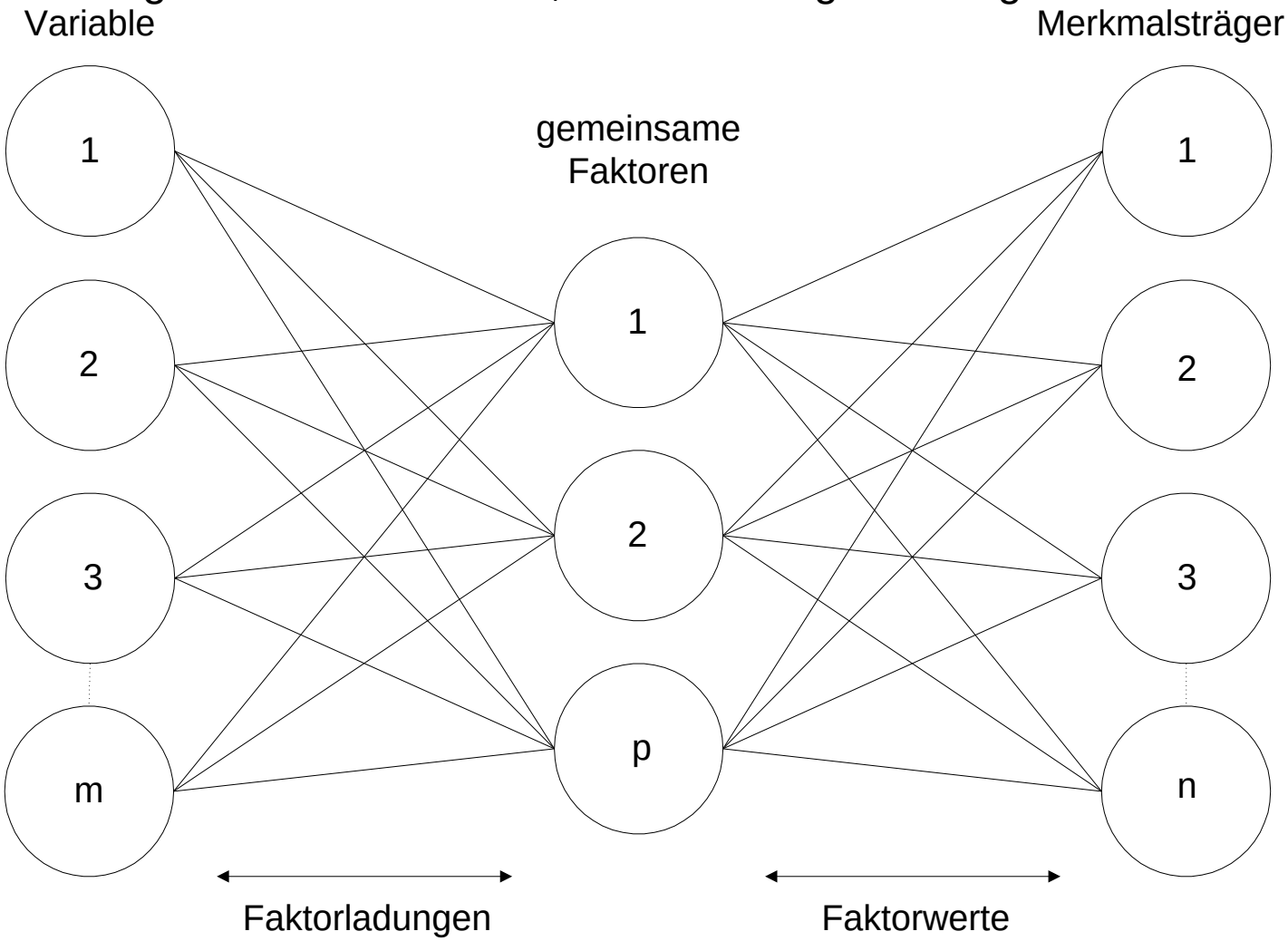
2.2 Faktorenextraktion

Abbildung: Faktorenanalytisches Modell

Zusammenhang zwischen Variablen, Merkmalsträgern und gemeinsamen Faktoren



$X_1, X_2, \ldots, X_m$: manifeste (beobachtbare) Variablen

$F_1, F_2, \ldots, F_p$: gemeinsame Faktoren („Supervariablen“)

● **Faktorenanalytisches Modell**

Strukturelle Gleichung für Merkmalsträger i und Variable X_j:

(2.5) $$z_{ij} = a_{j1} \cdot f_{i1} + a_{j2} \cdot f_{i2} + \cdots + a_{jp} \cdot f_{ip} + e_{ij}$$

Matrixform der strukturellen Gleichung:

(2.6) $$\underset{m \times n}{\mathbf{Z'}} = \underset{m \times p}{\mathbf{A}} \cdot \underset{p \times n}{\mathbf{F'}} + \underset{m \times n}{\mathbf{E'}}$$ ausführlich:

$$\underbrace{\begin{bmatrix} z_{11} & z_{21} & \cdots & z_{n1} \\ z_{12} & z_{22} & \cdots & z_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ z_{1m} & z_{2m} & \cdots & z_{nm} \end{bmatrix}}_{m \times n} =$$

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mp} \end{bmatrix}}_{m \times p} \cdot \underbrace{\begin{bmatrix} f_{11} & f_{21} & \cdots & f_{n1} \\ f_{12} & f_{22} & \cdots & f_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ f_{1p} & f_{2p} & \cdots & f_{np} \end{bmatrix}}_{p \times n} + \underbrace{\begin{bmatrix} e_{11} & e_{21} & \cdots & e_{n1} \\ e_{12} & e_{22} & \cdots & e_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ e_{1m} & e_{2m} & \cdots & e_{nm} \end{bmatrix}}_{m \times n}$$

$F_1, F_2, \ldots, F_p$: gemeinsame Faktoren

$E_1, E_2, \ldots, E_m$: Einzelrestfaktoren

f_{ik}: Faktorwert (Wert des gemeinsamen Faktors F_k bei dem Merkmalsträger i)

- Die Faktorwerte f_{ik} sind wie die Größen z_{ij} standardisiert, d.h. sie haben einen Mittelwert von 0 und eine Varianz von 1.
- Die gemeinsamen Faktoren sind nicht miteinander korreliert, d.h. die Korrelation zwischen F_1 und F_2 ist gleich 0.

e_{ij}: Merkmalsspezifischer Wert (Wert des Einzelrestfaktors) für das Merkmal X_j bei dem Merkmalsträger i

- Die merkmalsspezifischen Werte haben einen Mittelwert von 0.
- Die Einzelrestfaktoren sind nicht mit den gemeinsamen Faktoren korreliert.

a_{jk}: Faktorladung (Ladung der Variablen X_j auf dem gemeinsamen Faktor F_k)

- Die Faktorladung a_{jk} lässt sich als Korrelation zwischen der Variablen X_j und dem gemeinsamen Faktor F_k interpretieren.

Annahmen: $\frac{1}{n-1}\mathbf{F'F} = \mathbf{I}$ und $\frac{1}{n-1}\mathbf{E'F} = \mathbf{0}$

● **Fundamentaltheorem der Faktorenanalyse (=** Zerlegung der Korrelationsmatrix):

(2.7) $$\underset{m\times m}{\mathbf{R}} = \underset{m\times p}{\mathbf{A}} \cdot \underset{p\times m}{\mathbf{A'}} + \frac{1}{n-1} \cdot \underset{m\times n}{\mathbf{E'}} \cdot \underset{n\times m}{\mathbf{E}}$$

mxp-Matrix **A**: Matrix der Faktorladungen (Faktormatrix)

Nach dem Fundamentaltheorem der Faktorenanalyse lässt sich die Korrelations matrix **R** in zwei Teile zerlegen, nämlich in

- einen durch die gemeinsamen Faktoren $F_1, F_2, \ldots, F_p$ erklärten Teil, der aus dem Produkt aus der Faktormatrix **A** und ihrer Transponierten **A**' besteht
- und einem Rest $\frac{1}{n-1} \cdot \mathbf{E'} \cdot \mathbf{E}$, der den Einzelrestfaktoren (merkmalsspezifischen Faktoren) zuzuschreiben ist.

