Ann. Inst. Statist. Math.
35 (1983), Part A, 461-473

OPTIMAL ALLOCATION OF UNITS IN EXPERIMENTAL DESIGNS WITH HIERARCHICAL AND CROSS CLASSIFICATION

CZESŁAW STĘPNIAK

(Received Mar. 12, 1982; revised Mar. 12, 1983)

Abstract

Consider the class of random linear models induced by possible allocations of units in an experimental design with hierarchical or cross classification. Assuming a balanced model belongs to the class, it is shown that this model is optimal for estimation of mean.

1. Introduction and summary

Optimal allocation of units in an experimental design resolves itself into the comparison of linear models by linear estimation, or the comparison of linear *normal* models by sufficiency. A result of Torgersen [5] shows that the two orderings, of linear models and of corresponding linear normal models, are equivalent, providing the variance components are known. On the other hand, the comparison of linear models with unknown variance components is reduced to the same problem for known variance components (cf. Stępniak and Torgersen [4]).

One way random normal models with known variance components have been considered by DeGroot [1] and Stępniak [3]. It was shown that the balanced (or almost balanced if there are no balanced) one way random normal model is optimal. The same result may be obtained by comparison of one way random models with unknown variance components by linear estimation. In both cases the considerations are restricted, in fact, to only one parameter—just the mean.

Consider the class of random linear models induced by possible allocations of units in an experimental design with hierarchical or cross classification. Assuming a balanced model belongs to the class, it will be shown that this model is optimal for estimation of the mean.

AMS 1970 subject classifications: Primary 62 J05; Secondary 62 B15, 62 J10.
Key words and phrases: Comparison of models, hierarchical classification, cross classification, balanced model, optimality.

2. Some general results

Consider models where the expectation and the covariances of the observable random vector X depend linearly on some unknown parameters.

The assumption can be written in

$$\mathrm{E}\, X = A\beta \tag{1}$$

and

$$\mathrm{Cov}\, X = V\gamma = \sum_{i=1}^{q} \gamma_i V_i \,, \tag{2}$$

where X is observable $n\times 1$ random vector, A and V_i, $i=1, \cdots, q$, are known matrices of, respectively, $n\times p$ and $n\times n$, while $\beta=(\beta_1, \cdots, \beta_p)'$ and $\gamma=(\gamma_1, \cdots, \gamma_q)'$ are unknown column vectors. It is assumed that the prior possible values of β and γ constitute, respectively, the p-dimensional real linear space R^p, and a set Γ in the q-dimensional r.l.s. R^q.

It will be assumed that $V\gamma$ is nonnegative definite (n.n.d.) for all $\gamma \in \Gamma$. The integers p, q and the set Γ are fixed, but arbitrary, while n may vary from model to model.

We shall denote this structure by $L(A\beta, V\gamma\,;\, \gamma \in \Gamma)$. Say that X is subject to linear model $L(A\beta, V\gamma\,;\, \gamma \in \Gamma)$ if (1) and (2) hold. If $\Gamma=\{\gamma_0\}$ i.e. the case of known covariances we may write $L(A\beta, V)$, where $V=V\gamma_0$, instead of $L(A\beta, V\gamma\,;\, \gamma=\gamma_0)$.

If T is a matrix then T' and $\mathcal{C}(T)$ will denote, respectively, the transposition and the column space of T.

Suppose X and Y are subject, respectively, to $L(A\beta, V\gamma\,;\, \gamma \in \Gamma)$ and $L(B\beta, W\gamma\,;\, \gamma \in \Gamma)$. We shall then say that the model $L(A\beta, V\gamma\,;\, \gamma \in \Gamma)$ is at least as good as the model $L(B\beta, W\gamma\,;\, \gamma \in \Gamma)$ *w.r.t. linear estimation with squared risk* if for any function Ψ on $R^p\times\Gamma$ and for any estimator $b'Y$ there is an estimator $a'X$ such that

$$\mathrm{E}_{\beta,\gamma}\,(a'X-\Psi)^2 \leqq \mathrm{E}_{\beta,\gamma}\,(b'Y-\Psi)^2 \tag{3}$$

uniformly for $\beta \in R^p$ and $\gamma \in \Gamma$.

If this condition is satisfied then we shall write $L(A\beta, V\gamma\,;\, \gamma \in \Gamma) \succ L(B\beta, W\gamma\,;\, \gamma \in \Gamma)$.

Another way of ordering of linear models was presented by Stępniak and Torgersen [4]. Namely, model $L(A\beta, V\gamma\,;\, \gamma \in \Gamma)$ is said to be at least as good as the model $L(B\beta, W\gamma\,;\, \gamma \in \Gamma)$ *w.r.t. unbiased estimation with squared risk* if for any unbiased estimator $b'Y$ of a parameter $c'\beta$ there is an unbiased estimator $a'X$ of this parameter such that $\mathrm{Var}\,(a'X)$

$\leqq \mathrm{Var}(b'Y)$ for each $\gamma \in \Gamma$.

