e_i^2 = Restvarianz,

r_{ii}^2 = Reliabilität ($r_{ii}^2 \geq h_i^2$).

Die Durchführung einer Faktorenanalyse beginnt mit einer Datenmatrix: Für jede der n Personen (Fälle) sind Meßwerte auf i Variablen vorhanden. Aus dieser Matrix wird eine Korrelationsmatrix berechnet. In die Diagonale dieser Korrelationsmatrix werden Schätzungen der Kommunalitäten eingesetzt (Kommunalitätenproblem, vgl. 2. Option). Die neue Matrix wird nun faktorisiert, wobei als Ergebnis eine bestimmte Anzahl von voneinander unabhängigen Faktoren resultiert. Die Zahl der Faktoren kann nach unterschiedlichen Kriterien bestimmt werden (Faktorenproblem, vgl. 3. Option). Die sich nach der Extraktion ergebenden orthogonalen Faktoren sind nicht eindeutig. Um hier zu sinnvollen Ergebnissen zu gelangen, muß entschieden werden, nach welchem Kriterium die Faktorenmatrix (orthogonal oder schiefwinklig) rotiert werden soll (Rotationsproblem, vgl. 4. Option). Abschließend ist es noch möglich, zu einer Schätzung der Faktorenwerte für jede Person zu gelangen, um die Faktorenwerte für weitere Berechnungen nutzbar zu machen (Faktorwertschätzung, vgl. 5. Option).

Für die Lösung dieser Probleme und Fragen stehen jeweils mehrere Möglichkeiten zur Verfügung (vgl. 4.Optionentabelle).

1.1 Kommunalitätenproblem

(a) Diagonalelemente = 1. Es wird eine Hauptkomponentenanalyse gerechnet, bei der so viele Faktoren extrahiert werden als Variable verwendet wurden (siehe Beispiel 1). Dies ist zur Bestimmung des Scree-Tests (siehe Beispiel 1) notwendig.

(b) Iterative Kommunalitätenschätzung nach Festlegung der Faktorenanzahl (siehe Beispiel 2).

(c) Quadrate der multiplen Korrelationskoeffizienten (SMC) zwischen der Variablen i und den restlichen Variablen (siehe Beispiel 3). Die SMC stellen eine Schätzung der unteren Grenze der Kommunalitäten dar.

(d) Iterative Kommunalitätenschätzung nach KAISER-GUTTMAN, ausgehend von SMC-Schätzungen (siehe Beispiel 4) bzw. einem modifizierten Iterationskriterium.

1.2 Faktorenproblem

Als gebräuchliche Extraktionsmethode wird heute die von Hotelling entwickelte Hauptachsenmethode angewandt. Dabei wird die Lage der Faktoren so bestimmt, daß sie sukzessiv jeweils ein Maximum an Varianz abdecken. Algebraisch gesehen, entspricht dies der Eigenwertbestimmung (λ_i) der Korrelationsmatrix, wobei die Wurzel aus den Eigenwerten der Länge der jeweiligen Achsen entspricht. Die Summe der Eigenwerte entspricht bei einer vollständigen Lösung wiederum der Summe der Einheitsvarianz sämtlicher Variablen, und die einzelnen Eigenwerte verhalten sich proportional zur Einheitsvarianz.

Wesentlich sind aber Kriterien zur Bestimmung der Zahl der als bedeutsam erachteten Faktoren. Die verschiedenen Möglichkeiten werden im Zusammenhang mit Beispiel 1 beschrieben.

1.3 Rotationsproblem

Hier sind zwei grundsätzliche Rotationsmöglichkeiten zu unterscheiden, und zwar orthogonale und schiefwinkelige (oblique) Rotationsverfahren. Ferner können diese Verfahren von einem analytischen Rotationskriterium Gebrauch machen (z.B. Varimaxrotation, vgl. Beispiel 1 und 3), oder man kann versuchen, mittels einer rechtwinkligen oder schiefwinkeligen Rotation zur maximalen Deckung einer Faktorenstruktur auf vorher bestimmten Variablen zu kommen (kriteriumsbezogene Rotation, vgl. Beispiel 5). Wird ein schiefwinkeliges Rotationsverfahren gewählt, so bedeutet dies, daß die rotierten Faktoren miteinander korreliert sind. Die Korrelationsmatrix (zwischen den Faktoren) kann dann einer neuerlichen FA unterzogen werden (hierarchische FA; FAM realisiert diese Möglichkeit automatisch, siehe unter Beispiel 5). Bei der Bildung von Untergruppen und der Durchführung von mehreren FAs hintereinander wird mittels FISCHER-ROPPERT-Transformationen die Ähnlichkeit der Faktorenlösungen bestimmt (vgl. Beispiel 7). Mittels des Bargmann-Tests wird geprüft, ob nach der Rotation Einfachstruktur der Faktorlösung gegeben ist (vgl. Beispiel 1).

1.4 Faktorwertbestimmung

Eine FA ohne Bestimmung der Faktorwerte für jede Versuchsperson ist unvollständig. Die Werte auf den neuen nicht direkt meßbaren Variablen (Faktoren) sind es ja, die angeblich eine ökonomischere Beschreibung der ursprünglichen Daten ermöglichen und die als Grundlage für weitere Analysen (als abhängige Variablen) verwendet werden können.Die Faktorwerte können entweder über eine OLS - (ordinary least square) - oder eine GLS (generalized least square)-Schätzung berechnet werden (vgl. Optionentabelle). Die OLS-Schätzung setzt Gleichheit der Fehlervarianzen voraus (Homoskedastizität), was praktisch gleiche Kommunalitäten bedeutet; die GLS-Schätzung berücksichtigt zusätzlich unterschiedliche Fehlervarianzen (=Heteroskedastizität).Bei Durchführung einer Hauptkomponentenanalyse können ebenfalls Hauptkomponentenwerte berechnet werden. Als Ausgangsmatrix dient dabei die unrotierte Ladungsmatrix mit den Ladungen auf den Hauptkomponenten.

2 Dateneingabe und Beispiele

Die Faktorenanalyse kann über alle Einheiten (z.B. Versuchspersonen) gerechnet werden, deren Daten in einer Datei abgelegt sind (vgl. Beispiel 1), oder für bestimmmte Untergruppen (vgl. Beispiel 2). Nimmt man eine Aufteilung nach Untergruppen vor, wird für jede Gruppe eine eigene Faktorenanalyse gerechnet, die Faktorstrukturen der Gruppen werden verglichen, und es wird eine Ähnlichkeitstransformation durchgeführt (vgl. Beispiel 7).

2.1 Beispiel 1: Hauptkomponentenanalyse: Anzahl Faktoren gleich Variablenzahl, orthogonale Rotation.

Steuersätze:

```
EDATEI,2,0,0,0                              (1)
(1477-1481,1537,1538,1551-1555)             (2)
/*                                          (3)
```

In diesem Beispiel wird über alle Fälle der Datei DATEI eine Faktorenanalyse mit den Variablen gerechnet, deren Nummern als "abhängige" Variablen eingegeben werden (vgl. Satz 2). Aufgrund der Optionen (vgl. Satz

1) wird eine Hauptkomponentenanalyse durchgeführt (2. Option = 0) mit so vielen Faktoren wie Variablen (3. Option = 0), (orthogonaler) Varimaxrotation (4. Option = 0), ohne Berechnung von Faktorwerten (5. Option = 0) und bei Berücksichtigung von Faktorladungen >= .20 (6. Option = 0, daher Voreinstellung .20).

Abbildung XI-1

Beispiel 1, Hauptkomponentenanalyse: protokollierte Steuersätze und Registerauszug.

```
EINGABE... EDATEI N300, 2,0,0          Optionen
                                       N=Stichprobengröße
                                       E=Eingabe, Dateiname
Verzeichnis der Untergruppen
     1. Gruppe                         Variablen, die faktoren-
Eingabe...(1477-1481,1537,1538,1551-1555) analysiert werden sollen

Eingabe.../*                           Spaltenzahl
                                       untere Variablengrenze
                                       obere Variablengrenze
Neue   Alte Nr.

  1.    1477.  2  0  21   42  000  Leistungsangst'
  2.    1478.  2  0  10   20  000  Lernmoral'
  3.    1479.  2  0  10   20  000  Arbeitskapazität'
  4.    1480.  2  0   6   12  000  Schulinvolvement Einstellung zu Lehrern'
  5.    1481.  2  0   9   18  000  Schulinvolvement Wohlbefinden'
  6.    1537.  2  0   1   47  000  D9-1
  7.    1538.  2  0   1   46  000  F9                 Variablenbezeichnung
  8.    1551.  2  0   1   55  000  E9-5
  9.    1552.  2  0   1   19  000  M9-3-9-7
 10.    1553.  2  0   1   16  000  N9-6-PH
 11.    1554.  2  0   1   59  000  E9-4-Gesamt
 12.    1555.  3  0   1  115  000  D9-2 Gesamt
```

Im ersten Teil des Ausdruckes werden die Angaben der Steuersätze wiederholt. Zusätzlich werden die Variablen, welche in die Faktorenanalyse eingehen, aufgelistet, wobei die ursprüngliche Nummer auf der Datei (vgl. "alte Nr.") durch eine neue Nummer ersetzt wird. Es werden dabei Variablen, welche in dieser Analyse verwendet werden, nach der Reihenfolge in der Datei durchnummeriert.

Abbildung XI-2

Beispiel 1, Hauptkomponentenanalyse: Verteilungsparameter und Besetzungen der Variablen

```
neue Va-      Mittel-
riablen-      werte(AM)                          reale Be-
nummer   vgl.                  Stichproben-      setzung der
      Beispiel 2                  größe          Variablen
      1. Gruppe                N =     300
                                                              Standardabweichungen
Var. Nr   Mittelwert         Standardabw.    Besetzung        (s)
     1.       28.496              5.395        266
     2.       15.105              3.038        266
     3.       15.079              2.728        266
     4.        9.244              1.656        266
     5.       12.526              2.393        266
     6.       22.197              7.865        264
     7.       26.307              6.113        244
     8.       22.669              8.039        269
     9.        5.423              2.488        274
    10.        7.483              2.655        261
    11.       31.519              9.276        270
    12.       58.800             16.251        260
```

Als nächstes wird die Interkorrelationsmatrix der Variablen ausgedruckt. Bei der Berechnung der Korrelationskoeffizienten werden i.a. nur die Fälle (Personen) berücksichtigt, die bei den beiden jeweils korrelierenden Variablen zulässige Werte haben (d.h. innerhalb des Intervalls zwischen unterer und oberer Grenze liegen), Fälle mit K.A. auf einer der beiden Variablen bleiben also unberücksichtigt (Voreinstellung B-Option 1 = 4). Für eine aktzentuierte Optimierung hinsichtlich der Laufzeit läßt sich mittels B-Option 1 eine Berechnungsvariante wählen, die bei K.A. den Mittelwert der beantworteten Fälle der fraglichen Variable einsetzt. Da dieses Verfahren zur Folge hat, daß die Varianz tendenziell unterschätzt wird, empfiehlt es sich in erster Linie für umfangreiche Datensätze mit nur wenigen unbeantworteten Fällen in der aktuell analysierten (Teil-) Stichprobe.

Abbildung XI-3

Beispiel 1, Hauptkomponentenanalyse: Interkorrelationsmatrix

Interkorrelationsmatrix

KORRELATIONEN

Variablennummern

	1	2	3	4	5	6	7	8	9	10	11	12
1	1.00											
2	-0.04	1.00										
3	-0.51	0.44	1.00									
4	0.07	0.32	0.22	1.00								
5	-0.10	0.58	0.45	0.40	1.00							
6	-0.09	0.04	0.14	0.13	0.19	1.00						
7	-0.08	-0.04	0.09	0.15	0.10	0.28	1.00					
8	0.01	0.20	0.10	0.15	0.21	0.45	0.32	1.00				
9	-0.06	-0.08	0.18	0.03	0.02	0.07	0.22	0.22	1.00			
10	-0.26	-0.09	0.18	-0.06	-0.02	0.21	0.36	0.11	0.34	1.00		
11	0.04	0.16	0.13	0.15	0.16	0.38	0.32	0.71	0.28	0.11	1.00	
12	0.02	0.15	0.09	0.25	0.10	0.41	0.21	0.55	0.16	0.03	0.51	1.00

Da bei einer Hauptkomponentenanalyse keine Kommunalitätenschätzung erfolgt, werden nur die Eigenwerte (λ_i) ausgedruckt.

