

# Zur Verwendung dieses Buches

<b>1</b>	<b>Zielsetzung des Buches . . . . .</b>	<b>8</b>
<b>2</b>	<b>Strukturierung der fortgeschrittenen multivariaten Analyse-</b> <b>methoden . . . . .</b>	<b>11</b>
2.1	Strukturen-prüfende Verfahren . . . . .	11
2.2	Strukturen-entdeckende Verfahren . . . . .	14
2.3	Zusammenfassende Betrachtung . . . . .	15
<b>3</b>	<b>Die verwendeten Programmsysteme . . . . .</b>	<b>16</b>
	<b>Literaturhinweise . . . . .</b>	<b>19</b>

## 1 Zielsetzung des Buches

### Einsatzfeld

Einhergehend mit der breiten Verfügbarkeit von Softwareprogrammen für multivariate Analysemethoden hat sich auch das Einsatzfeld dieser Verfahren sowohl in der Wissenschaft als auch in der Praxis enorm verbreitet. Das gilt nicht nur für die Häufigkeit, mit der multivariate Verfahren angewendet werden, sondern gleichzeitig auch für die Vielzahl der unterschiedlichsten Methoden. Darüber hinaus bilden die multivariaten Analysemethoden, die sich immer noch in einer „stürmischen“ Entwicklung befinden, heute eines der Fundamente der empirischen Forschung in den Realwissenschaften. Allerdings ist damit auch die Einschätzung, welche multivariaten Analysemethoden denn in der Ausbildung an Hochschulen eine besonders hohe Verbreitung gefunden haben, nicht einfach.

### Veränderung der Hochschullandschaft

Gleichzeitig hat in den letzten Jahren das Bildungssystem der Hochschulen mit der Einführung von Bachelor- und Master-Studiengängen eine grundlegende Veränderung erfahren, so dass es auch hier schwer fällt, eine allgemeine Einschätzung vorzunehmen, welche multivariaten Verfahren eher in der Bachelor-Ausbildung und welche eher in der Master- und/oder Doktoranden-Ausbildung eingesetzt werden. Wird allerdings der Überlegung gefolgt, dass in einem Bachelor-Studium eher die grundlegenden Varianten multivariater Verfahren behandelt werden und erst in der Master- und Doktoranden-Ausbildung Verfahrensvarianten und komplexere Verfahren zur Anwendung gelangen, so kann eine erste Annäherung an die Differenzierung von multivariaten Analysemethoden gefunden werden. Obwohl eine Einteilung der Vielzahl an multivariaten Verfahren in „Grundlegende Verfahren“ und „Fortgeschrittene Verfahren“ weder leicht noch eindeutig ist, haben die Autoren dieses Buches eine solche Unterscheidung wie folgt vorgenommen:

### Grundlegende Verfahren

#### (1) Multivariate Analysemethoden: Grundlegende Verfahren

- Regressionsanalyse (Lineare Einfachregression und multiple Regression)
- Zeitreihenanalyse
- Varianzanalyse
- Diskriminanzanalyse
- Logistische Regression
- Kreuztabellierung und Kontingenzanalyse
- Faktorenanalyse
- Clusteranalyse
- (Traditionelle) Conjoint-Analyse

Die hier als „Grundlegende Verfahren“ bezeichneten multivariaten Analysemethoden werden behandelt in dem Buch:

**Backhaus, Klaus/Erchson, Bernd/Plinke, Wulff/Weiber, Rolf:**  
Multivariate Analysemethoden. Eine anwendungsorientierte Einführung, 14. Aufl., Berlin 2015.

## (2) Fortgeschrittene Verfahren der multivariaten Analyse

Fortgeschrittene  
Verfahren

- Nichtlineare Regressionsanalyse
- Strukturgleichungsanalyse
- Konfirmatorische Faktorenanalyse
- Auswahlbasierte Conjoint-Analyse
- Neuronale Netze
- Multidimensionale Skalierung
- Korrespondenzanalyse

Das vorliegende Buch widmet sich ausschließlich den „Fortgeschrittenen Verfahren“ und verfolgt dabei das Ziel, auch hier eine anwendungsorientierte Einführung in diese Verfahren für den Anwender zu geben. Trotz umfangreicher und auch weitgehend benutzerfreundlicher Softwarelösungen ist der Zugang zu den fortgeschrittenen Verfahren der multivariaten Datenanalyse für den Einsteiger i. d. R. nicht einfach, was sich häufig begründet in

- Vorbehalten gegenüber den mathematischen Darstellungen,
- einer gewissen Scheu vor dem Einsatz des Computers und
- mangelnder Kenntnis der Methoden und ihrer Anwendungsmöglichkeiten.

