

REVUE DE STATISTIQUE APPLIQUÉE

E. M. QANNARI

PH. COURCOUX

M. LEJEUNE

O. MAYSTRE

Comparaison de trois stratégies de détermination d'un compromis en évaluation sensorielle

Revue de statistique appliquée, tome 45, n° 1 (1997), p. 61-74

http://www.numdam.org/item?id=RSA_1997__45_1_61_0

© Société française de statistique, 1997, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

COMPARAISON DE TROIS STRATÉGIES DE DÉTERMINATION D'UN COMPROMIS EN ÉVALUATION SENSORIELLE

E.M. Qannari, Ph. Courcoux, M. Lejeune, O. Maystre

*ENITIAA/INRA, Unité de Statistique Appliquée à la Caractérisation des Aliments, Domaine
de la Géraudière, 44322 Nantes cedex 03, France*

RÉSUMÉ

Les méthodes Statis et Procruste sont bien adaptées pour l'analyse des données de profils sensoriels avec un vocabulaire fixé ou libre. Afin de permettre une comparaison directe des résultats de Statis et Procruste, nous proposons une reformulation de la méthode Statis qui conduit à des facteurs d'échelles et à un compromis comparables à ceux de Procruste. Nous proposons aussi un nouvel algorithme pour la méthode Procruste qui s'inspire de la méthode Statis. Celui-ci est plus direct que l'algorithme classique et montre explicitement, à l'instar de Statis, que chaque facteur d'échelle est en fait le produit d'un poids associé au juge et d'un facteur de normalisation. Les trois stratégies (Procruste, Statis et la variante de Procruste) sont comparées sur la base de deux exemples concrets relatifs à la perception sensorielle de yaourts et de pommes.

ABSTRACT

Procrustes and Statis are two statistical methods which are appropriate for the analysis of sensory profiling data. We propose a new presentation of Statis that leads to results similar to those obtained by means of Procrustes. We also discuss a new algorithm for Procrustes which explicitly shows that, similarly to Statis, each isotropic scaling factor is actually a product of a weight associated with the panellist and a scalar factor that adjusts for differences in scaling. The three strategies are compared on the basis of real data sets.

I. Introduction

Dans beaucoup de situations, l'utilisateur est intéressé par la détermination d'un compromis pour un ensemble de tableaux de données. Nous nous situons ici dans le contexte de l'évaluation sensorielle afin de souligner les types de problèmes rencontrés et les objectifs poursuivis en relation avec les méthodes de traitement des données.

L'évaluation sensorielle par un jury consiste souvent à déterminer la nature et le degré des différences organoleptiques entre des produits. L'épreuve du profil sensoriel vise à attribuer des notes d'intensité aux différents produits selon une liste de

descripteurs (variables). Il existe deux procédures : profil conventionnel et profil libre (Williams et Langron, 1984). Avec la première, les juges utilisent une liste préétablie de descripteurs afin d'évaluer les produits considérés. Pour les profils libres, les listes des descripteurs ne sont pas nécessairement les mêmes d'un juge à un autre. Dans les deux cas, la détermination d'un compromis permet d'établir une carte sensorielle des produits ou de relier les aspects organoleptiques aux aspects physico-chimiques ainsi qu'aux données de préférences...

Bien que les juges soient sélectionnés sur la base de leurs aptitudes sensorielles et entraînés pour harmoniser leurs évaluations, la méthode de détermination d'un compromis doit prendre en considération les dérives constatées dans ce genre d'épreuves. En particulier :

- le centrage de chaque tableau permet de supprimer l'effet juge qui se traduit par des niveaux de notation différents;

- l'introduction d'un facteur d'échelle est dictée par le fait que certains juges ont tendance à utiliser toute l'étendue de l'échelle alors que d'autres n'en utilisent qu'une partie restreinte;

- l'ajustement des tableaux par rotation ou en considérant les matrices des produits scalaires entre les produits permet de tenir compte d'éventuelles différences d'interprétations des descripteurs (profil conventionnel) ou du fait que les descripteurs ne sont pas les mêmes pour les différents juges (profil libre).

Il peut aussi être judicieux de tenir compte de la performance des juges en affectant un poids faible dans la détermination du compromis aux juges identifiés comme peu performants. Cette préoccupation est explicitement déclarée dans la méthode Statis (Lavit, 1988; Schlich, 1993); elle est, par contre, seulement «pressentie» dans la méthode Procruste (Collins, 1991).

