Nested row-column design – direct approach to ANOVA

Tadeusz Caliński, Agnieszka Łacka, Idzi Siatkowski

Department of Mathematical and Statistical Methods
Poznań University of Life Sciences

XII Working Seminar on Statistical Methods in Variety Testing
04 – 08.07.2022

Introduction

Definition 1. (from Section 2.2 in Houtman and Speed, 1983).

An experiment is said to have the **orthogonal block structure** (OBS) if the covariance (dispersion) matrix of the random variables observed on the experimental units (plots), $\boldsymbol{y} = [y_1, y_2, ..., y_n]'$, has a representation of the form

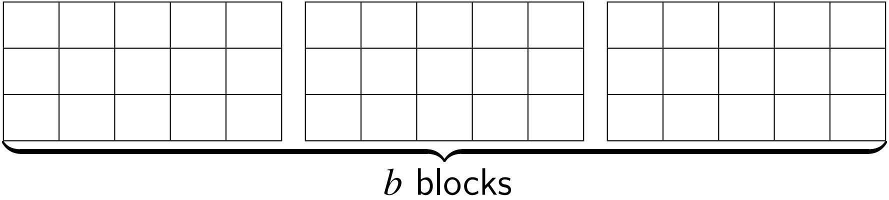
$$\mathrm{D}(\boldsymbol{y}) = \sigma_1^2 \phi_1 + \sigma_2^2 \phi_2 + \cdots + \sigma_t^2 \phi_t,$$

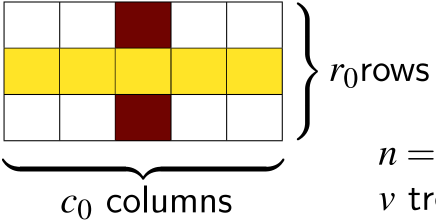
where the $\{\phi_\alpha\}$, $\alpha = 1, 2, ..., t$, are known symmetric, idempotent and pairwise orthogonal matrices, summing to the identity matrix, the last usually being of the form $\phi_t = n^{-1} \mathbf{1}_n \mathbf{1}_n'$.

Introduction

1. Caliński, T.; Siatkowski, I. On a new approach to the analysis of variance for experiments with orthogonal block structure. I. Experiments in proper block designs. *Biometrical Letters* **2017**, *54*, 91–122.
2. Caliński, T.; Siatkowski, I. On a new approach to the analysis of variance for experiments with orthogonal block structure. II. Experiments in nested block designs. *Biometrical Letters* **2018**, *55*, 147–178.
3. Caliński, T.; Łacka, A.; Siatkowski, I. On a new approach to the analysis of variance for experiments with orthogonal block structure. III. Experiments in row-column designs. *Biometrical Letters* **2019**, *56*, 183–213.
4. Caliński, T.; Łacka, A.; Siatkowski, I. On a new approach to the analysis of variance for experiments with orthogonal block structure. IV. Experiments in split-plot designs. *Biometrical Letters* **2020**, *57*, 183–213.
5. Łacka, A. NRC Designs – New Tools for Successful Agricultural Experiments. *Agronomy* **2021**, *11(12)*, 2406.

NRC design





$n = br_0c_0$

v treatments

A randomization-derived model (NRC design)

$$\boldsymbol{y} = \boldsymbol{X}_1\tau + \boldsymbol{X}_B\beta + \boldsymbol{X}_{R(B)}\rho + \boldsymbol{X}_{C(B)}\gamma + \eta + \boldsymbol{e}, \tag{1}$$

$\boldsymbol{y} = [\boldsymbol{y}_1', \boldsymbol{y}_2', ..., \boldsymbol{y}_b']'$ – $n \times 1$ vector of data observed on $n = br_0c_0$ plots,
$\tau = [\tau_1, \tau_2, ..., \tau_v]'$ – unobservable treatment parameters (their fixed effects),
β – block random effects,
ρ – row random effects,
γ – column random effects,
η and $\boldsymbol{e}$ – unit and technical random errors ($n \times 1$).

$$\boldsymbol{X}_1 = [\boldsymbol{X}_{11}' : \boldsymbol{X}_{12}' : \cdots : \boldsymbol{X}_{1b}']', \quad \boldsymbol{X}_B = \boldsymbol{I}_b \otimes \boldsymbol{1}_{n_0}, \quad \boldsymbol{X}_{R(B)} = \boldsymbol{I}_b \otimes \boldsymbol{I}_{r_0} \otimes \boldsymbol{1}_{c_0},$$

$$\boldsymbol{X}_{C(B)} = \boldsymbol{I}_b \otimes \boldsymbol{1}_{r_0} \otimes \boldsymbol{I}_{c_0}$$

$\boldsymbol{X}_1'\boldsymbol{1}_n = \boldsymbol{r} = [r_1, ..., r_v]'$ – vector of treatment replications,
$\boldsymbol{X}_1'\boldsymbol{X}_1 = \boldsymbol{r}^\delta = \text{diag}[r_1, r_2, ..., r_v]$,
$\boldsymbol{r}^{-\delta}$

A randomization-derived model (NRC design)

Units (plots) $\rightarrow$ (Rows $\times$ Columns) $\rightarrow$ Blocks $\rightarrow$ Total exp. area

Thus, the observed vector $\boldsymbol{y}$ can be decomposed as

$\boldsymbol{y} = \boldsymbol{y}_1 + \boldsymbol{y}_2 + \boldsymbol{y}_3 + \boldsymbol{y}_4 + \boldsymbol{y}_5$, where $\boldsymbol{y}_1 = \boldsymbol{\phi}_1 \boldsymbol{y}$, $\boldsymbol{y}_2 = \boldsymbol{\phi}_2 \boldsymbol{y}$, $\boldsymbol{y}_3 = \boldsymbol{\phi}_3 \boldsymbol{y}$, $\boldsymbol{y}_4 = \boldsymbol{\phi}_4 \boldsymbol{y}$, $\boldsymbol{y}_5 = \boldsymbol{\phi}_5 \boldsymbol{y}$.

