

Numerisches Rechnen und Lineare Algebra

Ao.Univ.-Prof. Dr. P. Berglez

Institut für Analysis und Zahlentheorie
Technische Universität Graz

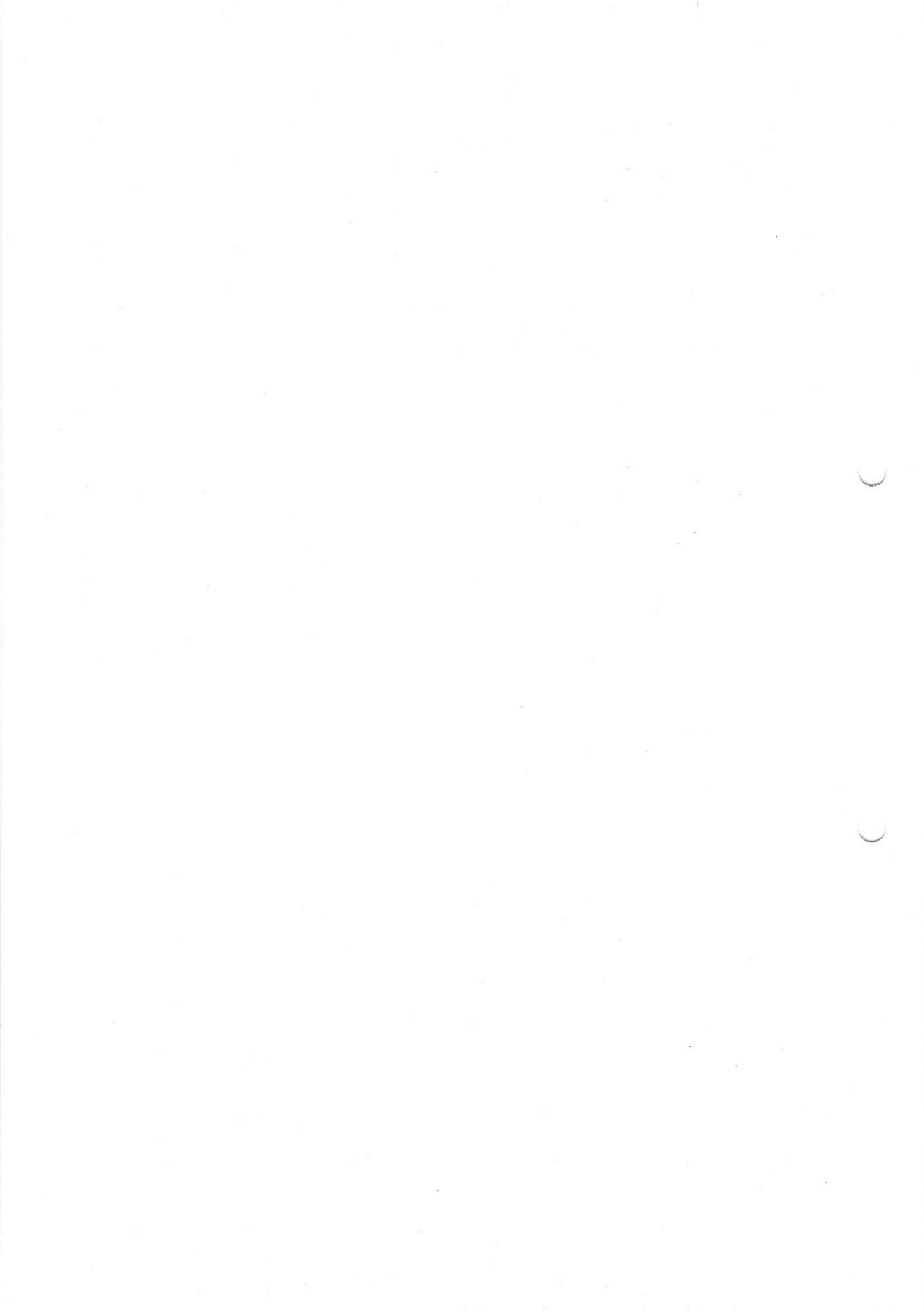
© 2016 P. Berglez
2017 Skript unverändert
Vorlesung 2017 gelesen von C. Elsholtz und A. Rosenmann

Inhaltsverzeichnis

0 Grundlagen	1
I Lineare Algebra	5
1 Lineare Gleichungssysteme – GAUSS'sche Elimination	6
1.1 Einführung	6
1.2 Matrizen	7
1.2.1 Definitionen	7
1.2.2 Rechenoperationen	8
1.2.3 Die Transponierte, die Inverse	10
1.2.4 Anwendung der Inversen bei linearen Gleichungssystemen	12
1.3 GAUSS'sche Elimination	13
1.3.1 Einführung - GAUSS'scher Algorithmus	13
1.3.2 GAUSS - JORDAN - Elimination	16
1.3.3 Frage der Existenz von Lösungen	17
1.3.4 Matrizen von Zeilenstufenform	18
1.3.5 Existenz und Struktur der Lösungen	19
1.3.6 Bestimmung der inversen Matrix	20
1.4 LR-Zerlegung einer regulären Matrix	22
1.4.1 LR-Zerlegung ohne und mit Zeilenvertauschung	22
1.4.2 Anwendungen	24
1.5 Determinanten	25
2 Vektoren im anschaulichen Raum	28
2.1 Einführung	28
2.2 Das Skalarprodukt	30
2.3 Das Vektorprodukt	31
2.4 Das Spatprodukt	32
3 Allgemeine Vektorräume	34
3.1 Grundlagen	34
3.2 Lineare Unabhängigkeit	38
3.3 Basis von Vektorräumen	40
4 Lineare Abbildungen	42
4.1 Einleitung	42
4.2 Kern und Bild einer linearen Abbildung	44
4.3 Summe, Vielfaches und Verknüpfung linearer Abbildungen	45
4.4 Injektive, surjektive und bijektive Abbildungen	45
4.5 Konstruktion linearer Abbildungen mittels Basen	46
4.6 Koordinaten, Koordinaten-Abbildung	48

4.7 Basiswechsel - Koordinatentransformation	49
4.8 Die Matrix einer linearen Abbildung	51
4.9 Spezielle lineare Abbildungen im \mathbb{R}^2	55
4.10 Spezielle lineare Abbildungen im \mathbb{R}^3	59
4.11 Affine Abbildungen im \mathbb{R}^2 und homogene Koordinaten	60
5 Unitäre Räume	63
5.1 Längen und Abstände	63
5.2 Das Skalarprodukt	64
5.3 Die SCHWARZ'sche Ungleichung	65
5.4 Winkelmessung, Orthonormierung nach GRAM-SCHMIDT	66
5.5 QR-Zerlegung	69
6 Eigenwerte und Eigenvektoren, Diagonalisierung	72
6.1 Eigenwerte und Eigenvektoren	72
6.2 Ähnliche Matrizen und Diagonalisierung	76
6.3 Symmetrische Matrizen und orthogonale Diagonalisierung	78
6.4 Singulärwertzerlegung	81
6.4.1 Singulärwerte einer Matrix	82
6.4.2 Singulärwertzerlegung einer Matrix	82
6.4.3 Anwendungen der Singulärwertzerlegung	84
II Numerische Methoden	87
7 Numerische Behandlung von linearen Gleichungssystemen	88
7.1 Einleitung	88
7.2 GAUSS'scher Algorithmus – LR–Zerlegung	88
7.3 CHOLESKY–Zerlegung	90
7.4 Anwendungen	91
7.5 Rundungsfehler, Pivotstrategien	97
7.6 Auswahl geeigneter Pivots	98
7.7 Skalierung der Gleichungen	100
7.8 Iterative Methoden	101
7.8.1 JACOBI–Iteration	102
7.8.2 GAUSS–SEIDEL–Iteration	103
7.8.3 Zusammenfassung	104
7.8.4 SOR–Methode	105
7.9 Überbestimmte lineare Gleichungssysteme	107
7.9.1 Bestimmung von Näherungslösungen	107
7.9.2 Die Pseudoinverse	108
7.9.3 Näherungslösung mittels QR–Zerlegung	109
7.9.4 Näherungslösung mittels Singulärwertzerlegung	110
8 Näherungsweise Bestimmung von Nullstellen – NEWTON–Iteration	112
8.1 Der eindimensionale Fall	112
8.2 NEWTONSches Verfahren für Systeme von Gleichungen	114
9 Numerische Bestimmung von Eigenwerten	117
9.1 Rückblick und Problemstellung	117
9.2 Bestimmung der Eigenwerte durch Iteration – Potenzmethode	118
9.3 Erweiterte Potenzmethode	120
9.4 GERSCHGORIN–Kreise	121

9.5 Approximation von Eigenwerten mittels QR -Zerlegung	124
10 Interpolation und Polynomapproximation	126
10.1 Einleitung	126
10.2 LAGRANGE'sche Polynome	127
10.3 Stückweise Interpolation mittels kubischer Splines	129
11 Approximationstheorie	135
11.1 Methode der kleinsten Quadrate	135
11.2 Gleichmäßige Approximation	140
11.3 Diskrete Approximation – Schnelle FOURIER-Transformation	142
11.3.1 Approximation durch trigonometrische Polynome	142
11.3.2 Interpolation durch trigonometrische Polynome - FFT	144
Literatur	149



Kapitel 0

Grundlagen

Hier sollen zunächst einige – eventuell schon aus der Schule bekannte – Instrumente und Sprechweisen der Mathematik zusammengestellt werden, die in dieser Lehrveranstaltung verwendet werden.

Sprechweisen und Symbole der Mathematik

In der Mathematik werden Aussagen formuliert und auf ihren Wahrheitsgehalt hin untersucht. Unter einer **Aussage** stellen wir uns hierbei vereinfacht einen feststellenden Satz vor, dem eindeutig einer der beiden Wahrheitswerte FALSCH oder WAHR zugeordnet werden kann. Als Beispiele dienen

$$Es regnet \quad \text{oder} \quad \sqrt{2} > 1.12$$

Junktoren

Mit Junktoren werden einfache Aussagen zu einer komplexen Aussage verknüpft. Wir betrachten die fünf (wichtigsten) Junktoren

NICHT UND ODER Implikation Äquivalenz

Junktoren

- Ist A eine Aussage, so ist $\neg A$ die **Negation** von A .
- Sind A und B Aussagen, so kann man $A \wedge B$ betrachten; man nennt “ \wedge ” den **UND-Junktor**. Es gilt
 - $A \wedge B$ ist wahr, wenn beide Aussagen erfüllt sind
 - $A \wedge B$ ist falsch, wenn eine der beiden Aussagen falsch ist
- Sind A und B Aussagen, so kann man $A \vee B$ betrachten; man nennt “ \vee ” den **ODER-Junktor**. Es gilt
 - $A \vee B$ ist wahr, wenn eine der beiden Aussagen erfüllt ist
 - $A \vee B$ ist falsch, wenn beide Aussagen falsch sind
- Wenn A gilt, **dann** gilt auch B , kurz $A \Rightarrow B$. Man nennt “ \Rightarrow ” **Implikation**
- **Genau dann** gilt A , wenn B gilt, kurz ($A \Rightarrow B$ und $B \Rightarrow A$), noch kürzer $A \Leftrightarrow B$. Man nennt “ \Leftrightarrow ” **Äquivalenz**.

Bemerkung.

Das ODER ist nicht ausschließend – es dürfen auch beide Aussagen erfüllt sein.
Ausschließend ist ENTWEDER-ODER.

Beispiel.

1. $\neg(\text{Heute regnet es})$ heißt: Heute regnet es nicht.
2. $\neg(x \geq 5)$ heißt: $x < 5$
3. $\neg(\text{Für alle } x, y \in M \text{ gilt } f(x+y) = f(x)+f(y))$ heißt: Es gibt $x, y \in M$ mit $f(x+y) \neq f(x)+f(y)$
4. Sind $A : x \leq 5$, $B : x \in \mathbb{N}$, so heißt $A \wedge B : x \in \{1, 2, 3, 4, 5\}$
5. Sind $A : x \in \mathbb{R} \wedge x \geq 2 \wedge x \leq 4$, $B : x \in \{2, 3, 7\}$, so heißt $A \vee B : x \in [2, 4] \cup \{7\}$
6. **Wenn** es regnet, **dann** ist die Straße nass; kurz

$$\text{Es regnet} \Rightarrow \text{Die Straße ist nass}$$

7. **Wenn** m eine gerade Zahl ist, **dann** ist $m \cdot n$ für jedes $n \in \mathbb{N}$ eine gerade Zahl; kurz

$$m \text{ gerade} \Rightarrow m \cdot n \text{ gerade } (n \in \mathbb{N})$$

Begründung. m gerade $\Rightarrow m = 2 \cdot m'$, $m' \in \mathbb{N} \Rightarrow m \cdot n = 2 \cdot m' \cdot n$, $m', n \in \mathbb{N} \Rightarrow m \cdot n$ gerade

8. Für $x \in \mathbb{R}$ gilt

$$x \leq 5 \wedge x \in \mathbb{N} \Leftrightarrow x \in \{1, 2, 3, 4, 5\}$$

9. Für $x \in \mathbb{R}$ gilt

$$x \in \mathbb{Q} \Leftrightarrow (\text{es gibt ein } n \in \mathbb{N} \text{ mit } n \cdot x \in \mathbb{Z})$$

Begründung. $x \in \mathbb{Q} \Leftrightarrow x = \frac{p}{q}$ mit $p \in \mathbb{Z}$ und $q \in \mathbb{N} \Leftrightarrow q \cdot x = p$ mit $p \in \mathbb{Z}$ und $q \in \mathbb{N}$. Wähle etwa $n = q$.

10. Sind $m, n \in \mathbb{N}$, so gilt

$$m \cdot n \text{ ist gerade} \Leftrightarrow (m \text{ ist gerade}) \vee (n \text{ ist gerade})$$

Begründung.

“ \Leftarrow ”: m gerade oder n gerade $\Rightarrow m \cdot n$ gerade (siehe oben)

“ \Rightarrow ”: $m \cdot n$ gerade; angenommen, weder m noch n sind gerade. Dann gilt

$$m = 2 \cdot m' + 1 \quad \text{und} \quad n = 2 \cdot n' + 1$$

mit $m', n' \in \mathbb{N}$. Es folgt $m \cdot n = 4 \cdot m' \cdot n' + 2 \cdot (m' + n') + 1 = 2 \cdot k + 1$ mit einem $k \in \mathbb{N}$. Das ist im Widerspruch zu: $m \cdot n$ gerade.

Quantoren

Quantoren erfassen Variablen mengenmäßig. Wir betrachten vier Quantoren

Quantoren

- \forall : "zu jedem" bzw. "für alle"
- \exists : "es gibt"
- \exists_1 : "es gibt genau ein"
- \nexists : "es gibt kein"

Beispiel.

1. Die Aussage "Zu jeder reellen Zahl x gibt es eine natürliche Zahl n , die größer als x ist" kann man kurz schreiben als

$$\forall x \in \mathbb{R} \quad \exists n \in \mathbb{N} : n > x$$

2. Sind $A = \{1, 2, 3\}$ und $B = \{1, 4, 9\}$, so gilt

$$\forall b \in B \quad \exists_1 a \in A : a^2 = b$$

Summen- und Produktzeichen

Das Summenzeichen Σ und das Produktzeichen Π sind nützliche Abkürzungen; man setzt

$$a_1 + a_2 + \cdots + a_n = \sum_{i=1}^n a_i \quad \text{und} \quad a_1 \cdot a_2 \cdot \cdots \cdot a_n = \prod_{i=1}^n a_i$$

Beispiel.

$$1. \sum_{i=1}^{100} 2^i = 2 + 2^2 + 2^3 + \cdots + 2^{100}$$

$$2. \prod_{i=1}^{100} \frac{1}{i^2} = 1 \cdot \frac{1}{4} \cdot \frac{1}{9} \cdots \frac{1}{10000}$$

$$3. \sum_{i=0}^n a_k = a_0 + \sum_{k=1}^{n-1} a_k + a_n$$

Gelegentlich braucht man auch die **leere Summe** bzw. das **leere Produkt** und meint damit, dass die obere Grenze (der Summation bzw. der Produktbildung) kleiner ist als die untere. Man definiert die leere Summe als 0, das leere Produkt als 1, z.B.

$$\sum_{k=1}^0 a_k = 0 \quad \text{und} \quad \prod_{k=2}^{-1} b_k = 1$$

Symbole der Mengenlehre

Unter einer **Menge** verstehen wir eine Zusammenfassung bestimmter, wohlunterschiedener Objekte, die wir **Elemente** der Menge nennen:

$$A = \{a, b, c, \dots\}$$

↑ ↑↑↑
Menge Elemente

Es gibt zwei verschiedene Arten Mengen anzugeben:

- Man kann Mengen beschreiben, indem man die Elemente explizit angibt

$$A = \{a, b, c\} \quad \text{oder} \quad \mathbb{N} = \{1, 2, 3, \dots\}$$

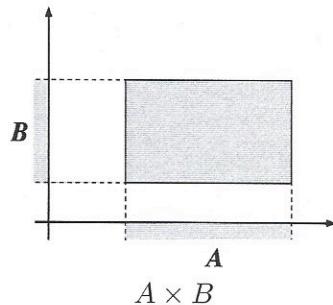
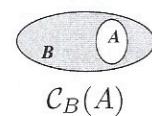
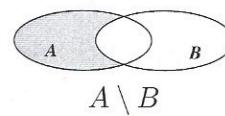
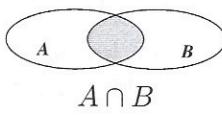
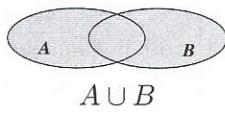
- Man kann die Menge durch die Eigenschaften der Elemente kennzeichnen

$$A = \{n \in \mathbb{N} \mid 1 \leq n \leq 5\} \quad \text{oder} \quad B = \{n \in \mathbb{N} \mid 2^n + 1 \text{ ist eine Primzahl}\}$$

Der senkrechte Strich leitet die Bedingung ein, die die Elemente erfüllen müssen, und wird gelesen als "für die gilt" oder "mit der Eigenschaft".

Begriffe und Notationen zu Mengen

- $a \in A$: a ist Element von A
- $a \notin A$: a ist kein Element von A
- $A \subseteq B$: A ist Teilmenge von B , d.h.: $a \in A \Rightarrow a \in B$
- $A \not\subseteq B$: A ist keine Teilmenge von B , d.h.: $\exists a \in A : a \notin B$
- $A = B$: A ist gleich B : $A \subseteq B \wedge B \subseteq A$
- \emptyset bezeichnet die leere Menge, das ist eine Menge ohne Elemente,
z.B.: $\emptyset = \{n \in \mathbb{N} \mid n < -1\}$
- $A \cap B = \{x \mid x \in A \wedge x \in B\}$ bezeichnet den Durchschnitt von A und B
- $A \cup B = \{x \mid x \in A \vee x \in B\}$ bezeichnet die Vereinigung von A und B
- $A \setminus B = \{x \mid x \in A \wedge x \notin B\}$ bezeichnet die Mengendifferenz von A und B
- $C_B(A) = B \setminus A$, falls $A \subseteq B$ bezeichnet das Komplement von A in B
- $A \times B = \{(a, b) \mid a \in A \wedge b \in B\}$ bezeichnet das kartesische Produkt von A und B



Teil I

Lineare Algebra

Kapitel 1

Lineare Gleichungssysteme – GAUSS'sche Elimination

1.1 Einführung

Beispiel 1. Wir betrachten das folgende System von 3 linearen Gleichungen für die 3 Unbestimmten u, v und w :

$$\begin{array}{rcl} 2u & + & v & + & w = 1 \\ 4u & + & v & & = -2 \\ -2u & + & 2v & + & w = 7 \end{array}$$

Zur Ermittlung der Lösung (= Berechnung der Unbestimmten) subtrahiert man ein Vielfaches einer Gleichung von einer anderen:

- Subtrahiere das 2-fache der ersten von der zweiten,
- subtrahiere das (-1) -fache der ersten von der dritten.

Resultat:

$$\begin{array}{rcl} 2u & + & v & + & w = 1 \\ - & v & - & 2w & = -4 \\ + & 3v & + & 2w & = 8 \end{array}$$

Definition 1. Der Koeffizient **2** in der 1. Gleichung heißt **Pivot**.

Subtraktion der (-3) -fachen 2. Gleichung von der dritten liefert:

$$\begin{array}{rcl} 2u & + & v & + & w = 1 \\ -1v & - & 2w & = & -4 \\ -4w & = & -4 \end{array}$$

$$\begin{array}{lcl} 3. \text{ Gleichung} & \Rightarrow & w = 1 \\ 2. \text{ Gleichung} & \Rightarrow & v = 2 \\ 1. \text{ Gleichung} & \Rightarrow & u = -1 \end{array}$$

d.h. man gewinnt die Lösung durch *Rückwärts-Einsetzen*.

Die Pivots sind: $2, -1, -4$

Fragen:

- Führt dieser Eliminationsprozess stets zu einer Lösung?
D.h. unter welchen Umständen bricht der Prozeß zusammen?
- Wie groß ist der Rechenaufwand für ein System von n Gleichungen für n Unbestimmte?

Antworten:

zu 1. Wenn keiner der Pivots Null ist, gibt es nur eine Lösung und die kann durch *Rückwärts-Einsetzen* gewonnen werden. Wenn aber einer der Pivots Null ist, muß das Eliminationsverfahren beendet werden, entweder vorläufig oder endgültig. Es kann nicht vorausgesagt werden, ob und welche Pivots Null sein werden.

zu 2. Bestimmung der Anzahl der notwendigen Prozesse (nur für die linken Seiten der Gleichungen). Notwendige Schritte:

- Eine Division, um herauszufinden, mit welchem Faktor multipliziert werden muß.
- Multiplikation und Addition, um in der darunterliegenden Gleichung eine Null zu erzeugen.

1. Schritt: 1 Operation zur Bestimmung des Faktors. $n - 1$ Operationen zur Bestimmung der übrigen Faktoren.

Insgesamt also n Operationen.

Für die $n - 1$ Gleichungen (unter der ersten) sind daher $n(n - 1) = n^2 - n$ Prozesse notwendig.

Insgesamt ergibt sich die Anzahl P der notwendigen Operationen:

$$\begin{aligned} P &= n^2 + (n-1)^2 + \dots + 1^2 - n - (n-1) - \dots - 1 \\ &= \frac{n(n+1)(2n+1)}{6} - \frac{n(n+1)}{2} = \frac{n^3 - n}{3} \\ P &\approx \frac{n^3}{3} \end{aligned}$$

Fürs Rückwärts-Einsetzen sind noch

$$Q = 1 + 2 + \dots + n = \frac{n(n+1)}{2} \approx \frac{n^2}{2}$$

Operationen notwendig.

1.2 Matrizen

1.2.1 Definitionen

Mit Hilfe von Matrizen können umfangreiche Berechnungen systematisch ausgeführt werden, da man damit eine kompakte Schreibweise

- zur Speicherung von Informationen und
- für komplizierte Beziehungen

zur Hand hat.

Definition 2. Eine $m \times n$ - Matrix ist eine rechteckige Anordnung von $m \cdot n$ Zahlen $a_{ij} \in \mathbb{K}$:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Mit \mathbb{K} wird die Menge der reellen Zahlen \mathbb{R} oder die Menge der komplexen Zahlen \mathbb{C} bezeichnet, d.h.

$$\mathbb{K} = \mathbb{R} \quad \text{oder} \quad \mathbb{K} = \mathbb{C}.$$

In der Matrix nebeneinander stehende Elemente können zu **Zeilen** zusammengefaßt werden, übereinander stehende Elemente zu **Spalten**.

Schreibweise: $A = (a_{ij})$

Bezeichnungen: a_{ij} : $i \dots$ Zeilenindex,
 $j \dots$ Spaltenindex

$M(m \times n) \dots$ Menge der $m \times n$ Matrizen.

$A \in M(m \times 1) \dots$ Spaltenmatrix (-vektor)

$A \in M(1 \times n) \dots$ Zeilenmatrix (-vektor)

$A \in M(n \times n) \dots$ quadratische Matrix
 $a_{ii} \dots$ Diagonalelemente

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} =: \text{diag}(a_{11}, \dots, a_{nn}) \quad \text{Diagonalmatrix}$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix} \quad \text{obere Dreiecksmatrix}$$

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \quad \text{untere Dreiecksmatrix}$$

1.2.2 Rechenoperationen

Definition 3.

1. $A, B \in M(m \times n)$ heißen **gleich**, falls

$$a_{ij} = b_{ij} \quad \text{für alle } 1 \leq i \leq m, 1 \leq j \leq n$$

2. $A, B \in M(m \times n)$

$$A + B := (a_{ij} + b_{ij}) \quad \text{Addition,}$$

d.h. es wird gliedweise addiert.

$$3. 0 := \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \quad \dots \text{ Nullmatrix}$$

$-A := (-a_{ij}) \quad \dots$ die zu A **negative Matrix**

4. $A + 0 = 0 + A = A$

$$A + (-A) = 0$$

5. $A - B := A + (-B)$ **Subtraktion**

6. $\lambda \in \mathbb{K}, A \in M(m \times n)$

$\lambda \cdot A := (\lambda a_{ij})$ Multiplikation mit Skalaren

Rechengesetze: Für $A, B, C \in M(m \times n)$ und $\lambda, \mu \in \mathbb{K}$ gilt:

$$\begin{aligned} A + B &= B + A \dots \text{Kommutativgesetz} \\ A + (B + C) &= (A + B) + C \dots \text{Assoziativgesetz} \\ (\lambda + \mu)A &= \lambda \cdot A + \mu \cdot A \\ \lambda \cdot (\mu A) &= (\lambda\mu) \cdot A \quad \text{Distributivgesetze} \\ \lambda(A + B) &= \lambda A + \lambda B \\ (-1)A &= -A \\ \lambda \cdot 0 &= 0 \end{aligned}$$

Beispiel 2.

$$A = \begin{pmatrix} 9 & x+2 \\ -3 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} y & -2 \\ 4 & 6 \end{pmatrix}, \quad 0 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$A + B = \begin{pmatrix} 9+y & x \\ 1 & 8 \end{pmatrix} = B + A, \quad -A = \begin{pmatrix} -9 & -x-2 \\ 3 & -2 \end{pmatrix}, \quad A + (-A) = 0$$

$$2A = \begin{pmatrix} 18 & 2x+4 \\ -6 & 4 \end{pmatrix}, \quad 3A - 2B = \begin{pmatrix} 27-2y & 3x+10 \\ -17 & -6 \end{pmatrix}$$

Definition 4. 1. $A \in M(1 \times n), B \in M(n \times 1)$

$$A \cdot B \equiv (a_1, \dots, a_n) \cdot \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} := a_1 b_1 + \dots + a_n b_n$$

2. $A \in M(m \times n), B \in M(n \times r)$

$$C = A \cdot B = (c_{ik}) \in M(m \times r)$$

mit

$$c_{ik} := a_{i1}b_{1k} + \dots + a_{in}b_{nk} = \sum_{j=1}^n a_{ij}b_{jk}$$

heißt **Produkt** der Matrizen A und B .

Bemerkung 1. Die Anzahl der Spalten von A muß gleich sein der Anzahl der Zeilen von B !

$$\text{Beispiel 3. } 1. \begin{pmatrix} -1 & 5 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 4 & 3 & 6 \\ 0 & -1 & 2 \end{pmatrix} = \begin{pmatrix} -4 & -8 & 4 \\ 8 & 5 & 14 \end{pmatrix}$$

$$2. \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$$

$$3. A = (1, 2), \quad B = \begin{pmatrix} -3 \\ 4 \end{pmatrix}, \quad A \cdot B = 5, \quad B \cdot A = \begin{pmatrix} -3 & -6 \\ 4 & 8 \end{pmatrix}$$

Rechenregeln

1. $A \cdot B \neq B \cdot A$ (i.a.)

2. Für $A \in M(n \times n)$ gilt $A^1 := A, \quad A^2 = A \cdot A, \quad A^3 = A \cdot A \cdot A, \quad A^k, k \in \mathbb{N}$, entsprechend.

Definition 5.

$$I \equiv I_n \equiv E_n := \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix} \in M(n \times n)$$

heißt n -reihige Einheitsmatrix.

Rechenregeln: A, B, C seien geeignete Matrizen

1. $A(BC) = (AB)C$ **Assoziativgesetz**
2. $A(B \pm C) = AB \pm AC$ **Distributivgesetz**
3. $(A \pm B)C = AC \pm BC$ **Distributivgesetz**
4. $A \cdot I = I \cdot A = A$
5. $c(AB) = (cA)B = A(cB), c \in \mathbb{K}$
6. $A \cdot 0 = 0 \cdot A = 0$
7. $A \cdot B = A \cdot C \neq B = C !!$
8. $(A + B)(A - B) = A^2 - AB + BA - B^2$

Beispiel 4. Zu Rechenregel 7.

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, C = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\begin{array}{l} A \cdot B = 0 \\ A \cdot C = 0 \end{array} \quad \not\Rightarrow \quad B = C$$

1.2.3 Die Transponierte, die Inverse

Definition 6. Für $A \in M(m \times n)$ heißt

$$A^T \equiv A^t := (a'_{ij}) \quad \text{mit} \quad a'_{ij} = a_{ji}$$

die zu A transponierte Matrix.

Beispiel 5.

$$A = \begin{pmatrix} -1 & 2 & 4 \\ 2 & 6 & 3 \end{pmatrix}, A^T = \begin{pmatrix} -1 & 2 \\ 2 & 6 \\ 4 & 3 \end{pmatrix}$$

Rechenregeln:

1. $(A^T)^T = A$
2. $(A + B)^T = A^T + B^T$
3. $(c \cdot A)^T = cA^T, c \in \mathbb{K}$
4. $(A \cdot B)^T = B^T \cdot A^T$

Definition 7. Gilt für $A \in M(n \times n)$

$$A^T = A$$

so nennt man A eine **symmetrische Matrix**.

Beispiel 6.

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \quad \text{ist symmetrisch.}$$

Definition 8. Sei $A \in M(n \times n)$. Dann heißt die Matrix $X \in M(n \times n)$, für die gilt

$$A \cdot X = X \cdot A = I$$

die **Inverse zu A** .

Beispiel 7.

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}; \quad \text{Ansatz für } X : \quad X = \begin{pmatrix} x & z \\ y & w \end{pmatrix}$$

$$\begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} x & z \\ y & w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \Rightarrow \quad \begin{array}{l} x - y = 1 \\ x + 2y = 0 \end{array} \quad \begin{array}{l} z - w = 0 \\ z + 2w = 1 \end{array}$$

Lösung: $x = \frac{2}{3}, y = -\frac{1}{3}, z = \frac{1}{3}, w = \frac{1}{3}$

$$X = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} \end{pmatrix} \quad \text{Probe: } X \cdot A = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Beispiel 8.

$$A = \begin{pmatrix} 1 & -1 \\ -3 & 3 \end{pmatrix}; \quad \text{Ansatz: } X = \begin{pmatrix} x & z \\ y & w \end{pmatrix}$$

$$A \cdot X = \begin{pmatrix} 1 & -1 \\ -3 & 3 \end{pmatrix} \begin{pmatrix} x & z \\ y & w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \Rightarrow \quad \begin{array}{l} x - y = 1 \\ -3x + 3y = 0 \end{array} \quad \begin{array}{l} z - w = 0 \\ -3z + 3w = 1 \end{array}$$

Widerspruch! $\Rightarrow X$ existiert nicht!

Bemerkung 2. Die Berechnung der Inversen zu $A \in M(n \times n)$ folgt in Abschnitt 1.3.6.

Satz 1. Falls die Inverse zu A existiert, so ist sie eindeutig bestimmt.

Definition 9. Eine Matrix $A \in M(n \times n)$ heißt **regulär**, falls sie eine Inverse besitzt. Eine **singuläre** Matrix besitzt keine Inverse.

Satz 2. Seien $A, B \in M(n \times n)$ regulär. Dann gilt:

1. $A \cdot B$ ist regulär, es gilt

$$(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$$

2. A^{-1} ist regulär, es gilt

$$(A^{-1})^{-1} = A$$

3. A^T ist regulär, es gilt

$$(A^T)^{-1} = (A^{-1})^T$$

Beweis.

Zu 1: Behauptung: $X = B^{-1} \cdot A^{-1}$ sei die Inverse zu AB

- X existiert, da B^{-1} und A^{-1} existieren.

- $(AB)X = (AB)(B^{-1}A^{-1}) = A(\underbrace{BB^{-1}}_I)A^{-1} = AA^{-1} = I$
- $X(AB) = (B^{-1}A^{-1})(AB) = B^{-1}(\underbrace{A^{-1}A}_I)B = B^{-1}B = I$

Zu 2: $X = A$ sei Inverse zu A^{-1}

$$\begin{aligned} X \cdot A^{-1} &= AA^{-1} = I \\ A^{-1} \cdot &= A^{-1}A = I \end{aligned}$$

Zu 3: $X = (A^{-1})^T$ sei Inverse zu A^T

$$\begin{aligned} X \cdot A^T &= (A^{-1})^T A^T = (A \cdot A^{-1})^T = I^T = I \\ A^T \cdot X &= A^T (A^{-1})^T = (A^{-1}A)^T = I^T = I \end{aligned}$$

1.2.4 Anwendung der Inversen bei linearen Gleichungssystemen

Beispiel 9.

$$\begin{aligned} 3x &= 7 && | \cdot (3^{-1}) \\ \Rightarrow x &= 7 \cdot \frac{1}{3} = \frac{7}{3} \end{aligned}$$

Jetzt:

$$A \cdot x = b \quad \text{mit} \quad A \in M(m \times n), x \in M(n \times 1), b \in M(m \times 1)$$

ausführlich

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ \vdots &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

System von m Gleichungen für die n Unbestimmten x_1, \dots, x_n (= **lineares Gleichungssystem**).

Annahme: $A \in M(n \times n)$ sei regulär, d.h. A^{-1} existiert. Dann könnte man durch Multiplikation der Gleichung $Ax = b$ von links mit A^{-1} den Vektor x gewinnen:

$$\underbrace{A^{-1} \cdot A}_{=I} \cdot x = A^{-1}b$$

$$x = A^{-1}b$$

Beispiel 10.

$$\begin{aligned} u - v &= -4 \\ u + 2v &= 5 \end{aligned}$$

$$A = \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}, b = \begin{pmatrix} -4 \\ 5 \end{pmatrix}, x = \begin{pmatrix} u \\ v \end{pmatrix}$$

$$\text{mit } A^{-1} = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{-1}{3} & \frac{1}{3} \end{pmatrix} \quad (\text{siehe Bsp. 7}) \quad \text{folgt } x = A^{-1}b = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{-1}{3} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} -4 \\ 5 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$$

$$\Rightarrow u = -1, \quad v = 3$$

(Probe!)

1.3 GAUSS'sche Elimination

- Zur Berechnung der Lösung eines linearen Gleichungssystems.
- Zur Bestimmung der Inversen einer regulären Matrix.

1.3.1 Einführung - GAUSS'scher Algorithmus

Beispiel 11.

$$\begin{array}{rclclcl} -2x_1 & + & 2x_2 & - & 4x_3 & - & 6x_4 = -4 \\ -3x_1 & + & 6x_2 & + & 3x_3 & - & 15x_4 = -3 \\ 5x_1 & - & 8x_2 & - & x_3 & + & 17x_4 = 9 \\ x_1 & + & x_2 & + & 11x_3 & + & 7x_4 = 7 \end{array}$$

Ziel: Man eliminiert der Reihe nach die einzelnen Unbestimmten aus den Gleichungen.

1. Schritt: Mit Hilfe der 1. Gleichung eliminiert man x_1 aus den übrigen Gleichungen:

1. Dividiere 1. Gleichung durch -2 :

$$x_1 - x_2 + 2x_3 + 3x_4 = 2$$

2. Addiere das 3-fache dieser Gleichung zur 2. Gleichung
3. Addiere das (-5) -fache dieser Gleichung zur 3. Gleichung
4. Addiere das (-1) -fache dieser Gleichung zur 4. Gleichung

Ergebnis:

$$\begin{array}{rclclcl} x_1 & - & x_2 & + & 2x_3 & + & 3x_4 = 2 \\ & & 3x_2 & + & 9x_3 & - & 6x_4 = 3 \\ & & -3x_2 & - & 11x_3 & + & 2x_4 = -1 \\ & & 2x_2 & + & 9x_3 & + & 4x_4 = 5 \end{array}$$

Die letzten 3 Gleichungen enthalten nur mehr die 3 Unbestimmten x_2, x_3, x_4 . Die 1. Gleichung kann bei bekanntem x_2, x_3, x_4 die Variable x_1 liefern.

2. Schritt: Aus den letzten 3 Gleichungen werden jetzt 2 Gleichungen gewonnen, in denen nur x_3 und x_4 vorkommen:

1. Multiplikation der 2. Gleichung mit $\frac{1}{3}$:

$$x_2 + 3x_3 - 2x_4 = 1$$

2. Addiere das 3-fache dieser Gleichung zur 3. Gleichung
3. Addiere das (-2) -fache dieser Gleichung zur 4. Gleichung

Ergebnis:

$$\begin{array}{rclclcl} x_1 & - & x_2 & + & 2x_3 & + & 3x_4 = 2 \\ & & x_2 & + & 3x_3 & - & 2x_4 = 1 \\ & & -2x_3 & - & 4x_4 & = & 2 \\ & & 3x_3 & + & 8x_4 & = & 3 \end{array} \tag{1.1}$$

3. Schritt: Aus den letzten beiden Gleichungen gewinne eine Gleichung für x_4 :

1. Multipliziere die 3. Gleichung mit $(-\frac{1}{2})$:

$$x_3 + 2x_4 = -1$$

2. Addiere das (-3) -fache dieser Gleichung zur 4. Gleichung

Ergebnis:

$$\begin{array}{rcl} x_1 - x_2 + 2x_3 + 3x_4 & = & 2 \\ x_2 + 3x_3 - 2x_4 & = & 1 \\ x_3 + 2x_4 & = & -1 \\ x_4 & = & 3 \end{array} \quad (1.2)$$

Dieses System heißt **reduziertes Gleichungssystem**.

4. Gleichung liefert: $x_4 = 3$

In 3. Gleichung eingesetzt: $x_3 + 6 = 1 \Rightarrow x_3 = -7$

In 2. Gleichung eingesetzt: $x_2 + 3(-7) - 2 \cdot 3 = 1 \Rightarrow x_2 = 28$

In 1. Gleichung eingesetzt: $x_1 - 28 - 14 + 9 = 2 \Rightarrow x_1 = 35$

Eventuell Probe durch Einsetzen in System.

Im vorigen Eliminationsprozeß wurden die Berechnungen mit den Faktoren (=Zahlen) gemacht, nicht mit den Symbolen x_i ! Wesentlich waren nur die Faktoren und die Zahlen auf den rechten Seiten der Gleichungen.

Man bedient sich deshalb der sogenannten **erweiterten Koeffizienten-Matrix** (A, b) :

$$(A, b) = \left(\begin{array}{cccc|c} -2 & 2 & -4 & -6 & -4 \\ -3 & 6 & 3 & -15 & -3 \\ 5 & -8 & -1 & 17 & 9 \\ 1 & 1 & 11 & 7 & 7 \end{array} \right)$$

1. Schritt des Verfahrens:

Die erste Gleichung wird durch (-2) dividiert

\triangleq man ersetzt die 1. Zeile dieser Matrix durch das $(-\frac{1}{2})$ -fache dieser Zeile:

$$(1 \ -1 \ 2 \ 3 \mid 2)$$

Dann benutzen wir die 1. Zeile um in den übrigen Zeilen die Elemente in der ersten Spalte zu Null zu machen:

$$\rightsquigarrow \left(\begin{array}{cccc|c} 1 & -1 & 2 & 3 & 2 \\ 0 & 3 & 9 & -6 & 3 \\ 0 & -3 & -11 & 2 & -1 \\ 0 & 2 & 9 & 4 & 5 \end{array} \right)$$

2. Schritt:

1. Division der 2. Zeile durch 3
2. Addition der 3-fachen 2. Zeile zur 3. Zeile
3. Addition der (-2) -fachen 2. Zeile zur 4. Zeile

$$\rightsquigarrow \left(\begin{array}{cccc|c} 1 & -1 & 2 & 3 & 2 \\ 0 & 1 & 3 & -2 & 1 \\ 0 & 0 & -2 & -4 & 2 \\ 0 & 0 & 3 & 8 & 3 \end{array} \right)$$

(vgl. das entsprechende Gleichungssystem (1.1))

3. Schritt:

1. Multiplikation der 3. Zeile mit $(-\frac{1}{2})$
2. Addition der (-3) -fachen 3. Zeile zur 4. Zeile

$$\rightsquigarrow \left(\begin{array}{cccc|c} 1 & -1 & 2 & 3 & 2 \\ 0 & 1 & 3 & -2 & 1 \\ 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{array} \right)$$

(vgl. das reduzierte Gleichungssystem (1.2))

Beispiel 12.

$$\begin{aligned} -3x_1 - 3x_2 - 3x_3 &= -3 \\ -2x_1 + 2x_2 + x_3 &= 0 \\ x_1 - 3x_2 + 3x_3 &= 0 \end{aligned}$$

$$(A, b) = \left(\begin{array}{ccc|c} -3 & -3 & -3 & -3 \\ -2 & 2 & 1 & 0 \\ 1 & -3 & 3 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 4 & 3 & 2 \\ 0 & -4 & 2 & -1 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & \frac{3}{4} & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{5} \end{array} \right)$$

Das entsprechende reduzierte Gleichungssystem lautet:

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ x_2 + \frac{3}{4}x_3 &= \frac{1}{2} \\ x_3 &= \frac{1}{5} \end{aligned}$$

Man gewinnt, beginnend bei der letzten Gleichung, schrittweise durch Rückwärtseinsetzen:
 $x_3 = 0.2$, $x_2 = 0.35$, $x_3 = 0.45$.

Beispiel 13. Vertauschung von 2 Zeilen $\stackrel{\wedge}{=} \text{Vertauschung von 2 Gleichungen im System.}$

Wir betrachten die erweiterte Koeffizientenmatrix von Bsp. 12:

$$\begin{aligned} (A, b) &= \left(\begin{array}{ccc|c} -3 & -3 & -3 & -3 \\ -2 & 2 & 1 & 0 \\ 1 & -3 & 3 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & -3 & 3 & 0 \\ -2 & 2 & 1 & 0 \\ -3 & -3 & -3 & -3 \end{array} \right) \\ &\rightsquigarrow \left(\begin{array}{ccc|c} 1 & -3 & 3 & 0 \\ 0 & -4 & 7 & 0 \\ 0 & -12 & 6 & -3 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & -3 & 3 & 0 \\ 0 & 1 & -\frac{1}{2} & \frac{1}{4} \\ 0 & 0 & 1 & \frac{1}{5} \end{array} \right) \\ &\quad \begin{aligned} x_1 - 3x_2 + 3x_3 &= 0 \\ \Rightarrow x_2 - \frac{1}{2}x_3 &= \frac{1}{4} \\ x_3 &= \frac{1}{5} \end{aligned} \\ x_3 &= \frac{1}{5}, \quad x_2 = \frac{7}{20}, \quad x_1 = \frac{9}{20} \end{aligned}$$

Bemerkung 3. Man kann Zeilen vertauschen, wenn es das Rechnen vereinfacht. Manchmal ist es aber notwendig, Zeilen zu vertauschen!

Beispiel 14.

$$\begin{aligned} -2x_1 + 4x_2 - 2x_3 - 6x_4 &= 4 \\ 3x_1 - 6x_2 + 6x_3 + 10x_4 &= -1 \\ -2x_1 + 6x_2 - x_3 + x_4 &= 1 \\ 2x_1 - 5x_2 + 4x_3 + 8x_4 &= -3 \end{aligned}$$

$$(A, b) = \left(\begin{array}{cccc|c} -2 & 4 & -2 & -6 & 4 \\ 3 & -6 & 6 & 10 & -1 \\ -2 & 6 & -1 & 1 & 1 \\ 2 & -5 & 4 & 8 & -3 \end{array} \right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & -2 & 1 & 3 & -2 \\ 0 & 0 & 3 & 1 & 5 \\ 0 & 2 & 1 & 7 & -3 \\ 0 & -1 & 2 & 2 & 1 \end{array} \right)$$

Hier ist es, um das Eliminationsverfahren fortsetzen zu können, notwendig, die 2. Zeile z.B. mit der 4. Zeile zu vertauschen.

Anschließend wird die neue 2. Zeile mit dem Faktor (-1) multipliziert:

$$\rightsquigarrow \left(\begin{array}{cccc|c} 1 & -2 & 1 & 3 & -2 \\ 0 & 1 & -2 & -2 & -1 \\ 0 & 2 & 1 & 7 & -3 \\ 0 & 0 & 3 & 1 & 5 \end{array} \right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & -2 & 1 & 3 & -2 \\ 0 & 1 & -2 & -2 & -1 \\ 0 & 0 & 5 & 11 & -1 \\ 0 & 0 & 3 & 1 & 5 \end{array} \right) \rightsquigarrow \dots$$

Allgemein: Die GAUSS’sche Elimination an einem $m \times n$ -System folgt analog zu diesen Beispielen.

1.3.2 GAUSS - JORDAN - Elimination

Beim GAUSS’schen Algorithmus waren nur die Elemente unterhalb der betrachteten Zeile eliminiert worden, jetzt wollen wir versuchen, auch Elemente über der betrachteten Zeile zu eliminieren.

Beispiel 15. Wir betrachten das Gleichungssystem von Bsp. 11 und führen wie dort angegeben zunächst die Elimination nach GAUSS durch:

$$(A, b) = \left(\begin{array}{cccc|c} -2 & 2 & -4 & -6 & -4 \\ -3 & 6 & 3 & -15 & -3 \\ 5 & -8 & -1 & 17 & 9 \\ 1 & 1 & 11 & 7 & 7 \end{array} \right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & -1 & 2 & 3 & 2 \\ 0 & 1 & 3 & -2 & 1 \\ 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{array} \right)$$

Jetzt werden durch geeignete Addition der letzten Zeile (bzw. der vorletzten, dann der zweiten Zeile) zu den darüberliegenden Zeilen weitere Nullen erzeugt:

$$\rightsquigarrow \left(\begin{array}{cccc|c} 1 & -1 & 2 & 0 & 7 \\ 0 & 1 & 0 & 0 & 28 \\ 0 & 0 & 1 & 0 & -7 \\ 0 & 0 & 0 & 1 & 3 \end{array} \right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 35 \\ 0 & 1 & 0 & 0 & 28 \\ 0 & 0 & 1 & 0 & -7 \\ 0 & 0 & 0 & 1 & 3 \end{array} \right)$$

Das reduzierte Gleichungssystem lautet nun

$$\begin{aligned} x_1 &= 35 \\ x_2 &= 28 \\ x_3 &= -7 \\ x_4 &= 3 \end{aligned}$$

d.h. die Lösung ist direkt ablesbar!

Dieses Verfahren nennt man **GAUSS-JORDAN-Elimination**. Es führt dazu, dass im reduzierten Gleichungssystem in jeder Gleichung nur mehr eine Variable allein steht. Damit kann die Lösung stets direkt aus dem reduzierten Gleichungssystem abgelesen werden.

1.3.3 Frage der Existenz von Lösungen

Beispiel 16. $a \cdot x = b$

1. $a \neq 0 : x = \frac{b}{a}$ Lösung ist eindeutig bestimmt.

2. $a = 0 :$

a) $b \neq 0 \quad 0 \cdot x = b$ Widerspruch! Es existiert keine Lösung.

β) $b = 0 \quad 0 \cdot x = 0$ ist wahr für alle $x \in \mathbb{R}$. Es gibt unendlich viele Lösungen.

Beispiel 17. Wir betrachten drei Gleichungssysteme:

1.

$$\begin{array}{rcl} x_1 & + & x_2 = 2 \\ x_1 & - & x_2 = 0 \end{array}$$

Es gibt eine eindeutig bestimmte Lösung: $x_1 = x_2 = 1$

2.

$$\begin{array}{rcl} x_1 & + & x_2 = 2 \\ x_1 & + & x_2 = 1 \end{array}$$

Es gibt keine Lösung.

3.

$$\begin{array}{rcl} x_1 & + & x_2 = 2 \\ 2x_1 & + & 2x_2 = 4 \end{array}$$

Es existieren unendlich viele Lösungen: $x_1 = k, x_2 = 2 - k, k \in \mathbb{R}$.

Beispiel 18.

$$(A, b) = \left(\begin{array}{ccc|c} 1 & 2 & -5 & 2 \\ 2 & -3 & 4 & 4 \\ 4 & 1 & -6 & 8 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 2 & -5 & 2 \\ 0 & -7 & 14 & 0 \\ 0 & -7 & 14 & 0 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 2 & -5 & 2 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

$$\begin{array}{rcl} x_1 & + & 2x_2 - 5x_3 = 2 \\ x_2 & - & 2x_3 = 0 \\ 0 \cdot x_3 & = & 0 \end{array} \implies \begin{cases} x_3 = k, k \in \mathbb{R} & \text{erfüllt 3. Gleichung} \\ x_2 = 2k \\ x_1 = 2 + k \end{cases}$$

Die Lösung x kann also angegeben werden in der Form:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} + k \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, k \in \mathbb{R}$$

Es gibt unendliche viele Lösungen, für jeden Wert von $k \in \mathbb{R}$ eine andere.

Beispiel 19.

$$\begin{array}{rcl} x_1 & + & 2x_2 - x_3 + 2x_4 = 4 \\ 2x_1 & + & 7x_2 + x_3 + x_4 = 14 \\ 3x_1 & + & 8x_2 - x_3 + 4x_4 = 17 \end{array}$$

$$\left(\begin{array}{cccc|c} 1 & 2 & -1 & 2 & 4 \\ 2 & 7 & 1 & 1 & 14 \\ 3 & 8 & -1 & 4 & 17 \end{array} \right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 2 & -1 & 2 & 4 \\ 0 & 3 & 3 & -3 & 6 \\ 0 & 2 & 2 & -2 & 5 \end{array} \right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 2 & -1 & 2 & 4 \\ 0 & 1 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right)$$

$$\begin{array}{rcl} x_1 & + & 2x_2 - x_3 + 2x_4 = 4 \\ x_2 & + & x_3 - x_4 = 2 \\ 0 \cdot x_4 & = & 1 \end{array} \quad \text{Widerspruch!}$$

Es existiert keine Lösung!

1.3.4 Matrizen von Zeilenstufenform

Definition 10. Eine Matrix $B \in M(m \times n)$ heißt von **Zeilenstufenform**, wenn sie von folgender Gestalt ist:

$$B = \left(\begin{array}{cccc|ccccc} 0 & \cdots & 0 & | & b_{1j_1} & \cdots & \cdots & \cdots \\ 0 & \cdots & 0 & | & 0 & \cdots & | & b_{2j_2} \\ \vdots & & \vdots & & \ddots & & & \\ 0 & | & \cdots & & 0 & & | & b_{rj_r} & \cdots \\ 0 & | & \cdots & & & & | & 0 & \\ 0 & | & \cdots & & & & | & 0 & \end{array} \right) \quad \text{mit } b_{1j_1} \neq 0, \dots, b_{rj_r} \neq 0,$$

Beachte: Die ersten r Zeilen sind ungleich Null.

Die an den Stufenkanten stehenden Elemente sind $\neq 0$, d.h.

$$b_{lj_l} \neq 0 \quad \text{für } l = 1, \dots, r.$$

Die unterhalb der Stufenlinie stehenden Elemente sind alle Null, die restlichen Elemente können beliebig sein.

Definition 11. Sei r die Anzahl der von Null verschiedenen Zeilen der Matrix B von Zeilenstufenform. Dann heißt r **Rang** von B , d.h.

$$\text{rang } B := r.$$

Definition 12. An einer Matrix A kann man folgende **elementare Zeilenumformungen (Zeilenoperationen)** durchführen:

Typ I: Vertauschung von zwei Zeilen.

Typ II: Multiplikation der i -ten Zeile mit dem Faktor $\lambda, \lambda \neq 0$.

Typ III: Addition der λ -fachen i -ten Zeile zur k -ten Zeile, $\lambda \neq 0$.

Satz 3.

1. Jede Matrix $A \in M(m \times n)$ kann durch endlich viele elementare Zeilenumformungen auf Zeilenstufenform gebracht werden.
2. Sei B eine Matrix von Zeilenstufenform, die aus der Matrix A durch elementare Zeilenumformungen entstanden ist. Dann gilt:

$$\text{rang } A = \text{rang } B$$

d.h. elementare Zeilenumformungen verändern den Rang einer Matrix nicht.

3. Sei (A, b) die erweiterte Koeffizientenmatrix des Gleichungssystems

$$A x = b$$

die durch elementare Zeilenumformungen auf die Zeilenstufenform (A', b') gebracht worden ist. Dann gilt:

$$x \text{ ist Lösung von } Ax = b \iff x \text{ ist Lösung von } A'x = b'.$$

D.h. das ursprüngliche Gleichungssystem und das reduzierte Gleichungssystem haben (falls sie nicht leer ist) dieselbe Lösungsmenge.

1.3.5 Existenz und Struktur der Lösungen

Zur Erinnerung: In den Beispielen 11, 18 und 19 hatten wir folgende Matrizen von Zeilenstufenform gewonnen:

Beispiel 20.

$$(A', b') = \left(\begin{array}{cccc|c} 1 & -1 & 2 & 0 & -7 \\ 0 & 1 & 3 & 0 & 7 \\ 0 & 0 & 1 & 0 & -7 \\ 0 & 0 & 0 & 1 & 3 \end{array} \right)$$

$\text{rang } A = 4, \quad \text{rang } (A, b) = 4$. Das System $Ax = b$ war eindeutig lösbar.

Beispiel 21.

$$(A', b') = \left(\begin{array}{ccc|c} 1 & 2 & -5 & 2 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

$\text{rang } A = 2, \quad \text{rang } (A, b) = 2$. Das System besaß unendlich viele Lösungen.

Beispiel 22.

$$(A', b') = \left(\begin{array}{cccc|c} 1 & 2 & -1 & 2 & 4 \\ 0 & 1 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right)$$

$\text{rang } A = 2, \quad \text{rang } (A, b) = 3$. Das System hatte keine Lösung.

Satz 4. Lösbarkeit von Gleichungssystemen

Seien

$$A \in M(m \times n), \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m, \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n.$$

- Das Gleichungssystem $Ax = b$ ist lösbar, falls $\text{rang } A = \text{rang } (A, b)$
- Das Gleichungssystem $Ax = b$ ist eindeutig lösbar, falls $\text{rang } A = \text{rang } (A, b) = n$

Beispiel 23.

$$\begin{array}{rcl} x & - & 3y = -2 \\ 2x & + & y = 3, \quad \alpha \in \mathbb{R}. \\ 3x & - & 2y = \alpha \end{array}$$

Für welche $\alpha \in \mathbb{R}$ gibt es eine Lösung?

$$(A, b) = \left(\begin{array}{cc|c} 1 & -3 & -2 \\ 2 & 1 & 3 \\ 3 & -2 & \alpha \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|c} 1 & -3 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & \alpha - 1 \end{array} \right)$$

$$\text{rang } A = 2, \quad \text{rang } (A, b) = \begin{cases} 3 & \text{für } \alpha \neq 1 \\ 2 & \text{für } \alpha = 1 \end{cases}$$

\Rightarrow System eindeutig lösbar für $\alpha = 1$; System nicht lösbar für $\alpha \neq 1$.

Zur **Struktur** der Lösung: Zwei Fälle sind klar

- Es existiert keine Lösung
- Es existiert eine eindeutig bestimmte Lösung.

Was ist im Fall

- Es existieren unendlich viele Lösungen?

Seien y und z Lösungen von $Ax = b$, d.h. $Ay = b$ und $Az = b$:

$$\Rightarrow A(y - z) = Ay - Az = b - b = 0$$

d.h. die Spalte $h = y - z$ erfüllt das sogenannte **zugehörige homogene System**

$$Ah = 0.$$

Ist umgekehrt h eine Lösung des zugehörigen homogenen Systems und x_0 eine spezielle Lösung des inhomogenen Systems $Ax = b$, dann gilt:

$$A(x_0 + h) = Ax_0 + Ah = b + 0 = b$$

$\Rightarrow x_0 + h$ ist Lösung des inhomogenen Systems.

Satz 5. Sei x_0 eine partikuläre (= spezielle) Lösung des Systems $Ax = b$. Dann ist die Menge aller Lösungen x von $Ax = b$ gleich der Menge aller Vektoren von der Form

$$x = x_0 + h$$

wobei h die Menge der Lösungen des zugehörigen homogenen Systems $Ah = 0$ durchläuft.

Folgerungen:

1. $Ax = b$ eindeutig lösbar $\Leftrightarrow \begin{cases} Ax = b & \text{ist lösbar} \\ Ax = 0 & \text{besitzt nur die (triviale) Lösung } x = 0 \end{cases}$
2. $A \in M(n \times n)$:
 $Ax = b$ eindeutig lösbar $\Leftrightarrow \text{rang } A = n \Leftrightarrow A$ ist regulär $\Leftrightarrow Ax = 0$ besitzt nur die triviale Lösung.
3. $\text{rang } A = n \Leftrightarrow A^{-1}$ existiert.

1.3.6 Bestimmung der inversen Matrix

Gesucht ist die Inverse der regulären Matrix $A \in M(n \times n)$. Es gilt

$$A \cdot A^{-1} = I$$

d.h. die Spalten y_i von A^{-1} erfüllen die Gleichungssysteme

$$Ay_i = e_i, \quad i = 1, \dots, n, \quad \text{mit} \quad e_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow \text{i-te Komponente.}$$

Wir lösen diese Gleichungssysteme gleichzeitig mit Hilfe des GAUSS-JORDAN-Algorithmus:

$$(A, I) = \left(\begin{array}{ccc|cccc} a_{11} & \cdots & a_{1n} & 1 & 0 & \cdots & 0 \\ a_{21} & \cdots & a_{2n} & 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} & 0 & 0 & \cdots & 1 \end{array} \right)$$

Für jedes System ist nur eine andere rechte Seite zu nehmen.

1. Schritt: Wir bringen (A, I) auf Zeilenstufenform:

$$\left(\begin{array}{ccc|cccc} b_{11} & \cdots & b_{1n} & c_{11} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & b_{nn} & c_{n1} & c_{n2} & \cdots & c_{nn} \end{array} \right) = (B, C)$$

Hier kann man schon erkennen, ob A invertierbar ist. Ist $\text{rang } B < n$, so ist A singulär! Das Verfahren ist zu stoppen.

2. Schritt: Durch weitere elementare Zeilenumformungen führt man die Teilmatrix B in die Einheitsmatrix über:

$$\left(\begin{array}{ccc|ccc} 1 & 0 & \cdots & 0 & d_{11} & \cdots & d_{1n} \\ 0 & 1 & \cdots & 0 & d_{21} & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & d_{n1} & \cdots & d_{nn} \end{array} \right) = (I, D) \quad (1.3)$$

Jetzt betrachten wir das erste Gleichungssystem für die erste Spalte von A^{-1} in reduzierter Form:

$$\begin{aligned} y_1 &= d_{11} \\ y_2 &= d_{21} \\ &\vdots \\ y_n &= d_{n1} \end{aligned}$$

d.h. die erste Spalte von A^{-1} stimmt mit der ersten Spalte von D überein.

Entsprechend stimmt die i -te Spalte von A^{-1} ($i = 1, \dots, n$) mit der i -ten Spalte von D überein. D.h. es gilt

$$A^{-1} = D$$

was bedeutet, daß rechts von der Trennlinie in (1.3) die Inverse A^{-1} von A steht.

Beispiel 24.

$$A = \begin{pmatrix} 4 & -5 \\ -3 & 5 \end{pmatrix}$$

$$\begin{aligned} \left(\begin{array}{cc|cc} 4 & -5 & 1 & 0 \\ -3 & 5 & 0 & 1 \end{array} \right) &\rightsquigarrow \left(\begin{array}{cc|cc} 1 & -\frac{5}{4} & \frac{1}{4} & 0 \\ 0 & \frac{5}{4} & \frac{3}{4} & 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & -\frac{5}{4} & \frac{1}{4} & 0 \\ 0 & 1 & \frac{3}{5} & \frac{4}{5} \end{array} \right) \rightsquigarrow \\ &\rightsquigarrow \left(\begin{array}{cc|cc} 1 & 0 & \frac{1}{3} & \frac{4}{5} \\ 0 & 1 & \frac{3}{5} & \frac{4}{5} \end{array} \right) \Rightarrow A^{-1} = \begin{pmatrix} \frac{1}{3} & \frac{4}{5} \\ \frac{3}{5} & \frac{4}{5} \end{pmatrix} \end{aligned}$$

Beispiel 25.

$$\begin{array}{l}
 \left(\begin{array}{ccc|ccc} -1 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 1 & -1 & -1 & 0 & 0 & 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & -2 & -1 & -1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{array} \right) \rightsquigarrow \\
 \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & -2 & -1 & -1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 & -1 & 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & -2 & 0 & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{array} \right) \rightsquigarrow \\
 \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{3}{2} & -\frac{1}{2} & \frac{5}{2} \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{array} \right) \implies A^{-1} = \left(\begin{array}{ccc} \frac{3}{2} & -\frac{1}{2} & \frac{5}{2} \\ 1 & 0 & 1 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{array} \right)
 \end{array}$$

Beispiel 26.

$$A = \begin{pmatrix} 4 & -3 \\ -8 & 6 \end{pmatrix} : \quad \left(\begin{array}{cc|cc} 4 & -3 & 1 & 0 \\ -8 & 6 & 0 & 1 \end{array} \right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & -\frac{3}{4} & \frac{1}{4} & 0 \\ 0 & 0 & 2 & 1 \end{array} \right)$$

$$\text{rang } A = 1 < 2 \Rightarrow A^{-1} \text{ existiert nicht!}$$

1.4 LR-Zerlegung einer regulären Matrix

1.4.1 LR-Zerlegung ohne und mit Zeilenvertauschung

Satz 6. LR-Zerlegung ohne Zeilenvertauschung

Eine reguläre Matrix $A \in M(n \times n)$, die ohne Zeilenvertauschungen auf Zeilenstufenform gebracht werden kann, kann als Produkt einer linken unteren Dreiecksmatrix L und einer rechten oberen Dreiecksmatrix R angegeben werden:

$$A = L \cdot R$$

Dabei ist

- R die Matrix von Zeilenstufenform, die aus A durch elementare Zeilenumformungen vom Typ III (siehe Def. 12) (d.h. ohne Zeilenvertauschungen!) entstanden ist.
- L die quadratische Matrix, die man erhält, wenn man in der Hauptdiagonale jeweils 1 einträgt und in die Spalten darunter jeweils die negativen Faktoren, die beim GAUSS'schen Verfahren notwendig waren, um A in Zeilenstufenform überzuführen.

Beispiel 27.

$$\begin{array}{l}
 A = \begin{pmatrix} 1 & 2 & -6 \\ -3 & 4 & 7 \\ 2 & 4 & 3 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 2 & -6 \\ 0 & 10 & -11 \\ 0 & 0 & 15 \end{pmatrix} = R \\
 L = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}
 \end{array}$$

Probe:

$$L \cdot R = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & -6 \\ 0 & 10 & -11 \\ 0 & 0 & 15 \end{pmatrix} = A$$

Sind beim GAUSS'schen Eliminationsprozeß jedoch Zeilenvertauschungen notwendig, so ist die Zerlegung zu modifizieren.

Satz 7. LR-Zerlegung mit Zeilenvertauschung

Jede Matrix $A \in M(n \times n)$ kann zerlegt werden in ein Produkt der Gestalt

$$A = P^T L R \quad (\Leftrightarrow P A = L R)$$

Dabei ist

- R die Matrix, die aus A durch elementare Zeilenumformungen (einschließlich Zeilenvertauschungen) entsteht und die Zeilenstufenform hat.
- L eine quadratische Matrix, in deren Hauptdiagonale nur 1 stehen. Nach jedem Eliminationsprozeß trägt man die negativen Faktoren, die notwendig waren, um die Elemente in der i -ten Spalte mit einer Umformung vom Typ III zu Null zu machen, in die i -te Spalte unter den Diagonalelementen einer Matrix L^* ein (in der Hauptdiagonalen der unteren Dreiecksmatrix stehen lauter 1). Bei jeder Zeilenvertauschung werden die Elemente in den entsprechenden Zeilen von L^* , die bisher bestimmt worden sind und die links von der 1 stehen, vertauscht. Daraus ergibt sich am Schluß des Eliminationsprozesses die Matrix L .
- Die Matrix P (mit der Eigenschaft $P^T = P^{-1}$) enthält man dadurch, daß man die Zeilen der n -reihigen Einheitsmatrix in derselben Weise vertauscht, wie es für das GAUSS'sche Verfahren notwendig war. Man nennt P die zugehörige **Permutationsmatrix**.

