

PAUL KLEMMER

Die Faktorenanalyse im Rahmen der Regionalforschung

Möglichkeiten und Grenzen ihrer Anwendung

I. Vorbemerkung

Wie in vielen anderen Wissenschaftszweigen (z. B. Medizin oder Psychologie), so finden auch in den Wirtschaftswissenschaften in der letzten Zeit die multivariaten statistischen Methoden zunehmend Anwendung¹. Dies ist z. T. auf den zunehmenden Einsatz der elektronischen Datenverarbeitung zurückzuführen, vor allem aber auch darauf, daß diese Verfahren (z. B. die Varianz- oder Diskriminanzanalyse) dazu dienen können, komplexe Sachverhalte, wie sie gerade auch für den Wirtschaftsbereich typisch sind, zu durchleuchten. Dies gilt insbesondere für die Faktorenanalyse, die aufgrund der ihr zugrunde liegenden formal-mathematischen Konzeption helfen kann, bestimmte theoretische und empirische Problemstellungen besser zu lösen².

Da dieses statistische Verfahren auch in der Bundesrepublik allmählich Eingang findet, ist es sinnvoll, in zusammenfassender Weise die grundsätzlichen Möglichkeiten und Grenzen seines Einsatzes darzustellen, und zwar vor allem im Hinblick auf den Bereich der Regionalforschung. Eine solche „Bestandsaufnahme“ erscheint auch deshalb notwendig, weil in vielen Fällen mit dem Einsatz der Faktorenanalyse Erwartungen verbunden werden, die über ihre eigentlichen Möglichkeiten hinausgehen und sie darum überfordern.

Im folgenden wird zunächst ein knapper Überblick über das der Faktorenanalyse zugrunde liegende allgemeine Konzept geboten; denn nur bei Verständnis der relevanten Annahmen sowie der notwendigen Rechenschritte kann

eine sachgemäße Anwendung erwartet werden. Vor diesem Hintergrund wird dann anhand einiger Beispiele aus der Regionalforschung ein Überblick über wichtige Einsatzmöglichkeiten dieses Verfahrens versucht, wobei gleichzeitig die Begrenztheit der mit ihrer Hilfe gewonnenen Aussagen betont werden soll.

II. Das Grundkonzept der Faktorenanalyse

Liegen für bestimmte Beobachtungseinheiten (etwa Städte oder Regionen) jeweils mehrere Merkmale vor, so läßt sich dieses statistische Ausgangsmaterial mit verschiedenen Verfahren analysieren. Man kann z. B. die zwischen den einzelnen Variablen relevanten Korrelationszusammenhänge ermitteln oder mit Hilfe regressionsanalytischer Methoden die Beziehungen zwischen „abhängigen“ und „unabhängigen“ Größen bestimmen. Damit sind aber noch nicht alle im Ausgangsmaterial enthaltenen Informationen ausgeschöpft. Es besteht vielmehr noch die Möglichkeit, daß die beobachteten Variablen gewisse Strukturen beschreiben, die nicht ohne weiteres sichtbar sind. Über die Zusammenfassung zu Variablengruppierungen in Gestalt von Faktoren lassen sich jedoch diese Gruppierungen herausarbeiten und für bestimmte Aufgabenstellungen verwerten.

Geht man nämlich von einer Korrelationsmatrix aus, so müssen bei hohen Korrelationen die im Datensatz enthaltenen Beobachtungsmerkmale nicht nur untereinander, sondern vielmehr auch mit hypothetischen dritten Größen (den sogenannten Faktoren) in einem engen Zusammenhang stehen. Das bedeutet mit anderen Worten: Hinter der mit den Beobachtungsmerkmalen errechneten Korrelationsmatrix stehen (weniger) Größen, die schwerpunktmäßig mit bestimmten Variablen korrelieren. Bei diesen Größen handelt es sich nicht um gewöhnliche Variablen, sondern um Variablen „höherer Ordnung“, die die in ihnen zusammengefaßten Merkmale repräsentieren. Mit Hilfe der zwischen den Variablen und diesen Faktoren ermittelten Korrelationswerte kann man bei Zugrundelegung bestimmter Funktionalzusammenhänge die Ausgangsmatrix reproduzieren.

Eigentliche Aufgabe der Faktorenanalyse ist die Herausarbeitung und Bestimmung dieser hinter der Korrelationsmatrix bzw. dem Ausgangsmaterial stehenden und bis jetzt noch unbekannten Faktoren. Hierbei ist zu beachten, daß diese Überlegungen bis jetzt nur auf einem rein formal-mathematischen Modell beruhen, das allein die rechnerische Reproduktion der Korrelationsmatrix gestattet³. Wichtige Modellannahmen sind Linearität der Beziehungen zwischen

¹ „Man spricht von multivariaten Methoden, wenn an jeder Person (Objekt) mehrere Variablen gemessen werden und man alle diese Größen für eine Reihe von Personen in der statistischen Analyse gleichzeitig betrachtet und auswertet“. *Überla, K.*: Faktorenanalyse. Eine systematische Einführung für Psychologen, Mediziner, Wirtschafts- und Sozialwissenschaftler. – Berlin; Heidelberg; New York 1968. S. 81. Einen guten Überblick über die multivariaten statistischen Methoden bieten z. B. *Kendall, M. G.*: A course in multivariate analysis. 3. Aufl. – London 1965; *Morrison, D. F.*: Multivariate statistical methods. – New York; St. Louis; San Francisco; Toronto; London; Sydney 1967 oder *Cooley, W. W.*; *Lohnes, P. R.*: Multivariate procedures for the behavioral sciences. 4. Aufl. – New York; London; Sydney 1967.

