A Mathematische Grundlagen

A.1 Matrizenrechnung

Das Erlernen und die Anwendung multivariater Verfahren setzt insbesondere Grundkenntnisse der Matrizenrechnung voraus. So sind zum Beispiel auf Seite 207 folgende Umformungen zu finden:

$$
\begin{aligned}
S(\boldsymbol{\beta}) &= (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})' \, (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) && \text{(A.1)} \\
&= (\mathbf{y}' - (\mathbf{X}\boldsymbol{\beta})') \, (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) && \text{(A.2)} \\
&= (\mathbf{y}' - \boldsymbol{\beta}'\mathbf{X}') \, (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) && \text{(A.3)} \\
&= \mathbf{y}'\mathbf{y} - \mathbf{y}'\mathbf{X}\boldsymbol{\beta} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} && \text{(A.4)} \\
&= \mathbf{y}'\mathbf{y} - (\mathbf{y}'\mathbf{X}\boldsymbol{\beta})' - \boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} && \text{(A.5)} \\
&= \mathbf{y}'\mathbf{y} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{y} - \boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} && \text{(A.6)} \\
&= \mathbf{y}'\mathbf{y} - 2\boldsymbol{\beta}'\mathbf{X}'\mathbf{y} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta} \, . && \text{(A.7)}
\end{aligned}
$$

Dabei ist $\mathbf{y}$ ein n-dimensionaler Vektor, $\mathbf{X}$ eine $(n, k+1)$-Matrix und $\boldsymbol{\beta}$ ein $k+1$-dimensionaler Vektor. Der Übergang von einer zur nächsten Zeile erfolgt jeweils nach einer bestimmten Regel. Diese Regeln werden im weiteren Verlauf dieses Kapitels nach und nach dargestellt. Vorausgeschickt sei dabei, dass folgende Gleichungen benutzt werden, um von einer Zeile zur nächsten zu gelangen:

von (A.1) zu (A.2) : (A.14)
von (A.2) zu (A.3) : (A.24)
von (A.3) zu (A.4) : (A.23)
von (A.4) zu (A.5) : (A.10)
von (A.5) zu (A.6) : (A.9) und (A.24)

Die Regel beim Übergang von (A.6) zu (A.7) kennt man aus der elementaren Mathematik.

A. Handl, *Multivariate Analysemethoden*, Statistik und ihre Anwendungen, 2nd ed.,
DOI 10.1007/978-3-642-14987-0,

A.1.1 Definitionen und spezielle Matrizen

Definition 24. *Eine (n, p)-Matrix $\mathbf{A}$ ist ein rechteckiges Schema reeller Zahlen, das aus n Zeilen und p Spalten besteht:*

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & \dots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{np} \end{pmatrix}. \tag{A.8}$$

Dabei heißt (n, p) die Ordnung der Matrix, i Zeilenindex, j Spaltenindex und a_{ij} Element der Matrix, das in der i-ten Zeile und j-ten Spalte steht. Wir schreiben kurz $\mathbf{A} = (a_{ij})$. Eine $(1, 1)$-Matrix ist ein Skalar a.

Beispiel 70. Wir betrachten im Folgenden die Matrizen

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}.$$

□

Definition 25. *Vertauscht man bei einer (n, p)-Matrix $\mathbf{A}$ Zeilen und Spalten, so erhält man die transponierte Matrix $\mathbf{A}'$ von $\mathbf{A}$.*

Beispiel 70. (fortgesetzt) Es gilt

$$\mathbf{C}' = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \end{pmatrix}.$$

□

Es gilt

$$(\mathbf{A}')' = \mathbf{A}. \tag{A.9}$$

Für einen Skalar a gilt

$$a' = a\,. \tag{A.10}$$

Definition 26. *Eine Matrix $\mathbf{A}$, für die $\mathbf{A}' = \mathbf{A}$ gilt, heißt symmetrisch.*

Beispiel 70. (fortgesetzt) Die Matrizen $\mathbf{A}$ und $\mathbf{D}$ sind symmetrisch. □

Definition 27. *Eine $(n, 1)$-Matrix heißt n-dimensionaler Spaltenvektor $\mathbf{a}$ und eine $(1, n)$-Matrix n-dimensionaler Zeilenvektor $\mathbf{b}'$.*

Beispiel 70. (fortgesetzt) Im Folgenden betrachten wir die beiden Spaltenvektoren

$$\mathbf{a}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

□

Die beiden Vektoren des Beispiels sind die Spalten der Matrix $\mathbf{A}$. Eine (n, p)-Matrix $\mathbf{A}$ besteht aus p n-dimensionalen Spaltenvektoren $\mathbf{a}_1, \ldots, \mathbf{a}_p$. Wir schreiben hierfür auch

$$\mathbf{A} = (\mathbf{a}_1, \ldots, \mathbf{a}_p).$$

Entsprechend können wir die Matrix aus Zeilenvektoren aufbauen.

Eine Matrix, die aus lauter Nullen besteht, heißt *Nullmatrix* $\mathbf{0}$. Wir bezeichnen den n-dimensionalen Spaltenvektor, der aus lauter Nullen besteht, ebenfalls mit $\mathbf{0}$ und nennen ihn den *Nullvektor*. Eine Matrix, die aus lauter Einsen besteht, bezeichnen wir mit $\mathbf{E}$. Der n-dimensionale Spaltenvektor $\mathbf{1}$, der aus lauter Einsen besteht, heißt *Einservektor* oder *summierender Vektor*. Wir werden später eine Begründung für die letzte Bezeichnungsweise liefern. Ein Vektor, bei dem die i-te Komponente gleich 1 und alle anderen Komponenten gleich 0 sind, heißt i-ter *Einheitsvektor* $\mathbf{e}_i$.

Definition 28. *Eine (n, p)-Matrix heißt quadratisch, wenn gilt $n = p$.*

Beispiel 70. (fortgesetzt) Die Matrizen $\mathbf{A}$, $\mathbf{B}$ und $\mathbf{D}$ sind quadratisch. □

Definition 29. *Eine quadratische Matrix, bei der alle Elemente außerhalb der Hauptdiagonalen gleich Null sind, heißt Diagonalmatrix.*

Beispiel 70. (fortgesetzt) Die Matrix $\mathbf{D}$ ist eine Diagonalmatrix. □

Definition 30. *Sind bei einer (n, n)-Diagonalmatrix alle Hauptdiagonalelemente gleich 1, so spricht man von der Einheitsmatrix $\mathbf{I}_n$.*

A.1.2 Matrixverknüpfungen

Definition 31. *Sind $\mathbf{A}$ und $\mathbf{B}$ (n, p)-Matrizen, dann ist die Summe $\mathbf{A} + \mathbf{B}$ definiert durch*

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} a_{11} + b_{11} & \ldots & a_{1p} + b_{1p} \\ \vdots & \ddots & \vdots \\ a_{n1} + b_{n1} & \ldots & a_{np} + b_{np} \end{pmatrix}.$$

Die Differenz $\mathbf{A} - \mathbf{B}$ ist definiert durch

$$\mathbf{A} - \mathbf{B} = \begin{pmatrix} a_{11} - b_{11} & \ldots & a_{1p} - b_{1p} \\ \vdots & \ddots & \vdots \\ a_{n1} - b_{n1} & \ldots & a_{np} - b_{np} \end{pmatrix}.$$

Beispiel 70. (fortgesetzt) Es gilt

$$\mathbf{A}+\mathbf{B}=\begin{pmatrix}2&0\\2&2\end{pmatrix}$$

und

$$\mathbf{A}-\mathbf{B}=\begin{pmatrix}2&2\\0&2\end{pmatrix}.$$

□

Schauen wir uns einige Rechenregeln für die Summe und Differenz von Matrizen an. Dabei sind $\mathbf{A}$, $\mathbf{B}$ und $\mathbf{C}$ (n,p)-Matrizen. Es gilt

$$\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}, \tag{A.11}$$

$$(\mathbf{A}+\mathbf{B})+\mathbf{C}=\mathbf{A}+(\mathbf{B}+\mathbf{C}), \tag{A.12}$$

$$(\mathbf{A}+\mathbf{B})'=\mathbf{A}'+\mathbf{B}' \tag{A.13}$$

und

$$(\mathbf{A}-\mathbf{B})'=\mathbf{A}'-\mathbf{B}'\,. \tag{A.14}$$

Eine Matrix kann mit einem Skalar multipliziert werden.

Definition 32. *Ist* $\mathbf{A}$ *eine* (n,p)*-Matrix und* $k\in\mathbb{R}$ *ein Skalar, dann ist das Produkt* $k\mathbf{A}$ *definiert durch*

$$k\mathbf{A}=\begin{pmatrix}k\,a_{11}&\dots&k\,a_{1p}\\\vdots&\ddots&\vdots\\k\,a_{n1}&\dots&k\,a_{np}\end{pmatrix}.$$

Beispiel 70. (fortgesetzt) Es gilt

$$2\mathbf{A}=\begin{pmatrix}4&2\\2&4\end{pmatrix}.$$

□

Sind $\mathbf{A}$ und $\mathbf{B}$ (n,p)-Matrizen und k und l Skalare, dann gilt

$$k(\mathbf{A}+\mathbf{B})=k\mathbf{A}+k\mathbf{B} \tag{A.15}$$

und

$$(k+l)\mathbf{A}=k\mathbf{A}+l\mathbf{A}\,. \tag{A.16}$$

Geeignet gewählte Matrizen kann man miteinander multiplizieren. Das Produkt von Matrizen beruht auf dem inneren Produkt von Vektoren.

Definition 33. *Seien*

$$\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

und

$$\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

n-dimensionale Spaltenvektoren. Das innere Produkt von $\mathbf{a}$ *und* $\mathbf{b}$ *ist definiert durch*

$$\mathbf{a}'\mathbf{b} = \sum_{i=1}^{n} a_i\, b_i .$$

Beispiel 70. (fortgesetzt) Es gilt

$$\mathbf{a}_1'\mathbf{a}_2 = 2 \cdot 1 + 1 \cdot 2 = 4 .$$

□

Offensichtlich gilt

$$\mathbf{a}'\mathbf{b} = \mathbf{b}'\mathbf{a}. \tag{A.17}$$

Bildet man $\mathbf{a}'\mathbf{a}$, so erhält man gerade die Summe der quadrierten Komponenten von $\mathbf{a}$:

$$\mathbf{a}'\mathbf{a} = \sum_{i=1}^{n} a_i\, a_i = \sum_{i=1}^{n} a_i^2 . \tag{A.18}$$

Beispiel 70. (fortgesetzt) Es gilt

$$\mathbf{a}_1'\mathbf{a}_1 = 2^2 + 1^2 = 5 .$$

□

Die Länge $||\mathbf{a}||$ eines n-dimensionalen Vektors $\mathbf{a}$ ist definiert durch

$$||\mathbf{a}|| = \sqrt{\mathbf{a}'\mathbf{a}} . \tag{A.19}$$

Dividiert man einen Vektor durch seine Länge, so spricht man von einem normierten Vektor. Die Länge eines normierten Vektors ist gleich 1.

Beispiel 70. (fortgesetzt) Es gilt

$$\frac{1}{\sqrt{\mathbf{a}_1'\ \mathbf{a}_1}}\, \mathbf{a}_1 = \frac{1}{\sqrt{5}} \begin{pmatrix} 2 \\ 1 \end{pmatrix} .$$

□

Das innere Produkt des n-dimensionalen Einservektors $\mathbf{1}$ mit einem n-dimensionalen Vektor $\mathbf{a}$ liefert die Summe der Komponenten von $\mathbf{a}$. Deshalb heißt $\mathbf{1}$ auch summierender Vektor.

Nun können wir uns dem Produkt zweier Matrizen zuwenden.

Definition 34. *Das Produkt* $\mathbf{AB}$ *einer* (n, p)*-Matrix* $\mathbf{A}$ *und einer* (p, q)*-Matrix* $\mathbf{B}$ *ist definiert durch*

$$\mathbf{AB} = \begin{pmatrix} \sum_{k=1}^{p} a_{1k}b_{k1} & \dots & \sum_{k=1}^{p} a_{1k}b_{kq} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{p} a_{nk}b_{k1} & \dots & \sum_{k=1}^{p} a_{nk}b_{kq} \end{pmatrix}.$$

Das Element in der i-ten Zeile und j-ten Spalte von $\mathbf{AB}$ erhält man, indem man das innere Produkt aus dem i-ten Zeilenvektor von $\mathbf{A}$ und dem j-ten Spaltenvktor von $\mathbf{B}$ bildet. Das Produkt $\mathbf{AB}$ einer (n, p)-Matrix $\mathbf{A}$ und einer (p, q)-Matrix $\mathbf{B}$ ist eine (n, q)-Matrix.

Beispiel 70. (fortgesetzt) Es gilt

$$\mathbf{AB} = \begin{pmatrix} 1 & -2 \\ 2 & -1 \end{pmatrix}$$

und

$$\mathbf{BA} = \begin{pmatrix} -1 & -2 \\ 2 & 1 \end{pmatrix}.$$

□

Das Beispiel zeigt, dass $\mathbf{AB}$ nicht notwendigerweise gleich $\mathbf{BA}$ ist.