Beispiel 28.

$$\begin{aligned} A &= \begin{pmatrix} 2 & 6 & -4 \\ -4 & -12 & 11 \\ 3 & 14 & -16 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & 6 & -4 \\ 0 & 0 & 3 \\ 0 & 5 & -10 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & 6 & -4 \\ 0 & 5 & -10 \\ 0 & 0 & 3 \end{pmatrix} = R \\ L^* &= \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ \frac{3}{2} & * & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ \frac{3}{2} & 1 & 0 \\ -2 & * & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ \frac{3}{2} & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} = L \\ P &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad P^{-1} = P^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \end{aligned}$$

Probe: $A = P^T L R$

Beispiel 29.

$$\begin{aligned} A &= \begin{pmatrix} 2 & 4 & 0 & -2 \\ -4 & -8 & 0 & 3 \\ 3 & 7 & 2 & 4 \\ 5 & 6 & 1 & -8 \end{pmatrix} \\ &\rightsquigarrow \begin{pmatrix} 2 & 4 & 0 & -2 \\ 0 & 0 & 0 & -1 \\ 0 & 1 & 2 & 7 \\ 0 & -4 & 1 & -3 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & 4 & 0 & -2 \\ 0 & 1 & 2 & 7 \\ 0 & -4 & 1 & -3 \\ 0 & 0 & 0 & -1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & 4 & 0 & -2 \\ 0 & 1 & 2 & 7 \\ 0 & 0 & 9 & 25 \\ 0 & 0 & 0 & -1 \end{pmatrix} = R \\ L^* &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ \frac{3}{2} & * & 1 & 0 \\ \frac{5}{2} & * & * & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{3}{2} & 1 & 0 & 0 \\ \frac{5}{2} & * & 1 & 0 \\ -2 & * & * & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{3}{2} & 1 & 0 & 0 \\ \frac{5}{2} & -4 & 1 & 0 \\ -2 & 0 & 0 & 1 \end{pmatrix} = L \end{aligned}$$

$$I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} = P$$

Probe: $P^{-1} = P^T$; $P^T L R = A$

1.4.2 Anwendungen

Lösen von linearen Gleichungssystemen der Form $Ax = b$:

- Mit $A = LR$ gilt:

$$L \cdot \underbrace{R \cdot x}_{=y} = b$$

d.h. man hat zwei Gleichungssysteme

$$\begin{aligned} Ly &= b \\ Rx &= y \end{aligned}$$

zu lösen, allerdings ist jedes davon durch Rückwärtssubstitution zu behandeln:

1. $Ly = b$ beginnend bei der 1. Gleichung nach unten fortschreitend;
2. $Rx = y$ beginnend bei der letzten Gleichung nach oben fortschreitend.

- Mit $A = P^T LR$ (d.h. LR -Zerlegung mit Zeilenumtauschung) gilt

$$A \cdot x = b \Leftrightarrow P^T L R x = b \Leftrightarrow \underbrace{P P^T}_{=I} L R x = \underbrace{P b}_{=b'}$$

Man hat also $LRx = b'$ mit $b' = Pb$ zu lösen:

$$Ly = b'$$

$$Rx = y$$

Beispiel 30.

$$\begin{array}{rclll} x_1 & - & x_2 & + & x_3 = 2 \\ 2x_1 & - & 2x_2 & + & 3x_3 = 7 \\ -x_1 & + & 2x_2 & + & x_3 = 6 \end{array}$$

$$A = \begin{pmatrix} 1 & -1 & 1 \\ 2 & -2 & 3 \\ -1 & 2 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} = R$$

$$L^* = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & * & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & * & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} = L$$

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}; \quad b' = Pb = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 7 \\ 6 \end{pmatrix} = \begin{pmatrix} 2 \\ 6 \\ 7 \end{pmatrix}$$

Das Gleichungssystem $Ly = b'$ lautet:

$$\begin{array}{rcl} y_1 & & = 2 \\ -y_1 + y_2 & & = 6 \\ 2y_1 & + & y_3 = 7 \end{array}$$

$$\Rightarrow y_1 = 2, \quad y_2 = 8, \quad y_3 = 3$$

Das Gleichungssystem $Rx = y$ lautet nun:

$$\begin{array}{rcl} x_1 - x_2 + x_3 & = & 2 \\ x_2 + 2x_3 & = & 8 \\ x_3 & = & 3 \end{array}$$

$$\Rightarrow x_3 = 3, \quad x_2 = 2, \quad x_1 = 1$$

Vergleich des Rechenaufwandes zur Lösung eines $n \times n$ -Systems:

1. Über Inverse von A : n^3 Multiplikationen/Divisionen
2. GAUSS'sches Eliminationsverfahren: $\frac{n^3-n}{3}$ Multiplikationen/Divisionen
3. LR-Zerlegung: $\frac{n^3-n}{3}$ Multiplikationen/Divisionen für Zerlegung
 n^2 Multiplikationen/Divisionen für Lösen der Systeme

Die Kosten für ein einzelnes System sind beim GAUSS'schen Verfahren niedriger.

Liegen allerdings mehrere Gleichungssysteme mit derselben Koeffizientenmatrix A , aber verschiedenen rechten Seiten b_i , $i = 1, \dots, k$, vor, so muß die LR-Zerlegung nur einmal erfolgen.

Weitere Einsparungen ergeben sich dann bei der sogenannten Cholesky-Zerlegung, die in Abschnitt 7.3 besprochen werden soll.

1.5 Determinanten

Wir ordnen einer $n \times n$ -Matrix eine (reelle oder komplexe) Zahl zu.

Definition 13. 1. Die Determinante einer 1×1 Matrix $A = (a)$ ist definiert als $\det A := a$

2. Die Determinante einer 2×2 Matrix $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ ist definiert als

$$\det A \equiv \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} := a_{11}a_{22} - a_{12}a_{21}$$

3. Die Determinante einer 3×3 Matrix $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ ist definiert als

$$\det A \equiv \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} := a_{11} \cdot \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \cdot \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \cdot \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} =$$

$$= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{21}a_{12}a_{33} + a_{21}a_{13}a_{32} + a_{31}a_{12}a_{23} - a_{31}a_{13}a_{22}$$

Regel von Sarrus: Man schreibt die erste und die zweite Spalte von A noch einmal neben die Matrix

$$\begin{array}{ccccc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}$$

Jetzt bildet man die Produkte der drei Diagonalen, die von links oben nach rechts unten verlaufen und addiert diese Zahlen; davon zieht man die Produkte der in den von rechts oben nach links unten verlaufenden Diagonalen stehenden Elemente ab.

Beispiel 31.

$$\det A = \begin{vmatrix} 1 & 2 & -1 \\ 0 & 1 & 3 \\ -1 & 2 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 2 & -1 \\ 0 & 1 & 3 \\ -1 & 2 & 1 \end{vmatrix} \begin{vmatrix} 1 & 2 \\ 0 & 1 \\ -1 & 2 \end{vmatrix}$$

$$= 1 \cdot 1 \cdot 1 + 2 \cdot 3 \cdot (-1) + (-1) \cdot 0 \cdot 2 - [(-1) \cdot 1 \cdot (-1) + 1 \cdot 3 \cdot 2 + 2 \cdot 0 \cdot 1] = -12$$

Definition 14. 1. Für eine Matrix $A = (a_{ij}) \in M(n \times n)$ ist das **algebraische Komplement** A'_{ij} die Determinante der $(n-1) \times (n-1)$ -Matrix, die man aus A durch Streichen der i -ten Zeile und der j -ten Spalte erhält.

2. Die Determinante einer Matrix $A = (a_{ij}) \in M(n \times n)$ ist definiert als

$$\det A = \sum_{j=1}^n (-1)^{1+j} a_{1j} A'_{1j}$$

Beispiel 32.

$$\begin{vmatrix} 1 & 2 & -1 \\ 0 & 1 & 3 \\ -1 & 2 & 1 \end{vmatrix} = 1 \cdot \begin{vmatrix} 1 & 3 \\ 2 & 1 \end{vmatrix} - 2 \cdot \begin{vmatrix} 0 & 3 \\ -1 & 1 \end{vmatrix} + (-1) \cdot \begin{vmatrix} 0 & 1 \\ -1 & 2 \end{vmatrix} = 1 \cdot (-5) - 2 \cdot 3 - 1 \cdot 1 = -12$$

Satz 8. Entwicklung nach einer Zeile bzw. Spalte

Für die Determinante von $A \in M(n \times n)$ gilt:

- $\det A = \sum_{j=1}^n (-1)^{i+j} a_{ij} A'_{ij}, \quad i = 1, \dots, n, \quad (\text{Entwicklung nach der } i\text{-ten Zeile})$
- $\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} A'_{ij}, \quad j = 1, \dots, n, \quad (\text{Entwicklung nach der } j\text{-ten Spalte})$

Beispiel 33. Wir betrachten die Determinante einer oberen Dreiecksmatrix

$$\det A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix}$$

Wir entwickeln wiederholt nach der 1. Spalte:

$$\begin{aligned} \det A &\stackrel{\text{Entw.n.1.Sp.}}{=} a_{11} \begin{vmatrix} a_{22} & a_{23} & \dots & a_{24} \\ 0 & a_{33} & \dots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & a_{nn} \end{vmatrix} \stackrel{\text{Entw.n.1.Sp.}}{=} \\ &= a_{11} \cdot a_{22} \begin{vmatrix} a_{33} & \dots & a_{3n} \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{nn} \end{vmatrix} = \dots = a_{11} \cdot a_{22} \cdots a_{nn} = \prod_{k=1}^n a_{kk} \end{aligned}$$

D.h. die Determinante einer oberen Dreiecksmatrix ergibt sich als Produkt der Diagonalelemente. Dasselbe gilt für eine untere Dreiecksmatrix.

Folgerung:

$$\det I_n = 1$$

Weitere Eigenschaften:

1. $\det A^T = \det A$.
2. Ist eine Zeile oder eine Spalte von A Null, so ist $\det A = 0$.
3. $\det(\lambda \cdot A) = \lambda^n \det A$ für $A \in M(n \times n)$, $\lambda \in \mathbb{K}$.
4. Sind zwei Zeilen (oder Spalten) von A gleich, so ist $\det A = 0$.
5. $\det(A \cdot B) = \det A \cdot \det B$.
6. $\det(A + B) \neq \det A + \det B$!!

Wie sich die Determinante einer Matrix $A \in M(n \times n)$ ändert, wenn man an ihr elementare Zeilenumformungen (siehe Definition 12) durchführt, wird in folgendem Satz zusammengefaßt.

Satz 9.

1. Entsteht A' aus A durch Vertauschen von zwei Zeilen (oder Spalten), so gilt

$$\det A' = (-1) \cdot \det A$$

2. Entsteht A' aus A durch Multiplikation der i -ten Zeile (Spalte) mit dem Faktor $\lambda \in \mathbb{K}$, so gilt

$$\det A' = \lambda \cdot \det A$$

3. Entsteht A' aus A durch Addition der λ -fachen i -ten Zeile (Spalte) zur j -ten Zeile (Spalte), so gilt

$$\det A' = \det A.$$

Bemerkung: Um Determinanten möglichst einfach zu berechnen, kombiniert man das GAUSS'sche Verfahren mit dem Entwicklungssatz. Zunächst erzeugt man z.B. durch elementare Zeilen- bzw. Spaltenumformungen möglichst viele Nullen in einer Zeile bzw. Spalte und entwickelt anschließend nach dieser Zeile bzw. Spalte.

Beispiel 34.

$$\begin{aligned} \det A &= \begin{vmatrix} 2 & -3 & 2 & 5 \\ 1 & -1 & 1 & 2 \\ 3 & 2 & 2 & 1 \\ 1 & 1 & -3 & -1 \end{vmatrix} = \begin{vmatrix} 0 & -1 & 0 & 1 \\ 1 & -1 & 1 & 2 \\ 0 & 5 & -1 & -5 \\ 0 & 2 & -4 & -3 \end{vmatrix} \\ &\stackrel{\text{Entw.n.1.Spalte}}{=} (-1) \begin{vmatrix} -1 & 0 & 1 \\ 5 & -1 & -5 \\ 2 & -4 & -3 \end{vmatrix} = - \begin{vmatrix} -1 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & -4 & -1 \end{vmatrix} \\ &\stackrel{\text{Entw.n.2.Zeile}}{=} -(-1) \cdot \begin{vmatrix} -1 & 1 \\ 0 & -1 \end{vmatrix} = 1 \cdot (-1)(-1) = 1 \end{aligned}$$

Satz 10.

- Für $A \in M(n \times n)$ sind folgende Aussagen äquivalent:

$$A \text{ ist regulär} \Leftrightarrow \text{rang } A = n \Leftrightarrow \det A \neq 0$$

- Ist A regulär, so gilt

$$\det(A^{-1}) = \frac{1}{\det A}$$

Kapitel 2

Vektoren im anschaulichen Raum

2.1 Einführung

In diesem Abschnitt geht es um Vektoren, mit denen geometrische Fragestellungen im anschaulichen Raum oder in der Ebene beschrieben und gelöst werden können. Ihre Eigenschaften werden zusammengefaßt dargestellt und anschaulich erklärt, damit bei einer abstrakten Behandlung von allgemeinen Vektorräumen auf den Spezialfall des dreidimensionalen Raumes unserer Anschauung zurückgeriffen werden kann.

Definition 15.

1. Unter einem **Pfeil** \vec{AB} in der Ebene oder im Raum versteht man ein Paar von verschiedenen Punkten (in der Ebene oder im Raum), die durch eine Strecke verbunden sind.
2. Ein **Vektor** ist eine Äquivalenzklasse gleichlanger paralleler Pfeile.

Schreibweise: $a, \vec{a}, \overrightarrow{AB}$

$\|a\| \dots$ Länge von a .

D.h.: Alle Pfeile gleicher Länge und Richtung sind Veranschaulichung des gleichen Vektors, z.B. Kraft, Geschwindigkeit, Verschiebung.

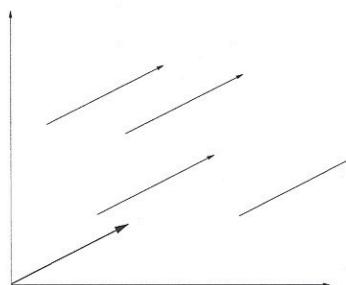
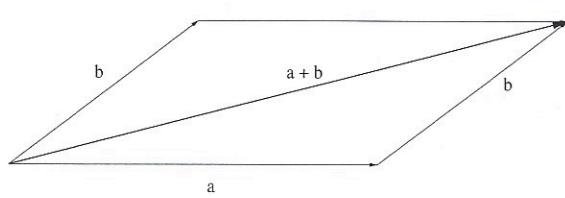
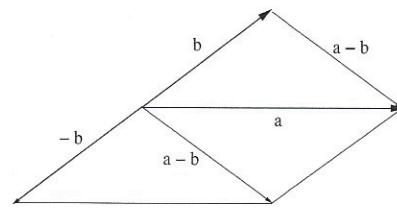


Abbildung 2.1:

Definition 16. Die **Summe** zweier Vektoren a und b ist die gerichtete Diagonale des durch a und b aufgespannten Parallelogramms. (Siehe Abb. 2.2)

D.h.: Vektoren lassen sich wie Verschiebungen durch Hintereinanderausführen addieren. Man hängt den Vektorpfeil des zweiten Summanden b an die Spitze des ersten Summanden a und zeichnet den Summenvektor $a + b$ vom Anfangspunkt von a zur Spitze von b ein.

Abbildung 2.2: Summe $a + b$ Abbildung 2.3: Differenz $a - b$

Definition 17. Die **Differenz** $a - b$ ist gleich der Summe des Vektors a und des zu b entgegengesetzten Vektors $-b$. (Siehe Abb. 2.3)

$$a - b := a + (-b)$$

Definition 18. Unter dem λ -fachen $b = \lambda \cdot a$, $\lambda \in \mathbb{R}$, **eines Vektors** versteht man den Vektor, der die $|\lambda|$ -fache Länge von a und die gleiche Richtung wie a hat, falls $\lambda > 0$ ist, die entgegengesetzte Richtung von a , falls $\lambda < 0$ ist.

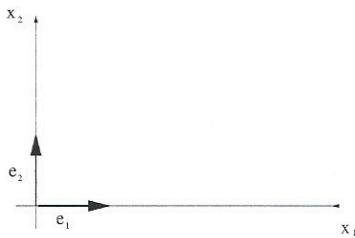
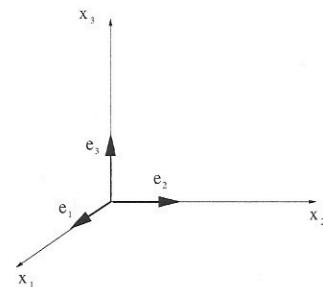
Definition 19. Ein Vektor a mit der Länge 0 heißt **Nullvektor** 0.

Rechenregeln.

$a + b = b + a$	$a + 0 = a$
$(a + b) + c = a + (b + c)$	$0 \cdot a = 0$
$\ \lambda \cdot a\ = \lambda \cdot \ a\ $	
$\lambda \cdot (a + b) = \lambda \cdot a + \lambda \cdot b$	$\lambda, \mu \in \mathbb{R}$
$(\lambda + \mu) \cdot a = \lambda \cdot a + \mu \cdot a$	
$\ a + b\ \leq \ a\ + \ b\ $	Dreiecksungleichung

Um mit Vektoren numerisch rechnen zu können, braucht man außer ihren Beträgen noch eine zahlenmäßige Erfassung ihrer Richtung und dazu ein geeignetes Bezugssystem. Als solches wählt man ein rechtwinkeliges (x_1, x_2) - oder ein (x_1, x_2, x_3) -Koordinatensystem. Dabei bilden die x_1 -, x_2 - und x_3 -Koordinatenachsen ein Rechtssystem (Rechte-Hand-Regel).

Wir legen die zwei Einheitsvektoren e_1, e_2 bzw. die drei Einheitsvektoren e_1, e_2, e_3 fest. Sie besitzen die Länge 1 und zeigen in Richtung der drei Koordinatenachsen.

Abbildung 2.4: Koordinatensystem im \mathbb{R}^2 Abbildung 2.5: Koordinatensystem im \mathbb{R}^3

Definition 20. Man nennt

$$a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

wobei die a_i die vorzeichenbehafteten Projektionen von a auf die x_i -Achsen ($i = 1, 2, 3$) sind, die **Koordinatendarstellung** des Vektors a .

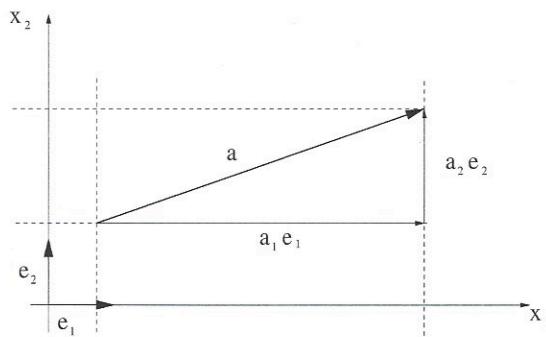


Abbildung 2.6: Die Koordinaten \$a_1, a_2\$ von \$a\$

Die Koordinatendarstellung der Einheitsvektoren lautet also

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Definition 21. Unter Verwendung der Einheitsvektoren \$e_1, e_2, e_3\$ ergibt sich mittels der Vektorsumme die **Komponentendarstellung**

$$a = a_1 e_1 + a_2 e_2 + a_3 e_3$$

(siehe Abb. 2.6)

Satz 11.

$$a = b \Leftrightarrow a_1 = b_1, a_2 = b_2, a_3 = b_3$$

$$\lambda \cdot a = \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ \lambda a_3 \end{pmatrix}, \quad \lambda \in \mathbb{R}, \quad a \pm b = \begin{pmatrix} a_1 \pm b_1 \\ a_2 \pm b_2 \\ a_3 \pm b_3 \end{pmatrix}, \quad 0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

2.2 Das Skalarprodukt

Definition 22. Das **Skalarprodukt** \$\langle a, b \rangle\$ zweier Vektoren \$a\$ und \$b\$ ist definiert als

$$\langle a, b \rangle = \|a\| \cdot \|b\| \cdot \cos \varphi$$

\$\varphi = \angle(a, b)\$... **Winkel** zwischen \$a\$ und \$b\$

\$b' = \|b\| \cdot \cos \varphi\$... **Projektion** von \$b\$ auf \$a\$

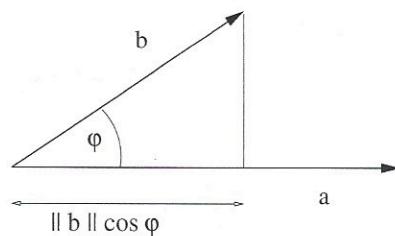


Abbildung 2.7: Winkel zwischen 2 Vektoren

Rechenregeln. 1. $\langle a, b \rangle = \langle b, a \rangle$

$$2. \langle a, b + c \rangle = \langle a, b \rangle + \langle a, c \rangle$$

$$3. \langle a, b \rangle = 0 \Leftrightarrow a \perp b \quad (a \text{ ist orthogonal zu } b, \text{ d.h. } a \text{ steht senkrecht auf } b)$$

$$4. \langle a, a \rangle = \|a\|^2$$

Aus obiger Rechenregel folgt:

$$\begin{aligned} \langle a, b \rangle &= \langle a_1 e_1 + a_2 e_2 + a_3 e_3, b_1 e_1 + b_2 e_2 + b_3 e_3 \rangle \\ &= a_1 b_1 \langle e_1, e_1 \rangle + a_1 b_2 \langle e_1 e_2 \rangle + \dots \\ &= a_1 b_1 \cdot 1 + a_1 b_2 \cdot 0 + \dots \\ &= a_1 b_1 + a_2 b_2 + a_3 b_3 \end{aligned}$$

Satz 12. 1. $\langle a, b \rangle = a_1 b_1 + a_2 b_2 + a_3 b_3$

$$2. \|a\| = \sqrt{\langle a, a \rangle} = \sqrt{a_1^2 + a_2^2 + a_3^2}$$

2.3 Das Vektorprodukt

Definition 23. Das **Vektorprodukt** (= Kreuzprodukt) der beiden Vektoren $a, b \in \mathbb{R}^3$ ist ein Vektor $c := a \times b$ mit folgenden Eigenschaften:

1. $\|c\| = \|a\| \cdot \|b\| \cdot \sin \varphi \stackrel{\wedge}{=} \text{Flächeninhalt } F \text{ des von } a \text{ und } b \text{ aufgespannten Parallelogramms (siehe Bemerkung 5).}$
2. c ist orthogonal zu a und b .
3. a, b und c bilden ein Rechtssystem.

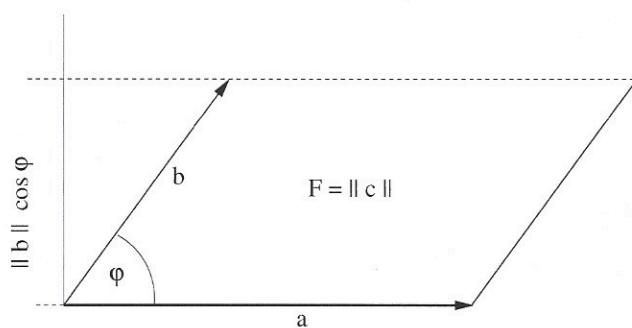


Abbildung 2.8: Das Vektorprodukt

Rechenregeln. 1. $a \times b = -b \times a$

$$2. a \times (b + c) = a \times b + a \times c$$

$$3. 0 \times a = a \times 0 = 0$$

$$4. a \times b = 0 \Leftrightarrow a \text{ und } b \text{ sind kollinear (d.h. } \varphi = 0)$$

Zunächst gilt:

$$\begin{aligned} e_i \times e_i &= 0, \quad i = 1, 2, 3 \\ e_1 \times e_2 &= -e_2 \times e_1 = e_3 \\ e_2 \times e_3 &= -e_3 \times e_2 = e_1 \\ e_3 \times e_1 &= -e_1 \times e_3 = e_2 \end{aligned}$$

Damit folgt:

$$\begin{aligned} a \times b &= (a_1 e_1 + a_2 e_2 + a_3 e_3) \times (b_1 e_1 + b_2 e_2 + b_3 e_3) \\ &= a_1 b_1 (e_1 \times e_2) + a_1 b_2 (e_1 \times e_2) + \dots \\ &= a_1 b_1 \cdot 0 + a_1 b_2 \cdot e_3 + \dots \\ &= (a_2 b_3 - a_3 b_2) e_1 + (a_3 b_1 - a_1 b_3) e_2 + (a_1 b_2 - a_2 b_1) e_3 \\ &= \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix} \end{aligned}$$

Bemerkung 4. Das Vektorprodukt kann mit Hilfe einer Determinante (siehe Abschnitt 1.5) auch angegeben werden in der Form

$$a \times b = \begin{vmatrix} e_1 & e_2 & e_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

Bemerkung 5. Betrachtet man die beiden Vektoren

$$a = \begin{pmatrix} a_1 \\ a_2 \\ 0 \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix}$$

die in der xy -Ebene liegen, so wird der Flächeninhalt F des von ihnen aufgespannten Parallelogramms gegeben durch

$$F = \|a \times b\| = \left\| \begin{pmatrix} a_1 \\ a_2 \\ 0 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix} \right\| = \left\| \begin{pmatrix} 0 \\ 0 \\ a_1 b_2 - a_2 b_1 \end{pmatrix} \right\| = |a_1 b_2 - a_2 b_1| = \left| \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \right|$$

d.h. der Flächeninhalt eines Parallelogramms im \mathbb{R}^2 wird bis auf das Vorzeichen durch die Determinante der Matrix, in deren Spalten man die Vektoren a und b einträgt, angegeben.

2.4 Das Spatprodukt

Definition 24. Das **Spatprodukt** (a, b, c) der drei Vektoren a , b und c ist die skalare Größe

$$(a, b, c) = \langle a \times b, c \rangle$$

Bemerkung 6. Der Betrag des Spatprodukts gibt das Volumen des von den Vektoren a , b und c aufgespannten Spats (=Parallelepipeds) an, denn

$$|\langle a \times b, c \rangle| = \underbrace{\|a \times b\|}_{\text{Grundfläche}} \cdot \underbrace{\|c\| \cdot \cos \angle(a \times b, c)}_{\text{Höhe}}$$

Für (a, b, c) erhalten wir

$$(a, b, c) = \left\langle \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}, \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} \right\rangle = a_1 b_2 c_3 + a_2 b_3 c_1 + a_3 b_1 c_2 - a_1 b_3 c_2 - a_2 b_1 c_3 - a_3 b_2 c_1$$

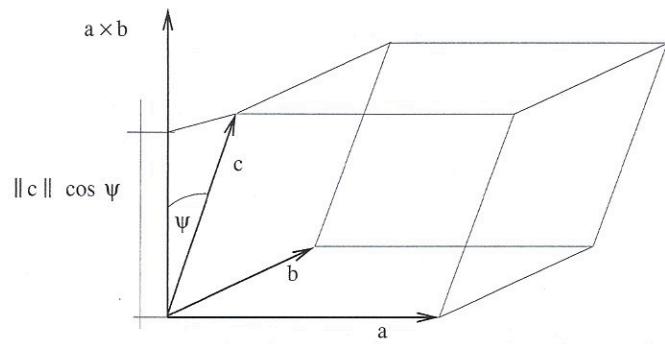


Abbildung 2.9: Das Spatprodukt

bzw. unter Verwendung der Determinante

$$(a, b, c) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

Kapitel 3

Allgemeine Vektorräume

3.1 Grundlagen

Definition 25. Sei K ein beliebiger Körper. Eine nichtleere Menge V heißt **Vektorraum über K** oder **K -Vektorraum (K -VR)**, wenn

1. für $a, b \in V$ die Summe $a + b$ stets erklärt ist und wieder in V liegt,
2. für $\lambda \in K$ und $a \in V$ das Produkt $\lambda \cdot a$ erklärt und wieder Element von V ist.

(*Abgeschlossenheit* bezüglich der Addition und der Multiplikation mit Skalaren.)
Dabei müssen folgende Rechengesetze gelten:

- Addition:** (A1) $a + b = b + a$ (Kommutativgesetz)
(A2) $a + (b + c) = (a + b) + c$ (Assoziativgesetz)
(A3) Es gibt einen **Nullvektor** $0 \in V$ so dass für alle $a \in V$ gilt $a + 0 = a$
(A4) Zu jedem Vektor $a \in V$ existiert der **inverse Vektor** (= **negative Vektor**)
 $(-a) \in V$ mit $a + (-a) = 0$

Multiplikation mit Skalaren: (M1) $\lambda \cdot (\mu \cdot a) = (\lambda\mu) \cdot a$, $\lambda, \mu \in K$

(M2) Für das Element $1 \in K$ und für alle $a \in V$ gilt $1 \cdot a = a$

- Distributivgesetz:** (D1) $\lambda \cdot (a + b) = \lambda \cdot a + \lambda \cdot b$ $a, b \in V, \lambda, \mu \in K$
(D2) $(\lambda + \mu) \cdot a = \lambda \cdot a + \mu \cdot a$

Schreibweise: Für die Elemente des Vektorraumes (= Vektoren) verwenden wir kleine lateinische Buchstaben ($a, b, c, \dots, v, w, \dots$), für Skalare aus dem Körper K kleine griechische Buchstaben ($\alpha, \beta, \dots, \lambda, \mu, \dots$).

Beispiel 35. Die Mengen \mathbb{R}^2 und \mathbb{R}^3 bilden einen Vektorraum (siehe ev. Kapitel 2).

Beispiel 36. Wir betrachten **geordnete n -tupel reeller Zahlen**

$$a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

und definieren (analog wie bei den bisher betrachteten Paaren im \mathbb{R}^2 und Tripeln im \mathbb{R}^3 , aber auch wie bei den $n \times 1$ -Matrizen):

Addition:

$$a + b = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} := \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}.$$

Multiplikation mit Skalaren

$$\lambda \cdot a = \lambda \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} := \begin{pmatrix} \lambda a_1 \\ \vdots \\ \lambda a_n \end{pmatrix} \quad \text{für } \lambda \in \mathbb{R}.$$

Damit wird die Menge

$$\mathbb{R}^n := \left\{ a \mid a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, a_i \in \mathbb{R} \right\}$$

zu einem \mathbb{R} -Vektorraum.

Beispiel 37. Die Menge der $m \times n$ -Matrizen mit den Operationen $A + B$ und $\lambda \cdot A$ ($A, B \in M(m \times n)$, $\lambda \in \mathbb{R}$) wie in Abschnitt 1.2.2 besprochen bildet einen \mathbb{R} -Vektorraum.

Beispiel 38. Sei $\mathcal{C}[a, b]$ die Menge der auf dem Intervall $[a, b] \subset \mathbb{R}$ stetigen Funktionen, d.h.

$$\mathcal{C}[a, b] = \{f : [a, b] \rightarrow \mathbb{R} \mid f \text{ stetig}\}.$$

Wir definieren:

Addition: Für $f, g \in \mathcal{C}[a, b]$ gilt

$$(f + g)(x) := f(x) + g(x) \quad \forall x \in [a, b]$$

Multiplikation mit Skalaren: Für $f \in \mathcal{C}[a, b]$ und $\lambda \in \mathbb{R}$ gilt

$$(\lambda \cdot f)(x) := \lambda \cdot f(x) \quad \forall x \in [a, b]$$

Bemerkung 7. Zwei Funktionen f und g heißen gleich auf dem Intervall $[a, b]$, wenn Ihre Funktionswerte für alle $x \in [a, b]$ gleich sind, d.h.

$$f = g \iff f(x) = g(x) \quad \forall x \in [a, b]$$

$\mathcal{C}[a, b]$ ist ein \mathbb{R} -Vektorraum.

Definition 26. Sei V ein K -Vektorraum. Eine nichtleere Teilmenge $U \subset V$ heißt **linearer Teilraum** von V , wenn für alle $\lambda \in K$ und $a, b \in U$ gilt:

$$a + b \in U, \quad \lambda \cdot a \in U$$

Bemerkung 8. 1. Ein linearer Teilraum ist wieder ein Vektorraum.

2. $U = \{0\}$, also die Menge, die nur aus dem Nullvektor besteht, ist ein linearer Teilraum jedes Vektorraumes.

Beispiel 39. $V = \mathbb{R}^2$

$$A = \left\{ \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \in \mathbb{R}^2 \mid 3a_1 - 4a_2 = 0 \right\} \quad \dots \quad \text{Gerade im } \mathbb{R}^2$$

Für $a, b \in A$: $a = \begin{pmatrix} a_1 \\ \frac{3}{4}a_1 \end{pmatrix}$, $b = \begin{pmatrix} b_1 \\ \frac{3}{4}b_1 \end{pmatrix}$, $\lambda \in \mathbb{R}$ gilt:

$$a + b = \begin{pmatrix} a_1 + b_1 \\ \frac{3}{4}(a_1 + b_1) \end{pmatrix} = \begin{pmatrix} d_1 \\ \frac{3}{4}d_1 \end{pmatrix} \in A, \quad \lambda \cdot a = \begin{pmatrix} \lambda a_1 \\ \frac{3}{4}\lambda a_1 \end{pmatrix} = \begin{pmatrix} c_1 \\ \frac{3}{4}c_1 \end{pmatrix} \in A$$

$\Rightarrow A$ ist linearer Teilraum.

Allgemein: Geraden im \mathbb{R}^2 und \mathbb{R}^3 , die den Ursprung enthalten, sind lineare Teilräume.

Beispiel 40. $V = \mathbb{R}^3$

$$A = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \mid 2x_1 - 2x_2 + x_3 = 0 \right\} \quad \dots \quad \text{Ebene im } \mathbb{R}^3$$

Für $x, y \in A$: $x = \begin{pmatrix} x_1 \\ x_2 \\ -2x_1 + 2x_2 \end{pmatrix}$, $y = \begin{pmatrix} y_1 \\ y_2 \\ -2y_1 + 2y_2 \end{pmatrix}$ und $\lambda \in \mathbb{R}$ gilt:

$$x + y = \begin{pmatrix} x_1 \\ x_2 \\ -2x_1 + 2x_2 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ -2y_1 + 2y_2 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ -2(x_1 + y_1) + 2(x_2 + y_2) \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \\ -2z_1 + 2z_2 \end{pmatrix} \in A$$

$$\lambda \cdot x = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ -2(\lambda x_1) + 2(\lambda x_2) \end{pmatrix} \in A$$

$\Rightarrow A$ ist linearer Teilraum

Allgemein: Ebenen im \mathbb{R}^3 , die den Ursprung enthalten, sind lineare Teilräume des \mathbb{R}^3 .

Beispiel 41. $\mathbb{P}_m \dots$ Vektorraum der Polynome, deren Grad höchstens gleich m ist.

$$\mathbb{P}_m = \{p_m(x) = a_m x^m + \dots + a_1 x + a_0 \mid m \in \mathbb{N}, a_0, \dots, a_m \in \mathbb{R}, x \in \mathbb{R}\}$$

$\mathbb{P}_m \subset \mathcal{C}(\mathbb{R})$ ist ein linearer Teilraum.

Beispiel 42.

$$C = \left\{ A = \begin{pmatrix} a_1 & -a_2 \\ a_2 & a_1 \end{pmatrix} \mid a_1, a_2 \in \mathbb{R} \right\} \subset M(2 \times 2)$$

ist ein linearer Teilraum, denn:

$$\begin{aligned} A_1 + A_2 &= \begin{pmatrix} a_1 & -a_2 \\ a_2 & a_1 \end{pmatrix} + \begin{pmatrix} a'_1 & -a'_2 \\ a'_2 & a'_1 \end{pmatrix} = \\ &= \begin{pmatrix} a_1 + a'_1 & -a_2 - a'_2 \\ a_2 + a'_2 & a_1 + a'_1 \end{pmatrix} = \begin{pmatrix} b_1 & -b_2 \\ b_2 & b_1 \end{pmatrix} \in C \\ \lambda \cdot A_1 &= \lambda \cdot \begin{pmatrix} a_1 & -a_2 \\ a_2 & a_1 \end{pmatrix} = \begin{pmatrix} \lambda a_1 & -\lambda a_2 \\ \lambda a_2 & \lambda a_1 \end{pmatrix} = \begin{pmatrix} c_1 & -c_2 \\ c_2 & c_1 \end{pmatrix} \in C \end{aligned}$$

Eigenschaften: Es sei V ein K -Vektorraum. Sind U, U' lineare Teilräume, so ist $U \cap U'$ wieder ein linearer Teilraum.

Beweis. $\lambda \in K, a, b \in U \cap U'$ d.h. $a, b \in U \wedge a, b \in U'$

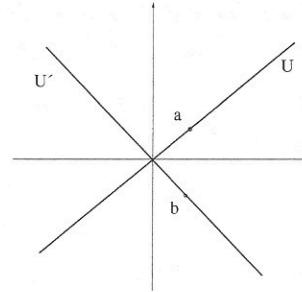
$$\Rightarrow a + b \in U, a + b \in U' \Rightarrow a + b \in U \cap U'$$

$$\lambda a \in U, \lambda a \in U' \Rightarrow \lambda a \in U \cap U'$$

Beispiel 43. $V = \mathbb{R}^2$; lineare Teilräume sind:

$$\begin{aligned} U &= \{x|x = \lambda \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \lambda \in \mathbb{R}\} \\ U' &= \{x|x = \lambda' \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \lambda' \in \mathbb{R}\}, \end{aligned}$$

Z.B. mit $a = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \in U$, $b = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \in U'$ gilt:
 $a + b = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \notin U, U', U \cup U'$



D.h. die Summe $a + b$ ist in keinem der Teilräume enthalten.

Bemerkung 9. Seien $U, U' \subset V$ Teilräume: Dann ist $U \cup U'$ i.a. kein linearer Teilraum von V .

Definition 27. U, U' seien lineare Teilräume von V . Ihre **Summe** ist definiert durch

$$U + U' = \{x|x = a + a', a \in U, a' \in U'\}$$

Bemerkung 10. $U + U'$ ist wieder ein Vektorraum.

Beweis. $\lambda \in K, x, y \in U + U'$

$$x = a + a', y = b + b', a, b \in U, a', b' \in U'$$

$$x + y = a + a' + b + b' = \underbrace{a + b}_{\in U} + \underbrace{a' + b'}_{\in U'} \in U + U'$$

$$\lambda x = \lambda(a + a') = \underbrace{\lambda a}_{\in U} + \underbrace{\lambda a'}_{\in U'} \in U + U'$$

Beispiel 44. (siehe Bsp. 43)

$$U = \left\{ x|x = \lambda \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \lambda \in \mathbb{R} \right\}, U' = \left\{ x|x = \lambda' \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \lambda' \in \mathbb{R} \right\}$$

$U + U'$ = Ebene, die von den Geraden U und U' aufgespannt wird ($\stackrel{\Delta}{=} \mathbb{R}^2$)

Beispiel 45. $V = \mathbb{R}^3$

$$U_1 = \left\{ x|x = \begin{pmatrix} a \\ b \\ 2b \end{pmatrix}, a, b \in \mathbb{R} \right\}, \quad U_2 = \left\{ x|x = \begin{pmatrix} a \\ b \\ a \end{pmatrix}, a, b \in \mathbb{R} \right\}, \quad U_3 = \left\{ x|x = \begin{pmatrix} 0 \\ 0 \\ a \end{pmatrix}, a \in \mathbb{R} \right\}$$

Die Menge U_1 stellt die Ebene $2y = z$ dar, die Menge U_2 die Ebene $x = z$ und die Menge U_3 die z -Achse (= Gerade) dar.

$$\begin{aligned} U_1 + U_2 &= \{y \in \mathbb{R}^3 | y = x' + x'', x' \in U_1, x'' \in U_2\} \\ &= \left\{ \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} | y_1 = a' + a'', y_2 = b' + b'', y_3 = 2b' + a'', a', b', a'', b'' \in \mathbb{R} \right\} \\ &= \mathbb{R}^3 \end{aligned}$$

Analog: $U_1 + U_3 = \mathbb{R}^3, U_2 + U_3 = \mathbb{R}^3$

3.2 Lineare Unabhängigkeit

Definition 28. Es seien Skalare $\lambda_1, \dots, \lambda_m \in K$ und Vektoren $a_1, \dots, a_m \in V$ gegeben. Ein Vektor a der Form

$$a = \lambda_1 a_1 + \dots + \lambda_m a_m$$

heißt **Linearkombination** der Vektoren a_1, \dots, a_m . Eine Linearkombination heißt **trivial**, wenn $\lambda_1 = \dots = \lambda_m = 0$, sonst heißt sie **nichttrivial** (d.h. mindestens ein $\lambda_k \neq 0$).

Definition 29. Sei $U \subset V$ eine nichtleere Teilmenge. Die Menge aller Linearkombinationen von Vektoren aus U

$$L(U) = \left\{ a = \sum_i \lambda_i a_i, \lambda_i \in K, a_i \in U \right\}$$

heißt der von den Vektoren $a_i \in U$ **aufgespannte Raum**. Ist $U = \{a_1, \dots, a_m\}$, so schreiben wir

$$L(U) = L(a_1, \dots, a_m).$$

Bemerkung 11. $L(U)$ ist wieder ein Vektorraum.

Beispiel 46. $V = \mathbb{R}^2$

$$a = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \Rightarrow L(a) = \{x \in \mathbb{R}^2 | x = \lambda \cdot a\}$$

$\hat{=}$ Gerade durch $P(2, 1)$ und $(0, 0)$.

Beispiel 47. $V = \mathbb{R}^3$

$$a_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, a_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \Rightarrow L(a_1, a_2) = \{x \in \mathbb{R}^3 | x = \lambda_1 a_1 + \lambda_2 a_2\}$$

$\hat{=}$ Ebene durch $P_1(1, 1, 0), P_2(0, 0, 1)$ und $(0, 0, 0)$.

Definition 30. Die Vektoren $a_1, \dots, a_m \in V$ heißen

1. **linear abhängig**, (l.a.), wenn es eine nichttriviale Linearkombination gibt mit

$$\lambda_1 a_1 + \dots + \lambda_m a_m = 0,$$

2. **linear unabhängig**, (l.u.), wenn sie nicht linear abhängig sind.

Beispiel 48. $V = \mathbb{R}^n$

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

Aus

$$\lambda_1 e_1 + \dots + \lambda_n e_n = 0$$

folgt

$$\begin{pmatrix} \lambda_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \lambda_2 \\ \vdots \\ 0 \end{pmatrix} + \dots + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\Rightarrow \lambda_1 = \dots = \lambda_n = 0 \Rightarrow e_1, \dots, e_n \text{ sind l.u.}$$

Beispiel 49. $V = \mathbb{R}^3$

$$a_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, a_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, a_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3 = 0 \Rightarrow \begin{array}{rcl} \lambda_1 & -\lambda_2 & +\lambda_3 = 0 \\ 2\lambda_1 & & +\lambda_3 = 0 \\ \lambda_2 & & +\lambda_3 = 0 \end{array}$$

$$\stackrel{\wedge}{=} \begin{pmatrix} 1 & -1 & 1 \\ 2 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -1 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 3/2 \end{pmatrix} \Rightarrow \lambda_1 = \lambda_2 = \lambda_3 = 0 \Rightarrow a_1, a_2, a_3 \text{ sind l.u.}$$

Beispiel 50. $V = \mathbb{R}^3$

$$a_1 = \begin{pmatrix} 1 \\ -2 \\ 8 \end{pmatrix}, a_2 = \begin{pmatrix} 4 \\ 5 \\ -3 \end{pmatrix}, a_3 = \begin{pmatrix} -2 \\ -9 \\ 19 \end{pmatrix}$$

Analog wie im Bsp. 49 ergibt sich hier die Koeffizientenmatrix zu

$$\begin{pmatrix} 1 & 4 & -2 \\ -2 & 5 & -9 \\ 8 & -3 & 19 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 4 & -2 \\ 0 & 13 & -13 \\ 0 & 0 & 0 \end{pmatrix}$$

\Rightarrow es existiert eine nichtriviale Lösung, z.B. $\lambda_1 = 2, \lambda_2 = -1, \lambda_3 = -1$, d.h.

$$2a_1 - a_2 - a_3 = 0 \Rightarrow a_1, a_2, a_3 \text{ sind l.a.!}$$

Satz 13. Seien $a_1, \dots, a_m \in V$, dann sind die folgenden Aussagen äquivalent:

1. Die Vektoren a_1, \dots, a_m sind linear abhängig.
2. Mindestens ein Vektor $a_i, 1 \leq i \leq m$, ist Linearkombination der übrigen, d.h.

$$\exists i \in \{1, \dots, m\} : a_i = \sum_{\substack{j=1 \\ j \neq i}}^m \lambda_j a_j$$

Beispiel 51. In Bsp. 50 gilt (u.a.) die Relation:

$$a_3 = 2a_1 + (-1)a_2$$

Definition 31. Eine nichtleere Teilmenge $U \subset V$ heißt **linear unabhängig**, wenn je endlich viele Vektoren aus U linear unabhängig sind, sonst heißt sie **linear abhängig**.

Beispiel 52.

$$V = \mathbb{P}_m = \{p_m(x) = a_m x^m + \dots + a_1 x + a_0\}$$

Frage: Ist $U = \{1, x, x^2, \dots, x^m\} \subset \mathbb{P}_m$ linear unabhängig?

$$\lambda_0 \cdot 1 + \lambda_1 \cdot x + \lambda_2 x^2 + \dots + \lambda_m x^m = 0 \quad \forall x \in \mathbb{R}$$

$$\Rightarrow \lambda_0 = 0, \dots, \lambda_m = 0 \Rightarrow U \text{ ist l.u.}$$

Beispiel 53. $V = \mathcal{C}[-\pi, \pi]$

$$f_1(x) = \sin x, f_2(x) = \cos x \in \mathcal{C}[-\pi, \pi]$$

sind linear unabhängig, denn aus

$$\lambda_1 \sin x + \lambda_2 \cos x = 0 \quad \forall x \in [-\pi, \pi]$$

$$\begin{aligned} \text{folgt z.B.} \quad & \text{für } x = 0: \quad \lambda_1 \cdot 0 + \lambda_2 \cdot 1 = 0 \Rightarrow \lambda_2 = 0 \\ & \text{für } x = \pi/2: \quad \lambda_1 \cdot 1 + \lambda_2 \cdot 0 = 0 \Rightarrow \lambda_1 = 0 \end{aligned}$$

Beispiel 54. $V = \mathcal{C}[a, b], U = \{e^x, e^{-x}, \cosh x\} \subset \mathcal{C}[a, b]$

$$\cosh = \frac{1}{2}e^x + \frac{1}{2}e^{-x} \quad \xrightarrow{\text{Satz 13}} \quad U \text{ ist l.a.}$$

3.3 Basis von Vektorräumen

Definition 32. Es sei V ein K-Vektorraum. Eine Teilmenge $U \subset V$ von linear unabhängigen Vektoren heißt **Basis** von V , wenn gilt $L(U) = V$.

Ein Vektorraum heißt **endlichdimensional** wenn er eine endliche Basis besitzt.

Satz 14. In einem endlichdimensionalen Vektorraum V hat jede Basis dieselbe Anzahl von Vektoren. Diese Zahl heißt **Dimension von V** , $\dim V$.

Beispiel 55. Der Vektorraum \mathbb{R}^n hat die Vektoren e_1, \dots, e_n als Basis $\Rightarrow \dim \mathbb{R}^n = n$

Definition 33. Die Vektoren $e_1, \dots, e_n \in \mathbb{R}^n$ von Beispiel 48 heißen **kanonische Einheitsvektoren**. Sie bilden die **Standardbasis \mathcal{K}** (oder **kanonische Basis**) des \mathbb{R}^n .

Beispiel 56. $\dim \mathbb{P}_m = m + 1$ da $\{1, x, \dots, x^m\}$ eine Basis ist.

Beispiel 57.

1. $V = \mathbb{R}^3, a \in \mathbb{R}^3, a \neq 0$. Dann stellt

$$L(a) = \{x \in \mathbb{R}^3 \mid x = \lambda \cdot a, \lambda \in \mathbb{R}\}$$

eine Gerade durch 0 und a dar.

2. $V = \mathbb{R}^3, a_1, a_2 \in \mathbb{R}^3, a_1, a_2 \text{ l.u.}$ Dann stellt

$$L(a_1, a_2) = \{x \in \mathbb{R}^3 \mid x = \lambda_1 \cdot a_1 + \lambda_2 \cdot a_2, \lambda_1, \lambda_2 \in \mathbb{R}\}$$

eine Ebene durch 0, a_1 und a_2 dar.

Im Fall der Geraden $A = L(a)$ stellt $\{a\}$ eine Basis von A dar, es ist $\dim A = 1$, im Fall der Ebene $B = L(a_1, a_2)$ stellt $\{a_1, a_2\}$ eine Basis von B dar, es ist $\dim B = 2$.

Beispiel 58.

$$C = \left\{ A = \begin{pmatrix} a_1 & -a_2 \\ a_2 & a_1 \end{pmatrix} \mid a_1, a_2 \in \mathbb{R} \right\} \subset M(2 \times 2)$$

ist ein Vektorraum (siehe Bsp. 42).

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in C$$

$\{A_1, A_2\}$ ist eine Basis von C , da

1. A_1, A_2 sind l.u., denn

$$\lambda_1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \lambda_1 & -\lambda_2 \\ \lambda_2 & \lambda_1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \Rightarrow \lambda_1 = \lambda_2 = 0$$

2. Eine beliebige Matrix A aus C

$$A = \begin{pmatrix} a_1 & -a_2 \\ a_2 & a_1 \end{pmatrix} \in C$$

kann dargestellt werden als

$$A = a_1 A_1 + a_2 A_2$$

d.h.

$$L(A_1, A_2) = C$$

Satz 15. Sei V ein Vektorraum mit der Basis $\mathcal{B} = \{v_1, \dots, v_n\}$. Dann kann jeder Vektor $v \in V$ mit eindeutig bestimmten Skalarfaktoren $\lambda_1, \dots, \lambda_n$ in der Form

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n$$

dargestellt werden.

Beweis. Da \mathcal{B} eine Basis ist, gilt $V = L(\mathcal{B})$, d.h. jeder Vektor v von V kann (zumindest auf eine Art) als Linearkombination der Vektoren v_1, \dots, v_n angegeben werden:

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n, \quad \lambda_k \in K \quad (3.1)$$

Wir wollen nun zeigen, dass die Darstellung (3.1) eindeutig ist.

Sei

$$v = \mu_1 v_1 + \dots + \mu_n v_n, \quad \mu_k \in K$$

eine weitere Darstellung von v . Subtraktion beider Gleichungen liefert unter Verwendung der Rechengesetze von Definition 25

$$(\lambda_1 - \mu_1)v_1 + \dots + (\lambda_n - \mu_n)v_n = 0$$

Da \mathcal{B} eine Basis ist, sind die Vektoren v_1, \dots, v_n linear unabhängig. Deshalb folgt

$$\lambda_1 - \mu_1 = 0, \dots, \lambda_n - \mu_n = 0 \iff \lambda_1 = \mu_1, \dots, \lambda_n = \mu_n$$

Damit gibt es genau eine Art den Vektor v als Linearkombination der Basisvektoren von \mathcal{B} anzugeben.

Kapitel 4

Lineare Abbildungen

4.1 Einleitung

Die linke Seite eines linearen Gleichungssystems $A \cdot x = b$ ist von der Gestalt

$$A \cdot x = \begin{pmatrix} a_{11}x_1 & + & \dots & + & a_{1n}x_n \\ & \vdots & & & \\ a_{m1}x_1 & + & \dots & + & a_{mn}x_n \end{pmatrix}$$

Die einzelnen Komponenten sind lineare Ausdrücke in den Variablen x_1, \dots, x_n . Man kann dies auch in Form einer Abbildung F schreiben:

$$F : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto y = A \cdot x = \begin{pmatrix} a_{11}x_1 & + & \dots & + & a_{1n}x_n \\ & \vdots & & & \\ a_{m1}x_1 & + & \dots & + & a_{mn}x_n \end{pmatrix}$$

Allgemein kann folgende Abbildung zwischen den \mathbb{K} -Vektorräumen V und W betrachtet werden:

$$\begin{aligned} F : V &\rightarrow W \\ a &\mapsto F(a) \end{aligned}$$

Beispiel 59.

1. Parameterdarstellung einer Geraden im \mathbb{R}^3 , die durch den Ursprung geht:

$$\begin{aligned} F : \mathbb{R} &\rightarrow \mathbb{R}^3 \\ \alpha &\mapsto F(\alpha) = \alpha \cdot u \end{aligned}$$

$u \in \mathbb{R}^3$, fest, (Richtungsvektor), $\alpha \in \mathbb{R}$, (Parameter).

$$F(\alpha + \beta) = (\alpha + \beta) \cdot u = \alpha \cdot u + \beta \cdot u = F(\alpha) + F(\beta)$$

$$F(\lambda \cdot \alpha) = (\lambda \cdot \alpha) \cdot u = \lambda \cdot (\alpha \cdot u) = \lambda \cdot F(\alpha) \quad \text{für } \lambda \in \mathbb{R}$$

2. Parameterdarstellung einer Ebene im \mathbb{R}^3 , die den Ursprung enthält:

$$\begin{aligned} F : \mathbb{R}^2 &\rightarrow \mathbb{R}^3 \\ \begin{pmatrix} \alpha \\ \beta \end{pmatrix} &\mapsto F\left(\begin{pmatrix} \alpha \\ \beta \end{pmatrix}\right) = \alpha \cdot u + \beta \cdot v \end{aligned}$$

$u, v \in \mathbb{R}^3$, fest; $\alpha, \beta \in \mathbb{R}$, (Parameter).

$$\begin{aligned} F\left(\begin{pmatrix}\alpha \\ \beta\end{pmatrix} + \begin{pmatrix}\alpha' \\ \beta'\end{pmatrix}\right) &= (\alpha + \alpha') \cdot u + (\beta + \beta') \cdot v = (\alpha u + \beta v) + (\alpha' u + \beta' v) \\ &= F\left(\begin{pmatrix}\alpha \\ \beta\end{pmatrix}\right) + F\left(\begin{pmatrix}\alpha' \\ \beta'\end{pmatrix}\right) \end{aligned}$$

$$F\left(\lambda \begin{pmatrix}\alpha \\ \beta\end{pmatrix}\right) = \lambda \alpha u + \lambda \beta v = \lambda \cdot F\left(\begin{pmatrix}\alpha \\ \beta\end{pmatrix}\right) \quad \text{für } \lambda \in \mathbb{R}$$

3. Skalarprodukt im \mathbb{R}^3 mit festem Vektor $b = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \in \mathbb{R}^3$:

$$\begin{aligned} F : \quad \mathbb{R}^3 &\rightarrow \mathbb{R} \\ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &\mapsto \langle x, b \rangle = x_1 b_1 + x_2 b_2 + x_3 b_3 \end{aligned}$$

$$F(x+y) = F(x) + F(y), \quad F(\lambda \cdot x) = \lambda \cdot F(x), \quad \lambda \in \mathbb{R}$$

4. Matrix $A \in M(n \times n)$ mal Spaltenvektor $x \in \mathbb{R}^n$:

$$\begin{aligned} F : \quad \mathbb{R}^n &\rightarrow \mathbb{R}^n \\ x &\mapsto F(x) = A \cdot x \end{aligned}$$

$$F(x+y) = F(x) + F(y), \quad F(\lambda \cdot x) = \lambda \cdot F(x), \quad \lambda \in \mathbb{R}$$

Definition 34. Wir betrachten die \mathbb{K} -Vektorräume V und W . Eine Abbildung

$$F : V \rightarrow W$$

heißt **linear**, wenn gilt

- (L1): $F(x+y) = F(x) + F(y)$ für alle $x, y \in V, \lambda \in \mathbb{K}$.
(L2): $F(\lambda \cdot x) = \lambda \cdot F(x)$

Bemerkung:

- (L1) \wedge (L2) \iff
 $(L): F(\lambda x + \mu y) = \lambda F(x) + \mu F(y) \quad \forall x, y \in V, \lambda, \mu \in \mathbb{K}$

- Eine lineare Abbildung respektiert also die Vektorraumstruktur von V und W :

1. Das Bild einer Summe ist die Summe der Bilder.
2. Das Bild eines Vielfachen ist das Vielfache des Bildes.

Eigenschaften:

1. $F(0) = 0$, denn:

$$\left. \begin{aligned} F(0) &= F(0) + 0 \\ F(0) &= F(0+0) = F(0) + F(0) \end{aligned} \right\} \Rightarrow F(0) = 0$$

2. $F(a-b) = F(a) - F(b)$, denn:

$$F(a-b) = F(a+(-b)) = F(a) + F(-b) = F(a) + (-1)F(b) = F(a) - F(b)$$

4.2 Kern und Bild einer linearen Abbildung

Definition 35.

1. $\text{Kern}(F) = \{v \in V | F(v) = 0\}$
2. $\text{Bild}(F) = \{F(v) | v \in V\}$

heißen **Kern** bzw. **Bild** der linearen Abbildung F .

Bemerkung:

$\text{Kern}(F) \subseteq V$ ist linearer Teilraum

$\text{Bild}(F) \subseteq W$ ist linearer Teilraum

Definition 36.

$$\mathcal{L}(V, W) := \{F : V \rightarrow W | F \text{ ist linear}\}$$

bezeichnet die Menge aller linearen Abbildungen von V nach W .

Beispiele:

1. Nullabbildung

$$\begin{aligned} \mathbb{O} : & \quad V \rightarrow W \\ & v \mapsto \mathbb{O}(v) = 0 \quad \forall v \in V \end{aligned}$$

2. Identische Abbildung

$$\begin{aligned} \text{id} : & \quad V \rightarrow V \\ & v \mapsto \text{id}(v) = v \quad \forall v \in V \end{aligned}$$

$$\text{Kern } (\mathbb{O}) = V, \quad \text{Bild } (\mathbb{O}) = 0$$

$$\text{Kern } (\text{id}) = 0, \quad \text{Bild } (\text{id}) = V$$

3. $w \in \mathbb{R}^3$ fest

$$\begin{aligned} F : & \quad \mathbb{R} \rightarrow \mathbb{R}^3 \\ & \alpha \mapsto \alpha \cdot w \end{aligned}$$

$$w = 0 : \quad \text{Kern } (F) = \mathbb{R}, \quad \text{Bild } (F) = 0$$

$$w \neq 0 : \quad \text{Kern } (F) = 0, \quad \text{Bild } (F) = \text{Gerade durch } 0$$

4. $w_1, w_2 \in \mathbb{R}^3$ fest, linear unabhängig

$$\begin{aligned} F : & \quad \mathbb{R}^2 \rightarrow \mathbb{R}^3 \\ & \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \mapsto \alpha w_1 + \beta w_2 \end{aligned}$$

$$\text{Kern } (F) = \{0\}, \quad \text{Bild } (F) = \text{Ebene durch } 0$$

5. $A \in M(m \times n)$

$$\begin{aligned} F : & \quad \mathbb{R}^n \rightarrow \mathbb{R}^m \\ & x \mapsto A \cdot x \end{aligned}$$

$$\text{Kern } (F) = \text{Lösungsmenge des linearen homogenen Gleichungssystems}$$

$$\begin{array}{rcl} a_{11}x_1 + \dots + a_{1n}x_n & = & 0 \\ \vdots & & \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n & = & 0 \end{array}$$

4.3 Summe, Vielfaches und Verknüpfung linearer Abbildungen

Definition 37. Für $F, G \in \mathcal{L}(V, W)$, $H \in \mathcal{L}(U, V)$ gilt:

$$\begin{aligned} F + G : & \quad V \rightarrow W \\ v \mapsto & (F + G)(v) := F(v) + G(v) \end{aligned}$$

$$\begin{aligned} \alpha \cdot F : & \quad V \rightarrow W \\ v \mapsto & (\alpha \cdot F)(v) := \alpha \cdot F(v), \quad \alpha \in \mathbb{K} \end{aligned}$$

$$\begin{aligned} F \circ H : & \quad U \rightarrow W \\ u \mapsto & (F \circ H)(u) := F(H(u)) \end{aligned}$$

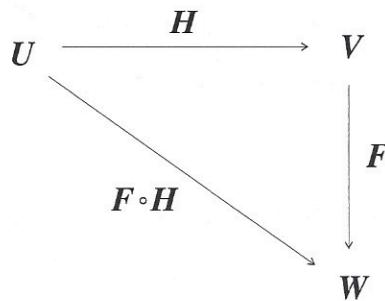


Abbildung 4.1: Komposition von Abbildungen

Bemerkung:

$F + G, \alpha \cdot F, F \circ H$ sind wieder lineare Abbildungen.

4.4 Injektive, surjektive und bijektive Abbildungen

Zur Erinnerung:

Seien X, Y beliebige Mengen. Die Vorschrift, die jedem Element $x \in X$ genau ein Element aus Y (wir nennen es $f(x)$) zuordnet, heißt **Abbildung** von X nach Y :

$$\begin{aligned} f : & \quad X \rightarrow Y \\ x \mapsto & f(x) \end{aligned}$$

Die Menge X heißt **Definitionsmenge**, die Menge Y nennt man **Bildbereich** oder **Wertebereich**. Für $M \subseteq X$ ist

$$f(M) := \{y \in Y \mid \text{es gibt ein } x \in M \text{ mit } y = f(x)\} \subset Y.$$

das **Bild** von M unter der Abbildung f .

Definition 38. Eine Abbildung $f : X \rightarrow Y$ heißt

- **surjektiv** falls $f(X) = Y$, d.h. falls es zu jedem $y \in Y$ ein $x \in X$ mit $y = f(x)$ gibt.
- **injektiv**, falls aus $x, x' \in X$ und $f(x) = f(x')$ stets $x = x'$ folgt.
- **bijektiv**, falls f surjektiv und injektiv ist.

Bemerkung 12. Zu einer bijektiven Abbildung $f : X \rightarrow Y$ existiert stets die **Umkehrabbildung**

$$f^{-1} : Y \rightarrow X$$

die jedem $y \in Y$ das durch $f(x) = y$ eindeutig bestimmte $x \in X$ zuordnet:

$$f^{-1}(y) = x$$

Satz 16. Für eine lineare Abbildung $F : V \rightarrow W$ gilt:

1. F ist injektiv $\Leftrightarrow \text{Kern } (F) = \{0\}$
2. F ist surjektiv $\Leftrightarrow \text{Bild } (F) = W$
3. Ist F bijektiv, so ist die Umkehrabbildung $F^{-1} : W \rightarrow V$ ebenfalls linear.

Beweis.

- Zu 1. $\alpha)$

$$\begin{array}{l} F \text{ injektiv} \\ v \in \text{Kern } (F) \end{array} \left. \right\} \Rightarrow \begin{array}{l} F(0) = 0 \\ F(v) = 0 \end{array} \xrightarrow[F \text{ inj.}]{\quad} v = 0$$

$$\Rightarrow \text{Kern } (F) = \{0\}$$

$\beta)$ $\text{Kern } (F) = \{0\}$
 $F(u) = F(v), u, v \in V$

$$\begin{aligned} 0 &= F(u) - F(v) = F(u - v) \Rightarrow u - v \in \text{Kern}(F) \Rightarrow u - v = 0 \\ &\Rightarrow u = v \Rightarrow F \text{ injektiv} \end{aligned}$$

- Zu 2. Wiederholung der Definition von "surjektiv".

- Zu 3. $u, w \in W$ mit $F(x) = u, F(y) = w$ bzw. $x = F^{-1}(u), y = F^{-1}(w)$

$$\Rightarrow F^{-1}(\alpha u + \beta w) = F^{-1}(\alpha F(x) + \beta F(y)) = F^{-1}(F(\alpha x + \beta y)) = \alpha x + \beta y = \alpha F^{-1}(u) + \beta F^{-1}(w)$$

Definition 39. Eine bijektive lineare Abbildung $F : V \rightarrow W$ heißt **Isomorphismus**. Die beiden \mathbb{K} -Vektorräume V und W heißen dann **isomorph unter F** .

4.5 Konstruktion linearer Abbildungen mittels Basen

Bisher haben wir lineare Abbildungen $F : V \rightarrow W$ durch Angabe des Bildes eines beliebigen Vektors $v \in V$ angegeben. Jetzt soll gezeigt werden, wie man jede lineare Abbildung $F : V \rightarrow W$ allein durch die Angabe der Bilder der Basisvektoren von V angeben kann. Es sei $\mathcal{B} = \{b_1, \dots, b_n\}$ eine Basis des Vektorraumes $V, V \neq \{0\}$, und $F \in \mathcal{L}(V, W)$.

Satz 17. Zur Berechnung von $F(v) \in W$ (für $v \in V$ beliebig) braucht man nur die n Werte $w_1 = F(b_1), \dots, w_n = F(b_n)$ zu kennen.

Beweis. Der Vektor $v \in V$ ist mit Hilfe von \mathcal{B} eindeutig darstellbar:

$$v = \alpha_1 b_1 + \dots + \alpha_n b_n, \quad \alpha_i \in \mathbb{K}$$

$$\Rightarrow F(v) = F(\alpha_1 b_1 + \dots + \alpha_n b_n) = \alpha_1 F(b_1) + \dots + \alpha_n F(b_n) = \alpha_1 w_1 + \dots + \alpha_n w_n$$

Sind umgekehrt die Elemente $w_1, \dots, w_n \in W$ vorgegeben, so läßt sich ein und nur ein $F \in \mathcal{L}(V, W)$ angeben, so daß $F(b_1) = w_1, \dots, F(b_n) = w_n$: Wir setzen einfach für beliebiges $v = \alpha_1 b_1 + \dots + \alpha_n b_n \in V$ fest:

$$F(v) = \alpha_1 w_1 + \dots + \alpha_n w_n.$$

Man kann also die Werte einer zu konstruierenden linearen Abbildung $F : V \rightarrow W$ auf einer Basis vorschreiben und erhält dann F durch "lineare Fortsetzung".

