# On the General Gauss−Markov Model for Experiments in Block Designs

## Tadeusz Caliński

Department of Mathematical and Statistical Methods, Poznań University of Life Sciences, Wojska Polskiego 28, 60-637 Poznań, e-mail: calinski@up.poznan.pl

Dedicated to the memory of Professor Wiktor Oktaba

### Summary

The main estimation and hypothesis testing results related to the Gauss−Markov model, in its general form, are recalled and the application of these results to the analysis of experiments in block designs is considered. Special attention is given to the randomization-derived model for a general block design, and for a proper block design in particular. The question whether the randomization-derived model can be considered as a particular general Gauss−Markov model is discussed. It is indicated that the former, as a mixed model, is in fact an extension of the general Gauss−Markov model. Thus, the analysis based on the randomization-derived model requires a more extended methodical approach. The present paper has been inspired by one of the last papers of Professor Wiktor Oktaba.

**Key words:** block designs, estimation, general Gauss−Markov model, hypothesis testing, mixed model, randomization-derived model.

## 1. Introduction

As recalled by Oktaba (2003), methods of linear statistical inference are related to the Gauss−Markov model, originally considered by Gauss (1809, 1855) and Markov (1900) in a simple form, later extended and studied by many prominent statisticians (see also Rao, 1971, Section 1.1). One of the generalizations of particular interest is the so-called general uni- and multivariate Gauss−Markov model on which the analyses of experiments are based. Various aspects of this general model were studied by Oktaba (e.g., 1984, 1989, 1996, 1998, 2003) and Oktaba, Kornacki and Wawrzosek (1986, 1988). For practical applications of special interest is the adoption of the theory based on the general Gauss−Markov model to some common

experimental designs. In his last paper devoted to that theory, Oktaba (2003) drew attention to its application to the analysis of experimental data from an incomplete block design. The purpose of the present paper is to continue his considerations in this direction.

## 2. Main results under the general Gauss−Markov model

A linear model for an $n \times 1$ random vector $\boldsymbol{y}$ with $\mathrm{E}(\boldsymbol{y}) = \boldsymbol{\theta} = \boldsymbol{X\beta}$ and $\mathrm{D}(\boldsymbol{y}) = \sigma^2 \boldsymbol{V}$, where $\boldsymbol{V}$ is a known n.n.d. matrix (p.d. in particular) is usually denoted by the triplet

$$\{\boldsymbol{y}, \boldsymbol{X\beta}, \sigma^2 \boldsymbol{V}\} \tag{2.1}$$

and called the general Gauss−Markov model (see, e.g., Oktaba, 1984). It will be always assumed here that it is consistent, i.e., that $\boldsymbol{y} \in \mathcal{R}[\boldsymbol{X} : \boldsymbol{V}]$, where $\mathcal{R}[\boldsymbol{X} : \boldsymbol{V}]$ denotes the range space of the indicated matrix. (See also Rao, 1973, Section 4i.2.)

### 2.1. Main estimation results

**Theorem 2.1** (3 in Zyskind, 1967). Under the assumptions of the general Gauss−Markov model (2.1), a linear function $\boldsymbol{a'y}$ is the best linear unbiased estimator (BLUE) of its expectation, $\boldsymbol{a'X\beta}$, if and only if the condition $(\boldsymbol{I}_n - \boldsymbol{P}_X)\boldsymbol{Va} = \boldsymbol{0}$ holds, where $\boldsymbol{P}_X$ denotes the orthogonal projector on $\mathcal{R}(\boldsymbol{X})$, i.e., if and only if the vector $\boldsymbol{Va}$ belongs to $\mathcal{R}(\boldsymbol{X})$.

**Theorem 2.2** (4 in Zyskind, 1967). Under the assumptions of the general Gauss−Markov model (2.1), a linear function $\boldsymbol{a'y}$ is both the simple least squares estimator (SLSE) and the BLUE of its expectation, $\boldsymbol{a'X\beta}$, if and only if the conditions $(\boldsymbol{I}_n - \boldsymbol{P}_X)\boldsymbol{a} = \boldsymbol{0}$ and $(\boldsymbol{I}_n - \boldsymbol{P}_X)\boldsymbol{Va} = \boldsymbol{0}$ hold simultaneously, i.e., if and only if both vectors $\boldsymbol{a}$ and $\boldsymbol{Va}$ belong to $\mathcal{R}(\boldsymbol{X})$.

**Corollary 2.1** (4.2 in Zyskind, 1967). Under the assumptions of Theorem 2.2, if $\boldsymbol{a'y}$ is the SLSE of $\mathrm{E}(\boldsymbol{a'y}) = \boldsymbol{a'X\beta}$, then it is also the BLUE of $\boldsymbol{a'X\beta}$ if and only if $(\boldsymbol{I}_n - \boldsymbol{P}_X)\boldsymbol{Va} = \boldsymbol{0}$, i.e., if and only if $\boldsymbol{Va}$ belongs to $\mathcal{R}(\boldsymbol{X})$.

**Corollary 2.2** (3 in Section 8.2 of Rao and Mitra, 1971). Under the assumptions of Theorem 2.2, any SLSE of its expectation is also the BLUE of that if and only if, for any vector $\boldsymbol{s}$, $\boldsymbol{VXs}$ belongs to $\mathcal{R}(\boldsymbol{X})$, i.e., if and only if

$$\mathcal{R}(\boldsymbol{VX}) \subset \mathcal{R}(\boldsymbol{X}).$$

(See also Theorem 1.1 in Baksalary and Kala, 1983.)

**Theorem 2.3** (from Section 2 in Baksalary and Puntanen, 1990). Under the assumptions of the general Gauss-Markov model (2.1), a linear function $\boldsymbol{Fy}$ is the BLUE of $\boldsymbol{X\beta}$ if and only if the three conditions

    (a)  $\boldsymbol{Fy}$ is unbiased, i.e., $\boldsymbol{FX} = \boldsymbol{X}$,
    (b)  $\mathcal{R}(\boldsymbol{FV}) \subset \mathcal{R}(\boldsymbol{X})$ and
    (c)  $\boldsymbol{FV} = \boldsymbol{VF'}$

hold. (See also Theorem 5.1 in Rao, 1971).

**Corollary 2.3** Under the assumptions of Theorem 2.3, with $\boldsymbol{V}$ in (2.1) being p.d., the statistic $\boldsymbol{P}_{X(V^{-1})}\boldsymbol{y}$, with

$$\boldsymbol{P}_{X(V^{-1})} = \boldsymbol{X}(\boldsymbol{X'V^{-1}X})^{-}\boldsymbol{X'V^{-1}},$$

is the BLUE of $\mathrm{E}(\boldsymbol{y}) = \boldsymbol{X\beta}$.

Applying Theorem 3.2 of Rao (1974), this can be extended to the case where $\boldsymbol{V}$ is n.n.d. Then the projector to be used for $\boldsymbol{F}$ in Theorem 2.3 (above) is

$$\boldsymbol{P}_{X(T^{-})} = \boldsymbol{X}(\boldsymbol{X'T^{-}X})^{-}\boldsymbol{X'T^{-}},$$

where $\boldsymbol{T}^{-}$ is an n.n.d. $g$-inverse of the matrix $\boldsymbol{T} = \boldsymbol{V} + \boldsymbol{XUX'}$, for any symmetric $\boldsymbol{U}$ such that $\boldsymbol{T}$ is n.n.d. and $\mathrm{rank}(\boldsymbol{T}) = \mathrm{rank}[\boldsymbol{V} : \boldsymbol{X}]$. (See also Rao, 1973, Section 4i.4.) This extension was utilized in several papers by Oktaba (e.g. 1984, 1996, 1998, 2003).

## 2.2. Main hypothesis testing results

If not otherwise stated, it will be assumed here that $\boldsymbol{V}$ is p.d.

A hypothesis to be tested under the model $\{\boldsymbol{y}, \boldsymbol{X\beta}, \sigma^2\boldsymbol{V}\}$ can be presented as

$$H : \boldsymbol{\theta} \in \omega = \mathcal{R}(\boldsymbol{X}_*) \subset \Omega = \mathcal{R}(\boldsymbol{X}). \tag{2.2}$$

Thus under the general assumption, $\mathrm{E}(\boldsymbol{y}) = \boldsymbol{\theta} = \boldsymbol{X\beta}$, whereas under the hypothesis, it is reduced to $\mathrm{E}(\boldsymbol{y}) = \boldsymbol{\theta} = \boldsymbol{X}_*\boldsymbol{\beta}_*$, with $\boldsymbol{X}_* = \boldsymbol{XL}$ for some matrix $\boldsymbol{L}$. Here the notation $r = \mathrm{rank}(\boldsymbol{X})$ and $r - q = \mathrm{rank}(\boldsymbol{X}_*)$ will be used. The appropriate $F$ statistic for testing (2.2) can be derived assuming the multivariate normal distribution of $\boldsymbol{y}$, i.e., that

$$\boldsymbol{y} \sim N_n(\boldsymbol{\theta}, \sigma^2\boldsymbol{V}). \tag{2.3}$$

Under (2.3), the likelihood function for the observations $\boldsymbol{y} = [y_1, y_2, ..., y_n]'$ is given by their joint probability density (see, e.g., Scheffé, 1959, p. 416; Seber, 1980, p. 77), i.e., in the form

$$
\begin{aligned}
L_y(\boldsymbol{\theta}, \sigma^2) &= (2\pi)^{-n/2} |\sigma^2 \boldsymbol{V}|^{-1/2} \exp\{-(\boldsymbol{y} - \boldsymbol{\theta})'(\sigma^2 \boldsymbol{V})^{-1}(\boldsymbol{y} - \boldsymbol{\theta})/2\} \\
&= (2\pi\sigma^2)^{-n/2} |\boldsymbol{V}|^{-1/2} \exp\{-(\boldsymbol{y} - \boldsymbol{\theta})' \boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{\theta})/(2\sigma^2)\}. \quad (2.4)
\end{aligned}
$$

Now, one has to derive the maximum of $L_y(\boldsymbol{\theta}, \sigma^2)$ under the general assumption, $\mathcal{G}$, that $\boldsymbol{\theta} = \boldsymbol{X}\boldsymbol{\beta}$ (i.e., $\boldsymbol{\theta} \in \Omega$) and also under the hypothesis, $H$, that $\boldsymbol{\theta} = \boldsymbol{X}_*\boldsymbol{\beta}_*$ (i.e., $\boldsymbol{\theta} \in \omega$).

It appears that, under $\mathcal{G}$, the maximum of the likelihood (2.4) is attained by

$$
L_y(\boldsymbol{\theta}_0, \sigma_0^2) = (2\pi\sigma_0^2)^{-n/2} |\boldsymbol{V}|^{-1/2} \exp\{-n/2\},
$$

with $\boldsymbol{\theta}_0 = \boldsymbol{F}\boldsymbol{y}$ and with $\sigma_0^2$ obtained as the quadratic form

$$
\begin{aligned}
\sigma_0^2 &= (\boldsymbol{y} - \boldsymbol{\theta}_0)' \boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{\theta}_0)/n = (\boldsymbol{y} - \boldsymbol{F}\boldsymbol{y})' \boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{F}\boldsymbol{y})/n \\
&= \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{F})' \boldsymbol{V}^{-1}(\boldsymbol{I}_n - \boldsymbol{F})\boldsymbol{y}/n, \quad (2.5)
\end{aligned}
$$

where $\boldsymbol{F}$ satisfies the conditions of Theorem 2.3. Evidently, by that theorem $\boldsymbol{\theta}_0$ is the BLUE of $\boldsymbol{\theta} = \boldsymbol{X}\boldsymbol{\beta}$.

In general, as $\boldsymbol{F}$ one can take $\boldsymbol{P}_{X(V^{-1})} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X})^{-}\boldsymbol{X}'\boldsymbol{V}^{-1}$, i.e., the $\boldsymbol{V}^{-1}$-orthogonal projector on $\mathcal{R}(\boldsymbol{X})$. Then (2.5) is related to

$$
s^2 = \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{P}_{X(V^{-1})})' \boldsymbol{V}^{-1}(\boldsymbol{I}_n - \boldsymbol{P}_{X(V^{-1})})\boldsymbol{y}/(n - r)
$$

(the MINQUE of $\sigma^2$) by the equality $\sigma_0^2 = (n - r)s^2/n$.

Similarly, the maximum value of $L_y(\boldsymbol{\theta}, \sigma^2)$, under the hypothesis $H$, is attained by

$$
L_y(\boldsymbol{\theta}_1, \sigma_1^2) = (2\pi\sigma_1^2)^{-n/2} |\boldsymbol{V}|^{-1/2} \exp\{-n/2\},
$$

with $\boldsymbol{\theta}_1 = \boldsymbol{F}_1\boldsymbol{y}$ and

$$
\sigma_1^2 = (\boldsymbol{y} - \boldsymbol{\theta}_1)' \boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{\theta}_1)/n = \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{F}_1)' \boldsymbol{V}^{-1}(\boldsymbol{I}_n - \boldsymbol{F}_1)\boldsymbol{y}/n,
$$

where $\boldsymbol{F}_1$ satisfies the conditions (a)−(c) of Theorem 2.3 with $\boldsymbol{X}$ replaced by $\boldsymbol{X}_*$, such that $\mathcal{R}(\boldsymbol{X}_*) \subset \mathcal{R}(\boldsymbol{X})$. Hence $\boldsymbol{\theta}_1$ is the BLUE of $\boldsymbol{\theta} = \boldsymbol{X}_*\boldsymbol{\beta}_*$, specified by the hypothesis.

From these results, the likelihood ratio can be written as

$$L(H|\mathcal{G}) = \frac{\sup\limits_{\theta \in \omega} L_y(\boldsymbol{\theta}, \sigma^2)}{\sup\limits_{\theta \in \Omega} L_y(\boldsymbol{\theta}, \sigma^2)}, \tag{2.6}$$

where $H$ and $\mathcal{G}$ stand for the hypothesis and for the general assumption, respectively. Evidently, the ratio (2.6) reduces to $(\sigma_0^2/\sigma_1^2)^{n/2}$, which when transformed according to

$$F = \{[L(H|\mathcal{G})]^{-2/n} - 1\}(n-r)/q,$$

gives the $F$ statistic

$$F = \frac{n-r}{q}\,\frac{\sigma_1^2 - \sigma_0^2}{\sigma_0^2} = \frac{n-r}{q}\,\frac{Q_1 - Q_0}{Q_0}, \tag{2.7}$$

where $r = \mathrm{rank}(\boldsymbol{X})$, $r - q = \mathrm{rank}(\boldsymbol{X}_*)$ and

$$Q_1 = \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{F}_1)'\boldsymbol{V}^{-1}(\boldsymbol{I}_n - \boldsymbol{F}_1)\boldsymbol{y}, \quad Q_0 = \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{F})'\boldsymbol{V}^{-1}(\boldsymbol{I}_n - \boldsymbol{F})\boldsymbol{y}.$$

However, taking into account the conditions (a)−(c) of Theorem 2.3, which the matrices $\boldsymbol{F}$ and $\boldsymbol{F}_1$ have to satisfy in relation to $\boldsymbol{X}$ and $\boldsymbol{X}_*$, respectively, it can be shown that $Q_1 - Q_0 = \boldsymbol{y}'(\boldsymbol{F} - \boldsymbol{F}_1)'\boldsymbol{V}^{-1}(\boldsymbol{F} - \boldsymbol{F}_1)\boldsymbol{y}$. Hence the statistic (2.7) can be presented in a more simple form, as

$$F = \frac{n-r}{q}\,\frac{\boldsymbol{y}'(\boldsymbol{F} - \boldsymbol{F}_1)'\boldsymbol{V}^{-1}(\boldsymbol{F} - \boldsymbol{F}_1)\boldsymbol{y}}{\boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{F})'\boldsymbol{V}^{-1}(\boldsymbol{I}_n - \boldsymbol{F})\boldsymbol{y}}. \tag{2.8}$$

It can be shown that (2.8) really has an $F$ distribution by referring to a generalization of Fisher−Cochran theorem given in Rao and Mitra (1971, Section 9.3). In fact, the considered statistic $F$ has a noncentral distribution $F(q, n-r, \delta)$, i.e., with $q$ and $n-r$ d.f. and the noncentrality parameter

$$\delta \equiv \delta_1 = \boldsymbol{\theta}'(\boldsymbol{F} - \boldsymbol{F}_1)'\boldsymbol{V}^{-1}(\boldsymbol{F} - \boldsymbol{F}_1)\boldsymbol{\theta}/\sigma^2,$$

which becomes 0 if the hypothesis is true. The latter statement follows from the fact that, under $H : \boldsymbol{\theta} = \boldsymbol{X}_*\boldsymbol{\beta}_*$, the equalities $(\boldsymbol{F}-\boldsymbol{F}_1)\boldsymbol{\theta} = \boldsymbol{F}\boldsymbol{\theta}-\boldsymbol{F}_1\boldsymbol{\theta} = \boldsymbol{F}\boldsymbol{X}_*\boldsymbol{\beta}_* - \boldsymbol{F}_1\boldsymbol{X}_*\boldsymbol{\beta}_* = \boldsymbol{F}\boldsymbol{F}_1\boldsymbol{X}_*\boldsymbol{\beta}_* - \boldsymbol{F}_1\boldsymbol{X}_*\boldsymbol{\beta}_* = \boldsymbol{F}_1\boldsymbol{X}_*\boldsymbol{\beta}_* - \boldsymbol{F}_1\boldsymbol{X}_*\boldsymbol{\beta}_* = \boldsymbol{0}$ hold [on account of conditions (a) and (b) of Theorem 2.3].

The above results, leading to (2.7) and (2.8), can be summarized in the form of an ANOVA table, given in Table 1.

