



**HAL**  
open science

# Régression et analyse canonique sous contraintes linéaires, algorithmes et applications

Abdelkader Agha

► **To cite this version:**

Abdelkader Agha. Régression et analyse canonique sous contraintes linéaires, algorithmes et applications. Optimisation et contrôle [math.OC]. Université Paris 9 1991. Français. NNT : 1991PA090016 . tel-01579588

**HAL Id: tel-01579588**

**<https://theses.hal.science/tel-01579588>**

Submitted on 10 Sep 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**UNIVERSITE PARIS IX DAUPHINE**  
**UER SCIENCES DES ORGANISATIONS**

Thèse pour l'obtention du titre de

**DOCTEUR**

en Méthodes Scientifiques de Gestion

par

**Abdel Kader AGHA**

**SUJET**

Régression et Analyse Canonique sous contraintes  
linéaires, algorithmes & applications

**MEMBRES DU JURY**

<b>MM. Pierre CAZES</b>	<b>Professeur à PARIS IX</b>
<b>Alain MORINEAU</b>	<b>Directeur du CISIA</b>
<b>Gilbert SAPORTA</b>	<b>Professeur au CNAM</b>
<b>Michel TENENHAUS</b>	<b>Professeur à HEC</b>

Soutenue le 3 juillet

- 1 9 9 1 -

## **AVERTISSEMENT**

La qualité de numérisation de ce fichier dépendant de l'état général de la microfiche, l'A.N.R.T. ne peut garantir un résultat irréprochable.

Je remercie

Monsieur Pierre CAZES, Professeur à l'Université Paris IX Dauphine, pour avoir bien voulu me faire l'honneur de juger cette thèse qui a bénéficié de ses précieuses observations. Qu'il me soit permis de lui exprimer toute ma reconnaissance

Monsieur Alain MORINEAU, Directeur du Centre International de Statistiques et d'Informatique Appliquée, pour l'intérêt qu'il a porté à ce travail et l'honneur qu'il m'a fait en acceptant de participer au jury. Qu'il veuille bien trouver ici l'expression de toute ma gratitude.

Monsieur Gilbert SAPORTA, Professeur au Conservatoire National des Arts & Métiers, pour les précieux conseils et les encouragements qu'il m'a prodigués et pour avoir bien voulu me faire l'honneur de juger ce travail. Qu'il me soit permis de lui exprimer ma profonde reconnaissance.

Monsieur Michel TENENHAUS, Professeur à l'École des Hautes Etudes Commerciales, qui a inspiré et dirigé cette thèse, pour son soutien déterminant et ses conseils avisés. J'aurai toujours en mémoire son enseignement clair et rigoureux. Qu'il reçoive, par ce travail, un témoignage de ma déférente gratitude.

Je tiens à exprimer tous mes remerciements à Mesdames Odile BAUDET et Josette REY, pour leur aide constante et efficace.



## AVERTISSEMENT

L'Université n'entend donner aucune approbation ou improbation aux opinions émises dans les thèses: ces opinions doivent être considérées comme propres à leurs auteurs.

## TABLE DES MATIERES

TABLE DES MATIERES.....	IV
INTRODUCTION.....	2
CHAPITRE I. REGRESSION SOUS CONTRAINTES LINEAIRES SUR LES COEFFICIENTS.....	15
1. PRELIMINAIRES.....	16
1.1. NOTATIONS.....	16
1.2. CONES POLYEDRIQUES CONVEXES.....	17
1.3. CONDITIONS DE KUHN & TUCKER.....	20
1.4. QUALITE DE LA REGRESSION.....	21
1.5. OPERATEUR SWEEP.....	22
1.5.1. DESCRIPTION.....	22
1.5.2. LE PIVOTAGE.....	23
1.5.3. APPLICATIONS.....	24
1.5.3.1. INVERSION MATRICIELLE.....	25
1.5.3.2. REGRESSION MULTIPLE.....	27
2. CONTRAINTES DE POSITIVITE.....	29
2.1. LE PROBLEME NNLS.....	29
2.1.1. CONDITIONS ADOPT.....	29
2.1.2. DESCRIPTION.....	30
2.1.3. INTEGRATION DE L'OPERATEUR SWEEP.....	31
2.1.3.1. INTEGRATION DE TYPE I.....	31
2.1.3.2. INTEGRATION DE TYPE II.....	31
2.1.3.3. VERSION MCCP.....	32
2.2. ETUDE COMPARATIVE REGCAZ/MCCP.....	33
3. CONTRAINTES DE BORNES.....	37
3.1. LE PROBLEME MCCE.....	37
3.2. L'ALGORITHME PROPOSE.....	37
3.2.1. CONDITIONS ADOPT.....	37
3.2.2. DESCRIPTION.....	38
3.3. ETUDE NUMERIQUE.....	39

3.3.1.ETUDE COMPARATIVE REGCAZ/MCCP.....	40
3.3.2.APPLICATION DE MCCB.....	42
3.4.FONDEMENTS THEORIQUES DE L'ALGORITHME MCCB.....	45
<b>4.CONTRAINTES D'INEGALITES LINEAIRES QUELCONQUES.....</b>	<b>50</b>
4.1.LE PROBLEME LDP.....	50
4.1.1.DESCRPTION DE LA METHODE.....	50
4.1.2.JUSTIFICATION DE LAWSON & HANSON.....	51
4.1.3.JUSTIFICATION PROPOSEE.....	51
4.2.LE PROBLEME MCCIL.....	53
4.2.1.LA METHODE LSI DE LAWSON & HANSON.....	54
4.2.2.LA METHODE MCCIL.....	55
4.2.3.MCCIL GENERALISE.....	57
4.3.ETUDE NUMERIQUE.....	59
<b>CHAPITRE II.ANALYSE CANONIQUE SOUS CONTRAINTES</b>	
<b>DE POSITIVITE DES FACTEURS.....</b>	<b>63</b>
<b>1.PRELIMINAIRES.....</b>	<b>64</b>
1.1.NOTATIONS.....	64
1.2.NOTION DE STATIONNARITE.....	64
<b>2.ANALYSE CANONIQUE DE DEUX ESPACES VECTORIELS.....</b>	
2.1.POSITION DU PROBLEME.....	68
2.2.METHODES DE RESOLUTION.....	68
2.2.1.PRODUIT DE MATRICES DE COVARIANCES.....	68
2.2.2.ALGORITHME ALS.....	69
<b>3.ANALYSE CANONIQUE D'UN CONE POLYEDRIQUE CONVEXE</b>	
<b>ET D'UN ESPACE VECTORIEL.....</b>	<b>70</b>
3.1.POSITION DU PROBLEME.....	70
3.2.L'ALGORITHME ACCE.....	70
3.2.1.CONDITIONS ADOPT.....	70
3.2.2.DESCRPTION.....	72
3.3.ETUDE NUMERIQUE.....	73
3.3.1.ETAPES DE L'ALGORITHME ACCE.....	75
3.3.2.ANALYSE DES RESULTATS.....	76

3.3.3.ETUDE COMPARATIVE REGCAZ/MCCP.....	77
3.4.FONDEMENTS THEORIQUES DE L'ALGORITHME ACCE.....	79
<b>CHAPITRE III.CODAGE DE VARIABLE ORDINALE</b>	
<b>ET TAU DE KENDALL.....</b>	<b>87</b>
1.PRELIMINAIRES.....	88
1.1.NOTATIONS.....	88
1.2.CODAGE D'UNE VARIABLE QUALITATIVE.....	89
2.CODAGE OPTIMAL D'UNE VARIABLE ORDINALE.....	91
2.1.CODAGE ADMISSIBLE.....	91
2.2.CARACTERISATION DU CODAGE OPTIMAL.....	93
3.LE TAU DE KENDALL.....	96
3.1.DEFINITION DU TAU DE KENDALL.....	96
3.2.UNE VARIANTE RELATIVE AUX PREORDRES.....	97
3.2.1.AGREGATION LOGIQUE DE DEUX PREORDRES.....	99
3.2.2.INTERPRETATION GRAPHIQUE.....	103
3.2.3.ETUDE NUMERIQUE.....	105
4.CODAGE ET TAU DE KENDALL.....	106
4.1.UNE FORMULE ORIGINALE DU TAU DE KENDALL.....	106
4.2.TAU DE KENDALL OPTIMAL.....	107
<b>CHAPITRE IV.PRESENTATION DU PROGRAMME REGALS</b>	
<b>ET EXEMPLES D'APPLICATIONS.....</b>	<b>110</b>
1.LE PROGRAMME REGALS.....	111
1.1.DESCRPTION GENERALE.....	111
1.2.LES FICHIERS.....	111
1.3.LES METHODES.....	112
1.4.OPTIONS.....	112
1.5.ENVIRONNEMENT.....	112
2.REGRESSION SOUS CONTRAINTES.....	113
2.1.LE PROBLEME.....	113

2.2.LES METHODES.....	113
2.2.1.«PROTECTION» DE LA REGRESSION.....	113
2.2.2.REGRESSION MONOTONE.....	114
2.2.3.UN EXEMPLE DE REGRESSION MONOTONE.....	115
3.REGRESSION QUALITATIVE.....	117
3.1.LE PROBLEME.....	117
3.2.LES METHODES.....	119
3.2.1.LE PROGRAMME MORALS.....	119
3.2.2.ANALYSE DES MESURES CONJOINTES.....	120
3.2.2.1.LE PROGRAMME MONANOVA.....	120
3.2.2.2.UN EXEMPLE DE MONANOVA.....	122
4.LE TRADE-OFF.....	125
4.1.LE PROBLEME.....	125
4.2.LA METHODE.....	126
4.3 DESCRIPTION DU TRAITEMENT REGALS.....	126
4.4.UN EXEMPLE DE TRADE-OFF.....	127
CHAPITRE V.DESAGREGATION DE LA PREFERENCE GLOBALE EN ANALYSE MULTICRITERE.....	133
1.LE PROBLEME & LES METHODES.....	134
1.1.LE PROBLEME.....	134
1.2.METHODES DE RESOLUTION.....	136
1.2.1.LA METHODE UTA.....	137
1.2.1.1.DESCRPTION.....	137
1.2.1.2.EXEMPLE DE MISE EN OEUVRE.....	139
1.2.2.REGRESSION SOUS CONTRAINTES.....	142
1.2.3.LA METHODE REGALS.....	143
2.ETUDE COMPARATIVE UTA/REGALS.....	144
2.1.EXEMPLES ET RESULTATS.....	144
2.2.COMMENTAIRES.....	159
2.3.METRIQUES D'AGREGATION D'ECHELLES.....	162
2.4.CONCLUSION DE L'ETUDE COMPARATIVE.....	164
REFERENCES BIBLIOGRAPHIQUES.....	166

**INTRODUCTION**

## INTRODUCTION

La Régression Qualitative est mise en oeuvre selon deux approches distinctes: optimisation d'un critère au sens des moindres écarts quadratiques («moindres carrés») ou des moindres écarts absolus («moindres écarts»). Elle a suscité ces dernières années de nombreux travaux car, que ce soit en Analyse Multicritère ou en Analyse des Mesures Conjointes, ses méthodes trouvent un vaste champ d'application.

En nous situant dans une approche moindres carrés, nous considérons le problème général comme une Analyse Canonique sous Contraintes (Tenenhaus, 1979), requérant lors de sa mise en oeuvre, la détermination de la projection orthogonale d'un vecteur sur un polyèdre convexe.

Ce dernier problème rentre dans le cadre de la Régression Multiple sous Contraintes sur les coefficients et permet, en outre, d'apporter une réponse à la question de la «protection» de la régression, rencontrée lors du traitement de données en économie, géologie, physique, etc...

Notre étude nous a amené à définir la notion d'«approche de type ADOPT». Notion essentielle et féconde, elle est à la base des algorithmes MCCB et ACCE que nous proposons et qui résolvent les problèmes des contraintes, respectivement de bornes en régression et de positivité des facteurs en analyse canonique.

L'application de la régression qualitative en analyse multicritère est abordée dans le problème de la Désagrégation de la Préférence Globale. Nous y effectuons une étude comparative des approches moindres carrés et moindres écarts.

Le programme informatique REGALS que nous avons mis au point, a permis le traitement des exemples qui illustrent les méthodes étudiées tout au long de ce document.

Nous décrivons ci-après, par le détail, nos travaux.

Le chapitre I est consacré au problème de la Régression Multiple sous Contraintes, qui consiste à expliquer une variable numérique dite critère, par un ensemble de variables numériques explicatives dites prédicteurs, en lui adjoignant un système de contraintes linéaires sur les coefficients de régression.

De nombreux auteurs ont abordé la régression multiple dans une approche moindres écarts. Les idées de Singleton (1940) ont été à la base de la plupart des travaux ultérieurs. Charnes, Cooper et Ferguson ont identifié le problème primal (1955). Wagner a proposé la formulation du problème dual (1959). Barrodale et Roberts y ont apporté (1972, 1973) des améliorations notables. La souplesse de l'approche permet la prise en compte de contraintes linéaires sur les coefficients.

Dans le cadre de l'approche moindres carrés du problème des contraintes de positivité, les recherches n'ont pas été moins actives: Mantel (1969) a proposé une approche basée sur une recherche arborescente. Waterman a proposé d'effectuer toutes les régressions possibles (1974) puis a intégré un critère de sélection (1977) inspiré des travaux de Mantel. L'algorithme NNLS (NonNegative Least Squares) de Lawson & Hanson (1974) est basé sur les conditions de Kuhn & Tucker. "Branch-and-bound" de Armstrong & Frome (1976) est le plus achevé des modèles de recherche arborescente. Bremner a proposé (1982) un algorithme basé sur l'opérateur SWEEP de A.E Beaton (1964).

L'intégration de l'opérateur SWEEP dans l'algorithme NNLS a été réalisée par P.Cazes, D.Bergougnan et C.Mullon (programme REGCAZ 1984). Nous avons proposé une étude similaire dans le Cahier du Lamsad n°87 (1988) selon une approche différente. Une analyse générale de ce problème nous a amené à une variante de REGCAZ que nous nommons MCCP (Moindres Carrés sous Contraintes de Positivité).

Une description géométrique, élégante, de NNLS, a été proposée par Tenenhaus (1988). L'idée d'en «extraire» les principes fondateurs nous a amené à définir la notion d'algorithme de type ADOPT que nous allons présenter.



Soit un problème concernant dans un ensemble  $E$  donné, parmi les éléments, dits admissibles, d'une partie finie  $A$  de  $E$ , la recherche d'un ou plusieurs éléments... optimaux au sens d'un critère. Si l'on est assuré de l'existence d'un sous-ensemble optimal, s'il y'a un test permettant d'identifier le caractère optimal d'un élément admissible, et si on dispose d'une procédure permettant dans  $A$ , à partir d'un élément, de construire un élément «meilleur» au sens du critère, alors ce problème est résoluble par un algorithme de type ADOPT (ADmissibilité-OPTimalité).

Cette définition appelle deux remarques:

-La condition de finitude de  $A$  sera en pratique, généralement établie par celle de  $E$ . Elle impose un choix subtil de l'élément recherché : ainsi dans NNLS, on veut déterminer un vecteur parmi une infinité, mais on tourne la difficulté en recherchant un sous-ensemble optimal de vecteurs générateurs dont la détermination résoud le problème.

-La condition d'admissibilité n'est pas précisée; elle suppose l'existence de «candidats» en nombre fini possédant une «qualité» liée à la nature du problème posé: ainsi, dans NNLS, un sous-ensemble de vecteurs générateurs est admissible si la projection orthogonale, sur l'espace vectoriel qu'il engendre, appartient au cône polyédrique convexe.

Une condition nécessaire à l'existence d'un algorithme de type ADOPT, résolvant un problème donné, repose sur le principe fondamental suivant:

«L'élément optimal recherché fait partie d'un ensemble fini d'éléments-candidats possédant une particularité connue».

Dans le cas de NNLS, cette condition est:

«La projection orthogonale sur le cône polyédrique coïncide avec la projection sur un espace vectoriel engendré par un sous-ensemble de l'ensemble des vecteurs générateurs du cône».

Le problème des contraintes de bornes sur les coefficients a été traité par P.Cazes et P.Y.Turpin (1971) par une méthode de programmation non linéaire.

L'«approche ADOPT» a été appliquée avec succès à ce problème. Nous avons établi (proposition I.6, section 3.4) la condition ADOPT nécessaire:

«La projection orthogonale sur le polyèdre convexe, découlant des contraintes de bornes imposées aux coefficients de régression, coïncide avec la projection sur une variété linéaire engendrée par un sous-ensemble de l'ensemble des vecteurs générateurs du polyèdre».

Le test d'optimalité a suivi ce résultat (proposition I.7, section 3.4). Il ne restait plus qu'à déterminer la procédure d'amélioration et initialiser «convenablement» l'algorithme.

Nous avons, ainsi, mis au point l'algorithme MCCB (Moindres Carrés sous Contraintes de Bornes), réalisant une généralisation de NNLS au cas de contraintes de bornes autres que 0 et  $+\infty$ .

La méthode LDP (Least Distance Programming), de Lawson & Hanson (1974), basée sur NNLS, résoud le problème de la recherche d'un vecteur de norme minimale dans un polyèdre.

Nous décrivons la justification de Lawson & Hanson consistant à montrer que la solution obtenue vérifie les conditions de Kuhn & Tucker (1951) du problème posé.

Une explication géométrique originale de LDP est proposée; elle est fondée sur deux propositions: I.11, I.12 (section 4.1.3). Un résultat de R.Faure (1979) nous a mis sur la voie de la proposition I.11 qui caractérise la non-vacuité du polyèdre. Ce résultat devient un cas-limite de notre proposition.

Le problème de la régression sous contraintes d'inégalités linéaires sur les coefficients revient à la détermination de la projection orthogonale d'un vecteur sur l'intersection d'un espace vectoriel et d'un polyèdre convexe. Il a été également

traité par Lawson & Hanson (1974). Leur méthode, dite LSI (Least Squares Inequality), est basée sur la notion de décomposition orthogonale d'une matrice et utilise l'algorithme LDP.

Nous traitons le problème selon une autre approche: la méthode MCCIL (Moindres Carrés sous Contraintes d'Inégalités Linéaires) basée aussi sur LDP, est de portée plus générale puisqu'elle permet la détermination de la projection orthogonale d'un vecteur sur l'intersection d'un polyèdre convexe et d'une variété linéaire. Outre sa relative facilité de mise en oeuvre, elle présente l'intérêt de donner une expression de la solution.

Le chapitre II concerne le problème de l'Analyse Canonique sous Contraintes.

Tenenhaus en a développé (1988) une théorie générale et a proposé un algorithme de résolution de type «Moindres Carrés Alternés» n'assurant pas l'obtention d'une solution optimale.

Le problème est limité au cas de contraintes de positivité sur les facteurs de l'un des ensembles de variables et à la recherche du premier couple canonique.

Le fait suivant nous a inspiré l'idée de l'existence d'un algorithme possédant une structure analogue à celle de NNLS:

«L'analyse canonique est une généralisation de la régression multiple au cas de deux ensembles de variables».

L'approche ADOPT a constitué, naturellement, le cadre de notre recherche, la condition nécessaire étant acquise:

«Le premier couple de l'analyse canonique d'un espace vectoriel et d'un cône polyédrique convexe coïncide avec le premier couple de l'analyse canonique de l'espace vectoriel initial et d'un espace vectoriel engendré par un sous-ensemble de vecteurs générateurs du cône».

En définissant la stationnarité comme critère d'optimalité, nous

avons développé l'algorithme ACCE (Analyse Canonique d'un Cône polyédrique convexe et d'un Espace vectoriel) en nous attachant à établir les conditions «ADOPT». Le caractère premier du couple obtenu n'a pas été établi mais la mise en oeuvre de cet algorithme n'est pas dépendante d'un vecteur initial.

Le chapitre III traite du problème du codage d'une variable qualitative et aborde l'étude de la corrélation des rangs de deux préordres à l'aide du tau de Kendall (1962).

Fait remarquable, «les méthodes de codage de variables qualitatives permettent de faire de l'analyse non linéaire de données avec des méthodes linéaires» (Saporta, 1978). La quantification des caractères - leurs poids et les utilités de leurs niveaux - provient d'une normalisation effectuée, sous l'hypothèse d'additivité, à partir des valeurs de fonctions numériques, éventuellement monotones, de ces caractères et non de la recherche d'une liaison linéaire.

De plus, en analyse de données ordinales, la notion de codage est fondamentale, car «si l'on peut utiliser des techniques conçues pour les données nominales (test du khi deux d'indépendance, analyse factorielle des correspondances sur le codage disjonctif complet), on perd une information essentielle constituée par le caractère ordinal des variables» (Chauchat, 1985).

La recherche du codage d'une variable ordinale «approchant au mieux» le vecteur des valeurs observées d'une variable numérique, amène à la notion de codage monotone optimal. Plusieurs auteurs ont développé la technique de ce type de codage: Kruskal (1965), Judge & Takayama (1966), Barlow & al (1972), Bertier et Tenenhaus (1974), de Leeuw (1975), Tenenhaus (1983).

Le tau de Kendall est communément utilisé comme mesure de la corrélation des rangs de deux ordres totaux. L'analyse d'une variante de cet indicateur, relative à des préordres et donc plus générale (Siskos, 1979), a abouti à des inégalités reliant les taus de Kendall de trois préordres quelconques (proposition III.6, section 3.2).

L'analyse des cas particuliers où ces relations deviennent des égalités, nous a amené à définir l'un des préordres comme une «agrégation logique» des deux autres (section 3.2.1, propositions III.6, III.7). Il est remarquable que certaines des conséquences déduites de ces relations, rejoignent l'intuition.

Une étude finale a permis d'une part, d'établir une formule simple du tau de Kendall entre le préordre et un de ses codages et, d'autre part, de relier codage optimal et tau de Kendall.

Le chapitre IV concerne les applications de la régression et de l'analyse canonique sous contraintes.

La régression sous contraintes traite principalement des problèmes de la régression monotone et «protégée».

La régression qualitative que l'on peut définir comme une extension de la régression multiple au cas d'un critère qualitatif et de prédicteurs de type quelconque, est le champ d'application principal de l'analyse canonique sous contraintes.

La prédominance du qualitatif dans la modélisation de divers phénomènes, relevant des sciences aussi bien biologiques et médicales que celles relatives à l'Economie, la Psychologie mathématique, etc., a, très tôt, suscité l'intérêt des statisticiens et motivé leur recherche de méthodes adaptées à des caractères non quantitatifs.

Ainsi, Fisher (1948) dans une étude sur l'analyse de variance, a exposé et traité un problème de statistique médicale, où il s'agissait d'expliquer le tableau des réactions liées à douze échantillons de sérums et de sang. Ayant pris en compte un critère ordinal et tiré des conclusions à l'aide d'un test F, il peut être considéré, en cela, comme un initiateur de l'analyse de variance monotone.

En suivant cette voie, plusieurs auteurs ont dirigé leurs investigations vers la pratique du codage, l'extension de méthodes classiques ou la recherche de nouvelles techniques trai-

tant le cas de variables ordinales, nominales ou scalées.

Divers programmes de régression qualitative ont ainsi été mis au point. Sans prétendre à l'exhaustivité, nous en présentons quelques uns :

-**MONANOVA** (Kruskal, 1965) traite l'analyse de variance monotone qui concerne l'explication d'un préordre par des variables nominales. Basée sur un algorithme de programmation non linéaire, la méthode consiste à minimiser un critère appelé **STRESS**, de formulation analogue à celle de l'Analyse des Similarités.

-**ADDALS** et **MORALS** (Young, de Leeuw et Takane, 1976) sont basés sur l'algorithme des «projections alternées au sens des moindres carrés» (ALS) et traitent des variables de type quelconque.

-**DISQUAL** (Saporta, 1977) traite l'analyse discriminante (variable à expliquer nominale), avec prédicteurs qualitatifs, par une méthode itérative.

L'Analyse des Mesures Conjointes représente dans le modèle additif, une branche de la régression qualitative. De très actives recherches y sont menées, notamment en marketing.

Relatif à ce dernier domaine, le **TRADE-OFF** auquel nous attachons un intérêt particulier, est une technique actuellement en plein essor.

Il concerne principalement le positionnement d'un nouveau produit ou la relance d'un produit «en perte de vitesse», mais la richesse de la méthode apporte aussi une meilleure connaissance de la Concurrence et du Marché, l'analyse pouvant être affinée par une segmentation (détermination de groupes d'acheteurs à partir de critères comme l'âge, le sexe, le revenu, etc..).

L'étude détaillée d'un exemple, effectuée à l'aide du programme **REGALS**, en sera proposée à titre d'illustration.

Le chapitre V est consacré à l'Analyse Multicritère. Le problème de la Désagrégation de la Préférence Globale peut être considéré comme une régression ou une analyse canonique sous contraintes.

Considérons un préordre  $g_0$  à expliquer par une famille  $g = (g_1, g_2, \dots, g_p)$  de critères ordinaux ou quantitatifs monotones. La distinction proposée entre les deux catégories de méthodes précitées est basée sur le fait que l'on recherche une fonction d'utilité  $U(g)$  agrégeant, sous forme d'une combinaison linéaire, les critères explicatifs à partir d'un codage  $U_0(g_0)$ , aussi «proche» que possible de  $g_0$ . Il y a deux choix possibles, chacun privilégiant la corrélation, respectivement des rangs et linéaire, de  $U(g)$  avec  $U_0(g_0)$ :

1-Fixer a priori  $U_0(g_0) = g_0$ : c'est le choix de la régression.

2-Laisser à la méthode le soin de déterminer le codage  $U_0(g_0)$ : c'est le choix de l'analyse canonique.

Cette distinction permet de définir les deux catégories dans chacune des deux approches, moindres écarts et moindres carrés.

Une description originale de UTA (version améliorée par Siskos & Yannacopoulos, 1983) nous permet de conclure que cette méthode est basée sur le choix n°2, impose au codage  $U_0(g_0)$  d'être de structure ordinale identique à celle de  $g_0$  à un seuil de préférence fixé et nous amène à la classer dans la catégorie «analyse canonique aux moindres écarts».

Ainsi, nous étendons à l'approche moindres écarts, le résultat établi par Tenenhaus (1979), concernant la base commune des méthodes de la régression qualitative dans l'approche moindres carrés.

Dans cette dernière approche, nous proposons la méthode REGALS (REGression par ALgorithme aIS). Elle est basée sur une théorie de l'analyse canonique de deux cônes polyédriques convexes, développée par Tenenhaus (1988).

Nous effectuons une étude comparative des méthodes UTA et REGALS à partir de trois exemples.

Notre conclusion est fondée sur le fait que la désagrégation de la préférence globale, selon chacune des approches moindres carrés et moindres écarts, ne peut être séparée de la question du mode d'agrégation duquel a résulté le préordre à expliquer :

Si dans le traitement du problème d'agrégation, l'homme d'étude fonde son appréciation à partir d'une performance moyenne de chaque action sur l'ensemble des critères, l'approche moindres carrés appréhenderait mieux le mécanisme de cette agrégation.

Par contre, si l'homme d'étude se fonde sur un point de vue «modal» et donne ainsi plus d'importance à l'«efficacité» sur un critère, l'approche moindres écarts reconstituerait le mieux sa démarche.

L'agrégation peut être considérée comme une phase d'un processus de décision. Les conclusions de l'étude établissent donc un lien entre chacune des approches moindres écarts et moindres carrés et un modèle particulier de cette phase.

Nous allons mettre en relation chacune de ces approches et un modèle particulier d'Aide à la Décision.

Deux écoles principales de l'Aide à la Décision se font face de part et d'autre de l'Atlantique. Les deux modèles suivants peuvent en être considérés respectivement représentatifs.

Le premier (Keeney, Nair, Raiffa, 1974, 1976) repose sur la théorie de l'Utilité Multi-Attributs. Il est postulé dans ce modèle, que la réalité du problème de décision, objet de l'étude, est structurée selon l'axiomatique de cette théorie.

Deux actions quelconques sont nécessairement en relation d'indifférence ou de préférence stricte. Le problème de l'homme d'étude est de déterminer cette relation pour toute paire d'actions.



Une part d'arbitraire est, dès le départ, ignorée puisque la démarche est fondée sur un postulat.

Dans le second (méthodes ELECTRE, Roy, 1977, 1978), la réalité du problème de décision, sujet de l'étude, prend corps progressivement, dans une démarche interactive avec le décideur, et vient combler le vide axiomatique.

Deux actions A et A' sont en relation de surclassement. L'affirmation "A est au moins aussi bonne que A'" et sa réciproque "A' est au moins aussi bonne que A" ne sont pas exclusives. L'acceptation de la réciproque implique l'indifférence, son refus implique soit une préférence faible soit une préférence stricte de A, soit une indifférence. Enfin, le refus des deux affirmations implique l'incomparabilité de ces actions. On voit là toute la richesse du système relationnel que l'homme d'étude doit essayer d'appréhender.

La part d'arbitraire dans la mise en oeuvre de la méthode, est prise en compte dans une analyse de robustesse. Cette analyse permet de faire ressortir le noyau des actions dont les relations sont décrites, par le modèle, de la façon la plus stable.

*«Les différences constatées entre les deux approches au niveau de l'obtention des données ne sont, en fait, que la conséquence d'une césure beaucoup plus profonde entre un modèle fondant sa légitimité sur une attitude "descriptive", visant à représenter une relation pré-existante et un modèle trouvant sa légitimité dans une attitude "constructive" cherchant, en collaboration avec le décideur, à bâtir une relation de préférence acceptable» (Roy & Bouyssou, 1983).*

Ainsi, chacun de ces modèles a sa "philosophie" propre, fondée sur une approche qui est dite «descriptive» pour le premier et «constructive» pour le second.

Un lien peut être précisé entre les approches descriptive et moindres carrés: elles ont en commun un plus grand degré de liberté dans leur représentation, a priori, de la réalité.

Celà se traduit, respectivement, par une axiomatique riche et une modélisation géométrique multidimensionnelle. Elles procèdent d'une attitude qui peut être dite risquée ou «optimiste».

De même, on peut dire des approches constructive et moindres écarts qu'elles compensent l'austérité de leur représentation (absence d'axiomatique et base de modélisation unidimensionnelle), par leur souplesse d'adaptation et une meilleure prise en compte de la complexité. On peut parler à leur égard, d'attitude de prudence ou «pessimiste».

La plupart des tenants de l'approche descriptive diraient avec B. de Finetti (1937):

*«Je juge plus utile un schéma simple comme repère dont j'admets de m'écarter pour des raisons accessoires plutôt qu'une construction complexe qui, afin de tout embrasser, ferait perdre la perspective de ce qui est essentiel».*

Nous reprenons à notre compte ces propos avec la réserve suivante: en nous plaçant sur un plan méthodologique, nous pensons qu'une construction complexe serait plus utile si, regroupant en un tout cohérent, les différentes approches, les techniques de recherche et les méthodes d'étude, elle se donne pour objectif d'apporter, à chaque fois, la réponse la plus appropriée au problème posé.

**Chapitre I**

**REGRESSION SOUS CONTRAINTES**

**LINEAIRES SUR LES COEFFICIENTS**

## I. REGRESSION SOUS CONTRAINTES LINEAIRES SUR LES COEFFICIENTS

Nous abordons, dans ce qui suit, le problème de la Régression Multiple sous Contraintes. Il consiste à expliquer une variable numérique, dite critère par un ensemble de variables numériques explicatives, dites prédicteurs, auquel est adjoit un système de contraintes linéaires sur les coefficients de régression.

Dans la section 1, nous présentons un ensemble de notations, conventions et résultats, à la base de nos développements ultérieurs.

La section 2 traite du problème des contraintes de positivité. L'intégration de l'opérateur SWEEP dans NNLS a été réalisée par P.Cazes, D.Bergougnan et C.Mullon (programme REGCAZ, 1984). Nous effectuons une analyse générale de ce problème et proposons une variante de REGCAZ nommée MCCP.

La section 3 aborde le problème des contraintes de bornes, selon l'approche ADOPT (voir Introduction). L'algorithme qui en découle, nommé MCCB, coïncide avec NNLS dans le cas des bornes 0 et  $+\infty$ .

La dernière section est consacrée au cas de contraintes d'inégalités linéaires quelconques sur les coefficients. Le problème revient à la détermination de la projection orthogonale sur l'intersection d'un espace vectoriel et d'un polyèdre convexe.

La méthode MCCIL (Moindres Carrés sous Contraintes d'Inégalités Linéaires) que nous développons, est basée sur l'algorithme LDP de Lawson & Hanson (1974). Une extension en est proposée; elle permet la détermination de la projection orthogonale sur l'intersection d'une variété linéaire et d'un polyèdre convexe.

## 1. PRELIMINAIRES

Après avoir fixé des notations, reprises tout au long du document, nous commençons par rappeler des résultats fondamentaux concernant les cônes polyédriques convexes.

Les conditions de Kuhn & Tucker, à la base des travaux de Lawson & Hanson (1974) concernant la régression sous contraintes, seront rappelées dans le but de montrer clairement notre apport dans une explication originale de l'algorithme LDP traitant de la recherche du vecteur de norme minimale dans un polyèdre.

Nous précisons ensuite la notion de qualité de la régression.

Nous terminons par la présentation de l'opérateur SWEEP de A.E. Beaton (1964). Une variante PIVOTAGE en est définie et ses applications décrites selon deux champs principaux.

### 1.1. NOTATIONS

Soit  $y \in \mathbb{R}^n$  le vecteur des valeurs observées, sur  $n$  individus, de la variable à expliquer. Soit  $X = \{X_1, X_2, \dots, X_p\}$  un ensemble de  $p$  variables explicatives;  $X$  désignera aussi la matrice  $n \times p$ , de rang  $p$ , des vecteurs colonnes  $X_i$  et  $X'$  la matrice transposée de  $X$ .

Le sous-espace vectoriel de  $\mathbb{R}^n$ , engendré par  $X$  sera noté

$$L(X) = \{X\beta, \beta' = (\beta_1, \dots, \beta_p)\};$$

Soit  $a = \sum \{a_i X_i, X_i \in X\}$ ,  $b = \sum \{b_i X_i, X_i \in X\}$  dans  $\mathbb{R}^n$  tels que

$$a_i \leq b_i, \quad i=1, 2, \dots, p;$$

on note  $C_{a,b} = \{z = \sum \alpha_i X_i, a_i \leq \alpha_i \leq b_i, i=1, \dots, p\}$ .

A tout  $z$  dans  $C_{a,b}$ , on peut associer une partition  $\{U, V, W\}$  de  $X$  telle que

$$U = \{X_i \mid \alpha_i = a_i\}; \quad V = \{X_i \mid a_i < \alpha_i < b_i\}; \quad W = \{X_i \mid \alpha_i = b_i\};$$

d'où  $z = \sum \{a_i X_i, X_i \in U\} + \sum \{b_i X_i, X_i \in W\} + \sum \{\alpha_i X_i, X_i \in V\}$ .

Remarquons que :

1.  $z$  appartient à la variété linéaire

$$\text{VAR}(U, V, W) = h + L(V)$$

avec 
$$h = \sum \{a_i X_i, X_i \in U\} + \sum \{b_i X_i, X_i \in W\}$$

et  $L(V)$  le sous-espace vectoriel engendré par  $V$ .

2.  $C_0, +\infty = C(X) = \{X\alpha, \alpha \geq 0\}$  est le cône polyédrique convexe engendré par  $X$ .

On notera :

$\langle x, y \rangle$  le produit scalaire de  $x$  et  $y$ ,

$\|x - y\|$  la distance euclidienne entre  $x$  et  $y$

et, par souci d'allègement de l'écriture, le vecteur  $y$  ayant été fixé une fois pour toute :

$P(U) = U(U'U)^{-1}U'y$  la projection orthogonale de  $y$  sur  $L(U)$

$P(U, V, W)$  la projection orthogonale de  $y$  sur  $\text{VAR}(U, V, W)$  .

On conviendra enfin que

$$U+i = U \cup \{X_i\};$$

$$U-i = U \setminus \{X_i\};$$

$$U^p = \{x \in \mathbb{R}^n, \langle x, u \rangle \leq 0, \text{ pour tout } u \in U\}$$

$$U^+ = \{x \in \mathbb{R}^n, \langle x, u \rangle = 0, \text{ pour tout } u \in U\}.$$

### 1.2. CONES POLYEDRIQUES CONVEXES

Soit  $C = C(X)$ ;  $C^p$ , le cône polaire de  $C$ , joue un rôle analogue à celui de l'espace orthogonal  $E^\perp$  d'un sous-espace vectoriel  $E$ :  $C \cap C^p = \{0\}$  et  $C + C^p = \mathbb{R}^n$ .

La proposition suivante donne une caractérisation de la projection orthogonale d'un vecteur sur un cône polyédrique convexe:

**Proposition I.1**

Soit  $z$  appartenant à  $C$ ;  $z$  est la projection orthogonale de  $y$  sur  $C$  si et seulement si  $y-z$  appartient à  $C^p$ .

Le résultat suivant constitue le fondement des algorithmes de projection orthogonale sur un cône polyédrique convexe:

**Proposition I.2**

Il existe une partie  $U$  de  $X$  telle que la projection de  $y$  sur  $C$  soit égale à celle, notée  $P(U)$ , de  $y$  sur l'espace vectoriel  $L(U)$ .

**Proposition I.3**

Soit  $C_1$  et  $C_2$  deux cônes polyédriques convexes :  $C_1$  est contenu dans  $C_2$  si et seulement si  $C_2^p$  est contenu dans  $C_1^p$ .

A la section 4.1, nous établirons un résultat (I.11) duquel, la proposition ci-après (énoncée dans:Faure, 1979, pp 225,226) est un cas-limite. Soit  $R$  une matrice  $m \times p$ :

**Proposition I.4**

$\{\beta \mid R\beta > 0\} \neq \emptyset$  si et seulement si, pour tout  $\alpha$ , vecteur de  $R^m$ ,  $\alpha > 0$  (composantes positives, non toutes nulles), on a :  $R'\alpha \neq 0$ .

