Statistique et analyse des données

SOCIÉTÉ FRANÇAISE DE CLASSIFICATION

Résumés - Journées de Statistique, Nice 22-26 mai 1978

Statistique et analyse des données, tome 3, n° 2 (1978), p. 31-44.

<http://www.numdam.org/item?id=SAD_1978__3_2_31_0>

SOCIETE FRANÇAISE
DE CLASSIFICATION

(Résumés - Journées de Statistique, Nice 22-26 MAI 1978)

SUR LES TYPES DE PARTITIONS

J.P. BARTHELEMY

E.N.S.C.M.B.
25030 - BESANÇON Cédex

L'étude des types de partitions (suite ordonnée des cardinaux des classes), c'est à dire des partages d'un entier peut être envisagée des points de vue suivants :

- Etude mathématique d'un "problème fondamental lié à la démarche du taxinomiste"
- Comparaison de classifications (menées concurremment) selon la taille et le nombre des classes indépendamment des éléments qui constituent ces classes.

C'est dans cette optique que nous nous proposons de présenter quelques résultats récents (essentiellement des propriétés métriques) sur les partages d'un entier.

L'ensemble P_n des partages de n est muni d'une relation d'ordre (non latticielle) pour laquelle il est modulaire. Cette remarque conduit à la construction de métriques qui s'interprètent en termes de graphes, certaines d'entre elles étant liées à la taxinomie et la théorie de l'information.

LA BI

U.E.R. DES
UNIVERSITE

Les concepts de h
observation lors de dépouillemen
les mettent en évidence les oppo
d'un axe. Alors que les classifi
qui se ressemblent.

Il est souhaitabl
lyse de données les classes d'ob
les classes. Le concept de bi-cl

Une bi-classe est
dire que deux objets (individus,
même classe d'une bi-classe se r
différentes d'une bi-classe s'op

Le principe de la
tableau de données les bi-classes
du type ressemblance et du type

Nous avons dévelo
rechercher une telle structure :

- une adaptation
recherche d'une bi-partition (ty

- une bi-classifi
pour former des bi-classes

- une méthode qui
fournir des familles de k bi-cla
entre un axe factoriel et une bi-

CONSTRUCTION D'UNE ULTRAMETRIQUE LA PLUS PROCHE AU SENS DES MOINDRES CARRES APPROXIMATION VS OPTIMISATION

J.L. CHANDON

UNIVERSITE D'AIX MARSEILLE 3

I.A.E.

29, av. Robert Schumann

13670 - AIX en PROVENCE

Soit un ensemble X d'objets à classer et soit un indice de proximité d défini pour l'ensemble IP des paires d'objets de X. d est une application symétrique de IP dans R_+ , indicatrice du degré de dissimilarité existant entre deux objets.

L'objectif de la classification est de simplifier l'information contenue dans le tableau des proximités en remplaçant les proximités entre paires d'objets par des proximités entre classes d'objets.

Nous nous intéressons à la construction d'une classification ascendante hiérachique (CAH) sur l'ensemble X. JOHNSON (1967) et BENZECRI (1973) ont démontré que pour obtenir une hiérarchie totale indicée sur X, il suffit de transformer l'indice de dissimilarité d en dissimilarité ultramétrique S.

L'écart entre deux dissimilarités d et S étant mesuré par la norme euclidienne $\| d - S \|$, le problème de la construction d'une ultramétrique optimale S^x minimisant $\| d - S \|$ a été résolu par CHANDON, LEMAIRE, POUGET (1978). Toutefois, l'algorithme de séparation ét évaluation progressive proposé est relativement lent. Il devient impraticable dès que le nombre d'objets est supérieur à 15.

Pour classer un grand nombre d'objets il est nécessaire de disposer d'un algorithme rapide permettant de construire ou d'approximer S^x. Deux algorithmes basés sur le concept de préordonnance sont proposés. Le premier, très rapide, ne garantit pas l'obtention de l'optimum S^x. Néammoins, il conduit toujour à une bonne approximation et il améliore toujours l'ultramétrique obtenue par l'algorithme de la moyenne de LANCE, WILLIAMS (1967), lorsque celle ci n'est pas optimale. Le second, moins rapide, améliore encore cette estimation.

