Méthodes et programmes en analyse de données ordinales

Eric Jacquet-Lagrèze

metra

direction scientifique

16-20, rue barbès, 92128 montrouge - téléphone : 657-13-00

méthodes et programmes en analyse de données ordinales

éric jacquet-lagrèze

DIRECTION SCIENTIFIQUE
Synthèse et Formation n° 81
Décembre 1973

METHODES ET PROGRAMMES EN ANALYSE DE DONNEES ORDINALES

Eric JACQUET-LAGREZE

SEMA (Metra International)
16/20, rue Barbès, 92128 MONTROUGE

SOMMAIRE

Pages

Pages

AVANT-PROPOS

Depuis ELECTRE I et II, méthodes d'agrégation des préférences (multi-critères), depuis PREFMAP, méthode d'analyse des préférences conçue par J.C. CARROLL et introduite à la SEMA, un certain nombre de méthodes permettant une description ou une structuration des données ordinales ont été rassemblées ou développées dans le cadre de l'Action Programmée PSD III.

Tandis que J.M. BOUROCHE et P. LAYE ont rendu disponible la version PREFMAP II (1), E. JACQUET-LAGREZE présente dans cette note de synthèse l'ensemble des méthodes de description et de structuration de données ordinales actuellement disponibles à la SEMA.

Cette note de synthèse voudrait être :

- un guide pour le praticien cherchant une ou plusieurs méthodes pour aborder le problème qu'il doit résoudre ;
- un guide pour le chercheur ou pour le méthodologue qui voudrait trouver une présentation générale de différents axes de recherche avant d'aller plus loin dans l'approfondissement de certaines voies. A cette fin, on s'est efforcé de présenter une bibliographie importante sur le sujet.

Une introduction a pour objet de présenter au lecteurs les différents problèmes que les méthodes cherchent à résoudre.

Les deux premiers chapitresdonnent une présentation de méthodes :

- méthodes de description (première partie),
- méthodes de structuration (deuxième partie),

dont on a cherché à alléger l'exposé en renvoyant certains développements mathématiques en annexe.et en cherchant à mettre en évidence leurs princi-

---

(1) Une note de travail vient de paraître à ce sujet (cf. Annexe 3 : Bibliographie [19]).

pales différences. Une compréhension plus poussée de ces méthodes nécessiterait un texte beaucoup plus volumineux.

Un troisième chapitrecomprend une présentation générale des programmes de description et de structuration de données ordinales disponibles à la SEMA sur CDC 6600 et un manuel d'utilisation pour la plupart d'entre eux.

Dans ce domaine, le recueil des données pose des problèmes particulier qui sont abordés en annexe.

P. BERTIER

INTRODUCTION

Celle-ci a pour objet de définir ce qu'on entend par données ordinales, de présenter un certain nombre d'exemples où l'on recueille de telles données ordinales, enfin de présenter des problèmes que l'on peut chercher à résoudre à l'aide de ces données.

1. Données de proximité et données ordinales

Lorsqu'on se trouve en présence d'un ensemble d'objets [1], on est amené dans bien des problèmes à comparer ces objets entre eux. Une comparaison peut être de deux natures différentes et peut chercher à saisir l'une des notions suivantes :

- Une ressemblance ou une proximité :

"a et b se ressemblent" ou "a et b sont proches" par opposition à "a et b" sont différents.

Une modélisation d'une telle comparaison est alors une relation binaire (cf. Annexe I) entre ces objets dont la principale propriété est d'être symétrique, à savoir si "a et b sont proches", "b et a sont proches". Cette relation binaire peut être non valuée (relation vraie ou fausse) ou valuée (on fait correspondre à la relation un nombre indiquant le degré ou l'intensité de la relation). Une telle relation est ce qu'on appelle une donnée de proximité.

- Un classement ou une hiérarchisation :

"a est mieux que b", "je préfère a à b", "a a une influence sur b", "selon ce point de vue, a est préféré à b", "a domine b au cours d'un match".

---

(1) Le mot objet est pris dans un sens général. Il peut aussi bien s'agir d'objets concrets (produits nouveaux, voitures automobiles) que d'objets plus abstraits (dossiers de projets pour un concours d'architecture) ou encore de personnes, de proposition abstraites, etc.

Une modélisation de telles comparaisons est également une relation binaire entre ces objets dont la principale propriété est cette fois l'existence d'une non symétrie : si "a domine b", "b peut ne pas dominer a". De telles relations, en général non symétriques, sont ce qu'on appellera dans la suite des données ordinales.

2. Exemples de données ordinales

Comme exemples de données ordinales, on peut citer :

2.1 Les données de préférences globales

Celles-ci peuvent porter sur un ensemble

- d'oeuvres d'art (peintures, sculptures, ...) ;
- d'émissions de télévision, de films, ... ;
- de lieux de séjour de vacances ;
- de logements à louer, à acheter ;
- de voitures automobiles ;
- de différents produits concurrents destinés au même usage ;
- de différents constructeurs de matériel industriel ;
- de projets d'architecture présentés à un concours ;
- d'hommes politiques ;
- etc.

Cette liste pourrait comporter plusieurs pages. Notre univers de perception d'un ensemble d'objets peut très vite évoluer volontairement ou non vers une construction ou une définition de nos préférences sur ces objets. Ces préférences sur ces objets se traduisent par un classement complet ou non de l'ensemble des objets, nous dirons par la définition d'une relation de préférence. Nous verrons (Introduction, § 3) sous quelles formes on peut chercher à appréhender ces préférences. On trouvera en annexe II (§ 2) un exemple de recueil de données de préférence globale dans un problème de gestion du personnel.

2.2 Les données de jugement ou d'opinions sur plusieurs dimensions

On recueille ce type de données lorsqu'on demande à des experts, des jurés ou encore à des personnes interrogées au cours d'une enquête un classement d'objets relativement à une ou plusieurs dimensions.

- On pourra par exemple chercher à classer des professions [1] :

Architecte	Ingénieur	Officier
Avocat	Magistrat	Professeur d'Université
Banquier	Médecin	Expert comptable
Diplomate	Notaire	

. suivant la dimension "prestige" (les classer depuis la plus prestigieuse jusqu'à la moins prestigieuse) ;

. suivant la dimension "difficulté des études pour parvenir à ces professions".

- On trouvera en annexe II (§ 3) un exemple de recueil de données de classements de métiers suivant la dimension prestige.

- Dans une enquête auprès d'automobilistes, on pourra chercher à classer des voitures suivant des dimensions telles que :

sécurité	confort
puissance	etc.
vitesse	

- Pour un dépouillement de concours de projets d'architecture, on peut chercher à comparer les différents projets suivant différents <u>critères</u> [2]

---

(1) SEMA (Division E) - Etude sur le profil idéal des dirigeants d'entreprise. Juin 1973.

(2) Le mot critère s'emploie dans le cadre de l'aide à la décision pour définir une dimension telle que l'ordre des objets sur cette dimension soit compatible avec la préférence du décideur suivant cette dimension.

en s'appuyant sur des jugements d'experts.

Dans le cas du Programme d'Architecture Nouvelle (Ministère de l'Equipement et du Logement [1]), le jury évalue chaque projet suivant les critères :

- innovation ;
- qualité de la construction ;
- aptitude du projet à passer en logement social ;
- densité de l'habitat (se rapproche-t-elle des normes souhaitées ?)
- importance du marché potentiel ;
- réalisme (économique, technique, sociologique) du projet ;
- adéquation de l'équipe présentant le projet au principe du concours

- Dans un problème d'aide à la décision en présence de critères multiples, après avoir défini l'ensemble des actions réalisables, des critères et de leurs indicateurs (cf. B. ROY [81]), on évalue chaque action sur cha critère. Certains critères permettent une évaluation quantitative (coût, surface, temps), d'autres nécessitent une évaluation qualitative (aspect "stratégie de l'entreprise" dans un problème de choix de projets de recherche, indice d'une bonne solution politique dans un problème de décision d'une municipalité, etc.). On peut, si l'on veut, dans un premier temps, ne tenir compte que de la propriété ordinale de ces évaluations (c'est-à-dire retenir l'ordre et non plus les valeurs quantitatives), considérer le évaluations des actions sur chaque critère comme autant de préordres complets de ces actions que l'on a de critères.

---

(1) On pourra consulter à ce sujet la revue Tourisme Equipement Logement TEL (n° 202 en particulier) ou encore P. BERTIER et E. JACQUET-LAGREZE [14

2.3 Données de relation de pouvoir et de dominance

- On peut chercher à traduire la notion de pouvoir ou de dominance d'un individu (ou groupe d'individus) sur un autre individu (ou groupe) à l'aide d'une relation binaire valuée (c'est-à-dire ayant une intensité) ou non valuée.

- Les jeux et plus précisément ce qu'on appelle tournois constituent un exemple bien connu où l'on recueille des données de dominance. Un tournoi entre n joueurs est un jeu où chaque joueur rencontre dans une partie les $n - 1$ autres joueurs. Au cours d'un tournoi, on recueille des données ordinales non valuées (x bat y) ou valuées (x bat y par n_x contre n_y). Une telle relation binaire complète, non valuée, ainsi obtenue est un "tournoi" ou un "match" (cf. Annexe I, § 3).

- Dans l'étude d'organisations sociales, on peut être amené à recueillir des données de relation de pouvoir ou de dominance (cf. [4] et [76]). Les relations cherchées peuvent porter sur le pouvoir des individus les uns sur les autres, leur influence mutuelle, leur degré d'affinité, etc. Ces relations valuées ou non sont presque toujours non symétriques et peuvent être analysées comme des données ordinales.

3. Les divers modes de représentation de données ordinales

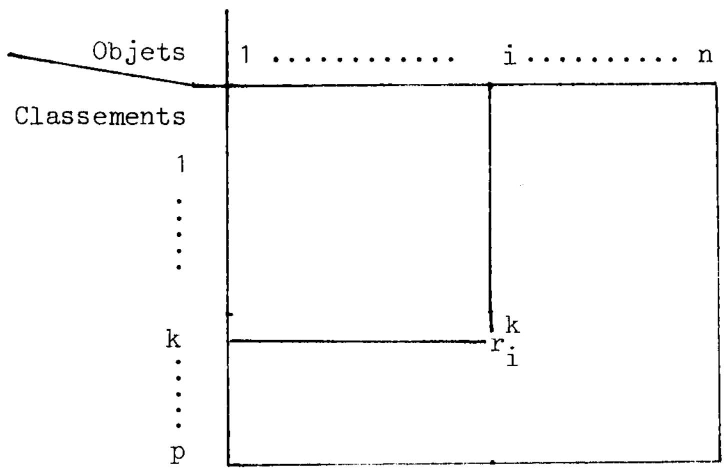
Après ces quelques exemples de données ordinales, précisons les formes sous lesquelles on peut recueillir ces données.

Nous distinguerons les rangs, les évaluations sur des échelles, les comparaisons par paires.

3.1 Classements d'objets à l'aide de rangs

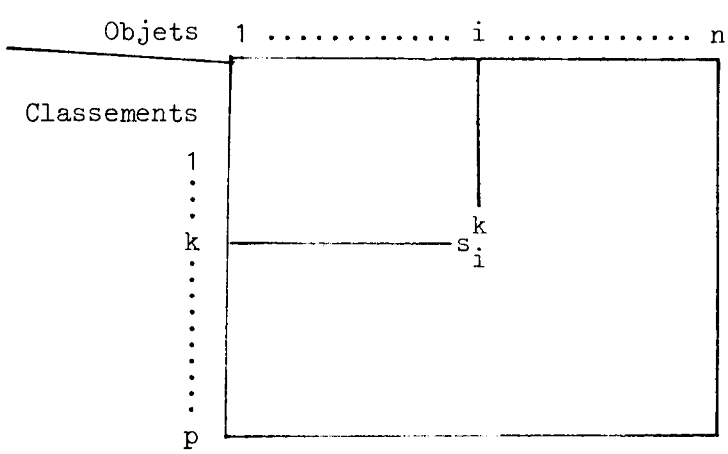
Les objets sont classés en un préordre complet (Annexe 1, § 21). Certains objets de la première classe d'équivalence reçoivent le rang 1, ceux de la seconde classe le rang 2, etc.

Dans l'exemple de recueil d'opinions sur des professions suivant les dimensions "prestige" et "difficulté des études" (§ 2.2 ci-dessus), on a adopté ce mode de recueil par rangs. Si p personnes ont été interrogées on a p classements sur les n objets. On possède alors un tableau de n x p rangs à analyser où l'élément r_i^k est le rang de l'objet i suivant le $k^{ième}$ classement.



3.2 Evaluation sur des échelles

Les objets sont évalués sur une échelle E. L'échelle peut être quantitative : une unité de mesure existe. C'est le cas d'une échelle de coût de surface, de temps, ... L'échelle peut être qualitative (note de 1 à 20 échelle à 4 positions A, B, C, D, etc.). Si on possède p classements o évaluations sur n objets, on obtient un tableau de n x p évaluations o l'élément s_i^k est l'évaluation de l'objet i suivant le classement k, évaluation faite à l'aide de l'échelle E_k :



- Dans le problème de dépouillement du concours d'architecture évoqué plus haut, chaque juré donne 7 x n évaluations (7 critères, n projets) en utilisant une échelle qualitative commune aux 7 critères, munie des six échelons (A, B, C, D, E, F).

- Dans un problème d'aide à la décision en présence de critères multiples, le plus souvent, certains critères sont quantitatifs, d'autres qualitatifs : on trouvera un exemple de problème de choix de magazines comme supports publicitaires dans le troisième chapitre, § II.5. Un critère, le coût, est évalué en Francs au 1 000 lecteurs, les autres le sont à l'aide d'une échelle qualitative à 10 positions (note de 1 à 10).

- Un cas particulier est celui des échelles dichotomiques (0,1). Chaque classement est alors un préordre en deux classes d'équivalences. Si les deux classes ne sont pas hiérarchisables, mais sont du type (0 : ne possède pas telle propriété, 1 : possède cette propriété), on ne peut pas parler de données ordinales. Cependant, dans certains problèmes, les deux classes peuvent être hiérarchisées en utilisant par exemple le fait de posséder la propriété comme un moyen d'exprimer sa préférence sur un ensemble d'objets.

<u>Exemples</u> :

- réussite à un examen, à un test ;
- choix de films - 0 : film en noir et blanc, 1 : film en couleur : un film en couleur peut, dans certains cas, être considéré comme étant préférable à un film en noir et blanc.

On verra une application de cette approche avec les techniques d'analyse hiérarchique (deuxième chapitre, § 4).

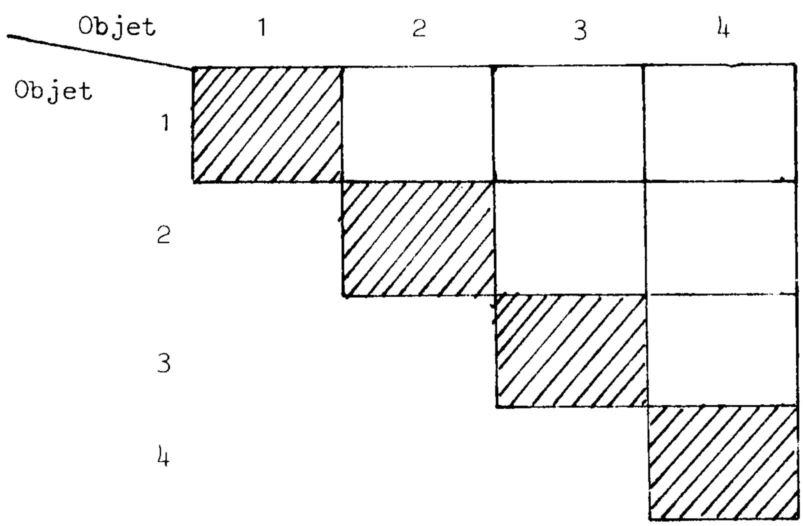
3.3 <u>Comparaisons par paires</u> (Annexe II, § 2)

Les comparaisons entre objets se fait directement en portant un jugement sur chaque paire d'objets.

Si on a n objets, il existe $C_n^2 = n(n - 1)/2$ paires d'objets.

- Exemple : Il existe 6 paires d'objets :

(1, 2), (1, 3), (1, 4)
(2, 3), (2, 4), (3, 4)



- Dans un problème de recueil de données de préférences, on retiendra par exemple pour chaque paire d'objets (i, j) l'une des quatre possibilités suivantes pour chaque classement k [1] :

$$\left.\begin{array}{l} a_{ij}^k = 1 \\ a_{ij}^k = 0 \end{array}\right\} \quad i \text{ classé avant } j$$

$$\left.\begin{array}{l} a_{ij}^k = 0 \\ a_{ij}^k = 1 \end{array}\right\} \quad j \text{ classé avant } i$$

$$a_{ij}^k = a_{ji}^k = 1/2 \qquad i \text{ et } j \text{ dans une même classe d'indifférence}$$

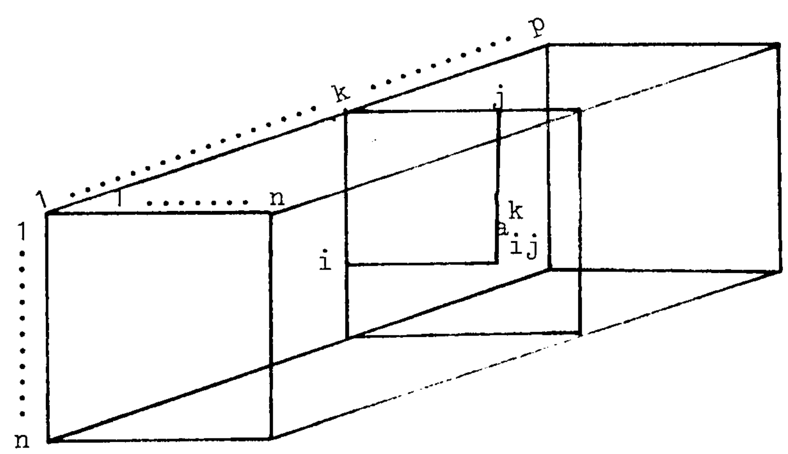
$$a_{ij}^k = a_{ji}^k = 0 \qquad i \text{ et } j \text{ non comparables ou non comparés.}$$

---

(1) Sauf précision, lorsqu'on parlera dans la suite de comparaisons par paires, il s'agira de la représentation présentée ici.

En annexe II (§ 2), on trouvera un exemple de questionnaire permettant de collecter de telles données.

- Les comparaisons peuvent se faire selon plusieurs dimensions (plusieurs sujets répondent à un questionnaire en comparaison suivant plusieurs critères). Si on a p relations de comparaison par paires portant sur n objets, on notera par a_{ij}^{k} la relation entre les objets i et j pour la $k^{ième}$ dimension. On se trouve alors en présente d'un cube de données :



- Les données que l'on recueille à l'issue d'un tournoi sont des comparaisons par paires ; elles peuvent être valuées ; par exemple, si l'équipe i bat l'équipe j par 3 contre 2, on aura :

$a_{ij} = 3$, $a_{ji} = 2$.

- Dans le cas où le nombre d'objets est trop grand, le nombre de paires d'objets devient trop important (pour $n = 25$, $n(n-1)/2 = 300$) ; on a alors parfoir recours à des <u>comparaisons par blocs</u>.

. Exemple : $n = 9$, avec 12 blocs de 3 objets, on peut reconstituer les $9.8/2 = 36$ paires d'objets.

Si les objets sont identifiés par les chiffres 1, 2, 3, 4, 5, 6, 7, 8, 9, les 12 blocs sont :

	1	4	7	1	2	3	1	1	2	2	3	3
	2	5	8	3	5	6	5	6	4	6	4	5
	3	6	9	7	8	9	9	8	9	7	8	7
bloc n°	1	2	3	4	5	6	7	8	9	10	11	12

Chaque bloc permet la comparaison de 3 paires (exemple : le bloc n° 1 permet de comparer les paires (1, 2), (1, 3), (2, 3)) ; les 12 blocs permettent la comparaison de 36 paires. On demandera par exemple un ordre complet sur chaque bloc.

Exemple : bloc n° 2 : on aura les rangs :

4	2
5	1
6	3

ce qui signifie que l'on a les comparaisons

$5 \succ 4$
$5 \succ 6$
$4 \succ 6$

. Dans le tableau ci-dessous (objet x objet), on a figuré à l'intersection de la ligne i et de la colonne j le numéro du bloc où figure la paire (i,j)

	1	2	3	4	5	6	7	8	9
1		1	1	4	7	8	4	8	7
2			1	9	5	10	10	5	9
3				11	12	6	12	11	6
4					2	2	4	11	9
5						2	12	5	7
6							10	8	6
7								3	3
8									3
9									

. Les comparaisons par blocs sont une forme de recueil de données intermédiaire entre la comparaison par rangs (1 bloc de n objets) et la comparaison par paires ($n(n - 1)/2$ blocs de 2 objets).

. <u>Remarque</u> : <u>Dans toute la suite, on désignera par classement une relation de comparaison ordinale pouvant avoir l'une des trois formes ci-dessus</u> (§ 3.1 à 3.3).

4. Les problèmes en analyse ordinale

- Les exemples cités dans les paragraphes précédents ont montré que l'on peut se trouver devant des données ordinales dans des contextes variés :

. analyse et dépouillement d'enquêtes par sondage (préférences sur des professions, études de marché) ;

. animation d'un séminaire en gestion du personnel (recherche de motivations) ;

. aide à la décision en présence de critères multiples (choix de projets de recherche ;

. aide à la décision d'un jury (dépouillement du Concours d'Architecture Nouvelle) ;

. analyse d'un tissu de relations dans une organisation sociale ;

. analyse d'un tournoi (tournoi des cinq nations, ...) ;

. etc.

- Il est certain que, dans tous ces contextes, les problèmes posés ne sont pas les mêmes.

- Dans les problèmes d'aide à la décision, une optique normative est davantage recherchée. Il s'agit de déceler le ou les meilleurs objets (actions) d'un ensemble. On cherchera un préordre complet sur les objets, ou encore pour rester plus fidèle au caractère souvent incertain des données et au caractère flou de la préférence du décideur, on sera conduit à construire une <u>relation de surclassement</u> (cf. [15] et [81]), modèle mathématique d'une relation de préférence du décideur. Dans cette optique, on utilisera des <u>méthodes de structuration</u> des données ordinales. La structure recherchée est alors un préordre complet (ou une relation de surclassement) établie à partir de p classements sur n objets. Ce problème est celui de l'<u>agrégation des préférences</u>.

- Dans l'étude d'une organisation sociale, on cherchera dans quelle mesure le tissu de relations observées est proche de structures particulières ou structures types. G. RIBEILL [75] a développé une telle approche Le problème peut être posé comme celui de l'<u>approximation d'une relation binaire quelconque</u> par une relation binaire donnée (exemple : approximer une relation binaire quelconque par un préordre complet).

Ce problème d'approximation de relations binaires se retrouve également dans le cas de l'agrégation des préférences (approximer une relation de surclassement par un préordre complet, par exemple. Cette démarche est une phase de la méthode ELECTRE II).

On retrouve également ce problème dans l'analyse de tournois.

- Dans un problème d'analyse de données d'enquêtes par sondage, on s'intéressera principalement à une description de l'ensemble des préférences ou des jugements, la recherche dans les différentiations des préférences étant au moins aussi importante sinon plus que la recherche d'ordres sur des objets. On utilisera surtout dans ces problèmes des méthodes de description.

Ces méthodes peuvent être basées sur la recherche d'indices (corrélations entre classements) mais, le plus souvent, il s'agit de méthodes de visualisation utilisant des projections dans R^2 ou R^3 de nuages représentant l'ensemble des classements et l'ensemble des objets. Ces méthodes de visualisation ont pour but de donner une représentation, déformée certes mais synthétique et globale, de l'ensemble des données ordinales.

De plus, dans ces problèmes d'analyse de données d'enquêtes, une structuration portant sur les classements, la recherche par exemple de types de consommateurs ayant des préférences voisines, est également souvent utile.

- Dans le contexte de l'animation d'un séminaire ou de l'établissement d'une procédure de décision par jury, on sera de plus intéressé par l'évolution des préférences ou des jugements. Les méthodes de visualisation et de structuration permettent alors non seulement de mesurer cette évolution mais semblent de plus être un outil efficace permettant un dialogue plus riche et une évolution plus rapide et mieux comprise des préférences.
Dans ces problèmes, très liés à des phénomènes d'échange d'informations, établir un processus d'évolution des préférences est aussi important, sinon plus, que capter les préférences à un instant donné.

CHAPITRE I

LES METHODES DE DESCRIPTION

Ces méthodes ont essentiellement pour objet de mesurer des écarts entre différents classements d'objets.

Lorsqu'il s'agit de données de préférences ou de jugements sur des dimensions, ces écarts s'interprètent en terme de divergence-convergence de jugements.

La quasi-totalité des méthodes de description de données ordinales sont des méthodes métriques, la comparaison de deux classements se faisant en calculant une distance entre ces classements.

Les indices permettant une analyse de 2 ou plusieurs classements sont, on le verra, liés à des distances entre classements (§ I).

Dès que l'on sait établir des distances, on est en mesure de construire un espace géométrique permettant de situer tous les classements comme points ou vecteurs (direction) de cet espace. D'où la naissance des diverses méthodes de visualisation des classements qui projettent dans (R^2) ou dans R^3 le nuage des points "classements" correspondant (§II).

Dans les méthodes de visualisation, on cherche également à représenter l'ensemble des points objets dans le même espace de façon à obtenir la double représentation des points "classements" et des points objets.

C'est dans la façon d'obtenir cette double représentation que se différencient principalement les méthodes.

I - Description de classements par l'utilisation d'indices

Il existe un certain nombre d'indices permettant une description d'un ensemble de classements. Ces indices sont tous reliés à des distances entre classements. Ils diffèrent par conséquent soit par le mode de repré sentation des classements choisi (rangs ou comparaisons par paires), soit par la distance utilisée (distance euclidienne ou distance somme des valeurs absolues).

De plus, il faut distinguer les indices permettant une comparaison de classements deux à deux et les indices permettant une description globale de p classements.

1. Indices de description de classements deux à deux

- On peut utiliser directement les distances entre classements :

. distance de SPEARMAN (cf. KENDALL [54]) (distance euclidienne calculée à partir de différences de rangs) ;

. distance de KENDALL [54] (distance somme des valeurs absolues calculée à partir des comparaisons par paires).

- On peut également utiliser les coefficients de corrélation des ran établis à partir de ces distances (ρ de SPEARMAN, τ de KENDALL). Les diverses distances et coefficients couramment utilisés sont présentés dans le tableau ci-après.

- Lorsque le nombre d'objets est faible, on calcule facilement la distance de KENDALL de la façon suivante :

. on donne les deux classements par la suite ordonnée des objets - exemple : k : a b c d e (a en 1er, b en 2e, ...)

l : c d a e b (c en 1er, d en 2e, ...) ;

Indice	Mode de représentation des classements	Formules	Plage de variation	
			accord	désaccord
Distance de SPEARMAN	rangs $\underline{r}^k$, $\underline{r}^l$	$d_s = \sqrt{\sum_i (r_i^k - r_i^l)^2}$	0	$\sqrt{(n^3 - n)/3} = d_s \max$
Distance de KENDALL	Comparaisons par paires $\underline{a}^k$, $\underline{a}^l$	$d_k = \frac{1}{2} \sum_{i,j} \lvert a_{ij}^k - a_{ij}^l \rvert$	0	$n(n - 1)/2 = d_k \max$
Coefficient de corrélation des rangs de SPEARMAN	rangs $\underline{r}^k$, $\underline{r}^l$	$\rho_{kl} = 1 - 2 \frac{d_s}{d_s \max} = 1 - \frac{6 d_s^2}{n^3 - n}$	+ 1	- 1
Coefficient de corrélation des rangs de KENDALL	rangs $\underline{r}^k$, $\underline{r}^l$	$\tau_{kl} = \frac{\sum_{i,j} x_{ij}^k x_{ij}^l}{n(n - 1)}$ où $x_{ij}^k = + 1$ si $r_i^k - r_i^l > 0$; $x_{ij}^k = - 1$ si $r_i^k - r_i^l < 0$	+ 1	- 1
	Comparaisons par paires $\underline{a}^k$, $\underline{a}^l$	$\tau_{kl} = 1 - 2 \frac{d_k}{d_k \max} = 1 - \frac{4 d_k}{n(n - 1)}$	+ 1	- 1

<u>Remarques</u> : - Les notations de ce tableau sont celles présentées dans l'introduction, § 3 - d_s max et d_k max sont les valeurs maximales qui prennent respectivement les distances de SPEARMAN et de KENDALL.

- On trouvera dans KENDALL [54] des comparaisons entre les coefficients de corrélation (τ/ρ est voisin de 3/2 lorsqu'ils ne sont pas trop proches de + 1 ou - 1).

. on relie les objets entre eux et le nombre d'intersections que l'o obtient est le nombre d'inversions nécessaire pour passer d'un classement un autre ; ce nombre est la distance de KENDALL (cf. BARBUT et FREY | 6 |)

. on en déduit alors facilement le coefficient de corrélation :
$\tau = 1 - \frac{4 \cdot 5}{5 \cdot 4} = 0.$

2. Indices de description de p classements

Ces indices ont pour objet de mesurer des accords ou divergences entr p classements.

- Certains indices sont des coefficients d'ajustement à un ordre, ce ordre étant une représentation moyenne des p classements obtenue par une agrégation de ces derniers (concordance W, KENDALL [54] et cohésion C_o, JACQUET-LAGREZE [50]). Il est par conséquent difficile de présenter ces indices en les séparant du problème de l'agrégation de p classements en un classement unique, problème étudié par la suite (2e chapitre, § II), cela notamment pour le coefficient de cohésion (ajustement à l'ordre média

- D'autres indices sont des coefficients d'accords entre classements l'accord sous-jacent recherché ne devant nullement être nécessairement un ordre comme pour les coefficients ci-dessus.

La plupart de ces coefficients se définissent à partir du tableau des comparaisons par paires (A_{ij}) obtenu en sommant les comparaisons par paires déduites de chaque classement :

$$A_{ij} = \sum_k a^k_{ij}.$$

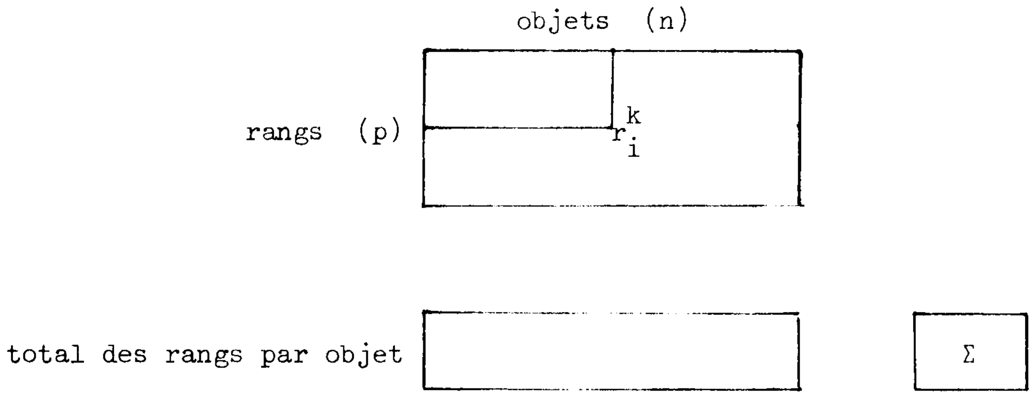
On a présenté les divers coefficients dans le tableau ci-après.

	Indices	Distance sous-jacente	Mode de représentation des classements	Formules	Plage de variation accord	Plage de variation désaccord
Ajustement à un ordre	Concordance W KENDALL [54]	d_s (SPEARMAN) (ρ)	rangs	$W=\frac{12\ S}{p^2(n^3-n)}=\frac{1}{p^2}\sum_{k,l}\rho_{kl}$ $=\rho_{moy}$ (1)	1	0
	Cohésion C_o JACQUET-LAGREZE [50]	d_k (KENDALL) (τ)	- rangs - comparaison par paires	$C_o=\frac{\sum_{(i,j)\in O} A_{ij}-A_{ji}}{\sum_{i,j} A_{ij}}$ $=\frac{1}{p}\sum_k \tau_{k,O}$ (2)	1	0
Coefficients d'accord	u KENDALL [54]	d_k (KENDALL) (τ)	- rangs - comparaisons par paires	$u=\frac{\sum_{i\neq j} A_{ij}(A_{ij}-1)}{\frac{n(n-1)}{2}\cdot\frac{p(p-1)}{2}}-1$ $=\frac{2}{p(p-1)}\sum_{k\neq l}\tau_{kl}$	1	$\frac{1}{p-1}$ si p est pair $\frac{1}{p}$ si p est impair
	v DEGENNE [28]	d_k (KENDALL) (τ)	- rangs - comparaisons par paires	$v=\frac{\frac{1}{2}\sum_{i\neq j}(2\ A_{ij}-p)^2}{p^2\ \frac{n(n-1)}{2}}$ $=\frac{1}{p^2}\sum_{kl}\tau_{kl}=\tau_{moy}$	1	0

(1) Pour la définition de S, voir les remarques ci-après.

(2) $\sum_{(i,j)\in O}$ signifie somme pour les couples d'objets compatibles avec l'ordre O, O étant l'ordre agrégé obtenu par la méthode de l'ordre médian. On trouvera un exemple de calcul de ce coefficient ainsi qu'une présentation de la méthode dans le deuxième chapitre (§ I, 3.3 et 3.4).

- Définition de la somme S dans le calcul du W de KENDALL :



$$S = \sum_{i=1}^{n} (p.\frac{n+1}{2} - \sum_{k=1}^{p} r_i^k)^2$$

Comme $\Sigma = p(1 + 2 + \ldots + n) = pn(n+1)/2$, alors $p(n+1)/2 = \Sigma/n$ et S est la somme des carrés des différences pour chaque objet entre le total des rangs et le total moyen des rangs Σ/n. W mesure l'ajustement à l'ordre obtenu par la méthode des sommes de rangs (deuxième chapitre, § I, 3.1).

- Signalons enfin d'autres indices utilisés dans la méthode d'analys des préférences ANAPREF (cf. II, 2.3) et définie dans [50] :

. L'<u>intensité</u> d'un classement k défini par des comparaisons pa paires est :

$$d_k = \frac{\sum_{i,j} a_{ij}^k}{n(n-1)/2}$$

Elle varie entre 0 et 1.

. L'<u>intensité</u> moyenne de p classements est :

$$D = \frac{1}{p} \sum_k d_k = \frac{\sum_{i,j} A_{ij}}{pn(n-1)/2} \quad \text{(varie entre 0 et 1).}$$

Dans le cas de réponses à des enquêtes où l'absence de comparaisons sur certaines paires d'objet peut signifier de la part du sujet une ignorance du ou des objets, l'intensité représente un indice de notoriété moyen. La notoriété d'un objet serait :

$$D_i = \frac{\sum_{j \neq i} A_{ij} + A_{ji}}{(n-1) \cdot p} \quad \text{(on a alors } D = \frac{1}{n} \sum_i D_i\text{).}$$

. La <u>puissance</u> d'un classement agrégé (collectif) est un indice (variant entre 0 et p) qui tient compte de la cohésion, de l'intensité et de l'effectif (nombre de classements) :

$$\boxed{P_o = C_o \cdot D \cdot p}$$

3. Conclusion sur la description par indices

L'ensemble des indices présentés ci-dessus permet déjà une description très fine d'un ensemble de classements. Si certains coefficients de corrélation ou distances entre classements contiennent déjà beaucoup d'information, ils en contiennent bien trop dès que le nombre de classements croît.

Si l'on connaît les distances entre 3 classements, on pourra visualiser ces trois classements sur le plan en construisant un triangle dont la longueur des côtés est proportionnelle aux distances mais comment procéder si on en possède davantage ?

De même, un coefficient d'ajustement de p ordre à un ordre moyen (concordance W ou cohésion C_o) peut, dans certains cas, être faible (voisin de 0) révélant ainsi d'importantes différences entre les classements. Peut-on avoir une idée de ces divergences, plus nuancée, c'est-à-dire moins globale qu'un simple indice ne le permet mais également suffisamment globale pour pouvoir les observer du même et seul coup d'oeil ? Répondre à ces questions est l'objet des méthodes de description par visualisation présentées dans le paragraphe suivant.

II - Description de classements par des méthodes de visualisation

L'objet des méthodes de visualisation est d'obtenir une représentation dans un même espace R^r de faible dimension ($r = 1, 2$ ou 3) les deux ensembles de points représentant les classements et les objets.

Pour obtenir cette visualisation, il faut être en mesure de définir d'une part les coordonnées de ces points dans un espace donné et d'autre part de projeter l'ensemble de ces points dans R^r.

On peut distinguer deux classes de méthodes de visualisation :

- Les méthodes de détermination des configurations dans R^r : elles consistent à situer les points objets et les points classements (nous dirons les configurations des objets et des classements) directement dans R^r de faible dimension. On cherche donc :

  . $X = \{\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_i, \ldots, \underline{x}_r\}$ la configuration des objets où $\underline{x}_i$ est un vecteur de R^r dont les composantes sont les coordonnées du point objet i et

  . $Y = \{\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_k, \ldots, \underline{y}_r\}$ la configuration des classements où $\underline{y}_k$ est un vecteur de R^r dont les composantes sont les coordonnées du point classement k.

- Les méthodes de projection du permutoèdre : elles consistent en une définition d'une double configuration des n objets et des $n!$ classements (ordres complets) possibles dans un espace de dimension élevée (R^{n-1}) (cette double configuration est appelée permutoèdre) puis en une projection dans R^r (R^2 ou R^3) de la partie du permutoèdre qui comprend les n objets et les p classements.

Les méthodes présentées sont :

- PREFMAP, PREFMAP II, MDPREF de J.D. CARROLL ; ce sont des méthode de détermination des configurations dans R^r ;

- l'analyse en composantes principales des classements de J.P. BENZ et J.P. FENELON, ANAPREF de E. JACQUET-LAGREZE ; ce sont des méthodes de projection du permutoèdre.

1. Les méthodes de détermination des configurations dans R^r

Ces méthodes sont caractérisées, on l'a vu, par une estimation direc dans R^r (r faible) des configurations des objets et des classements. distingue alors entre méthodes d'analyse externe et méthodes d'analyse interne (cf. J.M. BOUROCHE [18] et J.D. CARROLL [22]).

- Dans les premières (analyse externe), on se donne a priori la configuration des points objets obtenue par exemple à la suite d'une analyse des proximités sur les objets (1).

- Dans les secondes (analyse interne), on calcule la configuration des objets sans faire appel à une configuration extérieure.

1.1 La méthode PREFMAP

Nous n'exposerons que très succinctement les principes de cette méthode très largement développée par ailleurs (cf. J.M. BOUROCHE et P. LAY [18]).

PREFMAP est une méthode d'analyse externe (PREFMAP et PREFMAP II) ou interne (PREFMAP II).

---

(1) Il faut recueillir alors en plus des données de classements des données de proximité entre les objets (cf. Introduction). Les méthodes d'analyse des proximités utilisées peuvent être MDSCAL, TORSCA (cf. J.M. BOUROCHE [18]).

Cette méthode est basée sur un modèle où la distance entre les points objets et les points classements s'interprètent en terme de préférence (si les classements sont des données de préférence). Dans ce modèle, l'ordre des distances des n points objets à un point classement k doit être autant que possible compatible avec ce classement.

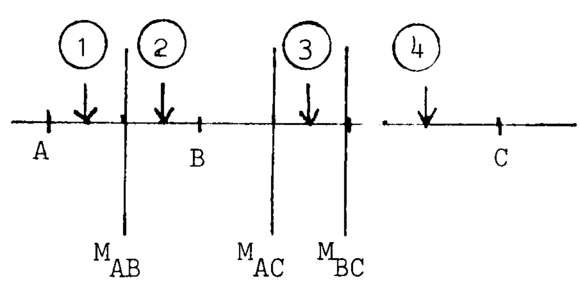
Si i et j sont deux objets, k un classement, d_{ik} et d_{jk} les distances des points i et j au point k, ce modèle est tel que :

$$d_{ik} < d_{jk} \Longleftrightarrow \begin{cases} r_i^k < r_j^k \\ s_i^k > s_j^k \\ a_{ij}^k = 1 \end{cases}$$

Le premier modèle de distance objet-classement ainsi introduit fut le modèle unidimensionnel de COOMBS (1950) (cf. COOMBS [23], TORGERSON [86], GREEN et CARMONE [41]).

