

STATISTIQUE ET ANALYSE DES DONNÉES

ALAIN PATRIS

Analyse de la variance et indicatrices de modalités

Statistique et analyse des données, tome 6, n° 3 (1981), p. 65-101

http://www.numdam.org/item?id=SAD_1981__6_3_65_0

© Association pour la statistique et ses utilisations, 1981, tous droits réservés.

L'accès aux archives de la revue « Statistique et analyse des données » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

ANALYSE DE LA VARIANCE
ET INDICATRICES DE MODALITES

Alain PATRIS

Laboratoire de Probabilité et Statistique

Université de NANCY I

Case Officielle 140

54037 NANCY cedex

Résumé : On présente l'analyse de la variance à l'aide du modèle de régression linéaire. On étudie de manière formelle le système générateur des indicatrices des modalités engendrant le sous espace W sous lequel s'effectue la projection. On en étudie en particulier le rang et on en extrait une base. Lorsque le plan est orthogonal, on examine les propriétés géométriques de W rendant très simple l'opérateur de projection sur W . On montre enfin une propriété d'optimalité globale des plans orthogonaux.

Abstract : We are presenting the variance analysis in the frame of the linear regression model. We study in a formal way the generating system of the indicator variables of the modalities. These indicator variables generate a subspace W on which we are projecting. We study the rank of W and we construct a basis. If the experimental design is orthogonal we examine the geometrical properties of W in order to simplify the projection operator on W . Finally we show a global optimality property for orthogonal designs.

Mots clés : Analyse de la variance, système générateur, projecteur orthogonal.

0 - INTRODUCTION

Cet article contient une formalisation de l'analyse de la variance prenant délibérément appui sur le modèle de régression linéaire. On cherche à en tirer le maximum d'idées pouvant donner au praticien une possibilité de mieux comprendre son modèle. Certes, dès les premiers modèles, l'aspect régression linéaire avait été vu, mais il restait un peu marginal. Une évolution importante a eu lieu lorsqu'on a commencé à utiliser des plans non orthogonaux quelconques, où la notion de projecteur a pris un sens opérationnel.

Mais il semble qu'on se satisfait souvent de l'exhibition d'une matrice, sans trop s'inquiéter de l'opérateur sous-jacent. Ce type de méthode, à mon sens, ne favorise pas la compréhension des problèmes. C'est pourquoi je m'intéresserai par-

ticulièrément au système générateur associé aux matrices des opérateurs utilisés. Il ne sera donc question ici ni d'étudier des procédures de test raffinées ni de chercher des plans d'expérience particulièrement adaptés à telle ou telle expérience. Il s'agira essentiellement d'utiliser le formalisme de l'algèbre linéaire et de la géométrie euclidienne pour regarder certains problèmes d'un oeil un peu neuf. On donnera cependant en dernière partie une propriété d'optimalité des plans orthogonaux. Elle semble être plus ou moins sous-jacente dans de nombreux articles. L'auteur pense néanmoins qu'il était bon de la démontrer et de l'exposer dans toute sa généralité.

Remarque 1 : Pour des raisons de pure commodité, on notera avec un tilda les variables aléatoires définies sur la population de référence (appelées par la suite v.a. théoriques). Ainsi les n-échantillons correspondant (appelés par la suite v.a. empiriques) pourront être différenciés par la simple suppression du tilda, sans nécessiter, comme c'est l'usage, le passage à la lettre minuscule. On pourra ainsi utiliser les symboles : \tilde{g} (variable aléatoire théorique), g (variable aléatoire empirique), \tilde{G} (ensemble de v.a. théoriques), G (ensemble de v.a. empiriques).

Remarque 2 : Nous dirons qu'un estimateur T d'un paramètre β est de matrice de covariance minimale dans un ensemble τ d'estimateurs sans biais de β si T est élément de τ et si : $\forall T' \in \tau$, $\text{cov}(T') - \text{cov}(T)$ est une matrice définie positive.

1 - NOTATIONS

Le modèle de régression

L'analyse de la variance, comme la régression, est un problème d'estimation des paramètres β_1, \dots, β_m dans l'expression d'une espérance conditionnelle :

$$E^{\tilde{X}}(\tilde{Y}) = \sum_{i=1}^m \beta_i \tilde{g}_i(\tilde{X})$$

où $\tilde{X} = (\tilde{X}_1, \dots, \tilde{X}_r)$ et \tilde{Y} sont des v.a. définies sur $(\tilde{\Omega}, \tilde{\mathcal{A}}, \tilde{P})$, les composantes de \tilde{X} étant quantitatives ou qualitatives, $\tilde{g}_1(\tilde{X}), \dots, \tilde{g}_m(\tilde{X})$ étant des fonctions connues de \tilde{X} .

Pour estimer ces paramètres, on échantillonne conditionnellement à $\tilde{X} = x^1, \dots, \tilde{X} = x^n$ fixés. L'ensemble $\{x^1, \dots, x^n\}$ est appelé plan d'expérience. On note \tilde{Y}_i la v.a. associée au $i^{\text{ème}}$ tirage de l'échantillon et on note X la matrice (n, m) : $X(i, j) = \tilde{g}_j(x^i)$. On peut alors modéliser les résultats de l'expérience par :

$$\begin{bmatrix} \tilde{Y}_1 \\ \vdots \\ \tilde{Y}_n \end{bmatrix} = [\tilde{Y}] = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} + \begin{bmatrix} \tilde{\varepsilon}_1 \\ \vdots \\ \tilde{\varepsilon}_n \end{bmatrix} \text{ avec } \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix} = X\beta = \begin{bmatrix} E(\tilde{Y}/\tilde{X} = x^1) \\ \vdots \\ E(\tilde{Y}/\tilde{X} = x^n) \end{bmatrix}$$

Nous supposons par la suite que $\tilde{\varepsilon}_1, \dots, \tilde{\varepsilon}_n$ ont des lois normales de moyenne nulle, d'écart type σ et sont indépendantes.

Variables empiriques

Considérons \mathbb{R}^n en tant qu'espace vectoriel. Notons $I = \{\vec{i}_1, \dots, \vec{i}_n\}$ sa base canonique, et munissons-là de la mesure de probabilité : $P(\vec{i}_j) = \frac{1}{n}$.

Soit $u \in \mathbb{R}^n$. Ses coordonnées dans I seront notées :

$$[u]_I = [u(\vec{i}_1), \dots, u(\vec{i}_n)]^t$$

Ainsi nous assimilerons u à une v.a. empirique définie sur I .

Réciproquement toute fonction réelle définie sur I sera assimilée à un élément de \mathbb{R}^n par le processus inverse. Nous munirons \mathbb{R}^n du produit scalaire :

$$\langle u, v \rangle = \frac{1}{n} \sum_k u(\vec{i}_k) v(\vec{i}_k)$$

Base et système générateur

En particulier nous pouvons définir les variables empiriques associées aux colonnes de la matrice X :

g_i : telle que $[g_i]_I$ est la $i^{\text{ème}}$ colonne de X

μ : telle que $[\mu]_I = X\beta$

Notons W le sous espace de \mathbb{R}^n engendré par $G = \{g_1, \dots, g_m\}$.

On a :

$$[\mu]_I = X\beta \Leftrightarrow \mu = \sum_i \beta_i g_i \Leftrightarrow \mu \in W.$$

β est donc une représentation de μ dans G . Notons donc que tout élément de W possède deux représentations intéressantes :

- ses coordonnées dans la base I , notées $[u]_I$
- ses coordonnées dans le système générateur G , notées $[u]$ liées par :
 $[u]_I = X[u]$. On a en particulier : $\beta = [\mu]$.

On remarque tout de suite la différence entre ces deux systèmes de représentation :

- les coordonnées dans la base I sont uniques et sont vraiment pertinentes vis à vis de l'expérience puisque chaque composante représente une moyenne conditionnelle.

- les coordonnées dans le système générateur G offrent en général un "résumé" de l'expérience, permettent d'en extrapoler les résultats et de comprendre comment s'exerce l'influence des \tilde{X}_i sur \tilde{Y} . Il est cependant évident que la différence entre deux coordonnées dans G du même vecteur (dans le cas où G n'est pas une base) n'a aucune signification physique. Il ne s'agirait en effet que de deux manières différentes d'écrire le même ensemble de moyennes conditionnelles. Résumons-nous : la différence entre deux vecteurs de W est réellement significative, mais le choix entre plusieurs représentations dans G du même vecteur de W relève de conventions.

Variable aléatoire à valeurs dans W

Nous noterons Y^* la variable d'échantillonnage à valeur dans \mathbb{R}^n en tant qu'espace vectoriel (donc non liée à une base) définie par :

$$[Y^*]_I = [\tilde{Y}] = \begin{bmatrix} \tilde{Y}_1 \\ \vdots \\ \tilde{Y}_n \end{bmatrix}, \quad \tilde{Y}_i \text{ étant associée au } i^{\text{ème}} \text{ tirage.}$$

On peut alors présenter les techniques du calcul de régression à l'aide de calculs géométriques dans \mathbb{R}^n . Par exemple considérons :

P_W : projecteur orthogonal de \mathbb{R}^n sur W .

$\|\cdot\|$: l'application norme dans \mathbb{R}^n

$\langle g_i, \cdot \rangle$: l'application $u \in \mathbb{R}^n \rightarrow \langle g_i, u \rangle$

On pourra alors définir par composition :

$$\langle g_i, Y^* \rangle = \langle g_i, \cdot \rangle \circ Y^*$$

$$P_W(Y^*) = P_W \circ Y^*$$

$$\|P_W(Y^*)\|^2 = \|\cdot\|^2 \circ P_W \circ Y^*$$

2 - ESTIMATION ET FONCTIONS ESTIMABLES

2.1 - Estimation de β

En analyse de la variance G n'est pas une base, de sorte que la relation $\mu = \sum_i \beta_i g_i$ admet une infinité de solution. Chaque solution sera appelée une version de β . Pour retrouver l'unicité, il suffit d'imposer à β un système de contraintes linéaires du type $A\beta = [0]$ où A est une matrice de rang égal à $m - \dim W$ et telle que : quel que soit $u \in W$, il existe une version $[u_0]$ de la représentation de u dans G telle que $A[u_0] = [0]$.

Soit Q la "métrique" de W associée à G , c'est-à-dire que Q est la matrice (m, m) telle que $Q(i, j) = \langle g_i, g_j \rangle$. D'où $Q = \frac{1}{n} X^t X$. On montre alors que $\forall u \in R^n$:

1) Toute version des coordonnées de $P_W(u)$ dans G s'écrit :

$$[P_W(u)] = Q^{-1} \gamma_u$$

où Q^{-1} est un inverse généralisé de Q

γ_u est la matrice des coordonnées covariantes de $P_W(u)$:

$$\gamma_u^t = [\langle g_1, u \rangle, \dots, \langle g_m, u \rangle] = \frac{1}{n} (X^t [u]_I)^t = (Q [P_W(u)])^t$$

2) Le choix d'une contrainte $A\beta = [0]$ est équivalent au choix d'un inverse généralisé Q_0^{-1} de Q tel que $AQ_0^{-1}Q = [0]$. Nous dirons alors que la version β_0 de β telle que $A\beta_0 = [0]$, qui peut s'écrire : $\beta_0 = Q_0^{-1}Q\beta$ quelle que soit la version β , est associée à Q_0^{-1} .

L'estimation de β passe par l'intermédiaire de l'estimation des coordonnées covariantes de μ . Définissons :

$$\tilde{\phi}_i = \langle g_i, Y^* \rangle = \langle g_i, P_W(Y^*) \rangle = \frac{1}{n} [g_i]_I [\tilde{Y}]^t$$

$$\tilde{\phi}^t = [\tilde{\phi}_1, \dots, \tilde{\phi}_m] = \frac{1}{n} [\tilde{Y}]^t X$$

On peut alors écrire le théorème de Gauss-Markov ainsi :

Théorème 1 : Soit β_0 la version de β associée à Q_0^{-1} (telle que $A\beta_0 = [0]$). Alors :

1) $\hat{\beta}_0 = Q_0^{-1}\bar{\phi}$ est un estimateur de β_0 non biaisé, au sens du maximum de vraisemblance, et de matrice de covariance minimum dans la classe des estimateurs T de β_0 non biaisés et tels que $AT = [0]$.

$$2) \text{Cov } \hat{\beta}_0 = \frac{\sigma^2}{n} Q_0^{-1} Q_0 Q_0^{-1}$$

3) $\hat{\mu} = P_W(Y^*) = \sum_{i=1}^m \hat{\beta}_i g_i$ est un estimateur de μ non biaisé, au sens du maximum de vraisemblance, indépendant du choix de la contrainte $A\beta_0 = [0]$.

On constate en particulier que $\hat{\beta}_0$ n'est autre que la représentation dans G de $\hat{\mu} = P_W(Y^*)$ vérifiant les mêmes conditions que β_0 .