**Table 1.** The analysis of variance related to the $F$ test (2.7).

| Source of variation | Degrees of freedom | Sum of squares | Mean square | Expected mean square |
|---|---|---|---|---|
| Deviations from $H$ | $q$ | $Q_1 - Q_0$ | $(Q_1 - Q_0)/q$ | $\sigma^2 + \sigma^2 \delta/q$ |
| Residuals | $n - r$ | $Q_0$ | $Q_0/(n-r)$ | $\sigma^2$ |
| Total | $n - r + q$ | $Q_1$ | — | — |

To complete this section, it may be interesting to draw attention to the test of a hypothesis specified in the form

$$H : \boldsymbol{\theta} \in \omega = \Omega \cap \mathcal{N}(\boldsymbol{B}_1),$$

where $\boldsymbol{B}_1$ is a $q \times n$ matrix of rank $q$, such that $\mathcal{N}(\boldsymbol{B}_1) = \omega \oplus \Omega^{\perp}$, where $\mathcal{N}(\boldsymbol{B}_1)$ denotes the null space (kernel) of $\boldsymbol{B}_1$. Referring to the model $\{\boldsymbol{y}, \boldsymbol{X}\boldsymbol{\beta}, \sigma^2 \boldsymbol{V}\}$, this can be written as

$$H : \boldsymbol{\theta} \in \omega = \Omega \cap \{\boldsymbol{\theta} : \boldsymbol{B}_1 \boldsymbol{\theta} = \boldsymbol{0}\}, \quad \text{where} \quad \Omega = \{\boldsymbol{\theta} : \boldsymbol{\theta} = \boldsymbol{X}\boldsymbol{\beta}\},$$

or shortly

$$H : \boldsymbol{B}_1 \boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{0}. \tag{2.9}$$

This hypothesis can be tested using an $F$ statistic of the type

$$F = (\widehat{\boldsymbol{B}_1 \boldsymbol{\theta}})'[\mathrm{D}(\widehat{\boldsymbol{B}_1 \boldsymbol{\theta}})]^{-1}(\widehat{\boldsymbol{B}_1 \boldsymbol{\theta}})/q,$$

which here can be written as

$$F = \frac{n - r}{q} \frac{\boldsymbol{y}'\boldsymbol{F}'\boldsymbol{B}_1'(\boldsymbol{B}_1 \boldsymbol{F} \boldsymbol{V} \boldsymbol{F}' \boldsymbol{B}_1')^{-1}\boldsymbol{B}_1 \boldsymbol{F}\boldsymbol{y}}{\boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{F})'\boldsymbol{V}^{-1}(\boldsymbol{I}_n - \boldsymbol{F})\boldsymbol{y}}, \tag{2.10}$$

where $r = \mathrm{rank}(\boldsymbol{X})$, $q = \mathrm{rank}(\boldsymbol{B}_1) = \mathrm{rank}(\boldsymbol{B}_1 \boldsymbol{F})$, the matrix $\boldsymbol{F}$ being the same as that in (2.5), i.e., such which satisfies the conditions (a)$-$(c) of Theorem 2.3. The equality $\mathrm{rank}(\boldsymbol{B}_1) = \mathrm{rank}(\boldsymbol{B}_1 \boldsymbol{F})$ can be proved by referring to Lemma 4.4.3 in Seber (1980). That the statistic (2.10) has the $F$ distribution with $q$ and $n - r$ d.f. and the noncentrality parameter

$$\delta = \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{B}_1'(\boldsymbol{B}_1 \boldsymbol{F} \boldsymbol{V} \boldsymbol{F}' \boldsymbol{B}_1')^{-1}\boldsymbol{B}_1 \boldsymbol{X}\boldsymbol{\beta}/\sigma^2,$$

can be proved referring to Rao and Mitra (1971, Sections 9.2 and 9.3). One may also check that if $\boldsymbol{B}_1 = \boldsymbol{F} - \boldsymbol{F}_1$, then the statistic (2.8) can be seen as a special case of (2.10).

Furthermore, if the hypothesis to be tested is of the type

$$H : \boldsymbol{B}_1 \boldsymbol{X} \boldsymbol{\beta} = \boldsymbol{t}, \tag{2.11}$$

instead of (2.9), then the appropriate test statistic to be used is

$$F = \frac{n-r}{q} \, \frac{(\boldsymbol{B}_1 \boldsymbol{F} \boldsymbol{y} - \boldsymbol{t})'(\boldsymbol{B}_1 \boldsymbol{F} \boldsymbol{V} \boldsymbol{F}' \boldsymbol{B}_1')^{-1}(\boldsymbol{B}_1 \boldsymbol{F} \boldsymbol{y} - \boldsymbol{t})}{\boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{F})' \boldsymbol{V}^{-1}(\boldsymbol{I}_n - \boldsymbol{F}) \boldsymbol{y}}, \tag{2.12}$$

which has the same distribution as the statistic (2.10), except that here the noncentrality parameter is of the form

$$\delta = (\boldsymbol{B}_1 \boldsymbol{X} \boldsymbol{\beta} - \boldsymbol{t})'(\boldsymbol{B}_1 \boldsymbol{F} \boldsymbol{V} \boldsymbol{F}' \boldsymbol{B}_1')^{-1}(\boldsymbol{B}_1 \boldsymbol{X} \boldsymbol{\beta} - \boldsymbol{t})/\sigma^2, \tag{2.13}$$

becoming 0 if (2.11) is true.

Finally, it may be noted that the $F$ statistic (2.12) is in agreement with the relevant statistic given by Rao (1971, Theorem 3.2). One may check this by showing that

$$\begin{bmatrix} \boldsymbol{V} & \boldsymbol{X} \\ \boldsymbol{X}' & \boldsymbol{O} \end{bmatrix}^{-} = \begin{bmatrix} \boldsymbol{C}_1 & \boldsymbol{C}_2 \\ \boldsymbol{C}_3 & -\boldsymbol{C}_4 \end{bmatrix}, \tag{2.14}$$

where the matrices $\boldsymbol{C}_1$, $\boldsymbol{C}_2$, $\boldsymbol{C}_3$ and $\boldsymbol{C}_4$ are defined by the equalities

$$\begin{aligned} \boldsymbol{C}_1 &= (\boldsymbol{I}_n - \boldsymbol{F})' \boldsymbol{V}^{-1}(\boldsymbol{I}_n - \boldsymbol{F}), \\ \boldsymbol{X} \boldsymbol{C}_2' &= \boldsymbol{X} \boldsymbol{C}_3 = \boldsymbol{F}, \\ \boldsymbol{X} \boldsymbol{C}_4 \boldsymbol{X}' &= \boldsymbol{F} \boldsymbol{V} \boldsymbol{F}' \ (= \boldsymbol{F} \boldsymbol{V}), \end{aligned}$$

in accordance with Theorem 3.1 of Rao (1971). In examining this, it is helpful to refer to the conditions (a)−(c) given for $\boldsymbol{F}$ in Theorem 2.3.

Note, however, that Rao's $F$ statistic is more general. In the notation used here, it can be written in the form

$$F = q^{-1}(\boldsymbol{B}_1 \boldsymbol{F} \boldsymbol{y} - \boldsymbol{t})'(\boldsymbol{B}_1 \boldsymbol{F} \boldsymbol{V} \boldsymbol{F}' \boldsymbol{B}_1')^{-}(\boldsymbol{B}_1 \boldsymbol{F} \boldsymbol{y} - \boldsymbol{t})/s^2, \tag{2.15}$$

where $q = \text{rank}(\boldsymbol{B}_1 \boldsymbol{F} \boldsymbol{V} \boldsymbol{F}' \boldsymbol{B}_1')$ and $s^2$ is the MINQUE of $\sigma^2$ (see Rao, 1974, Section 3). This formula is more general than (2.12), because it allows the matrix $\boldsymbol{V}$ to be n.n.d. (not necessarily p.d.), and the matrix $\boldsymbol{B}_1$, unlike that used in (2.10) and (2.12), to have more rows than its rank (not necessarily to have all rows linearly independent). The statistic (2.15) is comparable with that in (2.4) of Oktaba (1984).

### 3. Application to the classic randomized block design

Following Nelder (1954), suppose that the experimenter randomizes the available blocks of number $N_B$, say, and then randomizes the experimental units (plots) within the blocks (of size $v$ each). Furthermore, following Scheffé (1959), let a distinction be made between the "unit errors" and the "technical errors". Then, introducing for the observations on the $n = vb$ experimental units the $n \times 1$ vector

$$\boldsymbol{y} = [y_{11}, ..., y_{v1}, y_{12}, ..., y_{v2}, ..., y_{1b}, ..., y_{vb}]' = [\boldsymbol{y}_1', \boldsymbol{y}_2', ..., \boldsymbol{y}_b']',$$

where $\boldsymbol{y}_j = [y_{1j}, y_{2j}, ..., y_{vj}]'$, $j = 1, 2, ..., b \leq N_B$, the model can be expressed in matrix notation as

$$\boldsymbol{y} = \boldsymbol{X}_1 \boldsymbol{\tau} + \boldsymbol{X}_2 \boldsymbol{\beta} + \boldsymbol{\eta} + \boldsymbol{e}, \tag{3.1}$$

with

$$\mathrm{E}(\boldsymbol{y}) = \boldsymbol{X}_1 \boldsymbol{\tau},$$
$$\mathrm{D}(\boldsymbol{y}) = (\boldsymbol{I}_b - N_B^{-1} \mathbf{1}_b \mathbf{1}_b') \otimes \mathbf{1}_v \mathbf{1}_v' \sigma_B^2 + \boldsymbol{I}_b \otimes (\boldsymbol{I}_v - v^{-1} \mathbf{1}_v \mathbf{1}_v') \sigma_U^2 + \boldsymbol{I}_b \otimes \boldsymbol{I}_v \sigma_e^2,$$

where

$\boldsymbol{X}_1 = \mathbf{1}_b \otimes \boldsymbol{I}_v$ is the design matrix for treatments,
$\boldsymbol{X}_2 = \boldsymbol{I}_b \otimes \mathbf{1}_v$ is the design matrix for blocks,
$\boldsymbol{\tau} = [\tau_1, \tau_2, ..., \tau_v]'$ is the vector of treatment parameters,
$\boldsymbol{\beta} = [\beta_1, \beta_2, ..., \beta_b]'$ is that of block random effects,
$\boldsymbol{\eta} = [\eta_{11}, ..., \eta_{v1}, \eta_{12}, ..., \eta_{v2}, ..., \eta_{1b}, ..., \eta_{vb}]'$ is that of unit errors,
$\boldsymbol{e} = [e_{11}, ..., e_{v1}, e_{12}, ..., e_{v2}, ..., e_{1b}, ..., e_{vb}]'$ is that of technical errors,
and where $\sigma_B^2$, $\sigma_U^2$ and $\sigma_e^2$ stand for variances of the block random effects, the experimental unit errors and the technical errors, respectively.

The model (3.1) differs from the usually assumed linear model, considered for example by Pearce (1983, Chapter 3) and Oktaba (2003, Section 6), in two aspects. In that model

(a) $\boldsymbol{\beta}$ is a vector of constant block parameters,

(b) $\boldsymbol{\eta}$ and $\boldsymbol{e}$ are reduced to one term with the dispersion matrix of the form $\boldsymbol{I}_{bv} \sigma^2$.

Now, by Theorem 2.1 applied to the present model, one finds that a necessary and sufficient condition for $\boldsymbol{a}'\boldsymbol{y}$ ($\boldsymbol{a} \neq \boldsymbol{0}$) to be the BLUE of $\mathrm{E}(\boldsymbol{a}'\boldsymbol{y}) = \boldsymbol{a}'\boldsymbol{X}_1 \boldsymbol{\tau} = \boldsymbol{a}'(\mathbf{1}_b \otimes \boldsymbol{I}_v)\boldsymbol{\tau}$ is the equality

$$[(\boldsymbol{I}_b - b^{-1} \mathbf{1}_b \mathbf{1}_b') \otimes \boldsymbol{I}_v][(\boldsymbol{I}_b - N_B^{-1} \mathbf{1}_b \mathbf{1}_b') \otimes \mathbf{1}_v \mathbf{1}_v' \sigma_B^2$$
$$+ \boldsymbol{I}_b \otimes (\boldsymbol{I}_v - v^{-1} \mathbf{1}_v \mathbf{1}_v') \sigma_U^2 + \boldsymbol{I}_{bv} \sigma_e^2]\boldsymbol{a} = \boldsymbol{0}.$$

It holds for any values of $\sigma_B^2, \sigma_U^2$ and $\sigma_e^2$ if and only if

$$[(\boldsymbol{I}_b - b^{-1}\mathbf{1}_b\mathbf{1}_b') \otimes \mathbf{1}_v\mathbf{1}_v']\boldsymbol{a} = \boldsymbol{0} \quad \text{and} \quad [(\boldsymbol{I}_b - b^{-1}\mathbf{1}_b\mathbf{1}_b') \otimes \boldsymbol{I}_v]\boldsymbol{a} = \boldsymbol{0}$$

hold simultaneously, i.e., if and only if $\boldsymbol{a} \in \mathcal{R}[\mathbf{1}_b \otimes \boldsymbol{I}_v]$.

Thus, for any $v \times 1$ vector $\boldsymbol{c}$ and any scalar $s$, $(s\mathbf{1}_b' \otimes \boldsymbol{c}')\boldsymbol{y}$ is the BLUE of $(s\mathbf{1}_b' \otimes \boldsymbol{c}')(\mathbf{1}_b \otimes \boldsymbol{I}_v)\boldsymbol{\tau} = (sb)\boldsymbol{c}'\boldsymbol{\tau}$. In particular, $(b^{-1}\mathbf{1}_b' \otimes \boldsymbol{c}')\boldsymbol{y}$ is for any $v \times 1$ vector $\boldsymbol{c}$ the BLUE of $\boldsymbol{c}'\boldsymbol{\tau}$, which can be written

$$\widehat{\boldsymbol{c}'\boldsymbol{\tau}} = b^{-1}\sum_{j=1}^{b} \boldsymbol{c}'\boldsymbol{y}_j = \boldsymbol{c}'\left(b^{-1}\sum_{j=1}^{b} \boldsymbol{y}_j\right).$$

This implies that the BLUE of $\boldsymbol{\tau}$ is of the form $\hat{\boldsymbol{\tau}} = b^{-1}\sum_{j=1}^{b} \boldsymbol{y}_j = b^{-1}\boldsymbol{X}_1'\boldsymbol{y}$, with

$$\mathrm{D}(\hat{\boldsymbol{\tau}}) = b^{-1}[(1 - b/N_B)\mathbf{1}_v\mathbf{1}_v'\sigma_B^2 + (\boldsymbol{I}_v - v^{-1}\mathbf{1}_v\mathbf{1}_v')\sigma_U^2 + \boldsymbol{I}_v\sigma_e^2].$$

Hence

$$\begin{aligned}
\mathrm{Var}(\widehat{\boldsymbol{c}'\boldsymbol{\tau}}) &= b^{-1}[(1 - b/N_B)\boldsymbol{c}'\mathbf{1}_v\mathbf{1}_v'\boldsymbol{c}\sigma_B^2 \\
&\quad + \boldsymbol{c}'(\boldsymbol{I}_v - v^{-1}\mathbf{1}_v\mathbf{1}_v')\boldsymbol{c}\sigma_U^2 + \boldsymbol{c}'\boldsymbol{c}\sigma_e^2].
\end{aligned} \tag{3.2}$$

To compare with the above-mentioned assumed linear model of the type

$$\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\tau} + \boldsymbol{X}_2\boldsymbol{\beta} + \boldsymbol{e}, \quad \text{with} \quad \mathrm{D}(\boldsymbol{y}) = \boldsymbol{I}_{bv}\sigma^2$$

and $\boldsymbol{\beta}$ considered as a vector of (constant) parameters, note that then $\mathrm{E}(\boldsymbol{y}) = \boldsymbol{X}_1\boldsymbol{\tau} + \boldsymbol{X}_2\boldsymbol{\beta}$ and the condition of Theorem 2.1 is $[\boldsymbol{I}_{bv} - \boldsymbol{P}_{[X_1:X_2]}]\boldsymbol{a} = \boldsymbol{0}$, where $\boldsymbol{I}_{bv} - \boldsymbol{P}_{[X_1:X_2]} = (\boldsymbol{I}_b - b^{-1}\mathbf{1}_b\mathbf{1}_b') \otimes (\boldsymbol{I}_v - v^{-1}\mathbf{1}_v\mathbf{1}_v')$. Hence the condition is then satisfied by any vector $\boldsymbol{a} = \boldsymbol{a}_1 \otimes \boldsymbol{a}_2$ such that $\boldsymbol{a}_1$ is proportional to $\mathbf{1}_b$, $\boldsymbol{a}_2$ is proportional to $\mathbf{1}_v$, or both. In particular, it is satisfied by $\boldsymbol{a} = b^{-1}\mathbf{1}_b \otimes \boldsymbol{c}$ for any $v \times 1$ vector $\boldsymbol{c}$. This implies that $b^{-1}(\mathbf{1}_b' \otimes \boldsymbol{c}')\boldsymbol{y} = b^{-1}\boldsymbol{c}'\boldsymbol{X}_1'\boldsymbol{y}$ is the BLUE of its expectation

$$(b^{-1}\mathbf{1}_b' \otimes \boldsymbol{c}')(\boldsymbol{X}_1\boldsymbol{\tau} + \boldsymbol{X}_2\boldsymbol{\beta}) = \boldsymbol{c}'\boldsymbol{\tau} + b^{-1}\boldsymbol{c}'\mathbf{1}_v\mathbf{1}_b'\boldsymbol{\beta}.$$

The estimator $b^{-1}\boldsymbol{c}'\boldsymbol{X}_1'\boldsymbol{y}$ is certainly the same as $b^{-1}\sum_{j=1}^{b} \boldsymbol{c}'\boldsymbol{y}_j$, obtained under the randomization-derived model (3.1), but now it cannot, in general, be denoted by $\widehat{\boldsymbol{c}'\boldsymbol{\tau}}$, because under the assumed linear model it estimates the function $\boldsymbol{c}'\boldsymbol{\tau} + b^{-1}\boldsymbol{c}'\mathbf{1}_v\mathbf{1}_b'\boldsymbol{\beta}$, not just $\boldsymbol{c}'\boldsymbol{\tau}$. Also, its variance under this

model is simply $\mathrm{Var}(b^{-1}\boldsymbol{c}'\boldsymbol{X}_1'\boldsymbol{y}) = b^{-1}\boldsymbol{c}'\boldsymbol{c}\sigma^2$, which differs from the $\mathrm{Var}(\widehat{\boldsymbol{c}'\boldsymbol{\tau}})$ obtained in (3.2).

