Orthogonal families of one and two strata prime basis factorials

Paulo C. Rodrigues¹, Elsa E. Moreira¹, Vera M. Jesus¹ and João T. Mexia¹

¹CMA - Center in Mathematics and Applications, Faculty of Sciences and Technology, Universidade Nova de Lisboa
2829-516 Caparica, Portugal

¹Corresponding author: paulocanas@gmail.com; paulocanas@fct.unl.pt
Abstract

Structured families of models constitute an alternative to nesting in which interactions between factors that nest are usually not considered. In this alternative, for each treatment of a fixed effects design $\pi$, there is a linear model and the action of factors in $\pi$ on the fixed effects parameters of the models is studied. When $\pi$ has orthogonal structure, the family of models will be said to be orthogonal. The models in the family can have one, two or more strata. Models with two strata are obtained through nesting of one strata models, derived from base factorials. These are very well known models in which all factors have the same number of levels. We will use a general framework to consider orthogonal families of models derived from prime basis factorials. These are very well known models in which all factors have the same prime number of levels.

**Keywords:** orthogonal models; families of models; orthogonal block structure; double strata; factorial designs

1 Introduction

Let us consider a model with two disjoint sets of factors. The level combinations of the factors of each sets will be the treatments for that set. If each treatment of the first set nests every treatment of the second set, we could say that the model was obtained through balanced nesting. Then, the action of the two sets of factors could studied, but we would have to assume have no interactions between both sets of factors.

A different situation arises when we are interested in studying the action of the factors on the first set on the models with the second set of factors. We point out we have one such model for each treatment defined by the first set of factors. We will say that

1. these models constitute a family,
2. the treatments constitute a base design $\pi$.

We will assume that both $\pi$ and the models in the family have fixed effects and the factors in $\pi$ only act on the fixed effects part of the models. Then, the factors in $\pi$ will act on the estimable vectors of the models. To study that action, we will assume that both $\pi$ and the family models satisfy certain orthogonality conditions. As for the family, besides one strata models, we consider two strata models obtained through nesting.

We will apply our results to an interesting case in which both $\pi$ and the linear models are obtained through nesting from prime bases factorials. These models have over the years, (Raktoe, 1981; Montgomery, 2005; Mukerjee and Wu, 2006; Bailey, 2008), been widely used to study the effect of each factor on the response variable, as well as the interaction effects between factors on
the response variable. The case with a prime or a power of prime number of levels, especially for \( p = 2 \) (Cheng et al., 1998; Fontana et al., 2000; Cheng and Wu, 2002; Ye, 2003) or \( p = 3 \) (Hoke, 1974; Toman, 1994; Morris, 2000; Pistone and Rogantin, 2008) levels, has been widely studied.

We will start with some results on orthogonality for one and two strata models. Next we consider the case of prime base factorials and models obtained from them through nesting. Last, we see how to carry the joint analysis of families of models. The base model \( \pi \) of the families we consider will still enjoy orthogonality, so we say that they are orthogonal families.

## 2 Orthogonality

### 2.1 One strata models

We assume the mean vector of the observation vector \( y \) to be

\[
\mu = \mu_0 \otimes 1_r
\]

(1)

this is, there are \( \hat{n} \) groups of \( r \) contiguous observations with the same mean value and variance-covariance matrix \( \sigma^2 I_n \), where \( \otimes \) represents the Kronecker product of matrices. This model will be associated, (Fonseca et al., 2006), to an orthogonal partition

\[
R_{\hat{n}} = \oplus_{j=1}^m \omega_j.
\]

(2)

The hypothesis

\[
H_{0,j}: \mu_0 \in \omega_j^1, j = 1, ..., m
\]

(3)

holding when factors [groups of factors] have no effects [interactions].