Inhaltlich bedeutet Gleichung (2.7), dass sich die Korrelationen zwischen den manifesten (beobachtbaren) Variablen bis auf einen „Rest“ auf die Korrelationen zwischen den gemeinsamen Faktoren und den manifesten Variablen zurückführen lassen.

Faktormatrix (=Matrix der Faktorladungen):

$$\underset{m \times p}{\mathbf{A}} = \begin{array}{c} \begin{array}{cccc} \text{Faktor 1} & \text{Faktor 2} & \ldots & \text{Faktor p} \end{array} \\ \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \vdots & a_{mp} \end{bmatrix} \end{array} \begin{array}{l} \\ \text{Variable 1} \\ \text{Variable 2} \\ \vdots \\ \text{Variable m} \end{array}$$

$$= [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_p]$$

Faktor – vektor 1	Faktor – vektor 2	…	Faktor – vektor p

Bestimmung der Faktorladungen und Faktorwerte

Formal sieht die strukturelle Gleichung (2.5)/(2.6) wie eine Regressionsgleichung aus, in der man die manifesten Variablen $X_1, X_2, \ldots, X_m$ als abhängige Variablen und die gemeinsamen Faktoren $F_1, F_2, \ldots, F_p$ als unabhängige Variablen betrachtet. Ein Schätzproblem besteht aber darin, dass nicht nur die **Koeffizienten (=Faktor-ladungen)** a_{jk}, sondern auch die **Faktorwerte** f_{ik} **unbekannt** sind. Die Bestimmung der Faktorladungen und Faktorwerte erfolgt aus diesem Grund innerhalb einer **zwei-stufigen Prozedur**.

<u>1. Stufe:</u> **Bestimmung der Faktorladungen**

Die Bestimmung der Faktorladungen setzt nicht unmittelbar an der strukturellen Gleichung (2.5)/(2.6) des faktorenanalytischen Modells, sondern an der **Zerlegung der Korrelationsmatrix** (2.7) an.

Man versucht hierbei, die **Faktormatrix A**, die das „Faktormuster“ wiedergibt, aus der Korrelationsmatrix **R** „herauszuziehen“. Aus diesem Grund spricht man von einer **„Extraktion“ der Faktoren**.

In dieser Stufe sollten die extrahierten Faktoren inhaltlich interpretiert werden. Die inhaltliche Interpretation wird anhand der Faktorladungen vorgenommen, die als Korrelationen zwischen den gemeinsamen Faktoren und den beobachtbaren Variablen verstanden werden können.

2. Stufe: **Bestimmung der Faktorwerte**

Nach Extraktion der Faktoren liegt die Faktormatrix **A** vor. In der strukturellen Gleichung (2.5)/(2.6) sind dann nur noch die Faktorwerte f_{ik}, die die Ausprägungen der gemeinsamen Faktoren bei den Merkmalsträgern wiedergeben, unbekannte Größen. Aus diesem Grund lässt sich die **Matrix der Faktorwerte**, **F**, aus der **strukturellen Gleichung** (2.6) schätzen.

● **Bestimmung der Faktorladungen (Hauptkomponentenmethode)**

Die Hauptkomponentenmethode vernachlässigt im Lösungsalgorithmus den durch die gemeinsamen Faktoren nicht erklärten Teil der Korrelationsmatrix **R**, den wir mit

$\frac{1}{n-1} \cdot \mathbf{E'} \cdot \mathbf{E}$ bezeichnet haben. Sie geht bei der **Faktorenextraktion** von der

Beziehung $\underset{m \times m}{\mathbf{R}} = \underset{m \times p}{\mathbf{A}} \cdot \underset{p \times m}{\mathbf{A'}}$ mit $p = m$ aus. In unserer Lösung werden nur

diejenigen Faktoren [(Haupt-) Komponenten] berücksichtigt, die einen bedeutsamen Beitrag zur Erklärung der Korrelationen r_{jh} leisten.

Zur Bestimmung des ersten Faktors wird zunächst ein mx1-Vektor (Ladungsvektor) $\mathbf{a}_1$ gesucht,

$$\underset{mx1}{\mathbf{a}_1} = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} \begin{array}{l} \text{Ladung der Variablen } X_1 \text{ auf dem Faktor } F_1 \\ \text{Ladung der Variablen } X_2 \text{ auf dem Faktors } F_1 \\ \\ \text{Ladung der Variablen } X_m \text{ auf dem Faktors } F_1 \end{array}$$

dessen Elemente die Ladungen zwischen dem ersten Faktor (F_1) und den einzelnen Variablen wiedergeben. Er soll so gewählt werden, dass er die Korrelationsmatrix **R** mit einer größtmöglichen Genauigkeit approximiert. Aus der Approximation

$$\underset{m \times m}{\mathbf{R}} \cong \underset{m \times 1}{\mathbf{a}_1} \cdot \underset{1 \times m}{\mathbf{a}_1'}$$

folgt nach Multiplikation von rechts mit $\mathbf{a}_1$

$$\mathbf{R} \cdot \mathbf{a}_1 \cong \mathbf{a}_1 \cdot \mathbf{a}_1' \cdot \mathbf{a}_1 = \mathbf{a}_1' \cdot \mathbf{a}_1 \cdot \mathbf{a}_1$$

Unter Verwendung des Skalars

(2.8) $\lambda_1 = \mathbf{a}_1' \cdot \mathbf{a}_1$

ist damit die Gleichung

$$\underset{m\times m}{\mathbf{R}} \cdot \underset{m\times 1}{\mathbf{a}_1} = \underset{1\times 1}{\lambda_1} \cdot \underset{m\times 1}{\mathbf{a}_1}$$

motiviert, aus der wir den ersten Faktor „extrahieren“, d.h. den Faktorvektor $\mathbf{a}_1$ bestimmen.
Multiplizieren wir die rechte Seite dieser Gleichung mit der mxm-Einheitsmatrix **I** (Multiplikation mit der Einheitsmatrix lässt einen Vektor oder eine Matrix ebenso unverändert wie die Multiplikation einer Zahl mit 1),

$$\underset{\text{mxm}}{\mathbf{I}} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

erhält man

$$\mathbf{R} \cdot \mathbf{a}_1 = \lambda_1 \cdot \mathbf{I} \cdot \mathbf{a}_1$$

Wir bringen jetzt noch die rechte Seite auf die linke Seite:

$$\mathbf{R} \cdot \mathbf{a}_1 - \lambda_1 \cdot \mathbf{I} \cdot \mathbf{a}_1 = 0$$

und klammern $\mathbf{a}_1$ aus:

(2.9) $$\left(\mathbf{R} - \lambda_1 \cdot \mathbf{I}\right) \cdot \mathbf{a}_1 = 0$$

Die Berechnung von λ_1 und a_1 stellt mathematisch ein **Eigenwertproblem** dar. λ_1 bezeichnet man als den ersten **Eigenwert** (=größter Eigenwert) und $\mathbf{a}_1$ ist der zugehörige **Eigenvektor**.