Ist $\mathbf{A}$ eine (n, p)-Matrix, $\mathbf{B}$ eine (p, q)-Matrix und $\mathbf{C}$ eine (q, r)-Matrix, dann gilt

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})\,. \tag{A.20}$$

Ist $\mathbf{A}$ eine (n, p)-Matrix, $\mathbf{B}$ eine (p, q)-Matrix und k ein Skalar, dann gilt

$$k\mathbf{AB} = \mathbf{A}k\mathbf{B} = \mathbf{AB}k. \tag{A.21}$$

Sind $\mathbf{A}$ und $\mathbf{B}$ (n, p)-Matrizen und $\mathbf{C}$ und $\mathbf{D}$ (p, q)-Matrizen, dann gilt

$$(\mathbf{A} + \mathbf{B})(\mathbf{C} + \mathbf{D}) = \mathbf{AC} + \mathbf{AD} + \mathbf{BC} + \mathbf{BD} \tag{A.22}$$

und

$$(\mathbf{A} - \mathbf{B})(\mathbf{C} - \mathbf{D}) = \mathbf{AC} - \mathbf{AD} - \mathbf{BC} + \mathbf{BD}. \tag{A.23}$$

Ist $\mathbf{A}$ eine (n, p)-Matrix und $\mathbf{B}$ eine (p, q)-Matrix, dann gilt

$$(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'. \tag{A.24}$$

Der Beweis ist bei Zurmühl & Falk (1997), S.21-22 zu finden.

Definition 35. *Das äußere Produkt des n-dimensionalen Spaltenvektors* $\mathbf{a}$ *mit dem p-dimensionalen Spaltenvektor* $\mathbf{b}$ *ist definiert durch*

$$\mathbf{ab}' = \begin{pmatrix} a_1 b_1 & a_1 b_2 & \ldots & a_1 b_p \\ a_2 b_1 & a_2 b_2 & \ldots & a_2 b_p \\ \vdots & \vdots & \ddots & \vdots \\ a_n b_1 & a_n b_2 & \ldots & a_n b_p \end{pmatrix}. \tag{A.25}$$

Man nennt das äußere Produkt auch das *dyadische Produkt*. Ist $\mathbf{a}$ ein n-dimensionaler Spaltenvektor und $\mathbf{b}$ ein p-dimensionaler Spaltenvektor, so ist $\mathbf{ab}'$ eine (n,p)-Matrix und $\mathbf{ba}'$ eine (p,n)-Matrix. In der Regel ist $\mathbf{ab}'$ ungleich $\mathbf{ba}'$.

Beispiel 70. (fortgesetzt) Es gilt

$$\mathbf{a}_1\mathbf{a}_2' = \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix}$$

und

$$\mathbf{a}_2\mathbf{a}_1' = \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix}.$$

□

A.1.3 Die inverse Matrix

Definition 36. *Die (n,n)-Matrix* $\mathbf{A}$ *heißt invertierbar, wenn eine (n,n)-Matrix* $\mathbf{A}^{-1}$ *existiert, sodass gilt*

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}_n. \tag{A.26}$$

Man nennt $\mathbf{A}^{-1}$ *auch die inverse Matrix von* $\mathbf{A}$.

Beispiel 70. (fortgesetzt) Die inverse Matrix von

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

ist

$$\mathbf{A}^{-1} = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}.$$

Es gilt nämlich

$$\mathbf{A}^{-1}\mathbf{A} = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} = \mathbf{I}_2.$$

□

Für uns sind folgende Eigenschaften wichtig:

1. Die inverse Matrix $\mathbf{A}^{-1}$ der Matrix $\mathbf{A}$ ist eindeutig. Der Beweis ist bei Strang (1988), S.42 zu finden.
2. Sei $\mathbf{A}$ eine invertierbare (n, n)-Matrix. Dann gilt:
$$\left(\mathbf{A}^{-1}\right)^{-1} = \mathbf{A}. \tag{A.27}$$
3. Sei $\mathbf{A}$ eine invertierbare (n, n)-Matrix. Dann gilt:
$$\left(\mathbf{A}^{-1}\right)' = \left(\mathbf{A}'\right)^{-1}. \tag{A.28}$$
Der Beweis ist bei Zurmühl & Falk (1997), S.38 zu finden.
4. Sind die (n, n)-Matrizen $\mathbf{A}$ und $\mathbf{B}$ invertierbar, so gilt
$$\left(\mathbf{AB}\right)^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}. \tag{A.29}$$
Der Beweis ist bei Zurmühl & Falk (1997), S. 38 zu finden.

A.1.4 Orthogonale Matrizen

Definition 37. *Die n-dimensionalen Spaltenvektoren $\mathbf{a}$ und $\mathbf{b}$ heißen orthogonal, wenn gilt $\mathbf{a}'\mathbf{b} = 0$.*

Definition 38. *Eine (n, n)-Matrix $\mathbf{A}$ heißt orthogonal, wenn gilt*

$$\mathbf{AA}' = \mathbf{A}'\mathbf{A} = \mathbf{I}_n. \tag{A.30}$$

Beispiel 70. (fortgesetzt) Die Matrix $\mathbf{B}$ ist orthogonal. □

In einer orthogonalen Matrix haben alle Spaltenvektoren die Länge 1, und die Spaltenvektoren sind paarweise orthogonal.

Man kann einen n-dimensionalen Spaltenvektor als Punkt in einem kartesischen Koordinatensystem einzeichnen. Multipliziert man eine orthogonale (n, n)-Matrix $\mathbf{T}$ mit einem Spaltenvektor $\mathbf{x}$, so wird der Vektor bezüglich des Nullpunkts gedreht. Schauen wir uns dies in einem zweidimensionalen kartesischen Koordinatensystem an. Multipliziert man die orthogonale Matrix

$$\mathbf{T} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

mit dem Vektor

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

so wird der Vektor $\mathbf{x}$ um α Grad im Gegenzeigersinn gedreht. Eine Begründung hierfür ist bei Zurmühl & Falk (1997), S.6 zu finden.

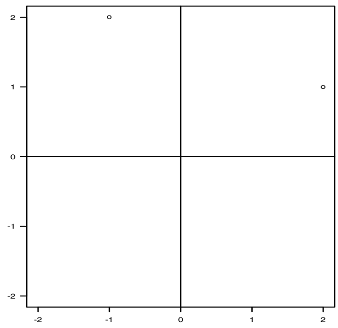

Abb. A.1. Drehung eines Punktes um 90 Grad im Gegenzeigersinn

Beispiel 70. (fortgesetzt) Es gilt $cos(0.5\pi) = 0$ und $sin(0.5\pi) = 1$. Somit erhalten wir für $\alpha = 0.5\pi$ die Matrix $\mathbf{B}$. Wir bilden

$$\mathbf{B}\mathbf{a}_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}.$$

Abbildung A.1 verdeutlicht den Zusammenhang.

□

A.1.5 Spur einer Matrix

Definition 39. *Sei* $\mathbf{A}$ *eine* (n, n)*-Matrix. Die Spur* $tr(\mathbf{A})$ *ist gleich der Summe der Hauptdiagonalelemente:*

$$tr(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}. \tag{A.31}$$

Dabei steht tr für trace.

Beispiel 70. (fortgesetzt) Es gilt

$$tr(\mathbf{A}) = 4.$$

□

Sind $\mathbf{A}$ und $\mathbf{B}$ (n, n)-Matrizen, so gilt

$$tr(\mathbf{A} + \mathbf{B}) = tr(\mathbf{A}) + tr(\mathbf{B}). \tag{A.32}$$

Ist $\mathbf{A}$ eine (n, p)-Matrix und $\mathbf{B}$ eine (p, n)-Matrix, so gilt

$$tr(\mathbf{A}\mathbf{B}) = tr(\mathbf{B}\mathbf{A}). \tag{A.33}$$

Der Beweis ist bei Zurmühl & Falk (1997), S.22 zu finden.

A.1.6 Determinante einer Matrix

Man kann einer (n,n)-Matrix $\mathbf{A}$ eine reelle Zahl zuordnen, die $\mathbf{A}$ charakterisiert. Dies ist die Determinante $|\mathbf{A}|$.

Definition 40. *Seien $\mathbf{A}$ eine (n,n)-Matrix und $\mathbf{A}_{ij}$ die $(n-1,n-1)$-Matrix, die man dadurch erhält, dass man die i-te Zeile und j-te Spalte von $\mathbf{A}$ streicht. Die Determinante $|\mathbf{A}|$ von $\mathbf{A}$ ist definiert durch*

$$|\mathbf{A}| = \begin{cases} a_{11} & \text{für } n = 1 \\ \sum_{j=1}^{n} (-1)^{i+j} a_{ij} |\mathbf{A}_{ij}| & \text{für } n \geq 2,\ i \text{ fest, aber beliebig, } 1 \leq i \leq n\,. \end{cases}$$

Sei

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

eine $(2,2)$-Matrix. Wir bestimmen $|\mathbf{A}|$ von $\mathbf{A}$ für $i = 1$:

$$|\mathbf{A}| = (-1)^{1+1} a_{11} |\mathbf{A}_{11}| + (-1)^{1+2} a_{12} |\mathbf{A}_{12}| = a_{11}a_{22} - a_{12}a_{21}\,.$$

Beispiel 70. (fortgesetzt) Es gilt

$$|\mathbf{A}| = 2 \cdot 2 - 1 \cdot 1 = 3.$$

□

Für uns sind drei Eigenschaften der Determinante wichtig:

1. Für eine (n,n)-Diagonalmatrix
$$\mathbf{D} = \begin{pmatrix} d_1 & 0 & \ldots & 0 \\ 0 & d_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & d_n \end{pmatrix}$$
gilt
$$|\mathbf{D}| = d_1 \cdot d_2 \cdot \ldots \cdot d_n. \tag{A.34}$$
Der Beweis ist bei Wetzel et al. (1981), S. 112 zu finden.
2. Sind $\mathbf{A}$ und $\mathbf{B}$ (n,n)-Matrizen, so gilt
$$|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|. \tag{A.35}$$
Der Beweis ist bei Jänich (2000), S. 148 zu finden.
3. Die Invertierbarkeit einer (n,n)-Matrix $\mathbf{A}$ kann über die Determinante von $\mathbf{A}$ charakterisiert werden. Die (n,n)-Matrix $\mathbf{A}$ ist genau dann invertierbar, wenn die Determinante von $\mathbf{A}$ ungleich Null ist. Der Beweis ist bei Jänich (2000), S.147-148 zu finden.

A.1.7 Lineare Gleichungssysteme

Wir betrachten ein in den p Unbekannten $x_1, \ldots, x_p$ lineares Gleichungssystem :

$$\begin{aligned} a_{11}\, x_1 + a_{12}\, x_2 + \ldots + a_{1p}\, x_p &= b_1 \\ a_{21}\, x_1 + a_{22}\, x_2 + \ldots + a_{2p}\, x_p &= b_2 \\ \vdots \quad &\vdots \\ a_{n1}\, x_1 + a_{n2}\, x_2 + \ldots + a_{np}\, x_p &= b_n \end{aligned} \tag{A.36}$$

Gilt $b_i = 0$ für $i = 1, \ldots, n$, so spricht man von einem *linear homogenen Gleichungssystem*, ansonsten von einem *linear inhomogenen Gleichungssystem.* Gesucht sind Werte von $x_1, \ldots, x_p$, die das Gleichungssystem erfüllen.

Mit

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1p} \\ a_{21} & a_{22} & \ldots & a_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{np} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

können wir (A.36) auch folgendermaßen schreiben:

$$\mathbf{A}\mathbf{x} = \mathbf{b}. \tag{A.37}$$

Um die Lösbarkeit von (A.37) diskutieren zu können, benötigen wir den Begriff der linearen Unabhängigkeit.

Definition 41. *Die n-dimensionalen Vektoren $\mathbf{a}_1, \ldots, \mathbf{a}_p$ heißen linear unabhängig, wenn aus*

$$x_1\mathbf{a}_1 + \ldots + x_p\mathbf{a}_p = \mathbf{0}$$

folgt

$$x_1 = x_2 = \ldots = 0.$$

Matrizen sind aus Spaltenvektoren beziehungsweise Zeilenvektoren aufgebaut. Die Maximalzahl linear unabhängiger Spaltenvektoren einer Matrix $\mathbf{A}$ nennt man den *Spaltenrang* von $\mathbf{A}$. Die Maximalzahl linear unabhängiger Zeilenvektoren einer Matrix $\mathbf{A}$ nennt man den *Zeilenrang* von $\mathbf{A}$. Die Maximalzahl linear unabhängiger Spaltenvektoren ist gleich der Maximalzahl linear unabhängiger Zeilenvektoren. Der Beweis ist bei Jänich (2000), S. 116-117 zu finden. Diese Zahl bezeichnet man als *Rang* $rg(\mathbf{A})$ von $\mathbf{A}$.

