

Lineare Algebra für Naturwissenschaften

Thomas P. Wihler

Mathematisches Institut
Universität Bern

Frühling 2009

U
ni
B
erlin

Inhaltsverzeichnis

Vorwort	6
1 Lineare Gleichungssysteme	7
1.1 Gauss-Verfahren	7
1.2 Lineare Gleichungssysteme und Matrizen	15
1.3 Anwendungen	17
1.3.1 Interpolation von Daten	17
1.3.2 Mischungen und Konzentrationen	18
1.3.3 Ebene und räumliche Geometrie	19
1.4 Reguläre und singuläre Matrizen, Rang einer Matrix	20
1.5 Determinanten	26
1.6 MATLAB und OCTAVE	28
1.7 Überbestimmte und unterbestimmte lineare Gleichungssysteme	30
1.8 Übungsaufgaben	34
2 Vektorräume	39
2.1 Der Vektorraum \mathbb{R}^n	39
2.1.1 Grundoperationen für Vektoren	41
2.1.2 Linearkombinationen	42
2.2 Euklidische Norm und Skalarprodukt	44
2.2.1 Längenmessung mit Normen in \mathbb{R}^n	44
2.2.2 Skalarprodukt	46
2.2.3 Projektionseigenschaft	49
2.3 Unterräume von \mathbb{R}^n	51
2.3.1 Beispiele und Definition	51
2.3.2 Unterräume und Lösungen von homogenen Gleichungssystemen	55
2.3.3 Linearkombinationen in Unterräumen	57
2.4 Lineare Unabhängigkeit	59
2.5 Basen und Dimension	63
2.6 Koordinaten	68
2.7 Anwendung: Composition Space für Mineralien	70
2.8 Übungsaufgaben	72
3 Lineare Abbildungen	77
3.1 Vektorfunktionen	77

INHALTSVERZEICHNIS

3.2	Definition und Beispiele	79
3.3	Lineare Abbildungen und Matrizen	82
3.4	Verknüpfte lineare Abbildungen	87
3.5	Umkehrung von linearen Abbildungen	91
3.6	Übungsaufgaben	94
4	Kleinste Quadrate und diskrete Fouriertransformation	97
4.1	Orthogonalprojektionen und kleinste Quadrate	99
4.2	Diskrete Fouriertransformation und Analyse von periodischen Daten	106
4.3	Übungsaufgaben	113
5	Eigenwertprobleme	117
5.1	Iterative Prozesse	117
5.2	Eigenwerte und Eigenvektoren	123
5.3	Berechnung von Eigenwerten und Eigenvektoren in der Praxis	127
5.4	Weitere Anwendungen	132
5.4.1	Markovmodelle in Ökologie und Wirtschaft	132
5.4.2	Leslie-Modelle in der Biologie	136
5.5	Übungsaufgaben	140
	Ergänzende Literatur	144
	Abbildungsverzeichnis	145
	Index	147

Vorwort

Dieser Kurs ist eine *Einführung in die lineare Algebra für Studentinnen und Studenten der Naturwissenschaften*. Er wurde als Teil der Mathematikvorlesungen im Grundstudium der Chemie, Biochemie, Pharmazie, Erdwissenschaften und Geographie an der Universität Bern entwickelt.

Neben den theoretischen Ausführungen werden die einzelnen Themen durch die Besprechung von Anwendungen aus den entsprechenden naturwissenschaftlichen Disziplinen — sowohl in der Vorlesung als auch in den Übungen — ergänzt und bereichert. Es soll dabei die praktische Nützlichkeit der Mathematik als Werkzeug im Alltag der Naturwissenschaftlerinnen und Naturwissenschaftler deutlich gemacht werden. Ferner werden wir uns an verschiedenen Stellen mit einigen einfachen numerischen Berechnungen und Aspekten unter Einbezug der Programme MATLAB und OCTAVE befassen.

Dieses Skript ist ausschliesslich für die Studentinnen und Studenten der Universität Bern gedacht. Jede Vervielfältigung ist untersagt.

Anregungen und Hinweise zu Tippfehlern sind herzlich willkommen und können direkt an wihler@math.unibe.ch geschickt werden.

Thomas P. Wihler
20. August 2009

U
ni
B
erlin

1

Lineare Gleichungssysteme

1.1 Gauss-Verfahren

Die einfachsten elektrischen Kreise enthalten typischerweise zwei verschiedene Komponenten:

- Stromquellen (z. B. Batterien):



- Widerstände (z. B. Glühlampen):



Stromquellen erzeugen Strom, Widerstände dämpfen die Stromstärke. Im Zusammenhang mit elektrischen Kreisen sind die folgenden drei physikalischen Größen besonders wichtig:

- Spannung V , gemessen in Volt V ;
- Widerstand R , gemessen in Ohm Ω ;
- Stromstärke I , gemessen in Ampère A .

Für den Stromfluss in einem elektrischen Kreis gelten die folgenden Gesetze:

- Ohm's Gesetz: Der Spannungsabfall durch einen Widerstand ist für gewisse leitende Materialien gegeben durch $V = I \cdot R$.
- 1. Kirchhoff'sches Gesetz: Die Summe der zufließenden Ströme in einem Knotenpunkt ist gleich der Summe der abfließenden Ströme.
- 2. Kirchhoff'sches Gesetz: Die Summe aller Spannungsunterschiede in einem geschlossenen (Teil-) Kreislauf ist gleich Null.

Betrachten wir als Beispiel den Stromkreis in Abbildung 1.1. Gesucht sind die drei Stromstärken I_1 , I_2 , I_3 . Um diese Größen zu bestimmen, stellen wir zunächst drei geeignete Gleichungen auf.

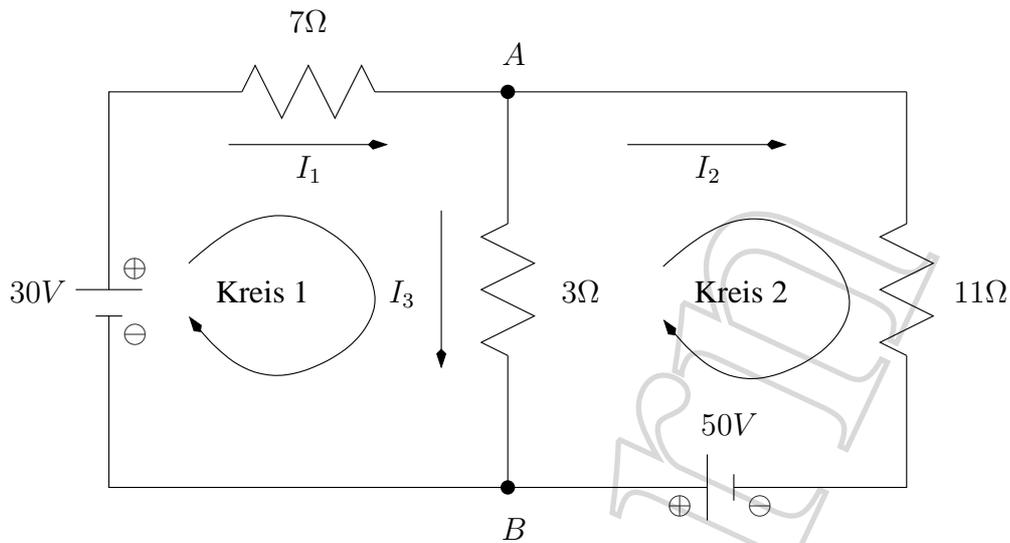


Abbildung 1.1: Beispiel eines elektrischen Stromkreises.

1. Nach dem 1. Kirchhoff'schen Gesetz, angewandt auf Punkt A , gilt:

$$I_1 = I_2 + I_3,$$

oder

$$I_1 - I_2 - I_3 = 0.$$

Für den Punkt B ergibt sich genau dieselbe Gleichung.

2. Betrachten wir Kreis 1: Die beiden Widerstände ergeben mit dem Ohm'schen Gesetz einen Spannungsverlust von $7I_1 + 3I_3$. Nach dem 2. Kirchhoff'schen Gesetz muss dies genau gleich der Spannung sein, welche von der Spannungsquelle geliefert wird, also:

$$7I_1 + 3I_3 = 30.$$

3. Genau gleich erhalten wir für Kreis 2:

$$11I_2 - 3I_3 = 50.$$

Fassen wir die so gefundenen drei Gleichungen zusammen, so ergibt sich das folgende *lineare Gleichungssystem*:

$$\begin{array}{l} (I) \quad I_1 - I_2 - I_3 = 0 \\ (II) \quad 7I_1 \quad \quad + 3I_3 = 30 \\ (III) \quad \quad \quad 11I_2 - 3I_3 = 50 \end{array} \quad (1.1)$$

Um diese Gleichungen nach den Unbekannten I_1, I_2, I_3 aufzulösen, bringen wir das System in eine einfachere Form. Dabei beachten wir:

- Multiplizieren einer Gleichung mit einer festen Zahl (ungleich 0) ändert den Inhalt der Gleichung und damit die Lösung des Systems nicht.
- Addieren oder Subtrahieren einer Gleichung zu/von einer anderen Gleichung ändert die Lösung des Systems ebenfalls nicht.

Kommen wir auf das System (1.1) zurück. Wir lassen die erste Gleichung (I) unberührt und betrachten zunächst die zweite Gleichung. Unser Ziel ist es, die Unbekannte I_1 zu eliminieren. Zu diesem Zweck subtrahieren wir 7-mal die erste Gleichung von der zweiten Gleichung:

$$\begin{array}{r}
 (II) - 7 \cdot (I) : \quad 7I_1 \quad \quad \quad + \quad 3I_3 = 30 \\
 7(I_1 - I_2 - I_3) = 0 \quad \ominus \\
 \hline
 7I_1 \quad \quad \quad + \quad 3I_3 = 30 \\
 7I_1 - 7I_2 - 7I_3 = 0 \quad \ominus \\
 \hline
 \quad \quad + 7I_2 + 10I_3 = 30
 \end{array}$$

Das System (1.1) hat jetzt die neue Form

$$\begin{array}{l}
 (I) \quad I_1 - I_2 - I_3 = 0 \\
 (II) \quad \quad 7I_2 + 10I_3 = 30 \\
 (III) \quad \quad 11I_2 - 3I_3 = 50
 \end{array} \quad (1.2)$$

Die Unbekannte I_1 ist nun aus der zweiten Gleichung verschwunden! Die Lösung des Gleichungssystems bleibt aber unverändert. Im nächsten Schritt lassen wir sowohl die Gleichung (I) als auch die Gleichung (II) unberührt und arbeiten in Gleichung (III), wobei wir die Unbekannte I_2 eliminieren wollen. Subtrahieren von $\frac{11}{7}$ -mal die zweite Gleichung von der dritten Gleichung macht dies möglich:

$$\begin{array}{r}
 (III) - \frac{11}{7} \cdot (II) : \quad 11I_2 - 3I_3 = 50 \\
 \frac{11}{7}(7I_2 + 10I_3) = 30 \quad \ominus \\
 \hline
 11I_2 - 3I_3 = 50 \\
 11I_2 + \frac{110}{7}I_3 = \frac{330}{7} \quad \ominus \\
 \hline
 \quad \quad - \frac{131}{7}I_3 = \frac{20}{7}
 \end{array}$$

Damit erhalten wir das System

$$\begin{array}{l}
 (I) \quad I_1 - I_2 - I_3 = 0 \\
 (II) \quad \quad 7I_2 + 10I_3 = 30 \\
 (III) \quad \quad - \frac{131}{7}I_3 = \frac{20}{7}
 \end{array} \quad (1.3)$$

Wir stellen fest: Wir haben das Gleichungssystem (1.1) so umgewandelt, dass

- Gleichung (I) ist unverändert,
- Gleichung (II) enthält nur noch die Unbekannten I_2 und I_3 ,
- Gleichung (III) enthält nur noch die Unbekannte I_3 und kann gelöst werden.

Ähnlich arbeiten wir in der dritten Gleichung und subtrahieren 3-mal die erste Gleichung von der dritten Gleichung:

$$\begin{array}{r}
 (III) - 3 \cdot (I) : \\
 \begin{array}{r}
 (3 \quad 6 \quad -5 \quad | \quad 0) \\
 \underline{3 \cdot (1 \quad 1 \quad 2 \quad | \quad 9) \ominus} \\
 (3 \quad 6 \quad -5 \quad | \quad 0) \\
 \underline{(3 \quad 3 \quad 6 \quad | \quad 27) \ominus} \\
 (0 \quad 3 \quad -11 \quad | \quad -27)
 \end{array}
 \end{array}$$

Jetzt ist das ursprüngliche lineare Gleichungssystem (1.6) so umgewandelt, dass die zweite und dritte Gleichung die Unbekannte x_1 nicht mehr enthalten. Die entsprechenden Koeffizienten sind 0. Wir schreiben dies wie folgt:

$$\begin{array}{l}
 (I) \\
 (II) \\
 (III)
 \end{array}
 \begin{pmatrix}
 1 & 1 & 2 & | & 9 \\
 \boxed{2} & 4 & -3 & | & 1 \\
 \boxed{3} & 6 & -5 & | & 0
 \end{pmatrix}
 \rightsquigarrow
 \begin{array}{l}
 (I) \\
 (II) - 2 \cdot (I) \\
 (III) - 3 \cdot (I)
 \end{array}
 \begin{pmatrix}
 1 & 1 & 2 & | & 9 \\
 0 & 2 & -7 & | & -17 \\
 0 & 3 & -11 & | & -27
 \end{pmatrix}$$

ursprüngliches System transformiertes System

Schritt 2: Um das lineare Gleichungssystem aufzulösen, eliminieren wir noch die Variable x_2 aus der dritten Gleichung des zuletzt erhaltenen (transformierten) Systems (mit Hilfe der Gleichung (II)). Dies geschieht durch Subtrahieren von $\frac{3}{2}$ -mal der zweiten Gleichung von der dritten. Die erste und zweite Gleichung bleiben unberührt. Kurz,

$$\begin{array}{l}
 (I) \\
 (II) \\
 (III)
 \end{array}
 \begin{pmatrix}
 1 & 1 & 2 & | & 9 \\
 0 & 2 & -7 & | & -17 \\
 0 & \boxed{3} & -11 & | & -27
 \end{pmatrix}$$

$$\rightsquigarrow
 \begin{array}{l}
 (I) \\
 (II) \\
 (III) - \frac{3}{2} \cdot (II)
 \end{array}
 \begin{pmatrix}
 1 & 1 & 2 & | & 9 \\
 0 & 2 & -7 & | & -17 \\
 0 & 3 - \frac{3}{2} \cdot (2) & -11 - \frac{3}{2} \cdot (-7) & | & -27 - \frac{3}{2} \cdot (-17)
 \end{pmatrix}$$

$$= \begin{pmatrix}
 1 & 1 & 2 & | & 9 \\
 0 & 2 & -7 & | & -17 \\
 0 & 0 & -\frac{1}{2} & | & -\frac{3}{2}
 \end{pmatrix}.$$

Durch Gausselimination haben wir das lineare Gleichungssystem (1.5) nun umgewandelt in

$$\begin{array}{l}
 (I) \\
 (II) \\
 (III)
 \end{array}
 \begin{array}{r}
 x_1 + x_2 + 2x_3 = 9 \\
 2x_2 - 7x_3 = -17 \\
 -\frac{1}{2}x_3 = -\frac{3}{2}
 \end{array}
 \quad (1.7)$$

Rückwärtseinsetzen: Aus (III) erhalten wir

$$x_3 = 3.$$

Einsetzen in (II) liefert

$$2x_2 - 21 = -17 \quad \Rightarrow \quad x_2 = 2.$$

Schliesslich, mit (I) :

$$x_1 + 2 + 6 = 9 \quad \Rightarrow \quad x_1 = 1.$$

□

Beispiel 1.8 Betrachten wir weiter das System

$$\begin{array}{l} (I) \quad x_1 + x_2 + 2x_3 = 6 \\ (II) \quad 3x_1 + 3x_2 - x_3 = 11 \\ (III) \quad x_1 + 2x_2 - 4x_3 = 2 \end{array} \quad (1.9)$$

Wir gehen gleich vor wie zuvor:

Schritt 1: Gleichung (I) bleibt unverändert; aus den Gleichungen (II) und (III) eliminieren wir zunächst die Variable x_1 mit Hilfe der Gleichung (I) :

$$\begin{array}{l} (I) \\ (II) \\ (III) \end{array} \left(\begin{array}{ccc|c} 1 & 1 & 2 & 6 \\ \boxed{3} & 3 & -1 & 11 \\ \boxed{1} & 2 & -4 & 2 \end{array} \right) \rightsquigarrow \begin{array}{l} (I) \\ (II) \\ (III) \end{array} \left(\begin{array}{ccc|c} 1 & 1 & 2 & 6 \\ 0 & 0 & -7 & -7 \\ 0 & 1 & -6 & -4 \end{array} \right).$$

Schritt 2: Im nächsten Schritt würden wir nun gerne, wie zuvor, die Variable x_2 aus Gleichung (III) entfernen. Vorher geschah dies durch Subtrahieren eines Vielfaches der zweiten Gleichung von der dritten. Weil die Variable x_2 in Gleichung (II) aber nicht auftritt (0 an der zweiten Stelle!), lässt sich dies hier nicht realisieren. Das Problem ist aber einfach lösbar: Wir vertauschen die zweite und die dritte Gleichung. Dies ändert die Lösung des linearen Gleichungssystems nicht:

$$\begin{array}{l} (I) \\ (II) \\ (III) \end{array} \left(\begin{array}{ccc|c} 1 & 1 & 2 & 6 \\ 0 & 0 & -7 & -7 \\ 0 & \boxed{1} & -6 & -4 \end{array} \right) \rightsquigarrow \begin{array}{l} (I) \\ (II) \\ (III) \end{array} \left(\begin{array}{ccc|c} 1 & 1 & 2 & 6 \\ 0 & 1 & -6 & -4 \\ 0 & 0 & -7 & -7 \end{array} \right)$$

Nun hat das lineare Gleichungssystem genau die gewünschte Form!

Rückwärtseinsetzen: Aus der dritten Zeile lesen wir:

$$-7x_3 = -7 \quad \Rightarrow \quad x_3 = 1.$$

Einsetzen in die zweite Gleichung ergibt:

$$1 \cdot x_2 - 6 = -4 \quad \Rightarrow \quad x_2 = 2.$$

Schliesslich erhalten wir aus Gleichung (I) :

$$x_1 + 2 + 2 \cdot 1 = 6 \quad \Rightarrow \quad x_1 = 2.$$

Durch Einsetzen der Lösungswerte in das ursprüngliche Gleichungssystem kann *überprüft* werden, ob die Lösung richtig gefunden worden ist. □

Wir fassen zusammen:

Algorithmus 1.10 (Gaussalgorithmus) Gegeben sei ein lineares Gleichungssystem mit n Gleichungen und n Unbekannten $x_1, x_2, x_3, \dots, x_n$:

$$\begin{array}{cccccc} a_{11}x_1 & + & a_{12}x_2 & + \cdots + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + \cdots + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \ddots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + \cdots + & a_{nn}x_n & = & b_n \end{array} \quad (1.11)$$

Oder in Matrixform:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{array} \right)$$

Hier sind die Zahlen a_{ij} die Koeffizienten der Unbekannten in den Gleichungen. Genauer steht a_{ij} für den Koeffizient von x_j in Gleichung i . Die Zahlen b_i sind die Zahlen auf der rechten Seite der Gleichheitszeichen im linearen Gleichungssystem.

Das Gauss-Verfahren zur Lösung von (1.11) ist dann wie folgt definiert:

- Schritt 1: Die erste Gleichung bleibt unverändert. Aus den restlichen Gleichungen wird jeweils die Unbekannte x_1 eliminiert. Dies geschieht durch Subtraktion eines geeigneten Vielfachen der *ersten* Gleichung von der entsprechenden Gleichung. Das Gleichungssystem hat dann typischerweise die folgende Form:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\ 0 & \star & \star & \cdots & \star & \star \\ 0 & \star & \star & \cdots & \star & \star \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \star & \star & \cdots & \star & \star \end{array} \right)$$

Hier sind \star irgendwelche Zahlen, die sich aus den Manipulationen in den Zeilen ergeben.