- Wie lässt sich der Eigenwert λ_1 in der Faktorenanalyse interpretieren?

Nach Gleichung (2.8) lässt sich λ_1 unter Verwendung der Faktorladungen in der Form

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} = \mathbf{a}_1$$

$$\mathbf{a}_1' = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \end{pmatrix} \begin{pmatrix} \sum_{j=1}^{m} a_{j1}^2 \end{pmatrix} = \lambda_1$$

darstellen.

Der erste Eigenwert λ_1 ergibt sich also als Summe der quadrierten Faktorladungen des ersten Faktors. Aus der einfachen Regressionsanalyse ist bekannt, dass das Quadrat des Korrelationskoeffizienten dem Determinationskoeffizienten (Bestimmtheitsmaß) entspricht: $R^2 = r^2$. Er drückt den Anteil der Varianz der abhängigen Variablen aus, der durch die unabhängige Variable erklärt wird.

Da die Faktorladungen Korrelationskoeffizienten darstellen, sind ihre quadrierten Werte als **Determinationskoeffizienten** zu interpretieren:

a_{11}^2: Teil der Varianz der Variablen X_1, der durch den Faktor F_1 erklärt wird,

a_{21}^2: Teil der Varianz der Variablen X_2, der durch den Faktor F_1 erklärt wird,

$\vdots$

a_{m1}^2 : Teil der Varianz der Variablen X_m, der durch den Faktor F_1 erklärt wird

Der erste Eigenwert λ_1 ergibt sich als Summe von m Determinationskoeffizienten (quadrierten Faktorladungen), die mit dem ersten Faktor F_1 verbunden sind. Er gibt damit die Summe der durch den ersten Faktor F_1 erklärten Varianzanteile wieder. Dividiert man den ersten Eigenwert durch m, dann erhält man ein arithmetisches Mittel – einen Durchschnitt – der erklärten Varianzanteile. Insgesamt gibt es m verschiedene Eigenwerte, die ebenso zu interpretieren sind.

Da manifeste Variablen standardisiert sind, hat jede von ihnen eine Varianz von 1. Bei m manifesten Variablen beträgt die Gesamtvarianz damit m:

Gesamtvarianz der manifesten Variablen: m

Maximal können die Faktoren damit eine Varianz von m erklären. Wenn m Faktoren extrahiert werden, könnten m Eigenwerte bestimmt werden, die die Teile der Varianz wiedergeben, die durch die Faktoren erklärt werden:

λ_1 : Teil der Varianz der manifesten Variablen, die durch den ersten Faktor F_1 erklärt wird,

λ_2 : Teil der Varianz der manifesten Variablen, die durch den zweiten Faktor F_2 erklärt wird,

usw.

Bei Extraktion von m Faktoren (p=m) gilt daher: $\sum_{k=1}^{m} \lambda_k = m$.

Dividiert man die Eigenwerte λ_k durch die Gesamtvarianz m [=Spur der Korrelationsmatrix [tr(**R**) = m)], dann erhält man die durch die einzelnen Faktoren erklärten Anteile der Gesamtvarianz der manifesten Variablen:

$$\frac{\lambda_k}{m} \text{ (zwischen 0 und 1)} \quad \text{oder} \quad \frac{\lambda_k}{m} \cdot 100\% \quad \text{(zwischen 0 und 100\%)}$$

• Wie können die Eigenwerte λ_1, λ_2, … berechnet werden?

Eine **triviale Lösung** des Gleichungssystems (2.9) liegt vor, wenn der Faktorvektor $\mathbf{a}_1$ gleich dem Nullvektor **o** ist: $\mathbf{a}_1 = \mathbf{o}$. Diese Lösung scheidet aber aus, da sie inhaltlich keinerlei Informationen liefert.

Inhaltlich kommt allein eine nichttriviale Lösung in Betracht, die dann gegeben ist, wenn die Determinante der Matrix ($\mathbf{R} - \lambda 1 \cdot \mathbf{I}$) den Wert 0 annimmt:

(2.10) $$\left|\mathbf{R} - \lambda_1 \cdot \mathbf{I}\right| = 0$$

[Anmerkung: Die Matrix ($\mathbf{R} - \lambda_1 \cdot \mathbf{I}$) darf nicht invertierbar sein, d.h. sie muss singulär sein. Dann muss aber ihre Determinante gleich 0 sein.]
Gleichung (2.10) stellt die **charakteristische Gleichung** dar, die ausführlich durch

$$\left|\begin{pmatrix} 1 & r_{12} & \cdots & r_{1m} \\ r_{21} & 1 & \cdots & r_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & 1 \end{pmatrix} - \lambda_1 \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}\right| = \begin{vmatrix} 1-\lambda_1 & r_{12} & \cdots & r_{1m} \\ r_{21} & 1-\lambda_1 & \cdots & r_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \cdots & 1-\lambda_1 \end{vmatrix} = 0$$

gegeben ist. Aus ihr lassen sich alle m Eigenwerte λ_k bestimmen, die den m Wurzeln entsprechen. Tatsächlich bestimmt man jedoch in der Faktorenanalyse nur die größten Eigenwerte, da nur diese Faktoren mit einem hohen Erklärungsgehalt begründen.

• Wie kann der Faktorvektor $\mathbf{a}_1$ berechnet werden?

Nach Berechnung der Eigenwerte wird zunächst der größte Eigenwert λ_1 in das Gleichungssystem (2.9) eingesetzt:

$$(\mathbf{R} - \lambda_1 \cdot \mathbf{I}) \cdot \mathbf{a}_1 = 0 \; .$$

Da jedoch nur m-1 der m Gleichungen

$$\begin{array}{lllllll}
(1-\lambda_1)\cdot a_{11} & +r_{12}\cdot a_{21} & +r_{13}\cdot a_{31} & \cdots & +r_{1m}\cdot a_{m1} & = & 0 \\
r_{21}\cdot a_{11} & +(1-\lambda_1)\cdot a_{21} & +r_{23}\cdot a_{31} & \cdots & +r_{2m}\cdot a_{m1} & = & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\
r_{(m-1)1}\cdot a_{11} & +r_{(m-1)2}\cdot a_{21} & +r_{(m-1)3}\cdot a_{31} & \cdots & +r_{(m-1)m}\cdot a_{m1} & = & 0 \\
r_{m1}\cdot a_{11} & +r_{m2}\cdot a_{21} & +r_{m3}\cdot a_{31} & \cdots & +r_{mm}\cdot a_{m1} & = & 0
\end{array}$$

linear unabhängig sind, benötigt man eine zusätzliche Bedingung, um den Faktor-vektor $\mathbf{a}_1$, d.h. die Ladungen des ersten Faktors F_1 bestimmen zu können. Diese zusätzliche Bedingung ist durch (2.8)

$$\lambda_1 = \mathbf{a}_1' \cdot \mathbf{a}_1$$

gegeben.

Man ersetzt z.B. die letzte Gleichung durch diese Beziehung und erhält das Gleichungssystem,

$$
\begin{array}{lllllll}
(1-\lambda_1)\cdot a_{11} & +r_{12}\cdot a_{21} & +r_{13}\cdot a_{31} & \dots & +r_{1m}\cdot a_{m1} & = & 0 \\
r_{21}\cdot a_{11} & +(1-\lambda_1)\cdot a_{21} & +r_{23}\cdot a_{31} & \dots & +r_{2m}\cdot a_{m1} & = & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\
r_{(m-1)1}\cdot a_{11} & +r_{(m-1)2}\cdot a_{21} & +r_{(m-1)3}\cdot a_{31} & \dots & +r_{(m-1)m}\cdot a_{m1} & = & 0 \\
a_{11}^2 & +a_{21}^2 & +a_{31}^2 & \dots & +a_{m1}^2 & = & \lambda_1
\end{array}
$$

das damit eindeutig lösbar wird.

