

# Eine Uebersicht

Objekttyp: **Chapter**

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

Band (Jahr): **90 (1986)**

PDF erstellt am: **25.08.2024**

## **Nutzungsbedingungen**

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

## **Haftungsausschluss**

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

## 1. Eine Uebersicht

### 1.1 Besonderheiten statistischer Analysen

Die Botanik und insbesondere die Vegetationskunde sind Wissenschaften, die für die Anwendung von Analysen der explorativen Statistik bekannt sind. Viele ihrer Fachleute haben sich allerdings kaum je solcher Hilfsmittel bedient. In der Tat waren es stets kleinere Forschergruppen innerhalb verschiedenster vegetationskundlicher Schulen, welche die Anwendung statistischer Modelle auf vegetationskundliche Daten studierten und entwickelten. Viele Praktiker verwendeten bisher mangels leistungsfähiger Rechner subjektiv-manuelle Methoden der Datenverarbeitung, die mit der Zeit etwas vereinheitlicht wurden. ELLENBERG's (1956) Vorschriften zum Ordnen einer Vegetationstabelle sind ein Beispiel. Mit der Verfügbarkeit schneller Rechenautomaten hat sich nun die Situation geändert. Selbst aus der Sicht des reinen Anwenders lohnt es sich, nach Sinn und Leistungsfähigkeit statistischer Methoden zu fragen.

In allen nachfolgenden Ueberlegungen treten immer wieder die Begriffe Pflanzensoziologie und Pflanzenökologie auf. Deren Definition kann jedem einschlägigen Lehrbuch entnommen werden. Wesentlich ist, dass die Pflanzenökologie die Verbindung sucht zwischen Standortdaten einerseits und pflanzensoziologischen Daten andererseits. Wenn gelegentlich von speziellen Datenstrukturen die Rede ist, so bezieht sich dies nur auf die letzteren. Standortliche Daten dagegen unterscheiden sich von solchen anderer naturwissenschaftlicher Disziplinen kaum und lassen sich mit dem normalen statistischen Instrumentarium bearbeiten. Ein ausgesprochen fachspezifisches Problem bildet dagegen wiederum - immer aus der Sicht der Analyse - die Untersuchung des Zusammenhanges zwischen pflanzensoziologischen und standörtlichen Daten.

Wer erstmals statistische Analysemethoden einsetzt, wird feststellen, dass - im Gegensatz zu manuellen Methoden - die Resultate exakt reproduzierbar sind. Meist spielt es keine Rolle, in welcher Reihenfolge die Daten in den

Analyseprozess eingeschleust werden. Sofern die Methode genau beschrieben wird, ist die Verarbeitung für Drittpersonen bis ins Detail einsichtig. Ein weiterer Vorteil liegt in der hohen Effizienz. Wird ein multivariater Datensatz analysiert, also zum Beispiel eine Vegetationstabelle oder eine Tabelle von Standortdaten, so bearbeiten gute Methoden alle Aufnahmen gleichzeitig und gleichberechtigt. Die Resultate widerspiegeln deshalb die Gesamtähnlichkeitsstruktur des Datensatzes. Dies wiederum erklärt den erheblichen rechnerischen Aufwand, mit dem solche Analysen verbunden sind. Dass dafür Computer eingesetzt werden müssen, mag einerseits als Nachteil gewertet werden. Andererseits hat dies den willkommenen Nebeneffekt, dass sich langwierige Arbeitsschritte wie das Umschreiben von Tabellen und das Zeichnen von Diagrammen automatisieren lassen. In der vegetationskundlichen Forschung wird davon zunehmend und mit bemerkenswertem Erfolg Gebrauch gemacht. Die neuere Fachliteratur bestätigt diese Feststellung (vgl. z.B. PODANI 1984). Sie wirft aber zugleich Fragen auf, mit denen sich auch der Leser dieser Ausführungen konfrontiert sieht.

Auffällig ist zunächst, dass es ausserordentlich viele verschiedene Methoden gibt und dass laufend neue entwickelt werden. Warum diese Vielfalt? Jeder Methode entspricht ein mehr oder weniger eigenständiges geometrisches oder statistisches Modell. Solche Modelle sind in fast beliebiger Zahl formulierbar. Der Vegetationskundler sucht für seine Daten ein Modell, das die Struktur seiner Befunde möglichst klar abzubilden vermag. Seine Analysen sind deshalb erfolgversprechender, wenn er unter vielen Modellen geschickt auszuwählen versteht. Aus diesem Grunde ist es auch nicht sinnvoll, im Interesse einer Vereinheitlichung ein einziges Verfahren zu propagieren. Vielmehr ist immer ein - natürlich überblickbares - Spektrum von Möglichkeiten in Erwägung zu ziehen.

In Anbetracht der Vielfalt der Methoden drängt sich auch die Frage auf, weshalb für die Pflanzenökologie spezielle Methoden erforderlich sind. Die Antwort ergibt sich leicht aus Erfahrung im Umgang mit solchen Daten. In der Praxis tritt

immer wieder ein ähnliches Spektrum von Datenstrukturen auf, welches eben für das Fachgebiet typisch zu sein scheint. In der Tat weisen Vegetationsdaten - im Vergleich zu Datensätzen anderer Wissenschaften (geografische, sozialwissenschaftliche, ökonomische Daten usw.) - einen hohen Standardisierungsgrad auf. Dies liegt an den allgemein gültigen Naturgesetzen, denen die Vegetation unterworfen ist, aber auch an der Verwendung weit verbreiteter, teilweise standardisierter Erhebungsverfahren (vgl. z.B. MUELLER-DOMBOIS und ELLENBERG 1974). Es ist auch als Vorteil zu werten, dass die Daten nicht bereits bestehenden Statistiken entstammen, sondern normalerweise auf einem räumlich und zeitlich nach wissenschaftlichen Kriterien festgelegten Stichprobenverfahren basieren. Insbesondere multivariate Verfahren arbeiten unter solchen Voraussetzungen optimal.

