

Grundzüge der Faktorenanalyse

Bei der Faktorenanalyse handelt es sich um ein Verfahren, mehrere Variablen durch möglichst wenige gemeinsame, hinter ihnen stehende Faktoren zu beschreiben.

Beispiel: Betrachtet werden die Schulnoten nach Fächern für n Schüler. Es stellt sich die Frage, ob hinter den Leistungen in den Fächern gemeinsame Faktoren stehen. Im Ergebnis könnten diese Faktoren z.B. die „naturwissenschaftliche Begabung“ und die „sprachliche Begabung“ ausdrücken. Meist werden die Faktoren voneinander unabhängig angenommen.

Hinter der Faktorenanalyse verbirgt sich kein bestimmtes Rechenverfahren, sondern es handelt sich vielmehr um eine Sammelbezeichnung für eine Vielzahl von Verfahren, die Lösungen für das o.g. Problem liefern. Es gibt jedoch ein gemeinsames Modell (Thurstone, 1931). Einzelheiten hierzu lassen sich z.B. bei Clauß/Ebner¹ nachlesen.

Im folgenden soll daher ein Weg vorgestellt werden, eine Faktorenanalyse durchzuführen. Es handelt sich um die sog. Hauptkomponentenmethode. Dieses Verfahren findet z.B. im Programm SPSS Anwendung.

Die Durchführung der Faktorenanalyse führt auf vier Aufgabenstellungen:

- 1) Bestimmung der Anzahl von Faktoren: Ermittlung der Anzahl der Faktoren, die benötigt werden, um die Variablen zu erklären.
- 2) Bestimmung der Faktorenladungen: Die Faktorenladungen geben an, wie stark die Faktoren auf eine Variable wirken.
- 3) Ermittlung der Kommunalitäten: Die Kommunalität gibt an, welcher Anteil der Varianz einer Variablen durch die „gemeinsamen“ Faktoren aufgeklärt wird.

¹ Clauß/Ebner, STATISTIK für Soziologen, Pädagogen, Psychologen und Mediziner, Band 1, 6. Auflage, Thun und Frankfurt am Main 1989

- 4) Rotation der Faktorenladungen: Zur besseren Interpretation der Faktoren ist es im Allgemeinen erforderlich, eine „Rotation“ der Faktorenladungsmatrix durchzuführen. Das Ziel dieser Rotation besteht darin, eine „Einfachstruktur“ der Faktorenladungsmatrix zu erzielen.

Durchführung der Faktorenanalyse nach der Hauptkomponentenmethode

Als Ausgangsdaten werden metrisch skalierte Variablen verwendet. Im allgemeinen handelt es sich um Daten aus empirischen Untersuchungen, die zeilenweise nach n Personen (Fällen) und spaltenweise nach m Variablen (Merkmalen) dargestellt werden.

Beispiel: Noten in vier Fächern für zehn Schüler. Betrachtet werden sollen die Korrelationen zwischen den Fächern in einer Schülerpopulation. Damit die Zufallsfehler nicht zu groß werden, sollte die Zahl der Schüler möglichst groß sein. Zu Anschauungszwecken wird jedoch ein Datensatz mit nur zehn Fällen verwendet.

X	Deutsch	Philosophie	Mathematik	Physik
Schüler 1	2	3	1	2
Schüler 2	4	4	3	2
Schüler 3	1	1	2	2
Schüler 4	3	2	4	5
Schüler 5	2	2	1	2
Schüler 6	4	3	2	2
Schüler 7	4	3	2	3
Schüler 8	1	2	4	4
Schüler 9	2	2	3	3
Schüler 10	3	2	1	3

Die Ausgangsdatenmatrix wird folgend als X bezeichnet.

Es soll zunächst dargestellt werden, welche Korrelation zwischen den Variablen besteht. Die entsprechende Korrelationsmatrix R ist im Beispiel:

R	Deutsch	Philosophie	Mathematik	Physik
Deutsch	1,00	0,74	-0,07	-0,07
Philosophie	0,74	1,00	-0,02	-0,28
Mathematik	-0,07	-0,02	1,00	0,71
Physik	-0,07	-0,28	0,71	1,00

Anhand der Korrelationsmatrix lässt sich im konkreten Fall schon erkennen, dass ein Zusammenhang zwischen den Noten der Fächer Deutsch und Philosophie einerseits sowie zwischen den Noten der Fächer Mathematik und Physik andererseits besteht.

Es gibt eine Reihe von Verfahren, um vor Durchführung einer Faktorenanalyse zu prüfen, ob das Ausgangsdatenmaterial überhaupt geeignet ist. Hierauf soll an dieser Stelle nicht eingegangen werden.

