Osaka Mathematical Journal
Vol. 6, No. 1, June, 1954

On a Certain Type of Matrices with an Application to Experimental Design

By Masashi OKAMOTO

Summary. In the first two sections there are stated some basic properties concerning the direct sum decomposition of matrices. They are preliminary to Section 3 which together with Section 5 constitutes the main part of the paper. There is introduced the notion of the "type D" in Section 3. Section 4 is supplemental and devoted to other results related to the preceding section. In the last section we deal with an application to the 2-way classification design with unequal number of replications. It is shown that every block and treatment comparison can be estimated if and only if the replication matrix is mixing, i.e., that the experiment does not split into more than one scheme.

1. Direct sum decomposition of matrices. Let us say that a matrix A is decomposed into the direct sum of components $A_1, A_2, \cdots, A_p$ and write

$$A \approx A_1 \dotplus A_2 \dotplus \cdots \dotplus A_p ,$$

whenever A can be transformed into the form

$$\begin{pmatrix} A_1 & & & 0 \\ & A_1 & & \\ & & \ddots & \\ 0 & & & A_p \end{pmatrix}$$

by the appropriate permutations, if necessary, between rows and between columns. A is called *mixing* when it cannot be decomposed into the direct sum of more than one component.

We shall investigate the method to decompose a matrix $A=(a_{ij})$, $i=1, \cdots, m$, $j=1, \cdots, n$. Put

$$\begin{aligned} (1.1) \qquad & M=\{1,2,\cdots,m\}, \quad N=\{1,2,\cdots,n\}, \\ & R=\{i : a_{ij} \neq 0 \quad \text{for some} \quad j \in N\}, \\ & C=\{j : a_{ij} \neq 0 \quad \text{for some} \quad i \in M\}. \end{aligned}$$

If R is null, then C is also null and $a_{ij}=0$ for every i, j, which case is trivial. Otherwise, take any element i_0 of R. Let

$$I_0=\{i_0\}, \qquad J_0=\{j: a_{i_0 j}\neq 0\},$$
$$I_k=\{i: a_{ij}\neq 0 \quad \text{for some} \quad j\in J_{k-1}\},$$
$$J_k=\{j: a_{ij}\neq 0 \quad \text{for some} \quad i\in I_k\}, \qquad k=1,2,\cdots.$$

The sequence $\{I_k, J_k\}$, $k=0,1,2,\cdots$, will be called the I, J-sequence starting from i_0 (-th row) (with respect to the matrix A). It is easily seen that

$$I_0\subset I_1\subset I_2\subset\cdots\subset R, \qquad J_0\subset J_1\subset J_2\subset\cdots\subset C.$$

Since R and C are finite sets, these two sequences cannot increase indefinitely and hence there is a subscript r such that

$$(1.2) \qquad I_r=I_{r+1}=\cdots=I \text{ (say)}, \qquad J_r=J_{r+1}=\cdots=J \text{ (say)}.$$

I, J is called the limit of the I, J-sequence starting from i_0 and is sometimes written as $I(i_0)$, $J(i_0)$. Clearly $I=R$ if and only if $J=C$, and

$$(1.3) \qquad \begin{aligned} I&=\{i: a_{ij}\neq 0 \quad \text{for some} \quad j\in J\},\\ J&=\{j: a_{ij}\neq 0 \quad \text{for some} \quad i\in I\}.\end{aligned}$$

Lemma 1. *For any $i'\in I(i_0)$ it holds that*

$$I(i')=I(i_0), \qquad J(i')=J(i_0).$$

Proof. Let $\{I_k', J_k'\}$ be the I, J-sequence starting from i'. Since $i'\in I(i_0)$, it follows from (1.3) by the mathematical induction that

$$I_k'\subset I(i_0), \quad J_k'\subset J(i_0), \qquad k=0,1,2,\cdots.$$

Hence $I(i')\subset I(i_0)$, $J(i')\subset J(i_0)$.