**Preuve**

1. Supposons qu'il existe  $\beta \in R^p$  tel  $R\beta > 0$  et qu'il existe  $\alpha \in R^m, \alpha > 0$ ,  $R'\alpha = 0$  donc  $\beta'(R'\alpha) = 0 = (\beta'R')\alpha > 0$ : contradiction.

$$2. \{ \beta \mid R\beta > 0 \} = \emptyset \text{ implique } [C(R')]^p = \text{KER}(R)$$

implique  $C(R') = [\text{KER}(R)]^\perp$ ; on a donc l'alternative:

$$2.1. C(R') = \{0\} \text{ donc } R' = 0 \text{ et, pour tout } \alpha > 0, R'\alpha = 0.$$

2.2. Il existe  $\alpha > 0$  avec  $z = R'\alpha$  non nul; comme  $-z \in C(R')$ , on a  $-z = R'\mu$ ,  $\mu > 0$  et  $R'(\alpha + \mu) = 0$ .

L'interprétation de ce résultat est intéressante: étant donné un cône  $C$  engendré par un système  $G$  de  $p$  vecteurs de  $\mathbb{R}^n$  dont  $m$  sont linéairement indépendants, on a l'alternative suivante:

-Il existe une combinaison linéaire nulle à coefficients positifs des  $p$  vecteurs et le cône  $C$  emplit l'espace engendré par les  $m$  vecteurs linéairement indépendants et coïncide avec  $[\text{KER}(G')]^\perp$ .

-Il existe  $\beta$  tel que  $G'\beta > 0$  et le plan  $\{h, h'\beta = 0\}$ , de dimension  $m-1$ , est strictement latéral à  $C$ .

#### Lemme de Minkowski-Farkās

Soit  $\mu$  un vecteur de  $\mathbb{R}^p$ ; les deux propositions suivantes sont équivalentes:

-Pour tout  $\beta$  de  $\mathbb{R}^p$  tel que  $R\beta \geq 0$ , on a  $\beta'\mu \geq 0$

- $\mu$  appartient à  $C(R)$ .

#### Preuve

Nous proposons la démarche directe suivante:

$\mu \in C(R)$  équivalent à  $C(R:\mu)$  inclus dans  $C(R)$  équivalent à  $[C(R)]^p$  inclus dans  $[C(R:\mu)]^p$  équivalent à: pour tout  $\beta$  de  $\mathbb{R}^p$  tel que  $R\beta \geq 0$ ,  $\beta'\mu \geq 0$ .

CQFD



## 1.3. CONDITIONS DE KUHN &amp; TUCKER

Soit  $f$  et  $g_i, i=1, \dots, m$ , des fonctions dérivables de  $\mathbb{R}^p$  dans  $\mathbb{R}$ .

On considère le problème:

$$\text{MIN } f(\beta)$$

$$\text{SOUS LES CONTRAINTES } g_i(\beta) \leq 0, \quad i=1, \dots, m;$$

soit  $\text{grad}f(\beta) = (f'_{\beta_1}(x), f'_{\beta_2}(x), \dots, f'_{\beta_p}(x))$ , le gradient de  $f$ .

Kuhn & Tucker ont établi (1951) les résultats suivants:

## Résultat 1

La solution optimale  $\beta^*$  vérifie les conditions (dites de Kuhn & Tucker): il existe un vecteur  $u \in \mathbb{R}^m$  tel que

$$(1) \quad u_i \geq 0, \quad i=1, \dots, m;$$

$$(2) \quad \text{grad}f(\beta^*) + \sum \{u_i \text{grad}g_i(\beta^*), i=1, \dots, m\} = 0;$$

$$(3) \quad u_i g_i(\beta^*) = 0, \quad i=1, \dots, m;$$

## Résultat 2

Si la fonction  $f$  est convexe les conditions de Kuhn & Tucker sont suffisantes pour assurer l'optimalité d'un vecteur  $\beta \in \mathbb{R}^p$ .

Dans le cas du problème non linéaire (problème MCCIL, section 4.2):

$$\text{MIN } \|X\beta - y\|^2$$

$$\text{SOUS LES CONTRAINTES LINEAIRES } R\beta \geq r$$

avec  $R$  matrice  $m \times p$  et  $r \in \mathbb{R}^m$ , problème que l'on peut écrire

$$\text{MIN } (\frac{1}{2} \|X\beta - y\|^2) \text{ SOUS LES CONTRAINTES LINEAIRES } G\beta + r \leq 0$$

avec  $G = -R$  et  $G^i = -R^i$  la  $i$ ème ligne de  $G$ , la fonction  $f$  est convexe:

$$f(\beta) = \frac{1}{2}(X\beta - y)'(X\beta - y) = \frac{1}{2}(\beta'X'X\beta + y'y - \beta'X'y - y'X\beta)$$

$$\text{grad}f(\beta) = X'X\beta - X'y$$

$$\text{grad}g_i(\beta) = G^i$$

$$\text{grad}f(\beta) + \sum\{u_i \text{grad}g_i(\beta), i=1, \dots, m\} = X'X\beta - X'y + G'u$$

les conditions de Kuhn & Tucker sont nécessaires et suffisantes:

$$(1) u_i \geq 0, i=1, \dots, m;$$

$$(2) X'(X\beta - y) = -G'u = R'u$$

$$(3) u_i (R^i\beta - r) = -u_i (G^i\beta + r) = 0, i=1, \dots, m;$$

il existe un sous-ensemble  $I$  de  $\{1, \dots, m\}$  telle que

$$u_i = 0, R^i\beta - r > 0 \text{ si } i \in I$$

$$u_i \geq 0, R^i\beta - r = 0 \text{ sinon.}$$

#### 1.4. QUALITE DE LA REGRESSION

Notons  $y^L, y^C, y^P$  les projections de  $y$  sur  $L, C, P$ , respectivement l'espace, le cône et le polyèdre convexe engendrés par  $X$ ;  $P$  découle de contraintes de bornes ou d'inégalités linéaires quelconques.

Une mesure de la qualité de la régression multiple, est le coefficient de détermination

$$r^2 = |y^L|^2 / |y|^2$$

posons  $s^2 = |y - y^L|^2 / |y|^2$  (rapport résiduel). On a la relation

$$r^2 + s^2 = 1.$$

On définit les mêmes mesures dans le cas de la régression sous contraintes.

a)  $r^2 = \|y^c\|^2 / \|y\|^2$  dans le cas de contraintes de positivité;

on a toujours  $r^2 + s^2 = 1$  car  $y^c$  est la projection de  $y$  sur un sous-espace de  $L(X)$ .

b)  $r^2 = \|y^p\|^2 / \|y\|^2$  dans le cas de contraintes linéaires quelconques et dans ce cas:

$$\|y\|^2 = \|y - y^p\|^2 + \|y^p\|^2 + 2\langle y - y^p, y^p \rangle$$

d'où  $r^2 + s^2 = 1 - (2\langle y - y^p, y^p \rangle / \|y\|^2)$ .

#### Remarque

Si  $P$  contient l'origine, comme  $\langle y - y^p, 0 - y^p \rangle \leq 0$ , on a

$$r^2 + s^2 \leq 1.$$

### 1.5. OPERATEUR SWEEP

C'est un opérateur matriciel d'usage classique en régression multiple itérative (programmes BMDP et SAS).

#### 1.5.1. DESCRIPTION

Etant donné une matrice carrée  $A = \{a_{ij}\}$  telle que le  $k$ ème élément  $a_{kk}$  soit non nul, le résultat d'un "sweeping" de  $A$  par rapport à  $a_{kk}$  est une matrice  $A^{s_k}$  définie comme suit:

$$a^{s_{kk}} = -1/a_{kk} \quad ; \quad a^{s_{ik}} = a_{ik}/a_{kk} \quad ;$$

$$a^{s_{kj}} = a_{kj}/a_{kk} \quad ; \quad a^{s_{ij}} = a_{ij} - (a_{ik}a_{kj}/a_{kk}) \quad .$$

On peut, de même, définir l'opération "sweeping inverse": appliquée à  $A^{s_k}$  par rapport à  $a^{s_{kk}}$ , elle permet de retrouver  $A$  :

$$a_{kk} = -1/a^{s}_{kk} \quad ; \quad a_{ik} = -a^{s}_{ik}/a^{s}_{kk}$$

et

$$a_{kj} = -a^{s}_{kj}/a^{s}_{kk} \quad ; \quad a_{ij} = a^{s}_{ij} - (a^{s}_{ik}a^{s}_{kj}/a^{s}_{kk})$$

pour  $i$  et  $j$  différents de  $k$ .

L'intérêt de l'opérateur SWEEP repose sur les théorèmes suivants.

#### Théorème I.1

S'il est possible de "sweeper" la matrice  $A$  par rapport à tous les éléments de la diagonale, alors  $A$  est inversible et la matrice résultante  $A^s$  vérifie  $A^s = -A^{-1}$ .

#### Théorème I.2

S'il est possible de "sweeper" la matrice partitionnée

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

par rapport à chaque élément de la diagonale de  $A_{11}$ , alors  $A_{11}$

est inversible et

$$A^s / \text{DIAG}(A_{11}) = \begin{bmatrix} A^{s}_{11} & -A^{s}_{11} A_{12} \\ -A_{21} A^{s}_{11} & A_{22} + A_{21} A^{s}_{11} A_{12} \end{bmatrix}$$

#### 1.5.2. LE PIVOTAGE

Il est possible de définir une variante de

l'opérateur SWEEP possédant des propriétés intéressantes :

$$a^{p_{kk}} = 1/a_{kk} \quad ; \quad a^{p_{1k}} = -a_{1k}/a_{kk} \quad ;$$

$$a^{p_{kj}} = a_{kj}/a_{kk} \quad ; \quad a^{p_{1j}} = a_{1j} - (a_{1k}a_{kj}/a_{kk}).$$

Dans ce cas, "sweeping" et "sweeping inverse" sont identiques: on dira "pivotage" par rapport à  $a_{kk}$ .

On a des théorèmes analogues à ceux précédemment cités:

#### Théorème I.1b

S'il est possible de faire pivoter la matrice A par rapport à tous les éléments de la diagonale, alors A est inversible et la matrice résultante  $A^p$  vérifie  $A^p = A^{-1}$ .

#### Théorème I.2b

S'il est possible de faire pivoter la matrice partitionnée

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

par rapport à chaque élément de la diagonale de  $A_{11}$ , alors  $A_{11}$  est inversible et

$$A^p / \text{DIAG}(A_{11}) = \begin{bmatrix} A^{p_{11}} & A^{p_{11}} A_{12} \\ -A_{21} A^{p_{11}} & A_{22} - A_{21} A^{p_{11}} A_{12} \end{bmatrix}$$

**Remarque**

Si  $A$  est symétrique,  $A^s / \text{DIAG}(A_{11})$  est symétrique mais  $A^p / \text{DIAG}(A_{11})$  ne l'est pas.

**1.5.3.APPLICATIONS**

Deux champs principaux d'application de l'opérateur SWEEP sont décrits. Le premier correspond au problème de l'inversion d'une matrice. Le deuxième est relatif à la régression multiple itérative.

**1.5.3.1.INVERSION MATRICIELLE**

Nous présentons quatre procédures correspondant à trois situations d'inversion matricielle: les deux premières correspondent au problème classique de la détermination d'une matrice inverse. Les deux dernières s'appliquent au cas où la matrice à inverser est déduite d'une matrice dont l'inverse est connue, par respectivement l'ajout d'une ligne et d'une colonne et la suppression de lignes et de colonnes de même rang.

**Procédure  $P_1$ -Inversion** d'une matrice donnée par une suite de "sweeping" sur les éléments diagonaux, tous non nuls.

**Procédure  $P_2$ -Inversion** d'une matrice donnée en "construisant" progressivement la matrice à inverser par l'ajout d'une ligne et d'une colonne à chaque étape. C'est le processus récurrent:

$$A^1 = \begin{bmatrix} a_{11} \end{bmatrix} ; A^{1s} = \begin{bmatrix} -1/a_{11} \end{bmatrix}$$

$$A^k = \begin{bmatrix} A^{k-1} & A^{k, 12} \\ A^{k, 21} & a_{kk} \end{bmatrix}$$

avec  $A^{k,12} = (a_{1k}, a_{2k}, \dots, a_{k-1k})'$ ;  $A^{k,21} = (a_{k1}, a_{k2}, \dots, a_{kk-1})$ .

On a alors d'après le théorème 2:

$$A^{ks} / \text{DIAG}(A^{k-1}) = \begin{bmatrix} A^{(k-1)s} & -A^{(k-1)s} A^{k,12} \\ -A^{k,21} A^{s k-1} & a_{kk} + A^{k,21} A^{(k-1)s} A^{k,12} \end{bmatrix}$$

et, à l'étape courante  $k$ , on détermine  $A^{ks}$  en "sweepant"  $A^{ks} / \text{DIAG}(A^{k-1})$  par rapport à  $a_{kk} + A^{k,21} A^{(k-1)s} A^{k,12}$ , supposé non nul pour tout  $k$ ; cette condition est toujours vérifiée si  $A$  est symétrique définie positive.

**Procédure  $P_3$** -Etant donné une matrice  $A(n,n)$  inversible dont l'inverse  $A^{-1}$  est connue, on se pose le problème de la détermination de l'inverse de la matrice  $B$  suivante:

$$B = \begin{bmatrix} A & v \\ u' & b \end{bmatrix}$$

avec  $u, v$  dans  $R^n$  et  $b$  un réel.

C'est l'application d'une étape du processus  $P_2$  :

$$B^s / \text{DIAG}(A) = \begin{bmatrix} -A^{-1} & A^{-1} v \\ u' A^{-1} & b - u' A^{-1} v \end{bmatrix}$$

on effectue un "sweeping" de  $B^s$ /DIAG(A) par rapport à  $b-u'A^{-1}v$  et l'on a

$$B^{-1} = -B^s .$$

Procédure P<sub>4</sub>-Etant donné une matrice  $A(n,n)$  inversible dont l'inverse  $A^{-1}$  est connue, on se pose le problème de la détermination de l'inverse de la matrice  $B$  obtenue en supprimant de  $A$  un nombre  $r$  de lignes et de colonnes de même rang :  $k_1, k_2, \dots, k_r$ .

On effectue, dans un ordre arbitraire, une suite de "sweepings inverses" par rapport aux éléments diagonaux correspondants de  $A^s = -A^{-1}$ ; on obtient  $B^s$  en supprimant de la matrice résultante les lignes et les colonnes de rang  $k_i, i = 1, \dots, r$ .

#### 1.5.3.2. REGRESSION MULTIPLE

Soit  $m < p$  et  $A = [1 \ X \ Y]$  la matrice du vecteur unité, des critères  $X_j, j=1, \dots, p$  et du prédicteur  $Y$ .

On s'intéresse au modèle :

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_m X_m + \varepsilon$$

et au rôle marginal des variables  $X_{m+1}, \dots, X_p$ .

Ecrivons  $A = [X^1 \ X^2]$  avec  $X^1 = [1 \ X_1 \ \dots \ X_m]$  et  $X^2 = [X_{m+1} \ \dots \ X_p \ Y]$  et posons

$$B = A'A = \begin{bmatrix} X^1 ' X^1 & X^1 ' X^2 \\ X^2 ' X^1 & X^2 ' X^2 \end{bmatrix}$$

En "sweepant" par rapport à la diagonale de  $X^1 ' X^1$ , on obtient



$$B^s = \begin{bmatrix} -(X^1 ' X^1)^{-1} & (X^1 ' X^1)^{-1} X^1 ' X^2 \\ X^2 ' X^1 (X^1 ' X^1)^{-1} & X^2 ' X^2 - X^2 ' X^1 (X^1 ' X^1)^{-1} X^1 ' X^2 \end{bmatrix}$$

Notons  $y$  l'indice de la dernière colonne  $Y$  de  $X$ ,  $Y^* = X^1 \beta^*$  la projection de  $Y$  sur  $L(X^1)$ .

On a les statistiques classiques :

$$\Sigma e_i^2 = \|Y - Y^*\|^2 = Y'(Y - X^1 (X^1 ' X^1)^{-1} X^1 ' Y) = b^s_{yy}$$

$$\sigma^{*2} = (\Sigma e_i^2) / (n - m - 1) = b^s_{yy} / (n - m - 1).$$

Pour  $j = 1, \dots, m$  :

$$\beta^*_j = ((X^1 ' X^1)^{-1} X^1 ' Y)_j = b^s_{jy}$$

$$s_j = \sigma^* \sqrt{-b^s_{jj}}$$

$$t_j = \beta^*_j / s_j = b^s_{jy} / \sigma^* \sqrt{-b^s_{jj}}$$

et pour  $j = m+1, \dots, p$  :

$$\langle X_j, Y - Y^* \rangle = b^s_{jy}$$

$$\begin{aligned} (\Sigma e_i^2)_{1, \dots, p} - (\Sigma e_i^2)_{1, \dots, p, j} &= b^s_{yy} - (b^s_{yy})^s \text{ [sur } b^s_{jj}] \\ &= b^s_{yy} - (b^s_{yy} - (b^s_{jy} b^s_{yj} / b^s_{jj})) \\ &= b^s_{jy} b^s_{yj} / b^s_{jj}. \end{aligned}$$

Notons enfin, que:

$$b^p_{yj} = -b^s_{yj}, \quad b^p_{jy} = b^s_{jy}, \quad b^p_{jj} = -b^s_{jj}, \quad b^p_{yy} = b^s_{yy}$$

$$\text{d'où } (\Sigma e_i^2)_{1, \dots, p} - (\Sigma e_i^2)_{1, \dots, p, j} = b^p_{jy} b^p_{yj} / b^p_{jj}.$$

## 2.CONTRAINTES DE POSITIVITE

Nous décrivons dans cette section l'algorithme NNLS (Non-Negative Least Squares), proposé et justifié à l'aide des conditions de Kuhn & Tucker (section 1.3) par Lawson et Hanson (1974).

Une amélioration notable de cette algorithme a été apportée par l'intégration de l'opérateur SWEEP, réalisée par P.Cazes, D.Bergougnan et C.Mullon (programme REGCAZ, 1984).

Nous faisons une analyse générale de ce problème et proposons une variante de REGCAZ, nommée MCCP, fondée sur la prise en compte de coefficients libres et sur un choix original de la variable à entrer, à chaque étape, dans la régression.

Une étude numérique est effectuée. Elle permet de conclure qu'une nette amélioration est apportée par ce choix.

### 2.1.LE PROBLEME NNLS

Il consiste en la détermination de la projection au sens des moindres carrés du vecteur  $y$  sur le cône polyédrique convexe  $C(X)$ .

Ce problème revient (proposition I.2, section 1.2) à déterminer une partie optimale  $U$  de  $X$ . L'algorithme NNLS est de type ADOPT.

#### 2.1.1.CONDITIONS ADOPT

-E = {U, partie de X};

-A = {U  $\in$  E telle que P(U)  $\in$  C(X)};

-Critère : distance  $\|y - P(U)\|$  à minimiser pour  $U \in A$ ;

-Test : U optimale si et seulement si

$$y - P(U) \in (X \setminus U)^{\circ}.$$

-Procédure: Partant d'une partie admissible, U non optimale, on pose  $z^0 = P(U)$  ; il existe  $X_r \in X \setminus U$  tel que  $\langle y - z^0, X_r \rangle > 0$ .

Soit  $P=U+r$ ; si  $P$  est admissible, elle est «meilleure» que  $U$  et la procédure est terminée car  $L(U)$  est inclus dans  $L(P)$ . Sinon, décrivons le procédé suivant:

Posons  $y^1 = P(P)$  ; on considère un élément  $z^1$  du segment  $[y^1 \ z^0]$  appartenant à  $C(X)$ ;

on a  $\|y-y^1\| < \|y-z^1\| < \|y-z^0\|$  car  $y^1-z^0 \in L(P)$ .

Soit  $U^1$  la partie de  $X$  associée à  $z^1$  ( $L(U^1)$  le plus petit espace vectoriel contenant  $z^1$ ); si  $U^1$  est admissible la procédure est terminée; sinon on remplace  $z^0$  par  $z^1$ ,  $y^1$  par  $y^2=P(U^1)$  et on réitère le procédé.

On construit ainsi une suite  $\{z^j\}$  dans  $C(X)$ , telle que la distance  $\|y-z^j\|$  soit de plus en plus petite;  $CARD(E)$  étant fini,  $\{z^j\}$  converge dans  $C(X)$  en un nombre fini d'étapes.

#### 2.1.2.DESCRPTION

Les différentes étapes de NNLS sont:

1.Initialisation:

$U=\emptyset$ ;  $\alpha_i=0, i=1, \dots, p$ ;  $y^*=\sum \alpha_i X_i$ ;

2.si  $\text{MAX}\{\langle y-y^*, X_i \rangle, X_i \in X \setminus U\} \leq 0$  alors aller à 5

sinon, soit  $X_r$  réalisant le maximum et  $P=U+r$ ;

3.Soit  $y_1=\sum \alpha_i X_i = P(P)$ ; si  $y_1$  est admissible alors  $y^* = y_1$ ,

$\alpha_i = \alpha_{i1}, i=1, \dots, p, U = P$ , aller à 2 sinon

4.Déterminer

$\mu = \text{MAX}\{ \alpha_{i1} / (\alpha_{i1} - \alpha_i), \alpha_i < 0 \}$ ;

$I = \{X_i \in P \text{ tel que } \alpha_{i1} + \mu(\alpha_i - \alpha_{i1}) = 0\}$

Soit  $P=P \setminus I$ ,  $\alpha_i = \alpha_{i1} + \mu(\alpha_i - \alpha_{i1}), i = 1, \dots, p$ ; aller à 3;

5. $y^*$  est la projection de  $y$  sur  $C(X)$ .

### 2.1.3. INTEGRATION DE L'OPERATEUR SWEEP

Deux types d'intégration de l'opérateur SWEEP sont décrits. Ils correspondent respectivement, aux deux champs d'application précités (sections 1.5.3.1 et 1.5.3.2).

#### 2.1.3.1. INTEGRATION DE TYPE I

Aux étapes 2 et 4 de l'algorithme ci-dessus on rajoute respectivement, les procédures P3 et P4 de la section 1.5.3.1: c'est l'intégration proposée et décrite en détail dans le Cahier du Lamsade n°87 (1988).

#### 2.1.3.2. INTEGRATION DE TYPE II

Elle a été effectuée par Cazes, Bergougnan et Mullon (1984), sur des variables normées, à l'aide de la variante PIVOTAGE. Le résultat en est le programme REGCAZ dont nous faisons ci-dessous la description:

1. Initialisation:

$$B = [X \ Y]' [X \ Y];$$

2. Calculer  $v_i = b_{iy} b_{yi} / b_{ii}$ ,  $i = 1, \dots, p$ ; poser  $U = \{X_i \mid v_i < 0\}$

et, pour tout  $i$ :  $\alpha_i = b_{iy}$  si  $X_i \in U$  et  $\alpha_i = 0$  sinon.

Si  $U = \emptyset$  ou si  $\alpha_i \geq 0$  pour tout  $i$  aller à 4 sinon

3. Déterminer

$$\mu = \text{MAX}\{\alpha_i / (\alpha_i - a_i) \mid \alpha_i < 0\};$$

$$I = \{b_{ii} \text{ tel que } X_i \in U \text{ et } \alpha_i + \mu(\alpha_i - a_i) = 0\}$$

Soit  $a_i = \alpha_i + \mu(\alpha_i - a_i)$ ,  $i = 1, \dots, p$ ; aller à 6;

4. Poser  $a_i = a_i$ ,  $i = 1, \dots, p$ . Si  $U = X$ , aller à 7 sinon

5. Si  $\text{MAX}\{b_{1y} \mid X_1 \in X \setminus U\} \leq 0$  alors aller à 7 sinon soit  $b_{ry}$  réalisant le maximum et  $I = \{b_{rr}\}$ .

6. Faire le pivotage de la matrice B par rapport à I; aller à 2.

7.  $y^* = \sum \alpha_i X_i$  est la projection de y sur C(X).

### 2.1.3.3. VERSION MCCP

Sans imposer aux variables d'être normées, nous proposons une variante de REGCAZ, que nous nommons MCCP (Moindres Carrés sous Contraintes de Positivité) permettant :

(1) Une prise en compte de coefficients libres: soit  $\Phi = \{X_i \text{ tel que } \alpha_i \text{ libre}\}$ .

(2) Un choix optimal de la variable à entrer: nous entrons la variable  $X_r$  produisant la plus grande réduction de somme des carrés résiduelle (section 1.5.3.2).

A notre sens, ceci donne une version définitivement optimisée de l'algorithme NNLS:

1. Initialisation:

$B = [X \ Y]' [X \ Y]$ ;  $I = \{b_{11} \text{ tel que } X_1 \in \Phi\}$ ; aller à 6.

2. Calculer  $v_i = b_{1y} b_{y1} / b_{11}$ ,  $i = 1, \dots, p$ ; poser  $U = \{X_i \mid v_i < 0\}$

et, pour tout  $i$ :  $\alpha_i = b_{1y}$  si  $X_i \in U$  et  $\alpha_i = 0$  sinon.

Si  $U = \Phi$  ou si  $\alpha_i \geq 0$  pour tout  $i$  tel que  $X_i \in X \setminus \Phi$ , aller à 4 sinon

3. Déterminer  $\mu = \text{MAX}\{\alpha_1 / (\alpha_1 - \alpha_i), X_i \in U \setminus \Phi, \alpha_i < 0\}$ ;

$I = \{b_{11} \text{ tel que } X_1 \in U \setminus \Phi \text{ et } \alpha_1 + \mu(\alpha_1 - \alpha_i) = 0\}$ .

Soit  $\alpha_i = \alpha_1 + \mu(\alpha_1 - \alpha_i)$ ,  $i = 1, \dots, p$ ; aller à 6;

4. Poser  $a_i = a_{1y}$ ,  $i=1, \dots, p$ . Si  $U=X$ , aller à 7 sinon

5. Si  $\text{MAX}\{b_{1y} \mid X_i \in X \setminus U\} \leq 0$  alors aller à 7 sinon

soit  $v_r$  réalisant le maximum de  $\{v_i \mid X_i \in X \setminus U, b_{1y} > 0\}$  et  $I = \{b_{r-r}\}$ .

6. Faire le pivotage de la matrice B par rapport à I.

Aller à 2.

7.  $y^* = \sum a_i X_i$  est la projection de y sur  $L(\Phi) + C(X \setminus \Phi)$ .

#### Remarques

1. Dans le cas où toutes les variables sont centrées-réduites, le critère d'entrée de REGCAZ signifie l'introduction de la variable dont la corrélation partielle avec le critère, conditionnellement aux variables dans la régression, est positive et la plus élevée.

2. Tenenhaus (1983) a proposé l'amélioration (1) dans NNLS et a suggéré d'incorporer dans U, la variable  $X_r$  qui réalise le minimum de  $\{\|y - P(U+i)\|, X_i \in X \setminus U, \langle y - P(U+i), X_i \rangle > 0\}$ .

#### 2.2. ETUDE COMPARATIVE REGCAZ/MCCP

Elle a été effectuée sur un exemple traité par Kennedy & Gentle (1980, pp. 354-357), en illustration de l'algorithme "branch-and-bound" de R.D. Armstrong & E.L. Frome (1976), avec le modèle

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

sous la condition  $\beta_i \geq 0$ ,  $i = 0, 1, 2, 3$  et les variables non normées. Nous présentons ci-après les données et les résultats obtenus, respectivement par le programme REGCAZ et la version MCCP.

**Données de Kennedy & Gentle**

Y	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>
378.4	0.4	94	39
310.0	6.7	77	85
99.0	9.3	8	44
206.0	5.6	41	43
179.0	8.5	33	20
285.2	10.0	71	60
351.9	6.3	88	0
256.7	0.6	54	37
266.9	8.8	63	64
282.6	7.1	66	26
270.3	7.5	63	39
284.2	2.7	65	50
129.3	8.9	17	35
176.0	8.3	31	1
196.4	8.9	39	16
286.1	1.9	66	78
311.6	8.9	77	15
352.9	5.1	89	35
266.1	2.8	57	6
152.9	3.5	22	53

## 1. Programme REGCAZ

Initialisation :  $\alpha$  est le vecteur nul.

$$y'X = [5041.5 \quad 28953.3 \quad 318760.1 \quad 190577.9]$$

Pivotage n°1 : on entre la variable  $X_2$

$$\alpha = [0 \quad 0 \quad 4.277 \quad 0]$$

$$(y - X\alpha)'X = [246.9895 \quad 1740.0691 \quad 0 \quad 7201.7741]$$

Pivotage n°2 : on entre la variable  $X_3$

$$\alpha = [0 \quad 0 \quad 3.9787 \quad 0.5185]$$

$$(y - X\alpha)'X = [194.5603 \quad 1394.227 \quad 0 \quad 0]$$

Pivotage n°3 : on entre la variable  $X_1$

$$\alpha = [0 \quad 4.0111 \quad 3.7471 \quad 0.3259]$$

$$(y - X\alpha)'X = [199.3558 \quad 0 \quad 0 \quad 0]$$

Pivotage n°4 : on entre la variable  $X_0$

$$\alpha = [98.9971 \quad -2.01389 \quad 3.022 \quad -0.1043]$$

Pivotage n°5 : on retire la variable  $X_1$

$$\alpha = [80.9723 \quad 0 \quad 3.0999 \quad -0.0710]$$

Pivotage n°6 : on retire la variable  $X_3$

$$\alpha = [78.6867 \quad 0 \quad 3.0935 \quad 0]$$

$$(y - X\alpha)'X = [0 \quad -313.4827 \quad 0 \quad -754.3745]$$

l'optimum  $\alpha^* = \alpha$  est atteint.



## 2.Version M CCP

Rappelons que  $v$  désigne le vecteur des «écarts de sommes de carrés résiduelles»:

$$v_i = b_{1y} b_{y1} / b_{11}, \quad i = 0, 1, 2, 3.$$

Initialisation :  $\alpha$  est le vecteur nul

$$y'X = [5041.5 \quad 28953.3 \quad 318760.1 \quad 190577.9]$$

$$v' = [1270836.1125 \quad 908404.2185 \quad 1363335.0958 \quad 942053.6382]$$

Pivotage n°1 : on entre la variable  $X_2$

$$\alpha = [0 \quad 0 \quad 4.277 \quad 0]$$

$$(y - X\alpha)'X = [246.9895 \quad 1740.0691 \quad 0 \quad 7201.7741]$$

$$v' = [19434.7828 \quad 7975.9201 \quad -1363335.0958 \quad 3734.3180]$$

Pivotage n°2 : on entre la variable  $X_0$

$$\alpha = [78.6867 \quad 0 \quad 3.0935 \quad 0]$$

$$(y - X\alpha)'X = [0 \quad -313.4827 \quad 0 \quad -754.3745]$$

$$v' = [-19434.7828 \quad 604.2385 \quad -111933.7661 \quad 53.5261]$$

l'optimum  $\alpha^* = \alpha$  est atteint (!).

### Remarque

Avec les variables normées, M CCP converge en deux pivotages et REGCAZ présente une amélioration en convergeant en quatre pivotages. L'équation de régression est:

$$Y = 0.29918 + 0.71799X_2.$$

Avec les variables centrées-réduites, M CCP et REGCAZ convergent en deux pivotages.

### 3.CONTRAINTES DE BORNES

L'algorithme MCCB (Moindres Carrés sous Contraintes de Bornes) découle d'une approche ADOPT (voir Introduction) dont la condition nécessaire est le résultat fondamental suivant (démontré à la section 3.4) : «La projection orthogonale d'un vecteur sur un polyèdre convexe, résultant des contraintes de bornes, coïncide avec celle de ce vecteur, sur une variété linéaire optimale ou "face" particulière du polyèdre.»

On peut considérer l'algorithme MCCB comme une extension de NNLS, au cas de bornes quelconques, autres que 0 et  $+\infty$ , imposées aux coefficients de la régression. Nous terminons la section par une étude numérique.

#### 3.1.LE PROBLEME MCCB

Il consiste en la détermination de la projection au sens des moindres carrés du vecteur  $y$  sur le polyèdre  $C_{a,b}$ .

$$C_{a,b} = \{z = \sum \alpha_i X_i, a_i \leq \alpha_i \leq b_i, i=1, \dots, p\}$$

A tout vecteur du polyèdre  $C_{a,b}$ , on peut associer une partition de  $X$  et définir une variété linéaire le contenant (section 1.1). On démontre (proposition I.6, section 3.4) qu'il existe une partition  $\{U,V,W\}$  telle que la projection recherchée soit égale à celle, notée  $P(U,V,W)$ , de  $y$  sur la variété linéaire  $VAR(U,V,W)$ . Ce problème revient donc à déterminer la partition optimale  $\{U,V,W\}$ .

#### 3.2.L'ALGORITHME PROPOSE

Nous allons en donner les conditions ADOPT et la description.

##### 3.2.1.CONDITIONS ADOPT

-E =  $\{\{U,V,W\}$  partition de  $X\}$ ;

-A =  $\{\{U,V,W\} \in E$  telle que  $P(U,V,W) \in C_{a,b}\}$ ;

-Critère: distance  $\|y - P(U,V,W)\|$  à minimiser pour  $\{U,V,W\} \in A$ ;

-Test :  $\{U, V, W\}$  optimale si et seulement si

$$y - P(U, V, W) \in U^p \cap V^1 \cap (-W)^p.$$

-Procédure: Partant d'une partition admissible,  $\{U, V, W\}$  non optimale, on pose  $z^0 = P(U, V, W)$ ; il existe  $X_r \in X \setminus V$  tel que

$$\langle y - z^0, cX_r \rangle > 0$$

avec  $c = 1$  si  $X_r \in U$  ou  $c = -1$  si  $X_r \in W$ .

Soit  $P = U - r, Q = V + r, R = W - r$ . Si  $\{P, Q, R\}$  admissible, elle est «meilleure» que  $\{U, V, W\}$  et la procédure est terminée car  $\text{VAR}(U, V, W)$  est incluse dans  $\text{VAR}(P, Q, R)$ .

Sinon, décrivons le procédé suivant: posons  $y^1 = P(P, Q, R)$ ; on considère un élément  $z^1$  du segment  $[y^1, z^0]$  appartenant à  $C_{a, b}$ ; on a:

$$\|y - y^1\| < \|y - z^1\| < \|y - z^0\|$$

car  $y^1 - z^0 \in L(Q)$ .

Soit  $\{P, Q, R\}$  la partition associée à  $z^1$  ( $z^1 \in \text{VAR}(P, Q, R)$  et  $L(Q)$  le plus petit espace vectoriel possible); si  $\{P, Q, R\}$  est admissible la procédure est terminée; sinon on remplace  $z^0$  par  $z^1, y^1$  par  $y^2 = P(P, Q, R)$  et on réitère le procédé.

On construit ainsi une suite  $\{z^j\}$  dans  $C_{a, b}$ , telle que la distance  $\|y - z^j\|$  soit de plus en plus petite;  $\text{CARD}(E)$  étant fini,  $\{z^j\}$  converge dans  $C_{a, b}$  en un nombre fini d'étapes.

### 3.2.2. DESCRIPTION

Nous décrivons, ci-dessous, les étapes de l'algorithme MCCB:

#### 1. Initialisation

$V = \emptyset$ ; si  $\|y - a\| < \|y - b\|$  alors  $U = X, W = \emptyset; y^* = a; \alpha_i = a_i, i = 1, \dots, p$ ;

sinon  $U = \emptyset, W = X, y^* = b; \alpha_i = b_i, i = 1, \dots, p$ ;

2. si  $\text{MAX}(\{\langle y-y^*, X_i \rangle, X_i \in U \} \cup \{\langle y-y^*, -X_i \rangle, X_i \in W \}) \leq 0$  alors aller à 5

sinon soit  $X_r$  réalisant le maximum;  $P=U-r$ ;  $Q=V+r$ ;  $R=W-r$ .

3. Soit  $y_1=P(P,Q,R)$ ;

$y_1 - \Sigma\{a_i X_i, X_i \in P\} - \Sigma\{b_i X_i, X_i \in R\} = \Sigma\{a_i X_i, X_i \in Q\}$  ; si  $y_1$  est admissible alors  $y^*=y_1$ ;  $a_i = a_{i1}$ ,  $i=1, \dots, p$ ;  $U=P, V=Q, W=R$ ; aller à 2 sinon

4. Déterminer

$\mu = \text{MAX}(\{(a_i - a_{i1}) / (a_i - a_{i1}), a_{i1} < a_i\} \cup \{(b_i - a_{i1}) / (a_i - a_{i1}), a_{i1} > b_i\})$ ;

$I = \{X_i \in X \setminus P \text{ tel que } a_{i1} + \mu(a_i - a_{i1}) = a_i\}$

$J = \{X_i \in X \setminus R \text{ tel que } a_{i1} + \mu(a_i - a_{i1}) = b_i\}$ .

Soit  $P=P+I, Q=Q-(I+J), R=R+J$ ;  $a_i = a_{i1} + \mu(a_i - a_{i1}), i=1..p$ ; aller à 3

5.  $y^*$  est la projection de  $y$  sur  $C_{a,b}$ .

#### Remarques

1. L'algorithme MCCB permet de traiter les cas de coefficients fixes (bornes supérieure et inférieure égales) et libres.

2. Son implémentation informatique a été réalisée avec l'intégration de l'opérateur SWEEP du type I (section 2.1.3.1).

#### 3.3. ETUDE NUMERIQUE

Nous reprenons le problème de Walpole & Myers (1972) étudié par R.D. Armstrong et E.L. Frome (1976) en application de l'algorithme branch-and-bound de projection sur un cône.

Trois substances sont utilisées à différentes concentrations dans le but d'augmenter le taux de survie d'un certain type de semence animale.