Für die weitere Betrachtung werden die Variablen standardisiert. Die Matrix, die standardisierten Werte enthält, wird als \underline{Z} bezeichnet.

In unserem Beispiel erhalten wir als standardisierte Werte

Z	Deutsch	Philosophie	Mathematik	Physik
Schüler 1	-0,51	0,71	-1,12	-0,77
Schüler 2	1,19	1,90	0,60	-0,77
Schüler 3	-1,36	-1,66	-0,26	-0,77
Schüler 4	0,34	-0,47	1,47	2,13
Schüler 5	-0,51	-0,47	-1,12	-0,77
Schüler 6	1,19	0,71	-0,26	-0,77
Schüler 7	1,19	0,71	-0,26	0,19
Schüler 8	-1,36	-0,47	1,47	1,16
Schüler 9	-0,51	-0,47	0,60	0,19
Schüler 10	0,34	-0,47	-1,12	0,19

Unter Verwendung von \underline{Z} lässt sich nun wiederum die Korrelationsmatrix bestimmen:

$$\underline{R} = \frac{1}{n-1} \underline{Z}'\underline{Z}$$

(Natürlich ist dies dieselbe Korrelationsmatrix, die Sie auch für die Spalten von \underline{X} erhalten).

Ziel ist es, die Variablen z_1, \dots, z_m als Linearkombination aus $K < m$ Faktoren p_1, \dots, p_K und deren Ladungen a_{ij} darzustellen. Die Faktoren werden alle als standardisiert vorausgesetzt, dies ist keine Einschränkung der Allgemeinheit. Der i -te Variablenwert der j -ten Variablen ist

$$z_{ij} = p_{i1}a_{j1} + \dots + p_{iK}a_{jK}$$

Dies ist die Grundgleichung der Faktorenanalyse. In Matrixschreibweise ausgedrückt ergibt dies

$$\underline{Z} = \underline{PA}'$$

Es geht nun darum, die Matrix \underline{A} , die die Faktorenladungen enthält, zu bestimmen. Weiterhin sollen auch die Kommunalitäten, die angeben, welcher Varianzanteil einer Variablen durch die Faktoren aufgeklärt wird, bestimmt werden², dazu später mehr.

Die Korrelationsmatrix lässt sich nun ausdrücken als

$$\begin{aligned} \underline{R} &= \frac{1}{n-1} \underline{Z}'\underline{Z} \\ &= \frac{1}{n-1} (\underline{PA}')'(\underline{PA}') \end{aligned}$$

² Hinweis: Zu Beginn der Faktorenanalyse werden üblicherweise die „anfänglichen Kommunalitäten“ geschätzt, die dann anstelle der Einsen in die Hautdiagonale der Korrelationsmatrix eingetragen werden, bevor mit der eigentlichen rechnerischen Durchführung der Faktorenanalyse begonnen wird. Für die Schätzung der Kommunalitäten gibt es bestimmte Verfahren, die aber hier nicht weiter dargestellt werden.

In der Art, wie die Faktorenanalyse hier vorgestellt wird und wie sie von Programmen wie SPSS durchgeführt wird, taucht das Problem der Schätzung der Kommunalitäten praktisch nicht auf: es wird die „unbehandelte“ Korrelationsmatrix zugrunde gelegt – die anfänglichen Kommunalitäten sind also alle Einsen.

$$= \frac{1}{n-1} \underline{A} \underline{P}' \underline{P} \underline{A}'.$$

Da man davon ausgeht, dass die Faktoren nicht korreliert sind, ist die Korrelationsmatrix der Faktoren $\frac{1}{n-1} \underline{P}' \underline{P} = \underline{I}$, also

$$\underline{R} = \underline{A} \underline{A}'$$

Somit lässt sich \underline{A} durch Eigenwertzerlegung von \underline{R} bestimmen:

$$\underline{R} = \underline{E} \underline{\Lambda} \underline{E}'$$

$$= \underline{E} \underline{\Lambda}^{1/2} (\underline{E} \underline{\Lambda}^{1/2})', \text{ also}$$

$$\underline{A} = \underline{E} \underline{\Lambda}^{1/2} \underline{I},$$

wobei \underline{I} eine orthogonale Matrix bezeichnet. Dies ist der Grund für das Rotationsproblem – hierzu später mehr.

Für unser Beispiel erhält man nach einer Eigenwertzerlegung der Matrix \underline{A} die Eigenwerte λ_j

1,95 1,51 0,39 0,14

und die Matrix \underline{E} mit den Eigenvektoren von \underline{R}

-0,50	-0,50	0,52	0,48
-0,55	-0,44	-0,42	-0,57
0,43	-0,57	-0,55	0,44
0,52	-0,47	0,50	-0,50

Nun gilt es, die Zahl der erforderlichen Faktoren zu ermitteln. Als übliches Maß wird hier das „Kaiser-Kriterium“ verwendet: Die Anzahl der beizubehaltenden Faktoren entspricht der Anzahl der Eigenwerte > 1 . Dieses Kriterium ergibt sich daraus, dass

nur diese Faktoren einen Varianzanteil > 1 einer standardisierten Ausgangsvariablen aufklären.

