REVUE DE STATISTIQUE APPLIQUÉE

J. L. PETIT

Généralisation de la méthode des partitions centrales

Revue de statistique appliquée, tome 41, n° 3 (1993), p. 49-72

<http://www.numdam.org/item?id=RSA_1993__41_3_49_0>

Rev. Statistique Appliquée, 1993, **XLI** (3), 49-72

GÉNÉRALISATION DE LA MÉTHODE DES PARTITIONS CENTRALES

J.L. Petit

Département de Mathématiques et Informatique Appliquées
Université Paul Valéry
Route de Mende – B.P. 5043
34032 – MONTPELLIER CEDEX 1

«En hommage à Simon Régnier»

RÉSUMÉ

La méthode des partitions centrales de S. Régnier est une technique de classification. Nous généralisons cette méthode en utilisant un point de vue bayésien. Nous testons les résultats obtenus sur différents problèmes concrets.

Mots clés : *Classification, Agrégation, Partition Centrale, Méthode Bayésienne.*

SUMMARY

The method of central partition of S. Régnier is a cluster's technique. We generalize this method in introducing a bayesian point of view. We verify ours results over some concrete problems.

Key Words : *Cluster's Analysis, Aggregation, Binary Relation, Bayes Method.*

Introduction

La méthode des *Partitions Centrales* est une technique de classification qui, à la différence des *Méthodes Hiérarchiques* de classification, produit une seule partition à partir d'observations faites sur une certaine population. (Il est souvent plus simple d'interpréter une seule partition qu'une chaîne de partitions). Cette méthode a été introduite par S. Régnier (cf : [4],[8]).

Nous allons généraliser cette méthode d'un *point-de-vue bayésien* en introduisant deux mesures *a-priori* : une mesure sur les individus de la population que l'on cherche à classifier et une mesure sur les observations faites sur cette population.

Ce travail comporte deux parties :

– Dans la première partie théorique, nous exposons la généralisation de la méthode des partitions centrales et donnons un certain nombre de résultats nouveaux sur cette méthode.

– Dans la deuxième partie pratique, nous testons les résultats obtenus sur un certain nombre de problèmes concrets.

I. Partie théorique

I.1 Rappels

Espaces Euclidiens

Soient :

E un ensemble fini

μ une mesure de probabilité sur E

Nous pouvons définir une distance sur l'ensemble des sous-ensembles de E :

$$d_\mu(A, B) = \mu(A \Delta B)$$

où Δ note la différence symétrique $((A - B) + (B - A))$.

Sur l'espace fonctionnel $\mathbb{R}^E$, nous pouvons introduire une structure euclidienne, en considérant le produit scalaire suivant :

$$\langle f, g \rangle_\mu = \sum_{e \in E} f(e) g(e) \mu(e);\ f \in \mathbb{R}^E,\ g \in \mathbb{R}^E$$

En notant I_A la fonction indicatrice d'un sous-ensemble A, nous avons alors :

Proposition (1)

$$d_\mu(A, B) = \|I_A - I_B\|_\mu^2 = \sum_{e \in E} (I_A(e) - I_B(e))^2 \mu(e) = \sum_{e \in E} |I_A(e) - I_B(e)| \mu(e)$$

□

Treillis des Partitions

Soit $\mathbb{P}(E)$, l'ensemble des partitions de E. Une partition $P \in \mathbb{P}(E)$ est une décomposition en classes de l'ensemble E :

$$E = \sum_{\alpha=1}^{\alpha=k} E_\alpha$$

Nous noterons $I_P \in \{0,1\}^{E*E}$, la fonction indicatrice de la relation d'équivalence associée à la partition P :

$$e\,P\,e' \text{ (e et e' dans la même classe)} \Leftrightarrow I_P(e,e') = 1$$

Sur $\mathbb{P}(E)$, il existe une relation d'ordre classique, notée $P < Q$, et se lisant de droite à gauche Q *est plus grossière que* P ou de gauche à droite P *est plus fine que* Q :

$$P < Q \Leftrightarrow I_P \leqslant I_Q$$

Pour cette relation d'ordre, $\mathbb{P}(E)$ est un treillis possèdant une borne supérieure, P_{sup}, la partition la *plus grossière* et une borne inférieure, P_{inf}, la partition la *plus fine* :

$$P_{\text{inf}} < P < P_{\text{sup}}; P \in \mathbb{P}(E)$$

Soit un «caractère» (c'est-à-dire une fonction) défini sur E et à valeurs dans un ensemble $\mathcal{X}$:

$$X : E \to \mathcal{X}$$

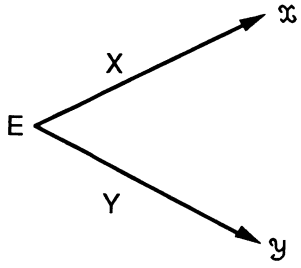
A ce caractère, nous pouvons associer une partition P_X définie par :

$$e\,P_X\,e' \Leftrightarrow X(e) = X(e')$$

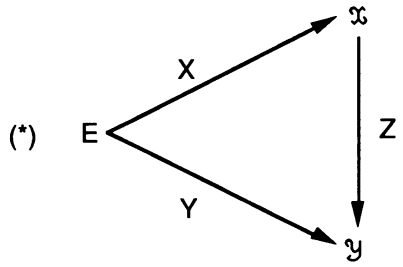
La technique des partitions centrales ne conserve des observations X que les partitions P_X correspondantes. Etudions cette *perte d'information.*

Codages

Considérons deux caractères définis sur E :



Le caractère Y est dit un *codage* du caractère X si on a le schéma commutatif suivant (c'est-à-dire s'il existe Z tel que $Y = Z \circ X$) :



Pour simplifier, nous supposerons les caractères X et Y surjectifs. Nous avons alors :

Proposition (2)

Les deux propriétés suivantes sont équivalentes :

(1) $P_X < P_Y$

(2) Y *est un codage de* X

□

Corollaire

Les deux propriétés suivantes sont équivalentes :

(1) $P_X = P_Y$

(2) Le diagramme commutatif (*) est construit à l'aide d'une bijection.