PRINCIPAL COMPONENTS - MODEL

EIGENWERTE

3.2641	1.9681	1.6591	0.9779	0.8922	0.7044
0.6259	0.5277	0.4347	0.3755	0.3130	0.2575

Die Eigenwerte sind der Größe nach geordnet und werden nach der Hauptachsenmethode berechnet.

2.1.1 Der Scree-Test

Zur Durchführung des graphischen Scree-Tests wird ein Eigenwertdiagramm ausgegeben.

Abbildung XI-4

Beispiel 1, Hauptkomponentenanalyse: Graphischer Scree-Test

```
EIGENWERTDIAGRAMM
  BIS    3.264*
  BIS    3.114
  BIS    2.963
  BIS    2.B13
  BIS    2.663
  BIS    2.512
  BIS    2.362
  BIS    2.212
  BIS    2.061
  BIS    1.911 *
  BIS    1.761
  BIS    1.610  *
  BIS    1.460
  BIS    1.310
  BIS    1.159
  BIS    1.009
---------------------------
  BIS    0.859    **
  BIS    0.708
  BIS    0.558      **
  BIS    0.408        **
  BIS    0.257          ***
```

"*" = Eigenwerte

Die Skalenweite des Eigenwertdiagramms wird bestimmt, indem von dem größten Eigenwert der kleinste abgezogen und diese Zahl durch 20 dividiert wird. Die so berechnete Schrittweite wird von dem größten Eigenwert 21 mal abgezogen.

Oberhalb dieser Linie liegen alle Eigenwerte größer/gleich 1.

Verlauf der Geröllgerade nach dem Knick im Eigenwertdiagramm.

Faktorisiert man Zufallsvariablen, so ergibt sich ein geradliniger bzw. kontinuierlicher Eigenwertverlauf mit Werten von 1,5 >= λ <= 0,5. Bei sinnvollen Daten liegen die Punkte des Eigenwertdiagramms nicht auf einer Geraden, sondern auf einer j-förmigen Kurve mit einem mehr oder weniger deutlich erkennbaren Knick. Erst nach dem Knick liegen die restlichen Eigenwerte wieder auf einer Geraden.

CATTELL (1966) schlug wegen dieses charakteristischen Verlaufes zur Bestimmung der Zahl der als wesentlich zu erachtenden Faktoren vor, die Anzahl der Eigenwerte bis einschließlich des ersten auf der Geraden liegenden heranzuziehen (Scree-Test). Durch die zusätzliche Berücksichtigung des auf der Geraden liegenden Eigenwertes sollte gewährleistet sein, daß nicht zu wenig Faktoren extrahiert werden. Nach CATTELL und JASPERS (1967) reicht es jedoch aus, für die Bestimmung der Faktorenzahl nur diejenigen heranzuziehen, deren Eigenwerte vor dem Knick liegen.

Dieser graphische Scree-Test führt zu identischen Ergebnissen hinsichtlich der zu extrahierenden Faktorenanzahl, wenn anstatt einer Hauptkomponentenanalyse (mit 1 als Diagonalelementen) eine andere Kommunalitätenschätzung verwendet wird. In diesen Fällen kann das Eigenwertkriterium (λ >= 1) nicht mehr zur Bestimmung der Faktorenzahl verwendet werden, da die Eigenwerte wesentlich niedriger als bei einer Hauptkomponentenanalyse ausfallen.

Die graphische Beurteilung des Eigenwertverlaufes kann dann problematisch werden, wenn (a) kein klarer Knick im Eigenwertverlauf erkennbar oder (b) mehr als nur ein Knick vorhanden ist. In diesen Fällen ist auf andere Kriterien zur Bestimmung der Faktorenanzahl zurückzugreifen.

2.1.2 Scree-Test zur numerischen Bestimmung der Abweichung von der Geröllgeraden

Abbildung XI-5

Beispiel 1, Hauptkomponentenanalyse: Numerischer Scree-Test

SCREE-TEST

ANZ.FAKT.	ABW. VON GERÖLLGERADE	
8	0.000002	0.000092
7	0.000094	0.000149
6	0.000243	0.000023
5	0.000267	0.001140
4	0.001407	0.000217
3	0.001623	0.026297
2	0.027920	0.019722
1	0.047642	0.137482

Differenz zwischen den Abweichungswerten bei der k-ten und k-1-ten Berechnung

Summe der Abweichungen von der geschätzten Geröllgeraden dividiert durch k - 1. Die Schätzung der Geröllgeraden beruht auf k Punkten (Eigenwerten).

Der Scree-Test kann nach einem Vorschlag von W.NAGL noch weitergeführt werden, indem aufgrund der Summe der durchschnittlichen quadrierten Abweichungen von der Geröllgeraden der Eigenwert gefunden wird, ab dem keine gute Anpassung an die Geröllgerade mehr möglich ist. Dabei wird, ausgehend von den Eigenwerten einer Hauptkomponentenanalyse, zuerst durch das letzte Drittel der Punkte des Eigenwertdiagramms eine Gerade gelegt. Für die Bestimmung der Geraden nach der Methode der kleinsten quadrierten Abweichungen werden die Eigenwerte als Y-Werte und die ordinalen Positionen der Eigenwerte als X-Werte verwendet. Sodann wird die Abweichung der k Eigenwerte von der Geraden wie folgt bestimmt:

$$\text{Abw} = \frac{\sum_{1}^{k} (\lambda_i - \hat{\lambda}_i)^2}{k - 1}$$

k = Anzahl der Eigenwerte, auf denen die Schätzung der Geröllgeraden beruht.

λ_i = tatsächliche Eigenwerte

$\hat{\lambda}_i$ = geschätzte Eigenwerte aufgrund des Verlaufes der Geröllgeraden

In Beispiel 1 beruht die Schätzung der Gerölleraden bei Faktorenanzahl 8 auf 4 Punkten (letztes Drittel der insgesamt 12 Eigenwerte). Die Anpassung an die Geröllgerade mit einem Abweichungswert von 0,000 002 ist gut, d.h. bei Annahme von 8 Faktoren würde man noch Faktoren interpretieren, die vor dem Knick im Eigenwerteverlauf liegen.

Das Verfahren wird so lange fortgesetzt, bis die Abweichung aller Eigenwerte von der sukzessiv neu berechneten Geröllgeraden bestimmt ist. Als bedeutsam kann nach einer Faustregel die Anzahl der Faktoren angesehen werden, ab der ein deutlicher Sprung in den Abweichungswerten auftritt (siehe Differenzwerte), die also von einem Verlauf der Geröllgeraden unter Zufallsbedingungen augenfällig abweichen. In dem Beispiel liegt mit Ausnahme des letzten Wertes der größte Sprung bei 3 Faktoren, d.h. man wird 3 Faktoren als wesentlich erachten.

2.1.3 Beurteilung der Residualkorrelationen nach schrittweiser Faktorenextraktion - Kriterium von SOKAL (vgl. PAWLIK 1971, S. 170 f.), modifiziert und erweitert durch W. NAGL

Abbildung XI-6

Beispiel 1, Hauptkomponentenanalyse: Sukzessive Faktorenextraktion nach Sokal

BEI	PROZENT RESTK.ZW. .1 U.-.1	POSITIVE RESTKORR.	SIGN.PAR T.KORR. (C.AKT)	SIGN.PAR T.KORR. (C.KONST.)	SIGN.PART. KORR.
11 FAKT.	100.0	47.0	89.4	100.0	0.114
10 FAKT.	98.5	43.9	93.9	100.0	0.114
9 FAKT.	97.0	43.9	80.3	100.0	0.114
8 FAKT.	97.0	34.8	84.8	100.0	0.114
7 FAKT.	92.4	39.4	72.7	100.0	0.114
6 FAKT.	90.0	34.8	69.7	100.0	0.113
5 FAKT.	80.3	40.9	62.1	100.0	0.113
4 FAKT.	74.2	40.9	63.6	100.0	0.113
3 FAKT.	75.8	36.4	63.3	100.0	0.113
2 FAKT.	62.1	33.3	62.1	100.0	0.113
1 FAKT.	34.8	37.9	66.7	100.0	0.112
0 FAKT.	33.3	81.8	62.1	100.0	0.112

(6) Höhe der signifikanten partiellen Korrelation

(5) Prozentsatz der signifikanten partiellen Korrelations-Koeffizienten ($r_{gi}.k$)

(4) Prozentsatz der signifikanten partiellen Korrelationskoeffizienten ($r_{gi.k}$) nach Extraktion von k Faktoren

(3) Prozentsatz der positiven Restkorrelationen ($_k r_{gi}$) nach Extraktion von k Faktoren

(2) Prozentsatz der Restkorrelationen zwischen +0,10 und -0,10 nach Extraktion von k Faktoren

(1) Anzahl der extrahierten Faktoren (k)

Erläuterung:

r_{gi}: Korrelation der standardisierten Variablen z_g und z_i.

r_{gi}: Durch k Faktoren "erklärte" Korrelation zwischen den Variablen g und i.

$_k r_{gi}$: Kovarianz der Restvariablen, die durch die Extraktion von k Faktoren noch nicht geklärt ist (=Restkorrelation)

$$_k r_{gi} = r_{gi} - \tilde{r}_{gi} = r_{gi} - \sum_{p=1}^{k} a_{gp} \cdot a_{ip}$$

$r_{gi.k}$: partielle Korrelation zwischen den Variablen g und i unter Konstanthaltung (= Extraktion) von k Faktoren

$$r_{gi.k} = \frac{_k r_{gi}}{\sqrt{u_g \cdot u_i}}$$

u_g, u_i: Einzelvarianz der Variablen g und i nach Extraktion von k Faktoren

$$u_g = 1 - h_g^2$$
$$u_i = 1 - h_i^2$$

(1) Anzahl (k) der extrahierten Faktoren

(2) Prozentsatz der Restkorrelationen ($_k r_{gi}$) zwischen +0,10 und -0,10 nach Extraktion von k Faktoren. Je größer dieser Prozentsatz, desto besser kann die ursprüngliche Korrelationsmatrix reproduziert werden bzw. desto weniger lohnend ist die Extraktion eines weiteren Faktors.

(3) Prozentsatz der positiven Restkorrelationen ($_k r_{gi}$) nach Extraktion von k Faktoren. Sind in der Restkorrelationsmatrix nur mehr zufällig von Null abweichende Koeffizienten enthalten, sollten sich jeweils gleich viele positive wie negative ergeben (~ 50 %).

(4) Prozentsatz der signifikanten partiellen Korrelationskoeffizienten ($r_{gi.k}$) unter Kontanthaltung von k Faktoren.

C.AKT: es wird die jeweils nach k Faktoren berechnete Kommunalität bzw. Einzelvarianz ($1 - h_i^2$) in die Formel für ($r_{gi} \cdot k$) eingesetzt.

Dieses Kriterium (Prozentsätze) steigt nicht linear an, da bei geringer Varianzaufklärung durch die gemeinsamen Faktoren die Einzelvarianz groß ist und es deshalb vorkommen kann, daß trotz hoher Restkorrelation im Zähler wegen der großen Einzelvarianz im Nenner eine kleine partielle Korrelation resultiert und umgekehrt. Zu erwarten ist deshalb ein kurvi-linearer Verlauf dieses Kriteriums. Nach einer Faustregel könnte man die Zahl der Faktoren heranziehen, bei denen das Minimum des Prozentsatzes an signifikanten partiellen Korrelationskoeffizienten liegt (in Beispiel 1: 2 Faktoren).

(5) Prozentsatz signifikanter Korrelationskoeffizienten nach Extraktion von k Faktoren.

C.CONST: es wird die Gesamtkommunalität in die Formel für $r_{gi.k}$ eingesetzt und nicht die nach k Faktoren extrahierte. In Beispiel 1 ergibt sich wegen der Hauptkomponentenanalyse und der damit verbundenen vollständigen Varianzaufklärung immer 100 %.

(6) Größe der signifikanten partiellen Korrelationskoeffizienten ($r_{gi.k}$).