Satz 18. Sei $F : V \rightarrow W$ eine lineare Abbildung und $\mathcal{B} = \{b_1, \dots, b_n\}$ eine Basis von V und gelte $F(b_i) = w_i, i = 1, \dots, n$.

Dann ist

$$\begin{aligned}\text{Bild}(F) &= L(w_1, \dots, w_n) \\ &= L(F(b_1), \dots, F(b_n))\end{aligned}$$

d.h. das Bild von F wird von den Bildern der Basisvektoren aufgespannt.

Bemerkung 13. Damit kann man Eigenschaften einer linearen Abbildung untersuchen.

Beispiel 60. Man untersuche, welche der folgenden Abbildungen injektiv, surjektiv bzw. bijektiv sind.

a)

$$\begin{aligned}F_1 : \quad \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \end{pmatrix} &\mapsto \begin{pmatrix} 2x - y \\ x + 2y \end{pmatrix}\end{aligned}$$

- $\text{Kern}(F_1) : F_1 \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x - y \\ x + 2y \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies x = y = 0$

$$\implies \text{Kern}(F_1) = \{0\} \implies F_1 \text{ injektiv}$$

- $\text{Bild}(F_1) : F_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, F_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix} \dots \text{ Bilder der Basisvektoren}$

$$\text{Bild}(F_1) = L \left(\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix} \right) = \mathbb{R}^2 \implies F_1 \text{ surjektiv}$$

b)

$$\begin{aligned}F_2 : \quad \mathbb{R}^3 &\rightarrow \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \\ z \end{pmatrix} &\mapsto \begin{pmatrix} x \\ y \end{pmatrix} \quad (\text{Projektion in die } xy\text{-Ebene})\end{aligned}$$

- $\text{Kern}(F_2) : F_2 \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 0 \end{pmatrix} \implies x = y = 0, z \text{ beliebig(!)}$

$$\implies \text{Kern}(F_2) = z\text{-Achse} \implies F_2 \text{ nicht injektiv}$$

- $\text{Bild}(F_2) : F_2 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, F_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, F_2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \dots \text{ Bilder der Basisvektoren}$

$$\text{Bild}(F_2) = L \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right) = \mathbb{R}^2 \implies F_2 \text{ surjektiv}$$

c)

$$F_3 : \quad \mathbb{R} \rightarrow \mathbb{P}_3$$

$$a \mapsto F_3(a) := a + ax + ax^2 + ax^3$$

- $\text{Kern}(F_3) : \quad F_3(a) = 0 \implies a + ax + ax^2 + ax^3 = 0 \implies a = 0$
 $\implies \text{Kern}(F_3) = \{0\} \implies F_3 \text{ injektiv}$
- $\text{Bild}(F_3) : \quad F_3(1) = 1 + x + x^2 + x^3 \dots \text{Bild des Basisvektors}$
 $\text{Bild}(F_3) = \{p \in \mathbb{P}_3 \mid p(x) = \lambda(1 + x + x^2 + x^3), \lambda \in \mathbb{R}\} \neq \mathbb{P}_3 \implies F_3 \text{ nicht surjektiv}$

4.6 Koordinaten, Koordinaten-Abbildung

Sei V ein \mathbb{K} -Vektorraum mit der Basis $\mathcal{B} = \{b_1, \dots, b_n\}$.

Die Darstellung eines Vektors $v \in V$ bezüglich dieser Basis ist eindeutig (siehe Satz 15):

$$v = \alpha_1 b_1 + \dots + \alpha_n b_n, \quad \alpha_i \in \mathbb{K} \quad (4.1)$$

Definition 40.

1. Die **i-te Koordinate** eines Vektors $v \in V$ bezüglich der Basis \mathcal{B} ist das Skalar α_i in der Darstellung (4.1).
2. Der **Koordinatenvektor** eines Vektors v bezüglich der Basis \mathcal{B} ist der Spaltenvektor $c_{\mathcal{B}}(v)$ mit

$$c_{\mathcal{B}}(v) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{K}^n$$

wenn für v die Darstellung (4.1) gilt.

3. Sei $v = \alpha_1 b_1 + \dots + \alpha_n b_n$. Die Abbildung

$$c_{\mathcal{B}} : \quad V \rightarrow \mathbb{K}^n$$

$$v \mapsto c_{\mathcal{B}}(v) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

wird **Koordinaten-Abbildung** genannt. Sie ist bijektiv und weist jedem Vektor $v \in V$ seinen bezüglich der Basis \mathcal{B} eindeutig bestimmten Koordinatenvektor zu.

Beispiel 61. $V = \mathbb{R}^3$

$$\mathcal{B}' = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\} \dots \text{ Basis des } \mathbb{R}^3$$

$$v = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

$$\begin{array}{rcl} 1 & = & \alpha_1 + \alpha_2 \\ \Rightarrow 2 & = & \alpha_1 + \alpha_3 \quad \Rightarrow \quad \alpha_1 = 0, \quad \alpha_2 = 1, \quad \alpha_3 = 2 \\ 3 & = & \alpha_2 + \alpha_3 \end{array}$$

$$\Rightarrow c_{\mathcal{B}'}(v) = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$$

Beispiel 62. $V = \mathbb{P}_2$, $\mathcal{B}' = \{1, 1+t, 1+t+t^2\}$

$$\begin{aligned} v &= a + bt + ct^2 \in \mathbb{P}_2 \\ &= \alpha_1 \cdot 1 + \alpha_2 \cdot (1+t) + \alpha_3 (1+t+t^2) \end{aligned}$$

Der Koeffizientenvergleich liefert:

$$\left. \begin{array}{lcl} a &= \alpha_1 &+ \alpha_2 &+ \alpha_3 \\ b &= &\alpha_2 &+ \alpha_3 \\ c &= && \alpha_3 \end{array} \right\} \Rightarrow \begin{array}{lcl} \alpha_1 &= a - b \\ \alpha_2 &= b - c \\ \alpha_3 &= c \end{array} \Rightarrow c_{\mathcal{B}'}(v) = \begin{pmatrix} a - b \\ b - c \\ c \end{pmatrix}$$

Speziell gilt z.B.:

$$c_{\mathcal{B}'}(2 - 3t + 6t^2) = \begin{pmatrix} 5 \\ -9 \\ 6 \end{pmatrix}.$$

Der Vorteil von Koordinaten und Koordinatensystemen liegt in der Tatsache, dass man (z.B. auf Computern) konkret mit abstrakten Vektoren rechnen kann.

Bemerkung 14. Die Reihenfolge, in der die Vektoren in der Basis \mathcal{B} auftreten, bestimmt die Reihenfolge der Elemente des Koordinatenvektors.

4.7 Basiswechsel - Koordinatentransformation

Seien $\mathcal{B} = \{b_1, \dots, b_n\}$ und $\mathcal{B}' = \{b'_1, \dots, b'_n\}$ Basen von V .

Man schreibt die (neuen) Basisvektoren b'_i als Linearkombination der (alten) Basisvektoren b_i an:

$$b'_i = m_{1i}b_1 + m_{2i}b_2 + \dots + m_{ni}b_n, \quad i = 1, \dots, n,$$

d.h.

$$c_{\mathcal{B}}(b'_i) = \begin{pmatrix} m_{1i} \\ \vdots \\ m_{ni} \end{pmatrix}, \quad i = 1, \dots, n.$$

Sei nun ein Vektor $v \in V$ beliebig vorgegeben. Angenommen, wir haben die Koordinaten von v bezüglich \mathcal{B}' :

$$c_{\mathcal{B}'}(v) = \begin{pmatrix} \alpha'_1 \\ \vdots \\ \alpha'_n \end{pmatrix} \quad (\text{d.h. } v = \alpha'_1 b'_1 + \dots + \alpha'_n b'_n)$$

und wollen die Koordinaten bezüglich \mathcal{B} berechnen, d.h. wir suchen

$$c_{\mathcal{B}}(v) = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

Es gilt:

$$\begin{aligned}
 c_{\mathcal{B}}(v) &= c_{\mathcal{B}}(\alpha'_1 b'_1 + \dots + \alpha'_n b'_n) \\
 &= \alpha'_1 c_{\mathcal{B}}(b'_1) + \dots + \alpha'_n c_{\mathcal{B}}(b'_n) \\
 &= \alpha'_1 \begin{pmatrix} m_{11} \\ \vdots \\ m_{n1} \end{pmatrix} + \dots + \alpha'_n \begin{pmatrix} m_{1n} \\ \vdots \\ m_{nn} \end{pmatrix} \\
 &= \begin{pmatrix} \alpha'_1 m_{11} + \dots + \alpha'_n m_{1n} \\ \vdots \\ \alpha'_1 m_{n1} + \dots + \alpha'_n m_{nn} \end{pmatrix} \\
 &= \underbrace{\begin{pmatrix} m_{11} & \dots & m_{1n} \\ \vdots & & \vdots \\ m_{n1} & \dots & m_{nn} \end{pmatrix}}_{=:M} \cdot \underbrace{\begin{pmatrix} \alpha'_1 \\ \vdots \\ \alpha'_n \end{pmatrix}}_{c'_{\mathcal{B}}(v)}
 \end{aligned}$$

d.h. es gilt $c_{\mathcal{B}}(v) = M \cdot c'_{\mathcal{B}}(v)$ mit $M = (m_{ij}) \in M(n \times n)$.

Behauptung: Die Matrix M ist regulär.

Beweis. Sei $v \in V$ mit $c'_{\mathcal{B}}(v) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und es gelte $M \cdot x = 0$, d.h. die \mathcal{B} -Koordinaten von v sind alle Null, d.h.

$$v = 0 \cdot b_1 + \dots + 0 \cdot b_n = 0.$$

Dann folgt $v = 0 = x_1 b'_1 + \dots + x_n b'_n \Rightarrow x_1 = \dots = x_n = 0$ da $\{b'_1, \dots, b'_n\}$ linear unabhängig. Aus $M \cdot x = 0$ folgt also $x = 0$. D.h. das homogene Gleichungssystem $Mx = 0$ besitzt nur die triviale Lösung $x = 0$. Nach Folgerung 2 von 1.3.5 ist damit M regulär und M^{-1} existiert.

Wir erhalten die Formel für die Berechnung des Koordinatenvektors $c'_{\mathcal{B}}(v)$ bezüglich der (neuen) Basis \mathcal{B}' aus den Koordinatenvektor $c_{\mathcal{B}}(v)$ bezüglich der (alten) Basis \mathcal{B} :

$$c'_{\mathcal{B}}(v) = M^{-1} c_{\mathcal{B}}(v) \quad (\text{Koordinatentransformation})$$

Definition 41. Die Matrix $T := M^{-1}$ heißt **Transformationsmatrix** des Basiswechsels $\mathcal{B} \mapsto \mathcal{B}'$.

Beispiel 63. $V = \mathbb{R}^3$

$$\mathcal{B} = \{e_1, e_2, e_3\}, \quad \mathcal{B}' = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\}, \quad v = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = c_{\mathcal{B}}(v)$$

Frage: Wie sieht der Koordinatenvektor $c'_{\mathcal{B}}(v)$ aus?

$$M = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad M^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix}$$

$$\Rightarrow c'_{\mathcal{B}}(v) = M^{-1} \cdot c_{\mathcal{B}}(v) = \frac{1}{2} \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \quad (\text{siehe Bsp. 61})$$

Beispiel 64. $V = \mathbb{P}_2$

$$\mathcal{B} = (1, t, t^2), \quad \mathcal{B}' = (1, 1+t, 1+t+t^2)$$

$$\left. \begin{array}{lcl} b'_1 & = & 1 \\ b'_2 & = & 1+t \\ b'_3 & = & 1+t+t^2 \end{array} \right\} = \left. \begin{array}{l} 1 \cdot 1 + 0 \cdot t + 0 \cdot t^2 \\ 1 \cdot 1 + 1 \cdot t + 0 \cdot t^2 \\ 1 \cdot 1 + 1 \cdot t + 1 \cdot t^2 \end{array} \right\} \Rightarrow M = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad M^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

Für $v = 2 - 3t + 6t^2 \in \mathbb{P}_2$ gilt $c_{\mathcal{B}}(v) = \begin{pmatrix} 2 \\ -3 \\ 6 \end{pmatrix}$

$$\Rightarrow c_{\mathcal{B}'}(v) = M^{-1}c_{\mathcal{B}}(v) = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -3 \\ 6 \end{pmatrix} = \begin{pmatrix} 5 \\ -9 \\ 6 \end{pmatrix}.$$

Allgemein gilt für $v = a + bt + ct^2 \in \mathbb{P}_2$: $c_{\mathcal{B}}(v) = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$ und damit

$$c_{\mathcal{B}'}(v) = M^{-1} \cdot c_{\mathcal{B}}(v) = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} a-b \\ b-c \\ c \end{pmatrix} \quad (\text{siehe Bsp. 62})$$

4.8 Die Matrix einer linearen Abbildung

Von früher hat man folgende Aussagen:

- Jeder endlichdimensionale Vektorraum V ist gleichwertig zum \mathbb{R}^n oder \mathbb{C}^n (vgl. Koordinatenisomorphismus, Def. 40.3).
- Eine Matrix $A \in M(m \times n)$ definiert eine lineare Abbildung zwischen dem \mathbb{K}^n und dem \mathbb{K}^m .

Damit kann man vermuten, dass eine lineare Abbildung im wesentlichen nichts anderes ist als eine Multiplikation mit einer Matrix.

Gegeben:

- \mathbb{K} -Vektorraum V mit der Basis $\mathcal{A} = \{v_1, \dots, v_n\}$
- \mathbb{K} -Vektorraum W mit der Basis $\mathcal{B} = \{w_1, \dots, w_m\}$
- $F : V \rightarrow W$ linear, definiert durch

$$F(v_j) = a_{1j}w_1 + a_{2j}w_2 + \dots + a_{mj}w_m, \quad j = 1, \dots, n$$

Wir setzen fest

$$M_{\mathcal{B}}^{\mathcal{A}}(F) := A = (a_{ij}),$$

d.h. die Spaltenvektoren der die Abbildung F darstellenden Matrix $M_{\mathcal{B}}^{\mathcal{A}}(F)$ sind die Koordinatenvektoren der Bilder der Basisvektoren v_1, \dots, v_n .

Sei weiters $v = x_1v_1 + \dots + x_nv_n \in V$, d.h. $c_{\mathcal{A}}(v) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ ist der Koordinatenvektor von v .

Wir berechnen nun $F(v)$:

$$\begin{aligned} F(v) &= F(x_1v_1 + \dots + x_nv_n) \\ &= x_1F(v_1) + \dots + x_nF(v_n) \\ &= x_1(a_{11}w_1 + \dots + a_{m1}w_m) + \dots \\ &\quad \dots + x_n(a_{1n}w_1 + \dots + a_{mn}w_m) \\ &= (a_{11}x_1 + \dots + a_{1n}x_n) \cdot w_1 + \dots \\ &\quad \dots + (a_{m1}x_1 + \dots + a_{mn}x_n)w_m \\ &= y_1w_1 + \dots + y_mw_m \end{aligned}$$

wenn

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = c_{\mathcal{B}}(F(v))$$

den Koordinatenvektor des Bildes von $v \in V$ bezeichnet. Man erhält also

$$y = Ax.$$

Daher ist folgende Vorgangsweise empfehlenswert:

Statt $F(v)$ direkt zu berechnen kann man in 3 Schritten vorgehen:

1. Berechne den Koordinatenvektor $x = c_{\mathcal{A}}(v)$ von $v \in V$.
2. Berechne $y = Ax$
3. Verwende y als Koordinatenvektor von $w \in W$, wobei w dasselbe ist wie $F(v)$.

Definition 42. Man nennt $M_{\mathcal{B}}^{\mathcal{A}}(F)$ die **darstellende Matrix**.

Satz 19. Seien V und W jeweils \mathbb{K} -Vektorräume mit den Basen $\mathcal{A} = \{v_1, \dots, v_n\}$ und $\mathcal{B} = \{w_1, \dots, w_m\}$ sowie $F : V \rightarrow W$ eine lineare Abbildung.

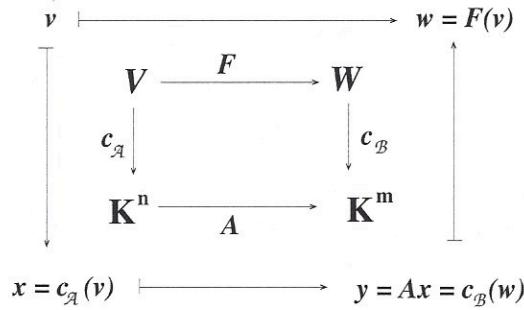
Man erhält $w = F(v) \in W$ mit $v = x_1v_1 + \dots + x_nv_n \in V$ in der Form

$$w = y_1w_1 + \dots + y_mw_m$$

wobei der Koordinatenvektor $y = c_{\mathcal{B}}(w)$ gegeben ist durch

$$y = A \cdot x = M_{\mathcal{B}}^{\mathcal{A}}(F) \cdot x$$

mit $x = c_{\mathcal{A}}(v)$ (siehe Abb. 4.2).

Abbildung 4.2: Berechnung von $F(v)$ **Beispiel 65.**

$$F : \quad \mathbb{R}^2 \rightarrow \mathbb{R}$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto ax_1 + bx_2, \quad a, b \in \mathbb{R} \quad \text{fest}$$

Mit den Standardbasen $\mathcal{A} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$ und $\mathcal{B} = \{1\}$ bestimmt man nun $A = M_{\mathcal{B}}^{\mathcal{A}}(F)$.

$$F \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) = a, \quad F \left(\begin{pmatrix} 0 \\ 1 \end{pmatrix} \right) = b \quad \Rightarrow \quad M_{\mathcal{B}}^{\mathcal{A}}(F) = (a, b) \in M(1 \times 2) \quad \Rightarrow$$

$$\Rightarrow \quad y = A \cdot x = (a, b) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = ax_1 + bx_2$$

Beispiel 66.

$$F : \quad \mathbb{P}_2 \rightarrow \mathbb{P}_1$$

$$v(t) \mapsto \frac{v(t) - v(0)}{t}$$

Wir verwenden die Basen $\mathcal{A} = \{1+t, t+t^2, 1+t^2\}$ und $\mathcal{B} = \{1+t, 1-t\}$ und erhalten wegen $F(a+bt+ct^2) = \frac{a+bt+ct^2-a}{t} = b+ct$:

$$\left. \begin{array}{rcl} F(1+t) & = & 1 \\ F(t+t^2) & = & 1+t \\ F(1+t^2) & = & t \end{array} \right\} = \left. \begin{array}{rcl} & = & \frac{1}{2}(1+t) + \frac{1}{2}(1-t) \\ & = & 1(1+t) + 0(1-t) \\ & = & \frac{1}{2}(1+t) - \frac{1}{2}(1-t) \end{array} \right\} \Rightarrow A = M_{\mathcal{B}}^{\mathcal{A}}(F) = \begin{pmatrix} 1/2 & 1 & 1/2 \\ 1/2 & 0 & -1/2 \end{pmatrix} \in M(2 \times 3)$$

Zur Probe berechnen wir $F(1+t+t^2)$ auf zwei verschiedene Arten:

a) $v_1 = 1+t+t^2 = \frac{1}{2}(1+t) + \frac{1}{2}(t^2+t) + \frac{1}{2}(1+t^2) \Rightarrow$

$$\Rightarrow x = c_{\mathcal{A}}(v_1) = \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \end{pmatrix} \Rightarrow y = M_{\mathcal{B}}^{\mathcal{A}}(F) \cdot x = \begin{pmatrix} 1/2 & 1 & 1/2 \\ 1/2 & 0 & -1/2 \end{pmatrix} \begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$F(v_1) = w = y_1 w_1 + y_2 w_2 = 1 \cdot (1+t) + 0 \cdot (1-t) = 1+t$$

b) Direkt: $F(v_1) = F(1+t+t^2) = 1+t$.

Beispiel 67. Die Null-Abbildung:

$$\mathbb{O} : V \rightarrow W$$

$$v \mapsto 0$$

$$M_{\mathcal{B}}^{\mathcal{A}}(\mathbb{O}) = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix} = 0 \in M(m \times n) \quad \text{Nullmatrix}$$

Beispiel 68.

$$F : \mathbb{P}_3 \rightarrow \mathbb{P}_3$$

$$v(t) \mapsto F(v) := v'(t) = \frac{d}{dt} v(t) \quad (\text{Ableitung})$$

Basen: $\mathcal{A} = \mathcal{B} = \{1, t, t^2, t^3\}$

$$\left. \begin{array}{rcl} F(1) & = & 0 \\ F(t) & = & 1 \\ F(t^2) & = & 2t \\ F(t^3) & = & 3t^2 \end{array} \right\} = \left. \begin{array}{rcl} 0 \cdot 1 + 0 \cdot t + 0 \cdot t^2 + 0 \cdot t^3 \\ 1 \cdot 1 + 0 \cdot t + 0 \cdot t^2 + 0 \cdot t^3 \\ 0 \cdot 1 + 2 \cdot t + 0 \cdot t^2 + 0 \cdot t^3 \\ 0 \cdot 1 + 0 \cdot t + 3 \cdot t^2 + 0 \cdot t^3 \end{array} \right\} \Rightarrow A = M_{\mathcal{B}}^{\mathcal{A}}(F) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Berechnung von $F(2 - 3t + t^2 - t^3)$ auf zwei Wegen:

a) Über die darstellende Matrix:

$$x = c_{\mathcal{A}}(v) = \begin{pmatrix} 2 \\ -3 \\ 1 \\ -1 \end{pmatrix}, \quad y = Ax = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ -3 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -3 \\ 2 \\ -3 \\ 0 \end{pmatrix}$$

$$F(v) = -3 \cdot 1 + 2 \cdot t + (-3) \cdot t^2 + 0 \cdot t^3 = -3 + 2t - 3t^2$$

$$\text{b) Direkt: } v'(t) = \frac{d}{dt}(2 - 3t + t^2 - t^3) = -3 + 2t - 3t^2.$$

Satz 20.

1. **Komposition von Abbildungen:**

Seien V, W und U \mathbb{K} -Vektorräume mit den Basen \mathcal{A}, \mathcal{B} und \mathcal{C} und seien

$$F : V \rightarrow W, \quad G : W \rightarrow U$$

lineare Abbildungen. Dann gilt:

$$M_{\mathcal{C}}^{\mathcal{A}}(G \circ F) = M_{\mathcal{C}}^{\mathcal{B}}(G) \cdot M_{\mathcal{B}}^{\mathcal{A}}(F).$$

D.h. die Komposition von Abbildungen wird durch das Produkt der darstellenden Matrizen beschrieben.

2. **Inverse eines Isomorphismus:**

Sei $F : V \rightarrow W$ ein Isomorphismus (= lineare, bijektive Abbildung) und seien \mathcal{A} bzw. \mathcal{B} Basen von V und W . Dann gilt:

(i) $M_{\mathcal{B}}^{\mathcal{A}}(F)$ ist regulär.

(ii) $M_{\mathcal{A}}^{\mathcal{B}}(F^{-1}) = (M_{\mathcal{B}}^{\mathcal{A}}(F))^{-1}$

d.h. die Matrix der Umkehrabbildung (= inversen Abbildung) F^{-1} ergibt sich als Inverse der die Abbildung F darstellenden Matrix.

Satz 21. Sei $F : V \rightarrow W$ eine lineare Abbildung, \mathcal{A} eine Basis von V , $\dim V = n$, \mathcal{B} eine Basis von W , $\dim W = m$, sowie $M_{\mathcal{B}}^{\mathcal{A}}(F) \in M(m \times n)$ die die Abbildung F darstellende Matrix.

Dann gilt:

1. F ist injektiv $\Leftrightarrow \text{Kern}(F) = \{0\} \Leftrightarrow \text{rang } M_{\mathcal{B}}^{\mathcal{A}}(F) = n$
2. F ist surjektiv $\Leftrightarrow \text{Bild}(F) = W \Leftrightarrow \text{rang } M_{\mathcal{B}}^{\mathcal{A}}(F) = m$
3. F ist bijektiv $\Leftrightarrow \text{rang } M_{\mathcal{B}}^{\mathcal{A}}(F) = n \Leftrightarrow M_{\mathcal{B}}^{\mathcal{A}}(F)$ ist regulär
($\dim V = \dim W = n$)

Bemerkung 15. Unter Verwendung der darstellenden Matrix $A := M_{\mathcal{B}}^{\mathcal{A}}(F)$ kann das Problem der Bestimmung einer Basis von $\text{Bild}(F)$ bzw. $\text{Kern}(F)$ für eine lineare Abbildung $F : V \rightarrow W$ zwischen den Vektorräumen V und W mit den Basen \mathcal{A} bzw. \mathcal{B} wie folgt behandelt werden.

- Bestimmung von $\text{Kern}(F)$:

1. Sei $v \in \text{Kern}(F)$ mit dem Koordinatenvektor $x = c_{\mathcal{A}}(v)$
2. $F(v) = 0 \Leftrightarrow A \cdot x = 0$. D.h. man muß das lineare Gleichungssystem $Ax = 0$ lösen.
3. $v = c_{\mathcal{A}}^{-1}(x)$ liegt im Kern von F . Ist $(x_1, \dots, x_k) \in \mathbb{K}^n$ eine Basis des Lösungsraumes von $Ax = 0$, so stellt $(c_{\mathcal{A}}^{-1}(x_1), \dots, c_{\mathcal{A}}^{-1}(x_k))$ eine Basis von $\text{Kern}(F)$ dar.

- Bestimmung von $\text{Bild}(F)$:

1. Das Bild von F wird aufgespannt von den Vektoren $c_{\mathcal{B}}^{-1}(u_1), \dots, c_{\mathcal{B}}^{-1}(u_n)$, wobei die Vektoren u_1, \dots, u_n die Spaltenvektoren von $M_{\mathcal{B}}^{\mathcal{A}}(F)$ sind.
2. Gesucht ist nun eine Basis des von den Vektoren u_1, \dots, u_n aufgespannten Vektorraumes. Dazu bildet man die Matrix $\tilde{A} = (M_{\mathcal{B}}^{\mathcal{A}}(F))^t$ und führt diese auf Zeilenstufenform über.
3. Die nun vom Nullvektor verschiedenen Zeilenvektoren $\tilde{u}_1, \dots, \tilde{u}_l$ dieser Matrix bilden eine Basis des
 - von den Zeilenvektoren von $(M_{\mathcal{B}}^{\mathcal{A}}(F))^t$ aufgespannten Vektorraumes,
 - des von den Spaltenvektoren von $M_{\mathcal{B}}^{\mathcal{A}}(F)$ aufgespannten Vektorraumes.
4. Damit erhält man eine Basis von $\text{Bild}(F)$ durch $(c_{\mathcal{B}}^{-1}(\tilde{u}_1), \dots, c_{\mathcal{B}}^{-1}(\tilde{u}_l))$.

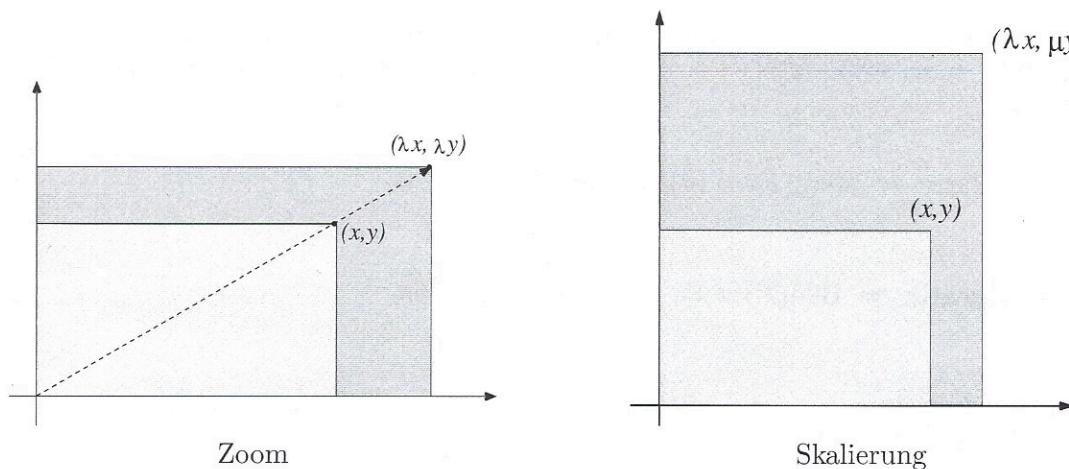
4.9 Spezielle lineare Abbildungen im \mathbb{R}^2

1. **Zoom** (= zentrische Streckung bzw. Stauchung)

$$Z_{\lambda} : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \lambda \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \lambda \cdot x \\ \lambda \cdot y \end{pmatrix}, \quad \lambda \in \mathbb{R}$$

Darstellende Matrix (bzgl. der kanonischen Basis):

$$M_{Z_{\lambda}} = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$$



2. Skalierung

zwei unabhängige Skalierungsfaktoren: $\lambda \in \mathbb{R}$ in x -Richtung,
 $\mu \in \mathbb{R}$ in y -Richtung

$$S_{\lambda,\mu} : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \lambda x \\ \mu y \end{pmatrix}$$

$$M_{S_{\lambda,\mu}} = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$$

3. Spiegelung an einer Koordinatenachse

- an der x -Achse

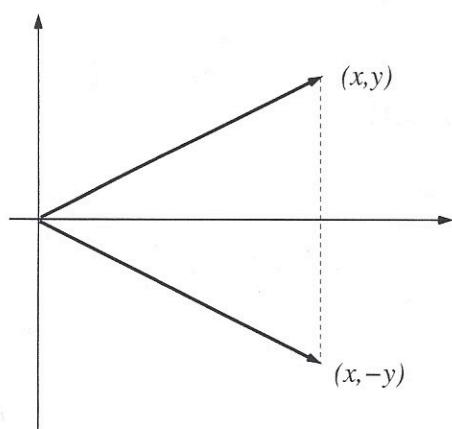
$$Sp_x : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ -y \end{pmatrix}$$

$$M_{Sp_x} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

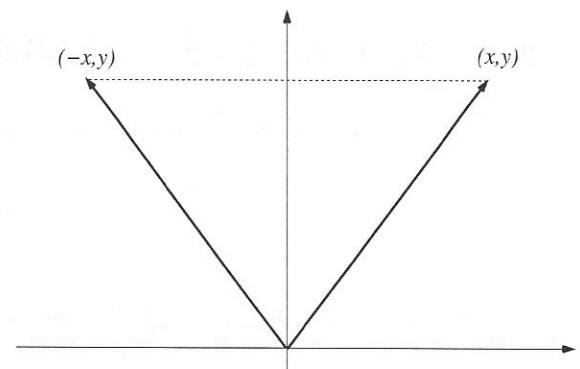
- an der y -Achse

$$Sp_y : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} -x \\ y \end{pmatrix}$$

$$M_{Sp_y} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$



Spiegelung an der x -Achse



Spiegelung an der y -Achse

4. Scherung

- in Richtung der x -Achse

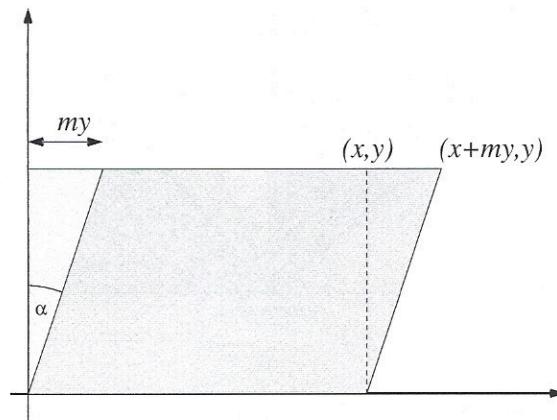
$$Sh_x(\alpha) : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x + m \cdot y \\ y \end{pmatrix} \text{ mit } m = \tan \alpha$$

$$M_{Sh_x} = \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix}$$

- in Richtung der y -Achse

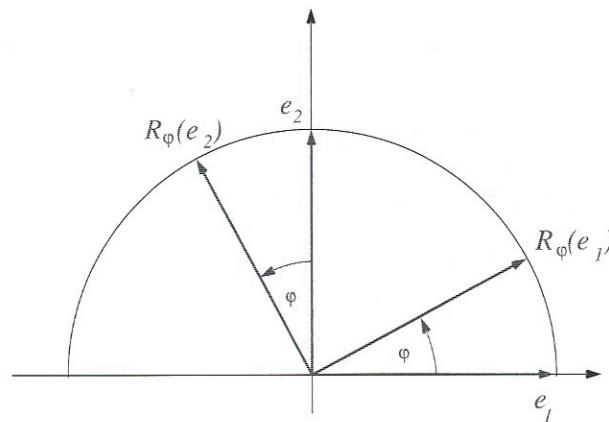
$$Sh_y(\beta) : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ y + n \cdot x \end{pmatrix} \text{ mit } n = \tan \beta$$

$$M_{Sh_y} = \begin{pmatrix} 1 & 0 \\ n & 1 \end{pmatrix}$$



Scherung in Richtung der x -Achse

5. Rotation (Drehung) um den Ursprung



Die darstellende Matrix (= Drehmatrix) lautet

$$M_{R_\varphi} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

6. Orthogonale Projektion

- auf die x -Achse

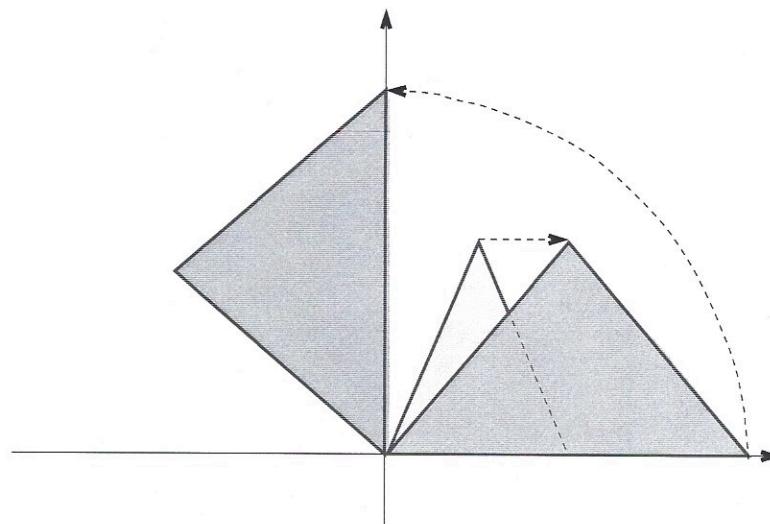
$$\Pi_x : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ 0 \end{pmatrix}$$

- auf die y -Achse

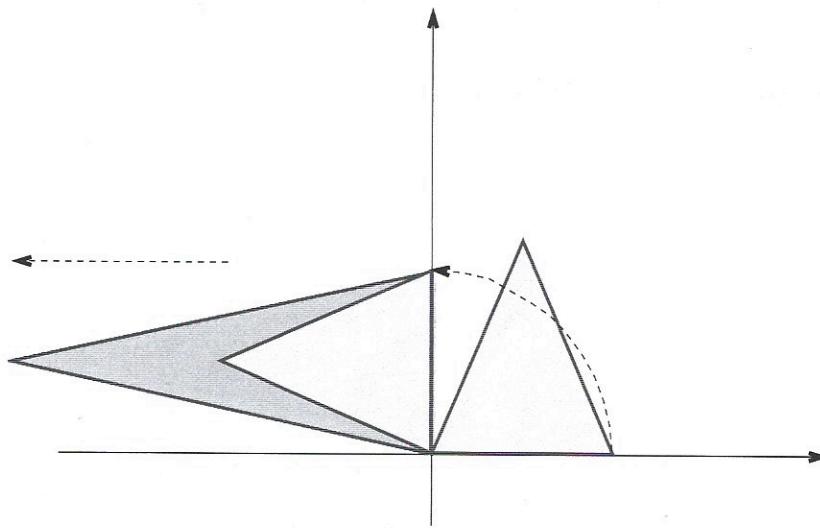
$$\Pi_y : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ y \end{pmatrix}$$

$$M_{\Pi_x} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad M_{\Pi_y} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Beispiel 69. Ein Dreieck wird zuerst in x -Richtung um den Faktor $\lambda = 2$ gestreckt und dann um $\varphi = \pi/2$ gedreht.



Wird hingegen zunächst gedreht und dann in x -Richtung gestreckt, so ergibt sich das folgende Bild.



Die darstellenden Matrizen sind

$$M_1 = \begin{pmatrix} 0 & -1 \\ 2 & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix}$$

Satz 22. Jede bijektive lineare Abbildung im \mathbb{R}^2 kann angegeben werden als Komposition von Skalierungen, Scherungen und von Spiegelungen.

D.h.: Jede invertierbare Matrix $A \in M(2 \times 2)$ kann aufgefasst werden als Produkt von Matrizen vom Typ

$$M_{S_{\lambda,\mu}}, M_{Sh_x}, M_{Sh_y}, M_{Sp_x} \text{ und } M_{Sp_y}$$

Beispiel 70.

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}}_{\text{Scherung}} \cdot \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}}_{\text{Spiegelung}} \cdot \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}}_{\text{Skalierung}} \cdot \underbrace{\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}}_{\text{Scherung}}$$

4.10 Spezielle lineare Abbildungen im \mathbb{R}^3

1. Skalierung

$$S_{\kappa,\lambda,\mu} : \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} \kappa x \\ \lambda y \\ \mu z \end{pmatrix}$$

$$M_{S_{\kappa,\lambda,\mu}} = \begin{pmatrix} \kappa & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \mu \end{pmatrix}$$

2. Rotation um die x -Achse

$$M_{R_{x,\varphi}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix}$$

3. Rotation um die y -Achse

$$M_{R_{y,\varphi}} = \begin{pmatrix} \cos \varphi & 0 & -\sin \varphi \\ 0 & 1 & 0 \\ \sin \varphi & 0 & \cos \varphi \end{pmatrix}$$

4. Rotation um die z -Achse

$$M_{R_{z,\varphi}} = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

5. Spiegelung an der xy -Ebene

$$M_{Sp_{xy}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

6. Spiegelung an der yz -Ebene

$$M_{Sp_{yz}} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

7. Spiegelung an der xz -Ebene

$$M_{Sp_{xz}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

4.11 Affine Abbildungen im \mathbb{R}^2 und homogene Koordinaten

Die Abbildungen

1. Skalierung
2. Spiegelung an einer Koordinatenachse
3. Drehung um den Ursprung
4. Scherung
5. orthogonale Projektion auf eine Koordinatenachse

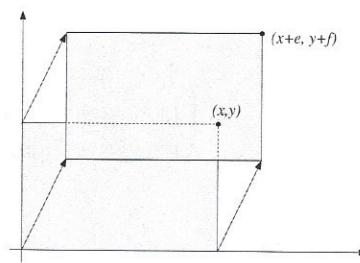
sind lineare Abbildungen.

Bemerkung 16. Alle Abbildungen bis auf die letzte sind bijektiv.

Weitere Abbildungen:

6. Translation

$$T_{(e,f)} : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} = \begin{pmatrix} x+e \\ y+f \end{pmatrix}$$



7. Drehung um einen anderen Punkt als den Ursprung
8. Spiegelung an einer Achse, die nicht durch den Ursprung geht
9. Skalierung relativ zu einem Punkt, der nicht der Ursprung ist

Bemerkung 17. Die Abbildungen 6 bis 9 sind **nicht linear**, da der Ursprung nicht erhalten bleibt.

Definition 43. Die linearen Abbildungen 1 bis 5 und die Abbildungen 6 bis 9 heißen zusammen **affine Abbildungen**.

Bemerkung 18. Die Abbildungen 7 bis 9 kann man als Komposition von linearen Abbildungen und der Translation erhalten.

Ziel: Auch diese nichtlinearen Abbildungen sollen durch Matrizen dargestellt werden.

Beispiel 71.

1. Translation

$$T_{(e,f)} : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad T_{(e,f)} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}$$

2. Lineare Abbildung

$$L : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad L \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$$

3. Lineare Abbildung **und** Translation

$$A : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by + e \\ cx + dy + f \end{pmatrix}$$

Definition 44. Eine **affine Abbildung** ist von der Form

$$A : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by + e \\ cx + dy + f \end{pmatrix}$$

Ist $e = f = 0$, so handelt es sich um eine lineare Abbildung.

Matrzenschreibweise:

$$A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b & e \\ c & d & f \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} ax + by + e \\ cx + dy + f \end{pmatrix}$$

Dem "Bildvektor" fehlt die 3. Komponente!

\Rightarrow wir erweitern die Matrix A um eine 3. Zeile:

$$\begin{pmatrix} a & b & e \\ c & d & f \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} ax + by + e \\ cx + dy + f \\ 1 \end{pmatrix}$$

Definition 45. Die **homogenen Koordinaten** des Vektors $v = \begin{pmatrix} x \\ y \end{pmatrix}$ sind

$$v = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Die affine Abbildung

$$A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by + e \\ cx + dy + f \end{pmatrix}$$

wird dann durch die Matrix

$$M = \begin{pmatrix} a & b & e \\ c & d & f \\ 0 & 0 & 1 \end{pmatrix}$$

in homogenen Koordinaten dargestellt.

Beispiel 72. Translation

$$M_{T(e,f)} = \begin{pmatrix} 1 & 0 & e \\ 0 & 1 & f \\ 0 & 0 & 1 \end{pmatrix}$$

$$T_{(e,f)} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & e \\ 0 & 1 & f \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} x + e \\ y + f \\ 1 \end{pmatrix}$$

Bemerkung 19. Alle Transformationen im \mathbb{R}^2 können auf diese Weise in homogenen Koordinaten eingebettet werden.

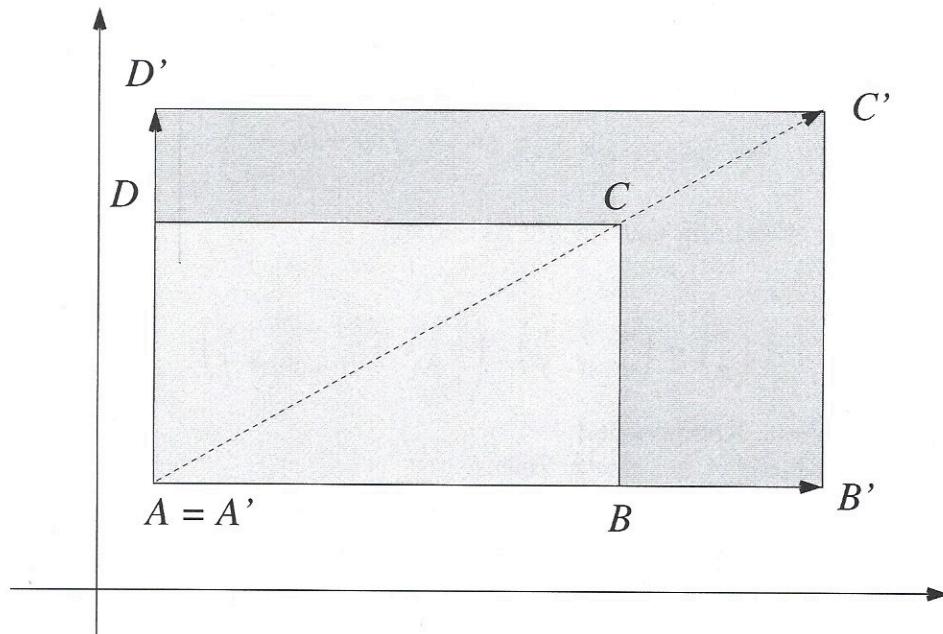
$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Matrix einer
linearen Abb.
in gew. Koord.

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

Matrix einer
linearen Abb.
in homog. Koord.

Beispiel 73. Ein achsenparalleles Rechteck $ABCD$ mit $A(a, b)$ soll relativ zu seinem linken unteren Eckpunkt A um den Faktor λ gezoomt werden.



1. Schritt: $T_{(-a, -b)}$

$$M_{T_{(-a, -b)}} = \begin{pmatrix} 1 & 0 & -a \\ 0 & 1 & -b \\ 0 & 0 & 1 \end{pmatrix}$$

2. Schritt: Z_λ

$$M_{Z_\lambda} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

3. Schritt: $T_{(a, b)}$

$$M_{T_{(a, b)}} = \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}$$

Komposition:

$$M_{\text{gesamt}} = \begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -a \\ 0 & 1 & -b \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \lambda & 0 & a(1-\lambda) \\ 0 & \lambda & b(1-\lambda) \\ 0 & 0 & 1 \end{pmatrix}$$

Kapitel 5

Unitäre Räume

5.1 Längen und Abstände

Im \mathbb{R}^2 gibt der Ausdruck

$$\|v\| = \sqrt{x_1^2 + x_2^2}$$

die (physikalische) Länge des Vektors $v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ an. Man spricht auch vom Betrag des Vektors v .

Wir wollen uns hier mit einer Verallgemeinerung des Betragsbegriffs, der sogenannten **Norm** eines Vektors, befassen, die in den Anwendungen notwendig ist.

Definition 46. Sei V ein \mathbb{K} -Vektorraum. Jedem Vektor $v \in V$ sei eine reelle Zahl $\|v\|$ zugeordnet, wobei folgende Axiome gelten sollen:

$$(N1) \quad \|v\| \geq 0 \quad \forall v \in V \quad \text{und} \quad (\|v\| = 0 \Leftrightarrow v = 0)$$

$$(N2) \quad \|\lambda v\| = |\lambda| \cdot \|v\| \quad \forall v \in V, \quad \forall \lambda \in \mathbb{K}$$

$$(N3) \quad \|v + w\| \leq \|v\| + \|w\| \quad \forall v, w \in V \quad (\text{Dreiecksungleichung})$$

Durch diese Abbildung

$$\begin{aligned} \|\cdot\| : & \quad V \rightarrow \mathbb{R} \\ & v \mapsto \|v\| \end{aligned}$$

ist eine **Norm** auf V definiert.

Die Zahl $\|v\|$ heißt **Norm** oder **Betrag** oder **Länge** des Vektors v .

Ein Vektor $v \in V$ heißt **Einheitsvektor** oder **normiert**, falls

$$\|v\| = 1$$

Bemerkung 20. Einen Vektor $v \in V, v \neq 0$, normiert man dadurch, dass man ihn mit dem Kehrwert seines Betrages multipliziert, d.h. der Vektor

$$w = \frac{1}{\|v\|} v$$

besitzt die Norm 1 ($\|w\| = 1$) und er weist in dieselbe Richtung wie v .

Definition 47. Für $v, w \in V$ definiert

$$d(v, w) := \|v - w\|$$

den **Abstand von v und w** .

Beispiel 74. $V = \mathbb{K}^n : v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n$

$$\|v\|_1 := |x_1| + \dots + |x_n| \quad \dots \text{ Betragssnorm (= Manhattan-Norm)}$$

$$\|v\|_2 := \sqrt{|x_1|^2 + \dots + |x_n|^2} \quad \dots \text{ Euklidische Norm}$$

$$\|v\|_\infty := \max\{|x_1|, \dots, |x_n|\} \quad \dots \text{ Maximumsnorm}$$

Beispiel 75. $V = \mathbb{K}^3 : v = \begin{pmatrix} -5 \\ 3 \\ -2 \end{pmatrix}, w = \begin{pmatrix} 2+i \\ -1-i \\ 2 \end{pmatrix}$

$$\|v\|_1 = |-5| + |3| + |-2| = 10$$

$$\|v\|_2 = \sqrt{25 + 9 + 4} = \sqrt{38}$$

$$\|v\|_\infty = \max\{5, 3, 2\} = 5$$

$$\|w\|_1 = |2+i| + |-1-i| + |2| = \sqrt{5} + \sqrt{2} + 2$$

$$\|w\|_2 = \sqrt{5+2+4} = \sqrt{11}$$

$$\|w\|_\infty = \max\{\sqrt{5}, \sqrt{2}, 2\} = \sqrt{5}$$

5.2 Das Skalarprodukt

Definition 48. Sei V ein \mathbb{K} -Vektorraum. Eine Abbildung

$$\begin{aligned} s : & V \times V \rightarrow \mathbb{K} \\ & (v, w) \mapsto s(v, w) \end{aligned}$$

heißt ein **Skalarprodukt** oder **inneres Produkt** in V , wenn für alle $v, w, u \in V$ und alle $\lambda \in \mathbb{K}$ gilt:

$$(S1) \quad s(v, v) > 0 \quad \text{falls} \quad v \neq 0$$

$$(S2) \quad s(v, w) = \overline{s(w, v)}$$

$$(S3) \quad s(\lambda v, w) = \lambda s(v, w)$$

$$(S4) \quad s(u + v, w) = s(u, w) + s(v, w)$$

Ein Vektorraum mit einem Skalarprodukt heißt ein **unitärer Raum**.

Schreibweise:

$$\langle v, w \rangle \quad \text{statt} \quad s(v, w)$$

Folgerungen:

$$(i) \quad \mathbb{K} = \mathbb{R} : s(v, w) = s(w, v)$$

$$(ii) \quad \mathbb{K} = \mathbb{C} : s(v, \lambda w) = \overline{s(\lambda w, v)} = \overline{\lambda} \overline{s(w, v)} = \overline{\lambda} s(v, w)$$

$$(iii) \quad (v = 0 \text{ oder } w = 0) \Rightarrow s(v, w) = 0$$

$$(iv) \quad v = 0 \Leftrightarrow s(v, v) = 0$$

Beispiel 76. $V = \mathbb{R}^2 : v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, w = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$

$$\langle v, w \rangle := x_1 y_1 + x_2 y_2$$

aber auch

$$\langle v, w \rangle := x_1 y_1 - x_1 y_2 - x_2 y_1 + 2 x_2 y_2$$

sind Skalarprodukte.

Beispiel 77. Wir betrachten den \mathbb{K}^n mit den beiden n -Tupeln $v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $w = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$.

Im Fall $V = \mathbb{R}^n$ definieren wir das Skalarprodukt in der Form

$$\langle v, w \rangle := \sum_{i=1}^n x_i y_i.$$

im Vektorraum $V = \mathbb{C}^n$ in der Form

$$\langle v, w \rangle := \sum_{i=1}^n \overline{x_i} y_i.$$

Man nennt diese Skalarprodukte auch die **kanonischen Skalarprodukte** im \mathbb{R}^n bzw. \mathbb{C}^n .

Beispiel 78. $V = \mathbb{P}_2$

$$f = a_0 + a_1 t + a_2 t^2, \quad g = b_0 + b_1 t + b_2 t^2$$

$\langle f, g \rangle = a_0 b_0 + a_1 b_1 + a_2 b_2$ ist ein Skalarprodukt

Speziell: $f = t, \quad g = 1 - t + t^2$

$$\langle f, g \rangle = -1, \quad \langle f, f \rangle = 1, \quad \langle g, g \rangle = 3$$

Beispiel 79. $V = C[a, b], f, g \in C[a, b]$

$$\langle f, g \rangle := \int_a^b f(t) \cdot g(t) dt$$

ist ein Skalarprodukt.

Im Speziellen stellt für $V = \mathbb{P}_2$ der Ausdruck $\langle f, g \rangle = \int_0^1 f(t) \cdot g(t) dt$ ein Skalarprodukt dar.

Für $f = t, \quad g = 1 - t + t^2$ gilt

$$\langle f, g \rangle = \int_0^1 t \cdot (1 - t + t^2) dt = \frac{5}{12}, \quad \langle f, f \rangle = \int_0^1 t^2 dt = \frac{1}{3}, \quad \langle g, g \rangle = \int_0^1 (1 - t + t^2) dt = \frac{7}{10}$$

5.3 Die SCHWARZ'sche Ungleichung

Satz 23. Sei V ein unitärer Vektorraum. Dann gilt für alle $v, w \in V$ die **SCHWARZ'sche Ungleichung**

$$|\langle v, w \rangle|^2 = \langle v, w \rangle \cdot \overline{\langle v, w \rangle} \leq \langle v, v \rangle \langle w, w \rangle$$

Beweis. (i) $w = 0$: linke Seite: $\langle v, 0 \rangle = 0$ w.A.
rechte Seite: $\langle 0, 0 \rangle = 0$

(ii) $w \neq 0$, d.h. $\langle w, w \rangle > 0$:

$$\begin{aligned} 0 &\leq \langle v - \lambda w, v - \lambda w \rangle = \\ &= \langle v, v \rangle - \lambda \langle w, v \rangle - \bar{\lambda} \langle v, w \rangle + \lambda \bar{\lambda} \langle w, w \rangle \\ &= \langle v, v \rangle - \lambda \overline{\langle v, w \rangle} - \bar{\lambda} \langle v, w \rangle + \lambda \bar{\lambda} \langle w, w \rangle \end{aligned}$$

Wählt man speziell

$$\lambda = \frac{\langle v, w \rangle}{\langle w, w \rangle}, \quad \text{also} \quad \bar{\lambda} = \frac{\overline{\langle v, w \rangle}}{\langle w, w \rangle}$$

so folgt

$$\begin{aligned} 0 &\leq \langle v, v \rangle - \frac{\langle v, w \rangle}{\langle w, w \rangle} \overline{\langle v, w \rangle} - \frac{\overline{\langle v, w \rangle}}{\langle w, w \rangle} \langle v, w \rangle + \frac{\langle v, w \rangle}{\langle w, w \rangle} \frac{\overline{\langle v, w \rangle}}{\langle w, w \rangle} \langle w, w \rangle \\ &\Rightarrow 0 \leq \langle v, v \rangle \cdot \langle w, w \rangle - \langle v, w \rangle \overline{\langle v, w \rangle} \end{aligned}$$

d.h.

$$|\langle v, w \rangle|^2 \leq \langle v, v \rangle \langle w, w \rangle$$

Mit Hilfe eines Skalarprodukts kann man nun eine Norm definieren. Es gilt der folgende

Satz 24. Sei V ein unitärer Raum. Dann stellt die Abbildung

$$\begin{aligned}\|\cdot\| : \quad V &\rightarrow \mathbb{R} \\ v &\mapsto \|v\| := \sqrt{\langle v, v \rangle}\end{aligned}$$

eine **Norm** auf V dar. Man nennt sie **induzierte Norm**.

Beweis. Es ist zu zeigen, dass die Normaxiome von Definition 46 erfüllt sind.

(N1) $\|v\| \geq 0$ und $\|v\| = 0 \Leftrightarrow v = 0$ folgen unmittelbar aus Folgerung (iv) von Seite 64.

(N2) $\|\lambda v\|^2 = \langle \lambda v, \lambda v \rangle = \lambda \bar{\lambda} \langle v, v \rangle = |\lambda|^2 \langle v, v \rangle \geq 0$
 $\Rightarrow \|\lambda v\| = \sqrt{|\lambda|^2 \langle v, v \rangle} = |\lambda| \|v\|$

(N3) Die Schwarz'sche Ungleichung ist äquivalent zur Ungleichung $|\langle v, w \rangle| \leq \|v\| \cdot \|w\|$. Damit erhalten wir:

$$\begin{aligned}0 \leq \|v + w\|^2 &= \langle v + w, v + w \rangle = \\&= \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle \\&= \langle v, v \rangle + \langle v, w \rangle + \overline{\langle v, w \rangle} + \langle w, w \rangle \\&= \langle v, v \rangle + 2 \operatorname{Re} \langle v, w \rangle + \langle w, w \rangle \\&\leq \langle v, v \rangle + 2|\langle v, w \rangle| + \langle w, w \rangle \\&\leq \|v\|^2 + 2\|v\| \cdot \|w\| + \|w\|^2 \quad (\text{Schwarz'sche Ungl.}) \\&= (\|v\| + \|w\|)^2 \quad \text{q.e.d.}\end{aligned}$$

5.4 Winkelmessung, Orthonormierung nach GRAM-SCHMIDT

Im reellen unitären Raum V gilt wegen der Schwarz'schen Ungleichung für alle $v, w \in V \setminus \{0\}$:

$$0 \leq \frac{|\langle v, w \rangle|^2}{\langle v, v \rangle \langle w, w \rangle} \leq 1 \iff -1 \leq \frac{\langle v, w \rangle}{\|v\| \|w\|} \leq 1$$

Damit ist folgende Definition sinnvoll:

Definition 49. Im reellen unitären Raum \mathbb{R}^n betrachten wir zwei Vektoren $v, w \in \mathbb{R}^n \setminus \{0\}$. Als **Winkel** zwischen v und w , $\angle(v, w)$, erklären wir die reelle Zahl

$$\varphi \equiv \angle(v, w) := \arccos \frac{\langle v, w \rangle}{\|v\| \cdot \|w\|}, \quad 0 \leq \varphi \leq \pi.$$

Für $v = 0$ oder $w = 0$ wird $\angle(v, w)$ nicht definiert.

Eigenschaften:

- (i) $\angle(v, w) = \angle(w, v)$
- (ii) $\langle v, w \rangle = 0 \iff \angle(v, w) = \pi/2$
- (iii) $\angle(v, -v) = \arccos \frac{\langle v, -v \rangle}{\|v\| \cdot \|v\|} = \arccos \frac{-\langle v, v \rangle}{\|v\| \cdot \|v\|} = \arccos(-1) = \pi$

Beispiel 80. $V = \mathbb{R}^2$: Für $v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, w = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ erklären wir

$$\langle v, w \rangle = x_1 y_1 + x_2 y_2$$

Speziell für $v = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$, $w = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$ erhalten wir:

$$\langle v, w \rangle = 10$$

$$\langle v, v \rangle = 10, \quad \|v\| = \sqrt{10}$$

$$\langle w, w \rangle = 20, \quad \|w\| = \sqrt{20}$$

und damit

$$\angle(v, w) = \arccos \frac{10}{\sqrt{10}\sqrt{20}} = \arccos \frac{\sqrt{2}}{2} = \frac{\pi}{4}$$

Beispiel 81. $V = \mathbb{R}^2$, $v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ $w = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ mit dem Skalarprodukt

$$\langle v, w \rangle = x_1y_1 - x_1y_2 - x_2y_1 + 2x_2y_2$$

Mit den speziellen Vektoren von Beispiel 80 erhalten wir jetzt

$$\langle v, w \rangle = 6 - 12 - 2 + 8 = 0$$

$$\langle v, v \rangle = 9 - 6 + 2 = 5$$

$$\langle w, w \rangle = 4 + 32 - 16 = 20$$

Damit folgt

$$\angle(v, w) = \arccos 0 = \frac{\pi}{2}$$

D.h. die Wahl des Skalarprodukts bestimmt die Größe des Winkels!

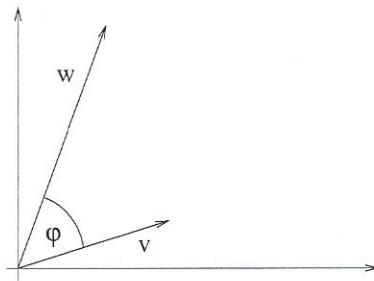


Abbildung 5.1: Winkel zwischen 2 Vektoren

In komplexen unitären Räumen wollen wir keinen Winkel sondern lediglich den Begriff der **Orthogonalität** einführen.

Definition 50. Im unitären Raum V heißen zwei Vektoren $v, w \in V$ **orthogonal** (bzgl. des gewählten Skalarprodukts), wenn

$$\langle v, w \rangle = 0.$$

Eine nichtleere Teilmenge $A \subset V$ heißt **orthonormiert**, wenn alle Vektoren von A normiert und paarweise orthogonal sind.

Beispiel 82. $V = \mathbb{P}_2$. Man zeige, dass $f = 1$, $g = -\frac{1}{2} + t$, bzgl. des Skalarprodukts

$$\langle f, g \rangle = \int_0^1 f(t) \cdot g(t) dt$$

orthogonal sind. Man bilde daraus eine orthonormale Menge.

$$\begin{aligned}\langle f, g \rangle &= \int_0^1 1 \cdot \left(-\frac{1}{2} + t\right) dt = \left(-\frac{t}{2} + \frac{t^2}{2}\right) \Big|_0^1 = 0 \Rightarrow f \perp g. \\ \langle f, f \rangle &= \int_0^1 1 dt = 1, \quad \langle g, g \rangle = \int_0^1 \left(-\frac{1}{2} + t\right)^2 dt = \frac{1}{12} \Rightarrow \|f\| = 1, \|g\| = \frac{1}{\sqrt{12}} \\ \Rightarrow \tilde{f} &= \frac{1}{\|f\|} f = 1, \quad \tilde{g} = \frac{1}{\|g\|} g = \sqrt{12} \left(-\frac{1}{2} + t\right) \\ \Rightarrow \tilde{f}, \tilde{g} &\text{ sind orthonormal.}\end{aligned}$$

Definition 51. Wir definieren das **Kronecker'sche Delta-Symbol** δ_{ik} wie folgt:

$$\delta_{ik} := \begin{cases} 1 & \text{für } i = k \\ 0 & \text{für } i \neq k \end{cases}$$

Bemerkung 21.

1. Die Einheitsmatrix I kann damit auch geschrieben werden in der Form

$$I = (\delta_{ik})$$

2. Für eine orthonormierte Teilmenge $\mathcal{A} = \{v_1, \dots, v_k\}$ eines unitären Raumes gilt also

$$\langle v_i, v_j \rangle = \delta_{ij} \quad \forall i, j \in \{1, 2, \dots, k\}$$

Satz 25. Eine endliche orthonormierte Teilmenge $\mathcal{A} = \{v_1, \dots, v_n\}$ eines unitären Raumes V ist stets linear unabhängig. Ist zudem V endlich dimensional und gilt $\dim V = n$, so ist \mathcal{A} eine Basis von V .

Beweis. Es gilt

$$\langle v_i, v_k \rangle = \delta_{ik} \quad \forall i, k \in \{1, 2, \dots, n\}$$

Aus $\lambda_1 v_1 + \dots + \lambda_n v_n = 0$ folgt:

$$\begin{aligned}0 &= \langle 0, v_k \rangle = \langle \lambda_1 v_1 + \dots + \lambda_n v_n, v_k \rangle \\ &= \lambda_1 \langle v_1, v_k \rangle + \dots + \lambda_k \langle v_k, v_k \rangle + \dots + \lambda_n \langle v_n, v_k \rangle = \lambda_k \\ \Rightarrow \lambda_k &= 0, \quad k = 1, \dots, n \quad \Rightarrow \{v_1, \dots, v_n\} \text{ sind l.u.}\end{aligned}$$

Satz 26. Orthonomierungsverfahren von GRAM-SCHMIDT

Sei V ein unitärer Raum mit der induzierten Norm $\|\cdot\|$ und $\{v_1, \dots, v_n\}$ eine linear unabhängige Teilmenge von V . Dann ist die Teilmenge $\{w_1, \dots, w_n\}$ mit

$$\begin{aligned}w_1 &= \frac{1}{\|v_1\|} \cdot v_1 \\ w_i &= \frac{1}{\|u_i\|} \cdot u_i, \quad i = 2, \dots, n \\ \text{wobei} \quad u_i &= v_i - \sum_{k=1}^{i-1} \langle v_i, w_k \rangle w_k, \quad i = 2, \dots, n\end{aligned}$$

orthonomiert.

Beispiel 83. $V = \mathbb{R}^3$

$$v_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$\begin{aligned}
w_1 &= \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \\
u_2 &= v_2 - \langle v_2, w_1 \rangle w_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2} \left\langle \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right\rangle \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix} \\
w_2 &= \frac{1}{\sqrt{6}} \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} \\
u_3 &= v_3 - \langle v_3, w_1 \rangle w_1 - \langle v_3, w_2 \rangle w_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2} \cdot 1 \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{6} \cdot 1 \cdot \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ \frac{3}{3} \\ \frac{-2}{3} \end{pmatrix} \\
w_3 &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}
\end{aligned}$$

$\{w_1, w_2, w_3\}$ ist eine Orthonormalbasis des \mathbb{R}^3 .

 **Beispiel 84.** $V = \mathbb{P}_2$ mit dem Skalarprodukt

$$\langle f, g \rangle = \int_{-1}^1 f(t) \cdot g(t) dt$$

$$f_1(t) = 1, \quad f_2(t) = t, \quad f_3(t) = t^2$$

$$\begin{aligned}
\|f_1\|^2 &= \langle f_1, f_1 \rangle = \int_{-1}^1 1^2 dt = 2 \\
&\Rightarrow w_1(t) = \frac{1}{\|f_1\|} \cdot f_1 = \frac{1}{\sqrt{2}}
\end{aligned}$$

$$u_2(t) = f_2(t) - \langle f_2, w_1 \rangle w_1(t) = t - \frac{1}{2} \int_{-1}^1 t dt = \left(t - \frac{1}{2} \frac{t^2}{2} \right) \Big|_{-1}^1 = t$$

$$\begin{aligned}
\|u_2\|^2 &= \int_{-1}^1 t^2 dt = \frac{2}{3} \\
&\Rightarrow w_2(t) = \frac{1}{\|u_2\|} \cdot u_2 = \sqrt{\frac{3}{2}} t
\end{aligned}$$

 Analog bildet man

$$w_3(t) = \sqrt{\frac{5}{2}} \left(\frac{3}{2} t^2 - \frac{1}{2} \right)$$

Bemerkung 22. Man nennt die Polynome w_1, w_2, w_3 die **Legendre'schen Polynome**, sie spielen in vielen Bereichen der Technik eine Rolle.

