

## INCIDENCE MATRICES, INTERVAL GRAPHS AND SERIATION IN ARCHAEOLOGY

DAVID G. KENDALL

The work of Fulkerson and Gross on incidence matrices shows that the question, whether a given incidence matrix  $A$  can be so re-arranged by rows as to bring together all the 1's in each separate column, can be settled if one merely knows  $A$  through the symmetrised product  $A^T A$ . Suppose it is known that such a row re-arrangement exists; it is proved here that  $A$  can then be re-arranged in the required way if one merely knows  $A$  through the dual symmetrised product,  $AA^T$ .

Thus  $A^T A$  and  $AA^T$  contain respectively (i) information sufficient to decide on the possibility or otherwise of such a re-arrangement, and (ii) information sufficient to determine a sorting algorithm.

Implications for archaeology are briefly discussed.

For the mathematical background to this paper the reader is referred to D. R. Fulkerson and O. A. Gross [1]; for archaeological motivations he is referred to the author's papers [3] and [4], and especially to the latter. There it is pointed out that the mathematical problems of recognizing incidence matrices with the 'consecutive 1's property' (i.e., zero-one matrices permitting a re-arrangement of rows which bunches the 1's in each separate column), and of sorting the rows of such a matrix so as to bring the 1's together, have much in common with the archaeological problem of 'sequence dating' first formulated in 1899 by Flinders Petrie. What distinguishes the two is that in the archaeological situation one is concerned with zero-one matrices permitting a re-arrangement of rows which only *approximately* bunches the 1's in each separate column. (Normally the rows represent say graves, and the columns represent objects or aspects of objects which may or may not be present in a given grave. The re-arrangement of rows determines the ordinal chronology of the graves, and this in turn assigns a range of 'sequence dates' to each object or feature). Despite this important difference, however, it is maintained here that the mathematical problems can be a convenient source of tentative heuristic algorithms for use in the archaeological one.

Note that incidence matrices having the consecutive 1's property are those which *after row re-arrangement* have a single bunch of 1's (if any) in each column. It is useful also to have a name for incidence matrices which display this pattern of 0's and 1's *as they stand*, no re-arrangement of rows being necessary; we shall therefore call such a matrix a *Petrie matrix*. Theorem 2.1 of [1] can then be formulated

as follows: *if  $A$  and  $B$  have the same shape and if  $A^tA = B^tB$ , then  $A$  can be row re-arranged so as to become a Petrie matrix if and only if this is true of  $B$ .* Thus, if we are only interested in the possibility (or otherwise) of casting  $A$  into Petrie form by sorting rows, then the question can be answered even if  $A$  has been lost, provided that the symmetrised product  $A^tA$  has been recorded. In the archaeological context the corresponding question is not of primary interest. One normally knows (or at any rate believes) that some kind of chronological seriation is possible, and on the other hand one can be virtually sure that no re-arrangement of the rows will create a perfect Petrie pattern. The problem of primary interest here is not that of deciding whether we are dealing with seriated material, but rather that of constructing an acceptable seriation. (See, however, the remarks in the final paragraph below.)

A graph-theoretic algorithm for casting a 'consecutive 1's' matrix into Petrie form was given by Fulkerson and Gross in [1]. Nothing is known yet about the performance of this algorithm in the presence of perturbations (i.e., when the best row re-arrangement of  $A$  is only *approximately* Petrie). It seems likely that other procedures, sub-optimal in the mathematical problem, may have a compensating robustness which will make them preferable when we come to deal with the archaeological problem. One such sub-optimal procedure will now be described, but first we must prove a new result which complements the Fulkerson-Gross theorem.

We shall find the following definition useful; bearing in mind an earlier mathematical approach [6] to sequence dating, we shall say that a square symmetric matrix  $S$  is a *Robinson matrix* when, in going (i) to the left, or (ii) down, from any position on the main diagonal, the elements never increase. We now prove the

**THEOREM.** *Let  $A$  be a row-permutation of a Petrie matrix. Then the row-permutations which give to  $A$  the Petrie form are exactly those which, when applied simultaneously to the rows and columns of  $S = AA^t$ , give to that the Robinson form.*

*Proof.* (1°) Because  $(PA)(PA)^t = PSP^t$ , it is clear that any permutation applied to the rows of  $A$  acts simultaneously on the rows and columns of  $S$ . Thus it suffices to show under the stated hypothesis that  $A$  is Petrie, if and only if  $S$  is Robinson.