Let the row vectors of \( A_j \) constitute an orthogonal basis for \( \omega_j, j = 1, ..., m \), thus

\[
S_j = \|A_j \otimes (\frac{1}{\sqrt{r}} 1_r') y\|^2, j = 1, ..., m
\]

(4)

will be the product of \( \sigma^2 \) by \( \chi^2_{g_j, \delta_j} \), this is, by a chi-square with \( g_j = \text{rank}(A_j) = \text{dim}(\omega_j), j = 1, ..., m \) degrees of freedom and non-centrality parameter

\[
\delta_j = \|A_j \otimes (\frac{1}{\sqrt{r}} 1_r') \mu\|^2 = r \|A_j \otimes \mu\|^2, j = 1, ..., m
\]

(5)

which is null when and only when \( H_{0,j}, j = 1, ..., m \) holds.

Moreover \( S_j, j = 1, ..., m \), will be independent from

\[
S = \|I_n \otimes (I_r - \frac{1}{r} J_r)y\|^2
\]

(6)

where \( J_r = 1_r 1_r' \), which is the product of \( \sigma^2 \) by a central chi-square with \( g = \hat{n}(r - 1) \) degrees of freedom. Thus, the statistics

\[
F_j = \frac{g S_j}{g_j S}, j = 1, ..., m
\]

(7)
have \( F \) distribution with \( g_j \) and \( g \) degrees of freedom and non-centrality parameters \( \delta_j, j = 1, \ldots, m \).

Since the non-centrality parameters are null when and only when the corresponding hypothesis hold, these \( F \) testes are unbiased.

Now, with \( \hat{\mu} = X\beta \),

\[
\psi = G\beta
\]  
(8)

is estimable, Mexia(1990), when

\[
G = CX,
\]  
(9)

thus

\[
\psi = CX\beta = C\hat{\mu} = C\left(\sum_{j=1}^{m} A'_j A_j\right)\hat{\mu},
\]  
(10)

since

\[
\sum_{j=1}^{m} A'_j A_j = I_n,
\]  
(11)

so with

\[
\eta_j = A_j\hat{\mu}, j = 1, \ldots, m
\]  
(12)

we get

\[
\psi = \sum_{j=1}^{m} C_j \eta_j
\]  
(13)

where \( C_j = CA'_j, j = 1, \ldots, m \). So, the estimable vectors are generalized linear combinations of the canonical estimable vectors \( \eta_j, j = 1, \ldots, m \). Moreover

\[
\delta_j = r\|\eta_j\|^2, j = 1, \ldots, m,
\]  
(14)

so we may write the hypothesis as

\[
H_{0,j} : \eta_j = 0, j = 1, \ldots, m.
\]  
(15)

### 2.2 Two strata models

When for the two initial models we have matrices \( A_j(l), j = 1, \ldots, m(l) ; l = 1, 2 \), with

\[
\begin{align*}
A_j(l) &= \frac{1}{\sqrt{n(l)}}1'_{n(l)}, l = 1, 2 \\
m(l) &= \sum_{j=1}^{m(l)} A'_j(l) A_j(l) = I_{n(l)}, l = 1, 2
\end{align*}
\]  
(16)

for the model obtained through nesting we have, (Fonseca et al., 2006),

\[
\begin{align*}
A_1 &= \frac{1}{\sqrt{n}}1'_{n} \\
A_j &= A_j(1) \otimes \frac{1}{\sqrt{n(2)}}1'_{n(2)}, j = 1, \ldots, m(1) \\
A_j &= I_{n(1)} \otimes A_{j-m(1)+1}, j = 1, \ldots, m = m(1) + m(2) - 1
\end{align*}
\]  
(17)
with \( \dot{n} = \dot{n}(1) + \dot{n}(2) \). So, it is straightforward to apply the results of the preceding section.

Actually, if you have for a model, the corresponding matrices \( A \) there is no difficulty in testing the relevant hypothesis and obtaining the canonical estimable vectors.

### 2.3 Prime basis factorials

We now present these models basing ourselves in Jesus et al. (2009).