Weiterhin ist eine Kluft zwischen interessierten Fachleuten und Methodenexperten festzustellen, die bisher nicht genügend durch das Angebot der Fachliteratur überbrückt wird. Die Autoren dieses Buches haben sich deshalb das Ziel gesetzt, zur Überwindung dieser Kluft beizutragen. Daraus ist ein Text entstanden, der folgende Charakteristika besonders herausstellt:

1. Es ist größte Sorgfalt darauf verwendet worden, die Methoden *allgemeinverständlich* darzustellen. Der Zugang zum Verständnis durch den mathematisch ungeschulten Leser hat in allen Kapiteln Vorrang gegenüber dem methodischen Detail. Dennoch wird der rechnerische Gehalt der Methoden in den wesentlichen Grundzügen erklärt, damit sich der Leser, der sich in die Methoden einarbeitet, eine Vorstellung von der Funktionsweise, den Möglichkeiten und Grenzen der Methoden verschaffen kann.
2. Das Verständnis wird erleichtert durch die ausführliche Darstellung von *Beispielen*, die es erlauben, die Vorgehensweise der Methoden leicht nachzuvollziehen und zu verstehen.
3. Darüber hinaus wurde – soweit die Methoden das zulassen – ein Beispiel durchgehend für mehrere Methoden benutzt, um das Einarbeiten zu erleichtern und um die Ergebnisse der Methoden besser vergleichen und deren unterschiedliche Fragstellungen leichter unterscheiden zu können. Die Rohdaten der Beispiele können über den Bestellschein am Anfang des Buches oder über die Internetadresse [www.multivariate.de](http://www.multivariate.de) angefordert werden. Mit Ausnahme der Nichtlinearen Regression, bei der ein Beispiel aus dem Mobilfunk verwendet wurde, greifen alle anderen Verfahren im ausführlichen Fallbeispiel auf die Situation eines Margarineherstellers zurück, die auch für die im ersten Band behandelten Grundlegenden Verfahren betrachtet wird. Da wohl die meisten Leserinnen und

Allgemein-  
verständlich

Beispiel

Rohdaten

Leser über Erfahrungen sowohl mit dem Mobilfunk als auch mit dem Margarinekauf verfügen dürften, erhoffen sich die Autoren, dass die Beispiele – ohne besondere Kenntnis des Anwendungsfeldes „Margarinevermarktung“ – leicht zu verstehen sind und dann auch problemlos auf die spezifische Fragestellung der eigenen Anwendungsfelder übertragen werden können.

Software-  
unterstützung

SPSS

4. Der Umfang des zu verarbeitenden Datenmaterials ist in aller Regel so groß, dass die Rechenprozeduren der einzelnen Verfahren mit vertretbarem Aufwand nur computergestützt durchgeführt werden können. Deshalb erstreckt sich die Darstellung der Methoden sowohl auf die Grundkonzepte der Methoden als auch auf die *Nutzung geeigneter Computer-Programme* als Arbeitshilfe. Es existiert heute eine Reihe von Programmpaketen, die die Anwendung multivariater Analysemethoden nicht nur dem Computer-Spezialisten erlauben. Insbesondere bedingt durch die zunehmende Verbreitung und Leistungsfähigkeit des PCs sowie die komfortablere Gestaltung von Benutzeroberflächen wird auch die Nutzung der Programme zunehmend erleichtert. Damit wird der Fachmann für das Sachproblem unabhängig vom Computer-Spezialisten. Das Programmpaket bzw. Programmsystem, mit dem die meisten Beispiele durchgerechnet werden, ist *IBM SPSS Statistics* (ursprünglich: Statistical Package for the Social Sciences).<sup>1</sup> Als Programmsystem wird dabei eine Sammlung von Programmen mit einer gemeinsamen Benutzeroberfläche bezeichnet. SPSS Statistics hat besonders im Hochschulbereich, aber auch in der Praxis eine sehr weite Verbreitung gefunden. Es ist unter vielen Betriebssystemen auf Großrechnern, Workstations und PC verfügbar.

