ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
20,2 (1990), pp. 211–217

M. KOZŁOWSKA and R. WALKOWIAK (Poznań)

INCOMPLETE HOMOGENEOUS MULTIRESPONSE MODELS: ESTIMATION

1. Introduction. Models being a generalization of those described by Srivastava [5], [6] and by Caliński and Kozłowska [2], referred to as incomplete homogeneous multiresponse models, are further considered. The simple least squares estimators of treatment contrasts are presented.

The purpose of this paper is the development of a method of estimation in multivariate experiments in which a different subset of the variables under study is observed on each of the disjoint subsets of experimental units. If all t variables are observed on each unit, the experiment is called a *complete multiresponse experiment*. Otherwise, it is called an *incomplete multiresponse experiment*. So, one can say that the term *incomplete* refers here to variables.

Let $\boldsymbol{Y}_i$ denote the $(n_i \times t_i)$-matrix of observations from the i-th subset of experimental units, $i = 1, 2, \ldots, u$. The adopted model is

$$(1.1)\qquad \left(\boldsymbol{Y}_i, [\boldsymbol{U}_i', \boldsymbol{\Delta}_i']\begin{bmatrix} \boldsymbol{\Xi}_i \\ \boldsymbol{\Gamma M}_i \end{bmatrix}, \boldsymbol{M}_i'\boldsymbol{\Sigma M}_i \otimes \boldsymbol{I}_{n_i}\right),$$

where $[\boldsymbol{U}_i', \boldsymbol{\Delta}_i']$ is the design matrix in which the $(n_i \times s_i)$-matrix $\boldsymbol{U}_i'$ corresponds to the $(s_i \times t_i)$-matrix $\boldsymbol{\Xi}_i$ of unknown parameters, and the $(n_i \times v)$-matrix $\boldsymbol{\Delta}_i'$ corresponds to the $(v \times t_i)$-matrix $\boldsymbol{\Gamma M}_i$. Here $\boldsymbol{\Gamma}$ is the $(v \times t)$-matrix of treatment parameters, and $\boldsymbol{M}_i$ denotes the $(t \times t_i)$-matrix obtained from the identity matrix $\boldsymbol{I}_t$ through elimination of columns corresponding to variables which are not observed on the i-th subset of variables. The matrix $\boldsymbol{\Sigma}$ is the unknown covariance matrix of the elements in each row of the matrix of observations if these rows are complete, $\boldsymbol{I}_{n_i}$ is the $n_i \times n_i$ identity matrix, and $\otimes$ denotes the Kronecker product.

The above i-th model describes observations of units from the i-th subset of units only. Now it is convenient to use the operation vec, which transforms the matrix $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_g]$ into the vector $[\boldsymbol{a}_1', \boldsymbol{a}_2', \ldots, \boldsymbol{a}_g']'$, to construct the model of observation of all experimental units. We assume that this model is of the form

$$\left(y, [U', \Delta']\begin{bmatrix}\xi \\ \gamma\end{bmatrix}, \Sigma^*\right), \tag{1.2}$$

where

$$y = \begin{bmatrix} \operatorname{vec} Y_1 \\ \dots\dots \\ \operatorname{vec} Y_u \end{bmatrix}, \quad U' = \begin{bmatrix} I_{t_1}\otimes U'_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & I_{t_u}\otimes U'_u \end{bmatrix},$$

$$\Delta' = \begin{bmatrix} M'_1\otimes\Delta'_1 \\ \dots\dots \\ M'_u\otimes\Delta'_u \end{bmatrix}, \quad \xi = \begin{bmatrix} \operatorname{vec} \Xi_1 \\ \dots\dots \\ \operatorname{vec} \Xi_u \end{bmatrix}, \quad \gamma = \operatorname{vec} \Gamma, \tag{1.3}$$

$$\Sigma^* = \begin{bmatrix} M'_1\Sigma M_1\otimes I_{n_1} & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & M'_u\Sigma M_u\otimes I_{n_u} \end{bmatrix}.$$

