Zeitschrift: Veröffentlichungen des Geobotanischen Institutes der Eidg. Tech. Hochschule, Stiftung Rübel, in Zürich

Herausgeber: Geobotanisches Institut, Stiftung Rübel (Zürich)

Band: 90 (1986)

Artikel: Analyse vegetationskundlicher Daten : Theorie und Einsatz statistischer Methoden = Analysis of vegetation data : theory and application of numerical methods

Autor: Wildi, Otto

Kapitel: 6: Gewichtungsverfahren (Rangierung)

DOI: https://doi.org/10.5169/seals-308832

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: 11.07.2025

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

6. Gewichtungsverfahren (Rangierung)

Der Gewichtung von Merkmalen kommt in der Vegetationskunde besondere Bedeutung zu. Vegetationsaufnahmen besitzen die Besonderheit, dass sie durch sehr viele Merkmale (Arten) charakterisiert sind. Ist die Zahl der Merkmale gar grösser als diejenige der Aufnahmen, so kann ein Datensatz überbestimmt sein. In einem solchen Falle lässt sich jede beliebige Gruppen- oder Gradientenstruktur anhand ausgewählter Merkmale begründen. Auf die traditionelle Tabellenarbeit übertragen (ELLENBERG 1956) bedeutet dies, dass sich dieselbe Vegetationstabelle in verschiedenster Weise durchaus sinnvoll ordnen lässt. Es ist daher schon aus formalen Ueberlegungen heraus erstrebenswert, Aufnahmen mithilfe möglichst weniger Arten oder Standortsfaktoren charakterisieren zu können. Dies erfordert aber meistens eine Gewichtung der Merkmale.

Eine solche kann vorgenommen werden, um Zeigerarten zu finden. Letztere dienen dann einerseits der Klassifikation der Aufnahmen. Andererseits erlauben sie die Zuordnung neuer Aufnahmen zu beschriebenen Pflanzengesellschaften (Identifikation). Schliesslich dienen sie dem Aufbau von Vegetationsschlüsseln zur pflanzensoziologischen Kartierung im Felde.

ORLOCI (1978) schlägt eine in der Praxis bislang kaum realisierte Anwendung der Gewichtungsverfahren vor. Seine Ueberlegungen zielen auf eine Minimalisierung der Feldarbeit bei vegetationskundlichen Kartierungen ab. Traditionellerweise werden in Vegetationsaufnahmen aus verschiedenen Gründen alle Arten von Gefässpflanzen berücksichtigt (POORE 1962, MUELLER-DOMBOIS und ELLENBERG 1974), was aus der Sicht der Grundlagenforschung auch sinnvoll ist. Namentlich bei den heute aktuellen Anwendungen der Vegetationskunde im Bereich der Umweltforschung ist dieses Vorgehen aber nicht zwingend. Der Aufwand für die Feldarbeiten fällt geringer aus, wenn nur Zeigerarten erhoben werden müssen. Diese sind aber bei echten Neuuntersuchungen nicht hinreichend bekannt. Mit dem folgenden Vorgehen könnte das Problem im Rahmen einer Vorun-

tersuchung gelöst werden:

1. Man macht einige Vegetationsaufnahmen im zukünftigen Untersuchungsgebiet (z.B. 20).

2. Durch ein Rangierungs- oder Gewichtungsverfahren werden die besten Zeigerarten innerhalb des Datensatzes bestimmt.

3. Es werden weitere Aufnahmen zugefügt (z.B. 10). Eine erneute Rangierung ergibt wahrscheinlich einige Aenderungen in der Gewichtung der Arten.

4. Die Erweiterung des Datensatzes ist so lange durchzuführen, bis sich in der Gewichtung kaum noch Verschiebungen ergeben. Die Arten hohen Gewichtes (Zeigerarten) beschreiben die Aehnlichkeitsstruktur der gesamten zu untersuchenden Vegetation hinreichend.

Wird nach Abschluss dieser Vorabklärung eine Kartierung in Angriff genommen, so brauchen keine vollständigen Aufnahmen mehr durchgeführt zu werden. Es kann unmittelbar mit der flächendeckenden Erhebung auf Grund der gefundenen Zeigerarten begonnen werden. Selbstverständlich kann dieses Vorgehen nur dann zum Erfolg führen, wenn eine zweckmässige Auswahl der Aufnahmeflächen getroffen wurde, sodass die Erhebung für das Untersuchungsobjekt repräsentativ ist.

Man kann grundsätzlich zwei Typen von Gewichtungsverfahren unterscheiden. Im einen Fall wird die Gesamtstichprobe untersucht, ohne dass ihre Gruppenstruktur bereits bekannt wäre. Dieser Weg führt in der Regel über die Analyse einer Aehnlichkeitsmatrix. Anders liegen die Verhältnisse, wenn von bekannten Pflanzengesellschaften ausgegangen wird. Hier stellt sich die Frage, welche Arten die besten Zeiger für eine vorgegebene Gruppenstruktur sind. Wir werden im folgenden Vertreter beider Typen diskutieren.

