MATHÉMATIQUES ET SCIENCES HUMAINES

JEAN-PIERRE BARTHELEMY

Trois propriétés des médianes dans un treillis modulaire

Mathématiques et sciences humaines, tome 75 (1981), p. 83-91

<http://www.numdam.org/item?id=MSH_1981__75__83_0>

Math. Sci. hum. (19^{e}année, n°75, 1981, p.83-91)

TROIS PROPRIETES DES MEDIANES DANS UN TREILLIS MODULAIRE

Jean-Pierre BARTHELEMY*

1 - INTRODUCTION

En statistique descriptive la médiane d'une série (totalement) ordonnée est une valeur centrale classique. L'extension de cette notion aux ordres partiels, si elle se rencontre parfois en Sciences Humaines (nous pensons, en particulier, au choix collectif, c.f. Barthélemy - Monjardet [2]), n'a été étudiée que pour les treillis distributifs (Barbut [1], c.f. aussi Monjardet [4]) où de très beaux résultats ont été obtenus. Les treillis distributifs forment, d'une certaine manière, la généralisation la plus immédiate des ordres totaux. Une autre généralisation serait constituée par les arbres (cette fois-ci, l'ordre est oublié mais la "linéarité" est conservée). Cette dernière voie empruntée par de nombreux chercheurs (c.f. en particulier Slater [6]) a conduit aussi à d'intéressants résultats.

Ainsi, il ne semble pas y avoir à l'heure actuelle de résultats sur les médianes dans des ensembles ordonnés plus généraux que les treillis distributifs. Le but de cet article est de commencer à combler cette lacune en examinant le cas des treillis modulaires.

Explicitons les trois propriétés que nous allons établir ; pour ce faire, mettons au clair quelques notations.
Le graphe (non orienté) de couverture $G(E) = (E, U(E))$ d'un ensemble ordonné fini $(E,\leqslant)$ admet E comme ensemble de sommets et xy est une arête si et seulement si y couvre x (c'est-à-dire $x < y$ et $x < t \leqslant y$ entraîne $t = y$) ou x couvre y.

* Ecole Nationale Supérieure des Télécommunications, Paris.

Une géodésique entre deux éléments x, y de E est un chemin, entre x et y, de G(E) de longueur minimale. La distance canonique d sur E est définie par : d (x, y) = longueur d'une géodésique entre x et y.

Posons $O(E) = \bigcup_{n\in N} E^n$. On définit une application D de E x O (E) dans l'ensemble des réels par la formule :

$$D(x, \pi) = \sum_{i=1}^{p} d(x, t_i),$$

pour $\pi = (t_1, \ldots, t_p)$.

On dit que $m \in E$ est une médiane de π lorsque :

$$D(m, \pi) = \min_{x \in E} D(x, \pi)$$

On note Med (π) l'ensemble des médianes de π.

On dit que (E, $\leqslant$) est semi-modulaire supérieurement lorsque, pour tout triplet (x,y,t) tel que x et y couvrent t, il existe $z \in E$ couvrant x et y. Dualement, E est semi-modulaire inférieurement lorsque pour tout triplet (x,y,z) tel que z couvre x et y il existe $t \in E$ couvert par x et y. E est modulaire s'il est semi-modulaire supérieurement et inférieurement.

Lorsque (E,$\leqslant$) est un inf. demi - treillis, pour chaque $x,y \in E$ et $\pi = (t_1,\ldots,t_p) \in O(E)$ on désigne par $\pi^{\wedge}(x,y)$ le nombre d'indices i tels que $x \wedge t_i = y \wedge t_i$ $(1 \leqslant i \leqslant p)$. Dualement, lorsque (E,$\leqslant$) est un sup. demi . treillis, on pose : $\pi^{\vee}(x,y) = |\{i \mid x \vee t_i = y \vee t_i\}|$.