2.2 - Hypothèses sur β

Examinons des hypothèses sur β du type $K\beta = k$ où K est une matrice (ℓ, m) . Puisque β n'est pas unique, on peut donner trois sens à une telle hypothèse :

- il existe au moins une version de β telle que $K\beta = k$
- la version de β associée à Q_0^{-1} (connue) vérifie $K\beta = k$
- nous avons peut-être choisi K telle que $\forall u \in W$, alors $K[u]$ ne dépend pas du choix de la version $[u]$ de la représentation de u dans G . De sorte que $K[u]$ est effectivement fonction uniquement de u .

Définition : Soit K une matrice (ℓ, m) telle que $\forall u \in W$, alors $K[u]$ ne dépende pas du choix de la version $[u]$ de la représentation de u dans G . Nous dirons alors que :

- la fonction $\psi(K, \cdot) : u \rightarrow K[u]$ de W dans R^ℓ est une fonction estimable.
- $K\beta$ est estimable.

Propriétés des fonctions estimables

a) Soit Z une matrice $(m, m - \dim W)$ dont les colonnes sont des représentations dans G , linéairement indépendantes, du vecteur nul. Alors les quatre affirmations suivantes sont équivalentes :

- $K\beta$ est estimable.
- K est de la forme aQ , a étant une matrice (ℓ, m)
- K est de la forme bX , b étant une matrice (ℓ, n)
- $KZ = [0], [0]$ étant une matrice nulle.

En effet, dire que $\psi(K, \cdot)$ est une fonction (évidemment linéaire) de W dans R^l est équivalent à dire qu'il existe une matrice $b(l, n)$ telle que $\psi(K, u) = b[u]_{\Gamma}$. Or $[u]_{\Gamma} = X u$. De plus, puisque $Q = \frac{1}{n} X^t X$, on peut montrer que quel que soit l'inverse généralisé Q^- de Q on a : $X = XQ^-Q$. La quatrième équivalence vient du fait que la différence entre deux représentations dans G du même vecteur est nécessairement de la forme Zx puisque c'est une représentation du vecteur nul. Et on a : $\psi(K, \vec{0}) = [0]$.

On retrouve donc bien la notion habituelle de fonction estimable. On pourra voir une définition donnée en terme assez semblable dans [4]. On y verra d'ailleurs une démonstration plus directe de l'équivalence entre 1 et 3 ainsi que l'introduction de géométrie (non euclidienne) dans l'estimation non paramétrique dans le modèle linéaire. L'équivalence de ces quatre affirmations permet de montrer que :

b) Chaque version β_0 de β peut se mettre sous forme de fonction estimable : $\beta_0 = Q_0^- Q \beta$.

c) L'ensemble des fonctions estimables $\psi(K, \cdot)$ où K n'a qu'une ligne est le dual de W . Le fait de prendre des fonctions estimables ne limite donc pas notre choix parmi les fonctions linéaires du vecteur de R^n , μ , formé des moyennes conditionnelles.

d) Si $K\beta$ est estimable, et quels que soient les inverses généralisés Q_1^- et Q_2^- de Q on a :

$$1) KQ_1^-Q = K \text{ et } KQ_1^-QQ_1^-K^t = KQ_2^-K^t$$

$$2) \text{cov}(K\hat{\beta}) = \frac{\sigma^2}{n} KQ_1^-K^t \text{ qui est inversible si } K \text{ est de rang plein.}$$

e) Soient K_1, \dots, K_ρ des matrices (l, m) telles que $K_1\beta$ soit estimable. Alors pour que $\psi(K_1, \cdot), \dots, \psi(K_\rho, \cdot)$ soient linéairement indépendantes, il faut et il suffit que K_1, \dots, K_ρ le soient.

f) Si $K\beta$ est estimable, si $\hat{\beta}$ est un estimateur de β de Gaus-Markov, alors $K\hat{\beta}$ est un estimateur de $K\beta$ non biaisé, au sens du maximum de vraisemblance et de matrice de covariance minimale parmi les estimateurs sans biais de $K\beta$.

Relations entre les trois types d'hypothèses

Soit $K\beta$ non estimable, Q_0^- un inverse généralisé de Q . Considérons les deux hypothèses :

H_1 : la version de β , notée β_0 , associée à Q_0^- vérifie $K\beta_0 = k$

H_2 : il existe au moins une version de β telle que $K\beta = k$.

On peut alors montrer qu'il existe K' , K'' et k'' tels que :

- 1) $K'\beta$ et $K''\beta$ sont estimables
- 2) $H_1 \Leftrightarrow K'\beta = k$ (avec $K' = KQ_0^{-1}Q$)
- 3) $H_2 \Leftrightarrow K''\beta = k''$

Par la suite nous caractérisons une hypothèse sur β par un sous espace affine de W en nous servant du résultat bien connu :

- 1) Soit $\psi(K, \cdot)$ une fonction estimable. Alors il existe un sous espace affine \mathcal{E} de W tel que $\psi(K, u) = k \Leftrightarrow u \in \mathcal{E}$.
- 2) Soit \mathcal{E} un sous espace affine de W . Alors il existe une fonction estimable $\psi(K, \cdot)$ et une matrice k telles que : $u \in \mathcal{E} \Leftrightarrow \psi(K, u) = k$.

Test d'hypothèse

Soit \mathcal{E} le sous espace affine de W formé des éléments u de W tels que $K[u_0] = k$ où :

$K\beta$ n'est pas nécessairement estimable

$[u_0]$ est la représentation de u dans G associée à Q_0^{-1}

Soit $s = \dim \mathcal{E}$, $q = \dim W$

Soit H_0 l'hypothèse : $\mu \in \mathcal{E}$

Soit H_1 l'alternative : $\mu \notin \mathcal{E}$

Soit $\tilde{F} = \frac{n-q}{q-s} \frac{\|P_W(Y^*) - P_{\mathcal{E}}(Y^*)\|^2}{\|Y^* - P_W(Y^*)\|^2}$. On rappelle que :

1) Le test de H_0 contre H_1 obtenu par la méthode du rapport de vraisemblance correspond à la procédure :

On accepte H_0 si $\tilde{F} < C_\alpha$

On rejette H_0 si $\tilde{F} > C_\alpha$

$$\begin{aligned} 2) \|P_W(Y^*) - P_{\mathcal{E}}(Y^*)\|^2 &= (K\hat{\beta}_0 - k)^t (KQ_0^{-1}QQ_0^{-t}K^t)^{-1} (K\hat{\beta}_0 - k) \\ &= \|P_W(Y^*)\|^2 - \|P_{\mathcal{E}}(Y^*)\|^2 \text{ si } \mathcal{E} \text{ est un sous} \end{aligned}$$

espace vectoriel, donc si $k = [0]$.

$$\|Y^* - P_W(Y^*)\|^2 = \|Y^*\|^2 - \|P_W(Y^*)\|^2$$

$$\|Y^*\|^2 = \frac{1}{n} [\tilde{Y}]^t [\tilde{Y}]$$

$$\|P_W(Y^*)\|^2 = \hat{\beta}_0^t Q \hat{\beta}_0 = \tilde{\phi} Q^{-1} \tilde{\phi}, \text{ pour } \tilde{\phi} \text{ définie Th. 1.}$$

3) Sous H_0 , \tilde{F} suit un $F_{q-s, n-q}$ centré.

4) Sous une alternative $K\beta_0 = k'$, alors \tilde{F} suit un $F_{q-s, n-q}$ décentré de paramètre :

$\delta^2 = \frac{n}{\sigma^2} d^2(\mathcal{E}, \mathcal{E}') = \frac{n}{\sigma^2} (k-k')^t (KQ_0^- Q_0^- K^t)^{-1} (k-k')$ où :
 \mathcal{E}' est le sous espace affine parallèle à \mathcal{E} défini par :
 $u \in \mathcal{E}' \Leftrightarrow K[u_0] = k'$
 $d(\mathcal{E}, \mathcal{E}')$ est la distance entre \mathcal{E} et \mathcal{E}' .

3 - LE MODELE D'ANALYSE DE LA VARIANCE

Pour ne pas compliquer davantage les notations, nous n'allons pas parler du modèle d'analyse de la covariance. Nous allons donc supposer que toutes les variables explicatives $\tilde{X}_1, \dots, \tilde{X}_r$ sont qualitatives. Nous pouvons toujours supposer que \tilde{X}_i est codée $1, \dots, m_i$. Notons :

$\tilde{\Omega}_i^j = \tilde{X}_i^{-1}(j)$, appelée $j^{\text{ème}}$ modalité de \tilde{X}_i

\tilde{M}_i^j l'indicatrice de $\tilde{\Omega}_i^j$

$\tilde{\Omega}_{i_1 \dots i_p}^{j_1 \dots j_p} = \bigcap_{\ell=1}^p \tilde{\Omega}_{i_\ell}^{j_\ell}$ appelé noeud d'ordre p

$\tilde{M}_{i_1 \dots i_p}^{j_1 \dots j_p} = \prod_{\ell=1}^p \tilde{M}_{i_\ell}^{j_\ell}$ l'indicatrice du noeud précédent.

$\tilde{\Omega}$ sera appelé noeud d'ordre zéro, et son indicatrice notée \tilde{J}_0 .

\tilde{G}^p sera l'ensemble des indicatrices des noeuds d'ordre inférieur ou égal à p

\tilde{W}^p sera l'espace vectoriel engendré par les éléments de \tilde{G}^p .

Définition 2 : Nous dirons que nous avons un modèle d'analyse de la variance avec interactions d'ordre p si ce modèle est :

$$E^{\tilde{X}}(\tilde{Y}) \text{ est élément de } \tilde{W}^p.$$

Ce modèle peut aussi s'écrire des deux manières suivantes :

$$E^{\tilde{X}}(\tilde{Y}) = \beta_0 \tilde{J}_0 + \sum_i \sum_j \beta_i^j \tilde{M}_i^j + \dots + \sum_{i_1 < i_2 < \dots < i_p} \sum_{j_1 \dots j_p} \beta_{i_1 \dots i_p}^{j_1 \dots j_p} \tilde{M}_{i_1 \dots i_p}^{j_1 \dots j_p}$$

$$E(\tilde{Y} / \tilde{X}_1 = \ell_1, \dots, \tilde{X}_r = \ell_r) = \beta_0 + \sum_i \beta_i^{\ell_i} + \dots + \sum_{i_1 < i_2 < \dots < i_p} \beta_{i_1 \dots i_p}^{\ell_{i_1} \dots \ell_{i_p}}$$

Remarquons que l'on peut toujours utiliser le modèle avec interactions d'ordre r (il faut cependant vérifier les hypothèses sur la loi des $\tilde{\epsilon}_i$).

Bien entendu, à chacune des indicatrices de G^p , est associée par le plan d'expérience une variable empirique notée sans tilda. Par exemple à \tilde{M}_i^j est associée la variable empirique M_i^j telle que pour tout $\vec{i}_\ell \in I$ on ait :

$$M_i^j(\vec{i}_\ell) = \begin{cases} 1 & \text{si le } \ell^{\text{ème}} \text{ individu du plan d'expérience est dans } \tilde{\Omega}_i^j \\ 0 & \text{sinon} \end{cases}$$

M_i^j pourra donc être considérée comme l'indicatrice d'un sous-ensemble de I noté Ω_i^j , qui est en fait la trace de $\tilde{\Omega}_i^j$ sur I . De plus, si j varie de 1 à m_i , on construit ainsi une partition de $I : \{\Omega_i^1, \dots, \Omega_i^{m_i}\}$, partition que l'on peut associer à une variable qualitative empirique notée X_i et telle que :

$$X_i(\vec{i}_\ell) = j \text{ si } \vec{i}_\ell \in \Omega_i^j, \text{ donc si le } \ell^{\text{ième}} \text{ individu est dans } \tilde{\Omega}_i^j.$$

On constatera encore que les variables empiriques associées aux indicatrices des noeuds d'ordre k entre $\tilde{X}_{i_1}, \dots, \tilde{X}_{i_k}$ sont justement les indicatrices des mêmes noeuds entre les variables empiriques X_{i_1}, \dots, X_{i_k} . C'est-à-dire que :

$$M_{i_1 \dots i_k}^{j_1 \dots j_k} = \prod_{\ell=1}^k M_{i_\ell}^{j_\ell}, \text{ indicatrice du noeud } \Omega_{i_1 \dots i_k}^{j_1 \dots j_k} = \prod_{\ell=1}^k \Omega_{i_\ell}^{j_\ell}$$

Ainsi le modèle avec interactions d'ordre p se traduit dans l'espace des variables empiriques par la relation :

$$\mu = \beta_0 J_0 + \sum_i \sum_j \beta_{ij}^j M_i^j + \dots + \sum_{i_1 < \dots < i_p} \sum_{j_1 \dots j_p} \beta_{i_1 \dots i_p}^{j_1 \dots j_p} M_{i_1 \dots i_p}^{j_1 \dots j_p}$$

Nous noterons :

J_0 l'indicatrice de I , considéré comme noeud empirique d'ordre zéro

G^p l'ensemble des indicatrices des noeuds empiriques d'ordre inférieur ou égal à p .