Inhomogenes Datenmaterial erfordert die Verwendung stark standardisierender Methoden. Der weit verbreitete Produktmomentkorrelationskoeffizient ist ein Beispiel dafür. Gerade bei diesem sind aus formalen Gründen fast immer Informationsverluste in Kauf zu nehmen. Bei der Analyse von Vegetationsdaten kann demgegenüber eine Auswahl wenig standardisierender und deshalb gut differenzierender Verfahren eingesetzt werden. Weil die Stichprobenplanung als Voraussetzung für ein gutes Gelingen eine entscheidende Rolle spielt, kann durch sie zugleich eine weitere Verbesserung der Effizienz vieler Methoden erzielt werden. Dies ist der einzige Weg, um die Reproduzierbarkeit numerischer Analysen voll auszunützen.

Schliesslich stellt sich die Frage, ob die beobachteten Datenstrukturen allgemeingültig oder bloss ein Abbild des einigermaßen willkürlich gewählten Analyseverfahrens sind. Die letzte Vermutung kann wohl für strukturelle Details eines multivariaten Datensatzes zutreffen. Für die Interpretation ist es in diesem Falle besonders wichtig, die genaue Funktionsweise des Analyseverfahrens zu kennen. Hingegen zeigt die Erfahrung, dass die meisten Methoden eindeutige Zusammenhänge und Strukturen aufzudecken vermögen. Wer die Möglichkeit zur Variation seiner Methoden besitzt, verfügt

also zugleich über ein Instrument, um die Bestimmtheit seiner Befunde abzuschätzen.

Statistische Analysen können nicht isoliert betrachtet werden. Sie sind nur ein Schritt auf dem Weg zu wissenschaftlichen Erkenntnissen. Es ist nicht Absicht dieser Arbeit, auf alle methodischen Probleme einzugehen. So werden z.B. die pflanzensoziologischen Aufnahmemethoden nicht diskutiert. Die Planung von Stichprobenkonzepten ist jedoch ein so wichtiger Schritt, dass ihr ein eigenes Kapitel gewidmet ist. Doch zunächst sind einige einführende Überlegungen zu Datenstrukturen und Datenmodellen zu behandeln.

## **1.2 Das Untersuchungsobjekt: Die Stichprobe**

Multivariate Analysemethoden beschäftigen sich mit Stichproben. Jede Stichprobe besteht aus einer kleineren oder grösseren Anzahl von Individuen (Stichprobenelementen), welche sich mehr oder weniger voneinander unterscheiden. Beispiele sind die Schüler einer Schulklasse, die Bäume eines Waldstückes, die Häuser einer Siedlung usw. In der Pflanzensoziologie bildet oft die Vegetationstabelle die Stichprobe. Individuen sind floristische und ökologische Aufnahmen. Ein Beispiel zeigt Abb. 1.1. Wir haben hier einen multivariaten Datensatz vor uns. Im Gegensatz zu univariaten Datensätzen sind die Individuen grundsätzlich durch eine Vielzahl von Merkmalen (vgl. GAUCH 1982, S.4 und 5) charakterisiert. Bei Stichproben menschlicher Populationen wären analoge Merkmale Eigenschaften wie Körpergrösse, Haarfarbe, Alter usw., bei soziologischen Studien auch Bildung, Religion, Weltanschauung usw. In der Pflanzenökologie sind es Artvorkommen (Abundanz, Dominanz), Standortsfaktoren, Lebensformen, phänologischer Zustand und anderes mehr, welche die Individuen charakterisieren. Entscheidend an der eben dargestellten Sichtweise ist, dass Aufnahmen und Tabellen die fundamentalen Einheiten der Analyse bilden. Im Gegensatz etwa zur Auffassung von BRAUN-BLANQUET (1964) spielen Pflanzengesellschaften erst im Zusammenhang mit Ergebnissen eine wesentliche Rolle.

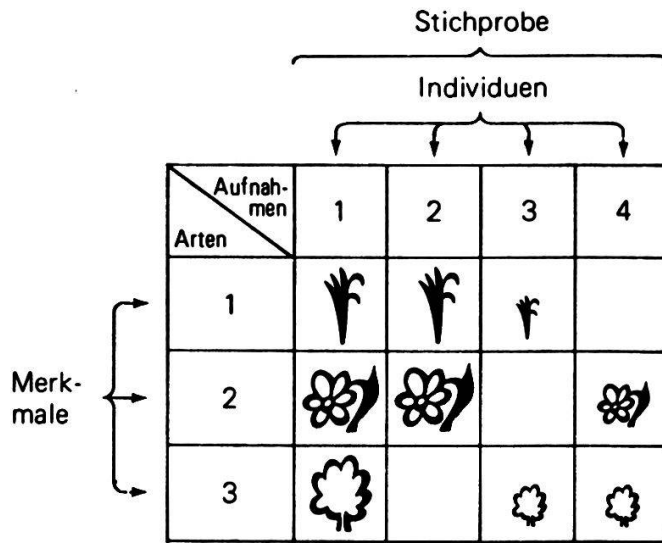


Abb. 1.1 Beispiel einer Stichprobe, bestehend aus 4 Individuen, welche durch je 3 Merkmale charakterisiert sind.