(2°) Suppose then that  $A$  is Petrie, and let us prove that  $S$  is Robinson. For *all* zero-one matrices,  $S$  necessarily has a weakly dominant main diagonal, and therefore we have only to show that neither of

$$(1) \quad s_{ij} > s_{ik} \quad \text{with} \quad j < k < i$$

and

$$(2) \quad s_{ij} < s_{ik} \quad \text{with} \quad i < j < k$$

is possible. Now (1) implies that, of those columns which have 1's in the  $i$ th row, strictly more have 1's in the  $j$ th than in the  $k$ th row, and this in turn implies the existence within  $A$  of a pattern

$$\begin{aligned} (j) & \quad \dots 1 \dots \\ (k) & \quad \dots 0 \dots \\ (i) & \quad \dots 1 \dots \end{aligned}$$

contrary to hypothesis. Similarly (2) leads to a contradiction.

(3°) Now suppose that  $S = AA^T$  is a Robinson matrix. We must show that  $A$ , if it admits a Petrie sorting, is already a Petrie matrix. First observe that if  $B$  is obtained from  $A$  by suppression of rows, then  $BB^T$  will be Robinson if  $AA^T$  is so. Now if  $A$  is not already Petrie, it must contain three rows displaying the pattern

$$\begin{aligned} & \dots 1 \dots \\ & \dots 0 \dots \\ & \dots 1 \dots \end{aligned}$$

Suppress all rows but these, and call the result  $B$ . There are eight possibilities for the columns of  $B$ :-

1	1	1	1	0	0	0	0
1	1	0	0	1	1	0	0
1	0	1	0	1	0	1	0
$a$	$b$	$c$	$d$	$e$	$f$	$g$	$h$

Suppose that each occurs the indicated number of times, and bear in mind that  $c$  is strictly positive. The lower triangle of  $BB^T$  will be

$$\begin{matrix} & \alpha & & & & & & & \\ & \beta & & \gamma & & & & & \\ & \delta & & \epsilon & & \zeta & & & \end{matrix}$$

where

$$\begin{aligned} \alpha &= a + b + c + d, \\ \beta &= a + b, \\ \gamma &= a + b + e + f \\ \delta &= a + c, \\ \epsilon &= a + e, \\ \zeta &= a + c + e + g. \end{aligned}$$

We know that  $BB^T$  is Robinson, and so we have the inequalities

$$\begin{aligned} a + c &\leq a + b \leq a + b + c + d, \\ a + c &\leq a + e \leq a + c + e + g, \\ a + b &\leq a + b + e + f, \\ a + e &\leq a + b + e + f. \end{aligned}$$

These tell us that  $0 < c \leq \min(b, e)$ , so that each of  $b$  and  $e$  is positive. Thus  $A$  contains the pattern

$$\begin{array}{ccc} (i) & 1 & 1 & 0 \\ (j) & 1 & 0 & 1 \\ (k) & 0 & 1 & 1. \end{array}$$

Now  $A$  admits a Petrie sorting. Imagine such a sorting carried out. The new row orderings must be  $(ik \text{ or } ki)j$  or  $j(ik \text{ or } ki)$ , in order to bunch the 1's in the central column in the pattern displayed above, but all four of these orderings wreck the bunching in either the first or the last column. This contradiction implies that the matrix  $A$  must already possess the Petrie property, as was required to complete the proof of the theorem.

From the theorem it follows that  $S = AA^T$  retains enough of the structure of  $A$  to permit the construction of a row-permutation which will restore the Petrie character to  $A$ , whenever such a permutation exists; that is, in view also of the Fulkerson-Gross theorem,  $A^T A$  tells us whether there is a solution, and  $AA^T$  then enables us to find it.