- Wie können die Faktorvektoren $\mathbf{a}_2$, $\mathbf{a}_3$, … berechnet werden?

<u>Verfahren 1</u>

Der zweite Faktor F_2 ist so zu bestimmen, dass seine Ladungen die **Restkorrelationsmatrix** $\mathbf{R}-\mathbf{a}_1\cdot\mathbf{a}_1'$, d.h. den Anteil der Korrelationsmatrix **R**, der noch nicht durch den ersten Faktor erklärt worden ist, bestmöglich approximiert:

$$
\underbrace{\underbrace{\mathbf{R}}_{m\times m} - \underbrace{\mathbf{a}_1\cdot\mathbf{a}_1'}_{m\times m}}_{m\times m} \cong \mathbf{a}_2\cdot\mathbf{a}_2'
$$

Um den Faktorvektor $\mathbf{a}_2$ des zweiten Faktors F_2 zu bestimmen, ersetzt man im Gleichungssystem (2.9)

- die Korrelationsmatrix **R** durch die Restkorrelationsmatrix $\mathbf{R} - \mathbf{a}_1 \cdot \mathbf{a}_1'$ und
- den größten Eigenvektor λ_1 durch den zweitgrößten Eigenvektor λ_2.

Das sich dadurch ergebende Gleichungssystem

(2.11) $$\left(\mathbf{R} - \mathbf{a}_1 \cdot \mathbf{a}_1' - \lambda_2 \cdot \mathbf{I}\right) \cdot \mathbf{a}_2 = 0$$

wird hierbei noch um zwei Bedingungen ergänzt. Zum einen wird erneut von der Bedingung Gebrauch gemacht, dass die Summe der quadrierten Faktorladungen dem Eigenwert entsprechen muss:

(2.12) $$\lambda_2 = \mathbf{a}_2' \cdot \mathbf{a}_2$$

Zum anderen spiegelt sich die Unabhängigkeit der Faktoren in der Orthogonalitätsbedingung

(2.13) $$\mathbf{a}_1' \cdot \mathbf{a}_2 = 0$$

wider. Danach steht der durch den Faktorvektor $\mathbf{a}_2$ repräsentierte Faktor F_2 senkrecht auf dem durch den Faktorvektor $\mathbf{a}_1$ repräsentierten Faktor F_1.

Auf diese Art und Weise lassen sich bei Bedarf weitere Faktoren aus der jeweils verbleibenden Restkorrelationsmatrix extrahieren, die sukzessive die größten Varianzanteile der manifesten Variablen erklären.

Verfahren 2

Die Bestimmung des Faktorvektors $\mathbf{a}_2$, $\mathbf{a}_3$, … aus der jeweiligen Restkorrelationsmatrix, d.h. der jeweils nach der Extraktion des 1. Faktors, 2. Faktors, etc. verbleibenden Korrelationsmatrix, vollzieht die Vorstellung nach, dass jeder zusätzlich extrahierte Faktor den größtmöglichen Teil der Restkorrelationen zwischen den Variablen erklären soll. Formal gleichwertig können alle Faktorvektoren aber auch aus der originären Korrelationsmatrix berechnet werden. So kann der Faktorvektor $\mathbf{a}_2$ ebenfalls auf der Grundlage des Gleichungssystems

(2.14) $$\left(\mathbf{R} - \lambda_2 \cdot \mathbf{I}\right) \cdot \mathbf{a}_2 = 0$$

mit der Bedingung (2.12) bestimmt werden. Erneut ist dann die Orthogonalitätsbedingung (2.13) erfüllt.

Auch bestehende weitere Faktorvektoren $\mathbf{a}_3$, $\mathbf{a}_4$, … lassen sich ohne explizite Berechnung der Restkorrelationsmatrizen aus dem Gleichungssystem (2.14) ermitteln, wenn man den Eigenwert λ_2 durch die Eigenwerte λ_3, λ_4, … ersetzt. Aufgrund der einfacheren Rechentechnik wird diesem Verfahren häufig der Vorzug vor dem intuitiv einleuchtenden ersten Verfahren gegeben.

Berechnung der Eigenwerte im 12-Regionen-Beispiel

Bei mehr als drei manifesten Variablen ist die Berechnung der Eigenwerte ohne ein Computerprogramm bereits zu aufwendig. In dem 12-Regionen-Beispiel mit 6 Variablen wird zur Illustration der Eigenwertberechung daher auf den Rechner zurückgegriffen.

Zuerst sind die Nullstellen der Determinante $\left|\mathbf{R} - \lambda_1 \cdot \mathbf{I}\right|$ zu bestimmen:

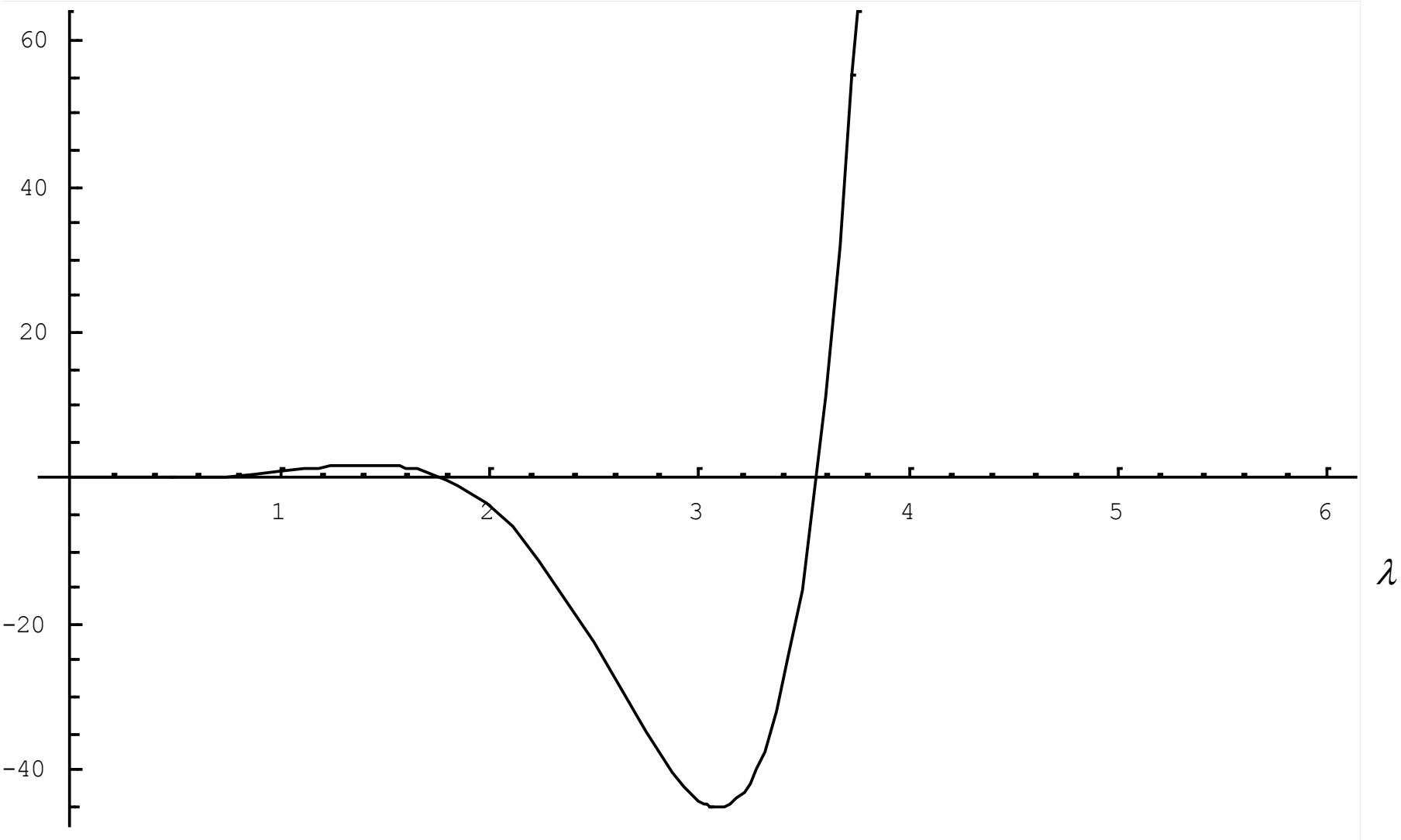
$$\begin{vmatrix} 1-\lambda & 0{,}907 & -0{,}834 & -0{,}161 & -0{,}787 & -0{,}309 \\ 0{,}907 & 1-\lambda & -0{,}845 & -0{,}054 & -0{,}711 & -0{,}220 \\ -0{,}834 & -0{,}845 & 1-\lambda & -0{,}067 & 0{,}719 & 0{,}039 \\ -0{,}161 & -0{,}054 & -0{,}067 & 1-\lambda & 0{,}226 & 0{,}832 \\ -0{,}787 & -0{,}711 & 0{,}719 & 0{,}226 & 1-\lambda & 0{,}454 \\ -0{,}309 & -0{,}220 & 0{,}039 & 0{,}832 & 0{,}454 & 1-\lambda \end{vmatrix} = 0$$