Arbeitsbuch

5. Das vorliegende Buch hat den Charakter eines *Arbeitsbuches*. Die Darstellungen sind so gewählt, dass der Leser in jedem Fall alle Schritte der Lösungsfindung nachvollziehen kann. Die Syntaxkommandos für die SPSS Prozeduren werden im einzelnen aufgeführt, so dass der Leser durch eigenes Probieren sehr schnell erkennen kann, wie leicht letztlich der Zugang zur Anwendung der Methoden unter Einsatz des Computers ist. Dabei kann er seine eigenen Ergebnisse gegen die im vorliegenden Buch ausgewiesenen kontrollieren. Alle Ausgangsdaten, die den Beispielen zugrunde liegen und ggf. die zugehörigen Syntaxdateien zu den einzelnen Analysemethoden, können für die umfangreicheren Fallbeispiele über [www.multivariate.de](http://www.multivariate.de) bestellt werden.

Interpretation

6. Die Ergebnisse der computergestützten Rechnungen in den einzelnen Methoden werden jeweils anhand der betreffenden *Programmausdrucke* erläutert und kommentiert. Dadurch kann der Leser, der sich in die Handhabung der Methoden einarbeitet, schnell in den eigenen Ergebnissen eine Orientierung finden.
7. Besonderes Gewicht wurde auf die *inhaltliche Interpretation* der Ergebnisse der einzelnen Verfahren gelegt. Die Autoren haben es sich zur Aufgabe gemacht, die Ansatzpunkte für *Ergebnismanipulationen* in den Verfahren offen zu legen und die Gestaltungsspielräume aufzuzeigen, damit der Anwender der Methoden objektive und subjektive Bestimmungsfaktoren der Ergebnisse unterscheiden kann.

---

<sup>1</sup>Zeitweilig wurde SPSS auch interpretiert als „Statistical Product and Service Solutions“ oder „Superior Performing Software System“. Im Folgenden verwenden wir für IBM SPSS Statistics weitgehend nur die Kurzform „SPSS“.

Dies macht u. a. erforderlich, dass methodische Details offen gelegt werden. Dabei wird auch deutlich, dass dem Anwender der Methoden eine Verantwortung für seine Interpretation der Ergebnisse zukommt.

Fastet man die genannten Merkmale des Buches zusammen, dann ergibt sich ein Konzept, das geeignet ist, sowohl dem Anfänger, der sich in die Handhabung der Methoden einarbeitet, als auch demjenigen, der mit den Ergebnissen dieser Methoden arbeiten muss, die erforderliche Hilfe zu geben. Die Konzeption lässt es dabei zu, dass *jede dargestellte Methode für sich verständlich* ist. Der Leser ist also an keine Reihenfolge der Kapitel gebunden.

## 2 Strukturierung der fortgeschrittenen multivariaten Analysemethoden

Wird eine Einteilung der fortgeschrittenen multivariaten Verfahren nach anwendungsbezogenen Fragestellungen angestrebt, so bietet es sich an, eine Unterscheidung in primär strukturen-prüfende Verfahren und primär strukturen-entdeckende Verfahren vorzunehmen:

1. *Strukturen-prüfende Verfahren* sind solche multivariaten Verfahren, deren primäres Ziel in der *Überprüfung von Zusammenhängen* zwischen Variablen liegt. Der Anwender besitzt eine auf sachlogischen oder theoretischen Überlegungen basierende Vorstellung über die Zusammenhänge zwischen Variablen und möchte diese mit Hilfe multivariater Verfahren überprüfen. Verfahren, die diesem Bereich der multivariaten Datenanalyse zugeordnet werden können, sind die nicht-lineare Regressionsanalyse, die konfirmatorische Faktorenanalyse, die Strukturgleichungsanalyse und die auswahlbasierte Conjoint-Analyse.
2. *Strukturen-entdeckende Verfahren* sind solche multivariaten Verfahren, deren Ziel in der *Entdeckung von Zusammenhängen* zwischen Variablen oder zwischen Objekten liegt. Der Anwender besitzt zu Beginn der Analyse noch keine Vorstellungen darüber, welche Beziehungszusammenhänge in einem Datensatz existieren. Verfahren, die primär zur Aufdeckung möglicher Beziehungszusammenhänge eingesetzt werden, sind die Multidimensionale Skalierung, die Korrespondenzanalyse und die Neuronalen Netze.