But if $\boldsymbol{c}'\boldsymbol{\tau}$ is a contrast, then the moments under both models are exactly the same, with $\sigma^2 = \sigma_U^2 + \sigma_e^2$. This follows also directly from Theorem 2.2, by which, for any vector $\boldsymbol{c}$, the function $b^{-1}(\mathbf{1}_b' \otimes \boldsymbol{c}')\boldsymbol{y} = b^{-1}\boldsymbol{c}'\boldsymbol{X}_1'\boldsymbol{y}$ is both the SLSE and the BLUE of $\boldsymbol{c}'\boldsymbol{\tau}$ under the randomization-derived model (3.1). But, by Theorem 2.2, the same is true under the assumed linear model, if $\boldsymbol{c}'\boldsymbol{\tau}$ is a contrast.

It will be convenient to write the dispersion matrix of model (3.1) in the form

$$\mathrm{D}(\boldsymbol{y}) = \boldsymbol{\phi}_1\sigma_1^2 + \boldsymbol{\phi}_2\sigma_2^2 + \boldsymbol{\phi}_3\sigma_3^2, \tag{3.3}$$

where

$$\begin{aligned}
\boldsymbol{\phi}_1 &= \boldsymbol{I}_n - \boldsymbol{P}_{X_2} = \boldsymbol{I}_n - v^{-1}\boldsymbol{X}_2\boldsymbol{X}_2' = \boldsymbol{I}_b \otimes (\boldsymbol{I}_v - v^{-1}\mathbf{1}_v\mathbf{1}_v'), &(3.4)\\
\boldsymbol{\phi}_2 &= \boldsymbol{P}_{X_2} - \boldsymbol{P}_{1_n} = (\boldsymbol{I}_b - b^{-1}\mathbf{1}_b\mathbf{1}_b') \otimes v^{-1}\mathbf{1}_v\mathbf{1}_v', &(3.5)\\
\boldsymbol{\phi}_3 &= \boldsymbol{P}_{1_n} = b^{-1}\mathbf{1}_b\mathbf{1}_b' \otimes v^{-1}\mathbf{1}_v\mathbf{1}_v', &(3.6)
\end{aligned}$$

and

$$\sigma_1^2 = \sigma_U^2 + \sigma_e^2, \quad \sigma_2^2 = v\sigma_B^2 + \sigma_e^2, \quad \sigma_3^2 = (1 - N_B^{-1}b)v\sigma_B^2 + \sigma_e^2.$$

The matrices $\{\boldsymbol{\phi}_\alpha\}$ appearing in (3.3) can be seen as orthogonal projectors on relevant subspaces, $\boldsymbol{\phi}_1$ on $\mathcal{R}^\perp(\boldsymbol{X}_2)$, $\boldsymbol{\phi}_2$ on $\mathcal{R}^\perp(\mathbf{1}_n) \cap \mathcal{R}(\boldsymbol{X}_2)$ and $\boldsymbol{\phi}_3$ on $\mathcal{R}(\mathbf{1}_n)$. These subspaces correspond to three "strata" into which the experimental units are grouped. Adopting the terminology used by Pearce (1983, p. 109), these strata may be specified as follows:

($\alpha = 1$) 1st stratum—of units within blocks, called "intra-block,"

($\alpha = 2$) 2nd stratum—of blocks within the total area, called "inter-block,"

($\alpha = 3$) 3rd stratum—of the total experimental area.

The variances $\sigma_1^2$, $\sigma_2^2$ and $\sigma_3^2$ can then be called the "stratum variances" (see e.g. John, 1987, p. 187). Evidently the three orthogonal projectors, (3.4), (3.5) and (3.6), satisfy the conditions

$$\boldsymbol{\phi}_\alpha = \boldsymbol{\phi}_\alpha', \quad \boldsymbol{\phi}_\alpha\boldsymbol{\phi}_\alpha = \boldsymbol{\phi}_\alpha, \quad \boldsymbol{\phi}_\alpha\boldsymbol{\phi}_{\alpha'} = \boldsymbol{O} \quad \text{for } \alpha \neq \alpha',$$

where $\alpha, \alpha' = 1, 2, 3$, and also the condition

$$\boldsymbol{\phi}_1 + \boldsymbol{\phi}_2 + \boldsymbol{\phi}_3 = \boldsymbol{I}_b \otimes \boldsymbol{I}_v \quad (\equiv \boldsymbol{I}_n).$$

Moreover, note that $\text{rank}(\boldsymbol{\phi}_1) = b(v-1)$, $\text{rank}(\boldsymbol{\phi}_2) = b-1$ and $\text{rank}(\boldsymbol{\phi}_3) = 1$.

From formula (3.3), one can write $\text{D}(\boldsymbol{y}) = \sigma^2 \boldsymbol{V}$, with $\sigma^2 = \sigma_1^2$ and

$$\boldsymbol{V} = \boldsymbol{\phi}_1 + \boldsymbol{\phi}_2(\sigma_2^2/\sigma_1^2) + \boldsymbol{\phi}_3(\sigma_3^2/\sigma_1^2).$$

The inverse of $\boldsymbol{V}$ has then the form

$$\boldsymbol{V}^{-1} = \boldsymbol{\phi}_1 + \boldsymbol{\phi}_2(\sigma_1^2/\sigma_2^2) + \boldsymbol{\phi}_3(\sigma_1^2/\sigma_3^2). \tag{3.7}$$

This allows the present model to be written in the notation of the general Gauss−Markov model as $\{\boldsymbol{y}, \boldsymbol{X}_1\boldsymbol{\tau}, \sigma^2\boldsymbol{V}\}$, with $\boldsymbol{V}$ being p.d. here.

Suppose now that one is interested to test the hypothesis that the treatment parameters are all equal. With the notation $\text{E}(\boldsymbol{y}) = \boldsymbol{\theta}$, the hypothesis can be written as in (2.2), i.e., as

$$H : \boldsymbol{\theta} \in \omega = \mathcal{R}(\boldsymbol{1}_n) \subset \Omega = \mathcal{R}(\boldsymbol{X}_1), \tag{3.8}$$

or equivalently as $H : \tau_1 = \tau_2 = \cdots = \tau_v$. The question is whether the $F$ statistic presented in (2.8) can be used under the present model for testing $H$. First it should be noted, by referring to Theorem 2.3, that the matrix

$$\boldsymbol{F} = \boldsymbol{P}_{X_1} = b^{-1}\boldsymbol{X}_1\boldsymbol{X}_1' = b^{-1}\boldsymbol{1}_b\boldsymbol{1}_b' \otimes \boldsymbol{I}_v$$

satisfies all the three conditions of that theorem. In fact,

(a) $\boldsymbol{P}_{X_1}\boldsymbol{X}_1 = \boldsymbol{X}_1$,

(b) $\boldsymbol{P}_{X_1}\boldsymbol{V} = b^{-1}\boldsymbol{X}_1\boldsymbol{X}_1'\boldsymbol{V}$, i.e., $\mathcal{R}(\boldsymbol{P}_{X_1}\boldsymbol{V}) \subset \mathcal{R}(\boldsymbol{X}_1)$,

(c) $\boldsymbol{P}_{X_1}\boldsymbol{V} = \boldsymbol{V}\boldsymbol{P}_{X_1}$, because $\boldsymbol{P}_{X_1}\boldsymbol{\phi}_\alpha = \boldsymbol{\phi}_\alpha\boldsymbol{P}_{X_1}$, $\alpha = 1, 2, 3$.

Thus $\boldsymbol{F}\boldsymbol{y} = \boldsymbol{P}_{X_1}\boldsymbol{y} = (b^{-1}\boldsymbol{1}_b\boldsymbol{1}_b' \otimes \boldsymbol{I}_v)\boldsymbol{y}$ is the BLUE of $\boldsymbol{\theta} = \boldsymbol{X}_1\boldsymbol{\tau}$ under the present model. Similarly, it can be shown that $\boldsymbol{F}_1\boldsymbol{y} = \boldsymbol{P}_{1_n}\boldsymbol{y} = (b^{-1}\boldsymbol{1}_b\boldsymbol{1}_b' \otimes v^{-1}\boldsymbol{1}_v\boldsymbol{1}_v')\boldsymbol{y}$ is the BLUE of $\boldsymbol{1}_n\boldsymbol{\tau}$, to which $\boldsymbol{\theta}$ is reduced if $H$ is true. Also note that $\text{rank}(\boldsymbol{X}_1) = v$ and $\text{rank}(\boldsymbol{1}_n) = 1$.

In view of these results, the projectors $\boldsymbol{F} = \boldsymbol{P}_{X_1} = b^{-1}\boldsymbol{1}_b\boldsymbol{1}_b' \otimes \boldsymbol{I}_v$ and $\boldsymbol{F}_1 = \boldsymbol{P}_{1_n} = b^{-1}\boldsymbol{1}_b\boldsymbol{1}_b' \otimes v^{-1}\boldsymbol{1}_v\boldsymbol{1}_v'$ can be used in (2.8), to provide the appropriate $F$ statistic for testing the hypothesis (3.8), in the form

$$
\begin{aligned}
F &= \frac{n-v}{v-1} \frac{\boldsymbol{y}'(\boldsymbol{P}_{X_1} - \boldsymbol{P}_{1_n})\boldsymbol{V}^{-1}(\boldsymbol{P}_{X_1} - \boldsymbol{P}_{1_n})\boldsymbol{y}}{\boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{P}_{X_1})\boldsymbol{V}^{-1}(\boldsymbol{I}_n - \boldsymbol{P}_{X_1})\boldsymbol{y}} \\
&= \frac{n-v}{v-1} \frac{\boldsymbol{y}'[b^{-1}\boldsymbol{1}_b\boldsymbol{1}_b' \otimes (\boldsymbol{I}_v - v^{-1}\boldsymbol{1}_v\boldsymbol{1}_v')]\boldsymbol{V}^{-1}[b^{-1}\boldsymbol{1}_b\boldsymbol{1}_b' \otimes (\boldsymbol{I}_v - v^{-1}\boldsymbol{1}_v\boldsymbol{1}_v')]\boldsymbol{y}}{\boldsymbol{y}'[(\boldsymbol{I}_b - b^{-1}\boldsymbol{1}_b\boldsymbol{1}_b') \otimes \boldsymbol{I}_v]\boldsymbol{V}^{-1}[(\boldsymbol{I}_b - b^{-1}\boldsymbol{1}_b\boldsymbol{1}_b') \otimes \boldsymbol{I}_v]\boldsymbol{y}},
\end{aligned}
$$

where $\boldsymbol{V}^{-1}$ is as in (3.7). Note, however, that because of the formulae (3.4), (3.5) and (3.6), this $F$ statistic simplifies to

$$F = \frac{n-v}{v-1}\frac{\boldsymbol{y}'\boldsymbol{A}_1\boldsymbol{y}}{\boldsymbol{y}'\boldsymbol{A}_2\boldsymbol{y}}, \tag{3.9}$$

where

$$\begin{aligned}
\boldsymbol{A}_1 &= (\boldsymbol{P}_{X_1} - \boldsymbol{P}_{1_n})\boldsymbol{\phi}_1(\boldsymbol{P}_{X_1} - \boldsymbol{P}_{1_n}) = \boldsymbol{P}_{X_1}\boldsymbol{\phi}_1\boldsymbol{P}_{X_1} \\
&= b^{-1}\boldsymbol{1}_b\boldsymbol{1}_b' \otimes (\boldsymbol{I}_v - v^{-1}\boldsymbol{1}_v\boldsymbol{1}_v'), \\
\boldsymbol{A}_2 &= (\boldsymbol{I}_n - \boldsymbol{P}_{X_1})[\boldsymbol{\phi}_1 + \boldsymbol{\phi}_2(\sigma_1^2/\sigma_2^2)](\boldsymbol{I}_n - \boldsymbol{P}_{X_1}) \\
&= (\boldsymbol{I}_b - b^{-1}\boldsymbol{1}_b\boldsymbol{1}_b') \otimes (\boldsymbol{I}_v - v^{-1}\boldsymbol{1}_v\boldsymbol{1}_v') + (\boldsymbol{I}_b - b^{-1}\boldsymbol{1}_b\boldsymbol{1}_b') \otimes v^{-1}\boldsymbol{1}_v\boldsymbol{1}_v'(\sigma_1^2/\sigma_2^2) \\
&= (\boldsymbol{I}_b - b^{-1}\boldsymbol{1}_b\boldsymbol{1}_b') \otimes [\boldsymbol{I}_v - (1 - \sigma_1^2/\sigma_2^2)v^{-1}\boldsymbol{1}_v\boldsymbol{1}_v']
\end{aligned}$$

and

$$\begin{aligned}
\boldsymbol{A} = \boldsymbol{A}_1 + \boldsymbol{A}_2 &= \boldsymbol{I}_b \otimes (\boldsymbol{I}_v - v^{-1}\boldsymbol{1}_v\boldsymbol{1}_v') + (\boldsymbol{I}_b - b^{-1}\boldsymbol{1}_b\boldsymbol{1}_b') \otimes v^{-1}\boldsymbol{1}_v\boldsymbol{1}_v'(\sigma_1^2/\sigma_2^2) \\
&= \boldsymbol{\phi}_1 + \boldsymbol{\phi}_2(\sigma_1^2/\sigma_2^2).
\end{aligned}$$

This shows that the relevant ANOVA is based on the equation

$$\boldsymbol{y}'\boldsymbol{A}\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{A}_1\boldsymbol{y} + \boldsymbol{y}'\boldsymbol{A}_2\boldsymbol{y}, \tag{3.10}$$

where the quadratic forms $\boldsymbol{y}'\boldsymbol{A}_1\boldsymbol{y}/\sigma_1^2$ and $\boldsymbol{y}'\boldsymbol{A}_2\boldsymbol{y}/\sigma_1^2$ are distributed independently,

$$\boldsymbol{y}'\boldsymbol{A}_1\boldsymbol{y}/\sigma_1^2 \sim \chi^2(\nu_1, \delta_1), \quad \boldsymbol{y}'\boldsymbol{A}_2\boldsymbol{y}/\sigma_1^2 \sim \chi^2(\nu_2, \delta_2),$$

with

$$\begin{aligned}
\nu_1 = v - 1, \quad \delta_1 &= \boldsymbol{\tau}'\boldsymbol{X}_1'[b^{-1}\boldsymbol{1}_b\boldsymbol{1}_b' \otimes (\boldsymbol{I}_v - v^{-1}\boldsymbol{1}_v\boldsymbol{1}_v')]\boldsymbol{X}_1\boldsymbol{\tau}/\sigma_1^2 \\
&= b\boldsymbol{\tau}'(\boldsymbol{I}_v - v^{-1}\boldsymbol{1}_v\boldsymbol{1}_v')\boldsymbol{\tau}/\sigma_1^2, \\
\nu_2 = n - v, \quad \delta_2 &= 0.
\end{aligned}$$

Thus the statistic (3.9) has a noncentral $F$ distribution with $v-1$ and $n-v$ d.f. and the noncentrality parameter $\delta \equiv \delta_1$, becoming central if $H$ is true (as then $\delta_1 = 0$).

It should be noted, however, that the application of the test statistic $F$ of the form (3.9) requires knowledge of the variance ratio $\sigma_1^2/\sigma_2^2$. In practice, this ratio is usually unknown, as the involved variances are not known in advance. They can be estimated, but then the $F$ statistic so obtained will

not follow the indicated $F$ distribution exactly. To solve this problem, it may be instructive to partition the equation (3.10) into two equations,

$$\boldsymbol{y}'\boldsymbol{\phi}_1\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{P}_{X_1}\boldsymbol{\phi}_1\boldsymbol{P}_{X_1}\boldsymbol{y} + \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{P}_{X_1})\boldsymbol{\phi}_1(\boldsymbol{I}_n - \boldsymbol{P}_{X_1})\boldsymbol{y}, \qquad (3.11)$$
$$\boldsymbol{y}'\boldsymbol{\phi}_2\boldsymbol{y} = \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{P}_{X_1})\boldsymbol{\phi}_2(\boldsymbol{I}_n - \boldsymbol{P}_{X_1})\boldsymbol{y}. \qquad (3.12)$$

The first represents the so-called intra-block analysis, the second is related to the so-called inter-block analysis. Of course, in the present design case, the classic randomized block design, each block contains as many units as there are treatments to be compared in the experiment. This gives the advantage that in each of the $b$ blocks there is room for comparing all the $v$ treatments, so that no contrast of treatment parameters is confounded with blocks. As a consequence, the intra-block analysis (3.11) comprises the complete treatment sum of squares,

$$\begin{aligned}\text{Treatment SS} &= \boldsymbol{y}'\boldsymbol{P}_{X_1}\boldsymbol{\phi}_1\boldsymbol{P}_{X_1}\boldsymbol{y} = \boldsymbol{y}'[b^{-1}\boldsymbol{1}_b\boldsymbol{1}_b' \otimes (\boldsymbol{I}_v - v^{-1}\boldsymbol{1}_v\boldsymbol{1}_v')]\boldsymbol{y} \\ &= \boldsymbol{y}'(\boldsymbol{P}_{X_1} - \boldsymbol{P}_{1_n})\boldsymbol{y}\end{aligned}$$

and the intra-block error sum of squares,

$$\begin{aligned}\text{Error (intra) SS} &= \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{P}_{X_1})\boldsymbol{\phi}_1(\boldsymbol{I}_n - \boldsymbol{P}_{X_1})\boldsymbol{y} \\ &= \boldsymbol{y}'[(\boldsymbol{I}_b - b^{-1}\boldsymbol{1}_b\boldsymbol{1}_b') \otimes (\boldsymbol{I}_v - v^{-1}\boldsymbol{1}_v\boldsymbol{1}_v')]\boldsymbol{y} \\ &= \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{P}_{X_2} - \boldsymbol{P}_{X_1} + \boldsymbol{P}_{1_n})\boldsymbol{y}.\end{aligned}$$

On the other hand, the inter-block analysis (3.12) is confined to the inter-block error sum of squares only, i.e., to

$$\begin{aligned}\text{Error (inter) SS} &= \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{P}_{X_1})\boldsymbol{\phi}_2(\boldsymbol{I}_n - \boldsymbol{P}_{X_1})\boldsymbol{y} \\ &= \boldsymbol{y}'[(\boldsymbol{I}_b - b^{-1}\boldsymbol{1}_b\boldsymbol{1}_b') \otimes v^{-1}\boldsymbol{1}_v\boldsymbol{1}_v']\boldsymbol{y} \\ &= \boldsymbol{y}'\boldsymbol{\phi}_2\boldsymbol{y} = \boldsymbol{y}'(\boldsymbol{P}_{X_2} - \boldsymbol{P}_{1_n})\boldsymbol{y},\end{aligned}$$

equal here to the block sum of squares, Block SS. It will be seen later that this is not necessarily the case in a general block design, and in an incomplete block design in particular.