Im vorliegenden Beispiel gibt es zwei Eigenwerte > 1 , daher ist die Zahl der beizubehaltenden Faktoren $K = 2$.

Damit lässt sich die Faktorenladungsmatrix A durch Eigenwertzerlegung von \underline{R} bestimmen:

$$\underline{A} = \underline{E} \underline{\Lambda}_K^{1/2}$$

In unserem Beispiel ergibt sich für \underline{A}

A	Faktor 1	Faktor 2
Deutsch	-0,69	-0,62
Philosophie	-0,77	-0,54
Mathematik	0,60	-0,70
Physik	0,72	-0,58

Unter Verwendung von \underline{A} lässt sich die reproduzierte Korrelationsmatrix $\tilde{\underline{R}}$ bestimmen:

$$\tilde{\underline{R}} = \underline{A} \underline{A}'$$

Für unser Beispiel³ ist $\tilde{\underline{R}}$

R~	Deutsch	Philosophie	Mathematik	Physik
Deutsch	<u>0,86</u>	0,87	0,02	-0,14
Philosophie	0,87	<u>0,88</u>	-0,08	-0,24
Mathematik	0,02	-0,08	<u>0,86</u>	0,84
Physik	-0,14	-0,24	0,84	<u>0,86</u>

³ Rundungsdifferenzen möglich

In der Diagonalen der reduzierten Korrelationsmatrix steht der Anteil der Varianz der Variablen, der durch die Faktoren aufgeklärt wird – dies sind die Kommunalitäten. Anders ausgedrückt: Die Kommunalitäten h_j^2 sind die Summen aller quadrierten Faktorladungen pro Variablen, also

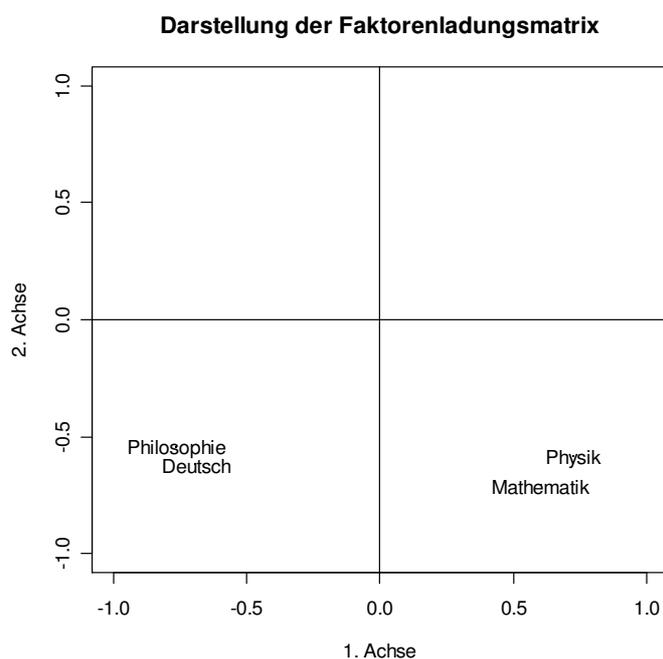
$$h_j^2 = \sum_{m=1}^K a_{jm}^2$$

Für alle vier Variablen klären also die beiden ermittelten Faktoren jeweils mehr als 80% der Varianz auf.

Die Abweichungen zwischen der „echten“ und der reproduzierten Korrelationsmatrix sind relativ gering:

R - R~	Deutsch	Philosophie	Mathematik	Physik
Deutsch	0,14	-0,13	-0,08	0,07
Philosophie	-0,13	0,12	0,06	-0,04
Mathematik	-0,08	0,06	0,14	-0,14
Physik	0,07	-0,04	-0,14	0,14

Die Komponenten der Faktorenladungsmatrix lassen sich entsprechend graphisch darstellen:



Als mögliche Interpretation ergibt sich ein bipolarer Faktor (Trennung in sprachliche und naturwissenschaftliche Begabung) und ein Generalfaktor („wer schlecht in der Schule ist, der hat in allen Fächern eine schlechtere Note als ein guter Schüler“).

Damit sind die Aufgabestellungen 1) bis 3) praktisch gelöst.