Strukturen-prüfende  
Verfahren

Strukturen-  
entdeckende  
Verfahren

### 2.1 Strukturen-prüfende Verfahren

Die strukturen-prüfenden Verfahren werden primär zur Durchführung von *Kausalanalysen* eingesetzt, um herauszufinden, ob und wie stark sich z. B. das Wetter, die Bodenbeschaffenheit sowie unterschiedliche Düngemittel und -mengen auf den Ernteertrag auswirken oder wie stark die Nachfrage eines Produktes von dessen Qualität, dem Preis, der Werbung und dem Einkommen der Konsumenten abhängt. Voraussetzung für die Anwendung der entsprechenden Verfahren ist, dass der Anwender *a priori (vorab)* eine sachlogisch möglichst gut fundierte Vorstellung über den Kausalzusammenhang zwischen den Variablen entwickelt hat, d. h. er weiß bereits oder vermutet, welche der Variablen auf andere Variablen einwirken. Zur Überprüfung seiner (theoretischen) Vorstellungen werden die von ihm betrachteten Variablen i. d. R. in *abhängige* und *unabhängige Variablen* eingeteilt und dann

Kausalanalyse

Hypothesen

mit Hilfe von multivariaten Analysemethoden an den empirisch erhobenen Daten überprüft.

### Nichtlineare Regression

#### Schwierigkeiten

Durch die Nichtlineare Regression wird das Anwendungsspektrum der Regressionsanalyse erheblich erweitert. Es lassen sich nahezu beliebige Modellstrukturen schätzen. Anwendungen finden sich z. B. im Rahmen der Werbewirkungsforschung (Abhängigkeit der Werbeerinnerung von der Zahl der Werbekontakte, Abhängigkeit der Absatzmenge von der Höhe des Werbebudgets oder in der Marktforschung bei der Untersuchung des Wachstums von neuen Produkten). Die Nichtlineare Regression ist allerdings mit einer Reihe von Schwierigkeiten verbunden. Der Rechenaufwand ist um ein Vielfaches größer als bei der traditionellen Regressionsanalyse, da iterative Algorithmen für die Berechnung der Schätzwerte verwendet werden müssen. Ob diese Algorithmen konvergieren, hängt u. a. davon ab, welche Startwerte der Untersucher vorgibt. Es werden somit auch erhöhte Anforderungen an den Untersucher gestellt. Ein weiterer Nachteil ist, dass die statistischen Tests, die bei der linearen Regressionsanalyse zur Prüfung der Güte des Modells oder der Signifikanz der Parameter verwendet werden, für die nichtlineare Regression nicht anwendbar sind. Der Untersucher sollte daher, wenn möglich, der linearen Regressionsanalyse den Vorzug geben. Wie gezeigt werden wird, lassen sich auch mit Hilfe der linearen Regressionsanalyse vielfältige nichtlineare Problemstellungen behandeln.

### Strukturgleichungsanalyse

#### Latente Variablen

Bei einer Vielzahl von Kausalbetrachtungen werden Zusammenhänge zwischen Variablen vermutet, die sich einer direkten empirischen Beobachtbarkeit entziehen. Solche Variablen werden auch als hypothetische Konstrukte oder latente Variablen bezeichnet. Beispiele hierfür sind etwa psychologische Konstrukte wie Einstellung und Motivation oder soziologische Konstrukte wie Kultur und soziale Schicht. In solchen Fällen kann die Analyse von Strukturgleichungen mit latenten Variablen zur Anwendung kommen. Zur Behandlung von Strukturgleichungsmodellen wird in diesem Buch auf das Programmpaket *AMOS* (Analysis of Moment Structures) zurückgegriffen, das Datenmatrizen aus SPSS analysiert und Ergebnisse mit SPSS austauschen kann. Mit Hilfe von AMOS lassen sich komplexe Kausalstrukturen überprüfen. Insbesondere können Beziehungen mit mehreren abhängigen Variablen, mehrstufigen Kausalbeziehungen und mit nicht beobachtbaren (latenten) Variablen überprüft werden. Der Benutzer muss, wenn er latente Variable in die Betrachtungen einbeziehen will, zwei Modelle spezifizieren:

#### AMOS

#### Messmodell

- Das *Messmodell*, das die Beziehungen zwischen den latenten Variablen und geeigneten Indikatoren vorgibt, mittels derer sich die latenten Variablen indirekt messen lassen.

#### Strukturmodell

- Das *Strukturmodell*, welches die Kausalbeziehungen zwischen den latenten Variablen vorgibt, die letztlich dann zu überprüfen sind.

Die Variablen des Strukturmodells können alle latent sein, müssen es aber nicht. Ein Beispiel, bei dem nur die unabhängigen Variablen latent sind, wäre die Abhängigkeit der Absatzmenge von der subjektiven Produktqualität und Servicequalität eines Anbieters.

### Konfirmatorische Faktorenanalyse

Die Konfirmatorische Faktorenanalyse (KFA) stellt einen Spezialfall eines (vollständigen) Strukturgleichungsmodells mit latenten Variablen dar, da sie „lediglich“ die sachlogisch formulierten Messmodelle von hypothetischen Konstrukten analysiert. Mit Hilfe der KFA werden eine Güteprüfung der Operationalisierung hypothetischer Konstrukte vorgenommen und ggf. auch Abhängigkeiten zwischen mehreren Konstrukten untersucht. Dabei werden immer sog. *reflektive Messmodelle* unterstellt, die ein Konstrukt über empirisch direkt messbare Variablen (sog. Indikatorvariable) operationalisieren. Die Indikatorvariablen müssen so definiert werden, dass ihre Messwerte jeweils beispielhafte Manifestierungen des betrachteten hypothetischen Konstruktes darstellen. Im Gegensatz zur explorativen Faktorenanalyse wird bei der KFA die Faktorenstruktur, d.h. die Zuordnung von Indikatorvariablen zu Faktoren, *vorgegeben* und dann die Stärke des Zusammenhangs durch Schätzung der Faktorladungen überprüft.

Reflektive  
Messmodelle

### Auswahlbasierte Conjoint-Analyse

Während bei der traditionellen Conjoint-Analyse zwecks Analyse von Nutzenstrukturen die Präferenzen von Probanden bezüglich alternativer Objekte (Stimuli) auf ordinalem Skalenniveau gemessen werden (mittels Ranking- oder Ratingskalen), erfolgt bei der Auswahlbasierten Conjoint-Analyse (Choice-Based Conjoint) nur eine Abfrage von Auswahlentscheidungen. Aus einer Menge von Alternativen (Choice Set) muss der Proband nur jeweils die am meisten präferierte Alternative auswählen, wobei meist auch die Option besteht, keine der Alternativen zu wählen. Dies ist für ihn nicht nur einfacher, sondern kommt auch seinem realen Entscheidungsverhalten (z. B. in Kaufsituationen) sehr viel näher, als das Ranking oder Rating aller Alternativen im Choice Set, wie es die klassische Conjoint-Analyse verlangt. Die erhöhte Realitätsnähe wird allerdings mit einem Verlust an Information „erkauft“, da bei dieser Vorgehensweise die Präferenz nur noch auf nominalem Skalenniveau gemessen wird. Zur Schätzung der Nutzenbeiträge einzelner Merkmale (Teilnutzenwerte) muss daher ein anderes Schätzverfahren verwendet werden. Während bei der traditionellen Conjoint-Analyse die Schätzung meist durch Regression mit Dummy-Variablen erfolgt, kommt bei der Auswahlbasierten Conjoint-Analyse die Maximum-Likelihood-Methode zur Anwendung. Dabei wird dem Verhalten der Probanden ein probabilistisches Entscheidungsmodell zugrunde gelegt. Wegen des geringeren Informationsgehalts ist es meist nur möglich, die Teilnutzenwerte aggregiert zu schätzen, während es bei der traditionellen Conjoint-Analyse üblich ist, sie individuell für jeden Probanden zu schätzen. Allerdings können auch im Rahmen der auswahlbasierten Conjoint-Analyse mit Hilfe des Latent Class-Ansatzes zielgruppenspezifische und mit Hilfe des sog. Hierarchical Bayes-Ansatzes individuelle Teilnutzenwerte geschätzt werden.

