COWLES COMMISSION DISCUSSION PAPER: MATHEMATICS NO. 411

Some Theorems on the Combinatorial Ranks of Matrices

Herbert A. Simon
Carnegie Institute of Technology

November 29, 1951

The rank of a matrix whose elements are real numbers is the order of the largest nonzero determinant contained in it. Suppose that in a matrix, $\|M\|$, it is known that certain specified elements are zero, but nothing is known of the values of the other elements. Then the rank of matrix is not determined.

Consider, however, the set of determinants contained in $\|M\|$ and suppose them to be expanded as sums of the individual terms. If, for a particular determinant, $|D|$, of this set, each term contains at least one of the known zeros of $\|M\|$ then it follows that $|D| \equiv 0$ for any prescribed values whatsoever of the unknown elements of $\|M\|$.

If, however, at least one term in $|D|$, does not contain any of the known zeros of $\|M\|$, then it is easy to select particular values for the unknown elements of $\|M\|$, such that we will have $|D| \neq 0$. In fact, if we consider the entire space of possible sets of values of the unknown elements of $\|M\|$, we will have $|D| \neq 0$ for all points in this space except a subset of measure zero, i.e., we will have $|D| \neq 0$ almost everywhere in the space.

These considerations suggest the possible interest of properties of $\|M\|$ which depend only on the specification that certain elements are zero. Knowledge of the locations of these zero elements will permit us to make assertions about the rank of $\|M\|$ which will hold almost everywhere in the space of the remaining elements. Since these properties of $\|M\|$ are purely combinatorial we may introduce the notion of the <u>combinatorial rank</u> (<u>c-rank</u>) of a matrix.

<u>Definition 1.</u> A matrix, $\|M\|$, certain of whose elements, $\{E_o\}$, are specified to be zero will be said to have <u>c-rank</u> <u>n</u> with respect to $\{E_o\}$ if n is the highest order of any determinant contained in $\|M\|$ at least one of whose developmental terms contains no member of $\{E_o\}$.

For example, the matrix:

$$\begin{Vmatrix} 0 & x & x & 0 \\ 0 & 0 & 0 & x \\ 0 & 0 & 0 & x \end{Vmatrix},$$

where the x's designate the unspecified elements, has c-rank 2.

<u>Literature.</u> The combinatorial properties of determinants with a known arrangement of zero and nonzero elements have been extensively studied by Konig [2], who was able to show that a number of theorems of Frobenius on determinants with nonnegative elements were of a purely combinatorial character. Theorem 1 of this paper, which is central to the present investigation was first published by P. Hall [1], in 1935. Recently, work by Koopmans and Rubin [3] on the identifiability of equations, has suggested an important area of application of the concepts of c-rank.

<u>Consistency of Equations.</u>

Consider a set of linear equations $Mx = m$ with matrix $\|M\|$ and augmented matrix $\|M'\| \equiv \|M\ m\|$. This system will be consistent if and only if $\|M\|$ has the same rank as $\|M'\|$. We introduce an analogous definition of c-consistency.

<u>Definition 2.</u> Consider a set of linear equations with matrix $\|M\|$ and augmented matrix $\|M'\|$ having specified zero elements, $\{E_o\}$. We will say that these equations are <u>c-consistent</u> with respect to $\{E_o\}$ if $\|M\|$ and $\|M'\|$ have the same c-rank with respect to $\{E_o\}$.

It is clear that if a set of linear equations with augmented matrix $\|M'\|$ is c-consistent, it is consistent almost everywhere in the space of the unspecified elements of $\|M'\|$.

<u>Definition 3.</u> Let $\|M\|$ be the matrix and $\|M'\|$ the augmented matrix of a system of n linear equations. We will say that the system is a <u>linear structure</u> if $\|M\|$ has as many rows as columns, and if $\|K\|$, the matrix of any subset of k equations of the system ($k \leq n$), has at least as many columns not identically zero as it has rows.

