WEIGHTED RESTRICTED MODELS: A NEW INSIGHT ON SUMS OF SQUARES AND HYPOTHESES TESTING

Daniel Furtado FERREIRA

• ABSTRACT: Gauss-Markov’s linear models have been used in many experimentation problems. To the unbalanced model, the sums of squares associated hypothesis are not so easily interpreted. The purpose of this paper was to present the weighted parametric restricted model and to obtain the sums of squares associated hypotheses. A 2-way crossed classification model was used. We conclude that researchers not so familiarized with statistical analysis of unbalanced data easily interpreted the presented hypotheses. $W$-restricted model lead to the sums of squares and associated hypotheses equivalent to those presented in type II SAS® sums of squares and estimable functions.

• KEYWORDS: Parametric restrictions, testable hypothesis, and linear models.

1 Introduction

On experimentation, many problems have been solved by Gauss-Markov’s linear models, i.e., $Y=X\theta + e$, where $Y$ is a observational vector of random variables; $X$ is a known coefficients matrix (model matrix) with rank $k \leq \min\{n, p\}$; $\theta$ is a vector of unknown parameters; and $e$ is a vector of random non-observable variables, with $e \sim N(0, \sigma^2I)$. In these cases it is common to adopt an overparametrized model, which explicitly shows one parameter to each factor levels.
This feature of having more parameters in a model than there are observed cell means to estimate them from is well described in literature (Searle, 1987). To circumvent this situation we usually use estimable functions, which has us confined attention to only certain functions of the parameters that can be estimated from the data, or we use reparametrization, wherein we define relationships among parameters of an overparametrized model. The “new” parameters can be estimated from the data.

These overparametrized linear models are not, in general, of full rank. Due to that, in least square solution of the normal equations, it is necessary to obtain a generalized inverse matrix. For unbalanced models, i.e., for those ones in which the number of observation per cell (where the intersection of one level of every factor is being considered) is not equal, the sum of squares associated hypotheses are not easily interpreted (Searle, 1987). The unbalanced models sometimes include empty cells, which raise difficulties to obtain the sum of squares and associated hypothesis. Latin square, partial balanced incomplete blocks and balanced incomplete blocks experiments must not be considered unbalanced since they have planned structure, though occurring empty cells. These cases are what may be called planned unbalancedness.

Within unbalanced data two divisions are available. One is for data in which all cells contain data, all-cells-filled data. Complementary to this are some-cells-empty data, wherein there are some cells that have no data.

Testable hypotheses on balanced analysis of variance are of equal effects to the factor levels that have been studied. Although, in unbalanced case, the testable hypothesis using \( F \) statistic is not so clear for the understanding. In these cases, many kinds of sums of squares are available, which account for different hypothesis and have different utilities (Searle, 1987). The available methods are often not as easily interpreted as methods for balanced data; neither are they well known, neither are so widely documented.

Due to the hypothesis complexity in the unbalanced models, an alternative for their simplification is to use a restricted parameter space. The most usual restrictions for this purpose are those known as \( \Sigma \) restrictions, which adopt the sum of factors effects equal to zero and transform the overparametrized model of incomplete rank into a full rank model. In this case all parameters and linear combinations of them are estimable.
To obtain the restricted model sums of squares, with column full rank, Searle (1987) presented the invert part of inverse algorithm. Those sums of squares test the equality of each parameter under restriction to zero.

There are alternatives to the $\Sigma$-restrictions; one that is used what will be called the $W$-restrictions or weighted restrictions-using as weights the number of observations (Searle et al., 1981; Carlson and Tim, 1974; Speed and Hocking, 1976). These authors have studied the $W$-restriction and they indicate that relationships to the classical analysis of variance are given by $R(\cdot)$ notation, of a factor adjusted for the constant term in the model. Comparisons between $W$ and $\Sigma$-restrictions were not done.

The purpose of this paper is to present the utilization of weighted parametric restricted models and to obtain model sums of squares and associated hypothesis and also to compare them with those of $\Sigma$-restricted models.

