SCHRIFTEN DER GESELLSCHAFT FÜR WIRTSCHAFTS- UND SOZIALWISSENSCHAFTEN DES LANDBAUES E.V.

Heiland, H.W.: Zur Entwicklung und Transformation orthogonaler Faktoren-Modelle. In: Reisch, E.: Quantitative Methoden in den Wirtschafts- und Sozialwissenschaften des Landbaues. Schriften der Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e.V., Band 4, Münster-Hiltrup: Landwirtschaftsverlag (1967), S. 205-217.

Zur Entwicklung und Transformation orthogonaler Faktoren-Modelle

Von Dipl.-Math. H. W. HEILAND, München-Weihenstephan

Einführung

Bei technischen Untersuchungen ist es im allgemeinen möglich derart zu experimentieren, daß man einen Faktor variiert, während alle übrigen konstant gehalten werden. Damit läßt sich der Einfluß eines einzigen Faktors auf das Untersuchungsergebnis eindeutig bestimmen. In der Betriebswirtschaft hingegen führt die Variation eines Faktors automatisch zur Änderung anderer Faktoren. Die Landwirtschaft ist überdies noch in besonderem Maße von klimatischen Verhältnissen abhängig, die der Unternehmer selbst in keiner Weise beeinflussen kann. Das Betriebsergebnis ist stets eine Folge des gesamten Komplexes der Änderungen aller Faktoren.
Faktorielle Modelle vermögen nun zur Beschreibung solcher komplexer Faktorenstrukturen beizutragen. Sie ermöglichen eine Antwort auf folgende Fragen zu finden: welches sind die wichtigsten Zusammenhänge innerhalb des Untersuchungsmaterials, welche Faktoren bilden eine in sich abgeschlossene Gruppe und wieviele solcher voneinander unabhängiger Teilbereiche existieren im gesamten Untersuchungskomplex? Der Modellvorstellung liegt im allgemeinen zu Grunde, daß jedes Merkmal des Untersuchungsobjektes als Linearkombination einer bestimmten Anzahl von — zunächst hypothetischen — Faktoren darstellbar ist. Man unterscheidet hierbei, entsprechend der bezüglich jedes Merkmals zugelassenen Menge von Faktoren, ein- zwei- und mehrfaktorielle Modelle, entsprechend der Eigenschaft der Faktoren, „orthogonale“ und

„schiefe“ Modelle. Bei orthogonalen Modellen sind sie Faktoren voneinander unabhängig. Die Einwirkungen der Faktoren auf die Merkmale, das sind die Koeffizienten des linearen Modells, geben an, welche Merkmale zusammenhängen und wie stark ihre Verbindungen untereinander sind.
Zur Bestimmung dieser Modellkoeffizienten gibt es eine große Anzahl von Methoden und Techniken, die unter dem Sammelbegriff „Faktoren-Analyse“ zusammengefaßt sind. Ausgangsdaten der Faktoren-Analyse sind die Varianz- und Kovarianzmaße bezw. die einfachen Korrelationskoeffizienten der beobachteten Merkmale. Merkmale einer Faktorenuntersuchung können Variablen des Einkommens, der Arbeitseffizienz, der Produktion, der Mechanisierung und vieler ähnlicher Größen mehr sein, sofern sie stochastischer Natur und meßbar sind. Aus dem resultierenden Faktoren-Schema lassen sich Strukturzusammenhänge entnehmen, die eine vertiefte Kenntnis der Zusammenhänge der Variablen innerhalb des Betriebsgeschehens vermitteln können. Jedoch — und das muß ausdrücklich betont werden — verlangen faktorielle Modelle ein beachtliches Maß an Geschicklichkeit sowohl bei der Wahl der Untersuchungsmerkmale als auch bei der Interpretation der Ergebnisse. Nur ein mit dem Untersuchungsgegenstand verstrauter Fachmann kann die Kette von Ursachen und Wirkungen aus dem Lösungsmodell der faktorisierten Variablen herauslesen. FRUCHTER [5, S. 4] äußert sich im Hinblick auf die Aussage, die aus dem Faktoren-Modell gewonnen werden kann, wie folgt: „Die Interpretation der Ergebnisse der Faktoren-Analyse ist, wie dies für jede wissenschaftliche Interpretation gilt, ein Versuch. Faktoren sind keine ewigen Wahrheiten.“

1 Begriffe und Bezeichnungen

Im Interesse einer auch für den „Nicht-Mathematiker“ verständlichen Schreibweise wird bei Gleichungssystemen auf die vereinfachte Darstellung mit Hilfe von Matrizen und Vektoren verzichtet. Statistische Variablen, das heißt nach TINTNER [16, S. 19] „Variablen, die mit gewissen Wahrscheinlichkeiten bestimmte Werte annehmen können,“ werden mit großen lateinischen Buchstaben bezeichnet und zur Unterscheidung voneinander mit einfachen Indizes versehen. Meß- oder Schätzwerte, die über solche Zufallsvariablen vorliegen, werden mit kleinen lateinischen Buchstaben bezeichnet und doppelt indiziert. Ist X_j eine bestimmte statistische Variable, so ist x_{ji} der i-te Beobachtungswert diesert Variablen.
Unter dem Mittelvert der n Beobachtungswerte der Variablen X_j wird die Größe

$$\bar{X}_j = \sum_{i=1}^{n} x_{ji}/n \tag{1.1}$$

verstanden. Die Abweichung vom Mittelwert wird für den i-ten Beobachtungswert der j-ten Variablen definiert durch

$$v_{ji} = x_{ji} - \bar{X}_j. \tag{1.2}$$

Diese Beziehung transformiert die Beobachtungswerte x_{ji} derart, daß die neuen Werte v_{ji} den Mittelwert null haben. Schließlich wird die Varianz einer Variablen X_j definiert durch

$$\sigma_j^2 = \sum_{i=1}^{n} v_{ji}^2/n, \tag{1.3}$$

wobei unter der Wurzel aus der Varianz in allgemeinen die Streuung verstanden wird.