This allows the expectation vector and the covariance (dispersion) matrix of $\boldsymbol{y}$ to be written as

$$\begin{aligned} \mathsf{E}(\boldsymbol{y}) &= \boldsymbol{\phi}_1 \boldsymbol{X}_1 \boldsymbol{\tau} + \boldsymbol{\phi}_2 \boldsymbol{X}_1 \boldsymbol{\tau} + \boldsymbol{\phi}_3 \boldsymbol{X}_1 \boldsymbol{\tau} + \boldsymbol{\phi}_4 \boldsymbol{X}_1 \boldsymbol{\tau} + \boldsymbol{\phi}_5 \boldsymbol{X}_1 \boldsymbol{\tau} = \boldsymbol{X}_1 \boldsymbol{\tau}, \\ \mathsf{D}(\boldsymbol{y}) \equiv \boldsymbol{V} &= \sigma_1^2 \boldsymbol{\phi}_1 + \sigma_2^2 \boldsymbol{\phi}_2 + \sigma_3^2 \boldsymbol{\phi}_3 + \sigma_4^2 \boldsymbol{\phi}_4 + \sigma_5^2 \boldsymbol{\phi}_5 \end{aligned}$$

where the matrices

$$\boldsymbol{\phi}_1 = \boldsymbol{I}_n - c_0^{-1} \boldsymbol{X}_{R(B)} \boldsymbol{X}'_{R(B)} - r_0^{-1} \boldsymbol{X}_{C(B)} \boldsymbol{X}'_{C(B)} + n_0^{-1} \boldsymbol{X}_B \boldsymbol{X}'_B,$$
$$\boldsymbol{\phi}_2 = c_0^{-1} \boldsymbol{X}_{R(B)} \boldsymbol{X}'_{R(B)} - n_0^{-1} \boldsymbol{X}_B \boldsymbol{X}'_B, \ \boldsymbol{\phi}_3 = r_0^{-1} \boldsymbol{X}_{C(B)} \boldsymbol{X}'_{C(B)} - n_0^{-1} \boldsymbol{X}_B \boldsymbol{X}'_B,$$
$$\boldsymbol{\phi}_4 = n_0^{-1} \boldsymbol{X}_B \boldsymbol{X}'_B - n^{-1} \mathbf{1}_n \mathbf{1}'_n \text{ and } \boldsymbol{\phi}_5 = n^{-1} \mathbf{1}_n \mathbf{1}'_n$$

are symmetric, idempotent and pairwise orthogonal, summing to the identity matrix, and the scalars $\sigma_1^2, \sigma_2^2, \sigma_3^2$, σ_4^2 and σ_5^2 represent the relevant unknown stratum variances.

Theoretical background of the analysis

$\tau = [\tau_1, \tau_2, ..., \tau_v]'$,

$(\boldsymbol{I}_v - n^{-1}\mathbf{1}_v\boldsymbol{r}')\tau = [\tau_1 - \tau_., \tau_2 - \tau_., ..., \tau_v - \tau_.]'$, where $\tau_. = n^{-1}\sum_{i=1}^{v}(r_i\tau_i)$,

$(\boldsymbol{V}^{-1}$-orthogonal) projector

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

We can decompose the analyzed data vector $\boldsymbol{y}$ into two uncorrelated parts, as

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

Theoretical background of the analysis

$$y = P_{X_1(V^{-1})} y + (I_n - P_{X_1(V^{-1})}) y$$

Under the model (1), with presented properties, the first term of the partition provides the best linear unbiased estimator (BLUE) of $X_1\tau$, which can be expressed as

$$\widehat{X_1\tau} = P_{X_1(V^{-1})} y.$$

With regard to the second term, it can be seen as the residual vector, giving the residual sum of squares in the form

$$\begin{aligned} \|(I_n - P_{X_1(V^{-1})})y\|^2_{V^{-1}} &= y'(I_n - P_{X_1(V^{-1})})' V^{-1} (I_n - P_{X_1(V^{-1})}) y \\ &= y'[V^{-1} - V^{-1} X_1 (X_1' V^{-1} X_1)^{-1} X_1' V^{-1}] y \\ &= y' V^{-1} (I_n - P_{X_1(V^{-1})}) y, \end{aligned}$$

with the residual degrees of freedom given by $\mathrm{rank}(V : X_1) - \mathrm{rank}(X_1) = n - v$.

Theoretical background of the analysis

Also note that, as $\tau = \boldsymbol{r}^{-\delta} \boldsymbol{X}_1' \boldsymbol{X}_1 \tau$, the BLUE of τ can be obtained as

$$\hat{\tau} = (\boldsymbol{X}_1' \boldsymbol{V}^{-1} \boldsymbol{X}_1)^{-1} \boldsymbol{X}_1' \boldsymbol{V}^{-1} \boldsymbol{y}.$$

Its covariance (dispersion) matrix then takes the form

$$\begin{aligned} \mathsf{D}(\hat{\tau}) &= (\boldsymbol{X}_1' \boldsymbol{V}^{-1} \boldsymbol{X}_1)^{-1} \boldsymbol{X}_1' \boldsymbol{V}^{-1} \mathsf{D}(\boldsymbol{y}) \boldsymbol{V}^{-1} \boldsymbol{X}_1 (\boldsymbol{X}_1' \boldsymbol{V}^{-1} \boldsymbol{X}_1)^{-1} \\ &= (\boldsymbol{X}_1' \boldsymbol{V}^{-1} \boldsymbol{X}_1)^{-1}. \end{aligned}$$