- Schritt 2: Die erste und zweite Gleichung bleiben unberührt. Aus den restlichen Gleichungen wird jeweils die Unbekannte x_2 eliminiert. Dies geschieht durch Subtraktion eines geeigneten Vielfachen der *zweiten* Gleichung von der entsprechenden Gleichung. Das Gleichungssystem hat dann typischerweise die folgende Form:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} & b_1 \\ 0 & \star & \star & \cdots & \star & \star \\ 0 & 0 & \star & \cdots & \star & \star \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \star & \cdots & \star & \star \end{array} \right)$$

- Schritt 3: Falls das Gleichungssystem mehr als nur 3 Gleichungen und Unbekannte hat wird nun genau gleich weitergefahren: Die erste, zweite und dritte Gleichung bleiben unberührt. Aus den restlichen Gleichungen wird jeweils die Unbekannte x_3 eliminiert. Dies geschieht durch Subtraktion eines geeigneten Vielfachen der *dritten* Gleichung von der entsprechenden Gleichung. Das Gleichungssystem hat dann typischerweise die folgende Form:

$$\left(\begin{array}{cccccc|c} a_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1n} & b_1 \\ 0 & * & * & * & \cdots & * & * \\ 0 & 0 & * & * & \cdots & * & * \\ 0 & 0 & 0 & * & \cdots & * & * \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & * & \cdots & * & * \end{array} \right)$$

- Nach dem gleichen Prinzip wird nun weiterverfahren, bis das Gleichungssystem nach n Schritten typischerweise (nicht immer!) die folgende Form aufweist:

$$\left(\begin{array}{cccccc|c} a_{11} & a_{12} & a_{13} & a_{14} & \cdots & a_{1n} & b_1 \\ 0 & * & * & * & \cdots & * & * \\ 0 & 0 & * & * & \cdots & * & * \\ 0 & 0 & 0 & * & \cdots & * & * \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & * & * \end{array} \right)$$

Diese Form wird auch *rechts-obere Dreiecksform* genannt.

- Auflösen durch Rückwärtseinsetzen: Ausgehend von der rechts-oberen Dreiecksform, kann das Gleichungssystem nun “von unten her” aufgelöst werden. In der Tat lässt sich die letzte Gleichung direkt nach x_n auflösen. Einsetzen dieses Wertes in die zweitletzte Gleichung ermöglicht dann die Lösung nach x_{n-1} , etc.

Unter Umständen ist es nötig in den einzelnen Schritten geeignete *Zeilenvertauschungen* vorzunehmen.

Wir betrachten nochmals ein Beispiel.

Beispiel 1.12 Gegeben ist das folgende lineare Gleichungssystem mit 4 Unbekannten und 4 Gleichungen:

$$\begin{array}{l} (I) \\ (II) \\ (III) \\ (IV) \end{array} \left(\begin{array}{cccc|c} 1 & 3 & 0 & 1 & 2 \\ \boxed{1} & 4 & 2 & 0 & 1 \\ \boxed{0} & -2 & -2 & -1 & -1 \\ \boxed{2} & -4 & 1 & 1 & 3 \end{array} \right)$$

Schritt 1: Erste Gleichung unverändert, Elimination von x_1 aus den restlichen Gleichungen mit Hilfe von Gleichung (I).

$$\rightsquigarrow \begin{array}{l} (II) - (I) \\ (III) \\ (IV) - 2 \cdot (I) \end{array} \left(\begin{array}{cccc|c} 1 & 3 & 0 & 1 & 2 \\ 0 & 1 & 2 & -1 & -1 \\ 0 & -2 & -2 & -1 & -1 \\ 0 & -10 & 1 & -1 & -1 \end{array} \right).$$

Hier konnten wir die Gleichung (III) unverändert lassen, da sie bereits die gewünschte Form hat.

Schritt 2: Erste und zweite Gleichung unverändert. Elimination von x_2 aus den restlichen Gleichungen mit Hilfe von Gleichung (II).

$$\rightsquigarrow \begin{array}{l} (III) - (-2) \cdot (II) \\ (IV) - (-10) \cdot (II) \end{array} \left(\begin{array}{cccc|c} 1 & 3 & 0 & 1 & 2 \\ 0 & 1 & 2 & -1 & -1 \\ 0 & 0 & 2 & -1 & -1 \\ 0 & 0 & 21 & -11 & -11 \end{array} \right).$$

Schritt 3: Erste, zweite und dritte Gleichung bleiben unverändert. Elimination von x_3 aus der vierten Gleichung mit Hilfe von Gleichung (III).

$$\rightsquigarrow \begin{array}{l} (IV) - 21/2 \cdot (III) \end{array} \left(\begin{array}{cccc|c} 1 & 3 & 0 & 1 & 2 \\ 0 & 1 & 2 & -1 & -1 \\ 0 & 0 & 2 & -1 & -1 \\ 0 & 0 & 0 & -1/2 & -1/2 \end{array} \right).$$

Nun hat das Gleichungssystem rechts-obere Dreiecksform.

Rückwärtseinsetzen:

$$\begin{array}{l} (IV) \quad -\frac{1}{2}x_4 = -\frac{1}{2} \quad \Rightarrow \quad x_4 = 1 \\ (III) \quad 2x_3 - 1 \cdot 1 = -1 \quad \Rightarrow \quad x_3 = 0 \\ (II) \quad x_2 + 2 \cdot 0 - 1 \cdot 1 = -1 \quad \Rightarrow \quad x_2 = 0 \\ (I) \quad x_1 + 3 \cdot 0 + 0 \cdot 0 + 1 \cdot 1 = 2 \quad \Rightarrow \quad x_1 = 1. \end{array}$$

□

1.2 Lineare Gleichungssysteme und Matrizen

Wir haben bereits gesehen, dass sich lineare Gleichungssysteme elegant durch Matrizen darstellen lassen:

$$\begin{array}{rcl} x_1 & + & x_2 & + & 2x_3 & = & 6 \\ 3x_1 & + & 3x_2 & - & x_3 & = & 11 \\ x_1 & + & 2x_2 & - & 4x_3 & = & 2 \end{array} \Leftrightarrow \left(\begin{array}{ccc|c} 1 & 1 & 2 & 6 \\ 3 & 3 & -1 & 11 \\ 1 & 2 & -4 & 2 \end{array} \right).$$

Die Matrix, welche zu einem linearen Gleichungssystem gehört lässt sich in zwei verschiedene Teile “aufspalten”. Hierbei ist der erste Teil eine Matrix (beispielsweise bezeichnet mit A^1), die die Koeffizienten der Unbekannten enthält. Der zweite Teil ist ein Vektor (zum Beispiel bezeichnet mit b), der die Zahlen auf der rechten Seite des Gleichheitszeichens im Gleichungssystem beinhaltet. Wir zeigen dies anhand des obigen Beispiels:

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 6 \\ 3 & 3 & -1 & 11 \\ 1 & 2 & -4 & 2 \end{array} \right) \rightsquigarrow A = \begin{pmatrix} 1 & 1 & 2 \\ 3 & 3 & -1 \\ 1 & 2 & -2 \end{pmatrix}, \quad b = \begin{pmatrix} 6 \\ 11 \\ 2 \end{pmatrix}.$$

Oftmals werden auch die Unbekannten in einem Vektor gespeichert:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

Im Zusammenhang mit linearen Gleichungssystem heisst

- die Matrix A die **Koeffizientenmatrix**,
- der Vektor b der **Rechtseite-Vektor**,
- der Vektor x der **Unbekannten-Vektor**

des linearen Gleichungssystems. Die Matrix

$$(A \mid b)$$

heisst **erweiterte Matrix** des linearen Gleichungssystems.

Beispiel 1.13 Betrachten wir die Koeffizientenmatrix $A = \begin{pmatrix} 0 & -1 \\ 3 & 2 \end{pmatrix}$ und den Rechtseite-Vektor $b = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$ eines linearen Gleichungssystems. Wie lautet die Lösung?

Die erweiterte Matrix ist

$$\left(\begin{array}{cc|c} 0 & -1 & 3 \\ 3 & 2 & 5 \end{array} \right).$$

Hier lässt sich die Gaußelimination nicht direkt anwenden, da der Eintrag in Zeile 1 und Spalte 1 gleich 0 ist. Wir vertauschen daher die Zeilen:

$$\left(\begin{array}{cc|c} 3 & 2 & 5 \\ 0 & -1 & 3 \end{array} \right).$$

Nun hat die Matrix bereits rechts-obere Dreiecksform (ohne Eliminationsschritte) und kann durch backward solve gelöst werden:

$$x_2 = -3, \quad 3x_1 - 5 \cdot 3 = 5 \quad \Rightarrow \quad x_1 = \frac{20}{3}.$$

□

¹In diesem Skript schreiben wir Matrizen mit fetten Grossbuchstaben und Vektoren mit fetten Kleinbuchstaben.

1.3 Anwendungen

Lineare Gleichungssysteme kommen vielerorts in den Naturwissenschaften und anderen Gebieten vor. Anwendungen, welche die Lösung von linearen Gleichungssystemen erfordern, gibt es unzählige. Ein wichtiges Beispiel haben wir bereits in Abschnitt 1.1 beim Berechnen von einfachen Stromkreisen kennengelernt. Wir wollen nun noch einige weitere Anwendungsbereiche betrachten.

1.3.1 Interpolation von Daten

Gegeben sei eine Menge von Daten, welche zu einem gewissen, beispielsweise physikalischen, Vorgang gehört. Zu diesem Gesetz gebe es ein mathematisches Modell, das von gewissen Parametern bestimmt ist. Wie müssen die Parameter gewählt werden, damit die Daten vom entsprechenden Modell erzeugt werden? Mit dieser Frage befasst sich die Interpolation.

Beispiel 1.14 Die Lufttemperatur nimmt typischerweise mit wachsender Höhe ab. Der genaue Zusammenhang hängt von verschiedenen Faktoren ab und ist sehr kompliziert. Wir betrachten das folgende (stark) vereinfachte Modell:

- Auf der Meereshöhe $h = 0$ beträgt die Temperatur $T(0) = 15^\circ\text{C}$.
- Im Höhenbereich von 0 bis 11000m nimmt die Temperatur gleichmässig ab bis auf $T(11000) = -56.5^\circ\text{C}$.
- Über 11000m bleibt die Temperatur konstant.

Aufgabe: Finde die Funktion $T(h)$, welche die Temperatur in Abhängigkeit der Höhe h angibt.

Zeichnen wir die Temperatur in einem Koordinatensystem in Abhängigkeit der Höhe auf, so erhalten wir das folgende Bild: Die gleichmässige Abnahme der Temperatur unter 11000m entspricht einer Geraden mit negativer Steigung. Für noch grössere Höhen bleibt die Temperatur konstant, was eine horizontale Gerade ergibt; vgl. Abbildung 1.2. Zwischen $0 \leq h \leq 11000\text{m}$ hat die gesuchte Gerade die Form

$$T(h) = ah + b,$$

wobei die Parameter a und b zu bestimmen sind. Wir wissen, dass

$$\begin{aligned} 15 &= T(0) = a \cdot 0 + b = b \\ -56.5 &= T(11000) = 11000a + b. \end{aligned}$$

Somit ergibt sich das lineare Gleichungssystem

$$\begin{aligned} b &= 15 \\ 11000a + b &= -56.5 \end{aligned}$$

Die Lösung lautet:

$$a = \frac{-71.5}{11000} = -0.0065, \quad b = 15.$$

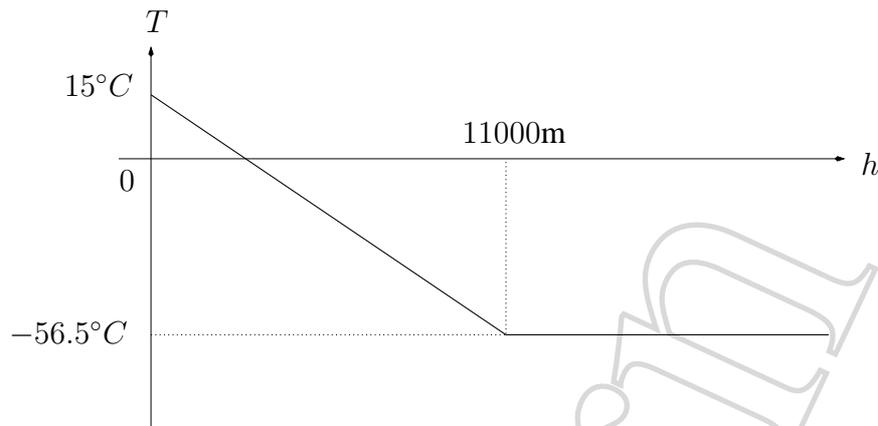


Abbildung 1.2: Vereinfachtes Atmosphärenmodell: Temperatur in Abhängigkeit der Höhe.

Die gesuchte Temperaturfunktion ist somit gegeben durch

$$T(h) = \begin{cases} -0.0065h + 15 & \text{wenn } h \text{ zwischen } 0 \text{ und } 11000\text{m} \text{ liegt} \\ -56.5 & \text{wenn } h \geq 11000\text{m} \end{cases}$$

□

Beispiel 1.15 Auf einer Parabel

$$f(x) = x^2 + ax + b$$

liegen die zwei Punkte $(-1, 1)$ und $(2, 2)$. Zu finden sind die Parameter a und b . Es gilt:

$$1 = f(-1) = 1 - a + b, \quad 2 = f(2) = 4 + 2a + b.$$

Dies führt auf das lineare Gleichungssystem

$$\begin{cases} -a + b = 0 \\ 2a + b = -2 \end{cases}$$

welches die Lösung $a = -\frac{2}{3}$ und $b = \frac{2}{3}$ hat.

□

1.3.2 Mischungen und Konzentrationen

Beispiel 1.16 Zwei Flüssigkeiten F_1 und F_2 enthalten 3% respektive 6% Alkohol. Es soll eine Mischung von 5 Litern hergestellt werden, die 5% Alkohol enthält.

Lösung: Wir bezeichnen die Anzahl Liter von jeder Flüssigkeit im Gemisch mit l_1 und l_2 . Sicherlich muss gelten:

$$l_1 + l_2 = 5.$$

Nun leiten wir noch eine Gleichung für die Alkoholanteile her:

Menge (in Litern) Alkohol in l_1 Litern von F_1 : $0.03l_1$;

Menge Alkohol in l_2 Litern von F_2 : $0.06l_2$;

Menge Alkohol in $l_1 + l_2$ Litern vom Gemisch soll sein: $0.05(l_1 + l_2)$.

Vor und nach dem Mischen ist die Gesamtmenge des Alkohols gleich. Daher:

$$0.03l_1 + 0.06l_2 = 0.05(l_1 + l_2) \quad \Rightarrow \quad 3l_1 + 6l_2 = 5(l_1 + l_2) \quad \Rightarrow \quad -2l_1 + l_2 = 0.$$

Wir erhalten das lineare Gleichungssystem

$$\begin{aligned} l_1 + l_2 &= 5 \\ -2l_1 + l_2 &= 0. \end{aligned}$$

Die gesuchten Anteile der Flüssigkeiten für die richtige Mischung entspricht der Lösung dieses linearen Gleichungssystems:

$$l_1 = \frac{5}{3} \text{ Liter}, \quad l_2 = \frac{10}{3} \text{ Liter.}$$

□

1.3.3 Ebene und räumliche Geometrie

Beispiel 1.17 Im drei-dimensionalen Raum seien die folgenden drei Ebenen gegeben:

$$\begin{aligned} E_1 : & \quad x_1 + x_2 + x_3 = 1 \\ E_2 : & \quad 2x_1 + 2x_2 + x_3 = -3 \\ E_3 : & \quad -3x_1 + 5x_2 + 6x_3 = 1 \end{aligned}$$

Sie schneiden sich in einem Punkt P . Dieser Punkt liegt auf allen drei Ebenen, d.h., seine Koordinaten erfüllen das Gleichungssystem

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 2 & 2 & 1 & -3 \\ -3 & 5 & 6 & 1 \end{array} \right).$$

Es hat die Lösung

$$x_1 = \frac{9}{8}, \quad x_2 = -\frac{41}{8}, \quad x_3 = 5,$$

und dies sind die Koordinaten des Schnittpunktes.

□

Beispiel 1.18 Gesucht ist die Gleichung einer Ebene E im Raum,

$$E : \quad ax_1 + bx_2 + cx_3 = 1,$$

die durch die drei Punkte

$$P_1 : (-2, 1, 1), \quad P_2 : (4, 0, -1), \quad P_3 : (2, -2, -1)$$

geht. Da der Punkt P_1 auf E liegt, muss gelten

$$-2a + b + c = 1.$$

Analog erhalten wir für P_2 und P_3 :

$$4a - c = 1, \quad 2a - 2b - c = 1.$$

Diese drei Gleichungen formen ein lineares Gleichungssystem für die drei Unbekannten a, b, c . Es lässt sich mit dem Gaußverfahren lösen. Als Lösung errechnet man $a = 2, b = -2, c = 7$. Die gesuchte Gleichung der Ebene E ist durch

$$E : \quad 2x_1 - 2x_2 + 7x_3 = 1$$

gegeben. □

1.4 Reguläre und singuläre Matrizen, Rang einer Matrix

In den bisherigen Beispielen haben die linearen Gleichungssysteme immer genau eine Lösung gehabt. Dies muss nicht immer so sein, wie die folgenden zwei Beispiele zeigen.

Beispiel 1.19 Betrachten wir das lineare Gleichungssystem

$$\begin{array}{l} (I) \\ (II) \\ (III) \end{array} \quad \left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 1 \end{array} \right)$$

Gaußelimination liefert:

$$\begin{array}{l} \rightsquigarrow \\ \rightsquigarrow \end{array} \quad \begin{array}{l} (II) - 4 \cdot (I) \\ (III) - 7 \cdot (I) \end{array} \quad \left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & -3 & -6 & 0 \\ 0 & -6 & -12 & 1 \end{array} \right)$$

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & -3 & -6 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right)$$

Die letzte Gleichung lautet jetzt:

$$0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 = 7.$$

Offenbar gibt es keine Wahl von x_1, x_2, x_3 welche diese Gleichung erfüllen kann. Folglich hat dieses Gleichungssystem *keine Lösung*. □

Beispiel 1.20 Wir ändern im vorherigen Beispiel den Rechenseite-Vektor:

$$\begin{array}{l} (I) \\ (II) \\ (III) \end{array} \quad \left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 0 \end{array} \right) \quad (1.21)$$

Die einzelnen Schritte der Gaußelimination sind nun die gleichen wie vorher, mit dem Unterschied, dass die Zahlen auf der rechten Seite des linearen Gleichungssystems immer gleich 0 bleiben. Somit,

$$\rightsquigarrow \quad (III) - (-2) \cdot (II) \quad \left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & -3 & -6 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) \quad (1.22)$$

Die letzte Gleichung lautet jetzt

$$0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 = 0. \quad (1.23)$$

Sie enthält keine Information, und jede beliebige Kombination von Zahlen für x_1, x_2, x_3 löst diese Gleichung. Die dritte Zeile in der Matrix (1.22) wird deshalb auch **Nullzeile** genannt. Das System reduziert sich damit auf die erste und zweite Gleichung im transformierten System:

$$\begin{array}{l} x_1 + 2x_2 + 3x_3 = 0 \\ -3x_2 - 6x_3 = 0. \end{array}$$

Wiederum lösen wir von “unten her” auf. Aus der zweiten Gleichung folgt

$$x_2 = -2x_3.$$

Einsetzen in die erste Gleichung führt zu

$$x_1 + 2 \cdot (-2x_3) + 3x_3 = 0 \quad \Rightarrow \quad x_1 = x_3.$$

Wir sehen, dass der Wert von x_3 unbestimmt bleibt. Wir haben keine Information mehr, um ihn festzusetzen. Dies liegt daran, dass wir durch das “Verlorengehen” von Gleichung (1.23) im Prinzip nur zwei Gleichungen zur Verfügung haben, um die (drei) Unbekannten zu bestimmen. Der Wert von x_3 ist nun frei wählbar. In diesem Fall führt man häufig einen sogenannten *Parameter* ein, der verwendet wird, um die Gesamtmenge der Lösungen auszudrücken.