² Es gibt mehrere gute Einführungen in die Methodik der Faktorenanalyse, z. B.: *Fruchter, B.*: Introduction to factor analysis. – Princeton; New York; Toronto; London 1954; *Thurstone, L. L.*: Multiple-factor analysis. A development and expansion of the "vectors of mind". 7. Aufl. – Chicago; London 1965; oder *Harman, H. H.*: Modern factor analysis. 2. Aufl. – Chicago; London 1967. Über eine Einführung hinaus geht: *Horst, P.*: Factor analysis of data matrices. – New York; Chicago; San Francisco; Toronto; London 1965. Im deutschsprachigen Bereich ist auf das bereits zitierte Werk von *Überla, K.* zu verweisen. Vgl. aber auch *Richter, P.*: Anwendungen der Faktorenanalyse auf ökonomische Daten. In: Allgemeines Statistisches Archiv. Bd. 52 (1968) S. 125 ff. bzw. *Schäfer, K.-A.*: Faktorenanalyse und ihre Anwendungsmöglichkeiten. In: Allgemeines Statistisches Archiv. Bd. 53 (1969) S. 51 ff.

³ Ausführliche Erläuterungen dieses mathematisch-formalen Grundmodells finden sich in allen oben zitierten Einführungen. Vgl. insbesondere *Harman, H. H.*, a. a. O., S. 12 ff.; *Thurstone, L. L.*, a. a. O., S. 68, oder *Überla, K.*, a. a. O., S. 50 ff.

Faktoren und Variablen (d. h. konstante Koeffizienten) sowie additive Verbundenheit und Orthogonalität der Faktoren. Orthogonalität bedeutet in diesem Zusammenhang, daß die Faktoren voneinander unabhängig sind, d. h. daß die zwischen ihnen relevanten Korrelationswerte den Wert 0 annehmen.

Grundsätzlich besteht auch die Möglichkeit, von einem anderen Fundamentaltheorem auszugehen und z. B. mit einem nichtlinearen Beziehungssystem oder mit sich gegenseitig beeinflussenden Faktoren zu arbeiten. Eine Antwort auf die Frage, welches formal-mathematische Ausgangsmodell zu bevorzugen ist, kann nicht von vornherein gegeben werden. Dies hängt in starkem Maße von der mit der Anwendung verbundenen Aufgabenstellung und dem dort relevanten theoretischen Hintergrund ab. Während die Linearität angesichts des Aufbaus vieler wirtschaftstheoretischer Modelle relativ unproblematisch ist, muß man gegen die Aufrechterhaltung der Orthogonalitätsbedingung in vielen Fällen Bedenken anmelden. Es besteht jedoch die Möglichkeit, über eine schiefwinklige Rotation die Abhängigkeit der Faktoren voneinander zu prüfen.

Hauptaufgabe der Faktorenanalyse ist es somit, jene zunächst noch unbekannten Faktoren herauszuarbeiten, die den dem Fundamentaltheorem zugrundeliegenden Annahmen genügen. Es besteht hier nicht die Möglichkeit, eine detaillierte Beschreibung und Begründung der einzelnen Rechenvorgänge für die Faktorenextraktion zu geben (diese Aufgabe führt auf das sogenannte Eigenwertproblem einer reellen symmetrischen Matrix), die bei größeren Variablenzahlen ohne elektronische Rechanlage nicht mehr zu bewältigen sind.

Im Hinblick auf das Verständnis der späteren Ausführungen ist jedoch noch auf die Unterscheidung zwischen Komponentanalyse und Faktorenanalyse im engeren Sinne hinzuweisen. Im ersteren Fall geht man zur Ermittlung der fiktiven Größen (auch Komponenten genannt) von der Korrelationsmatrix (Diagonalwerte gleich 1) aus und unterstellt damit, daß diese auf alle Variablen des Ausgangsmaterials zurückgeführt werden können. Bei der Faktorenanalyse im engeren Sinne werden auch fiktive Größen zugelassen, die jeweils nur in einer Variablen vorkommen, woraus sich eine Unterscheidung in gemeinsame Faktoren und Einzelfaktoren ergibt⁴.

Nach Herausarbeitung der Faktoren sowie Entscheidung über ihre endgültig beizubehaltende Zahl (hierfür gibt es verschiedene Kriterien, wie z. B. den Scree-Test) sowie anschließender Rotation des (zumeist orthogonalen) Koordinatensystems zur „Einfachstruktur“ beginnt die für den Fachwissenschaftler relevante Aufgabe der Interpretation der Rechenergebnisse. Diese Interpretation setzt eine möglichst eindeutige Zuordnung von Variablen und Faktoren zueinander voraus, wobei die Beziehungen zwischen den Variablen und Faktoren durch eigene Korrelationswerte (die sog. Faktorenladungen) zum Ausdruck gebracht werden.