Entsprechend versteht man unter der Kovarianz zweier Variablen X_j und X_k:

$$\sigma_{jk}^2 = \sum_{i=1}^{n} v_{ji} v_{ki}/n. \tag{1.4}$$

Dividiert man v_{ji} durch die Streuung σ_j, so erhält man folgende Transformationsbeziehung

$$z_{ji} = v_{ji}/\sigma_j, \tag{1.5}$$

die — wie man leicht einsieht — eine Menge von Werten z_{ji} $(i = 1, 2 \ldots n)$ erzeugt, deren Mittelwert null und deren Varianz die Einheit ist. Die Variable Z_j heißt normierte oder standardisierte Zufallsvariable.
Von grundlegender Bedeutung für die Faktoren-Analyse ist der Korrelationskoeffizient zweier Variablen X_j und X_k:

$$r(X_j, X_k) = \sum_{i=1}^{n} v_{ji} v_{ki} \Bigg/ \sqrt{\sum_{i=1}^{n} v_{ji}^2 \cdot \sum_{i=1}^{n} v_{ki}^2} \tag{1.6}$$

oder für Variablen in standardisierter Form:

$$r(Z_j, Z_k) = \sum_{i=1}^{n} z_{ji} z_{ki}/n. \tag{1.7}$$

Vergleicht man die Beziehungen (1.3) und (1.4) mit (1.7), so erkennt man, daß für standardisierte Variablen die Varianzmaße mit den Korrelationskoeffizienten übereinstimmen. Der Korrelationskoeffizient is eine Maßzahl für den Zusammenhang zwischen zwei Variablen und bewegt sich im Intervall von -1 bis $+1$ [12, S. 175].

2 Das Modell orthogonaler Faktoren

Obwohl man, wie bereits in der Einführung gesagt wurde, mannigfaltige Faktorenmodelle kennt [6, S. 119–227)], beschränkt sich die folgende Darstellung auf das Model orthogonaler Faktoren. Es wird also die Annahme unterstellt, daß die beobachteten Variablen als Linearkombinationen von wechselseitig unabhängigen Faktoren darstellbar sind. Diese Beschränkung des Modells besagt jedoch nicht, daß die folgenden grundlegenden Bemerkungen nur für orthogonale Modelle zutreffen. Orthogonalität heißt lediglich, daß die hypothetischen Faktoren nicht korrelieren.

2.1 Die logische Grundlage

Wenn zwei Merkmale korrelieren, so darf stets das Vorhandensein gemeinsamer Entstehungsbedingungen angenommen werden. Diese, die Korrelation bedingenden Ursachen sind die hypothetischen Faktoren. Im statistischen Sinne sind diese Faktoren ebenso wie die Merkmale als Zufallsvariablen aufzufassen. Ohne Einschränkung der Allgemeingültigkeit kann angenommen werden, daß Merkmale und Faktoren in standardisierter Form gegeben sind. Man definiert dann die Korrelation zweier variabler Merkmale Z_j und Z_k durch folgende Gleichung:

$$r(Z_j, Z_k) = a_{j1}a_{k1} + a_{j2}a_{k2} + \ldots + a_{jm}a_{km}. \tag{2.1}$$

Hierbei bedeutet für $p = 1, 2 \ldots m$
a_{jp} die „Einwirkung" des p-ten Faktors auf die Variable Z_j und
a_{kp} die „Einwirkung" des p-ten Faktors auf die Variable Z_k.

Diese Einwirkungen werden allgemeiner auch Faktorkoeffizienten genannt. Die Korrelation zweier Variablen wird demnach durch die Summe der Produkte der Koeffizienten ihrer gemeinsamen Faktoren gegeben.
Bezeichnet man diese Faktoren mit F_p, so entsprechen ihnen die hypothetischen Beobachtungswerte f_{pi} ($i = 1, 2 \ldots n$). Sie gestatten für jeden Beobachtungswert der zu untersuchenden Variablenmenge eine lineare Darstellung der Form

$$\begin{aligned} \hat{z}_{ji} &= a_{j1}f_{1i}+a_{j2}f_{2i}+\ldots+a_{jm}f_{mi} \\ \hat{z}_{ki} &= a_{k1}f_{1i}+a_{k2}f_{2i}+\ldots+a_{km}f_{mi} \end{aligned} \tag{2.2}$$

Das Zeichen wird benützt, um die theoretische, das heißt angenäherte Form der Beobachtungswerte zu verdeutlichen. Addiert man auf beiden Seiten von (2.2) Hilfsgrößen $a_j \cdot u_{ji}$ und $a_k \cdot u_{ki}$ derart, daß gilt

$$z_{ji} = \hat{z}_{ji}+a_j \cdot u_{ji}, \quad z_{ki} = \hat{z}_{ki}+a_k \cdot u_{ki},$$

so erhält man eine vollständige Beschreibung der wahren Beobachtungswerte durch

$$z_{ji} = \sum_{p=1}^{m} a_{jp}f_{pi}+a_j u_{ji}, \quad z_{ki} = \sum_{p=1}^{m} a_{kp}f_{pi}+a_k u_{ki}. \tag{2.3}$$

Den Werten u_{ji} und u_{ki} entsprechen zwei Variablen U_j und U_k. Da diese Hilfsvariablen U_j und U_k entsprechend der Annahme (2.1) nicht zum Aufbau der Korrelation zweier verschiedener Variablen Z_j, Z_k beitragen, können sie weder untereinander noch mit den Faktoren F_p korrelieren.
Bildet man numehr für alle n Beobachtungswerte die Produkte $z_{ji}\,z_{ki}$ (in der Darstellung (2.3)) und summiert über alle $i = 1, 2 \ldots n$ und beachtet, daß gilt $r(U_j, U_k) = r(U_j, F_p) = r(F_j, F_k) = 0$ für $j \neq k$ und $j \neq p$, dann erhält man:

$$\sum_{i=1}^{n} z_{ji}z_{ki} = a_{j1}a_{k1}\sum_{i=1}^{n} f_{1i}^2+\ldots+a_{jm}a_{km}\sum_{i=1}^{n} f_{mi}^2 \tag{2.4}$$

Dividert man diese Gleichung durch n und beachtet, daß die Varianz für die standardisierten Variablen F_p gleich der Einheit ist, so erhält man genau die Beziehung (2.1).
Das Vorhaben der Faktoren-Analyse besteht nun darin, ein Minimum an Faktoren F_p zu ermitteln, die das Zustandekommen der Korrelationen einer bestimmten Menge von Variablen hinreichend erklären. Diese Faktoren werden als „Mengenfaktoren" bezeichnet, da sie zur Beschreibung der Korrelationen nötig sind. Sie gehören stets mehreren Variablen an. Die Größen U_j werden „Einzelfaktoren" genannt, da sie gerade den Varianzanteil einer Variablen angeben, der nicht durch die Mengenfaktoren erklärt werden kann.