Mit Hilfe des Rangs kann man die Lösbarkeit von (A.37) diskutieren. Wir beginnen mit dem linear homogenen Gleichungssystem.

Das linear homogene Gleichungssystem

$$\mathbf{Ax} = \mathbf{0} \tag{A.38}$$

besitzt für $n \geq p$ genau eine Lösung, wenn der Rang von $\mathbf{A}$ gleich p ist. Dies folgt aus der Definition der linearen Unabhängigkeit. Ist der Rang von $\mathbf{A}$ kleiner als p, so besitzt das linear inhomogene Gleichungssystem mehr als eine Lösung. Die Struktur des Lösungsraums ist bei Wetzel et al. (1981), S.72-74 beschrieben. Gilt $p = n$, so kann man die Lösbarkeit von (A.38) auch über die Determinante der Matrix $\mathbf{A}$ charakterisieren. Die Determinante $|\mathbf{A}|$ von $\mathbf{A}$ ist genau dann ungleich 0, wenn der Rang von $\mathbf{A}$ gleich n ist (siehe dazu Wetzel et al. (1981), S. 114). Ist die Determinante der (n, n)-Matrix $\mathbf{A}$ also gleich 0, so ist der Rang von $\mathbf{A}$ kleiner als n und das linear homogene Gleichungssystem (A.38) hat mehr als eine Lösung.

Wir betrachten nun das linear inhomogene Gleichungssystem

$$\mathbf{Ax} = \mathbf{b}, \tag{A.39}$$

wobei $\mathbf{A}$ eine (n, n)-Matrix ist. Ist der Rang von $\mathbf{A}$ gleich n, so existiert die inverse Matrix $\mathbf{A}^{-1}$ von $\mathbf{A}$ (siehe dazu Wetzel et al. (1981), S. 97). Wir multiplizieren (A.39) von links mit $\mathbf{A}^{-1}$ und erhalten die Lösung

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}. \tag{A.40}$$

Da die inverse Matrix $\mathbf{A}^{-1}$ eindeutig ist, ist diese Lösung eindeutig.

Wir betrachten an einigen Stellen in diesem Buch eine (n, p)-Matrix $\mathbf{X}$, wobei gilt $rg(\mathbf{X}) = p$, und bilden die Matrix $\mathbf{X}'\mathbf{X}$. Schauen wir uns diese (p, p)-Matrix genauer an. Sie ist symmetrisch. Dies sieht man folgendermaßen:

$$(\mathbf{X}'\mathbf{X})' = \mathbf{X}'(\mathbf{X}')' = \mathbf{X}'\mathbf{X}. \tag{A.41}$$

Der Rang von $\mathbf{X}'\mathbf{X}$ ist gleich p. Da $\mathbf{X}$ den Spaltenrang p besitzt, gilt

$$\mathbf{Xy} = \mathbf{0} \Rightarrow \mathbf{y} = \mathbf{0}.$$

Wir haben zu zeigen

$$\mathbf{X}'\mathbf{Xy} = \mathbf{0} \Rightarrow \mathbf{y} = \mathbf{0}.$$

Es gilt

$$\begin{aligned}
\mathbf{X}'\mathbf{Xy} = \mathbf{0} &\Rightarrow \mathbf{y}'\mathbf{X}'\mathbf{Xy} = \mathbf{0} \\
&\Rightarrow (\mathbf{Xy})'\mathbf{Xy} = \mathbf{0} \\
&\Rightarrow \mathbf{Xy} = \mathbf{0} \\
&\Rightarrow \mathbf{y} = \mathbf{0}.
\end{aligned}$$

Da $\mathbf{X}'\mathbf{X}$ eine (p, p)-Matrix mit Rang p ist, existiert $(\mathbf{X}'\mathbf{X})^{-1}$. Die Matrix $(\mathbf{X}'\mathbf{X})^{-1}$ ist symmetrisch. Mit (A.28) und (A.41) gilt nämlich

$$\left((\mathbf{X}'\mathbf{X})^{-1}\right)' = \left((\mathbf{X}'\mathbf{X})'\right)^{-1} = (\mathbf{X}'\mathbf{X})^{-1}. \tag{A.42}$$

A.1.8 Eigenwerte und Eigenvektoren

Bei einer Reihe multivariater Verfahren benötigt man die *Eigenwerte* und *Eigenvektoren* einer symmetrischen (n,n)-Matrix $\mathbf{A}$.

Definition 42. *Sei* $\mathbf{A}$ *eine* (n,n)*-Matrix. Erfüllen ein Skalar* λ *und ein* n*-dimensionaler Spaltenvektor* $\mathbf{u}$ *mit* $\mathbf{u} \neq \mathbf{0}$ *das Gleichungssystem*

$$\mathbf{A}\mathbf{u} = \lambda\mathbf{u}, \tag{A.43}$$

so heißt λ *Eigenwert von* $\mathbf{A}$ *und* $\mathbf{u}$ *zugehöriger Eigenvektor von* $\mathbf{A}$*.*

Erfüllt ein Vektor $\mathbf{u}$ die Gleichung (A.43), so erfüllt auch jedes Vielfache von $\mathbf{u}$ die Gleichung (A.43).

Um eine Lösung von (A.43) zu erhalten, formen wir (A.43) um zu

$$(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{u} = \mathbf{0}. \tag{A.44}$$

Für festes λ ist (A.44) ein linear homogenes Gleichungssystem. Dieses besitzt genau dann Lösungen, die ungleich dem Nullvektor sind, wenn die Spalten von $\mathbf{A} - \lambda\mathbf{I}_n$ linear abhängig sind. Dies ist genau dann der Fall, wenn gilt

$$|\mathbf{A} - \lambda\mathbf{I}_n| = 0. \tag{A.45}$$

Gleichung (A.45) ist ein Polynom n-ten Grades in λ. Dieses besitzt genau n Nullstellen. Die Nullstellen $\lambda_1, \ldots, \lambda_n$ des Polynoms sind also die Eigenwerte der Matrix $\mathbf{A}$. Diese Eigenwerte müssen nicht notwendigerweise verschieden sein. Im Folgenden seien die Eigenwerte der Größe nach durchnummeriert, wobei der erste Eigenwert der größte ist.

Beispiel 71. Wir bestimmen die Eigenwerte der Matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Es gilt

$$\mathbf{A} - \lambda\mathbf{I}_2 = \begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix}.$$

Somit gilt

$$|\mathbf{A} - \lambda\mathbf{I}_2| = (2-\lambda)^2 - 1\,.$$

Ein Eigenwert λ erfüllt also die Gleichung

$$(2-\lambda)^2 - 1 = 0. \tag{A.46}$$

Die Nullstellen von (A.46) und somit die Eigenwerte von $\mathbf{A}$ sind $\lambda_1 = 3$ und $\lambda_2 = 1$. □

Die Eigenvektoren zum Eigenwert λ_i, $i = 1, \ldots, n$ erhalten wir dadurch, dass wir λ_i in Gleichung (A.44) für λ einsetzen und die Lösungsmenge des dadurch entstandenen linear homogenen Gleichungssystems bestimmen.

Beispiel 71. (fortgesetzt) Beginnen wir mit $\lambda_1 = 3$. Der zu $\lambda_1 = 3$ gehörende Eigenvektor

$$\mathbf{u}_1 = \begin{pmatrix} u_{11} \\ u_{21} \end{pmatrix}$$

erfüllt also das Gleichungssystem

$$(\mathbf{A} - 3\mathbf{I}_2)\mathbf{u} = \mathbf{0}.$$

Wegen

$$\mathbf{A} - 3\mathbf{I}_2 = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$$

ergibt sich

$$\begin{aligned} -u_{11} + u_{21} &= 0, \\ u_{11} - u_{21} &= 0\,. \end{aligned}$$

Für die Komponenten des Eigenvektors $\mathbf{u}_1$ zum Eigenwert $\lambda_1 = 3$ muss also gelten $u_{11} = u_{21}$. Der Vektor

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

und alle Vielfachen dieses Vektors sind Eigenvektoren zum Eigenwert $\lambda_1 = 3$. Analoge Berechnungen zum Eigenwert $\lambda_2 = 1$ ergeben, dass für die Komponenten u_{12} und u_{22} des zu $\lambda_2 = 1$ gehörenden Eigenvektors $\mathbf{u}_2$ die Beziehung $u_{12} = -u_{22}$ gelten muss. Der Vektor

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

und alle Vielfachen dieses Vektors sind Eigenvektoren zum Eigenwert $\lambda_2 = 1$.
□

In der multivariaten Analyse sollen die Eigenvektoren normiert sein.

Beispiel 71. (fortgesetzt) Die normierten Eigenvektoren von $\mathbf{A}$ sind

$$\mathbf{u}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

und

$$\mathbf{u}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

□

Folgende Eigenschaften von Eigenwerten und Eigenvektoren sind wichtig:

1. Ist λ ein Eigenwert von $\mathbf{A}$ und k eine reelle Zahl, so ist $k\lambda$ ein Eigenwert von $k\mathbf{A}$. Dies ist offensichtlich.
2. Der Rang der symmetrischen (n, n)-Matrix $\mathbf{A}$ ist gleich der Anzahl der von 0 verschiedenen Eigenwerte von $\mathbf{A}$. Der Beweis ist bei Basilevsky (1983), S. 201 zu finden.
3. Die Eigenwerte einer symmetrischen Matrix sind alle reell. Der Beweis ist bei Basilevsky (1983), S. 199 zu finden.
4. Die Eigenvektoren einer symmetrischen Matrix, die zu unterschiedlichen Eigenwerten gehören, sind orthogonal. Der Beweis ist bei Basilevsky (1983), S. 200 zu finden.
5. $$tr(\mathbf{A}) = \sum_{i=1}^{n} \lambda_i \tag{A.47}$$
6. $$|\mathbf{A}| = \lambda_1 \cdot \lambda_2 \cdot \ldots \cdot \lambda_n \tag{A.48}$$

Wir beweisen die Eigenschaften 5. am Ende des nächsten Abschnitts für eine symmetrische Matrix $\mathbf{A}$. Der Beweis von 6. ist bei Basilevsky (1983), S.200 zu finden.

A.1.9 Die Spektralzerlegung einer symmetrischen Matrix

Wir gehen zunächst davon aus, dass die Eigenwerte λ_i, $i = 1, \ldots, n$ der symmetrischen (n, n)-Matrix $\mathbf{A}$ alle unterschiedlich sind. Sei $\mathbf{u}_i$ der normierte Eigenvektor zum Eigenwert λ_i, $i = 1, \ldots, n$. Die Eigenwerte λ_i und Eigenvektoren $\mathbf{u}_i$ erfüllen für $i = 1, \ldots, n$ die Gleichungen

$$\mathbf{A}\mathbf{u}_i = \lambda_i \mathbf{u}_i . \tag{A.49}$$

Mit $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ und

$$\boldsymbol{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}$$

können wir diese Gleichungen auch folgendermaßen kompakt schreiben:

$$\mathbf{A}\mathbf{U} = \mathbf{U}\boldsymbol{\Lambda} . \tag{A.50}$$

Da bei einer symmetrischen Matrix Eigenvektoren zu unterschiedlichen Eigenwerten orthogonal sind, ist die Matrix $\mathbf{U}$ eine Orthogonalmatrix. Es gilt also

$$\mathbf{U}\mathbf{U}' = \mathbf{U}'\mathbf{U} = \mathbf{I}_n.$$

Multiplizieren wir (A.50) von rechts mit $\mathbf{U}'$, so erhalten wir:

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}'. \tag{A.51}$$

Gleichung (A.51) nennt man die *Spektralzerlegung* der symmetrischen (n, n)-Matrix $\mathbf{A}$. Diese Zerlegung ist auch möglich, wenn nicht alle Eigenwerte unterschiedlich sind. Ein Beweis des allgemeinen Falles ist bei Jänich (2000), S.218 ff. zu finden.

Beispiel 71. (fortgesetzt) Es gilt

$$\boldsymbol{\Lambda} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$$

und

$$\mathbf{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Somit gilt

$$\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}' = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

$$= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 3 & 3 \\ 1 & -1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

□

Wir können die Gleichung (A.51) in Abhängigkeit von den Eigenwerten $\lambda_1, \ldots, \lambda_n$ und den Eigenvektoren $\mathbf{u}_1, \ldots, \mathbf{u}_n$ auch folgendermaßen schreiben:

$$\mathbf{A} = \sum_{i=1}^{n} \lambda_i \, \mathbf{u}_i \mathbf{u}_i'. \tag{A.52}$$

Dies sieht man folgendermaßen:

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}' = (\mathbf{u}_1, \, \mathbf{u}_2, \, \ldots, \, \mathbf{u}_n) \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix} \begin{pmatrix} \mathbf{u}_1' \\ \mathbf{u}_2' \\ \vdots \\ \mathbf{u}_n' \end{pmatrix}$$

$$= (\mathbf{u}_1, \, \mathbf{u}_2, \, \ldots, \, \mathbf{u}_n) \begin{pmatrix} \lambda_1 \mathbf{u}_1' \\ \lambda_2 \mathbf{u}_2' \\ \vdots \\ \lambda_n \mathbf{u}_n' \end{pmatrix} = \sum_{i=1}^{n} \lambda_i \, \mathbf{u}_i \mathbf{u}_i'.$$

Wir können die Matrix $\mathbf{A}$ also als Summe von Matrizen darstellen. Sind die ersten beiden Eigenwerte groß im Verhältnis zu den restlichen Eigenwerten, so reichen vielleicht schon die ersten beiden Summmanden und somit die ersten beiden Eigenvektoren zur Approximation von $\mathbf{A}$.