Theoretical background of the analysis

$$\boldsymbol{\tau}_* = (\boldsymbol{I}_v - n^{-1}\mathbf{1}_v\boldsymbol{r}')\boldsymbol{\tau}$$

$$H_0 : \boldsymbol{\tau}_* = \mathbf{0} \tag{2}$$

Theoretical background of the analysis

Assumpion: $\boldsymbol{y} \sim N_n(\boldsymbol{X}_1\boldsymbol{\tau}, \boldsymbol{V})$ and, hence $\hat{\boldsymbol{\tau}}_* \sim N_v[\boldsymbol{\tau}_*, \mathrm{D}(\hat{\boldsymbol{\tau}}_*)]$

$$F = \frac{n-v}{v-1}\,\frac{\mathrm{SS}_V}{\mathrm{SS}_R} = \frac{n-v}{v-1}\,\frac{\hat{\boldsymbol{\tau}}_*'\boldsymbol{X}_1'\boldsymbol{V}^{-1}\boldsymbol{X}_1\hat{\boldsymbol{\tau}}_*}{\boldsymbol{y}'\boldsymbol{V}^{-1}(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}} \qquad (3)$$

$$\mathrm{SS}_V = \boldsymbol{y}'\boldsymbol{V}^{-1}\boldsymbol{X}_1(\boldsymbol{I}_v - n^{-1}\mathbf{1}_v\boldsymbol{r}')(\boldsymbol{X}_1'\boldsymbol{V}^{-1}\boldsymbol{X}_1)^{-1}(\boldsymbol{I}_v - n^{-1}\boldsymbol{r}\mathbf{1}_v')\boldsymbol{X}_1'\boldsymbol{V}^{-1}\boldsymbol{y} \qquad (4)$$

$$\mathrm{SS}_R = \boldsymbol{y}'[\boldsymbol{V}^{-1} - \boldsymbol{V}^{-1}\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{V}^{-1}\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{V}^{-1}]\boldsymbol{y} \qquad (5)$$

Referring now to Theorems 9.2.1 and 9.4.1 in Rao and Mitra (1971), one can show that, independently,

$$\mathrm{SS}_V \sim \chi^2(v-1,\delta), \quad \text{with} \quad \delta = \boldsymbol{\tau}_*'\boldsymbol{X}_1'\boldsymbol{V}^{-1}\boldsymbol{X}_1\boldsymbol{\tau}_*, \qquad (6)$$

$$\mathrm{SS}_R \sim \chi^2(n-v,0). \qquad (7)$$

Evidently, the distribution in (6) is central if H_0 is true, whereas that in (7) is central whether H_0 is true or not. These results imply that the statistic (3) has a noncentral F distribution with $v-1$ and $n-v$ d.f., and with the noncentrality parameter δ as in (6). Thus, the distribution is central if H_0 is true.

But what to do if stratum variances σ_1^2, σ_2^2, σ_3^2, σ_4^2 and σ_5^2 are unknown?

$$\begin{aligned}\|(\boldsymbol{I}_n-\boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2_{V^{-1}} &= \boldsymbol{y}'(\boldsymbol{I}_n-\boldsymbol{P}_{X_1(V^{-1})})'\boldsymbol{V}^{-1}(\boldsymbol{I}_n-\boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y} \\ &= \sigma_1^{-2}\boldsymbol{y}'(\boldsymbol{I}_n-\boldsymbol{P}_{X_1(V^{-1})})'\phi_1(\boldsymbol{I}_n-\boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y} \\ &\quad+\sigma_2^{-2}\boldsymbol{y}'(\boldsymbol{I}_n-\boldsymbol{P}_{X_1(V^{-1})})'\phi_2(\boldsymbol{I}_n-\boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y} \\ &\quad+\sigma_3^{-2}\boldsymbol{y}'(\boldsymbol{I}_n-\boldsymbol{P}_{X_1(V^{-1})})'\phi_3(\boldsymbol{I}_n-\boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y} \\ &\quad+\sigma_4^{-2}\boldsymbol{y}'(\boldsymbol{I}_n-\boldsymbol{P}_{X_1(V^{-1})})'\phi_4(\boldsymbol{I}_n-\boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y} \end{aligned} \tag{8}$$

$\mathrm{D}(\boldsymbol{y})\equiv\boldsymbol{V}=\sigma_1^2\phi_1+\sigma_2^2\phi_2+\sigma_3^2\phi_3+\sigma_4^2\phi_4+\sigma_5^2\phi_5$
$\phi_5=n^{-1}\mathbf{1}_n\mathbf{1}_n'=n^{-1}\mathbf{1}_n\mathbf{1}_v'\boldsymbol{X}_1'$

$$\phi_5(\boldsymbol{I}_n-\boldsymbol{P}_{X_1(V^{-1})})=\sigma_5^2\phi_5\boldsymbol{V}^{-1}(\boldsymbol{I}_n-\boldsymbol{P}_{X_1(V^{-1})})=\mathbf{O}$$

Theoretical background of the analysis (NRC design)

$$
\begin{aligned}
\mathsf{E}\{\|(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2_{V^{-1}}\} &= \sigma_1^{-2}\mathsf{E}\{\|\phi_1(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2\} \\
&+ \sigma_2^{-2}\mathsf{E}\{\|\phi_2(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2\} \\
&+ \sigma_3^{-2}\mathsf{E}\{\|\phi_3(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2\} \\
&+ \sigma_4^{-2}\mathsf{E}\{\|\phi_4(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2\} \\
&= d_1' + d_2' + d_3' + d_4' = n - v \qquad (9)
\end{aligned}
$$

$$\mathsf{E}\{\|\phi_1(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2\} = \sigma_1^2 d_1', \quad \text{where} \quad d_1' = \mathrm{tr}[\phi_1(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})] \qquad (10)$$

$$\mathsf{E}\{\|\phi_2(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2\} = \sigma_2^2 d_2', \quad \text{where} \quad d_2' = \mathrm{tr}[\phi_2(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})] \qquad (11)$$

$$\mathsf{E}\{\|\phi_3(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2\} = \sigma_3^2 d_3', \quad \text{where} \quad d_3' = \mathrm{tr}[\phi_3(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})] \qquad (12)$$

$$\mathsf{E}\{\|\phi_4(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2\} = \sigma_4^2 d_4', \quad \text{where} \quad d_4' = \mathrm{tr}[\phi_4(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})] \qquad (13)$$

$$\|\phi_1(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2 = \sigma_1^2 d_1' \qquad \|\phi_2(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2 = \sigma_2^2 d_2' \qquad (14)$$

$$\|\phi_3(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2 = \sigma_3^2 d_3' \qquad \|\phi_4(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})})\boldsymbol{y}\|^2 = \sigma_4^2 d_4' \qquad (15)$$

