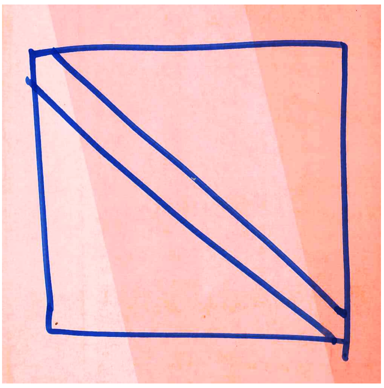
Decomposing Square Matrices



Jean-Baptiste Denis

Rapport Technique 2014-3, 34 pp.

Unité Mathématiques et Informatique Appliquées

INRA

Domaine de Vilvert

78352 Jouy-en-Josas cedex

France

Jean-Baptiste.Denis@Jouy.Inra.Fr

.

Ce travail est dédié à *Isabelle Albert*, en reconnaissance à tous ceux que nous avons menés ensemble. C'est en 2001 que nous avons rédigé une première étude, et nous partageons à l'heure actuelle pas moins de 23 cosignatures ; Isabelle est ainsi devenue mon premier coauteur sur toute ma carrière. J'arrête mais je suis persuadé que sa propre série continuera longtemps encore !

This is a resumption of a study left in 1999 due to a shift in research interests.

(July 2014)

Contents

Résumé : Dans cette étude, quelques développements matriciels sont proposés dans le but d'aboutir à la décomposition de matrices carrées, prenant en compte des effets de symétrie et d'anti-symétrie à l'aide de covariables communes aux lignes et aux colonnes de la matrice. La motivation initiale de cette recherche est la proposition de modélisations associées aux croisements diallèles.

Ce faisant, un certain nombre de résultats sur les modèles linéaires sont reformulés (en particulier l'introduction de contraintes pour lever la surparamétrisation qui rend les termes orthogonaux). La décomposition tensorielle de tables rectangulaires (régression factorielle) est rappelée dans ce cadre, comme base à la proposition.

Mots-Clef : diallèle, covariable, produit tensoriel, matrice carrée, matrice symétrique, matrice anti-symétrique, espace vectoriel, modèle linéaire, modèles avec contraintes, vectorisation de matrice.

Abstract: In this paper are proposed some matricial considerations supposed to be useful when looking for statistical decompositions of square matrices. The initial aim was the decomposition of effects in plant breeding diallel designs but the decomposition could be applied to any type of matrices, symmetrical or not, like dissimilarity tables between a series of items, or even correlation matrices. The unique requirement is that rows and columns of the matrix be in univocal correspondance.

Doing so some developpement are made about constrained linear models producing orthogonal decompositions, and tensor products of matrices. Also the factorial regression decomposition is taken as a basis for the construction of the proposals.

Key-Words: diallel, covariable, tensor product, square matrix, symmetrical matrix, antisymmetrical matrix, vector spaces, linear model, constrained models, matrix vectorization.

The initial aim of the study was to give some statistical answers suited to the analysis of diallel data when a natural structure of groups appears on the parents. Progressively, the document was developed and became a kind of recapitulation of ideas written without constraints for any future use... That explains why the notations still are those of diallel context.

It is an attempt to adapt the tensor decomposition of rectangular matrices initiated in [1] and summarized in [3] to the case of square matrices where one can take advantage of possible symmetries. It is also a complement in the linear framework of the adaptation for bilinear models to square matrices proposed in [2].

An **R** package denominated TENSO have been written, especially to numerically check the proposed forumulae, it is available under request *as it is*.

1 Motivation

In this first section are presented three basic decompositions which will be illuminating with the algebra developped in further sections. Each decomposition is briefly indicated through model terms, degrees of freedom and names for corresponding vector subspaces when retaining the orthocomplements of the preceding terms; this is done with adequate constraints on the parameters. They correspond to a particular viewpoint on the diallel table. Here f designates the female parent in rows, m designates the male parent in columns, both are varying from 1 until P. The data set is supposed complete, *i.e.* that all P by P crosses are observed (each being identified by the couple (f, m) and the associated performances denoted with τ_{fm}). $\mathbf{R}^{P^2}$ is therefore the vector space associated to the vector of observations. Each decomposition (or model) is given with four lines describing respectively: (i) the parameterization in scalar form, (ii) the parameterization in matricial form, (iii) the degrees of freedom and (iv) the orthogonal vector subspaces associated; also are added the necessary constraints to provide the orthogonality between vector subspaces. These statements are well known for the additive decomposition, they will be established further for the Hayman's decomposition and are obvious for the last decomposition.

1.1 Additive decomposition

The first decomposition is the addtive decomposition which exhibits general mean (μ), main effects of female parent (α_f), main effect of male parent (β_m) and their interaction (θ_{fm}).

$$\begin{array}{ccccccccc} \tau_{fm} & = & \mu & + & \alpha_f & + & \beta_m & + & \theta_{fm} \\ \tau & = & 1\mu1' & + & \alpha1' & + & 1\beta' & + & \theta \\ P^2 & = & 1 & + & P-1 & + & P-1 & + & (P-1)^2 \\ \mathbf{R}^{P^2} & = & \mathcal{M} & \oplus & \mathcal{A} & \oplus & \mathcal{B} & \oplus & \mathcal{C} \end{array} \quad (1)$$

with the following constraints on the parameters (the last column is the number of free constraints)

$$\begin{array}{ccccccc} \sum_u \alpha_u & = & 0 & \Leftrightarrow & 1'\alpha & = & 0 & 1 \\ \sum_u \beta_u & = & 0 & \Leftrightarrow & 1'\beta & = & 0 & 1 \\ \sum_u \theta_{um} & = & 0 & \Leftrightarrow & 1'\theta & = & 0 & P \\ \sum_u \theta_{fu} & = & 0 & \Leftrightarrow & 1'\theta' & = & 0 & P-1 \end{array} \quad (2)$$

The additive modelling is one of the basic anova models, many references are possible, a good one is [7].

Parameters are given by the following expressions of the performances of the P^2 crosses, that is the matrix τ:

$$\begin{array}{rclcrcl} \mu & = & \frac{1}{P^2}\sum_{f,m}\tau_{fm} & \Leftrightarrow & \mu & = & (1'1)^{-1}\,1'\tau 1\,(1'1)^{-1} \\ \alpha_f & = & \frac{1}{P}\sum_m(\tau_{fm}-\mu) & \Leftrightarrow & \alpha & = & \left(\mathbf{I}-1\,(1'1)^{-1}\,1'\right)\tau 1\,(1'1)^{-1} \\ \beta_m & = & \frac{1}{P}\sum_f(\tau_{fm}-\mu) & \Leftrightarrow & \beta' & = & (1'1)^{-1}\,1'\tau\left(\mathbf{I}-1\,(1'1)^{-1}\,1'\right) \\ \theta_{fm} & = & \tau_{fm}-(\mu+\alpha_f+\beta_m) & \Leftrightarrow & \theta & = & \left(\mathbf{I}-1\,(1'1)^{-1}\,1'\right)\tau\left(\mathbf{I}-1\,(1'1)^{-1}\,1'\right) \end{array} \qquad (3)$$

- μ is the average of all performances;
- α_f is the deviation of the average of performances where genotype f is used as female, from the general average (μ). It is called main effect of factor female;
- β_m is the deviation of the average of performances where genotype m is used as male, from the general average (μ). It is called main effect of factor male;
- θ_{fm} is the so-called interaction effect due to the combination of female f crossed with male m, to be added to the general average and corresponding main effects to reproduce the exact performance of the cross.

1.2 Hayman decomposition

In the Hayman decomposition [5] the symmetric role of the two parents is taken into account and for the additive and interactive parts the decomposition is parental effect + female/male effect.

$$\begin{array}{ccccccccccc} \tau_{fm} & = & \mu & + & [\pi_f+\pi_m] & + & [\lambda_f-\lambda_m] & + & \omega_{fm} & + & \phi_{fm} \\ \tau & = & 1\mu 1' & + & [\pi 1'+1\pi'] & + & [\lambda 1'-1\lambda'] & + & \omega & + & \phi \\ P^2 & = & 1 & + & P-1 & + & P-1 & + & P(P-1)/2 & + & (P-1)(P-2)/2 \\ \mathbf{R}^{P^2} & = & \mathcal{M} & \oplus & \mathcal{P} & \oplus & \mathcal{L} & \oplus & \mathcal{W} & \oplus & \mathcal{F} \end{array} \qquad (4)$$

with the following constraints onto the parameters:

$$\begin{array}{ccccc} \begin{array}{rcl} \sum_u \pi_u & = & 0 \\ \sum_u \lambda_u & = & 0 \end{array} & \begin{array}{c}\Leftrightarrow \\ \Leftrightarrow\end{array} & \begin{array}{rcl} 1'\pi & = & 0 \\ 1'\lambda & = & 0 \end{array} & & \begin{array}{c} 1 \\ 1 \end{array} \\ \left\{\begin{array}{rcl} \sum_u \omega_{um} & = & 0 \\ \omega_{fm} & = & \omega_{mf} \end{array}\right\} & \Leftrightarrow & \left\{\begin{array}{rcl} 1'\omega & = & 0 \\ \omega & = & \omega' \end{array}\right\} & & \frac{P(P+1)}{2} \\ \left\{\begin{array}{rcl} \sum_u \phi_{um} & = & 0 \\ \phi_{fm} & = & -\phi_{mf} \end{array}\right\} & \Leftrightarrow & \left\{\begin{array}{rcl} 1'\phi & = & 0 \\ \phi & = & -\phi' \end{array}\right\} & & \frac{P(P+1)}{2}+(P-1) \end{array} \qquad (5)$$

The numbers of free constraints associated to the interaction terms ω and θ are easily obtained by direct construction, for instance taking the case of $P=4$, the braced terms are redundant:

$$\omega = \left(\begin{array}{cccc} a & b & c & \{-(a+b+c)\} \\ \{b\} & d & e & \{-(b+d+e)\} \\ \{c\} & \{e\} & f & \{-(c+e+f)\} \\ \{-(a+b+c)\} & \{-(b+d+e)\} & \{-(c+e+f)\} & \{-(a+2b+2c+d++2e+f)\} \end{array}\right)$$

and

$$\phi = \begin{pmatrix} \{0\} & b & c & \{-b-c\} \\ \{-b\} & \{0\} & e & \{b-e\} \\ \{-c\} & \{-e\} & \{0\} & \{c+e\} \\ \{b+c\} & \{-b+e\} & \{-(c+e)\} & \{0\} \end{pmatrix}.$$

Parameters are given by the following expressions of the performances of the P^2 crosses:

$$\begin{array}{rcl} \mu & = & \frac{1}{P^2}\sum_{f,m}\tau_{fm} \\ \pi_p & = & \frac{1}{2P}\left(\sum_m \tau_{pm} + \sum_f \tau_{fp}\right) - \mu \\ \lambda_p & = & \frac{1}{2P}\left(\sum_m \tau_{pm} - \sum_f \tau_{fp}\right) \\ \omega_{fm} & = & \frac{1}{2}(\tau_{fm} + \tau_{fm}) - (\mu + \pi_f + \pi_m) \\ \phi_{fm} & = & \frac{1}{2}(\tau_{fm} - \tau_{fm}) - (\lambda_f - \lambda_m) \end{array} \tag{6}$$

which gives in matricial form:

$$\begin{array}{rcl} \mu & = & (1'1)^{-1}1'\tau 1(1'1)^{-1} \\ 2\pi & = & \left(\mathbf{I} - 1(1'1)^{-1}1'\right)\tau 1(1'1)^{-1} + \left(\mathbf{I} - 1(1'1)^{-1}1'\right)\tau' 1(1'1)^{-1} \\ 2\lambda & = & \left(\mathbf{I} - 1(1'1)^{-1}1'\right)\tau 1(1'1)^{-1} - \left(\mathbf{I} - 1(1'1)^{-1}1'\right)\tau' 1(1'1)^{-1} \\ 2\omega & = & \left(\mathbf{I} - 1(1'1)^{-1}1'\right)\tau\left(\mathbf{I} - 1(1'1)^{-1}1'\right) + \left(\mathbf{I} - 1(1'1)^{-1}1'\right)\tau'\left(\mathbf{I} - 1(1'1)^{-1}1'\right) \\ 2\phi & = & \left(\mathbf{I} - 1(1'1)^{-1}1'\right)\tau\left(\mathbf{I} - 1(1'1)^{-1}1'\right) - \left(\mathbf{I} - 1(1'1)^{-1}1'\right)\tau'\left(\mathbf{I} - 1(1'1)^{-1}1'\right) \end{array} \tag{7}$$

- μ, identical to the one of the additive model, is the average of all performances;
- π_p is the deviation of the average performance where genotype p is used as parent (female or male) from the general average (μ). Indeed $\pi_p = \frac{1}{2}(\alpha_p + \beta_p)$. Note that a double weight is given when p is simultaneously female and male. It is called the parental effect;
- λ_p is the average difference when genotype p is used as female or as male. It is called the maternal effect. Also $\lambda_p = \frac{1}{2}(\alpha_p - \beta_p)$;
- ω_{fm} is the symmetrical interaction of genotypes f and m, that is the complement to the additive part when it is supposed that no maternal effect exists at the interactive level, that is $(\tau_{ac} + \tau_{da} + \tau_{cb} + \tau_{bd} = \tau_{ad} + \tau_{bc} + \tau_{ca} + \tau_{db})$. It can easily be expressed from the interaction term of the additive decomposition: $2\omega_{fm} = \theta_{fm} + \theta'_{mf}$;
- ϕ_{fm} is the antisymmetrical interaction of genotypes f and m, that is the last part to add to obtain the exact performance of every cross. Also $2\phi_{fm} = \theta_{fm} - \theta'_{mf}$.

It is worth noting that

1. $\mu + \alpha_f + \beta_m = \mu + [\pi_f + \pi_m] + [\lambda_f - \lambda_m]$;
2. the model without heterosis $\tau_{fm} = \frac{1}{2}(\tau_{ff} + \tau_{mm})$ is equivalent to the following submodel of Hayman's model $\tau_{fm} = \mu + [\pi_f + \pi_m]$.