Let us assume that are \( N \) factors with \( p \) (prime) levels, numbered from 0 to \( p - 1 \). Then, we have a \( p^N \) factorial with \( n = p^N \) treatments. The treatments may be represented by the vectors \( x = (x_1, ..., x_N) \), with components \( x_j = 0, ..., p - 1, j = 1, ..., N. \)

Considering the family of linear applications

\[
L(x|a) = \left( \sum_{j=1}^{N} a_j x_j \right)_{(p)},
\]

where \( a_j = 0, ..., j - 1; j = 1, ..., N \) and \( (p) \) indicates the use of module \( p \) arithmetic, there will be

\[
k_N(p) = \frac{p^N - 1}{p - 1}
\]

linear applications whose first non null coefficient is 1 (Jesus et al., 2009). The linear applications \( L(x|a_l), l = 1, ..., m \) are called reduced and are independent if \( (\sum_{l=1}^{m} c_l a_l)_{(p)} = 0 \) implies \( c_1 = ... = c_m = 0. \)

We order the reduced linear applications from 1 to \( k_N(p) \) and the treatments according to the indexes

\[
f(x) = 1 + \sum_{j=1}^{N} x_j p^{j-1}.
\]

With \( L(x|a_i), i = 1, 2 \) linearly independent reduced linear applications, the pair of equations

\[
L(x|a_i) = L_i(x) = b_i;
\]

with \( b_i = 1, ..., p - 1, i = 1, 2 \), enables us to express 2 components of \( x \) as a linear functions of the remaining. Thus

(i) \( L(x|a) \) takes each of its values for \( p^{N-1} \) vectors \( x \);

(ii) \( L(x|a_1) \) and \( L(x|a_2) \) take, each of the possible \( p^2 \) pairs of values, for \( p^{N-2} \) vectors \( x \).

To each linear application \( L_h = L_h(x), h = 1, ..., k_N(p) \), we can associate a \( p \times p^N \) matrix \( C(L_h), h = 1, ..., k_N(p) \), with elements

\[
c_{i,j}(L_h) = \begin{cases} 
0; & L_h(x_j) \neq i - 1 \\
1; & L_h(x_j) = i - 1,
\end{cases}
\]
\[ j = 1, \ldots, p^N, h = 1, \ldots, k_N(p). \]

We will have,
\[
\begin{align*}
    & C(L_h)C(L_h)' = p^{N-1}1_p, \quad h = 1, \ldots, k_N(p) \\
    & C(L_h)C(L_{h'})' = p^{N-2}1_p, \quad h \neq h',
\end{align*}
\]
where \( J_p = \frac{1}{p}1_p1_p' \).

Let \( T_p \) be a matrix obtained deleting the first row equal to \( \sqrt{\frac{1}{p}}1_p' \) of a \( p \times p \) orthogonal matrix \( P_p \) and \( q = p^{\frac{N-2}{2}} \). We take \( A(L_h) = \frac{1}{q}T_pC(L_h), \ h = 1, \ldots, k_N(p) \). It is easy to see that the sums of the elements in any row of \( A(L_h) \) is null, \( h = 1, \ldots, k_N(p) \).

### 3 Families of models

Let us assume the models in the family to be associated to the orthogonal partition
\[ R^h = \oplus_{j=1}^m \omega_j, \]
for which we had the matrices \( A_j, \ j = 1, \ldots, m \), with ranks \( g_j, j = 1, \ldots, m \). With
\[
\begin{align*}
    & d_0 = 0 \\
    & d_t = \sum_{j=1}^t g_j, t = 1, \ldots, m.
\end{align*}
\]
we put
\[ A_j = [\alpha_{d_{j-1}} \ldots \alpha_{d_j}]', j = 1, \ldots, m \]
where \( \alpha_i, i = 1, \ldots, d_m = \hat{m} \) represent the row vectors of \( A_j \) which constitute an orthogonal basis for \( \omega_j, j = 1, \ldots, m \).