2 Methodology

To obtain the sums of squares and the associated hypotheses under parametric restrictions, a with-interactions overparametrized model (Gauss-Markov) was considered for the 2-way crossed classification (factors A and B) with different number of observation per cell.

\[ Y_{ijk} = \mu + \alpha_i + \beta_j + \delta_{ij} + e_{ijk} \]  

(1)

where $\mu$ is a general mean, $\alpha_i$ is an effect due to the $i$th level of the row factor (A), $\beta_j$ is an effect due to $j$th level of the column factor (B) and $\delta_{ij}$ represents an effect for the interaction of row $i$ with column $j$.

By the matrix notation,

\[ \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix} = \begin{pmatrix} \mathbf{X} \end{pmatrix} \begin{pmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_k \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} \]  

(2)

where $Y$, $X$, $\theta$ and $e$ were previously defined.

To exemplify and easily demonstrate the results obtained using $\Sigma$ and $W$ restrictions the data given in Table 1 were used, without loss of generality.
Table 1 – Observed data obtained in a 2-way (A and B) experiment with different number of observation per cell. Factor A with two levels and factor B with three ones.

<table>
<thead>
<tr>
<th></th>
<th>j=1</th>
<th>j=2</th>
<th>j=3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i=1</td>
<td>4, 6</td>
<td>8</td>
<td>2, 4</td>
<td>24(5)4.8</td>
</tr>
<tr>
<td></td>
<td>10*(2)5</td>
<td>8(1)8</td>
<td>6(2)3</td>
<td></td>
</tr>
<tr>
<td>i=2</td>
<td>2, 4</td>
<td>1, 2, 3</td>
<td>-</td>
<td>12(5)2.4</td>
</tr>
<tr>
<td></td>
<td>6(2)3</td>
<td>6(3)2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>16(4)4</td>
<td>14(4)3,5</td>
<td>6(2)3</td>
<td>36(10)3.6</td>
</tr>
</tbody>
</table>

* Total, (number of replicates) and mean.

To distinguish the parameters in the model that have $\Sigma$-restrictions from those from (1) we use the symbols $\mu^\cdot$, $\alpha^\cdot$, $\beta^\cdot$, and $\delta^\cdot$. To the $\Sigma$-model, considering the data structure from Table 1, the following parametric restrictions were considered.

$$ \sum_i \alpha_i = 0 $$
$$ \sum_j \beta_j = 0 $$
$$ \sum_j \delta_{ij} = \sum_i \delta_{ij} = 0 $$
$$ \delta_{13} = 0 $$
$$ \delta_{11} = -\delta_{12} = -\delta_{21} = \delta_{22} $$

(3)

For the weighted parametric restricted model ($W$-Model) the restrictions are presented below. In this model we use the symbols $\mu^\$, $\alpha^\$, $\beta^\$ and $\delta^\$. 

12

\[
\sum_i n_i \alpha_i^* = 0 \\
\sum_j n_j \beta_j = 0 \\
\sum_j n_j \delta_{ij} = \sum_i n_j \delta_{ij} = 0 \\
\delta_{13} = 0 \\
\delta_{11} = -\frac{1}{2} \delta_{21} = -\frac{1}{2} \delta_{22}
\] (4)

The sums of squares were, in the R notation of Searle (1987),
\[ R(\alpha/\mu, \beta, \delta) \quad R(\beta/\mu, \alpha, \delta) \quad R(\delta/\mu, \alpha, \beta) \] in the \( \Sigma \)-model; and
\[ R(\alpha^* / \mu^*, \beta^*, \delta^*) \quad R(\beta^* / \mu^*, \alpha^*, \delta^*) \quad R(\delta^* / \mu^*, \alpha^*, \beta^*) \] in the \( W \)-model, to the factor \( A \), factor \( B \) and interaction \( (A*B) \), respectively.

