

Eine nichtlineare Analyse der sozioökonomischen Entwicklung von Nationen

Von Gerhard Kockläuner

Die vergleichende Länderforschung arbeitet auf der Basis multidimensionaler sozioökonomischer Entwicklungsbegriffe. Der Entwicklungsstand einzelner Nationen wird dabei mit Hilfe multivariater statistischer Verfahren analysiert. Diese Verfahren erfassen häufig nur lineare Beziehungen zwischen einzelnen Indikatoren für Entwicklung bzw. Unterentwicklung. Als Alternative dazu wird hier eine nichtlineare verallgemeinerte Hauptkomponentenanalyse vorgestellt und exemplarisch durchgeführt.

Einleitung

Ländervergleiche auf der Basis BSP/Kopf bleiben notwendig auf rein ökonomische Kategorien beschränkt. Als Alternative sind daher synthetische Indikatoren erarbeitet worden, die die sozioökonomische Entwicklung von Nationen beschreiben sollen¹. Solche Indikatoren reduzieren den Entwicklungsbegriff auf eine einzige synthetische Dimension. Der Multidimensionalität von Entwicklung Rechnung tragend, werden multivariate statistische Verfahren genutzt, um synthetische Indikatoren von Unterentwicklung bzw. Entwicklung zu gewinnen. Wie ein nachfolgend gegebener Überblick zeigt, geht die Mehrzahl der zu diesem Zweck in der vergleichenden Länderforschung genutzten multivariaten Verfahren von bestimmten Linearitätsannahmen aus. Entwicklungsprozesse und Beziehungen zwischen Indikatoren für Entwicklung bzw. Unterentwicklung zeigen sich jedoch in der Regel als nichtlinear.

So wird hier die Anwendung der nichtlinearen verallgemeinerten Hauptkomponentenanalyse vorgeschlagen, um multidimensionale sozioökonomische Entwicklungsbegriffe auf niedrige Dimensionen zu reduzieren. Zur Illustration dient eine quadratische Hauptkomponentenanalyse, die für fünf von Sato (1982) abgeleitete Indikatorengruppen vorgenommen wird. Wie die durchgeführte Analyse zeigt, sind Interaktionseffekte zwischen originären sozioökonomischen Entwicklungs- und Strukturindikatoren bestimmend für die schließlich gewonnene Dimensionshöhe. Gleichzeitig wird deutlich, daß das untersuchte „Ausmaß an Dimensionalität“ größer als Eins bleibt.

¹ Vgl. z.B. *Mc Granahan et al.* (1977).

Eindimensionale Ländervergleiche, die z.B. auf der Basis des BSP/Kopf, aber auch beruhend auf eindimensionalen synthetischen Indikatoren erfolgen, können daher nur Teilaspekte von Unterentwicklung bzw. Entwicklung beleuchten. Dabei werden besonders solche Aspekte ausgeblendet, die, z.B. bezogen auf politische Variablen, qualitativen Charakter haben. Es ergibt sich somit ein Plädoyer für eine mehrdimensionale, auch qualitative Analyse des Entwicklungsstandes einzelner Nationen.

1. Zur Multidimensionalität von Entwicklung

*Nohlen/Nuscheler*² definieren Entwicklung über die Elemente „Wachstum“, „Arbeit“, „Gleichheit/Gerechtigkeit“, „Partizipation“ und „Unabhängigkeit“. Diesem mittels zu definierender Ziele entstandenen Entwicklungsbegriff steht ein entsprechend multidimensionaler und komplexer Begriff von Unterentwicklung gegenüber³. Multidimensionale Konzeptionen von Entwicklung erlauben, „die Reduktion des Entwicklungsbegriffs auf das meßbare BSP zu überwinden“⁴. Eine Quantifizierung dieser Konzeptionen erfordert Indikatoren von Unterentwicklung und Entwicklung, die z.B. „repräsentative Einzelaspekte“⁵ sozialer Entwicklung indirekt messen. Dabei gilt: „Auswahl und Gewichtung der Indikatoren lassen auch dann, wenn sie nicht theoretisch expliziert werden, Rückschlüsse auf theoretische Prämissen zu“⁶. Bestimmte Indikatoren korrespondieren jeweils zu bestimmten Entwicklungszielen. Die Anzahl der aufgenommenen Indikatoren gibt die Dimension der gewählten Entwicklungskonzeption an. So arbeitet *Scholing*⁷ „mit 19 gängigen sozioökonomischen Entwicklungsindikatoren“ unter der folgenden Prämisse: „Es gibt keinen festumrissenen Entwicklungsbegriff und damit auch kein empirisch-operationales Meßkonzept, anhand dessen sich Entwicklungsindikatoren eindeutig festlegen ließen“⁸. Das UN-Research Institute for Social Development (UNRISD) unterscheidet zwischen Entwicklungs- und Strukturindikatoren, die im sozialen/demographischen sowie im wirtschaftlichen Bereich einander gegenübergestellt werden⁹. So steht z.B. der für das Entwicklungsziel „Befriedigung der Grundbedürfnisse“ wesentliche Indikator „Durchschnittlicher Kalorienverbrauch pro Kopf und Tag“ neben dem Strukturindikator „Kalorienverbrauch an Getreide und Stärkemehl im Verhältnis zum Ge-

² Vgl. *Nohlen / Nuscheler* (1982), 55, 68.

³ Vgl. *Nohlen / Nuscheler* (1982), 47.

⁴ Vgl. *Nohlen / Nuscheler* (1974), 325.

