REVUE DE STATISTIQUE APPLIQUÉE

CATHERINE DUPONT-GATELMAND

Un modèle individuel de décomposition additive des préférences : le modèle Benito

Revue de statistique appliquée, tome 29, nº 1 (1981), p. 21-41

<http://www.numdam.org/item?id=RSA_1981__29_1_21_0>

UN MODELE INDIVIDUEL DE DECOMPOSITION ADDITIVE DES PREFERENCES : LE MODELE BENITO

Catherine DUPONT-GATELMAND

Attachée de Recherche au C.N.R.S.
Lamsade, Université Paris-Dauphine

RESUME

Ce modèle est présenté dans un contexte d'analyse et d'explication des préférences afin de dégager les liens qui peuvent exister avec la classe des modèles d'"Analyse Conjointe" (ou Conjoint Measurement).

A partir d'un ensemble d'opinions individuelles, il met en évidence les différentes influences spécifiques à des profils identiques de juges. Autrement dit, l'objectif poursuivi n'est pas d'agréger des relations individuelles pour obtenir un comportement collectif mais au contraire de désagréger les préférences individuelles en fonction d'attributs propres aux juges et aux objets.

Le modèle repose sur une hypothèse fondamentale : on suppose en effet que les choix d'un individu peuvent être expliqués par ses propres caractéristiques (par exemple : variables socio-économiques). Ainsi, on décompose la note y_{ij} donnée par le juge i à l'objet j en une somme de termes correspondant aux interactions entre les attributs de i et ceux de j.

Le problème est alors de trouver ce tableau de poids ou coefficients d'influence : chacun d'eux permet de corriger la préférence globale moyenne en tenant compte des éléments qui décrivent le couple (i, j) auquel on s'intéresse.

L'interprétation des résultats est rapide et simple.

INTRODUCTION

L'analyse des préférences a suscité la construction de deux approches complémentaires. En effet, du concept individuel de "préférence" qui traduit une détermination rationnelle puisque comparative, on a cherché d'une part à décomposer, à désagréger les choix et, d'autre part, à synthétiser, à agréger les jugements.

Dans le cadre des méthodes de décomposition sont nées différentes modélisations.

D'une manière simplificatrice, on distingue les modèles descriptifs des modèles explicatifs (J.M. BOUROCHE [1]). Dans le premier cas, on cherche à visualiser les individus et les objets dans un espace de faible dimension $\mathbf{R}^2$ ou $\mathbf{R}^3$: la proximité de deux individus peut ainsi se traduire par deux jugements voisins et les distances entre individus et objets s'interprètent en terme de préférence. Les méthodes dites "d'analyse interne" ou "externe" permettent de répondre à cette intention de description et confèrent un rôle symétrique aux variables : elles ont

Mots-Clé : **Préférence, Mesure conjointe, Analyse de la variance, Trade-off.**

été développées pour la plupart par des équipes américaines : J.D. CARROLL et J.J. CHANG [2], [3], V. SRINIVASAN et A.D. SHOCKER [14].

Dans le deuxième cas, le problème de l'explication des préférences est plus délicat et les tentatives de résolution qui en ont été faites plus récentes. On considère en général plusieurs ensembles de données : par exemple (cf. E. JACQUET-LAGREZE [9]), un ensemble d'individus émet des jugements sur un ensemble d'objets en tenant compte d'un ensemble de critères. Il est parfois nécessaire d'aborder la question par étapes successives : après une typologie des individus en fonction de leurs réponses aux différents critères, on visualisera les préférences moyennes obtenues.

On distingue différents modèles explicatifs selon la nature de la (ou des) variable(s) qu'ils privilégient. Ainsi ORDREG [15] et ADEPREF [9] traitent le cas où les variables à expliquer sont des relations binaires. Toutefois, pour le premier, les variables explicatives sont quantitatives et, pour le second, celles-ci sont également des relations binaires.

L'analyse conjointe non métrique (MONANOVA) [11] travaille sur une variable à expliquer qualitative ordinale et des variables explicatives qualitatives. UTA [10] travaille sur des variables à expliquer et explicatives qualitatives ordinales.

Dans cette recherche de relations entre plusieurs ensembles de données, un nouveau chapitre vient d'être ouvert par J.D. CARROLL. En tenant compte de trois groupes de variables et en développant ainsi CANDELINC (CANonical DEcomposition under LINear Constraints) [5], [6], il s'est situé à mi-chemin entre les deux pôles de visualisation et d'explication.

Le modèle BENITO, que nous présentons ici, s'appuie sur la définition de tableaux d'entrées analogues au précédent modèle mais, répondant à des objectifs différents, ne peut être comparé avec celui-ci.

1. POSITION DU PROBLEME

1.1. Les données

On considère un ensemble d'individus I de cardinal n et un ensemble d'objets J de cardinal p.

Chaque individu détermine ses préférences sur les p objets sous forme d'évaluations sur des échelles numériques. Lorsque les jugements sont recueillis sous forme de rang (ordre ou préordre total), ceux-ci sont assimilés à des vecteurs de $\mathbf{R}^p$.

Il convient de spécifier que la méthode s'applique actuellement sur des données complètes et ne se prête pas aux techniques des plans d'expériences.

Le modèle BENITO est construit à partir des matrices suivantes :

– *La matrice Y* de dimension (n , p) qui rassemble les n jugements de préférences émis sur les p objets. La réponse d'un individu est donc représentée par une ligne de la matrice.

– *La matrice* X_1 de dimension (n, N_1) sert à caractériser les n juges : N_1 désigne le nombre de variables qualitatives considérées pour les représenter. Elle est transformée en tableau disjonctif complet et devient alors de dimension (n, n_1) où n_1 définit le nombre total de modalités possibles de X_1 sur l'ensemble des N_1 variables.

– *La matrice* X_2 de dimension (p, N_2) caractérise les p objets, N_2 désignant le nombre de variables qualitatives prises en compte pour les définir. Comme X_1, elle est transformée en tableau disjonctif complet et sa dimension est alors (p, n_2).

1.2. Formulation mathématique du problème

Le problème est alors de *trouver* la matrice T de dimension (n_1, n_2) et le vecteur t de dimension $(1, n_2)$ tels que :

$$\textcircled{P_1}\left\{\begin{array}{l} \underset{(n,p)}{Y} \simeq \underset{(n,n_1)}{X_1}\ \underset{(n_1,n_2)}{T}\ \underset{(n_2,p)}{X_2'} + \underset{(n,1)}{\underline{1}}\ \underset{(1,n_2)}{\tilde{t}}\ \underset{(n_2,p)}{X_2'} \\ \text{où} \simeq \text{désigne l'approximation au sens des moindres carrés.} \end{array}\right.$$

De manière évidente, on peut montrer que le problème P_1 est équivalent au problème P_2 suivant.

Trouver la matrice T^* de dimension $(n_1 + 1, n_2)$ telle que :

$$\textcircled{P_2}\left\{\begin{array}{l} \underset{(n,p)}{Y} \simeq \underset{(n,n_1+1)}{X_1^*}\ \underset{(n_1+1,n_2)}{T^*}\ \underset{(n_2,p)}{X_2'} \\ \text{avec :} \\ \underset{(n,n_1+1)}{X_1^*} = (\underset{(n,n_1)}{X_1} \mid \underset{(n,1)}{\underline{1}}) \quad \text{et} \quad \underset{(n_1+1,n_2)}{T^*} = \underbrace{\begin{pmatrix} T \\ \overline{\tilde{t}} \end{pmatrix}}_{n_2} \begin{array}{l} \}n_1 \\ \}1 \end{array} \end{array}\right.$$

N.B. : Dans toute la suite, on posera $m_1 = n_1 + 1$.