Theoretical background of the analysis (NRC design)

$$\hat{\sigma}_1^{-2}\|\phi_1(\boldsymbol{I}_n-\boldsymbol{P}_{X(V^{-1})})\boldsymbol{y}\|^2 \;+\; \hat{\sigma}_2^{-2}\|\phi_2(\boldsymbol{I}_n-\boldsymbol{P}_{X(V^{-1})})\boldsymbol{y}\|^2$$
$$\hat{\sigma}_3^{-2}\|\phi_3(\boldsymbol{I}_n-\boldsymbol{P}_{X(V^{-1})})\boldsymbol{y}\|^2 \;+\; \hat{\sigma}_4^{-2}\|\phi_4(\boldsymbol{I}_n-\boldsymbol{P}_{X(V^{-1})})\boldsymbol{y}\|^2$$
$$= d_1' + d_2' + d_3' + d_4' = n - v \tag{16}$$

Now, returning to SS_R given in (5) note that, on account of the relations $\boldsymbol{V}^{-1} = \sigma_1^{-2}\phi_1 + \sigma_2^{-2}\phi_2 + \sigma_3^{-2}\phi_3 + \sigma_4^{-2}\phi_4 + \sigma_5^{-2}\phi_5$ and $\phi_5(\boldsymbol{I}_n - \boldsymbol{P}_{X_1(V^{-1})}) = \mathbf{O}$, it can be written in the form

$$\begin{aligned}\mathrm{SS}_R \;=\; & \sigma_1^{-2}\|\phi_1(\boldsymbol{I}_n-\boldsymbol{P}_{X(V^{-1})})\boldsymbol{y}\|^2 + \sigma_2^{-2}\|\phi_2(\boldsymbol{I}_n-\boldsymbol{P}_{X(V^{-1})})\boldsymbol{y}\|^2 \\ & + \sigma_3^{-2}\|\phi_3(\boldsymbol{I}_n-\boldsymbol{P}_{X(V^{-1})})\boldsymbol{y}\|^2 + \sigma_4^{-2}\|\phi_4(\boldsymbol{I}_n-\boldsymbol{P}_{X(V^{-1})})\boldsymbol{y}\|^2.\end{aligned} \tag{17}$$

A comparison of formulae (16) and (17) shows that, if the stratum variances are estimated by solutions of the equations (14) – (15), the result

$$\begin{aligned}\widehat{\mathrm{SS}}_R \;=\; & \hat{\sigma}_1^{-2}\|\phi_1(\boldsymbol{I}_n-\boldsymbol{P}_{X(V^{-1})})\boldsymbol{y}\|^2 + \hat{\sigma}_2^{-2}\|\phi_2(\boldsymbol{I}_n-\boldsymbol{P}_{X(V^{-1})})\boldsymbol{y}\|^2 \\ & + \hat{\sigma}_3^{-2}\|\phi_3(\boldsymbol{I}_n-\boldsymbol{P}_{X(V^{-1})})\boldsymbol{y}\|^2 + \hat{\sigma}_4^{-2}\|\phi_4(\boldsymbol{I}_n-\boldsymbol{P}_{X(V^{-1})})\boldsymbol{y}\|^2 = n - v\end{aligned} \tag{18}$$

then follows.

Theoretical background of the analysis (NRC design)

Now, the statistic F in (3) is reduced to the form

$$\widehat{F} = \frac{n-v}{v-1}\,\frac{\widehat{\mathsf{SS}}_V}{n-v} = \frac{\widehat{\mathsf{SS}}_V}{v-1}, \tag{19}$$

where $\widehat{\mathsf{SS}}_V$ is as in (4), but with σ_1^2, σ_2^2, σ_3^2 and σ_4^2 there replaced by their estimates.

However, the χ^2 distribution of SS_V, indicated in (6), is valid exactly only if the true stratum variances are used in the applied matrix $\boldsymbol{V}^{-1} = \sigma_1^{-2}\phi_1 + \sigma_2^{-2}\phi_2 + \sigma_3^{-2}\phi_3 + \sigma_4^{-2}\phi_4 + \sigma_5^{-2}\phi_5$. As for the component $\sigma_5^{-2}\phi_5$, it does not in fact play any role in the application of formula (6) given for SS_V. Thus, when using in $\boldsymbol{V}^{-1}$ the estimates of σ_1^2, σ_2^2, σ_3^2 and σ_4^2 obtained from (14) – (15), the distribution can be regarded as approximate only.

$$V^{-1} = \sigma_1^{-2}\phi_1 + \sigma_2^{-2}\phi_2 + \sigma_3^{-2}\phi_3 + \sigma_4^{-2}\phi_4 + \sigma_5^{-2}\phi_5$$

A desirable simplification can be obtained when the dispersion matrix $\boldsymbol{V}$ is replaced by the matrix

$$\boldsymbol{V}_* = \sigma_1^2\phi_1 + \sigma_2^2\phi_2 + \sigma_3^2\phi_3 + \sigma_4^2(\phi_4 + \phi_5),$$

$$\text{where} \quad \phi_4 + \phi_5 = n_0^{-1}\boldsymbol{X}_B\boldsymbol{X}_B' = \boldsymbol{I}_b \otimes n_0^{-1}\mathbf{1}_{n_0}\mathbf{1}_{n_0}'$$