$r^*_{gi.k}$: niedrigste partielle Korrelation, die nach Extraktion von k Faktoren signifikant von Null verschieden ist

$$r^*_{si.k} = \frac{t_\alpha - df}{t_\alpha^2 + 1}$$

df (Freiheitsgrade) = n - (k+2)

n = Stichprobengröße
k = Anzahl der extrahierten Faktoren

t_α = t-Wert, der für die Anzahl der Freiheitsgrade bei gewähltem Signifikanzniveau α eben noch signifikant ist.

Zur Rechenvereinfachung wird die Signifikanz der Restkorrelationen $_k r_{si}$ nach dem Verfahren von SOKAL (1959) bestimmt:

$_k r^*_{si}$ = untere Grenze für signifikante Restkorrelationen

$$_k r^*_{si} = r^*_{si \cdot k} \cdot \sqrt{u_a^2 \cdot u_b^2}$$

u_a, u_b = niedrigste Einzelrestvarianzen bei sämtlichen Variablen.

Alle Restkorrelationen <= $_k r^*_{si}$ sind sicher von Null verschieden. Die Restkorrelationen >=$_k r_{si}$ werden in partielle Korrelationskoeffizienten ($r_{si.k}$) umgerechnet. Ist der Prozentsatz signifikanter Restkorrelationen größer als α, muß mindestens ein weiterer Faktor extrahiert werden.

2.1.3.1 Weitere Methoden zur Bestimmung der Faktorenanzahl
Eigenwertkriterien (siehe auch Option 3)

Es werden nur die Faktoren extrahiert, deren Eigenwerte λ >= 1 sind. Bei kleiner Variablenanzahl werden bei diesem Kriterium eher zu wenig Faktoren, bei großer eher zu viele extrahiert. Im allgemeinen ist dieses Kriterium als Schätzung der unteren Grenze der Faktorenanzahl anzusehen (Ausnahme: sehr viele Variablen, z.B. bei der Analyse psychometrischer Skalen).

Bisweilen wird empfolen, nicht mehr als halb so viele Faktoren zu extrahieren als Variablen verwendet worden sind. Diese Regel kann (mit Ausnahmen) als eine Schätzung der oberen Grenze für die Faktorenzahl angesehen werden (s.u. 3.).

Faustregel: k >= m/2

k Faktorenzahl
m Anzahl Variablen

2.1.4 Varianzextraktion

Die Eigenwerte entsprechen dem Anteil der durch die ihnen korrespondierenden Faktoren aufgeklärten Anteile der Gesamtvarianz (s_i^2).

$$s_i^2 = \frac{\lambda_i * 100\ \%}{m}, \quad m = \text{Anzahl der Variablen}$$

Willkürlich wird dabei felstgelegt, daß 75 %, 80 %, 90 % etc. der Varianz der Korrelationsmatrix durch die Faktorenanalyse aufgeklärt werden soll. Bei einer Hauptkomponentenanalyse entspricht der Anteil der Gesamtvarianz dem der Gesamtkommunalität

$$(s^2 = \sum_{i=1}^{m} h_i^2),$$

bei anderen Kommunalitätenschätzungen ist die Gesamtvarianz i.a. größer als die durch die gemeinsamen Faktoren aufgeklärte Varianz

$$(s^2 > \sum_{i=1}^{m} h_i^2),$$

Ein hoher Anteil eines Faktors an der Gesamtkommunalität bedeutet in einem solchen Fall nicht unbedingt einen hohen Anteil an der Gesamtvarianz.

2.1.5 Die Faktormatrix (unrotiert) wird iterativ mittels der Hauptachsenmethode nach dem Verfahren von JACOBI berechnet.

Abbildung XI-7

Beispiel 1, Hauptkomponentenanalyse: Unrotierte Faktormatrix

FAKTORMATRIX (12 FAKTOREN) — unrotierte Faktormatrix (Anzahl der extrahierten Faktoren)

Kommunalitäten $(h_i^2 = \sum_{j=1}^{k} a_{ij}^2)$

Faktorenbezeichnung

VARI-ABLE	KOMM. H**2	FA.-LADUNGEN 1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
1.	1.00	-0.19	0.24	0.71	0.45	0.00	0.22	-0.25	0.05	-0.12	0.20	0.17	-0.12
2.	1.00	0.43	-0.69	0.18	0.03	0.15	0.30	0.0	0.28	-0.12	0.13	-0.29	0.12
3.	1.00	0.48	-0.56	-0.47	-0.12	0.16	-0.10	0.04	-0.05	-0.08	0.31	0.23	-0.19
4.	1.00	0.43	-0.38	0.27	0.44	-0.29	-0.49	0.04	0.04	0.27	0.04	-0.07	-0.01
5.	1.00	0.51	-0.64	0.06	0.15	-0.08	0.23	-0.16	-0.22	-0.05	-0.36	0.19	0.03
6.	1.00	0.60	0.22	0.02	-0.35	-0.39	-0.04	-0.47	-0.21	-0.06	0.15	-0.12	0.06
7.	1.00	0.50	0.33	-0.23	0.31	-0.45	0.19	0.43	-0.12	-0.24	0.06	-0.04	0.01
8.	1.00	0.76	0.29	0.25	-0.16	0.15	0.17	0.10	-0.02	0.18	-0.12	-0.14	-0.34
9.	1.00	0.36	0.31	-0.36	0.47	0.53	-0.17	-0.18	-0.21	-0.13	-0.03	-0.15	0.04
10.	1.00	0.30	0.29	-0.66	0.21	-0.16	0.17	-0.25	0.43	0.17	-0.06	0.08	-0.01
11.	1.00	0.74	0.31	0.23	-0.08	0.25	0.13	0.16	-0.07	0.27	0.13	0.18	0.27
12.	1.00	0.65	0.24	0.31	-0.19	0.08	-0.36	0.03	0.32	-0.33	-0.15	0.12	0.03
	% VAR.=100.0	27.2	16.4	13.8	8.1	7.4	5.9	5.2	4.4	3.6	3.1	2.6	2.1

Faktorenladungen (a_{ij})

Variablenbezeichnung

Varianzaufklärung durch die gemeinsamen Faktoren $(= \frac{\sum_{i=1}^{m} h_i^2}{m} \cdot 100)$

Varianzanteil, der durch einen Faktor aufgeklärt wird $(= \frac{\sum_{i=1}^{m} a_{ij}^2}{m} \cdot 100)$

ITERATION CYCLE	VARIANCES
0	0.102584
1	0.423510
2	0.616142
3	0.666352
4	0.668100
5	0.668116
6	0.668116
7	0.668116
8	0.668116
9	0.668116

Werte für das Varimaxkriterium nach i-ter Iteration (vgl. PAWLIK 1971, S. 207)

Die Kommunalität stellt den Varianzanteil dar, der durch die gemeinsamen Faktoren bei einer Variablen aufgeklärt wird. Dieser ist bei einer Hauptkomponentenanalyse immer 1 (=100%); bei anderen Analysen ist $1 = h_i^2 + u_i^2$ (u_i^2 = Einzel- oder Restvarianz einer Variablen, die nicht durch die gemeinsamen Faktoren aufgeklärt wird).

2.1.6 Varimaxrotierte Faktormatrix (siehe Option 4)

Die Varimax-Rotation wird nach dem von KAISER vorgeschlagenen normierten Varimax-Kriterium durchgeführt (vgl. PAWLIK 1971, S. 107). Die rechnerische Lösung erfolgt iterativ.

Abbildung XI-8

Beispiel 1, Hauptkomponentenanalyse: Rotierte Faktormatrix

ROTIERTE FAKTORMATRIX (12 FAKTOREN) — varimaxrotierte Faktormatrix (Anzahl der Faktoren)

Kommunalitäten nach der Rotation

Faktorenladungen nach der Rotation

Faktorenbezeichnung

VARI-ABLE	KOMM. H**2	FA.-LADUNGEN 1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
1.	1.00	0.01	0.03	0.96	-0.01	0.03	-0.05	0.04	-0.12	-0.01	-0.23	0.00	0.03
2.	1.00	0.08	-0.27	-0.00	-0.06	0.04	-0.14	0.01	-0.05	-0.06	0.20	-0.92	0.06
3.	1.00	0.01	-0.20	-0.30	0.10	-0.02	-0.10	-0.05	0.08	-0.02	0.89	-0.22	0.04
4.	1.00	0.03	-0.17	0.04	0.01	-0.07	-0.96	-0.04	-0.04	-0.10	0.08	-0.13	0.04
5.	1.00	0.07	-0.91	-0.04	0.00	-0.04	-0.20	-0.09	-0.02	-0.01	0.19	-0.28	0.05
6.	1.00	0.15	-0.08	-0.04	0.0	-0.11	-0.05	-0.95	0.09	-0.17	0.05	0.01	0.14
7.	1.00	0.10	-0.03	-0.03	0.09	-0.96	-0.07	-0.11	0.17	-0.07	0.02	0.03	0.12
8.	1.00	0.84	-0.08	0.01	0.10	-0.14	-0.04	-0.21	0.03	-0.26	0.01	-0.10	0.37
9.	1.00	0.07	-0.00	-0.01	0.97	-0.08	-0.01	0.0	0.16	-0.06	0.07	0.05	0.11
10.	1.00	0.03	0.02	-0.12	0.17	-0.17	0.04	-0.09	0.95	0.00	0.06	0.04	0.02
11.	1.00	0.29	-0.05	0.03	0.13	-0.14	-0.05	-0.16	0.03	-0.23	0.04	-0.07	0.89
12.	1.00	0.20	-0.01	0.01	0.07	-0.07	-0.12	-0.18	-0.01	-0.93	0.02	-0.05	0.21
	% VAR.=100.0	7.3	8.3	8.7	8.5	8.5	8.5	8.6	8.3	8.6	7.8	8.4	8.5

Matrix der varimax-orientierten Faktor-ladungen (a_{ij})

Variablenbezeichnung

Varianzaufklärung durch sämtliche gemeinsame Faktoren (bei Hauptkomponentenanalyse immer 100%).

Varianzanteil, der durch jeweils einen varimaxrotierten Faktor aufgeklärt wird (in % der Gesamtvarianz)

Die Varimax-Rotation wird nach dem von KAISER vorgeschlagenen normierten Varimax-Kriterium durchgeführt (vgl. PAWLIK 1971, S. 107 f). Die rechnerische Lösung erfolgt iterativ.

2.1.7 Bargmann-Test

Abbildung XI-9

Beispiel 1, Hauptkomponentenanalyse: Bargmann-Test mit Einfachstruktur

```
BARGMANN-TEST
*************

IN DEN FAKTOREN   ANZ.   NULLADUNGEN      P

          1               7        1.000
          2               8        1.000
          3               9        1.000
          4               9        1.000
          5               7        1.000
          6               8        1.000
          7               7        1.000
          8               8        1.000
          9               7        1.000
         10               8        1.000
         11               8        1.000
         12               6        1.000

                                     Signifikanzangabe
                                    über die Anzahl der Null-
                                    ladungen pro Faktor
                                    (Signifikanzniveau: 0,05
                                    bzw. darunter)

                              Anzahl der Nullladungen auf der
                          dem jeweiligen varimaxrotierten
                          Faktor entsprechenden Hyperebene

            Faktorenbezeichnung
```

Durch den Bargmann-Test wird geprüft, ob ein Faktor durch genügend viele Variablen definiert ist und eine Variable möglichst nur einen Faktor bedeutend lädt (Einfachstruktur). Dabei wird die Zahl der Variablen bestimmt, deren Ladungen auf der Hyperebene dieses Faktors in dem Bereich

$$\left|\frac{a_{ij}}{h_i}\right| < 0{,}10$$

liegen. Ist die Zahl der ermittelten Nullladungen größer als nach dem festgelegten Signifikanzniveau zu erwarten, kann dieser Faktor als gesichert und interpretierbar gelten.

In dem vorliegenden Beispiel kann nicht davon ausgegangen werden, daß Einfachstruktur vorliegt, da kein einziger Faktor durch genügend Nulladungen in der entsprechenden Hyperebene definiert ist.

