

REVUE DE STATISTIQUE APPLIQUÉE

MARIE CHAVENT

CHRISTIANE GUINOT

YVES LECHEVALLIER

MICHEL TENENHAUS

Méthodes divisives de classification et segmentation non supervisée : recherche d'une typologie de la peau humaine saine

Revue de statistique appliquée, tome 47, n° 4 (1999), p. 87-99.

http://www.numdam.org/item?id=RSA_1999__47_4_87_0

© Société française de statistique, 1999, 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/legal.php>). 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/>

MÉTHODES DIVISIVES DE CLASSIFICATION ET SEGMENTATION NON SUPERVISÉE : RECHERCHE D'UNE TYPOLOGIE DE LA PEAU HUMAINE SAINÉ

Marie Chavent^{1,4}, Christiane Guinot², Yves Lechevallier¹, Michel Tenenhaus³

¹ INRIA, Rocquencourt, France

² CERIES, Neuilly-sur-Seine, France

³ Groupe HEC, Jouy-en-Josas, France

⁴ MAB, Université Bordeaux I, France

RÉSUMÉ

Nous proposons deux méthodes de classification hiérarchique descendante, qui se rapprochent, dans un cadre non supervisé, des méthodes de segmentation. En effet, lors de la construction de la hiérarchie par divisions successives de l'ensemble des individus, une classe est toujours divisée en deux sous-classes. Ces deux sous-classes sont obtenues en choisissant une variable et sa transformation en variable binaire. On obtient ainsi une classification facilement interprétable, l'arbre hiérarchique ayant la particularité d'être ici un arbre de décision. La première méthode, DIVAF, sélectionne la variable par analyse des correspondances et utilise les scores factoriels pour transformer cette variable en variable binaire. La deuxième méthode, DIVOP, choisit simultanément la variable candidate et sa transformation par optimisation d'un critère d'évaluation des partitions. L'aspect simple et interprétable de ces deux méthodes a motivé leur application dans le cadre d'une étude sur la peau humaine saine. Le but de cette étude est la recherche d'une classification de la peau du visage, élaborée à l'aide d'outils d'analyse des données, susceptible de contribuer à une approche rationnelle de la cosmétologie et d'intéresser les dermatologues.

Mots-clés : Classification automatique, classification hiérarchique descendante, segmentation, classification monothétique, analyse des correspondances multiples.

ABSTRACT

Two descendant hierarchical clustering methods are proposed in this paper. These methods are, in a non supervised context, closed to the segmentation methods. During the hierarchy construction process, each cluster is always divided in two sub-clusters. The two sub-clusters are obtained by the selection of a particular variable and its binary transformation. An easily interpretable classification is obtained, and, in this case, the resulting hierarchical

tree is a decision tree. In the first method, DIVAF, the variable is selected by using the correspondence analysis method; then, the factorial scores conduct to a binary transformation of the variable. In the second method, DIVOP, the candidate variable and its binary transformation are simultaneously selected by optimization of a criterion which evaluates the resulting partitions. The apparent simplicity of the two methods and the easiness of interpretation of their results conducted to apply these methods to a set of healthy human skin data, to search for a scientifically based classification, which would contribute to a rational approach in cosmetology and which would be potentially useful for dermatologists.

Keywords : *divisive hierarchical clustering methods, unsupervised segmentation, monothetic clustering, multiple correspondence analysis.*

1. Introduction

Les méthodes de classification divisive sont des méthodes de classification hiérarchique. Elles partent de l'ensemble des individus et procèdent par divisions successives des classes jusqu'à l'obtention de classes vérifiant certaines règles d'arrêt. On les appelle aussi méthodes descendantes de classification hiérarchique pour les différencier des méthodes ascendantes qui partent des singletons et qui procèdent par agrégations successives. De plus, la complexité d'un algorithme ascendant est généralement polynomiale, tandis que celle d'un algorithme descendant est généralement exponentielle. En effet, lors de la première étape d'une méthode ascendante, il faut évaluer toutes les agrégations possibles de deux individus parmi n , soit $n(n-1)/2$ possibilités, tandis qu'un algorithme descendant basé sur l'énumération complète évalue toutes les divisions des n individus en deux sous-ensembles non vides, soit $2^{n-1} - 1$ possibilités. Cette stratégie de l'énumération complète, adoptée par Edwards et Cavalli-Sforza (1965) pour le critère de la variance intra-classe, est bien sûr difficilement applicable dès que le nombre n d'individus est supérieur à 20. Différentes approches ont été envisagées pour palier à ce problème de complexité.

