Allgemeines Lineares Modell

BSc Psychologie, SoSe 2024

Joram Soch

(3) Matrizen

Motivation

Matrizen sind die Worte der Sprache der Datenanalyse.

Vektoren sind nur spezielle Matrizen. (Skalare sind nur spezielle Vektoren.)

Matrizen können als Tabellen der Datenrepräsentation dienen.

Matrizen können lineare Abbildungen repräsentieren.

Matrizen können Vektorräume repräsentieren.

$\Rightarrow$ Ein sicherer Umgang mit Matrizen ist für das Verständnis moderner datenanalytischer Verfahren unverzichtbar.

Definition

Operationen

Determinanten

Rang und Spur

Spezielle Matrizen

Selbstkontrollfragen

Definition (Matrix)

Eine Matrix ist eine rechteckige Anordnung von Zahlen, die wie folgt bezeichnet wird:

$$A := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} := (a_{ij})_{1 \le i \le n, \, 1 \le j \le m}. \tag{1}$$

Bemerkungen

- Matrizen bestehen aus *Zeilen (rows)* und *Spalten (columns)*.
- Die Matrixeinträge a_{ij} werden mit einem *Zeilenindex* i und einem *Spaltenindex* j indiziert.
- Zum Beispiel gilt für $A := \begin{pmatrix} 2 & 7 & 5 & 2 \\ 8 & 2 & 5 & 6 \\ 6 & 4 & 0 & 9 \\ 9 & 2 & 1 & 2 \end{pmatrix}$, dass $a_{32} = 4$.

Definition

Bemerkungen (fortgeführt)

- Die *Größe* oder *Dimension* einer Matrix ergibt sich aus der Anzahl ihrer Zeilen $n \in \mathbb{N}$ und Spalten $m \in \mathbb{N}$.
- Matrizen mit $n = m$ heißen *quadratische Matrizen*.
- In der Folge benötigen wir nur Matrizen mit reellen Einträgen, also $a_{ij} \in \mathbb{R}$ für alle $i = 1, ..., n, j = 1, ..., m$.
- Wir nennen die Matrizen mit reellen Einträge *reelle Matrizen*.
- Die Menge der reellen Matrizen mit n Zeilen und m Spalten bezeichnen wir mit $\mathbb{R}^{n \times m}$.
- Aus dem Ausdruck $A \in \mathbb{R}^{2 \times 3}$ lesen wir ab, dass A eine reelle Matrix mit zwei Zeilen und drei Spalten ist.
- Wir identifizieren die Menge $\mathbb{R}^{1 \times 1}$ mit der Menge $\mathbb{R}$.
- Wir identifizieren die Menge $\mathbb{R}^{n \times 1}$ mit der Menge $\mathbb{R}^n$.
- Reelle Matrizen mit einer Spalte und n Zeilen sind also dasselbe wie n-dimensionale reelle Vektoren.

Definition

Operationen

Determinanten

Rang und Spur

Spezielle Matrizen

Selbstkontrollfragen

Matrixoperationen

Man kann mit Matrizen rechnen.

In der Folge betrachten wir folgende grundlegende Matrixoperationen:

- Addition und Subtraktion von Matrizen gleicher Größe (Matrixaddition und Matrixsubtraktion)
- Multiplikation einer Matrix mit einem Skalar (Skalarmultiplikation)
- Vertauschen der Zeilen- und Spaltenanordnung (Matrixtransposition)
- Multiplikation einer Matrix mit einer passenden zweiten Matrix (Matrixmultiplikation)
- "Division" durch eine quadratische Matrix (Matrixinversion)

Definition (Matrixaddition)

Es seien $A, B \in \mathbb{R}^{n\times m}$. Dann ist die *Addition* von A und B definiert als die Abbildung

$$+ : \mathbb{R}^{n\times m} \times \mathbb{R}^{n\times m} \to \mathbb{R}^{n\times m}, \, (A, B) \mapsto +(A, B) := A + B \tag{2}$$

mit

$$\begin{aligned} A + B &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{pmatrix} \\ &:= \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1m} + b_{1m} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2m} + b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} + b_{n1} & a_{n2} + b_{n2} & \cdots & a_{nm} + b_{nm} \end{pmatrix}. \end{aligned} \tag{3}$$

Bemerkungen

- Nur Matrizen identischer Größe können miteinander addiert werden.
- Die Addition zweier gleich großer Matrizen ist *elementweise* definiert.

Definition (Matrixsubtraktion)

Es seien $A, B \in \mathbb{R}^{n\times m}$. Dann ist die *Subtraktion* von A und B definiert als die Abbildung

$$-: \mathbb{R}^{n\times m} \times \mathbb{R}^{n\times m} \to \mathbb{R}^{n\times m}, \; (A,B) \mapsto -(A,B) := A - B \tag{4}$$

mit

$$\begin{aligned}
A - B &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} - \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ b_{21} & b_{22} & \cdots & b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{pmatrix} \\
&:= \begin{pmatrix} a_{11} - b_{11} & a_{12} - b_{12} & \cdots & a_{1m} - b_{1m} \\ a_{21} - b_{21} & a_{22} - b_{22} & \cdots & a_{2m} - b_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} - b_{n1} & a_{n2} - b_{n2} & \cdots & a_{nm} - b_{nm} \end{pmatrix}.
\end{aligned} \tag{5}$$

Bemerkungen

- Nur Matrizen identischer Größe können voneinander subtrahiert werden.
- Die Subktration zweier gleich großer Matrizen ist *elementweise* definiert.

Operationen

Beispiel

Es seien $A, B \in \mathbb{R}^{2\times 3}$ definiert als

$$A := \begin{pmatrix} 2 & -3 & 0 \\ 1 & 6 & 5 \end{pmatrix} \text{ und } B := \begin{pmatrix} 4 & 1 & 0 \\ -4 & 2 & 0 \end{pmatrix}. \tag{6}$$

Da A und B gleich groß sind, können wir sie addieren

$$\begin{aligned} C := A + B &= \begin{pmatrix} 2 & -3 & 0 \\ 1 & 6 & 5 \end{pmatrix} + \begin{pmatrix} 4 & 1 & 0 \\ -4 & 2 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 2+4 & -3+1 & 0+0 \\ 1-4 & 6+2 & 5+0 \end{pmatrix} \\ &= \begin{pmatrix} 6 & -2 & 0 \\ -3 & 8 & 5 \end{pmatrix} \end{aligned} \tag{7}$$

und voneinander subtrahieren

$$\begin{aligned} D := A - B &= \begin{pmatrix} 2 & -3 & 0 \\ 1 & 6 & 5 \end{pmatrix} - \begin{pmatrix} 4 & 1 & 0 \\ -4 & 2 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 2-4 & -3-1 & 0-0 \\ 1+4 & 6-2 & 5-0 \end{pmatrix} \\ &= \begin{pmatrix} -2 & -4 & 0 \\ 5 & 4 & 5 \end{pmatrix}. \end{aligned} \tag{8}$$

Beispiel

```
# spaltenweise Definition von A (R-Default)
A = matrix(c(2,1,-3,6,0,5), nrow = 2)
print(A)
```

```
>      [,1] [,2] [,3]
> [1,]    2   -3    0
> [2,]    1    6    5
```

```
# zeilenweise Definition von B
B = matrix(c(4,1,0,-4,2,0), nrow = 2, byrow = TRUE)
print(B)
```

```
>      [,1] [,2] [,3]
> [1,]    4    1    0
> [2,]   -4    2    0
```