Man erhält ein Polynom sechsten Grades,

$$\lambda^6 - 6{,}0000 \cdot \lambda^5 + 9{,}9985 \cdot \lambda^4 - 4{,}9536 \cdot \lambda^3 + 0{,}9931 \cdot \lambda^2 - 0{,}0844 \cdot \lambda + 0{,}0025 = 0$$

das sechs Nullstellen aufweist:

$$\left|\mathbf{R} - \lambda \cdot \mathbf{I}\right|$$



Das Polynom hat sechs Wurzeln, wobei die größte der gesuchte erste Eigenwert λ_1 ist. Die zweitgrößte Wurzel ist der zweite Eigenwert λ_2 usw.

Die sechs Eigenwerte gehen aus folgender Tabelle hervor:

Eigenwerte	Erklärte Varianz	Kumulierte erklärte Varianz
$\lambda_1 = 3{,}562$	59,4 %	59,4 %
$\lambda_2 = 1{,}782$	29,7 %	89,1 %
$\lambda_3 = 0{,}301$	5,0 %	94,1 %
$\lambda_4 = 0{,}182$	3,0 %	97,1 %
$\lambda_5 = 0{,}102$	1,7 %	98,5 %
$\lambda_6 = 0{,}071$	1,2 %	100 %

Berechung der Ladungen des Faktors F_1 im 12-Regionen-Beispiel

Der zum ersten Eigenwert λ_1 gehörende Eigenvektor wird berechnet, indem das Gleichungssystem

$$
\begin{array}{lllllll}
-2{,}562^{*}\cdot a_{11} & +0{,}907\cdot a_{21} & -0{,}834\cdot a_{31} & -0{,}161\cdot a_{41} & -0{,}787\cdot a_{51} & -0{,}309\cdot a_{61} & =0 \\
0{,}907\cdot a_{11} & -2{,}562\cdot a_{21} & -0{,}845\cdot a_{31} & -0{,}054\cdot a_{41} & -0{,}711\cdot a_{51} & -0{,}220\cdot a_{61} & =0 \\
-0{,}834\cdot a_{11} & -0{,}845\cdot a_{21} & -2{,}562\cdot a_{31} & -0{,}067\cdot a_{41} & +0{,}719\cdot a_{51} & +0{,}039\cdot a_{61} & =0 \\
-0{,}161\cdot a_{11} & -0{,}054\cdot a_{21} & -0{,}067\cdot a_{31} & -2{,}562\cdot a_{41} & +0{,}226\cdot a_{51} & +0{,}832\cdot a_{61} & =0 \\
-0{,}787\cdot a_{11} & -0{,}711\cdot a_{21} & +0{,}719\cdot a_{31} & +0{,}226\cdot a_{41} & -2{,}562\cdot a_{51} & +0{,}454\cdot a_{61} & =0 \\
a_{11}^2 & +a_{21}^2 & +a_{31}^2 & +a_{41}^2 & +a_{51}^2 & +a_{61}^2 & =3{,}562
\end{array}
$$

*1 – λ1 =1 – 3,562 = – 2,562

nach den unbekannten Koeffizienten a_{11}, a_{21}, …, a_{61} gelöst wird. Hierzu werden im ersten Schritt die ersten fünf Gleichungen verwendet und vorläufige Lösungswerte (mit einem Stern * gekennzeichnet) ermittelt, indem a_{11} gleich 1 setzt.

Das sich hierbei ergebende Gleichungssystem

$$
\begin{array}{llllll}
-2{,}562\cdot 1 & +0{,}907\cdot a^*_{21} & -0{,}834\cdot a^*_{31} & -0{,}161\cdot a^*_{41} & -0{,}787\cdot a^*_{51} & -0{,}309\cdot a^*_{61} & =0 \\
0{,}907\cdot 1 & -2{,}562\cdot a^*_{21} & -0{,}845\cdot a^*_{31} & -0{,}054\cdot a^*_{41} & -0{,}711\cdot a^*_{51} & -0{,}220\cdot a^*_{61} & =0 \\
-0{,}834\cdot 1 & -0{,}845\cdot a^*_{21} & -2{,}562\cdot a^*_{31} & -0{,}067\cdot a^*_{41} & +0{,}719\cdot a^*_{51} & +0{,}039\cdot a^*_{61} & =0 \\
-0{,}161\cdot 1 & -0{,}054\cdot a^*_{21} & -0{,}067\cdot a^*_{31} & -2{,}562\cdot a^*_{41} & +0{,}226\cdot a^*_{51} & +0{,}832\cdot a^*_{61} & =0 \\
-0{,}787\cdot 1 & -0{,}711\cdot a^*_{21} & +0{,}719\cdot a^*_{31} & +0{,}226\cdot a^*_{41} & -2{,}562\cdot a^*_{51} & +0{,}454\cdot a^*_{61} & =0
\end{array}
$$

führt zu den vorläufigen Lösungswerten

$$a^*_{11}=1,\ a^*_{21}=0{,}9609,\ a^*_{31}=-0{,}9054,\ a^*_{41}=-0{,}2979,\ a^*_{51}=-0{,}9390,\ a^*_{61}=-0{,}4801$$

Die endgültigen Faktorladungen erhält man unter Berücksichtigung der sechsten Gleichung. Die quadrierten Faktorladungen müssen aufsummiert den ersten Eigenwert λ_1 ergeben:

$$\lambda_1 = a_1'\cdot a_1 = a^2_{11}+a^2_{21}+a^2_{31}+a^2_{41}+a^2_{51}+a^2_{61}=3{,}562$$