2.2 *Das Modell als System linearer Gleichungen*

Für s Variablen mit m gemeinsamen Faktoren und s Einzelfaktoren erhält man das vollständige Modell multipler Faktoren in folgendem System von s linearen Gleichungen:

$$\begin{aligned} Z_1 &= a_{11}F_1+a_{12}F_2+\ldots+a_{1m}F_m+a_1U_1 \\ Z_2 &= a_{21}F_1+a_{22}F_2+\ldots+a_{2m}F_m+a_2U_2 \\ &\vdots \\ Z_s &= a_{s1}F_1+a_{s2}F_2+\ldots+a_{sm}F_m+a_sU_s \end{aligned} \tag{2.5}$$

Für den i-ten Beobachtungswert ($i = 1, 2 \ldots n$) einer Variablen Z_j existiert im System (2.5) eine Beziehung der Form

$$z_{ji} = a_{j1} f_{1i} + a_{j2} f_{2i} + \ldots + a_{jm} f_{m1} + a_j u_{ji}, \tag{2.6}$$

welche der Darstellung (2.3) entspricht. Es wird folglich angenommen, daß es für jeden Faktor n hypothetische Beobachtungswerte gibt, die den Beobachtungswerten der Variablen zuordenbar sind. Demnach kann jede Gleichung des Systems (2.5) wiederum als System von n Gleichungen aufgefaßt werden. Multipliziert man nun das zu Z_j gehörende System mit dem Faktor F_p, das heißt multipliziert man zugeordnete Beobachtungswerte von Z_j und F_p, summiert über alle so erhaltenen n Gleichungen und dividiert das Ergebnis schließlich durch n, so erhält man

$$r(Z_j, F_p) = a_{j1} r(F_p, F_1) + a_{j2} r(F_p, F_2) + \ldots + a_{jm} r(F_p, F_m). \tag{2.7}$$

Berücksichtigt man noch die Annahme, daß die Mengenfaktoren nicht korrelieren, so verschwinden die Ausdrücke $r(F_p, F_q)$ für $p \neq q$ und der Ausdruck (2.7) vereinfacht sich zu:

$$r(Z_j, F_p) = a_{jp} r(F_p, F_p) = a_{jp}. \tag{2.8}$$

Wendet man diese Operation auf alle $Z_j (j = 1, 2 \ldots s)$ mit allen Mengenfaktoren F_p ($p = 1, 2 \ldots m$) und allen Einzelfaktoren U_j ($j = 1, 2 \ldots s$) an, so wird offensichtlich, daß die Modellkoeffizienten a_{jp} des Systems (2.5) gerade die einfachen Korrelationskoeffizienten der Variablen mit den Faktoren sind.

2.3 *Varianzkomponenten*

Die Varianz einer Variablen Z_j, deren Werte durch die Beziehung (2.5) als Linearkombinationen der Faktoren gegeben werden, läßt sich gemäß der Definition (1.3) durch Quadrieren, Summieren über alle n Beobachtungswerte und Dividieren durch n bestimmen zu:

$$\begin{aligned} \sigma_{Z_j}^2 = {} & a_{j1}^2 \cdot \sigma_{F_1}^2 + a_{j2}^2 \cdot \sigma_{F_2}^2 + \ldots + a_{jm}^2 \cdot \sigma_{F_m}^2 + a_j^2 \cdot \sigma_{u_j}^2 \\ & + 2 \cdot (a_{j1} a_{j2} r(F_1, F_2) + \ldots + a_{jm} a_j r(F_m, U_j)). \end{aligned} \tag{2.9}$$

Da alle Variablen in normierter Form gegeben sind, werden sämtliche Varianzmaße gleich der Einheit. Berücksichtigt man noch die Annahme, daß die Faktoren untereinander nicht korrelieren, so erhält man:

$$\sigma_{Z_j}^2 = 1 = a_{j1}^2 + a_{j2}^2 + \ldots + a_{jm}^2 + a_j^2. \tag{2.10}$$

Die Glieder der rechten Seite von (2.10) geben die Anteile der Einheitsvarianz von Z_j an, die den einzelnen Faktoren zugeschrieben werden können. Entsprechend der Unterscheidung von Mengen- und Einzelfaktoren, unterteilt man die Gesamtvarianz von Z_j in die Mengenvarianz, die durch die F_p erklärt wird, und die Einzelvarianz, die allein dem Faktor U_j zuordenbar ist. Die Mengenvarianz einer Variablen Z_j wird dann einfach gegeben durch die Summe der quadrierten Mengenfaktorkoeffizienten

$$h_j^2 = \sum_{p=1}^{m} a_{jp}^2. \tag{2.11}$$

Den gesamten Varianzanteil, den ein einziger Mengenfaktor F_p zur Erklärung der Varianz aller Variablen des Systems (2.5) beiträgt, wird gegeben zu:

$$V_{F_p} = \sum_{j=1}^{s} a_{jp}^2. \tag{2.12}$$

2.4 *Die Schätzung der Mengenvarianz*

Wie bereits gesagt wurde, besteht das wesentliche Problem der Faktoren-Analyse in der Bestimmung der Koeffizienten a_{jp} des Modells (2.5). Diese Koeffizienten lassen sich gemäß der Beziehung (2.1) aus den einfachen Korrelationen der Variablen ermitteln. Hierbei stellt sich allerdings folgendes Problem: für die Korrelation einer Variablen mit sich selbst gilt gemäß der Beziehung (2.1):

$$r(Z_j,\, Z_j) = \sigma^2_{Z_j} = \sum_{p=1}^{m} a^2_{jp}, \tag{2.13}$$

während die Beziehung (2.10) liefert

$$r(Z_j,\, Z_j) = \sigma^2_{Z_j} = \sum_{p=1}^{m} a^2_{jp} + a^2_j = 1 \tag{2.14}$$

