RAIRO. RECHERCHE OPÉRATIONNELLE

FARID BENINEL

Dissimilarités de type sphérique et positionnement multidimensionnel normé

RAIRO. Recherche opérationnelle, tome 33, nº 4 (1999), p. 569-581

<http://www.numdam.org/item?id=RO_1999__33_4_569_0>

RAIRO Operations Research
RAIRO Oper. Res. **33** (1999) 569-581

DISSIMILARITÉS DE TYPE SPHÉRIQUE ET POSITIONNEMENT MULTIDIMENSIONNEL NORMÉ (*)

par Farid BENINEL ([1])

Communiqué par Jean-Yves JAFFRAY

Résumé. – *On s'intéresse à la caractérisation des mesures de dissimilarité sur ensembles finis admettant une représentation euclidienne sphérique. On propose, en conséquence, une méthodologie de détermination de la représentation euclidienne sphérique associée à un ensemble d'items et restituant au mieux les dissimilarités inter items. Cette méthodologie présente l'avantage de la lisibilité graphique des qualités individuelles de projection à l'instar de l'ACP normée dont elle constitue une généralisation. En outre, elle évite l'arbitraire du codage sphérique que suppose l'usage des fonctions de similitude couramment utilisées en MDS.*

Mots clés : Dissimilarité, image euclidienne, MDS, analyse métrique, dissimilarité sphérique, transformation euclidienne, fonctions de similitude.

Abstract. – *Our concern here, is the characterization of dissimilarity indexes defined over finite sets, whose spatial representation is spherical. Consequently, we propose a methodology (Normed MultiDimensional Scaling) to determine the spherical euclidean representation of a set of items best accounting for the initial dissimilarity between items. This methodology has the advantage of being graphically readable on individual qualities of projection like the normed PCA, of which it constitutes a generalization. Moreover, it avoids the arbitrary character of spherical encoding which the use of similitude functions currently used in MDS, implies.*

Keywords: Dissimilarity, Euclidean image, MDS, metric analysis, spherical dissimilarity, Euclidean transformation, similitude functions.

1. PRÉLIMINAIRES

Considérons un ensemble fini d'items E de cardinal $n > 1$, qu'on assimilera à l'ensemble d'index associé $\{1, 2, \ldots, n\}$ et δ une mesure de dissimilarité sur E. Rappelons que δ vérifie les axiomes qui suivent :

$$(*) \quad \forall\, i \in E \qquad \delta_{ii} = 0,$$
$$(**) \quad \forall\, i, j \in E \qquad \delta_{ij} = \delta_{ji}.$$

(*) Reçu en décembre 1998.
([1]) IUT Département STID, Centre d'Activités de Noron, Université de Poitiers, 8 rue Archimède, 79000 Niort.

Recherche opérationnelle/Operations Research, 0399-0559/99/04/$ 7.00

En analyse métrique (MultiDimensional Scaling), on s'intéresse à la représentation des objets de E par un système ponctuel $\{X_i \,;\, i \in E\}$ d'un espace métrique (Ω, d) tel que :

$$\forall\, i, j \in E \qquad \delta_{ij} \cong d(X_i, X_j).$$

Cette étude se restreint souvent au cas où $\Omega = R^p$ et d la distance euclidienne classique ; dans ce cas, le système $\{X_i \,;\, i \in E\}$ est appelé image euclidienne de δ.

Lorsque $\delta = d$ l'image euclidienne est dite exacte, sinon on parle d'image approchée. δ étant donnée, cette image euclidienne s'obtient par minimisation d'une fonction coût mesurant l'écart entre δ et d. On trouvera, dans Critchley [7] et De Leuw-Heiser [9], les fonctions d'écart couramment utilisées ainsi que les algorithmes d'optimisation correspondant.