Assuming the observations vectors \( y_1, \ldots, y_n \) of the models in the family to be normal independent with mean vectors \( \mu_1, \ldots, \mu_n \) and variance-covariance matrix \( \sigma^2I_{n,r} \), the values
\[ z_{i,l} = \left( \alpha_i \otimes \frac{1}{\sqrt{r}} 1_r \right)' y_l, i = 1, \ldots, \hat{n}, l = 1, \ldots, n, \]
will be normal independent with mean values
\[ \mu_{i,l} = \left( \alpha_i \otimes \frac{1}{\sqrt{r}} 1_r \right)' \mu_l, i = 1, \ldots, \hat{n}, l = 1, \ldots, n, \]
and variance \( \sigma^2 \). They will also be independent of the sum of squares
\[ S_l = \left\| \left( I_{\hat{n}} \otimes (I_r - \frac{1}{r}J_r) \right) y_l \right\|^2, l = 1, \ldots, n, \]
which will be the products by \( \sigma^2 \) of a central chi-squares with \( \hat{n}(r - 1) \) degrees of freedom. So
\[ S = \sum_{l=1}^n S_l \]
will be the product by $\sigma^2$ of a central chi-square with $g = n\hat{n}(r - 1)$ degrees of freedom. We point out that $S$ is independent of the $z_{i,l}$, $i = 1, \ldots, \hat{n}$, $l = 1, \ldots, n$. The $z_{i,1}, \ldots, z_{i,n}$ will be the components of the vectors $z_i, i = 1, \ldots, \hat{n}$.

The joint analysis may now be carried out at two levels. The first level is a row by row analysis. This analysis is performed by assuming that the models correspond to the treatments in a base design. We will consider the case in which the base design is associated to an orthogonal partition

$$\mathbb{R}^n = \bigoplus_{j=1}^{h} \nabla_j.$$  \hspace{1cm} (31)

where the row vectors of $K_0, K_1, \ldots, K_h$ constitute an orthogonal basis for $\nabla_0, \nabla_1, \ldots, \nabla_h$. Namely, we will take $K_0 = \frac{1}{\sqrt{n}}1'_{\hat{n}}$. The nullity spaces of the matrices $K_1, \ldots, K_h$ will correspond to the absence of effects and interactions for the factors in the base design. Thus if $\lambda$ is the mean vector for the base design we have the hypothesis

$$H_{0,v}: K_v \lambda = 0, v = 1, \ldots, h,$$  \hspace{1cm} (32)

which are special cases of the

$$H_{0,v}(b): K_v \lambda = K_v b, v = 1, \ldots, h.$$  \hspace{1cm} (33)

We now have the normal vectors $z_i = (z_{i,1}, \ldots, z_{i,n})$, with mean vectors $\mu_i = (\mu_{i,1}, \ldots, \mu_{i,n})$, and variance covariance matrix $\sigma^2 I_n$ independent of $S$. It is now straightforward to show that to test

$$H_{0,v,i}(b): K_v \mu_i = K_v b, v = 1, \ldots, h, i = 1, \ldots, \hat{n}$$  \hspace{1cm} (34)

we must use the statistic

$$F_{v,i}(b) = \frac{g}{g_v} \frac{\|K_v z_i - K_v b\|^2}{S}, v = 1, \ldots, h, i = 1, \ldots, \hat{n},$$  \hspace{1cm} (35)

which has $F$ distribution with $g_v$ and $g$ degrees of freedom and non-centrality parameter

$$\delta_{v,i}(b) = \frac{1}{\sigma^2} \|K_v \mu_i - K_v b\|^2, v = 1, \ldots, h, i = 1, \ldots, \hat{n}.$$  \hspace{1cm} (36)

Let the row vectors of $B$ be linearly independent, and put

$$\begin{cases}
\tilde{\Psi}_i(B) = B z_i, i = 1, \ldots, \hat{n} \\
\Psi_i(B) = B \mu_i, i = 1, \ldots, \hat{n}
\end{cases}.$$  \hspace{1cm} (37)