Wir wollen nun noch (A.47) beweisen für den Fall, dass die Matrix $\mathbf{A}$ symmetrisch ist. Es gilt

$$
\begin{aligned}
tr(\mathbf{A}) &= tr(\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}') && \text{(A.53)}\\
&= tr(\mathbf{U}'\mathbf{U}\boldsymbol{\Lambda}) && \text{(A.54)}\\
&= tr(\boldsymbol{\Lambda}) = \sum_{i=1}^{n} \lambda_i .
\end{aligned}
$$

Beim Übergang von (A.53) zu (A.54) wird (A.33) benutzt. Beim Übergang von (A.54) zu (A.55) wird (A.30) benutzt.

A.1.10 Die Singulärwertzerlegung

Wir haben im letzten Abschnitt gesehen, dass man eine symmetrische Matrix so in das Produkt von drei Matrizen zerlegen kann, dass man die Matrix durch eine Summe von einfachen Matrizen schreiben kann. Eine ähnlich nützliche Zerlegung ist für jede (n,p)-Matrix $\mathbf{A}$ mit $rg(\mathbf{A}) = r$ möglich. Zu jeder (n,p)-Matrix $\mathbf{A}$ mit $rg(\mathbf{A}) = r$ existiert eine orthogonale (n,n)-Matrix $\mathbf{U}$, eine orthogonale (p,p)-Matrix $\mathbf{V}$ und eine (n,p)-Matrix $\mathbf{D}$ mit

$$
\mathbf{D} = \begin{pmatrix}
d_1 & 0 & \ldots & 0 & 0 & \ldots & 0\\
0 & d_2 & \ldots & 0 & 0 & \ldots & 0\\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\
0 & 0 & \ldots & d_r & 0 & \ldots & 0\\
0 & 0 & \ldots & 0 & 0 & \ldots & 0\\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\
0 & 0 & \ldots & 0 & 0 & \ldots & 0
\end{pmatrix},
$$

sodass gilt

$$
\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}'. \tag{A.55}
$$

Dabei sind die Spalten der Matrix $\mathbf{U}$ die Eigenvektoren der Matrix $\mathbf{A}\mathbf{A}'$ und die Spalten der Matrix $\mathbf{V}$ die Eigenvektoren der Matrix $\mathbf{A}'\mathbf{A}$. $d_1, \ldots, d_r$ sind die positiven Quadratwurzeln aus den positiven Eigenwerten der Matrix $\mathbf{A}\mathbf{A}'$ beziehungsweise $\mathbf{A}'\mathbf{A}$. Man nennt (A.55) auch die *Singulärwertzerlegung* der Matrix $\mathbf{A}$. Eine ausführliche Darstellung der Singulärwertzerlegung unter Berücksichtigung von Anwendungen ist bei Watkins (1991), S. 390-430 zu finden.

A.1.11 Quadratische Formen

Wir benötigen in diesem Buch an einigen Stellen quadratische Formen.

Definition 43. *Sei* $\mathbf{x}$ *ein* n*-dimensionaler Vektor und* $\mathbf{A}$ *eine symmetrische* (n,n)*-Matrix. Dann heißt*

$$Q = \mathbf{x}'\mathbf{A}\mathbf{x} \tag{A.56}$$

quadratische Form in den Variablen $x_1, \ldots, x_n$*.*

Beispiel 72. Sei

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Es gilt

$$\begin{aligned}\mathbf{x}'\mathbf{A}\mathbf{x} &= (x_1 \; x_2) \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = (x_1 \; x_2) \begin{pmatrix} 2x_1 + x_2 \\ x_1 + 2x_2 \end{pmatrix} \\ &= 2x_1^2 + 2x_1x_2 + 2x_2^2 \,.\end{aligned}$$

□

In einer Reihe von Situationen benötigen wir die Definitheit einer Matrix.

Definition 44. *Die Matrix* $\mathbf{A}$ *heißt positiv definit, wenn* $\mathbf{x}'\mathbf{A}\mathbf{x} > 0$ *für alle* $\mathbf{x} \neq \mathbf{0}$ *gilt.*
Die Matrix $\mathbf{A}$ *heißt positiv semidefinit, wenn* $\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0$ *für alle* $\mathbf{x} \neq \mathbf{0}$ *gilt, wobei* $\mathbf{x}'\mathbf{A}\mathbf{x} = 0$ *für mindestens ein* $\mathbf{x} \neq \mathbf{0}$ *gilt.*
Die Matrix $\mathbf{A}$ *heißt negativ definit, wenn* $\mathbf{x}'\mathbf{A}\mathbf{x} < 0$ *für alle* $\mathbf{x} \neq \mathbf{0}$ *gilt.*
Die Matrix $\mathbf{A}$ *heißt negativ semidefinit, wenn* $\mathbf{x}'\mathbf{A}\mathbf{x} \leq 0$ *für alle* $\mathbf{x} \neq \mathbf{0}$ *gilt, wobei* $\mathbf{x}'\mathbf{A}\mathbf{x} = 0$ *für mindestens ein* $\mathbf{x} \neq \mathbf{0}$ *gilt.*

Beispiel 72. (fortgesetzt) Es gilt

$$2x_1^2 + 2x_1x_2 + 2x_2^2 = x_1^2 + 2x_1x_2 + x_2^2 + x_1^2 + x_2^2 = (x_1 + x_2)^2 + x_1^2 + x_2^2.$$

Dieser Ausdruck ist nichtnegativ. Er wird nur Null, wenn gilt $x_1 = x_2 = 0$. Somit ist $\mathbf{A}$ positiv definit. □

Die Definitheit einer Matrix lässt sich auch über die Eigenwerte charakterisieren. Es gilt

$$\mathbf{A} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}',$$

wobei die Spalten von $\mathbf{U}$ die normierten Eigenvektoren enthalten und die Hauptdiagonalelemente der Diagonalmatrix $\boldsymbol{\Lambda}$ die Eigenwerte $\lambda_1, \ldots, \lambda_n$ von $\mathbf{A}$ sind. Somit gilt

$$\mathbf{x}'\mathbf{A}\mathbf{x} = \mathbf{x}'\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}'\mathbf{x} = (\mathbf{U}'\mathbf{x})'\boldsymbol{\Lambda}\mathbf{U}'\mathbf{x} = \mathbf{z}'\boldsymbol{\Lambda}\mathbf{z} = \sum_{i=1}^{n} \lambda_i z_i^2$$

mit

$$\mathbf{z} = \mathbf{U}'\mathbf{x}.$$

Somit ist die symmetrische (n, n)-Matrix

– positiv definit, wenn alle Eigenwerte größer als Null sind,
– positiv semidefinit, wenn alle Eigenwerte größer gleich Null sind und mindestens ein Eigenwert gleich Null ist,
– negativ definit, wenn alle Eigenwerte kleiner als Null sind,
– negativ semidefinit, wenn alle Eigenwerte kleiner gleich Null sind und mindestens ein Eigenwert gleich Null ist.

Beispiel 72. (fortgesetzt) Wir haben in Beispiel 71 die Eigenwerte von $\mathbf{A}$ bestimmt. Da diese positiv sind, ist $\mathbf{A}$ positiv definit. □

A.2 Extremwerte

Wir müssen an einigen Stellen in diesem Buch Extremwerte von Funktionen mehrerer Veränderlicher bestimmen. Dabei betrachten wir Funktionen

$$f : D \subset \mathbb{R}^n \to \mathbb{R}\,.$$

Beispiel 73. Wir betrachten die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit

$$\mathbf{x} \mapsto f(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$$

mit

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Wir können $f(\mathbf{x})$ auch explizit in Abhängigkeit von den Komponenten x_1 und x_2 von $\mathbf{x}$ schreiben. Es gilt

$$f(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = 2x_1^2 + 2x_1x_2 + 2x_2^2.$$

□

Die ϵ-Umgebung $U_\epsilon(\mathbf{x}_0)$ eines Punktes $\mathbf{x}_0 \in \mathbb{R}^n$ ist definiert durch

$$U_\epsilon(\mathbf{x}_0) = \{\mathbf{x}\,|\,\|\mathbf{x} - \mathbf{x}_0\| < \epsilon\}\ .$$

Definition 45. *Die Funktion $f : D \subset \mathbb{R}^n \to \mathbb{R}$ besitzt in $\mathbf{x}_0$ ein lokales Minimum, wenn eine ϵ-Umgebung $U_\epsilon(\mathbf{x}_0)$ von $\mathbf{x}_0$ existiert, sodass für alle $\mathbf{x} \in U_\epsilon(\mathbf{x}_0)$ mit $\mathbf{x} \neq \mathbf{x}_0$ gilt*

$$f(\mathbf{x}_0) < f(\mathbf{x}).$$

Die Funktion $f : D \subset \mathbb{R}^n \to \mathbb{R}$ besitzt in $\mathbf{x}_0$ ein lokales Maximum, wenn eine ϵ-Umgebung $U_\epsilon(\mathbf{x}_0)$ von $\mathbf{x}_0$ existiert, sodass für alle $\mathbf{x} \in U_\epsilon(\mathbf{x}_0)$ mit $\mathbf{x} \neq \mathbf{x}_0$ gilt

$$f(\mathbf{x}_0) > f(\mathbf{x}).$$

A.2.1 Der Gradient und die Hesse-Matrix

Eine notwendige Bedingung für einen Extremwert in x_0 einer in x_0 differenzierbaren Funktion $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$ ist $f'(x_0) = 0$. Dabei ist $f'(x_0)$ die erste Ableitung von f an der Stelle x_0. Diese ist folgendermaßen definiert:

$$f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Dieses Konzept kann auf eine Funktion $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ übertragen werden. Die partielle Ableitung von f nach x_i an der Stelle $\mathbf{x}_0$ ist definiert durch

$$\frac{\partial f(\mathbf{x}_0)}{\partial x_i} = \lim_{h \rightarrow 0} \frac{f(\mathbf{x}_0 + h\mathbf{e}_i) - f(\mathbf{x}_0)}{h}.$$

Dabei ist $\mathbf{e}_i$ der i-te Einheitsvektor. Wir sagen, dass die Funktion $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ in $\mathbf{x}_0$ nach der i-ten Komponente partiell differenzierbar ist, wenn $\frac{\partial f(\mathbf{x}_0)}{\partial x_i}$ existiert.

Beispiel 73. (fortgesetzt) Es gilt

$$\frac{\partial f(\mathbf{x})}{\partial x_1} = 4x_1 + 2x_2$$

und

$$\frac{\partial f(\mathbf{x})}{\partial x_2} = 2\,x_1 + 4\,x_2\,.$$

□

Ist die Funktion $f : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ nach jeder Komponente von $\mathbf{x}$ partiell differenzierbar, dann heißt der Vektor

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \begin{pmatrix} \frac{\partial f(\mathbf{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\mathbf{x})}{\partial x_n} \end{pmatrix}$$

Gradient der Funktion.

Beispiel 73. (fortgesetzt) Es gilt

$$\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} = \begin{pmatrix} 4\,x_1 + 2\,x_2 \\ 2\,x_1 + 4\,x_2 \end{pmatrix}.$$

□

Schauen wir uns den Gradienten spezieller Funktionen an. Sei

$$\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

Dann gilt

$$\frac{\partial\, \mathbf{a}'\mathbf{x}}{\partial \mathbf{x}} = \mathbf{a}\,. \tag{A.57}$$

Es gilt nämlich

$$\frac{\partial\, \mathbf{a}'\mathbf{x}}{\partial x_i} = \frac{\partial\, a_1 x_1 + \ldots + a_n x_n}{\partial x_i} = a_i\,.$$

Ist $\mathbf{A}$ eine symmetrische Matrix, so gilt

$$\frac{\partial\, \mathbf{x}'\mathbf{A}\mathbf{x}}{\partial \mathbf{x}} = 2\,\mathbf{A}\mathbf{x}\,. \tag{A.58}$$

Der Beweis ist bei Büning et al. (2000), S. 144-145 zu finden.