⁴ Bei einer größeren Zahl von Variablen unterscheiden sich jedoch die Ergebnisse der beiden Modellansätze immer weniger, so daß man in der Regel zur Faktorenextraktion auf die weniger komplizierte Komponentanalyse zurückgreift.

Es ist in diesem Zusammenhang wichtig, darauf hinzuweisen, daß ohne eine solche Rotation keine Interpretation der Ergebnisse durchgeführt werden darf⁵. Die zunächst ermittelte Matrix der Faktorenladungen stellt nämlich nur eine von unendlich vielen Lösungen dar, von denen mit Hilfe zusätzlicher Bedingungen die für die Interpretation geeignete ausgesucht werden muß. Eine besondere Rolle spielt hierbei das oben erwähnte Konzept der Einfachstruktur.

Die Interpretation der extrahierten Faktoren ist eine der entscheidenden Schwachstellen der Faktorenanalyse. Zum einen erfüllen gerade im sozioökonomischen Bereich die meisten Datensätze nicht die strengen Kriterien der Einfachstruktur (z. B. die Bedingungen des Bargmann-Tests), zum anderen kommt je nach Lückenhaftigkeit des theoretischen Horizonts ein subjektives Element ins Spiel⁶. Insbesondere besteht die Möglichkeit, daß gewisse „Wunschvorstellungen“ des Interpreten die inhaltliche Ausdeutung der Variablenkonfigurationen beeinflussen. Dies gilt vor allem für die Analyse kleiner Datensätze, wo häufig eine Interpretation der Faktoren mittels weniger Faktorladungen vorgenommen werden muß.

Mit dieser Interpretation hören viele Analysen auf, was zwar für manche Aufgabenstellungen genügen mag, aber in der Regel doch den Verzicht auf weitere mögliche Informationen bedeutet. Gerade im Hinblick auf viele wissenschaftliche Fragestellungen ist es nämlich zweckmäßig, aufgrund der für die einzelnen Beobachtungseinheiten gegebenen Variablenwerte diesen Merkmalsträgern durch lineare multiple Regressionsverfahren Faktorenwerte zuzuordnen. Mit den so ermittelten Schätzungen der Faktorenwerte lassen sich dann die Unterschiede der einzelnen Beobachtungseinheiten quantifizieren, wobei diese neuen Unterscheidungsmerkmale die Informationen des Ausgangsmaterials in verdichteter Form enthalten. Darüber hinaus werden auf diese Weise noch die Regressionskoeffizienten der Variablen auf die Faktoren ermittelt, und es wird somit deutlich, in welchem Ausmaß die Variationen ausgewählter Variablen die Faktorenwerte beeinflussen. Sieht man einmal von dem Testen der so ermittelten Ergebnisse über verschiedene Stichproben, Variablengruppierungen usw. ab, so ist mit der Schätzung der Faktorenwerte bzw. der Bestimmung der Regressionskoeffizienten der Variablen auf die einzelnen Faktoren das übliche Verfahren der Faktorenanalyse beendet.

Zusätzliche Überlegungen werden erforderlich, wenn die Untersuchungsobjekte durch mehrere bedeutsame Faktoren beschrieben werden. Dann steht man nämlich vor dem ent-

⁵ Man findet in der Literatur immer wieder Beispiele, wo ohne Rotation eine inhaltliche Deutung der Faktoren versucht wird; etwa neuerdings in: Beirat für Raumordnung beim Bundesminister des Innern. Empfehlungen. Folge 2. Hrsg. vom Bundesminister des Innern. – Bonn 1969. S. 28.

⁶ Vgl. auch Mälich, W.: Einige Schwierigkeiten bei der Anwendung der Faktorenanalyse, dargestellt am Beispiel einer Untersuchung des Entwicklungsstandes der Stadt- bzw. Landkreise Baden-Württembergs, S. 32 f. in: Datenverarbeitung in der Stadt- und Regionalplanung – Darmstadt 1969. = Schriftenreihe des Deutschen Rechenzentrums, H. S – 10.

scheidenden Problem, mehrere Faktorenwerte zu einer einzigen Maßzahl zusammenfassen zu müssen, was bis heute mathematisch-statistisch noch nicht geklärt ist. Insbesondere die Frage, welche Gewichte den einzelnen Faktoren bei einer Kombination zu einer gemeinsamen Größe zugeordnet werden sollen, wird zum entscheidenden Problem. Zwei Auswege bieten sich in der Regel an: 1. der Versuch einer mehrstufigen Faktorenanalyse oder 2. die alleinige Verwendung des Faktors mit dem größten Varianzanteil.

Im ersten Fall werden die Faktoren aus einer Faktorenanalyse einer weiteren Faktorenanalyse unterzogen⁷. Angesichts der Tatsache, daß bereits bei der „einfachen“ Faktorenanalyse der Nachweis der Einfachstruktur bzw. die Interpretation der Faktoren Schwierigkeiten bereitet, können die so ermittelten Ergebnisse oft sehr problematisch sein. Außerdem besteht keineswegs die Gewißheit, daß sich alle Faktoren zu einem einzigen „Superfaktor“ zusammenfassen lassen.