5.5 QR-Zerlegung

Jede Matrix $A \in M(m \times n)$ mit $m \geq n$ und linear unabhängigen Spaltenvektoren kann in ein Produkt $Q \cdot R$ zerlegt werden, wobei Q eine Matrix mit orthonormalen Spaltenvektoren ist und R eine reelle obere Dreiecksmatrix. Diese Zerlegung findet Anwendungen bei der Bestimmung von Näherungslösungen von Gleichungssystemen (siehe Abschnitt 7.9.3) und in der Approximation von Eigenwerten (siehe Abschnitt 9.5).

Seien a_1, \dots, a_n die linear unabhängigen Spaltenvektoren von A und q_1, \dots, q_n die orthonormalen Vektoren, die man mit Hilfe des GRAM-SCHMIDT-Verfahrens aus a_1, \dots, a_n gewinnt.

Dann gilt:

$$W_i = L(a_1, \dots, a_i) = L(q_1, \dots, q_i)$$

für $i = 1, 2, \dots, n$.

Es existieren daher skalare Faktoren $r_{1i}, r_{2i}, \dots, r_{ii}$ so, dass

$$a_i = r_{1i}q_1 + r_{2i}q_2 + \dots + r_{ii}q_i, \quad i = 1, \dots, n,$$

d.h.

$$\begin{aligned} a_1 &= r_{11}q_1 \\ a_2 &= r_{12}q_1 + r_{22}q_2 \\ &\vdots \\ a_n &= r_{1n}q_1 + r_{2n}q_2 + \dots + r_{nn}q_n \end{aligned}$$

Diese Beziehungen können jetzt in Matrzenschreibweise angegeben werden:

$$A = (a_1, \dots, a_n) = (q_1, \dots, q_n) \cdot \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_{nn} \end{pmatrix} = QR$$

Bemerkung 23.

1. Die Matrix Q besitzt orthonormale Spaltenvektoren, d.h. es gilt

$$\langle q_i, q_j \rangle = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j \end{cases}$$

oder in Matrzenschreibweise

$$Q^T Q = I \tag{5.1}$$

2. Alle Diagonalelemente $r_{ii}, i = 1, \dots, n$, von R sind ungleich Null.
Daher ist R regulär.

Satz 27. QR-Zerlegung

Sei $A \in M(m \times n)$ mit linear unabhängigen Spaltenvektoren. Dann kann A zerlegt werden in der Form $A = QR$ wobei $Q \in M(m \times n)$ eine Matrix mit orthonormalen Spaltenvektoren ist und $R \in M(n \times n)$ eine invertierbare obere Dreiecksmatrix darstellt.

Beispiel 85. Man ermittle die QR-Zerlegung der Matrix

$$A = \begin{pmatrix} 1 & 2 & 2 \\ -1 & 1 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

Das Orthonormierungsverfahren von GRAM-SCHMIDT auf die Spaltenvektoren von A angewandt liefert die Vektoren

$$q_1 = \begin{pmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{pmatrix}, q_2 = \begin{pmatrix} 3\sqrt{5}/10 \\ 3\sqrt{5}/10 \\ \sqrt{5}/10 \\ \sqrt{5}/10 \end{pmatrix}, q_3 = \begin{pmatrix} -\sqrt{6}/6 \\ 0 \\ \sqrt{6}/6 \\ \sqrt{6}/3 \end{pmatrix}$$

Damit erhält man

$$Q = \begin{pmatrix} 1/2 & 3\sqrt{5}/10 & -\sqrt{6}/6 \\ -1/2 & 3\sqrt{5}/10 & 0 \\ -1/2 & \sqrt{5}/10 & \sqrt{6}/6 \\ 1/2 & \sqrt{5}/10 & \sqrt{6}/3 \end{pmatrix}$$

Um nun R zu bestimmen beachtet man die Relation (5.1). Damit folgt aus $A = QR$

$$Q^T A = Q^T QR = IR = R$$

In unserem Fall erhalten wir also

$$R = Q^T A = \begin{pmatrix} 1/2 & -1/2 & -1/2 & 1/2 \\ 3\sqrt{5}/10 & 3\sqrt{5}/10 & \sqrt{5}/10 & \sqrt{5}/10 \\ -\sqrt{6}/6 & 0 & \sqrt{6}/6 & \sqrt{6}/3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 2 \\ -1 & 1 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 1/2 \\ 0 & \sqrt{5} & 3\sqrt{5}/2 \\ 0 & 0 & \sqrt{6}/2 \end{pmatrix}$$

Kapitel 6

Eigenwerte und Eigenvektoren, Diagonalisierung

6.1 Eigenwerte und Eigenvektoren

Sei $F : V \rightarrow V$ eine lineare Abbildung des Vektorraumes V auf sich selbst. In vielen Anwendungen ist es notwendig, einen Vektor $v \in V$ zu finden, dessen Bild $F(v)$ parallel zu v ist. D.h. man sucht nach einem Vektor v und einem Skalar λ mit der Eigenschaft

$$F(v) = \lambda \cdot v$$

Wenn $v \neq 0$ und λ diese Bedingung erfüllen, werden λ **Eigenwert** (EW) und v **Eigenvektor** (EV) genannt.

Ist V endlich dimensional, so kann F durch eine $n \times n$ -Matrix A angegeben werden. Deshalb behandeln wir hier Eigenwerte und Eigenvektoren von $n \times n$ -Matrizen.

Definition 52. Sei $A \in M(n \times n)$ mit reellen Komponenten. Die Zahl $\lambda \in \mathbb{K}$ heißt **Eigenwert** von A , wenn es einen Vektor $v \in \mathbb{C}^n$ gibt, so daß

$$A \cdot v = \lambda \cdot v$$

Der Vektor $v \neq 0$ wird **Eigenvektor** von A zum **Eigenwert** λ genannt.

Beispiel 86. $A = \begin{pmatrix} 10 & -18 \\ 6 & -11 \end{pmatrix}$

$$(i) A \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 10 & -18 \\ 6 & -11 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \Rightarrow \lambda_1 = 1 \text{ ist Eigenwert, } v_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \text{ ist Eigenvektor}$$

$$(ii) A \cdot \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 10 & -18 \\ 6 & -11 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} -6 \\ -4 \end{pmatrix} = (-2) \begin{pmatrix} 3 \\ 2 \end{pmatrix} \Rightarrow \lambda_2 = -2 \text{ ist Eigenwert, } v_2 = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \text{ ist Eigenvektor.}$$

Frage: Wie bestimmt man die Eigenwerte und die Eigenvektoren?

Die Relationen

$$Av = \lambda v \Leftrightarrow Av - \lambda Iv = 0 \Leftrightarrow (A - \lambda I)v = 0$$

stellen ein lineares homogenes Gleichungssystem für die Komponenten von $v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{C}^n$ dar. Es besitzt eine nichttriviale Lösung, falls

$$\det(A - \lambda I) = 0$$

Ist $\det(A - \lambda I) = 0$ so folgt: $\exists v \neq 0 : A \cdot v = \lambda v$, d. h. v ist Eigenvektor zum Eigenwert λ .

Satz 28. Sei $A \in M(n \times n)$. Ein $\lambda \in \mathbb{K}$ ist Eigenwert von A genau dann, wenn

$$P(\lambda) = \det(A - \lambda I) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0 \quad (6.1)$$

Definition 53. Die Gleichung (6.1) nennt man **charakteristische Gleichung**, das Polynom $P(\lambda)$ heißt **charakteristisches Polynom**.

Bemerkung 24. Die charakteristische Gleichung

$$P(\lambda) = \det(A - \lambda I) = 0$$

kann auch geschrieben werden in der Form

$$P(\lambda) = \lambda^n + b_{n-1}\lambda^{n-1} + \dots + b_1\lambda + b_0 = 0, \quad b_i \in \mathbb{R}.$$

Nach dem Fundamentalsatz der Algebra besitzt jedes Polynom n -ten Grades mit reellen oder komplexen Koeffizienten genau n (reelle oder komplexe Nullstellen), wobei eine Nullstelle mit der Vielfachheit r entsprechend oft gezählt wird. Z.B. hat das Polynom $(\lambda - 1)^5$ 5 Nullstellen, alle sind 1.

⇒ Jede Matrix $A \in M(n \times n)$ hat genau n Eigenwerte.

Definition 54. Sei λ ein Eigenwert von A . Die Menge $E_\lambda = \{v \in \mathbb{C}^n \mid Av = \lambda v, v \neq 0\}$ wird **Eigenraum** von A zum Eigenwert λ genannt.

Bemerkung 25. Ist λ Eigenwert von $A \in M(n \times n)$ und sei $E_\lambda = \{v \mid Av = \lambda v, v \neq 0\}$. Dann ist E_λ ein linearer Teilraum von \mathbb{C}^n .

Beweis. $v \in E_\lambda \Leftrightarrow Av = \lambda v \Leftrightarrow (A - \lambda I)v = 0 \Leftrightarrow v \in \text{Kern}(A - \lambda I)$.

Da der Kern einer linearen Abbildung stets ein linearer Teilraum ist (siehe Abschnitt 4.2), ist E_λ auch einer.

Beispiel 87. $A = \begin{pmatrix} 4 & 2 \\ 3 & 3 \end{pmatrix}$

$$\begin{aligned} P(\lambda) &= \det(A - \lambda I) = \begin{vmatrix} 4 - \lambda & 2 \\ 3 & 3 - \lambda \end{vmatrix} = \lambda^2 - 7\lambda + 6 = 0 \\ \lambda_{1/2} &= \frac{7}{2} \pm \sqrt{\frac{49}{4} - 6} = \frac{7}{2} \pm \frac{5}{2} \Rightarrow \lambda_1 = 6, \quad \lambda_2 = 1 \end{aligned}$$

Eigenvektoren: zu $\lambda_1 = 6$: $(A - \lambda_1 I)v = 0$

$$\begin{pmatrix} -2 & 2 \\ 3 & -3 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \xi = \eta \Rightarrow v_1 = \xi \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \xi \in \mathbb{R}, \quad \text{oder} \quad \tilde{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

zu $\lambda_2 = 1$: $(A - \lambda_2 I)v = 0$

$$\begin{pmatrix} 3 & 2 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow 3\xi + 2\eta = 0 \quad \text{z.B. } \xi = 2, \quad \eta = -3$$

$$v_2 = \begin{pmatrix} 2 \\ -3 \end{pmatrix} \quad \text{oder} \quad \tilde{v}_2 = t \begin{pmatrix} 2 \\ -3 \end{pmatrix}, t \in \mathbb{R}.$$

Die Vektoren v_1 und v_2 sind linear unabhängig!

$$\Rightarrow E_6 = L\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right), \quad E_1 = L\left(\begin{pmatrix} 2 \\ -3 \end{pmatrix}\right)$$

Beispiel 88. $A = \begin{pmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \\ 2 & 1 & -1 \end{pmatrix}$

$$P(\lambda) = \det(A - \lambda I) = \begin{vmatrix} 1-\lambda & -1 & 4 \\ 3 & 2-\lambda & -1 \\ 2 & 1 & -1-\lambda \end{vmatrix} = -(\lambda^3 - 2\lambda^2 - 5\lambda + 6) = 0$$

$$\Rightarrow \lambda_1 = 1 \quad \text{geraten}, \quad (\lambda^3 - 2\lambda^2 - 5\lambda + 6) : (\lambda - 1) = \lambda^2 - \lambda - 6 \quad (\text{Polynomdivision})$$

$$\lambda_{2/3} = \frac{1}{2} \pm \sqrt{\frac{1}{4} + 6} = \frac{1}{2} \pm \frac{5}{2} \quad \Rightarrow \quad \lambda_2 = 3, \quad \lambda_3 = -2$$

Eigenvektoren: zu $\lambda_1 = 1$:

$$\begin{pmatrix} 0 & -1 & 4 \\ 3 & 1 & -1 \\ 2 & 1 & -2 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Das Gauß'sche Eliminationsverfahren liefert:

$$\begin{pmatrix} 0 & -1 & 4 \\ 3 & 1 & -1 \\ 2 & 1 & -2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & 1 & -2 \\ 0 & -\frac{1}{2} & 2 \\ 0 & 0 & 0 \end{pmatrix}, \quad \text{wähle: } \zeta = t \quad \Rightarrow \quad \eta = 4t, \quad \xi = -t$$

$$v_1 = t \begin{pmatrix} -1 \\ 4 \\ 1 \end{pmatrix}, \quad t \in \mathbb{R}, \quad \Rightarrow \quad E_1 = L \left\{ \begin{pmatrix} -1 \\ 4 \\ 1 \end{pmatrix} \right\}$$

zu $\lambda_2 = 3$:

$$\begin{pmatrix} -2 & -1 & 4 \\ 3 & -1 & -1 \\ 2 & 1 & -4 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} -2 & -1 & 4 \\ 3 & -1 & -1 \\ 2 & 1 & -4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} -2 & -1 & 4 \\ 0 & -\frac{5}{2} & 5 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\text{wähle: } \zeta = t \quad \Rightarrow \quad \eta = 2t, \quad \xi = t \quad \Rightarrow \quad v_2 = t \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad t \in \mathbb{R}, \quad \Rightarrow \quad E_2 = L \left\{ \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \right\}$$

zu $\lambda_3 = -2$:

$$\begin{pmatrix} 3 & -1 & 4 \\ 3 & 4 & -1 \\ 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} 3 & -1 & 4 \\ 3 & 4 & -1 \\ 2 & 1 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 3 & -1 & 4 \\ 0 & 5 & -5 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\text{wähle: } \zeta = t \quad \Rightarrow \quad \eta = t, \quad \xi = -t \quad \Rightarrow \quad v_3 = t \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \quad t \in \mathbb{R}, \quad E_{-2} = L \left\{ \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

Die Eigenvektoren v_1, v_2, v_3 sind linear unabhängig!

Beispiel 89. $A = \begin{pmatrix} 3 & -5 \\ 1 & -1 \end{pmatrix}$

$$P(\lambda) = \begin{vmatrix} 3-\lambda & -5 \\ 1 & -1-\lambda \end{vmatrix} = \lambda^2 - 2\lambda + 2 = 0 \quad \Rightarrow \quad \lambda_{1/2} = 1 \pm i \in \mathbb{C}$$

Eigenvektoren: zu $\lambda_1 = 1 + i$:

$$\begin{pmatrix} 2-i & -5 \\ 1 & -2-i \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}; \quad \text{die 2. Gleichung liefert: } \xi = (2+i)\eta$$

$$v_1 = t \begin{pmatrix} 2+i \\ 1 \end{pmatrix}, \quad t \in \mathbb{R}, \quad E_{1+i} = L \left\{ \begin{pmatrix} 2+i \\ 1 \end{pmatrix} \right\}$$

zu $\lambda_2 = 1 - i$:

$$\begin{pmatrix} 2+i & -5 \\ 1 & -2+i \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}; \quad \text{aus 2. Gleichung folgt: } \xi = (2-i)\eta$$

$$v_1 = t \begin{pmatrix} 2-i \\ 1 \end{pmatrix}, \quad t \in \mathbb{R}, \quad E_{1-i} = L \left\{ \begin{pmatrix} 2-i \\ 1 \end{pmatrix} \right\}$$

Bemerkung: $\lambda_2 = \overline{\lambda_1}$, $v_2 = \overline{v_1}$.

Allgemein gilt:

Die komplexen Eigenwerte einer reellen Matrix treten stets paarweise als konjugiert komplexe Zahlen auf, die zugehörigen Eigenvektoren sind dann auch zueinander konjugiert komplex.

Satz 29. Die Eigenwerte $\lambda_1, \dots, \lambda_m$ der Matrix $A \in M(n \times n)$ seien verschieden. Dann sind die zugehörigen Eigenvektoren v_1, \dots, v_m linear unabhängig.

Beispiel 90. $A = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}$

$$P(\lambda) = \begin{vmatrix} 4-\lambda & 0 \\ 0 & 4-\lambda \end{vmatrix} = (4-\lambda)^2 = 0 \Rightarrow \lambda_{1/2} = 4 \text{ doppelter Eigenwert!}$$

Eigenvektoren: zu $\lambda = 4$:

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \xi \text{ und } \eta \text{ können beliebig gewählt werden:}$$

$$v = \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \xi \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \eta \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ ist Eigenvektor}$$

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ sind linear unabhängige Eigenvektoren} \Rightarrow E_4 = L \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$

Beispiel 91. $A = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}$ Eigenwerte: $\lambda_{1/2} = 4$ ist doppelter Eigenwert!

Eigenvektoren: zu $\lambda = 4$:

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \eta = 0, \quad \xi \text{ kann beliebig gewählt werden.}$$

$$v_1 = t \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad t \in \mathbb{R} \Rightarrow E_4 = L \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$$

Definition 55.

- (i) Tritt eine Nullstelle x_0 eines Polynoms P r -fach auf, so sagt man, x_0 habe die **algebraische Vielfachheit** $r = \mu(P, x_0)$
- (ii) Sei λ ein Eigenwert der Matrix A . Dann ist die **geometrische Vielfachheit** von λ die Dimension des zu λ gehörenden Eigenraumes E_λ , d.h.

geometrische Vielfachheit := $\dim E_\lambda$

Satz 30. Sei λ ein Eigenwert von A . Dann ist stets

$$\dim E_\lambda \leq \mu(P, \lambda)$$

d.h. die geometrische Vielfachheit ist stets kleiner oder gleich der algebraischen Vielfachheit.

Frage: Wann besitzt $A \in M(n \times n)$ n linear unabhängige Eigenvektoren?

Satz 31. Eine Matrix $A \in M(n \times n)$ besitzt genau dann n linear unabhängige Eigenvektoren, wenn für jeden Eigenwert λ gilt

$$\dim E_\lambda = \mu(P, \lambda)$$

Insbesonders besitzt A n linear unabhängige Eigenvektoren, wenn alle n Eigenwerte von A voneinander verschieden sind.

6.2 Ähnliche Matrizen und Diagonalisierung

Definition 56. Zwei Matrizen $A, B \in M(n \times n)$ heißen **ähnlich**, wenn es eine reguläre Matrix $C \in M(n \times n)$ gibt, so daß

$$B = C^{-1}AC$$

Ziel: Wir werden zeigen, daß

- (i) ähnliche Matrizen einige gemeinsame Eigenschaften besitzen,
- (ii) die meisten Matrizen ähnlich sind einer Diagonalmatrix.

Satz 32. Seien $A, B \in M(n \times n)$ ähnliche Matrizen. Dann besitzen A und B dasselbe charakteristische Polynom und daher auch dieselben Eigenwerte.

Beweis. $B = C^{-1}AC$

$$\begin{aligned} \det(B - \lambda I) &= \det(C^{-1}AC - \lambda I) = \det[C^{-1}AC - C^{-1}(\lambda I)C] = \det[C^{-1}(A - \lambda I)C] \\ &= \det C^{-1} \cdot \det(A - \lambda I) \cdot \det C = \frac{1}{\det C} \cdot \det(A - \lambda I) \cdot \det C = \det(A - \lambda I) \end{aligned}$$

Definition 57. Eine Matrix $A \in M(n \times n)$ heißt **diagonalisierbar**, falls es eine Diagonalmatrix D gibt, so daß A ähnlich zu D ist.

Bemerkung 26. Die Eigenwerte einer Diagonalmatrix sind ihre Diagonalelemente, d.h. eine diagonalisierbare Matrix A ist dann einer Diagonalmatrix ähnlich, in deren Diagonale die Eigenwerte von A stehen.

Satz 33. $A \in M(n \times n)$ ist genau dann diagonalisierbar, wenn A n linear unabhängige Eigenvektoren besitzt. In diesem Fall ist D gegeben durch

$$D = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}, \quad \lambda_1, \dots, \lambda_n \dots \text{ Eigenwerte von } A.$$

Bildet man $C \in M(n \times n)$ dadurch, daß man die linear unabhängigen Eigenvektoren von A als Spalten von C einträgt, so gilt

$$D = C^{-1}AC$$

Beweis.

(i) $v_1 = \begin{pmatrix} c_{11} \\ \vdots \\ c_{n1} \end{pmatrix}, v_2 = \begin{pmatrix} c_{12} \\ \vdots \\ c_{n2} \end{pmatrix}, \dots, v_n = \begin{pmatrix} c_{1n} \\ \vdots \\ c_{nn} \end{pmatrix}$ seien die n linear unabh. Eigenvektoren von A .

Wir bilden die Matrix

$$C = \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \dots & c_{nn} \end{pmatrix}$$

C ist regulär, da die Spalten linear unabhängig sind. Jetzt folgt:

$$A \cdot C = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \dots & c_{nn} \end{pmatrix};$$

die i -te Spalte von $A \cdot C$ ist daher

$$A \cdot \begin{pmatrix} c_{1i} \\ \vdots \\ c_{ni} \end{pmatrix} = A \cdot v_i = \lambda_i v_i \implies A \cdot C = \begin{pmatrix} \lambda_1 c_{11} & \dots & \lambda_n c_{1n} \\ \vdots & \ddots & \vdots \\ \lambda_1 c_{n1} & \dots & \lambda_n c_{nn} \end{pmatrix}$$

Wir betrachten nun

$$C \cdot D = \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \dots & c_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix} = \begin{pmatrix} \lambda_1 c_{11} & \dots & \lambda_n c_{1n} \\ \vdots & \ddots & \vdots \\ \lambda_1 c_{n1} & \dots & \lambda_n c_{nn} \end{pmatrix}$$

Daraus folgt $A \cdot C = C \cdot D$ und schließlich $D = C^{-1}AC$, d.h. A ist diagonalisierbar.

(ii) Sei A diagonalisierbar, d.h. $\exists C \in M(n \times n)$ regulär, so daß $D = C^{-1}AC$. Seien v_1, \dots, v_n die Spaltenvektoren wie oben. Es folgt zunächst $A \cdot C = C \cdot D$. Die i -te Spalte von $A \cdot C$ liefert $A v_i$, die i -te Spalte von $C \cdot D$ liefert $\lambda_i v_i$. Daraus folgt

$$Av_i = \lambda_i v_i$$

d.h. v_i ist Eigenvektor von A zum Eigenwert λ_i .

Korollar: Besitzt $A \in M(n \times n)$ n verschiedene Eigenwerte, so ist A diagonalisierbar.

Beispiel 92. $A = \begin{pmatrix} 4 & 2 \\ 3 & 3 \end{pmatrix}$ mit den Eigenvektoren $v_1 = \begin{pmatrix} 2 \\ -3 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ (siehe Bsp. 87.)

$$C = \begin{pmatrix} 2 & 1 \\ -3 & 1 \end{pmatrix}, C^{-1} = \frac{1}{5} \begin{pmatrix} 1 & -1 \\ 3 & 2 \end{pmatrix} \Rightarrow C^{-1}AC = \frac{1}{5} \begin{pmatrix} 1 & -1 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 4 & 2 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 6 \end{pmatrix}$$

$\lambda_1 = 1$ und $\lambda_2 = 6$ sind Eigenwerte von A .

Beispiel 93. $A = \begin{pmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \\ 2 & 1 & -1 \end{pmatrix}$

$$v_1 = \begin{pmatrix} -1 \\ 4 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} \text{ sind Eigenvektoren (siehe Bsp. 88.)}$$

$$C = \begin{pmatrix} -1 & 1 & 1 \\ 4 & 2 & -1 \\ 1 & 1 & -1 \end{pmatrix} \Rightarrow C^{-1}AC = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -2 \end{pmatrix}; \quad \lambda_1 = 1, \lambda_2 = 3, \lambda_3 = -2 \text{ sind EW von } A.$$

6.3 Symmetrische Matrizen und orthogonale Diagonalisierung

Satz 34. Sei $A \in M(n \times n)$ reell und symmetrisch, d.h. es gilt $A^t = A$.

Dann sind alle Eigenwerte von A reell.

Beweis. Sei λ Eigenwert von A mit dem zugehörigen Eigenvektor $v \in \mathbb{C}^n$, d.h. es gilt $Av = \lambda v$. Im \mathbb{C}^n betrachten wir das kanonische Skalarprodukt (siehe Abschnitt 5.2, Bsp. 77)

$$\langle v, w \rangle := \sum_{i=1}^n x_i \bar{y}_i \quad \text{für } v = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, w = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

Mit der Interpretation $v, w \in M(n \times 1)$ kann dieses auch angegeben werden in der Form

$$\langle v, w \rangle := \sum_{i=1}^n x_i \bar{y}_i = v^t \cdot \bar{w} = (x_1, \dots, x_n) \cdot \begin{pmatrix} \bar{y}_1 \\ \vdots \\ \bar{y}_n \end{pmatrix}$$

Damit erhalten wir

$$\begin{aligned} \langle Av, v \rangle &= \langle \lambda v, v \rangle = \lambda \langle v, v \rangle \\ \langle Av, v \rangle &= (Av)^t \cdot \bar{v} = (v^t A^t) \bar{v} = v^t (A^t \bar{v}) = v^t \cdot (\bar{A} v) \quad \text{da } A \text{ reell, symmetrisch} \\ &= \langle v, Av \rangle = \langle v, \lambda v \rangle = \bar{\lambda} \langle v, v \rangle \end{aligned}$$

Da $\langle v, v \rangle \neq 0$, folgt $\lambda = \bar{\lambda}$, d.h. mit $\lambda = a + ib$, $\bar{\lambda} = a - ib$ folgt

$$a + ib = a - ib \Rightarrow b = 0 \Rightarrow \lambda \text{ ist reell.}$$

Satz 35. Sind λ_1 und λ_2 verschiedene Eigenwerte einer reellen symmetrischen Matrix mit den entsprechenden reellen Eigenvektoren v_1 und v_2 . Dann sind v_1 und v_2 orthogonal.

Beweis. Wir betrachten hier das kanonische Skalarprodukt im \mathbb{R}^n :

$$\begin{aligned} \langle Av_1, v_2 \rangle &= \lambda_1 \langle v_1, v_2 \rangle \\ \langle Av_1, v_2 \rangle &= \langle v_1, A^t v_2 \rangle = \langle v_1, Av_2 \rangle = \langle v_1, \lambda_2 v_2 \rangle = \lambda_2 \langle v_1, v_2 \rangle \\ &\Rightarrow \lambda_1 \langle v_1, v_2 \rangle = \lambda_2 \langle v_1, v_2 \rangle \\ &\text{wegen } \lambda_1 \neq \lambda_2 \text{ folgt } \langle v_1, v_2 \rangle = 0. \end{aligned}$$

Satz 36. Eine reelle symmetrische Matrix $A \in M(n \times n)$ besitzt stets n orthonormierte Eigenvektoren.

Anmerkung: Die Matrix Q , die die orthonormierten Eigenvektoren von A als Spalten enthält, hat folgende Eigenschaften:

$$(i) Q^{-1} A Q = D, \quad D \dots \text{Diagonalmatrix}$$

$$(ii) Q^{-1} = Q^t$$

Definition 58. Eine reguläre Matrix Q mit der Eigenschaft

$$Q^{-1} = Q^t$$

heißt **orthogonal**.

Satz 37. Eine Matrix $A \in M(n \times n)$ ist genau dann orthogonal, wenn ihre Spaltenvektoren eine Orthonormalbasis des \mathbb{R}^n bilden.

Beweis. 1. Aus der Definition 58 sehen wir:

$$A \text{ orthogonal} \iff A^{-1} = A^t \iff A^t \cdot A = I = (\delta_{jk}) \quad (6.2)$$

2. Bezeichnet a_j den j -ten Spaltenvektor von A , so erhalten wir

$$\begin{aligned} A^t \cdot A &= \begin{pmatrix} a_1^t \\ a_2^t \\ \vdots \\ a_n^t \end{pmatrix} (a_1 \ a_2 \ \dots \ a_n) = \\ &= \begin{pmatrix} a_1^t \cdot a_1 & a_1^t \cdot a_2 & \dots & a_1^t \cdot a_n \\ a_2^t \cdot a_1 & a_2^t \cdot a_2 & \dots & a_2^t \cdot a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n^t \cdot a_1 & a_n^t \cdot a_2 & \dots & a_n^t \cdot a_n \end{pmatrix} = \begin{pmatrix} \langle a_1, a_1 \rangle & \langle a_1, a_2 \rangle & \dots & \langle a_1, a_n \rangle \\ \langle a_2, a_1 \rangle & \langle a_2, a_2 \rangle & \dots & \langle a_2, a_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle a_n, a_1 \rangle & \langle a_n, a_2 \rangle & \dots & \langle a_n, a_n \rangle \end{pmatrix} \quad (6.3) \end{aligned}$$

3. Aus (6.2) folgt nun mit (6.3) und Bemerkung 21

$$\begin{aligned} A \text{ ist orthogonal} &\iff \langle a_j, a_k \rangle = \delta_{jk} \quad \forall j, k \in \{1, 2, \dots, n\} \\ &\iff \dim \mathbb{R}^n = n \quad (a_1, a_2, \dots, a_n) \text{ ist eine Orthonormalbasis des } \mathbb{R}^n \end{aligned}$$

Definition 59. Eine Matrix $A \in M(n \times n)$ heißt **orthogonal diagonalisierbar**, wenn es eine orthogonale Matrix $Q \in M(n \times n)$ gibt, so dass

$$Q^t A Q = D$$

wobei D eine Diagonalmatrix ist, in deren Diagonale die Eigenwerte von A stehen.

Satz 38. Eine reelle Matrix $A \in M(n \times n)$ ist genau dann orthogonal diagonalisierbar, wenn sie symmetrisch ist.

Programm zur Bestimmung der diagonalisierenden Matrix Q :

1. Ermittle eine Basis des Eigenraumes von A .
2. Berechne eine Orthonormalbasis jedes Eigenraumes von A mittels Gram-Schmidt-Verfahren (siehe Satz 26, Abschnitt 5.4).
3. Trage in die Spalten von Q die orthonormierten Eigenvektoren aus Schritt 2 ein.

Beispiel 94. $A = \begin{pmatrix} 19 & -3 \\ -3 & 11 \end{pmatrix}$ ist symmetrisch

(i) Eigenwerte:

$$P(\lambda) = \begin{vmatrix} 19 - \lambda & -3 \\ -3 & 11 - \lambda \end{vmatrix} = \lambda^2 - 30\lambda + 200 = 0 \Rightarrow \lambda_1 = 20, \lambda_2 = 10$$

(ii) Eigenvektoren: zu $\lambda_1 = 20$:

$$\begin{pmatrix} -1 & -3 \\ -3 & -9 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow v_1 = \begin{pmatrix} -3 \\ 1 \end{pmatrix} \Rightarrow w_1 = \frac{1}{\|v_1\|} \cdot v_1 = \frac{1}{\sqrt{10}} \begin{pmatrix} -3 \\ 1 \end{pmatrix}$$

zu $\lambda_2 = 10$:

$$\begin{pmatrix} 9 & -3 \\ -3 & 1 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow v_2 = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \Rightarrow w_2 = \frac{1}{\|v_2\|} v_2 = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

$$(iii) Q = \begin{pmatrix} \frac{-3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & \frac{3}{\sqrt{10}} \end{pmatrix} = \frac{1}{\sqrt{10}} \begin{pmatrix} -3 & 1 \\ 1 & 3 \end{pmatrix}$$

$$Q^t A Q = \frac{1}{\sqrt{10}} \begin{pmatrix} -3 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 19 & -3 \\ -3 & 11 \end{pmatrix} \frac{1}{\sqrt{10}} \begin{pmatrix} -3 & 1 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 20 & 0 \\ 0 & 10 \end{pmatrix}$$

Beispiel 95. $A = \begin{pmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix}$ ist symmetrisch

$$(i) \text{ Eigenwerte: } P(\lambda) = \begin{vmatrix} 5-\lambda & 4 & 2 \\ 4 & 5-\lambda & 2 \\ 2 & 2 & 2-\lambda \end{vmatrix} = (\lambda-1)^2(\lambda-10) = 0 \Rightarrow \lambda_{1/2} = 1, \quad \lambda_3 = 10$$

(ii) Eigenvektoren: zu $\lambda_{1/2} = 1$

$$\begin{pmatrix} 4 & 4 & 2 \\ 4 & 4 & 2 \\ 2 & 2 & 1 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \text{Koeff.-Matrix} \rightsquigarrow \begin{pmatrix} 4 & 4 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\alpha) \text{ Wähle: } \zeta = 0 : \xi = -\eta \Rightarrow v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

$$\beta) \text{ Wähle: } \eta = 0 : \zeta = -2\xi \Rightarrow v_2 = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}$$

zu $\lambda_3 = 10$:

$$\begin{pmatrix} -5 & 4 & 2 \\ 4 & -5 & 2 \\ 2 & 2 & -8 \end{pmatrix} \begin{pmatrix} \xi \\ \eta \\ \zeta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \text{Koeff.-Matrix} \rightsquigarrow \begin{pmatrix} 1 & 1 & -4 \\ 0 & -9 & 18 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\eta = 2\zeta, \quad \xi = 2\zeta \Rightarrow v_3 = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$$

(iii) Bestimmung einer Orthonormalbasis der Eigenräume nach Gram-Schmidt (siehe Abschnitt 5.4):

$$E_1 = L \left\{ \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} \right\}$$

$$w_1 = \frac{1}{\|v_1\|} \cdot v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

$$u_2 = v_2 - \langle v_2, w_1 \rangle w_1 = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} - \frac{1}{2} \cdot 1 \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -2 \end{pmatrix}$$

$$w_2 = \frac{1}{\|u_2\|} \cdot u_2 = \frac{1}{3\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ -4 \end{pmatrix}$$

Probe: $\langle w_1, w_2 \rangle = 0$

$$w_3 = \frac{1}{\|v_3\|} v_3 = \frac{1}{3} \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$$

(iv)

$$Q = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \\ -\frac{1}{\sqrt{2}} & \frac{3\sqrt{2}}{3\sqrt{2}} & \frac{2}{3} \\ 0 & \frac{-4}{3\sqrt{2}} & \frac{1}{3} \end{pmatrix} \Rightarrow Q^t A Q = \dots = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 10 \end{pmatrix}$$

Bemerkung 27. Eine Matrix $A \in M(n \times n)$ kann genau dann diagonalisiert werden, wenn für jeden Eigenwert λ von A gilt

$$\dim E_\lambda = \mu(P, \lambda).$$

Im Fall $\dim E_\lambda < \mu(P, \lambda)$ kann die Matrix A auf Jordan'sche Normalform transformiert werden.

Definition 60. Eine **Jordan'sche Blockmatrix** $B \in M(k \times k)$ besitzt die Form

$$B(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & 0 & \dots & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \lambda \end{pmatrix}$$

d.h. in der Hauptdiagonalen steht λ , in der schrägen Reihe darüber die 1er. Eine **Jordan-Matrix** J besitzt die Form

$$J = \begin{pmatrix} B_1(\lambda_1) & & & & & \\ & B_2(\lambda_2) & & & & 0 \\ 0 & & \ddots & & & \\ & & & & & B_r(\lambda_r) \end{pmatrix}$$

d.h. in ihrer Diagonalen stehen Jordan-Blockmatrizen, der Rest ist mit Nullen aufgefüllt.

Beispiel 96. Die folgenden Matrizen sind Beispiele für Jordan-Matrizen:

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}, \quad \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & -3 & 1 & 0 & 0 \\ 0 & 0 & -3 & 1 & 0 \\ 0 & 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 & 7 \end{pmatrix}$$

Satz 39. Zu jeder Matrix $A \in M(n \times n)$ existiert eine reguläre Matrix $C \in M(n \times n)$, so daß

$$C^{-1} A C = J$$

d.h. A kann stets auf Jordansche Normalform transformiert werden. In der Diagonalen der einzelnen Jordan-Blöcke steht jeweils ein (mehrerer) Eigenwert von A .

Hinweis: Details findet man in der einschlägigen Literatur.

6.4 Singulärwertzerlegung

In Kapitel 6.3 wurde gezeigt, dass jede symmetrische Matrix A zerlegt werden kann in das Produkt

$$A = Q \cdot D \cdot Q^t$$

wobei Q eine orthogonale Matrix ist und D eine Diagonalmatrix, in deren Diagonale die Eigenwerte von A stehen. Ist A nicht symmetrisch, so gibt es nach Satz 36 keine derartige Zerlegung, eventuell kann A gemäß Satz 33 in der Form $A = C \cdot D \cdot C^{-1}$ zerlegt werden.

Wir haben also gesehen, dass nicht jede Matrix diagonalisierbar ist, im Folgenden wird aber gezeigt, dass jede Matrix $A \in M(m \times n)$ zerlegt werden kann in der Form

$$A = U\Sigma V^t$$

wobei U und V orthogonale Matrizen sind und Σ eine Diagonalmatrix als Teilmatrix enthält. Diese Zerlegung wird Singulärwertzerlegung (engl.: singular value decomposition, kurz SVD) genannt.

6.4.1 Singulärwerte einer Matrix

Für jede Matrix $A \in M(m \times n)$ ist die Matrix $A^t \cdot A$ eine symmetrische Matrix, die nach Satz 36 orthogonal diagonalisiert werden kann. Die Eigenwerte von $A^t \cdot A$ sind alle reell und nicht negativ.

Definition 61. Für $A \in M(m \times n)$ sind die Quadratwurzeln der (positiven) Eigenwerte von $A^t \cdot A$ die **Singulärwerte von A** . Sie werden mit $\sigma_1, \dots, \sigma_n$ bezeichnet. Wir vereinbaren $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$.

Beispiel 97. Man bestimme die Singulärwerte von

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

$$A^t A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

Die Eigenwerte von $A^t A$ sind: $\lambda_1 = 3, \lambda_2 = 1$

Die Singulärwerte von A sind demnach: $\sigma_1 = \sqrt{\lambda_1} = \sqrt{3}, \quad \sigma_2 = \sqrt{\lambda_2} = 1$.

6.4.2 Singulärwertzerlegung einer Matrix

Satz 40. Sei $A \in M(m \times n)$ mit den Singulärwerten $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ und $\sigma_{r+1} = \sigma_{r+2} = \dots = \sigma_n = 0$. Dann existieren eine orthogonale Matrix $U \in M(m \times m)$, eine orthogonale Matrix $V \in M(n \times n)$ und eine Matrix $\Sigma \in M(m \times n)$ von der Form

$$\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \quad \text{mit} \quad D = \begin{pmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_r \end{pmatrix} \in M(r \times r),$$

sodass A zerlegt werden kann in der Form

$$A = U\Sigma V^t. \tag{6.4}$$

Diese Zerlegung wird **Singulärwertzerlegung** von A genannt. Die Spalten von U heißen linke Singulärvektoren von A , die Spalten von V nennt man rechte Singulärvektoren von A . Die Matrizen U und V sind nicht eindeutig bestimmt.

Es gilt:

- Die orthonomierten Eigenvektoren $\{v_1, \dots, v_n\}$ von $A^t A$ bilden die Spalten von V .
- Die Spaltenvektoren von U werden wie folgt bestimmt:
Man berechnet zunächst

$$u_i = \frac{1}{\sigma_i} A v_i, \quad i = 1, \dots, r.$$

Ist $r < m$, so ergänzt man die Vektoren u_1, \dots, u_r zu einer orthonormalen Basis $\{u_1, \dots, u_m\}$ des \mathbb{R}^m . Die Vektoren $\{u_1, \dots, u_m\}$ bilden dann die Spalten von U .

- Der Rang von A ist r .

Bemerkung 28. Ist A eine symmetrische, positiv definite Matrix, so reduziert sich die Singulärwertzerlegung auf die orthogonale Diagonalisierung.

Beispiel 98.

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

$$A^t A = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Die Eigenwerte von $A^t A$ sind: $\lambda_1 = 2, \lambda_2 = 1, \lambda_3 = 0$

Die zugehörigen Eigenvektoren sind: $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$.

Diese Vektoren sind orthogonal, wir normieren sie noch und erhalten

$$v_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}.$$

Die Singulärwerte von A sind also $\sigma_1 = \sqrt{2}, \sigma_2 = 1, \sigma_3 = 0$.

Damit erhalten wir $V = \begin{pmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \end{pmatrix}$ und $\Sigma = \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$.

Um U zu bestimmen, berechnen wir

$$u_1 = \frac{1}{\sigma_1} Av_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

$$u_2 = \frac{1}{\sigma_2} Av_1 = \frac{1}{1} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

$\{u_1, u_2\}$ sind schon orthonormal, wir erhalten also $U = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Damit ergibt sich die Singulärwertzerlegung von A zu

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{pmatrix} = U \Sigma V^t.$$

Beispiel 99.

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (\text{siehe Beispiel 97}).$$

Die Singulärwerte von A sind: $\sigma_1 = \sqrt{3}, \sigma_2 = 1$.

Die orthonomisierte Eigenvektoren von $A^t A$ lauten: $v_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}, v_2 = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$.

Damit erhalten wir $V = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}, \Sigma = \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$.

Um U zu bestimmen, berechnen wir

$$u_1 = \frac{1}{\sigma_1} Av_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix}$$

$$u_2 = \frac{1}{\sigma_2} Av_2 = \frac{1}{1} \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

Nun muß $\{u_1, u_2\}$ zu einer orthonormalen Basis des \mathbb{R}^3 ergänzt werden.

Wir nehmen zunächst einen dritten Vektor, der von u_1 und u_2 linear unabhängig ist, dazu, z.B. $\{u_1, u_2, e_3\}$. Dann wird mit dem Orthonormalisierungsverfahren von Gram-Schmidt daraus eine Orthonormalbasis des \mathbb{R}^3 . Man benötigt hier nur den letzten Schritt und erhält

$$u_3 = \begin{pmatrix} -1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}.$$

$$\text{Somit ist } U = \begin{pmatrix} 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{pmatrix}$$

und wir erhalten die Singulärwertzerlegung von A in der Form

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = U\Sigma V^t.$$

Man kann die rechte Seite in der Zerlegung (6.4) noch umformen und erhält so die **äußere Produktform** der Singulärwertzerlegung:

Satz 41. Für $A \in M(m \times n)$ mit den Singulärwerten $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ und $\sigma_{r+1} = \sigma_{r+2} = \dots = \sigma_n = 0$ seien u_1, \dots, u_r die entsprechenden Spaltenvektoren von U und v_1, \dots, v_r die entsprechenden Spaltenvektoren von V . Dann gilt

$$A = \sigma_1(u_1 \cdot v_1^t) + \sigma_2(u_2 \cdot v_2^t) + \dots + \sigma_r(u_r \cdot v_r^t).$$

6.4.3 Anwendungen der Singulärwertzerlegung

a) Bestimmung des Ranges einer Matrix

Bei der Rangbestimmung mit Hilfe des Gauß'schen Algorithmus kann es für schlecht konditionierte Matrizen zu numerischen Problemen kommen (siehe Abschnitt 7). Bei den Umformungen können Einträge, die eigentlich Null sein müßten, als sehr kleine Zahlen ungleich Null auftreten. Damit kann die Rangbestimmung falsch ausfallen.

Beispiel 100.

$$A = \begin{pmatrix} 8.1650 & -0.0041 & -0.0041 \\ 4.0826 & -3.9960 & 4.0042 \\ 4.0825 & 4.0042 & -3.9960 \end{pmatrix}, \quad B = \begin{pmatrix} 8.17 & 0 & 0 \\ 4.08 & -4 & 4 \\ 4.08 & 4 & -4 \end{pmatrix}$$

Die Matrix B entsteht aus der Matrix A durch Runden auf 2 Dezimalen.

Es gilt: $\text{rang } A = 3$, $\text{rang } B = 2$,

d.h. der Rang dieser beiden annähernd gleichen Matrizen ist unterschiedlich.

Daraus folgt z.B., dass A invertierbar ist und B singulär ist. Die Erklärung dieses Unterschiedes liegt in den Singulärwerten der beiden Matrizen:

Singulärwerte von A : $\sigma_1 = 10$, $\sigma_2 = 8$, $\sigma_3 = 0.01$

Singulärwerte von B : $\sigma_1 = 10$, $\sigma_2 = 8$, $\sigma_3 = 0$

Daher verwendet man die Singulärwertzerlegung in der Praxis oft zur Rangbestimmung, da sie gegenüber Rundungsfehlern wesentlich unempfindlicher ist. Dabei benutzt man folgenden

Satz 42. Seien $\sigma_1, \dots, \sigma_r$ alle von Null verschiedenen Singulärwerte von A . Dann gilt

$$\text{rang } A = r.$$

b) Bilddatenkompression

Unter den vielen Anwendungen der Singulärwertzerlegung ist die Möglichkeit der Datenkompression bei der Übertragung digitaler Bilder (z.B. via Satelliten, Fax, Internet) wohl am beeindruckendsten. Hier will man den Aufwand der zu übermittelnden Daten reduzieren, ohne wesentliche Informationen zu verlieren. Wir betrachten z.B. ein Bild, das in 340×280 Pixel zerlegt wurde, jedem Pixel wird ein Grauwert auf einer 256-teiligen Skala zugeordnet. Dieser wird durch eine entsprechende Zahl zwischen 0 und 255 dargestellt. Man kann diese Bildinformation in Form einer 340×280 Matrix A abspeichern, ihre Einträge sind die den Grauwerten entsprechenden Zahlen. Insgesamt sind 95.200 Zahlen zu speichern bzw. zu übertragen, was sehr kostenaufwendig ist.

Bei einem Bild, z.B. einem Foto einer Person, die im Freien steht, sind viele Teile mit Himmel und Hintergrund bedeckt, während z.B. das Gesicht sehr viele Details enthält. Die Idee ist es nun, bei der Übertragung der Bilddaten jedes zweite oder jedes dritte Pixel vom Hintergrund wegzulassen, jedoch vom Gesicht alle Pixel zu nehmen. Es stellt sich nun heraus, dass die kleineren Singulärwerte der Matrix A von den "uninteressanten" Teilen des Bildes stammen. Wir könnten sie ignorieren.

Wir betrachten die äußere Produktform der Singulärwertzerlegung der Matrix A (vgl. Satz 41):

$$A = \sigma_1(u_1 \cdot v_1^t) + \sigma_2(u_2 \cdot v_2^t) + \dots + \sigma_r(u_r \cdot v_r^t).$$

Für $k \leq r$ definieren wir

$$A_k = \sigma_1(u_1 \cdot v_1^t) + \sigma_2(u_2 \cdot v_2^t) + \dots + \sigma_k(u_k \cdot v_k^t),$$

d.h. A_k ist eine Approximation von A , bei der nur die ersten k Singulärwerte und die ersten k Singulärvektoren berücksichtigt werden. Nun enthält diese Approximation z.T. schon recht viele Informationen über die Matrix A , d.h. man kann näherungsweise statt A die Matrix A_k , $k < r$, verwenden.

In unserem Beispiel mit den 340×280 Pixeln kann man schon mit der Übertragung der ersten 20 Singulärwerte (d.h. mit $k = 20$) ein akzeptables Bild bekommen. Dann müssen statt 95.000 Zahlen nur die 20 Singulärwerte, die 20 Vektoren $u_1, \dots, u_{20} \in \mathbb{R}^{340}$ und die 20 Vektoren $v_1, \dots, v_{20} \in \mathbb{R}^{280}$ übertragen werden. Das sind also insgesamt

$$20 + 20 \cdot 340 + 20 \cdot 280 = 12.420$$

Zahlen, was eine erhebliche Einsparung bedeutet.

Bemerkung: Weiter Anwendungen der Singulärwertzerlegung zur Bestimmung von Näherungslösungen von überbestimmten Gleichungssystemen werden im Abschnitt 7.9.4 behandelt.

Teil II

Numerische Methoden

Kapitel 7

Numerische Behandlung von linearen Gleichungssystemen

7.1 Einleitung

Bei vielen Aufgaben im Bereich der Linearen Algebra übernehmen numerische Methoden unter Verwendung eines Computers die entscheidende Arbeit. Dazu müssen geeignete numerische **Algorithmen** entwickelt, analysiert und implementiert werden. Dabei versteht man unter einem numerischen Algorithmus eine Folge von Rechenoperationen, durch deren Abarbeitung man aus **Eingangsdaten** die **Ergebnisse** erhält. Für ein lineares Gleichungssystem von der Form $Ax = b$ stellen die Elemente der Matrix A und die des Vektors b die Eingangsdaten dar, der Vektor x das Ergebnis.

Die Auswahl eines geeigneten Algorithmus, mit dem die numerische Lösung einer speziellen Aufgabe unter Verwendung der *moderen Rechentechnik* realisiert werden kann, sollte unter Berücksichtigung von *ökonomischen Gesichtspunkten* mit der *geforderten Genauigkeit* erfolgen.

Bemerkung 29. Zu den hervorgehobenen Begriffen:

Rechentechnik: In diesem Zusammenhang spielt die Rechengenauigkeit, die Rechengeschwindigkeit und der verfügbare Speicherplatz eine Rolle.

ökonomische Gesichtspunkte: Es muß angestrebt werden, dass ein hinreichend gutes Ergebnis mit geringem Aufwand (kurze Rechenzeit) berechnet wird.

Geforderte Genauigkeit: Es muß gesichert sein, dass mit den gegebenen Eingangsdaten das gewünschte Ergebnis mit der für die Praxis notwendigen Genauigkeit berechnet werden kann.

In den folgenden Abschnitten sollen für einige häufig auftretende Probleme einige wesentliche Gesichtspunkte erörtert werden und geeignete Verfahren vorgestellt werden.

7.2 Der GAUSS'sche Algorithmus – Die *LR*-Zerlegung

Der GAUSS'sche Algorithmus ist schon besprochen worden. Er eignet sich zwar prinzipiell gut zur Bestimmung der Lösung eines linearen inhomogenen Gleichungssystems von der Form

$$Ax = b,$$

bei der praktischen numerischen Behandlung treten aber einige Probleme auf, die durch verfeinerte Methoden vermieden bzw. gemindert werden können. Zum Beispiel ist im Fall einer regulären Matrix A die Anzahl der zur Gewinnung einer Lösung mit Hilfe der inversen Matrix A^{-1} in der Form

$$x = A^{-1} b$$

notwendigen Schritte noch relativ hoch. Man beachte, daß man in den meisten Fällen bei der Bestimmung der Lösung eines inhomogenen Gleichungssystems gar nicht alle Elemente der inversen Matrix A^{-1} kennen muß, man interessiert sich ja nur für die Lösung x . Beim GAUSS'schen Algorithmus wird die erweiterte Koeffizientenmatrix durch elementare Zeilenumformungen auf Zeilenstufenform gebracht, wobei auch der Vektor b eine Veränderung erfährt. Dies ist zum Beispiel dann sehr arbeitsintensiv, wenn man – wie es in der Praxis oft der Fall ist – eine Reihe von Gleichungssystemen mit derselben Koeffizientenmatrix A aber unterschiedlichen Vektoren b lösen möchte.

Um die Lösung von

$$Ax = b \quad \text{mit } A \in M(n \times n) \text{ invertierbar, } b \in \mathbb{R}^n \quad (7.1)$$

mit Hilfe von A^{-1} berechnen zu können, benötigt man i.a. n^3 wesentliche Rechenoperationen (= Multiplikationen bzw. Divisionen). Mit der *LR*-Zerlegung (siehe Abschnitt 1.4) kann der Rechenaufwand auf ca. $n^3/3$ Operationen gesenkt werden. Wir haben gesehen, dass eine Matrix $A \in M(m \times n)$ die durch elementare Zeilenumformungen ohne Zeilenumtauschung auf Zeilenstufenform gebracht werden kann, in ein Produkt einer linken unteren Dreiecksmatrix L und einer rechten oberen Dreiecksmatrix R zerlegt werden kann:

$$A = L \cdot R$$

Definition 62. Für $A \in M(m \times n)$ und $k = 1, \dots, \min(m, n)$, bezeichnen wir mit $A_k \in M(k \times k)$ die linke obere **Teilmatrix** von A .

Ihre Determinante $\det A_k$ heißt **Hauptminor** von A .

Satz 43. Sind alle Hauptminoren von $A \in M(m \times n)$ ungleich Null, so besitzt A eine *LR*-Zerlegung.

Beispiel 101. Für $A \in M(3 \times 4)$ gilt die folgende Zerlegung

$$A = \begin{pmatrix} 2 & 1 & 0 & 4 \\ 4 & 5 & 1 & 7 \\ 2 & -8 & 1 & 12 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & -3 & 1 \end{pmatrix}}_L \cdot \underbrace{\begin{pmatrix} 2 & 1 & 0 & 4 \\ 0 & 3 & 1 & -1 \\ 0 & 0 & 4 & 5 \end{pmatrix}}_R$$

Beim praktischen Lösen von Gleichungssystemen der Form

$$Ax = b$$

mit $A \in M(m \times n)$

$$Ax = b \implies \underbrace{L \cdot R}_{=:y} x = b$$

hat man dann zwei Gleichungssysteme

$$Ly = b \quad \text{und} \quad Rx = y \quad (7.2)$$

zu lösen. 2 Schritte: (i) Löse $Ly = b$ beginnend bei der ersten Gleichung, dann
(ii) löse $Rx = y$ beginnend bei der letzten Gleichung.

Vergleich des Rechenaufwandes für ein System (7.1) mit einer regulären Matrix $A \in M(n \times n)$:

- GAUSS'scher Algorithmus direkt: $\frac{1}{2}n^3 + \frac{1}{2}n^2$ wesentliche Operationen,
- Mit Hilfe der *LR*-Zerlegung: $\frac{1}{3}n^3 - \frac{1}{3}n$ wesentliche Operationen für die *LR*-Zerlegung und n^2 Operationen für das Lösen der Systeme (7.2).

Hinweise:

1. Für L und R muß nicht extra Speicherplatz reserviert werden. Ihre Elemente können in einer Matrix abgespeichert werden, da ja nur zwei Dreiecksmatrizen benötigt werden.

2. Dieses Lösungsverfahren ist besonders dann interessant, wenn es sich um p Systeme von Gleichungen der Form

$$A x_i = b_i, \quad i = 1, \dots, p$$

handelt, bei denen sich die Matrix A nicht ändert. Die LR -Zerlegung ist dann nur einmal notwendig!

Bei 10 Systemen mit $n = 100$ liegt der Unterschied schon bei $\approx 4.6 \cdot 10^6$ Operationen.

Bei der Überführung einer Matrix A in Zeilenstufenform kann es unter Umständen notwendig sein, einzelne Zeilen zu vertauschen. Dann gilt z.B. folgender Satz:

Satz 44. Zu jeder regulären Matrix A existiert eine Permutationsmatrix P , so daß PA in das Produkt LR zerlegt werden kann, d.h.

$$P \cdot A = L \cdot R$$

Dabei ist P die Permutationsmatrix, die die notwendigen Zeilenvertauschungen bewirkt.

D.h. die Matrix PA kann jetzt ohne Zeilenvertauschung in ein Produkt LR zerlegt werden. Für das Gleichungssystem bedeutet dies nun

$$P \cdot | \quad Ax = b \quad \Rightarrow \quad \underbrace{PA}_{LR} \cdot x = Pb \quad \Rightarrow \quad LRx = Pb$$

Man hat also die Systeme

$$Ly = Pb \quad \text{und} \quad Rx = y$$

sukzessive zu lösen.

Beispiel 102.

$$A = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 6 & 9 \\ 1 & 5 & 6 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 3 & 2 \\ 0 & 0 & 5 \\ 0 & 2 & 4 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 3 & 2 \\ 0 & 2 & 4 \\ 0 & 0 & 5 \end{pmatrix} = R$$

D.h. die Permutationsmatrix lautet

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

Hätte man im System $Ax = b$ von vorneherein die 2. und die 3. Gleichung vertauscht, so wäre der GAUSS'sche Algorithmus ohne Zeilenvertauschung möglich gewesen. Damit gewinnt man dann

$$P \cdot A = \begin{pmatrix} 1 & 3 & 2 \\ 1 & 5 & 6 \\ 2 & 6 & 9 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}}_L \cdot \underbrace{\begin{pmatrix} 1 & 3 & 2 \\ 0 & 2 & 4 \\ 0 & 0 & 5 \end{pmatrix}}_R$$

7.3 CHOLESKY-Zerlegung

Definition 63. Eine symmetrische Matrix $A \in M(n \times n)$ heißt **positiv definit** genau dann, wenn eine der folgenden Bedingungen erfüllt ist:

1. $x^T A x > 0 \quad \forall x \in \mathbb{R}^n, x \neq 0$.
2. Alle Eigenwerte λ_i von A sind größer als Null.
3. Alle Hauptminoren von A sind > 0 .
4. Alle Pivots d_i der aus A ohne Zeilenvertauschung entstandenen Matrix von Zeilenstufenform sind > 0 .

Für eine quadratische, symmetrische und positiv definite Matrix $A \in M(n \times n)$ kann man auch eine Zerlegung in der Form

$$A = L \cdot L^T$$

angeben, wobei $L \in M(n \times n)$ eine linke untere Dreiecksmatrix ist und L^T die zu L transponierte Matrix bezeichnet. Man nennt diese Zerlegung **CHOLESKY-Zerlegung**. Die Elemente l_{ik} von L können der Reihe nach spaltenweise berechnet werden, wobei man bei der ersten Spalte beginnt. Für die Berechnung der Elemente der k -ten Spalte lauten die Formeln

$$l_{kk} = \begin{cases} \sqrt{a_{11}} & \text{für } k = 1 \\ \sqrt{a_{kk} - \sum_{\mu=1}^{k-1} l_{k\mu}^2} & \text{für } k = 2, \dots, n \end{cases}$$

$$l_{ik} = \begin{cases} \frac{a_{i1}}{l_{11}} & \text{für } \begin{cases} k = 1 \\ i = 2, \dots, n \end{cases} \\ \frac{a_{ik} - \sum_{\mu=1}^{k-1} l_{i\mu} l_{k\mu}}{l_{kk}} & \text{für } \begin{cases} k = 2, 3, \dots, n \\ i = k + 1, \dots, n \end{cases} \end{cases}$$

Beispiel 103. Man führe die CHOLESKY-Zerlegung für folgende Matrix durch

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 3 & 7 & 26 \end{pmatrix}$$

$$l_{11} = \sqrt{a_{11}} = 1, \quad l_{21} = \frac{a_{21}}{l_{11}} = \frac{2}{1} = 2, \quad l_{31} = \frac{a_{31}}{l_{11}} = \frac{3}{1} = 3, \quad l_{22} = \sqrt{a_{22} - l_{21}^2} = \sqrt{5 - 4} = 1$$

$$l_{32} = \frac{1}{l_{22}}(a_{32} - l_{31}l_{21}) = 7 - 3 \cdot 2 = 1, \quad l_{33} = \sqrt{a_{33} - l_{31}^2 - l_{32}^2} = \sqrt{26 - 9 - 1} = 4$$

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 1 & 4 \end{pmatrix}$$

Bemerkung 30. Sollte eine CHOLESKY-Zerlegung möglich sein, so kann das im vorigen Abschnitt geschilderte Verfahren zur Lösung eines linearen Gleichungssystems natürlich auch entsprechend angewandt werden, wobei weitere Einsparungen den Speicherplatz und die Rechenzeit betreffend möglich sind.

7.4 Anwendungen

- Ziele:**
1. An Hand einfacher Beispiele aus dem Bereich der Differentialgleichungen soll die Herkunft großer Gleichungssystem erklrt werden.
 2. Es soll gezeigt werden, welche speziellen Eigenschaften die Koeffizientenmatrizen oft haben (z.B. Bandmatrizen).

Beispiel 104. Gesucht ist eine Funktion $u \in C^2[0, 1]$ mit der Bedingung

$$-\frac{d^2 u}{dx^2} = f(x) \quad \text{für } x \in [0, 1] \tag{7.3}$$

wobei f eine vorgelegte Funktion sei.

Bemerkung 31.

1. Dieses Problem kann natürlich sofort mit den Mitteln der Analysis gelöst werden, es soll aber hier als einfaches Beispiel für die Gewinnung eines diskreten Problems (= lineares Gleichungssystem) dienen.

2. Ist $u(x)$ eine Lösung von (3), so ist auch

$$v(x) = u(x) + C + Dx, \quad C, D \in \mathbb{R}$$

Lösung. Mit den Randwerten

$$u(0) = 0, \quad u(1) = 0 \quad (7.4)$$

wird u eindeutig festgelegt (= **Zwei-Punkt-Randwertproblem**).

3. Dieses RWP entspricht dem Wärmeleitungsproblem in einem Stab der Länge 1 mit den beiden Enden auf der Temperatur 0° und vorgegebener Wärmequelle $f(x)$.

Wir approximieren dieses Problem durch ein diskretes Problem, wodurch wir auf ein lineares Gleichungssystem geführt werden (= **Methode der finiten Differenzen**):

Wir suchen nicht nach den Funktionswerten von u in allen Punkten von $[0, 1]$ (dies würde jeden Computer überfordern), sondern nur nach den Funktionswerten an diskreten Stellen des Intervalls. Dazu wird das Intervall $[0, 1]$ in $n+1$ Teilintervalle unterteilt, d.h. mit $h = \frac{1}{n+1}$ gilt:

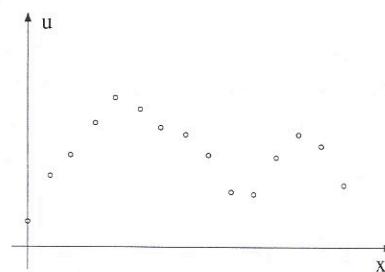
$$[0, 1] = [0, h] \cup [h, 2h] \cup \dots \cup [nh, (n+1)h]$$

Die Funktion f sei an den Zwischenstellen (= Stützstellen)

$$x_k = k h, \quad k = 0, \dots, n+1, \quad \text{mit} \quad x_0 = 0, x_{n+1} = (n+1)h = 1$$

gegeben. Wir können nun näherungsweise die wahren Werte von u an diesen Stellen berechnen und die Punkte $P_k(x_k, u(x_k))$ in eine Graphik eintragen. Verbindet man z.B. diese Punkte durch einen Polygonzug, so erhält man ein ungefähres Bild vom Verlauf des Graphen von u . Bezeichnungsweise:

$$\begin{aligned} u_0 &= u(0) = 0 \\ u_1 &= u(1 \cdot h) \\ &\vdots \\ u_n &= u(n \cdot h) \\ u_{n+1} &= u((n+1)h) = 0 \end{aligned}$$



Approximation der Ableitung $\frac{d^2 u}{dx^2}$:

Jede Ableitung ist der Grenzwert eines Differenzenquotienten:

$$\frac{du}{dx} = \lim_{h \rightarrow 0} \frac{u(x+h) - u(x)}{h}$$

Wir approximieren nun:

$$\frac{du}{dx} \approx \frac{u(x+h) - u(x)}{h}$$

oder

$$\frac{du}{dx} \approx \frac{u(x) - u(x-h)}{h} \quad \text{oder} \quad \frac{du}{dx} \approx \frac{u(x+h) - u(x-h)}{2h}$$

Wegen ihrer Symmetrie wählen wir die letzte Näherung. Für die Approximation der zweiten Ableitung sei nur eine der vielen Möglichkeiten angegeben:

$$\begin{aligned}\frac{d^2u}{dx^2} &= \frac{d}{dx} \left(\frac{du}{dx} \right) \\ &\approx \frac{u'(x+h) - u'(x)}{h} \\ &\approx \left(\frac{u(x+h) - u(x)}{h} - \frac{u(x) - u(x-h)}{h} \right) / h \\ &= \frac{u(x+h) - 2u(x) + u(x-h)}{h^2}\end{aligned}$$

An der Stelle $x_k = k h$ wird die Differentialgleichung (7.3) jetzt durch die diskrete Gleichung

$$-\frac{u(x_k+h) - 2u(x_k) + u(x_k-h)}{h^2} = f(x_k)$$

approximiert, d.h. man erhält

$$-u_{k+1} + 2u_k - u_{k-1} = h^2 f(x_k), \quad k = 1, \dots, n.$$

Das sind nun n lineare Gleichungen für die n Unbestimmten u_1, \dots, u_n . Die erste und die letzte Gleichung enthalten noch u_0 bzw. u_{n+1} , die aber aus den Randbedingungen (7.4) bekannt sind: $u_0 = u(0) = 0$, $u_{n+1} = u(1) = 0$. Speziell: Wähle $n = 5$, d.h. $h = 1/6$

$$\begin{aligned}k = 1 : \quad -u_2 + 2u_1 - \underbrace{u_0}_{=0} &= \frac{1}{6^2} f\left(\frac{1}{6}\right) \\ k = 2 : \quad -u_3 + 2u_2 - u_1 &= \frac{1}{6^2} f\left(\frac{2}{6}\right) \\ &\vdots \\ k = 5 : \quad -\underbrace{u_6}_{=0} + 2u_5 - u_4 &= \frac{1}{6^2} f\left(\frac{5}{6}\right)\end{aligned}$$

In Matrizenbeschreibweise haben wir jetzt folgendes Gleichungssystem:

$$\begin{pmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{pmatrix} = \frac{1}{6^2} \begin{pmatrix} f\left(\frac{1}{6}\right) \\ \vdots \\ f\left(\frac{5}{6}\right) \end{pmatrix} \quad (7.5)$$

Die Koeffizientenmatrix A hat (auch für den Fall $n \in \mathbb{N}$) folgende Eigenschaften:

1. $A = (a_{ij})$ ist **tridiagonal**, d.h. alle Elemente außerhalb der Hauptdiagonalen und der beiden anschließenden Nebendiagonalen sind Null (= **Bandmatrix**), d.h.

$$a_{ij} = 0 \quad \text{für } |i - j| > 1$$

2. A ist *symmetrisch*, d.h. $A^T = A$. Dies ist nicht der Fall, wenn in der Differentialgleichung Ableitungen ungerader Ordnung (u' bzw. u'' etc.) auftreten.
3. A ist *positiv definit* (siehe Def. 63).