□

Remarques

L'opération de codage correspond à *une perte d'injectivité*. (cf : par exemple, le découpage en classes). Cette technique est souvent utilisée en statistique. Pour une étude et une bibliographie détaillées sur ce sujet, nous renvoyons le lecteur à [2]. La technique des partitions centrales ne gardant des observations que les partitions associées, est *invariante* par l'introduction de bijections sur ces observations. Dans le cadre où nous nous sommes placés, nous pourrions *quantifier* cette perte d'information en introduisant la quantité d'information de Shannon :

$$I(P) = -\sum_{\alpha=1}^{\alpha=k} \mu(E_\alpha) \mathrm{Log}\, \mu(E_\alpha)$$

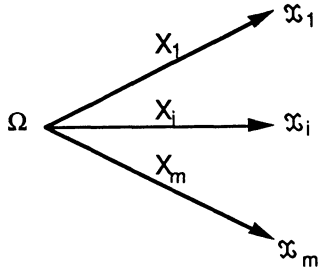
(Si nous étions placés dans un cadre inférentiel, nous serions conduits à la notion d'exhaustivité). □

I.2 Notions de Partitions Centrales

Les données du problème sont les suivantes :

– une population Ω à n éléments dont on cherche à classifier les éléments.

– m observations sur cette population, $X = \{X_i / 1 \leqslant i \leqslant m\}$:



– une mesure de probabilité μ strictement positive sur Ω

– une mesure de probabilité ν strictement positive sur X

A chaque observation X_i est associée une partition que nous noterons P_i (pour simplifier l'écriture, nous notons P_i à la place de P_{X_i}). Les m observations fournissent ainsi une famille de partition :

$$\wp = (\wp(X)) = \{P_i / 1 \leqslant i \leqslant m\}$$

Le problème que l'on se pose est de chercher une partition C *proche* de la famille $\wp$. Pour cela définissons un indice de distance δ par :

$$\delta(P; \wp) = \sum_{i=1}^{i=m} \nu(i) \| I_P - I_i \|^2_{\mu^2};\ P \in \mathbb{P}(\Omega)$$

où μ^2 note la mesure produit sur Ω^2 et I_i la fonction indicatrice de la relation d'équivalence associée à la partition P_i.

Définition

Une partition $C \in \mathbb{P}(\Omega)$ est dite centrale *si elle minimise le critère δ :*

$$\delta(C; \wp) = \min_P \delta(P; \wp) \tag{1}$$

□

Remarques

(1) Il existe toujours une partition centrale, car l'ensemble Ω est fini, mais elle n'est pas nécessairement unique.

(2) Si nous avons $P_1 = ... = P_m$, la partition $C = P_1 = ... = P_m$ est centrale car $\delta(C; \wp) = 0$. Nous sommes en particulier dans ce cas si les caractères X_i sont injectifs (dans ce cas C est la partition la plus fine) ou si les caractères X_i sont constants (dans ce cas C est la partition la plus grossière).

□

I.3 Programmation Linéaire en Nombres Entiers

Nous allons montrer que le problème des partitions centrales se ramène à un problème de programmation linéaire en nombres entiers. Pour cela, introduisons le barycentre I des points I_i :

$$I = (I(\wp)) = \sum_{i=1}^{i=m} \nu(i) I_i$$

Un calcul classique nous conduit au résultat suivant :

Proposition (3)

Une partition centrale minimise le critère suivant :

$$\delta'(P, \wp) = \|I_p - I\|^2_{\mu^2} \tag{2}$$

□

En introduisant le point $H \in \mathbb{R}^{\Omega^2}$ dont toutes les coordonnées sont 1/2, nous obtenons le lemme suivant :

Lemme

Pour toute partition $P \in \mathbb{P}(\Omega)$, nous avons :

$$\|I_P - H\|^2_{\mu^2} = 1/4$$

□

Introduisons la forme linéaire suivante :

$$\begin{aligned} L(P; \wp) &= \langle I_P, I - H \rangle_{\mu_2} \\ &= \sum_{\omega, \omega'} I_P(\omega, \omega')(I(\omega, \omega') - 1/2)\mu(\omega)\mu(\omega') \end{aligned} \tag{3}$$

L'utilisation du lemme nous conduit au résultat suivant :

Théorème (1)

Une partition centrale C maximise la forme linéaire $L(P;\wp)$:

$$L(C;\wp) = \max_P L(P;\wp) = L(\wp) \tag{4}$$

□

Remarques

(1) Le problème de la recherche d'une partition centrale est donc équivalent au problème de programmation linéaire en nombres entiers suivant :
recherche d'une famille de nombres $\{x(\omega,\omega')/(\omega,\omega') \in \Omega^2\}$ maximisant le critère :

$$\sum_{\omega,\omega'} x(\omega,\omega')(I(\omega,\omega') - 1/2)\mu(\omega)\mu(\omega')$$

et vérifiant les contraintes :

$$(x(\omega,\omega'))^2 = x(\omega,\omega'); x(\omega,\omega') = x(\omega',\omega); x(\omega,\omega) = 1;$$
$$x(\omega,\omega') + x(\omega',\omega'') - x(\omega,\omega'') \leq 1$$

Barthélemy et Monjardet ([1]) ainsi que Marcotorchino et Michaud ([6]) signalent les analogies existantes entre la méthode des partitions centrales (agrégation de relations d'équivalence) et certaines méthodes de choix social (agrégation de relations d'ordres linéaires). En effet, dans ce dernier cas, on cherche à optimiser une forme linéaire $L(O)$, O étant un ordre linéaire. Les nombres $\{x(\omega,\omega')/(\omega,\omega') \in \Omega^2\}$ caractérisant un tel ordre vérifient :

$$(x(\omega,\omega'))^2 = x(\omega,\omega'); x(\omega',\omega') + x(\omega',\omega) = 1(\omega \neq \omega'); x(\omega,\omega) = 1;$$
$$x(\omega,\omega') + x(\omega',\omega'') - x(\omega,\omega'') \leq 1$$

Nous allons préciser, au paragraphe I-8, cette notion.