Les deux algorithmes sont appliqués aux données classiques de RAO (1952), FITCH, MARGOLIASH (1967), KAMEN (1971) et MILLER, NICELY, afin d'illustrer l'amélioration du critère des moindres carrés obtenue par rapport à plusieurs autres méthodes de classification hiérarchique.

RECONNAISSANCE D
D'APPROXIMATION PO

C

UNIVE
LASAD
Place

Y.
IF
Dé

Lorsque nous désir
fournie par une fonction g(x) écha
$[a,b]$, nous cherchons une fonctio
approxime au mieux g(x), au sens
les plus utilisées sont les polyn
pentes infinies suivies de pentes
sont faibles, ce qui est nécessair
avons donc choisi une approximatio
Ces fonctions sont définies par di
dont les limites sont variables.
et qu'on ne possède de solution gé
puisque la clé du succès d'une app
ces points limites.

Les algorithmes que
classification automatique (algori
techniques de régression (norme L_2
mensionné au sens de Chebyshev (no

Ces algorithmes pré
ralisés aux cas de contours et per
de dérivabilité (fonctions splines
à la fonction approximante globale

UNE METHODOLOGIE DE CLASSIFICATION : L'ANALYSE PAR CRITERES ORDONNES SUR UN ENSEMBLE MUNI D'UNE RELATION DE PROXIMITE.

C. COCHET

IRIA, CEPIA
Domaine de Voluceau
BP 105
78150 LE CHESNAY

Nous proposons, dans le cadre de cette communication, une méthodologie d'approche, "l'analyse par critères ordonnés sur un ensemble muni d'une relation de proximité" (APCO), qui vise des problèmes se posant en des termes complexes à propos du morcellement de certains systèmes. Le déroulement d'une méthode qui permet de discerner des sous parties à l'intérieur d'un système peut prendre deux formes essentielles.

Dans un premier cas il est possible d'identifier un procédé de morcellement dont les résultats seront conformes aux spécifications auxquelles doivent répondre les sous-systèmes discernés. La méthode qu'il faut mettre en oeuvre se réduit alors au procédé qui est ainsi mis en évidence.

Dans tous les autres cas, il n'est pas possible de découvrir une procédure unique qui satisfasse, en fonction des spécificités du domaine étudié, les besoins de l'analyste. L'investigation scientifique auquel est soumis le système sera composé de plus d'un procédé de morcellement.

Nous avons choisi d'établir la synthèse des résultats de morcellement en exploitant les partitions obtenues pour chaque procédé utilisé. Ce moyen autorise l'usage d'une structure mathématique riche : le treillis géométrique des partitions.

Un ensemble de procédures méthodologiques doivent être utilisées pour mener à bonne fin l'analyse des problèmes qui sont soulevés ici, c'est à dire ceux justifiables de différentes analyses de morcellement.

Dans un premier temps, chaque analyse de classification doit être appliquée au champs d'observation du système qui la concerne. Pour développer cette opération, il faut disposer d'une théorie des systèles dont les concepts soient suffisamment évolués afin de rendre possible cette procédure.

Ensuite, nous avons établi une méthodologie qui permet d'envisager une synthèse des différents morcellements utilisés. Pour atteindre cet objectif, nous avons exprimé les résultats des procédés de morcellement sous forme de "proximités" affectant les points faisant l'objet de cette étude.

Tous les résultats obtenus, sous forme de partition des parties morcellées sont alors synthétisés dans une analyse unique qui est le réflet des différents points de vue présents dans l'élaboration de la stratégie de classification. On débouche alors, par l'exploitation des résultats centraux de cette analyse et des produits intermédiaires, sur des quantifications qui permettent de mesurer l'intervention des éléments de la stratégie et même, à la limite, de mettre en évidence certaines composantes qui sont le résumé d'autres, plus complexe.

ANALYSE ET CLAS
OBTENUES A PARTIR D

J.

CENTRE DE CALCUL I

UER Mathématiques

Il s'agit de la rec
géographique de stèles thessalienn
consiste en deux spirales symétriq
spirales définit un important crit
caractériser les ateliers et de su
sans, en relation avec des situati

Les questions qui n
en Archéologie sont les suivantes

. Peut-on distingue
rales à partir des documents photo

. Si oui, peut-on d
rales et représenter chaque classe

Dans une première p
fecter l'ensemble des spirales à t
"à centres, sommets d'un polygone

Une analyse plus él
rayon de courbure le long de la sp
des différentes modalités de cette
la forme $R = f(\theta)$, où (R, θ) sont
tions ont été retenues :

$R = a\theta + b$, R en es

Nous cherchons main
et à augmenter le nombre de types
permettre, d'après nos dernières
de spirales reconnues.