Operationen

Beispiel

```
# Addition
C = A + B
print(C)
```

```
>      [,1] [,2] [,3]
> [1,]    6   -2    0
> [2,]   -3    8    5
```

```
# Subtraktion
D = A - B
print(D)
```

```
>      [,1] [,2] [,3]
> [1,]   -2   -4    0
> [2,]    5    4    5
```

Definition (Skalarmultiplikation)

Es sei $c \in \mathbb{R}$ ein Skalar und $A \in \mathbb{R}^{n \times m}$. Dann ist die *Skalarmultiplikation* von c und A definiert als die Abbildung

$$\cdot : \mathbb{R} \times \mathbb{R}^{n \times m} \to \mathbb{R}^{n \times m}, \, (c, A) \mapsto \cdot(c, A) := cA \tag{9}$$

mit

$$cA = c \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} := \begin{pmatrix} ca_{11} & ca_{12} & \cdots & ca_{1m} \\ ca_{21} & ca_{22} & \cdots & ca_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ ca_{n1} & ca_{n2} & \cdots & ca_{nm} \end{pmatrix}. \tag{10}$$

Bemerkungen

- Die Skalarmultiplikation ist *elementweise* definiert.

Operationen

Beispiel

Es seiein $c := -3$ und $A \in \mathbb{R}^{4\times 3}$ definiert als

$$A := \begin{pmatrix} 3 & 1 & 1 \\ 5 & 2 & 5 \\ 2 & 7 & 1 \\ 3 & 4 & 2 \end{pmatrix}. \tag{11}$$

Dann ergibt sich

$$B := cA = -3\begin{pmatrix} 3 & 1 & 1 \\ 5 & 2 & 5 \\ 2 & 7 & 1 \\ 3 & 4 & 2 \end{pmatrix} = \begin{pmatrix} -3\cdot 3 & -3\cdot 1 & -3\cdot 1 \\ -3\cdot 5 & -3\cdot 2 & -3\cdot 5 \\ -3\cdot 2 & -3\cdot 7 & -3\cdot 1 \\ -3\cdot 3 & -3\cdot 4 & -3\cdot 2 \end{pmatrix} = \begin{pmatrix} -9 & -3 & -3 \\ -15 & -6 & -15 \\ -6 & -21 & -3 \\ -9 & -12 & -6 \end{pmatrix}. \tag{12}$$

Operationen

Beispiel

```
# Definitionen
c = -3                      # Skalar c
A = matrix(c(3,1,1,         # Matrix A
             5,2,5,
             2,7,1,
             3,4,2),
           nrow  = 4,       # Anzahl Zeilen
           byrow = TRUE)    # zeilenweise

# Skalarmultiplikation
B = c*A
print(B)
```

```
>      [,1] [,2] [,3]
> [1,]   -9   -3   -3
> [2,]  -15   -6  -15
> [3,]   -6  -21   -3
> [4,]   -9  -12   -6
```

Theorem (Vektorraum $\mathbb{R}^{n\times m}$)

Das Tripel $(\mathbb{R}^{n\times m}, +, \cdot)$ mit der oben definierten Matrixaddition und Skalarmultiplikation ist ein Vektorraum. Insbesondere gelten also für $A, B, C \in \mathbb{R}^{n\times m}$ und $r, s, t \in \mathbb{R}$ folgende Rechenregeln:

(1) Kommutativität der Addition	$A + B = B + A$
(2) Assoziativität der Addition	$(A + B) + C = A + (B + C)$
(3) Existenz eines neutralen Elements der Addition	$\exists 0 \in \mathbb{R}^{n\times m}$ mit $A + 0 = 0 + A = A$
(4) Existenz inverser Elemente der Addition	$\forall A \,\exists -A$ mit $A + (-A) = 0$
(5) Existenz eines neutralen Elements der Skalarmultiplikation	$\exists 1 \in \mathbb{R}$ mit $1 \cdot A = A$
(6) Assoziativität der Skalarmultiplikation	$r \cdot (s \cdot t) = (r \cdot s) \cdot t$
(7) Distributivität hinsichtlich der Matrixaddition	$r \cdot (A + B) = r \cdot A + r \cdot B$
(8) Distributivität hinsichtlich der Skalaraddition	$(r + s) \cdot A = r \cdot A + s \cdot A$

Bemerkungen

- Wir verzichten auf einen Beweis.
- Der Beweis ergibt sich mit dem elementweisen Charakter von $+, -, \cdot$ und den Rechenregeln in $(\mathbb{R}, +, \cdot)$.
- Das neutrale Element der Addition heißt *Nullmatrix*; wir schreiben $0_{nm} := (0)_{1\le i\le n, 1\le j\le m}$ mit $0 \in \mathbb{R}$.
- Die inversen Elemente der Addition sind durch $-A := (-a_{ij})_{1\le i\le n, 1\le j\le m}$ gegeben.
- Das neutrale Element der Skalarmultiplikation ist $1 \in \mathbb{R}$.

Definition (Matrixtransposition)

Es sei $A \in \mathbb{R}^{n\times m}$. Dann ist die *Transposition* von A definiert als die Abbildung

$$\cdot^{\mathrm{T}} : \mathbb{R}^{n\times m} \to \mathbb{R}^{m\times n},\ A \mapsto \cdot^{\mathrm{T}}(A) := A^{\mathrm{T}} \tag{13}$$

mit

$$A^{\mathrm{T}} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}^{\mathrm{T}} := \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{pmatrix} \tag{14}$$

Bemerkungen

- Die Matrixtransposition “vertauscht” Zeilen und Spalten.
- Das Transponieren einer Matrix wird auch als *Stürzen einer Matrix* bezeichnet (Animation).
- Für $A \in \mathbb{R}^{n\times m}$ gilt immer $A^{\mathrm{T}} \in \mathbb{R}^{m\times n}$.
- Für $A \in \mathbb{R}^{1\times 1}$ gilt immer $A^{\mathrm{T}} = A$.
- Es gilt $\left(A^{\mathrm{T}}\right)^{\mathrm{T}} = A$.
- Es gilt $(a_{ii})_{1\le i\le \min(n,m)} = (a_{ii})^{\mathrm{T}}_{1\le i\le \min(n,m)}$.
- Matrixelemente auf der Hauptdiagonalen einer Matrix bleiben bei Transposition also unberührt.