Das GAUSS'sches Verfahren auf A angewandt liefert nun:

$$A \rightsquigarrow \begin{pmatrix} 2 & -1 & & \\ 3/2 & -1 & & \\ -1 & 2 & -1 & \\ & -1 & 2 & -1 \\ & & -1 & 2 \end{pmatrix}$$

Die neue Matrix ist wieder tridiagonal! Weitere Eliminationsschritte liefern wieder Matrizen mit derselben Eigenschaft. Schließlich erhält man die LR -Zerlegung von A in der Form

$$A = \underbrace{\begin{pmatrix} 1 & & & & \\ -\frac{1}{2} & 1 & & & \\ & -\frac{2}{3} & 1 & & \\ & & -\frac{3}{4} & 1 & \\ & & & -\frac{4}{5} & 1 \end{pmatrix}}_L \underbrace{\begin{pmatrix} 2 & -1 & & & \\ \frac{3}{2} & -1 & & & \\ \frac{4}{3} & -1 & & & \\ \frac{5}{4} & -1 & & & \\ \frac{6}{5} & -1 & & & \end{pmatrix}}_R$$

Bemerkung 32. L, R sind bidiagonal, sie liegen in derselben Bandbreite wie A .

Rechenaufwand: Will man ein lineares Gleichungssystem mit einer tridiagonalen Matrix lösen, so ist der Rechenaufwand

$$N_{\text{Band}} = \frac{2}{3}(3n - 2) \quad \text{Mult./Div.}$$

hingegen benötigt man bei einer voll besetzte Matrix

$$N_{\text{Voll}} = \frac{1}{3}(n^3 + 3n^2 - n) \quad \text{Mult./Div.}$$

Für ein typisches Problem mit $n = 150$ hat man

$$N_{\text{Voll}} \approx 1.1 \cdot 10^6 \quad \text{gegenüber} \quad N_{\text{Band}} \approx 300$$

Beispiel 105. Wir betrachten nun das spezielle Randwertproblem

$$\begin{aligned} -u'' &= 4\pi^2 \sin 2\pi x \\ u(0) &= u(1) = 0 \end{aligned}$$

a) Wähle $n = 3$, d.h. $h = \frac{1}{4}$

$$\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \frac{4\pi^2}{4^2} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

$$\Rightarrow u_1 = \frac{\pi^2}{8}, u_2 = 0, u_3 = -\frac{\pi^2}{8}$$

In Abb. 7.1 sind die drei Funktionswerte durch gerade Stecken verbunden.

b) Wähle $n = 9$, d.h. $h = \frac{1}{10}$. Die Berechnung führt auf den in Abb. 7.2 gezeichneten Polygonzug.
In Abb. 7.3 sind die Näherungslösungen der exakten Lösung

$$u(x) = \sin 2\pi x$$

gegenübergestellt.

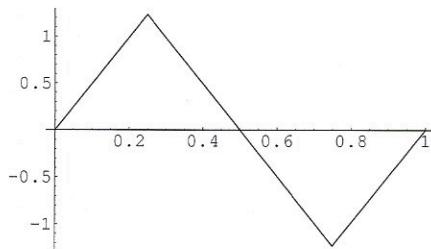
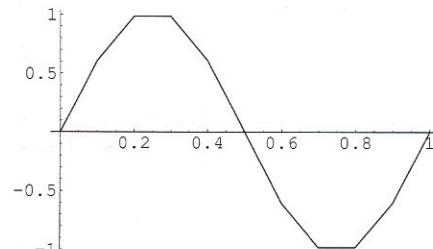
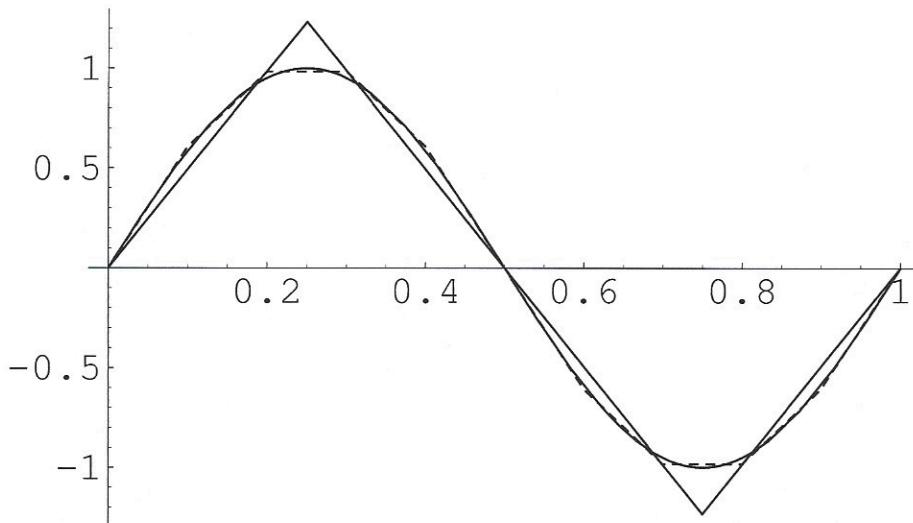
Abbildung 7.1: $n = 3$ Abbildung 7.2: $n = 9$ 

Abbildung 7.3: Vergleich mit exakter Lösung

Man kann erkennen, dass die Näherungslösung mit $n = 9$ (gestrichelte Kurve) schon recht gut an die exakte Lösung herankommt.

Beispiel 106. Temperaturverteilung in einer trapezförmigen Eisenplatte

Gesucht ist die Temperaturverteilung u in einer trapezförmigen Platte, die sich nach längerer Zeit einstellt, wenn man den Rand der Platte gemäß Abb. 7.4 erwärmt. Dieses Problem wird durch eine partielle Differentialgleichung 2. Ordnung mit gewissen Zusatzbedingungen (sogenannten Randbedingungen) beschrieben.

Die Lösung dieser Aufgabe ist überraschend schwierig. Mit Hilfe der Methode der finiten Differenzen gelingt aber relativ einfach eine näherungsweise Bestimmung der Temperaturwerte im Inneren der Platte.

Aus den Überlegungen zum Beispiel mit der Wärmeleitung in seinem Stab (siehe Bsp. 104) kann man sehen, dass dort bei fehlender Wärmequelle ($f(x) \equiv 0$) die Temperatur im Stab an der Stelle x_k gleich ist dem Mittelwert der Temperaturwerte in den beiden Nachbarpunkten (s. Gleichungssystem (7.5)).

Wir überziehen nun die vorliegende Platte mit einem diskreten Raster mit quadratischen Zellen (siehe untenstehende Abbildung). Es folgt in analoger Weise, dass die Temperatur in jedem Rasterpunkt das arithmetische Mittel der 4 angrenzenden Rasterpunkte ist (Mittelwertegenschaft). Für die Temperaturwerte u_1, \dots, u_9 in den Punkten P_1, \dots, P_9 ergibt sich damit das Gleichungssystem (7.6).

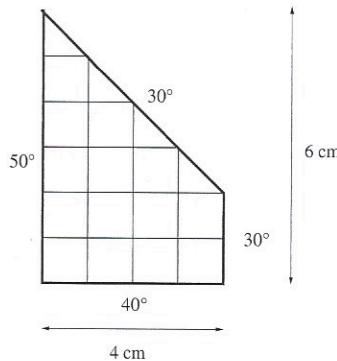
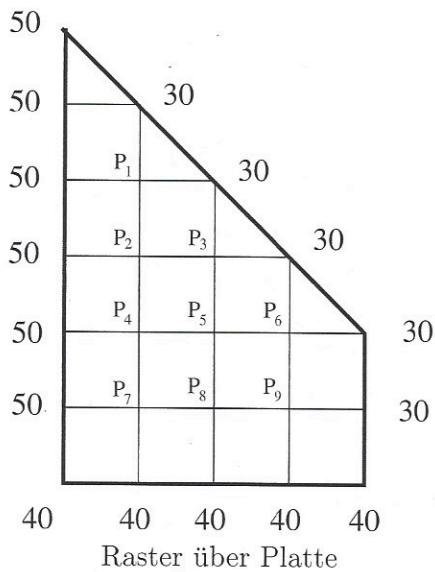


Abbildung 7.4: Trapezförmige Platte mit Randtemperatur



$$\begin{aligned}
 4u_1 &= u_2 + 50 + 30 + 30 \\
 4u_2 &= u_1 + u_3 + u_4 + 50 \\
 4u_3 &= u_2 + u_5 + 30 + 30 \\
 4u_4 &= u_2 + u_5 + u_7 + 50 \\
 4u_5 &= u_3 + u_4 + u_6 + u_8 \\
 4u_6 &= u_5 + u_9 + 30 + 30 \\
 4u_7 &= u_4 + u_8 + 40 + 50 \\
 4u_8 &= u_5 + u_7 + u_9 + 40 \\
 4u_9 &= u_6 + u_8 + 40 + 30
 \end{aligned} \tag{7.6}$$

Mit der Notation

$$u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \\ u_7 \\ u_8 \\ u_9 \end{pmatrix}, \quad b = \begin{pmatrix} 27.5 \\ 12.5 \\ 15 \\ 12.5 \\ 0 \\ 15 \\ 22.5 \\ 10 \\ 17.5 \end{pmatrix}, \quad A = \frac{1}{4} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$

gilt es jetzt, das Gleichungssystem

$$u = A \cdot u + b \tag{7.7}$$

zu lösen. Die Lösung lautet

$$\begin{aligned}
 u_1 &= 37.8^\circ, & u_4 &= 43.0^\circ, & u_7 &= 43.0^\circ \\
 u_2 &= 41.4^\circ, & u_5 &= 37.5^\circ, & u_8 &= 39.0^\circ \\
 u_3 &= 34.7^\circ, & u_6 &= 33.3^\circ, & u_9 &= 35.6^\circ
 \end{aligned} \tag{7.8}$$

Gerundet hat man also die in Abb. 7.5 angegebene Temperaturverteilung.

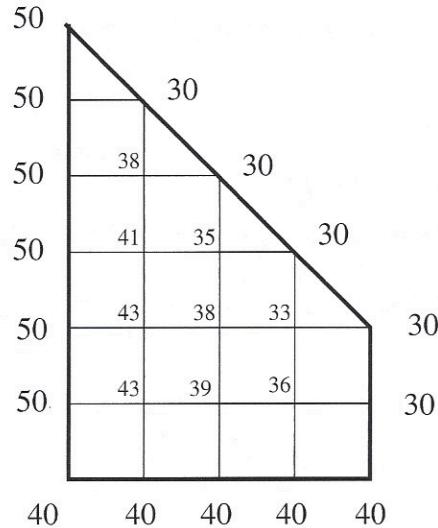


Abbildung 7.5: Temperatur am Rand und in den Knoten

Das Gleichungssystem (7.7) könnte man übrigens auch iterativ lösen, z.B. mit Hilfe des JACOBI-Verfahrens (siehe Abschnitt 7.8.1). Dabei zeigt es sich, dass unter Verwendung des Startvektors $u_0 = 0$ die Iteration nach 30 Schritten die Lösung (7.8) auf drei signifikante Stellen genau approximiert. Wählt man als Startvektor allerdings $u_0 = (40, 40, 40, 40, 40, 40, 40, 40, 40)^T$, so wird bereits nach 11 Schritten die Lösung auf drei signifikante Stellen genau approximiert.

7.5 Rundungsfehler, Pivotstrategien

In einem Rechner wird jede reelle Zahl a in Form einer Gleitkommazahl abgespeichert, z.B.

$$a = \underbrace{-}_{\text{Vorzeichen}} \underbrace{0.1234567}_{\text{Mantisse}} \cdot \underbrace{10^{-9}}_{\text{Exponent}}$$

Dabei ist die Mantisse M eine vorzeichenbehaftete t -stellige Zahl mit $0.1 \leq M < 1$. Die größte Anzahl von Dezimalstellen, die gespeichert werden kann, hängt vom Rechner bzw. vom ComputeralgebraSystem (CAS) ab. Ist die maximale Anzahl der gespeicherten Dezimalstellen t , so sagt man, es gibt t "signifikante Stellen":

Taschenrechner: $t = 10$

PC (Matlab): $t = 16$

Im Allgemeinen wird jede Zahl $a \in \mathbb{R}$ gerundet angegeben, speziell Messdaten.

Definition 64. $[a + b]$ bezeichnet die auf t Stellen gerundete Gleitkommazahl der Summe von a und b . Die gerundeten Ergebnisse von Differenz, Produkt oder Quotient werden analog bezeichnet.

Daraus ergeben sich i.a. **Rundungsfehler!**

Beispiel 107. Mit $t = 2$ berechne man

$$\left(\frac{2}{3} + \frac{2}{3} \right) - \frac{1}{3}$$

$$\left[\frac{1}{3} \right] = 0.33, \quad \left[\frac{2}{3} \right] = 0.67, \quad [0.67 + 0.67] = [1.34] = 1.3 \\ [1.3 - 0.33] = [0.97] = 0.97 !!$$

7.6 Auswahl geeigneter Pivots

Bemerkung 33. Manche Matrizen sind extrem sensibel gegenüber kleinen Änderungen ihrer Elemente, andere wiederum nicht.

Beispiel 108.

$$A = \begin{pmatrix} 1. & 1. \\ 1. & 1.0001 \end{pmatrix}, B = \begin{pmatrix} 0.0001 & 1. \\ 1. & 1. \end{pmatrix}$$

A ist sensibel, man sagt auch **schlecht konditioniert**. B ist **gut konditioniert**.

Begründung: A ist „fast“ singulär ($\det A = 0.0001$), ändert man a_{22} auf 1.0 ab, so ist $\det A = 0$. Die Matrix B ist regulär. Ein Maß dafür, wie gut eine Matrix konditioniert ist, stellt die **Konditionszahl** $\text{cond}(A)$ dar.

Definition 65. Für $A \in M(n \times n)$ seien $\lambda_1, \dots, \lambda_k$ die verschiedenen Eigenwerte von A . Die Zahl

$$\varrho(A) := \max_i |\lambda_i|$$

heißt **Spektralradius** von A .

Definition 66. Durch folgende Größen ist jeweils eine **Norm** $\|A\|$ einer Matrix $A = (a_{ij}) \in M(n \times n)$ definiert

$$\begin{aligned} \|A\|_2 &:= \sqrt{\varrho(A^t A)} \\ \|A\|_\infty &:= \max \left\{ \sum_{j=1}^n |a_{ij}| \mid i = 1, \dots, n \right\} \\ \|A\|_F &:= \sqrt{\text{Spur}(A^t A)} = \sqrt{\sum_{i,j=1}^n |a_{ij}|^2} \\ &= \sqrt{\sigma_1^2 + \dots + \sigma_r^2} \quad \text{FROBENIUS-Norm} \end{aligned}$$

Dabei bezeichnen $\sigma_1, \dots, \sigma_r$ die Singulärwerte der Matrix A . Die **Konditionszahl** $\text{cond}(A)$ einer regulären Matrix A ist definiert als

$$\text{cond}(A) := \|A^{-1}\| \cdot \|A\|$$

Bemerkung 34. Je kleiner die Konditionszahl einer Matrix ist, desto besser ist die Matrix konditioniert, desto robuster reagiert sie auf Änderungen in ihren Einträgen.

Für die hier vorliegenden Matrizen haben wir folgende Konditionszahlen.

$$\text{cond}(A)_\infty \approx 40004, \text{cond}(B)_\infty \approx 4$$

Beispiel 109. Betrachte zwei verschieden rechte Seiten für das Gleichungssystem $A \cdot x = b$:

$$b = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, b' = \begin{pmatrix} 2 \\ 2.0001 \end{pmatrix}$$

1. System:

$$\begin{aligned} u + v &= 2 \\ u + 1.0001v &= 2 \end{aligned} \implies \begin{aligned} u &= 2 \\ v &= 0 \end{aligned}$$

2. System:

$$\begin{aligned} u + v &= 2 \\ u + 1.0001v &= 2.0001 \end{aligned} \implies \begin{aligned} u &= 1 \\ v &= 1 \end{aligned}$$

Bemerkung 35. Eine Änderung in der 5. Stelle der 2. Komponente von b hat eine Vergrößerung der Änderung in der ersten Stelle von u und v zur Folge.

Bemerkung 36. Geometrisch gesehen stellen diese beiden Gleichungssysteme jeweils zwei Geraden dar, die sich "schleifend" schneiden.

Keine numerische Methode kann diese Sensibilität gegenüber kleinen Änderungen in den Daten verkleinern!

Bemerkung 37. Auch eine gut konditionierte Matrix kann durch einen schlechten Algorithmus ruiniert werden!

Beispiel 110. Wir betrachten das Gleichungssystem

$$Bx = c \quad \text{mit} \quad B \text{ von Bsp. 108} \quad \text{und} \quad c = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Wird der GÄUSS'scher Algorithmus unmittelbar angewandt, so werden wir sehen, dass sich die Rundungsfehler fatal auswirken:

$$\begin{array}{l} 0.0001u + v = 1 \\ u + v = 2 \end{array} \implies \left(\begin{array}{cc|c} 0.0001 & 1 & 1 \\ 0 & -9999 & -9998 \end{array} \right) \implies v_{\text{exakt}} = 0.99990$$

Rundung auf 3 Dezimalen liefert:

$$-10000v = -10000 \implies v_{\text{ger}} = 1$$

d.h. v_{ger} ist korrekt auf 3 Dezimalen. 1. Gleichung ungerundet:

$$0.0001u + 0.9999 = 1 \implies u_{\text{exakt}} = 1$$

1. Gleichung gerundet:

$$0.0001u + 1 = 1 \implies u_{\text{ger}} = 0 !!$$

u_{ger} ist gänzlich falsch! Obwohl B gut konditioniert ist, ist das direkte Verfahren extrem instabil!
LR-Zerlegung von B :

$$B = \underbrace{\begin{pmatrix} 1 & 0 \\ 10000 & 1 \end{pmatrix}}_L \underbrace{\begin{pmatrix} 0.0001 & 1 \\ 0 & -9999 \end{pmatrix}}_R$$

Achtung. Die Pivots in der Matrix R sind von extrem unterschiedlicher Größenordnung!
Der „kleine“ Pivot 0.0001 brachte das Malheur.

Abhilfe: Zeilenumtauschung.

Bemerkung 38. So wie ein „Null-Pivot“ eine Zeilenumtauschung erfordert, so erfordert ein sehr kleiner Pivot eine praktische Zeilenumtauschung. Ein gutes Computerprogramm sollte alle möglichen Pivots in einer Spalte miteinander vergleichen. Wählt man den größten der Kandidaten und vertauscht man dann die Zeilen entsprechend, so spricht man von einem **partiellen Pivotieren**.

Beispiel 111.

$$\begin{aligned} B &= \begin{pmatrix} 0.0001 & 1. \\ 1. & 1. \end{pmatrix}, \quad B' = \begin{pmatrix} 1. & 1. \\ 0.0001 & 1. \end{pmatrix} \\ B' &= \underbrace{\begin{pmatrix} 1. & 0 \\ 0.0001 & 1. \end{pmatrix}}_L \underbrace{\begin{pmatrix} 1 & 1 \\ 0 & 0.9999 \end{pmatrix}}_R \end{aligned}$$

Die Pivots in R sind jetzt alle von der selben Größenordnung.

Damit haben nun Rundungsfehler in den Eingangsdaten keinen so großen Einfluss mehr auf die Lösung des Gleichungssystems.

Eine mögliche Erweiterung der Strategie zur Vermeidung von ungewünscht hohen Einflüssen von Rundungsfehlern bildet eine geeignete

7.7 Skalierung der Gleichungen

Beispiel 112.

$$\begin{aligned} 2 \cdot x_1 + x_2 + x_3 &= 1 \\ x_1 + \varepsilon \cdot x_2 + \varepsilon \cdot x_3 &= 2 \cdot \varepsilon \quad \varepsilon \ll 1 \\ x_1 + \varepsilon \cdot x_2 - \varepsilon \cdot x_3 &= \varepsilon \end{aligned}$$

$$\begin{aligned} (A, b) &= \left(\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 1 & \varepsilon & \varepsilon & 2\varepsilon \\ 1 & \varepsilon & -\varepsilon & \varepsilon \end{array} \right) \mid \cdot(-1/2) \mid \cdot(-1/2) \\ &\rightsquigarrow \left(\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & \varepsilon - \frac{1}{2} & \varepsilon - \frac{1}{2} & 2\varepsilon - \frac{1}{2} \\ 0 & \varepsilon - \frac{1}{2} & -\varepsilon - \frac{1}{2} & \varepsilon - \frac{1}{2} \end{array} \right) \mid \cdot(-1) \\ &\rightsquigarrow \left(\begin{array}{ccc|c} 2 & 1 & 1 & 1 \\ 0 & \varepsilon - \frac{1}{2} & \varepsilon - \frac{1}{2} & 2\varepsilon - \frac{1}{2} \\ 0 & 0 & -2\varepsilon & -\varepsilon \end{array} \right) \quad \text{exakte Lösung: } \begin{aligned} x_1 &= \frac{-\varepsilon}{2\varepsilon-1} \\ x_2 &= \frac{6\varepsilon-1}{(2\varepsilon-1)\cdot 2} \\ x_3 &= \frac{1}{2} \end{aligned} \end{aligned}$$

Rechner: $\varepsilon = 0.01, t = 2:$

$$(A, b) \rightsquigarrow \left(\begin{array}{cccc} 2 & 1 & 1 & 1 \\ 0 & -0.5 & -0.5 & -0.5 \\ 0 & -0.5 & -0.5 & -0.5 \end{array} \right) \Rightarrow \begin{aligned} x_1 &= 0 \\ x_2 &= 1 - \sigma \\ x_3 &= \sigma, \sigma \in \mathbb{R} \end{aligned} \quad \text{keine eindeutige Lösung!}$$

Läßt man hingegen in der exakten Lösung $\varepsilon \rightarrow 0$ gehen, so folgt:

$$x_1 = 0, x_2 = 1/2, x_3 = 1/2.$$

Dies erhält man in der Gleitkommadarstellung mit $\sigma = 1/2!$

D.h. diese Lösung ist unbefriedigend.

Ausweg: Skalierung der einzelnen Gleichungen

d.h. jede Gleichung wird zunächst mit einem geeigneten Faktor multipliziert und anschließend wird eine Variablentransformation durchgeführt.

$$\begin{aligned} 2 \cdot x_1 + x_2 + x_3 &= 1 & \frac{x_1}{\varepsilon} &= z_1 \\ \frac{x_1}{\varepsilon} + x_2 + x_3 &= 2 & \text{Transf.: } x_2 &= z_2 \\ \frac{x_1}{\varepsilon} + x_2 - x_3 &= 1 & x_3 &= z_3 \end{aligned}$$

$$\left(\tilde{A}, \tilde{b} \right) = \left(\begin{array}{ccc|c} 2\varepsilon & 1 & 1 & 1 \\ 1 & 1 & 1 & 2 \\ 1 & 1 & -1 & 1 \end{array} \right) \text{ Zeilenvert. } \rightsquigarrow \left(\begin{array}{cccc} 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 2 \\ 2\varepsilon & 1 & 1 & 1 \end{array} \right) \mid \cdot(-1) \mid \cdot(-2\varepsilon)$$

$$\left(\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & [1-2\varepsilon] & [1+2\varepsilon] & [1-2\varepsilon] \end{array} \right) \text{ Zeilenvert. } \rightsquigarrow \left(\begin{array}{cccc} 1 & 1 & -1 & 1 \\ 0 & [1-2\varepsilon] & [1+2\varepsilon] & [1-2\varepsilon] \\ 0 & 0 & 2 & 1 \end{array} \right)$$

$$\begin{aligned} \text{exakte Lösung: } z_3 &= 1/2, z_2 = \frac{6\varepsilon-1}{4\varepsilon-2}, z_1 = \frac{-1}{2\varepsilon-1} \\ &\Rightarrow x_3 = 1/2, x_2 = \frac{6\varepsilon-1}{4\varepsilon-2}, x_1 = \frac{-\varepsilon}{2\varepsilon-1} \end{aligned}$$

Mit $t = 2$ und $\varepsilon = 0.01$ liefert der Rechner:

$$\left(\begin{array}{cccc} 1 & 1 & -1 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 1 & 1 & 1 \end{array} \right) \Rightarrow \begin{aligned} z_1 &= 1 & x_1 &= \varepsilon \\ z_2 &= \frac{1}{2} & x_2 &= \frac{1}{2} \\ z_3 &= \frac{1}{2} & x_3 &= \frac{1}{2} \end{aligned}$$

Günstiger Weg: Skaliere das Gleichungssystem so, daß im neuen System $\tilde{A} \cdot z = \tilde{b}$ gilt:

$$\max_{1 \leq j \leq n} \{ |\tilde{a}_{ij}| \} = 1 \quad \text{mit } i = 1, 2, \dots, n \quad (7.9)$$

Zusammenfassung: Geeignete Skalierung des Systems führt i.a. zu einer Verbesserung des Ergebnisses. Gute Programme zur numerischen Lösung von Gleichungssystemen skalieren die Gleichungen automatisch in bestimmter Weise, z.B. unter der Bedingung (7.9).

7.8 Iterative Methoden zur Lösung von linearen Gleichungssystemen

Will man mit Hilfe des GAUSS'schen Algorithmus bzw. der *LR*-Zerlegung ein System von der Form

$$Ax = b, \quad A \in M(n \times n) \quad \text{regulär}$$

lösen, so wächst die Anzahl der erforderlichen wesentlichen Rechenschritte mit der dritten Potenz von n was natürlich viel Rechenzeit kosten kann. Exakte Methoden können aber, was den Rechenaufwand betrifft, i.a. nicht mehr wesentlich verbessert werden. Gesucht sind also schnellere Verfahren zur Bestimmung einer Näherungslösung. Man startet bei einem anfänglich beliebig gewählten Vektor x_0 , berechnet dann x_1 aus x_0 und allgemein eine verbesserte Lösung x_{k+1} aus der vorangegangenen Approximation x_k . Man erhält so eine Folge von Approximationen $\{x_k\}$. Nun zur Vorgangsweise: Man spaltet die Matrix A auf in eine Summe von zwei Matrizen

$$A = S - T \tag{7.10}$$

und erhält dann die Gleichung

$$Sx = Tx + b$$

Damit kann man eine Iteration formulieren:

$$Sx_{k+1} = Tx_k + b, \quad k = 0, 1, 2, \dots \tag{7.11}$$

Es gibt aber keine Garantie dafür, daß dieser Weg zum Ziel führt. Die Aufspaltung (7.10) muß 2 Bedingungen erfüllen:

- Der neue Vektor x_{k+1} soll leicht berechenbar sein. Deshalb soll S einfach aufgebaut und invertierbar sein, S sollte z.B. diagonal oder tridiagonal sein.
- Die Folge $\{x_k\}$ von Vektoren soll gegen die wahre Lösung konvergieren:

$$\left. \begin{array}{l} Sx = Tx + b \\ Sx_{k+1} = Tx_k + b \end{array} \right\} - \quad S(x - x_{k+1}) = T(x - x_k)$$

Mit dem **Fehlervektor** $v_k := x - x_k$ gilt also

$$Sv_{k+1} = Tv_k, \quad k = 0, 1, 2, \dots$$

$$v_k = S^{-1}Tv_{k-1} = \dots = (S^{-1}T)^k v_0$$

Konvergenz für $k \rightarrow \infty$: $x_k \rightarrow x \iff v_k \rightarrow 0$

Satz 45. Die Iteration (7.11) ist genau dann konvergent, wenn für jeden Eigenwert λ von $S^{-1}T$ gilt:

$$|\lambda| < 1$$

Die Konvergenzgüte hängt von der maximalen Größe von $|\lambda|$ ab, die als **Spektralradius** ρ von $S^{-1}T$ bezeichnet wird:

$$\rho(S^{-1}T) := \max_i |\lambda_i|$$

Die Konvergenzrate r ist z.B. gegeben durch

$$r = -\log_{10} \rho$$

Unter anderem hat man zwei (extreme) Möglichkeiten, A aufzuspalten:

- $A = A + 0$, d.h. $S = A, T = 0$:

Die Iteration lautet dann

$$Ax_1 = b$$

$$\Rightarrow S^{-1}T = 0, \quad \rho(S^{-1}T) = 0, \quad r = -\log_{10} 0 = \infty$$

Nachteil: $S = A$ kann u.U. schwer invertiert werden. Das war ja das ursprüngliche Problem!

- Man wählt

$$S = \text{diag}(a_{11}, \dots, a_{nn})$$

d.h. man nimmt für die Diagonalmatrix S die Elemente aus der Diagonalen von A . Für T gilt dann $T = S - A$.

Dies führt jetzt zur sogenannten

7.8.1 JACOBI-Iteration

Aus dem Gleichungssystem $Ax = b$ gewinnt man jetzt mit der Bezeichnung

$(x_i)_k$... i -te Komponente des Vektors x_k , der im k -ten Schritt berechnet worden ist folgendes System

$$\begin{aligned} a_{11}(x_1)_{k+1} &= (-a_{12}x_2 - \dots - a_{1n}x_n)_k + b_1 \\ &\vdots \\ a_{nn}(x_n)_{k+1} &= (-a_{n1}x_1 - \dots - a_{n,n-1}x_{n-1})_k + b_n \end{aligned} \tag{7.12}$$

Ist $a_{ii} \neq 0, i = 1, \dots, n$, und ist A „dünn“ besetzt (d.h. nur „wenige“ Elemente von A sind $\neq 0$), dann funktioniert diese JACOBI-Approximation recht gut.

Fragen:

- Konvergiert das Verfahren gegen eine (bzw. gegen die) Lösung des Gleichungssystems?
- Wie steht es um die Güte der Konvergenz (=Konvergenzrate)?

Beispiel 113.

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, \quad S = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad T = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad S^{-1}T = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix}$$

$$\text{Mit } x = \begin{pmatrix} v \\ w \end{pmatrix} \quad \text{gilt jetzt} \quad \begin{aligned} 2v_{k+1} &= w_k + b_1 \\ 2w_{k+1} &= v_k + b_2 \end{aligned} \quad \text{oder}$$

$$\begin{pmatrix} v \\ w \end{pmatrix}_{k+1} = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 0 \end{pmatrix} \begin{pmatrix} v \\ w \end{pmatrix}_k + \begin{pmatrix} b_1/2 \\ b_2/2 \end{pmatrix}$$

Die Eigenwerte von $S^{-1}T$ sind $\lambda_{1/2} = \pm 1/2$, der Spektralradius ist demnach $\rho = 1/2$; mit $r = -\log_{10} \rho \approx 0.3$ liegt eine gute Konvergenz vor. Der Fehler halbiert sich bei jedem Schritt, eine weitere binäre Stelle wird korrekt.

Nachteil: Bei großen Systemen muß immer der gesamte Vektor x_k abgespeichert werden.

7.8.2 GAUSS-SEIDEL-Iteration

Neue Idee: Man verwendet bei der Iteration jede Komponente des neu berechneten Vektors x_{k+1} sobald sie ermittelt worden ist (= GAUSS-SEIDEL-Iteration):

1. Gleichung (wie vorher):

$$a_{11}(x_1)_{k+1} = (-a_{12}x_2 - \dots - a_{1n}x_n)_k + b_1$$

2. Gleichung (verwendet schon den neuen Wert von x_1):

$$a_{22}(x_2)_{k+1} = -a_{21}(x_1)_{k+1} + (-a_{23}x_3 - \dots - a_{2n}x_n)_k + b_2$$

⋮

n -te Gleichung:

$$a_{nn}(x_n)_{k+1} = (-a_{n1}x_1 - \dots - a_{n,n-1}x_{n-1})_{k+1} + b_n$$

Beispiel 114.

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}, \quad S = \begin{pmatrix} 2 & 0 \\ -1 & 2 \end{pmatrix}, \quad T = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad S^{-1}T = \begin{pmatrix} 0 & 1/2 \\ 0 & 1/4 \end{pmatrix}$$

$$2v_{k+1} = w_k + b_1$$

$$2w_{k+1} = v_{k+1} + b_2$$

$$\begin{pmatrix} 2 & 0 \\ -1 & 2 \end{pmatrix} x_{k+1} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} x_k + b$$

Die Eigenwerte von $S^{-1}T$ sind $\lambda_1 = 0, \lambda_2 = 1/4$, der Spektralradius ist demnach $\rho = 1/4$, die Konvergenzrate $r = -\log_{10} \rho \approx 0.6$.

Der Fehler wird bei jedem Schritt mit dem Faktor $1/4$ multipliziert. Ein GAUSS-SEIDEL-Schritt ist soviel wert wie zwei JACOBI-Schritte.

Satz 46. Für eine symmetrische Matrix $A \in M(n \times n)$ konvergiert die GAUSS-SEIDEL-Iteration genau dann, wenn A positiv definit ist.

Beispiel 115.

$$Ax = b \quad \text{mit} \quad A = \begin{pmatrix} 3 & -1 & 0 & 0 \\ -1 & 3 & -1 & 0 \\ 0 & -1 & 3 & -1 \\ 0 & 0 & -1 & 3 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix}, \quad \text{Startvektor } x_0 = \begin{pmatrix} 1/3 \\ 2/3 \\ 0 \\ 1/3 \end{pmatrix} \quad (7.13)$$

$$\begin{array}{rcl} x_1 & = & 0.333x_2 & +0.333 \\ x_2 & = & 0.333x_1 & +0.333x_3 & +0.667 \\ x_3 & = & 0.333x_2 & & +0.333x_4 \\ x_4 & = & & 0.333x_3 & +0.333 \end{array}$$

JACOBI-Iteration für das System (7.13):

Iterations-schritt	x_1	x_2	x_3	x_4
0	0.333	0.667	0.000	0.333
1	0.556	0.778	0.333	0.333
2	0.593	0.963	0.370	0.444
3	0.654	0.988	0.469	0.457
4	0.662	1.041	0.481	0.490
5	0.680	1.048	0.510	0.494
⋮	⋮	⋮	⋮	⋮
10	0.690	1.072	0.526	0.509
11	0.691	1.072	0.527	0.509
12	0.691	1.073	0.527	0.509
13	0.691	1.073	0.527	0.509

GAUSS–SEIDEL–Iteration für das System (7.13):

Iterations-schritt	x_1	x_2	x_3	x_4
0	0.333	0.667	0.000	0.333
1	0.556	0.852	0.395	0.465
2	0.617	1.004	0.490	0.497
3	0.668	1.053	0.516	0.505
4	0.684	1.067	0.524	0.508
5	0.689	1.071	0.526	0.509
6	0.690	1.072	0.527	0.509
7	0.691	1.073	0.527	0.509
8	0.691	1.073	0.527	0.509

Beispiel 116.

$$\begin{array}{lcl} x + 2y & = & 1 \\ 2x + y + 2z & = & 0 \\ 2y + z & = & 1 \end{array}, \quad A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 2 \\ 0 & 2 & 1 \end{pmatrix} \quad (7.14)$$

Exakte Lösung:

$$x = -\frac{1}{7}, \quad y = \frac{4}{7}, \quad z = -\frac{1}{7}$$

JACOBI– und GAUSS–SEIDEL–Iteration für das System (7.14):

Iterations-schritt	Jacobi			GAUSS–SEIDEL		
	x	y	z	x	y	z
1	1	0	1	1	0	1
2	1	-4	1	1	-4	9
3	9	-4	9	9	-36	73
4	9	-36	9	73	-292	585
5	73	-36	73	585	-2340	4681
6	73	-292	73			

Die Iteration ist stark divergent!

Definition 67. Eine Matrix $A \in M(n \times n)$ heißt **streng diagonal dominant** (s.d.d.), wenn

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \quad \forall i = 1, \dots, n.$$

Bemerkung 39. Ist die Matrix A ist streng diagonal dominant so folgt daraus, dass A invertierbar ist.

Satz 47. Die JACOBI–Iteration und die GAUSS–SEIDEL–Iteration für $Ax = b$ konvergieren jeweils falls A streng diagonal dominant ist.

Bemerkung 40. Für (7.13) war A s.d.d., für (7.14) war A nicht s.d.d.!

7.8.3 Zusammenfassung

Zunächst zerlegt man die Matrix A in 3 Summanden

$$A = L + D + R$$

mit den folgenden Eigenschaften:

- $L \dots$ strikt linke untere Dreiecksmatrix, d.h. die Diagonalelemente sind Null
- $R \dots$ strikt rechte obere Dreiecksmatrix, d.h. die Diagonalelemente sind Null

- $D \dots$ Diagonalmatrix

Die JACOBI–Iteration liefert für die i -te Komponente von x (vgl. Gleichung 7.12) die Relation

$$\begin{aligned} a_{ii}(x_i)_{k+1} &= (-a_{i1}x_1 - \dots - a_{in}x_n)_k + b_i \\ &= -\sum_{\substack{j=1 \\ j \neq i}}^n a_{ij}(x_j)_k + b_i, \quad i = 1, \dots, n \end{aligned}$$

In Matrzenschreibweise:

$$Dx_{k+1} = -(L + R)x_k + b$$

Wegen der Voraussetzung $a_{ii} \neq 0, i = 1, \dots, n$, ist die Matrix $D = \text{diag}(a_{11}, \dots, a_{nn})$ regulär.

$$\Rightarrow x_{k+1} = \underbrace{-D^{-1}(L + R)}_{T_J} x_k + D^{-1}b$$

$$x_{k+1} = T_J x_k + D^{-1}b$$

JACOBI–Verfahren (Gesamtschrittverfahren)

$T_J \dots$ JACOBI–Iterationsmatrix.

Das GAUSS–SEIDEL–Verfahren lieferte für die i -te Komponente von x die Relation

$$a_{ii}(x_i)_{k+1} = (-a_{i1}x_1 - \dots - a_{i,i-1}x_{i-1})_{k+1} + (-a_{i,i+1}x_{i+1} - \dots - a_{in}x_n)_k + b_i$$

d.h. in Matrzenschreibweise

$$Dx_{k+1} = -Lx_{k+1} - Rx_k + b \quad \Rightarrow \quad (D + L)x_{k+1} = -Rx_k + b$$

$D + L$ ist linke untere Dreiecksmatrix, deren Diagonalelemente $\neq 0$ sind. Daher ist $D + L$ invertierbar.

$$\Rightarrow x_{k+1} = \underbrace{-(D + L)^{-1}R}_{T_{GS}} x_k + (D + L)^{-1}b$$

$$x_{k+1} = T_{GS} x_k + (D + L)^{-1}b$$

GAUSS–SEIDEL–Verfahren (Einzelschrittverfahren)

$T_{GS} \dots$ GAUSS–SEIDEL–Iterationsmatrix.

7.8.4 SOR–Methode

Eine unter Umständen wesentliche Verbesserung des Konvergenzverhaltens liefert die sogenannte **SOR–Methode**: (= Methode der sukzessiven Überrelaxation = successive overrelaxation).

Dazu betrachten wir das GAUSS–SEIDEL–Verfahren und berechnen die Korrektur der i -ten Komponente im k -ten Schritt:

$$\begin{aligned} (\Delta x_i)_{k+1} &:= (x_i)_{k+1} - (x_i)_k \\ &= -\frac{1}{a_{ii}} \left[\sum_{j=1}^{i-1} a_{ij}(x_j)_{k+1} + \sum_{j=i+1}^n a_{ij}(x_j)_k \right] + \frac{b_i}{a_{ii}} - (x_i)_k \\ &= -\frac{1}{a_{ii}} \left[\sum_{j=1}^{i-1} a_{ij}(x_j)_{k+1} + \sum_{j=i}^n a_{ij}(x_j)_k \right] + \frac{b_i}{a_{ii}} \end{aligned}$$

Die Idee ist es nun, die Korrektur $(\Delta x_i)_{k+1}$ durch den Ausdruck

$$\omega \cdot (\Delta x_i)_{k+1}, \omega \in \mathbb{R},$$

zu ersetzen. Man nennt ω den **Relaxationsfaktor**. Ist $\omega > 1$, so spricht man von einer **Überrelaxation**. Damit erhält man nun

$$\begin{aligned} (x_i)_{k+1} &= (x_i)_k + \omega(\Delta x_i)_{k+1} \\ &= (x_i)_k - \frac{\omega}{a_{ii}} \left[\sum_{j=1}^{i-1} a_{ij}(x_j)_{k+1} + \sum_{j=i}^n a_{ij}(x_j)_k - b_i \right] \end{aligned}$$

Nach einer Multiplikation mit a_{ii} erhalten wir in Matrzenschreibweise

$$Dx_{k+1} = Dx_k - \omega Lx_{k+1} - \omega(D + R)x_k + \omega b$$

oder

$$(D + \omega L)x_{k+1} = [(1 - \omega)D - \omega R]x_k + \omega b$$

$\omega = 1$ liefert wieder die GAUSS-SEIDEL-Iteration, d.h. keine Beschleunigung. Man erhält eine Beschleunigung der Konvergenz erst mit einem Faktor $\omega > 1$!

Beispiel 117.

$$\begin{aligned} A &= \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix}}_L + \underbrace{\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}}_D + \underbrace{\begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix}}_R \\ D + \omega L &= \begin{pmatrix} 2 & 0 \\ -\omega & 2 \end{pmatrix} \\ (1 - \omega)D - \omega R &= \begin{pmatrix} 2(1 - \omega) & \omega \\ 0 & 2(1 - \omega) \end{pmatrix} \end{aligned}$$

Die Iteration lautet jetzt:

$$\begin{aligned} \begin{pmatrix} 2 & 0 \\ -\omega & 2 \end{pmatrix} x_{k+1} &= \begin{pmatrix} 2(1 - \omega) & \omega \\ 0 & 2(1 - \omega) \end{pmatrix} x_k + \omega b \\ x_{k+1} &= \underbrace{\begin{pmatrix} 1 - \omega & \frac{\omega}{2} \\ \frac{\omega(1 - \omega)}{2} & 1 - \omega + \frac{\omega^2}{4} \end{pmatrix}}_{T_\omega} x_k + \frac{\omega}{4} \begin{pmatrix} 2 & 0 \\ \omega & 2 \end{pmatrix} b \end{aligned}$$

Frage: Wie kann man ω optimal wählen? Die Bedingung für ein geeignetes ω lautet:

$$\rho(T_{\omega_{\text{opt}}}) \leq \rho(T_\omega), 0 < \omega < 2$$

Die Berechnung von ω wird i.a. schwierig!

In unserem Fall ist

$$\omega_{\text{opt}} = 4(2 - \sqrt{3}) \approx 1.07 \implies \rho \approx 0.07 \implies r = 1.15$$

Definition 68. Eine Matrix $A \in M(n \times n; \mathbb{R})$ mit der speziellen tridiagonalen Gestalt

$$A = \begin{pmatrix} D_1 & H_1 & & & & \\ K_1 & D_2 & H_2 & & & O \\ & K_2 & D_3 & & & \\ & & K_3 & \ddots & & \\ & & & \ddots & \ddots & H_{s-1} \\ O & & & & K_{s-1} & D_s \end{pmatrix}$$

mit den quadratischen Diagonalmatrizen D_i und den i.a. rechteckigen und beliebigen Matrizen H_i, K_i heißt **T-Matrix**.

Satz 48. Ist A eine T -Matrix mit $a_{ii} \neq 0$ und besitzt die JACOBI-Iterationsmatrix

$$T_J = D^{-1}(L + R)$$

nur reelle Eigenwerte λ_j und gilt

$$\rho(T_J) < 1,$$

so ist der optimale Relaxationsparameter ω_{opt} des SOR-Verfahrens gegeben durch

$$\omega_{\text{opt}} = \frac{2}{1 + \sqrt{1 - \rho(T_J)^2}} \quad (7.15)$$

Beispiel 118. Wir betrachten die Differentialgleichung $-u'' = f(x)$ von Bsp. 104 und approximieren sie mit $n = 21$, d.h. $h = \frac{1}{22}$, durch ein lineares Gleichungssystem mit der Koeffizientenmatrix

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 & 0 & \cdots \\ -1 & 2 & -1 & 0 & \cdots & \ddots \\ & & \ddots & \ddots & & \ddots \\ & & & -1 & 2 & -1 \\ & & & 0 & -1 & 2 \end{pmatrix} \in M(21 \times 21)$$

Die Matrix A ist nun eine T -Matrix. Wendet man das SOR-Verfahren an, so erhält man mit Glg. (7.15) für $\omega_{\text{opt}} \approx 1.75$.

In diesem Fall ist dann

$$\rho_{\text{SOR}} \approx 0.75$$

im Vergleich zu

$$\rho_J \approx 0.99 \quad \text{und} \quad \rho_{\text{GS}} \approx 0.98$$

Mit der SOR-Methode wird der Fehler bei jedem Schritt um 25% verbessert.

Da $\rho_J^{30} \approx \rho_{\text{SOR}}$ gilt, ist ein einziger SOR-Schritt so viel wert wie 30 JACOBI-Schritte.

7.9 Überbestimmte lineare Gleichungssysteme

7.9.1 Bestimmung von Näherungslösungen

Eine Matrix $A \in M(n \times n)$ mit linear unabhängigen Spaltenvektoren ist invertierbar, das Gleichungssystem $A \cdot x = b$ ist eindeutig lösbar mit $x = A^{-1} \cdot b$. Ist $A \in M(m \times n)$ mit $m > n$ eine Matrix mit linear unabhängigen Spaltenvektoren, so besitzt $A \cdot x = b$ i.a. keine Lösung.

Definition 69. Sei $A \in M(m \times n), b \in \mathbb{R}^m$. Man nennt $\bar{x} \in \mathbb{R}^n$ eine **Näherungslösung** (im quadratischen Mittel) oder Lösung nach der Methode der kleinsten Quadrate oder **approximative Lösung** von $Ax = b$, wenn für alle $x \in \mathbb{R}^n$ gilt

$$\|b - A\bar{x}\| \leq \|b - Ax\|.$$

Es gilt der folgende

Satz 49. Sei $A \in M(m \times n), b \in \mathbb{R}^m$. Dann besitzt $Ax = b$ stets mindestens eine approximative Lösung (im quadratischen Mittel).

Es gilt:

- \bar{x} ist eine Näherungslösung von $Ax = b$ genau dann, wenn \bar{x} eine Lösung ist des Systems

$$A^t A \bar{x} = A^t b.$$

- A besitzt genau dann linear unabhängige Spaltenvektoren, wenn $A^t A$ invertierbar ist. In diesem Fall ist die Näherungslösung von $Ax = b$ eindeutig bestimmt durch

$$\bar{x} = (A^t A)^{-1} A^t b.$$

Beispiel 119. $A \cdot x = b$ mit $A = \begin{pmatrix} 1 & 5 \\ 2 & -2 \\ -1 & 1 \end{pmatrix}$, $b = \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix}$

$$(A, b) = \begin{pmatrix} 1 & 5 & 3 \\ 2 & -2 & 2 \\ -1 & 1 & 5 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 5 & 3 \\ 0 & 3 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$\text{rang } A = 2$, $\text{rang}(A, b) = 3 \Rightarrow A \cdot x = b$ nicht lösbar.

Näherungslösung:

$$A^t A = \begin{pmatrix} 1 & 2 & -1 \\ 5 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 5 \\ 2 & -2 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 30 \end{pmatrix}$$

$$A^t b = \begin{pmatrix} 1 & 2 & -1 \\ 5 & -2 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 2 \\ 16 \end{pmatrix}$$

$$\begin{pmatrix} 6 & 0 \\ 0 & 30 \end{pmatrix} \bar{x} = \begin{pmatrix} 2 \\ 16 \end{pmatrix} \Rightarrow \bar{x} = \begin{pmatrix} 1/3 \\ 8/15 \end{pmatrix} \dots \text{ Näherungslösung}$$

7.9.2 Die Pseudoinverse

Definition 70. Sei $A \in M(m \times n)$ mit linear unabhängigen Spaltenvektoren. Man nennt

$$A^\# = (A^t A)^{-1} A^t$$

die **Pseudoinverse** von A . Es gilt $A^\# \in M(n \times m)$.

Beispiel 120. Man bestimme die Pseudoinverse von

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}$$

$$A^t A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 6 \\ 6 & 14 \end{pmatrix} \Rightarrow (A^t A)^{-1} = \frac{1}{6} \begin{pmatrix} 14 & -6 \\ -6 & 3 \end{pmatrix}$$

$$A^\# = (A^t A)^{-1} \cdot A^t = \frac{1}{6} \begin{pmatrix} 14 & -6 \\ -6 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 8 & 2 & -4 \\ -3 & 0 & 3 \end{pmatrix}$$

Bemerkung 41. Einige Eigenschaften der Pseudoinversen von A :

- (i) $AA^\# A = A$
- (ii) $A^\# AA^\# = A^\#$
- (iii) $AA^\#$ und $A^\# A$ sind symmetrisch.

7.9.3 Näherungslösung mittels QR-Zerlegung

Für Matrizen, die eine QR -Zerlegung besitzen (siehe Abschnitt 5.5), kann die Näherungslösung des Gleichungssystems $Ax = b$ schneller auf folgendem Weg berechnet werden. Verwendet man in der Relation

$$A^T A \bar{x} = A^T b$$

für A die Zerlegung $A = QR$, so erhält man

$$\begin{aligned} (QR)^T (QR) \bar{x} &= (QR)^T b \\ R^T Q^T Q R \bar{x} &= R^T Q^T b \end{aligned}$$

woraus wegen $Q^T Q = I$ folgt

$$R \bar{x} = Q^T b$$

oder

$$\bar{x} = R^{-1} Q^T b.$$

Bemerkung 42. Da R eine obere Dreiecksmatrix ist, ist es in der Praxis leichter, das System $R\bar{x} = Q^T b$ direkt durch Rückwärtseinsetzen zu lösen, als R zu invertieren und dann $R^{-1} Q^T b$ zu berechnen.

Beispiel 121. Man verwende die QR -Zerlegung um die Näherungslösung von $Ax = b$ mit

$$A = \begin{pmatrix} 1 & 2 & 2 \\ -1 & 1 & 2 \\ -1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix}, b = \begin{pmatrix} 2 \\ -3 \\ -2 \\ 5 \end{pmatrix}$$

zu gewinnen.

Von Beispiel 85 übernehmen wir die QR -Zerlegung $A = QR$ mit

$$Q = \begin{pmatrix} 1/2 & 3\sqrt{5}/10 & -\sqrt{6}/6 \\ -1/2 & 3\sqrt{5}/10 & 0 \\ -1/2 & \sqrt{5}/10 & \sqrt{6}/6 \\ 1/2 & \sqrt{5}/10 & \sqrt{6}/3 \end{pmatrix}, R = \begin{pmatrix} 2 & 1 & 1/2 \\ 0 & \sqrt{5} & 3\sqrt{5}/2 \\ 0 & 0 & \sqrt{6}/2 \end{pmatrix}$$

Damit folgt

$$Q^T b = \begin{pmatrix} 6 \\ 0 \\ \sqrt{6} \end{pmatrix}$$

und das System $R\bar{x} = Q^T b$ lautet

$$\begin{pmatrix} 2 & 1 & 1/2 \\ 0 & \sqrt{5} & 3\sqrt{5}/2 \\ 0 & 0 & \sqrt{6}/2 \end{pmatrix} \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \\ \sqrt{6} \end{pmatrix}$$

Leicht gewinnt man jetzt durch Rückwärtseinsetzen die Näherungslösung

$$\bar{x} = \begin{pmatrix} 4 \\ -3 \\ 2 \end{pmatrix}.$$

7.9.4 Näherungslösung mittels Singulärwertzerlegung

Ist $A^t A$ nicht invertierbar, dann existiert die Pseudoinverse im Sinne von Definition 70 nicht. Man kann aber unter Verwendung der Singulärwertzerlegung für jede Matrix eine Pseudoinverse definieren, die eine Verallgemeinerung der vorigen Definition darstellt:

Definition 71. Sei $A = U\Sigma V^t$ die Singulärwertzerlegung von $A \in M(m \times n)$ mit $\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}$, wobei $D \in M(r \times r)$ eine Diagonalmatrix darstellt, in deren Diagonale die von Null verschiedenen Singulärwerte $\sigma_1 \geq \sigma_2 \geq \dots \sigma_r > 0$ von A stehen. Die **Pseudoinverse (MOORE-PENROSE-Inverse)** $A^\#$ von A ist definiert als

$$A^\# = V\Sigma^\# U^t$$

$$\text{mit } \Sigma^\# = \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix} \in M(n \times m).$$

Beispiel 122. Man bestimme die Pseudoinverse von $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Die Singulärwertzerlegung von A ist (siehe Beispiel 98)

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ -1/\sqrt{2} & 1/\sqrt{2} & 0 \end{pmatrix} = U\Sigma V^t$$

Daraus folgt:

$$\Sigma^\# = \begin{pmatrix} 1/\sqrt{2} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

und weiters

$$A^\# = V\Sigma^\# U^t = \begin{pmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1/2 & 0 \\ 1/2 & 0 \\ 0 & 1 \end{pmatrix}$$

Beispiel 123. Man bestimme die Pseudoinverse von $A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$ (siehe Beispiel 99).

$$\text{SVD: } A = \begin{pmatrix} 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = U\Sigma V^t$$

$$\Sigma^\# = \begin{pmatrix} 1/\sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

$$A^\# = V\Sigma^\# U^t = \dots = \begin{pmatrix} 1/3 & 2/3 & -1/3 \\ 1/3 & -1/3 & 2/3 \end{pmatrix}.$$

Anmerkung: In diesem Beispiel gilt auch $A^\# = (A^t A)^{-1} A^t$.

Satz 50. Das Gleichungssystem $Ax = b$ mit $A \in M(m \times n), b \in \mathbb{R}^n, m > n$, besitzt genau eine Näherungslösung \bar{x} mit minimaler Länge. Sie ist gegeben durch

$$\bar{x} = A^\# b.$$

Beispiel 124. $Ax = b$ mit $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

$$\begin{array}{rcl} r + s = 0 \\ r + s = 1 \end{array} \quad \text{besitzt keine Lösung, die Spalten von } A \text{ sind linear abhängig}$$

$$\text{SVD von } A: A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} = U\Sigma V^t$$

$$A^\# = V\Sigma^\# U^t = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 1/2 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$\bar{x} = A^\# b = \frac{1}{4} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Diese Lösung erfüllt die Gleichung

$$r + s = \frac{1}{2}$$

Sie stellt in gewissem Sinn einen Kompromiß zwischen den beiden Gleichungen des ursprünglichen Systems dar.

Kapitel 8

Näherungsweise Bestimmung von Nullstellen – NEWTON-Iteration

Für die meisten Funktionen gibt es keine Formeln, die zu den Nullstellen dieser Funktion führen. In diesen Fällen gibt es die Möglichkeit, die Nullstellen durch das NEWTON'sche Verfahren näherungsweise zu bestimmen.

8.1 Der eindimensionale Fall

Wir betrachten eine auf dem Intervall $[a, b] \subset \mathbb{R}$ definierte Funktion $f(x)$ unter den folgenden

- Voraussetzungen:**
- $f(a) < 0 < f(b)$ oder $f(a) > 0 > f(b)$, d.h. nach dem Mittelwertsatz existiert in $[a, b]$ (mindestens) eine Nullstelle von $f(x)$
 - $f(x)$ sei auf $[a, b]$ zweimal stetig differenzierbar
 - $f'(x) \neq 0$ auf $[a, b]$

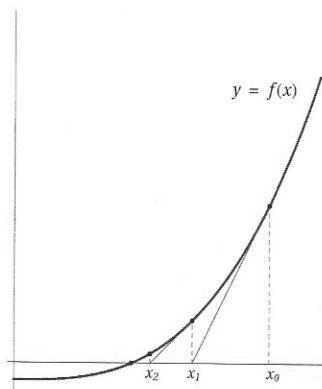
Zur näherungsweisen Berechnung der in (a, b) liegenden Nullstelle \tilde{x} (d.h. es gilt $f(\tilde{x}) = 0$) wählt man nun folgende Vorgangsweise:

1. Errichte in $(x_0, f(x_0))$, $x_0 \in (a, b)$, die Tangente t_0 an die Kurve $y = f(x)$

$$t_0 : y - f(x_0) = f'(x_0)(x - x_0)$$

2. Ermittle den Schnittpunkt der Tangente mit der x -Achse (d.h. $y = 0$)

$$-f(x_0) = f'(x_0)(x - x_0) \implies x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$



3. Tangente t_1 an die Kurve im Punkt $(x_1, f(x_1))$

$$t_1 : y - f(x_1) = f'(x_1)(x - x_1)$$

4. Ermittle den Schnittpunkt der Tangente mit der x -Achse, (d.h. $y = 0$)

$$\Rightarrow x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$$

5. Wiederhole diesen Vorgang bis zur gewünschten Genauigkeit.

Die **NEWTON-Iteration** lautet also

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} , \quad n = 0, 1, 2, \dots \quad (8.1)$$

Man erhält so eine Folge von Werten $\{x_n\}$.

Satz 51. (ohne Beweis)

Die durch (8.1) definierte Folge $\{x_n\}$ konvergiert (im allgemeinen) unter den getroffenen Voraussetzungen gegen ein \tilde{x} mit $f(\tilde{x}) = 0$, d.h. \tilde{x} ist eine Nullstelle der Funktion $f(x)$.

Beispiel 125. Gesucht ist die Nullstelle der Funktion $f(x) = \cos x - x$.

$$\begin{aligned} f'(x) &= -\sin x - 1 \\ \Rightarrow x_{n+1} &= x_n - \frac{\cos x_n - x_n}{-\sin x_n - 1} \end{aligned}$$

Aus der grafischen Darstellung der beiden Funktionen $f_1(x) = \cos x$ und $f_2(x) = x$ sieht man, dass deren Schnittpunkt in der Nähe von $x_0 = 0.5$ liegt. Wir starten also mit $x_0 = 0.5$ und erhalten folgende Werte

n	x_n	$f(x_n)$	$f'(x_n)$	$\frac{f(x_n)}{f'(x_n)}$	$x_n - \frac{f(x_n)}{f'(x_n)}$
0	0,5	0,3776	-1,4794	-0,2552	0,7552
1	0,7552	-0,0271	-1,6855	0,0161	0,7391
2	0,7391	-0,0000	-1,6737	0,0000	0,7391

Man erkennt, dass die Nullstelle bereits ab dem zweiten Schritt auf 4 Dezimalstellen genau bestimmt ist. Somit liegt die Nullstelle der Funktion $f(x)$ bei $\tilde{x} \approx 0,7391$.

Beispiel 126. Näherungsweise Bestimmung von \sqrt{a} für $a \in \mathbb{R}$:

$$\begin{aligned} f(x) &= x^2 - a = 0 \quad (x = \pm\sqrt{a}) \\ f'(x) &= 2x \\ \Rightarrow x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - a}{2x_n} = \frac{2x_n^2 - x_n^2 + a}{2x_n} = \frac{x_n^2 + a}{2x_n} \Rightarrow \end{aligned}$$

$$x_{n+1} = \frac{1}{2} \left(x_n + \frac{a}{x_n} \right) \quad \dots \quad \text{allgemeine Iteration zur Berechnung quadratischer Wurzeln}$$

Speziell zur näherungsweisen Berechnung von $\sqrt{3}$ wählt man $x_0 = 3$. Man erhält folgende Werte:

n	x_n	$x_{n+1} = \frac{1}{2} \left(x_n + \frac{a}{x_n} \right)$
0	3	2
1	2	1,75
2	1,75	1,732142
3	1,732142	1,732050
4	1,732050	1,732050

$$\Rightarrow \sqrt{3} \approx 1,732050$$

Eigenschaften des NEWTON-Verfahrens

- Sehr schnelle Konvergenz (nur wenige Iterationsschritte nötig)
- Es gilt die Fehlerabschätzung

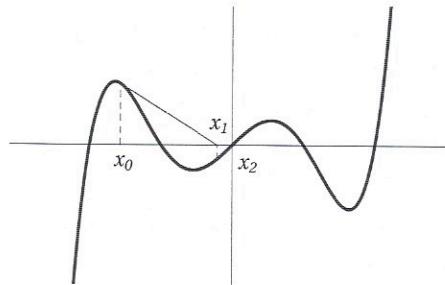
$$|x_{n+1} - \tilde{x}| \leq L^n |x_{n+1} - x_n| \leq L^n (b - a)$$

wenn $|f'(x)| \leq L$ für alle $x \in [a, b]$

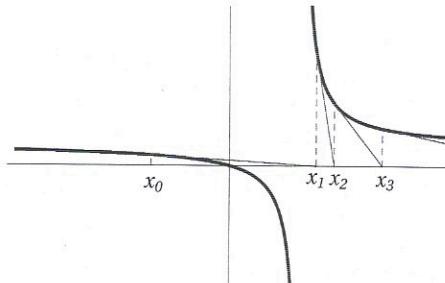
- Für die Ableitung der Funktion f muss ein analytischer Ausdruck vorliegen

Achtung: Das Verfahren kann divergieren bzw. eine andere als die gesuchte Nullstelle approximieren, falls der Startwert x_0 nicht nahe genug an der Nullstelle liegt.

- Man bekommt die “falsche” Nullstelle



- Man bewegt sich von der Nullstelle weg



8.2 NEWTONsches Verfahren für Systeme von Gleichungen

Nun betrachtet man ein System von n nichtlinearen Gleichungen in n Variablen x_1, \dots, x_n von der Form

$$\begin{aligned} f_1(x_1, \dots, x_n) &= 0 \\ f_2(x_1, \dots, x_n) &= 0 \\ &\vdots \\ f_n(x_1, \dots, x_n) &= 0 \end{aligned} \tag{8.2}$$

Mit dem Vektor $x = (x_1, \dots, x_n)^T$ und der Vektorfunktion

$$F(x) := \begin{pmatrix} f_1(x_1, \dots, x_n) \\ f_2(x_1, \dots, x_n) \\ \vdots \\ f_n(x_1, \dots, x_n) \end{pmatrix}$$

kann das System (8.2) auch geschrieben werden in der Form

$$F(x) = 0$$

In Verallgemeinerung der Iteration (8.1) wird nun die Ableitung der Funktion $f(x)$ durch die JACOBI-Matrix

$$J(x) = \begin{pmatrix} \frac{\partial f_1(x)}{\partial x_1} & \frac{\partial f_1(x)}{\partial x_2} & \cdots & \frac{\partial f_1(x)}{\partial x_n} \\ \frac{\partial f_2(x)}{\partial x_1} & \frac{\partial f_2(x)}{\partial x_2} & \cdots & \frac{\partial f_2(x)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n(x)}{\partial x_1} & \frac{\partial f_n(x)}{\partial x_2} & \cdots & \frac{\partial f_n(x)}{\partial x_n} \end{pmatrix}$$

ersetzt, wobei zur Abkürzung $f_k(x) := f_k(x_1, \dots, x_n)$ gesetzt wurde.

Bemerkung. Die JACOBI-Matrix findet auch Anwendung bei der Berechnung von Mehrfachintegralen.

Mit einem vorgegebenen Startwert

$$x^{(0)} = \begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \\ \vdots \\ x_n^{(0)} \end{pmatrix}$$

lautet die NEWTON-Iteration jetzt

$$x^{(k+1)} = x^{(k)} - (J(x^{(k)}))^{-1} \cdot F(x^{(k)}), \quad k = 0, 1, 2, \dots \quad (8.3)$$

Eine Schwachstelle des NEWTON-Verfahrens für Systeme ist die Notwendigkeit, die JACOBI-Matrix $J(x^{(k)})$ bei jedem Iterationsschritt zu invertieren. Das kann man vermeiden, wenn man folgendermaßen vorgeht:

1. Schritt: Bestimme den Vektor $y^{(k)}$ aus dem Gleichungssystem

$$J(x^{(k)}) \cdot y^{(k)} = -F(x^{(k)})$$

2. Schritt: Berechne

$$x^{(k+1)} = x^{(k)} + y^{(k)}$$

Beispiel 127.

$$\begin{aligned} 3x_1 - \cos(x_2 x_3) - \frac{1}{2} &= 0 \\ x_1^2 - 81(x_2 + 0.1)^2 + \sin x_3 + 1.06 &= 0 \\ e^{-x_1 x_2} + 20x_3 + \frac{10\pi - e}{3} &= 0 \end{aligned}$$

Startwert:

$$x^{(0)} = \begin{pmatrix} 0.1 \\ 0.1 \\ -0.1 \end{pmatrix}$$

Die JACOBI-Matrix lautet

$$J(x_1, x_2, x_3) = \begin{pmatrix} 3 & x_3 \sin(x_2 x_3) & x_2 \sin(x_2 x_3) \\ 2x_1 & -162(x_2 + 0.1) & \cos x_3 \\ -x_2 e^{-x_1 x_2} & -x_1 e^{-x_1 x_2} & 20 \end{pmatrix}$$

Die über die NEWTON-Iteration (8.3) erhaltenen Ergebnisse sind in folgender Tabelle zusammengestellt.

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\ x^{(k)} - x^{(k-1)}\ _\infty$
0	0.10000000	0.10000000	-0.10000000	
1	0.50003702	0.01946686	-0.52152047	0.422
2	0.50004593	0.00158859	-0.52355711	$1.79 \cdot 10^{-2}$
3	0.50000034	0.00001244	-0.52359845	$1.58 \cdot 10^{-3}$
4	0.50000000	0.00000000	-0.52359877	$1.24 \cdot 10^{-5}$
5	0.50000000	0.00000000	-0.52359877	0

Die exakte Lösung des Gleichungssystems ist

$$x_1 = \frac{1}{2}, x_2 = 0, x_3 = -\frac{\pi}{6}$$

Man sieht, dass das NEWTONSche Verfahren hier sehr rasch konvergiert.