Im Falle der alleinigen Verwendung des wichtigsten Faktors ergibt sich automatisch ein Informationsverlust, da die Faktoren mit den geringeren Varianzanteilen nicht mehr berücksichtigt werden. Dieser Informationsverlust wird zwar in vielen Fällen bewußt in Kauf genommen, weil man nur an der Auswertung eines einzigen Faktors⁸ interessiert ist. Man sollte jedoch stets versuchen, den Varianzanteil dieses Faktors so groß wie möglich werden zu lassen⁹.

Zum Schluß dieses allgemeinen Teils sei noch auf einen vielversprechenden neuen Ansatz verwiesen, nämlich auf die drei-modale Faktorenanalyse von Tucker¹⁰. Bei vielen Aufgabenstellungen ergibt sich z. B. das Problem, daß die Faktorenwerte einer bestimmten Anzahl von Beobachtungseinheiten (z. B. die Stadt- und Landkreise der Bundesrepublik oder eines Bundeslandes) aus verschiedenen Jahren stammen und daher nicht direkt vergleichbar sind. Hier kann der dreidimensionale Ansatz weiterhelfen, wobei man sich vorstellen muß, daß die Beobachtungen über die Variablen, über die Regionen und die Jahre vorgenommen werden. Jeder Art dieser Beobachtungen kann nun eine Faktormatrix zugeordnet werden, wobei sich diese drei Matrizen wiederum durch eine übergeordnete dreidimensionale Matrix, die sogenannte Zentralmatrix, verbinden lassen. Führt man jetzt für die ursprünglichen Faktormatrizen Rotationen durch, so ändert sich auch entsprechend die Zentralmatrix. Hieraus läßt sich nach einigen Transformationen zeigen, wie sich die Faktoren über die einzelnen Jahre hinweg verändern.

III. Die Haupteinsatzmöglichkeiten der Faktorenanalyse unter besonderer Berücksichtigung regionalwirtschaftlicher Fragestellungen

1. Die Faktorenanalyse als Mittel der Datenreduktion

Die erste Verwendungsmöglichkeit der Faktorenanalyse liegt in der damit verbundenen Chance der Datenreduktion umfassender Ausgangsbeobachtungen. Sobald nämlich die Zahl der Faktoren kleiner ist als die Zahl der Variablen, d. h., sobald sich typische Gruppierungen von Variablen herausbilden, die eindeutig einem Faktor zugeordnet werden können, erfolgt eine Verdichtung des statistischen Ausgangsmaterials in Form weniger meßbarer Größen.

Ein solches Vorgehen ist somit dann sinnvoll, wenn für die Beobachtungseinheiten jeweils mehrere Merkmale vorliegen, die angesichts ihrer Vielfalt den Überblick über die Untersuchungsobjekte erschweren. Die zu analysierenden Merkmalsträger lassen sich dann mittels einer Variablen, einer Variablenkombination oder mittels aller Angaben gliedern. Was nun das letztere Vorgehen angeht, so bietet die Faktorenanalyse die einzigartige Möglichkeit, die im Ausgangsmaterial enthaltene Datenfülle zu verengen, ohne daß auf irgendeine vielleicht wichtige Detailinformation verzichtet werden muß. Gegenüber der häufig willkürlichen Bevorzugung bestimmter Einzelbeobachtungen oder der problematischen Verknüpfung von Größen zu Merkmalskombinationen kann man die so ermittelten Ergebnisse als einen echten Fortschritt bezeichnen.

Solange nur die Aufgabe der Datenreduktion im Vordergrund steht, kann man sich auch auf die Hauptkomponentenanalyse beschränken und auf Rotation zur Einfachstruktur verzichten. Das bedeutet jedoch, daß man von einer inhaltlichen Deutung der Faktoren absehen muß, da ohne Rotation nicht von einer eindeutigen Zuordnung der Variablen zu den Faktoren gesprochen werden kann. Die so ermittelten Ergebnisse entsprechen somit nur den formal-mathematischen Erfordernissen des der Faktorenanalyse zugrunde liegenden Fundamentaltheorems, d. h., es handelt sich um rein „mathematische Konstruktionen mit bestimmten optimalen Eigenschaften, denen in der Realität nichts zu entsprechen braucht“¹¹.

Trotz dieser Einschränkung im Hinblick auf eine mögliche wirtschaftstheoretische Interpretation kann die Datenreduktion mittels der Hauptkomponentenanalyse wertvolle Hilfe leisten. Gerade die Gruppierung komplexer Beobachtungseinheiten wird auf diese Weise befriedigender gelöst als durch das Herausgreifen einer oder einiger weniger Variablen. Aus diesem Grund nimmt es nicht wunder, daß diese Anwendungsform der Faktorenanalyse sich in der praktischen Regionalforschung zunehmend größerer Beliebtheit erfreut. Es ist zu erwarten, daß sie sich gerade im Hinblick auf die Versuche der Gemeinde-, Stadt- oder Regionstypisierung bei uns stärker durchsetzen wird. Ein Nachteil dieses Verfahrens bleibt jedoch der Tatbestand, daß die auf diese Weise herausgearbeiteten Komponenten

⁷ Vgl. Schraeder, W. F.: Programme der Faktorenanalyse. Anwendung in der Stadt- und Regionalplanung. In: Datenverarbeitung in der Stadt- und Regionalplanung, a. a. O., S. 28 ff.