Dans ce travail, on présente un résultat de caractérisation d'une dissimilarité δ sur E admettant une image exacte portée par une sphère de R^p ; rappelons que la sphéricité d'une dissimilarité est entendue au sens de la distance euclidienne ou distance corde, ceci pour distinguer ce travail de ceux de Shoenberg [20] et Critchley et Fichet [8] donnant notamment des caractérisations des dissimilarités sphériques au sens de la distance de Riemman ou distance arc. La caractérisation que l'on propose dispense de fixer au préalable le rayon de la sphère support comme c'est le cas pour les caractérisations données par Fichet et Le Calvé [11]. On présentera, en conséquence, des exemples de transformations d'une dissimilarité quelconque sur E en une dissimilarité admettant une image sphérique.

En application, on propose une méthode d'analyse métrique baptisée **NMDS** (**N**ormed **M**ulti**D**imensional **S**caling) utilisant le caractère sphérique ou presque sphérique d'une mesure de dissimilarité quelconque. Cette méthode offre l'avantage de la lisibilité graphique des qualités individuelles de projection et permet d'interpréter les éloignements entre items en terme d'opposition, de corrélation et d'indépendance.

2. DISSIMILARITÉS DE TYPE SPHÉRIQUE

Rappelons, dans cette section, la définition et la caractérisation d'une structure de dissimilarité euclidienne. On appelle structure de dissimilarité, un couple (E, δ) où E est un ensemble d'items et δ la dissimilarité associée.

On appellera image euclidienne exacte de (E, δ) dans R^p, lorsqu'elle existe, un système de points $\{X_i \,;\, i \in E\} \subset R^p$ tel que

$$\forall\, i, j \in E \qquad \delta_{ij} = \|X_i - X_j\|_2.$$

Lorsqu'une structure de dissimilarité admet une image euclidienne exacte, on parlera de structure de dissimilarité euclidienne.

Soit $k \in E$ et considérons la forme $W^k(\delta)$ définie par:

$$\forall\, i,j \in E \qquad W^k(\delta)_{ij} = \frac{1}{2}\,(\delta_{ki}^2 + \delta_{kj}^2 - \delta_{ij}^2). \tag{2.1}$$

THÉORÈME 1 [Shoenberg 1937]: *La structure de dissimilarité (E, δ) est euclidienne si et seulement si $\exists k \in E$ tel que la forme $W^k(\delta)$ est Semi Définie Positive.*

Donnons brièvement la trame de la démonstration:

Lorsque (E, δ) est euclidienne on a

$$\forall\, i,j \in E \qquad W^k(\delta)_{ij} = \langle X_k - X_i,\, X_k - X_j \rangle. \tag{2.2}$$

Ainsi donc, $W^k(\delta)$ est une matrice de produits scalaires, donc SDP.

Lorsque $W^k(\delta)$ est SDP, la décomposition de Cholesky garantit l'existence de Y, matrice réelle de type (n, p), et telle que $W^k(\delta) = Y^tY$.

L'image euclidienne de (E, δ) est donnée par $X_k = 0 \in R^p$ et pour $i \in E/\{k\}$ $X_i = Y_i$, avec Y_i vecteur donné par la i-ème ligne de Y. ■

COROLLAIRE 1: *S'il existe k tel que $W^k(\delta)$ soit SDP, alors $W^{k'}(\delta)$ est SDP $\forall\, k' \in E$.*

Cette caractérisation est tournée dans différentes versions équivalentes. On donne dans ce qui suit une version plus générale et intéressante de par son utilisation dans la suite.

Soit $s \in R^n$ et considérons la forme $F(s, \delta)$ définie par:

$$F(s,\delta)_{ij} = \frac{1}{2}(\delta_{i.}^{(2)} + \delta_{j.}^{(2)} - \delta_{ij}^2) \tag{2.3}$$

avec

$$\delta_{i.}^{(2)} = \sum_j s_j \delta_{ij}^2 - \frac{1}{2} \sum_i \sum_j s_i s_j \delta_{ij}^2.$$

En prenant $s = e_k$ (k-ième vecteur de la base canonique de R^n) on obtient la forme donnée en (2.1) *i.e.* $F(s, \delta) = W^k(\delta)$.