Rotation der Faktorenladungen

Die ermittelten Faktorenladungen stellen noch kein „Optimum im Sinn einer Einfachstruktur“ dar. Um eine solche Einfachstruktur zu erreichen, wird eine Rotation der Faktoren durchgeführt.

Wir hatten gesehen, dass $\tilde{\mathbf{R}} = \mathbf{A}\mathbf{A}'$. Multipliziert man nun \mathbf{A} v.r. mit einer orthogonalen Matrix \mathbf{I} , so erhält man für das Produkt aus $\mathbf{A}\mathbf{I}$ und seiner Transponierten $(\mathbf{A}\mathbf{I})(\mathbf{A}\mathbf{I})' = \mathbf{A}\mathbf{I}\mathbf{I}'\mathbf{A}' = \mathbf{A}\mathbf{A}' = \tilde{\mathbf{R}}$.

Also ist auch $\mathbf{A}\mathbf{I}$ eine Lösung für die o.g. Aufgabe. Die Lösungsgesamtheit von $\tilde{\mathbf{R}} = \mathbf{A}\mathbf{A}'$ ergibt sich aus einer Partikulärlösung und Rechtsmultiplikation mit einer orthogonalen Matrix \mathbf{I} passenden Formats⁴.

Eine Partikulärlösung wurde bereits gefunden durch $\mathbf{A} = \mathbf{E}\mathbf{\Lambda}_K^{1/2}$

Durch die zulässige Multiplikation von \mathbf{A} mit der orthogonalen Matrix \mathbf{I} wird eine (orthogonale) Rotation der Vektoren von \mathbf{A} durchgeführt. Eine im Sinn der Einfachstruktur passende Rotationsmatrix \mathbf{I} wird gefunden, indem eine Maximierung der Varianz der Spalten von \mathbf{A} vorgenommen wird. Dieses Vorgehen wird als „Varimaxmethode“ (nach Kaiser) bezeichnet. Vielfach wird die sog. „Kaiser-Normalisierung“ angewendet – dies bedeutet, dass die quadrierten Faktorladungen durch die Kommunalitäten dividiert werden.

⁴ Der Beweis dafür wird hier nicht explizit angegeben.

Man erhält eine Rotationsmatrix mit einer Drehung um den Winkel⁵ α im Gegenuhrzeigersinn (hier: für den zweidimensionalen Fall) durch

$$\underline{I} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}.$$

Praktisch erfolgt die Varimax-Optimierung, indem man für Rotationswinkel ($-45^\circ \dots +45^\circ$) die entsprechende Rotationsmatrix \underline{I} bestimmt. Die Rotationsmatrix wird von rechts mit der Faktorenladungsmatrix \underline{A} multipliziert, so dass

$$\underline{A}_r = \underline{A}\underline{I}$$

Durch die „Kaiser-Normalisierung“ werden die quadrierten Faktorladungen durch die Kommunalitäten geteilt. Also ist für jede vorgenommene Rotation die Summe der Varianz der Spalten für den Ausdruck

$$\underline{D}_K^{-1}(\underline{A}_r * \underline{A}_r)$$

zu bestimmen, wobei \underline{D}_K die Diagonalmatrix mit den Kommunalitäten bezeichnet. (Bei der Schreibweise $\underline{Y}*\underline{Z}$ handelt es sich um das sog. Hadamardprodukt – das Hadamardprodukt zweier Matrizen erhält man, indem man die Elemente der beiden Matrizen in entsprechender Position miteinander multipliziert.⁶)

Die Rotationsmatrix, bei der die Varianzsumme am größten wird, wird beibehalten und auf dieser Basis die Matrix der rotierten Faktorladungen bestimmt.

Im vorliegenden Beispiel ergibt sich ein Varianzmaximum bei einem Rotationswinkel von $42,15^\circ$, die Summe der Spaltenvarianzen ist hier 0,641. Dies ergibt eine Rotationsmatrix \underline{I}

⁵ Angabe des Winkels im Bogenmaß. Bei Angabe in Grad Umrechnung durch $w = \alpha/180*\pi$

⁶ Beispiel: $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} * \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 5 & 12 \\ 21 & 32 \end{pmatrix}$

0,74 -0,67
0,67 0,74

Die rotierte Faktorenladungsmatrix ist

A_r	Faktor 1	Faktor 2
Deutsch	-0,93	0,01
Philosophie	-0,93	0,11
Mathematik	-0,03	-0,92
Physik	0,15	-0,92

Zur besseren Lesbarkeit werden häufig Faktorladungen unterhalb eines bestimmten Wertes (z.B. Faktorladungen $< 0,2$) gestrichen:

A_r	Faktor 1	Faktor 2
Deutsch	-0,93	
Philosophie	-0,93	
Mathematik		-0,92
Physik		-0,92

Graphische Darstellung:

