SÉMINAIRE DUBREIL.
ALGÈBRE ET THÉORIE
DES NOMBRES

WALTER LEDERMANN

Representation theory and statistics

Séminaire Dubreil. Algèbre et théorie des nombres, tome 20, nº 2 (1966-1967), exp. nº 15, p. 1-8
<http://www.numdam.org/item?id=SD_1966-1967__20_2_A3_0>

10 avril 1967

REPRESENTATION THEORY AND STATISTICS

by Walter LEDERMANN

1. Introduction.

In the statistical method, known as the analysis of variance, one is concerned with n random variables $x_1 , x_2 , \ldots , x_n$, whose joint distribution function $F(x_1 , x_2 , \ldots , x_n)$ is invariant under a permutation group G , which is a subgroup of the symmetric group S_n . The group G is the symmetry group of the design in which the variables are arranged.

For example, in the randomized block design, the variables are arranged in a p by q rectangle, where $n = pq$, and it is convenient to denote the variables by x_{ij} $(i = 1 , 2 , \ldots , p , \ j = 1 , 2 , \ldots , q)$. Thus we have the design

$$\begin{matrix} x_{11} & x_{12} & \cdots & x_{1q} \\ x_{21} & x_{22} & \cdots & x_{2q} \\ \cdots & \cdots & \cdots & \cdots \\ x_{p1} & x_{p2} & \cdots & x_{pq} \end{matrix} .$$

In some agricultural experiments, x_{ij} is the yield in the plot (i , j) . Now, p different treatments with fertilizers are tried in such a way that plots of one particular row are all given the same treatment. Furthermore, it is assumed that soil conditions are constant for all plots lying in the same column. But the treatments are chosen at random, and all distributions of soil conditions are supposed to be equally likely. Under these conditions, all permutations of the rows are permissible, and all permutations of the columns are permissible, that is

$$G = S_p \times S_q .$$

Now, for each design the analysis of variance is based on a certain algebraical identity which expresses the total estimated variance as a sum of quadratic forms of special significance. In the case of the block design this identity is as follows : let

$$\bar{x}_{i.} = \frac{1}{q} \sum_{j=1}^{q} x_{ij} , \qquad \bar{x}_{.j} = \frac{1}{p} \sum_{i=1}^{p} x_{ij} , \qquad \bar{\bar{x}} = \frac{1}{pq} \sum_{i=1}^{p} \sum_{j=1}^{q} x_{ij} .$$

Then

$$(1) \quad \sum_{i,j} x_{ij}^2 = pq\,\bar{x}^2 + p\sum_{j=1}^{q} (\bar{x}_{.j} - \bar{x})^2 + q\sum_{i=1}^{p} (\bar{x}_{i.} - \bar{x})^2 + \sum_{i=1}^{p}\sum_{j=1}^{q} (x_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{x})^2 \ldots$$

Actually, it is customary to transpose the first term on the right, and to express

$$\sum_{i,j} x_{ij}^2 - pq\,\bar{x}^2 = \sum_{i,j} (x_{ij} - \bar{x})^2$$

as the three remaining terms on the right. However, for our purposes, it is more convenient to consider the identity given in (1). It is, of course, not difficult, though a little laborious to establish (1) by elementary algebraical manipulation, but we are interested to see how this identity arises, and we shall find that it is intimately related to the irreducible representations of the underlying symmetry group.

2. Outline of the general theory.

We introduce the vector space $X = (x_1, x_2, \ldots, x_n)$ generated by the n variables, now called vectors, and we assume that X is endowed with the usual inner product. The permutation group G is made to act on X in the obvious way. Thus, if

$$g = \begin{pmatrix} 1 & 2 & \ldots & n \\ \gamma_1 & \gamma_2 & \ldots & \gamma_n \end{pmatrix}$$

is an element of G, then

$$(2) \qquad x_t\, g = x_{\gamma_t} \qquad (t = 1, 2, \ldots, n) \quad ,$$

operators being written on the right. The field of scalars will be taken to be the real field (or one of its subfields), a fact which will be of some importance later on. The linear map of X onto itself which is induced by the equations (2) will simply be denoted by $g : X \to X$, and we identify the group G with the group of maps induced by the elements $g \in G$.