W^p le sous espace de \mathbb{R}^n engendré par les éléments de G^p

Nous remarquerons que la métrique Q de W^p associée à G^p se calcule aisément, puisque si $n_{i_1 \dots i_k}^{j_1 \dots j_k}$ désigne l'effectif du noeud $\Omega_{i_1 \dots i_k}^{j_1 \dots j_k}$ dans le plan d'expérience alors :

$$\langle M_{i_1 \dots i_k}^{j_1 \dots j_k}, M_{i'_1 \dots i'_\ell}^{j'_1 \dots j'_\ell} \rangle = \frac{1}{n} n_{i_1 \dots i_k i'_1 \dots i'_\ell}^{j_1 \dots j_k j'_1 \dots j'_\ell} = P(\Omega_{i_1 \dots i_k}^{j_1 \dots j_k} \cap \Omega_{i'_1 \dots i'_\ell}^{j'_1 \dots j'_\ell})$$

D'autre part les coordonnées covariantes de $P_W(Y^*)$ sont aisées à calculer puisque :

$$\langle M_{i_1 \dots i_k}^{j_1 \dots j_k}, Y^* \rangle = \frac{1}{n} \sum_{i_1 \dots i_k \in \Omega_{i_1 \dots i_k}} j_1 \dots j_k \tilde{Y}_\ell$$

Nous pouvons donc utiliser tous les résultats du § 3 pour l'estimation des paramètres et pour les tests d'hypothèses, et ceci quel que soit le plan d'expérience. On se trouve cependant en présence de deux problèmes importants :

a) les éléments de G^P (et de \tilde{G}^P) ont de nombreuses liaisons linéaires.

Ainsi :

$$M_{i_1 \dots i_k}^{j_1 \dots j_k} = \sum_{j_{k+1}} M_{i_1 \dots i_k}^{j_1 \dots j_k j_{k+1}}$$

Ceci pose le problème de trouver les dimensions des sous espaces associés aux hypothèses à tester, dimensions qui sont les degrés de liberté des F . Il nous faut aussi savoir comment caractériser les fonctions estimables.

b) Dès que r dépasse 3 et que p atteint 2, G^P contient de très nombreux éléments (171 pour $r = 4, p = 2, m = m = m = m = 5$) de sorte qu'il devient difficile de trouver numériquement un inverse généralisé de Q , d'autant que cette matrice est de rang bien inférieur à sa dimension (113 au maximum dans l'exemple précédent).

Plusieurs voies de recherche s'offrent alors. Par exemple :

- pour simplifier le calcul de Q^- , la méthode la plus utilisée dans la littérature consiste à inverser la matrice Q par morceaux. Ce qui revient à utiliser la méthode suivante :

soit $G = \{g_1, \dots, g_\ell, g_{\ell+1}, \dots, g_m\}$, W_1 engendré par g_1, \dots, g_ℓ et W_2 engendré par $g_{\ell+1}, \dots, g_m$. Pour tout $u \in \mathbb{R}^n$ notons : $u = P_{W_2}(u)$ et $u^- = u - P_{W_2}(u)$. Alors :

$$\langle u, g_i^- \rangle = \langle g_i^-, u^- \rangle = \langle g_i^-, u \rangle - \langle P_{W_2}(g_i^-), u \rangle \quad (1)$$

Supposons que Q s'écrive $Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}$ et que $P_{W_2}(u) = \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}$

Supposons, pour simplifier, que Q_{22} soit inversible. Alors :

$$\langle g_i^-, g_j^- \rangle = Q'(i, j) \text{ avec } Q' = Q_{11} - Q_{21} Q_{22}^{-1} Q_{12}.$$

De (1) on obtient alors :

$$Q' \begin{bmatrix} u_1 \\ \vdots \\ u_\ell \end{bmatrix} = \begin{bmatrix} \langle u, g_1^- \rangle \\ \vdots \\ \langle u, g_\ell^- \rangle \end{bmatrix} - Q_{21} Q_{22}^{-1} \begin{bmatrix} \langle u, g_{\ell+1}^- \rangle \\ \vdots \\ \langle u, g_m^- \rangle \end{bmatrix}$$

Cette méthode s'utilise surtout dans les modèles sans interactions et quand la matrice Q_{22} s'inverse facilement, c'est-à-dire dans le cas où $r = 2$ (on prend Q_{22} diagonale) et dans le cas où le plan est partiellement orthogonal.

- construire le projecteur par un algorithme itératif comme dans [2]. On n'obtient pas la matrice de covariance de l'estimateur, mais on peut cependant effectuer les tests usuels en utilisant l'algorithme pour rechercher les normes des projections de Y sur les différents sous espaces vectoriels associés aux différentes hypothèses.

- rechercher une base de W^p , ce qui permet d'utiliser une métrique plus petite et inversible.

- trouver des plans d'expérience où W^p se décompose en sous espaces orthogonaux dont les projecteurs s'expriment simplement. C'est le cas des plans orthogonaux. On pourra aussi chercher (cas $r = 2$) des plans où W^1 se décompose en deux sous espaces simples dont les deux projecteurs associés P_1 et P_2 sont tels que $P_1 \circ P_2$ et $P_2 \circ P_1$ n'ont qu'une seule valeur propre non nulle, ce qui est le cas des plans équilibrés, notion que l'on pourra étendre pour $r > 2$, mais dont nous ne parlerons pas ici.

4 - BASE ET DIMENSIONS

Puisque $\hat{\beta}$ et β sont les représentations dans G^p d'éléments de W^p ($\hat{\mu}$ et μ) les problèmes liés à leurs multiples versions sont ceux que pose toute étude des représentations dans G^p d'un élément quelconque de W^p .

4.1 - Plans sans biais

Reprenons les notations du second paragraphe :

$$\tilde{G} = \{\tilde{g}_1(\tilde{X}), \dots, \tilde{g}_m(\tilde{X})\} \text{ et } G = \{g_1, \dots, g_m\}$$

$$E^{\tilde{X}}(\tilde{Y}) = \sum_{i=1}^m \beta_i \tilde{g}_i(\tilde{X}) \quad (1) \text{ et } \mu = \sum_i \beta_i g_i \quad (2)$$

Dans notre modèle ni \tilde{G} ni G ne sont des bases. Les équations (1) et (2) ont donc plusieurs solutions. Le problème est de s'assurer que toute solution d'une équation est également solution de l'autre. De par la construction de μ il est aisé de voir que toute solution de (1) est aussi solution de (2). Mais la réciproque n'est pas nécessairement vraie. En effet, prenons un exemple où :

$$E^{\tilde{X}}(\tilde{Y}) = \tilde{g}_1 + \tilde{g}_2, \quad \tilde{g}_1 \text{ et } \tilde{g}_2 \text{ étant linéairement indépendantes. Si l'on choisit un}$$

plan d'expérience où $g_1 = g_2$ on a alors : $\mu = \beta_1 g_1 + \beta_2 g_2$ dès que $\beta_1 + \beta_2 = 2$.

Il est aisé de voir que pour que toute solution de (2) soit aussi solution de (1) il faut et il suffit que \tilde{G} et G soient de même rang. C'est d'ailleurs aussi la condition nécessaire et suffisante pour pouvoir se servir de l'estimation de β pour faire de la prévision sur \tilde{Y} ainsi que pour pouvoir tester toutes les hypothèses sur β ayant un sens (ce qui est en rapport avec la notion de plan sans biais).

Nous allons faire d'autre part l'hypothèse que $\tilde{X}_1, \dots, \tilde{X}_r$ ont leurs noeuds d'ordre r tous non vides. Puisque les éléments de \tilde{G}^r sont engendrés par les indicatrices des noeuds d'ordre r , indicatrices qui seront alors linéairement indépendantes, notre hypothèse implique donc que : $\text{rang}(\tilde{G}^r) = m_1 \cdot m_2 \dots m_r$. Les éléments de \tilde{G}^p sont aussi des combinaisons linéaires des indicatrices des noeuds d'ordre r . On pourra donc aussi calculer le rang de \tilde{G}^p en tenant compte uniquement des liaisons formelles entre les éléments de \tilde{G}^p . C'est pourquoi dans la suite de cet article nous n'allons utiliser que des plans "sans biais", notion usuelle que nous allons employer ici dans le sens suivant :

Définition 3 : Un plan d'expérience est sans biais pour le modèle avec interactions d'ordre p si le rang de G^p est celui que l'on obtiendrait avec un plan dont tous les noeuds d'ordre r sont non vides.

On remarque alors qu'un plan est sans biais pour le modèle d'ordre p si et seulement si \tilde{G}^p et G^p sont de même rang. D'autre part, pour qu'un plan soit sans biais pour le modèle d'ordre p :

- il est nécessaire que tous les noeuds d'ordre p soient non vides
- il est suffisant que tous les noeuds d'ordre r soient non vides.

4.2 - Base de W^p

Notons B^p l'ensemble des éléments de G^p formés de J_0 et de toutes les indicatrices des noeuds d'ordre inférieur ou égal à p , exceptées celles où interviennent $\Omega_1^1, \Omega_2^1, \dots, \Omega_r^1$. Ainsi B^p est formé des éléments de G^p dans lesquels l'indice de modalité "1" n'intervient pas. Par exemple :

$$B^1 = \{J_0, M_1^2, \dots, M_1^{m_1} \cdot M_2^2, \dots, M_2^{m_2}, \dots, M_r^2, \dots, M_r^{m_r}\}$$

Notons $\pi(k)$ l'ensemble des k -uplets d'entiers : $\varepsilon = (i_1, \dots, i_k)$ tels que

$0 < i_1 < \dots < i_k \leq r$. Pour tout $\varepsilon = (i_1, \dots, i_k) \in \pi(k)$ notons

$$m(\varepsilon) = \prod_{j=1}^k m_{i_j} \quad \text{et} \quad m'(\varepsilon) = \prod_{j=1}^k (m_{i_j} - 1).$$

Proposition 4 :

1) Pour qu'un plan soit sans biais pour le modèle d'ordre p il faut et il suffit que B^p soit une base de W^p .

2) Dans ce cas :

$$\begin{aligned} \dim W^p &= 1 + \sum_{k=1}^p \sum_{\varepsilon \in \pi(k)} m'(\varepsilon) \\ &= (-1)^p \binom{p}{r-1} + \sum_{k=1}^p (-1)^{p-k} \binom{p-k}{r-k-1} \sum_{\varepsilon \in \pi(k)} m(\varepsilon) \quad \text{si } p < r \\ &= \prod_{i=1}^r m_i \quad \text{si } p = r. \end{aligned}$$

Ainsi $\dim W^3 = \frac{-(r-1)(r-2)(r-3)}{6} + \frac{(r-2)(r-3)}{2} \sum_{i < j} m_i m_j + \sum_{i < j < k} m_i m_j m_k$

La démonstration se fait en deux étapes :

a) On montre que B^p est toujours un système générateur de W^p , en construisant les éléments de G^p à partir de ceux de B^p . Ainsi par exemple :

$$\begin{aligned} M_1^1 &= J_0 - \sum_{j=2}^{m_1} M_1^j \\ M_{12}^{1j} &= M_2^j - \sum_{k=2}^m M_{12}^{kj} \quad \text{pour } j \neq 1 \\ M_{12}^{11} &= J_0 - \sum_{(k,\ell) \neq (1,1)} M_{12}^{k\ell} \end{aligned}$$

b) On montre que B^r est une base de W^r si le plan est sans biais et ceci en montrant que $\text{Card}(B^r) = \dim W^r$. Puisque $B^p \subset B^r$, on en déduit que les éléments de B^p sont linéairement indépendants et forment donc une base de W^p . Ce qui donne directement la dimension de W^p . La réciproque est immédiate.

D'un point de vue pratique, la métrique de W^p associée à la base B^p s'obtient à partir de la métrique de W^p associée à G^p en enlevant les lignes et les colonnes correspondant aux noeuds comportant au moins un indice de modalité égal à 1.

On remarquera que les coordonnées d'un vecteur u de W^p dans la base B^p sont en fait la représentation de u dans G^p associée à la contrainte : la

composante de u relative au noeud $\Omega_{i_1 \dots i_k}^{j_1 \dots j_k}$ est nulle dès que l'un des indices j_1, \dots, j_k est égal à 1. Il s'agit donc d'une contrainte se mettant sous la forme $A[u] = [0]$ dont nous avons déjà parlé. Nous avons vu d'autre part que toute version de β , donc en particulier celle liée à la base B^P pouvait s'écrire sous forme de fonction estimable. On peut donc obtenir la représentation dans B^P d'un élément u de W^P , notée β , connaissant une représentation quelconque δ de u dans G^P . Nous allons donner la formule pour les noeuds d'ordre p seulement. Utilisons quelques notations :

1) Soit $\varepsilon = (i_1, \dots, i_k) \in \pi(k)$. Nous noterons $K(\varepsilon)$ l'ensemble des $\eta = (j_1, \dots, j_k)$ tels que M_ε^η existe. Nous pourrions donc écrire pour tout $u \in W^P$ et pour toute représentation $[u]$ de u dans G^P :

$$u = u_o J_o + \sum_{k=1}^P \sum_{\varepsilon \in \pi(k)} \sum_{\eta \in K(\varepsilon)} u_\varepsilon^\eta M_\varepsilon^\eta$$

2) Notons $H_p(k)$ l'ensemble des p -uplets formés de k indices égaux à 1 et de $p-k$ indices égaux à zéro. Soit $h = (h_1, \dots, h_p)$, élément de $H_p(k)$, et soit $\eta = (j_1, \dots, j_p)$. Nous noterons $h(\eta)$ le p -uplet (j'_1, \dots, j'_p) défini par : $j'_i = j_i$ si $h_i = 0$; $j'_i = 1$ si $h_i = 1$.