These full rank models test the sum of squares associated hypothesis of equality of each parameter to zero, as presented below.

\[ H : \alpha_i = 0, \quad H : \beta_j = 0, \quad H : \delta_{ij} = 0 \quad \Sigma \text{-mod el} \]
\[ H : \alpha_i^* = 0, \quad H : \beta_j^* = 0, \quad H : \delta_{ij}^* = 0 \quad W \text{-mod el} \]

The sums of squares were obtained from the invert part of inverse algorithm presented in Searle (1987). To do so, let \( X \) be a model matrix under \( \Sigma \)-restriction (Equations 3), and then take the partitions of \( (X^t, X)^t \) as follows:

\[
(X^t, X)^t = 
\begin{bmatrix}
X_{11} & X_{12} & X_{13} & X_{14} \\
X_{21} & X_{22} & X_{23} & X_{24} \\
X_{31} & X_{32} & X_{33} & X_{34} \\
X_{41} & X_{42} & X_{43} & X_{44}
\end{bmatrix}
\] (6)

Where \( X_{11} \) (1x1) are related to \( \hat{\mu} \), \( X_{22} \) (1x1) are related to \( \hat{\alpha}_j \), \( X_{33} \) (2x2) are related to \( \hat{\beta}_1 \), \( \hat{\beta}_2 \), and \( X_{44} \) (1x1) are related to \( \hat{\delta}_{11} \). The same partitions of \( \hat{\theta} \) were:

\[
\hat{\theta} = \begin{bmatrix}
\hat{\mu} \\
\hat{\alpha}_j \\
\hat{\beta}_1 \\
\hat{\beta}_2 \\
\hat{\delta}_{11}
\end{bmatrix}
\]

Then the sums of squares obtained by the algorithm, are:

\[
R(\alpha/\mu, \beta, \delta) = \hat{\alpha}_j' X_{22}^{-1} \hat{\alpha}_j \tag{7}
\]

\[
R(\beta/\mu, \alpha, \delta) = (\hat{\beta}_1, \hat{\beta}_2)' X_{33}^{-1} \begin{bmatrix}
\hat{\beta}_1 \\
\hat{\beta}_2
\end{bmatrix} \tag{8}
\]

\[
R(\delta/\mu, \alpha, \beta) = \hat{\delta}_{11}' X_{44}^{-1} \hat{\delta}_{11} \tag{9}
\]

In the same way it was obtained the \( W \)-model sums of squares, replacing \( X \), by \( X_W \) and \( \hat{\Theta} \) by \( \hat{\theta} \).

### 3 Results and discussion

#### 3.1 \( \mathcal{E} \)-parametric restricted model

Adopting model (1) and using data showed in Table 1, the \( (X, X)' \) matrix and the least square estimator vector were:
\[ (X_r' X_r)^{-1} = \frac{1}{144} \begin{bmatrix} 17 & -3 & -5 & -5 & -3 \\ -3 & 21 & 3 & 15 & -3 \\ -5 & 3 & 29 & -7 & 3 \\ -5 & 15 & -7 & 41 & -9 \\ -3 & 3 & -9 & 21 & \end{bmatrix} \]

\[ \text{and } \theta = \frac{1}{3} \begin{bmatrix} 10 \\ 6 \\ 2 \\ 5 \\ -3 \end{bmatrix} \]

The sums of squares in the R notation were:

\[
R(\alpha/\mu, \beta, \delta) = \frac{1}{3} \begin{bmatrix} 21 \\ 144 \end{bmatrix}^{-1} 2 = 27.428571
\]

\[
R(\beta/\mu, \alpha, \delta) = 14.442105
\]

\[
R(\delta/\mu, \alpha, \beta) = 6.857143
\]

These sums of squares test the associated hypothesis presented in (5). Using a procedure describe by Searle (1987) \( W_r \) and \( W \) were defined as matrix of \( r \) distinctly rows of \( X \) and \( X_r \) respectively. In this model, \( r \) was also considered the number of non-empty cells. Then \( \theta \) and \( \hat{\theta} \) are related by equating cell means: \( W_r \hat{\theta} = W \theta \). The solutions are:

\[
\hat{\theta} = W_r^{-1}W \theta \quad (10)
\]

