2.3 Bestimmung der Faktorenzahl

Ziel der Faktorenanalyse:
Dimensionsreduktion (Informationsverdichtung) durch Extraktion von Faktoren, die einen großen Varianzanteil erklären und inhaltlich interpretierbar sind.

Faktoren mit geringem Erklärungsgehalt stehen einer Dimensionsreduktion im Wege. Da sie keinen substanziellen Beitrag liefern („Zufallsfaktoren“), haben sie in der Anwendung keinerlei Relevanz („Störfaktoren“). Eine sachlich gehaltvolle Faktorenanalyse bedingt daher die Festlegung der **Anzahl der zu extrahierenden Faktoren**.

Kriterien zur Bestimmung der Faktorenzahl

a) Kumuliertes Varianzkriterium

Man beschränkt den Informationsverlust, indem man eine minimal erklärte Varianz von x% vorgibt (**kumuliertes Varianzkriterium**). Verlangt man z. B. mindestens 90% erklärte Varianz, dann sind so viele Faktoren zu extrahieren, bis der durchschnittlich erklärte Varianzanteil die 90% gerade erreicht oder übersteigt.

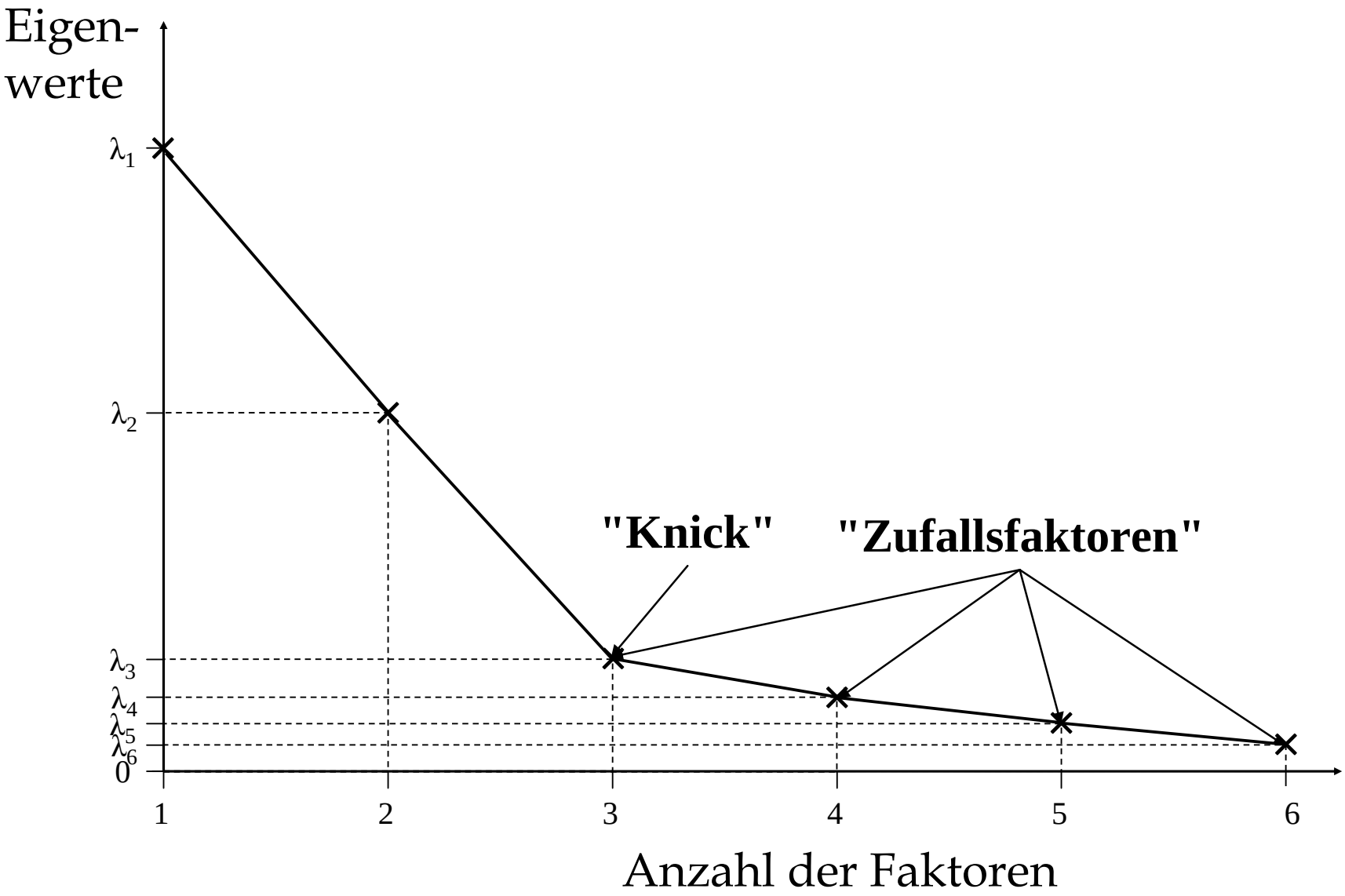
b) Scree-Test

Nach dem sog. **Scree-Test** ist die Anzahl der gemeinsamen Faktoren aus einem Eigenwertdiagramm zu ermitteln, in dem die der Größe nach geordneten Eigenwerte gegen ihre Ordnungsnummer abgetragen werden. Die Faktorenanzahl wird hiernach durch die Anzahl der Eigenwerte bestimmt, die links vom „Knick“ der Verbindungslinien liegen. Wenn die Eigenwerte ab dem „Knick“ approximativ auf einer Parallelen zur Abszisse liegen, lassen sie sich in der Regel als Resultat von „Zufallsfaktoren“ erklären.