Ersetzen wir x_3 durch einen Parameter s , so lässt sich die Lösung des linearen Gleichungssystems wie folgt schreiben:

$$x_1 = s, \quad x_2 = -2s, \quad x_3 = s. \quad (1.24)$$

Hier kann s einen beliebigen Wert annehmen. Beispielsweise erhalten wir für $s = -4$ die Lösung

$$x_1 = -4, \quad x_2 = 8, \quad x_3 = -4.$$

In der Mathematik verwendet man häufig die Schreibweise

$$\mathbb{L} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : x_1 = s, x_2 = -2s, x_4 = s, \text{ wobei } s \text{ eine beliebige Zahl ist} \right\},$$

um die Menge \mathbb{L} aller Lösungen (Lösungsmenge) zu beschreiben. In Worten bedeutet dies: "Die Menge aller Vektoren $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, für die gilt: $x_1 = s, x_2 = -2s, x_4 = s$, wobei s eine beliebige Zahl ist."

Wir können einfach nachprüfen, ob die Lösung in (1.24) tatsächlich für alle s eine Lösung des ursprünglichen linearen Gleichungssystems ist. Dazu setzen wir sie in (1.21) ein. Für die erste Gleichung gilt

$$x_1 + 2x_2 + 3x_3 = s + 2 \cdot (-2s) + 3s = 0,$$

d.h., die Gleichung ist erfüllt. Genau gleich überprüfen wir die Gleichungen (I) und (II) :

$$4x_1 + 5x_2 + 6x_3 = 4s - 10s + 6s = 0$$

$$7x_1 + 8x_2 + 9x_3 = 7s - 16s + 9s = 0.$$

Somit ist die gefundene Lösung korrekt und es folgt, dass dieses lineare Gleichungssystem unendlich viele Lösungen (nämlich eine Lösung für jede Wahl des Parameters s) hat. \square

Wir halten fest: Ein Gleichungssystem kann *keine*, *genau eine*, oder *unendlich viele* Lösungen haben. Dies hängt sowohl von der Koeffizientenmatrix als auch vom Rechenseite-Vektor ab.

Betrachten wir ein lineares Gleichungssystem, welches den Nullvektor $\mathbf{0}$ als Rechenseite-Vektor hat, d.h.,

$$(\mathbf{A} \mid \mathbf{0}).$$

Schreiben wir die Koeffizientenmatrix des linearen Gleichungssystems in der Form

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$

so gilt:

$$\begin{array}{cccccc} a_{11}x_1 & + & a_{12}x_2 & + \cdots + & a_{1n}x_n & = & 0 \\ a_{21}x_1 & + & a_{22}x_2 & + \cdots + & a_{2n}x_n & = & 0 \\ \vdots & & \vdots & & \ddots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + \cdots + & a_{nn}x_n & = & 0 \end{array} \quad (1.25)$$

Ein solches lineares Gleichungssystem heisst **homogen**. Falls es auf der rechten Seite mindestens einen Wert gibt, der ungleich 0 ist, dann heisst das System **inhomogen**.

Offensichtlich hat das Gleichungssystem (1.25) *immer mindestens eine Lösung*, nämlich die Lösung

$$x_1 = x_2 = \dots = x_n = 0.$$

Diese Lösung wird auch **Null-Lösung** genannt.

Wir sagen:

- Eine Matrix A ist **regulär**, falls das zugehörige homogene Gleichungssystem (1.25) *genau eine Lösung* hat, nämlich die Null-Lösung.
- Eine Matrix A ist **singulär**, falls das zugehörige homogene Gleichungssystem (1.25) *mehr als eine Lösung* hat.

Beispiel 1.26 Sind die folgenden Matrizen regulär oder singulär?

$$\text{a) } A = \begin{pmatrix} -1 & 2 \\ 4 & 5 \end{pmatrix}, \quad \text{b) } B = \begin{pmatrix} 2 & 1 \\ 8 & 4 \end{pmatrix},$$

Betrachten wir zunächst die Matrix A und das zugehörige homogene Gleichungssystem:

$$\begin{array}{l} (I) \\ (II) \end{array} \left(\begin{array}{cc|c} -1 & 2 & 0 \\ \boxed{4} & 5 & 0 \end{array} \right).$$

Mit dem Gauss-Verfahren finden wir

$$\rightsquigarrow \begin{array}{l} (I) \\ (II) - (-4) \cdot (I) \end{array} \left(\begin{array}{cc|c} -1 & 2 & 0 \\ 0 & 13 & 0 \end{array} \right).$$

Die letzte Gleichung heisst damit

$$13x_2 = 0 \quad \Rightarrow \quad x_2 = 0.$$

Einsetzen in die erste Gleichung ergibt dann

$$-1 \cdot x_1 + 2 \cdot 0 = 0 \quad \Rightarrow \quad x_1 = 0.$$

Somit ist $x_1 = x_2 = 0$ die einzige Lösung des homogenen Gleichungssystem. Die Matrix A ist deshalb regulär.

Analysieren wir die Matrix B . Wiederum betrachten wir das zugehörige homogene Gleichungssystem

$$\begin{array}{l} (I) \\ (II) \end{array} \left(\begin{array}{cc|c} 2 & 1 & 0 \\ \boxed{8} & 4 & 0 \end{array} \right).$$

Gausselimination liefert

$$\rightsquigarrow \begin{array}{l} (I) \\ (II) - 4 \cdot (I) \end{array} \left(\begin{array}{cc|c} 2 & 1 & 0 \\ 0 & 0 & 0 \end{array} \right).$$

Wie in Beispiel 1.20 ist auch hier die letzte Zeile eine *Nullzeile*, d.h., sie enthält keine Information. Der Wert von x_2 ist somit frei wählbar. Wir ersetzen x_2 durch einen Parameter s ,

$$x_2 = s.$$

Aus der ersten Gleichung folgt dann

$$2x_1 + 1 \cdot x_2 = 0 \quad \Rightarrow \quad x_1 = -\frac{1}{2}s.$$

Das homogene Gleichungssystem hat also unendlich viele Lösungen. Die Matrix B ist somit singulär. \square

Aus den obigen Beispielen 1.20 und 1.26 b) sehen wir, dass beim Anwenden des Gaussverfahrens auf ein *homogenes* lineares Gleichungssystem Nullzeilen entstehen können (insbesondere kann es auch mehr als nur eine geben). Die Differenz zwischen der Anzahl Zeilen einer Matrix und der Anzahl Nullzeilen, die bei der Gausselimination entstehen, wird **Rang** einer Matrix genannt und mit rang bezeichnet.

Für eine $m \times n$ -Matrix A haben wir demnach:

$$\begin{aligned} \text{rang}(A) &= (\text{Anzahl Zeilen } m \text{ von } A) \\ &\quad - (\text{Anzahl Nullzeilen die entstehen beim vollständigen Anwenden} \\ &\quad \text{des Gaussverfahrens auf das homogene lineare} \\ &\quad \text{Gleichungssystem } (A | 0)). \end{aligned}$$

Wir stellen fest: Ist $\text{rang}(A) = 0$, dann gibt es nach abgeschlossener Gausselimination keine Nullzeilen, d.h., die Matrix A ist regulär. Falls $\text{rang}(A) \geq 1$, dann ist A singulär.

Beispiel 1.27 Wir bestimmen den Rang der Matrix

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & -1 & 2 & 3 \\ 3 & -1 & 4 & 7 \\ -3 & 3 & -6 & -9 \end{pmatrix}.$$

Dazu führen wir die Gausselimination vollständig durch. Dies ergibt nach dem ersten Schritt

$$\begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & -2 & 2 & 2 \\ 0 & -4 & 4 & 4 \\ 0 & 6 & -6 & -6 \end{pmatrix},$$

und nach dem zweiten Schritt

$$\begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & -2 & 2 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Hier endet das Gaussverfahren und es werden 2 Nullzeilen sichtbar. Da die Matrix \mathbf{A} genau 4 Zeilen hat, gilt

$$\text{rang}(\mathbf{A}) = 4 - 2 = 2.$$

Insbesondere ist $\text{rang}(\mathbf{A}) \geq 1$ und die Matrix \mathbf{A} ist singulär. \square

Die Klassifikation in reguläre und singuläre Matrizen ist in der Praxis sehr nützlich. Es lässt sich der folgende Satz zeigen:

Satz 1.28 Gegeben sei ein Gleichungssystem mit n Gleichungen und n Unbekannten

$$(\mathbf{A} \mid \mathbf{b}). \quad (1.29)$$

Hier ist \mathbf{A} eine Koeffizientenmatrix (mit n Zeilen und n Spalten) und \mathbf{b} ist ein beliebiger Rechenseite-Vektor (mit n Einträgen).

- Falls \mathbf{A} regulär ist, dann hat das lineare Gleichungssystem (1.29) immer genau eine Lösung.
- Falls \mathbf{A} singulär ist, dann hat das lineare Gleichungssystem (1.29) entweder keine oder unendlich viele Lösungen.

Beispiel 1.30 Ist die Lösung des linearen Gleichungssystems

$$\begin{aligned} -x_1 + 2x_2 &= 10 \\ 4x_1 + 5x_2 &= 21 \end{aligned} \quad (1.31)$$

eindeutig?

Satz 1.28 zeigt uns, dass es genügt, die entsprechende Koeffizientenmatrix des Systems zu betrachten:

$$\mathbf{A} = \begin{pmatrix} -1 & 2 \\ 4 & 5 \end{pmatrix}.$$

Dies ist die Matrix aus Beispiel 1.26 a). Wir haben bereits festgestellt, dass sie regulär ist. Folglich hat das Gleichungssystem (1.31) *genau eine Lösung*. Sie lautet

$$x_1 = -\frac{8}{13}, \quad x_2 = \frac{61}{13}$$

und kann mit dem Gaussverfahren gefunden werden. \square

1.5 Determinanten

Im letzten Abschnitt haben wir reguläre und singuläre Matrizen im Zusammenhang mit der Anzahl von Lösungen linearer Gleichungssysteme definiert. Wir wollen uns nun die Frage stellen, ob es andere Möglichkeiten gibt, die Regularität bzw. Singularität einer Matrix festzustellen.

Betrachten wir dazu ein beliebiges lineares Gleichungssystem mit 2 Gleichungen und 2 Unbekannten:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned} \quad (1.32)$$

Hier sind die Zahlen $a_{11}, a_{12}, a_{21}, a_{22}$ und b_1, b_2 gegeben. Die Zahlen x_1 und x_2 sind die gesuchten Unbekannten. Die Koeffizientenmatrix ist

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}. \quad (1.33)$$

Wann ist sie regulär, wann ist sie singulär? Dazu betrachten wir, nach Definition, das zugehörige homogene Gleichungssystem

$$\begin{aligned} (I) & \quad \left(a_{11} \quad a_{12} \mid 0 \right) \\ (II) & \quad \left(a_{21} \quad a_{22} \mid 0 \right). \end{aligned} \quad (1.34)$$

Wir lösen es mit Hilfe des Gaussverfahrens. Dazu unterscheiden wir die folgenden zwei Fälle:

Fall 1: Es sei $a_{11} \neq 0$. Dann erhalten wir nach einem Schritt mit der Gausselimination:

$$\rightsquigarrow \quad (II) - a_{21}/a_{11} \cdot (I) \quad \left(\begin{array}{cc|c} a_{11} & a_{12} & 0 \\ 0 & a_{22} - a_{21}/a_{11} \cdot a_{12} & 0 \end{array} \right).$$

Die letzte Gleichung lautet nun

$$\left(a_{22} - \frac{a_{21}}{a_{11}} a_{12} \right) x_2 = 0.$$

Sie hat die eindeutige Lösung $x_2 = 0$, genau falls

$$a_{22} - \frac{a_{21}}{a_{11}} a_{12} \neq 0, \quad \text{d.h.,} \quad a_{11}a_{22} - a_{12}a_{21} \neq 0.$$

Einsetzen in Gleichung (I) ergibt

$$a_{11}x_1 = 0.$$

Da $a_{11} \neq 0$ erhalten wir $x_1 = 0$. Die Lösung des homogenen linearen Gleichungssystems hat also genau und nur die Null-Lösung, falls

$$a_{11}a_{22} - a_{12}a_{21} \neq 0. \quad (1.35)$$

Fall 2: Es sei $a_{11} = 0$. Vertauschen der Zeilen im linearen Gleichungssystem liefert

$$\rightsquigarrow \begin{array}{l} (I) \leftrightarrow (II) \\ (II) \leftrightarrow (I) \end{array} \left(\begin{array}{cc|c} a_{21} & a_{22} & 0 \\ 0 & a_{12} & 0 \end{array} \right). \quad (1.36)$$

Die letzte Gleichung lautet nun

$$a_{12}x_2 = 0.$$

Sie hat die eindeutige Lösung $x_2 = 0$, genau falls $a_{12} \neq 0$. Einsetzen in die erste Gleichung von (1.36) impliziert

$$a_{21}x_1 = 0.$$

Wir erhalten die eindeutige Lösung $x_1 = 0$, genau falls $a_{21} \neq 0$. Im Fall $a_{11} = 0$ hat das homogene System also genau die Null-Lösung, falls $a_{12} \neq 0$ und $a_{21} \neq 0$. Im Vergleich mit (1.35) finden wir dann ebenfalls

$$\underbrace{\underbrace{a_{11}}_{=0} a_{22}}_{=0} - \underbrace{\underbrace{a_{12}}_{\neq 0} \underbrace{a_{21}}_{\neq 0}}_{\neq 0} \neq 0.$$

Ausserdem bemerken wir, falls $a_{12} = 0$ und/oder $a_{21} = 0$, dass gilt

$$\underbrace{\underbrace{a_{11}a_{22}}_{=0}}_{=0} - \underbrace{\underbrace{a_{12}a_{21}}_{=0}}_{=0} = 0.$$

Fassen wir die beiden Fälle zusammen, so halten wir fest: Das Gleichungssystem (1.34) hat genau die Null-Lösung (und nur diese), falls

$$a_{11}a_{22} - a_{12}a_{21} \neq 0.$$

In Bezug auf die Matrix \mathbf{A} in (1.33) bedeutet dies:

Die 2×2 -Matrix \mathbf{A} ist *regulär*, falls

$$a_{11}a_{22} - a_{12}a_{21} \neq 0,$$

und *singulär*, falls

$$a_{11}a_{22} - a_{12}a_{21} = 0.$$

Die Zahl $a_{11}a_{22} - a_{12}a_{21}$ wird **Determinante** von \mathbf{A} genannt und geschrieben als

$$\det(\mathbf{A}) = a_{11}a_{22} - a_{12}a_{21}. \quad (1.37)$$

Beispiel 1.38 Ist die Matrix

$$\mathbf{M} = \begin{pmatrix} -12 & 22 \\ -5 & 23 \end{pmatrix}$$

regulär oder singular? Wir berechnen die Determinante:

$$\det(\mathbf{M}) = (-12) \cdot 23 - (-5) \cdot 22 = -276 - (-110) = -166 \neq 0.$$

Die Determinante ist ungleich 0; die Matrix \mathbf{M} ist daher regulär. Betrachten wir beispielsweise das lineare Gleichungssystem

$$\left(\begin{array}{cc|c} -12 & 22 & -443 \\ -5 & 23 & 5453 \end{array} \right),$$

folgt aus der Regularität der Koeffizientenmatrix (d.h., \mathbf{M}), dass es eine eindeutige Lösung gibt. \square

Das Konzept der Determinante lässt sich auf Matrizen, welche mehr als zwei Zeilen resp. Spalten haben, erweitern. Determinanten sind aber immer nur für sogenannte **quadratische Matrizen** definiert, d.h., Matrizen die gleich viele Zeilen wie Spalten haben. Ferner lassen sich Determinanten im Allgemeinen nicht mehr so einfach wie im Fall der 2×2 -Matrizen berechnen. Wie bei den 2×2 -Matrizen ist die Determinante eine Zahl, die uns sagt, ob eine Matrix regulär oder singular ist: Falls $\det(\mathbf{A}) = 0$, dann bedeutet dies, dass die Matrix \mathbf{A} singular ist; falls $\det(\mathbf{A}) \neq 0$, dann ist die Matrix regulär.

Zusammenfassend gilt:

Um festzustellen, ob ein Gleichungssystem

$$(\mathbf{A} \mid \mathbf{b}) \quad (*)$$

genau eine Lösung hat, überprüfen wir eine der beiden folgenden Bedingungen:

(1) Das zugehörige homogene Gleichungssystem

$$(\mathbf{A} \mid \mathbf{0})$$

hat nur die Nulllösung oder, äquivalent,

(2) es gilt

$$\det(\mathbf{A}) \neq 0,$$

d.h., die Matrix \mathbf{A} ist regulär.

Ansonsten ist die Matrix \mathbf{A} singular und das Gleichungssystem $(*)$ hat entweder keine oder unendlich viele Lösungen (das homogene System hat dann immer unendlich viele Lösungen).

1.6 MATLAB und OCTAVE

Viele Berechnungen in der linearen Algebra, wie beispielsweise das Bestimmen von Determinanten oder das Lösen von linearen Gleichungssystemen, sind oftmals sehr aufwendig. Dies ist

vor allem dann der Fall, wenn Berechnungen mit sehr grossen Matrizen nötig sind. Hier wird in der Praxis häufig auf geeignete Computerprogramme, wie beispielsweise MATLAB² oder die frei verfügbare Software OCTAVE³, zurückgegriffen.

Beispiel 1.39 Wir berechnen die Determinante von

$$A = \begin{pmatrix} 1 & 5 & 3 & 7 & -3 \\ -2 & 3 & 1 & 3 & 1 \\ 0 & 0 & -1 & 1 & 7 \\ 3 & -5 & 2 & 3 & 1 \\ 1 & -1 & -3 & -2 & 5 \end{pmatrix}$$

in OCTAVE:

```
octave:1> A = [ 1  5  3  7 -3
>             -2  3  1  3  1
>             0  0 -1  1  7
>             3 -5  2  3  1
>             1 -1 -3 -2  5]
```

A =

```
  1  5  3  7 -3
-2  3  1  3  1
  0  0 -1  1  7
  3 -5  2  3  1
  1 -1 -3 -2  5
```

```
octave:2> det(A)
```

```
ans = 659
```

Es gilt also

$$\det(\mathbf{A}) = 659 \neq 0,$$

d.h., die Matrix \mathbf{A} ist regulär. □

Mit MATLAB oder OCTAVE lassen sich auch Gleichungssysteme einfach lösen.

Beispiel 1.40 Betrachte

$$A = \left(\begin{array}{ccccc|c} 1 & 5 & 3 & 7 & -3 & 107 \\ -2 & 3 & 1 & 3 & 1 & 34 \\ 0 & 0 & -1 & 1 & 7 & -65 \\ 3 & -5 & 2 & 3 & 1 & -19 \\ 1 & -1 & -3 & -2 & 5 & -78 \end{array} \right).$$

²MATLAB is a trademark of The MathWorks™

³© John W. Eaton and others, siehe <http://www.gnu.org/software/octave/index.html>

Die Koeffizientenmatrix ist genau die Matrix aus Beispiel 1.39. Da sie regulär ist, hat das obige lineare Gleichungssystem genau eine Lösung. Um diese mit OCTAVE zu berechnen, müssen wir zuerst die Koeffizientenmatrix eingeben wie zuvor. Anschliessend geben wir den Rechtseitevektor ein:

```
octave:3> b=[107
>          34
>         -65
>         -19
>         -78]
```

```
b =
    107
     34
    -65
    -19
    -78
```

Nun wird das Gleichungssystem mit der Operation `\` (“backslash operator”) gelöst:

```
octave:4> A\b
ans =
   -3.0000
    5.0000
    2.0000
    7.0000
   -10.0000
```

□

1.7 Überbestimmte und unterbestimmte lineare Gleichungssysteme

Bis jetzt haben wir uns auf quadratische lineare Gleichungssystem beschränkt, d.h., Gleichungssysteme welche gleich viele Gleichungen wie Unbekannte haben. Es kann aber auch sein, dass ein lineares Gleichungssystem

- mehr Gleichungen wie Unbekannte hat. In diesem Fall ist das Gleichungssystem **überbestimmt**. Wir werden uns mit solchen System in Kapitel 4 genauer beschäftigen.

- weniger Gleichungen als Unbekannte hat. Dann sprechen wir von einem **unterbestimmten** linearen Gleichungssystem. Solche Systeme haben nie eine eindeutige Lösung; entweder sie haben unendlich viele Lösungen oder keine. Wir geben im folgenden einige Beispiele.

Anwendung 1.41 (Chemische Reaktionen)

Wir betrachten die folgende chemische Reaktion zur Herstellung von Eisen:



Hier sind a, b, c, d die unbekanntenen *stöchiometrischen Koeffizienten*.