6.1 Feolis Methode

Diese einfache und leicht verständliche Methode soll einführend dargestellt werden (vgl. FEOLI 1973). Grundsätzlich basiert sie auf Präsenz und Absenz der Arten. Sie kann auf folgende zwei Annahmen zurückgeführt werden:

1. Eine Art ist dann ein guter Zeiger, wenn die Aufnahmen, in denen sie auftritt, möglichst ähnlich sind.
2. Aufnahmen, in denen sie vorkommt, sollen im Vergleich zu solchen, in denen sie fehlt, möglichst verschieden sein.

Das Prinzip der Methode zeigt Tab. 6.1, A bis C. In Tab. 6.1, A, sind die Rohdaten des nachfolgenden Beispiels zu finden. Zunächst muss eine Aehnlichkeitsmatrix S der Aufnahmen verfügbar gemacht werden. In Tab. 6.1, B und C ist eine solche dargestellt. Verwendet wird van der Maarels Koeffizient, doch ist die Methode auch mit jedem anderen Mass zu realisieren. Feoli bezeichnet nun die Aehnlichkeitskoeffizienten zwischen Aufnahmen, in denen die zu gewichtende Art vorkommt, mit SI. SZ werden jene genannt, welche Aufnahmen mit und solche ohne die betreffende Art vergleichen. Für die Berechnung des Gewichts der Art, SF, sind die Mittelwerte massgebend:

$$SF = 1 - \frac{\bar{S}Z}{\bar{S}I}$$

Tab. 6.1, B illustriert das Vorgehen für die Art 1. Diese kommt nur in den Aufnahmen 1 und 2 vor, die eine Aehnlichkeit von 0.8 aufweisen. Für SZ berücksichtigt man alle Koeffizienten, welche die Aufnahmen 1 und 2 mit 3, 4 und 5 vergleichen. Diese sind punktiert eingerahmt. Wir erhalten

$$SF_1 = 1 - \frac{(0.143+0.25)/6}{0.8} = 1 - \frac{0.066}{0.8} = 0.9175$$

Da Art 2 genau das gleiche Vorkommen hat, gilt dieses

Tabelle 6.1

Feolis Artrangierung. Vegetationstabelle (A), Aehnlichkeitsmatrix der Aufnahmen (B,C). Fall B illustriert die Berechnung des Gewichtes für Art 1 oder 2, Fall C für die Art 3.

A

Art	Aufn. 1	2	3	4	5
1	2	1			
2	1	1			
3	1	1	1		
4			1	1	1
5				1	2

B

Aufnahme	1	2	3	4	5
1	1	0.8	0.143	0	0
2		1	0.25	0	0
3			1	0.33	0.17
4				1	0.75
5					1

C

Aufnahme	1	2	3	4	5
1	1	0.8	0.143	0	0
2		1	0.25	0	0
3			1	0.33	0.17
4				1	0.75
5					1

Gewicht auch für sie. Art 3 hingegen kommt in den Aufnahmen 1, 2 und 3 vor (Tab. 6.1, C). Wir erhalten

$$SF_3 = 1 - \frac{(0.33+0.17)/6}{(0.8+0.143+0.25)/3} = 1 - \frac{0.083}{0.397} = 0.79$$

Ihre Zeigereigenschaft ist somit deutlich schlechter als jene der Arten 1 und 2.

Analog werden nun die Gewichte der Arten 4 und 5 berechnet. Schliesslich erhält man folgende Rangordnung:

Art	SF-Wert	Rang
1	0.9175	1
2	0.9175	1
5	0.889	2
4	0.843	3
3	0.79	4

Anhand dieser Ergebnisse lässt sich leicht erkennen, welche Arten beispielsweise für die Zusammenstellung eines Vegetationsschlüssels gut geeignet sind (z.B. 1, 2 und 5). Zur Datenreduktion eignet sich jedoch das Verfahren weniger. Aus den Arten 1 und 2 wird deutlich, weshalb dem so ist. Beide besitzen hohe SF-Werte. Wird auf eine von beiden verzichtet, so bleibt die Differenzierung zwischen Aufnahmen (1,2) und (3,4,5) dank der verbleibenden Art erhalten. Trotz hohen SF-Wertes würde somit der Verzicht auf Art 2 die Gruppenstruktur der Tabelle nicht zerstören. Will man den Informationsverlust durch Weglassen von Arten untersuchen, so bedarf es komplizierterer Verfahren.

6.2 Rangierung nach erklärter Varianz

Dieses von ORLOCI (1973) unter dem Namen RANK vorgeschlagene Verfahren ist dann angezeigt, wenn ein äusserstes Minimum an Zeigerarten gesucht wird. Die Analyse sehr grosser Tabellen

gestaltet sich einfacher, falls es gelingt, die Anzahl Arten auf einen Bruchteil zu reduzieren. Natürlich darf dabei die ursprüngliche Aehnlichkeitsstruktur der Aufnahmen nicht allzu sehr verändert werden. Eine erfolgreiche Anwendung dieses Verfahrens zeigen FEOLI und LAGONEGRO (1982). Sie revidierten die Klassifikation italienischer Buchenwälder anhand von 602 Aufnahmen. Die ursprünglich 475 Arten wurden mit Hilfe des RANK-Verfahrens vorgängig auf 108 reduziert, was die Analyse entscheidend vereinfachte.