If G is reducible, and hence completely reducible by Maschke's theorem, then X can be broken up into the direct sum of orthogonal subspaces, thus

$$(3) \qquad X = U_1 \oplus U_2 \oplus \ldots \oplus U_s \ldots \quad ,$$

where each U_α is invariant under each g, that is

$$(4) \qquad U_\alpha\, g = U_\alpha \qquad (\alpha = 1, 2, \ldots, s, \quad g \in G) \quad .$$

On the other hand, a decomposition of the type (3) implies that there exist projection maps $e_1, e_2, \dots, e_s$ such that

$$e_\alpha : X \to X ,$$

where $Xe_\alpha = U_\alpha$, $e_\alpha^2 = e_\alpha$, which means that each vector of U_α remains fixed under e_α and

(5) $$1 = e_1 + e_2 + \dots + e_s ,$$

where 1 is the identity map of X and

(6) $$e_\alpha e_\beta = e_\beta e_\alpha = 0 \qquad (\alpha \neq \beta) .$$

Conversely, any decomposition of 1 as a sum of orthogonal idempotents leads to a decomposition of X into subspaces. If the subspaces are invariant under G, then

(7) $$e_\alpha g = g e_\alpha \qquad (g \in G , \alpha = 1, 2, \dots, s) .$$

For applications, it will be necessary to compute the idempotent matrices E_α which correspond to the projections e_α. Generally, if $\ell : X \to X$ is a linear map of X into itself with matrix $L = (\ell_{ij})$, then

$$x_i \ell = \sum_{j=1}^{n} \ell_{ij} x_j \qquad (i = 1, 2, \dots, n) ,$$

and it is convenient to summarize this information by using the column "vector of vectors"

$$\underset{\sim}{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} ,$$

so that we can write more concisely

(8) $$\underset{\sim}{x}\ell = L\underset{\sim}{x} , \qquad L = (\ell_{ij}) .$$

Suppose now that $U = [u_1, u_2, \dots, u_m]$ is a subspace of X of dimension m, spanned by the basis $u_1, u_2, \dots, u_m$. Then

$$u_k = \sum_{j=1}^{n} a_{kj} x_j \qquad (k = 1, 2, \dots, m) ,$$

and we may say that U is given by the $m \times n$ matrix

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{m1} & \dots & a_{mn} \end{pmatrix} ,$$

which is of rank m. Since we are working over a real field, A is real, and AA' is an $m \times m$ non-singular matrix. We assert that the projection, e, of X onto U is given by the matrix

$$E = A'(AA')^{-1} A \quad . \tag{9}$$

First, we observe that any collection of vectors in U can be expressed in matrix form as

$$KA\underset{\sim}{x} \quad ,$$

where K is a suitable matrix. Hence, it is plain that

$$E\underset{\sim}{x} = A'(AA')^{-1} A\underset{\sim}{x}$$

maps X <u>into</u> U. Next we verify that

$$E^2 = A'(AA')^{-1} AA'(AA')^{-1} A = A'(AA')^{-1} A = E \quad .$$

In order to show that E maps X <u>onto</u> U, it suffices to prove that E is of rank m. But for an idempotent matrix E, rank E = trace E (= tr E), and we find that

$$\mathrm{tr}\{A'(AA')^{-1} A\} = \mathrm{tr}\{AA'(AA')^{-1}\} = \mathrm{tr}\, I_m = m \quad .$$

Thus, every vector of U is of the form $u = xe$, for some x, which proves that $ue = u$.

We notice the additional fact that

$$E' = E \quad , \tag{10}$$

which will be required later. Incidentally, this construction of E is unique, that is, it does not depend on the basis of U.

The decomposition of the representation space X is now expressed by the matrix equations

$$I = \sum_{\alpha=1}^{s} E_\alpha \quad , \tag{11}$$

where $E_\alpha E_\beta = \delta_{\alpha\beta} E_\alpha$, and

$$P(g)E_\alpha = E_\alpha P(g) \quad , \tag{12}$$

where $P(g)$ is the permutation matrix that corresponds to the permutation $g \in G$. Since permutation matrices are orthogonal, we have that

$$P'(g)\, P(g) = I \quad . \tag{13}$$

Let us now return to the analysis of variance, that is the braking up of

$$\underset{\sim}{x}'\underset{\sim}{x} = \sum_{i=1}^{n} x_i^2$$

into quadratic forms. By (11),

$$\underset{\sim}{x} = \sum_{\alpha=1}^{s} E_\alpha \underset{\sim}{x} \quad ,$$

whence

$$\underset{\sim}{x}'\underset{\sim}{x} = \sum_{\alpha=1}^{s} \underset{\sim}{x}'E_\alpha \underset{\sim}{x} = \sum_{\alpha=1}^{s} Q_\alpha \quad ,$$

say. We note that, since $E'_\alpha = E_\alpha$, $Q_\alpha = \underset{\sim}{x}'E_\alpha \underset{\sim}{x}$ is indeed a quadratic form.