On distingue généralement les méthodes de type polythétique des méthodes de type monothétique. Les méthodes monothétiques imposent aux individus d'une même classe de posséder au moins une caractéristique en commun. Les méthodes descendantes de type polythétique sont généralement, comme les méthodes ascendantes, définies à partir d'un tableau de dissimilarités et cherchent des classes d'individus se ressemblant sur une majorité de caractéristiques. Une des plus anciennes méthodes de type polythétique est celle de MacNaughton-Smith et coll. (1964). Ces méthodes cherchent généralement à optimiser un critère local, c'est-à-dire le critère de la bipartition, et non celui de la partition induite sans faire une énumération complète de toutes les bipartitions. Différents critères ont été étudiés comme le critère du diamètre (Hubert (1973), Guénoche, Hansen et Jaumard (1991)), le critère de la distance moyenne intergroupe (Matula (1986), Guénoche (1994), Roux (1995), Hunstman-Labed (1997)) ou encore le critère d'inertie intra-classe (Reinert (1983), Roux (1985)). En 1978, Gowda et Krishna proposent une stratégie différente utilisant la notion des k plus proches voisins mutuels. Un problème intéressant et peu abordé jusqu'à présent est certainement celui de la validité d'une division et par là même du bon nombre de classes (Har-Even et Brailovsky, 1995).

Les méthodes de type monothétique diffèrent des méthodes polythétiques en imposant aux individus d'une même classe de posséder au moins une caractéristique en commun. Cette propriété est généralement obtenue en divisant une classe en fonction d'une variable et d'une dichotomie des valeurs de cette variable. L'intérêt de cette approche est la facilité d'interprétation des classes obtenues. En effet, chaque classe est définie par une conjonction de caractéristiques traduisant une condition nécessaire et suffisante d'appartenance à la classe. La première méthode divisive de type monothétique a été proposée par Williams et Lambert (1959). Cette méthode, définie initialement pour des variables binaires, sépare les individus de modalité 0 des individus de modalité 1 sur une variable choisie en maximisant un critère d'association comme par exemple la somme des khi-deux de la variable aux autres variables. Cette méthode a été étendue au cas des variables qualitatives (Volle, 1976), ce qui crée un nouveau problème de complexité, le nombre de dichotomies augmentant de manière exponentielle avec le nombre de modalités. Depuis, les méthodes monothétiques de classification divisive ont été étudiées en intelligence artificielle dans le cadre des méthodes de classification conceptuelle (Michalsky, Diday et Stepp (1982), Michalsky et Stepp (1983)).

En parallèle, les méthodes divisives de type monothétique ont été développées dans un objectif prédictif sous le nom de méthodes de segmentation. Dans ce cadre, on cherche une liaison entre une variable à expliquer et un ensemble de variables explicatives. La méthode AID (Automatic Interaction Detector) de Sonquist et Morgan (1963) cherche à expliquer une variable quantitative à l'aide d'un ensemble de variables qualitatives. Les méthodes de segmentation, qui cherchent à expliquer une variable qualitative représentant des classes a priori à partir d'un ensemble quelconque de variables, ont été développées en statistique par Breiman, Friedman, Olshen et Stone (1984) à travers le logiciel CART (Classification and Regression Tree). En intelligence artificielle, Quinlan (1986) propose le logiciel ID3 (Induction Decision Tree).

En intelligence artificielle, l'apprentissage non supervisé correspond à la problématique de la classification, tandis que l'apprentissage supervisé correspond à celle de la discrimination. Ainsi les méthodes usuelles de segmentation sont des méthodes supervisées. Nous proposons dans cet article deux méthodes de segmentation non supervisées. Ces deux méthodes divisives font appel à des notions et à des critères issus de l'analyse des données. La première méthode s'appuie à chaque étape sur une analyse factorielle. On l'appellera DIVAF. La deuxième méthode est basée sur l'optimisation de l'inertie intra-classe des bipartitions. On l'appellera DIVOP.

Ces deux méthodes ont été utilisées sur les données d'une étude menée au C.E.R.I.E.S. dont l'objectif était de proposer une typologie fiable de la peau humaine saine reposant sur un petit nombre de caractères cutanés pertinents. Les données ont été recueillies entre avril et mai 1996, sur 212 femmes volontaires d'Ile-de-France présentant une peau apparemment saine et d'âge compris entre 20 et 50 ans. Ces données résultent d'un examen clinique appréciant 17 caractéristiques de la peau de la joue, évaluées sur des échelles qualitatives. Ces signes cliniques peuvent être visuels comme «aspect gras» ou encore «régularité du grain de la peau». Ils peuvent également être tactiles comme «toucher rêche» ou encore «incapacité à rosir même après un léger pincement». Ces variables sont toutes binaires ou ordinales.