En prenant $S = \frac{1}{n} 1_n$ (avec 1_n le vecteur de R^n dont toutes les composantes sont égales à l'unité) on obtient la forme dite de Torgeson *i.e.* $F(s, \delta) = W^g(\delta)$ où g est un item fictif vérifiant: $\forall\, i \in E$ $\delta_{gi}^2 = \delta_{i.}^{(2)}$.

L'item g a pour point image, lorsque l'image euclidienne existe, le centre de gravité de celle-ci. Plus généralement, le vecteur de pondération s détermine l'item $k(s)$ réel ou fictif en lequel la forme $W^{k(S)}(\delta)$ est prise. Cet item est celui ayant pour point image, lorsque la structure de dissimilarité est euclidienne, le barycentre du système ponctuel $\{(X_i, s_i)\,;\ 1 \leq i \leq n\}$. Il vérifie, dans ce cas, $F(S, \delta)_{ii} = \delta_{ki}^2$ pour tout $i \in E$.

Ainsi, il convient que s vérifie la condition d'un vecteur de pondération:

$$\sum_i s_i = 1. \tag{2.4}$$

COROLLAIRE 2: *La structure de dissimilarité (E, δ) est euclidienne si et seulement si $\exists s \in R^n$ vérifiant la condition* (2.4) *et tel que $F(s, \delta)$ est Semi Définie Positive.*

La démonstration de ce corollaire tient du fait que $W^k(\delta)$ et $F(s, \delta)$ sont SDP ou non en même temps.

On montre sans peine que s'il existe $s \in R^n$ tel que $F(s, \delta)$ est Semi Définie Positive, alors $F(s, \delta)$ est SDP $\forall\, s \in R^n$ vérifiant (2.4).

DÉFINITION 1: *Une structure de dissimilarité (E, δ) sera dite de type sphérique si et seulement si elle admet une image euclidienne portée par une sphère de R^p.*

Remarques:

r1) p désigne la dimension de l'espace engendré par l'image euclidienne associée à (E, δ). Rappelons que $p \leq n - 1$. Par la suite, on fera l'abus d'appeler rang de δ, le rang ou la dimension de l'image associée.

r2) L'image d'une sphère par isométrie étant une sphère de même rayon, s'il existe une image exacte sphérique de (E, δ), toute autre image est sphérique.

Complétons la structure de données (E, δ) par un item fictif o obtenant la structure $(E \cup \{o\}, \delta')$ vérifiant: $\delta'_{ij} = \delta_{ij}$ pour i, j dans E et $\delta'_{oi} = r$ pour i dans E.

DÉFINITION 2: *Une structure (E, δ) sera dite de type sphérique si et seulement si, il existe $r \in R^+$ tel que la structure complétée $(E \cup \{o\}, \delta')$ soit euclidienne.*

Dans le cas où $(E \cup \{o\}, \delta')$ est euclidienne et si l'on note $\{X_o\} \cup \{X_i;\ i \in E\}$ l'image euclidienne associée, on aurait: $\|X_o - X_i\|_2 = r$ pour tout $i \in E$.

X_o et r désignant le centre et le rayon de la sphère portant l'image euclidienne $\{X_i;\ i \in E\}$ associée à (E, δ).

Il est bien connu que le simplexe est inscriptible sur une sphère (*cf.* par exemple Gbegan [12]). On en donne, dans ce qui suit, une autre démonstration.