Ce modèle est celui d'une configuration d'objets et de classements unidimensionnel, c'est-à-dire linéaire. Les points objets sont sur une droite, un point classement compatible avec la configuration objet est alors un point de la droite.

Exemple $n = 3$
(3 objets A, B, C)



Les milieux M_{AB}, M_{AC}, M_{BC} des segments AB, AC, BC partagent la droite en 4 régions 1 , 2 , 3 , 4 . Chaquepoint de ces 4 régions représente un ordre compatible avec la configuration ABC :

1 : classement $A \succ B \succ C$
2 : classement $B \succ A \succ C$
3 : classement $B \succ C \succ A$
4 : classement $C \succ B \succ A$

On constate que les classements $A \succ C \succ B$ et $C \succ A \succ B$ ne peuven être représentés par ce modèle.

Pour pouvoir représenter les $3!$ c'est-à-dire les 6 ordres possible il faudrait avoir recours à deux dimensions.

De plus, si n croût, le nombre d'ordres compatibles à une configuration unidimensionnelle devient très faible devant les $n!$ ordres possibles : ce nombre est $n(n - 1)/2 + 1$.

- BENNET et HAYS [9] en 1960 ont cherché à généraliser le modèle d COOMBS en tâchant de répondre à la question suivante : quelle est la dimension minimale d d'un espace qui permet de construire une configurati d'objets compatible avec p classements donnés sachant que, pour $d = n$ on peut toujours représenter les $n!$ ordres complets ?

La démarche suivie par J.D. CARROLL dans PREFMAP est différente. O s'impose la dimension $r << n - 1$ de l'espace et on cherche une configuration des objets et classements qui s'ajuste au mieux.

Dans la version d'analyse externe de PREFMAP II et dans PREFMAP, on connaît dans R^r la configuration des <u>points objets X</u>.

Soit $X = (\underline{x}_1, \ldots, \underline{x}_i, \ldots, \underline{x}_n)$ cette configuration ; $\underline{x}_i$ est un vecteur de R^r et représente les coordonnées du point objet i.

On cherche alors à évaluer dans R^r la configuration des <u>points cla sements y</u>.

$Y = (\underline{y}_1, \ldots, \underline{y}_k, \ldots, \underline{y}_p)$ où $\underline{y}_k$ est un vecteur de R^r et représente les coordonnées du point classement k.

Si r_i^k est le rang de l'objet i suivant le classement k, on cherche Y tel que :

$$r_i^k = M_k(a_k\, d_{ik}^2 + b_k + e_{ik}) \quad \text{où :}$$

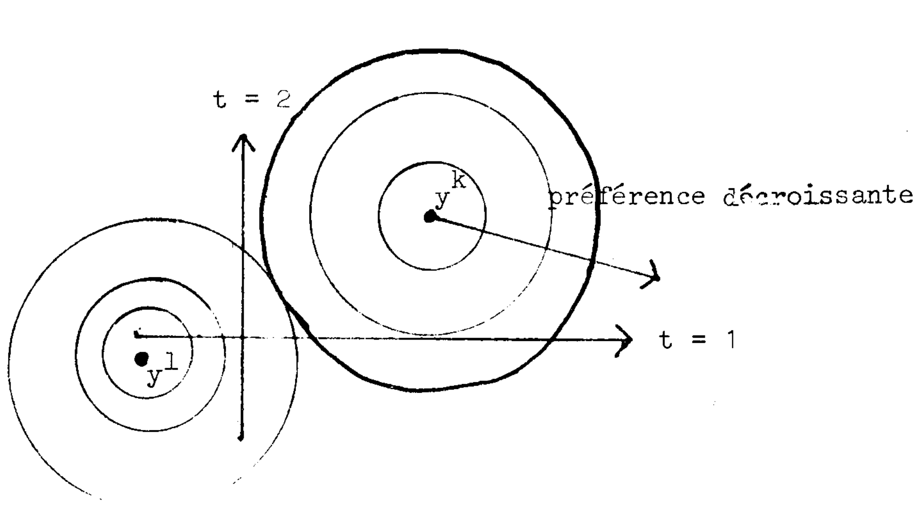
. M_k est une fonction monotone croissante à estimer dans les modèles appelés non métriques et où $M_k(x) = x$ dans les modèles appelés métriques ;
. d_{ik} est la distance dans R^r de l'objet i au classement k ;
. $a_k > 0$ et b_k sont des constantes à estimer ;
. e_{ik} est un terme d'erreur que l'on veut aussi faible que possible.

La méthode d'analyse de PREFMAP s'appuie sur plusieurs modèles qui diffèrent par la façon dont on calcule la distance d_{ik}. Sans entrer dans les détails, précisons qu'on distingue les modèles suivants :

- <u>Modèle de distance proprement dit</u>

Les courbes d'isopréférence d'un point classement sont des cercles (sphères ou hypersphères)

$$\boxed{d_{ik}^2 = \sum_{t=1}^{r} (y_k^t - x_i^t)^2}$$



- <u>Modèle de distance avec pondération des axes</u>

Dans ce modèle, on suppose que chaque classement k possède sa prop pondération des axes t de R^r.

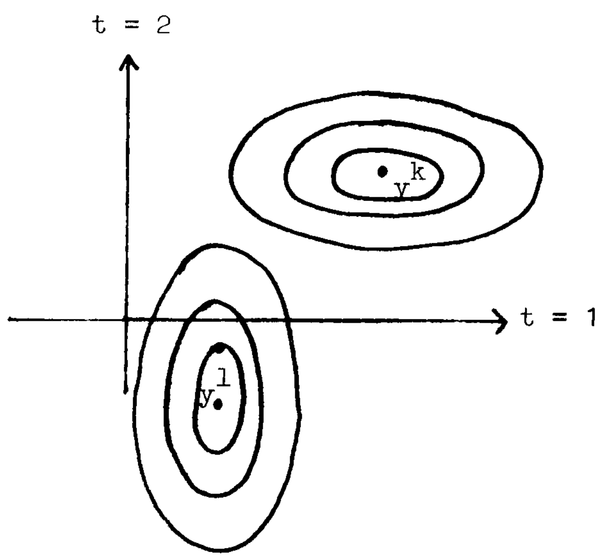
Si w_k^t est le poids donné par k pour l'axe t, on a :

$$\boxed{d_{ik}^2 = \sum_{t=1}^{r} w_{kt}(y_k^t - x_i^t)^2}$$

Les surfaces d'indifférences sont des ellipsoïdes dans R^r dont les axes principaux d'inertie sont les axes t (et sont donc les mêmes pour tous les classements).

Pour y^k : $w_k^1 > w_k^2$

Pour y^1 : $w_1^1 < w_1^2$



- <u>Modèle de distance avec pondération et rotation des axes</u>

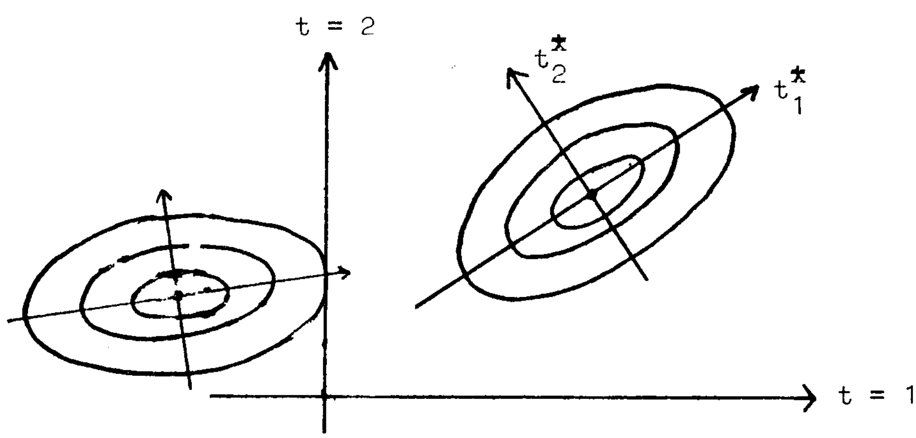
Dans ce modèle, on suppose que chaque classement pondère des axes déduits des axes t par une rotation propre au classement. On a :

$$d_{ik}^2 = \sum_{t^*=1}^{r} w_i^{t^*}(y_k^{*t^*} - x_i^{*t^*})^2$$

où $\underline{y}_k^*$ et $\underline{x}_i^*$ sont les coordonnées des points k et i dans le nouveau repère déduit du repère commun par rotation.

Matriciellement, si T_1 est la matrice orthogonale de rotation du classement k, on a

$$X^* = T_i X$$
$$Y^* = T_i Y$$



- <u>Modèle de distance avec poids négatifs</u>

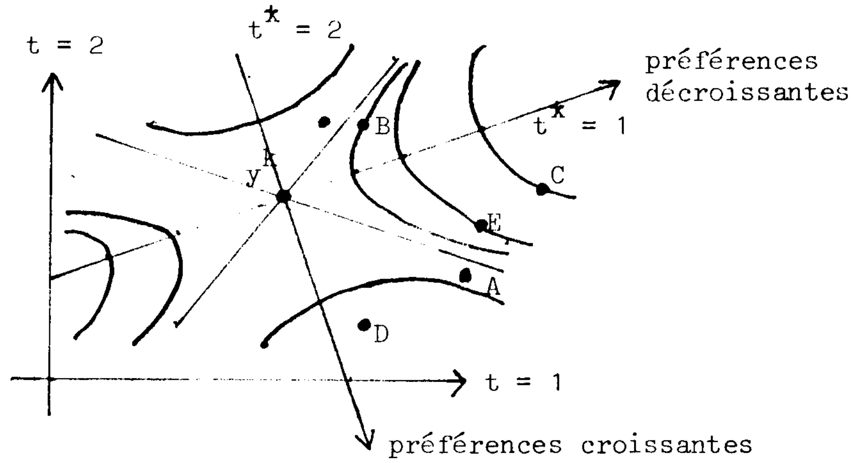
Lorsqu'on cherche dans les 3 modèles ci-dessus à estimer les poids w_i^t, on rencontre parfois des poids négatifs. Dans ce cas, pour l'axe t considéré, l'interprétation se fait en sens inverse : plus faible est la distance, moins bonne est la préférence (moins bon est le classement).

Dans le cas général, les courbes d'isopréférences dans R^r peuvent être des hyperboles :

$$w_i^{t_1^*} > 0$$

$$w_i^{t_2^*} < 0$$

Exemple : pour k on a l'ordre $D \succ A \succ B \succ E \succ C$.



Le modèle le plus général est donc le 3e modèle avec possibilité de poids négatifs.

On trouvera dans [18] et [19] la méthode de résolution permettant d'estimer la configuration X ainsi que les poids w_i^t et les matrices d rotation des axes dans le cas général.

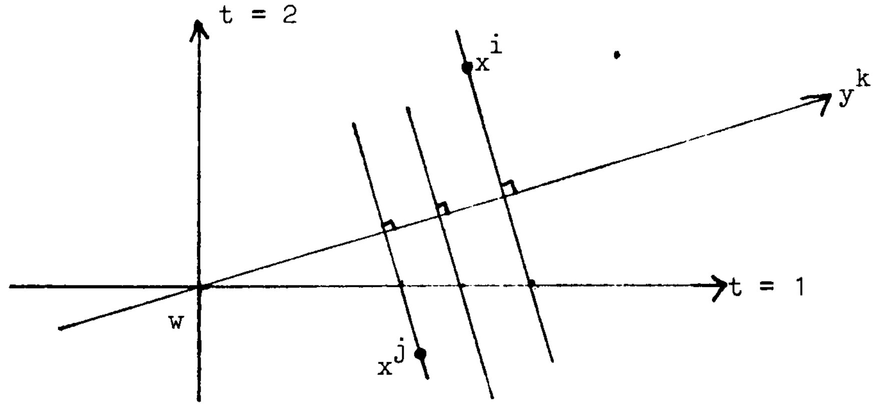
- La méthode PREFMAP II comprend également un modèle d'analyse externe. On ne se donne pas la configuration des points objets ; il s'agit donc d'estimer simultanément la configuration des points objets et des points classements dans R^r pour r donné

$$X = (\underline{x}_1, \ldots, \underline{x}_i, \ldots, \underline{x}_n)$$
$$Y = (\underline{y}_1, \ldots, \underline{y}_k, \ldots, \underline{y}_p)$$

Ce modèle est un modèle de distance car on cherche comme dans PREFMA une liaison monotone entre les rangs r_i^k et les carrés des distances d_i^2

Les modèles vectoriels

Ces modèles sont en fait un cas particulier des modèles de distance. En effet, les points classements sont rejetés à l'infini. Un classement est donc représenté par une direction, droite passant par w, le centre de gravité dans R^r de la configuration des objets $X = (\underline{x}_1, \ldots, \underline{x}_i, \ldots, \underline{x}_n)$



Les courbes d'isopréférences deviennent alors des droites perpendiculaires à la direction (w, y^k), c'est-à-dire au vecteur y^k (sur la figure, i est avant j pour le classement k).

Le modèle PREFMAP peut être utilisé comme modèle vectoriel. Dans ce cas, les vecteurs $\underline{y}^k$ recherchés sont de norme égale à 1 : $\sum_{t=1}^{r} y_k^{t2} = 1$.

Les directions sont en fait représentées par des points sur l'hypersphère de centre w et de rayon 1 dans R^r.

1.2 La méthode MDPREF (1) [18], [22]

Contrairement aux méthodes précédentes, la méthode MDPREF permet d'analyser des classements obtenus sous forme de comparaisons par paires.

---

(1) Multi-Dimensional analysis of PREFerence data.

Les données sont alors les suivantes [1] :

$a^k_{ij} = +1$ si i est avant j $\quad i \longrightarrow j$

$a^k_{ij} = -1$ si j est avant i $\quad i \longleftarrow j$

$a^k_{ij} = 0$ si i et j sont ex-aequo $\quad i \rightleftarrows j$

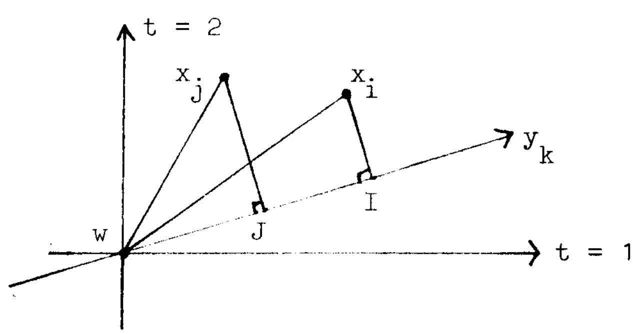
Ce modèle est un modèle d'analyse interne. Il s'agit donc d'évaluer

$X = (\underline{x}_1, \ldots, \underline{x}_i, \ldots, \underline{x}_n)$

$Y = (\underline{y}_1, \ldots, \underline{y}_k, \ldots, \underline{y}_p)$

Ce modèle est vectoriel. Par conséquent, on aura :

$$\sum_{t=1}^{r} y_k^{t^2} = 1 \qquad \forall \; k$$



Critère d'une "bonne configuration"

Pour le triplet (i, j, k), la configuration est satisfaisante si

$$a^k_{ij} = +1 \Longleftrightarrow \overline{wI} > \overline{wJ} \text{ sur l'axe } w_{yk}.$$

---

(1) La façon de représenter les comparaisons par paires est propre à cett méthode et ne correspond pas à la représentation donnée dans l'introductic

Or, dans R^r, $\overline{wI} = < \underline{y}_k, \underline{x}_i > = \hat{s}_{ki}$ produit scalaire des deux vecteurs ($\|y_k\| = 1$). Par conséquent :

$a^k_{ij} = +1 \Longleftrightarrow \delta^k_{ij} = (\hat{s}_{ki} - \hat{s}_{kj}) > 0$ si la configuration est correcte pour (i, j, k).

Une bonne configuration des n points x_i et des p points y_k est une configuration qui maximise

$$\sum_k \sum_{i \neq j} \delta^k_{ij} \, a^k_{ij}.$$

Pour comparer deux configurations, il faut normaliser ce critère. Dans MDPREF, on cherche à maximiser

$$C = \frac{\left(\sum_{i \neq j} \delta^k_{ij} \, a^k_{ij} \right)^2}{\sum_{i \neq j} (\delta^k_{ij})^2}$$

<u>Méthode d'estimation de X et Y</u>

On calcule une matrice de score (s_{ki}) par

$$s_{ki} = \sum_{j \neq i} a^k_{ij} - a^k_{ji}$$

J.D. CARROLL [22] montre que rechercher X, Y maximisant le critère C revient à rechercher X et Y minimisant :

$$\|S - \hat{S}\|^2 \quad \text{si} \quad \hat{S} = Y' X \quad \text{(produits scalaires).}$$

Cette estimation peut se faire par une méthode analytique : la méthode d'ECKART et YOUNG [32].

Remarque sur la visualisation obtenue par la méthode

Si on calcule les configurations dans R^r $(r = 2)$, les points classements se trouveront situés sur le cercle de centre w et de rayon 1.

Si on calcule la configuration dans R^3 (ou R^r, $r > 3$), les points classements se trouveront situés sur la sphère (hypersphère) de centre w et de rayon 1. Dans ce dernier cas, une projection sur R^r de points si tués sur une sphère donnera des points situés sur le cercle et à l'intéri du cercle. Plus le sujet sera proche de w, plus mauvaise sera sa représentation dans le plan de projection considéré.

2. Les méthodes de projection du permutoèdre

Les deux méthodes que nous allons voir maintenant reposent sur des modèles où la configuration commune X et Y est définie dans R_{n-1} (le permutoèdre). La configuration est définie a priori, indépendemment des données que l'on possède mais la projection de la configuration dans R^2 ou R^3 dépend, par contre, des données ordinales que l'on possède.

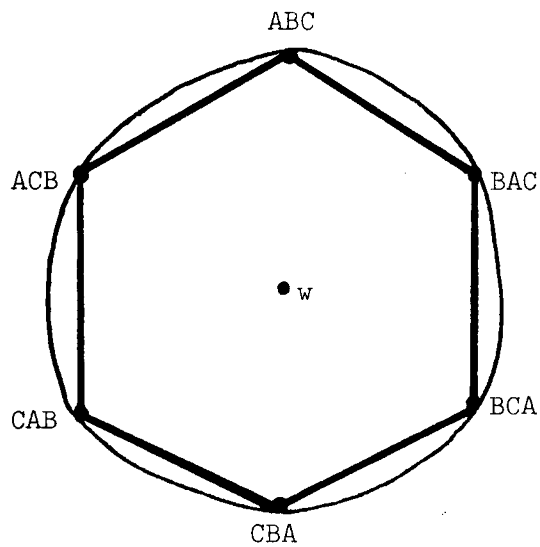
2.1 Les définitions du permutoèdre

Le permutoèdre est une figure géométrique permettant la représentatio des n! ordres complets (permutations) sur n objets. Cette notion, ainsi que le nom de permutoèdre, ont été introduits par G.T. GUILBAUD et P. ROSENSTIEHL en 1963 [44].

Cette figure est un polyèdre à n! sommets inscrits dans une hypersphère dans R^{n-1} de centre w et de rayon 1.

Dans le permutoèdre, chaque sommet (permutation) est symétrique par rapport à w du sommet représentant la permutation inverse.

Pour 3 objets, on a la figure suivante :



Deux permutations voisines ne diffèrent que par l'inversion de deux objets adjacents.

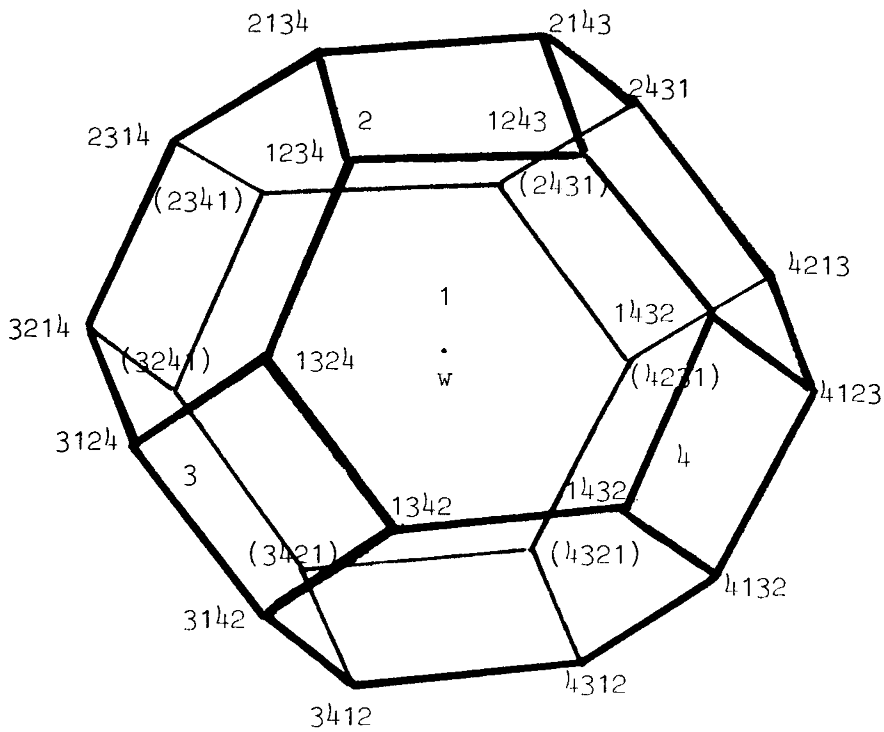
2.1.1 Définition du permutoèdre dans R^{n-1} muni de la distance euclidienne

La définition mathématique que nous allons donner a été proposée par J.P. BENZECRI [10] et conduit à un hexagone régulier dans R^r pour $n = 3$ et à la figure régulière présentée ci- après pour $n = 4$.

- Configuration des points objets $X = (\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_i, \ldots, \underline{x}_n)$

Dans un espace à n dimensions $(0, R^n)$, on situe chaque point objet i sur l'axe i au point d'abscisse 1. Les coordonnées de $\underline{x}_i$ sont donc dans $(0, R^n)$

$$\begin{cases} x_t^i = 0 & (t \neq i) \\ x_i^i = 1 & (t = i) \end{cases}$$



Permutoèdre à 4 objets {1, 2, 3, 4}

Chaque sommet du polyèdre représente une permutation de la suite 1, 3, 4. Quatre des faces hexagonales sont marquées d'un chiffre : chacune correspond aux permutations où son chiffre est le premier. Les éléments non-vus sont marqués en trait fin (arêtes) ou entre parenthèses (chiffres

Soit w le centre de gravité des points objets. Ses coordonnées sont dans $(0, R^n)$:

$$\left(\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}\right).$$

On effectue un changement d'origine de façon à ce que le nouveau repère (w, R^n) ait pour origine le centre de gravité w des points objets.

Dans ce nouveau repère (w, R^n), la configuration des points objets est :

$$(\underline{x}_i) \quad \begin{cases} x_t^i = -\frac{1}{n} & (t \neq i) \\ x_i^i = 1 - \frac{1}{n} & (t = i) \end{cases}$$

Remarques

- Tout vecteur x_i vérifie la relation :

$$\sum_{t=1}^{n} x_t^i = 1 - \frac{1}{n} + (n - 1)\left(-\frac{1}{n}\right) = 0$$

Par conséquent, tous les points objets appartiennent à l'hyperplan $\sum_{t=1}^{n} x_t = 0$, c'est-à-dire l'hyperplan passant par w et orthogonal au vecteur $(1, 1, 1, 1, \ldots, 1)$.

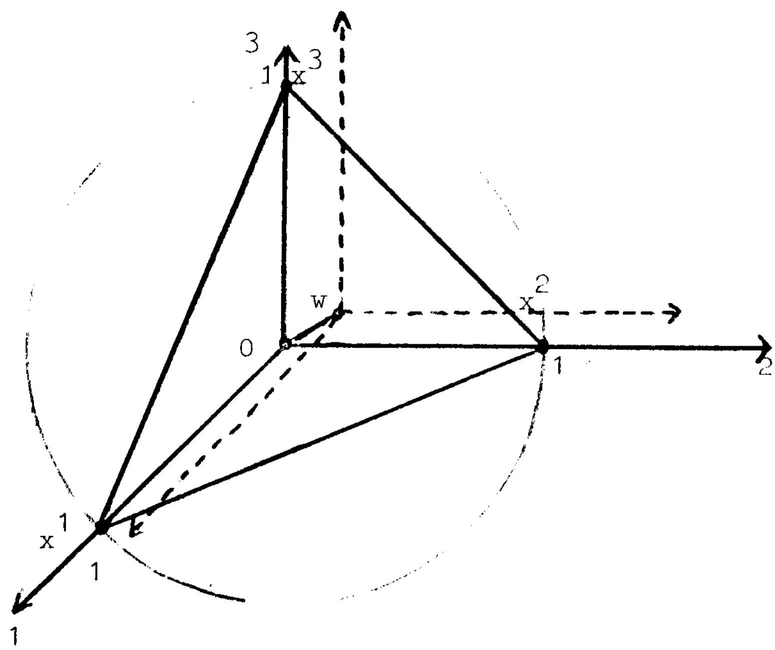
Les points objets sont donc tous situés dans un espace R^{n-1}.

- Tout vecteur x^i a pour norme $\sqrt{1 - \frac{1}{n}}$ dans R^n. En effet :

$$\sum_{t=1}^{n} x_t^{i^2} = \left(-\frac{1}{n}\right)^2 (n - 1) + \left(1 - \frac{1}{n}\right)^2 = \frac{n - 1}{n^2} + \frac{(n - 1)^2}{n^2}$$
$$= 1 - \frac{1}{n}$$

Donc tous les points objets i appartiennent à une hypersphère dans R^n de centre w et de rayon $\sqrt{1 - \frac{1}{n}}$. Comme les points objets appartiennent à un hyperplan passant par w et une hypersphère de centre w dans R^n, ils appartiennent à une hypersphère de centre w dans R^{n-1} et de rayon $\sqrt{1 - \frac{1}{n}}$.

Pour $n = 3$, les points objets appartiennent à une sphère de centre et de rayon $\sqrt{\frac{2}{3}}$ dans (w, R^3) et au cercle de centre w et de rayon $\sqrt{}$ dans (w, R^2)



- <u>Configuration des points classements</u> $Y = (\underline{y}_1, \ldots, \underline{y}_k, \ldots, \underline{y}_p)$

Dans (w, R^n), on situe le classement k au point de coordonnées

$$(\underline{y}_k) : y_k^t = (\frac{n + 1}{2} - r_t^k)/\sqrt{M} \quad \text{où} \quad M = n(n^2 - 1)/12$$

où r_t^k est le rang de l'objet t dans le classement y_k.

- Les points classements appartiennent au même hyperplan de (w, R^n) que les points objets. En effet :

$$\forall\, k, \sum_{t=1}^{n} y_k^t = \left[\frac{n(n + 1)}{2} - \sum_t r_t^k\right] / \sqrt{M} = 0$$

- Les points classements appartiennent à une hypersphère de centre w et de rayon 1 dans (w, R^n). En effet :

$$\sum_{t=1}^{n} (y_k^t)^2 = \frac{1}{M} \sum_{t=1}^{n} \left(\frac{n+1}{2} - r_t^k\right)^2$$

$$= \frac{1}{M} \left[\frac{n(n+1)^2}{4} - 2\,\frac{n+1}{2} \sum_t r_t^k + \sum_t (r_t^k)^2\right]$$

$$= \frac{1}{M} \left[\frac{n(n+1)^2}{4} - \frac{n(n+1)^2}{2} + \frac{n(n+1)\,(2\,n+1)}{6}\right]$$

$$= \frac{1}{M}\,\frac{n(n^2-1)}{12}$$

$$= 1$$

- Les points classements appartenant à une hypersphère de centre w et un hyperplan passant par w appartiennent donc à une hypersphère de centre w et de rayon 1 dans R^{n-1}.

On trouvera ci-dessus le permutoèdre pour n = 4.

<u>Permutoèdre et ρ de SPEARMAN</u>

Le coefficient de corrélation des rangs de SPEARMAN (cf. § I.1) s'interprète comme le cosinus de l'angle sous lequel on voit deux classements du permutoèdre.

En effet, entre deux classements k l, on a :

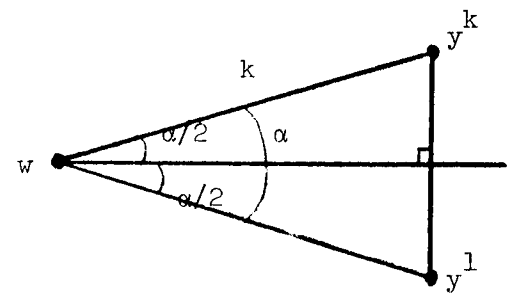
- d'une part : $\rho_{(k,l)} = 1 - \dfrac{6}{n^3 - n} \displaystyle\sum_{t=1}^{n} (r_t^k - r_t^l)^2$;

- d'autre part, dans le permutoèdre, la distance entre deux classements est :

$$d^2(k,l) = \sum_{t=1}^{n} (y_t^k - y_t^l)^2 = \frac{1}{M} \sum_t (r_t^k - r_t^l)^2$$

$$= \frac{12}{n^3 - n} \sum_t (r_t^k - r_t^l)^2$$

(1) $d^2(k,l) = 2\,(1 - \rho_{kl})$



Or, si on considère le triangle (w, y^l, y^k), on a :

$d(k,l) = 2 \sin \frac{\alpha}{2}$

(2) $d^2(k,l) = 4 \sin^2 \frac{\alpha}{2} = 2(1 - \cos \alpha)$

(1) et (2) montrent que :

$$\boxed{\cos \alpha = \rho_{k,l}}$$

2.1.2 Définition du permutoèdre dans la méthode ANAPREF [50]

La définition du permutoèdre dans R^{n-1} muni de la norme euclidienne permet une approche cohérente avec la définition du coefficient de corrélation des rangs de SPEARMAN ρ.

Il semble légitine de vouloir proposer une démarche analogue en liaison avec la distance de KENDALL et le coefficient de corrélation τ.

G.T. GUILBAUD et M. ROSENSTIEHL [44] proposent de définir l'angle de deux classements la quantité α telle que :

$$\cos \alpha = \tau = 1 - 2 \frac{d_K}{d_{K\ max}}.$$

Cependant, cette définition d'angle entre 2 classements ne permet pas de reconstituer l'hexagone régulier pour $n = 3$.

Afin de pouvoir reconstituer l'hexagone régulier pour $n = 3$ et en remarquant que la distance de KENDALL vaut 1 sur chaque arête du permutoèdre, on peut interpréter (cf. [50]) la distance de KENDALL comme une distance angulaire en posant :

angle entre deux classements :

$$\boxed{\alpha = \pi \frac{d_K}{d_K \max}}$$

Cette définition permet de reconstruire le permutoèdre pour $n = 3$

$$\boxed{\alpha = \frac{\pi}{2} (1 - \tau)}$$

La construction du permutoèdre est alors une construction sphérique.

Soit G1 et G2 deux ordres complets particuliers appelés pôles de construction du permutoèdre.

. Construction de la configuration des points classements

Pour chaque point classement k, on calcule les deux angles :

$$(k, G1) = \pi \frac{2\, d_k(k, G1)}{n(n-1)} \qquad \left(d_K \max = \frac{n(n-1)}{2}\right)$$

$$(k, G2) = \pi \frac{2\, d_k(k, G2)}{n(n-1)}$$

. Ces angles permettent de situer tout point k par rapport aux pôles G1 et G2 et permettent notamment de situer la projection de k sur le plan G1 G2.

- Construction de la configuration des points objets

On introduit la notion d'angle entre un objet i et un pôle G1 de la façon suivante. Considérons les ordres complets G1 et G2 suivants

$$G_1 : a \ldots\ldots d \boxed{i} l \ldots\ldots pq \ldots\ldots r$$

$$G_1^{\star} : r \ldots\ldots q \boxed{i} p \ldots\ldots ld \ldots\ldots a$$

où $G_1 - \{i\}$ et $G_1^{\star} - \{i\}$ sont deux ordres inverses définis sur les n - autres objets et où i occupe le même rang r_i dans les deux ordres.

Par définition, on appellera angle entre i et G1

$$\boxed{(i, G_1) = (i, G_1^{\star}) = \frac{1}{2} (G_1, G_1^{\star})}$$

On montre alors que [50] :

$$\boxed{(i, G_1) = \frac{\pi}{2} - \frac{\pi}{n(n-1)} (n + 1 - 2 r_i^{G_1})}$$

. Projection de relations de comparaisons par paires valuées

ANAPREF permet l'analyse et la visualisation de comparaisons par paires valuées :

$$\underline{a}^k = (a_{ij}^k) \quad \text{avec} \quad a_{ij}^k \in [0, 1] \quad \text{et}$$
$$a_{ij}^k + a_{ji}^k \leq 1 \quad \forall\, i \neq j$$

On obtient de telles relations après avoir agrégé par exemple un ensemble de classements. Le tableau $\underline{a}_k$ représente alors un classement moyen. On trouvera dans le troisième chapitre (§ II.4) un exemple d'une relation valuée $\underline{a}^k$, résultat de l'agrégation des préférences sur 25 métiers de 30 enfants.

Pour visualiser un classement $\underline{a}^k$, il faut pouvoir définir ses angles avec les pôles G1 et G2 ; on a par exemple :

$$k\,G_1 = \frac{2\,\pi}{n(n-1)}\,d_K(k, G_1) = \frac{\pi}{n(n-1)} \sum_{i,j} |a_{ij}^{k} - a_{ij}^{G_1}|$$

où $a_{ij}^{G_1} \in \{0, 1\}$.

- Soit k un classement représenté par $\underline{a}_k$. Un <u>classement k^* inverse</u> de k sera, par définition, représenté par :

$$\forall\, i,j \qquad a_{ij}^{k^*} = a_{ji}^{k}$$

On voit qu'en particulier, si k est un ordre complet, $a_{ij}^{k} \in \{0, 1\}$ et $a_{ij}^{k} + a_{ji}^{k} = 1\ \forall i \neq j$. Dans ce cas, $a_{ij}^{k^*} = 1 - a_{ij}^{k}$. k^* et k sont alors deux ordres inverses.

- Dans la méthode de visualisation ANAPREF, deux <u>classements inverses sont toujours représentés par deux points symétriques par rapport au centre w du permutoèdre</u>.

En effet, comme G_1 est un ordre complet :

$$\begin{aligned}(k, G1) &= \frac{\pi}{n(n-1)} \sum_{i,j} |a_{ij}^{k} - a_{ij}^{G_1}| \\ &= \frac{\pi}{n(n-1)} \left[\sum_{(i,j)\in G1} (1 - a_{ij}^{k}) + \sum_{(i,j)\notin G1} (a_{ij}^{k} - 0)\right] \\ &= \frac{\pi}{n(n-1)} \sum_{(i,j)\in G1} (1 - a_{ij}^{k} + a_{ji}^{k}) \quad \text{car}\end{aligned}$$

$$\sum_{(i,j)\notin G1} a_{ij}^{k} = \sum_{(i,j)\in G1} a_{ji}^{k}$$

De même $(k^*, G1) = \dfrac{\pi}{n(n-1)} \displaystyle\sum_{(i,j)\in G1} (1 - a_{ij}^{k^*} + a_{ji}^{k^*})$

$$(k, G1) + (k^*, G1) = \frac{\pi}{n(n-1)} \sum_{(i,j)\in G1} (2 - a_{ij}^{k} + a_{ji}^{k} - a_{ij}^{k^*} + a_{ji}^{k^*})$$

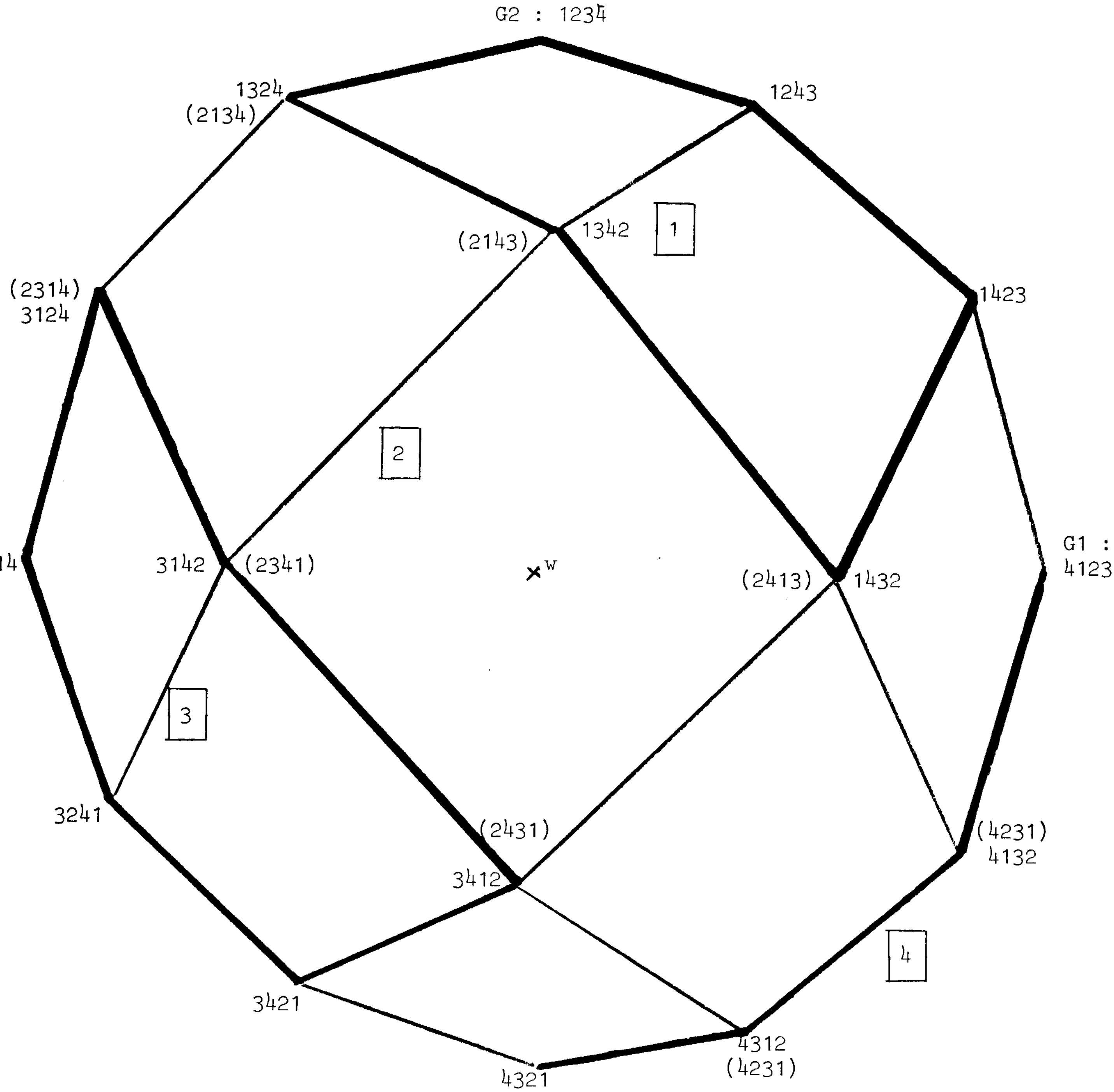
Or $a_{ij}^{k^*} = a_{ji}^{k}$; donc :

$$(k, G1 + k^*, G1) = \frac{\pi}{n(n-1)} \sum_{(i,j)\in G1} (2 + 0 + 0)$$

$$= \pi$$

De même : $(k, G2) + (k^*, G2) = \pi$.

Ces deux dernières égalités montrent que, dans les projections du permutoèdre, les relations inverses k et k^* apparaîtront toujours comme symétriques l'une de l'autre par rapport à w.



Projection du permutoèdre à 4 objets {1, 2, 3, 4} construit à partir des pôles G1 : 4123 et G2 : 1234.

2.2 La méthode d'analyse factorielle des classements de J.P. BENZECR et J.P. FENELON

Cette méthode d'analyse est due à J.P. BENZECRI [10] et à J.P. FENELO [36].