2.1.8 Geordnete Ladungsmatrix

Die Variablen werden nach der Ladungshöhe auf den einzelnen Faktoren geordnet ausgegeben. Eine Variable wird dem Faktor zugeordnet, auf dem sie die höchste Ladung hat. Faktorenladungen, die unter dem nach Option 6 gewählten Minimum liegen (Voreinstellung 0,20), werden Null gesetzt. Die Variablen, welche einen Faktor am deutlichsten bestimmen, sind durch gepunktete Linien voneinander getrennt. Im Beispiel wird jeder Faktor durch nur eine Variable gebildet (z.B. Faktor 1 durch Variable 8, die darüber hinaus Faktor 7, 9 und 12 nennenswert lädt).

Abbildung XI-10

Beispiel 1, Hauptkomponentenanalyse: Geordnete Ladungsmatrix

DAS MINIMUM DER BERDCKSICHTIGTEN LADUNGEN LIEGT BEI: 0.20

Das Minimum der zu berück-sichtigenden Ladungen kann über Option 6 bestimmt werden.

Geordnete Matrix

Ladungen kleiner 0,20 werden Null gesetzt

Faktorbezeichnung

FAKTOR:	1	2	3	4	5	6	7	8	9	10	11	12
VAR.NR.												
8.	0.84.	D.0	0.0	0.0	0.0	0.0	-0.21	0.0	-0.26	0.0	0.0	0.37
5.	0.0	-0.91	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-0.28	0.0
1.	0.0	0.0	0.96.	0.0	0.0	0.0	0.0	0.0	0.0	-0.23	0.0	0.0
9.	0.0	0.0	0.0	0.97.	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
7.	0.0	0.0	0.0	0.0	-0.96.	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4.	0.0	0.0	0.0	0.0	0.0	-0.96.	0.0	0.0	0.0	0.0	0.0	0.0
6.	0.0	0.0	0.0	0.0	0.0	0.0	-0.95.	0.0	0.0	0.0	0.0	0.0
10.	0.0	0.0	0.0	0.0	0.0	9.0	0.0	-0.95.	0.0	0.0	0.0	0.0
12.	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-0.93.	0.0	0.0	0.21
3.	0.0	-0.20	-0.30	0.0	0.0	0.0	0.0	0.0	0.0	0.89.	-0.22	0.0
2.	0.0	-0.27	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-0.92.	0.0
11.	0.29	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-0.23	0.0	0.0	0.89.
% VAR.	7.3	8.3	8.7	8.5	8.5	8.5	8.6	8.3	8.6	7.8	8.4	8.5

Variablen-bezeichnung

Varianzanteil, der durch den jeweiligen Faktor aufgeklärt wird

2.2 Beispiel 2: Faktorenanalyse mit iterativer Kommunalitätenschätzung

Steuersätze:

```
(1)EDATEI,N300,2,1,0                      ( 1 )
(1-14,24)'TS-NW'                          ( 2 )
(1477-1481,1537,1538,1551-1555)           ( 3 )
/*                                        ( 4 )
```

In diesem Beispiel werden nur die Einheiten (Personen) in die Faktorenanalyse einbezogen, welche auf der Auswahlvariable (Variable 1, "unabhängige" Variable, "Filter" (s. Satz 1) die Ausprägungen 1,2, ... 14, 24 (s. Satz 2 Klammer) aufweisen. Diese sind als Gruppe unter der Bezeichnung 'TS-NW' zusammengefaßt (s. Satz 2, Text). Sollen aufgrund der Ausprägung auf der Auswahlvariablen mehrere Gruppen unterschieden werden, so sind die Ausprägungen durch "/" zu trennen, ebenso der Text zur Gruppenbezeichnung.

Beispiel: (1-14,24/15-23)'TS-NW/GS-NW'.

In diesem Fall wird jeweils eine getrennte Faktorenanalyse für die so definierte erste und zweite Gruppe gerechnet mit anschließender Ähnlichkeitsrotation (s.Beispiel 7). Die Kombination von mehreren Auswahlvariablen oder die Bildung weiterer Gruppen erfolgt nach der im Kapitel III, Steuersprache beschriebenen Syntax.
Steuersatz 3 gibt die alten Nummern der Variablen, über die die FA für die definierte Teilstichprobe gerechnet werden soll. Es sind dies dieselben zwölf Variablen wie in Beispiel 1.

2.2.1 Programmausgaben

Als erstes wird die durch FAM interpretierte Information der Steueransätze protokolliert, die Eingabe unter (1) wird als "unabhängige Variable, Optionen und Eingabedatei" ausgewertet usw. Im Anschluß daran wird ein Verzeichnis sämtlicher Variablen erstellt, die in die Analyse eingehen werden. Die Variablen werden dabei in der Reihenfolge, in der sie im ursprünglichen Datensatz enthalten sind, durchnummeriert.

Abbildung XI-11

Beispiel 2, FA einer Teilstichprobe und iterativer Kommunalitäten-schätzung: Protokollierte Steuersätze und Registerauszug

```
EINGABE..(1)EDATEI,N300,2,1,0                       Auswahlvariable
EINGABE...(1-14,24)'TS-NW'                          Text zur Bezeichnung der
      TS-NW                                         ausgewählten Einheiten
                                                    Ausprägungen auf der Auswahlvariab-
VERZEICHNIS DER UNTERGRUPPEN                        len für die Zuweisung zur 1. Gruppe
    1. GRUPPE: TS-NW
EINGABE...(1477-1481,1537,1538,1551-1555)
EINGABE.../*
NEUE   ALTE NR.
***************
   1.     1.35   3  0   0  150  000 SCHULNUMMER
   2.  1477.     2  0  21   42  000 LEISTUNGSANGST'
   3.  1478      2  0  10   20  000 LERNMORAL'
   4.  1479.     2  0  10   20  000 ARBEITSKAPAZITAT'
   5.  1480.     2  0   6   12  000 SCHULINVOLVEMENT EINSTELLUNG ZU LEHRERN'
   6.  1481.     2  0   9   18  000 SCHULINVOLVEMENT WOHLBEFINOEN'
   7.  1537.     2  0   1   47  000 09-1
   8.  1538.     2  0   1   46  000 F9
   9.  1551.     2  0   1   55  000 E9-5
  10.  1552.     2  0   1   19  000 M9-3-9-7
  11.  1553.     2  0   1   16  000 N9-6-PH
  12.  1554.     2  0   1   59  000 E9-4-GESAMT
  13.  1555.     3  0   1  115  000 D9-2 GESAMT
```

Die folgende Angaben zu den Berechnungsmodalitäten, die Ausgabe der Mittelwerte, Streuungen und Besetzungen pro Variable wurde weggelassen (siehe hierzu Beispiel 1), ebenso die Interkorrelationsmatrix.

Abbildung XI-12

Beispiel 2, FA einer Teistichprobe und iterativer Kommunalitäten und Eigenwerte

```
FAKTORENANALYSE
***************
KOMMUNALITATENSCHATZUNG
              ITERAT. NACH SCHATZEN DER FAKTOREN  ( FA= 12 )  k=Anzahl der Faktoren
EIGENWERTE VOR DER ITERATION
     4.055    2.190    1.274    0.930    0.701    0.672    0.558    0.411    0.374    0.350
     0.296    0.189
  1  0.0  1.00  1.00  1.00 1.00  1.00 1.00  1.00 1.00  1.00 1.00  1.00 1.00
              UBER ITERATION,  KRIT.= 0.1636594E-07  KRIT.MAX.= 0.5000000E-01   IT.CYCL.= 1
WERTE FUR DIE VAR.
 1.00  1.00  1.00  1.00  1.00  1.00  1.00  1.00  1.00  1.00  1.00  1.00

SUMME DER GESCHATZTEN KOMMUNALITATEN = 12.000


EIGENWERTE
    4.0553   2.1896   1.2736   0.9300   0.7006   0.6724   0.5581   0.4114   0.3740   0.3498
    0.2962   0.1890
```

2.3 Beispiel 3: Faktorenanalyse mit Kommunalitätenschätzung aus multiplen Korrelationskoeffizienten

Steuersätze:

Wie in Beispiel 2, aber Optionen 2,2,0 (1. Option = Faktoranalyse, 2. Option = Schätzung der Kommunalitäten über die Quadrate der multiplen Korrelationskoeffizienten, 3. Option = Abbruchkriterium: alle Eigenwerte größer Null).

Programmausgaben: bis einschließlich der Interkorrelationsmatrix analog zu Beispiel 2.

Abbildung XI-13

Beispiel 3, FA einer Teilstichprobe und iterativer Kommunalitätenschätzung: Kommunalitäten und Eigenwerte

FAKTORENANLYSE
****

KOMMUNALITÄTENSCHÄTZUNG
(MULT. KORR.-KOEFF.)**2
WERTE FÜR DIE VAR.

0.32 0.45 0.53 0.20 0.46 0.46 0.22 0.71 0.44 0.35 0.70 0.54

SUMME DER GESCHÄTZTEN KOMMUNALITÄTEN = 5.372

EIGENWERTE

3.5851 1.6465 0.6533 0.2787 0.0774 -0.0123

-0.0305 -0.0836 -0.1149 -0.1824 -0.1934 -0.2517

quadrierter multipler Korrelationskoeffizient der 1.Variable mit den restlichen Variablen als Schätzung der Kommunalität (= untere Grenze)

Summe der quadrierten multiplen Korrelationskoeffizient=Summe der geschätzten Kommunalitäten

Eigenwertdiagramm und numerischer Scree-Test werden ausgelassen (siehe hierzu Beispiel 1). Ein Ausdruck der schrittweisen Restkorrelationen erfolgt nicht, da die Größe der Eigenwerte als Abbruchkriterium festgelegt wurde (3. Option = 0.0).

Abbildung XI-14

Beispiel 3, FA einer Teilstichprobe und iterativer Kommunalitätenschätzung: unrotierte und varimaxrotierte Faktormatrix, Bargmann-Test

FAKTORMATRIX (5 FAKTOREN)

VARI-ABLE	KOMM. H**2	FA.-LADUNGEN 1.	2.	3.	4.	5.
2.	0.39	-0.22	-0.28	0.47	0.21	0.00
3.	0.52	0.26	0.61	0.25	0.04	0.07
4.	0.65	0.31	0.69	-0.23	-0.14	0.03
5.	0.25	0.25	0.33	0.16	0.21	-0.08
6.	0.56	0.30	0.66	0.16	0.07	-0.05
7.	0.51	0.67	-0.16	-0.01	-0.12	-0.17
8.	0.28	0.44	-0.12	-0.21	0.17	-0.06
9.	0.77	0.83	-0.19	0.17	-0.12	0.08
10.	0.51	0.65	-0.13	-0.15	0.20	0.11
11.	0.43	0.52	-0.07	-0.33	0.22	-0.01
12.	0.76	0.83	-0.18	0.14	-0.08	0.08
13.	0.60	0.72	-0.19	0.19	-0.09	-0.07
% VAR.=	52.0	29.9	13.7	5.4	2.3	0.6

Varianzanteil, der durch die gemeinsamen Faktoren aufgeklärt wird

ITERATION CYCLE	VARIANCES
0	0.222593
1	0.391914
2	0.413325
3	0.419882
4	0.429701
5	0.444354
6	0.449035
7	0.449079
8	0.449080
9	0.449080
10	0.449080
11	0.449080

Werte des Varimaxkriteriums

ROTIERTE FAKTORMATRIX (5 FAKTOREN)

VARI-ABLE	KOMM. H**2	FA.-LADUNGEN 1.	2.	3.	4.	5.
2.	0.39	-0.03	-0.07	0.60	-0.16	0.01
3.	0.52	0.09	0.70	-0.11	-0.08	0.10
4.	0.65	0.02	0.55	-0.58	0.06	0.04
5.	0.25	0.08	0.48	0.05	0.12	-0.06
6.	0.56	0.07	0.73	-0.18	0.0	-0.02
7.	0.51	0.64	0.04	-0.12	0.25	-0.16
8.	0.28	0.27	-0.00	-0.08	0.45	-0.06
9.	0.77	0.83	0.11	-0.01	0.23	0.10
10.	0.51	0.45	0.08	-0.06	0.53	0.12
11.	0.43	0.26	0.04	-0.17	0.58	-0.01
12.	0.76	0.81	0.12	-0.02	0.27	0.10
13.	0.60	0.74	0.10	0.03	0.19	-0.05
% VAR.=	52.0	22.3	13.2	6.7	9.1	0.7

BARGMANN-TEST
****

IN DEN FAKTOREN	ANZ. NULLADUNGEN	P
1	3	1.000
2	3	1.000
3	5	0.726
4	2	1.000
5	5	0.726

nicht hinreichend bestimmte Faktoren

2.4 Beispiel 4: Kommunalitätenschätzung nach Iterationsverfahren von KAISER und GUTTMAN

Steuersätze:

Wie in Beispiel 2, aber Optionen 2,3,0 (2. Option = Schätzung der Kommunalitäten nach dem Iterationsverfahren von KAISER und GUTTMAN vgl. PAWLIK 1971, S. 121 f., modifiziert nach W.NAGL).