Le travail présenté dans cet article fait suite à une première analyse qui avait mis en parallèle deux méthodologies différentes : la classification par la méthode des centres mobiles et la classification par la méthode DIVAF (Guinot, Tenenhaus, Dubourgeat, Le Fur, Morizot, Tschachler, 1997). Dans cet article, les deux méthodes divisives DIVAF et DIVOP sont présentées et mises en parallèle dans le cadre de cette application. Elles sont aussi comparées à d'autres méthodes de classification.

2. La méthode DIVAF

Cette méthode construit une hiérarchie d'un ensemble d'individus, en divisant à chaque étape une classe en fonction d'une variable qualitative à deux ou plus modalités. Cette méthode utilise à chaque étape l'analyse factorielle des correspondances multiples (AFCM), d'une part, pour sélectionner une variable qualitative et, d'autre part, pour choisir le découpage en deux groupes des modalités de cette variable qualitative afin de construire les deux classes.

2.1 Présentation de la méthode

La première étape consiste à effectuer une analyse factorielle sur l'ensemble des individus. On sélectionne dans un premier temps la variable X la plus liée à la première composante principale. Pour mesurer cette liaison, on peut, par exemple, utiliser le rapport de corrélation entre la variable X et la première composante principale ou encore le niveau de signification du F de l'analyse de la variance associée. Dans un deuxième temps, la dichotomie du domaine d'observation de cette variable est définie en opposant les modalités à valeurs positives sur la première composante principale à celles à valeurs négatives. Le critère de choix doit aussi tenir compte des effectifs des classes obtenues qui ne doivent pas être trop faibles.

Au cours des étapes suivantes, on répète l'étape précédente sur les deux classes obtenues, et ainsi de suite. Chaque classe est divisée systématiquement.

Les divisions s'arrêtent lorsque les classes obtenues sont homogènes ou lorsque la classe est trop petite. On considère qu'une classe est homogène lorsque le lien entre chaque variable et la première composante principale est trop faible.

Une particularité intéressante de cette méthode permet à l'expert d'intervenir dans la construction de la hiérarchie : l'expert peut choisir une des variables les plus liées à la première composante pour définir la bipartition.

2.2 Application de la méthode

La méthode DIVAF a permis de définir une typologie des femmes en six classes. La figure 1 indique les divisions successives ayant permis la construction des six classes.

Au début, les 212 femmes se divisent en fonction de la variable «capacité à rosir dès l'attouchement», la première question binaire étant :

[capacité à rosir dès l'attouchement = oui ?]
ou [capacité à rosir dès l'attouchement = non ?]

Puis, les 156 femmes répondant oui à la question binaire précédente se divisent en fonction de la variable «toucher gras», tandis que les 56 femmes ayant répondu non, se divisent en fonction de la variable «couleur jaune», etc. Le critère d'arrêt utilisé ici a été la taille des classes.

A chaque division, les trois variables les plus liées à la première composante principale sont indiquées par ordre décroissant et la variable retenue par l'expert est soulignée. Ainsi la variable «couleur jaune» a été choisie par l'expert bien qu'elle ne soit que la deuxième variable la plus liée à la première composante.

La représentation choisie est celle des arbres de décision car cette représentation est simple et lisible.

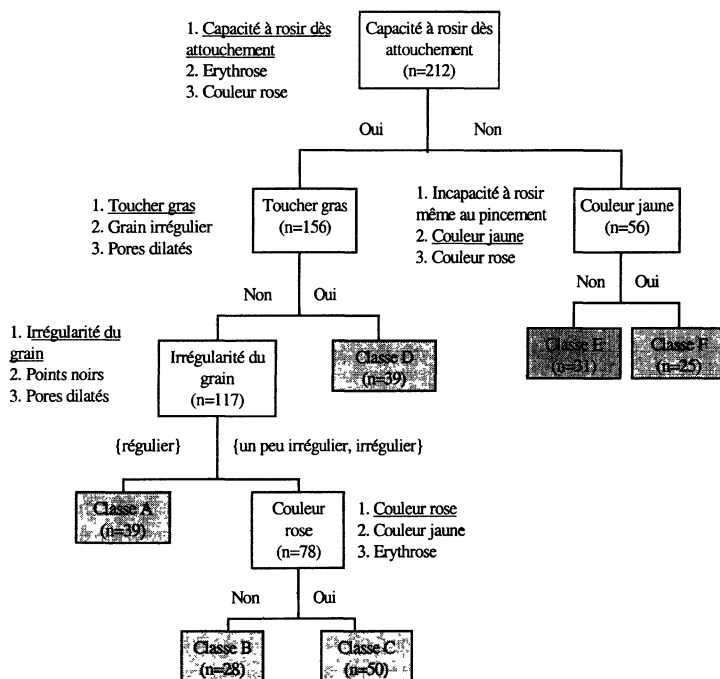


FIGURE 1

Arbre hiérarchique obtenu avec DIVAF.