Next we shall prove the inverse inclusion relation. Since $i'\in I(i_0)$, there exist by means of (1.2) $i_k\in I_k$, $k=1,\cdots,r-1$, and $j_k\in J_k$, $k=0,1,\cdots,r-1$, such as

$$a_{i_k j_k}\neq 0, \quad a_{i_{k+1} j_k}\neq 0, \qquad k=0,1,\cdots,r-1,$$

where $i_r=i'$. This implies that $i_0\in I_r'\subset I(i')$. Again referring to (1.3), we have

$$I_k\subset I(i'), \quad J_k\subset J(i'), \qquad k=0,1,2,\cdots.$$

Thus $I(i_0)\subset I(i')$, $J(i_0)\subset J(i')$.

The I, J-sequence $\{I_k, J_k\}$ starting from the j_0-th column (with respect to A) is defined as follows:

$$J_0 = \{j_0\}, \quad I_0 = \{i : a_{ij_0} \neq 0\},$$
$$J_k = \{j : a_{ij} \neq 0 \quad \text{for some} \quad i \in I_{k-1}\},$$
$$I_k = \{i : a_{ij} \neq 0 \quad \text{for some} \quad j \in J_k\}, \qquad k = 1, 2, \cdots.$$

Its limit which exists as before is denoted by $I[j_0]$, $J[j_0]$. We have, quite similarly to Lemma 1,

Lemma 2. *For any $j_0 \in J(i_0)$ it holds that*

$$I[j_0] = I(i_0), \quad J[j_0] = J(i_0).$$

For an arbitrary $i_1 \in R$ let I_1, J_1 be the limit of the I, J-sequence starting from i_1. If $I_1 = R$, then it is a happy end. Otherwise, for an arbitrary $i_2 \in R - I_1$ let $I_2 = I(i_2)$ and $J_2 = J(i_2)$. Provided that I_1 and I_2 together do not exhaust R, we start again from an $i_3 \in R - (I_1 \cup I_2)$ to get I_3, J_3 and so on. Finally we shall have $I_1, \cdots, I_p$ which together exhaust R and have corresponding $J_1, \cdots, J_p$.

Lemma 3. *If $k \neq l$, then I_k and I_l are disjoint as well as J_k and J_l.*

Proof. Suppose that I_k and I_l intersect and $k < l$. Take an $i \in I_k \cap I_l$. By Lemma 1 we have $I_k = I(i) = I_l \ni i_l$ which contradicts the fact that $i_l \in R - (I_1 \cup \cdots \cup I_{l-1})$.

Thus we have

$$R = I_1 + I_2 + \cdots + I_p, \qquad (1.4)$$
$$C = J_1 + J_2 + \cdots + J_p.$$

Denoting by A_k, $k = 1, \cdots, p$, the matrix corresponding to rows I_k and columns J_k and by A_0 the (zero) matrix corresponding to rows $M - R$ and columns $N - C$ (A_0 vanishes when $R = M$ and $C = N$), we get the direct sum decomposition of A:

$$A \approx A_1 \dotplus A_2 \dotplus \cdots \dotplus A_p \dotplus A_0. \qquad (1.5)$$

It is readily seen that A_k, $k = 1, \cdots, p$, cannot be decomposed further and hence is mixing after our definition. Denoting by $r(X)$ the rank of the matrix X, we have

Lemma 4. *If (1.5) holds, then $r(A) = \sum_{k=1}^{p} r(A_k)$.*

2. Symmetric matrices. Our aim is the direct sum decomposition of matrices of type D. For that purpose we shall first consider the symmetric matrix $A=(a_{ij})$, $i, j=1, \cdots, n$. In order to let the zero-component A_0 in (1.5) vanish we assume

$$R=C=N, \tag{2.1}$$

R, C and N being defined in (1.1). This means that for any $i \in N$ there exists a $j \in N$ such that $a_{ij} \neq 0$ or in other words that every row has at least one non-zero element.

Lemma 5. *If A is symmetric, then in the decomposition* (1.4) *of R and C it holds that $I_k=J_k$ or $I_k \cap J_k=0$, $k=1, \cdots, p$.*

Proof. We shall show that $I_k \cap J_k \neq 0$ implies $I_k=J_k$. Take arbitrarily $i \in I_k \cap J_k$. Since $i \in J_k$, it follows from Lemma 2 that the I, J-sequence starting from the i-th column has the limit $I[i]=I_k$, $J[i]=J_k$. Because of the symmetry of A the I, J-sequence starting from the i-th row has the limit $I(i)=J_k$, $J(i)=I_k$. On the other hand $i \in I_k$ implies by Lemma 1 that $I(i)=I_k$. Therefore $I_k=J_k$. This proves the lemma.