PROPOSITION 1: *Lorsque δ est euclidienne de rang maximum (i.e.* Rang $(\delta) = n - 1$*) alors (E, δ) est de type sphérique.*

En effet, dans le cas où (E, δ) est de type sphérique, les points $\{X_i; i \in E\}$ seraient inscriptibles sur une sphère. Notons X_0 le centre de cette sphère, on a

$$\forall\, i \in E \quad \|X_o - X_i\| = \|X_o - X_1\|.$$

Il s'ensuit

$$\forall\, i \in E \quad \|X_i - X_1\|^2 = 2\,\langle X_o - X_1,\, X_i - X_1\rangle.$$

Ce qui donne $\forall\, i \in E/\{1\}$, $\exists \alpha \in R^{n-1}$ tel que

$$\|X_i - X_1\|^2 = 2\sum_j \alpha_j \langle X_j - X_1, X_i - X_1\rangle. \tag{2.5}$$

Ici α est le vecteur donnant les coordonnées de $(X_0 - X_1)$ dans la base $(X_i - X_1)_{i=2,\ldots,n}$. La détermination de α, revient à la détermination du centre de la sphère. Pour montrer le caractère sphérique de δ, il suffit de montrer l'existence et l'unicité de α.

Soit $W^1(\delta)$ la forme définie en (2.1) prise pour $k = 1$. En vertu de la relation donnée en (2.2) on a:

$$\langle X_j - X_1,\, X_i - X_1\rangle = W^1(\delta)_{ij}.$$

En substituant dans la relation (2.5), on obtient:

$$\exists \alpha \in R^{n-1} \text{ tel que } \forall\, i \in E/\{1\} \quad W^1(\delta)_{ii} = +2\sum_j \alpha_j W^1(\delta)_{ij}.$$

Notons W^{11} la matrice obtenue de $W^1(\delta)$ par suppression de la $1^{\text{ère}}$ ligne et de la $1^{\text{ère}}$ colonne composées de termes nuls. On a $rang\,(W^{11}) =$

$rang\,(W^1(\delta)) = n - 1$ (car δ est de rang maximal) et donc W^{11} est inversible. Plus précisément α est donné de façon unique par l'équation :

$$Diag(W^{11}) = -2W^{11} \cdot \alpha.$$

■

En conséquence de ce résultat, toute perturbation d'une dissimilarité δ sur E conduisant à une dissimilarité euclidienne de rang maximal, est une perturbation sphérique. On donne dans ce qui suit deux exemples de ce type de perturbation.

Exemple 1 : Puissances terme à terme d'une dissimilarité.
Considérons la famille de transformations paramétrées $(T_\beta)_{\beta \in R^+}$ définie par :

$$\forall\, i, j \in E \quad (T_\beta(\delta))_{ij} = \delta_{ij}^\beta.$$

Notons β_e l'indice « d'euclidiennité » de la dissimilarité δ ; cela signifie que pour $\beta \leq \beta_e$ la dissimilarité $T_\beta(\delta)$ est euclidienne, et non euclidienne pour $\beta > \beta_e$.

On sait, à partir du travail de Joly-Le Calvé [15], que pour $\beta < \beta_e$, $T_\beta(\delta)$ est euclidienne de rang maximal.

En application de la proposition 1 précédente, l'ensemble $\{T_\beta\,;\ 0 < \beta < \beta_e\}$ constitue une famille de transformations sphériques pour la dissimilarité δ.

Exemple 2 : Constantes additives.

Soit $r \in R^+$ et considérons la famille de transformations $(T_\lambda)_{\lambda \in R^+}$ définie par :

$$\forall\, i, j \in E \quad (T_\lambda(\delta))_{ij} = (\delta_{ij}^r + \lambda(1 - k_i^j))^{1/r}$$
$$k_i^j\ : \text{ coefficient de Kronecker.}$$

On montre que $\exists \lambda^*(r) \in R^+$ tel que $\forall\, \lambda \geq \lambda(r)\ T_\lambda(\delta)$ est euclidienne ; on trouve dans Lingoes [17] et Cailliez [6] respectivement les solutions $\lambda^*(2)$ et $\lambda^*(1)$.

On montre sans peine que $\forall\, \lambda > \lambda^*(r)\ T_\lambda(\delta)$ est euclidienne de rang maximal et donc, en vertu de la proposition 1, sphérique.

Les dissimilarités de rang maximal n'englobent pas toutes les dissimilarités sphériques sur E. Il convient donc d'établir une caractérisation des dissimilarités sphériques de type quelconque.