Nous avons utilisé deux approches différentes pour résoudre ces deux problèmes équivalents :

– une méthode de moindres carrés alternés pour P_1 consistant à chercher $\tilde{t}$ puis T ;
– une méthode directe pour P_2.

Nous détaillons au paragraphe 2 la résolution du problème P_2 et nous donnons ensuite les différences essentielles correspondant à celle du problème P_1.

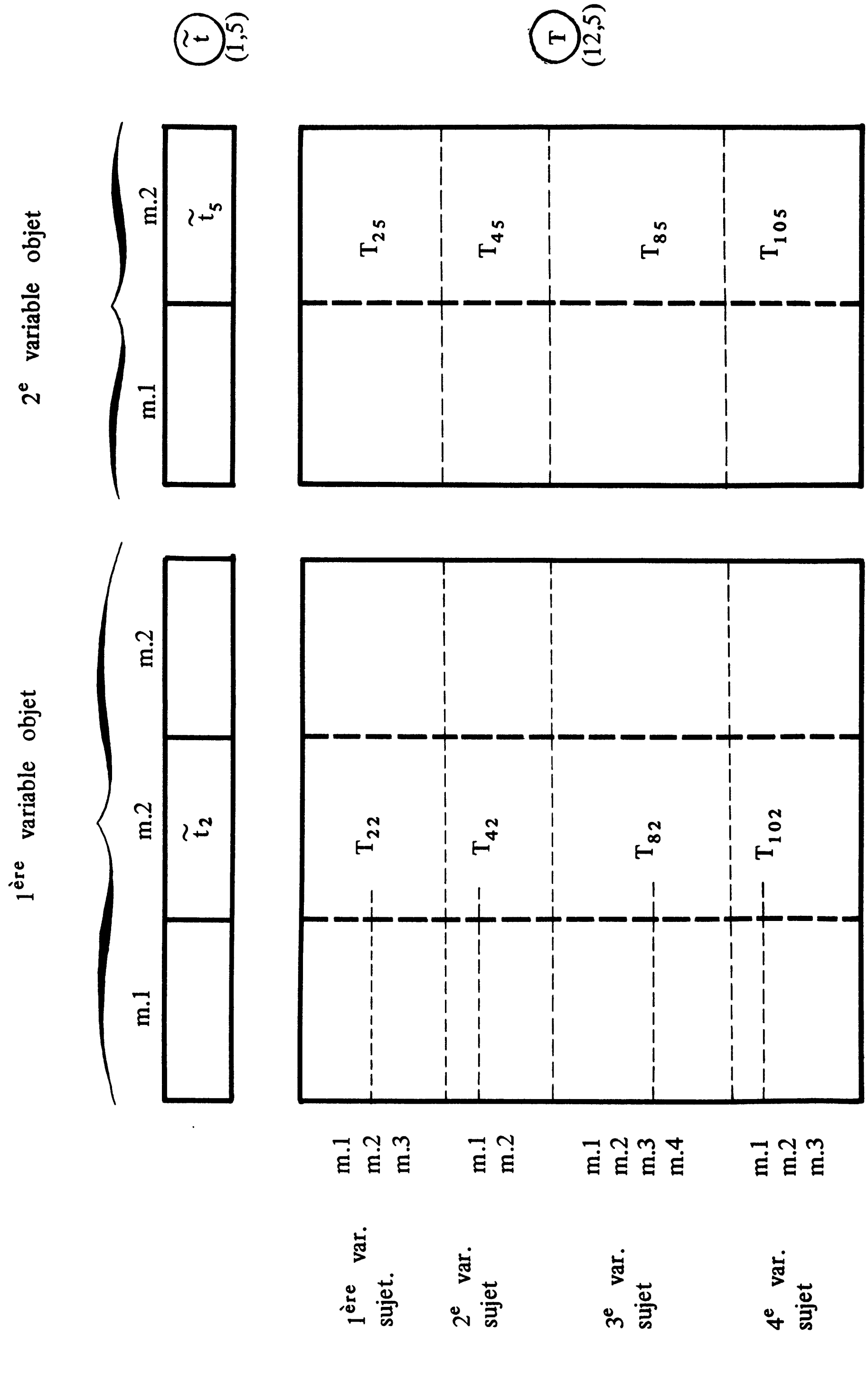
Auparavant, il convient d'éclairer cette formulation mathématique en tentant d'interpréter les éléments recherchés et en s'interrogeant sur l'intérêt d'un tel modèle.

1.3. Quelques règles d'interprétation

Reprenons le problème sous sa forme P_1 : on cherche T et $\tilde{t}$ tels que :

$$\underset{(n,p)}{Y} \simeq \underset{(n,n_1)}{X_1}\ \underset{(n_1,n_2)}{T}\ \underset{(n_2,p)}{X_2'} + \underset{(n,n_2)}{\tilde{T}}\ \underset{(n_2,p)}{X_2'} \quad \text{avec} \quad \underset{(n,n_2)}{\tilde{T}} = \underset{(n,1)}{\underline{1}}\ \underset{(1,n_2)}{\tilde{t}}$$

Que dire de ces deux matrices T et $\tilde{t}$?



1ère variable objet
m.1
m.2
m.2
$\tilde{t}_2$
2e variable objet
m.1
m.2
$\tilde{t}_5$
$\tilde{t}$ (1,5)
T_{22}
T_{42}
T_{82}
T_{102}
T_{25}
T_{45}
T_{85}
T_{105}
T (12,5)
1ère var. sujet.
m.1
m.2
m.3
2e var. sujet
m.1
m.2
3e var. sujet
m.1
m.2
m.3
m.4
4e var. sujet
m.1
m.2
m.3

• $\tilde{t}$ correspond à la préférence globale moyenne des n sujets, indépendamment de leurs caractéristiques propres, sur les caractéristiques des objets qu'ils ont eu à ordonner. Tout se passe comme si l'on avait fait une analyse de la variance sous contrainte (toutes les lignes de $\tilde{T}$ étant identiques) expliquant Y à l'aide de N_2 variables qualitatives déterminant les p objets.

Le modèle serait alors le suivant :

$$\underset{(n,p)}{Y} \simeq \underset{(n,n_2)}{\tilde{T}} \ \underset{(n_2,p)}{X_2'}$$

• T est la matrice des coefficients d'influence induits par les interactions entre les caractéristiques sujets et les caractéristiques objets.

D'une manière globale, $\tilde{t}$ est une approximation de la moyenne que viennent moduler les coefficients de T. Ainsi le profil d'un individu vient corriger la note moyenne d'une caractéristique objet déterminée sur l'ensemble de la population.

Prenons un exemple schématique et considérons un ensemble de sujets caractérisés par 4 variables qualitatives à respectivement 3, 2, 4 et 3 modalités. Supposons qu'ils aient noté plusieurs objets selon leurs préférences et que ces objets soient définis par 2 variables qualitatives à respectivement 3 et 2 modalités.