Im Gegensatz zu Feolis Verfahren geht Orlocis RANK schrittweise vor. Das Prinzip ist folgendes (Tab. 6.2): Zu Beginn jedes Schrittes wird eine provisorische Gewichtung der Arten gesucht. Als Gewichtungskriterium sind verschiedene Masse denkbar. Würde z.B. die Summe der Abundanzwerte verwendet, so erhielte Art 1 Rang 1 (dass das RANK-Verfahren ein anderes Kriterium verwendet, ist aus Tab. 6.2 zu schliessen). Ausgehend von diesem Zwischenresultat ist nun der Datensatz zu reduzieren. Dazu muss die gesamte Information, welche die Art mit Rang 1 trägt, aus der Tabelle entfernt werden. Die Information der Art 2 müsste dabei fast verschwinden, da sie mit derjenigen von Art 1 annähernd identisch ist. Der nächste Schritt besteht in einer erneuten provisorischen Gewichtung der verbleibenden Arten. Art 4 dürfte dabei den höchsten Wert erhalten und mit dem definitiven Rang 2 ausgeschieden werden. Doch wenden wir uns nun dem eigentlichen RANK-Algorithmus zu.

Das Verfahren beginnt mit der Berechnung einer Aehnlichkeitsmatrix der Arten. Dabei kann ein beliebiges Mass ausgewählt werden. Verwendet man den Korrelationskoeffizienten, so erhält man für die p = 4 Arten in Tab. 6.2

$$R = \begin{pmatrix} 1.0 & 0.85 & -0.91 & -0.64 \\ 0.85 & 1.0 & -0.71 & -0.43 \\ -0.91 & -0.71 & 1.0 & 0.91 \\ -0.64 & -0.43 & 0.91 & 1.0 \end{pmatrix}$$

Wie erwartet, heben sich als mögliche Gruppen (1,2) und (3,4) mit hohen Korrelationen von 0.85 und 0.91 ab. Nun wird

Tabelle 6.2

Einfache Vegetationstabelle zur Rangierung der Arten. Die Ränge ergeben sich aufgrund Orlocis RANK.

Aufnahme	1	2	3	4	Rang
Art 1	2	2	1		4.
Art 2	2	1	1		2.
Art 3			1	1	1.
Art 4			2	1	3.

eine provisorische Rangordnung gesucht. Massgebend ist die Summe der Korrelationen jeder Art mit allen andern, also

$$SS_i = \sum_{h=1}^{p} \Sigma r_{hi}^2 / r_{ii}$$

Darin ist SS_i die Summe der total p quadrierten Aehnlichkeitswerte, welche die Art i betreffen. r_{ii} ist der Diagonalwert, der im Falle der Korrelationsmatrix 1.0 beträgt und damit vernachlässigt werden kann. Für Art 1 finden sich alle Aehnlichkeitswerte r_{hi} in der ersten Zeile (oder in der ersten Kolonne) von R. Somit ergibt sich für sie

$$SS_1 = ((1.0)^2+(0.85)^2+(-0.91)^2+(-0.64^2))/1 = 2.95$$

Entsprechend erhält man für die übrigen Arten

$$SS_2 = 2.41$$
$$SS_3 = 3.14$$
$$SS_4 = 2.41$$

Die Art mit der höchsten Summe, Art 3, erhält den definitiven Rang 1. Bevor die Bedeutung dieses Ergebnisses genauer untersucht wird, soll aber der Algorithmus weiter verfolgt werden. Zu diesem Zwecke ist der Einfluss der Art 3 auf die Korrelationen zu eliminieren. Die Elemente der neuen, reduzierten Matrix R berechnen sich nach der Regel

$$r'_{hi} = r_{hi} - y_{hm}\, y_{im} .$$

Symbol m steht für die zu entfernende Art, hier also Art 3. Die Formel gilt für die Korrelation jeder Art h und i. Bei den y-Werten handelt es sich um Elemente der Korrelationsmatrix, denn sie sind folgendermassen definiert:

$$y_{hm} = r_{hm}/(r_{mm})^{1/2} \text{ , sowie } y_{im} = r_{im}/(r_{mm})^{1/2} .$$

r_{mm} ist in unserem Falle die Korrelation von Art 3 mit sich

selbst, also 1.0. Mit dem ersten Element beginnend, ergibt die Reduktion:

$$
\begin{aligned}
r'_{11} &= 1 - (-0.91)^2 = 0.17 \quad (h=1,i=1)\\
r'_{12} &= 0.85 - (-0.71*-0.91) = 0.20 \quad (h=1,i=2)\\
r'_{13} &= -0.91 - (1.0*-0.91) = 0 \quad (h=1,i=3)\\
r'_{14} &= -0.64 - (-0.91*+0.91) = 0.19 \quad (h=1,i=4)
\end{aligned}
$$