Nous discutons une présentation originale de la méthode Statis (normée) qui permet de déterminer des facteurs d'échelle et un compromis directement comparables à ceux donnés par Procruste. Puis, nous présentons une variante de la méthode Procruste qui s'inspire de la méthode Statis. Cet algorithme est plus simple que celui proposé par Gower (1975) et montre, à l'instar de Statis, que chaque facteur d'échelle est le produit d'un poids associé à chaque juge et d'un facteur de normalisation.

L'analyse conjointe de plusieurs tableaux a fait l'objet de nombreuses études. Lavit (1988) donne une revue bibliographique et développe en particulier la méthode Statis. Glaçon (1981) entreprend une comparaison de différentes méthodes d'analyse conjointe de tableaux de données. Cette étude met en évidence les liens méthodologiques qui existent entre les différentes techniques : représentation globale, compromis, représentation détaillée. Nous pouvons aussi citer le travail de Bourgeois (1980) qui fait la synthèse des propriétés mathématiques liées à la méthode Procruste et discute quelques applications. Enfin, Lafosse (1985) développe de nouvelles techniques autour de l'analyse procrustéenne de deux tableaux.

II. DONNÉES

Soit m le nombre de juges impliqués dans une épreuve de détermination de profils sensoriels relatifs à n produits. A chaque juge i , est associé le tableau X_i

dont les lignes correspondent aux n produits évalués et les colonnes aux descripteurs sensoriels. Lorsqu'il s'agit d'une épreuve de détermination de profils sensoriels avec un vocabulaire libre, le nombre des colonnes des tableaux $X_i (i = 1, 2, \dots, m)$ n'est pas toujours le même. Nous supposons, cependant, que tous les tableaux sont de dimension $n \times p$, en complétant, le cas échéant, certains tableaux par des colonnes de zéros. Ceci n'a pas d'incidence sur les résultats car les configurations ne sont pas altérées du point de vue des distances entre les individus (les produits).

Toutes les configurations considérées sont supposées centrées.

III. ANALYSE DE PROCRUSTE GÉNÉRALISÉE

Cette méthode est bien connue sous l'acronyme GPA (*Generalised Procrustes Analysis*). Le problème de détermination d'un compromis C des configurations X_1, X_2, \dots, X_m après ajustement par des rotations et des facteurs d'échelle peut être traduit par le problème d'optimisation suivant :

$$\text{Min}_{\{C, \alpha_i, H_i\}} \sum \|\alpha_i X_i H_i - C\|^2; \quad (P1)$$

où $\|\cdot\|$ est la norme de Hilbert-Schmidt définie pour une matrice A par $\|A\|^2 = \text{trace}(A^t A)$, C désigne la configuration compromise, H_i la rotation qui ajuste la configuration X_i au compromis et enfin, α_i le facteur d'échelle associé à X_i .

Une contrainte de détermination doit être imposée aux coefficients $\alpha_i (i = 1, 2, \dots, m)$ afin d'éviter des solutions triviales. Gower (1975) a choisi la contrainte suivante :

$$\sum \alpha_i^2 \text{trace}(X_i^t X_i) = \sum \text{trace}(X_i^t X_i) (*).$$

Cette contrainte impose, en définitive, que l'inertie totale (somme des inerties de toutes les configurations) soit conservée lorsque l'on affecte les facteurs d'échelle aux configurations.

Pour la résolution du problème (P1), Gower propose une heuristique consistant en une procédure itérative qui ajuste de proche en proche les configurations à l'aide de rotations et de facteurs d'échelle.

Pour une description détaillée de cet algorithme, nous renvoyons à l'article de Gower (1975) et à un article écrit par Arnold et Williams (1986). Schlich (1989) a écrit un programme en langage SAS/IML pour la mise en oeuvre de GPA.

IV. MÉTHODE STATIS

Le recours aux matrices des produits scalaires au lieu des configurations elles mêmes présente l'avantage de se dispenser de déterminer explicitement les rotations qui ajustent les configurations. Le problème de détermination d'un compromis est, par conséquent, nettement simplifié.

Soit $W_i = X_i {}^t X_i$ la matrice des produits scalaires associée à la configuration X_i . La matrice W_i est caractéristique du positionnement mutuel des individus (produits) (Escoufier, Pagès, Cazes, 1976; L'Hermier des Plantes, 1976; Glaçon, 1981).

Soit $Y_i = \gamma_i X_i$ la configuration obtenue à partir de X_i en multipliant par le scalaire γ_i (facteur d'échelle). $W_i^* = Y_i {}^t Y_i = \gamma_i^2 W_i$ désigne la matrice des produits scalaires associée à la configuration pondérée Y_i . On propose de déterminer des scalaires γ_i et un opérateur compromis W , de sorte à minimiser :

$$\text{Min}_{\{\gamma_i, W\}} \Sigma \|\gamma_i^2 W_i - W\|^2; \quad (P2)$$

sous la contrainte :

$$\Sigma \text{trace} (\gamma_i^2 W_i \gamma_i^2 W_i) = \Sigma \text{trace} (W_i W_i) \quad (**)$$

Le choix de la contrainte (**) se justifie par la commodité des calculs qu'elle induit à l'instar de la contrainte (*) de Procruste. Ce choix se justifie aussi par un souci de conservation de l'«inertie» totale.