Thus what is of interest here is only the intra-block part of the analysis, based on (3.11). It can explicitly be written as

$$\boldsymbol{y}'\tilde{\boldsymbol{A}}\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{A}_1\boldsymbol{y} + \boldsymbol{y}'\tilde{\boldsymbol{A}}_2\boldsymbol{y}, \qquad (3.13)$$

where $\tilde{\boldsymbol{A}} = \boldsymbol{\phi}_1 = \boldsymbol{I}_b \otimes (\boldsymbol{I}_v - v^{-1}\boldsymbol{1}_v\boldsymbol{1}_v')$, $\boldsymbol{A}_1 = \boldsymbol{P}_{X_1}\boldsymbol{\phi}_1\boldsymbol{P}_{X_1} = b^{-1}\boldsymbol{1}_b\boldsymbol{1}_b' \otimes (\boldsymbol{I}_v - v^{-1}\boldsymbol{1}_v\boldsymbol{1}_v')$, $\tilde{\boldsymbol{A}}_2 = (\boldsymbol{I}_n - \boldsymbol{P}_{X_1})\boldsymbol{\phi}_1(\boldsymbol{I}_n - \boldsymbol{P}_{X_1}) = (\boldsymbol{I}_b - b^{-1}\boldsymbol{1}_b\boldsymbol{1}_b') \otimes (\boldsymbol{I}_v - v^{-1}\boldsymbol{1}_v\boldsymbol{1}_v')$.

To obtain the $F$ type statistic for testing the hypothesis $H$ in this analysis, it should be noted that, independently,

$$y'A_1y/\sigma_1^2 \sim \chi^2(\nu_1, \delta_1), \quad \nu_1 = v - 1, \quad \delta_1 = b\tau'(I_v - v^{-1}1_v1_v')\tau/\sigma_1^2,$$
$$y'\tilde{A}_2y/\sigma_1^2 \sim \chi^2(\nu_2, \delta_2), \quad \nu_2 = (b-1)(v-1), \quad \delta_2 = 0.$$

Hence the appropriate $F$ statistic is

$$\begin{aligned} F &= \frac{\nu_2}{\nu_1} \frac{y'A_1y}{y'\tilde{A}_2y} \\ &= \frac{(b-1)(v-1)}{v-1} \frac{y'[b^{-1}1_b1_b' \otimes (I_v - v^{-1}1_v1_v')]y}{y'[(I_b - b^{-1}1_b1_b') \otimes (I_v - v^{-1}1_v1_v')]y}. \end{aligned} \quad (3.14)$$

It follows that (3.14) has an $F(\nu_1, \nu_2, \delta)$ distribution, i.e., a noncentral $F$ distribution, with $\nu_1 = v - 1$ and $\nu_2 = (b-1)(v-1)$ d.f. and the noncentrality parameter $\delta \equiv \delta_1 = b\tau'(I_v - v^{-1}1_v1_v')\tau/\sigma_1^2$. The latter becomes 0 if $H$ is true.

The results presented here can be summarized as in Table 2.

**Table 2.** The analysis of variance for the randomized block design.

| Source of variation | Degrees of freedom | Sum of squares | Expected mean square |
|---|---|---|---|
| Blocks | $b-1$ | $y'(P_{X_2} - P_{1_n})y$ | $\sigma_2^2 = v\sigma_B^2 + \sigma_e^2$ |
| Treatments | $v-1$ | $y'(P_{X_1} - P_{1_n})y$ | $\sigma_1^2 + \sigma_1^2\delta/(v-1)$ |
| Error (intra) | $(b-1)(v-1)$ | $y'(I_n - P_{X_2} - P_{X_1} + P_{1_n})y$ | $\sigma_1^2 = \sigma_U^2 + \sigma_e^2$ |
| Total | $bv-1$ | $y'(I_n - P_{1_n})y$ | — |

## 4. A randomization-derived model for a general block design

According to one of the basic principles of experimental design, the randomization principle (see Fisher, 1925, Section 48), the experimental units (plots) are to be randomized before they enter the experiment. Suppose that to apply a general block design, randomization is performed as described by Nelder (1954), i.e. by randomly permuting blocks within their total area and by randomly permuting units within the blocks. Then, assuming the usual unit-treatment additivity (in the sense of Nelder, 1965, p. 168; see also White, 1975, p. 560; Bailey, 1981, p. 215, 1991, p. 30; Kala, 1991, p. 7; Hinkelmann and Kempthorne, 2008, Section 9.2.2), and, as

usual, that the technical errors are uncorrelated, each with zero expectation and a finite variance, and that they are independent of the unit responses to treatments (see Neyman, 1935, pp. 110-114 and 145; Kempthorne, 1952, p. 132 and Section 8.4; Ogawa, 1961, 1963; Hinkelmann and Kempthorne, 2008, Section 9.2.6), the model of the variables observed on the $n$ units actually used in the experiment can be written in matrix notation as

$$\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\tau} + \boldsymbol{X}_2\boldsymbol{\beta} + \boldsymbol{\eta} + \boldsymbol{e}, \tag{4.1}$$

where $\boldsymbol{y}$ is an $n \times 1$ vector of observations, $\boldsymbol{\tau}$ is a $v \times 1$ vector of treatment parameters, $\boldsymbol{\beta}$ is a $b \times 1$ vector of block random effects, $\boldsymbol{\eta}$ is an $n \times 1$ vector of unit errors and $\boldsymbol{e}$ is an $n \times 1$ vector of technical errors, the matrices $\boldsymbol{X}_1$ (traditionally as $\boldsymbol{\Delta}'$) and $\boldsymbol{X}_2$ (traditionally as $\boldsymbol{D}'$) being the relevant design matrices for treatments and blocks respectively. Properties of the model (4.1) can be established by following its derivation from the randomizations involved, as shown in Caliński and Kageyama (2000, Section 3.1.1).

With the results of this derivation, one can express the model as in (4.1), and the corresponding moments in the form of the expectation vector

$$\mathrm{E}(\boldsymbol{y}) = \boldsymbol{X}_1\boldsymbol{\tau} \tag{4.2}$$

and the covariance (dispersion) matrix

$$\mathrm{D}(\boldsymbol{y}) = (\boldsymbol{X}_2\boldsymbol{X}_2' - N_B^{-1}\mathbf{1}_n\mathbf{1}_n')\sigma_B^2 + (\boldsymbol{I}_n - K_H^{-1}\boldsymbol{X}_2\boldsymbol{X}_2')\sigma_U^2 + \boldsymbol{I}_n\sigma_e^2, \tag{4.3}$$

respectively, where the matrices $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are as defined above, whereas $\sigma_B^2$, $\sigma_U^2$ and $\sigma_e^2$ are respectively the variances of the block random effects, the experimental unit errors and the technical errors, $N_B$ denoting the available number of blocks and $K_H$ standing for the weighted harmonic average of the potential numbers of units in the available blocks.

The model (4.1), with properties (4.2) and (4.3), coincides with that of Patterson and Thompson (1971), when their matrix $\boldsymbol{\Gamma}$ in the formula

$$\mathrm{D}(\boldsymbol{y}) = \sigma^2(\boldsymbol{X}_2\boldsymbol{\Gamma}\boldsymbol{X}_2' + \boldsymbol{I}_n)$$

is taken equal to $\boldsymbol{I}_b\gamma - N_B^{-1}\mathbf{1}_b\mathbf{1}_b'\sigma_B^2/\sigma^2$, where $\gamma = (\sigma_B^2 - K_H^{-1}\sigma_U^2)/\sigma^2$ and $\sigma^2 = \sigma_U^2 + \sigma_e^2$. In its general form, the present model coincides completely with the model obtained by Kala (1991) under more general considerations. Furthermore, if $N_B = b$, and the block sizes are equal, i.e., $k_1 = k_2 = \cdots = k_b = k$ (say), and $k = K_H$, then this model coincides essentially with that considered by Rao (1959) and by Shah (1992).

Under the model (4.1) with the above properties, the following main results concerning the linear estimation of treatment parametric functions can be obtained.

**Theorem 4.1** (3.1.1 in Caliński and Kageyama, 2000). Under the model (4.1), with the properties (4.2) and (4.3), a function $\boldsymbol{a}'\boldsymbol{y}$ is uniformly the BLUE of $\boldsymbol{c}'\boldsymbol{\tau}$ if and only if $\boldsymbol{a} = \boldsymbol{X}_1\boldsymbol{s}$, where $\boldsymbol{s} = \boldsymbol{r}^{-\delta}\boldsymbol{c}$ satisfies the condition

$$(\boldsymbol{k}^{\delta} - \boldsymbol{N}'\boldsymbol{r}^{-\delta}\boldsymbol{N})\boldsymbol{N}'\boldsymbol{s} = \boldsymbol{0}, \tag{4.4}$$

where $\boldsymbol{k}^{\delta} = \mathrm{diag}[k_1, k_2, ..., k_b]$ and $\boldsymbol{r}^{\delta} = \mathrm{diag}[r_1, r_2, ..., r_v]$, $\boldsymbol{r}^{-\delta} = (\boldsymbol{r}^{\delta})^{-1}$.

**Corollary 4.1** (3.1.1 in Caliński and Kageyama, 2000). For the estimation of $\boldsymbol{c}'\boldsymbol{\tau} = \boldsymbol{s}'\boldsymbol{r}^{\delta}\boldsymbol{\tau}$ under the model considered in Theorem 4.1 the following applies:

(a) If $\boldsymbol{N}'\boldsymbol{s} = \boldsymbol{0}$, then (4.4) is satisfied and the estimated function is a contrast.

(b) If $\boldsymbol{N}'\boldsymbol{s} \neq \boldsymbol{0}$, then to satisfy (4.4) it is necessary and sufficient that the elements of $\boldsymbol{N}'\boldsymbol{s}$ obtained from the same connected subdesign are all equal, i.e., that $\boldsymbol{N}'\boldsymbol{s} \in \mathcal{R}\{\mathrm{diag}[\boldsymbol{1}_{b_1} : \boldsymbol{1}_{b_2} : \cdots : \boldsymbol{1}_{b_g}]\}$.

Now the question is under which design conditions any function $\boldsymbol{s}'\boldsymbol{X}_1'\boldsymbol{y}$ is the BLUE of its expectation. An answer to this question can be given as follows.

**Theorem 4.2** (3.1.2 in Caliński and Kageyama, 2000). Under the model as in Theorem 4.1, any function $\boldsymbol{a}'\boldsymbol{y} = \boldsymbol{s}'\boldsymbol{X}_1'\boldsymbol{y}$, i.e., with any $\boldsymbol{s}$, is uniformly the BLUE of $\mathrm{E}(\boldsymbol{a}'\boldsymbol{y}) = \boldsymbol{s}'\boldsymbol{r}^{\delta}\boldsymbol{\tau}$ if and only if

$$(\boldsymbol{I}_n - \boldsymbol{P}_{X_1})\boldsymbol{X}_2\boldsymbol{X}_2'\boldsymbol{X}_1 = \boldsymbol{O}, \tag{4.5}$$

and this condition holds if and only if

(i) the design is orthogonal, i.e., satisfies the condition $\boldsymbol{N}\boldsymbol{k}^{-\delta}\boldsymbol{N}'\boldsymbol{r}^{-\delta}\boldsymbol{N} = \boldsymbol{N}$, and

(ii) the block sizes of the design are constant within any of its connected subdesigns, i.e., $\boldsymbol{N}_{\ell}'\boldsymbol{1}_{v_{\ell}} = k_{\ell}\boldsymbol{1}_{b_{\ell}}$, $\ell = 1, 2, ..., g$ (if the incidence matrix is of the form $\boldsymbol{N} = \mathrm{diag}[\boldsymbol{N}_1 : \boldsymbol{N}_2 : \cdots : \boldsymbol{N}_g]$).

Note that if equality (4.5) holds, then the BLUE of the expectation vector (4.2) is obtainable by an SLSE procedure, in the form $\boldsymbol{P}_{X_1}\boldsymbol{y}$. This implication can be checked by noting that (a) $\boldsymbol{P}_{X_1}\boldsymbol{X}_1 = \boldsymbol{X}_1$, (b) $\mathcal{R}[\boldsymbol{P}_{X_1}\mathrm{D}(\boldsymbol{y})] \subset \mathcal{R}(\boldsymbol{X}_1)$ and that (c) $\boldsymbol{P}_{X_1}\mathrm{D}(\boldsymbol{y}) = \mathrm{D}(\boldsymbol{y})\boldsymbol{P}_{X_1}$, as required in Theorem 2.3.

## 5. Resolving into stratum submodels

The results of Section 4 appear discouraging, as in many block designs the BLUEs will exist under the model (4.1) for only a few parametric functions of interest, or for none of them. For example, in the case of a balanced incomplete block (BIB) design, for which (see e.g. Raghavarao and Padgett, 2005, p. 55) the equality $\boldsymbol{NN'} = (r-\lambda)\boldsymbol{I}_v + \lambda \boldsymbol{1}_v \boldsymbol{1}_v'$ holds, none of the contrasts of treatment parameters will have a BLUE (on account of Corollary 4.1).

The apparent difficulty with the model (4.1) is usually avoided by resolving it into three submodels (two for contrasts), in accordance with the stratification of the experimental units. In fact, the units of a block experiment can be seen as being grouped according to a nested classification with three strata. Their specification given in Section 3 applies to any experiment in a block design, not in the classic RB design only. Because of this stratification, the observed vector $\boldsymbol{y}$ can be decomposed as

$$\boldsymbol{y} = \boldsymbol{y}_1 + \boldsymbol{y}_2 + \boldsymbol{y}_3, \tag{5.1}$$

where each of the three components is related to one of the strata, intra-block, inter-block or total. The component vectors $\boldsymbol{y}_\alpha, \alpha = 1, 2, 3$, are thus obtainable by projecting $\boldsymbol{y}$ orthogonally onto relevant mutually orthogonal subspaces. The first component in (5.1) can be written as

$$\boldsymbol{y}_1 = \boldsymbol{\phi}_1 \boldsymbol{y}, \tag{5.2}$$

where

$$\boldsymbol{\phi}_1 = \boldsymbol{I}_n - \boldsymbol{X}_2 \boldsymbol{k}^{-\delta} \boldsymbol{X}_2' = \boldsymbol{I}_n - \boldsymbol{P}_{X_2} = \boldsymbol{P}_{X_2^\perp}, \tag{5.3}$$

i.e., $\boldsymbol{y}_1$ is the orthogonal projection of $\boldsymbol{y}$ on $\mathcal{R}^\perp(\boldsymbol{X}_2)$, the orthogonal complement of $\mathcal{R}(\boldsymbol{X}_2)$. The second component is

$$\boldsymbol{y}_2 = \boldsymbol{\phi}_2 \boldsymbol{y}, \tag{5.4}$$

where

$$\boldsymbol{\phi}_2 = \boldsymbol{X}_2 \boldsymbol{k}^{-\delta} \boldsymbol{X}_2' - n^{-1} \boldsymbol{1}_n \boldsymbol{1}_n' = \boldsymbol{P}_{X_2} - \boldsymbol{P}_{\boldsymbol{1}_n}, \tag{5.5}$$

i.e., $\boldsymbol{y}_2$ is the orthogonal projection of $\boldsymbol{y}$ on $\mathcal{R}^\perp(\boldsymbol{1}_n) \cap \mathcal{R}(\boldsymbol{X}_2)$, the orthogonal complement of $\mathcal{R}(\boldsymbol{1}_n)$ in $\mathcal{R}(\boldsymbol{X}_2)$. The third is

$$\boldsymbol{y}_3 = \boldsymbol{\phi}_3 \boldsymbol{y}, \tag{5.6}$$

where

$$\phi_3 = n^{-1}\mathbf{1}_n\mathbf{1}'_n = \boldsymbol{P_{1_n}}, \tag{5.7}$$

i.e., $\boldsymbol{y}_3$ is the orthogonal projection of $\boldsymbol{y}$ on $\mathcal{R}(\mathbf{1}_n)$. Evidently the three matrices (5.3), (5.5) and (5.7) satisfy the conditions

$$\phi_\alpha = \phi'_\alpha, \quad \phi_\alpha\phi_\alpha = \phi_\alpha, \quad \phi_\alpha\phi_{\alpha'} = \mathbf{O} \ \ \text{for} \ \alpha \neq \alpha', \tag{5.8}$$

where $\alpha, \alpha' = 1, 2, 3$, and the condition

$$\phi_1 + \phi_2 + \phi_3 = \boldsymbol{I}_n. \tag{5.9}$$

Note that the third equality in (5.8) implies in particular that $\phi_1\boldsymbol{X}_2 = \mathbf{O}$ and $\phi_\alpha\mathbf{1}_n = \mathbf{0}$ for $\alpha = 1, 2$, whereas the first two equalities in (5.8) imply that $\text{rank}(\phi_1) = n - b$, $\text{rank}(\phi_2) = b - 1$ and $\text{rank}(\phi_3) = 1$.