Proposition 5 : Compte tenu des notations précédentes, quel que soit $u \in W^P$, quelle que soit sa représentation δ dans G^P , sa représentation β dans la base B^P est telle que : $\forall \varepsilon \in \pi(p)$ on a :

$$\beta_\varepsilon^\eta = \delta_\varepsilon^\eta + \sum_{k=1}^p (-1)^k \sum_{h \in H_p(k)} \delta_\varepsilon^{h(\eta)}$$

Ainsi pour $p = 3$ on a :

$$\beta_{ijk}^{\ell mn} = \delta_{ijk}^{\ell mn} - \delta_{ijk}^{\ell mn} - \delta_{ijk}^{\ell ln} - \delta_{ijk}^{\ell ml} + \delta_{ijk}^{\ell ln} + \delta_{ijk}^{\ell ml} + \delta_{ijk}^{\ell ll} - \delta_{ijk}^{\ell ll}$$

Cette proposition pourra en particulier s'appliquer à μ et $\hat{\mu}$.

La démonstration se fait en deux parties : en notant β_ε^η le nombre défini par le deuxième membre de l'égalité dans la proposition,

a) on montre que β_ε^η est nul dès que η contient un indice égal à 1.

b) on montre que $u' = u - \sum_{\varepsilon \in \pi(p)} \sum_{\eta \in K(\varepsilon)} \beta_\varepsilon^\eta M_\varepsilon^\eta$ est un élément de W^{P-1} , pouvant donc s'écrire à l'aide de B^{P-1} .

Remarques :

1) β_ϵ^η s'exprime uniquement à l'aide des termes sur les interactions d'ordre p pour $\epsilon \in \pi(p)$.

2) Pour $r = 4, p = 2, m_1 = m_2 = m_3 = m_4 = 3$ alors :

G^2 contient 67 éléments

B^2 contient 40 éléments.

On voit donc l'intérêt de cette base pour écrire le projecteur au lieu de garder G : on passe du calcul d'une inverse généralisée d'une matrice (67,67) de rang 40 au calcul de l'inverse d'une matrice (40,40) de rang 40. De plus on peut écrire très simplement toutes les hypothèses habituelles en fonction des coordonnées dans cette base.

3) Si toutes les variables $\tilde{X}_1, \dots, \tilde{X}_r$ ont une modalité "sans traitement" (qui sera codée "1") la représentation de μ (et de $\hat{\mu}$) dans la base B^p qui est une version particulière de β , (et de $\hat{\beta}$) a un sens plus physique que toute autre version de β (et de $\hat{\beta}$).

4) D'après le théorème 1, l'estimateur de Gauss-Markov de la représentation de μ dans B^p est la représentation de $\hat{\mu} = P_{W^p}(Y^*)$ dans cette même base.

4.3 - Codages centrés

Nous allons maintenant étudier un autre type de contraintes sur β . Comme dans le paragraphe précédent, dans le cadre d'un modèle avec interactions d'ordre p , nous allons étudier des contraintes portant uniquement sur les coordonnées relatives aux noeuds d'ordre p . Pour ceci, nous allons décomposer : $\mu = \mu_p + \mu_{p-1}$ avec :

a) μ_p est combinaison linéaire des indicatrices des noeuds d'ordre p dont les coefficients vérifient les contraintes recherchées.

b) μ_{p-1} est élément de W^{p-1} .

Une récurrence utilisant μ_{p-1} permettrait d'écrire β et $\hat{\beta}$ avec un vaste choix de contraintes.

Fixons $\epsilon = (i_1, \dots, i_p) \in \pi(p)$ et choisissons un système de poids $d_\epsilon^\eta = \{d_\epsilon^\eta; \eta \in K(\epsilon)\}$ tel que :

$$(1) d_\epsilon^\eta > 0$$

$$(2) \sum_{\eta \in K(\epsilon)} d_\epsilon^\eta = 1$$

On peut alors montrer que tout vecteur u de W^p possède une représentation

[u] dans G^p telle que :

$$\sum_{j_\ell} d_{i_1 \dots i_\ell \dots i_p}^{j_1 \dots j_\ell \dots j_p} u_{i_1 \dots i_\ell \dots i_p}^{j_1 \dots j_\ell \dots j_p} = 0, \quad \forall j_1 \dots j_p, \quad \forall \ell = 1, \dots, p$$

La démonstration se fait en utilisant le projecteur sur le sous-espace engendré par les noeuds d'ordre $p-1$ entre X_{i_1}, \dots, X_{i_p} , en mettant sur R^n une métrique associée à des probabilités sur I telles que :

$$P(\Omega_\varepsilon^\eta) = \delta_\varepsilon^\eta \quad \text{pour tout } \eta \in K(\varepsilon).$$

Cependant ces contraintes sont peu maniables, sauf si le système de poids vérifie certaines contraintes "d'orthogonalité", c'est-à-dire si :

(3) il existe un système de poids $d = \{d_1^1, \dots, d_1^{m_1}, d_2^1, \dots, d_r^{m_r}\}$ tq :

$$d_i^j > 0$$

$$d_{i_1 \dots i_p}^{j_1 \dots j_p} = \prod_{\ell=1}^p d_{i_\ell}^{j_\ell}$$

On remarquera qu'alors : $\sum_{j_\ell} d_{i_1 \dots i_p}^{j_1 \dots j_p} u_{i_1 \dots i_p}^{j_1 \dots j_p} = 0 \Leftrightarrow \sum_{j_\ell} d_{i_\ell}^{j_\ell} u_{i_1 \dots i_p}^{j_1 \dots j_p} = 0$

et que $\sum_{j=1}^{m_i} d_i^j = 1 \quad \forall i = 1, \dots, r$. D'où la définition :

Définition 6 : Si le système de poids $d = \{d_1^1, \dots, d_r^{m_r}\}$ vérifie : $d_i^j > 0$ et $\sum_j d_i^j = 1$ pour $i = 1, \dots, r$; si γ est une représentation de u dans G^p telle que $\forall \varepsilon \in \pi(p)$ on ait :

$$\sum_{j_\ell} d_{i_\ell}^{j_\ell} \gamma_{i_1 \dots i_\ell \dots i_p}^{j_1 \dots j_\ell \dots j_p} = 0 \quad \forall \ell = 1, \dots, p$$

nous dirons que γ est un codage centré de u associé à d .

Pour étudier les propriétés des codages centrés, nous avons besoin des notations suivantes :

1) Pour $\varepsilon = (\varepsilon_1, \dots, \varepsilon_{k_1})$ et $\varepsilon' = (\varepsilon'_1, \dots, \varepsilon'_{k_2})$ on note : $\varepsilon \varepsilon' = (\varepsilon_1 \dots \varepsilon_{k_1}, \varepsilon'_1, \dots, \varepsilon'_{k_2})$.

2) Pour $\varepsilon' = (\varepsilon'_1 \dots \varepsilon'_{k_2})$ et $\eta' = (j'_1, \dots, j'_{k_2})$ on note : $d_{\varepsilon'}^{\eta'} = \prod_{\ell=1}^{k_2} d_{\varepsilon'_\ell}^{j'_\ell}$ et pour toute représentation δ de u dans G^p :

$$\delta_{\varepsilon \varepsilon'}^{\eta \cdot} = \sum_{\eta' \in K(\varepsilon')} d_{\varepsilon'}^{\eta'} \delta_{\varepsilon \varepsilon'}^{\eta \eta'}$$

3) Notons toujours $H_p(k)$ l'ensemble des p -uplets formés de k indices égaux à 1 et de $p-k$ indices égaux à zéro. Pour $h = (h_1, \dots, h_p) \in H_p(k)$ et $\varepsilon = (i_1, \dots, i_p)$ notons :

$$h^+(\varepsilon) = (i_{j_1}, \dots, i_{j_k}) \text{ si } h_{j_i} = 1 \text{ pour } i = 1, \dots, k$$

$$h^-(\varepsilon) = (i_{\ell_1}, \dots, i_{\ell_{p-k}}) \text{ si } h_{\ell_i} = 0 \text{ pour } i = 1, \dots, p-k$$

On décompose ainsi ε en un k -uplet et en un $(p-k)$ uplet.

Proposition 7 : Soit $u \in W^P$. Si le plan est sans biais :

1) u possède un et un seul codage centré pour chaque système de poids d (unicité uniquement des coordonnées relatives aux noeuds d'ordre p).

2) Soit γ un codage centré de u , et soit δ une représentation quelconque de u dans G^P . Alors :

$$\gamma_\varepsilon^\eta = \sum_{\ell=0}^p (-1)^{p-\ell} \sum_{h \in H_p(\ell)} \delta^{h^+(\eta)} \cdot h^+(\varepsilon) h^-(\varepsilon) \quad \forall \varepsilon \in \pi(p) \text{ et } \forall \eta \in K(\varepsilon)$$

Ainsi :

$$\gamma_{ijk}^{\ell mn} = \delta_{ijk}^{\ell mn} - \delta_{ijk}^{\ell mn} - \delta_{ijk}^{\ell n} - \delta_{ijk}^{\ell m} + \delta_{ijk}^{\ell n} + \delta_{ijk}^{\ell m} + \delta_{ijk}^{\ell} - \delta_{ijk}^{\ell}$$

La démonstration de la proposition peut se faire d'après l'étude des plans orthogonaux que nous verrons plus loin.

On remarquera en particulier que γ_ε^η , pour $\varepsilon \in \pi(p)$, est ainsi mis sous forme de fonction estimable ne dépendant que des coordonnées relatives aux seuls noeuds d'ordre p . De plus, le théorème 1 nous indique que l'estimateur de Gauss-Markov d'un codage centré de μ est le codage centré analogue de $\hat{\mu} = P_{W^P}(Y^*)$.

4.4 - Fonctions estimables

Soit K une matrice $(1, m)$ où $m = \text{Card}(G^P)$. Nous allons chercher à quelles conditions $K\beta$ est estimable. Notons $q = \dim W^P$.

Nous allons utiliser une propriété des fonctions estimables énoncée au paragraphe 3 : $K\beta$ est estimable si et seulement si $KZ = [0]$ où Z est une matrice dont les colonnes engendrent toutes les représentations du vecteur nul dans G^P . Ici on remarque que l'on parcourt toutes les écritures possibles du vecteur nul en écrivant l'ensemble des relations :

$$M_{i_1 \dots i_{\ell-1} i_{\ell+1} \dots i_k}^{j_1 \dots j_{\ell-1} j_{\ell+1} \dots j_k} - \sum_{j_\ell} M_{i_1 \dots i_{\ell-1} i_{\ell+1} \dots i_k}^{j_1 \dots j_{\ell-1} j_\ell \dots j_k} = 0$$

Supposons que les éléments de la matrice K soient numérotés de telle sorte que : $K\beta = K_0 \beta_0 + \sum_{k=1}^p \sum_{\epsilon \in \eta} \sum_{\epsilon} K^{\eta} \beta^{\eta}$. Alors :

Proposition 8 : Si le plan d'expérience est sans biais, $K\beta$ est estimable si et seulement si :

1) $\forall i_1 < i_2 < \dots < i_p, \forall \ell \leq k \leq p, \forall j_1, \dots, j_k :$

$$K_{i_1 \dots i_{\ell-1} i_{\ell+1} \dots i_k}^{j_1 \dots j_{\ell-1} j_{\ell+1} \dots j_k} = \sum_{j_\ell} K_{i_1 \dots i_{\ell-1} i_{\ell+1} \dots i_k}^{j_1 \dots j_{\ell-1} j_\ell \dots j_k}$$

2) $\forall i = 1, \dots, r : K_0 = \sum_{j=1}^{m_i} K_i^j$

Dans le modèle avec interactions d'ordre p , toute fonction estimable non identiquement nulle contient donc des coefficients non nuls relativement aux noeuds d'ordre p . Et il existe des fonctions estimables portant uniquement sur les coordonnées relatives aux noeuds d'ordre p .

