

B. CERANKA (Poznań)

AFFINE RESOLVABLE INCOMPLETE BLOCK DESIGNS

1. Introduction. A general method of the analysis of variance for block designs has been presented by Tocher [8]. Rees [6] has extended this method to the case of many non-orthogonal classifications. The chief numerical calculation in the method of Tocher [8] is the inversion of a matrix that yields the covariance matrix of the adjusted treatment means. This can be simple when there is some pattern in the design. Sometimes the pattern of the design itself is not very helpful, but that of its dual has certain desirable properties. The affine resolvable incomplete block designs are of that type. Their analysis can be simplified by considering their duals first. The general formula of Rees [6] is then helpful in carrying out the analysis.

2. General method for a design with two classifications. Assuming the same model as that adopted by Plackett [5] and Rees [6] let us write the linear observational equations in the matrix form

$$(1) \quad \mathbf{y} = \mathbf{A}\boldsymbol{\theta} + \mathbf{e},$$

where \mathbf{y} is an $(N \times 1)$ -vector of observations, \mathbf{A} — an $(N \times p)$ -matrix of known coefficients (with $N > p$), called the *design matrix*, $\boldsymbol{\theta}$ — a $(p \times 1)$ -vector of parameters, and \mathbf{e} — an $(N \times 1)$ -vector of random errors with $E(\mathbf{e}) = \mathbf{0}$ and $E(\mathbf{e}\mathbf{e}') = \sigma^2\mathbf{I}$, where \mathbf{I} is the $(N \times N)$ -identity matrix.

For a design with two classifications we can write

$$(2) \quad \boldsymbol{\theta} = [\mu, \boldsymbol{\theta}'_1, \boldsymbol{\theta}'_2]' \quad \text{and} \quad \mathbf{A} = [\mathbf{1}, \mathbf{A}_1, \mathbf{A}_2],$$

where $\boldsymbol{\theta}_1$ is an $(l_1 \times 1)$ -parameter vector and \mathbf{A}_1 is an $(N \times l_1)$ -design matrix for the first classification, $\boldsymbol{\theta}_2$ is an $(l_2 \times 1)$ -parameter vector and \mathbf{A}_2 is an $(N \times l_2)$ -design matrix for the second classification, μ is the general parameter and $\mathbf{1}$ is the $(N \times 1)$ -unit vector. Thus model (1) has the form

$$(3) \quad E(\mathbf{y}) = [\mathbf{1}, \mathbf{A}_1, \mathbf{A}_2][\mu, \boldsymbol{\theta}'_1, \boldsymbol{\theta}'_2]'.$$

It is assumed that the rank of the partitioned design matrix $[\mathbf{1}, \mathbf{A}_1, \mathbf{A}_2]$ in (3) is $l_1 + l_2 - 1 = p - 2$, which means that disconnected designs are

excluded from the present consideration. This is the only restriction of the design. Since the normal equations for the general model (1) are

$$\mathbf{A}'\mathbf{A}\boldsymbol{\theta} = \mathbf{A}'\mathbf{y},$$

the least-squares estimate of $\boldsymbol{\theta}$ is

$$\hat{\boldsymbol{\theta}} = (\mathbf{A}'\mathbf{A})^{-}\mathbf{A}'\mathbf{y},$$

where $(\mathbf{A}'\mathbf{A})^{-}$ is the generalized inverse of the matrix $\mathbf{A}'\mathbf{A}$, so that

$$(\mathbf{A}'\mathbf{A})(\mathbf{A}'\mathbf{A})^{-}(\mathbf{A}'\mathbf{A}) = \mathbf{A}'\mathbf{A}.$$

Since the matrix $\mathbf{A}'\mathbf{A}$ is of rank $p-2$, and hence there exists a $(p \times 2)$ -matrix \mathbf{D} of rank 2 such that $\mathbf{A}\mathbf{D} = \mathbf{0}$. Furthermore, there exists a $(2 \times p)$ -matrix \mathbf{C} of rank 2 such that $|\mathbf{C}\mathbf{D}| \neq 0$. A suitable generalized inverse of $\mathbf{A}'\mathbf{A}$ is then the true inverse of the matrix $\mathbf{A}'\mathbf{A} + \mathbf{C}'\mathbf{C}$ with \mathbf{C} given by

$$\mathbf{C} = \frac{1}{\sqrt{N}} \begin{bmatrix} 0 & \mathbf{r}'_1 & \mathbf{0}' \\ 0 & \mathbf{0}' & \mathbf{r}'_2 \end{bmatrix},$$