(2) Nous avons les cas particuliers suivants :

$I \leqslant 1/2 \Rightarrow$ la partition la plus fine est centrale
$I \geqslant 1/2 \Rightarrow$ la partition la plus grossière est centrale
$I = 1/2 \Rightarrow$ toute partition est centrale

où $I \leq 1/2$ (resp. $\geq 1/2$; $= 1/2$) signifie que $\forall\omega,\omega'(\omega \neq \omega') : I(\omega,\omega') \leq 1/2$ (resp. $I(\omega,\omega') \geq 1/2$; $I(\omega,\omega') = 1/2$)

(3) Si la partition P correspond à la décomposition suivante de Ω :

$$\Omega = \sum_{\alpha=1}^{\alpha=k} \Omega_\alpha$$

nous avons :

$$L(P;\wp) = \sum_{\alpha=1}^{\alpha=k} \sum_{(\omega,\omega')\in\Omega_\alpha^2} (I(\omega,\omega') - 1/2)\mu(\omega)\mu(\omega') \tag{5}$$

□

I.4 Partitions Centrales et Codage

Supposons que nous remplacions les observations $X = \{X_i/1 \leq i \leq m\}$ par des observations *codées* $Y = \{Y_i/1 \leq i \leq m\}$. En notant P_i' la partition associée à Y_i, nous avons donc :

$$P_i < P_i'; i = 1, ..., m$$

Nous obtenons alors les inégalités suivantes :

Théorème (2)

$$1/2 \sum_{\omega} \mu^2(\omega) \leqslant L(\wp) \leqslant L(\wp') \leqslant 1/2 \tag{6}$$

où $\wp' = \{P_i'/1 \leq i \leq m\}$

Démonstration

Nous avons :

$$\text{, } P_i < P_i' \text{; pour tout } i \Rightarrow I_i(P_i) \leq I_i(P_i') \text{; pour tout } i$$
$$\Rightarrow I(\wp) \leq I(\wp') \Rightarrow L(P;\wp) \leq L(P;\wp')$$

d'où

$$L(\wp) = \max_P L(P;\wp) \leq L(\wp') = \max_P L(P;\wp')$$

La borne inférieure $1/2 \sum_{\omega} \mu^2(\omega)$ *est obtenue en prenant :*

$$P_1 = ... = P_m = P_{\text{inf}}$$

La borne supérieure 1/2 est obtenue en prenant :

$$P_1 = ... = P_m = P_{\text{sup}}$$

□

1.5 Attraction

Définissons, sur $\Omega \times \Omega$, *la fonction d'attraction* symétrique :

$$A : \Omega \times \Omega \to R$$
$$(\omega, \omega') \to A(\omega, \omega') = \text{ attraction entre } \omega \text{ et } \omega'$$
$$= 2\mu(\omega)\mu(\omega')(I(\omega, \omega') - 1/2)$$

Remarques

(1) $A(\omega, \omega) = \mu(\omega)^2$

(2) L'attraction peut être positive ou négative

□

Nous pouvons, par intégration, définir l'attraction sur $\wp(\Omega) * \wp(\Omega)$ en posant :

$$A(\Omega', \Omega'') = \sum_{\omega' \in \Omega', \omega'' \in \Omega''} A(\omega', \omega'') \tag{7}$$

Par convention, nous posons :

$$A(\Omega', \Omega'') = 0 \text{ si } \Omega' = \varnothing \text{ ou } \Omega'' = \varnothing$$

Considérons une partition P :

$$\Omega = \sum_{\alpha=1}^{\alpha=k} \Omega_\alpha$$

un calcul simple permet d'exprimer la forme linéaire $L(P; \wp)$ à partir de la fonction d'attraction A :

Proposition (4)

$$L(P; \wp) = 1/2 \sum_{\alpha=1}^{\alpha=k} A(\Omega_\alpha, \Omega_\alpha) \tag{8}$$
$$= 1/2 \sum_{\alpha=1}^{\alpha=k} \sum_{\omega \in \Omega_\alpha} A(\omega, \Omega_\alpha)$$
$$= 1/2 \sum_{\omega \in \Omega} \mu^2(\omega) + 1/2 \sum_{\alpha=1}^{\alpha=k} \sum_{\omega \in \Omega_\alpha} A(\omega, \Omega_\alpha - \{\omega\})$$

□

I.6 Algorithme des Transferts

La recherche d'une partition centrale est un problème très complexe. Ceci tient au fait que le nombre de partitions sur un ensemble à n éléments (nombres de Bell; cf [3]) a une croissance gigantesque avec n (c'est le drame de la classification!). Il existe différents algorithmes pour chercher une *approximation* de la notion de partition centrale (cf [1], [5],[6]).

Remarques

Par exemple, Marcotorchino et Michaud (cf [6]) utilisent une technique de programmation linéaire en nombres entiers. Ils remplacent la liaison :

$$(x(\omega,\omega'))^2 = x(\omega,\omega')$$

par la contrainte :

$$0 \leq x(\omega,\omega') \leq 1$$

et emploient une méthode de coupes (cf [9])

□

Nous allons généraliser *l'algorithme des transferts* dû à S. Régnier. Cet algorithme est simple et conduit à un *optimum local.*

Considérons une partition P :

$$\Omega = \sum_{\alpha=1}^{\alpha=k} \Omega_\alpha$$

et soient $1 \leq \beta \leq k$, $1 \leq \gamma \leq k+1$, $\omega \in \Omega_\beta$. Nous appellerons P' la partition obtenue à partir de P en *transférant* l'élément ω de la classe Ω_β dans la classe Ω_γ :

$$\Omega = \sum_{\alpha=1}^{\alpha=k'} \Omega'_\alpha$$

où :

$$\Omega'_\beta = \Omega_\beta - \{\omega\}$$
$$\Omega'_\gamma = \Omega_\gamma + \{\omega\}$$
$$\Omega'_\delta = \Omega_\delta \text{ pour tout } \delta \neq \beta, \gamma$$

Remarques

L'indice $\gamma = k + 1$ correspond à la création d'une nouvelle classe :

$$\Omega'_{k+1} = \{\omega\}$$

□

Calculons le gain obtenu par ce transfert :

$$\Delta L(\omega : \Omega_\beta \to \Omega_\gamma) = L(P'; \wp) - L(P; \wp)$$

Théorème (3)

$$\Delta L(\omega : \Omega_\beta \to \Omega_\gamma) = A(\omega, \Omega_\gamma) - A(\omega_\beta - \{\omega\}) \tag{9}$$

Démonstration

L'équation (9) découle directement de l'équation (8).