For a subscript k such as $I_k \cap J_k=0$ there exists another subscript k' which satisfies

$$I_{k'}=J_k, \quad J_{k'}=I_k. \tag{2.2}$$

To see this we need only choose such k' as $I_{k'}$ and J_k intersect. Thus, redenoting if necessary the subscripts of I, J's in (1.4), we have

$$\begin{aligned} R &= I_1+\cdots+I_m+I_{m+1}+\cdots+I_{m+2l}, \\ C &= J_1+\cdots+J_m+J_{m+1}+\cdots+J_{m+2l}, \end{aligned} \tag{2.3}$$

where

$$\begin{aligned} I_k &= J_k, & k&=1, \cdots, m, \\ I_{m+2k-1}=J_{m+2k}, \quad I_{m+2k} &= J_{m+2k-1}, & k&=1, \cdots l. \end{aligned}$$

Let A_k, $k=1, \cdots, m$, be the matrix corresponding to rows I_k and columns J_k and let P_k, $k=1, \cdots, l$, be the matrix with rows I_{m+2k-1} and columns J_{m+2k-1}. Then the matrix with rows I_{m+2k} and columns J_{m+2k} is P_k', transposed matrix of P_k. Putting $A_{m+k}=\begin{pmatrix} 0 & P_k \\ P_k' & 0 \end{pmatrix}$, $k=1, \cdots, l$, we obtain the direct sum decomposition of A:

$$A \approx A_1 \dotplus \cdots \dotplus A_m \dotplus A_{m+1} \dotplus \cdots \dotplus A_{m+l}. \tag{2.4}$$

Clearly all A's in the right hand side are symmetric, A_k, $k=1, \cdots, m$, are mixing but A_{m+k}, $k=1, \cdots, l$, are not.

3. Matrices of type *D*. Various notions will be introduced here. Square matrix $A=(a_{ij})$, $i,j=1,\cdots,n$, is called of Type D whenever

(i) it is symmetric;
(ii) $a_{ij}\geq 0$ for every $i,j=1,\cdots,n$;
(iii) $a_{ii}=\sum_{j\neq i} a_{ij}>0$ for $i=1,\cdots,n$.

We postulate the condition $a_{ii}>0$ in (iii) only to exclude the zero-component in the decomposition of A as we have done by (2.1) in the preceding section and hence this restriction is not essential.

Denote by $A^*=(a^*_{ij})$ the matrix obtained by substituting zeroes in the principal diagonal of A and call it the *kernel* of A. The matrix of type D which is mixing will be called of type D_1 if its kernel is also mixing and of type D_2 if not.

Theorem 1. *Any matrix of type D is semi-difinite positive.*

Proof. Let $A=(a_{ij})$, $i,j=1,\cdots,n$, be of type D. Given any real vector $(x_1\cdots,x_n)$, we consider the quadratic form

$$\begin{aligned} Q&=\sum_i\sum_j a_{ij}x_ix_j=\sum_i\sum_{j\neq i} a_{ij}x_ix_j+\sum_i a_{ii}x_i^2\\ &=2\sum_i\sum_{j>i} a_{ij}x_ix_j+\sum_i x_i^2\sum_{j\neq i} a_{ij}=\sum_i\sum_{j>i} a_{ij}(x_i+x_j)^2\geq 0\,. \end{aligned}$$

Thus A is semi-definite.