⁵ Vgl. *Nohlen / Nuscheler* (1974), 326.

⁶ Vgl. *Nohlen / Nuscheler* (1982), 457.

⁷ Vgl. *Scholing* (1982), 202.

⁸ Vgl. *Scholing* (1982), 208.

⁹ Vgl. *Mc Granahan et al.* (1977), 31 ff.

samtkalorienverbrauch“. Entwicklungsindikatoren sind jeweils vom Prozent-Typ oder vom Pro-Kopf-Typ, Strukturindikatoren jeweils vom Prozent-Typ¹⁰. Qualitative Indikatoren, die politische, kulturelle, institutionelle oder psychologische Elemente von Entwicklung kennzeichnen, bleiben nicht nur bei *Mc Granahan et al.* (1977) unberücksichtigt. Die damit verbundene Dimensionsreduktion des Entwicklungsbegriffes läßt sich umgehen, wenn z.B. qualitative politische Variablen auf geeignete Weise kardinalisiert werden¹¹. Verfahren zur Skalierung kategoriieller Merkmale sind bei *Hartung/Elpelt*¹² beschrieben.

2. Multivariate statistische Verfahren in der vergleichenden Länderforschung

Ein gegebenes System quantitativer Entwicklungs- und Strukturindikatoren erlaubt nach 1. eine mehrdimensionale Messung des sozioökonomischen Entwicklungsstandes verschiedener Länder. Für die vergleichende Länderforschung wird im folgenden davon ausgegangen, daß für eine Vielzahl von Ländern (größer als die Zahl der Indikatoren) Indikatorwerte mit Bezug auf ein bestimmtes Kalenderjahr vorliegen. Zu Kriterien und Problemen der Indikatoreauswahl sei auf *Mc Granahan/Pizarro/Richard* (1979), zur Ersetzung fehlender Indikatorwerte auf *Mc Cranahan et al.* (1974) verwiesen.

Soll nun der sozioökonomische Entwicklungsstand einzelner Länder verglichen werden, so ist der Multidimensionalität von Entwicklung Rechnung zu tragen. Ein eindimensionaler Vergleich auf der Basis einzelner originärer Entwicklungs- bzw. Strukturindikatoren unterliegt her der gleichen Kritik, wie sie der Verwendung des BSP/Kopf als alleinigem Entwicklungsindikator zukommt¹³.

Als Alternative bietet sich eine Anwendung multivariater statistischer Verfahren an, wie sie, wenn auch unter irreführendem Titel, von *Adelman/Morris* (1971) diskutiert werden. Diese Verfahren erlauben mehrdimensionale Vergleiche des Entwicklungsstandes verschiedener Länder, können aber auch genutzt werden, um multidimensionale Entwicklungsbegriffe auf wenige Dimensionen zu reduzieren. Nach *Wold* (1979) lassen sich solche Verfahren in die Linien R bzw. S. einreihen. Die Linie R umfaßt multiple Regressionsansätze mit erklärenden Kausalstrukturen. Da jedoch „development is to large extent by its nature a system of interdependent variables“¹⁴,

¹⁰ Vgl. *Mc Granahan / Pizarro / Richard* (1979), 3 f.

¹¹ Vgl. *Adelman / Morris* (1965).

¹² Vgl. *Hartung / Elpelt* (1984), 269 ff.

¹³ Vgl. *Nohlen / Nuscheler* (1982), 458.

¹⁴ Vgl. *Mc Granahan et al.* (1977), 12.

kommen *Mc Granahan et al.* (1977), 2, folgerichtig zu der Annahme, „that developmental data ... do not as a rule justify the use of correlation analysis or regression analysis to determine directions and amounts of causal influence or functional dependency in the development process“. Entsprechend sind die bei *Adelman/Morris* (1971) aufgeführten multivariaten Verfahren vor allem deskriptiv, damit der Linie S zuzuordnen. Diese Verfahren – wie auch diejenigen der Linie R – zerfallen nun wiederum in solche, die variablen-, d. h. indikatororientiert sind (R-Techniken) und solche, die auf Merkmalsträger, d. h. Länder abheben (Q-Techniken)¹⁵.

Für eine vergleichende Länderforschung müssen insbesondere Q-Techniken von Interesse sein. Entsprechende Methoden der geometrischen Datenanalyse finden sich als nichtmetrische multidimensionale Skalierung in den Arbeiten von *Adelman/Morris* (1974) und *Tosstorff* (1983), als hierarchische Clusteranalysen bei *Scholing* (1981) und auch *Tosstorff* (1983). Nach obiger Diskussion erscheint dabei problematisch, das *Scholing* (1981) eine von den von ihm gewählten Indikatoren abhängige Einkommensvariable zur Bildung von Ländergruppen heranzieht. Analog benötigt man einen Merkmalsträger als Referenzland, wenn die von *Siegel* (1982) entwickelte robuste Version der Procrustes Analyse (Matching Configurations) auf Ländervergleiche übertragen wird¹⁶. Robuste, d. h. vor allem gegenüber Ausreißern resistente Q-, aber auch R-Techniken, sind besonders in der vergleichenden Länderforschung gefordert, da viele Entwicklungsindikatoren „extreme Verteilungen“¹⁷ aufweisen. Solche Techniken gehen in Form der ‚best fitting line‘¹⁸ auch in die Gewinnung graphischer Länderprofile ein¹⁹. Diese Profile erlauben wie andere multivariate graphische Displays – z. B. in der Form von Gesichtern wie bei *Abel* (1985) – einen bilateralen Ländervergleich auf der Basis aller vorhandenen Indikatoren. Eine Dimensionsreduktion des Länderraumes in Anlehnung an die multidimensionale Skalierung bzw. Clusteranalyse findet hier nicht statt. Andererseits lassen sich die Ergebnisse hierarchischer Clusteranalysen wiederum graphisch darstellen²⁰.