Afin d'illustrer cet arbre de classification, le centre de gravité des 6 classes et les modalités des variables ont été représentés simultanément sur le plan factoriel de la figure 2. Par souci de lisibilité seuls les libellés des modalités les plus contributives sont indiqués. On distingue déjà deux familles de variables : les variables liées à l'état de la microcirculation de la peau, qui correspond, globalement au premier axe, et les variables liées à la production de sébum correspondant au deuxième axe. On note également que la première variable de coupure «capacité à rosir dès l'attouchement» résume la variation sur le premier axe factoriel tandis que la variable de coupure «toucher gras» résume la variation sur le deuxième axe factoriel.

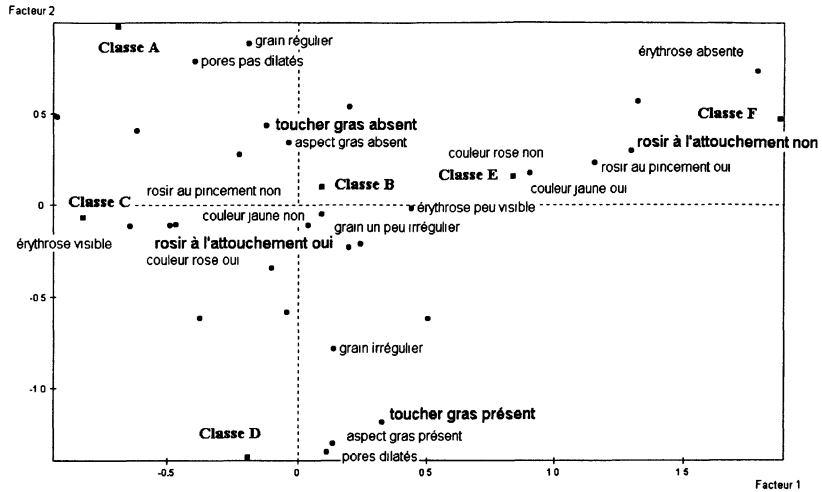


FIGURE 2

Visualisation simultanée des modalités et des centres de gravité des 6 classes (Plan 1-2 de l'analyse factorielle du tableau disjonctif complet initial croisant les 212 femmes avec les modalités de toutes les variables).

3. La méthode DIVOP

La méthode DIVOP (Chavent 1997, 1998), présentée ici dans le cas particulier des variables qualitatives, divise à chaque étape une classe en fonction d'une question binaire et du critère d'inertie. Dans le cas de variables qualitatives ordinales, la méthode utilise soit la distance euclidienne usuelle sur le tableau individus \times variables, où les modalités d'une variable sont codées par leur rang (1 pour la première modalité, 2 pour la seconde, etc.) soit la distance du khi-deux sur le tableau disjonctif complet. A chaque étape, la méthode définit la question binaire qui induit la bipartition d'inertie intra-classe minimum.

3.1 Présentation de la méthode

Rappelons la formule de Huygens qui relie l'inertie de l'union de deux classes $I(C_1 \cup C_2)$ à l'inertie intra-classe $W(C_1, C_2)$ et à l'inertie inter-classe $B(C_1, C_2)$:

$$I(C_1 \cup C_2) = W(C_1, C_2) + B(C_1, C_2)$$

L'inertie intra-classe W est égale à l'inertie de C_1 plus l'inertie de C_2 :

$$W(C_1, C_2) = I(C_1) + I(C_2)$$

On recherche tout d'abord une bipartition (C_1, C_2) de l'ensemble des individus de plus petite inertie intra-classe.

Pour cela, on évalue toutes les bipartitions induites par toutes les questions binaires. Si la variable X est ordinale, on évalue au maximum $m - 1$ bipartitions, m étant le nombre de modalités de X . Dans le cas d'une variable qualitative nominale, on se heurte à un problème de complexité, le nombre de dichotomies du domaine d'observation étant alors égal à $2^{m-1} - 1$. Dans la pratique, le temps de calcul devient important dès que m est supérieur à 10. On sélectionne la variable X et la dichotomie (S, \bar{S}) du domaine d'observation de X qui induit la bipartition de plus petite inertie intra-classe.