1.3 Design decomposition

Perhaps "design" is not the best term, what is meant is that information can be commonly supplied in two steps from experiments before achieving the P^2 crosses. First, all $P(P+1)/2$ different crosses when female and male roles are not distinguished, and second the $P(P-1)/2$ complementary crosses.

$$\begin{array}{ccccc} \tau_{fm} & = & \sigma_{fm} & + & \rho_{fm} \\ \tau & = & \sigma & + & \rho \\ P^2 & = & (P+1)P/2 & + & P(P-1)/2 \\ \mathbf{R}^{P^2} & = & \mathcal{S} & \oplus & \mathcal{R} \end{array} \tag{8}$$

with the following constraints onto the parameters:

$$\begin{array}{ccccccc} \sigma_{fm} & = & \sigma_{mf} & \Leftrightarrow & \sigma & = & \sigma' & \frac{P(P-1)}{2} \\ \rho_{fm} & = & -\rho_{mf} & \Leftrightarrow & \rho & = & -\rho' & \frac{(P+1)P}{2} \end{array} \tag{9}$$

Parameters are given by the following expressions of the performances of the P^2 crosses:

$$\begin{array}{ccccccc} \sigma_{fm} & = & \frac{1}{2}(\tau_{fm}+\tau_{mf}) & \Leftrightarrow & \sigma & = & \frac{1}{2}(\tau+\tau') \\ \rho_{fm} & = & \frac{1}{2}(\tau_{fm}-\tau_{mf}) & \Leftrightarrow & \rho & = & \frac{1}{2}(\tau-\tau') \end{array} \tag{10}$$

- σ_{pp} is the performance of genotype p crossed with itself, it is called the own value of genotype p. In some crops (like maize) these performances are at a very low level with respect to performance of true hybrids (heterosis effect).
- σ_{fm} (when $f \neq m$ *i.e.* in case of hybrids) is the average of the two reciprocal crosses with parent f and m;
- ρ_{fm} (when $f \neq m$ *i.e.* in case of hybrids) is the half of difference of the two reciprocal crosses with parent f and m; notice that the constraints on this part of the model imply that $\rho_{pp} = 0$.

Remark For interpretation purpose, it may of interest to further extract the intrinsic effect of the genotypes from the symmetrical terms, that is something like:

$$\begin{array}{ccccccc} \tau_{fm} & = & \delta_f 1_{[f=m]} & + & \sigma_{fm} 1_{[f\neq m]} & + & \rho_{fm} 1_{[f\neq m]} \\ P^2 & = & P & + & P(P-1)/2 & + & P(P-1)/2 \end{array} \tag{11}$$

where the diagonal terms of matrix σ have been vanished and δ_p is the own effect of genotype p.

2 Fixed Linear Models

2.1 Full rank case

2.1.1 Definition

When the parameterization of a fixed linear models is defined without additional constraints, standard formulae are straightforward.

$$E[Y] = \mathbf{X}\theta \quad ; \qquad Var[Y] = \sigma^2 \mathbf{I}_N \tag{12}$$

where Y is the N-vector of observations, $\mathbf{X}$ is the $N \times p$ design-model matrix supposed to be full rank by columns, and θ is the p-vector of unknown parameters.

2.1.2 Properties

The θ vector can be expressed in a unique way from the expectation vector:

$$\theta = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' E[Y] \tag{13}$$

2.1.3 Estimation

The unique LS estimators of the parameters are given by

$$\widehat{\theta} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' Y. \tag{14}$$

They are obviously unbiased and their variance-covariance matrix is

$$Var\left(\widehat{\theta}\right) = \sigma^2 (\mathbf{X}'\mathbf{X})^{-1}. \tag{15}$$

2.2 Overparameterized case

When the rank of $\mathbf{X}$ is less than its column number, say $q < p$, to be defined, the parameterization of a fixed linear model needs additional linear constraints. This is often the case for interpretation purpose of the parameters. In that case, the model is overparameterized, nevertheless standard formulae are easily extended.

2.2.1 Definition

With similar notations introduced in the previous section and a new matrix $\mathbf{C}$ (like constraints) of size $(p-q) \times p$, let the model defined by

$$\left\{ \begin{array}{lcc} E[Y] & = & \mathbf{X}\theta \\ Var[Y] & = & \sigma^2 \mathbf{I}_N \\ rk(\mathbf{X}) & = & q < p \\ rk\begin{pmatrix} \mathbf{X} \\ \mathbf{C} \end{pmatrix} & = & p \\ \mathbf{C}\theta & = & \mathbf{0}_{q,1} \; . \end{array} \right. \tag{16}$$

This implies that matrix $\mathbf{C}$ is full rank by rows ; this is often the case but additional consistent constraints could be added as well[1].

[1] More precisely $\mathbf{C} \longmapsto \begin{pmatrix} \mathbf{C} \\ \mathbf{MC} \end{pmatrix}$ where $\mathbf{M}$ is any matrix.

2.2.2 Properties

It can be shown that the constraint on the parameter vector can be equivalently expressed:

$$\begin{aligned}
\mathbf{C}\theta &= \mathbf{0} \\
&\Longleftrightarrow \\
\theta &= \left(\mathbf{X}'\mathbf{X} + \mathbf{C}'\mathbf{C}\right)^{-1} \left(\mathbf{X}'\mathbf{X}\right) \theta \ .
\end{aligned} \tag{17}$$

An interesting symmetrical relationship between the two matrices $\mathbf{X}$ and $\mathbf{C}$ is:

$$\mathbf{C} \left(\mathbf{X}'\mathbf{X} + \mathbf{C}'\mathbf{C}\right)^{-1} \mathbf{X}' = \mathbf{0} \tag{18}$$

rows of both matrices are orthogonal for the positive matrix $\left(\mathbf{X}'\mathbf{X} + \mathbf{C}'\mathbf{C}\right)^{-1}$.

Expression (13) can be generalized as:

$$\theta = \left(\mathbf{X}'\mathbf{X} + \mathbf{C}'\mathbf{C}\right)^{-1} \mathbf{X}' E\left[Y\right] \tag{19}$$

2.2.3 Estimation

The least-square estimator is

$$\widehat{\theta} = \left(\mathbf{X}'\mathbf{X} + \mathbf{C}'\mathbf{C}\right)^{-1} \mathbf{X}'Y. \tag{20}$$

It is unbiased and its variance matrix is

$$Var\left[\widehat{\theta}\right] = \sigma^2 \left(\mathbf{X}'\mathbf{X} + \mathbf{C}'\mathbf{C}\right)^{-1} \left(\mathbf{X}'\mathbf{X}\right) \left(\mathbf{X}'\mathbf{X} + \mathbf{C}'\mathbf{C}\right)^{-1}. \tag{21}$$

As a consequence of the unbiasedness and the linear constraints on θ, it is consistent to check that :

$$\begin{aligned}
Var\left[\mathbf{C}\widehat{\theta}\right] &= \mathbf{C} \left(\mathbf{X}'\mathbf{X} + \mathbf{C}'\mathbf{C}\right)^{-1} \left(\mathbf{X}'\mathbf{X}\right) \left(\mathbf{X}'\mathbf{X} + \mathbf{C}'\mathbf{C}\right)^{-1} \mathbf{C}' \\
&= \mathbf{0}_{p-q,p-q}
\end{aligned}$$

2.2.4 Basic matrices

To characterize some specific overparameterized linear models, we will try to explicit the following matrices:

- the design-model matrix: $\mathbf{X}$,
- the constraint matrix: $\mathbf{C}$,
- the $\mathbf{X}'\mathbf{X}$ matrix,
- the $\mathbf{C}'\mathbf{C}$ matrix,
- the estimator matrix: $EST = \left(\mathbf{X}'\mathbf{X} + \mathbf{C}'\mathbf{C}\right)^{-1} \mathbf{X}'$ and
- the variance structure matrix of the estimators:

$$VAR = (EST)\,(EST)' = \left(\mathbf{X}'\mathbf{X} + \mathbf{C}'\mathbf{C}\right)^{-1} \left(\mathbf{X}'\mathbf{X}\right) \left(\mathbf{X}'\mathbf{X} + \mathbf{C}'\mathbf{C}\right)^{-1} \ .$$

2.3 Nested case

Sometimes, the constrained models introduced in §2.2 are generated from the nesting of two effects which split the design matrix into two blocks. The interest of the proposed constraints is to obtain an orthogonal decomposition of the expections according to the two effects. So we denominate them *orthogonalizing constraints*.

2.3.1 Definition

Model (16) is precised by the nesting structure in the following way:

$$\begin{cases} \mathbf{X} = \begin{pmatrix} \mathbf{X}_1 & \mathbf{X}_2 \end{pmatrix} \quad \text{and} \quad \theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \\ \{\mathbf{X}_1\} \subseteq \{\mathbf{X}_2\} \quad \Longleftrightarrow \quad \mathbf{X}_2 \left(\mathbf{X}_2'\mathbf{X}_2\right)^{-1} \mathbf{X}_2'\mathbf{X}_1 = \mathbf{X}_1 \\ \quad \text{and} \\ \mathbf{C} = \begin{pmatrix} \mathbf{0}_{p_1,p_1} & \mathbf{X}_1'\mathbf{X}_2 \end{pmatrix} . \end{cases} \tag{22}$$

where $\mathbf{X}_1$ and $\mathbf{X}_2$ are the $N \times p_1$ and $N \times p_2$ design-model matrices, θ_1 and θ_2 are p_1 and p_2-vectors of unknown parameters. The fact that $\{\mathbf{X}_1\} \subseteq \{\mathbf{X}_2\}$ be satisfied justifies the denomination of nested case since the two models $E[Y] = \mathbf{X}_1\theta_1$ and $E[Y] = \mathbf{X}_2\theta_2$ are nested. As it will be see below, the interest of such a constraint is to provide to θ_1 the same meaning when estimated in the submodel or in the complete model. We will also suppose that

$$\begin{aligned} rk\,(\mathbf{X}_1) &= p_1 \\ rk\,(\mathbf{X}_2) &= p_2 \end{aligned}$$

which implie that

$$rk \begin{pmatrix} \mathbf{X}_1 & \mathbf{X}_2 \end{pmatrix} = p_2 \ .$$

A common example of such a setup is the one-way balanced ANOVA model $E[Y_{ij}] = \mu + \alpha_i$ where $\mathbf{X}_1 = \mathbf{1}_J \otimes \mathbf{1}_I$ and $\mathbf{X}_2 = \mathbf{1}_J \otimes \mathbf{I}_I$.

2.3.2 Properties

An interesting consequence of the proposed constraint is that the two components of the expectation are orthogonal ensuring an independent interpretation of each.

$$(\mathbf{X}_1\theta_1)'\,(\mathbf{X}_2\theta_2) \;=\; \theta_1'\mathbf{X}_1'\mathbf{X}_2\theta_2 = 0.$$

Expression (19) turns to be:

$$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} \mathbf{X}_1'\mathbf{X}_1 & \mathbf{X}_1'\mathbf{X}_2 \\ \mathbf{X}_2'\mathbf{X}_1 & \mathbf{X}_2'\left(\mathbf{I}_N + \mathbf{X}_1\mathbf{X}_1'\right)\mathbf{X}_2 \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{X}_1' \\ \mathbf{X}_2' \end{pmatrix} E[Y]. \tag{23}$$

From now, we will use the notation:

$$\mathbf{S}_{ij} \;=\; \mathbf{X}_i'\mathbf{X}_j$$

and Formula reads (23)

$$\begin{aligned} \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} &= \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} + \mathbf{S}_{21}\mathbf{S}_{12} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{X}_1' \\ \mathbf{X}_2' \end{pmatrix} E[Y] \\ &= \begin{pmatrix} \mathbf{S}_{11}^{-1}\mathbf{X}_1' \\ \left(\mathbf{S}_{22}^{-1} - \mathbf{S}_{22}^{-1}\mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{S}_{12}\mathbf{S}_{22}^{-1}\right)\mathbf{X}_2' \end{pmatrix} E[Y] \end{aligned} \tag{24}$$

From that expression, one can notice that

$$\theta_1 = \mathbf{S}_{11}^{-1}\mathbf{X}_1 E\left[Y\right]$$

that is that the used constraint leaves to θ_1 the same meaning in the complete model (22) that it have got in the submodel with only the first effect.

2.3.3 Estimation

Applying formulae from the general case (§2.2), the estimators are:

$$\begin{pmatrix} \widehat{\theta_1} \\ \widehat{\theta_2} \end{pmatrix} = \begin{pmatrix} \mathbf{S}_{11}^{-1}\mathbf{X}_1' \\ \left(\mathbf{S}_{22}^{-1} - \mathbf{S}_{22}^{-1}\mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{S}_{12}\mathbf{S}_{22}^{-1}\right)\mathbf{X}_2' \end{pmatrix} Y. \tag{25}$$

The estimators are still unbiased and their variance-covariance matrix is:

$$\begin{aligned} Var\begin{pmatrix} \widehat{\theta_1} \\ \widehat{\theta_2} \end{pmatrix} &= \sigma^2 \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} + \mathbf{S}_{21}\mathbf{S}_{12} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} + \mathbf{S}_{21}\mathbf{S}_{12} \end{pmatrix}^{-1} \\ &= \sigma^2 \begin{pmatrix} \mathbf{S}_{11}^{-1} & \mathbf{0}_{p_1,p_2} \\ \mathbf{0}_{p_2,p_1} & \mathbf{S}_{22}^{-1} - \mathbf{S}_{22}^{-1}\mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{S}_{12}\mathbf{S}_{22}^{-1} \end{pmatrix}. \end{aligned} \tag{26}$$

another path for this kind of derivation is proposed in §2.3.5.

