Prof. Dr. Reinhold Kosfeld
Fachbereich Wirtschaftswissenschaften
Universität Kassel

Lösungen zu den Aufgaben zur Multivariaten Statistik
Teil 2: Aufgaben zur Faktorenanalyse

1. Worin besteht der Unterschied zwischen einer Hauptkomponentenanalyse und der Faktorenanalyse im engeren Sinne?

   Lösung:

   Die Hauptkomponentenmethode ist ein vereinfachtes Verfahren der Faktorenextraktion auf der Basis eines faktorenanalytischen Modells, das von einer Kommunalitätenschätzung absieht. Sie setzt bei der Faktorenextraktion an der Korrelationsmatrix **R** an, obwohl in den Einsen auf ihrer Hauptdiagonale die merkmalsspezifischen Varianzen enthalten sind, die nicht durch die gemeinsamen Faktoren erklärt werden können.

   Die Faktorenanalyse im engeren Sinne (z.B. Hauptachsen-Faktorenanalyse) berücksichtigt bei der Faktorenextraktion dagegen das Vorhandensein merkmalsspezifischer Varianzen durch eine Kommunalitätenschätzung. Die Einsen auf der Hauptdiagonale der Korrelationsmatrix **R** werden bei diesem Verfahren durch die geschätzten Kommunalitäten ersetzt. Die Korrelationsmatrix **R** geht dadurch in die reduzierte Korrelationsmatrix $\mathbf{R}_h$ über, aus der die gemeinsamen Faktoren extrahiert werden.

2. Was geben Eigenwerte in einer Faktorenanalyse wieder?

   Lösung:

   In der Faktorenanalyse gibt der Eigenwert eines Faktors den Teil der Gesamtvarianz der manifesten Variablen an, der durch ihn erklärt wird. Er muss der Summe seiner quadrierten Ladungen auf den beobachtbaren Variablen entsprechen. Dividiert man die Eigenwerte durch m (= Gesamtvarianz bei standardisierten Variablen), erhält man die durch die jeweiligen gemeinsamen Faktoren erklärten Anteile der Gesamtvarianz der Variablen.

3. Was spiegeln Kommunalitäten wider?

   Lösung:

   Die Kommunalitäten spiegeln die Varianzanteile der Variablen wider, die durch alle gemeinsamen Faktoren erklärt werden. Man erhält sie, in dem man die Summe der quadrierten Ladungen der jeweiligen Variablen auf allen gemeinsamen Faktoren bildet.

4. Was drückt das Fundamentaltheorem der Faktorenanalyse aus?

   Lösung:

   Das Fundamentaltheorem der Faktorenanalyse besagt, dass sich die Korrelationsmatrix **R** in einen durch die gemeinsamen Faktoren erklärten Teil und in einen durch sog. Einzelrestfaktoren (merkmalsspezifische Faktoren) erklärten Teil zerlegen lässt. Der auf die gemeinsamen Faktoren zurückzuführende Teil der Korrelationsmatrix, der durch das Produkt der Faktormatrix **A** und ihrer Transponierten **A**' gegeben ist, enthält Kommunalitäten und die substanziellen Korrelationen zwischen den Variablen. Der auf die Einzelrestfaktoren zurückführbare Teil wird als Restkorrelationsmatrix $\mathbf{R}^{Rest}$ bezeichnet. Hierin befinden sich – abgesehen von unbedeutenden Restkorrelationen (≈ 0) – die merkmalsspezifischen Varianzen.

5. Worin besteht das Kommunalitätenproblem?

   Lösung:

   Bei der Faktorenanalyse im engeren Sinne werden die gemeinsamen Faktoren aus der reduzierten Korrelationsmatrix $\mathbf{R}_h$ extrahiert, die sich aus der Korrelationsmatrix **R** ergibt, wenn man die Einsen auf der Hauptdiagonale um die merkmalsspezifischen Varianzen vermindert. Auf der Hauptdiagonale von $\mathbf{R}_h$ stehen die Kommunalitäten, d.h. die Varianzanteile der standardisierten Variablen, die auf die gemeinsamen Faktoren zurückgeführt werden können. Die Kommunalitäten sind im Allgemeinen jedoch a priori unbekannt und ergeben sich erst nach der Faktorenextraktion aus der Faktormatrix (=Matrix der Faktorladungen). Das Kommunalitätenproblem besteht darin, dass man die Kommunalitäten bereits für die Faktorenextraktion benötigt, obgleich sie sich erst nachher aus den ermittelten Faktorladungen berechnen lassen.