⁸ Etwa weil er sich in theoretisch relevanter Weise interpretieren läßt.

⁹ Z. B. durch Datenselektion oder Datentransformation.

¹⁰ Vgl. Tucker, L. R.: Implications of factor analysis of three-way matrices for measurement of change. In: Problems in measuring change. Hrsg. von Ch. W. Harris. – Madison; London 1967. S. 122 ff.; vgl. auch derselbe: The extension of factor analysis to three-dimensional matrices. In: Contribution to mathematical psychology. Hrsg. von F. Frederiksen und H. Gulliksen. – New York 1964. S. 110 ff.; sowie Levin, J.: Three-mode factor analysis. In: Psychological Bulletin. Bd. 64 (1965) S. 442 ff.

¹¹ Überla, K., a. a. O., S. 356.

quasi anonym bleiben und damit als Faktoren ohne interpretierbaren Inhalt auftreten.

2. Die Faktorenanalyse als Mittel zur Formulierung komplexer Indikatoren

Verbindet man die Aufgabe der Datenreduktion mit einer Interpretation der Faktoren, so kommt man zu einem Anwendungsbereich, in dem die Faktorenanalyse eine dominierende Rolle einnehmen kann. Diese inhaltliche Deutung erfordert aber zusätzliche Rechenschritte, vor allem – wie bereits mehrfach erwähnt – die Rotation zur Einfachstruktur. Das sich hieraus ergebende Faktorenmuster bietet sich dann der Interpretation an, d. h., aus der Eigenart und inhaltlichen Ausrichtung der für den jeweiligen Faktor relevanten hochladenden Beobachtungsmerkmale versucht man, den Faktor zu deuten.

Für den so gedeuteten Faktor kann man im Hinblick auf die einzelnen Beobachtungseinheiten die Faktorenwerte schätzen sowie die Regressionskoeffizienten der Variablen auf die Faktoren bestimmen. Auf dieser Ebene lassen sich Probleme lösen, die vor allem bei der quantitativen Erfassung komplexer bzw. qualitativer Tatbestände auftreten. Diese Möglichkeiten sollen anhand der folgenden Beispiele aus der Regionalforschung näher erläutert werden¹².

Eine entscheidende Aufgabe der regionalen Wirtschaftspolitik wird z. B. im Abbau interregionaler Entwicklungsunterschiede gesehen. Das hierbei angesprochene Kriterium des „Entwicklungsniveaus“ ist jedoch ein recht komplexer Ausdruck, dessen inhaltliche Deutung von vornherein keineswegs eindeutig festliegt. Man kann diese Größe einmal mehr in Richtung auf den Aspekt der „allgemeinen Lebensbedingungen“ interpretieren oder durch alleinige Verwendung der jeweiligen Werte des Bruttoinlandsprodukts mehr auf die Erfassung des Leistungsprozesses abstellen. Im Fall der erweiterten Ausdeutung des „Entwicklungsstandes“ bleibt jedoch die entscheidende Frage, welche Merkmale mit welchen Gewichten in die zu bildende Merkmalskombination eingehen sollen, ungelöst. Zumeist geht man daher von der ungeprüften Annahme aus, das Bruttoinlandsprodukt

repräsentiere das allgemeine Entwicklungsniveau am besten.

Hier bietet nun die Faktorenanalyse eine Möglichkeit, die insbesondere im Hinblick auf das bisher recht unbefriedigend gelöste Problem der Selektion und Gewichtung der Beobachtungsmerkmale zu besseren Ergebnissen führen kann. Dies zeigt u. a. eine im Institut für Regionalpolitik und Verkehrswissenschaft der Universität Freiburg entstandene Studie über den Entwicklungsstand der 72 Stadt- und Landkreise Baden-Württembergs für die Jahre 1957, 1961 und 1964. Ausgangsmaterial waren 21 Kennziffern aus der Kreismappe des Instituts für Raumordnung, die, wie Testrechnungen mit einem umfassenden Datensatz von bis zu 47 Kennziffern für 1961 zeigten, für diese Aufgabe völlig ausreichen.

Mit Hilfe der Faktorenanalyse konnte aus diesem Ausgangsmaterial ein Faktor abgeleitet werden, der aufgrund des Faktorenusters eine inhaltliche Ausdeutung als „Entwicklungsstand“ zuließ. Die für die einzelnen Beobachtungseinheiten (Stadt- und Landkreise Baden-Württembergs) ermittelten Faktorenwerte ließen eine tabellarische Anordnung der Merkmalsträger zu, wobei diese Gruppenbildung in 10 Fällen (d. h. beinahe 14%) von jener mittels des Bruttoinlandsproduktes je Person der Wohnbevölkerung abwich¹³.