2.3.4 Basic matrices

$$\begin{aligned} \mathbf{X} &= \begin{pmatrix} \mathbf{X}_1, & \mathbf{X}_2 \end{pmatrix} \\ \mathbf{C} &= \begin{pmatrix} \mathbf{0}_{p_1,p_1} & \mathbf{X}_1'\mathbf{X}_2 \end{pmatrix} \\ \mathbf{X}'\mathbf{X} &= \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{pmatrix} \\ \mathbf{C}'\mathbf{C} &= \begin{pmatrix} \mathbf{0}_{p_1,p_1} & \mathbf{0}_{p_1,p_2} \\ \mathbf{0}_{p_2,p_1} & \mathbf{S}_{21}\mathbf{S}_{12} \end{pmatrix} \\ EST &= \begin{pmatrix} \mathbf{S}_{11}^{-1}\mathbf{X}_1' \\ \left(\mathbf{S}_{22}^{-1} - \mathbf{S}_{22}^{-1}\mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{S}_{12}\mathbf{S}_{22}^{-1}\right)\mathbf{X}_2' \end{pmatrix} \\ VAR &= \begin{pmatrix} \mathbf{S}_{11}^{-1} & \mathbf{0}_{p_1,p_2} \\ \mathbf{0}_{p_2,p_1} & \mathbf{S}_{22}^{-1} - \mathbf{S}_{22}^{-1}\mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{S}_{12}\mathbf{S}_{22}^{-1} \end{pmatrix} \end{aligned}$$

2.3.5 Orthonormalized parameterization

The previous exhibited properties can, perhaps, be better understood and easily obtained using a new parameterization for Model (22). Let us consider a consistent orthonormalization[2] of matrices $\mathbf{X}_1$ and $\mathbf{X}_2$:

$$\begin{aligned} \mathbf{X}_1 &= \mathbf{U}_1\mathbf{P}_1 \\ \mathbf{X}_2 &= (\mathbf{U}_1, \mathbf{U}_2)\begin{pmatrix} \mathbf{P}_{12} \\ \mathbf{P}_{22} \end{pmatrix} \end{aligned} \tag{27}$$

[2]Of course, it is not unique but this doesn't matter.

where $\mathbf{U}_1$ is a $p_1 \times N$ matrix and $\mathbf{U}_2$ is a $(p_2 - p_1) \times N$ matrix such[3] that $(\mathbf{U}_1, \mathbf{U}_2)' (\mathbf{U}_1, \mathbf{U}_2) = \mathbf{I}_{p_2}$. This implies that $\mathbf{P}_1$ is a square invertible matrix, as well is $\mathbf{P}_2 = \begin{pmatrix} \mathbf{P}_{12} \\ \mathbf{P}_{22} \end{pmatrix}$.

We can then define the new parameter vector $\rho' = (\rho_1', \rho_2')$ by:

$$\begin{aligned} E[Y] &= \mathbf{X}_1\theta_1 + \mathbf{X}_2\theta_2 \\ &= \mathbf{U}_1(\mathbf{P}_1\theta_1) + (\mathbf{U}_1, \mathbf{U}_2)\begin{pmatrix} \mathbf{P}_{12} \\ \mathbf{P}_{22} \end{pmatrix}\theta_2 \\ &= \mathbf{U}_1(\mathbf{P}_1\theta_1) + (\mathbf{U}_1, \mathbf{U}_2)\begin{pmatrix} \mathbf{P}_{12}\theta_2 \\ \mathbf{P}_{22}\theta_2 \end{pmatrix} \\ &= \mathbf{U}_1\rho_1 + (\mathbf{U}_1, \mathbf{U}_2)\begin{pmatrix} \rho_{12} \\ \rho_{22} \end{pmatrix} \\ &= \mathbf{U}_1\rho_1 + (\mathbf{U}_1, \mathbf{U}_2)\rho_2 \ . \end{aligned}$$

therefore matrices $\mathbf{U}_1$ and $(\mathbf{U}_1, \mathbf{U}_2)$ are substituted to $\mathbf{X}_1$ and $\mathbf{X}_2$.

Now, the constraint is just

$$\begin{pmatrix} \mathbf{0}_{p_1,p_1} & \mathbf{I}_{p_1,p_1} & \mathbf{0}_{p_1,p_2-p_1} \end{pmatrix} \begin{pmatrix} \rho_1 \\ \rho_{12} \\ \rho_{22} \end{pmatrix} = 0$$

that is the nullity of the component ρ_{12}.

One can shift from one parameterization to the other one with the simple formulae:

$$\begin{array}{ccc} \theta_1 = \mathbf{P}_1^{-1}\rho_1 & ; & \theta_2 = \mathbf{P}_2^{-1}\rho_2 \\ \rho_1 = \mathbf{P}_1\theta_1 & ; & \rho_2 = \mathbf{P}_2\theta_2 \ . \end{array}$$

Notice also that

$$\mathbf{S}_{11} = \mathbf{P}_1'\mathbf{P}_1 \quad ; \quad \mathbf{S}_{12} = \mathbf{P}_1'\mathbf{P}_{12} \quad ; \quad \mathbf{S}_{22} = \mathbf{P}_2'\mathbf{P}_2 \ .$$

In the same way, the basic matrices simplify a lot:

$$\begin{aligned} \mathbf{X}'\mathbf{X} &= \begin{pmatrix} \mathbf{I}_{p_1} & \mathbf{I}_{p_1} & \mathbf{0} \\ \mathbf{I}_{p_1} & \mathbf{I}_{p_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{p_2} \end{pmatrix} \\ \mathbf{C}'\mathbf{C} &= \begin{pmatrix} \mathbf{0}_{p_1,p_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{p_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0}_{p_2,p_2} \end{pmatrix} \\ (\mathbf{X}'\mathbf{X} + \mathbf{C}'\mathbf{C})^{-1} &= \begin{pmatrix} 2\mathbf{I}_{p_1} & -\mathbf{I}_{p_1} & \mathbf{0} \\ -\mathbf{I}_{p_1} & \mathbf{I}_{p_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{p_2} \end{pmatrix} \\ EST &= \begin{pmatrix} \mathbf{U}_1' \\ \mathbf{0} \\ \mathbf{U}_2' \end{pmatrix} \\ VAR &= \begin{pmatrix} \mathbf{I}_{p_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{p_2} \end{pmatrix} \end{aligned}$$

Notice the last expression which is, up to σ^2, the variance of the LS estimator; as expected a zero variance is attributed to ρ_{12} which is null.

[3] This is possible, for instance with a Grand-Schmit procedure on $(\mathbf{X}_1, \mathbf{X}_2)$ followed by the elimination of zero columns.

2.4 Twice nested case

Results of the previous section can be extended with several levels in the hierarchy of he modelling. We will briefly consider the twice nested case.

2.4.1 Definition

The model comprises three embedded $\mathbf{X}$ matrices and three associated components of the vector of parameters

$$\left\{\begin{array}{l} \mathbf{X} = \left(\begin{array}{ccc} \mathbf{X}_1 & \mathbf{X}_2 & \mathbf{X}_3 \end{array}\right) \quad \text{and} \quad \theta = \left(\begin{array}{c} \theta_1 \\ \theta_2 \\ \theta_3 \end{array}\right) \\ \{\mathbf{X}_1\} \subseteq \{\mathbf{X}_2\} \subseteq \{\mathbf{X}_3\} \\ \quad \text{and} \\ \mathbf{C} = \left(\begin{array}{ccc} \mathbf{0}_{p_1,p_1} & \mathbf{X}_1'\mathbf{X}_2 & \mathbf{0}_{p_1,p_3} \\ \mathbf{0}_{p_2,p_1} & \mathbf{0}_{p_2,p_2} & \mathbf{X}_2'\mathbf{X}_3 \end{array}\right) . \end{array}\right. \tag{28}$$

with adapted dimensions, θ_1, θ_2 and θ_3 are p_1, p_2 and p_3-vectors of the parameters. Also

$$\begin{array}{rcl} rk\,(\mathbf{X}_1) & = & p_1 \\ rk\,(\mathbf{X}_2) & = & p_2 \\ rk\,(\mathbf{X}_3) & = & p_3 \end{array}$$

which implie that

$$\begin{array}{rcl} rk\left(\begin{array}{cc} \mathbf{X}_1 & \mathbf{X}_2 \end{array}\right) & = & p_2 \\ rk\left(\begin{array}{cc} \mathbf{X}_2 & \mathbf{X}_3 \end{array}\right) & = & p_3 \\ rk\left(\begin{array}{cc} \mathbf{X}_1 & \mathbf{X}_3 \end{array}\right) & = & p_3 \\ rk\left(\begin{array}{ccc} \mathbf{X}_1 & \mathbf{X}_2 & \mathbf{X}_3 \end{array}\right) & = & p_3. \end{array}$$

2.4.2 Basic matrices

$$\begin{array}{rcl} \mathbf{X} & = & \left(\mathbf{X}_1,\ \mathbf{X}_2,\ \mathbf{X}_3\right) \\ \mathbf{C} & = & \left(\begin{array}{ccc} \mathbf{0}_{p_1,p_1} & \mathbf{X}_1'\mathbf{X}_2 & \mathbf{0}_{p_1,p_3} \\ \mathbf{0}_{p_2,p_1} & \mathbf{0}_{p_2,p_2} & \mathbf{X}_2'\mathbf{X}_3 \end{array}\right) \\ \mathbf{X}'\mathbf{X} & = & \left(\begin{array}{ccc} \mathbf{S}_{11} & \mathbf{S}_{12} & \mathbf{S}_{13} \\ \mathbf{S}_{21} & \mathbf{S}_{22} & \mathbf{S}_{23} \\ \mathbf{S}_{31} & \mathbf{S}_{32} & \mathbf{S}_{33} \end{array}\right) \\ \mathbf{C}'\mathbf{C} & = & \left(\begin{array}{ccc} \mathbf{0}_{p_1,p_1} & \mathbf{0}_{p_1,p_2} & \mathbf{0}_{p_1,p_3} \\ \mathbf{0}_{p_2,p_1} & \mathbf{S}_{21}\mathbf{S}_{12} & \mathbf{0}_{p_2,p_3} \\ \mathbf{0}_{p_3,p_1} & \mathbf{0}_{p_3,p_2} & \mathbf{S}_{32}\mathbf{S}_{23} \end{array}\right) \\ EST & = & \left(\begin{array}{c} \mathbf{S}_{11}^{-1}\mathbf{X}_1' \\ \left(\mathbf{S}_{22}^{-1} - \mathbf{S}_{22}^{-1}\mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{S}_{12}\mathbf{S}_{22}^{-1}\right)\mathbf{X}_2' \\ \left(\mathbf{S}_{33}^{-1} - \mathbf{S}_{33}^{-1}\mathbf{S}_{32}\mathbf{S}_{22}^{-1}\mathbf{S}_{23}\mathbf{S}_{33}^{-1}\right)\mathbf{X}_3' \end{array}\right) \\ VAR & = & \left(\begin{array}{ccc} \mathbf{S}_{11}^{-1} & \mathbf{0}_{p_1,p_2} & \mathbf{0}_{p_1,p_3} \\ \mathbf{0}_{p_2,p_1} & \mathbf{S}_{22}^{-1} - \mathbf{S}_{22}^{-1}\mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{S}_{12}\mathbf{S}_{22}^{-1} & \mathbf{0}_{p_2,p_3} \\ \mathbf{0}_{p_3,p_1} & \mathbf{0}_{p_3,p_2} & \mathbf{S}_{33}^{-1} - \mathbf{S}_{33}^{-1}\mathbf{S}_{32}\mathbf{S}_{22}^{-1}\mathbf{S}_{23}\mathbf{S}_{33}^{-1} \end{array}\right) \end{array}$$

2.5 Crossed nested case

With some modifications, to give to the second and third effects symmetrical roles, the twice nested case can become a crossed case. Let see how.

2.5.1 Definition

$$\left\{\begin{array}{l} E\left[Y\right]=\left(\begin{array}{ccc}\mathbf{X}_1 & \mathbf{X}_2 & \mathbf{X}_3\end{array}\right)\left(\begin{array}{c}\theta_1\\ \theta_2\\ \theta_3\end{array}\right) \quad ; \qquad Var\left[Y\right]=\sigma^2\mathbf{I}_N \\ \{\mathbf{X}_1\}\subseteq\{\mathbf{X}_2\} \text{ and } \{\mathbf{X}_1\}\subseteq\{\mathbf{X}_3\} \\ \quad\text{and} \\ \left(\begin{array}{ccc}\mathbf{0}_{p_1,p_1} & \mathbf{X}_1'\mathbf{X}_2 & \mathbf{0}_{p_1,p_3}\\ \mathbf{0}_{p_1,p_1} & \mathbf{0}_{p_1,p_2} & \mathbf{X}_1'\mathbf{X}_3\end{array}\right)\left(\begin{array}{c}\theta_1\\ \theta_2\\ \theta_3\end{array}\right)=\mathbf{0}_{2p_1,1}\end{array}\right. \tag{29}$$

with similar notation of Model (22). The fact that $\{\mathbf{X}_1\}\subseteq\{\mathbf{X}_2\}$ and $\{\mathbf{X}_1\}\subseteq\{\mathbf{X}_3\}$ justifies the denomination of crossed nested case since the two nesting starts from the same nucleus of $E\left[Y\right]=\mathbf{X}_1\theta_1$. Again, we will also suppose that

$$\begin{array}{rcl} rk\left(\mathbf{X}_1\right) & = & p_1 \\ rk\left(\mathbf{X}_2\right) & = & p_2 \\ rk\left(\mathbf{X}_3\right) & = & p_3 \end{array}$$

and additionally

$$\{\mathbf{X}_2\}\cap\{\mathbf{X}_3\}=\{\mathbf{X}_1\}$$

which implie that

$$rk\left(\begin{array}{ccc}\mathbf{X}_1 & \mathbf{X}_2 & \mathbf{X}_3\end{array}\right)=rk\left(\begin{array}{cc}\mathbf{X}_2 & \mathbf{X}_3\end{array}\right)=p_2+p_3-p_1 \ .$$

A common example of such a setup is the balanced additive two-way ANOVA model $E\left[Y_{ij}\right]=\mu+\alpha_i+\beta_j$ where $\mathbf{X}_1=\mathbf{1}_J\otimes\mathbf{1}_I$, $\mathbf{X}_2=\mathbf{1}_J\otimes\mathbf{I}_I$ and $\mathbf{X}_3=\mathbf{I}_J\otimes\mathbf{1}_I$.