Dans un deuxième temps, il faut choisir la classe que l'on va diviser à l'étape suivante. On peut en effet choisir de diviser la classe C_1 en deux classes C_1^1 et C_1^2 . On obtient alors une partition en trois classes (C_1^1, C_1^2, C_2) . On peut également choisir de diviser la classe C_2 en deux classes C_2^1 et C_2^2 . On obtient alors la partition en trois classes (C_1, C_2^1, C_2^2) . La méthode DIVOP divise la classe C_i qui induit la partition en trois classes de plus petite inertie intra-classe. On montre que cela revient à choisir la classe C_i qui maximise :

$$\Delta(C_i) = I(C_i) - I(C_i^1) + I(C_i^2)$$

On répète ensuite les deux étapes précédentes jusqu'à ce que le nombre de classes fixé au départ soit atteint. Les divisions sont arrêtées après K itérations, ce nombre étant fixé au départ par l'utilisateur. En conséquence, les feuilles de l'arbre de décision sont les $K + 1$ classes de la dernière partition construite et les nœuds sont les questions binaires sélectionnées par la méthode.

L'arbre de décision est une hiérarchie sur les $K + 1$ classes. Cette hiérarchie indiquée par Δ ne possède pas d'inversion. Ainsi, une classe divisée avant une autre est représentée plus haut dans l'arbre hiérarchique et les partitions de 2 à $K + 1$ classes obtenues après chaque division sont bien les partitions de la hiérarchie indiquée.

Par exemple, l'arbre de décision de la figure 3.1 représente une partition en 5 classes (C_5, C_6, C_7, C_8, C_9). Dans cette représentation, on ne connaît pas l'ordre du découpage. On ne sait pas si C_2 a été découpé avant C_3 . En revanche, en associant à cet arbre de décision la hiérarchie indiquée par Δ de la figure 3.2, l'indétermination est levée.

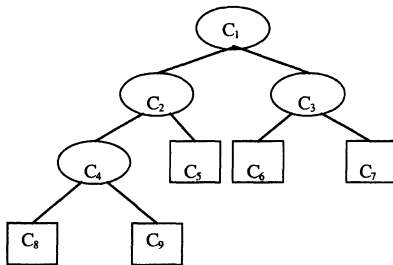


FIGURE 3.1
Arbre de décision
d'une partition en 5 classes.

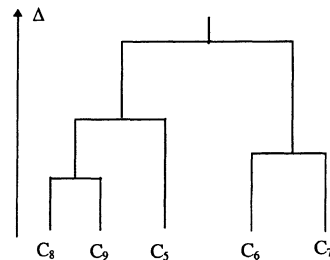


FIGURE 3.2
Hiérarchie indiquée
d'une partition en 5 classes.

3.2 Application de la méthode

On a appliqué la méthode DIVOP sur le tableau disjonctif complet des données sur la peau saine en utilisant dans un premier temps la distance du khi-deux. L'arbre de décision obtenu alors est proche de celui de la méthode DIVAF puisque les deux premières coupures sont identiques.

Dans un deuxième temps, la méthode DIVOP a été appliquée en utilisant la distance euclidienne usuelle sur le tableau individus \times variables, les modalités d'une variable étant codées par leur rang. L'arbre hiérarchique obtenu après 5 divisions est présenté figure 4. La dernière partition est une partition en 6 classes. Cet arbre donne les différentes questions binaires ayant induit les partitions de 2 à 6 classes. On lit ainsi que la première question binaire est :

{érythrose \in {absente, peu visible}} ou {érythrose \in {visible}}

Les 212 femmes sont donc divisées en 102 femmes pour lesquelles une «érythrose» est absente ou peu visible et 110 femmes pour lesquelles une «érythrose» est visible. Puis ces 110 femmes sont de nouveau divisées en fonction de la variable «régularité du grain» de la peau. A la différence de l'arbre de décision de la figure 1, cette représentation indique l'ordre de construction des partitions de 2 à 6 classes.

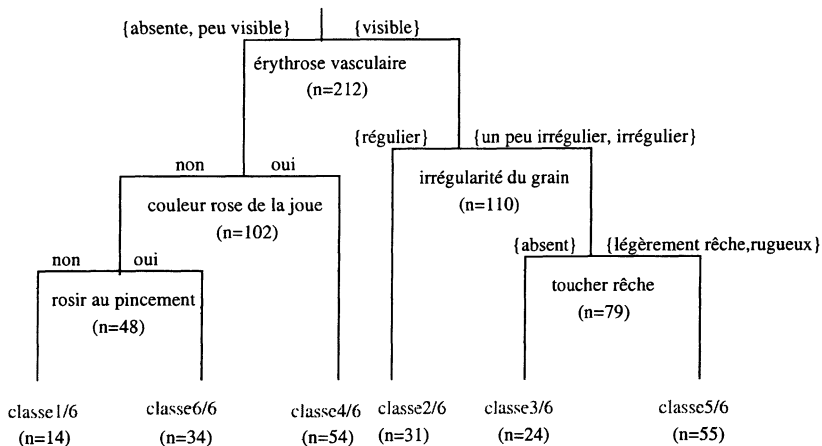


FIGURE 4
Arbre hiérarchique obtenu avec DIVOP.