Let A be of type D. Since the kernel A^* is symmetric and satisfies (2.1) as well as A does, it is decomposed as in (2.4):

$$A^*\approx A_1^*\dotplus\cdots\dotplus A_m^*\dotplus A_{m+1}^*\dotplus\cdots\dotplus A_{m+l}^*\,, \tag{3.1}$$

where A_k^*, $k=1,\cdots,m$, are mixing and

$$A_{m+k}^*=\begin{pmatrix}0, & P_k\\ P_k', & 0\end{pmatrix},\ P_k \text{ being mixing,}\qquad k=1,\cdots,l, \tag{3.2}$$

and all A_k^*, $k=1,\cdots,m+l$, are symmetric. Denote by A_k, $k=1,\cdots,m+l$, the matrix obtained by performing the inverse operation upon A_k^* as taking the kernel. Thus every A_k is of type D and its kernel is A_k^*. Corresponding to (3.1), we have the decomposition of A:

$$A\approx A_1\dotplus\cdots\dotplus A_m\dotplus A_{m+1}\dotplus\cdots\dotplus A_{m+l}\,. \tag{3.3}$$

Lemma 6. *A necessary and sufficient condition that a matrix A of type D be of type D_2 is that the kernel A^* satisfies*

$$A^*\approx\begin{pmatrix}0 & P\\ P' & 0\end{pmatrix},\ P \text{ being mixing.} \tag{3.4}$$

Proof. Necessity. Suppose A is of type D_2. Since A is mixing, there remains only one term in the right hand side of (3.3). Consider the corresponding relation (3.1). Because of the assumption that A^* is not mixing the unique term in the right side of (3.1) must be of the form (3.2). Hence follows (3.4).

Sufficiency. We have only to prove that A is mixing. Because the mixingness is invariant under permutations between rows and between columns, we may assume that the equality holds instead of $\approx$ in (3.4). Let r be the number of rows in P. Put $I_1 = J_1 = \{1, \cdots, r\}$ and $I_2 = J_2 = \{r+1, \cdots, n\}$. Since P is mixing, I_1, J_2 is the limit of the I, J-sequence starting from the first row with respect to A^*. Let I, J be the limit with respect to A. To $a^*_{ij} \neq 0$ corresponds $a_{ij} \neq 0$, so that $I \supset I_1$, and $J \supset J_2$.

Now by means of (1.3)

$$I = \{i : a_{ij} \neq 0 \text{ for some } J \in J\} \supset \{i : a_{ij} \neq 0 \text{ for some } j \in J_2\} \supset I_2,$$

where the last inclusion follows from the fact that $a_{ii} \neq 0$ for $i \in I_2 = J_2$. Hence $I = N$. This shows that A is mixing and completes the proof.

We can see thus that in the decomposition (3.3) $A_1, \cdots, A_m$ are of type D_1 and $A_{m+1}, \cdots, A_{m+l}$ are of type D_2.

Theorem 2. (a) *Any matrix of type D_1 is definite positive.* (b) *The rank of a matrix of type D_2 is smaller than its order by one.*

Proof. (a) Let $A = (a_{ij})$, $i, j = 1, \cdots, n$, be of type D_1. From the proof of Theorem 1

$$Q = \sum_i \sum_j a_{ij} x_i x_j = \sum_i \sum_{j > i} a_{ij} (x_i + x_j)^2 \geq 0.$$

We have to prove that $Q = 0$ implies $x_i = 0$, $i = 1, \cdots, n$. Suppose $Q = 0$. Then

$$a_{ij} \neq 0 \text{ implies } x_i + x_j = 0, \qquad i, j = 1, \cdots, n \ (i < j).$$

This is equivalent to the statement

$$\text{(3.5)} \qquad a^*_{ij} \neq 0 \text{ implies } x_i + x_j = 0, \qquad i, j = 1, \cdots, n.$$

Let $\{I_k, J_k\}$ be the I, J-sequence starting from $i = 1$ with respect to A^*. If $j \in J_0$, then $a^*_{1j} \neq 0$ and by (3.5) $x_1 + x_j = 0$. Therefore

$$x_j = -x_1 \qquad \text{for } j \in J_0.$$

For any $i \in I_1$ there exists a $j \in J_0$ such as $a^*_{ij} \neq 0$. Hence $x_i + x_j = 0$ and

$$x_i = -x_j = x_1 \qquad \text{for} \quad i \in I_1 .$$

By the induction it holds that

$$\text{(3.6)} \qquad \begin{aligned} x_i &= x_1 && \text{for} \quad i \in I_k , \\ x_i &= -x_1 && \text{for} \quad i \in J_k , \qquad k = 0, 1, 2, \cdots . \end{aligned}$$

Since A^* is mixing, there is a subscript r such as

$$\text{(3.7)} \qquad I_r = J_r = N .$$

(3.6) and (3.7) imply $x_1 = -x_1$ and $x_1 = 0$. This in turn implies with (3.6), (3.7) that $x_i = 0$ for all i.