ANALYSE CANONIQUE DU POINT DE VUE DE LA CLASSIFICATION AUTOMATIQUE

E. DIDAY

UNIVERSITE PARIS IX-DAUPHINE
Place de Lattre de Tassigny
75775 - PARIS Cédex 16

Quand le tableau des données est de grande taille, il est légitime de chercher des combinaisons linéaires dépendant des tendances locales qui peuvent apparaître dans la population. Il s'agit de détecter ces tendances et simultanément les composantes canoniques qui leur sont le mieux associées.

Suivant que les données sont centrées ou non on propose plusieurs algorithmes qui tendent à minimiser le critère. Dans le cas où toutes les variables sont qualitatives, le problème se pose en terme d'analyse factorielle des correspondances et revient à minimiser le critère. Dans le cas où toutes les variables sont qualitatives, le problème se pose en terme d'analyse factorielle des correspondances et revient à chercher les classes d'objets qui induisent les plus grands X^2 de contingences entre les variables. Si l'un des deux paquets de variables est formé de variables d'incidence on aboutit à des méthodes intéressantes d'analyse discriminante locale.

LA NOTION DE DISPERS

UER
UNIV

Alors que générale
sont fondées sur la notion de di
à traiter, il est proposé ici un

A toute partie A d
persion D_A caractérisée par des
tivité : $D_{A\cup B} \geq D_A + D_B$ si $A \cap B = \emptyset$.

Différents cas son
des dispersions avec les types d
dispersion permet de construire
tatifs (notamment la pondération
sible de généraliser des algorit
mun.

Enfin est présenté
Méthode Non Hiérarchique Descend
de dispersion. Cette méthode rec
etc classes sans imposer de struc
elle résulte alors d'une structu
apportée par la méthode.

(+) les programmes de calcul auto
1974, par différents laboratoires

SELECTION ET DISCRETISATION OPTIMALES DE VARIABLES CONTINUES EN VUE D'UN PROBLEME DE RECONNAISSANCE DE FORMES

R. FAGES

UER de Mathématiques
UNIVERSITE DE LYON I

Lorsque les variables sont en partie ou totalité des variables continues, on se ramène à des variables discrètes (ou de classification) par le choix arbitraire de seuils afin d'utiliser les avantages de processus interrogatifs, comme par exemple les pseudo-questionnaires.

Il est abordé ici le problème du choix optimal pour chacune des variables continues, du nombre de classes et des seuils correspondants, afin de minimiser la probabilité d'erreur de l'identification par la règle de décision de BAYES.

La sélection des variables les plus discriminantes s'en déduit naturellement par le rejet des variables discrétisées par une seule classe. (Cette sélection est étendue aux variables initialement qualitatives).

La technique proposée utilise la majoration la plus fine de la probabilité d'erreur par une mesure d'entropie dérivant du coefficient de BHATTACHARRYA (+).

Un exemple concret est présenté, montrant l'efficacité des sélections et discrétisations obtenues, même dans le cas où les hypothèses nécessaires à la justification du critère d'optimisation ne sont pas vérifiées.

(+) M. TERRENOIRE, D. TOUNISSOUX. "Inequalities using BHATTACHARRYA distance and application to decision process".

3 rd International Joint Conference on Pattern Recognition, Coronado November 1976.

UN NOUVEAU TEST D'U
ALTER

L. FA

W. FE

Centre Nati
LABORA
POUR LE
31,chemin Jose

Si F(x) est la for
unimodale dont le support est inc
points u et v auxquels la différe
ximale respectivement se suivent
sation, pour tester l'uniformité
la statistique

$$S = \mathrm{Max}_{0<a<b<1} \left[a-\right.$$

où $F^{x}(.)$ désigne la fonction de r

On a déterminé, en
30 000 échantillons simulés indép
effectifs égaux à 10, 20 et 40. D
ternatives du type

$$F(z) = \frac{z^k}{a^{k-1}} \quad , \quad z<a$$

$$F(z) = 1 - \frac{(1-z)^k}{(1-a)^{k-}}$$

indiquent que S se compare favorab

DONNEES QUANTITATIVES INCOMPLETES ET CLASSIFICATION

P.P. FEVRE

IRIA - LABORIA
Domaine de Voluceau
78150 - LE CHESNAY

Lorsque des variables sont observées sur une population, il est fréquent que, pour certains individus, des variables ne soient pas relevées.