Les figures 5, 6 et 7 permettent de commenter ces résultats sur le même plan factoriel 1-2 que celui de la figure 2. Sur la figure 5, on remarque que les modalités des variables «érythrose» et «régularité du grain», qui induisent les deux premières divisions, sont respectivement contributives au premier et au deuxième axe factoriel.

Sur le plan factoriel de la figure 6, on visualise la partition en trois classes et l'évolution des centres de gravité de la partition en 2 classes à la partition en 3 classes.

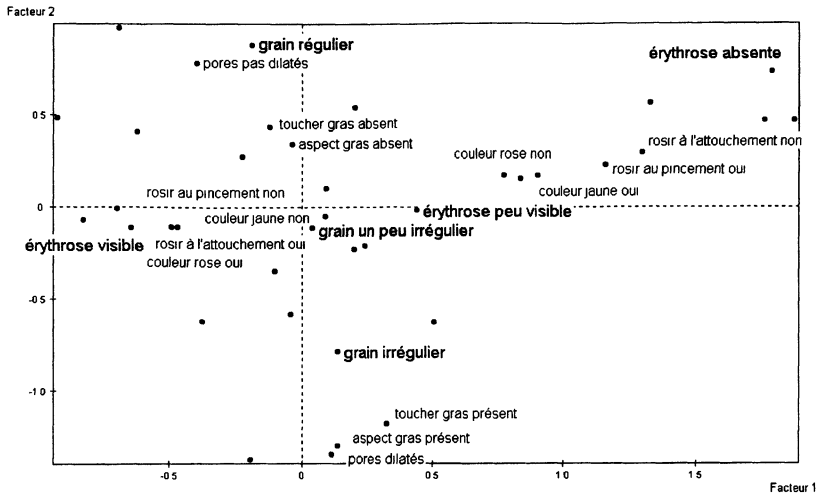


FIGURE 5
Modalités des deux premières divisions.

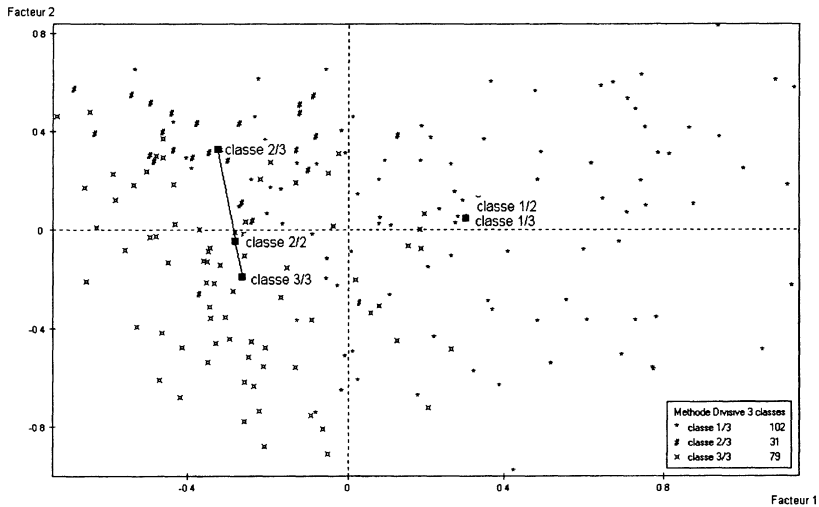


FIGURE 6
Visualisation de la partition en trois classes.

On lit classe i/j où i est le numéro de la classe de la partition en j classes. Par exemple la classe 1/2 est la classe 1 de la partition en 2 classes.

Dans un premier temps, on note sur cette figure que femmes de la classe 1/2 pour lesquelles une érythrope est absente ou peu visible sont à droite du second axe factoriel, tandis que les femmes de la classe 2/2, pour lesquelles une érythrope est

visible, sont à gauche du second axe factoriel. Dans un deuxième temps, on note que les femmes de la classe 2/2 sont divisées en deux classes 2/3 et 3/3, c'est à dire la deuxième et la troisième classe de la partition en trois classes. La classe 1/2, qui n'a pas été divisée, est aussi la classe 1 de la partition en trois classes c'est à dire la classe 1/3. Les femmes de la classe 2/3 dont le grain est régulier sont au-dessus du premier axe factoriel tandis que celles dont le grain est irrégulier sont au-dessous du premier axe factoriel.