(b) Let A be of type D_2. By Lemma 6 A^* is written as (3.4). Because of the invariance of the rank under permutations between rows and columns we may replace $\approx$ in (3.4) by the equality. Let r be the number of rows in P. Put $I = \{1, \cdots, r\}$ and $J = \{r+1, \cdots, n\}$.

The rank of A is k if and only if the equation $Q = \sum_i \sum_j a_{ij} x_i x_j = 0$ is equivalent to a set of k independent linear relations in $x_1, \cdots, x_n$. Assume $Q = 0$. As in the proof of (a) we have

$$\text{(3.8)} \qquad \begin{aligned} x_i &= x_1 && \text{for} \quad i \in I , \\ x_i &= -x_1 && \text{for} \quad i \in J , \end{aligned}$$

for I, J is the limit of the I, J-seqnence starting from $i = 1$ with respect to A^*. $n-1$ linear equations in (3.8) excluding the trivial one $x_1 = x_1$ are linearly independent.

Conversely (3.8) implies $Q = 0$. This proves the theorem.

Theorem 3. *The rank of a matrix of type D is smaller than its order by the number of components (in the direct sum decomposition) of type* D_2.

Theorem follows readily from Lemma 5 and Theorem 2.

4. Further results and generalization. Every property stated in the preceding section has its analogue with respect to another type of matrices defined as follows: matrix $A = (a_{ij})$, $i, j = 1, \cdots, n$, is called of type D' when

(i) it is symmetric;

(ii′) $a_{ij} \leq 0$ for $i, j = 1, \cdots, n$ $(i \neq j)$;

(iii′) $a_{ii} = \sum_{j \neq i} |a_{ij}| > 0$.

Definitions of the kernel A^*, type D_1', D_2' are quite the same as before. That is $A^* = (a_{ij} - a_{ii}\delta_{ij})$ and the mixing matrix of type D' is of type D_1' if its kernel is mixing and of type D_2' if not.

The analogues to Theorems 1 and 2 (b) hold unaltered but it is not the case with 2 (a), because the rank of a matrix of type D_1' is smaller than its order by one just as well as in the case of D_2'. These two types are thus dealt with together in Theorem 2'. Thereby Theorem 3' also slightly differs from the corresponding Theorem 3.

Theorem 1'. *Any matrix of type D' is semi-definite positive.*

Theorem 2'. *The rank of a matrix which is mixing and of type D' is smaller than its order by one.*

Theorem 3'. *The rank of a matrix of type D' is smaller than its order by the number of mixing components in the direct sum decomposition.*

Results stated in the preceding and the present sections can be generalized in some respects. In the first we can adopt the condition

$$\text{(iii}''\text{)} \qquad a_{ii} \geq \sum_{j \neq i} |a_{ij}| > 0$$

in place of either (iii) in the case of type D or (iii') in D'. As the second generalization we may postulate only (i) and (iii''), omitting (ii) or (ii'). While the denotation is wide enough to contain both the types D and D', the general theory will be somewhat complicated, at least in the second generalization.

5. Application to the experimental design. Some little explanation concerning the experimental design will be needed to justify the application of the results obtained above. For details references are to be made to O. Kempthorne [1]. Let the model be

$$y_\alpha = \sum_{i=1}^{p} x_{\alpha i}\beta_i + e_\alpha . \qquad \alpha = 1, \cdots, n \ (\geq p) ,$$

where $x_{\alpha i}$ are known coefficients, β_i unknown parameters and e_α are random variables independently distributed according to $N(o, \sigma^2)$, σ^2 being unknown. In the matrix notation we write

$$y = X\beta + e ,$$

where y, X, β and e are $n \times 1$, $n \times p$, $p \times 1$ and $n \times 1$ matrices, respectively. Put $S = X'X$, X' being the transposed matrix of X. A linear form in β_i which admits the best linear unbiased estimate is called estimable. The number of linearly independent estimable functions is $r(S)$, the rank of the matrix S.