Wir gehen von den folgenden chemischen Voraussetzungen aus:

1. Auf beiden Seiten der Reaktionsgleichung hat es gleich viele Atome jeder Sorte.
2. Die stöchiometrischen Koeffizienten sind ganzzahlig und minimal.

Der erste Punkt lässt sich mit Hilfe von linearen Gleichungssystemen lösen. Der zweite Punkt kann oftmals schwieriger sein und wird beispielsweise mit Hilfe von Plausibilitätsargumenten gelöst.

Wenden wir nun Punkt 1 auf die obige chemische Reaktion an. Wir haben:

Atom	Anzahl vorher	Anzahl nachher
C	a	d
Fe	$2b$	c
O	$3b$	$2d$

Vor und nach der Reaktion muss Gleichheit gelten. Dies führt auf das lineare Gleichungssystem:

$$\begin{aligned}
 a &= d \\
 2b &= c \\
 3b &= 2d
 \end{aligned} \tag{1.42}$$

oder

$$\begin{array}{rcl}
 a & - & d = 0 \\
 2b & - & c = 0 \\
 3b & - & 2d = 0
 \end{array}$$

Wir stellen fest, dass dieses lineare Gleichungssystem 4 Unbekannte, jedoch nur 3 Gleichungen besitzt. Dies liegt hier daran, dass wir die anfängliche Anzahl von Molekülen (d.h., vor der Reaktion) nicht festgelegt haben. Trotzdem wenden wir das Gaussverfahren auf das gegebene lineare Gleichungssystem an. In der erweiterten Matrixform haben wir:

$$\begin{array}{l}
 (I) \\
 (II) \\
 (III)
 \end{array}
 \left(\begin{array}{cccc|c}
 1 & 0 & 0 & -1 & 0 \\
 0 & 2 & -1 & 0 & 0 \\
 0 & \boxed{3} & 0 & -2 & 0
 \end{array} \right).$$

In der zweiten und dritten Zeile ist der erste Eintrag jeweils 0, d.h., der erste Eliminationsschritt ist bereits vollzogen. Im zweiten Schritt lassen wir die erste und zweite Zeile unberührt und eliminieren den zweiten Eintrag in der dritten Zeile:

$$\begin{array}{l} (I) \\ (II) \\ (III) - \frac{3}{2} \cdot (II) \end{array} \quad \left(\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 2 & -1 & 0 & 0 \\ 0 & 0 & \frac{3}{2} & -2 & 0 \end{array} \right).$$

Die vierte Unbekannte d ist frei wählbar. Wir führen deshalb einen Parameter, beispielsweise s , ein:

$$d = s.$$

Einsetzen in (III) ergibt:

$$\frac{3}{2}c - 2s = 0 \quad \Rightarrow \quad c = \frac{4s}{3}.$$

Dann erhalten wir mit Gleichung (II) :

$$2b - \frac{4s}{3} = 0 \quad \Rightarrow \quad b = \frac{2s}{3}.$$

Schliesslich folgt aus (I) :

$$a - s = 0 \quad \Rightarrow \quad a = s.$$

Da der Parameter s beliebig gewählt werden darf, hat das lineare Gleichungssystem (1.42) unendlich viele Lösungen. Beispielsweise erhalten wir für $s = 15$:

$$a = 15, \quad b = 10, \quad c = 20, \quad d = 15.$$

Nun wollen wir noch Punkt 2 in den obigen Voraussetzungen erfüllen. Dazu suchen wir eine ganzzahlige, minimale Lösung. Dies ist dann der Fall, wenn $s = 3$, d.h.

$$a = 3, \quad b = 2, \quad c = 4, \quad d = 3$$

ist die gesuchte Lösung. Nochmals unterstreichen wir, dass das ursprüngliche Gleichungssystem (1.42) unendlich viele (auch nicht-ganzzahlige Lösungen) besitzt. Ausserdem sei noch die wichtige Tatsache erwähnt, dass es bei einer chemischen Reaktion auch *mehrere* ganzzahlige, minimale Lösungen geben kann (was in der obigen Situation allerdings nicht zutrifft). Es kann sogar sein, dass gar keine Lösung existiert; dann ist die Reaktionsgleichung falsch gestellt und die chemische Reaktion läuft auf eine andere Weise ab. Wir werden dazu weitere Beispiele in den Übungen betrachten. \diamond

Beispiel 1.43 Betrachten wir zum Üben noch das lineare Gleichungssystem

$$\begin{array}{rccccrcr} x_1 & + & 2x_2 & - & 5x_3 & - & 2x_4 & = & 3 \\ 2x_1 & + & x_2 & + & x_3 & + & 4x_4 & = & 3 \\ -2x_1 & - & 5x_2 & + & x_3 & - & x_4 & = & -7 \end{array} \quad (1.44)$$

Die erweiterte Matrix ist gegeben durch

$$\begin{array}{l} (I) \\ (II) \\ (III) \end{array} \left(\begin{array}{cccc|c} 1 & 2 & -5 & -2 & 3 \\ 2 & 1 & 1 & 4 & 3 \\ -2 & -5 & 1 & -1 & -7 \end{array} \right).$$

Auch hier können wir das Gauss-Verfahren anwenden.

$$\begin{array}{l} \rightsquigarrow \\ \rightsquigarrow \\ \rightsquigarrow \end{array} \begin{array}{l} (I) \\ (II) \\ (III) \end{array} \left(\begin{array}{cccc|c} 1 & 2 & -5 & -2 & 3 \\ \boxed{2} & 1 & 1 & 4 & 3 \\ \boxed{-2} & -5 & 1 & -1 & -7 \end{array} \right)$$

$$\begin{array}{l} \rightsquigarrow \\ \rightsquigarrow \end{array} \begin{array}{l} (II) - 2 \cdot (I) \\ (III) - (-2) \cdot (I) \end{array} \left(\begin{array}{cccc|c} 1 & 2 & -5 & -2 & 3 \\ 0 & -3 & 11 & 8 & -3 \\ 0 & \boxed{-1} & -9 & -5 & -1 \end{array} \right)$$

$$\rightsquigarrow \begin{array}{l} (III) - \frac{1}{3} \cdot (II) \end{array} \left(\begin{array}{cccc|c} 1 & 2 & -5 & -2 & 3 \\ 0 & -3 & 11 & 8 & -3 \\ 0 & 0 & -\frac{38}{3} & -\frac{23}{3} & 0 \end{array} \right)$$

Die letzte Gleichung lautet

$$-\frac{38}{3}x_3 - \frac{23}{3}x_4 = 0.$$

Daraus folgt

$$x_3 = -\frac{23}{38}x_4.$$

Die Unbekannte x_4 ist frei wählbar und wird durch einen Parameter s ersetzt:

$$x_4 = s, \quad x_3 = -\frac{23}{38}s.$$

Einsetzen in Gleichung (II) des transformierten Systems liefert

$$-3x_2 + 11 \cdot \left(-\frac{23}{38}s\right) + 8s = -3 \quad \Rightarrow \quad x_2 = \frac{17}{38}s + 1.$$

Aus der ersten Gleichung ergibt sich weiter

$$x_1 + 2 \cdot \left(\frac{17}{38}s + 1\right) + (-5) \cdot \left(-\frac{23}{38}s\right) - 2s = 3 \quad \Rightarrow \quad x_1 = 1 - \frac{73}{38}s.$$

Durch Einsetzen in das ursprüngliche lineare Gleichungssystem kann nun überprüft werden, dass die gefundenen x_1, x_2, x_3, x_4 für jedes beliebige s eine Lösung darstellen. Das System hat folglich unendlich viele Lösungen. \square

Bemerkung 1.45 Der Begriff des Rangs lässt sich einfach auf Matrizen übertragen, welche eine ungleiche Anzahl von Zeilen und Spalten haben. Allerdings sind die Begriffe "regulär" und

“singulär” dann nicht mehr sinnvoll definiert. Auch hier soll ein Beispiel Klarheit schaffen: Wir bestimmen den Rang der Matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & 5 \\ 1 & 2 & 4 \\ 3 & 1 & 7 \\ -4 & 0 & -8 \\ 2 & 6 & 10 \end{pmatrix}$$

Durchführen des Gaußalgorithmus ergibt nach dem ersten Schritt (erste Zeile unverändert) die Matrix

$$\begin{pmatrix} 1 & 3 & 5 \\ 0 & -1 & -1 \\ 0 & -8 & -8 \\ 0 & 12 & 12 \\ 0 & 0 & 0 \end{pmatrix},$$

und nach dem zweiten Schritt (erste und zweite Zeile unverändert)

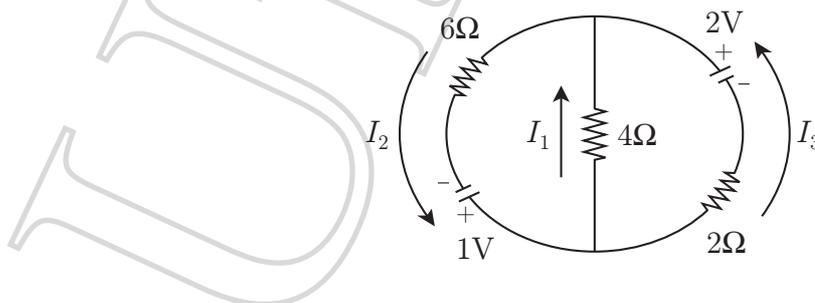
$$\begin{pmatrix} 1 & 3 & 5 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Hier stoppt das Gaußverfahren. Es werden 3 Nullzeilen sichtbar und somit gilt

$$\begin{aligned} \text{rang}(\mathbf{A}) &= (\text{Anzahl Zeilen von } \mathbf{A}) - (\text{Anzahl Nullzeilen nach vollständiger Gausselimination}) \\ &= 5 - 3 = 2. \end{aligned}$$

1.8 Übungsaufgaben

- 1.1. Bestimmen Sie das Gleichungssystem für die Ströme I_1 , I_2 und I_3 , welches zum folgenden Stromkreis gehört:



1.2. Bestimmen Sie die Lösungsmenge der folgenden Gleichungssysteme mit Gauss-Elimination.

$$\begin{array}{l} x + 2y = 3 \\ 3x + 4y = 5 \end{array} \qquad \begin{array}{l} 2u^2 - v^2 = 0 \\ u^2 + 2v^2 = 10 \end{array}$$

1.3. Verschiedene lineare Gleichungssysteme mit jeweils drei Unbekannten wurden in Matrixform mit dem Gauss-Verfahren bearbeitet. Finden Sie für jedes der Systeme die Lösungsmenge in Parameterform und beschreiben Sie diese in Worten.

$$\left(\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right), \quad \left(\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & 3 & -1 \\ 0 & 0 & 0 & 4 \end{array} \right), \quad \left(\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array} \right).$$

1.4. Bestimmen Sie je die Lösungsmenge der folgenden Gleichungssysteme.

$$\begin{array}{l} r + s + t = 1 \\ r + 2s + 4t = 2 \\ r + 3s + 9t = 6 \end{array} \qquad \begin{array}{l} u + 2v = 1 \\ v + 2w = 2 \\ w + 2u = 3. \end{array}$$

1.5. Durch die Punkte $(1|1)$ und $(-2|3)$ in der x_1 - x_2 -Ebene soll eine Ellipse der Form

$$ax_1^2 + bx_2^2 = 1$$

gelegt werden. Bestimmen Sie die Koeffizienten a und b . Stellen Sie dazu zunächst ein Gleichungssystem für a und b auf und lösen Sie es dann mit Gauss-Elimination.

1.6. Bekanntlich lässt sich jeder Kreis in der xy -Ebene durch eine Gleichung der Art

$$x^2 + y^2 + ax + by + c = 0$$

beschreiben. Bestimmen Sie die Gleichung des Kreises, welcher durch die drei Punkte $P_1 = (7, -2)$, $P_2 = (5, -4)$ und $P_3 = (-3, 4)$ verläuft.

1.7. a) Gegeben sind die beiden Geraden g_1 und g_2 in der x_1 - x_2 -Ebene:

$$\begin{array}{l} g_1: \quad 2x_1 - x_2 = 5 \\ g_2: \quad -5x_1 + 3x_2 = 2 \end{array}$$

Bestimmen Sie den Schnittpunkt von g_1 und g_2 .

b) Gegeben sind die beiden Geraden

$$\begin{array}{l} h_1: \quad x_1 - x_2 = 3 \\ h_2: \quad 2x_1 - 2x_2 = k \end{array}$$

Für welche Werte des Parameters k

1 LINEARE GLEICHUNGSSYSTEME

- i) sind h_1 und h_2 *identische Geraden*?
- ii) haben h_1 und h_2 *keinen* Schnittpunkt? Was bedeutet dies geometrisch?
- iii) haben h_1 und h_2 *genau einen* Schnittpunkt?

1.8. Lösen Sie die folgenden linearen Gleichungssysteme mit dem Gauss-Verfahren.

$$\begin{aligned} x_1 + 2x_2 + 3x_3 &= 12 \\ 4x_1 + 5x_2 + 6x_3 &= 24 \\ 7x_1 + 8x_2 + 10x_3 &= 41 \end{aligned}$$

$$\begin{aligned} a + b - c + 4d &= 2 \\ 2a - 3b - 2c + 9d &= 5 \\ a - 4b - c + 3d &= 1 \\ 2a + 12b + c + 11d &= 16. \end{aligned}$$

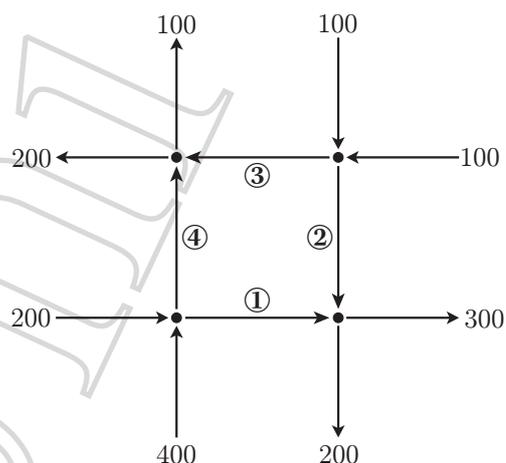
Bringen Sie die Gleichungssysteme zuerst in Rechtsobere-Dreiecks-Form und lösen Sie sie dann durch Rückwärtseinsetzen.

1.9. Wir betrachten die drei Alkohol-Ammoniak-Wasser-Mischungen F_1 , F_2 und F_3 mit den folgenden Konzentrationen:

	Alkohol	Ammoniak	Wasser
F_1	85%	10%	5%
F_2	75%	15%	10%
F_3	60%	20%	20%

Hergestellt werden soll 1 Liter eines Gemischs mit 70% Alkohol, 16% Ammoniak und 14% Wasser. Wie gross müssen die Anteile der Flüssigkeiten F_1 , F_2 bzw. F_3 sein? Stellen Sie ein lineares Gleichungssystem auf und lösen Sie es mit dem Gauss-Verfahren.

1.10. Gegeben sei das folgende Abwasserröhrensystem.



Die Pfeile zeigen die jeweilige Flussrichtung des Wassers an; die Zahlen stehen für die Durchflussmenge in Liter pro Sekunde.

- a) Wie gross müssen die Kapazitäten (in Liter/Sekunde) der Röhren ①, ② und ④ sein, wenn die Röhre ③ geschlossen ist?
 b) Was passiert, wenn die Röhre ④ versperrt ist?

Stellen Sie für (a) und (b) je ein geeignetes lineares Gleichungssystem auf und lösen Sie es.

- 1.11. Für welche Zahlen p werden die folgenden Matrizen singular?

$$\begin{pmatrix} 1-p & 2 \\ 3 & 4 \end{pmatrix}, \quad \begin{pmatrix} 1-p & 2 \\ 3 & 4-p \end{pmatrix}.$$

- 1.12. Bestimmen Sie jeweils den Rang der folgenden Matrizen:

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \\ 3 & 2 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & -5 & 3 & 1 \\ 3 & -2 & 2 & 2 & 2 \\ -1 & -1 & 1 & 2 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 2 & 3 & 2 & 1 \\ 3 & 7 & 2 & -2 \\ 4 & -3 & 2 & 9 \\ 8 & 4 & -4 & 0 \end{pmatrix}.$$

- 1.13. Gegeben sind die folgenden drei Ebenen im dreidimensionalen Raum.

$$\begin{aligned} E_1: & \quad x - 5y + z = 2 \\ E_2: & \quad 5x + 2y = 1 \\ E_3: & \quad -3x + 15y - 3z = -6 \end{aligned}$$

- a) Bestimmen Sie die Schnittmenge dieser drei Ebenen.
 b) Bestimmen Sie die Schnittmenge der Ebene E_1 mit der xy -Ebene.

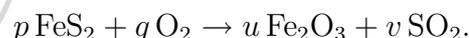
Geben Sie jeweils eine Parameterdarstellung der Schnittmenge an und interpretieren Sie das Resultat geometrisch.

- 1.14. Von drei Alkohollösungen seien folgende Angaben bekannt.

Flüssigkeit	Alkoholgehalt	Preis pro Liter
F_1	3%	4
F_2	4%	6
F_3	6%	8

- a) Wie lassen sich die Flüssigkeiten mischen, um eine 5%-ige Lösung zu erhalten?
 b) Welche davon ist die Mischung mit dem preisgünstigsten Literpreis, und wie gross ist ihr Preis pro Liter?

- 1.15. Pyrit FeS_2 und Sauerstoff O_2 verbrennen zu Schwefeldioxid SO_2 und Roteisenstein Fe_2O_3 . Finden Sie die stöchiometrischen Koeffizienten p, q, u, v für diese Reaktion:

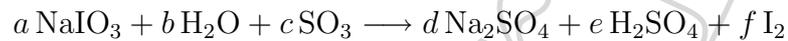


1.16. Betrachten Sie das Auflösen einer Brausetablette mit Zitronensäure:



Bestimmen Sie die stöchiometrischen Koeffizienten a , b , c , d und e nach den bekannten Regeln. Lösen Sie das entsprechende Gleichungssystem mit Gauss-Elimination. Bestimmen Sie zunächst die *allgemeine* Lösung und daraus dann die minimalen ganzzahligen Koeffizienten.

1.17. Betrachten Sie folgende "Knacknuss":



Was geschieht hier, und wie lässt sich das Problem lösen?

2

Vektorräume

2.1 Der Vektorraum \mathbb{R}^n

Mit \mathbb{R} bezeichnen wir die Menge aller reeller Zahlen. Dazu gehören Zahlen wie

$$2, \quad 0, \quad -5, \quad \frac{2}{3}, \quad \sqrt{2}, \quad \pi, \quad \text{etc.}$$

Diese Zahlen lassen sich geometrisch auf der Zahlengeraden darstellen:

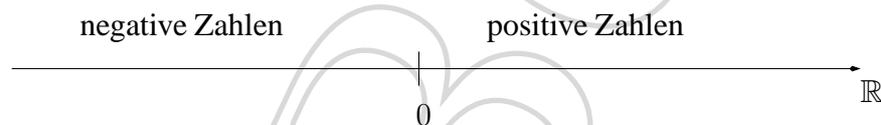


Abbildung 2.1: Zahlengerade.

In der ebenen Geometrie tritt häufig die Menge \mathbb{R}^2 auf. Dort ist dies die Menge aller *Punkte* im zwei-dimensionalen Koordinatensystem. Identifiziert man einen solchen Punkt P mit Koordinaten (x_1, x_2) in der Ebene mit dem zugehörigen *Ortsvektor*

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

d.h., mit dem Vektor der vom Ursprung des Koordinatensystems zum Punkt P zeigt, so lässt sich \mathbb{R}^2 auch als die Menge aller Ortsvektoren, oder noch allgemeiner, der Menge aller Vektoren der Ebene, definieren; vgl. Abbildung 2.3. Wie in Kapitel 1 erwähnt, verwenden wir in der Mathematik zur Beschreibung von Mengen typischerweise die Notation $\{\dots\}$. Die Menge \mathbb{R}^2 können wir dann wie folgt beschreiben:

$$\mathbb{R}^2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : x_1, x_2 \text{ sind zwei beliebige reelle Zahlen} \right\}.$$

Analog entspricht die Menge \mathbb{R}^3 dem dreidimensionalen Raum.