The inverted matrix $\boldsymbol{V}_*^{-1}$ can be obtained as

$$\boldsymbol{V}_*^{-1} = \sigma_1^{-2}\phi_1 + \sigma_2^{-2}\phi_2 + \sigma_3^{-2}\phi_3 + \sigma_4^{-2}(\phi_4 + \phi_5).$$

Some simplifying reformulations (NRC design)

$$\boldsymbol{V} = \boldsymbol{V}_* + (\sigma_5^2 - \sigma_4^2) n^{-1} \mathbf{1}_n \mathbf{1}_n' \quad \text{and} \quad \boldsymbol{V}^{-1} = \boldsymbol{V}_*^{-1} + (\sigma_5^{-2} - \sigma_4^{-2}) n^{-1} \mathbf{1}_n \mathbf{1}_n'$$

Now it can be shown that the BLUE of $\boldsymbol{\tau}_* = (\boldsymbol{I}_v - n^{-1}\mathbf{1}_v \boldsymbol{r}')\boldsymbol{\tau}$, i.e.,

$$\hat{\boldsymbol{\tau}}_* = (\boldsymbol{I}_v - n^{-1}\mathbf{1}_v \boldsymbol{r}')\hat{\boldsymbol{\tau}} = (\boldsymbol{I}_v - n^{-1}\mathbf{1}_v \boldsymbol{r}')(\boldsymbol{X}_1' \boldsymbol{V}^{-1} \boldsymbol{X}_1)^{-1} \boldsymbol{X}_1' \boldsymbol{V}^{-1} \boldsymbol{y},$$

can equivalently be written as

$$\hat{\boldsymbol{\tau}}_* = (\boldsymbol{I}_v - n^{-1}\mathbf{1}_v \boldsymbol{r}')(\boldsymbol{X}_1' \boldsymbol{V}_*^{-1} \boldsymbol{X}_1)^{-1} \boldsymbol{X}_1' \boldsymbol{V}_*^{-1} \boldsymbol{y}_*, \tag{20}$$

where $\boldsymbol{y}_* = (\boldsymbol{I}_n - n^{-1}\mathbf{1}_n \mathbf{1}_n')\boldsymbol{y}$, for which
$\mathsf{E}(\boldsymbol{y}_*) = (\boldsymbol{I}_n - n^{-1}\mathbf{1}_n \mathbf{1}_n')\boldsymbol{X}_1 \boldsymbol{\tau} = \boldsymbol{X}_1(\boldsymbol{I}_v - n^{-1}\mathbf{1}_v \boldsymbol{r}')\boldsymbol{\tau} = \boldsymbol{X}_1 \boldsymbol{\tau}_*$ and
$\mathsf{D}(\boldsymbol{y}_*) = (\boldsymbol{I}_n - n^{-1}\mathbf{1}_n \mathbf{1}_n')\boldsymbol{V}_*(\boldsymbol{I}_n - n^{-1}\mathbf{1}_n \mathbf{1}_n')$.

Some simplifying reformulations (RC design)

The dispersion matrix of $\hat{\tau}_*$, can be presented as

$$\mathrm{D}(\hat{\tau}_*) = (\boldsymbol{I}_v - n^{-1}\mathbf{1}_v\boldsymbol{r}')(\boldsymbol{X}_1'\boldsymbol{V}_*^{-1}\boldsymbol{X}_1)^{-1}(\boldsymbol{I}_v - n^{-1}\boldsymbol{r}\mathbf{1}_v'). \tag{21}$$

Furthermore, the formulae of SS_V and SS_R, given in (4) for treatments (varieties) and in (5) for residuals, can equivalently be written as

$$\mathrm{SS}_V = \hat{\tau}_*'\boldsymbol{X}_1'\boldsymbol{V}_*^{-1}\boldsymbol{X}_1\hat{\tau}_* = \boldsymbol{y}_*'\boldsymbol{V}_*^{-1}\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{V}_*^{-1}\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{V}_*^{-1}\boldsymbol{y}_*, \tag{22}$$

$$\mathrm{SS}_R = \boldsymbol{y}_*'[\boldsymbol{V}_*^{-1} - \boldsymbol{V}_*^{-1}\boldsymbol{X}_1(\boldsymbol{X}_1'\boldsymbol{V}_*^{-1}\boldsymbol{X}_1)^{-1}\boldsymbol{X}_1'\boldsymbol{V}_*^{-1}]\boldsymbol{y}_*, \tag{23}$$

with $\boldsymbol{y}_* = (\boldsymbol{I}_n - n^{-1}\mathbf{1}_n\mathbf{1}_n')\boldsymbol{y}$,. The formulae (22) and (23) provide the sum

$$\mathrm{SS}_V + \mathrm{SS}_R = \boldsymbol{y}_*'\boldsymbol{V}_*^{-1}\boldsymbol{y}_* = \mathrm{SS}_T \quad \text{(say)}, \tag{24}$$

which can be called the total sum of squares. Referring again to Rao and Mitra (1971, Theorem 9.2.1), it can be shown that

$$\mathrm{SS}_T \sim \chi^2(n-1,\delta), \quad \text{with} \quad \delta = \tau_*'\boldsymbol{X}_1'\boldsymbol{V}_*^{-1}\boldsymbol{X}_1\tau_*$$

equivalent to δ as given in (6).

