Les ultramétriques supérieures minimales d'une dissimilarité à valeurs dans un treillis distributif

Minimal superior ultrametrics of a dissimilarity valued into a distributive lattice

Taoufik Benkaraache

Édition électronique

URL : <http://journals.openedition.org/msh/2799>

DOI : 10.4000/msh.2799

ISSN : 1950-6821

Éditeur

Centre d'analyse et de mathématique sociales de l'EHESS

Édition imprimée

Date de publication : 1 septembre 2000

ISSN : 0987-6936

Référence électronique

Taoufik Benkaraache, « Les ultramétriques supérieures minimales d'une dissimilarité à valeurs dans un treillis distributif », *Mathématiques et sciences humaines* [En ligne], 151 | Automne 2000, mis en ligne le 10 février 2006, consulté le 23 juillet 2020. URL : <http://journals.openedition.org/msh/2799> ; DOI : <https://doi.org/10.4000/msh.2799>

LES ULTRAMÉTRIQUES SUPÉRIEURES MINIMALES D'UNE DISSIMILARITÉ À VALEURS DANS UN TREILLIS DISTRIBUTIF

Taoufik BENKARAACHE¹

RÉSUMÉ – *L'objectif de ce papier est de montrer comment se généralise la construction des ultramétriques supérieures minimales d'une dissimilarité quand cette dernière à ses valeurs dans un ensemble L non nécessairement totalement ordonné. Nous étudions spécialement le cas des dissimilarités à valeurs dans un treillis distributif.*

MOTS-CLÉS – Ensemble ordonné, Dissimilarité, Ultramétrique, Ultramétrique supérieure minimale.

SUMMARY – Minimal superior ultrametrics of a dissimilarity valued into a distributive lattice
The aim of this paper is the generalisation of the calculus of dissimilarity approximation by minimal superior ultrametrics when the given dissimilarity takes values in a set L , not necessarily linearly ordered. We especially consider the case where L is a distributive lattice.

KEYWORDS – Ordered set, Dissimilarity, Ultrametric, Minimal superior ultrametric.

Article communiqué par B. Leclerc.

1. INTRODUCTION

L'importance des considérations ordinales en analyse des données et plus particulièrement en classification est certaine. Dès que les objets et les caractères étudiés sont de types variés, la description complète d'un objet se traduit par sa valeur dans l'ensemble L ordonné, qui est un produit d'ensembles de valeurs de chaque caractère étudié. On est donc amené à considérer des modèles de classification plus généraux que les modèles habituels où les valeurs prises par les objets (ou paires d'objets) sont dans un ensemble totalement ordonné (LTO) qui est souvent une partie de $\mathbb{R}^+$.

Plusieurs auteurs ont développé une méthodologie qui vise à obtenir des modèles de classifications à partir de dissimilarités en mettant l'accent surtout sur l'aspect ordinal des données. Ainsi le modèle «résiduel» introduit par Janowitz(1978), puis par Barthélemy et al.(1984) et Leclerc (1994) permet de retrouver divers modèles de sériation et de classification de base, et de rendre compte de divers types d'objets valués de la littérature. En particulier, les applications à seuils (ou résiduelles), qui associent à une valeur (ou un seuil) $\square$ de L un sous-ensemble d'objets dont les valeurs sont au plus égales à $\square$, généralisent par exemple les ultramétriques. Dans le cas des dissimilarités à valeurs dans

¹ Faculté d'économie et de droit, B.P. 145, Mohammédia, Maroc, e-mail : btaoufik@hotmail.com.

$\mathbb{R}^2$, nous avons montré (Benkaraache & Van Cutsem 1994) qu'une ultramétrie à valeurs dans $\mathbb{R}^2$ est équivalente à un couple d'ultramétries habituelles (à valeurs dans $\mathbb{R}$), et nous avons appliqué ce résultat à la comparaison des ultramétries et par suite à la comparaison des structures hiérarchiques.

Dans le cadre de ces généralisations, nous avons montré (Benkaraache 1998) comment se généralise le calcul de l'ultramétrie inférieure maximum d'une dissimilarité à valeurs dans un inf-demi-treillis fini. Il est naturel de se poser la question de la possibilité d'étendre le calcul des ultramétries supérieures minimales à une dissimilarité d qui est à valeurs dans un ensemble L ordonné d'une certaine manière. Le cas habituel des dissimilarités à valeurs dans $\mathbb{R}^+$ a été étudié par plusieurs chercheurs. Gondran (1976) qui a donné une caractérisation algébrique et qui permet, en calculant un nombre fini des puissances successives de la matrice associée à d , de trouver simplement une ultramétrie supérieure minimale. Defays (1977) propose un algorithme pour le calcul de l'ultramétrie du lien complet dont la minimalité est prouvée par exemple dans Van Cutsem (1984). Leclerc (1986) propose une caractérisation descendante de telles ultramétries et étudie en détail le cas où la dissimilarité est injective : dans ce cas il y a $(n - 1)!$ ultramétries supérieures minimales distinctes (n étant le nombre d'objets à classer). etc.