□

Le théorème (3) nous fournit donc le principe de l'algorithme des transferts : on part d'une partition initiale (par exemple, P_{inf} ou P_{sup} ou une des partitions P_i) et nous cherchons à l'améliorer en effectuant des transferts.

Remarques

(1) L'algorithme des transferts nous conduit à un optimum local : une partition que l'on ne peut améliorer par le transfert d'un seul élément.

(2) Pour trouver une partition centrale, il suffit de maximiser le critère suivant :

$$L'(P; \wp) = 1/2 \sum_{\alpha=1}^{\alpha=k} \sum_{\omega \in \Omega_\alpha} A(\omega, \Omega_\alpha - \{\omega\}) \tag{10}$$

Pour ce critère, les inégalités du théorème (2) deviennent :

$$0 \leqslant L'(\wp) \leqslant L'(\wp') \leqslant 1/2(1 - \sum_{\omega} \mu^2(\omega)) \tag{11}$$

(3) Nous pouvons généraliser l'algorithme des transferts en transférant de la classe Ω_β dans la classe Ω_γ, non un seul élément mais un sous-ensemble $\Omega'_\beta \subset \Omega_\beta$. Nous avons alors en utilisant l'équation (8) :

$$\begin{aligned} \Delta L(\Omega'_\beta : \Omega_\beta \to \Omega_\gamma) &= A(\Omega'_\beta, \Omega_\gamma) - A(\Omega'_\beta, \Omega_\beta - \Omega'_\beta) \\ &= \Delta L'(\Omega'_\beta : \Omega_\beta \to \Omega_\gamma) \end{aligned} \tag{12}$$

□

I.7 Eléments Equivalents

Définition

Deux éléments ω et ω' de Ω sont dit équivalents *s'ils vérifient les conditions suivantes :*

$$A(\omega, \omega'') = A(\omega', \omega'') \text{ pour tout } \omega'' \in \Omega$$

Nous pouvons donner une autre caractérisation de cette équivalence.

Proposition (5)

Deux éléments ω et ω' sont équivalents si et seulement si ils vérifient les conditions suivantes :

$$\mu(\omega) = \mu(\omega'), \omega P_i \omega' \text{ pour } i = 1, ..., m$$

□

Nous avons alors le résultat suivant :

Théorème (4)

Deux éléments équivalents sont toujours dans la même classe pour une partition centrale.

Démonstration

Raisonnons par l'absurde. Soit P une partition telle que :

$$\omega \in \Omega_\beta, \omega' \in \Omega, \ \omega \text{ et } \omega' \text{ équivalents}$$

Montrons que P n'est pas une partition centrale. Considérons deux partitions voisines *de P (au sens des transferts) :*

$$P' \text{ obtenue à partir de } P \text{ par le transfert } \omega : \Omega_\beta \to \Omega_\gamma$$
$$P'' \text{ obtenue à partir de } P \text{ par le transfert } \omega' : \Omega_\gamma \to \Omega_\beta$$

Nous avons alors :

$$\begin{aligned}
&[L(P'; \wp) - L(P; \wp)] + [L(P''; \wp) - L(P; \wp)] \\
&\quad = \Delta L(\omega : \Omega_\beta \to \Omega_\gamma) + \Delta L(\omega' : \Omega_\gamma \to \Omega_\beta) \\
&\quad = A(\omega, \Omega_\gamma) + A(\omega', \Omega_\beta) - A(\omega, \Omega_\beta - \{\omega\}) - A(\omega', \Omega_\gamma - \{\omega'\}) \\
&\quad = 2A(\omega, \omega') > 0
\end{aligned}$$

Ainsi, une des deux partitions P' ou P'' est strictement meilleure que la partition P vis-à-vis du critère L.

□

Ainsi, le problème de la recherche d'une partition centrale se simplifie en utilisant le passage au quotient :

$$\pi : \Omega \longrightarrow \pi(\Omega) = \Omega/\sim$$
$$\omega \longrightarrow \pi(\omega) = \rho = \text{classe d'équivalence de } \omega$$

Dans cet espace quotient, nous cherchons à maximiser le critère $L_\pi(P;\wp)$, $P \in \mathbb{P}(\pi(\Omega))$:

$$L_\pi(P;\wp) = \sum_{(\rho,\rho')\in\pi(\Omega)} (I_\pi(\rho,\rho') - 1/2)\pi_p(\rho)\mu_\pi(\rho') \tag{13}$$

où :

$$I_\pi(\rho,\rho') = I(\omega,\omega');\omega \in \rho,\omega' \in \rho'$$
$$\mu_\pi(\rho) = \sum_{\omega\in\rho} \mu(\omega) = \mu(\pi^{-1}(\rho))$$

Ceci revient à travailler avec la fonction d'attraction A_π :

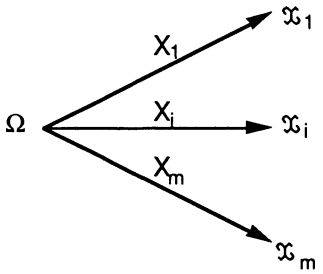
$$A_\pi(\rho,\rho') = 2\mu_\pi(\rho)\mu_\pi(\rho')(I_\pi(\rho,\rho') - 1/2)$$

I.8 Ordres Centraux

Nous allons remplacer la notion de relation d'équivalence par la notion de relation d'ordre total. Les données du problème sont maintenant les suivantes :

– une population Ω à n éléments dont on cherche à ordonner les éléments.