2.5.2 Basic matrices

$$\begin{array}{rcl}
\mathbf{X} & = & \left(\mathbf{X}_1,\ \ \mathbf{X}_2,\ \mathbf{X}_3\right) \\
\mathbf{C} & = & \left(\begin{array}{ccc}\mathbf{0}_{p_1,p_1} & \mathbf{X}_1'\mathbf{X}_2 & \mathbf{0}_{p_1,p_3}\\ \mathbf{0}_{p_1,p_1} & \mathbf{0}_{p_1,p_2} & \mathbf{X}_1'\mathbf{X}_3\end{array}\right) \\
\mathbf{X}'\mathbf{X} & = & \left(\begin{array}{ccc}\mathbf{S}_{11} & \mathbf{S}_{12} & \mathbf{S}_{13}\\ \mathbf{S}_{21} & \mathbf{S}_{22} & \mathbf{S}_{23}\\ \mathbf{S}_{31} & \mathbf{S}_{32} & \mathbf{S}_{33}\end{array}\right) \\
\mathbf{C}'\mathbf{C} & = & \left(\begin{array}{ccc}\mathbf{0}_{p_1,p_1} & \mathbf{0}_{p_1,p_2} & \mathbf{0}_{p_1,p_3}\\ \mathbf{0}_{p_2,p_1} & \mathbf{S}_{21}\mathbf{S}_{12} & \mathbf{0}_{p_2,p_3}\\ \mathbf{0}_{p_3,p_1} & \mathbf{0}_{p_3,p_2} & \mathbf{S}_{31}\mathbf{S}_{13}\end{array}\right) \\
EST & = & \left(\begin{array}{c}\mathbf{S}_{11}^{-1}\mathbf{X}_1'\\ \left(\mathbf{S}_{22}^{-1}-\mathbf{S}_{22}^{-1}\mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{S}_{12}\mathbf{S}_{22}^{-1}\right)\mathbf{X}_2'\\ \left(\mathbf{S}_{33}^{-1}-\mathbf{S}_{33}^{-1}\mathbf{S}_{31}\mathbf{S}_{11}^{-1}\mathbf{S}_{13}\mathbf{S}_{33}^{-1}\right)\mathbf{X}_3'\end{array}\right) \\
VAR & = & \left(\begin{array}{ccc}\mathbf{S}_{11}^{-1} & \mathbf{0}_{p_1,p_2} & \mathbf{0}_{p_1,p_3}\\ \mathbf{0}_{p_2,p_1} & \mathbf{S}_{22}^{-1}-\mathbf{S}_{22}^{-1}\mathbf{S}_{21}\mathbf{S}_{11}^{-1}\mathbf{S}_{12}\mathbf{S}_{22}^{-1} & \mathbf{0}_{p_2,p_3}\\ \mathbf{0}_{p_3,p_1} & \mathbf{0}_{p_3,p_2} & \mathbf{S}_{33}^{-1}-\mathbf{S}_{33}^{-1}\mathbf{S}_{31}\mathbf{S}_{11}^{-1}\mathbf{S}_{13}\mathbf{S}_{33}^{-1}\end{array}\right)
\end{array}$$

2.6 Factorial regression

Factorial Regression was defined in [1], it is based on the tensor decomposition of a rectangular matrix. We will base the decomposition of square matrices on it, the reason why, it is included here.

2.6.1 Definition

We will give three equivalent definitions using scalar, matricial and vectorized presentations. Let a matrix $\mathbf{Y}$ of size $I \times J$, a matrix $\mathbf{W}$ of size $I \times K$ and a matrix $\mathbf{Z}$ of size $J \times H$; the model comprises three sets of parameters, also presented in matrix form: μ of size $K \times H$, α of size $I \times H$ and β of size $J \times K$.

Scalar form

$$E\left[Y_{ij}\right] = \sum_{kh} W_{ik}\mu_{kh}Z_{jh} + \sum_{h} \alpha_{ih}Z_{jh} + \sum_{k} W_{ik}\beta_{jk} \ .$$

Matricial form

$$E\left[\mathbf{Y}\right] = \mathbf{W}\mu\mathbf{Z}' + \alpha\mathbf{Z}' + \mathbf{W}\beta' \ .$$

Vectorized form

$$E\left[vec\left(\mathbf{Y}\right)\right] = \left(\mathbf{Z}\otimes\mathbf{W}\right)vec\left(\mu\right) + \left(\mathbf{Z}\otimes\mathbf{I}_I\right)vec\left(\alpha\right) + \left(\mathbf{I}_J\otimes\mathbf{W}\right)vec\left(\beta'\right) \ . \tag{30}$$

Easy to check that it is a crossed nested case presented in §2.5.

2.6.2 Basic matrices

The vectorized form is the most convenient to provide the basic results of this model since we can applied the general formulae of §2.5, finding for this specific model more simplifications. Here are the basic matrices:

$$\begin{aligned}
\mathbf{X} &= \left(\mathbf{Z}\otimes\mathbf{W}\ ,\ \mathbf{Z}\otimes\mathbf{I}_I\ ,\ \mathbf{I}_J\otimes\mathbf{W}\right)\ ,\\
\mathbf{C} &= \begin{pmatrix} \mathbf{0}_{KH,KH} & \mathbf{I}_H\otimes\mathbf{W}' & \mathbf{0}_{KH,KJ} \\ \mathbf{0}_{KH,KH} & \mathbf{0}_{KH,IH} & \mathbf{I}_K\otimes\mathbf{Z}' \end{pmatrix}\ ,\\
\mathbf{X}'\mathbf{X} &= \begin{pmatrix} \mathbf{Z}'\mathbf{Z}\otimes\mathbf{W}'\mathbf{W} & \mathbf{Z}'\mathbf{Z}\otimes\mathbf{W}' & \mathbf{Z}'\otimes\mathbf{W}'\mathbf{W} \\ \mathbf{Z}'\mathbf{Z}\otimes\mathbf{W}' & \mathbf{Z}'\mathbf{Z}\otimes\mathbf{I}_I & \mathbf{Z}'\otimes\mathbf{W} \\ \mathbf{Z}\otimes\mathbf{W}'\mathbf{W} & \mathbf{Z}\otimes\mathbf{W}' & \mathbf{I}_J\otimes\mathbf{W}'\mathbf{W} \end{pmatrix}\ ,\\
\mathbf{C}'\mathbf{C} &= \begin{pmatrix} \mathbf{0}_{KH,KH} & \mathbf{0}_{KH,IH} & \mathbf{0}_{KH,KJ} \\ \mathbf{0}_{IH,KH} & \mathbf{I}_H\otimes\mathbf{W}\mathbf{W}' & \mathbf{0}_{IH,KJ} \\ \mathbf{0}_{KJ,KH} & \mathbf{0}_{KJ,IH} & \mathbf{Z}\mathbf{Z}'\otimes\mathbf{I}_K \end{pmatrix}\ ,\\
EST &= \begin{pmatrix} \left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\otimes\left(\mathbf{W}'\mathbf{W}\right)^{-1}\mathbf{W}' \\ \left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\mathbf{Z}'\otimes\mathbf{P}_{\{\mathbf{W}\}^\perp} \\ \mathbf{P}_{\{\mathbf{Z}\}^\perp}\otimes\left(\mathbf{W}'\mathbf{W}\right)^{-1}\mathbf{W}' \end{pmatrix}\ ,\\
VAR &= \begin{pmatrix} \left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\otimes\left(\mathbf{W}'\mathbf{W}\right)^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \left(\mathbf{Z}'\mathbf{Z}\right)^{-1}\otimes\mathbf{P}_{\{\mathbf{W}\}^\perp} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{P}_{\{\mathbf{Z}\}^\perp}\otimes\left(\mathbf{W}'\mathbf{W}\right)^{-1} \end{pmatrix}
\end{aligned}$$

where $\mathbf{P}_{\{\mathbf{U}\}^{\perp}} = \left(\mathbf{I}_I - \mathbf{U}\left(\mathbf{U}'\mathbf{U}\right)^{-1}\mathbf{U}'\right)$.

2.6.3 Generalization

A complete decomposition into IJ matrices of rank one each, can be easily obtained. Let matrices $\mathbf{W}$ and $\mathbf{Z}$ be orthonormal matrices, denoting by $\{\mathbf{w}_k\}$ and $\{\mathbf{z}_h\}$ their columns one can write that

$$\mathbf{Y} = \sum_{k=1}^{I}\sum_{h=1}^{J} \mathbf{w}_k \mu_{kh} \mathbf{z}'_h$$

and check that the IJ matrices, $\mathbf{w}_k\mathbf{z}'_h$, are orthogonal.

3 Useful Matrices

3.1 Commutation matrices to transpose vectorized matrices

For convenience, following the standard linear model notation, matrices of data, estimators... will be dealt as vector using the "*vec*" operator defined in §A.3. That is why, we need the linear operator, described here as a matrix, transposing such a vectorized matrix. For instance, let us consider a matrix $\mathbf{M}$ of size 2×3, we want to transform

$$vec\,(\mathbf{M}) = \begin{pmatrix} M_{11} \\ M_{21} \\ M_{12} \\ M_{22} \\ M_{13} \\ M_{23} \end{pmatrix} \text{ into } vec\,(\mathbf{M}') = \begin{pmatrix} M_{11} \\ M_{12} \\ M_{13} \\ M_{21} \\ M_{22} \\ M_{23} \end{pmatrix}.$$

This is done by a commutation matrix (0 everywhere except one 1 in each row and each column)[4] of size 6 :

$$\mathbf{T}_{2,3} = \begin{pmatrix} 1 & . & . & . & . & . \\ . & . & 1 & . & . & . \\ . & . & . & . & 1 & . \\ \\ . & 1 & . & . & . & . \\ . & . & . & 1 & . & . \\ . & . & . & . & . & 1 \end{pmatrix}$$

indeed $vec\,(\mathbf{M}') = \mathbf{T}_{2,3} vec\,(\mathbf{M})$.

Definition Let $\mathbf{T}_{P,Q}$ be a $PQ \times PQ$ matrix constituted by P times Q blocks of size $Q \times P$; block (p,q) being defined by $\mathbf{c}_{q/Q}\mathbf{c}'_{p/P}$; it is denominated the *commutation matrix* of size P and Q.

[4] To better see the structure of such matrices 0 have been replaced with ".".

Properties $\mathbf{T}_{P,Q}$ matrices satisfy the following properties

- $\mathbf{T}_{P,1} = \mathbf{T}_{1,P} = \mathbf{I}_P$.
- As it is an orthonormal matrix and as transposing twice gives the identity transformation

  $$(\mathbf{T}_{P,Q})^{-1} = \mathbf{T}'_{P,Q} = \mathbf{T}_{Q,P}.$$

  Indeed

  $$\mathbf{T}_{3,2} = \begin{pmatrix} 1 & . & . & . & . & . \\ . & . & . & 1 & . & . \\ & & & & & \\ . & 1 & . & . & . & . \\ . & . & . & . & 1 & . \\ & & & & & \\ . & . & 1 & . & . & . \\ . & . & . & . & . & 1 \end{pmatrix}$$

  is the transposed matrix of previously proposed $\mathbf{T}_{2,3}$ and also its inverse.
- When $Q = P$

  $$\mathbf{T}_{P,P} = \mathbf{T}'_{P,P}$$

  for instance

  $$\mathbf{T}_{2,2} = \begin{pmatrix} 1 & . & . & . \\ . & . & 1 & . \\ & & & \\ . & 1 & . & . \\ . & . & . & 1 \end{pmatrix};$$

  the two eigenvalues of these matrices are 1 and -1 with multiplicities $\frac{P(P+1)}{2}$ and $\frac{P(P-1)}{2}$. The associated eigenvector spaces are

  $$\left\{ \left\{ c_{i/P} \otimes c_{i/P} \right\}_{i=1,\ldots,P}, \left\{ c_{i/P} \otimes c_{j/P} + c_{j/P} \otimes c_{i/P} \right\}_{i<j<P} \right\} \quad \text{for} \quad 1\ ,$$

  $$\left\{ \left\{ c_{i/P} \otimes c_{j/P} - c_{j/P} \otimes c_{i/P} \right\}_{i<j<P} \right\} \quad \text{for} \quad -1\ .$$

- The commutation matrices take their name from their ability to commute the tensor product of matrices, indeed if $\mathbf{A}$ and $\mathbf{B}$ are respectively $P \times Q$ and $R \times S$ matrices, then

  $$\mathbf{T}_{R,P} (\mathbf{A} \otimes \mathbf{B}) \mathbf{T}_{Q,S} = \mathbf{B} \otimes \mathbf{A} \tag{31}$$
  $$(\mathbf{A} \otimes \mathbf{B}) \mathbf{T}_{Q,S} = \mathbf{T}_{P,R} (\mathbf{B} \otimes \mathbf{A})$$

3.2 Symmetrical and anti-symmetrical matrices

Definition Let $\mathbf{S}_P$ and $\mathbf{A}_P$ be the following $P^2 \times P^2$ matrices:

$$\begin{aligned} \mathbf{S}_P &= \frac{1}{2} (\mathbf{I}_{P^2} + \mathbf{T}_{P,P}) \\ \mathbf{A}_P &= \frac{1}{2} (\mathbf{I}_{P^2} - \mathbf{T}_{P,P}) \end{aligned}$$

respectively called symmetrical and anti-symmetrical matrices. For instance

$$\mathbf{S}_2 = \begin{pmatrix} 1 & \cdot & \cdot & \cdot \\ \cdot & \frac{1}{2} & \frac{1}{2} & \cdot \\ \cdot & \frac{1}{2} & \frac{1}{2} & \cdot \\ \cdot & \cdot & \cdot & 1 \end{pmatrix} \quad ; \quad \mathbf{A}_2 = \begin{pmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & -\frac{1}{2} & \frac{1}{2} & \cdot \\ \cdot & \frac{1}{2} & -\frac{1}{2} & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

$$\mathbf{S}_3 = \frac{1}{2} \begin{pmatrix} 2 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & 1 & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 2 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & 1 & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 2 \end{pmatrix}$$

$$\mathbf{A}_3 = \frac{1}{2} \begin{pmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 1 & \cdot & -1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot & \cdot & -1 & \cdot & \cdot \\ \cdot & -1 & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & -1 & \cdot \\ \cdot & \cdot & -1 & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & -1 & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

Properties

- The justification of their names is due to the fact that if $\mathbf{M}$ is a square matrix of dimension P:

$$\begin{aligned} \mathbf{S}_P \, vec\,(\mathbf{M}) &= vec\left(\frac{1}{2}(\mathbf{M} + \mathbf{M}')\right) \\ \mathbf{A}_P \, vec\,(\mathbf{M}) &= vec\left(\frac{1}{2}(\mathbf{M} - \mathbf{M}')\right) \end{aligned}$$

  since $\frac{1}{2}(\mathbf{M} + \mathbf{M}')$ is a symmetrical matrix: $(\mathbf{M} + \mathbf{M}')' = (\mathbf{M} + \mathbf{M}')$ with the same diagonal as $\mathbf{M}$ and, its counterpart, $\left(\frac{1}{2}(\mathbf{M} - \mathbf{M}')\right)$, is anti-symmetrical since $(\mathbf{M} - \mathbf{M}')' = -(\mathbf{M} - \mathbf{M}')$.