Dimensionsreduzierende statistische Verfahren in der vergleichenden Länderforschung sind vor allem auch danach zu beurteilen, ob sie Linearitätsannahmen enthalten. So gilt nach *Mc Granahan et al.*²¹: „the process of development involving both economic and social variables is not a uniform linear progression of variables but rather a changing complex of factors.“ Die metrische multidimensionale Skalierung mit ihrer linearen Hauptkoor-

¹⁵ Vgl. *Hartung / Elpelt* (1984), 15.

¹⁶ Vgl. auch *Seber* (1984), 253.

¹⁷ Vgl. *Tosstorff* (1983), 146.

¹⁸ Vgl. *Mc Granahan et al.* (1974).

¹⁹ Vgl. *Mc Granahan et al.* (1977), 119 ff.

²⁰ Vgl. *Seber* (1984), 135.

²¹ Vgl. *Tosstorff* (1983), 114.

dinatenanalyse kommt daher für Ländervergleiche kaum in Betracht²². Solen keine Annahmen über die funktionale Form des gemeinsamen Zusammenhangs der Indikatoren getroffen werden, können Ländervergleiche über Rangordnungen einzelner Indikatorwerte stattfinden. So verwendet *Tossorff* (1983) Rangkorrelationskoeffizienten zur Prüfung monotoner Zusammenhänge im Rahmen seiner *Unfolding-Analyse* als Spezialfall der nicht-metrischen multidimensionalen Skalierung. Das *Unfolding-Verfahren* stellt Länder und Indikatoren räumlich dergestalt dar, daß die Rangfolge der Länderabstände von einem Indikatorpunkt nach Möglichkeit mit der Rangfolge der Indikatorwerte übereinstimmt. Es steht damit als statistisches Verfahren zwischen den Q- und R-Techniken. Dies gilt auch für die von *Adelman/Morris/Wold* (1980) vorgenommene disjunkte Hauptkomponentenanalyse. Darin werden die Clusteranalyse als Q-Technik und die der Hauptkoordinatenanalyse im Bereich der R-Techniken gegenüberstehende Hauptkomponentenanalyse miteinander verknüpft²³. Die zur Verknüpfung genutzten Regressionsbeziehungen unterliegen für Entwicklungsindikatoren als abhängige Variablen der oben geäußerten Kritik.

R-Techniken wie die Hauptkomponentenanalyse bieten sich dann zur vergleichenden Länderforschung an, wenn es gilt, die Vielzahl von einzelnen Indikatoren auf die eine Dimension eines daraus abgeleiteten synthetischen „general index of socio-economic development“²⁴ zu reduzieren. Ein solcher Index soll die genannte Multidimensionalität von Entwicklung bzw. Unterentwicklung repräsentieren. Anhand seiner Werte aufgestellte Länderreihenfolgen ergeben einen Vergleichsmaßstab zu entsprechenden Rangfolgen auf der Basis des ebenfalls eindimensionalen Indikators BSP/Kopf. Synthetische eindimensionale Entwicklungsindikatoren sind von verschiedenen Autoren bestimmt worden. Dabei haben unterschiedliche R-Techniken zur Dimensionsreduktion Eingang gefunden. So benutzt *Scholing* (1982) das auch von *Adelman/Morris* (1965) herangezogene lineare Regressionsmodell der Faktorenanalyse. *Mc Granahan et al.* (1977) und *Sato* (1982) verbleiben mit ihren Ansätzen dagegen im Bereich der Linie S. Der von *Mc Granahan et al.* (1977) entwickelte Indikator beruht auf einem Korrespondenzsystem für die originären Entwicklungs- und Strukturindikatoren, das vom robusten Ansatz der ‚best fitting line‘ ausgeht. *Sato* (1982) berechnet synthetische Indikatorwerte als Elemente der ersten Hauptkomponente einer linearen Hauptkomponentenanalyse.

Gerade die Indikatororientierung der R-Techniken macht eine Berücksichtigung nichtlinearer Zusammenhänge zwischen einzelnen Indikatoren erforderlich. So „The relationship between two percentage-type indicators

²² Vgl. *Seber* (1984), 235 ff.

²³ Vgl. auch *Seber* (1984), 241.

²⁴ Vgl. *Mc Granahan et al.* (1977), 19.