## Données de Walpole &amp; Myers

%SURVIE	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>
25.50	1.74	5.30	10.80
31.20	6.32	5.42	9.40
25.90	6.22	8.41	7.20
38.40	10.52	4.63	8.50
18.40	1.19	11.60	9.40
26.70	1.22	5.85	9.90
26.40	4.10	6.62	8.00
25.90	6.32	8.72	9.10
32.00	4.08	4.42	8.73
25.20	4.15	7.60	9.20
39.70	10.15	4.83	9.40
35.70	1.72	3.12	7.60
26.50	1.70	5.30	8.20

La variable à expliquer Y est le taux de survie, les variables explicatives X<sub>1</sub> , taux de concentration de la substance i avec le modèle

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon.$$

## 3.3.1. ETUDE COMPARATIVE REGCAZ/MCCP

Le traitement de ces données est effectué sous contrainte de positivité sur les coefficients.

Nous donnons, ci-dessous, les résultats obtenus respectivement par REGCAZ et MCCP.

## REGCAZ

Initialisation :  $\alpha$  est le vecteur nul

$$(y - X\alpha)'X = [377.0000 \quad 1877.5670 \quad 2246.6610 \quad 3337.7800]$$

Pivotage n°1 : on entre  $X_3$

$$\alpha = [0 \ 0 \ 0 \ 3.2219]$$

$$(y - X\alpha)'X = [5.6905 \ 195.4736 \ -99.8954 \ 0.0000]$$

Pivotage n°2 : on entre  $X_1$

$$\alpha = [0 \ 1.4851 \ 0 \ 2.4735]$$

$$(y - X\alpha)'X = [3.799915 \ 0 \ -90.4294 \ 0]$$

Pivotage n°3 : on entre  $X_0$

$$\alpha = [28.6123 \ 1.2083 \ 0 \ -0.5743]$$

Pivotage n°4 : on retire  $X_3$

$$\alpha = [23.3979 \ 1.23385 \ 0 \ 0]$$

$$(y - X\alpha)'X = [0.0000 \ 0.0000 \ -112.7556 \ -6.5000]$$

l'optimum est atteint.

#### MCCP

Initialisation :  $\alpha$  est le vecteur nul.

$$(y - X\alpha)'X = [377.0000 \ 1877.5670 \ 2246.6610 \ 3337.7800]$$

$$v' = [10962.0192 \ 8930.9098 \ 8751.9587 \ 10754.0594]$$

Pivotage n°1 : on entre  $X_0$

$$\alpha = [29.03846 \ 0 \ 0 \ 0]$$

$$(y - X\alpha)'X = [0.0000 \ 151.8112 \ -129.2659 \ -13.2585]$$

$$v' = [-10962.0192 \ 187.3118 \ 270.5412 \ 15.2024]$$

Pivotage n°2 : on entre  $X_1$

$$\alpha = [23.3979 \quad 1.23385 \quad 0 \quad 0]$$

$$(y - X\alpha)'X = [0.0000 \quad 0.0000 \quad -112.7556 \quad -6.5000]$$

l'optimum est atteint.

On a l'équation de régression :

$$\%ESTIME = 23.3979 + 1.2339X_1$$

et  $r^2 = 0.97800$  ,  $s^2 = 0.02200$  ,  $r^2 + s^2 = 1$  .

### 3.3.2. APPLICATION DE MCCB

Nous effectuons le traitement de ces données en imposant les contraintes:

$$0 \leq \beta_0 \leq 23 \text{ , et } 0 \leq \beta_i \leq 2, i=1,2,3.$$

Nous donnons ci-dessous, les différentes étapes du traitement par l'algorithme MCCB avec:

$$a_0 = 0 \text{ , } b_0 = 23 \text{ ; } a_i = 0, b_i = 2 \text{ , } i=1,2,3;$$

$$c = 1 \text{ si } X_i \in U \text{ et } c = -1 \text{ si } X_i \in W, i=0,1,2,3.$$

Initialisation

$$\|y\| = 106.77149 < \|y - b\| = 124.87580 : y^* = 0, \alpha_i = 0, U = X, V = \emptyset, W = \emptyset;$$

$$\langle y - y^*, cX_i \rangle_{i=0} = [377.5 \quad 1877.567 \quad 2246.661 \quad 3337.78]$$

Etape 1 : on entre dans V la variable 3

$$U = \{X_0, X_1, X_2\}, W = \emptyset, \alpha = [0 \quad 0 \quad 0 \quad 3.22192]$$

il y a non-admissibilité:

Etape 2 : on retire de V la variable 3

$$U=\{X_0, X_1, X_2\}, W=\{X_3\}, \alpha = [0 \ 0 \ 0 \ 2]$$

$$\langle y-y^*, cX_1 \rangle_{1,2,0} = [146.7 \ 833.411 \ 790.041 \ 0 \ -1265.86]$$

Etape 3 : on entre dans V la variable 1

$$U=\{X_0, X_2\}, W=\{X_3\}, \alpha = [0 \ 2.11137 \ 0 \ 2]$$

il y a non-admissibilité:

Etape 4 : on retire de V la variable 1

$$U=\{X_0, X_2\}, W=\{X_1, X_3\}, \alpha = [0 \ 2 \ 0 \ 2]$$

$$\langle y-y^*, cX_1 \rangle_{1,2,0} = [27.84 \ -43.96 \ 68.7168 \ -221.704]$$

Etape 5 : on entre dans V la variable 2

$$U=\{X_0\}, W=\{X_1, X_3\}, \alpha = [0 \ 2 \ 0.11915 \ 2]$$

$$\langle y-y^*, cX_1 \rangle_{1,2,0} = [18.09117 \ -0.9872 \ 0 \ -134.92605]$$

Etape 6 : on entre dans V la variable 0

$$U=\emptyset, W=\{X_1, X_3\}, \alpha = [12.99446 \ 2 \ -1.72437 \ 2]$$

il y a non-admissibilité:

Etape 7 : on retire de V la variable 2

$$U=\{X_2\}, W=\{X_1, X_3\}, \alpha = [2.14154 \ 2 \ 0 \ 0 \ 2]$$

$$\langle y-y^*, cX_1 \rangle_{1,2,0} = [0 \ 83.31163 \ -106.50388 \ 25.42954]$$

Etape 8 : on entre dans V la variable 1

$$U=\{X_2\}, W=\{X_3\}, \alpha = [5.237 \ 1.32288 \ 0 \ 2]$$

$$\langle y-y^*, cX_1 \rangle_{1,2,0} = [0 \ 0 \ -115.56448 \ 29.13847]$$



Etape 9 : on entre dans V la variable 3

$$U=\{X_2\}, W=\emptyset, \alpha = [28.61229 \quad 1.20828 \quad 0 \quad -0.57425]$$

il y a non-admissibilité:

Etape 10 : on retire de V la variable 0

$$U=\{X_2\}, W=\{X_0\}, \alpha = [23 \quad 1.26258 \quad 0 \quad 0.02357]$$

$$\langle y-y^*, cX_1 \rangle_{1,2,0} = [-0.7452 \quad 0 \quad -107.72801 \quad 0]$$

l'optimum est atteint.

L'équation de la régression est:

$$y^* = 23 + 1.26258X_1 + 0.02357X_3$$

avec  $r^2 = 0.97495$  ,  $s^2 = 0.02204$  ,  $r^2+s^2 = 0.997$

et les moyennes :  $(\Sigma y_1)/13 = 29.03846$  ,  $(\Sigma y^*_1)/13 = 28.98114$ .

Avec les contraintes imposées, nous concluons, en définitive, que la substance 3 a une influence moindre que celle de la substance 1, la substance 2 étant éliminée du modèle.

## 3.4. FONDEMENTS THEORIQUES DE L'ALGORITHME MCCC

Une caractérisation de la projection orthogonale d'un vecteur sur une variété linéaire est contenue dans le résultat suivant.

Soit  $h+L(V)$  une variété linéaire ( $V$  une partie de  $X$ ,  $h$  appartenant à  $R^n$ ) et  $z$  appartenant à  $h+L(V)$ :

**Proposition I.5**

$z$  est la projection orthogonale de  $y$  sur  $h+L(V)$  si et seulement si  $z-h$  est la projection orthogonale de  $y-h$  sur  $L(V)$ .

**Corollaire I.5**

$z$  est la projection de  $y$  sur  $h+L(V)$  si et seulement si  $y-z \in V^\perp$ .

Soit une partition  $\{U, V, W\}$  de  $X$ . Rappelons (section 1) que  $P(U, V, W)$  désigne la projection au sens des moindres carrés de  $y$  sur la variété linéaire  $VAR(U, V, W) = h+L(V)$  avec:

$$h = \sum \{a_i X_i, X_i \in U\} + \sum \{b_i X_i, X_i \in W\}.$$

Soit  $y^*$  la projection, au sens des moindres carrés de  $y$  sur  $C_{a, b}$ ; il existe une partition  $\{U, V, W\}$  de  $X$  telle que :

$$y^* = \sum \{a_i X_i, X_i \in U\} + \sum \{b_i X_i, X_i \in W\} + \sum \{\alpha_i X_i, a_i < \alpha_i < b_i, X_i \in V\}$$

**Proposition I.6**

$y^*$  est la projection  $P(U, V, W)$  de  $y$  sur la variété linéaire  $h+L(V)$ .

**Preuve**

Il suffit de montrer que  $y-y^* \in V^\perp$  (corollaire I.5). Soit  $X_i \in V$  :

1. Comme  $y^* - (\alpha_i - a_i) X_i \in C_{a, b}$  on a

$$\langle y - y^*, (y^* - (\alpha_i - a_i) X_i) - y^* \rangle \leq 0$$

d'où  $\langle y-y^*, X_1 \rangle \geq 0$

2. Comme  $y^* - (\alpha_1 - b_1)X_1 \in C_{a,b}$  on a

$$\langle y-y^*, (y^* - (\alpha_1 - b_1)X_1) - y^* \rangle \leq 0$$

d'où  $\langle y-y^*, X_1 \rangle \leq 0$ .

**CQFD**

Une partition  $\{U, V, W\}$  est dite admissible si  $P(U, V, W)$  appartient à  $C_{a,b}$ ; la condition d'optimalité d'une partition  $\{U, V, W\}$  provient de la proposition suivante :

**Proposition I.7**

$y^* = P(U, V, W)$  est la projection, au sens des moindres carrés de  $y$  sur  $C_{a,b}$  si et seulement si

$$y - y^* \in U^p \cap V^+ \cap (-W)^p.$$

**Preuve**

-Supposons que  $y^*$  est la projection de  $y$  sur  $C_{a,b}$ ; il suffit de montrer que  $y - y^* \in U^p \cap (-W)^p$  :

1. Soit  $X_1 \in U$ ;  $y^* - (\alpha_1 - b_1)X_1 \in C_{a,b}$  donc

$$\langle y - y^*, X_1 \rangle \leq 0$$

2. Soit  $X_1 \in W$ ;  $y^* - (b_1 - \alpha_1)X_1 \in C_{a,b}$  donc

$$\langle y - y^*, -X_1 \rangle \leq 0.$$

-Soit  $y^*$  tel que  $y - y^* \in U^p \cap V^+ \cap (-W)^p$ .

Pour tout  $z$  dans  $C_{a,b}$ ,  $z = \sum \{\beta_i X_i, i=1, \dots, p\}$

$$\text{on a } \langle y - y^*, z - y^* \rangle = \langle y - y^*, \sum \{(\beta_i - \alpha_i) X_i, X_i \in U\} \rangle +$$

$$+ \langle y - y^*, \sum \{(\beta_i - \alpha_i) X_i, X_i \in V\} \rangle + \langle y - y^*, \sum \{(\beta_i - b_i) X_i, X_i \in W\} \rangle$$

donc  $\langle y - y^*, z - y^* \rangle \leq 0$ .

Appelons (H) l'hypothèse suivante :

Soit  $\{U, V, W\}$  une partition de  $X$  admissible et non optimale; il existe au moins un vecteur  $X_r$  de  $X \setminus V$  tel que :

$$\langle y - P(U, V, W), cX_r \rangle > 0$$

si  $X_r \in U$  alors  $c=1$  sinon  $X_r \in W$  et  $c=-1$ .

Proposition I.8

Sous l'hypothèse (H), soit  $Q$  une partie de  $V$ ; posons

$$P(U-r, Q+r, W-r) = \sum \{\beta_i X_i, i=1, \dots, p\};$$

Si  $\|y - P(U-r, Q+r, W-r)\| < \|y - P(U, V, W)\|$  alors

si  $X_r \in U$  on a  $\beta_r > \alpha_r$  sinon  $X_r \in W$  et  $\beta_r < \alpha_r$ .

Preuve

On a

$$\begin{aligned} \|y - P(U-r, Q+r, W-r)\|^2 &= \|y - P(U, V, W)\|^2 + \|P(U, V, W) - P(U-r, Q+r, W-r)\|^2 + \\ &+ 2\langle y - P(U, V, W), P(U, V, W) - P(U-r, Q+r, W-r) \rangle \end{aligned}$$

Par ailleurs, si  $X_r \in U$  alors

$$\langle y - P(U, V, W), P(U, V, W) - P(U-r, Q+r, W-r) \rangle = \langle y - P(U, V, W), (\alpha_r - \beta_r) X_r \rangle$$

sinon  $X_r \in W$  et l'on a

$$\langle y - P(U, V, W), P(U, V, W) - P(U-r, Q+r, W-r) \rangle = \langle y - P(U, V, W), (\beta_r - \alpha_r) X_r \rangle.$$

CQFD

Notons  $z^0 = P(U, V, W) = \sum \{\alpha_i X_i, i=1, \dots, p\}$ .

Remarquons que, si  $Q = V$  et  $\{U-r, \setminus W-r\}$  non admissible, alors  $\text{VAR}(U-r, V+r, W-r)$  contient  $\text{VAR}(U, V, W)$  et la projection

$$y^1 = P(U-r, V+r, W-r) = \Sigma\{\beta_i X_i, i=1, \dots, p\}$$

vérifie les conditions suivantes :

$$C1. \|y-y^1\| < \|y-z^0\|$$

$$C2. \text{il existe } X_i \in V+r \text{ tel que } \beta_i > b_i \text{ ou } \beta_i < a_i$$

$$C3. \text{si } X_r \in U \text{ on a } \beta_r > a_r \text{ sinon } X_r \in W \text{ et } \beta_r < b_r.$$

Sous ces conditions, on énonce:

**Proposition I.9**

Il existe  $z^1 \in C_{a,b}$  tel que

$$\|y-y^1\|^2 < \|y-z^1\|^2 < \|y^1-z^0\|^2$$

**Preuve**

Pour tout  $z^1$  dans le segment  $]y^1 \quad z^0[$ , il existe  $\mu \in ]0 \quad 1[$  tel que

$$z^1 = y^1 + \mu(z^0 - y^1)$$

Remarquons que  $y^1 - z^0 \in L(V+r)$  et  $\langle y - y^1, y^1 - z^0 \rangle = 0$

$$\text{d'où} \quad \|y - y^1\|^2 < \|y - z^1\|^2 = \|y - y^1\|^2 + \mu^2 \|y^1 - z^0\|^2 < \|y^1 - z^0\|^2$$

Soit  $\mu^*$  tel que  $z^1$  soit dans  $C_{a,b}$ ; cela revient à imposer :

$$a_i \leq \beta_i + \mu^* (a_i - \beta_i) \leq b_i, \text{ pour tout } i = 1, \dots, p;$$

$$\text{d'où} \quad \mu^* = \text{MAX}\{((a_i - \beta_i) / (a_i - \beta_i)), \beta_i < a_i\} \cup \{(b_i - \beta_i) / (a_i - \beta_i), \beta_i > b_i\}$$

**CQFD**

Soit  $I = \{X_i \in X \mid \beta_i + \mu^*(\alpha_i - \beta_i) = a_i\}$ ;  $J = \{X_i \in X \mid \beta_i + \mu^*(\alpha_i - \beta_i) = b_i\}$ .

Soit  $P_1 = U - r + I$ ,  $Q_1 = V + r - (I + J)$ ,  $R_1 = W - r + J$ .

En partant de la partition  $\{U - r, V + r, W - r\}$  non admissible et  $z^0$ , nous avons obtenu par le procédé décrit ci-dessus, que nous nommons, en raccourci,  $\text{PROC}(U - r, V + r, W - r; z^0)$ , une partition  $\{P_1, Q_1, R_1\}$  et un vecteur  $z^1$  de  $\text{VAR}(P_1, Q_1, R_1) \cap C_{a,b}$  tels que  $P(P_1, Q_1, R_1)$  et  $z^1$  vérifient

$$\|y - P(P_1, Q_1, R_1)\|^2 \leq \|y - z^1\|^2 < \|y - z^0\|^2$$

L'algorithme MCCB repose sur la proposition suivante :

**Proposition I.10**

Sous l'hypothèse (H):

Il existe une partition  $\{P, Q, R\}$  de  $X$  admissible et meilleure que  $\{U, V, W\}$ .

**Preuve**

Notons  $z^0 = P(U, V, W) = \sum\{\alpha_i X_i, i=1, \dots, p\}$ .

Si  $\{U - r, V + r, W - r\}$  est admissible, elle est meilleure que  $\{U, V, W\}$  car  $\text{VAR}(U - r, V + r, W - r)$  contient  $\text{VAR}(U, V, W)$  et la proposition est démontrée.

Sinon, soit  $\{P_1, Q_1, R_1\}$  et  $z^1$  obtenus par  $\text{PROC}(U - r, V + r, W - r; z^0)$ ; on a

$$\|y - P(P_1, Q_1, R_1)\|^2 \leq \|y - z^1\|^2 < \|y - z^0\|^2.$$

Si  $\{P_1, Q_1, R_1\}$  est admissible la proposition est démontrée; sinon par  $\text{PROC}(P_1, Q_1, R_1; z^1)$ , nous obtenons  $\{P_2, Q_2, R_2\}$  et  $z^2$  vérifiant

$$\|y - P(P_2, Q_2, R_2)\|^2 \leq \|y - z^2\|^2 < \|y - z^1\|^2.$$

Nous construisons ainsi une suite  $\{z^i\}_{i \geq 0}$  d'éléments de  $C_{a,b}$  tels que la suite de leurs distances respectives à  $y$  soit strictement décroissante. La suite  $\{P_i, Q_i, R_i\}$  converge, en un nombre fini d'étapes, vers une partition admissible.

#### 4. CONTRAINTES D'INEGALITES LINEAIRES QUELCONQUES

La méthode LDP («Least Distance Programming») de Lawson & Hanson (1974) résoud le problème de la recherche d'un vecteur de norme minimale dans un polyèdre convexe. Nous en proposons une explication géométrique originale, basée sur la démonstration de la proposition I.11 qui caractérise la non-vacuité du polyèdre convexe.

Dans le problème de la régression sous contraintes d'inégalités linéaires quelconques sur les coefficients, Lawson & Hanson (1974) ont proposé la méthode LSI, fondée sur LDP et la notion de décomposition orthogonale d'une matrice.

Nous avons mis au point la méthode MCCIL basée sur une approche différente. De portée plus générale, elle résoud le problème de la détermination de la projection orthogonale d'un vecteur sur l'intersection d'un polyèdre convexe et d'une variété linéaire.

##### 4.1. LE PROBLEME LDP

Le problème LDP est:

$$\text{MIN}(\|\beta\|, \beta \in \text{POL}(R, r))$$

R étant une matrice  $m \times p$ , r un vecteur de  $R^m$  et

$$\text{POL}(R, r) = \{\beta \in R^p \text{ tel que } R\beta \geq r\}$$

Nous allons présenter la méthode de Lawson & Hanson

##### 4.1.1. DESCRIPTION DE LA METHODE

Notons  $z^1 = (0, 0, \dots, 0, 1)'$ , vecteur de  $R^{p+1}$  et  $A' = [R:r]$  :

1. Résolution du problème NNLS

$$\text{MIN}(\|z^1 - A\alpha\|, \alpha \geq 0) = \|z^1 - A\alpha^*\|;$$

2. Si  $\|z^1 - A\alpha^*\| = 0$  alors  $\text{POL}(R, r)$  est vide sinon

$$\beta^* = (1/\|z^1 - A\alpha^*\|^2) R' \alpha^*$$

est la solution du problème LDP.

#### 4.1.2. JUSTIFICATION DE LAWSON & HANSON

Les conditions de Kuhn & Tucker (voir section 1.3) du problème LDP sont:

$$u_i \geq 0, \quad i=1, \dots, m;$$

$$\beta^* = R'u$$

$$u_i (R^i \beta^* - r) = 0, \quad i=1, \dots, m;$$

il existe un sous-ensemble  $I$  de  $\{1, \dots, m\}$  telle que

$$u_i = 0, \quad R^i \beta^* - r > 0 \text{ si } i \in I$$

$$u_i \geq 0, \quad R^i \beta^* - r = 0 \text{ sinon.}$$

Si  $\|z^1 - A\alpha^*\| = 0$  i.e.  $[R:r]' \alpha^* - z^1 = 0$ : soit  $\beta \in \text{POL}(R, r)$ :

$$0 = [\beta' : -1] ([R:r]' \alpha^* - z^1) = (R\beta - r)' \alpha^* + 1 > 0 : \text{contradiction}$$

donc l'inégalité  $R\beta > r$  n'est jamais vérifiée.

Si  $\|z^1 - A\alpha^*\|$  différent de 0, Lawson & Hanson posent

$$u = (1/\|z^1 - A\alpha^*\|^2) \alpha^*$$

et établissent que  $\beta^*$  vérifie la contrainte  $R\beta^* \geq r$ , les conditions de Kuhn & Tucker précitées et est unique (pp 166-167).

#### 4.1.3. JUSTIFICATION PROPOSEE

La proposition ci-dessous caractérise la non-vacuité de  $\text{POL}(R, r)$  et la suivante justifie complètement la méthode LDP.

##### Proposition I.11

Le polyèdre  $\text{POL}(R, r)$  est non vide si et seulement si, pour tout  $\alpha$  vecteur de  $\mathbb{R}^n$ ,  $\alpha > 0$  (composantes non toutes nulles), on a  $R'\alpha \neq 0$  ou  $r'\alpha \leq 0$ .



**Preuve**

-Supposons  $POL(R,r)$  non vide; il existe donc  $\beta \in R^p$  tel que  $R\beta \geq r$ ; soit  $\alpha > 0$  tel que  $R'\alpha = 0$  et  $r'\alpha > 0$ , on a :

$$\beta'R' - r' \geq 0$$

et par suite  $\beta'R'\alpha - r'\alpha = -r'\alpha \geq 0$  : contradiction.

-Supposons que, pour tout  $\alpha > 0$ , on a :

$$R'\alpha \neq 0 \text{ ou } r'\alpha \leq 0.$$

Soit  $A$  définie par  $A' = [R:r]$ ,  $c$  un nombre réel positif et  $z^c$  le vecteur de  $R^{p+1}$  défini par :

$$z^c = (0, 0, \dots, 0, c)'$$

L'hypothèse est équivalente à :  $z^c \in R^{p+1} \setminus C(A)$ .

Soit  $\alpha^c$  réalisant le minimum de  $\{\|A\alpha - z^c\|, \alpha \in R^p, \alpha \geq 0\}$ .  
Il existe un ensemble  $B$  de vecteurs colonnes de  $A$  tel que :

$$\langle A\alpha^c - z^c, A_i \rangle = 0 \text{ et } \alpha_i^c > 0 \text{ si } A_i \in B.$$

$$\langle A\alpha^c - z^c, A_i \rangle \geq 0 \text{ et } \alpha_i^c = 0 \text{ sinon.}$$

Posons  $v = A\alpha^c - z^c$ ; remarquons que  $A'v \geq 0$ .

Comme  $\|v\|^2 = v'v = -v'z^c = -cv_{p+1}$ , on a  $-v_{p+1} > 0$ .

Posons  $\beta_i = (-1/v_{p+1})v_i$ ,  $i=1, \dots, p$ .

On a  $(-1/v_{p+1})A'v = R\beta - r \geq 0$  et  $POL(R,r)$  est donc non vide.

**Remarques**

1. Pour tout  $c > 0$ ,  $\alpha^c = c\alpha^1$  avec  $\alpha^1$  réalisant le minimum de  $\{\|A\alpha - z^1\|, \alpha \in R^p, \alpha \geq 0\}$  et par suite :

$$\beta = (-1/v_{p+1})R'\alpha^c = (1/(c-r'\alpha^c))R'\alpha^c = (1/(1-r'\alpha^1))R'\alpha^1$$

est indépendant de  $c$  et  $1-r'\alpha^1 = \|A\alpha^1 - z^1\|^2 > 0$ .

2. Si  $r$  a toutes ses composantes négatives ou nulles, la proposition devient triviale:  $z^c$  appartient au cône polaire de  $C(A)$ ; par suite  $A\alpha^c = 0$  et  $\beta = 0$ .

Dans ce cas la démonstration amène à construire l'élément de  $POL(R,r)$  de norme minimale. Nous allons montrer que cette propriété est, en fait, vraie pour tout  $r$  de  $R^m$ .

**Proposition I.12**

La solution du problème LDP :

$$\text{MIN}\{\|\beta\|, \beta \in \text{POL}(R,r)\} \text{ est } \beta^* = (1/(1-r'\alpha^*))R'\alpha^*$$

avec  $\alpha^*$  solution du problème NNLS:

$$\text{MIN}\{\|z^1 - A\alpha\|, \alpha \geq 0\};$$

$z^1 = (0, 0, \dots, 0, 1)'$ , vecteur de  $R^{p+1}$  et  $A' = [R:r]$ .

**Preuve**

Soit  $\beta \in \text{POL}(R,r)$ ; posons  $\tau = \beta - \beta^*$ ; notons:

$R^j$ ,  $j=1, \dots, m$ , les vecteurs lignes de  $R$ ;

$$B = \{A_j \mid \langle A\alpha^* - z^1, A_j \rangle = 0 \text{ et } \alpha_j^* > 0\}.$$

Pour tout  $j$  tel que  $A_j \in B$ , on a:

$$R^j \beta^* = r_j, \quad R^j \beta \geq r_j \text{ et } R^j \tau \geq 0$$

d'où  $\|\beta\|^2 = \|\beta^*\|^2 + \|\tau\|^2 + (2/(1-r'\alpha^*)) \sum \{\alpha_j^* R^j \tau, j \mid A_j \in B\} > \|\beta^*\|^2$ .

**4.2. LE PROBLEME MCCIL**

Le problème MCCIL (Moindres Carrés sous Contraintes d'Inégalités Linéaires) concerne la détermination de la projection orthogonale d'un vecteur sur un polyèdre convexe:

$$\text{MIN}\{\|X\beta - y\|, \beta \in \text{POL}(R,r)\}.$$

Lawson et Hanson (1974) ont proposé une méthode de résolution que nous allons présenter.

#### 4.2.1. LA METHODE LSI DE LAWSON & HANSON

Elle est basée sur la procédure LDP de recherche de vecteur de norme minimale dans un polyèdre et la notion de décomposition orthogonale d'une matrice. Soit A une matrice (n x p) de rang k.

Proposition (Lawson et Hanson, 1974, p 13)

Il existe une matrice orthogonale (n x n) H, une matrice orthogonale (p x p) K telles que

$$A = HGK'$$

avec

$$G = \begin{bmatrix} G_{11} & 0 \\ 0 & 0 \end{bmatrix}$$

telle que  $G_{11}$  matrice (k x k), non singulière, pouvant être choisie triangulaire ou diagonale.

On détermine une décomposition de  $X = HGK'$ ; posons

$$H = \begin{bmatrix} H_1 & H_2 \end{bmatrix}$$

avec  $H_1$  (n x p); comme X de rang p, on peut écrire

$$X = \begin{bmatrix} H_1 & H_2 \end{bmatrix} \begin{bmatrix} G_{11} \\ 0 \end{bmatrix} K'$$

on pose  $\beta = K\alpha$ ;  $H$  étant orthogonale, on peut écrire

$$\|X\beta - y\|^2 = \|H'(X\beta - y)\|^2 = \left\| \begin{bmatrix} H_1' y \\ H_2' y \end{bmatrix} - \begin{bmatrix} G_{11} \alpha \\ 0 \end{bmatrix} \right\|^2 = \|G_{11} \alpha - y_1\|^2 + \|y_2\|^2$$

avec  $y_i = H_i y$ ,  $i=1,2$ ; soit  $z = G_{11} \alpha - y_1$ :

$$\|X\beta - y\|^2 = \|z\|^2 + \|y_2\|^2$$

Le problème MCCIL revient au problème LDP:

$$\text{MIN} \{ \|z\|, RKG_{11}^{-1} z \geq r - RKG_{11}^{-1} y_1 \}$$

#### 4.2.2. LA METHODE MCCIL

Basée aussi sur LDP, elle résoud le problème, équivalent, de la recherche du vecteur  $X\beta$ ,  $\beta \in \text{POL}(R, r)$ , le plus «proche» de

$$y^L = X\beta^L = X(X'X)^{-1}X'y$$

constituant la solution du problème MC:  $\text{MIN} \{ \|y - X\beta\|, \beta \in R^p \}$ ,

car on a, pour tout  $\beta \in R^p$ :

$$\|y - X\beta\|^2 = \|y - y^L\|^2 + \|y^L - X\beta\|^2;$$

posons  $\tau = \beta - \beta^L$ ; le problème MCCII devient

$$\text{MIN} \{ \|X\tau\|, R\tau \geq r - R\beta^L \}$$

posons  $X\tau = \delta$ ; cela revient à  $\tau = (X'X)^{-1}X'\delta$  et  $\delta \in L(X)$ ;

le problème MCCIL devient

$$\text{MIN} \{ \|\delta\|, R(X'X)^{-1}X'\delta \geq r - R\beta^L \text{ et } \delta \in L(X) \}$$

équivalent au problème LDP :

$$\text{MIN} \{ \|\delta\|, R(X'X)^{-1}X'\delta \geq r - R\beta^L \}$$

car la solution de ce dernier problème est dans  $L(X)$ .

C'est 
$$\delta^* = (1/(1-\alpha^{*'}(r-R\beta^L)))X(X'X)^{-1}R'\alpha^*$$

avec  $\alpha^*$ , vecteur de  $R^n$ , solution du problème NNLS :

$$\text{MIN} \{ \|z_1 - A\alpha\|, \alpha \geq 0 \}$$

$z_1' = (0, 0, \dots, 0, 1)$ , vecteur de  $R^{n+1}$  et la matrice  $(n+1) \times n$

$$A = \begin{bmatrix} X(X'X)^{-1}R' \\ (r - R\beta^L)' \end{bmatrix}$$

Ecrivons  $\delta^* = X\tau^*$  ; finalement, la solution du problème MCCIL est

$$\beta^{\text{MCCIL}} = \beta^L + \tau^* = \beta^L + 1/[1 - \alpha^{*'}(r - R\beta^L)](X'X)^{-1}R'\alpha^*$$

et 
$$y^{\text{MCCIL}} = y^L + \delta^* .$$

#### Remarques

1. Si  $r - R\beta^L \leq 0$  alors  $\{\delta; R(X'X)^{-1}X'\delta \geq r - R\beta^L\}$  contient le vecteur nul et  $\delta^* \equiv 0$  est optimum absolu. Comme cela signifie  $R\beta^L \geq r$ , on a bien dans ce cas,  $\beta^{\text{MCCIL}} = \beta^L$ .

2. Si  $R\beta^{\text{MCCIL}} = r$  alors

$$R\beta^L + (1/(1 - \alpha^{*'}(r - R\beta^L)))R(X'X)^{-1}R'\alpha^* = r$$

d'où 
$$(1/(1 - \alpha^{*'}(r - R\beta^L)))\alpha^* = (R(X'X)^{-1}R')^{-1}(r - R\beta^L)$$

et finalement 
$$\beta^{\text{MCCIL}} = \beta^L + (X'X)^{-1}R'(R(X'X)^{-1}R')^{-1}(r - R\beta^L) .$$

Nous retrouvons ainsi l'expression classique de la solution du problème MCCEL (Moindres Carrés sous Contraintes d'Égalités Linéaires):

$$\text{MIN}\{\|y-X\beta\|, R\beta=r, \beta \in R^p\}$$

que donnent par exemple, Judge et Takayama (1966, page 168).

#### 4.2.3. MCCIL GENERALISE

En suivant le procédé décrit à la précédente section, nous allons donner une expression de la solution du problème MCCIL généralisé:

$$\text{MIN}\{\|z-y\|, z \in (h+L) \cap \text{POL}(Q,r)\} \quad (1)$$

où  $h \in R^n$ ,  $Q$  une matrice  $m \times n$ ,  $r \in R^m$  fixés et  $L$  un sous-espace vectoriel de  $R^n$ .

Notons  $\Omega$  l'opérateur de projection sur  $L$  et  $y^{vL}$  la projection de  $y$  sur la variété linéaire  $h+L$ :

$$y^{vL} = h + \Omega(y-h);$$

$$\delta = z - y^{vL}.$$

Le problème (1) revient à rechercher:

$$\text{MIN}\{\|\delta\|, \delta \in L, Q\delta \geq r - Qy^{vL}\} \quad (2)$$

Nous énonçons tout d'abord:

#### Proposition I.13

La solution du problème

$$\text{MIN}\{\|\delta\|, \delta \in L, R\delta \geq r\} \text{ est } \Omega\tau^* \text{ avec}$$

$$\tau^* \text{ solution du problème: } \text{MIN}\{\|\tau\|, RQ\tau \geq r\}.$$

Preuve

Soit  $\delta^*$  la solution de  $\text{MIN}\{\|\delta\|, \delta \in L, R\delta \geq r\}$ .

Supposons  $\|\delta^*\| < \|\Omega\tau^*\|$  alors  $\|\delta^*\| < \|\Omega\|\|\tau^*\| = \|\tau^*\|$

car l'opérateur  $\Omega$  est de norme 1, d'où la contradiction.

CQFD

Le problème (2) revient donc au problème LDP:

$$\text{MIN}\{\|\delta\|, \Omega\delta \geq r - Qy^{vL}\}$$

car la solution de ce dernier problème est dans  $L$ .

On a finalement la solution :  $z^* = y^{vL} + 1/(1 - \alpha^{*'}(r - Qy^{vL}))\Omega Q'\alpha^*$

avec  $\alpha^*$  vecteur de  $R^n$ , solution du problème NNLS :

$$\text{MIN}\{\|z_1 - A\alpha\|, \alpha \geq 0\}$$

$z_1' = (0, 0, \dots, 0, 1)$ , vecteur de  $R^{n+1}$  et la matrice  $(n+1) \times n$

$$A = \begin{bmatrix} \Omega Q' \\ (r - Qy^{vL})' \end{bmatrix}$$

Remarques

1. Nous retrouvons, bien entendu, la solution du problème MCCIL «restreint» en posant:

$$h=0, L = L(X) \text{ avec } \Omega = X(X'X)^{-1}X', Q = R(X'X)^{-1}X'.$$

2. Une application possible de ce résultat est la résolution du problème:

$$\text{MIN}\{\|X\beta - y\|, \beta \geq 0, R\beta = r\}$$

équivalent à  $\text{MIN}\{\|z-y\|, z \in (h+L) \cap \text{POL}(Q,r)\}$

avec la variété linéaire  $\{X\beta \mid R\beta = r\} = h+L$

en posant  $h = X(X'X)^{-1}R'(R(X'X)^{-1}R')^{-1}r$

$L = \text{KER}(R(X'X)^{-1}X')$  d'opérateur de projection

$Q = X(X'X)^{-1}[I - R'(R(X'X)^{-1}R')^{-1}R(X'X)^{-1}]X'$

et  $Q = (X'X)^{-1}X'$ ,  $r = 0$ ;

la solution est:

$$y^* = y^{\text{MCCCEL}} + 1/1 + (y^{\text{MCCCEL}}'X(X'X)^{-1}\alpha^*)QX(X'X)^{-1}\alpha^*$$

avec  $\alpha^*$  vecteur de  $\mathbb{R}^m$ , solution du problème NNLS :

$$\text{MIN}\{\|z_1 - A\alpha\|, \alpha \geq 0\}$$

$z_1' = (0, 0, \dots, 0, 1)$ , vecteur de  $\mathbb{R}^{m+1}$  et la matrice  $(n+1) \times m$ .

$$A = \begin{bmatrix} Q'X(X'X)^{-1} \\ -y^{\text{MCCCEL}}'X(X'X)^{-1} \end{bmatrix}$$

Notons que, dans le cas où  $Q$  est la matrice uniligne formée de 1 et  $r=1$ :

$$(h+L) \cap \text{POL}(Q,r) = \{X\beta \mid \beta \geq 0 \text{ et } \sum \beta_i = 1\}$$

est l'enveloppe convexe de  $X$ .

#### 4.3. ETUDE NUMERIQUE

En application de la méthode LSI, Lawson et Hanson (1974, pp.169-173) traitent le problème suivant:



étant donné

$$t' = [0.25 \quad 0.50 \quad 0.50 \quad 0.80] , \quad w' = [0.5 \quad 0.6 \quad 0.7 \quad 1.2],$$

on recherche une fonction affine  $F : F(x) = \beta_1 x + \beta_2$  telle que  $(F(t_i))_{i=1, \dots, 4}$ , «approche» au mieux  $w$ , au sens des moindres carrés, sous les contraintes:  $F'(x) \geq 0$ ,  $F(0) \geq 0$  et  $F(1) \leq 1$ .

Nous présentons, ci-dessous, la formalisation du problème, le détail des calculs et les résultats obtenus.