That is, we define a linear structure as a system of equations with as many equations as variables, and with at least as many variables appearing in any subset of equations as there are equations in that subset.

A theorem proved by Hall can be employed to show that every linear structure is c-consistent.

<u>Theorem 1.</u> (Hall). Let $\|M\|$ be the matrix of a linear structure with n equations in n variables. Then a set of n nonzero elements, $\{a_{ij}\}$, can be selected from $\|M\|$ such that no two elements belong to the same column or the same row.

This theorem is simply a translation into the terminology of this paper of Hall's Theorem 1. A proof will be found in König, pp. 234-5.

<u>Theorem 2.</u> Every linear structure is c-consistent.

<u>Proof:</u> The product of the elements $\{a_{ij}\}$ of Theorem 1 is a developmental term of the determinant $|M|$, of order n. This term contains no element of $\{E_o\}$. Hence $\|M\|$ has c-rank n. But $\|M'\|$ has n rows and (n+1) columns, and contains the determinant $|M|$. Hence $\|M'\|$ also has c-rank n.

The converse of Theorem 2 follows directly from Definitions 2 and 3.

<u>Theorem 2a.</u> Let $\|M\|$ be the (nxn) matrix, and $\|M'\|$ the augmented matrix of a system of linear equations. Then if the system is c-consistent and of c-rank

n, it is a linear structure.

Proof: If the system were not a linear structure, then $\|M\|$ could be partitioned as follows:

$$\|M\| = \begin{Vmatrix} A & O \\ B & D \end{Vmatrix}$$

where $\|A\|$ is a (pxm) matrix with $p > m$, and $\|O\|$ is a null matrix. But, since $\|M\|$ is of c-rank n, we can select from it a set of n nonvanishing elements, one from each row and column. Since $\|B \;\; D\|$ has only (n-p) rows, at least p of these elements must lie in $\|A\|$. But A has only m columns, a contradiction.

From Theorems 2 and 2a we see that the c-consistency of a system of equations can be established by counting equations and variables in the system and subsystems.

We now prove two additional theorems that are applicable to problems of identifiability.

Theorem 3. Let $\|M\|$ be a (p+q)x(p+r) matrix. Suppose that $\|M\|$ has c-rank (p+q), and that it can be partitioned as follows:

$$\|M\| = \begin{Vmatrix} A & O \\ B & D \end{Vmatrix},$$

where $\|A\|$ is pxp, $\|B\|$ is qxp, and $\|D\|$ is qxr. Then the c-rank of $\|A\|$ is p and of $\|D\|$ is q.

Proof: By hypothesis $\|M\|$ contains a determinant, $|D|$ of order (p+q) with at least one developmental term, T, containing no elements of $\{E_o\}$. Since T contains (p+q) elements, hence one from each row of $\|M\|$, it must contain one element from each row and each column of $\|A\|$. Therefore $\|A\|$ has c-rank p.

None of the remaining elements of T can lie in B, since there is already one element in each column of B. Hence q elements must lie in $\|D\|$, and this matrix must be of c-rank q or greater. But $\|D\|$ has exactly q rows, hence has c-rank q.

Theorem 4. Let $\|M\|$ be a (p+r)x(q+s) matrix, with $q \geq p$, $s \geq r$. Suppose that $\|M\|$ can be partitioned as follows:

$$\|M\| = \begin{Vmatrix} A & 0 \\ B & D \end{Vmatrix} ,$$

where $\|A\|$ is (pxq) and of c-rank p, $\|B\|$ is (rxq), $\|D\|$ is rxs and of c-rank r. Then the c-rank of $\|M\|$ is (p+r).

Proof: From $\|A\|$, p nonzero elements, $\{a\}$, can be selected, each from a different row and column; while from $\|D\|$, r nonzero elements $\{d\}$, can be selected, each from a different row and column. The set of elements $\{a\} \cup \{d\}$ consists of (p+r) nonzero elements from $\|M\|$, each belonging to a different row and column. Since $\|M\|$ has only (p+r) rows, its c-rank must therefore be exactly (p+r).