tends to be a wavy line often approximating an S curve“²⁵. Daneben zeigt sich die Beziehung zwischen Prozent-Typ-Indikatoren und dem BSP/Kopf tendenziell als geometrisch mit hoher Progression²⁶. Vor diesem Hintergrund verwundert es, das *Mc Granahan et al.* (1977) auf lineare Zusammenhänge ausgerichtete gewöhnliche Korrelationskoeffizienten zur Indikatorauswahl nutzen. Man vergleiche dazu auch die kontroverse Diskussion in *Mc Granahan/Pizarro/Richard* (1981) sowie die Gewichtung einzelner originärer Indikatoren mit gewöhnlichen Korrelationskoeffizienten bei der Gewinnung des ‚general index of socio-economic development‘²⁷. Die mögliche Alternative in Form einer Verwendung von Rangkorrelationskoeffizienten wurde bereits in Zusammenhang mit der multidimensionalen Skalierung aufgezeigt. Rangstatistiken können aber auch an die Stelle originärer Indikatorwerte bei allen R-Techniken treten²⁸. So ist eine auf Rängen basierende lineare Hauptkomponentenanalyse zur Dimensionsreduktion in der Lage, Nichtlinearitäten zwischen originären Indikatoren einzubeziehen. Ranganalysen bieten sich auch an, wenn besonders für Indikatorwerte aus Ländern der Dritten Welt gilt, daß ihre „Zuverlässigkeit meist in direktem Verhältnis zur Entwicklungsstufe steht“²⁹. Qualität, Aktualität, Verfügbarkeit und Relevanz von Länderstatistiken der Dritten Welt³⁰ weisen Ranganalysen bei Ländervergleichen eine besondere Bedeutung zu. Andererseits ist der Übergang von quantitativen Indikatorwerten zu Rangstatistiken jedoch immer mit einem Informationsverlust verbunden. Dieser Informationsverlust läßt sich vermeiden, wenn eine Dimensionsreduktion im Indikatorraum über eine nichtlineare, d.h. verallgemeinerte Hauptkomponentenanalyse erfolgt³¹. Dieser R-Technik entspricht im Bereich der Q-Technik eine bisher in der Literatur nicht diskutierte nichtlineare, d.h. verallgemeinerte Hauptkoordinatenanalyse.

3. Verallgemeinerte Hauptkomponentenanalyse sozioökonomischer Entwicklungs- und Strukturindikatoren

Multivariate Verfahren zur Dimensionsreduktion in der vergleichenden Länderforschung sollen nach 2. deskriptiv bleiben und Nichtlinearitäten berücksichtigen. Beiden Ansprüchen genügt die bei *Gnanadesikan* (1977) entwickelte verallgemeinerte Hauptkomponentenanalyse. Angewandt auf Entwicklungs- und Strukturindikatoren, besteht dieses Verfahren in einer

²⁵ Vgl. *Mc Granahan / Pizarro / Richard* (1979), 15.

²⁶ Vgl. *Mc Granahan / Pizarro / Richard* (1979), 15.

²⁷ Vgl. *Mc Granahan et al.* (1977), 138.

²⁸ Vgl. *Conover / Iman* (1981).

²⁹ Vgl. *Nohlen / Nuscheler* (1974), 329.

³⁰ Vgl. *Hofmann* (1977).

³¹ Vgl. *Gnanadesikan* (1977), 53 ff.

linearen Hauptkomponentenanalyse mit nicht ausschließlich linearen Indikatoren. Bei einer quadratischen Hauptkomponentenanalyse können die einbezogenen Variablen neben den originären Indikatoren z.B. deren Quadrate sowie Produkte sein. Letztere repräsentieren Interaktionseffekte zwischen unterschiedlichen Indikatoren und stellen wie auch Indikatorquadrate zusätzliche Dimensionen von Entwicklung dar. Ob diese Dimensionen allerdings die genannten nichtlinearen Zusammenhänge zwischen einzelnen originären Entwicklungsindikatoren erfassen können, bleibt zu überprüfen. Ein auf *Wilk/Gnanadesikan/Huyett* (1962) zurückgehender Test gegen Nichtlinearität erweist sich hier als wenig hilfreich, da nach obiger Diskussion Bündel von Indikatoren nicht als multivariat normalverteilt angenommen werden können. Ebenso wenig läßt sich davon ausgehen, daß zwischen verschiedenen originären Indikatoren eine exakte nichtlineare Abhängigkeit besteht. Dieser von *Gnanadesikan*³² betrachtete Fall erlaubt, bei einer quadratischen Hauptkomponentenanalyse diejenigen Quadrate bzw. Produkte von Indikatoren zu identifizieren, welche in den Sätzen vorliegender Indikatorwerte keine eigene Dimension von Entwicklung aufspannen. So müssen alternative deskriptive Ansätze gefunden werden, mit denen vorhandene Interaktionseffekte zwischen Entwicklungs- bzw. Strukturindikatoren nachweisbar sind.

Besonders ein solcher Ansatz soll nachfolgend anhand eines Datensatzes von *Sato* (1982) illustriert werden. *Sato* (1982) führt auf der Basis von 32 originären Entwicklungs- und Strukturindikatoren mit Werten für 102 Länder eine lineare Hauptkomponentenanalyse durch. Dabei werden die standardisierten Indikatorwerte von 24 Indikatoren auf sechs Hauptkomponenten, die jeweils eine Gruppe von Indikatoren repräsentieren, verdichtet. Eine dieser Hauptkomponenten umfaßt die ‚Anteile des Exports und Imports am BSP‘ und steht damit für einen Handelsindex. Wie eine neue lineare Hauptkomponentenanalyse aller sechs Komponenten zeigt, spannt der Handelsindex eine eigene zweite Dimension von Entwicklung auf, während sich die anderen Indikatorgruppen unter den Stichworten Landwirtschaftsindex, Lebensstandardindex I, Ernährungsindex, Lebensstandardindex II und Bevölkerungsindex auf eine gemeinsame Dimension von Entwicklung reduzieren lassen³³. Dieses Resultat entspricht der Sonderrolle, die der Indikator ‚Exporte von verarbeiteten Gütern in Prozent des Gesamtexports‘ in der Unfolding-Analyse von *Tosstorff* (1983) spielt. Dort „wird die volle Struktur der Daten erst in der dreidimensionalen Konfiguration sichtbar“ ... und es ist die genannte Exportvariable „sehr stark mit der dritten Dimension assoziiert“³⁴. Entsprechend weist dieselbe Exportvariable für den ersten Faktor

³² Vgl. *Gnanadesikan* (1977), 55.