Die Gleichungen (2.13) und (2.14) stimmen nur für $a_j = 0$ überein, das heißt, wenn keine Einzelvarianz vorliegt. Existiert hingegen ein Einzelfaktor, dann kann die Beziehung (2.13) nicht gelten. An ihre Stelle tritt die sogenannte „reduzierte" Form

$$r(Z_j,\, Z_j)' = r(Z_j,\, Z_j) - a^2_j = \sum_{p=1}^{m} a^2_{jp}, \tag{2.15}$$

das heißt, es wird ein um die Einzelvarianz reduzierter Korrelationskoeffizient angegeben. Hier nun tritt das bis heute noch nicht völlig gelöste Problem der Schätzung des Reduktionsmaßes a^2_j oder, was das Komplement ist, der Schätzung der Mengenvarianz auf [6, S. 87 ff.). Nur bei Kenntnis der Größen $r(Z_j, Z_j)'$ für alle $j = 1, 2 \ldots s$ läßt sich das vollständige System (2.1) für alle Variablen Z_j, Z_k (j, $k = 1, 2 \ldots s$) angeben. Der wohl am häufigsten benutzte Schätzwert für die Mengenvarianz einer Variablen Z_j ist der quadrierte multiple Korrelationskoeffizient $R^2_{Z_j} = R^2_{Z_j}, \ldots {}_{Z_{j-1} Z_{j+1} \ldots Z_s}$, dessen Berechnung bei ANDERSON [1, Seite 31–32] zu finden ist und für den DWYER [3] die Ungleichung

$$R^2_{z_j} \leqq h^2_j \tag{2.16}$$

nachgewiesen hat, die für s gegen Unendlich das Gleichheitszeichen annimmt.

2.5 *Die Nicht-Eindeutigkeit der Lösung*

In irgendeinem Untersuchungsbereich lassen im allgemeinen die beobachteten Phänomene mehrere Erklärungen zu, die alle ihre wohlbegründeten Vermutungen besitzen. Die Wahl einer bestimmten Theorie muß dabei stets von ihrer sinnvollen Anwendbarkeit auf die Wirklichkeit abhängen.

Das orthogonale Faktorenmodell ist nun in dem Sinne „unbestimmt", als bei gegebenen Korrelationen einer Menge von s Variablen (System (2.1)) die Koeffizienten nicht eindeutig bestimmt werden können. Der Beweis hierfür findet sich bei HOTELLING [8]. Im Hinblick auf die im folgenden Abschnitt 3 beschriebenen Transformationsverfahren ist folgende geometrische Interpretation dieses Sachverhaltes von Interesse. Für eine Menge von s nichtkorrelierenden Mengenfaktoren existieren im s-dimensionalen Raum, den diese Faktoren aufspannen, $(s^2-s)/2$ Freiheitsgrade für eine starre Rotation dieser Faktormenge in eine andere.

2.6 *Das Haupt-Faktoren-Modell*

Das Modell der Haupt-Faktoren entspricht der Darstellung (2.5) auf der Grundlage der Beziehung (2.1). Es wird also angenommen, daß jede der s Variablen als Linearkombination einer Menge von m wechselseitig unabhängigen Mengenfaktoren darstellbar ist. Die Einzelfaktoren werden im Lösungsmodell ausgelassen. Ihre Größe muß gemäß der Darstellung (2.15) bestimmt werden. Die Grundlage zum Modell und zur Lösungsmethode hat bereits 1933 HOTELLING [8] gelegt. Das Verfahren bestimmt die Koeffizienten der Mengenfaktoren derart, daß gilt: der erste Faktor F_1 hat den größten Anteil an der Mengenvarianz aller Variablen, das heißt, die Größe V_{F_1} der Beziehung (2.12) ist möglichst groß und größer als alle V_{F_p} für $p > 1$. Der zweite Faktor F_2 hat den größten Anteil an der möglichen Mengenvarianz des Restes, der verbleibt, wenn man den Varianzbeitrag des ersten Faktors unberücksichtigt läßt. Der dritte Faktor hat dann entsprechend den größten Varianzanteil an der dann noch verbleibenden Restvarianz. Der letzte Faktor schließlich enthält den kleinsten noch übrig gebliebenen Mengenvarianzanteil. Für die Varianzbeiträge aller Faktoren zu allen Variablen gilt dann im Haupt-Faktoren-Modell folgende Ungleichungskette

$$V_{F_1} > V_{F_2} > \ldots > V_{F_{m-1}} > V_{F_m}.$$

Bei HARMAN [6] sind Grundlage, Theorie und Technik der Haupt-Faktoren-Methode ausführlich dargestellt, sodaß hier auf eine Wiederholung verzichtet werden kann. Ebenso wird bei HARMAN auf die Bedingung hingewiesen, die bezüglich der Wahl der Mengenvarianz einzuhalten ist. Das in einer Matrix angeordnete System der in die Rechnung eingehenden Korrelationskoeffizienten darf nur nichtnegative Eigenwerte besitzen. Unterstellt man, daß es keine Einzelvarianzen gibt, das heißt, daß die Variation der beobachteten Variablen vollständig durch die Mengenfaktoren erklärt wird, dann ist diese Bedingung stets erfüllt. Andernfalls muß ihre Gültigkeit nachgeprüft werden. Das resultierende Haupt-Faktoren Modell kann dann etwa in folgender Form angegeben werden:

	F_1	F_2	...	F_m
Z_1	a_{11}	a_{12}	...	a_{1m}
Z_2	a_{21}	a_{22}	...	a_{2m}
⋮	⋮	⋮	...	⋮
Z_s	a_{s1}	a_{s2}		a_{sm}

(2.17)

3 Die Transformation orthogonaler Faktoren

Im Abschnitt 2.5 wurde dargelegt, daß die Lösung einer Faktoren-Analyse nicht eindeutig ist. Man kann folglich beliebig viele orthogonale Lösungsmodelle entwickeln, die vom theoretischen Standpunkt aus betrachtet, alle „gleich gut" sind. Es ist jedoch ein anerkanntes wissenschaftliches Vorgehen, stets die einfachste Form von möglichen Theorien zu wählen, wenn sich die Erscheinungsformen damit widerspruchsfrei erklären lassen. Einige spezielle Lösungsmodelle der Faktoren-Analyse können einfache Strukturverbindungen aufweisen, während andere ein kaum überschaubares Bild für den Zusammenhang der Variablen entwerfen. Deshalb war und ist das Ziel der Faktoren-Analyse, unter allen möglichen Modellen ein möglichst einfach strukturiertes zu bestimmen, dessen Anwendung auf die Wirklichkeit zu sinnvollen Ergebnissen führt. Der Begriff der „einfachen Struktur" ist dabei nicht eindeutig definiert. Von den fünf von

THURSTONE [14, S. 335] aufgestellten Bedingungen einer Einfachstruktur bis zu dem von HEILAND [7] vorgeschlagenen Normenkriterium existiert eine weite Skala von Definition für einfache Strukturen.