The resulting projections (5.2), (5.4) and (5.6) can be considered as submodels of the overall model (4.1). This can be seen by substituting (4.1) for $\boldsymbol{y}$ in (5.2), (5.4) and (5.6), respectively. The submodel (5.2) leads to the so-called intra-block analysis, resulting from the elimination of block effects, whereas (5.4) provides the so-called inter-block analysis, based on block totals. The submodel (5.6) underlies the total-area analysis, suitable mainly for estimating the general parametric mean.

To simplify the considerations, from now on only the so-called proper block designs (i.e., such for which $k_1 = k_2 = \cdots = k_b = k$), as in Oktaba (2003), will be taken into account.

It can easily be shown that for a proper block design

$$\text{D}(\boldsymbol{y}) = \phi_1\sigma_1^2 + \phi_2\sigma_2^2 + \phi_3\sigma_3^2, \tag{5.10}$$

with $\phi_1$, $\phi_2$ and $\phi_3$ defined above, and with the variance components (called stratum variances)

$$\begin{aligned}
\sigma_1^2 &= \sigma_U^2 + \sigma_e^2, \\
\sigma_2^2 &= k\sigma_B^2 + (1 - K_H^{-1}k)\sigma_U^2 + \sigma_e^2, \\
\sigma_3^2 &= (1 - N_B^{-1}b)k\sigma_B^2 + (1 - K_H^{-1}k)\sigma_U^2 + \sigma_e^2.
\end{aligned}$$

On account of (5.10), to follow the theory of the Gauss−Markov model, one would need to know the true values of these variances, which in practice are usually not available. Therefore, one has to consider the three submodels, resulting from the decomposition (5.1), separately. Each of them can be treated as a Gauss−Markov model with a simple dispersion matrix.

### 5.1. Intra-block submodel

The submodel (5.2) has the properties

$$\mathrm{E}(\boldsymbol{y}_1) = \boldsymbol{\phi}_1 \boldsymbol{X}_1 \boldsymbol{\tau} = \boldsymbol{X}_1 \boldsymbol{\tau} - k^{-1} \boldsymbol{X}_2 \boldsymbol{X}_2' \boldsymbol{X}_1 \boldsymbol{\tau} \qquad (5.11)$$

and

$$\mathrm{D}(\boldsymbol{y}_1) = \boldsymbol{\phi}_1 (\sigma_U^2 + \sigma_e^2) = \boldsymbol{\phi}_1 \sigma_1^2, \qquad (5.12)$$

where $\sigma_1^2 = \sigma_U^2 + \sigma_e^2$. From these, the main result concerning estimation under (5.2) can be stated as follows.

**Theorem 5.1** (3.2.1 in Caliński and Kageyama, 2000). Under (5.2), a function $\boldsymbol{a}' \boldsymbol{y}_1 = \boldsymbol{a}' \boldsymbol{\phi}_1 \boldsymbol{y}$ is uniformly the BLUE of $\boldsymbol{c}' \boldsymbol{\tau}$ if and only if $\boldsymbol{\phi}_1 \boldsymbol{a} = \boldsymbol{\phi}_1 \boldsymbol{X}_1 \boldsymbol{s}$, where the vectors $\boldsymbol{c}$ and $\boldsymbol{s}$ are in the relation $\boldsymbol{c} = \boldsymbol{X}_1' \boldsymbol{\phi}_1 \boldsymbol{X}_1 \boldsymbol{s}$ (i.e., $\boldsymbol{c} = \boldsymbol{C}_1 \boldsymbol{s}$, where $\boldsymbol{C}_1 = \boldsymbol{X}_1' \boldsymbol{\phi}_1 \boldsymbol{X}_1$).

Note that because $\boldsymbol{1}_v' \boldsymbol{X}_1' \boldsymbol{\phi}_1 = \boldsymbol{0}'$, the only parametric functions for which the BLUEs may exist under (5.2) are contrasts. If $\boldsymbol{c}' \boldsymbol{\tau}$ is a contrast, and the condition of Theorem 5.1 is satisfied, then the variance of its BLUE under (5.2), i.e., of $\widehat{\boldsymbol{c}' \boldsymbol{\tau}} = \boldsymbol{s}' \boldsymbol{X}_1' \boldsymbol{y}_1 = \boldsymbol{c}' \boldsymbol{C}_1^- \boldsymbol{X}_1' \boldsymbol{\phi}_1 \boldsymbol{y}$, is of the form

$$\mathrm{Var}(\widehat{\boldsymbol{c}' \boldsymbol{\tau}}) = \boldsymbol{s}' \boldsymbol{C}_1 \boldsymbol{s} \sigma_1^2 = \boldsymbol{c}' \boldsymbol{C}_1^- \boldsymbol{c} \sigma_1^2, \qquad (5.13)$$

where $\boldsymbol{C}_1^-$ is any $g$-inverse of the matrix $\boldsymbol{C}_1 = \boldsymbol{X}_1' \boldsymbol{\phi}_1 \boldsymbol{X}_1 = \boldsymbol{r}^\delta - k^{-1} \boldsymbol{N} \boldsymbol{N}'$.

Also note that the BLUE of the expectation vector $\boldsymbol{\phi}_1 \boldsymbol{X}_1 \boldsymbol{\tau}$ can be obtained by a simple least-squares procedure, in the form

$$\widehat{\mathrm{E}(\boldsymbol{y}_1)} = \boldsymbol{P}_{\phi_1 X_1} \boldsymbol{y}_1 = \boldsymbol{P}_{\phi_1 X_1} \boldsymbol{y}, \qquad (5.14)$$

where $\boldsymbol{P}_{\phi_1 X_1} = \boldsymbol{\phi}_1 \boldsymbol{X}_1 \boldsymbol{C}_1^- \boldsymbol{X}_1' \boldsymbol{\phi}_1$. To check this, in view of Theorem 2.3, note that $\mathrm{E}(\boldsymbol{P}_{\phi_1 X_1} \boldsymbol{y}_1) = \boldsymbol{\phi}_1 \boldsymbol{X}_1 \boldsymbol{\tau} = \mathrm{E}(\boldsymbol{y}_1)$ and that $\boldsymbol{P}_{\phi_1 X_1} \boldsymbol{\phi}_1 = \boldsymbol{\phi}_1 \boldsymbol{P}_{\phi_1 X_1}$. Furthermore, it follows that the vector $\boldsymbol{y}_1$ can be decomposed as

$$\begin{aligned} \boldsymbol{y}_1 &= \boldsymbol{P}_{\phi_1 X_1} \boldsymbol{y}_1 + (\boldsymbol{I}_n - \boldsymbol{P}_{\phi_1 X_1}) \boldsymbol{y}_1 \quad \text{(in terms of } \boldsymbol{y}_1 \text{)} \\ &= \boldsymbol{P}_{\phi_1 X_1} \boldsymbol{y} + (\boldsymbol{\phi}_1 - \boldsymbol{P}_{\phi_1 X_1}) \boldsymbol{y} \quad \text{(in terms of } \boldsymbol{y} \text{)}, \end{aligned}$$

where the second component on the right-hand side in each form of the decomposition is the residual vector providing the MINQUE of $\sigma_1^2$ (see Rao, 1974, Section 3). More precisely, taking the squared norm on both sides of the above decomposition of $\boldsymbol{y}_1$, one can write

$$\|\boldsymbol{y}_1\|^2 = \|\boldsymbol{P}_{\phi_1 X_1} \boldsymbol{y}_1\|^2 + \|(\boldsymbol{I}_n - \boldsymbol{P}_{\phi_1 X_1}) \boldsymbol{y}_1\|^2,$$

as $\boldsymbol{P}_{\phi_1 X_1}(\boldsymbol{I}_n - \boldsymbol{P}_{\phi_1 X_1}) = \boldsymbol{O}$. This equation provides the intra-block analysis of variance, which in terms of the observed vector $\boldsymbol{y}$ can be expressed in a more customary way as

$$\begin{aligned} \boldsymbol{y}'\phi_1\boldsymbol{y} &= \boldsymbol{y}'\phi_1\boldsymbol{X}_1\boldsymbol{C}_1^-\boldsymbol{X}_1'\phi_1\boldsymbol{y} + \boldsymbol{y}'(\phi_1 - \phi_1\boldsymbol{X}_1\boldsymbol{C}_1^-\boldsymbol{X}_1'\phi_1)\boldsymbol{y} \\ &= \boldsymbol{Q}_1'\boldsymbol{C}_1^-\boldsymbol{Q}_1 + \boldsymbol{y}'\psi_1\boldsymbol{y}, \end{aligned}$$

where $\boldsymbol{Q}_1 = \boldsymbol{X}_1'\phi_1\boldsymbol{y} = \boldsymbol{X}_1'\boldsymbol{y}_1$ and $\psi_1 = \phi_1 - \phi_1\boldsymbol{X}_1\boldsymbol{C}_1^-\boldsymbol{X}_1'\phi_1 = \phi_1(\boldsymbol{I}_n - \boldsymbol{X}_1\boldsymbol{C}_1^-\boldsymbol{X}_1')\phi_1$, with the properties $\psi_1' = \psi_1$, $\psi_1\psi_1 = \psi_1$ and $\psi_1\boldsymbol{X}_1 = \boldsymbol{O}$. The quadratic form $\boldsymbol{y}'\phi_1\boldsymbol{y}$ can be called the intra-block total sum of squares, and its components, $\boldsymbol{Q}_1'\boldsymbol{C}_1^-\boldsymbol{Q}_1$ and $\boldsymbol{y}'\psi_1\boldsymbol{y}$, can be called the intra-block treatment sum of squares and the intra-block residual sum of squares respectively. The corresponding d.f. are $n - b = \text{rank}(\phi_1)$ for the total, $h = \text{rank}(\boldsymbol{C}_1)$ for the treatment component and $n - b - h = \text{rank}(\psi_1)$ for the residual component. The expectations of these component sums of squares are

$$\text{E}(\boldsymbol{Q}_1'\boldsymbol{C}_1^-\boldsymbol{Q}_1) = h\sigma_1^2 + \boldsymbol{\tau}'\boldsymbol{C}_1\boldsymbol{\tau} \quad \text{and} \quad \text{E}(\boldsymbol{y}'\psi_1\boldsymbol{y}) = (n - b - h)\sigma_1^2.$$

It follows that the intra-block residual mean square $s_1^2 = \boldsymbol{y}'\psi_1\boldsymbol{y}/(n-b-h)$ is an unbiased estimator of $\sigma_1^2$. Moreover, $s_1^2$ is the MINQUE of $\sigma_1^2$ under the submodel (5.2), as may be seen from Theorem 3.4 of Rao (1974).

Thus $s_1^2$ can be used to obtain an unbiased estimator of the variance (5.13), in the form

$$\widehat{\text{Var}(\boldsymbol{c}'\boldsymbol{\tau})} = \boldsymbol{s}'\boldsymbol{C}_1\boldsymbol{s}s_1^2 = \boldsymbol{c}'\boldsymbol{C}_1^-\boldsymbol{c}s_1^2.$$

Furthermore, because under the multivariate normal distribution of $\boldsymbol{y}$, and hence of $\boldsymbol{y}_1$, the quadratic functions $\boldsymbol{Q}_1'\boldsymbol{C}_1^-\boldsymbol{Q}_1/\sigma_1^2$ and $\boldsymbol{y}'\psi_1\boldsymbol{y}/\sigma_1^2$ have independent $\chi^2$ distributions, the first noncentral with $h$ d.f. and the non-centrality parameter $\delta_1 = \boldsymbol{\tau}'\boldsymbol{C}_1\boldsymbol{\tau}/\sigma_1^2$, the second central with $n - b - h$ d.f., the hypothesis $\boldsymbol{\tau}'\boldsymbol{C}_1\boldsymbol{\tau} = 0$, equivalent to $\text{E}(\boldsymbol{y}_1) = \boldsymbol{0}$ [or $\text{E}(\boldsymbol{y}) \in \mathcal{R}(\boldsymbol{X}_2)$], can be tested by the statistic

$$F = h^{-1}\boldsymbol{Q}_1'\boldsymbol{C}_1^-\boldsymbol{Q}_1/s_1^2, \tag{5.15}$$

which under the normality assumption then has the $F$ distribution with $h$ and $n - b - h$ d.f., central when the hypothesis is true. Note that the $F$ statistic coincides with (2.15) taken for $\boldsymbol{B}_1 = \boldsymbol{I}_n$ and $\boldsymbol{t} = \boldsymbol{0}$, and with $\boldsymbol{F} = \boldsymbol{P}_{\phi_1 X_1}$, on account of (5.14).

### 5.2. Inter-block submodel

As to the submodel (5.4), it has the properties

$$\mathrm{E}(\boldsymbol{y}_2) = \boldsymbol{\phi}_2 \boldsymbol{X}_1 \boldsymbol{\tau} = k^{-1} \boldsymbol{X}_2 \boldsymbol{X}_2' \boldsymbol{X}_1 \boldsymbol{\tau} - n^{-1} \boldsymbol{1}_n \boldsymbol{r}' \boldsymbol{\tau} \tag{5.16}$$

and

$$\mathrm{D}(\boldsymbol{y}_2) = \boldsymbol{\phi}_2 \boldsymbol{X}_2 \boldsymbol{X}_2' \boldsymbol{\phi}_2 (\sigma_B^2 - K_H^{-1} \sigma_U^2) + \boldsymbol{\phi}_2 \sigma_1^2 = \boldsymbol{\phi}_2 \sigma_2^2. \tag{5.17}$$

The main estimation result under (5.4) can be expressed as follows.

**Theorem 5.2** (3.2.2 in Caliński and Kageyama, 2000). Under (5.4), a function $\boldsymbol{a}' \boldsymbol{y}_2 = \boldsymbol{a}' \boldsymbol{\phi}_2 \boldsymbol{y}$ is uniformly the BLUE of $\boldsymbol{c}' \boldsymbol{\tau}$ if and only if $\boldsymbol{\phi}_2 \boldsymbol{a} = \boldsymbol{\phi}_2 \boldsymbol{X}_1 \boldsymbol{s}$, where the vectors $\boldsymbol{c}$ and $\boldsymbol{s}$ are in the relation $\boldsymbol{c} = \boldsymbol{X}_1' \boldsymbol{\phi}_2 \boldsymbol{X}_1 \boldsymbol{s}$ (i.e., $\boldsymbol{c} = \boldsymbol{C}_2 \boldsymbol{s}$, where $\boldsymbol{C}_2 = \boldsymbol{X}_1' \boldsymbol{\phi}_2 \boldsymbol{X}_1$).

Now it may be noted that if the conditions of Theorem 5.2 are satisfied, then $\widehat{\boldsymbol{c}' \boldsymbol{\tau}} = \boldsymbol{s}' \boldsymbol{X}_1' \boldsymbol{y}_2 = \boldsymbol{c}' \boldsymbol{C}_2^- \boldsymbol{X}_1' \boldsymbol{\phi}_2 \boldsymbol{y}$ is the BLUE of the contrast $\boldsymbol{c}' \boldsymbol{\tau} = \boldsymbol{s}' \boldsymbol{C}_2 \boldsymbol{\tau}$ under (5.4), and that its variance has the form

$$\mathrm{Var}(\widehat{\boldsymbol{c}' \boldsymbol{\tau}}) = k^{-1} \boldsymbol{s}' \boldsymbol{N}_0 \boldsymbol{N}_0' \boldsymbol{s} \sigma_2^2 = \boldsymbol{c}' (k^{-1} \boldsymbol{N}_0 \boldsymbol{N}_0')^- \boldsymbol{c} \sigma_2^2, \tag{5.18}$$

where $\boldsymbol{N}_0 = \boldsymbol{N} - b^{-1} \boldsymbol{r} \boldsymbol{1}_b'$.

Furthermore, it should be noted that the BLUE of the expectation vector (5.16) is obtainable by a simple least-squares procedure, i.e., has the form $\widehat{\mathrm{E}(\boldsymbol{y}_2)} = \boldsymbol{P}_{\phi_2 X_1} \boldsymbol{y}_2 = \boldsymbol{P}_{\phi_2 X_1} \boldsymbol{y}$, where $\boldsymbol{P}_{\phi_2 X_1} = \boldsymbol{\phi}_2 \boldsymbol{X}_1 \boldsymbol{C}_2^- \boldsymbol{X}_1' \boldsymbol{\phi}_2$. This result can be checked, referring to Theorem 2.3, by noting that

(a) $\boldsymbol{P}_{\phi_2 X_1} \boldsymbol{\phi}_2 \boldsymbol{X}_1 = \boldsymbol{\phi}_2 \boldsymbol{X}_1$,
(b) $\mathcal{R}[\boldsymbol{P}_{\phi_2 X_1} \mathrm{D}(\boldsymbol{y}_2)] \subset \mathcal{R}(\boldsymbol{\phi}_2 \boldsymbol{X}_1)$,
(c) $\boldsymbol{P}_{\phi_2 X_1} \mathrm{D}(\boldsymbol{y}_2) = \mathrm{D}(\boldsymbol{y}_2) \boldsymbol{P}_{\phi_2 X_1}$.