c) Kaiser-Kriterium

Das **Kaiser-Kriterium** stellt darauf ab, dass die erklärte Streuung eines Faktors nicht kleiner sein darf als die durch eine manifeste Variable hervorgerufene Streuung. Wenn ein Faktor („Supervariable“) einen geringeren Erklärungsgehalt hätte als eine beobachtbare Variable, würden die Informationen durch ihn nicht "verdichtet". Da die Variablen in standardisierter Form vorliegen, ist ihre Varianz eins. Man könnte daher jeder manifesten Variablen einen Erklärungsbeitrag von 1 zuordnen. Deshalb sind nach dem Kaiser-Kriterium Faktoren zu extrahieren, sofern der mit ihnen korrespondierende Eigenwert größer als 1 ist.

Eigenwertdiagramm für Scree-Test



Bestimmung der Faktorenzahl

Beispiel 2.6: Wie viele Faktoren sind im Beispiel der 12 Regionen zu extrahieren? Würde man einen Erklärungsgehalt von 90 % vorgegeben, dann wären nach dem **kumulierten Varianzkriterium** drei Faktoren zu extrahieren. Ab dem dritten Faktor übersteigt die kumulierte erklärte Varianz die 90 %.

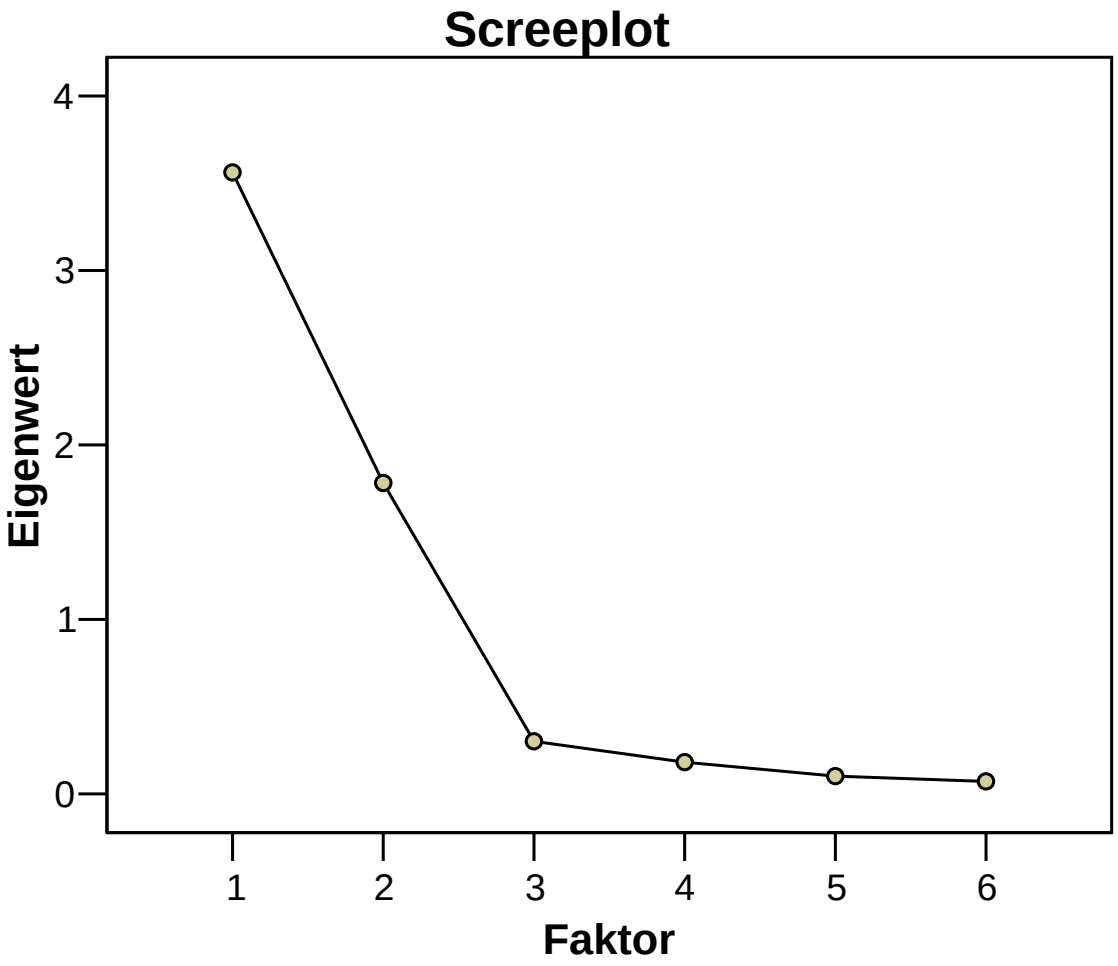
Erklärte Gesamtvarianz

	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion		
Komponente	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	3,562	59,366	59,366	3,562	59,366	59,366
2	1,782	29,704	89,070	1,782	29,704	89,070
3	,301	5,021	94,090	,301	5,021	94,090
4	,182	3,027	97,117	,182	3,027	97,117
5	,102	1,698	98,815	,102	1,698	98,815
6	,071	1,185	100,000	,071	1,185	100,000

Extraktionsmethode: Hauptkomponentenanalyse.