CBC

Kaufsimulation

## 2.2 Strukturen-entdeckende Verfahren

Die hier den strukturen-entdeckenden Verfahren zugeordneten Analysemethoden werden primär zur *Entdeckung von Zusammenhängen* zwischen Variablen oder zwischen Objekten eingesetzt. Es erfolgt daher vorab durch den Anwender *keine* Zuteilung der Variablen in abhängige und unabhängige Variablen, wie es bei den strukturen-prüfenden Verfahren der Fall ist.

### Neuronale Netze

Substitut

Methodik

Nicht-Linearität

Lernprozesse

Neuronale Netze werden heute in der Praxis in zunehmendem Maße sowohl ergänzend zu den klassischen multivariaten Methoden eingesetzt, als auch in den Fällen, in denen die klassischen Methoden versagen. Anwendungsgebiete sind Klassifikationen von Objekten, Prognosen von Zuständen oder Probleme der Gruppenbildung. Insofern bestehen hinsichtlich der Aufgabenstellungen Ähnlichkeiten zur Diskriminanzanalyse und zur Clusteranalyse. Die Methodik neuronaler Netze lehnt sich an biologische Informationsverarbeitungsprozesse im Gehirn an (daher der Name). Es werden künstliche neuronale Netze gebildet, die in der Lage sind, selbständig aus Erfahrung zu lernen. Insbesondere vermögen sie, komplexe Muster in vorhandenen Daten (z. B. Finanzdaten, Verkaufsdaten) zu erkennen und eröffnen so eine sehr einfache Form der Datenanalyse. Besonders vorteilhaft lassen sie sich zur Behandlung von schlecht strukturierten Problemstellungen einsetzen. Innerhalb neuronaler Netze werden künstliche Neuronen (Nervenzellen) als Grundelemente der Informationsverarbeitung in Schichten organisiert, wobei jedes Neuron mit denen der nachgelagerten Schicht verbunden ist. Dadurch lassen sich auch hochgradig nicht-lineare und komplexe Zusammenhänge ohne spezifisches Vorwissen über die etwaige Richtung und das Ausmaß der Wirkungsbeziehungen zwischen einer Vielzahl von Variablen modellieren. Zum Erlernen von Strukturen wird das Netz zunächst in einer sog. *Trainingsphase* mit beobachteten Daten „gefüttert“. Dabei wird unterschieden zwischen Lernprozessen, bei denen die richtigen Ergebnisse bekannt sind und diese durch das Netz reproduziert werden sollen (*überwachtes Lernen*), und solchen, bei denen die richtigen Ergebnisse nicht bekannt sind und lediglich ein konsistentes Verarbeitungsmuster erzeugt werden soll (*unüberwachtes Lernen*). Nach der Trainingsphase ist das Netz konfiguriert und kann für die Analyse neuer Daten eingesetzt werden.

### Multidimensionale Skalierung

Positionierung

Ähnlichkeiten

Den Hauptanwendungsbereich der Multidimensionalen Skalierung (MDS) bilden Positionierungsanalysen, d. h. die *Positionierung von Objekten im Wahrnehmungsraum* von Personen. Sie bildet somit eine Alternative zur faktoriellen Positionierung mit Hilfe der Faktorenanalyse. Im Unterschied zur faktoriellen Positionierung werden bei Anwendung der MDS nicht die subjektiven Beurteilungen von Eigenschaften der untersuchten Objekte erhoben, sondern es werden nur wahrgenommene globale Ähnlichkeiten zwischen den Objekten erfragt. Mittels der MDS werden die diesen Ähnlichkeiten zugrundeliegenden Wahrnehmungsdimensionen abgeleitet. Wie schon bei der faktoriellen Positionierung lassen sich sodann die Objekte im Raum dieser Dimensionen positionieren und grafisch darstellen. Die MDS findet insbesondere dann Anwendung, wenn der Forscher keine oder nur vage Kenntnisse darüber hat, welche