It is convenient to use the matrices

$$C_i = \Delta_i\Phi_i\Delta'_i, \quad i = 1, 2, \dots, u, \tag{1.4}$$

where

$$\Phi_i = I_{n_i} - U'_i(U_iU'_i)^- U_i. \tag{1.5}$$

DEFINITION. The incomplete multiresponse model (1.2) is called *homogeneous* with respect to the positive definite symmetric matrix X if there exists a set of X-orthonormal vectors w_j, $j\in Z$, which satisfies the relations

$$C_i w_j = \lambda_{ij} X w_j \tag{1.6}$$

for every i, $i = 1, 2, \dots, u$, where

$$Z = \{j: \bigvee_i \lambda_{ij} > 0\}.$$

In the above definition the matrix X is described. It is a certain positive definite symmetric matrix. Although various suitable matrices can be chosen, there are two of particular interest: $X = I$ or $X = (\Delta\Delta')^{-1}$.

2. Least squares estimators. The simple least squares estimators of ξ and γ in the model (1.2) are obtainable by solving the normal equation

$$\begin{bmatrix} U \\ \Delta \end{bmatrix}[U', \Delta']\begin{bmatrix}\xi \\ \gamma\end{bmatrix} = \begin{bmatrix} U \\ \Delta \end{bmatrix} y.$$

Applying the lemma from [4], p. 27, and the theorem from [3], p. 22, we obtain the following forms of these estimators:

$$\mathring{\xi} = (UU')^{-} U[I - \Delta' C^{-} \Delta \Phi] y, \quad \mathring{\gamma} = C^{-} \Delta \Phi y, \tag{2.1}$$

where

$$\Phi = \begin{bmatrix} I_{t_1} \otimes \Phi_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & I_{t_u} \otimes \Phi_u \end{bmatrix}, \tag{2.2}$$

Φ_i is given by (1.5), and C^{-} is a certain generalized inverse of the matrix

$$C = \Delta \Phi \Delta' = \sum_{i=1}^{u} (M_i M_i' \otimes C_i), \tag{2.3}$$

C_i is given by (1.4). It is evident from equalities (2.3) and (1.6) that the vectors $(e_l \otimes w_j)$, $l = 1, 2, \ldots, t$, $j \in Z$, are eigenvectors of C with respect to $I_t \otimes X$ corresponding to the eigenvalues $\sum_{i=1}^{u} m_{il} \lambda_{il}$, where m_{il} is the (i, l)-th element of the matrix $M_i M_i'$, and e_l is the $(t \times 1)$-vector with the l-th element equal to 1 and other elements equal to 0. It is convenient to use a decomposition of C according to the formula

$$C = \sum_{j \in Z} \sum_{l \in L_j} [(\sum_{i=1}^{u} m_{il} \lambda_{ij})(e_l e_l' \otimes X w_j w_j' X)], \tag{2.4}$$

where

$$L_j = \{l: \bigvee_i m_{il} \lambda_{ij} \neq 0\}, \quad j \in Z.$$

Hence one of the possibilities of the generalized inverse of the matrix (2.4) can be

$$C^{-} = \sum_{j \in Z} \sum_{l \in L_j} [(\sum_{i=1}^{u} m_{il} \lambda_{ij})^{-1} (e_l e_l' \otimes w_j w_j')]. \tag{2.5}$$

We want to estimate the parametric function

$$c' \Gamma d = (d' \otimes c') \gamma = z' \gamma,$$

where c is a certain $(v \times 1)$-vector whose components sum up to zero, and d is a $(t \times 1)$-vector. It is known (see, e.g., [1]) that each function $z'\gamma$, if it is estimable, may be written in the form

$$z' \gamma = \sum_{j \in Z} \sum_{l \in L_j} g_{jl} z_{lj}' \gamma, \tag{2.6}$$

where g_{jl} are certain numbers and $z_{lj} = e_l \otimes X w_j$. The above formula shows that the estimator of $z'\gamma$ may be obtained as the linear combination of the estimators of the contrasts $z_{jl}'\gamma$, $j \in Z$, $l \in L_j$, which are called *basic contrasts.*

