MATHÉMATIQUES ET SCIENCES HUMAINES

SUZANNE TAPONIER
CLAUDE FLAMENT

Analyse de la structure d'un matériel utilisé dans des expériences de tri

Mathématiques et sciences humaines, tome 19 (1967), p. 3-11
<http://www.numdam.org/item?id=MSH_1967__19__3_0>

Suzanne TAPONIER et Claude FLAMENT *

ANALYSE DE LA STRUCTURE D'UN MATERIEL UTILISE DANS DES EXPERIENCES DE TRI

I - LE PROBLEME

Les expériences psychologiques sur les processus intellectuels de classification utilisent fréquemment un matériel défini de façon combinatoire. Par exemple, des cartes ne se différenciant que par les dessins qu'elles portent, ces dessins variant systématiquement selon plusieurs descripteurs (1). Par exemple, les descripteurs sont: "couleur" et "forme". Soit D le nombre de descripteurs. Chaque descripteur peut prendre un certain nombre de modalités (2) (carré, triangle, .. pour la forme; vert, rouge, .. pour la couleur). On peut associer à chaque descripteur un ensemble Ω_i ($i = 1, 2, \ldots, D$), ayant V_i éléments ou modalités: le matériel M est alors défini comme le produit cartésien des ensembles Ω_i. (Hovland, 1952).

Pour diverses raisons, l'expérimentateur choisit dans M un sous-ensemble P dont les éléments sont utilisés comme étiquettes pour des boîtes (P est souvent appelé "ensemble de référence", nous dirons ici "ensemble de pivots", un élément de P étant un pivot). On donne alors au sujet l'ensemble T des objets restants (ou une partie de T) et on lui demande de les classer (trier) dans les boîtes étiquetées par les pivots (cf. BERG, 1948; OLERON, 1953). Toutes les hypothèses relatives aux stratégies classificatoires du Sujet tournent autour de l'idée que le Sujet utilise, d'un façon ou d'une autre, les ressemblances existant entre les objets à trier et les pivots.

Il semble donc utile, pour construire un matériel de ce type, pour planifier une expérience, et pour préparer la voie à des modèles du comportement des Sujets, d'analyser la structure formelle de T dans ses relations avec P.

L'ensemble P a pour fonction de présenter au sujet les modalités des descripteurs: une modalité peut figurer sur un seul pivot, sur plusieurs ou sur

---

* Département de Psychologie
Faculté des Lettres et Sciences Humaines d'Aix-en-Provence.

(1) ROUANET (1966) utilise les termes attributs à la place de descripteurs.

(2) TAPONNIER (1967) utilisait le terme valeur à la place de modalité.

aucun) (1). Les processus psychologiques mis en jeu seront très différents dans les trois cas. La plupart des auteurs considèrent cependant le cas où chaque modalité apparaît une fois et une seule dans l'ensemble P. Supposons que P ait V éléments, chaque élément portant une modalité et une seule de chaque descripteur; pour que chaque modalité apparaisse une fois et une seule en P, comme on le désire, il est donc nécessaire que chaque descripteur ait exactement V modalités. Tous les V_i sont égaux à V. Dans ces conditions, le matériel M comporte V^D éléments (2).

II - L'ENSEMBLE A

Soit $\Delta = \{1, 2, ..., i, ..., D\}$ l'ensemble des D premiers nombres entiers, et $\Phi = \{1, 2, ..., j, ..., V\}$ l'ensemble des V premiers nombres entiers (Δ est l'ensemble des indices des descripteurs; Φ est l'ensemble des indices des valeurs dans chaque descripteur, tous les descripteurs ayant le même nombre de valeurs).

Soit A l'ensemble des applications α de Δ en Φ : $A = \{\alpha\}$ et $\alpha : \Delta \longrightarrow \Phi$.

1. Correspondance bi-univoque entre M et A

Un objet a de M est, tel que nous l'avons défini plus haut, un D-uple $a = (a_1^{j_1}, a_2^{j_2}, a_i^{j_i}, ..., a_D^{j_D})$, les i étant les indices de descripteur; les j_k, les indices de valeur dans le descripteur considéré. A un objet a, on fait facilement correspondre une application α de Δ en Φ, en posant $j_i = \alpha(i)$. Inversement à une application α de Δ en Φ, on fait correspondre un objet a de M en posant: $a = \left(a_i^{\alpha(i)}\right)$.

Donc, pour étudier M, on peut étudier A.

