

K. Überla

Faktorenanalyse

*Eine systematische Einführung
für Psychologen, Mediziner, Wirtschafts- und Sozial-
wissenschaftler*

Nachdruck der zweiten Auflage

Mit 110 Abbildungen



Springer-Verlag Berlin Heidelberg New York 1977

Prof. Dr. med. Dipl. Psych. K. ÜBERLA

Leiter der Abteilung für Medizinische Statistik, Dokumentation und Datenverarbeitung der Universität Ulm

AMS Subject Classifications (1970):

62 H 30, 62 H 25, 62 P 10, 62 P 15, 62 P 20, 62 P 25, 62 P 99

ISBN-13: 978-3-642-61986-1 e-ISBN-13: 978-3-642-61985-4

DOI: 10.1007/978-3-642-61985-4

Das Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersetzung, des Nachdruckes, der Entnahme von Abbildungen, der Funksendung, der Wiedergabe auf photomechanischem oder ähnlichem Wege und der Speicherung in Datenverarbeitungsanlagen bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Bei Vervielfältigungen für gewerbliche Zwecke ist gemäß § 54 UrhG eine Vergütung an den Verlag zu zahlen, deren Höhe mit dem Verlag zu vereinbaren ist. © by Springer-Verlag Berlin Heidelberg 1968 and 1971. Library of Congress Catalog Card Number 73-157071. Printed in Germany. Druck und Bindearbeiten: Brühlsche Universitätsdruckerei, Gießen.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, daß solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

2144/3140-543210

Vorwort zur zweiten Auflage

Die rasche Verbreitung des Buches zeigt das wachsende Interesse an den Techniken der Faktorenanalyse. Diese zweite Auflage ist im wesentlichen ein unveränderter Abdruck der ersten, wobei freilich die bekannt gewordenen Fehler verbessert wurden. Ich bedanke mich bei allen, die durch Zuschriften, Rezensionen oder mündliche Hinweise zur Fehlerkorrektur beigetragen haben.

Ulm, im Januar 1971

KARL ÜBERLA

Vorwort zur ersten Auflage

Die Faktorenanalyse spielt in vielen psychologischen Untersuchungen eine zentrale Rolle. Sie ermöglicht die Reduktion eines umfangreichen Datenmaterials auf wenige unabhängige und einfache Faktoren. Die Aufgabe, wesentliche Komponenten als erste Hypothesen aus einem großen Material zu gewinnen, tritt in vielen empirischen Wissenschaften auf. Im deutschen Sprachraum fehlte bislang eine zusammenfassende Darstellung des Verfahrens.

Die Faktorenanalyse wurde in den angelsächsischen Ländern entwickelt. In den letzten Jahrzehnten ist eine umfangreiche Literatur entstanden, deren gesicherte Anteile in den deutschen Sprachraum übernommen werden sollten. Um einen Anschluß an die Diskussionen im Ausland zu gewinnen, ist ein breiter gestreutes Wissen über die Faktorenanalyse notwendig. Ein weiteres Motiv für das Erscheinen des Buches in der vorliegenden Form ist die Absicht, die Faktorenanalyse aus ihrer engen Verflechtung mit psychologischen Theorien zu lösen. Sie ist als eine statistische Methodik unabhängig von einem bestimmten Sachgebiet zu sehen.

Das Buch ist aus der Anwendungserfahrung entstanden und entspricht im Aufbau mehr der Denkweise des Praktikers als einer systematischen mathematischen Ableitung. Es ist für Leser aus den Verhaltenswissenschaften geschrieben, die sich ernsthaft mit der Faktorenanalyse auseinandersetzen möchten. Auf Rezepte und Vereinfachungen wurde weitgehend verzichtet. Die Genauigkeit wurde gegenüber der Verständlichkeit bevorzugt, obwohl alle Anstrengungen gemacht wurden, gerade das Verständnis zu fördern.