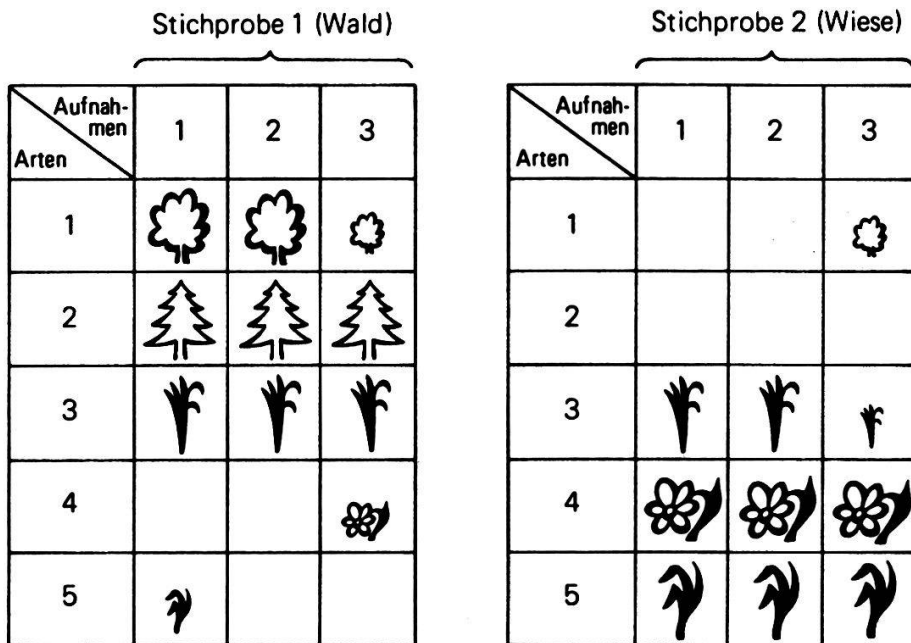


Abb. 1.2 Zwei klar unterscheidbare Stichproben. Sie werden sinnvollerweise separat analysiert.

Die Abgrenzung der Grundgesamtheit als Gegenstand einer Analyse ist ein wichtiger eigenständiger Schritt im Rahmen einer Untersuchung (vgl. Kap. 2.2.1). Manchmal ist eine einzige unstrukturierte Stichprobe Gegenstand einer Analyse. Bei umfangreicheren Forschungsprojekten sind kompliziertere Verhältnisse jedoch die Regel. Dabei kann es sich als sinnvoll erweisen, das gesamte Datenmaterial a priori in Stichproben zu unterteilen, welche einzeln analysiert werden. Dies zeigt Abb. 1.2. Die beiden Datensätze lassen keinen Zweifel offen, dass sie in jeder Hinsicht (d.h. bezüglich aller Merkmale) verschieden sind und deshalb ohne weiteres getrennt analysiert werden können. Es gibt aber öfters Situationen, bei denen die Aufteilung in verschiedene Stichproben nicht von vornherein gegeben ist. Dabei lassen wohl einige Merkmale eine bestimmte Unterteilung als sinnvoll erscheinen, andere widersprechen dem jedoch. Gleichzeitig mag es aber wenig sinnvoll sein, so zu handeln, als ob eine unstrukturierte Stichprobe vorläge. Die Untersuchung einer vermuteten Gruppenstruktur ist dann Gegenstand der Analyse. Im Beispiel von Abb. 1.3 liegt die Gruppierung dem vorgegebenen Merkmal "Exposition" zugrunde. Die daran anschließende Analyse soll dann zeigen, inwiefern die Kombination der anderen Merkmale diese Vorgabe bestätigen.

### 1.3 Strukturmodelle von Stichproben

Ziel multivariater Analysen ist die Untersuchung der Struktur, dem Vorhandensein von Gesetzmässigkeiten innerhalb der Stichprobe. Man bildet zu diesem Zwecke geometrische Modelle der Daten. Alle diese Modelle zielen darauf ab, die Datenstruktur vereinfacht abzubilden. Der Preis, welcher für jede Vereinfachung zu bezahlen ist, besteht in einem Informationsverlust. Der Gewinn besteht darin, dass in der Stichprobe vorhandene Gesetzmässigkeiten klarer hervortreten als in den Originaldaten.

Soll nun die Struktur einer multivariaten Stichprobe untersucht werden, so geschieht dies durch den Vergleich ihrer Individuen oder ihrer Merkmale. Dafür stehen sehr unterschiedliche Methoden zur Verfügung. Im allgemeinen wird man



















Stichprobe						
Gruppe 1						
Gruppe 1			Gruppe 2			
Aufnahmen	1	2	3	4	5	6
Exposition	N	N	N	SE	SE	SE
Art 1						
Art 2						
Art 3						

Abb. 1.3 Eine Stichprobe, welche zum vornherein in zwei Gruppen unterteilt worden ist. Die Prüfung der Unterschiedlichkeit dieser Gruppen kann ein Analyseziel sein.





	Aufn. 1	Aufn. 2	Übereinstimmung s	Unterschied d
Art 1			1	—
Art 2			—	1
Art 3			—	1
			$\Sigma s = 1$	$\Sigma d = 2$

Abb. 1.4 Vergleich zweier Aufnahmen.  $\sum_i x_i$  ist ein Ähnlichkeitsmass,  $\sum_i d_i$  ein Distanzmass (Unähnlichkeitsmass).



zunächst die Merkmale eines Individuenpaares einzeln vergleichen und die so erhaltenen Ergebnisse zusammenfassen. Abb. 1.4 zeigt zwei mögliche Vorgehensweisen: Je nach Definition des Vergleichskoeffizienten liegt am Schluss ein Mass für die Aehnlichkeit oder aber für die Unähnlichkeit (Distanz) der Individuen vor.