Pour traiter de telles données, la plupart des auteurs cherchent à "reconstituer" tout d'abord les données non disponibles, afin de pouvoir appliquer, sur ces données complétées, les méthodes usuelles d'analyse.

Le point de vue adopté ici est différent : nous cherchons à mettre en oeuvre directement les méthodes classiques de traitement des données, en ne tenant compte que des observations connues, et sans chercher à reconstituer les observations manquantes.

Pour ce faire, nous calculons, à partir des données disponibles, des approximations des quantités nécessaires au traitement habituel et nous travaillons en nous servant de ces approximations.

En application, afin de montrer ce que permet cette méthodologie, nous montrons comment une méthode de classification, la méthode de Nuées Dynamiques, peut être mise en oeuvre sur des données quantitatives incomplètes.

ANALYSE CLASSIFICATOIRE D'UN TEST SCOLAIRE

R. GRAS

Département de Mathématiques et I.R.E.M. de RENNES
Campus de Beaulieu
35042 RENNES CEDEX

Cette communication présente quelques résultats didactiques d'une analyse en classification hiérarchique d'un test mathématique proposé à près de 1 100 élèves de 13 à 15 ans (fin de 3ème de C.E.S.). La classification C1, obtenue à l'aide de l'algorithme de la vraisemblance du lien de I.C. Lerman, est comparée à 2 autres classifications :

- classification C2 par rapport au contenu et à la nature de la tâche de l'item
- classification C3 par rapport à une taxinomie d'objectifs cognitifs de R. Gras.

L'hypothèse d'indépendance entre C1 et C2, puis C2 et C3 est rejetée par un test du χ^2, au seuil de $1°/_{oo}$.

Les 5 classes de C1 conduisent à des interprétations confirmées par une analyse factorielle des correspondances :

- classe α de découverte de règle de production logique
- classe β de fonctions numériques et géométriques
- classe γ de nature numérique, très didactique
- classe δ d'observations de propriétés affines de l'espace
- classe ε d'observations de propriétés métriques de l'espace.

Les niveaux cognitifs croissent puis décroissent de α à ε, en passant par un maximum dans γ. Les classes $\{\delta, \varepsilon\}$ et $\{\alpha, \beta, \gamma\}$ sont relatives aux deux derniers stades de développement cognitif selon Piaget : "opérations concrètes" et "opérations logico-formelles".

PRESENTATION
DE DISCRETISATIO

J.

I.U.
Rue
KERCADO

Le codage de variab
données pose d'une façon générale
de variation de telles variables e
densité soit unimodale.

Dans le cadre des p
de mélanges de lois de probabilité
définition d'une fonction de densi
larges la séparation en composante

On rend finalement
médical ainsi que sur des mélanges

CLASSIFICATION VISUALISEE DE GRANDS ENSEMBLES SOUS DOUBLE CONTRAINTE

L. LEBART

C.N.R.S.
CREDOC - 140, rue du Chevaleret
75013 - PARIS

C. ROCHE

S.P.A.F.
Direction des Télécommunications

La procédure présentée répond aux préoccupations suivantes :

Construire une partition d'objet respectant une contrainte de contiguité géographique (un zonage) et dont les effectifs des classes soient bornés par une quantité imposée ; faciliter au maximum la compréhension du programme en multipliant les aides à l'interprétation, en introduisant de nouvelles procédures de visualisation, de façon à permettre une utilisation de routine aisée de la procédure ; assurer un encombrement mémoire réduit et une exécution rapide.

L'algorithme de base de construction d'une classification ascendante hiérarchique adaptée au sous-ensemble réactualisé des couples d'objets contigus n'est pas original (cf par exemple la thèse de A. THAURONT, PARIS, 1975).