Now we shall show that the two orderings of linear models are equivalent.

LEMMA 1. *$L(A\beta, V\gamma; \gamma \in \Gamma) \succ L(B\beta, W\gamma; \gamma \in \Gamma)$ if and only if the first model is at least as good as the second w.r.t. unbiased estimation with squared risk.*

PROOF. Let X and Y are subject, respectively, to $L(A\beta, V\gamma; \gamma \in \Gamma)$ and $L(B\beta, W\gamma; \gamma \in \Gamma)$.

Sufficiency. Suppose the first model is at least as good as the second w.r.t. unbiased estimation. Then for each $b'Y$ there is $a'X$ such that $\mathrm{E}(a'X) = \mathrm{E}(b'Y)$ and $\mathrm{E}[a'X - \mathrm{E}(a'X)]^2 \leqq \mathrm{E}[b'Y - \mathrm{E}(b'Y)]^2$. Thus, by the identity $\mathrm{E}(b'Y - \Psi)^2 = \mathrm{E}[b'Y - \mathrm{E}(b'Y)]^2 + [\mathrm{E}(b'Y) - \Psi]^2$, $b'Y$ satisfies the condition (3).

Necessity. Let $L(A\beta, V\gamma; \gamma \in \Gamma) \succ L(B\beta, W\gamma; \gamma \in \Gamma)$. We only need to show that if $b'Y$ is an unbiased estimator of $c'\beta$ and $a'X$ satisfies the condition (3) then $a'X$ is also unbiased. Suppose not. Then $\mathrm{E}(a'X) - c'\beta = k'\beta$, where $k \neq 0$. Thus, given $\gamma \in \Gamma$, the risk $\mathrm{E}(b'Y - c'\beta)^2 = \mathrm{Var}(b'Y)$ is bounded, whereas the risk $\mathrm{E}(a'X - c'\beta)^2 = \mathrm{Var}(a'X) + (k'\beta)^2$ is not. This contradicts $L(A\beta, V\gamma; \gamma \in \Gamma) \succ L(B\beta, W\gamma; \gamma \in \Gamma)$ and completes the proof.

We need the following theorems.

THEOREM 1. *$L(A\beta, V\gamma; \gamma \in \Gamma) \succ L(B\beta, W\gamma; \gamma \in \Gamma)$ if and only if $L(A\beta, V\gamma) \succ L(B\beta, W\gamma)$ for all γ belonging to the convex hull of Γ.*

THEOREM 2. *$L(A\beta, V) \succ L(B\beta, W)$ if and only if there is a matrix G such that $B = GA$ and $W - GVG'$ is n.n.d.*

Theorem 1 reduces the comparison of linear models with unknown variance components to the same problem for known variance components. This theorem was established by Stępniak and Torgersen [4].

Theorem 2 was proved by Stępniak [2] under the assumption that V and W are non singular. In the presented form it was stated by Torgersen [5]. For completeness we shall give a short

PROOF OF THEOREM 2. Let X and Y are subject, respectively, to $L(A\beta, V)$ and $L(B\beta, W)$.

Sufficiency. Suppose $B = GA$ and $W - GVG'$ is n.n.d. Then

$$\begin{aligned}\mathrm{E}(b'Y - \Psi)^2 &= b'Wb + (b'B\beta - \Psi)^2 \\ &\geqq b'GVG'b + (b'B\beta - \Psi)^2 \\ &= b'GVG'b + (b'GA\beta - \Psi)^2 \\ &= \mathrm{E}(a'X - \Psi)^2,\end{aligned}$$

where $a=G'b$.

Necessity. First we notice that $\{a'X: Va \in \mathcal{C}(A)\}$ is the set of all best linear unbiased estimators (BLUE's) in $L(A\beta, V)$ (cf. Zyskind [6]). Denote by P the orthogonal projector on $\{a: Va \in \mathcal{C}(A)\}$. Then, by Lemma 1, $L(PA\beta, PVP) \succ L(A\beta, V)$. In particular

$$\mathcal{C}(A') \subseteqq \mathcal{C}(A'P). \tag{4}$$

Suppose $L(A\beta, V) \succ L(B\beta, W)$. Then, by Lemma 1, $\mathcal{C}(A') \supseteqq \mathcal{C}(B')$. Next, as a consequence of this, $B=DA$ for some matrix D. Thus, by (4), $B=DCPA$ for some matrix C. Putting $G=DCP$ we satisfy the conditions $B=GA$ and

$$\mathcal{C}(VG') \subseteqq \mathcal{C}(A). \tag{5}$$

Suppose, by contradiction, that $W-GVG'$ is not n.n.d. Then there is a vector b such that

$$b'(W-GVG')b<0. \tag{6}$$

Let $\Psi=b'B$. It follows that $b'Y$ is an unbiased estimator of Ψ and, by (6),

$$\mathrm{Var}\,(b'Y)=b'Wb<b'GVG'b=\mathrm{Var}\,(a'X),$$

where $a=G'b$. On the other hand, by (5), $a'X$ is a BLUE of Ψ. This contradicts to $L(A\beta, V) \succ L(B\beta, W)$ and completes the proof.