Man kann bei einer Funktion $f : D \subset I\!R^n \to I\!R$ auch partielle Ableitungen höherer Ordnung betrachten. Existieren die partiellen Ableitungen zweiter Ordnung und sind sie stetig, so heißt die Matrix der partiellen Ableitungen zweiter Ordnung *Hesse-Matrix*:

$$\mathbf{H}(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{x})}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(\mathbf{x})}{\partial x_n^2} \end{pmatrix}. \tag{A.59}$$

Beispiel 73. (fortgesetzt) Es gilt

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_1^2} = \frac{\partial\, 4x_1 + 2x_2}{\partial x_1} = 4,$$

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_1\, \partial x_2} = \frac{\partial\, 4x_1 + 2x_2}{\partial x_2} = 2,$$

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_2\, \partial x_1} = \frac{\partial\, 2x_1 + 4x_2}{\partial x_1} = 2,$$

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_2^2} = \frac{\partial\, 2x_1 + 4x_2}{\partial x_2} = 4\,.$$

Es gilt also

$$\mathbf{H}(\mathbf{x}) = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}.$$

□

A.2.2 Extremwerte ohne Nebenbedingungen

Wir suchen in diesem Buch Extremwerte von Funktionen $f : D \subset I\!R^n \to I\!R$, deren erste und zweite partielle Ableitungen existieren und stetig sind. Eine notwendige Bedingung dafür, dass die Funktion $f : D \subset I\!R^n \to I\!R$ einen Extremwert an der Stelle $\mathbf{x}_0$ hat, ist

$$\frac{\partial f(\mathbf{x}_0)}{\partial \mathbf{x}} = \mathbf{0}.$$

Der Beweis ist bei Khuri (1993), S. 283 zu finden.

Beispiel 73. (fortgesetzt) Notwendige Bedingungen für Extremwerte von

$$f(x_1, x_2) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = 2x_1^2 + 2x_1x_2 + 2x_2^2$$

sind

$$\begin{aligned} 4\,x_1 + 2\,x_2 &= 0, \\ 2\,x_1 + 4\,x_2 &= 0. \end{aligned}$$

Dieses linear homogene Gleichungssystem hat die Lösung $x_1 = x_2 = 0$. □

Um zu überprüfen, ob in $\mathbf{x}_0$ ein Extremwert vorliegt, bestimmt man $\mathbf{H}(\mathbf{x}_0)$. Ist $\mathbf{H}(\mathbf{x}_0)$ negativ definit, so liegt ein lokales Maximum vor. Ist $\mathbf{H}(\mathbf{x}_0)$ positiv definit, so liegt ein lokales Minimum vor. Der Beweis ist bei Khuri (1993), S. 283-284 zu finden.

Beispiel 73. (fortgesetzt) Es gilt

$$\mathbf{H}(x_1, x_2) = \begin{pmatrix} 4 & 2 \\ 2 & 4 \end{pmatrix}.$$

Es gilt $\mathbf{H}(\mathbf{x}) = 2\mathbf{A}$ mit

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Wir haben die Eigenwerte von $\mathbf{A}$ bereits im Beispiel 71 bestimmt. Da die Eigenwerte von $2\mathbf{A}$ doppelt so groß wie die Eigenwerte von $\mathbf{A}$ sind, sind auch beide Eigenwerte von $2\mathbf{A}$ positiv. Also liegt ein lokales Minimum vor. □

A.2.3 Extremwerte unter Nebenbedingungen

Bei der Optimierung von $f(\mathbf{x})$ müssen oft Nebenbedingungen der Form $g(\mathbf{x}) = 0$ berücksichtigt werden. Zur Bestimmung der Extremwerte stellen wir die Lagrange-Funktion

$$L(\mathbf{x}, \lambda) = f(\mathbf{x}) - \lambda\, g(\mathbf{x})$$

auf.

Eine notwendige Bedingung eines Extremwerts von $f(\mathbf{x})$ in $\mathbf{x}_0$ unter der Nebenbedingung $g(\mathbf{x}) = 0$ ist

$$\frac{\partial L(\mathbf{x}_0, \lambda_0)}{\partial \mathbf{x}} = \mathbf{0}$$

und

$$\frac{\partial L(\mathbf{x}_0, \lambda_0)}{\partial \lambda} = \mathbf{0}.$$

Der Beweis ist bei Khuri (1993), S. 287-290 zu finden.

Beispiel 74. Wir suchen den Extremwert von

$$f(x_1, x_2) = 2x_1^2 + 2x_1x_2 + 2x_2^2$$

unter der Nebenbedingung

$$x_1^2 + x_2^2 = 1.$$

Wir stellen die Lagrange-Funktion auf:

$$L(x_1, x_2, \lambda) = 2x_1^2 + 2x_1x_2 + 2x_2^2 - \lambda(x_1^2 + x_2^2 - 1).$$

Die partiellen Ableitungen lauten:

$$\frac{\partial L(x_1, x_2, \lambda)}{\partial x_1} = 4x_1 + 2x_2 - 2\lambda x_1,$$

$$\frac{\partial L(x_1, x_2, \lambda)}{\partial x_2} = 2x_1 + 4x_2 - 2\lambda x_2$$

und

$$\frac{\partial L(x_1, x_2, \lambda)}{\partial \lambda} = -x_1^2 - x_2^2 + 1.$$

Ein Extremwert

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

muss also die folgenden Gleichungen erfüllen:

$$4x_1 + 2x_2 - 2\lambda x_1 = 0, \tag{A.60}$$

$$2x_1 + 4x_2 - 2\lambda x_2 = 0 \tag{A.61}$$

und

$$x_1^2 + x_2^2 = 1. \tag{A.62}$$

Wir können die Gleichungen (A.60) und (A.61) auch schreiben als:

$$2x_1 + x_2 = \lambda x_1, \tag{A.63}$$

$$x_1 + 2x_2 = \lambda x_2. \tag{A.64}$$

Mit

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

lauten diese Gleichungen in Matrixform

$$\mathbf{Ax} = \lambda \mathbf{x}.$$

Dies ist aber ein Eigenwertproblem. Ein Eigenvektor von **A** erfüllt also die notwendigen Bedingungen für einen Extremwert. Da dieser auch die Nebenbedingung erfüllen muss, müssen wir ihn normieren. Die notwendigen Bedingungen erfüllen also die Punkte

$$\mathbf{x}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

und

$$\mathbf{x}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

□

Auf die hinreichenden Bedingungen für einen Extremwert unter Nebenbedingungen wollen wir hier nicht eingehen. Sie sind bei Wetzel et al. (1981) zu finden.

A.3 Matrizenrechnung in S-PLUS

In S-PLUS sind alle beschriebenen Konzepte der Matrizenrechnung implementiert. Wir schauen uns die Beispiele aus Kapitel A.1 in S-PLUS an und beginnen mit Vektoren. Wir geben zunächst die Vektoren

$$\mathbf{a}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

aus Beispiel 70 ein. Dazu verwenden wir die Funktion c:

```
> a1<-c(2,1)
> a2<-c(1,2)
```

Schauen wir uns zunächst Verknüpfungen von Vektoren an. In S-PLUS kann man die Vektoren a1 und a2 mit einem Operator verknüpfen, wenn sie die gleiche Länge besitzen. Das Ergebnis ist ein Vektor a3, der die gleiche Länge besitzt wie die Vektoren a1 und a2. Jedes Element des Vektors a3 erhält man dadurch, dass man die entsprechenden Elemente der Vektoren a1 und a2 mit dem Operator verknüpft. Die Operatoren sind dabei nicht wie in der Matrizenrechnung beschränkt auf + und -. Man kann also auch den Multiplikationsoperator * verwenden. Dieser liefert dann aber nicht das aus der Matrizenrechnung bekannte innere Produkt der Vektoren. Zuerst überprüfen wir aber, ob die beiden Vektoren die gleiche Länge besitzen:

```
> length(a1)
[1] 2
> length(a2)
[1] 2
```

Zum Vergleich der Längen der beiden Vektoren verwenden wir den Vergleichsoperator ==. Beim Vergleich von zwei Skalaren liefert dieser den Wert T, wenn beide identisch sind, ansonsten den Wert F:

```
> 3==(2+1)
[1] T
> 3==(3+1)
[1] F
```

Wir geben also ein

```
> length(a1)==length(a2)
[1] T
```

Schauen wir uns einige Beispiele für Verknüpfungen von Vektoren an:

```
> a1+a2
[1] 3 3
> a1-a2
[1]  1 -1
```

```
> a1*a2
[1] 2 2
> a1==a2
[1] F F
```

Beim letzten Beispiel war das Ergebnis ein logischer Vektor. Um zu überprüfen, ob irgendeine Komponente eines Vektors mit der entsprechenden Komponente eines anderen Vektors übereinstimmt, verwenden wir die Funktion `any`:

```
> any(a1==a2)
[1] F
```

Mit der Funktion `all` können wir überprüfen, ob alle Komponenten übereinstimmen:

```
> all(a1==a2)
[1] F
```

Das innere Produkt der gleich langen Vektoren `a1` und `a2` liefert der Operator `%*%`:

```
> a1%*%a2
     [,1]
[1,]    4
```

Das Ergebnis ist eine Matrix und kein Vektor. Man kann einen Vektor mit einem Skalar verknüpfen. Dabei wird jedes Element jeder Komponente des Vektors mit dem Skalar verknüpft:

```
> 1+a1
[1] 3 2
> 2*a1
[1] 4 2
```

Das äußere Produkt der Vektoren `a1` und `a2` gewinnt man mit der Funktion `outer`:

```
> outer(a1,a2)
      [,1] [,2]
 [1,]    2    4
 [2,]    1    2
> outer(a2,a1)
      [,1] [,2]
 [1,]    2    1
 [2,]    4    2
```

Schauen wir uns Matrizen an. Wir geben die Matrizen

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$$

aus Beispiel 70 mit der Funktion `matrix` ein. Dabei beachten wir, dass Matrizen spaltenweise aufgefüllt werden:

```
> A<-matrix(c(2,1,1,2),2,2)
> B<-matrix(c(0,1,-1,0),2,2)
> C<-matrix(c(1,1,1,2,1,2),3,2)
> D<-matrix(c(3,0,0,1),2,2)
```

Man kann zwei Matrizen mit einem Operator verknüpfen, wenn sie die gleiche Dimension besitzen. Die Dimension einer Matrix erhält man mit der Funktion `dim`:

```
> dim(C)
[1] 3 2
```

Das Ergebnis der Verknüpfung der Matrizen `A` und `B` ist eine Matrix `C` mit der Dimension der Objekte, die durch den Operator verknüpft werden. Jedes Element der Matrix `C` erhält man dadurch, dass man die entsprechenden Elemente der Matrizen `A` und `B` mit dem Operator verknüpft:

```
> A+B
      [,1] [,2]
 [1,]    2    0
 [2,]    2    2
> A-B
      [,1] [,2]
 [1,]    2    2
 [2,]    0    2
> A*B
      [,1] [,2]
 [1,]    0   -1
 [2,]    1    0
```

Um das Produkt der Matrizen $\mathbf{A}$ und $\mathbf{B}$ bilden zu können, muss die Anzahl der Spalten von $\mathbf{A}$ gleich der Anzahl der Zeilen von $\mathbf{B}$ sein. Ist dies der Fall, so liefert der Operator `%*%` das Produkt:

```
> A%*%B
      [,1] [,2]
 [1,]    1   -2
 [2,]    2   -1
```

Die Transponierte $\mathbf{A}'$ einer Matrix $\mathbf{A}$ erhält man mit der Funktion `t`:

```
> t(C)
      [,1] [,2] [,3]
 [1,]    1    1    1
 [2,]    2    1    2
```

Um zu überprüfen, ob eine quadratische Matrix symmetrisch ist, geben wir ein

```
> all(A==t(A))
[1] T
```

Mit der Funktion `diag` kann man eine Diagonalmatrix erzeugen. Der Aufruf

```
> diag(c(3,1))
```

liefert als Ergebnis

```
     [,1] [,2]
[1,]    3    0
[2,]    0    1
```

Die Einheitsmatrix $\mathbf{I}_3$ erhält man durch

```
> diag(3)
     [,1] [,2] [,3]
[1,]    1    0    0
[2,]    0    1    0
[3,]    0    0    1
```

Außerdem kann man mit der Funktion `diag` die Hauptdiagonalelemente einer Matrix extrahieren.

```
> diag(A)
[1] 2 2
```

Die inverse Matrix $\mathbf{A}^{-1}$ erhält man mit der Funktion `solve`:

```
> solve(A)
           [,1]       [,2]
[1,]  0.6666667 -0.3333333
[2,] -0.3333333  0.6666667
```

Der Aufruf

```
> solve(A)%*%A
```

liefert im Rahmen der Rechengenauigkeit die Einheitsmatrix. Mit der Funktion `solve` kann man auch lineare Gleichungssysteme lösen. Hierauf wollen wir aber nicht eingehen.