4.5 - Non existence d'interactions

Considérons toujours le modèle avec interactions d'ordre p . Etant données les liaisons linéaires entre les éléments de G^p , tout élément de W^p , même le vecteur nul, possède une infinité de représentations pour lesquelles les composantes $\beta_{i_1 \dots i_p}^{j_1 \dots j_p}$ sont non nulles. C'est pourquoi nous avons besoin de la définition suivante :

Définition 9 :

1) Nous dirons qu'il n'y a pas d'interactions d'ordre p s'il existe une version de β telle que :

$$\beta_{i_1 \dots i_p}^{j_1 \dots j_p} = 0 \quad \forall i_1 \dots i_p \quad \text{et} \quad \forall j_1 \dots j_p$$

2) Nous dirons qu'il n'y a pas d'interactions d'ordre p entre $\tilde{X}_{i_1}, \dots, \tilde{X}_{i_p}$ s'il existe une version de β telle que :

$$\beta_{i_1 \dots i_p}^{j_1 \dots j_p} = 0 \quad \forall j_1 \dots j_p$$

Des propositions 5 et 7 on tire, en supposant le plan sans biais :

Proposition 10 :

1) Les trois affirmations suivantes sont équivalentes :

a) il n'y a pas d'interactions d'ordre p entre $\tilde{X}_{i_1}, \dots, \tilde{X}_{i_p}$

b) la représentation β de μ dans la base B^p vérifie :

$$\beta_{i_1 \dots i_p}^{j_1 \dots j_p} = 0 \quad \forall j_1 \dots j_p$$

c) un codage centré quelconque γ de μ vérifie :

$$\gamma_{i_1 \dots i_p}^{j_1 \dots j_p} = 0 \quad \forall j_1 \dots j_p$$

2) Les deux affirmations suivantes sont équivalentes :

a) il n'y a pas d'interactions d'ordre p

b) quel que soient $i_1 < i_2 < \dots < i_p \leq r$, il n'y a pas d'interactions d'ordre p entre $\tilde{X}_{i_1}, \dots, \tilde{X}_{i_p}$.

En effet, dans les propositions 5 et 7, $\beta_{i_1 \dots i_p}^{j_1 \dots j_p}$ et $\gamma_{i_1 \dots i_p}^{j_1 \dots j_p}$ sont exprimées en fonction d'éléments du type $\delta_{i_1 \dots i_p}^{\ell_1 \dots \ell_p}$ uniquement.

De cette proposition on déduit facilement les degrés de liberté associés à ces hypothèses, puisqu'on les a exprimées en fonction des coordonnées de μ dans la base B^p . Ainsi notons : $q = \dim W^p$.

$m'(\epsilon) = \prod_{\ell=1}^p (m_{i_\ell} - 1)$ pour $\epsilon = (i_1, \dots, i_p)$, qui est le nombre d'éléments de B^p relatifs aux noeuds d'ordre p entre $\tilde{X}_{i_1}, \dots, \tilde{X}_{i_p}$.

Alors le nombre de degrés de liberté associés à l'hypothèse "il n'y a pas d'interactions d'ordre p entre $\tilde{X}_{i_1}, \dots, \tilde{X}_{i_p}$ " est égal à $q - m'(\epsilon)$, et celui associé à l'hypothèse "il n'y a pas d'interactions d'ordre p " est égal à $q - \sum_{\epsilon \in \pi(p)} m'(\epsilon)$.

Il est possible de généraliser ces résultats à des réunions d'hypothèses du type "il n'y a pas d'interactions d'ordre supérieur à k entre $\tilde{X}_{i_1}, \dots, \tilde{X}_{i_k}$ " pour un certain $k \leq p$.

5 - PLANS ORTHOGONAUX

On sait que pour certains plans d'expérience particulièrement bien conçus, appelés "orthogonaux", on peut obtenir un estimateur de β et les statistiques des tests usuels sans inverser une seule matrice, même pour p et r grands. Il ne nous semble pas qu'il y ait d'exposé général des propriétés géométriques de W^p expliquant ce phénomène. Seuls les plans sans interactions nous semblent avoir été abordés.

Notons $W_{i_1 \dots i_k}$ le sous espace de W^p engendré par les indicatrices des noeuds d'ordre k entre X_{i_1}, \dots, X_{i_k} . Notons $P_{W_{i_1 \dots i_k}}$ le projecteur orthogonal de R^n sur $W_{i_1 \dots i_k}$.

Notons $P(\Omega_{i_1 \dots i_k}^{j_1 \dots j_k})$ la probabilité (empirique) de $\Omega_{i_1 \dots i_k}^{j_1 \dots j_k}$ et $n_{i_1 \dots i_k}^{j_1 \dots j_k}$ l'effectif de ce noeud.

$$\text{Notons } E(Y^*/\Omega_{i_1 \dots i_k}^{j_1 \dots j_k}) = \frac{1}{P(\Omega_{i_1 \dots i_k}^{j_1 \dots j_k})} \langle Y^*, M_{i_1 \dots i_k}^{j_1 \dots j_k} \rangle = \frac{1}{n_{i_1 \dots i_k}^{j_1 \dots j_k}} \sum_{i_\ell \in \Omega_{i_1 \dots i_k}^{j_1 \dots j_k}} \tilde{Y}_\ell$$

Alors :

1) les indicatrices des noeuds d'ordre k entre X_{i_1}, \dots, X_{i_k} forment une base orthogonale de $W_{i_1 \dots i_k}$.

2) $\|M_\epsilon^\eta\|^2 = P(\Omega_\epsilon^\eta)$.

On en déduit donc que :

$$P_{W_\epsilon} (Y^*) = \sum_{\eta \in K(\epsilon)} E(Y^*/\Omega_\epsilon^\eta) M_\epsilon^\eta$$

qui est un projecteur facile à mettre en oeuvre. Nous allons alors décomposer W^p en sous espaces orthogonaux dont le projecteur s'exprime simplement en fonction des projecteurs du type précédent. Mais pour cela nous devons faire certaines hypothèses sur le plan d'expérience.

Nous en déduirons une version particulière de β et de $\hat{\beta}$ que nous appellerons "codage centré adapté" qui sera facile à calculer et dont la matrice de covariance est simple. C'est la version de β qui est habituellement utilisée.

5.1 - Décomposition de W^p

Définition 11 : Un plan d'expérience est orthogonal pour le modèle avec interac-

tions d'ordre p , si pour $t = \inf(2p, r)$, pour tout $i_1 < i_2 < \dots < i_t \leq r$ et pour tout $j_1 \dots j_t$ on a : $P(\Omega_{i_1 \dots i_t}^{j_1 \dots j_t}) = \prod_{\ell=1}^t P(\Omega_{i_\ell}^{j_\ell})$.

Ainsi :

Un plan orthogonal possède tous ses noeuds d'ordre t non vides

Un plan orthogonal pour le modèle d'ordre p l'est aussi pour le modèle d'ordre $p-1$.

Les plans complets ayant même effectif dans chaque noeud d'ordre r sont orthogonaux quel que soit $p \leq r$.

Les plans obtenus par la méthode des carrés latins et des carrés latins orthogonaux (pour $m_1 = m_2 = \dots = m_r$) sont orthogonaux pour le modèle sans interactions.

Il existe d'autres types assez simples de plans orthogonaux. Les plans incomplets équilibrés ne sont pas orthogonaux.

Notons :

$G_{i_1 \dots i_k}$ l'ensemble des indicatrices des noeuds d'ordre k entre X_{i_1}, \dots, X_{i_k}

$W_{i_1 \dots i_k}$ le sous espace engendré par les éléments de $G_{i_1 \dots i_k}$

$$W^k = \sum_{\epsilon \in \pi(k)} W_\epsilon$$

W_\perp^{k-1} l'orthogonal de W^{k-1} dans W^k

$W_{i_1 \dots i_k}^{k-1} = \left\{ \begin{array}{l} j_1 < \dots < j_{k-1} \\ j_\ell \in \{i_1, \dots, i_k\} \end{array} \right\} W_{j_1 \dots j_{k-1}}$ qui est le sous espace engendré par les indicatrices des noeuds d'ordre $k-1$ entre X_{i_1}, \dots, X_{i_k} . Il est donc dans $W_{i_1 \dots i_k}$.

$V_{i_1 \dots i_k}$ l'orthogonal de $W_{i_1 \dots i_k}^{k-1}$ dans $W_{i_1 \dots i_k}$

$V'_{i_1 \dots i_k}$ la trace dans $W_{i_1 \dots i_k}$ de l'orthogonal de W^{k-1}

J l'axe des variables constantes sur I . V_i et V'_i désignerons l'orthogonal de J dans W_i .

Proposition 12 : Si le plan est orthogonal pour le modèle d'ordre p :

- 1) le plan est sans biais pour ce modèle
- 2) $V_{i_1 \dots i_k} = V'_{i_1 \dots i_k} \quad \forall k \leq p$
- 3) les sous espaces de type $V_{i_1 \dots i_k}$ pour $k \leq p$ sont orthogonaux

$$4) W_{\perp}^{k-1} = \bigoplus_{i_1 < i_2 < \dots < i_k} V_{i_1 \dots i_k} \text{ pour } k \leq p$$

$$5) W^p = \bigcup_{k=1}^p \bigoplus_{i_1 < \dots < i_k} V_{i_1 \dots i_k} \text{ est une décomposition de } W^p \text{ en sous espaces orthogonaux.}$$

En particulier on déduit de cette proposition que pour $\varepsilon \neq \varepsilon'$ et si on note : $H = W_{\varepsilon} \cap W_{\varepsilon'}$,

H_1 : l'orthogonal de H dans W_{ε}

H_2 : l'orthogonal de H dans $W_{\varepsilon'}$,

alors H_1 et H_2 sont orthogonaux si le plan est orthogonal.

La démonstration peut se faire selon les étapes 2-3-4-5-1.

Dans ce cas la difficulté essentielle est de montrer que le 2^{ème} membre de la relation indiquée en 4 contient effectivement W_{\perp}^{k-1} . Démontrons par exemple la 2^{ème} affirmation, c'est-à-dire que $V_{\varepsilon} = V'_{\varepsilon}$ pour $\varepsilon = (\varepsilon_1, \dots, \varepsilon_k)$.

1) par construction $V'_{\varepsilon} \subset V_{\varepsilon}$

2) pour montrer la réciproque, il suffit de montrer que tout élément $u \in V_{\varepsilon}$ est orthogonal aux indicatrices du type $M_{i_1 \dots i_{k-1}}^{j_1 \dots j_{k-1}}$ qui engendrent W^{p-1} .

a) Supposons que i_1, \dots, i_{k-1} soient éléments de $\{\varepsilon_1, \dots, \varepsilon_k\}$. Dans ce cas tout élément de $V_{\varepsilon_1 \dots \varepsilon_k}$ sera par définition orthogonal à $M_{i_1 \dots i_{k-1}}^{j_1 \dots j_{k-1}}$.

b) Supposons que pour $q \neq k-1$ on ait : i_1, \dots, i_q sont éléments de $\{\varepsilon_1, \dots, \varepsilon_k\}$, et pour plus de simplicité supposons $i_1 = \varepsilon_1; \dots; i_q = \varepsilon_q$.

i_{q+1}, \dots, i_k ne sont pas éléments de $\{\varepsilon_1, \dots, \varepsilon_k\}$.

Soit $u \in V_{\varepsilon}$. On peut écrire $u = \sum_{l_1 \dots l_k} u_{l_1 \dots l_k} M_{\varepsilon}^{l_1 \dots l_k}$. Donc :

$$\begin{aligned} \langle u, M_{i_1 \dots i_{k-1}}^{j_1 \dots j_{k-1}} \rangle &= \sum_{l_1 \dots l_k} u_{l_1 \dots l_k} \langle M_{\varepsilon}^{l_1 \dots l_k}, M_{i_1 \dots i_{k-1}}^{j_1 \dots j_{k-1}} \rangle \\ &= \sum_{l_{q+1} \dots l_k} u_{j_1 \dots j_q l_{q+1} \dots l_k} P(\Omega_{i_1 \dots i_{k-1}}^{j_1 \dots j_{k-1}} \cap \Omega_{\varepsilon_{q+1} \dots \varepsilon_k}^{l_{q+1} \dots l_k}) \end{aligned}$$

et les indices $i_1, \dots, i_{k-1}, \varepsilon_{q+1}, \dots, \varepsilon_k$, en nombre inférieur à $2p$ sont tous différents. Si la plan est orthogonal on aura donc :

$$(1) \langle u, M_{i_1 \dots i_{k-1}}^{j_1 \dots j_{k-1}} \rangle = P(\Omega_{i_1 \dots i_{k-1}}^{j_1 \dots j_{k-1}}) \sum_{l_{q+1} \dots l_k} u_{j_1 \dots j_q l_{q+1} \dots l_k} P(\Omega_{\varepsilon_{q+1} \dots \varepsilon_k}^{l_{q+1} \dots l_k})$$

Or par hypothèse u est orthogonal à $M_{i_1 \dots i_q}^{j_1 \dots j_q}$ puisque cette indicatrice est élément de W_C^{p-1} . Après un calcul semblable au précédent on obtient :

$$0 = \langle u, M_{i_1 \dots i_q}^{j_1 \dots j_q} \rangle = P(\Omega_{i_1 \dots i_q}^{j_1 \dots j_q}) \sum_{l_{q+1} \dots l_k} u_{j_1 \dots j_q l_{q+1} \dots l_k} P(\Omega_{l_{q+1} \dots l_k}^{l_{q+1} \dots l_k})$$

Et en se reportant dans (1) on en déduit que $\langle u, M_{i_1 \dots i_{k-1}}^{j_1 \dots j_{k-1}} \rangle = 0$ puisque $P(\Omega_{i_1 \dots i_q}^{j_1 \dots j_q})$ est non nul.