C'est une méthode de projection du permutoèdre tel qu'il est défini au § 2.1.1 ci-dessus.

L'analyse se fait en deux étapes. Au cours de la première étape, on étudie la position du centre de gravité G du nuage des points classements $Y = (y_k)$. La position de ce centre de gravité permet d'étudier un classement moyen. Au cours de la seconde étape, on effectue une analyse en composantes principales du nuage des points classements.

2.2.1 Etude du centre de gravité, étude d'un classement moyen

On a vu que, par construction de la configuration X des points objets, leur centre de gravité est le point w, origine du repère (w, R^n).

Dans ce même repère, les coordonnées des points classements sont :

$$y_t^k = \left(\frac{n + 1}{2} - r_t^k\right) / \sqrt{M}$$

Les coordonnées du centre de gravité G sont :

$$y_t^G = \frac{1}{p} \sum_{k=1}^{p} y_t^k.$$

. Il est intéressant d'étudier la norme du vecteur wG. Le point w centre de gravité des points objets, centre du permutoèdre, représente l'i différence. Si G est en w, cela signifie qu'il n'existe pas de classement moyen, ou plus précisément que le classement moyen est un préordre à une classe d'équivalence contenant tous les objets. A l'inverse, si G est sur la sphère, tous les classements k sont identiques. La norme de wG représente par conséquent un indice de concordance des classements.

Le carré de la norme de wG est le coefficient de concordance W de KENDALL

En effet, calculons $\sum_{t=1}^{n} {y_t^G}^2$

$$y_t^G = \frac{1}{p} \sum_{k=1}^{p} y_t^k = \frac{1}{p} \left[p \frac{n+1}{2} - \sum_{k=1}^{p} r_t^k \frac{1}{\sqrt{M}} \right]$$

$$\sum_{t=1}^{n} {y_t^G}^2 = \frac{1}{Mp^2} \sum_{t=1}^{n} (p \frac{n+1}{2} - \sum_{k=1}^{p} r_t^k)^2$$

On reconnaît l'expression du coefficient W (§ I, 1.3).

. Le vecteur wG apparaît comme un axe privilégié dans le permutoèdre. Il est alors intéressant de projeter le permutoèdre sur cet axe moyen

- projection des points objets

L'abscisse du point objet sur l'axe wG sera :

$$\frac{1}{\sqrt{W}} < wx^i, wG > = \frac{1}{\sqrt{W}} \sum_{t=1}^{n} x_t^i \, y_t^G$$

où W est le coefficient de KENDALL

- projection des points classements

L'abscisse du point classement sera de même :

$$\frac{1}{\sqrt{W}} < wy^k, wG > = \frac{1}{\sqrt{W}} \sum_{t=1}^{n} y_t^k \, y_t^j$$

- La projection des objets sur l'axe moyen correspond au classement moyen obtenu par la méthode de la somme pondérée (cf. deuxième chapitre, § I, 3.1). En effet, on peut montrer que :

$$\frac{1}{\sqrt{W}} < wx^i, wG > = \frac{1}{\sqrt{MW}} (\frac{n+1}{2} - \frac{1}{p} \sum_{k=1}^{p} r_i^k) = \frac{y_i^G}{\sqrt{W}}$$

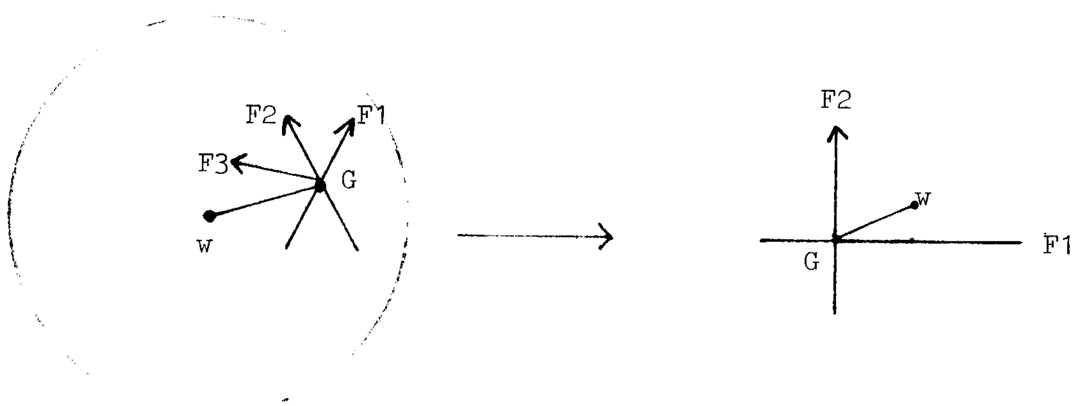
ce qui montre que cette abscisse est une fonction linéaire du score obtenu par la méthode de la somme des rangs. Plus grande est l'abscisse, plus faible est le rang moyen, donc meilleure est la position de l'objet dans le classement moyen.

- La projection des classements s'interprète en terme d'accord avec le classement moyen. Plus le classement est éloigné de w dans la direc tion de G, plus grand est l'accord entre ce classement et le classement moyen.

2.2.2 Analyse en composantes principales du nuage des points classe

Après avoir étudié le classement moyen, la seconde étape consiste à décrire l'ensemble des classements dans leurs différences. On effectue pour cela une analyse en composantes principales du nuage des points clas sements. On projette ensuite ce nuage dans le nouvel espace des facteurs (G, R^r) où les axes $F_1, \ldots, F_r$ sont les axes principaux d'inertie du nuage.

Schématiquement, on a :



Comme on connaît dans (w, R^n) les coordonnées des points objets, il est possible de projeter également l'ensemble des points objets, ce qui donne une double visualisation des ensembles d'objets et des ensembles de classements. L'interprétation se fait alors en terme de proximité entre points classements entre eux et de proximité entre points classements et points objets.

- Ces deux étapes constituent la méthode d'analyse factorielle des rangs de J.P. BENZECRI et J.P. FENELON. J.P. BENZECRI propose également une extension de cette méthode d'analyse au cas où l'on connaîtrait plusieurs classements de plusieurs ensembles d'objets donnés par les mêmes individus.

2.3 La méthode ANAPREF [50]

ANAPREF est une méthode de description et de structuration de données ordinales :

- c'est une méthode de description car elle permet de projeter le permutoèdre, donc de visualiser l'ensemble des classements et l'ensemble des objets ;
- c'est une méthode de structuration :
. des classements : on recherche une typologie en 2 ou 3 groupes de classements,
. des objets: pour l'ensemble des classements et, pour chaque groupe de la typologie, on recherche des classements représentatifs.

Les données sont, comme dans MDPREF, des rangs, des évaluations sur des échelles ou des comparaisons par paires. Dans ce dernier cas, les comparaisons peuvent être des relations valuées (à condition que $a^k_{ij} + a^k_{ji} \leq 1 \quad \forall$ ijk ; cf. § 2.1.2).

La méthode d'analyse comprend les phases suivantes :

2.3.1 Recherche d'un classement moyen par la méthode de l'ordre médian (cf. chapitre II, § I, 3.3 et 3.4) pour l'ensemble des p classements.

2.3.2 Constitution d'une première typologie des classements en éliminant les classements corrélés négativement (tau de KENDALL) avec l'ordre médian du groupe principal. Cette première typologie a pour but de mettre en évidence un groupe de classements dont la puissance (cf. I, 2) soit plus grande que celle de l'ensemble des classements. On trouvera dans [50] un organigramme détaillé de l'algorithme conduisant à cette première typologie.

Cette typologie conduit à 1, 2 ou 3 groupes dont le groupe principal (le plus puissant) comprend des classements appartenant tous à la même moitié du permutoèdre.

Pour chacun des groupes, on détermine le classement moyen (ordre médian) et les indices de cohésion et de puissance (cf. § I, 2).

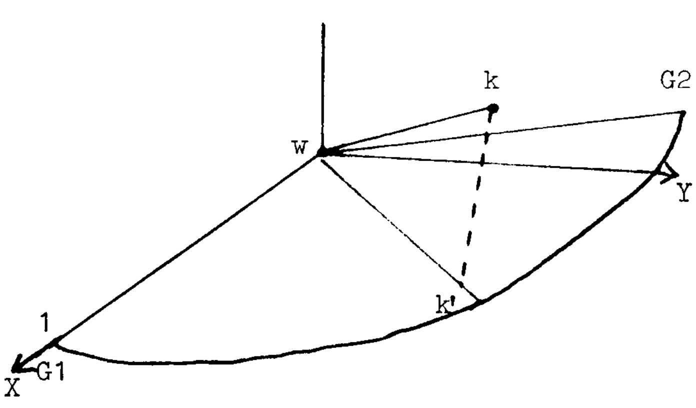
2.3.3 Constitution d'une deuxième typologie des classements en partant de la première lorsqu'il y a 2 ou 3 groupes et en créant artificiellement, lorsqu'il n'y a qu'un groupe, un second groupe qui comprend le classement le plus divergent du classement moyen. Cette seconde typologie a pour but de rééquilibrer les groupes numériquement. On affecte chaque classement au groupe dont le classement moyen est le plus proche du classement considéré puis on réitère le processus jusqu'à ce qu'il y ait équilibre. Chaque classement est alors affecté à un groupe tel qu'il se trouv plus proche du classement moyen de ce groupe que des autres classements moyens. Pour les nouveaux groupes ainsi constitués, on détermine également les classements moyens par la procédure de l'ordre médian et les indices de cohésion et de puissance qui lui sont associés.

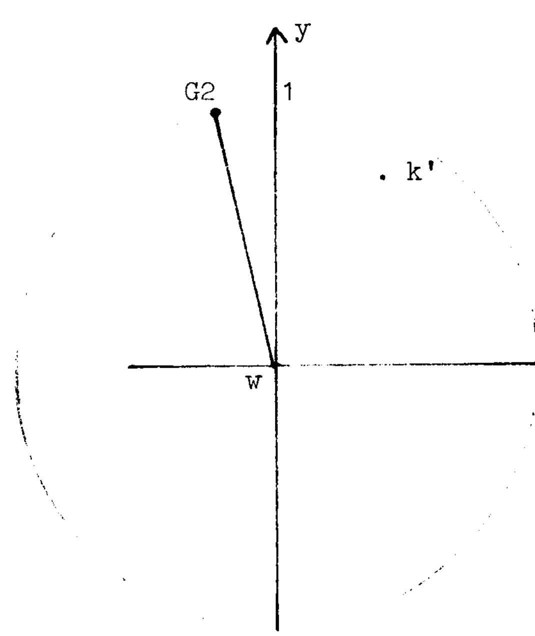
Cette seconde typologie conduit le plus souvent à 2 groupes (parfois

2.3.4 Projection du permutoèdre sur le plan passant par w (centre du permutoèdre) et les deux ordres médians G1 et G2 des deux groupes les plus importants déterminés par la typologie précédente.

Pour tout classement k, on sait évaluer les angles (k, G1) et (k, G2).

Pour tout objet i, on sait de même évaluer les angles (i, G1) et (i, G2) (cf. §II, 2.1.2). Dans ce plan, on situe le point G1 au point d coordonnées (0, 1) et le point G2 sur le cercle de centre w et de rayon 1 à un angle (G1, G2) :





Les formules de projection de k sont les suivantes [50] :

$X'_k = \cos(k, G1)$

$Y'_k = [\cos(k, G2) - \cos(k, G1)\cos(G1, G2)] / \sin(G1, G2).$

3. Remarques sur les méthodes de visualisation

Si nous avons essayé de différencier les méthodes de visualisation au niveau des modèles, il nous paraît beaucoup plus difficile de guider l'utilisateur dans le choix qu'il sera amené à faire lorsqu'il aura à analyser des préférences.

Néanmoins, au niveau des méthodes, on peut faire les quelques remarques suivantes :

3.1 Si l'on possède des données de proximités en plus des données ordinales, c'est alors PREFMAP I ou II, en version d'analyse externe, qu'il convient d'utiliser.

3.2 La méthode d'analyse de BENZECRI est une méthode d'analyse en composantes principales de classements donnés à partir de rangs et semble donc particulièrement bien adaptée lorsque les classements sont obtenus sous cette dernière forme.

3.3 Les méthodes MDPREF et PREFMAP II conviennent particulièrement bien pour une analyse très fine des classements. MDPREF est relativement souple puisqu'il permet d'analyser des classements sous la forme de comparaisons par paires.

3.4 ANAPREF, méthode d'analyse et d'agrégation de classements, semble adaptée lorsqu'on se préoccupe davantage de problèmes de structuration (recherche d'ordre sur les objets, typologie de classements). Cette méthode semble être également très souple au niveau des entrées puisqu'elle permet d'analyser des comparaisons par paires valuées (relations agrégées par exemple). Par contre, la visualisation est moins fine que dans les autres méthodes puisque seule une visualisation dans R^2 est possible.

Au niveau des programmes, on pourra guider son choix à l'aide de ces quelques remarques et surtout à l'aide de la grille de présentation proposée dans le troisième chapitre.

Sinon, seule l'expérience et le temps seront à même de révéler les différents intérêts de ces méthodes qui présentent certaines particularités complémentaires.

A un niveau plus théorique, il reste des recherches à poursuivre sur les différents modèles de visualisation des préférences, les hypothèses psychologiques sous-jacentes à ces modèles, une approche plus approfondie du lien existant entre préférence et proximité étant peut-être une recherche préalable nécessaire.

CHAPITRE II

LES METHODES DE STRUCTURATION

Ces méthodes ont pour objet de dégager des structures particulières portant soit sur l'ensemble des objets, soit sur l'ensemble des classements.

La structure recherchée pourra être :

- une relation d'ordre (§ I) : c'est le problème de l'agrégation de plusieurs classements en une relation unique (relation d'ordre le plus souvent) mais c'est également le problème d'approximation d'une relation quelconque (relation de tournoi par exemple) par une relation binaire donnée (relation d'ordre le plus souvent) (cf. § I, 3.2.2) ;

- une équivalence (§ II) : c'est le problème de détermination de typologies. On ne s'intéressera à ce problème que dans la mesure où cette typologie portera sur un ensemble de classements, c'est-à-dire de relations de comparaisons ordinales ;

- une structure particulière (§ III) se référant à un modèle ; on a vu l'exemple du modèle unidimensionnel de COOMBS, on verra d'autres exemples de structures particulières liées aux propriétés algébriques des relations binaires (M. BARBUT et L. FREY [5]) ;

- une relation d'ordre (§ IV) à partir de données le plus souvent dichotomiques (réponses : oui-non) : c'est le modèle d'analyse hiérarchique de L. GUTTMAN [45].

I - Méthodes d'agrégation ordinales

Les méthodes proposées ont pour objet d'agréger différentes relations d'ordre en une relation unique, un préordre complet en général.

Le problème peut être posé dans ces termes pour la recherche :

- d'une opinion ou préférence collective ;
- d'une agrégation multicritère ;
- d'un problème mixte dans lequel il existe des critères et des juge

L'objet n'est pas d'exposer ici la démarche qui peut conduire à proposer des critères et leurs indicateurs. Les méthodes ci-dessous supposen les étapes préalables effectuées. Le problème peut alors être posé dans les termes suivants : peut-on dégager un préordre unique à partir de plusieurs préordres ?

Il peut être intéressant de rapprocher toutes les méthodes d'agrégation au célèbre théorème d'ARROW [1] qui nous montre l'impossibilité d'imaginer des méthodes respectant un certain nombre de conditions jugées raisonnables.

Parmi les méthodes, nous distinguerons entre des méthodes purement ordinales, celles qui ne tiennent compte que des relations d'ordre sur les objets et des méthodes mixtes qui tiennent compte de distances sur des échelles initiales (ELECTRE, agrégation de quasi-ordres : § I, 4).

1. La relation de dominance

Si un objet i est préféré à un objet j sur tous les classements, on dit que <u>i domine j</u> :

$$i \succ j \Longleftrightarrow \{\forall\, k,\ r_i^k \leq r_j^k\} \qquad \text{(rangs)}$$

ou encore :

$$i \succcurlyeq j \iff \{\forall\, k,\ s_i^k \geq s_j^k\} \qquad \text{(évaluations)}$$

ou encore :

$$i \succcurlyeq j \iff \{\forall\, k,\ a_{ij}^k \neq 0\} \qquad \text{(paires)}$$

Si pour un classement k on a préférence stricte, on dit que <u>i domine strictement j</u> ($i \succ j$).

La relation de dominance est un ordre partiel sur les objets

- c'est un ordre parce que cette relation est transitive : $i \succcurlyeq j$ et $j \succcurlyeq l \Longrightarrow i \succcurlyeq l$ (1) ;
- cet ordre est partiel car tous les objets ne sont pas nécessairement comparables.

Déterminer l'ordre partiel obtenu par la relation de dominance est ce qu'on appelle également déterminer l'ordre partiel produit des p classements initiaux.

Le problème inverse a été posé par BEN DUSHNIK et E.W. MILLER [8] : peut-on, étant donné un ordre partiel, trouver un nombre minimum d'ordres complets tels que leur ordre produit (leur intersection) soit l'ordre partiel donné ? Ce nombre minimum est appelé <u>dimension de l'ordre partiel</u>.

Cette démarche est en fait une méthode d'analyse. Elle revient à essayer de dégager des dimensions sous-jacentes (les ordres complets recherchés) sur lesquels se référeraient les sujets en donnant leurs classements. Le problème dans sa forme algébrique est complexe et à notre connaissance mal résolu. A. DUCAMP [30] a donné un exposé récent du problème.

---

(1) En effet : $i \succcurlyeq j \iff r_i^k \leq r_j^k \quad \forall\, k$

$j \succcurlyeq l \iff r_j^k \leq r_l^k \quad \forall\, k$

or : $r_i^k \leq r_j^k$ et $r_j^k \leq r_l^k \Longrightarrow r_i^k \leq r_l^k$ donc

$r_i^k \leq r_l^k \ \forall k$ et $i \succcurlyeq l$.

La relation de dominance permet donc d'obtenir un classement partiel des objets. Cependant, l'incomparabilité demeure souvent très grande. Comment enrichir la relation de dominance ? Ne peut-on pas obtenir un préordre complet sur les objets ? Il est intéressant alors de suivre K.J. ARROW.

2. Les conditions et le théorème d'ARROW

On peut se demander avec K.J. ARROW [1] (ATTALI [3]) à quelles conditions il est possible d'agréger plusieurs préordres (classements) en un préordre unique.

ARROW s'est posé le problème dans le cadre de la théorie des choix collectifs, chaque préordre représentant une préférence individuelle. Cependant, que les préordres représentent des préférences individuelles ou des préférences d'un décideur sur des critères ne changent pas la nature de la question.

ARROW a présenté un certain nombre de conditions que devrait satisfair toute procédure d'agrégation :

C1 - Une préférence individuelle (ou sur un critère) est un préordre complet sur les objets et tous les préordres sont admissibles (1).

C2 - La préférence collective (ou multicritère) est un préordre complet sur les objets.

C3 - Si on améliore le classement d'un objet dans un préordre, toutes choses étant égales par ailleurs, le classement de cet objet ne doit pas baisser dans le préordre de la préférence collective.

C4 - La préférence collective sur un sous-groupe d'objets ne dépend que des préférences individuelles relatives à ce sous-groupe (en particulier, si un objet est retiré de l'ensemble des objets, l'ordre de préférenc collective sur les autres objets ne doit pas être modifié). Cette conditic est parfois appelée indépendance vis-à-vis des choix extérieurs.

---

(1) La condition ne porte en fait que sur 3 objets au moins.

C5 - Aucune préférence collective n'est a priori interdite (il existe toujours un système de préférences individuelles donnant, à l'aide de la procédure d'agrégation, un préordre collectif donné).

C6 - Aucun individu (ou critère) ne peut imposer son préordre quels que soient les préordres des autres individus (critères) (condition de non-dictature d'un préordre).

Il semble "raisonnable" d'exiger de toute méthode ou procédure d'agrégation qu'elle respecte chacune de ces conditions.

Pourtant ARROW a démontré que ces six conditions sont incompatibles (1).

Toute méthode d'agrégation existante ou imaginable violera par conséquent une ou plusieurs de ces six conditions. En général, les conditions non respectées sont :

C1 - On restreint l'ensemble des préordres admissibles.
C2 - La préférence collective n'est pas un préordre complet.
C4 - La préférence collective dépend des choix extérieurs.

M. VICKREY [87] a exposé une légère variante des conditions d'ARROW en faisant intervenir la relation de dominance :

C1 - Tous les préordres sont admissibles.
C2 - La préférence collective est un préordre complet.
C4 - Indépendance vis-à-vis des choix extérieurs.
C6 - Condition de non dictature d'un préordre.
C7 - La préférence collective respecte la relation de dominance (principe d'optimalité de PARETO).

Dans cette variante, les conditions C3 et C5 sont remplacées par la condition C7 d'optimalité.

---

(1) On pourra se reporter à J. ATTALI [3] pour un exposé récent et une démonstration du théorème en français.

Ces conditions sont légèrement plus fortes que celles d'ARROW. Elles sont également incompatibles.

Nous allons présenter dans le paragraphe suivant un certain nombre de méthodes violant toute une et une seule de trois conditions C1, C2, C4 exceptée la méthode d'agrégation lexicographique qui viole la condition C

3. Exemples de méthodes ordinales

Rappelons la relation de dominance. Cette relation est un ordre partiel et ne conduit pas en général à un préordre complet. Adopter la relation de dominance comme règle d'agrégation, c'est choisir une méthode qui ne respecte pas la condition C2.

Dans les autres méthodes que nous allons voir, on utilise éventuellement des poids p_k pour chaque préordre k. Dans le cadre d'une agrégati de préférences individuelles, ces poids seront le plus souvent pris égaux à 1. Dans celui d'une agrégation de plusieurs critères en un critère uni les poids seront le moyen pour le décideur d'exprimer sa préférence entre les différents critères.

3.1 Somme pondérée des rangs

Cette méthode est très simple et très connue. Pour chaque objet i, on calcule son score :

$$s(i) = \sum_{k=1}^{p} p_k \, r_i^k .$$

s(i) représente alors une fonction de valeur permettant de classer tous les objets en un préordre complet : plus s(i) est faible, plus grande est la préférence pour i.

La condition C4 (indépendance vis-à-vis des choix extérieurs) n'est pas satisfaite (1).

---

(1) Exemple : Soit 3 objets a, b, c et les 7 préordres sur ces objets :

$a \succ b \succ c$ 3 fois
$b \succ c \succ a$ 2 fois
$c \succ a \succ b$ 2 fois.

Les rangs sont les suivants :

	a	b	c	p_k
a b c	1	2	3	3
b c a	3	1	2	2
c a b	2	3	1	2
s	13	14	15	

L'ordre collectif est :

$\boxed{a \succ b \succ c}$

Supprimons l'objet b. Les rangs sont alors les suivants :

	a	c	p_k
a c	1	2	3
c a	2	1	2
c a	2	1	2
s	11	10	

On en déduit l'ordre collectif $\boxed{c \succ a}$

Par conséquent supprimer b modifie le classement collectif entre a et b.

3.2 Méthode des majorités par paires : procédure CONDORCET (G.T. GUILBAUD [42])

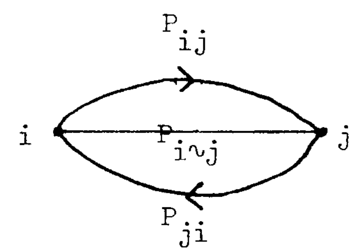
3.2.1 Définition de la méthode

Pour chaque paire d'objet (i,j) on calcule les 3 quantités :

$$P_{ij} = \sum_{k \in K_{ij}} p_k$$

$$P_{ji} = \sum_{k \in K_{ji}} p_k$$

$$P_{i\sim j} = \sum_{k \in K_{i\sim j}} p_k \quad (P_{i\sim j} = P_{j\sim i})$$



où K_{ij} est l'ensemble des classements k où i est préféré à j
K_{ji} est l'ensemble des classements où j est préféré à i
$K_{i\sim j}$ est l'ensemble des classements où i et j sont indifférents.

La méthode de CONDORCET (cf. [42]) consiste alors à construire une re lation de préférence collective antisymétrique P telle que :

$$i \; P \; j \iff P_{ij} > P_{ji}.$$

Si $P_{ij} \neq P_{ji}$ $\forall i, j, i \neq j$, alors la relation P est complète, antisymétrique. C'est ce qu'on appelle une relation de tournoi (cf. Annex § 3).

En général, cette relation de tournoi n'est pas transitive et n'est donc pas un ordre complet sur les objets ; le graphe associé à P contient des circuits. C'est le <u>paradoxe de CONDORCET</u> appelé également effet CONDORCET.

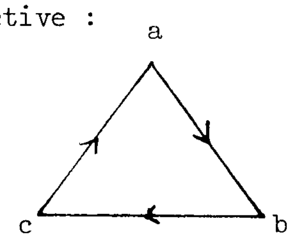
<u>Exemple</u> : En reprenant l'exemple précédent, on a : $P_{i\sim j} = 0$ $\forall i,j$.

P_{ij} est donné par le tableau carré suivant :

$$
(P_{ij}) = \begin{array}{c|ccc} & a & b & c \\ \hline a & 0 & 5 & 3 \\ b & 2 & 0 & 5 \\ c & 4 & 2 & 0 \end{array}
$$

On en déduit la relation de préférence collective :

$$
\begin{cases} a \; P \; b \\ b \; P \; c \\ c \; P \; a \end{cases}
$$
ou son graphe associé :



La procédure de CONDORCET respecte toutes les conditions d'ARROW, exceptée la seconde : P n'est pas en général un préordre complet.

3.2.2 <u>Méthode d'approximation d'une relation de tournoi (obtenue par la procédure CONDORCET) par un préordre complet</u>

Le but de ces méthodes est de chercher à approximer une relation de tournoi par un préordre complet. Dans ce cas, on ne violera plus la seconde condition d'ARROW : P sera un préordre complet mais, dans ce cas, la 4e condition ne sera plus respectée (indépendance vis-à-vis des choix extérieurs.

- On peut approximer une relation de tournoi par un préordre en calculant pour chaque objet un score. Le score d'un objet est alors la somme des demi-degrés extérieurs de chaque sommet dans le graphe associé à la relation.

<u>Exemple</u> : En reprenant le graphe associé à la relation de tournoi, on constate que chaque sommet a un demi-degré extérieur égal à 1. Par conséquent, l'approximation de cette relation de tournoi est le préordre $(a \sim b \sim c)$ (cf. MOON [66] et Annexe II, § 1).

- On peut rechercher un ordre complet à distance minimum de la relation de tournoi, la distance étant par exemple la différence symétrique : si P_{ij} et O_{ij} sont les matrices booléennes associées au tournoi P e à l'ordre complet O recherché, le problème est : trouver O minimisant $d(O, P) = \sum_{i,j} |P_{ij} - O_{ij}|$. Pour trouver cet ordre, on se reportera aux travaux de M. BARBUT [5], J.C. BERMOND [12], G. RIBEILL [75], P. SLATER [84], J.P.N. PHILLIPS [65], REMAGE et THOMPSON [73], HEUCHENNE [46]. On peut également adopter sur la matrice booléenne (P_{ij}) la méthode de maximisation des avis partiels en accord (cf. § I, 3.3). C'est cette dernière méthode qui est utilisée dans une des possibilités d'AGREPREF.

3.2.3 Restriction de l'ensemble des préordres admissibles pour rend la procédure CONDORCET transitive (cf. SEN [82])

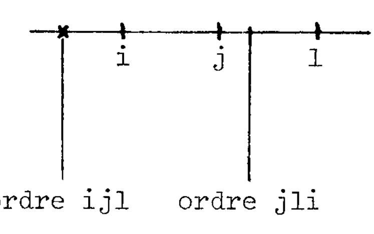
En ne respectant plus la condition C1, c'est-à-dire en restreignant l'ensemble des préordres possibles, on peut être assuré d'obtenir un ordr complet par la procédure de majorité.

- Condition de D. BLACK [17]

Pour tout triplet d'objet i, j, l, les préordres admissibles sont ceux compatibles avec un ordre de référence latent.

Autrement dit, pour tout triplet (i, j, l), l'un des trois objets n peut jamais être classé en queue.

Exemple :



Les ordres ilj et lij sont interdits.

Lorsque la condition de BLACK est satisfaite, la relation de majorité par paires est transitive.

- 1ère condition d'INADA [48]

Pour tout triplet (i, j, l) l'un des 3 objets ne peut jamais être classé en tête de ce triplet.

Exemple : jil et jli sont interdits.

Lorsque cette condition est satisfaite, la relation de majorité par paires est transitive.

- 2e condition d'INADA [48]

Tout triplet admissible doit être séparable en deux groupes A B tels que tous les objets de A soient classés avant B ou l'inverse.

Exemple : A = {i} B = {j, l}.

Les triplets admissibles seront :

$$\begin{cases} i\ j\ l \\ i\ l\ j \end{cases} \quad \begin{cases} j\ l\ i \\ l\ j\ i \end{cases} \quad \text{par contre} \quad \begin{cases} j\ i\ l \\ l\ i\ j \end{cases} \quad \text{sont interdits.}$$

Cette condition s'exprime également par le fait que, pour tout triplet, l'un des objets (i dans l'exemple ci-dessus), ne peut jamais occuper une position intermédiaire).

- Condition de WARD [88]

Parmi les triplets admissibles, il ne faut pas que l'on puisse former un carré latin.

Exemple : $\begin{cases} i\ j\ l \\ l\ i\ j \\ j\ l\ i \end{cases}$ sont trois triplets interdits.

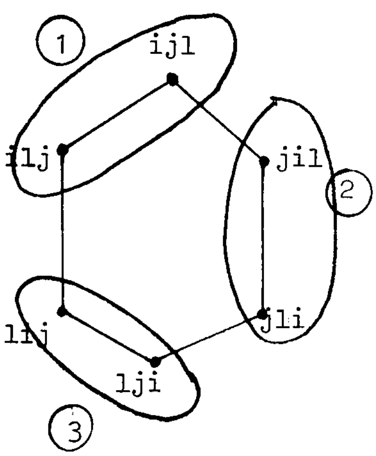
Cette condition ne s'applique qu'au cas des ordres (permutations). Autrement dit, dans l'ensemble des triplets admissibles, ne peuvent figure 3 permutations circulaires de ce triplet.

- <u>Condition de SEN</u> [82]

Pout tout triplet (i, j, l), il existe un des 3 objets qui ne soit jamais, soit en tête, soit au milieu, soit en queue.

Lorsque cette condition est remplie, la relation de majorité est tran sitive.

. un objet jamais en tête



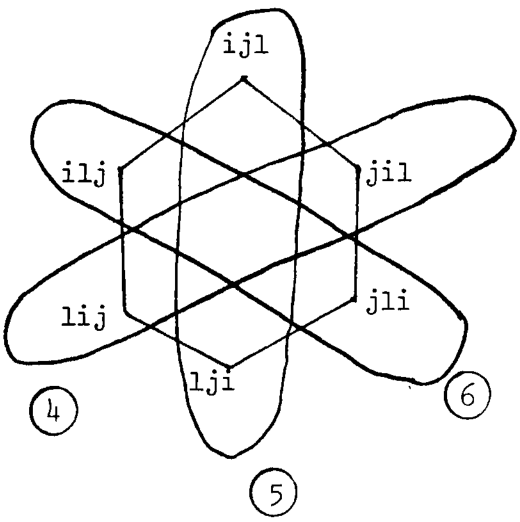
1 : i jamais en tête

2 : j jamais en tête

3 : l jamais en tête

C'est la 1ère condition d'INADA.

. un objet jamais au milieu



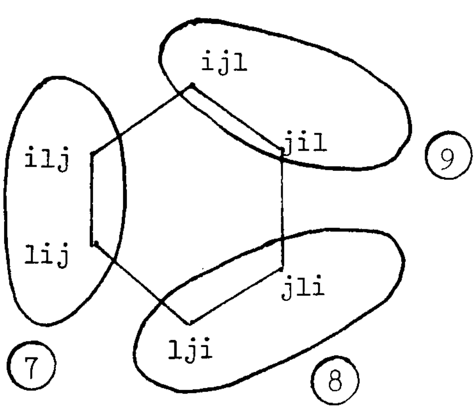
4 : i jamais au milieu

5 : j jamais au milieu

6 : l jamais au milieu

C'est la deuxième condition d'INADA.

. un objet jamais en queue



7 : j jamais en queue

8 : i jamais en queue

9 : l jamais en queue

C'est la condition de BLACK.

On constate que cette condition généralise aux préordres la condition de WARD, valable sur des ordres et qui interdit la présence de 3 permutations circulaires dans l'ensemble des triplets admissibles.

3.3 Recherche d'un ordre maximisant les avis partiels en accord (JACQUET-LAGREZE [49])

Pour chaque paire d'objet (i, j), on calcule comme précédemment les trois quantités P_{ij}, P_{ji}, $P_{i\sim j}$.

On calcule alors le tableau A_{ij} suivant :

$$A_{ij} = P_{ij} + \frac{1}{2} P_{i\sim j} \quad (A_{ji} = P_{ji} + \frac{1}{2} P_{i\sim j}).$$

Autrement dit, si on considère chaque classement k (préordre ou comparaison par paires) comme un graphe dont la matrice associée est $a^k_{ij} \in \{0, 1, 1/2\}$ comme défini dans le premier chapitre (§ III), on a :

$$A_{ij} = \sum_{k=1}^{p} c_k \, a^k_{ij}.$$

(A_{ij}) représente alors la matrice associée à un graphe valué que l'on peut qualifier d'opinion collective. C'est une relation de préférence valuée à rapprocher des relations floues au sens de ZADEH (cf. A. KAUFMAN [53]).

A_{ij} est l'intensité de l'avis $i \succ j$ et A_{ji} est l'intensité de l'avis contraire.

Dans le tableau A, les termes situés au-dessus de la diagonale représentent des avis partiels en accord avec l'ordre des lignes et des colonnes de A et les termes situés au-dessous de la diagonale représentent des avis en désaccord avec cet ordre.

La méthode consiste alors à trouver un ordre complet O sur les lignes et les colonnes tel que la somme des avis au-dessus de la diagonale soit maximum.

Trouver O tel que $\sum_{(i,j)\in O} A_{ij}$: maximum

$\sum_{(i,j)\in O}$ désignant la somme des termes au-dessus de la diagonale, l'ordre des lignes et des colonnes étant O.

On appelle cohésion la quantité :

$$C_0 = \frac{\sum\limits_{(i,j)\in 0} A_{ij} - A_{ji}}{\sum\limits_{i,j} A_{ij}}$$

L'ordre recherché 0 maximise également la cohésion C_0 qui est un indice alors compris entre 0 et 1 (cf. Chapitre I, § 1.2).

<u>Exemple</u> : En prenant les données du § I, 3.1, on a le tableau A_{ij} suivant :

$$(A_{ij}) = \begin{array}{c|ccc} & a & b & c \\ \hline a & 0 & 5 & 3 \\ b & 2 & 0 & 5 \\ c & 4 & 2 & 0 \end{array}$$

On vérifie facilement sur cet exemple que l'ordre recherché est ici unique : c'est l'ordre a b c.

<u>Cohésion</u> :

$$C = \frac{5+3+5 - (2+4+2)}{3 \times 7} = 0{,}24$$

Pour rechercher cet ordre, il existe un algorithme [49] et un programme (inclu dans les programmes ANAPREF et AGREPREF).

J.S. de CANI [26] puis [27] a également mis au point un algorithme qui pourrait être utilisé dans la recherche de cet ordre.

N. LECLERE et C. PEYROUX [38] ont repris cette méthode, ont proposé un nouvel algorithme et un programme correspondant qu'il serait intéressant de comparer à [49] (précision, rapidité).

Plus récemment, en 1971, V.J. BOWMAN et C.S. COLANTONI [20] ont redécouvert cette méthode ; ils proposent une solution se résolvant par une programmation mathématique en variables booléennes.

Remarques sur cette méthode

- Elle donne un (ou plusieurs) ordre(s) complet(s) mais jamais de préordre (ce qui est une faiblesse).

- Lorsque la procédure CONDORCET conduit à un ordre, cet ordre est également celui donné par la méthode ci-dessus. En effet, le critère de CONDORCET est en fait : trouver une relation antisymétrique P maximisant le nombre des avis partiels en accord. Ce critère conduit à prendre la règle de majorité mais peut également s'interpréter comme :

trouver P maximisant $\sum_{(i,j)\in P} A_{ij}$.

La méthode ci-dessus est analogue (trouver 0 maximisant $\sum_{(i,j)\in 0} A_{ij}$). Cependant, elle introduit des contraintes sur la relation recherchée.

- Cette méthode respecte toutes les conditions d'ARROW, exceptée C4 (indépendance vis-à-vis des choix extérieurs).

En effet, dans l'exemple ci-dessus, on constate que, si on supprime l'objet b, l'ordre $a \succ b \succ c$ devient l'ordre $c \succ a$.

- Lorsque la relation P est intransitive, on a vu qu'il était possible de l'approximer par un ordre complet (cf. § I, 3.2.2) en minimisant par exemple le nombre d'arcs en désaccord (i de SLATER). On peut pour cela utiliser sur la matrice booléenne associée à P la méthode exposée ci-dessus.

En effet, trouver 0 tel que $\sum_{(i,j)\in 0} P_{ij}$ soit maximum revient à trouver un ordre 0 tel que le nombre d'arcs en accord avec P soit maximum.

Il est à remarquer que les deux ordres trouvés par les deux méthodes ne sont pas nécessairement les mêmes.

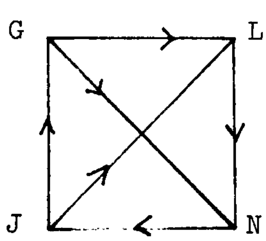
<u>Exemple</u> (tiré de données de préférences réelles) : Soit le tableau A_{ij} suivant :

$A_{ij} =$

	G	N	J	L
G	0	14.5	10	13
N	9.5	0	16.5	11.5
J	14	7.5	0	16
L	11	12.5	8	0

On peut montrer que l'ordre GNJL est l'ordre maximisant les avis partiels en accord.

Si on cherche à établir la relation de tournoi de la règle des majorités par paires, on obtient le graphe suivant :



ou

	J	G	L	N
J	0	1	1	0
G	0	0	1	1
L	0	0	0	1
N	1	0	0	0

Le i de SLATER vaut dans ce cas 1 (on sait par ailleurs que sa valeur maximale est 1 pour 4 objets - cf. JACQUET-LAGREZE [50] et BERMOND [12]) alors que la méthode de maximisation des arcs en accord revient à contredire deux arcs de majorité par paires ($J \rightarrow G$ et $L \rightarrow N$).

3.4 <u>Recherche d'un ordre médian</u>

Soit $P_1, P_2, \ldots, P_k, \ldots, P_p$ les p préordres. La recherche d'un ordre O médian consiste à trouver O tel que :

$$p_1\ d(P_1,\ 0) + p_2\ d(P_2,\ 0) + \ldots + p_k\ d(P_k,\ 0) + \ldots + p_p\ d(P_p,\ 0)$$

soit minimum sachant que p_k est le poids du préordre k et d la différence symétrique telle qu'elle a été définie au § I, 3.2.2.

On peut montrer que cette méthode est identique à celle du paragraphe précédent, autrement dit :

$$(1)\quad \text{trouver } 0,\ \sum_{(i,j)\in 0} A_{ij}\ \text{max.} \Longleftrightarrow \text{trouver } 0,\ \sum_{k=1}^{p} p_k\ d(P_k,\ 0)\ \text{min}$$

On trouvera la démonstration à la fin de ce paragraphe.

B. MONJARDET [65] et J.F. FELDMANN [35] ont récemment montré certains résultats concernant les ordres médians.

<u>Exemple</u> : Reprenons les données du § I,3.1 :

	a	b	c	P_k
P_1 : a b c	1	2	3	3
P_2 : b c a	3	1	2	2
P_3 : c a b	2	3	1	2

Pour trouver 0 parmi les 3! ordres possibles, calculons les distances entre les préordres P_1, P_2, P_3 et les 6 ordres possibles à l'aide du tableau ci-dessous :

P_k	p_k \ 0	a b c	b a c	b c a	c b a	c a b	a c
a b c	3	0	2	4	6	4	2
b c a	2	4	2	0	2	4	6
c a b	2	4	6	4	2	0	2
$\Sigma\ p_k\ d(0,\ P_k)$		16	22	20	26	20	20

L'ordre recherché est l'ordre $a \succ b \succ c$.