³³ Vgl. *Sato* (1982), 91.

³⁴ Vgl. *Tosstorff* (1983), 140

in der von *Scholing* (1982) vorgenommenen Faktorenanalyse eine im Vergleich zu anderen Indikatoren niedrige Faktorladung auf. „Dennoch überrascht das beobachtete ‚Ausmaß an Eindimensionalität‘“³⁵ von Entwicklung, welches sich bei *Sato* (1982) bzw. *Scholing* (1982) in durch die erste Hauptkomponente/den ersten Faktor erklärten Anteilen der gesamten Stichprobenvarianz aller einbezogenen Indikatoren von jeweils ca. 70 % äußert. Dieses Ergebnis veranlaßt *Sato* (1982), bei der Konstruktion seines sozialwissenschaftlichen Entwicklungsindex den Handelsindex zu unterdrücken. Analog finden acht in den Indizes der ersten Hauptkomponente nicht repräsentierte Entwicklungs- und Strukturindikatoren mit der von ihnen aufgespannten dritten Dimension von Entwicklung keine weitere Berücksichtigung. Der sozialwirtschaftliche Entwicklungsindex konzentriert sich auf die Werte und damit die eine Dimension der ersten ermittelten Hauptkomponente. Die Werte dieser Hauptkomponente werden jedoch ohne Beachtung zweifellos vorhandener Nichtlinearitäten aus den fünf verbliebenen Indexgruppen über eine lineare Hauptkomponentenanalyse gewonnen. Diese Hauptkomponentenanalyse, d. h. die Verdichtung der fünf hochkorrelierten Hauptkomponenten zu einer einzigen, soll nun als verallgemeinerte Hauptkomponentenanalyse wiederholt werden.

Ausgangsbasis für die nichtlineare Hauptkomponentenanalyse sind die standardisierten Komponentenwerte des Landwirtschaftsindex, der beiden Lebensstandardindizes I und II, des Ernährungs- und des Bevölkerungsindex³⁶. Von diesen $p = 5$ Variablen ausgehend, ermöglicht eine quadratische Hauptkomponentenanalyse die Einbeziehung von $p + p(p-1)/2$ zusätzlichen, daraus abgeleiteten Interaktionsvariablen bzw. Indikatorquadraten. Es kann also eine lineare Hauptkomponentenanalyse von zwanzig teilweise nichtlinearen Variablen vorgenommen werden. Bezogen auf die von *Sato* (1982) übernommenen Daten, besteht die gesamte Analyse in „first performing a linear principal components analysis and then pursuing nonlinear analysis“³⁷. Was Nichtlinearitäten anbetrifft, so sind Interaktionseffekte z. B. zwischen den beiden Lebensstandardindizes mit Indikatoren wie ‚Anzahl der Zeitungsabonnements je 1000 Einwohner‘ und ‚Einschulungsquote 5 - 19-jähriger‘ durch die Gruppierung originärer Indikatoren angelegt. Es bietet sich also an, die verallgemeinerte Hauptkomponentenanalyse unter Einbeziehung aller möglichen Indikatorquadrate und Interaktionen zu beginnen. Dem stehen auch absolut hohe gewöhnliche Korrelationen zwischen den fünf eingehenden Indizes³⁸ nicht entgegen. Diese Korrelationen bewirken, daß von den im Rahmen der Hauptkomponentenanalyse zu untersuchenden bis zu zwanzig Eigenwerten der Korrelationsmatrix (stan-

³⁵ Vgl. *Scholing* (1982), 215.

³⁶ Vgl. *Sato* (1982), 208 ff.

³⁷ Vgl. *Gnanadesikan* (1977), 60.

³⁸ Vgl. *Sato* (1982), 90.

dardisierte Daten!) viele wesentlich kleiner als Eins ausfallen. Eine Möglichkeit zur Dimensionsreduktion des nichtlinearen Variablenraumes ergibt sich daraus wie folgt: Man betrachtet den zum jeweils kleinsten Eigenwert gehörenden Eigenvektor und eliminiert diejenige Variable, welche in diesem Eigenvektor den absolut größten Koeffizienten aufweist. „This principle is consistent with the notion that we regard a component with small eigenvalue as of less importance, and consequently, the variable which dominates it should be of less importance or redundant“³⁹. Zur Auswirkung dieses Ansatzes auf die ‚Dimensionalität‘ des untersuchten Entwicklungsbegriffs vergleiche unten.