Pour démontrer la 3^{ème} partie, on utilise le même type de raisonnement. On choisit $u \in V_\varepsilon$, $v \in V_{\varepsilon'}$. On développe le produit $\langle u, v \rangle$ à l'aide des indicatrices du type M_ε^η et $M_{\varepsilon'}^\eta$. On utilise enfin le fait que $V_\varepsilon = V_{\varepsilon'}$, donc que u est orthogonal au sous espace engendré par les indicatrices des noeuds d'ordre $k-1$ si $\varepsilon = (\varepsilon_1, \dots, \varepsilon_k)$.

Pour démontrer la 4^{ème} partie, on utilise le lemme :

Lemme :

Soit H un espace de Hilbert

Soient H_1, \dots, H_k des sous espace fermés engendrant H

Soient E_1, \dots, E_k des sous espaces fermés de H_1, \dots, H_k engendrant un sous espace de H noté E .

Soient $E_1^\perp, \dots, E_k^\perp$ les orthogonaux de E_1, \dots, E_k dans H_1, \dots, H_k respectivement.

Soient $E_1'^\perp, \dots, E_k'^\perp$ les traces dans H_1, \dots, H_k de l'orthogonal de E dans H .

Soit E^\perp l'orthogonal de E dans H

Alors, si $E_i^\perp = E_i'^\perp$ pour $i = 1, \dots, k$, on a $E^\perp = \sum_{i=1}^k E_i^\perp$.

On en déduit facilement la 5^{ème} partie. Quant à la première partie elle s'obtient en calculant la dimension de V_ε et en remarquant qu'alors la dimension de W^p est bien celle indiquée dans la proposition 4.

De cette proposition on déduit facilement l'expression des projecteurs recherches. Notons :

$\pi(k)$ l'ensemble des k -uplets $\varepsilon = \{i_1, \dots, i_k\}$ avec $0 < i_1 < \dots < i_k \leq r$.

$\pi(0) = \{0\}$ et $J = V_0 = W_0$

$\pi(l, \varepsilon)$ l'ensemble des l -uplets formés d'indices de ε , placés en ordre croissant

Pour tout espace vectoriel E de \mathbb{R}^n , P_E désignera le projecteur orthogonal de \mathbb{R}^n sur E .

Proposition 13 : Si le plan est orthogonal pour le modèle d'ordre p :

1) pour tout $k \leq p$ et tout $\varepsilon \in \pi(k)$:

$$P_{V_\varepsilon} = P_{W_\varepsilon} - \sum_{\ell=0}^{k-1} \sum_{\varepsilon' \in \pi(\ell, \varepsilon)} P_{V_{\varepsilon'}} = \sum_{\ell=0}^k (-1)^{k-\ell} \sum_{\varepsilon' \in \pi(\ell, \varepsilon)} P_{W_{\varepsilon'}}$$

2) pour tout $k < p$ et tout $k < r$:

$$P_{W^k} = \sum_{j=0}^k (-1)^{k-j} \binom{k-j}{r-j-1} \sum_{\varepsilon' \in \pi(j)} P_{W_{\varepsilon'}} = \sum_{j=0}^k \sum_{\varepsilon' \in \pi(j)} P_{V_{\varepsilon'}}$$

La première forme de P_{V_ε} et la deuxième forme de P_{W^k} sont des conséquences directes de la proposition précédente. La deuxième forme de P_{V_ε} se démontre par récurrence en partant de la première forme : on utilise la formule de récurrence, on regroupe les termes en P_{W_ε} , et on remarque que : $\sum_{\ell=j}^k (-1)^{\ell-j} \binom{\ell-j}{k-j+1} = (-1)^{k-j}$. Enfin la première forme de P_{W^k} s'obtient en utilisant la première forme de P_{V_ε} dans la seconde forme de P_{W^k} et en regroupant les termes semblables.

Pour $p = r$, P_{W^p} ne pose aucun problème, puisque les indicatrices des noeuds d'ordre r sont une base orthogonale de W^r . Comme applications, on obtient :

$$P_{W^1} = \sum_{i=1}^r P_{W_i} - (r-1)P_J$$

$$P_{W^2} = \sum_{i < j} P_{W_{ij}} - (r-2) \sum_i P_{W_i} + \frac{(r-1)(r-2)}{2} P_J$$

$$P_{W^3} = \sum_{i < j < k} P_{W_{ijk}} - (r-3) \sum_{i < j} P_{W_{ij}} + \frac{(r-2)(r-3)}{2} \sum_i P_{W_i} - \frac{(r-1)(r-2)(r-3)}{6} P_J$$

5.2 - Codages centrés adaptés

Nous avons déjà défini la notion de codage centré. Nous allons ici choisir un système de poids adapté au plan d'expérience.

Définition 14 : Soit $u \in W^p$ et soit γ une représentation de u dans G^p . Nous dirons que γ est un codage centré adapté de u si :

$$\forall k = 1, \dots, p, \forall i_1, \dots, i_k, \forall j_1, \dots, j_k, \forall \ell = 1, \dots, k \text{ alors :}$$

$$\sum_{j_\ell} P(\Omega_{i_\ell}^{j_\ell}) \gamma_{i_1 \dots i_\ell \dots i_k}^{j_1 \dots j_\ell \dots j_k} = 0$$

On peut alors facilement montrer que le codage centré adapté γ de u est lié à la décomposition de W^p en sous espaces orthogonaux. Plus exactement :

$$P_{V_\epsilon}(u) = \sum_{\eta \in K(\epsilon)} \gamma_\epsilon^\eta M_\epsilon^\eta$$

Des deux écritures de P_{V_ϵ} dans la proposition 13 on déduit que :

Proposition 15 : Si le plan est orthogonal pour le modèle d'ordre p :

- 1) Tout élément u de W^p possède un et un seul codage centré adapté.
- 2) Ce codage centré adapté s'obtient par :
 - a) $\gamma_0 = E(u/I) = \langle u, J_0 \rangle$ ($\gamma_0 =$ coefficient de J_0)
 - b) pour $\epsilon \in \pi(k)$ et $\eta \in K(\epsilon)$:

$$\begin{aligned} \gamma_\epsilon^\eta &= E(u/\Omega_\epsilon^\eta) + \sum_{\ell=1}^{k-1} (-1)^{k-\ell} \sum_{h \in H_k(\ell)} E(u/\Omega_{h^+}^{\eta^+}(\eta)) \\ &= E(u/\Omega_\epsilon^\eta) - \sum_{\ell=0}^{k-1} \sum_{h \in H_k(\ell)} \gamma_{h^+}^{\eta^+}(\eta) \end{aligned}$$

Nous avons employé les notations suivantes :

$$1) E(u/\Omega_\epsilon^\eta) = \frac{1}{n_\epsilon^\eta} \sum_{i_\ell \in \Omega_\epsilon^\eta} u(i_\ell) = \frac{1}{P(\Omega_\epsilon^\eta)} \langle u, M_\epsilon^\eta \rangle$$

2) H_k et h^+ sont définis à la proposition 7.

De plus, si δ est une représentation quelconque de u dans G^p , alors le codage centré adapté de u s'obtient facilement. Utilisons toujours les notations de la proposition 7, en notant de plus : $\pi^c(\ell, \epsilon)$ l'ensemble des ℓ -uplets formés d'indices $0 < i_1 \dots < i_\ell < r$ non inclus dans ϵ . Alors :

Proposition 16 : $\forall k \leq p, \forall \epsilon \in \pi(k)$ et $\forall \eta \in K(\epsilon)$:

$$\gamma_\epsilon^\eta = \sum_{\ell=0}^k (-1)^{k-\ell} \sum_{h \in H_k(\ell)} \sum_{\ell'=0}^{p-k} \sum_{\epsilon'' \in \pi^c(\ell', \epsilon)} \delta_{h^+}^{\eta^+}(\eta) \cdot \delta_{h^+}^{\eta^+}(\epsilon) \epsilon''$$

Pour obtenir cette écriture, on utilise la première forme de γ_ϵ^η donnée dans la proposition 15. On écrit ensuite que :

$$E(u/\Omega_{\epsilon'}^{\eta'}) = \frac{1}{P(\Omega_{\epsilon'}^{\eta'})} \langle u, M_{\epsilon'}^{\eta'} \rangle$$

et on développe le produit scalaire en utilisant les coordonnées δ .

En application des deux dernières propositions on a :

- quel que soit $p \geq 2$:

$$\gamma_{ij}^{kl} = E(u/\Omega_{ij}^{kl}) - E(u/\Omega_i^k) - E(u/\Omega_j^l) + E(u/I) = E(u/\Omega_{ij}^{kl}) - \gamma_i^k - \gamma_j^l - \gamma_0$$

- pour $p = 3$:

$$\gamma_{ij}^{lm} = \sum_{k \neq i, j} (\delta_{ijk}^{lm} - \delta_{ijk}^{lm} - \delta_{ijk}^{lm} + \delta_{ijk}^{lm}) + \delta_{ij}^{lm} - \delta_{ij}^{lm} - \delta_{ij}^{lm} + \delta_{ij}^{lm}$$

De plus, c'est à partir des propositions 15 et 16 que l'on peut démontrer la proposition 7. On utilise pour cela un plan hypothétique, orthogonal, ayant justement comme probabilités marginales les poids utilisés dans la définition du codage centré.

Nous allons bien entendu appliquer les propositions 15 et 16 à μ et $\hat{\mu}$: μ a un codage centré adapté, qui est donc une version de β que nous noterons γ . Nous savons que son estimateur $\hat{\gamma}$ de Gauss-Markov sera le codage centré adapté de $\hat{\mu} = P_{WP}(Y^*)$. Pour $\varepsilon = (i_1, \dots, i_k) \in \pi(k)$ notons :

$$\hat{\gamma}_\varepsilon^t = [\hat{\gamma}_\varepsilon^{11\dots 1}, \dots, \hat{\gamma}_\varepsilon^{m_{i_1} \dots m_{i_k}}]$$

Q_ε la matrice diagonale dont les éléments diagonaux sont les probabilités empiriques de $\Omega_\varepsilon^{1\dots 1}, \dots, \Omega_\varepsilon^{m_{i_1} \dots m_{i_k}}$. Q_ε est donc la métrique de W_ε associée à G_ε .

Proposition 17 : $\forall k \leq p, \forall \varepsilon \in \pi(k)$ et $\forall \eta \in K(\varepsilon)$:

$$1) \hat{\gamma}_\varepsilon^\eta = \sum_{\ell=0}^k (-1)^{k-\ell} \sum_{h \in H_k(\ell)} E(Y^*/\Omega_{h^+(\varepsilon)}^{h^+(\eta)}) = E(Y^*/\Omega_\varepsilon^\eta) - \sum_{\ell=0}^{k-1} \sum_{h \in H_k(\ell)} \hat{\gamma}_{h^+(\varepsilon)}^{h^+(\eta)}$$

2) $\hat{\gamma}_\varepsilon^\eta$ est une variable aléatoire normale, $\hat{\gamma}_\varepsilon$ et $\hat{\gamma}_{\varepsilon'}$ sont indépendantes dès que $\varepsilon \neq \varepsilon'$.

$$3) \text{cov}(\hat{\gamma}_\varepsilon^\eta, \hat{\gamma}_{\varepsilon'}^{\eta'}) = \frac{\sigma^2}{n} \sum_{\ell=0}^k (-1)^{k-\ell} \sum_{\left\{ \begin{array}{l} h \in H_k(\ell) \\ h^+(\eta) = h^+(\eta') \end{array} \right.} \frac{1}{P(\Omega_{h^+(\varepsilon)}^{h^+(\eta)})}$$

On note ici $H_k(0) = \{h_0\}$ et $h_0^+(\varepsilon) = 0, \Omega_0^0 = I$.

$$4) Q_\varepsilon \text{ est un inverse généralisé de } \frac{n}{\sigma^2} \text{cov}(\hat{\gamma}_\varepsilon)$$

Ainsi, si $\delta_{kk'}$ est le symbole de Kroneker :

$$\text{cov}(\hat{\gamma}_{ij}^{k\ell}, \hat{\gamma}_{ij}^{k'\ell'}) = \frac{\sigma^2}{n} \left(\frac{\delta_{kk'} \delta_{\ell\ell'}}{P(\Omega_{ij}^{k\ell})} - \frac{\delta_{\ell\ell'}}{P(\Omega_j^{\tilde{\ell}})} - \frac{\delta_{kk'}}{P(\Omega_i^{k'})} + 1 \right)$$

La démonstration se fait en remarquant que :

$$a) E(Y^*/\Omega_\varepsilon^\eta) = \frac{1}{P(\Omega_\varepsilon^\eta)} \langle Y^*, M_\varepsilon^\eta \rangle = \frac{1}{P(\Omega_\varepsilon^\eta)} \langle P_{WP}(Y^*), M_\varepsilon^\eta \rangle$$

b) Par définition $P_{V_\varepsilon}(Y^*) = \sum_{\eta} \hat{\gamma}_{\varepsilon}^{\eta} M_{\varepsilon}^{\eta}$ (18). Un résultat général permet alors de montrer que :

$$\text{cov}(\hat{\gamma}_{\varepsilon}^{\eta}, \hat{\gamma}_{\varepsilon'}^{\eta'}) = \frac{\sigma^2}{n} \frac{1}{P(\Omega_\varepsilon^\eta) P(\Omega_{\varepsilon'}^{\eta'})} \langle P_{V_\varepsilon}(M_\varepsilon^\eta), P_{V_{\varepsilon'}}(M_{\varepsilon'}^{\eta'}) \rangle$$

qu'il suffit alors de développer (uniquement dans le cas où $\varepsilon = \varepsilon'$).