Thus the vector $\boldsymbol{y}_2$ can then be decomposed as

$$\begin{aligned} \boldsymbol{y}_2 &= \boldsymbol{P}_{\phi_2 X_1} \boldsymbol{y}_2 + (\boldsymbol{I}_n - \boldsymbol{P}_{\phi_2 X_1}) \boldsymbol{y}_2 \quad \text{(in terms of } \boldsymbol{y}_2) \\ &= \boldsymbol{P}_{\phi_2 X_1} \boldsymbol{y} + (\boldsymbol{\phi}_2 - \boldsymbol{P}_{\phi_2 X_1}) \boldsymbol{y} \quad \text{(in terms of } \boldsymbol{y}), \end{aligned}$$

with $(\boldsymbol{I}_n - \boldsymbol{P}_{\phi_2 X_1}) \boldsymbol{y}_2 = (\boldsymbol{\phi}_2 - \boldsymbol{P}_{\phi_2 X_1}) \boldsymbol{y}$ as the residual vector.

The above decomposition yields the inter-block analysis of variance, of the form

$$\|\boldsymbol{y}_2\|^2 = \|\boldsymbol{P}_{\phi_2 X_1} \boldsymbol{y}_2\|^2 + \|(\boldsymbol{I}_n - \boldsymbol{P}_{\phi_2 X_1}) \boldsymbol{y}_2\|^2,$$

expressible with the use of the observed vector $\boldsymbol{y}$ as

$$\begin{aligned} \boldsymbol{y}' \boldsymbol{\phi}_2 \boldsymbol{y} &= \boldsymbol{y}' \boldsymbol{\phi}_2 \boldsymbol{X}_1 \boldsymbol{C}_2^- \boldsymbol{X}_1' \boldsymbol{\phi}_2 \boldsymbol{y} + \boldsymbol{y}' (\boldsymbol{\phi}_2 - \boldsymbol{\phi}_2 \boldsymbol{X}_1 \boldsymbol{C}_2^- \boldsymbol{X}_1' \boldsymbol{\phi}_2) \boldsymbol{y} \\ &= \boldsymbol{Q}_2' \boldsymbol{C}_2^- \boldsymbol{Q}_2 + \boldsymbol{y}' \boldsymbol{\psi}_2 \boldsymbol{y}, \end{aligned}$$

where $\boldsymbol{Q}_2 = \boldsymbol{X}_1'\boldsymbol{\phi}_2\boldsymbol{y}$, $\boldsymbol{C}_2 = \boldsymbol{X}_1'\boldsymbol{\phi}_2\boldsymbol{X}_1 = k^{-1}\boldsymbol{N}\boldsymbol{N}' - n^{-1}\boldsymbol{r}\boldsymbol{r}'$ and $\boldsymbol{\psi}_2 = \boldsymbol{\phi}_2 - \boldsymbol{\phi}_2\boldsymbol{X}_1\boldsymbol{C}_2^-\boldsymbol{X}_1'\boldsymbol{\phi}_2 = \boldsymbol{\phi}_2(\boldsymbol{I}_n - \boldsymbol{X}_1\boldsymbol{C}_2^-\boldsymbol{X}_1')\boldsymbol{\phi}_2$, the matrix $\boldsymbol{\psi}_2$ having the readily seen properties

$$\boldsymbol{\psi}_2\boldsymbol{\phi}_2 = \boldsymbol{\psi}_2, \ \boldsymbol{\psi}_2\boldsymbol{1}_n = \boldsymbol{0}, \ \boldsymbol{\psi}_2\boldsymbol{X}_1 = \boldsymbol{O}, \ \boldsymbol{\psi}_2' = \boldsymbol{\psi}_2, \ \text{and} \ \boldsymbol{\psi}_2\boldsymbol{\psi}_2 = \boldsymbol{\psi}_2.$$

Also note that $\mathrm{rank}(\boldsymbol{C}_2) = v - \rho - 1$, $\mathrm{rank}(\boldsymbol{\phi}_2\boldsymbol{X}_1\boldsymbol{C}_2^-\boldsymbol{X}_1'\boldsymbol{\phi}_2) = \mathrm{rank}(\boldsymbol{C}_2)$ and $\mathrm{rank}(\boldsymbol{\psi}_2) = \mathrm{rank}(\boldsymbol{\phi}_2) - \mathrm{rank}(\boldsymbol{C}_2) = b - v + \rho$, if $\mathrm{rank}(\boldsymbol{N}) = v - \rho$.

The quadratic form $\boldsymbol{y}'\boldsymbol{\phi}_2\boldsymbol{y}$ can be called the inter-block total sum of squares, whereas its components, $\boldsymbol{Q}_2'\boldsymbol{C}_2^-\boldsymbol{Q}_2$ and $\boldsymbol{y}'\boldsymbol{\psi}_2\boldsymbol{y} = \boldsymbol{y}'\boldsymbol{\phi}_2(\boldsymbol{I}_n - \boldsymbol{X}_1\boldsymbol{C}_2^-\boldsymbol{X}_1')\boldsymbol{\phi}_2\boldsymbol{y}$, can be called the inter-block treatment sum of squares and the inter-block residual sum of squares respectively. The expectations of these component sums of squares are

$$\mathrm{E}(\boldsymbol{Q}_2'\boldsymbol{C}_2^-\boldsymbol{Q}_2) = (v - \rho - 1)\sigma_2^2 + \boldsymbol{\tau}'\boldsymbol{C}_2\boldsymbol{\tau}$$

and

$$\mathrm{E}(\boldsymbol{y}'\boldsymbol{\psi}_2\boldsymbol{y}) = (b - v + \rho)\sigma_2^2.$$

It follows from the latter equality that the inter-block residual mean square $s_2^2 = \boldsymbol{y}'\boldsymbol{\psi}_2\boldsymbol{y}/(b - v + \rho)$ is an unbiased estimator of

$$\sigma_2^2 = k\sigma_B^2 + (1 - K_H^{-1}k)\sigma_U^2 + \sigma_e^2, \tag{5.19}$$

in fact the MINQUE of $\sigma_2^2$. It should be noted, however, that $b - v + \rho = 0$ if $b = v - \rho$. In that case, no estimator for $\sigma_2^2$ exists in the inter-block analysis.

Thus, in the case of equal $k_j$'s and $b > v - \rho$, the mean square $s_2^2$ can be used to obtain an unbiased estimator of the variance (5.18), in the form

$$\widehat{\mathrm{Var}(\boldsymbol{c}'\boldsymbol{\tau})} = k^{-1}\boldsymbol{s}'\boldsymbol{N}_0\boldsymbol{N}_0'\boldsymbol{s}s_2^2 = k\boldsymbol{c}'(\boldsymbol{N}_0\boldsymbol{N}_0')^-\boldsymbol{c}s_2^2.$$

Furthermore, as for the intra-block analysis, it can be shown that under the multivariate normality assumption the quadratic functions $\boldsymbol{Q}_2'\boldsymbol{C}_2^-\boldsymbol{Q}_2/\sigma_2^2$ and $\boldsymbol{y}'\boldsymbol{\psi}_2\boldsymbol{y}/\sigma_2^2$ have independent $\chi^2$ distributions, the first noncentral with $v - \rho - 1$ d.f. and the noncentrality parameter $\delta_2 = \boldsymbol{\tau}'\boldsymbol{C}_2\boldsymbol{\tau}/\sigma_2^2$, the second central with $b - v + \rho$ d.f. Hence the hypothesis $\boldsymbol{\tau}'\boldsymbol{C}_2\boldsymbol{\tau} = 0$, equivalent to $\mathrm{E}(\boldsymbol{y}_2) = \boldsymbol{0}$ [or $\boldsymbol{P}_{X_2}\mathrm{E}(\boldsymbol{y}) \in \mathcal{R}(\boldsymbol{1}_n)$], can be tested by the statistic

$$F = (v - \rho - 1)^{-1}\boldsymbol{Q}_2'\boldsymbol{C}_2^-\boldsymbol{Q}_2/s_2^2, \tag{5.20}$$

which under the assumed normality then has the $F$ distribution with $v - \rho - 1$ and $b - v + \rho$ d.f., central when the hypothesis is true.

### 5.3. Total-area submodel

Considering the third submodel, given in (5.6), note that its properties are

$$\mathrm{E}(\boldsymbol{y}_3) = \boldsymbol{\phi}_3 \boldsymbol{X}_1 \boldsymbol{\tau} = n^{-1} \mathbf{1}_n \boldsymbol{r}' \boldsymbol{\tau} \tag{5.21}$$

and

$$\mathrm{D}(\boldsymbol{y}_3) = \boldsymbol{\phi}_3 [(k - N_B^{-1} n) \sigma_B^2 + (1 - K_H^{-1} k) \sigma_U^2 + \sigma_e^2] = \boldsymbol{\phi}_3 \sigma_3^2. \tag{5.22}$$

These lead to the following main result concerning estimation under (5.6).

**Theorem 5.3** (3.2.3 in Caliński and Kageyama, 2000). Under (5.6), a function $\boldsymbol{a}' \boldsymbol{y}_3 = \boldsymbol{a}' \boldsymbol{\phi}_3 \boldsymbol{y}$ is uniformly the BLUE of $\boldsymbol{c}' \boldsymbol{\tau}$ if and only if $\boldsymbol{\phi}_3 \boldsymbol{a} = \boldsymbol{\phi}_3 \boldsymbol{X}_1 \boldsymbol{s}$, where the vectors $\boldsymbol{c}$ and $\boldsymbol{s}$ are in the relation $\boldsymbol{c} = \boldsymbol{X}_1' \boldsymbol{\phi}_3 \boldsymbol{X}_1 \boldsymbol{s}$ $(= n^{-1} \boldsymbol{r} \boldsymbol{r}' \boldsymbol{s}$, as $\boldsymbol{C}_3 = \boldsymbol{X}_1' \boldsymbol{\phi}_3 \boldsymbol{X}_1 = n^{-1} \boldsymbol{r} \boldsymbol{r}')$.

It may be interesting to note the following.

(a) The only parametric functions for which the BLUEs under (5.6) exist are those defined as $\boldsymbol{c}' \boldsymbol{\tau} = (\boldsymbol{s}' \boldsymbol{r}) n^{-1} \boldsymbol{r}' \boldsymbol{\tau}$, where $\boldsymbol{s}' \boldsymbol{r} \neq 0$, i.e., the general parametric mean and its multiplicities, contrasts being excluded *a fortiori* (as it follows that $\mathbf{1}_v' \boldsymbol{c} = \boldsymbol{r}' \boldsymbol{s}$ here).

(b) Because of the equality $\mathrm{D}(\boldsymbol{y}_3) \boldsymbol{\phi}_3 \boldsymbol{X}_1 = \boldsymbol{\phi}_3 \boldsymbol{X}_1 \sigma_3^2$, the BLUEs under (5.6) and the SLSEs are the same (on account of Theorem 2.2, applied to Theorem 5.3).

If $\boldsymbol{c}' \boldsymbol{\tau} = (\boldsymbol{s}' \boldsymbol{r}) n^{-1} \boldsymbol{r}' \boldsymbol{\tau} = (\boldsymbol{c}' \mathbf{1}_v) n^{-1} \boldsymbol{r}' \boldsymbol{\tau}$, then the variance of its BLUE under (5.6), i.e., of $\widehat{\boldsymbol{c}' \boldsymbol{\tau}} = \boldsymbol{s}' \boldsymbol{X}_1' \boldsymbol{y}_3 = (\boldsymbol{s}' \boldsymbol{r}) n^{-1} \mathbf{1}_n' \boldsymbol{y} = (\boldsymbol{c}' \mathbf{1}_v) n^{-1} \mathbf{1}_n' \boldsymbol{y}$, is of the form

$$\begin{aligned}
\mathrm{Var}(\widehat{\boldsymbol{c}' \boldsymbol{\tau}}) \quad &= n^{-1} (\boldsymbol{c}' \mathbf{1}_v)^2 [(1 - N_B^{-1} b) k \sigma_B^2 + (1 - K_H^{-1} k) \sigma_U^2 + \sigma_e^2] \\
&= n^{-1} (\boldsymbol{c}' \mathbf{1}_v)^2 \sigma_3^2.
\end{aligned}$$

Finally, it may be noticed that because $\boldsymbol{P}_{\boldsymbol{\phi}_3 X_1} = \boldsymbol{\phi}_3$ (as $n^{-1} \mathbf{1}_v \mathbf{1}_v'$ is a $g$-inverse of $\boldsymbol{C}_3 = n^{-1} \boldsymbol{r} \boldsymbol{r}'$), the function $\boldsymbol{P}_{\boldsymbol{\phi}_3 X_1} \boldsymbol{y}_3$ satisfies the conditions of Theorem 2.3, and hence the vector $\boldsymbol{P}_{\boldsymbol{\phi}_3 X_1} \boldsymbol{y}_3 = \boldsymbol{y}_3 = n^{-1} \mathbf{1}_n \mathbf{1}_n' \boldsymbol{y}$ is itself the BLUE of its expectation, $n^{-1} \mathbf{1}_n \boldsymbol{r}' \boldsymbol{\tau}$, leaving no residuals.

## 6. Recovery of inter-block information

The intra-block submodel considered in Section 5.1 can well be seen as a Gauss−Markov model, with the properties indicated in (5.11) and (5.12). However, it does not provide complete information concerning the original

model considered in Section 4. Therefore it is of interest to recover the information contained in the inter-block submodel considered in Section 5.2, with the properties presented in (5.16) and (5.17). Following Section 4, that recovery will be considered here for proper block designs only.

## 6.1. BLUEs under known stratum variances

First note that if for a proper block design the stratum variances $\sigma_1^2$, $\sigma_2^2$ and $\sigma_3^2$ are known, or at least the ratio $\sigma_1^2/\sigma_2^2$ and the difference $\sigma_2^2 - \sigma_3^2 = N_B^{-1} n \sigma_B^2$ are known, then the BLUE of $\boldsymbol{\tau}$ can be written as

$$\hat{\boldsymbol{\tau}} = \boldsymbol{C}_c^{-1} \boldsymbol{Q}_c, \tag{6.1}$$

where

$$\boldsymbol{C}_c = \boldsymbol{X}_1'[\boldsymbol{\phi}_1 + (\sigma_1^2/\sigma_2^2)(\boldsymbol{I}_n - \boldsymbol{\phi}_1)]\boldsymbol{X}_1 \quad \text{and} \quad \boldsymbol{Q}_c = \boldsymbol{X}_1'[\boldsymbol{\phi}_1 + (\sigma_1^2/\sigma_2^2)(\boldsymbol{I}_n - \boldsymbol{\phi}_1)]\boldsymbol{y},$$

and its dispersion matrix as

$$\mathrm{D}(\hat{\boldsymbol{\tau}}) = \sigma_1^2 \boldsymbol{C}_c^{-1} - N_B^{-1}\sigma_B^2 \mathbf{1}_v \mathbf{1}_v' = \sigma_1^2 \boldsymbol{C}_c^{-1} - (\sigma_2^2 - \sigma_3^2)n^{-1}\mathbf{1}_v \mathbf{1}_v'. \tag{6.2}$$

Thus, under this knowledge of stratum variances, for any proper block design, the BLUE of a function $\boldsymbol{c}'\boldsymbol{\tau}$ is

$$\widehat{\boldsymbol{c}'\boldsymbol{\tau}} = \boldsymbol{c}'\boldsymbol{C}_c^{-1}\boldsymbol{Q}_c, \tag{6.3}$$

with the variance

$$\mathrm{Var}(\widehat{\boldsymbol{c}'\boldsymbol{\tau}}) = \sigma_1^2 \boldsymbol{c}'\boldsymbol{C}_c^{-1}\boldsymbol{c} - (\sigma_2^2 - \sigma_3^2)n^{-1}\boldsymbol{c}'\mathbf{1}_v \mathbf{1}_v'\boldsymbol{c}. \tag{6.4}$$

Evidently, if $\boldsymbol{c}'\boldsymbol{\tau}$ is a contrast, then the variance (6.4) reduces to

$$\mathrm{Var}(\widehat{\boldsymbol{c}'\boldsymbol{\tau}}) = \sigma_1^2 \boldsymbol{c}'\boldsymbol{C}_c^{-1}\boldsymbol{c}.$$

These results follow from Theorem 3.7.1 in Caliński and Kageyama (2000), which for proper block designs can be restated as follows.