Wird auf diese Weise über eine quadratische Hauptkomponentenanalyse die Dimension des Variablenraumes auf die von *Sato* (1982) genutzte Zahl fünf reduziert, verbleiben die folgenden Variablen: der Landwirtschaftsindex, die Interaktionsvariable zwischen dem Lebensstandardindex I und dem Bevölkerungsindex sowie die Quadrate des Landwirtschaftsindex, des Ernährungsindex und des Lebensstandardindex II. Gleichzeitig reduziert sich der durch die jeweils erste Hauptkomponente erklärte Anteil der gesamten Stichprobenvarianz von 0,55 bei zwanzig Variablen auf unter 0,45. Erst durch Hinzunahme der zweiten Hauptkomponente ergeben sich wieder erklärte Anteile von ca. 70 % wie bei der linearen Hauptkomponentenanalyse für sechs Indikatorgruppen von *Sato* (1982). Das zu beobachtende ‚Ausmaß an Zweidimensionalität‘ mit dominierenden quadratischen Variablen mag auf deren Berücksichtigung in der nichtlinearen Hauptkomponentenanalyse zurückzuführen sein. Fehlspezifikationen sind hier nicht auszuschließen. Deswegen erfolgt eine weitere verallgemeinerte Hauptkomponentenanalyse, bei der zu Beginn lediglich sämtliche zehn möglichen Interaktionsterme als nichtlineare Variablen eingehen. Solche Interaktionsvariablen stellen nach obigen Aussagen ‚natürliche‘ Größen für die Erklärung aufsummierter Stichprobenvarianzen dar. Eine nach obigem Muster vorgenommene Dimensionsreduktion des nun von fünfzehn Variablen aufgespannten Variablenraumes auf die Dimension fünf läßt folgende Variablen übrig: den Landwirtschaftsindex, daneben die Interaktionsvariablen zwischen diesem und sowohl dem Ernährungs- als auch dem Lebensstandardindex II, aber auch die Interaktionsvariablen zwischen dem Bevölkerungsindex und dem Lebensstandardindex I bzw. dem Ernährungsindex. Es dominieren jetzt also die gegenüber einer Hauptkomponentenanalyse mit lediglich linearen Indikatoren zusätzlich aufgenommenen Interaktionseffekte. Der durch die erste Hauptkomponente erklärte Varianzanteil liegt nach der Dimensionsreduktion bei ca. 0,55, vorher bei 0,60. In beiden Fällen werden durch Hinzunahme einer zweiten Hauptkomponente jedoch über 77 % der gesamten Stichprobenvarianz erklärt. Gegenüber der Berücksich-

³⁹ Vgl. *Mardia / Kent / Bibby* (1979), 242.

tigung quadratischer Terme steigt also das genannte Ausmaß von Zweidimensionalität. Im Vergleich dazu liefert eine lineare Hauptkomponentenanalyse der bei *Sato* (1982) verbleibenden fünf Indexgruppen erklärte Varianzanteile von 0,81 bzw. 0,90 für die erste bzw. die ersten beiden Hauptkomponenten.

Wie sind diese Ergebnisse nun zu interpretieren? Die im Rahmen der verallgemeinerten Hauptkomponentenanalyse vorgenommene Dimensionsreduktion des Variablenraumes auf die Dimension fünf hat jeweils nur zu einer geringen Verminderung des durch die erste Hauptkomponente erklärten Varianzanteils geführt. Es sind also hinsichtlich des Ziels, die gesamte Stichprobenvarianz aller einbezogenen Variablen durch möglichst wenige Hauptkomponenten zu erklären, lediglich redundante Variablen eliminiert worden. Die nach der Beschränkung des Variablenraumes auf die Dimension fünf verbliebenen Variablen erweisen sich aber als extrem spezifikationsabhängig. Ein Spezifikationstest steht jedoch für die hier genutzte deskriptive R-Technik ohne weitere Modell- und Verteilungsannahmen nicht zur Verfügung.

Woraus ergeben sich dennoch Anhaltspunkte für die notwendige Aufnahme von Interaktionsvariablen, wenn deren Berücksichtigung in einer nichtlinearen Hauptkomponentenanalyse zu kleineren erklärten Varianzanteilen führt als im linearen Fall. Da ist zuerst das Ergebnis einer linearen Hauptkomponentenanalyse nach Rangtransformation der fünf von *Sato* (1982) untersuchten Indexgruppen zu nennen. Hier liegen die erklärten Anteile der gesamten Rangvarianz bei 0,83 bzw. 0,91 für die erste bzw. die ersten beiden Hauptkomponenten und damit oberhalb der Werte vor dem Übergang auf Rangdaten. Die Nichtlinearität einer solchen Rangtransformation stellt eine Alternative zur Berücksichtigung von Interaktionsvariablen in einer nichtlinearen Hauptkomponentenanalyse der nicht transformierten Indikatoren dar.

Ein weiterer Anhaltspunkt für eine Spezifikation mit Interaktionen ergibt sich, wenn eine Dimensionsreduktion des Variablenraumes bei der betrachteten verallgemeinerten Hauptkomponentenanalyse alternativ zum bisher geschilderten Ansatz erfolgt. Mit Blick auf *Gnanadesikan*⁴⁰ können diejenigen Variablen eliminiert werden, die zu den absolut kleinsten Komponenten des zum jeweils kleinsten Eigenwert gehörenden Eigenvektors korrespondieren. Sind diese Komponenten wie auch der jeweilige Eigenwert ungefähr Null, dann können die zugehörigen Variablen aus der Liste der zwanzig im Rahmen einer verallgemeinerten Hauptkomponentenanalyse anfänglich betrachteten Variablen gestrichen werden. Eine derart vorgenommene Dimensionsreduktion des Variablenraumes auf die Dimension fünf läßt fol-

⁴⁰ *Gnanadesikan* (1977), 55.