5.3 - Quelques tests

On retrouve bien entendu les expressions habituelles pour les calculs du F de Snédécour. Calculons en effet $\|Y^* - P_{WP}(Y^*)\|^2$ et $\|P_{WP}(Y^*) - P_{\mathcal{L}}(Y^*)\|^2$ pour les espaces \mathcal{L} associés aux hypothèses les plus simples :

1) Compte tenu de la relation (18) précédente :

$$\|P_{V_\varepsilon}(Y^*)\|^2 = \sum_{n \in K(\varepsilon)} P(\Omega_\varepsilon^n) (\hat{\gamma}_\varepsilon^n)^2$$

2) Puisque les sous espaces V_ε sont orthogonaux :

$$\|Y^* - P_{WP}(Y^*)\|^2 = \|Y^*\|^2 - \sum_{k=0}^p \sum_{\varepsilon \in \Pi(k)} \|P_{V_\varepsilon}(Y^*)\|^2$$

3) L'hypothèse "il n'y a pas d'interactions d'ordre p entre $\tilde{X}_{i_1}, \dots, \tilde{X}_{i_p}$ " est équivalente à : " μ est orthogonal à V_ε ", puisque γ_ε^η sera nul d'après la proposition 10. Le sous espace affine \mathcal{L} associé à cette hypothèse sera donc l'orthogonal de V_ε dans W^p . D'où :

$$\|P_{WP}(Y^*) - P_{\mathcal{L}}(Y^*)\|^2 = \|P_{V_\varepsilon}(Y^*)\|^2 \quad \text{si } \varepsilon = (i_1, \dots, i_p)$$

4) L'hypothèse "il n'y a pas d'interactions d'ordre p" est associée au sous espace vectoriel $\mathcal{L} = W^{p-1}$. On aura donc :

$$\|P_{WP}(Y^*) - P_{\mathcal{L}}(Y^*)\|^2 = \sum_{\epsilon \in \pi(p)} \|P_{V_\epsilon}(Y^*)\|^2$$

5) Pour les tests d'hypothèses plus générales $K\beta = k$, nous possédons la matrice de covariance de $\hat{\gamma}$ et un inverse généralisé de celle-ci, ce qui simplifie les calculs si K est diagonale. En effet si on exprime l'hypothèse en fonction du codage centré γ et sous forme non nécessairement estimable, nous avons vu au § 3 que :

$$\|P_W(Y^*) - P_{\mathcal{L}}(Y^*)\|^2 = (K\hat{\gamma} - k)^t (K \cdot \text{cov}(\hat{\gamma}) K^t)^{-} (K\hat{\gamma} - k)$$

6 - MAXIMISATION DE LA PUISSANCE DE CERTAINS TESTS

Nous allons chercher des plans d'expérience maximisant la puissance des tests construits au paragraphe 3. En général, un plan d'expérience maximisant, pour n fixé, la puissance du test de : " $K\beta = k_0$ " contre " $K\beta = k_1$ " ne donne pas entière satisfaction, car les plans obtenus :

- 1) sont fonction de l'alternative si K a plus d'une ligne.
- 2) sont en général liés à un seul test, de sorte qu'il est généralement impossible de satisfaire aux contraintes de plusieurs tests.
- 3) sont parfois trop particuliers. Ainsi pour maximiser la puissance du test de " $\beta_1^1 = \beta_1^2$ " pour un modèle sans interactions, il faudrait répartir tous les individus du plan dans les deux premières modalités de X_1 .

C'est pourquoi, comme c'est le cas dans la littérature anglosaxonne et ailleurs [3], nous allons nous intéresser à un problème de maximisation sous contraintes. Nous allons fixer certains effectifs marginaux du plan, et nous allons chercher des plans maximisant la puissance de certaines classes de tests, indépendamment de l'alternative. Nous allons donner quelques résultats généraux, que nous appliquerons ensuite pour montrer que les plans orthogonaux ont de bonnes propriétés dans le sens.

6.1 - Notations

Reprenons les notations générales : μ est par hypothèse élément d'un sous espace W de R^n engendré par $G = g_1, \dots, g_m$. β est la dénomination des différentes représentations de μ dans G .

1) Soit $\psi(K, \cdot)$ une fonction estimable. Nous noterons :

$H(K, k)$ l'hypothèse : $K\beta = k$

$\mathcal{E}(K, k)$ le sous espace de W défini par : $x \in \mathcal{E}(K, k) \Leftrightarrow \psi[x] = k$

$x(K, k)$ la projection du vecteur nul sur $\mathcal{E}(K, k)$

2) Nous noterons $\mathcal{H}(\ell)$ l'ensemble des hypothèses $H(K, k)$ où :

$K\beta$ est estimable

seules les ℓ premières colonnes de K peuvent être non nulles. Nous noterons alors K_1 les ℓ premières colonnes de K .

3) Fixons ℓ et notons :

W_1 le sous espace engendré par $G_1 = \{g_1, \dots, g_\ell\}$

W_2 le sous espace engendré par $G_2 = \{g_{\ell+1}, \dots, g_m\}$

Q_1 la métrique de W_1 associée à G_1

Q_2 la métrique de W_2 associée à G_2

4) Pour toute hypothèse $H(K, k) \in \mathcal{H}(\ell)$ notons :

$\mathcal{E}_1(K, k)$ le sous espace affine de W_1 : $\mathcal{E}_1(K, k) = \mathcal{E}(K, k) \cap W_1$

$x_1(K, k)$ la projection de l'origine sur $\mathcal{E}_1(K, k)$.

5) Fixons n . Soit Q_1 une métrique possible de W_1 . Nous noterons \mathcal{P}_{Q_1} l'ensemble des plans d'expérience tels que :

a) tout plan de \mathcal{P}_{Q_1} est sans biais.

b) pour tout plan de \mathcal{P}_{Q_1} , la métrique de W_1 associée à G_1 est Q_1 .

Notre objectif va être de rechercher dans \mathcal{P}_{Q_1} des plans maximisant la puissance des tests de $H(K, k_0)$ contre $H(K, k_1)$, pour tout $H(K, k_0)$ et $H(K, k_1)$ éléments de $\mathcal{H}(\ell)$.

6.2 - Résultats généraux

Proposition 19 :

1) La puissance du test de $H(K, k_0)$ contre $H(K, k_1)$ ne dépend que de n, σ , $\dim W$, $\dim \mathcal{E}(K, k_0)$ et de $\|x(K, k_1 - k_0)\|$. De plus c'est une fonction croissante de $\|x(K, k_1 - k_0)\|$.

2) $x_1(K, k_1 - k_0)$ et $\|x_1(K, k_1 - k_0)\|^2 = (k_1 - k_0)^t (K_1 Q_1^{-1} K_1^t)^{-1} (k_1 - k_0)$ ne dépendent pas du plan de \mathcal{P}_{Q_1} .

3) $\|x_1(K, k_1 - k_0)\| \geq \|x(K, k_1 - k_0)\|$

4) $\|x_1(K, k_1 - k_0)\| = \|x(K, k_1 - k_0)\|$ si et seulement si $x_1(K, k_1 - k_0)$ est orthogonal à W_2 . On a alors : $x(K, k_1 - k_0) = x_1(K, k_1 - k_0)$.

La démonstration de la première partie tient au fait que le F de Snédécour utilisé pour le test est non centré de paramètre :

$$\delta^2 = \frac{n}{\sigma^2} d^2(\mathcal{E}(K, k_0), \mathcal{E}(K, k_1)) = \frac{n}{\sigma^2} \|\mathbf{x}(K, k_1 - k_0)\|^2$$

Le reste découle du fait que $\mathcal{E}_1(K, k)$ est inclu dans $\mathcal{E}(K, k)$ et que projeter orthogonalement revient à minimiser une norme.

Si la matrice K n'a qu'une ligne, on remarque que :

$$\text{Var}(K\hat{\beta}) = \frac{\sigma^2}{n} KQ_1^{-1}K^t = \frac{\sigma^2}{n} \frac{1}{\|\mathbf{x}(K, 1)\|^2} \quad \text{et le rapport :}$$

$$\rho(K) = \frac{\|\mathbf{x}(K, 1)\|^2}{\|\mathbf{x}_1(K, 1)\|^2} \quad \text{entre le minimum possible de } \text{Var}(K\hat{\beta})$$

et la variance de $K\hat{\beta}$ pour un certain plan est appelée "efficacité" de ce plan relativement à $K\beta$. On trouve une très abondante littérature utilisant différents critères pour définir une efficacité relativement à $\beta_1, \dots, \beta_\ell$, et ceci pour des plans d'expérience ayant des contraintes structurelles sur les facteurs auxquels $\xi_{\ell+1}, \dots, \xi_m$ sont associés. Ces contraintes empêchent généralement $\mathbf{x}(K, k)$ d'être égal à $\mathbf{x}_1(K, k)$ pour tout $H(K, k) \in \mathcal{H}(\ell)$. Les auteurs utilisent souvent un certain nombre de fonctions estimables de $\beta_1, \dots, \beta_\ell$, notées $(\psi_1, \dots, \psi_s)^t = \psi$ et cherchent par exemple à minimiser soit le déterminant de $\text{cov}(\hat{\psi})$ soit sa trace, soit sa plus grande valeur propre. Voir par exemple [3] et [5].

Dans cette recherche, on impose souvent une recherche d'équilibre du plan. Par exemple on fera en sorte que les variances de $\hat{\psi}_1, \dots, \hat{\psi}_s$ soient égales, ou encore que l'efficacité du plan soit la même pour ψ_1, \dots, ψ_s . Les résultats obtenus dans ces directions sont assez techniques et souvent très spécialisés : chaque contrainte structurelle doit être abordée avec ses méthodes particulières. Voir par exemple [1] ou encore [6].

Nous allons ici nous intéresser à un problème plus simple. Puisque $\|\mathbf{x}_1(K, k_1 - k_0)\|^2$ correspond au décentrage maximum parmi tous les plans de \mathcal{P}_{Q_1} , nous allons nous intéresser au cas où ce maximum est atteint, et nous allons rechercher des plans où ce maximum est atteint simultanément pour tout $H(K, k_0)$ et $H(K, k_1)$ dans $\mathcal{H}(\ell)$. Nous verrons que cela revient à rendre $\text{Cov}(K\hat{\beta})$ minimum (au sens de A-B définie négative) quel que soit $H(K, k) \in \mathcal{H}(\ell)$ et donc à minimiser $\text{Det}[\text{cov } \hat{\psi}], \text{Tr}[\text{cov } \hat{\psi}]$, ainsi que la plus grande valeur propre de $\text{cov } \hat{\psi}$, et ceci quel que soit ψ fonction estimable de $\beta_1, \dots, \beta_\ell$. Nous sommes donc amenés à introduire les définitions suivantes :

Définition 20 :

- 1) Notons $E = W_1 \cap W_2$
 E^\perp = orthogonal de E dans W_1
 W_2^\perp = trace dans W_1 de l'orthogonal de W_2
- 2) On dit que P est un plan optimum dans \mathcal{P}_{Q_1} si :
- P est dans \mathcal{P}_{Q_1}
 - Quel que soit $H(K, k_0)$ et $H(K, k_1)$ dans $\mathcal{H}(\ell)$, alors P maximise dans \mathcal{P}_{Q_1} la puissance du test de $H(K, k_0)$ contre $H(K, k_1)$, c'est-à-dire si pour ce plan : $x(K, k) = x_1(K, k)$ $\forall H(K, k) \in \mathcal{H}(\ell)$.

Voici le résultat essentiel :

Proposition 21 :

- 1) Pour qu'un plan soit optimum dans \mathcal{P}_{Q_1} , il faut et il suffit que pour ce plan : $E^\perp = W_1 \cap W_2^\perp$.
- 2) Si P est optimum dans \mathcal{P}_{Q_1}
 Si K , de dimension (s, m) est de rang s
 Si $H(K, 0)$ est dans $\mathcal{H}(\ell)$. Alors :
- $KQ^{-1}Q_1^{-1}K^t = K_1Q_1^{-1}K_1^t$ quel que soient les inverses généralisés Q^{-1} et Q_1^{-1} de Q et de Q_1 respectivement.
 - $\text{Cov}(K\hat{\beta}) = \frac{\sigma^2}{n} K_1Q_1^{-1}K_1^t$ est inversible.
 - Si P' est un autre plan de \mathcal{P}_{Q_1} , si $K\hat{\beta}'$ est l'estimateur de $K\beta$ pour ce plan P' et si $K\hat{\beta}$ est l'estimateur de $K\beta$ pour le plan P , alors

$\text{Cov}(K\hat{\beta}') - \text{cov}(K\hat{\beta})$ est définie positive.