Allgemeiner definieren wir nun die Menge \mathbb{R}^n als die *Menge aller Vektoren mit n Einträgen* (auch Komponenten genannt):

$$\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} : x_1, x_2, \dots, x_n \text{ sind } n \text{ beliebige reelle Zahlen} \right\}.$$

Beispiel 2.1

1. Der Vektor

$$\begin{pmatrix} -3 \\ 2.4 \\ 4 \\ 0 \end{pmatrix}$$

hat vier Komponenten. Er gehört daher zur Menge \mathbb{R}^4 aller Vektoren mit vier Komponenten.

2. Angenommen, ein Wissenschaftler führt ein Experiment 20 mal durch und macht jedesmal eine Messung (zum Beispiel eine Temperatur). Dann können wir die Gesamtheit der Resultate auffassen als einen Vektor im Raum \mathbb{R}^{20} , dessen Einträge die 20 Ergebnisse aus den Experimenten sind.
3. Ein optisches Signal sei eine Mischung aus drei Farben: rot, blau und grün. Jede Farbe kommt in dem Signal mit einer gewissen Intensität vor. Dies entspricht drei Zahlen R , B und G . Sie können in einem Vektor aus \mathbb{R}^3 angegeben werden:

$$\begin{pmatrix} R \\ B \\ G \end{pmatrix}.$$

So entspricht der Vektor

$$\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

einem Signal, welches gleich viel rot und blau enthält, jedoch kein grün. Dies entspricht also in etwa der Farbe lila.

□

Diese Beispiele zeigen, dass Vektoren nicht immer einen Bezug zur Geometrie haben müssen, sondern ganz allgemeine “Datencontainer” sein können.

2.1.1 Grundoperationen für Vektoren

Für Vektoren in \mathbb{R}^n sind einige bekannte *Operationen* definiert, wie beispielsweise die *Addition* $+$, die *Subtraktion* $-$, oder die *Skalarmultiplikation mit einer Zahl*.

Beispiel 2.2

a) Betrachten wir die zwei Vektoren

$$\mathbf{v} = \begin{pmatrix} 3 \\ -1 \\ 6 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} 4 \\ -1 \\ 1 \end{pmatrix}$$

in \mathbb{R}^3 . Dann bilden wir die Summe dieser beiden Vektoren durch das jeweilige Addieren der entsprechenden Einträge. Das Resultat ist wiederum ein Vektor in \mathbb{R}^3 :

$$\mathbf{v} + \mathbf{w} = \begin{pmatrix} 3+4 \\ -1+(-1) \\ 6+1 \end{pmatrix} = \begin{pmatrix} 7 \\ -2 \\ 7 \end{pmatrix},$$

Genau analog berechnen wir die Differenz:

$$\mathbf{v} - \mathbf{w} = \begin{pmatrix} 3-4 \\ -1-(-1) \\ 6-1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 5 \end{pmatrix},$$

b) Sei

$$\mathbf{u} = \begin{pmatrix} 3 \\ 3 \\ -1 \\ 0 \end{pmatrix}$$

ein Vektor in \mathbb{R}^4 und $\alpha = 5$ eine Zahl. Dann wird die Skalarmultiplikation von \mathbf{u} mit α wiederum eintragsweise durchgeführt:

$$\alpha \cdot \mathbf{v} = 5 \cdot \mathbf{v} = \begin{pmatrix} 5 \cdot 3 \\ 5 \cdot 3 \\ 5 \cdot (-1) \\ 5 \cdot 0 \end{pmatrix} = \begin{pmatrix} 15 \\ 15 \\ -5 \\ 0 \end{pmatrix}.$$

Die Skalarmultiplikation ist *nicht zu verwechseln* mit dem Skalarprodukt (welches wir im folgenden besprechen werden)!

□

Die obigen Beispiele zeigen, dass die Menge \mathbb{R}^n unter anderen die folgenden Eigenschaften besitzt:

(E1) Falls zwei Vektoren v und w zu \mathbb{R}^n gehören, so befindet sich auch ihre Summe

$$v + w$$

in \mathbb{R}^n . Gleiches gilt für die Differenz zweier Vektoren.

(E2) Gegeben ein Vektor v in \mathbb{R}^n und eine beliebige reelle Zahl α . Dann ist

$$\alpha \cdot v$$

ebenfalls ein \mathbb{R}^n -Vektor (wir bemerken, dass der Vektor $\alpha \cdot v$ die gleiche Richtung hat wie v ; falls $\alpha > 0$, dann hat er sogar dieselbe Orientierung wie v , für $\alpha < 0$ dreht er sich um 180° , wobei die Richtung aber dennoch gleich bleibt).

Die Menge \mathbb{R}^n ist ein sogenannter *Vektorraum*. Alle Vektorräume haben (nebst weiteren) die Eigenschaften (E1) und (E2).

2.1.2 Linearkombinationen

Aus den vorherigen Beobachtungen resultiert eine wichtige Verallgemeinerung: Angenommen, wir betrachten eine Menge von Vektoren v_1, v_2, \dots, v_p in \mathbb{R}^n . Ausserdem seien beliebige Zahlen $\alpha_1, \alpha_2, \dots, \alpha_p$ gegeben. Dann sind die Vektoren

$$\alpha_1 \cdot v_1, \quad \alpha_2 \cdot v_2, \quad \dots, \quad \alpha_p \cdot v_p$$

ebenfalls Vektoren in \mathbb{R}^n . Überdies gehört auch deren Summe

$$\alpha_1 \cdot v_1 + \alpha_2 \cdot v_2 + \dots + \alpha_p \cdot v_p \tag{2.3}$$

zum Raum \mathbb{R}^n .

Ein Ausdruck der Form (2.3) nennt man **Linearkombination** der Vektoren v_1, v_2, \dots, v_p . Zusammenfassend können wir sagen: Jede beliebige Linearkombination von Vektoren aus \mathbb{R}^n ist wiederum ein Vektor in \mathbb{R}^n

Beispiel 2.4

a) Seien

$$v_1 = \begin{pmatrix} -3 \\ 2 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix}$$

zwei Vektoren in \mathbb{R}^3 . Dann ist beispielsweise der Vektor

$$2 \cdot v_1 + (-5) \cdot v_2 = \begin{pmatrix} -6 \\ 4 \\ 0 \end{pmatrix} + \begin{pmatrix} -5 \\ -20 \\ -5 \end{pmatrix} = \begin{pmatrix} -11 \\ -16 \\ -5 \end{pmatrix}$$

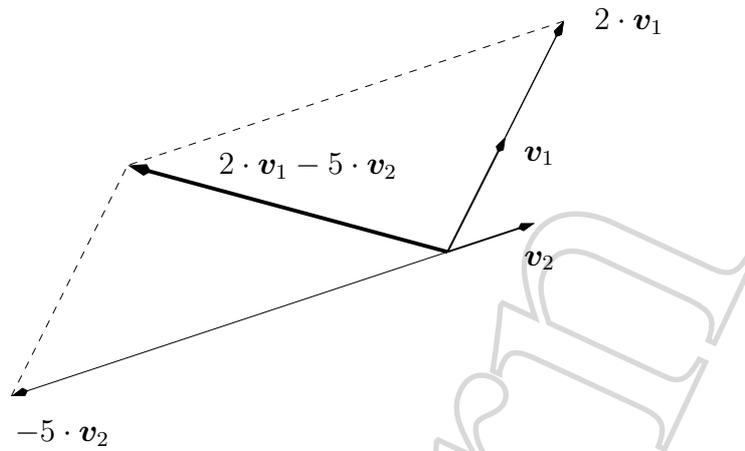


Abbildung 2.2: Linearkombination zweier Vektoren.

eine Linearkombination von v_1 und v_2 . Geometrisch bedeutet dies, dass der Vektor

$$\begin{pmatrix} -11 \\ -16 \\ -5 \end{pmatrix}$$

in der Ebene durch den Ursprung liegt, die durch die Vektoren v_1 und v_2 gebildet wird; vgl. Abbildung 2.2.

b) Ist der Vektor

$$\begin{pmatrix} -2 \\ 6 \\ 2 \end{pmatrix}$$

eine Linearkombination von v_1 und v_2 ? Um diese Frage mit “ja” beantworten zu können, müssen wir zwei Zahlen α_1 und α_2 finden, sodass

$$\alpha_1 v_1 + \alpha_2 v_2 = \begin{pmatrix} -2 \\ 6 \\ 2 \end{pmatrix}.$$

Ausgeschrieben heisst dies

$$\alpha_1 \cdot \begin{pmatrix} -3 \\ 2 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 6 \\ 2 \end{pmatrix},$$

oder nach komponentenweisem Ausmultiplizieren

$$\begin{aligned} -3\alpha_1 + \alpha_2 &= -2 \\ 2\alpha_1 + 4\alpha_2 &= 6 \\ \alpha_2 &= 2 \end{aligned} \quad (2.5)$$

Dies ist ein überbestimmtes lineares Gleichungssystem für die Unbekannten α_1 und α_2 . Aus der dritten Gleichung folgt direkt, dass

$$\alpha_2 = 2.$$

Wir setzen dies in die zweite Gleichung ein:

$$2\alpha_1 + 8 = 6 \quad \Rightarrow \quad \alpha_1 = -1.$$

Nun müssen wir noch überprüfen, ob auch die erste Gleichung für diese Lösung erfüllt ist:

$$-3\alpha_1 + \alpha_2 = 3 + 2 = 5 \neq -2.$$

Die erste Gleichung ist nicht erfüllt! Somit hat das Gleichungssystem (2.5) keine Lösung. Der Vektor

$$\begin{pmatrix} -2 \\ 6 \\ 2 \end{pmatrix}$$

lässt sich somit *nicht* als Linearkombination von v_1 und v_2 darstellen.

□

2.2 Euklidische Norm und Skalarprodukt

2.2.1 Längenmessung mit Normen in \mathbb{R}^n

Jede Zahl α auf der Zahlengerade \mathbb{R} hat einen bestimmten *positiven* Abstand vom Nullpunkt. Diesen Abstand nennen wir **Absolutbetrag** oder **Betrag** von α . Er wird mit $|\alpha|$ bezeichnet¹. Beispielsweise haben wir:

$$|5| = 5, \quad |-3| = 3, \quad |-\pi| = \pi, \quad \text{etc.}$$

Für einen Punkt P in \mathbb{R}^n mit zugehörigem Ortsvektor (vgl. Abbildung 2.3)

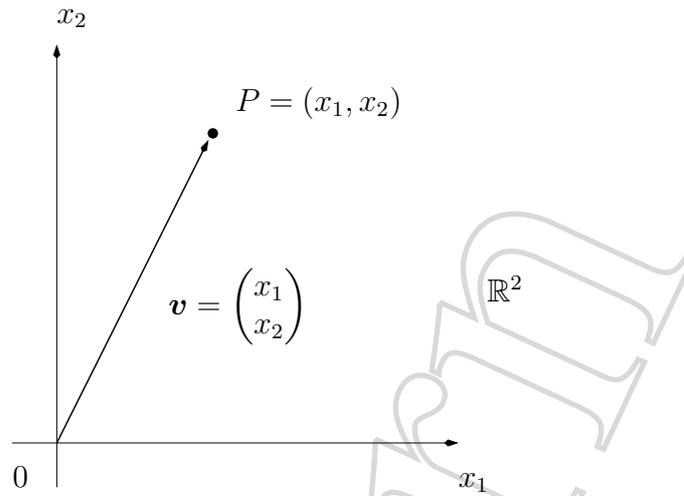
$$\mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{2.6}$$

lässt sich ebenfalls der Abstand vom Ursprung messen. Genauer definieren wir hier mit Hilfe des Satzes von Pythagoras, genau wie in \mathbb{R}^2 und \mathbb{R}^3 , die (Euklidische) **Länge** oder auch **Norm** des Ortsvektors \mathbf{v} durch

$$\|\mathbf{v}\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}.$$

¹Genauer gilt:

$$|\alpha| = \begin{cases} \alpha & \text{falls } \alpha \geq 0 \\ -\alpha & \text{falls } \alpha < 0 \end{cases}.$$

Abbildung 2.3: Punkt P und Ortsvektor v in \mathbb{R}^2 .

Beispiel 2.7 Sei

$$v = \begin{pmatrix} -2 \\ 3 \\ 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}$$

ein Vektor in \mathbb{R}^6 . Dann berechnet sich seine Länge als

$$\|v\| = \sqrt{(-2)^2 + 3^2 + 1^2 + 0^2 + (-1)^2 + 1^2} = \sqrt{16} = 4.$$

□

Der Begriff der Norm lässt sich verallgemeinern. Insbesondere gibt es auch andere Möglichkeiten, Längen von Vektoren in \mathbb{R}^n zu messen². Ganz allgemein haben Normen immer die folgenden drei Eigenschaften:

²Beispielsweise lassen sich für beliebige reelle Zahlen $p \geq 1$ die sogenannten p -Normen definieren: Sei v ein Vektor wie in (2.6), dann schreiben wir:

$$\|v\|_p = (|x_1|^p + |x_2|^p + \dots + |x_n|^p)^{\frac{1}{p}}.$$

Diese Normen haben die obigen drei Eigenschaften. Betrachten wir den Vektor v aus Beispiel 2.7 so berechnen wir

$$\|v\|_1 = |-2| + |3| + |1| + |0| + |-1| + |1| = 2 + 3 + 1 + 0 + 1 + 1 = 8$$

oder

$$\|v\|_4 = (|-2|^4 + |3|^4 + |1|^4 + |0|^4 + |-1|^4 + |1|^4)^{\frac{1}{4}} = 100^{\frac{1}{4}} = \sqrt[4]{100} \approx 3.1623.$$

Offenbar entspricht $p = 2$ (d.h., die 2-Norm) der Euklidischen Norm.

1. Für einen beliebigen Vektor \mathbf{v} in \mathbb{R}^n und eine beliebige Zahl α gilt die Gleichheit

$$\|\alpha \cdot \mathbf{v}\| = |\alpha| \|\mathbf{v}\|.$$

2. Für zwei beliebige Vektoren \mathbf{v}, \mathbf{w} aus \mathbb{R}^n gilt immer

$$\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|.$$

Diese Eigenschaft heisst *Dreiecksungleichung*.

3. Falls $\|\mathbf{v}\| = 0$, dann folgt $\mathbf{v} = \mathbf{0}$.

Skalierung: Jeder Vektor lässt sich in einen Vektor der Länge 1 “umwandeln”. Betrachten wir als Beispiel den Vektor

$$\mathbf{v} = \begin{pmatrix} 3 \\ 3 \\ -1 \end{pmatrix}.$$

Er hat die Länge

$$\|\mathbf{v}\| = \sqrt{3^2 + 3^2 + (-1)^2} = \sqrt{19}.$$

Teilen wir den Vektor \mathbf{v} durch seine Länge, dann hat der Vektor

$$\mathbf{w} = \frac{1}{\|\mathbf{v}\|} \cdot \mathbf{v} = \frac{1}{\sqrt{19}} \cdot \begin{pmatrix} 3 \\ 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 3/\sqrt{19} \\ 3/\sqrt{19} \\ -1/\sqrt{19} \end{pmatrix}$$

die Länge 1, denn

$$\|\mathbf{w}\| = \sqrt{\left(\frac{3}{\sqrt{19}}\right)^2 + \left(\frac{3}{\sqrt{19}}\right)^2 + \left(\frac{-1}{\sqrt{19}}\right)^2} = \sqrt{\frac{9}{19} + \frac{9}{19} + \frac{1}{19}} = \sqrt{\frac{19}{19}} = 1.$$

Also: Dividieren eines Vektors durch seine eigene Länge ergibt einen Vektor, der die gleiche Richtung und Orientierung hat wie dieser Vektor, jedoch auf Länge 1 *skaliert* ist. Ganz allgemein entspricht die Skalarmultiplikation eines Vektors mit einer Zahl einer Längenänderung des Vektors; dieser Vorgang heisst **Skalierung**.

2.2.2 Skalarprodukt

Für zwei Vektoren

$$\mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

definieren wir das folgende **Skalarprodukt**:

$$\langle \mathbf{v}, \mathbf{w} \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n. \quad (2.8)$$

Das Skalarprodukt ist immer eine *Zahl*.

Beispiel 2.9 Das Skalarprodukt der beiden Vektoren

$$\mathbf{v} = \begin{pmatrix} 3 \\ -2 \\ 0 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} 4 \\ 1 \\ -1 \end{pmatrix}$$

ist die Zahl

$$\langle \mathbf{v}, \mathbf{w} \rangle = 3 \cdot 4 + (-2) \cdot 1 + 0 \cdot (-1) = 12 - 2 + 0 = 10.$$

□

Anwendung 2.10 (Datenübermittlung I)

Computerprogramme zur Datenübermittlung versenden oftmals Zusatzinformationen zu der zu übermittelnden Nachricht. Diese können vom Empfänger zur Überprüfung der Richtigkeit einer Nachricht verwendet werden.

Ein Beispiel ist die “International Standard Book Number” oder kurz ISBN-Nummer zur Identifikation von Büchern. Dies ist eine 10-ziffrige Zahl, die für jedes Buch verschieden ist und ein solches eindeutig identifiziert. Die ersten 9 Ziffern der ISBN-Nummer sind aufgeteilt in 3 Zahlengruppen – die erste Gruppe steht für die Gruppe oder das Land, aus dem das Buch stammt, die zweite Gruppe entspricht dem Verlag, und die dritte Gruppe bezieht sich auf den Buchtitel. Die zehnte und letzte Zahl ist eine sogenannte Prüfziffer (Englisch “check digit”), die aus den ersten 9 Zahlen berechnet wird. Sie wird benutzt, um die Richtigkeit einer elektronischen Übermittlung einer ISBN-Nummer (beispielsweise via Internet) zu überprüfen.

Der Algorithmus zur Berechnung der Prüfziffer bei ISBN-Nummer funktioniert wie folgt:

1. Speichere die 9 ersten Ziffern der ISBN-Nummer als Einträge in einem Vektor \mathbf{a} in \mathbb{R}^9 . Definiere weiter den Vektor

$$\mathbf{b} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ \vdots \\ 9 \end{pmatrix}$$

in \mathbb{R}^9 .

2. Berechne das Skalarprodukt $\langle \mathbf{a}, \mathbf{b} \rangle$.
3. Teile $\langle \mathbf{a}, \mathbf{b} \rangle$ durch 11; dabei entsteht ein Rest r zwischen 0 und 10 (um zweistellige Reste zu vermeiden, wird 10 durch X ersetzt).
4. Der Rest r ist die Prüfziffer.

Betrachten wir als Beispiel das bekannte Schweizer Kinderbuch “Mein Name ist Eugen” mit der ISBN-Nummer

3290114708.

Die Prüfziffer, also die zehnte Ziffer, ist 8. Wir rechnen sie mit dem obigen Algorithmus nach:

$$\mathbf{a} = \begin{pmatrix} 3 \\ 2 \\ 9 \\ 0 \\ 1 \\ 1 \\ 4 \\ 7 \\ 0 \end{pmatrix}.$$

Dann gilt:

$$\langle \mathbf{a}, \mathbf{b} \rangle = 3 \cdot 1 + 2 \cdot 2 + 9 \cdot 3 + 0 \cdot 4 + 1 \cdot 5 + 1 \cdot 6 + 4 \cdot 7 + 7 \cdot 8 + 0 \cdot 9 = 129.$$

Weiter haben wir

$$129 : 11 = 11 \text{ Rest } 8,$$

also $r = 8$, und dies ist genau die Prüfziffer. \diamond

Im folgenden fassen wir einige wichtige Eigenschaften des Skalarproduktes zusammen. Hier sind $\mathbf{u}, \mathbf{v}, \mathbf{w}$ beliebige Vektoren aus \mathbb{R}^n , α ist eine beliebige Zahl, und $\|\cdot\|$ bezeichnet die *Euklidische Norm* in \mathbb{R}^n .

1. Symmetrie:

$$\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle.$$

2. Bilinearität:

$$\langle \alpha \cdot \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, \alpha \cdot \mathbf{w} \rangle = \alpha \langle \mathbf{v}, \mathbf{w} \rangle.$$

und

$$\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle, \quad \langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle.$$

3. Es gilt die Gleichheit

$$\langle \mathbf{v}, \mathbf{w} \rangle = \|\mathbf{v}\| \|\mathbf{w}\| \cos(\theta), \tag{2.11}$$

wobei θ der Zwischenwinkel³ der beiden Vektoren \mathbf{v} und \mathbf{w} ist; vgl. Abbildung 2.4. Wir betrachten zwei Spezialfälle:

³In diesem Text vereinbaren wir, dass Winkel immer im Gradmass gemessen werden, d.h. der Vollwinkel beträgt 360° . Wir bemerken jedoch, dass in der Mathematik häufig das Bogenmass verwendet, mit Vollwinkel 2π .