2. Intérêt de A

Considérons la famille Π des applications <u>constantes</u> de Δ en Φ ($\Pi \subset A$). Ces applications peuvent être indiciées de 1 à V, Π_j étant l'application constante qui applique <u>tous</u> les éléments de Δ sur le même élément j de Φ. Donc, la première modalité de chaque descripteur apparaît en Π_1, la deuxième

---

(1) Cf. l'étude de LEPINE (1966), qui utilise deux pivots pour un matériel dont certains descripteurs comportent plus de deux modalités.

(2) - Notons que l'on considère très généralement que les descripteurs sont des ensembles non ordonnés. Si chaque descripteur était un ensemble totalement ordonné M aurait une structure de treillis distributif - problème que nous n'abordons pas ici.

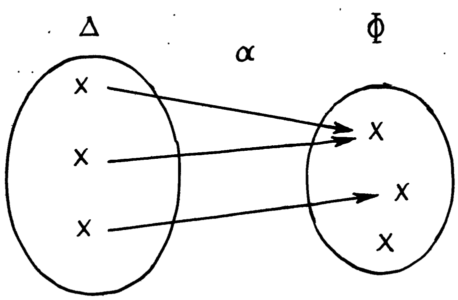
modalité de chaque descripteur apparaît en Π_2, etc .. Et on voit que les applications de Π correspondent aux pivots de P, pour peu qu'ayant choisit P, on ait pris la peine de bien indicier les modalités des descripteurs (règle: choisir P; ordonner arbitrairement les pivots; pour chaque descripteur indicier par j la modalité qui figure sur le $j^{ième}$ pivot. A partir de là, on peut facilement calculer le nombre d'ensembles P possibles qui est $(V!)^{(D-1)}$).

Chemin faisant, nous avons montré une relation bi-univoque entre Π et P, mais aussi une relation bi-univoque entre Π et Φ , d'où il résulte une relation bi-univoque entre P et Φ. En d'autres termes, les applications de Δ en Φ peuvent être considérées comme des applications de Δ en P; et, remplacer un objet a de M par une application α de A, revient donc à définir l'objet a par ses relations avec P. Cela devrait faciliter notre travail.

III - ORGANISATION DE A

1. Rappel: image, ensemble quotient, décomposition canonique d'une application.

Chaque application α de Δ en Φ peut être considérée comme un ensemble de flèches allant des éléments de Δ vers les éléments de Φ . Une flèche et une seule part de chaque élément de Δ ; un élément de Φ peut être l'extrêmité d'une seule flèche, de plusieurs, ou d'aucune.



L'image $\alpha(\Delta)$ de Δ par α est une partie de Φ : celle qui comprend les points extrémités d'au moins une flèche.

Sur Δ , on peut définir une relation d'équivalence, réunissant les points de Δ d'où partent des flèches aboutissant à un même point de Φ (ayant même image en Φ). Les classes résultant de cette relation constituent une "partition" de Δ : l'ensemble quotient Δ / α est celui dont les éléments sont les classes.

L'application α de Δ en Φ est canoniquement décomposée en une application α_1 de Δ sur Δ/α , et une application α_2 de Δ/α en Φ :

$$\begin{tikzcd} \Delta \arrow[rr, "\alpha"] \arrow[dr, "\alpha_1"'] & & \Phi \\ & \Delta/\alpha \arrow[ur, "\alpha_2"'] & \end{tikzcd}$$

- α_1 : à chaque élément de Δ on fait correspondre, très naturellement, la classe de Δ / α qui contient l'élément considéré;

- α : à chaque classe de Δ / α , on fait correspondre l'élément de Φ qui est l'image commune de tous les éléments de la classe considérée. Il est clair que α_2 se réduit à une application bijective (bi-univoque) de Δ / α sur $\alpha(\Delta)$. Le nombre de classes en Δ / α , et le nombre d'éléments de $\alpha(\Delta)$, sont égaux; ce nombre sera désigné par $N(\alpha)$.

2. Organisation de A selon l'image

Soit deux applications α et β de Δ en Φ ; les images $\alpha(\Delta)$ et $B(\Delta)$ sont deux parties de Φ ; on peut donc considérer la relation d'inclusion entre parties de Φ , et la structure simpliciale de l'ensemble des parties de Φ. On en induit sur A une première relation d'ordre:

$\alpha \prec \beta \Longleftrightarrow \alpha(\Delta)$ est une <u>(sous)-partie</u> de $\beta(\Delta)$;

et une relation d'équivalence I (selon l'image):

$$\alpha \ I \ \beta \Longleftrightarrow \alpha < \beta \text{ et } \beta < \alpha$$
$$\Longleftrightarrow \alpha(\Delta) = \beta(\Delta)$$

L'ensemble $\mathscr{C}_I$ des classes selon I est donc organisé comme le simplexe des parties de Φ .