**Theorem 6.1** (3.8.1 in Caliński and Kageyama, 2000). If the design is proper, then under the model (4.1), with properties (4.2) and (5.10), assuming that the variance components defined there are known and positive (or at least the ratio $\sigma_1^2/\sigma_2^2$ is known and positive), a function $\boldsymbol{a}'\boldsymbol{y}$ is the BLUE of a function $\boldsymbol{c}'\boldsymbol{\tau}$ if and only if

$$\begin{aligned} \boldsymbol{a} &= [\boldsymbol{\phi}_1 + (\sigma_1^2/\sigma_2^2)(\boldsymbol{I}_n - \boldsymbol{\phi}_1)]\boldsymbol{X}_1\boldsymbol{s} \\ &= [\boldsymbol{\phi}_1 + (\sigma_1^2/\sigma_2^2)(\boldsymbol{\phi}_2 + \boldsymbol{\phi}_3)]\boldsymbol{X}_1\boldsymbol{s}, \end{aligned} \tag{6.5}$$

where the vectors $\boldsymbol{c}$ and $\boldsymbol{s}$ are in the relation

$$
\begin{aligned}
\boldsymbol{c} &= \boldsymbol{X}_1'[\boldsymbol{\phi}_1 + (\sigma_1^2/\sigma_2^2)(\boldsymbol{I}_n - \boldsymbol{\phi}_1)]\boldsymbol{X}_1\boldsymbol{s} \\
&= \boldsymbol{X}_1'[\boldsymbol{\phi}_1 + (\sigma_1^2/\sigma_2^2)(\boldsymbol{\phi}_2 + \boldsymbol{\phi}_3)]\boldsymbol{X}_1\boldsymbol{s}.
\end{aligned} \tag{6.6}
$$

Note that the equalities (6.5) and (6.6) give jointly the result (6.3), with $\boldsymbol{C}_c$ and $\boldsymbol{Q}_c$ as given above. That the matrix $\boldsymbol{C}_c = \boldsymbol{X}_1'[\boldsymbol{\phi}_1 + (\sigma_1^2/\sigma_2^2)(\boldsymbol{I}_n - \boldsymbol{\phi}_1)]\boldsymbol{X}_1$ is p.d. follows from the fact that $\boldsymbol{\phi}_1 + (\sigma_1^2/\sigma_2^2)(\boldsymbol{I}_n - \boldsymbol{\phi}_1)$ is p.d. and that the matrix $\boldsymbol{X}_1$ is of rank $v$. As (6.3) is true for any conformable vector $\boldsymbol{c}$, formula (6.1) follows immediately. It can also be seen as the solution of the normal equations $\boldsymbol{C}_c\hat{\boldsymbol{\tau}} = \boldsymbol{Q}_c$, i.e., the equations

$$
\boldsymbol{X}_1'[\boldsymbol{\phi}_1 + (\sigma_1^2/\sigma_2^2)(\boldsymbol{I}_n - \boldsymbol{\phi}_1)]\boldsymbol{X}_1\boldsymbol{\tau} = \boldsymbol{X}_1'[\boldsymbol{\phi}_1 + (\sigma_1^2/\sigma_2^2)(\boldsymbol{I}_n - \boldsymbol{\phi}_1)]\boldsymbol{y}.
$$

Now it is interesting to compare the results in (6.3) and (6.4) with the corresponding results in Section 5.1. It then becomes evident that in a proper block design the matrix $(\sigma_1^2/\sigma_2^2)(\boldsymbol{I}_n - \boldsymbol{\phi}_1) = (\sigma_1^2/\sigma_2^2)k^{-1}\boldsymbol{X}_2\boldsymbol{X}_2'$ is responsible for the recovery of inter-block information . More precisely, the amount of the recovery depends very much on the ratio $\sigma_1^2/\sigma_2^2$, the gain from the recovery increasing with an increase in the value of this ratio (or a decrease in the value of $\sigma_2^2/\sigma_1^2$). However, to get a better insight into the process of recovery of inter-block information in a proper block design, it may be interesting to look at the problem in terms of the estimators of the so-called basic contrasts. For this reason, it will be useful to consider first the following corollary, presented with $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$ as defined in Sections 5.1 and 5.2, respectively.

**Corollary 6.1** (3.8.1 in Caliński and Kageyama, 2000). If the design is proper, then for any $\boldsymbol{c} = \boldsymbol{r}^\delta\boldsymbol{s}$ such that $\boldsymbol{s}$ satisfies the eigenvector condition

$$
\boldsymbol{C}_1\boldsymbol{s} = \varepsilon\boldsymbol{r}^\delta\boldsymbol{s} \quad \text{with} \ \ 0 < \varepsilon < 1, \tag{6.7}
$$

or its equivalence

$$
\boldsymbol{C}_2\boldsymbol{s} = (1 - \varepsilon)\boldsymbol{r}^\delta\boldsymbol{s} \quad \text{with} \ \ 0 < \varepsilon < 1, \tag{6.8}
$$

the BLUE of the contrast $\boldsymbol{c}'\boldsymbol{\tau}$ is, under the assumptions of Theorem 6.1, of the form

$$
\begin{aligned}
\widehat{\boldsymbol{c}'\boldsymbol{\tau}} = \widehat{\boldsymbol{s}'\boldsymbol{r}^\delta\boldsymbol{\tau}} &= w_1\varepsilon^{-1}\boldsymbol{s}'\boldsymbol{Q}_1 + w_2(1 - \varepsilon)^{-1}\boldsymbol{s}'\boldsymbol{Q}_2 \\
&= w_1(\widehat{\boldsymbol{c}'\boldsymbol{\tau}})_1 + w_2(\widehat{\boldsymbol{c}'\boldsymbol{\tau}})_2,
\end{aligned} \tag{6.9}
$$

where the components $(\widehat{c'\tau})_1 = \varepsilon^{-1}s'Q_1 = \varepsilon^{-1}c'r^{-\delta}Q_1$ and $(\widehat{c'\tau})_2 = (1-\varepsilon)^{-1}s'Q_2 = (1-\varepsilon)^{-1}c'r^{-\delta}Q_2$, with $Q_\alpha = X'_1\phi_\alpha y$, $\alpha = 1, 2$, are uncorrelated, and their "weights" are

$$w_1 = \frac{\varepsilon\sigma_2^2}{\varepsilon\sigma_2^2 + (1-\varepsilon)\sigma_1^2} \quad \text{and} \quad w_2 = \frac{(1-\varepsilon)\sigma_1^2}{\varepsilon\sigma_2^2 + (1-\varepsilon)\sigma_1^2}. \tag{6.10}$$

It may be seen that the result of Corollary 6.1 coincides with that given by Houtman and Speed (1983, p. 1078) for BIB designs, except that their model is different, with (in the present notation) $\sigma_1^2 = \sigma_e^2$ and $\sigma_2^2 = k\sigma_B^2 + \sigma_e^2$, i.e., with $\sigma_U^2$ ignored. In general, a function $c'\tau$ for which the BLUE is of the combined form (6.9) can be said, after Martin and Zyskind (1966), to be best combinable by simple weighting of intra-block and inter-block BLUEs, or shortly, to be "simply combinable."

It may also be noted that the equivalent conditions (6.7) and (6.8) are not only sufficient (as stated in Corollary 6.1), but also necessary for the contrast $c'\tau = s'r^\delta\tau$ to be simply combinable (a result originally established by Martin and Zyskind, 1966, Theorem 2.6). Moreover, it follows from Corollary 6.1 and the corresponding formula

$$\text{Var}[(\widehat{c'\tau})_\alpha] = \varepsilon_\alpha^{-1}s'r^\delta s\sigma_\alpha^2 = \varepsilon_\alpha^{-1}c'r^{-\delta}c\sigma_\alpha^2, \quad \alpha = 1, 2, \tag{6.11}$$

where $\varepsilon_1 = \varepsilon$ and $\varepsilon_2 = 1 - \varepsilon$, that, in the case of a proper block design and under the assumptions of Theorem 6.1, for any contrast $c'\tau = s'r^\delta\tau$ such that $s$ satisfies the equivalent conditions (6.7) and (6.8), the variance of the BLUE (6.9) is

$$\text{Var}(\widehat{c'\tau}) = \frac{\sigma_1^2\sigma_2^2}{\varepsilon\sigma_2^2 + (1-\varepsilon)\sigma_1^2}c'r^{-\delta}c, \tag{6.12}$$

for which the inequalities

$$\text{Var}(\widehat{c'\tau}) \leq \text{Var}[(\widehat{c'\tau})_1], \tag{6.13}$$

$$\text{Var}(\widehat{c'\tau}) \leq \text{Var}[(\widehat{c'\tau})_2] \tag{6.14}$$

hold, the equality in (6.13) holding if and only if $\varepsilon = 1$, and that in (6.14) if and only if $\varepsilon = 0$.

The results indicated above show that a combination of the intra-block and inter-block BLUEs of a contrast for which the conditions of Corollary 6.1 are satisfied (with $0 < \varepsilon < 1$) may lead to an improved estimator. This improvement will certainly be achieved when the stratum variances

$\sigma_1^2$ and $\sigma_2^2$ are known (or their ratio is known). The gain obtainable from using the BLUE (6.9) instead of the intra-block BLUE, $(\widehat{c'\tau})_1 = \varepsilon^{-1}s'Q_1 = \varepsilon^{-1}c'r^{-\delta}Q_1$, can be measured by considering the ratio of precisions

$$\frac{\mathrm{Var}[(\widehat{c'\tau})_1]}{\mathrm{Var}(\widehat{c'\tau})} = 1 + \frac{(1-\varepsilon)\sigma_1^2}{\varepsilon\sigma_2^2} \quad \text{(for a proper design)}.$$

Thus the gain, $\varepsilon^{-1}(1-\varepsilon)(\sigma_1^2/\sigma_2^2) = w_2/w_1$, depends on the eigenvalue $\varepsilon$ and the variance ratio

$$\frac{\sigma_1^2}{\sigma_2^2} = \frac{\sigma_U^2 + \sigma_e^2}{k\sigma_B^2 + (1 - K_H^{-1}k)\sigma_U^2 + \sigma_e^2} \quad (= \frac{\sigma_U^2 + \sigma_e^2}{k\sigma_B^2 + \sigma_e^2} \text{ if } k = K_H).$$

The gain would be 0 in the case of $\varepsilon = 1$, i.e., if the incidence matrix $N$ of the design were such that $N's = 0$. Hence it becomes evident that $\varepsilon$ can be interpreted as the "efficiency factor" of the analysed design for the contrast $c'\tau = s'r^\delta\tau$ when it is estimated in the intra-block analysis. On the other hand, $1-\varepsilon$ can be interpreted as the relative "loss of information", incurred in that analysis, due to partially confounding the contrast with blocks. This terminology is from Jones (1959, p. 176).

The above observation means that the recovery of inter-block information (as it is called after Yates, 1939, 1940) is worth undertaking, particularly when the efficiency factor in the intra-block analysis, $\varepsilon$, is small and the inter-block variation does not much exceed the variation of units within blocks. Thus the gain to be expected is larger, the less one is successful in (i) choosing a design with high intra-block efficiency for the contrasts of interest, and (ii) blocking the available experimental units aimed at removing their intra-block variation.

Before proceeding further, it will be useful to recall the concept of basic contrasts introduced by Pearce, Caliński and Marshall (1974).

**Definition 6.1** (3.4.1 in Caliński and Kageyama, 2000). For any block design, contrasts $\{c_i'\tau = s_i'r^\delta\tau, s_i'r = 0, i = 1, 2, ..., v-1\}$ are said to be basic contrasts of the design if the vectors $\{s_i\}$ are $r^\delta$-orthonormal eigenvectors of the matrix $C_1 = X_1'\phi_1 X_1$ of the design with respect to $r^\delta$.

In connection with Definition 6.1, the matrix $C_1$ can be written in spectral decomposition form as

$$C_1 = X_1'\phi_1 X_1 = r^\delta \sum_{\beta=0}^{m-1} \varepsilon_\beta H_\beta r^\delta = r^\delta \sum_{\beta=0}^{m-1} \varepsilon_\beta L_\beta, \tag{6.15}$$

where

$$\boldsymbol{H}_\beta = \sum_{j=1}^{\rho_\beta} \boldsymbol{s}_{\beta j}\boldsymbol{s}'_{\beta j} = \boldsymbol{S}_\beta \boldsymbol{S}'_\beta \ \ \text{and} \ \ \boldsymbol{L}_\beta = \boldsymbol{H}_\beta \boldsymbol{r}^\delta,$$

the columns of the matrix $\boldsymbol{S}_\beta = [\boldsymbol{s}_{\beta 1} : \boldsymbol{s}_{\beta 2} : \cdots : \boldsymbol{s}_{\beta\rho_\beta}]$ being $\boldsymbol{r}^\delta$-orthonormal eigenvectors of $\boldsymbol{C}_1$ with respect to $\boldsymbol{r}^\delta$ corresponding to a common eigenvalue $\varepsilon_\beta$ of multiplicity $\rho_\beta$, for $\beta = 0, 1, ..., m-1$ [with $\rho_0 \equiv \rho$ if $\mathrm{rank}(\boldsymbol{N}) = v - \rho$], and where $m-1$ is the number of distinct, less than 1, nonzero (positive) eigenvalues of $\boldsymbol{C}_1$ with respect to $\boldsymbol{r}^\delta$. It is customary to order these eigenvalues as $1 = \varepsilon_0 > \varepsilon_1 > \cdots > \varepsilon_{m-1} > \varepsilon_m = 0$. Similarly to (6.15), one can also write

$$\boldsymbol{C}_2 = \boldsymbol{X}'_1 \boldsymbol{\phi}_2 \boldsymbol{X}_1 = \boldsymbol{r}^\delta \sum_{\beta=1}^m (1 - \varepsilon_\beta) \boldsymbol{H}_\beta \boldsymbol{r}^\delta = \boldsymbol{r}^\delta \sum_{\beta=1}^m (1 - \varepsilon_\beta) \boldsymbol{L}_\beta, \qquad (6.16)$$

and

$$\boldsymbol{C}_3 = \boldsymbol{X}'_1 \boldsymbol{\phi}_3 \boldsymbol{X}_1 = \boldsymbol{r}^\delta \boldsymbol{H}_{m+1} \boldsymbol{r}^\delta \quad (= n^{-1}\boldsymbol{r}\boldsymbol{r}'), \qquad (6.17)$$

where $\boldsymbol{H}_1, ..., \boldsymbol{H}_{m-1}$ and $\boldsymbol{L}_1, ..., \boldsymbol{L}_{m-1}$ are as in (6.15), whereas

$$\boldsymbol{H}_m = \sum_{j=1}^{\rho_m} \boldsymbol{s}_{mj}\boldsymbol{s}'_{mj} = \boldsymbol{S}_m \boldsymbol{S}'_m \ \ \text{and} \ \ \boldsymbol{L}_m = \boldsymbol{H}_m \boldsymbol{r}^\delta,$$

the columns of the matrix $\boldsymbol{S}_m = [\boldsymbol{s}_{m1} : \boldsymbol{s}_{m2} : \cdots : \boldsymbol{s}_{m\rho_m}]$ being $\boldsymbol{r}^\delta$-orthonormal eigenvectors of $\boldsymbol{C}_1$ with respect to $\boldsymbol{r}^\delta$ corresponding to the common eigenvalue $\varepsilon_m = 0$ of multiplicity $\rho_m$, and where

$$\boldsymbol{H}_{m+1} = \boldsymbol{s}_v \boldsymbol{s}'_v, \quad \text{with} \quad \boldsymbol{s}_v = n^{-1/2}\boldsymbol{1}_v.$$

Thus it follows that the matrices $\{\boldsymbol{S}_\beta\}$ in $\boldsymbol{H}_\beta = \boldsymbol{S}_\beta \boldsymbol{S}'_\beta$ for $\beta = 0, 1, ..., m, m+1$ ($\boldsymbol{H}_{m+1} = \boldsymbol{s}_v \boldsymbol{s}'_v$) are such that $\boldsymbol{S}'_\beta \boldsymbol{r}^\delta \boldsymbol{S}_\beta = \boldsymbol{I}_{\rho_\beta}$ for any $\beta$ and $\boldsymbol{S}'_\beta \boldsymbol{r}^\delta \boldsymbol{S}_{\beta'} = \boldsymbol{O}$ for $\beta \neq \beta'$.

   Now it will be interesting to find, still under the assumptions of Theorem 6.1, the BLUE of $\mathrm{E}(\boldsymbol{y}) = \boldsymbol{X}_1 \boldsymbol{\tau}$ and hence the residuals $\boldsymbol{y} - \widehat{\boldsymbol{X}_1 \boldsymbol{\tau}}$. The BLUE of $\boldsymbol{X}_1 \boldsymbol{\tau}$ can be expressed as

$$\widehat{\boldsymbol{X}_1 \boldsymbol{\tau}} = \boldsymbol{P}_{X_1(V^{-1})}\boldsymbol{y}, \qquad (6.18)$$

where $\boldsymbol{V} = \mathrm{D}(\boldsymbol{y}) = \boldsymbol{\phi}_1 \sigma_1^2 + \boldsymbol{\phi}_2 \sigma_2^2 + \boldsymbol{\phi}_3 \sigma_3^2$ and

$$\boldsymbol{P}_{X_1(V^{-1})} = \boldsymbol{X}_1 (\boldsymbol{X}'_1 \boldsymbol{V}^{-1} \boldsymbol{X}_1)^{-1} \boldsymbol{X}'_1 \boldsymbol{V}^{-1}.$$

On account of the form of $\boldsymbol{V}$, the formulae $(6.15)-(6.17)$ and the properties of the matrices $\boldsymbol{H}_\beta = \boldsymbol{S}_\beta \boldsymbol{S}'_\beta$, $\beta = 0, 1, ..., m, m+1$, appearing there, it is justified to write

$$(\boldsymbol{X}'_1 \boldsymbol{V}^{-1} \boldsymbol{X}_1)^{-1} = \sigma_1^2 \boldsymbol{H}_0$$
$$+ \sum_{\beta=1}^{m-1} \frac{\sigma_1^2 \sigma_2^2}{\varepsilon_{1\beta} \sigma_2^2 + \varepsilon_{2\beta} \sigma_1^2} \boldsymbol{H}_\beta + \sigma_2^2 \boldsymbol{H}_m + \sigma_3^2 \boldsymbol{H}_{m+1}. (6.19)$$

Hence for any proper block design, the projector $\boldsymbol{P}_{X_1(V^{-1})}$ can be written in decomposed form as

$$\boldsymbol{P}_{X_1(V^{-1})} = \sum_{\beta=0}^{m-1} w_{1\beta} \varepsilon_{1\beta}^{-1} \boldsymbol{P}_{X_1 S_\beta} \boldsymbol{\phi}_1 + \sum_{\beta=1}^{m} w_{2\beta} \varepsilon_{2\beta}^{-1} \boldsymbol{P}_{X_1 S_\beta} \boldsymbol{\phi}_2 + \boldsymbol{\phi}_3, (6.20)$$

where the weights are defined in accordance with (6.10), i.e., as

$$w_{1\beta} = \frac{\varepsilon_{1\beta} \sigma_2^2}{\varepsilon_{1\beta} \sigma_2^2 + \varepsilon_{2\beta} \sigma_1^2} \quad \text{and} \quad w_{2\beta} = \frac{\varepsilon_{2\beta} \sigma_1^2}{\varepsilon_{1\beta} \sigma_2^2 + \varepsilon_{2\beta} \sigma_1^2}, \tag{6.21}$$

with $\varepsilon_{1\beta} = \varepsilon_\beta$ and $\varepsilon_{2\beta} = 1 - \varepsilon_\beta$, $\varepsilon_\beta$ being the eigenvalue of $\boldsymbol{C}_1$ with respect to $\boldsymbol{r}^\delta$ corresponding to all $\boldsymbol{s}_{\beta j}$, $j = 1, 2, ..., \rho_\beta$, and where $\boldsymbol{P}_{X_1 S_\beta} = \boldsymbol{X}_1 \boldsymbol{S}_\beta \boldsymbol{S}'_\beta \boldsymbol{X}'_1$ is the orthogonal projector on the subspace $\mathcal{R}(\boldsymbol{X}_1 \boldsymbol{S}_\beta)$, for $\beta = 0, 1, ..., m$ (where the matrices $\boldsymbol{S}_0, \boldsymbol{S}_1, ..., \boldsymbol{S}_m$ represent basic contrasts of the design, according to Definition 6.1 and the comments following it).