Bemerkung 43.

- Es ist nicht immer einfach, geeignete Startwerte zu bestimmen, die zu einer Lösung führen.
- Diese Iteration ist i.a. sehr rechenaufwendig.

Es gibt aber geeignete Verfahren, die bei diesen beiden Problemen helfen.

Kapitel 9

Numerische Bestimmung von Eigenwerten

9.1 Rückblick und Problemstellung

Die Eigenwerte einer Matrix $A = (a_{ij}) \in M(n \times n)$, $a_{ij} \in \mathbb{R}$, ergeben sich nach Abschnitt 6 als Nullstellen des charakteristischen Polynoms, d.h. sie sind Lösungen der charakteristischen Gleichung

$$P_n(\lambda) = \det(A - \lambda I) = 0$$

Nach dem Fundamentalsatz der Algebra besitzt $P_n(\lambda)$ in \mathbb{C} genau n Nullstellen, d.h. $P_n(\lambda)$ kann angegeben werden in der Form

$$P_n(\lambda) = a(\lambda - \lambda_1)^{k_1} \cdot (\lambda - \lambda_2)^{k_2} \cdot \dots \cdot (\lambda - \lambda_r)^{k_r}$$

$\lambda_1, \dots, \lambda_r$ verschiedene Nullstellen von $P_n(\lambda)$
 k_1, \dots, k_r ihre Vielfachheit

Zunächst erscheint das charakteristische Polynom von A der geeignete Ausdruck zur Bestimmung der Eigenwerte von A zu sein. Jedoch ist die Konditionierung eines Polynoms oft nicht gut, d.h. kleine Abänderungen in den Koeffizienten können große Auswirkungen auf die Nullstellen haben.

Beispiel 128.

1. Das Polynom $P_{20}(\lambda) = (\lambda - 1) \cdot (\lambda - 2) \cdot \dots \cdot (\lambda - 20)$ besitzt die Nullstellen $\lambda_i = i, i = 1, \dots, 20$.
2. Das Polynom $Q_{20}(\lambda) = P_{20}(\lambda) - 2^{-23} \cdot \lambda^{19}$, das sich nur geringfügig von $P_{20}(\lambda)$ unterscheidet, besitzt die in Abb. 9.1 als Punkte in der komplexen Zahlenebene dargestellten Nullstellen. Die

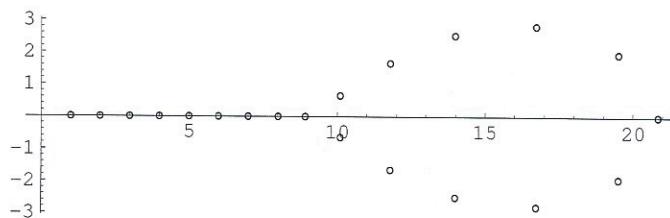


Abbildung 9.1: Nullstellen von $Q_{20}(\lambda)$

Lage der Nullstellen hat sich jetzt (obwohl das Polynom nur geringfügig abgeändert worden ist) dramatisch verändert!

9.2 Bestimmung der Eigenwerte durch Iteration – Potenzmethode

Die Matrix $A = (a_{ij}) \in M(n \times n)$ besitze die Eigenwerte $\lambda_1, \dots, \lambda_n \in \mathbb{C}$

Voraussetzungen:

1. Es gelte

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|$$

λ_1 wird als **dominanter Eigenwert** bezeichnet.

2. Es existieren n linear unabhängige Eigenvektoren der Matrix A . Dies ist z.B. bei reellen symmetrischen Matrizen immer der Fall. Die Eigenvektoren u_1, \dots, u_n von A bilden dann eine Basis des \mathbb{R}^n .

Die im folgenden beschriebene Methode liefert eine gute Näherung des dominanten Eigenwertes λ_1 : Wähle ein $x_0 \in \mathbb{R}^n$ beliebig und bilde

$$\begin{aligned} x_1 &= A \cdot x_0 \\ x_2 &= A \cdot x_1 = A^2 \cdot x_0 \\ x_3 &= A \cdot x_2 = A^3 \cdot x_0 \\ &\vdots \\ x_s &= A^s \cdot x_0, \quad s \in \mathbb{N} \end{aligned}$$

Da $\{u_1, \dots, u_n\}$ eine Basis ist, kann der Vektor x_0 mit Hilfe dieser Basisvektoren dargestellt werden:

$$\exists \alpha_1, \dots, \alpha_n \in \mathbb{R} : x_0 = \alpha_1 \cdot u_1 + \dots + \alpha_n \cdot u_n$$

$$\begin{aligned} \Rightarrow x_s &= A^s \cdot x_0 = A^s \cdot (\alpha_1 \cdot u_1 + \dots + \alpha_n \cdot u_n) \\ &= \alpha_1 \cdot A^s \cdot u_1 + \dots + \alpha_n \cdot A^s \cdot u_n \end{aligned}$$

Wegen

$$A \cdot u_i = \lambda_i \cdot u_i, \quad i = 1, \dots, n$$

folgt nun

$$\begin{aligned} A \cdot u_1 &= \lambda_1 \cdot u_1 \\ A^2 \cdot u_i &= A \cdot (\underbrace{A \cdot u_i}_{\lambda_i \cdot u_i}) = \lambda_i^2 \cdot u_i \\ A^3 \cdot u_i &= \lambda_i^3 \cdot u_i, \\ &\vdots \\ A^s \cdot u_i &= \lambda_i^s \cdot u_i, \quad i = 1, \dots, n, \quad s \in \mathbb{N} \end{aligned}$$

und weiters

$$\begin{aligned} \Rightarrow x_s &= \alpha_1 \cdot \lambda_1^s \cdot u_1 + \alpha_2 \cdot \lambda_2^s \cdot u_2 + \dots + \alpha_n \cdot \lambda_n^s \cdot u_n, \quad \lambda_1 \neq 0 \\ &= \lambda_1^s \cdot \left(\alpha_1 \cdot u_1 + \alpha_2 \cdot \left(\frac{\lambda_2}{\lambda_1} \right)^s \cdot u_2 + \dots + \alpha_n \cdot \left(\frac{\lambda_n}{\lambda_1} \right)^s \cdot u_n \right) \end{aligned}$$

Wegen $\lambda_k = |\lambda_k| e^{i \arg \lambda_k}$ und $|\lambda_1| > |\lambda_2|$ folgt nun:

$$\left(\frac{\lambda_2}{\lambda_1} \right)^s = \left(\frac{|\lambda_2|}{|\lambda_1|} \right)^s \cdot \underbrace{e^{is(\arg \lambda_2 - \arg \lambda_1)}}_{\text{beschränkt}} \rightarrow 0 \text{ für } s \rightarrow \infty$$

D.h. für große Werte von s gilt:

$$x_s \approx \lambda_1^s \cdot \alpha_1 \cdot u_1$$

Analog folgt

$$x_{s+1} \approx \lambda_1^{s+1} \cdot \alpha_1 \cdot u_1 = \lambda_1 \cdot x_s \quad (9.1)$$

Folgerung: Die Folge $\{x_s\}$ konvergiert (i.a.) gegen einen Eigenvektor u_1 bzw. ein skalares Vielfaches davon.

Multipliziert man die Ausdrücke in (9.1) mit dem Vektor x_s so erhält man

$$\langle x_{s+1}, x_s \rangle \approx \lambda_1 \cdot \langle x_s, x_s \rangle, \quad \|x_s\| \neq 0$$

$$\Rightarrow \lambda_1 \approx \frac{\langle x_{s+1}, x_s \rangle}{\langle x_s, x_s \rangle}$$

Definition 72. Der Quotient

$$\mu_s = \frac{\langle x_{s+1}, x_s \rangle}{\langle x_s, x_s \rangle}$$

wird als RAYLEIGH–Quotient bezeichnet.

Folgerung: Der RAYLEIGH–Quotient approximiert den dominanten Eigenwert!

Beispiel 129. Man bestimme den dominanten Eigenwert der Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 1 \end{pmatrix}$$

näherungsweise.

Wir beginnen mit

$$x_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Dann berechnet man

$$x_1 = A \cdot x_0 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad x_2 = A \cdot x_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 14 \\ 28 \\ 18 \end{pmatrix}$$

$$x_3 = A \cdot x_2 = \begin{pmatrix} 124 \\ 248 \\ 228 \end{pmatrix}, \quad x_4 = A \cdot x_3 = \begin{pmatrix} 1304 \\ 2608 \\ 2088 \end{pmatrix}$$

Damit folgt

$$\mu_3 = \frac{\langle x_4, x_3 \rangle}{\langle x_3, x_3 \rangle} \approx 9.968 \quad \Rightarrow \quad \lambda_1 \approx 9.968$$

Exakte Rechnung liefert folgende EW:

$$\lambda_1 = 10, \quad \lambda_2 = 0, \quad \lambda_3 = -4$$

Hätte man λ_1 mit Hilfe von μ_2 angenähert, so wäre

$$\lambda_1 \approx \mu_2 = \frac{\langle x_3, x_2 \rangle}{\langle x_2, x_2 \rangle} \approx 9.80$$

d.h. die Genauigkeit steigt (i.a.) mit wachsendem s .

Achtung. Es könnte Probleme geben, wenn $\alpha_1 = 0$ ist. Das wäre dann der Fall, wenn bei einer orthonormierten Basis $\{u_1, \dots, u_n\}$ der Startvektor x_0 orthogonal zu u_1 steht:

$$\begin{aligned} x_0 &= \alpha_1 \cdot u_1 + \dots + \alpha_n \cdot u_n \quad | \cdot u_1 \\ < x_0, u_1 > &= \underbrace{\alpha_1 < u_1, u_1 >}_1 + \underbrace{\alpha_2 < u_2, u_1 >}_0 + \dots + \underbrace{\alpha_n < u_n, u_1 >}_0 \\ \alpha_1 &= < x_0, u_1 > = 0 \Leftrightarrow u_1 \perp x_0 \end{aligned}$$

Dieser Fall kommt allerdings sehr selten vor.

Hinweis: Ernsthaft Probelme kann es geben, wenn ein Eigenwert λ_2 dem Betrag nach sich nur wenig von $|\lambda_1|$ unterscheidet.

In diesem Fall wird die Approximation mit wachsendem s ungenau.

Beispiel 130.

$$\begin{aligned} A &= \begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix}, \quad x_0 = \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \quad x_1 = A \cdot x_0 = \begin{pmatrix} 5 \\ 15 \end{pmatrix} \\ x_2 &= \begin{pmatrix} 75 \\ -25 \end{pmatrix}, \quad x_3 = \begin{pmatrix} 125 \\ 375 \end{pmatrix}, \quad x_4 = \begin{pmatrix} 1875 \\ -625 \end{pmatrix} \\ \mu_3 &= \frac{< x_4, x_3 >}{< x_3, x_3 >} = 0 \end{aligned}$$

Direkte Rechnung liefert:

$$\begin{vmatrix} 3 - \lambda & 4 \\ 4 & -3 - \lambda \end{vmatrix} = \lambda^2 - 25 = 0 \Rightarrow \lambda_{1,2} = \pm 5$$

D.h. hier gilt $|\lambda_1| = |\lambda_2| = 5$

Satz 52. Sei A eine n -reihige, symmetrische reelle Matrix. Sei $x_0 \neq 0$ beliebig. Dann konvergiert der Rayleigh-Quotient

$$\mu_s = \frac{< x_{s+1}, x_s >}{< x_s, x_s >} = \frac{< A \cdot x_s, x_s >}{< x_s, x_s >}$$

gegen den dominanten Eigenwert λ_1 . Setzt man $\mu_s = \lambda_1 + \varepsilon$, d.h. ε ist der Fehler in μ_s , so gilt:

$$|\varepsilon| \leq \sqrt{\frac{< x_{s+1}, x_{s+1} >}{< x_s, x_s >} - \mu_s^2}$$

9.3 Erweiterte Potenzmethode

Sei $A = (a_{ij}) \in M(n \times n)$ eine reelle Matrix mit den Eigenwerten $\lambda_1, \dots, \lambda_n$, mit $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$ und den linear unabhängigen Eigenvektoren u_1, \dots, u_n . Die Potenzmethode liefert eine gute Näherung des dominanten Eigenwertes λ_1 . Nun soll gezeigt werden, wie man noch weitere Eigenwerte approximieren kann. Dazu wählt man $\alpha \in \mathbb{R}$ so dass $\alpha \neq \lambda_i$, $i = 1, \dots, n$, und betrachtet die Matrix

$$B := A - \alpha I$$

Die Eigenwertgleichung lautet nun

$$\begin{aligned} B \cdot u_i &= (A - \alpha I) \cdot u_i = \underbrace{A \cdot u_i}_{\lambda_i u_i} - \alpha \cdot u_i = (\lambda_i - \alpha) u_i \\ \Rightarrow A - \alpha I &\text{ besitzt die Eigenwerte } \lambda_i - \alpha \end{aligned}$$

Bemerkung 44. Ist λ ein Eigenwert der regulären Matrix $A \in m(n \times n)$, so ist

$$\frac{1}{\lambda} \text{ Eigenwert von } A^{-1}$$

$$\Rightarrow (A - \alpha I)^{-1} \text{ besitzt die Eigenwerte } \frac{1}{\lambda_i - \alpha}$$

Wir wenden nun die Potenzmethode auf die Matrix $(A - \alpha I)^{-1}$ an:
Ihre Eigenwerte sind

$$\frac{1}{\lambda_1 - \alpha}, \frac{1}{\lambda_2 - \alpha}, \dots, \frac{1}{\lambda_n - \alpha}$$

Wir approximieren hier den Eigenwert λ_j , für den gilt:

$$|\frac{1}{\lambda_j - \alpha}| > |\frac{1}{\lambda_i - \alpha}|, i = 1, \dots, n, \quad i \neq j, \quad \text{d.h.} \quad |\lambda_i - \alpha| > |\lambda_j - \alpha|$$

D.h. es wird der Eigenwert λ_j approximiert, der von α den geringsten Abstand hat.

Beispiel 131.

$$A = \begin{pmatrix} 1 & -1 & 2 \\ -2 & 0 & 5 \\ 6 & -3 & 6 \end{pmatrix}$$

1.Schritt: Berechnung des dominanten Eigenwerts λ_1 von A :

$$x_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mu_s \rightarrow \lambda_1 = 5 \quad (\text{Berechnung z.B. mit MAPLE})$$

2.Schritt: Berechnung des kleinsten Eigenwerts von A :

Setze $B = (A - \alpha I)^{-1}$ mit $\alpha = 0$, d.h. wir verwenden A^{-1} .

x_0 wählen wir wie oben:

$$x_1 = A^{-1} \cdot x_0, \dots, x_s = (A^{-1})^s \cdot x_0 \\ \mu_s \rightarrow \lambda_B = -1 \quad \text{Eigenwert von } B = A^{-1} \quad (\text{Berechnung mit MAPLE})$$

$$\Rightarrow \lambda_2 = \frac{1}{\lambda_B} = \frac{1}{-1} = -1 \quad \text{Eigenwert von } A$$

3.Schritt: Berechnung des 3. Eigenwerts λ_3 : Er liegt (betragmäßig) irgendwo

zwischen den Eigenwerten λ_1 und λ_2 :

$$\text{Wähle also } \alpha = \frac{\lambda_1 + \lambda_2}{2} = \frac{5 - 1}{2} = 2$$

Mit $B = (A - 2 \cdot I)^{-1}$ und x_0 wie oben folgt:

$$x_1 = B \cdot x_0, \dots, x_s = B^s \cdot x_0 \\ \mu_s \rightarrow \hat{\lambda}_3 = 1 \quad (\text{EW von } B = (A - 2 \cdot I)^{-1}) \quad (\text{Berechnung mit MAPLE}) \\ \Rightarrow \hat{\lambda}_3 = \frac{1}{\lambda_3} = 1 \quad \text{EW von } A - 2 \cdot I \quad \Rightarrow \frac{1}{\lambda_3} = \lambda_3 - 2 \quad \dots \lambda_3 \text{ EW von } A \\ \Rightarrow \lambda_3 = 3 \quad 3. \text{ EW von } A.$$

9.4 GERSCHGORIN-Kreise

Frage: Wie kann man sich – in Vorbereitung auf ein numerisches Verfahren z.B. zur Bestimmung der Größe α in der verallgemeinerten Potenzmethode – einen Überblick darüber verschaffen, wo die Eigenwerte einer Matrix liegen?

Definition 73. Für $A = (a_{ij}) \in M(n \times n)$ mit reellen oder komplexen Elementen sei

$$r_i = \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|, c_i = \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ji}|, i = 1, \dots, n$$

Für $i = 1, \dots, n$, heißen dann die Mengen

$$R_i = \{z \in \mathbb{C} \mid |z - a_{ii}| \leq r_i\} \quad \text{bzw.} \quad C_i = \{z \in \mathbb{C} \mid |z - a_{ii}| \leq c_i\}$$

GERSCHGORIN-Kreisscheiben. Das sind Kreisscheiben in der komplexen Zahlenebene mit den Mittelpunkten in a_{ii} und den Radien r_i bzw. c_i .

Satz 53. Ist λ ein Eigenwert von A , so gilt

$$(i) \quad \lambda \in R_i \quad \text{für ein } i \in \{1, \dots, n\} \quad \text{und}$$

$$(ii) \quad \lambda \in C_i \quad \text{für ein } i \in \{1, \dots, n\}$$

Beweis. Für (i): Sei $\lambda \in \mathbb{C}$ ein Eigenwert von $A = (a_{ij}) \in M(n \times n)$. Dann gibt es einen Vektor $x \in \mathbb{C}^n$ mit $x \neq 0$ (= Eigenvektor zum Eigenwert λ) sodass

$$A \cdot x = \lambda \cdot x \tag{9.2}$$

Sei die Komponente x_j des Vektors x die betragsmäßig größte, d.h. für ein $j \in \{1, \dots, n\}$ gilt

$$|x_j| = \max\{|x_1|, \dots, |x_n|\} \neq 0 \tag{9.3}$$

dann folgt $\left| \frac{x_i}{x_j} \right| \leq 1$ für alle $i \in \{1, \dots, n\}$. Wegen (9.2) gilt jetzt

$$\sum_{k=1}^n a_{jk} x_k = \lambda x_j$$

woraus folgt

$$\lambda x_j - a_{jj} x_j = \sum_{\substack{k=1 \\ k \neq j}}^n a_{jk} x_k$$

bzw.

$$(\lambda - a_{jj}) x_j = \sum_{\substack{k=1 \\ k \neq j}}^n a_{jk} x_k$$

Damit gilt

$$|\lambda - a_{jj}| |x_j| = \left| \sum_{\substack{k=1 \\ k \neq j}}^n a_{jk} x_k \right| \leq \sum_{\substack{k=1 \\ k \neq j}}^n |a_{jk}| |x_k| \leq \sum_{\substack{k=1 \\ k \neq j}}^n |a_{jk}| |x_j| = r_j |x_j|$$

letzteres wegen (9.3). Somit erhält man schließlich

$$|\lambda - a_{jj}| \leq r_j \Rightarrow \lambda \in R_j$$

Da die Matrix A und ihre Transponierte dieselben Eigenwerte besitzen, gilt nun auch die Aussage (ii).

Beispiel 132.

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 7 & 0 \\ -1 & 0 & -5 \end{pmatrix}$$

ist symmetrisch und reell. Daher sind auch alle Eigenwerte reell. Wir berechnen zunächst

$$\begin{array}{ll} r_1 = |2| + |-1| = 3 & c_1 = |2| + |-1| = 3 \\ r_2 = |2| + |0| = 2 & c_2 = |2| + |0| = 2 \\ r_3 = |-1| + |0| = 1 & c_3 = |-1| + |0| = 1 \end{array}$$

Die Kreisscheiben sind nun:

$$R_1 = C_1 = \{z \in \mathbb{C} \mid |z - a_{11}| \leq r_1\} = \{z \in \mathbb{C} \mid |z - 1| \leq 3\}$$

$$R_2 = C_2 = \{z \in \mathbb{C} \mid |z - 7| \leq 2\}$$

$R_3 = C_3 = \{z \in \mathbb{C} \mid |z + 5| \leq 1\}$ Die Eigenwerte liegen alle auf der reellen Achse und damit in den



Abbildung 9.2: Intervalle für die Eigenwerte

in der Abb. 9.2 angegebenen Intervallen, die sich als Schnittmenge der GERSCHGORIN-Kreisscheiben mit der Menge der reellen Zahlen ergeben. Die Eigenwerte von A sind angenähert:

$$\lambda_1 \approx 7.61, \lambda_2 \approx 0.559, \lambda_3 \approx -5.17$$

sie ergeben sich als die Nullstellen des charakteristischen Polynoms:

$$\det(A - \lambda I) = \lambda^3 - 3\lambda^2 - 38\lambda + 22 = 0.$$

Bemerkung 45. 1. Ist die Vereinigung von k solchen Kreisscheiben disjunkt von den anderen Kreisscheiben, so liegen in dieser Vereinigung genau k Eigenwerte.

2. Ist jede der n GERSCHGORIN-Kreisscheiben disjunkt von den übrigen, so gibt es in jeder genau einen Eigenwert.

Beispiel 133.

$$A = \begin{pmatrix} 1 & 0.1 & 2 \\ 0.01 & 10 & 10 \\ -0.01 & 1 & 100 \end{pmatrix}$$

Mit Hilfe der Radien

$$\begin{array}{ll} r_1 = 2.1 & c_1 = 0.02 \\ r_2 = 10.01 & c_2 = 1.1 \\ r_3 = 1.01 & c_3 = 12 \end{array}$$

werden die GERSCHGORIN-Kreise

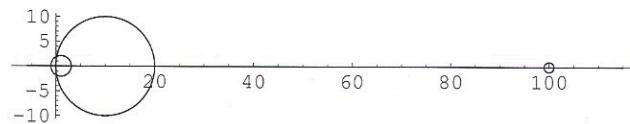


Abbildung 9.3: Die GERSCHGORIN-Kreise R_i

$$\begin{array}{ll} R_1 = \{z \mid |z - 1| \leq 2.1\} & C_1 = \{z \mid |z - 1| \leq 0.02\} \\ R_2 = \{z \mid |z - 10| \leq 10.01\} & C_2 = \{z \mid |z - 10| \leq 1.1\} \\ R_3 = \{z \mid |z - 100| \leq 1.01\} & C_3 = \{z \mid |z - 100| \leq 12\} \end{array}$$

angegeben.

Ein Eigenwert liegt in R_3 , zwei Eigenwerte liegen in $R_1 \cup R_2$ (siehe Abb. 9.3).

In C_1, C_2 und C_3 liegt jeweils ein Eigenwert (siehe Abb. 9.4): Die numerische Berechnung der Eigen-

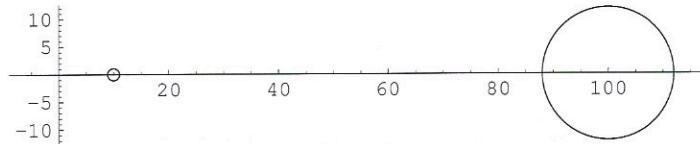


Abbildung 9.4: Die GERSCHGORIN-Kreise C_i

werte liefert folgendes Ergebnis:

$$\lambda_1 \approx 1.0001 ; \lambda_2 \approx 9.8891 ; \lambda_3 \approx 100.11 .$$

9.5 Approximation von Eigenwerten mittels QR-Zerlegung

Sei $A \in M(m \times n)$, $m \geq n$, eine Matrix mit linear unabhängigen Spaltenvektoren. Dann kann A stets in ein Produkt von der Form

$$A = QR \quad (9.4)$$

zerlegt werden (siehe Abschnitt 5.5). Nun bildet man eine Folge $\{A_1, A_2, A_3, \dots\}$ von Matrizen nach dem folgenden Bildungsgesetz:

$$A_1 = RQ \quad (9.5)$$

Die QR-Zerlegung von A_1 sei

$$A_1 = Q_1 R_1$$

Setze nun

$$A_2 = R_1 Q_1$$

Die QR-Zerlegung von A_2 sei

$$A_2 = Q_2 R_2$$

Setze nun

$$A_3 = R_2 Q_2$$

D.h. für $k \geq 1$ zerlegt man also A_k in der Form

$$A_k = Q_k R_k \quad (9.6)$$

und setzt dann

$$A_{k+1} = R_k Q_k \quad (9.7)$$

Bemerkung 46. Die Matrizen $A_k, k \geq 1$, sind ähnlich zu A , d.h. sie besitzen dieselben Eigenwerte wie A (siehe Abschnitt 6.2, Satz 32).

Beweis. (Durch vollständige Induktion)

1. A_1 ist ähnlich zu A , denn es gilt:

Da R regulär ist, folgt aus (9.5) mit (9.4)

$$A_1 = RQ = R(AR^{-1}) = RAR^{-1}$$

d.h. A_1 ist ähnlich zu A (siehe Def. 56).

2. A_{k+1} ist ähnlich zu A_k , $k \geq 1$, denn es gilt wegen (9.6) und (9.7):

$$A_{k+1} = R_k Q_k = R_k A_k R_k^{-1}.$$

Satz 54. Sind alle Eigenwerte von A reell und betragsmäßig voneinander verschieden, so konvergiert die Folge $\{A_k\}$ der gemäß (9.7) gebildeten Matrizen gegen eine obere Dreiecksmatrix U . In der Diagonale von U stehen dann die Eigenwerte von A .

Beispiel 134.

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

Zunächst gewinnt man

$$Q = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad R = Q^T A = \frac{1}{\sqrt{2}} \begin{pmatrix} 2 & 3 \\ 0 & 3 \end{pmatrix}$$

und damit

$$A_1 = RQ = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 0 & 3 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 5 & 1 \\ 3 & 3 \end{pmatrix}$$

Mit der Zerlegung

$$A_1 = Q_1 R_1$$

mit

$$Q_1 = \begin{pmatrix} 0.8575 & -0.5145 \\ 0.5145 & 0.8575 \end{pmatrix}, \quad R_1 = \begin{pmatrix} 2.9155 & 1.2005 \\ 0 & 1.0290 \end{pmatrix}$$

gewinnt man nun

$$A_2 = R_1 Q_1 = \begin{pmatrix} 3.1176 & -0.4706 \\ 0.5294 & 0.8824 \end{pmatrix}$$

Analog erhält man nun

$$A_3 = \begin{pmatrix} 3.0647 & 0.8412 \\ -0.1588 & 0.9353 \end{pmatrix}$$

$$A_4 = \begin{pmatrix} 3.0237 & -0.9494 \\ 0.0506 & 0.9763 \end{pmatrix}$$

$$A_5 = \begin{pmatrix} 3.0081 & 0.9834 \\ -0.0166 & 0.9919 \end{pmatrix}$$

$$A_6 = \begin{pmatrix} 3.0027 & -0.9945 \\ 0.0055 & 0.9973 \end{pmatrix}$$

Die Eigenwerte von A sind $\lambda_1 = 3$ und $\lambda_2 = 1$, sie werden durch die Diagonalelemente von A_6 schon recht gut approximiert.

Kapitel 10

Interpolation und Polynomapproximation

10.1 Einleitung

Im allgemeinen gehen Ingenieure und Wissenschaftler davon aus, dass die Beziehungen zwischen Variablen in einem Problem analytisch dargestellt werden können. Diese Darstellung kann mit den Daten des Problems ungefähr reproduziert werden. Letztendlich möchte man damit die Werte an den dazwischen liegenden Stellen bestimmen, das Integral oder die Ableitung der zugrunde liegenden Funktion approximieren oder einfach eine stetige glatte Darstellung des interessierenden Phänomens geben. Die Interpolation befaßt sich mit der Problematik, eine Funktion zu bestimmen, die eine Menge von gegebenen Werten genau darstellt. Der elementarste Typ der Interpolation besteht darin, ein Polynom auf eine Menge von Werten zurückzuführen. Polynome besitzen Ableitungen und Integrale, die wiederum Polynome sind, so daß sie eine natürliche Auswahl zur Approximation von Ableitungen und Integralen bilden.

Hier wird gezeigt, dass approximierende Polynome leicht konstruiert werden können und dass die Menge der Polynome hinreichend groß ist, um jede stetige Funktion beliebig genau darzustellen.

Satz 55. (TAYLOR'scher Satz)

Vorgelegt sei eine Funktion $f \in C^n[a, b]$, $f^{(n+1)}$ existiere auf $[a, b]$, x_0 sei eine Zahl in $[a, b]$. Für jedes x in $[a, b]$ existiert eine Zahl $\xi(x)$ zwischen x_0 und x mit

$$f(x) = P_n(x) + R_n(x)$$

wobei

$$\begin{aligned} P_n(x) &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n \\ &= \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k \end{aligned}$$

und

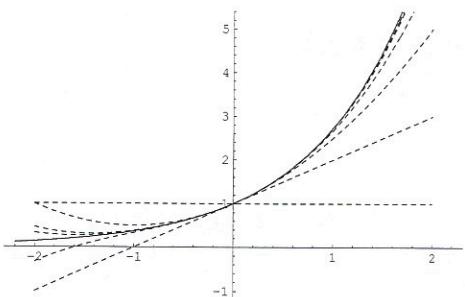
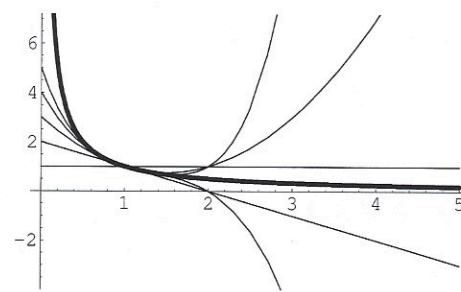
$$R_n(x) = \frac{f^{(n+1)}(\xi(x))}{(n+1)!}(x - x_0)^{n+1}$$

Man nennt $P_n(x)$ *n*-tes TAYLOR'sches Polynom, die Funktion $R_n(x)$ heißt **Restglied** (oder **Abbruchfehler**).

Bemerkung 47. Die TAYLOR'schen Polynome einer gegebenen Funktion stimmen zwar in einer Umgebung einer festen Stützstelle x_0 mit dieser Funktion sehr gut überein, diese Genauigkeit bezieht sich allerdings nur auf diese Stelle.

In Abb.10.1 sind die Funktion

$$f(x) = e^x$$

Abbildung 10.1: $f(x) = e^x$ Abbildung 10.2: $f(x) = \frac{1}{x}$

und die ersten 6 TAYLOR-Polynome $P_0(x), \dots, P_5(x)$ über dem Intervall $[-2, 2]$ dargestellt. Man sieht, dass der Fehler auch für Polynome höherer Ordnung größer wird, wenn man sich weiter von Null entfernt.

Beispiel 135. $f(x) = \frac{1}{x}$, Entwicklungspunkt $x_0 = 1$.

$$P_n(x) = \sum_{k=0}^n \frac{f^{(k)}(1)}{k!} (x-1)^k = \sum_{k=0}^n (-1)^k (x-1)^k$$

In der Abbildung 10.2 sieht man die TAYLOR-Polynome $P_0(x), P_1(x), P_2(x), P_3(x)$ und $P_4(x)$ und die Funktion $f(x)$ über dem Intervall $(0, 5]$ gezeichnet. Approximiert man z.B. $f(3) = \frac{1}{3}$ durch $P_n(3)$ mit aufsteigendem n , so erhält man

$$P_0(3) = 1, \quad P_1(3) = -1, \quad P_2(3) = 3, \quad P_3(3) = -5, \quad P_4(3) = 11,$$

d.h. dieses Verfahren versagt dramatisch. Daher sind TAYLOR-Polynome zur Polynomapproximation nur dort anwendbar, wo die Approximation an Stellen nahe an x_0 benötigt wird.

Viele Anwendungen verlangen, dass vorgegebene Wertepaare, z.B. Messwerte, in gewisser Weise durch eine Formel beschrieben werden. Mit der Formel möchte man dann in einer Weise weiterarbeiten, die mit den Wertepaaren selbst nicht möglich ist, z.B. Nullstellen, Integrale oder ähnliche Größen berechnen. Fordert man, dass diese Formel die Wertepaare exakt reproduziert, so hat man ein **Interpolationsproblem** vorliegen. Genauer gesagt suchen wir eine Funktion f , deren Graph genau durch die vorgegebenen Wertepaare verläuft.

Definition 74. Gegeben sind $n+1$ Wertepaare $(x_i, f_i), i = 0, \dots, n$, mit $x_i \neq x_j$ für $i \neq j$. Gesucht ist eine stetige Funktion f mit der Eigenschaft $f(x_i) = f_i$ für alle $i = 0, \dots, n$. Man nennt die x -Werte der Wertepaare auch **Stützstellen**, die y -Werte der Wertepaare auch **Stützwerte** und die Wertepaare selbst **Stützpunkte**.

Eine stetige Funktion f mit der Eigenschaft $f(x_i) = f_i$ für $i = 0, \dots, n$ heißt **Interpolierende** der Wertepaare (x_i, f_i) . Man sagt auch f **interpoliert** diese Wertepaare.

Im Folgenden setzen wir stets voraus, dass die x_i paarweise verschieden sind.

10.2 LAGRANGE'sche Polynome

In diesem Abschnitt werden approximierende Polynome diskutiert, die dadurch festgelegt sind, dass sie durch bestimmte vorgegebene Punkte in der Ebene gehen müssen.

1. Fall: Gegeben sind 2 Punkte (x_0, y_0) und (x_1, y_1) , durch die ein Polynom 1. Grades gehen soll. Dem entspricht die Approximation einer Funktion f durch ein Polynom 1. Grades, wobei $f(x_0) = P(x_0)$ und $f(x_1) = P(x_1)$ gelten soll, d.h. beide Funktionen stimmen in den zwei Punkten (x_0, y_0) und (x_1, y_1) überein (siehe Abb. 10.3).

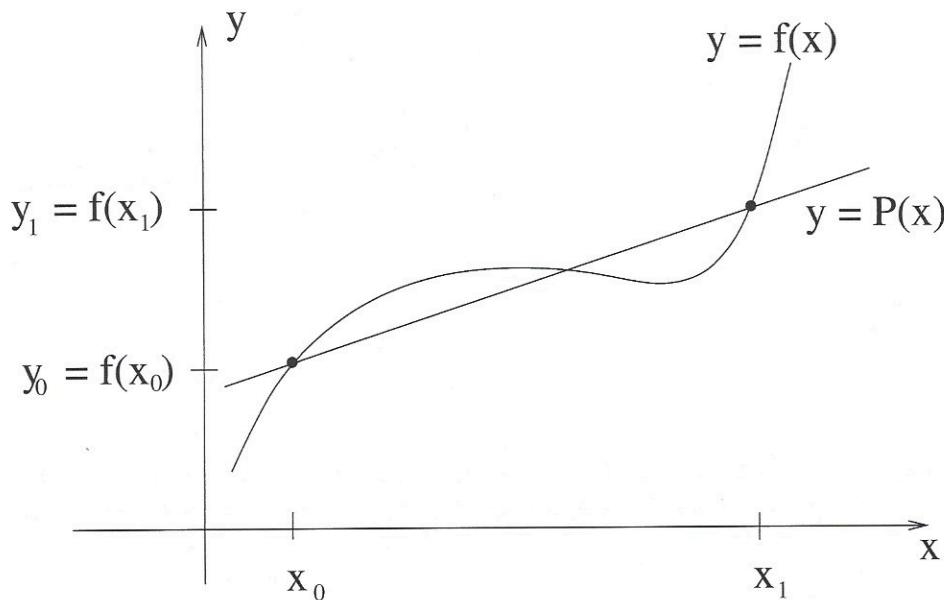


Abbildung 10.3: Interpolation durch eine lineare Funktion

Dazu bilden wir das Polynom

$$P(x) = \frac{x - x_1}{x_0 - x_1} f(x_0) + \frac{x - x_0}{x_1 - x_0} f(x_1). \quad (10.1)$$

Für $x = x_0$ gilt

$$P(x_0) = f(x_0) = y_0,$$

für $x = x_1$ ist

$$P(x_1) = f(x_1) = y_1.$$

D.h. das Polynom P in (10.1) besitzt die geforderten Eigenschaften.

Allgemeiner Fall: Wir betrachten $n+1$ Punkte $(x_0, f(x_0)), (x_1, f(x_1)), \dots, (x_n, f(x_n))$ und suchen ein Polynom n -ten Grades, das durch diese Punkte geht (siehe Abb. 10.4).

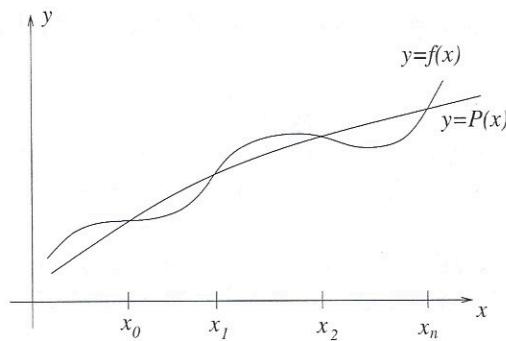


Abbildung 10.4: Interpolation durch ein Polynom

Das lineare Polynom in (10.1), das durch $(x_0, f(x_0))$ und $(x_1, f(x_1))$ geht, verwendet die Faktoren

$$L_0(x) = \frac{x - x_1}{x_0 - x_1} \quad \text{und} \quad L_1(x) = \frac{x - x_0}{x_1 - x_0}.$$

Für sie gilt:

$$L_0(x_0) = 1, \quad L_1(x_0) = 0, \quad L_0(x_1) = 0, \quad L_1(x_1) = 1.$$

Im allgemeinen Fall verwendet man die Faktoren

$$L_{n,k}(x) = \frac{(x - x_0) \dots (x - x_{k-1})(x - x_{k+1}) \dots (x - x_n)}{(x_k - x_0) \dots (x_k - x_{k-1})(x_k - x_{k+1}) \dots (x_k - x_n)}, \quad k = 0, \dots, n. \quad (10.2)$$

Für sie gilt:

$$L_{n,k}(x_k) = 1, \quad L_{n,k}(x_i) = 0 \quad \text{für } i \neq k.$$

Damit kann nun das Interpolationspolynom beschrieben werden, das durch $n + 1$ vorgeschrriebene Punkte $(x_0, f(x_0)), \dots, (x_n, f(x_n))$ geht:

Definition 75. Das n -te LAGRANGE'sche Interpolationspolynom P_n lautet:

$$\begin{aligned} P_n(x) &= f(x_0)L_{n,0}(x) + \dots + f(x_n)L_{n,n}(x) \\ &= \sum_{k=0}^n f(x_k)L_{n,k}(x) \end{aligned}$$

mit $L_{n,k}(x)$ gemäß (10.2).

Bemerkung 48. $P_n(x)$ ist das einzige Polynom vom Grad $\leq n$, das an den Stellen x_0, x_1, \dots, x_n mit der Funktion $f(x)$ übereinstimmt.

Beispiel 136. Gesucht ist das zweite Interpolationspolynom von $f(x) = \frac{1}{x}$, das an den Stellen $x_0 = 2, x_1 = 2.5$ und $x_2 = 4$ mit f übereinstimmt.

$$\begin{aligned} L_{2,0}(x) &= \frac{(x - x_1)(x - x_2)}{(x_0 - x_1)(x_0 - x_2)} = \frac{(x - 2.5)(x - 4)}{(2 - 2.5)(2 - 4)} = (x - 6.5)x + 10 \\ L_{2,1}(x) &= \frac{(x - x_0)(x - x_2)}{(x_1 - x_0)(x_1 - x_2)} = \frac{(x - 2)(x - 4)}{(2.5 - 2)(2.5 - 4)} = \frac{1}{3}[(-4x + 24)x - 32] \\ L_{2,2}(x) &= \frac{(x - x_0)(x - x_1)}{(x_2 - x_0)(x_2 - x_1)} = \frac{(x - 2)(x - 2.5)}{(4 - 2)(4 - 2.5)} = \frac{1}{3}[(x - 4.5)x + 5] \end{aligned}$$

Mit $f(x_0) = f(2) = 0.5, f(x_1) = f(2.5) = 0.4$ und $f(x_2) = f(4) = 0.25$ folgt nun

$$\begin{aligned} P_2(x) &= \sum_{k=0}^2 f(x_k)L_{2,k}(x) = 0.5[(x - 6.5)x + 10] + \frac{0.4}{3}[(-4x + 24)x - 32] + \frac{0.25}{3}[(x - 4.5)x + 5] \\ &= (0.05x - 0.425)x + 1.15 \end{aligned}$$

Damit kann man nun z.B. $f(3)$ durch $P_2(3)$ approximieren:

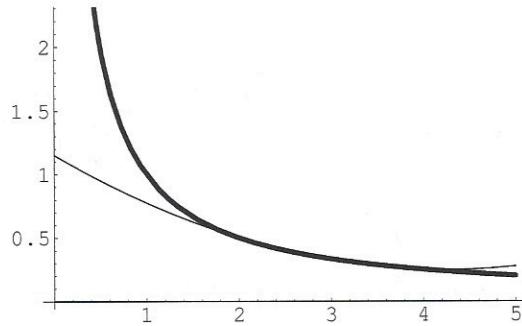
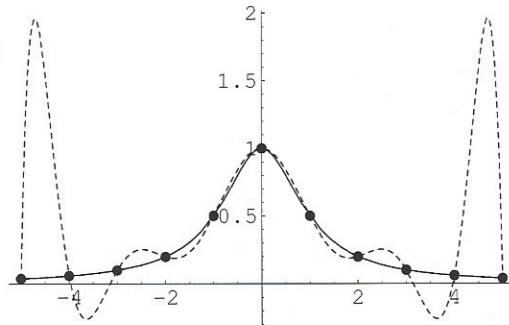
$$f(3) \approx P_2(3) = 0.325$$

Das ist wesentlich besser als die Approximation durch das entsprechende Taylorpolynom (vgl. Bsp. 135).

In Abb. 10.5 sind die Funktion $f(x) = \frac{1}{x}$ und das zugehörige Interpolationspolynom $P_2(x)$ dargestellt.

10.3 Stückweise Interpolation mittels kubischer Splines

Bei der Interpolation durch Polynome kann man durch geschickte Umformungen zwar für den Computer besser geeignete Algorithmen als den hier angegebenen finden, jedoch treten in der Praxis oft numerische Verfahrensfehler auf. Dies soll an folgendem Beispiel demonstriert werden.

Abbildung 10.5: $f(x) = \frac{1}{x}$ Abbildung 10.6: $f(x) = \frac{1}{1+x^2}$

Beispiel 137. Die Funktion

$$f(x) = \frac{1}{1+x^2}$$

soll auf dem Intervall $[-5, 5]$ durch ein Polynom 10. Grades interpoliert werden. Wählt man die 11 Stützstellen x_i äquidistant, d.h. $x_i = -5 + i, i = 0, 1, \dots, 10$, so erhält man die in Abb. 10.6 gezeichneten Graphen. Die durchgezeichnete Linie stellt die Funktion f dar, die strichlierte Kurve das Interpolationspolynom P_{10} .

Das Polynom P_{10} beschreibt also die zu interpolierende Funktion nur im Bereich $-1 \leq x \leq 1$ gut, an den Rändern des Interpolationsintervalls ist der Fehler sehr groß. Die Ursache dafür ist die Neigung von Interpolationspolynomen, zwischen den Stützstellen zu oszillieren. Da diese Oszillationen mit steigendem Grad des Polynoms i.a. zunehmen, ist die Polynominterpolation in vielen Fällen unbrauchbar. Hier soll nun gezeigt werden, dass es meist zielführender ist, eine gegebene Punktmenge stückweise unter Verwendung von Polynomen niedriger Grade zu interpolieren.

Die einfachste stückweise Polynomapproximation ist eine stückweise lineare Interpolation. Sie erhält man, wenn man die Punkte

$$(x_0, f(x_0)), (x_1, f(x_1)), \dots, (x_n, f(x_n))$$

durch eine Folge von Geradenstücken verbindet, wie es in Abb. 10.7 gezeigt wird.

Ein Nachteil dieser linearen Approximation liegt darin, dass die interpolierende Funktion in den Endpunkten der Teilintervalle nicht differenzierbar ist, d.h. nicht glatt ist. Oft ist aber z.B. aus physikalischen Gründen klar, dass Glattheit zu fordern ist, d.h. die interpolierende Funktion muß stetig differenzierbar sein.

Die am weitesten verbreitete stückweise Polynomapproximation verwendet Polynome 3. Grades zwischen den einzelnen Knoten und wird **kubische Spline-Interpolation** genannt. Ein kubisches Polynom $a + bx + cx^2 + dx^3$ enthält vier Konstanten, so daß hier genug Freiraum besteht, um nicht nur die stetige Differenzierbarkeit der interpolierenden Funktion auf dem gesamten Intervall fordern zu können, sondern auch eine stetige zweite Ableitung. (Siehe Abb. 10.8)

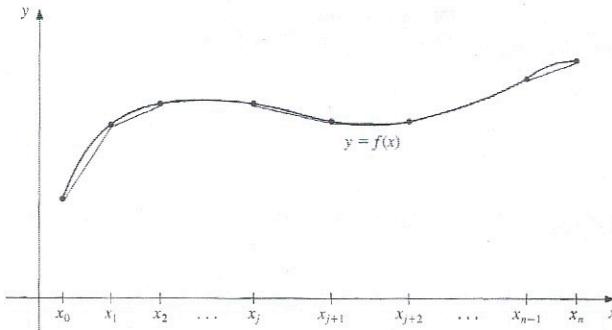


Abbildung 10.7: Interpolation durch lineare Funktionen

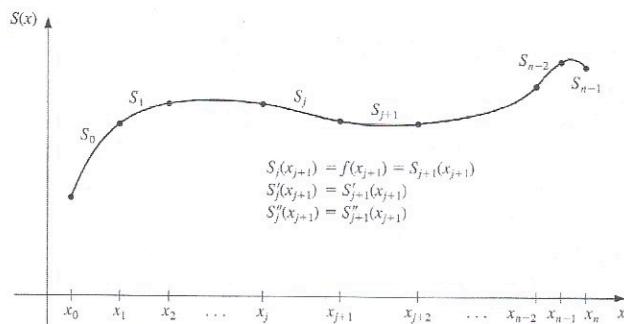


Abbildung 10.8: Interpolation durch kubische Splines

Hinweis: Die interpolierende Funktion besitzt i.a. nicht dieselben Ableitungen wie die zu approximierende Funktion, nicht einmal in den Knoten.

Definition 76. Sei $f \in C[a, b]$ und $\{x_i\}$ eine Menge von Knoten mit $a = x_0 < x_1 < \dots < x_n = b$. Eine **kubische Spline-Interpolierende** S der Funktion f ist eine Funktion mit folgenden Eigenschaften:

- a) Mit den kubischen Polynomen $S_j(x)$, die jeweils auf dem Teilintervall $[x_j, x_{j+1}]$, $j = 0, 1, \dots, n-1$, definiert sind, gilt:

$$S(x) := \begin{cases} S_0(x) & \text{für } x \in [x_0, x_1] \\ S_1(x) & \text{für } x \in [x_1, x_2] \\ \vdots \\ S_{n-1}(x) & \text{für } x \in [x_{n-1}, x_n] \end{cases}$$

- b) $S(x_j) = f(x_j)$ für $j = 0, 1, \dots, n$
c) $S_{j+1}(x_{j+1}) = S_j(x_{j+1})$ für $j = 0, 1, \dots, n-2$
d) $S'_{j+1}(x_{j+1}) = S'_j(x_{j+1})$ für $j = 0, 1, \dots, n-2$
e) $S''_{j+1}(x_{j+1}) = S''_j(x_{j+1})$ für $j = 0, 1, \dots, n-2$
f) $S''(x_0) = S''(x_n) = 0$

Um die kubische Spline-Interpolierende S zu einer gegebenen Funktion f zu bestimmen, werden nun die Bedingungen a) bis f) in der Definition auf die kubischen Polynome

$$S_j(x) = a_j + b_j(x - x_j) + c_j(x - x_j)^2 + d_j(x - x_j)^3, \quad j = 0, 1, \dots, n-1,$$

angewandt. Wegen

$$S_j(x_j) = a_j = f(x_j), \quad j = 0, \dots, n-1 \quad (10.3)$$

folgt aus c) nun

$$\begin{aligned} a_{j+1} &= S_{j+1}(x_{j+1}) = S_j(x_{j+1}) = \\ &= a_j + b_j(x_{j+1} - x_j) + c_j(x_{j+1} - x_j)^2 + d_j(x_{j+1} - x_j)^3, \quad j = 0, 1, \dots, n-2. \end{aligned} \quad (10.4)$$

Zur Abkürzung setzen wir

$$h_j = x_{j+1} - x_j, \quad j = 0, 1, \dots, n-1, \quad (10.5)$$

weiters verwenden wir $a_n := f(x_n)$. Damit geht Gleichung (10.4) über in

$$a_{j+1} = a_j + b_j h_j + c_j h_j^2 + d_j h_j^3, \quad j = 0, 1, \dots, n-1. \quad (10.6)$$

Mit $b_n := S'(x_n)$ und $S'_j(x) = b_j + 2c_j(x - x_j) + 3d_j(x - x_j)^2$ gilt

$$S'_j(x_j) = b_j \quad \text{für } j = 0, 1, \dots, n-1.$$

Aus der Bedingung d) folgt nun

$$b_{j+1} = b_j + 2c_j h_j + 3d_j h_j^2, \quad j = 0, 1, \dots, n-1. \quad (10.7)$$

Mit $c_n := \frac{1}{2}S''(x_n)$ folgt aus der Bedingung e)

$$c_{j+1} = c_j + 3d_j h_j, \quad j = 0, 1, \dots, n-1. \quad (10.8)$$

Löst man Gleichung (10.8) nach d_j auf und setzt dies in die Gleichungen (10.6) und (10.7) ein, so erhält man

$$a_{j+1} = a_j + b_j h_j + \frac{h_j^2}{3}(2c_j + c_{j+1}) \quad (10.9)$$

und

$$b_{j+1} = b_j + h_j(c_j + c_{j+1}). \quad (10.10)$$

Gleichung (10.9) liefert nun

$$b_j = \frac{1}{h_j}(a_{j+1} - a_j) - \frac{h_j}{3}(2c_j + c_{j+1}) \quad (10.11)$$

was nach der Indextransformation $j \mapsto j-1$

$$b_{j-1} = \frac{1}{h_{j-1}}(a_j - a_{j-1}) - \frac{h_{j-1}}{3}(2c_{j-1} + c_j) \quad (10.12)$$

liefert. Jetzt erniedrigt man auch in Glg (10.10) den Index j um 1 und setzt in die so erhaltene Gleichung die Ausdrücke b_j von Gleichung (10.11) und b_{j-1} von Gleichung (10.12) ein. Dies liefert

$$h_{j-1}c_{j-1} + 2(h_{j-1} + h_j)c_j + h_jc_{j+1} = \frac{3}{h_j}(a_{j+1} - a_j) - \frac{3}{h_{j-1}}(a_j - a_{j-1}) \quad (10.13)$$

für $j = 1, 2, \dots, n-1$.

Dieses lineare Gleichungssystem enthält nur mehr die Konstanten c_0, \dots, c_n als Unbestimmte, da die Größen h_0, \dots, h_{n-1} und a_0, \dots, a_n schon bekannt sind (vgl. (10.3) und (10.5)).

Die Randbedingung f) liefert

$$0 = S''(x_0) = 2c_0 \Rightarrow c_0 = 0, \quad 0 = S''(x_n) = 2c_n \Rightarrow c_n = 0.$$

Damit kann nun das Gleichungssystem (10.13) mit der $(n+1) \times (n+1)$ -Matrix

$$A = \begin{pmatrix} 1 & 0 & 0 & \dots & \dots & 0 \\ h_0 & 2(h_0 + h_1) & h_1 & 0 & \dots & 0 \\ 0 & h_1 & 2(h_1 + h_2) & h_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & h_{n-2} & 2(h_{n-2} + h_{n-1}) & h_{n-1} \\ & & & 0 & 0 & 1 \end{pmatrix}$$

und den Vektoren

$$x = \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} 0 \\ \frac{3}{h_1}(a_2 - a_1) - \frac{3}{h_0}(a_1 - a_0) \\ \vdots \\ \frac{3}{h_{n-1}}(a_n - a_{n-1}) - \frac{3}{h_{n-2}}(a_{n-1} - a_{n-2}) \\ 0 \end{pmatrix}$$

angegeben werden in der Form

$$A \cdot x = b. \quad (10.14)$$

Die Matrix A ist streng diagonal dominant. Das Gleichungssystem (10.14) besitzt eine eindeutig bestimmte Lösung c_0, \dots, c_n .

Bei Kenntnis der c_k können nun die übrigen Größen leicht berechnet werden. Man erhält aus (10.11) die Konstanten b_1, \dots, b_{n-1} , und aus (10.8) die Konstanten d_0, \dots, d_{n-1} . Damit kann man jetzt die kubischen Polynome $S_0(x), S_1(x), \dots, S_{n-1}(x)$ konstruieren.

Beispiel 138. Man bestimme die Spline-Interpolierende für die Funktion

$$f(x) = \cos x, \quad x \in [0, 2\pi] \quad \text{mit} \quad n = 6.$$

Mit $x_j = j \cdot \frac{\pi}{3}$, $j = 0, 1, \dots, 6$, erhält man

$$a_j = f(x_j) = \cos \frac{j\pi}{3} = \begin{cases} 1 & \text{für } j = 0 \\ \frac{1}{2} & \text{für } j = 1 \\ -\frac{1}{2} & \text{für } j = 2 \\ -1 & \text{für } j = 3 \\ -\frac{1}{2} & \text{für } j = 4 \\ \frac{1}{2} & \text{für } j = 5 \\ 1 & \text{für } j = 6 \end{cases}$$

und weiters

$$h_j = x_{j+1} - x_j = \frac{\pi}{3}, \quad j = 0, 1, \dots, 5.$$

Das Gleichungssystem (10.14) lautet damit

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{3}\pi & \frac{4}{3}\pi & \frac{1}{3}\pi & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3}\pi & \frac{4}{3}\pi & \frac{1}{3}\pi & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3}\pi & \frac{4}{3}\pi & \frac{1}{3}\pi & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3}\pi & \frac{4}{3}\pi & \frac{1}{3}\pi & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{3}\pi & \frac{4}{3}\pi & \frac{1}{3}\pi \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{9}{2}\pi^{-1} \\ \frac{9}{2}\pi^{-1} \\ 9\pi^{-1} \\ \frac{9}{2}\pi^{-1} \\ -\frac{9}{2}\pi^{-1} \\ 0 \end{pmatrix}$$

Der Lösungsvektor c ist damit gegeben durch

$$c = \begin{pmatrix} 0 \\ -\frac{54}{13}\pi^{-2} \\ \frac{81}{26}\pi^{-2} \\ \frac{135}{26}\pi^{-2} \\ \frac{81}{26}\pi^{-2} \\ -\frac{54}{13}\pi^{-2} \\ 0 \end{pmatrix}$$

Damit erhält man nun die Splines:

$$\begin{aligned} S_0 &= 1 - \frac{27}{26} \frac{x}{\pi} - \frac{54}{13} \frac{x^3}{\pi^3} \\ S_1 &= 1/2 - \frac{63}{26} \frac{x - 1/3\pi}{\pi} - \frac{54}{13} \frac{(x - 1/3\pi)^2}{\pi^2} + \frac{189}{26} \frac{(x - 1/3\pi)^3}{\pi^3} \\ S_2 &= -1/2 - \frac{36}{13} \frac{x - 2/3\pi}{\pi} + \frac{81}{26} \frac{(x - 2/3\pi)^2}{\pi^2} + \frac{27}{13} \frac{(x - 2/3\pi)^3}{\pi^3} \\ S_3 &= -1 + \frac{135}{26} \frac{(x - \pi)^2}{\pi^2} - \frac{27}{13} \frac{(x - \pi)^3}{\pi^3} \\ S_4 &= -1/2 + \frac{36}{13} \frac{x - 4/3\pi}{\pi} + \frac{81}{26} \frac{(x - 4/3\pi)^2}{\pi^2} - \frac{189}{26} \frac{(x - 4/3\pi)^3}{\pi^3} \\ S_5 &= 1/2 + \frac{63}{26} \frac{x - 5/3\pi}{\pi} - \frac{54}{13} \frac{(x - 5/3\pi)^2}{\pi^2} + \frac{54}{13} \frac{(x - 5/3\pi)^3}{\pi^3} \end{aligned}$$

In den Abb. 10.9 und 10.10 sind die lineare und die kubische Spline-Interpolation dargestellt, in Abb. 10.11 der absolute Fehler $|\cos x - S(x)|$, der bei der kubischen Spline-Interpolation auftritt.

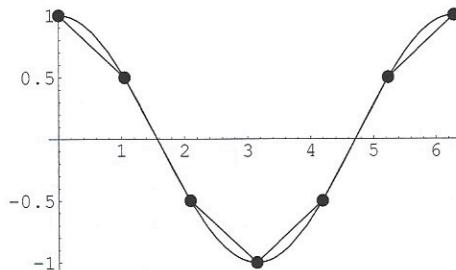


Abbildung 10.9: Lineare Interpolation

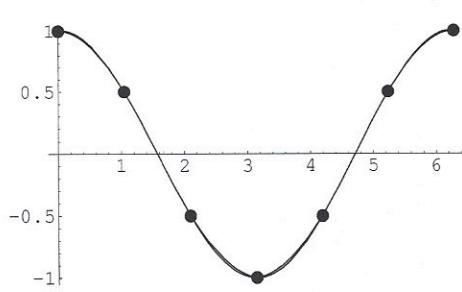


Abbildung 10.10: Kubische Splines

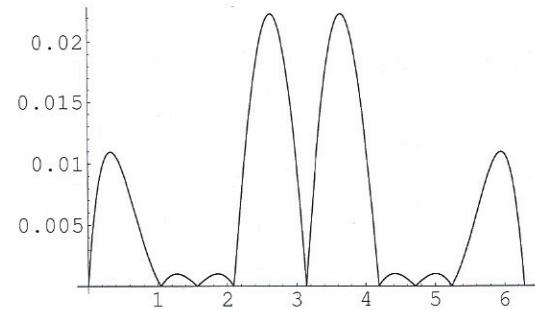


Abbildung 10.11: Fehler $|\cos x - S(x)|$

Kapitel 11

Approximationstheorie

11.1 Methode der kleinsten Quadrate

Hier wollen wir uns mit dem Problem beschäftigen, zu gegebenen Punkten in der xy -Ebene die “beste” Funktion aus einer bestimmten Klasse zu ermitteln, mit der der funktionale Zusammenhang dieser Punkte dargestellt werden kann. Damit möchte man dann die Werte einer Funktion an nichttabellierten Stellen abschätzen.

Beispiel 139. Vorgelegt seien die folgenden Punkte:

x_i	y_i	x_i	y_i
1	1.3	6	8.8
2	3.5	7	10.1
3	4.2	8	12.5
4	5.0	9	13.0
5	7.0	10	15.6

Tabelle 11.1:

Bei der Interpolation (siehe Kapitel 10) sucht man eine Funktion, die genau durch die Punkte (x_i, y_i) , $i = 1, \dots, 10$, geht.

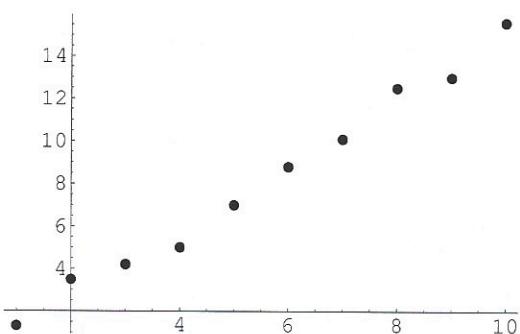


Abbildung 11.1: Die Punkte aus der Tabelle

Hier sucht man nach einer “einfachen” Funktion, die den Zusammenhang zwischen den x_i und den y_i “bestmöglich” beschreibt. Auf Grund der Abbildung 11.1 scheint die Beziehung zwischen x und y linear zu sein. Die Werte in Tabelle 1 sind z.B. aus Messungen entstanden und damit fehlerbehaftet. Wir wollen hier die (in einem gewissen Sinn) “beste” approximierende Gerade angeben, auch wenn sie

nicht in jedem Punkt x_i mit dem Wert y_i exakt zusammenfällt.

Wir wählen also eine Gerade

$$y = ax + b$$

und bezeichnen mit $ax_i + b$ den i -ten Wert auf dieser Näherungsgeraden, sowie mit y_i den i -ten gegebenen y -Wert. Die Konstanten a und b sollen nun so gewählt werden, dass der sogenannte **quadratische Fehler**

$$E(a, b) = \sum_{i=1}^{10} [y_i - (ax_i + b)]^2$$

minimal wird (= Methode der kleinsten Quadrate).

Dieses Minimalproblem zur Bestimmung von a und b soll nun diskutiert werden. Wir betrachten allgemein m verschiedene Punkte $\{(x_1, y_1), \dots, (x_m, y_m)\}$. Notwendig für die Existenz eines lokalen Minimums der Funktion

$$E(a, b) = \sum_{i=1}^m [y_i - (ax_i + b)]^2$$

ist das Verschwinden der ersten beiden partiellen Ableitungen

$$\frac{\partial E}{\partial a} = \frac{\partial}{\partial a} \sum_{i=1}^m [y_i - (ax_i + b)]^2 = 2 \sum_{i=1}^m [y_i - (ax_i + b)] \cdot (-x_i) = 0,$$

$$\frac{\partial E}{\partial b} = \frac{\partial}{\partial b} \sum_{i=1}^m [y_i - (ax_i + b)]^2 = 2 \sum_{i=1}^m [y_i - (ax_i + b)] \cdot (-1) = 0.$$

Daraus erhält man die folgenden beiden Gleichungen in den beiden Unbestimmten a und b :

$$\begin{aligned} a \sum_{i=1}^m x_i^2 + b \sum_{i=1}^m x_i &= \sum_{i=1}^m x_i y_i \\ a \sum_{i=1}^m x_i + b \cdot m &= \sum_{i=1}^m y_i \end{aligned}$$

Die Lösung dieses Problems lautet:

Satz 56. Methode der kleinsten Quadrate

Die lineare Lösung einer gegebenen Wertemenge $(x_i, y_i), i = 1, \dots, m$, nach der Methode der kleinsten Quadrate besitzt die Form

$$y = ax + b$$

wenn

$$\begin{aligned} a &= \frac{m \sum_{i=1}^m x_i y_i - (\sum_{i=1}^m x_i)(\sum_{i=1}^m y_i)}{m \sum_{i=1}^m x_i^2 - (\sum_{i=1}^m x_i)^2} \\ b &= \frac{(\sum_{i=1}^m x_i^2)(\sum_{i=1}^m y_i) - (\sum_{i=1}^m x_i y_i)(\sum_{i=1}^m x_i)}{m \sum_{i=1}^m x_i^2 - (\sum_{i=1}^m x_i)^2} \end{aligned} \tag{11.1}$$

gilt.

Beispiel 140. Wir betrachten die in Tabelle 11.1 aufgelisteten Punkte und erweitern diese Tabelle um zwei weitere Spalten, in denen x_i^2 und $x_i y_i$ angegeben sind. Die Summen der einzelnen Spalten liefern dann die in (11.1) zur Berechnung von a und b benötigten Größen.

x_i	y_i	x_i^2	$x_i y_i$
1	1.3	1	1.3
2	3.5	4	7.0
3	4.2	9	12.6
4	5.0	16	20.0
5	7.0	25	35.0
6	8.8	36	52.8
7	10.1	49	70.7
8	12.5	64	100.0
9	13.0	81	117.0
10	15.6	100	156.0
55	81.0	385	572.4

Damit erhält man

$$\begin{aligned} a &= \frac{10 \cdot 572.4 - 55 \cdot 81}{10 \cdot 385 - 55^2} = 1.538 \\ b &= \frac{385 \cdot 81 - 55 \cdot 572.4}{10 \cdot 385 - 55^2} = -0.360 \end{aligned}$$

In Abbildung 11.2 ist nun die Gerade (= Ausgleichsgerade)

$$y = 1.538x - 0.360$$

zu den Punkten $(x_i, y_i), i = 1, \dots, 10$, eingefügt.

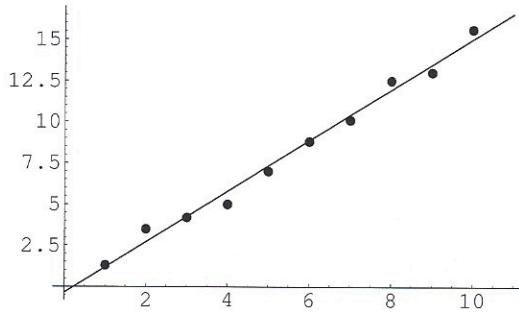


Abbildung 11.2: Ausgleichsgerade

Nun soll das Problem behandelt werden, eine Datenmenge $\{(x_i, y_i) | i = 1, \dots, m\}$ durch ein Polynom

$$P_n(x) = \sum_{k=0}^n a_k x^k \quad (11.2)$$

vom Grad $n < m - 1$ nach der Methode der kleinsten Quadrate zu approximieren. Die Konstanten a_0, a_1, \dots, a_n in (11.2) müssen nun derart gewählt werden, dass der Gesamtfehler, das ist die Summe der Quadrate der Differenzen zwischen den y -Werten auf der approximierenden Kurve und den gegebenen y -Werten, minimiert wird (= **Methode der kleinsten Quadrate**).