Supposons que (E,$\leqslant$) est un treillis modulaire. Les trois propriétés que nous nous proposons d'établir sont les suivantes :

Propriété 1 (Bornes pour Med (π)). Pour tout $\pi = (t_1,\ldots, t_p) \in O(E)$ et pour tout $m \in$ Med (π) :

$$\bigwedge_{i=1}^{p} t_i \leqslant m \leqslant \bigvee_{i=1}^{p} t_i$$

PROPRIETE 2 (Condition nécessaire de type majoritaire).

Pour tout $\pi = (t_1,\ldots,t_p)$ pour tout $m \in$ Med (π)

<u>et pour tout</u> x <u>tel que</u> mx $\in$ U (E) , <u>il vient</u> :

- si m $<$x , $\pi^{\wedge}(x,m) \leqslant p/2$ <u>et</u> $\pi^{\vee}(x,m) \leqslant p/2$;
- si m $>$x , $\pi^{\wedge}(x,m) \geqslant p/2$ <u>et</u> $\pi^{\vee}(x,m) \geqslant p/2$.

<u>PROPRIETE 3</u> (stabilité latticielle de Med). <u>Pour tout</u> $\pi \in$ 0 (E), Med (π) <u>définit un sous treillis</u> de (E,$\leqslant$).

En fait, nous allons établir chacune de ces propriétés dans le cadre "le plus général possible". Cela nécessite, dans un premier temps, de compléter les rappels sur les ordres que comporte cette introduction.

2 - PRELIMINAIRES

Les quelques rappels que nous allons effectuer concernent les propriétés métriques des ensembles ordonnés. Pour plus de détails et pour des références bibliographiques, le lecteur pourra consulter Monjardet [5] .

(E,$\leqslant$) désigne toujours un ensemble ordonné fini. Soit v une fonction définie sur E à valeurs réelles. Pour chaque chemin c : x = $x_o x_1 \ldots x_p$ = y de G(E) on pose :

$$\tilde{v}(c) = \sum_{i=0}^{p-1} |v(x_{i+1}) - v(x_i)|$$

et on appelle v <u>-géodésique</u> entre x et y tout chemin c entre x et y tel que $\tilde{v}(c)$ soit minimum.
On obtient alors la v- distance d_v :

$d_v(x,y)$ = longueur d'une v-géodésique entre x et y ; et on définit la fonction D_v, de E x O(E) dans l'ensemble des réels par :

$$D_v\,(x,\pi) = \sum_{i=1}^{p} d_v\,(x,t_i),$$

pour tout $\pi = (t_1,\ldots,t_p)$.

On dit que m est une v-médiane de $\pi \in$ O(E), lorsque :

$$D_v\,(m,\pi) = \operatorname*{Min}_{x\,\in E} D_v\,(x,\pi)$$

On note Med_v (π) l'ensemble des v-médianes de π.

Si v est une fonction de rang (i.e $v(y) = v(x) + 1$ lorsque y couvre x), $d_v = d$ et les notions de médiane et de v- médiane coïncident.

Lorsque E est un inf. demi-treillis on dit que v est une <u>valuation inférieure</u> lorsque :
(i) v est strictement croissante ($x < y$ entraine $v(x) < (y)$),
(ii) pour tout $x,y,z \in E$ tels que $x \leqslant z$, $y \leqslant z$: $v(x) + v(y) \leqslant v(x \wedge y) + v(z)$.

Dualement lorsque E est un sup. demi-treillis, v est une <u>valuation supérieure</u> lorsqu'elle est strictement croissante et :
(ii)★ pour tout x, y, $t \in E$ tels que $t \leqslant x$, $t \leqslant y$
$v(x) + v(y) \geqslant v(t) + v(x \vee y)$.

Enfin, sur un treillis, v, strictement croissante, est une <u>valuation</u> quand :
pour tout x, $y \in E$ $v(x) + v(y) = v(x \wedge y) + v(x \vee y)$.

Notons que :

Un treillis peut être muni d'une valuation si et seulement si il est modulaire.