6. Was ist in der Faktorenanalyse mit einer Einfachstruktur gemeint?

   Lösung:

   Die gemeinsamen Faktoren werden so aus der Korrelationsmatrix extrahiert, dass der erste Faktor ein Maximum der Variablenvarianz erklärt, der zweite Faktor ein Maximum der durch den Faktor 1 noch nicht erklärten Varianz etc. Die Extraktion von Faktoren, die sukzessive maximale Varianzanteile der verbleibenden Variablenstreuung erklären, führt im Allgemeinen noch nicht zu einer bestmöglichen Interpretation. Vor allem mittlere Ladungen „verwischen" das Faktormuster, das deutlicher hervortreten würde, wenn die Faktorladungen in Richtung 0 (kein Zusammenhang zwischen einer Variablen und einem Faktor) oder ±1 (perfekter Zusammenhang zwischen einer Variablen und einem Faktor) tendieren würden. Ein solches Faktormuster bezeichnet man als Einfachstruktur.

7. Was versteht man unter dem Rotationsproblem?

   Lösung:

   In einem Koordinatensystem, dessen Achsen durch die gemeinsamen Faktoren gebildet werden (=Faktorraum), werden die Variablen durch die Variablenvektoren (=Zeilen der Faktormatrix) abgebildet, deren Längen durch die Wurzeln der Kommunalitäten gegeben sind. Dreht man das Koordinatensystem (Faktorraum) um einen Winkel α, ändern sich zwar die Koordinaten (=Faktorladungen), jedoch bleiben die Kommunalitäten unverändert. Bei gleichem Erklärungsgehalt der Faktoren kann der Faktorraum daher so gedreht („rotiert“) werden, dass seine Achsen möglichst durch die Variablencluster verlaufen. Auf diese Weise führt man die Faktorladungen näher an 0 oder ± 1, d.h. an die Werte einer Einfachstruktur, heran. Das Rotationsproblem besteht darin, festzulegen, um welchen Winkel das Koordinatensystem gedreht („rotiert“) werden soll, um eine Lösung zu erhalten, die möglichst nahe an eine Einfachstruktur herankommt und damit bestmöglich interpretierbar ist.

8. Auf welche Art führt die Varimax-Methode zu einer Einfachstruktur?

   Lösung:

   Mit der Varimax-Methode wird der Rotationswinkel α so bestimmt, dass die Streuung der quadrierten Ladungen der Faktoren um ihre Mittelwerte (=spaltenweise Mittelwerte der Ladungsquadrate) maximiert wird. Die Anzahl der Variablen mit einer hohen Ladung auf einem Faktor wird hierdurch minimiert. Auf diese Weise zielt die Varimax-Methode auf eine Einfachstruktur ab, in der möglichst wenige Variablen, dafür aber hoch, auf den einzelnen Faktoren laden.

9. Welche Voraussetzung im Hinblick auf das Skalenniveau wird hier bei der Durchführung einer Faktorenanalyse gemacht?

   Lösung:

   Die Anwendung der Faktorenanalyse setzt ein metrisches Skalenniveau voraus. Werden Merkmale herangezogen, die auf einer sog. Ratingskala (z.B. von 1 = sehr wichtig oder vollständige Zustimmung bis 7 = völlig unwichtig oder vollständige Ablehnung) gemessen werden, muss unterstellt werden können, dass die Abstände zwischen den Ratingwerten inhaltlich als gleich groß interpretierbar sind.

10. Gegeben sind die aus den Testergebnissen (=Punkte) von 20 Schülern in sechs Schulfächern ermittelten Korrelationen:

Korrelationen

		deutsch	mathe	englisch	bio	franz	physik
deutsch	Korrelation nach Pearson	1	-.580**	.794**	-.315	.681**	-.273
	Signifikanz (2-seitig)		.007	.000	.176	.001	.245
	N	20	20	20	20	20	20
mathe	Korrelation nach Pearson	-.580**	1	-.492*	.772**	-.516*	.546*
	Signifikanz (2-seitig)	.007		.028	.000	.020	.013
	N	20	20	20	20	20	20
englisch	Korrelation nach Pearson	.794**	-.492*	1	-.339	.857**	-.256
	Signifikanz (2-seitig)	.000	.028		.144	.000	.276
	N	20	20	20	20	20	20
bio	Korrelation nach Pearson	-.315	.772**	-.339	1	-.477*	.592**
	Signifikanz (2-seitig)	.176	.000	.144		.034	.006
	N	20	20	20	20	20	20
franz	Korrelation nach Pearson	.681**	-.516*	.857**	-.477*	1	-.248
	Signifikanz (2-seitig)	.001	.020	.000	.034		.291
	N	20	20	20	20	20	20
physik	Korrelation nach Pearson	-.273	.546*	-.256	.592**	-.248	1
	Signifikanz (2-seitig)	.245	.013	.276	.006	.291	
	N	20	20	20	20	20	20

**. Die Korrelation ist auf dem Niveau von 0,01 (2-seitig) signifikant.

*. Die Korrelation ist auf dem Niveau von 0,05 (2-seitig) signifikant.

a) Welche Anhaltspunkte für die Durchführung einer Faktorenanalyse liefert die Korrelationsmatrix?