3. Organisation de A selon l'ensemble quotient (cf. ROSENSTIEHL et MOTHES, 1965, ch. 1-4 et 2-3)

Δ/α et Δ/β sont deux partitions de l'ensemble Δ ; l'une des deux peut être une partition plus fine que l'autre. On en induit sur A une relation d'ordre:

$\alpha \dashv \beta \Longleftrightarrow \Delta/\alpha$ est une <u>sous-partition</u> de Δ/β, et une relation d'équivalence Q (selon le quotient):

$$\alpha \ Q \ \beta \Longleftrightarrow \alpha \dashv \beta \text{ et } \beta \dashv \alpha$$

$$\Longleftrightarrow \Delta/\alpha = \Delta/\beta$$

L'ensemble des partitions d'un ensemble s'organise, par la relation d'affinage que nous venons d'envisager, selon un treillis semi-modulaire (cf. BARBUT, 1967) *, et donc l'ensemble $\mathscr{C}_Q$ des classes selon Q s'organise comme ce treillis.

4. Organisation de A selon N

L'effectif des éléments de $\alpha(\Delta)$, ou des classes de Δ/α , est désigné par $N(\alpha)$, ce qui induit un ordre total sur A, et une équivalence selon N:

$$\alpha \ N \ \beta \Longleftrightarrow N(\alpha) = N(\beta).$$

Les classes de $\mathscr{C}_N$ s'ordonnent totalement.

* Cet article, dont le titre est donnée dans la bibliographie et qui devait paraître dans le présent bulletin, ne sera publié qu'ultérieurement en raison de l'abondance de matière de ce numéro. N.D.L.R.

5. Relations entre ces organisations

Des ensembles égaux ayant même effectif, alors que des ensembles de même effectif ne sont pas forcément égaux, il s'ensuit que les classifications $\mathscr{C}_I$ et $\mathscr{C}_Q$ sont des sous-classifications de $\mathscr{C}_N$. En fait, $N(\alpha)$ correspond au <u>niveau</u> où se trouve α dans le treillis simplicial selon I (si on numérote ses niveaux de 0 à V) et le treillis semi-modulaire selon Q (si on numérote ses niveaux de I à D). Mais tous les niveaux de ces treillis ne sont pas représentés en A. L'image de Δ en Φ comporte au moins un élément: donc le niveau $N = o$ du simplexe de Φ n'est pas représenté en A. Par ailleurs, le niveau maximum du simplexe des parties de Φ est V (V étant l'effectif de Φ), d'où: $N \leqslant V$; et la partition maximum de Δ, qui comporte D éléments, comprend D classes, toutes ponctuelles, et donc $N \leqslant D$; en bref, $N \leqslant \min\{D, V\}$, et les niveaux supérieurs des treillis peuvent ne pas être présents, suivant la relation entre V et D.

Pour N donné, les classifications selon I et selon Q peuvent être croisées. On obtient donc le schéma:

$N = 1$	$N = 2$	... N ...	$N = \min V, D$
pivots	paires dans Φ ; bi-partitions de Δ	C_V^N ; P_D^N	. si $V \leqslant D$: C_V^V et P_D^V . si $V \geqslant D$: C_V^D et P_D^D

où C_V^N désigne les <u>parties</u> de N objets dans un ensemble de V objets, et P_D^N les <u>partitions</u> en N classes d'un ensemble de D objets. Chaque colonne du tableau $\left(C_V^N \times P_D^N\right)$ pour N donné constitue le niveau N du treillis des parties de Φ ; chaque ligne du même tableau constitue le niveau N du treillis des partitions de Δ.