La matrice de contiguité ne figure ici que sous la forme de tableau de codage réduit (pour chaque sommet du graphe : adresse des sommets adjacents). C'est sous cette forme qu'elle est actualisée après chaque agrégation, et après chaque intervention du seuil de taille maximale des classes. C'est également à partir de cette forme condensée que cette matrice est soumise à une analyse des correspondances impliquant une technique de diagonalisation particulière, de façon à faire apparaître sur l'imprimante une reconstitution de la carte géographique, sur laquelle seront positionnées les différents noeuds de l'arbre intermédiaire, puis les classes finales. Le principe de cette dernière opération permet d'analyser en quelques secondes des matrices binaires clairsemées d'ordre 1000x1000. Une analyse des correspondances classique effectuées cette fois sur le tableau de données de départ permet également de suivre les évolutions des noeuds et la position des classes non plus dans l'espace géographique, mais dans l'espace des variables. On a ainsi tous les éléments pour suivre et comprendre les mécanismes de formation des classes et le caractère plus ou moins prégnant des contraintes.

ARBRES VALUES ET ULTRAMETRIQUES

B. LECLERC

C. M. S.
54, bd Raspail
75270 - PARIS Cédex 06

Il est maintenant bien connu que tout arbre (graphe connexe et sans cycle) valué, défini sur un ensemble X de cardinal n, induit naturellement une ultramétrique r sur X, donc une classification hiérarchique sur X. Quelques travaux ont commencé à paraître, cherchant à étudier l'ensemble des ultramétriques ainsi définies à partir de tous les arbres valués donnés par un indice de distance d sur X. Inversement, on a aussi posé le problème suivant : soit r une ultramétrique sur X : représenter r par un arbre (Benzécri et Jambu, 1976).

Après avoir rappelé les résultats antérieurs, nous précisons le lien entre arbres valués et ultramétriques. On établit d'abord que celles-ci se caractérisent, parmi les indices de distance, par des propriétés ou interviennent uniquement leurs arbres minimaux. Mais, s'il est vrai qu'une ultramétrique r est parfaitement définie par l'un quelconque de ses arbres minimaux (valué par la restriction de r), le nombre N(r) de ceux-ci est compris, dans le cas général, entre $4^{n-1}/n^2$ et (n-1)!. Ceci pose le problème du choix d'un arbre particulier pour représenter r, qui n'a pas de réponse évidente que lorsque r a été obtenue à partir d'un arbre valué lisible directement dans les données (c'est le cas dans certaines méthodes classificatoires : lien simple et lien complet).

On s'intéresse ensuite aux propriétés du nombre N(r) qui paraît être un descripteur intéressant de la classification hiérarchique associée à l'ultramétrique r.

CLASSIFICATION DE GRANDS ENSEMBL

FACULTE UNIVERSITAIR
FACULTE DE DROIT ET

FACULTE UNIVERSITAIR
INSTITUT D'ECONOMIE

La méthode "single
est très utilisée en classificati
à classer sont grands, une métho
minimum d'un graphe, exige de O
où n désigne le nombre d'objets
à classer sont des distances de
sans que toutes les dissimilarit
rapide pour ce problème, ainsi q
semble de 10 000 étoiles.

LA METHODE DES POLES D'AGREGATION

H. LERFONDE

Département de Mathématiques
Centre Scientifique et Polytechnique
UNIVERSITE PARIS-NORD (XIII)
Av. J.B. Clément
93430 - VILLETANEUSE

Le principe général de cette méthode de classification non hiérarchique est de constituer pas à pas des agrégats parmi les éléments à classer.

Pour former un agrégat nous choisissons un élément "pertinent", appelé pôle, auquel nous agrégeons les éléments qui lui sont le plus proche, fonction d'un certain seuil. Cet agrégat constitué, nous recommençons la même opération avec les éléments non encore agrégés, et ceci tant qu'il reste des éléments à classer. A la fin nous obtenons une partition de l'ensemble des éléments, chaque classe de la partition étant l'un des agrégats (pour l'idée de base de cet algorithme, on pourra consulter I.C. LERMAN, reconnaissance et classification de structures finies en analyse des données, Université de RENNES I, 1977, rapport IRISA n° 70).

Nous avons mis en oeuvre cette méthode sous forme d'un algorithme appelé MPAGD (Méthode des Pôles d'Agrégation sur les Distances) où l'on étudie la distribution des distances entre éléments à classer.

La principale difficulté de ce type d'algorithmes est de parvenir à un système permettant d'arrêter la formation d'un agrégat que l'on puisse appliquer indépendamment de la distribution des distances et de la nature des données. Cet algorithme est une tentative dans cette voie.