La figure 7 permet de visualiser, à partir des centres de gravités des classes des partitions de 2 à 6 classes, la construction de l'arbre hiérarchique. Par exemple, la classe 3/4 est découpée lors de la quatrième division en deux classes 3/5 et 5/5.

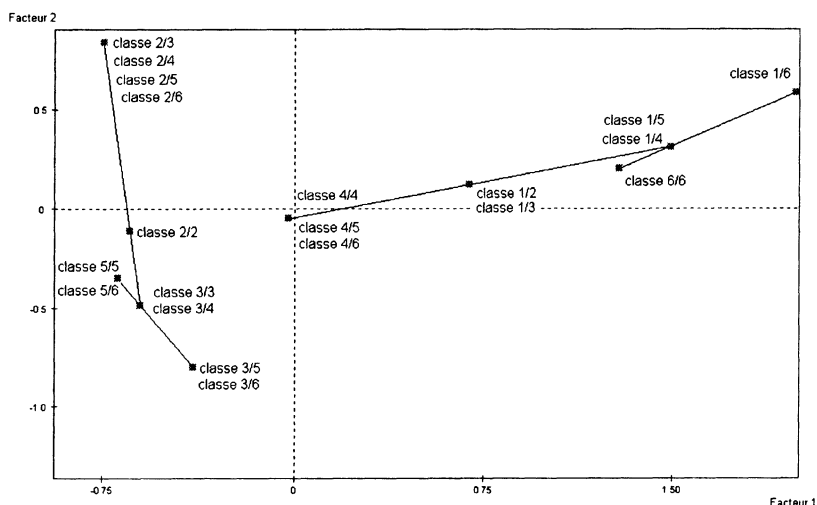


FIGURE 7

Visualisation des centres de gravité des classes des partitions de 2 à 6 classes.

4. Discussion

La discussion portera sur la comparaison des méthodes DIVAF et DIVOP au niveau méthodologique. La qualité d'optimisation de ces méthodes sera comparée à d'autres méthodes de classification automatique.

4.1 Comparaison des méthodes DIVAF et DIVOP

La méthode DIVAF adopte une approche locale, l'analyse factorielle des correspondances étant recalculée localement sur chaque classe. La méthode DIVOP adopte une approche plus globale puisque à chaque étape, elle cherche à obtenir non pas seulement une bonne bipartition mais également une bonne partition induite. C'est pourquoi elle choisit à chaque étape la classe qui va être divisée. La méthode DIVAF définit la question binaire qui va induire la bipartition en choisissant dans un premier

temps la variable et, dans un deuxième temps, la dichotomie des modalités de cette variable. La méthode DIVOP définit la question binaire en choisissant simultanément la variable et la dichotomie des modalités de cette variable qui induisent la bipartition de plus petite inertie intra-classe.

DIVOP se heurte à un problème de complexité lorsque les modalités ne sont pas ordonnées. La méthode DIVAF résout ce problème en opposant les modalités à score négatif aux modalités à score positif. En revanche, la dichotomie des modalités n'est pas définie comme dans la méthode DIVOP, en fonction d'un critère d'évaluation des bipartitions que l'on cherche à optimiser. La complémentarité des deux méthodes permet d'envisager une troisième approche où les modalités sont ordonnées en fonction de la première composante principale et la dichotomie de ces modalités choisie pour optimiser la bipartition induite.

Enfin, la méthode DIVOP est définie pour tous types de variables tandis que la méthode DIVAF n'est définie que pour des variables qualitatives.

4.2 Comparaison avec les autres méthodes de classification

Nous avons obtenu avec la méthode DIVOP deux typologies en 6 classes en utilisant respectivement la distance du khi-deux sur le tableau disjonctif complet et la distance euclidienne sur le tableau des variables ordinales (où chaque modalité est codée par son rang). Sur ces deux tableaux, et avec les mêmes métriques, nous avons également utilisé la méthode de classification ascendante hiérarchique avec le critère de Ward et la méthode des Nuées Dynamiques.

Le tableau 1 présente le pourcentage d'inertie expliquée en fonction de la méthode et de la distance

TABLEAU 1
Pourcentage d'inertie expliquée des partitions en six classes.

Distance	DIVOP	DIVAF	WARD	Nuées Dynamiques
Euclidienne	31,10	×	32,12	35,29
Khi-deux	30,15	26,67	29,10	31,44

Les pourcentages d'inertie expliquée, obtenus avec la méthode DIVOP sont, naturellement, légèrement moins bons qu'avec la méthode des nuées dynamiques. En revanche, ils sont tout à fait comparable aux résultats obtenus avec la méthode ascendante de Ward. Les 26,67 % d'inertie expliquée, obtenu avec la méthode DIVAF, s'expliquent par l'intervention experte, au niveau du choix de la seconde variable de coupure.