3. Optimal allocation of experimental units in hierarchical classification

Formal definition of allocation will be preceded by an example. Suppose 8 experimental units are submitted to a two-stage classification. The first stage of the classification is more general and includes 2 subclasses, say $S_{11}=\{1, 2, 3, 4, 5\}$ and $S_{12}=\{6, 7, 8\}$, while the second stage is more detailed and includes 4 subclasses, say $S_{21}=\{1, 2\}$, $S_{22}=\{3, 4, 5\}$, $S_{23}=\{6\}$ and $S_{24}=\{7, 8\}$. Then the observations $X_1, \cdots, X_8$ corresponding to the experimental units may be presented in the form

$$\begin{aligned}
X_1&=\mu+a_{11}+a_{21}+e_1\\
X_2&=\mu+a_{11}+a_{21}+e_2\\
X_3&=\mu+a_{11}+a_{22}+e_3\\
X_4&=\mu+a_{11}+a_{22}+e_4\\
X_5&=\mu+a_{11}+a_{22}+e_5
\end{aligned}$$

$$X_6=\mu \quad +a_{12} \quad +a_{23} \quad +e_6$$
$$X_7=\mu \quad +a_{12} \quad +a_{24}+e_7$$
$$X_8=\mu \quad +a_{12} \quad +a_{24}+e_8\,,$$

where μ is the general mean, a_{1j}, $j=1, 2$, are the effects of the first classification, a_{2j}, $j=1, 2, 3, 4$, are the effects of the second classification, and e_j, $j=1, \cdots, 8$, are the experimental errors.

Assuming all these effects are independent random variables with the expectations zero and the variances

$$\mathrm{Var}\,(e_j) = \gamma_0\,, \qquad j=1, \cdots, 8\,,$$
$$\mathrm{Var}\,(a_{1j}) = \gamma_1\,, \qquad j=1, 2\,,$$
$$\mathrm{Var}\,(a_{2j}) = \gamma_2\,, \qquad j=1, 2, 3, 4\,,$$

we can see that the random vector $X=(X_1, \cdots, X_8)'$ is subject to the linear model

$$L(1_8\mu,\ \gamma_0 I_8+\gamma_1 V_1+\gamma_2 V_2;\ \gamma_0>0,\ \gamma_1, \gamma_2 \geqq 0)\,,$$

where 1_8 is the column of 8 ones, μ is an unknown scalar, and V_i, $i=1, 2$, are 8×8 matrices of the form

$$V_1=\mathrm{diag}\,(1_5 1_5', 1_3 1_3')\,,$$

and

$$V_2=\mathrm{diag}\,(1_2 1_2', 1_3 1_3', 1, 1_2 1_2')\,.$$

Now let n, q and $k_1, \cdots, k_q$ be positive integers such that $q \leqq n$ and $k_1 \leqq k_2 \leqq \cdots \leqq k_q \leqq n$. Moreover let $n_{11}, \cdots, n_{1k_1}$; $n_{21}, \cdots, n_{2k_2}$; $\cdots$; $n_{q1}, \cdots, n_{qk_q}$ be non negative integers such that

$$(7) \qquad \sum_{j=1}^{k_i} n_{ij}=n\,, \qquad i=1, \cdots, q\,,$$

and

$$(8) \qquad \mathrm{diag}\,(1_{n_{i1}}1'_{n_{i1}}, \cdots, 1_{n_{ik_i}}1'_{n_{ik_i}})-\mathrm{diag}\,(1_{n_{i+1,1}}1'_{n_{i+1,1}}, \cdots, 1_{n_{i+1,k_{i+1}}}1'_{n_{i+1,k_{i+1}}})\,,$$

$i=1, \cdots, q-1$, are matrices with non negative elements.

Here n denotes the total number of experimental units, q is the number of classifications, k_i, $i=1, \cdots, q$, is the number of subclasses in the i-th classification, and n_{ij}, $i=1, \cdots, q$; $j=1, \cdots, k_i$, is the number of experimental units in the j-th subclass of the i-th classification.