Diese Operation ist für alle r_{hi}-Werte vorzunehmen. Damit verschwindet jener Anteil der Korrelation aus der Matrix, der durch Art 3 verursacht wird. Zeile und Spalte 3 von R müssen nun zwangsläufig mit Nullen besetzt sein:

$$
R = \begin{matrix}
0.17 & 0.20 & 0 & 0.19\\
0.20 & 0.50 & 0 & 0.21\\
0 & 0 & 0 & 0\\
0.19 & 0.21 & 0 & 0.17
\end{matrix}
$$

Die erneute provisorische Rangordnung ergibt:

$$
\begin{aligned}
SS_1 &= (0.17^2+0.20^2+0.19^2)/0.17 = 0.61\\
SS_2 &= (0.20^2+0.50^2+0.21^2)/0.50 = 0.67\\
SS_4 &= (0.19^2+0.21^2+0.17^2)/0.17 = 0.64
\end{aligned}
$$

Art 2 erreicht nun den maximalen Wert und erhält den endgültigen Rang 2. Die Aehnlichkeitsmatrix kann damit weiter reduziert werden, wobei Zeile und Kolonne 2 und 3 mit Nullen besetzt sind:

$$
R' = \begin{matrix}
0.14 & 0 & 0 & 0.16\\
0 & 0 & 0 & 0\\
0 & 0 & 0 & 0\\
0.16 & 0 & 0 & 0.14
\end{matrix}
$$

Den Rangierungsprozess mit dieser Präzision weiterzuführen, lohnt sich nicht, da nunmehr immer grössere Rundungsfehler auftreten! Mit den Arten 3 und 2 sind aber offensichtlich zwei typische Vertreter der Tabelle 6.2 gefunden worden. Die Korrelationssummen SS bilden ausserdem ein Mass für deren Güte als Zeigerwert. Dieser ist jedoch leichter zu interpre-

tieren, wenn er am maximal möglichen Wert gemessen wird, wie er bei einer absolut idealen Zeigerart auftreten würde. Solche Maximalkorrelationen finden sich in der Diagonalen der Matrix, wo jede Art mit sich selbst verglichen wird. Der Maximalwert beträgt

$$SS_{max} = \sum_{i=1}^{p} r_{ii}^2$$

Symbol p steht für die Anzahl Arten. In unserem Beispiel ist $SS_{max} = 4$. Damit ergeben sich folgende Anteile erklärter Korrelation:

$$\text{Art 3:} \quad \frac{3.14}{4.0} * 100 = 78.5\%$$

$$\text{Art 2:} \quad \frac{0.67}{4.0} * 100 = 17\%$$

Die Artkombination (Art 3, Art 2) erklärt damit 78.5% + 17% = 95.5% der gesamten Korrelation. Dies bestätigt die oft gemachte Erfahrung, dass schon sehr wenige Arten die Struktur einer Vegetationstabelle weitgehend darzustellen vermögen (vgl. WILDI 1979).

Rechnet man RANK mit grösserer Genauigkeit für alle Arten, so erhält man

Rang	Art	Erklärte Varianz in %
1	3	78.5
2	2	17.0
3	4	4.5
4	1	0

Die Information, welche die Art 1 trägt, ist also völlig redundant, sodass eine Reduktion auf drei Arten verlustfrei erfolgen kann!

Die Interpretation der Ergebnisse ist mit einigen Schwierigkeiten verbunden. Zunächst ist zu beachten, dass Arten aufeinanderfolgenden Ranges immer sehr unterschiedliches Vorkommen aufweisen, wie hier 3 und 2, 2 und 4 usw. Sodann kann es ohne weiteres geschehen, dass an sich gute Zeigerarten (hier z.B. Art 1) wegfallen, weil ihre Varianz (oder Korrelation) bereits durch andere Arten vollständig erklärt wird. Das hat oft den Nachteil, dass im Feld gut sichtbare und leicht bestimmbare Arten aus dem Datensatz entfernt werden. Um dem vorzubeugen, wurde ein interaktives RANK Verfahren entwickelt, bei welchem der Experimentator aus einer kleinen Auswahl günstig gelegener Arten diejenige auswählt, welche definitiv zu rangieren ist (WILDI 1984). Damit steht dem Pflanzensoziologen ein Instrument zur Verfügung, das sehr rasch einen guten Einblick in die Aehnlichkeitsverhältnisse komplexer Datensätze vermittelt. Aus früher erwähnten Gründen sollte die Analyse nicht auf dem Korrelationskoeffizienten, sondern auf dem Skalarprodukt oder der Kovarianz basieren. Sehr gut interpretierbare Ergebnisse erzielt man mit den Kontingenzmassen von Jaccard, Soerensen bzw. van der Maarel. Allerdings muss dabei der Rechenvorgang wegen fehlender Metrik frühzeitig abgebrochen werden.