Lösung:

Aus der Korrelationsmatrix geht hervor, dass die Testleistungen in den sechs Schulfächern nicht unabhängig voneinander sind. Vielmehr lässt sich aus der Reihe signifikanter Korrelationen auf substanzielle Zusammenhänge zwischen den Testleistungen schließen. So ist z.B. die Korrelation zwischen „franz“ und „englisch“ mit 0,857 beträchtlich. Der Korrelationskoeffizient zwischen „deutsch“ und „mathe“ von –0,580 zeigt dagegen einen negativen Zusammenhang zwischen diesen Schulfächern an. Aufgrund der vorhandenen substanzieller Korrelationen ist die Faktorenanalyse ist geeignet, die hinter den Testleistungen stehenden gemeinsamen Dimensionen (Faktoren oder „Supervariablen“) aus der Korrelationsmatrix zu extrahieren.

b) Testen Sie den Korrelationskoeffizienten der Mathematik und Englischpunkte auf Signifikanz (α=0,05)!

Lösung:

Prüfgröße: t = -2,398; Kritischer Wert: $t_{18;0,975}$ = 2,101; Testentscheidung: H_0 ablehnen

11. Bei einer Faktorenanalyse der Punkte in Schulfächern (Korrelationsmatrix s. Aufg. 10) unter Verwendung der Hauptkomponentenmethode gibt SPSS die folgende Tabelle „Erklärte Varianz“ aus:

Erklärte Gesamtvarianz

	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion		
Komponente	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	3.617	60.276	60.276	3.617	60.276	60.276
2	1.280	21.332	81.608	1.280	21.332	81.608
3	.477	7.958	89.566			
4	.384	6.405	95.971			
5	.139	2.325	98.296			
6	.102	1.704	100.000			

Extraktionsmethode: Hauptkomponentenanalyse.

Erläutern Sie anhand der SPSS-Tabelle „Erklärte Gesamtvarianz“ das Kaiser-Kriterium! Welches kumulierte Varianzkriterium würde hier zum selben Ergebnis im Hinblick auf die Faktorenextraktion führen?

Lösung:

Kaiser-Kriterium:
Nach dem Kaiser-Kriterium sind diejenigen Faktoren zu extrahieren, deren Eigenwert größer als 1 ist. Die Ratio dieses Kriteriums besteht darin, dass die extrahierten Faktoren (“Supervariablen“) einen größeren Erklärungsbeitrag leisten sollten als die beobachtbaren Variablen, da mit ihnen sonst keine Informationsverdichtung einhergehen würde. Der Schwellenwert von 1 ergibt sich daraus, dass die Varianz jeder Variablen bei einer Standardisierung gleich 1 ist. Gemeinsame Faktoren werden nach dem Kaiser-Kriterium somit nur dann extrahiert, wenn sie einen größeren Teil der Gesamtvarianz als eine manifeste Variable erklären. In der SPSS-Tabelle „Erklärte Gesamtvarianz“ sind die Eigenwerte der ersten beiden Faktoren > 1, während die Eigenwerte vom dritten Faktor ab < 1 sind. Nach dem Kaiser-Kriterium sind daher 2 gemeinsame Faktoren zu extrahieren.

Kumulatives Varianzkriterium:
Das hierzu kompatible kumulative Varianzkriterium würde die Forderung enthalten, dass durch die extrahierten Faktoren mindestens 80% der Gesamtvarianz der manifesten Variablen erklärt werden sollten.

12. Gegeben ist die von SPSS ausgewiesene Faktormatrix (Komponentenmatrix):

Komponentenmatrix[a]

	Komponente	
	1	2
deutsch	.804	.392
mathe	-.838	.324
englisch	.828	.477
bio	-.736	.532
franz	.835	.371
physik	-.586	.610

Extraktionsmethode: Hauptkomponentenanalyse.

a. 2 Komponenten extrahiert

a) Wie lassen sich die beiden größten Eigenwerte aus der nachstehenden Tabelle der Faktorladungen (=Komponentenladungen) reproduzieren?

Lösung:

Eigenwert λ_1: Summe der quadrierten Ladungen der 1. Spalte der Faktormatrix (Komponentenmatrix):

$$\lambda_1 = \sum_{j=1}^{m=6} a_{j1}^2 = 0{,}804^2 + (-0{,}838)^2 + 0{,}828^2 + (-0{,}736)^2 + 0{,}835^2 + (-0{,}586)^2 = 3{,}617$$

Eigenwert λ_2: Summe der quadrierten Ladungen der 2. Spalte der Faktormatrix (Komponentenmatrix):

$$\lambda_2 = \sum_{j=1}^{m=6} a_{j2}^2 = 0{,}392^2 + 0{,}324^2 + 0{,}477^2 + 0{,}532^2 + 0{,}371^2 + 0{,}610^2 = 1{,}279$$

b) Wie lassen sich die beiden extrahierten Faktoren (Komponenten) aus der in Teil b) angegebenen Faktormatrix (Komponentenmatrix) interpretieren?