3.1 *Analytisch-geometrische Grundlagen*

Die $m+s$ unabhängigen Mengenfaktoren und Einzelfaktoren spannen einen $m+s$ dimensionalen Euklid'schen Raum auf. Jede Variable ist als Linearkombination dieser $m+s$ Faktoren ausgedrückt zu:

$$Z_j = a_{j1}F_1+a_{j2}F_2+\ldots+a_{jm}F_m+a_jU_j. \tag{3.1}$$

Die Vektordarstellung einer Variablen in diesem Raum wird dann gegeben durch:

$$Z_j = \{a_{j1}, a_{j2}, \ldots a_{jm}, o, \ldots a_j, \ldots o\} \tag{3.2}$$

Die Länge eines solchen Variablen-Vektors ist:

$$L_{Z_j} = \sqrt{a_{j1}^2+a_{j2}^2+\ldots+a_{jm}^2+a_j^2} = 1 \tag{3.3}$$

Beschränkt man die Darstellung auf den m-dimensionalen Mengenfaktorraum, so ist die Länge des Variablenvektors genau der Quadratwurzel aus der Mengenvarianz dieser Variablen. Die Richtungscosinuse zweier Variablen Z_j und Z_k werden dann für den gesamten Raum einfach gegeben durch die Endpunktkoordinaten a_{jp} und a_{kp} ($p = 1, 2 \ldots m+s$), für den Mengenfaktorraum hingegen zu:

$$\alpha_{jp} = a_{jp}/h_j, \alpha_{kp} = a_{kp}/h_k, p = 1, 2 \ldots m. \tag{3.4}$$

Der Cosinus des Winkels γ_{jk} zwischen zwei Variablen-Vektoren Z_j und Z_k wird im gesamten Faktorraum durch das Skalarprodukt ihrer Richtungscosinuse gegeben zu:

$$\cos\gamma_{jk} = \sum_{p=1}^{m+s} \alpha_{jp}\alpha_{kp} = \sum_{p=1}^{m+s} a_{jp}a_{kp} = r(Z_j, Z_k) \tag{3.5}$$

für den Mengenfaktorraum allein zu:

$$\cos\gamma_{jk} = \sum_{p=1}^{m} a_{jp}a_{kp}/h_jh_k \tag{3.6}$$

Im gesamten Faktorraum wird demnach die Korrelation zwischen zwei Variablen gegeben durch den Cosinus des Winkels, den sie einschließen. Der Ausdruck (3.6) gibt dann offensichtlich die durch die Mengenfaktoren erklärte Verknüpfung der Variablen an. Die Beziehungen (3.1) bis (3.6) lassen sich für den zweidimensionalen Mengenfaktorraum an Hand der folgenden Abb. 1 verdeutlichen:

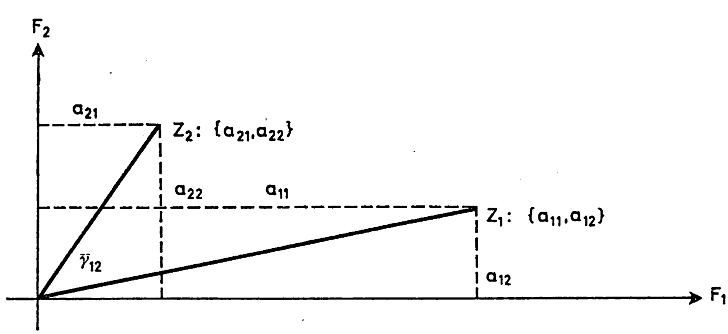

Abb. 1

Eine starre Drehung zweier Achsen in einem beliebig- dimensionalen Orthogonalsystem um einen Winkel δ ändert weder die Dimension noch die Orthogonalität des Systems. Man kann demnach jedes Hauptachsensystem durch Drehung zweier Achsen um den Ursprung in ein anderes Orthogonalsystem überführen.
Es seien F_k und F_l zwei Faktoren, die im Faktorraum Hauptachsen sind, und $a_{1k}, a_{2k}, \ldots a_{sk}$ bezw. $a_{1l}, a_{2l}, \ldots a_{sl}$ in Übereinstimmung mit dem Modell (2.5) ihre Koeffizienten. Jedes Koeffizientenpaar $\{a_{jk}, a_{jl}\}$ für $j = 1, 2 \ldots s$ liefert die Koordinaten der Variablen Z_j in der (F_k, F_l)-Ebene. Die Drehung der beiden Achsen F_k und F_l um den Ursprung mit dem Drehwinkel δ führt die Punkte $\{a_{jk}, a_{jl}\}$ in die neuen Punkte $\{a'_{jk}, a'_{jl}\}$ nach folgenden Transformationsgleichungen über

$$\begin{aligned} a'_{jk} &= a_{jk} \cdot \cos\delta + a_{jl} \cdot \sin\delta \\ a'_{jl} &= -a_{jk} \cdot \sin\delta + a_{jl} \cdot \cos\delta \end{aligned} \tag{3.7}$$

Die Abb. 2 veranschaulicht diesen Sachverhalt:

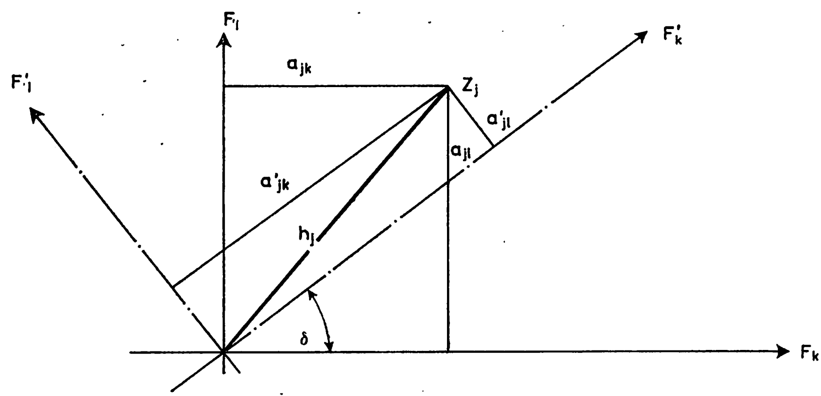

Abb. 2

Die Abb. 2 erlaubt noch den Beweis der Behauptung: die Mengenvarianzen $h_j^2 = \sum_{i=1}^{m} a_{ji}^2$ sind gegenüber Drehungen invariant. Die Anwendung des Pythagoräischen Lehrsatzes auf die Verhältnisse der Abb. 2 beweist, daß stets gilt:

$$h_j^2 = a_{jk}^2 + a_{jl}^2 = a'^2_{jk} + a'^2_{jl}. \tag{3.8}$$

3.2 *Die Quartimax-Methode*

Das erste analytische Kriterium zur Bestimmung einer Transformationslösung wurde 1953 von Carrol [2] aufgestellt.
Die fünf Bedingungen der Thurstoneschen Einfachstruktur [14, S. 335], und dabei vornehmlich die Forderung, daß für je zwei Faktoren möglichst viele Variablen durch den einen Faktor hoch durch den anderen jedoch nur gering erklärt werden sollen, mögen Carroll folgende Zielfunktion nahegelegt haben:

$$S_1 = \sum_{j}^{s} \sum_{s<t}^{m} a_{js}^2 a_{jt}^2 = \text{MIN}. \tag{3.9}$$

In der Folgezeit wurden von Neuhaus und Wrigley [11], Ferguson [4] und Saunders [13] Kriterien vorgestellt, die im Falle einer orthogonalen Faktorenstruktur der Forderung von Carrol äquivalent sind.

NEUHAUS und WRIGLEY gehen davon aus, daß ein Faktorenschema wohl dann am leichtesten zu interpretieren sei, wenn die Varianz aller $s \,.\, m$ quadrierten Koeffizienten des Modells (2.17) ein Maximum wird, das heißt, wenn

$$S_2 = \frac{1}{m \cdot s} \sum_j^s \sum_p^m (a_{jp}^2)^2 - \frac{1}{m^2 s^2} \left(\sum_j^s \sum_p^m a_{jp}^2 \right)^2 = \text{MAX.} \tag{3.10}$$

gilt.
FERGUSON simplifiziert die Forderung (3.10) zu

$$S_3 = \sum_j^s \sum_p^m a_{jp}^4 = \text{MAX.} \tag{3.11}$$

und SAUNDERS verlangt, daß das Verhältnis der vierten Momente zu den quadrierten zweiten Momenten ein Maximum wird:

$$S_4 = m \cdot s \sum_j^s \sum_p^m a_{jp}^4 \Big/ \left(\sum_j^s \sum_p^m a_{jp}^2 \right)^2 = \text{MAX} \tag{3.12}$$

Alle Kriterien dienen dem Zweck, ein Faktorenschema zu entwickeln, dessen Koeffizienten entweder sehr klein oder sehr groß sind. Daß die Forderungen (3.9) bis (3.12) für orthogonale Faktoren gleichwertig sind, folgt aus der Transformationsinvarianz der Summe der Quadrate aller Koeffizienten einer Zeile des Modells (2.17). (Man vgl. die Beziehung (3.8)). Es gilt nämlich:

$$\begin{aligned} \sum_j^s \left(\sum_p^m a_{jp}^2 \right)^2 &= \sum_j^s \sum_p^m a_{jp}^4 + 2 \sum_j^s \sum_{s<t}^m a_{js}^2 a_{jt}^2 = \text{constant} \\ &= S_3 + 2S_1 = \text{constant} \end{aligned} \tag{3.13}$$

und die Maximierung der FERGUSONschen Zielfunktion (3.11) ist der Minimierung der CARROLLschen Zielfunktion (3.9) äquivalent. Da mit $\sum_j \sum_i a_{ij}^2$ auch $\left(\sum_i \sum_j a_{ij}^2 \right)^2$ invariant ist, entsprechen die Zielfunktionen (3.10) und (3.12) ebenfalls der FERGUSONschen Zielfunktion (3.11). Alle Kriterien lassen sich demnach für ein orthogonales Faktorenschema etwa in der von FERGUSON vorgeschlagenen Form erfassen.
Dieses Vorgehen wird allgemein als „Quartimax Methode" bezeichnet, da die Summe der vierten Potenzen aller Koeffizienten maximiert wird.

3.3 *Die Varimax-Methode*

Die Quartimax-Methode schließt zwar alle Koeffizienten in ihre Betrachtung ein, jedoch wird das Kriterium der Varienzmaximierung auf jede Zeile des Modells (2.17) angewendet und dann über alle Zeilen summiert.
Während demnach bei den Kriterien (3.9) bis (3.12) die Betonung auf einer Vereinfachung jeder Zeile des Faktorenmodells liegt, das heißt, für jede Variable eine einfachere Darstellung durch die hypothetischen Faktoren gesucht wird, setzt sich KAISER [9] bei der Entwicklung der „Varimax Methode" das umgekehrte Ziel, nämlich durch geeignete Transformationen eine einfachere Darstellung der Faktoren selbst zu erhalten. Folgerichtig wird bei ihm nicht die Summe der Varianzen in den einzelnen Zeilen, sondern die Summe der Varianzen in den einzelnen Spalten maximiert:

$$\frac{1}{m} \sum_p^m \sum_j^s a_{jp}^4 - \frac{1}{m^2} \sum_p^m \left(\sum_j^s a_{jp}^2 \right)^2 = \text{MAX.} \tag{3.14}$$

Wurde die Annahme unterstellt, daß in der Beziehung (2.14) die Einzelvarianz $a_j^2 \neq 0$ ist, das hießt, daß die Mengenfaktoren nur einen Teil der Gesamtvarianz erklären, dann wird die Zielfunktion (3.14) durch die gemäß der Beziehung (3.8) gegebenen Mengenvarianzen dividiert:

$$\frac{1}{m}\sum_{p}^{m}\sum_{j}^{s}(a_{jp}^2/h_j^2)^2-\frac{1}{m^2}\sum_{p}^{m}\left(\sum_{j}^{s}a_{jp}^2/h_j^2\right)^2 = \text{MAX.} \tag{3.15}$$

KAISER [9] hat festgestellt, daß im allgemeinen die Quartimax- und die nicht gewichtete Varimax-Methode zu Faktorenschemen führen, in denen die Gesamtvarianzerklärungen derjenigen Faktorspalten, die im Ausgangssystem bereits relativ große Varianzbeiträge lieferten, noch vergrößert werden, während die Varianzsummen in den Spalten, die kleine Gesamtbeiträge zusteuerten, infolge der Transformation noch stärker verkleinert werden. Mit anderen Worten, auch die Streuung in den Gesamtvarianzbeiträgen der einzelnen Faktoren wird auf Grund einer der Forderungen (3.9) bis (3.14) vergrößert.