Eine besondere Bedeutung kommt in diesem Zusammenhang der Analyse der Faktorenladungen zu, deren Quadrat die durch die einzelnen Variablen erklärten Varianzanteile des Faktors angibt. Je höher hierbei der Varianzanteil einer Variablen ist, desto eher repräsentiert eine solche einzelne Merkmalsgröße einen Faktor. Interessanterweise ergab sich im Hinblick auf die Stadt- und Landkreise Baden-Württembergs, daß die Kennziffer „Gesamtumsatz je Einwohner in DM“ in der Regel den höchsten Repräsentationsgrad (zwischen 78,25% und 85,21%) aufwies und das „Bruttoinlandsprodukt je Person der Wohnbevölkerung“ mit einem Varianzanteil zwischen 58,22% und 80,51% erst auf dem zweiten Platz folgte.

Entscheidend für eine befriedigende Anwendung dieses Verfahrens zur Bestimmung des regionalen Entwicklungsstandes ist die Bedingung, daß möglichst alle relevanten Variablen in die Analyse eingehen. Als „relevant“ werden in diesem Zusammenhang jene Merkmale bzw. Merkmals Transformationen angesehen, für die nur der geringste Verdacht besteht, daß sie in einer Beziehung zum Entwicklungsstand stehen könnten. Das setzt normalerweise bereits das Vorhandensein eines theoretischen Hintergrunds voraus, ohne den in der Regel auch die Interpretation der Faktoren sehr erschwert wird. Es mag auch vorkommen, daß bestimmte Variablen im Hinblick auf ausgewählte Faktoren hohe Ladungen aufweisen. Dann kann dies ein Hinweis für eine eventuelle Modifikation des bisherigen theoretischen Modellansatzes sein.

Alle mit Hilfe der Faktorenanalyse ermittelten Merkmalskombinationen beziehen sich immer auf das in die statistische Analyse eingehende Ausgangsmaterial. Insofern

¹² Als typische Beispiele dieser Anwendungsrichtung mögen dienen: z. B. *Berry, Brian J. L.*: A method for deriving multi-factor uniform regions. In: *Przegląd Geograficzny*. Bd. 33 (1961) S. 263 ff.; derselbe: Identification of declining regions: an empirical study of the dimensions of rural poverty. In: *Areas of Economic Stress in Canada*. Hrsg. von R. S. Thoman und W. D. Wood. – Kingston (Ontario) 1965. S. 22 ff.; *Müller, J. Heinz; Klaus Rittenbruch*: Die Methoden der Regionalanalyse unter besonderer Berücksichtigung ihrer Verwendbarkeit für die Beurteilung verschiedener Projekte. Forschungsbericht, erstellt im Auftrag des Bundesministers für Wirtschaft. – Freiburg 1968. (Vervielf. Manuskript); *Smith, D. M.*: Identifying the "grey" areas: multivariate approach. In: *Regional Studies*. Bd. 2 (1968) S. 183 ff.; *Geisenberger, S.; W. Mälich; J. H. Müller; G. Strassert*: Zur Bestimmung wirtschaftlichen Notstandes und wirtschaftlicher Entwicklungsfähigkeit von Regionen. Eine theoretische und empirische Analyse anhand von Kennziffern unter Verwendung von Faktoren- und Diskriminanzanalyse. – Hannover 1970. = Abhandlungen der Akademie für Raumforschung und Landesplanung, Bd. 59. *Geisenberger, S.; W. Mälich*: Verbesserte Maßstäbe zur Bestimmung unterdurchschnittlich entwickelter Gebiete. In: *Institut für Raumordnung. Informationen*. 20. Jg. (1970) H. 10, S. 301 ff.

¹³ Vgl. *Geisenberger, S.; W. Mälich*, a. a. O., S. 308.

gibt dieses Verfahren weniger eine Antwort auf die Frage, welche Variable in einem vom theoretischen Standpunkt aus befriedigenden Index eingehen müßte, sondern vielmehr nur eine Lösung für die Problemstellung, ob z. B. unter den statistisch ausgewiesenen Merkmalsgrößen Gruppierungen vorhanden sind, die sich zu einem als „Entwicklungsstand“ interpretierbaren Faktor zusammenfassen lassen. Die Selektions- und Gewichtungsaufgabe wird somit nur für die im Ausgangsmaterial enthaltenen Informationen gelöst, und zwar jeweils nur für die vorgegebene Stichprobe, d. h. für die Gesamtheit der genau fixierten Beobachtungseinheiten zu einem bestimmten Zeitpunkt.

Solche Problemstellungen sind in der Regionalforschung häufig anzutreffen, und zwar insbesondere dann, wenn es sich um stark qualitativ geprägte Erscheinungen handelt. Nimmt man z. B. den „Verstädterungs-“ oder auch den „Zentralitätsgrad“, so handelt es sich hier um zwei Tatbestände, die in der regionalen Wirtschaftstheorie eine große Rolle spielen, die sich aber in der Regel einer quantitativen Erfassung entziehen. Aus diesem Grunde werden sowohl für den Verstädterungs- als auch für den Zentralitätsgrad in der Literatur zumeist mehrere Ersatzindikatoren verwendet, ohne daß von vornherein feststeht, daß diese auch im Hinblick auf den von ihnen vertretenen Tatbestand einen hohen Repräsentationsgrad aufweisen.