Die umfassende Analyse der Stichprobe erfordert nun eine Bestimmung der gesamten Stichprobenstruktur. Man kommt dabei nicht umhin, jedes Individuum mit jedem anderen zu vergleichen. Schon bei bescheidenen Stichprobengrössen ist daher eine grosse Zahl von Vergleichskoeffizienten zu berechnen. Abb. 1.5 zeigt ein einfaches Beispiel. Die Vegetationstabelle oben besteht aus vier Individuen. Unten ist die sich ergebende Stichprobenstruktur dargestellt. Es handelt sich um eine Aehnlichkeitsmatrix mit 16 Koeffizienten. Da jeder Koeffizient je zwei mal erscheint, ist nur etwas mehr als die Hälfte der Matrix zu berechnen. Die Zahl der paarweisen Vergleiche,  $M$ , ergibt sich nach der Formel

$$M = N(N+1)/2.$$

Darin ist  $N$  die Anzahl Individuen. Für die Tabelle in Abb. 1.5 erhalten wir:

$$M = 4(4+1)/2 = 10.$$

Der Aufwand zur Bestimmung der Stichprobenstruktur nimmt mit steigender Stichprobengrösse rasch zu. So ergibt sich für eine Vegetationstabelle mit 100 Aufnahmen

$$M = 100(100+1)/2 = 5050.$$

Die Aehnlichkeitsstruktur einer Stichprobe ist mit einer solchen Matrix abschliessend definiert. Sie ist aber ihrer Grösse wegen praktisch nicht überblickbar. Um sie zu analysieren, bedarf es spezieller Methoden, die in den späteren Kapiteln behandelt werden.

Für pflanzenökologische Untersuchungen bedeutungsvoll sind

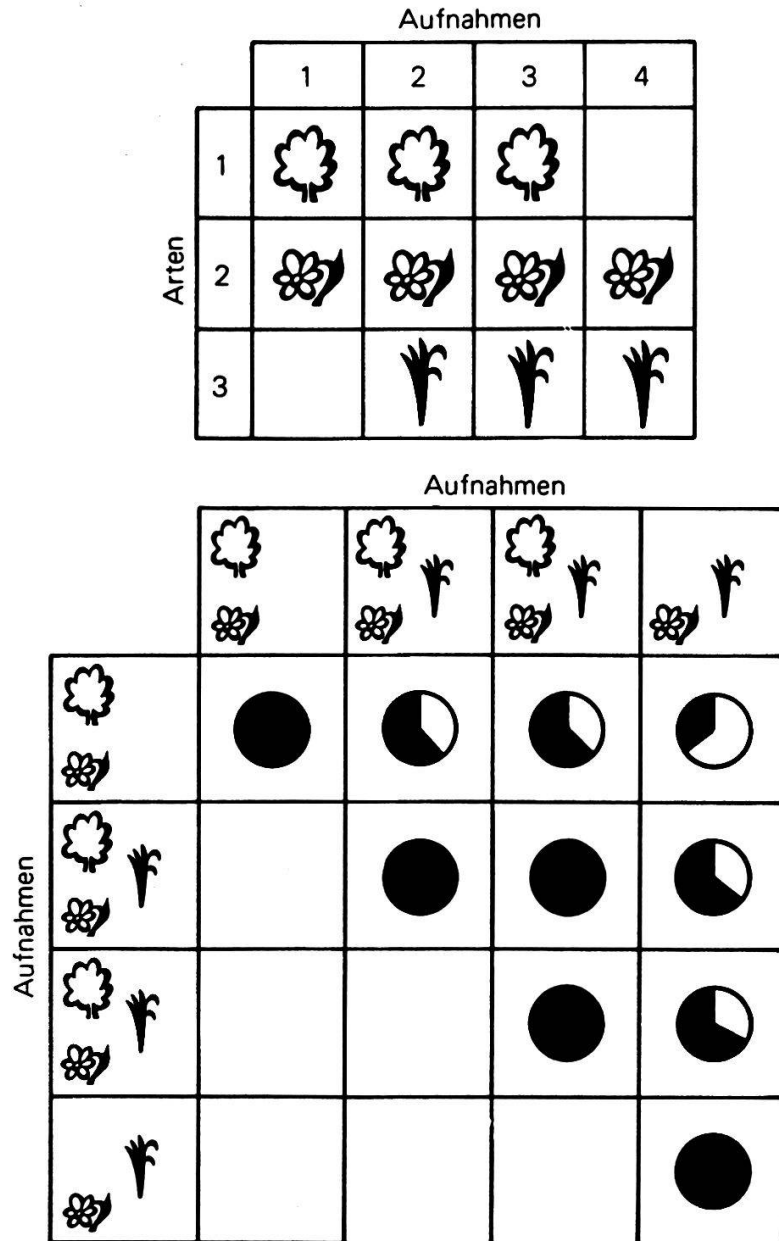


Abb. 1.5 Bestimmung der Aehnlichkeitsstruktur einer Stichprobe. Aus 4 Aufnahmen (oben) wird eine Aehnlichkeitsmatrix berechnet (unten). Die Grösse der schwarzen Kreissektoren symbolisiert das Mass des Zusammenhanges zwischen den Aufnahmen.

Gradienten- und Gruppenstrukturen. Solche lassen sich recht gut aus einer ausgewerteten, geordneten Vegetationstabelle herauslesen. Abb. 1.6 zeigt einige einfache Beispiele. Ganz eindeutige Verhältnisse (A,C) treten in Wirklichkeit kaum auf. Die Fälle B und D sind eher typisch.

Ob sich eine Stichprobe in Gruppen aufteilen lässt, wird z.B. mit Hilfe sogenannter Gruppierungsanalysen (Clusteranalysen) untersucht (Kap.5). Manchmal werden deren Resultate in Form von Dendrogrammen präsentiert (Abb. 1.7). Darin sind auf der X-Achse die Individuen aufgetragen, während die Y-Achse die Unterschiedlichkeiten der Gruppen wiedergibt. Die Stärke der Gruppenstruktur ergibt sich aus dem Verhältnis der Aehnlichkeiten innerhalb der Gruppen zu jenen zwischen den Gruppen. In Abb. 1.7, A ist eine klar zweigeteilte Stichprobe dargestellt. Dendrogramm B zeigt dagegen eine schwach ausgeprägte Struktur. Die Unterschiede innerhalb der Gruppen sind hier fast ebenso gross wie diejenigen zwischen denselben.

