Zeitschrift: Mémoires de la Société Vaudoise des Sciences Naturelles

Herausgeber: Société Vaudoise des Sciences Naturelles

Band: 18 (1987-1991)

Heft: 2

Artikel: Essai de syntaxonomie numérique appliqué aux pelouses à Laserpitium siler

Autor: Ziergler, Pascal

Kapitel: 3: L'analyse factorielle des correspondances

DOI: https://doi.org/10.5169/seals-259819

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 17.07.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

implicitement «autant» de poids à la dominance d'une espèce qu'à sa présence:

Indices de BRAUN-BLANQUET:	r	+	1	2	3	4	5
Codes:	4	4	5	6	7	8	9

Ce codage intermédiaire se rapproche beaucoup de l'indice de Moore 1966 (VAN DER MAAREL 1979).

Pour les tableaux de constance, le codage des classes a été établi comme suit:

Classes:	1-9%	10-19%	20-39%	40-59%	60-79%	80-100%
Codes:	1	2	3	5	7	9

Pour appliquer l'analyse factorielle aux facteurs écologiques, nous subdivisons le domaine de chaque paramètre en classes pour obtenir une matrice binaire; par exemple, l'altitude est divisée en 10 classes:

Classe	1	400 - 599 m
Classe	2	600 - 799 m
Classe	3	800 - 999 m
Classe	4	1'000 - 1'199 m
Classe	5	1'200 - 1'399 m
Classe	6	1'400 - 1'599 m
Classe	7	1'600 - 1'699 m
Classe	8	1'700 - 1'799 m
Classe	9	1'800 - 1'899 m
Classe	10	1'900 m et plus

Un relevé à l'altitude 700 m est codé 1 pour la classe 2 et 0 pour les autres. Les classes sont découpées de manière à contenir des effectifs voisins et leur étendue n'est pas forcément constante. Le tableau 1 fournit la liste des classes des paramètres écologiques utilisés. Cette technique de codage correspond au «tableau de description logique» de JAMBU (1978, p. 34 sq). Dans les projections AFC, les différentes classes d'un facteur se disposent sur une ligne brisée qui peut révéler un gradient écologique dans le nuage des relevés (voir par ex. fig. 11). Cette technique est utilisée par DUPOUEY (1983) en taxonomie des espèces, et a déjà été proposée par ROMANE (1972) en phytosociologie.

3. L'ANALYSE FACTORIELLE DES CORRESPONDANCES

3.1. *Généralités*

Le détail de la technique est donné par BENZÉCRI *et al.* (1980 a,b), LACOSTE et ROUX (1971), LACOSTE (1975), et LEBART *et al.* (1982). Pratiquement, c'est le programme CORRES (EPFL, selon BENZÉCRI *et al.* 1980) qui a été utilisé.

Remarques préliminaires:

–le fait qu'un ensemble d'individus soient proches sur les premiers axes

AFC usuels est une condition nécessaire pour pouvoir en faire un groupe taxonomique, mais non suffisante. On peut toujours imaginer un axe ultérieur qui fasse éclater ce groupe; cependant, ce nouvel axe correspondra à un facteur écologique certes discriminant, mais pas assez influent dans l'ensemble des relevés considérés pour provoquer une disjonction de groupes dans les premiers axes; retenons donc qu'une proximité en AFC est toujours relative au référentiel particulier de l'analyse.

–l'AFC déforme l'«espace des données» en choisissant des axes de projection préférentiels: des relevés situés en bordure du nuage sont souvent moins excentriques en réalité que ne le suggère leur projection (MONOD et ZIEGLER 1983, p. 106). Ce point sera discuté au chapitre 5.1.1.

code des classes	altitude (m)	exposition (°)	pente (°)	profondeur du sol	argile (%)	sable (%)
1	400-599	N-NE 1-45	0-15	très mince (roche)	0-5	0-5
2	600-799	NE-E 46-90	16-25	mince (-> 5-6 cm)	5.1-10	5.1-10
3	800-999	NW-N 316-0	26-35	moyen (-> 20 cm)	10.1-15	10.1-15
4	1000-1199	E-SE 91-135	36-45	profond (> 20 cm)	15.1-20	15.1-20
5	1200-1399	W-NW 271-315	46-55		20.1-25	20.1-25
6	1400-1599	SE-S 136-180	56-70		25.1-30	25.1-30
7	1600-1699	SW-W 226-270	> 70		30.1-40	30.1-40
8	1700-1799	S-SW 181-225			40.1-50	40.1-50
9	1800-1899	S-SW 181-225			50.1-60	50.1-60
10	> 1900				> 60	< 60

code des classes	humidité à la cap. rétension	humidité au pt flétriss. (%)	carbone (%) organique	azote (‰)	C/N	pH
1	0-30	0-20	0-3	0-3	0-6	< 4.50
2	31-40	21-30	3.1-6	3.1-6	6.1-10	4.51-5.50
3	41-50	31-40	6.1-8	6.1-8	10.1-14	5.51-6.00
4	51-60	41-50	8.1-10	8.1-10	14.1-20	6.01-6.50
5	61-70	51-60	10.1-12	10.1-12	> 20	6.51-6.75
6	71-80	61-70	12.1-15	12.1-15		7.76-7.00
7	81-90	71-80	15.1-20	15.1-20		7.00-7.25
8	91-100	81-90	20.1-25	20.1-25		7.26-7.50
9	101-120	91-100	25.1-30	25.1-30		7.51-8.00
10	> 120	> 100	> 30	> 30		> 8.00

code des classes	calcaire (%)	cations échangeables au pH du sol : Ca, Mg et Na, K				
		calcium (ppm)	magnésium (ppm)	potassium (ppm)	sodium (ppm)	acidité d'échange (ml NaOH 0.01N)
1	0	0-400	0-10	0-10	0-3	0-0.50
2	traces	401-600	10.1-20	10.1-30	3.1-5	0.51-1.00
3	< 2	601-800	20.1-30	30.1-50	5.1-7	1.01-1.50
4	2.1-20	801-1000	30.1-40	50.1-70	7.1-9	1.51-2.00
5	> 20	> 1000	40.1-50	> 70	> 9	2.01-2.50
6						> 2.50