– m observations sur cette population, $X = \{X_i/1 \leq i \leq m\}$:



où $\mathfrak{X}_1 = ... = \mathfrak{X}_i = ... = \mathfrak{X}_m = \{1,...,n\}$ et les fonctions X_i sont des bijections associées aux ordres O_i :

$$\omega O_i \omega' \Leftrightarrow X_i(\omega) \leq X_i(\omega')$$

– une mesure de probabilité μ strictement positive sur Ω

– une mesure de probabilité ν strictement positive sur X

En notant I_i la fonction indicatrice associée à l'ordre O_i, nous cherchons à minimiser, sur l'ensemble des ordres totaux le critère suivant :

$$\delta_1(O, X) = \sum_{i=1}^{i=m} \nu(i)||I_O - I_i||^2_{\mu^2} \tag{14}$$

Définition

Un ordre total C est dit central s'il minimise le critère δ_1

□

Remarques

Monjardet (cf : [7]) appelle un tel ordre, un *ordre médian* et pense que cette notion a été introduite par Condorcet.

□

En introduisant le barycentre I, nous obtenons le résultat suivant : Un ordre central C minimise le critère δ'_1 :

$$\delta'_1(0, X) = ||I_O - I||^2_{\mu^2} \tag{15}$$

Les ordres totaux sont sur une sphère centrée à l'origine car nous avons le résultat suivant :

Lemme

Pour tout ordre total O, nous avons :

$$||I_O||^2_{\mu^2} = \frac{1 + \mu^2(\Delta)}{2}$$

où

$$\mu^2(\Delta) = \sum_{\omega} \mu^2(\omega)$$

□

Ainsi, un ordre central maximise le critère suivant :

$$L_1(O, X) = \langle I_O, I \rangle_{\mu^2} = \sum_{\omega,\omega'} I_O(\omega, \omega') I(\omega, \omega') \mu(\omega) \mu(\omega') \tag{16}$$

Nous pouvons généraliser le τ de Kendall en posant, pour deux ordres totaux O et O' :

$$\tau_\mu(O, O') = 1 - \frac{2\mu^2(I_O \Delta I_{O'})}{1 - \mu^2(\Delta)} \tag{17}$$

Remarques :

(1) Si on prend pour m la mesure uniforme, on retrouve le τ de Kendall.

(2) Il est curieux de constater qu'on ne rencontre pas, dans les ouvrages classiques, la définition (17). Et pourtant, la statistique est une science qui cherche à extraire de *l'information* c'est-à-dire à échapper à la mesure la moins informative, la mesure uniforme.

□

Avec cette définition, nous obtenons le résultat suivant : un ordre central maximise le critère :

$$L_1''(O, X) = \sum_i \nu(i)\tau_\mu(O, O_i) \tag{18}$$

Nous pouvons étendre l'algorithme des transferts à la recherche des ordres centraux. Soit O l'ordre total suivant :

$$\omega_1 < \ldots < \omega_i < \omega_{i+1} < \ldots < \omega_n$$

et soit O', l'ordre total obtenu en échangeant ω_i et ω_{i+1} :

$$\omega_1 < \ldots < \omega_{i+1} < \omega_i < \ldots < \omega_n$$

Nous avons donc :

$$L_1(O', X) - L_1(O, X) = \mu(\omega_i)\mu(\omega_{i+1})(I(\omega_{i+1}, \omega_i) - I(\omega_i, \omega_{i+1})) \tag{19}$$

II. Partie pratique

II.1 Etude du Système Scolaire Camerounais

Le tableau (1) donne les ratios *Elèves/Salles* pour les 10 provinces du Cameroun et les trois systèmes scolaires

TABLEAU 1
Ratios Elèves/Salles pour le Cameroun

	Primaire	Secondaire Général	Secondaire Technique
Adamaoua	56	54	34
Centre	48	70	43
Est	40	43	35
Extrême Nord	66	55	35
Littoral	59	61	44
Nord	67	71	43
Nord Ouest	50	62	36
Ouest	50	50	28
Sud	36	47	37
Sud Ouest	55	56	44

(a) Quartiles, Mesures uniformes

En utilisant les quartiles, nous transformons les trois caractères numériques en 4 classes (cf tableau (2)). Nous utilisons pour μ et ν des mesures uniformes.

TABLEAU 2
Transformation du tableau (1) en utilisant les quartiles

	Primaire	Secondaire Général	Secondaire Technique
Adamaoua	3	2	1
Centre	1	4	3
Est	1	1	1
Extrême Nord	4	2	1
Littoral	3	3	4
Nord	4	4	3
Nord Ouest	2	3	2
Ouest	2	1	1
Sud	1	1	3
Sud Ouest	3	3	4

En partant de la partition la plus fine, nous obtenons la partition :

$$P_1 : \{A, EN\} + \{C, N, S\} + \{E, O\} + \{L, SO\} + \{NO\}$$
$$L(P_1; \wp) = 0.020$$

En partant de la partition la plus grossière, nous obtenons la partition :

$$P_2 : \{A, EN\} + \{C, N\} + \{E, O, S\} + \{L, SO\} + \{NO\}$$
$$L(P_2; \wp) = 0.020$$

(b) Médianes, Mesures uniformes

En utilisant la médiane, nous transformons les trois caractères numériques en 3 caractères à 2 classes (dans le tableau (2), on effectue les transformations : 1,2 $\rightarrow$ 1 et 3,4 $\rightarrow$ 2. Nous utilisons pour μ et ν des mesures uniformes.