Um die Gradientenstruktur einer Stichprobe darzustellen, greift man öfters zur Ordination. Die Individuen werden dabei in einem zwei- oder dreidimensionalen Koordinatensystem so dargestellt, dass sich ihre Aehnlichkeitsbeziehungen aus den gegenseitigen Abständen möglichst genau herauslesen lassen. Abb. 1.8 zeigt zwei Beispiele. Darin werden je zwei Merkmale zur zweidimensionalen Darstellung von 8 Individuen verwendet. Verfügt man über drei oder mehr Merkmale, so lassen sich in dieser Weise nur Teile der Gesamtinformation wiedergeben. Die Information ist stets an die Achsen gebunden. Die Beiträge der Achsen zur Gesamtinformation verhalten sich additiv. Trägt beispielsweise die erste Achse 50%, die zweite 30% und die dritte 20% der Information, so ist die maximale Effizienz einer zweidimensionalen Ordination gleich 80%.

Oft wird erwartet, dass Vegetationsaufnahmen, die im Felde einem einfachen Gradienten entstammen, in einer Ordination auf einer geraden Linie erscheinen müssen (GAUCH 1982). Dies ist nicht der Fall. Das Erscheinungsbild in Abb. 1.8 ist

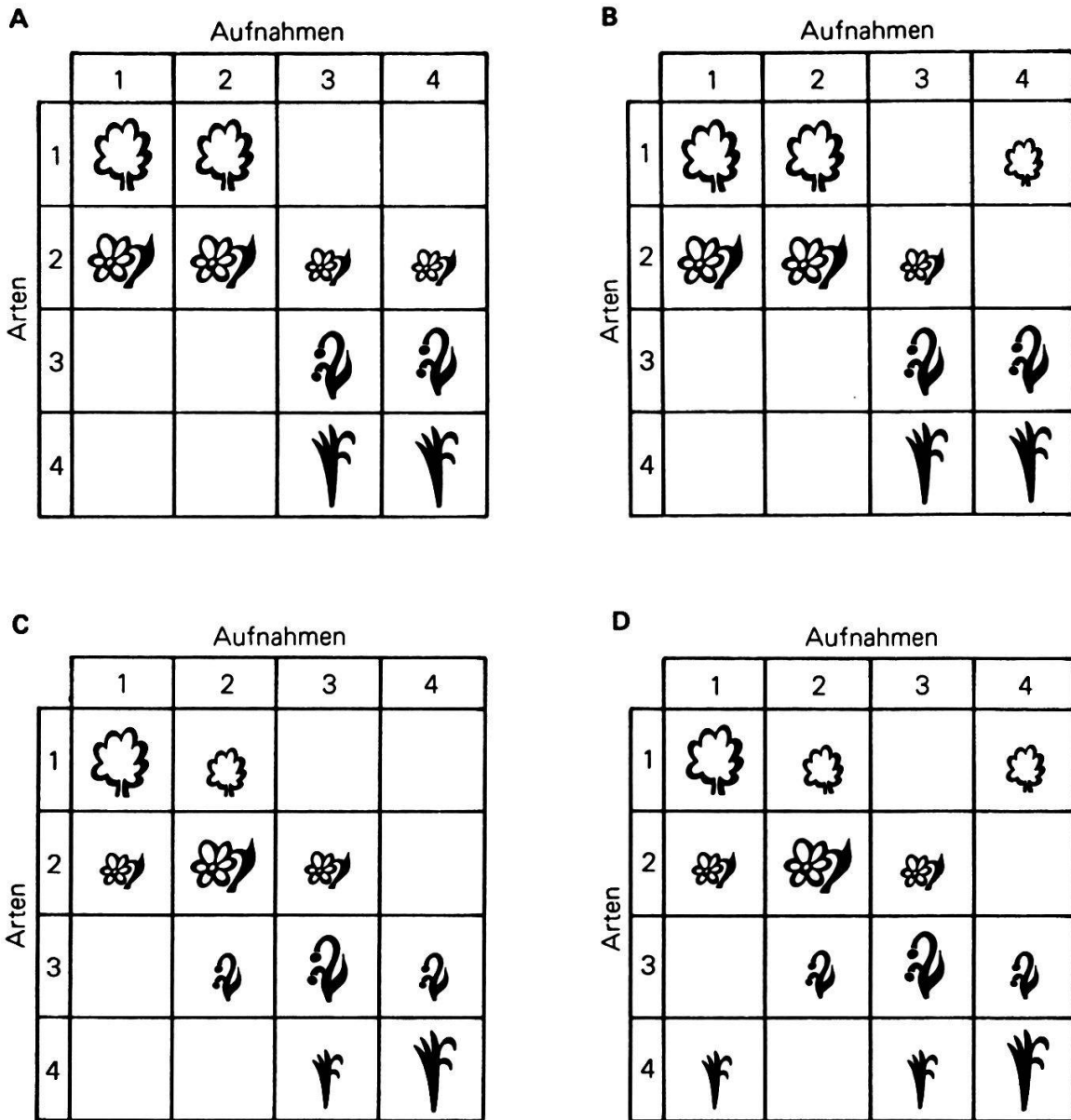


Abb. 1.6 Unterschiedlich strukturierte Vegetationstabellen. Deutliche Gruppen (A), schwache Gruppen (B), deutlicher Gradient (C), schwacher Gradient (D).