Now $\tilde{\Psi}_i(B)$ will be normal with mean vectors $\Psi_i(B)$ and variance covariance matrix $B (\sigma^2 I_n) B' = \sigma^2 BB'$ which, (Silvey, 1975), is regular since the $\text{rank}(BB') = \text{rank}(B)$ and the row vectors of $B$ are linearly independents. Thus to test

$$H_{0,i}(B, b): \Psi_i(B) = B b, i = 1, \ldots, \hat{n},$$  \hspace{1cm} (38)
we may use the statistic
\[
\mathcal{F}_i(B, b) = \frac{g}{r(B)} \left( \bar{\Psi}_i(B) - Bb \right)' \left( BB' \right)^{-1} \left( \bar{\Psi}_i(B) - Bb \right), v = 1, \ldots, n, i = 1, \ldots, \hat{n}.
\] (39)

Since the quadratic form \( Q_i(B, b) \) in the numerator is the product by \( \sigma^2 \) of a chi-square with 
\( r(B) = \text{rank}(B) \) degrees of freedom and non-centrality parameter
\[
\delta_j(B, b) = \frac{1}{\sigma^2} (\Psi_i(B) - Bb)' (BB')^{-1} (\Psi_i(B) - Bb), \quad i = 1, \ldots, \hat{n},
\] (40)
\( \mathcal{F}_i(B, b) \) will be \( F \) distributed with \( r(B) \) and \( g \) degrees of freedom and non-centrality parameter \( \delta_i(B, b), i = 1, \ldots, \hat{n} \).

Moreover, since \( H_{0,j}(B, b), i = 1, \ldots, \hat{n} \) can be written as
\[
H_{0,i}(B, b) : \delta_i(B, b) = 0, \quad i = 1, \ldots, \hat{n},
\] (41)
The test is strictly unbiased.

Now the pivot variable
\[
\mathcal{F}_{\hat{i}}(B, b) = \frac{g}{r(B)} \left( \bar{\Psi}_i(B) - \bar{\Psi}_i(B) \right)' \left( BB' \right)^{-1} \left( \bar{\Psi}_i(B) - \bar{\Psi}_i(B) \right), v = 1, \ldots, n, i = 1, \ldots, \hat{n}.
\] (42)
has central \( F \) distribution with \( r(B) \) and \( g \) degrees of freedom. So, with \( f_{1-q, r(B), g} \) the \( (1-q) \)-th quantile for that distribution, we get the \( 1-q \) level confidence ellipsoid for \( \Psi_i(B) \) given by
\[
\left( \bar{\Psi}_i(B) - \bar{\Psi}_i(B) \right)' \left( BB' \right)^{-1} \left( \bar{\Psi}_i(B) - \bar{\Psi}_i(B) \right) \leq r(B)f_{1-q, r(B), g} S g, i = 1, \ldots, \hat{n}.
\] (43)
Since the \( q \) level \( F \) test accepts \( H_{0,i}(B, b) \) when \( Bb \) is covered by the \( 1-q \) level confidence ellipsoid
the \( F \) test enjoys duality.

Writing \( \hat{d} \) to indicate that all possible vectors \( d \) are considered, according to the Scheffé theorem,
(Scheffé, 1959), we will have
\[
pr \left[ \hat{d} \left( d' \bar{\Psi}_i(B) - d' \bar{\Psi}_i(B) \right) \leq \sqrt{r(B)f_{1-q, r(B), g} d'B'B \frac{S}{g}} \right] = 1 - q, i = 1, \ldots, \hat{n}.
\] (44)