Lösung:

Faktor 1: Gegenpol sprachliche vs. naturwissenschaftlich-mathematische Begabung

Faktor 2: Naturwissenschaftliche Begabung

Anmerkung: Die beiden Faktoren treten nach einer Rotation der Faktormatrix deutlicher hervor (→ Aufg. 14).

13. Welche anfänglichen Werte für die Kommunalitäten sind bei der Faktorenextraktion mit der Hauptkomponentenmethode und Hauptachsen-Faktorenanalyse verwendet worden?

Inverse Korrelationsmatrix

	deutsch	mathe	englisch	bio	franz	physik
deutsch	3.396	1.557	-2.143	-.944	-.108	.060
mathe	1.557	3.723	.050	-2.387	-.391	-.280
englisch	-2.143	.050	5.752	-.768	-3.716	.392
bio	-.944	-2.387	-.768	3.524	1.538	-.854
franz	-.108	-.391	-3.716	1.538	4.661	-.519
physik	.060	-.280	.392	-.854	-.519	1.646

Kommunalitäten

	Anfänglich	Extraktion
deutsch	?	.800
mathe	?	.807
englisch	?	.913
bio	?	.825
franz	?	.836
physik	?	.715

Extraktionsmethode: Hauptkomponentenanalyse.

Kommunalitäten

	Anfänglich	Extraktion
deutsch	?	.659
mathe	?	.774
englisch	?	.999
bio	?	.834
franz	?	.732
physik	?	.413

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

Lösung:

- Hauptkomponentenmethode:

Bei der Hauptkomponentenmethode wird keine Kommunalitätenschätzung durchgeführt. Die „anfänglichen Kommunalitäten“ entsprechen daher den Hauptdiagonalelementen der Korrelationsmatrix **R**, die alle gleich 1 sind.

- Hauptachsen-Faktorenanalyse:

SPSS verwendet bei der Hauptachsen-Faktorenanalyse (Faktorenanalyse i.e.S.) die multiplen Determinationskoeffizienten der Variablen als anfängliche Kommunalitäten, die sich durch $1 - 1/r^{jj}$ (r^{jj}: j-tes Hauptdiagonalelement der inversen Korrelationsmatrix $\mathbf{R}^{-1}$) berechnen lassen.

Anfängliche Kommunalitäten:

$$h_1^2 = 1 - \frac{1}{3{,}396} = 0{,}706; \quad h_2^2 = 1 - \frac{1}{3{,}723} = 0{,}731; \quad h_3^2 = 1 - \frac{1}{5{,}752} = 0{,}826;$$

$$h_4^2 = 1 - \frac{1}{3{,}524} = 0{,}716; \quad h_5^2 = 1 - \frac{1}{4{,}661} = 0{,}785; \quad h_6^2 = 1 - \frac{1}{1{,}646} = 0{,}392;$$

14. Bei Durchführung einer Faktorenanalyse i. e. S. (Hauptachsen-Faktorenanalyse) gibt SPSS im Beispiel der Punktnoten in Schulfächern (Korrelationsmatrix s. Aufg. 10) die nachfolgende Tabelle „Erklärte Gesamtvarianz“ aus:

Erklärte Gesamtvarianz

	Anfängliche Eigenwerte			Summen von quadrierten Faktorladungen für Extraktion			Rotierte Summe der quadrierten Ladungen		
Faktor	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %	Gesamt	% der Varianz	Kumulierte %
1	3.617	60.276	60.276	3.387	56.447	56.447	2.426	40.429	40.429
2	1.280	21.332	81.608	1.023	17.056	73.502	1.984	33.074	73.502
3	.477	7.958	89.566						
4	.384	6.405	95.971						
5	.139	2.325	98.296						
6	.102	1.704	100.000						

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

a) Warum unterscheiden sich die Werte in der Spalte „Kumulierte %“ vor und nach der Extraktion?

Lösung:

Die Spalte „Kumulierte %“ vor der Extraktion bezieht sich auf die Korrelationsmatrix **R**, während sich die Angaben in dieser Spate nach der Extraktion auf die reduzierte Korrelationsmatrix $\mathbf{R}_h$ bezieht.

b) Vergleichen Sie die Kommunalitäten der unrotierten und rotierten Faktormatrix miteinander! Erläutern Sie den Befund!