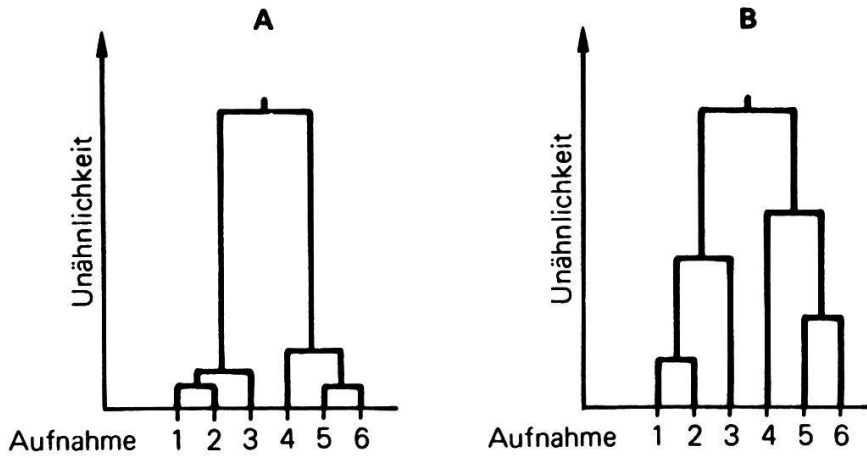


Abb. 1.7 Dendrogramme. Klare Gruppenstruktur mit zwei Gruppen (A), schwache Gruppenstruktur (B).

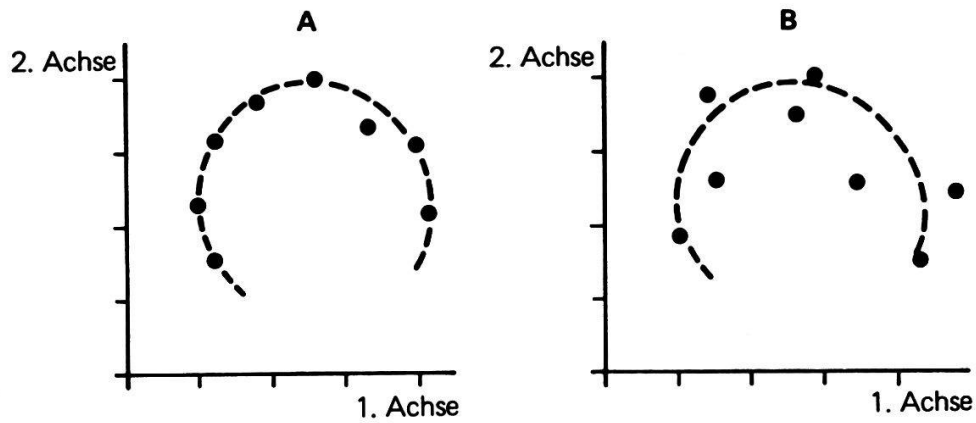


Abb. 1.8 Ordinationen. Klare Gradientenstruktur (A), schwache Gradientenstruktur (B). Umweltgradienten entstammende Vegetationsdaten zeigen in der Ordination oft die hier dargestellte Krümmung.

dagegen typisch, indem die Aufnahmen auf gekrümmten, oft fast kreisförmigen oder - im dreidimensionalen Raum - spiralförmigen Linien angeordnet sind. Die genauen Ursachen dafür werden in Kapitel 7 gezeigt. Weniger klare Gradienten ergeben Muster wie in Abb. 1.8, rechts, dargestellt. Die Aufnahmen sind hier bandförmig angeordnet. Strukturen dieser Art sind in der Praxis sehr häufig zu finden. Schliesslich können Ordinationen auch Gruppenstrukturen sehr differenziert wiedergeben. Beispiele finden sich in Kap. 5.

#### 1.4 Regressionsmodelle

Es gibt vegetationskundliche Untersuchungen, die sich mit der Erforschung der Datenstruktur begnügen. Dank guter Kenntnis der Standortsansprüche der Arten können solche Resultate ökologisch interpretiert und nutzbar gemacht werden. Eine beliebte Möglichkeit bietet die Verwendung der Zeigerwerte der Arten nach LANDOLT (1977) oder nach ELLENBERG (1979). Man stützt sich in dieser Weise auf alt bekanntes Wissen ab. Pflanzenökologische Forschung umfasst aber auch die Messung von Standortfaktoren. Ziel der Analyse des Zusammenhanges zwischen Vegetation und Standort ist die Bildung eines prädiktiven Modelles. Ein solches erlaubt die Voraussage des Standortes (auf Grund der Pflanzengesellschaft) oder der Pflanzengesellschaft (auf Grund des Standortes) oder aber eines Standortfaktors auf Grund der Ausprägung eines andern.

Ein einfaches prädiktives Modell ist die lineare Regression. Sind die Ausgangsdaten als Punkteschwarm in einem zweidimensionalen Koordinatensystem darstellbar, so wird versucht, die Abhängigkeiten der Individuen durch eine Gerade zu beschreiben. Abb. 1.9 zeigt einen solchen Fall. Wird dieses Modell als gültig akzeptiert, so kann bei vorgegebener Neigung der zugehörige pH-Wert abgelesen werden. Die horizontalen Abstände der Punkte von der Regressionsgeraden zeigen, welche Fehler bei einer solchen Aussage toleriert werden müssen. Abb. 1.9, unten, zeigt ein anderes Modell, das dieselbe Datenstruktur genauer wiedergibt. Es handelt sich um eine nichtlineare Regression. Die Pfeile weisen auf