where \mathbf{r}_i ($i = 1, 2$) is the vector of replications for the i -th classification. The parameters $\boldsymbol{\theta} = [\mu, \theta'_1, \theta'_2]'$ are then estimable under the side conditions $\mathbf{C}\boldsymbol{\theta} = \mathbf{0}$, and the least-squares estimate of $\boldsymbol{\theta}$ (see [7]) is

$$\hat{\boldsymbol{\theta}} = (\mathbf{A}'\mathbf{A} + \mathbf{C}'\mathbf{C})^{-1}\mathbf{A}'\mathbf{y}$$

which, for the design with two classifications defined by (2), reads explicitly as

$$\hat{\boldsymbol{\theta}} = \begin{bmatrix} \hat{\mu} \\ \hat{\theta}'_1 \\ \hat{\theta}'_2 \end{bmatrix} = \begin{bmatrix} G/N \\ \mathbf{r}_1^{-\delta}(\mathbf{Q}_1 - \mathbf{n}\boldsymbol{\Omega}\mathbf{Q}_2 + \mathbf{r}_1\mathbf{r}'_2\boldsymbol{\Omega}\mathbf{Q}_2/N) \\ \boldsymbol{\Omega}\mathbf{Q}_2 \end{bmatrix}$$

with

$$(4) \quad \begin{aligned} \mathbf{Q}_1 &= \mathbf{T}_1 - (G/N)\mathbf{r}_1, & \mathbf{Q}_2 &= \mathbf{T}_2 - \mathbf{n}\mathbf{r}_1^{-\delta}\mathbf{T}_1, \\ \boldsymbol{\Omega}^{-1} &= \mathbf{r}_2^{\delta} - \mathbf{n}'\mathbf{r}_1^{-\delta}\mathbf{n} + \mathbf{r}_2\mathbf{r}'_2/N, \end{aligned}$$

where \mathbf{T}_i ($i = 1, 2$) is the vector of totals for the i -th classification, G — the overall total, \mathbf{r}_i^{δ} — the vector \mathbf{r}_i expressed as a diagonal matrix, and \mathbf{n} ($\mathbf{n} = \mathbf{A}'_1\mathbf{A}_2$) is the incidence matrix with a row for each level of the first classification and a column for each level of the second classification, and with elements representing the number of times each level of the first classification occurs with each level of the second classification.

The main numerical difficulty is the inversion of the matrix $\boldsymbol{\Omega}^{-1}$. For some types of experimental designs it is possible to give the final formula of $\boldsymbol{\Omega}$, so that the numerical inversion can be avoided in practical applications.

The adjusted means for the second classification are given by

$$\mathbf{a}_2 = (G/N)\mathbf{1} + \mathbf{\Omega}\mathbf{Q}_2.$$

The covariance matrix of the vector \mathbf{a}_2 is equal to $\sigma^2\mathbf{\Omega}$. The error sum of squares is given by the formula

$$\begin{aligned} E &= (\mathbf{y} - \mathbf{A}\hat{\boldsymbol{\theta}})'(\mathbf{y} - \mathbf{A}\hat{\boldsymbol{\theta}}) \\ &= (\mathbf{y}'\mathbf{y} - G^2/N) - (\mathbf{T}'_1\mathbf{r}_1^{-\delta}\mathbf{T}_1 - G^2/N) - \mathbf{Q}'_2\mathbf{\Omega}\mathbf{Q}_2 \\ &= \mathbf{y}'\{\mathbf{I} - \mathbf{A}_1\mathbf{r}_1^{-\delta}\mathbf{A}'_1 - (\mathbf{A}_2 - \mathbf{A}_1\mathbf{r}_1^{-\delta}\mathbf{n})\mathbf{\Omega}(\mathbf{A}'_2 - \mathbf{n}'\mathbf{r}_1^{-\delta}\mathbf{A}'_1)\}\mathbf{y}. \end{aligned}$$

The sum of squares for i -th classification ($i = 1, 2$) ignoring the others, is given by

$$H_i = \mathbf{y}'\{\mathbf{A}_i(\mathbf{r}_i^{-\delta} - (1/N)\mathbf{1}\mathbf{1}')\mathbf{A}'_i\}\mathbf{y}.$$