Programmausgaben: Bis einschließlich der Interkorrelationsmatrix analog zu Beispiel 2.

Abbildung XI-15

Beispiel 4, FA einer Teilstichprobe und Kommunalitätenschätzung nach Kauser-Guttman: Kommunalitäten und Eigenwerte

```
FAKTORENANALYSE
***************
KOMMUNALITÄTENSCHÄTZUNG
 1  0.11 0.86 0.89 0.91 0.84 0.89 0.89 0.84 0.94 0.89 0.87 0.94 0.91
 2  0.09 0.76 0.81 0.84 0.72 0.81 0.81 0.72 0.90 0.80 0.77 0.90 0.84
 3  0.06 0.70 0.76 0.80 0.62 0.76 0.75 0.63 0.87 0.74 0.70 0.87 0.79
 4  0.04 0.66 0.73 0.77 0.55 0.72 0.71 0.56 0.85 0.70 0.65 0.85 0.75
CON= 0.2000
              ÜBER ITERATION, KRIT. =  0.4375706E-01
                 KRIT.MAX. =  0.5000000E-01   IT.CYCL. = 4
WERTE FÜR DIE VAR.
  0.66  0.73  0.77  0.55  0.72  0.71  0.56  0.85  0.70  0.65  0.85  0.75
SUMME DER GESCHÄTZTEN KOMMUNALITÄTEN =     8.493

EIGENWERTE
     3.8038      1.9106      0.9494      0.5727      0.3453      0.2894
     0.2371      0.1527      0.0891      0.0833      0.0414      0.0183
```

geschätzte h_i^2 (zu 0.65)

$\sum_{i=1}^{m} h_i^2$ (zu 8.493)

Konstante für die Bestimmung der Kommunalitäten nach dem KAISER-GUTTMANN-Verfahren. Im Gegensatz zu den Angaben von PAWLIK ist es bei Verwendung der quadrierten multiplen Korrelationskoeffizienten als Ausgangsschätzung nicht möglich, eine Konvergenz zu erzielen, da die Matrix singulär wird. Deshalb wird als Schätzung der Ausgangswerte von dem Wert 1 ein Fünftel der Differenz zwischen 1 und den SMC's abgezogen. Für die nächsten Iterationszyklen wird analog verfahren, bis eine hinreichend gute Konvergenz erzielt wird.

Summe der quadrierten Differenzen zwischen Kommunalitäten der n-ten und n+1-ten Iteration. Unterschreitet diese Summe das gesetzte Kriterium (=0,05), so wird mit der Iteration a ge rochen.

Unrotierte, rotierte Faktormatrix und BARGMANN-Test werden ausgelassen, ebenso die geordnete Ladungsmatrix

2.5 Beispiel 5: Kriteriumsbezogene Rotation (Analyse B) als Versuch der Replikation der Faktorenstruktur bei einer anderen Stichprobe (Analyse A).

Analyse A: Zur Bestimmung der Faktorenstruktur der Variablen wird an einer ersten Stichprobe eine Faktorenanalyse gerechnet.

Abbildung XI-16

Beispiel 5, kriteriumsbezogene FA: protokollierte Steuersätze und Registerauszug von Analyse A (Varimaxrotation).

iterative Kommunalitätenschätzung (zu 3)

Abbruchkriterium: Eigenwerte größer 1 (zu 1.0)

```
EINGABE...(1)EDATEI,2,3,1.0,N200

EINGABE...(1-14,24)'TS-NW'

          TS-NW


VERZEICHNIS DER UNTERGRUPPEN
     1. GRUPPE: TS-NW
EINGABE...(1477-1481,1537,1538,1551-1555)

EINGABE.../*

NEUE    ALTE NR.
****************
    1.      1.35   3   0    0  150   000  SCHULNUMMER
    2.    1477.    2   0   21   42   000  LEISTUNGSANGST
    3.    1478.    2   0   10   20   000  LERNMORAL'
    4.    1479.    2   0   10   20   000  ARBEITSKAPAZITAT
    5.    1480.    2   0    6   12   000  SCHULINVOLV.
    6.    1481.    2   0    9   18   000  SCHULINVOLV.
    7.    1537.    2   0    1   47   000  09-1
    8.    1538.    2   0    1   46   000  F9
    9.    1551.    2   0    1   55   000  E9-5
   10.    1552.    2   0    1   19   000  M9-3-9-7
   11.    1553.    2   0    1   16   000  N9-6-PH
   12.    1554.    2   0    1   59   000  E9-4-GESAMT
   13.    1555.    3   0    1  115   000  D9-2 GESAMT
```

Analyse B: Nachdem an einer Stichprobe die Faktorenstruktur der Variablen ermittelt wurde, wird diese an einer anderen Stichprobe "GS-NW" (s. Steuersatz 2) zu replizieren versucht.

Abbildung XI-17

Beispiel 5, kriteriumsbezogene FA: protokollierte Steuersätze und Registerauszug von Analyse B (kriteriumsbezogene Rotation)

```
                                 iterative Kommunalitätenschätzung
                                   Eigenwertkriterium entfällt
                                     schiefwinkelige kriteriumsbezogene Rot.
EINGABE...(1)EDATEI,N200,2,3,0,2

EINGABE...(15-23)'GS-NW'

         GS-NW
                        "/" Trennungskriterium
VERZEICHNIS DER UNTERGRUPPEN
     1. GRUPPE: GS-NW
EINGABE...(1537,1538,1551-1555/1477-1481)

EINGABE.../*    Variable,          Variable, welche
             welche auf dem     auf dem 2. Faktor
             ersten Faktor      hoch laden sollen
NEUE    ALTE Nr.
****************
    1.      1.35   3   0    0   150   000  SCHULNUMMER
    2.   1477.     2   0   21    42   000  LEISTUNGSANGST
    3.   1478.     2   0   10    20   000  LERNMORAL'
    4.   1479.     2   0   10    20   000  ARBEITSKAP.
    5.   1480.     2   0    6    12   000  SCHULINVOLV.
    6.   1481.     2   0    9    18   000  SCHULINVOLV.
    7.   1537.     2   0    1    47   000  D9-1
    8.   1538.     2   0    1    46   000  F9
    9.   1551.     2   0    1    55   000  E9-5
   10.   1552.     2   0    1    19   000  M9-3-9-7
   11.   1553.     2   0    1    16   000  N9-6-PH
   12.   1554.     2   0    1    59   000  E9-4-GESAMT
   13.   1555.     3   0    1   115   000  D9-2 GESAMT
```

Ausdrucke der Mittelwerte, Standardabweichung und Besetzungen pro Variable werden ausgelassen, ebenso die beiden Interkorrelationsmatrizen.

Ausdrucke der Mittelwerte, Standardabweichung und Besetzungen pro Variable werden ausgelassen, ebenso die beiden Interkorrelationsmatrices.

Abbildung XI-18

Beispiel 5, kriteriumsbezogene FA: Kommunalitäten und Eigenwerte von Analyse A und B

```
FAKTORENANALYSE
***************
KOMMUNALITÄTENSCHÄTZUNG
1   0.11  0.86  0.89  0.91  0.84  0.89  0.89  0.84  0.94  0.89  0.87  0.94  0.91
2   0.09  0.76  0.81  0.84  0.72  0.81  0.81  0.72  0.90  0.80  0.77  0.90  0.84
3   0.06  0.70  0.76  0.80  0.62  0.76  0.75  0.63  0.87  0.74  0.70  0.87  0.79
4   0.04  0.66  0.73  0.77  0.55  0.72  0.71  0.56  0.85  0.70  0.65  0.85  0.75
CON = 0.2000
                OBER ITERATION, KRIT.= 0.4375706E-01   KRIT.MAX= 0.5000000E-01   IT.CYCL.=    4
WERTE FOR DIE VAR.
0.66  0.73  0.77  0.55  0.72  0.71  0.56  0.85  0.70  0.65  0.85  0.75

SUMME DER DER GESCHÄTZTEN KOMMUNALITÄTEN=  8.493

EIEGENWERTE
    3.8038    1.9106    0.9494    0.5727    0.3453    0.2894    0.2371    0.1527    0.0891
    0.0833    0.0414    0.0183

Analyse B: Kommunalitätenschätzungen und Eigenwerte
FAKTORENANALYSE
***************
KOMMUNALITÄTENSCHÄTZUNG
1   0.12  0.87  0.88  0.90  0.85  0.90  0.90  0.85  0.91  0.87  0.86  0.93  0.90
2   0.09  0.77  0.80  0.83  0.74  0.82  0.83  0.74  0.84  0.77  0.76  0.88  0.83
3   0.06  0.71  0.74  0.79  0.65  0.77  0.78  0.66  0.79  0.69  0.68  0.84  0.78
4   0.04  0.67  0.71  0.76  0.60  0.74  0.74  0.59  0.76  0.63  0.62  0.81  0.74
CON=  0.2000
                OBER ITERATION, KRIT.=  0.4316797E-01   KRIT.MAX.= 0.5000000E-01   IT.CYCL.=   4
WERTE FOR DIE VAR.
0.67   0.71   0.76   0.60   0.74   0.74   0.59   0.76   0.63   0.62   0.81   0.74
SUMME DER DER GESCHÄTZTEN KOMMUNALITÄTEN=  8.387
EIGENWERTE    3.6764  1.8279  0.9281  0.6071  0.3934  0.2942  0.2584  0.1601   0.0970  0.0726  0.0521  0.0198
```

Eigenwertediagramme, numerischer Scree-Test und die Indices zur Beurteilung der Restkorrelationen werden ausgelassen

Abbildung XI-19

Beispiel 5, kriteriumsbezogene FA: unrotierte Faktormatrizes von Analyse A und B

Analyse A: unrotierte Faktormatrix

FAKTORMATRIX (2 FAKTOREN)

VARIABLE	KOMM. H**2	FA.-LADUNGEN 1.	2.
2.	0.15	-0.24	-0.30
3.	0.52	0.28	0.66
4.	0.64	0.33	0.73
5.	0.21	0.27	0.38
6.	0.60	0.32	0.71
7.	0.51	0.69	-0.18
8.	0.24	0.46	-0.15
9.	0.73	0.83	-0.21
10.	0.47	0.67	-0.15
11.	0.31	0.55	-0.09
12.	0.74	0.84	-0.20
13.	0.58	0.74	-0.21
% VAR. =	47.6	31.7	15.9

ITERATION CYCLE	VARIANCES
0	0.245189
1	0.440096
2	0.440096
3	0.440096
4	0.440096
5	0.440096

Analyse B: unrotierte Faktormatrix

FAKTORMATRIX (2 FAKTOREN)

VARIABLE	KOMM. H**2	FA.-LADUNGEN 1.	2.
2.	0.14	-0.33	-0.16
3.	0.50	0.04	0.71
4.	0.50	0.31	0.63
5.	0.25	0.19	0.46
6.	0.64	0.28	0.75
7.	0.60	0.76	-0.14
8.	0.32	0.53	-0.20
9.	0.54	0.73	-0.06
10.	0.40	0.63	-0.09
11.	0.34	0.57	-0.12
12.	0.70	0.83	-0.09
13.	0.57	0.74	-0.16
% VAR. =	45.9	30.6	15.2

Abbildung XI-20

Beispiel 5, kriteriumsbezogene FA: rotierte Faktormatrix, Bargmann-Test und geordnete Ladungsmatrix von Analyse A

Analyse A: Varimaxrotierte Faktormatrix und Bargmann-Test

ROTIERTE FAKTORMATRIX (2 FAKTOREN)

VARIABLE	KOMM. H**2	FA.-LADUNGEN 1.	2.
2.	0.15	-0.11	-0.37
3.	0.52	0.02	0.72
4.	0.64	0.05	0.80
5.	0.21	0.12	0.45
6.	0.60	0.04	0.78
7.	0.51	0.71	0.08
8.	0.24	0.49	0.03
9.	0.73	0.85	0.11
10.	0.47	0.68	0.11
11.	0.31	0.54	0.12
12.	0.74	0.85	0.12
13.	0.58	0.76	0.07
% VAR. =	47.6	29.6	18.0

BARGMANN-TEST
****

IN DEN FAKTOREN	ANZ. NULLADUNGEN	P
1	3	0.153
2	2	0.516

Analyse A: Geordnete Ladungsmatrix

DAS MINIMUM DER BEROCKSICHTIGTEN LADUNGEN LIEGT BEI: 0.20

GEORDNETE MATRIX
****

FAKTOR:	1	2
VAR.NR.		
12.	0.85.	0.0
9.	0.85.	0.0
13.	0.76.	0.0
7.	0.71.	0.0
10.	0.68.	0.0
11.	0.54.	0.0
8.	0.49.	0.0
4.	0.0	0.80.
6.	0.0	0.78.
3.	0.0	0.72.
5.	0.0	0.45.
2.	0.0	-0.37.
%VAR.	29.6	18.0

Ende von Analyse A.