6. Les classes IQ

Soit la relation d'équivalence IQ, produit des relations I et Q:

$$\alpha \text{ IQ } \beta \Longleftrightarrow \alpha \text{ I } \beta \quad \text{et } \alpha \text{ Q } \beta$$

α et β ont alors même ensemble quotient et même image; deux telles applications figurent dans une même petite case du tableau ci-dessus. Dans leur décomposition canonique, α et β ne diffèrent qu'en α_2 et β_2, qui sont deux applications <u>bijectives</u> différentes d'un même ensemble quotient sur une même ensemble image. Traditionnellement, l'ensemble des applications bijectives d'un ensemble de

N éléments sur un ensemble de N éléments constitue l'une des définitions de l'ensemble des permutations de N éléments. Donc une classe IQ (de niveau N) peut être organisée comme le permutoèdre de base N (cf. Guilbaud & Rosenstiehl, Analyse Algébrique d'un scrution, M.S.H. n° 4).

IV - INTERPRETATION DE L'ORGANISATION DE A

Revenons au matériel M. Ayant choisi un ensemble P de pivots, nous dirons que $p \in P$ est un pivot proche de $a \in M$ si et seulement si p et a, pour au moins un descripteur, ont la même modalité - en d'autres termes, si p et a ont un minimum de similitude. L'hypothèse dominante ici est que le sujet, pour classer a, hésitera entre les boîtes étiquetées par des pivots proches, à l'exclusion des autres boîtes.

Il est facile de voir que, si à l'objet a on fait correspondre l'application α , et à l'ensemble de pivots P, l'ensemble Φ , aux pivots proches de a correspondent les éléments de l'image $\alpha\ (\Delta) \subset \Phi$. Donc, deux applications équivalentes selon I ont le même ensemble de pivots proches; deux applications équivalentes selon N ont même nombre de pivots proches.

L'interprétation des équivalences selon Q est moins directe: il faut introduire un degré de similitude. Si une classe de Δ / α, comportant k éléments de Δ , a pour image par α le même élément j, cela veut dire que l'objet a correspondant à α, a exactement les mêmes valeurs que le pivot p_j sur les k descripteurs. En d'autres termes, les effectifs des classes de Δ/α mesurent le degré de similitude de a à ses divers pivots proches.

Il est clair que la répartition de ces degrés de similitude peuvent influer fortement sur la classification faite par le Sujet: il risque de classer l'objet avec le pivot proche le plus similaire. D'où l'idée de considérer des objets ayant même similitude avec tous leurs pivots proches (le nombre N de pivots proches étant alors la mesure de l'incertitude). De tels objets correspondent à des applications dont l'ensemble quotient constitue une équipartition de Δ . Notons que si l'on veut une équipartition de D objets en N classes, cela suppose que D est un multiple de N.

V - STATISTIQUES (cf. ROSENSTIEHL et MOTES, 1965, chap. 2-3, et KREWERAS, 1963)

Il peut être utile de compter le nombre d'objets dans chaque classe.

Chaque classe IQ, pour N donné a pour effectif N !, et peut être munie de la structure d'un permutoèdre de base N.

Pour N donné, le nombre de classes selon l'image est le nombre de combinaisons de V termes N à N, soit:

$$C_V^N = \frac{V!}{N!\ (V-N)!}$$

Pour N donné, le nombre de classes selon le quotient est le nombre P_D^N de partitions de D objets en N classes; il ne semble pas exister de formule commode donnant directement P_D^N en fonction de N et D; mais, au moins pour les petites valeurs de N et D (ce qui est toujours le cas dans notre problème), une formule récurrente donne commodément ces nombres de proche en proche (triangle de Stirling; cf. KREWERAS, 1963).

Et le nombre d'objets dans une classe selon N est alors:

$$P_D^N \cdot C_V^N \cdot N!.$$

Considérons les <u>équipartitions</u>: soit EP_D^N le nombre d'équipartitions de D objets en N classes. Deux cas sont à considérer: si $k = \frac{D}{N}$ n'est <u>pas</u> un nombre entier, alors $EP_D^N = 0$; sinon, on montre facilement (cf. ROSENSTIEHL et MOTHES, 1965, p. 152):

$$EP_D^N = \frac{D!}{N!\ (k!)^N}$$

Et le nombre d'objets ayant N pivots proches équidistants est alors:

$$EP_D^N \cdot C_V^N \cdot N! = \frac{D!\ V!}{N!\left(\frac{D}{N}!\right)^N \cdot (V-N)!}$$

VI - ETUDE DE QUELQUES CLASSES PARTICULIERES

1. Cas N = 1

Nous retrouvons l'ensemble Π des applications constantes, c'est-à-dire, l'ensemble des pivots.