Table 1. Analysis of variance for an experiment in a nested row-column design with orthogonal block structure

Source of variation	Degrees of freedom	Sum of squares	Expected mean square
Treatments	$v-1$	SS_V	$1+\delta/(v-1)$
Residuals	$n-v$	SS_R	1
Total	$n-1$	SS_T	—

Table 2. Analysis of variance for an experiment in an NRC design

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square
Treatments	$v-1$	$\widehat{\mathsf{SS}}_V = \hat{\boldsymbol{\tau}}'_* \boldsymbol{X}'_1 \hat{\boldsymbol{V}}_*^{-1} \boldsymbol{X}_1 \hat{\boldsymbol{\tau}}_*$	$\widehat{\mathsf{MS}}_V = \frac{\widehat{\mathsf{SS}}_V}{(v-1)}$
Residuals	$n-v$	$\widehat{\mathsf{SS}}_R = \boldsymbol{y}'_*[\hat{\boldsymbol{V}}_*^{-1} - \hat{\boldsymbol{V}}_*^{-1}\boldsymbol{X}_1(\boldsymbol{X}'_1\hat{\boldsymbol{V}}_*^{-1}\boldsymbol{X}_1)^{-1}\boldsymbol{X}'_1\hat{\boldsymbol{V}}_*^{-1}]\boldsymbol{y}_* = n-v$	1
Total	$n-1$	$\widehat{\mathsf{SS}}_T = \boldsymbol{y}'_*\hat{\boldsymbol{V}}_*^{-1}\boldsymbol{y}_*$	—

Verification of the hypothesis $H_0 : \tau_* = \mathbf{0}$ will be based on the formulae presented in Table 2, which correspond to the statistics

$$\widehat{F} = \frac{n-v}{v-1}\,\frac{\widehat{\mathsf{SS}}_V}{n-v} = \frac{\widehat{\mathsf{SS}}_V}{v-1}, \qquad (25)$$

where the estimated mean square has, under H_0, an approximated $\chi^2(v-1,0)/(v-1)$ distribution.

Hypothesis for a Set of Contrasts – Two Ways

Let us consider one hypothesis concerning a set of contrasts (or a single contrast) among treatment parameters

$$H_{0,\mathrm{L}} : \boldsymbol{U}_{\mathrm{L}}' \boldsymbol{\tau}_* = \boldsymbol{0}, \quad \text{where} \quad \boldsymbol{U}_{\mathrm{L}}' \boldsymbol{1}_v = \boldsymbol{0}. \tag{26}$$

Caliński et al. indicate that the BLUE of $\boldsymbol{U}_{\mathrm{L}}' \boldsymbol{\tau}_*$ is of the form

$$\boldsymbol{U}_{\mathrm{L}}' \hat{\boldsymbol{\tau}}_* = \boldsymbol{U}_{\mathrm{L}}' \hat{\boldsymbol{\tau}} = \boldsymbol{U}_{\mathrm{L}}' (\boldsymbol{X}_1' \boldsymbol{V}_*^{-1} \boldsymbol{X}_1)^{-1} \boldsymbol{X}_1' \boldsymbol{V}_*^{-1} \boldsymbol{y}_*. \tag{27}$$

and the relevant sum of squares can then be obtained in the form

$$\mathsf{SS}(\boldsymbol{U}_{\mathrm{L}}) = \hat{\boldsymbol{\tau}}_*' \boldsymbol{U}_{\mathrm{L}} [\boldsymbol{U}_{\mathrm{L}}' (\boldsymbol{X}_1' \boldsymbol{V}_*^{-1} \boldsymbol{X}_1)^{-1} \boldsymbol{U}_{\mathrm{L}}]^{-} \boldsymbol{U}_{\mathrm{L}}' \hat{\boldsymbol{\tau}}_*, \tag{28}$$

with the d.f. equal to $\mathrm{rank}(\boldsymbol{U}_{\mathrm{L}})$.

CASE 1 – FACTORIAL EXPERIMENT

Two factors with a cross structure: F_1 with f_1 levels and F_2 with f_2 levels

$$\mathsf{SS}(\boldsymbol{U}_{\mathrm{F}_1}) + \mathsf{SS}(\boldsymbol{U}_{\mathrm{F}_2}) + \mathsf{SS}(\boldsymbol{U}_{\mathrm{F}_1\mathrm{F}_2}) = \mathsf{SS}_V. \tag{29}$$

$\boldsymbol{U}'_{\mathrm{L}}$ for $\mathrm{L} \in \{\mathrm{F}_1, \mathrm{F}_2, \mathrm{F}_1\mathrm{F}_2\}$ of the form:
$\boldsymbol{U}'_{\mathrm{F}_1} = (\boldsymbol{I}_{f_1} - \frac{1}{f_1}\mathbf{1}_{f_1}\mathbf{1}'_{f_1}) \otimes \frac{1}{f_2}\mathbf{1}_{f_2}\mathbf{1}'_{f_2},$
$\boldsymbol{U}'_{\mathrm{F}_2} = \frac{1}{f_1}\mathbf{1}_{f_1}\mathbf{1}'_{f_1} \otimes (\boldsymbol{I}_{f_2} - \frac{1}{f_2}\mathbf{1}_{f_2}\mathbf{1}'_{f_2}),$
$\boldsymbol{U}'_{\mathrm{F}_1\mathrm{F}_2} = (\boldsymbol{I}_{f_1} - \frac{1}{f_1}\mathbf{1}_{f_1}\mathbf{1}'_{f_1}) \otimes (\boldsymbol{I}_{f_2} - \frac{1}{f_2}\mathbf{1}_{f_2}\mathbf{1}'_{f_2}).$

It satisfies the condition $\boldsymbol{U}'_{\mathrm{L}}(\boldsymbol{X}'_1\boldsymbol{V}_*^{-1}\boldsymbol{X}_1)^{-1}\boldsymbol{U}_{\mathrm{L}^*} = \mathbf{O}$ for $\mathrm{L} \neq \mathrm{L}^*$ necessary for the partition SS_V.

CASE 2 – a letter-based representation of all-pairwise comparisons

- analysis of all $(v-1)v/2$ contrasts between pairs of treatments
- Based on all-pairwise significance statements (P values) for v treatments obtained during the verification of the hypotheses for all simple contrasts, a familiar letter display is obtained, where as usual treatments that do not differ significantly share a common letter.
- Algorithm: Piepho, H.-P. An Algorithm for a Letter-Based Representation of All-Pairwise Comparisons. *J. Comput. Graph. Stat.* **2004**, *13*, 456–466.
- multcompView package in R

EXAMPLE

Ratajkiewicz et al. (2018) analyzed a field experiment carried out in Poznań, Poland in 2011. Studies were conducted to improve the use of fungicides against potato late blight (*Phytophthora infestans* (Mont.) De Bary) (PLB) in tomato processing.