Der Stoff ist in acht größere Abschnitte aufgeteilt. In den Vorbemerkungen wird in die Fragestellung eingeführt. Außerdem werden die Formeln der Korrelations- und Regressionsrechnung zusammengestellt und eine sehr knappe Einführung in das Umgehen mit Matrizen, Vektoren und Determinanten gegeben. Im zweiten Abschnitt wird eine Übersicht über die Faktorenanalyse von verschiedenen Ansätzen aus versucht. Diese Übersicht soll das Gebiet abstecken und die Probleme formulieren. Die vier Hauptprobleme der Faktorenanalyse werden dann in je einem Abschnitt behandelt, wobei die wichtigsten Lösungsverfahren angegeben sind. Im Abschnitt VII werden Modelluntersuchungen zur Faktorenanalyse dargestellt. Solche Untersuchungen sind eine wichtige Quelle von Informationen über die Gültigkeit und Genauigkeit des Verfahrens. Im letzten Abschnitt werden Einzelprobleme herausgegriffen.

Dem Leser, der noch wenig mit der Faktorenanalyse in Berührung gekommen ist, sei folgender Leseplan empfohlen: Die Kapitel 1.1 und 1.2 sollten gelesen, das Kapitel 1.3 wirklich durchgearbeitet werden, soweit es nicht schon bekannt ist. Das Kapitel 1.4 sollte man zunächst überfliegen und die Definitionen und Rechenregeln hinnehmen. Wenn im Verlauf der weiteren Lektüre Rechenoperationen auftauchen, die in Kapitel 1.4 behandelt wurden, sollte man dorthin zurückgehen und sich genau mit diesen Rechenoperationen beschäftigen. Der Abschnitt II, die Übersicht über die Faktorenanalyse, sollte ganz gelesen werden. Dabei sind die Kapitel 2.1 und 2.2 genau zu studieren. Das Kapitel 2.3 sagt dasselbe wie 2.2, es ist jedoch für manche anschaulicher als die algebraische Darstellung. Die Kapitel 2.4 und 2.5 sind als Ergänzung gedacht. Aus dem Abschnitt III ist für eine Übersicht das Kapitel 3.2 über die Zentroidmethode zu empfehlen, man kann aber auch direkt mit Kapitel 3.1 beginnen. Die Kapitel 3.3.1, 3.3.2, 3.3.5, 3.3.6 sowie 3.6 und 3.7 sollte der Anfänger ebenfalls lesen. Das Kommunalitätenproblem (Abschnitt IV) sollte ganz studiert werden. Ebenso ist der Abschnitt V über die Rotation ganz durchzuarbeiten bis auf die Kapitel 5.7 und 5.8, die speziellere Fragen behandeln. Aus dem Abschnitt VI sind für eine erste Information die einleitenden Bemerkungen sowie die Kapitel 6.2 und 6.3 notwendig. Die Abschnitte VII und VIII können für eine erste

Übersicht ausgeklammert werden – bis auf das Schlußkapitel 8.9. In den beiden letzten Abschnitten sind zahlreiche Informationen enthalten, die das Verständnis der Faktorenanalyse fördern. Derjenige, der schon einige Erfahrung im Umgang mit statistischen Methoden und Matrizenrechnung hat, sollte das Buch ohne Teile auszulassen durcharbeiten. Er kann sich nach dem angegebenen Leseplan zunächst eine erste Übersicht verschaffen.

Ich bin vielen für ihre Hilfe zu Dank verpflichtet. Prof. R. B. CATTELL hat mein Interesse an der Faktorenanalyse begründet. Seine Schule prägte das Konzept des Buches entscheidend mit. Prof. S. KOLLER hat in großzügiger Weise das Entstehen des Manuskripts gefördert. Die Diskussionen mit ihm haben dem Werk eine besondere Richtung gegeben. Sein Streben nach einer verständlichen Darstellung der Grundproblematik und nach einer Systematisierung der Ansatzprobleme hat zahlreiche Teile beeinflusst. Meine Habilitationsschrift „Faktorenanalyse in der Medizin. Beiträge zur Methodik und Probleme der Anwendung“ ist mit in die vorliegende Darstellung eingegangen. Ich danke den Kollegen am Institut für Medizinische Statistik und Dokumentation in Mainz sowie den Kollegen aus anderen Instituten, die mir in vielen Einzeldiskussionen Anregungen vermittelten. Eine Gruppe von Psychologiestudenten aus Mainz hat das Manuskript gelesen und zahlreiche Hinweise gegeben. Herrn PIETSCHMANN und Herrn BREITLAUCH, die die Abbildungen anfertigten, sowie Fr. HENNECKEN und Frau THOMAS, die das Manuskript schrieben, möchte ich für ihre Hilfe danken. Herr KNODT hat das Literaturverzeichnis durchgesehen und zusammen mit Frau HEISER Korrekturen gelesen. Allen anderen, die in verschiedener Weise mitgeholfen haben, sei zusammen Dank gesagt. Der Springer-Verlag ist auf meine Wünsche freundlich eingegangen. Besonders danke ich meiner Frau, die mir mit viel Geduld und Charme bei der Arbeit zur Seite stand.