Die vorläufigen quadrierten Lösungswerte ergeben aber aufaddiert:

$$\left(a_{11}^{*}\right)^2+\left(a_{21}^{*}\right)^2+\left(a_{31}^{*}\right)^2+\left(a_{41}^{*}\right)^2+\left(a_{51}^{*}\right)^2+\left(a_{61}^{*}\right)^2$$

$$=1^2+0{,}9609^2+(-0{,}9054)^2+(-0{,}2979)^2+(-0{,}9390)^2+(-0{,}4801)^2=3{,}9440\,.$$

Man bestimmt zunächst den auf die Länge 1 normierten Eigenvektor $\mathbf{a}_1^n$, indem man die Komponenten des vorläufigen Lösungsvektors $\mathbf{a}_1^*$ durch $\sqrt{3{,}9440}$ dividiert:

$$a_{11}^n=\frac{1}{\sqrt{3{,}9440}}=0{,}5035\,,\quad a_{21}^n=\frac{0{,}9609}{\sqrt{3{,}9440}}=0{,}4838\,,\quad a_{31}^n=\frac{-0{,}9054}{\sqrt{3{,}9440}}=-0{,}4559\,,$$

$$a_{41}^n=\frac{-0{,}2979}{\sqrt{3{,}9440}}=-0{,}1500,\quad a_{51}^n=\frac{-0{,}9390}{\sqrt{3{,}9440}}=-0{,}4728\,,\quad a_{61}^n=\frac{-0{,}4801}{\sqrt{3{,}9440}}=-0{,}2417\,.$$

Die Summe der quadrierten normierten Lösungswerte ergibt eins:

$$\left(a_{11}^n\right)^2+\left(a_{21}^n\right)^2+\left(a_{31}^n\right)^2+\left(a_{41}^n\right)^2+\left(a_{51}^n\right)^2+\left(a_{61}^n\right)^2$$

$$0{,}5035^2+0{,}4838^2+(-0{,}4559)^2+(-0{,}1500)^2+(-0{,}4728)^2+(-0{,}2417)^2=1.$$

Wir erhalten nun den gesuchten Faktorvektor $\mathbf{a}_1$, in dem wir den auf die Länge 1 normierten Eigenvektor $\mathbf{a}_1^n$ mit der Wurzel des ersten Eigenwerts multiplizieren:

$$\mathbf{a}_1 = \mathbf{a}_1^n \cdot \sqrt{\lambda_1} = \mathbf{a}_1^n \cdot \sqrt{3{,}562}$$

Komponentenweise Berechnung:

$$a_{11} = a_{11}^n \cdot \sqrt{\lambda_1} = 0{,}5035 \cdot \sqrt{3{,}562} = 0{,}950\,,\ a_{21} = 0{,}4838 \cdot \sqrt{3{,}562} = 0{,}913\ ,$$

$$a_{31} = (-0{,}4559) \cdot \sqrt{3{,}562} = -0{,}860\ ,\quad a_{41} = (-0{,}1500) \cdot \sqrt{3{,}562} = -0{,}283\,,$$

$$a_{51} = (-0{,}4728) \cdot \sqrt{3{,}562} = -0{,}892\ ,\quad a_{61} = (-0{,}2417) \cdot \sqrt{3{,}562} = -0{,}456\ .$$

Wie eine Probe zeigt, erfüllen die berechneten Faktorladungen die sechste Gleichung:

$$\begin{aligned} & a_{11}^2 + a_{21}^2 + a_{31}^2 + a_{41}^2 + a_{51}^2 + a_{61}^2 \\ & = 0{,}950^2 + 0{,}913^2 + (-0{,}860)^2 + (-0{,}283)^2 + (-0{,}892)^2 + (-0{,}456)^2 \\ & = 3{,}562 = \lambda_1 \end{aligned}$$

(Anmerkung: Man beachte, dass die sechste Gleichung auch bei einer Vertauschung der Vorzeichen erfüllt sein würde.)

Wie könnten wir den ersten Faktor F_1, der den Ladungsvektor

$$\mathbf{a}_1' = \begin{pmatrix} 0{,}950 & 0{,}913 & -0{,}860 & -0{,}283 & -0{,}892 & -0{,}456 \end{pmatrix}$$

$$\begin{matrix} r_{ED;F_1} & r_{BIP;F_1} & r_{EL;F_1} & r_{WBIP;F_1} & r_{GEB;F_1} & r_{WS;F_1} \end{matrix}$$

besitzt, interpretieren? Als Faustregel kann man zur Interpretation der Faktoren Faktorladungen, die absolut größer als 0,5 sind, heranziehen. Der erste Faktor F_1 weist eine hohe positive Korrelation mit der Einwohnerdichte und dem BIP je Einwohner auf. Die erklärten Varianzanteile liegen bei:

$0{,}950^2 \cdot 100\% = 90{,}2\%$ (Einwohnerdichte)
$0{,}913^2 \cdot 100\% = 83{,}4\%$ (BIP pro Kopf).

Der erste Faktor F_1 hängt gleichzeitig stark negativ mit dem Anteil der Erwerbstätigen in der Landwirtschaft und der Geburtenrate zusammen.

Der Faktor F_1 nimmt also in Regionen mit

hoher Einwohnerdichte,
hohem BIP,
niedriger Beschäftigung in der Landwirtschaft und
geringer Geburtenrate

tendenziell große Werte an. Hierbei handelt es sich um verstädterte Regionen. Der erste Faktor ist deshalb als "Verstädterung" zu interpretieren.

Berechung der Ladungen des Faktors F_2 im 12-Regionen-Beispiel

<u>Verfahren 1</u>

Zur Bestimmung der Ladungen des zweiten Faktors F_2 gehen wir von dem Gleichungssystem (2.11) aus. Wir bilden zunächst die Restkorrelationsmatrix $\mathbf{R} - \mathbf{a}_1 \cdot \mathbf{a}_1'$:

$$\begin{array}{cc} & \begin{pmatrix} 0{,}950 & 0{,}913 & -0{,}860 & -0{,}283 & -0{,}892 & -0{,}456 \end{pmatrix} \mathbf{a}_1' \\ \mathbf{a}_1 \begin{pmatrix} 0{,}950 \\ 0{,}913 \\ -0{,}860 \\ -0{,}283 \\ -0{,}892 \\ -0{,}456 \end{pmatrix} & \begin{pmatrix} 0{,}903 & 0{,}867 & -0{,}817 & -0{,}269 & -0{,}847 & -0{,}433 \\ 0{,}867 & 0{,}834 & -0{,}785 & -0{,}258 & -0{,}814 & -0{,}416 \\ -0{,}817 & -0{,}785 & 0{,}740 & 0{,}243 & 0{,}767 & 0{,}392 \\ -0{,}269 & -0{,}258 & 0{,}243 & 0{,}080 & 0{,}252 & 0{,}129 \\ -0{,}847 & -0{,}814 & 0{,}767 & 0{,}252 & 0{,}796 & 0{,}407 \\ -0{,}433 & -0{,}416 & 0{,}392 & 0{,}129 & 0{,}407 & 0{,}208 \end{pmatrix} \mathbf{a}_1 \cdot \mathbf{a}_1' \end{array},$$