6.3 Rangierung nach Gruppenstruktur

Bei allen bisher behandelten Rangierungsverfahren wurde versucht, die Arten oder Standortsfaktoren vorgängig einer weiterführenden Analyse zu gewichten. Dabei wurde stets die gesamte Information innerhalb der Aehnlichkeitsmatrix verwendet. Hat man sich jedoch einmal zu einer Klassifikation entschlossen, so hat man auch auf einen Teil der Gesamtinformation des ursprünglichen Datensatzes verzichtet. Statt einzelner Aufnahmen stehen jetzt nur noch Gruppen im Zentrum des Interesses. Es lohnt sich, die einzelnen Arten oder Standortsfaktoren erneut zu gewichten, und zwar nach der Zuverlässigkeit, mit welcher sie auf eine Gruppenzugehörigkeit hinweisen. JANCEY (1979) schlägt dazu ein Verfahren vor, welches auf der Varianzanalyse basiert.

Die Varianzanalyse beruht auf dem Prinzip, dass sich die Varianz einer gruppierten Stichprobe wie folgt aufteilen lässt:

Totalvarianz der Stichprobe	=	Varianz innerhalb der Gruppen	+	Varianz zwischen den Gruppen

Bekanntlich berechnen sich die Varianzen als Summen von Abweichungsquadraten (Abschnitt 4.3). Da obige Beziehungen auch für die nicht quadrierten Abweichungen gelten, lässt sich der Sachverhalt leicht grafisch veranschaulichen (Abb. 6.1).

Ausgangspunkt bilde eine Stichprobe, bestehend aus drei Gruppen zu je drei Individuen. Die Indices in den nachfolgenden Formeln bedeuten:

i	für die laufende Aufnahme,	$i = 1,\ldots,n$
j	für die laufende Art,	$j = 1,\ldots,p$
g	für die laufende Gruppe,	$g = 1,\ldots,m$
t_g	für die laufende Aufnahme in der Gruppe g,	$t_g = 1,\ldots,m_g$

Man berechnet zuerst die Mittelwerte $\bar{x}_j$ für jede Art j. Für die Summe der Abweichungsquadrate ST_j^2 ergibt sich dann

$$ST_j^2 = \sum_{i=1}^{n} (x_{ji} - \bar{x}_j)^2$$

für jede Art j. Für die Varianz innerhalb der Gruppen wird analog vorgegangen, wobei die Abweichungsquadrate aller m Gruppen noch zu summieren sind:

$$SI_j^2 = \sum_{g=1}^{m} \sum_{t_g=1}^{m_g} (x_{jt_g} - \bar{x}_g)^2$$

Darin ist g die Gruppennummer, m_g die Anzahl Individuen in

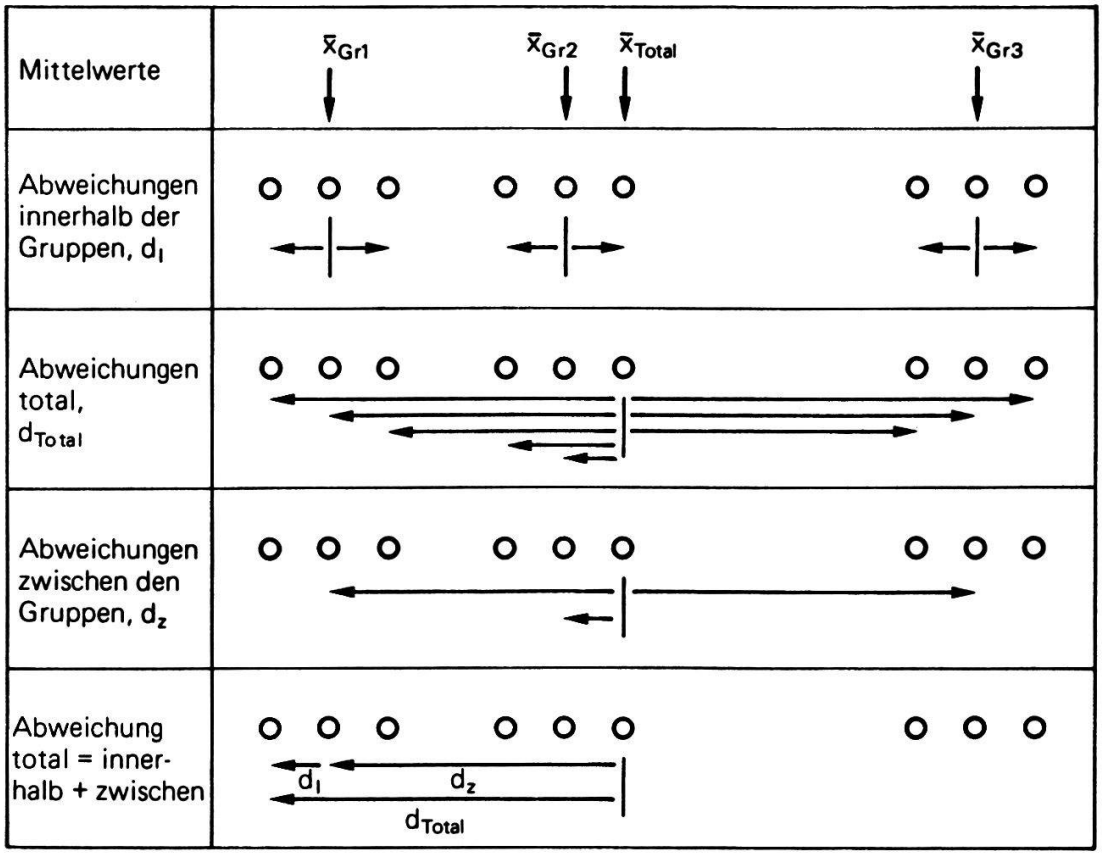

Abb. 6.1 Beziehungen zwischen der Gesamtabweichung d_{Total}, der Abweichung innerhalb einer Gruppe d_I und der Abweichung zwischen den Gruppen d_Z. Der Gesamtmittelwert ist $\bar{x}_{Total}$, der Mittelwert der Gruppe g ist $\bar{x}_{Gr\ g}$.

Gruppe g und x_{jt} der Messwert (Abundanz) der Art j in der Aufnahme t der Gruppe g. Die Summe der Abweichungsquadrate zwischen den Gruppen beträgt

$$SZ_j^2 = \sum_{g=1}^{m} m_g (\bar{x}_g - \bar{x}_j)^2.$$

Man ersieht daraus, dass die Abweichungen zwischen dem Gesamtmittelwert und den Gruppenmittelwerten jeweils mit der Gruppengrösse m_g zu multiplizieren sind.

Am folgenden Beispiel soll nun demonstriert werden, dass sich die verschiedenen Summen der Abweichungsquadrate als Messgrösse für den Zeigerwert von Arten eignen. Daten und Resultate sind zusammengefasst in Tab. 6.3. Art 1 ist so gleichmässig auf alle Gruppen verteilt, dass sie als Zeigerart sicher nicht in Frage kommt.

Genau gegenteilig verhält sich Art 3, bei welcher die gruppeninternen Abweichungen klein, diejenigen zwischen den Gruppen dafür gross sind. Art 3 ist somit ein optimaler Zeiger - zumindest für Gruppe 1 gegenüber 2 und 3. Art 2 schliesslich steht zwischen den beiden Extremen.

Man könnte nun der Idee verfallen, die Summe der Abweichungsquadrate zwischen den Gruppen direkt als Mass für die Güte des Zeigerwertes zu verwenden. In Realität treten aber nicht so einfache Fälle auf wie in Tab. 6.3. Insbesondere unterscheidet sich die Gesamtvarianz ST_j^2 von Art zu Art. Stichproben- und Gruppengrössen beeinflussen ausserdem die Zuverlässigkeit der Ergebnisse. JANCEY (1979) standardisiert deshalb zuerst die Artvektoren j,

$$x'_{ji} = (x_{ji} - \bar{x}_j)/s_j.$$

x'_{ji} ist der transformierte Artwert (Abundanz) der Art j in der Aufnahme i. s_j ist die Standardabweichung der Art j. Um die verminderte Zuverlässigkeit des Ergebnisses bei kleinen Gruppen auszugleichen, wird darin durch n-1 und nicht wie in

Tabelle 6.3.

Rangierung von Arten auf Grund einer Klassifikation. A: Rohwerte. B: Standardisierte Artvektoren. C: Gruppeninterne Statistik. D: Summen der Abweichungsquadrate, Varianzen und F-Werte sowie die resultierende Rangfolge.

A

Aufn.	1	2	3	4	5	6	7	8	9	$\bar{x}_j$	$\Sigma(x_{ji}-\bar{x}_j)^2$	s_j
Art 1	3	-	-	2	-	-	1	-	-	0.666	10	1.118
Art 2	3	2	-	1	-	-	-	-	-	0.666	10	1.118
Art 3	3	2	1	-	-	-	-	-	-	0.666	10	1.118

B

Aufn.	1	2	3	4	5	6	7	8	9	ST_j^2
Art 1	2.09	-0.6	-0.6	1.19	-0.6	-0.6	0.3	-0.6	-0.6	8.0
Art 2	2.09	1.19	-0.6	0.3	-0.6	-0.6	-0.6	-0.6	-0.6	8.0
Art 3	2.09	1.19	0.3	-0.6	-0.6	-0.6	-0.6	-0.6	-0.6	8.0

C

	$\bar{x}_1$	$ST_{j_1}^2$	$\bar{x}_2$	$ST_{j_2}^2$	$\bar{x}_3$	$ST_{j_3}^2$
Art 1	0.30	4.82	0	2.14	-0.3	0.54
Art 2	0.89	3.75	-0.3	0.54	-0.6	0
Art 3	1.19	1.6	-0.6	0	-0.6	0