4 An example

Let us assume that \( \pi \) is a two factor model, where the 1st and the 2nd factors cross and have
respectively 2 and 3 levels. Associated to these factors and their interaction, we have the matrices
\[
A_1 = \left[ \begin{array}{cc} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{array} \right] \otimes \left[ \begin{array}{cc} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{array} \right]
\]
\[
A_2 = \left[ \begin{array}{cc} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} \end{array} \right] \otimes \left[ \begin{array}{ccc} -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{array} \right]
\]
\[ A_3 = \left[ -\frac{1}{\sqrt{2}} \quad \frac{1}{\sqrt{2}} \right] \otimes \begin{bmatrix} -\frac{1}{\sqrt{3}} & 0 & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} \]

obtained from deleting the first row from orthogonal matrices \( P_2 \) and \( P_3 \), since when we have balanced crossing of \( u \) factors, the matrices “\( A \)” are given by Kronecker products of \( u \) matrices one per factor (Fonseca et al., 2003). We have one such matrix for all possible factor effects and interaction. When the \( l \)-th factor, with \( a_l \) levels is not considered, the corresponding matrix will be \( \frac{1}{\sqrt{a_l}} I_{a_l} \), while when that factor is considered, that matrix will be obtained deleting the first row equal to \( \frac{1}{\sqrt{a_l}} I_{a_l}' \) from a \( a_l \times a_l \) orthogonal matrix. For instance, the second factor matrix in \( A_2 \) and \( A_3 \) is obtained in this way from

\[
\begin{bmatrix}
-\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\
\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\
\frac{1}{\sqrt{6}} & -\frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}}
\end{bmatrix}
\]

which is a \( 3 \times 3 \) orthogonal matrix.

The factors in \( \pi \) could be, for instance, the type of soil and the crop variety. For each of the 6 possible level combinations we could have a fertilization trial. Suppose now that the fertilization trials are for the macronutrients \( N \) (nitrogen), \( P \) (phosphorus) and \( K \) (potassium) each with 3 levels. So, we would have \( 3^3 \) experiments. If \( x_1, x_2 \) and \( x_3 \) are the levels of \( N, P \) and \( K \), we now would have the reduced linear functions

\[
\begin{align*}
L_1(x) &= x_1 \quad \text{- effect of } N; \\
L_2(x) &= x_2 \quad \text{- effect of } P; \\
L_3(x) &= x_1 + x_2 \quad \text{- Factorial interactions between } N \text{ and } P; \\
L_4(x) &= x_1 + 2x_2 \\
L_5(x) &= x_3 \quad \text{- effect of } K; \\
L_6(x) &= x_1 + x_3 \quad \text{- Factorial interactions between } N \text{ and } K; \\
L_7(x) &= x_1 + 2x_3 \\
L_8(x) &= x_2 + x_3 \quad \text{- Factorial interactions between } P \text{ and } K; \\
L_9(x) &= x_2 + 2x_3 \\
L_{10}(x) &= x_1 + x_2 + x_3 \\
L_{11}(x) &= x_1 + x_2 + 2x_3 \\
L_{12}(x) &= x_1 + 2x_2 + x_3 \\
L_{13}(x) &= x_1 + 2x_2 + 2x_3 \quad \text{- Factorial interactions between } N, P \text{ and } K.
\end{align*}
\]

whose values are obtained using modulus 3 arithmetic.
For each of the $L_h(x)$, $h = 1, \ldots, 13$ (= $k_3(3)$) applications, we have two degrees of freedom. In this case we have an algorithm due to Frank Yates (Montgomery, 2005), to obtain the vectors $A_h y$, $h = 1, \ldots, 13$, and the sum of squares of the errors $S$.

5 Concluding remarks

Through structured families of models, we showed how to use the algorithms for orthogonal models to study the action of the factors in the base model on the set of field trials one per treatment of the base design. Both the base design and the model for the field trials were assumed to have orthogonal structure. The field trials we considered were prime basis factorials. In this way, it was developed an analysis going beyond the usual nesting approach in which interactions between factors that nest are usually discarded.

6 Acknowledgements

This work was partially supported by CMA/FCT/UNL, under the project PEst-OE/MAT/UI0297/2011. The authors would like to thank Miguel Fonseca and the three anonymous referees for their constructive comments from which our paper has benefited greatly.

References