gende Variablen übrig: die Lebensstandardindizes I und II, deren Interaktionseffekt, das Quadrat des Lebensstandardindex I sowie die Interaktionsvariable zwischen dem Lebensstandardindex I und dem Ernährungsindex. Die durch die erste bzw. die ersten beiden Hauptkomponenten erklärten Anteile an der gesamten Stichprobenvarianz der verbliebenen Variablen liegen hier bei 0,75 bzw. 0,96 und damit im Bereich derjenigen einer linearen Hauptkomponentenanalyse. Im Gegensatz zur linearen und auch zu den vorher betrachteten verallgemeinerten Hauptkomponentenanalysen liegt der kleinste Eigenwert der zu untersuchenden Korrelationsmatrix jetzt mit 0,018 weit unterhalb von 0,1, der vorgefundenen Untergrenze bei fünf einbezogenen Variablen. Damit ist eine weitere Reduktion der Variablenzahl möglich, wie sich zeigt, sogar bis zur Dimension drei. Hier verbleiben dann die beiden Lebensstandardindizes zusammen mit ihrer Interaktionsvariablen. Unter diesen Variablen ist bei Verwendung des geschilderten Kriteriums keine auszumachen, deren korrespondierendes Element im zum kleinsten Eigenwert (0,022) gehörenden Eigenvektor als Null betrachtet werden könnte. Für eine weitere Dimensionsreduktion auf die Dimension zwei sind der Lebensstandardindex II und die Interaktionsvariable gleichrangige Kandidaten. Die erste Hauptkomponente erklärt hier 70 % der Gesamtvarianz von drei. Somit verbleibt eine wesentliche Interaktionsvariable als für die erste Hauptkomponente bestimmende Variable und beleuchtet damit die Rolle von Interaktionsvariablen bei der Dimensionsreduktion des Variablenraumes.

Betrachtet man die unterschiedlichen Vorgehensweisen zur Dimensionsreduktion, so ist festzuhalten⁴¹: Es wird jeweils diejenige Variable mit der größten bzw. diejenige mit der kleinsten gewöhnlichen Korrelation zwischen sich und der letzten Hauptkomponente eliminiert. Bei nahe Null liegendem Eigenwert sind beide Verfahrensweisen als äquivalent aufzufassen. Das zweite Verfahren entfernt immer Variablen, die mit den restlichen nahezu unkorreliert sind. Es beläßt also wesentlich diejenigen Variablen in der weiteren Analyse, welche einen Raum aufspannen, dessen ‚Ausmaß an Dimensionalität‘ möglichst klein ist. Dieses zeigt sich u. a. im Anstieg des durch die erste Hauptkomponente erklärten Varianzanteils, wenn die Dimension des Variablenraumes von zwanzig auf fünf reduziert wird. Folglich ‚minimieren‘ die am Ende der verallgemeinerten Hauptkomponentenanalyse verbliebenen Variablen das bereits mehrfach genannte ‚Ausmaß an Dimensionalität‘ auf nahezu Eins. Eine diese Dimension aufspannende Interaktionsvariable zeigt die zentrale Rolle von Interaktionseffekten für die Erklärung von Entwicklung als multidimensionalem Begriff. Entwicklung stellt sich somit als Begriff dar, dessen Dimension die Anzahl einbezogener Indikatoren übersteigt. Daß Entwicklung nicht eindimensional beschrieben

⁴¹ Vgl. *Seber* (1984), 176 ff.

werden kann, ergibt sich aus dem zuerst geschilderten Vorgehen, bei dem die eliminierten Variablen jeweils einen Variablenraum mit möglichst großem ‚Ausmaß an Dimensionalität‘ für die weitere Analyse belassen. Zudem ist daran zu erinnern, das *Sato* (1982) bei seiner linearen Hauptkomponentenanalyse den eine eigene Dimension von Entwicklung aufspannenden Handelsindex bewußt unterdrückt.

Die Wichtigkeit der durch die beiden Lebensstandardindizes und deren Interaktion aufgespannten Dimensionen von Entwicklung zeigt auch ein Blick auf die darin aggregierten einzelnen Indikatoren. So enthält der Lebensstandardindex I hauptsächlich ökonomische Variablen, neben dem kritisierten BSP/Kopf z.B. auch per capita Größen für den privaten Verbrauch und den Energieverbrauch, zusätzlich aber auch Angaben zur Ausstattung mit Kraftfahrzeugen und Radio- bzw. Fernsehgeräten. Der Lebensstandardindex II ist eher sozial orientiert und bündelt Indikatoren wie die Einschulungsquote, die Lebenserwartung bei Geburt, die Einwohner pro Arzt oder den Anteil der Bevölkerung in Großstädten⁴². Zusammen mit verschiedenen möglichen sozioökonomischen Interaktionen spannen diese Indikatoren, wie die verallgemeinerte Hauptkomponentenanalyse gezeigt hat, eine synthetische Dimension von Entwicklung auf. Das Ergebnis stimmt damit mit dem Resultat von *Scholing* (1982) überein, der allerdings davon ausging, „daß es zwei Faktoren gibt, einen ‚sozialen‘, und einen ‚ökonomischen‘“⁴³.

Trotz dieses Resultats haben die verschiedenen durchgeführten verallgemeinerten Hauptkomponentenanalysen bestätigt, daß die Reduktion von Entwicklung auf einen einzigen synthetischen und durch die erste Hauptkomponente repräsentierten Indikator problematisch ist. Auch wenn diese Hauptkomponente sozioökonomische Interaktionseffekte widerspiegelt, kann es ihr nicht gelingen, das vollständige Ausmaß multidimensionaler Entwicklungsbegriffe in sich zu vereinigen. Das gilt besonders, solange eine Analyse sozioökonomischer Entwicklungs- und Strukturindikatoren auf die Einbeziehung qualitativer Indikatoren verzichtet. Werden damit verknüpfte z.B. politische Dimensionen von Entwicklung ausgeschaltet, müssen Vergleiche von Länderrangfolgen auf der Basis synthetischer Entwicklungsindikatoren mit denen gemäß des BSP/Kopf⁴⁴ fruchtlos bleiben. Anstatt die Dimension untersuchter Entwicklungsbegriffe durch Anwendung multivariater statistischer Verfahren auf Eins zu reduzieren, sollte der Multidimensionalität von Entwicklung Rechnung getragen werden. Für Ländervergleiche ist dabei an die von *Nohlen/Nuscheler* vorgeschlagene quantifizierte Zeitreihenanalyse von Unterentwicklung bzw. Entwicklung zu erinnern⁴⁵.