Soit $Q = \Sigma \|\gamma_i^2 W_i - W\|^2$, le minimum de Q est atteint pour :

$$W = \frac{1}{m} \Sigma \gamma_i^2 W_i;$$

ce qui permet d'écrire Q sous la forme :

$$Q = \sum_i \text{trace} (\gamma_i^2 W_i \gamma_i^2 W_i) - \frac{1}{m} \sum_i \sum_j \gamma_i^2 \gamma_j^2 \text{trace} (W_i W_j).$$

Nous sommes par conséquent conduits à maximiser, relativement à γ_i ($i = 1, 2, \dots, m$), la quantité :

$$S = \sum_i \sum_j \gamma_i^2 \gamma_j^2 \text{trace} (W_i W_j);$$

sous la contrainte (**).

Le maximum de S est atteint pour :

$$\gamma_i^2 = \frac{v_i \sqrt{T_w}}{\sqrt{\text{trace} (W_i W_i)}};$$

où $T_w = \Sigma \text{trace} (W_i W_i)$ et v_i est la $i^{\text{ème}}$ coordonnée du premier vecteur propre de la matrice RV dont l'élément générique est donné par :

$$RV(i, j) = RV(X_i, X_j) = \frac{\text{trace}(W_i W_j)}{\sqrt{\text{trace} (W_i W_i)} \sqrt{\text{trace} (W_j W_j)}}.$$

$RV(X_i, X_j)$ est un indice d'association entre les configurations X_i et X_j qui a été introduit par Robert et Escoufier (1976).

Nous pouvons noter que γ_i^2 a bien un sens car dans la formule donnant γ_i^2 , tous les v_i peuvent, d'après le théorème de Perron-Frobenius, être choisis positifs (ou nuls).

Etant semi-défini positif, W peut s'écrire sous la forme : $W = C^t C$. Le tableau C ainsi obtenu constitue une configuration compromis de X_1, X_2, \dots, X_m . Il existe en fait une infinité de tableaux tels que $W = C^t C$.

Le problème d'optimisation (P2) conduit par conséquent à la méthode Statis normée. La contrainte (**) induit *de facto* une normalisation du compromis qui le rend comparable aux configurations initiales.

Il apparaît que le facteur d'échelle γ_i est le produit d'un facteur de normalisation, $\left(\frac{\sqrt{TW}}{\sqrt{\text{trace}(W_i W_i)}}\right)^{\frac{1}{2}}$, et d'un poids qui pondère la configuration du juge i , donné par $\sqrt{v_i}$. Du fait de cette pondération, le compromis de Statis a été qualifié comme étant un résumé majoritaire des configurations (Lavit, 1988).

En résumé, la méthode Statis (normée) pourrait être présentée dans les mêmes termes que la méthode Procruste : détermination de facteurs d'échelle $\gamma_i (i = 1, 2, \dots, m)$ et d'un compromis C , le critère d'optimisation étant donné par :

$$\text{Min}_{\{\gamma_i, C\}} \sum \|(\gamma_i X_i)^t (\gamma_i X_i) - C^t C\|^2; \tag{P2}$$

sous la contrainte (**):

$$\sum \gamma_i^4 \text{trace} [(X_i^t X_i)(X_i^t X_i)] = \sum \text{trace} [(X_i^t X_i)(X_i^t X_i)].$$

V. UNE VARIANTE DE PROCRUSTE

L'heuristique proposée par Gower pour résoudre le problème (P1) conduit à une solution satisfaisante bien que ce ne soit pas toujours la solution optimale. Ten Berge (1977), Ten Berge et Bekker (1993) et Bourgeois (1980) ont discuté de nouvelles procédures pour déterminer les facteurs d'échelle associés aux configurations. A chaque étape de la procédure de Ten Berge, les facteurs d'échelles sont déterminés à partir du premier vecteur propre de la matrice dont l'élément générique est :

$$\frac{\text{trace} ({}^t Y_i Y_j)}{\sqrt{\text{trace} ({}^t Y_i Y_i)} \sqrt{\text{trace} ({}^t Y_j Y_j)}}$$

où Y_i est la matrice obtenue à partir de X_i en appliquant des rotations calculées à l'étape précédente de l'algorithme. Cette procédure a été critiquée par Gower sur la base de la complexité des calculs qu'elle entraîne (Gower, 1984, p. 773).