LES PRESENTATIONS FA

I

Laboratoire de
UNIVERSITE DE R
35031 - RENNES

Différentes tentat
et la Classification ont été prop
M. GONDRAN (1975), M. JAMBU suiva
correspondent en fait, comme nous
"factorielles" du problème de la
tations est en effet grande compt
de l'Analyse Factorielle en Compo
critère de l'inertie expliquée.

Notre but dans cet
pour préciser sa véritable nature
nouvelles. Ce faisant, nous contr
des approches en les situant, dan
aux autres. Ce qui nous permettra
ces différentes tentatives.

Les différents typ
et la Classification se distingue
la recherche d'une classification
che d'un arbre binaire des Classi
la nature du critère optimisé : d
est basé sur l'inertie expliquée
proximité entre parties disjointe
mum). Ces différentes approches s
factorielle retenue : s'agit-il d
l'espace de représentation du nua
trouve défini comme une fonction
Ces approches se distinguent enfi
l'espace engendré par la solution

CONTRIBUTION DE LA THEORIE DES GRAPHES A LA TAXINOMIE MATHEMATIQUE

B. MONJARDET
UNIVERSITE DE PARIS V
Centre de Mathématique Sociale (E.H.E.S.S.)

Il est bien connu que deux méthodes classiques en classification hiérarchique, celle du "lien simple" (ou de l'ultramétrique sous-dominante) et celle du "lien complet" s'interprètent aisément en termes de théories des graphes. Les classes des partitions de l'arbre hiérarchique sont en effet, dans le premier cas, des classes connexes, dans le second cas, des cliques, des graphes "seuils" associés à la dissimilarité considérée. La considération de ces graphes seuils ajoutée à une condition de cohérence, ramène le problème de la classification hiérarchique à celui de la classification des sommets d'un graphe. A cet égard, les classements formés par les classes connexes d'une part, des cliques d'autre part, apparaissent comme deux solutions extrêmes maximisant respectivement un critère de séparation entre classes et un critère d'homogénéité à l'intérieur des classes. Mais en fait, la théorie des graphes permet d'envisager bien d'autres possibilités intermédiaires entre les deux solutions extrêmes. Ces possibilités sont d'ailleurs apparues dans des contextes variés. Par exemple, l'analyse de réseaux sociométriques et la recherche de leur décomposition en groupements homogènes amène à définir des concepts de "cliques généralisées" d'un graphe (Luce, 1950). Inversement, des études théoriques sur le nombre chromatique ou la connectivité d'un graphe conduisent à des notions pouvant être utiles en classification.

Le but de l'exposé est de présenter ces apports de la théorie des graphes à la taxinomie mathématique, qu'il s'agisse ou non de classification hiérarchique. La littérature sur le sujet étant déjà fort vaste (plus de 150 références), on ne tentera pas d'être exhaustif, mais on essaiera d'indiquer les principales contributions, en distinguant les apports conceptuels de ceux plus techniques.

CLASSIFICATION DE MAXIMUM DE VRAISEMBLANCE DE DONNEES BINAIRES

P. ROUSSEAU

Centre de Recherches Mathématiques
UNIVERSITE DE MONTREAL
Casier Postal 6128
MONTREAL (CANADA)

Une famille d'ensembles de données binaires est par hypothèse représentée par K modèles logistique-linéaires où chaque modèle représente une sous famille de cette famille.

On définit un algorithme qui trouve simultanément la partition de maximum de vraisemblance en K sous familles et les estimateurs de maximum de vraisemblance des paramètres de chaque modèle.

Cet algorithme est une extension des méthodes de partition itérative selon l'approche de Diday.

ETUDE DE L'AGR
PAR UNE METHODE

Laborat
U.E.R.
UNIVERS

On cherche à class
façon à dégager les principaux
la Vraisemblance du lieu dO à I
descriptifs ; chaque départemen
1 - surfaces par
2 - importance du
3 - structure d'e

On a défini un in
lonnes) d'un tableau de continge
I.C. LERMAN et respectant la mé
cas où les données sont une juxt
ayant, tous, le même ensemble de
l'application de l'A.V.L. pour c
tements français selon un seul c
deux ou trois caractères réunis,

Nous avons aussi a
de la Classification Ascendante
grégation par la maximisation, à
velle classe formée par réunion
gorithmes sont comparés dans les
a) classification
et
b) classification