Any choice of such integers is called an allocation of n experimental units in q-stage hierarchical classification with $k_1, \cdots, k_q$ subclasses and is denoted by $H(n, q;\ k_1, \cdots, k_q | n_{ij},\ i=1, \cdots, q;\ j=1, \cdots, k_i)$.

The class of all allocations of not more than n units in q-stage hierarchical classification with not more than $k_1, \cdots, k_q$ subclasses will be denoted by $\mathcal{H}(n, q;\ k_1, \cdots, k_q)$.

To each allocation $H(n, q;\ k_1, \cdots, k_q | n_{ij},\ i=1, \cdots, q;\ j=1, \cdots, k_i)$ corresponds a linear model

$$L\Big(1_n\mu,\ \gamma_0 I_n + \sum_{i=1}^{q} \gamma_i V_i;\ \gamma_0 > 0,\ \gamma_i \geqq 0,\ i=1, \cdots, q\Big),$$

where 1_n is the column of n ones, μ is an unknown scalar, and V_i, $i=1, \cdots, q$, are $n \times n$ matrices of the form

$$V_i = \operatorname{diag}(1_{n_{i1}} 1'_{n_{i1}}, \cdots, 1_{n_{ik_i}} 1'_{n_{ik_i}}).$$

Now let the numbers n and $k_1, \cdots, k_q$ satisfy the conditions $n = k_i r_i$, $i=1, \cdots, q$, for some integers $r_1, \cdots, r_q$. Then we can put

$$(9) \qquad n_{ij} = r_i, \qquad i=1, \cdots, q;\ j=1, \cdots, k_i.$$

An allocation $H(n, q;\ k_1, \cdots, k_q | n_{ij},\ i=1, \cdots, q;\ j=1, \cdots, k_i)$ satisfying the conditions (9) for some integers $r_1, \cdots, r_q$ is said to be *balanced*. We note that for the balanced allocation

$$V_i = I_{k_i} \otimes 1_{r_i} 1'_{r_i}, \qquad i=1, \cdots, q,$$

where $\otimes$ denotes the Kronecker product of matrices.

We shall write $H(n, q;\ k_1, \cdots, k_q | n_{ij},\ i=1, \cdots, q;\ j=1, \cdots, k_i) \geqq H(m, q;\ l_1, \cdots, l_q | m_{ij},\ i=1, \cdots, q;\ j=1, \cdots, l_i)$ if the linear model corresponding to the first allocation is at least as good as the model corresponding to the second allocation.

An allocation $H(n, q;\ k_1, \cdots, k_q | n_{ij},\ i=1, \cdots, q;\ j=1, \cdots, k_i)$ is said to be *optimal* in the class $\mathcal{H}(n, q;\ k_1, \cdots, k_q)$ if

$$H(n, q;\ k_1, \cdots, k_q | n_{ij},\ i=1, \cdots, q;\ j=1, \cdots, k_i)$$
$$\geqq H(m, q;\ l_1, \cdots, l_q | m_{ij},\ i=1, \cdots, q;\ j=1, \cdots, l_i)$$

for all $H(m, q;\ l_1, \cdots, l_q | m_{ij},\ i=1, \cdots, q;\ j=1, \cdots, l_i)$ being members of $\mathcal{H}(n, q;\ k_1, \cdots, k_q)$.

THEOREM 3. *Let n and $k_1, \cdots, k_q$ be positive integers such that $n = k_i r_i$, $i=1, \cdots, q$, for some integers $r_1, \cdots, r_q$. Then the balanced allocation $H(n, q;\ k_1, \cdots, k_q | n_{ij} = r_i,\ i=1, \cdots, q;\ j=1, \cdots, k_i)$ is optimal in the class $\mathcal{H}(n, q;\ k_1, \cdots, k_q)$.*

The proof of the theorem will be preceded by

LEMMA 2. *Let k, l, m and $m_1, \cdots, m_l$ be positive integers such that*

$l \leqq k$ and $\sum_{i=1}^{l} m_i = m$. Then the matrix

$$D = \operatorname{diag}(1_{m_1}1'_{m_1}, \cdots, 1_{m_l}1'_{m_l}) - \frac{1}{k}1_m 1'_m \tag{10}$$

is n.n.d.

PROOF OF THE LEMMA. Note that D is a pivotal submatrix of the matrix

$$I_l \otimes 1_r 1'_r - \frac{1}{k}1_{lr}1'_{lr} = \left(I_l - \frac{1}{k}1_l 1'_l\right) \otimes 1_r 1'_r\,, \tag{11}$$

where $r = \max\{m_1, \cdots, m_l\}$. Since $l \leqq k$ and $I_l - \frac{1}{l}1_l 1'_l$ is n.n.d., the matrix (11) is n.n.d. too. This implies the desired result.