L'objectif du présent travail est d'étendre les méthodes de calcul des ultramétries supérieures minimales au cas où la dissimilarité est à valeurs dans ensemble ordonné plus général (un treillis distributif).

Le deuxième paragraphe reprend des notions essentielles sur les ensembles ordonnés et les ultramétries à valeurs dans un treillis distributif qui nous serviront dans la suite. Le troisième paragraphe sera consacré à l'étude des ultramétries supérieures minimales. Nous proposerons une méthode descendante de construction de telles ultramétries et nous terminerons par des exemples.

2. DISSIMILARITÉS ET ULTRAMÉTRIQUES À VALEURS DANS UN ENSEMBLE ORDONNÉ

2.1. Rappels sur les ensembles ordonnés et les ultramétries

Soit S un ensemble fini de cardinal n et L un ensemble partiellement ordonné admettant un minimum noté 0 . La relation d'ordre partiel sur L (réflexive, transitive et antisymétrique) est notée $\leq$. Un majorant (resp. minorant) d'une partie de L est un élément de L qui est supérieur (resp. inférieur) ou égal à tous les éléments de cette partie. Le supremum (resp. infimum) d'une partie L (quand il existe) est le plus petit des majorants (resp. plus grand des minorants) communs à cette partie.

Un treillis est un ensemble ordonné où toute paire d'éléments admet un infimum (inf-demi-treillis) et un supremum (sup-demi-treillis). Nous utilisons dans la suite les notations $\inf(x, y) = x \wedge y$ et $\sup(x, y) = x \vee y$.

La notation $\square_1 \not\leq \square_2$ signifie $\square_1$ n'est pas inférieur ni égal à $\square_2$ pour l'ordre $\leq$.

Nous dirons que l'élément $\square_2$ couvre l'élément $\square_1$ (ou que $\square_1$ est couvert par $\square_2$) si $\square_1 < \square_2$ et il n'y a pas d'éléments compris entre $\square_1$ et $\square_2$. Un prédécesseur (resp. successeur) d'un élément $\square$ est alors un élément de L qui est couvert par (resp. qui

couvre) $\square$. L'ensemble des prédécesseurs (resp. successeurs) de $\square$ est noté $\text{Pred}(\square)$ (resp. $\text{Succ}(\square)$).

Une dissimilarité d sur S et à valeurs dans L (ou L -dissimilarité) est une application de $S \times S$ dans L qui vérifie $d(a,a) = 0$ pour tout $a \in S$. Nous ne considérons que les L -dissimilarités symétriques, i.e. les L -dissimilarités d vérifiant $d(a,b) = d(b,a)$ pour tous $a, b \in S$.

Une L -dissimilarité d est dite propre si $d(a,b) = 0$ implique $a = b$.

L'ensemble des L -dissimilarités sur S est muni de l'ordre usuel défini comme suit : si d_1 et d_2 sont deux L -dissimilarités sur S , $(d_1 \leq d_2)$ si $d_1(a,b) \leq d_2(a,b)$ pour tous $a, b \in S$. La dominance est stricte s'il existe un couple d'éléments de S pour lequel l'inégalité précédente est stricte.

DÉFINITION 1. Soit S un ensemble fini et L un ensemble ordonné.

Une L -dissimilarité u sur S est une L -ultramétrie sur S si : pour tout $\square \in L$, pour tout $(a,b,c) \in S \times S \times S$

$$[u(a,b) \leq \square \text{ et } u(b,c) \leq \square] \implies u(a,c) \leq \square$$

On dit aussi que le triplet (a,b,c) est u -ultramétrique.

Cette définition est équivalente à ce que tout majorant de deux des trois éléments $u(a,b)$, $u(b,c)$ et $u(a,c)$ est un majorant du troisième.

Dans le cas d'une dissimilarité à valeurs dans un ensemble totalement ordonné (LTO), la définition de l'ultramétrie devient équivalente à la définition habituelle suivante : pour tous $a,b,c \in S$, $u(a,c) \leq \max(u(a,b), u(b,c))$.

Si L est un produit fini d'ensembles LTO, on montre facilement qu'une L -dissimilarité est une L -ultramétrie si et seulement si toutes ses composantes sont des ultramétries (Benkaraache 1998).

Une L -ultramétrie peut-être caractérisée en utilisant les partitions induites à chaque valeur de L (Critchley et Van Cutsem 1994).