Le problème revient en ce qui concerne la forme $F(s, \delta)$ donnée en 2.4 de déterminer, lorsqu'il existe, le vecteur de pondération s assurant sa positivité et tel que $F(s, \delta)_{ii} = c$ (c: constante positive) pour tout $i \in E$.

Notons B la matrice de terme général $B_{ij} = -\frac{1}{2}\delta_{ij}^2$.

$U_1 U_2 \ldots U_n$: ses vecteurs propres normés à l'unité ; ils sont orthogonaux du fait du caractère symétrique de B.

$\lambda_1 \lambda_2 \ldots \lambda_n$: les valeurs propres correspondantes ; elles sont réelles.

On donne dans ce qui suit l'expression analytique de l'inverse généralisé au sens de Moore-Penrose [19] de la matrice B. Notons B^- un tel inverse.

LEMME 1 :

$$B^- = \sum_{\lambda_k \neq 0} \frac{1}{\lambda_k} U_k U_k^t.$$

En effet, en utilisant la décomposition spectrale de B, l'on a :

$$B = \sum_k \lambda_k U_k U_k^t.$$

Il est aisé de vérifier que B^- vérifie les propriétés caractéristiques de l'inverse généralisé de Moore-Penrose, en l'occurrence :

$$(1) \quad BB^-B = B \qquad (2) \quad B^-BB^- = B^-$$
$$(3) \quad (BB^-)^t = BB^- \qquad (4) \quad (B^-B)^t = B^-B.$$

LEMME 2 : *Si δ est une dissimilarité sur E, non nulle, alors $B^- 1_n \neq 0$.*

En utilisant le lemme 1, on a

$$B^- = \sum_{\lambda_k \neq 0} \frac{1}{\lambda_k} U_k U_k^t.$$

Ce qui donne

$$B^- 1_n = \sum_{\lambda_k \neq 0} \frac{1}{\lambda_k} U_k \langle U_k, 1_n \rangle.$$

Les vecteurs propres U_k, $k = 1, \ldots, n$, étant orthogonaux et donc indépendants, à cause de la symétrie de B, on a

$$B^- 1_n = 0 \quad \text{implique} \quad \langle U_k, 1_n \rangle = 0$$
$$\text{pour tout } k \text{ tel que } \lambda_k \neq 0.$$

Ce qui donne

$$^t1_nB^-1_n = \sum_k \lambda_k \langle U_k, 1_n\rangle^2 = 0.$$

Ou encore, en revenant à la définition initiale de B

$$^t1_nB1_n = \sum_{i,j}\left(-\frac{1}{2}\delta_{ij}^2\right) = 0.$$

Ceci n'est possible que si δ est supposée nulle. ■

On donne dans ce qui suit une caractérisation d'une dissimilarité sphérique quelconque. Considérons, pour cela, le vecteur de pondération donné par $s^* = (1_n^t B^- 1_n)^{-1} B^- 1_n$; il vérifie bien la condition posée en (2.4).

En effet

$$\sum_i s_i^* = \langle s^*, 1_n\rangle = {}^t1_n({}^t1_nB^-1_n)^{-1}B^-1_n = 1.$$

■

THÉORÈME 2 : *Soit δ une mesure de dissimilarité sur E, non nulle.*

δ est sphérique si et seulement si la forme $F(s^*, \delta)$ est Semi Définie Positive.

La condition de positivité de $F(s^*, \delta)$ étant nécéssaire et suffisante pour l'existence d'une image euclidienne de (E, δ), on la suppose vérifiée; dans ce cas, la sphéricité de cette image euclidienne est vérifiée si:

$$\exists c > 0 \text{ tel que } \forall\, i \in E \quad F(S^*, \delta)_{ii} = C.$$

La constante c ainsi définie, correspond au carré du rayon de la sphère support de l'image euclidienne.