PROOF OF THE THEOREM. Let $G = \frac{1}{n}1_m 1'_n$. Note that G satisfies the condition $G1_n = 1_m$. Thus, by Theorems 1 and 2, we only need to show that the matrix

$$\begin{aligned} &\gamma_0 I_m + \sum_{i=1}^{q} \gamma_i \operatorname{diag}(1_{m_{i1}}1'_{m_{i1}}, \cdots, 1_{m_{il_i}}1'_{m_{il_i}}) - G\left[\gamma_0 I_n + \sum_{i=1}^{q} \gamma_i (I_{k_i} \otimes 1_{r_i}1'_{r_i})\right]G' \\ &= \gamma_0[I_m - GG'] + \sum_{i=1}^{q} \gamma_i[\operatorname{diag}(1_{m_{i1}}1'_{m_{i1}}, \cdots, 1_{m_{il_i}}1'_{m_{il_i}}) - G(I_{k_i} \otimes 1_{r_i}1'_{r_i})G'] \end{aligned} \tag{12}$$

is n.n.d. for all $\gamma_0 > 0$ and $\gamma_i \geqq 0$, $i = 1, \cdots, q$.

As $l_i \leqq k_i$, $i = 1, \cdots, q$, we can write the matrix (12) in the form $\gamma_0(I_n - GG') + \sum_{i=1}^{q} \gamma_i D_i$, where D_i, $i = 1, \cdots, q$, are some matrices of the form (10), and hence are n.n.d.

To complete the proof we only need to see that, under the condition $m \leqq n$, the matrix

$$I_m - GG' = I_m - \frac{1}{n}1_m 1'_m$$

is n.n.d.

Applying the theorem to the example considered before we obtain that the balanced allocation $H(8, 2;\ 2, 4|4, 4;\ 2, 2, 2, 2)$ is optimal in the class $\mathcal{H}(8, 2;\ 2, 4)$ i.e. in the class of all allocations of 8 units in two-stage hierarchical classification with not more than 2 subclasses on the first stage and not more than 4 subclasses on the second stage.

4. Optimal allocation of experimental units in cross classification

As in Section 3, we start from example. Suppose 24 experimental units are submitted to tree independent classifications with 2, 3 and 2 subclasses, respectively. Denote by S_{ij} the set of experimental units belonging to the j-th subclass in the i-th classification. For example

$$S_{11}=\{1, 2, 5, 8, 10, 13, 18, 20, 21, 22, 23\}$$

$$S_{12}=\{3, 4, 6, 7, 9, 11, 12, 14, 15, 16, 17, 19, 24\}$$

$$S_{21}=\{4, 6, 8, 15, 16, 17, 18\}$$

$$S_{22}=\{10, 11, 12, 13, 14, 20\}$$

$$S_{23}=\{1, 2, 3, 5, 7, 9, 19, 21, 22, 23, 24\}$$

$$S_{31}=\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14\}$$

$$S_{32}=\{15, 16, 17, 18, 19, 20, 21, 22, 23, 24\}\,.$$

To each S_{ij} is defined a column vector $N_{ij}=(n_1, \cdots, n_{24})'$ of zeros and ones, where

$$n_k=\begin{cases} 1 & \text{if } k \in S_{ij} \\ 0 & \text{otherwise}\,, \end{cases}$$

for $k=1, \cdots, 24$. In our example

$$N_{11}=(1, 1, 0, 0, 1, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 0, 1, 0, 1, 1, 1, 1, 0)'$$

$$N_{12}=(0, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 0, 0, 0, 0, 1)'$$

$$N_{21}=(0, 0, 0, 1, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0)'$$

$$N_{22}=(0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0)'$$

$$N_{23}=(1, 1, 1, 0, 1, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 1, 1, 1)'$$

$$N_{31}=(1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)'$$

$$N_{32}=(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)'\,.$$

By using the Hadamard product $*$ of vectors, defined as

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} * \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 y_1 \\ \vdots \\ x_n y_n \end{bmatrix},$$

the observation vector $X=(X_1, \cdots, X_{24})'$, corresponding to the experimental units may be presented in the form

$$X=1_{24}\mu+\sum_{j=1}^{2} N_{1j}a_j+\sum_{j=1}^{3} N_{2j}b_j+\sum_{j=1}^{2} N_{3j}c_j$$

$$+\sum_{\substack{j_1=1,2\\ j_2=1,2,3}} N_{1j_1} * N_{2j_2}(ab)_{j_1j_2} + \sum_{\substack{j_1=1,2\\ j_3=1,2}} N_{1j_1} * N_{3j_3}(ac)_{j_1j_3}$$
$$+\sum_{\substack{j_2=1,2,3\\ j_3=1,2}} N_{2j_2} * N_{3j_3}(bc)_{j_2j_3} + \sum_{\substack{j_1=1,2\\ j_2=1,2,3\\ j_3=1,2}} N_{1j_1} * N_{2j_2} * N_{3j_3}(abc)_{j_1j_2j_3}$$