Eigenschaften für die subjektive Beurteilung von Objekten (z. B. Produktmarken, Unternehmen oder Politiker) von Relevanz sind. Zwischen der Multidimensionalen Skalierung und der Conjoint-Analyse besteht sowohl inhaltlich wie auch methodisch eine enge Beziehung, obgleich wir sie hier unterschiedlich zum einen den strukturen-entdeckenden und zum anderen den strukturen-prüfenden Verfahren zugeordnet haben. Beide Verfahren befassen sich mit der Analyse psychischer Sachverhalte und bei beiden Verfahren können auch ordinale Daten analysiert werden, weshalb sie z.T. auch identische Algorithmen verwenden. Ein gewichtiger Unterschied besteht jedoch darin, dass der Forscher bei Anwendung der Conjoint-Analyse bestimmte Merkmale auszuwählen hat.

Beziehung zu  
anderen Verfahren

### Korrespondenzanalyse

Die Korrespondenzanalyse dient, ebenso wie die Multidimensionale Skalierung (MDS), zur Visualisierung komplexer Daten. Sie wird daher in der Marktforschung ebenfalls zur Durchführung von Positionierungsanalysen verwendet. Insbesondere kann sie als ein Verfahren der multidimensionalen Skalierung von nominal skalierten Variablen charakterisiert werden. Sie ermöglicht es, die Zeilen und Spalten einer zweidimensionalen Kreuztabelle (Kontingenztafel) grafisch in einem gemeinsamen Raum darzustellen.

Visualisierung

*Beispiel:* : Gegeben sei eine Häufigkeitstabelle, deren Zeilen Automarken betreffen und in deren Spalten wünschenswerte Merkmale von Autos (z. B. hohe Sicherheit, schönes Design) stehen. Die Zellen der Matrix sollen beinhalten, mit welcher Häufigkeit ein bestimmtes qualitatives Merkmal den verschiedenen Automarken im Rahmen einer Käuferbefragung zugeordnet wurde. Marken und Merkmale lassen sich sodann mit Hilfe der Korrespondenzanalyse in einem gemeinsamen Raum als Punkte darstellen. Dadurch lässt sich dann erkennen, wie die Automarken relativ zueinander und in Bezug auf die Merkmale von den Käufern beurteilt werden. Für die Korrespondenzanalyse spielt es dabei keine Rolle (im Unterschied zur explorativen Faktorenanalyse), welche Elemente in den Zeilen und welche in den Spalten angeordnet werden.

Beispiel

Ein besonderer Vorteil der Korrespondenzanalyse liegt darin, dass sie kaum Ansprüche an das Skalenniveau der Daten stellt. Die Daten müssen lediglich nicht-negativ sein. Die Korrespondenzanalyse kann daher auch zur Quantifizierung qualitativer Daten verwendet werden. Da sich qualitative Daten leichter erheben lassen als quantitative Daten, kommt diesem Verfahren eine erhebliche praktische Bedeutung zu.

Vorteil

### 2.3 Zusammenfassende Betrachtung

Die vorgenommene Zweiteilung der multivariaten Verfahren in strukturen-prüfende und strukturen-entdeckende Verfahren kann keinen Anspruch auf Allgemeingültigkeit erheben, sondern kennzeichnet nur den vorwiegenden Einsatzbereich der Verfahren. So können und werden z. B. Strukturgleichungsanalysen oder Konfirmatorische Faktorenanalysen zur Identifikation empirisch relevanter Kausalpfade oder die Eliminierung sachlogisch vermuteter Kausalstrukturen herangezogen, wobei leider viel zu häufig dabei vergessen wird, dass sich damit der konfirmatorische Charakter dieser Verfahren in einen explorativen verwandelt. Darüber hinaus kann der „gedankenlose Einsatz“ von multivariaten Verfahren leicht zu einer Quelle von Fehlinterpretationen

Fehlinterpretation

werden, da ein statistisch signifikanter Zusammenhang noch keine hinreichende Bedingung für das Vorliegen eines kausal bedingten Zusammenhangs bildet. („Erst denken, dann rechnen!“) Es sei daher generell empfohlen, die strukturen-prüfenden Verfahren auch in diesem Sinne, d. h. zur empirischen Überprüfung von theoretisch oder sachlogisch begründeten Hypothesen, einzusetzen. In Abbildung 1 sind die oben skizzierten multivariaten Verfahren noch einmal mit jeweils einer typischen Anwendungsfrage zusammengefasst.