We list some of the properties of these forms :

1° $Q_\alpha = (E_\alpha \underset{\sim}{x})' (E_\alpha \underset{\sim}{x})$, which shows that Q_α is non-negative definite. It is, in fact, expressed as the sum of n squares of linear forms

$$\underset{\sim}{y}_\alpha = E_\alpha \underset{\sim}{x} \quad ;$$

these forms are, of course, linearly dependent when $s > 1$.

2° Let n_α be the rank of Q_α (statisticians call n_α the number of degrees of freedom in Q_α). Then

$$n_\alpha = \mathrm{tr}\ E_\alpha \ , \qquad n = \sum_{\alpha=1}^{s} n_\alpha \quad .$$

3° Each Q_α is invariant under the action of G , that is $Q_\alpha g = Q_\alpha$ for all $g \in G$. For

$$Q_\alpha g = (P(g)\underset{\sim}{x})' E_\alpha(P(g)\underset{\sim}{x}) = \underset{\sim}{x}'P'(g)\ E_\alpha\ P(g) = \underset{\sim}{x}'E_\alpha \underset{\sim}{x} \quad ,$$

by virtue of (12) and (13).

3. Applications to Block design.

Since the variables are labelled by two suffixes, it is more convenient to collect the basis vectors of X in a matrix

$$\underset{\sim}{X} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1q} \\ x_{21} & x_{22} & \cdots & x_{2q} \\ \cdots & \cdots & \cdots & \cdots \\ x_{p1} & x_{p2} & \cdots & x_{pq} \end{pmatrix} \qquad (pq = n) \quad .$$

The linear transformations of X belong to the matrix ring M_n , but we know that

$$M_n = M_p \otimes M_q \quad ,$$

and, for technical reasons, it is desirable to introduce two auxiliary vector spaces Y , Z of dimensions p and q respectively, say

$$Y = (y_1 , \dots , y_p) , \qquad Z = (z_1 , \dots , z_q) \quad ,$$

and to identify X with $Y \otimes Z$, thus $x_{ij} = y_i \otimes z_j$. If

$$y_i\, k = \sum_{a=1}^{p} k_{ia}\, y_a , \qquad z_j\, \ell = \sum_{b=1}^{q} \ell_{jb}\, z_b \qquad (i = 1 , \dots , p , \quad j = 1 , \dots , q)$$

are linear transformations k , ℓ with matrices $K = (k_{ia})$ and $L = (\ell_{jb})$ respectively, then

$$x_{ij}(k \otimes \ell) = \sum_{a,b} k_{ia}\, \ell_{jb}(y_a \otimes z_b) = \sum_{a,b} k_{ia}\, \ell_{jb}\, x_{ab} \quad ,$$

so that

$$\underset{\sim}{X}(k \otimes \ell) = K\underset{\sim}{X}L'$$

gives the transformation of $\underset{\sim}{X}$ in matrix form.

Now, the symmetry group of the block design is

$$G = S_p \times S_q \quad ,$$

where S_p and S_q range over all permutation matrices of degrees p and q respectively. We let S_p act on Y , and let S_q act on Z .