Beispiel

Es sei $A \in \mathbb{R}^{2\times 3}$ definiert durch

$$A := \begin{pmatrix} 2 & 3 & 0 \\ 1 & 6 & 5 \end{pmatrix} . \tag{15}$$

Dann gilt $A^{\mathrm{T}} \in \mathbb{R}^{3\times 2}$ und speziell

$$A^{\mathrm{T}} = \begin{pmatrix} 2 & 1 \\ 3 & 6 \\ 0 & 5 \end{pmatrix} . \tag{16}$$

Weiterhin gilt offenbar $\min(m, n) = 2$ und folglich

$$(a_{11}) = (a_{11})^{\mathrm{T}} \text{ und } (a_{22}) = (a_{22})^{\mathrm{T}} . \tag{17}$$

Operationen

Beispiel

```
# Definition
A = matrix(c(2,3,0,
             1,6,5),
           nrow  = 2,
           byrow = TRUE)
print(A)
```

```
>      [,1] [,2] [,3]
> [1,]    2    3    0
> [2,]    1    6    5
```

```
# Transposition
AT = t(A)
print(AT)
```

```
>      [,1] [,2]
> [1,]    2    1
> [2,]    3    6
> [3,]    0    5
```

Definition (Matrixmultiplikation)

Es seien $A \in \mathbb{R}^{n\times m}$ und $B \in \mathbb{R}^{m\times k}$. Dann ist die *Matrixmultiplikation* von A und B definiert als die Abbildung

$$\cdot : \mathbb{R}^{n\times m} \times \mathbb{R}^{m\times k} \to \mathbb{R}^{n\times k},\ (A,B) \mapsto \cdot(A,B) := AB \tag{18}$$

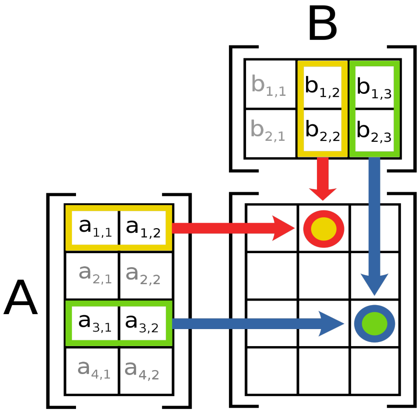
mit

$$\begin{aligned}
AB &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1k} \\ b_{21} & b_{22} & \cdots & b_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mk} \end{pmatrix} \\
&:= \begin{pmatrix} \sum_{i=1}^m a_{1i}b_{i1} & \sum_{i=1}^m a_{1i}b_{i2} & \cdots & \sum_{i=1}^m a_{1i}b_{ik} \\ \sum_{i=1}^m a_{2i}b_{i1} & \sum_{i=1}^m a_{2i}b_{i2} & \cdots & \sum_{i=1}^m a_{2i}b_{ik} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^m a_{ni}b_{i1} & \sum_{i=1}^m a_{ni}b_{i2} & \cdots & \sum_{i=1}^m a_{ni}b_{ik} \end{pmatrix} \\
&= \left(\sum_{i=1}^m a_{ji}b_{il}\right)_{1\le j\le n, 1\le l\le k}
\end{aligned} \tag{19}$$

Bemerkungen

- Das Matrixprodukt AB ist nur dann definiert, wenn A genau so viele Spalten hat wie B Zeilen.
- Informell gilt für die beteiligten Matrixgrößen immer $(n \times m)(m \times k) = (n \times k)$.
- In AB ist $(AB)_{ij}$ die Summe der multiplizierten $i$ten Zeilen von A und $j$ten Spalten von B.
- Zum Berechnen von $(AB)_{ij}$ für $1 \leq i \leq n, 1 \leq j \leq k$ geht man also wie folgt vor:
  1. Man legt in Gedanken die Tranposition der $i$ten Zeile von A über die $j$te Spalte von B.
  2. Weil A genau m Spalten hat und B genau m Zeilen hat, gibt es zu jedem Element der Zeile aus A ein korrespondierendes Element in der Spalte von B.
  3. Man multipliziert die korrespondierenden Elemente miteinander.
  4. Man summiert die Ergebnisse auf.
  5. Dies ist der Eintrag mit Index ij in AB.
- Die Multiplikation von Matrizen ist im Allgemeinen nicht kommutativ (also meist $AB \neq BA$).

Visualisierung der Matrixmultiplikation



(Quelle: *Wikimedia Commons*: "Matrix_multiplication_diagram_2.svg"; Lizenz: GNU Free, CC-BY-SA 3.0.)

Beispiel

$A \in \mathbb{R}^{2\times 3}$ und $B \in \mathbb{R}^{3\times 2}$ seien definiert als

$$A := \begin{pmatrix} 2 & -3 & 0 \\ 1 & 6 & 5 \end{pmatrix} \text{ und } B := \begin{pmatrix} 4 & 2 \\ -1 & 0 \\ 1 & 3 \end{pmatrix}. \tag{20}$$

Wir wollen $C := AB$ und $D := BA$ berechnen.

Mit $n = 2, m = 3$ und $k = 2$ wissen wir schon, dass $C \in \mathbb{R}^{2\times 2}$ und $D \in \mathbb{R}^{3\times 3}$, weil

$$(2 \times 3)(3 \times 2) = (2 \times 2) \tag{21}$$

und

$$(3 \times 2)(2 \times 3) = (3 \times 3) \tag{22}$$

Es gilt hier also sicher $AB \neq BA$.

Beispiel (fortgeführt)

Es ergibt sich zum einen

$$
\begin{aligned}
C &= AB \\
&= \begin{pmatrix} 2 & -3 & 0 \\ 1 & 6 & 5 \end{pmatrix} \begin{pmatrix} 4 & 2 \\ -1 & 0 \\ 1 & 3 \end{pmatrix} \\
&= \begin{pmatrix} 2 \cdot 4 + (-3) \cdot (-1) + 0 \cdot 1 & 2 \cdot 2 + (-3) \cdot 0 + 0 \cdot 3 \\ 1 \cdot 4 + 6 \cdot (-1) + 5 \cdot 1 & 1 \cdot 2 + 6 \cdot 0 + 5 \cdot 3 \end{pmatrix} \\
&= \begin{pmatrix} 8 + 3 + 0 & 4 + 0 + 0 \\ 4 - 6 + 5 & 2 + 0 + 15 \end{pmatrix} \\
&= \begin{pmatrix} 11 & 4 \\ 3 & 17 \end{pmatrix}.
\end{aligned} \tag{23}
$$

Operationen

Beispiel (fortgeführt)

```
# Definitionen
A = matrix(c( 2,-3, 0,
              1, 6, 5),
           nrow  = 2,
           byrow = TRUE)
B = matrix(c( 4, 2,
             -1, 0,
              1, 3),
           nrow  = 3,
           byrow = TRUE)

# Matrixmultiplikation
C = A %*% B                     # (2 x 3)(3 x 2) = (2 x 2)
print(C)
```

```
>      [,1] [,2]
> [1,]   11    4
> [2,]    3   17
```

Beispiel (fortgeführt)

Es ergibt sich zum anderen

$$
\begin{aligned}
D &= BA \\
&= \begin{pmatrix} 4 & 2 \\ -1 & 0 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 2 & -3 & 0 \\ 1 & 6 & 5 \end{pmatrix} \\
&= \begin{pmatrix} 4\cdot 2 + 2\cdot 1 & 4\cdot(-3) + 2\cdot 6 & 4\cdot 0 + 2\cdot 5 \\ (-1)\cdot 2 + 0\cdot 1 & (-1)\cdot(-3) + 0\cdot 6 & (-1)\cdot 0 + 0\cdot 5 \\ 1\cdot 2 + 3\cdot 1 & 1\cdot(-3) + 3\cdot 6 & 1\cdot 0 + 3\cdot 5 \end{pmatrix} \\
&= \begin{pmatrix} 8+2 & -12+12 & 0+5 \\ -2+0 & 3+0 & 0+0 \\ 2+3 & -3+18 & 0+15 \end{pmatrix} \\
&= \begin{pmatrix} 10 & 0 & 10 \\ -2 & 3 & 0 \\ 5 & 15 & 15 \end{pmatrix}.
\end{aligned}
\tag{24}
$$

