. M. P. Drazin and E.V. Haynsworth, A theorem on matrices of 0:s and 1's. Pacific Journal of Mathematics, Vol. 13, No. 2, 487-495, (1963).
4. J. Quirk and R. Ruppert, Qualitative economics and the stability of equilibrium, forthcoming in Review of Economic Studies, Jan., 1966.
5. F.R. Gantmacher and M.G. Krein, Oscillation matrices and kernels and small vibrations of dynamic systems, 2nd ed., Moskow, Gostekhizdat, 1950 (in Russian) also available in German translation.
6. A.S. Housholder, Principles of numerical analysis, McGraw-Hill Book Company, New York- Toronto-London, (1953).
7. M. Moroshima, On the laws of change of the price system in an economy which contains complementary commodities, Osaka Economic Papers, Vol. 1,101-113, 1952.
8. G. Debreu and I. Herstein, Nonnegative square matrices, Econometrica 21, 597-607, 1953.
9. S. Karlin, Total positivity, absorption probabilities and applications, Trans. Amer. Math. Soc., Vol. 111, No. 1, 33-107, (1964).
10. ___________, Determinants of eigenfunctions of Sturm-Liouville equations, Jour. D'Analyse Math., Vol. IX, 365-397, (1961)(1962).

11. ____________, Total positivity and applications, forthcoming book, Stanford University Press.

12. ________, Oscillation properties of eigenvectors of strictly totally positive matrices, Technical report No. 37, Math. Dept., Stanford University, (1964).

13. K. J. Lancaster, The scope of qualitative economics, Review of Economic Studies, XXIX, No. 2, 99-123.

14. W. M. Gorman, More scope for qualitative economics, Review of Economic Studies, XXXI, No. 1, 65-68.

15. J. S. Maybee, New generalizations of Jacobi matrices, (to appear)

16. ____________, Some structural theorems for partial difference operators, Numerische Math. 7, 66-72 (1965).

17. R. S. Varga, Matrix iterative analysis, Prentice-Hall, Englewood Cliffs, New Jersey, (1962).

18. Claude Berge, Theorie des Graphs et ses Applications, Dunod, Paris (1958).