PROPOSITION 1. Soit S un ensemble quelconque, L un inf-demi-treillis complet, et u une L -dissimilarité sur S . Les deux propositions suivantes sont équivalentes :

1. u est une L -ultramétrie.
2. Pour tout $\square \in L$, la relation binaire $P_\square$ définie par $a P_\square b$ si et seulement si $u(a,b) \leq \square$ est une relation d'équivalence.

Dans le cas où L est un treillis complet, on retrouve la caractérisation d'une L -ultramétrie par les applications à seuils (ou résiduelles) (Leclerc 1994).

Une L -ultramétrie définit donc, pour chaque élément $\square$ de L , une partition dont les classes sont les classes d'équivalence de la relation $P_\square$. Cependant, l'ensemble de partitions $P_\square$ ne définit une L -ultramétrie que si pour tout couple d'éléments $a,b \in S$, l'ensemble $\{\square \in L / a P_\square b\}$ a un minimum. L'ultramétrie induite dans ce cas est définie par $u(a,b) = \min \{\square \in L / a P_\square b\}$.

Dans ce qui suit, nous assimilerons toute relation d'équivalence $P_\square$ à la partition en classes d'équivalence qu'elle induit sur S .

Un infimum de partitions, noté $\sqcap P_{\square}$, est la partition obtenue par croisement des partitions $P_{\square}$, c'est-à-dire l'intersection de leurs classes. En particulier, quand l'ensemble L est un produit d'ensembles LTO ($L = L_1 \times L_2$), la partition $P_{\square}$ induite par une ultramétrie u au seuil $\square = (\square_1, \square_2)$ est $P_{\square} = P_{\square_1} \sqcap P_{\square_2}$.

Pour un élément a de S , la classe de P_x qui contient a est la boule (au sens de u) de centre a et de rayon $\square$: $B_u(a, \square) = \{b \in S / u(a, b) \leq \square\}$.

Nous avons aussi la caractérisation suivante des L -ultramétries.

PROPOSITION 2. Soit S un ensemble fini et L un treillis. Les propositions suivantes sont équivalentes :

(P1) u est une L -ultramétrie sur S .

(P2) Pour tous éléments $a, b, c \in S$

$$\begin{aligned} u(a, b) \vee u(a, c) &= u(a, b) \vee u(b, c) \\ &= u(a, c) \vee u(b, c) \\ &= u(a, b) \vee u(a, c) \vee u(b, c) \end{aligned}$$

(P3) Pour tous éléments $a, b, c \in S$, $u(a, b) \leq u(a, c) \vee u(b, c)$

Démonstrations.

(P1) $\Leftrightarrow$ (P3) : il suffit de prendre dans la définition 1 : $\square = u(a, c) \vee u(b, c)$.

(P3) $\Leftrightarrow$ (P2) : en effet, $u(a, b) \leq u(a, c) \vee u(b, c) \Leftrightarrow u(a, c) \vee u(a, b) \leq u(a, c) \vee u(b, c)$

Et en interchangeant les rôles de a, b , et c on obtient l'inégalité inverse. De même pour la deuxième égalité de (P2).

(P2) $\Leftrightarrow$ (P1) : évidente. □

2. 2. Configurations d'un triplet L -ultramétrie

Si L n'est pas nécessairement LTO, les trois éléments $u(a, b)$, $u(b, c)$ et $u(a, c)$ peuvent présenter des configurations géométriques autres que celle du triangle ultramétrique habituel : en effet, d'après les propriétés précédentes, un triplet (a, b, c) est u -ultramétrique si on a l'une des trois configurations suivantes (nous représentons deux éléments de L comparables pour l'ordre $\leq$ comme étant liés par un chemin d'éléments de L qui est totalement ordonné par $\leq$, l'élément inférieur se trouvant plus bas, avec $\{u_1, u_2, u_3\} = \{u(a, b), u(a, c), u(b, c)\}$) :

Toute autre configuration ne représente pas un triplet u -ultramétrique.

3. ULTRAMÉTRIQUES SUPÉRIEURES MINIMALES D'UNE L -DISSIMILARITÉ

DÉFINITION 2. (ultramétrie supérieure minimale)

Soit d une L -dissimilarité sur S et u une L -ultramétrie telle que $u \geq d$.

L'ultramétrie u est supérieure minimale de d si pour toute ultramétrie v qui vérifie $d \leq v \leq u$, on a $u = v$.

3.1 Cas où L est un produit d'ensembles totalement ordonnés

Supposons que $L = L_1 \square L_2$ où L_i est LTO ($i = 1, 2$).

La L -dissimilarité d admet alors deux composantes $d = (d_1, d_2)$ où d_i est à valeurs dans L_i ($i = 1, 2$).

On a alors le théorème suivant qui caractérise les ultramétries supérieures minimales quand L est un produit d'ensembles LTO.