- They are symmetrical matrices

$$\begin{aligned} \mathbf{S}'_P &= \mathbf{S}_P \\ \mathbf{A}'_P &= \mathbf{A}_P \,. \end{aligned}$$

- Every row and column of $\mathbf{S}_P$ ($\mathbf{A}_P$) sums to one (zero)

$$\begin{aligned} \mathbf{S}_P \mathbf{1}_{P^2,1} &= \mathbf{1}_{P^2,1} \\ \mathbf{A}_P \mathbf{1}_{P^2,1} &= \mathbf{0}_{P^2,1} \,. \end{aligned}$$

- They are projectors since

$$\begin{aligned} \mathbf{S}_P^2 &= \mathbf{S}_P \\ \mathbf{A}_P^2 &= \mathbf{A}_P \end{aligned}$$

- They are projectors onto two orthocomplements of $\mathbf{R}^{P^2}$ since

$$\begin{aligned} \mathbf{S}_P + \mathbf{A}_P &= \mathbf{I}_{P^2} \\ \mathbf{S}_P \mathbf{A}_P &= \mathbf{0}_{P^2,P^2} \end{aligned}$$

  As a consequence, whatever are $\mathbf{M}$ and $\mathbf{N}$, two $P \times P$ matrices:

$$\left(\mathbf{S}_P vec\left(\mathbf{M}\right)\right)' \left(\mathbf{A}_P vec\left(\mathbf{N}\right)\right) = 0 \tag{32}$$

- The eigenvector subspaces of the two matrices $\mathbf{S}_P$ and $\mathbf{A}_P$ are no more that the two eigen vector spaces of $\mathbf{T}_{P,P}$ and $\mathbf{S}_P$ have got a unique eigen value of 1, and $\mathbf{A}_P$ also a unique eigenvalue but of -1. $\mathbf{S}_P$ is of rank $\frac{P(P+1)}{2}$ and $\mathbf{A}_P$ is of rank $\frac{P(P-1)}{2}$.

4 Symmetrical and Anti-Symmetrical Decomposition of Square Matrices

4.1 Preliminaries

4.1.1 From the tensorial decomposition

It is not the unique way, but it is a convenient one: we will introduce the symmetrical and anti-symmetrical decomposition of square matrices from the tensor decomposition of rectangular matrices as introduced in §2.6. In this section, matrix $\mathbf{Y}$ is square, of size $P \times P$, and its rows and columns are associated to the same series of items, so only one covariable matrix $\mathbf{W}$ of size $P \times K$ will be considered. In that framework, Equation (30) reads:

$$E\left[vec\left(\mathbf{Y}\right)\right] = \left(\mathbf{W} \otimes \mathbf{W}\right) vec\left(\mu\right) + \left(\mathbf{W} \otimes \mathbf{I}_P\right) vec\left(\alpha\right) + \left(\mathbf{I}_P \otimes \mathbf{W}\right) vec\left(\beta'\right) \tag{33}$$

μ is sized $K \times K$, and α as well as β are sized $I \times K$.

4.1.2 Sym./antisym. reformulation of the additive part

New parameterization Let consider the sum of the last two terms:

$$\left(\mathbf{W} \otimes \mathbf{I}_P\right) vec\left(\alpha\right) + \left(\mathbf{I}_P \otimes \mathbf{W}\right) vec\left(\beta'\right) \tag{34}$$

when matrices α and β are free. It is equivalent to

$$\left(\mathbf{W} \otimes \mathbf{I}_P + \left(\mathbf{I}_P \otimes \mathbf{W}\right) \mathbf{T}_{P,K}\right) vec\left(\pi\right) + \left(\mathbf{W} \otimes \mathbf{I}_P - \left(\mathbf{I}_P \otimes \mathbf{W}\right) \mathbf{T}_{P,K}\right) vec\left(\lambda\right) \tag{35}$$

where π and λ are free sized $P \times K$ matrices. To prove it, it is sufficient to notice that one can go from (35) to (34) with

$$\begin{aligned} \alpha &= \pi + \lambda \\ \beta &= \pi - \lambda \end{aligned}$$

and conversely with

$$\pi = \frac{1}{2}(\alpha + \beta) \quad (36)$$

$$\lambda = \frac{1}{2}(\alpha - \beta) \ . \quad (37)$$

Orthonormality The interest of formulation (35) is[5] that

$$\begin{aligned}(\mathbf{W} \otimes \mathbf{I}_P + (\mathbf{I}_P \otimes \mathbf{W}) \mathbf{T}_{P,K}) \, vec(\pi) &= (\mathbf{W} \otimes \mathbf{I}_P) \, vec(\pi) + (\mathbf{I}_P \otimes \mathbf{W}) \, vec(\pi') \\ &= vec(\pi \mathbf{W}') + vec(\mathbf{W}\pi') \\ &= vec(\pi \mathbf{W}' + \mathbf{W}\pi')\end{aligned}$$

is associated to a symmetrical matrix meanwhile

$$(\mathbf{W} \otimes \mathbf{I}_P - (\mathbf{I}_P \otimes \mathbf{W}) \mathbf{T}_{P,K}) \, vec(\lambda) = (\mathbf{W} \otimes \mathbf{I}_P) \, vec(\lambda) - (\mathbf{I}_P \otimes \mathbf{W}) \, vec(\lambda')$$

is associated to an antisymmetrical one. Remind that any symmetrical matrix is orthogonal to any antisymmetrical matrix since $\mathbf{S} = \mathbf{S}'$ and $\mathbf{A} = -\mathbf{A}'$ implies that

$$tr(\mathbf{AS}) = tr(\mathbf{SA}) = tr(\mathbf{A}'\mathbf{S}) = -tr(\mathbf{AS}) \ .$$

Constraints To keep the decomposition of the first term of (33), due to the equivalence of the two parameterizations, it suffices to transport the constraint from (α, β) to (π, λ) by means of the equations (36, 37) that is

$$\begin{pmatrix} \mathbf{I}_K \otimes \mathbf{W}' & \mathbf{0}_{KK,KI} \\ \mathbf{0}_{KK,IK} & \mathbf{I}_K \otimes \mathbf{W}' \end{pmatrix} \begin{pmatrix} vec(\pi) \\ vec(\lambda) \end{pmatrix} = \mathbf{0}$$

Decomposition the diagonal block The first term of (33), $(\mathbf{W} \otimes \mathbf{W}) \, vec(\mu)$, can also be decomposed in symmetrical and antisymmetrical parts:

$$\begin{aligned}(\mathbf{W} \otimes \mathbf{W}) \, vec(\mu) &= (\mathbf{W} \otimes \mathbf{W}) \, \mathbf{S}_K vec(\mu) + (\mathbf{W} \otimes \mathbf{W}) \, \mathbf{A}_K vec(\mu) \\ &= (\mathbf{W} \otimes \mathbf{W}) \, vec(\mu_S) + (\mathbf{W} \otimes \mathbf{W}) \, vec(\mu_A)\end{aligned}$$

where μ_S and μ_A are matrices of size $K \times K$ constrained to be symmetrical and antisymmetrical:

$$\begin{pmatrix} \mathbf{A}_K & \mathbf{0}_{K^2,K^2} \\ \mathbf{0}_{K^2,K^2} & \mathbf{S}_K \end{pmatrix} \begin{pmatrix} vec(\mu_S) \\ vec(\mu_A) \end{pmatrix} = \mathbf{0}$$

leading the parametric dimensions of μ_S and μ_A to respectively be $\frac{K(K+1)}{2}$ and $\frac{K(K-1)}{2}$.

4.2 Definition

With these preliminaries, we can safely introduce the symmetrical - antisymmetrical decomposition of a square matrix. It is now proposed under the three forms.

[5] Using Identity (40).

Scalar form

$$
\begin{aligned}
E\left[Y_{ij}\right] &= \sum_{kh} W_{ik}\mu_{S,kh}W_{jh} + \sum_{kh} W_{ik}\mu_{A,kh}W_{jh} \\
&\quad + \left(\sum_{h}\pi_{ih}W_{jh} + \sum_{k} W_{ik}\pi_{jk}\right) + \left(\sum_{h}\lambda_{ih}W_{jh} - \sum_{k} W_{ik}\lambda_{jk}\right) \ .
\end{aligned}
$$

Matricial form

$$
E\left[\mathbf{Y}\right] = \mathbf{W}\mu_S\mathbf{W}' + \mathbf{W}\mu_A\mathbf{W}' + \left(\pi\mathbf{W}' + \mathbf{W}\pi'\right) + \left(\lambda\mathbf{W}' - \mathbf{W}\lambda'\right) \ .
$$

Vectorized form

$$
\begin{aligned}
E\left[vec\left(\mathbf{Y}\right)\right] &= \left(\mathbf{W}\otimes\mathbf{W}\right)vec\left(\mu_S\right) + \left(\mathbf{W}\otimes\mathbf{W}\right)vec\left(\mu_A\right) \\
&\quad + \left(\mathbf{W}\otimes\mathbf{I}_I + \left(\mathbf{I}_I\otimes\mathbf{W}\right)\mathbf{T}_{P,K}\right)vec\left(\pi\right) + \left(\mathbf{W}\otimes\mathbf{I}_I - \left(\mathbf{I}_I\otimes\mathbf{W}\right)\mathbf{T}_{P,K}\right)vec\left(\lambda\right) \ .
\end{aligned}
$$

Degrees of freedom The parametric dimensions of the different terms are:

$$
\begin{aligned}
\mu_S &\rightarrow \frac{K\left(K+1\right)}{2} \\
\mu_A &\rightarrow \frac{K\left(K-1\right)}{2} \\
\pi &\rightarrow K\left(P-K\right) \\
\lambda &\rightarrow \left(P-K\right)K \ .
\end{aligned}
$$

4.3 Basic matrices

There is no much to add because we just have now to follow the standard path. So just the basic matrices are necessary. To shorten the formulae, we will denote $\left(\mathbf{W}'\mathbf{W}\right)^{-1}\mathbf{W}'$ by $\mathbf{U}$.

The design matrix with dimensions $P^2 \times \left(K^2 + K^2 + PK + PK\right)$:

$$
\begin{aligned}
\mathbf{X} &= \left(\mathbf{X}_1\ ,\ \mathbf{X}_2\ ,\ \mathbf{X}_3\ ,\ \mathbf{X}_4\right) \\
\mathbf{X}_1 = \mathbf{X}_2 &= \left(\mathbf{W}\otimes\mathbf{W}\right) \\
\mathbf{X}_3 &= \left(\mathbf{W}\otimes\mathbf{I}_P + \left(\mathbf{I}_P\otimes\mathbf{W}\right)\mathbf{T}_{P,K}\right) \\
\mathbf{X}_4 &= \left(\mathbf{W}\otimes\mathbf{I}_P - \left(\mathbf{I}_P\otimes\mathbf{W}\right)\mathbf{T}_{P,K}\right)
\end{aligned}
$$

The constraint matrix with dimensions $\left(K^2 + K^2 + K^2 + K^2\right) \times \left(K^2 + K^2 + PK + PK\right)$:

$$
\mathbf{C} = \begin{pmatrix}
\mathbf{A}_K & \mathbf{0} & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{S}_K & \mathbf{0} & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{I}_K\otimes\mathbf{W}' & \mathbf{0} \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_K\otimes\mathbf{W}'
\end{pmatrix}
$$

The scalar products of the $\mathbf{X}$ colums with dimensions $\left(K^2 + K^2 + PK + PK\right)\times\left(K^2 + K^2 + PK + PK\right)$

$$
\begin{array}{rcl}
\mathbf{X}'\mathbf{X} & = & \begin{pmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} & \mathbf{S}_{13} & \mathbf{S}_{14} \\ \mathbf{S}_{21} & \mathbf{S}_{22} & \mathbf{S}_{23} & \mathbf{S}_{24} \\ \mathbf{S}_{31} & \mathbf{S}_{32} & \mathbf{S}_{33} & \mathbf{S}_{34} \\ \mathbf{S}_{41} & \mathbf{S}_{42} & \mathbf{S}_{34} & \mathbf{S}_{44} \end{pmatrix} \\
\mathbf{S}_{11} = \mathbf{S}_{12} = \mathbf{S}_{22} & = & (\mathbf{W}'\mathbf{W} \otimes \mathbf{W}'\mathbf{W}) \\
\mathbf{S}_{33} & = & 2\left[(\mathbf{W}'\mathbf{W} \otimes \mathbf{I}) + (\mathbf{W}' \otimes \mathbf{W})\,\mathbf{T}_{P,K}\right] \\
\mathbf{S}_{44} & = & 2\left[(\mathbf{W}'\mathbf{W} \otimes \mathbf{I}) - (\mathbf{W}' \otimes \mathbf{W})\,\mathbf{T}_{P,K}\right] \\
\mathbf{S}_{13} = \mathbf{S}_{23} & = & (\mathbf{W}'\mathbf{W} \otimes \mathbf{W}' + (\mathbf{W}' \otimes \mathbf{W}'\mathbf{W})\,\mathbf{T}_{P,K}) \\
\mathbf{S}_{14} = \mathbf{S}_{24} & = & (\mathbf{W}'\mathbf{W} \otimes \mathbf{W}' - (\mathbf{W}' \otimes \mathbf{W}'\mathbf{W})\,\mathbf{T}_{P,K}) \\
\mathbf{S}_{34} & = & \mathbf{0}
\end{array}
$$

The scalar products of the $\mathbf{C}$ colums with dimensions $(K^2 + K^2 + PK + PK) \times (K^2 + K^2 + PK + PK)$

$$
\mathbf{C}'\mathbf{C} = \begin{pmatrix} \mathbf{A}_K & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_K & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_K \otimes \mathbf{W}\mathbf{W}' & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_K \otimes \mathbf{W}\mathbf{W}' \end{pmatrix}
$$