Le problème revient à rechercher

$$\text{MIN} \{ \|X\beta - y\| , \beta \in \text{POL}(R, r) \} ;$$

avec  $n = 4$ ,  $p = 2$ ,  $m = 3$  et

$$X = \begin{bmatrix} 0.25 & 1 \\ 0.50 & 1 \\ 0.50 & 1 \\ 0.80 & 1 \end{bmatrix} ; \quad y = \begin{bmatrix} 0.5 \\ 0.6 \\ 0.7 \\ 1.2 \end{bmatrix} ;$$

$$R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{bmatrix} ; \quad r = \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix} ;$$

$$\beta^{\perp} = \begin{bmatrix} 1.30041 \\ 0.08354 \end{bmatrix} ; \quad r - R\beta^{\perp} = \begin{bmatrix} -1.30041 \\ -0.08354 \\ 0.38395 \end{bmatrix} ;$$

$$(X'X)^{-1}R' = \begin{bmatrix} 6.58436 & -3.37449 & -3.20988 \\ -3.37449 & 1.97942 & 1.39506 \end{bmatrix} ;$$

$$X(X'X)^{-1}R' = \begin{bmatrix} -1.72840 & 1.13580 & 0.59259 \\ -0.08230 & 0.29218 & -0.20988 \\ -0.08230 & 0.29218 & -0.20988 \\ 1.89300 & -0.72016 & -1.17284 \end{bmatrix} ;$$

La solution NNLS :  $\alpha^* = [0.00000 \quad 0.00000 \quad 0.19567]$ ;

on a  $1/[1-\alpha^*(r-R\beta^L)] = 1.08123$

et finalement:

$$\beta^{MCCIL} = \beta^L + 1.08123(X'X)^{-1}R'\alpha^* = \begin{bmatrix} 0.6213 \\ 0.3787 \end{bmatrix} ;$$

le vecteur des résidus est:

$$e' = (y - X\beta^{MCCIL})' = [-0.03401 \quad -0.08934 \quad 0.01066 \quad 0.32426]$$

de norme  $\|e\| = 0.33823$ , la qualité de la régression étant mesurée par

$$r^2 = 0.78837$$

$$s^2 = 0.04504 \text{ et } r^2 + s^2 = 0.83341.$$

Lawson et Hanson donnent la solution:

$$\beta^{LS1} = [0.621 \quad 0.379].$$

et  $\|e\| = 0.338$ .

**Chapitre II**

**ANALYSE CANONIQUE SOUS CONTRAINTES**

**DE POSITIVITE DES FACTEURS**

## II. ANALYSE CANONIQUE SOUS CONTRAINTES DE POSITIVITE DES FACTEURS

Le problème avec contraintes de positivité sur les facteurs canoniques, dont Tenenhaus (1988) a développé une théorie générale, est traité par un algorithme de type «moindres carrés alternés» (ALS), n'assurant pas l'obtention d'une solution optimale. En nous limitant au cas de contraintes de positivité sur les facteurs de l'un des ensembles de variables, nous en proposons une résolution originale, basée sur une approche ADOPT (voir Introduction).

Une section préliminaire fixe des notations et rappelle la notion de couple stationnaire et ses propriétés principales.

La section 2 traite de l'analyse canonique classique et de ses méthodes de résolution.

L'analyse canonique sous contraintes est abordée dans la dernière section. Nous développons un algorithme, nommé ACCE (Analyse Canonique d'un Cône polyédrique convexe et d'un Espace vectoriel), qui peut être considéré comme une généralisation de l'algorithme NNLS (Lawson & Hanson, 1974), au cas de plusieurs variables à expliquer. L'optimalité de la solution n'a pas été démontrée mais la mise en oeuvre de cet algorithme n'est pas dépendante d'un vecteur initial.

## 1. PRELIMINAIRES

Nous commençons par fixer quelques notations. Nous rappelons ensuite une caractérisation de la stationnarité d'un couple de vecteurs de deux espaces vectoriels, à partir de laquelle, peut être définie une caractérisation originale de la stationnarité dans le cas d'un cône polyédrique convexe et d'un espace vectoriel.

### 1.1. NOTATIONS

Soit  $S = \{S_1, S_2, \dots, S_p\}$ ,  $T = \{T_1, T_2, \dots, T_q\}$  deux familles de vecteurs de  $\mathbb{R}^n$ , linéairement indépendants.

L'espace vectoriel  $L(T)$  (resp le cône  $C(S)$ ) engendré par  $T$  (resp. par  $S$ ) a pour opérateur de projection, au sens des moindres carrés,  $P_L(T)$  (resp.  $P_C(S)$ ).

Notons

$\alpha^h, h=1, 2, \dots, m$  (resp.  $\beta^h$ ),  $m$  vecteurs de  $\mathbb{R}^p$  (resp. de  $\mathbb{R}^q$ )

avec  $m = \text{MIN}(p, q)$ . Soit  $x \in \mathbb{R}^n$ ; rappelons que

$$P_L(T)(x) = T(T'T)^{-1}T'x$$

et convenons que  $\langle x, y \rangle_T = \langle P_L(T)(x), P_L(T)(y) \rangle$ .

### 1.2. NOTION DE STATIONNARITE

Un couple  $(A, B)$  de  $L(S) \times L(T)$  est dit stationnaire s'il vérifie

$$P_L(S) \circ P_L(T)A = \pi A \text{ et } P_L(T) \circ P_L(S)B = \pi B$$

avec  $\pi = \cos^2(A, B)$ .

Une caractérisation de la stationnarité est contenue dans la proposition suivante:

#### Proposition II.1

Le couple  $(A, B)$  de  $L(S) \times L(T)$  est stationnaire si et seulement si pour tout  $s \in L(S)$ ,  $t \in L(T)$  :

$$\langle A, s \rangle_T = \pi \langle A, s \rangle ; \langle B, t \rangle_S = \pi \langle B, t \rangle.$$

Preuve

1. Supposons (A,B) stationnaire, on a

$$\langle A, s \rangle_r = \langle P_L(\tau)A, s \rangle = \langle P_L(s) \circ P_L(\tau)A, s \rangle = \langle \pi A, s \rangle = \pi \langle A, s \rangle$$

$$\langle B, t \rangle_s = \langle P_L(s)B, t \rangle = \langle P_L(\tau) \circ P_L(s)B, t \rangle = \langle \pi B, t \rangle = \pi \langle B, s \rangle$$

2. Supposons que, pour tout  $s \in L(S)$ ,  $t \in L(T)$ :

$$\langle A, s \rangle_r = \pi \langle A, s \rangle, \quad \langle B, t \rangle_s = \pi \langle B, t \rangle;$$

$$\langle P_L(\tau)A, s \rangle = \langle A, s \rangle_r = \langle \pi A, s \rangle$$

on a :

$$\langle P_L(s)B, t \rangle = \langle B, t \rangle_r = \langle \pi B, t \rangle.$$

Q.E.D.

Proposition II.2

Le couple (A,B) de  $L(S) \times L(T)$  est stationnaire si et seulement si

$$\cos^2(A,B) = \cos^2(A, P_L(\tau)A) = \cos^2(B, P_L(s)B).$$

Preuve

Notons  $\pi = \cos^2(A,B)$ .

1. Supposons (A,B) stationnaire. On a, pour tout  $s \in L(S)$ ,  $t \in L(T)$  :

$$\langle A, s \rangle_r = \pi \langle A, s \rangle ; \quad \langle B, t \rangle_s = \pi \langle B, t \rangle$$

et pour  $s=A, t=B$ :

$$\langle A, A \rangle_r = \pi \langle A, A \rangle ; \quad \langle B, B \rangle_s = \pi \langle B, B \rangle$$

2. Supposons  $\pi = \cos^2(A, P_L(\tau)A) = \cos^2(B, P_L(s)B)$ ;

il existe donc  $a, b \in \mathbb{R}^+$  tels que  $B = bP_L(\tau)A$ ,  $A = aP_L(s)B$

$$\text{d'où} \quad \cos(A,B) = b \|P_L(\tau)A\|^2 / (\|A\| \|B\|) = a \|P_L(s)B\|^2 / (\|A\| \|B\|)$$

donc  $ab = 1/\pi$ .

Notons, au passage, que  $a$  et  $b$  doivent être de même signe.

Finalement  $P_L(s) \circ P_L(\tau) A = P_L(s) ((1/b)B) = (1/ab)A = \pi A$

et  $P_L(\tau) \circ P_L(s) B = P_L(\tau) ((1/a)A) = (1/ab)B = \pi B$ .

### Proposition II.3

Si  $(A, B)$  est stationnaire et normé alors

$$A = (1/\sqrt{\pi}) P_L(s) B, \quad B = (1/\sqrt{\pi}) P_L(\tau) A$$

Preuve

Il suffit de remarquer que

$$\|P_L(\tau)(A)\| = \|P_L(s)(B)\| = \sqrt{\pi}.$$

### Proposition II.4

Si  $(A, B)$  est un couple de  $L(S) \times L(T)$  vérifiant

$$\cos^2(A, B) = \text{MAX}\{\cos^2(s, t), s \in L(S), t \in L(T)\}$$

alors le couple  $(A, B)$  est stationnaire.

Preuve

L'hypothèse implique

$$\cos^2(A, B) = \cos^2(A, P_L(\tau)A) = \cos^2(B, P_L(s)B) \quad \text{car}$$

$$\cos^2(A, P_L(\tau)A) = 1 - (\|A - P_L(\tau)A\|^2 / \|A\|^2) \geq 1 - (\|A - P_L(B)A\|^2 / \|A\|^2)$$

$$\cos^2(B, P_L(s)B) = 1 - (\|B - P_L(s)B\|^2 / \|B\|^2) \geq 1 - (\|B - P_L(A)B\|^2 / \|B\|^2).$$

Remarque

Si  $A \in L(S)$  vérifie la stationnarité :  $P_L(s) \circ P_L(\tau) A = \pi A$ , alors pour tout  $b \in \mathbb{R}^*$  et en notant  $B = b P_L(\tau) A$ , le couple  $(A, B)$  est stationnaire.

Nous noterons  $(A, (1/\sqrt{\pi})P_L(\tau)A)$  un couple stationnaire. Si  $A$  est normé, le couple est normé.

Un couple  $(A, B)$  de  $C(S) \times L(T)$  est dit stationnaire s'il vérifie

$$P_C(s) \circ P_L(\tau) A = \pi A$$

$$P_L(\tau) \circ P_C(s) B = \pi B$$

Il existe un sous-ensemble  $U$  de  $S$  tel que  $P_C(s) B = P_L(u) B$ ; d'après la remarque précédente,  $(A, B)$  est un couple stationnaire de  $L(U) \times L(T)$ .

#### Proposition II.5

Si  $(A, B)$  est un couple de  $C(S) \times L(T)$  vérifiant

$$\cos^2(A, B) = \text{MAX}\{\cos^2(s, t), s \in C(S), t \in L(T)\},$$

le couple  $(A, B)$  est stationnaire dans  $C(S) \times L(T)$ .

#### Preuve

Il existe  $U$  inclus dans  $S$  tel que

$$P_C(s) B = P_L(u) B.$$

L'hypothèse implique

$$\cos^2(A, B) = \cos^2(A, P_L(\tau) A) = \cos^2(B, P_C(s) B) = \cos^2(B, P_L(u) B)$$

car 
$$\cos^2(A, P_L(\tau) A) = 1 - (\|A - P_L(\tau) A\|^2 / \|A\|^2)$$

$$\cos^2(B, P_C(s) B) = 1 - (\|B - P_C(s) B\|^2 / \|B\|^2)$$

Les propositions II.2 et II.4 permettent de conclure que  $(A, B)$  est un couple stationnaire de  $L(U) \times L(T)$ . D'où

$$P_L(\tau) \circ P_C(s) B = \pi B \quad \text{et} \quad P_C(s) \circ P_L(\tau) A = \pi A.$$



## 2. ANALYSE CANONIQUE DE DEUX ESPACES VECTORIELS

Nous effectuons un bref rappel du problème et des méthodes de résolution de l'analyse canonique de deux espaces vectoriels.

### 2.1. POSITION DU PROBLEME

Proposé par H. Hotelling (1936), le problème consiste à rechercher

$$\text{MAX}\{\sum\{\cos^2(S\alpha^h, T\beta^h), h=1, 2, \dots, m\}\}$$

sous les contraintes

$$\alpha^h 'S' S\alpha^k = \delta_{hk} \quad , \quad \beta^h 'T' T\beta^k = \delta_{hk} \quad h, k = 1, 2, \dots, m$$

$\delta_{hk}$  étant le symbole de Kronecker.

Les vecteurs  $A^h = S\alpha^h, B^h = T\beta^h$  pour lesquels l'optimum est atteint, appelés caractères canoniques, sont respectivement les vecteurs propres de  $P_L(s) \circ P_L(r)$  et  $P_L(r) \circ P_L(s)$ , associés aux  $m$  plus grandes valeurs propres  $\pi_1, \pi_2, \dots, \pi_m$ . Les vecteurs de coefficients  $\alpha^h$  et  $\beta^h$  sont dits facteurs (ou variables) canoniques.

### 2.2. METHODES DE RESOLUTION

Nous décrivons, ci-après, les deux méthodes classiques de résolution de l'analyse canonique.

#### 2.2.1. PRODUIT DE MATRICES DE COVARIANCES

On suppose les vecteurs  $\{S_j\}, \{T_j\}$  centrés.

Notons  $A = S\alpha, B = T\beta,$

on a  $P_L(s) \circ P_L(r) A = \pi A ; P_L(r) \circ P_L(s) B = \pi B.$

Notons  $V_{11} = S'S, V_{22} = T'T, V_{12} = S'T = V_{21}'.$

On peut écrire

$$V_{11}^{-1} V_{12} V_{22}^{-1} V_{21} \alpha = \pi \alpha \quad \text{et} \quad V_{22}^{-1} V_{21} V_{11}^{-1} V_{12} \beta = \pi \beta,$$

les conditions de normalisation étant :

$$A'A = \alpha'V_{11}\alpha = 1 ; B'B = \beta'V_{22}\beta .$$

Si  $p < q$ , on recherchera d'abord  $\alpha$  et  $\beta$  sera déterminé par

$$\beta = (1/\sqrt{\pi})V_{22}^{-1}V_{21}\alpha .$$

### 2.2.2. ALGORITHME ALS

C'est une méthode classique de résolution de ce problème, basée sur le principe des «projections alternées au sens des moindres carrés» (Alternating Least Squares). Nous en proposons une description, inspirée de Banzécri (1974).

Appliqué avec l'opérateur  $P_{L(s)} \circ P_{L(r)}$ , l'algorithme produit une suite de systèmes orthonormés :

$$A(t) = \{A^h(t), h = 1, 2, \dots, m\}$$

avec  $A^h(t) = \{A_i^h(t), i = 1, 2, \dots, n\}$ .

Le système initial  $A(0)$  est une partie quelconque de  $R^n \setminus T^1$ ; on passe de  $A(t-1)$  à  $A(t)$  en orthogonalisant (par Gram-Schmidt, par exemple) le système  $\{P_{L(s)} \circ P_{L(r)} A^h(t-1), h=1, 2, \dots, m\}$  et en normant les vecteurs résultants.

Les vecteurs propres  $A^h$  sont les limites  $A^h(\infty)$  et les valeurs propres  $\pi_h$  sont les limites  $\|P_{L(s)} \circ P_{L(r)} A^h(\infty)\|$ . On en déduit le système  $\{B^h, h = 1, 2, \dots, m\}$  :

$$B^h = (1/\sqrt{\pi_h}) P_{L(r)} (A^h) .$$

Cet algorithme est approprié à la recherche des caractères canoniques et garantit l'obtention de la solution optimale.

Dans la suite, on s'intéressera uniquement au problème de l'obtention du premier couple canonique par l'algorithme simplifié noté, en raccourci, ALS(S,T,u), appliqué avec l'opérateur  $P_{L(s)} \circ P_{L(r)}$  et de vecteur initial u.

### 3. ANALYSE CANONIQUE D'UN CONE POLYEDRIQUE CONVEXE ET D'UN ESPACE VECTORIEL

On se limite à la recherche du premier couple de l'analyse canonique de l'espace vectoriel  $L(T)$  et du cône polyédrique convexe  $C(S)$ . L'algorithme ALS ne garantit plus, dans ce cas, l'obtention d'une solution optimale mais seulement stationnaire.

Tenenhaus (1988) a montré que dans le cas, plus général, de deux cônes, la solution optimale est le premier couple canonique de deux sous-espaces vectoriels de  $L(S)$  et  $L(T)$ .

Dans le cas qui nous occupe, nous donnons (proposition II.8, section 3.4) une démonstration originale de ce résultat, constituant la condition ADOPT nécessaire (voir Introduction) de l'algorithme proposé, nommé ACCE (Analyse Canonique d'un Cône et d'un Espace vectoriel). Nous effectuons une étude numérique de cet algorithme et proposons un complément à l'étude comparative REGCAZ/MCCP (chapitre I, section 2).

#### 3.1. POSITION DU PROBLEME

Il s'agit dans le problème que l'on s'est posé, de déterminer un sous-ensemble  $U$  de  $S$ , optimal au sens de la stationnarité, la solution étant un couple canonique, normé,  $(c, (1/\sqrt{n})P_L(T)(c))$  de  $L(U) \times L(T)$ ,  $c$  appartenant à  $C(S)$  et vérifiant :

$$P_{C(S)} \circ P_{L(T)}(c) = nc.$$

#### 3.2. L'ALGORITHME ACCE

L'optimalité étant la stationnarité sur  $C(S)$  et  $L(T)$ , l'algorithme ACCE est développé dans une approche ADOPT, la condition nécessaire étant acquise (proposition II.8, section 3.4).

##### 3.2.1. CONDITIONS ADOPT

$$E = \{U, \text{sous-ensemble de } S\};$$

$\Lambda = \{U \text{ de } E \text{ tel que le premier couple canonique de } L(U) \times L(T), (c, (1/\sqrt{n})P_L(\tau)(c)), \text{ appartienne à } C(S) \times L(T)\}.$

**Critère:** maximiser le carré du cosinus associé au premier couple canonique de  $L(U) \times L(T)$  pour  $U \in \Lambda$ .

**Test:**  $U$  est dit optimal si et seulement si

$$\langle S_1, c \rangle_\tau \leq \pi \langle S_1, c \rangle$$

pour tout  $S_1 \in S \setminus U$ ,  $(c, (1/\sqrt{n})P_L(\tau)(c))$  étant le premier couple canonique de  $L(U) \times L(T)$  (proposition II.9, section 3.4).

**Procédure:** Partant d'une partie admissible, non optimale  $U$  de  $S$ , on note  $(c^0, (1/\sqrt{n})P_L(\tau)(c^0))$  le premier couple canonique de  $L(U) \times L(T)$ ; il existe  $S_1 \in S \setminus U$  tel que

$$\langle S_1, c^0 \rangle_\tau > \pi \langle S_1, c^0 \rangle$$

Soit  $V = UU\{S_1\}$ ; si  $V$  est admissible, elle est «meilleure» que  $U$  sinon décrivons le procédé suivant:

Notons  $(y^1, (1/\sqrt{n_1})P_L(\tau)(y^1))$  le premier couple canonique de  $L(V) \times L(T)$ , obtenu par ALS( $V, T, c^0$ ). Soit  $c^1$  pris sur le segment  $[c^0, y^1]$  et appartenant à  $C(S)$ . D'après les propositions II.6, II.7 (section 3.4), on a

$$n_1 > \cos^2(c^1, P_L(\tau)(c^1)) > \pi;$$

soit  $L(V)$  le plus petit espace vectoriel contenant  $c^1$ ; si  $V$  est admissible alors la procédure est terminée sinon, on réitère le procédé en remplaçant  $y^1$  par  $y^2$  avec  $(y^2, (1/\sqrt{n_2})P_L(\tau)(y^2))$  le premier couple canonique de  $L(V) \times L(T)$ , obtenu par ALS( $V, T, c^1$ ); on prend  $c^2$  sur le segment  $[c^1, y^2]$  et appartenant à  $C(S)$ .

On a  $n_2 > \cos^2(c^2, P_L(\tau)(c^2)) > \cos^2(c^1, P_L(\tau)(c^1)) > \pi$ .

La suite  $\{\cos^2(c^i, P_L(\tau)(c^i))\}$  est croissante et majorée; comme  $\text{CARD}(E)$  est fini, la suite  $\{(y^i, (1/\sqrt{n_i})P_L(\tau)(y^i))\}$  converge, en un nombre fini d'étapes, vers un premier couple canonique admissible.

## 3.2.2.DESCRPTION

## 1.Initialisation

$$\pi = \cos^2 (S_k, P_L(\tau)(S_k)) = \text{MAX}\{\cos^2 (S_j, P_L(\tau)(S_j))\}, j=1, \dots, p\};$$

$$U = \{S_k\}, c = \sum\{\beta_j S_j, S_j \in U\} = S_k, z=c; V=U.$$

2. Si  $\langle S_j, P_L(\tau)(c) \rangle \leq \pi \langle S_j, c \rangle$  pour  $S_j \in S \setminus U$  aller à 5

sinon, soit  $S_1$  réalisant  $\text{MAX}\{\langle S_j, P_L(\tau)(c) \rangle - \pi \langle S_j, c \rangle, S_j \in S \setminus U\}$ .

3. Soit  $(c^1, (1/\sqrt{n_1})P_L(\tau)(c^1))$  déterminé par ALS(V+i, T, z),

$$\text{avec } c^1 = \sum\{\alpha_j S_j, S_j \in V\} + \alpha_1 S_1.$$

Si le couple est admissible alors  $c=c^1, \beta=\alpha, n=n_1, z=c$  et  $U=V+i$ ,

aller à 2 sinon

$$4. \mu = \text{MIN}\{\beta_j / (\beta_j - \alpha_j), \alpha_j < 0\}; V = V \setminus \{S_j, \beta_j + \mu(\alpha_j - \beta_j) = 0\};$$

$$z = c + \mu(c^1 - c); \beta_j = \beta_j + \mu(\alpha_j - \beta_j), j=1, \dots, p; \text{aller à 3.}$$

5. Le couple canonique  $(c, (1/\sqrt{n})P_L(\tau)(c))$  déterminé par

ALS(U, T, z) est admissible et optimal.

## Remarques

1. Une caractéristique importante de l'algorithme ACCE est la possibilité d'intégration de l'opérateur SWEEP (type I, voir chapitre I, section 2.1.3.1).

2. Le sous-ensemble optimal (ensemble des solutions stationnaires) n'est, bien entendu, pas un singleton. Il n'a pas été établi que la solution obtenue est la meilleure, mais ACCE ne dépend pas d'un vecteur initial, contrairement à l'algorithme ALS; cela n'exclut donc pas, pour ce dernier, la convergence vers un optimum local dépendant de ce vecteur initial.

3. L'algorithme proposé peut être modifié de façon à prendre en compte le cas de coefficients libres. Soit  $\Phi$  le sous-ensemble des variables de  $S$  correspondant à ces coefficients, l'admissibilité étant la positivité, au sens large, des coefficients de  $S \setminus \Phi$  :

#### 1. Initialisation

Si  $\Phi = \emptyset$  alors  $\pi = \cos^2(S_k, P_L(\tau)(S_k)) = \text{MAX}\{\cos^2(S_j, P_L(\tau)(S_j)), j \geq 1\}$  et  $U = \{S_k\}$  sinon  $U = \Phi$ . Soit  $(c, (1/\sqrt{\pi})P_L(\tau)(c))$  premier couple canonique de  $L(U) \times L(T)$ ,  $c = \sum\{\beta_j S_j, S_j \in U\}$ ,  $z = c$ ,  $V = U$ . Si  $S \setminus \Phi = \emptyset$  aller à 5 sinon

2. Si  $\langle S_j, P_L(\tau)(c) \rangle \leq \pi \langle S_j, c \rangle$  pour  $S_j \in S \setminus U$  aller à 5 sinon, soit  $S_1$  réalisant  $\text{MAX}\{\langle S_j, P_L(\tau)(c) \rangle - \pi \langle S_j, c \rangle, S_j \in S \setminus U\}$ .

3. Soit  $(c^1, (1/\sqrt{\pi_1})P_L(\tau)(c^1))$  déterminé par  $\text{ALS}(V+i, T, z)$ , avec  $c^1 = \sum\{\alpha_j S_j, S_j \in V\} + \alpha_1 S_1$ ;

Si le couple est admissible alors  $c = c^1$ ,  $\pi = \pi_1$ ,  $z = c$  et  $U = V+i$ , aller à 2 sinon

$$4. \mu = \text{MIN}\{\beta_j / (\beta_j - \alpha_j), j \in S \setminus \Phi, \alpha_j < 0\};$$

$$V = V \setminus \{S_j, \beta_j + \mu(\alpha_j - \beta_j) = 0\};$$

$$z = c + \mu(c^1 - c); \beta_j = \beta_j + \mu(\alpha_j - \beta_j), j = 1, \dots, p; \text{aller à 3.}$$

5. Le couple canonique  $(c, (1/\sqrt{\pi})P_L(\tau)(c))$  est admissible et optimal.

### 3.3. ETUDE NUMERIQUE

On considère les données de Berce et Wilbaux, décrites par P. Dagnelie (1975):

$Y$  = rendement moyen de froment d'hiver (quintaux par ha)

$T_1$  = précipitation des mois de novembre et décembre (mm)

$T_2$  = précipitation du mois de juillet (mm)

$T_3$  = température moyenne de juillet (degrés centigrades)

$T_4$  = radiation du mois de juillet (ml d'alcool mesurés à l'actinomètre de Bellani).

Nous nous intéressons uniquement à l'ordre des valeurs prises par  $Y$ . Soit  $Z$  la variable ordinale associée à  $Y$  (structure ordinale identique).

Le cône polyédrique convexe des codages de  $Z$  (voir chapitre III, section 1.2) est engendré par les vecteurs  $S_j, j=1, \dots, 10$  :

$$S_{j,i} = 1 \text{ si } Z(i) \geq j, S_{j,i} = 0 \text{ sinon.}$$

Faisons remarquer que  $Z = \Sigma\{S_j, j=1, \dots, 10\}$ .

Nous effectuons l'analyse canonique du cône polyédrique convexe  $L(S_1)+C(S_2, S_3, \dots, S_{10})$  et de l'espace vectoriel  $L(T_1, T_2, T_3, T_4)$  par l'algorithme ACCE.

#### Données de Berce & Wilbaux

Années	Y	$T_1$	$T_2$	$T_3$	$T_4$
1920-21	28.37	87.9	1.0	19.6	1661
1921-22	23.77	89.9	90.1	15.2	968
1922-23	26.04	153.0	56.6	19.7	1353
1923-24	25.74	132.1	91.0	17.0	1293
1924-25	26.68	88.8	93.7	18.3	1153
1925-26	24.29	220.9	106.9	17.8	1286
1926-27	28.00	117.7	65.5	17.8	1104
1927-28	28.37	109.0	41.8	18.3	1574
1928-29	24.96	156.1	57.4	17.8	1222
1929-30	21.66	181.5	140.6	16.8	902
1930-31	24.37	181.4	74.3	17.0	1150

Nous notons  $c = \sum \beta_j S_j$  ,  $P_L(\tau)c = \sum \alpha_j T_j$

$$\pi = \cos^2(c, P_L(\tau)c), \Phi = \{S_i\}$$

et  $\Omega$  le vecteur d'optimalité :

$$\Omega_j = \langle S_j, P_L(\tau)(c) - \pi c \rangle, j=1, \dots, 10.$$

### 3.3.1. ETAPES DE L'ALGORITHME ACCE

#### Initialisation

$$U = \emptyset; \pi = 0.997765;$$

$$\beta' = (1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0);$$

$$\alpha' = (-0.00031 \ 0.00168 \ 0.04674 \ 0.00007);$$

$$\Omega' = (0.00000 \ -0.02994 \ 0.06710 \ 0.03293 \ 0.08816 \ 0.12169 \\ 0.1238 \ 0.0599 \ -0.0070 \ 0.0091).$$

Etape 1: on entre la variable 7

$$U = \{S_1, S_7\}; \pi = 0.9990766;$$

$$\beta' = (0.91889 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0.11310 \ 0 \ 0 \ 0);$$

$$\alpha' = (-0.00091 \ 0.00083 \ 0.05972 \ -0.00002)$$

$$\Omega' = (0.00000 \ -0.01701 \ 0.02138 \ 0.01687 \ 0.04824 \ 0.02435 \\ 0.00000 \ -0.02369 \ -0.05744 \ -0.01263).$$

Etape 2: on entre la variable 5

$$U = \{S_1, S_5, S_7\}; \pi = 0.9992869;$$

$$\beta' = (0.87491 \ 0 \ 0 \ 0 \ 0.04423 \ 0 \ 0.09657 \ 0 \ 0 \ 0);$$



$\alpha' = (-0.00102 \quad 0.00066 \quad 0.05973 \quad -0.00002)$   
 $\alpha' = (0.00000 \quad -0.02073 \quad -0.00446 \quad -0.01554 \quad 0.00000 \quad -0.01280$   
 $0.00000 \quad -0.01846 \quad -0.04671 \quad 0.00196).$

Etape 3: on entre la variable 10

$U = \{S_1, S_5, S_7, S_{10}\}; \pi = 0.9992882;$   
 $\beta' = (0.84561 \quad 0 \quad 0 \quad 0 \quad 0.04481 \quad 0 \quad 0.09174 \quad 0 \quad 0 \quad 0.00650);$   
 $\alpha' = (-0.0010 \quad 0.00064 \quad 0.05715 \quad -0.00001)$   
 $\alpha' = (0.00000 \quad -0.01775 \quad -0.00282 \quad -0.01451 \quad 0.00000 \quad -0.00043$   
 $0.00000 \quad -0.01907 \quad -0.04704 \quad 0.00000).$

L'optimum est atteint.

#### Remarque

Il est à noter que l'algorithme ALS (projections alternées entre le cône  $L(S_1)+C(S_2, S_3, \dots, S_{10})$  et l'espace vectoriel  $L(T_1, T_2, T_3, T_4)$  de vecteur initial  $Y$  converge vers la même solution en 110 itérations.

#### 3.3.2. ANALYSE DES RESULTATS

Nous effectuons une étude de la validité du modèle (explication de la variable «rendement» à partir des valeurs  $c_i, i=1, \dots, 11$ , par les variables  $T_j, j=1, \dots, 4$ ).

#### Validité du modèle

Source de var.	D.L.	Som. des car.	Carrés moy.	F.Global
Expliquée	4	9.2880	2.3220	2456.7384
Résiduelle	7	0.0066	0.0009	
Totale	11	9.2946		

## Variables dans la régression

Variable	Coefficient	Ecart-type	T Student	F partiel
T <sub>1</sub>	-0.000997	0.00026	-3.79697	14.41697
T <sub>2</sub>	0.000636	0.00042	1.51564	2.29718
T <sub>3</sub>	0.057145	0.00699	8.17319	66.80110
T <sub>4</sub>	-0.000006	0.00008	-0.07702	0.00593

Nous concluons à l'influence prépondérante du facteur «température en juillet» sur l'ordre des valeurs prises par la variable «rendement» et dans une moindre mesure du facteur «précipitation en novembre/décembre».

Nous terminons cette analyse par ces remarques, justifiant le modèle et les résultats:

1. En 1921/22 et 1929/30: températures les plus faibles et rendements les plus faibles mais l'année la plus mauvaise est 1929/30, les précipitations ayant été trop fortes.

2. En 1925/26 et 1926/27 d'une part, 1924/25 et 1927/28 d'autre part: températures égales mais la différence de rendement enregistrée, provient de l'écart important entre les précipitations respectives en novembre/décembre et amène à conclure, en toute logique, que les niveaux de précipitations ne doivent être ni trop élevés ni trop faibles, le seuil optimal se situant autour de 110 mm environ.

## 3.3.3. ETUDE COMPARATIVE REGCAZ/MCCP

Le problème est maintenant considéré comme une régression multiple (variable "RENDEMENT" numérique) sous contrainte de positivité des coefficients:

$$Y = \sum \{ \alpha_i T_i, \alpha_i \geq 0, i=0, \dots, 4 \} + \varepsilon$$

en notant T<sub>0</sub> le vecteur unité et T la matrice des vecteurs colonnes T<sub>i</sub>, i=0, ..., 4.

Nous présentons, à titre de comparaison, les résultats obtenus par REGCAZ et M CCP (sections 2.1.3.2 et 2.1.3.3 chapitre I).

programme REGCAZ

Initialisation:  $\alpha$  est le vecteur nul

$(Y - T\alpha)'T = [282.2500 \quad 38386.0590 \quad 20391.4250 \quad 5028.5670$   
 $354275.1700]$

Pivotage n°1: on entre la variable  $T_4$

$\alpha = [0.0000 \quad 0.0000 \quad 0.0000 \quad 0.0000 \quad 0.0202]$   
 $(Y - T\alpha)'T = [5.7066 \quad 709.15528 \quad 1184.6662 \quad 74.7132 \quad 0.0000]$

Pivotage n°2: on entre la variable  $T_2$

$\alpha = [0.0000 \quad 0.0000 \quad 0.0516 \quad 0.0000 \quad 0.0174]$   
 $(Y - T\alpha)'T = [1.6842 \quad -326.5197 \quad 0.0000 \quad 24.4664 \quad 0.0000]$

Pivotage n°3: on entre la variable  $T_3$

$\alpha = [0.0000 \quad 0.0000 \quad -0.00202 \quad 1.2634 \quad 0.0027]$

Pivotage n°4: on retire la variable  $T_2$

$\alpha = [0.0000 \quad 0.0000 \quad 0.0000 \quad 1.2309 \quad 0.0030]$   
 $(Y - T\alpha)'T = [0.5276 \quad -364.9827 \quad -14.8247 \quad 0.0000 \quad 0.0000]$

Pivotage n°5: on entre la variable  $T_0$

$\alpha = [12.9263 \quad 0.0000 \quad 0.0000 \quad 0.3349 \quad 0.0055]$   
 $(Y - T\alpha)'T = [0.0000 \quad -422.4018 \quad -150.0362 \quad 0.0000 \quad 0.0000]$

l'optimum est atteint.

## version M CCP

Initialisation:  $\alpha$  est le vecteur nul

$$(Y - T\alpha)'T = [282.2500 \quad 38386.0590 \quad 20391.4250 \quad 5028.5670 \\ 354275.1700]$$

$$v' = [7242.2784 \quad 6426.0396 \quad 5586.4464 \quad 7258.2434 \quad 7169.0656]$$

Pivotage n°1: on entre la variable  $T_3$

$$\alpha = [0.0000 \quad 0.0000 \quad 0.0000 \quad 1.4434 \quad 0.0000]$$

$$(Y - T\alpha)'T = [0.3536 \quad -452.3077 \quad -171.8709 \quad 0.0000 \quad 922.2907]$$

$$v' = [2.4194 \quad 9.5250 \quad 1.8663 \quad -7258.2434 \quad 2.7888]$$

Pivotage n°2: on entre la variable  $T_4$

$$\alpha = [0.0000 \quad 0.0000 \quad 0.0000 \quad 1.2309 \quad 0.0030]$$

$$(Y - T\alpha)'T = [0.5276 \quad -364.9827 \quad -14.8247 \quad 0.0000 \quad 0.0000]$$

$$v' = [6.8203 \quad 7.1069 \quad 0.0299 \quad -91.9665 \quad -2.7888]$$

Pivotage n°3: on entre la variable  $T_0$

$$\alpha = [12.9263 \quad 0.0000 \quad 0.0000 \quad 0.3349 \quad 0.0055]$$

$$(Y - T\alpha)'T = [0.0000 \quad -422.4018 \quad -150.0362 \quad 0.0000 \quad 0.0000]$$

l'optimum est atteint.

## 3.4. FONDEMENTS THEORIQUES DE L'ALGORITHME ACCE

Nous commençons par énoncer un résultat fondamental. Soient  $V$  une partie de  $S$ ,  $x \in L(V)$  et  $(x^*, y^*)$  premier couple canonique obtenu par ALS( $V, T, x$ ). Notons  $\eta^* = \cos^2(x^*, P_{L(T)}(x^*))$ .

**Proposition II.6**

Sous ces hypothèses, on a :

$$1. \langle x^*, x \rangle_T = \pi^n \langle x^*, x \rangle;$$

2. Le produit scalaire  $\langle x^*, x \rangle$  est positif ou nul.

**Preuve**

1. Découle de la proposition II.1 (section 1.2).

2.  $x^*$  est la limite de la suite  $\{x^n\}$  découlant du processus suivant :

$$x^0 = x; y^n = P_{L(\tau)}(x^n); x^{n+1} = P_{L(v)}(y^n) / \|P_{L(v)}(y^n)\|, \quad n \geq 0.$$

Notons  $\|P_{L(v)}(y^n)\| = \mu_n$  et, pour  $p \leq k \leq m$ :

$$\mu_{p, k, m} = (\mu_p \mu_{p+1} \dots \mu_k) / (\mu_{k+1} \mu_{k+2} \dots \mu_m).$$

On montre que pour tout  $n \geq 1$  :

$$(1) \quad \langle x, x^{2^n-1} \rangle = \mu_{0, n-2, 2^n-2} \langle x^{n-1}, x^{n-1} \rangle_T > 0$$

$$(2) \quad \langle x, x^{2^n} \rangle = \mu_{0, n-1, 2^n-1} \langle x^n, x^n \rangle = \mu_{0, n-1, 2^n-1} > 0.$$

Soit  $X = \{\alpha x^n, n \geq 0, \alpha \in \mathbb{R}\}$  ;  $Y = \{\alpha y^n, n \geq 0, \alpha \in \mathbb{R}\}$  ;

notons  $f$  la forme bilinéaire symétrique, de  $\mathbb{R}^n \times \mathbb{R}^n$  dans  $\mathbb{R}$  telle que  $f(x, y) = \langle x, y \rangle$ . Nous considérons deux opérateurs  $a$  et  $r$ , définis respectivement sur  $(X \times Y) \cup (Y \times X)$  à valeurs dans  $(X \times X) \cup (Y \times Y)$  et sur  $(X \times X) \cup (Y \times Y)$  à valeurs dans  $(X \times Y) \cup (Y \times X)$  telles que  $f[a(\cdot, \cdot)] = f(\cdot, \cdot)$ ;  $f[r(\cdot, \cdot)] = f(\cdot, \cdot)$ ;  $f[ao_r(\cdot, \cdot)] = f(\cdot, \cdot)$  :

$$a(x^p, y^q) = (y^p, y^q) \text{ et } a(y^p, x^q) = (\mu_p x^{p+1}, x^q) \quad \text{si } p < q$$

$$a(x^p, y^p) = a(y^p, x^p) = (y^p, y^p) \quad \text{si } p = q$$

$$a(x^p, y^q) = (x^p, \mu_q x^{q+1}) \text{ et } a(y^p, x^q) = (y^p, y^q) \quad \text{si } p > q$$

$$r(x^p, x^q) = (x^p, (1/\mu_{q-1}) y^{q-1}) \text{ et } r(y^p, y^q) = (y^p, y^q) \quad \text{si } p < q$$

$$r(x^p, x^q) = ((1/\mu_{p-1})y^{p-1}, x^q) \text{ et } r(y^p, y^q) = (x^p, y^q) \quad \text{si } p > q$$

$$r(x^p, x^p) = (x^p, x^p) \quad \text{et } r(y^p, y^p) = (y^p, x^p) \quad \text{si } p=q.$$

Les opérateurs  $a$  et  $r$  servent à exprimer une suite d'égalités de produits scalaires: nous les munissons donc, conventionnellement, des propriétés de symétrie et de bilinéarité. Les définitions énoncées ci-dessus, précisent leur mode d'application afin d'aboutir aux égalités (1) et (2) recherchées.