Operationen

Beispiel (fortgeführt)

```
# Definitionen
A = matrix(c( 2,-3, 0,
              1, 6, 5),
           nrow  = 2,
           byrow = TRUE)
B = matrix(c( 4, 2,
             -1, 0,
              1, 3),
           nrow  = 3,
           byrow = TRUE)

# Matrixmultiplikation
D = B %*% A                     # (3 x 2)(2 x 3) = (3 x 3)
print(D)
```

```
>      [,1] [,2] [,3]
> [1,]   10    0   10
> [2,]   -2    3    0
> [3,]    5   15   15
```

```
# Beispiel für eine nicht-definierte Matrixmultiplikation
E = t(A) %*% B                  # (3 x 2)(3 x 2)
```

```
> Error in t(A) %*% B: non-conformable arguments
```

Theorem (Matrixmultiplikation und Skalarprodukt)

Es seien $x, y \in \mathbb{R}^n$. Dann gilt:

$$\langle x, y\rangle = x^{\mathrm{T}} y. \tag{25}$$

Weiterhin seien für $A \in \mathbb{R}^{m\times n}$ die Spalten von A^{T} für $i = 1, ..., m$ gegeben durch

$$\bar{a}_i := (a_{ji})_{1\le j\le n} \in \mathbb{R}^n \tag{26}$$

und für $B \in \mathbb{R}^{n\times k}$ die Spalten von B für $i = 1, ..., k$ gegeben durch

$$\bar{b}_j := (b_{ij})_{1\le j\le n} \in \mathbb{R}^n\ , \tag{27}$$

also

$$A^{\mathrm{T}} = \begin{pmatrix} \bar{a}_1 & \bar{a}_2 & ... & \bar{a}_m \end{pmatrix} \text{ und } B = \begin{pmatrix} \bar{b}_1 & \bar{b}_2 & ... & \bar{b}_k \end{pmatrix}\ . \tag{28}$$

Dann gilt

$$AB = \left(\langle \bar{a}_i, \bar{b}_j\rangle\right)_{1\le i\le m, 1\le j\le k} \tag{29}$$

Bemerkungen

- Der Eintrag $(AB)_{ij}$ entspricht dem Skalarprodukt von $i$ter Spalte von A^{T} und $j$ter Spalte von B.
- Die erste Aussage folgt mit der Identifikation von $\mathbb{R}^n = \mathbb{R}^{n\times 1}$.
- Wir verzichten auf einen ausführlichen Beweis.

Theorem (Summe der Quadrate aller Einträge eines Vektors)

Es sei $x \in \mathbb{R}^n$ ein n-dimensionaler Vektor. Dann kann die Summe der Quadrate aller Einträge dieses Vektors durch die Matrixmultiplikation des transponierten Vektors mit sich selbst berechnet werden:

$$\sum_{i=1}^{n} x_i^2 = x^{\mathrm{T}} x \; . \tag{30}$$

Bemerkungen

- Es handelt sich hierbei um einen Spezialfall der ersten Aussage des vorherigen Theorems.
- Im Rahmen des ALM ist das Theorem nützlich, um z.B. die residuelle Quadratsumme auszudrücken.

Beweis

$$x^{\mathrm{T}} x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1^2 + x_2^2 + \dots + x_n^2 = \sum_{i=1}^{n} x_i^2 \tag{31}$$

$\square$

Theorem (Matrixmultiplikation und Transposition)

Es seien $A \in \mathbb{R}^{m \times n}$ und $B \in \mathbb{R}^{n \times k}$. Dann gilt:

$$(AB)^{\mathrm{T}} = B^{\mathrm{T}} A^{\mathrm{T}}. \tag{32}$$

Beweis

$$\begin{aligned}
(AB)^{\mathrm{T}} &= \left(\left(\sum_{i=1}^{n} a_{ji} b_{il} \right)_{1 \le j \le m, 1 \le l \le k} \right)^{\mathrm{T}} \\
&= \left(\sum_{i=1}^{n} a_{ij} b_{li} \right)_{1 \le j \le m, 1 \le l \le k} \\
&= \left(\sum_{i=1}^{n} b_{li} a_{ij} \right)_{1 \le l \le k, 1 \le j \le m} \\
&= B^{\mathrm{T}} A^{\mathrm{T}}
\end{aligned} \tag{33}$$

$\square$

Motivation für Begriff der Inversen einer quadratischen Matrix

- Es seien $A \in \mathbb{R}^{n \times n}, x \in \mathbb{R}^n$ und $b \in \mathbb{R}^n$, A und b seien als bekannt vorausgesetzt, x sei unbekannt.
- Zum Beispiel sei $A := \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ und $b := \begin{pmatrix} 5 \\ 11 \end{pmatrix}$
- In diesem Fall gilt $Ax = b \quad \Leftrightarrow \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 11 \end{pmatrix} \quad \Leftrightarrow \quad \begin{matrix} 1x_1 + 2x_2 & = 5 \\ 3x_1 + 4x_2 & = 11 \end{matrix}$
- Wir haben also ein *lineares Gleichungssystem* (LGS) mit zwei Gleichungen und zwei Unbekannten.
- Wir stellen uns vor, dass wir wissen möchten, für welche(s) x das LGS erfüllt ist.
- Wären $A = a \in \mathbb{R}$, $x \in \mathbb{R}$ und $b \in \mathbb{R}$, also $ax = b$ gegeben, so würden wir mit dem *multiplikativem Inversen* von a multiplizieren, also dem Wert, der mit a multipliziert 1 ergibt und durch $a^{-1} = \frac{1}{a}$ gegeben ist.
- Dann würde nämlich gelten $ax = b \quad \Leftrightarrow \quad a^{-1}ax = a^{-1}b \quad \Leftrightarrow \quad 1 \cdot x = a^{-1}b \quad \Leftrightarrow \quad x = \frac{b}{a}$
- Konkret etwa $2x = 6 \quad \Leftrightarrow \quad 2^{-1}2x = 2^{-1}6 \quad \Leftrightarrow \quad \frac{1}{2}2x = \frac{1}{2}6 \quad \Leftrightarrow \quad x = 3$.
- Analog möchte man mit dem *multiplikativen Inversen* A^{-1} von A multiplizieren können, sodass "$A^{-1}A = 1$".
- Dann hätte man nämlich $Ax = b \quad \Leftrightarrow \quad A^{-1}Ax = A^{-1}b \quad \Leftrightarrow \quad x = A^{-1}b$
- Diese Idee des multiplikativen Inversen wird im folgenden als *Inverse eine quadratischen Matrix* formalisiert.

Definition (Einheitsmatrix)

Die Matrix A mit $a_{ij} = 1$ für $i = j$ und $a_{ij} = 0$ für $i \neq j$

$$I_n := (a_{ij})_{1 \le i \le n, 1 \le j \le n} \in \mathbb{R}^{n \times n} := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \tag{34}$$

heißt n*-dimensionale Einheitsmatrix*.

Bemerkungen

- Die Einheitsmatrix is eine quadratische Matrix, bei alle diagonalen Einträge 1 und alle anderen Einträge 0 sind.
- I_n wird in R mit dem Befehl `diag(n)` erzeugt.