The sum of squares for the second classification, eliminating the first, is given by

$$H'_2 = \mathbf{y}'\{(\mathbf{A}_2 - \mathbf{A}_1\mathbf{r}_1^{-\delta}\mathbf{n})\mathbf{\Omega}(\mathbf{A}'_2 - \mathbf{n}'\mathbf{r}_1^{-\delta}\mathbf{A}'_1)\}\mathbf{y},$$

and the sum of squares for the first classification, eliminating the second, H'_1 say, can be evaluated from the equation

$$H_1 + H'_2 = H'_1 + H_2.$$

The quotient

$$(5) \quad F = \frac{H'_i/(l_i - 1)}{E/(N - l_1 - l_2 + 1)},$$

is a proper statistics for testing the hypothesis $H_0 : \boldsymbol{\theta}_i = \mathbf{0}$. If the hypothesis H_0 is true, then the quotient has the F -distribution with $l_i - 1$ and $N - l_1 - l_2 + 1$ degrees of freedom.

2. Construction of the design. Let t be the number of treatments, k — the number of experimental units per block, r — the number of replications for each treatment, b — the number of blocks, and λ — the number of times any 2 treatments occur together in a block. Then, evidently, $N = tr = bk$.

In some cases it is possible to group the blocks into c sets such that each set consists of exactly one or more replicates of the treatments. In the special case where $c = r$ the blocks are grouped into a single replication. Balanced incomplete block designs that can be arranged in sets of blocks, each containing a complete replicate, are known as *resolvable balanced incomplete block designs*. Bose [1] has shown that if a balanced incomplete block design with parameters t, b, r, k and λ is resolvable, then the following inequality must hold:

$$b \geq t + r - 1.$$

Moreover, if for a resolvable balanced incomplete block design the equality

$$(6) \quad b = t + r - 1$$

holds, the design has the property that two blocks belonging to different sets have the same number of treatments in common. This number is

$$(7) \quad m = k^2/t.$$

The designs for which properties (6) and (7) hold simultaneously are called *affine resolvable balanced incomplete block designs*. The analysis of them has been discussed by Bose [1]. It is clear that if we omit some of the replications of the affine resolvable balanced incomplete block design, property (7) will be preserved, though equality (6) will not remain true. The resulting design will be called an *affine resolvable partially balanced incomplete block design*. Such designs exist for $r = 2, 3, \dots, b - t$, where b denotes the number of blocks in the appropriate balanced design. Since, for some numbers of treatments, affine resolvable balanced incomplete block designs exist for very large numbers of replications only, affine resolvable partially balanced incomplete block designs can quite extensively be used in practice. Lists of partially balanced incomplete block designs with a small number of replications, known up to now, do not contain experiments with all number of treatments. It turns out that the gaps in the lists can partially be fulfilled by constructing plans of experiments based on condition (7), but not necessarily balanced. The designs so constructed will be called the *affine resolvable partially balanced designs*, i.e. designs with more than one association class. New plans for designs of this type can be built for the following numbers t ($t \leq 200$) of treatments: 18, 24, 27, 32, 45, 48, 50, 54, 75, 80, 98, 108, 128, 162, 200. They supplement the list of incomplete block designs given by Cochran and Cox [3] for a small number of replications.

Let us consider affine resolvable balanced incomplete block designs with parameters t_1, b, r, k_1 and λ . Replicating the designs n times, we get again affine resolvable partially balanced incomplete block designs with parameters $t = nt_1, b, r, k = nk_1, \lambda_1 = r, \lambda_2 = \lambda$, and such that two blocks belonging to different sets have the number of common treatments equal to $m = n^2 k_1^2 / nt_1 = nm_1$. If from such designs we eliminate one or more replicates, then we still have affine resolvable partially balanced incomplete block designs, but with a smaller number of blocks. The resulting design has three or more association classes. These designs can be obtained also by taking some of the first replications from affine resolvable balanced incomplete block designs and repeating this procedure n times.

4. Analysis of variance. In this section the analysis of variance for affine resolvable balanced and partially balanced incomplete block designs

is evaluated. Let the first classification be due to treatments and the second one due to blocks. Then $l_1 = t$, $l_2 = b$, $r_1 = r\mathbf{1}$, $r_2 = k\mathbf{1}$, $r_2^\delta = k\mathbf{I}$, $r_1^{-\delta} = (1/r)\mathbf{I}$, $r_2 r_2' = k^2 \mathbf{1}\mathbf{1}'$, and the matrix Ω^{-1} , defined in (4), is expressed by the equation

$$(8) \quad \Omega^{-1} = k\mathbf{I} - \frac{1}{r} \mathbf{n}'\mathbf{n} + \frac{k^2}{N} \mathbf{1}\mathbf{1}'.$$