The model of the 2-way classification with unequal number of replications is

$$y_{ijk} = \mu + b_i + t_j + e_{ijk},$$

$i=1,\cdots,r$; $j=1,\cdots,s$; $k=1,\cdots,n_{ij}$. Putting

$$\beta = (\mu, b_1, \cdots, b_r, t_1, \cdots, t_s)',$$

$$N_{i\cdot} = \sum_{j=1}^{s} n_{ij}, \quad N_{\cdot j} = \sum_{i=1}^{r} n_{ij}, \quad N_{\cdot\cdot} = \sum_{i=1}^{r}\sum_{j=1}^{s} n_{ij},$$

we have

$$S = \begin{pmatrix} N_{\cdot\cdot} & N_{1\cdot} \cdots N_{r\cdot} & N_{\cdot 1} \cdots N_{\cdot s} \\ N_{1\cdot} & N_{1\cdot} \quad\quad 0 & n_{11} \cdots n_{1s} \\ \vdots & \ddots & \vdots \quad\quad \vdots \\ N_{r\cdot} & 0 \quad\quad N_{r\cdot} & n_{r1} \cdots n_{rs} \\ N_{\cdot 1} & n_{11} \cdots n_{r1} & N_{\cdot 1} \quad\quad 0 \\ \vdots & \vdots \quad\quad \vdots & \ddots \\ N_{\cdot s} & n_{1s} \cdots n_{rs} & 0 \quad\quad N_{\cdot s} \end{pmatrix}$$

Let S_0 be the matrix obtained by deleting the first row and the first column of S. Since in S the first row is equal to the sum of r rows from the second to the $(r+1)$-th, we have $r(S)=r(S_0)$.

It is easily seen that S_0 is of type D. The restriction $a_{ii}>0$ in (iii) in the definition of the type D means here $N_{i\cdot}>0$, $N_{\cdot j}>0$ for all i, j, which we shall be permitted to assume from the practical meaning of the experiment. The kernel S_0^* of S_0 is $\begin{pmatrix} 0 & N \\ N' & 0 \end{pmatrix}$, where $N=(n_{ij})$, $i=1,\cdots,r$; $j=1,\cdots,s$, which will be called the replication matrix. Whenever N is decomposed into exactly m components, S_0 is decomposed into m components of type D_2 and by Theorem 3 we have $r(S_0)=r+s-m$. Therefore a necessary and sufficient condition that $r(S)=r+s-1$ is that N is mixing. In this case every block comparison $b_i-b_{i'}$ and every treatment comparison $t_j-t_{j'}$ are estimable. That the replication matrix is not mixing means that the set $\mathfrak{B}=\{B_1,\cdots,B_r\}$ of blocks and the set $\mathfrak{T}=\{T_1,\cdots,T_s\}$ of treatments split into two sets $\mathfrak{B}_1, \mathfrak{B}_2$ and $\mathfrak{T}_1, \mathfrak{T}_2$ respectively such that any variate corresponding to a $B\in\mathfrak{B}_1$ and a $T\in\mathfrak{T}_2$ or to a $B\in\mathfrak{B}_2$ and a $T\in\mathfrak{T}_1$ is not observed. To this effect our result will be plausible.

In order that N is mixing, the total number $N_{\cdot\cdot}$ of plots must be larger than or equal to $r+s-1$. This minimum is always attainable by putting, for instance,

$$n_{ij}=1, \quad \text{if } i=1 \text{ or } j=1,$$
$$n_{ij}=0, \quad \text{otherwise.}$$

In the analysis of variance, however, the degrees of freedom of the error term is $N_{..}-(r+s-1)$ and it vanishes, to our regret, under the most economical design above, which makes it unable to test the significance of any comparison. Thus, at least $r+s$ plots are required to perform the significance test. This gives the simplest type of the incomplete block design.

(Received February 26, 1954)

Reference

[1] O. Kempthorne, The Design and Analysis of Experiments, New York, 1952.