Un inf. demi-treillis (resp. un sup. demi-treillis) est semi-modulaire inférieurement (resp. supérieurement) si et seulement si il possède une fonction de rang qui est une valuation inférieure (resp. supérieure).

Sur un inf. demi-treillis (resp. un sup. demi-treillis, resp., un treillis) une fonction v, strictement croissante est une valuation inférieure (resp. une valuation supérieure, resp. une valuation) si et seulement si :

$$d_v(x,y) = d_v(x, x \wedge y) + d_v(x \wedge y, y)$$
$$(\text{resp. } d_v(x,y) = d_v(x, x \vee y) + d_v(x \vee y, y),$$
$$\text{resp. } d_v(x,y) = d_v(x \wedge y, x \vee y)).$$

On dit qu'une fonction v, sur un treillis, est sur-modulaire lorsque, pour tout x, $y \in E$

$$v(x) + v(y) \geqslant v(x \wedge y) + v(x \vee y).$$

(Ainsi, sur un treillis, une valuation supérieure est une fonction sur-modulaire strictement croissante).

3 PROPRIETE 1

La propriété 1 s'étend à des "éléments centraux" plus généraux que les médianes.

On appelle v- éloignement sur l'ensemble ordonné $(E,\leqslant)$ toute application Δ définie sur $E \times O(E)$ et à valeurs réelles telle que : pour tout $x,y \in E$ et pour tout $\pi = (t_1,\ldots,t_p) \in O(E)$: (i) si $d_v(x,t_i) = d_v(y,t_i)$, pour $1 \leqslant i \leqslant p$, $\Delta(x,\pi) = \Delta(y,\pi)$;

(ii) si $d_v(x,t_i) < d_v(y,t_i)$, pour $1 \leqslant i \leqslant p$, $\Delta(x,\pi) < \Delta(y,\pi)$.

On dit alors que $z \in E$ est Δ- central pour $\pi \in O(E)$, lorsque :

$$\Delta(z,\pi) = \operatorname*{Min}_{x \in E} \Delta(x,\pi).$$

Exemples de v- éloignements $\Delta_q(x,\pi) = (\sum_{i=1}^{p} d_v(x,t_i)^q)^{1/q}$, en particulier $\Delta_1 = D$ et les éléments Δ_1 -centraux pour π sont les médianes de π.

$$\Delta_\infty(x,\pi) = \operatorname*{Max}_{1 \leqslant i \leqslant p} d(x,t_i).$$

PROPOSITION 1 Soit $(E,\leqslant)$ un treillis, soit v une fonction définie sur E et à valeurs réelles, soit Δ un v- éloignement sur E, soit $\pi = (t_1,\ldots,t_p) \in O(E)$ et soit z un élément de E, Δ- central pour π.

(i) Si v est une valuation supérieure : $\bigwedge_{i=1}^{p} t_i \leqslant z$,

(ii) Si v est une valuation inférieure : $z \leqslant \bigvee_{i=1}^{p} t_i$,

(iii) Si v est une valuation $\bigwedge_{i=1}^{p} t_i \leqslant z \leqslant \bigvee_{i=1}^{p} t_i$.

Preuve : Il suffit d'établir (i), (ii) s'en déduira par dualité et (iii) par juxtaposition de (i) et (ii).

Supposons que z est tel que $\bigwedge_{i=1}^{p} t_i \not\leqslant z$ et posons $z' = z \vee (\bigwedge_{i=1}^{p} t_i)$. Il vient, pour chaque t_i : $d_v(z',t_i) = v(z' \vee t_i) - v(z') + v(z' \vee t_i) - v(t_i)$. Or $v(z' \vee t_i) = v(z \vee t_i)$ et $v(z') > v(z)$. Donc : $d_v(z',t_i) < d(z,t_i)$, on en déduit : $\Delta(z',\pi) < \Delta(z,\pi)$. Ceci contredit le fait que z est Δ- central pour π.//

La propriété 1 n'est autre que l'assertion (iii), dans le cas particulier $\Delta = D_v$, lorsque v est une fonction de rang sur le treillis (modulaire) $(E,\leqslant)$.