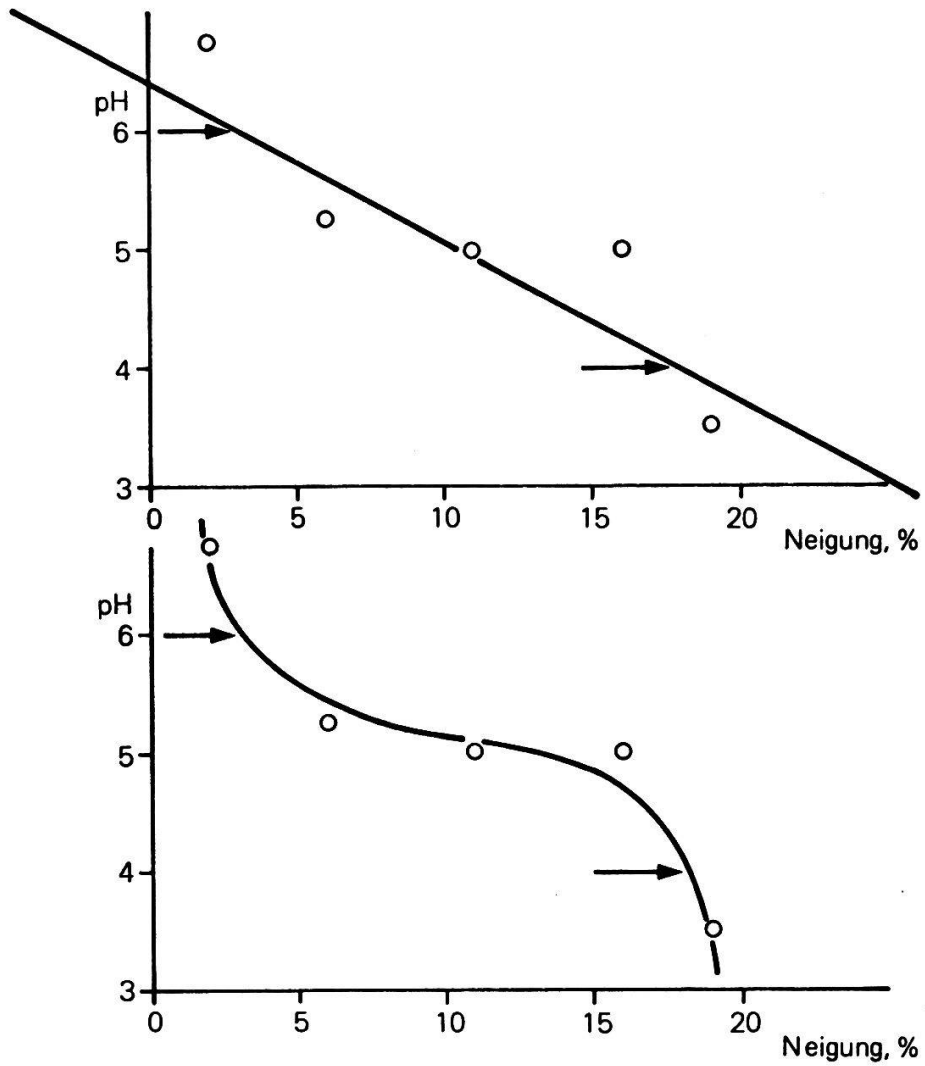


Abb. 1.9 Regressionsmodelle. Einfache, lineare Regression (oben), einfache nichtlineare Regression (unten).

Stellen hin, an welchen die Voraussage mit derjenigen des linearen Modelles identisch ist. Leider treten in der Pflanzenökologie selten Datensätze mit nur zwei Variablen auf, sodass sich mit der einfachen Regression die wenigsten Probleme bewältigen lassen.

Eine Möglichkeit, einen einfachen Vegetationsgradienten mit dem Standort zu verknüpfen, bietet die multiple Regressionsanalyse. In Abb. 1.10 werden zwei Faktoren, der pH-Wert und eine Mitteltemperatur, mit der Position von Vegetationsaufnahmen entlang eines Gradienten in Beziehung gesetzt. Die Lösung ist geometrisch darstellbar. Sie besteht aus einer Ebene, welche analog der Geraden einer einfachen Regressionsanalyse optimal einem Punkteschwarm (Aufnahmen) angepasst sein soll. Die eingezeichnete Aufnahme P demonstriert die Aussagemöglichkeiten des Modells: Werden ein pH-Wert und eine Temperatur vorgegeben, so lässt sich der zu erwartende Vegetationstyp (d.h. seine Entfernung vom Anfangspunkt des Gradienten) voraussagen.

Ein anderes, verwandtes Modell bietet die Möglichkeit, Stichproben mit Gruppenstruktur zu analysieren. In Abb. 1.11 wird diese diskriminanzanalytische Lösung dargestellt. Das besondere an der Situation ist, dass sich der Vegetationstyp weder aus dem pH-Wert allein noch aus der Temperatur voraussagen lässt. Beide Faktoren weisen innerhalb derselben Vegetationstypen gleichzeitig hohe und niedrige Werte auf. Wird jedoch eine günstige Kombination von Messwerten gefunden, so lassen sie sich eindeutig auseinander halten. Das Ergebnis lässt sich geometrisch als neue Achse,

$$y = f(\text{pH}, \text{ }^{\circ}\text{C})$$

darstellen. Es kann aber auch verbal, mit Hilfe einer "Grammatik" (DALE 1980) ausgedrückt werden. Im vorliegenden Beispiel könnte man formulieren:

- Bei hoher Temperatur und gleichzeitig niedrigem pH-Wert tritt bevorzugt Typ A auf.