For the example of Table 2 (10) is given as follows:

\[
\mu = \mu + \frac{1}{2}(\alpha_1 + \alpha_2) + \frac{1}{3}(\beta_1 + \beta_2 + \beta_3) + \frac{1}{12}(\delta_{11} + \delta_{12} + 4\delta_{13} + 3\delta_{21} + 3\delta_{22})
\]

\[
\alpha_1 = \frac{1}{2}(\alpha_1 - \alpha_2) + \frac{1}{4}(\delta_{11} + \delta_{12} - \delta_{21} - \delta_{22})
\]

\[
\beta_1 = \frac{1}{3}(2\beta_1 - \beta_2 - \beta_3) + \frac{1}{12}(5\delta_{11} - \delta_{13} - 4\delta_{12} + 3\delta_{21} - 3\delta_{22})
\]

\[
\beta_2 = \frac{1}{3}(-\beta_1 + 2\beta_2 - \beta_3) + \frac{1}{12}(-\delta_{11} + 5\delta_{12} - 4\delta_{13} - 3\delta_{21} + 3\delta_{22})
\]

\[
\delta_{11} = \frac{1}{4}(\delta_{11} - \delta_{12} - \delta_{21} + \delta_{22})
\]
With this relationships it was possible to verify that in the non-restricted overparametrized model (S-model), the sums of squares associated hypotheses were respectively:

\[
H_0^1 : \alpha_1 - \alpha_2 + \frac{1}{2}(\delta_{12} + \delta_{21} - \delta_{22}) = 0
\]

\[
H_0^2 : \left[2\beta_1 - \beta_2 - \beta_3 + \frac{1}{4} \left(5\delta_{11} - \delta_{12} - 4\delta_{13} + 3\delta_{21} - 3\delta_{22}\right) = 0 \right.
\]

\[
\left. - \beta_1 + 2\beta_2 - \beta_3 + \frac{1}{4} \left(-\delta_{11} - 5\delta_{12} - 4\delta_{13} - 3\delta_{21} + 3\delta_{22}\right) = 0 \right]
\]

(11)

\[
H_0^3 : \delta_{11} - \delta_{12} - \delta_{21} + \delta_{22} = 0
\]

An important observation related to the hypotheses of (11) was that it is the same as SAS\textsuperscript{®} type III estimable functions.

### 3.2 Weighted restricted model (W-model)

With the latter procedure, the \((X_w’X_w)^{-1}\) matrix and the \(\theta^*\) vector were obtained:

\[
(X_w’X_w)^{-1} = \frac{1}{980}
\begin{bmatrix}
98 & 0 & 0 & 0 & 0 \\
0 & 140 & 0 & 70 & 0 \\
0 & 0 & 147 & -98 & 0 \\
0 & 70 & -98 & 182 & 0 \\
0 & 0 & 0 & 0 & 105
\end{bmatrix}
\begin{bmatrix}
25.2 \\
13.0 \\
2.8 \\
5.8 \\
-6.0
\end{bmatrix}
\]

and \(\theta^* = \frac{1}{7}\)

Using the invert part of inverse algorithm, the sums of squares were:

\[
R(\alpha^*/\mu^*\beta^*\delta^*) = \frac{13}{7} \left(\frac{140}{980}\right)^{-1} = 24.142857
\]

\[
R(\beta^*/\mu^*\alpha^*\delta^*) = 11.142857
\]

\[
R(\delta^*/\mu^*\alpha^*\beta^*) = 6.857143
\]
The sums of squares associated hypotheses were presented in (5). Using the procedure showed by Searle (1987), \( W_w \) was defined as matrix of \( r \) distinctly rows of \( X_w \). Hence:

\[
\theta^* = W_w^{\dagger} W \theta
\]  

(12)

Using the data presented in Table 1, \( \theta^* \) and \( \theta \) are related as follows:

\[
\mu^* = \mu + \frac{1}{2}(\alpha_1 + \alpha_2) + \frac{1}{5}(2\beta_1 + 2\beta_2 + \beta_3) + \frac{1}{10}(2\delta_{11} + \delta_{12} + 2\delta_{13} + 2\delta_{21} + 3\delta_{22})
\]

\[
\alpha_1^* = \frac{1}{2}(\alpha_1 - \alpha_2) + \frac{1}{14}(4\delta_{11} + 3\delta_{12} - 4\delta_{21} - 3\delta_{22})
\]

\[
\beta_1^* = \frac{2}{10}(3\beta_1 - 2\beta_2 - \beta_3) + \frac{1}{10}(3\delta_{11} - \delta_{12} - 3\delta_{13} - 3\delta_{21} - 3\delta_{22})
\]

\[
\beta_2^* = \frac{2}{10}(-2\beta_1 + 3\beta_2 - \beta_3) + \frac{1}{35}(-2\delta_{11} + 9\delta_{12} - 7\delta_{13} - 12\delta_{21} + 12\delta_{22})
\]

\[
\delta_{11}^* = \frac{3}{14}(\delta_{11} - \delta_{12} - \delta_{21} + \delta_{22})
\]