where μ is the general mean, a_{j_1}, b_{j_2}, and c_{j_3}, $j_1=1, \cdots, k_1$; $j_2=1, \cdots, k_2$; $j_3=1, \cdots, k_3$, are the effects of the respective subclasses in the first, second and third classification, while $(ab)_{j_1j_2}$, $(ac)_{j_1j_3}$, $(bc)_{j_2j_3}$ and $(abc)_{j_1j_2j_3}$, $j_1=1, \cdots, k_1$; $j_2=1, \cdots, k_2$; $j_3=1, \cdots, k_3$, are the effects of double and triple interactions, and $e=(e_1, \cdots, e_{24})'$ is the vector of the experimental errors.

Assuming all these effects are independent random variables with the expectations zero and the variances

$$\begin{aligned}
&\mathrm{Var}\,(e_j) = \gamma_0 && j=1, \cdots, 24\,,\\
&\mathrm{Var}\,(a_{j_1}) = \gamma_1 && j_1=1, 2\,,\\
&\mathrm{Var}\,(b_{j_2}) = \gamma_2 && j_2=1, 2, 3\,,\\
&\mathrm{Var}\,(c_{j_3}) = \gamma_3 && j_3=1, 2\,,\\
&\mathrm{Var}\,((ab)_{j_1j_2}) = \gamma_{12}\,, && j_1=1, 2\,;\ j_2=1, 2, 3\,,\\
&\mathrm{Var}\,((ac)_{j_1j_3}) = \gamma_{13}\,, && j_1=1, 2\,;\ j_3=1, 2\,,\\
&\mathrm{Var}\,((bc)_{j_2j_3}) = \gamma_{23}\,, && j_2=1, 2, 3\,;\ j_3=1, 2\,,\\
&\mathrm{Var}\,((abc)_{j_1j_2j_3}) = \gamma_{123}\,, && j_1=1, 2\,;\ j_2=1, 2, 3\,;\ j_3=1, 2\,,
\end{aligned}$$

we can see that the random vector X is subject to the linear model

$$L\Big(1_{24}\mu,\ \gamma_0 I_{24} + \sum_{i=1}^{3} \gamma_i V_i + \sum_{i_1<i_2} \gamma_{i_1i_2} V_{i_1i_2} + \gamma_{123} V_{123};\ \gamma_0>0,\ \gamma_{i_1i_2}\geqq 0,\ \gamma_{123}\geqq 0\Big)\,,$$

where

$$V_1 = \sum_{j=1}^{2} N_{1j} N'_{1j}\,,$$

$$V_2 = \sum_{j=1}^{3} N_{2j} N'_{2j}\,,$$

$$V_3 = \sum_{j=1}^{2} N_{3j} N'_{3j}\,,$$

$$V_{12} = \sum_{\substack{j_1=1,2\\ j_2=1,2,3}} (N_{1j_1} * N_{2j_2})(N_{1j_1} * N_{2j_2})'\,,$$

$$V_{13} = \sum_{\substack{j_1=1,2\\ j_3=1,2}} (N_{1j_1} * N_{3j_3})(N_{1j_1} * N_{3j_3})'\,,$$

$$V_{23}=\sum_{\substack{j_2=1,2,3\\ j_3=1,2}}(N_{2j_2}*N_{3j_3})(N_{2j_2}*N_{3j_3})',$$

$$V_{123}=\sum_{\substack{j_1=1,2\\ j_2=1,2,3\\ j_3=1,2}}(N_{1j_1}*N_{2j_2}*N_{3j_3})(N_{1j_1}*N_{2j_2}*N_{3j_3})'.$$

Now let n, q and $k_1, \cdots, k_q$ be positive integers such that $q \leqq n$ and $k_i \leqq n$, $i=1, \cdots, q$. Moreover, let N_{ij}, $i=1, \cdots, q$; $j=1, \cdots, k_i$, be n-dimensional columns of zeros and ones such that

$$\sum_{j=1}^{k_i} N_{ij}=1_n, \qquad i=1, \cdots, q \tag{13}$$

and

$$N'_{ij}N_{ij'}=0, \qquad i=1, \cdots, q;\ j\neq j'. \tag{14}$$

Any choice of such columns is called an allocation of n experimental units in q-way cross classification with $k_1, \cdots, k_q$ subclasses and is denoted by $C(n, q;\ k_1, \cdots, k_q | N_{ij},\ i=1, \cdots, q;\ j=1, \cdots, k_i)$.

The class of all allocations of not more than n units in q-way cross classification with not more than $k_1, \cdots, k_q$ subclasses will be denoted by $C(n, q;\ k_1, \cdots, k_q)$.