La détermination d'une solution directe du problème (P1) est rendue difficile par le nombre important de paramètres à estimer. La démarche suivante, inspirée de Statis, contourne cette difficulté en procédant en deux étapes :

- étape 1 : détermination de facteurs d'échelle optimaux (procédure non itérative);
 étape 2 : détermination d'un compromis par application de l'algorithme de Kristof et Wingersky (1971); cet algorithme est connu pour son efficacité et sa rapidité de convergence.

Etape 1 : détermination des facteurs d'échelle

Nous proposons de déterminer des facteurs d'échelle α_i ($i = 1, 2, \dots, m$) de sorte à minimiser la quantité :

$$F = \sum_i \sum_j \|\alpha_i X_i - \alpha_j X_j A_{ij}\|^2;$$

où A_{ij} est une rotation optimale qui ajuste X_j à X_i . La détermination de cette rotation est basée sur la décomposition spectrale de $X_j {}^t X_i$ (voir par exemple, Sibson, 1978; Glaçon, 1981).

Sur les coefficients α_i , nous imposons la contrainte de Procruste (Gower, 1975), à savoir :

$$\Sigma \alpha_i^2 \text{trace} (X_i {}^t X_i) = \Sigma \text{trace} (X_i {}^t X_i) (*).$$

Il est facile de vérifier que le problème de minimisation de F sous la contrainte (*) équivaut à maximiser la quantité :

$$\sum_i \sum_j \alpha_i \alpha_j \text{trace} ({}^t X_i X_j A_{ij});$$

sous la même contrainte.

La solution est donnée par :

$$\alpha_i = \frac{u_i \sqrt{T}}{\sqrt{\text{trace} ({}^t X_i X_i)}};$$

où $T = \Sigma \text{trace} ({}^t X_i X_i)$ et u_i est la $i^{\text{ème}}$ coordonnée du premier vecteur propre de la matrice symétrique R dont l'élément générique est :

$$r_{ij} = \frac{\text{trace} ({}^t X_i X_j A_{ij})}{\sqrt{\text{trace} ({}^t X_i X_i)} \sqrt{\text{trace} ({}^t X_j X_j)}}.$$

Cet indice, compris entre 0 et 1, a été utilisé par Lingoes et Schönemann (1974). Il est nul si et seulement si les configurations X_i et X_j sont dans des espaces

orthogonaux; il est égal à 1 si et seulement si les deux configurations peuvent être ajustées par une rotation et un facteur d'échelle (Glaçon, 1981, p. 63).

Les coefficients $\alpha_i (i = 1, 2, \dots, m)$ peuvent tous être choisis non négatifs car les éléments de la matrice R sont positifs ou nuls (théorème de Perron-Frobenius).

Il apparaît, là aussi, que chaque facteur d'échelle α_i est le produit d'un facteur de normalisation : $\frac{\sqrt{T}}{\sqrt{\text{trace}({}^t X_i X_i)}}$ et d'un poids, donné par u_i , qui pondère chaque configuration en fonction de son degré d'accord avec les autres. Les configurations des juges peuvent être comparées sur la base de la matrice des similarités $R = (r_{ij})(i, j = 1, 2, \dots, m)$, de la même manière que Statis les compare à l'aide de la matrice contenant les coefficients RV . De même, la première valeur propre de R peut servir d'indice de concordance globale des juges.

Etape 2 : Détermination d'un compromis

Pour déterminer un tableau compromis, nous proposons d'appliquer l'algorithme de Kristof et Wingersky (1971) sur les configurations $Y_i = \alpha_i X_i (i = 1, 2, \dots, m)$. Cet algorithme procède de la manière suivante :

Pas n° 1 : Ajuster à l'aide d'une rotation chaque configuration Y_1, Y_2, \dots, Y_m à Y_1 et déterminer un compromis C en considérant la moyenne des configurations ajustées;

Pas n° 2 : Ajuster de nouveau les configurations Y_1, Y_2, \dots, Y_m à C , et réinitialiser le compromis en considérant la moyenne des configurations ainsi ajustées.

Par la suite le *pas n° 2* est réitéré jusqu'à ce que l'algorithme converge (variation non significative de C).