THÉORÈME 1. On suppose que $L = L_1 \square L_2$ où L_i est (LTO). Soit $d = (d_1, d_2)$ une L -dissimilarité et $u = (u_1, u_2)$ une L -ultramétrie telle que $u \geq d$. On a l'équivalence suivante : u est supérieure minimale de d $\square$ u_i est supérieure minimale de d_i ($i = 1, 2$).

Démonstrations.

$\square$) Soit v une ultramétrie supérieure minimale de d , $v = (v_1, v_2)$. Supposons que v_1 par exemple ne soit pas une ultramétrie supérieure minimale de d_1 : il existe alors une L_1 -ultramétrie w_1 qui vérifie : $d_1 \leq w_1 < v_1$.

L'ultramétrie $w = (v_1, w_1)$ vérifie alors $d \leq w < v$ ce qui contredit le fait que v est minimale.

$\square$) Soient u_1 et u_2 deux ultramétries supérieures minimales de d_1 et d_2 respectivement. On sait que (u_1, u_2) est une L -ultramétrie supérieure à d . Supposons qu'elle ne soit pas minimale : il existe alors une L -ultramétrie v vérifiant $d \leq v < u$, donc aussi un couple $(a_0, b_0) \square S \square S$ tel que :

$$v(a_0, b_0) < u(a_0, b_0)$$

ce qui est équivalent à $[v_1(a_0, b_0) < u_1(a_0, b_0)]$ ou $[v_2(a_0, b_0) < u_2(a_0, b_0)]$.

Donc l'une des deux composantes de u n'est pas minimale. D'où la contradiction. $\square$

Il est facile de voir que le théorème précédent reste vrai dans le cas d'un produit quelconque d'ensembles (LTO) et S fini.

3.2. Cas d'une L -dissimilarité à valeurs dans un treillis distributif fini

Dans toute la suite, on suppose que L est un treillis distributif fini, S est un ensemble fini de cardinal n et d une L -dissimilarité sur S .

On remarque que l'ensemble $F_d = \{u \text{ } L\text{-ultramétrie} / u \geq d\}$ est non vide parce qu'il contient l'ultramétrie constante partout égale à $d_0 = \{d(a, b) / (a, b) \square S \square S\} = \sup d$ (d_0 est bien définie car S est fini et L est un treillis fini).

Il est facile d'établir l'existence des ultramétriques supérieures minimales à d : en effet, soit (u_n) une suite décroissante d'ultramétriques dans F_d . Comme L et S sont supposés finis, cette suite constitue donc un ensemble ordonné fini, elle admet par conséquent au moins un élément minimal.

LEMME 1. Soit u une L -ultramétrique supérieure minimale de d .

Si on pose $u_0 = \{u(a,b) / (a,b) \in S \times S\}$, alors : $d_0 = u_0$.

Démonstration. On a d'abord, par définition de u , pour tout couple $(a,b) \in S \times S$, $u_0(a,b) \leq u(a,b)$. Donc $d_0 \leq u_0$.

Il reste à montrer l'inégalité inverse : $d_0 \geq u_0$. Pour cela, on va montrer que $u(a,b) \leq d_0$ pour tout couple (a,b) . Supposons qu'il existe un couple (a_0, b_0) de S tel que $u(a_0, b_0) \not\leq d_0$: soit alors la partition $P = (B_1, \dots, B_p)$ définie par u au niveau d_0 . Il existe alors deux indices i_0 et j_0 ($i_0 \neq j_0$) tels que $a_0 \in B_{i_0}$ et $b_0 \in B_{j_0}$ puisqu'on a $u(a_0, b_0) \not\leq d_0$. Soit alors la dissimilarité w définie par :

$$w(x, y) = \begin{cases} d_0 & \text{si } x \text{ et } y \text{ sont séparées par les classes de } P \\ u(x, y) & \text{sinon} \end{cases}$$

On va montrer que w est une L -ultramétrique qui vérifie $d \leq w < u$, ce qui contredit le fait que u est une ultramétrique supérieure minimale.

Tout d'abord, pour montrer que w est L -ultramétrique, nous utilisons la propriété P_2 de la proposition 2. Soit $(a, b, c) \in S^3$. Deux cas peuvent se présenter :

$$\begin{aligned} 1^{\text{er}} \text{ cas : Si } (a, b, c) \in B_{i_0} \times B_{j_0} \times B_{i_0} \\ w(a, b) &= u(a, b) \not\leq d_0 \\ w(b, c) &= u(b, c) \not\leq d_0 \\ w(a, c) &= u(a, c) \end{aligned}$$