Notons  $(aor)^n = (aor) \circ (aor) \circ \dots \circ (aor) \quad n \text{ fois}$

et  $E(\alpha) =$  la partie entière d'un réel  $\alpha$ .

1. Nous montrons:

$$(1.1) \quad (aor)^{2n} (x^0, x^{2n}) = \mu_{0, n-1, 2n-1} (x^n, x^n);$$

$$(1.2) \quad (aor)^{2n-2} (x^0, x^{2n-1}) = (\mu_{0, n-2, 2n-2} \mu_{n-1}) (x^{n-1}, x^n).$$

Il suffit de montrer que, pour tout  $k \geq 2$ ,  $p \in \{1, \dots, E(k/2)\}$ , on a

$$(aor)^2 (x^{p-1}, x^{k-(p-1)}) = (\mu_{p-1} / \mu_{k-p}) (x^p, x^{k-p}) :$$

$$(aor)^2 (x^{p-1}, x^{k-(p-1)}) = (1/\mu_{k-p}) aor \circ a (x^{p-1}, y^{k-(p-1)-1}) =$$

$$= (1/\mu_{k-p}) aor (y^{p-1}, y^{k-p}) = (1/\mu_{k-p}) a (y^{p-1}, x^{k-p}) =$$

$$= (\mu_{p-1} / \mu_{k-p}) (x^{p-1+1}, x^{k-p}).$$

2. Nous aboutissons aux égalités (1) et (2) en notant tout d'abord que (1.2) implique

$$ro(aor)^{2n-2} (x^0, x^{2n-1}) = \mu_{0, n-2, 2n-2} (x^{n-1}, y^{n-1});$$

on a, pour tout  $n \geq 1$  :

$$\langle x^0, x^{2n} \rangle = f[(aor)^{2n} (x^0, x^{2n})] = \mu_{0, n-1, 2n-1} \langle x^n, x^n \rangle;$$

$$\langle x^0, x^{2n-1} \rangle = f[ro(aor)^{2n-2} (x^0, x^{2n-1})] = \mu_{0, n-2, 2n-2} \langle x^{n-1}, y^{n-1} \rangle;$$

comme  $\langle x^{n-1}, y^{n-1} \rangle = \langle x^{n-1}, P_{L(\tau)}(x^{n-1}) \rangle = \langle x^{n-1}, x^{n-1} \rangle_{\tau}$  :

$$\langle x^0, x^{2n-1} \rangle = \mu_{0, n-2, 2n-2} \langle x^{n-1}, x^{n-1} \rangle_{\tau}.$$

### Proposition II.7

Soient  $V$  une partie de  $S$ ,  $x \in L(V)$ ;  $(x^*, y^*)$  premier couple canonique de  $L(V) \times L(T)$  et  $\langle x^*, x \rangle \geq 0$ .

Soit  $z \in ]x \ x^* [$ ; on a

$$\pi^* = \cos^2(x^*, P_{L(\tau)}(x^*)) > \cos^2(z, P_{L(\tau)}(z)) \text{ et}$$

$$\cos^2(z, P_{L(\tau)}(z)) > \cos^2(x, P_{L(\tau)}(x)) = \pi.$$

### Preuve

Soit  $\mu \in ]0 \ 1[$ ,  $z = x + \mu(x^* - x)$  :

$$\begin{aligned} \langle z, z \rangle_{\tau} &= (1-\mu)^2 \langle x, x \rangle_{\tau} + \mu^2 \langle x^*, x^* \rangle_{\tau} + 2\mu(1-\mu) \langle x^*, x \rangle_{\tau} \\ &= (1-\mu)^2 \pi \langle x, x \rangle + \mu^2 \pi^* \langle x^*, x^* \rangle + 2\mu(1-\mu) \pi^* \langle x^*, x \rangle \\ &> (1-\mu)^2 \pi \langle x, x \rangle + \mu^2 \pi \langle x^*, x^* \rangle + 2\mu(1-\mu) \pi \langle x^*, x \rangle = \pi \langle z, z \rangle. \end{aligned}$$

### Proposition II.8

Si le couple  $(c, \tau)$  de  $C(S) \times L(T)$  vérifie

$$\cos^2(c, \tau) = \text{MAX}[\cos^2(s, t), s \in C(S), t \in L(T)]$$

alors il existe  $U$  sous-ensemble de  $S$  tel que  $(c, \tau)$  soit le premier couple canonique de  $L(U) \times L(T)$ .

### Preuve

D'après la proposition II.5 (section 1.2),  $(c, \tau)$  est un couple stationnaire de  $C(S) \times L(T)$ . Il existe une partie  $U$  de  $S$ , telle que  $(c, \tau)$  soit stationnaire dans  $L(U) \times L(T)$  :

$$c = \sum \{\alpha_j S_j, \alpha_j > 0, S_j \in U\}.$$

Supposons que  $(c, \tau)$  ne soit pas le premier couple canonique de  $L(U) \times L(T)$ .

Soit  $(c^*, \tau^*)$  premier couple canonique de  $L(U) \times L(T)$ ,  $c^* = \sum \alpha_j^* S_j$  non admissible ( $c^*$  n'appartient pas à  $C(S)$ ) et  $\langle c^*, c \rangle = 0$  ( $c$  et  $c^*$  appartiennent respectivement à deux sous-espaces propres orthogonaux distincts).

Soit  $z \in ]c^*, c[ : \mu \in ]0, 1[$  et  $z = c + \mu(c^* - c)$ ; choisissons

$$\mu = \text{MIN}\{\alpha_j / (\alpha_j - \alpha_j^*), \alpha_j^* < 0\};$$

$z \in C(S)$  et d'après la proposition II.7

$$\cos^2(c^*, \tau^*) > \cos^2(z, P_{L(T)}(z)) > \cos^2(c, \tau)$$

d'où la contradiction.

**CQFD**

Soit  $c$  dans  $C(S)$ ,  $\pi = \cos^2(c, P_{L(T)}(c))$ ; la proposition suivante donne une caractérisation de la stationnarité:

$$\forall c \in C(S) \quad \langle P_{L(T)}(c), c \rangle = \pi c.$$

**Proposition II.9**

Le couple  $(c, (1/\sqrt{\pi})P_{L(T)}(c))$  est stationnaire si et seulement si, pour tout  $z$  dans  $C(S)$  :

$$\langle c, z \rangle_T \leq \pi \langle c, z \rangle$$

$$\text{avec } \langle c, c \rangle_T = \pi \langle c, c \rangle.$$

**Preuve**

1. Soit  $(c, (1/\sqrt{\pi})P_{L(T)}(c))$  stationnaire; pour tout  $z$  dans  $C(S)$ , on a

$$\langle c, z \rangle_T = \langle P_{L(T)}(c), z \rangle \leq \langle P_{C(S)}[P_{L(T)}(c)], z \rangle = \langle \pi c, z \rangle.$$

2. Soit  $c$  dans  $C(S)$  tel que, pour tout  $z$  dans  $C(S)$



$$\langle c, z \rangle_T \leq \pi \langle c, z \rangle$$

avec  $\pi = \langle c, c \rangle_T / \langle c, c \rangle$  ; on a donc :

$$\langle P_{L(T)}(c), z \rangle = \langle c, z \rangle_T \leq \langle \pi c, z \rangle$$

et  $\langle P_{L(T)}(c), c \rangle = \langle c, c \rangle_T = \langle \pi c, c \rangle$

conditions vérifiées par  $\pi c$ , caractéristiques de la projection de  $P_{L(T)}(c)$  sur  $C(S)$  d'où :

$$P_{C(S)}[P_{L(T)}(c)] = \pi c.$$

#### Remarque

Il existe  $U$  dans  $S$  tel que  $(c, (1/\sqrt{\pi})P_{L(T)}(c))$  couple canonique de  $L(U) \times L(T)$  ; la caractérisation devient :

$$\text{pour tout } S_1 \in U : \langle c, S_1 \rangle_T = \pi \langle c, S_1 \rangle$$

$$\text{pour tout } S_1 \in S \setminus U : \langle c, S_1 \rangle_T \leq \pi \langle c, S_1 \rangle.$$

Soit  $U$  dans  $S$  tel que  $(c, (1/\sqrt{\pi})P_{L(T)}(c))$  premier couple canonique de  $L(U) \times L(T)$ , admissible ( $c \in C(S)$ ) non stationnaire sur  $C(S) \times L(T)$ . Il existe  $S_1 \in S \setminus U$  tel que

$$\langle c, S_1 \rangle_T > \pi \langle c, S_1 \rangle.$$

Soit  $(c^1, (1/\sqrt{\pi^1})P_{L(T)}(c^1))$  premier couple canonique de  $L(V+i) \times L(T)$ ,  $V$  étant une partie de  $U$ ; notons :

$$c^1 = \sum \{\alpha_j S_j, S_j \in V\} + \alpha_1 S_1.$$

#### Proposition II.10

Si  $\langle c^1, c \rangle_T > \pi \langle c^1, c \rangle$  alors  $\alpha_1 > 0$ .

#### Preuve

$$\langle c^1, c \rangle_T = \langle \sum \{\alpha_j S_j, S_j \in V\} + \alpha_1 S_1, c \rangle_T$$

$$= \langle \sum \{\alpha_j S_j, S_j \in V\}, c \rangle_T + \alpha_1 \langle S_1, c \rangle_T$$

$$= \pi \langle \sum \{ \alpha_j S_j, S_j \in V \}, c \rangle + \alpha_1 \langle S_1, c \rangle_T;$$

comme  $\langle c^1, c \rangle_T > \pi \langle c^1, c \rangle = \pi [ \langle \sum \{ \alpha_j S_j, S_j \in V \}, c \rangle + \alpha_1 \langle S_1, c \rangle ]$

on a  $\alpha_1 [ \langle S_1, c \rangle_T - \pi \langle S_1, c \rangle ] > 0$ .

#### Corollaire II.10

Si  $V = U$  et  $(c^1, (1/\sqrt{\pi^1}) P_{L(T)}(c^1))$  obtenu par ALS( $U+i, T, c$ ):

alors  $\langle c^1, c \rangle_T = \pi^1 \langle c^1, c \rangle > \pi \langle c^1, c \rangle$  et  $\alpha_1 > 0$ .

#### Preuve

$\pi^1 > \pi$  (U contenu dans  $U+i$ )

et  $\langle c^1, c \rangle > 0$  (proposition II.6).

#### Proposition II.11

Soit  $S_j$  un vecteur de  $S, \pi_j = \cos^2(S_j, P_{L(T)}(S_j))$ .  
Le premier couple canonique de  $L(\{S_j\}) \times L(T)$  est  
 $(S_j, (1/\sqrt{\pi_j}) P_{L(T)}(S_j))$  et il est admissible.

#### Preuve

Notons  $z = P_{L(T)}(S_j); P_{L(\{S_j\})}(z) = \alpha S_j;$

comme  $\langle z - \alpha S_j, S_j \rangle = 0$ , on a  $\alpha = \langle z, S_j \rangle / \langle S_j, S_j \rangle;$

du fait  $\langle S_j - z, z \rangle = 0$ , on a  $\langle z, S_j \rangle = \langle z, z \rangle;$

d'où  $\alpha = \pi_j > 0$ .

On a  $P_{L(\{S_j\})} \circ P_{L(T)}(S_j) = P_{L(\{S_j\})}(z) = \pi_j S_j$ .

Si on suppose  $S_j$  normé alors

$$\|z\|^2 = \pi_j$$

et le couple canonique  $(S_j, (1/\sqrt{\pi_j}) P_{L(T)}(S_j))$  est normé.

**Chapitre III**

**CODAGE DE VARIABLE ORDINALE**

**ET TAU DE KENDALL**

### III. CODAGE DE VARIABLE ORDINALE ET TAU DE KENDALL

La notion de codage est essentielle en analyse de données ordinales et un indicateur de la corrélation des rangs de deux ordres totaux, communément utilisé, est le tau de M.G.Kendall (1962).

Nous proposons à la section 1, un ensemble de notations et une étude générale de la notion de codage.

La section 2 traite des notions d'admissibilité et d'optimalité dans le problème de la recherche du codage d'une variable ordinaire approchant «au mieux» le vecteur des valeurs observées d'une variable numérique.

La section 3 est consacrée à une analyse du tau de Kendall. Des inégalités reliant les taus de trois préordres seront établies. L'étude des cas particuliers où ces relations deviennent des égalités, nous a amené à définir la notion d'«agrégation logique» de deux préordres.

Reliant codage de variable ordinaire et tau de Kendall, nous effectuons à la section 4 une étude dont deux résultats découleront :

-Une formule simple du tau de Kendall entre un préordre et un, quelconque, de ses codages.

-Parmi les codages admissibles, le codage optimal (voir section 2.2) est aussi le mieux corrélé, au sens du tau de Kendall, avec le préordre.

## 1. PRELIMINAIRES

Nous commençons par fixer des notations et définitions. Nous présentons ensuite la notion générale de codage de variable qualitative, en précisant la structure de l'ensemble des codages, relatif à chacun des types nominal et ordinal.

### 1.1. NOTATIONS

Soit  $A$  un ensemble de  $n$  individus et  $y$  une variable numérique observée sur  $A$ .

Soit  $Z_1, Z_2, \dots, Z_k$ ,  $k$  vecteurs de  $R^n$  et  $J$  un sous-ensemble de  $\{2, \dots, k\}$ :

$$J = \{j_1, j_2, \dots, j_r\}$$

les  $j_i$  étant en ordre croissant.

Posons  $j_0 = 1$  et  $j_{r+1} = k+1$ .

$J$  induit une partition de  $\{1, 2, \dots, k\}$  définie comme suit:

$$B_q = \{j_q, j_q+1, \dots, j_{q+1}-1\}, \quad q=0, \dots, r.$$

$L^J$  désigne l'espace vectoriel  $L(Z_{j_0}, Z_{j_1}, \dots, Z_{j_r})$ .

On note  $My(B_q) = \sum \{y_i 1_{B_q} \circ X(i), i=1, \dots, n\} / \text{CARD}\{i | X(i) \in B_q\}$

la moyenne de  $y$  restreinte aux indices  $i$  tels que  $X(i) \in B_q$  avec  $1_{B_q}$  la désignation classique de la variable indicatrice de l'ensemble  $B_q$ :

$$1_{B_q}(j) = 1 \text{ si } j \in B_q \text{ et } 1_{B_q}(j) = 0 \text{ sinon.}$$

On note  $y^J = \sum \{My(B_q) 1_{B_q} \circ X, q=0, \dots, r\}$ .

Quelques notations complémentaires seront nécessitées par certains développements ultérieurs et introduites au fur et à mesure.

## 1.2. CODAGE D'UNE VARIABLE QUALITATIVE

Soit  $X$  une variable observée sur  $A$  et prenant ses valeurs dans l'ensemble  $M$  des  $k$  premiers nombres entiers naturels.

Notons  $X_j$  la variable indicatrice de  $X$  :

$$X_j(i) = 1 \text{ si } X(i) = j \text{ et } X_j(i) = 0 \text{ sinon.}$$

Pour toute application  $C$  de  $M$  sur  $R$ , on peut écrire

$$CoX = C(1)X_1 + C(2)X_2 + \dots + C(k)X_k$$

$C$  est appelée codage des modalités de  $X$  et  $CoX$  est dit codage de  $X$ .

Nous noterons  $COD(X)$  l'ensemble des codages de  $X$ .

1. Si  $X$  est nominale, l'ensemble  $COD(X)$  des codages de  $X$  est l'espace vectoriel  $L(X_1, X_2, \dots, X_k)$ .

2. Si  $X$  est ordinale,  $M$  est muni de la structure d'ordre naturel.

L'application  $C$  doit satisfaire à

$$i < j \text{ équivalent à } C(i) \leq C(j), \text{ pour tout } i, j \text{ de } M.$$

La contrainte sur l'ordre des valeurs prises par  $C$  nous permet d'écrire

$$C(j) = c_1 + c_2 + \dots + c_j, \text{ avec } c_2, c_3, \dots, c_k \text{ non négatifs.}$$

Finalement on obtient  $CoX = c_1 Z_1 + c_2 Z_2 + \dots + c_k Z_k$ , avec

$$Z_j = X_j + X_{j+1} + \dots + X_k.$$

L'ensemble  $COD(X)$  des codages de  $X$  est le cône polyédrique convexe  $L(Z_1) + C(Z_2, \dots, Z_k)$ .

Faisons remarquer que  $Z_1$  est le vecteur de  $\mathbb{R}^n$  dont toutes les composantes sont égales à 1.

**Remarques**

1. Rechercher un codage centré de  $X$  revient à rechercher  $CoX$  dans  $COD(X) \cap Z_1^\perp$ .

Soit  $S$  la sphère de  $\mathbb{R}^n$  centrée en 0, de rayon  $\sqrt{n}$ .

Rechercher un codage centré-réduit de  $X$  revient à rechercher  $CoX$  dans  $COD(X) \cap Z_1^\perp \cap S$ .

2. La notion de codage peut être étendue au cas d'une variable numérique mais deux choix sont possibles:

$$COD(X) = L(X) \text{ ou } COD(X) = \{X_{\text{centré-réduit}}\}.$$

## 2. CODAGE OPTIMAL D'UNE VARIABLE ORDINALE

La recherche d'un codage d'une variable ordinale approchant «au mieux» le vecteur des valeurs observées d'une variable numérique amène à la notion de codage monotone optimal. Nous en effectuons une analyse complète.

### 2.1. CODAGE ADMISSIBLE

Soit une variable ordinale  $X$  observée sur  $A$  et prenant ses valeurs dans l'ensemble  $M$  des  $k$  premiers nombres entiers naturels.

L'ensemble des codages de  $X$  est:

$$\text{COD}(X) = L(Z_1) + C(Z_2, \dots, Z_k) \text{ (section 1.2);}$$

Soit  $J$  un sous-ensemble de  $M \setminus \{1\}$ :

$$J = \{j_1, j_2, \dots, j_r\}$$

les  $j_i$  étant en ordre croissant.

En reprenant les notations et conventions de la section 1.1,  $J$  induit la partition de  $M$ :

$$B_q = \{j_q, j_q+1, \dots, j_{q+1}-1\}, \quad q=0, \dots, r.$$

On a  $L^J = L(Z_{j_0}, Z_{j_1}, \dots, Z_{j_r})$

et  $y^J = \sum \{My(B_q) 1_{B_q} \circ X, q=0, \dots, r\}$

comme  $1_{B_q} \circ X = Z_{j_q} - Z_{j_{q+1}}, q=0, \dots, r-1$  et  $1_{B_r} \circ X = Z_{j_r}$

on a  $y^J = My(B_0)Z_1 + \sum \{[My(B_q) - My(B_{q-1})]Z_{j_q}, q=1, \dots, r\}$  (1)

A noter aussi que l'on a:

$$\|y^J\|^2 = \sum \{[My(B_q)]^2 \text{CARD}\{X \in B_q\}, q=0, \dots, r\} \quad (2)$$

#### Proposition III.1

La projection orthogonale de  $y$  sur  $L^J$  est  $y^J$ .



Preuve

Il suffit de montrer que

$$\begin{aligned} \langle y^J, Z_{J^c} \rangle &= \langle y, Z_{J^c} \rangle = \sum \{ y_i \mid X(i) \geq_{J^c} \}, \quad q=0, \dots, r; \\ \langle y^J, Z_{J^c} \rangle &= \sum \{ Z_{J^c}(i) \sum \{ My(B_p) 1_{B_p} \circ X(i), p=0, \dots, r \}, i=1, \dots, n \} \\ &= \sum \{ My(B_p) \sum \{ 1_{B_p} \circ X(i) Z_{J^c}(i), i=1, \dots, n \}, p=0, \dots, r \} \\ &= \sum \{ My(B_p) \sum \{ 1_{B_p} \circ X(i), i=1, \dots, n \}, p \geq q \} \end{aligned}$$

CQFD

Si  $J$  est un sous-ensemble de  $M \setminus \{1\}$  tel que

$$My(B_q) \geq My(B_{q-1}), \quad q=1, \dots, r$$

c'est à dire  $y^J \in \text{COD}(X)$  d'après (1) et le résultat fondamental énoncé dans la proposition II.2, section 1.2,  $y^J$  est dit codage admissible de  $X$ . Partant d'un sous-ensemble  $J$ , soit  $j_m \in J$  et

$$k_q = \text{CARD}\{X \in B_q\}, \quad q=0, \dots, r, \quad J^m = J \setminus \{j_m\}.$$

Soit  $\{B^{m,q}, q=0, \dots, r-1\}$  la partition induite par  $J^m$  :

$$B^{m,q} = B_q \text{ si } q < m-1, \quad B^{m,m-1} = B_{m-1} \cup B_m, \quad B^{m,q} = B_{q+1}, \quad q \in \{m, \dots, r-1\}.$$

Un évaluation de la variation de  $\|y - y^J\|^2$  est apportée par la proposition suivante:

Proposition III.2

En passant de  $J$  à  $J^m$ , la distance  $\|y - y^J\|^2$  augmente de la quantité

$$d_{J^m} = [My(B_{m-1}) - My(B_m)]^2 [(k_{m-1} k_m / (k_{m-1} + k_m))]$$

Preuve

$$\begin{aligned} \text{On a } \|y^{J^m}\|^2 &= \sum \{ [My(B_q)]^2 k_q, q=0, \dots, m-2 \} + \\ &+ [My(B_{m-1} \cup B_m)]^2 (k_{m-1} + k_m) + \sum \{ [My(B_q)]^2 k_q, q=m+1, \dots, r \}. \end{aligned}$$

Comme  $\|y - y^J\|^2 - \|y - y^J\|^2 = \|y^J\|^2 - \|y^J\|^2 = d_{J_n}$ , on a, d'après (2):

$$d_{J_n} = k_{n-1} [\text{My}(B_{n-1})]^2 + k_n [\text{My}(B_n)]^2 - (k_{n-1} + k_n) [\text{My}(B_{n-1} \cup B_n)]^2.$$

Posons  $\Sigma_1 = k_{n-1} \text{My}(B_{n-1})$ ,  $\Sigma_0 = k_n \text{My}(B_n)$ ;

on a  $d_{J_n} = (\Sigma_1^2/k_{n-1}) + (\Sigma_0^2/k_n) - (\Sigma_0 + \Sigma_1)^2 / (k_{n-1} + k_n)$

$$= [(k_{n-1} + k_n)(k_n \Sigma_1^2 + k_{n-1} \Sigma_0^2) - k_{n-1} k_n (\Sigma_0 + \Sigma_1)^2] / [(k_{n-1} + k_n) k_{n-1} k_n]$$

$$= (k_n^2 \Sigma_1^2 + k_{n-1}^2 \Sigma_0^2 - 2k_{n-1} k_n \Sigma_0 \Sigma_1) / [(k_{n-1} + k_n) k_{n-1} k_n]$$

$$= (k_{n-1} k_n)^2 (\text{My}(B_{n-1}) - \text{My}(B_n))^2 / [(k_{n-1} + k_n) k_{n-1} k_n]$$

$$= (\text{My}(B_{n-1}) - \text{My}(B_n))^2 [(k_{n-1} k_n) / (k_{n-1} + k_n)].$$

**CQFD**

Soit  $J$  un sous-ensemble de  $M \setminus \{1\}$  et  $y^J$  un codage admissible. S'il existe  $q$  tel que  $\text{My}(B_q) = \text{My}(B_{q-1})$ , on retire  $j_q$  de  $J$ : d'après le résultat établi ci-dessus, la distance  $\|y - y^J\|$  n'augmente pas.

#### Définition

$J$  est dit minimal si et seulement si, pour tout  $q \in \{1, \dots, r\}$ ,  $\text{My}(B_q) > \text{My}(B_{q-1})$ .

### 2.2. CARACTERISATION DU CODAGE OPTIMAL

Il existe un sous-ensemble  $J$  optimal au sens où  $y^J$  est la projection orthogonale de  $y$  sur  $\text{COD}(X)$ .

Etant donné  $J$  minimal, une condition nécessaire et suffisante d'optimalité est que

$$\langle y - y^J, Z_j \rangle \leq 0, \text{ pour tout } j \in M \setminus J.$$

On se propose de donner une caractérisation de la projection orthogonale de  $y$  sur  $\text{COD}(X)$ .

Notons, pour  $j \in (M \setminus J) \cap B_q$  :

$$My(B_q, j^+) = \sum \{y_i \mid_{B_q} \circ X(i), X(i) \geq j, i=1, \dots, n\} / \text{CARD}\{X \in B_q, X \geq j\}$$

$$My(B_q, j^-) = \sum \{y_i \mid_{B_q} \circ X(i), X(i) < j, i=1, \dots, n\} / \text{CARD}\{X \in B_q, X < j\}.$$

Un résultat préliminaire simple est contenu dans la proposition suivante.

**Proposition III.3**

Soit  $q \in \{0, \dots, r\}$  et  $j \in (M \setminus J) \cap B_q$  :

$My(B_q, j^+) \leq My(B_q)$  équivalent à  $My(B_q, j^+) \leq My(B_q, j^-)$

**Preuve**

Posons  $k^+ = \text{CARD}\{X \in B_q, X \geq j\}$ ,  $\Sigma^+ = k^+ My(B_q, j^+)$

$k^- = \text{CARD}\{X \in B_q, X < j\}$ ,  $\Sigma^- = k^- My(B_q, j^-)$ ; on a

$\Sigma^+ / k^+ \leq (\Sigma^+ + \Sigma^-) / (k^+ + k^-)$  équivalent à  $\Sigma^+ (k^+ + k^-) \leq k^+ (\Sigma^+ + \Sigma^-)$

équivalent à  $\Sigma^+ k^- \leq k^+ \Sigma^-$  équivalent à  $\Sigma^+ / k^+ \leq \Sigma^- / k^-$ .

**Proposition III.4**

$y^j$  est la projection orthogonale de  $y$  sur  $\text{COD}(X)$  si et seulement si  $J$  est minimal et

$My(B_q, j^+) \leq My(B_q)$ , pour tout  $j \in (M \setminus J) \cap B_q, q=1, \dots, r$ .

**Preuve**

Soit  $q \in \{1, \dots, r\}$  et  $j \in (M \setminus J) \cap B_q$ .

On a  $\langle y, Z_j \rangle = \sum \{y_i \mid_{X(i) \geq j}\}$

$$= \sum \{y_i \mid_{X(i) \in B_q, X(i) \geq j}\} + \sum \{y_i \mid_{X(i) \geq j_{q+1}}\}.$$

Par ailleurs

$$\begin{aligned} \langle y^J, Z_j \rangle &= \sum \{ Z_j(i) \sum \{ My(B_p) 1_{B_p} \circ X(i), p=0, \dots, r \}, i=1, \dots, n \} \\ &= My(B_q) \sum \{ 1_{B_q} \circ X(i) Z_j(i), i \geq 1 \} + \sum \{ My(B_p) \sum \{ 1_{B_p} \circ X(i), i \geq 1 \}, p \geq q+1 \} \end{aligned}$$

$$\text{D'où} \quad \sum \{ y_i \mid X(i) \in B_q, X(i) \geq j \} \leq My(B_q) \sum \{ 1_{B_q} \circ X(i) Z_j(i), i \geq 1 \}$$

$$\text{et finalement} \quad \sum \{ y_i \mid X(i) \in B_q, X(i) \geq j \} \leq My(B_q) \text{CARD}\{X \in B_q, X \geq j\}.$$

#### Corollaire III.4

Un sous-ensemble  $J$  est optimal si et seulement si  $J$  est minimal et contient tous les sous-ensembles minimaux.

#### Preuve

-Supposons  $J$  optimal. La minimalité de  $J$  est évidemment un condition nécessaire.

Soit  $J' = \{j_1, j_2, \dots, j_{r'}\}$ , minimal non contenu dans  $J$  et  $\{B'_q\}$  la partition induite par  $J'$ .

$J'$  n'étant pas optimal, il existe donc  $q \in \{1, \dots, r'\}$  et  $j \in (M \setminus J') \cap B'_q$  tels que:

$$My(B'_q, j^+) > My(B'_q).$$

D'après la proposition III.3, cela revient à

$$My(B'_q, j^+) > My(B'_q, j^-)$$

et  $J'' = J' \cup \{j\}$  est minimal et meilleur que  $J'$  d'après la proposition III.2.

le sous-ensemble  $J^*$  résultant de la réunion de  $J'$  et de l'ensemble de tous les entiers  $j$  du type précédent est forcément optimal d'où la contradiction.

-La réciproque est évidente.

### 3. LE TAU DE KENDALL

Le Tau de M.G. Kendall (1962), désigné par « $\tau$ », est un indicateur de corrélation ordinale entre deux ordres totaux ( $k=n$ ) d'usage classique.

Nous précisons sa définition et celle d'une variante plus générale (Siskos, 1979), traitant le cas de préordres. Nous avons établi des inégalités (proposition III.6, section 3.2) reliant les taus de Kendall de trois préordres donnés.

Une analyse des cas particuliers où ces relations deviennent des égalités nous a amené à définir la notion d'«agrégation logique» de deux préordres (propositions III.7, III.8, section 3.2.1).

Nous terminons par une interprétation graphique de ces relations et une étude numérique.

#### 3.1. DEFINITION DU TAU DE KENDALL

Soit  $u, v$  deux vecteurs de  $\mathbb{R}^n$ . Soit  $K$  l'application de  $\{1, \dots, n\}^2$  dans  $\{-1, 0, 1\}$  définie comme suit:

$$K(i, j) = 1 \text{ si } (u(i) - u(j))(v(i) - v(j)) > 0$$

$$K(i, j) = -1 \text{ si } (u(i) - u(j))(v(i) - v(j)) < 0$$

$$K(i, j) = 0 \text{ sinon;}$$

$K$  est symétrique; il y a  $n(n-1)$  couples  $(i, j)$  avec  $i \neq j$

$$\text{et } \tau(u, v) = (\sum\{K(i, j); i \neq j\}) / n(n-1)$$

$\tau \in [-1, 1]$ ; notons que

$$\tau(u, v) = 1 : \text{ordres identiques}$$

$$\tau(u, v) = 0 : \text{ordres indépendants}$$

$$\tau(u, v) = -1 : \text{ordres inverses.}$$

### 3.2. UNE VARIANTE RELATIVE AUX PREORDRES

Dans le cas de préordres ( $k \leq n$ ), on définit le coefficient de corrélation des rangs  $\tau'$  (Siskos, 1979): soit  $R_u$  l'application de  $\{1, \dots, n\}^2$  dans  $\{0, 1\}$  définie par

$$R_u(i, j) = 1 \text{ et } R_u(j, i) = 0 \text{ si } u(i) > u(j)$$

$$R_u(i, j) = 1 \text{ et } R_u(j, i) = 1 \text{ si } i \neq j \text{ et } u(i) = u(j)$$

$$R_u(i, i) = 0;$$

Définissons l'application  $R_{uv}$  de  $\{1, \dots, r\} \times \{1, \dots, r\}$  dans  $\{0, 1\}$ :

$$R_{uv}(i, j) = |R_u(i, j) - R_v(i, j)|;$$

Remarquons que  $R_{uv}$  n'est pas symétrique.

Notons  $K'_{uv} = \sum \{R_{uv}(i, j); i \neq j\}$

on a  $\tau'(u, v) = 1 - (2K'_{uv}/n(n-1))$ .

#### Proposition III.5

$\tau'$  est un variante de  $\tau$ , coïncidant avec  $\tau$  dans le cas d'ordres totaux et résultant de la définition suivante de l'application  $K$ :

$$K(i, j) = 1 \text{ si } (u[i] - u[j])(v[i] - v[j]) > 0$$

$$K(i, j) = 1 \text{ si } i \neq j, u[i] = u[j] \text{ et } v[i] = v[j]$$

$$K(i, j) = -1 \text{ si } (u[i] - u[j])(v[i] - v[j]) < 0$$

$$K(i, j) = 0 \text{ sinon;}$$

Preuve

On doit montrer que

$$2K'_{uv} + (\sum \{K(i, j); i \neq j\}) = n(r-1).$$

$$\begin{aligned}
\text{Soit } n_1 &= \text{CARD}\{(i, j), i < j, (u[i]-u[j])(v[i]-v[j]) > 0\} \\
n_2 &= \text{CARD}\{(i, j), i < j, (u[i]-u[j])(v[i]-v[j]) < 0\} \\
n_3 &= \text{CARD}\{(i, j), i < j, u[i]=u[j] \text{ et } v[i]=v[j]\} \\
n_4 &= \text{CARD}\{(i, j), i < j, u[i]=u[j] \text{ et } v[i]>v[j]\} \\
n_5 &= \text{CARD}\{(i, j), i < j, u[i]=u[j] \text{ et } v[i]<v[j]\} \\
n_6 &= \text{CARD}\{(i, j), i < j, u[i]<u[j] \text{ et } v[i]=v[j]\} \\
n_7 &= \text{CARD}\{(i, j), i < j, u[i]>u[j] \text{ et } v[i]=v[j]\}
\end{aligned}$$

on a  $2\sum\{n_k, k=1, \dots, 7\} = n(n-1)$ ;

$$\begin{aligned}
2K'_{uv} &= 2[(0n_1 + 0n_1) + (n_2 + n_2) + 0n_3 + (0n_4 + n_4) + (n_5 + 0n_5) + \\
&\quad + (n_6 + 0n_6) + (0n_7 + n_7)] = 2n_2 + 2(n_4 + n_5 + n_6 + n_7);
\end{aligned}$$

et  $\sum\{K(i, j); i \neq j\} = 2n_1 - 2n_2 + 2n_3 + 0n_4 + 0n_5 + 0n_6 + 0n_7 = 2(n_1 + n_3) - 2n_2$ .

CQFD

L'indicateur  $\tau'$ , de portée plus générale, est retenu comme coefficient de corrélation des rangs, sera dit Tau de Kendall (généralisé) et désigné par  $\tau$ .

Soit  $u$  et  $v$  deux préordres quelconques sur  $A$ ; nous avons établi le résultat suivant:

**Proposition III.6**

Pour tout préordre  $w$  sur  $A$ , on a les relations

$$(1) \quad 1 + \tau(u, v) \geq \tau(u, w) + \tau(w, v)$$

$$(2) \quad 1 - \tau(u, v) \geq |\tau(u, w) - \tau(w, v)|.$$

Preuve

(1): Il suffit de remarquer que, pour tout  $i, j$ :

$$|R_u(i, j) - R_v(i, j)| \leq |R_u(i, j) - R_w(i, j)| + |R_w(i, j) - R_v(i, j)|$$

ce qui permet d'écrire

$$-2K'_{uv}/[n(n-1)] \geq -2(K'_{uw} + K'_{vw})/[n(n-1)]$$

d'où  $2-2K'_{uv}/[n(n-1)] \geq 2-2(K'_{uw} + K'_{vw})/[n(n-1)]$

puis

$$1 + [1 - 2(K'_{uv}/(n^2 - n))] \geq [1 - 2(K'_{uw}/(n^2 - n))] + [1 - 2(K'_{vw}/(n^2 - n))].$$

CQFD

(2) est équivalente à

$$1 + \tau(u, w) \geq \tau(u, v) + \tau(v, w) \text{ et } 1 + \tau(v, w) \geq \tau(v, u) + \tau(u, w).$$

Conséquences

1. Si  $\tau(u, v) = 0$  alors  $\tau(u, w) + \tau(w, v) \leq 1$

2. Si  $\tau(u, v) = 1$  alors  $\tau(u, w) = \tau(v, w)$

3. Si  $\tau(u, v) = -1$  alors  $\tau(u, w) + \tau(w, v) \leq 0$

4. Si  $\tau(u, w) = \tau(w, v) = 1$  alors  $\tau(u, v) = 1$

5. Si  $|\tau(u, w) - \tau(w, v)| = 2$  alors  $\tau(u, v) = -1$ .

### 3.2.1. AGREGATION LOGIQUE DE DEUX PRÉORDRES

Dans le but de préciser les conditions dans lesquelles est réalisée l'égalité dans chacune des relations (1) et (2) ci-dessus, nous sommes amenés à proposer une définition



«naturelle» d'un préordre issu de deux préordres donnés et respectant, sur toute paire d'individus, l'indifférence et l'ordre de préférence.

#### Définition

Nous dirons qu'un préordre  $w$  est une agrégation logique de deux préordres  $u$  et  $v$  si et seulement si, pour toute paire d'individus  $i$  et  $j$ , on a :

$$u(i) > u(j) \text{ et } v(i) > v(j) \text{ alors } w(i) > w(j)$$

$$u(i) > u(j) \text{ et } v(i) = v(j) \text{ alors } w(i) \geq w(j)$$

$$u(i) = u(j) \text{ et } v(i) > v(j) \text{ alors } w(i) \geq w(j)$$

$$u(i) = u(j) \text{ et } v(i) = v(j) \text{ alors } w(i) = w(j).$$

Notons  $AL(u,v)$  l'ensemble des agrégations logiques de  $u$  et  $v$  et remarquons que  $AL(u,v)$  n'est jamais vide.

#### Proposition III.7

$w \in AL(u,v)$  si et seulement si

$$1 + \tau(u,v) = \tau(u,w) + \tau(w,v).$$

#### Preuve

Les conditions de l'agrégation logique sont exactement celles qui réalisent, pour toute paire d'individus  $i$  et  $j$  :

$$|R_u(i,j) - R_v(i,j)| = |R_u(i,j) - R_w(i,j)| + |R_w(i,j) - R_v(i,j)|.$$

Rappelons que  $R_{u \vee v}(i,j) = |R_u(i,j) - R_v(i,j)|$ .