Les inconnues du problème (T et $\tilde{t}$) sont présentées sous la forme suivante. On a ici $N_1 = 4$ $(n_1 = 12)$, $N_2 = 2$ $(n_2 = 5)$.

Supposons qu'un des objets classés soit caractérisé par :

– la seconde modalité de la première variable-objet ;
– la seconde modalité de la seconde variable-objet.

La note moyenne déterminée par l'ensemble des juges pour cet objet sera $\tilde{t}_2 + \tilde{t}_5$.

Supposons maintenant que l'on s'intéresse à un sujet caractérisé par :

– la seconde modalité de la première variable-sujet ;
– la première modalité de la seconde variable-sujet ;
– la troisième modalité de la troisième variable-sujet ;
– la première modalité de la quatrième variable-sujet.

$\tilde{t}_2 + T_{22} + T_{42} + T_{82} + T_{102} + \tilde{t}_5 + T_{25} + T_{45} + T_{85} + T_{105}$ est l'approximation de la note donnée par un tel individu à cet objet. Le préordre moyen établi sur les caractéristiques objets est ainsi modifié selon les catégories des interviewés auxquels on s'intéresse.

Le modèle BENITO détermine ainsi, par approximation au sens des moindres carrés, la note y_{ij} donnée par le juge i à l'objet j. Celle-ci est approchée par une note moyenne $\tilde{t}$ calculée sur l'ensemble des juges et corrigée par une somme de coefficients – ceux de la matrice T – correspondant simultanément aux caractéristiques de i et à celles de j.

La transcription de P_1 ou P_2 sous forme décomposée permet de le décrire explicitement :

$$\hat{y}_{ij} = \sum_{k=1}^{n_2} x_{2_{jk}} \left[\tilde{t}_k + \sum_{s=1}^{n_1} t_{sk}\, x_{1_{is}} \right]$$

$$\hat{y}_{ij} = \sum_{s=1}^{m_1} \sum_{k=1}^{n_2} t^*_{sk}\, x_{1_{is}}\, x_{2_{jk}}$$

Une telle interprétation est rendue possible par la totale ADDITIVITE du modèle (les matrices X_1 et X_2 étant sous forme disjonctive complète).

L'intérêt du modèle réside essentiellement dans son aspect "décomposition" : il donne "l'éclatement" des interactions entre les caractéristiques sujets et objets déterminant ainsi les contributions respectives de chacun des couples "modalité variable-sujet – modalité variable-objet".

2. RESOLUTION DU PROBLEME P_2 : LA METHODE DIRECTE

Notre objectif est donc de trouver une matrice T^* de dimension (m_1, n_2) qui minimise la quantité :

$$\|Y - \hat{Y}\|^2 = \|Y - X_1^* T^* X_2'\|^2 \quad \text{avec} \quad \|A\|^2 = \operatorname{tr}(AA')$$

2.1. Formulation du critère et équivalence du critère

* *Proposition 1 :*

Si les matrices X_1^* et X_2 sont orthonormales, c'est-à-dire :

$$\begin{cases} X_1^{*\prime} X_1^* = I_{m_1} \\ X_2' X_2 = I_{n_2} \end{cases}$$

une solution du problème (P_2) est :

$$T^* = X_1^{*\prime} Y X_2 \,.$$

Démonstration :

Posons $(E)^2 = \operatorname{Tr}[(Y - \hat{Y})(Y - \hat{Y})']$.

Le problème (P_2) est équivalent à $\underset{T^*}{\operatorname{Min}}\, E^2$.

En développant, on obtient :

$$\begin{aligned}(E)^2 &= \operatorname{Tr}[(Y - X_1^* T^* X_2')(Y - X_1^* T^* X_2')'] \\ &= \operatorname{Tr}(YY' - YX_2 T^{*\prime} X_1^{*\prime} - X_1^* T^* X_2' Y' + X_1^* T^* X_2' X_2 T^{*\prime} X_1^{*\prime}) \\ &= \operatorname{Tr}(YY') - 2\operatorname{Tr}(X_1^{*\prime} Y X_2 T^{*\prime}) + \operatorname{Tr}(X_1^* T^* X_2' X_2 T^{*\prime} X_1^{*\prime})\end{aligned}$$

Comme $X_1^{*\prime} X_1^* = I_{m_1}$ et $X_2' X_2 = I_{n_2}$

$$(E)^2 = \operatorname{Tr}(YY') - 2\operatorname{Tr}(X_1^{*\prime} Y X_2 T^{*\prime}) + \operatorname{Tr}(T^* T^{*\prime})$$

On pose $\quad (E^*)^2 = \operatorname{Tr}[(X_1^{*\prime} Y X_2 - T^*)(X_1^{*\prime} Y X_2 - T^*)']$

En développant comme précédemment, on a :

$$\begin{aligned}(E^*)^2 &= \operatorname{Tr}(X_1^{*\prime} Y X_2 X_2' Y' X_1^*) - 2\operatorname{Tr}(X_1^{*\prime} Y X_2 T^{*\prime}) + \operatorname{Tr}(T^* T^{*\prime}) \\ &= E^2 + K\end{aligned}$$

avec $$K = \mathrm{Tr}(X_1^{*\prime} Y X_2 X_2' Y' X_1^*) - \mathrm{Tr}(YY') .$$

K est indépendant de T^* : il est donc équivalent de minimiser $(E^*)^2$ ou $(E)^2$.

Or $(E^*)^2$ est minimisé avec $T^* = X_1^{*\prime} Y X_2$.

C.Q.F.D.

Afin d'obtenir la condition d'orthonormalité des matrices X_1^* et X_2, on utilise le théorème suivant (connu par les anglo-saxons sous le nom de "singular value decomposition").

* *Théorème 1 :*

Soit X une matrice de dimension (m , n) et de rang p ; on peut toujours trouver :

– une matrice U orthogonale de dimension (m , m) ;
– une matrice V orthogonale de dimension (n , n) ;
– et une matrice β diagonale de dimension (m , n) telles que :

$$U' X V = \beta \quad \text{ou} \quad X = U \beta V' .$$

Les éléments diagonaux non nuls de la matrice β sont les valeurs propres non nulles rangées par ordre décroissant de XX' ou $X'X$; tous ces éléments sont non négatifs et p d'entre eux sont strictement positifs.

La démonstration de ce théorème peut être trouvée dans [7].