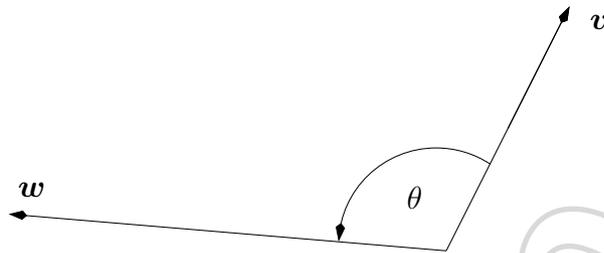


Abbildung 2.4: Zwischenwinkel zweier Vektoren.

- a) Falls $w = v$, dann gilt $\theta = 0^\circ$ und $\cos(0^\circ) = 1$. Daher

$$\langle v, v \rangle = \|v\|^2.$$

Diese Eigenschaft ergibt sich ebenso durch direktes Einsetzen in (2.8).

- b) Falls v und w senkrecht sind, dann ist $\cos(90^\circ) = 0$, und somit folgt aus (2.11), dass

$$\langle v, w \rangle = 0,$$

d.h., das Skalarprodukt zweier orthogonaler Vektoren ist immer gleich 0. Hier gilt auch der *Satz von Pythagoras*:

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

4. Da der Absolutbetrag von $\cos(\theta)$ immer kleiner gleich 1 ist, folgt aus der obigen Formel die sogenannte *Cauchy-Schwarz-Ungleichung*:

$$|\langle v, w \rangle| \leq \|v\| \|w\|.$$

5. Es gilt die *Parallelogrammidentität*

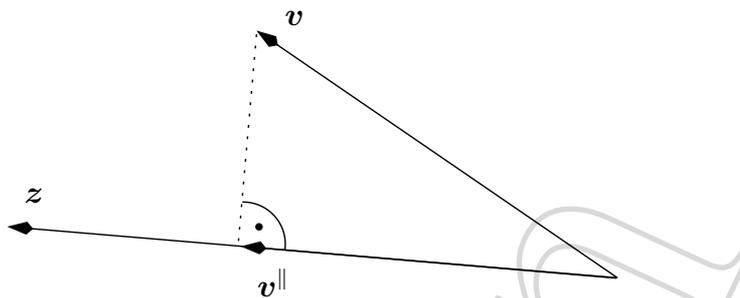
$$\langle v, w \rangle = \frac{1}{4} \|v + w\|^2 - \frac{1}{4} \|v - w\|^2.$$

2.2.3 Projektionseigenschaft

Betrachten wir zwei Vektoren v und z des \mathbb{R}^n . Wir wollen den Vektor v *senkrecht* auf den Vektor z projizieren; siehe Abbildung 2.5. Wie lässt sich der entstandene Vektor v^\parallel berechnen?

Sicherlich hat der gesuchte Vektor v^\parallel dieselbe Richtung wie der Vektor z . Er ist also ein skalares Vielfaches von z und lässt sich somit schreiben als

$$v^\parallel = \alpha \cdot z,$$

Abbildung 2.5: Projektion von v auf z .

wo α eine noch unbekannte Zahl ist. Ausserdem steht der Differenzvektor $v - v^{\parallel}$ senkrecht auf z . Aus den Eigenschaften des Skalarproduktes folgt dann:

$$0 = \langle v - v^{\parallel}, z \rangle = \langle v - \alpha \cdot z, z \rangle = \langle v, z \rangle - \alpha \langle z, z \rangle = \langle v, z \rangle - \alpha \|z\|^2.$$

Daher,

$$\alpha = \frac{\langle v, z \rangle}{\|z\|^2}.$$

Somit haben wir gezeigt:

Projizieren wir einen Vektor v senkrecht auf einen Vektor z , so lässt sich der projizierte Vektor v^{\parallel} berechnen als

$$v^{\parallel} = \frac{\langle v, z \rangle}{\|z\|^2} \cdot z. \quad (2.12)$$

Diese Formel nennt man **Projektionseigenschaft**.

Beispiel 2.13 Wir projizieren den Vektor

$$v = \begin{pmatrix} 11 \\ -3 \end{pmatrix}$$

auf den Vektor

$$z = \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Der projizierte Vektor berechnet sich dann mit der Projektionseigenschaft als

$$v^{\parallel} = \frac{\left\langle \begin{pmatrix} 11 \\ -3 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \end{pmatrix} \right\rangle}{\left\| \begin{pmatrix} 2 \\ -1 \end{pmatrix} \right\|^2} \cdot \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \frac{11 \cdot 2 + (-3) \cdot (-1)}{2^2 + (-1)^2} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \frac{25}{5} \cdot \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 10 \\ -5 \end{pmatrix}.$$

Das Resultat lässt sich in diesem Beispiel einfach grafisch nachprüfen. □

2.3 Unterräume von \mathbb{R}^n

In diesem Abschnitt wollen wir uns mit Teilmengen von \mathbb{R}^n beschäftigen.

2.3.1 Beispiele und Definition

Wir betrachten zunächst einige Beispiele.

Beispiel 2.14

- a) Es sei U die Menge aller Vektoren in \mathbb{R}^2 mit Länge 1:

$$U = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : \sqrt{x_1^2 + x_2^2} = 1 \right\}.$$

Dies ist eine Teilmenge von \mathbb{R}^2 .

- b) Betrachten wir die folgende Teilmenge von \mathbb{R}^3 :

$$V = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : 2x_1 + 4x_2 - 3x_3 = 0 \right\}.$$

- c) Schliesslich sei

$$W = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} : x_4 = 0 \right\} = \left\{ \text{alle Vektoren in } \mathbb{R}^4 \text{ der Form } \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ 0 \end{pmatrix} \right\}$$

eine Teilmenge von \mathbb{R}^4 .

Was bedeuten diese Mengen geometrisch? □

Erinnern wir uns an die wichtigen zwei Eigenschaften (E1) und (E2) des Vektorraums \mathbb{R}^n (siehe Seite 42). Sie besagen: Für zwei Vektoren in \mathbb{R}^n sind immer auch deren Summe und beliebige Skalierungen Vektoren in \mathbb{R}^n .

Wir stellen die folgende Frage: Gelten diese Eigenschaften auch für Teilmengen von \mathbb{R}^n ?
Genauer: Angenommen, es sind zwei beliebige Vektoren aus einer Teilmenge U von \mathbb{R}^n gegeben; sind dann deren Summe und beliebige Skalierungen ebenfalls Vektoren aus dieser Teilmenge U ?
Um Einblick in diese Frage zu erhalten, betrachten wir die obigen Beispiele a) und c):

2 VEKTORRÄUME

- a) Die Menge U besteht aus allen Vektoren des \mathbb{R}^2 , die Länge 1 haben. Zählen wir zwei Vektoren der Länge 1 zusammen, so kann es sein, dass deren Summe nicht mehr Länge 1 hat. Dies sehen wir am folgenden Beispiel:

$$\begin{aligned} \mathbf{v}_1 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} && \Rightarrow && \|\mathbf{v}_1\| = 1 \\ \mathbf{v}_2 &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} && \Rightarrow && \|\mathbf{v}_2\| = 1 \\ \mathbf{v}_1 + \mathbf{v}_2 &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} && \Rightarrow && \|\mathbf{v}_1 + \mathbf{v}_2\| = \sqrt{2} \neq 1. \end{aligned}$$

Somit folgt, dass die Teilmenge U die besagten Eigenschaften *nicht* immer erfüllt.

- c) Betrachten wir zwei *beliebige* Vektoren aus der Teilmenge W von \mathbb{R}^4 :

$$\mathbf{v}_1 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ 0 \end{pmatrix}.$$

Dann gilt:

$$\mathbf{v}_1 + \mathbf{v}_2 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ 0 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ 0 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \\ 0 \end{pmatrix}$$

und für eine beliebige Zahl α :

$$\alpha \cdot \mathbf{v}_1 = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \alpha x_3 \\ 0 \end{pmatrix}.$$

Entscheidend ist hier, dass bei den Resultatvektoren die 4. Komponente immer 0 bleibt und damit Summen und Skalierungen von beliebigen Vektoren in W ebenfalls zu W gehören. Die erwähnten Eigenschaften sind hier also erfüllt. In der linearen Algebra sind solche Teilmengen von besonderem Interesse. Wir nennen sie *Unterräume* von \mathbb{R}^n .

Genauer definieren wir:

Eine Teilmenge W von \mathbb{R}^n heisst **linearer Unterraum** von \mathbb{R}^n , falls die folgenden zwei Eigenschaften erfüllt sind:

(U1) Falls zwei beliebige Vektoren v und w zu W gehören, so befindet sich auch ihre Summe

$$v + w$$

in W .

(U2) Für einen beliebigen Vektor v in W und eine beliebige reelle Zahl α gehört auch der Vektor

$$\alpha \cdot v$$

zu W .

Wir illustrieren diese Definition nochmals mit ein paar Beispielen:

Beispiel 2.15

a) Wir betrachten alle Punkte mit Koordinaten $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ im dreidimensionalen Raum, welche die folgende Gleichung erfüllen:

$$x_1 + 4x_2 - 2x_3 = 0. \quad (2.16)$$

Diese Punkte bilden eine Ebene E durch den Ursprung. Ist E ein linearer Unterraum von \mathbb{R}^3 ? Wir überprüfen (U1) und (U2):

(U1) Betrachte zwei *beliebige* (Orts-) Vektoren

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

in der Ebene E . Es gilt also:

$$x_1 + 4x_2 - 2x_3 = 0 \quad \text{und} \quad y_1 + 4y_2 - 2y_3 = 0.$$

Der Summenvektor

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{pmatrix}$$

liegt ebenfalls in E , denn durch Einsetzen in die Gleichung (2.16) sehen wir, dass

$$(x_1 + y_1) + 4(x_2 + y_2) - 2(x_3 + y_3) = \underbrace{(x_1 + 4x_2 - 2x_3)}_{=0} + \underbrace{(y_1 + 4y_2 - 2y_3)}_{=0} = 0,$$

d.h., (U1) ist erfüllt.

(U2) Betrachte einen *beliebigen* Vektor

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}$$

in der Ebene E . Dann gilt

$$z_1 + 4z_2 - 2z_3 = 0,$$

und für ein beliebiges skalares Vielfaches

$$\alpha \cdot \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} \alpha z_1 \\ \alpha z_2 \\ \alpha z_3 \end{pmatrix}$$

haben wir, wiederum durch Einsetzen in (2.16),

$$(\alpha z_1) + 4(\alpha z_2) - 2(\alpha z_3) = \alpha \underbrace{(z_1 + 4z_2 - 2z_3)}_{=0} = 0.$$

Somit gehört jedes skalare Vielfache ebenso zu E , und (U2) ist auch erfüllt.

b) Wir betrachten alle Punkte mit Koordinaten $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ im dreidimensionalen Raum, welche die folgende Gleichung erfüllen:

$$x_1 + 4x_2 - 2x_3 = 1.$$

Diese Punkte bilden eine Ebene E , die jetzt aber nicht durch den Ursprung geht! Ist E ein linearer Unterraum von \mathbb{R}^3 ? Wir überprüfen (U1) und (U2):

(U1) Betrachte zwei *beliebige* Vektoren

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}.$$

in der Ebene E . Es gilt also:

$$x_1 + 4x_2 - 2x_3 = 1, \quad y_1 + 4y_2 - 2y_3 = 1.$$

Der Summenvektor

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{pmatrix}$$

liegt *nicht* in E , denn

$$(x_1 + y_1) + 4(x_2 + y_2) - 2(x_3 + y_3) = \underbrace{(x_1 + 4x_2 - 2x_3)}_{=1} + \underbrace{(y_1 + 4y_2 - 2y_3)}_{=1} = 2 \neq 1.$$

d.h., (U1) ist *nicht* erfüllt. Somit ist E *kein* linearer Unterraum von \mathbb{R}^3 (aber trotzdem eine Teilmenge).

(U2) Auch (U2) ist nicht erfüllt. Betrachte einen *beliebigen* Vektor

$$\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}$$

in der Ebene E . Dann gilt

$$z_1 + 4z_2 - 2z_3 = 1,$$

und für ein beliebiges skalares Vielfaches

$$\alpha \cdot \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} \alpha z_1 \\ \alpha z_2 \\ \alpha z_3 \end{pmatrix}$$

haben wir

$$(\alpha z_1) + 4(\alpha z_2) - 2(\alpha z_3) = \alpha \underbrace{(z_1 + 4z_2 - 2z_3)}_{=1} = \alpha.$$

Falls $\alpha \neq 1$, dann gehört das entsprechende skalare Vielfache *nicht* zu E . Somit ist (U2) im Allgemeinen nicht erfüllt, und wiederum folgt, dass E kein linearer Unterraum von \mathbb{R}^3 ist.

□

Bemerkung 2.17 Ein linearer Unterraum von \mathbb{R}^n enthält *immer* den Nullvektor (wieso?). Umgekehrt gilt: Enthält eine Teilmenge des \mathbb{R}^n den Nullvektor nicht, so ist diese Teilmenge kein linearer Unterraum. Allerdings gibt es Teilmengen, die den Nullvektor enthalten und dennoch keine Unterräume sind.

2.3.2 Unterräume und Lösungen von homogenen Gleichungssystemen

Lösungsmengen von homogenen linearen Gleichungssystemen bilden immer Unterräume. Die folgenden Beispiele sollen dies illustrieren.

Beispiel 2.18

a) Es sei \mathbb{L} die Menge aller Lösungen des *homogenen* linearen Gleichungssystems (1.21),

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 0 \\ 7 & 8 & 9 & 0 \end{array} \right).$$

Schreiben wir die allgemeine Lösung als Vektor, so ist sie gegeben durch

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} s \\ -2s \\ s \end{pmatrix},$$

wobei s eine beliebige Zahl sein darf; vgl. (1.24). Somit ist die Lösungsmenge

$$\mathbb{L} = \left\{ s \cdot \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} : s \text{ eine beliebige Zahl} \right\}$$

die Menge aller beliebigen Vielfachen des Vektors $\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$. Ist \mathbb{L} ein linearer Unterraum von \mathbb{R}^3 ? Überprüfen wir die zwei Eigenschaften (U1) und (U2):

(U1) Seien $\mathbf{v}_1 = s_1 \cdot \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ und $\mathbf{v}_2 = s_2 \cdot \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ zwei beliebige Vektoren in \mathbb{L} . Dann gilt:

$$\mathbf{v}_1 + \mathbf{v}_2 = s_1 \cdot \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + s_2 \cdot \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} = (s_1 + s_2) \cdot \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

Der Resultatvektor ist wieder ein Vielfaches von $\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ und gehört somit zu \mathbb{L} .

Folglich ist (U1) erfüllt.

(U2) Betrachte einen beliebigen Vektor $\mathbf{v} = s \cdot \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ in \mathbb{L} und eine beliebige Zahl α .

Dann haben wir

$$\alpha \cdot \mathbf{v} = \alpha s \cdot \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

Auch dies ist ein Vielfaches von $\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$, d.h., eine beliebige Skalierung führt nicht zu einem "Verlassen" der Menge \mathbb{L} . Also ist auch (U2) wahr.

Zusammenfassend sind sowohl (U1) wie auch (U2) erfüllt, d.h., \mathbb{L} ist ein linearer Unterraum von \mathbb{R}^3 .

c) Definieren wir U als die Menge aller Lösungen des *inhomogenen* Gleichungssystems (1.44). Hier ist also

$$U = \left\{ s \cdot \begin{pmatrix} -73/38 \\ 17/38 \\ -23/38 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} : s \text{ eine beliebige Zahl} \right\}.$$

Ist U ein linearer Unterraum von \mathbb{R}^4 ? Hier sind weder (U1) noch (U2) erfüllt! Nehmen wir nämlich beispielsweise den Vektor

$$\mathbf{v} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix},$$

der zu U gehört (mit $s = 0$) und die Zahl $\alpha = 2$, so müsste mit (U2) auch der Vektor

$$2 \cdot \mathbf{v} = 2 \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 0 \\ 0 \end{pmatrix}$$

in U sein, was aber nicht der Fall ist. Die Menge U ist somit zwar eine Teilmenge, nicht aber ein linearer Unterraum von \mathbb{R}^4 .

□

Bemerkung 2.19 Allgemein gilt: Die Lösungsmenge eines *homogenen* linearen Gleichungssystems ist *immer* ein linearer Unterraum. Auf der anderen Seite ist die Lösungsmenge eines *inhomogenen* linearen Gleichungssystems *nie* ein linearer Unterraum (in der Tat spricht man hier von einem *affinen* Unterraum).

2.3.3 Linearkombinationen in Unterräumen

In Abschnitt 2.1.2 haben wir bereits den Begriff der *Linearkombination* von Vektoren in \mathbb{R}^n kennengelernt. Dieser Begriff lässt sich ganz natürlich auf beliebige Unterräume erweitern. Dazu erinnern wir uns an die Eigenschaften (U1) und (U2) der linearen Unterräume und bemerken, genau wie auf Seite 42, dass gilt:

Jede beliebige Linearkombination von Vektoren aus einem linearen Unterraum U von \mathbb{R}^n ist wiederum ein Vektor in demselben linearen Unterraum U .

Diese Beobachtung zieht die folgende Definition nach sich:

Die Menge aller möglichen Linearkombinationen, die sich aus einer gegebenen Menge von Vektoren

$$\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$$

bilden lassen, heisst **lineare Hülle** diese Vektoren, oder auch **Raum aller Vektoren, der durch die gegebenen Vektoren aufgespannt wird**. Diese Menge bezeichnen wir mit

$$\text{span}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p).$$

Sie erfüllt immer die Eigenschaften (U1) und (U2).

Beispiel 2.20 Der vom Vektor

$$\mathbf{v}_1 = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$$

aufgespannte lineare Unterraum $\text{span}(\mathbf{v}_1)$ von \mathbb{R}^2 ist der Raum aller Linearkombinationen von \mathbf{v}_1 . Dies ist die Menge aller Vektoren, die die Form

$$s \cdot \begin{pmatrix} 3 \\ 5 \end{pmatrix}$$

haben, wobei s beliebig gewählt werden darf. Geometrisch ist dies eine Gerade durch den Nullpunkt mit Richtung \mathbf{v}_1 . Die entsprechende Geradengleichung lautet

$$5x_1 - 3x_2 = 0.$$

Ähnlich kann jede Ebene durch den Ursprung im dreidimensionalen Raum \mathbb{R}^3 aufgespannt werden durch zwei Vektoren $\mathbf{v}_1, \mathbf{v}_2$. □

Beispiel 2.21 Wir betrachten das unterbestimmte homogene lineare Gleichungssystem

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

für fünf Unbekannte x_1, x_2, x_3, x_4, x_5 . Das System hat bereits rechts-obere-Dreiecksform und kann durch Rückwärtseinsetzen gelöst werden. Die Unbekannte x_5 ist frei wählbar und wird durch einen Parameter s ersetzt:

$$x_5 = s.$$

Dann folgt aus der dritten Gleichung:

$$x_4 = -5x_5 = -5s.$$

Einsetzen in die zweite Gleichung liefert:

$$x_3 = -3x_5 = -3s.$$

Aus der ersten Gleichung erhalten wir schliesslich

$$x_1 = -6x_2 - 4x_5 = -6x_2 - 4s.$$

Wir sehen, dass auch x_2 frei wählbar ist, wofür wir einen Parameter t einführen:

$$x_2 = t.$$

Daher

$$x_1 = -6t - 4s.$$

Damit ist die Menge aller Lösungen des obigen linearen Gleichungssystem gegeben durch

$$\mathbb{L} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} : x_1 = -6t - 4s, x_2 = t, x_3 = -3s, x_4 = -5s, x_5 = s \right\}.$$

Hier dürfen s und t beliebig gewählt werden. Die Lösungsvektoren haben also alle die Form

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} -6t - 4s \\ t \\ -3s \\ -5s \\ s \end{pmatrix} = s \cdot \begin{pmatrix} -4 \\ 0 \\ -3 \\ -5 \\ 1 \end{pmatrix} + t \cdot \begin{pmatrix} -6 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

für beliebige Zahlen s und t . Daraus folgt, dass sich jeder Lösungsvektor des obigen Gleichungssystems als Linearkombination der beiden Vektoren

$$\begin{pmatrix} -4 \\ 0 \\ -3 \\ -5 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} -6 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

darstellen lässt. Umgekehrt ist jede beliebige Linearkombination dieser beiden Vektoren eine Lösung des linearen Gleichungssystems. Also wird der Lösungsraum \mathbb{L} durch die beiden Vektoren \mathbf{v}_1 und \mathbf{v}_2 aufgespannt, d.h., wir können schreiben:

$$\mathbb{L} = \text{span}(\mathbf{v}_1, \mathbf{v}_2).$$

□

2.4 Lineare Unabhängigkeit

Das Beispiel 2.4 zeigt, dass sich Vektoren manchmal als Linearkombinationen von anderen (gegebenen) Vektoren schreiben lassen, manchmal aber auch nicht. So lässt sich, wie im Beispiel gezeigt, der Vektor

$$\begin{pmatrix} -11 \\ -16 \\ -5 \end{pmatrix}$$

als Linearkombination der Vektoren

$$\mathbf{v}_1 = \begin{pmatrix} -3 \\ 2 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix}$$

darstellen, nämlich

$$\begin{pmatrix} -11 \\ -16 \\ -5 \end{pmatrix} = 2 \cdot \begin{pmatrix} -3 \\ 2 \\ 0 \end{pmatrix} + (-5) \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix}.$$

Jedoch kann der Vektor

$$\begin{pmatrix} -2 \\ 6 \\ 2 \end{pmatrix}$$

nicht als Linearkombination dieser zwei Vektoren \mathbf{v}_1 und \mathbf{v}_2 geschrieben werden.