$$\mathbf{R} - \mathbf{a}_1 \cdot \mathbf{a}_1' =$$

$$\begin{pmatrix} 1-0{,}903 & 0{,}907-0{,}867 & -0{,}834-(-0{,}817) & -0{,}161-(-0{,}269) & -0{,}787-(-0{,}847) & -0{,}309-(-0{,}433) \\ 0{,}907-0{,}867 & 1-0{,}834 & -0{,}845-(-0{,}785) & -0{,}054-(-0{,}258) & -0{,}711-(-0{,}814) & -0{,}220-(-0{,}416) \\ -0{,}834-(-0{,}817) & -0{,}845-(-0{,}785) & 1-0{,}740 & -0{,}067-0{,}243 & 0{,}719-0{,}767 & 0{,}039-0{,}392 \\ -0{,}161-(-0{,}269) & -0{,}054-(-0{,}258) & -0{,}067-0{,}243 & 1-0{,}080 & 0{,}226-0{,}252 & 0{,}832-0{,}129 \\ -0{,}787-(-0{,}847) & -0{,}711-(-0{,}814) & 0{,}719-0{,}767 & 0{,}226-0{,}252 & 1-0{,}796 & 0{,}454-0{,}407 \\ -0{,}309-(-0{,}433) & -0{,}220-(-0{,}416) & 0{,}039-0{,}392 & 0{,}832-0{,}129 & 0{,}454-0{,}407 & 1-0{,}208 \end{pmatrix}$$

$$= \begin{pmatrix} 0{,}097 & 0{,}040 & -0{,}017 & 0{,}108 & 0{,}060 & 0{,}124 \\ 0{,}040 & 0{,}166 & -0{,}060 & 0{,}204 & 0{,}103 & 0{,}196 \\ -0{,}017 & -0{,}060 & 0{,}260 & -0{,}310 & -0{,}048 & -0{,}353 \\ 0{,}108 & 0{,}204 & -0{,}310 & 0{,}920 & -0{,}026 & 0{,}703 \\ 0{,}060 & 0{,}103 & -0{,}048 & -0{,}026 & 0{,}204 & 0{,}047 \\ 0{,}124 & 0{,}196 & -0{,}353 & 0{,}703 & 0{,}047 & 0{,}792 \end{pmatrix}$$

Hiermit erhält das Gleichungssystem:

$$\begin{array}{lllllll}
-1{,}685^{*} \cdot a_{12} & +0{,}040 \cdot a_{22} & -0{,}017 \cdot a_{32} & +0{,}108 \cdot a_{42} & +0{,}060 \cdot a_{52} & +0{,}142 \cdot a_{62} & = 0 \\
0{,}040 \cdot a_{12} & -1{,}685 \cdot a_{22} & -0{,}060 \cdot a_{32} & +0{,}204 \cdot a_{42} & +0{,}103 \cdot a_{52} & +0{,}196 \cdot a_{62} & = 0 \\
-0{,}017 \cdot a_{12} & -0{,}060 \cdot a_{22} & -1{,}685 \cdot a_{32} & -0{,}310 \cdot a_{42} & -0{,}048 \cdot a_{52} & -0{,}353 \cdot a_{62} & = 0 \\
0{,}108 \cdot a_{12} & +0{,}204 \cdot a_{22} & -0{,}310 \cdot a_{32} & -1{,}685 \cdot a_{42} & -0{,}026 \cdot a_{52} & +0{,}703 \cdot a_{62} & = 0 \\
0{,}060 \cdot a_{12} & +0{,}103 \cdot a_{22} & -0{,}048 \cdot a_{32} & -0{,}026 \cdot a_{42} & -1{,}685 \cdot a_{52} & +0{,}047 \cdot a_{62} & = 0 \\
a_{12}^2 & +a_{22}^2 & +a_{32}^2 & +a_{42}^2 & +a_{52}^2 & +a_{62}^2 & = 1{,}782
\end{array}$$

* $0{,}097 - \lambda 2 = 0{,}097 - 1{,}782 = -1{,}685$

Unter Verwendung der ersten fünf Gleichungen erhält man die vorläufigen Faktor-ladungen

$$a^*_{12} = 1\,,\ a^*_{22} = 1{,}8134\,,\ a^*_{31} = -2{,}9955,\ a^*_{42} = 6{,}8830\,,\ a^*_{52} = 0{,}3259\,,\ a^*_{62} = 6{,}4531\ .$$

Sie müssen wegen

$$\left(a^*_{12}\right)^2 + \left(a^*_{22}\right)^2 + \left(a^*_{31}\right)^2 + \left(a^*_{41}\right)^2 + \left(a^*_{51}\right)^2 + \left(a^*_{62}\right)^2$$

$$= 1^2 + 1{,}8134^2 + \left(-2{,}9955\right)^2 + 6{,}8830^2 + 0{,}3259^2 + 6{,}4531^2 = 102{,}3858$$

durch den Faktor $\sqrt{102{,}3858}$ dividiert werden, um den auf 1 normierten Faktorvektor $\mathbf{a}^n_2$ zu erhalten:

$$a^n_{12} = \frac{1}{\sqrt{102{,}3858}} = 0{,}0988,\quad a^n_{22} = \frac{1{,}8134}{\sqrt{102{,}3858}} = 0{,}1792,\quad a^n_{32} = \frac{-2{,}9955}{\sqrt{102{,}3858}} = -0{,}2960,$$

$$a^n_{42} = \frac{6{,}8830}{\sqrt{102{,}3858}} = 0{,}6802,\quad a^n_{52} = \frac{0{,}3259}{\sqrt{102{,}3858}} = 0{,}0322,\quad a^n_{62} = \frac{6{,}4531}{\sqrt{102{,}3858}} = 0{,}6377.$$

Der Faktorvektor $\mathbf{a}_2$ ergibt sich denn nach Multiplikation von $\mathbf{a}_2^n$ mit der Wurzel des zweiten Eigenwertes:

$$\mathbf{a}_2 = \mathbf{a}_2^n \cdot \sqrt{\lambda_2} = \mathbf{a}_2^n \cdot \sqrt{1{,}782}$$

Komponentenweise Berechnung:

$$\mathbf{a}_{12} = a_{12}^n \cdot \sqrt{\lambda_2} = 0{,}0988 \cdot \sqrt{1{,}7822} = 0{,}132\,, \quad \mathbf{a}_{22} = 0{,}1792 \cdot \sqrt{1{,}7822} = 0{,}239,$$

$$\mathbf{a}_{32} = (-0{,}2960) \cdot \sqrt{1{,}7822} = -0{,}395\,, \quad \mathbf{a}_{42} = 0{,}6802 \cdot \sqrt{1{,}7822} = 0{,}908,$$

$$\mathbf{a}_{52} = 0{,}0322 \cdot \sqrt{1{,}7822} = 0{,}043\,, \quad \mathbf{a}_{62} = 0{,}6377 \cdot \sqrt{1{,}7822} = 0{,}851.$$

Verfahren 2

Alternativ berechnen wir jetzt den Faktorvektor $\mathbf{a}_2$ aus dem Gleichungssystem (2.14), das ohne die Berechnung der Restkorrelationsmatrix auskommt:

$$
\begin{array}{llllllll}
-0{,}782^{*}\cdot a_{12} & +0{,}907\cdot a_{22} & -0{,}834\cdot a_{32} & -0{,}161\cdot a_{42} & -0{,}787\cdot a_{52} & -0{,}309\cdot a_{62} & =0 \\
0{,}907\cdot a_{12} & -0{,}782\cdot a_{22} & -0{,}845\cdot a_{32} & -0{,}054\cdot a_{42} & -0{,}711\cdot a_{52} & -0{,}220\cdot a_{62} & =0 \\
-0{,}834\cdot a_{12} & -0{,}845\cdot a_{22} & -0{,}782\cdot a_{32} & -0{,}067\cdot a_{42} & +0{,}719\cdot a_{52} & +0{,}039\cdot a_{62} & =0 \\
-0{,}161\cdot a_{12} & -0{,}054\cdot a_{22} & -0{,}067\cdot a_{32} & -0{,}782\cdot a_{42} & +0{,}226\cdot a_{52} & +0{,}832\cdot a_{62} & =0 \\
-0{,}787\cdot a_{12} & -0{,}711\cdot a_{22} & +0{,}719\cdot a_{32} & +0{,}226\cdot a_{42} & -0{,}782\cdot a_{52} & +0{,}454\cdot a_{62} & =0 \\
a_{12}^2 & +a_{22}^2 & +a_{32}^2 & +a_{42}^2 & +a_{52}^2 & +a_{62}^2 & =1{,}782
\end{array}
$$

* $1 - \lambda 2 = 1 - 1{,}782 = -0{,}782$

Unter Verwendung der ersten fünf Gleichungen erhält man folgende vorläufige Lösungswerte:

$a_{12}^{*} = 1$, $a_{22}^{*} = 1{,}8134$, $a_{31}^{*} = -2{,}9955$, $a_{42}^{*} = 6{,}8830$, $a_{52}^{*} = 0{,}3259$, $a_{62}^{*} = 6{,}4531$.