The product $\mathbf{n}'\mathbf{n}$ can be expressed in the form of the partitioned matrix

$$(9) \quad \mathbf{n}'\mathbf{n} = \begin{bmatrix} \mathbf{S} & \mathbf{R} & \dots & \mathbf{R} \\ \mathbf{R} & \mathbf{S} & \dots & \mathbf{R} \\ \dots & \dots & \dots & \dots \\ \mathbf{R} & \mathbf{R} & \dots & \mathbf{S} \end{bmatrix}$$

with submatrices $\mathbf{S} = k\mathbf{I}$ and $\mathbf{R} = m\mathbf{1}\mathbf{1}'$, where \mathbf{I} is an identity matrix $(t/k) \times (t/k)$, and $\mathbf{1}$ is a unit vector $(t/k) \times 1$. Then the matrix Ω^{-1} is of the form

$$\Omega^{-1} = \frac{1}{N} \begin{bmatrix} \mathbf{U} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{U} & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{U} \end{bmatrix},$$

where the $((t/k) \times (t/k))$ -matrix \mathbf{U} is $\mathbf{U} = k(N-t)\mathbf{I} + k^2 \mathbf{1}\mathbf{1}'$. The matrix Ω^{-1} has a simple pattern and can be inverted by one of the methods described by Pearce [4]. Then we obtain

$$(10) \quad \Omega = \frac{1}{N-t} \begin{bmatrix} \mathbf{u} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{u} & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{u} \end{bmatrix},$$

where the $((t/k) \times (t/k))$ -matrix \mathbf{u} is $\mathbf{u} = b\mathbf{I} - \mathbf{1}\mathbf{1}'$.

The components of the vector $\mathbf{Q}_2 = \mathbf{T}_2 - (1/r)\mathbf{n}\mathbf{T}_1$, denoted by q_j , can be obtained practically in the following way. Let \mathbf{T}_2 be the vector of block totals, and \mathbf{T}_1 the vector of treatment totals. Among the components of \mathbf{T}_1 we choose sums of those treatments which appear in the j -th block, take the sum of them, divide it by the number of replications r , and subtract the result from the j -th component of \mathbf{T}_2 . This is performed, in turn, for each block, i.e. for $j = 1, 2, \dots, b$. Evidently, $\mathbf{1}'\mathbf{Q}_2 = 0$.

The analysis of variance can be summarized in the following table:

Source of variation	Degrees of freedom	Sum of squares
Blocks eliminating treatments	$b - 1$	H'_2
Treatments ignoring blocks	$t - 1$	H_1
Error	$N - t - b + 1$	E
Blocks ignoring treatments	$b - 1$	H_2
Treatments eliminating blocks	$t - 1$	H'_1
Total	$N - 1$	S

In this table we have

$$\begin{aligned}
 H_1 &= (1/r)\mathbf{T}'_1\mathbf{T}_1 - G^2/N, & H'_2 &= \mathbf{Q}'_2\boldsymbol{\Omega}\mathbf{Q}_2, \\
 S &= \mathbf{y}'\mathbf{y} - G^2/N, & E &= S - H_1 - H'_2 = S - H'_1 - H_2, \\
 H_2 &= (1/k)\mathbf{T}'_2\mathbf{T}_2 - G^2/N, & H'_1 &= H_1 + H'_2 - H_2.
 \end{aligned}$$

The statistics (5) for testing the hypothesis $H_0: \boldsymbol{\theta}_1 = \mathbf{0}$ is

$$F = \frac{H'_1/(t-1)}{E/(N-t-b+1)}.$$

If the hypothesis H_0 is true, then the quotient has the distribution F with $t-1$ and $N-t-b+1$ degrees of freedom. The estimates of block parameters are given by the vector $\hat{\boldsymbol{\theta}}_2 = \boldsymbol{\Omega}\mathbf{Q}_2$. The adjusted blocks means are given by

$$\mathbf{a}_2 = (G/N)\mathbf{1} + \boldsymbol{\Omega}\mathbf{Q}_2.$$

The estimates of treatment parameters are given by the vector

$$\hat{\boldsymbol{\theta}}_1 = (1/r)(\mathbf{Q}_1 - \mathbf{n}\boldsymbol{\Omega}\mathbf{Q}_2).$$

The adjusted treatments means are given by

$$\mathbf{a}_1 = (G/N)\mathbf{1} + (1/r)(\mathbf{Q}_1 - \mathbf{n}\boldsymbol{\Omega}\mathbf{Q}_2).$$

The covariance matrix of the vector \mathbf{a}_1 is equal to $\sigma^2\boldsymbol{\Omega}_t$, where

$$\boldsymbol{\Omega}_t = (1/r^2)(r\mathbf{I} + \mathbf{n}_0'\boldsymbol{\Omega}\mathbf{n}_0) \quad \text{with } \mathbf{n}_0 = \mathbf{n} - (rk/N)\mathbf{1}\mathbf{1}'.$$

The standard error of the difference between two adjusted treatment means is evaluated from the formula

$$S_d = \sqrt{\frac{E}{N-t-b+1} \frac{2b(kr-k+r-\lambda_i)}{r^2(N-t)}},$$

where $\lambda_i = 0, 1, 2, \dots, r$ is the number of times any two treatments occur together in a block.

We show now a different way of getting the matrix Ω . Caliński [2] has introduced the matrix $M_0 = r^{-\delta}nk^{-\delta}n' - \mathbf{1}r'/N$ and has shown that, using M_0 , the matrix Ω_i for treatments can be constructed without any inversion. This will be particularly in the case where

$$(11) \quad M_0^h = \nu^{h-1}M_0.$$

Caliński [2] showed also that affine resolvable incomplete block designs have property (11).

To evaluate the matrix Ω for blocks, let us first define the matrix

$$\tilde{M}_0 = k^{-\delta}n'r^{-\delta}n - \mathbf{1}k'/N,$$

which in the case of affine resolvable incomplete block designs is of the form

$$\tilde{M}_0 = (1/kr)n'n - (k/N)\mathbf{1}\mathbf{1}'.$$

In view of property (9), the matrix \tilde{M}_0 can be written as

$$\tilde{M}_0 = \frac{1}{N} \begin{bmatrix} D & 0 & \dots & 0 \\ 0 & D & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & D \end{bmatrix},$$

where $D = tI - k\mathbf{1}\mathbf{1}'$ is a $((t/k) \times (t/k))$ -matrix. The matrix Ω^{-1} , defined by (8), can be presented as

$$\Omega^{-1} = k(I - \tilde{M}_0),$$

and hence

$$\Omega = (1/k)(I - \tilde{M}_0)^{-1} = (1/k)\left(I + \sum_{h=1}^{\infty} M_0^h\right).$$

The matrix \tilde{M}_0 satisfies condition (11) with $\nu = 1/r$. Then

$$\Omega = (1/k)\{I + [r/(r-1)]\tilde{M}_0\}$$

which is equivalent to (10).

It can be shown that the matrix Ω_i for treatments is equal to

$$\Omega_i = (1/r)\{I + [r/(r-1)]M_0\}.$$

References

[1] R. C. Bose, *A note on the resolvability of balanced incomplete block designs*, Sankhyā 8 (1947), p. 249-256.
 [2] T. Caliński, *On some desired patterns in block designs*, Biometrics 27 (1971), p. 275-292.

- [3] W. G. Cochran and G. M. Cox, *Experimental designs*, J. Wiley, New York 1950.
[4] S. C. Pearce, *The use and classification of non-orthogonal designs*, J. Royal Statist. Soc. A 126 (1963), p. 353-369.
[5] R. L. Plackett, *Principles of regression analysis*, Clarendon Press, Oxford 1960.
[6] D. H. Rees, *The analysis of variance of designs with many non-orthogonal classifications*, J. Royal Statist. Soc. B 28 (1966), p. 110-117.
[7] H. Scheffé, *The analysis of variance*, J. Wiley, New York 1959.
[8] K. D. Tocher, *The design and analysis of block experiments*, J. Royal Statist. Soc. B 14 (1952), p. 45-91.

DEPT. OF MATHEMATICAL AND STATISTICAL METHODS
ACADEMY OF AGRICULTURE, POZNAŃ

Received on 26. 5. 1973

B. CERANKA (Poznań)

AFINICZNIE ROZKŁADALNE UKŁADY O BLOKACH NIEKOMPLETNYCH

STRESZCZENIE

W pracy podano definicję oraz zasadę konstrukcji afinicznie rozkładalnych układów o blokach niekompletnych. Przedstawiono ogólną metodę uzyskiwania estymatorów dla parametrów obiektowych oraz metodę uzyskiwania macierzy kowariancji tych estymatorów. Podstawowe obliczenia numeryczne w analizie doświadczeń układów blokowych związane są ze znalezieniem odwrotności tej macierzy kowariancji. Wykorzystując rozwiązanie dualne pokazano, w jak prosty sposób można znaleźć poszukiwaną macierz kowariancji.