2. Cas N = 2

Considérons une classe I, comprenant toutes les applications α ayant pour image $\{j, k\}$; une telle application α est complètement décrite par la partie X de Δ qui a pour image j, puisque la partie $\overline{X}$, complémentaire de X en Δ a alors k pour image (on a X et $\overline{X} \neq \phi$). On peut alors munir l'ensemble des applications de cette classe I de la structure simpliciale correspondant à l'ensemble des parties X de Δ , les pôles de ce simplexe étant constitués par les deux applications constantes sur les deux points de l'image commune aux applications de la classe I. Au total, la classe $N = 2$ s'organise comme une famille de simplexes de base D, un pour chaque paire d'éléments de V.

Notons que si $V = 2$, tout le matériel se réduit à un seul simplexe de base D.

3. Cas N = D ⩽ V

Il n'y a alors qu'une seule classe Q: l'unique D-partition possible pour D objets (les classes sont toutes ponctuelles; c'est un cas extrême d'équipartition). Une classe I est alors constituée par un permutoèdre, mais l'important est que tous les objets sont équidistants de tous les pivots proches. Si $N = D = V$, la classe N se réduit à une seule classe I.

VII - RELATIONS ENTRE OBJETS ET RELATIONS AUX PIVOTS

Nous avons étudié les relations des objets aux pivots, parce que nous supposons que le sujet classe un objet en fonction de sa similitude aux pivots. Mais lorsque le Sujet doit classer successivement plusieurs objets, on peut penser que la similitude existant entre deux objets successifs interferera avec la similitude du deuxième de ces objets et les pivots.

Le langage adopté ici n'est pas très commode pour exprimer les relations entre objets: il revient en fait à exprimer la relation entre les relations des objets aux pivots. Mais on peut espérer que ce langage facilitera néanmoins l'étude de l'interférence de deux types de similitude (entre objets, et entre objets et pivots).

VIII - MODELES DE COMPORTEMENT

La formalisation présentée suggère toute une famille de modèles pouvant rendre compte de certains types de comportement des sujets dans la situation considérée.

A chaque élément de Δ , attachons un poids $w_i \geqslant 0$, de telle sorte que $\Sigma w_i = 1$. Chaque application α permet alors de définir une pondération $v_\alpha(j)$ sur les éléments de Φ , en posant:

$$v_\alpha(j) = \sum_{i \in \alpha^{-}(j)} w_i$$

On peut penser que les Sujets classeront l'objet a correspondant à α avec le pivot v_j correspondant à l'élément j et Φ , soit parce que $v_\alpha(j)$ est la plus forte pondération sur Φ obtenue par α ; soit avec une probabilité $v_\alpha(j)$.

On voit donc que les w_i sont des caractéristiques constantes des descripteurs i. Si tous les w_i sont égaux, les Sujets classent en fonction du nombre de ressemblances entre les objets et les pivots, ce qui est proche de ce que nous avons décrit comme stratégie du modèle (TAPONIER, 1967). Si un w_i égale l'unité, les autres étant nuls, les Sujets classent en n'utilisant que le descripteur i (cf. stratégie du critère). D'une manière générale, les w_i mesurent le poids que les Sujets accordent à chaque descripteur, que ce soit pour des raisons décisionnelles, perceptives ou autres.

- - - - - -

BIBLIOGRAPHIE

BARBUT, M., 1967 - Treillis de partitions et leurs représentations géométriques. M.S.H. n° 20, à paraître.

BERG, E.A., 1948 - A simple objective technique for measuring flexibility in thinking, J. gen. Psychol., 39, 15-22.

HOVLAND, C.I., 1952 - A "communication analysis" of concept learning, Psycholog. Rev., 59, 461-472.

KREWERAS, G., 1963 - Une dualité élémentaire souvent utile dans les problèmes combinatoires, M.S.H. n° 3, 31-41. 1963.

LEPINE, D., 1966 - Critères de la réponse d'"identité" chez l'enfant, Année psycholo., 66, 417-446.

OLERON, P., 1953 - Classement multiple et langage, J. Psychol. norm. path., 46, 299-315.

ROSENSTIEHL, P., MOTHES, J., 1965 - Mathématiques de l'action, Paris, Dunod.

ROUANET, H., 1966 - Mathématiques et méthodes psychologiques: la description des situations expérimentales, M.S.H., n° 16, 25-37. 1966.

TAPONIER, S., 1967- Utilisation du critère de "similitude maximale" dans une tâche de tri, Cah. Psycholo., 10, 3-23.

\- - - - - -