Nach dem **Kaiser-Kriterium** wäre eine zweifaktorielle Lösung zu wählen, da nur zwei Eigenwerte eins überstiegen. Die übrigen Faktoren würden einen geringeren Erklärungsbeitrag leisten als eine manifeste Variable.

Ein Eigenwertdiagramm lässt sich mit SPSS ebenfalls darstellen. Hierzu ist das Item „Screeplot“ im Feld „Anzeigen“ des Fensters „Faktorenanalyse: Extraktion“ mit einem Häkchen zu versehen. Im SPSS-Viewer erscheint dann nach Betätigen der Schaltflächen „Weiter“ und „OK“ für das Regionenbeispiel der Screeplot. Hier zeigt sich ein eindeutiger "Knick" beim dritten Eigenwert. Insofern legt der **Scree-Test** die Extraktion von zwei Faktoren nahe.



Wir wählen hier die zweifaktorielle Lösung, die der Scree-Test und das Kaiser-Kriterium empfehlen. Der dritte Faktor liefert nämlich nur einen Varianzbeitrag von 5,02 % (siehe Tabelle "erklärte Gesamtvarianz").

Die bereits berechneten Ladungen der beiden Faktoren (Hauptkomponenten) werden in der Faktormatrix (Komponentenmatrix) **A** angeordnet:

$$\underset{6\times 2}{\mathbf{A}} = \begin{array}{cc} \text{Faktor 1} & \text{Faktor 2} \\ \end{array}$$

$$\begin{array}{cl} \begin{bmatrix} 0{,}950 & 0{,}132 \\ 0{,}913 & 0{,}239 \\ -0{,}860 & -0{,}395 \\ -0{,}283 & 0{,}908 \\ -0{,}892 & 0{,}043 \\ -0{,}456 & 0{,}851 \end{bmatrix} & \begin{array}{l} \text{ED} \\ \text{BIP} \\ \text{EL} \\ \text{WBIP} \\ \text{GEB} \\ \text{WS} \end{array} \end{array}$$

♦

2.4 Das Kommunalitätenproblem

Als Ergebnis der Anwendung der Hauptkomponentenmethode zur Bestimmung der Matrix der Faktorladungen (Faktormatrix) **A** hat sich die unvollständige Hauptkomponentenlösung

$$\underset{m \times m}{\mathbf{R}} \cong \underset{m \times p}{\mathbf{A}} \cdot \underset{p \times m}{\mathbf{A}'} \quad \text{mit } p < m$$

ergeben. Bei ihr wurde im Lösungsalgorithmus die Existenz von Einzelrestvarianzen und damit der Ausdruck $\frac{1}{n-1} \cdot \mathbf{E}' \cdot \mathbf{E}$ vernachlässigt.

- Was stellt der Ausdruck $\frac{1}{n-1} \cdot \mathbf{E}' \cdot \mathbf{E}$ dar?

Man kann ihn als „Restkorrelationsmatrix“ bezeichnen,

$$\mathbf{R}^{\text{Rest}} = \frac{1}{n-1} \cdot \mathbf{E}' \cdot \mathbf{E}$$

da er die Komponenten der Korrelationsmatrix **R** enthält, die auf das Wirken der Einzelrestfaktoren (merkmalsspezifische Faktoren) zurückzuführen sind.

- Welche Elemente enthält die Restkorrelationsmatrix $\mathbf{R}^{\text{Rest}}$?
  - Hauptdiagonalelemente von $\mathbf{R}^{\text{Rest}}$:

Einzelrestvarianzen (merkmalsspezifische Varianzen) u_j^2 der manifesten Variablen, d.h. der Teil der Varianzen der beobachtbaren Variablen, der nicht durch die gemeinsamen Faktoren F_1, F_2, …, F_p erklärt wird.

- Nicht-Hauptdiagonalelemente von $\mathbf{R}^{Rest}$:
  Da die Einzelrestfaktoren merkmalsspezifisch sind, sind sie in der Grundgesamtheit nicht miteinander korreliert. Die Nicht-Hauptdiagonalelemente von $\mathbf{R}^{Rest}$ sind daher bis auf Zufallsabweichungen gleich 0

Damit hat die **Restkorrelationsmatrix** $\mathbf{R}^{Rest}$ die Gestalt

$$\underset{m \times m}{\mathbf{R}^{\mathbf{Rest}}} = \begin{pmatrix} u_1^2 & \approx 0 & \dots & \approx 0 \\ 0 & u_2^2 & \dots & \approx 0 \\ \vdots & \vdots & \ddots & \vdots \\ \approx 0 & \approx 0 & \dots & u_m^2 \end{pmatrix}$$

Dagegen stellt das Produkt aus der Faktormatrix **A** und ihrer Transponierten **A**', **A·A**', den Teil der Korrelationsmatrix **R** dar, der durch die gemeinsamen Faktoren F_1, F_2, …, F_p reproduziert (erklärt) wird. Man bezeichnet diesen Teil daher als **reproduzierte Korrelationsmatrix**:

$$\tilde{\mathbf{R}}_h = \mathbf{A} \cdot \mathbf{A}'$$

Im Kern unterscheiden sich die Korrelationsmatrix **R** und die reproduzierte Korrelationsmatrix $\tilde{\mathbf{R}}_h$ darin, dass

- in **R** auf der Hauptdiagonale die Varianzen der standardisierten Variablen stehen, die alle gleich 1 sind, während

- auf der Hauptdiagonale von $\tilde{\mathbf{R}}_h$ die Summen der quadrierten Faktorladungen $\sum_{k=1}^{p} a_{jk}^2$, d.h. die durch die gemeinsamen Faktoren erklärten Varianzen der standardisierten Variablen, stehen (≤ 1).