Solche Ersatzgrößen für die Messung der Urbanisierung sind z. B. die Bevölkerungsgröße, die Zahl der Wohnungen oder Gebäude, die Bevölkerungs- oder Bebauungsdichte, die Zahl der Erwerbspersonen in nichtlandwirtschaftlichen Arbeitsstätten, der Anteil der in nichtlandwirtschaftlichen Arbeitsstätten Beschäftigten an der Gesamtbevölkerung oder die Gesamtzahl der Beschäftigten oder auch die örtliche Häufung „zentraler“ Einrichtungen (wobei man häufig nicht weiß, wo die Grenze zwischen zentralen und nichtzentralen Einrichtungen zu ziehen ist), um nur einige typische Indikatoren zu nennen.

Ein eng mit der Verstädterung verknüpft Problem ist somit die quantitative Erfassung des Zentralitätsgrades, insbesondere im Hinblick auf eine Politik der zentralen Orte, die an einem vorgegebenen Siedlungsnetz anknüpfen muß. Gäbe es, wie es der theoretische Ansatz von *Walter Christaller* impliziert, ein streng hierarchisches System zentraler Güter und Dienste, und zwar derart, daß in jedem zentralen Ort einer bestimmten Ordnung auch immer alle Güter angeboten werden, die in Orten niedriger Ordnung vorhanden sind, dann ließe sich über eine Skalogrammanalyse, d. h. über die alleinige Erfassung des Gutes mit der jeweils höchsten Ordnung, gemessen an einer vorgegebenen und bekannten Güterhierarchie, das Problem der Zentralitätsmessung lösen. Nun zeigt sich aber, daß in der Realität diese strenge vertikale Anordnung und Schichtung der zentralen Güter nicht gegeben ist und sich außerdem die Frage, welcher Dienst noch als zentral angesehen werden kann, auch nicht exakt von vornherein beantworten läßt.

Gerade was die höheren Dienstleistungen angeht, kann sich vielmehr eine Spezialisierung der Städte auf bestimmte Funktionen ergeben, die bei Lagerung in den Verdich-

tungsräumen häufig sogar den Charakter einer zwischenörtlichen Arbeitsteilung annimmt. In solchen Fällen versagt die Skalogrammanalyse, und das alte Problem der Selektion und Gewichtung von Zentralitätsmerkmalen taucht wieder auf. Angesichts dieses ungelösten Problems sind auch die meisten bisherigen Verfahren der Messung des Zentralitätsgrads wenig befriedigend. Zumeist greift man sich nämlich nur einzelne Beobachtungsmerkmale heraus und unterstellt ihnen ohne nähere Prüfung im Hinblick auf die Zentralität einen recht hohen Repräsentationsgrad. Beispiele solcher Ersatzindikatoren sind etwa die Zahl der Einwohner, die Zahl der Telefonanschlüsse, die Zahl der Einzelhandels- oder Großhandelsbeschäftigten, die Stärke der Zugfrequenz oder Kombinationen solcher Angaben. Gerade hier vermag die Faktorenanalyse zu helfen – insbesondere, wenn es gelingt, einen als Zentralitätsgrad interpretierbaren Faktor herauszuarbeiten, der über eine Schätzung der Faktorenwerte auch eine kardinale Einstufung der Beobachtungseinheiten erlaubt¹⁴.

3. Die Faktorenanalyse als Mittel der Hypothesenfindung und -überprüfung

Die Faktorenanalyse wird weitgehend auch als eine Methodik der Hypothesenfindung angesehen. Hier erhofft man sich bei Anwendung auf komplexe und wenig erforschte Gebiete interpretierbare Faktoren, die aufgrund der ermittelten Variablenkonfigurationen Hinweise auf mögliche Beziehungszusammenhänge geben. Gerade diese Interpretation der Faktoren anhand des Faktorenmusters setzt jedoch bereits ein Mindestmaß an theoretischem Wissen voraus, da ohne diesen theoretischen Hintergrund die inhaltliche Ausdeutung der ermittelten Faktoren durch einen zu großen Spielraum gekennzeichnet wäre.

Es wurde bereits darauf hingewiesen, daß sich bei den meisten Interpretationsversuchen subjektiv bestimmte Plausibilitätsannahmen nicht vermeiden lassen. Im Falle unerforschter Sachzusammenhänge wird man darum darauf achten müssen, daß ein Minimum an „Objektivität“ gewahrt bleibt, wobei die „Plausibilität“ der inhaltlichen Ausdeutung der Faktoren stets auf bereits bekanntes theoretisches Wissen zurückgreifen muß.

Der Einsatz der Faktorenanalyse für den Zweck der Hypothesenformulierung verlangt daher ein sehr vorsichtiges Vorgehen, insbesondere eine schiefwinklige Rotation, da den Beziehungen zwischen den Faktoren besondere Aufmerksamkeit zugewendet werden muß. Weiter sollte man bestrebt sein, möglichst viele und verschiedenartige Variablen in die Analyse einzubeziehen, da man sonst in die Gefahr kommt, bestimmte Größen zu bevorzugen und damit die Ergebnisse durch bewußte Selektion der Ausgangsbeobachtungen zu beeinflussen.