Auf Grund der gewichteten Beziehung (3.15) wird hingegen eine gleichmäßigere Verteilung der Varianzbeiträge aller Faktoren zur Varianz aller Variablen erreicht. KAISER [9, S. 191] sieht darin einen großen Vorteil, weil die Interpretation eines Faktorenschemas wohl dann wesentlich erleichtert wird, wenn in jeder Faktorenspalte sowohl sehr große als auch sehr kleine Elemente vorkommen.

3.4 *Das Normenkriterium*

Ein weiteres Kriterium wurde von HEILAND [7] vorgeschlagen, das im Unterschied zu den Varimax- und Quartimaxmethoden nicht zur Maximierung einer Varianz der Koeffizienten des Modells (2.17) führt, sondern fordert, daß jede Variable des Faktorenschemas vermöge eines Faktors so hoch wie möglich erklärt wird. In jeder Zeile des Modells soll demnach ein Element seinem Betrage noch möglichst groß gemacht werden, wobei — um eine simultane Betrachtung aller Variablen zu sichern — die Summe aller dieser Zeilenelemente selbst möglichst hoch werden soll.

Eine geometrische Interpretation vermag den gewünschten Sachverhalt anschaulich zu erklären. Die s Zeilenvektoren $(a_{i1}, a_{i2}, \ldots a_{im})$ haben die Längen h_i. Die Lage der m Faktorachsen ist dann so zu bestimmen, daß die Summe

$$\sum_{j}^{s}\left(h_j^2-\max_{p} a_{jp}^2\right) \tag{3.16}$$

möglichst klein wird, das heißt, daß die Summe der Abstände jedes Punktes im m-dimensionalen Faktorraum zur jeweils nächstgelegenen Achse ein Minimum wird. Da die Zeilenlängen h_i gegenüber Orthogonaltransformationen invariant sind, läßt sich folgende Zielfunktion angeben:

$$S = \|A\| = \left(\sum_{j}^{s}\max_{p} a_{jp}^2\right)^{\frac{1}{2}} = \text{MAX.} \tag{3.17}$$

Da hierbei eine spezielle Norm einer Matrix, die durch das Koeffizientensystem des Modells (2.17) gegeben ist, maximiert werden soll, wird das Kriterium (3.17) „Normenkriterium" genannt.

Bei der Behandlung dieses Transformationsproblems erhebt sich offensichtlich zunächst die Frage, welche Zeilenelemente des Modells (2.17) geeignet sind, damit die Zielfunktion (3.17) das absolut größte Maximum annehmen kann. Jede Auswahl von Zeilenelementen führt zu einem relativen Maximum. Jedoch kann diese — auf den ersten

Blick nachteilig erscheinende — Eigenschaft gerade in der praktischen Anwendung von besonderem Nutzen sein. Wie man zum absoluten Maximum gelangt, wird bei HEILAND [7] beschrieben.

Die Aussage eines Faktorenschemas mit absolut größter Norm (3.17) muß nicht notwendig besser und für die Beschreibung der Wirklichkeit geeigneter sein, als ein Faktorenschema, dessen zugehörige Norm auf Grund einer anderen Auswahl von Zeilenelementen des Modells (2.17) maximiert wurde. Hier wird dem Praktiker die Möglichkeit geboten, die Zeilenelemente nach Gesichtspunkten auszuwählen, die vom Untersuchungsgegenstand her nahegelegt werden. Sucht man zum Beispiel den Faktor, der den Zusammenhang zwischen einer bestimmten Auswahl von Merkmalen am stärksten zum Ausdruck bringt, so wird man die zu den Merkmalen gehörenden Elemente der entsprechenden Spalte des Schemas (2.17) auswählen und die zugehörige Norm (3.17) maximieren. Umgekehrt kann man sich fragen, welcher Faktor die Streuung einer Variablen möglichst hoch zu erklären vermag, ohne daß er gleichzeitig besonders stark auf andere Merkmale wirkt. Man wählt dann in einer Spalte des Schemas (2.17) nur das Element in der Zeile, welche dieser Variablen entspricht. So kann man bei der praktischen Anwendung dieser Transformationsmethode Elemente nach verschiedenen Aspekten auswählen und gewisse Charakteristika besonders herausarbeiten, wobei jedesmal ein — im Sinne des Normenkriteriums (3.17) — optimales Faktorenschema entwickelt wird.

3.5 *Ein Beispiel*

Das folgende Beispiel beruht auf einer Untersuchung der Zusammenhänge physischer Kenngrößen des Menschen, die bereits 1939 von MULLEN [10] mit Hilfe der Faktoren-Analyse vorgenommen wurde. Allerdings standen MULLEN zu diesem Zeitpunkt die beschriebenen Transformationskriterien noch nicht zur Verfügung. Es kann deshalb sehr interessant sein, auf einer von MULLEN angegebenen Hauptfaktoren-Lösung aufbauend, die entsprechend der drei beschriebenen Transformationskriterien entwickelten Rotationslösungen zu diskutieren. Tab. 1 enthält die Koeffizienten der Haupt-Faktoren sowie die auf Grund der drei Kriterien entwickelten Transformationslösungen. Offensichtlich lassen sich die untersuchten Merkmale in zwei Gruppen einteilen, wobei die ersten vier Merkmale typisch sein dürften für „Längenmaße", während die letzten vier Merkmale Eigenschaften von „Breitenmaßen" besitzen. Man erkennt, daß die angegebenen Lösungen in etwa diesen Vorstellungen entsprechen. Allerdings wirkt bei der Hauptfaktorenlösung der erste Faktor fast gleich stark auf alle Merkmalsgrößen, während der zweite Faktor auf die ersten vier Merkmale einen negativen, auf die letzten vier einen positiven Einfluß nimmt. Daß sich daraus zwei Faktoren entwickeln lassen, welche offensichtlich den Merkmalsträgern „Längenwachstum" und „Breitenwachstum" entsprechen, zeigen die Transformationslösungen. Die Faktoren dieser Lösungen lassen sich sehr leicht im angegebenen Sinne interpretieren, während die Eigenschaften der Faktoren der Hauptfaktorenlösung wesentlich unklarer sind. Noch deutlicher tritt dieser Sachverhalt in den Vordergrund, wenn man anstelle der Koeffizienten die Prozentsätze angibt, welche die Faktoren jeweils zur Erklärung der Gesamtvarianz der einzelnen Merkmale beitragen. (Das sind die quadrierten und mit 100 multiplizierten Faktorkoeffizienten). Tab. 2 weist diese Prozentzahlen aus, wobei Prozente unterhalb einer Grenze von 10 durch einen Punkt ersetzt wurden. Von Interesse für die Beurteilung der Güte der Transformationslösungen können auch die in der letzten Zeile von Tab. 1 angegebenen Varianzsummen sein, welche die einzelnen Faktoren zur Varianzerklärung aller Variablen beitragen. Während bei der Hauptfaktoren-Lösung der erste Faktor eine wesentlich größere Varianzsumme erklärt als der zweite Faktor, sind bei den