Then the associated hypotheses in the non-restricted over-parametrized model were:

\[
H_0^1: \alpha_1 - \alpha_2 + \frac{1}{7}(4\delta_{11} + 3\delta_{12} - 4\delta_{21} - 3\delta_{22}) = 0
\]

\[
H_0^2: \begin{cases} 
3\beta_1 - 2\beta_2 - \beta_3 + \frac{1}{2}(3\delta_{11} - \delta_{12} - 2\delta_{13} + 3\delta_{21} - 3\delta_{22}) = 0 \\
-2\beta_1 + 3\beta_2 - \beta_3 + \frac{1}{7}(-2\delta_{11} + 9\delta_{12} - 7\delta_{13} - 12\delta_{21} + 12\delta_{22}) = 0
\end{cases}
\]

(13)

\[
H_0^3: \delta_{11} - \delta_{12} - \delta_{21} + \delta_{22} = 0
\]

Researchers not familiarized with Statistics easily interpret the hypotheses of the restricted models presented in (5). These hypotheses, although equivalent ones represent complex linear combinations of the parameters from \( S \)-model, as presented in (11) and (13). Due to that, these ones are not so easily interpreted, even so, for skilled statisticians. An important observation, related to the hypotheses of (13), obtained
Vectors \( \hat{\theta} \) and \( \theta^* \) are least square estimates of \( \theta \), with

\[
\text{Var}(\hat{\theta}) = (X_w' X_w)^{-1} \hat{\sigma}^2 \quad \text{and} \quad \text{Var}(\theta^*) = (X_r' X_r)^{-1} \hat{\sigma}^2.
\]

The best unbiased estimator of a linear combination of \( \theta \), say \( \lambda' \theta \), under normality, is obtained from

\[
\text{BLUE}(\lambda' \theta) = \lambda' \hat{\theta}
\]

in the \( \Sigma \)-model or from

\[
\text{BLUE}(\lambda' \theta) = \lambda' \theta^*
\]

in the \( W \)-model, for any non-null vector \( \lambda \). Since, \( X_r \) and \( X_w \) are of column full rank, then for any \( \lambda \neq 0 \) the BLUE will exist (Graybill, 1976) although, \( \lambda' \hat{\theta} \) and \( \lambda' \theta^* \) represent distinct linear function of the parameters in the unrestricted parameter space, as showed in (10) and (12).

4 Conclusions

The hypotheses presented in this paper were easily interpreted by researchers not familiarized with statistical analysis of unbalanced models. A weighted restricted model (\( W \)-model) lead to sums of squares which test the hypothesis \( H: \theta_i = 0, \forall i \) and \( \theta^*_i \in \theta^* \), besides they are equivalent to the type II SAS\(^{\circ}\) sums of squares and with associated hypotheses equivalent to the type II estimable function of that program. Comparing the results of the restricted models, it seems that the hypotheses related to \( W \)-model are more realistic for a practical purpose, since they consider as weights the number of observations in the parametric restrictions. Although these remarks
were made from an example, they can be extended to several unbalanced models if the latter ones have no disconnection.


- **RESUMO:** Os modelos lineares Gauss-Markov têm sido usados em muitos problemas de experimentação. Para os modelos desbalanceados as somas de quadrados, as hipóteses associadas não são facilmente interpretadas. Esse trabalho teve por objetivos apresentar restrições paramétricas ponderadas e obter as hipóteses associadas as somas de quadrados obtidas. Um modelo cruzado de dois fatores foi usado para ilustrar. Concluiu-se que os pesquisadores não muito familiarizados com as análises estatísticas de dados desbalanceados facilmente são capazes de interpretar essas hipóteses. O Modelo com restrição W conduz a somas de quadrados e hipóteses associadas equivalentes àsquelas apresentadas pelas somas de quadrados tipo II e funções estimáveis do SAS.

- **PALAVRAS-CHAVE:** Restrição paramétrica, hipóteses testadas, modelos lineares.

## 5 References


Recebido em 10.1.2000