Faktorenmatrix[a]

	Faktor	
	1	2
deutsch	.757	.293
mathe	-.807	.350
englisch	.863	.505
bio	-.719	.563
franz	.806	.287
physik	-.502	.400

Extraktionsmethode: Hauptachsen-Faktorenanalyse.

a. Es wurde versucht, 2 Faktoren zu extrahieren. Es werden mehr als 13 Iterationen benötigt. (Konvergenz= 002). Die Extraktion wurde abgebrochen.

Rotierte Faktorenmatrix[a]

	Faktor	
	1	2
deutsch	.770	-.257
mathe	-.398	.785
englisch	.986	-.161
bio	-.194	.892
franz	.804	-.292
physik	-.132	.629

Extraktionsmethode: Hauptachsen-Faktorenanalyse.
Rotationsmethode: Varimax mit Kaiser-Normalisierung.

a. Die Rotation ist in 3 Iterationen konvergiert.

Lösung:

Kommunalität der Variablen X_1 („deutsch“):	
$h_1^2 = \sum_{k=1}^{p=2} a_{1k}^2 = 0{,}757^2 + 0{,}293^2 = 0{,}659$	$h_1^2 = \sum_{k=1}^{p=2} a_{1k}^2 = 0{,}770^2 + (-0{,}257)^2 = 0{,}659$

Wenn Sie die Kommunalitäten der übrigen Variablen berechnen, stellen Sie fest, dass es hierbei keinen Unterschied zwischen der unrotierten und rotierten Lösung gibt (bitte überprüfen!). Der durch die beiden extrahierten Faktoren erklärte Teil der Varianz einer manifesten Variablen (=Kommunalität) ändert sich nicht durch eine Rotation der Faktormatrix.

c) Interpretieren Sie das in Teil b) durch die rotierte Faktormatrix wiedergegebene Faktormuster!

Lösung:

Die Variablen „deutsch“, „englisch“ und „franz“ laden auf dem ersten Faktor hoch. Es liegt nahe, Faktor 1 als „Sprachliche Begabung“ zu interpretieren. Dagegen laden die Variablen „mathe“, „bio“ und „physik“ substanziell auf dem zweiten Faktor, der deshalb als „Mathematisch-naturwissenschaftliche Begabung“ interpretiert werden kann.

15. Welche Informationen können Sie der Transformationsmatrix entnehmen?

Faktor-Transformationsmatrix

Faktor	1	2
1	.770	-.638
2	.638	.770

Extraktionsmethode: Hauptachsen-Faktorenanalyse.
Rotationsmethode: Varimax mit Kaiser-Normalisierung.

Lösung:

Die Transformationsmatrix enthält die Sinus- und Kosinus-Funktionen des Drehwinkels, d.h. des Winkels, mit dem der Faktorraum gedreht („rotiert“) wird, um eine Einfachstuktur zu erhalten. Die beiden Hauptdiagonalelemente geben den Wert der Kosinusfunktion 0,770 für den Drehwinkel α an. Man erhält den Drehwinkel α aus der Arcus-Kosinus-Funktion (=inverse Kosinusfunktion): arccos(0,770) = 39,6°. Die beiden Nicht-Hauptdiagonalelemente geben den Wert der Sinusfunktion 0,638 bzw. negativen Sinusfunktion -0,638 an. Hiermit lässt sich ebenfalls der Drehwinkel bestimmen α bestimmen: arcsin(0,638) = 39,6°.

16. Interpretieren Sie die Koeffizientenmatrix der Faktorwerte:

Koeffizientenmatrix der Faktorwerte

	Faktor	
	1	2
deutsch	-.029	-.081
mathe	-.048	.320
englisch	1.114	.311
bio	.130	.651
franz	-.075	.003
physik	.076	.126

Extraktionsmethode: Hauptachsen-Faktorenanalyse.
Rotationsmethode: Varimax mit Kaiser-Normalisierung.
Methode für Faktorwerte: Regression.

a) Woraus besteht die Koeffizientenmatrix der Faktorwerte? Wie erhält man hiermit die Matrix der Faktorwerte?

Lösung:

Bei Wahl der „Regressionsmethode“ zur Schätzung der Faktorwertematrix **F** in der Hauptachsen-Faktorenanalyse besteht die Koeffizientenmatrix der Faktorwerte aus dem Produkt der transponierten Faktormatrix, **A**’, und der inversen Korrelationsmatrix $\mathbf{R}^{-1}$. Man erhält die transponierte Faktorwertematrix **F**’ nach Multiplikation der Koeffizientenmatrix $\mathbf{A}'\mathbf{R}^{-1}$ mit der transponierten standardisierten Beobachtungsmatrix **Z**’. SPSS gibt die Koeffizientenmatrix in der Form $(\mathbf{A}'\mathbf{R}^{-1})'=\mathbf{R}^{-1}\mathbf{A}$ aus.

b) Für drei ausgewählte Schüler sind folgende Faktorwerte ermittelt worden:

	Faktor 1	Faktor 2
Schüler 1	-0,708	-0,642
Schüler 4	0,890	1,214
Schüler 20	1,196	0,082

Hinweis: Die hier neu ausgewiesenen Faktorwerte ergeben sich nach Rotation der Faktormatrix.