En partant de la partition la plus fine, nous obtenons la partition :

$$P_3 : \{A, EN\} + \{C, L, N, SO\} + \{E, NO, O, S\}$$
$$L(P_3; \wp') = 0.070$$

En partant de la partition la plus grossière, nous obtenons la partition :

$$P_4 : \{C, L, N, SO\} + \{A, E, EN, O, S\} + \{NO\}$$
$$L(P_4; \wp') = 0.073$$

(c) Quartiles, mesures non uniformes

Nous utilisons les quartiles et les mesures suivantes (avec des notations évidentes) :

$$\mu(A) = 0.0718, \mu(C) = 0.0722, \mu(E) = 0.1161, \mu(EN) = 0.1965,$$
$$\mu(L) = 0.0232, \mu(N) = 0.2902, \mu(NO) = 0.0175, \mu(O) = 0.0266,$$
$$\mu(S) = 0.1528, \mu(SO) = 0.0330$$
$$\nu(P) = 0.2659, \nu(SG) = 0.3851, \nu(ST) = 0.3490$$

Remarques

La mesure μ est construite à partir du *ratio Elèves-francophones/Elèves anglophones* et la mesure ν à partir du *ratio Elèves Publics/Elèves Privés.*

□

En partant de la partition la plus fine ou de la plus grossière, nous obtenons la partition :

$$P_5 : \{A, EN\} + \{C, N\} + \{E, O, S\} + \{L, SO\} + \{NO\}$$
$$L(P_5; \wp) = 0.023$$

(d) Médianes, mesures non uniformes

Nous utilisons finalement la médiane et les mesures μ et ν du cas (c).

En partant de la partition la plus fine ou de la plus grossière, nous obtenons la partition :

$$P_6 : \{A, E, EN, O\} + \{C, L, N, SO\} + \{NO\} + \{S\}$$
$$L(P_6; \wp') = 0.064$$

Commentaires

1) Nous vérifions bien sur ces cas le théorème (2), par exemple :

$$L(P_1; \wp) = L(P_2; \wp) = 0.020 < L(P_3; \wp') = 0.070 < L(P_4; \wp') = 0.073$$

Remarques

Les partitions $P_1, ..., P_4$ obtenues sont des optima locaux et non globaux. Or le théorème (2) s'applique aux optima globaux. Ainsi, on aurait pu ne pas avoir ces inégalités.

□

2) L'introduction de mesures non uniformes n'a pas tellement de conséquences sur le résultat, par exemple :

$$P_2 = P_5$$

3) Dans le cas (a), nous obtenons deux optima locaux avec même valeur du critère. Par contre, dans le cas (b), nous obtenons deux optima locaux avec des valeurs différentes du critère.

4) Quand on codifie les caractères (Quartiles → Médianes), on tend à diminuer le nombre de classes, ce qui semble normal. Mais on n'a pas de relation précise entre les différentes partitions obtenues.

II.2 Etude Energétique du Cameroun

Le tableau (3) donne un certain nombre de *caractères énergétiques* mesurés sur 32 villes du Cameroun.

(a) Quartiles, Mesures uniformes

Nous transformons les 4 caractères numériques en 4 caractères à 4 classes en utilisant les quartiles et nous prenons pour μ et ν, les mesures uniformes.

TABLEAU 3
Mesures Energétiques sur 32 villes du Cameroun

	VILLES	T.S.M.	H.R.M.	H.A.	ALT.	M.P.C.	I.S.R.	VENT
1	ABONG-MBANG	29,9	80	15,8	693	Mars	Faible	Moyen
2	AKONOLINGA	30,5	79	16,9	671	Février	Faible	Moyen
3	BAFIA	32,7	78	16,6	500	Février	Faible	Moyen
4	BAMENDA	26,2	74	12,6	1608	Février	Fort	Moyen
5	BANYO	32	57	11,2	1110	Février	Fort	Moyen
6	BATOURI	31,2	79	15	650	Février	Faible	Moyen
7	BERTOUA	31,1	80	16	663	Février	Faible	Moyen
8	BETARE-OYA	31,7	77	15,3	815	Février	Fort	Fort
9	DOUALA	31,8	83	18,1	5	Février	Faible	Moyen
10	DSCHANG	27,3	77	13,3	1407	Février	Faible	Moyen
11	EBOLOWA	29,8	82	16,3	628	Février	Faible	Moyen
12	EDEA	32,8	82	18,1	31	Février	Faible	Calme
13	ESEKA	31	84	21,1	398	Février	Faible	Calme
14	GAROUA	39,8	56	13,6	241	Mars	Fort	Fort
15	KAELE	39,1	49	12,1	389	Mars	Fort	Fort
16	KOUNDJA	29,9	73	13,3	1208	Février	Fort	Calme
17	KRIBI	30,4	86	18,3	10	Mars	Faible	Moyen
18	LOMIE	29,6	83	15,9	624	Avril	Faible	Moyen
19	MANFE	33,6	81	17,9	126	Février	Faible	Calme
20	MAROUA	39,3	48	11,8	423	Avril	Fort	Fort
21	MEIGANGA	32	68	13,4	1027	Février	Fort	Fort
22	MANGA-EBOKO	31,6	74	12,2	622	Février	Faible	Moyen
23	NGAMBE	29,7	86	17,1	610	Février	Faible	Moyen
24	NGAOUNDERE	31,7	65	12,4	1115	Mars	Fort	Fort
25	NKONGSAMBA	28,3	83	16	816	Février	Faible	Moyen
26	POLI	37,4	65	14,2	436	Mars	Fort	Fort
27	SANGMELIMA	29,5	81	16	712	Février	Faible	Moyen
28	TIBATI	33,3	68	13,6	873	Février	Fort	Moyen
29	TIKO	32	84	18,1	50	Février	Faible	Moyen
30	YAOUNDE	29,9	79	15,8	753	Février	Faible	Moyen
31	YOKADOUMA	30,8	81	16,7	634	Février	Faible	Moyen
32	YOKO	30,5	75	15,3	1027	Février	Fort	Moyen

T.S.M. : Température Sèche Maximum (°C)
H.R.M. : Humidité Relative Moyenne (%)
H.A. : Humidité Absolue (g/kg air sec)
ALT : Altitude (m)
M.P.C. : Mois le Plus Chaud
I.S.R. : InSolation Relative

En partant de la partition la plus grossière, nous obtenons la partition suivante :

$$P_1 : \{1,2,3,6,7,11,18,22,23,25,27,30,31\} + \{4,5,10,16,21,28,32\} + \{8\} + \{9,12,13,17,19,29\} + \{14,15,20,24,26\}$$
$$L(P_1;\wp) = 0.028$$

La figure (1) représente cette partition (cf. Appendice).