Et on a donc :

$$\begin{aligned} 1. \quad w(a, b) \quad w(a, c) &= (u(a, b) \not\leq d_0) \quad u(a, c) \\ &= (u(a, b) \quad u(a, c)) \not\leq (u(a, c) \quad d_0) \quad (\text{distributivité de } L) \\ &= (u(a, b) \quad u(a, c)) \not\leq d_0 \\ 2. \quad w(b, c) \quad w(a, c) &= (u(b, c) \not\leq d_0) \quad u(a, c) \\ &= (u(b, c) \quad u(a, c)) \not\leq (u(a, c) \quad d_0) \\ &= (u(b, c) \quad u(a, c)) \not\leq d_0 \\ 3. \quad w(a, b) \quad w(b, c) &= (u(a, b) \not\leq d_0) \quad (u(b, c) \not\leq d_0) \\ &= (u(a, b) \quad u(b, c)) \not\leq (u(a, b) \quad d_0) \not\leq d_0 \not\leq (u(b, c) \not\leq d_0) \\ &= (u(a, b) \quad u(b, c)) \not\leq d_0 \end{aligned}$$

Comme (a, b, c) est un triangle u -ultramétrique, on a, en vertu de la propriété P_2 :

$$u(a, b) \quad u(a, c) = u(a, b) \quad u(b, c) = u(a, c) \quad u(b, c)$$

donc aussi :

$$w(a, b) \quad w(a, c) = w(a, b) \quad w(b, c) = w(a, c) \quad w(b, c)$$

On a donc montré que le triplet (a, b, c) est w -ultramétrique.

2^{ème} cas : Si $(a, b, c) \sqsubseteq B_{i_0} \sqsubseteq B_{j_0} \sqsubseteq B_k : (k \neq i_0 \text{ et } k \neq j_0)$

$$w(a, b) = u(a, b) \sqsubseteq d_0$$

$$w(b, c) = u(b, c) \sqsubseteq d_0$$

$$w(a, c) = u(a, c) \sqsubseteq d_0$$

On a alors :

$$\begin{aligned} w(a, b) \quad w(a, c) &= (u(a, b) \sqsubseteq d_0) \quad (u(a, c) \sqsubseteq d_0) \\ &= (u(a, b) \quad u(a, c)) \sqsubseteq (u(a, c) \quad d_0) \sqsubseteq (u(a, b) \quad d_0) \sqsubseteq (d_0 \quad d_0) \\ &= (u(a, b) \quad u(a, c)) \sqsubseteq d_0 \end{aligned}$$

par un développement analogue, on a aussi

$$w(a, b) \quad w(b, c) = (u(a, b) \quad u(b, c)) \sqsubseteq d_0$$

$$w(b, c) \quad w(a, c) = (u(b, c) \quad u(a, c)) \sqsubseteq d_0$$

et comme (a, b, c) est un triangle u -ultramétrique, on a, en vertu de la propriété P_2 :

$$u(a, b) \quad u(a, c) = u(a, b) \quad u(b, c) = u(a, c) \quad u(b, c)$$

donc aussi :

$$w(a, b) \quad w(a, c) = w(a, b) \quad w(b, c) = w(a, c) \quad w(b, c)$$

Les autres cas ne présentent pas de difficulté pour voir que (a, b, c) est w -ultramétrique. On conclut que w est une L -ultramétrie.

Montrons maintenant que w vérifie $d \leq w < u$: On a $w \geq d$ puisque pour tout couple (a, b) , on a : soit $w(a, b) = u(a, b) \geq d(a, b)$, soit $w(a, b) = u(a, b) \sqsubseteq d_0 \geq d$ puisque $d_0 \geq d(a, b)$.

On a aussi $w \leq u$ par définition de w . De plus, $w(a_0, b_0) < u(a_0, b_0)$; en effet : $w(a_0, b_0) = u(a_0, b_0) \sqsubseteq d_0$. Or, par hypothèse, on a $u(a_0, b_0) \not\sqsubseteq d_0$, donc :

$$u(a_0, b_0) \sqsubseteq d_0 < u(a_0, b_0) \text{ et } u(a_0, b_0) \sqsubseteq d_0 < d_0$$

On conclut enfin que, nécessairement, $u(a, b) \leq d_0$ pour tout $(a, b) \sqsubseteq S \sqsubseteq S$. D'où $u_0 \leq d_0$.

Comme $d_0 \leq u_0$, on a donc l'égalité cherchée. $\square$

Par conséquent, pour la recherche des L -ultramétriques supérieures minimales d'une dissimilarité il suffit de se placer dans l'ensemble $F_0 = \{u \text{ ultramétrique} / d \leq u \leq d_0\}$.

Pour énoncer une condition suffisante pour qu'une ultramétrie supérieure à d soit minimale. Nous avons besoin des deux lemmes suivants :

LEMME 2. Si u et v sont deux L -ultramétriques telles que $u \leq v$, alors : pour tous $\square \sqsubseteq L$ et $a \sqsubseteq S$, on a $B_v(a, \square) \sqsubseteq B_u(a, \square)$.