Wir wollen uns nun mit einer speziellen Frage befassen: Gegeben seien Vektoren $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$ in \mathbb{R}^n . Lässt sich der Nullvektor

$$\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

in \mathbb{R}^n darstellen als Linearkombination dieser Vektoren? Anders gefragt: Gibt es Zahlen $\alpha_1, \alpha_2, \dots, \alpha_p$, sodass

$$\alpha_1 \cdot \mathbf{v}_1 + \alpha_2 \cdot \mathbf{v}_2 + \dots + \alpha_p \cdot \mathbf{v}_p = \mathbf{0}? \quad (2.22)$$

Die Antwort ist “ja”, denn mit

$$\alpha_1 = \alpha_2 = \dots = \alpha_p = 0 \quad (2.23)$$

ist eine Lösung gefunden. Interessanter ist also die Frage: Ist die Lösung (2.23) die *einzig*e Lösung der Gleichung (2.22), oder lässt sich der Nullvektor $\mathbf{0}$ auf verschiedene Arten als Linearkombination der Vektoren $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$ darstellen?

Wir betrachten ein Beispiel.

Beispiel 2.24

a) Betrachten wir die zwei Vektoren

$$\mathbf{v}_1 = \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} -1 \\ 2 \\ 4 \end{pmatrix}$$

in \mathbb{R}^3 . Auf wieviele Arten kann der Nullvektor als Linearkombination dieser Vektoren dargestellt werden? Wir lösen Gleichung (2.22):

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 = \mathbf{0},$$

d.h.,

$$\begin{aligned} 2\alpha_1 + (-1)\alpha_2 &= 0 \\ 2\alpha_1 + 2\alpha_2 &= 0 \\ 3\alpha_1 + 4\alpha_2 &= 0 \end{aligned} \quad (2.25)$$

Wir schreiben dieses lineare Gleichungssystem als erweiterte Matrix

$$\begin{array}{l} (I) \\ (II) \\ (III) \end{array} \left(\begin{array}{cc|c} 2 & -1 & 0 \\ 2 & 2 & 0 \\ 3 & 4 & 0 \end{array} \right)$$

und führen einen Eliminationsschritt mit dem Gaussverfahren durch:

$$\rightsquigarrow \begin{array}{l} (II) - (I) \\ (III) - 3/2 \cdot (I) \end{array} \left(\begin{array}{cc|c} 2 & -1 & 0 \\ 0 & 3 & 0 \\ 0 & 5.5 & 0 \end{array} \right)$$

Aus der zweiten und dritten Gleichung folgt, dass

$$\alpha_2 = 0.$$

Einsetzen in die erste Gleichung, ergibt dann

$$\alpha_1 = 0.$$

Die einzige Lösung ist in diesem Fall also $\alpha_1 = \alpha_2 = 0$, d.h., die Lösung (2.23).

b) Wir lösen nun dieselbe Aufgabe für die zwei Vektoren

$$\mathbf{w}_1 = \begin{pmatrix} -2 \\ 4 \\ 8 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} -1 \\ 2 \\ 4 \end{pmatrix}$$

in \mathbb{R}^3 , und stellen erneut die Frage: Auf wieviele Arten kann der Nullvektor als Linearkombination dieser Vektoren dargestellt werden? Auch hier lösen wir die Gleichung (2.22). In Form der erweiterten Matrix bedeutet dies

$$\begin{array}{l} (I) \\ (II) \\ (III) \end{array} \left(\begin{array}{cc|c} -2 & -1 & 0 \\ 4 & 2 & 0 \\ 8 & 4 & 0 \end{array} \right).$$

Ein Schritt mit dem Gaussverfahren liefert:

$$\begin{array}{l} (II) - (-2) \cdot (I) \\ (III) - (-4) \cdot (I) \end{array} \left(\begin{array}{cc|c} -2 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right).$$

Die zweite und dritte Gleichung beinhalten keine Information; sie sind für jede beliebige Wahl von α_1, α_2 erfüllt. Aus der ersten Gleichung erhalten wir

$$-2\alpha_1 + (-1)\alpha_2 = 0 \quad \Rightarrow \quad \alpha_1 = -\frac{1}{2}\alpha_2.$$

Die Unbekannte α_2 ist frei wählbar. Wir führen einen Parameter ein:

$$\alpha_2 = s.$$

Daraus ergibt sich

$$\alpha_1 = -\frac{1}{2}s.$$

Also lässt sich der Nullvektor wie folgt als Linearkombination der Vektoren w_1 und w_2 darstellen:

$$-\frac{1}{2}s \cdot w_1 + s \cdot w_2 = \mathbf{0}.$$

Hier darf s beliebig gewählt werden, d.h., es gibt unendlich viele Linearkombinationen aus den Vektoren w_1, w_2 (insbesondere mehr als eine!), um den Nullvektor zu bilden.

□

Diese Betrachtungen führen uns zur folgenden Definition:

Gegeben seien Vektoren

$$v_1, v_2, \dots, v_p.$$

Dann lässt sich der Nullvektor immer als Linearkombination (2.22) dieser Vektoren darstellen, nämlich mit

$$\alpha_1 = \alpha_1 = \dots = \alpha_p = 0.$$

Falls dies die *einzigste Möglichkeit* ist, den Nullvektor als Linearkombination dieser Vektoren zu schreiben, dann heißen die Vektoren v_1, \dots, v_p **linear unabhängig**. Wenn es mehr als eine Möglichkeit gibt, dann heißen sie **linear abhängig**.

Beispiel 2.26 Aus Beispiel 2.24 wissen wir, dass die beiden Vektoren

$$v_1 = \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -1 \\ 2 \\ 4 \end{pmatrix}$$

linear *unabhängig* sind. Im Gegensatz dazu sind die zwei Vektoren

$$v_1 = \begin{pmatrix} -2 \\ 4 \\ 8 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -1 \\ 2 \\ 4 \end{pmatrix}.$$

linear *abhängig*.

□

Bemerkung 2.27 Es gibt in diesem Zusammenhang eine wichtige Tatsache: Schreiben wir eine Menge von Vektoren v_1, v_2, \dots, v_p als Spalten in eine Matrix A , d.h.,

$$A = (v_1 \mid v_2 \mid \cdots \mid v_p),$$

so ist der *Rang* der Matrix A (vgl. Abschnitt 1.4) gleich der maximalen Anzahl von linear unabhängigen Vektoren, die sich in der Menge der Vektoren v_1, v_2, \dots, v_p finden lassen. Beispielsweise seien

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 3 \\ -4 \\ 2 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 3 \\ 2 \\ 1 \\ 0 \\ 6 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 5 \\ 4 \\ 7 \\ -8 \\ 10 \end{pmatrix}.$$

Schreiben wir diese Vektoren als Spalten in eine Matrix, so erhalten wir:

$$A = \begin{pmatrix} 1 & 3 & 5 \\ 1 & 2 & 4 \\ 3 & 1 & 7 \\ -4 & 0 & -8 \\ 2 & 6 & 10 \end{pmatrix}.$$

In Bemerkung 1.45 haben wir festgestellt, dass $\text{rang}(A) = 2$. Folglich gibt es unter den drei Vektoren v_1, v_2, v_3 nie mehr als zwei, die linear unabhängig sind. In der Tat sind in diesem Beispiel immer je zwei linear unabhängig. Alle drei zusammen sind aber linear abhängig (sonst wäre ja $\text{rang}(A) = 3$, was nicht der Fall ist).

2.5 Basen und Dimension

Betrachten wir einen linearen Unterraum W von \mathbb{R}^n . Weiter nehmen wir an, dass es p Vektoren v_1, v_2, \dots, v_p gibt, welche diesen Raum aufspannen, d.h.

$$W = \text{span}(v_1, v_2, \dots, v_p).$$

In anderen Worten: Jeder Vektor in W lässt sich als Linearkombination der Vektoren v_1, v_2, \dots, v_p darstellen.

Beispiel 2.28

a) In Beispiel 2.20 wurde gezeigt, dass sich die Gerade

$$g : 5x_1 - 3x_2 = 0$$

aufspannen lässt durch den Vektor

$$v_1 = \begin{pmatrix} 3 \\ 5 \end{pmatrix},$$

d.h.,

$$g = \text{span}(v_1).$$

- b) In Beispiel 2.21 konnten wir den Lösungsraum \mathbb{L} des betreffenden linearen Gleichungssystems (der ein linearer Unterraum von \mathbb{R}^5 ist) schreiben als

$$\mathbb{L} = \text{span} \left(\begin{pmatrix} -4 \\ 0 \\ -3 \\ -5 \\ 1 \end{pmatrix}, \begin{pmatrix} -6 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right).$$

Hier lässt sich ferner zeigen, dass die beiden aufspannenden Vektoren linear unabhängig sind. □

Wir halten die folgende wichtige Definition fest.

Angenommen, es seien ein linearer Unterraum W von \mathbb{R}^n und p Vektoren $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_p$ gegeben. Dann heißen diese Vektoren **Basis** von W , falls die folgenden zwei Bedingungen erfüllt sind:

(B1) Die Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_p$ spannen den linearen Unterraum W auf, d.h.,

$$W = \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_p).$$

(B2) Die Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_p$ sind *linear unabhängig*.

In diesem Fall heisst die Zahl p **Dimension** von W , und wir schreiben

$$p = \dim(W).$$

Es gilt die wichtige Tatsache: Jeder lineare Unterraum von \mathbb{R}^n hat eine Basis (sogar unendlich viele).

Beispiel 2.29

- a) Bilden die Vektoren

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

eine Basis von \mathbb{R}^3 ? Es lässt sich sofort sehen, dass die beiden Vektoren linear unabhängig sind, d.h., (B2) ist erfüllt. Die Bedingung (B1) ist jedoch nicht gegeben: Beispielsweise kann der Vektor

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

unmöglich als Linearkombination von \mathbf{v}_1 und \mathbf{v}_2 dargestellt werden, was aber mit Bedingung (B1) für *jeden* Vektor in \mathbb{R}^3 möglich sein müsste.

b) Bilden die Vektoren

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

eine Basis von \mathbb{R}^2 ? Hier ist (B1) erfüllt, aber (B2) nicht. Zum Beispiel lässt sich der Nullvektor $\mathbf{0}$ darstellen in der Form

$$1 \cdot \mathbf{v}_1 + 1 \cdot \mathbf{v}_2 + (-1) \cdot \mathbf{v}_3 = \mathbf{0}.$$

Insbesondere ist

$$0 \cdot \mathbf{v}_1 + 0 \cdot \mathbf{v}_2 + 0 \cdot \mathbf{v}_3 = \mathbf{0}$$

nicht die einzige Möglichkeit, den Nullvektor als Linearkombination der Vektoren $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ darzustellen, d.h., die Vektoren sind linear abhängig.

c) Im vorherigen Beispiel 2.28 a), wird die Gerade g durch den Vektor $\mathbf{v}_1 = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$ erzeugt. Ein einzelner Vektor ist immer linear unabhängig, und somit ist \mathbf{v}_1 eine Basis von g . Das es nur einen Basisvektor gibt, hat g die Dimension 1, was auch unserer geometrischen Vorstellung entspricht.

d) Betrachten wir die Ebene

$$E : x_1 + 3x_2 - 2x_3 = 0$$

durch den Nullpunkt in \mathbb{R}^3 . Dies ist ein linearer Unterraum von \mathbb{R}^3 . Folglich muss sich eine Basis von E finden lassen. Um dies zu tun, stellen wir E zunächst in Parameterschreibweise dar. Dazu lösen wir die Gleichung für E nach x_1 auf:

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

Die Variablen x_2 und x_3 sind frei wählbar und können durch Parameter ersetzt werden:

$$x_2 = s, \quad x_3 = t.$$

Damit folgt

$$x_1 = -3s + 2t.$$

Also kann die Ebene E geschrieben werden als die Menge aller (Orts-) Vektoren, die die Form

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -3s + 2t \\ s \\ t \end{pmatrix} = \begin{pmatrix} -3s \\ s \\ 0 \end{pmatrix} + \begin{pmatrix} 2t \\ 0 \\ t \end{pmatrix} = s \cdot \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix} + t \cdot \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

haben, wobei s, t beliebige Zahlen sind. Mit anderen Worten: Die Ebene E wird gebildet durch alle möglichen Linearkombinationen der beiden Vektoren

$$\mathbf{v}_1 = \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix},$$

d.h.,

$$E = \text{span}(\mathbf{v}_1, \mathbf{v}_2).$$

Es lässt sich zeigen, dass \mathbf{v}_1 und \mathbf{v}_2 linear unabhängig sind, und somit bilden sie eine Basis von E . Es folgt auch $\dim(E) = 2$.

- e) Im obigen Beispiel 2.28 b) spannen die Vektoren $\begin{pmatrix} -4 \\ 0 \\ -3 \\ -5 \\ 1 \end{pmatrix}$, $\begin{pmatrix} -6 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ den linearen Unterraum \mathbb{L} auf. Ausserdem sind sie linear unabhängig. Damit bilden sie eine Basis von \mathbb{L} . Ferner gilt $\dim(L) = 2$.

□

Anwendung 2.30 (Datenübermittlung II)

Es sollen 3 Zahlen übermittelt werden. Zwecks Überprüfung der Richtigkeit wird zusätzlich noch eine Prüfziffer (vgl. Anwendung 2.10) gesendet, die sich als die Summe der drei Zahlen berechnet. Bezeichnen wir die drei Zahlen mit a, b, c , so schickt der Sender folgende Nachricht:

$$\boxed{a \mid b \mid c \mid \mid a + b + c}.$$

Eine solche Originalnachricht bestehend aus den drei Zahlen zusammen mit der Prüfziffer heisst *Codewort*. Der Empfänger erhält die Nachricht in der Form

$$\boxed{u \mid v \mid w \mid \mid z}.$$

Wenn er feststellt, dass $u + v + w \neq z$, dann weiss er, dass mindestens eine der vier gesendeten Zahlen falsch ist. Wenn Gleichheit gilt, so ist dies zwar keine Garantie für die Richtigkeit, jedoch ein Indikator. Die Menge aller "Codewörter" (Originalnachrichten) erfüllt nach Definition die folgende Gleichung:

$$u + v + w - z = 0.$$

Interpretieren wir eine Nachricht als einen Vektor in \mathbb{R}^4 , so haben alle Codewörter die Form

$$\begin{pmatrix} u \\ v \\ w \\ u + v + w \end{pmatrix},$$

wobei u, v, w beliebige Zahlen sein dürfen. Ein solcher Vektor lässt sich schreiben in der Form

$$\begin{pmatrix} u \\ v \\ w \\ u + v + w \end{pmatrix} = \begin{pmatrix} u \\ 0 \\ 0 \\ u \end{pmatrix} + \begin{pmatrix} 0 \\ v \\ 0 \\ v \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ w \\ w \end{pmatrix} = u \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} + v \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} + w \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$$

Somit ist jedes Codewort eine Linearkombination der drei (linear unabhängigen) Vektoren

$$\mathbf{r}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{r}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{r}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix},$$

und umgekehrt ist jede solche Linearkombination ein Codewort. Die Menge der Codewörter ist also ein 3-dimensionaler linearer Unterraum von \mathbb{R}^4 . Zu bemerken ist, dass auch eine falsch übermittelte Nachricht wiederum ein Codewort sein kann; sicherlich ist aber eine empfangene Nachricht, die nicht im Sinne der Definition ein Codewort ist, falsch übermittelt worden und muss erneut gesendet werden. \diamond

Standardbasis in \mathbb{R}^n : Auch der Raum \mathbb{R}^n selbst besitzt eine Basis, nämlich die Vektoren

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \quad \dots, \quad \mathbf{e}_{n-1} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}. \quad (2.31)$$

In der Tat lässt sich jeder beliebige Vektor

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

in \mathbb{R}^n schreiben als Linearkombinationen dieser Basisvektoren:

$$\mathbf{x} = x_1 \cdot \mathbf{e}_1 + x_2 \cdot \mathbf{e}_2 + \dots + x_n \cdot \mathbf{e}_n. \quad (2.32)$$

Also spannen die Vektoren $\mathbf{e}_1, \dots, \mathbf{e}_n$, die auch **Standardbasis von \mathbb{R}^n** genannt werden, den Raum \mathbb{R}^n auf. Ausserdem sind sie linear unabhängig. Die Dimension von \mathbb{R}^n ist natürlicherweise n .

Beispiel 2.33 Die Standardbasis in \mathbb{R}^4 ist gegeben durch die Vektoren

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

\square

2.6 Koordinaten

Wir betrachten die zwei Vektoren

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}.$$

Sie sind linear unabhängig und bilden damit eine Basis des linearen Unterraums

$$U = \text{span}(\mathbf{v}_1, \mathbf{v}_2)$$

von \mathbb{R}^3 . Geometrisch ist U eine Ebene des \mathbb{R}^3 durch den Ursprung, die aufgespannt wird von den zwei Ortsvektoren \mathbf{v}_1 und \mathbf{v}_2 . Wir können nun beliebige Linearkombinationen der beiden Basisvektoren bilden, um weitere Ortsvektoren in der Ebene zu finden. Zum Beispiel:

$$\mathbf{w}_1 = 5 \cdot \mathbf{v}_1 + 2 \cdot \mathbf{v}_2 = \begin{pmatrix} 11 \\ 10 \\ 2 \end{pmatrix}$$

oder

$$\mathbf{w}_2 = (-1) \cdot \mathbf{v}_1 + 4 \cdot \mathbf{v}_2 = \begin{pmatrix} 11 \\ -2 \\ 4 \end{pmatrix}.$$

In der bekannten Weise können wir zeigen, dass die beiden Vektoren \mathbf{w}_1 und \mathbf{w}_2 linear unabhängig sind und deshalb ebenfalls die Ebene U aufspannen:

$$U = \text{span}(\mathbf{v}_1, \mathbf{v}_2) = \text{span}(\mathbf{w}_1, \mathbf{w}_2).$$

Wir haben also zwei verschiedene Basen von U

$$\text{Basis 1: } \mathcal{B}_1 = \{\mathbf{v}_1, \mathbf{v}_2\}, \quad \text{Basis 2: } \mathcal{B}_2 = \{\mathbf{w}_1, \mathbf{w}_2\}$$

gefunden.