Die auf der Hauptdiagonale von $\tilde{\mathbf{R}}_h$ stehenden Summen quadrierter Faktorladungen heißen **Kommunalitäten**:

$$h_j^2 = \sum_{k=1}^{p} \cdot a_{jk}^2, \quad j = 1,2\ldots,m$$

Die Kommunalitäten h_j^2 geben an, welcher Anteil der Varianz der Variablen X_j durch die extrahierten p Faktoren erklärt wird.

Die gemeinsamen Faktoren $F_1, F_2, \ldots, F_p$, $p<m$, zielen nicht darauf ab, die Varianzen der manifesten Variablen $X_1, X_2, \ldots, X_m$ vollständig zu erklären, da im Allgemeinen **Einzelrestvarianzen** (**merkmalsspezifische Varianzen**) u_j^2 vorhanden sind.

Letztere werden durch die **Einzelrestfaktoren** (**merkmalsspezifische Faktoren**) erklärt. Bei einer verfeinerten Prozedur der Faktorenextraktion müssten daher vorab die merkmalsspezifischen Varianzen u_j^2 aus den (Einheits-)Varianzen der manifesten Variablen „herausgerechnet“ werden, womit man die Kommunalitäten h_j^2 erhält:

$$h_j^2 = 1 - u_j^2$$

Ersetzt man die Einsen („1sen“) auf der Hauptdiagonale der Korrelationsmatrix **R** durch die Kommunalitäten h_j^2, erhält man die Matrix

$$\mathbf{R}_{h\,mxm} = \begin{bmatrix} h_1^2 & r_{12} & r_{13} & \cdots & r_{1m} \\ r21 & h_2^2 & r_{23} & \cdots & r2m \\ r_{31} & r_{32} & h_3^2 & \cdots & r_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & r_{m3} & \cdots & h_m^2 \end{bmatrix}$$

die die **reduzierte Korrelationsmatrix** (Einheitsvarianzen auf der Hauptdiagonale auf h_j^2 „reduziert“). auf darstellt. Die **Faktorenanalyse im engeren Sinne** setzt bei der

Faktorenextraktion an der reduzierten Korrelationsmatrix $\mathbf{R}_h$ an. Wendet man bei der Bestimmung der Faktorladungen die gleiche Verfahrensweise wie bei der Hauptkomponentenmethode an, spricht man von der **Hauptachsen-Faktorenanalyse**.

[Anmerkung: Die reduzierte Korrelationsmatrix $\mathbf{R}_h$ stimmt bis auf die Restkorrelationen (≈ 0) mit der vorher betrachteten reproduzierten Korrelationsmatrix $\mathbf{R}_h = \mathbf{A} \cdot \mathbf{A}'$ überein.]

Während Faktoren (Hauptkomponenten) in der **Hauptkomponentenanalyse** aus der **Korrelationsmatrix R** extrahiert werden, erfolgt die Faktorenextraktion in der **Faktorenanalyse im engeren Sinne** also auf der Grundlage der **reduzierten Korrelationsmatrix $\mathbf{R}_h$**.

Extraktionsverfahren der Faktorenanalyse	
Hauptkomponentenmethode	Faktorenanalyse im engeren Sinne (Hauptachsen-Faktorenanalyse)
- Die gemeinsamen Faktoren (Hauptkomponenten) werden aus der **Korrelationsmatrix R** extrahiert. - Es erfolgt **keine Kommunalitätenschätzung**	- Die gemeinsamen Faktoren werden aus der **reduzierten Korrelationsmatrix $\mathbf{R}_h$** extrahiert. - Hierzu müssen die **Kommunalitäten** h_j^2 **vorab geschätzt** werden.
$\underset{mxm}{\mathbf{R}} = \begin{bmatrix} 1 & r_{12} & r_{13} & \cdots & r_{1m} \\ r21 & 1 & r_{23} & \cdots & r2m \\ r_{31} & r_{32} & 1 & \cdots & r_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & r_{m3} & \cdots & 1 \end{bmatrix}$	$\underset{mxm}{\mathbf{R}_h} = \begin{bmatrix} h_1^2 & r_{12} & r_{13} & \cdots & r_{1m} \\ r21 & h_2^2 & r_{23} & \cdots & r2m \\ r_{31} & r_{32} & h_3^2 & \cdots & r_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & r_{m3} & \cdots & h_m^2 \end{bmatrix}$

Das **Kommunalitätenproblem** besteht darin, dass die Kommunalitäten a priori unbekannt sind und sich erst nach der Faktorextraktion ergeben. Man muss daher die **Kommunalitäten vorab schätzen**, um die Faktorladungen bei einer **Faktorenanalyse im engeren Sinne** (**Hauptachsen-Faktorenanalyse**) bestimmen zu können.

Schätzverfahren für Kommunalitäten

- **Maximaler absoluter Korrelationskoeffizient**

  Relativ einfach ist die Bestimmung der Kommunalität, wenn man auf der Hauptdiagonale der reduzierten Korrelationsmatrix $\mathbf{R}_h$ für jede Variable den höchsten **absoluten Korrelationskoeffizienten** mit einer anderen Variablen einsetzt:

$$h_j^2 = \max\left\{\left|r_{jh}\right|\right\}$$

  Dieses Verfahren hat den Nachteil, dass nur die Beziehung zu *einer* anderen Variablen berücksichtigt wird.