4 - PROPRIETE 2

Désignons par $A(x)$ l'ensemble des éléments de E adjacents à x dans le graphe $G(E)$. On dit que $s \in E$ est une médiane locale de $\pi \in O(E)$, lorsque :

$$D(s,\pi) = \min_{x \in A(s)} D(x,\pi).$$

Toute médiane de π est une médiane locale de π (l'assertion réciproque serait, en général, fausse).

PROPOSITION 2: Soit $(E,\leqslant)$ un sup. demi-treillis semi-modulaire supérieurement soit, $\pi = (t_1,\dots,t_p) \in O(E)$ et soit $s \in E$. Les deux assertions ci-dessous sont équivalentes :

(i) s est une médiane locale de π ;

(ii) pour tout $x \in A(s)$:

- si $s < x$, $\pi^V(s,x) \leqslant p/2$,
- si $s > x$, $\pi^V(s,x) \geqslant p/2$.

Preuve : Considérons d'abord une valuation supérieure v sur E; pour tout y, z, $u \in E$, on a :

$$d_v(y,z) \geqslant d_v(yVu, zVu) \text{ lorsque } y \leqslant z.$$

En effet, la définition d'une valuation supérieure appliquée aux trois éléments z, yVu,y ($y \leqslant z$ et $y \leqslant yVu$) conduit à l'inégalité :

$$v(z) + v(yVu) \geqslant v(z\,V y\,Vu) + v(y). \text{ Donc } v(z) - v(y) \geqslant v(zVu) - v(yVu).$$

On suppose maintenant que E est semi-modulaire supérieurement et que la valuation supérieure v est une fonction de rang sur E.
De l'inégalité précédente, on déduit que, pour $x \in A(s)$ on a soit $d(xVt_i, sVt_i) = 0$ c'est-à-dire $xVt_i = sVt_i$, soit $d(xVt_i, sVt_i) = 1$, et ceci pour tout i, $1 \leqslant i \leqslant p$.

Calculons la différence $I = D(x,\pi) - D(s,\pi)$.

$I : \sum_{i=1}^{p} (d(x,t_i) - d(s,t_i)) = \sum_{i=1}^{p} (2(v(xVt_i) - v(sVt_i)) + v(s) - v(x))$. Distinguons deux cas :

1) $s < x$, il vient : $I = p - 2\,\pi^V(s,x)$;

2) $s > x$, il vient : $I = 2\,\pi^V(s,x) - p$.

Le résultat découle alors du fait que s est une médiane locale de π si et seulement si $I \geqslant 0$, pour tout $x \in A(s)$. //

On obtient la propriété 2, en juxtaposant, lorsque E est un treillis, la proposition 2 et sa duale. Indiquons une autre application.

Un ensemble ordonné arborescent (en abrègé E.O.A.) est est un sup. demi -treillis où tout élément est couvert par, au plus, un élément. Lorsque $(E \leqslant)$ est arborescent, G(E) est un arbre (la réciproque serait fausse !). Tout E.O.A. $(E,\leqslant)$ est évidemment semi-modulaire supérieurement et il est clair que, pour $s \in E$ et $x \in A(s)$:

1) Si $s < x$, $tVs = tVx$ si et seulement si t appartient à la composante connexe $C(x,sx)$ de $G(E)-sx$ qui contient x (condition équivalente à $d(x,t) < d(s,t)$). On notera $\pi(x,sx)$ le nombre d'indices i tels que $t_i \in C(x,sx)$: $\pi(x,sx) = \pi^V(x,s)$.

2) Si $s > x$, $tVs = tVx$ si et seulement si t appartient à la composante connexe $C(s,sx)$ de $G(E)-sx$ qui contient s. $\pi(s,sx)$ désignera le nombre d'indices i tels que $t_i \in C(s,sx)$ (donc $\pi(s,sx) = \pi^V(x,s)$).