Nous obtenons un compromis qui s'écrit, *in fine*, comme suit :

$$C = \frac{1}{m} \sum_i \alpha_i X_i H_i = \frac{1}{m} \sum_i u_i X_i^* H_i$$

où X_i^* désigne la configuration normée associée à X_i et H_i , la rotation optimale obtenue à l'issue de l'étape 2. Cette écriture met en évidence le rôle des poids u_i car une configuration pour laquelle ce poids est relativement important contribue davantage à la détermination du compromis qu'une configuration pour laquelle ce poids est faible. Cette écriture est à mettre en parallèle avec la formule donnant l'opérateur compromis de Statis en tant que moyenne pondérée des opérateurs associés aux configurations $X_i (i = 1, 2, \dots, m)$.

VI. COMPARAISON DES MÉTHODES

6.1 Rang du compromis

Le compromis obtenu à partir de Procruste que ce soit à l'aide de l'algorithme classique ou de l'algorithme que nous proposons s'écrit :

$$C = \frac{1}{m} \sum_i \alpha_i X_i H_i$$

Soit p_i le nombre de variables (non nulles) du tableau X_i ; le rang de C est au plus égal à $p = \max(p_i)$.

Le compromis de Statis est obtenu à partir de la décomposition spectrale de l'opérateur :

$$W = \frac{1}{m} \sum_i \gamma_i^2 X_i {}^t X_i.$$

Cela revient, par conséquent, à effectuer une analyse en composantes principales (ACP) du tableau Y obtenu en juxtaposant horizontalement les tableaux $Y_i = \frac{\gamma_i}{\sqrt{m}} X_i (i = 1, 2, \dots, m)$. Ainsi le rang du compromis de Statis peut être dans certains cas aussi grand que $P = \sum p_i$. Bien évidemment, les premières composantes de ce compromis sont liées de manière privilégiée aux variables «communes» des tableaux $X_i (i = 1, 2, \dots, m)$ par le fait des pondérations et par le fait qu'étant liées, ces variables contribuent beaucoup à la détermination des premières composantes de Y . Le compromis de Statis apparaît donc comme étant un tableau synthétique car il contient toutes les dimensions sous-jacentes à tous les tableaux pour lesquels les poids sont non nuls.

La démarche de Statis rappelle l'Analyse Factorielle Multiple (AFM) développée par Escofier et Pagès (1988) qui consiste à effectuer l'ACP du tableau Z obtenu en juxtaposant horizontalement les tableaux $Z_i = \frac{1}{\sqrt{\lambda_i}} X_i (i = 1, 2, \dots, m)$ où λ_i est la première valeur propre de $X_i {}^t X_i$. Ainsi, seule la normalisation des tableaux est préconisée en AFM (pas de pondération en fonction des accords mutuels entre les tableaux). Pagès (1995) a comparé les caractéristiques techniques de Statis et de l'AFM.

Nous déduisons de ce qui précède que les méthodes Statis et Procruste diffèrent beaucoup par les pourcentages d'inertie restituée par les composantes principales du compromis. L'utilisateur doit avoir à l'esprit que ces différences proviennent essentiellement du fait que Statis rend compte de plus d'axes que Procruste. En tout état de cause, les pourcentages d'inertie expliquée par les axes associés à l'ACP du compromis doivent être interprétés avec précaution. Ce point sera étayé par une application concrète.

6.2. Choix des indices de dissimilarité et de similarité

Pour Statis, la dissimilarité entre deux configurations X et Y est évaluée par la distance euclidienne :

$$d(X, Y) = \|X {}^t X - Y {}^t Y\|;$$

pour Procruste, cette dissimilarité est évaluée par la distance :

$$\delta(X, Y) = \|X - Y A\|$$

où A est une rotation procrustéenne de Y par rapport à X . Cette dernière distance n'est pas euclidienne (Sibson, 1978). La différence entre les distances d et δ est dans une certaine mesure de même nature que celle existant entre les distances usuelles associées respectivement aux espaces L^2 et L^1 .

Parallèlement, les similarités entre les configurations X et Y sont évaluées respectivement pour Statis et Procruste par :

$$\text{Cov}(X, Y) = \text{trace}(X^t X Y^t Y)$$

et

$$T(X, Y) = \text{trace}(^t X Y A)$$

ou par les indices normalisés associés : coefficient RV (Robert et Escoufier, 1976) et indice de Lingoes Schönemann (1974). En désignant par $\mu_i (i = 1, 2, \dots, p)$ les valeurs singulières de la matrice $^t X Y$ classées par ordre décroissant, $\text{Cov}(X, Y)$ et $T(X, Y)$ peuvent s'écrire (voir par exemple Lingoes et Schönemann (1974) et Glaçon (1981)) :

$$\text{Cov}(X, Y) = \sum_{i=1}^p \mu_i^2$$

$$T(X, Y) = \sum_{i=1}^p \mu_i$$

Il est facile de montrer, en supposant que $T(X, Y) > 0$, les inégalités suivantes :