<u>Démonstration de (1)</u>

Dans la sommation $\sum_{i,j}$, prenons comme ordre sur les indices i et j l'ordre 0. Dans ce cas :

$$0_{ij} = 1 \quad \text{si} \quad i < j, 0$$
$$0_{ji} = 0 \quad \text{si} \quad i > j, 0$$

$$\sum_{i,j} |a_{ij}^k - 0_{ij}| = \sum_{i<j,0} |a_{ij}^k - 1| + \sum_{j>i,0} |a_{ij}^k - 0|$$

Comme $a_{ij}^k \in \{0, 1/2, 1\}$, on a :

$$\sum_{i,j} |a_{ij}^k - 0_{ij}| = \sum_{i<j,0} (1 - a_{ij}^k) + \sum_{j>i,0} a_{ij}^k$$
$$= \sum_{i<j,0} (1 - a_{ij}^k + a_{ji}^k)$$
$$= \frac{n(n-1)}{2} - \sum_{i<j,0} (a_{ij}^k - a_{ji}^k)$$

d'où :

$$\sum_{k=1}^{p} p_k\, d(P_k, 0) = \sum_{k=1}^{p} p_k \frac{n(n-1)}{2} - \sum_{k=1}^{p} p_k \sum_{i<j,0} (a_{ij}^k - a_{ji}^k)$$
$$= \frac{n(n-1)}{2} (\Sigma\, p_k) - \sum_{i<j,0} \left[\sum_{k=1}^{p} p_k\, a_{ij}^k - \sum_{k=1}^{p} p_k\, a_{ji}^k \right]$$

$$\boxed{\sum_{k=1}^{p} p_k\, d(P_k, 0) = \frac{n(n-1)}{2} \Sigma\, p_k - \sum_{i<j,0} (A_{ij} - A_{ji})}$$

Minimiser le premier membre revient donc à maximiser la somme $\sum_{i<j,0} (A_{ij} - A_{ji})$ c.q.f.d.

<u>Exemple</u> : Avec les données ci-dessus, on a calculé le premier membre ; il vaut 16 pour 0 = a b c. Le second vaut :

$$\frac{3.2}{2} \times 7 - [5 + 3 + 5 - (2 + 4 + 2)]$$
$$= 21 - (5) = 16.$$

3.5 Méthode lexicographique

Dans cette méthode, toutes les conditions d'ARROW sont respectées, exceptée la 6e (condition de non dictature d'un préordre).

Les préordres sont ordonnés de façon hiérarchique. Le 1er dans cet ordre impose un premier préordre sur les objets.

Le second départage éventuellement les ex-aequo lorsqu'ils existent dans le premier préordre. On obtient ainsi un nouveau préordre.

Cette méthode revient à prendre un jeu de poids du type $\{\frac{1}{2}, \frac{1}{4}, \ldots, \frac{1}{2}k, \ldots, \frac{1}{2}p\}$ pour les p préordres. On a alors $p_1 > \sum_{i=2}^{p} p_i, \ldots, p_k > \sum_{i=k+1}^{n} p_i, \ldots)$.

Lorsqu'on applique l'une quelconque des méthodes exposées ci-dessus avec ce jeu de poids, on obtient la méthode lexicographique.

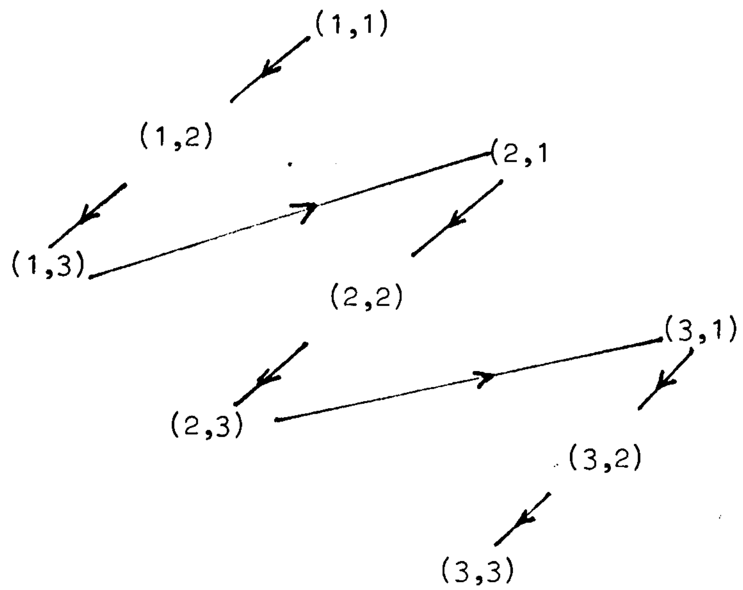
Une variante de la méthode lexicographique repose sur une notion de hiérarchisation partielle entre les critères. On en trouvera un développement dans le rapport du C.R.U. [24].

Donnons seulement ici un exemple. Soit deux échelles permettant de donner des préordres à 3 classes, les échelons étant (1, 2, 3).

Un objet i est caractérisé par un couple de valeur (x_i^1, x_i^2) si le premier classement 1 est supérieur hiérarchiquement au second , l'ordre lexicographique des évaluations possibles est le suivant :

.

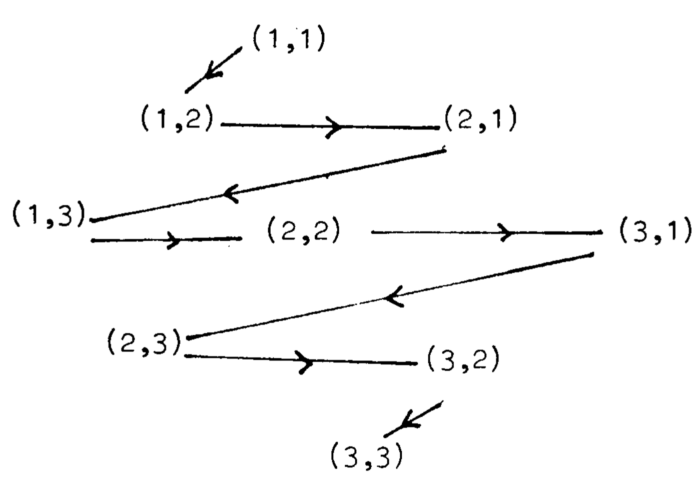
.



L'ordre des évaluations possibles est alors :

(1,1) > (1,2) > (1,3) > (2,1) > (2,2) > (2,3) > (3,1) > (3,2) > (3,3).

Un ordre correspondant à une <u>hiérarchisation partielle</u> serait par exemple :



3.6 Introduction aux méthodes à seuil

Avec la méthode ELECTRE que nous étudierons plus loin apparaît une no tion nouvelle : la relation de surclassement (P. BERTIER et J. de MONTGOLF [15], B. ROY [71]). Celle-ci est née dans le cadre de modèles d'aide à la décision en présence de critères multiples. Une relation de surclassement est le modèle mathématique d'une relation réelle mais difficile à appréhen der : la relation de préférence d'un décideur qui se trouve confronté à un problème de choix.

"On dit que <u>i surclasse j</u> si on peut prendre le risque de décider que i est préféré à j pour le décideur". Construire une relation de surclassement apparaît donc comme une procédure d'agrégation. Dans le cas général, cette relation n'est pas complète ; l'incomparabilité demeure une caractéristique essentielle des démarches d'agrégation s'appuyant sur le concept de relation de surclassement. Il est bien naturel que, dans cette optique, ce soit la seconde condition d'ARROW qui ne soit pas vérifiée (la préférence collective ou multicritère n'est pas un préordre complet sur le objets, c'est une relation de surclassement dont l'incomparabilité demeure capitale pour l'interprétation des résultats).

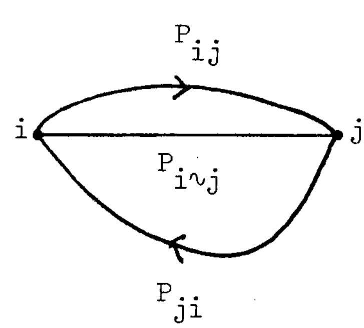
Il est possible de construire des relations de surclassement en ne tenant compte que du caractère ordinal des données de préférence sur les objets.

Dans une méthode à seuil, on construit donc, comme dans ELECTRE et comme dans la méthode des comparaisons par paires, les trois quantités (cf. § 3.2.1) :

$$P_{ij} = \sum_{k \in K_{ij}} p_k$$

$$P_{ji} = \sum_{k \in K_{ji}} p_k$$

$$P_{i\sim j} = \sum_{k \in K_{i\sim j}} p_k$$



Suivant les valeurs de ces trois nouveaux critères, on va chercher à se prononcer sur la relation de surclassement entre i et j : S = (P, I). Les cas possibles seront :

$i \longrightarrow j$	i P j	i préféré à j	i et j comparables
$i \longleftarrow j$	j P i	j préféré à i	
$i \rightleftarrows j$	i I j	i et j indifférents	
$i \bullet \quad \bullet j$		i et j non comparables	

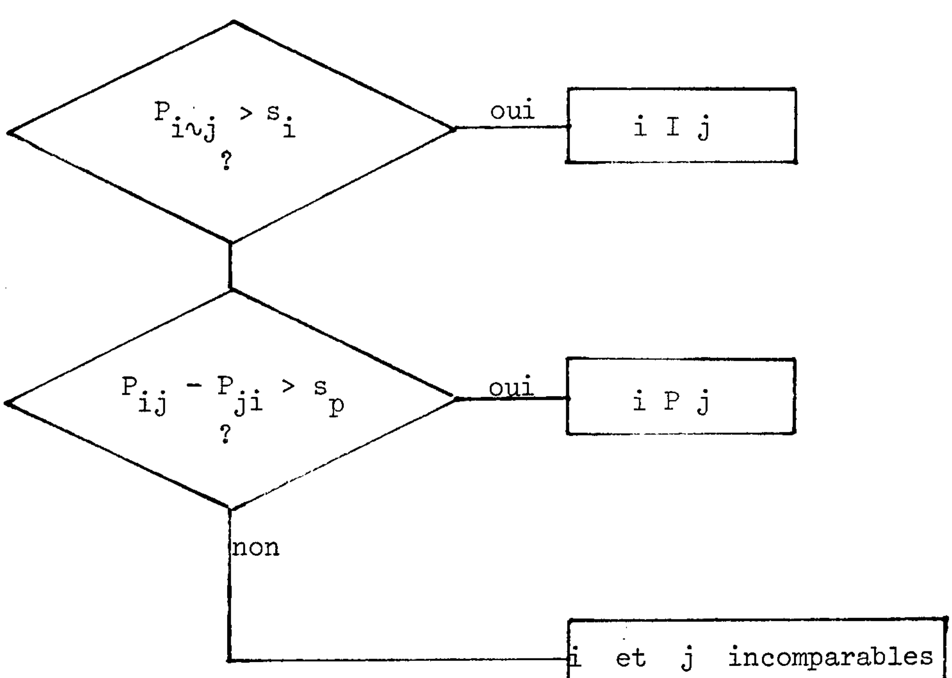
Pour construire une relation de surclassement (P, I), on peut retenir les idées suivantes :

- si $P_{i \sim j}$ est grand, on peut décider en faveur de l'indifférence ;
- dans le cas opposé, on peut tester l'incomparabilité. Supposons que $P_{ij} + P_{ji} + P_{i \sim j} = \Sigma\, p_k = 1$.

Si P_{ij} est voisin de P_{ji}, on peut vouloir se prononcer sur l'incomparabilité.

Par contre, si P_{ij} est suffisamment grand devant P_{ji}, on peut se prononcer sans prendre trop de risques sur la relation de préférence stricte.

Si s_i et s_p sont deux seuils d'indifférence et de préférence stricte, on pourra adopter pour chaque couple (i, j) la séquence de tests suivante :



Pour éviter des contradictions sur les conclusions, on prendra $s_i + s_p \geq 1$.

Remarques - cas particuliers :

- Si on prend $s_i = 1$ et $s_p = 0$, on retrouve la méthode des majorités par paires : i P j si $P_{ij} > P_{ji}$.

- Si on prend $s_i = 1$ et $s_p = 1$, on retrouve la relation de dominance : i P j si $P_{ij} = 1$ ($= \Sigma p_k$). Les procédures à seuil apparaissent comme un intermédiaire heureux entre deux méthodes extrêmes :

. une méthode trop prudente pour être opérationnelle : la relation de dominance ;

. une méthode opérationnelle mais risquée : la relation de majorité par paires.

Pour de tels développements, on pourra se reporter à [52], notamment dans l'étude des conditions dans lesquelles on obtient des circuits dans la relation de surclassement.

3.7 Conclusion sur les méthodes purement ordinales

Les méthodes ordinales sont variées et il est intéressant de les confronter au théorème d'ARROW.

Il faut cependant remarquer que ces méthodes peuvent très vite atteindre leurs limites, notamment dans les problèmes d'agrégation multicritères.

En effet, dans ces problèmes, il est rare que des critères soient tous purement ordinaux. Dès qu'un critère permet une évaluation cardinale (coût d'un projet, surface d'un terrain, etc.), il est légitime de vouloir tenir compte de distance sur des échelles, autrement dit de tenir compte des différences $s_i^k - s_j^k$ et non plus seulement du signe de ces différences. Nous allons voir comment, en restant dans l'esprit de ces méthodes ordinales, on peut néanmoins, dans certains cas, tenir compte de distances sur les échelles.

4. Exemples de méthodes non purement ordinales

On a vu que l'on peut appliquer les méthodes précédentes non seulement avec des données ordinales mais également avec des données quantitatives. Plus généralement, lorsque certains classements k sont définis par l'intermédiaire d'une évaluation s_i^k sur une échelle E_k, on peut vouloir tenir compte des différences $s_i^k - s_j^k$ sur une même échelle. Nous allons présenter rapidement deux méthodes qui utilisent ces différences.

4.1 Introduction à la méthode ELECTRE [16], [81]

ELECTRE est une méthode qui, dans une première phase, construit une relation de surclassement S sur l'ensemble des objets puis, dans une seconde phase, construit, ou bien deux préordres complets à partir de la relation de surclassement (ELECTRE II), ou bien un noyau, c'est-à-dire un sous-ensemble d'objets parmi lesquels doivent se trouver le ou les meilleurs (ELECTRE I).

Pour construire la relation de surclassement, on introduit deux tests : un test de concordance et un test de discordance.

Le test de concordance correspond à une méthode à seuils. Il est satisfait si $\dfrac{P_{ij} + P_{i\sim j}}{P_{i\sim j} + P_{ij} + P_{ji}} \geq c$ où c est un seuil de concordance et si $\dfrac{P_{ij}}{P_{ji}} > 1$. Dans ces conditions, i S j. Cependant, on peut refuser ce surclassement (ordinal) i S j si il y a discordance, c'est-à-dire si il existe des classements sur lesquels j domine i avec un trop grand écart : autrement dit, on introduit sur chaque classement (critère) des ensembles de discordance. On pourra par exemple refuser le surclassement i S j si $s_j^k - s_i^k \geq D_k$ où D_k est une différence importante sur l'échelle E_k.

Dans la méthode ELECTRE, on introduit donc la notion de distance sur une échelle au niveau du test de non-discordance.

4.2 Introduction aux méthodes d'agrégation de quasi-ordres

Nous renvoyons à [52] pour ce problème. Présentons-le cependant en quelques lignes.

La notion de quasi-ordre a été introduite par LUCE [59] pour pouvoir traduire une relation de préférence portant sur des objets pouvant être tr voisins (cf. Annexe I, § 3.3). La principale propriété d'une relation de quasi-ordre est de se décomposer en une double relation (I, P) où I est une relation d'indifférence non-transitive et P une relation de pré- férence stricte transitive.

Une façon simple d'obtenir des relations de quasi-ordres est d'introduire une notion de seuil d'indifférence sur une échelle (quantitative). La relation k prend ainsi une certaine "distance" vis-à-vis de la relation de préordre induite par les valeurs s_i^k.

Soit s_i^k l'évaluation de i, d_k le seuil d'indifférence. On aura :

$$
\begin{aligned}
|s_i^k - s_j^k| \le d_k &\iff i\ I_k\ j \\
s_i^k - s_j^k > d_k &\iff i\ P_k\ j \\
s_i^k - s_j^k < - d_k &\iff j\ P_k\ i.
\end{aligned}
$$

On peut alors utiliser une méthode d'agrégation de quasi-ordres analogue à celle décrite au § I 3.6 en construisant les indicateurs P_{ij}, P_{ji}, $P_{i\sim j}$ de la façon suivante :

$$
\begin{aligned}
\text{Soit}\quad K_{i\sim j} &= \{k,\ |s_i^k - s_j^k| \le d_k\} \\
K_{ij} &= \{k,\ s_i^k - s_j^k > d_k\} \\
K_{ji} &= \{k,\ s_i^k - s_j^k < - d_k\}
\end{aligned}
$$

On construit alors les trois valeurs :

$$P_{ij} = \sum_{k \in K_{ij}} p_k , \; P_{ji} = \sum_{k \in K_{ji}} p_k , \; P_{i \sim j} = \sum_{k \in K_{i \sim j}} p_k$$

et on peut ensuite utiliser la série de tests proposée au § I ,3.6. Cette méthode a été programmée dans le programme AGREPREF [1] (cf. troisième chapitre).

Notons que cette méthode pourrait également utiliser la notion de discordance telle qu'elle a été définie pour la méthode ELECTRE.

(1) AGREgation des PREFérences.

II - Méthodes de structuration portant sur les classements

On a vu dans le premier chapitre (§ I.1) différentes définitions de distances entre classements.

Dans les modèles géométriques de description des données ordinales, on a vu des méthodes de visualisation des nuages de points que l'on pouvait construire pour représenter les classements.

On peut cependant essayer d'aller plus loin en cherchant une structure portant sur les classements, la structure recherchée étant alors une relation d'équivalence entre classements. Autrement dit, on peut cherche à fabriquer des typologies de classements. Il n'existe pas de méthodes, donc de programmes, spécialement conçus pour réaliser ces typologies mais nous présenterons quelques voies possibles, tant sur le plan des réalisations pratiques que sur celui de nouvelles recherches.

1. Typologie de classements à partir de données de rangs ou d'évaluation sur des échelles

Dans ce cas, les données sont les rangs r_i^k ou les évaluations s_i^k. On peut calculer des distances euclidiennes entre les rangs (distance de SPEARMAN) ou entre les évaluations sur les échelles.

On appliquera alors une méthode de classification métrique (cf. P. B et J.M. BOUROCHE [13]).

Pour chaque groupe donné par la typologie, on pourra rechercher une relation d'ordre qui soit un classement moyen par l'une des méthodes du § II ci-dessus. Les individus associés (W de KENDALL, coefficient de cohésion C_0) seront alors des indices permettant de mesurer la bonne homogénéité des groupes de la typologie ainsi constituée.

2. Typologie de relations de classements quelconques

- Dans le cas où les données de classement sont les matrices $\underline{a}^k = (a^k_{ij})$ (comparaisons par paires, préférences floues, etc.), on peut utiliser la distance : $d(k, l) = \sum_{i,j} |a^k_{ij} - a^l_{ij}|$, le centre de gravité d'un ensemble de classements étant (comme dans la méthode ANAPREF) la relation valuée : $\underline{a}^G = (a^G_{ij})$ avec $a^G_{ij} = \frac{1}{p} \sum_{k=1}^{p} a^k_{ij}$.

- Dans la méthode ANAPREF, on a proposé deux algorithmes permettant de définir une typologie de classements en 1, 2 ou 3 groupes. On pourrait très bien cependant utiliser une méthode de classification métrique avec la distance d ci-dessus. On peut également penser à utiliser la méthode de typologie proposée par C. MAYER [62].

III - Modèles de structuration particuliers

Les méthodes de structuration vues dans les deux paragraphes précédents portent, soit sur les objets (agrégation, recherche d'une relation d'ordre...), soit sur les classements (recherche d'une typologie de classements).

Présentons quelques modèles permettant de déceler une structure particulière reliant l'ensemble des objets et l'ensemble des classements lorsque ces derniers sont des ordres totaux. M. BARBUT et L. FREY [6] présentent plusieurs de ces modèles.

1. Le modèle de COOMBS et ordres blackiens

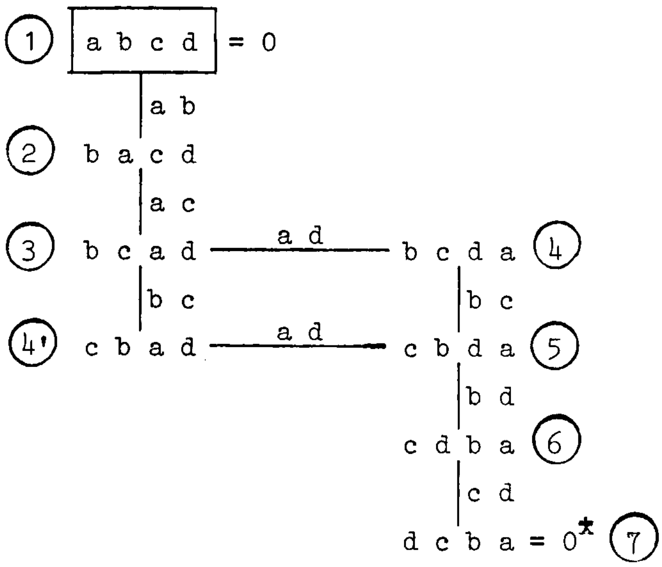
Nous avons présenté le modèle de COOMBS (premier chapitre, §II, 1.1) et on a vu que $n(n-1)/2+1$ ordres sont compatibles avec une configuration latente unidimensionnelle.

La définition d'un ordre blackien est plus générale :

Compatibilité entre ordres

Un ordre k est compatible avec un ordre de référence 0 si, pour tout triplet (i, j, l) d'objets non nécessairement consécutifs mais classés dans cet ordre dans 0, alors $i \succ j \Longrightarrow j \succ l$. Un ordre blackien est un ordre compatible avec un ordre de référence donné 0.

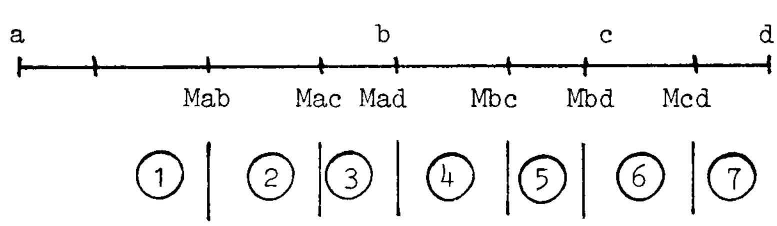
Exemple : Soit $0 = a\ b\ c\ d$ un ordre sur 4 objets. Les ordres blackiens sont :



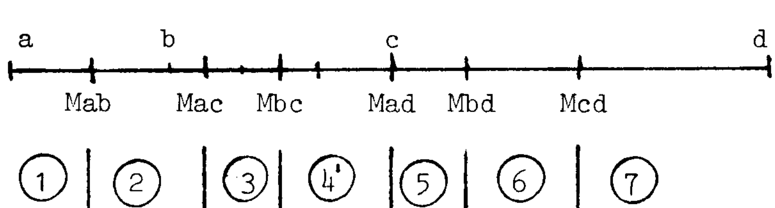
L'ensemble des ordres blackiens représente ici deux "chaînes" sur le permutoèdre reliant les deux permutations inverses 0 et 0^*.

Une configuration de COOMBS unidimensionnelle ne correspond elle qu'à une chaîne.

Autrement dit, dans notre exemple, relativement au même ordre de référence 0 = a b c d, on peut avoir <u>deux</u> configurations unidimensionnelles qui sont :



Les 7 ordres compatibles sont ceux de la chaîne supérieure passant par 4 .



Les 7 ordres sont ceux de la chaîne inférieure passant par 4' .

Les deux configurations diffèrent par les positions relatives des mi-lieux Mad et Mbc des segments ad et bc.

Le nombre d'ordres blackiens est 2^{n-1} (cf. [6]). Le nombre d'ordr de COOMBS, c'est-à-dire le nombre d'ordres blackiens une fois la configura-tion des proximités des objets donnée, est : $n(n - 1)/2 + 1$.

nombre d'objets n	nombre d'ordres de COOMBS $n(n - 1)/2 + 1$	nombre d'ordres blackiens 2^{n-1}	nombre d'ordre $n!$
2	2	2	2
3	4	4	6
4	7	8	24
5	11	16	120

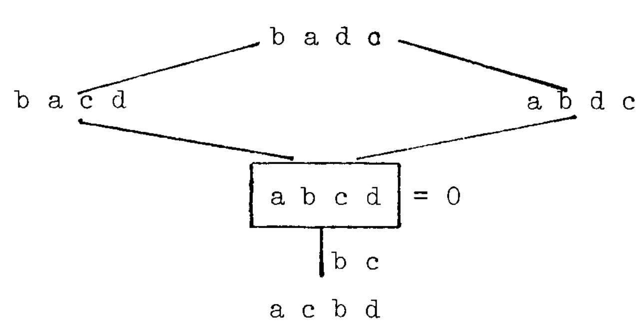
L'intérêt des ordres blackiens est, on l'a vu, de permettre d'obtenir toujours un ordre par la procédure d'agrégation des majorités par paires.

2. Faisceaux d'indifférence de centre 0

Soit un ordre 0, r_i et r_j les rangs de deux objets pour cet ordre A cet ordre, on associe l'ensemble F(0) des ordres k tel que si $r_i - r_j > 1$ pour 0, alors $r_i^k - r_j^k > 0$.

Autrement dit, on appelle faisceau d'indifférence de centre 0 l'en-semble des ordres qui ne diffèrent de 0 que sur les seuls couples d'ob-jets consécutifs de 0.

Exemple : Soit 0 = a b c d ; le faisceau de centre 0 est :



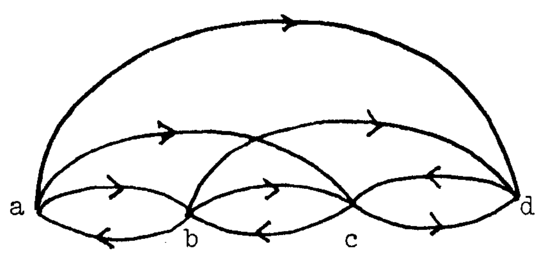
Faisceau d'indifférence de centre 0

Le terme d'indifférence est utilisé pour représenter deux objets adjacents (i, j) trop proches pour être discernés. Lorsque des sujets classent les objets, ils peuvent alors donner nimporte quel ordre du faisceau d'indifférence de centre O.

A tout faisceau d'indifférence de centre O, on peut associer un quasi-ordre construit à partir de O.

Soit (I, P) le quasi-ordre défini par $i \; P \; j \Longleftrightarrow r_i - r_j > 1$. L'intersection de 2 des ordres d'un faisceau d'indifférence de centre O est un ordre partiel P compatible avec ce quasi-ordre (I, P).

Le quasi-ordre associé au faisceau ci-dessus est :



Le nombre d'ordres totaux d'un faisceau d'indifférence, c'est-à-dire le nombre d'ordres totaux compatibles avec un quasi-ordre associé, est [6] :

$$|F(n)| = \sum_{k=0}^{k \leq n-k} \binom{n-k}{k}.$$

On a la relation $|F_n| = |F_{n-2}| + |F_{n-1}|$, ce qui donne :

n	$\|F(n)\|$	n!
2	2	2
3	3	6
4	5	24
5	8	120

Si on applique une procédure d'agrégation à seuil sur un ensemble d'ordres appartenant à un faisceau d'indifférence de centre 0, on doit trouver un quasi-ordre particulier (celui construit comme ci-dessus).

3. Fuseaux d'ordres

- Relation d'intermédiaire sur le permutoèdre

Un ordre l est intermédiaire de deux ordres k et m (notée $k/l/m$) si on a la relation :

$$d_K(k, l) + d_K(l, m) = d_K(k, m)$$

où d_K est la distance de KENDALL sur le permutoèdre.

Cette relation d'intermédiaire permet de définir une relation d'ordre partiel sur le permutoèdre dès que l'on choisit un ordre (permutation) k_o donné, $k \geq l \Longleftrightarrow k_o/k/l$.

Cette relation d'ordre permet de munir le permutoèdre d'une structure de treillis (cf. Annexe I, § 2.3.2) dont le supremum est la permutation k_o et l'infimum la permutation inverse.

On avait présenté le permutoèdre comme figure géométrique. On voit qu'il est possible de lui donner une définition comme ensemble des permutations munies d'une structure de treillis. On l'appelle parfois treillis des ordres totaux. Pour un développement de ces questions, on consultera l'ouvrage de M. BARBUT et B. MONJARDET [7].

- Fuseau de permutation

C'est l'ensemble des ordres intermédiaires de deux ordres k_1, k_2. On notera $F(k_1, k_2)$. k_1 et k_2 sont les pôles du fuseau.

Le fuseau de deux permutations inverses est alors le permutoèdre entier.

Remarque : Dans ANAPREF, le fuseau de permutation des pôles G1 et G2 se projette sur l'arc de cercle G1, G2.

- Fuseau bipolaire d'insertion

Soit A, B deux ordres totaux définis sur deux sous-ensembles disjoints d'objets de n_1 et n_2 éléments chacuns.

Une permutation (ordre) d'insertion est une permutation définie sur les $n_1 + n_2$ objets et respectant l'ordre A et l'ordre B.

Exemple : A : a b c
B : i j m.

a i b c j m est une permutation d'insertion.

Soit k_1 = A B (l'ordre a b c i j m)
k_2 = B A (l'ordre i j m a b c).

On appelle fuseau d'insertion des séquences A, B le fuseau des permutations k_1 et k_2, c'est-à-dire l'ensemble des intermédiaires de k_1 et k_2.

Le nombre de permutations d'un fuseau d'insertion est $\dfrac{(n_1 + n_2)!}{n_1!\, n_2!}$.
Dans l'exemple précédent ($n_1 = 3$, $n_2 = 3$), on a 20 permutations d'insertion.

- <u>Interprétation de ce modèle</u>

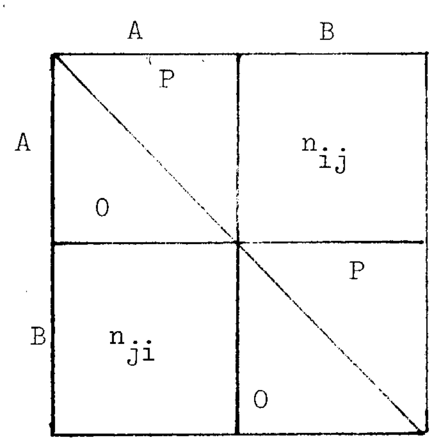
Supposons que l'on demande au cours d'une enquête de comparer en terme de hiérarchie un ensemble d'objets dont certains sous-ensembles sont relativement peu comparables.

Sur un sous-ensemble d'objets comparables, il peut y avoir un consensus sur une hiérarchie admise. Par contre, des interclassements peuvent révéler des échelles de valeurs ou de préférences différentes.

<u>Exemple</u> : Supposons que l'on demande des classements sur les 4 métiers suivants : architecte, maçon, ingénieur, ouvrier et que l'on demande à des personnes un classement de ces métiers en terme de hiérarchie de prestige. Il est fort possible que l'ensemble des classements soit le fuseau d'insertion des deux séquences :

A : architecte $\succ$ maçon
B : ingénieur $\succ$ ouvrier.

L'agrégation d'un ensemble de p classements appartenant à un fuseau d'insertion de séquence A, B conduit à un tableau de comparaisons par paires n_{ij} de la forme :



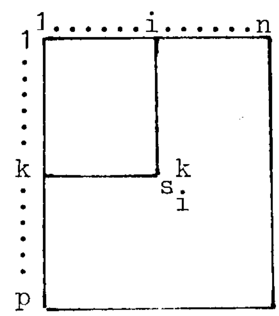
Le problème pourrait être généralisé à plusieurs séquences d'insertion A, B, C, ... Dans chaque sous-ensemble A, B, les classements seraient identiques sur les objets de ces sous-ensembles.

Cette démarche soulève un problème intéressant concernant le recueil de données de comparaisons portant sur un ensemble d'objets non tous comparables. On pourrait penser à recueillir sur un même ensemble d'objets des comparaisons soit de proximité, soit de classements sans imposer l'un des types de comparaisons aux personnes interrogées mais nous reviendrons sur cette question dans la conclusion.

IV - L'analyse hiérarchique

Dans l'introduction (§ 3.2), nous avons présenté des données ordinal d'un caractère particulier : les échelles dichotomiques (1, 0).

Il faut remarquer que, bien souvent au départ, ces données ne sont pas des données de classement. Elles expriment le plus souvent une propriété i qu'un sujet k peut posséder ou ne pas posséder. On a le tableau :



$s_i^k = 1$ si le sujet k possède la propriété i

$s_i^k = 0$ sinon.

Posséder la propriété peut signifier :

- succès à l'examen ou au test i ;
- réponse favorable (oui) à la question i.

On trouvera dans le troisième chapitre un exemple de problème traité par l'analyse hiérarchique (BMD05S) en gestion du personnel.

Cependant, on peut attribuer une valeur sur des données dichotomique et considérer qu'elles déterminent sur l'ensemble des p sujets une relation d'ordre : un préordre en deux classes (classe des sujets qui possède la propriété i, ceux qui ne la possèdent pas).

L'objet de l'analyse hiérarchique est de classer des sujets suivant <u>une dimension</u> que l'on appréhende au-travers d'un ensemble de n questio dichotomiques.

1. Relation d'ordre définie sur les sujets

Soit deux sujets k et l. On définit une relation d'ordre partiel entre ces sujets :

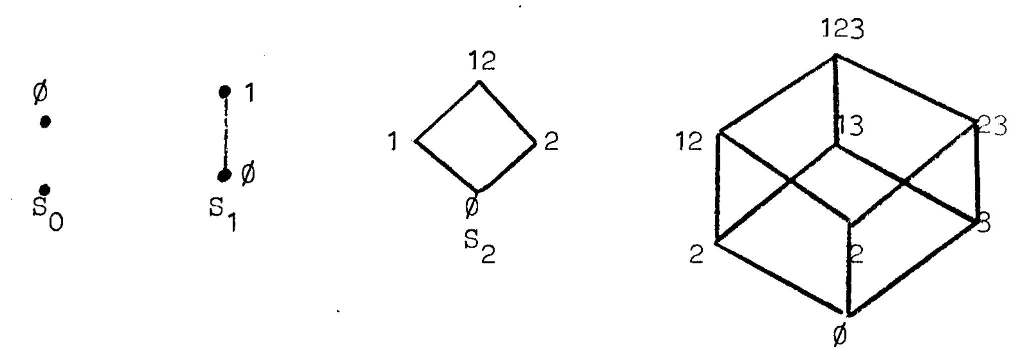
$$k \succcurlyeq l \Longleftrightarrow \forall i,\ s_i^l = 1 \Longrightarrow s_i^k = 1.$$

Autrement dit, k est avant l si k possède au moins toutes les propriétés que possède l. On s'aperçoit que la relation ainsi définie est la relation de dominance si on considère chaque question i comme un classement des sujets :

$$k \succcurlyeq l \Longleftrightarrow \forall i,\ s_i^k \geq s_i^l.$$

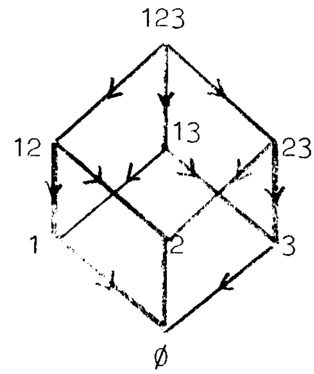
Interprétation algébrique

Soit le simplexe S_n (cf. Annexe II § 2.3.2) d'ordre n :



Chaque patron, c'est-à-dire chaque ensemble de réponses données par un sujet, est un sommet du simplexe. Ce simplexe possède une structure de treillis dont le supremum est le patron (1 1 ... 1) et l'infimum le patron (0 0 0 ... 0).

Ce treillis (cf. Annexe I, § 2.3.2) est un ordre partiel sur les 2^n patrons possibles.



Exemple : 1 2 3 > 1. 12 et 23 sont non comparables.

On trouvera un développement de ces questions dans DEGENNE [28].

2. Le modèle de GUTTMAN [45]

On appelle échelle de GUTTMAN une suite de patrons totalement ordonnés

Une échelle de GUTTMAN est donc une chaîne dans le simplexe.

Exemple : L'ensemble des patrons 123, 23, 3, ∅ constituent une échelle de GUTTMAN.

	1	2	3
123	1	1	1
23	0	1	1
3	0	0	1
∅	0	0	0

Considérons alors un ensemble de réponses s_i^k (i = 1, n et k = 1, p
L'ensemble des patrons peut-il être totalement ordonné par la relation d'ordre définie ci-dessus ?

On constate que si les patrons peuvent être totalement ordonnés les questions i sont également totalement ordonnés.

Ainsi le modèle de GUTTMAN se présente comme une recherche simultanée de deux ordres complets :

- un ordre sur les patrons (les sujets) ;
- un ordre sur les questions.

Dans un problème où $s_i^k = 1$, si un élève k réussit l'exercice i et où les données sont ajustables à un modèle de GUTTMAN, alors l'ordre des questions représente un ordre unidimensionnel de difficulté des différents exercices et l'ordre des sujets représente un classement des élèves du meilleur au moins bon suivant cette dimension.

3. <u>L'ajustement d'un ensemble de réponses au modèle de GUTTMAN</u>

Ce problème a fait l'objet d'une littérature abondante [45], [28], [33], [60], [89].

Un certain nombre d'indices permet de mesurer l'ajustement au modèle Guttmanien. Citons le coefficient de reproductibilité de GUTTMAN :

$$\boxed{r = 1 - \frac{\sigma}{np}}$$

où σ est le nombre d'erreurs, une erreur au sens de GUTTMAN étant un 0 ou un 1 mal situé dans le tableau.

<u>Exemple</u> (MATALON [59], p. 38) :

Questions

Sujets		c	a	b	e	d
	7	1	1	1	1	1
	4	0	1	1	1	1
	8	1	0	1	1	1
	5	0	0	1	1	1
	2	1	0	0	1	1
	9	0	1	0	1	1
	10	0	0	0	1	1
	1	0	0	1	0	1
	6	0	1	0	0	0
	3	0	0	0	0	0

Pour déterminer les erreurs, GUTTMAN fixe des points de coupure question par question. Une erreur est alors un 0 se trouvant dans une zone de 1 ou un 1 se trouvant dans une zone de 0.

Ici $\sigma = 5$ et $r = 1 - \frac{5}{50} = 0{,}9$.

La méthode pour trouver une échelle reste cependant assez empirique. Nous exposerons sommairement la méthode proposée par L. GUTTMAN en présentant le programme BMD05S dans le troisième chapitre.

Pour rester dans l'esprit de la relation de dominance, on pourrait chercher un ordre total sur les sujets minimisant le nombre d'avis en désaccord (cf. I, 3.3) en procédant de la façon suivante :

Pour chaque paire (k, l) de sujets, calculons les 3 quantités P_{kl}, P_{lk}, $P_{k \sim l}$ telles qu'on les a définies au § I 3.2. En appliquant la procédure de l'ordre médian, ou du nombre maximum d'avis en accord, on trouve un ordre total (programme AGREPREF). L'exemple ci-dessus donne pour P_{kl} et P_{lk} le tableau :

	7	4	8	5	2	9	10	1	6	3
7		1	1	2	2	2	3	3	4	5
4	0		1	1	2	1	2	2	3	4
8	0	1		1	1	2	2	2	4	4
5	0	0	0		1	1	1	1	3	3
2	0	1	0	1		1	1	2	3	3
9	0	0	1	1	1		1	2	2	3
10	0	0	0	0	0	0		1	2	2
1	0	0	0	0	1	1	1		2	2
6	0	0	1	1	1	0	1	1		1
3	0	0	0	0	0	0	0	0	0	

Lecture du tableau : (2 est avant 4 1 fois et 4 avant 2 2 fois). La même méthode appliquée à la recherche d'un ordre total sur les questions conduit à l'ordre sur les questions suivant :

	d	e	b	a	c
d		1	3	5	5
e	0		3	4	4
b	0	1		3	3
a	1	1	2		3
c	0	0	1	2	

On peut alors définir deux coefficients d'ajustement à un ordre. En effet, si on avait un ajustement parfait, on n'aurait que des 0 sous la diagonale.