Ce théorème permet ainsi de décomposer les matrices X_1^* et X_2 sous la forme suivante :

$$\underset{(n, m_1)}{X_1^*} = \underset{(n, n)}{U_1} \; \underset{(n, m_1)}{\beta_1} \; \underset{(m_1, m_1)}{V_1'}$$

avec $$\begin{cases} U_1 U_1' = U_1' U_1 = I_n \\ V_1 V_1' = V_1' V_1 = I_{m_1} \end{cases}$$

pour $n > m_1$
et $r_1 = R_g X_1^*$

$$\beta_1 = n\left\{\overbrace{\left(\begin{array}{ccc|cc} b_1 & & & & \\ & b_1 & & & O \\ & & b_{r1} & & \\ \hline & & & 0 & \\ & O & & & 0 \end{array}\right)}^{m_1}\right. = \overbrace{\left(\begin{array}{c|c} \bar{\beta}_1 & O \\ \hline O & O \end{array}\right)}^{r_1} \begin{array}{l} \}\, r_1 \\ \\ \end{array}$$

De même :

$$\underset{(p\,,\,n_2)}{X_2} = \underset{(p\,,\,p)}{U_2}\ \underset{(p\,,\,n_2)}{\beta_2}\ \underset{(n_2\,,\,n_2)}{V'_2}$$

avec
$$\left\{\begin{array}{l} U_2\,U'_2 = U'_2\,U_2 = I_p \\ V'_2\,V_2 = V_2\,V'_2 = I_{n_2} \end{array}\right\}$$

De manière équivalente, on peut écrire ces deux décompositions en considérant les matrices diagonales $\overline{\beta}_1$ et $\overline{\beta}_2$, issues de β_1 et β_2 et ne contenant que des valeurs strictement positives. On a alors :

$$\underset{(n\,,\,m_1)}{X_1^*} = \underset{(n\,,\,r_1)}{U_1}\ \underset{(r_1\,,\,r_1)}{\overline{\beta}_1}\ \underset{(r_1\,,\,m_1)}{V'_1}$$

avec
$$\left\{\begin{array}{l} r_1 = R_g\,X_1^* \\ U'_1\,U_1 = I_{r_1} \\ V'_1\,V_1 = I_{r_1} \end{array}\right.$$

$$\underset{(p\,,\,n_2)}{X_2} = \underset{(p\,,\,r_2)}{U_2}\ \underset{(r_2\,,\,r_2)}{\overline{\beta}_2}\ \underset{(r_2\,,\,n_2)}{V'_2}$$

avec
$$\left\{\begin{array}{l} r_2 = R_g X_2 \\ U'_2\,U_2 = I_{r_2} \\ V'_2\,V_2 = I_{r_2} \end{array}\right.$$

Le problème à résoudre est alors :

$$\underset{T^*}{\mathrm{Min}}\ \|\ \underset{(n\,,\,p)}{Y} - \underset{(n\,,\,r_1)}{U_1}\ \underset{(r_1\,,\,r_1)}{\overline{\beta}_1}\ \underset{(r_1\,,\,m_1)}{V'_1}\ \underset{(m_1\,,\,n_2)}{T^*}\ \underset{(n_2\,,\,r_2)}{V_2}\ \underset{(r_2\,,\,r_2)}{\overline{\beta}_2}\ \underset{(r_2\,,\,p)}{U'_2}\ \|^2$$

On pose $\qquad \theta = \overline{\beta}_1\,V'_1\,T^*V_2\overline{\beta}_2$.

Proposition 2 :

Soit (P) le problème : $\underset{T^*}{\mathrm{Min}}\ \|\,Y - U_1\overline{\beta}_1\,V'_1\,T^*V_2\overline{\beta}_2\,U'_2\,\|^2$.

Soit (P′) le problème : $\underset{\theta}{\mathrm{Min}}\ \|\,Y - U_1\,\theta\,U'_2\,\|^2$.

Si $\hat{\theta}$ est solution de (P′), alors $\hat{T}^*$ tel que :

$\hat{\theta} = \overline{\beta}_1\,V'_1\,\hat{T}^*\,V_2\overline{\beta}_2$ est solution de (P) .

Démonstration (évidente).

Remarque : Pour tout θ, il existe toujours un T^* de dimension $(m_1\,,n_2)$ vérifiant $\theta = \overline{\beta}_1 V'_1 T^* V_2 \overline{\beta}_2$.

De même, pour tout T^*, il existe toujours un θ de dimension $(r_1\,,\,r_2)$ satisfaisant cette condition.

On raisonne par l'absurde.

Ainsi, d'après la proposition 1, une solution du problème (P′) est

$\theta = U'_1 YU_2$. La proposition 2 détermine une solution T^* au problème (P) telle que :

$$\boxed{\overline{\beta}_1 V'_1 T^* V_2 \overline{\beta}_2 = U'_1 YU_2} \qquad (*)$$

Remarque : Une autre façon d'exprimer les propositions 1 et 2 et de montrer l'optimalité analytique de la solution est d'utiliser la dérivation matricielle (cf. [7]).

2.2. Détermination d'une solution T^*

La solution de (*) s'obtient en deux étapes par deux systèmes linéaires "en cascade" :

On pose :

$$\begin{cases} \underset{(r_1, m_1)}{W_1} = \underset{(r_1, r_1)}{\overline{\beta}_1} \; \underset{(r_1, m_1)}{V'_1} \\ \underset{(n_2, r_2)}{W_2} = \underset{(n_2, r_2)}{V_2} \; \underset{(r_2, r_2)}{\overline{\beta}_2} \end{cases}$$

• Dans un premier temps, on résout le système :

$$U'_1 YU_2 = [W_1 T^*] W_2$$

ce qui revient à chercher une matrice W_2^- de dimension (r_2, n_2) telle que :

$$\underset{(r_2, n_2)}{W_2^-} \; \underset{(n_2, r_2)}{W_2} = I_{r_2}$$

La solution n'est pas unique puisque le système a $r_2 n_2$ inconnues et seulement $r_2 r_2$ équations.

• Dans un deuxième temps, on résout le système :

$$[U'_1 YU_2] W_2^- = W_1 T^*$$

ce qui revient à chercher une matrice W_1^- de dimension (m_1, r_1) telle que :

$$\underset{(r_1, m_1)}{W_1} \; \underset{(m_1, r_1)}{W_1^-} = I_{r_1}$$

Comme précédemment, la solution n'est pas unique, le système ayant $r_1 m_1$ inconnues et seulement $r_1 r_1$ équations.

Par ces deux étapes, on détermine ainsi une solution :

$$T^* = W_1^- [U'_1 YU_2] W_2^-$$

Remarque : W_1^- et W_2^- sont des g-inverses de W_1 et W_2 (cf. RAO [12]).

Il existe un ensemble de solutions admissibles vérifiant ($*$). C'est la non unicité des inverses généralisées W_1^- et W_2^- qui entraîne celle de la solution en T^* du système que l'on résout.

Il convient maintenant de déterminer *une* solution parmi celles-ci (on a vu que l'on ne recherchait pas la solution de norme minimum).

Pour des raisons d'interprétation, on *CHOISIT* la solution (le tableau T^*)

dont les lignes et les colonnes sont au mieux centrées par bloc de variables, c'est-à-dire on *CHOISIT* le tableau T^* centré au mieux d'une part sur chaque variable objet et, d'autre part, sur chaque variable sujet.

W_1^- et W_2^- ne correspondent pas à *la* pseudo-inverse de MOORE-PENROSE de W_1 et de W_2.

En effet : $(W_1 W_1^-)' = W_1 W_1^-$ et $(W_2^- W_2)' = W_2^- W_2$

mais : $(W_1^- W_1)' \neq W_1^- W_1$ mais $(W_2 W_2^-)' \neq W_2 W_2^-$

De plus, on ne recherche pas la solution T^* de norme minimum – solution que l'on aurait obtenue en utilisant les pseudo inverses W_1^+ et W_2^+ de MOORE-PENROSE.