Interpretieren Sie die Faktorwerte!

Lösung:

Schüler 1: $f_{11} = -0{,}708$; $f_{12} = -0{,}642$

Sowohl die sprachliche Begabung (Faktor 1) als auch die mathematisch-naturwissenschaftliche Begabung (Faktor 2) sind beim Schüler 1 unterdurchschnittlich ausgeprägt.

Schüler 4: $f_{41} = 0{,}890$; $f_{42} = 1{,}214$

Der Schüler 4 hat eine überdurchschnittliche sprachliche und mathematisch-naturwissenschaftliche Begabung.

Schüler 20: $f_{20,1} = 1{,}196$; $f_{20,2} = 0{,}082$

Der Schüler 20 hat eine hohe sprachliche Begabung (Faktor 1), aber nur eine durchschnittliche mathematisch-naturwissenschaftliche Begabung (Faktor 2).

c) Der Schüler 1 hat in den 6 Fächern folgende Punktwerte erzielt:

deutsch	mathe	englisch	bio	franz	physik
-0,671	0,087	-0,481	-0,638	0,125	-1,257

Reproduzieren Sie die für diesen Schüler in Teil b) ausgewiesenen Faktorwerte!

Lösung:

Faktorwert $f_{11} = -0{,}708$

Spalte der Koeffizientenmatrix der Faktorwerte x Vektor der standardisierten Beobachtungswerte des 1. Schülers:

$$\begin{bmatrix} -0{,}029 & -0{,}048 & 1{,}114 & 0{,}130 & -0{,}075 & 0{,}076 \end{bmatrix} \cdot \begin{bmatrix} -0{,}0671 \\ 0{,}087 \\ -0{,}481 \\ -0{,}638 \\ 0{,}125 \\ -1{,}257 \end{bmatrix}$$

= (-0,029)·(-0,671) + (-0,048)·0,087 + 1,114·(-0,481) + 0,130·(-0,638) + (-0,075)·0,125 + 0,076·(1,257) = -0,708

Faktorwert f_{12} = -0,642

$$\begin{bmatrix} -0{,}081 & 0{,}320 & 0{,}311 & 0{,}651 & 0{,}003 & 0{,}126 \end{bmatrix} \cdot \begin{bmatrix} -0{,}0671 \\ 0{,}087 \\ -0{,}481 \\ -0{,}638 \\ 0{,}125 \\ -1{,}257 \end{bmatrix}$$

= (-0,081)·(-0,671) + (0,320·0,087 + 0,311·(-0,481) + 0,651·(-0,638) + 0,003·0,125 + 0,126·(1,257) = -0,641

17. Gegeben ist die Korrelationsmatrix

$$\mathbf{R} = \begin{bmatrix} 1 & 0{,}8 \\ 0{,}8 & 1 \end{bmatrix}$$

der beiden Variablen X_1 und X_2, aus der zwei Faktoren nach der Hauptkomponentenmethode zu extrahieren sind.

a) Welche Eigenwerte hat die Korrelationsmatrix?

<u>Lösung:</u>
Charakteristische Gleichung:

$$|\mathbf{R} - \lambda\mathbf{I}| = \begin{vmatrix} 1-\lambda & 0{,}8 \\ 0{,}8 & 1-\lambda \end{vmatrix} = 0$$

$$|\mathbf{R} - \lambda\mathbf{I}| = (1-\lambda)^2 - 0{,}8^2 = 1 - 2\lambda + \lambda^2 - 0{,}64$$

$$\lambda^2 \underbrace{-2}_{=a} \cdot \lambda + \underbrace{0{,}36}_{=b} = 0$$

$$\lambda_{1/2} = -\frac{a}{2} \pm \sqrt{\left(\frac{a}{2}\right) - b} = -\frac{-2}{2} \pm \sqrt{\left(\frac{-2}{2}\right)^2 - 0{,}36}$$
$$= 1 \pm \sqrt{1 - 0{,}36} = 1 \pm \sqrt{0{,}64} = 1 \pm 0{,}8$$
$$\Rightarrow \lambda_1 = 1{,}8 \quad \text{und} \quad \lambda_2 = 0{,}2$$

b) Bestimmen Sie den ersten Faktorvektor!