(b) Médianes, Mesures uniformes

Nous utilisons maintenant les médianes à la place des quartiles, en conservant pour μ et ν les mesures uniformes.

En partant de la partition associée au caractère H.R.M., nous obtenons la partition suivante :

$$P_2 : \{1,2,3,7,9,11,12,13,17,18,19,23,25,27,29,31\} + \{4,5,6,8,10,14,15,16,20,21,22,24,26,28,30,32\}$$
$$L(P_2;\wp') = 0.089$$

La figure (2) représente cette partition (cf. Appendice).

Commentaires

La partition P1 correspond à la division climatique suivante :

– climat équatorial de type camerounien. Précipitations : 2000-10000 mm $\{9,12,...,29\}$

– climat équatorial de type guinéen. Précipitations : 1500-2000 mm $\{1,2,...,31\}$

– climat tropical de type soudanien. Précipitations : 900-1500 mm $\{4,5,...,32\}$

– climat tropical de type sahélien. Précipitations : 400-900 mm $\{14,15,...,26\}$

La partition P_2 correspond à la division climatique suivante :

– climat équatorial : $\{1,2,3,...,31\}$

– climat tropical : $\{4,5,6,...,32\}$

II.3 Classification des Félidés

Considérons l'exemple traité par Marcotorchino et Michaud (cf [6]). Ω est une population de 30 félins sur laquelle sont mesurés 14 caractères.

En partant des partitions associées aux caractères X_1, X_9, X_{10} X_{11}, X_{12}, X_{13} ainsi que de la partition la plus fine nous obtenons la partition suivante :

$$P_1 : \{1,2\} + \{3,4,5,7,8,11\} + \{6,9\} + \{10,12,13,14,15,16,17,18,$$
$$19,20,21,22,23,24,25,26,27,28,29,30\}$$
$$L(P_1;\wp) = 0.103$$

En partant des partitions associées aux caractères X_2, X_3, X_5 X_7, X_8, X_{14} ainsi que de la partition la plus grossière nous obtenons la partition suivante :

$$P_2 : \{1,2,3,4\} + \{5,7,8,11\} + \{6,9\} + \{10,12,13,14,15,16,17,18,$$
$$19,20,21,22,23,24,25,26,27,28,29,30\}$$
$$L(P_2;\wp) = 0.103$$

En partant de la partition associée au caractère X_4, nous obtenons la partition suivante :

$$P_3 : \{1,2\} + \{3,4,5,7,8,11\} + \{6\} + \{9,10,12,13,14,15,16,17,18,$$
$$19,20,21,22,23,24,25,26,27,28,29,30\}$$
$$L(P_3;\wp) = 0.103$$

En partant de la partition associée au caractère X_6, nous obtenons la partition suivante :

$$P_4 : \{1,2,3,4,5,7,8\} + \{6\} + \{9,11\} + \{10,12,13,14,15,16,17,18,$$
$$19,20,21,22,23,24,25,26,27,28,29,30\}$$
$$L(P_4;\wp) = 0.103$$

La partition P_3 est celle obtenue par Marcotorchino et Michaud. Si cette partition est centrale, alors il y a trois autres partitions centrales.

Remerciements

L'auteur remercie P. Delfour qui a programmé l'algorithme des transferts, P. Cazes pour son amical et constant soutien et un REFEREE qui a considérablement amélioré la première version de ce travail.

Bibliographie

[1] BARTHELEMY J. P. and MONJARDET B. (1988), "The median procedure in data analysis : new results and open problems" in classification and related methods in data analysis. Bock (ed). North Holland.

[2] CAZES P. (1990), «Codage d'une variable continue en vue de l'analyse des correspondances». Rev. Stat. App. 38 – n° 3 – pp 35-51.

[3] COMTET L. (1970), «Analyse Combinatoire». P.U.F. Tome II.

[4] LERMAN I. C. (1981), «Classification et analyse ordinale des données». Dunod.

[5] MARCOTORCHINO J. F. and MICHAUD P. (1981), "Heuristic Approach to the similarity aggregation problem". Methods of Operation Research 43 – pp 395-404.

[6] MARCOTORCHINO J. F. et MICHAUD P. (1982), «Agrégation des similarités en classification automatique». Rev. Stat. App. 30 – n°2 – pp 21-44.

[7] MONJARDET B. (1990), «Sur diverses formes de la 'Règle de Condorcet' d'agrégation des préférences». Math. Sci. Hum. 111-Automne-pp 61-71.

[8] REGNIER S. (1983), «Sur quelques aspects mathématiques des problèmes de classification automatique». Math. Sci. Hum. 82 pp 13-29.

[9] SAKAROVITCH M.(1984), «Optimisation combinatoire. Programmation discrète». Hermann