Theorem (Neutrales Element der Matrixmultiplikation)

I_n ist das neutrale Element der Matrixmultiplikation, d.h. es gilt für $A \in \mathbb{R}^{n \times m}$, dass

$$I_n A = A \text{ und } A I_m = A. \tag{35}$$

Bemerkungen

- Ist die Matrixmultiplikation der Einheitsmatrix mit einer anderen Matrix definiert, so bleibt letztere dabei unverändert und die Einheitsmatrix "verschwindet" aus dem Produkt.

Beweis

Es sei $B = (b_{ij}) = I_n A \in \mathbb{R}^{n \times m}$. Dann gilt für alle $1 \le i \le n$ und alle $1 \le j \le n$

$$b_{ij} = 0 \cdot a_{1j} + 0 \cdot a_{2j} + \cdots + 0 \cdot a_{i-1,j} + 1 \cdot a_{ij} + +0 \cdot a_{i+1,j} + \cdots + 0 \cdot a_{nj} = a_{ij} \tag{36}$$

und analog für AI_m. $\square$

Definition (Invertierbare Matrix und inverse Matrix)

Eine quadratische Matrix $A \in \mathbb{R}^{n \times n}$ heißt *invertierbar*, wenn es eine quadratische Matrix $A^{-1} \in \mathbb{R}^{n \times n}$ gibt, sodass

$$A^{-1}A = AA^{-1} = I_n \tag{37}$$

ist. Die Matrix A^{-1} heißt die *inverse Matrix von* A.

Bemerkungen

- Invertierbarkeit und inverse Matrizen beziehen sich nur auf quadratische Matrizen.
- Die inverse Matrix von A heißt auch einfach *Inverse* von A.
- Quadratische Matrizen können, müssen aber nicht invertierbar sein.
- Nicht-invertierbare Matrizen nennt man *singuläre* Matrizen.
- Für $A = a \in \mathbb{R}^{1 \times 1}$ gilt $A^{-1} = \frac{1}{a}$.
- Die Definition sagt nur aus, was eine inverse Matrix ist, aber nicht, wie man sie berechnet.

Beispiel für eine invertierbare Matrix

Die Matrix $A = \begin{pmatrix} 2.0 & 1.0 \\ 3.0 & 4.0 \end{pmatrix}$ ist invertierbar mit inverser Matrix $A^{-1} = \begin{pmatrix} 0.8 & -0.2 \\ -0.6 & 0.4 \end{pmatrix}$, denn

$$\begin{pmatrix} 2.0 & 1.0 \\ 3.0 & 4.0 \end{pmatrix} \begin{pmatrix} 0.8 & -0.2 \\ -0.6 & 0.4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0.8 & -0.2 \\ -0.6 & 0.4 \end{pmatrix} \begin{pmatrix} 2.0 & 1.0 \\ 3.0 & 4.0 \end{pmatrix}, \tag{38}$$

wovon man sich durch Nachrechnen überzeugt.

Beispiel für eine nicht-invertierbare Matrix

Die Matrix $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ ist nicht invertierbar, denn wäre B invertierbar, dann gäbe es $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ mit

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{39}$$

Das würde aber bedeuten, dass $0 = 1$ in $\mathbb{R}$ und das ist ein Widerspruch. Also kann B nicht invertierbar sein.

Berechnen inverser Matrizen

- 2×2 bis etwa 5×5 Matrizen kann man prinzipiell per Hand invertieren.
- Dazu lernt man im Bachlor-Studium Mathematik verschiedene Verfahren.
- Wir verzichten auf eine Einführung in die Matrizeninvertierung per Hand.
- Ein kurzes (30 min) Erklärvideo (englisch) findet sich hier.
- In der Anwendung werden Matrizen standardmäßig numerisch invertiert.
- Matrixinversion ist ein weites Feld in der numerischen Mathematik.
- Es gibt sehr viele Algorithmen zur Invertierung invertierbarer Matrizen.
- In R berechnet man inverse Matrizen zum Beispiel mit dem Paket `matlib`.

Berechnen inverser Matrizen

```
# Definition
A = matrix(c(2,1,
             3,4),
           nrow  = 2,
           byrow = TRUE)

# Inverse Matrix
solve(A)                        # A^{-1}
```

```
>      [,1] [,2]
> [1,]  0.8 -0.2
> [2,] -0.6  0.4
```

Operationen

Berechnen inverser Matrizen

```
print(solve(A) %*% A)
```

```
>      [,1] [,2]
> [1,]    1    0
> [2,]    0    1
```

```
print(A %*% solve(A))
```

```
>      [,1]      [,2]
> [1,]    1 -5.55e-17
> [2,]    0  1.00e+00
```

```
# nicht-invertierbare Matrizen sind auch numerisch nicht-invertierbar (singulär)
B = matrix(c(1,0,
             0,0),
           nrow  = 2,
           byrow = 2)
solve(B)                        # B^{-1}
```

```
> Error in solve.default(B): Lapack routine dgesv: system is exactly singular: U[2,2] = 0
```

Definition

Operationen

Determinanten

Rang und Spur

Spezielle Matrizen

Selbstkontrollfragen

Definition (Determinante)

Für $A = (a_{ij})_{1\leq i,j\leq n} \in \mathbb{R}^{n\times n}$ mit $n > 1$ sei $A_{ij} \in \mathbb{R}^{(n-1)\times(n-1)}$ die Matrix, die aus A durch Entfernen der $i$ten Zeile und der $j$ten Spalte entsteht. Dann heißt die Zahl

$$\begin{aligned} |A| &:= a_{11} && \text{für } n = 1 \\ |A| &:= \textstyle\sum_{j=1}^{n} a_{1j}(-1)^{1+j}|A_{1j}| && \text{für } n > 1 \end{aligned} \tag{40}$$

die *Determinante* von A.

Bemerkungen

- Für

$$A := \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \tag{41}$$

  ergeben sich zum Beispiel

$$A_{11} = \begin{pmatrix} 5 & 6 \\ 8 & 9 \end{pmatrix}, A_{12} = \begin{pmatrix} 4 & 6 \\ 7 & 9 \end{pmatrix}, A_{21} = \begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix}, A_{22} = \begin{pmatrix} 1 & 3 \\ 7 & 9 \end{pmatrix} \tag{42}$$

- Determinanten sind nichtlineare Abbildungen der Form $|\cdot| : \mathbb{R}^{n\times n} \to \mathbb{R}, A \mapsto |A|$.