On a, en exploitant la relation (2.3),

$$F(s^*, \delta)_{ii} = 2(Bs^*)_i - {}^ts^*Bs^*.$$

En remplaçant s^* par $B^-1_n/{}^t1_nB^-1_n$, on obtient:

$$c = -({}^t1_nB^-1_n)^{-1}.$$ ■

Remarques :

r1) Ce choix de s détermine de façon unique le centre et le rayon de la sphère support de l'image associée à (E, δ). Des choix d'inverses généralisés autres que celui de Moore-Penrose (*cf.* 4), conduisent à un autre centre et donc à un autre rayon.

r2) Dans le cas où ${}^t1_n B^- 1_n = 0$, cela signifie dans le cas où (E, δ) est euclidienne que le rayon de la sphère circonscrivant l'image euclidienne correspondante est infini.

3. POSITIONNEMENT MULTIDIMENSIONNEL NORMÉ

L'intérêt de la représentation sphérique se justifie par sa lisibilité graphique lors du passage par projection orthogonale à des espaces affines de faible dimension.

En effet, si l'on projette un item donné sur un hyperplan de dimension $q \leq p$ passant par X_0 le centre de la sphère portant l'image associée à (E, δ), celui ci sera d'autant mieux représenté, que son projeté se trouve à distance maximale de X_0.

Mieux encore, si l'on travaille sur l'image euclidienne transformée par l'homothétie affine $H_{X_0}, 1/r$, ce qui ne change rien aux positions relatives des points items, un item sera d'autant bien représenté que la distance de son projeté à X_0 est proche de l'unité.

La démarche classique pour parvenir à ce type de résultat s'apparente au MDS non métrique [9]. Elle repose sur l'utilisation des fonctions de similitudes [16]; cela consiste à choisir comme formes de pseudo produits scalaires $G_{c,\varphi}(\delta)$ définies comme suit:

$$\forall\, i, j \in E \quad G_{c,\varphi}(\delta)_{ij} = c - \varphi(\delta_{ij})$$
$$\text{avec } c \in R \text{ et } \varphi : R^+ \to R^+, \text{ croissante.}$$

Étant donné δ, le couple (c, φ) est choisi de telle sorte que G soit SDP; cela permet la représentation sphérique de l'étude transformée (E, δ) avec δ^* défini par:

$$\forall\, i, j \in E \qquad \delta^*_{ij} = \sqrt{\varphi(\delta_{ij})}.$$

La sphère support dans ce cas est de rayon $\sqrt{c}$.

La démarche que nous proposons n'utilise aucun codage préalable de la dissimilarité initiale et de ce fait permet de conserver davantage que l'information relative au préordre induit sur les couples d'items par les termes de δ.

On s'intéresse dans ce qui suit à l'analyse sphérique d'une mesure de dissimilarité: étant donnée une structure de donnée de dissimilarité (E, δ), on en cherche dans un premier temps une approximation sphérique (E, δ^*) et dans un deuxième temps, une image dans un hyperplan de dimension inférieure ou égale à q, restituant au mieux les proximités entre items.

Du fait de la difficulté à déterminer l'optimum global, différentes approches numériques de ce problème peuvent être employées. On présente dans ce qui suit, deux approches nous semblant satisfaisantes.

3.1. Approximation au sens des produits scalaires

Posons $L(s^*, \delta) = \frac{1}{r^2} F(s^*, \delta)$ avec $r^2 = -(1_n^t B^{-} 1_n)^{-1}$.
Cette transformation permet de ramener la représentation sphérique, lorsqu'elle est possible, à une représentation sur la sphère unité ce qui ne change pas les positions relatives.

Intégrer le caractère sphérique ou presque dans la lecture de l'image projetée revient à se poser le problème d'approximation suivant :

Il s'agit de déterminer $X \in M_{n,q}(R)$ minimisant la fonction coût

$$Stress1(q) = \|L(S^*, \delta) - X^t X\|_2.$$

Ce problème consiste donc en l'approximation d'une matrice symétrique par une matrice S.D.P de rang q. Le théorème d'Eckart-Young [10] en fournit la solution, en l'occurrence :

– Si l'on note X_i^k la k-ième coordonnée de l'item indexé par i, on aurait

$$X_i^k = U_k^i \sqrt{\max\,(0, \mu_k)}.$$

– Avec : μ_k la k-ième valeur parmi les valeurs propres de $L(s^*, \delta)$ rangées par ordre décroissant.