- <u>Un coefficient d'ajustement des sujets</u>

$$\boxed{r_s = 1 - \frac{\Sigma_s}{n \cdot \frac{p(p-1)}{2}}}$$

où $\Sigma_s = \sum_{k>l} P_{lk}$ (somme des erreurs, c'est-à-dire des termes se trouvant sous la diagonale).

Dans l'exemple :

$$r_s = 1 - \frac{14}{5 \cdot \frac{10.9}{2}} = 0{,}938.$$

- Un coefficient d'ajustement des questions

$$\boxed{r_q = 1 - \frac{\Sigma q}{p \, \frac{n(n-1)}{2}}}$$

où $\Sigma_q = \sum_{i>j} P_{ji}$ (somme des erreurs, c'est-à-dire des termes se trouvant sous la diagonale).

Dans l'exemple :

$$r_q = 1 - \frac{8}{10 \cdot \frac{5.4}{2}} = 0{,}92.$$

Les deux coefficients sont différents.

Il semblerait que, sur les deux tableaux précédents, on puisse retrouver le nombre d'erreurs au sens de GUTTMAN en sommant pour chaque ligne la plus grande des valeurs située à gauche de la diagonale. Sur l'exemple, on a, pour le premier tableau, 5 lignes où il y a des 1, donc 5 erreurs. Dans le second tableau, la troisième ligne a un 1, la quatrièm un 2 et la cinquième un 2. Donc 5 erreurs également.

D'autres coefficients d'ajustement existent. Citons le coefficient d J. LOEVINGER, l'indice de consistance de GREEN mais nous renvoyons à A. DEGENNE [28], B. MATALON [60] et TORGERSON [78] pour de tels développements.

Citons également une extension du modèle de GUTTMAN en un modèle probabiliste, le modèle de l'ogive normale. Nous renvoyons à TORGERSON [86] pour un exposé de ce modèle. Dans [34], on trouvera un exposé en français ainsi qu'une application dans une enquête d'attitudes.

4. Extension à des échelles à plusieurs modalités

Les réponses ne sont plus dichotomiques mais se font à l'aide d'échelles à plusieurs modalités. On suppose ces modalités ordonnées (toujours, très souvent, assez souvent, rarement, jamais) par exemple.

En général, on cherche toujours un ordre sur les sujets mais on abandonne la recherche d'un ordre sur les questions.

La relation d'ordre cherchée est toujours la relation de dominance :

$$k \succcurlyeq l \Longleftrightarrow s_i^k \geq s_i^l \quad \forall i.$$

Exemple : Supposons trois questions a b c aux trois modalités ordonnées 1, 2, 3. Un ajustement parfait est, par exemple :

	a	b	c
k	3	2	3
l	3	1	2
m	1	1	2

On a la relation : $k \succ l \succ m$.

Par contre, une telle relation de dominance sur les question (à supposer qu'elle ait un sens) n'est plus complète : a et c sont incomparables.

La technique de GUTTMAN, programmée dans le programme BMD05S, s'applique également aux questions à plusieurs modalités.

Remarquons que la méthode exposée ci-dessus (§ 3) s'applique également.

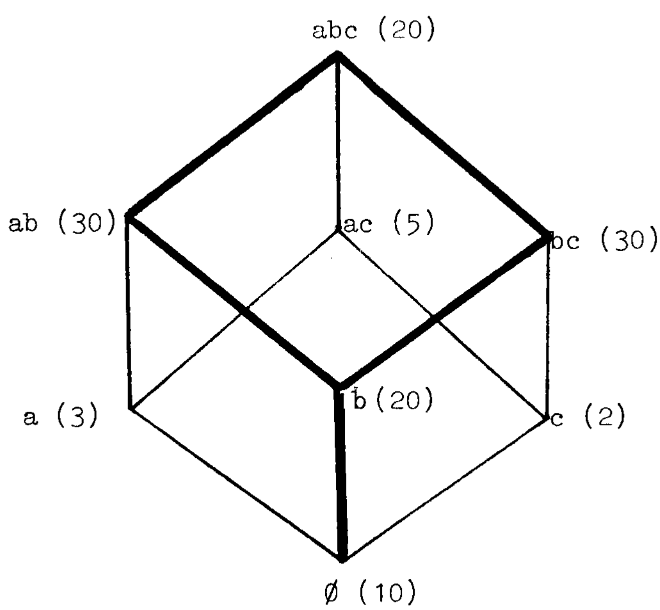
5. Ajustement à un ordre partiel et extension multidimensionnelle

Au lieu de chercher un ordre total sur les sujets, on peut chercher un ordre partiel. En effet, la relation d'ordre que l'on a définie sur l'ensemble des sujets est une relation d'ordre partiel.

Lorsqu'on ajuste un ensemble de patrons à un ordre total sur les sujets, on choisit une chaîne dans le simplexe (cas des données dichotomiques) et on considère comme erreur tout patron qui ne se situe pas sur cette chaîne.

Si au lieu de retenir une chaîne unique on en retient plusieurs, alors on retient un ensemble de patrons faisant partie du modèle, ces patrons étant ordonnés de façon partielle.

Exemple :



Supposons que nos données soient celles représentées sur le simplexe ci-dessus (p = 120 sujets). On constate que l'ajustement à une échelle de GUTTMAN sera nécessairement pas très bon, que ce soit l'échelle (abc, ab, b, ∅) ou l'échelle (abc, bc, b, ∅). On a 40 patrons en-dehors de l'échelle.

Par contre l'ajustement à l'ordre partiel représenté en traits forts est très bon puisque seuls 10 sujets sont en-dehors de ce modèle.

Dans DEGENNE [28] et MATALON [60], on trouvera un développement de ces questions.

Ces méthodes ont donné lieu à des programmes d'analyse hiérarchique multidimensionnelle (GUTTMAN and LINGOES - Multidimensional Scalogram Analysis - I, II and III) (cf. R.E. GREEN et F.J. CARMONE [41]).

CHAPITRE III

LES PROGRAMMES

Tous les programmes cités ci-dessous sont disponibles à la SEMA sur CDC 6600.

Une présentation générale des programmes a pour but de permettre une différentiation rapide quant à la nature des données en entrée, la nature des résultats en sortie, les méthodes utilisées.

Pour toute information sur les méthodes, on pourra se reporter aux deux premiers chapitres.

Les programmes présentés sont :

- programmes de description :

  . CORAN
  . PREFMAP I
  . PREFMAP II
  . MDPREF
  . FENELON
  . ANAPREF I
  . ANAPREF II

- programmes de structuration :

  . TRICON
  . ELECTRE I
  . ELECTRE II
  . AGREPREF
  . BMD05S

Puis on présentera les manuels de référence et des exemples d'utilisation pour les six programmes : MDPREF, FENELON, ANAPREF I, ANAPREF II, AGREPREF, BMD05S.

I - Présentation générale des programmes

On distinguera entre programmes de description et programmes de stru turation.

1. Programmes de description

1.1 Programmes de calcul d'indice

- CORAN [1]

Ce programme calcule :

- le coefficient de corrélation des rangs de SPEARMAN : ρ ;
- le coefficient de corrélation des rangs de KENDALL : τ ;
- le coefficient de concordance de KENDALL W.

Les entrées sont des classements sous forme de rangs et les sorties les valeurs des coefficients.

1.2 Programmes de visualisation

Dans le tableau ci-après, six programmes de description de données o dinales sont caractérisés suivant la nature des entrées, des sorties et de méthodes utilisées.

---

(1) Cf. Programmes de Structuration de Données - Direction Scientifique, Rapport de Recherche n° 52.

		PREFMAP I	PREFMAP II	MDPREF	FENELON	ANAPREF I	ANAPREF II
ENTREES Nature des données	rangs	X	X		X	X	
	évaluation sur des échelles	X	X	X		X	
	comparaisons par paires			X			X
	comparaisons par paires valuées						X
METHODE	analyse interne		X	X	X	X	X
	analyse externe	X	X				
	configuration déterminée dans R^r	X	X	X			
	projection du permutoèdre				X	X	X
SORTIES Nature des résultats	classement moyen avec coefficient associé	X	X		X somme pondérée W de KENDALL	X ordre médian C : cohésion	X ordre médian C : cohésion
	visualisation dans R^r	X	X	X	X	X r=2 seulement	X r=2 seulement

Une X signifie que le programme possède la caractéristique en ligne.

2. Programmes de structuration

2.1 Programmes de structuration d'une relation binaire quelconque

Les programmes suivants approximent une relation binaire quelconque par un ordre complet.

- TRICON [(1)]

Ce programme recherche à déterminer une ordonnance (ordre total sur distance entre objets) à partir de données de proximités par pivot.

- AGREPREF

Voir ci-dessous.

2.2 Programmes d'agrégation de plusieurs classements

- ELECTRE I [(2)]

Pour chaque classement (évaluation sur des échelles), on lit la nature de l'échelle, les évaluations des objets, les poids de chaque classement (critère), les ensembles de discordance et les seuils de concordance. Le programme donne la relation de surclassement et le noyau (ensembles des objets dans lequel doit se trouver le ou les meilleurs).

- ELECTRE II [(3)]

Les données sont les mêmes que celles d'ELECTRE I mais, en sortie, on obtient 3 préordres sur les objets : un préordre moyen et deux préordres extrêmes.

---

(1) Programmes de Structuration de Données - Annexe III - Direction Scientifique - Rapport de Recherche n° 52.

(2) ELECTRE (Manuel de Référence) - Direction Scientifique - Synthèse et Formation n° 25.

(3) Manuel de Référence du Programme ELECTRE II - Direction Scientifique Document Technique n° 24.

- AGREPREF

Ce programme agrège des relations de classement, quelle que soit la nature des données ordinales (rangs, évaluation sur des échelles, comparaisons par paires valuées ou non) selon différentes méthodes (somme pondérée, majorité par paires, ordre médian, relations de surclassement).

2.3 Programme d'analyse hiérarchique : BMD05S

Ce programme structure un tableau sujets x questions de données dichotomiques : il cherche un préordre (classement avec ex-aequo) des sujets (lignes). Il permet également de définir un ordre sur les questions.

Lorsque les réponses sont à plus de deux modalités, on obtient encore un préordre mais seulement sur les questions.

II - Manuel d'utilisation de certains programmes

Pour chacun des programmes MDPREF, FENELON, ANAPREF I, ANAPREF II, AGREPREF, BMD05S, on donnera :

- une présentation sommaire ;
- une description des cartes d'entrées ;
- une présentation des sorties ;
- un exemple d'utilisation.

Pour les autres programmes, on se reportera aux documents suivants :

- CORAN : Programme de Structuration de données - Direction Scientifique - Rapport de Recherche n° 52 ;
- TRICON : idem ;
- PREFMAP I : référence bibliographique [19] ;
- PREFMAP II : idem ;
- ELECTRE I : ELECTRE (Manuel de Référence) - Direction Scientifique Synthèse et Formation n° 25 ;
- ELECTRE II : Manuel de Référence du Programme ELECTRE II - Direction Scientifique - Document Technique n° 24.

1. MDPREF (1)

Ce programme a été écrit par J.J. CHANG selon la méthode de J.D. CARR exposée dans le premier chapitre (§ II, 1.2).

a) Présentation sommaire

MDPREF est un programme de description de données ordinales pouvant être des évaluations sur des échelles (des notes ou des scores) ou des comparaisons par paires. Il utilise une méthode d'analyse interne et un modèle vectoriel. Une configuration de points objets et de vecteurs sujets est calculée dans R^r $(r = 2, 3\ldots)$.

---

(1) Multi-Dimensional analysis of PREFerence data.

b) <u>Description des cartes d'entrée</u>

Les cartes à lire sont les suivantes :

- <u>Carte TITRE</u> - Toute information écrite sur cette carte sera reproduite en tête du programme.

- <u>Carte paramètre</u> - Cette carte contient les 10 entiers suivants perforés sur 4 colonnes (FORMAT 10I4) :

. NP - nombre de sujets (classements) $\leq$ 100 ;

. NS - nombre de stimuli (objets) $\leq$ 100 ; NP x NS $\leq$ 5 000 ;

. NF - nombre de facteurs (dimension r choisie) ; NF < MIN (NP, NS), NP x NF $\leq$ 1 000 ;

. NPF - nombre de facteurs à visualiser ;

. IREAD - nature des données

= 0 - lire des comparaisons par paires

= 1 - lire la matrice des scores (tableau des notes) à NP lignes et NS colonnes ; chaque ligne est centrée (on soustrait la valeur moyenne)

= 2 - même chose que pour IREAD = 1 mais chaque colonne est divisée par l'écart type de la colonne ;

. MDATA - facultatif, à introduire seulement si on a des comparaisons par paires

= 0 - pas de données manquantes, pas de poids par sujet

= 1 - des données sont manquantes, pas de poids

= 2 - introduction de poids

= 3 - des données sont manquantes et introduction de poids ;

. NS1 - facultatif, à introduire seulement si MDATA = 1 ou 2 ; nombre de sujets ayant les mêmes données manquantes ;

. NPUNCH = 0 - ne pas perforer les solutions sur cartes

= 1 - perforer les solutions sur cartes ;

. IPUNF = 0 - ne pas perforer la matrice des scores

= 1 - perforer la matrice des scores ;

. NBVIS = 0 - nombre de visualisations demandées.

- <u>Lecture des cartes sujets</u>

. <u>Lecture des comparaisons par paires</u> (si IREAD = 0)

+ <u>Carte format</u> - donnée à lire en Format A : la carte spécification et les comparaisons par paires de chaque sujet

+ <u>Carte des 4 spécifications</u> :

1ère spécification : code : ligne préférée à colonne
2e spécification : code : colonne préférée à ligne
3e spécification : code : pas de préférence
4e spécification : code : absence de données

+ <u>Comparaisons par paires de chaque sujet</u> - Chaque matrice de comparaisons par paires est lue dans le Format indiqué avec les codes indiqués.

. <u>Lecture de la matrice des scores</u> (si IREAD = 1 ou 2)

+ Carte Format
+ La matrice des scores ; il faut au moins une carte par sujet.

- <u>Une carte blanche ou un autre jeu de données</u>

c) <u>Présentation des sorties</u>

On obtient en sortie :

- le score moyen de chaque sujet (c'est-à-dire de chaque ligne ou chaque classement) ;
- la matrice des scores lus centrée ;
- la matrice des scores estimée ;
- les coordonnées des points sujets dans l'espace R^r (configuration
- les coordonnées des points objets dans l'espace R^r (configuration
- les visualisations dans les plans demandés :

. de l'ensemble des sujets,
. de l'ensemble des objets,
. des deux-ensembles superposés.

d) Exemple d'utilisation [1]

Pour une étude de marché, on a recueilli des préférences par paires sur 11 produits : on a placé 11 x 10/2 = 55 paires distinctes pour chaque sujet.

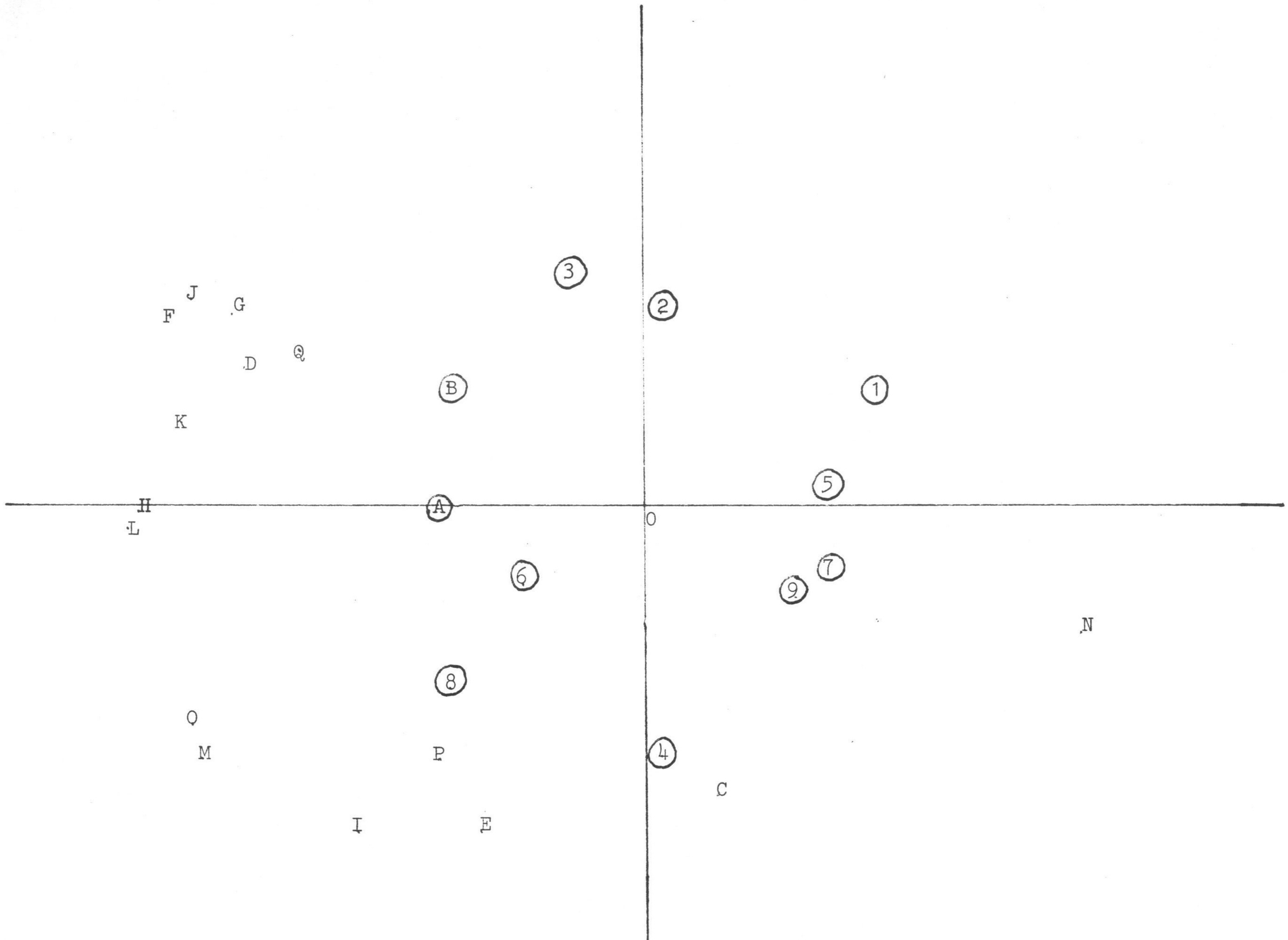
Pour chaque paire, on a demandé d'indiquer :

- si le produit de droite serait acheté de préférence au produit de gauche ;
- si le produit de gauche serait acheté de préférence au produit de droite ;
- si les produits seraient indifféremment achetés ;
- si il semble impossible de répondre.

On a obtenu la configuration moyenne dans R^3 dont on trouvera ci-après une projection sur les deux premiers axes. Il est à noter que la configuration n'a pas été calculée dans R^2. Si cela avait été le cas, les extrémités des vecteurs sujets (lettres C à Q) seraient toutes sur le cercle de centre 0 et de rayon 1.

---

(1) Utilisation du programme MDPREF - Note interne - 5 décembre 1972.



Les produits sont représentés par les codes 1 2 ... 9 A B

Les sujets sont représentés par les codes C, D ... Q.

2. FENELON - Programme d'analyse en composantes principales des rangs

Ce programme est une version CDC 6600 du programme PRENC 1 de J.P. FENELON que nous remercions vivement de nous avoir communiqué une version IBM 360 de son programme.

On trouvera un listing du programme et de nombreux exemples d'application dans la thèse de J.P. FENELON [36].

a) Présentation sommaire

C'est un programme de description de données ordinales qui sont nécessairement des rangs et il ne doit pas y avoir de données manquantes. Il s'agit d'une méthode d'analyse interne et le modèle est une projection du permutoèdre sur les premiers axes d'inertie du nuage des points classements. Le permutoèdre est également projeté sur l'axe moyen wG (cf. deuxième chapitre, § I 2.2.2).

b) Description des cartes d'entrée

Les cartes à lire sont les suivantes :

- Carte Paramètre - 14 paramètres à lire en Format (14I4)

. JJ - nombre d'objets (n) ;

. II - nombre de sujets (classements : p) ;

. FF - nombre de facteurs demandés ;

. LEC - numéro logique du lecteur de cartes (prendre LEC = 5) ;

. IMP - numéro logique de l'imprimante (à préciser dans la carte PROGRAM, prendre IMP = 6) ;

. FISUJ - numéro logique du fichier support des sujets (FISUJ = LEC = 5 si lecture par cartes) ;

. JD - paramètre dimension des tableaux pour le nombre d'objets (prendre JD = JJ) ;

. ID - paramètre dimension des tableaux pour le nombre de sujets (prendre ID = II + 1) ;

. FD - paramètre dimension des tableaux pour le nombre de facteurs (prendre FD = FF + 1) ;

. IMPSUJ = 1 - édition du tableau des données (rangs) ;

. PROPRE = 1 - édition du tableau des valeurs et vecteurs propres et des statistiques associées (pourcentage d'inertie, ...) ;

. FACI = 1 - une projection des coordonnées des points sujets sur les facteurs est demandée ;

. GRAPHE = 1 - la visualisation est demandée ;

. NC - nombre de cartes TITRE ; si NC = 0 ou blanc, le programme lit une carte titre.

- Cartes Titres - Le programme lit 1 carte titre ou NC cartes (si NC > 1).

- Carte Nom des objets - Le nom des objets est à lire en Format 20A4 (4 lettres par objet au plus).

- Carte Visualisation demandée (si GRAPHE = 1) - Lire en Format 20I4 couples d'axes pour lesquels on veut une visualisation

code 1 : axe moyen (appelé dans l'édition axe 0)
code 2 : 1er facteur (appelé dans l'édition axe 1)
code 3 : 2e facteur (appelé dans l'édition axe 2).

Exemple : bbb1bbb1bbb2bbb3 donnera :

. une projection sur l'axe moyen ;
. une projection sur le plan des 2 premiers facteurs.

- Carte Format de lecture des sujets.

- Cartes Sujets (classement) - Pour chaque sujet, (il faut au moins une carte par sujet), on lit d'après le format de lecture indiqué :

. le nom du sujet (format A) ;
. les rangs donnés à chaque objet (format F).

Remarque importante : La somme des rangs pour un sujet doit être égale à JJ(JJ + 1)/2

Exemple : 3 sujets A, B, C et 3 objets I, J, L

	I	J	L
A : (ordre J $\succ$ I $\succ$ L)	2	1	3
B : (préordre I $\sim$ J $\succ$ L)	1.5	1.5	3
C : (préordre I $\sim$ J $\sim$ L)	2	2	2

c) Présentation des sorties

On obtient en sortie :

- le tableau des données (rangs) ;
- le tableau des valeurs propres (valeurs, pourcentages d'inertie, pourcentages cumulés, histogramme) ;
- étude du vecteur O-G (wG) : origine centre du permutoèdre - centre de gravité des sujets. On obtient les coordonnées de ce vecteur sur les JJ (n) axes d'origine, la norme de ce vecteur, le carré de cette norme (W de KENDALL) ;
- coordonnées des points objets
  . sur l'axe moyen wG,
  . sur les axes factoriels ;
- coordonnées des points sujets
  . sur l'axe moyen wG,
  . sur les axes factoriels ;
- les visualisations demandées.

d) Exemple d'utilisation

On a demandé à des étudiants de l'Université PARIS-DAUPHINE de classer par ordre de préférence les 9 formules de vacances ci-dessous :

1) voyage en URSS
2) voyage en Amérique Latine
3) stage ouvrier
4) stage de voile en Bretagne
5) stage d'alpinisme
6) camp de travail dans un village

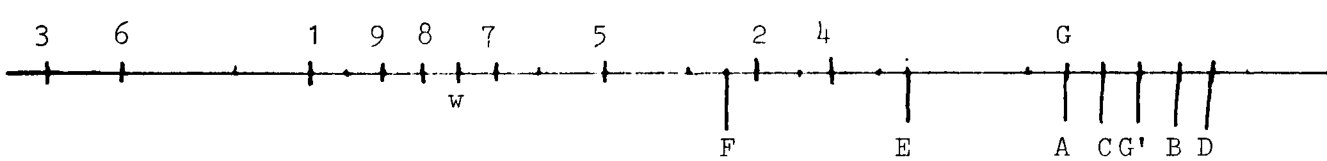
7) voyage en bicyclette
8) vacances artistiques (initiation au théâtre, ...)
9) stage agricole.

Ayant organisé les étudiants en 7 groupes (A, B, C, D, E, F, G), on demandé à chaque groupe de donner un classement.

Les données sont les suivantes :

	1	2	3	4	5	6	7	8	9
A	2	1	9	3	4	8	5	6	7
B	3	1	9	2	4	8	5	6	7
C	4	1.5	9	1.5	5	8	6	7	3
D	6	2	9	1	4	8	3	7	5
E	7	3	9	1	6	4.5	4.5	2	8
F	9	6	4	1	2	8	7	5	3
G	9	1	8	2	3	7	4	5	6

La projection sur l'axe moyen wG est :

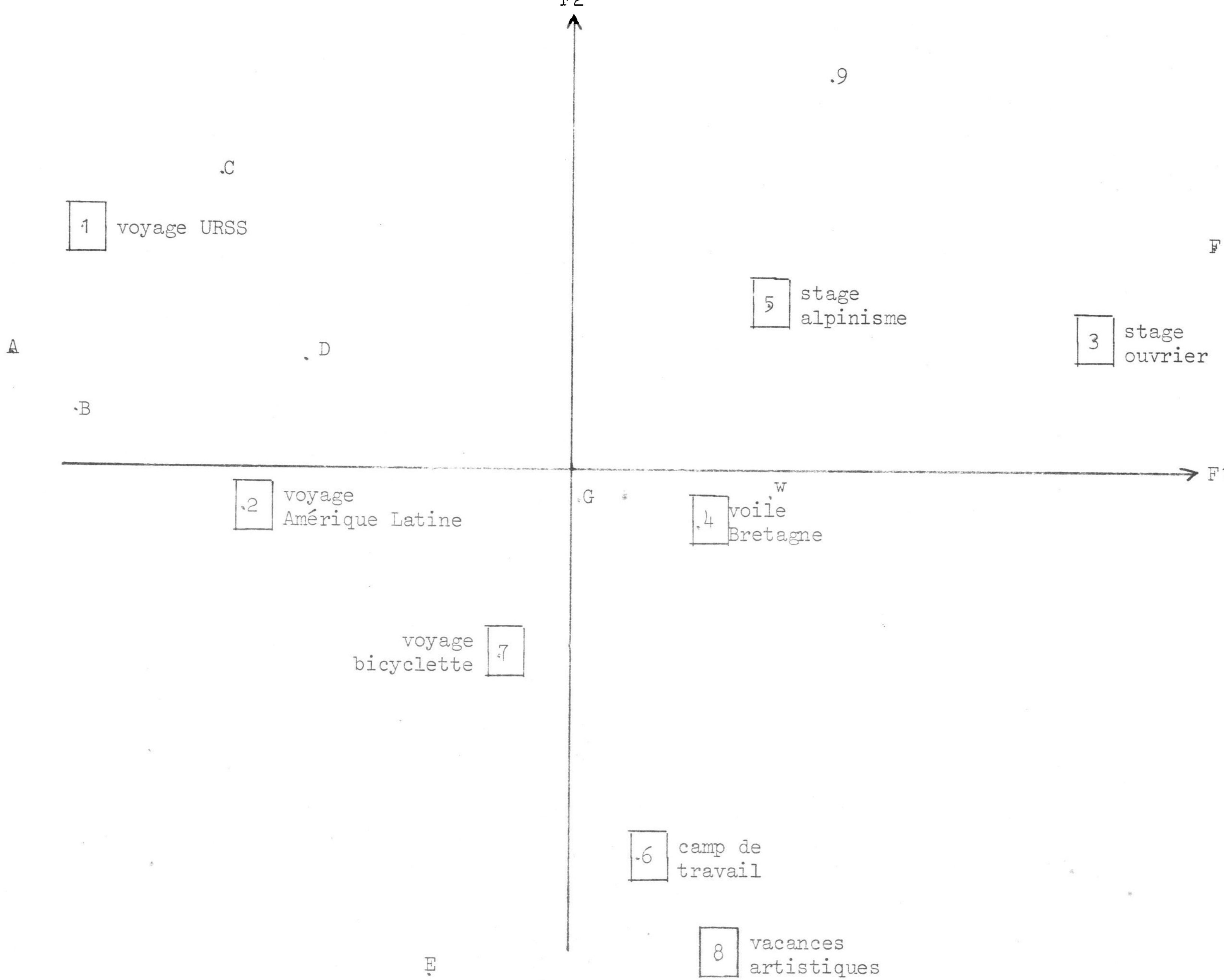


Le classement moyen est de l'ordre :

4, 2, 5, 7, 8, 9, 1, 6, 3.

Le coefficient de KENDALL est W = 0,61.

La projection sur les deux premiers axes factoriels (51 % et 29 % d'inertie) a donné la visualisation suivante :


F2
9
C
1 voyage URSS
F
5 stage alpinisme
3 stage ouvrier
A
D
B
F1
W
2 voyage Amérique Latine
G
4 voile Bretagne
voyage bicyclette 7
6 camp de travail
8 vacances artistiques
E

3. ANAPREF I [1]

a) Présentation sommaire

ANAPREF I est un programme d'analyse et de structuration de données ordinales qui sont des rangs ou des évaluations sur des échelles. La méthode d'analyse est interne et consiste en une projection du permutoèdre dans R^2. La méthode de structuration porte :

- sur les objets : recherche d'un ordre par la méthode de l'ordre médian ;
- sur les classements (sujets) : affectation en 2 groupes des classements.

Il est de plus possible de visualiser des classements passifs, c'est-à-dire des classements qui ne contribuent pas à la détermination des recherches d'ordres, des typologies sur les classements mais qui sont visualisés dans la projection.

b) Description des cartes d'entrée

- Carte NBRE - Nombre d'analyses à effectuer (nombre de jeu de données distinctes), Format I4.

- Carte Titre - Le programme lit une carte titre.

- Carte Paramètres - 5 paramètres sont à lire en Format 5I4 :

. N - nombre d'objets
. P - nombre de classements (y compris les éventuels classements passifs ;

---

(1) ANAlyse des PREFérences. Cf. [50] et Manuel de Référence du Programme ANAPREF - Direction Scientifique - Document Technique n° 30.

. LECT - nature des données
= 0 - données sous forme de rangs
= 1 - données sous forme d'évaluation sur des échelles ;

. NPA - nombre de classements passifs ;

. ICK = 1 si on veut lire des poids correspondant à chaque classement ; sinon, tout classement a un poids égal à 1.

- Carte Nom des objets - Le nom des objets est à lire en Format 25A3 (3 lettres au plus par nom).

- Carte poids des classements (seulement si ICK ≠ 0) - Les poids doivent correspondre à l'ordre des cartes sujets ; ils sont lus en 25A3, seulement pour les classements actifs.

- Carte Format de lecture des classements.

- Cartes classements - Pour chaque classement (il faut au moins une carte par classement), on lit d'après le format de lecture indiqué :

. le nom du classement (format A) ;
. les rangs ou évaluations sur des échelles (format F).

Remarque importante : Les classements passifs doivent venir après les autres.

c) Présentation des sorties

On obtient en sortie :

- un classement moyen (appelé tendance moyenne) obtenu par la méthode de l'ordre médian ainsi que les coefficients de cohésion et puissance ;
- la première typologie des classements (algorithme 1) et, pour chaque groupe, on détermine son classement moyen (méthode de l'ordre médian) et les indices associés ;
- la seconde typologie des classements (algorithme 2) ;

- un tableau donnant pour chaque classement en ligne :
. son affectation dans la seconde typologie,
. ses angles (divergences) avec les ordres médians G1, G2 des deux premiers groupes de la seconde typologie,
. ses coordonnées dans le plan (wG1, wG2),
. ses rangs ou évaluations des N objets ;
- un tableau donnant pour chaque objet :
. son rang sur G1,
. son rang sur G2,
. ses coordonnées dans le plan (wG1, wG2) ;
- une visualisation de la projection du permutoèdre sur le plan (wG1, wG2).

d) Exemple d'utilisation

Il s'agit de choisir, dans un ensemble de projets de recherche, ceux qui vont être financés et lancés. L'exemple ci-dessous a été traité à la SEMA sur des projets de recherche d'une entreprise. La méthode de déclassement comparé et ELECTRE II ont été utilisées sur les mêmes données, ce qui a permis de faire figurer leurs résultats comme classements ayant un caractère passif dans l'analyse. Figurent également comme classements passifs les classements de quatre experts B, C, D, E interrogés par ailleurs.

Dans cette analyse, on a :

29 projets N = 29
12 classements P = 12, dont
7 classements passifs NPA = 7.

Les 5 classements non passifs sont les critères :

ECO : économie
CHA : chances d'aboutissement
P.I. : propriété industrielle
STR : stratégie
COU : coût.

Les classements sur les 29 projets sont exprimés sous forme d'évaluations sur des échelles (LECT = 1).

Les données sont les suivantes :

Nom du projet	ECO	CHA	P.I.	STR	COU
12	1	9	7	10	3
13	1	5	7	10	3
211	12	9	4	10	2
212	10	5	7	10	2
221	7	5	4	7	1
222	1	3	7	10	1
312	10	5	4	7	2
331	12	7	1	4	2
341	7	5	4	10	2
74	17	9	7	10	3
114	7	7	4	7	3
11	17	5	7	4	3
42	10	7	7	7	2
51	10	3	1	7	1
53	17	3	4	7	2
54	10	7	4	7	3
55	12	7	4	10	1
56	1	7	7	7	3
71	10	3	4	10	1
73	12	5	7	10	1
81	10	9	4	7	2
83	10	5	7	7	3
84	7	3	4	7	2
85	1	7	4	7	3
91	17	3	1	7	2
101	12	3	7	7	2
102	7	5	4	4	3
103	1	5	4	1	3
106	12	7	7	10	3

Chaque classement est un préordre complet contenant un plus ou moins grand nombre de classes d'équivalence (3 classes pour le critère COU).

Un jeu de poids a été adopté pour les 5 critères (ICK ≠ 0) : c'est l même jeu (10, 6, 5, 4, 2, total : 27) qui est utilisé avec ELECTRE II et donne le classement EL 1 (un autre jeu de poids diminuant l'importance relative du critère COU a donné le classement noté EL 2).

Le classement obtenu par le programme est l'ordre noté GM : 74, 106, 11, 211, 55, 73, 212, 12, 53, 91, 101, 81, 42, 54, 331, 83, 312, 71, 114, 341, 13, 56, 51, 221, 102, 84, 222, 85, 103.

La cohésion est : 0.36.

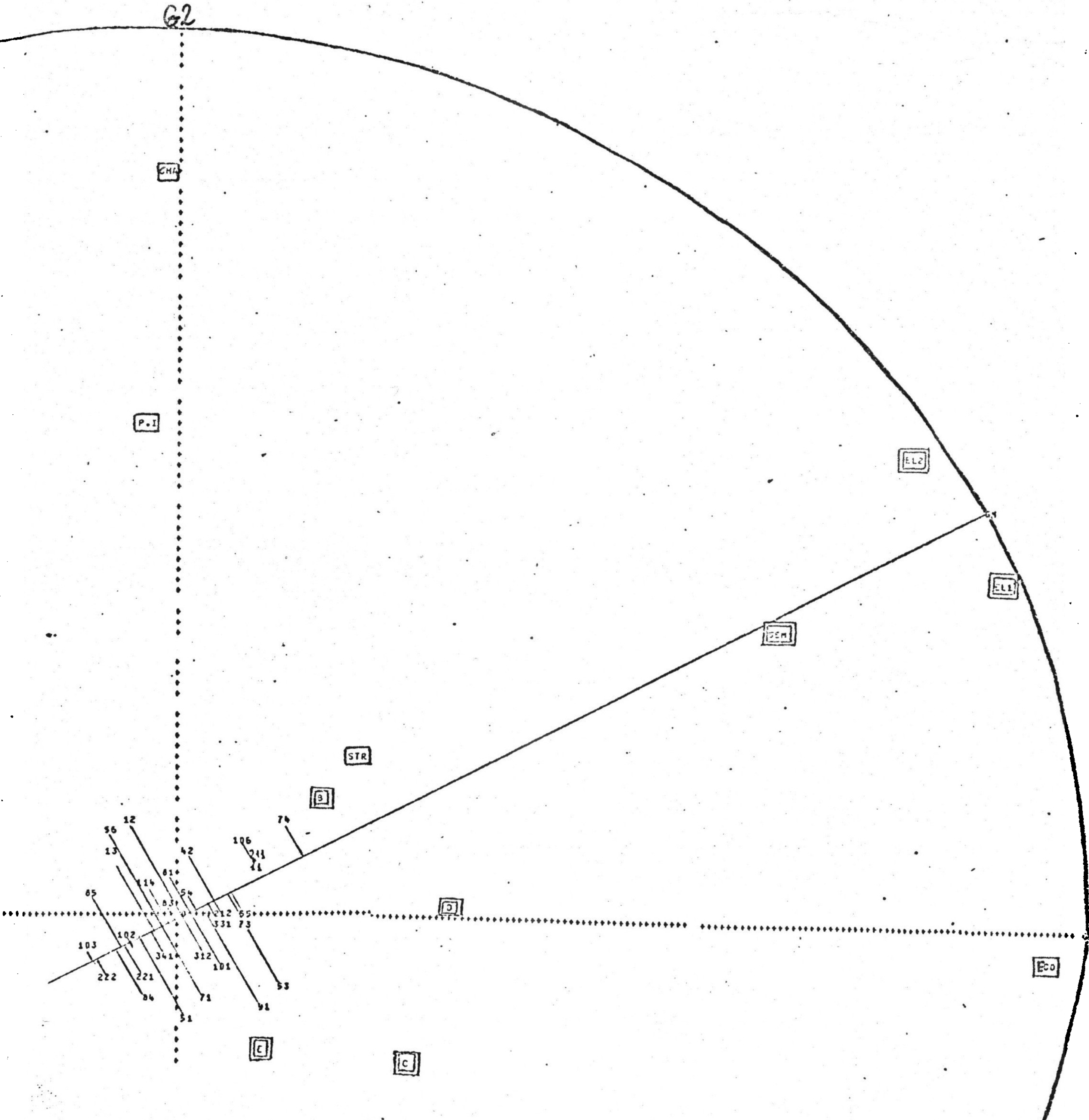
La visualisation (cf. graphique ci-après) représente les 29 projets (autour de l'origine O) et les 10 classements.

En projetant les projets sur un axe joignant l'origine à un point re présentatif d'un classement, on obtient une approximation du classement. Cette approximation sera d'autant meilleure que le point sera décentré (STR est mal représenté par exemple).

Une lecture de la visualisation permet de dire : si on augmente l'im portance du critère ECO, les projets 53, 91 pourraient passer en tête et, d'une façon plus générale, tous les projets situés du même côté de l'axe O, GM que ECO verraient leur classement s'améliorer au détriment des projets de l'autre côté de cet axe (12 et 56 par exemple). Au contraire, si on diminue l'importance du critère ECO, l'importance relative des autres critères va augmenter et le point GM va s'éloigner de ECO pour se rapprocher des autres critères. Les projets situés au-dessus (12, 56) vont alors obtenir un meilleur classement.

Les quatre experts B, C, D, E sont assez proches du centre et leur classement est donc mal représenté sur cette visualisation.

Les résultats de la méthode des déclassements comparés (SEM), de ELECTRE II (EL 1, EL 2) et d'ANAPREF I (GM) sont assez voisins.