To each allocation $C(n, q;\ k_1, \cdots, k_q | N_{ij},\ i=1, \cdots, q;\ j=1, \cdots, k_i)$ corresponds a linear model

$$L\Big(1_n\mu,\ \gamma_0 I_n+\sum_{i=1}^{q}\gamma_i V_i+\sum_{i_1<i_2}\gamma_{i_1i_2}V_{i_1i_2}+\cdots+\sum_{i_1<i_2<\cdots<i_{q-1}}\gamma_{i_1i_2\cdots i_{q-1}}V_{i_1i_2\cdots i_{q-1}}$$
$$+\gamma_{1\cdots q}V_{1\cdots q};\ \gamma_0>0,\ \gamma_i\geqq 0,\ \gamma_{i_1i_2}\geqq 0,\ \cdots,\ \gamma_{1\cdots q}\geqq 0\Big),$$

where

$$V_i=\sum_{j=1}^{k_i}N_{ij}N'_{ij},$$

$$V_{i_1i_2}=\sum_{\substack{j_1=1,\cdots,k_{i_1}\\ j_2=1,\cdots,k_{i_2}}}(N_{i_1j_1}*N_{i_2j_2})(N_{i_1j_1}*N_{i_2j_2})',$$

$$\vdots$$

$$V_{i_1\cdots i_{q-1}}=\sum_{\substack{j_1=1,\cdots,k_{i_1}\\ \vdots\\ j_{q-1}=1,\cdots,k_{i_{q-1}}}}(N_{i_1j_1}*\cdots*N_{i_{q-1}j_{q-1}})(N_{i_1j_1}*\cdots*N_{i_{q-1}j_{q-1}})',$$

$$V_{1\cdots q}=\sum_{\substack{j_1=1,\cdots,k_1\\ \vdots\\ j_q=1,\cdots,k_q}}(N_{1j_1}*\cdots*N_{qj_q})(N_{1j_1}*\cdots*N_{qj_q})'.$$

Now let the integers n and $k_1, \cdots, k_q$ satisfy the condition $n=r\prod_{i=1}^{q}k_i$ for some integer r. Then we can put

$$
(15)\qquad
\begin{aligned}
N_{1j} &= E_{1j} \otimes 1_{k_2} \otimes \cdots \otimes 1_{k_q} \otimes 1_r\,, & j&=1, \cdots, k_1\,,\\
N_{2j} &= 1_{k_1} \otimes E_{2j} \otimes \cdots \otimes 1_{k_q} \otimes 1_r\,, & j&=1, \cdots, k_2\,,\\
&\vdots\\
N_{qj} &= 1_{k_1} \otimes 1_{k_2} \otimes \cdots \otimes E_{qj} \otimes 1_r\,, & j&=1, \cdots, k_q\,,
\end{aligned}
$$

where E_{ij} is k_i-dimensional column vector with one on the j-th place and zeros besides.

Any allocation $C(n, q;\ k_1, \cdots, k_q | N_{ij},\ i=1, \cdots, q;\ j=1, \cdots, k_i)$ satisfying the condition (15) for some integer r is said to be *balanced.*

Refering to the example considered before, we note that the balanced allocation of 24 units in 2, 3 and 2 subclasses, respectively, is defined by

$$N_{11}=(1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)'$$

$$N_{12}=(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)'$$

$$N_{21}=(1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0)'$$

$$N_{22}=(0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0)'$$

$$N_{23}=(0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1)'$$

$$N_{31}=(1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0)'$$

$$N_{32}=(0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1, 0, 0, 1, 1)'\,.$$

We shall write $C(n, q;\ k_1, \cdots, k_q | N_{ij},\ i=1, \cdots, q;\ j=1, \cdots, k_i) \geqq C(m, q;\ l_1, \cdots, l_q | M_{ij},\ i=1, \cdots, q;\ j=1, \cdots, k_i)$ if the linear model corresponding to the first allocation is at least as good as the linear model corresponding to the second allocation.

An allocation $C(n, q;\ k_1, \cdots, k_q | N_{ij},\ i=1, \cdots, q;\ j=1, \cdots, k_i)$ is said to be *optimal* in the class $C(n, q;\ k_1, \cdots, k_q)$ if

$$
\begin{aligned}
&C(n, q;\ k_1, \cdots, k_q | N_{ij},\ i=1, \cdots, q;\ j=1, \cdots, k_i)\\
&\quad \geqq C(m, q;\ l_1, \cdots, l_q | M_{ij},\ i=1, \cdots, q;\ j=1, \cdots, l_i)
\end{aligned}
$$

for all $C(m, q;\ l_1, \cdots, l_q | M_{ij},\ i=1, \cdots, q;\ j=1, \cdots, l_i)$ being members of $C(n, q;\ k_1, \cdots, k_i)$.