Cependant, les méthodes de segmentation non supervisée ont l'avantage d'avoir une interprétation facile des classes grâce à leur caractère monothétique. Elles rejoignent les méthodes de classification conceptuelles, connues en intelligence artificielle, en présentant l'avantage de se placer dans un cadre d'optimisation.

Par ailleurs, comme elles utilisent aussi la logique de la segmentation classique, ces méthodes peuvent s'intégrer facilement dans les outils de data-mining.

Enfin, les tests statistiques d'arrêt et l'intervention d'experts au niveau du choix de la variable sont deux points forts de ces méthodes.

Remerciements

Les auteurs remercient le professeur E. Tschachler pour ses encouragements et toute l'équipe du CE.R.I.E.S. pour leur contribution aux données et en particulier M. Dubourgeat, F. Morizot et I. Le Fur.

Bibliographie

- BREIMAN L., FRIEDMAN J.H., OLSHEN R.A., STONE C.J. (1984). Classification And Regression Trees, Chapman & Hall.
- CHAVENT M. (1997). Analyse des données symboliques; une méthode division de classification, Thèse, Université Paris IX-Dauphine.
- CHAVENT M. (1998). A monothetic clustering method, Pattern Recognition Letters 19, pp. 989-996.
- GOWDA K., KRISHNA G. (1978). Disaggregative clustering using the concept of mutual nearest neighborhood, IEE Transaction on Systems, Man and Cybernetics 8, pp. 888-895.
- EDWARDS A.W.F., CAVALLI-SFORZA L.L. (1965). A method for cluster analysis, Biometrics 21, pp. 362-375.
- GUENOCHÉ A. (1994). Spanning trees and average linkage clustering, New approaches in classification and data analysis, éd. Diday E. et Lechevallier Y., Springer Verlag, pp. 119-127.
- GUENOCHÉ A., HANSEN P., JAUMARD B. (1991). Efficient algorithms for divisive hierarchical clustering, Journal of classification 8, pp. 5-50.
- GUINOT C., TENENHAUS M., DUBOURGEAT M., LE FUR I., MORIZOT F., TSCHACHLER E. (1997). Recherche d'une classification de la peau humaine saine : méthode de classification et méthode de segmentation, XXVIII^{èmes} journées de statistiques, Carcassonne, pp. 429-432.
- HAR-EVEN M. et BRAILOVSKY V.L. (1995). Probabilistic validation approach for clustering, Pattern Recognition Letters, pp. 1189-1196.
- HUBERT L. (1973). Monotone invariant clustering procedure, *Psychometrika* vol. 38, n° 1, pp. 47-62.
- HUNTSMAN-LABED A. (1997). Algorithmes de construction hiérarchiques cherchant à optimiser le critère des moindres carrés, Thèse, Université Aix-Marseille III.
- MAC-NAUGHTON-SMITH P., WILLIAMS W.T., DALE M.B., MOCKETT L.G. (1964). Dissimilarity analysis : a new technique of hierarchical subdivision, *Nature* 202, pp. 1034-1035.

- MATULA D.W. (1986). Divisive versus agglomerative average linkage hierarchical clustering, *Classification as a tool of research*, éd. Gaul W. et Shader M., Elsevier, pp. 289-301.
- MICHALSKY R., DIDAY E., STEPP R. (1982). A recent advance in data analysis : clustering objects into classes characterized by conjunctive concepts, *Progress in Pattern Recognition*, éd. Kanal L.N. et Rosenfeld A., pp. 33-56.
- MICHALSKY R. et STEPP R. (1983). Learning from observations : conceptual clustering, *Machine Learning : an artificial intelligence approach*, éd. Michalsky R.S. et Corbonell J.G., Morgan Kaufman.
- QUINLAN J.R. (1986). Induction of decision trees, *Machine Learning* 1, pp. 86-106.
- REINERT M. (1983), Une méthode de classification descendante hiérarchique, *Cahiers de l'analyse des données* 3, pp. 187-198.
- ROUX M. (1985). *Algorithmes de classification*, Masson.
- ROUX M. (1995). About divisive methods in hierarchical clustering, *Data Science and its applications*, éd. Escoffier Y. et Hayashi C., Academic press, pp. 101-106.
- SONQUIST J.A. et MORGAN J.J. (1963). Problems in the analysis of survey data and proposal, *J.A.S.A.* 58, pp. 415-435.
- VOLLE M. (1976). *Analyse des données*, Economica.
- WILLIAMS W.T. et LAMBERT J.M. (1959), Multivariate methods in plant ecology, *Journal of Ecology* 47, pp. 83-101.