Theorem (Determinanten von 2×2 und 3×3 Matrizen)

(1) Es sei $A = (a_{ij})_{1 \le i,j \le 2} \in \mathbb{R}^{2\times 2}$. Dann gilt

$$|A| = a_{11}a_{22} - a_{12}a_{21}. \tag{43}$$

(2) Es sei $A = (a_{ij})_{1 \le i,j \le 3} \in \mathbb{R}^{3\times 3}$. Dann gilt

$$|A| = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31}. \tag{44}$$

Bemerkungen

- Die Determinante einer 1×1 Matrix ist identisch mit ihrem einzigen Eintrag.
- Für 2×2 und 3×3 Matrizen (und nur für diese) gilt die *Sarrusche Merkregel*:

  "Summe der Produkte auf den Diagonalen minus Summe der Produkte auf den Gegendiagonalen"

- Bei 3×3 Matrizen bezieht sich die Merkregel auf das Schema

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & | & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & | & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & | & a_{31} & a_{32} \end{pmatrix}. \tag{45}$$

Determinanten

Beweis

Für $A \in \mathbb{R}^{2\times 2}$ gilt nach Definition

$$
\begin{aligned}
|A| &= \sum_{j=1}^{n} a_{1j}(-1)^{1+j}|A_{1j}| \\
&= a_{11}(-1)^{1+1}|A_{11}| + a_{12}(-1)^{1+2}|A_{12}| \\
&= a_{11}|(a_{22})| - a_{12}|(a_{21})| \\
&= a_{11}a_{22} - a_{12}a_{21}.
\end{aligned} \tag{46}
$$

Für $A \in \mathbb{R}^{3\times 3}$ gilt nach Definition und mit der Formel für die Determinante von 2×2 Matrizen

$$
\begin{aligned}
|A| &= \sum_{j=1}^{n} a_{1j}(-1)^{1+j}|A_{1j}| \\
&= a_{11}(-1)^{1+1}|A_{1j}| + a_{12}(-1)^{1+2}|A_{12}| + a_{13}(-1)^{1+3}|A_{13}| \\
&= a_{11}|A_{11}| - a_{12}|A_{12}| + a_{13}|A_{13}| \\
&= a_{11}\left|\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}\right| - a_{12}\left|\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix}\right| + a_{13}\left|\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}\right| \\
&= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\
&= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \\
&= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31}.
\end{aligned} \tag{47}
$$

$\square$

Determinanten

Beispiel 1

Es seien

$$A := \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} \text{ und } B := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} . \tag{48}$$

Dann ergeben sich

$$|A| = 2 \cdot 4 - 1 \cdot 3 = 8 - 3 = 5 \tag{49}$$

und

$$|B| = 1 \cdot 0 - 0 \cdot 0 = 0 - 0 = 0 . \tag{50}$$

Beispiel 2

Es sei

$$C := \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix} . \tag{51}$$

Dann ergibt sich

$$|C| = 2 \cdot 1 \cdot 3 + 0 \cdot 0 \cdot 0 + 0 \cdot 0 \cdot 0 - 0 \cdot 0 \cdot 3 - 0 \cdot 0 \cdot 0 - 0 \cdot 1 \cdot 0 = 2 \cdot 1 \cdot 3 = 6 . \tag{52}$$

Determinanten

```r
# Beispiel 1
A = matrix(c(2,1,                     # Matrixdefinition
             3,4),
           nrow = 2,
           byrow = TRUE)
det(A)                                # Determinantenberechnung
```

```
> [1] 5
```

```r
# Beispiel 1
B = matrix(c(1,0,                     # Matrixdefinition
             0,0),
           nrow = 2,
           byrow = TRUE)
det(B)                                # Determinantenberechnung
```

```
> [1] 0
```

```r
# Beispiel 2
C = matrix(c(2,0,0,                   # Matrixdefinition
             0,1,0,
             0,0,3),
           nrow = 3,
           byrow = TRUE)
det(C)                                # Determinantenberechnung
```

```
> [1] 6
```

Theorem (Rechenregeln für Determinanten)

(*Determinantenmultiplikationssatz*) Für $A, B \in \mathbb{R}^{n \times n}$ gilt

$$|AB| = |A||B|. \tag{53}$$

(*Transposition*) Für $A \in \mathbb{R}^{n \times n}$ gilt

$$|A| = \left|A^{\mathrm{T}}\right|. \tag{54}$$

(*Inversion*) Für eine invertierbare Matrix $A \in \mathbb{R}^{n \times n}$ gilt

$$\left|A^{-1}\right| = \frac{1}{|A|} \tag{55}$$

(*Dreiecksmatrizen*) Für Dreiecksmatrizen, d.h. eine Matrix $A = (a_{ij})_{1 \le i,j \le n} \in \mathbb{R}^{n \times n}$ mit $a_{ij} = 0$ für $i > j$ oder $a_{ij} = 0$ für $j > i$, ist die Determinante gleich dem Produkt der Diagonalelemente:

$$|A| = \prod_{i=1}^{n} a_{ii} \tag{56}$$

Bemerkungen

- Wir verzichten auf einen Beweis.
- Bei Dreiecksmatrizen sind alle Elemente unterhalb ($i > j$) oder oberhalb ($j > i$) der Diagonalen 0.
- Bei I_n sind alle nicht-diagonalen Elemente 0 und alle diagonalen Elemente 1, also folgt $|I_n| = 1$.

Theorem (Invertierbarkeit und Determinante)

$A \in \mathbb{R}^{n\times n}$ ist dann und nur dann invertierbar, wenn $|A| \neq 0$ ist. Es gilt also

$$A \text{ ist invertierbar} \Leftrightarrow |A| \neq 0 \text{ und} A \text{ ist nicht invertierbar} \Leftrightarrow |A| = 0. \tag{57}$$

Beweisandeutung

Wir zeigen lediglich, dass aus der Invertierbarkeit von A folgt, dass $|A|$ nicht null sein kann. Nehmen wir also an, dass A invertierbar ist. Dann gibt es eine Matrix B mit $AB = I_n$ und mit dem Determinantenmultiplikationssatz folgt

$$|AB| = |A||B| = |I_n| = 1. \tag{58}$$

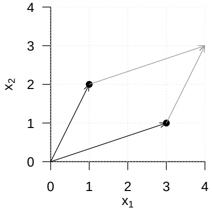
Also kann nicht $|A| = 0$ gelten, denn sonst wäre $0 = 1$. □

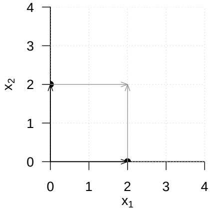
Visuelle Intuition

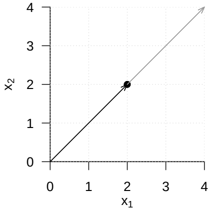
$a_1, ..., a_n \in \mathbb{R}^n$ seien die Spalten von $A \in \mathbb{R}^{n \times n}$.

$\Rightarrow |A|$ entspricht dem signierten Volumen des von $a_1, ..., a_n \in \mathbb{R}^n$ aufgespannten Parallelotops.

$$A_1 = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \qquad A_3 = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix}$$

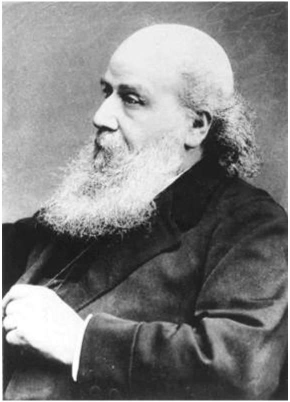






$$|A_1| = 3 \cdot 2 - 1 \cdot 1 = 5 \qquad |A_2| = 2 \cdot 2 - 0 \cdot 0 = 4 \qquad |A_3| = 2 \cdot 2 - 2 \cdot 2 = 0$$

Anekdote: Herkunft des Ausdrucks "Matrix"



James Joseph Sylvester (1814 – 1897)

(Quelle: *Wikimedia Commons*: "James_Joseph_Sylvester.jpg"; Lizenz: gemeinfrei.)