Nous donnons, ci-après, le tableau de tous les cas possibles, en  $i$  différent de  $j$ , et les valeurs correspondantes de  $R_{u \vee v}(i,j)$ ,  $R_{u \wedge v}(i,j)$  et  $R_{w \vee v}(i,j)$ .

i différent de j	$R_{uv}(i, j)$	$R_{uw}(i, j)$	$R_{vw}(i, j)$
$u(i) > u(j)$ et $v(i) > v(j)$	0	0	0
$u(i) < u(j)$ et $v(i) < v(j)$	0	0	0
$u(i) < u(j)$ et $v(i) > v(j)$	1	0 ou 1	1 ou 0
$u(i) > u(j)$ et $v(i) < v(j)$	1	1 ou 0	0 ou 1
$u(i) > u(j)$ et $v(i) = v(j)$	0	0	0
$u(i) = u(j)$ et $v(i) > v(j)$	0	0	0
$u(i) < u(j)$ et $v(i) = v(j)$	1	0 ou 1	1 ou 0
$u(i) = u(j)$ et $v(i) < v(j)$	1	1 ou 0	0 ou 1
$u(i) = u(j)$ et $v(i) = v(j)$	0	0	0

#### Remarque

Une relation équivalente, caractéristique de l'agrégation logique, est

$$K'_{uv} = K'_{vw} + K'_{uw}$$

#### Conséquences

1. Si  $w \in AL(u, v)$  et  $\tau(u, v) = -1$  alors

$$\tau(u, w) = -\tau(v, w).$$

2.  $w \in AL(u, v)$  et  $\tau(u, v) = 1$  si et seulement si

$$\tau(u, w) = \tau(v, w) = 1.$$

3.  $w \in AL(u, v)$  et  $u \in AL(v, w)$  si et seulement si

$$\tau(u, w) = 1 \text{ et } \tau(u, v) = \tau(v, w).$$

4.  $w \in AL(u, v)$ ,  $u \in AL(v, w)$  et  $v \in AL(u, w)$  si et seulement si

$$\tau(u, v) = \tau(v, w) = \tau(u, w) = 1.$$

**Proposition III.8**

$u \in AL(v,w)$  ou  $v \in AL(u,w)$  si et seulement si

$$1-\tau(u,v) = |\tau(u,w)-\tau(w,v)|$$

**Preuve**

$u \in AL(v,w)$  ou  $v \in AL(u,w)$  revient à dire que:

$1+\tau(v,w) = \tau(v,u) + \tau(u,w)$  ou  $1+\tau(u,w) = \tau(u,v) + \tau(v,w)$  équivalent à  $\tau(u,w)-\tau(v,w) = 1-\tau(v,u)$  ou  $\tau(v,w)-\tau(u,w) = 1-\tau(v,u)$  équivalent à  $1-\tau(u,v) = |\tau(u,w)-\tau(w,v)|$ .

**CQFD**

**Remarque**

Si  $w \in AL(u,v)$ , il existe  $\mu \geq 1$  tel que

$$K'_{uv} = \mu |K'_{vw} - K'_{uw}| \text{ avec } K'_{uv} = K'_{vw} + K'_{uw}.$$

-Le niveau  $\mu=1$  correspondant à l'agrégation logique triviale  $w$  identique à l'un des préordres initiaux: on peut parler de classe triviale à deux représentants.

-Si  $K'_{uv}$  est pair, le minimum de  $|\tau(u,w)-\tau(w,v)|$  est 0 atteint en  $K'_{vw}=K'_{uw}$ ;  $\mu$  n'existe pas, on peut convenir  $\mu=\infty$  et parler de classe d'équidistance, de niveau maximum.

-Si  $K'_{uv}$  est impair, le minimum de  $|\tau(u,w)-\tau(w,v)|$  est  $2/n(n-1)$  correspondant à  $|K'_{vw}-K'_{uw}| = 1$  d'où  $\mu=K'_{uv}$ : c'est la classe de niveau maximum dans ce cas.

La classe de niveau  $\mu$  maximum contient l'agrégation logique  $w$  optimale au sens de la structure ordinale la plus proche respectivement de  $u$  et de  $v$ .

## 3.2.2.INTERPRETATION GRAPHIQUE

Etant donné  $u$  et  $v$  deux préordres fixés, sur un ensemble  $A$  d'individus et  $\tau = \tau(u, v)$  leur tau de Kendall, soit  $w$  un préordre quelconque sur  $A$  et  $\tau_1 = \tau(u, w)$ ,  $\tau_2 = \tau(v, w)$ ; le point du plan  $W(\tau_1, \tau_2)$  représente  $w$ .

La représentation graphique de l'ensemble

$$\{W(\tau_1, \tau_2) \text{ tel que } \tau_1, \tau_2 \in [-1, 1]; 1 + \tau \geq \tau_1 + \tau_2; 1 - \tau \geq |\tau_1 - \tau_2|\}$$

est le polyèdre convexe  $ABB'C'A'$ .

Notons que si  $w \in AL(u, v)$  alors  $W \in [A, B]$ . Le segment  $[A, B]$  constitue la frontière des agrégations logiques et les préordres  $a$  et  $b$ , associés aux points  $A$  et  $B$ , sont les deux représentants de la «classe triviale» (voir section précédente).

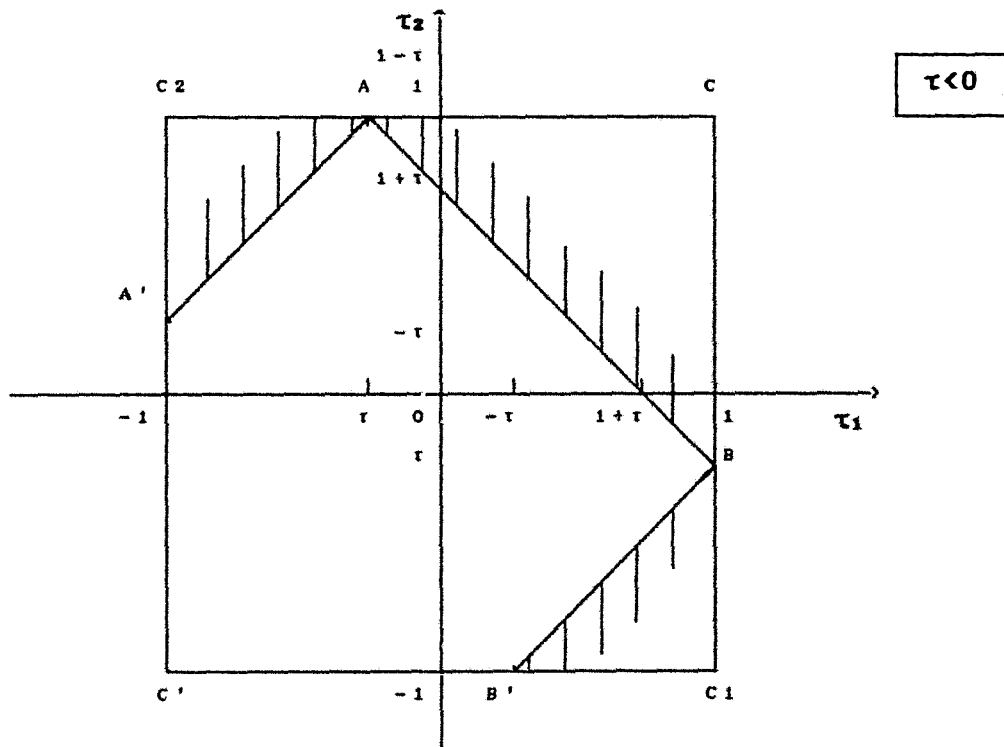
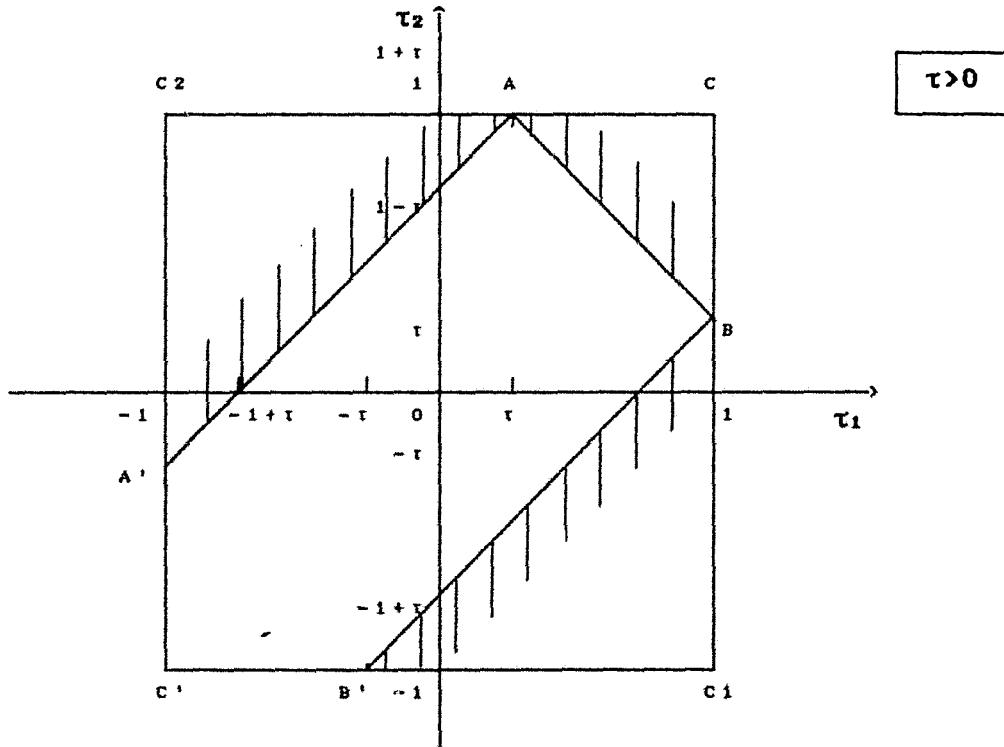
Nous distinguons deux cas :

1.  $\tau > 0$ :

Si  $\tau$  se rapproche de 1 alors le segment  $[A, B]$ , frontière des agrégations logiques, tend vers le point  $C$  de coordonnées  $\tau_1 = \tau_2 = 1$  et le polyèdre vers le segment  $[C', C]$ .

2.  $\tau < 0$ :

Si  $\tau$  se rapproche de -1 alors le segment  $[A, B]$ , frontière des agrégations logiques, tend vers le segment  $[C_1, C_2]$  dont le support est d'équation  $\tau_1 + \tau_2 = 0$  et le polyèdre vers  $C_1 C' C_2$ .



## 3.2.3. ETUDE NUMERIQUE

Considérons, sur sept individus, les préordres

$$u' = (1 \ 2 \ 4 \ 3 \ 5 \ 6 \ 7) ; v' = (1 \ 2 \ 3 \ 1 \ 4 \ 5 \ 5).$$

Notons que:  $K'_{uv} = 4$  ,  $\tau(u,v) = 0.809524$

et la décomposition  $4 = 2(3-1)$ ; donc, hormis les classes

triviale:  $\mu = 1$  ,  $|K'_{uw} - K'_{vw}| = 4$

et d'équidistance:  $\mu = \infty$  ,  $\tau(u,w) = \tau(v,w)$  ,

une seule classe d'agrégation logique possible:  $\mu=2$ .

Nous construisons arbitrairement, à partir de  $u$  et  $v$ , une agrégation logique:

$$w' = (1 \ 3 \ 4 \ 2 \ 5 \ 6 \ 6)$$

on obtient  $\tau(u,w) = 0.857143$  ,  $\tau(v,w) = 0.952381$

et les relations:  $1 + \tau(u,v) = \tau(u,w) + \tau(w,v)$

$$1 - \tau(u,v) = 2 |\tau(u,w) - \tau(w,v)|.$$

Notons que l'agrégation logique optimale au sens des rangs (niveau  $\mu=\infty$ ) est

$$w^{*'} = (1 \ 2 \ 3 \ 2 \ 4 \ 5 \ 5)$$

correspondant à  $K'_{uw} = K'_{vw} = 2$ ;

on a donc  $\tau(u,w) = \tau(v,w) = 0.904762$ .

#### 4. CODAGE ET TAU DE KENDALL

Nous effectuons dans cette section, une étude reliant la notion de codage optimal d'un préordre et le tau de Kendall.

Deux résultats en découlent:

-Une formule du tau de Kendall entre le préordre et un, quelconque, de ses codages; elle établit la positivité du tau et en donne une borne inférieure.

-La projection orthogonale sur le cône des codages (section 2.2) est, de tous les codages admissibles, celui qui est le mieux corrélé, au sens du tau de Kendall, avec le préordre.

##### 4.1. UNE FORMULE ORIGINALE DU TAU DE KENDALL

Soit  $u = \sum \{\alpha_i Z_i, i=1, \dots, k\}$ , un codage de  $X$ . Considérons l'ensemble

$$J = \{j \mid \alpha_j \neq 0, j \in M \setminus \{1\}\} = \{j_1, j_2, \dots, j_r\};$$

posons  $j_0 = 1$  et  $j_{r+1} = k+1$ ;  $J$  induit la partition de  $M$ :

$$B_q = \{j_q, j_q+1, \dots, j_{q+1}-1\}, q=0, \dots, r.$$

Remarquons que  $u = \sum \{\alpha_{j_q} Z_{j_q}, q=0, \dots, r\} = \sum \{\beta_q 1_{B_q} \circ X, q=0, \dots, r\}$

avec  $\alpha_{j_0} = \alpha_1 = \beta_0$  et  $\alpha_{j_q} = \beta_q - \beta_{q-1}$ ,  $q=1, \dots, r$ .

Notons  $b_q = \text{CARD}(\{(i, j), X(i), X(j) \in B_q, X(i) < X(j)\})$ .

Soit  $\text{COD}_J(X)$  l'ensemble des codages de  $X$  auxquels est associé le même sous-ensemble  $J$ , induisant donc la même partition  $\{B_q\}$ .

##### Propositions III.10

Le Tau de Kendall entre  $X$  et tout codage appartenant à  $\text{COD}_J(X)$  est constant, toujours positif ou nul et il vaut:

$$1 - (2 \sum \{b_q, q=0, \dots, r\} / (n^2 - n)).$$

Preuve

On doit montrer que

$$\sum \{b_q, q=0, \dots, r\} = K'_{X_u}$$

Remarquons que les seules différences possibles entre  $X$  et  $u$  sont des indifférences entre certains individus  $i, j$  dans le classement  $u$  :  $u(i)=u(j)$  alors que  $X(i) \neq X(j)$ .

A chaque classe d'indifférence induite par  $u$  dans  $\{1, \dots, n\}$ , correspond un ensemble  $B_q$  et  $b_q$  en est le nombre de couples distincts, égal à  $\sum \{|R_X(i, j) - R_u(i, j)|; i \neq j\}$ , somme restreinte aux individus de cette classe.

CQFD

#### 4.2. TAU DE KENDALL OPTIMAL

Rappelons que  $k_q = \text{CARD}\{X \in B_q\}$ ,  $J^* = J \setminus \{j_m\}$ ;

$J^*$  induit la partition  $\{B_q^*, q=0, \dots, r-1\}$  (voir section 2.1).

Soit  $v \in \text{COD}_{J^*}(X)$  et  $\tau^*$  le tau de Kendall entre  $X$  et  $v$ . Nous nous proposons d'évaluer la variation du Tau de Kendall:

Proposition III.11

En passant de  $J$  à  $J^*$ , le Tau de Kendall diminue de la quantité

$$\tau - \tau^* = 2k_m k_{m-1} / (n^2 - n)$$

Preuve

Il suffit de développer

$$\tau - \tau^* = 2[(k_{m-1} k_m + b_m + b_{m-1}) - (b_m + b_{m-1})] / (n^2 - n).$$

Remarque

Parmi les couples de blocs  $(B_{q-1}, B_q)$  tels que  $\text{CARD}(X \in B_{q-1} \cup B_q) = c$  fixé, notons que la diminution est maximale pour  $k_q \approx c/2$ .



**Corollaire III.11.1**

La valeur minimale de  $\{\tau(X,u), u \in \text{COD}(X)\}$  est

$1 - (2b_0 / (n^2 - n))$ , atteinte pour  $u \in \text{COD}_n(X)$ .

**Preuve**

Il suffit de noter que la partition induite par  $J = \emptyset$  est réduite au seul ensemble  $B_0 = M$ .

CQFD

Soit  $y^J$  la projection orthogonale de  $y$  sur  $\text{COD}(X)$ .

**Corollaire III.11.2**

Le tau de Kendall entre  $X$  et  $y^{J'}$ ,  $J'$  un sous-ensemble minimal, est maximal pour  $J'$  optimal:

$$\text{MAX}\{\tau(X, y^{J'}) \mid y^{J'} \in \text{COD}(X)\} = \tau(X, y^J).$$

**Preuve**

Le corollaire III.4 implique que  $J$ , optimal, contient tous les sous-ensembles minimaux; les propositions III.10 et III.11 induisent ce corollaire.

**Remarque**

Ce résultat peut aussi être établi en considérant que pour tout sous-ensemble minimal  $J'$ ,  $y^J$  est une agrégation logique (section 3.2.1) de  $X$  et  $y^{J'}$ , ce qui amène à

$$\tau(X, y^{J'}) = \tau(X, y^J) + \tau(y^J, y^{J'}) - 1 \leq \tau(X, y^J).$$

Plus généralement, pour tout codage  $u$  de  $X$  tel que  $y^J$  constitue une agrégation logique de  $X$  et  $u$ , on a:

$$\tau(X, u) = \tau(X, y^J) + \tau(y^J, u) - 1 \leq \tau(X, y^J).$$

**Chapitre IV**

**PRESENTATION DU PROGRAMME REGALS**

**ET EXEMPLES D'APPLICATIONS**

#### IV. PRESENTATION DU PROGRAMME REGALS ET EXEMPLES D'APPLICATIONS

Il sera fait référence dans ce chapitre à des notations et résultats du chapitre III, principalement: la définition et notation de l'ensemble des codages d'une variable ordinale (section III.2) et les résultats concernant le codage optimal d'une variable ordinale (section III.2).

La section 1 présente de façon succincte le programme REGALS, l'outil informatique que nous avons développé et qui a permis le traitement des exemples proposés.

La régression sous contraintes est abordée dans la section 2. Elle a deux champs principaux d'application: la régression monotone et les problèmes de «protection de la régression».

La section 3 concerne la régression qualitative. Nous la décrivons comme une analyse canonique sous contraintes et en présentons deux méthodes classiques: MORALS et MONANOVA.

Le TRADE-OFF est traité dans la dernière section. Nous en précisons le champ d'application et donnons une description et une illustration de la méthode.

## 1.LE PROGRAMME REGALS

Le programme REGALS est le résultat d'une structuration logicielle d'un ensemble de routines, écrites en langage TURBO-PASCAL et accumulées depuis nos premières investigations dans le domaine de la régression qualitative.

Ces routines ont bénéficié d'une continuelle mise à jour au rythme du progrès de notre recherche d'amélioration des algorithmes aussi bien informatiques que mathématiques.

### 1.1.DESCRPTION GENERALE

Ecrit en TURBO-PASCAL (Borland,version 5.5),le programme REGALS, version 1.0, est pourvu de menus déroulants, d'une aide (touche F1) et est régi par des règles simples. Le menu principal est présenté ci-dessous.

FICHER	TRAITEMENT	OPTIONS	ENVIRONNEMENT
-Ouvrir	-Regals	-Répertoire	-Info/Pc
-Imprimer	-Monanova	-Impression	-Mémoire
-Renommer	-Trade-Off	-Coul/Mono	-Video
-Supprimer	-Régression	-Réinit.	-Dos

### 1.2.LES FICHIERS

Les fichiers REGALS sont de type ASCII.

#### 1.En entrée:

-Extension :.REG par défaut

-(Lec/ouver)ture : par exploration de répertoires ou spécification de nom de fichier.

-Variables: de quatre types à spécifier: NUM,NOM,ORD et TOF qui est une «super-variable»,groupe de variables nominales (dix au plus,observées sur au plus vingt et un individus ), réservée au TRADE-OFF.

## 2. En sortie:

-Extension: .UTL (fichier des utilités du TRADE-OFF)  
.RES (autre).

-Un menu est proposé en fin de traitement et comporte, entre autres choix (liés à la méthode appliquée), une redirection des résultats vers un support magnétique ou l'imprimante.

### 1.3. LES METHODES

Elles couvrent les analyses : multicritère (REGALS), de variance monotone (MONANOVA), des préférences sur un échantillon de classements (TRADE-OFF) et de REGRESSION (multiple classique et sous contraintes linéaires sur les coefficients).

### 1.4. OPTIONS

Deux seulement des quatre options proposées nécessitent des commentaires:

-Répertoire: permet de spécifier, pour le fichier de sortie, le répertoire de sauvegarde qui est, par défaut, celui du fichier des données traitées.

-Impression: permet de définir les paramètres d'impression: largeur, longueur, marges, type de caractère.

### 1.5. ENVIRONNEMENT

Un ensemble d'informations est proposé dans cette rubrique concernant le mode video actif, l'état de la mémoire, la présence d'un coprocesseur numérique, etc...

## 2. REGRESSION SOUS CONTRAINTES

Cette section traite du problème de la régression sous contraintes linéaires sur les coefficients et de ses deux applications principales : la «protection» de la régression et la régression monotone. Ce dernier sujet fait référence aux notations et résultats établis au chapitre III, section 1, dont une lecture préalable est nécessaire.

### 2.1. LE PROBLEME

On considère un ensemble  $A$  de  $n$  individus, sur lesquels on observe  $p+1$  variables :  $X_0$  la variable à expliquer et  $X_1, X_2, \dots, X_p$ ,  $p$  variables explicatives, toutes numériques.

Soit  $R$  une matrice  $m \times p$ , de vecteurs colonnes  $R_j$ ,  $r \in \mathbb{R}^m$ .

On recherche :

$$\text{MIN}(\|X_0 - z\|, z \in \text{POL}(R, r))$$

avec 
$$\text{POL}(R, r) = \{\sum \beta_j X_j \mid \sum \beta_j R_j \geq r\}.$$

### 2.2. LES METHODES

Elles concernent respectivement deux champs principaux d'application :

-La «protection» de la régression découlant de problèmes numériques, le plus souvent liés à la nature des variables (signification des coefficients en physique, géologie, etc..).

-La régression monotone découlant de problèmes sur des données ordinales (de préférences en marketing, par exemple).

#### 2.2.1. «PROTECTION» DE LA REGRESSION

«Dans un certain nombre de problèmes économiques physiques, géologiques, statistiques, etc., on ne peut toujours appliquer directement la méthode des moindres carrés pour expliquer une variable  $y$  en fonction d'autres variables  $x_1, x_2, \dots, x_p$ . C'est le cas si les coefficients de régression  $a_1, a_2, \dots, a_p$  doivent être compris entre des bornes données pour avoir un sens» (Cazes et Turpin, 1971).

Dans leur étude de «l'Estimation de la Courbe Granulométrique d'un Aérosol», Cazes et Turpin (1971) utilisent une méthode de programmation quadratique.

L'algorithme MCCB (section I.3.2) apporte une solution élégante et rapide à ce type de problème.

### 2.2.2. REGRESSION MONOTONE

On se pose le problème de la recherche du codage de la variable ordinaire  $X$  «approchant» au mieux, au sens des moindres carrés, une variable numérique  $y$ .

La solution est la projection orthogonale de  $y$  sur le cône polyédrique convexe  $COD(X)$ :  $R$  est la matrice identité  $p \times p$ ,  $r$  est le vecteur nul.

Ce problème peut être résolu par NNLS mais, parfaitement adapté à ce cas, l'algorithme PAV (Pool Adjacent Violators) de Barlow & al. (1972) en effectue le meilleur traitement.

Simple et rapide, il est basé sur la structure particulière (section III.1.2) du cône polyédrique convexe engendré par la variable ordinaire.

Soit une partition de  $M$ ,  $\{B_q, q=0, \dots, r\}$ , induite par un sous-ensemble  $J = \{j_1, \dots, j_r\}$  de  $\{2, \dots, k\}$  (voir section III.1.1). Notons  $C^J$  l'application de  $M$  dans  $R$ , définie par :

$$C^J(j) = \sum \{My(B_q) 1_{B_q}(j)\}$$

Notons que  $C^J(X) = \sum \{My(B_q) 1_{B_q}(X), q=0, \dots, r\} = y^J$ ,

est la projection de  $y$  sur  $L^J$ .

Il a été établi (section III.2.2) que  $J$  est optimal si et seulement si  $C^J$  est un codage de  $X$  et tout sous-ensemble minimal est contenu dans  $J$ ; ce qui justifie la démarche "descendante" de l'algorithme PAV:

1. Initialisation :  $J = \{2, \dots, k\}$

2. Soit  $\{B_q, q=0, \dots, r\}$ , induite par  $J$ ; calculer

$$\alpha_q = My(B_q) - My(B_{q-1}), \quad q=1, \dots, r.$$

3. Si  $J = \emptyset$  ou  $\text{MIN}\{\alpha_q, q=1, \dots, r\} > 0$  aller à 5 sinon

4.  $J = J \setminus \{j_q \mid \alpha_q \leq 0\}$ ; aller à 2.

5.  $J$  est optimal:  $y^J = P_{\text{con}(X)}(y)$ .

### 2.2.3. UN EXEMPLE DE REGRESSION MONOTONE

Les données portent sur un échantillon de 24 offres de ventes d'appartements, situés dans le V<sup>e</sup> et le VI<sup>e</sup> arrondissement de Paris, en 1975 (Saporta, 1978), et sont décrites par deux variables: PRIX (en millier de francs) et SURFACE (en m<sup>2</sup>). Une régression multiple, appliquée avec le critère PRIX, a donné l'équation:  $\text{PRIX} = 3.9844[\text{SURFACE}] + 30.0921$ .

Nous allons montrer l'intérêt d'une prise en compte de la structure ordinale du prédicteur SURFACE.

Malgré un coefficient de corrélation élevé (0.9733), l'intervalle de confiance, au risque  $\alpha=5\%$ , associé à l'estimation du prix pour une surface de 100 m<sup>2</sup> est imprécis: [334.89 522.17] (Saporta, 1978).

Nous considérons la variable SURFACE comme ordinale et cherchons à expliquer le prix. La description des données et les résultats sont présentés ci-après (sortie du programme REGALS).

Huit appartements ont leurs prix (marqués d'un asterisque) exactement restitués par le codage. Notons que  $\text{PRIX}_{\text{calculé}}$  n'est pas une «agrégation logique» (section III.3.2.1) de SURFACE et PRIX.

Une régression multiple, appliquée avec le vecteur  $\text{PRIX}_{\text{calculé}}$ , donne l'équation:  $\text{PRIX} = 4.0049[\text{SURFACE}] + 28.6568$ . Le coefficient de corrélation est de 0.9877. L'intervalle de confiance, au risque  $\alpha=5\%$ , associé à l'estimation du prix pour une surface de 100 m<sup>2</sup> est [366.04 492.25] et présente une amélioration notable.



## SORTIE REGALS

Traitement      Régression Monotone  
Données        VTEAPP.RGM

Variable à expliquer : PRIX  
Variable explicative : SURFACE

Tau(PRIX,SURFACE)      = 0.8406  
Cor(PRIX,SURFACE)      = 0.9733  
Tau(PRIXcalc,SURFACE) = 0.9348  
Cor(PRIXcalc,SURFACE) = 0.9877  
Tau(PRIXcalc,PRIX)     = 0.8696  
Cor(PRIXcalc,PRIX)     = 0.9905  
Tau au seuil 5%        = 0.2863

SURFACE *	CLASSEMENT *	PRIX *	PRIXcalculé
28.000	20	130.000	97.667
50.000	14	280.000	243.750
196.000	1	800.000	800.000*
55.000	12	268.000	268.000*
190.000	2	790.000	790.000*
110.000	3	500.000	500.000*
60.000	11	320.000	302.000
48.000	15	250.000	243.750
90.000	6	378.000	378.000*
35.000	17	250.000	225.000
36.000	7	350.000	350.000*
65.000	10	300.000	302.000
32.000	18	155.000	155.000*
52.000	13	245.000	243.750
40.000	16	200.000	225.000
70.000	9	325.000	302.000
28.000	20	85.000	97.667
30.000	19	78.000	97.667
105.000	4	375.000	435.000
52.000	13	200.000	243.750
80.000	8	270.000	302.000
60.000	11	295.000	302.000
20.000	21	85.000	85.000*
100.000	5	495.000	435.000

### 3. REGRESSION QUALITATIVE

Comme méthode d'ajustement d'un modèle à des données, la régression qualitative est une généralisation de la régression multiple au cas d'une variable nominale ou ordinale, à expliquer par un ensemble  $\{X_1, X_2, \dots, X_p\}$  de variables de type quelconque.

Le problème peut être décrit comme une analyse canonique sous contraintes et la plupart des méthodes sont basées sur le principe des «projections alternées au sens des moindres carrés» (ALS).

Dans le cas de l'analyse canonique d'un cône polyédrique convexe et d'un espace vectoriel, nous avons proposé un algorithme original (ACCE, chapitre II).

Deux méthodes classiques de la régression qualitative sont décrites: les programmes MORALS et MONANOVA.

#### 3.1. LE PROBLEME

Ayant déjà abordé et décrit le problème général (chapitre II), nous le posons, dans cette section, en termes de recherche de fonctions d'utilités et de façon appropriée à l'explication d'une variable qualitative.

On considère un ensemble  $A$  de  $n$  individus sur lesquels on a observé  $p+1$  variables  $g_j$ , la variable à expliquer étant  $g_0$ :

$$g_j(i) = \text{modalité de la variable } j \text{ sur l'individu } i$$

$i=1, 2, \dots, n; j=0, 1, 2, \dots, p; g_j$  prenant ses valeurs dans

$$\mu_j(M_j) = \{\mu_j(1), \dots, \mu_j(k_j)\}$$

ensemble des  $k_j$  modalités du facteur  $j$ , en notant  $M_j$  l'ensemble des  $k_j$  premiers entiers naturels.

Notons que si  $g_j$  est numérique,  $k_j=n$ .

Il s'agit de déterminer les fonctions d'utilité  $U_j$  sur  $R$  définissant l'ensemble des utilités optimales des modalités de

chaque variable  $j$ ,  $j \geq 0$  et les poids  $\beta_j$ ,  $j \geq 1$ :

$$U_j \circ \mu_j (M_j) = \{U_j(\mu_j(1)), \dots, U_j(\mu_j(k_j))\}, j=1, 2, \dots, p$$

avec la normalisation :  $0 \leq U_j(\mu_j(k)) \leq 1$ ,  $k=1, 2, \dots, k_j$

et  $\beta_j \geq 0$ ,  $\sum \beta_j = 1$ .

Notons que si  $g_j$  est muni d'une structure ordinale,  $U_j \circ \mu_j$  doit être monotone.

Soit  $X_j$  l'application définie sur  $A$  et prenant ses valeurs dans  $M_j$  telle que

$$\mu_j \circ X_j = g_j$$

$X_j$  est la variable d'"état" associée à  $g_j$ .

La variable  $C_j = U_j \circ \mu_j$  est une application de  $M_j$  dans  $R$  donc  $C_j$  est un codage des modalités de  $X_j$

et  $C_j \circ X_j = (U_j \circ \mu_j) \circ X_j = U_j \circ (\mu_j \circ X_j) = U_j \circ g_j$

est un codage de la variable  $X_j$ . Si  $g_j$  est muni d'une structure d'ordre,  $X_j$  est ordinale.

Déterminer les utilités optimales  $U_j$  revient donc à rechercher les codages optimaux  $C_j$  des modalités des  $X_j$ .

Notons que:

-Si  $\mu$  est l'Identité de  $M_j$  dans  $M_j$ :

$$\mu_j(k) = k, k=1, \dots, k_j;$$

alors  $g_j = X_j$  est ordinale ou nominale.

-Si  $g_j$  est numérique,  $X_j$  est la bijection de  $A$  dans  $M_j$ :

$$X_j(i) = i, i=1, \dots, n; g_j = \mu_j \text{ et } U_j \text{ linéaire.}$$

Pour  $j \geq 1$ , on peut décider de fixer  $C_j = U_j \circ \mu_j = \mu_j$  :

$$C_j \circ X_j = \mu_j \circ X_j = g_j$$

et  $\text{COD}(X_j) = g_j$ .

L'objectif est la recherche de  $\text{MAX}\{\cos^2(U_0(g_0), \sum \beta_j U_j(g_j))\}$ .

La méthode de résolution consiste à effectuer l'analyse canonique de  $\text{COD}(X_0)$  et de  $\text{COD}(X_1, X_2, \dots, X_p) = \sum \{\text{COD}(X_j), j \geq 1\}$  engendrés respectivement par la variable  $X_0$  associée au critère  $g_0$  et les variables  $X_j$ , associées aux prédicteurs  $g_j$ .

#### Remarque

Nous posons le problème en termes de recherche du cosinus carré maximal mais si l'on cherche à maximiser la corrélation linéaire, il suffit de centrer les variables. En pratique, dans le but d'homogénéiser les données (variables explicatives mixtes, par exemple), on recherche des codages centrés-réduits.

### 3.2. LES METHODES

Deux programmes principaux sont décrits : **MORALS** et **MONANOVA** qui traite l'analyse de variance monotone que nous présentons comme un cas particulier de l'analyse des mesures conjointes.

#### 3.2.1. LE PROGRAMME MORALS

Proposé par F.W.Young, J.de Leeuw et Y.Takane (1976), c'est une méthode de résolution du problème de la régression qualitative étendue à tous les types de variables dont elle détermine des codages centrés-réduits. Le critère à optimiser est équivalent au critère général de l'analyse canonique sous contraintes.

#### L'algorithme:

1. A la première étape, on fixe, arbitrairement, des codages centrés-réduits  $X_j^0$  ( $\text{COD}(X_j)$ ) de toutes les variables; on définit

$$Z^0 = \sum \beta_j^0 X_j^0, \text{ projection de } X_0^0 \text{ dans } L(X_j^0, j \geq 1).$$

2. A l'étape courante  $t$ , trois phases distinctes:

$$1. X_0^t = P_{\text{con}}(x_0) (Z^{t-1})$$

2. Détermination de  $X_j^t$  ( $X_1^t, \dots, X_{j-1}^t$  connus):

$$U_j^t = X_0^t - \sum \{\beta^{t-1}_k X_k^t, k \leq j-1\}$$

$$X_j^t = P_{\text{con}}(x_j) \left( (1/\beta^{t-1}_j) (U_j^t - \sum \{\beta^{t-1}_k X_k^{t-1}, k \geq j\}) \right)$$

3.  $Z^t = \sum \beta^t_j X_j^t$ , projection de  $X_0^t$  dans  $L(X_j^t, j \geq 1)$ .

### 3.2.2. ANALYSE DES MESURES CONJOINTES

Issue de la psychologie mathématique, cette méthode sert «à mesurer l'effet conjoint de plusieurs variables indépendantes sur l'ordre des valeurs prises par une variable à expliquer, cette liaison pouvant être sous forme additive ou polynômiale». Cette définition est empruntée à Y. Evrard et P. Lemaire (1976).

Dans l'hypothèse de lien additif, on entre dans le cadre des problèmes de la régression qualitative, avec la variable à expliquer ordinaire et où l'objectif est l'élaboration de «modèles de compromis», postulant l'existence d'utilités additives. Cette existence est justifiée par le système axiomatique implicitement supposé, de la Théorie de l'Utilité Multi-Attribut (von Neumann & Morgenstern, 1947).

L'analyse de variance monotone en est un cas particulier, traitant le cas de variables explicatives nominales.

#### 3.2.2.1. LE PROGRAMME MONANOVA

C'est une méthode d'analyse de variance monotone (MONotone ANalysis Of VARIance), proposée par J.B. Kruskal (1964), programmée en FORTRAN IV par Kruskal et Carmone, (1968).

La variable à expliquer est ordinaire et les variables explicatives nominales. Le but recherché est la minimisation du STRESS (analogue à celui de l'analyse des similarités):

$$\text{STRESS}(u, v) = |u - v| / |v - E(v)|.$$

Il est intéressant de rappeler l'interprétation proposée par Kruskal (1964a, page 3), de différents niveaux de STRESS :

STRESS (%)	Qualité de l'ajustement
20	faible
10	moyenne
5	bonne
2½	excellente
0	"parfaite"

L'algorithme MONANOVA recherche un codage centré de  $X_0$  et des codages centrés-réduits des  $X_j$ .

La méthode de minimisation du STRESS choisie par Kruskal est basée sur la programmation non linéaire (méthode du GRADIENT).

Tenenhaus en a montré l'équivalence avec un algorithme de type ALS, fondé sur le résultat suivant:

Soit  $u \in \mathbb{R}^n$  fixé,  $L$  un sous-espace de  $\mathbb{R}^n$ .

Proposition (Tenenhaus, 1979)

Le minimum de  $\{\|u-v\|/\|v\|, v \in L \cap \mathbb{R}^n\}$  est

$\sqrt{1 - \cos^2(u, P_{L \perp}(u))}$ , atteint en

$v^* = (1/\cos^2(u, P_{L \perp}(u)))P_{L \perp}(u)$ .

L'algorithme:

1. A la première étape, on fixe, arbitrairement, un codage centré  $X_0^0$  de  $X_0$ ; on détermine la projection  $U^0$  de  $X_0^0$  sur  $\text{COD}(X_1, \dots, X_p)$  et

$$Z^0 = (1/\cos^2(X_0^0, U^0))U^0 = \sum X_j^0$$

2.A l'étape courante  $t$ , 2 phases distinctes:

$$1. X_0^t = P_{\text{COD}}(x_0) (Z^{t-1})$$

2.  $U^t$  est la projection de  $X_0^t$  sur  $\text{COD}(X_1, \dots, X_p)$

et

$$Z^t = (1/\cos^2(X_0^t, U^t)) U^t = \Sigma X_J^t.$$

### 3.2.2.2. UN EXEMPLE DE MONANOVA

Nous proposons le traitement par **MONANOVA** de l'étude de la «résistance d'un fil de laine à la traction» de Barella et Sust, repris par Box et Cox (1964) et Kruskal (1965).