- **Multipler Determinationskoeffizient**

  Eine bessere Möglichkeit zur Schätzung der Kommunalitäten besteht über den **multiplen Determinationskoeffizienten** R_j^2:

$$h_j^2 = R_j^2$$

  Er drückt den Varianzanteil einer Variablen aus, den diese mit *allen* anderen Variablen gemeinsam hat.

Die multiplen Determinationskoeffizienten lassen sich einfach bestimmen, wenn man vorab die Inverse der Korrelationsmatrix, $\mathbf{R}^{-1}$, berechnet. Es gilt:

$$R_j^2 = 1 - \frac{1}{r^{jj}}$$

r^{jj}: j-tes Diagonalelement der inversen Korrelationsmatrix $\mathbf{R}^{-1}$
Man kann nachweisen, dass die Kommunalität einer Variablen stets zwischen R_j^2 und 1 liegt. Man verwendet bei diesem Schätzverfahren damit den niedrigsten Wert, den die "wahre" Kommunalität annehmen kann.

- **Iterationsverfahren**

Noch aufwendiger ist das **Iterationsverfahren,** das beispielsweise mit SPSS durchgeführt werden kann. Hier werden im 1. Schritt **Anfangsschätzer** für die Kommunalität ermittelt wie z.B. die multiplen Determinationskoeffizienten bei einer Faktorenanalyse im engeren Sinne (Methode: „Hauptachsen-Faktorenanalyse“) mit SPSS. Die Anfangsschätzer werden in die Hauptdiagonale der reduzierten Korrelationsmatrix $\mathbf{R}_h$ eingesetzt (Anfangsschätzer ersetzen die Einsen auf der Hauptdiagonale von **R**). Mit diesen Anfangsschätzern werden die Faktorladungen der 1. Iteration bestimmt. Jetzt können die Kommunalitäten aus den Faktorladungen unter

Verwendung der Formel $h_j^2 = \sum_{k=1}^{p} a_{jk}^2, j = 1,2,...,m,$ berechnet werden.

Weichen sie von den Anfangsschätzern merklich ab, wird die Extraktion mit den zuletzt ermittelten Kommunalitäten erneut durchgeführt.
Dieser Prozess kann wiederholt werden, bis sich die Kommunalitäten zweier aufeinander folgender Iterationen kaum noch ändern, d.h. bis das Verfahren konvergiert. Tatsächlich muss das Verfahren aber auch vor einer Iteration beendet werden, in der eine Kommunalität den Wert 1 [Varianz der standardisierten Variablen (=Einheitsvarianz)] überschreiten würde. SPSS stoppt den Iterationsprozess in diesem Fall nicht automatisch, so dass der Anwender hier selbst die maximal zulässige Zahl der Iterationen herausfinden muss.

Kommunalitätenschätzung

Beispiel 2.7: In dem Beispiel der 12 Regionen werden Anfangsschätzer für die Kommunalitäten unter Verwendung des **maximalen Korrelationskoeffizienten,** des **multiplen Determinationskoeffizienten** und des **Iterationsverfahrens** ermittelt.
Um das Verfahren des **maximalen Korrelationskoeffizienten** nachvollziehen zu können, versehen wir das Item „Koeffizienten“ im Feld „Korrelationsmatrix“ des Fensters „Faktorenanalyse: Deskriptive Statistiken“ mit einem Häkchen. Beim Verfahren des **multiplen Determinationskoeffizienten** erhält dagegen das Item „Inverse“ im selben Feld ein Häkchen. Für beide Verfahren setzen wir die „Maximalzahl der Iterationen für Konvergenz“ im Fenster „Faktorenanalyse: Extraktion“ auf 1.

Bei Anwendung des Iterationsverfahrens mit SPSS wählen wir aus dem Pull-down-Menü „Methode“ im Fenster „Faktorenanalyse: Extraktion“ „Hauptachsen-Faktoren-analyse“. Außerdem begrenzen wir hier die „Maximalzahl der Iterationen für Konvergenz“ auf 10, da bei mehr als zehn Iterationen die Kommunalität der Variablen Wanderungssaldo (WS) größer als 1 werden würde.
In Abhängigkeit vom Verfahren zeigen sich zum Teil deutliche Unterschiede.

Geschätzte Kommunalitäten mit unterschiedlichen Verfahren

Variable	Maximaler Korrelationskoeffizient	Multipler Determinations-koeffizient	Iterationsverfahren (Anfangsschätzer R_j^2, 10 Iterationen)
ED	0,907	0,875	0,920
BIP	0,907	0,857	0,858
EL	0,845	0,820	0,859
WBIP	0,832	0,736	0,708
GEB	0,787	0,746	0,701
WS	0,832	0,809	1,000

♦

- **Maximaler absoluter Korrelationskoeffizient**

Die maximalen absoluten Korrelationskoeffizienten können der Korrelationsmatrix **R** entnommen werden. Nach diesem Verfahren wird der absolut größte Korrelationskoeffizient einer Zeile oder Spalte der jeweiligen Variablen zugeordnet. In der unten stehenden Tabelle („Korrelationsmatrix“) sind die zeilenweise absolut maximalen Korrelationskoeffizienten unterstrichen.