$$\frac{1}{p} \leq \frac{\text{Cov}(X, Y)}{T^2(X, Y)} \leq 1$$

Cette inégalité a été démontrée et interprétée dans le cas $X = Y$ par Kazi-Aoual (1993). La borne inférieure est atteinte si et seulement si toutes les valeurs $\mu_i (i = 1, 2, \dots, p)$ sont égales et la borne supérieure est atteinte si et seulement si μ_1 est la seule valeur singulière non nulle. Ainsi le rapport :

$$c(X, Y) = \frac{\text{Cov}(X, Y)}{T^2(X, Y)}$$

peut être considéré comme un indicateur de la concentration du spectre $\mu_i (i = 1, 2, \dots, p)$. En d'autres termes, pour $T(X, Y)$ fixé, la quantité $\text{Cov}(X, Y)$ est d'autant plus grande que les valeurs $\mu_i (i = 1, 2, \dots, p)$ sont dispersées. Ceci montre que la quantité $\text{Cov}(X, Y)$ mesure plus qu'un ajustement par rotation entre X et Y : elle rend aussi compte de la qualité d'ajustement sur des espaces de faibles dimensions (grande dispersion des valeurs singulières).

Nous en déduisons que Statis a tendance à privilégier davantage que Procruste les accords entre configurations sur les axes d'inertie relativement importante. Ce point sera illustré dans le cadre d'une application.

VII. APPLICATION

Première application

Les données traitées sont extraites de Dijksterhuis et Gower (1991), où elles ont servi pour illustrer la méthode Procruste. Elles concernent l'évaluation sensorielle de 8 variétés de yaourts. Le protocole expérimental utilisé est le profil libre (free choice profiling) où les juges utilisent leurs propres descripteurs pour évaluer les aspects organoleptiques.

Le tableau 1 donne les poids associés aux juges obtenus à partir de Statis et de la variante de Procruste. Ces indices sont relativement proches.

Le tableau 1 donne aussi les facteurs d'échelle associés aux différentes configurations obtenus par les 3 stratégies. Là aussi, on s'aperçoit que les résultats sont comparables.

TABLEAU 1

*Indices de performance et facteurs d'échelle.
(Les indices de performance ont été relativisés
en divisant par l'indice de performance moyen)*

Juge	Poids des juges		Facteurs d'échelle		
	Variante Procruste	Statis	Procruste	Statis	Variante Procruste
A	1,008	1,014	1,412	1,449	1,416
B	1,020	1,047	1,036	1,036	1,032
C	1,028	1,054	0,776	0,819	0,774
D	1,007	1,010	1,225	1,216	1,231
E	1,002	1,043	0,992	0,989	0,992
F	0,965	0,924	0,926	0,969	0,929
G	0,971	0,910	0,987	1,051	0,983

Par la suite, nous avons effectué une ACP sur les compromis obtenus à l'aide des trois stratégies. La figure 1 donne la représentation graphique des yaourts sur la base du premier plan principal des trois ACP. Les représentations graphiques sont dans une large mesure similaires.

Deuxième application

Dans la première application, nous avons davantage mis en exergue les points communs entre les méthodes. Dans cette deuxième application, nous nous attachons à mettre en évidence leurs spécificités. Les données analysées sont extraites de la thèse de Langron (1981) et concernent l'évaluation sensorielle de 27 types de pommes sur la base d'un vocabulaire fixé comprenant 14 descripteurs (profil conventionnel). Huit juges ont participé à l'épreuve. Nous avons traité les données à l'aide de la méthode

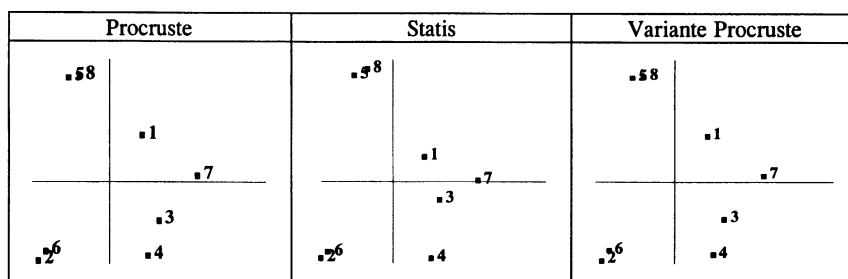


FIGURE 1

Représentation des produits sur la base du premier plan principal des compromis.

Statis et de la méthode Procruste en utilisant l’algorithme que nous avons discuté. Les résultats se recourent dans une large mesure et recourent ceux de GPA discutés de manière détaillée par Langron (1981). Soulignons cependant que les pourcentages d’inertie restituée par les axes sont très différents (tableau 2). Ceci provient du fait que Statis génère plus d’axes que Procruste. En effet, comme nous l’avons signalé

TABLEAU 2
Valeurs propres et inerties restituées (%)
par les composantes principales des compromis.