Diese werden unter Verwendung von:

$$\left(a_{12}^{*}\right)^2+\left(a_{22}^{*}\right)^2+\left(a_{31}^{*}\right)^2+\left(a_{41}^{*}\right)^2+\left(a_{51}^{*}\right)^2+\left(a_{62}^{*}\right)^2=$$

$$1^2+1{,}8134^2+\left(-2{,}9955\right)^2+6{,}8830^2+0{,}3259^2+6{,}4531^2=102{,}3858$$

normiert,

$$a_{12}^{n}=\frac{1}{\sqrt{102{,}3858}}=0{,}0988,\quad a_{22}^{n}=\frac{1{,}8134}{\sqrt{102{,}3858}}=0{,}1792,\quad a_{32}^{n}=\frac{-2{,}9955}{\sqrt{102{,}3858}}=-0{,}2960,$$

$$a_{42}^{n}=\frac{6{,}8830}{\sqrt{102{,}3858}}=0{,}6802,\quad a_{52}^{n}=\frac{0{,}3259}{\sqrt{102{,}3858}}=0{,}0322,\quad a_{62}^{n}=\frac{6{,}4531}{\sqrt{102{,}3858}}=0{,}6377.$$

und mit der Wurzel des zweiten Eigenwertes multipliziert:

$$\mathbf{a}_{12}=a_{12}^{n}\cdot\sqrt{\lambda_2}=0{,}0988\cdot\sqrt{1{,}7822}=0{,}132,\quad \mathbf{a}_{22}=0{,}1792\cdot\sqrt{1{,}7822}=0{,}239,$$

$$\mathbf{a}_{32}=\left(-0{,}2960\right)\cdot\sqrt{1{,}7822}=-0{,}395,\quad \mathbf{a}_{42}=0{,}6802\cdot\sqrt{1{,}7822}=0{,}908,$$

$$\mathbf{a}_{52}=0{,}0322\cdot\sqrt{1{,}7822}=0{,}043,\quad \mathbf{a}_{62}=0{,}6377\cdot\sqrt{1{,}7822}=0{,}851.$$

Zur Interpretation des zweiten Faktors F_2 betrachten wir die Komponenten des Faktorvektors

$$\mathbf{a}_2' = \begin{pmatrix} 0{,}132 & 0{,}239 & -0{,}395 & 0{,}908 & 0{,}043 & 0{,}851 \end{pmatrix}$$

$$r_{ED;F_2} \quad r_{BIP;F_2} \quad r_{EL;F_2} \quad r_{WBIP;F_2} \quad r_{GEB;F_2} \quad r_{WS;F_2}$$

Offenbar hängt F_2 stark positiv mit

der Wachstumsrate des BIP (WBIP) und

dem Wanderungssaldo (WS)

zusammen. Regionen mit hohen Werten des zweiten Faktors weisen tendenziell einen starken Anstieg des BIP und hohe Zuzüge von Einwohnern auf. Anscheinend sind diese Regionen sehr attraktiv, und der zweite Faktor könnte damit als "Attraktivität" interpretiert werden.

Faktorenextraktion mit SPSS (Hauptkomponentenmethode)

Beispiel 2.5: Die Faktorladungen lassen sich im Beispiel der 12 Regionen ebenfalls mit SPSS berechnen. Hierzu laden wir die Datei regionen(12).sav in den SPSS-Daten-Editor und wählen die Menüpunkte

Analysieren
    Dimensionsreduktion
        Faktorenanalyse …

Wir markieren die sechs quantitativen Variablen und übertragen Sie durch Anklicken der oberen Pfeil-Schaltfläche in das Feld „Variablen“. Ansonsten nutzen wir zunächst nur die Voreinstellungen. Im Fenster „Faktorenanalyse: Extraktion“ befindet sich ein Pull-down-Menü „Methode:“. Dort ist die Extraktionsmethode „Hauptkomponenten“ voreingestellt.

Nach Betätigen der Schaltflächen „Weiter“ und „OK“ gibt SPSS im SPSS-Viewer verschiedene Outputtabellen aus. Wir beschränken uns hier auf die beiden Tabellen „Erklärte Gesamtvarianz“ und „Komponentenmatrix“.

Die SPSS-Tabelle „Erklärte Gesamtvarianz“ gibt die Eigenwerte und die kumulierten Varianzanteile wieder. Die letzten vier Faktoren erklären zusammen nur noch 100% - 94,1% = 5,9% der gesamten Varianz der Variablen. Die Summe der quadrierten Faktorladungen eines Faktors stimmt stets mit seinem Eigenwert überein.

Erklärte Gesamtvarianz

Komponente	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion		
	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.

Die Matrix der Faktorladungen finden wir bei einer Faktorextraktion mit der Hauptkomponentenmethode unter der Bezeichnung „Komponentenmatrix“ wieder. SPSS bezeichnet die mit dieser Methode extrahierten Faktoren als Komponenten.

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

Wie bereits angesprochen, ist die Faktorlösung bis auf das Vorzeichen eindeutig. SPSS hat die Vorzeichen des ersten Faktors anders als wir gesetzt. Der erste Faktor, den SPSS bestimmt hat, wäre dann als "Gegenteil der Verstädterung" zu interpretieren.

Auf die gleiche Weise lassen sich die Ladungen der übrigen Faktoren bestimmen. Man könnte natürlich im Prinzip alle 6 Faktoren (Komponenten) extrahieren. Doch erkennt man hier, dass die Faktorladungen der Faktoren F_3 bis F_6 absolut den Wert 0,5 unterschreiten. Sie liegen also alle zwischen − 0,5 und 0,5 und weisen damit nur einen geringen Zusammenhang zu den Variablen auf. Dies deckt sich mit dem vorher festgestellten Befund, dass die Komponenten F_3 bis F_6 nur einen geringen Beitrag zur Erklärung der Varianz der manifesten Variablen haben.

Komponentenmatrix[a]

	Komponente					
	1	2	3	4	5	6
Einwohnerdichte	-,950	,132	,121	,093	,189	-,144
BIP/Einwohner	-,913	,239	,214	,184	-,064	,159
Anteil der Erwerbstätigen in der Landwirtschaft (in %)	,860	-,395	-,068	,264	,156	,068
Wachstum des BIP (in %)	,283	,908	-,203	,202	-,088	-,074
Geburten je 1000 Einwohner	,892	,043	,434	,051	-,069	-,080
Fort- minus Zuwanderungen je 1000 Einwohner	,456	,851	,084	-,161	,159	,092

Extraktionsmethode: Hauptkomponentenanalyse.

a. 6 Komponenten extrahiert

♦