Il découle immédiatement de la proposition 2 que dans un E.O.A. $(E,\leqslant)$, s est une médiane locale de π si et seulement si, pour tout $x \in A(s)$:

$$\pi(s,sx) \geqslant \pi(x,sx)$$

(ou, ce qui est équivalent : $\pi(s,sx) \geqslant p/2$)

(notons que Zelinka ([8]) a montré un résultat très proche de celui-ci dans le cas des médianes dans un arbre, la proposition constitue donc, d'une certaine manière une version - étendue aux ensembles ordonnés semi-modulaires du résultat de Zelinka c.f. aussi Leclerc [3]). Par ailleurs, on peut montrer que, dans un E.O.A., les notions de médiane, de médiane locale et de v-médiane coïncident.

5 PROPRIETE 3

PROPOSITION 3. Soit v une valuation sur le treillis modulaire $(E,\leqslant)$. Pour tout

$\pi \in O(E)$, $Med_v(\pi)$ définit un sous-treillis de $(E,\leqslant)$.

Preuve : Montrons, dans un premier temps que, pour chaque $t \in E$, la fonction δ, définie par : $\delta(s) = d_v(s,t)$ est sur-modulaire. Pour cela, évaluons l'expression $I = \delta(s) + \delta(x) - \delta(sVx) - \delta(s\Lambda x)$.

$I = v(s) + v(x) - v(s\Lambda x) - v(sVx) + 2(v(s\Lambda x\Lambda t) + v((sVx)\Lambda t) - v(x\Lambda t) - v(s\Lambda t))$.

Or $(xVs)\Lambda t \geqslant (x \Lambda t) V (s \Lambda t)$. Donc :

$I \geqslant v(s) + v(x) - v(s\Lambda x) - v(sVx) + 2v((s\Lambda t)\Lambda(x\Lambda t) + v((s\Lambda t)V(x\Lambda t)) - v(x\Lambda t) - v(s\Lambda t))$ et cette dernière expression est nulle, puisque v est une valuation. $I \geqslant 0$: δ est sur-modulaire.

Il s'ensuit, par additivité, que pour chaque $\pi \varepsilon O(E)$, la fonction $D_v(-,\pi)$ est également surmodulaire.

Le résultat découle alors immédiatement de la remarque (Topkis [7]) : l'ensemble des éléments d'un treillis qui minimisent une fonction surmodulaire f forment un sous-treillis de ce treillis (en effet si x et y minimisent f, $0 \leqslant f(xVy) - f(x) \leqslant f(y) - f(x\Lambda y) \leqslant 0$). //

En choisissant pour v une fonction de rang, on obtient la propriété 3.

BIBLIOGRAPHIE

[1] BARBUT M., "Médianes, distributivité, éloignements", multigraphié (1961), repris dans *Math. Sci. hum.* 70 (1980), 5-31.

[2] BARTHELEMY J.P., MONJARDET B., "The median procedure in cluster analysis and social choice theory", *Mathematical social sciences*, 1, 3 (1981), 235-268.

[3] LECLERC B., "Les arbres et les indices de centralité et de compacité de P. Parlebas", *Math. Sci. hum.*, 39 (1972), 27-35.

[4] MONJARDET B., "Théorie de la médiane dans les treillis distributifs finis et applications", *Ann. of discrete Math.*, 9 (1980), 87-91.

[5] MONJARDET B., "Metrics on partially ordered sets - a survey", *Discrete Math.*, 35 (1981), 173-184.

[6] SLATER P.J., "Centers to centroids in graphs", *Journal of graph theory*, 2 (1978), 209-222.

[7] TOPKIS M.T., "Minimizing a submodular function on a lattice", *Operation Researchs*, 26, 2 (1978), 305-321.

[8] ZELINKA B., "Medians and peripherians of trees", *Archivum Mathematicum* (1968), 87-95.