The estimator matrix with dimensions $(K^2 + K^2 + PK + PK) \times (P^2)$

$$
EST = \begin{pmatrix} \mathbf{S}_K\left[\mathbf{U}' \otimes \mathbf{U}\right] \\ \mathbf{A}_K\left[\mathbf{U}' \otimes \mathbf{U}\right] \\ \left(\mathbf{U}' \otimes P_{\{W\}^\perp}\right)\mathbf{S}_K\left[\left(\mathbf{P}_{\{W\}} \otimes \mathbf{P}_{\{W\}^\perp}\right) + \left(P_{\{W\}^\perp} \otimes \mathbf{P}_{\{W\}}\right)\right] \\ \left(\mathbf{U}' \otimes P_{\{W\}^\perp}\right)\mathbf{A}_K\left[\left(\mathbf{P}_{\{W\}} \otimes P_{\{W\}^\perp}\right) + \left(P_{\{W\}^\perp} \otimes \mathbf{P}_{\{W\}}\right)\right] \end{pmatrix}
$$

The variance matrix with dimensions $(K^2 + K^2 + PK + PK) \times (K^2 + K^2 + PK + PK)$

$$
\begin{array}{rcl}
VAR & = & \begin{pmatrix} \mathbf{V}_{11} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_{22} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{V}_{33} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{V}_{44} \end{pmatrix} \\
\mathbf{V}_{11} & = & \mathbf{S}_K\left[(\mathbf{W}'\mathbf{W})^{-1} \otimes (\mathbf{W}'\mathbf{W})^{-1}\right]\mathbf{S}_K \\
\mathbf{V}_{22} & = & \mathbf{A}_K\left[(\mathbf{W}'\mathbf{W})^{-1} \otimes (\mathbf{W}'\mathbf{W})^{-1}\right]\mathbf{A}_K \\
\mathbf{V}_{33} & = & \left(\mathbf{U}' \otimes P_{\{W\}^\perp}\right)\mathbf{S}_K\left[\left(\mathbf{P}_{\{W\}} \otimes P_{\{W\}^\perp}\right) + \left(P_{\{W\}^\perp} \otimes \mathbf{P}_{\{W\}}\right)\right]\mathbf{S}_K\left(\mathbf{U} \otimes P_{\{W\}^\perp}\right) \\
\mathbf{V}_{44} & = & \left(\mathbf{U}' \otimes P_{\{W\}^\perp}\right)\mathbf{A}_K\left[\left(\mathbf{P}_{\{W\}} \otimes P_{\{W\}^\perp}\right) + \left(P_{\{W\}^\perp} \otimes \mathbf{P}_{\{W\}}\right)\right]\mathbf{A}_K\left(\mathbf{U} \otimes P_{\{W\}^\perp}\right)
\end{array}
$$

4.4 Parameter estimators

Parameter estimators can be proposed in a matricial form.

$$
\begin{array}{rcl}
\mu_S & = & \frac{1}{2}\left(\mathbf{U}\mathbf{Y}\mathbf{U}' + \mathbf{U}\mathbf{Y}'\mathbf{U}'\right) \\
\mu_A & = & \frac{1}{2}\left(\mathbf{U}\mathbf{Y}\mathbf{U}' - \mathbf{U}\mathbf{Y}'\mathbf{U}'\right) \\
\pi & = & \frac{1}{2}\left(\mathbf{U}\mathbf{Y}\mathbf{P}_{\{W\}^\perp} + \mathbf{P}_{\{W\}^\perp}\mathbf{Y}\mathbf{U}'\right) \\
\lambda & = & \frac{1}{2}\left(\mathbf{U}\mathbf{Y}\mathbf{P}_{\{W\}^\perp} - \mathbf{P}_{\{W\}^\perp}\mathbf{Y}\mathbf{U}'\right)
\end{array}
$$

4.5 Symmetrical and additive decomposition

The direct splitting of the vector subspace associated to the term $(\mathbf{W}\otimes\mathbf{W})\,vec\,(\mu)$ into a symmetrical part, $(\mathbf{W}\otimes\mathbf{W})\,vec\,(\mu_S)$ and an antisymmetrical one, $(\mathbf{W}\otimes\mathbf{W})\,vec\,(\mu_A)$ asks for the possibility to split the other terms into four orhtogonal pieces like

$$\mathbf{X}_{\alpha\pi}\theta_{\alpha\pi}+\mathbf{X}_{\alpha\lambda}\theta_{\alpha\lambda}+\mathbf{X}_{\beta\pi}\theta_{\beta\pi}+\mathbf{X}_{\beta\lambda}\theta_{\beta\lambda}$$

such that

$$\begin{aligned}
\mathbf{X}_{\alpha\pi}\theta_{\alpha\pi}+\mathbf{X}_{\alpha\lambda}\theta_{\alpha\lambda} &= \left(\mathbf{W}\otimes\mathbf{P}_{\{W\}^\perp}\right)vec\,(\alpha)\\
\mathbf{X}_{\beta\pi}\theta_{\beta\pi}+\mathbf{X}_{\beta\lambda}\theta_{\beta\lambda} &= \left(\mathbf{P}_{\{W\}^\perp}\otimes\mathbf{W}\right)vec\,(\beta')\\
\mathbf{X}_{\alpha\pi}\theta_{\alpha\pi}+\mathbf{X}_{\beta\pi}\theta_{\beta\pi} &= \left(\mathbf{W}\otimes\mathbf{P}_{\{W\}^\perp}+\left(\mathbf{P}_{\{W\}^\perp}\otimes\mathbf{W}\right)\mathbf{T}_{P,K}\right)vec\,(\pi)\\
\mathbf{X}_{\alpha\lambda}\theta_{\alpha\lambda}+\mathbf{X}_{\beta\lambda}\theta_{\beta\lambda} &= \left(\mathbf{W}\otimes\mathbf{P}_{\{W\}^\perp}-\left(\mathbf{P}_{\{W\}^\perp}\otimes\mathbf{W}\right)\mathbf{T}_{P,K}\right)vec\,(\lambda)\ ?
\end{aligned}$$

The answer is negative. To see it better, we have to investigate the canonical correlation analyses of the four subspaces

$$\begin{aligned}
\{\alpha\} &= \left\{\mathbf{W}\otimes\mathbf{P}_{\{W\}^\perp}\right\}\\
\{\beta\} &= \left\{\mathbf{P}_{\{W\}^\perp}\otimes\mathbf{W}\right\}\\
\{\pi\} &= \left\{\mathbf{W}\otimes\mathbf{P}_{\{W\}^\perp}+\left(\mathbf{P}_{\{W\}^\perp}\otimes\mathbf{W}\right)\mathbf{T}_{P,K}\right\}\\
\{\lambda\} &= \left\{\mathbf{W}\otimes\mathbf{P}_{\{W\}^\perp}-\left(\mathbf{P}_{\{W\}^\perp}\otimes\mathbf{W}\right)\mathbf{T}_{P,K}\right\}\ .
\end{aligned}$$

Clearly, we already know that:

$$\begin{aligned}
\{\{\alpha\},\{\beta\}\} &= \{\{\pi\},\{\lambda\}\}\\
\{\alpha\} &\perp \{\beta\}\\
\{\pi\} &\perp \{\lambda\}\ .
\end{aligned}$$

Due to the symmetry of the roles of terms associated to α and β, we only have to consider the relationships between $\{\alpha\}$ and $\{\pi\}$, and between$\{\alpha\}$ and $\{\lambda\}$. Numerical experiments are not easy because they are very unstable. Nevertheless, it appears that most of the canonical correlations of both subspace pairs are $\frac{1}{\sqrt{2}}$ which can be seen when $K=1$.

Indeed, let $w_1,w_2,...,w_P$ be an orthonormalized basis of R^P such that the matrix $\mathbf{W}$ be proportional to w_1, then

$$\{\alpha\}=\{w_1\otimes w_2,w_1\otimes w_3,...,w_1\otimes w_P\}$$

and

$$\{\pi\}=\{w_1\otimes w_2+w_2\otimes w_1,...,w_1\otimes w_P+w_P\otimes w_1\}\,.$$

It occurs that these $2\,(P-1)$ vectors are mutually orthogonal to the exception of the $P-1$ pairs $(w_1\otimes w_i,w_1\otimes w_i+w_i\otimes w_1)$ which are, as a consequence, the canonical vectors: all of them having a correlation of $\frac{1}{\sqrt{2}}$.

4.6 Back to classical models

4.6.1 Hayman decomposition

Just considering the case of $\mathbf{W}=\mathbf{1}_P$.

4.6.2 Design decomposition

Just considering the case of $\mathbf{W} = \mathbf{I}_P$ (the π and λ components of the model disappear since their parametric dimension is zero).

4.6.3 A classification covariable set

Just for the fun, have a graphical look to the case when the matrix $\mathbf{W}$ is constituted with indicators of a classification of the row/column items, completed with the within constrasts.

$$\mathbf{W} = \begin{pmatrix} 1 & . & . & 1 & . \\ 1 & . & . & -1 & . \\ . & 1 & . & . & 1 \\ . & 1 & . & . & -1 \\ . & . & 1 & . & . \end{pmatrix}. \tag{38}$$

Figure 1 gives the tensor decomposition, just obtained with the products of the columns of $\mathbf{W}$; it is no more than the application of §2.6.3 formula. The construction of its symmetrical / antisymmetrical counterpart (Figure 2) is not so apparent but after a while of observation, it looks very natural, the symmetrical components being the diagonal and the contrasts above the diagonal.

5 To continue

Many things could be developped from the previous proposals. Among them:

- Explicitation, in the general case, of the canonical correlations as started in §4.5.
- Introduction of a diagonal effect to free the symmetrical part of it and give it the same parametric dimension that the antisymmetrical part.
- Generalization to more than two dimensions: arrays and not matrices (using **R** terminology); of course, not only $\left(R^P\right)^L$ cases have to be considered but rather $\prod_a \left(R^{P_a}\right)^{L_a}$ with at least one L_a greater than 1.
- From a statistical view, what happens when the square matrix is not complete:
  - half of the crosses with diagonal components,
  - half of the crosses without diagonal components,
  - general missing value configuration (including identifiability considerations),...
- Introduce a Bayesian perspective and look for adapted and flexible priors as done in [4] for bilinear models.

Figure 1: Tensor decomposition of a square matrix with contrasts generated from **W** matrix (38). Blue means +1, red means −1 and nothing 0. For an easier interpretation, a contrast is given in a matricial form by a square matrix of size 5 × 5. The 25 contrasts are themselves arranged in a 5 × 5 matrix, according to the tensor products of **W** columns drawn in the margins.