Korrelationsmatrix

		Einwohnerdichte	BIP/ Einwohner	Anteil der Erwerbstätigen in der Landwirtschaft (in %)	Wachstum des BIP (in %)	Geburten je 1000 Einwohner	Fort- minus Zuwanderungen je 1000 Einwohner
Korrelation	Einwohnerdichte	1,000	,907_	-,834	-,161	-,787	-,309
	BIP/Einwohner	,907_	1,000	-,845	-,054	-,711	-,220
	Anteil der Erwerbstätigen in der Landwirtschaft (in %)	-,834	-,845_	1,000	-,067	,719	,039
	Wachstum des BIP (in %)	-,161	-,054	-,067	1,000	,226	,832_
	Geburten je 1000 Einwohner	-,787_	-,711	,719	,226	1,000	,454
	Fort- minus Zuwanderungen je 1000 Einwohner	-,309	-,220	,039	,832_	,454	1,000

Die reduzierte Korrelationsmatrix $\mathbf{R}_h$, aus der die Faktorenextraktion vorgenommen wird, hat hier folgende Gestalt:

$$\mathbf{R}_h = \begin{pmatrix} \underline{0{,}907} & 0{,}907 & -0{,}834 & -0{,}161 & -0{,}787 & -0{,}309 \\ 0{,}907 & \underline{0{,}907} & -0{,}845 & -0{,}054 & -0{,}711 & -0{,}220 \\ -0{,}834 & -0{,}845 & \underline{0{,}845} & -0{,}067 & 0{,}719 & 0{,}039 \\ -0{,}161 & -0{,}054 & -0{,}067 & \underline{0{,}832} & 0{,}226 & 0{,}832 \\ -0{,}787 & -0{,}711 & 0{,}719 & 0{,}226 & \underline{0{,}787} & 0{,}454 \\ -0{,}309 & -0{,}220 & 0{,}039 & 0{,}832 & 0{,}454 & \underline{0{,}832} \end{pmatrix}$$

- **Multipler Determinationskoeffizient**

Um die multiplen Determinationskoeffizienten zu bestimmen, lassen wir uns von SPSS die inverse Korrelationsmatrix $\mathbf{R}^{-1}$ ausgeben:

Inverse Korrelationsmatrix

	Einwohnerdichte	BIP/ Einwohner	Anteil der Erwerbstätigen in der Landwirtschaft (in %)	Wachstum des BIP (in %)	Geburten je 1000 Einwohner	Fort- minus Zuwanderungen je 1000 Einwohner
Einwohnerdichte	8.029	-4.817	1.469	.822	1.687	-.083
BIP/Einwohner	-4.817	7.010	2.399	-1.088	-.851	1.252
Anteil der Erwerbstätigen in der Landwirtschaft (in %)	1.469	2.399	5.545	-.652	-1.978	2.207
Wachstum des BIP (in %)	.822	-1.088	-.652	3.792	1.131	-3.628
Geburten je 1000 Einwohner	1.687	-.851	-1.978	1.131	3.940	-2.318
Fort- minus Zuwanderungen je 1000 Einwohner	-.083	1.252	2.207	-3.628	-2.318	5.233

Die Kommunalitätenschätzung für die Variable Einwohnerdichte (ED) erhält man dann unter Verwendung des ersten Hauptdiagonalelements von $\mathbf{R}^{-1}$, also dem Wert 8,029:

$$h_1^2 = 1 - \frac{1}{r^{11}} = 1 - \frac{1}{8{,}029} = 1 - 0{,}125 = 0{,}875\,.$$

Berechnet man auf diese Weise die Kommunalitätenschätzer der übrigen Variablen, erhält man die reduzierte Korrelationsmatrix

$$\mathbf{R}_h = \begin{pmatrix} \underline{0{,}875} & 0{,}907 & -0{,}834 & -0{,}161 & -0{,}787 & -0{,}309 \\ 0{,}907 & \underline{0{,}857} & -0{,}845 & -0{,}054 & -0{,}711 & -0{,}220 \\ -0{,}834 & -0{,}845 & \underline{0{,}820} & -0{,}067 & 0{,}719 & 0{,}039 \\ -0{,}161 & -0{,}054 & -0{,}067 & \underline{0{,}736} & 0{,}226 & 0{,}832 \\ -0{,}787 & -0{,}711 & 0{,}719 & 0{,}226 & \underline{0{,}746} & 0{,}454 \\ -0{,}309 & -0{,}220 & 0{,}039 & 0{,}832 & 0{,}454 & \underline{0{,}809} \end{pmatrix}$$

- **Iterationsverfahren**

Führt man das Iterationsverfahren mit SPSS durch, werden die sich mit dem Verfahren des multiplen Determinationskoeffizienten ergebenden Kommunalitätenschätzer an Stelle der Einsen (1sen) in die Hauptdiagonale der Korrelationsmatrix eingesetzt. Anschließend werden die Kommunalitäten iteriert. Da die Kommunalität des Wanderungssaldos in der 11. Iteration den Wert 1 überschreiten würde, muss das Verfahren nach der 10. Iteration abgebrochen werden. SPSS gibt eine Tabelle mit den Anfangsschätzern („Anfänglich“) und den endgültigen Schätzern („Extraktion“) der Kommunalitäten aus.

Kommunalitäten

	Anfänglich	Extraktion
Einwohnerdichte	.875	.920
BIP/Einwohner	.857	.858
Anteil der Erwerbstätigen in der Landwirtschaft (in %)	.820	.859
Wachstum des BIP (in %)	.736	.708
Geburten je 1000 Einwohner	.746	.701
Fort- minus Zuwanderungen je 1000 Einwohner	.809	1.000

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

Vergleich des Verfahrens des multiplen Determinationskoeffizienten und des Iterationsverfahrens

Faktorenmatrix (mit multiplen Determinationskoeffizien als Anfangsschätzer für Kommunalitäten)[a]

	Faktor	
	1	2
Einwohnerdichte	-.941	.118
BIP/Einwohner	-.901	.223
Anteil der Erwerbstätigen in der Landwirtschaft (in %)	.840	-.368
Wachstum des BIP (in %)	.264	.831
Geburten je 1000 Einwohner	.850	.052
Fort- minus Zuwanderungen je 1000 Einwohner	.437	.818

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

a. Es wurde versucht, 2 Faktoren zu extrahieren. Es werden mehr als 1 Iterationen benötigt. (Konverge 051). Die Extraktion wurde abgebrochen.