Abbildung XI-21

Beispiel 5, kriteriumsbezogene FA: Gefügematrix und geordnete Ladungsmatrix (Analyse B)

Analyse B: Kriteriumsbezogen rotierte schiefwinkelige Faktormatrix (Faktorladungen der Variablen)

KRITERIUM-ROTATION

GEFOGEMATRIX 2 FAKTOREN)

	0.14	-0.34	-0.18
	0.50	0.05	0.71
4.	0.50	0.33	0.65
5.	0.25	0.20	0.48
6.	0.64	0.30	0.77
7.	0.60	0.76	-0.10
8.	0.32	0.53	-0.17
9.	0.54	0.73	-0.01
10.	0.40	0.63	-0.05
11.	0.34	0.57	-0.09
12.	0.70	0.83	-0.04
13.	0.57	0.74	-0.11

Analyse B. Geordnete Ladungsmatrix (Faktorladungen auf den schiefwinkeligen Faktoren)

DAS MINIMUM DER BEROCKSICHTIGTEN LADUNGEN LIEGT BEI:0.20

GEORDNETE MATRIX
****

FAKTOR:	1	2
VAR.NR:		
12.	0.83.	0.0
7.	0.76.	0.0
13.	0.74.	0.0
9.	0.73.	0.0
10.	0.63.	0.0
11.	0.57.	0.0
8.	0.53.	0.0
2.	-0.34.	0.0
6.	0.30	0.77.
3.	0.0	0.71.
4.	0.33	0.65.
5.	0.0	0.48.

Abbildung XI-22

Beispiel 5, kriteriumsbezogene FA: Strukturmatrix und geordnete Faktorkorrelationsmatrix (Analyse B)

ktor ...ukturmatrix (Faktorkorrelatio- ...riablen)

RUKTURMATRIX

2.	0.14	-0.32	-0.15
3.	0.50	-0.01	0.71
4.	0.50	0.27	0.63
5.	0.25	0.16	0.46
6.	0.64	0.23	0.74
7.	0.60	0.77	-0.16
8.	0.32	0.54	-0.21
9.	0.54	0.74	-0.07
10.	0.40	0.63	-0.11
11.	0.34	0.57	-0.14
12.	0.70	0.83	-0.11
13.	0.57	0.75	-0.18

Analyse B: Matrix der geordneten Faktorenkorrelationen der Variablen

DAS MINIMUM DER BEROCKSICHTIGTEN LADUNGEN LIEGT BEI: 0.20

GEORDNETE MATRIX
****

FAKTOR:	1	2
VAR.NR:		
12.	0.83.	0.0
7.	0.77.	0.0
13.	0.75.	0.0
9.	0.74.	0.0
10.	0.63.	0.0
11.	0.57.	0.0
8.	0.54.	-0.21
2.	-0.32.	0.0
6.	0.23	0.74.
3.	0.0	0.71.
4.	0.27	0.63.
5.	0.0	0.46.

Abbildung XI-23

Beispiel 5, kriteriumsbezogene FA: Faktorinterkorrelation und Bargmann-Test (FA 2. Ordnung)

```
                             FAKTOREN-INTER-KORR.
                                        1            2
Interkorrelations-     1.        1.00
matrix zwischen        2.       -0.09         1.00
den schiefwinke-
ligen Faktoren


                       BARGMANN-TEST
                       *************
                       IN DEN FAKTOREN     ANZ. NULLADUNGEN     P
                                  1              12          0.0
                                  2               4          0.029

nach dem Bargmann-Test
hinreichend bestimmte
Faktoren
```

2.6 <u>Beispiel 6</u>: Berechnung von Faktorwerten nach schiefwinkeliger Faktorenrotation (Promax)

Abbildung XI-24

Beispiel 6, FA mit Ausgabe von Faktorwerten: protokollierte Steuersätze und Registerauszug

```
       Eingabedatei: DATEI
          N=50 Fälle (Personen) werden berücksichtigt
             Faktorenanalyse (1.Opt. = 2)
               Kommunalitätenschätzung: Kaiser-Guttman
                 Eigenwerte größer/gleich 1.0
                   Promaxrotation (schiefwinkelig)
                     Faktorscores: OLS-Schätzung
                  1  2  3  4  5 Ausgabedatei: FAKT

EINGABE...E DATEI,N50, 2, 3, 1.0, 1, 1, AFAKT

EINGABE...(721-726,741-747)               ABH. VAR. BZW. KAT-TRANSF. BZW. -TEXT

EINGABE.../*                              ABH. VAR. BZW. KAT-TRANSF. BZW. -TEXT


NEUE    ALTE NR.
*****************
    1.  721.    2  0    7   35  000  ARBEITSZUFRIEDENHEIT-SS'
    2.  722.    2  0    5   25  000  BERUFSENGAGEMENT-SS'
    3.  723.    2  0    8   48  000  ANLAGEORIENTIERUNG (KSE)-SS'
    4.  724.    2  0    8   48  000  KONSERVATISMUS (MK-3)-SS'
    5.  725.    2  0    4   20  000  ROLLENAMBIGUITAET-SS'
    6.  726.    2  0   14   70  000  AMBIGUITAETSTOLERANZ-SS'
    7.  741.    2  0   10   50  000  KONSENS/SCHULBETRIEB-MS'
    8.  742.    2  0   10   50  000  AUSSTATTUNG-MS'
    9.  743.    2  0   10   50  000  SCHUELERZENTR./PAEDAGOGIK-MS'
   10.  744.    2  0   10   50  000  STOERFAKTOR SCHUELER-MS'
   11.  745.    2  0   10   50  000  SCHULORG. BELASTUNG-MS'
   12.  746.    2  0   10   50  000  BASISFUNKT. SCHULE/FACHFOERDG.-MS'
   13.  747.    2  0   10   50  000  SCHULORG. LEHRER-MS'
```

Abbildung XI-25

Beispiel 6, FA mit Ausgabe von Faktorwerten: Kommunalitäten und Eigenwerte

```
FAKTORENANALYSE
***************

KOMMUNALITAETENSCHAETZUNG
 1  0.12  0.91  0.84  0.87  0.90  0.87  0.82  0.92  0.88  0.92  0.89  0.87  0.91  0.85
 2  0.09  0.85  0.72  0.78  0.84  0.78  0.69  0.86  0.79  0.86  0.81  0.78  0.84  0.75
 3  0.06  0.82  0.64  0.72  0.80  0.72  0.59  0.82  0.73  0.82  0.76  0.72  0.80  0.67
 4  0.03  0.80  0.59  0.70  0.79  0.69  0.51  0.80  0.70  0.81  0.74  0.70  0.79  0.63
                     Kommunalitäten der Variablen 1 bis 13
                                Iteration
CON=  0.2000

          UEBER ITERATION, KRIT.= 0.3449909E+00  KRIT.MAX.= 0.5000000E-01  IT.CYCL.= 4

WERTE FUER DIE VAR.
 0.80  0.59  0.70  0.79  0.69  0.51  0.80  0.70  0.81  0.74  0.70  0.79  0.63

SUMME DER DER GESCHAETZTEN KOMMUNALITAETEN=   9.245


EIGENWERTE
     3.4353     1.6134     0.9459     0.7639     0.6838     0.5639     0.4089
     0.3668     0.2727     0.1252     0.0465     0.0186     0.0002
```

Abbildung XI-26

Beispiel 6, FA mit Ausgabe von Faktorwerten: Unrotierte, varimaxrotierte Faktormatrix, Bargmann-Test und geordnete Ladungsmatrix

FAKTORMATRIX (2 FAKTOREN)

VARIABLE	KOMM. H**2	FA.-LADUNGEN 1.	2.
1.	0.43	0.65	0.04
2.	0.05	0.23	0.03
3.	0.33	0.01	0.57
4.	0.43	0.22	0.62
5.	0.23	-0.44	0.17
6.	0.10	-0.07	-0.31
7.	0.50	-0.71	-0.01
8.	0.36	-0.60	-0.08
9.	0.47	-0.66	-0.20
10.	0.36	-0.59	-0.09
11.	0.22	-0.43	0.19
12.	0.42	-0.54	0.36
13.	0.19	-0.30	0.31
% VAR.=	31.4	22.6	8.8

ROTIERTE FAKTORMATRIX (2 FAKTOREN)

VARIABLE	KOMM. H**2	FA.-LADUNGEN 1.	2.
1.	0.43	0.65	0.09
2.	0.05	0.23	0.05
3.	0.33	-0.03	0.57
4.	0.43	0.17	0.63
5.	0.23	-0.46	0.13
6.	0.10	-0.04	-0.31
7.	0.50	-0.71	-0.07
8.	0.36	-0.59	-0.13
9.	0.47	-0.64	-0.25
10.	0.36	-0.59	-0.13
11.	0.22	-0.44	0.15
12.	0.42	-0.56	0.32
13.	0.19	-0.32	0.29
% VAR.=	31.4	22.5	8.9

ITERATION CYCLE	VARIANCES
0	0.280070
1	0.283351
2	0.283351
3	0.283351
4	0.283351
5	0.283351

BARGMANN-TEST
****

IN DEN FAKTOREN	ANZ. NULLADUNGEN	P
1	1	1.000
2	1	1.000

GEORDNETE MATRIX
****

VAR.NR.	FAKTOR: 1	2
7.	-0.71.	0.00
1.	0.65.	0.00
9.	-0.64.	-0.25
8.	-0.59.	0.00
10.	-0.59.	0.00
12.	-0.56.	0.32
5.	-0.46.	0.00
11.	-0.44.	0.00
13.	-0.32.	0.29
2.	0.23.	0.00
4.	0.00	0.63.
3.	0.00	0.57.
6.	0.00	-0.31.
% VAR.	22.5	8.9

Abbildung XI-27

Beispiel 6, FA mit Ausgabe von Faktorwerten: Promaxrotation, Gefüge- und Strukturmatrix sowie Faktorinterkorrelation und Bargmann-Test für die FA 2. Ordnung

PROMAX ROTATION
GEFUEGEMATRIX (2 FAKTOREN)

1.	0.43	0.66	0.10
2.	0.05	0.23	0.05
3.	0.33	0.06	0.58
4.	0.43	0.28	0.64
5.	0.23	-0.44	0.13
6.	0.10	-0.09	-0.32
7.	0.50	-0.72	-0.07
8.	0.36	-0.61	-0.13
9.	0.47	-0.68	-0.26
10.	0.36	-0.61	-0.14
11.	0.22	-0.42	0.15
12.	0.42	-0.51	0.32
13.	0.19	-0.28	0.29

Gefügematrix: Faktorladungen der Variablen

STRUKTURMATRIX

1.	0.43	0.65	-0.01
2.	0.05	0.23	0.01
3.	0.33	-0.04	0.57
4.	0.43	0.17	0.60
5.	0.23	-0.46	0.20
6.	0.10	-0.04	-0.30
7.	0.50	-0.71	0.04
8.	0.36	-0.59	-0.03
9.	0.47	-0.64	-0.15
10.	0.36	-0.58	-0.04
11.	0.22	-0.44	0.22
12.	0.42	-0.57	0.40
13.	0.19	-0.32	0.33

Strukturmatrix: Faktorkorrelationen der Var.