⁴² Vgl. *Sato* (1982), 78 ff.

⁴³ Vgl. *Scholing* (1982), 210.

⁴⁴ Vgl. z.B. *Scholing* (1982), 217 ff und *Tosstorff* (1983), 138.

⁴⁵ Vgl. *Nohlen / Nuscheler* (1974), 356 ff.

Zusammenfassung

Multidimensionale sozioökonomische Entwicklungsbegriffe lassen sich u. a. über die Anwendung einer Hauptkomponentenanalyse auf niedrige Dimensionen reduzieren. Die hier genutzte quadratische verallgemeinerte Hauptkomponentenanalyse zeigt am Beispiel eines Ländervergleichs, daß Entwicklung notwendig mehrdimensional ist. Insbesondere spannen Interaktionen zwischen einzelnen Indikatoren für Entwicklung bzw. Unterentwicklung eigene Entwicklungsdimensionen auf.

Summary

Principal components can be used to represent essential dimensions of multidimensional socioeconomic notions of development. Here, a quadratic generalized analysis of principal components is performed when comparing the state of development for different nations. As a result, development turns out to be multidimensional indeed. Special dimensions of development are introduced by interactions between original indicators for development and underdevelopment, respectively.

Literatur

- Abel, P. (1985), Gesichter einer grundbedürfnisorientierten Ländergruppierung. Allgemeines Statistisches Archiv 69, 151 - 165.
- Adelmann, I. / Morris, C. T. (1965), Factor Analysis of the Interrelationship between Social and Political Variables and Per Capita Gross National Product. Quarterly Journal of Economics 79, 555 - 578.
- / — (1971), Analysis of Variance Techniques for the Study of Economic Development. Journal of Development Studies 8, 91 - 115.
- / — (1974), The Derivation of Cardinal Scales from Ordinal Data: An Application of Multidimensional Scaling to Measure Levels of National Development, in: W. Sellekaerts (Hrsg.): Economic Development and Planning. London, 1 - 39.
- Adelmann, I. / Morris, C. T. / Wold, S. (1980), Society, Politics and Economic Development Revisited: A Disjoint Principal-Components Analysis, in: L. R. Klein et al. (Hrsg.), Quantitative Economics and Development. New York, 2 - 18.
- Conover, W. J. / Iman, R. L. (1981), Rank Transformation as a Bridge Between Parametric and Nonparametric Statistics. American Statistician 35, 124 - 129.
- Gnanadesikan, R. (1977), Methods for Statistical Data Analysis of Multivariate Observations. New York.
- Hartung, J. / Elpelt, B. (1984), Multivariate Statistik. München.
- Hofmann, H. (1977), Die Statistik in der Dritten Welt: Probleme und Perspektiven. München.
- Mardia, K. V. / Kent, J. T. / Bibby, J. M. (1979), Multivariate Analysis. London.
- Mc Granahan, D. V. et al. (1974), Regression and the Best Fitting Line. UNRISD 73.2., Genf.
- (1977), Contents and Measurement of Socio-Economic Development. UNRISD 77 - 12485, Genf.

- Mc Granahan, D. / Pizarro, E. / Richard, C. (1979), Methodological Problems in Selection and Analysis of Socio-Economic Development Indicators. UNRISD 79.4., Genf.
- / — / — (1981), Development Statistics and Correlations: A Comment on Hicks and Streeten. *World Development* 9, 389 - 397.
- Nohlen, D. / Nuscheler, F. (Hrsg.) (1974), Handbuch der Dritten Welt. Bd. 1: Theorien und Indikatoren von Unterentwicklung. Hamburg.
- / — (1982), Handbuch der Dritten Welt. Bd. 1: Unterentwicklung und Entwicklung. Theorien - Strategien - Indikatoren. Hamburg.
- Sato, M. (1982), Sozialwirtschaftlicher Entwicklungsindex auf der Basis der Hauptkomponentenanalyse. Frankfurt.
- Scholing, E. (1981), Die hierarchische Interaktionsanalyse – Ein empirisches Klassifikationsverfahren für die vergleichende Entwicklungsländerforschung. *Statistische Hefte* 22, 280 - 315.
- (1982), Zur mehrdimensionalen Messung des wirtschaftlichen Entwicklungsstandes. *Kyklos* 35, 201 - 222.
- Seber, G. A. F. (1984), *Multivariate Observations*. New York.
- Siegel, A. F. (1982), Geometric Data Analysis. An Interactive Graphics Program for Shape Comparison, in: R. L. Launer / Siegel (Hrsg.), *Modern Data Analysis*. New York, 103 - 122.
- Tosstorff, G. (1983), Methoden der geometrischen Datenanalyse und ihre Anwendung bei der Untersuchung des Entwicklungsprozesses. Würzburg.
- Wilk, M. B. / Gnanadesikan, R. / Huyett, M. J. (1962), Probability Plots for the Gamma Distribution. *Technometrics* 4, 1 - 20.
- Wold, H. (1979), Multivariate Analyse, in: G. Menges (Hrsg.), *Ökonometrie und Statistik. Handwörterbuch der Mathematischen Wirtschaftswissenschaften* 2, 85 - 95.