Les données sont constituées par le nombre  $g_0$  de cycles (**NUMBCYCL**), nécessaire à la rupture du fil dans 27 conditions expérimentales, classées par ordre croissant et définies par les variables à 3 niveaux:

$$g_1 = \text{longueur du fil (LENGYARN:250,300,350 mm)}$$

$$g_2 = \text{amplitude du cycle (AMPLCYCL:8,9,10 mn)}$$

$$g_3 = \text{poids (LOAD:40,45,50 gr)}$$

L'analyse de variance effectuée en complément à l'étude (sortie du programme **REGALS** ci-après) amène à la conclusion que le nombre de cycles  $g_0$  est bien expliqué, par ordre décroissant d'importance, par  $g_1$ ,  $g_2$  et  $g_3$ .

Kruskal donne le **STRESS** optimal de 7.1% ; nous retrouvons ce résultat (7.085% , plus précisément).

D'après l'échelle de Kruskal, on peut conclure que la qualité de l'ajustement est assez bonne.

## Données de Barella &amp; Sust

NUMBCYCL	LENGYARN	AMPLCYCL	LOAD
90	1	3	3
118	1	3	2
170	1	3	1
210	1	2	3
220	2	3	3
266	1	2	2
292	1	1	3
332	2	3	2
338	1	2	1
360	3	3	3
370	1	1	2
438	2	2	3
442	2	3	1
566	3	2	3
620	2	2	2
634	2	1	3
674	1	1	1
884	3	3	2
1022	2	2	1
1070	3	2	2
1140	3	3	1
1198	2	1	2
1414	2	1	1
1568	3	2	1
2000	3	1	3
3164	3	1	2
3636	3	1	1



## SORTIE REGALS

Traitement Monanova  
Données SUSDATA.SON

Classement à expliquer : ORDEE  
Stress de Kruskal = 7.0848 %  
Tau de Kendall = 0.9202 (Tau(α=5%) = 0.2679)  
Corrélation linéaire = 0.9975  
Nombre d'itérations = 10

## POIDS (%) des Facteurs &amp; UTILITES des Niveaux

LENGYARN AMPLCYCL LOAD  
46.14911 32.21496 21.63593

0.000000 0.677386 0.652272  
0.375150 0.322614 0.347728  
0.624850 0.000000 0.000000

R/P	* PRODUIT*	UTILITE *	N/R/P
1	P27	0.66771	1
2	P26	0.58182	2
3	P25	0.50658	5
4	P24	0.53342	3
5	P23	0.53247	4
6	P22	0.46658	7
7	P21	0.42949	8
8	P20	0.46753	6
9	P19	0.41818	9
10	P18	0.36360	12
11	P17	0.35934	13
12	P16	0.39135	11
13	P15	0.35229	14
14	P14	0.39229	10
15	P13	0.31425	15
16	P12	0.27706	18
17	P11	0.29245	16
18	P10	0.28836	17
19	P9	0.24505	20
20	P8	0.24836	19
21	P7	0.21822	21
22	P6	0.17916	22
23	P5	0.17313	23
24	P4	0.10393	25
25	P3	0.14113	24
26	P2	0.07523	26
27	P1	0.00000	27

## Analyse de Variance Monotone

Source de Variation	D.f.	Som. des Carrés	Carrés Moyens	Pseudo-F
LENGYARN	2	880.1404	440.0702	1099.9217
AMPLCYCL	2	497.7274	248.8637	622.0157
LOAD	2	208.3148	104.1574	260.3334
RESIDUELLE	20	8.0018	0.4001	

#### 4. LE TRADE-OFF

Le problème principal auquel s'applique le TRADE-OFF concerne en Marketing, le positionnement d'un nouveau produit ou la relance d'un produit «en perte de vitesse» par la définition ou la redéfinition de ses caractéristiques.

Mais le TRADE-OFF est une technique permettant aussi de:

- simuler des «scénarios» (tester des produits sans avoir à les élaborer physiquement).

- mesurer les qualités du produit sur une échelle numérique.

- simuler toutes les situations de concurrence.

- calculer en parts de marchés les résultats de chacun des produits proposés à l'échantillon.

- etc..

De plus, la richesse de la méthode apporte une meilleure connaissance de la Concurrence et du Marché, l'analyse pouvant être affinée par une segmentation (détermination de groupes d'acheteurs à partir de critères comme l'âge, le sexe, le revenu, etc..).

##### 4.1. LE PROBLEME

L'idée de base est la constitution d'un échantillon représentatif de «barèmes», chaque barème étant un système d'utilités propre à chaque individu, par lequel il évalue les différentes caractéristiques du produit et effectue son choix, «qui s'opère entre une série d'options imparfaites et implique le sacrifice de certaines qualités du produit en vue de l'obtention d'autres bénéfices plus attendus» (J. Schworer & J. Ph. Faivre, Revue Française de Marketing, n°55). Ceci signifie donc que le produit idéal n'existe pas en général.

#### 4.2.LA METHODE

Elle consiste à soumettre à chaque individu d'un échantillon représentatif du marché visé, un ensemble de «scénarios» ou stimuli (combinaisons de différents niveaux d'attributs du produit formant un plan factoriel), l'interviewé devant en effectuer un classement. Le traitement de ce classement par un algorithme de type MONANOVA fournit le barème ou système de valeur de cet individu.

#### 4.3.DESCRPTION DU TRAITEMENT REGALS

Deux possibilités sont offertes: effectuer des simulations à partir d'un fichier d'utilités ou traiter une variable TOF (section 1.2) à partir d'un fichier de classements. Nous allons décrire cette dernière option.

##### En entrée

Un écran de paramétrage est proposé en début de traitement. Il permet les choix suivants:

1.Type de classement (1/2/3): numéros de produits ou codages «lexicographiques» par ordre décroissant ou vecteur des rangs.

2.Traitement automatique (O/N): offre la possibilité d'effectuer le traitement détaillé par classement (analyse de variance, simulations,etc..).

3.Fichier des utilités (0/1/2): le fichier (extension UTL) est, au choix 1 ou 2, enregistré sur disque ou imprimé.

4.Nombre maximal d'itérations: du processus MONANOVA appliqué à chaque classement.

5.Niveau de Cohérence (1/.../9): désigne le tau de Kendall (x10) à partir duquel les classements sont sélectionnés.

##### En sortie

En fin de traitement, un menu propose des analyses, globale et sélective séparées mais, comme l'abaissement du ni-

veau de «cohérence» permet de donner le «global», les résultats imprimés concernent uniquement le «sélectif»:

-Effectif total du TRADE-OFF.

-Nombre de classements retenus au niveau de cohérence préalablement fixé.

-Nombre maximal d'itérations, préalablement fixé, du processus MONANOVA appliqué à chaque classement.

-Cohérence moyenne globale: le tau de Kendall entre les vecteurs des moyennes respectives des rangs et des utilités dans le traitement global.

-M/rg: le vecteur des rangs moyens.

-M/ut: le vecteur des utilités moyennes.

-Cohérence moyenne sélective: le tau de Kendall entre les vecteurs M/rg et M/ut.

-Corrélation linéaire entre les vecteurs M/rg et M/ut.

-Le tableau du nombre d'occurrences par rang de chaque produit.

-Le tableau des poids et des utilités en moyennes, relatifs respectivement, aux facteurs et à leurs niveaux.

#### 4.4. UN EXEMPLE DE TRADE-OFF

A titre d'illustration, nous considérons une étude (SAS/STAT User's Guide, 1989) portant sur l'appréciation donnée par six personnes sur dix-huit tasses de thé, chacune combinant les différents niveaux de quatre facteurs: la température (TEMP: chaud, tiède et froid), le sucre (SUCRE: zéro, un, deux), la force (FORCE: fort, moyen, léger) et le citron (CITRON: avec, sans). L'étude consiste en l'analyse des six classements donnés par ces personnes.

Le nombre maximal d'itérations étant fixé à 3, cinq classements sur six ont été retenus au niveau 0.8. La cohérence de -0.876 est bonne. L'étude amène à conclure que le facteur «température» est le plus important. Il est à noter, d'après les utilités moyennes des niveaux, que les personnes préfèrent le thé chaud ou froid mais non tiède.

Le thé idéal est donc chaud, sans sucre, fort ou moyen, avec du citron.

La cohérence moyenne dans le traitement sélectif, peut être inférieure à celle du traitement global si les classements cohérents au seuil fixé, sont très contrastés.

C'est le cas dans cette étude. Le programme donne la possibilité d'un traitement unitaire détaillé permettant de mettre en évidence ce fait. L'ajout des classements non retenus peut ainsi avoir un effet positif sur le tau de Kendall.

Dans le traitement global, la cohérence moyenne est de -0.922. L'ordre d'importance des facteurs est préservé mais SUCRE et FORCE ont des poids très proches. Relativement à ce dernier facteur, le niveau "moyen" domine les deux autres. Nous donnons, ci-dessous, le tableau des poids des facteurs et des taus de Kendall, en pourcentage, par individu.

## INDIVIDUS

	1	2	3	4	5	6	MOY.
TEMP	79.89	57.05	41.78	16.27	18.30	23.91	39.53
SUCRE	09.55	31.69	07.76	77.51	13.84	20.24	26.77
FORCE	06.19	08.32	49.83	04.55	35.72	49.09	25.62
CITRON	04.37	02.94	00.63	01.67	32.14	06.76	08.08
TAU	86.93	93.46	89.54	92.16	93.46	68.63	92.16

## Les données

Produits	Classements					
chaud, zéro, fort, avec	4	2	4	3	13	5
chaud, un, moyen, avec	2	8	1	9	10	8
chaud, deux, léger, sans	6	10	13	18	5	6
tiède, zéro, moyen, sans	13	13	10	5	2	12
tiède, un, léger, avec	14	16	17	12	16	9
tiède, deux, fort, avec	15	18	12	15	8	16
froid, zéro, léger, avec	7	3	14	2	18	2
froid, un, fort, sans	11	6	5	7	3	17
froid, deux, moyen, avec	10	11	6	13	12	7
chaud, zéro, léger, sans	3	1	11	4	6	4
chaud, un, fort, avec	1	7	2	10	7	14
chaud, deux, moyen, avec	5	12	3	17	9	13
tiède, zéro, fort, avec	17	14	16	6	11	18
tiède, un, moyen, sans	18	15	9	11	1	11
tiède, deux, léger, avec	16	17	18	16	15	10
froid, zéro, moyen, avec	8	4	8	1	14	1
froid, un, léger, avec	9	5	15	8	17	3
froid, deux, fort, sans	12	9	7	14	4	15

## SORTIE REGALS 1

Traitement Trade-Off  
Données THE.TOP

Effectif Total 6  
Nombre maximal d'itérations 3  
Niveau de Cohérence 0.8  
Classements Retenus 5  
Cohérence Globale Moyenne -0.922  
Cohérence Sélective Moyenne -0.876  
Corrélation linéaire -0.977

Prod	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
M/rg	5.20	6.00	10.40	8.60	15.00	13.60	8.80	6.40	10.40	5.00	5.40	9.20	12.80	10.80	16.40	7.00	10.80	9.20	
M/ut	0.50	0.45	0.34	0.35	0.15	0.14	0.36	0.44	0.30	0.48	0.45	0.37	0.28	0.29	0.06	0.44	0.30	0.36	
BANG																			SOM
1	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	0.00	5
2	1.00	1.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5
3	1.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00	1.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	5
4	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	5
5	0.00	0.00	1.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	1.00	0.00	5
6	0.00	0.00	1.00	0.00	0.00	0.00	3.00	1.00	1.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	5
7	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5
8	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.00	1.00	0.00	5
9	0.00	1.00	0.00	0.00	0.00	3.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	0.00	0.00	1.00	1.00	5
10	0.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5
11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00	0.00	0.00	1.00	1.00	1.00	0.00	0.00	0.00	5
12	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	1.00	5
13	1.00	0.00	1.00	2.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5
14	0.00	0.00	0.00	0.00	1.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	5
15	0.00	0.00	0.00	0.00	0.00	2.00	0.00	0.00	0.00	0.00	0.00	3.00	0.00	1.00	1.00	0.00	1.00	0.00	5
16	0.00	0.00	0.00	0.00	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	2.00	0.00	0.00	0.00	5
17	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00	1.00	0.00	1.00	5
18	0.00	0.00	1.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	5

Poids (%) des Facteurs & Utilités des Niveaux

TEMP SUCRE FORCE CITRON  
42.65801 28.06973 20.92075 8.35152

0.237691 0.157363 0.093730 0.012097

0.017063 0.104827 0.099526 0.071418

0.171826 0.018507 0.015952

## SORTIE REGALS 2

Traitement Trade-Off  
Données THE.TOF

Effectif Total 6  
Nombre maximal d'itérations 3  
Niveau de Cohérence 0.6  
Classements Retenus 6  
Cohérence Globale Moyenne -0.922  
Cohérence Sélective Moyenne -0.922  
Corrélation linéaire -0.969

Prod	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
M/rg	5.17	6.33	9.67	9.17	14.00	14.00	7.67	8.17	9.83	4.83	6.83	9.83	13.67	10.83	15.33	6.00	9.50	10.17	
M/ut	0.47	0.45	0.35	0.35	0.19	0.13	0.41	0.40	0.32	0.49	0.41	0.37	0.27	0.29	0.11	0.46	0.34	0.31	
RANG																			SOM
1	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00	0.00	1.00	0.00	2.00	0.00	0.00	6
2	1.00	1.00	0.00	1.00	0.00	0.00	2.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6
3	1.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00	1.00	0.00	1.00	0.00	0.00	0.00	0.00	1.00	0.00	6
4	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	6
5	1.00	0.00	1.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	1.00	0.00	6
6	0.00	0.00	2.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	6
7	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00	0.00	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	6
8	0.00	2.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.00	1.00	0.00	6
9	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	0.00	0.00	1.00	1.00	6
10	0.00	1.00	1.00	1.00	0.00	0.00	0.00	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	6
11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00	0.00	0.00	1.00	2.00	0.00	0.00	0.00	0.00	6
12	0.00	0.00	0.00	1.00	1.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	1.00	6
13	1.00	0.00	1.00	2.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	6
14	0.00	0.00	0.00	0.00	1.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	6
15	0.00	0.00	0.00	0.00	0.00	2.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	0.00	1.00	1.00	6
16	0.00	0.00	0.00	0.00	2.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	2.00	0.00	0.00	0.00	6
17	0.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	1.00	0.00	1.00	0.00	1.00	0.00	6
18	0.00	0.00	1.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00	0.00	0.00	0.00	6

Poids (\*) des Facteurs & Utilités des Niveaux

TEMP SUCRE FORCE CITRON  
39.53272 26.76548 25.61641 8.08539

0.214920 0.157106 0.078108 0.021339  
0.014220 0.095126 0.116620 0.059515  
0.166188 0.015423 0.061435



**Chapitre V**

**DESAGREGATION DE LA PREFERENCE**

**GLOBALE EN ANALYSE MULTICRITERE**

## V. DESAGREGATION DE LA PREFERENCE GLOBALE EN ANALYSE MULTICRITERE

La Désagrégation de la Préférence Globale est un cas particulier important de la régression qualitative, concernant l'explication du classement d'un ensemble d'actions jugées sur plusieurs critères, ordinaux ou quantitatifs monotones.

La première section pose le problème et décrit les méthodes, classées en deux catégories: régression et analyse canonique sous contraintes de positivité respectivement, des coefficients et des facteurs. La distinction proposée entre ces deux catégories permet de les définir dans chacune des deux approches, moindres écarts et moindres carrés.

Une description originale de la méthode UTA de Jacquet-Lagrèze et Siskos (1982), version améliorée par Siskos et Yannacopoulos (1983), nous amène à la considérer comme une «analyse canonique aux moindres écarts».

Nous proposons la méthode REGALS (REGression par ALgorithme als); elle est basée sur une théorie de l'analyse canonique de deux cônes polyédriques convexes, développée par Tenenhaus (1988).

Dans la deuxième section, une étude comparative des méthodes UTA et REGALS est effectuée à partir de trois exemples. Nous terminons par une synthèse des résultats et une conclusion fondée sur le fait que le problème de la désagrégation, selon chacune des approches moindres carrés et moindres écarts, ne peut être séparé de celui de l'agrégation.

## 1. LE PROBLEME ET LES METHODES

Deux formulations équivalentes du problème sont proposées. L'une est basée sur la recherche de fonctions d'utilité et l'autre sur la détermination de codages de variables ordinales.

Les méthodes de résolution sont classées en deux catégories, régression et analyse canonique. Cette dernière est étendue à l'approche moindres écarts.

La méthode UTA, selon une description originale qui en est faite, est une analyse canonique. Nous en présentons un exemple de mise en oeuvre.

Dans l'approche moindres carrés, la méthode REGALS que nous proposons est une analyse canonique des deux cônes polyédriques convexes, engendrés respectivement par le préordre à expliquer et l'ensemble des critères explicatifs.

### 1.1. LE PROBLEME

On considère un ensemble  $A$  de  $n$  actions, classées par ordre de préférence décroissante selon un préordre  $g_0$  à  $k_0$  modalités dit critère 0, et jugées sur une famille  $g$  de critères représentés par  $p$  variables  $g_j$  :

$$g_j(i) = \text{performance de l'action } i \text{ sur le critère } j$$

$$i=1,2,\dots,n; j=0,1,2,\dots,p.$$

Chaque critère  $g_j$  est à  $k_j$  modalités.

Notons  $M_j$  l'ensemble des  $k_j$  premiers entiers naturels. Le critère  $g_j$  prend ses valeurs dans

$$\mu_j(M_j) = \{\mu_j(1), \dots, \mu_j(k_j)\}$$

ensemble des  $k_j$  modalités ordonnées (notes de performance) du critère  $j$  et

$$g_j^* = \mu_j(1) \quad , \quad g_j^* = \mu_j(k_j).$$

Il s'agit de déterminer les fonctions optimales d'utilité  $U_j$ , monotones:

$$U_j \circ \mu_j (M_j) = \{U_j(\mu_j(1)), \dots, U_j(\mu_j(k_j))\}, j=0,1,2,\dots,p$$

avec  $U_j \circ \mu_j$  croissante et, pour  $j \geq 1$ , la normalisation:

$$U_j(\mu_j(1)) = U_j(g_{j*}) = 0, \quad U_j(\mu_j(k_j)) = U_j(g_{j*}) = 1;$$

$$\beta_j \geq 0, \quad \sum \{\beta_j, j=1, \dots, p\} = 1;$$

enfin, l'agrégation en un critère unique:

$$U(g) = U(g_1, g_2, \dots, g_p) = \sum \{\beta_j U_j(g_j), j \geq 1\}.$$

Le but recherché est le suivant: les préordres  $g_0$ ,  $U_0(g_0)$  et  $U(g)$  doivent être de structures ordinales très proches et la corrélation linéaire  $\text{cor}(g_0, U(g))$  suffisamment élevée pour que les poids  $\beta_j$  puissent être interprétés.

#### Remarque

Le problème admet une formulation équivalente, en termes de recherche de codages ordinaux (section III.1.2):

Soit  $X_j$  l'application définie sur  $A$  et prenant ses valeurs dans  $M_j$  telle que  $\mu_j \circ X_j = g_j$ .  $X_j$  est la variable ordinale associée à  $g_j$ . Notons que  $\mu_0$  est l'identité de  $M_0$  dans  $M_0$ :

$$\mu_0(k) = k, \quad k=1, \dots, k_0,$$

car  $g_0$  est ordinale.

La variable  $C_j = U_j \circ \mu_j$  est une application de  $M_j$  dans  $\mathbb{R}$  et

$$m_1 < m_2 \quad \text{équivalent à} \quad C_j(m_1) \leq C_j(m_2)$$

donc  $C_j$  est un codage des modalités de  $X_j$  et

$$C_j \circ X_j = (U_j \circ \mu_j) \circ X_j = U_j \circ (\mu_j \circ X_j) = U_j \circ g_j$$

est un codage de la variable ordinale  $X_j$ .

Déterminer les fonctions d'utilité optimales  $U_j$  des  $g_j$ , revient donc à rechercher les codages optimaux  $C_j$  des modalités des  $X_j$ .

Finalement, nous posons le problème à résoudre comme suit:

$$\text{MIN} \|U_0(g_0) - U(g)\|$$

SOUS LES CONTRAINTES

$$(1) U_0(g_0) \in \text{COD}(X_0); U(g) \in \text{COD}(X_1, X_2, \dots, X_p)$$

$$(2) U_j(\mu_j(1)) = U_j(g_j^*) = 0, U_j(\mu_j(k_j)) = U_j(g_j^*) = 1$$

$$(3) \beta_j \geq 0, \sum \beta_j = 1.$$

## 1.2.METHODES DE RESOLUTION

On peut aborder ce problème dans une approche :

1. "moindres écarts":  $R^n$  est muni de la norme

$$\|u\| = \sum \{|u_i|, i=1, \dots, n\}$$

La méthode UTA de Jacquet-Lagrèze et Siskos (1982), en constitue un modèle.

2. "moindres carrés" :  $R^n$  est muni de la norme

$$\|u\| = \sqrt{\sum \{u_i^2, i=1, \dots, n\}}$$

On peut poser le problème comme:

-une régression: en fixant  $U_0(g_0) = g_0$

-une analyse canonique:  $U_0(g_0) \in \text{COD}(X_0)$ .

Selon cette distinction, on classera la méthode UTA dans cette dernière catégorie.

## 1.2.1.LA METHODE UTA

Nous allons décrire la version de la méthode UTA proposée par Siskos et Yannacopoulos (1983), et en présenter un exemple d'application.

## 1.2.1.1.DESCRPTION

On distingue plusieurs étapes:

1. On détermine pour chaque intervalle  $[g_j^* \quad g_j^*]$ , un ensemble de niveaux  $\Omega_j(N_j)$ ,  $N_j = \{1, 2, \dots, l_j\}$ , à partir de  $\mu_j(M_j)$ , ceci pour doter  $g_j$ ,  $j \geq 1$ , d'échelons significatifs, en nombre restreint (technique du point de valeur médiane). Cette étape amène à écrire:

$$U_{og}(i) = \sum \{ \sum \{ c_{1,j,k} U_j \circ \Omega_j(k), k=1, \dots, l_j \}, j=1, \dots, p \}.$$

On impose à  $U_0(g_0)$  d'être un codage parfait de  $X_0$ :

$$\tau(g_0, U_0(g_0)) = 1$$

et aux  $U_j$ ,  $j \geq 1$ , de vérifier:

$$U_j \circ \Omega_j(k+1) - U_j \circ \Omega_j(k) = w_{j,k} \geq 0, \text{ pour } k=1, \dots, l_j-1, j \geq 1 \quad (1)$$

$$\text{et } U_j \circ \Omega_j(1) = 0, \text{ pour tout } j \geq 1.$$

Les poids  $\beta_j$  seront déterminés a posteriori, après normalisation des utilités optimales donnant  $U_j \circ \Omega_j(l_j) = 1$ , pour tout  $j$ , ce qui amène à la condition de normalisation:

$$\sum \{ U_j \circ \Omega_j(l_j), j=1, \dots, p \} = 1; \quad (2)$$

d'après (1) :  $U_j \circ \Omega_j(k) = \sum \{ w_{j,m}, m=1, \dots, k-1 \}$ , pour  $k > 1$ , d'où

$$U_{og}(i) = \sum \{ \sum \{ c_{1,j,k} \sum \{ w_{j,m}, m=1, \dots, k-1 \}, k=1, \dots, l_j \}, j=1, \dots, p \} \quad (3)$$

$$\text{D'après (2) : } \sum \{ \sum \{ w_{j,m}, m=1, \dots, l_j-1 \}, j=1, \dots, p \} = 1 \quad (4)$$

2. On introduit une borne  $\delta$  et une double fonction d'erreurs positives  $\sigma = (\sigma^+, \sigma^-)$  définie comme suit sur l'ensemble  $A$ :

$$\sigma^+(i) = \frac{1}{2}[|U_0 \circ g_0(i) - U_{og}(i)| + (U_0 \circ g_0(i) - U_{og}(i))]$$

$$\sigma^-(i) = \frac{1}{2}[|U_0 \circ g_0(i) - U_{og}(i)| - (U_0 \circ g_0(i) - U_{og}(i))].$$

Notons que  $|U_0 \circ g_0(i) - U_{og}(i)| = \sigma^+(i) + \sigma^-(i)$  (5)

et pour tout  $i$   $\sigma^+(i)\sigma^-(i) = 0$ ;

$\sigma^+ - \sigma^-$  représente, pour chaque action  $i$ , l'écart entre l'utilité estimée  $U_{og}(i)$  et celle qu'il conviendrait de lui attribuer pour préserver son rang dans le préordre  $U_0(g_0)$ , identique à  $g_0$  au seuil de préférence  $\delta$ :

$$U_0 \circ g_0(i) = U_{og}(i) + \sigma^+(i) - \sigma^-(i)$$

avec  $\sigma^+(i) \geq 0$ ,  $\sigma^-(i) \geq 0$ ,  $i=1, \dots, n$  (6)

et pour toute paire d'actions consécutives:

$$D(i, i+1) = [U_{og}(i) + \sigma^+(i) - \sigma^-(i)] - [U_{og}(i+1) + \sigma^+(i+1) - \sigma^-(i+1)]$$

avec  $D(i, i+1) \geq \delta$  si  $X_0(i) > X_0(i+1)$  (7)

et  $D(i, i+1) = 0$  si  $X_0(i) = X_0(i+1)$

On écrira  $D(i, i+1)$  à l'aide de la relation (3).

3. On résoud le programme linéaire

$$\text{MIN}(\sum\{|U_0 \circ g_0(i) - U_{og}(i)|, i=1, \dots, n\})$$

$$\text{SOUS LES CONTRAINTES (7); (6); (4); (1).}$$

D'après la relation (5), ce programme est équivalent à:

$$\text{MIN}(\sum\{\sigma^+(i) + \sigma^-(i), i=1, \dots, n\})$$

$$\text{SOUS LES CONTRAINTES (7); (6); (4); (1);}$$

4. On effectue une analyse post-optimale dans le but

d'étudier les problèmes de dégénérescence et d'instabilité des codages et améliorer la dimension explicative de la méthode par la prise en compte du critère statistique de corrélation linéaire.

#### 1.2.1.2. EXEMPLE DE MISE EN OEUVRE

Nous considérons l'exemple de Siskos et Yannacopoulos (1983), concernant l'analyse du choix par un individu du mode de transport domicile-lieu de travail, durant les heures d'affluence.

Trois critères sont retenus: PRIX (en francs), TEMPS (du trajet, en minutes) et CONFORT (ordinal à quatre modalités).

MODE DE TRANSPORT	RANG ( $X_0$ )	PRIX (FRS)	TEMPS (MNS)	CONFORT (RANGS)
RER	1	3	10	3
METRO 1 <sup>ERE</sup>	2	4	20	2
METRO 2 <sup>EME</sup>	2	2	20	4
BUS	3	6	40	4
TAXI	4	30	30	1

#### Analyse pré-optimale

1. On détermine pour chaque intervalle  $[g_j^* \quad g_j^*]$  un ensemble de niveaux  $\Omega_j(N_j)$ ,  $N_j = \{1, 2, \dots, l_j\}$ , à partir de  $\mu_j(M_j)$ :

$$[g_1^* \quad g_1^*] : \Omega_1(N_1) = \{30, 16, 2\}, l_1 = 3;$$

$$[g_2^* \quad g_2^*] : \Omega_2(N_2) = \mu_2(M_2) = \{40, 30, 20, 10\}, l_2 = 4;$$

$$[g_3^* \quad g_3^*] : \Omega_3(N_3) = \mu_3(M_3) = \{4, 3, 2, 1\}, l_3 = 4;$$

On détermine les  $c_{1jk}$  tels que

$$U_{og}(i) = \sum \{ \sum \{ c_{1jk} U_j \circ \Omega_j(k), k=1, \dots, l_j \}, j=1, \dots, p \}.$$

La technique de l'interpolation linéaire est utilisée.



Pour  $i, j$  fixés, les  $c_{ijk}$  sont nuls pour tout  $k$  différent d'une valeur  $k_0$  telle que  $g_j(i) = Q_j(k_0)$ :  $c_{ijk_0}$  égale à 1; ou  $k$  différent de deux valeurs consécutives  $k_0$  et  $k_0+1$  telles que  $Q_j(k_0)$  et  $Q_j(k_0+1)$  encadrent  $g_j(i)$ :  $c_{ijk_0} + c_{ijk_0+1} = 1$ .

Dans notre cas, pour tout  $j \geq 2$ ,  $i = 1, \dots, 5$ :

$$c_{ijk} = 1 \text{ si } g_j(i) = Q_j(k) \text{ et } c_{ijk} = 0 \text{ sinon.}$$

Détaillons pour  $i=j=1$ :  $g_1(1) = 3$ ; comme  $Q_1(1) = 30$ ,  $c_{111} = 0$ .

On a l'encadrement  $16 > 3 > 2$  et par interpolation linéaire attribuant à la borne la plus proche, la valeur la plus élevée:

$$c_{112} = 0.07 \text{ et } c_{113} = 0.93.$$

On a finalement les  $c_{ijk}$  (indice  $i$  en ligne,  $j=1, k$  en colonne):

	1	2	3	4	5
1	0	0	0	0	1
2	0.07	0.14	1	0.29	0
3	0.93	0.86	0	0.71	0

On définit  $w_{jm} \geq 0$ ,  $j = 1, \dots, 3$  et  $m = 1, \dots, l_j - 1$  tels que

$$U_{og}(i) = \sum \{ \sum \{ c_{ijk} \sum \{ w_{jm}, m=1, \dots, k-1 \}, k=1, \dots, l_j \}, j=1, \dots, p \}$$

$$\text{et } \sum \{ \sum \{ w_{jm}, m=1, \dots, l_j - 1 \}, j=1, \dots, p \} = 1$$

$$\text{c'est à dire } U_{og}(1) = w_{11} + 0.93w_{12} + w_{21} + w_{22} + w_{23} + w_{31}$$

$$\text{et } U_{og}(2) = w_{11} + 0.86w_{12} + w_{21} + w_{22} + w_{31} + w_{32}$$

$$Uog(3) = w_{11} + w_{12} + w_{21} + w_{22}$$

$$Uog(4) = w_{11} + 0.71w_{12}$$

$$Uog(5) = w_{21} + w_{31} + w_{32} + w_{33}$$

et  $w_{11} + w_{12} + w_{21} + w_{22} + w_{23} + w_{31} + w_{32} + w_{33} = 1;$

on effectue la comparaison par paire au seuil  $\delta=0.05$ :

$$D(i,i+1) = [Uog(i) + \sigma^+(i) - \sigma^-(i)] - [Uog(i+1) + \sigma^+(i+1) - \sigma^-(i+1)]$$

avec  $D(i,i+1) \geq 0.05$  si  $X_0(i) > X_0(i+1)$

et  $D(i,i+1) = 0$  si  $X_0(i) = X_0(i+1).$

On obtient le programme linéaire:

$w_{11}$	$w_{12}$	$w_{21}$	$w_{22}$	$w_{23}$	$w_{31}$	$w_{32}$	$w_{33}$	$\sigma^+_1$	$\sigma^+_2$	$\sigma^+_3$	$\sigma^+_4$	$\sigma^+_5$	$\sigma^-_1$	$\sigma^-_2$	$\sigma^-_3$	$\sigma^-_4$	$\sigma^-_5$	sgn	2 <sup>nd</sup> nombre
0	0.07	0	0	1	0	-1	0	1	-1	0	0	0	-1	1	0	0	0	$\geq$	0.05
0	-0.14	0	0	0	1	1	0	0	1	-1	0	0	0	-1	1	0	0	=	0
0	0.29	1	1	0	0	0	0	0	0	1	-1	0	0	0	-1	1	0	$\geq$	0.05
1	0.71	-1	0	0	-1	-1	-1	0	0	0	1	-1	0	0	0	-1	1	$\geq$	0.05
1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	=	1
Fonction à minimiser								1	1	1	1	1	1	1	1	1	1	=	F

La résolution de ce programme linéaire donne  $F^*=0$  donc une restitution parfaite du préordre initial ( $\tau(Uog, g_0)=1$ ) et

$$w_{11}=0.5, w_{12}=0;$$

$$w_{21}=0.05, w_{22}=0, w_{23}=0.05;$$

$$w_{31}=0, w_{32}=0, w_{33}=0.4.$$

Le tableau suivant donne les utilités (d'après la relation (1)) après normalisation:

utilité ( $U_j$ )	modalité (k)				poïds ( $\beta_j$ )
	1	2	3	4	
$U_1$	0	1	1		0.5
$U_2$	0	0.5	0.5	1	0.1
$U_3$	0	0	0	1	0.4

#### Analyse post-optimale

Trois programmes linéaires, découlant des maximisations respectives des poids des trois critères et imposant aux erreurs d'être nulles, sont résolus. Les trois solutions obtenues, très contrastées, prouvent l'instabilité des codages. La solution retenue est le centre de gravité des codages résultant de ces solutions (Siskos, 1984).

#### 1.2.2. REGRESSION SOUS CONTRAINTES

Soit  $\{Z_{jk}, k=1, \dots, k_j\}$  le système générateur de  $COD(X_j)$  (voir section III.1.2).

Remarquons que pour tout  $j$ ,  $Z_{j1}=Z_1$  est le vecteur unité.

D'où  $COD(X_1, X_2, \dots, X_p) = L(Z_1) + C(U\{Z_{jk}, k=2, \dots, k_j, j=1, \dots, p\})$ .

Nous posons le problème :  $\text{MIN}\{\|X_0 - Z\|, Z \in \text{COD}(X_1, X_2, \dots, X_p)\}$ ;

soit  $Z^0$  la solution de ce problème:

$$Z^0 = \alpha^0_1 Z_1 + \sum_{j=1, \dots, p} \{ \sum_{k=2, \dots, k_j} \alpha^0_{jk} Z_{jk} \};$$

posons  $\Sigma_0 = \Sigma\{\Sigma\{\alpha^0_{jk}, k=2, \dots, k_j\}, j=1, \dots, p\}$ .

Soit  $Z^1 = (Z^0 - \alpha^0_{j1} Z_1) / \Sigma_0 = \Sigma\{\Sigma\{\alpha^1_{jk} Z_{jk}, k=2, \dots, k_j\}, j=1, \dots, p\}$

avec  $\alpha^1_{jk} = \alpha^0_{jk} / \Sigma_0$ ,  $k=2, \dots, k_j$ ,  $j=1, \dots, p$

posons  $\beta_j = \Sigma\{\alpha^1_{jk}, k=2, \dots, k_j\}$

et  $\alpha^2_{jk} = \alpha^1_{jk} / \beta_j$ ,  $k=2, \dots, k_j$  et  $\alpha^2_{j1} = 0$ ,  $j=1, \dots, p$ ;

on note  $U_j(g_j) = \{\Sigma\{\alpha^2_{jk} Z_{jk}, k=1, \dots, k_j\}$ .

On a  $U_0 g = \Sigma\{\beta_j U_j(g_j), j=1, \dots, p\}$

avec  $U_j(\mu_j(1)) = U_j(g_j^*) = \alpha^2_{j1} = 0$

(2)

$U_j(\mu_j(k_j)) = U_j(g_j^*) = \Sigma\{\alpha^2_{jk}, k=2, \dots, k_j\} = 1$

et (3)  $\beta_j \geq 0$ ,  $\Sigma\beta_j = 1$ .

### 1.2.3. LA METHODE REGALS

Basée sur la théorie développée par Tenenhaus (1988), REGALS (REGression par ALgorithme ALS) est une analyse canonique, sur le principe des «projections alternées», des cônes polyédriques convexes  $COD(X_0)$  et  $COD(X_1, X_2, \dots, X_p)$  engendrés respectivement par la variable ordinaire  $X_0$  et les variables ordinaires  $X_j$ , associées aux critères  $g_j$ .

La méthode REGALS est basée sur l'utilisation alternative de deux algorithmes performants:

PAV: (Barlow & al., 1972) détermine la projection orthogonale sur le cône polyédrique convexe engendré par le préordre  $X_0$  (section IV.2.2.2).

MCCP: version optimisée de REGCAZ que nous proposons (voir section I.2.1.3.3), elle détermine la projection orthogonale d'un vecteur sur  $COD(X_1, X_2, \dots, X_p)$ .

Comme REGALS est initialisée avec le préordre  $X_0$ , la régression

sous contraintes proposée ci-dessus en constitue, en fait, une étape, à la différence que la normalisation décrite n'est effectuée qu'à la fin processus.

## 2. ETUDE COMPARATIVE UTA/REGALS

Nous l'effectuons sur trois exemples traités par R. Savouandaradja (1987), dans le cadre d'une étude comparative des méthodes UTA et MORALS, et dont nous reprenons les résultats.

### 2.1. EXEMPLES ET RESULTATS

Nous présentons, ci-dessous, les trois exemples étudiés et pour chacun, le tableau des données initiales, la sortie du programme REGALS et une synthèse regroupant les résultats obtenus par les méthodes UTA, MORALS et REGALS.

#### I. Evaluation des projets stratégiques d'une entreprise

De J.L. Richard (cité par Jacquet-Lagrèze et Siskos, 1983).

Dix-sept projets stratégiques sont notés sur quatre critères : compétitivité commerciale (COMPCOM), efficacité commerciale (EFFICOM), efficacité financière (EFFINAN), flexibilité commerciale (FLEXCOM).

#### II. Octroi de crédits

De B. Roy (aide à la décision, critères multiples et optimisation pour choisir, trier et ranger, polycopies de cours de DEA, extrait de Siskos (1979), repris par Roy (1985, pp 108, 109).