<u>Lösung:</u>

Gleichungssystem mit 1. Eigenwert $(\lambda_1 = 1{,}8)$:

$$(\mathbf{R} - \lambda_1 \mathbf{I}) \cdot a_1 = 0$$
$$\begin{bmatrix} 1-1{,}8 & 0{,}8 \\ 0{,}8 & 1-1{,}8 \end{bmatrix} \cdot \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

oder

$$\left.\begin{aligned} -0{,}8 \cdot a_{11} + 0{,}8 \cdot a_{21} &= 0 \\ 0{,}8 \cdot a_{11} - 0{,}8 \cdot a_{21} &= 0 \end{aligned}\right\}$$

Beide Gleichungen sind nicht linear unabhängig, da sich die 2. Gleichung durch Multiplikation mit (-1) in die 1. Gleichung überführen lässt.

Für eine vorläufige Lösung (a_{11}^*, a_{21}^*) setzt man daher eine der beiden Unbekannten fest, z.B. $a_{11}^* = 1$. Man kann dann die zweite Unbekannte a_{21}^* aus einer der beiden Gleichungen bestimmen. Mit der ersten Gleichung erhält man:

$$-0{,}8 \cdot (a_{11}^* = 1) + 0{,}8 \cdot a_{21}^* = 0$$
$$0{,}8 \cdot a_{21}^* = 0{,}8 \Rightarrow a_{21}^* = 1$$

Als zusätzliche Gleichung wählt man:

$$a_{11}^{*2} + a_{21}^{*2} = 1^2 + 1^2 = 2$$

Normierter 1. Eigenvektor (auf Länge 1):

$$a_{11}^n = a_{11}^* / \sqrt{2} = 1/\sqrt{2} = 0{,}707, \qquad a_{21}^n = a_{21}^* / \sqrt{2} = 1/\sqrt{2} = 0{,}707$$

$$\mathbf{a}_1^n = \begin{bmatrix} a_{11}^n = 0{,}707 \\ a_{21}^n = 0{,}707 \end{bmatrix}$$

Erster Faktorvektor:

$$\mathbf{a_1} = \begin{bmatrix} a_{11} = 0{,}949 \\ a_{21} = 0{,}949 \end{bmatrix} \left| \begin{aligned} a_{11} &= a_{11}^n \cdot \sqrt{\lambda_1} = 0{,}707 \cdot \sqrt{1{,}8} = 0{,}949 \\ a_{21} &= a_{21}^n \cdot \sqrt{\lambda_1} = 0{,}707 \cdot \sqrt{1{,}8} = 0{,}949 \end{aligned} \right.$$

$$\left[\begin{array}{c} \text{Anmerkung: Es gilt} \\ \mathbf{a}_1' \mathbf{a}_1 = a_{11}^2 + a_{21}^2 = \left(\sqrt{1{,}8}/\sqrt{2}\right)^2 + \left(\sqrt{1{,}8}/\sqrt{2}\right)^2 = 1{,}8/2 + 1{,}8/2 \\ = 2 \cdot (1{,}8/2) = 1{,}8 = \lambda_1 \end{array} \right]$$

c) Bestimmen Sie den zweiten Faktorvektor!

<u>Lösung</u>:

Gleichungssystem mit 2. Eigenwert $(\lambda_2 = 0{,}2)$:

$$(\mathbf{R} - \lambda_2 \mathbf{I}) \cdot \mathbf{a_2} = \mathbf{0}$$

$$\begin{bmatrix} 1-0{,}2 & 0{,}8 \\ 0{,}8 & 1-0{,}2 \end{bmatrix} \cdot \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

oder

$$\left.\begin{aligned} 0{,}8 \cdot a_{12} + 0{,}8 \cdot a_{22} &= 0 \\ 0{,}8 \cdot a_{12} + 0{,}8 \cdot a_{22} &= 0 \end{aligned}\right\} \text{Beide Gleichungen sind identisch}$$

Vorläufige Lösung:

$$a_{12}^* = 1 \quad (\text{gesetzt})$$

$$0{,}8 \cdot \left(a_{12}^* = 1\right) + 0{,}8 \cdot a_{22}^* = 0$$

$$0{,}8 \cdot a_{22}^* = -0{,}8 \Rightarrow a_{22}^* = -1$$

Als zusätzliche Gleichung erhält man:

$$a_{12}^{*2} + a_{22}^{*2} = 1^2 + (-1)^2 = 2$$

Normierter 2. Eigenvektor (auf Länge 1):

$$a_{12}^n = a_{12}^* / \sqrt{2} = 1/\sqrt{2} = 0{,}707, \qquad a_{22}^n = a_{22}^* / \sqrt{2} = -1/\sqrt{2} = -0{,}707$$

$$\mathbf{a}_2^n = \begin{bmatrix} a_{12}^n = 0{,}707 \\ a_{22}^n = -0{,}707 \end{bmatrix}$$