Definition

Operationen

Determinanten

Rang und Spur

Spezielle Matrizen

Selbstkontrollfragen

Rang und Spur

Überblick

- Der Rang einer Matrix ist eine Zahl, an der bestimmte Eigenschaften der Matrix abgelesen werden können.
- In dieser Hinsicht ist der Rang einer Matrix sehr ähnlich zur Determinante einer Matrix.
- Viele Resultate in der linearen Algebra beruhen auf Annahmen über den Rang einer Matrix.
- Wir werden z.B. durchgängig annehmen, dass die Designmatrix des ALMs *vollen Spaltenrang* hat.
- Der Rang einer Matrix ist ein tiefgehendes Konzept, das wir hier nur oberflächlich behandeln können.
- Für ausführlichere Einführungen, siehe z.B. Searle (1982), Kapitel 6 und Strang (2009), Kapitel 3.2.

Definition (Spaltenraum einer Matrix)

$X \in \mathbb{R}^{n\times p}$ sei eine Matrix mit $n > p$ und es seien $x_1, ..., x_p \in \mathbb{R}^n$ die Spalten von X. Dann wird die lineare Hülle der $x_1, ..., x_p \in \mathbb{R}^n$ *Spaltenraum von* $X \in \mathbb{R}^{n\times p}$ genannt und mit $\mathcal{S}(X)$ bezeichnet. Insbesondere gilt

$$\mathcal{S}(X) = \{X\beta \in \mathbb{R}^n \mid \beta \in \mathbb{R}^p\}. \tag{59}$$

Bemerkungen

- Ohne Beweis halten wir fest, dass $\mathcal{S}(X)$ ein Untervektorraum von $\mathbb{R}^n$ ist.
- Die lineare Hülle der $x_1, ..., x_p \in \mathbb{R}^n$ ist die Menge aller Linearkombinationen $x_1, ..., x_p \in \mathbb{R}^n$.
- Die Menge aller Linearkombinationen der $x_1, ..., x_p \in \mathbb{R}^n$ hat die Form

$$\left\{\sum_{i=1}^p \beta_i x_i \,\middle|\, x_i \in \mathbb{R}^n, \beta_i \in \mathbb{R}\right\}. \tag{60}$$

- Die Form des Spaltenraums in der obigen Definition ergibt sich für $x_{ij} \in \mathbb{R}, 1 \le i \le n, 1 \le j \le p$ durch

$$X\beta = \begin{pmatrix} \sum_{j=1}^p x_{1j}\beta_j \\ \vdots \\ \sum_{j=1}^p x_{nj}\beta_j \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^p \beta_j x_{1j} \\ \vdots \\ \sum_{j=1}^p \beta_j x_{nj} \end{pmatrix} = \beta_1 \begin{pmatrix} x_{11} \\ \vdots \\ x_{n1} \end{pmatrix} + \cdots + \beta_p \begin{pmatrix} x_{1p} \\ \vdots \\ x_{np} \end{pmatrix} = \sum_{i=1}^p \beta_i x_i.$$

Definition (Rang einer Matrix)

$X \in \mathbb{R}^{n \times p}$ sei eine Matrix. Dann wird die Dimension des Spaltenraums von X der *Rang von X* genannt und mit rg(X) bezeichnet. Gilt dabei $\text{rg}(X) = p$, so sagt man, dass X *vollen Spaltenrang* hat.

Bemerkungen

- Ohne Beweis halten wir untenstehend einige wichtige Eigenschaften des Rangs einer Matrix fest.
- Der Rang einer Matrix entspricht der Anzahl der linear unabhängigen Spaltenvektoren der Matrix.
- Es gilt: $\text{rg}(X) = \text{rg}(X^{\text{T}})$.
- Für $A \in \mathbb{R}^{p \times p}$ gilt: $\text{rg}(A) = p \Leftrightarrow A$ ist invertierbar.
- Für $X \in \mathbb{R}^{n \times p}$ gilt: $\text{rg}(X) = \text{rg}(X^{\text{T}}X) = \text{rg}(XX^{\text{T}})$.
- Für $X \in \mathbb{R}^{n \times p}$ gilt damit: $\text{rg}(X) = p \Leftrightarrow X^{\text{T}}X \in \mathbb{R}^{p \times p}$ ist invertierbar.
- Die Funktion `Rank()` aus dem R-Paket `pracma` erlaubt die Berechnung des Rangs einer Matrix.
- Das Spaltenraum- und Rangkonzept einer Matrix erlauben einen geometrischen Zugang zur Theorie des ALMs.
- Wir verfolgen diesen geometrischen Zugang hier nicht. Für Details, siehe z.B. Christensen (2011).

Beispiel zu linearer Unabhängigkeit

Die Vektoren

$$a_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},\ a_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},\ a_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{61}$$

sind linear unabhängig, weil sich keiner der Vektoren als Linearkombination der anderen Vektoren darstellen lässt.

Die Vektoren

$$b_1 = \begin{pmatrix} 2 \\ -1 \\ 1 \\ -2 \end{pmatrix},\ b_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix},\ b_3 = \begin{pmatrix} 4 \\ -1 \\ 3 \\ -2 \end{pmatrix} \tag{62}$$

sind hingegen linear abhängig, da sie sich gegenseitig als Linearkombinationen darstellen lassen. So ist z.B. der dritte Vektor die Summe des ersten Vektors und des mit 2 multiplizierten zweiten Vektors:

$$b_1 + 2 \cdot b_2 = \begin{pmatrix} 2 \\ -1 \\ 1 \\ -2 \end{pmatrix} + 2 \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 + 2 \cdot 1 \\ -1 + 2 \cdot 0 \\ 1 + 2 \cdot 1 \\ -2 + 2 \cdot 0 \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \\ 3 \\ -2 \end{pmatrix} = b_3\ . \tag{63}$$

Definition (Spur einer Matrix)

$A \in \mathbb{R}^{n\times n}$ sei eine quadratische Matrix. Dann ist die *Spur von* A die Summe der Elemente auf der Hauptdiagonale dieser Matrix und wird mit sp(A) bezeichnet:

$$\mathrm{sp}(A) = \sum_{i=1}^{n} a_{ii} \,. \tag{64}$$

Bemerkungen

- Ohne Beweis halten wir untenstehend einige wichtige Eigenschaften der Spur einer Matrix fest.
- Es gilt: $\mathrm{sp}(A) = \mathrm{sp}(A^{\mathrm{T}})$.
- Für $A, B \in \mathbb{R}^{n\times n}$ gilt: $\mathrm{sp}(A+B) = \mathrm{sp}(A) + \mathrm{sp}(B)$.
- Für $A, B, C \in \mathbb{R}^{n\times n}$ gilt: $\mathrm{sp}(ABC) = \mathrm{sp}(CAB) = \mathrm{sp}(BCA)$.
- Für $A \in \mathbb{R}^{n\times n}$ gilt: $AA = A \Rightarrow \mathrm{sp}(A) = \mathrm{rg}(A)$.
- Die Spur der Identitätsmatrix ist: $\mathrm{sp}(I_n) = n$.
- In R kann die Spur einer Matrix mit `sum(diag(A))` berechnet werden.
- Der Begriff der Spur ist beim Beweis der Verteilung der residuellen Quadratsumme im ALM von Bedeutung.