Using (2.1) and (2.5) we obtain the unbiased estimator of $z'_{lj}\gamma$ in the form

$$z'_{lj}\gamma = \Big(\sum_{i=1}^{u} m_{il}\lambda_{ij}\Big)^{-1}(e'_l \otimes w'_j)\Delta\Phi y, \tag{2.7}$$

which, using (1.3) and (2.2), can be written as follows:

$$z'_{lj}\gamma = \Big(\sum_{i=1}^{u} m_{il}\lambda_{ij}\Big)^{-1}(e'_l \otimes w'_j)\sum_{i=1}^{u}(M_i \otimes \Delta'_i\Phi_i)\operatorname{vec} Y_i. \tag{2.8}$$

This equation shows that the estimator of $z'_{lj}\gamma$ is the linear combination of estimators of this function in the model (1.1) for $i = 1, 2, \ldots, u$.

Not in every model of the form (1.2) the best linear unbiased estimator of the parametric function $z'_{lj}\gamma$ exists. When this estimator exists, it is the simple least squares estimator (2.8). The condition of the existence of the best estimator of the parametric function in the general linear model was given by Walkowiak [7]. For the case considered here this condition is the following:

$$\mathscr{C}\left(\begin{bmatrix} M'_1\Sigma_{lg}M_1M'_1 \otimes \Phi_1\Delta'_1 \\ \cdots\cdots\cdots\cdots \\ M_u\Sigma_{lg}M_uM'_u \otimes \Phi_u\Delta'_u \end{bmatrix} z_{lj}\right) \subset \mathscr{C}\left(\begin{bmatrix} M'_1 \otimes \Phi_1\Delta'_1 \\ \cdots\cdots \\ M'_u \otimes \Phi_u\Delta'_u \end{bmatrix}\right) \tag{2.9}$$

for each $g \neq l$, $g, l = 1, 2, \ldots, t$, where $\mathscr{C}(\cdot)$ denotes the column space of the matrix $(\cdot)$, and $\Sigma_{lg} = (e_l + e_g)(e_l + e_g)'$ if the g-th and l-th variables are correlated, and $\Sigma_{lg} = 0$ if these variables are uncorrelated. From (2.9) it follows that the best unbiased estimator of $z'_{lj}\gamma$ exists in the model (1.2) if:

1. all variables are uncorrelated or
2. this model is complete, i.e., when $t_i = t$, $i = 1, 2, \ldots, u$, or
3. each of variables which is correlated with the l-th variable and observed together with it at least in one subset of experimental units is not observed in a subset without the l-th variable or if that variable is observed in such a set, say k, then $\lambda_{kj} = 0$.

THEOREM. *In the incomplete homogeneous multiresponse model* (1.2) *the simple least squares estimator of the function* $z'_{lj}\gamma$ *is of the form* (2.8). *This estimator is the best one if and only if the condition* (2.9) *holds.*

Due to (2.8), the unbiased estimator of the parametric function $z'\gamma$ having the form (2.6) may be given as follows:

$$z'\gamma = \sum_{j \in Z}\sum_{l \in L_j} g_{jl} z'_{lj}\gamma. \tag{2.10}$$

It is the best estimator of $z'\gamma$ when each basic contrast from the formula (2.6) has the best estimator in the considered model.

3. Example. To make these concepts clear we illustrate the above considerations with a simple example.

Consider an experiment carried out in a row-and-column design with

3 treatments (A, B, C) applied to $n = 30$ experimental units which are divided into two disjoint sets of $n_1 = 21$ and $n_2 = 9$. Units from the first and second set are arranged into 7 and 3 rows and into 7 and 3 columns, respectively. In the described experiment, 3 distinct variables are observed but for each n_1 unit the first and the second of the variables are observed and for each n_2 unit the second and the third ones, that is

$$M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Patterns of a dislocation of treatments and experimental units are the following:

–	*A*	*B*	–	–	*C*	–
	1	2			3	
–	–	*A*	*B*	–	–	*C*
		4	5			6
C	–	–	*A*	*B*	–	–
7			8	9		
–	*C*	–	–	*A*	*B*	–
	10			11	12	
–	–	*C*	–	–	*A*	*B*
		13			14	15
B	–	–	*C*	–	–	*A*
16			17			18
A	*B*	–	–	*C*	–	–
19	20			21		