U_k^i la ième composante du vecteur propre associé à μ_k et normé à l'unité.

3.2. Approximation au sens des distances

Il s'agit de déterminer $X \in M_{n,q}(R)$ minimisant la fonction coût :

$$Stress2(q) = \|\delta - d(X)\|_2$$

Sous la contrainte $\forall\, i \in E \quad \sum_{k=1}^{q} (X_i^k)^2 = -({}^t 1_n B^{-} 1_n)^{-1}$.
Ici $d(X)$ la distance euclidienne entre items, associée à la matrice de coordonnées X *i.e.* $d_{ij}^2(X) = \sum_{k=1}^{q} (X_i^k - X_j^k)^2$.

Une solution minimum local à ce problème est donnée par l'algorithme **SMACOF**, proposé par Heiser-De Leuw [14] ou l'algorithme **PROXSCAL** [5] qui en constitue une version améliorée du point de vue informatique. Du point de vue purement mathématique, une amélioration de ces algorithmes, basée sur l'optimisation de la différence de deux fonctions convexes, est donnée par Pham Dinh Tao-Le Ti hoai An [18].

4. EXEMPLE D'APPLICATION

L'exemple que l'on présente est basé sur un jeu de données artificielles consistant en 5 items fictifs **A**, **B**, **C**, **E**, **F** et une matrice de dissimilarité associée δ.

Les termes de la matrice δ sont obtenus, du calcul de la distance euclidienne entre 5 points portés par la sphère unité de R^5 et dont les composantes sont données par le tableau 1 ci-après.

Le choix de cet exemple permet de juger plus facilement de la qualité de la solution obtenue.

TABLEAU 1

Coordonnées exactes de points items.

A	**B**	**C**	**E**	**F**
1.000	.950	.950	.950	.000
.000	.312	.221	.000	.000
.000	.000	.221	.000	.000
.000	.000	.000	.312	.000
.000	.000	.000	.000	1.000

Ainsi donc la dissimilarité δ dont on cherche l'image euclidienne est donnée par le tableau ci-après.

TABLEAU 2

Dissimilarités inter-items, constituant les données initiales.

B	.315981			
C	.3165154	.2390021		
E	.315981	.4412347	.4416175	
F	1.414214	1.414158	1.414278	1.414158
	A	**B**	**C**	**E**

Lorsque $q \geq 4$, les méthodes d'approximation sphériques présentées en 3.1 et 3.2 restituent l'image exacte ; dans ces cas $Stress1 = Stress2 = 0$.

On obtient des résultats (*cf.* Tab. 3), assez fidèles aux dissimilarités initiales quand on projette les items sur le plan. Les coordonnées (X_1, Y_1) sont obtenues par minimisation de Stress1(2) ; tandis que (X_2, Y_2) sont obtenues par minimisation de Stress2(2).

TABLEAU 3

Projections planes fournies par les 2 méthodes.

	X_1	Y_1	X_2	Y_2
A	.694	.048	.694	.048
B	.686	.048	.694	.048
C	.686	.165	.682	.302
E	.666	.258	.682	.285
F	.714	.030	.812	.067
Stress1(2) = 0.048			Stress2(2) = 0.0053	

REMERCIEMENTS

Je tiens à remercier J.Y. Jaffray, G. Le Calvé et les referees anonymes pour les conseils ayant permis d'aboutir à cette mise en forme. Que soient aussi remerciés T. Foucart, G.Thauront et toute l'équipe du CIR-INRETS pour les moyens informatiques mis à disposition.

RÉFÉRENCES

[1] M. BENAYADE, *Distances of spherical type*, Distancia 92-Rennes- France, S. Joly et G. Le Calvé, Eds., 1990.