Faktorenmatrix (mit Iterationsverfahren)[a]

	Faktor	
	1	2
Einwohnerdichte	-.950	.134
BIP/Einwohner	-.897	.231
Anteil der Erwerbstätigen in der Landwirtschaft (in %)	.844	-.383
Wachstum des BIP (in %)	.268	.798
Geburten je 1000 Einwohner	.836	.039
Fort- minus Zuwanderungen je 1000 Einwohner	.466	.885

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

a. Es wurde versucht, 2 Faktoren zu extrahieren. Es werden mehr als 10 Iterationen benötigt. (Konvergenz= 010). Die Extraktion wurde abgebrochen.

Kommunalitäten (Anfangsschätzer: multipler Determinationskoeffizient)

	Anfänglich	Extraktion
Einwohnerdichte	.875	.900
BIP/Einwohner	.857	.861
Anteil der Erwerbstätigen in der Landwirtschaft (in %)	.820	.841
Wachstum des BIP (in %)	.736	.760
Geburten je 1000 Einwohner	.746	.724
Fort- minus Zuwanderungen je 1000 Einwohner	.809	.860

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

Kommunalitäten (mit Iterationsverfahren)

	Anfänglich	Extraktion
Einwohnerdichte	.875	.920
BIP/Einwohner	.857	.858
Anteil der Erwerbstätigen in der Landwirtschaft (in %)	.820	.859
Wachstum des BIP (in %)	.736	.708
Geburten je 1000 Einwohner	.746	.701
Fort- minus Zuwanderungen je 1000 Einwohner	.809	1.000

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

Berechnung der endgültigen Kommunalitäten aus den Faktorladungen	
$h_1^2 = a_{11}^2 + a_{12}^2 = 0{,}941^2 + 0{,}118^2 = 0{,}899$	$h_1^2 = a_{11}^2 + a_{12}^2 = 0{,}950^2 + 0{,}134^2 = 0{,}920$
$h_2^2 = a_{21}^2 + a_{22}^2 = 0{,}901^2 + 0{,}223^2 = 0{,}861$	$h_2^2 = a_{21}^2 + a_{22}^2 = 0{,}897^2 + 0{,}231^2 = 0{,}858$
$h_3^2 = a_{31}^2 + a_{32}^2 = (-0{,}840)^2 + (-0{,}368)^2 = 0{,}841$	$h_3^2 = a_{31}^2 + a_{32}^2 = (-0{,}844)^2 + (-0{,}383)^2 = 0{,}859$
$h_4^2 = a_{41}^2 + a_{42}^2 = (-0{,}264)^2 + 0{,}831^2 = 0{,}760$	$h_4^2 = a_{41}^2 + a_{42}^2 = 0{,}268^2 + 0{,}798^2 = 0{,}709$
$h_5^2 = a_{51}^2 + a_{52}^2 = (-0{,}850)^2 + 0{,}052^2 = 0{,}725$	$h_5^2 = a_{51}^2 + a_{52}^2 = (-0{,}836)^2 + 0{,}039^2 = 0{,}700$
$h_6^2 = a_{61}^2 + a_{62}^2 = (-0{,}437)^2 + 0{,}818^2 = 0{,}860$	$h_6^2 = a_{61}^2 + a_{62}^2 = (-0{,}466)^2 + 0{,}885)^2 = 1{,}000$

Während sich beim Verfahren des multiplen Determinationskoeffizienten 90,0% der Varianz der Einwohnerdichte (ED) auf die beiden extrahierten gemeinsamen Faktoren zurückführen lässt, steigt dieser Anteil beim Iterationsverfahren auf 92,0%. Nach dem Verfahren des multiplen Determinationskoeffizienten beträgt die Einzelrestvarianz (merkmalsspezifische Varianz) bei der Einwohnerdichte (ED)

$$u_1^2 = 1 - h_1^2 = 1 - 0{,}900 = 0{,}100$$

d.h. 10,0% der gesamten Varianz dieser Variablen, wohingegen sie beim Iterationsverfahren auf 0,080, d.h. 8,0%, zurückgeht.

Berechnung der Eigenwerte der Faktoren F_1 und F_2 aus den Faktorladungen
• Verfahren des multiplen Determinationskoeffizienten
$\lambda_1 = \sum_{j=1}^{6} a_{j1}^2 = 0{,}941^2 + 0{,}901^2 + (-0{,}840)^2 + (-0{,}264)^2 + (-0{,}850)^2 + (-0{,}437)^2 = 3{,}386$
$\lambda_2 = \sum_{j=1}^{6} a_{j2}^2 = 0{,}118^2 + 0{,}223^2 + (-0{,}368)^2 + 0{,}831^2 + 0{,}052^2 + 0{,}818^2 = 1{,}561$
• Iterationsverfahren
$\lambda_1 = \sum_{j=1}^{6} a_{j1}^2 = 0{,}950^2 + 0{,}897^2 + (-0{,}844)^2 + (-0{,}268)^2 + (-0{,}836)^2 + (-0{,}466)^2 = 3{,}405$
$\lambda_2 = \sum_{j=1}^{6} a_{j2}^2 = 0{,}134^2 + 0{,}231^2 + (-0{,}383)^2 + 0{,}798^2 + 0{,}039^2 + 0{,}885^2 = 1{,}640$