Un responsable fonde sa décision (octroi de crédits) sur cinq ratios : RATIO1, RATIO2, RATIO3, RATIO4 et RATIO5, automatiquement calculés à partir des informations contenus dans les dossiers de demande. Les limites A1, A2 du «sûrement acceptable» (resp. A3, A4 du «sûrement inacceptable») sont deux combinaisons de différents niveaux de ratios relatives à deux populations principales de demandeurs ( $D_{13}$ ,  $D_{24}$ ) et sont difficilement comparables mais elles seront considérées comme des «actions» indifférentes.

Cinq dossiers sont classés en fonction de ces ratios. On cherche à expliquer le préordre (des neuf actions) par les cinq ratios.

III. Sélection de points de vente (Etude du lancement d'un nouveau produit, Siskos, 1983).

Cinquante points de vente sont classés selon douze critères: qualité de la localisation ( $g_1=LOCAL$ ), degré de présentation ( $g_2=PRSENTA$ ), existence d'un stand ( $g_3=STAND$ ), dynamisme1 ( $g_4=DYNAM1$ ), dynamisme2 ( $g_5=DYNAM2$ ), adéquation à l'innovation ( $g_6=INNOVAT$ ), agressivité publicitaire et commerciale ( $g_7=PUBLIC$ ), mode de règlement ( $g_8=REGLEMT$ ), qualité d'exercice de vente ( $g_9=EXRVTE$ ), service après-vente ( $g_{10}=SAPVTE$ ), type de produits commercialisés ( $g_{11}=PROCOMM$ ), et chiffre d'affaires en 1979 en millier de francs ( $g_{12}=CA1979$ ).

Cette étude est intéressante par son ampleur.

Notons que l'on doit effectuer un découpage du critère  $g_{12}$  ( $CA1979$ ) en classes permettant de le considérer comme ordinal avec un nombre restreint de modalités:

-Dans  $UTA, \Omega_j = \mu_j, j=1, \dots, 11$  et

$\Omega_{12} = \{13.161, 554.662, 1096.163, 1637.664, 2179.165\}$

-Dans  $REGALS$ , on retient 4 niveaux significatifs de  $CA1979$  découlant du découpage en classes suivant:

rang	classe
4	[ 13.161 625.500[
3	[ 625.500 1041.526[
2	[1041.526 2179.165[
1	[2179.165 3000.000[

## I. EVALUATION DE PROJETS

PROJETS STRATEG. SUR LE CHAMP 1	g <sub>1</sub>	g <sub>2</sub>	g <sub>3</sub>	g <sub>4</sub>
PP1-A1 Disponibilité pour le 31/12/1980 dans tte agglo. de plus de 20 000 hab...	3.75	1.50	2.00	2.00
PP1-A2 Nouveau mailing à 25 000 profes- sionnels (dessinat., BET, secrét).....	3.75	1.75	1.75	2.00
PP1-A3 même que préc. à 100 000 prof.....	4.00	1.50	1.25	2.00
PP1-A4 Prospection auprès des fournis- sistes et conditions de lancement.....	3.75	1.75	1.50	2.25
PP1-A5 Vte directe par mail. aux 2000 premières entreprises.....	3.50	1.50	1.75	2.25
PP1-A6 Promo. d'assistance au stockage...	3.75	1.75	2.00	2.00
PP1-A7 Présence sur 2 foires nationales.	3.50	1.75	2.00	2.00
PP1-A8 Système d'aide au réassort.....	4.25	2.25	2.00	2.25
PP1-A9 Conditionnement double.....	3.50	2.25	2.00	2.00
PP1-A10 Concours détaillant (prod. prof.)	3.75	1.50	1.50	2.25
PP1-A11 Présence sur 3 foires région....	3.50	1.25	1.25	2.00
PP1-A12 Création d'un dépôt régional de vente directe.....	3.75	1.75	1.00	2.25
PP1-A13 Démonstration par équipes sur lieu de vente.....	4.00	1.00	1.25	2.25
PP1-A14 Vte en grande surf. (prod. prof)...	3.50	1.00	1.50	2.00
PP1-A15 Relèvement important des prix...	2.50	2.00	2.00	2.00
PP1-A16 Création d'un dépôt régional de revente classique.....	3.50	1.00	1.00	2.25
PP1-A17 Publicité télévisée.....	4.50	0.75	0.75	2.00

g<sub>1</sub> = COMPCOMg<sub>2</sub> = EFFICOMg<sub>3</sub> = EFFINANg<sub>4</sub> = FLEXCOM

## SORTIE REGALS I

Traitement      Regals  
Données          EVALPROJ.ALS

Préordre à expliquer    : RANG  
Tau de Kendall          = 0.8456 (Tau( $\alpha=5\%$ ) = 0.3499)  
Corrélation linéaire    = 0.9965  
Nombre d'itérations    = 5

## POIDS (%) des Critères &amp; UTILITES des Niveaux

COMPCOM    EFFICOM    EFFINAM    FLEXCOM  
46.22129    26.06577    27.71294    0.00000

0.000000 0.000000 0.000000 0.000000  
0.682105 0.559794 0.000000 0.000000  
0.849761 0.725082 0.498878  
0.849761 0.725082 0.605150  
1.000000 1.000000 0.605150  
1.000000 1.000000 1.000000  
1.000000

R/A	ACTION	UTILITE	N/R/A
1	A8	1.00000	1
2	A6	0.93056	2
3	A1	0.85890	3
4	A7	0.85306	4
5	A4	0.82113	5
6	A9	0.85306	4
7	A2	0.82113	5
8	A3	0.72002	7
9	A10	0.74947	6
10	A11	0.64253	11
11	A12	0.65343	10
12	A13	0.67694	8
12	A14	0.62890	12
13	A5	0.67198	9
14	A15	0.53779	13
15	A16	0.46119	15
16	A17	0.46221	14



## SYNTHESE I

## Classements

INIT	RANG	UTA	MORALS	REGALS
A8	1	1	1	1
A6	2	3	2	2
A1	3	4	3	3
A7	4	3	6	4
A4	5	5	4	5
A9	6	2	6	4
A2	7	5	5	5
A3	8	6	8	7
A10	9	6	7	6
A11	10	7	13	11
A12	11	8	11	10
A13	12	9	9	8
A14	12	9	12	12
A5	13	6	10	9
A15	14	10	14	13
A16	15	11	15	15
A17	16	12	16	14

## Poids des Critères &amp; Indicateurs de qualité

CRITERE	UTA	MORALS	REGALS
COMPCOM	0.495	6.137	0.462
EFFICOM	0.460	5.040	0.261
EFFINAN	0.045	6.032	0.277
FLEXCOM	0.000	-3.338	0.000
Tau (Kendall)	0.8200	0.8235	0.8456
Corrélation linéaire		0.9989	0.9965
Nombre d'itérations		10	5

## II.OCTROI DE CREDITS

ACTIONS		RATIO1	RATIO2	RATIO3	RATIO4	RATIO5
Limites du «sûrement acceptable»	A1	0.7	0.7	0.5	0.5	0.5
	A2	0.5	0.5	0.7	0.7	0.7
Limites du «sûrement inacceptable»	A3	0.5	0.5	0.5	0.5	0.5
	A4	0.5	0.5	0.3	0.7	0.5
cinq	A5	0.7	0.7	0.6	0.5	0.5
dossiers	A6	0.5	0.5	0.6	0.7	0.7
de	A7	0.7	0.7	0.3	0.7	0.5
demande	A8	0.6	0.6	0.3	0.7	0.1
de crédits	A9	0.6	0.6	0.6	0.5	0.1

## SORTIE REGALS II

Traitement      Regals  
Données          OCTCRED.ALS

Préordre à expliquer    : RANG  
Tau de Kendall          = 0.8889 (Tau(a=5%) = 0.5222)  
Corrélation linéaire    = 0.9999  
Nombre d'itérations    = 8

## POIDS (%) des Critères &amp; UTILITES des Niveaux

RATIO1	RATIO2	RATIO3	RATIO4	RATIO5
22.60471	0.00000	22.31013	21.37217	34.31299

0.000000	0.000000	0.000000	0.000000	0.000000
1.000000	0.000000	0.967085	1.000000	1.000000
1.000000	0.000000	0.973817		1.000000
		1.000000		

R/A	ACTION	UTILITE	N/R/A
1	A5	0.78044	1
2	A1	0.77893	3
2	A2	0.77995	2
3	A6	0.77411	5
3	A7	0.77690	4
4	A3	0.55889	6
4	A4	0.55685	7
5	A8	0.43377	9
5	A9	0.43731	8

## SYNTHESE II

## Classements

INIT	RANG	UTA	MORALS	REGALS
A5	1	1	1	1
A1	2	2	1	3
A2	2	4	3	2
A6	3	4	3	5
A7	3	3	1	4
A3	4	5	4	6
A4	4	6	4	7
A8	5	8	2	9
A9	5	7	2	8

## Poids des Critères &amp; Indicateurs de qualité

CRITERE	UTA	MORALS	REGALS
RATIO 1	0.395	2424.420	0.220
RATIO 2	0.000	-2093.460	0.000
RATIO 3	0.210	0.223	0.223
RATIO 4	0.185	-5.642	0.214
RATIO 5	0.210	9.242	0.343
Tau (Kendall)	0.8100	0.3333	0.8889
Corrélation linéaire		0.9999	0.9999
Nombre d'itérations		7	8

## III. SELECTION DE POINTS DE VENTE

POINTS DE VENTE	g <sub>1</sub>	g <sub>2</sub>	g <sub>3</sub>	g <sub>4</sub>	g <sub>5</sub>	g <sub>6</sub>	g <sub>7</sub>	g <sub>8</sub>	g <sub>9</sub>	g <sub>10</sub>	g <sub>11</sub>	g <sub>12</sub> (KF)
1	2	4	1	5	4	3	2	3	4	3	3	1024.526
2	4	5	3	5	1	2	3	3	4	3	3	816.230
3	4	5	1	5	2	2	3	3	3	2	2	65.470
4	1	3	1	5	4	3	3	3	4	3	3	63.145
5	4	3	1	5	4	2	3	3	4	3	3	328.595
6	2	2	1	3	4	3	3	3	4	3	3	372.420
7	4	5	1	5	4	3	3	2	3	3	3	164.387
8	2	3	1	2	2	1	2	3	4	3	2	89.105
9	2	4	1	4	4	2	3	3	4	2	3	761.874
10	4	5	1	5	4	2	3	3	4	3	2	79.113
11	2	4	3	5	4	3	3	3	3	2	3	182.270
12	1	2	1	3	4	2	3	3	3	3	2	163.780
13	4	4	2	4	4	3	1	3	1	2	3	157.281
14	1	5	3	5	3	3	5	3	3	3	3	342.141
15	2	4	3	2	4	2	5	3	3	3	3	195.471
16	1	4	1	5	4	1	2	3	4	3	2	150.000
17	2	4	3	3	4	2	3	3	3	3	2	326.825
18	2	4	1	3	4	1	2	3	4	3	2	306.631
19	4	4	3	5	4	3	1	3	3	2	3	625.500
20	2	2	1	3	4	1	3	3	3	3	3	187.641
21	2	4	1	4	4	1	3	3	3	3	2	70.000
22	3	4	1	4	4	1	2	3	3	2	2	43.395
23	4	4	1	4	4	1	1	3	3	3	2	300.000
24	2	4	1	5	2	2	2	3	4	3	2	30.000
25	2	5	1	5	3	2	5	3	4	3	3	122.476
26	2	4	1	2	3	3	1	3	1	2	2	150.000
27	4	3	1	5	4	3	3	3	4	3	3	115.475
28	2	4	2	4	4	1	3	3	2	3	3	392.000
29	1	4	3	5	3	3	3	3	3	2	3	222.457
30	2	4	3	5	3	3	1	2	2	2	3	401.545
31	2	4	3	4	3	2	5	3	3	3	3	124.843
32	2	4	1	1	4	1	2	3	1	3	2	250.000
33	1	4	3	1	3	1	2	3	1	3	3	2179.165
34	2	4	3	2	3	1	2	2	1	2	3	729.000
35	2	4	3	4	3	1	2	3	1	3	2	300.000

## III. SELECTION DE POINTS DE VENTE (suite)

36	1	4	1	2	4	1	3	3	4	2	3	821.739
37	1	4	1	5	1	1	2	3	1	2	1	13.600
38	1	1	1	2	4	2	1	2	2	3	3	127.661
39	1	4	1	4	3	1	1	3	1	2	2	210.000
40	2	4	3	5	3	3	2	3	3	2	3	250.865
41	1	3	1	2	4	2	1	3	1	1	3	172.440
42	1	2	3	4	4	2	3	3	3	3	3	114.207
43	2	2	1	4	4	2	2	3	1	2	3	101.740
44	2	1	3	4	3	2	2	3	1	2	3	145.251
45	4	5	1	4	2	3	3	3	4	2	3	30.570
46	2	3	1	4	2	1	2	3	3	3	2	13.161
47	2	4	1	4	4	3	3	3	3	2	3	57.440
48	1	3	1	2	4	2	2	3	3	3	3	54.330
49	2	4	1	5	4	3	3	3	3	3	2	139.974
50	1	4	1	4	4	1	1	3	1	2	1	155.000

---

## SORTIE REGALS III

Traitement Regals  
Données LANPROD.ALS

Préordre à expliquer : RANG  
Tau de Kendall = 0.4269 ( $T_{\alpha}(n=54) = 0.1913$ )  
Corrélation linéaire = 0.8083  
Nombre d'itérations = 11

## POIDS (%) des Critères &amp; UTILITES des Niveaux

LOCAL	PRESENTA	STAND	DYNAM1	DYNAM2	INNOVAT	PUBLIC	REGLENT	EXRYTE	SAPYTE	PROCOMM	CA1979
8.98370	3.08850	0.06748	16.77920	0.00000	0.00000	4.02474	2.17089	3.33925	4.14373	34.50948	22.89303
0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
0.476196	1.000000	1.000000	0.513827	0.000000	0.000000	0.833679	1.000000	0.000000	0.000000	1.000000	0.469418
1.000000	1.000000	1.000000	0.513827	0.000000	0.000000	1.000000		0.000000	1.000000	1.000000	1.000000
1.000000	1.000000		0.513827	0.000000		1.000000		1.000000			1.000000
	1.000000		1.000000								

## R/A \* ACTION \* UTILITE \* N/R/A

1	A1	0.94557	1
1	A2	0.83710	2
2	A3	0.69557	10
2	A4	0.68056	12
2	A5	0.77039	3
3	A6	0.64176	20
3	A7	0.71529	7
3	A8	0.63507	21
3	A9	0.70779	8
4	A10	0.77039	3
4	A11	0.64918	15
4	A12	0.56559	33
5	A13	0.57442	30
5	A14	0.64784	16
6	A15	0.60904	23
6	A16	0.67386	13
7	A17	0.60904	23
7	A18	0.63507	21
8	A19	0.76346	4
8	A20	0.60837	24
8	A21	0.60837	24
9	A22	0.60730	25
9	A23	0.61518	22
9	A24	0.71664	6
9	A25	0.72334	5
10	A26	0.52668	37

## SORTIE REGALS III (suite)

10	A27	0.77039	3
11	A28	0.60904	25
11	A29	0.60640	26
11	A30	0.58723	29
11	A31	0.60904	23
12	A32	0.51546	38
13	A33	0.70228	9
13	A34	0.64667	18
13	A35	0.60235	27
13	A36	0.66501	14
14	A37	0.25394	41
15	A38	0.47275	40
16	A39	0.48390	39
16	A40	0.64249	19
16	A41	0.48390	39
17	A42	0.56626	32
17	A43	0.56024	34
17	A44	0.53003	36
18	A45	0.64738	17
18	A46	0.60168	28
18	A47	0.56693	31
18	A48	0.55890	35
18	A49	0.68995	11
19	A50	0.13881	42



## SYNTHESE III

## Classements

INIT	RANG	UTA	MORALS	REGALS
A1	1	3	13	1
A2	1	5	4	2
A3	2	10	11	10
A4	2	3	25	12
A5	2	2	3	3
A6	3	5	14	20
A7	3	7	7	7
A8	3	6	20	21
A9	3	8	12	8
A10	4	2	5	3
A11	4	5	2	15
A12	4	11	40	33
A13	5	13	15	30
A14	5	2	28	16
A15	6	4	16	23
A16	6	1	32	13
A17	7	5	26	23
A18	7	6	18	21
A19	8	6	6	4
A20	8	10	22	24
A21	8	10	26	24
A22	9	14	9	25
A23	9	11	9	22
A24	9	3	19	6

## SYNTHESE III (suite)

A25	9	1	8	5
A26	10	17	36	37
A27	10	2	1	3
A28	11	5	29	23
A29	11	6	35	26
A30	11	6	33	29
A31	11	4	17	23
A32	12	20	38	38
A33	13	10	42	9
A34	13	13	34	18
A35	13	10	39	27
A36	13	9	26	14
A37	14	15	46	41
A38	15	16	43	40
A39	16	18	44	39
A40	16	6	23	19
A41	16	19	41	39
A42	17	6	36	32
A43	17	17	30	34
A44	17	17	31	36
A45	18	8	2	17
A46	18	11	27	28
A47	18	13	21	31
A48	18	12	37	35
A49	18	7	24	11
A50	19	18	45	42

## SYNTHESE III (suite)

## Poids des Critères &amp; Indicateurs de qualité

CRITERE	UTA	MORALS	REGALS
LOCAL	0.001	1.833	0.090
PRSENTA	0.005	-0.038	0.031
STAND	0.005	0.001	0.001
DYNAM1	0.480	0.043	0.168
DYNAM2	0.000	0.197	0.000
INNOVAT	0.000	0.068	0.000
PUBLIC	0.002	0.311	0.040
REGLEMT	0.000	-0.259	0.022
EXRVTE	0.009	0.803	0.033
SAPVTE	0.008	-0.004	0.041
PROCOMM	0.000	-0.267	0.345
CA1979	0.478	0.000	0.229
Tau (Kendall)	0.3900	0.3829	0.4269
Corrélation linéaire		0.8720	0.8083
Nombre d'itérations		6	11

## 2.2.COMMENTAIRES

Faisons tout d'abord remarquer que le tau de Kendall, donné dans le tableau de synthèse, est celui du préordre initial  $g_0$  avec le préordre reconstitué  $\Sigma\{\beta_j U_j (g_j)\}$  et la corrélation linéaire est celle du couple canonique  $(U_0 (g_0), \Sigma\{\beta_j U_j (g_j)\})$ .

### I.EVALUATION DE PROJETS

REGALS restitue l'ordre des cinq premiers projets et apporte deux indifférences : (A4,A2) donnée par UTA, (A7,A9) donnée par MORALS.

UTA accorde, aux deux critères COMPCOM et EFFICOM, 95.5% de l'explication totale du préordre  $g_0$ . L'ordre des poids affectés aux quatre critères est identique pour MORALS et REGALS. L'effet conjoint des niveaux élevés du Tau de Kendall et de la corrélation linéaire atteints par REGALS explique la cohérence de l'ordre des critères EFFICOM, EFFINAN, FLEXCOM avec celui de leurs corrélations respectives avec le préordre initial (calculées à partir des données initiales):

$$\text{COR}(2,0)=0.644 ; \text{COR}(3,0)=0.674 ; \text{COR}(4,0)=-0.133.$$

### II.OCTRDI DE CREDITS

REGALS restitue globalement la structure ordinale du préordre initial en éliminant toutes les indifférences (la virgule sépare les actions indifférentes):

INIT : A5/A2,A1/A7,A6/A3,A4/A9,A8

REGALS : A5/A2/A1/A7/A6/A3/A4/A9/A8

UTA : A5/A1/A7/A2,A6/A3/A4/A9/A8

MORALS : A5,A1,A7/A8,A9/A2,A6/A3,A4

Les trois méthodes éliminent le critère RATIO2, en fait identique à RATIO1. On constate une divergence importante entre UTA, MORALS d'une part et REGALS d'autre part: le critère RATIO1 est considéré par les premières comme le plus important alors que RATIO5 a le poids le plus élevé pour REGALS.

Mais, pour cette dernière, notons que l'ordre d'importance des critères coïncide avec celui de leurs corrélations respectives avec le préordre initial (calculées à partir des données initiales):

$$\text{COR}(1,0)=0.365; \text{COR}(3,0)=0.454; \text{COR}(4,0)=-0.151; \text{COR}(5,0)=0.702$$

Cette cohérence s'explique par les niveaux excellents du Tau de Kendall ( $\tau=0.89$ ) et de la corrélation linéaire (0.9999).

REGALS ne remet donc pas en cause le classement initial. De plus, les limites A1,A2 d'une part et A3,A4 d'autre part sont départagées dans le classement et permettent de conclure que la population de demandeurs (voir section 2.1) D<sub>24</sub> est considérée comme étant plus «à risque» que la population D<sub>13</sub>: les dossiers y nécessitent plus souvent une analyse approfondie.

### III. SELECTION DE POINTS DE VENTE

Les trois méthodes donnent un tau de Kendall faible, donc une reconstitution médiocre du préordre initial. On observe une divergence nette au niveau de la pondération des critères:

UTA accorde aux deux critères DYNAM1 et CA1979, 96% de l'explication totale du préordre  $\rho_0$ .

MORALS donne la corrélation linéaire la plus élevée et affecte aux critères les mieux corrélés avec le préordre initial, LOCAL et EXRVTE, les poids les plus élevés mais dans un ordre inverse:

$$\text{COR}(1,0) = 0.367 ; \text{COR}(9,0) = 0.510.$$

CA1979, traité comme critère numérique, est éliminé: le poids qui lui est donné par les autres méthodes—qui le traitent comme qualitatif—est la traduction d'une forte liaison non linéaire.

REGALS dégage trois critères significatifs: DYNAM1, CA1979 (comme UTA) et PROCOMM (éliminé par UTA et affecté par MORALS

du poids le plus faible).

#### IV.SYNTHESE

Notons une tendance de la méthode UTA à la surpondération des critères les plus importants au détriment d'autres critères dont l'influence peut être significative.

Une conclusion, fondée sur le fait que le tau de Kendall obtenu par REGALS est systématiquement le plus élevé, consacrant cette méthode (et l'approche moindres carrés !) serait non seulement simpliste mais trompeuse : dans l'application de la méthode REGALS , nous avons recherché un compromis entre un tau de Kendall suffisamment élevé et une corrélation linéaire satisfaisante par le contrôle pas à pas du processus itératif:

#### Exemple I:Evaluation de projets

Itération	5	10	11	12
Tau	0.84554	0.83088	0.80147	0.80147
Corrélation	0.99650	0.99745	0.99750	0.99753

A partir de la 11<sup>e</sup> itération, le tau de Kendall ne varie plus. La corrélation linéaire évolue très faiblement et justifie , à notre sens, l'arrêt du processus à la 5<sup>e</sup> itération. Il faut noter qu'à la stationnarité, l'ordre des poids affectés aux critères devient identique à celui de UTA qui donne, en définitive, un meilleur tau de Kendall.

#### Exemple II:Octroi de crédits

La corrélation linéaire très élevée (0.9999) justifie l'arrêt du processus. Les résultats sont effectivement meilleurs.

#### Exemple III:Sélection de points de vente

La corrélation linéaire (0.8083) ne varie pratiquement plus (11 itérations) et justifie l'arrêt du processus à un tau de Kendall meilleur que celui des autres méthodes.

Pour éclairer les remarques précédentes, et faire passer l'analyse du niveau du constat à celui de l'explication, nous devons l'affiner par les considérations suivantes.

### 2.3. METRIQUES D'AGREGATION D'ECHELLES

Marcotorchino et Michaud définissent (1979) deux métriques simples d'agrégation d'échelles de rangs ( ou de valeurs ou de notes).

Le problème : Etant donné un ensemble de critères ordinaux

$$g_j, j = 1, \dots, p,$$

observés sur  $n$  individus, il s'agit de déterminer un préordre  $X^*$  agrégeant les  $p$  préordres initiaux en un seul.

On définit pour cela une métrique  $\mu(X, g_j)$  et on recherche  $X^*$  tel que

$$\text{MIN}\{\sum\{\mu(X, g_j), j=1, \dots, p\}, X \geq 0\} = \sum\{\mu(X^*, g_j), j=1, \dots, p\}.$$

$$\text{Métrique } d_1 : d_1(X^1, g_j) = \sum\{|X^1_{i1} - g_{j1}|, i=1, \dots, n\}.$$

$$\text{Métrique } d_2 : d_2(X^2, g_j) = \sum\{(X^2_{i1} - g_{j1})^2, i=1, \dots, n\}.$$

Marcotorchino et Michaud montrent que  $X^1$  est tel que:

$X^1_{i1}$  est la médiane de la distribution  $\{g_{j1}, j=1, \dots, p\}, i=1, \dots, n,$

et établissent que  $X^2$  est tel que:

$X^2_{i1}$  est la moyenne de la distribution  $\{g_{j1}, j=1, \dots, p\}, i=1, \dots, n.$

Dans chacune des deux métriques, il est possible d'affecter aux critères une pondération.

Une relation empirique classique de statistique descriptive, reliant mode médiane et moyenne, est:

$$\text{mode} \approx 3 \times \text{médiane} - 2 \times \text{moyenne}.$$

Elle permet de déduire que le classement  $X^0$  tel que

$X^0_i$  est le mode de la distribution  $\{g_{ji}, j=1, \dots, p\}$ ,  $i=1, \dots, n$ ,

serait relié à  $X^1$  et  $X^2$  par la même relation.

*Les tableaux ci-après donnent, pour chaque exemple étudié, le tau de Kendall et la corrélation linéaire entre le préordre issu de chacune des deux métriques d'agrégation et le préordre initial.*

Evaluation de projets	Métrique $d_1$	Métrique $d_2$
Tau de Kendall	0.6103	0.5882
Corrélation linéaire	0.8104	0.7746

Octroi de crédits	Métrique $d_1$	Métrique $d_2$
Tau de Kendall	0.4444	0.6667
Corrélation linéaire	0.5635	0.8167

Points de vente	Métrique $d_1$	Métrique $d_2$
Tau de Kendall	0.2571	0.3535
Corrélation linéaire	0.3827	0.5086



Les tableaux ci-après donnent , pour chaque exemple étudié, le tau de Kendall et la corrélation linéaire entre le préordre issu de chacune des deux métriques d'agrégation et sa reconstitution à l'aide d'une itération de REGALS.

Evaluation de projets	Métrique $d_1$	Métrique $d_2$
Tau de Kendall	0.8382	0.8897
Corrélation linéaire	0.9746	0.9992

Octroi de crédits	Métrique $d_1$	Métrique $d_2$
Tau de Kendall	0.5833	0.8889
Corrélation linéaire	0.9527	1.0000

Points de vente	Métrique $d_1$	Métrique $d_2$
Tau de Kendall	0.7045	0.9396
Corrélation linéaire	0.9768	0.9992

#### 2.4. CONCLUSION DE L'ETUDE COMPARATIVE

Au vu des deux groupes de tableaux présentés ci-dessus, deux remarques s'imposent:

-Le préordre issu de la métrique d'agrégation  $d_2$  est, systématiquement mieux expliqué par REGALS que celui issu de  $d_1$ .

-Le préordre à expliquer de l'exemple I est plus proche du préordre issu de  $d_1$ , contrairement aux préordres des exemples II et III.

La première remarque nous amène à conjecturer que, réciproquement, un préordre issu de la métrique d'agrégation  $d_1$  serait mieux expliqué par UTA.

La deuxième remarque et notre conclusion de la section 2.2 (dans la synthèse), concernant l'exemple I, motivent notre acceptation de cette conjecture.

UTA et REGALS sont issues de deux modèles différents mais elles ont, comme on l'a vu, un fondement commun: l'analyse canonique. On peut les juger représentatives des méthodes découlant de chacune des deux approches fondamentales. Le parallèle effectué entre ces deux méthodes peut donc être considéré comme une base fiable de comparaison de ces approches dans le problème de la désagrégation de la préférence globale.

Notre conclusion est fondée sur le fait que cette comparaison ne peut être séparée de la question du mode d'agrégation duquel a résulté le préordre à expliquer.

L'agrégation, considérée comme une phase particulière d'un processus de décision, est généralement effectuée à l'aide d'une méthode d'optimisation numérique (utilisation d'une métrique par exemple) et procède souvent d'un point de vue subjectif (de l'homme d'étude) pouvant être modal, médian, moyen, etc... Ces différents points de vue sont fondés sur l'approche tactique sous-jacente au problème.

Ainsi, une attitude «modale» adoptée par l'homme d'étude, privilégie «l'efficacité» sur un critère. Cette attitude, plus proche d'un point de vue «médian», serait mieux expliquée dans une approche moindres écarts.

Par contre, une attitude «moyenne» donne plus d'importance à l'«interchangeabilité» des actions sur les critères. L'approche moindres carrés en appréhenderait mieux le mécanisme.

## REFERENCES BIBLIOGRAPHIQUES

A.Agha (1988) : Analyse Canonique de deux cônes polyédriques convexes, Université de Paris-Dauphine, Cahier du Lamsade n°87.

R.D.Armstrong & E.L.Frome (1976) : Branch and bound Solution of Restricted Least Squares Problem, Technometrics vol.18, 447-450.

R.E.Barlow, D.J.Bartholomew, J.M.Bremner & H.D.Drunk (1972) : Statistical inference under order restrictions, New York, Wiley.

I.Barrodale & F.D.K.Roberts (1972) : An Improved Algorithm for Discrete  $L_1$  Linear Approximation, Technical Report TSR 1172, Mathematics Research Center, University of Wisconsin, Madison.

I.Barrodale & F.D.K.Roberts (1973) : An Improved Algorithm for Discrete  $L_1$  Linear Approximation, SIAM J.Num.Anal.10, 839-848.

A.E.Beaton (1964) : The use of special matrix operators in statistical calculus, Ed D.thesis, Harvard University, Reprinted as Educational Testing Service Res.Bull.64-51, Princeton, N.J.

J.P.Benzécri et Collaborateurs (1973) : L'Analyse des Données , tome II: L'analyse des correspondances, Dunod, Paris.

P.Bertier & M.Tenenhaus (1974) : Les méthodes de mesures conjointes, Séminaire AFCET. Paru dans L'Analyse des Données en Marketing, J.M.Bouroche (1977).

J.M.Bouroche, M.Marc et M.Tenenhaus (1974) : Application of Conjoint measurement to marketing problems and media research, Journal of Marketing Research, vol.XI, pp 109-129.

G.E.P.Box & D.R.Cox (1964) : An analysis of transformations, Journal of Royal Statistical Society, series B, 26, pp 211-252.

**J.M.Bremner (1982) : An Algorithm for Nonnegative Least Squares and Projection onto Cones, Compstat, Physica-Verlag, Vienna for IASC, pp 155-160.**

**P.Cazes (1975) : Protection de la régression par utilisation de contraintes linéaires et non linéaires, Revue de Statistique Appliquée, vol. XXIII, n°3, pp 37-56.**

**P.Cazes, D.Bergougnan et C.Mullon (1984) : Présentation de deux programmes de régression ,Data Analysis and Informatics, III, E. Diday & al (editors), Elsevier Science Publishers B.V. (North-Holland).**

**P.Cazes & P.Y.Turpin (1971): Régression sous contraintes, application à l'estimation de la courbe granulométrique d'un aérosol, Revue de Statistique Appliquée, vol. XIX, n°4, pp 23-43.**

**A.Charnes W.W.Cooper and R.O.Ferguson (1955) : Optimal estimation of executive compensation by linear programming, Management Science 1, pp 138-150.**

**J.H.Chauchat (1985) : Codage optimal ou codage a priori, Journées Internationales d'Analyse des Données, octobre.**

**P.Dagnelie (1975) : Analyse statistique à plusieurs variables, Ed. Les Presses Agronomiques de Gembloux.**

**Y.Evrard & P.Lemaire (1976) : Information et Décision en Marketing, Dalloz, Gestion.**

**R.Faure et Collaborateurs (1979) : Précis de Recherche Opérationnelle, Dunod, Paris.**

**B.de Finetti (1937) : La prévision: ses lois logiques, ses sources subjectives ,Annales de l'institut Henri-Poincaré, VII,1, in R.Charreton, J.M.Bourdaire: La décision économique, «Que sais-je ?», n°2235, Ed.PUF.**

R.A.Fisher (1948) : Statistical methods for research workers, 10<sup>e</sup> Ed., Oliver & Boyd, London.

H.Hotelling (1936) : Relations between two sets of variables, Biometrika, vol.28.

E.Jacquet-Lagrèze & J.Siskos (1982) : Assessing a set of additive utility fonctions of multicriteria decisions making, the UTA method, European Journal of Operational Research, 10, 151-164.

R.I.Jennrich (1977) : "Stepwise regression", in K.Enslein , A. Ralston & H.S.Wolf (Eds) , "Statistical methods for digital computers" , pp 58-75, J.Wiley & Sons.

G.G.Judge & T.Takayama (1966) : Inequality Restrictions in Regression Analysis , American Statistical Association Journal, 61, pp 166-181.

R.L.Keeney (1974) : Multiplicative Utility Fonctions, Operations Research, 22 pp 22-34.

R.L.Keeney & K.Nair (1976) : Evaluating Potential Nuclear Power Plan Sites in the Pacific Northwest using Decision analysis, IIASA Professional Paper n°76-1.

R.L.Keeney & H.Raiffa (1976) : Decision with Multiple Objectives-Preference and Value Tradeoffs, Wiley, New York.

M.G.Kendall (1962) : Rank correlation method , Ch.Griffin & co, Londres.

W.J.Kennedy & J.E.Gentle (1980) : Statistical Computing, Marcel Dekker inc., New York.

J.B.Kruskal (1964a) : Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis, Psychometrika, vol.29, pp 1-27.

J.B.Kruskal (1964b) : Nonmetric multidimensional scaling: a numerical method, *Psychometrika*, vol.29, 28-42.

J.B.Kruskal (1965) : Analysis of Factorial Experiments by Estimating Transformations of the Data, *Journal of the Royal Statistical Society* ,27, series B, 251-263.

J.B.Kruskal & F.J.Carmone (1968) : Monanova , a Fortran IV program for monotone analysis of variance , Unpublished Manuscript. Bell Telephone Laboratories, Murray Hill, New Jersey.

W.W.Kuhn & A.W.Tucker (1951) : Nonlinear Programming, Proc. 2<sup>nd</sup> Berkeley Symp. on Mathematical Statistics and Probability, University of California Press, Berkeley.

C.M.Lawson & R.J.Hanson (1974) : Solving Least Squares Problems Englewood Cliffs, Prentice Hall, New Jersey.

J.de Leeuw (1975) : Normalized cone regression, Department of data theory, University of Leiden.

J.de Leeuw, F.W.Young & Takane (1976) : Additive structure in qualitative data : an alternating least squares method with optimal scaling features. *Psychometrika*, vol.41, n°4, pp 471-503.

N.Mantel (1969) : Restricted Least Squares Regression and Convex Quadratic Programming, *Technometrics* vol.11, 763-773.

J.F.Marcotorchino & P.Michaud (1979) : Optimisation en analyse ordinaire des données, Collection Statistique et Décisions économiques, Masson, Paris.

J.von Neumann & O.Morgenstern (1947) : Theory of Games and Economic Behavior, 2<sup>nd</sup> ed, Princeton University Press, N.J.

J.L.Richard : Aide à la décision stratégique en PME, in E . Jacquet-Lagrèze et J.Siskos (1983), chapitre V, pp 119-142.

B.Roy (1978) : ELECTRE III : Un algorithme de classements fondé sur une représentation floue des préférences en présence de critères multiples, Cahiers du CERO, vol.20, n°1, pp 3-24.

B.Roy (1985) : Méthodologie Multicritère d'Aide à la Décision, Ed.Economica.

B.Roy & D.Bouyssou (1983) : Comparaison sur un cas précis de deux modèles concurrents d'Aide à la Décision, Université de Paris-Dauphine, Document du Lamsade n°22.

G.Saporta (1975) : Dépendance et codage de deux variables aléatoires, Revue de Statistique Appliquée, vol.XXIII, n°1, pp 43-63.

G.Saporta (1977) : Une méthode et un programme d'analyse discriminante pas à pas sur variables qualitatives, Colloques IRIA: Analyse des Données et Informatique, vol.1, pp 201-210.

G.Saporta (1978) : Théories et méthodes de la statistique, Publications de l'I.F.F., Ed.Technip.

SAS/STAT User's Guide (1989) : vol.2, version 6, CARY, NC: SAS Institute Inc.

R.Savoundararadja (1987) : Evaluation comparative des méthodes multicritères de régression ordinale : méthode UTA et méthode MORALS. Thèse de 3<sup>e</sup> cycle, Université Pierre et Marie Curie.

R.R.Singleton (1940) : A method of minimising the Sum of Absolute Values of Deviations, AMS 11, 301-310.

J.Siskos (1979) : La modélisation des préférences au moyen de fonctions d'utilités additives. Thèse de 3<sup>e</sup> cycle, Université Pierre et Marie Curie.

J.Siskos (1984) : Analyses de Régression et Programmation linéaire, Université de Paris-Dauphine, Cahier du Lamsade n°54.

J.Siskos & D.Yannacopoulos (1983) : Amélioration de la méthode UTA par introduction d'une double fonction d'erreurs, Université de Paris-Dauphine, Cahier du Lamsade n°49.

M.Tenenhaus (1979) : La régression qualitative , Revue de Statistique Appliquée, vol.XXVII, n°2.

M.Tenenhaus (1983) : L'analyse des données qualitatives par des méthodes de codage optimal, Cahiers de Recherche du CESA, n°229.

M.Tenenhaus (1988) : Canonical analysis of two convex polyhedral cones and applications, Psychometrika, vol.53, n°4, 503-524.

H.M.Wagner (1959) : Linear Programming Techniques for Regression Analysis, JASA 54, 206-212.

R.E.Walpole & R.H.Myers (1972) : Probability and Statistics for Engineers and Scientists, New York: The Macmillian Company.

M.S.Waterman (1974) : A Restricted Least Squares Problem, Technometrics vol.16, n°1, pp 135-136.

M.S.Waterman (1977) : Least Squares with Nonnegative Regression Coefficient, JSCS 6, 67-70.