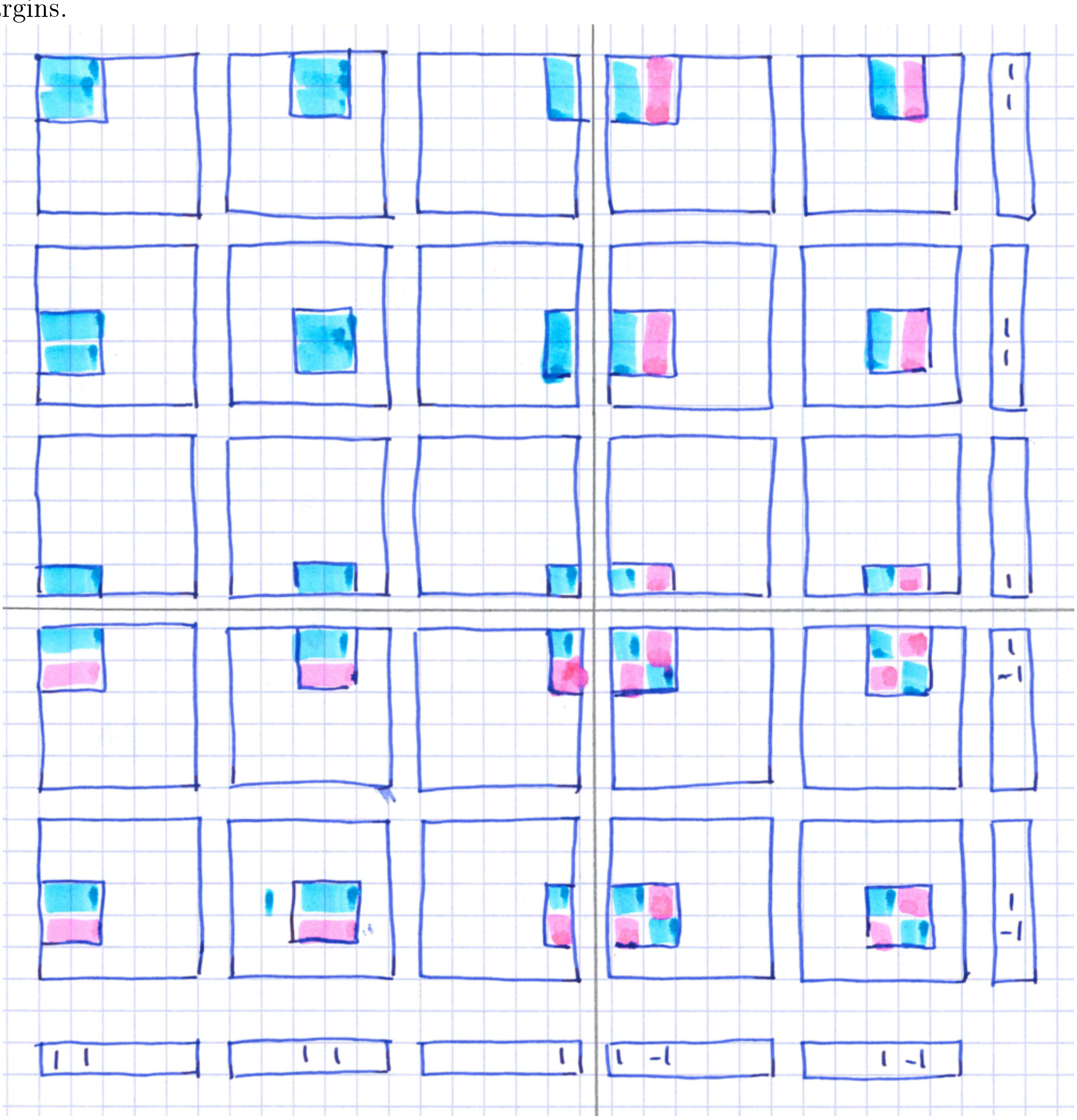
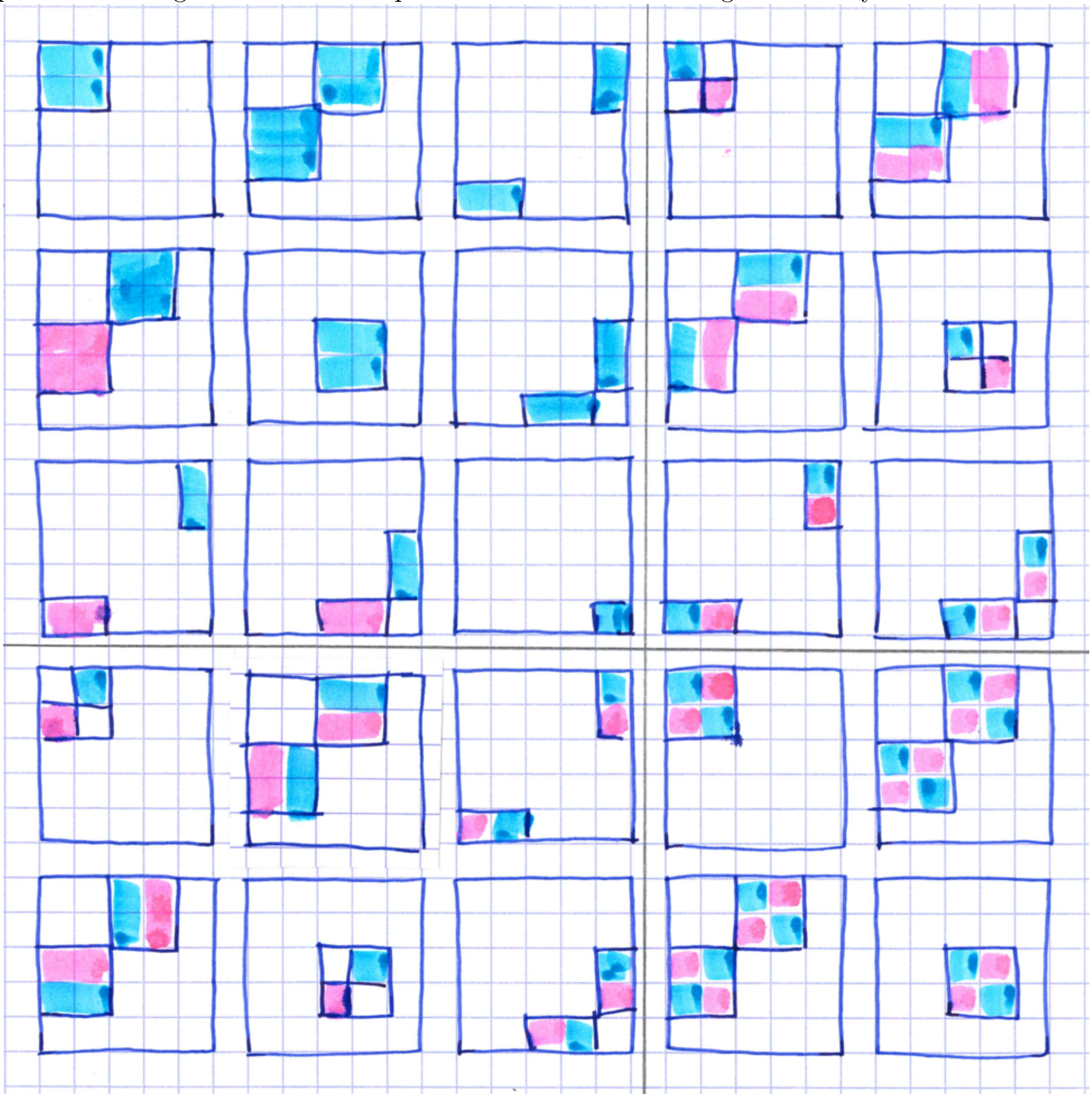


Figure 2: Symmetrical / antisymmetrical decomposition of a square matrix with contrasts generated from **W** matrix (38). Blue means +1, red means −1 and nothing 0. See further explanation if Figure 1 but tensor products are no more straightforwardly visible.



A Notations and Reminder

A.1 Writting

As far as possible, **matrices** are denoted by bold upper case letters, for instance: $\mathbf{X}$ of size (I, J) is a matrix with I rows and J columns. **Vectors** are indicated by lower cases, most of the time with indices. Vectors can be concatenated by brackets and comas to form a matrix: then if $\mathbf{X} = (x_1, x_2, ..., x_J)$, x_j are vectors of size I. More generally matrices can be exhibited in consistent blocks of submatrices.

Curly braces will be used for indicating the **vector space** generated either by a collection of vectors or the columns of matrices; for instance if $\mathbf{X}$ can be partitioned into $(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3)$ then $\{\mathbf{X}\} = \{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3\} = \{x_1, x_2, ..., x_J\}$ and designates the same vector subspace belonging to $\mathbf{R}^I$.

A.2 Special matrices and vectors

- $\mathbf{I}_N$ is the identity matrix of size N.
- $\mathbf{J}_N$ is the "all ones" matrix of size (N, N).
- $\mathbf{1}_N$ is the "all ones" vector of size N.
- $\mathbf{0}_{N,M}$ is the "all zeros" matrix of size (N, M).
- $\mathbf{c}_{i/N}$ is the ith canonical vector in $\mathbf{R}^N$, *i.e.* the column vector of size N, whose all components are null except the ith which is unity.
- $\overline{\mathbf{P}}_P$ is a $P \times (P-1)$ matrix such that its columns are orthonormal contrasts, *i.e.*

$$\left(\mathbf{1}_P, \overline{\mathbf{P}}\right)' \left(\mathbf{1}_P, \overline{\mathbf{P}}\right) = \begin{pmatrix} P & 0 \\ 0 & \mathbf{I}_{P-1} \end{pmatrix}$$

In all cases when the dimensions are obvious from the context, the indices can be dropped.

A.3 Operators

- $\mathbf{X}'$ is the matrix $\mathbf{X}$ transposed.
- $rk\,(\mathbf{X})$ is the rank of matrix $\mathbf{X}$.
- $vec\,(\mathbf{X})$ is the vector obtained from matrix $\mathbf{X}$ by stacking its columns. If $\mathbf{X}$ is a $I \times J$ matrix, $vec\,(\mathbf{X})$ is a IJ-vector whose component number $i + (j-1)\,I$ is $\mathbf{X}_{ij}$. For instance

$$\text{if } \mathbf{X} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \text{ then } vec\,(\mathbf{X}) = \begin{pmatrix} 1 \\ 4 \\ 2 \\ 5 \\ 3 \\ 6 \end{pmatrix} \text{ and } vec\,(\mathbf{X}') = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix}.$$

- $\mathbf{A} \otimes \mathbf{B}$ designates the Kronecker (or tensor) product of $\mathbf{A}$ by $\mathbf{B}$. If $\mathbf{A}$ is of size (N, M) and $\mathbf{B}$ of size (S, T), this product is defined as

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} A_{11}\mathbf{B} & A_{12}\mathbf{B} & \cdots & A_{1M}\mathbf{B} \\ A_{21}\mathbf{B} & A_{22}\mathbf{B} & \cdots & A_{2M}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1}\mathbf{B} & A_{N2}\mathbf{B} & \cdots & A_{NM}\mathbf{B} \end{pmatrix} \text{ then with size } (NS, MT) \ .$$

One can check that

$$(\mathbf{A} \otimes \mathbf{B})_{(n-1)S+s,(m-1)T+t} = \mathbf{A}_{nm}\mathbf{B}_{st}. \tag{39}$$

- The direct sum operator "$\oplus$" will be used for orthogonal vector subspaces. That is that $\{\mathbf{A}\} \oplus \{\mathbf{B}\}$ will be the vector space generated by $\{\mathbf{A}\} \cup \{\mathbf{B}\}$, $\{\mathbf{A}\}$ and $\{\mathbf{B}\}$ being supposed such that $\{\mathbf{A}\} \perp \{\mathbf{B}\}$ and not only that $\{\mathbf{A}\} \cap \{\mathbf{B}\} = \emptyset$.

A.4 Some definitions

- A square matrix, say $\mathbf{P}$, is said to be **orthonormal**[6] if and only if

$$\mathbf{P}'\mathbf{P} = \mathbf{I}$$

- A **basis of a vector (sub)space** is a minimal set of vectors such that any vector belonging to the vector (sub)space can be written as a linear combination of them. The number of such vectors is the dimension of the vector (sub)space.
- A basis of a vector (sub)space is said **orthogonal** (orthonormal) if its vectors are orthogonal (orthonormal).
- The **tensor product of two vector (sub)spaces** is the vector (sub)space generated by all tensor products of any two bases of them. It can be shown that it doesn't depend on the choice of the two bases, and that orthogonal (orthonormal) bases generate an orthogonal (orthonormal) basis.

A.5 Some properties

A.5.1 Miscellaneous

- All columns of an orthonormal matrix are normalized and orthogonal; also are its rows.
- $rk(\mathbf{X}) = rk(\mathbf{X}')$ gives the dimension of $\{\mathbf{X}\}$.

A.5.2 Kronecker product

- A lot of interesting properties are attached to the tensor product of matrices, among them
  - $\mathbf{I}_N \otimes \mathbf{I}_M = \mathbf{I}_{NM}$
  - $rk(\mathbf{A} \otimes \mathbf{B}) = rk(\mathbf{A})\, rk(\mathbf{B})$

[6] usually the adjective *orthogonal* is used but it is not appropriate because it does not imply the normalization of the vectors.

- $(\mathbf{A} \otimes \mathbf{B})' = (\mathbf{A}' \otimes \mathbf{B}')$
- $(\mathbf{A} \otimes \mathbf{B})^{-1} = \left(\mathbf{A}^{-1} \otimes \mathbf{B}^{-1}\right)$ when $\mathbf{A}$ and $\mathbf{B}$ are square non singular matrices.
- if $\mathbf{A}$ ($\mathbf{B}$) are square matrices and α (β) are one of their eigenvalues associated to eigenvectors a (b) then $\alpha\beta$ is an eigenvalue of $\mathbf{A} \otimes \mathbf{B}$ associated to eigenvector $a \otimes b$.
- if $\mathbf{A}$ and $\mathbf{B}$ are orthonormal matrices, so is $\mathbf{A} \otimes \mathbf{B}$.
- if $\mathbf{A}$ and $\mathbf{B}$ have respectively K and H columns and $\mathbf{C}$ is a $K \times H$ matrix, then

$$(\mathbf{A} \otimes \mathbf{B})\, vec\,(\mathbf{C}') = vec\,(\mathbf{B}\mathbf{C}'\mathbf{A}') \tag{40}$$

- if the dimensions are consistent

$$(\mathbf{A} \otimes \mathbf{B})\,(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{AC}) \otimes (\mathbf{BD})$$

- If $\mathbf{A}$ and $\mathbf{B}$ are orthonormal matrices of respective sizes I and J, then their sets of columns form respectively orthonormal bases of $\mathbf{R}^I$ and $\mathbf{R}^J$, and the set of columns of $\mathbf{A} \otimes \mathbf{B}$ form an orthonormal basis of $\mathbf{R}^{IJ}$.
- This can be proved with the previous rules

$$(\mathbf{A} \otimes \mathbf{B})'\,(\mathbf{A} \otimes \mathbf{B}) = (\mathbf{A}'\mathbf{A}) \otimes (\mathbf{B}'\mathbf{B}) = \mathbf{I}_I \otimes \mathbf{I}_J = \mathbf{I}_{IJ}$$

Here is a small numerical illustration. Let $I = 2$ and $J = 4$, and the two bases given by the columns of

$$\mathbf{A} = (a_1, a_2, a_3, a_4) = \tfrac{1}{2}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \quad \text{and} \quad \mathbf{B} = (b_1, b_2) = \tfrac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

then

$$\begin{aligned} \mathbf{C} &= \mathbf{A} \otimes \mathbf{B} \\ &= (a_1 \otimes b_1, a_1 \otimes b_2, a_2 \otimes b_1, a_2 \otimes b_2, a_3 \otimes b_1, a_3 \otimes b_2, a_4 \otimes b_1, a_4 \otimes b_2) \\ &= \frac{1}{2\sqrt{2}}\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{pmatrix} \end{aligned}$$

It can be easily checked that every column of $\mathbf{C}$ is normalized and that every couple of distinct columns of $\mathbf{C}$ are orthogonal.

A.5.3 Inverses

It can be checked that

$$\left(\mathbf{V} + \mathbf{X}'\mathbf{S}\mathbf{X}\right)^{-1} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}'\left(\mathbf{X}\mathbf{V}^{-1}\mathbf{X}' + \mathbf{S}^{-1}\right)^{-1}\mathbf{X}\mathbf{V}^{-1} \tag{41}$$

or after changing $\mathbf{S}$ for $-\mathbf{S}^{-1}$:

$$\left(\mathbf{V}-\mathbf{X}'\mathbf{S}^{-1}\mathbf{X}\right)^{-1}=\mathbf{V}^{-1}-\mathbf{V}^{-1}\mathbf{X}'\left(\mathbf{X}\mathbf{V}^{-1}\mathbf{X}'-\mathbf{S}\right)^{-1}\mathbf{X}\mathbf{V}^{-1} \tag{42}$$

also when $\mathbf{X}=\mathbf{I}$:

$$(\mathbf{V}+\mathbf{S})^{-1}=\mathbf{V}^{-1}-\mathbf{V}^{-1}\left(\mathbf{V}^{-1}+\mathbf{S}^{-1}\right)^{-1}\mathbf{V}^{-1}$$

Consider a $(n+m)\times(n+m)$ partitioned matrix, supposing that it is invertible (which implies that the two diagonal blocks are) it can be checked that:

$$\begin{pmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{pmatrix}^{-1}=\begin{pmatrix}\left(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1} & -\mathbf{A}^{-1}\mathbf{B}(\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1}\\ -\mathbf{D}^{-1}\mathbf{C}\left(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}\right)^{-1} & \left(\mathbf{D}-\mathbf{C}\mathbf{A}^{-1}\mathbf{B}\right)^{-1}\end{pmatrix}. \tag{43}$$

When the matrix is symmetric, the formula reads:

$$\begin{pmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{B}' & \mathbf{D}\end{pmatrix}^{-1}=\begin{pmatrix}\left(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{B}'\right)^{-1} & -\mathbf{A}^{-1}\mathbf{B}(\mathbf{D}-\mathbf{B}'\mathbf{A}^{-1}\mathbf{B})^{-1}\\ -\mathbf{D}^{-1}\mathbf{B}'\left(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{B}'\right)^{-1} & \left(\mathbf{D}-\mathbf{B}'\mathbf{A}^{-1}\mathbf{B}\right)^{-1}\end{pmatrix}. \tag{44}$$

It can be of use to notice that the expression of this last inverse does not immediately show it symmetry which gives us the following relationship:

$$\mathbf{A}^{-1}\mathbf{B}\left(\mathbf{D}-\mathbf{B}'\mathbf{A}^{-1}\mathbf{B}\right)^{-1}=\left(\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{B}'\right)^{-1}\mathbf{B}\mathbf{D}^{-1} \tag{45}$$

B Symmetric and Antisymetric Tensor Products of Matrices

These products have already been introduced[7] in [2] but here a more concise introduction is proposed. They could be used to avoid the commutation matrices but this alternative way was discarded to avoid the introduction of new notations.