G2
CHA
P.I
EL2
EL1
DEM
STR
B
D
C
C
Eco

Les divergences entre l'ordre trouvé par ANAPREF : GM et les autres classements exprimées en degrés et en tau de KENDALL sont :

	SEM	EL1	EL2	B	C	D	E
degrés α	28	12	15	79	82	75	9
tau de KENDALL τ	.69	.87	.83	.12	.09	.17	.-0

La relation entre α et τ étant :

$$\boxed{\tau = 1 - \frac{\alpha}{90^\circ}}$$

On constate donc un bon accord entre les trois méthodes. Par contre les classements directs des experts sont assez différents et plus proches cependant de ECO. Peut-être ont-ils donné implicitement plus de poids à ce critère ?

4. ANAPREF II

a) Présentation sommaire

ANAPREF II utilise les mêmes méthodes d'analyse et de structuration qu'ANAPREF I. La différence porte sur la nature des données qui sont ici des comparaisons par paires valuées ou non. Si elles sont valuées (préfé-rences floues, relations de comparaisons par paires agrégées), alors il f $a^k_{ij} + a^k_{ji} \leq 1$ pour toute paire d'objets (i, j) et pour tout classement (a^k_{ij} : relation valuée entre i et j suivant le classement k).

b) Description des cartes d'entrée

Les cartes à lire sont les suivantes :

- Carte NBRE - Nombre d'analyses à effectuer (format I4).

- Carte TITRE - Le programme lit 1 carte titre.

- Carte Paramètres - 2 paramètres sont à lire en format 2I4 :

. N - nombre d'objets ;
. P - nombre de classements.

- Carte Nom des objets - Le nom des objets est à lire en format 25A3 (3 lettres par nom).

- Carte Format de lecture des classements.

- Cartes classements - Il faut pour chaque classement au moins une carte distincte. On lit :

. le nom du classement ;
. le tableau des comparaisons par paires correspondant au classement (on a P tableaux à lire).

c) Présentation des sorties

- Pour chaque classement, le programme donne l'ordre à distance minimum du classement. Si on a des comparaisons par paires non valuées (match, tournoi), cet ordre est obtenu en inversant un nombre minimum d'arcs (cf. deuxième chapitre, § II, 3.2.2). Si les comparaisons sont valuées, l'ordre est celui qui maximise les avis en accord (cf. deuxième chapitre, § I, 3.3) (c'est également l'ordre à distance minimum de la relation valuée). Le programme donne également la cohésion qui représente ici un indice d'ajustement de la relation à un ordre complet.

- La suite du programme donne les mêmes résultats qu'ANAPREF I, excepté le fait que GM, appelé opinion moyenne, est la relation valuée moyenne $(\underline{a}^{GM} = \frac{1}{p} \sum_{k=1}^{p} \underline{a}^{k})$ et non plus l'ordre à distance minimum de cette relation valuée, c'est-à-dire l'ordre obtenu par la méthode de la recherche d'un ordre maximisant les avis partiels en accord (cf. deuxième chapitre, § I, 3.3).

d) <u>Exemple d'utilisation</u>

L'exemple est emprunté à M. HUTEAU [47].

480 enfants ont répondu au questionnaire de comparaisons par bloc présenté au § 3.1 de l'Annexe II et portant sur les 25 métiers suivants :

MED : médecin
JOU : journaliste
PRØ : professeur
ART : artiste
ECR : écrivain

AVI : aviateur
EXP : explorateur (exploratrice)
IGM : ingénieur mécanicien
IGA : ingénieur agronome
CHI : chimiste

INF : infirmier (ère)
INS : instituteur (trice)
COM : comptable
TEC : technicien (ienne)
ELE : électronicien (ienne)

DEI : dessinateur industriel
EMP : employé (e) de bureau
VEN : vendeur (euse)
TAX : chauffeur de taxi
MCA : mécanicien (ienne)

HOR : horloger (ère)
IMP : imprimeur
COI : coiffeur (euse)
COU : couturier (ère)
CVA : cultivateur

On a constitué 16 groupes de 30 enfants en utilisant les critères de segmentation suivants :

	Filles					Garçons			
	6e	5e	4e	3e		6e	5e	4e	3e
cadre	6FC	5FC	4FC	3FC	cadre	6GC	5GC	4GC	3GC
ouvrier	6FO	5FO	4FO	3FO	ouvrier	6GO	5GO	4GO	3GO

<u>Exemple</u> : 6FC : Filles en classe de 6e d'origine cadre
4GO : Garçons en classe de 4e d'origine ouvrière.

Pour chacun des 16 groupes ainsi constitués, on a calculé une <u>relation de préférence valuée</u> :

$$a_{ij} = \frac{1}{30} \sum_{k=1}^{p} a_{ij}^{k}.$$

Cette relation se traduit par un tableau où $a_{ij} + a_{ji} = 1 \ \forall i \neq j$.

Les données étaient donc :

P = 16 (16 classements à analyser)

N = 25 (comparaisons par paires valuées portant sur 25 métiers).

Pour chaque groupe, on a obtenu l'ordre à distance minimum. Par exemple, pour le groupe 6GC (garçons en classe de 6e d'origine cadre), on a obtenu le tableau ci-après.

Les termes situés sous la diagonale et supérieurs à 0.5 correspondent à des effets CONDORCET (relation de majorité par paires intransitives).

Le classement moyen correspondant aux 16 groupes réunis (donc à l'ensemble des 480 enfants) fut le suivant :

MED, PRO, AVI, JOU, EXP, CHI, INS, DEI, INF, AET, ELE, IGM, IGA, TEC, COM, EMP, ECR, MCA, COI, VEN, COU, IMP, HOR, CVA, TAX

la cohésion étant 0.34.

La typologie a donné deux groupes : un groupe contient les 8 classements des 8 groupes garçons, le second contient les 8 autres classements des 8 groupes filles.

Le groupe 1 (garçons) a pour tendance l'ordre : AVI, EXP, CHI, ELE, IGM, JOU, MED, DEI, IGA, TEC, PRO, INS, MCA, ART, INF, COM, HOR, EMP, ECR, IMP, CVA, TAX, VEN, COI, COU.

Sa cohésion est 0.43.

PREFERENCES SUR DES METIERS - HUTEAU

$\boxed{6GC}$

GROUPE NO 1 COHESION 44.765 EFFECTIF 1 PUISSANCE 0.45

25	EXP	AVI	CHI	IGM	DEI	MED	ELE	TEC	IGA	MCA	PRO	JOU	INF	INS	COM	ART	EMP	CVA	ECR	IMP	HOR	VEN	TAX	COI	COU
EXP	0.0	0.5	0.6	0.7	0.7	0.8	0.6	0.6	0.9	0.8	0.8	0.8	0.9	1.0	0.9	0.9	1.0	0.9	0.8	0.9	0.9	1.0	0.9	0.9	1.0
AVI	0.5	0.0	0.6	0.6	0.6	0.7	0.6	0.6	0.7	0.7	0.8	0.8	0.7	0.8	1.0	0.8	0.9	0.8	0.8	0.8	1.0	0.9	0.9	0.9	0.9
CHI	0.4	0.4	0.0	0.6	0.7	0.5	0.6	0.7	0.6	0.6	0.7	0.8	0.8	0.7	0.9	0.9	0.8	0.9	0.9	0.9	0.9	0.8	0.9	0.9	0.9
IGM	0.3	0.4	0.4	0.0	0.6	0.5	0.6	0.5	0.7	0.8	0.7	0.6	0.5	0.8	0.8	0.7	0.9	0.8	0.7	0.8	0.9	0.9	0.8	0.9	0.9
DEI	0.3	0.4	0.3	0.3	0.0	0.5	0.6	0.5	0.7	0.7	0.6	0.6	0.7	0.8	0.9	0.8	0.9	0.9	0.7	0.9	0.9	0.9	0.9	0.9	1.0
MED	0.2	0.3	0.5	0.5	0.5	0.0	0.5	0.7	0.7	0.7	0.8	0.6	0.7	0.8	0.9	0.8	0.9	0.9	0.8	0.7	0.8	0.9	0.9	1.0	1.0
ELE	0.4	0.4	0.4	0.4	0.4	0.5	0.0	0.6	0.5	0.6	0.7	0.7	0.7	0.7	0.8	0.7	0.8	0.8	0.6	0.9	0.8	0.9	0.9	0.8	0.9
TEC	0.4	0.4	0.3	0.5	0.5	0.2	0.4	0.0	0.6	0.6	0.8	0.6	0.6	0.7	0.8	0.7	0.8	0.9	0.8	0.7	0.8	0.8	0.8	0.9	1.0
IGA	0.1	0.3	0.4	0.3	0.3	0.3	0.5	0.4	0.0	0.6	0.5	0.5	0.7	0.6	0.7	0.7	0.8	0.8	0.6	0.8	0.8	0.8	0.8	0.9	0.8
MCA	0.2	0.3	0.4	0.2	0.3	0.3	0.4	0.4	0.4	0.0	0.5	0.5	0.5	0.6	0.7	0.6	0.7	0.7	0.6	0.7	0.7	0.9	0.8	0.7	0.9
PRO	0.2	0.2	0.3	0.3	0.4	0.2	0.3	0.2	0.5	0.5	0.0	0.5	0.5	0.7	0.7	0.6	0.7	0.7	0.7	0.7	0.8	0.8	0.8	0.9	0.9
JOU	0.2	0.2	0.2	0.4	0.4	0.4	0.3	0.4	0.5	0.5	0.5	0.0	0.5	0.6	0.7	0.6	0.6	0.7	0.7	0.3	0.7	0.5	0.7	0.8	0.8
INF	0.1	0.3	0.2	0.5	0.3	0.3	0.3	0.4	0.3	0.5	0.5	0.5	0.0	0.5	0.5	0.7	0.7	0.6	0.5	0.7	0.8	0.7	0.8	0.8	0.9
INS	0.0	0.2	0.3	0.2	0.2	0.2	0.3	0.3	0.4	0.4	0.3	0.3	0.5	0.0	0.7	0.6	0.8	0.6	0.7	0.6	0.8	0.9	0.8	0.8	0.9
COM	0.1	0.0	0.1	0.2	0.1	0.1	0.2	0.2	0.3	0.3	0.3	0.3	0.5	0.3	0.0	0.5	0.6	0.5	0.4	0.6	0.5	0.5	0.5	0.7	0.8
ART	0.1	0.2	0.1	0.3	0.2	0.2	0.3	0.3	0.3	0.4	0.4	0.4	0.3	0.4	0.5	0.0	0.7	0.5	0.6	0.6	0.5	0.7	0.7	0.8	0.9
EMP	0.0	0.1	0.2	0.1	0.1	0.1	0.2	0.2	0.2	0.3	0.3	0.4	0.3	0.2	0.4	0.3	0.0	0.5	0.5	0.5	0.4	0.5	0.5	0.6	0.7
CVA	0.1	0.2	0.1	0.2	0.1	0.1	0.2	0.1	0.2	0.3	0.3	0.3	0.4	0.4	0.4	0.5	0.5	0.0	0.6	0.6	0.4	0.7	0.5	0.5	0.9
ECR	0.2	0.2	0.1	0.3	0.3	0.2	0.4	0.2	0.4	0.4	0.3	0.3	0.4	0.3	$\boxed{0.6}$	0.4	0.5	0.4	0.0	0.5	0.6	0.5	0.5	0.6	0.8
IMP	0.1	0.2	0.1	0.2	0.1	0.3	0.2	0.3	0.2	0.3	0.3	0.2	0.3	0.4	0.4	0.4	0.5	0.4	0.5	0.0	0.6	0.5	0.6	0.6	0.7
HOR	0.1	0.0	0.1	0.1	0.1	0.2	0.2	0.2	0.2	0.3	0.2	0.3	0.2	0.2	0.4	0.5	$\boxed{0.6}$	$\boxed{0.6}$	0.4	0.4	0.0	0.6	0.6	0.5	0.8
VEN	0.0	0.1	0.2	0.1	0.1	0.1	0.1	0.2	0.2	0.1	0.2	0.4	0.3	0.1	0.5	0.3	0.5	0.3	0.5	0.5	0.4	0.0	0.6	0.6	0.8
TAX	0.1	0.1	0.1	0.2	0.1	0.1	0.1	0.2	0.2	0.2	0.2	0.3	0.2	0.2	0.5	0.3	0.5	0.5	0.5	0.4	0.4	0.4	0.0	0.5	0.5
COI	0.1	0.1	0.1	0.1	0.1	0.0	0.2	0.1	0.1	0.3	0.1	0.2	0.2	0.2	0.3	0.2	0.4	0.5	0.4	0.4	0.4	0.4	0.5	0.0	0.8
COU	0.0	0.1	0.1	0.1	0.0	0.0	0.1	0.0	0.2	0.1	0.1	0.2	0.1	0.1	0.2	0.1	0.3	0.1	0.2	0.3	0.2	0.2	0.4	0.8	[illegible]

Le groupe 2 (filles) a pour tendance l'ordre : PRO, MED, INS, INF, JOU, EXP, CHI, ART, COI, DEI, AVI, COM, EMP, COU, VEN, ECR, TEC, ELE, IGA, IMP, HOR, IGM, MCA, CVA, TAX.

Sa cohésion est 0.42.

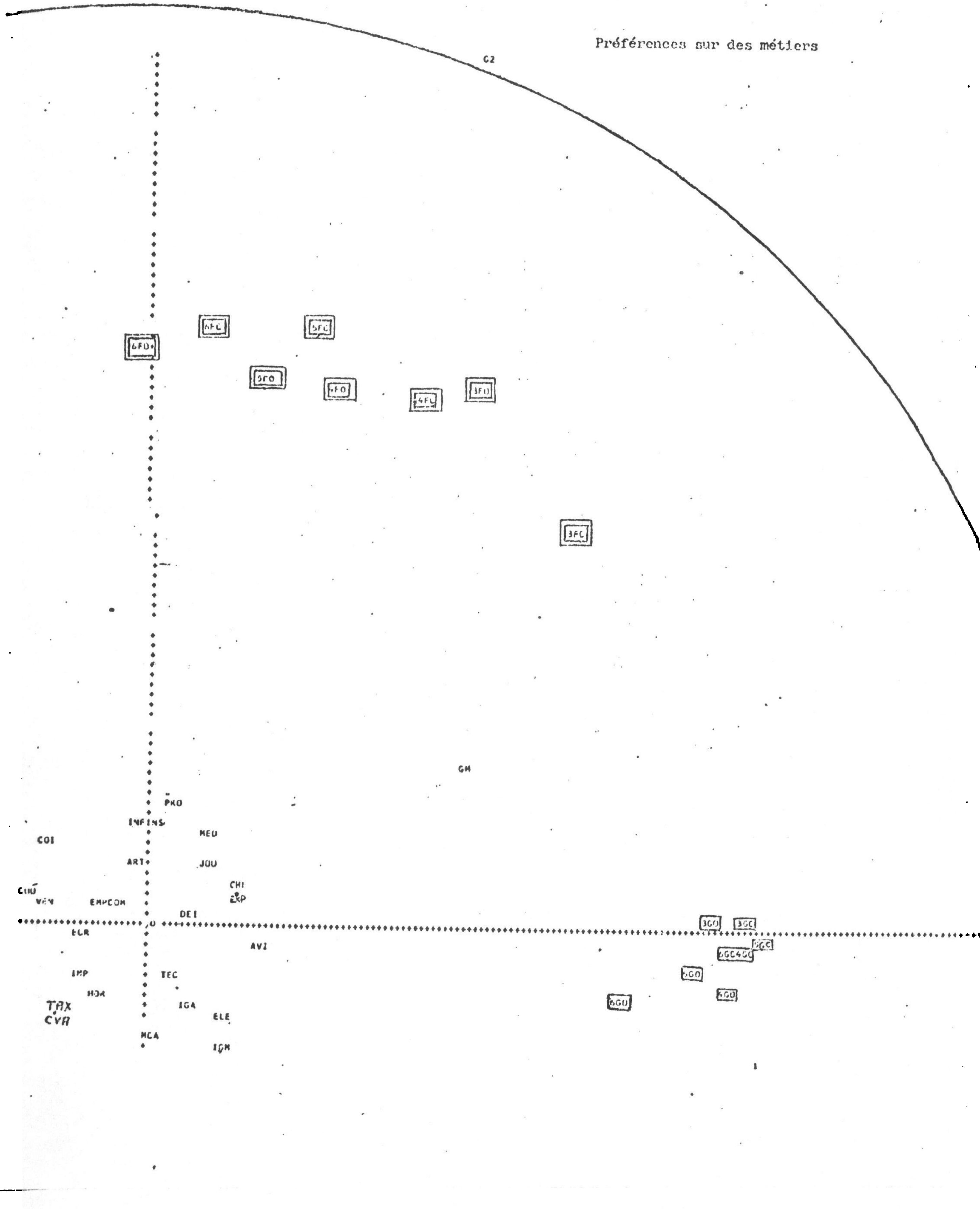
On constate sur la visualisation (voir graphique ci-après) que les filles et les garçons se différencient nettement sans être pour autant en opposition. La tendance de l'opinion collective est en effet assez nette : cohésion 0.34 et GM est assez décentré, donc distinct de l'indifférence (que représente l'origine 0).

<u>Remarque</u> : La cohésion des filles est moins grande que celle des garçons (0.43 et 0.42). La cohésion moyenne devrait être petite (0.34).

Les métiers se distinguent d'une part entre métiers appréciés et métiers dévalorisés (préférence collective, projection de ces métiers sur l'axe 0, GM) et d'autre part entre métiers féminins et métiers masculins. Plus un métier est éloigné de l'axe 0, GM, plus il est représentatif des filles ou des garçons (coiffeuse pour les métiers féminins, ingénieur mécanicien pour les métiers masculins). Par contre, les métiers voisins de l'axe GM ne sont pas typés (journaliste, chimiste, dessinateur industriel, imprimeur).

Sur les sujets, on remarque que les filles, en mûrissant, se rapprochent de l'opinion collective fille + garçons, ce qui n'est pas le cas des garçons (sauf peut-être pour ceux de la classe de 4e). Les enfants d'origine ouvrière ont des préférences plus typées.

En résumé, les filles, en mûrissant, et les enfants d'origine cadre, ont des préférences plus homogènes se rapprochant d'une certaine norme sociale exprimée par la tendance de l'opinion collective.


Préférences sur des métiers
G2
6FO
6FC
5FC
5FO
4FO
4FC
3FO
3FC
GM
PRO
INFINS
MED
COI
ART
JOU
CHI
EXP
VEN
EMPCOM
DEI
ECR
AVI
3GO
3GC
5GC
6GC
4GC
5GO
4GO
6GO
IMP
TEC
HOR
TAX
CVA
IGA
ELE
MCA
IGM

5. AGREPREF (1)

a) Présentation sommaire

AGREPREF est un programme d'agrégation de p classements regroupant un ensemble de méthodes qui sont présentées dans le tableau ci-après. Les données ordinales peuvent être :

- des rangs ;
- des évaluations sur des échelles ;
- des évaluations sur des échelles avec des seuils d'indifférence sur ces échelles (agrégation de quasi-ordres) ;
- des comparaisons par paires valuées ou non.

Méthodes / Nature des données	Somme pondérée des rangs	Ordre médian	Majorité par paires	Relations de surclassement (cf. deuxième chapitre, § I 3.6 et I 4.2)
Rangs	X	X	X	X
Evaluation sur des échelles		X	X	X
Evaluation sur des échelles avec seuils d'indifférence sur chaque échelle		X	X	X
Comparaisons par paires		X	X	X

Une X désigne que la méthode en ligne peut être utilisée avec les données en colonne.

---

(1) AGREgation des PREFérences - Programme écrit en juillet 1973 à la Direction Scientifique.

b) <u>Description des cartes d'entrée</u>

Les cartes à lire sont les suivantes :

- <u>Carte NBRE</u> - Nombre d'agrégations à effectuer (format I4).

- <u>Carte TITRE</u> - Le programme lit une carte titre.

- <u>Carte Paramètres</u> - 8 paramètres sont à lire en format 8I4 :

. N - nombre d'objets ;

. NP - nombre de classements ;

. LECT - nature des données

= 1 - rangs

= 2 - évaluations sur des échelles (avec ou sans seuils d'in-différence)

= 3 - comparaisons par paires (valuées ou non) ;

. ICK = 1 si on veut lire des poids correspondants à chaque classement. Si ICK = 0 ou blanc, alors tous les classements ont un poids égal à 1 ;

. IDK = 1 si on veut lire des seuils d'indifférence sur les échelles Si IDK = 0 ou blanc, les seuils d'indifférence sur tous les classements so égaux à 0 ;

. INVK = 1 si on veut lire une carte modifiant le sens des préférenc suivant une ou plusieurs échelles. Sinon (INVK = 0 ou blanc), le programm interprète les valeurs lues comme :

préférence allant en croissant avec ces valeurs si LECT = (évaluation sur des échelles)

préférence allant en décroissant avec ces valeurs si LECT = 1 (rangs) ;

. IEDITK = 1 si on veut une édition du tableau des données initiales (pour LECT = 1 et 2) et une édition de chaque tableau de comparaisons par paires réordonnée (si LECT = 3) ;

. ISEUIL = 1 si on veut construire une ou plusieurs relations de sur classement par une méthode d'agrégation à seuils.

. METH 1 méthode 1
2 " 1 + 2
3 " 1 + 2 + 3
4 " 3

- Carte Nom des objets - Le nom des objets est à lire en format 25A3 (3 lettres par nom).

- Carte Seuils (seulement si ISEUIL = 1) - On lit en format 10F3.2 5 paires de seuils d'indifférence et de préférence strictes au maximum (la somme des deux seuils doit être $\geq$ 1), chaque paire correspondant à la construction d'une relation de surclassement.

- Carte des poids (seulement si ICK = 1) - Lire le poids des NP classements en format 25F3.0.

- Carte des seuils d'indifférences par échelles (seulement si IDK = 1) - Lire le seuil d'indifférence sur chaque classement (à exprimer en grandeur absolue relativement à l'échelle utilisée) - Format 12F6.1.

- Carte Inversion sur les échelles (seulement si INVK = 1) - Lire en format 80F1.0 le sens des préférences voulu en utilisant le code :

. 1 : code "rangs", c'est-à-dire préférence allant en décroissant lorsque la valeur croît (exemple : un critère coût) ;

. 2 : code "notes", c'est-à-dire préférence allant en croissant lorsque la valeur croît.

- Carte Format de lecture des classements.

- Cartes Classements - Il faut au moins une carte par classement. Pour chaque classement, on lit :

. le nom du classement (format A) ;

. les rangs ou évaluations sur des échelles (N nombres à lire) si LECT = 1 ou 2 ;

. les tableaux de comparaisons par paires (N x N nombres à lire) si LECT = 3.

c) Présentation des sorties

- Si on a des comparaisons par paires (LECT = 3) et si on demande l'édition des données (IEDITK = 1), on obtient, pour chaque classement, le tableau des comparaisons par paires réordonné.

- Si on a des rangs ou des évaluations sur des échelles (LECT = 1 ou 2) et si on demande l'édition des données (IEDITK = 1), on obtient le tableau des données avec en colonne les objets (1 objet par colonne) et en ligne les classements pour lesquels figurent :

. le nom, le code d'inversion d'échelle, le poids, le seuil d'indifférence et enfin les évaluations ou les rangs de chaque objet ;

. éventuellement le classement par la méthode de la somme pondérée

. le tableau des préférences strictes (P_{ij} : cf. § II 1.3.6, 2e chapitre) ;

. le tableau des indifférences ($P_{i \sim j}$: cf. § II 1.3.6, 2e chapitre)

. le classement par la méthode de l'ordre médian ;

. la relation obtenue par la méthode des majorités par paires présentée sous une forme réordonnée de façon à la rendre lisible. Les effets CONDORCET (relation intransitive) se remarquent par des 1 situés au-dessous de la diagonale ;

. les différentes relations de surclassement demandé présentées sous une forme matricielle réordonnée de façon à les rendre plus lisibles. Un 0.5 au croisement de la ligne i et de la colonne j signifie la relation d'indifférence i I j (j I i). Un 1 au croisement de la ligne i et de la colonne j signifie la relation de préférence stricte i P j (non j P i).

d) Exemple d'utilisation

Nous avons appliqué la méthode d'agrégation de quasi-ordres sur le problème du média-planning exposé dans [16]. Rappelons brièvement le problème. Il s'agit de sélectionner des supports de presse (14 dans l'exemple) évalués suivant différents critères qualitatifs et quantitatifs en vue de faire passer des annonces publicitaires.

Six critères ont été retenus :

- COU : le coût aux mille lecteurs utiles exprimé en francs ;
- CON : le contexte rédactionnel : mesure, par une note de 0 à 10, une bonne adéquation entre le contenu du magazine et le contenu des annonces publicitaires de ce magazine ;
- PUI : la puissance mesure l'ampleur de l'audience et la fiabilité de l'information transmise ;
- AFF : ce critère mesure le degré d'affinité du support avec la cible visée par l'onnonce publicitaire que l'on veut faire passer ;
- REG : ce critère mesure la régularité de lecture du support ;
- PRE : le prestige du support ; ce critère est important si l'annonce publicitaire concerne des produits considérés comme prestigieux (note de 1 à 10).

Les données sont les suivantes :

CRITERES / SUPPORTS	Contexte CON	Coût aux 1 000 (en Fr.) COU	Régula-rité REG	Puissance PUI	Affinité AFF	Presti PRE
L'EXPRESS (Ex)	4	114	6	3	9	7
JOURS DE FRANCE (JF)	10	58	5	6	7	9
MODES DE PARIS (MP)	7	48	7	5	5	5
MLLE AGE TENDRE (MAT)	6	77	6	3	5	3
ELLE (E)	10	51	4	5	8	9
FEMMES D'AUJ. (FA)	7	62	6	5	6	5
INTIMITE (I)	5	74	8	2	5	3
NOUS DEUX (ND)	5	125	7	2	5	3
MODES & TRAVAUX (MT)	6	55	6	9	5	4
ECHO DE LA MODE (EM)	6	86	6	3	5	4
MARIE-CLAIRE (MC)	10	59	4	6	8	9
MARIE-FRANCE (MF)	9	59	4	5	7	7
FEMME PRATIQUE (FP)	7	51	4	4	7	6
JARDIN DES MODES (JM)	10	65	4	2	8	10

Les poids sur les critères CON, COU, REG, PUI, AFF, PRE sont 6, 4, 2, 3, 3, 2 et les seuils d'indifférence sont (2, 10 Francs, 2, 2, 2, 2).

La méthode d'agrégation de ces critères en une relation de surclassement présentée dans le deuxième chapitre, § I 4.2 a donné le résultat suivant en utilisant les seuils d'indifférence $s_i = 0{,}70$ et de préférence stricte $s_p = 0{,}30$:

	F	JF	MC	MF	JM	MP	FA	MT	FP	MAT	EM	I	EX	ND
E	0.0	.5	.5	.5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
JF	.5	0.0	.5	.5	.5	1.0	1.0	0.0	1.0	1.0	1.0	1.0	1.0	1.0
MC	.5	.5	0.0	.5	.5	0.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
MF	.5	.5	.5	0.0	.5	.5	.5	0.0	.5	1.0	1.0	1.0	1.0	1.0
JM	0.0	.5	.5	.5	0.0	0.0	0.0	1.0	0.0	1.0	1.0	1.0	1.0	1.0
MP	0.0	0.0	0.0	.5	0.0	0.0	.5	.5	.5	.5	.5	1.0	1.0	1.0
FA	0.0	0.0	0.0	.5	0.0	.5	0.0	.5	.5	.5	.5	1.0	1.0	1.0
MT	0.0	0.0	0.0	0.0	0.0	.5	.5	0.0	.5	1.0	1.0	1.0	0.0	1.0
FP	0.0	0.0	0.0	.5	0.0	.5	.5	.5	0.0	.5	.5	0.0	1.0	0.0
MAT	0.0	0.0	0.0	0.0	0.0	.5	.5	0.0	.5	0.0	.5	.5	0.0	.5
EM	0.0	0.0	0.0	0.0	0.0	.5	.5	0.0	.5	.5	0.0	.5	0.0	.5
I	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	.5	.5	0.0	0.0	.5
EX	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
ND	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	.5	.5	.5	0.0	0.0

Un 0.5 signifie indifférent - Exemple : E indifférent à JF

Un 1 signifie préféré - Exemple : E préféré à JM

Un 0 de part et d'autre de la diagonale signifie non comparable - Exemple : JF et MT.

Cette relation de surclassement est à comparer avec les deux préordr
(ν' et ν'') que l'on obtient par la méthode ELECTRE II avec le même jeu de poids :

ν' : E, JF, MC, JM, MF, MP, FP, FA, MT, MAT, EM, I, EX, ND
ν'' : E, JF, MC, MF, MP, FP, FA, JM, MT, MAT, EM, I, EX, ND

6. BMDO5S [1]

BMDO5S est un programme d'analyse hiérarchique unidimensionnel basé sur la technique de L. GUTTMAN.

L'objet du programme est de donner un ordre (préordre) sur des sujet qui ont répondu à un ensemble de questions qui sont supposées se référer une même dimension. Le préordre sur les sujets est alors un classement d ces sujets sur la dimension ainsi construite.

a) Présentation sommaire

p sujets ($k = 1, \ldots, p$) répondent à n questions pouvant chacun avoir de 2 à 7 modalités hiérarchisées.

On part d'un tableau à p lignes et n colonnes. Le programme cherche un nouveau classement des lignes de façon à pouvoir au mieux présenter pour chaque question (colonne) des points de coupures séparant une modalité de la suivante.

Si on a 2 modalités, on a 1 point de coupure.

Si on a 3 modalités, on a 2 points de coupure.

L'ajustement est parfait si on n'a pas d'erreurs, c'est-à-dire de modalités qui se trouvent en-dehors de leur zone.

---

(1) Biomedical Computer Programs - University of California.

Exemple 1

	Q_1	Q_2	Q_3
1	7	7	7
2	7	4	7
3	1	4	7
4	1	4	1
5	1	1	1

ajustement parfait

—— : point de coupure

Exemple 2

	Q_1	Q_2	Q_3
1	7	7	7
2	1	4	7
3	7	4	1
4	1	4	1
5	1	1	1

1 erreur entourée

Les points de coupures établis, on peut alors déterminer un préordre sur les sujets (lignes) en mettant dans une même classe d'équivalence des sujets non séparés par des points de coupures. Exemple :

	Q_1	Q_2	Q_3	rang (préordre sur les sujets)
1	7	7	7	1
2	1	7	7	1
3	7	4	1	2
4	1	4	1	3
5	1	1	1	4

Les rangs du préordre ainsi obtenus s'appellent le score de l'échelle de GUTTMAN (GUTTMAN SCALE SCORE).

Dans le cas où les questions sont dichotomiques, et seulement dans ce cas, on a 1 seul point de coupure par question et il est alors possible d'établir un ordre sur les questions. On en verra un exemple ci-dessous.

b) Présentation des données

Les cartes à lire sont les suivantes :

- Carte Paramètres - Les paramètres à lire sont les suivants :

col. 1-6 : PRØBLM (perforer ces 6 lettres)
col. 7-12 : Titre de l'analyse (6 colonnes)
col. 13 : Un blanc
col. 14-15 : Mois de l'année (1 à 12))
col. 16-17 : Jour du mois (1 à 31)) 6 chiffres facultatifs correspondant à la date
col. 18-19 : L'année (ex. : 73))
col. 20 : Un blanc
col. 21-22 : Nombre de questions (colonnes) (2 chiffres)
col. 23-27 : Nombre de sujets (lignes du tableau ou carte) (4 chiffres)
col. 28-30 : YES (perforer ces 3 lettres) si on désire l'inversion sur un ou plusieurs questions des modalités (inversion du sens des échelles)
col. 31-33 : YES si on veut une édition des données recodifiées
(une question à 2 modalités voit celles-ci recodifiées 7, 1

"	"	3	"	"	"	"	7, 4, 1
"	"	4	"	"	"	"	7, 5, 3, 1
"	"	5	"	"	"	"	7, 6, 4, 2, 1
"	"	6	"	"	"	"	7, 6, 5, 3, 2, 1
"	"	7	"	"	"	"	7, 6, 5, 4, 3, 2,

col. 34-36 : YES si on veut une édition
- du tableau des erreurs (questions x modalités)
- du tableau des fréquences des réponses (questions x modal

col. 37-39 : YES si on veut une édition des sujets avec leur score dans l'échelle de GUTTMAN (le rang du préordre trouvé)
col. 40-42 : YES si on veut une édition du tableau avec les sujets rangés par la technique de GUTTMAN
col. 43-45 : YES si on veut la même édition des sujets avec leur score da l'échelle de GUTTMAN mais présentée dans l'ordre de lecture de ces sujets
col. 46-48 : Un blanc
col. 69-70 : Numéro logique de lecture des données
col. 71-72 : Nombre de cartes format à lire.

- <u>Carte RESPØN</u> (facultative) - Indique le nombre de modalités que l'on doit avoir pour chaque question (permet ainsi de détecter des erreurs).

col. 1-6 : RESPØN (perforer ces 6 lettres)
col. 7-8 : Le nombre de modalités pour la question 1
col. 9-10 : Le nombre de modalités pour la question 2
etc. (2 colonnes par question).

- Carte REFLECT (si on a YES en col. 28-30 de la carte paramètre) - Cette carte dicte les inversions sur les questions
col. 1-6 : REFLECT (perforer ces 6 lettres)
col. 7-8 : 01 si on désire l'inversion sur la 1ère question, blanc ou 00 sinon
col. 9-10 : même chose pour la seconde question
etc. (2 colonnes par question)

- Format de lecture des sujets (des lignes).

- Lecture des sujets - Il faut au moins une carte par sujet. Pour chaque sujet, on lit dans le format indiqué par la carte précédente :

. le nom du sujet (format I) ;
. les réponses aux questions (format F).

c) Présentation des sorties

On obtient les résultats suivants :

- Le tableau des données recodifiées (éventuellement).

- Un contrôle des données, en particulier si le nombre de modalités indiqué dans la carte RESPON est différent du nombre rencontré.

- Une édition du tableau des données où :
. les lignes sont classées par score décroissant (le score est la somme des valeurs d'une ligne) ;
. les cobnnes sont classées par fréquence de 7 croissante.

- Une détermination sur le tableau ci-dessus des points de coupures on édite pour cela un tableau des erreurs :

		questions Q_1 Q_2 .... Q_n
modalités	1 2 ⋮ 7	fréquence des erreurs

A partir de ce tableau, on peut savoir où se trouvent les points de coupure (voir exemple d'utilisation).

- Une édition de la fréquence des réponses :

		modalités 1 2 7
questions	Q_1 ⋮ Q_n	fréquence des réponses

(la comparaison des deux tableaux permet de rejeter éventuellement certaines questions pour un prochain passage du programme).

- Une édition du même tableau mais avec les lignes réordonnées après une phase d'optimisation du nombre d'erreurs (rendu minimum) (les sujets ne sont plus nécessairement classés par score décroissant) :

. édition du nouveau tableau des erreurs ;

. indication sur des combinaisons regroupant certaines modalités (seulement si on en a plus de 2) de façon à améliorer l'ajustement ;

. édition du coefficient de reproductibilité de L. GUTTMAN

$$r = 1 - \frac{\sigma}{n.p}$$

σ : nombre d'erreurs

n : nombre de questions

p : nombre de sujets.

- Edition du tableau (identique au précédent) mais avec le "score de l'échelle de GUTTMAN", c'est-à-dire le classement des sujets en un pré-ordre d'après la position des points de coupure.

- Lorsqu'il existe des non réponses, le programme donne des valeurs à ces non réponses de façon à ne pas introduire de nouvelles erreurs.

d) Exemple d'utilisation

L'exemple est dû à C. RAUZIER [1] [2]. Le problème est de faire apprécier le personnel non cadre suivant certaines dimensions qui sont :

- L'activité : Dans ce facteur se trouvent réunies des situations professionnelles explorant - l'énergie, la vitalité physique - la volonté, la persévérance - l'intérêt pour le travail - l'efficacité.

- La sociabilité : La sympathie qu'on inspire, l'entregent, la bienveillance, la serviabilité, l'adaptation au caractère d'autrui traduisent cette qualité.

- La conscience professionnelle : Il s'agit ici du respect des consignes, de la tendance naturelle à se contrôler, à réaliser un travail fini et soigné.

- La stabilité émotionnelle : Ce trait analyse la tendance à dominer ses nerfs, ses impulsions explosives, son énergie sujette soit à l'éparpillement, soit à l'inhibition ou à l'effondrement devant l'imprévu. La maturité émotionnelle se traduit par un équilibre entre la franchise de l'expression personnelle et la considération qu'on éprouve pour autrui, en un mot par la maîtrise de soi.

---

(1) C. RAUZIER - Contribution à l'étude sur l'appréciation du personnel non-cadre, dans R. BENAYOUN, C. BOULIER : approches rationnelles dans la gestion du personnel (réflexions et expériences).

(2) Nous remercions C. RAUZIER qui nous a fourni les données réelles analysées ci-dessous.

- <u>La présentation</u> : Il s'agit d'apprécier l'attitude générale de l'individu dans sa tenue, son expression orale ou écrite.

- <u>L'éveil intellectuel</u> : Si l'enfant se contente de rêver sa vie, l'adulte s'efforce de vivre ses rêves car son intellect met tout en oeuvre pour construire le réel. Il est fait appel à des notions d'attention, de mémoire, de curiosité, d'intérêt pour la connaissance, d'adaptation, de jugement, d'imagination créatrice.

- <u>L'ascendant</u> : L'on veut mesurer ici l'intensité de cette expansivité que manifeste l'individu à exprimer une opinion, à affirmer sans crainte, à convaincre, à ne pas se soumettre, à diriger un groupe, à se faire obéir.

- <u>L'organisation</u> : Cette qualité s'apprécie à trois niveaux :

. l'ordre que l'on possède dans ses propres affaires ;

. la coordination nécessaire pour conduire simultanément plusieurs tâches ;

. la planification qui prolonge l'action dans le futur.

- <u>L'initiative</u> : Cette disposition caractérielle apparaît lorsque le sujet a besoin d'affronter le travail, va de l'avant, rejette le conformisme, lorsqu'il évite la présence d'autrui, prend la direction d'une organisation.

Chaque ouvrier est apprécié suivant chacune des 9 dimensions par un notateur qui lui-même est un agent de maîtrise mais connaît très bien le personnel qu'il encadre.

Une appréciation à l'aide d'une échelle qualitative à 4 ou 5 modalités a été jugée peu fiable pour mesurer ces appréciations, d'autant plus qu'il s'agit ensuite de comparer des profils d'ouvriers jugés par des notateurs différents.

Aussi a-t-il été décidé d'utiliser la technique d'analyse hiérarchique de GUTTMAN pour classer les sujets suivant chacune des 9 dimensions (9 analyses en tout). Pour chacune des dimensions, C. RAUZIER a proposé une liste de 25 items puis de 20 items après avoir éliminé les 5 items les plus aberrants (ceux qui entraînent un mauvais ajustement au modèle de

L. GUTTMAN). L'exemple analyse ci-dessous concerne la dimension activité.

Les 20 items ou questions proposés sont des questions très simples à deux modalités où le notateur peut répondre beaucoup plus facilement que s'il s'agissait d'apporter une notation unique sur cette dimension du type "Très actif" ou "Moyennement actif".

Les 20 items retenus sont les suivants :

1 Fait son travail quotidien avec la même constance
2 Est sujet à la rêverie
3 On peut le considérer comme un être sans grande vitalité
4 Semble manquer d'énergie pour faire autant de choses que les autres
5 Travaille avec peu d'enthousiasme et d'intérêt
6 Travaille généralement plus lentement que la plupart de ses camarades
7 S'accroche à un travail jusqu'à ce qu'il soit terminé
8 Aborde généralement ses problèmes avec confiance
9 Est suffisamment curieux pour apprendre
10 Cherche à parfaire ses connaissances du métier pour s'améliorer
11 Fait partie de ces gens qui sont toujours engagés dans un travail
12 Exécute sans entrain les travaux qui ne correspondent pas à ses goûts
13 Le responsable doit lui dire parfois "Pressez-vous!"
14 Est capable de persévérer pour réaliser ses projets professionnels
15 Se donne sans réserve à son activité professionnelle
16 Aime être engagé dans un travail qui demande de la rapidité d'exécution
17 Paraît toujours débordant de vigueur
18 Se presse pour entreprendre son travail même s'il en a le temps
19 L'échec partiel le stimule
20 Fournit une plus grosse quantité de travail que la moyenne des gens.