Consider first the decomposition of Y under the action of S_p . It is well-known that a doubly-transitive permutation group of degree p has precisely two invariant subspaces, of degrees 1 and $p - 1$ respectively. Thus

$$(14) \qquad Y = Y_1 \oplus Y_2 \quad .$$

In fact,

$$Y_1 = \{y_1 + y_2 + \dots + y_p\} , \qquad Y_2 = \{y_1 - y_p , y_2 - y_p , \dots , y_{p-1} - y_p\} \quad ,$$

and these spaces are absolutely irreducible. Let us find the idempotents D_1 and D_2 which correspond to this decomposition. As regards Y_1 , the matrix defining this space is the $1 \times p$ matrix

$$A = (1 , 1 , \dots , 1)_p \quad ,$$

and we obtain as the corresponding idempotent

$$D_1 = A'(AA')^{-1} A = \frac{1}{p} J_p \quad ,$$

where, for any p, we shall put

$$J_p = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & & 1 \\ \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} .$$

Since the decomposition (14) leads to the equation

$$I_p = D_1 + D_2 \quad ,$$

it follows that

$$D_2 = I_p - \frac{1}{p} J_p \quad .$$

Similarly, we have the decomposition

$$Z = Z_1 \oplus Z_2 \quad ,$$

with corresponding idempotents F_1, F_2 satisfying

$$I_q = F_1 + F_2 \quad ,$$

where

$$F_1 = \frac{1}{q} J_q \, , \qquad F_2 = I_q - \frac{1}{q} J_q \quad .$$

Hence the action of G on X causes X to break up into four invariant (and in fact absolutely irreducible) subspaces, which correspond to the mutually orthogonal idempotents

$$E_{11} = D_1 \times F_1 \, , \qquad E_{12} = D_1 \times F_2 \, , \qquad E_{21} = D_2 \times F_1 \, , \qquad E_{22} = D_2 \times F_2 \quad ,$$

whose sum is I_n. The corresponding decomposition of the space X is given by

$$\underset{\sim}{X} = \sum_{\alpha,\beta=1}^{2} D_\alpha \underset{\sim}{X} F'_\beta = \sum_{\alpha,\beta=1}^{2} \underset{\sim}{X}_{\alpha\beta} \quad ,$$

say. According to the general theory, this leads to a decomposition of the sum of squares of the elements of $\underset{\sim}{X}$. Thus, since

$$\text{tr } \underset{\sim}{X}\underset{\sim}{X}' = \sum_{i,j} x_{ij}^2 \quad ,$$

we conclude that

$$\operatorname{tr} \underset{\sim}{X}\underset{\sim}{X}' = \sum_{\alpha,\beta=1}^{2} \operatorname{tr}(X_{\alpha\beta} X'_{\alpha\beta}) \quad .$$

It is an easy matter to verify that this is precisely the identity which forms the basis of the analysis of variance. In fact, straightforward matrix calculations yield the following results :

$$D_1 \underset{\sim}{X}\underset{\sim}{F}'_1 = \frac{1}{pq} J_p \underset{\sim}{X} J_p = \begin{pmatrix} \bar{\bar{x}} & \cdots & \bar{\bar{x}} \\ \cdots & \cdots & \cdots \\ \bar{\bar{x}} & \cdots & \bar{\bar{x}} \end{pmatrix} ,$$

$$D_1 \underset{\sim}{X}\underset{\sim}{F}'_2 = \frac{1}{p} J_p \underset{\sim}{X}(I_q - \frac{1}{q} J_q) = \begin{pmatrix} \bar{x}_{.1} - \bar{\bar{x}} & \cdots & \bar{x}_{.q} - \bar{\bar{x}} \\ \cdots & \cdots & \cdots \\ \bar{x}_{.1} - \bar{\bar{x}} & \cdots & \bar{x}_{.q} - \bar{\bar{x}} \end{pmatrix} ,$$

$$D_2 \underset{\sim}{X}\underset{\sim}{F}'_1 = (I_p - \frac{1}{p} J_p)\underset{\sim}{X} \frac{1}{q} J_q = \begin{pmatrix} \bar{x}_{1.} - \bar{\bar{x}} & \cdots & \bar{x}_{1.} - \bar{\bar{x}} \\ \cdots & \cdots & \cdots \\ \bar{x}_{p.} - \bar{\bar{x}} & \cdots & \bar{x}_{p.} - \bar{\bar{x}} \end{pmatrix} ,$$

$$D_2 \underset{\sim}{X}\underset{\sim}{F}'_2 = (I_p - \frac{1}{p} J_p)\underset{\sim}{X}(I_q - \frac{1}{q} J_q) = (x_{ij} - \bar{x}_{ij} - \bar{x}_{i.} + \bar{\bar{x}}) \quad ,$$

whence the analysis of variance, given in (1), is obtained as the sum of the squares of the elements in each matrix.