Definition

Operationen

Determinanten

Rang und Spur

Spezielle Matrizen

Selbstkontrollfragen

Definition (Einheitsmatrizen und Einheitsvektoren)

- Wir bezeichnen die *Einheitsmatrix* als

$$I_n := (i_{jk})_{1\le j\le n, 1\le k\le n} \in \mathbb{R}^{n\times n} \text{ mit } i_{jk} = 1 \text{ für } j = k \text{ und } i_{jk} = 0 \text{ für } j \neq k. \tag{65}$$

- Wir bezeichnen die *Einheitsvektoren* $e_i, i = 1, ..., n$ als

$$e_i := (e_{i_j})_{1\le j\le n} \in \mathbb{R}^n \text{ mit } e_{i_j} = 1 \text{ für } i = j \text{ und } e_{i_j} = 0 \text{ für } i \neq j. \tag{66}$$

Bemerkungen

- I_n besteht nur aus Nullen und Diagonalelementen gleich Eins.
- $e_i, i = 1, ..., n$ besteht nur aus Nullen und einer Eins in der $i$ten Komponente.
- Es gilt

$$I_n = \begin{pmatrix} e_1 & e_2 & ... & e_n \end{pmatrix} \tag{67}$$

- Es gilt auch

$$I_n = \left(e_i^{\mathrm{T}} e_j\right)_{1\le j\le n, 1\le k\le n} \tag{68}$$

- Es gelten weiterhin für $1 \le i, j \le n$

$$e_i^{\mathrm{T}} e_j = 0 \text{ für } i \neq j \text{ sowie } e_i^{\mathrm{T}} e_i = 1 \quad \text{und} \quad e_i^{\mathrm{T}} v = v^{\mathrm{T}} e_i = v_i \text{ für } v \in \mathbb{R}^n. \tag{69}$$

Definition (Diagonalmatrix)

Eine Matrix $D \in \mathbb{R}^{n \times m}$ heißt *Diagonalmatrix*, wenn $d_{ij} = 0$ für $1 \le i \le n, 1 \le j \le m$ mit $i \ne j$.

Bemerkungen

- Eine Diagonalmatrix $D \in \mathbb{R}^{n \times n}$ mit Diagonalelementen $d_1, ..., d_n$ schreibt man auch als

$$D = \operatorname{diag}(d_1, ..., d_n). \tag{70}$$

- Diagonalmatrizen haben viele "gute" Eigenschaften.
- Zum Beispiel überzeugt man sich leicht davon, dass Multiplikation einer Matrix A von links mit einer Diagonalmatrix D der Multiplikation der Zeilen der Matrix A mit den entsprechenden Diagonaleinträgen von D entspricht. Die Multiplikation von rechts entspricht der Multiplikation der Spalten von A mit den entsprechenden Diagonaleinträgen von D.
- Eine weitere wichtige Eigenschaft ist

$$D := \operatorname{diag}(d_1, ..., d_n) \Rightarrow |D| = \prod_{i=1}^{n} d_i \tag{71}$$

- Für Beweise dieser Eigenschaften wird auf die einschlägige Literatur, z.B. Searle (1982), verwiesen.

Definition (Symmetrische Matrix)

Eine Matrix $S \in \mathbb{R}^{n \times n}$ heißt *symmetrisch*, wenn gilt dass $S^{\mathrm{T}} = S$.

Bemerkungen

- Eine symmetrische Matrix ist eine Matrix, deren Transponiertes gleich ihr selbst ist.
- Symmetrische Matrizen haben viele "gute" Eigenschaften.
- Beispielweise gilt für die Summe zweier symmetrischer Matrizen, dass auch diese wieder symmetrisch ist

$$A = A^{\mathrm{T}} \text{ und } B = B^{\mathrm{T}} \Rightarrow A + B = (A + B)^{\mathrm{T}} \tag{72}$$

  und dass die Inverse einer symmetrischen Matrix, sofern sie existiert, auch symmetrisch ist,

$$S^{\mathrm{T}} = S \Rightarrow \left(S^{-1}\right)^{\mathrm{T}} = S^{-1}. \tag{73}$$

- Für Beweise dieser Eigenschaften wird auf die einschlägige Literatur, z.B. Searle (1982), verwiesen.

Definition (Positiv-definite Matrix)

Eine quadratische Matrix $C \in \mathbb{R}^{n\times n}$ heißt positiv-definit (p.d.), wenn

- C eine symmetrische Matrix ist und
- für alle $x \in \mathbb{R}^n, x \neq 0_n$ gilt, dass $x^{\mathrm{T}}Cx > 0$ ist.

Bemerkungen

- Positiv-definite Matrizen sind für die Definition der multivariaten Normalverteilungen grundlegend.
- Positiv-definite Matrizen haben viele "gute" Eigenschaften.
- Beispielsweise gilt

$$C \text{ ist positiv-definit } \Rightarrow C^{-1} \text{ existiert und ist ebenfalls positiv-definit.} \tag{74}$$

- Für Beweise dieser Eigenschaften wird auf die einschlägige Literatur, z.B. Searle (1982), verwiesen.

Selbstkontrollfragen

1. Geben Sie die Definition einer Matrix wieder.
2. Nennen Sie sechs Matrixoperationen.
3. Geben Sie die Definitionen der Matrixaddition und -subtraktion wieder.
4. Geben Sie die Definition der Skalarmultiplikation für Matrizen wieder.
5. Geben Sie die Definition der Matrixtransposition wieder.
6. Es seien
$$A := \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix},\ B := \begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix} \text{ und } c := 2 \tag{75}$$
Berechnen Sie
$$D := c\left(A - B^{\mathrm{T}}\right) \text{ und } E := (cA)^{\mathrm{T}} + B \tag{76}$$
per Hand.
7. Geben Sie die Definition der Matrixmultiplikation wieder.
8. Es seien $A \in \mathbb{R}^{3\times 2}, B \in \mathbb{R}^{2\times 4}$ und $C \in \mathbb{R}^{3\times 4}$. Prüfen Sie, ob folgende Matrixprodukte definiert sind, und wenn ja, geben Sie die Größe der resultierenden Matrix an:
$$ABC, \qquad ABC^{\mathrm{T}}, \qquad A^{\mathrm{T}}CB^{\mathrm{T}}, \qquad BAC. \tag{77}$$

9. Es seien
$$A := \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 3 & 2 & 0 \end{pmatrix}, \; B := \begin{pmatrix} 1 & 2 & 2 \\ 1 & 3 & 1 \\ 2 & 0 & 0 \end{pmatrix} \text{ und } C := \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}. \tag{78}$$
Berechnen Sie die Matrixprodukte
$$AB, \quad B^{\mathrm{T}}A^{\mathrm{T}}, \quad \left(B^{\mathrm{T}}A^{\mathrm{T}}\right)^{\mathrm{T}}, \quad AC \tag{79}$$
per Hand.
10. Definieren Sie die Begriffe der invertierbaren Matrix und der inversen Matrix.
11. Geben Sie die Definition von Einheitsmatrix und Einheitsvektoren wieder.
12. Geben Sie die Definition von Einsmatrizen und Nullmatrizen wieder.
13. Geben Sie die Definition einer symmetrischen Matrix wieder.
14. Geben Sie die Definition einer Diagonalmatrix wieder.
15. Geben Sie die Definition einer positiv-definiten Matrix wieder.

Referenzen

Christensen, Ronald. 2011. *Plane Answers to Complex Questions*. Springer Texts in Statistics. New York, NY: Springer New York. https://doi.org/10.1007/978-1-4419-9816-3.

Searle, Shayle. 1982. *Matrix Algebra Useful for Statistics*. Wiley-Interscience.

Strang, Gilbert. 2009. *Introduction to Linear Algebra*. Cambridge University Press.