On trouvera les résultats pour 43 sujets pris dans un lot de 250 sujets dans les tableaux ci-après.

Tableau des fréquences des réponses

FREQUENCY OF OCCURRENCE OF SCORES 1 TO 7 AND SCORE ZERO

VARIABLE OR QUESTION	SCORE 1	2	3	4	5	6	7	NO RESPONSE
1	34	0	0	0	0	0	9	0
2	29	0	0	0	0	0	14	0
3	32	0	0	0	0	0	11	0
4	29	0	0	0	0	0	14	0
5	26	0	0	0	0	0	17	0
6	25	0	0	0	0	0	17	0
7	25	0	0	0	0	0	18	0
8	24	0	0	0	0	0	19	0
9	31	0	0	0	0	0	12	0
10	26	0	0	0	0	0	17	0
11	19	0	0	0	0	0	24	0
12	17	0	0	0	0	0	26	0
13	21	0	0	0	0	0	22	0
14	21	0	0	0	0	0	22	0
15	17	0	0	0	0	0	26	0
16	13	0	0	0	0	0	30	0
17	11	0	0	0	0	0	32	0
18	11	0	0	0	0	0	32	0
19	10	0	0	0	0	0	33	0
20	4	0	0	0	0	0	39	0

Il n'y a que deux modalités dans les réponses ; seuls les codes 1 et 7 ont été utilisés.

Tableau des erreurs

ERRORS FOR EACH SCALE SCORE
FINAL STEPS

NUMBER OF RESPONDENTS = 43 NUMBER OF VARIABLES = 20

STEP 4

VARIABLES OR QUESTIONS

SCORE OF	1	3	9	2	4	6	5	10	7	8	13	14	11	12	15	16	17	18	19	20
1	1	2	3	3	1	3	4	2	5	0	2	4	9	3	3	3	3	5	4	1
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	3	2	1	3	5	4	1	0	0	8	1	1	1	0	1	0	2	0	2	0
TOTAL ERROR	4	4	4	6	6	7	5	2	5	8	3	5	10	3	4	3	5	5	6	1

Le total des erreurs est 96, le nombre de réponses 20 x 43. Le coef ficient de reproductibilité de L. GUTTMAN est :

$$r = 1 - \frac{96}{860} = 0{,}888.$$

Le tableau des données est en fin d'analyse le tableau ci-après.

On a représenté pour chaque colonne le point de coupure que l'on retrouve facilement à partir du tableau des erreurs.

<u>Exemple</u> : Question 1 : on a 1 erreur pour la modalité 1, ce qui signifie que un "1" est mal placé. En descendant la colonne, on trouve ce "1" à la ligne du sujet 38, ce qui signifie que le prochain 1 rencontré sera bien situé et déterminera ainsi le point de coupure (entre le sujet 31 et 36). On vérifie aisément que 3 "7" sont alors mal placés, ce qu'in dique bien le tableau des erreurs.

<u>Autre exemple</u> : Question 7 : 5 "1" sont mal placés et 0 "7", par con séquent, en remontant la colonne, le point de coupure se situe au premier "7" rencontré.

Par ailleurs, un point de coupure doit correspondre à un changement dans le "score de l'échelle de GUTTMAN". C'est ce qu'on vérifie avec la question 1. Par exemple, le point de coupure sépare la classe 1 de la classe 2 du préordre de GUTTMAN.

Comme les questions sont dichotomiques (2 modalités), il est possibl de classer les questions.

Le classement des questions a été indiqué à gauche du préordre de GUTTMAN. Le nombre indique le numéro de la question. Ce préordre est :

1, 4, (3, 8), (2, 9), 6, 10, 5, (7, 13), 14, 15, 12, 11, (16, 17), 19, 18, 20.

des questions	SCORE																						
	1	42	1	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
	1	43	2	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
	1	40	3	7	7	7	7	7	7	7	1	7	7	7	7	7	7	7	7	7	7	7	7
1	1	41	4	7	7	7	1	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
	1	38	5	1	7	7	1	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
	1	39	6	7	1	7	7	7	1	7	7	7	7	7	7	7	7	7	7	7	7	7	7
	1	31	7	7	7	1	7	7	7	7	7	7	7	7	1	7	7	7	1	7	1	1	7
	2	36	8	1	1	7	7	1	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
4	2	37	9	1	7	7	7	7	1	1	7	7	7	7	7	7	7	7	7	7	7	7	7
	2	35	10	1	7	1	7	7	1	7	7	1	7	7	7	7	7	7	7	7	7	7	7
3,8	3	25	11	1	7	1	7	1	7	1	1	1	7	7	1	7	1	7	7	7	7	7	7
	4	34	12	7	1	7	7	1	7	7	7	7	1	7	7	1	7	7	7	7	7	7	7
2,9	4	32	13	1	1	7	1	1	7	7	7	7	1	7	7	7	7	7	7	7	7	7	7
	4	29	14	1	1	7	7	1	7	1	7	1	1	7	7	7	7	7	7	7	7	7	7
6	5	33	15	1	7	1	1	7	7	7	7	1	1	7	7	7	7	7	7	7	7	7	7
	5	28	16	1	1	1	1	7	7	1	7	7	7	7	7	1	7	1	7	7	7	7	7
	6	30	17	1	1	1	1	7	1	7	7	7	1	7	7	7	7	7	7	7	7	7	7
10	6	27	18	1	1	1	1	1	1	7	7	7	7	1	7	7	7	7	7	7	7	7	7
	6	23	19	1	1	1	1	7	1	7	7	1	7	1	7	1	7	7	7	7	7	1	7
5	7	26	20	7	1	1	7	1	1	7	1	7	1	7	1	1	7	7	7	7	7	7	7
	8	24	21	1	1	1	1	1	7	1	1	7	7	7	7	1	7	7	7	7	1	7	7
7,13	8	22	22	1	1	7	1	1	1	1	1	7	1	7	7	1	1	7	7	7	7	7	7
	8	21	23	7	1	1	1	1	1	1	1	7	1	7	1	7	7	7	7	1	7	7	7
14	9	20	24	1	1	1	1	1	1	1	1	1	7	1	7	7	7	1	7	7	7	7	7
	9	17	25	1	1	1	7	1	7	1	1	1	7	1	7	1	1	7	7	7	1	1	7
	10	19	26	1	1	1	1	7	1	7	1	1	1	7	1	1	7	7	7	7	7	1	7
15	10	18	27	1	1	1	7	1	1	1	1	1	7	1	1	7	7	1	1	7	7	7	7
	10	13	28	1	7	1	1	1	1	1	1	1	1	1	1	1	7	7	7	1	1	7	7
12	11	16	29	1	1	1	1	1	1	1	1	1	1	1	1	7	7	1	7	7	7	7	7
	12	15	30	1	1	1	1	1	1	1	1	1	7	1	1	7	1	1	7	7	7	7	7
11	12	14	31	1	1	1	1	1	7	1	1	1	1	1	1	7	1	1	7	1	7	7	7
	12	11	32	1	1	1	1	1	1	1	1	1	1	1	1	7	1	1	1	7	7	7	7
16,17	13	9	33	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	7	7	1	7	7
19	14	10	34	1	1	1	1	1	1	1	1	1	1	1	7	1	1	1	1	1	7	7	7
	14	12	35	1	1	1	1	1	7	1	1	1	1	1	1	1	1	7	1	1	7	7	1
18	15	7	36	1	1	1	1	1	1	1	1	1	1	1	1	7	1	1	1	1	7	1	7
	15	8	37	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	7	7	1	7
	16	5	38	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	7	7
20	16	6	39	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	7	1	1	7
	16	3	40	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	7
	17	4	41	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	7	1
	17	1	42	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Le sens des questions aurait dû être inversé. De ce fait, les ordres des questions et des sujets est à inverser pour l'interprétation.

Les questions 20, 18, 19 sont alors caractéristiques d'une activité élevée alors que les questions (3, 8), 4, 1 sont caractéristiques d'une faible activité.

Le classement des sujets est alors le classement inverse. Les sujets 2, 1, 4 (classe 17) sont les plus actifs ; 42, 43, 40, ... classe 1) sont les moins actifs.

Les méthodes d'analyse et de structuration de données ordinales sont extrêmement nombreuses. Certaines sont assez comparables quant à leur fondement mathématique mais il est certain que l'on discerne encore mal les conséquences du choix de telle ou telle distance dans une méthode de visualisation par exemple.

Les modèles eux-mêmes sont assez différents. Pourtant, les résultats d'analyse sur des données réelles sont eux assez comparables mais peut-être est-ce là une chose heureuse.

On peut citer, à titre de conclusion, quelques problèmes qui ne sont peut-être pas très bien abordés pour l'ensemble de ces méthodes et qui, à notre avis, mériteraient des approfondissements.

1. Le problème multijuge-multicritère

On l'a présenté dans le premier chapitre à propos du concours d'architecture nouvelle.

On possède un ensemble de n objets comparés par p juges suivant m dimensions ou critères.

Les données sont donc des données cubiques. s_{it}^{k} : évaluation par le juge k de l'objet i suivant le critère t.

Si on veut analyser, décrire de telles données, il nous faut actuellement faire :

- une analyse multijuge critère par critère (m analyses) ;
- une analyse multicritère juge par juge (p analyses) ;
- ou bien une analyse commune mais en agrégeant l'une des dimensions.

Dans certaines études de marché, on recueille souvent, en plus de données de préférences suivant différents critères, une préférence globale. Se pose alors le problème d'"expliquer" la préférence globale par les préférences suivant les différents critères.

Dans l'aide à la décision de jury pour des concours, il semble également intéressant de s'interroger sur la différence entre politique de choix et compétence d'évaluation, c'est-à-dire entre décideur et expert. Autrement dit, une politique en matière de choix peut être définie par une préférence pouvant s'exprimer par exemple par une pondération de différents critères.

Une compétence en matière d'évaluation est une aptitude reconnue au juge k pour évaluer ou comparer des objets sur la dimension k. Cette compétence pourrait être traduite par exemple par une pondération des divers juges relativement à un critère donné.

Développer de tels outils permettra peut-être de développer des méthodes de choix moins technocratiques où les experts auraient leur mot à dire sans que, sous le couvert d'expert, ils fassent en fait des choix politiques. Que le choix politique se fasse lui par des hommes (élus ou non) ou par un processus plus démocratique est là une autre question (entièrement politique cette fois).

2. Agréger ou ne pas agréger ?

Dans le premier chapitre, nous avons présenté la différence entre données ordinales et données de proximité.

Il existe actuellement des méthodes pour traiter soit les unes, soit les autres, soit les deux à condition d'avoir recueilli des données de proximité et des données de préférence.

Avec G. RIBEILL, nous pensons qu'il serait intéressant de développer des méthodes permettant d'analyser et de structurer des données de comparaisons entre objets, la relation de comparaison pouvant être une relation d'ordre, d'indifférence ou de différence. Autrement dit, devant deux objets présentés à une personne, cette personne les voit-elle en terme de hiérarchie ou en terme de simple différence, sans porter de jugement de valeur sur cette différence ? Il y aurait en sociologie un développement certainement intéressant à faire sur ce type de perception.

Les méthodes d'analyse et de structuration de données ordinales sont extrêmement nombreuses. Certaines sont assez comparables quant à leur fondement mathématique mais il est certain que l'on discerne encore mal les conséquences du choix de telle ou telle distance dans une méthode de visualisation par exemple.

Les modèles eux-mêmes sont assez différents. Pourtant, les résultats d'analyse sur des données réelles sont eux assez comparables mais peut-être est-ce là une chose heureuse.

On peut citer, à titre de conclusion, quelques problèmes qui ne sont peut-être pas très bien abordés pour l'ensemble de ces méthodes et qui, à notre avis, mériteraient des approfondissements.

1. Le problème multijuge-multicritère

On l'a présenté dans le premier chapitre à propos du concours d'architecture nouvelle.

On possède un ensemble de n objets comparés par p juges suivant m dimensions ou critères.

Les données sont donc des données cubiques. s_{it}^{k} : évaluation par le juge k de l'objet i suivant le critère t.

Si on veut analyser, décrire de telles données, il nous faut actuellement faire :

- une analyse multijuge critère par critère (m analyses) ;
- une analyse multicritère juge par juge (p analyses) ;
- ou bien une analyse commune mais en agrégeant l'une des dimensions.

Dans certaines études de marché, on recueille souvent, en plus de données de préférences suivant différents critères, une préférence globale. Se pose alors le problème d'"expliquer" la préférence globale par les préférences suivant les différents critères.

Dans l'aide à la décision de jury pour des concours, il semble également intéressant de s'interroger sur la différence entre politique de choix et compétence d'évaluation, c'est-à-dire entre décideur et expert. Autrement dit, une politique en matière de choix peut être définie par une préférence pouvant s'exprimer par exemple par une pondération de différents critères.

Une compétence en matière d'évaluation est une aptitude reconnue au juge k pour évaluer ou comparer des objets sur la dimension k. Cette compétence pourrait être traduite par exemple par une pondération des divers juges relativement à un critère donné.

Développer de tels outils permettra peut-être de développer des méthodes de choix moins technocratiques où les experts auraient leur mot à dire sans que, sous le couvert d'expert, ils fassent en fait des choix politiques. Que le choix politique se fasse lui par des hommes (élus ou non) ou par un processus plus démocratique est là une autre question (entièrement politique cette fois).

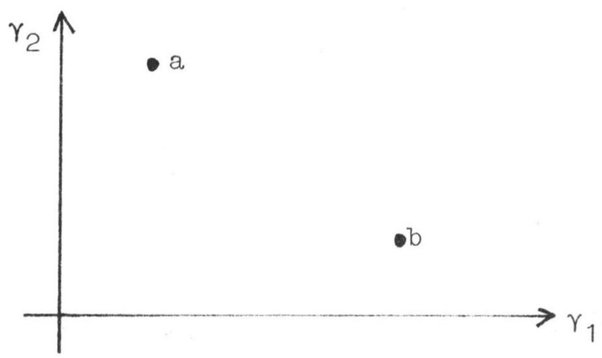
2. Agréger ou ne pas agréger ?

Dans le premier chapitre, nous avons présenté la différence entre données ordinales et données de proximité.

Il existe actuellement des méthodes pour traiter soit les unes, soit les autres, soit les deux à condition d'avoir recueilli des données de proximité et des données de préférence.

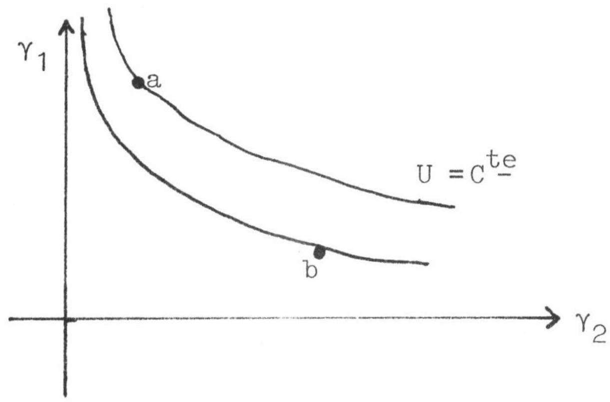
Avec G. RIBEILL, nous pensons qu'il serait intéressant de développer des méthodes permettant d'analyser et de structurer des données de comparaisons entre objets, la relation de comparaison pouvant être une relation d'ordre, d'indifférence ou de différence. Autrement dit, devant deux objets présentés à une personne, cette personne les voit-elle en terme de hiérarchie ou en terme de simple différence, sans porter de jugement de valeur sur cette différence ? Il y aurait en sociologie un développement certainement intéressant à faire sur ce type de perception.

Exemple : Soit deux objets supposés différentiés sur deux dimensions 1 et 2.



Je peux très bien, si on me demande de comparer a et b, affirmer la relation "a est différent de b", refusant par là même toute agrégation en un critère unique des deux dimensions sous-jacentes. Je peux également affirmer une relation "je préfère a à b" ou "a est mieux que b" à partir du moment où j'exprime une préférence sur les critères 1 et 2.

Mais cela suppose un jugement de valeur sur les deux dimensions, une agrégation des préférences au niveau individuel.



Dans quelles mesures les gens (les groupes sociaux) ont-ils une perception unidimensionnelle des choses en acceptant de les agréger en une relation de préférence, de les exprimer en une échelle de valeur unique, monétaire le plus souvent et dans quelles mesures en conservent-ils une vision multidimensionnelle, refusant de les agréger en une relation de préférence unique, en une échelle de valeur simple, bref refusant la possibilité de l'échange dans n'importe quelle limite "à condition d'y mettre le prix" ? Voilà une question qui semble importante et pour laquelle il serait utile de développer des outils d'expression et d'analyse.

3. L'analyse et la structuration de données ordinales : un instrument de dialogue et d'évolution des préférences

Il semble que ces outils puissent être utilisés de façon féconde dans les assemblées et en particulier dans des problèmes d'aide à la décision collective.

Loin d'être seulement des appareils à photographier les préférences, ils peuvent devenir des outils de dialogue, d'information et d'échange de points de vue, bref ils peuvent être des outils facilitant une évolution des préférences et peuvent être utilisés dans le champ du calcul négotiationnel. Il y a là à développer tout un domaine de la théorie des jeux applicable à des situations réelles où l'on s'intéresse à l'évolution des préférences de divers agents dans un processus décisionnel.

ANNEXE I

NOTIONS MATHEMATIQUES SUR LES RELATIONS

1. Relations binaires : Généralités

1.1 Définition

Soit E et E' deux ensembles distincts ou non. On a défini une relation R si on a précisé ce qu'est un lien entre certains couples de E x E' (B. ROY [79]).

Si E = E', on dit que la relation R est définie sur E. Dans toute la suite, on ne considérera que des relations définies sur E.

1.2 Graphe associé à une relation binaire

Soit G(X, U) un graphe où X est l'ensemble des sommets, U l'ensemble des lignes joignant les sommets.

La ligne est un arc si elle est orientée. C'est une arête si elle n'est pas orientée.

Un graphe orienté ne comprend que des arcs. Un graphe non orienté ne comprend que des arêtes.

A toute relation R définie sur un ensemble X d'éléments, on peut associer un graphe G(X, U) tel qu'il existe une ligne entre x et y si et seulement si x R y.

Exemple :

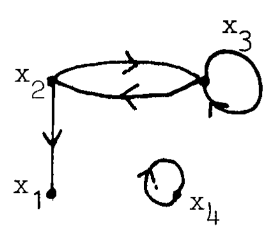


Figure 1

Le graphe de la figure 1 peut représenter la relation binaire R définie par :

$x_2 \; R \; x_1$ $\quad x_3 \; R \; x_2$ $\quad x_4 \; R \; x_4$
$x_2 \; R \; x_3$ $\quad x_3 \; R \; x_3$

1.3 Tableaux associés à un graphe ou à une relation binaire

On peut associer au graphe son dictionnaire : pour chaque sommet, on indique l'ensemble $\Gamma(x)$ des suivants de x.

Le dictionnaire du graphe de la figure 1 est :

x	$\Gamma(x)$
x_1	
x_2	x_1, x_3
x_3	x_2, x_3
x_4	x_4

sa matrice booléenne : soit $\{a_{ij}\}$ la matrice booléenne :

$$a_{ij} = 1 \qquad x_i \, R \, x_j \qquad (x_i, x_j) \in U$$
$$a_{ij} = 0 \qquad \text{non } x_i \, R \, x_j \qquad (x_i, x_j) \notin U$$

Exemple : La matrice booléenne du graphe de la figure 1 est la suivant

i \ j	1	2	3	4
1	0	0	0	0
2	1	0	1	0
3	0	1	1	0
4	0	0	0	1

1.4 Propriétés mathématiques des relations binaires

Pour caractériser diverses relations binaires et surtout pour les rendre plus intéressantes ou opérationnelles, on les caractérise par les propriétés suivantes :

- Réflexivité : $x \, R \, x, \forall x$

Exemples : - frère de, peu différent de, plus grand ou égal sont réflexives

- père de n'est pas réflexive.

- Symétrie : $x \, R \, y \Longrightarrow y \, R \, x \quad \forall (x, y)$

Exemples : - frère de, peu différent de sont symétriques

- père de, plus grand ou égal ne sont pas symétriques.

- Antisymétrie : $x \, R \, y \Longrightarrow$ non $y \, R \, x \; (x \neq y)$

Exemple : "plus grand".

- Transitivité : $x \, R \, y$ et $y \, R \, z \Longrightarrow x \, R \, z \quad (\forall \, x, y, z)$

Exemples : - frère de, plus grand ou égal sont transitives

- père de, peu différent de ne sont pas transitives

- domine, préféré à sont "normalement" considérées comme transitives mais la réalité montre que ces relations ne sont pas toujours transitives.

- Relation complète ou totale : $x \, R \, y$, $y \, R \, x$ ou $x \, R \, y$ et $y \, R \, x$ $\forall (x, y)$

Autrement dit, toute paire d'éléments x, y est comparable par R.

- Relation partielle : il existe des éléments non comparés par R

Lorsqu'une relation n'est pas complète ou totale, elle est partielle.

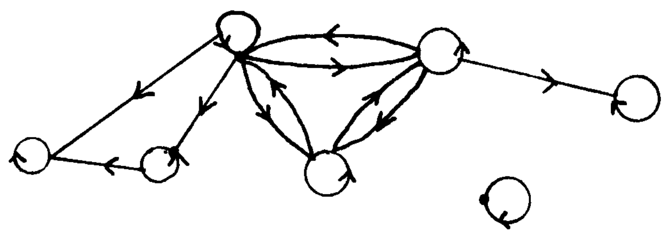
2. Relations binaires transitives

Certaines relations binaires caractérisées par des propriétés définies ci-dessus ont été étudiées plus particulièrement en raison de leur intérêt théorique ou pratique.

2.1 Préordres

Un préordre est une relation réflexive et transitive.

Exemple :



2.2 Equivalences

Une relation d'équivalence en équivalence est un préordre symétrique, c'est-à-dire une relation binaire réflexive, symétrique et transitive.

On appelle classe d'équivalence l'ensemble des éléments tels que $x \; R \; y$

$x \in C$ et $y \in C \Longrightarrow x \; R \; y$

$x \in C$ et $y \in C \Longrightarrow$ non $x \; R \; y$.

L'ensemble des classes d'équivalence est l'ensemble quotient $(E \mid R)$.

Les classes d'équivalence constituent une partition de E.

Exemple :

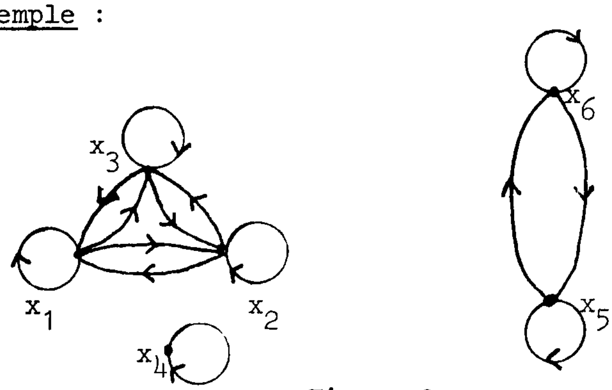


Figure 2

La matrice associée à une équivalence à k classes peut toujours se mettre sous la forme de k carrés de 1 le long de la diagonale.

	1	2	3	4	5	6
1	1	1	1	0	0	0
2	1	1	1	0	0	0
3	1	1	1	0	0	0
4	0	0	0	1	0	0
5	0	0	0	0	1	1
6	0	0	0	0	1	1

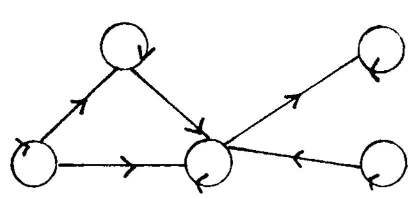
Une <u>clique</u> est une équivalence complète ou totale ou encore une équivalence à une classe.

La matrice booléenne associée ne contient que des 1.

2.3 <u>Ordres</u>

Un ordre est un <u>préordre antisymétrique</u>, c'est-à-dire une relation binaire <u>réflexive, antisymétrique et transitive</u>.

<u>Exemple</u> :



Le graphe associé à un ordre est sans circuit (autre que les boucles).

2.3.1 <u>Eléments particuliers</u>

- <u>Majorant</u> (notion duale : minorant)

x est majorant de y $\Longleftrightarrow$ x R y.

- <u>Majorant universel</u> (ou plus grand élément)

x est majorant universel de y $\Longleftrightarrow$ x R y, $\forall y \in E$.

- Intermédiaire

t est intermédiaire de x et y $\Longleftrightarrow$ x R t R y.

- Prédécesseur (notion duale : successeur)

y précède x si x R y et s'il n'y a pas d'intermédiaire entre x et y.

2.3.2 Ordres particuliers

- Ordre strict

Un ordre est strict s'il n'est pas réflexif. Le graphe associé ne co tient pas de boucles et la matrice booléenne associée ne contient pas de dans la diagonale.

- Ordre complet ou total

C'est un ordre où tous les éléments sont comparables. La matrice associée à un ordre total peut se représenter avec des 1 au-dessus de la diagonale et des 0 en-dessus en prenant l'ordre des lignes et des colonnes identique à l'ordre que l'on représente.

Exemple :

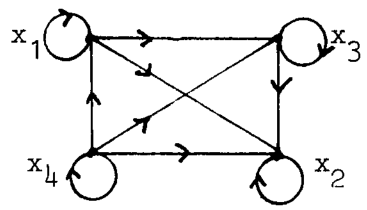


Figure 3

L'ordre x_4, x_1, x_3, x_2 peut se représenter par :

	4	1	3	2
4	1	1	1	1
1	0	1	1	1
3	0	0	1	1
2	0	0	0	1

On appelle permutation une suite ordonnée de n éléments. Cette suite ordonnée constitue un ordre total ou complet sur les n éléments. Une permutation sera donc un ordre total que l'on désignera généralement par une suite de lettres ou de chiffres (les identificateurs des éléments).

Exemple : $x_4\ x_1\ x_3\ x_2$ ou 4 1 3 2 est une permutation représentant l'ordre total de la figure 3.

Il y a $n!$ ordres totaux distincts sur un ensemble E à n éléments.

- Demi-treillis

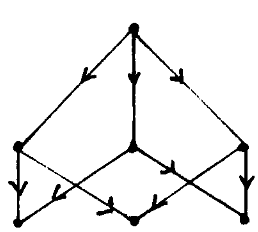
C'est un ordre partiel tel que tout couple (x, y) admette :

. un plus petit majorant commun (supremum : $x \vee y$) : c'est un sup-demi-treillis ;

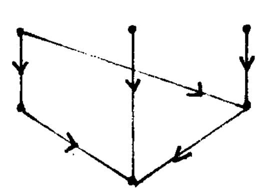
. un plus grand minorant commun (infimum : $x \wedge y$) : c'est un inf-demi-treillis.

Un treillis est à la fois un sup et inf-demi-treillis.

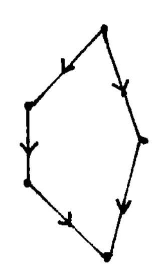
Exemples :



sup-demi-treillis

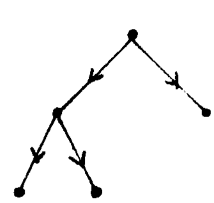


inf-demi-treillis



treillis

Un arbre est un sup-demi-treillis dont chaque élément n'a qu'un successeur, sauf la racine qui n'en n'a pas.



- Le treillis des parties d'un ensemble : le simplexe

On sait qu'il existe 2^n parties d'un ensemble à n éléments car, pour construire une partie, on a relativement à chaque élément 2 choix possibles : l'éliminer ou le retenir.

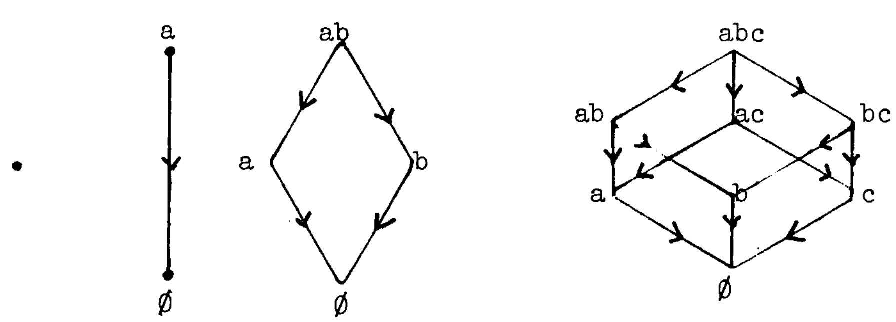
De plus, il existe C_n^k parties à k éléments (C_n^k : nombre de combinaisons à k éléments pris parmi n). Le nombre total de parties est alors :

$$C_n^o + C_n^1 + \dots + C_n^k + \dots + C_n^n = (1 + 1)^n = 2^n.$$

L'ensemble des parties d'un ensemble est muni d'une structure d'ordre partiel par la relation d'inclusion. Une partie P_1 domine une partie P_2 si elle contient tous les éléments de P_2.

L'ordre partiel obtenu est un treillis dont le supremum est la partie constituée par l'ensemble des éléments et l'infimum est la partie ensemble vide.

Exemples :



S_o S_1 S_2 S_3

2.3.3 Préordres complets

Un préordre complet est une relation binaire réflexive, transitive et complète.

<u>Exemple</u> :

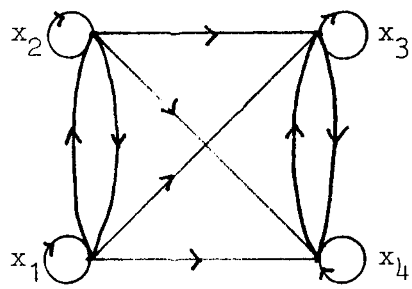


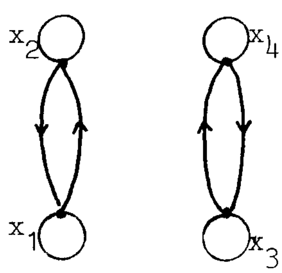
Figure 4

Un préordre R peut être décomposé en deux relations, l'une symétrique E, l'autre antisymétrique P

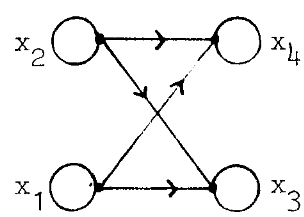
$$R = E \oplus P.$$

Dans l'exemple de la figure 4

E :



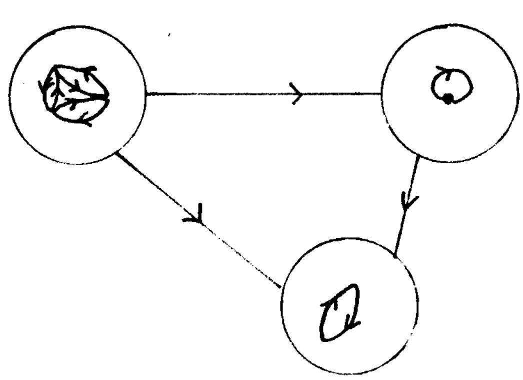
et P :



Les classes d'équivalence sont alors ordonnées par P.

Lorsque le préordre est complet et comprend k classes, il peut se décomposer en une équivalence contenant k classes et un ordre complet sur les classes d'équivalence.

<u>Exemple</u> :



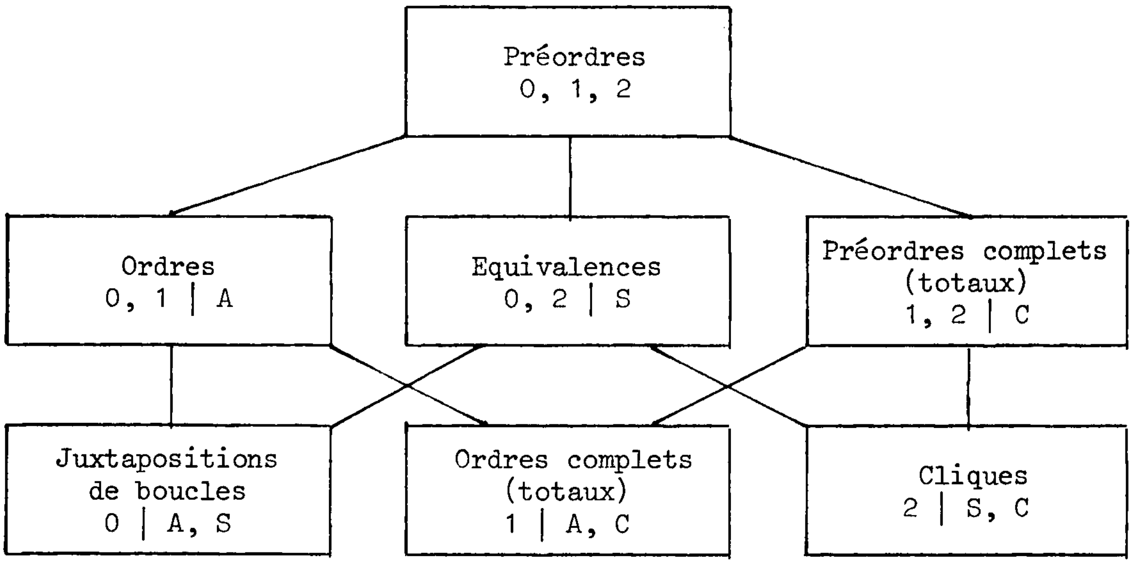
Un préordre complet défini sur n éléments est une <u>clique</u> s'il ne possède qu'une seule classe d'équivalence à n éléments et c'est un ordre complet s'il possède n classes d'équivalence à 1 élément.

2.4 Organisation des relations binaires transitives

Les différentes relations binaires transitives présentées peuvent se représenter selon le schéma suivant.

Le chiffre 0, 1, 2 indique le nombre d'arcs possible entre les pair et les lettres signifient les propriétés suivantes que possèdent les relations :

A : antisymétrique
S : symétrique
C : complète.



Remarques sur le schéma :

- Une juxtaposition de boucles est la seule relation binaire qui soit à la fois symétrique et antisymétrique. Exemple : x_1 x_2 x_3

- Toutes les relations binaires représentées sont transitives et réflexives. Certaines peuvent cependant ne pas être réflexives (cf. les ordres stricts). En particulier, la juxtaposition de boucles devient le graphe vide ou la relation vide. Si on supprime la réflexivité, la matrice associée ne comprend que des 0.

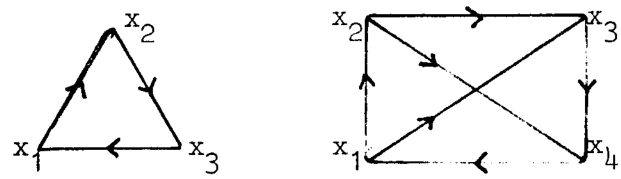
3. Relations binaires non transitives

Certaines relations binaires non transitives jouent un rôle particulier dans la pratique et suscitent des études mathématiques particulières.

3.1 Les tournois

Un tournoi est une relation binaire antisymétrique complète.

Exemple :



Matriciellement, la condition pour qu'une matrice booléenne soit celle d'un tournoi est que :

$$\forall\, i \neq j \qquad a_{ij} + a_{ji} = 1.$$

Un tel graphe comprend $n(n - 1)/2$ arcs si G possède n sommets (E n éléments).

Les tournois ont fait l'objet d'un développement mathématique poussé (cf. MOON [66] et plus récemment BERMOND [12]).

Citons quelques résultats et problèmes principaux.

- Le nombre de tournois à n sommets est $2^{n(n-1)/2}$.

- Un tournoi transitif est un ordre total (par définition : une relation complète, transitive et antisymétrique).

- Un tournoi sans circuit d'ordre 3 est un ordre total. En effet, on montre facilement qu'un tournoi sans circuit de longueur 3 ne peut contenir de circuit de longeur p $(3 \leq p \leq n)$.

Deux problèmes sont posés à propos des tournois :

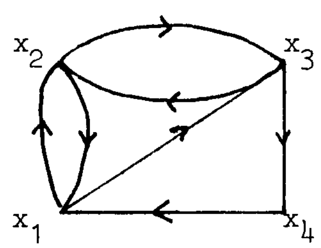
Pb. 1 - Etant donné un tournoi quelconque, quel est le nombre minimum d'arcs dont il faut inverser le sens pour le rendre transitif ? Divers algorithmes ont été proposés [84], [12], [72], [50] et mentionnés dans le deuxième chapitre (§ I, 3.2.2) dans les méthodes de recherche d'un ordre.

Pb. 2 - Quels sont les tournois à n sommets les plus "incohérents" ou "embrouillés", c'est-à-dire les plus "distants" de l'ordre total ? Nous renvoyons à J.C. BERMOND [12] pour une présentation récente du problème.

3.2 Les matchs

Le terme a été proposé par G. RIBEILL [75]. C'est une relation complète (réflexive ou non).

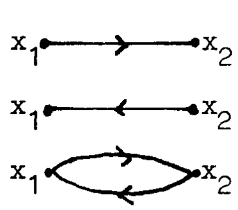
Exemple de match :



Un match est donc un tournoi qui permet les ex-aequo. C'est le graphe (ou la relation de dominance) que l'on obtient à l'issue d'un jeu où le match nul est possible et à l'issue de procédures d'agrégation à seuil (deuxième chapitre, § I, 3.6).

Le problème que l'on peut se poser est également l'approximation d'un match par une relation transitive (préordre complet).

Le nombre de matchs existant entre n sommets est $3^{n(n-1)/2}$. En effet, sur chaque paire de sommets (x_1, x_2), on a trois possibilités :



Ces graphes, tels qu'ils ont été introduits, sont complets (comme les tournois).

3.3 Les quasi-ordres

La relation de quasi-ordre a été introduite par R.D. LUCE [59] pour représenter une relation de préférence (P, I) qui soit transitive pour la relation de préférence stricte P et non transitive pour la relation d'indifférence I.

Par définition, (P, I) est un quasi-ordre sur un ensemble A de n objets si les 4 axiomes suivants sont vérifiés :

S1 : $\forall\, x_1, x_2 \in A$, on a une et une seule des relations $x_1\ P\ x_2$, $x_2\ P\ x_1$, $x_1\ I\ x_2$ (complète comparabilité). (P, I) est une relation complète.

S2 : $x_1\ I\ x_1$ (réflexivité).

S3 : $\forall\, x_1, x_2, x_3, x_4 \in A$

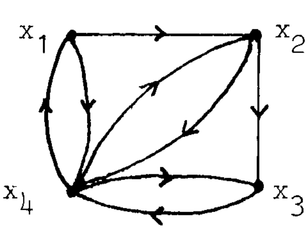
$$\left.\begin{array}{l} x_1\ P\ x_2 \\ x_2\ I\ x_3 \\ x_3\ P\ x_4 \end{array}\right\} \Rightarrow x_1\ P\ x_4$$

généralisation de la transitivité entraîne la transitivité de P avec la condition de réflexivité de I.

S4 : $\forall\, x_1, x_2, x_3, x_4 \in A$

$$\left\{\begin{array}{l} x_1\ P\ x_2,\ x_2\ P\ x_3 \\ x_4\ I\ x_1,\ I_4\ I\ x_2,\ I_4\ I\ x_3 \end{array}\right. \qquad \text{est interdite.}$$

On interdit la configuration :



Autrement dit, la relation

$$\begin{cases} x_1 \ P \ x_2 \\ x_4 \ I \ x_1, \ I_4 \ I \ x_2 \end{cases}$$

représente une relation : x_4 est intermédiaire entre x_1 et x_2.

Si $x_1 \ P \ x_2$ et $x_2 \ P \ x_3$, alors cet axiome nous dit que x_4 ne peut être à la fois intermédiaire entre x_1 et x_2 et intermédiaire entre x_2 et x_3.

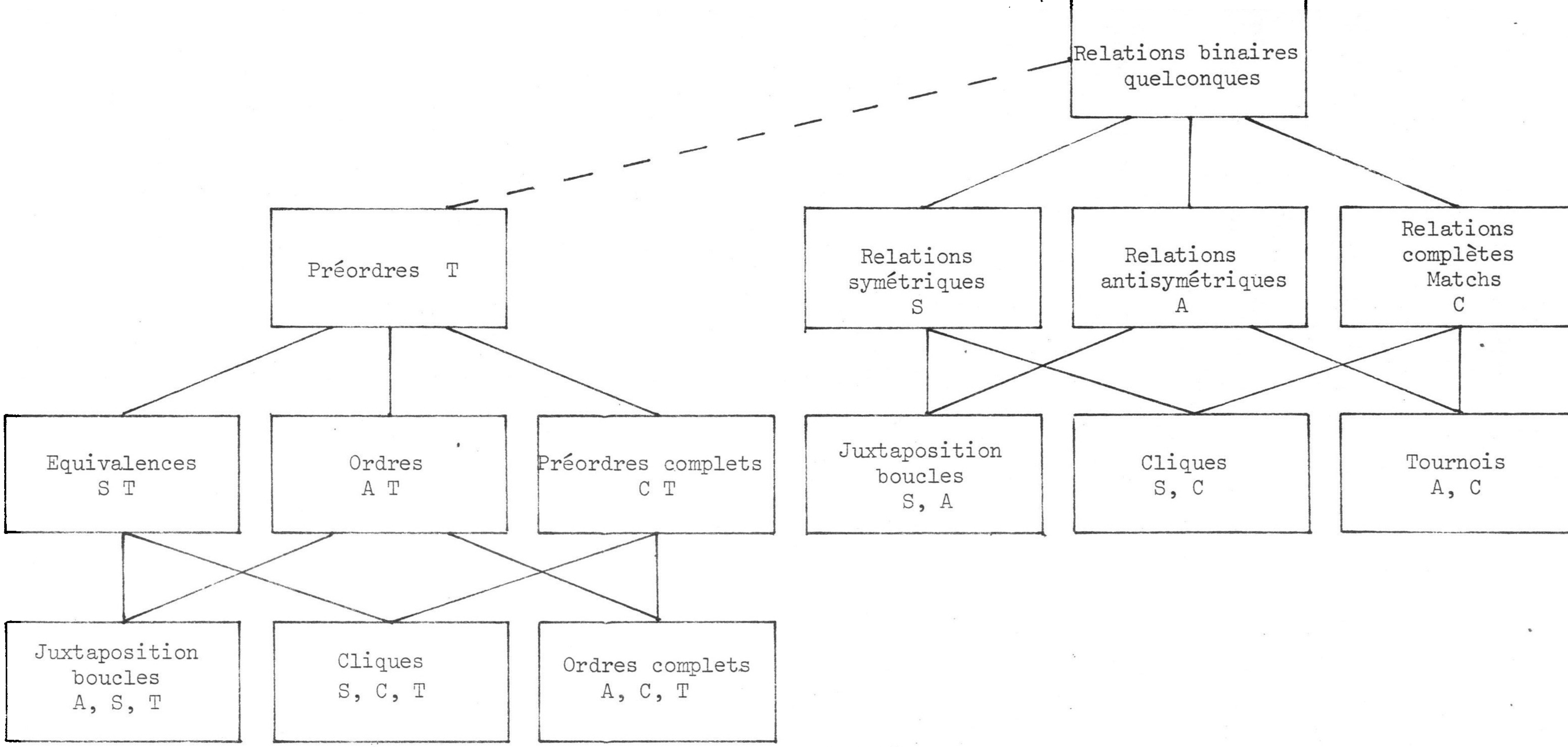
Pour des développements sur les notions de quasi-ordres, on se reportera à R.D. LUCE [59] et ROBERTS [77].

3.4 <u>Autres relations binaires et organisation des relations binaires</u>

D'autres relations pourraient être étudiées, notamment :

- les relations antisymétriques (A) qui comprennent :
  . les ordres (A, T) : il y en a $n!$,
  . les tournois (A, C) : il y en a $2^{n(n-1)/2}$;
- les relations symétriques (S) (non réflexives) : il y en a $2^{n(n-1}$

On a présenté l'ensemble des relations binaires sur le schéma suivant qui est à rapprocher du simplexe aux 4 éléments (T, A, S, C) :



T : Transitive

A : Antisymétrique

S : Symétrique

C : Complète ou totale

<u>Remarques</u> : $\boxed{\text{A, S, T}}$ et $\boxed{\text{S, A}}$ d'une part

$\boxed{\text{S C T}}$ et $\boxed{\text{S C}}$ d'autre part

sont équivalents.

4. Relations binaires valuées, relations binaires floues

4.1 Définition

Une relation binaire est valuée si le graphe associé à la relation binaire est lui-même valué.

La matrice associée au graphe n'est plus booléenne mais les termes de la matrice sont des membres réels quelconques.

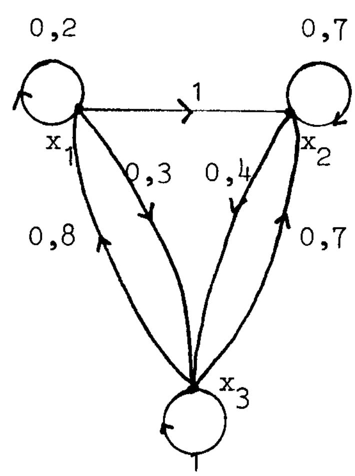
On obtient des relations de comparaisons valuées en agrégeant par exemple des relations binaires non valuées.

Dans le modèle ANAPREF, on a été conduit par exemple à introduire des relations de comparaisons valuées.

Le concept de relation floue est récent et se situe dans le contexte de la théorie des ensembles flous introduite par L.A. ZADEH (cf. A. KAUFMA [53]). Formellement, une relation floue et une relation valuée ne sont qu'une seule et même chose (la valuation d'un arc est cependant un nombre réel compris entre 0 et 1. Ce nombre $\mu(x, y)$ est un degré d'intensité correspondant à la relation.

Exemple :

	x_1	x_2	x_3
x_1	0,2	1	0,3
x_2	0	0,7	0,4
x_3	0,8	0,7	1



4.2 Propriétés des relations floues

On définit entre autres propriétés :

- la symétrie

$\forall\ (x_1, x_2) \in A\ \mu(x_1, x_2) = \mu \Longrightarrow \mu(x_2, x_1) = \mu$

- la réflexivité

$\forall\ x_1 \in A\ \mu(x_1, x_1) = 1$

- la transitivité

$\forall\ x_1, x_2, x_3 \in A$

$$\mu(x_1, x_3) \geq \underset{x_2}{\mathrm{MAX}} \left[\mathrm{MIN}(\mu(x_1, x_2), \mu(x_2, x_3))\right].$$

Cette notion généralise la transitivité établie sur les relations non floues où $\mu(x_1, x_2) \in \{0, 1\}$ au lieu de $[0, 1]$.

En effet, si il existe x_2 tel que $x_1\ R\ x_2$ et $x_2\ R\ x_3$, alors $\mu(x_1, x_2) = 1$, $\mu(x_2, x_3) = 1$. Le plus petit des deux est la valeur 1, donc $\mu(x_1, x_3) = 1$, ce qui signifie $x_1\ R\ x_3$.

Exemple (KAUFMAN [53], p. 80) : La relation ci-dessous est transitive :

	x_1	x_2	x_3	x_4
x_1	0,2	1	0,4	0,4
x_2	0	0,6	0,3	0
x_3	0	1	0,3	0
x_4	0,1	1	1	0,1

Une relation de préordre flou est une relation

- réflexive ;
- transitive.

Exemple : La relation ci-dessus avec des 1 dans la diagonale du tableau est une relation de préordre floue.

Une relation d'ordre flou est une relation floue

- réflexive ;
- transitive ;
- antisymétrique $(\mu(x_1, x_2) \neq \mu(x_2, x_1)$ ou $\mu(x_1, x_2) = \mu(x_2, x_1) = 0)$.

Exemple : La relation représentée ci-dessus est une relation d'ordre flou.

Graphe vulgaire associé à une relation floue antisymétrique : c'est le graphe $G(x_1, x_2)$ tel que :

$$\mu(x_1, x_2) > \mu(x_2, x_1) \Longrightarrow (x_1, x_2) \in G \text{ et } (x_2, x_1) \notin G$$
$$\mu(x_1, x_2) = \mu(x_2, x_1) = 0 \Longrightarrow (x_1, x_2) \notin G \text{ et } (x_2, x_1) \in G.$$

Tout graphe associé à une relation d'ordre flou est un ordre.

4.3 Relations vulgaires associées aux relations floues

Il serait intéressant d'étudier plus en détail les relations existant entre des relations floues et les relations binaires que l'on peut associer à ces relations floues, les relations vulgaires ayant l'avantage de les rendre plus "lisibles".

En particulier, des relations de comparaisons par paires (tournois) agrégées se présentent sous la forme de relation floue (A_{ij}) où $A_{ij} + A_{ji} \leq 1$.

Il serait intéressant de relier les notions de transitivité floue à celles de consistance en probabilité que l'on trouve dans F.J. ROBERTS [77].

Il serait également intéressant d'étudier dans quelle mesure on peut associer des relations binaires qui soient des quasi-ordres.

<u>Exemple de relation binaire associée</u> : Soit R_α la relation binaire associée à une relation floue $R_\alpha = (I, P)_\alpha$ où I est une relation symétrique, P une relation antisymétrique et α un seuil. R_α est défini de la façon suivante :

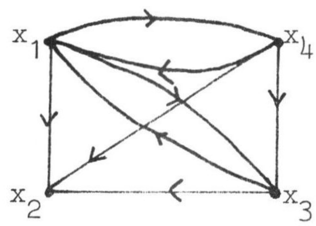
$$x_1 \; P \; x_2 \Longleftrightarrow \mu(x_1, x_2) > \alpha \quad \text{et} \quad \mu(x_2, x_1) \leq \alpha$$
$$x_1 \; I \; x_2 \Longleftrightarrow \mu(x_1, x_2) \leq \alpha \quad \text{et} \quad \mu(x_2, x_1) \leq \alpha.$$

La relation $R_\alpha = (I, P)_\alpha$ est-elle par exemple un quasi-ordre si la relation floue est transitive ?

<u>Exemple</u> : La relation d'ordre floue présentée ci-dessus s'écrit :

	x_1	x_4	x_3	x_2
x_1	1	0,4	0,4	1
x_4	0,1	1	1	1
x_3	0	0	1	1
x_2	0	0	0,3	1

La relation $R_{0,4} = (I, P)_{0,4}$ est la suivante :



La relation $R_{0,4}$ est ici un quasi-ordre.

ANNEXE II

PROBLEMES LIES AU RECUEIL DE DONNEES ORDINALES

On a présenté dans le premier chapitre trois types de données de classements :

- les rangs r_i^k : rang de l'objet i suivant le classement k ;
- les évaluations sur des échelles : s_i^k : évaluation de l'objet i suivant le classement k ;
- les comparaisons par paires : a_{ij}^k : avis sur la paire (ij) suivant le classement k.

1. <u>Problèmes liés au passage d'une forme à une autre dans l'utilisation des programmes</u>

On peut sans problème transformer des évaluations sur des échelles en des rangs. En effet, l'indicateur s_i^k donne à l'ensemble des objets une structure de préordre auquel on peut associer des rangs.

On peut également, dans certaines méthodes, transformer des rangs en notes en prenant, par exemple, le complément à $n + 1$ des rangs : $s_i^k = (n + 1) - r_i^k$ bien que cette transformation soit sans objet si on traite des données ordinales par des méthodes ordinales.

On peut passer des données de rangs ou d'évaluation sur des échelles à des données de comparaisons par paires.

<u>Exemple</u> : $\left\{\begin{array}{l} a_{ij}^k = 1 \\ a_{ji}^k = 0 \end{array}\right. \Longleftarrow r_i^k < r_j^k \Longleftrightarrow s_i^k > s_j^k$

$$\left\{\begin{array}{l} a_{ij}^k = 1/2 \\ a_{ji}^k = 1/2 \end{array}\right. \Longleftarrow r_i^k = r_j^k \Longleftrightarrow s_i^k = s_j^k$$

Le passage inverse est plus délicat et se fait le plus souvent avec une perte d'information.

<u>Exemple</u> : Soit le classement suivant obtenu par comparaisons par paires :

	x_1	x_2	x_3	x_4	score
x_1	0	1	1	0	2
x_2	0	0	1	1	2
x_3	0	0	0	1	1
x_4	1	0	0	0	1

Si on calcule l'ordre à distance minimum (cf. deuxième chapitre, § I, 3.4), on obtient l'ordre $x_1 \succ x_2 \succ x_3 \succ x_4$. Cette méthode peut être uti- lisée dans ANAPREF ou AGREPREF. On en trouvera un exemple au paragraphe : ci-dessous.

Si on calcule les "scores" de chaque objet, c'est-à-dire la somme des demi-degrés extérieurs, on obtient le préordre : $(x_1, x_2) \succ (x_3, x_4)$.

Ce passage devrait être réalisé si on voulait par exemple analyser des comparaisons par paires avec les programmes FENELON ou PREFMAP. Par contre il n'a pas lieu d'être avec des programme comme MDPREF ou ANAPREF.

2. Problèmes pratiques posés par le recueil de données

Demander une note, donc une évaluation sur une échelle, est peut-être le plus simple. Il est certain que nous sommes d'avantage habitués à répondre à ce genre de question mais quelle est la validité psychologique de ces données pour en déduire des comparaisons ordinales ?

Recueillir des rangs semble plus satisfaisant si l'on veut faire un traitement d'analyse de préférences ordinales. Cependant, classer des objets n'est pas toujours facile. Aussi peut-on conseiller de déterminer dans un premier temps une dichotomie des objets :

- les objets les plus préférés ;
- les objets les moins préférés ;
- les autres.

Dans un second temps, on demande de classer les objets au sein de deux premiers groupes puis de prendre chaque objet du 3e groupe et de l'insérer dans l'ordre déjà obtenu à l'aide des deux premiers groupes.

Pour ce genre de recueil de données ordinales, des petits cartons sur lesquels sont représentés chaque objet (nom, dessins, principales caractéristiques, etc.) aident beaucoup : l'interviewé passe plus de temps à répondre, réfléchit plus et s'amuse plus.

Recueillir des données de comparaisons par paires semble plus valide d'un point de vue psychologique. La transitivité n'est plus une contrainte. Déceler des intransitivités au niveau individuel est une source d'information. Nous donnons ci-dessous un exemple de questionnaire de comparaisons par paires et un exemple de recherche d'ordre à distance minimum de la relation de tournoi obtenue.

Le questionnaire porte sur 8 objets (8 propositions) caractérisant une attitude envers le travail). Il comprend par conséquent 28 = 8 x 7/2 comparaisons par paires.

RECHERCHE SUR LES MOTIVATIONS EN GESTION DU PERSONNEL

1) On vous demande de comparer VOS préférences en ce qui concerne diverses motivations.

2) Lire les définitions en page 2.

3) Prendre chaque question à tour de rôle et barrer le mot en majuscules qui ne vous convient pas.

4) Ne faire aucune correction aux questions déjà répondues, même si vous variez dans vos réponses.

Attachez-vous plus d'importance à :

1) des COLLEGUES de travail agréables ou une organisation plus EFFICAC

2) une plus grande SECURITE ou des COLLEGUES de travail plus agréables

3) plus de chances d'AVENIR ou une plus grande SECURITE dans votre travail

4) un plus grand PRESTIGE attaché à votre position ou de meilleures chances d'AVENIR

5) plus d'AUTORITE dans votre travail ou un plus grand PRESTIGE attaché à votre position

6) des HORAIRES de travail plus courts ou une plus grande AUTORITE dan votre travail

7) un SALAIRE plus élevé ou des HORAIRES de travail plus courts

8) plus d'AUTORITE dans votre travail ou un SALAIRE plus élevé

9) un plus grand PRESTIGE attaché à votre position ou des HORAIRES de travail plus courts

10) plus d'AUTORITE dans le travail ou plus de chance d'AVENIR

11) plus de SECURITE ou un plus grand PRESTIGE attaché au travail

12) de meilleures chances d'AVENIR ou des COLLEGUES de travail agréable

13) une plus grande SECURITE dans le travail ou une organisation plus EFFICACE

14) une organisation plus EFFICACE ou de meilleures chances d'AVENIR

15) un plus grand PRESTIGE attaché au travail ou des COLLEGUES de travail plus agréables

16) une AUTORITE plus grande ou une plus grande SECURITE dans le travail

17) des HORAIRES de travail plus courts ou de meilleures chances d'AVENIR

18) un PRESTIGE plus grand dans le travail ou un SALAIRE plus élevé

19) un SALAIRE plus élevé ou de meilleures chances d'AVENIR

20) une plus grande SECURITE ou des HORAIRES de travail plus courts

21) plus d'AUTORITE dans le travail ou des COLLEGUES de travail plus agréables

22) un plus grand PRESTIGE dans le travail ou une organisation plus EFFICACE

23) plus d'AUTORITE ou une organisation plus EFFICACE

24) des HORAIRES de travail plus courts ou des COLLEGUES de travail plus agréables

25) un SALAIRE plus élevé ou une SECURITE plus grande

26) des COLLEGUES de travail plus agréables ou un salaire plus ELEVE

27) des HORAIRES de travail plus courts ou une organisation plus EFFICACE

28) une organisation plus EFFICACE ou un SALAIRE plus élevé.

Définitions sommaires des motivations

- SALAIRE : Argent reçu pour une durée normale de travail.
- HORAIRE : Nombre d'heures ouvrées dans une semaine normale de travail.
- AUTORITE : Les limites dans lesquelles vous pouvez prendre des décisions et contrôler leur application.
- PRESTIGE : Reconnaissance par la direction et vos collègues de l'importance de votre contribution et de votre position, à l'exclusion de l'aspect financier.
- AVENIR : Chances d'améliorer votre position dans l'organisation.
- SECURITE : L'assurance d'une stabilité dans l'emploi et d'une retraite.
- COLLEGUES de travail : Tous les gens avec qui vous êtes en contact direct et quotidien au travail.
- EFFICACITE de l'organisation : Désigne plutôt ici l'efficacité administrative de votre organisation et son fonctionnement harmonieux.

3. Comparaisons par blocs

Dans le cas où n est élevé, il n'est plus question de demander des comparaisons par paires. En effet, leur nombre $n(n - 1)/2$ risque de devenir trop important, entraînant une fatigue chez l'interviewé et par là même des réponses sans signification

n	n(n - 1)/2
2	1
3	3
4	6
5	10
6	15
7	21
8	28
9	36
10	45 etc.

Si n est trop grand, on ne peut plus recueillir des comparaisons par paires. Demander des classements avec des rangs risque également de devenir délicat. Il existe alors des méthodes de comparaisons par blocs qui permettent d'éviter ces difficultés.

3.1 Définition et exemples

- Un bloc est un sous-ensemble des objets.
- Sur chaque bloc, on demande un classement complet des objets.
- Chaque paire (i, j) d'objet apparaît dans un bloc et un seul.

Exemple 1 : Un exemple de 12 blocs de 3 objets (36 paires) a été donné dans l'introduction.

Exemple 2 : Questionnaire de M. HUTEAU [47] pour recueillir des préférences par paires portant sur 25 métiers.

Il a fallu 50 blocs de 4 métiers chacun (les réponses sont très faciles, le questionnaire n'est pas trop fastidieux) au lieu de 25 x 24/2, soit 300 comparaisons par paires.

1 □ instituteur (trice)
□ couturier (rière)
□ technicien (ienne)
□ écrivain

2 □ ingénieur-mécanicien
□ comptable
□ couturier(ière)
□ artiste

3 □ infirmier(ière)
□ ingénieur-agronome
□ comptable
□ coiffeur(euse)

4 □ mécanicien (ienne)
□ employé(e) de bureau
□ ingénieur agronome
□ journaliste

5 □ médecin
□ technicien(ienne)
□ employé(e) de bureau
□ imprimeur

6 □ couturier(ière)
□ vendeur(euse)
□ cultivateur(trice)
□ horloger(ère)

7 □ vendeur(euse)
□ imprimeur
□ journaliste
□ ingénieur-mécanicien

8 □ imprimeur
□ électronicien(ienne)
□ dessinateur(trice) industriel
□ comptable

9 □ électronicien(ienne)
□ instituteur(trice)
□ médecin
□ chimiste

10 □ comptable
□ chimiste
□ vendeur(euse)
□ explorateur

11 □ chimiste
□ écrivain
□ imprimeur
□ infirmier(ère)

12 □ écrivain
□ horloger(ère)
□ électronicien(ienne)
□ ingénieur-agronome

13 □ horloger(ère)
□ ingénieur-mécanicien
□ instituteur(trice)
□ aviateur(trice)

14 □ ingénieur-agronome
□ aviateur(trice)
□ chimiste
□ professeur

15 □ aviateur(trice)
□ artiste
□ écrivain
□ mécanicien(ienne)

16 □ artiste
□ explorateur(trice)
□ horloger(ère)
□ employé(e) de bureau

17 □ explorateur(trice)
□ infirmier(ière)
□ ingénieur-mécanicien
□ chauffeur de taxi

18 □ employé(e) de bureau
□ chauffeur de taxi
□ aviateur(trice)
□ dessinateur(trice industriel

19
chauffeur de taxi
coiffeur(euse)
artiste
médecin

20
coiffeur(euse)
professeur
explorateur(trice)
technicien(ienne)

21
professeur
mécanicien(ienne)
infirmier(ière)
cultivateur(trice)

22
technicien(ienne)
cultivateur(trice)
chauffeur de taxi
électronicien(ienne)

23
cultivateur(trice)
journaliste
coiffeur(euse)
instituteur(trice)

24
journaliste
dessinateur(trice) industriel
professeur
couturier(ière)

25
dessinateur(trice) industriel
médecin
mécanicien(ienne)
vendeur(euse)

26
instituteur(trice)
chauffeur de taxi
ingénieur agronome
vendeur(euse)

27
ingénieur-mécanicien
cultivateur(trice)
employé(e) de bureau
chimiste

28
infirmier(ière)
vendeur(euse)
technicien(ienne)
aviateur(trice)

29
mécanicien(ienne)
chimiste
couturier(ière)
chauffeur de taxi

30
médecin
aviateur(trice)
comptable
cultivateur(trice)

31
couturier(ière)
coiffeur(euse)
aviateur(trice)
imprimeur

32
vendeur(euse)
professeur
artiste
électronicien(ienne)

33
imprimeur
mécanicien
explorateur(trice)
instituteur(trice)

34
électronicien(ienne)
employé(e) de bureau
infirmier(ière)
couturier(ière)

35
- comptable
- journaliste
- chauffeur de taxi
- écrivain

36
- chimiste
- dessinateur(trice) industriel
- coiffeur(euse)
- horloger(ère)

37
- écrivain
- médecin
- professeur
- ingénieur-mécanicien

38
- horloger(ère)
- technicien(ienne)
- mécanicien(ienne)
- comptable

39
- ingénieur-agronome
- imprimeur
- cultivateur(trice)
- artiste

40
- aviateur(trice)
- électronicien(ienne)
- journaliste
- explorateur(trice)

41
- artiste
- instituteur(trice)
- dessinateur(trice) industriel
- infirmier(ière)

42
- explorateur(trice)
- couturier(ière)
- médecin
- ingénieur-agronome

43
- employé(e) de bureau
- écrivain
- vendeur(euse)
- coiffeur(euse)

44
- chauffeur de taxi
- horloger(ère)
- imprimeur
- professeur

45
- coiffeur(euse)
- ingénieur-mécanicien
- électronicien(ienne)
- mécanicien(ienne)

46
- professeur
- comptable
- instituteur(trice)
- employé(e) de bureau

47
- technicien(ienne)
- artiste
- chimiste
- journaliste

48
- cultivateur(trice)
- explorateur(trice)
- écrivain
- dessinateur(trice) industriel

49
- journaliste
- infirmier(ière)
- horloger(ère)
- médecin

50
- dessinateur(trice) industriel
- ingénieur-agronome
- ingénieur-mécanicien
- technicien(ienne)

3.2 Nombre de blocs à construire

Soit n objets à comparer et m le nombres d'objets dans un bloc ($2 \leq m \leq n$).

Le nombre de comparaisons par paires recherché est :

$$C_n^2 = \frac{n(n-1)}{2}.$$

Le nombre de paires comparées dans un bloc de m objets est :

$$C_m^2 = \frac{m(m-1)}{2}.$$

Le nombre de blocs à construire est :

$$\boxed{x_{n,m} = \frac{C_n^2}{C_m^2} = \frac{n(n-1)}{m(m-1)}}$$

Ce nombre donne déjà une information sur la possibilité de construire les m blocs. Il est nécessaire que $x_{n,m}$ soit un entier.

Tableau $x_{n,m}$:

Objets n \ Taille des blocs m	2	3	4	5	6	7	8	9	10
2	1								
3	3	1							
4	6	2	1						
5	10	X	X	1					
6	15	5	X	X	1				
7	21	7	X	X	X	1			
8	28	X	X	X	X	X	1		
9	36	12	6	X	X	X	X	1	
10	45	15	X	X	3	X	X	X	1
⋮									
25	300	100	50	30	20	X	X	....	

X : nombre non entier

Deux blocs particuliers sont :

- les <u>blocs de 2 éléments</u> : il y en a $n(n - 1)/2$; c'est le nombre de comparaisons par paires ;
- les <u>blocs de n éléments</u> : il y en a 1 ; c'est le mode de recueil de données par rangs.

Les comparaisons par blocs généralisent donc les recueils de données par rangs et par paires et, d'un point de vue pratique, sont un mode de recueil de données intermédiaires.

3.3 Nombre de relations distinctes

Dans chaque bloc on a un ordre complet, donc $m!$ possibilités.

Tous les blocs sont indépendants, c'est-à-dire que le choix d'un classement dans un bloc ne détermine a priori aucunement le choix d'un classement dans un autre bloc.

On a $m!$ choix possibles pour le 1er bloc
$m!$ choix possibles pour le 2e bloc
$\vdots$
$m!$ choix possibles pour le $x_{n,m}{}^{\text{ième}}$ bloc.

Le nombre de relations de comparaisons par blocs est donc :

$$\boxed{\alpha_{n,m} = (m!)^{x_{n,m}} = (m!)^{\frac{n(n-1)}{m(m-1)}}}$$

Pour $m = n$ (recueils par rangs), on trouve $\alpha_n = n!$.

Pour $m = 2$ (recueils par paires), on trouve $\alpha_n = 2^{\frac{n(n-1)}{2}}$.

3.4 Probabilité d'obtenir une relation sans intransitivités

Parmi les solutions possibles, seules $n!$ sont des tournois transitifs (ce sont des ordres complets). Les autres sont des relations de tournois avec circuits. La probabilité d'avoir une relation sans circuit est donc

$$\boxed{P_{n,m} = \frac{n!}{(m!)^{\frac{n(n-1)}{m(m-1)}}}}$$

On retrouve les deux formules classiques :

$P_{n,n} = 1$; l'intransitivité est impossible avec des rangs

et $P_{n,2} = \dfrac{n!}{2^{n(n-1)/2}}$.

3.5 Construction des blocs

On a vu le tableau des valeurs $x_{n,m}$. Une condition nécessaire pour que le problème soit possible est que $x_{n,m}$ soit entier. Cette condition n'est pas suffisante.

Par exemple, pour $n = 4$, $m = 3$, il est impossible de construire 2 blocs de 3 éléments reconstituant les 6 comparaisons par paires.

Les problèmes mathématiques liés à de telles constructions ont donné lieu à de nombreux développements. On pourra consulter K.D. TOCHER [85], A. DEY [29], BRADLEY [21].

BIBLIOGRAPHIE

[1] ARROW, K.J. - "Social choice and individual values".
Wiley, 2e édition, 1962.

[2] ASTIE, A. - "Comparaisons par paires et problèmes de classements - Estimation et test statistiques".
Mathématiques et Sciences Humaines, n° 32, 1970.

[3] ATTALI, J. - "Analyse économique de la vie politique".
P.U.F., 1972.

[4] de BACKER, P. ; RIBEILL, G. ; GORGES, Y. -"OEDIPE".
SEMA (Metra International), Direction Scientifique, Rapport de Recherche n° 70, 1973.

[5] BARBUT, M. - "Note sur les ordres totaux à distance minimum d'une relation binaire donnée".
Mathématiques et Sciences Humaines, n° 17.

[6] BARBUT, M. ; FREY, L. - "Techniques ordinales en analyse des données - algèbre et combinatoire".
Hachette, 1971.

[7] BARBUT, M. ; MONJARDET, B. - "Ordre et classification, algèbre et combinatoire".
Hachette, Tomes 1 et 2, 1970.

[8] BEN DUSHNIK ; MILLER, E.W. - "Partially ordered sets".
Amer. Journ. of Math., 1941.

[9] BENNET, J.R. ; HAYS, W.L. - "Multidimensional unfolding : determining configuration from Complete Rank Ordering of Preference Data".
Psychometrika, Vol. 26, 1961.

[10] BENZECRI, J.P. - "Sur l'analyse des préférences".
ISUP, Polycopié, 1965.

[11] BERNARD, G. ; M.L. BESSON - "Douze méthodes d'analyse multicritères"
RIRO, 5e année, n° V-3, 1971.

[12] BERMOND, J.C. - "Ordres à distance minimum d'un tournoi et graphes partiels sans circuits maximaux".
Mathématiques et Sciences Humaines, n° 37.

[13] BERTIER, P. ; BOUROCHE, J.M. - "Analyse et structuration de données multidimensionnelles".
P.U.F., 1973.

[14] BERTIER, P. ; JACQUET-LAGREZE, E. - "Programme Architecture Nouvelle"
SEMA (Metra International), Direction Scientifique, Note de Travail n° 193, octobre 1973.

[15] BERTIER, P. ; de MONTGOLFIER, J. - "Comment choixir en tenant compte de points de vue non commensurables".
Analyse et Prévision, Tome XI, 1971.

[16] BERTIER, P. ; ROY, B. - "La méthode ELECTRE II : une application au média-planning".
VIe Conférence Internationale de Recherche Opérationnelle, Dublin, 1972.

[17] BLACK, D. - "The theory of committes and elections".
Cambridge University Press, 1958.

[18] BOUROCHE, J.M. - "Analyse des proximités et analyse des préférences".
METRA, Vol. X, n° 4, 1971.

[19] BOUROCHE, J.M. ; LAYE, P. - "Analyse des préférences : modèles PREFMAP I et II, utilisation des programmes".
SEMA(Metra International), Direction Scientifique, Note de Travail n° 190, 1973.

[20] BOWMAN, V.J. ; COLANTONI, C.S. - "Majority rule under transitivity constraints.
Management Science, Vol. 19, n° 9, 1973.

[21] BRADLEY - " Incomplete block rank analysis : on the appropriatness of the model for a method of paires comparaisons".
Biometrics, 10, n° 3, 1954.

[22] CARROLL, J.D. - "Individual differences and multidimensional scaling".
Bell Telephone Laboratories, Polycopié, 1970.

[23] COOMBS, C.H. - "A theory of data".
Wiley, 1964.

[24] C.R.U. (Centre de Recherche d'Urbanisme) - Etude sur la décision -
Tome 1 : Rapport sur l'agrégation des critères, 1970.

[25] DAVID, H.A. - "The method of paired comparaisons".
Griffin's statistical monographs, 1963.

[26] de CANI, J.S. - "A branch and bounded algorithm for maximum likehood paired comparaisons ranking".
Biometrika, 59, 1, 1972.

[27] de CANI, J.S. - "Maximum likehood paires comparaisons ranking by linear programming".
Biometrika, 56, 3, 1969.

[28] DEGENNE, A. - "Techniques ordinales en analyse des données statistiques".
Hachette, 1972.

[29] DEY, A. - "On construction of balanced n-ary glock designs".
Annals Inst. Statist. Mathematics, V 22, 1970.

[30] DUCAMP, A. - "Sur la dimension d'un ordre partiel".
Théorie des graphes, journées internationales, Rome 1966, Dunod, 1967, p. 103-112.

[31] DURAND, B. - "A propos du nombre minimal d'arcs à enlever pour supprimer les circuits d'un graphe".
Mathématiques et Sciences Humaines, n° 20, 1967.

[32] ECKART, C. ; YOUNG, G. - "The approximation of one matrix by another of lower rank".
Psychometrika, 1936.

[33] EDWARDS, A.L. ; KILPATRICK, F.P. - "Scale analysis and the measureme of social attitudes".
Psychometrika, Vol. 13, n° 2, 1948.

[34] Elaboration d'échelles de propension à dépenser et de progressisme dans les achats. SOBEMAP, 1969.

[35] FELDMAN, J. - "Pôles, intermédiaires et centres dans un groupe d'opinion".
Mathématiques et Sciences Humaines, n° 43, 1973.

[36] FENELON, J.P. - Thèse de 3e cycle, Université de Paris VI, 1973.

[37] FISHBURN, P.C. - "Intransitive individual indifference and transitiv majorities".
Econometrica, Vol. 38, n° 3, 1970.

[38] FLAMENT, C. - "Théorie des graphes et structures sociales".
Paris, Gauthier-Villars.

[39] FREY, L. - "Analyse ordinale des évangiles synoptiques".
Gauthier-Villars, Paris, 1972.

[40] GILLOT, C. - "Recherches sur un modèle statistique de comparaisons par paires".
Thèse, Toulouse, 1964.

[41] GREEN, P.E. ; CARMONE, F.J. - "Multidimensional scaling and related techniques in marketing analysis".
Allyn and Bacon, Inc., Boston, Massachusetts.

[42] GUILBAUD, G.T. - "Eléments de la théorie mathématique des jeux".
Monographies de Recherche Opérationnelle, n° 9, Dunod, 1964.

[43] GUILBAUD, G.T. - "Préférences stochastiques".
Mathématiques et Sciences Humaines, n° 32, 1970.

[44] GUILBAUD, G.T. ; ROSENSTIEHL, P. - "Analyse algébrique d'un scrutin".
Mathématiques et Sciences Humaines, n° 4, 1963.

[45] GUTTMAN, L. - "The Cornell technique for scale and intensity analysis".
Educational and Psychological Measurement, Vol. 7, 1947.

[46] HEUCHENNE, C. - "Un algorithme général pour trouver un sous-ensemble d'un certain type à distance minimum d'une partie donnée".
Mathématiques et Sciences Humaines, 1970.

[47] HUTEAU, M. - "Aspirations et représentations professionnelles chez les élèves du 1er cycle de l'enseignement secondaire".
Thèse de 3e cycle, Paris, 1972.

[48] INADA, K. - "A note on the Simple Majority Decision Rule".
Econometrica, 32, 1964.

[49] JACQUET-LAGREZE, E. - "L'agrégation des opinions individuelles".
Informatique en Sciences Humaines, n° 4, 1969.

[50] JACQUET-LAGREZE, E. - "Opinions valuées et graphes de préférence".
Mathématiques et Sciences Humaines, n° 33, 1971.

[51] JACQUET-LAGREZE, E. - Le problème de l'agrégation des préférences : une classe de procédures à seuil".
Mathématiques et Sciences Humaines, n° 43, 1973.

[52] JACQUET-LAGREZE, E. - "How we can use the notion of semiorders to bu
outranking relations in multicriteria decision making".
Fourth Research Conference on Subjective Probability, Utility and
Decision Making, Rome, September 3-6, 1973.

[53] KAUFMAN, A. - "Introduction à la théorie des sous-ensembles flous".
Masson et Cie, 1973.

[54] KENDALL, M.G. - "Rank Correlation Methods".
Londres, Griffin, 1962.

[55] KENDALL, M.G. ; BABINGTON SMITH - "On the method of paires comparais
Biometrika, 31, 1940.

[56] KREWERAS, G. - "Représentation polyédrique des préordres complets
finis et application à l'agrégation des préférences".
La Décision, CNRS, 1967.

[57] LEBART, L. ; FENELON, J.P. - "Statistique et informatique appliquées"
Dunod, 1971.

[58] LECLERE, N. ; PEYROUX, C. - "Session analyse des données et choix
multicritères", Méthode de JACQUET-LAGREZE, Cours CIRO, 1972.

[59] LUCE, R.D. - "Semiorders and a theory of utility discrimination".
Econometrica, V 24, 1956.

[60] MATALON - "L'analyse hiérarchique".
Gauthier-Villars, Paris, 1965.

[61] MAY, K.O. - "A set of independant, necessary and sufficient conditio
for simple majority decision".
Econometrica, V 20, 1952.

[62] MAYER, C. - "Variables qualitatives métrisables : mise en forme, méthodes de traitement".
SEMA (Metra International), Direction Scientifique, Note de Travail n° 183, 1973.

[63] MONJARDET, B. - "Probabilité de l'effet Condorcet".
Mathématiques et Sciences Humaines, n° 23, 1968.

[64] MONJARDET, B. - "Quelques problèmes relatifs à la méthode des comparaisons par paires".
Mathématiques et Sciences Humaines, n° 36, 1972.

[65] MONJARDET, B. - "Tournois et ordres médians pour une opinion".
Mathématiques et Sciences Humaines, n° 43, 1973.

[66] MOON, J.W. - "Topics on tournaments".
New-York, Holt, 1968.

[67] PARLEBAS - "Effet Condorcet et dynamique sociométrique".
Mathématiques et Sciences Humaines, n° 36, 1972.

[68] PATTANAIK, P.K. - "A note on Democratic Decision and the Existence of Choice Sets".
Review of Economic Studies, V 35, 1968.

[69] PHAN, T.H. - "Les méthodes INDSCAL et PREFMAP".
SEMA (Metra International), Direction Scientifique, Note de Travail n° 150, 1971.

[70] PHILLIPS, J.P.N. - "On a certain type of partial higher-ordered metric scaling".
The British Journal of Mathematical and Statistical Psychology, 1966.

[71] PHILLIPS, J.P.N. - "A procedure for determining SLATER's i and all nearest adjoining orders".
The British Journal of Mathematical and Statistical Psychology, 1967.

[72] PHILLIPS, J.P.N. - "A further procedure for determining SLATER's i and all nearest adjoining orders".
The British Journal of Mathematical and Statistical Psychology, 1969.

[73] REMAGE, R. ; THOMPSON, W.A. - "Ranking from paires comparisons".
Ann. Math. Statist., 35, n° 2, 1964.

[74] REMAGE, R. ; THOMPSON, W.A. - "Maximum likehood paired comparisons".
Biometrika, 53, n° 1 et 2, 1966.

[75] RIBEILL, G. - "Equilibres, équivalences, ordres et préordres à distanc minimum d'un graphe complet donné".
Mathématiques et Sciences Humaines, n° 43, 1973.

[76] RIBEILL, G. - "Modèles et Sciences Humaines".
METRA, Vol. XII, n° 2, 1973.

[77] ROBERTS, F.S. - "Homogeneous families of semiorders and the theory of probabilistic consistency".
Journ. of Math. Psych., 8, 1971.

[78] ROSKAM, E.I. - "Metric analysis of ordinal data in psychology : model and numerical methods for metric analysis of conjoint ordinal data in psychology".
VAN VOORSCHOTEN, 1968.

[79] ROY, B. - "Algèbre moderne et théorie des graphes orientées vers les sciences économiques et sociales".
Paris, Dunod, Tome 1 : 1969, Tome 2 : 1970.

[80] ROY, B. - "Décisions avec critères multiples : problèmes et méthodes"
METRA, Vol. XI, n° 1, 1972.

[81] ROY, B. - "Critères multiples et modélisation des préférences - l'apport des relations de surclassement".
SEMA (Metra International), Direction Scientifique, Synthèse et Formation n° 79, 1973.

[82] SEN, K.A. - "A possibility theorem on majority decisions".
Econometrica, Vol. 34, n° 2, 1966.

[83] SEN, K.A. - "Collective choice and social welfare".
Londres, Oliver and Boyd, 1970.

[84] SLATER, P. - "Inconsistencies in a schedule of paires comparisons".
Biometrika, 48, 3 and 4, 1961.

[85] TOCHER, K.D. - "Design and analysis of black experiments".
Journal. Roy. Statist. Soc., B.14, 1952.

[86] TORGERSON, W.S. - "Theory and methods of scaling".
Wiley, 1958.

[87] VICKREY, W. - "Utility, strategy, and social decision rules".
Quaterly Journal of Economics, V 74, 1960.

[88] WARD, B. - "Majority voting and alternative forms of public enterprises".
In J. Margolis (ed.), 2nd Conference in Urban Expenditure.

[89] WASSERVOGEL, F. ; LECHAT, J. - "Une méthode de classification d'items qualitatifs".
RIRO, V-2, 1971.

metra

bruxelles francfort genève londres madrid
milan montréal new york paris rome vienne