Diese Probleme sind auch sicherlich die Ursache für die Schwierigkeit, innerhalb der Regionalforschung Beispiele

¹⁴ Einen solchen Versuch hat der Verf. dieses Beitrages im Hinblick auf die Stadtregionen unternommen. *Klemmer, P.*: Der Metropolisierungs- und Verstädterungsgrad der westdeutschen Verdichtungsräume. – Freiburg 1970. (Unveröffentl. Mschr.)

für eine solche Anwendung der Faktorenanalyse zu finden. Meistens handelt es sich nämlich um den Versuch, bereits bekannte Hypothesen im Hinblick auf eine bestimmte Stichprobe zu überprüfen. Hier wird somit von einem bekannten theoretischen Hintergrund ausgegangen, und es gilt nur die Aufgabe zu lösen, einen im Hinblick auf die Problemstellung relevanten Faktor zu extrahieren. Dann läßt sich anhand des Faktorenmusters überprüfen, ob zwischen dem Faktor und ausgewählten Variablen ein straffer Zusammenhang besteht oder nicht¹⁵.

Nimmt man den bereits kurz diskutierten Fall der quantitativen Bestimmung des Zentralitäts- oder Verstärkungsgrads von Siedlungskörpern, so kann man – bei Vorliegen eines als Zentralitätsgrad interpretierbaren Faktors – anhand des Faktorenmusters nachweisen, ob Beobachtungsmerkmale theoretisch interessante hohe Faktorladungen aufweisen. Gleichzeitig läßt sich mit Hilfe der multiplen Regressionsanalyse der Einfluß dieser Variablen auf den Faktor bestimmen.

So wird in der Theorie der zentralen Orte z. B. der Entfernung zu anderen Siedlungskörpern eine große Bedeutung zugemessen, d. h. eine Beziehung zwischen Zentralität und Bevölkerungsgröße unterstellt. Anhand der Kerngebiete der Stadtregionen in der Bundesrepublik konnte der Verfasser dieses Beitrags aufzeigen, daß bei Städten dieser Größenordnung die Entfernung zur nächsten Stadtregion oder zur nächsten Stadt über 60 000 Einwohner keine Rolle spielt und die Bevölkerungsgröße in keinem Zusammenhang zur Zentralität steht¹⁶. Bei Ausschalten der großen Ballungsgebiete gewinnt die Bevölkerungsgröße – gemessen an der jeweiligen Faktorladung – an Gewicht, obwohl der Regressionskoeffizient immer noch sehr klein ist.

¹⁵ So versucht z. B. *Fischer*, mit Hilfe der Faktorenanalyse die Relevanz der Exportbasentheorie zu testen. Vgl. *Fischer, Alois*: Die Struktur von Wirtschaftsräumen. Ein Beitrag zur Anwendung statistischer Methoden in der Regionalforschung. – Wiesbaden 1969.

¹⁶ Vgl. *Klemmer, P.*, a. a. O.

IV. Zusammenfassung

Die Faktorenanalyse als Instrument der Regionalforschung bietet sich für mehrere Verwendungsmöglichkeiten an, insbesondere zur Datenreduktion, zur Formulierung komplexer Indikatoren sowie zum Finden und zur Überprüfung von Hypothesen. Die bisherigen Erfahrungen haben gezeigt, daß das Schwergewicht eindeutig beim zweiten Anwendungsbereich liegt, weil sich hier die Möglichkeiten dieser multivariaten statistischen Methode am besten ausnutzen lassen. Gerade im Hinblick auf das mit mehr traditionellen Verfahren kaum lösbare Problem der Selektion und Gewichtung von verschiedenen Beobachtungen zu einer einzigen komplexen Merkmalskombination bietet sie einen echten Ausweg. Sie erlaubt hierdurch die gleichzeitige Erfassung primär qualitativer Tatbestände sowie einen kardinalen Vergleich der Beobachtungseinheiten mit Hilfe der Faktorenwerte.

Geringe Bedeutung kommt der Faktorenanalyse als Mittel der Formulierung von Hypothesen zu. Vor allem das Interpretationsproblem ist aufgrund des großen Einflusses subjektiver Entscheidungen ein großes Hindernis. Ohne ein Mindestmaß an theoretischem Wissen wird hier der Willkür ein großer Spielraum eröffnet. Zweckmäßiger hingegen ist der Einsatz zur Überprüfung möglicher Zusammenhänge, insbesondere, wenn es sich hierbei um Größen handelt, die sich durch stark qualitativen Charakter auszeichnen.

Die bisherigen Anwendungen haben bereits gezeigt, daß sich die Faktorenanalyse in der Regionalforschung durchsetzen wird. Gleichzeitig wurde aber deutlich, daß dieses statistische Verfahren kein Generalrezept darstellt, sondern seine Anwendung genauere Kenntnisse der einzelnen Rechenschritte sowie der damit verbundenen Probleme voraussetzt. Nur so lassen sich jene Fehler vermeiden, die bei einer unüberlegten Anwendung auftreten können und damit häufig Anlaß zu einer nicht berechtigten Skepsis gegenüber dieser Methode werden.