Verfahren	Beispielhafte Fragestellungen
Nichtlineare Regression	Untersuchung des Wachstums von neuen Produkten, der Diffusion von Innovationen oder der Ausbreitung von Epidemien.
Strukturgleichungsanalyse	Abhängigkeit der Käufertreue von der subjektiven Produktqualität und Servicequalität eines Anbieters
Konfirmatorische Faktorenanalyse	Überprüfung der Eignung vorgegebener Indikatorvariablen für die Messung hypothetischer Konstrukte wie z. B. Loyalität, Vertrauen oder Reputation.
Auswahlbasierte Conjoint-Analyse	Schätzung der Nutzenbeiträge einzelner Merkmale von Produkten zur Gesamtpreferenz auf Basis simulierter Kaufentscheidungen.
Neuronale Netze	Untersuchung von Aktienkursen und möglichen Einflussfaktoren zwecks Prognose von Kursentwicklungen
Multidimensionale Skalierung	Positionierung von konkurrierenden Produktmarken im Wahrnehmungsraum der Konsumenten
Korrespondenzanalyse	Darstellung von Produktmarken und Produktmerkmalen in einem gemeinsamen Raum.

Abbildung 1: Synopsis der fortgeschrittenen multivariaten Analyseverfahren

### 3 Die verwendeten Programmsysteme

SPSS

Zur rechnerischen Durchführung der Analysen, die in diesem Buch behandelt werden, wurde vornehmlich das Programmsystem IBM SPSS Statistics verwendet, da dieses in Wissenschaft und Praxis eine besonders große Verbreitung gefunden hat.<sup>2</sup> Der Name „SPSS“ stand ursprünglich als Akronym für „Statistical Package for the Social Sciences“. Der Anwendungsbereich von SPSS wurde im Laufe der Zeit ständig erweitert und erstreckt sich inzwischen auf nahezu alle Bereiche der Datenanalyse. „SPSS“ steht daher heute als Markenname, der nicht weiter interpretiert wird. Während SPSS ursprünglich nur über eine Kommandodatei gesteuert werden konnte (Batch-Betrieb), besitzt es heute eine grafische Benutzeroberfläche, über die der Benutzer mit dem Programm kommunizieren kann (Dialog-Betrieb). Diese Benutzeroberfläche wird ständig verbessert und erweitert. Über die dort vorhandenen Menüs und Dialogfelder lassen sich auch komplexe Analysen sehr bequem durchführen. Die früher zur Steuerung des Programms benötigte Kommandosprache (Befehlssyntax) findet daher zunehmend weniger Anwendung, ist aber nicht überflüssig geworden. Intern wird sie weiterhin

Grafische  
Benutzeroberfläche

<sup>2</sup>Das Programmsystem IBM SPSS Statistics kann unter den Betriebssystemen Windows, Macintosh und Linux verwendet werden und ist zur Zeit in der Version 22 auf dem Markt. Neben der Vollversion von IBM SPSS Statistics 22 wird zu Lehrzwecken auch eine preiswertere Studentenversion angeboten. Diese weist einige Einschränkungen auf, die aber für die Mehrzahl der Nutzer kaum relevant sein dürfte: Datendateien dürfen maximal 50 Variablen und 1.500 Fälle enthalten und die SPSS-Befehlssyntax (Kommandosprache) sowie die Erweiterungsmodule sind nicht verfügbar.

Fortgeschrittene Multivariate Analysemethoden

Eine anwendungsorientierte Einführung

Backhaus, K.; Erichson, B.; Weiber, R.

2015, IX, 454 S., Softcover

ISBN: 978-3-662-46086-3