THEOREM 4. *Let n and $k_1, \cdots, k_q$ be positive integers such that $n= r \prod_{i=1}^{q} k_i$ for some integer r. Then the balanced allocation $C(n, q;\ k_1, \cdots, k_q | N_{ij},\ i=1, \cdots, q;\ j=1, \cdots, k_i)$, where N_{ij} is defined by* (15), *is optimal in the class $C(n, q;\ k_1, \cdots, k_q)$.*

PROOF. As in Theorem 3, we only need to show that the matrices $I_n - GG'$ and

(16) $$\sum_{j=1}^{l_i} M_{ij}M'_{ij} - GV_iG' , \qquad i=1, \cdots, q ,$$

(17) $$\sum_{\substack{j_1=1,\cdots,l_{i_1} \\ j_2=1,\cdots,l_{i_2}}} (M_{i_1j_1} * M_{i_2j_2})(M_{i_1j_1} * M_{i_2j_2})' - GV_{i_1i_2}G' ,$$
$$\vdots$$

(18) $$\sum_{\substack{j_1=1,\cdots,l_1 \\ \vdots \\ j_q=1,\cdots,l_q}} (M_{1j_1} * \cdots * M_{qj_q})(M_{1j_1} * \cdots * M_{qj_q})' - GV_{1\cdots q}G'$$

are n.n.d.

It was already shown that $I_n - GG'$ is n.n.d. To do the same for the other matrices we shall simplify them by simultaneous permutations of their rows and columns. We use the fact that any such permutation preserves the property of "being n.n.d.".

First, by setting the experimental units in order according to the subclasses of the i-th classification, $i=1, \cdots, q$, we reduce the matrix (16) to the form

(16′) $$\text{diag}\,(1_{m_1}1'_{m_1}, \cdots, 1_{m_{l_i}}1'_{m_{l_i}}) - \frac{1}{k_i} 1_m 1'_m$$

where m_j, $j=1, \cdots, l_i$, is the number of ones in the column M_{ij}.

Similarly, by setting the experimental units in order according to the subclasses of the classifications i_1 and i_2, the matrix (17) may be reduced to

(17′) $$\text{diag}\,(1_{m_{11}}1'_{m_{11}}, \cdots, 1_{m_{l_{i_1}l_{i_2}}}1'_{m_{l_{i_1}l_{i_2}}}) - \frac{1}{k_{i_1}k_{i_2}} 1_m 1'_m ,$$

where $m_{j_1j_2}$, $j_1=1, \cdots, l_{i_1}$; $j_2=1, \cdots, l_{i_2}$, is the number of ones in the column $M_{i_1j_1} * M_{i_2j_2}$.

Repeating this procedure we reduce, finally, the matrix (18) to

(18′) $$\text{diag}\,(1_{m_{1\cdots1}}1'_{m_{1\cdots1}}, \cdots, 1_{m_{l_1\cdots l_q}}1'_{m_{l_1\cdots l_q}}) - \frac{1}{k_1\cdots k_q} 1_m 1'_m ,$$

where $m_{j_1\cdots j_q}$, $j_i=1, \cdots, k_i$; $i=1, \cdots, q$, is the number of ones in the column $M_{1j_1} * \cdots * M_{qj_q}$.

It follows from Lemma 2 and from the assumptions of the theorem that the matrices (16′)–(18′) are n.n.d. This implies the desired result.

Acknowledgments

I am grateful to an anonymous person for his meticulous reading of an earlier version of the paper and for his very helpful suggestions. In fact, Lemma 3 is due to him. By the lemma the proofs of Theorems

3 and 4 could be decisively shortened.

Thanks are due also to the referee for suggesting improvements in the presentation of Sections 3 and 4.

AGRICULTURAL UNIVERSITY, LUBLIN

REFERENCES

[1] DeGroot, M. H. (1966). Optimal allocation of observations, *Ann. Inst. Statist. Math.*, **18**, 13-28.

[2] Stępniak, C. (1977). On comparison of linear models, VII Coll. Agro-Biometr., Academy of Agriculture, Lublin (In Polish).

[3] Stępniak, C. (1982). Optimal allocation of observations in one way random normal model, *Ann. Inst. Statist. Math.*, **34**, 175-181.

[4] Stępniak, C. and Torgersen, E. (1981). Comparison of linear models with partially known covariances with respect to unbiased estimation, *Scand. J. Statist.*, **8**, 183-184.

[5] Torgersen, E. (1980). Orderings of linear models, *Statist. Res. Report.*, Univ. of Oslo.

[6] Zyskind, G. (1967). On canonical forms, non-negative covariance matrices and best and simple least squares estimators in linear models, *Ann. Math. Statist.*, **38**, 1092-1109.