GEORDNETE MATRIX
****

VAR.NR.	FAKTOR: 1	2
7.	-0.72.	0.00
9.	-0.68.	-0.26
1.	0.66.	0.00
8.	-0.61.	0.00
10.	-0.61.	0.00
12.	-0.51.	0.32
5.	-0.44.	0.00
11.	-0.42.	0.00
2.	0.23.	0.00
4.	0.28	0.64.
3.	0.00	0.58.
6.	0.00	-0.32.
13.	-0.28	0.29.

GEORDNETE MATRIX
****

VAR.NR.	FAKTOR: 1	2
7.	-0.71.	0.00
1.	0.65.	0.00
9.	-0.64.	0.00
8.	-0.59.	0.00
10.	-0.58.	0.00
12.	-0.57.	0.40
5.	-0.46.	0.20
11.	-0.44.	0.22
2.	0.23.	0.00
4.	0.00	0.60.
3.	0.00	0.57.
13.	-0.32	0.33.
6.	0.00	-0.30.

FAKTOREN-INTER-KORR.

	1	2
1.	1.00	
2.	-0.16	1.00

BARGMANN-TEST
****

IN DEN FAKTOREN	ANZ. NULLADUNGEN	P
1	1	1.000
2	0	1.000

Abbildung XI-28

Beispiel 6, FA mit Ausgabe von Faktorwerten: Faktorscores

```
MATRIX ZUR BERECHNUNG DER FAKTORENWERTE
FAKT. VARIABLE
    1.     2.     3.     4.     5.     6.     7.     8.     9.    10.    11.    12.    13.

OLS SCHAETZUNG
1. 0.22   0.08 -0.04  0.03 -0.16  0.00 -0.24 -0.20 -0.21 -0.20 -0.16 -0.21 -0.12
2. 0.02   0.02  0.50  0.53  0.16 -0.26  0.01 -0.05 -0.15 -0.06  0.17  0.33  0.28
```

Koeffizienten der Gleichungen zur Berechnung der Scores auf Faktor 1 und Faktor 2 für jede der N = 50 einbezogenen Personen; ordinary least square-Schätzung (Annahme: Homoskedaszidität).

```
   1. GRUPPE   N=      49
```
(s.o. N=50; N=1 Fall mit K.A.)

```
MITTELWERTE DER WERTE
                0.00       -0.04
```
(Mittelwerte der Faktorscores; theoretisch erwartet: M = 0.0)

```
COV.-MATR.
                   1           2
         1.     1.25
         2.    -0.23        1.55
```
(Varianz-Kovarianz-Matrix)

```
MINIMA
               -2.28       -2.79
MAXIMA
                2.20        2.46
```
(weitere Verteilungsparameter der Faktorscores)

```
FAKTOR-WERTE WERDEN UEBERTRAGEN

FOLG. VAR. WURDEN DAZUGEFUEGT
   759.    4  0 -229  221  000  FAKTOR-WERT  1
   760.    4  0 -280  247  000  FAKTOR-WERT  2
```
(Registerauszug der Ausgabedatei)

Text der neugebildeten Variablen (Voreingestellung)

obere Grenze (Score multipliziert mit 100)

untere Grenze (Score multipliziert mit 100)

Anzahl Spalten einschließlich Vorzeichen

Variablennummern der neugebildeten Variablen (Faktorscores)

```
ES WERDEN ZU DEN 758 VARIABLEN DES 1.FILES   2 VARIBLE DES ZWEITEN FILES
DAZUGESPIELT
```
(Eingabe) *(Faktorscores)*

```
DAS ERGEBNIS STEHT AUF DEM AUSGABEFILE
ES WURDEN FUER     50 EINHEITEN  760 VAR. UEBERTRAGEN
```
(Ausgabedatei)

(N = 50 Fälle, N = 1 Fall K.A.; s.o.)

Wie aus dem Protokollauszug ersichtlich, werden die neuen Variablen im Ausgabefile an den bestehenden Variablensatz angehängt und mit "Faktor-Wert 1", "Faktor-Wert 2" usw. betextet. Will man statt dieses voreingestellten Textes einen anderen, frei gewählten eingeben - etwa um die Faktorwerte unterschiedlicher Analysen auch durch ihre Bezeichnung unter-

scheidbar zu machen - ist innerhalb der Klammern mit den abhängigen Variablen der jeweilige Variablentext in Hochkommata (Apostrophe) eingefaßt einzugeben. Dem Faktorwert 1 wird der erste gefundene Text zugewiesen, der Variable mit dem Faktorwert 2 der zweite usw. Dies setzt voraus, daß die Anzahl der zu bildenden Variablen und damit die Anzahl der Faktoren vorher bekannt ist. Es empfielt sich, die FA zunächst ohne die Bildung von Faktorscores zu rechnen und diese erst in einem weiteren Lauf in das Ausgabefile einzutragen, wenn andere als die voreingestellten Variablentexte zugewiesen werden sollen. Selbstverständlich läßt sich der Text von Faktorenvariablen jederzeit mit NIVES ändern (s.Kap.IV).

2.7 Beispiel 7: Faktorenanalyse an zwei Untergruppen mit anschließender Ähnlichkeitstransformation der Faktoren nach KAISER, FISCHER und ROPPERT

Erläuterungen zum Beispiel: (vgl. Abbildung XI-29)

Steuersatz 1: Die ersten N = 150 Fälle der Datei werden nach der unabhängigen Variable 170 aufgeteilt (s. Steuersatz 2). Über die Subgruppen wird je eine Faktorenanalyse (Optionen 1 und 2) gerechnet, wobei die Kommunalitäten iterativ nach dem Verfahren von KAISER und GUTTMAN bestimmt werden (Option 2). Extrahiert werden die Faktoren mit Eigenwerten größer gleich 1.0 (Option 3). Die Faktormatrix wird orthogonal (varimax-) rotiert (Option 4), und es werden keine Faktorwerte berechnet (Option 5).

Steuersatz 2: Die Variable 170 (Geschlecht) wird nach den Ausprägungen 1 "männlich" und 2 "weiblich" unterteilt. Es ergeben sich zwei Teilstichproben mit N=96 und N=48 Fällen (s. Abbildung XI-30).

Steuersatz 3: Die Faktorenanalyse wird über die Variablen 721 bis 726 und 741 bis 747 einschließlich gerechnet (s. Registerauszug).

Abbildung XI-29

Beispiel 7, FA mit Ähnlichkeitstransformation: protokollierte Steuersätze und Registerauszug

Unabhängige Variable Nr. 170, zwei Gruppen (s. Steuersatz 2): Es wird für jede Gruppe eine vollständige FA mit den geforderten Leistungen (Optionen) und anschließender Ähnlichkeitstransformation gerechnet (Ähnlichkeit der faktoriellen Struktur bei männlichen und weiblichen Pbn).

Eingabedatei: DATEI

N=150 Fälle (Personen) werden berücksichtigt

Faktorenanalyse (1.Opt. = 2)

Kommunalitätenschätzung: Kaiser-Guttman

Eigenwerte größer/gleich 1.0

Varimaxrotation

keine Faktorscores

```
                                 1   2   3    4   5
EINGABE...(170) E DATEI, N 150, 2, 3, 1.0, 0, 0

EINGABE...(1/2) 'MAENNL/WEIBL'           KAT-TRANSF.  BZW -TEXT D. UNABH. VAR.

EINGABE...(721-726,741-747)              ABH. VAR. BZW. KAT-TRANSF. BZW. -TEXT

EINGABE.../*                             ABH. VAR. BZW. KAT-TRANSF. BZW. -TEXT

NEUE     ALTE NR.
****************
   1.  170.   1  0    1    2  000  GESCHLECHT
   2.  721.   2  0    7   35  000  ARBEITSZUFRIEDENHEIT-SS'
   3.  722.   2  0    5   25  000  BERUFSENGAGEMENT-SS'
   4.  723.   2  0    8   48  000  ANLAGEORIENTIERUNG (KSE)-SS'
   5.  724.   2  0    8   48  000  KONSERVATISMUS (MK-3)-SS'
   6.  725.   2  0    4   20  000  ROLLENAMBIGUITAET-SS'
   7.  726.   2  0   14   70  000  AMBIGUITAETSTOLERANZ-SS'
   8.  741.   2  0   10   50  000  KONSENS/SCHULBETRIEB-MS'
   9.  742.   2  0   10   50  000  AUSSTATTUNG-MS'
  10.  743.   2  0   10   50  000  SCHUELERZENTR./PAEDAGOGIK-MS'
  11.  744.   2  0   10   50  000  STOERFAKTOR SCHUELER-MS'
  12.  745.   2  0   10   50  000  SCHULORG. BELASTUNG-MS'
  13.  746.   2  0   10   50  000  BASISFUNKT. SCHULE/FACHFOERDG.-MS'
  14.  747.   2  0   10   50  000  SCHULORG. LEHRER-MS'
```

Weggelassen werden die Angaben zu Berechnungsvarianten und Speicherart, die Ausgabe der Mittelwerte, Standardabweichungen und Besetzungen der abhängigen Variablen beider Teilstichproben, die Korrelationsmatrices, die Angaben zur Kommunalitätenberechnung und Faktorenextraktion sowie die unrotierten Faktormatrices.

Abbildung XI-30

Beispiel 7, FA mit Ähnlichkeitstransformation: varimaxrotierte Faktormatrices und Ähnlichkeitskoeffizienten

```
(Grp 1 männlich)                    (Grp 2 weiblich)

GEORDNETE MATRIX                    GEORDNETE MATRIX
****************                    ****************

...................                 ...................
FAKTOR:     1      2                FAKTOR:     1      2
...................                 ...................
VAR.NR.                             VAR.NR.
   8.  -0.76. 0.00                    13.  -0.69. 0.00
  10.  -0.72.-0.28                     2.   0.69. 0.00
  12.  -0.69. 0.00                    10.  -0.68. 0.00
  14.  -0.68. 0.00                     8.  -0.66.-0.24
  13.  -0.64. 0.00                    11.  -0.60.-0.23
  11.  -0.62. 0.00                    14.  -0.56.-0.49
   2.   0.56. 0.27                    12.  -0.49.-0.45
   6.  -0.53. 0.00                     6.  -0.42.-0.36
   3.   0.23. 0.00                     7.   0.23. 0.00
...................                 ...................
   5.   0.00  0.60.                    4.   0.00  0.66.
   4.   0.00  0.57.                    5.   0.00  0.62.
   7.   0.00 -0.47.                    9.  -0.25  0.45.
   9.   0.00 -0.21.                    3.   0.00  0.31.
...................                 ...................
% VAR.  27.4   9.3                  % VAR.  24.0  14.2
...................                 ...................

KORRELATION DER FAKTOREN VERSCH. GRUPPEN
NACH KAISER, FISCHER U. ROPPERT

          1. GRUPPE N=      96
      FAKTOREN

           1.    2.

2. GRUPPE N=   48  (    2 FAKTOREN)    VERGLEICH VON 1. MIT 2.GRUPPE

ÄHNLICHKEITSK. ZWISCHEN FAKTOREN DER   1. UND    2. GRUPPE
        1.    2.
  1. 0.986 0.587
  2. 0.268 0.541
```

(Gruppe 1: N=96 männliche Pbn)
(Gruppe 2: N=48 weibliche Pbn)
(N = 6 Fälle mit K.A.)

(Korrelationen zwischen korrespondierenden Faktoren in den Gruppen; Cosinus der Winkel der Faktoren)

```
M.Q.A.=   0.24362E+00   GLOBALE KORRELATION (GEBHARDT 1967)=  0.842
```

(Gebhardts Q)

(MQA: mittlere quadratische Abweichung der faktoriellen Strukturen)

Abbildung XI-30 enthält die rotierte geordnete Ladungsmatrix beider Teilstichproben (links für die N=96 Fälle mit Ausprägung 1 in Variable 170 - "männlich", Gruppe 1), rechts "weiblich" (Gruppe 2). Ladungen mit einem Betrag kleiner .20 sind Null gesetzt.