Démonstration. Soit b un élément de $B_v(a, \square)$, on a alors :

$$v(a, b) \leq \square \quad u(a, b) \leq \square \quad b \sqsubseteq B_u(a, \square) \quad \square$$

LEMME 3. Soit u une L -ultramétrie et $(a,b,c) \in S \times S \times S$; $\alpha \in L$:

$$(u(a,b) \leq \alpha \text{ et } u(b,c) \leq \alpha) \implies u(a,c) \leq \alpha.$$

Démonstration. Supposons $u(a,b) \leq \alpha$, $u(b,c) \leq \alpha$ et $u(a,c) > \alpha$. Les deux inégalités $u(a,c) \leq \alpha$ et $u(a,b) \leq \alpha$ impliquent l'inégalité $u(b,c) \leq \alpha$ car u est une ultramétrie. D'où la contradiction. $\square$

Soit u une ultramétrie supérieure à d . Pour tout élément a de S , $B_u(a, \alpha)$ désigne la classe de la partition associée à u au niveau α et qui contient l'élément a .

Nous allons donner une condition suffisante de la minimalité de u dans $B_u(a, \alpha)$ à l'aide des restrictions de u , définies dans les boules de u -diamètre égal à α qui est inférieur à α .

Si α et β sont deux éléments de L tels que $\alpha \leq \beta$, la partition de $B_u(a, \beta)$ en boules de u -diamètre inférieur à α s'écrit :

$$B_u(a, \beta) = B_u(a_1, \alpha) \sqcup B_u(a_2, \alpha) \dots \sqcup B_u(a_p, \alpha)$$

Pour $i = 1, \dots, p$, on note par :

B_i la boule $B_u(a_i, \alpha)$, (bien entendu on a, pour $i \neq j$, $B_i \cap B_j = \emptyset$)

u_i la restriction de u à la boule B_i

d_i la restriction de d à la boule B_i

THÉORÈME 2.

Soient d une L -dissimilarité sur S , u une ultramétrie supérieure à d , et un seuil $\alpha \in L$:

Alors u est supérieure minimale de d sur la boule $B(a, \alpha)$ si les conditions (i) et (ii) sont satisfaites pour tout $\beta < \alpha$:

(i) u_i est supérieure minimale de d_i sur B_i ; $i = 1, \dots, p$

(ii) Pour tous $i \neq j$, il existe $(b, b') \in B_i \cap B_j$ tel que $d(b, b') \leq \beta$

Démonstration. Tout d'abord, on peut se restreindre aux valeurs α et β tels que $\beta < \alpha$ et il existe un élément a de S tel que $B_u(a, \beta)$ est strictement incluse dans $B_u(a, \alpha)$, sinon, cela veut dire que la boule $B_u(a, \alpha)$ se retrouve en entier au niveau β .

Supposons que les conditions (i) et (ii) sont vérifiées et montrons alors que u est supérieure minimale de d sur la boule $B(a, \alpha)$.

Soit v une ultramétrie telle que $d \leq v \leq u$. Montrons alors que nécessairement $v = u$ (*).

D'après (i), la restriction u_i de u à chaque boule B_i est minimale de d_i , donc $u_i = v_i$ $i = 1, \dots, p$ (v_i étant la restriction de v à la boule B_i).

Pour établir l'égalité (*) il faut montrer que pour tout a de S :

$$B_u(a, \alpha) = B_v(a, \alpha)$$

car deux L -ultramétries sont égales si elles induisent les mêmes partitions en α .

On a d'après le Lemme 2, $B_i \sqsubset B_i^\vee$; $i = 1, \dots, p$, où $B_i^\vee$ est la boule (au sens de ν) de diamètre $\sqsupset$ et de centre a_i . Supposons que cette inclusion est stricte pour un indice i_0 : il existe alors $x \sqsubset S$ tel que $x \sqsubset B_{i_0}^\vee$ et $x \sqsubset B_{i_0}$. Ceci est équivalent à :

$$\nu(a_{i_0}, x) \leq \sqsupset \text{ et } u(a_{i_0}, x) \sqsupset \sqsupset \quad (**)$$

Il existe donc un indice $i_1 \neq i_0$ tel que $x \sqsubset B_{i_1}$.

La propriété (ii) affirme qu'il existe un couple $(x_0, y_0) \sqsubset B_{i_0} \sqsubset B_{i_1}$ qui vérifie $d(x_0, y_0) \sqsupset \sqsupset$, donc aussi $\nu(x_0, y_0) \sqsupset \sqsupset$. Il en résulte d'après le Lemme 3 que pour tous $(x, y) \sqsubset B_{i_0} \sqsubset B_{i_1}$, on a $\nu(x, y) \sqsupset \sqsupset$. En particulier, $\nu(a_i, x) \sqsupset \sqsupset$, ce qui contredit (**).