Autrement dit, à la résolution du premier système (calcul de W_2^-), on choisit parmi l'ensemble des solutions celle qui, pour chacun des r_2 systèmes, correspondant aux r_2 lignes de la matrice W_2^-, minimise la fonction suivante :

pour le $k^{\text{ème}}$ système $(k \in]r_2])$ et avec $W_2^- = (a_{ki})_{\substack{k=1,-r_2 \\ i=1,-n_2}}$

$$f(a_{k1}, \ldots, a_{ki}, \ldots, a_{kn_2}) = \sum_{j=1}^{N_2} \frac{1}{M(j)} \left(\sum_{q=A}^{B} a_{kq} \right)^2$$

où $M(j)$ désigne le nombre de modalités de la variable j

$$\begin{cases} A = \sum_{s=1}^{j-1} M(s) + 1 \\ B = \sum_{s=1}^{j} M(s) \end{cases}$$

sous les contraintes :

$$W_2^- W_2 = I_{r_2} \qquad (*)$$

En associant à la matrice $W_2'(r_2, n_2)$ l'application g de $\mathbf{R}^{n_2}$ dans $\mathbf{R}^{r_2}$, on peut écrire $(*)$ sous la forme de r_2 contraintes de la forme :

$$g[a_{k1}, \ldots, a_{kn_2}]' = [0, \ldots, \underset{\substack{\uparrow \\ k^{\text{ième}} \text{ place}}}{1}, \ldots, 0]' \qquad k \in]r_2]$$

Pour chacun de ces systèmes, on calcule le lagrangien $\mathcal{L}_k$ $(1 \leqslant k \leqslant r_2)$. Annulant les dérivées partielles de tous ces lagrangiens, on obtient un système linéaire d'équations qui permet de calculer les a_{ki}, i.e la matrice W_2^-. Pour plus de détails sur ce calcul, on pourra se reporter à [7].

De même, pour le calcul de W_1^-, on choisit parmi l'ensemble des solutions celle qui, pour chacun des r_1 systèmes, correspondant aux r_1 colonnes de la matrice W_1^-, minimise la fonction suivante :

Pour le $k^{\text{ième}}$ système $(k \in]r_1])$

et avec $W_1^- = (b_{ik})_{\substack{i=1,\ldots,m_1 \\ k=1,\ldots,r_1}}$

$$\tilde{f}(b_{1k}, \ldots, b_{ik}, \ldots, b_{m_1k}) = \sum_{j=1}^{N_1+1} \frac{1}{M(j)} \left(\sum_{q=A}^{B} b_{qk} \right)^2$$

(où A et B désignent les mêmes quantités que précédemment)

Sous les contraintes :

$$W_1 W_1^- = I_{r_1} \qquad (**)$$

En associant à la matrice $W_1(r_1, m_1)$ l'application $\tilde{g}$ de $\mathbf{R}^{m_1}$ dans $\mathbf{R}^{r_1}$, on peut écrire (**) sous la forme de r_1 contraintes :

$$\tilde{g}[b_{1k}, \ldots, b_{ik}, \ldots, b_{m_1k}]' = [0, \ldots, \underset{k^{\text{ième}} \text{ place}}{1}, \ldots, 0]' \quad k \in]r_1]$$

On calcule comme précédemment les b_{ik} en annulant les dérivées partielles des lagrangiens associés à chacun de ces systèmes.

Remarque importante : Ce critère de centrage n'est pas l'unique solution valable : il nous est apparu comme un moyen facilitant l'interprétation des résultats et respectant des procédés déjà utilisés (dans ANOVA par exemple). Mais ceci n'exclut pas le choix d'une autre adéquation.

Remarque 2 : Dans le centrage par bloc du tableau T^*, on a utilisé le fait suivant :
si B est une matrice centrée en ligne, alors AB est centrée en ligne
si A est une matrice centrée en colonne, alors AB est centrée en colonne.

3. DEUXIEME APPROCHE ET COMPARAISON

Cette approche est une méthode indirecte puisqu'elle met en œuvre une procédure analogue à celle des moindres carrés alternés. Par opposition à la première qui est analytique, nous dirons qu'elle est algorithmique. Après en avoir décrit les principales étapes, nous comparons les deux approches sur des exemples réalisés par simulation.

3.1. L'algorithme de la méthode indirecte

• On commence par calculer la moyenne par colonne de Y (c'est-à-dire par objet).

Soit $\underset{(n,1)}{\bar{Y}^{(0)}} = \underset{(n,1)}{1} \underset{(1,p)}{(\bar{y}_1, \ldots, \bar{y}_p)}$ où $\bar{y}_j = \frac{1}{n} \sum_{i=1}^{n} y_{ij}$ avec $j \in \{1, \ldots, p\}$

1) On cherche $\tilde{T}^{(1)}$ tel que $\underset{(n,p)}{\bar{Y}^{(0)}} \simeq \underset{(n,n_2)}{\tilde{T}^{(1)}} \underset{(n_2,p)}{X_2'}$ ce qui revient à chercher $\tilde{t}^{(1)}$ tel que :

$$\bar{Y}^{(0)} \simeq 1\, \tilde{t}^{(1)} X_2'$$

On résout donc le problème :

$$\underset{\tilde{T}}{\text{Min}} \| \bar{Y}^{(0)} - \tilde{T} X_2' \|^2$$

La décomposition de X_2 permet d'écrire :

$$\underset{\widetilde{T}}{\text{Min}} \| \underset{(n,p)}{\overline{Y}^{(0)}} - \underset{(n,n_2)}{\widetilde{T}} \underset{(n_2,r_2)}{V_2} \underset{(r_2,r_2)}{\overline{\beta}_2} \underset{(r_2,p)}{U'_2} \|^2$$

avec $$U'_2 U_2 = I_{r_2}$$

D'après la proposition 1, on sait qu'une solution de ce problème est $\widetilde{T}^{(1)}$ tel que :

$$\underset{(n,p)}{\overline{Y}^{(0)}} \underset{(p,r_2)}{U_2} = \widetilde{T}^{(1)} V_2 \overline{\beta}_2 = \underset{(n,n_2)}{\widetilde{T}^{(1)}} \underset{(n_2,r_2)}{W_2}$$

Résoudre ce système revient à chercher une matrice W_2^- telle que :

$$\underset{(r_2,n_2)}{W_2^-} \underset{(n_2,r_2)}{W_2} = I_{r_2}$$

la solution n'étant pas unique, on choisit alors celle qui centre au mieux le vecteur $\widetilde{t}^{(1)}$ par bloc de variable (cf. paragraphe 2).

2) On pose $\overline{Y}^{(1)} = \widetilde{T}^{(1)} X'_2$.

3) On cherche $T^{(1)}$ tel que $(Y - \overline{Y}^{(1)}) \simeq X_1 T^{(1)} X'_2$ (cf. paragraphe 2 en remplaçant T^* par T). La détermination de la solution se fait comme précédemment (§ 2) par une succession de deux systèmes linéaires "en cascade" puis par un centrage de T sur les variables objets et les variables sujets.

4) On détermine alors $\overline{Y}^{(2)}$ correspondant à la moyenne par colonne de $(Y - X_1 T^{(1)} X'_2)$.