Tableau 1. Codage en classes des paramètres choisis

3.2. *Fiabilité des premiers axes de projection*

Pour éprouver la fiabilité de l'AFC (3.2.2) et pour l'utiliser comme préparation d'un clustering (4.2), nous calculons des distances entre relevés dans l'AFC à partir des coordonnées sur les premiers axes.

3.2.1. *La pondération des coordonnées*

Pour tenir compte de la proportion d'information fournie par les divers axes, nous adoptons une pondération des coordonnées: le pourcent d'information que chaque axe factoriel exprime dans l'analyse est indiqué dans les statistiques de l'AFC («pourcentage d'inertie»; JAMBU 1978, p. 264); les coordonnées sur chaque axe sont multipliées par le rapport entre le pourcent d'inertie de cet axe et celui du premier axe. Cette technique a également été utilisée par BLOOM (1980).

Exemple AFC des relevés de Jorette (fig. 14):

Axe factoriel	Pourcent d'inertie %	Cumul %
1	18,39	18,39
2	9,56	27,95
3	9,00	36,95
4	5,47	42,42

Dans ce cas, les coordonnées factorielles sur l'axe 2 seront toutes multipliées par 9.56/18.39, les coordonnées sur l'axe 3 par 9.00/18.39, etc.

3.2.2. *Test statistique de fiabilité par le coefficient cophénétique*

SNEATH et SOKAL (1973, p. 278-279) proposent un coefficient cophénétique comme indice de la fiabilité d'une technique de taxonomie numérique. Il s'agit d'un coefficient de corrélation entre deux matrices de distances séparant les objets à classer: celle des distances brutes et celle des distances produites par la technique de taxonomie considérée. Plus le coefficient est proche de 1, plus la technique taxonomique est fidèle aux degrés de ressemblance observables.

En général, les deux matrices ne présentent pas une distribution normale, nous devons donc utiliser le coefficient de rang de Spearman (DIEM et LENTNER 1972, p. 184) comme coefficient cophénétique; il n'existe malheureusement pas de test de signification pour la différence entre deux coefficients de Spearman. Dans le cas présent, ce sont les 31 relevés (465 distances inter-relevés) du transect Jorette qui ont servi au test (tabl. 2).

3.2.3. *Conclusions et discussion*

Ces coefficients cophénétiques (tabl. 2) sont très élevés. Ils sont supérieurs à ceux de toutes les techniques de clustering présentées ultérieurement (comparer tabl. 2, 3 et 4). L'AFC produit donc une image taxonomique très fidèle à la réalité, même en ne considérant que ses trois premiers axes de

Axes AFC considérés	Pondération des coordonnées	Coefficient cophénétique
1 à 3	non	0.8419
1 à 7	oui	0.9036
1 à 30	oui	0.9047

Tableau 2. Coefficients cophénétiques des premiers axes de l'AFC
Matériel: 31 relevés du transect Jorette (chap. 7), donc 465 distances inter-relevés.
AFC par le programme CORRES. Calcul du coefficient cophénétique selon chapitre 3.2.2.

projection, comme on le fait couramment en phytosociologie.

Cette conclusion semble devoir être nuancée suivant le volume de données traitées, bien que nous ne l'ayons pas testé par le coefficient cophénétique. Lorsque l'on augmente le nombre de relevés soumis à l'analyse, la quantité d'information perçue par les premiers axes diminue. Alors que pour une trentaine de relevés les premiers axes reflètent une forte proportion de l'inertie (42% pour les 4 premiers axes de l'AFC Jorette, voir exemple sous 3.2.1), pour plus de cent relevés, cette proportion devient très réduite (par exemple 12,5 % de l'inertie pour les trois premiers axes de l'AFC, fig. 18 avec 131 unités).

Sa grande fidélité aux ressemblances floristiques, et le fait qu'elle ne propose pas de classification, font de l'AFC l'outil numérique primordial d'une syntaxonomie objective, devant les techniques de clustering.

4. La classification hiérarchique automatique (clustering)

4.1. *Introduction: les niveaux de choix*

Le détail des techniques est exposé par Sneath et Sokal (1973), Hartigan (1975), Lebart *et al.* (1982), Jambu (1978), Benzécri *et al.* (1980). Pratiquement, ce sont les programmes CLUSTAN (Wishart 1975) et CLTR (Wildi et Orloci 1983) qui ont été utilisés.

Remarques préliminaires:

–Un dendrogramme peut toujours être calculé, quelle que soit la structure taxonomique des données fournies: des classes sont mises en évidence de toute façon (Groenewoud 1983), même si par exemple une ordination selon un gradient était un bien meilleur reflet de la réalité.

–L'ordre des individus montré par un dendrogramme n'a aucun sens, car une rotation est possible autour de chaque liaison verticale (Jambu 1978, p. 71).

Plus que l'AFC, le clustering implique de nombreux choix à trois niveaux:

–préparation des données: normées ou non, soumises ou non au prétraitement par AFC (4.2);

–choix d'une mesure de ressemblance entre relevés (4.3);

–choix d'un algorithme de classification (4.4).