Procruste					
Dimension	1	2	3	4	5
valeur propre	35,86	19,13	12,37	11,80	9,02
inertie restituée (%)	28,92	15,43	9,97	9,51	7,28
Variante de Procruste					
Dimension	1	2	3	4	5
valeur propre	35,93	19,55	12,20	11,85	8,76
inertie restituée (%)	28,90	15,70	9,80	9,50	7,10
STATIS					
Dimension	1	2	3	4	5
valeur propre	37,68	22,38	17,87	16,20	12,20
inertie restituée (%)	18,20	10,80	8,60	7,80	5,90

au paragraphe 6.1, Statis rend compte de toutes les dimensions sous-jacentes à tous les tableaux. Pour cet exemple, le rang du compromis de Procruste était inférieur à 14 alors que qu'il était égal à 20 pour le compromis de Statis. Sur le tableau 2, nous pouvons aussi remarquer que les valeurs propres associées à l'ACP du compromis de Statis sont relativement supérieures à celles associées au compromis de Procruste.

Afin de mettre en évidence le fait que Statis attribue relativement plus de poids aux configurations qui s'accordent avec les autres sur la base des axes de plus grandes inerties, nous avons ajouté une neuvième configuration fictive aux 8 configurations associées respectivement aux 8 juges. Cette neuvième configuration est celle associée aux $k(k = 1, 2, \dots, 6)$ premières composantes principales du compromis obtenu à l'aide de GPA (les résultats discutés sont les mêmes quelque soit la méthode choisie pour la détermination du compromis). Pour chaque valeur de k allant de 1 à 6, nous avons effectué Statis et Procruste sur les 9 configurations. Pour la méthode Procruste, nous avons appliqué l'algorithme que nous avons proposé car il a l'avantage d'exhiber des poids associés aux juges. L'objectif est de comparer les poids accordés par les deux stratégies en fonction de la valeur de k . Nous remarquons (tableaux 3 et 4) que dès $k = 1$, le poids attribué à la neuvième configuration est le plus grand des poids de tous les juges pour la méthode Statis alors que ce poids est le plus petit de tous les poids pour la méthode Procruste. Ce constat permet d'étayer le point discuté au paragraphe 6.2 concernant le fait que Statis privilégie davantage que Procruste les accords entre configurations sur la base des axes de plus grandes inerties. Pour les deux méthodes le poids attribué à la neuvième configuration croît avec k mais c'est seulement à partir de $k = 4$ que ce poids devient le plus important des poids de tous les juges pour la méthode Procruste.

TABLEAU 3
*poids normalisés (somme égale à 9) obtenus par Statis
pour les 9 configurations en fonction du rang
de la neuvième configuration fictive.*

valeur de k juge	1	2	3	4	5	6
1	0,921	0,891	0,885	0,872	0,866	0,864
2	1,096	1,085	1,076	1,060	1,057	1,061
3	1,017	1,012	1,025	1,011	1,002	0,994
4	0,993	0,961	0,946	0,945	0,957	0,952
5	0,807	0,782	0,807	0,805	0,811	0,809
6	1,011	0,991	0,979	0,972	0,966	0,961
7	1,023	1,018	0,997	1,011	1,005	1,009
8	1,017	1,036	1,012	1,011	1,005	1,003
fictif : 9	1,114	1,224	1,272	1,314	1,332	1,346

TABLEAU 4
poids normalisés (somme égale à 9) obtenus par la variante de Procruste pour les 9 configurations en fonction du rang de la neuvième configuration fictive.

valeur de <i>k</i> juge	1	2	3	4	5	6
1	0,984	0,960	0,958	0,939	0,934	0,931
2	1,054	1,038	1,033	1,020	1,015	1,018
3	1,060	1,044	1,045	1,035	1,027	1,018
4	1,032	1,005	0,994	0,987	0,994	0,985
5	0,981	0,957	0,964	0,960	0,958	0,955
6	1,032	1,014	1,003	0,996	0,988	0,982
7	1,035	1,023	1,003	1,005	1,003	1,003
8	1,035	1,029	1,006	1,002	0,997	0,994
fictif : 9	0,786	0,927	0,997	1,053	1,087	1,114

VIII. CONCLUSION

Le travail discuté ici a pour intérêt de rapprocher les méthodes Statis et Procruste. D'un côté, la méthode Statis est présentée dans le même esprit que Procruste (facteurs d'échelle affectés aux configurations, détermination d'un compromis optimal, ...). De l'autre côté, une nouvelle procédure, proche de la démarche de Statis, est proposée pour la méthode Procruste. Outre sa simplicité, ce nouvel algorithme fournit des poids associés aux configurations au même titre que Statis. Nous avons aussi souligné quelques spécificités des méthodes en montrant notamment le fait que Statis a tendance à privilégier davantage que Procruste les accords entre configurations sur la base des axes de plus grandes inerties.