On conclut que pour tous $i = 1, \dots, p$: $B_i = B_i^\vee$. $\square$

3.3. Démarche descendante pour la construction des ultramétriques supérieures minimales

Connaissant la partition $P_\sqsupset$ en boules de u -diamètre inférieur à $\sqsupset$ d'une ultramétrique supérieure minimale u , le théorème 2 précédent permet d'avoir les partitions associées à u à un niveau $\sqsupset'$ tel que $\sqsupset < \sqsupset'$.

Nous allons montrer que pour un élément $\sqsupset'$ de L , la partition $P_{\sqsupset'}$ se déduit facilement à partir des partitions $(P_{\sqsupset} / \sqsupset' \sqsubset \text{Succ}(\sqsupset))$.

PROPOSITION 3. Soit u une ultramétrique supérieure minimale de d et $\sqsupset \sqsubset L$. $P_\sqsupset$ est la partition en boules de u -diam $\leq \sqsupset$.

Si l'ensemble $\text{Succ}(\sqsupset)$ contient plus d'un élément, alors $P_{\sqsupset'} = \sqcup \{P_{\sqsupset} / \sqsupset' \sqsubset \text{Succ}(\sqsupset)\}$ où $\sqcup$ désigne l'infimum de partitions défini dans le deuxième paragraphe.

Démonstration. Soient $\sqsupset_1, \dots, \sqsupset_p$ les éléments de $\text{Succ}(\sqsupset)$.

On a alors : $\sqsupset = \sqsupset_1 \sqsubset \dots \sqsubset \sqsupset_p$.

Montrons que pour tout a de S , on a :

$$B_u(a, \sqsupset) = B_u(a, \sqsupset_1) \sqsubset B_u(a, \sqsupset_2) \sqsubset \dots \sqsubset B_u(a, \sqsupset_p)$$

Inclusion directe

Comme u ultramétrique et $\sqsupset \leq \sqsupset_i$; $i = 1, \dots, p$ on a pour tout a de S :

$$B_u(a, \sqsupset) \sqsubset B_u(a, \sqsupset_i) ; i = 1, \dots, p$$

Inclusion inverse

$$x \sqsubset (B_u(a, \sqsupset_i) ; i = 1, \dots, p) \sqsubset u(a, x) \leq \sqsupset_i ; i = 1, \dots, p)$$

i.e. $u(a, x) \leq \sqsupset_1 \sqsubset \dots \sqsubset \sqsupset_p = \sqsupset$ donc $x \sqsubset B_u(a, \sqsupset)$. $\square$

Remarque. Si $\text{Succ}(\sqsupset)$ contient un seul élément $\sqsupset'$, une partition de la boule $B_u(a, \sqsupset)$ au seuil $\sqsupset'$ peut-être obtenue à l'aide du Théorème 2, en séparant tous les éléments de S qui ne sont pas à une distance inférieure à $\sqsupset'$ au sens de d .

En résumé, les résultats établis ci-dessus (Lemme 1, Théorème 2 et Proposition 3) nous permettent de construire une partie des ultramétriques supérieures minimales à une L -dissimilarité avec une démarche descendante : la construction des partitions associées à

une telle ultramétrie ne concerne que les seuils $\square$ compris entre $\sup d$ et $\min d = 0$: au seuil $\square = \sup d$ tous les éléments de S sont dans une même classe, et on parcourt les seuils $\square$ inférieurs à $\sup d$ en appliquant, soit la proposition 3 quand il y a plus d'un successeur, soit la condition suffisante du théorème 2 quand il y a un seul successeur.

On peut remarquer que le cas connu où L est un ordre total (Leclerc 1986) se retrouve dans cette généralisation : en effet, dans un ensemble LTO, tout élément a un seul successeur (la proposition 3 ne concerne donc pas le cas LTO) et on retrouve comme seules conditions de minimalité celles formulées dans le théorème 2, le signe « $\square$ » étant équivalent à « $>$ » dans le cas LTO.

3.4. Exemples

Exemple 1

Soit l'ensemble $S = \{a, b, c, d, e, f\}$ et la L -dissimilarité propre et symétrique suivante :

	b	c	d	e	f
a	6	7	12	3	10
b		10	8	2	8
c			11	5	4
d				7	5
e					7

où l'ensemble L est défini par la Figure 1 suivante :

Figure 1. L est produit de deux ensembles LTO (exemple 1)

On a donc $d_0 = \sup d = 15$.

Une méthode directe pour obtenir une ultramétrie supérieure minimale est d'appliquer le théorème 1 (L produit d'ensembles LTO). Mais utilisons plutôt la méthode descendante.