Alle Leser, die beim sorgfältigen Durchgehen die unvermeidlichen Fehler entdecken, bitte ich, diese mitzuteilen, damit Verbesserungen angebracht werden.

Mainz, im Januar 1968

KARL ÜBERLA

Inhaltsverzeichnis

I. Vorbemerkungen

1.1	Zur Fragestellung der Faktorenanalyse	1
1.2	Die historische Entwicklung	7
1.3	Korrelations- und Regressionsrechnung	10
1.4	Matrizen, Vektoren und Determinanten	22

II. Übersicht über die Faktorenanalyse

2.1	Zwei einführende Beispiele	43
2.2	Grundgleichungen und Grundkonzepte in algebraischer Form	50
2.2.1	Das Fundamentaltheorem	50
2.2.2	Mögliche Arten von Faktoren und ihre Beziehungen	54
2.2.3	Ablaufschema und Hauptprobleme der Faktorenanalyse	62
2.2.4	Veranschaulichung an Hand eines Zahlenbeispiels	63
2.3	Geometrische Darstellung des faktorenanalytischen Modells	68
2.3.1	Geometrie der Datenmatrix und des Testraumes	68
2.3.2	Gemeinsamer Faktorenraum und totaler Faktorenraum	71
2.3.3	Geometrische Veranschaulichung der Faktorenextraktion und Rotation	75
2.4	Partielle Korrelationsrechnung als Ausgangspunkt der Faktorenanalyse	78
2.5	Die Faktorenanalyse innerhalb der multivariaten statistischen Verfahren	81
2.6	Zusammenfassung und weitere Gliederung	89

III. Das Faktorenproblem

3.1	Die Hauptkomponentenmethode	93
3.1.1	Geometrische Veranschaulichung	93
3.1.2	Algebraische Darstellung	98
3.1.3	Rechenverfahren	100
3.1.4	Beispiele	108
3.2	Die Zentroidmethode	113
3.2.1	Ableitung der Zentroidmethode	113
3.2.2	Rechenverfahren	117
3.3	Kriterien zur Bestimmung der Zahl der zu extrahierenden Faktoren	123
3.3.1	Darstellung der Varianzanteile	124
3.3.2	Beurteilung der Residuen	129
3.3.3	Prüfung der Signifikanz von Hauptkomponenten	132
3.3.4	Prüfung der Signifikanz beim faktorenanalytischen Ansatz	133
3.3.5	Weitere Extraktionskriterien	136
3.3.6	Empfehlungen zur Bestimmung der Faktorenzahl	138
3.4	Historische und wenig gebräuchliche Extraktionsmethoden	139
3.4.1	Algebraische Lösung des Faktorenproblems	140
3.4.2	Die Uni-Faktor-Methode	140
3.4.3	Die Two-Faktor-Methode SPEARMANs	142
3.4.4	Die Bi-Faktor-Methode	143
3.4.5	Die Multiple Gruppenmethode	145

3.5	Neuere Lösungsansätze	146
3.5.1	Die Maximum-Likelihood-Schätzung der Faktorenladungen	147
3.5.2	Die Kanonische Faktorenanalyse	148
3.5.3	Die Alpha-Faktorenanalyse	149
3.6	Die Äquivalenz verschiedener Extraktionsverfahren	150
3.7	Empfehlungen zur Lösung des Faktorenproblems	153

IV. Das Kommunalitätenproblem

4.1	Einführung	155
4.2	Rechenverfahren	157
4.2.1	Der höchste Korrelationskoeffizient	158
4.2.2	Das Quadrat des multiplen Korrelationskoeffizienten	158
4.2.3	Die Iteration	159
4.2.4	Andere Techniken	160
4.3	Vergleiche an einem Beispiel	162

V. Das Rotationsproblem

5.1	Orthogonale und schiefwinkelige Rotation im gemeinsamen Faktorenraum	167
5.2	Der Begriff der Einfachstruktur	175
5.3	Die iterative Rotation zur Einfachstruktur in jeweils einer Ebene	188
5.4	Die Matrizen des Rotationsproblems und ihre Beziehungen	204
5.5	Die analytische Rotation zur Einfachstruktur	207
5.5.1	Orthogonale Verfahren	208
5.5.2	Schiefwinkelige Verfahren	211
5.6	Ein Beispiel für die Rotation	214
5.7	Andere Rotationsmethoden und Rotationskriterien	215
5.8	Faktoren zweiter und höherer Ordnung	227
5.9	Abschließende Bemerkungen	232

VI. Die Bestimmung von Faktorenwerten

6.1	Die Berechnung von Hauptkomponentenwerten	237
6.2	Die Schätzung von Faktorenwerten durch multiple Regressionsrechnung	241
6.3	Die multiple Korrelation zwischen Variablen und Faktoren	248
6.4	Andere Methoden der Schätzung von Faktorenwerten	251

VII. Die Überprüfung der Faktorenanalyse in überschaubaren Situationen

7.1	Zwei Beispiele aus der Literatur	256
7.1.1	Das Box-Problem THURSTONEs	256
7.1.2	Das Ball-Problem CATTELLs und DICKMANs	260
7.2	Weitere Modellrechnungen	264
7.2.1	Ein Beispiel bei Blutdruckmessungen	264
7.2.2	Ein Beispiel mit Einwohnerzahlen	269
7.3	Simulationen auf einer elektronischen Rechanlage	272
7.3.1	Ansatz und Ergebnisse bei Modellen mit gleichen Faktorenladungen	273
7.3.2	Die Genauigkeit der Faktorenanalyse bei ungleichen Faktorenladungen	282
7.3.3	Die Genauigkeit der Faktorenanalyse beim Vorliegen von Alternativdaten	284
7.3.4	Ein Vergleich der Schätzgenauigkeit der Faktorenanalyse mit der direkten Schätzung in der multiplen Regressionsrechnung	288
7.4	Aussagen und Grenzen von Modelluntersuchungen	292

VIII. Einzelprobleme

8.1	Unterscheidung verschiedener Techniken nach Art der Datenmatrix	296
8.2	Qualitative Daten in der Faktorenanalyse	301
8.3	Transformationen der Datenmatrix	303
8.4	Hinweise auf Verfahren, die mit der Faktorenanalyse verwandt sind	306
8.4.1	Cluster-Analyse	307
8.4.2	Image-Analyse	309
8.4.3	Analyse der latenten Struktur	309
8.5	Das Problem der Heterogenität	310
8.6	Das Problem der Klassifizierung	317
8.6.1	<i>Q</i> -Technik zum Auffinden von Gruppen	319
8.6.2	Andere Verfahren zur Klassifizierung	319
8.7	Die Prüfung der Genauigkeit des faktorenanalytischen Modells an der Datenmatrix	321
8.8	Elektronische Datenverarbeitung und Faktorenanalyse	330
8.8.1	Aufbau und Funktionsweise elektronischer Datenverarbeitungsanlagen	331
8.8.2	Auswirkungen auf die Faktorenanalyse	343
8.8.3	Programmbibliothek	345
8.8.4	Das Rotoplot-Programm als ein Beispiel	346
8.9	Ansatzpunkte für eine Systematisierung der Anwendungen der Faktorenanalyse	355

Anhang

Tafel A:	Signifikanz von Korrelationskoeffizienten	366
Tafel B:	Signifikanz von multiplen Korrelationskoeffizienten	368
Tafel C:	Z-Transformation	370
Tafel D:	Signifikanz der Einfachstruktur nach BARGMANN	373
Literaturverzeichnis		376
Namen- und Sachverzeichnis		393

Bezeichnung der wichtigsten Matrizen

$m = \text{Anzahl der Variablen}, n = \text{Anzahl der F\u00e4lle}, r = \text{Anzahl der Faktoren}$