5) On recommence en (1) en cherchant $\widetilde{T}^{(2)}$ tel que :

$$\overline{Y}^{(2)} \simeq \widetilde{T}^{(2)} X'_2$$

la solution étant donnée par $\widetilde{T}^{(2)} = [\overline{Y}^{(2)} U_2] W_2^-$.

On montre dans [7] que le processus converge en un nombre fini d'itérations.

3.2. Comparaison des deux approches

On a réalisé plusieurs exemples par simulation afin de tester les deux procédures.

Il est intéressant de constater que si les deux approches fournissent exactement la même approximation du tableau des préférences Y, les solutions obtenues (T et $\widetilde{t}$) n'en sont pas moins différentes. Ce résultat n'est guère étonnant puisque le modèle consiste à choisir parmi un ensemble de solutions admissibles celle qui centre au mieux par bloc de variables :

– dans un cas le tableau T^* complet ;
– dans l'autre cas successivement le vecteur $\widetilde{t}$ et le tableau T.

On obtient donc, dans les deux approches, des solutions admissibles distinctes : l'une est en effet centrée "au mieux" globalement et l'autre par morceaux.

L'exemple suivant concrétise bien ce résultat.

On considère le tableau de préférences Y de dimension (6,4) suivant :

6 sujets ont noté 4 objets.

	O_1	O_2	O_3	O_4
S_1	17	5	9	−3
S_2	5	1	13	9
S_3	7	7	7	7
S_4	−5	3	11	19
S_5	−5	3	11	19
S_6	−5	3	11	19

Y
(6,4)

On a volontairement pris trois sujets ayant des préférences semblables. Ce sont S_4, S_5 et S_6.

Le tableau X_1 des caractéristiques-sujets est le suivant. Avec deux variables qualitatives ayant chacune deux modalités et sous forme disjonctive complète.

	1^ère^ variable-sujet		2^ème^ variable-sujet	
	m.1	m.2	m.1	m.2
S_1	1	0	1	0
S_2	1	0	0	1
S_3	0	1	1	0
S_4	0	1	0	1
S_5	0	1	0	1
S_6	0	1	0	1

X_1
(6,4)

Toujours volontairement S_4, S_5, S_6 ont les mêmes caractéristiques.

Le tableau X_2 des variables-objets est de dimension (4,4) sous forme disjonctive complète (avec deux variables ayant chacune deux modalités).

	1ère variable-objet		2ème variable-objet	
	m.1	m.2	m.1	m.2
O_1	1	0	1	0
O_2	1	0	0	1
O_3	0	1	1	0
O_4	0	1	0	1

X_2
(4,4)

Les résultats fournis par les deux approches sont les suivants (il convient de remarquer que, dans la procédure alternée, le vecteur $\tilde{t}$ apparaît en première ligne du tableau tandis que, dans la méthode directe, il se trouve à la dernière ligne, ceci étant réalisé dans un souci pratique de distinction des résultats).

On constate que la solution fournie par la procédure alternée est imparfaitement centrée en colonne par bloc de variable. En revanche, le vecteur $\tilde{t}$ approche mieux la note moyenne établie sur l'ensemble des 6 sujets. En effet, prenons par exemple l'objet O_2. Il est caractérisé par :

– la première modalité de la première variable-objet ;
– la seconde modalité de la seconde variable-objet.

Procédure alternée

		1ère variable-objet modalité 1	1ère variable-objet modalité 2	2ème variable-objet modalité 1	2ème variable-objet modalité 2	
		-1.1	8.1	2.4	4.6	$\tilde{t}$ (1,4)
1ère var. sujet	modalité 1	3.3	-3.3	4.1	-4.1	T (4,4)
	modalité 2	-0.7	0.7	-1.9	1.9	
2ème var. sujet	modalité 1	5.3	-5.3	3.1	-3.1	
	modalité 2	-2.7	2.7	-0.9	0.9	

Procédure directe

1ère var. sujet	modalité 1	2.	$-2.$	3.	$-3.$	T* (5,4)
	modalité 2	$-2.$	2.	$-3.$	3.	
2ème var. sujet	modalité 1	4.	$-4.$	2.	$-2.$	
	modalité 2	$-4.$	4.	$-2.$	2.	
		1.5	5.5	4.5	2.5	

Grâce au tableau Y, on peut calculer la note moyenne exacte donnée à cet objet O_2 par les 6 sujets :

$$\frac{1}{6}(5 + 1 + 7 + 3 + 3 + 3) = 22/6 = 3.6$$

La note moyenne approchée obtenue par $\tilde{t}$ est :

$$-1.1 + 4.6 = 3.5$$

Elle est donc très voisine de la note réelle.

Par ailleurs, si la solution fournie par la procédure directe est parfaitement centrée en ligne et en colonne par bloc de variable, le vecteur $\tilde{t}$ approche de manière exacte les notes moyennes sur les objets calculées non pas en tenant compte des 6 sujets mais des 4 premiers : tout revient à négliger le fait que trois individus ont des préférences identiques.

Conservons l'objet O_2 comme exemple de référence. La note moyenne exacte donnée à cet objet par S_1, S_2, S_3 et S_4 est $1/4\ (5 + 1 + 7 + 3) = 4$. La note moyenne approchée fournie par $\tilde{t}$ est :

$$1.5 + 2.5 = 4.$$

L'approximation est parfaite mais elle ne tient pas compte de la pondération à attribuer au sujet S_4.

La non unicité de la solution provient, comme on l'a vu, du choix qui est fait parmi l'ensemble des solutions admissibles de celle qui respecte au mieux des conditions de centrage, ceci dans un souci d'interprétation des résultats.

Tout naturellement, on peut poser le problème de façon différente et s'intéresser à l'existence de solutions centrées minimisant le même critère : autrement dit, on peut chercher à résoudre le problème d'optimisation sous contraintes suivant (cf. [7]) :

$$(P_3)\ \underset{T^*}{\mathrm{Min}} \| \underset{(n,p)}{Y} - \underset{(n,m_1)}{X_1^*}\ \underset{(m_1,n_2)}{T^*}\ \underset{(n_2,p)}{X'_2} \|^2$$

sous :

$$\begin{cases} AT^* = 0 \\ T^*B = 0 \end{cases}$$

avec

$$\underset{(N_1+1,\,m_1)}{A} = \begin{array}{c} \begin{array}{ccccc} M(1) & M(2)\ldots & M(j)\ldots & M(N_1) & 1 \end{array} \\ \begin{bmatrix} [1\ldots1] & & & \bigcirc & \\ & [1\ldots1] & & & \\ & & [1\ldots1] & & \\ & \bigcirc & & [1\ldots1] & \\ & & & & 0 \end{bmatrix} \end{array}$$

et

$$\underset{(n_2,\,N_2)}{B} = \begin{array}{c|c} & 1\ \ldots\ j\ \ldots\ N_2 \\ \hline \begin{array}{c} M(1) \\ \\ M(2) \\ \vdots \\ \\ M(N_2) \end{array} & \begin{bmatrix} \boxed{\begin{matrix}1\\ \vdots \\ 1 \\ \vdots \\ 1\end{matrix}} & & & \bigcirc \\ & \boxed{\begin{matrix}1\\ \vdots \\ 1\end{matrix}} & & \\ & & \boxed{\begin{matrix}1\\ \vdots \\ 1\end{matrix}} & \\ \bigcirc & & & \boxed{\begin{matrix}1\\ \vdots \\ 1\end{matrix}} \end{bmatrix} \end{array}$$

Ce problème fournit certes une solution unique parfaitement centrée mais l'approximation du tableau des préférences Y qu'elle donne est au mieux celle qui est atteinte par (P_1) ou (P_2).