Tabelle 1 Hauptfaktorenlösung und drei Rotationslösungen (Faktor-Koeffizienten)

Variablen	Hauptfaktoren Faktoren I	II	Quartimax-kriterium Faktoren I	II	Varimax-kriterium Faktoren I	II	Normen-kriterium Faktoren I	II
1. Größe	,830	–,396	,899	,196	,879	,272	,887	,242
2. Spannweite	,818	–,469	,934	,131	,919	,210	,926	,179
3. Länge des Unterarms	,777	–,470	,902	,105	,890	,182	,895	,151
4. Länge des Unterschenkels	,789	–,401	,876	,172	,858	,246	,866	,217
5. Gewicht	,786	,500	,315	,877	,238	,900	,269	,892
6. Hüftweite	,672	,458	,250	,774	,183	,792	,210	,786
7. Brustumfang	,594	,444	,197	,715	,135	,729	,160	,724
8. Brustweite	,647	,333	,307	,660	,250	,684	,273	,675
$\sum_j a_{js}^2$	4,439	1,526	3,556	2,411	3,316	2,648	3,412	2,554

Tabelle 2 Hauptfaktorenlösung und drei Rotationslösungen (Prozente der Gesamtvarianz)

Variablen	Hauptfaktoren Faktoren I	II	Quartimax-kriterium Faktoren I	II	Varimax-kriterium Faktoren I	II	Normenkriterium Faktoren I	II
1. Größe	68,9	–15,7	80,8	.	77,3	.	78,7	.
2. Spannweite	66,9	–22,0	87,2	.	84,5	.	85,7	.
3. Länge des Unterarms	60,4	–22,1	81,4	.	79,2	.	80,1	.
4. Länge des Unterschenkels	63,7	–16,1	76,7	.	73,6	.	75,0	.
5. Gewicht	61,8	25,0	.	76,9	.	81,0	.	79,6
6. Hüftweite	45,2	20,2	.	59,9	.	62,7	.	61,8
7. Brustumfang	35,3	19,7	.	51,1	.	53,1	.	52,4
8. Brustweite	41,9	11,1	.	43,6	.	46,8	.	45,6

Rotationslösungen die Differenzen der Varianzsummen vom ersten und zweiten Faktor bedeutend kleiner. Bei den Transformationslösungen sind folglich die beiden ausgewiesenen Faktoren von nahezu gleichgroßer Aussagekraft, während dem zweiten Faktor bei der Hautfaktorenlösung eine nur geringe Bedeutung zukommt.

Zusammenfassung

1. Für die Untersuchung eines stochastischen mehrvariablen Phänomens kann die Faktoren-Analyse — mit ihren vielfältigen Methoden — stets dann Verwendung finden, wenn den Erscheinungsformen ein entsprechendes faktorielles Modell unterstellt werden darf.

2. Die Lösung eines orthogonalen Faktorensystems kann vermöge der HOTTELLING-schen Haupt-Faktoren-Methode ermittelt werden. Diese Lösung ist jedoch nur bis auf orthogonale Transformationen eindeutig.
3. Die auf speziellen Kriterien basierenden Transformationsverfahren führen zu Modellen, die „einfacher strukturiert„ und folglich im allgemeinen leichter zu interpretieren sind als das mehr oder weniger komplexe Ausganssystem.
4. Die Aussagekraft eines faktoranalytischen Resultates hängt sowohl von der Wirklichkeitsnähe des angenommenen faktoriellen Modells als auch von der Interpretationsgüte ab. Eine Orthogonaltransformation zerstört nicht die Eigenschaften eines Modells mit orthogonalen Faktoren. Da eine Transformation jedoch die Interpretationsmöglichkeiten zu verbessern vermag, führt sie unmittelbar zu einer verbesserten Gesamtaussage der faktoranalytischen Untersuchung.

Literatur

1. ANDERSON, O.: Probleme der statistischen Methodenlehre. Würzburg 1957
2. CARROLL, J. B.: An analytical solution for approximating simple structure in factoranalysis. Psychometrika *18*, 23 (1953)
3. DWYER, P. S.: The contribution of an orthogonal multiple factor solution to multiple correlation. Psychometrika *4*, 163, (1939)
4. FERGUSON, G. A.: The concept of parsimony in factor analysis. Psychometrika *19*, 281 (1954)
5. FRUCHTER: Introduction to factor-analysis. New York 1954
6. HARMAN, H. H.: Modern factor analysis. Chicago 1960
7. HEILAND, H. W.: Ein neuer Ansatz zur Orthogonaltransformation faktorieller Modelle, (unveröff. Manuskript). München 1966
8. HOTELLING, H.: Analysis of a complex of statistical variables into principal components. Educat. Psychol. *24*, 417 und 498, (1933)
9. KAISER, H. F.: The varimax criterion for analytic rotation in factor analysis. Psychometrika *23*, 187, (1958)
10. MULLEN, F.: Factors in the growth of girls seven to seventeen years of age. Diss. Chicago 1939
11. NEUHAUS, J. O.: The quartimax-method: an analytical approach to orthogonal simple structure. Brit. J. stat. Psychol. *7*, 81, (1954)
12. OSTLE, B.: Statistics in research. Ames, Iowa, 1960
13. SAUNDERS, D. R.: A computer programm to find the best-fitting orthogonal factors for a given hypothesis. Psychometrika *25*, 207, (1960)
14. THURSTONE, L. L.: Multiple factor analysis. Chicago 1947
15. DERS.: An analytical method for simple structure. Psychometrika *19*, 173, (1954)
16. TINTNER, G.: Handbuch der Ökonometrie. Berlin—Göttingen—Heidelberg 1960