Die Spur einer quadratischen Matrix erhält man durch

```
> sum(diag(A))
[1] 4
```

Die Funktion `eigen` liefert die Eigenwerte und Eigenvektoren einer quadratischen Matrix. Das Ergebnis von `eigen` ist eine Liste. Die erste Komponente der Liste enthält die Eigenwerte und die zweite die Eigenvektoren.

```
> e<-eigen(A)
> e[[1]]
[1] 3 1
> e[[2]]
          [,1]       [,2]
 [1,] 0.7071068  0.7071068
 [2,] 0.7071068 -0.7071068
```

Wir bilden die orthogonale Matrix `U` mit den Eigenvektoren von `A` in den Spalten:

```
> U<-e[[2]]
```

und eine Diagonalmatrix `L` mit den Eigenwerten von `A`:

```
> L<-diag(e[[1]])
```

Der Aufruf

```
> U%*%L%*%t(U)
```

liefert die Matrix `A`

```
      [,1] [,2]
 [1,]    2    1
 [2,]    1    2
```

Der Aufruf

```
> e<-svd(C)
```

liefert die Singulärwertzerlegung `UDV'`der Matrix `C`. Das Ergebnis ist eine Liste. Die erste Komponente enthält die positiven Elemente der Matrix `D`, die zweite Komponente die Matrix `V` und die dritte Komponente die Matrix `U`. Wir bilden die Matrizen `U`, `D` und `V`:

```
> U<-e[[3]]
> D<-diag(e[[1]])
> V<-e[[2]]
```

Der Aufruf

```
> U%*%D%*%t(V)
```

liefert die Matrix `C`

```
      [,1] [,2]
 [1,]    1    2
 [2,]    1    1
 [3,]    1    2
```

B S-PLUS-Funktionen

B.1 Quartile

Die Funktion berechnet für den Vektor `x` die Quartile nach der Methode auf Seite 19.

```
quartile<-function(x) {
  # berechnet Quartile
  # x ist Datensatz
  x <- sort(x)
  uq <- median(x[1:ceiling(length(x)/2)])
  x <- rev(x)
  oq <- median(x[1:ceiling(length(x)/2)])
  return(c(uq, oq))
}
```

B.2 Distanzmatrix

Die Funktion `distfull` bildet aus dem Vektor `dis` mit den Distanzen die Distanzmatrix. Dem Attribut `"Size"` von `dis` muss die Anzahl der Beobachtungen zugewiesen sein.

```
> distfull<-function(dis) {
    n <- attr(dis, "Size")
    full <- matrix(0, n, n)
    full[lower.tri(full)] <- dis
    full + t(full)
}
```

B.3 Monotone Regression

Die Funktion monreg führt eine monotone Regression für die Elemente des Vektors p durch.

```
monreg<-function(p)
{
  # Monotone Regression der Elemente des Vektors p
   g <- rep(1, length(p))
   while(!all(p[ - length(p)] <= p[-1])) {
       i <- pooladjacent(p)
       p <- miblock(p, i, g)
       g <- gew(i, g)
   }
   rep(p, g)
}
```

Sie ruft folgende Funktionen auf:

```
pooladjacent<-function(p)
{
   i <- numeric(0)
   j <- 1
   h <- p
   while(length(h) > 0) {
       a <- sum(cumprod(h[ - length(h)] > h[-1])) + 1
       i <- c(i, rep(j, a))
       h <- h[ - (1:a)]
       j <- j + 1
   }
   i
}

miblock<-function(p, i, g)
{
   m <- numeric(0)
   for(j in 1:max(i))
       m <- c(m, sum(g[i == j] * p[i == j])/sum(g[i == j]))
   m
}

> gew<-function(ind, ag)
{
   m <- numeric(0)
   for(i in 1:max(ind))
       m <- c(m, sum(ag[ind == i]))
   m
}
```

B.4 STRESS1

Die Funktion stress1 berechnet STRESS1. Ihre Argumente sind der Vektor d der Distanzen und der Vektor disp der Disparitäten.

```
stress1<-function(d,disp)
{sqrt(sum((d-disp)^2)/sum(d^2)) }
```

B.5 Bestimmung einer neuen Konfiguration

Die Funktion Neuekon bestimmt bei nichtmetrischer mehrdimensionaler Skalierung eine neue Konfiguration. Das Argument X ist die alte Konfiguration und das Argument delta ist die Matrix $\boldsymbol{\Delta}$. Das Ergebnis ist eine Liste. Die erste Komponente ist die neue Konfiguration xneu und die zweite Komponente der Wert stress von STRESS1.

```
> Neuekon<-function(X, delta)
 { # Neue Konfiguration bei einer nichtmetrischen MDS.
   # X: Startkonfiguration
   # delta: Matrix Delta
   # Ergebnis ist Liste mit
   # xneu: neue Konfiguration # stress:
          Wert von STRESS1 fuer diese Konfiguration
    n <- dim(X)[1]
    delta <- delta[lower.tri(delta)]
    d <- dist(X)
    disp <- monreg(d[order(delta)])
    dm <- matrix(0, n, n)
    dm[lower.tri(dm)] <- d
    dm <- dm + t(dm)
    dispm <- matrix(0, n, n)
    dispm[lower.tri(dispm)] <- disp[rank(delta)]
    dispm <- dispm + t(dispm)
    xneu <- matrix(0, n, 2)
    for(i in 1:n) {
        h <- matrix((dm[i,-i] - dispm[i,-i])/dm[i,-i],n-1,2)
                   * (X[-i,]-matrix(X[i,],n-1,2,b = T))
        xneu[i,] <- X[i,] + apply(h,2,mean)
                  }
    d <- dist(xneu)
    disp <- monreg(d[order(delta)])
    stress <- stress1(delta, d, disp)
    list(xneu, stress)
}
```

B.6 Kophenetische Matrix

Die Funktion `cophenetic` bestimmt die kophenetische Matrix. Ihre Argumente sind Ergebnisse der Funktion `hclust`. Das erste Argument ist die Komponente `merge` und das zweite die Komponente `height`.

```
cophenetic<-function(m, h) {
    k <- length(h) + 1
    co <- matrix(0, k, k)
    obj <- abs(m[1,  ])
    grp <- rep(1, 2)
    co[ - m[1, 1],  - m[1, 2]] <- h[1]
    co[ - m[1, 2],  - m[1, 1]] <- h[1]
    for(i in 2:(k - 1)) {
        if(all(m[i,  ] < 0)) {
            obj <- c(obj, abs(m[i,  ]))
            grp <- c(grp, rep(i, 2))
            co[ - m[i, 1],  - m[i, 2]] <- h[i]
            co[ - m[i, 2],  - m[i, 1]] <- h[i]
        }
        else if(all(m[i,  ] > 0)) {
            z <- abs(obj[grp == abs(m[i, 1])])
            s <- abs(obj[grp == abs(m[i, 2])])
            obj <- c(obj, z, s)
            grp <- c(grp, rep(i, length(z) + length(s)))
            co[z, s] <- h[i]
            co[s, z] <- h[i]
        }
        else {
            z <- abs(m[i,  ][m[i,  ] < 0])
            obj <- c(obj, z)
            grp <- c(grp, i)
            pos <- abs(m[i,  ][m[i,  ] > 0])
            s <- abs(obj[grp == pos])
            obj <- c(obj, s)
            grp <- c(grp, rep(i, length(s)))
            co[z, s] <- h[i]
            co[s, z] <- h[i]
        }
    }
    co
}
```

B.7 Gamma-Koeffizient

Die Funktion `gammakoeffizient` bestimmt den Gamma-Koeffizienten zwischen den Vektoren `v1` und `v2`.

```
gammakoeffizient<-function(v1, v2) {
    m1 <- outer(v1, v1, FUN = "<")
    m1 <- m1[lower.tri(m1)]
    m2 <- outer(v2, v2, FUN = "<")
    m2 <- m2[lower.tri(m2)]
    m3 <- outer(v1, v1, FUN = ">")
    m3 <- m3[lower.tri(m3)]
    m4 <- outer(v2, v2, FUN = ">")
    m4 <- m4[lower.tri(m4)]
    C <- sum((m1 + m2) == 2)
    C <- C + sum((m3 + m4) == 2)
    D <- sum((m1 + m4) == 2)
    D <- D + sum((m2 + m3) == 2)
    (C - D)/(C + D)
}
```

B.8 Bestimmung der Zugehörigkeit zu Klassen

Die Funktion `welche.cluster` gibt für jedes Objekt an, zu welcher Klasse es gehört. Ihre Argumente sind Ergebnisse der Funktion `hclust` und die Anzahl `anz` der Klassen. Das erste Argument ist die Komponente `merge` und das zweite die Komponente `height`. Das dritte Argument ist die Anzahl `anz` der Klassen.

```
welche.cluster<-function(mer,hei, anz) {
    co <- cophenetic(mer, hei)
    h <- hei[length(hei) + 1 - anz]
    n <- ncol(co)
    cl <- rep(0, n)
    k <- 1
    for(i in 1:n) {
        if(cl[i] == 0) {
            ind <- (1:n)[co[i,  ] <= h]
            cl[ind] <- k
            k <- k + 1
        }
    }
    cl
}
```

B.9 Silhouette

Die Funktion `silhouette` liefert die Informationen zum Zeichnen einer Silhouette. Das Argument `wo` ist ein Vektor, der für jedes Objekt die Nummer der Klasse enthält, zu der es gehört. Das Argument `d` ist die Distanzmatrix und das Argument `namen` ist ein Vektor mit den Namen der Objekte. Das Ergebnis der Funktion `silhouette` ist eine Liste. Die erste Komponente ist eine Matrix. In der ersten Spalte steht die Nummer der Klasse des Objekts, in der zweiten Spalte die Nummer der Klasse, die am nächsten liegt, und in der dritten Spalte der Wert von $s(i)$. Die Namen der Objekte sind die Namen der ersten Dimension der Matrix. Die zweite Komponente enthält den Vektor der Mittelwerte der $s(i)$ der Klasse, die dritte Komponente den Silhouettenkoeffizienten.

```
silhouette<-function(wo, d, namen) {
   if(is.numeric(namen))
       namen <- as.character(namen)
   anzgr <- max(wo)
   n <- length(wo)
   indgr <- matrix(0, anzgr, 2)
   gruppen <- numeric(0)
   mi <- 1
   for(k in 1:anzgr) {
       g <- (1:n)[wo == k]
       indgr[k,  ] <- c(mi, mi + length(g) - 1)
       mi <- mi + length(g)
       gruppen <- c(gruppen, g)
   }
   b <- rep(0, n)
   a <- rep(0, n)
   naechstes <- rep(0, n)
   for(i in 1:n) {
       andere <- (1:anzgr)[ - wo[i]]
       bgr <- rep(0, length(andere))
       for(j in 1:length(andere))
           bgr[j] <- mean(d[i, gruppen[indgr[andere[j], 1]:
                                    indgr[andere[j], 2]]])
       b[i] <- min(bgr)
       naechstes[i] <- andere[bgr == b[i]]
       eigene <- gruppen[indgr[wo[i], 1]:indgr[wo[i], 2]]
       if(length(eigene) == 1)
           a[i] <- b[i]
       else a[i] <- mean(d[i, eigene[eigene != i]])
   }
   si <- (b - a)/pmax(a, b)
```

```
    siclu <- rep(0, anzgr)
    for(l in 1:anzgr)
        siclu[l] <- mean(si[wo == l])
    m <- cbind(wo, naechstes, si)
    namen <- namen[order(m[, 1])]
    m <- m[order(m[, 1]),  ]
    ms <- numeric(0)
    namens <- numeric(0)
    clusmittel <- rep(0, anzgr)
    for(i in 1:anzgr) {
        h <- matrix(m[m[, 1] == i,  ], ncol = 3)
        clusmittel[i] <- mean(h[, 3])
        n <- namen[m[, 1] == i]
        ms <- rbind(ms, h[rev(order(h[, 3])),  ])
        namens <- c(namens, n[rev(order(h[, 3]))])
    }
    dimnames(ms) <- list(namens, dimnames(ms)[[2]])
    list(ms, clusmittel, mean(ms[, 3]))
}
```

B.10 Zeichnen einer Silhouette

Die Funktion `plotsilhouette` zeichnet eine Silhouette. Ihr Argument ist das Ergebnis der Funktion `silhouette`. Sie beruht auf der Funktion `plot.partition` aus der Library `cluster` und wurde für unsere Zwecke angepaßt.

```
 plotsilhouette<-function(silinfo) {
    S <- rev(silinfo[[1]][, 3])
    space <- c(0, rev(diff(silinfo[[1]][, 1])))
    names <- rev(dimnames(silinfo[[1]])[[1]])
    if(!is.character(names))
        names <- as.character(names)
    barplot(S, space = space, names = names,
    xlab = "Breite der Silhouette", ylab = "",
    xlim = c(min(0, min(S)), 1), horiz = T,
    mgp = c(2.5, 1, 0))
    invisible()
}
```