Plusieurs choix sont possibles à chaque élément de L . Par exemple, si $\square = 13$, il y a plusieurs choix possibles de partitions : choisissons la partition (d) (abcef) (voir Figure 2 suivante) :

Figure 2 : Partitions associées à une ultramétrie supérieure minimale de d (les éléments non mentionnés constituent des singletons)

L'ultramétrie supérieure minimale obtenue est :

	b	c	d	e	f
a	6	7	15	11	13
b		10	12	13	8
c			15	5	13
d				15	9
e					10

On peut retrouver ce résultat à l'aide du théorème 1 en écrivant $u = (u_1, u_2)$ et $d = (d_1, d_2)$.

Exemple 2. Treillis distributif quelconque)

Plaçons-nous dans le treillis distributif suivant (Figure 3) qui nous a servi d'exemple de référence dans Benkaraache (1998) :

Figure 3 : L est un treillis distributif (exemple 2)

Soit l'ensemble $S = \{a, b, c, d\}$, et D la L -dissimilarité suivante :

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

On a $\sup(D) = 6$.

Un choix de construction des partitions par la démarche descendante est donné dans la Figure 4 suivante (Π_0 désigne la partition en singletons) :

Figure 4 : Partitions induites par u aux différents niveaux de L .

Ce choix mène à l'ultramétrie supérieure minimale u_1 suivante :

	a	b	c	d
a	0	3	4	6
b		0	4	6
c			0	5
d				0

Un autre choix sur les partitions donne :

Figure 5 : Un deuxième choix possible pour les partitions.

Une deuxième ultramétrique supérieure minimale u_2 est donc :

	a	b	c	d
a	0	4	4	6
b		0	2	5
c			0	6
d				0

D'autres ultramétriques supérieures minimales peuvent être atteintes en variant les choix possibles des partitions, en voici quelques unes :

u_3 :

	b	c	d
a	5	4	4
b		6	6
c			1

u_4 :

	b	c	d
a	6	4	4
b		4	5
c			2

u_5 :

	b	c	d
a	4	4	4
b		6	5
c			2

u_6 :

	b	c	d
a	3	6	6
b		6	6
c			1

$u_7 :$

	b	c	d
a	6	6	4
b		2	5
c			6

 $u_8 :$

	b	c	d
a	5	6	6
b		4	6
c			5

BIBLIOGRAPHIE

BARTHÉLEMY, J.-P., LECLERC, B. et MONJARDET, B., «Ensembles ordonnés et taxonomie mathématique», M. Pouzet and D. Richard (editors), *Orders : Description and roles*, *Annals of Discrete Mathematics* (23), 1984, p. 523-548.

BENKARAACHE, T., *Problèmes de validité en classification et quelques généralisations aux ultramétries à valeurs dans un ensemble ordonné*, Thèse de doctorat, Université Joseph Fourier, Grenoble, 1993.

BENKARAACHE, T., «L'ultramétrie inférieure maximum d'une dissimilarité à valeurs dans un inf-demi-treillis», *Mathématiques, Informatique et Sciences humaines* 143, 1998, p. 27-40.

BENKARAACHE T, VAN CUTSEM, V., «Comparison of hierarchical classifications», E. Diday, Y. Lechevallier, M. Schader, P. Bertrand, B. Burtschy (éditeurs), *New approaches in classifications and data analysis*, Berlin, Springer-Verlag 1994, p. 85-92.

CRITCHLEY, F., VAN CUTSEM, B., «An order-theoretic and generalization of certain fundamental bijection in mathematical classification. I», in Bernard Van Cutsem (editor), *Classification and Dissimilarity Analysis, Lecture Notes in Statistics* 93, New-York, Springer-Verlag, 1994, p. 87-119.

DEFAYS, D., «An efficient algorithm for a complete link method», *The computer journal* (20), n° 4, 1977, p. 364-366.

GONDRAN, M., «La structure algébrique des classifications hiérarchiques», *Annales de l'INSEE*, n° 22-23, 1976, p. 181-190.

JANOWITZ, M.F., «An order theoretic model for cluster analysis», *SIAM J. Appl. Math.* (34), 1978, p. 55-72.

LECLERC, B., «Caractérisation, construction et dénombrement des ultramétries supérieures minimales», *Statistique et Analyse des données*, vol. 1, n° 2, 1986, p. 26-50.

LECLERC, B., «The residuation model for the ordinal construction of dissimilarities and other valued objects», in Bernard Van Cutsem (éditeur), *Classification and Dissimilarity Analysis, Lecture notes in Statistics* 93, New-York, Springer-Verlag, 1994, p. 149-172.

VAN CUTSEM, B., «Ultramétries supérieures minimales et algorithme du lien complet», in *Actes des journées de classification de la Grande-Motte*, C. Perruchet (éditeur), Paris, CNET et SFC, 1985.