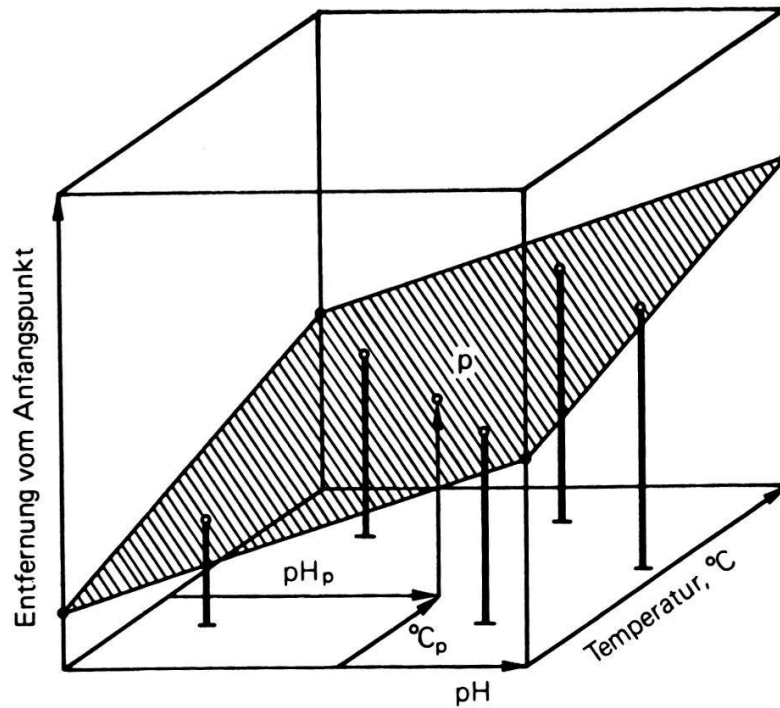


Abb. 1.10 Multiple Regression zwischen der Position von Aufnahmen entlang eines Transekten und den Standortsfaktoren pH-Wert und Temperatur.

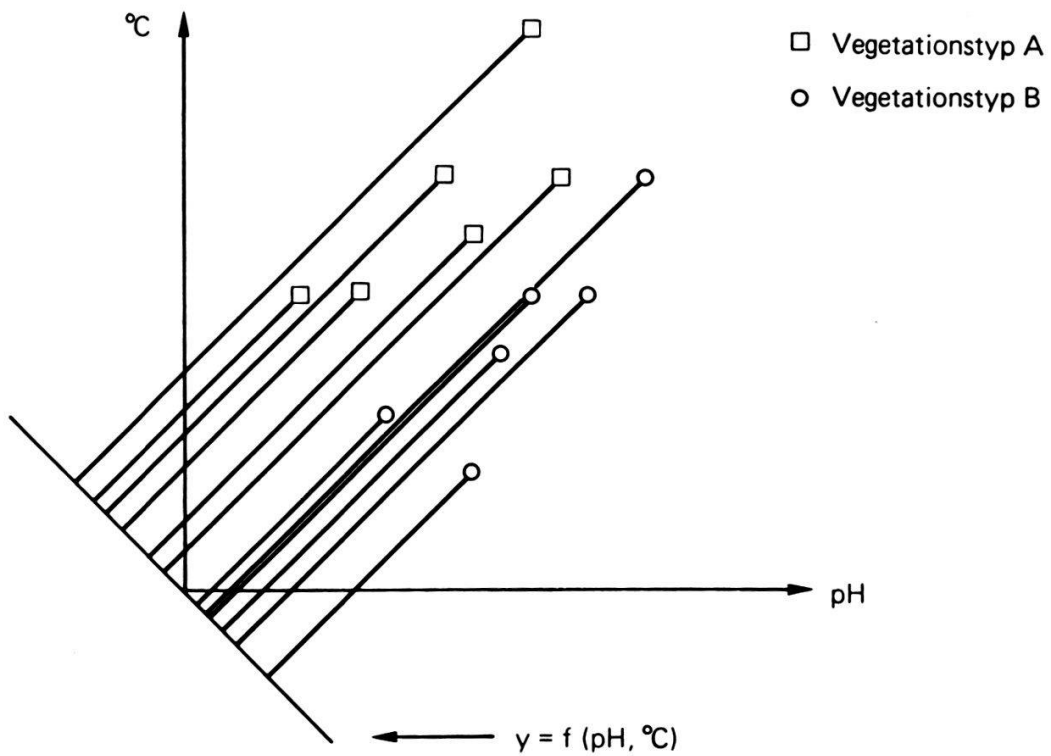


Abb. 1.11 Diskriminanzanalytische Trennung eines Punkteschwarmes mit Hilfe zweier Standortsfaktoren.

- Bei tiefer Temperatur und gleichzeitig tiefem pH- Wert lassen sich die Typen A und B nicht auseinanderhalten.
  
- usw.

Die Ableitung prädiktiver Modelle für pflanzenökologische Anwendungen ist eine recht schwierige Angelegenheit. Wie in Abb. 1.9 dargestellt, sind die wenigsten Strukturen linear. Im Falle der zweifachen Regression (Abb. 1.10) ist somit zu erwarten, dass die beste Lösung eine irgendwie gekrümmte Fläche und nicht eine Ebene sein wird. Werden zur Modellbildung eine grössere Zahl von Vegetationstypen verwendet, so bedarf die Darstellung der diskriminanzanalytischen Lösung einer Vielzahl von Dimensionen. Diese lassen sich nur schwer gleichzeitig veranschaulichen. Auch verbal ist eine solche Lösung schwierig zu fassen. Schliesslich werden in der Regel sowohl die Vegetation als auch der Standort durch mehrere Merkmale beschrieben (vgl. GAUCH 1982, S. 38). Es müssen also Modelle gefunden werden, die Beziehungen zwischen zwei multivariaten Datensätzen zu untersuchen erlauben. Dieses eigentliche Kernproblem der Vegetationkunde wird in Kapitel 9 näher erörtert.