C Tabellen

C.1 Standardnormalverteilung

Tabelle C.1. Quantil z_p der Standardnormalverteilung

p	.000	.001	.002	.003	.004	.005	.006	.007	.008	.009
0.50	0.000	0.002	0.005	0.008	0.010	0.012	0.015	0.018	0.020	0.023
0.51	0.025	0.028	0.030	0.033	0.035	0.038	0.040	0.043	0.045	0.048
0.52	0.050	0.053	0.055	0.058	0.060	0.063	0.065	0.068	0.070	0.073
0.53	0.075	0.078	0.080	0.083	0.085	0.088	0.090	0.093	0.095	0.098
0.54	0.100	0.103	0.106	0.108	0.110	0.113	0.116	0.118	0.121	0.123
0.55	0.126	0.128	0.131	0.133	0.136	0.138	0.141	0.143	0.146	0.148
0.56	0.151	0.154	0.156	0.159	0.161	0.164	0.166	0.169	0.171	0.174
0.57	0.176	0.179	0.182	0.184	0.187	0.189	0.192	0.194	0.197	0.199
0.58	0.202	0.204	0.207	0.210	0.212	0.215	0.217	0.220	0.222	0.225
0.59	0.228	0.230	0.233	0.235	0.238	0.240	0.243	0.246	0.248	0.251
0.60	0.253	0.256	0.258	0.261	0.264	0.266	0.269	0.272	0.274	0.277
0.61	0.279	0.282	0.284	0.287	0.290	0.292	0.295	0.298	0.300	0.303
0.62	0.306	0.308	0.311	0.313	0.316	0.319	0.321	0.324	0.327	0.329
0.63	0.332	0.334	0.337	0.340	0.342	0.345	0.348	0.350	0.353	0.356
0.64	0.358	0.361	0.364	0.366	0.369	0.372	0.374	0.377	0.380	0.383
0.65	0.385	0.388	0.391	0.393	0.396	0.399	0.402	0.404	0.407	0.410
0.66	0.412	0.415	0.418	0.421	0.423	0.426	0.429	0.432	0.434	0.437
0.67	0.440	0.443	0.445	0.448	0.451	0.454	0.456	0.459	0.462	0.465
0.68	0.468	0.470	0.473	0.476	0.479	0.482	0.484	0.487	0.490	0.493
0.69	0.496	0.499	0.501	0.504	0.507	0.510	0.513	0.516	0.519	0.522
0.70	0.524	0.527	0.530	0.533	0.536	0.539	0.542	0.545	0.548	0.550
0.71	0.553	0.556	0.559	0.562	0.565	0.568	0.571	0.574	0.577	0.580
0.72	0.583	0.586	0.589	0.592	0.595	0.598	0.601	0.604	0.607	0.610
0.73	0.613	0.616	0.619	0.622	0.625	0.628	0.631	0.634	0.637	0.640
0.74	0.643	0.646	0.650	0.653	0.656	0.659	0.662	0.665	0.668	0.671

Tabelle C.2. Quantil z_p der Standardnormalverteilung

p	.000	.001	.002	.003	.004	.005	.006	.007	.008	.009
0.75	0.674	0.678	0.681	0.684	0.687	0.690	0.694	0.697	0.700	0.703
0.76	0.706	0.710	0.713	0.716	0.719	0.722	0.726	0.729	0.732	0.736
0.77	0.739	0.742	0.745	0.749	0.752	0.755	0.759	0.762	0.766	0.769
0.78	0.772	0.776	0.779	0.782	0.786	0.789	0.793	0.796	0.800	0.803
0.79	0.806	0.810	0.813	0.817	0.820	0.824	0.827	0.831	0.834	0.838
0.80	0.842	0.845	0.849	0.852	0.856	0.860	0.863	0.867	0.870	0.874
0.81	0.878	0.882	0.885	0.889	0.893	0.896	0.900	0.904	0.908	0.912
0.82	0.915	0.919	0.923	0.927	0.931	0.935	0.938	0.942	0.946	0.950
0.83	0.954	0.958	0.962	0.966	0.970	0.974	0.978	0.982	0.986	0.990
0.84	0.994	0.999	1.003	1.007	1.011	1.015	1.019	1.024	1.028	1.032
0.85	1.036	1.041	1.045	1.049	1.054	1.058	1.062	1.067	1.071	1.076
0.86	1.080	1.085	1.089	1.094	1.098	1.103	1.108	1.112	1.117	1.122
0.87	1.126	1.131	1.136	1.141	1.146	1.150	1.155	1.160	1.165	1.170
0.88	1.175	1.180	1.185	1.190	1.195	1.200	1.206	1.211	1.216	1.221
0.89	1.226	1.232	1.237	1.243	1.248	1.254	1.259	1.265	1.270	1.276
0.90	1.282	1.287	1.293	1.299	1.305	1.311	1.316	1.322	1.328	1.335
0.91	1.341	1.347	1.353	1.360	1.366	1.372	1.379	1.385	1.392	1.398
0.92	1.405	1.412	1.419	1.426	1.432	1.440	1.447	1.454	1.461	1.468
0.93	1.476	1.483	1.491	1.498	1.506	1.514	1.522	1.530	1.538	1.546
0.94	1.555	1.563	1.572	1.580	1.589	1.598	1.607	1.616	1.626	1.635
0.95	1.645	1.655	1.665	1.675	1.685	1.695	1.706	1.717	1.728	1.739
0.96	1.751	1.762	1.774	1.787	1.799	1.812	1.825	1.838	1.852	1.866
0.97	1.881	1.896	1.911	1.927	1.943	1.960	1.977	1.995	2.014	2.034
0.98	2.054	2.075	2.097	2.120	2.144	2.170	2.197	2.226	2.257	2.290
0.99	2.326	2.366	2.409	2.457	2.512	2.576	2.652	2.748	2.878	3.090

C.2 χ^2-Verteilung

Tabelle C.3. Quantile der χ^2-Verteilung mit k Freiheitsgraden

k	$\chi^2_{k;0.95}$	$\chi^2_{k;0.975}$	$\chi^2_{k;0.9833}$	$\chi^2_{k;0.9875}$	$\chi^2_{k;0.99}$
1	3.84	5.02	5.73	6.24	6.63
2	5.99	7.38	8.19	8.76	9.21
3	7.81	9.35	10.24	10.86	11.34
4	9.49	11.14	12.09	12.76	13.28
5	11.07	12.83	13.84	14.54	15.09
6	12.59	14.45	15.51	16.24	16.81
7	14.07	16.01	17.12	17.88	18.48
8	15.51	17.53	18.68	19.48	20.09
9	16.92	19.02	20.21	21.03	21.67
10	18.31	20.48	21.71	22.56	23.21
11	19.68	21.92	23.18	24.06	24.72
12	21.03	23.34	24.63	25.53	26.22
13	22.36	24.74	26.06	26.98	27.69
14	23.68	26.12	27.48	28.42	29.14
15	25.00	27.49	28.88	29.84	30.58
16	26.30	28.85	30.27	31.25	32.00
17	27.59	30.19	31.64	32.64	33.41
18	28.87	31.53	33.01	34.03	34.81
19	30.14	32.85	34.36	35.40	36.19
20	31.41	34.17	35.70	36.76	37.57
21	32.67	35.48	37.04	38.11	38.93
22	33.92	36.78	38.37	39.46	40.29
23	35.17	38.08	39.68	40.79	41.64
24	36.42	39.36	41.00	42.12	42.98
25	37.65	40.65	42.30	43.45	44.31

C.3 t-Verteilung

Tabelle C.4. Quantile der t-Verteilung mit k Freiheitsgraden

k	$t_{k;0.90}$	$t_{k;0.95}$	$t_{k;0.975}$	$t_{k;0.99}$	$t_{k;0.995}$
1	3.0777	6.3138	12.7062	31.8205	63.6567
2	1.8856	2.9200	4.3027	6.9646	9.9248
3	1.6377	2.3534	3.1824	4.5407	5.8409
4	1.5332	2.1318	2.7764	3.7469	4.6041
5	1.4759	2.0150	2.5706	3.3649	4.0321
6	1.4398	1.9432	2.4469	3.1427	3.7074
7	1.4149	1.8946	2.3646	2.9980	3.4995
8	1.3968	1.8595	2.3060	2.8965	3.3554
9	1.3830	1.8331	2.2622	2.8214	3.2498
10	1.3722	1.8125	2.2281	2.7638	3.1693
11	1.3634	1.7959	2.2010	2.7181	3.1058
12	1.3562	1.7823	2.1788	2.6810	3.0545
13	1.3502	1.7709	2.1604	2.6503	3.0123
14	1.3450	1.7613	2.1448	2.6245	2.9768
15	1.3406	1.7531	2.1314	2.6025	2.9467
16	1.3368	1.7459	2.1199	2.5835	2.9208
17	1.3334	1.7396	2.1098	2.5669	2.8982
18	1.3304	1.7341	2.1009	2.5524	2.8784
19	1.3277	1.7291	2.0930	2.5395	2.8609
20	1.3253	1.7247	2.0860	2.5280	2.8453
21	1.3232	1.7207	2.0796	2.5176	2.8314
22	1.3212	1.7171	2.0739	2.5083	2.8188
23	1.3195	1.7139	2.0687	2.4999	2.8073
24	1.3178	1.7109	2.0639	2.4922	2.7969
25	1.3163	1.7081	2.0595	2.4851	2.7874
26	1.3150	1.7056	2.0555	2.4786	2.7787
27	1.3137	1.7033	2.0518	2.4727	2.7707
28	1.3125	1.7011	2.0484	2.4671	2.7633
29	1.3114	1.6991	2.0452	2.4620	2.7564
30	1.3104	1.6973	2.0423	2.4573	2.7500

C.4 *F*-Verteilung

Tabelle C.5. Das 0.95-Quantil $F_{m,n;0.95}$ der F-Verteilung mit m und n Freiheitsgraden

n \ m	1	2	3	4	5	6	7	8	9	10
1	161.45	199.5	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16

Tabelle C.6. Das 0.95-Quantil $F_{m,n;0.95}$ der F-Verteilung mit m und n Freiheitsgraden

n \ m	1	2	3	4	5	6	7	8	9	10
31	4.16	3.30	2.91	2.68	2.52	2.41	2.32	2.25	2.20	2.15
32	4.15	3.29	2.90	2.67	2.51	2.40	2.31	2.24	2.19	2.14
33	4.14	3.28	2.89	2.66	2.50	2.39	2.30	2.23	2.18	2.13
34	4.13	3.28	2.88	2.65	2.49	2.38	2.29	2.23	2.17	2.12
35	4.12	3.27	2.87	2.64	2.49	2.37	2.29	2.22	2.16	2.11
36	4.11	3.26	2.87	2.63	2.48	2.36	2.28	2.21	2.15	2.11
37	4.11	3.25	2.86	2.63	2.47	2.36	2.27	2.20	2.14	2.10
38	4.10	3.24	2.85	2.62	2.46	2.35	2.26	2.19	2.14	2.09
39	4.09	3.24	2.85	2.61	2.46	2.34	2.26	2.19	2.13	2.08
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08
41	4.08	3.23	2.83	2.60	2.44	2.33	2.24	2.17	2.12	2.07
42	4.07	3.22	2.83	2.59	2.44	2.32	2.24	2.17	2.11	2.06
43	4.07	3.21	2.82	2.59	2.43	2.32	2.23	2.16	2.11	2.06
44	4.06	3.21	2.82	2.58	2.43	2.31	2.23	2.16	2.10	2.05
45	4.06	3.20	2.81	2.58	2.42	2.31	2.22	2.15	2.10	2.05
46	4.05	3.20	2.81	2.57	2.42	2.30	2.22	2.15	2.09	2.04
47	4.05	3.20	2.80	2.57	2.41	2.30	2.21	2.14	2.09	2.04
48	4.04	3.19	2.80	2.57	2.41	2.29	2.21	2.14	2.08	2.03
49	4.04	3.19	2.79	2.56	2.40	2.29	2.20	2.13	2.08	2.03
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03
51	4.03	3.18	2.79	2.55	2.40	2.28	2.20	2.13	2.07	2.02
52	4.03	3.18	2.78	2.55	2.39	2.28	2.19	2.12	2.07	2.02
53	4.02	3.17	2.78	2.55	2.39	2.28	2.19	2.12	2.06	2.01
54	4.02	3.17	2.78	2.54	2.39	2.27	2.18	2.12	2.06	2.01
55	4.02	3.16	2.77	2.54	2.38	2.27	2.18	2.11	2.06	2.01
56	4.01	3.16	2.77	2.54	2.38	2.27	2.18	2.11	2.05	2.00
57	4.01	3.16	2.77	2.53	2.38	2.26	2.18	2.11	2.05	2.00
58	4.01	3.16	2.76	2.53	2.37	2.26	2.17	2.10	2.05	2.00
59	4.00	3.15	2.76	2.53	2.37	2.26	2.17	2.10	2.04	2.00
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99

Literaturverzeichnis

Agresti, A. (1990): Categorical data analysis. Wiley, New York

Andersen, E. B. (1991): The statistical analysis of categorical data. Springer, Berlin, 2nd edition

Anderson, T. W. (1984): An Introduction to Multivariate Statistical Analysis. Wiley, New York, 2nd edition

Bacher, J. (1994): Clusteranalyse: anwendungsorientierte Einführung. Oldenbourg, München

Bankhofer, U. (1995): Unvollständige Daten- und Distanzmatrizen in der Multivariaten Datenanalyse. Eul, Bergisch Gladbach

Basilevsky, A. (1983): Applied matrix algebra in the statistical sciences. North-Holland, New York

Basilevsky, A. (1994): Statistical factor analysis and related methods: theory and applications. Wiley, New York

Birkes, D., Dodge, Y. (1993): Alternative methods of regression. Wiley, New York

Bock, H. H. (1974): Automatische Klassifikation. Vandenhoeck & Ruprecht, Göttingen