Damit

$$E = \sum_{i=1}^m [y_i - P_n(x_i)]^2$$

minimal wird, ist es notwendig

$$\frac{\partial E}{\partial a_j} = 0, \quad j = 0, 1, \dots, n,$$

zu fordern. Dies liefert $n + 1$ lineare Gleichungen für die $n + 1$ Unbestimmten a_0, a_1, \dots, a_n :

$$\begin{aligned} a_0 \sum_{i=1}^m x_i^0 + a_1 \sum_{i=1}^m x_i^1 + a_2 \sum_{i=1}^m x_i^2 + \dots + a_n \sum_{i=1}^m x_i^n &= \sum_{i=1}^m y_i x_i^0, \\ a_0 \sum_{i=1}^m x_i^1 + a_1 \sum_{i=1}^m x_i^2 + a_2 \sum_{i=1}^m x_i^3 + \dots + a_n \sum_{i=1}^m x_i^{n+1} &= \sum_{i=1}^m y_i x_i^1, \\ &\vdots \\ a_0 \sum_{i=1}^m x_i^n + a_1 \sum_{i=1}^m x_i^{n+1} + a_2 \sum_{i=1}^m x_i^{n+2} + \dots + a_n \sum_{i=1}^m x_i^{2n} &= \sum_{i=1}^m y_i x_i^n. \end{aligned}$$

Dieses System besitzt eine eindeutig bestimmte Lösung falls die x_i verschieden voneinander sind.

Beispiel 141. Die Punkte

x_i	y_i
0	1.0000
0.25	1.2840
0.5	1.6487
0.75	2.1170
1.0	2.7183

sollen nach der Methode der kleinsten Quadrate mit dem Polynom

$$P_2(x) = a_0 + a_1 x + a_2 x^2$$

angepaßt werden.

Man erhält das Gleichungssystem

$$\begin{aligned} 5a_0 + 2.5a_1 + 1.875a_2 &= 8.7680 \\ 2.5a_0 + 1.875a_1 + 1.5625a_2 &= 5.4514 \\ 1.875a_0 + 1.5625a_1 + 1.3828a_2 &= 4.4015 \end{aligned}$$

Die Lösung dieses Systems lautet

$$a_0 = 1.0051, \quad a_1 = 0.86468, \quad a_2 = 0.84316.$$

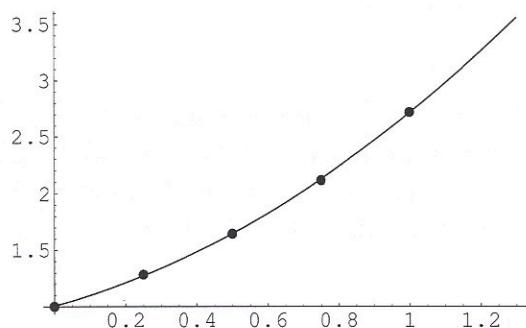


Abbildung 11.3: Ausgleichsparabel

In Abb. 11.3 sind die Punkte und das approximierende Polynom eingezeichnet.

Bemerkung 49. Bei der Suche nach der Gleichung der Geraden von der Form

$$y = a + b x$$

durch die Punkte $P_i(x_i, y_i)$, $i = 1, \dots, n$, hat man folgendes überbestimmtes Gleichungssystem für die Unbestimmten a und b zu "lösen"

$$\begin{aligned} a + x_1 b &= y_1 \\ a + x_2 b &= y_2 \\ a + x_3 b &= y_3 \\ &\vdots \\ a + x_n b &= y_n \end{aligned}$$

d.h. ein System von der Form

$$A \cdot z = c \quad (11.3)$$

mit

$$A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \in M(n \times 2), \quad c = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix}, \quad z = \begin{pmatrix} a \\ b \end{pmatrix}$$

Die mit der zur Matrix A Pseudoinversen $A^\#$ (vgl. Abschnitt 7.9, Def. 70) gewonnene Näherungslösung des Systems (11.3)

$$\bar{z} = A^\# \cdot c$$

stimmt mit der nach der Methode der kleinsten Fehlerquadrate gewonnenen Lösung überein.
Analog ergeben sich die Koeffizienten a_0, a_1, a_2 im Polynom

$$y = a_0 + a_1 x + a_2 x^2$$

das durch die Punkte $P_i(x_i, y_i)$, $i = 1, \dots, n$, gehen soll, als Näherungslösung des überbestimmten Gleichungssystems $A \cdot z = c$ mit

$$A = \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{pmatrix} \in M(n \times 3), \quad c = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad z = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix}$$

unter Verwendung der Pseudoinversen $A^\#$ in der Form

$$\bar{z} = A^\# \cdot c$$

in Übereinstimmung mit dem nach der Methode der kleinsten Fehlerquadrate gewonnenen Resultat.

11.2 Gleichmäßige Approximation nach der Methode der kleinsten Quadrate

Zu einer Funktion $f \in C[a, b]$ soll ein Polynom P_n vom Grad $\leq n$ so bestimmt werden, dass der (quadratische) Fehler

$$\int_a^b (f(x) - P_n(x))^2 dx$$

minimal wird.

Verfährt man hier wie im vorigen Abschnitt, so erhält man zur Bestimmung der Koeffizienten a_k des Polynoms $P_n(x) = \sum_{k=0}^n$ wiederum ein lineares Gleichungssystem von n Gleichungen. Dieses System besitzt jedoch i.a. eine Koeffizientenmatrix, die extrem anfällig auf Rundungsfehler ist. Man sucht daher nach einem Verfahren, das den Rechenaufwand verringert und das auf ein stabiles lineares Gleichungssystem führt.

Dazu betrachten wir eine Menge von Funktionen

$$\{\phi_0, \phi_1, \dots, \phi_n\}$$

im Vektorraum $C[a, b]$ der auf dem Intervall $[a, b]$ stetigen Funktionen, die dort linear unabhängig ist. Ferner seien diese Funktionen bezüglich des Skalarprodukts

$$\langle \varphi, \psi \rangle = \int_a^b \varphi(x) \psi(x) dx$$

orthogonal, d.h. es gilt

$$\int_a^b \phi_j(x) \cdot \phi_k(x) dx = \begin{cases} 0 & \text{für } j \neq k \\ \alpha_k & \text{für } j = k \end{cases}. \quad (11.4)$$

Wir bilden nun eine Linearkombination P dieser orthogonalen Funktionen

$$P(x) = \sum_{k=0}^n a_k \phi_k(x)$$

und versuchen mit dieser Funktion $P(x)$ eine vorgelegte Funktion $f \in C[a, b]$ so zu approximieren, dass der quadratische Fehler

$$E(a_0, \dots, a_n) = \int_a^b [f(x) - \sum_{k=0}^n a_k \phi_k(x)]^2 dx$$

minimal wird. Dies führt auf die Bedingungen

$$\frac{\partial E}{\partial a_j} = 2 \int_a^b [f(x) - \sum_{k=0}^n a_k \phi_k(x)] \phi_j(x) dx = 0, \quad j = 0, 1, \dots, n.$$

Daraus folgt

$$\int_a^b f(x) \phi_j(x) dx = \sum_{k=0}^n a_k \int_a^b \phi_k(x) \phi_j(x) dx, \quad j = 0, 1, \dots, n.$$

Wegen der Orthogonalität der Funktionen ϕ_j (vgl. Glg. (11.4)) gilt aber schließlich

$$\int_a^b f(x) \phi_j(x) dx = a_j \int_a^b \phi_j^2(x) dx = a_j \alpha_j.$$

Daraus können die Koeffizienten a_j jetzt einfach berechnet werden:

$$a_j = \frac{1}{\alpha_j} \int_a^b f(x) \phi_j(x) dx.$$

Es gilt also folgender

Satz 57. Approximation durch orthogonale Funktionen

Ist $\{\phi_0, \dots, \phi_n\}$ eine auf dem Intervall $[a, b]$ orthogonale Menge von Funktionen und $f \in \mathcal{C}[a, b]$. Dann approximiert das "Polynom"

$$P(x) = \sum_{k=0}^n a_k \phi_k(x)$$

die Funktion f nach der Methode der kleinsten Quadrate auf $[a, b]$, wenn gilt

$$a_k = \frac{\int_a^b \phi_k(x) f(x) dx}{\int_a^b \phi_k^2(x) dx}.$$

Bemerkung 50. Eine auf $[-1, 1]$ orthogonale Funktionenmenge liegt z.B. mit den LEGENDRE'schen Polynomen vor, die in Abschnitt 5.4 Bsp. 84 eingeführt worden sind. In vielen Anwendungen werden auch die TSCHEBYSCHEW-Polynome verwendet, die bezüglich des Skalarprodukts

$$\langle \varphi, \psi \rangle = \int_{-1}^1 \frac{1}{\sqrt{1-x^2}} \varphi(x) \psi(x) dx$$

auf $(-1, 1)$ orthogonal sind.

Hier sollen nur noch die **trigonometrischen Polynome** vom Grad $\leq n$ betrachtet werden (siehe auch die FOURIER-Reihen in der Vorlesung Analysis T2):

$$\begin{aligned}\phi_0 &= \frac{1}{2} \\ \phi_k &= \cos kx, \quad k = 1, \dots, n, \\ \phi_{n+k} &= \sin kx, \quad k = 1, \dots, n.\end{aligned}$$

Die Menge $\{\phi_0, \phi_1, \dots, \phi_{2n}\}$ ist orthogonal auf $[-\pi, \pi]$, denn es gilt z.B.

$$\begin{aligned}\int_{-\pi}^{\pi} \frac{1}{2} \cos kx dx &= 0, & \int_{-\pi}^{\pi} \cos^2 kx dx &= \pi, \\ \int_{-\pi}^{\pi} \frac{1}{2} \sin kx dx &= 0, & \int_{-\pi}^{\pi} \sin^2 kx dx &= \pi. \\ \int_{-\pi}^{\pi} \cos kx \sin kx dx &= 0,\end{aligned}$$

Man kann also eine auf $[-\pi, \pi]$ stetige Funktion f durch

$$\begin{aligned}S_n(x) &= \sum_{k=0}^{2n} \tilde{a}_k \phi_n(x) \\ &= \frac{a_0}{2} + \sum_{k=1}^n a_k \cos kx + \sum_{k=1}^n b_k \sin kx\end{aligned}$$

approximieren. Es gilt dann

$$\begin{aligned}a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx, \quad k = 0, 1, \dots, n \\ b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx, \quad k = 1, 2, \dots, n\end{aligned}$$

Das Polynom $S_n(x)$ nennt man das **FOURIER-Polynom** der Funktion $f(x)$, die Koeffizienten a_k und b_k heißen die **FOURIER-Koeffizienten**.

Beispiel 142. $f(x) = |x|, \quad -\pi \leq x \leq \pi$

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} |x| dx = \frac{1}{\pi} \int_{-\pi}^0 (-x) dx + \frac{1}{\pi} \int_0^{\pi} x dx = \pi$$

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} |x| \cos kx dx = \frac{2}{\pi} \int_0^{\pi} x \cos kx dx = \frac{2}{\pi} \frac{1}{k^2} [(-1)^k - 1] \quad \text{für } k = 1, 2, \dots, n.$$

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} |x| \sin kx dx = \dots = 0 \quad \text{für } k = 1, 2, \dots, n.$$

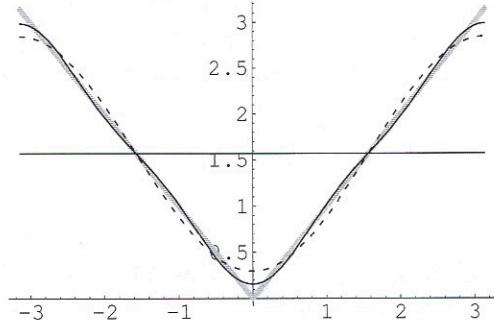


Abbildung 11.4: $f(x) = |x|$ approximiert durch S_0, S_1 und S_2

11.3 Diskrete Approximation – Schnelle FOURIER-Transformation

11.3.1 Approximation durch trigonometrische Polynome

Zunächst betrachten wir das diskrete Analogon zur gleichmäßigen Approximation nach der Methode der kleinsten Quadrate, das uns eine diskrete Approximation von Datenpunkten durch trigonometrische Polynome im Sinne der kleinsten Fehlerquadrate liefert. Dabei wird der Fehler nicht über das gesamte Intervall minimiert, sondern nur in den Interpolationspunkten (= diskrete Approximation). Wir wählen wieder das Intervall $[-\pi, \pi]$, das durch die Punkte

$$x_j = -\pi + \frac{j}{m}\pi, \quad j = 0, 1, \dots, 2m - 1,$$

in $2m$ Teilintervalle unterteilt wird. Dort seien die Stützwerte y_0, \dots, y_{2m-1} vorgelegt. Wir betrachten jetzt die Menge der Funktionen

$$\{\phi_0, \phi_1, \dots, \phi_{2n}\}$$

mit

$$\begin{aligned} \phi_0 &= \frac{1}{2} \\ \phi_n &= \cos kx, \quad k = 1, 2, \dots, n, \\ \phi_{n+k} &= \sin kx, \quad k = 1, 2, \dots, n-1. \end{aligned}$$

Mit ihrer Hilfe wird nun eine Funktion S_n gemäß

$$S_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n-1} (a_k \cos kx + b_k \sin kx) + \frac{a_n}{2} \cos nx \tag{11.5}$$

gebildet mit dem Ziel, die Summe der quadratischen Abweichungen der Funktionswerte $S_n(x_j)$ von den Stützwerten y_j an den Stellen x_j , $j = 0, 1, \dots, 2m - 1$, zu minimieren, d.h. wir fordern

$$E(a_0, \dots, a_n, b_1, \dots, b_{n-1}) = \sum_{j=0}^{2m-1} [y_j - S_n(x_j)]^2 \stackrel{!}{=} \text{Min.}$$

Man erhält dann (analog wie im kontinuierlichen Fall im vorigen Abschnitt) die Koeffizienten zu

$$\begin{aligned} a_k &= \frac{1}{m} \sum_{j=0}^{2m-1} y_j \cos kx_j, \quad k = 0, 1, \dots, n, \\ b_k &= \frac{1}{m} \sum_{j=0}^{2m-1} y_j \sin kx_j, \quad k = 1, 2, \dots, n-1. \end{aligned}$$

Beispiel 143. Man approximiere die Funktion

$$f(x) = x(x^2 - \pi^2)$$

auf dem Intervall $[-\pi, \pi]$ durch $S_4(x)$, wobei $m = 4$ sei.

$$S_4(x) = \frac{a_0}{2} + \sum_{k=1}^3 (a_k \cos kx + b_k \sin kx) + \frac{a_4}{2} \cos 4x$$

mit

$$\begin{aligned} a_k &= \frac{1}{4} \sum_{j=0}^7 y_j \cos kx_j, \quad k = 0, 1, \dots, 4, \\ b_k &= \frac{1}{4} \sum_{j=0}^7 y_j \sin kx_j, \quad k = 1, 2, 3. \end{aligned}$$

Mit $x_j = -\pi + \frac{j}{4}\pi$ erhält man

$$\begin{array}{ll} y_0 = f(x_0) = 0 & y_4 = f(x_4) = 0 \\ y_1 = f(x_1) = \frac{21\pi^3}{64} & y_5 = f(x_5) = -\frac{15\pi^3}{64} \\ y_2 = f(x_2) = \frac{3\pi^3}{8} & y_6 = f(x_6) = -\frac{3\pi^3}{8} \\ y_3 = f(x_3) = \frac{15\pi^3}{64} & y_7 = f(x_7) = -\frac{21\pi^3}{64} \end{array}$$

Die Koeffizienten a_k und b_k lauten nun

$$\begin{array}{lll} a_0 = 0 & a_3 = 0 & b_1 = -\frac{3\pi^3}{16} \left(1 + \frac{3}{2\sqrt{2}}\right) \approx -11.98 \\ a_1 = 0 & a_4 = 0 & b_2 = \frac{3\pi^3}{64} \approx 1.45 \\ a_2 = 0 & & b_3 = \frac{3\pi^3}{16} \left(1 - \frac{3}{2\sqrt{2}}\right) \approx -0.35 \end{array}$$

Das approximierende trigonometrische Polynom lautet somit

$$S_4(x) = -11.98 \sin x + 1.45 \sin 2x - 0.35 \sin 3x$$

In Abb. 11.5 sind die Funktion $f(x)$ und das Polynom $S_4(x)$ mit durchzogenen Linien dargestellt, der Fehler $f(x) - S_4(x)$ als strichlierte Linie.

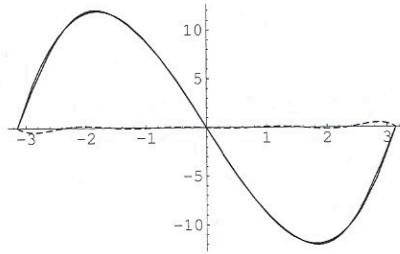


Abbildung 11.5: $f(x) = x(x^2 - \pi^2)$, $S_4(x)$ und Fehler

Beispiel 144. Man approximiere die Funktion

$$f(x) = x^4 - 3x^3 + 2x^2 - \tan[x(x-2)]$$

auf dem Intervall $[0, 2]$ durch $S_4(x)$, wobei $m = 4$ sei.

Wir setzen $x_j = j/4$, $j = 0, 1, \dots, 8$. Um obige Formeln anwenden zu können, muß das Intervall $[0, 2]$ auf das Intervall $[-\pi, \pi]$ transformiert werden. Die dazu benötigte Lineartransformation besitzt die Form

$$z_j = \pi(x_j - 1), \quad j = 0, \dots, 8,$$

und die neuen Punkte sind $(z_j, f(1 + \frac{z_j}{\pi}))$, $j = 0, \dots, 8$.

Das trigonometrische Polynom S_4 ist nun gegeben durch

$$S_4(z) = \frac{a_0}{2} + \sum_{k=1}^3 (a_k \cos kz + b_k \sin kz) + \frac{a_4}{2} \cos 4z$$

mit

$$\begin{aligned} a_k &= \frac{1}{4} \sum_{j=0}^7 f\left(1 + \frac{z_j}{\pi}\right) \cos kz_j, \quad k = 0, 1, 2, 3, 4, \\ b_k &= \frac{1}{4} \sum_{j=0}^7 f\left(1 + \frac{z_j}{\pi}\right) \sin kz_j, \quad k = 1, 2, 3. \end{aligned}$$

Damit erhält man

$$\begin{aligned} S_4(z) &= 0.761979 + 0.771841 \cos z + 0.0173037 \cos 2z + 0.00686304 \cos 3z \\ &\quad - 0.000578545 \cos 4z - 0.386374 \sin z + 0.0468750 \sin 2z - 0.0113738 \sin 3z \end{aligned}$$

Das trigonometrische Polynom $\tilde{S}_4(x)$ auf $[0, 2]$ erhält man jetzt dadurch, dass man $z = \pi(x-1)$ setzt.
Die Graphen $y = f(x)$ und $y = \tilde{S}_4(x)$ können in Abb. 11.6 verglichen werden.

In Abb. 11.7 ist der Fehler $f(x) - \tilde{S}_4(x)$ aufgetragen.

11.3.2 Interpolation durch trigonometrische Polynome - FFT

Die Interpolation von großen, äquidistant eingeteilten Datenmengen durch trigonometrische Polynome kann sehr genaue Ergebnisse bringen. Sie stellt das geeignete Approximationsverfahren auf den Gebieten der digitalen Filter, Antennenstrahlungsdiagrammen, Quantenmechanik, Optik und bei vielen Simulationsproblemen dar.

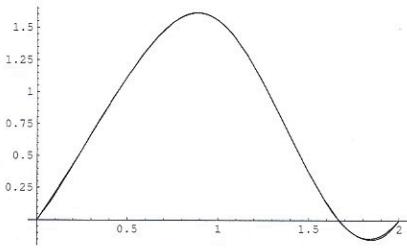


Abbildung 11.6: $f(x) = x^4 - 3x^3 + 2x^2 - \tan[x(x-2)]$
und $\tilde{S}_4(x)$

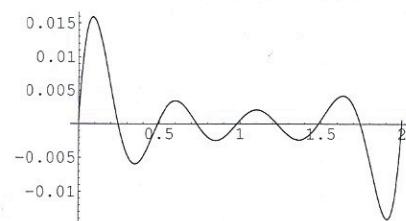


Abbildung 11.7: Fehler $f(x) - \tilde{S}_4(x)$

Man gewinnt das trigonometrische Interpolationspolynom $S_m(x)$ durch $2m$ Datenpunkte $\{(x_0, y_0), \dots, (x_{2m-1}, y_{2m-1})\}$ aus dem Polynom S_n gemäß (11.5) z.B. dadurch, dass man $n = m$ setzt und die Konstante a_m durch $a_m/2$ ersetzt. Man hat also für das trigonometrische Interpolationspolynom den Ausdruck

$$S_m(x) = \frac{a_0}{2} + \frac{a_m}{2} \cos mx + \sum_{k=1}^{m-1} (a_k \cos kx + b_k \sin kx) \quad (11.6)$$

Die Interpolation von n Datenpunkten erfordert ungefähr n^2 Multiplikationen und n^2 Additionen bei der direkten Berechnung. Die Approximation von vielen tausenden Datenpunkten ist in diesem Zusammenhang nicht unüblich, sodass die direkten Verfahren zur Berechnung der Konstanten Millionen von Operationen erfordern, was immense Rechenzeiten und auch einen unerwünschten Anstieg der Rundungsfehler bedeutet.

Man verwendet daher einen ausgeklügelten Algorithmus, der die Anzahl der erforderlichen Rechenschritte auf ungefähr $n \cdot \log_2 n$ reduziert. Dieses Verfahren wird als die **schnelle FOURIER-Transformation (Fast FOURIER-Transform = FFT)** bezeichnet. Es soll im folgenden beschrieben werden.

Dazu betrachten wir zunächst die Funktion

$$F(x) = \frac{1}{m} \sum_{k=0}^{2m-1} c_k e^{ikx} \quad (11.7)$$

mit

$$c_k = \sum_{j=0}^{2m-1} y_j e^{i \frac{kj}{m} \pi} \quad \text{für } k = 0, \dots, 2m-1, \quad (11.8)$$

wobei die im allgemeinen komplexen Konstanten c_k von der Form $\alpha + i\beta$, $\alpha, \beta \in \mathbb{R}$, $i^2 = -1$ sind.

Behauptung. Zwischen den reellen Koeffizienten a_k und b_k in (11.6) und den komplexen Koeffizienten c_k in (11.8) besteht folgender Zusammenhang:

$$a_k + ib_k = \frac{1}{m} c_k (-1)^k, \quad k = 0, \dots, m \quad (11.9)$$

Beweis. Für alle $k = 0, 1, \dots, m$, gilt

$$\frac{1}{m} c_k (-1)^k = \frac{1}{m} c_k e^{-i\pi k} = \frac{1}{m} \sum_{j=0}^{2m-1} y_j e^{i \frac{kj}{m} \pi} e^{-i\pi k} = \frac{1}{m} \sum_{j=0}^{2m-1} y_j e^{ik(\pi \frac{j}{m} - \pi)}.$$

Mit $\frac{j}{m} = x_j$ und $\pi(x_j - 1) = z_j$ und der (aus der Vorlesung Analysis T1) bekannten Beziehung

$e^{i\varphi} = \cos \varphi + i \sin \varphi$ folgt nun

$$\begin{aligned}\frac{1}{m} c_k (-1)^k &= \frac{1}{m} \sum_{j=0}^{2m-1} y_j e^{ikz_j} \\ &= \frac{1}{m} \sum_{j=0}^{2m-1} y_j (\cos kz_j + i \sin kz_j) \\ &= a_k + i b_k\end{aligned}$$

D.h. man berechnet sich zunächst die Konstanten c_k aus der Beziehung (11.8) und kann sich dann die Koeffizienten a_k und b_k über Gleichung (11.9) bestimmen.

Die Verringerung der Anzahl der Rechenoperationen bei der FFT folgt aus der Berechnung der Koeffizienten in Gruppen. Dies soll an Hand des folgenden Beispiels demonstriert werden.

Beispiel 145. Wir betrachten $2m = 2^3 = 8$ Datenpunkte $(x_j, y_j), j = 0, \dots, 7$ wobei $x_j = -\pi + \frac{j}{4}\pi$ ist. Dann gilt

$$S_4(x) = \frac{a_0}{2} + \sum_{k=1}^3 (a_k \cos kx + b_k \sin kx) + \frac{a_4}{2} \cos 4x$$

mit

$$a_k = \frac{1}{4} \sum_{j=0}^7 y_j \cos kx_j, \quad b_k = \frac{1}{4} \sum_{j=0}^7 y_j \sin kx_j \quad \text{für } k = 0, 1, \dots, 4.$$

Es ist

$$F(x) = \frac{1}{4} \sum_{j=0}^7 c_k e^{ikx}$$

mit

$$c_k = \sum_{j=0}^7 y_j e^{i \frac{kj}{4}\pi}, \quad k = 0, 1, \dots, 7.$$

Beim direkten Berechnen sind die komplexen Konstanten c_k gegeben durch

$$\begin{aligned}c_0 &= y_0 + y_1 + y_2 + y_3 + y_4 + y_5 + y_6 + y_7 \\ c_1 &= y_0 + \frac{i+1}{\sqrt{2}}y_1 + iy_2 + \frac{i-1}{\sqrt{2}}y_3 - y_4 - \frac{i+1}{\sqrt{2}}y_5 - iy_6 - \frac{i-1}{\sqrt{2}}y_7 \\ c_2 &= y_0 + iy_1 - y_2 - iy_3 + y_4 + iy_5 - y_6 - iy_7 \\ c_3 &= y_0 + \frac{i-1}{\sqrt{2}}y_1 - iy_2 + \frac{i+1}{\sqrt{2}}y_3 - y_4 - \frac{i-1}{\sqrt{2}}y_5 + iy_6 - \frac{i+1}{\sqrt{2}}y_7 \\ c_4 &= y_0 - y_1 + y_2 - y_3 + y_4 - y_5 + y_6 - y_7 \\ c_5 &= y_0 - \frac{i+1}{\sqrt{2}}y_1 + iy_2 - \frac{i-1}{\sqrt{2}}y_3 - y_4 - \frac{i+1}{\sqrt{2}}y_5 - iy_6 + \frac{i-1}{\sqrt{2}}y_7 \\ c_6 &= y_0 - iy_1 - y_2 + iy_3 + y_4 - iy_5 - y_6 + iy_7 \\ c_7 &= y_0 - \frac{i-1}{\sqrt{2}}y_1 - iy_2 - \frac{i+1}{\sqrt{2}}y_3 - y_4 + \frac{i-1}{\sqrt{2}}y_5 + iy_6 + \frac{i+1}{\sqrt{2}}y_7\end{aligned}$$

Aufgrund der kleinen Datenmenge sind viele Koeffizienten von y_i in diesen Gleichungen gleich 1 oder -1 . In einer größeren Anwendung wird dies weniger häufig der Fall sein, so dass, um die Zahl der Rechenoperationen genau zu bestimmen, auch die Multiplikation mit 1 oder -1 mitgezählt wird, obwohl es in diesem Beispiel nicht nötig wäre. Somit sind 64 Multiplikationen/Divisionen und 56 Additionen/Subtraktionen zum direkten Berechnen von c_0, c_1, \dots, c_7 notwendig.

Um die schnelle FOURIER-Transformation anzuwenden, sei zuerst

$$\begin{aligned}
 d_0 &= \frac{1}{2}(c_0 + c_4) = y_0 + y_2 + y_4 + y_6 \\
 d_1 &= \frac{1}{2}(c_0 - c_4) = y_1 + y_3 + y_5 + y_7 \\
 d_2 &= \frac{1}{2}(c_1 + c_5) = y_0 + iy_2 - y_4 - iy_6 \\
 d_3 &= \frac{1}{2}(c_1 - c_5) = \frac{i+1}{\sqrt{2}}(y_1 + iy_3 - y_5 - iy_7) \\
 d_4 &= \frac{1}{2}(c_2 + c_6) = y_0 - y_2 + y_4 - y_6 \\
 d_5 &= \frac{1}{2}(c_2 - c_6) = i(y_1 - y_3 + y_5 - y_7) \\
 d_6 &= \frac{1}{2}(c_3 + c_7) = y_0 - iy_2 - y_4 + iy_6 \\
 d_7 &= \frac{1}{2}(c_3 - c_7) = \frac{i-1}{\sqrt{2}}(y_1 - iy_3 - y_5 + iy_7)
 \end{aligned}$$

und danach

$$\begin{aligned}
 e_0 &= \frac{1}{2}(d_0 + d_4) = y_0 + y_4 \\
 e_1 &= \frac{1}{2}(d_0 - d_4) = y_2 + y_6 \\
 e_2 &= \frac{1}{2}(id_1 + d_5) = i(y_1 + y_5) \\
 e_3 &= \frac{1}{2}(id_1 - d_5) = i(y_3 + y_7) \\
 e_4 &= \frac{1}{2}(d_2 + d_6) = y_0 - y_4 \\
 e_5 &= \frac{1}{2}(d_2 - d_6) = i(y_2 - y_6) \\
 e_6 &= \frac{1}{2}(id_3 + d_7) = \frac{i-1}{\sqrt{2}}(y_1 - y_5) \\
 e_7 &= \frac{1}{2}(id_3 - d_7) = \frac{i-1}{\sqrt{2}}(y_3 - y_7)
 \end{aligned}$$

und letztlich

$$\begin{aligned}
 f_0 &= \frac{1}{2}(e_0 + e_4) = y_0 \\
 f_1 &= \frac{1}{2}(e_0 - e_4) = y_4 \\
 f_2 &= \frac{1}{2}(ie_1 + e_5) = iy_2 \\
 f_3 &= \frac{1}{2}(ie_1 - e_5) = iy_6 \\
 f_4 &= \frac{1}{2}(\frac{i+1}{\sqrt{2}}e_2 + e_6) = \frac{i-1}{\sqrt{2}}y_1 \\
 f_5 &= \frac{1}{2}(\frac{i+1}{\sqrt{2}}e_2 - e_6) = \frac{i-1}{\sqrt{2}}y_5 \\
 f_6 &= \frac{1}{2}(\frac{i-1}{\sqrt{2}}e_3 + e_7) = -\frac{i+1}{\sqrt{2}}y_3 \\
 f_7 &= \frac{1}{2}(\frac{i-1}{\sqrt{2}}e_3 - e_7) = -\frac{i+1}{\sqrt{2}}y_7
 \end{aligned}$$

definiert. Die Gestalt der Konstanten $c_0, \dots, c_7; d_0, \dots, d_7; e_0, \dots, e_7$ und f_0, \dots, f_7 ist von den speziellen Datenpunkten unabhängig. Sie hängen nur von $m = 4$ ab. Für jedes $m \in \mathbb{N}$ existiert eine eindeutige Menge von Konstanten $\{c_k\}, \{d_k\}, \{e_k\}$ und $\{f_k\}$, $k = 0, \dots, 2m-1$. Dieser Teil der Arbeit wird für eine spezielle Anwendung nicht benötigt, lediglich die folgenden Beziehungen.

1. $f_0 = y_0; f_1 = y_4; f_2 = iy_2; f_3 = iy_6;$
 $f_4 = \frac{i-1}{\sqrt{2}}y_1; f_5 = \frac{i-1}{\sqrt{2}}y_5; f_6 = -\frac{i+1}{\sqrt{2}}y_3; f_7 = -\frac{i+1}{\sqrt{2}}y_7.$
2. $e_0 = f_0 + f_1; e_1 = -i(f_2 + f_3); e_2 = \frac{-i+1}{\sqrt{2}}(f_4 + f_5);$
 $e_3 = \frac{-i-1}{\sqrt{2}}(f_6 + f_7); e_4 = f_0 - f_1; e_5 = f_2 - f_3; e_6 = f_4 - f_5; e_7 = f_6 - f_7.$
3. $d_0 = e_0 + e_1; d_1 = -i(e_2 + e_3); d_2 = e_4 + e_5; d_3 = -i(e_6 + e_7);$
 $d_4 = e_0 - e_1; d_5 = e_2 - e_3; d_6 = e_4 - e_5; d_7 = e_6 - e_7;$
4. $c_0 = d_0 + d_1; c_1 = d_2 + d_3; c_2 = d_4 + d_5; c_3 = d_6 + d_7;$
 $c_4 = d_0 - d_1; c_5 = d_2 - d_3; c_6 = d_4 - d_5; c_7 = d_6 - d_7.$

Berechnet man die Konstanten c_0, c_1, \dots, c_7 auf diese Art und Weise, so wird die im folgenden gezeigte Anzahl von Operationen benötigt:

Schritt	Multiplikationen/Divisionen	Additionen/Subtraktionen
1	8	0
2	8	8
3	8	8
4	0	8
Gesamt	24	24

Man beachte, dass wieder die Multiplikationen mit 1 oder -1 mitgezählt wurden, obwohl dies hier keinen Rechenaufwand erfordert. Das Fehlen von Multiplikationen/Divisionen in Schritt 4 folgt aus der Tatsache, dass für jedes m die Koeffizienten c_k aus den d_k in gleicher Weise berechnet werden:

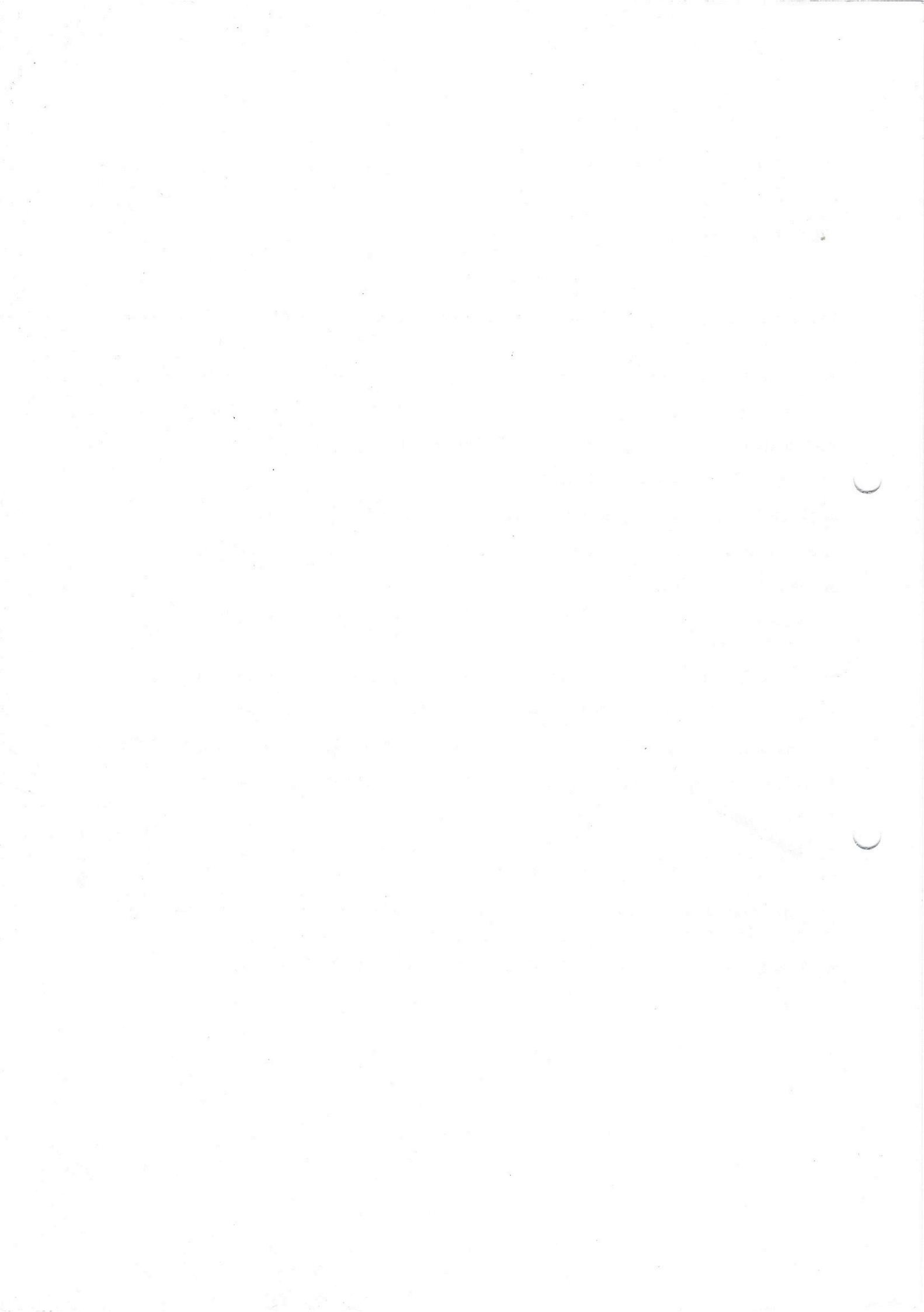
$$c_k = d_{2k} + d_{2k+1}, \quad c_{k+m} = d_{2k} - d_{2k+1}, \quad \text{für } k = 0, 1, \dots, m-1.$$

Insgesamt erfordert die direkte Berechnung der Koeffizienten c_0, \dots, c_7 64 Multiplikationen/Divisionen und 56 Additionen/Subtraktionen. Die schnelle FOURIER-Transformation reduziert die Berechnungen auf 24 Multiplikationen/Divisionen und 24 Additionen/Subtraktionen.

Im Allgemeinen kann der Rechenaufwand bei n Datenpunkten von n^2 (bei direkter Berechnung) auf $n \log_2 n$ Rechenschritte reduziert werden. Dies bedeutet z.B. bei $n = 2^{10} = 1024$ anstelle von 1 048 576 Rechenoperationen lediglich 10 240 Schritte.

Literatur

- M. Anthony, M. Harvey** Linear Algebra: Concepts and Methods, Cambridge University Press, Cambridge 2012
- R. Bronson, G.B. Costa, J.T. Saccoman** Linear Algebra, Academic Press, Oxford 2014
- W. Cheney, D. Kincaid** Numerical Mathematics and Computing, Thomson, Belmont 2004
- J.D. Faires, R.L. Burden** Numerische Methoden, Spektrum, Heidelberg 1994
- G. Fischer** Lineare Algebra, Vieweg Studium, Braunschweig 1995
- D.H. Griffel** Linear Algebra and Its Applications, Ellis Horwood Ltd., Chichester 1989
- S. Grossmann** Elementary Linear Algebra, Saunders College Publ., Philadelphia 1991
- K. Jänich** Lineare Algebra, Springer, Berlin 1993
- A. Jeffrey** Linear Algebra and Ordinary Differential Equations, Blackwell Scientific Publ., Boston 1990
- M. Knorrenschild** Numerische Mathematik, Fachbuchverlag Leipzig, München 2003
- R. Larson, D.C. Falvo** Elementary Linear Algebra, Brooks/Cole, Belmont 2010
- S. Lipschutz** Linear Algebra, Schaum's Outline Series, McGraw-Hill Book Co., New York 1974
- B. Noble, J. Daniel** Applied Linear Algebra, Prentice-Hall, Engelwood Cliffs 1988
- D. Poole** Linear Algebra, Books/Cole, Pacific Grove 2003
- W. Preuß, G. Wenisch** Lehr- und Übungsbuch Mathematik für Informatiker, Fachbuchverlag Leipzig, München 1997
- M. Scherfner, T. Senkbeil** Lineare Algebra für das erste Semester, Pearson Studium, München 2006
- G. Strang** Linear Algebra and Its Applications, Academic Press, New York 1984



Übungsbeispiele

zur Lehreranstaltung

Numerisches Rechnen und Lineare Algebra

Ao.Univ.-Prof. Dr. P. Berglez

1 Lineare Gleichungssysteme – Matrizen

1. Gegeben seien die Matrizen

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{und} \quad B = (b_{ij}) \in M(6 \times 6; \mathbb{R}) \quad \text{mit} \quad b_{ij} = \begin{cases} a_i & \text{für } i = j \\ 0 & \text{für } i \neq j \end{cases}$$

wobei $a_i \in \mathbb{R}$, $a_i \neq 0$, $0 \leq i \leq 6$. Man berechne A^3 und B^2 .

2. Man bestimme den Rang sowie die Summe und das Produkt der folgenden Matrizen (soweit möglich):

$$(a) A = \begin{pmatrix} 2 & 3 & 4 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 3 & 4 & 5 \end{pmatrix}$$

$$(b) A = \begin{pmatrix} 2 & 1 \\ 1 & 3 \\ 4 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

3. Man bestimme die allgemeine Lösung der folgenden Gleichungssysteme

$$\begin{aligned} (a) \quad & x_1 + x_2 + x_3 = 0 \\ & x_1 + 2x_2 - x_3 + x_4 = 0 \\ & x_1 - x_2 + 6x_3 = 0 \\ & x_1 + 3x_3 = 0 \end{aligned}$$

$$\begin{aligned} (b) \quad & x_1 + x_2 = 0 \\ & x_2 + x_3 = 0 \\ & x_3 + x_4 = 0 \\ & x_4 + x_5 = 0 \\ & x_1 - x_5 = 0 \end{aligned}$$

$$(c) \quad A \cdot x = 0 \quad \text{mit} \quad A = \begin{pmatrix} 1 & 3 & 2 & 4 \\ 0 & -1 & 1 & 2 \\ 1 & 5 & 0 & 0 \end{pmatrix}$$

$$(d) \quad A \cdot x = 0 \quad \text{mit} \quad A = \begin{pmatrix} 1 & 1 & -1 & 1 \\ 2 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix}$$

4. Welches der folgenden linearen Gleichungssysteme besitzt eine nichttriviale Lösung?

$$\begin{array}{ll} (a) \quad x + 2y + 3z = 0 & (b) \quad 2x + y - z = 0 \\ 2y + 2z = 0 & x - 2y - 3z = 0 \\ x + 2y + 3z = 0 & -3x - y + 2z = 0 \end{array}$$

5. Man bestimme die allgemeine Lösung der folgenden Gleichungssysteme

$$(a) \begin{array}{l} x_1 + x_2 + x_3 + x_4 = 2 \\ 2x_1 - x_2 + x_3 - x_4 = 12 \\ 3x_1 + 2x_3 = 14 \\ x_1 - 2x_2 - 2x_4 = 10 \end{array}$$

$$(b) \begin{array}{l} x_1 - x_2 + 2x_3 = -3 \\ 2x_1 + 4x_2 + x_3 = 6 \\ x_1 + 5x_2 - x_3 = 9 \end{array}$$

$$(c) A \cdot x = b \text{ mit } A = \begin{pmatrix} 1 & -1 & 2 & -2 \\ -1 & -4 & 8 & -3 \\ 1 & 1 & -2 & 0 \\ 3 & 2 & -4 & -1 \end{pmatrix}, \quad b = \begin{pmatrix} 3 \\ 2 \\ 1 \\ 4 \end{pmatrix}$$

6. Man bestimme alle Lösungen des linearen Systems

$$\begin{array}{l} x + y - z + w = 3 \\ 2x - y - z + 2w = 4 \\ -3y + z = -2 \\ -3x + 3y + z - 3w = -5 \end{array}$$

7. Man löse die folgenden linearen Gleichungssysteme mit dem Gaußschen Eliminationsverfahren:

$$(a) \begin{array}{l} x + y + z = 1 \\ x - y + z = 2 \\ x + y - z = 3 \end{array} \quad (b) \begin{array}{l} x + 2y + z = 5 \\ -x + 2y + 8z = -1 \\ x + 6y + 10z = 9 \end{array}$$

$$(c) \begin{array}{l} x - 3y + 8z = 3 \\ 2x + 2y + z = 0 \end{array}$$

8. Man löse die folgenden linearen Gleichungssysteme mit dem Gaußschen Eliminationsverfahren:

$$(a) \begin{array}{l} x + y + z = 1 \\ x - y - z = 0 \\ 2x + 2y = -1 \\ x + z = 5 \\ x - y + 4z = 3 \end{array} \quad (b) \begin{array}{l} x_1 + x_2 + x_3 + x_4 + x_5 = 5 \\ x_1 - x_2 + 5x_3 - x_4 - 2x_5 = 1 \\ 2x_1 + x_3 - x_5 = 0 \\ x_2 = 5 \end{array}$$

$$(c) \begin{array}{l} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 - x_2 - x_3 = 0 \\ -x_1 + x_2 + x_3 + 2x_4 = 0 \\ x_2 + x_4 = 0 \end{array}$$

9. Man bestimme alle $a \in \mathbb{R}$, für die das System

$$\begin{array}{rcl} x + y - z & = & 3 \\ x - y + 3z & = & 4 \\ x + y + (a^2 - 10)z & = & a \end{array}$$

- (a) keine Lösung,
 - (b) eine eindeutig bestimmte Lösung,
 - (c) beliebig viele Lösungen besitzt.
10. Für welche $b \in \mathbb{R}^3$ besitzt das inhomogene lineare Gleichungssystem $A \cdot x = b$ mit

$$A = \begin{pmatrix} 1 & 1 & -2 \\ 2 & -1 & -1 \\ 4 & 1 & -5 \end{pmatrix}$$

eine Lösung?

11. Man bestimme alle $a \in \mathbb{R}$, für die die folgenden Systeme lösbar sind und gebe in diesem Fall die Lösung an.

$$\begin{array}{ll} \text{(a)} \quad x + 2y + z = a^2 & \text{(b)} \quad x + 2y + z = a^2 \\ x + y + 3z = a & x + y + 3z = a \\ 3x + 4y + 7z = 8 & 3x + 4y + 8z = 8 \end{array}$$

12. Für welche $t \in \mathbb{R}$ ist $A \cdot x = b$ mit $A = \begin{pmatrix} 1 & 3 & 2 \\ -1 & -2 & 2 \\ 2 & 5 & t \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, lösbar?

13. Gegeben ist $A \cdot x = b$ mit $A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 5 & -1 \\ 5 & 7 & 3 \end{pmatrix}$, $b \in \mathbb{R}^3$.

- (a) Für welche $b \in \mathbb{R}^3$ existiert eine Lösung?
 - (b) Man bestimme die Lösung in Abhängigkeit von b .
14. Man bestimme alle $\alpha \in \mathbb{R}$, für die das Gleichungssystem $A \cdot x = b$ mit

$$A = \begin{pmatrix} 1 & -3 & 2 \\ 2 & \alpha & 0 \\ -1 & 6 & 4 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix},$$

- (a) eindeutig lösbar,
- (b) lösbar ist.

15. Gegeben ist $A = \begin{pmatrix} 2-t & 3 & -6 \\ 3 & 2-t & -6 \\ -6 & -6 & 11-t \end{pmatrix}$, $b = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$, $t \in \mathbb{R}$.

- (a) Für welche Werte von $t \in \mathbb{R}$ besitzt $A \cdot x = 0$ eine nichttriviale Lösung?
- (b) Für welche $t \in \mathbb{R}$ ist $A \cdot x = b$ lösbar?
- (c) Für $t = -1$ bestimme man die Lösung von $A \cdot x = b$.

16. Gegeben ist das Gleichungssystem

$$\begin{aligned} 2x_1 + x_2 + x_3 &= 0, \\ -2\lambda x_1 + \lambda x_2 + 9x_3 &= 6, \\ 2x_1 + 2x_2 + \lambda x_3 &= 1. \end{aligned}$$

- (a) Für welche $\lambda \in \mathbb{R}$ ist das Gleichungssystem eindeutig lösbar?
- (b) Für welche $\lambda \in \mathbb{R}$ existieren beliebig viele Lösungen?
- (c) Für welche $\lambda \in \mathbb{R}$ existieren keine Lösungen?
- (d) Man berechne die Lösung für $\lambda = 1$.
- (e) Man berechne die Lösung zu b).
- (f) Wie können die Ergebnisse von a), b) und c) geometrisch gedeutet werden?

17. Man bestimme - falls möglich - die Inverse folgender Matrizen

$$(a) A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \\ 0 & 1 & 2 \end{pmatrix} \quad (b) A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & -1 & 2 \\ 1 & -1 & 2 & 1 \\ 1 & 3 & 3 & 2 \end{pmatrix}$$

$$(c) A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 3 & 1 & 2 \\ 1 & 2 & -1 & 1 \\ 5 & 9 & 1 & 6 \end{pmatrix} \quad (d) A = \begin{pmatrix} 1 & 2 & -3 & 1 \\ -1 & 3 & -3 & -2 \\ 2 & 0 & 1 & 5 \\ 3 & 1 & -2 & 5 \end{pmatrix}$$

$$(e) A = \begin{pmatrix} 1 & 0 & 2 & 5 \\ 0 & 2 & 4 & 8 \\ 1 & 0 & 4 & 9 \\ 1 & 2 & 8 & 18 \end{pmatrix} \quad (f) A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 1 & 2 & 4 \end{pmatrix}$$

$$(g) A = \begin{pmatrix} 1 & 4 & -1 \\ 2 & 10 & -2 \\ -1 & 2 & 4 \end{pmatrix} \quad (h) A = \begin{pmatrix} 1 & 3 & 2 & -3 \\ 2 & 1 & -1 & 2 \\ -1 & 2 & 1 & -1 \\ 2 & -4 & -1 & 4 \end{pmatrix}$$

18. Man berechne - falls möglich - die Inverse der folgenden Matrizen

$$\begin{aligned} a) \begin{pmatrix} 4 & 7 \\ 1 & 2 \end{pmatrix} & b) \begin{pmatrix} 4 & -2 \\ 2 & 0 \end{pmatrix} & c) \begin{pmatrix} 3 & 4 \\ 6 & 8 \end{pmatrix} & d) \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} & e) \begin{pmatrix} \frac{3}{4} & \frac{3}{5} \\ \frac{5}{6} & \frac{2}{3} \end{pmatrix} & f) \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2} \\ 2\sqrt{2} & \sqrt{2} \end{pmatrix} \\ g) \begin{pmatrix} -1.5 & -4.2 \\ 0.5 & 2.4 \end{pmatrix} & h) \begin{pmatrix} 2.54 & 8.128 \\ 0.25 & 0.8 \end{pmatrix} & i) \begin{pmatrix} a & -b \\ b & a \end{pmatrix} & j) \begin{pmatrix} \frac{1}{a} & \frac{1}{b} \\ \frac{1}{c} & \frac{1}{d} \end{pmatrix}, a \cdot b \cdot c \cdot d \neq 0 \\ k) \begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix} & l) \begin{pmatrix} -2 & 4 \\ 3 & -1 \end{pmatrix} & m) \begin{pmatrix} 6 & -4 \\ -3 & 2 \end{pmatrix} & n) \begin{pmatrix} 1 & a \\ -a & 1 \end{pmatrix} & o) \begin{pmatrix} 2 & 3 & 0 \\ 1 & -2 & -1 \\ 2 & 0 & -1 \end{pmatrix} \\ p) \begin{pmatrix} 1 & -1 & 2 \\ 3 & 1 & 2 \\ 2 & 3 & -1 \end{pmatrix} & q) \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} & r) \begin{pmatrix} a & 0 & 0 \\ 1 & a & 0 \\ 0 & 1 & a \end{pmatrix} & s) \begin{pmatrix} 0 & a & 0 \\ b & 0 & c \\ 0 & d & 0 \end{pmatrix} \end{aligned}$$

$$t) \begin{pmatrix} 0 & -1 & 1 & 0 \\ 2 & 1 & 0 & 2 \\ 1 & -1 & 3 & 0 \\ 0 & 1 & 1 & -1 \end{pmatrix} \quad u) \begin{pmatrix} \sqrt{2} & 0 & 2\sqrt{2} & 0 \\ -4\sqrt{2} & \sqrt{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 3 & 1 \end{pmatrix} \quad v) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ a & b & c & d \end{pmatrix}$$

19. Man löse folgende Gleichungssysteme unter Verwendung der Inversen der Koeffizientenmatrix

$$a) \begin{array}{lcl} 2x + y = -1 \\ 5x + 3y = 2 \end{array} \quad b) \begin{array}{lcl} x_1 - x_2 = 1 \\ 2x_1 + x_2 = 2 \end{array} \quad c) \begin{array}{lcl} 2x_1 + x_2 = 3 \\ 3x_1 + 2x_2 = 2 \end{array}$$

20. Im folgenden bestimme man A^{-1} und berechne damit die Lösung der drei Systeme $Ax = b_1$, $Ax = b_2$ und $Ax = b_3$

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 6 \end{pmatrix}, \quad b_1 = \begin{pmatrix} 3 \\ 5 \end{pmatrix}, \quad b_2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \quad b_3 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

21. Man löse die folgenden Matrix-Gleichungen für X . Man vereinfache das Ergebnis soweit als möglich. Alle Matrizen seien regulär.

$$a) \ X A^2 = A^{-1} \quad c) \ (A^{-1} X)^{-1} = A(B^{-2} A)^{-1}$$

$$b) \ A X B = (B A)^2 \quad d) \ A B X A^{-1} B^{-1} = I + A$$

22. Man berechne die Inversen der angegebenen Matrizen:

$$(a) \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad (b) \begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad (c) \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

23. Für welche $a \in \mathbb{R}$ sind die Matrizen

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 2 & a \end{pmatrix} \quad \text{bzw.} \quad B = \begin{pmatrix} a-2 & 2 & 2 \\ 2 & a-2 & 2 \\ 2 & 2 & a-2 \end{pmatrix} \quad \text{regulär?}$$

Man berechne in diesem Fall A^{-1} und B^{-1} .

24. Man bestimme eine (3×3) -Matrix A so, daß gilt:

$$(a) \ A \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -4 \end{pmatrix} \quad (b) \ A \cdot \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 7 \\ -7 \end{pmatrix} \quad (c) \ A \cdot \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix}$$

25. Man bestimme die LR -Zerlegung der folgenden Matrizen und löse dann das Gleichungssystem $A \cdot x = b$ dadurch, dass man zuerst $Ly = b$ löst und anschließend $Rx = y$.

$$a) \ A = \begin{pmatrix} -2 & 1 \\ 2 & 5 \end{pmatrix}, \quad b = \begin{pmatrix} 5 \\ 1 \end{pmatrix} \quad b) \ A = \begin{pmatrix} 4 & -2 \\ 2 & 3 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 8 \end{pmatrix}$$

$$c) \quad A = \begin{pmatrix} 2 & 1 & -2 \\ -2 & 3 & -4 \\ 4 & -3 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix} \quad d) \quad A = \begin{pmatrix} 2 & -4 & 0 \\ 3 & -1 & 4 \\ -1 & 2 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 2 \\ 0 \\ -5 \end{pmatrix}$$

26. Für die folgenden Matrizen führe man die LR-Zerlegung ohne Zeilenumtauschung durch und mache die Probe

$$A = \begin{pmatrix} 2 & 4 \\ 2 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & -6 \\ -3 & 4 & 7 \\ 2 & 4 & 3 \end{pmatrix}$$

$$C = \begin{pmatrix} 2 & 1 & -2 \\ 1 & -1 & 2 \\ 3 & 1 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} -2 & -4 & 2 & -2 \\ 1 & 5 & 5 & -8 \\ -1 & 0 & 7 & -11 \\ 2 & 7 & 3 & -3 \end{pmatrix}$$

27. Für die folgenden Matrizen führe man die LR-Zerlegung mit Zeilenumtauschung durch und mache die Probe

$$A = \begin{pmatrix} 2 & 1 & 4 & 1 \\ 2 & 1 & 3 & 2 \\ 4 & 3 & 9 & 3 \\ 6 & 1 & 0 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 3 & -1 \\ 1 & 0 & 0 & 2 \\ 1 & 2 & -1 & 0 \end{pmatrix}$$

28. Man führe die LR-Zerlegung mit Zeilenumtauschung für folgende Matrizen durch und mache die Probe

$$a) \quad A = \begin{pmatrix} 0 & 1 & 4 \\ -1 & 2 & 1 \\ 1 & 3 & 3 \end{pmatrix} \quad b) \quad A = \begin{pmatrix} 2 & 1 & 4 & 1 \\ 2 & 1 & 3 & 2 \\ 4 & 3 & 9 & 3 \\ 6 & 1 & 0 & 4 \end{pmatrix} \quad c) \quad A = \begin{pmatrix} 0 & -1 & 1 & 3 \\ -1 & 1 & 1 & 2 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

$$d) \quad A = \begin{pmatrix} 2 & 6 & -4 \\ -4 & -12 & 11 \\ 3 & 14 & -16 \end{pmatrix} \quad e) \quad A = \begin{pmatrix} 2 & 4 & 0 & -2 \\ -4 & -8 & 0 & 3 \\ 3 & 7 & 2 & 4 \\ 5 & 6 & 1 & -8 \end{pmatrix}$$

29. Man löse das lineare Gleichungssystem $A\vec{x} = \vec{b}$ mit

$$A = \begin{pmatrix} 5 & 7 & 6 & 5 \\ 7 & 10 & 8 & 7 \\ 6 & 8 & 10 & 9 \\ 5 & 7 & 9 & 10 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 23 \\ 32 \\ 33 \\ 31 \end{pmatrix}$$

mittels LR-Zerlegung von A . Danach ersetze man den Vektor \vec{b} durch

$$\vec{b}_1 = \begin{pmatrix} 23.01 \\ 31.99 \\ 32.99 \\ 31.01 \end{pmatrix} \quad \text{bzw.} \quad \vec{b}_2 = \begin{pmatrix} 23.1 \\ 31.9 \\ 32.9 \\ 31.1 \end{pmatrix}$$

und löse die neuen Systeme. Vergleiche die Resultate.

30. Man verwende eine 2-stellige Gleitkommaarithmetik zur Lösung des Systems

$$\begin{aligned} 0.89x_1 + 0.53x_2 &= 0.36 \\ 0.47x_1 + 0.28x_2 &= 0.19 \end{aligned}$$

und vergleiche dies mit dem exakten Ergebnis.

31. Man berechne $\det A$ für

$$(a) A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 8 & 7 & 6 & 5 \\ -2 & 0 & 4 & 1 \\ 3 & 4 & -3 & 0 \end{pmatrix} \quad (b) A = \begin{pmatrix} 1 & 3 & 4 & 0 \\ 2 & 5 & 7 & 1 \\ -1 & 2 & -3 & 0 \\ 0 & 0 & 1 & 4 \end{pmatrix}.$$

32. Man berechne die folgenden Determinanten

$$(a) \begin{vmatrix} 1 & 0 & 0 & a \\ b & 0 & 1 & 0 \\ 0 & 1 & 0 & c \\ 0 & 0 & 0 & d \end{vmatrix} \quad (b) \begin{vmatrix} 0 & 2 & 0 & a \\ -3 & 0 & 0 & b \\ 0 & 0 & c & 4 \\ 0 & 0 & d & 0 \end{vmatrix} \quad a, b, c, d \in \mathbb{R}.$$

$$(c) \begin{vmatrix} 1 & 2 & 3 \\ -1 & 5 & 2 \\ 3 & 2 & 0 \end{vmatrix} \quad (d) \begin{vmatrix} 4 & -4 & 2 & 1 \\ 1 & 2 & 0 & 3 \\ 2 & 0 & 3 & 4 \\ 0 & -3 & 2 & 1 \end{vmatrix} \quad (e) \begin{vmatrix} 1 & 0 & a & 0 \\ 0 & 1 & 0 & a \\ b & 1 & 0 & 0 \\ 0 & 0 & b & 1 \end{vmatrix}$$

$$(f) \begin{vmatrix} 3 & 1 & 2 & -1 \\ 2 & 0 & 3 & -7 \\ 1 & 3 & 4 & -5 \\ 0 & -1 & 1 & -5 \end{vmatrix} \quad (g) \begin{vmatrix} 1 & -2 & 3 & -2 & -2 \\ 2 & -1 & 1 & 3 & 2 \\ 1 & 1 & 2 & 1 & 1 \\ 1 & -4 & -3 & -2 & -5 \\ 3 & -2 & 2 & 2 & -2 \end{vmatrix}$$

33. Man bestimme die Determinante folgender Matrizen ($a, b, c \in \mathbb{R}$):

$$\begin{pmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{pmatrix}, \begin{pmatrix} a & b+c & 1 \\ b & c+a & 1 \\ c & a+b & 1 \end{pmatrix}, \begin{pmatrix} a-b & b-c & c-a \\ b-c & c-a & a-b \\ c-a & a-b & b-c \end{pmatrix}, \begin{pmatrix} -a & a-1 & a+1 \\ 1 & 2 & 3 \\ 2-a & a+3 & a+7 \end{pmatrix}.$$

34. Man berechne die Determinanten der folgenden Matrizen

$$a) A = \begin{pmatrix} 1 & 0 & 3 \\ 5 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \quad b) A = \begin{pmatrix} 0 & 1 & -1 \\ 2 & 3 & -2 \\ -1 & 3 & 0 \end{pmatrix} \quad c) A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix}$$

$$d) A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \quad e) A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix} \quad f) A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \quad g) A = \begin{pmatrix} 5 & 2 & 2 \\ -1 & 1 & 2 \\ 3 & 0 & 0 \end{pmatrix}$$

$$h) A = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 0 & 1 \\ 3 & -2 & 1 \end{pmatrix} \quad i) A = \begin{pmatrix} -4 & 1 & 3 \\ 2 & -2 & 4 \\ 1 & -1 & 0 \end{pmatrix} \quad j) A = \begin{pmatrix} \cos \alpha & \sin \alpha & \tan \alpha \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{pmatrix}$$

k) $A = \begin{pmatrix} a & b & 0 \\ 0 & a & b \\ a & 0 & b \end{pmatrix}$ l) $A = \begin{pmatrix} 0 & a & 0 \\ b & c & d \\ 0 & e & 0 \end{pmatrix}$ m) $A = \begin{pmatrix} 1 & -1 & 0 & 3 \\ 2 & 5 & 2 & 6 \\ 0 & 1 & 0 & 0 \\ 1 & 4 & 2 & 1 \end{pmatrix}$

n) $A = \begin{pmatrix} 2 & 0 & 3 & -1 \\ 1 & 0 & 2 & 2 \\ 0 & -1 & 1 & 4 \\ 2 & 0 & 1 & -3 \end{pmatrix}$ o) $A = \begin{pmatrix} 0 & 0 & 0 & a \\ 0 & 0 & b & c \\ 0 & d & e & f \\ g & h & i & j \end{pmatrix}$ p) $A = \begin{pmatrix} 1 & 1 & 1 \\ 3 & 0 & -2 \\ 2 & 2 & 2 \end{pmatrix}$

q) $A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & -2 & 5 \\ 0 & 0 & 4 \end{pmatrix}$ r) $A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 5 & 2 \\ 3 & -1 & 4 \end{pmatrix}$ s) $A = \begin{pmatrix} 2 & 3 & -4 \\ 1 & -3 & -2 \\ -1 & 5 & 2 \end{pmatrix}$

t) $A = \begin{pmatrix} 2 & 3 & -4 \\ 1 & -3 & -2 \\ 1 & 5 & 2 \end{pmatrix}$ u) $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 1 \\ 1 & 6 & 4 \end{pmatrix}$ v) $A = \begin{pmatrix} 4 & 1 & 3 \\ -2 & 0 & -2 \\ 5 & 4 & 1 \end{pmatrix}$

w) $\begin{pmatrix} 0 & 2 & 0 & 0 \\ -3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ x) $\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$ y) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

2 Vektoren im anschaulichen Raum

35. Man überprüfe, ob der Vektor v eine Linearkombination der übrigen Vektoren ist.

a) $v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, u_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, u_2 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$

b) $v = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, u_1 = \begin{pmatrix} 4 \\ -2 \end{pmatrix}, u_2 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$

c) $v = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, u_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, u_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$

d) $v = \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix}, u_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, u_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$

e) $v = \begin{pmatrix} 1 \\ -2 \\ 0 \end{pmatrix}, u_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, u_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, u_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$

f) $v = \begin{pmatrix} 3.2 \\ 2.0 \\ -2.6 \end{pmatrix}, u_1 = \begin{pmatrix} 1.0 \\ 0.4 \\ 4.8 \end{pmatrix}, u_2 = \begin{pmatrix} 3.4 \\ 1.4 \\ -6.4 \end{pmatrix}, u_3 = \begin{pmatrix} -1.2 \\ 0.2 \\ -1.0 \end{pmatrix}$

36. Man berechne den Winkel zwischen u und v :

$$\text{a) } u = \begin{pmatrix} 3 \\ 0 \end{pmatrix}, \quad v = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \quad \text{b) } u = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}, \quad v = \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}$$

$$\text{c) } u = \begin{pmatrix} 4 \\ 3 \\ -1 \end{pmatrix}, \quad v = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \quad \text{d) } u = \begin{pmatrix} 0.9 \\ 2.1 \\ 1.2 \end{pmatrix}, \quad v = \begin{pmatrix} -4.5 \\ 2.6 \\ -0.8 \end{pmatrix}$$

37. Man berechne $u \times v$ für

$$\text{a) } u = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad v = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix} \quad \text{b) } u = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}, \quad v = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

$$\text{c) } u = \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix}, \quad v = \begin{pmatrix} 2 \\ -4 \\ -6 \end{pmatrix} \quad \text{d) } u = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad v = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

3 Allgemeine Vektorräume

38. Welche der folgenden Mengen sind lineare Teilräume des \mathbb{R}^n ?

- (a) $\{x = (x_1, x_2, \dots, x_n) \mid x_1 = a, a = \text{konstant}\}$,
- (b) $\{x = (x_1, x_2, \dots, x_n) \mid x_1 \geq 0\}$,
- (c) $\{x = (x_1, x_2, \dots, x_n) \mid x_1 \cdot x_2 = 0\}$,
- (d) $\{x = (x_1, x_2, \dots, x_n) \mid x_1 = 0\} \cup \{x = (x_1, x_2, \dots, x_n) \mid x_2 = 0\}$,
- (e) $\{x = (x_1, x_2, \dots, x_n) \mid x_1 = 0\} \cap \{x = (x_1, x_2, \dots, x_n) \mid x_2 = 0\}$.