D

	SI_j^2	SZ_j^2	ST_j^2	VI	VZ	F	Rang
Art 1	7.50	0.50	8.0	1.25	0.25	0.20	3
Art 2	4.29	3.71	8.0	0.715	1.86	2.60	2
Art 3	1.60	6.40	8.0	0.266	3.20	12.03	1

Kap. 3 durch n dividiert:

$$s_j = [\frac{1}{n-1} \sum_i (x_{ji} - \bar{x}_j)^2]^{1/2}$$

In Tab. 6.3, A lässt sich leicht nachrechnen, dass die Mittelwerte $\bar{x}_j$ aller Arten gleich 0.666 sind. Weiter beträgt die Summe der Abweichungsquadrate von diesem Mittelwert je genau 10.0. Somit ist auch s_j überall gleich, nämlich 1.118. In Tab. 6.3, B, ist dieselbe Tabelle in transformierter Form wiedergegeben. Das erste Element berechnet sich wie folgt:

$$x'_{ji} = (3-0.666)/1.118 = 2.09$$

Nun wird für die transformierten Werte in Abb. 6.3, B, die Summe der Abweichungsquadrate berechnet. Da der Mittelwert aller Arten nun null ist, brauchen nur die Elemente quadriert und addiert zu werden:

$$ST_j^2 = \sum_{i=1}^{n} x'^2_{ji}.$$

Die Ergebnisse betragen überall 8.0 (was n-1 entspricht!). Nun kann zur Berechnung der gruppeninternen Summen der Abweichungsquadrate geschritten werden (Tab. 6.3, C). Für die erste Art in der ersten Gruppe erhält man den Mittelwert

$$\bar{x}_1 = 1/3\ (2.09-0.6-0.6) = 0.30$$

Die Summe der Abweichungsquadrate ergibt

$$ST_{11}^2 = (2.09-0.3)^2+(-0.6-0.3)^2+(-0.6-0.3)^2$$
$$= 4.824$$

Wir kommen nun auf die in Tab. 6.2 dargestellten Beziehungen zurück und sind in der Lage, die Summe der Abweichungsquadrate innerhalb aller Gruppen zu berechnen. Für die erste Art findet man (Tab. 6.3, D):

$$SI_{11}^2 = 4.82+2.14+0.54 = 7.50$$

Die Summe der Abweichungsquadrate zwischen den Gruppen, SZ_j^2 wird aus den Gruppenmittelwerten berechnet, für Art 1 also 0.30, 0,-0.3. In Tab. 6.3, D ist auch die totale Summe der Abweichungsquadrate aus Tab. 6.3, B übernommen worden. Es bestätigt sich deren Zerlegbarkeit in SI_j^2 und SZ_j^2:

$$ST_j^2 = SI_j^2 + SZ_j^2$$

Für Art 1 gilt:

$$8.0 = 7.50 + 0.50.$$

Das Verhältnis zwischen SI_j^2 und SZ_j^2 ist tatsächlich charakteristisch für die Differenzierungskraft jeder Art, wie leicht in Tab. 6.3, D nachzuprüfen ist. JANCEY (1979) verwendet nun aber statt der Summen der Abweichungsquadrate die Varianzen. Damit wird ein eventueller Unterschied in der Grösse der Gruppen ausgeglichen. Sie berechnen sich gemäss der Theorie der Varianzanalyse wie folgt (vgl. z.B. GAENSSLEN und SCHUBÖ 1973):

$$Vz = \frac{1}{m-1} SZ_j^2$$

$$Vi = \frac{1}{n-m} SI_j^2$$

Darin ist n wiederum die Anzahl Aufnahmen, m die Anzahl Gruppen. Die Ergebnisse sind in Tab. 6.3, D dargestellt. Das endgültige Kriterium zur Rangierung der Arten ist nun der F-Wert der Varianzanalyse. Er ist nichts anderes als das Verhältnis der Varianzen zwischen und innerhalb der Gruppen:

$$F = \frac{VZ}{VI}$$

Grosse F-Werte deuten auf eine grosse Trennkraft der Arten hin, kleine auf ein unspezifisches Vorkommen der Arten innerhalb der Vegetationstabelle. Dementsprechend erhält Art 3 in Tab. 6.3 Rang 1. Die F-Werte sind der schliessenden Statistik gemäss F-verteilt mit df = 2 und 6 Freiheitsgraden. Anhand einer F-Tabelle kann jeder Wert auf Ueberzufälligkeit geprüft werden (BARTEL 1972, S. 187). Für eine Irrtumswahrscheinlichkeit von 5% finden wir

$$F(df = 2,6;p = 0.05) = 5.14$$

Damit ist der F-Wert der Art 3 signifikant von Null verschieden, diejenigen der Arten 1 und 2 nicht. Da auch hier die Voraussetzungen für die Durchführung des F-Testes nicht geklärt wurden, ist diese Aussage als reine Interpretationshilfe zu betrachten.