Zweiter Faktorvektor:

$$\mathbf{a}_2 = \begin{bmatrix} a_{12} = 0{,}316 \\ a_{22} = -0{,}316 \end{bmatrix} \left| \begin{aligned} a_{12} &= a_{12}^n \cdot \sqrt{\lambda_2} = 0{,}707 \cdot \sqrt{0{,}2} = 0{,}316 \\ a_{22} &= a_{22}^n \cdot \sqrt{\lambda_{21}} = -0{,}707 \cdot \sqrt{0{,}2} = -0{,}316 \end{aligned} \right.$$

$$\left[\begin{array}{l} \text{Anmerkung: Es gilt} \\ \qquad \mathbf{a}_2' \mathbf{a}_2 = a_{12}^2 + a_{22}^2 = \left(\sqrt{0{,}2}/\sqrt{2}\right)^2 + \left(-\sqrt{0{,}2}/\sqrt{2}\right)^2 \\ \qquad\qquad = 0{,}2/2 + 0{,}2/2 = 2 \cdot 0{,}2/2 = 0{,}2 = \lambda_2 \end{array} \right]$$

d) Zeigen Sie, dass die beiden extrahierten Faktoren orthogonal sind!

Lösung:

Im Falle orthogonaler Faktoren (unabhängiger Faktoren) ist das Skalarprodukt der Faktorvektoren $\mathbf{a}_1$ und $\mathbf{a}_2$ gleich 0:

$$\begin{aligned}\mathbf{a}_1'\mathbf{a}_2 &= a_{11}\cdot a_{12} + a_{21}\cdot a_{22}\\ &= \left(\sqrt{1{,}8}/\sqrt{2}\right)\cdot\left(\sqrt{0{,}2}/\sqrt{2}\right)+\left(\sqrt{1{,}8}/\sqrt{2}\right)\cdot\left(-\sqrt{0{,}2}/\sqrt{2}\right)\\ &= 2\cdot\sqrt{1{,}8}\cdot\sqrt{0{,}2} - 2\cdot\sqrt{1{,}8}\cdot\sqrt{0{,}2} = 0\end{aligned}$$

18. Der Preis eines Produktes A sei durch die Einflussgrößen Produktivität (X_1), Lohn (X_2) und Nachfrage (X_3) bestimmt).

Mit den Variablenwerten von 30 Unternehmen, die das Produkt A anbieten, ergibt sich die Korrrelationsmatrix der 3 Einflussgrößen:

$$\mathbf{R} = \begin{bmatrix} 1 & 0{,}7 & 0{,}3 \\ 0{,}7 & 1 & 0 \\ 0{,}3 & 0 & 1 \end{bmatrix}.$$

a) Ermitteln Sie die Eigenwerte der Korrelationsmatrix!

Lösung:

Lösen der charakteristischen Gleichung z.B. unter Anwendung der Sarrusschen Regel:

$\lambda_1 = 1{,}762$; $\lambda_2 = 1$; $\lambda_3 = 0{,}238$

b) Geben Sie die Anzahl der zu extrahierenden Faktoren nach dem kumulierten Varianzkriterium von 90% an!

Lösung:

Durch die Faktoren erklärte Varianzanteile:

$$\frac{\lambda_1}{m} = \frac{1{,}762}{3} = 0{,}587\,; \quad \frac{\lambda_2}{m} = \frac{1}{3} = 0{,}333\,; \quad \frac{\lambda_3}{m} = \frac{0{,}238}{3} = 0{,}079$$

$$\frac{\lambda_1}{m} + \frac{\lambda_2}{m} = 0{,}587 + 0{,}333 = 0{,}920 \quad (\hat{=}\ 92{,}0\%) \quad \Rightarrow \quad \text{2 Faktoren extrahieren}$$

c) Bestimmen Sie die Faktormatrix (=Matrix der Faktorladungen) für das in Teil b) angegebene kumulative Varianzkriterium nach der Hauptkomponentenmethode!

Lösung:

$$\text{Faktormatrix } \mathbf{A} = \begin{array}{c} \begin{array}{cc} \text{Faktor 1} & \text{Faktor 2} \end{array} \\ \begin{bmatrix} 0{,}939 & 0 \\ 0{,}863 & -0{,}394 \\ 0{,}370 & 0{,}919 \end{bmatrix} \end{array} \qquad \begin{array}{l} \text{Produktivität } (X_1) \\ \text{Lohn } (X_2) \\ \text{Nachfrage } (X_3) \end{array}$$

d) Wie lassen sich die extrahierten Faktoren (Hauptkomponenten) interpretieren?

Lösung:

Faktor 1: lädt hoch auf den Variablen Produktivität (X_1) und Lohn (X_2), d.h. auf angebotsseitig determinierte Variablen → Angebotsfaktor

Faktor 2: lädt hoch auf der Variablen Nachfrage (X_3) → Nachfragefaktor