39. Im \mathbb{R}^3 sind folgende Teilmengen gegeben:

- (a) $T_1 = \{a \in \mathbb{R}^3 \mid a = (a_1, a_2, a_2), a_1, a_2 \in \mathbb{R}\}$,
- (b) $T_2 = \{a \in \mathbb{R}^3 \mid a = (a_1, a_2, a_3), a_1 \geq 0\}$,
- (c) $T_3 = \{a \in \mathbb{R}^3 \mid a = (a_1, a_2, a_3), a_1 + a_2 + a_3 = 0\}$.

Welche dieser Teilmengen ist ein linearer Teilraum?

- 40. (a) Für welche $\alpha, \beta \in \mathbb{R}$ sind die Vektoren $(\alpha, 0, 1), (0, \beta, 2)$ linear unabhängig?
- (b) Man bestimme diejenigen Werte von $x \in \mathbb{R}$, für die die Vektoren $(1, 2, 1), (1, x, -1), (1, -2, -1)$ linear unabhängig sind.
- 41. Man prüfe folgende Vektoren auf lineare Unabhängigkeit:
 - (a) $a = (3, 6, 9), b = (-1, 8, 3), c = (11, 2, 21);$

- (b) $a = (3, 1, 7)$, $b = (2, -2, 18)$, $c = (-1, 1, -9)$;
(c) $a = (2, -1, 7, 6)$, $b = (-1, 1, -2, -4)$, $c = (1, 1, 8, 0)$, $d = (0, 1, 3, -2)$.

42. Man gebe die größte Anzahl von Vektoren aus den angegebenen Familien an, die linear unabhängig sind:

- (a) \mathbb{R}^5 : $u_1 = (1, 0, -1, 1, 2)$, $u_2 = (0, 1, 1, 2, 3)$ und $u_3 = (1, 1, 0, 1, 0)$.
(b) \mathbb{R}^4 : $u_1 = (1, 0, -1, 1)$, $u_2 = (0, 1, -1, -1)$, $u_3 = (1, 1, 0, -1)$, $u_4 = (0, -3, 0, 4)$.

43. Man untersuche, ob die angegebenen Vektoren linear unabhängig sind.

a) $\begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$, b) $\begin{pmatrix} 1 \\ 4 \\ 4 \end{pmatrix}$ b) $\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$, c) $\begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}$, c) $\begin{pmatrix} 1 \\ -5 \\ 2 \end{pmatrix}$, c) $\begin{pmatrix} -2 \\ 3 \\ 7 \end{pmatrix}$, d) $\begin{pmatrix} 4 \\ -1 \\ 5 \end{pmatrix}$, d) $\begin{pmatrix} 3 \\ 1 \\ 3 \end{pmatrix}$, d) $\begin{pmatrix} 5 \\ 0 \\ 2 \end{pmatrix}$

d) $\begin{pmatrix} -1 \\ 1 \\ 2 \\ 1 \end{pmatrix}$, e) $\begin{pmatrix} 3 \\ 2 \\ 2 \\ 4 \end{pmatrix}$, f) $\begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$, e) $\begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix}$, e) $\begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}$, f) $\begin{pmatrix} 1 \\ 0 \\ 1 \\ -1 \end{pmatrix}$, f) $\begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$

f) $\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$, g) $\begin{pmatrix} 0 \\ 0 \\ 2 \\ 1 \end{pmatrix}$, h) $\begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$, g) $\begin{pmatrix} 3 \\ -1 \\ 1 \\ -1 \end{pmatrix}$, h) $\begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}$, i) $\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$, i) $\begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}$, j) $\begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}$, j) $\begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix}$

h) $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$, i) $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, j) $\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ i) $\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$, j) $\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$, i) $\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$, j) $\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$

44. Man bestimme eine Basis des von den folgenden Vektoren aufgespannten Vektorraumes

a) $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, b) $\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$, c) $\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$, b) $\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, c) $\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$, d) $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$, d) $\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$

c) $\begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}$, d) $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, c) $\begin{pmatrix} 4 \\ -4 \\ 1 \end{pmatrix}$, d) $\begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix}$, c) $\begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}$, d) $\begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix}$

45. Die Vektoren $u = (1, -2, 5, -3)$, $v = (2, 3, 1, -4)$, $w = (3, 8, -3, -5)$ erzeugen einen Unterraum U des \mathbb{R}^4 . Bestimmen Sie die Dimension von U .

46. Bilden die folgenden Vektoren eine Basis des \mathbb{R}^3 bzw. des \mathbb{R}^4 ?

a) $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, b) $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, c) $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$, b) $\begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}$, c) $\begin{pmatrix} -1 \\ 5 \\ 1 \end{pmatrix}$, d) $\begin{pmatrix} 1 \\ -3 \\ 1 \end{pmatrix}$

$$c) \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad d) \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

47. Zu den folgenden zwei Vektoren des \mathbb{R}^3 gebe man einen dritten Vektor an, so dass diese drei Vektoren eine Basis des \mathbb{R}^3 bilden.

$$\begin{array}{lll} a) \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 2 \\ 5 \end{pmatrix} & b) \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} & c) \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \\ d) \begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix} & e) \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 4 \\ 2 \\ 0 \end{pmatrix} & f) \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 4 \\ 5 \end{pmatrix} \\ g) \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} & h) \begin{pmatrix} 2 \\ 0 \\ 3 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix} & i) \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} & j) \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \end{array}$$

4 Lineare Abbildungen

48. Man untersuche, ob die folgenden Abbildungen F linear sind.

Dabei ist $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$.

$$\begin{array}{lll} a) F(x) = \begin{pmatrix} x_1 + x_2 \\ x_1 - x_2 \end{pmatrix} & b) F(x) = \begin{pmatrix} -x_2 \\ x_1 + 2x_2 \\ 3x_1 - 4x_2 \end{pmatrix} & c) F(x) = \begin{pmatrix} 2x_1 \\ x_2 - x_1 \\ x_1 - x_2 \end{pmatrix} \\ d) F(y) = \begin{pmatrix} y_1 - y_2 + y_3 \\ 2y_1 + y_2 - 3y_3 \end{pmatrix} & e) F(y) = \begin{pmatrix} y_1 + y_3 \\ y_2 + y_3 \\ y_1 + y_2 \end{pmatrix} & f) F(y) = \begin{pmatrix} y_1 + y_2 \\ y_2 + y_3 \\ y_3 + y_1 \end{pmatrix} \\ g) F(x) = \begin{pmatrix} x_2 \\ x_1^2 \end{pmatrix} & h) F(x) = \begin{pmatrix} |x_1| \\ |x_2| \end{pmatrix} & i) F(x) = \begin{pmatrix} x_1 \cdot x_2 \\ x_1 + x_2 \end{pmatrix} & j) F(x) = \begin{pmatrix} x_1 + 1 \\ x_2 - 1 \end{pmatrix} \end{array}$$

49. Sei $F : M(2 \times 2) \rightarrow M(2 \times 2)$ eine lineare Abbildung mit

$$F \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}$$

Man führe für die nachstehend angegebenen Matrizen folgendes Programm durch:

- (i) Welche der folgenden Matrizen ist ein Element von $\text{Kern}(F)$.
(ii) Welche der folgenden Matrizen ist ein Element von $\text{Bild}(F)$.
(iii) Man beschreibe $\text{Kern}(F)$ und $\text{Bild}(F)$ jeweils durch die Angabe einer Basis.

$$\text{a)} \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}, \quad \text{b)} \begin{pmatrix} 0 & 4 \\ 2 & 0 \end{pmatrix}, \quad \text{c)} \begin{pmatrix} 3 & 0 \\ 0 & -3 \end{pmatrix}$$

50. Sei $F : M(2 \times 2) \rightarrow \mathbb{R}$, $F \begin{pmatrix} a & b \\ c & d \end{pmatrix} = a + d$, eine lineare Abbildung.

- (i) Welche Zahl liegt in $\text{Bild}(F)$?
a) 0 b) -2 c) $1/\sqrt{2}$
(ii) Welche der folgenden Matrizen liegt in $\text{Kern}(F)$?
a) $\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$, b) $\begin{pmatrix} 0 & 4 \\ 2 & 0 \end{pmatrix}$, c) $\begin{pmatrix} 3 & 0 \\ 0 & -3 \end{pmatrix}$
(iii) Man beschreibe $\text{Kern}(F)$ und $\text{Bild}(F)$ jeweils durch die Angabe einer Basis.

51. Sei $F : \mathbb{P}_2 \rightarrow \mathbb{R}^2$ eine lineare Abbildung definiert durch

$$F(a + bt + ct^2) = \begin{pmatrix} a - b \\ b + c \end{pmatrix}$$

- (i) Welches der folgenden Polynome liegt in $\text{Kern}(F)$?
a) $1 + t$ b) $t - t^2$, c) $1 + t - t^2$
(ii) Welcher der folgenden Vektoren liegt in $\text{Bild}(F)$?
a) $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, b) $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, c) $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$
(iii) Man beschreibe $\text{Kern}(F)$ und $\text{Bild}(F)$ jeweils durch die Angabe einer Basis.

52. Sei $F : \mathbb{P}_2 \rightarrow \mathbb{P}_2$ die lineare Abbildung definiert durch $F(p(t)) = t \cdot p'(t)$.

- (i) Welches der folgenden Polynome liegt in $\text{Kern}(F)$?
a) 2 , b) t^2 , c) $1 - t$
(ii) Welches der Polynome aus Teil (i) liegt in $\text{Bild}(F)$?
(iii) Man beschreibe $\text{Kern}(F)$ und $\text{Bild}(F)$ jeweils durch die Angabe einer Basis.

53. Die Matrix $A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ ordnet jedem Vektor $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ einen weiteren Vektor $x' = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}$ gemäß $x' = A \cdot x$ zu. Auf welche Vektoren werden die Basisvektoren $b_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $b_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ abgebildet? Worauf wird die Gerade $x_2 = kx_1 + d$ abgebildet,

worauf die Kreise $x_1^2 + x_2^2 = r^2$?

Hinweis: Man berechne das Bild eines beliebigen Vektors x , der auf der Geraden liegt, z.B. $x = \begin{pmatrix} x_1 \\ kx_1 + d \end{pmatrix}$, $x_1 \in \mathbb{R}$, bzw. das eines beliebigen Vektors x , der auf der Kreislinie liegt, z.B. $x = \begin{pmatrix} x_1 \\ \sqrt{r^2 - x_1^2} \end{pmatrix}$, $x_1 \in \mathbb{R}$ für die obere Halbkreislinie.

54. Welche geometrische Bedeutung haben die Koordinatentransformationen $y = A \cdot x$ mit

$$(a) \quad A = \begin{pmatrix} -\sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & -\sqrt{2}/2 \end{pmatrix}, \quad (b) \quad \begin{pmatrix} 1/2 & -\sqrt{3}/2 \\ \sqrt{3}/2 & 1/2 \end{pmatrix}?$$

Hinweis: Man ermittle zunächst die Bilder der kanonischen Basisvektoren und dann das Bild eines beliebigen Vektors.

55. Gegeben sind die Matrizen

$$A = \begin{pmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} \cos \psi & 0 & -\sin \psi \\ 0 & 1 & 0 \\ \sin \psi & 0 & \cos \psi \end{pmatrix}.$$

Man beschreibe geometrisch die folgenden linearen Abbildungen:

- (a) $L_1 : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $L_1(x) = A \cdot x$,
- (b) $L_2 : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $L_2(x) = B \cdot x$,
- (c) $L_3 : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $L_3(x) = A \cdot B \cdot x$,
- (d) $L_4 : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $L_4(x) = B \cdot A \cdot x$.

56. Sei $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ dargestellt durch $A = \begin{pmatrix} 2 & -1 \\ 3 & 4 \end{pmatrix}$

Man berechne $F(u)$ und $F(v)$ für

a) $u = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $v = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$ b) $u = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$, $v = \begin{pmatrix} 4 \\ -3 \end{pmatrix}$

57. Sei $F : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ dargestellt durch $A = \begin{pmatrix} 4 & 0 & -1 \\ -2 & 1 & 3 \end{pmatrix}$

Man berechne $F(u)$ und $F(v)$ für

a) $u = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$, $v = \begin{pmatrix} 0 \\ 5 \\ -1 \end{pmatrix}$ b) $u = \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix}$, $v = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$

58. Man bestimme die bezüglich der kanonischen Basen darstellenden Matrizen der Abbildungen a), b), c), d), e) und f) von Beispiel 48.

59. Man zeige, dass $w \in L(\mathcal{B})$ gilt und bestimme den Koordinatenvektor $c_{\mathcal{B}}(w)$

a) $\mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \right\}$, $w = \begin{pmatrix} 1 \\ 6 \\ 2 \end{pmatrix}$ b) $\mathcal{B} = \left\{ \begin{pmatrix} 3 \\ 1 \\ 4 \end{pmatrix}, \begin{pmatrix} 5 \\ 1 \\ 6 \end{pmatrix} \right\}$, $w = \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix}$

60. \mathcal{A}, \mathcal{B} und \mathcal{C} stellen die kanonischen Basen des \mathbb{R}^2 bzw. des \mathbb{R}^3 dar. Man verifiziere die Relation $M_{\mathcal{C}}^{\mathcal{A}}(G \circ F) = M_{\mathcal{C}}^{\mathcal{B}}(G) \cdot M_{\mathcal{B}}^{\mathcal{A}}(F)$ für die folgenden Abbildungen.

$$\text{a)} \quad F(x) = \begin{pmatrix} x_1 & - & x_2 \\ x_1 & + & x_2 \end{pmatrix}, \quad G(x) = \begin{pmatrix} 2x_1 \\ -x_2 \end{pmatrix}$$

$$\text{b)} \quad F(x) = \begin{pmatrix} x_1 & + & 2x_2 \\ -3x_1 & + & x_2 \end{pmatrix}, \quad G(x) = \begin{pmatrix} x_1 & + & 3x_2 \\ x_1 & - & x_2 \end{pmatrix}$$

$$\text{c)} \quad F(x) = \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix}, \quad G(x) = \begin{pmatrix} x_1 & - & 3x_2 \\ 2x_1 & + & x_2 \\ x_1 & - & x_2 \end{pmatrix}$$

$$\text{d)} \quad F(y) = \begin{pmatrix} y_1 & + & y_2 & - & y_3 \\ 2y_1 & - & y_2 & + & y_3 \end{pmatrix}, \quad G(x) = \begin{pmatrix} 4x_1 & - & 2x_2 \\ -x_1 & + & x_2 \end{pmatrix}$$

$$\text{e)} \quad F(y) = \begin{pmatrix} y_1 & + & 2y_2 \\ 2y_2 & - & y_3 \end{pmatrix}, \quad G(x) = \begin{pmatrix} x_1 & - & x_2 \\ x_1 & + & x_2 \\ -x_1 & + & x_2 \end{pmatrix}$$

$$\text{f)} \quad F(y) = \begin{pmatrix} y_1 & + & y_2 \\ y_2 & + & y_3 \\ y_1 & + & y_3 \end{pmatrix}, \quad G(y) = \begin{pmatrix} y_1 & - & y_2 \\ y_2 & - & y_3 \\ -y_1 & + & y_3 \end{pmatrix}$$

Bezüglich der Notation siehe Beispiel 48.

61. Man berechne die Koordinatenvektoren $c_{\mathcal{A}}(x)$ und $c_{\mathcal{B}}(x)$ des Vektors x bzgl. der Basen \mathcal{A} und \mathcal{B} . Anschließend bestimme man die Matrix des Basiswechsels $\mathcal{A} \rightarrow \mathcal{B}$, und bestimme damit $c_{\mathcal{B}}(x)$ aus $c_{\mathcal{A}}(x)$.

$$\text{a)} \quad x = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \mathcal{A} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}, \mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\} \quad \text{im } \mathbb{R}^2$$

$$\text{b)} \quad x = \begin{pmatrix} 4 \\ -1 \end{pmatrix}, \mathcal{A} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}, \mathcal{B} = \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 3 \end{pmatrix} \right\} \quad \text{im } \mathbb{R}^2$$

$$\text{c)} \quad x = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \mathcal{A} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}, \mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \quad \text{im } \mathbb{R}^3$$

$$\text{d)} \quad x = \begin{pmatrix} 3 \\ 1 \\ 5 \end{pmatrix}, \mathcal{A} = \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}, \mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\} \quad \text{im } \mathbb{R}^3$$

62. In den folgenden Beispielen ist die die lineare Abbildung $F : V \rightarrow W$ darstellende Matrix $A = M_{\mathcal{B}}^{\mathcal{A}}(F)$ gesucht. Dann berechne man $w = F(v)$ einmal direkt und einmal über die Relation $c_{\mathcal{B}}(w) = A \cdot c_{\mathcal{A}}(v)$.

- a) $F = \mathbb{P}_1 \rightarrow \mathbb{P}_1, \quad F(a + bt) = b - at, \quad \mathcal{A} = \mathcal{B} = \{1, t\}, \quad v = p(t) = 4 + 2t$
- b) $F = \mathbb{P}_1 \rightarrow \mathbb{P}_1, \quad F(a + bt) = b - at, \quad \mathcal{A} = \{1 + t, 1 - t\}, \quad \mathcal{B} = \{1, t\}, \quad v = 4 + 2t$
- c) $F = \mathbb{P}_2 \rightarrow \mathbb{P}_2, \quad F(p(t)) = p(t+2), \quad \mathcal{A} = \{1, t, t^2\}, \quad \mathcal{B} = \{1, t+2, (t+2)^2\},$
 $v = a + bt + ct^2$
- d) $F = \mathbb{P}_2 \rightarrow \mathbb{P}_2, \quad F(p(t)) = p(t+2), \quad \mathcal{A} = \{1, t+2, (t+2)^2\}, \quad \mathcal{B} = \{1, t, t^2\},$
 $v = a + bt + ct^2$
- e) $F = \mathbb{P}_2 \rightarrow \mathbb{R}^2, \quad F(p(t)) = \begin{pmatrix} p(0) \\ p(1) \end{pmatrix}, \quad \mathcal{A} = \{1, t, t^2\}, \quad \mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\},$
 $v = a + bt + ct^2$
- f) $F = \mathbb{P}_2 \rightarrow \mathbb{R}^2, \quad F(p(t)) = \begin{pmatrix} p(0) \\ p(1) \end{pmatrix}, \quad \mathcal{A} = \{t^2, t, 1\}, \quad \mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\},$
 $v = a + bt + ct^2$
- g) $F = \mathbb{R}^2 \rightarrow \mathbb{R}^3 \quad F \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a + 2b \\ -a \\ b \end{pmatrix},$
 $\mathcal{A} = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix} \right\}, \quad \mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}, \quad v = \begin{pmatrix} -7 \\ 7 \end{pmatrix}$
- h) Wie Beispiel g) mit $v = \begin{pmatrix} a \\ b \end{pmatrix}$
- i) $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad F \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a - b \\ 2a + 3b \end{pmatrix}, \quad \mathcal{A} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}, \quad \mathcal{B} = \left\{ \begin{pmatrix} -1 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 1 \end{pmatrix} \right\},$
 $v = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$
- j) $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad F \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 4a - b \\ 3a + 2b \end{pmatrix}, \quad \mathcal{A} = \mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 3 \end{pmatrix} \right\}, \quad v = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$

63. Man stelle die Vektoren der Basis $\mathcal{A} = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ bezüglich der Basis $\mathcal{B} = \{(0, 1, 1), (1, 0, 1), (1, 1, 0)\}$ dar (d.h. als Linearkombination der genannten Basisvektoren) und bestimme die Transformationsmatrix der Koordinatentransformation $\mathcal{B} \rightarrow \mathcal{A}$.

5 Unitäre Räume

64. Man zeige, daß mit den folgenden Abbildungen

$$\langle \cdot, \cdot \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$$

für $u = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, v = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, ein Skalarprodukt vorliegt.

- a) $\langle u, v \rangle = 4x_1y_1 - x_1y_2 - x_2y_1 + x_2y_2$
- b) $\langle u, v \rangle = x_1y_1 + 2x_1y_2 + 2x_2y_1 + 5x_2y_2$
- c) $\langle u, v \rangle = x_1y_1 - 2x_1y_2 - 2x_2y_1 + 5x_2y_2$
- d) $\langle u, v \rangle = 5x_1y_1 + 3x_1y_2 + 3x_2y_1 + 2x_2y_2$

65. Gegeben sind die Vektoren

$$a) u = \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \quad v = \begin{pmatrix} 3 \\ 1 \end{pmatrix} \quad b) u = \begin{pmatrix} 3 \\ -2 \end{pmatrix}, \quad v = \begin{pmatrix} 4 \\ 6 \end{pmatrix}$$

$$c) u = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad v = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} \quad d) u = \begin{pmatrix} 3.2 \\ -0.6 \\ -1.4 \end{pmatrix}, \quad v = \begin{pmatrix} 1.5 \\ 4.1 \\ -0.2 \end{pmatrix}$$

$$e) u = \begin{pmatrix} 1 \\ \sqrt{2} \\ \sqrt{3} \\ 0 \end{pmatrix}, \quad v = \begin{pmatrix} 4 \\ -\sqrt{2} \\ 0 \\ -5 \end{pmatrix} \quad f) u = \begin{pmatrix} 1.12 \\ -3.25 \\ 2.07 \\ -1.83 \end{pmatrix}, \quad v = \begin{pmatrix} -2.29 \\ 1.72 \\ 4.33 \\ -1.54 \end{pmatrix}$$

Man berechne $\langle u, v \rangle$, wobei in den Beispielen a) und b) einmal das kanonische Skalarprodukt im \mathbb{R}^2 zu verwenden ist und dann die in Beispiel 64 angegebenen Skalarprodukte. In den Beispielen c) – f) ist das kanonische Skalarprodukt im \mathbb{R}^n zu verwenden.

66. Für die in Beispiel 65 angegebenen Vektoren u berechne man $\|u\|$ und gebe einen Einheitsvektor in Richtung von u an. Dabei ist jeweils die Betragssnorm, die Euklidische Norm und die Maximumsnorm zu verwenden.
67. Für die in Beispiel 65 a) und b) angegebenen Vektoren v berechne man $\|v\|$, wobei die durch die in Beispiel 64 definierten Skalarprodukte induzierten Normen zu verwenden sind.
- a) Vektor v von Bsp. 65 a) mit den Skalarprodukten von Bsp. 64.
 b) Vektor v von Bsp. 65 b) mit den Skalarprodukten von Bsp. 64.
68. Man berechne den Abstand $d(u, v)$ für die in Beispiel 65 angegebenen Vektoren, wobei jeweils die Betragssnorm, die Euklidische Norm bzw. die Maximumsnorm zu verwenden sind.
69. Für die in Beispiel 65 a) und b) angegebenen Vektoren u und v berechne man den Abstand $d(u, v)$, wobei die durch die in Beispiel 64 definierten Skalarprodukte induzierten Normen zu verwenden sind.
- a) Vektoren u, v von Bsp. 65 a) mit den Skalarprodukten von Bsp. 64.
 b) Vektoren u, v von Bsp. 65 b) mit den Skalarprodukten von Bsp. 64.
70. Man untersuche, welche Menge von Vektoren bzgl. des kanonischen Skalarprodukts orthogonal ist.

$$a) \left\{ \begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \right\} \quad b) \left\{ \begin{pmatrix} 4 \\ 2 \\ -5 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} \right\}$$

$$c) \left\{ \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \\ 4 \end{pmatrix} \right\} \quad d) \left\{ \begin{pmatrix} 5 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix} \right\}$$

$$e) \left\{ \begin{pmatrix} 2 \\ 3 \\ -1 \\ 4 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -4 \\ -6 \\ 2 \\ 7 \end{pmatrix} \right\} \quad f) \left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \\ 2 \end{pmatrix} \right\}$$

71. Man bestimme den Winkel zwischen den zum kleinsten bzw. zum größten Eigenwert der Matrix A gehörigen Eigenvektoren für

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

72. Man zeige, daß die Vektoren (v_i) eine orthogonale Basis des \mathbb{R}^2 bzw. des \mathbb{R}^3 bilden. Dann gebe man den Vektor w als Linearkombination dieser Basisvektoren an und bestimme den Koordinatenvektor $c_{\mathcal{B}}(w)$ für $\mathcal{B} = (v_1, v_2)$ bzw. $\mathcal{B} = (v_1, v_2, v_3)$

$$a) v_1 = \begin{pmatrix} 4 \\ -2 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}; \quad w = \begin{pmatrix} 1 \\ -3 \end{pmatrix}$$

$$b) v_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -6 \\ 2 \end{pmatrix}; \quad w = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$c) v_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}; \quad w = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$d) v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}; \quad w = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

73. Man untersuche, ob die folgenden Vektoren eine orthonormale Menge von Vektoren bilden. Sind sie nicht orthonormal, so bilde man daraus eine orthonormale Menge.

$$a) \begin{pmatrix} \frac{3}{5} \\ \frac{4}{5} \end{pmatrix}, \quad \begin{pmatrix} -\frac{4}{5} \\ \frac{3}{5} \end{pmatrix} \quad b) \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \quad \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix} \quad c) \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{pmatrix}, \quad \begin{pmatrix} \frac{2}{3} \\ -\frac{1}{3} \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 2 \\ -\frac{5}{2} \end{pmatrix}$$

d) Die Vektoren v_1 und v_2 von Bsp. 72 a) und b).

e) Die Vektoren v_1, v_2 und v_3 von Bsp. 72 c) und d).

$$f) \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \quad \begin{pmatrix} 0 \\ \frac{1}{3} \\ \frac{2}{3} \\ \frac{1}{3} \end{pmatrix}, \quad \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{6} \\ \frac{1}{6} \\ -\frac{1}{6} \end{pmatrix} \quad g) \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}, \quad \begin{pmatrix} 0 \\ \sqrt{6}/3 \\ 1/\sqrt{6} \end{pmatrix}, \quad \begin{pmatrix} \sqrt{3}/2 \\ -\sqrt{3}/6 \\ \sqrt{3}/6 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 1/\sqrt{2} \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ -\sqrt{3}/6 \end{pmatrix}$$

74. Die folgenden Vektoren bilden eine Basis des \mathbb{R}^2 , des \mathbb{R}^3 , des \mathbb{R}^4 bzw. eines Teilraumes davon. Mit dem Verfahren von Gram-Schmidt ermittle man daraus eine orthonormale Basis.

a) $\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}$; b) $\begin{pmatrix} 3 \\ -3 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix}$; c) $\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix}$;

d) $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$; e) $\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \\ 2 \end{pmatrix}$; f) $\begin{pmatrix} 2 \\ -1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \\ 0 \\ 4 \end{pmatrix}$

75. Für $v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, w = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$ seien die Skalarprodukte

$$(S1) \quad \langle v, w \rangle = 2x_1y_1 + 3x_2y_2$$

und

$$(S2) \quad \langle v, w \rangle = 4x_1y_1 - 2x_1y_2 - 2x_2y_1 + 7x_2y_2$$

vorgelegt.

- a) Für $v = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$ und $w = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$ berechne man

$$\langle u, v \rangle, \|u\|, d(u, v)$$

unter Verwendung der beiden Skalarprodukte.

- b) Man bestimme in beiden Fällen einen Vektor $u \in \mathbb{R}^2, u \neq 0$, so daß u orthogonal zu v ist.

76. Im \mathbb{P}_2 sei das Skalarprodukt für $p(t) = a_0 + a_1t + a_2t^2$ und $q(t) = b_0 + b_1t + b_2t^2$ wie folgt definiert

$$\langle p(t), q(t) \rangle = a_0b_0 + a_1b_1 + a_2b_2$$

- a) Für $p(t) = 2 - 3t + t^2, q(t) = 1 - 3t^2$ berechne man

$$\langle p(t), q(t) \rangle, \|p(t)\|, d(p(t), q(t))$$

- b) Man bestimme einen Vektor (= Polynom), der orthogonal zu $p(t)$ ist.

77. Man führe das Programm von Beispiel 76 für das Skalarprodukt

$$\langle p(t), q(t) \rangle = \int_0^1 p(t) \cdot q(t) dt$$

durch.

78. Im $\mathcal{C}^2[0, 2\pi]$ sei das Skalarprodukt

$$\langle f, g \rangle = \int_0^{2\pi} f(t) \cdot g(t) dt$$

definiert.

Für $f(t) = \sin t, g(t) = \sin t + \cos t$ berechne man $\langle f, g \rangle$, $\|f\|$, $d(f, g)$ und bestimme einen Vektor $h \in \mathcal{C}^2[0, 2\pi], h \neq 0$, der orthogonal zu f ist.

79. Mit dem Verfahren von Gram-Schmidt berechne man eine orthonomiert Basis des von der Basis \mathcal{B} aufgespannten Vektorraumes. Dabei ist das angegebene Skalarprodukt zu verwenden.

a) $V = \mathbb{R}^2, \quad \mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$ mit dem Skalarprodukt (S1) von Beispiel 75.

b) $V = \mathbb{R}^2, \quad \mathcal{B} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$ mit dem Skalarprodukt (S2) von Beispiel 75.

c) $V = \mathbb{P}_2, \quad \mathcal{B} = \{1, 1+t, 1+t+t^2\}$ mit dem Skalarprodukt von Beispiel 77.

d) $V = \mathbb{P}_2, \quad \mathcal{B} = \{1, 1+t, 1+t+t^2\}$ mit dem Skalarprodukt von Beispiel 78.

80. Man bestimme die QR -Zerlegung der folgenden Matrizen

a) $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad$ b) $\begin{pmatrix} 2 & 8 & 2 \\ 1 & 7 & -1 \\ -2 & -2 & 1 \end{pmatrix}, \quad$ c) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \\ -1 & 1 & 0 \\ 1 & 5 & 1 \end{pmatrix}$

d) $\begin{pmatrix} 1 & 3 \\ 2 & 4 \\ -1 & -1 \\ 0 & 1 \end{pmatrix}, \quad$ e) $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$

6 Eigenwerte und Eigenvektoren

81. Man bestimme die Eigenwerte und die Eigenvektoren folgender Matrizen

(a) $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad$ (b) $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad$ (c) $A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}, \quad a, b \in \mathbb{R}$

(d) $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -3 & -3 \end{pmatrix} \quad$ (e) $A = \begin{pmatrix} 4 & 0 & 1 \\ 0 & 4 & 1 \\ 1 & 0 & 4 \end{pmatrix} \quad$ (f) $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 10 \end{pmatrix}$

82. Man bestimme die Eigenwerte und die Eigenvektoren der folgenden Matrizen.

$$\begin{array}{llll}
 \text{a)} A = \begin{pmatrix} 0 & 3 \\ 3 & 0 \end{pmatrix} & \text{b)} A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} & \text{c)} A = \begin{pmatrix} -1 & 1 \\ 6 & 0 \end{pmatrix} & \text{d)} A = \begin{pmatrix} 4 & -2 \\ 5 & -7 \end{pmatrix} \\
 \text{e)} A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & -2 \\ 1 & 0 & 1 \end{pmatrix} & \text{f)} A = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 1 & 1 \\ 1 & 2 & 0 \end{pmatrix} & \text{g)} A = \begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix} & \text{h)} A = \begin{pmatrix} 2 & 2 \\ 5 & -1 \end{pmatrix} \\
 \text{i)} A = \begin{pmatrix} 0 & 4 \\ -1 & 0 \end{pmatrix} & \text{j)} A = \begin{pmatrix} 0 & 4 \\ -1 & 5 \end{pmatrix} & \text{k)} A = \begin{pmatrix} 1 & 0 & 2 \\ -1 & 1 & 1 \\ 2 & 0 & 1 \end{pmatrix} & \text{l)} A = \begin{pmatrix} 3 & 1 & -1 \\ 1 & 1 & 1 \\ 4 & 2 & 0 \end{pmatrix} \\
 \text{m)} A = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} & \text{n)} A = \begin{pmatrix} 2 & 4 \\ 6 & 0 \end{pmatrix} & \text{o)} A = \begin{pmatrix} 2 & 5 \\ 0 & 2 \end{pmatrix} & \text{p)} A = \begin{pmatrix} 1 & 2 \\ -2 & 3 \end{pmatrix}
 \end{array}$$

83. Man berechne die Eigenwerte, die Eigenvektoren, eine Basis jedes Eigenraumes und die algebraische und die geometrische Vielfachheit jedes Eigenwerts.

$$\begin{array}{llll}
 \text{a)} A = \begin{pmatrix} 1 & 3 \\ -2 & 6 \end{pmatrix} & \text{b)} A = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} & \text{c)} A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & 3 \end{pmatrix} & \text{d)} A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} \\
 \text{e)} A = \begin{pmatrix} 1 & 2 & 0 \\ -1 & -1 & 1 \\ 0 & 1 & 1 \end{pmatrix} & \text{f)} A = \begin{pmatrix} 1 & 0 & 2 \\ 3 & -1 & 3 \\ 2 & 0 & 1 \end{pmatrix} & \text{g)} A = \begin{pmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ -1 & 0 & 2 \end{pmatrix} \\
 \text{h)} A = \begin{pmatrix} 1 & -1 & -1 \\ 0 & 2 & 0 \\ -1 & -1 & 1 \end{pmatrix} & \text{i)} A = \begin{pmatrix} 3 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 1 & 1 \end{pmatrix} & \text{j)} A = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 4 & 5 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \\
 \text{k)} A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 3 & 0 \\ -2 & 1 & 2 & -1 \end{pmatrix} & \text{l)} A = \begin{pmatrix} 4 & 0 & 1 & 0 \\ 0 & 4 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 3 & 0 \end{pmatrix}
 \end{array}$$

84. Man zeige, dass A und B keine ähnlichen Matrizen sind.

$$\begin{array}{llll}
 \text{a)} A = \begin{pmatrix} 4 & 1 \\ 3 & 1 \end{pmatrix} & B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, & \text{b)} A = \begin{pmatrix} 3 & -1 \\ -5 & 7 \end{pmatrix}, & B = \begin{pmatrix} 2 & 1 \\ -4 & 6 \end{pmatrix} \\
 \text{c)} A = \begin{pmatrix} 2 & 1 & 4 \\ 0 & 2 & 3 \\ 0 & 0 & 4 \end{pmatrix}, & B = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 4 & 0 \\ 2 & 3 & 4 \end{pmatrix} & \text{d)} A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}, & B = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}
 \end{array}$$

85. Man diagonalisiere folgende Matrizen, d.h. man bestimme eine Matrix C so, dass $C^{-1}AC = D$ gilt, wobei D eine Diagonalmatrix darstellt.

$$\text{a) } A = \begin{pmatrix} 5 & -1 \\ 2 & 2 \end{pmatrix} \quad \text{b) } A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad \text{c) } A = \begin{pmatrix} 1 & 3 & 3 \\ 2 & 0 & 2 \\ 3 & 3 & 1 \end{pmatrix}$$

$$\text{d) } A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix} \quad \text{e) } A = \begin{pmatrix} 1 & 1 & -2 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{pmatrix} \quad \text{f) } A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

$$\text{g) } A = \begin{pmatrix} 7 & -2 & -4 \\ 3 & 0 & -2 \\ 6 & -2 & -3 \end{pmatrix} \quad \text{h) } A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

86. Man untersuche, ob die folgenden Matrizen orthogonal sind und berechne gegebenenfalls die Inverse

$$\text{a) } \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{b) } \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \quad \text{c) } \begin{pmatrix} \frac{1}{2} & \frac{1}{3} & \frac{2}{5} \\ \frac{1}{2} & -\frac{1}{3} & \frac{2}{5} \\ -\frac{1}{2} & 0 & \frac{4}{5} \end{pmatrix} \quad \text{d) } \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & 1/\sqrt{6} & -1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{pmatrix}$$

$$\text{e) } \begin{pmatrix} \cos \varphi \sin \varphi & -\cos \varphi & -\sin^2 \varphi \\ \cos^2 \varphi & \sin \varphi & -\cos \varphi \sin \varphi \\ \sin \varphi & 0 & \cos \varphi \end{pmatrix} \quad \text{f) } \begin{pmatrix} 2/3 & 1/3 & 2/3 \\ 1/3 & 2/3 & -2/3 \\ -2/3 & 2/3 & 1/3 \end{pmatrix}$$

$$\text{g) } \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \quad \text{h) } \begin{pmatrix} 1 & 0 & 0 & 1/\sqrt{6} \\ 0 & \frac{2}{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 0 & -\frac{2}{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 0 & \frac{1}{3} & 0 & 1/\sqrt{2} \end{pmatrix}$$

87. Man bestimme die Eigenwerte und ein System paarweise orthogonaler Eigenvektoren von

$$\text{(a) } A = \begin{pmatrix} 3 & -1 & -1 \\ -1 & 3 & -1 \\ -1 & -1 & 3 \end{pmatrix} \quad \text{(b) } A = \begin{pmatrix} 0 & 2 & 4 \\ 1 & 1 & -2 \\ -2 & 0 & 5 \end{pmatrix} \quad \text{(c) } A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 3 & 2 \\ 1 & 2 & 3 \end{pmatrix}$$

88. Man diagonalisiere die folgenden symmetrischen Matrizen durch eine orthogonale Matrix Q , d.h. man bestimme Q so, dass gilt $Q^{-1}AQ = D$, wobei D eine Diagonalmatrix bezeichnet.

$$\text{a) } A = \begin{pmatrix} 4 & 1 \\ 1 & 4 \end{pmatrix} \quad \text{b) } A = \begin{pmatrix} -1 & 3 \\ 3 & -1 \end{pmatrix} \quad \text{c) } A = \begin{pmatrix} 1 & \sqrt{2} \\ \sqrt{2} & 0 \end{pmatrix} \quad \text{d) } A = \begin{pmatrix} 9 & -2 \\ -2 & 6 \end{pmatrix}$$

$$\text{e) } A = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 3 & 1 \end{pmatrix} \quad \text{f) } A = \begin{pmatrix} 2 & 3 & 0 \\ 3 & 2 & 4 \\ 0 & 4 & 2 \end{pmatrix} \quad \text{g) } A = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

h) $A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$ i) $A = \begin{pmatrix} 7 & -2 & 1 \\ -2 & 10 & -2 \\ 1 & -2 & 7 \end{pmatrix}$ j) $A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}, b \neq 0$

k) $A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$ l) $A = \begin{pmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{pmatrix}$ m) $A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 1 & 0 & 0 & 4 \end{pmatrix}$

89. Man bestimme eine Matrix A , deren Eigenwerte 1 und 4 sind und deren Eigenvektoren $v_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$ und $v_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ sind.

90. Man bestimme eine symmetrische 2×2 Matrix mit den Eigenwerten λ_1 und λ_2 und den dazugehörigen Eigenvektoren v_1 und v_2

a) $\lambda_1 = -1, \lambda_2 = 2, v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

b) $\lambda_1 = 3, \lambda_2 = -3, v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, v_2 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$

c) $\lambda_1 = 0, \lambda_2 = 1, v_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}, v_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$

d) $\lambda_1 = 1, \lambda_2 = -1, v_1 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, v_2 = \begin{pmatrix} -2 \\ 3 \end{pmatrix}$

91. Man bestimme eine symmetrische 3×3 Matrix mit den Eigenwerten λ_1, λ_2 und λ_3 und den entsprechenden Eigenvektoren v_1, v_2 und v_3

a) $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 3, v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}$

b) $\lambda_1 = 0, \lambda_2 = 1, \lambda_3 = 2, v_1 = \begin{pmatrix} -1 \\ 1 \\ 2 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$

c) $\lambda_1 = -2, \lambda_2 = 0, \lambda_3 = 2, v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$

d) $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = 4, v_1 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}$

7 Numerische Behandlung von Gleichungssystemen

92. Man ermittle die Cholesky-Zerlegung der folgenden Matrizen und mache die Probe.

$$a) \begin{pmatrix} 1 & 2 \\ 2 & 8 \end{pmatrix}, \quad b) \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 3 & 7 & 26 \end{pmatrix}, \quad c) \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 2 & 2 & 9 \end{pmatrix}, \quad d) \begin{pmatrix} 1 & -1 \\ -1 & 5 \end{pmatrix}$$

93. Man berechne die Norm $\|A\|_2$, $\|A\|_\infty$ und $\|A\|_F$ sowie die entsprechenden Konditionszahlen $\text{cond}(A)_2$, $\text{cond}(A)_\infty$ und $\text{cond}(A)_F$ der folgenden Matrizen

$$\begin{array}{llll} a) A = \begin{pmatrix} 2 & 3 \\ 4 & 1 \end{pmatrix} & b) A = \begin{pmatrix} 0 & -1 \\ -3 & 3 \end{pmatrix} & c) A = \begin{pmatrix} 1 & 5 \\ -2 & -1 \end{pmatrix} & d) A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 1 & 3 \end{pmatrix} \\ e) A = \begin{pmatrix} 0 & -5 & 2 \\ 3 & 1 & -3 \\ -4 & -4 & 3 \end{pmatrix} & f) A = \begin{pmatrix} 4 & -2 & -1 \\ 0 & -1 & 2 \\ 3 & -3 & 0 \end{pmatrix} & g) A = \begin{pmatrix} 3 & 1 \\ 4 & 2 \end{pmatrix} \\ h) A = \begin{pmatrix} 1 & -2 \\ -4 & 6 \end{pmatrix} & i) A = \begin{pmatrix} 1 & 0.99 \\ 1 & 1 \end{pmatrix} & j) A = \begin{pmatrix} 150 & 200 \\ 3001 & 4002 \end{pmatrix} \\ k) A = \begin{pmatrix} 1 & 1 & 1 \\ 5 & 5 & 6 \\ 1 & 0 & 0 \end{pmatrix} & l) A = \begin{pmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{pmatrix} \end{array}$$

94. Mit Hilfe der Jacobi-Iteration bestimme man einen Näherungswert der Lösung der folgenden Gleichungssysteme. Man verwende den Null-Vektor als Startvektor und iteriere solange, bis zwei aufeinander folgende Vektoren sich in allen Komponenten um weniger als 0.0001 unterscheiden. Man vergleiche die Antwort mit der exakten Lösung.

$$\begin{array}{ll} a) \begin{array}{l} 7x_1 - x_2 = 6 \\ x_1 - 5x_2 = -4 \end{array} & b) \begin{array}{l} 2x_1 + x_2 = 5 \\ x_1 - x_2 = 1 \end{array} \\ c) \begin{array}{l} 4.5x_1 - 0.5x_2 = 1 \\ x_1 - 3.5x_2 = -1 \end{array} & d) \begin{array}{l} 20x_1 + x_2 - x_3 = 17 \\ x_1 - 10x_2 + x_3 = 13 \\ -x_1 + x_2 + 10x_3 = 18 \end{array} \\ e) \begin{array}{l} 3x_1 + x_2 = 1 \\ x_1 + 4x_2 + x_3 = 1 \\ x_2 + 3x_3 = 1 \end{array} & f) \begin{array}{l} 3x_1 - x_2 = 1 \\ -x_1 + 3x_2 - x_3 = 0 \\ -x_2 + 3x_3 - x_4 = 1 \\ -x_3 + 3x_4 = 1 \end{array} \end{array}$$

95. Man rechne das Programm von Beispiel 94 mit der Gauß-Seidel-Iteration durch.

96. Für die folgenden Gleichungssysteme berechne man mit der Gauß-Seidel-Iteration die ersten vier Iterationsvektoren, wobei der Null-Vektor als Startvektor dienen soll, um zu zeigen, dass das Verfahren divergiert. Dann vertausche man die Gleichungen so, dass eine

streng diagonal dominante Koeffizientenmatrix entsteht und berechne dann eine Näherungslösung mit einem Fehler < 0.001 .

$$\text{a) } \begin{array}{rcl} x_1 - 2x_2 & = & 3 \\ 3x_1 + 2x_2 & = & 1 \end{array} \quad \text{b) } \begin{array}{rcl} x_1 - 4x_2 + 2x_3 & = & 2 \\ 2x_2 + 4x_3 & = & 1 \\ 6x_1 - x_2 - 2x_3 & = & 1 \end{array}$$

97. Die Koeffizientenmatrix der folgenden Gleichungssysteme ist nicht streng diagonal dominant (s.d.d.) auch durch Umordnen der Gleichungen erhält man keine s.d.d. Matrix. Trotzdem konvergieren hier die Jacobi- und die Gauß-Seidel-Iteration. Man zeige dies für die Gauß-Seidel-Iteration mit dem Null-Vektor als Startvektor und der Berechnung einer Näherungslösung mit dem Fehler < 0.01 .

$$\text{a) } \begin{array}{rcl} -4x_1 + 5x_2 & = & 14 \\ x_1 - 3x_2 & = & -7 \end{array} \quad \text{b) } \begin{array}{rcl} 5x_1 - 2x_2 + 3x_3 & = & -8 \\ x_1 + 4x_2 - 4x_3 & = & 102 \\ -2x_1 - 2x_2 + 4x_3 & = & -90 \end{array}$$

98. Man bestimme die Näherungslösung (im Sinne des kleinsten quadratischen Fehlers) der Gleichungssysteme $A \cdot x = b$ mit Hilfe der Pseudoinversen.

$$\text{a) } A = \begin{pmatrix} 3 & 1 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \text{ b) } A = \begin{pmatrix} 3 & -2 \\ 1 & -2 \\ 2 & -1 \end{pmatrix}, b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \text{ c) } A = \begin{pmatrix} 2 & 1 \\ 2 & 0 \\ 1 & 1 \end{pmatrix}, b = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix}$$

$$\text{d) } A = \begin{pmatrix} 1 & -2 \\ 0 & -3 \\ 2 & 5 \\ 3 & 0 \end{pmatrix}, b = \begin{pmatrix} 4 \\ 1 \\ -2 \\ 4 \end{pmatrix}, \text{ e) } A = \begin{pmatrix} 2 & 0 \\ 1 & -1 \\ 3 & 1 \\ -1 & 2 \end{pmatrix}, b = \begin{pmatrix} 5 \\ 1 \\ -1 \\ 3 \end{pmatrix}$$

99. Man berechne die Näherungslösung der Gleichungssysteme $Ax = b$ mit Hilfe der QR-Zerlegung der Koeffizientenmatrix A wobei A und b wie in Beispiel 98 angegeben zu wählen sind.

8 Numerische Bestimmung von Eigenwerten

100. Man approximiere den dominanten Eigenwert von

$$A = \begin{pmatrix} 3 & -8 & 4 \\ -1 & -2 & -2 \\ 0 & 6 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & -1 & 1 \\ -1 & 3 & -2 \\ 1 & -2 & 3 \end{pmatrix}$$

101. Man approximiere alle Eigenwerte von

$$A = \begin{pmatrix} 5 & 4 & 1 & 1 \\ 4 & 5 & 1 & 1 \\ 1 & 1 & 4 & 2 \\ 1 & 1 & 2 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 6 & 4 & 4 & 1 \\ 4 & 6 & 1 & 4 \\ 4 & 1 & 6 & 4 \\ 1 & 4 & 4 & 6 \end{pmatrix}, \quad C = \begin{pmatrix} 4 & -5 & 0 & 3 \\ 0 & 4 & -3 & -5 \\ 5 & -3 & 4 & 0 \\ 3 & 0 & 5 & 4 \end{pmatrix}$$

102. Vorgelegt sei die Matrix A und ein mit der Potenzmethode gewonnener iterierter Vektor x_5 . Damit approximiere man einen dominanten Eigenvektor auf 3 Dezimalen genau, dessen erste Komponente 1 ist, und den dominanten Eigenwert. Man vergleiche das Ergebnis mit den exakten Resultaten.

$$\text{a) } A = \begin{pmatrix} 1 & 2 \\ 5 & 4 \end{pmatrix}, \quad x_5 = \begin{pmatrix} 4443 \\ 11109 \end{pmatrix}, \quad \text{b) } A = \begin{pmatrix} 7 & 4 \\ -3 & -1 \end{pmatrix}, \quad x_5 = \begin{pmatrix} 7811 \\ -3904 \end{pmatrix}$$

$$\text{c) } A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad x_5 = \begin{pmatrix} 144 \\ 89 \end{pmatrix} \quad \text{d) } A = \begin{pmatrix} 1.5 & 0.5 \\ 2.0 & 3.0 \end{pmatrix}, \quad x_5 = \begin{pmatrix} 60.625 \\ 239.500 \end{pmatrix}$$

$$\text{e) } A = \begin{pmatrix} 2 & -3 \\ -3 & 10 \end{pmatrix}, \quad x_5 = \begin{pmatrix} -3.667 \\ 11.001 \end{pmatrix}$$

103. Mit der Potenzmethode approximiere man den dominanten Eigenwert und den zugehörigen Eigenvektor. Der Startvektor x_0 ist vorgelegt, ebenso die Anzahl k der Iterationen.

$$\text{a) } A = \begin{pmatrix} 14 & 12 \\ 5 & 3 \end{pmatrix}, \quad x_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad k = 5$$

$$\text{b) } A = \begin{pmatrix} -6 & 4 \\ 8 & -2 \end{pmatrix}, \quad x_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad k = 6$$

$$\text{c) } A = \begin{pmatrix} 7 & 2 \\ 2 & 3 \end{pmatrix}, \quad x_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad k = 6$$

$$\text{d) } A = \begin{pmatrix} 3.5 & 1.5 \\ 1.5 & -0.5 \end{pmatrix}, \quad x_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad k = 6$$

$$\text{e) } A = \begin{pmatrix} 9 & 4 & 8 \\ 4 & 15 & -4 \\ 8 & -4 & 9 \end{pmatrix}, \quad x_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad k = 5$$

$$\text{f) } A = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 3 \end{pmatrix}, \quad x_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad k = 6$$

$$\text{g) } A = \begin{pmatrix} 4 & 1 & 3 \\ 0 & 2 & 0 \\ 1 & 1 & 2 \end{pmatrix}, \quad x_0 = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \quad k = 5$$

$$\text{h) } A = \begin{pmatrix} 12 & -6 & -6 \\ 2 & 0 & -2 \\ -6 & 6 & 12 \end{pmatrix}, \quad x_0 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \quad k = 5$$

104. Man bestimme die einzelnen Rayleigh-Quotienten $\mu_s, s = 1, 2, \dots, k$, für die in Beispiel 103 angegebenen Matrizen.

105. Man bestimme die Gerschgorin-Kreisscheiben für

$$A = \begin{pmatrix} -5 & 0 & 1 \\ 1 & 1 & -2 \\ -3 & 0 & 8 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & 1 & -1 \\ -2 & 0 & 1 \\ 1 & 0 & -4 \end{pmatrix}, \quad C = \begin{pmatrix} 3 & 1 & 0 & 2i \\ 1 & 3 & -2i & 0 \\ 0 & 2i & 1 & 1 \\ -2i & 0 & 1 & 1 \end{pmatrix}$$

106. Man verwende Gerschgorin-Kreisscheiben, um die Lage der Eigenwerte der folgenden Matrizen anzugeben.

a) $\begin{pmatrix} 1 & -10^{-5} & 2 \cdot 10^{-5} \\ 4 \cdot 10^{-5} & 0.5 & -3 \cdot 10^{-5} \\ -10^{-5} & 3 \cdot 10^{-5} & 0.1 \end{pmatrix}$ b) $\begin{pmatrix} 2 & 0.1 & 0.2 \\ -0.1 & 2 & -0.1 \\ 1 & -1 & 10 \end{pmatrix}$

c) $\begin{pmatrix} 6 & 1 & 1 & -2 \\ 1 & 14 & 1 & 1 \\ 2 & 2 & -9 & -1 \\ 1 & -1 & 1 & 29 \end{pmatrix}$ d) $\begin{pmatrix} 3 & -\frac{1}{2} & -\frac{1}{3} & 0 \\ 0 & 6 & 1 & 0 \\ \frac{1}{3} & -\frac{1}{3} & 5 & \frac{1}{3} \\ -\frac{1}{2} & \frac{1}{4} & -\frac{1}{4} & 4 \end{pmatrix}$

9 Interpolation und Polynomapproximation

107. Approximieren Sie die folgenden Werte mit geeigneten Lagrangeschen Interpolationspolynomen vom Grad eins, zwei und drei. Dabei wähle man die Stützstellen jeweils geeignet aus. Stellen Sie die Polynome graphisch dar.

- (a) $f(8.4)$, wenn $f(8) = 16.63553$, $f(8.1) = 17.61549$,
 $f(8.3) = 17.56492$, $f(8.6) = 18.50515$, $f(8.7) = 18.82091$
- (b) $f(-\frac{1}{3})$, wenn $f(-1) = 0.1$, $f(-0.75) = -0.0718125$,
 $f(-0.5) = -0.02475000$, $f(-0.25) = 0.33493750$, $f(0) = 1.10100000$
- (c) $f(0.25)$, wenn $f(0) = -1$, $f(0.1) = -0.62049958$,
 $f(0.2) = -0.28398668$, $f(0.3) = 0.00660095$, $f(0.4) = 0.24842440$
- (d) $f(0.9)$, wenn $f(0.5) = -0.34409873$, $f(0.6) = -0.17694460$,
 $f(0.7) = 0.01375227$, $f(0.8) = 0.22363362$, $f(1.0) = 0.65809197$
- (e) $f(\pi)$, wenn $f(2.9) = -4.827866$, $f(3.0) = -4.240058$,
 $f(3.1) = -3.496909$, $f(3.2) = -2.596792$, $f(3.4) = -0.3330587$
- (f) $f(1.25)$, wenn $f(1.1) = 1.964760$, $f(1.2) = 2.572152$,
 $f(1.3) = 3.602102$, $f(1.4) = 5.797884$, $f(1.5) = 14.10142$
- (g) $f(1.15)$, wenn $f(1) = 1.684370$, $f(1.1) = 1.949477$,
 $f(1.2) = 2.199796$, $f(1.3) = 2.439189$, $f(1.4) = 2.670324$
- (h) $f(4.1)$, wenn $f(3.6) = 1.16164956$, $f(3.8) = 0.80201036$,
 $f(4) = 0.30663842$, $f(4.2) = -0.35916618$, $f(4.4) = -1.23926000$

(i) $f(0.2)$, wenn $f(-0.6) = -4.30789$, $f(-0.3) = -2.48886$,
 $f(0) = -1$, $f(0.3) = 0.666061$, $f(0.6) = 2.97862$

(j) $f(0.5)$, wenn $f(0) = 1$, $f(0.2) = 0.935897$,
 $f(0.4) = 0.802096$, $f(0.6) = 0.667463$, $f(0.8) = 0.57352$

108. Konstruieren Sie das Lagrangesche Interpolationspolynom $P_n(x)$ für die folgenden Funktionen. Zeichnen Sie den Graph von $f(x)$ und von $P_n(x)$.

(a) $f(x) = e^{2x} \cos 3x$, $x_0 = 0$, $x_1 = 0.3$, $x_2 = 0.6$, $n = 2$

(b) $f(x) = \sin(\ln x)$, $x_0 = 2.0$, $x_1 = 2.4$, $x_2 = 2.6$, $n = 2$

(c) $f(x) = \ln x$, $x_0 = 1$, $x_1 = 1.1$, $x_2 = 1.3$, $x_3 = 1.4$, $n = 3$

(d) $f(x) = \cos x + \sin x$, $x_0 = 0$, $x_1 = 0.25$, $x_2 = 0.5$, $x_3 = 1.0$, $n = 3$

(e) $f(x) = \frac{1}{1+x^2}$, $x_0 = -0.3$, $x_1 = -0.1$, $x_2 = 0.1$, $n = 2$

(f) $f(x) = \frac{1-x^2}{1+x^2}$, $x_0 = -0.1$, $x_1 = 0.1$, $x_2 = 0.3$, $n = 2$

(g) $f(x) = \frac{\sin x}{1+x^2}$, $x_0 = -0.5$, $x_1 = 0.5$, $x_2 = 1$, $n = 2$

(h) $f(x) = \frac{1+x}{\cos x}$, $x_0 = -1$, $x_1 = 0$, $x_2 = 1$, $n = 2$

(i) $f(x) = x^2 \ln x$, $x_0 = 0.1$, $x_1 = 0.5$, $x_2 = 1$, $n = 2$

(j) $f(x) = e^x \sin 2x$, $x_0 = 0$, $x_1 = 0.5$, $x_2 = 1$, $n = 2$

109. Konstruieren Sie jeweils die kubischen Spline-Interpolierenden für die folgenden Daten und fertigen Sei einen Graph der Spline-Funktion sowie der zugehörigen, in Beispiel 110 gegebenen Funktion $f(x)$.

(a)
$$\begin{array}{c|cccc} x & 8.0 & 8.1 & 8.3 & 8.6 \\ \hline f(x) & 16.636 & 16.944 & 17.565 & 18.505 \end{array}$$

(b)
$$\begin{array}{c|cccc} x & 0.6 & 0.8 & 1.0 & 1.2 \\ \hline f(x) & -0.1769 & 0.2236 & 0.6581 & 0.9687 \end{array}$$

(c)
$$\begin{array}{c|cccc} x & -0.5 & -0.25 & 0 & 0.25 \\ \hline f(x) & -0.02475 & 0.3349 & 1.1010 & 2.3672 \end{array}$$

(d)
$$\begin{array}{c|cccc} x & 3 & 3.1 & 3.2 & 3.3 \\ \hline f(x) & -4.2400 & -3.4969 & -2.5968 & -1.5408 \end{array}$$

(e)
$$\begin{array}{c|cccc} x & 0.1 & 0.2 & 0.3 & 0.4 \\ \hline f(x) & -0.6205 & -0.2840 & 0.0066 & 0.2484 \end{array}$$

(f)
$$\begin{array}{c|cccc} x & 1.2 & 1.3 & 1.4 & 1.5 \\ \hline f(x) & 2.5722 & 3.6021 & 5.7978 & 14.1014 \end{array}$$

(g)
$$\begin{array}{c|cccc} x & 1 & 1.1 & 1.2 & 1.3 \\ \hline f(x) & 1.6844 & 1.9495 & 2.1998 & 2.4392 \end{array}$$

(h)
$$\begin{array}{c|cccc} x & 3.6 & 3.8 & 4.0 & 4.2 \\ \hline f(x) & 1.1616 & 0.8020 & 0.3066 & -0.3592 \end{array}$$

(i)
$$\begin{array}{c|cccc} x & 1 & 1.2 & 1.4 & 1.6 \\ \hline f(x) & 2.5403 & 2.1548 & 1.5180 & 0.6333 \end{array}$$

(j)	x	2	2.3	2.6	2.9
	$f(x)$	-12.75	-8.645	-2.594	6.273

110. Die Daten in Beispiel 109 wurden mit den folgenden Funktionen erzeugt. Approximieren Sie $f(x_0)$ und $f'(x_0)$ mit den in Beispiel 109 konstruierten kubischen Splines für den gegebenen Wert von x_0 und berechnen Sie den tatsächlichen Fehler.

- (a) $f(x) = x \ln x, \quad x_0 = 8.4$
- (b) $f(x) = \sin(e^x - 2), \quad x_0 = 0.9$
- (c) $f(x) = x^3 + 4.001x^2 + 4.002x + 1.101, \quad x_0 = -\frac{1}{3}$
- (d) $f(x) = x \cos x - x^2 \sin x, \quad x_0 = \pi$
- (e) $f(x) = x \cos x - 2x^2 + 3x - 1, \quad x_0 = 0.25$
- (f) $f(x) = \tan x, \quad x_0 = \pi/2$
- (g) $f(x) = \ln(e^{2x} - 2), \quad x_0 = 1.15$
- (h) $f(x) = x - (\ln x)^x, \quad x_0 = 4.1$
- (i) $f(x) = 1 + 3x - 2x^2 + x \cos x, \quad x_0 = 1.1$
- (j) $f(x) = x \sinh x - 20, \quad x_0 = 2.5$

10 Approximationstheorie

111. Man bestimme ein approximierendes Polynom ersten und zweiten Grades (im Sinne der kleinsten Fehlerquadrate) für folgende Punkte. Plotten Sie die Punkte (x_i, y_i) sowie die Polynome.

(a)	x_i	1.0	1.1	1.3	1.5	1.9	2.1	
	y_i	1.84	1.96	2.21	2.45	2.94	3.18	
(b)	x_i	0	0.15	0.31	0.5	0.6	0.75	
	y_i	1.0	1.004	1.031	1.117	1.223	1.422	
(c)	x_i	4.0	4.5	5.1	5.9	6.3	7.1	
	y_i	102.56	130.11	167.53	224.87	256.73	326.72	
(d)	x_i	0.2	0.3	0.6	0.9	1.1	1.3	1.6
	y_i	0.050446	0.098426	0.33277	0.72660	1.0972	1.5697	2.5015
(e)	x_i	-3	-2	-1	0	1	2	3
	y_i	85	80	71	55	31	0	-22

112. Man bestimme ein approximierendes Polynom dritten Grades (im Sinne der kleinsten Fehlerquadrate) für die in Beispiel 111 angeführten Daten. Plotten Sie die Punkte (x_i, y_i) sowie die Polynome.

113. Gegeben sind die Daten

(a)	$\begin{array}{c ccccc} x_i & 0 & 1 & 2 & 3 & 4 \\ \hline y_i & 6 & 12 & 30 & 80 & 140 \end{array}$
-----	---

(b)	$\begin{array}{c ccccccc} x_i & -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ \hline y_i & 3 & 4 & 5 & 8 & 17 & 40 & 103 \end{array}$
-----	---

Man ermittle eine Funktion von der Form $f(x) = ae^x + b$, die diese Daten bestmöglich im Sinne der kleinsten Fehlerquadrate approximiert. Fertigen Sie eine graphische Darstellung der Punkte (x_i, y_i) und der gefundenen Funktion $f(x)$ an.

114. Man bestimme die **diskrete Approximation** der folgenden Funktionen durch ein trigonometrisches Polynom $S_n(x)$, wobei $2m$ die Anzahl der Teilintervalle angibt, in die das vorgelegte Intervall zu zerlegen ist. Plotten Sie die Funktion $f(x)$ und das Polynom $S_n(x)$.

- (a) $f(x) = |x| \sin x$, $-\pi < x < \pi$, $n = 4, m = 6$
- (b) $f(x) = x^2$, $-\pi < x < \pi$, $n = 4, m = 6$
- (c) $f(x) = x \cos x$, $-\pi/2 < x < \pi/2$, $n = 4, m = 6$
- (d) $f(x) = x + \sin x$, $-\pi < x < \pi$, $n = 4, m = 6$
- (e) $f(x) = 1 + x$, $-\pi < x < \pi$, $n = 4, m = 6$
- (f) $f(x) = e^x$, $-\pi < x < \pi$, $n = 4, m = 4$
- (g) $f(x) = 1 + x + x^2$, $-\pi < x < \pi$, $n = 4, m = 4$
- (h) $f(x) = x(\pi - x)$, $0 < x < \pi$, $n = 3, m = 4$
- (i) $f(x) = \frac{1}{2} - \frac{\pi}{4} \sin x$, $0 < x < \pi$, $n = 3, m = 4$
- (j) $f(x) = \frac{\pi}{4} \text{sign}(x)$, $-\pi < x < \pi$, $n = 8, m = 8$
- (k) $f(x) = x^2 + x$, $-1 < x < 1$, $n = 4, m = 4$
- (l) $f(x) = x^3$, $-2 < x < 2$, $n = 4, m = 4$
- (m) $f(x) = e^{-x}$, $0 < x < 1$, $n = 3, m = 4$
- (n) $f(x) = \cosh x$, $-1 < x < 1$, $n = 3, m = 4$

115. Berechnen Sie mit der schnellen Fouriertransformation das trigonometrische Interpolationspolynom vierten Grades auf $[-\pi, \pi]$ für folgende Funktionen und stellen Sie $f(x)$ und das zugehörige Interpolationspolynom graphisch dar. Dabei ist das Intervall in $2m = 8$ Teilintervalle zu zerlegen.

- (a) $f(x) = \pi(x - \pi)$
- (b) $f(x) = |x|$
- (c) $f(x) = \cos \pi x - 2 \sin \pi x$
- (d) $f(x) = x \cos(x^2) + e^x \cos(e^x)$
- (e) $f(x) = x^2 \cos x$
- (f) $f(x) = \text{sign}(x)$

$$(g) \quad f(x) = x(1 + x^2)$$

$$(h) \quad f(x) = x(1 - x^2)$$

$$(i) \quad f(x) = \begin{cases} 2x + 2\pi & \text{für } -\pi \leq x \leq 0 \\ -2x + 2\pi & \text{für } 0 \leq x \leq \pi \end{cases}$$

$$(j) \quad f(x) = \begin{cases} x + \pi & \text{für } -\pi \leq x \leq 0 \\ x & \text{für } 0 \leq x \leq \pi \end{cases}$$