6	1	3	0	5	2
5	4	0	1	6	3
3	0	1	5	2	4
2	5	4	6	0	1
4	3	6	2	1	0
0	2	5	3	4	6

0	2	3	6	4	1
6	3	2	5	0	4
1	6	5	4	2	0
3	1	4	0	5	6
2	0	1	3	6	5
4	5	0	1	3	2

TREATMENT T	**SV factor** F_1	**Adjuvant factor** F_2
0 (control)	–	–
1	SV300	NO
2	SV300	MULTI
3	SV300	PMH
4	PSV	NO
5	PSV	MULTI
6	PSV	PMH

Table: Experimental data for a field experiment concerning the infestation of tomato plants by potato late blight (*Phytophthora infestans* (Mont.) De Bary).

T	Block	Row_{NR}	$Col._{NR}$	Obs.	T	Block	Row_{NR}	$Col._{NR}$	Obs.
6	1	1	1	71.960	0	2	7	7	91.812
1	1	1	2	77.501	2	2	7	8	73.984
3	1	1	3	67.575	3	2	7	9	54.820
0	1	1	4	94.480	6	2	7	10	61.700
5	1	1	5	75.321	4	2	7	11	70.860
2	1	1	6	86.494	1	2	7	12	71.447
5	1	2	1	71.806	6	2	8	7	65.981
4	1	2	2	74.680	3	2	8	8	49.203
0	1	2	3	95.943	2	2	8	9	67.858
1	1	2	4	67.492	5	2	8	10	60.500
6	1	2	5	60.864	0	2	8	11	91.317
3	1	2	6	77.532	4	2	8	12	73.392
3	1	3	1	69.535	1	2	9	7	73.379
0	1	3	2	91.750	6	2	9	8	59.850
1	1	3	3	73.470	5	2	9	9	59.127
5	1	3	4	68.000	4	2	9	10	62.750
2	1	3	5	78.925	2	2	9	11	70.702
4	1	3	6	78.283	0	2	9	12	91.160
2	1	4	1	88.079	3	2	10	7	65.208

EXAMPLE

$$\begin{bmatrix} 0 & \mathbf{0}_6' \\ \mathbf{0}_6 & \left(\boldsymbol{I}_2 - \frac{1}{2}\mathbf{1}_2\mathbf{1}_2'\right) \otimes \frac{1}{3}\mathbf{1}_3\mathbf{1}_3' \end{bmatrix} \boldsymbol{\tau} = \boldsymbol{U}_{\mathrm{F}_1}' \boldsymbol{\tau} \equiv \boldsymbol{U}_{\mathrm{F}_1}' \boldsymbol{\tau}_*,$$

$$\begin{bmatrix} 0 & \mathbf{0}_6' \\ \mathbf{0}_6 & \frac{1}{2}\mathbf{1}_2\mathbf{1}_2' \otimes \left(\boldsymbol{I}_3 - \frac{1}{3}\mathbf{1}_3\mathbf{1}_3'\right) \end{bmatrix} \boldsymbol{\tau} = \boldsymbol{U}_{\mathrm{F}_2}' \boldsymbol{\tau} \equiv \boldsymbol{U}_{\mathrm{F}_2}' \boldsymbol{\tau}_*,$$

$$\begin{bmatrix} 0 & \mathbf{0}_6' \\ \mathbf{0}_6 & \left(\boldsymbol{I}_2 - \frac{1}{2}\mathbf{1}_2\mathbf{1}_2'\right) \otimes \left(\boldsymbol{I}_3 - \frac{1}{3}\mathbf{1}_3\mathbf{1}_3'\right) \end{bmatrix} \boldsymbol{\tau} = \boldsymbol{U}_{\mathrm{F}_1\mathrm{F}_2}' \boldsymbol{\tau} \equiv \boldsymbol{U}_{\mathrm{F}_1\mathrm{F}_2}' \boldsymbol{\tau}_*,$$

$$\frac{\sqrt{10}}{6} \begin{bmatrix} 6 & -\mathbf{1}_6' \end{bmatrix} \boldsymbol{\tau} = \boldsymbol{U}_{\mathrm{C}}' \boldsymbol{\tau} \equiv \boldsymbol{U}_{\mathrm{C}}' \boldsymbol{\tau}_*,$$

It can also be checked that

$$\mathsf{SS}(\boldsymbol{U}_{\mathrm{C}}) + \mathsf{SS}(\boldsymbol{U}_{\mathrm{F}_1}) + \mathsf{SS}(\boldsymbol{U}_{\mathrm{F}_2}) + \mathsf{SS}(\boldsymbol{U}_{\mathrm{F}_1\mathrm{F}_2}) = \mathsf{SS}_V,$$

Source	**Degrees of Freedom**	**Sum of Squares**	**Mean Square**	$\hat{F}$	**p Value**
Treatments	6	450.024	75.004	75.004	< 0.0001
F_1	1	14.3922	14.3922	14.3922	0.00015
F_2	2	35.9117	17.9558	17.9558	< 0.0001
$\mathrm{F}_1\mathrm{F}_2$	2	35.3518	17.6759	17.6759	< 0.0001
C	1	364.368	364.368	364.368	< 0.0001
Residuals	65	65	1		
Total	71	515.0241			

The results presented in the table are obtained with the use of the empirical estimates $\tilde{\tau}$ and $\tilde{\tau}_*$, i.e., based on:

$\hat{\sigma}_1^2 = 15.726$, $\hat{\sigma}_2^2 = 9.487$, $\hat{\sigma}_3^2 = 93.042$ and $\hat{\sigma}_4^2 = 1282.51$.

Table: Empirical estimates of τ and τ_*. Treatments that do not differ significantly share a common letter.

TREATMENT **T**	**SV** **Factor** F_1	**Adjuvant** **Factor** F_2	**Treatment Effect** $\tilde{\tau}$	**Main Effect** $\tilde{\tau}_*$	
0 (control)	–	–	93.125	(19.948)	a
1	SV300	NO	72.328	(−0.850)	c
2	SV300	MULTI	77.398	(4.221)	b
3	SV300	PMH	63.682	(−9.496)	d
4	PSV	NO	70.527	(−2.651)	c
5	PSV	MULTI	65.201	(−7.977)	d
6	PSV	PMH	65.993	(−7.185)	d

Concluding remarks

- The discovered results concerning the proposed approach to ANOVA for experiments with orthogonal block structure seem to be useful.
- The main advantage of the proposed approach is the fact that the ANOVA results are obtainable directly, not by performing first some partial analyses, under relevant stratum submodels, and then combining their results.

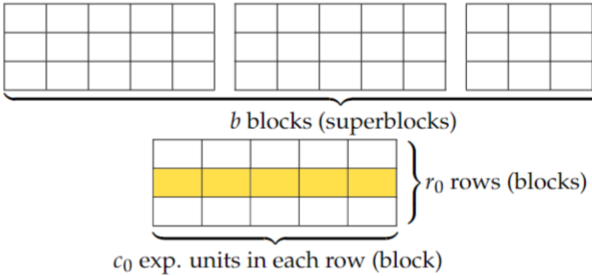
What if...

ROW-COLUMN DESIGN ($b = 1$)	
Parameters	
Model	$y = X_1\tau + \cancel{X_B\beta} + X_{R(B)}\rho + X_{C(B)}\gamma + \eta + e$ $V = \sigma_1^2\phi_1 + \sigma_2^2\phi_2 + \sigma_3^2\phi_3 + \sigma_5^2\phi_5$ taking $X_B = \mathbf{1}_n$ $V_* = \sigma_1^2(\phi_1 - \phi_5) + \sigma_2^2(\phi_2 + \phi_5) + \sigma_3^2(\phi_3 + \phi_5)$

What if...

NESTED BLOCK DESIGN

Parameters



Model

$$\boldsymbol{y} = \boldsymbol{X}_1\boldsymbol{\tau} + \boldsymbol{X}_B\boldsymbol{\beta} + \boldsymbol{X}_{R(B)}\boldsymbol{\rho} + \cancel{\boldsymbol{X}_{C(B)}\boldsymbol{\gamma}} + \boldsymbol{\eta} + \boldsymbol{e}$$

$$\boldsymbol{V} = \sigma_1^2\boldsymbol{\phi}_1 + \sigma_2^2\boldsymbol{\phi}_2 + \sigma_4^2\boldsymbol{\phi}_4 + \sigma_5^2\boldsymbol{\phi}_5 \text{ taking } \boldsymbol{\phi}_1 = \boldsymbol{I}_n - c_0^{-1}\boldsymbol{X}_{R(B)}\boldsymbol{X}'_{R(B)}$$

$$\boldsymbol{V}_* = \sigma_1^2\boldsymbol{\phi}_1 + \sigma_2^2\boldsymbol{\phi}_2 + \sigma_4^2(\boldsymbol{I}_n - \boldsymbol{\phi}_1 - \boldsymbol{\phi}_2)$$

What if...

PROPER BLOCK DESIGN

Parameters

b blocks ($n_0 = r_0 c_0$ plots in each block)

Model

$$y = X_1\tau + X_B\beta + \cancel{X_{R(B)}\rho} + \cancel{X_{C(B)}\gamma} + \eta + e$$

$$V = \sigma_1^2\phi_1 + \sigma_4^2\phi_4 + \sigma_5^2\phi_5 \text{ taking } \phi_1 = I_n - n_0^{-1}X_B X_B'$$

$$V_* = \sigma_1^2\phi_1 + \sigma_4^2(I_n - \phi_1)$$

References

- Caliński T., Siatkowski I. (2017): On a new approach to the analysis of variance for experiments with orthogonal block structure. I. Experiments in proper block designs. Biometrical Letters 54: 91-122.
- Caliński T., Siatkowski I. (2018): On a new approach to the analysis of variance for experiments with orthogonal block structure. II. Experiments in nested block designs. Biometrical Letters 55: 147-178.
- Caliński, T., Łacka, A., Siatkowski, I. (2019): On a new approach to the analysis of variance for experiments with orthogonal block structure. III. Experiments in row-column designs. Biometrical Letters 56: 183–213.
- Caliński, T., Łacka, A., Siatkowski, I. (2020): On a new approach to the analysis of variance for experiments with orthogonal block structure. IV. Experiments in split-plot designs. Biometrical Letters 57: 183–213.
- Łacka, A. NRC Designs – New Tools for Successful Agricultural Experiments. Agronomy **2021**, *11(12)*, 2406.

References

- Nelder, J.A. The combination of information in generally balanced designs. *J. R. Stat. Soc. Ser. B* **1968**, *30*, 303–311.
- Houtman, A.M.; Speed, T.P. Balance in designed experiments with orthogonal block structure. *Ann. Stat.* **1983**, *11*, 1069–1085.
- Piepho, H.-P. An Algorithm for a Letter-Based Representation of All-Pairwise Comparisons. *J. Comput. Graph. Stat.* **2004**, *13*, 456–466.
- Graves, S.; Piepho, H.-P.; Selzer, L. With Help from Dorai-Raj S. multcompView: Visualizations of Paired Comparisons. R package version 0.1-8. 2019. Available online: `https://CRAN.R-project.org/package=multcompView` (accessed on 19 December 2019).
- Ratajkiewicz, H.; Kierzek, R.; Raczkowski, M.; HoÅodyÅska-Kulas, A.; Åacka, A.; Szulc, T. The effect of coarse-droplet spraying with double flat fan air induction nozzle and spray volume adjustment model on the efficiency of fungicides and residues in processing tomato. *Span. J. Agric. Res.* **2018**, *16*, e1001.