Bödeker, M., Franke, K. (2001): Analyse der Potenziale und Grenzen von Virtual Reality Technologien auf industriellen Anwendermärkten. Diplomarbeit, Universität Bielefeld

Bollen, K. A. (1989): Structural equations with latent variables. Wiley, New York

Borg, I., Groenen, P. (1997): Modern multidimensional scaling: theory and applications. Springer, New York

Breiman, L., Friedman, J. H., Olshen, R. A., Stone, C. J. (1984): Classification and regression trees. Wadsworth, Belmont

Büning, H. (1991): Robuste und adaptive Tests. de Gruyter, Berlin

Brühl, O., Kahn, T. (2001): Analyse der Standortqualität zur Beurteilung der wirtschaftlichen Leistungsfähigkeit im interregionalen Vergleich. Diplomarbeit, Universität Bielefeld

Büning, H. (1996): Adaptive tests for the c-sample location problem - the case of two-sided alternatives. Communications in Statistics - Theory and Methods, **25**, 1569–1582

Büning, H., Naeve, P., Trenkler, G., Waldmann, K.-H. (2000): Mathematik für Ökonomen im Hauptstudium. Oldenbourg, München

Büning, H., Trenkler, G. (1994): Nichtparametrische statistische Methoden. de Gruyter, Berlin, 2nd edition

Calinski, T., Harabasz, J. (1974): A dendrite method for cluster analysis. Communications in Statistics - Theory and Methods A, **3**, 1–27

Carroll, J. D., Chang, J. J. (1970): Analysis of individual differences in multidimensional scaling via an N-way generalization of Eckart-Young decomposition. Psychometrika, **35**, 283–320

Carroll, R. J., Ruppert, D. (1988): Transformation and weighting in regression. Chapman & Hall, London

Cattell, R. B. (1966): The scree test for the number of factors. Multivariate Behavioral Research, **1**, 245–276

Christensen, R. (1997): Log-linear models and logistic regression. Springer, New York, 2nd edition

Clark, L. A., Pregibon, D. (1992): Tree-based models. In Chambers, J. M., Hastie, T. J. (eds.) Statistical models in S, Pacific Grove

Cook, R. D., Weisberg, S. (1982): Residuals and influence in regression. Chapman & Hall, New York

Cox, T. F., Cox, M. A. A. (1994): Multidimensional scaling. Chapman & Hall, London

Davison, M. L. (1983): Multidimensional scaling. Wiley, New York

Deutsches PISA-Konsortium (Hrsg.) (2001): PISA 2000. Leske + Budrich, Opladen

Doane, D. P. (1976): Aesthetic frequency classifications. American Statistician, **30**, 181–183

Draper, N. R., Smith, H. (1998): Applied regression analysis. Wiley, New York, 3rd edition

Everitt, B. (2001): Cluster analysis. Arnold, London, 4th edition

Fahrmeir, L., Hamerle, A., Tutz, G. (1996): Multivariate statistische Verfahren. de Gruyter, Berlin, 2nd edition

Fahrmeir, L., Künstler, R., Pigeot, I., Tutz, G. (2001): Statistik : der Weg zur Datenanalyse. Springer, Berlin, 3rd edition

Friedman, J. H., Tukey, J. W. (1974): A projection pursuit algorithm for exploratory data analysis. IEEE Transactions on Computers, **23**, 881–890

Goodman, L. A. (1971): The analysis of multidimensional contingency tables: stepwise procedures and direct estimation methods for building models for multiple classifications. Technometrics, **13**, 33–61

Goodman, L. A., Kruskal, W. H. (1954): Measures of association for cross-classification. Journal of the American Statistical Association, **49**, 732–764

Gordon, A. D. (1999): Classification. Chapman & Hall, Boca Raton, 2nd edition

Gower, J. C. (1971): A general coefficient of similarity and some of its properties. Biometrics, **27**, 857–872

Gower, J. C. (1975): Generalized procrustes analysis. Psychometrika, **40**, 33–51

Gower, J. C., Legendre, P. (1986): Metric and Euclidean properties of dissimilarity coefficients. Journal of Classification, **3**, 5–48

Guttman, L. (1954): Some necessary conditions for common factor analysis. Psychometrika, **19**, 149–161

Hand, D. J. (1997): Construction and assessment of classification rules. Wiley, Chichester

Härdle, W. (1990a): Applied nonparametric regression. Cambridge Univ. Press, Cambridge

Härdle, W. (1990b): Smoothing techniques: with implementation in S. Springer, Berlin

Hastie, T. J., Tibshirani, R. J. (1991): Generalized additive models. Chapman & Hall, London

Hastie, T. J., Tibshirani, R. J., Friedman, J. H. (2001): The elements of statistical learning : data mining, inference, and prediction. Springer, New York

Heiler, S., Michels, P. (1994): Deskriptive und explorative Datenanalyse. Oldenbourg, München

Hosmer, D. W., Lemeshow, S. (1989): Applied logistic regression. Wiley, New York

Huber, P. J. (1985): Projection pursuit (with discussion). Ann. Statist., **13**, 435–535

Hubert, L. (1974): Approximate evaluation techniques for the single-link and complete-link hierarchical clustering procedures. Journal of the American Statistical Association, **69**, 698–704

Huberty, C. J. (1994): Applied discriminant analysis. Wiley, New York

Hummel, J. (1996): Linked Bar Charts: Analysing Categorial Data Graphically. Computational Statistics, **11**, 23–33

Hyndman, R. J., Fan, Y. (1996): Sample quantiles in statistical packages. The American Statistician, **50**, 361–365

Jaccard, P. (1908): Nouvelles recherches sur la distribution florale. Bulletin de la Societe Vaudoise de Sciences Naturelles, **44**, 223–370

Jackson, J. E. (1991): A user's guide to principal components. Wiley, New York

Jänich, K. (2000): Lineare Algebra. Springer, Berlin, 8th edition

Jobson, J. D. (1992): Applied multivariate data analysis. Volume II. Categorical and multivariate methods. Springer, New York

Johnson, S. C. (1967): Hierarchical clustering scheme. Psychometrika **23**, 241–254

Johnson, R. A., Wichern, D. W. (1998): Applied multivariate statistical analysis. Prentice Hall, New Jersey, 4. Auflage

Jolliffe, I. T. (1972): Discarding variables in principal component analysis, I: Artificial data. Applied Statistics, **21**, 160–173

Jolliffe, I. T. (1986): Principal component analysis. Springer, New York

Jones, M. C., Sibson, R. (1987): What is projection pursuit? Journal of the Royal Statistical Society, Series A, **150**, 1–36

Kaiser, H. F. (1958): The varimax criterion for analytic rotation in factor analysis. Psychometrika, **23**, 187–200

Kaiser, H. F. (1960): The application of electronic computers to factor analysis. Educ. Psychol. Meas., **20**, 141–151

Kaufman, L., Rousseeuw, P. J. (1990): Finding groups in data. Wiley, New York

Kearsley, A. J., Tapia, R. A., Trosset, M. W. (1998): The solution of the metric STRESS and SSTRESS problems in multidimensional scaling using Newton's method. Computational Statistics, **13**, 369–396

Khuri, A. I. (1993): Advanced calculus with applications in statistics. Wiley, New York

Kleinbaum, D. G. (1994): Logistic regression: a self-learning text. Springer, New York

Krause, A., Olson, M. (2000): The basics of S and S-PLUS. Springer, New York

Kruskal, J. B. (1956): On the shortest spanning subtree of a graph and the travelling salesman problem. Poceedings AMS, **7**, 48–50

Kruskal, J. B. (1964): Nonmetric multidimensional scaling: a numerical method. Psychometrika, **29**, 115–129

Krzanowski, W. J. (2000): Principles of multivariate analysis: a user's perspective. Oxford University Press, Oxford, rev. edition

Lachenbruch, P. A., Mickey, M. R. (1968): Estimation of error rates in discriminant analysis. Technometrics, **10**, 1–11

Lance, G. N., W. T. Williams (1967): A general theory of classifactory sorting strategies I. Hierarchical systems. Comp. J., **9**, 373–380

Lasch, R., Edel, R. (1994): Einsatz multivariater Verfahren zur Analyse von Geschäftsstellen eines Kreditinstituts. Diskussionsarbeit am Institut für Statistik und mathematische Wirtschaftstheorie der Universität Augsburg

Loh, W. Y., Shih, Y. S. (1997): Split selection methods for classification trees. Statistica Sinica, **7**, 815–840

Mardia, K. V. (1978): Some properties of classical multidimensionale scaling. Communications in Statistics - Theory and Methods A, **7**, 1233–1241

Mardia, K. V., Kent, J. T., Bibby, J. M. (1979): Multivariate analysis. Academic Press, London

McLachlan, G. J. (1992): Discriminant analysis and statistical pattern recognition. Wiley, New York

Miller, R. G. (1981): Simultaneous statistical inference. Springer, New York, 2nd edition

Milligan, G. W., Cooper, M. C. (1985): An examination of procedures for determining the number of clusters in a data set. Psychometrika, **50**, 159–179

Mojena, R. (1977): Hierarchical grouping methods and stopping rules: an evaluation. Computer Journal, **20**, 359–363

Mood, A. M., Graybill, F. A., Boes, D. C. (1974): Introduction to the theory of statistics. McGraw-Hill, New York, 3rd edition

Neuhaus, J., Wrigley, C. (1954): The quartimax method: an analytical approach to orthogonal simple structure. British Journal of Mathematical and Statistical Psychology, **7**, 81–91

Rice, J. A. (1988): Mathematical statistics and data analysis. Wadsworth, Pacific Grove

Ripley, B. D. (1996): Pattern recognition and neural networks. Cambridge Univ. Press, Cambridge

Rogers, D. J., Tanimoto, T. T. (1960): A computer program for classifying plants. Science, **132**, 1115–1118

Rosenblatt, M. (1956): Remarks on some nonparametric estimates of a density function. Ann. Math. Statist., **27**, 832–837

Rousseeuw, P. J. (1984): Least median of squares regression. Journal of the American Statistical Association, **79**, 871–880

Rousseeuw, P. J. (1987): Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. Journal of Computational and Applied Mathematics, **20**, 53–65

Rousseeuw, P. J., Ruts, I., Tukey, J. W. (1999): The bagplot: a bivariate boxplot. The American Statistician, **53**, 382–387

Rousseeuw, P. J., van Driessen, K. (1999): A fast algorithm for the minimum covariance determinant estimator. Technometrics, **41**, 212–223

Schafer, J. L. (1997): Analysis of incomplete multivariate data. Chapman & Hall, London

Schlittgen, R. (1996): Statistische Inferenz. Oldenbourg, München

Schlittgen, R. (2000): Einführung in die Statistik. Oldenbourg, München, 9th edition

Scott, D. W. (1992): Multivariate Density Estimation. Wiley, New York

Seber, G. A. F. (1977): Linear regression analysis. Wiley, New York

Seber, G. A. F. (1984): Multivariate observations. Wiley, New York

Silverman, B. W. (1986): Density Estimation. Chapman & Hall, London

Small, C. G. (1990): A survey of multidimensional medians. International Statistical Review, **58**, 263–277

Smith, M. (1993): Neural networks for statistical modelling. Van Nostrand Reinhold, New York

Sneath, P. H., Sokal, R. R. (1973): Principles of numerical taxinomy. Freeman, San Francisco

Sokal, R. R., Michener, C. D. (1958): A statistical method for evaluation systematic relationships. The University of Kansas Scientific Bulletin, **38**, 1409–1438

Strang, G. (1988): Linear algebra and its applications. Harcourt Brace Jovanovich, San Diego, 3rd edition

Süselbeck, B. (1993): S und S-PLUS : Eine Einführung in Programmierung und Anwendung. Fischer, Stuttgart

Trippel, A. (2001): Ein Vergleich numerischer Methoden zur Lösung von nichtmetrischen MDS-Problemen. Diplomarbeit, Universität Bielefeld

Tufte, E. R. (2001): The visual display of quantitative information. Graphics Press, Cheshire, 2nd edition

Tukey, J. W. (1977): Exploratory data analysis. Addison-Wesley, Reading, Mass.

Venables, W. N., Ripley, B. D. (1999): Modern applied statistics with S-PLUS. Springer, Berlin, 3rd edition

Wainer, H. (1997): Visual revelations : graphical tales of fate and deception from Napoleon Bonaparte to Ross Perot. Copernicus, New York

Ward J. H. (1963): Hierarchical grouping to optimize an objective function. JASA **58**, 236–244

Watkins, D. S. (1991): Fundamentals of matrix computations. Wiley, New York

Wetzel, W., Skarabis, H., Naeve, P., Büning, H. (1981): Mathematische Propädeutik für Wirtschaftswissenschaftler. de Gruyter, Berlin, 4th edition

Wishart, D. (1969): An algorithm for hierarchical classifications. Biometrics **25**, 165–170

Zurmühl, R., Falk, S. (1997): Matrizen 1: Grundlagen. Springer, Berlin, 7th edition

Index