Application to Identifiability.

Let $\|M\|$ be a (nxm) matrix, with $m \geq n$, and let us partition $\|M\|$ as follows:

$$\|M\| = \begin{Vmatrix} a & 0 \\ B & D \end{Vmatrix}$$

where a is a (lxp) matrix ($p \leq m$) with nonzero terms.

Definition 4. The matrix $\|a \quad 0\|$ is called identifiable almost everywhere with respect to $\|M\|$ if $\|D\|$ is of c-rank (n-1).

The definition implies that for identifiability $\|a \quad 0\|$ must have at least (n-1) zeros, hence that $p \leq (m-n+1)$. In particular, if m=n, then p=1.

Theorem 5. Let $\|Q\|$ be a (m+q)x(n+q) matrix ($n \geq m$), with c-rank (m+q), partitioned as follows:

$$\|Q\| = \begin{Vmatrix} a & 0_1 & 0_2 \\ B & D & 0_3 \\ E & F & G \end{Vmatrix}$$

where $\|a\|$ is (lxp) with p nonzero elements, D is (m-1)x(n-p), G is (q)x(q);

and O_1, O_2, and O_3 are null matrices.

Finally, let $\|a \quad O_1\|$ be identifiable with respect to

$$\|M\| = \begin{Vmatrix} a & O_1 \\ B & D \end{Vmatrix} .$$

Then, $\|a \quad O_1 \quad O_2\|$ is identifiable with respect to $\|Q\|$.

Proof: Since $\|a \quad O_1 \quad O_2\|$ and hence, $\|M\|$ has c-rank m (by Theorem 4). Since $\|M\|$ has c-rank m and, by hypothesis, $\|Q\|$ has c-rank (m+q), $\|G\|$ has c-rank q (by Theorem 3, taking $\|G\|$ as the $\|A\|$ of that theorem, and $\|M\|$ as the $\|D\|$). Consider now $\|R\| = \begin{Vmatrix} D & O_3 \\ F & G \end{Vmatrix}$. By Theorem 4, since $\|G\|$ has c-rank q and $\|D\|$ has c-rank (m-1), $\|R\|$ has c-rank (q+m-1), which proves the theorem.

Theorem 6. Define $\|Q\|$, $\|M\|$, and $\|R\|$ as in Theorem 5, where a is now (1xq), D is (rxs), and G is (txu), and Q is (1+r+t)x(q+s+u); $(1+r) \leq (q+s)$, $t \leq u$. Then, if $\|a \quad O_1\|$ is unidentifiable with respect to $\|M\|$, $\|a \quad O_1 \quad O_2\|$ is unidentifiable with respect to $\|Q\|$.

Proof: Suppose, contrary to the theorem, that $\|a \quad O_1 \quad O_2\|$ is identifiable with respect to $\|Q\|$. Then $\|R\|$ has c-rank (r+t). But G has at most c-rank t, since it has only t rows; therefore D must have c-rank r. But if D has c-rank r, then $\|A \quad O_1\|$ is identifiable with respect to $\|M\|$, which contradicts the hypothesis.

In the language of the theory of identifiability, Theorem 5 shows (in the case where m=n), that if an equation is identifiable in a linear structure, its identifiability is not destroyed by imbedding the structure in a larger structure. Theorem 6 shows that if an equation is unidentifiable with respect to a system of equations, it remains unidentifiable if additional equations are adjoined to the system.

[1] P. Hall, "On Representatives of Subsets," Journal of the London Mathematical Society, 10, 26-30 (1935).

[2] Denes König, Theorie der Endlichen und Unendlichen Graphen

[3] T. C. Koopmans (ed), Statistical Inference in Dynamic Economic Models, esp. pp. 78-9, 82.