Definition Let $\mathbf{A}$ a $N\times M$ matrix and $\mathbf{B}$ a $S\times T$ matrix

- their symmetric tensor product[8], $\mathbf{A}\boxplus\mathbf{B}$, is a $NS\times MT$ matrix defined by

$$\mathbf{A}\boxplus\mathbf{B}=\frac{1}{\sqrt{2}}\left(\mathbf{A}\otimes\mathbf{B}+\mathbf{B}\otimes\mathbf{A}\right),$$

- their antisymmetric tensor product, $\mathbf{A}\boxminus\mathbf{B}$, is a $NS\times MT$ matrix defined by

$$\mathbf{A}\boxminus\mathbf{B}=\frac{1}{\sqrt{2}}\left(\mathbf{A}\otimes\mathbf{B}-\mathbf{B}\otimes\mathbf{A}\right).$$

The coefficient $\frac{1}{\sqrt{2}}$ has been added to preserve the normalization of the tensor products of vectors.

[7] In fact, the definition seems slightly different!

[8] This definition (and the antisymetric product below) is a definition of ours, possibly it already exists under another name and/or this name is already used for other purpose. If you are aware of the fact, please indicate it to me.

Remarks

1. Notice that contrary to the Kronecker tensor product, a labeling of the rows and columns of $\mathbf{A} \boxplus \mathbf{B}$ and $\mathbf{A} \boxminus \mathbf{B}$ cannot be associated to the row and column labels of matrices $\mathbf{A}$ and $\mathbf{B}$ since each cell of it is the sum of two different multiplications, with the exception of square matrices as considered in this report.

2. The inverse transformation exists and is given by

$$
\begin{aligned}
\mathbf{A} \otimes \mathbf{B} &= \frac{1}{\sqrt{2}} (\mathbf{A} \boxplus \mathbf{B} + \mathbf{A} \boxminus \mathbf{B}) \\
\mathbf{B} \otimes \mathbf{A} &= \frac{1}{\sqrt{2}} (\mathbf{A} \boxplus \mathbf{B} - \mathbf{A} \boxminus \mathbf{B}) \ .
\end{aligned}
$$

Properties

- First make clear that despite of the qualification of symmetric, in general $(\mathbf{A} \boxplus \mathbf{B}) \neq (\mathbf{A} \boxplus \mathbf{B})'$. The first reason being that it is not a square matrix. This is true only when $\mathbf{A} = \mathbf{A}'$ and $\mathbf{B} = \mathbf{B}'$ since the transposed matrices are:

$$
\begin{aligned}
(\mathbf{A} \boxplus \mathbf{B})' &= (\mathbf{A}' \boxplus \mathbf{B}') \\
(\mathbf{A} \boxminus \mathbf{B})' &= (\mathbf{A}' \boxminus \mathbf{B}') \ .
\end{aligned}
$$

- The symmetric product is commutative

$$\mathbf{A} \boxplus \mathbf{B} = \mathbf{B} \boxplus \mathbf{A}$$

and

$$\mathbf{A} \boxminus \mathbf{B} = -\mathbf{B} \boxminus \mathbf{A} \ .$$

- When $\mathbf{B} = \mathbf{A}$

$$
\begin{aligned}
\mathbf{A} \boxminus \mathbf{A} &= \mathbf{0}_{N^2, M^2} \\
\mathbf{A} \boxplus \mathbf{A} &= \sqrt{2} \mathbf{A} \otimes \mathbf{A}
\end{aligned}
$$

- Composition rules can easily be established, for instance if the four matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ have the same number of rows

$$
\begin{aligned}
(\mathbf{A} \boxplus \mathbf{B})' (\mathbf{C} \boxplus \mathbf{D}) &= \frac{1}{\sqrt{2}} ((\mathbf{A}'\mathbf{C} \boxplus \mathbf{B}'\mathbf{D}) + (\mathbf{A}'\mathbf{D} \boxplus \mathbf{B}'\mathbf{C})) \\
(\mathbf{A} \boxminus \mathbf{B})' (\mathbf{C} \boxminus \mathbf{D}) &= \frac{1}{\sqrt{2}} ((\mathbf{A}'\mathbf{C} \boxplus \mathbf{B}'\mathbf{D}) - (\mathbf{A}'\mathbf{D} \boxplus \mathbf{B}'\mathbf{C})) \\
(\mathbf{A} \boxplus \mathbf{B})' (\mathbf{C} \boxminus \mathbf{D}) &= \frac{1}{\sqrt{2}} ((\mathbf{A}'\mathbf{C} \boxminus \mathbf{B}'\mathbf{D}) - (\mathbf{A}'\mathbf{D} \boxminus \mathbf{B}'\mathbf{C})) \\
(\mathbf{A} \boxminus \mathbf{B})' (\mathbf{C} \boxplus \mathbf{D}) &= \frac{1}{\sqrt{2}} ((\mathbf{A}'\mathbf{C} \boxminus \mathbf{B}'\mathbf{D}) + (\mathbf{A}'\mathbf{D} \boxminus \mathbf{B}'\mathbf{C}))
\end{aligned}
$$

- Let $\mathbf{Y}$ be a $P \times P$ matrix, and $\mathbf{A}$ and $\mathbf{B}$ matrices with P rows

$$
\begin{aligned}
(\mathbf{A} \boxplus \mathbf{B})' \, vec\,(\mathbf{Y}) &= \frac{1}{\sqrt{2}} vec\,(\mathbf{B}'\mathbf{Y}\mathbf{A} + \mathbf{A}'\mathbf{Y}\mathbf{B}) \\
(\mathbf{A} \boxminus \mathbf{B})' \, vec\,(\mathbf{Y}) &= \frac{1}{\sqrt{2}} vec\,(\mathbf{B}'\mathbf{Y}\mathbf{A} - \mathbf{A}'\mathbf{Y}\mathbf{B})
\end{aligned}
$$

When $\mathbf{A}$ and $\mathbf{B}$ matrices have got P rows and that $\mathbf{A}'\mathbf{A} = \mathbf{I}_K$, $\mathbf{B}'\mathbf{B} = \mathbf{I}_H$ and $\mathbf{A}'\mathbf{B} = \mathbf{0}_{K,H}$ then the following orthogonal properties can be checked

$\mathbf{X}'\mathbf{Y}$ / $\mathbf{X}' =$	$\mathbf{Y} =$ (dim)	$\mathbf{A} \boxplus \mathbf{A}$ K^2	$\mathbf{A} \boxminus \mathbf{A}$ K^2	$\mathbf{A} \boxplus \mathbf{B}$ KH	$\mathbf{A} \boxminus \mathbf{B}$ KH
$(\mathbf{A} \boxplus \mathbf{A})'$	K^2	$2\mathbf{I}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$
$(\mathbf{A} \boxminus \mathbf{A})'$	K^2	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$
$(\mathbf{A} \boxplus \mathbf{B})'$	KH	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{I}$	$\mathbf{0}$
$(\mathbf{A} \boxminus \mathbf{B})'$	KH	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{I}$

Taking into account that these results imply that $(\mathbf{B} \boxplus \mathbf{B})'(\mathbf{B} \boxplus \mathbf{B}) = \mathbf{I}_{H^2}$, we obtain with the columns of the four matrices, $((\mathbf{A} \boxplus \mathbf{A}), (\mathbf{A} \boxplus \mathbf{B}), (\mathbf{A} \boxminus \mathbf{B})(\mathbf{B} \boxplus \mathbf{B}))$ an orthogonal vector basis of $\mathbf{R}^{(KH)^2}$.

- As a consequence

$$\begin{aligned} rk(\mathbf{A} \boxplus \mathbf{A}) &= K^2 \\ rk(\mathbf{A} \boxplus \mathbf{B}) &= KH \\ rk(\mathbf{A} \boxminus \mathbf{B}) &= KH \\ rk(\mathbf{B} \boxplus \mathbf{B}) &= H^2 \end{aligned}$$

- Straightforwardly from the definitions of the symmetrical product some properties can be stated concerning the vector subspaces generated by the columns of the resulting matrices

1. $$\{\mathbf{A} \boxplus \mathbf{B}\} = \{\mathbf{B} \boxplus \mathbf{A}\}$$
2. $$\{\mathbf{A}_1 \boxplus \mathbf{B}, \mathbf{A}_2 \boxplus \mathbf{B}\} = \{(\mathbf{A}_1, \mathbf{A}_2) \boxplus \mathbf{B}\}$$
3. if $\{(\mathbf{A}, \mathbf{B})\} = \{\mathbf{A}\} \oplus \{\mathbf{B}\}$, then
$$\{(\mathbf{A}, \mathbf{B}) \boxplus (\mathbf{A}, \mathbf{B})\} = \{(\mathbf{A} \boxplus \mathbf{A})\} \oplus \{(\mathbf{A} \boxplus \mathbf{B})\} \oplus \{(\mathbf{B} \boxplus \mathbf{B})\}.$$
As a consequence
$$\begin{aligned} \dim\{(\mathbf{A}, \mathbf{B}) \boxplus (\mathbf{A}, \mathbf{B})\} &= \dim\{\mathbf{A}\}(\dim\{\mathbf{A}\} + 1)/2 + \dim\{\mathbf{A}\}\dim\{\mathbf{B}\} \\ &\quad + \dim\{\mathbf{B}\}(\dim\{\mathbf{B}\} + 1)/2. \end{aligned}$$

Applying the previous formulae to the case of vectors gives

$$(\mathbf{a} \boxplus \mathbf{b})'(\mathbf{f} \boxplus \mathbf{g}) = (\mathbf{a}'\mathbf{f})(\mathbf{b}'\mathbf{g}) + (\mathbf{a}'\mathbf{g})(\mathbf{b}'\mathbf{f})$$

The **symmetric product of two vector (sub)spaces** is the vector (sub)space generated by the columns of a matrix obtained as symmetric product of two matrices whose columns are any two bases of them

- Let us show that the definition does not depend on the choosen bases.
- That is if $\{\mathbf{A}\} = \{\mathbf{C}\}$ and $\{\mathbf{B}\} = \{\mathbf{D}\}$ then $\{\mathbf{A} \boxplus \mathbf{B}\} = \{\mathbf{C} \boxplus \mathbf{D}\}$.

- That is if $\mathbf{AP} = \mathbf{C}$ and $\mathbf{BQ} = \mathbf{D}$ then $(\mathbf{A} \boxplus \mathbf{B})\, \mathbf{R} = (\mathbf{C} \boxplus \mathbf{D})$ where $\mathbf{P}$, $\mathbf{Q}$ and $\mathbf{R}$ are regular matrices.

- Now

$$
\begin{aligned}
\sqrt{2}\mathbf{C} \boxplus \mathbf{D} &= (\mathbf{AP} \otimes \mathbf{BQ}) + (\mathbf{BQ} \otimes \mathbf{AP})\, \mathbf{T}_{H,K} \\
&= ((\mathbf{A} \otimes \mathbf{B})\,(\mathbf{P} \otimes \mathbf{Q})) + ((\mathbf{B} \otimes \mathbf{A})\,(\mathbf{Q} \otimes \mathbf{P}))\, \mathbf{T}_{H,K} \\
&= (\mathbf{A} \otimes \mathbf{B})\,(\mathbf{P} \otimes \mathbf{Q}) + (\mathbf{B} \otimes \mathbf{A})\, \mathbf{T}_{H,K}\, (\mathbf{P} \otimes \mathbf{Q}) \\
&= ((\mathbf{A} \otimes \mathbf{B}) + (\mathbf{B} \otimes \mathbf{A})\, \mathbf{T}_{H,K})\, (\mathbf{P} \otimes \mathbf{Q}) \\
&= (\mathbf{A} \boxplus \mathbf{B})\, (\mathbf{P} \otimes \mathbf{Q})\,.
\end{aligned}
$$

References

[1] J.-B. Denis. Analyse de régression factorielle. Biométrie-Praximétrie (1980), 20, 1-34.

[2] J.-B. Denis. Biadditivity for square tables. Rapport technique 1999-1 (1999). Unité de Biométrie, INRA Versailles, 57pp.

[3] F. A. van Eeuwijk, J.-B. Denis, and M.S. Kang. Genotype-by-environnment interaction. CRC Press (1996). Chapter: Incorporating additional information on genotypes and environments in models for two-way tables. pp. 15-49.

[4] J. Josse, F. A. van Eeuwijk, H.-P. Piepho, and J.-B. Denis. Another look at Bayesian Analysis of AMMI models for Genotype-Environment Data. Journal of Agricultural, Biological and Environmental Statistics (2014).

[5] B. I. Hayman. The analysis of variance of diallel tables. Biometrics (1954), 10(2), 235-44.

[6] J. R. Magnus and H. Neudecker. The commutation matrix: some properties and applications. The Annals of Statistics (1979), 7(2), 381-94.

[7] H. Scheffé. The analysis of variance. John Wiley, New-York, 477 pp.