La démonstration repose sur deux résultats :

a) on montre que l'ensemble $x = \{x \in W_1 ; x \text{ est la projection du vecteur nul sur } \mathcal{E}_1(K, k) \text{ pour } H(K, k) \in \mathcal{H}(\ell)\}$ est égal à E^\perp . On en déduit la première partie de la proposition.

b) L'expression de $\text{cov}(K\hat{\beta})$ est simple à montrer. Pour montrer que cette matrice est minimum on se sert du résultat : si A et B sont deux matrices symétriques et définies positives telles que : $A^{-1} - B^{-1}$ est définie positive, alors $B - A$ est aussi définie positive.

6.3 - Application à la régression quantitative

Dans les modèles les plus courants on s'arrange pour que g_1, \dots, g_m soient linéairement indépendants. Dans ce cas, quel que soit ℓ on a :

$$\begin{aligned} E &= W_1 \cap W_2 = \vec{0} \\ E^\perp &= W_1 \end{aligned}$$

La proposition précédente nous indique alors qu'un plan de \mathcal{P}_{Q_1} est optimum si et seulement si W_1 et W_2 sont orthogonaux, c'est-à-dire si :

$$\frac{1}{n} X^t X = \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix} \text{ où } Q_2 \text{ est quelconque.}$$

D'autre part le résultat b) énoncé pour la démonstration de la proposition précédente nous indique que si l'on se fixe comme contrainte une certaine diagonale pour $X^t X$, alors il n'existe pas de plan rendant $\text{cov}(\hat{\beta})$ minimum. En effet si A et B sont les matrices de covariances de $\hat{\beta}$ pour deux plans différents, alors $\text{Diag}(A^{-1}) = \text{Diag}(B^{-1}) = \text{Diag}(\frac{1}{n} X^t X)$. Donc $\text{Tr}(B^{-1} - A^{-1}) = 0$ ce qui implique, si A^{-1} et B^{-1} ne sont pas égales, que $B^{-1} - A^{-1}$ a des valeurs propres positives et d'autres négatives. Ainsi $B^{-1} - A^{-1}$ ne sera ni définie positive ni définie négative. Le résultat énoncé en b) nous indique qu'il en sera de même de $B - A$. Cependant, si on se fixe comme critère d'efficacité de rendre minimum $\text{Det}[\text{cov} \hat{\beta}]$, l'optimum est atteint pour $X^t X$ diagonale, voir [3].

6.4 - Application à l'analyse de la variance : modèle sans interactions

Notre modèle s'écrit :

$$\begin{aligned} E^{\tilde{X}}(\tilde{Y}) &= \beta_0 \tilde{J}_0 + \sum_{i=1}^r \sum_{j=1}^{m_i} \beta_i^j \tilde{M}_i^j \\ E(\tilde{Y}/\tilde{X}_1 = \ell_1 ; \dots ; \tilde{X}_r = \ell_r) &= \beta_0 + \sum_{i=1}^r \beta_i^{\ell_i} \\ \mu &= \beta_0 \tilde{J}_0 + \sum_{i=1}^r \sum_j \beta_i^j \tilde{M}_i^j \end{aligned}$$

On notera :

1) $G_i = \{M_i^1, \dots, M_i^{m_i}\}$, W_i le sous espace engendré par G_i et Q_i la métrique de W_i associée à G_i :

$$Q_i = \frac{1}{n} \begin{vmatrix} n_i^1 & & \\ & 0 & \\ c & & m_i \\ & & & n_i^j \end{vmatrix} \quad n_i^j \text{ étant l'effectif du noeud } \Omega_i^j$$

2) $W^1 = \sum_{i=1}^r W_i$ et J l'axe des constantes. On remarquera que J est dans W_i pour tout i .

3) \mathcal{H}_i l'ensemble des hypothèses sur des fonctions estimables ne portant que sur $\beta_i^1, \dots, \beta_i^{m_i}$. On rappelle que $\sum_j K_j \beta_i^j$ est estimable si et seulement si : $\sum_j K_j = 0$.

4) \mathcal{P}_{Q_i} désignera l'ensemble des plans d'expérience sans biais pour lesquelles la métrique de W_i associée à G_i est égale à une certaine matrice Q_i .

Nous allons caractériser les plans optimaux dans \mathcal{P}_{Q_i} , qui maximisent donc la puissance des tests de \mathcal{H}_i pour $n_i^1, \dots, n_i^{m_i}$ fixés.

Proposition 22 :

1) Pour qu'un plan d'expérience soit optimum dans \mathcal{P}_{Q_i} il faut et il suffit que :

$$P(\Omega_i^j \cap \Omega_i^{j'}) = P(\Omega_i^j)P(\Omega_i^{j'}) \quad \forall i' \neq i, \forall j \text{ et } \forall j'.$$

2) Pour qu'un plan d'expérience soit optimum dans $\mathcal{P}_{Q_1}, \mathcal{P}_{Q_2}, \dots, \mathcal{P}_{Q_r}$ il faut et il suffit que :

- a) ce plan soit dans $\mathcal{P}_{Q_1}, \dots, \mathcal{P}_{Q_r}$
- b) il soit orthogonal.

La démonstration repose sur les faits suivants :

Notons $W_i^c = \sum_{i' \neq i} W_{i'}$; $E_i = W_i \cap W_i^c$; E_i^\perp l'orthogonal de E_i dans W_i et W_i^* la trace dans W_i de l'orthogonal de W_i^c . La proposition 21 nous fait rechercher les conditions sous lesquelles $E_i^\perp = W_i \cap W_i^*$. Or si le plan est sans biais $E_i^\perp = V_i$ (voir définition prop. 12). Il est évident que V_i contient toujours $W_i \cap W_i^*$. Pour étudier dans quelles conditions la réciproque est vraie, il suffit de remarquer que V_i est engendré par les éléments du type $M_{i-P_j}^j(M_i^j)$ et que pour $i' \neq i$:

$$\langle M_{i-P_j}^j(M_i^j), M_{i'-P_j}^{j'} \rangle = P(\Omega_{ii'}^{jj'}) - P(\Omega_i^j)P(\Omega_{i'}^{j'}).$$

Bien entendu, si on se pose Q_1, \dots, Q_r arbitrairement, on n'est pas certain de pouvoir construire un plan orthogonal.

6.5 - Application au modèle avec interactions d'ordre 2

Notre modèle s'écrit :

$$E\tilde{X}(\tilde{Y}) = \beta_0 \tilde{J}_0 + \sum_i \sum_k \beta_{iM_i^k} \tilde{k}^k + \sum_{i < j} \sum_{k, \ell} \beta_{ijM_{ij}^{k\ell}} \tilde{k}^k \tilde{\ell}^{\ell}$$

$$E(\tilde{Y}/\tilde{X}_1 = \ell_1 ; \dots ; \tilde{X}_r = \ell_r) = \beta_0 + \sum_i \beta_i^{\ell_i} + \sum_{i < j} \beta_{ij}^{\ell_i \ell_j}$$

$$\mu = \beta_0 \tilde{J}_0 + \sum_i \sum_k \beta_{iM_i^k} \tilde{k}^k + \sum_{i < j} \sum_{k, \ell} \beta_{ijM_{ij}^{k\ell}} \tilde{k}^k \tilde{\ell}^{\ell}$$

On notera :

1) $G_{ij} = \{M_{ij}^{11}, \dots, M_{ij}^{m_i m_j}\}$, W_{ij} le sous espace engendré par G_{ij} et Q_{ij} la métrique de W_{ij} associée à G_{ij} :

$$Q_{ij} = \frac{1}{n} \begin{bmatrix} n_{ij}^{11} & & & \\ & \ddots & & \\ & & 0 & \\ & & & m_i m_j \\ 0 & & & & n_{ij} \end{bmatrix}$$

2) $W^2 = \sum_{i < j} W_{ij}$. On remarquera que W_i est inclu dans W_{ij}

3) \mathcal{H}_{ij} l'ensemble des hypothèses sur des fonctions estimables ne portant que sur $\beta_{ij}^{11}, \dots, \beta_{ij}^{m_i m_j}$. On rappelle que $\sum_{k, \ell} K_{k\ell} \beta_{ij}^{k\ell}$ est estimable si et seulement si :

$$\sum_k K_{k\ell} = \sum_{\ell} K_{k\ell} = 0$$

\mathcal{H}_{ij} contient en particulier l'hypothèse de non existence d'interactions entre \tilde{X}_i et \tilde{X}_j .

4) $\mathcal{P}_{Q_{ij}}$ désignera l'ensemble des plans d'expérience sans biais tels que Q_{ij} soit la métrique de W_{ij} associée à G_{ij} .

5) Q_{ij} sera dite adaptable à un plan orthogonal si les éléments diagonaux de cette matrice vérifient :

$$P(\Omega_{ij}^{k\ell}) = P(\Omega_i^k)P(\Omega_j^{\ell})$$

6) Si Q_{ij} est adaptable à un plan orthogonal, on notera $\mathcal{P}_{Q_{ij}}^l$ l'ensemble

des plans de $\mathcal{P}_{Q_{ij}}$ qui sont orthogonaux pour le modèle avec interactions d'ordre deux.

7) $\mathcal{P}_{Q_{ij}}^*$ désignera l'ensemble des plans optimaux de $\mathcal{P}_{Q_{ij}}$.

Proposition 23 :

1) Pour que quel que soit $(i < j)$ un plan soit optimum dans $\mathcal{P}_{Q_{ij}}$ il faut qu'il soit orthogonal pour le modèle sans interactions, donc que $\mathcal{P}_{Q_{ij}}$ soit adaptable à un plan orthogonal pour tout $(i < j)$.

2) Si Q_{ij} est adaptable à un plan orthogonal :

$$\mathcal{P}_{Q_{ij}}^\perp \subset \mathcal{P}_{Q_{ij}}^* \quad (\text{orthogonalité} \Rightarrow \text{optimalité})$$

3) Si pour tout $(i < j)$ Q_{ij} est adaptable à un plan orthogonal :

$$\bigcap_{i < j} \mathcal{P}_{Q_{ij}}^\perp = \bigcap_{i < j} \mathcal{P}_{Q_{ij}}^* \quad (\text{orthogonalité} \Leftrightarrow \text{optimalité})$$

La démonstration utilise les mêmes principes que ceux de la proposition 22, cependant ici les calculs sont plus longs.

On remarquera en particulier que :

a) $E_{ij} = W_{ij} \cap \left(\bigcup_{(i',j') \neq (i,j)} W_{i'j'} \right) = W_{i+W_j}$ et donc que $E_{ij}^\perp = V_{ij}$.

b) si Q_{ij} est adaptable à un plan orthogonal :

$$M_{ij}^{k\ell - P_{W_i + W_j}} (M_{ij}^{k\ell}) = M_{ij}^{k\ell - P(\Omega_j^\ell)} M_i^{k - P(\Omega_i^k)} M_j^{\ell + P(\Omega_{ij}^{k\ell})} J_0$$

D'une manière générale, grâce à la décomposition de W^p énoncée à la proposition 12, on peut montrer qu'un plan orthogonal pour le modèle d'ordre p est optimum pour l'ensemble des hypothèses portant sur des fonctions estimables des coefficients β_ε^η , $\eta \in K(\varepsilon)$, $\varepsilon = (i_1, \dots, i_p)$, en supposant fixés les effectifs des noeuds d'ordre $p-1$ entre X_{i_1}, \dots, X_{i_p} .

7 - CONCLUSION

Pour tous les problèmes étudiés ici, l'introduction des noeuds et des notations géométriques aura, je l'espère, semblé utile, même si les démonstrations des différentes propositions à l'aide de ces techniques n'ont été qu'esquissées.

BIBLIOGRAPHIE

- [1] D.V. CHOPRA : "Balanced optimal 2^8 fractional factorial designs of resolution V, $52 \leq N \leq 59$ ". Dans "A survey of statistical Design and Linear Models". North-Holland Publishing Company. 1975.
- [2] J.J. DAUDIN : "An iterative algorithm for analysis of variance". Dans J. Bio-Medical Computing (10) (1979) 507-518.
- [3] FEDOROV : "Theory of Optimal Experiments". Academic Press, 1972.
- [4] J.W. Mc KEAN and R.M. SCHRADER : "The geometry of robust procedures in linear models". Jour. Roy. SS, B, 43, pp 366-371.
- [5] J. KIEFFER : "Construction and optimality of generalized Youden designs" dans "A survey of Statistical Design and Linear Models". North-Holland Publishing Company. 1975.
- [6] M. SINGH and A. DEY : "Block designs with nested rows and columns". Biometrika (1979), 66, 2, pp 321-326.