[2] F. BENINEL, *Problèmes de représentations sphériques des tableaux de dissimilarité*, Thèse de doctorat, Université Rennes I, 1987.

[3] L. M. BLUMENTHAL, *Theory and Applications of Distance geometry*, Clarendon Press, Oxford, 1953.

[4] J. BOUZITAT, F. FOURGEAUD et B. LENCLUD, *Inverses généralisés de matrices, définitions et propriétés*, Cahier du groupe de mathématiques économiques, Univ. Paris-I, 1980, 3.

[5] F. BUSING, J. COMMANDEUR et W. J. HEISER, *PROXSCAL: A Multidimensional Scaling Program for individual Differences Scaling with constraints*, Rapport Technique, Department of Data Theory, Leiden University, Holland, 1996.

[6] F. CAILLIEZ, The analytical solution of the additive constant problem, *Psychometrika*, 1983, *48*, n° 2, p. 305-308.

[7] F. CRITCHLEY, On certain linear mappings between inner-product and squared-distance matrices, *Linear Algebra Appl.*, 1988, *105*, p. 91-107.

[8] F. CRITCHLEY et B. FICHET, On (Super-)Spherical distance matrices and two results from Schoenberg, *Linear Algebra and its applications*, 1997, *251*, p. 145-165.

[9] J. DE LEUW et W. J. HEISER, *Theory of multidimensional scaling*, P. R. Krishnaiah et L. N. Kanal, Eds., Handbook of statistics, Vol. 2, Amsterdam: North-Holland company, 1982.

[10] C. ECKART et G. YOUNG, The approximation of one matrix by another of lower rank, *Psychometrika*, 1936, *1*, p. 211-218.

[11] B. FICHET et G. LE CALVÉ, Structure géométrique des principaux indices de dissimilarité sur signes de présence-absence, *Statist. Anal. Données*, 1984, *9*, p. 11-44.

[12] A. A. GBEGAN, *Une présentation globale de l'analyse factorielle des correspondances pour diverses mètriques*, Thèse de doctorat de l'université de Provence, 1984.

[13] J. C. GOWER, *Euclidean distance geometry*, Math. Scientist, 1982, *7*, p. 1-14.

[14] W. J. HEISER et J. DE LEUW, *How to use SMACOF-I*, Rapport de recherches, Departement of data theory, University of Leiden, Holland, 1977.

[15] S. JOLY et G. LE CALVÉ, Étude des puissances d'une distance, *Statist. Anal. Données*, 1986, *11*, p. 30-50.

[16] G. LE CALVÉ, Similarities functions, Edwards D., Raum N.E., Eds., *Proceedings of COMPSTAT 88*, Physica Verlag, Heidelberg, 1988, p. 341-347.

[17] J. C. LINGOES, Some boundary conditions for a monotone analysis of symetric matrices, *Psychometrika*, 1971, *36*, p. 195-203.

[18] Tao. PHAM DINH, An. LE THI HOAI, *Stability of lagrangian duality in D.C. optimization*, Solution of Multidimensional scaling problem by D.C. algorithms, in "Local and Global Approaches to Nonconvex Optimization", Book Series "Nonconvex Optimization and its Applications", Kluwer Academic Publishers, 1999.

[19] C. R. RAO, *Linear statistical inference and its applications*, Wiley series in probability and mathematical statistics, 1973.

[20] I. J. SHOENBERG, Remarks to Maurice Frechet article "Sur la définition axiomatique d'une classe d'espaces distanciés vectoriellement applicable sur l'espace de Hilbert", *Ann. of Math.*, 1935, *38*, p. 724-732.

[21] I. J. SHOENBERG, On certain metric spaces arising from Euclidean spaces by change of metric and their embedding in Hilbert space, *Ann. of Math.*, 1937, *38*, p. 787-793.

[22] B. VAN CUTSEM *et al.*, *Classification and Dissimilarity Analysis*, Springer Verlag, Heidelberg, Lectures Notes in Statistics 93, 1994, p. 341-347.