A	*A*	*A*
22	23	24
B	*B*	*C*
25	26	27
C	*C*	*C*
28	29	30

Number of variables	Number of experimental units														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	16	13	12	16	15	13	11	17	15	12	18	16	14	18	17
2	4	6	6	5	7	7	5	4	8	6	5	9	7	7	9

Number of variables	Number of experimental units														
	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
1	14	13	19	17	14	15	–	–	–	–	–	–	–	–	–
2	7	8	6	5	8	8	4	5	6	8	9	8	6	6	7
3	–	–	–	–	–	–	50	47	49	43	42	51	54	51	53

It can be seen that matrices (1.4) take here the forms

$$C_1 = \frac{7}{3}\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}, \quad C_2 = \frac{4}{9}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}.$$

Eigenvalues of these matrices with respect to $X = I_3$ are

$$\lambda_{11} = 7, \quad \lambda_{12} = 7, \quad \lambda_{13} = 0, \quad \lambda_{21} = 8/9, \quad \lambda_{22} = 0, \quad \lambda_{23} = 0.$$

These eigenvalues correspond to common eigenvectors

$$w_1 = (1/\sqrt{2})[0 \quad 1 \quad -1]', \quad w_2 = (1/\sqrt{6})[-2 \quad 1 \quad 1]',$$
$$w_3 = (1/\sqrt{3})[1 \quad 1 \quad 1]'$$

of matrices C_i with respect to I_3 $(i = 1, 2)$. Hence we can say that the model of the considered experiment is homogeneous with respect to the identity matrix I_3.

Let us consider estimability of two parametric functions

$$([0 \quad 0 \quad 1]\otimes[1 \quad -1 \quad 0])\gamma \quad \text{and} \quad ([-1 \quad 0 \quad 1]\otimes[0 \quad 1 \quad -1])\gamma.$$

These functions may be written as follows:

$$\begin{aligned}([0 \quad 0 \quad 1]\otimes[1 \quad -1 \quad 0])\gamma &= -(\sqrt{2}/2)([0 \quad 0 \quad 1]\otimes w_1')\gamma \\ &\quad -(\sqrt{6}/2)([0 \quad 0 \quad 1]\otimes w_2')\gamma, \\ ([-1 \quad 0 \quad 1]\otimes[0 \quad 1 \quad -1])\gamma &= \sqrt{2}([0 \quad 0 \quad 1]\otimes w_1')\gamma \\ &\quad -\sqrt{2}([1 \quad 0 \quad 0]\otimes w_1')\gamma.\end{aligned}$$

From the relation (2.6) we know that the first of the above functions is not estimable because for $l = 3$ we have $m_{i3}\lambda_{12} = 0$ $(i = 1, 2)$. The second function is estimable and its estimator is equal to -10. It is the best unbiased estimator if and only if $\Sigma_{12} = \Sigma_{32} = 0$.

In the considered case there exist functions, e.g., the functions $(e_2'\otimes c')\gamma$, with the best unbiased estimators.

References

[1] T. Caliński, *On the notion of balance in block designs*, pp. 365–374 in: J. R. Barra et al. (eds.), *Recent Developments in Statistics*, North Holland 1977.

[2] — and M. Kozłowska, *Estimation of contrasts in incomplete multiresponse block design* (in Polish), Mat. Stos. 14 (1979), pp. 33–47.

[3] C. R. Rao and S. K. Mitra, *Generalized Inverses of Matrices and Its Applications*, J. Wiley, New York 1971.

[4] S. R. Searle, *Linear Models*, J. Wiley, New York 1971.

[5] I. N. Srivastava, *Incomplete multiresponse designs*, Sankhya A 26 (1966), pp. 377–388.

[6] — *On a general class of designs for multiresponse experiments*, Ann. Math. Statist. 39 (1968), pp. 1825–1843.
[7] R. Walkowiak, *Estimation of linear parametric functions in multiresponse models* (in Polish), Mat. Stos. 29 (1987), pp. 73–91.

DEPARTMENT OF MATHEMATICAL AND STATISTICAL METHODS
ACADEMY OF AGRICULTURE
60-637 POZNAŃ

Received on 1987.03.05;
revised version on 1987.12.08