L'arbitrage est alors difficile à faire : doit-on gagner en centrage et perdre en approximation de Y ou l'inverse ? Une étude approfondie de (P_3) permettrait peut-être de moduler la réponse.

4. UNE APPLICATION DE BENITO

Une application du modèle BENITO a été réalisée sur des données réelles fournies par la SOFRES.

L'objet de l'étude est de déterminer la valeur qualitative spécifique conférée par un support de presse aux produits dont la publicité lui est confiée.

Cette valeur a été recherchée pour trois dimensions qualitatives, chacune d'elles pouvant constituer un thème sous-jacent à un message publicitaire :

- le thème de la nouveauté
- le thème de la personnalisation, c'est-à-dire de l'expression de soi dans la nouveauté
- le thème du choix judicieux dans la nouveauté.

4.1. Le recueil de l'information

Trois échantillons de 250 femmes ont été appareillés, en fonction de variables socio-démographiques, de manière que toute personne d'un échantillon ne réfléchisse que sur un thème.

Les interviewées sont placées dans le contexte d'un scénario qui a pour but de les mettre dans la meilleure situation de jugement.

Le critère est assimilé de la manière suivante : une interviewée doit s'identifier à une personne sensibilisée intrinsèquement par un thème.

Ainsi, par exemple :

- Pour le thème de la nouveauté, la phrase suivante est répétée lors de l'interview : "Et si l'on prenait le cas d'une femme qui, dans les choses nouvelles, aime ce qui est vraiment très nouveau".
- Pour le thème de la personnalisation : "Et si l'on prenait le cas d'une femme qui, dans les choses nouvelles, cherche la petite note personnelle pour les adapter à son style à elle".
- Pour le thème du choix judicieux : "Et si l'on prenait le cas d'une femme qui, dans les choses nouvelles, n'achète pas n'importe quoi, sait choisir ce qui est vraiment judicieux".

Les supports de presse considérés sont au nombre de quatre : les trois supports A, B, C font partie de la presse féminine tandis que le quatrième, le support D, appartient aux "news" (hebdomadaires d'information générale).

On a sélectionné d'autre part les 6 produits suivants :

- tenue de loisir ou de détente
- modèle de lingerie féminine
- living, séjour
- cuisine
- séjour de vacances
- appartement en multipropriété pour les vacances d'hiver ou d'été.

On a ainsi défini 24 couples produits × supports, chacun représentant une publicité qui aurait pu paraître pour le **produit** concerné dans le **support** correspondant.

Chacun de ces couples est matérialisé par un carton sur lequel est inscrit la nature du produit et le nom du support où est supposée avoir paru la publicité. Il est alors demandé à l'interviewée de classer ces publicités les unes par rapport aux autres en fonction de l'idée qu'elle se fait a priori de l'adéquation entre le thème publicitaire et le produit : ce jeu n'ayant pas été correctement compris par l'ensemble de la population, on n'a retenu dans l'analyse que les réponses de 456 femmes. (Cette sélection a pu être établie grâce à certaines questions-pièges).

Le tableau 1 décrit les caractéristiques des interviewées et celles des objets classés.

TABLEAU 1

Description des caractéristiques des juges et de celles des objets

Caractéristiques des juges	Caractéristiques des objets
critère	**publicité**
1. nouveau	1. tenue de loisir ou de détente
2. note personnelle	2. modèle de lingerie féminine
3. judicieux	3. séjours de vacances
lecture du support A	4. living
1. oui	5. appartement en multipropriété
2. non	6. cuisine
lecture du support C	**support**
1. oui	1. A
2. non	2. B
lecture du support D	3. C
1. oui	4. D
2. non	
lecture du support B	
1. oui	
2. non	
âge	
1. 15-17 ans	
2. 18-20 ans	
3. 21-24 ans	
4. 25-34 ans	
5. 35-49 ans	
6. 50-64 ans	
7. 65 ans ou plus	
activité professionnelle	
1. oui	
2. non	
profession du chef de famille	
1. profession libérale, cadre supérieur, gros patron	
2. petit patron, artisan	
3. cadre moyen	
4. employé, divers	
5. ouvrier	
6. inactif	

4.2. Problèmes posés

Les questions posées étaient les suivantes :

- Existe-t-il une valeur ajoutée propre au sujet indépendamment des produits publicités ?
- Comment les supports se hiérarchisent-ils par rapport à cette valeur ajoutée ?
- Existe-t-il une valeur ajoutée produit indépendante du support, et surtout : y a-t-il des produits dont l'image aurait tout intérêt à bénéficier d'un apport du titre pour appuyer le message publicitaire ?
- Quel gain d'image peut-on attendre pour chacun des produits dans différentes hypothèses de choix média ?

Remarque 1 : La méthode MONANOVA qui avait été utilisée auparavant n'avait pu répondre que partiellement aux questions posées.

Remarque 2 : On ne doit jamais oublier que les classements initiaux ont été faits sur une échelle de 1 à 24 correspondant pratiquement à des rangs puisque plus la note est faible et plus le couple produit-support est préféré.

4.3. Résultats

Notre objectif n'est pas d'expliciter les résultats obtenus dans cette étude spécifique. Nous cherchons simplement, au travers d'un exemple, à montrer le fonctionnement du modèle BENITO et l'exploitation que l'on peut en faire.

Le tableau 2 fournit un exemple de sorties du programme.

Essayons d'approximer la note donnée, indépendamment du thème publicitaire, à une publicité de modèle de lingerie féminine dans le support A par une femme qui lit exclusivement le support A : il suffit d'ajouter les coefficients entourés soit :

$$\underbrace{3.9 + 4.7}_{\substack{\text{note moyenne du}\\ \text{couple publicité-support}}} \quad \underbrace{+ .2 + .1 - .5 - .1 - .6 - .4 + .2 + .2 = 7.7}_{\substack{\text{corrections induites par la lecture}\\ \text{exclusive du support A}}}$$

Il est bien évident que toutes les combinaisons possibles peuvent être faites entre thème publicitaire – publicité – et support. On peut également s'intéresser à l'influence exclusive de certaines variables sur la note moyenne.

On peut reporter sur un graphique les valeurs ajoutées obtenues par telle ou telle variable ou combinaison de variables.