Finally, note that the vector $\boldsymbol{y}$ can be decomposed as

$$\boldsymbol{y} = \boldsymbol{P}_{X_1(V^{-1})} \boldsymbol{y} + (\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})}) \boldsymbol{y},$$

where the second term is the residual vector giving the residual sum of squares of the form

$$\|(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})}) \boldsymbol{y}\|_{V^{-1}}^2 = \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})' \boldsymbol{V}^{-1} (\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})}) \boldsymbol{y}$$
$$= \sigma_1^{-2} \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})' \boldsymbol{\phi}_1 (\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})}) \boldsymbol{y}$$
$$+ \sigma_2^{-2} \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})' \boldsymbol{\phi}_2 (\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})}) \boldsymbol{y},$$
$$(6.22)$$

as follows from the form of $\boldsymbol{V} = \mathrm{D}(\boldsymbol{y})$, which also implies that

$$\boldsymbol{\phi}_3 (\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})}) = \sigma_3^2 \boldsymbol{\phi}_3 \boldsymbol{V}^{-1} (\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})}) = \boldsymbol{O},$$

on account of (5.7), the equality $\boldsymbol{1}_n = \boldsymbol{X}_1 \boldsymbol{1}_v$ and the form of $\boldsymbol{P}_{X_1(V^{-1})}$. The residual degrees of freedom corresponding to (6.22) are given by the

rank difference $d = \mathrm{rank}(\boldsymbol{V} : \boldsymbol{X}_1) - \mathrm{rank}(\boldsymbol{X}_1) = n - v$ (see Rao, 1974, Theorem 3.4).

## 6.2. Estimation of unknown stratum variances

The results of Section 6.1 have been obtained under the assumptions of Theorem 6.1, which in particular imply that $\sigma_1^2$ and $\sigma_2^2$ are known. In real applications, however, these quantities are usually not known and need to be estimated. So the problem is how to estimate $\sigma_1^2$ and $\sigma_2^2$ so that the formulae appearing in Theorem 6.1 and Corollary 6.1, i.e., those of the type (6.1) and all others involving these stratum variances, could be applied in the analysis of real experimental data.

To solve the problem of estimating $\sigma_1^2$ and $\sigma_2^2$, first note that, from (6.22),

$$\mathrm{E}\{\|(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|_{V^{-1}}^2\}$$

$$= \sigma_1^{-2}\mathrm{E}\{\|\boldsymbol{\phi}_1(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2\} + \sigma_2^{-2}\mathrm{E}\{\|\boldsymbol{\phi}_2(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2\}$$

$$= d_1 + \sum_{\beta=0}^{m-1}(1 - w_{1\beta})\rho_\beta + d_2 + \sum_{\beta=1}^{m}(1 - w_{2\beta})\rho_\beta$$

$$= d_1 + d_2 + h - \rho_0 = n - v,$$

because, as can be shown,

$$\mathrm{E}\{\|\boldsymbol{\phi}_1(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2\} = \sigma_1^2 d_1',$$

where

$$
\begin{aligned}
d_1' &= \mathrm{tr}[\boldsymbol{\phi}_1(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})] \\
&= n - b - \sum_{\beta=0}^{m-1} w_{1\beta}\rho_\beta = d_1 + \sum_{\beta=0}^{m-1}(1 - w_{1\beta})\rho_\beta,
\end{aligned}
\tag{6.23}
$$

with $d_1 = n - b - h$, the residual d.f. in the intra-block analysis of variance (see Section 5.1), and

$$\mathrm{E}\{\|\boldsymbol{\phi}_2(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2\} = \sigma_2^2 d_2',$$

where

$$
\begin{aligned}
d_2' &= \operatorname{tr}[\phi_2(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})] \\
&= b - 1 - \sum_{\beta=1}^{m} w_{2\beta}\rho_\beta = d_2 + \sum_{\beta=1}^{m}(1 - w_{2\beta})\rho_\beta,
\end{aligned}
\tag{6.24}
$$

with $d_2 = b - v + \rho_0$, the residual d.f. in the inter-block analysis of variance (see Section 5.2), the results in (6.23) and (6.24) following from (6.20).

With these results, it is natural to consider as estimators of $\sigma_1^2$ and $\sigma_2^2$ the solutions of equations

$$
\|\phi_1(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2 = \sigma_1^2 d_1',
\tag{6.25}
$$

$$
\|\phi_2(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2 = \sigma_2^2 d_2',
\tag{6.26}
$$

respectively (as suggested by Nelder, 1968). However, it should be noted that the explicit forms of the sums of squares in (6.25) and (6.26) are

$$
\|\phi_1(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2 = \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})'\phi_1(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}
$$

$$
= \|\psi_1\boldsymbol{y}\|^2 + \sum_{\beta=1}^{m-1} w_{2\beta}^2 \varepsilon_{1\beta}\|\varepsilon_{1\beta}^{-1}\boldsymbol{P}_{X_1 S_\beta}\boldsymbol{y}_1 - \varepsilon_{2\beta}^{-1}\boldsymbol{P}_{X_1 S_\beta}\boldsymbol{y}_2\|^2
\tag{6.27}
$$

and

$$
\|\phi_2(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2 = \boldsymbol{y}'(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})'\phi_2(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}
$$

$$
= \|\psi_2\boldsymbol{y}\|^2 + \sum_{\beta=1}^{m-1} w_{1\beta}^2 \varepsilon_{2\beta}\|\varepsilon_{2\beta}^{-1}\boldsymbol{P}_{X_1 S_\beta}\boldsymbol{y}_2 - \varepsilon_{1\beta}^{-1}\boldsymbol{P}_{X_1 S_\beta}\boldsymbol{y}_1\|^2,
\tag{6.28}
$$

respectively, as follows from (6.20) and from the definitions $\psi_\alpha = \phi_\alpha - \boldsymbol{P}_{\phi_\alpha X_1}$ and $\boldsymbol{y}_\alpha = \phi_\alpha \boldsymbol{y}$, for $\alpha = 1, 2$, used in Section 5 (see also Caliński, 1996, Section 3.3).

Evidently, unlike the original intra-block and inter-block residual sums of squares, $\|\psi_1\boldsymbol{y}\|^2$ and $\|\psi_2\boldsymbol{y}\|^2$, introduced in Sections 5.1 and 5.2 respectively, their augmented forms (6.27) and (6.28) depend on the weights $\{w_{2\beta}\}$ or $\{w_{1\beta}\}$. The weights in turn depend on the stratum variances, $\sigma_1^2$ and $\sigma_2^2$, as shown in (6.21). Thus the solutions of (6.25) and (6.26) for $\sigma_1^2$ and $\sigma_2^2$ are not directly obtainable, unless the weights are known, which is usually not the case. Therefore, it has been suggested by Nelder (1968, Section 3.4) to use an iterative procedure, with a set of initial values $\{w_{\alpha\beta}^{(0)}\}$ for the

unknown values of $\{w_{\alpha\beta}\}$, in (6.27) and (6.28), from which the "working" estimates of $\sigma_1^2$ and $\sigma_2^2$ can be obtained by solving (6.25) and (6.26) respectively, giving then a revised set of the values of $\{w_{\alpha\beta}\}$ according to (6.21), and so on, until the process converges.

Alternatively, one can start the iterative procedure with a pair of initial values $\sigma_1^{2(0)}$ and $\sigma_2^{2(0)}$ for the unknown stratum variances, from which the "working" estimates of the weights $\{w_{\alpha\beta}\}$ can be obtained by (6.21), giving then a revised pair of values of $\sigma_1^2$ and $\sigma_2^2$ from the sums of squares (6.27) and (6.28) divided by $d_1'$ and $d_2'$, respectively, and so on (according to Fisher's method of scoring).

The above approach, adopted from Nelder (1968), which can be considered as a generalized alternative to that introduced by Yates (1939, 1940), is also advocated by Houtman and Speed (1983, Sections 4.5 and 5.1). For statistical consequences of this approach see e.g. Caliński and Kageyama (2000, Section 3.8).

## 6.3. Relation to the Gauss−Markov model

The model considered here, of the form

$$\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\tau} + \boldsymbol{X}_2\boldsymbol{\beta} + \boldsymbol{\eta} + \boldsymbol{e}, \tag{6.29}$$

with

$$\mathrm{E}(\boldsymbol{y}) = \boldsymbol{X}_1\boldsymbol{\tau} \quad \text{and} \quad \mathrm{D}(\boldsymbol{y}) = \boldsymbol{\phi}_1\sigma_1^2 + \boldsymbol{\phi}_2\sigma_2^2 + \boldsymbol{\phi}_3\sigma_3^2 \tag{6.30}$$

(for any proper block design), can be regarded as a general Gauss−Markov model, provided the stratum variances $\sigma_1^2$, $\sigma_2^2$ ans $\sigma_3^2$ are known. The question arises whether it can still be treated as a Gauss−Markov model when the true values of these variance components are replaced by their estimates.

This question can be answered by referring to Rao and Kleffe (1988, Section 3.1), where a model of this type is considered as an extension of the Gauss−Markov model. More precisely, it is called a mixed model, with fixed (here $\boldsymbol{\tau}$) and random (here $\boldsymbol{\beta}$, $\boldsymbol{\eta}$ and $\boldsymbol{e}$) effects. For more on the statistical properties of this extended model see, e.g., Demidenko (2004).

Thus when extending the analysis of an experiment in a block design (proper in particular) from the intra-block analysis to the more general analysis that combines information from the intra- and inter-block strata, one somehow proceeds from the Gauss−Markov model theory to a more extended mixed model theory. Note that Oktaba, in his paper devoted to

the Gauss−Markov model applicable to incomplete block designs, confines attention to the intra-block analysis (see Oktaba, 2003, p. 203).

## 7. Concluding remarks

• The theory and practice of the analysis of experimental data is based on the general Gauss−Markov model, described by $\{\boldsymbol{y}, \boldsymbol{X\beta}, \sigma^2\boldsymbol{V}\}$. The main estimation and hypothesis testing results originating from the theory of this model are well known.

• Application of this model to the classic randomized block design causes no difficulty, whether the block effects are considered as fixed or random.

• Of particular interest is the randomization-derived model for a general block design. In this model the block effects are to be considered as random. It appears that if uniform BLUEs of linear parametric functions are required, the model is very restrictive.

• This difficulty is usually avoided by resolving the model into three sub-models (two for treatment contrasts). Particularly interesting is the intra-block submodel. The analysis can well be restricted to this submodel in the case of (i) choice of a design with high intra-block efficiency factors for the contrasts of main interest and (ii) successful blocking of the available experimental units (plots) with the aim of removing their intra-block variation.

• In general, combining information from the intra-block and inter-block strata may be desirable. But then, if the stratum variances are not known, the model is extended from the general Gauss−Markov model to a mixed model, with serious consequences for the analysis.

### References

Bailey R.A. (1981): A unified approach to design of experiments. Journal of the Royal Statistical Society, Series A 144: 214–223.

Bailey R.A. (1991): Strata for randomized experiments. Journal of the Royal Statistical Society, Series B 53: 27–78.

Baksalary J.K., Kala R. (1983): On equalities between BLUEs, WLSEs, and SLSEs. The Canadian Journal of Statistics 11: 119–123.

Baksalary J.K., Puntanen S. (1990): Characterizations of the best linear unbiased estimator in the general Gauss−Markov model with the use of matrix partial orderings. Linear Algebra and its Applications 127: 363–370.

Caliński T. (1996): The basic contrasts of a block design with special reference to the recovery of inter-block information. In: A. Pázman and V. Witkovský (eds.), Tatra Mountains Mathematical Publications, Vol. 7: PROBA-STAT'94 Smolenice. Mathematical Institute, Bratislava, pp. 23–37.

Caliński T., Kageyama S. (2000): Block Designs: A Randomization Approach, Volume I: Analysis. Lecture Notes in Statistics, Volume 150, Springer, New York.

Demidenko E. (2004): Mixed Models: Theory and Applications. Wiley, Hoboken, New Jersey.

Fisher R.A. (1925): Statistical Methods for Research Workers. Oliver Boyd, Edinburgh.

Gauss C.F. (1809): Theoria Motus Corporum Coelestium. Perthes and Besser, Hamburg.

Gauss C.F. (1855): Méthode des Moindres Carrés. Mallet-Bachelier, Paris.

Houtman A.M., Speed, T.P. (2008): Balance in designed experiments with orthogonal block structure Annals of Statistics 11: 1069–1085.

Hinkelmann K., Kempthorne O. (2008): Design and Analysis of Experiments Volume I: Introduction to Experimental Design 2nd ed. Wiley, Hoboken, New Jersey.

John J.A. (1987): Cyclic Designs. Chapman and Hall, London.

Jones R.M. (1959): On a property of incomplete blocks. Journal of the Royal Statistical Society, Series B 21: 172–179.

Kala R. (1991): Elements of the randomization theory. III. Randomization in block experiments. Listy Biometryczne—Biometrical Letters 28: 3–23 (in Polish).

Kempthorne O. (1952): The Design and Analysis of Experiments. Wiley, New York.

Markoff A.A. (1900): Wahrscheinlichkeitsrechnung. Telner, Leipzig.

Martin F.B., Zyskind G. (1966): On combinability of information from uncorrelated linear models by simple weighting. Annals of Mathematical Statistics 37: 1338–1347.

Nelder J.A. (1954): The interpretation of negative components of variance. Biometrika 41: 544–548.

Nelder J.A. (1965): The analysis of randomized experiments with orthogonal block structure. Proceedings of the Royal Society, Series A 283: 147–178.

Nelder J.A. (1968): The combination of information in generally balanced designs. Journal of the Royal Statistical Society, Series B 30: 303–311.

Neyman J. (with cooperation of K. Iwaszkiewicz and S. Kołodziejczyk) (1935): Statistical problems in agricultural experimentation (with discussion). Journal of the Royal Statistical Society, Supplement 2: 107–180.

Ogawa J. (1961): The effect of randomization on the analysis of randomized block design. Annals of the Institute of Statistical Mathematics 13: 105–117.

Ogawa J. (1963): On the null-distribution of the $F$-statistic in a randomized balanced incomplete block design under the Neyman model. Annals of Mathematical Statistics 34: 1558–1568.

Oktaba W. (1984): Tests of hypotheses for the general Gauss−Markov model. Biometrical Journal 26: 415–424.

Oktaba W. (1989): F-tests for hypotheses with block matrices and under conditions of orthogonality in the general multivariate Gauss−Markov model. Biometrical Journal 31: 317–323.

Oktaba W. (1996): Asymptotically normal distributions in the multivariate Gauss–Markov model. Listy Biometryczne—Biometrical letters 33: 25–31.

Oktaba W. (1998): Characterization of the multivariate Gauss−Markov model with singular covariance matrix. Applications of Mathematics 43: 119–131.

Oktaba W. (2003): The general multivariate Gauss−Markov model of the incomplete block design. Australian & New Zealand Journal of Statistics 45: 195–205.

Oktaba W., Kornacki A., Wawrzosek J. (1986): Estimation of missing values in the general Gauss−Markov model. Statistics 17: 167–177.

Oktaba W., Kornacki A., Wawrzosek J. (1988): Invariant linearly sufficient transformations of the general Gauss−Markoff model. Estimation and testing. Scandinavian Journal of Statistics 17: 117–124.

Patterson H.D., Thompson R. (1971): Recovery of inter-block information when block sizes are unequal. Biometrika 58: 545–554.

Pearce S.C. (1983): The Agricultural Field Experiment: A Statistical Examination of Theory and Practice. Wiley, Chichester.

Pearce S.C., Caliński T., Marshall T.F. de C. (1974): The basic contrasts of an experimental design with special reference to the analysis of data. Biometrika 61: 449–460.

Raghavarao D., Padgett L.V. (2005): Block Designs: Analysis, Combinatorics and Applications. World Scientific Publishing Co., Singapore.

Rao C.R. (1959): Expected values of mean squares in the analysis of incomplete block experiments and some comments based on them. Sankhyã 21: 327–336.

Rao C.R. (1971): Unified theory of linear estimation. Sankhyã, Series A 33: 371–394.

Rao C.R. (1973): Linear Statistical Inference and Its Applications 2nd ed. Wiley, New York.

Rao C.R. (1974): Projectors, generalized inverses and the BLUEs. Journal of the Royal Statistical Society, Series B 36: 442–448.

Rao C.R., Kleffe J. (1988): Estimation of Variance Components and Applications. North-Holland, Amsterdam.

Rao C.R., Mitra S. K. (1971): Generalized Inverse of Matrices and Its Applications. Wiley, New York.

Scheffé H. (1959):  The Analysis of Variance. Wiley, New York.

Seber G.A.F. (1980):  The Linear Hypothesis: A General Theory. Charles Griffin, London.

Shah K.R. (1992): Recovery of interblock information: An update.  Journal of Statistical Planning and Inference 30: 163–172.

White R.F. (1975): Randomization in the analysis of variance.  Biometrics 31: 555–571.

Yates F. (1939): The recovery of inter-block information in variety trials arranged in three-dimensional lattices.  Annals of Eugenics 9: 136–156.

Yates F. (1940): The recovery of inter-block information in balanced incomplete block designs.  Annals of Eugenics 10: 317–325.

Zyskind G. (1967): On canonical forms, non-negative covariance matrices and best and simple least squares linear estimators in linear models.  Annals of Mathematical Statistics 38: 1092–1109.