Auf der Basis der reduzierten Korrelationsmatrix erklärt der Faktor F_1 (Verstädterung) bei Anwendung des Iterationsverfahrens somit einen Anteil von

$$\frac{\lambda_1}{m} = \frac{3{,}405}{6} = 0{,}568\,,$$

d.h. 56,8% der Variablenvarianz. Durch den Faktor F_2 lässt sich zusätzlich ein Anteil von

$$\frac{\lambda_2}{m} = \frac{1{,}640}{6} = 0{,}273,$$

d.h. 27,3% erklären. Die ersten beiden Faktoren erklären damit zusammen 84,1% der Variablenvarianz.

Die berechneten Eigenwerte und die Erklärungsanteile gehen auch aus den drei letzten Spalten der SPSS-Tabelle („Erklärte Gesamtvarianz“) hervor.

Erklärte Gesamtvarianz (mit Iterationsverfahren)

	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion		
Faktor	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	3.562	59.366	59.366	3.408	56.795	56.795
2	1.782	29.704	89.070	1.638	27.308	84.103
3	.301	5.021	94.090			
4	.182	3.027	97.117			
5	.102	1.698	98.815			
6	.071	1.185	100.000			

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

Ohne das Iterationsverfahren, d.h. allein mit dem Verfahren des multiplen Determinationskoeffizienten würde sich ein geringerer Erklärungsanteil der gemeinsamen Faktoren ergeben:

Faktor F_1: Erklärungsanteil 56,4% $\left(= \frac{3{,}386}{6} \cdot 100\%\right)$

Faktor F_2: Erklärungsanteil 26,0 % $\left(= \frac{1{,}561}{6} \cdot 100\%\right)$

Der Gesamterklärungsanteil der beiden Faktoren würde von 84,1% auf 82,4% sinken. Die Anwendung des aufwendigeren Iterationsverfahrens geht somit mit einer Erhöhung des Erklärungsgehalts des faktorenanalytischen Modells einher.

Vergleich der Lösungen ohne und mit Kommunalitätenschätzung

Unterschiede ergeben sich vor allem hinsichtlich des Anteils der durch die gemeinsamen Faktoren (Hauptkomponenten) erklärten Varianz der manifesten Variablen (→ SPSS-Tabellen Kommunalitäten). Das Faktorenmuster, das durch die Matrix der Faktorladungen (Faktormatrix) wiedergegeben wird, unterscheidet sich bei beiden Lösungen nur unwesentlich (→ SPSS-Tabelle Faktormatrix bzw. Komponentenmatrix).

Lösung ohne Kommunalitätenschätzung	Lösung mit Kommunalitätenschätzung
Hauptkomponentenmethode	**Faktorenanalyse im engeren Sinne (Hauptachsen-Faktorenanalyse)[a]**

Kommunalitäten

	Anfänglich	Extraktion
Einwohnerdichte	1.000	.920
BIP/Einwohner	1.000	.891
Anteil der Erwerbstätigen in der Landwirtschaft (in %)	1.000	.897
Wachstum des BIP (in %)	1.000	.905
Geburten je 1000 Einwohner	1.000	.798
Fort- minus Zuwanderungen je 1000 Einwohner	1.000	.933

Extraktionsmethode: Hauptkomponentenanalyse.

Kommunalitäten

	Anfänglich	Extraktion
Einwohnerdichte	.875	.920
BIP/Einwohner	.857	.858
Anteil der Erwerbstätigen in der Landwirtschaft (in %)	.820	.859
Wachstum des BIP (in %)	.736	.708
Geburten je 1000 Einwohner	.746	.701
Fort- minus Zuwanderungen je 1000 Einwohner	.809	1.000

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

a Im Folgenden wird stets die Anwendung des Iterationsverfahrens bei der Schätzung der Kommunalitäten unterstellt.

Lösung ohne Kommunalitätenschätzung	Lösung mit Kommunalitätenschätzung
Hauptkomponentenmethode	**Faktorenanalyse im engeren Sinne (Hauptachsen-Faktorenanalyse)[a]**

Komponentenmatrix[a]

	Komponente	
	1	2
Einwohnerdichte	-.950	.132
BIP/Einwohner	-.913	.239
Anteil der Erwerbstätigen in der Landwirtschaft (in %)	.860	-.395
Wachstum des BIP (in %)	.283	.908
Geburten je 1000 Einwohner	.892	.043
Fort- minus Zuwanderungen je 1000 Einwohner	.456	.851

Extraktionsmethode: Hauptkomponentenanalyse.

a. 2 Komponenten extrahiert

Faktorenmatrix[a]

	Faktor	
	1	2
Einwohnerdichte	-.950	.134
BIP/Einwohner	-.897	.231
Anteil der Erwerbstätigen in der Landwirtschaft (in %)	.844	-.383
Wachstum des BIP (in %)	.268	.798
Geburten je 1000 Einwohner	.836	.039
Fort- minus Zuwanderungen je 1000 Einwohner	.466	.885

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

a. Es wurde versucht, 2 Faktoren zu extrahieren. Es werden mehr als 10 Iterationen benötigt. (Konvergenz=.010). Die Extraktion wurde abgebrochen.

a Im Folgenden wird stets die Anwendung des Iterationsverfahrens bei der Schätzung der Kommunalitäten unterstellt.