TABLEAU 2

			publicité						support			
			1	2	3	4	5	6	A	B	C	D
$\tilde{t}$		456	3.6	3.9	5.0	4.6	6.7	4.5	4.7	5.0	7.3	11.2
	critère											
	1	165	.1	.0	-.2	.3	-.6	.5	-.2	-.0	.4	.0
	2	134	.0	.2	-.7	.0	.2	.1	.4	.1	-.6	-.0
	3	157	-.0	-.2	.8	-.3	.3	-.6	-.1	.1	.3	-.4
	lecteur A											
	oui	313	-.1	.2	.0	.0	-.1	.0	-.6	.3	.6	-.3
	non	143	.2	-.1	-.2	.0	.0	.1	.7	-.2	-.5	-.0
	lecteur C											
	oui	301	.1	.0	-.1	-.0	-.1	.2	.5	.3	-.8	.0
	non	155	-.0	.1	-.1	.1	.0	-.1	-.4	-.2	.9	-.3
	lecteur D											
	oui	141	.2	.6	-.6	-.1	-.3	.1	-.1	.2	.8	-.9
	non	315	-.2	-.5	.4	.1	.2	.0	.2	-.0	-.7	.6
	lecteur B											
	oui	351	-.0	.2	-.2	.1	-.2	.0	-.1	-.8	.4	.5
T	non	105	.1	-.1	-.0	-.1	.0	.1	.2	1.0	-.3	-.9
	âge											
	1	26	.4	.6	-.7	.7	-2.1	1.0	.4	-.5	.4	-.5
	2	30	-.6	-.3	.3	.6	-.5	.7	-.5	-.4	.2	1.0
	3	51	-.3	.3	-.2	-.1	.0	.2	-.5	-.3	.9	-.2
	4	118	-.4	-.9	.5	-.1	1.0	-.1	-.2	.0	-.1	.4
	5	145	.1	.1	-.0	-.8	.7	-.2	-.7	-.2	.1	.6
	6	61	.1	.1	.1	-.9	1.1	-.5	.5	.2	-.5	-.2
	7	25	.8	.3	-.3	.7	-.4	-1.0	1.1	1.2	-1.0	-1.3
	activité professionnelle											
	oui	203	.3	-.0	-.3	.1	-.2	.1	.1	-.0	.1	-.1
	non	253	-.3	.1	.2	-.0	.0	-.1	.0	.2	-.0	-.3
	profession chef de famille											
	1	118	-.2	.1	.5	.5	-.7	-.3	.4	-.4	-.1	-.1
	2	20	-.9	-.9	-.5	.5	1.0	1.1	-.0	.4	1.1	-1.2
	3	94	-.8	-.9	.5	.2	.7	.4	-.3	-.1	.2	.1
	4	79	.2	-.0	.3	-.2	-.3	.0	.0	.6	-.4	-.3
	5	95	.8	-.3	.4	-.2	.3	-1.0	.2	-.1	-.2	.2
	6	50	1.1	2.0	-1.3	-.6	-1.2	-.1	-.3	-.2	-.6	1.0
		Effectifs dans chaque catégorie	Tenue de loisir	Modèle de lingerie féminine	Circuits ou séjours de vacances	Living	Appartement en multipropriété	Cuisine				

CONCLUSIONS

Si l'on voulait résumer les propriétés du modèle BENITO, on dirait qu'il s'agit d'un modèle métrique, totalement additif et prenant en compte les interactions des caractéristiques sujets et objets dans l'explication des préférences de ces individus envers ces objets.

Cependant, ces particularités sous-entendent un certain nombre d'exigences quant à la nature des données qui sont traitées : ainsi, l'additivité du modèle réclame la mise sous forme disjonctive complète des matrices X_1 et X_2. De plus, le tableau Y doit rassembler les préférences complètes évaluées sur des échelles : tout jugement recueilli sous forme de rang (préordre total) est assimilé à une échelle.

Essentiellement descriptif, le modèle BENITO permet cependant d'approcher les notes de préférences qu'aurait donné un individu non intervenu dans l'analyse mais dont les caractéristiques sont celles de X_1. On définit ainsi son profil de réponse grâce au tableau T^*. Ce modèle permet également de situer un objet par rapport aux p autres grâce à ses seules caractéristiques, évitant ainsi la constitution de toutes les combinaisons des caractéristiques d'objets.

Des exemples de simulation ont été réalisés afin de tester la stabilité des coefficients du tableau T et du vecteur $\tilde{t}$. Pour cela, on a déterminé a priori les matrices T et $\tilde{t}$ et, compte tenu des matrices X_1 et X_2 que l'on avait choisies, on a calculé le tableau Y des préférences. C'est ce dernier que l'on a utilisé dans le modèle afin de pouvoir comparer les matrices T et $\tilde{t}$ évaluées par le modèle avec celles qui nous avaient servi de point de départ. On a constaté que la solution donnée par le modèle fournissait une "approximation exacte" de la matrice Y. Enfin, en introduisant volontairement une erreur dans Y, on a remarqué que l'ensemble des résultats restait stable en dépit de cette modification introduite dans les données.

L'intérêt de BENITO est à la fois de pouvoir moduler l'explication des préférences selon le profil des interviewés sans typologie a priori et de conserver une interprétation individuelle aux résultats.

REFERENCES

[1] BOUROCHE J.M. – "Analyse des données en marketing". *Monographies de l'AFCET*, Masson, 1977.

[2] CARROLL J.D., CHANG J.J. – "Non-parametric multidimensional analysis of paired comparisons data". Paper presented at the Joint Meeting of the Psychometric and Psychonomic Societies, Niagara Fall, 1964.

[3] CARROLL J.D., CHANG J.J. – "Relating preference data to multidimensional scaling solutions via a generalization of the Coomb's unfolding model". Paper presented at the meeting of Psychometric Society, Madison, Wi., 1967.

[4] CARROLL J.D., GREEN P.E., CARMONE F.J. – "Super Ordinate Factorial Designs in the Analysis of Consumer Judgments". Handout Bell Laboratories, Murray Hill, New Jersey, 1975.

[5] CARROLL J.D., GREEN P.E., CARMONE F.J. – "CANDELINC (CANonical DEcomposition with LINear Constraints)". International Congress of Psychology in Paris, July 18-25, 1976.

[6] CARROLL J.D., PRUZANSKY S. – "MULTILINC : MULTI-WAY CANDELINC". Meeting of the American Psychological Association in San Francisco, (California), August 26-30, 1977.

[7] DUPONT-GATELMAND C. – "Deux méthodes d'analyse des données qualitatives : typologie avec codage adaptatif et modèle de décomposition additive des préférences". Thèse de 3^e Cycle, Université Paris-Dauphine, avril 1978.

[8] GREEN P.E., WIND Y.E. – "Multiattribute Decisions in Marketing". The Dryden Press, 1973.

[9] JACQUET-LAGREZE E., BERGES J.C. – "ADEPREF : une méthode et un programme de données ordinales". Note de travail n° 20, COREF, 1977.

[10] JACQUET-LAGREZE E., SISKOS J. – "Une méthode de construction d'une fonction d'utilité additive, explicative d'une préférence globale". Université Paris-Dauphine, *Cahier LAMSADE*, n° 16, mai 1978.

[11] KRUSKAL J.B. – "Analysis of factorial experiments by estimating monotone transformations of the data". *Journal of the Royal Statistical Society*, Series B 27, 1965, p. 251-263.

[12] RAO C.R., MITRA S.K. – "Generalized Inverse of Matrices and its Applications". Wiley, New York, 1971.

[13] SEARLE S.R. – "Linear Models". Wiley, 1971.

[14] SRINIVASAN V., SHOCKER A.D. – "Linear programming techniques for multidimensional analysis of preferences". *Psychometrika*, Vol. 38, n° 3, 1973.

[15] SRINIVASAN V., SHOCKER A.D. – "Estimating the weights for multiple attributes in a composite criteria using pairwise judgments". *Psychometrika*, Vol. 38, n° 4, 1973.