Appendice

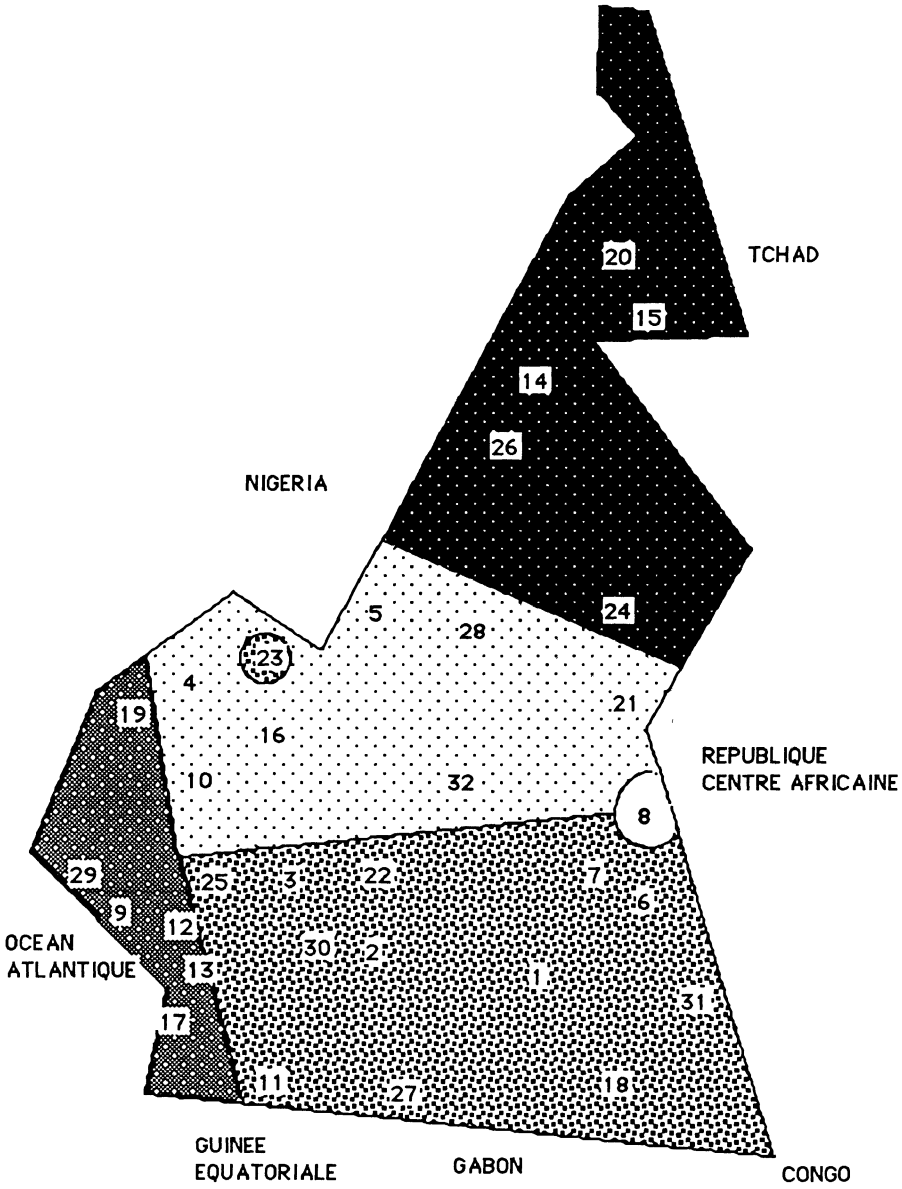


FIGURE 1
Partition Energétique du Cameroun en utilisant les Quartiles

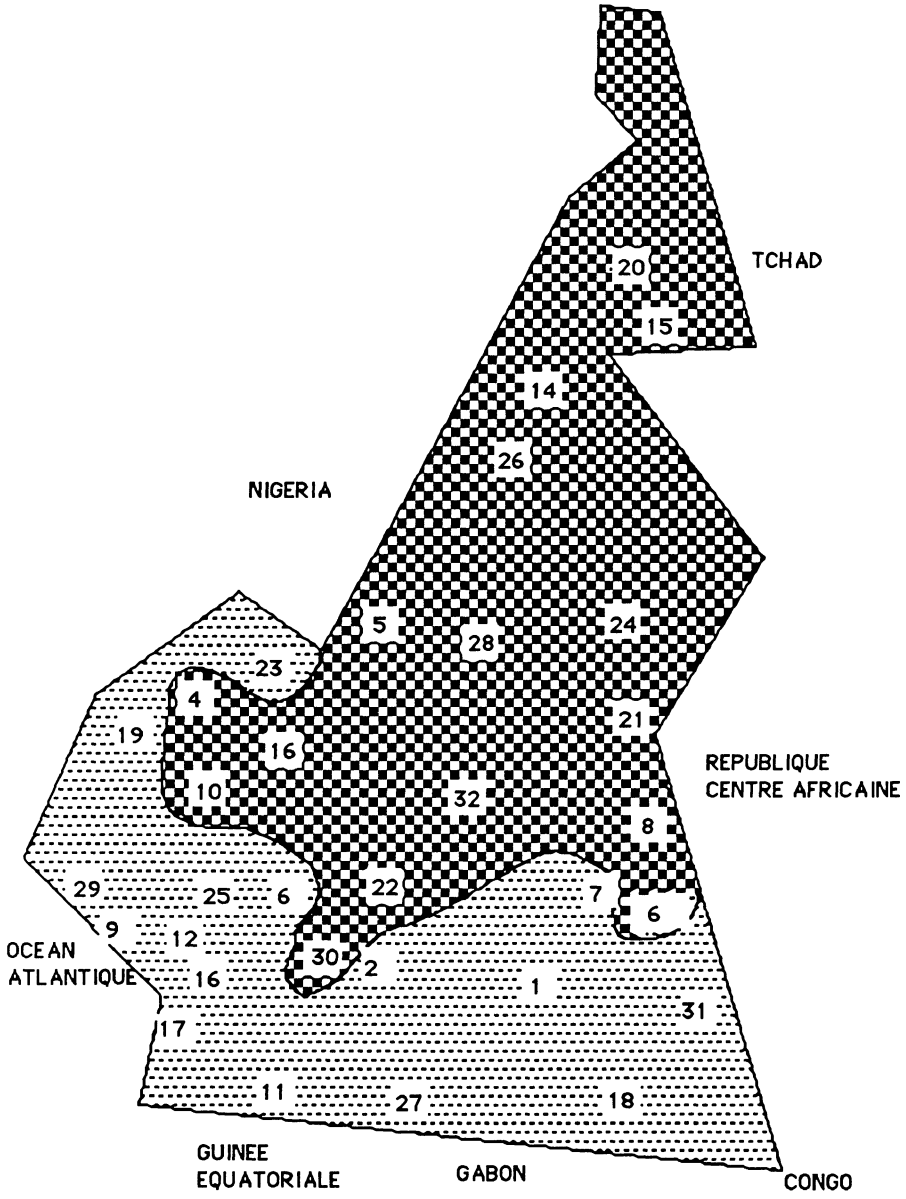


FIGURE 2
Partition Energétique du Cameroun en utilisant les Médianes