One way of finding a satisfactory permutation is of course to permute the rows and columns of  $S$  until that matrix acquires the Robinson form. This corollary therefore provides a justification for what is very close to Robinson's original procedure [6], although he did not envisage applying it to  $AA^T$ ; he worked with square symmetric matrices arrived at in a quite different way. The scanning of all double permutations of rows and columns is however a very tedious operation, and an impracticable one when the number of graves (i.e., of rows and columns of  $S$ ) is more than say 8 or 9. For a larger number of graves than this one could of course scan just a subset of the whole group of permutations, and a programme along these lines could be worked out using the ideas of [2].

Another method of procedure stems from the fact that  $S$  has the character of a *similarity* matrix; that is, we can think of  $s_{ij}$  as an indication of how like the  $i$ th row (grave) is to the  $j$ th row (grave), 'like' here being associated with nearness in time. This observation, combined with our knowledge that  $S$  does contain all the necessary information required for the construction of a 'petrifying' permutation,

strongly suggests that one should try the effect of using  $S$  as the similarity matrix initiating a Shepard-Kruskal multidimensional scaling (say in 2 dimensions). The details of multidimensional scaling will be found in [5]; for our present purpose it suffices to say that this procedure generates a two-dimensional representation in which the  $i$ th row (grave) is represented by a point  $P_i$  in such a way that, as far as possible,  $s_{gh} > s_{ij}$  when and only when  $P_g P_h < P_i P_j$ , the pairs  $(g, h)$  and  $(i, j)$  being distinct. When we are dealing with the mathematical problem,  $A$  being an incidence matrix with the consecutive 1's property, we should hope that the output of the multidimensional scaling would suggest a linear ordering of the points ( $P_i$ ) which corresponds to a permutation of rows restoring to  $A$  its Petrie character. When we are dealing with the archaeological problem,  $A$  now being the row-rearrangement of a matrix which is only approximately Petrie, we might expect in the same way to be led to a chronologically appropriate seriation of the graves. For a rather artificial example, see [4]; experiments with genuine data will be reported elsewhere.

Those who are familiar with multidimensional scaling will recall that that program makes no use of the diagonal elements  $s_{ii}$  of the similarity matrix. In the application suggested here this is entirely natural, because a row-and-column permutation which gives to  $S$  the Robinson form will do so whatever the diagonal elements may be, because they cannot fail to dominate. Thus the diagonal elements contain no information concerning the sorting permutations.

In view of the duality between the Fulkerson-Gross theorem and the one presented here it is natural to ask whether  $\Sigma = A^T A$  (which is also a similarity matrix, this time for *objects*) could be used to initiate a multidimensional scaling from the output of which one could decide whether  $A$  is or is not the row-rearrangement of a genuine or approximate Petrie matrix. This matter deserves careful study, but here we shall content ourselves with pointing out that the 'duality' referred to is slightly misleading. We cannot just exchange rows and columns and proceed as before, because rows represent interments which occur each at a single *point* of time, whereas columns represent objects or varieties and these persist throughout a *period* of time (during which they are 'fashionable'). Thus what is needed is a variant of the multidimensional scaling programme which will represent columns ( $i$ ) by say disks ( $D_i$ ) in such a way that the area of  $D_i \cap D_j$  is an increasing function of  $\sigma_{ij}$ . (Similarity is now associated with the extent to which the ranges of the two objects overlap.) No such programme has yet been devised. When one is available, then the complementary or pseudo-dual procedure implied by the Fulkerson-Gross theorem could be tried out.

## REFERENCES

1. D. R. Fulkerson and O. A. Gross, *Incidence matrices and interval graphs*, Pacific J. Math. **15** (1965), 835-855.
2. F. Hole and Mary Shaw, *Computer analysis of chronological seriation*, Rice Univ. Studies **53** (1967), No. 3.
3. D. G. Kendall, *A statistical approach to Flinders Petrie's sequence dating*, Bull. Intern. Statist. Inst. **40** (1963), 657-680.
4. ———, *Some problems and methods in statistical archaeology*, World Archaeology **1** (1969) (to appear)
5. J. B. Kruskal, *Multidimensional scaling*, Psychometrika **29** (1964), 1-27 and also 28-42.
6. W. S. Robinson, *A method for chronologically ordering archaeological deposits*, American Antiquity **16** (1951), 293-301.

Received January 11, 1968.

STATISTICAL LABORATORY,  
UNIVERSITY OF CAMBRIDGE, ENGLAND