Matrix	Spalten Zeilen	Bedeutung	Erkl\u00e4rt auf Seite
$\mathbf{Y} = (y_{ij})$	$i = 1, \dots, m$ $j = 1, \dots, n$	Matrix des Ausgangsdaten	50
$\mathbf{Z} = (z_{ij})$	$i = 1, \dots, m$ $j = 1, \dots, n$	Matrix der standardisierten Ausgangsdaten	51
$\mathbf{S} = (s_{ik})$	$i = 1, \dots, m$ $k = 1, \dots, m$	Kovarianzmatrix	51
$\mathbf{R} = (r_{ik})$	$i = 1, \dots, m$ $k = 1, \dots, m$	Korrelationsmatrix	51
$\mathbf{R}_h = (r_{ik}^h)$	$i = 1, \dots, m$ $k = 1, \dots, m$	Reduzierte Korrelationsmatrix	60, 61
$\mathbf{R}^+ = (r_{ik}^+)$	$i = 1, \dots, m$ $k = 1, \dots, m$	Reproduzierte Korrelationsmatrix	44, 101, 116
$\mathbf{R}_l = (r_{ik}^l)$	$i = 1, \dots, m$ $k = 1, \dots, m$	Residualmatrix nach Extraktion von l Faktoren	101, 116
$\mathbf{M} = (m_{ik})$	$i = 1, \dots, m$ $k = 1, \dots, m$	Diagonale Matrix der Eigenwerte von \mathbf{R}	238
$\mathbf{A} = (a_{il})$	$i = 1, \dots, m$ $l = 1, \dots, r$	Unrotiertes orthogonales Faktorenmuster	48, 52
$\mathbf{U}^2 = (u_{ik}^2)$	$i = 1, \dots, m$ $k = 1, \dots, m$	Diagonale Matrix mit Einzelvarianzen	60
$\mathbf{V}_{fp} = (v_{if}^p)$	$i = 1, \dots, m$ $l = 1, \dots, r$	Faktorenmuster	} der Prim\u00e4rfaktoren 204 ff.
$\mathbf{V}_{fs} = (v_{if}^s)$	$i = 1, \dots, m$ $l = 1, \dots, r$	Faktorenstruktur	
$\mathbf{V}_{rp} = (v_{if}^p)$	$i = 1, \dots, m$ $l = 1, \dots, r$	Faktorenmuster	} der Reference- Vektoren 204 ff.
$\mathbf{V}_{rs} = (v_{if}^s)$	$i = 1, \dots, m$ $l = 1, \dots, r$	Faktorenstruktur	
$\mathbf{T} = (t_{ip})$	$l = 1, \dots, r$ $p = 1, \dots, r$	Transformationsmatrix zu Prim\u00e4rfaktoren	204 ff.
$\mathbf{\Lambda} = (\lambda_{ip})$	$l = 1, \dots, r$ $p = 1, \dots, r$	Transformationsmatrix zu Referencevektoren	190, 197 ff., 204 ff.
$\mathbf{C}_f = (c_{fp}^f)$	$l = 1, \dots, r$ $p = 1, \dots, r$	Korrelationsmatrix zwischen Prim\u00e4rfaktoren	61, 204 ff.

Matrix	Spalten Zeilen	Bedeutung	Erklärt auf Seite
$\mathbf{C}_r = (c_{lp}^r)$	$l = 1, \dots r$ $p = 1, \dots r$	Korrelationsmatrix zwischen Reference-Vektoren	190, 197 ff., 204 ff.
$\mathbf{D} = (d_{lp})$	$l = 1, \dots r$ $p = 1, \dots r$	Diagonale Hilfsmatrix mit Korrelationen zwischen Primärfaktoren und Reference-Vektoren	206
$\mathbf{P} = (p_{ij})$	$i = 1, \dots m$ $j = 1, \dots n$	Faktorenwerte allgemein	52, 53
$\mathbf{P}' = (p'_{ij})$	$i = 1, \dots m$ $j = 1, \dots n$	Hauptkomponentenwerte	237 ff.
$\hat{\mathbf{P}} = (\hat{p}_{ij})$	$i = 1, \dots m$ $j = 1, \dots n$	Geschätzte Faktorenwerte	244 ff.
$\mathbf{B} = (\beta_{il})$	$i = 1, \dots m$ $l = 1, \dots r$	Regressionskoeffizienten der Variablen auf die Faktoren	244 ff.