Aus Abbildung XI-30 sind außerdem die Koeffizienten der Ähnlichkeit der faktoriellen Struktur in den beiden Gruppen ersichtlich. Hier nicht mehr dargestellt ist die Ähnlichkeitstransformation nach KAISER, FISCHER und ROPPERT. Dabei wird, ausgehend von der Gruppe mit der höchsten Zahl (hier Gruppe 2), die Rotation der Faktoren aus der letzten Analyse (hier 2) gesucht, die eine maximale Deckung pro Faktor der vorgangegangenen Analysen ergibt (vgl. PAWLIK 1971, S.263 f.). Für die Ähnlichkeiten der Faktorstrukturen (Gruppe 1 mit Gruppe 2, wie dargestellt; Gruppe 1 mit der ähnlichkeitstransformierten Struktur, hier nicht dargestellt usw.) berechnet FAM drei verschiedene Maße:

1. Korrelationskoeffizienten zwischen den einzelnen Faktoren in den zu vergleichenden Strukturen; korreliert wird über die Ladungen, die Koeffizienten werden als Matrix ausgegeben (Abbildung XI-30 unten). Im Beispiel korreliert Faktor 1 aus der ersten Gruppe eins r=.986 mit Faktor 1 aus der zweiten.

2. Mittlere quadratische Abweichung (MQA) zwischen den Ladungen beider Faktorstrukturen als ganze betrachtet. Bezeichnet man die Zahl der Faktoren mit k, die der Variablen mit l, und die Ladungen in den beiden Strukturen mit a_{ij} bzw. b_{ij}, so ist

$$MQA = \frac{1}{k \cdot l} \sum_{i=1}^{l} \sum_{j=1}^{k} (a_{ij} - b_{ij})^2$$

3. Analog zum weiter oben angegeben Vorgehen (s. 1.) hat GEBHARDT (1967) ein Ähnlichkeitsmaß vorgeschlagen, mit dem zwei Faktorstrukturen als ganze verglichen werden können:

$$Q(A,B) = \max_{L \in P} \frac{Sp (A'BL)}{\sqrt{Sp (A'A) \; Sp (B'B)}} ,$$

dabei steht Sp für die Spur einer Matrix (Summe der Elemente in der Hauptdiagonalen), P ist die Menge aller orthogonalen reellen kxk-Matrizen. Selbstverständlich stellt Q einen Korrelationskoeffizienten dar; korreliert werden die Ladungen der zu vergleichenden Strukturen, anders als bei (...) nun jedoch über alle Faktoren hinweg.

Für die Einschätzung solcher Gesamt-Koeffizienten gibt GEBHARDT (1967) folgende Richtlinien, die auf Monte-Carlo-Experimenten basieren:

- Für Matrizen aus Zufallszahlen erreicht Q Werte zwischen 0,2 und 0,5.

- wenn verschiedene Stichproben aus der gleichen Grundgesamtheit stammen, ist mit einer Ähnlichkeit der Faktormatrizen von über 0,9 zu rechnen.

- Werte unter 0,8 lassen kaum noch auf ähnliche Faktorstrukturen schließen.

- Einen Hinweis auf die Robustheit der faktorenanalytischen Methode liefert die Tatsache, daß sich für Strukturen, die mit unterschiedlichen Verfahren (andere Korrelationskoeffizienten und Kommunalitätenschätzung) aus gleichen Rohdaten errechnet wurden, sehr hohe Ähnlichkeiten ergaben.

3 Beschränkungen, Voraussetzungen

FAM speichert nur die Kovarianzmatrix (bzw. -matrices bei mehr als einer Gruppe), wobei die Summenwerte bereits beim Einlesen der Daten berechnet werden. Die Anzahl der Fälle (Personen) bzw. die der Ausprägungen der Variablen stellt demnach keine Beschränkung dar; dies wirkt sich lediglich in der benötigten Rechenzeit aus. Die Größe der Kovarianzmatrix wird einzig durch die Anzahl der (abhängigen) Variablen bestimmt. Derzeit können in einem FAM-Lauf bis zu 80 abhängige Variablen analysiert werden. Es existiert eine Großversion von FAM, die umfangreiche Variablensätze verarbeiten kann. Näheres hierzu auf Anfrage.

Wie alle Programmpakete prüfen die KOSTAS- Programme nicht, ob die Voraussetzungen für den Einsatz eines gewählten statistischen Verfahrens gegeben sind. Ob das je realisierte Modell (Hauptkomponentenanalyse, Modell gemeinsamer Faktoren ...) den Daten und der Fragestellung angemessen ist, kann durch das Verfahren selbst nicht geklärt werden. Die Faktorenanalyse geht von Produktmoment-Korrelationen aus, d.h. zunächst müssen die verteilungs- und skalierungsbezogenen Voraussetzungen für die Berechnung dieser Kennwerte gegeben sein. Faktorladungen haftet i.a. ein größerer Stichprobenfehler an als Korrelationskoeffizienten, so daß eine Mindeststichprobengröße von N=150 zu fordern ist. Darüber hinaus sollte für Faktorenanalysen die Personenzahl mindestens das Dreifache der Variablenzahl betragen (vgl. PAWLIK 1971, 275 ff.).

4 Überblick über die Steuermöglichkeiten (Optionentabelle)

Das Faktoranalyseprogramm FAM ermöglicht verschiedene Varianten einer faktorenanlytischen Auswertung von Daten, einschließlich der Berechnung von Faktorwerten. Zur Steuerung dieses Programmteils stehen 6 Optionen plus G- und B-Option(en) zur Verfügung:

1. Option: "2" aus Kompatibilitätsgründen mit älteren KOSTAS-Versionen; nur "2" ist zulässig.

2. Option: Methoden zur Schätzung der Kommunalitäten h^2

"0": Die Diagonalelemente der Korrelationsmatrix werden 1 gesetzt. Es wird somit eine Hauptkomponentenanalyse gerechnet, bei der angenommen wird, daß die gesamte Varianz der Einzelvariablen durch gemeinsame Faktoren erklärt wird. Die Zahl der extrahierten Faktoren entspricht dabei der Zahl der eingegebenen Variablen.

"1": Die Kommunalitäten werden iterativ geschätzt, nachdem die Zahl der Faktoren festgelegt ist (zur Bestimmung der Faktorenanzahl siehe Option 3).

"2": Als Diagonalelemente werden als Kommunalitätenschätzung die Quadrate der multiplen Korrelationskoeffizienten zwischen der Variablen i und den restlichen Variablen eingesetzt. Die R_i^2 stellen eine Schätzung der unteren Grenze der Kommunalitäten dar.

"3": Die Kommunalitätenschätzung erfolgt iterativ nach dem Verfahren von KAISER-GUTTMAN, wobei die quadrierten multiplen Korrelationskoeffizienten als Ausgangsschätzungen verwendet werden und ϵ gleich 1/2 gesetzt ist.

3. Option: Methoden zur Festlegung der Faktorenanzahl

"0" bis "1.99": Es werden so viele Faktoren extrahiert als Eigenwerte größer/gleich dem angegebenen Wert entsprechen (Standardeinstellung "1").

"2" bis "n": Es werden so viele Faktoren wie angegeben extrahiert (aus drucktechnischen Gründen maximal 18). Die Zahl ist sinvollerweise festzulegen, wenn eine erste Faktorenanalyse bereits durchgeführt worden ist und die Anzahl der als wesentlich betrachteten Faktoren (z.B. nach dem Scree-Test) bestimmt wurde.

4. Option: Rotationsverfahren

"0": Es erfolgt eine (orthogonale) Varimaxrotation bzw. bei Festlegung der Variablen, die auf einem Faktor hoch laden sollen, eine kriteriumsbezogene Rotation. Bei kriteriumsbezogener Rotation sind die (abhängigen) Variablen, die einen Faktor bilden sollen, durch "/" voneiander getrennt einzugeben (s. Beispiel 5).

"1": Zusätzlich zur Varimaxrotation wird eine schiefwinkelige Promaxrotation durchgeführt.

"2": Bei Festlegung der Variablen, die auf einem Faktor hoch laden sollen, wird eine schiefwinkelige kriteriumsbezogene Rotation durchgeführt.

5. Option: Berechnung von Faktorwerten

"0": Keine Berechnung von Faktorwerten.

"1": Bei Definition einer Ausgabedatei (vgl. Beispiel 6) wird ein File mit allen Variablen plus den Faktorwerten erstellt. Die Schätzung der Faktorwerte erfolgt nach dem OLS (ordinary least square)-Verfahren.

"2": Wie "1", aber die Faktorwerte werden nach dem GLS (generalized least square)-Verfahren geschätzt.

6. Option: Ordnung der Variablen nach der Höhe der Faktorladungen.

"0.00 - 1.00": Bei Ausgabe der geordneten Matrix werden nur solche Ladungen ausgedruckt, deren Betrag gleich oder größer dem angegebenen Kriterium ist (Voreinstellung: 0.20). Ladungen mit Absolutwerten kleiner oder gleich dem angegebenen Wert werden 0 gesetzt. Dies hat keine Wirkung auf den Rechengang, Option 6 bezieht sich nur auf die Ausgabe.

B-Optionen

Die B-Optionen dienen der Wahl unterschiedlicher Berechnungsvarianten. Im allgemeinen werden diese Varianten von FAM nach programmierten Optimierungen dynamisch ermittelt, d.h. man wird in der Regel mit den voreingestellten Werten operieren. Die in einem Lauf realisierten Berechnungsmodalitäten werden zu Beginn jeden Laufs ausgegeben.

B-Option 1: Korrelationen bzw. Kovarianzen, K.A.-Behandlung

"0" Voreinstellung (4) wird wirksam, falls der verfügbare Speicherplatz ausreicht, sonst 1.

"1" hat ein Fall in der fraglichen Variable keinen zulässigen Wert bei (K.A.), wird der Mittelwert dieser Variable aller einbezogener Fälle eingesetzt.

"2" wie 1, jedoch nur für Fälle (Einheiten) mit mindestens einer beantworteten Variable.

"3" Kovarianzen werden nur für solche Fälle (Einheiten) berechnet, die beide einzubeziehenden Variablen beantwortet haben.

"4" Korrelationen werden nur für solche Fälle (Einheiten) bestimmt, die in beiden zu korrelierenden Variablen zulässige Angaben haben (Voreinstellung, s.o.).

"5" wie "1" (Reserve)

"6" wie "2" (Reserve)

"7" Korrelationen werden nur für solche Fälle (Einheiten) berechnet, die in allen Variablen zulässige Angaben haben.

B-Option 2: Bestimmung des Gruppenumfangs (N).

"0" N ist tatsächliche Gruppen-N (Voreinstellung).

"1" N ist Mittelwert über die Besetzung der einbezogenen Variablen.

"2" N ist Minimum der Besetzungen der einbezogenen Variablen.

"3" N ist Maximum der Besetzungen der einbezogenen Variablen.

G-Option: Gewichtungsvariable (Ausgleich von Stichprobenverzerrungen, Auswertung mehrdimensionaler Daten).

"0" alle Fälle (Einheiten) werden gleich (1.00) gewichtet (Voreinstellung).

"n" Nummer der Variable, die die Gewichte enthält (z.B. G101 Variable 101 enthält die Gewichtungswerte).

Näheres zu B-Optionen und G-Optionen siehe in der Beschreibung des Programmelements LAM in Kapitel X.

Literatur:

BARGMANN, R. 1955; BORTZ, J. 1977; CATTELL, R.B. 1966; CATTELL, R.B. & JASPARS, J. 1967; CATTELL, R.B., VOGELMANN, S. 1977; FISCHER, G., ROPPERT, J. 1965a; FISCHER, G., ROPPERT, J. 1965b; FISCHER, G., ROPPERT, J. 1966; GEBHARDT, F. 1967; GUTTMAN, L. 1957; KAISER H.F. 1956;LUKESCH, H., KLEITER, G.D. 1974; PAWLIK, K. 1971; TIMM, N.H. 1975; ÜBERLA, K. 1968.