Mit dem dargelegten Beispiel wird die Trennkraft der Arten bezüglich aller drei Gruppen bestimmt. JANCEY (1979) weist darauf hin, dass die Rechnung auch nur für einen Teil einer Tabelle durchgeführt werden kann, also z.B. für die Gruppen 1 und 2. Eine mögliche Anwendung der Methode besteht nun darin, die ursprüngliche Vegetationstabelle auf die trennenden Arten zu reduzieren, womit man einen Vegetationsschlüssel erhält. Aus der Literatur sind leider noch keine grösseren Anwendungen bekannt.

6.4 Stressanalyse

Ein mögliches Ziel von Rangierungsanalysen besteht in der Reduktion der Artenzahl. Dies verbessert die Uebersicht über die Daten und erleichtert rechenintensive Analysen. Mit der Methode RANK haben wir bereits ein Verfahren erörtert, welches für eine reduzierte Vegetationstabelle geeignete Ar-

ten direkt bezeichnet. Wieviele Arten letztlich weggelassen werden können, bleibt eine offene Frage. Sicher ist, dass die Datenstruktur durch die Auswahl möglichst wenig Aenderungen erfahren sollte. Die Stressanalyse zeigt einen Weg auf, durch Artenreduktion auftretende Verzerrungen unter Kontrolle zu halten (ORLOCI 1978).

Als Beispiel verwenden wir die Vegetationstabelle in Tab. 6.2. Die Arten sollen dabei in der Reihenfolge verwendet werden, wie dies das Ergebnis der RANK-Analyse als zweckmässig erscheinen lässt. Zur Ermittlung der Aufnahmenstruktur wird eine Matrix der Euklidschen Distanzen zwischen den Aufnahmen gerechnet. Werden alle vier Arten berücksichtigt, so erhalten wir

$$D(1,2,3,4) = \begin{matrix} 0.00 & 1.00 & 2.65 & 3.16 \\ & 0.00 & 2.45 & 2.65 \\ & & 0.00 & 1.73 \\ & & & 0.00 \end{matrix}$$

Der untere Teil der Matrix braucht nicht geschrieben zu werden, da die selben d-Werte auftreten wie oben. Lässt man nun die letztrangierte Art (Nr. 1) weg, so werden die Distanzen generell kleiner. Wir berechnen also die neue Matrix und erhalten:

$$D(2,3,4) = \begin{matrix} 0 & 1.00 & 2.45 & 2.45 \\ & 0 & 2.24 & 1.73 \\ & & 0 & 1.41 \\ & & & 0 \end{matrix}$$

Da nur die relativen Veränderungen der Distanzen interessieren, wird zwischen allen von null verschiedenen N(N-1)/2 Distanzen der Korrelationskoeffizient p(DP;Dp) berechnet. Im obigen Beispiel gilt

	1.00	1.00	
	2.65	2.45	
D(1,2,3,4)=DP=	3.16	D(2,3,4)=Dp= 2.45	p(DP;Dp)=0.903
	2.45	2.24	
	2.65	1.73	
	1.73	1.41	

Nun wird auch noch Art 4 weggelassen. Die Aufnahmen weisen jetzt folgende Distanzen auf:

$$D(2,3) = \begin{matrix} 0 & 1.00 & 1.41 & 2.24 \\ & 0 & 1.00 & 1.41 \\ & & 0 & 1.00 \\ & & & 0 \end{matrix}$$

Wir berechnen wiederum die Korrelation mit der Aehnlichkeitsstruktur des vollständigen Datensatzes und erhalten

	1.00		1.00	
	2.65		1.41	
D(1,2,3,4) =	3.16	D(2,3) =	2.24	p(DP;Dp) = 0.751
	2.45		1.00	
	2.65		1.41	
	1.73		1.00	

Dass die Korrelation weiter gesunken ist, entspricht durchaus den Erwartungen. Es kommt aber auch vor, dass die Struktur mit einem Minimum an Arten hervorragend reproduziert wird. In unserem Beispiel ist dies der Fall. Nur noch auf der Art 3 basierend erhält man

	1.00		0.00	
	2.65		1.00	
D(1,2,3,4) =	3.16	D(3) =	1.00	p(DP;Dp) = 0.906
	2.45		1.00	
	2.65		1.00	
	1.73		0.00	

Dass Art 3 alleine die Struktur so gut zu reproduzieren ver-

mag, erstaunt weiter nicht. Sie differenziert klar zwischen den beiden Gruppen (1,2) und (3,4) in der Tabelle und qualifizierte sich auch als beste Differentialart bei RANK. Zusammenfassend lässt sich der Schluss ziehen, dass die Reduktion der Artenzahl die Datenstruktur in unerwarteter Weise verändern kann. Eine Untersuchung mit Hilfe der Stressanalyse ist daher empfehlenswert.