Références bibliographiques

- Arnold G. M. et Williams, A. A. (1986), *The Use of Generalised Procrustes Analysis in Sensory Analysis*. In *Statistical Procedures in Food Research*, ed. J.R. Piggott, North Holland, Amsterdam.
- Bourgeois P. (1980), *Recherche du déplacement minimisant la distance entre deux configurations de points indicés par un même ensemble fini, Méthodes et applications en reconnaissance des formes et en analyse des données cubiques*, Thèse, Université Pierre et Marie Curie – Paris VI.
- Collins A. J. (1991), *The Use of Generalised Procrustes Techniques in Sensory Analysis*, Comptes rendus de «Agro Industrie et Méthodes Statistiques» – Nantes.
- Dijksterhuis G. B. et Gower J. C. (1991), *The Interpretation of Generalised Procrustes Analysis and Allied Methods*, *Food Quality and Preference*, 3, 67-87.

- Escofier B., Pagès J. (1988), *Analyses factorielles simples et multiples*, Dunod, Paris.
- Escoufier Y., Pagès J. P., Cazes P. (1976), *Opérateurs et analyse des tableaux à plus de deux dimensions*, Cahiers du Buro, 25, 61-89.
- Glaçon F. (1981), *Analyse conjointe de plusieurs matrices de données; comparaisons de plusieurs méthodes*, Thèse, Université des Sciences de Grenoble.
- Gower J. C. (1975), *Generalised Procrustes Analysis*, Psychometrika, 40, 33-51.
- Gower J. C. (1984), *Multivariate Analysis : Ordination, Multidimensional Scaling and Allied Topics*, in E. Lloyd (Ed.) : Handbook of applicable mathematics VI, Statistics, B, 727-781, Wiley, New York.
- Kazi-Aoual F. (1993), *Approximations to Permutation Tests for Data Analysis*, Rapport n°93-06, Unité de Biométrie, Montpellier.
- Kristof, W. et Wingersky (1971), *Generalisation of the Orthogonal Procrustes Rotation Procedure for More than Two Matrices*, Proceedings of the 79th annual convention of the American Psychological Association, 89-90.
- Lafosse R. (1985), *Analyses Procrustéennes de deux tableaux*, Thèse, Université Paul Sabatier, Toulouse.
- Langron S. P. (1981), *The Statistical Treatment of Sensory Analysis Data*, Thèse Ph. D. Université de Bath.
- Lavit C. (1988), *Analyse conjointe de tableaux quantitatifs*, Masson, Paris.
- L'Hermier des Plantes H. (1976), *Structuration des tableaux à trois indices de la statistique*. Thèse, Univ. des Sciences et Techniques du Languedoc Montpellier II.
- Lingoes J. C., Schönemann P. H. (1974), *Alternative Measures of Fit for the Schönemann-Caroll Matrix Fitting Algorithm*, Psychometrika, 39, 4.
- Pagès J. (1995), *Eléments de comparaison entre l'analyse factorielle multiple et la méthode Statis*, Comptes rendus des XXVII Journées ASU, Jouy-en-Josas.
- Robert P. et Escoufier Y. (1976), *A Unifying Tool for Linear Multivariate Statistical Methods : the RV Coefficient*, Applied statistics, 25, 257-267.
- Schlich P. (1989), *A SAS/IML Program for Generalised Procrustes Analysis*, In Proceedings of the SAS European Users Group Internat, Conference, p. 529-537.
- Schlich P. (1993), *Contribution à la sensométrie*, Thèse, Université Paris Sud.
- Sibson R. (1978), *Studies in the Robustness of Multidimensional Scaling : Procrustes Statistics*, J. Royal. Statist. Soc. B, 40, 234-238.
- Ten Berge J. M. (1977), *Orthogonal Procrustes Rotation for Two or More Matrices*, Psychometrika, 42, 267-276.
- Ten Berge J. M. et Bekker P. A. (1993), *The Isotropic Scaling Factor in Generalised Procrustes Analysis*, Computational Statistics & Data Analysis, 16, North-Holland, 201-204
- Williams A. A. et Langron S. P. (1984), *The Use of Free-Choice Profiling for the Evaluation of Commercial Ports*, J. Sci. Food Agric., 35, 558-568.