Betrachten wir nun einen weiteren Vektor in U , beispielsweise

$$\mathbf{z} = \begin{pmatrix} 9 \\ 6 \\ 2 \end{pmatrix},$$

so können wir diesen als Linearkombination bezüglich beider Basen wie folgt schreiben (nachprüfen!):

$$\text{Basis } \mathcal{B}_1 : \mathbf{z} = 3 \cdot \mathbf{v}_1 + 2 \cdot \mathbf{v}_2 \tag{2.34}$$

und

$$\text{Basis } \mathcal{B}_2 : \mathbf{z} = \frac{7}{11} \cdot \mathbf{w}_1 + \frac{2}{11} \cdot \mathbf{w}_2$$

Der durch die obigen zwei Linearkombinationen dargestellte Vektor z ist in beiden Fällen genau der gleiche, jedoch sind die Koeffizienten bezüglich der beiden Basen verschieden, nämlich für Basis \mathcal{B}_1 haben wir 3 und 2, während wir bezüglich Basis \mathcal{B}_2 die Koeffizienten $\frac{7}{11}$ und $\frac{2}{11}$ verwenden (diese Koeffizienten sind übrigens eindeutig).

Um diesen Unterschied der Darstellungen bezüglich verschiedener Basen deutlich zu machen, benutzen wir die folgende Notation:

$$z = \begin{pmatrix} 9 \\ 6 \\ 2 \end{pmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}_{U, \mathcal{B}_1},$$

Dies bedeutet, dass wir den Vektor z bezüglich der Basis \mathcal{B}_1 des linearen Unterraums U repräsentieren können in der Form (2.34) Analog schreiben wir:

$$\begin{pmatrix} 9 \\ 6 \\ 2 \end{pmatrix} = \begin{bmatrix} 7/11 \\ 2/11 \end{bmatrix}_{U, \mathcal{B}_2}.$$

Die Zahlen in den eckigen Klammern nennen wir die **Koordinaten** des Vektors z . Die Koordinaten *beziehen sich immer auf eine gegebene Basis!*

Beispiel 2.35 Berechnen wir den Vektor y , der zu den Koordinaten

$$\begin{bmatrix} 1 \\ -2 \end{bmatrix}_{U, \mathcal{B}_1}$$

gehört. Mit der Definition von Koordinaten gilt:

$$y = 1 \cdot v_1 + (-2) \cdot v_2 = \begin{pmatrix} -5 \\ 2 \\ -2 \end{pmatrix}.$$

Wie lauten die Koordinaten dieses Vektors bezüglich der Basis \mathcal{B}_2 ? Wir müssen Zahlen c_1 und c_2 suchen, sodass

$$y = c_1 \cdot w_1 + c_2 \cdot w_2.$$

Dies bedeutet:

$$\begin{pmatrix} -5 \\ 2 \\ -2 \end{pmatrix} = c_1 \cdot \begin{pmatrix} 11 \\ 10 \\ 2 \end{pmatrix} + c_2 \cdot \begin{pmatrix} 11 \\ -2 \\ 4 \end{pmatrix} = \begin{pmatrix} 11c_1 + 11c_2 \\ 10c_1 - 2c_2 \\ 2c_1 + 4c_2 \end{pmatrix}$$

Dies entspricht dem linearen Gleichungssystem

$$\begin{aligned} 11c_1 + 11c_2 &= -5 \\ 10c_1 - 2c_2 &= 2 \\ 2c_1 + 4c_2 &= -2 \end{aligned}$$

Es hat die Lösung

$$c_1 = \frac{1}{11}, \quad c_2 = -\frac{6}{11},$$

was durch Einsetzen in die drei Gleichungen nachgeprüft werden kann. Somit finden wir

$$\mathbf{y} = \begin{bmatrix} 1/11 \\ -6/11 \end{bmatrix}_{U, B_2}.$$

□

2.7 Anwendung: Composition Space für Mineralien

Viele chemische Verbindungen können als einfache Formeln geschrieben werden, wenn eine *atomare (bzw. molare)* Masseinheit gewählt wird. So hat beispielsweise Wasser H_2O doppelt so viele Wasserstoff- wie Sauerstoffatome. Wird als Einheit jedoch die Masse gewählt, so gilt

$$\text{H} : \text{O} \approx \frac{2 \cdot 1.015}{1 \cdot 15.9994} \approx 0.126878,$$

was einem zahlenmässig komplizierteren Verhältnis entspricht.

In Atom- oder Moleinheiten lassen sich auch kompliziertere Verbindungen, wie beispielsweise viele Mineralien, recht einfach darstellen. Der Grund dafür liegt in der Stereochemie: Mineralien haben eine kristalline Struktur, d.h., sie bestehen aus einem dreidimensionalen repetierten Modul ("Molekül") von fester Struktur und Zusammensetzung. Dieser starre atomare Bau führt zu vielen makroskopischen Eigenschaften von Mineralien, hat aber auch interessante mathematische Aspekte.

Wir interessieren uns für die Menge aller Stoffe, welche sich aus Ca, Al, Si und O zusammensetzen. Hier identifizieren wir die einzelnen Elemente mit den Standardbasisvektoren in \mathbb{R}^4 :

$$\text{Ca} \cong \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \text{Al} \cong \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \text{Si} \cong \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \text{O} \cong \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Die Menge aller Linearkombinationen dieser Vektoren bilden den Raum \mathbb{R}^4 ; geologisch betrachtet, entspricht jede Linearkombination einer chemischen Verbindung basierend auf Ca, Al, Si und O. Wir nennen den Raum aller dieser Kombinationen darum auch den *Composition Space*. Dann können wir beispielsweise Quarz SiO_2 schreiben in der Form

$$1 \cdot \text{Si} + 2 \cdot \text{O} \cong 1 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \end{pmatrix}.$$

Wir erkennen, dass die Koeffizienten in der obigen Linearkombination (d.h., die Koordinaten bzgl. der Standardbasis) genau die stöchiometrischen Koeffizienten von Quarz sind. Analog ergibt sich:

$$\text{Disthen } \text{Al}_2\text{SiO}_5 \hat{=} \begin{pmatrix} 0 \\ 2 \\ 1 \\ 5 \end{pmatrix}$$

und

$$\text{Korund } \text{Al}_2\text{O}_3 \hat{=} \begin{pmatrix} 0 \\ 2 \\ 0 \\ 3 \end{pmatrix}.$$

Wir schreiben die drei Vektoren für Quarz, Disthen und Korund als *Spalten* in die sogenannten *Zusammensetzungsmatrix*:

$$M = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 2 \\ 1 & 1 & 0 \\ 2 & 5 & 3 \end{pmatrix}.$$

Bestimmen wir den Rang der Matrix M so erhalten wir

$$\text{rang}(M) = 2.$$

Dies bedeutet, dass die Spalten von M linear abhängig sind (vgl. Bemerkung 2.27), denn sonst wäre der Rang gleich 3. Insbesondere folgt nun, dass sich der Nullvektor so als Linearkombination der drei Spalten von M schreiben lässt, dass nicht alle Koeffizienten gleich Null sind. Wir suchen alle Koeffizienten, für die dies gilt:

$$a \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \end{pmatrix} + b \cdot \begin{pmatrix} 0 \\ 2 \\ 1 \\ 5 \end{pmatrix} + c \cdot \begin{pmatrix} 0 \\ 2 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Das ist ein lineares Gleichungssystem für die Unbekannten a, b, c . Wir haben die Parameterlösung

$$a = s, \quad b = -s, \quad c = s,$$

wo s eine beliebige Zahl sein darf. Beispielsweise gilt für $s = 1$:

$$1 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \end{pmatrix} - 1 \cdot \begin{pmatrix} 0 \\ 2 \\ 1 \\ 5 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 2 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

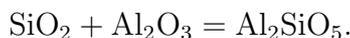
Dies ist eine Gleichheit im Composition Space, die sich jetzt wieder zurückübersetzen lässt in chemische Formeln:

$$1 \cdot \text{Quarz} - 1 \cdot \text{Disthen} + 1 \cdot \text{Korund} = 0, \quad (2.36)$$

oder



und daher



Diese drei Verbindungen sind also “linear abhängig”. Wichtig ist hier zu bemerken, dass die obigen Gleichheiten im Sinne von chemischen Verbindungen und nicht als Gemische gemeint sind.

Da der Rang von M gleich 2 ist, hat der lineare Unterraum

$$U = \text{span} \left(\begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 1 \\ 5 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \\ 3 \end{pmatrix} \right) \hat{=} \text{span} (\text{Quarz}, \text{Disthen}, \text{Korund})$$

von \mathbb{R}^4 die Dimension 2 (wegen $\text{rang}(M) = 2$ gibt es unter den drei Vektoren ja nur zwei linear unabhängige und damit kann höchstens ein Raum der Dimension 2 aufgespannt werden). In der Sprache des Composition Space hat der Raum aller möglichen “Linearkombinationen” der drei Stoffe Quarz, Disthen und Korund die Dimension 2. Eine Basis von U ist beispielsweise gegeben durch die beiden Vektoren

$$\mathbf{v}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 2 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 2 \\ 1 \\ 5 \end{pmatrix}.$$

Insbesondere gilt:

$$U = \text{span} (\mathbf{v}_1, \mathbf{v}_2) \hat{=} \text{span} (\text{Quarz}, \text{Disthen}).$$

Geologisch bedeutet diese mathematische Tatsache, dass sich jede chemische Verbindung, die sich aus den Molekülen Quarz, Disthen und Korund darstellen lässt, bereits allein nur mit Quarz und Disthen geschrieben werden kann.

Bemerkung 2.38 Es scheint klar zu sein, dass in einer chemischen Reaktionsgleichung grundsätzlich *keine negativen* stöchiometrischen Koeffizienten auftreten sollten. Insbesondere muss die entsprechende algebraische Gleichung zunächst so umgeschrieben werden, dass auf beiden Seiten der Gleichung nur positive Koeffizienten stehen (vgl. (2.36) und (2.37)). Erst dann ist eine Identifizierung mit einer (möglicherweise theoretischen) Reaktionsgleichung eigentlich sinnvoll.

2.8 Übungsaufgaben

2.1. Begründen oder widerlegen Sie folgende Aussagen.

- a) 3214667109 ist eine korrekte ISBN.

b) Die senkrechte Projektion des Vektors \mathbf{a} auf den Vektor \mathbf{b} ist gleich \mathbf{a}^{\parallel} , wobei

$$\mathbf{a} = \begin{pmatrix} 1 \\ -2 \\ 3 \\ -3 \\ 5 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} -2 \\ 4 \\ 4 \\ -4 \\ 4 \end{pmatrix} \quad \text{und} \quad \mathbf{a}^{\parallel} = \begin{pmatrix} -1 \\ 2 \\ 2 \\ -2 \\ 2 \end{pmatrix}.$$

c) Die Menge aller nicht-negativen reellen Zahlen ist ein linearer Unterraum von \mathbb{R}^1 .

d) Jede Gerade in \mathbb{R}^2 ist ein linearer Unterraum von \mathbb{R}^2 .

e) Die Menge aller Vektoren in \mathbb{R}^6 , die senkrecht auf den Vektor

$$\mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \\ 0 \\ -1 \end{pmatrix}$$

stehen, bilden einen linearen Unterraum von \mathbb{R}^6 der Dimension 5.

2.2. Es sei \mathbb{L} der Lösungsraum des linearen Gleichungssystems

$$\begin{aligned} 2x_1 + 3x_2 - x_3 + x_4 + 2x_5 &= 0 \\ 2x_1 + 4x_2 - x_3 - x_4 - x_5 &= 0 \\ 4x_1 + x_2 + x_5 &= 0. \end{aligned}$$

a) Schreiben Sie \mathbb{L} in Parameterdarstellung.

b) Geben Sie eine Basis von \mathbb{L} an und zeigen Sie, dass es sich dabei tatsächlich um eine Basis handelt.

c) Bestimmen Sie die Dimension von \mathbb{L} .

2.3. Betrachte die Ebene

$$E : 2x_1 + 3x_2 - x_3 = 0$$

in \mathbb{R}^3 .

a) Zeige, dass E ein linearer Unterraum von \mathbb{R}^3 ist.

b) Finde eine Basis von E .

c) Bilden die Vektoren

$$\mathbf{v} = \begin{pmatrix} -4 \\ 3 \\ 1 \end{pmatrix} \quad \text{und} \quad \mathbf{w} = \begin{pmatrix} -2 \\ 3 \\ 5 \end{pmatrix}$$

ebenfalls eine Basis von E ?

2.4. Welche der folgenden Mengen von Vektoren sind Basen von \mathbb{R}^3 ?

$$\begin{array}{ll} \text{a) } \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} & \text{b) } \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -7 \\ 1 \end{pmatrix} \\ \text{c) } \begin{pmatrix} 3 \\ 1 \\ -4 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \\ 6 \end{pmatrix}, \begin{pmatrix} 1 \\ 4 \\ 8 \end{pmatrix} & \text{d) } \begin{pmatrix} 1 \\ 6 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 5 \end{pmatrix}. \end{array}$$

2.5. Bestimmen Sie jeweils eine Basis und die Dimension der folgenden Unterräume:

$$\begin{array}{l} \text{a) } \text{span} \left(\begin{pmatrix} 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ -3 \end{pmatrix} \right) \\ \text{b) } \text{span} \left(\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix} \right) \\ \text{c) } \text{span} \left(\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right) \end{array}$$

2.6. Finden Sie für die folgenden Gleichungssysteme jeweils eine Basis und die Dimension des Lösungsraums.

$$\begin{array}{l} x_1 + x_2 - x_3 = 0 \\ -2x_1 - x_2 + 2x_3 = 0 \\ -x_1 + x_3 = 0 \end{array}$$

$$\begin{array}{l} 3x_1 + x_2 + x_3 + x_4 = 0 \\ 5x_1 - x_2 + x_3 - x_4 = 0 \end{array}$$

$$\begin{array}{l} x_1 - 4x_2 + 3x_3 - x_4 = 0 \\ 2x_1 - 8x_2 + 6x_3 - 2x_4 = 0 \end{array}$$

$$\begin{array}{l} x_1 - 3x_2 + x_3 = 0 \\ 2x_1 - 6x_2 + 2x_3 = 0 \\ 3x_1 - 9x_2 + 3x_3 = 0 \end{array}$$

$$\begin{array}{l} 2x_1 + x_2 + 3x_3 = 0 \\ x_1 + 5x_3 = 0 \\ x_2 + x_3 = 0 \end{array}$$

$$\begin{array}{l} x + y + z = 0 \\ 3x + 2y - 2z = 0 \\ 4x + 3y - z = 0 \\ 6x + 5y + z = 0. \end{array}$$

2.7. a) Finden Sie eine Basis der Ebene $E : 5x - 3y + 2z = 0$.

b) Projizieren Sie den Vektor $v = \begin{pmatrix} -2 \\ 16 \\ -9 \end{pmatrix}$ senkrecht auf die Ebene E .

2.8. Betrachten Sie die 6 Mineralien

Quarz:	SiO_2
Anorthit:	$\text{CaAl}_2\text{Si}_2\text{O}_8$
Disthen:	Al_2SiO_5
Grossular:	$\text{Ca}_3\text{Al}_2(\text{SiO}_4)_3$
Korund:	Al_2O_3
Wollastonit:	CaSiO_3

im Composition Space bestehend aus Ca, Al, Si, O.

- Schreiben Sie die 6 Mineralien in Form von Vektoren im Composition Space.
- Betrachten Sie die Matrix M , deren 6 Spalten die Vektoren aus a) sind und bestimmen Sie $\text{rang}(M)$ mit Hilfe des Gaußverfahrens. Was bedeutet Ihr Resultat?
- Finden Sie eine Basis von

$\text{span}(\text{Quarz, Anorthit, Disthen, Grossular, Korund, Wollastonit})$.

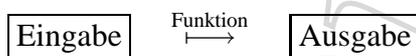
- Finden Sie die stöchiometrischen Koeffizienten (ganzzahlig und minimal) aller (theoretisch) möglichen Reaktionen zwischen den obigen 6 Mineralien.

U
ni
B
erlin

3

Lineare Abbildungen

Aus der Analysis wissen wir, dass eine Funktion eine *Vorschrift* ist, welche einer Zahl (Eingabe) eine andere Zahl (Ausgabe) zuordnet:



Beispielsweise ist die Länge des Umfangs eines Kreises mit Radius r gegeben durch

$$r \mapsto U(r) = 2\pi r.$$

Jedem Radius r (Eingabe) wird hier der entsprechende Kreisumfang $U(r)$ (Ausgabe) zugewiesen:

$$\begin{aligned} U(0) &= 0 \\ U(1) &= 2\pi = 6.283185 \dots \\ U(11.53) &= 72.445126 \dots \\ &\vdots \\ U(6371) &= 40030.173592 \dots \end{aligned}$$

3.1 Vektorfunktionen

Wir werden nun das Konzept der Funktionen von Zahlen auf Vektoren erweitern. Vektorfunktionen sind Vorschriften, die jedem Eingabe-Vektor in einem Raum \mathbb{R}^m einen Ausgabe-Vektor in \mathbb{R}^n zuweisen (hier ist $m \neq n$ durchaus zulässig).

Beispiel 3.1

a) Eingabe: ein Vektor $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ des \mathbb{R}^3 .

Ausgabe: Ein Vektor in \mathbb{R}^2 , dessen Komponenten sich aus dem Eingabevektor wie folgt berechnen:

$$\begin{pmatrix} x_1 + x_2 + x_3 \\ x_1 x_3 \end{pmatrix}.$$

Bezeichnen wir die Funktion, die diese Rechnung ausführt mit F , so schreiben wir:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto F \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 + x_3 \\ x_1 x_3 \end{pmatrix}.$$

So gilt also beispielsweise

$$F \begin{pmatrix} -1 \\ 3 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 + 3 + 2 \\ (-1) \cdot 2 \end{pmatrix} = \begin{pmatrix} 4 \\ -2 \end{pmatrix}.$$

Wir nennen F eine Funktion von \mathbb{R}^3 nach \mathbb{R}^2 .

b) Betrachten wir weiter eine Funktion G von \mathbb{R}^2 nach \mathbb{R}^5 :

$$G \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 - x_2 \\ x_1 + x_2 \\ 0 \\ 2x_1 \end{pmatrix}.$$

Zum Beispiel:

$$G \begin{pmatrix} 3 \\ -5 \end{pmatrix} = \begin{pmatrix} -5 \\ 3 - (-5) \\ 3 + (-5) \\ 0 \\ 2 \cdot 3 \end{pmatrix} = \begin{pmatrix} -5 \\ 8 \\ -2 \\ 0 \\ 6 \end{pmatrix}.$$

c) Die Länge eines Vektors in \mathbb{R}^4 (oder eines Vektors in einem beliebigen Raum \mathbb{R}^n) kann als Funktion von \mathbb{R}^4 (respektive \mathbb{R}^n) nach $\mathbb{R}^1 = \mathbb{R}$ aufgefasst werden:

$$L(\mathbf{v}) = \|\mathbf{v}\|,$$

oder ausführlicher

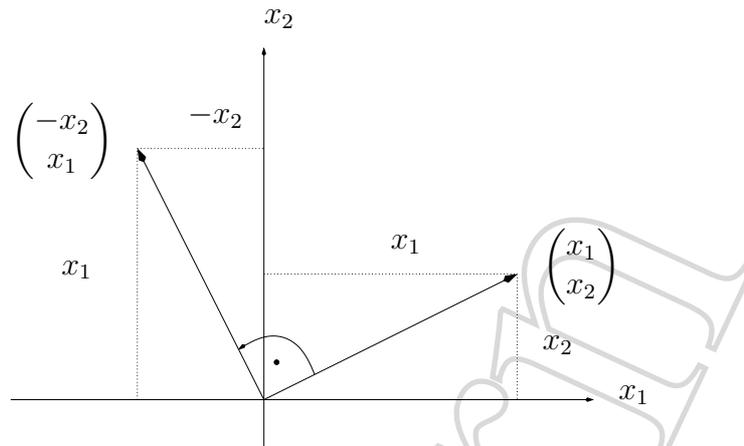
$$L \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \sqrt{x_1^2 + x_2^2 + x_3^2 + x_4^2}.$$

d) Schliesslich betrachten wir die Drehung R eines (Orts-) Vektors in \mathbb{R}^2 um 90° im Gegen-
uhrzeigersinn:

$$R \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix};$$

vgl. Abbildung 3.1.

□

Abbildung 3.1: Drehung um 90° .

3.2 Definition und Beispiele

Wir nennen eine Funktion T von \mathbb{R}^m nach \mathbb{R}^n wie wir sie in den obigen Beispielen gesehen haben eine **lineare Abbildung**, falls die folgenden zwei Bedingungen erfüllt sind:

(L1) Für alle Eingabevektoren v und w gilt

$$T(v + w) = T(v) + T(w).$$

In Worten: Anwenden einer linearen Abbildung auf eine Summe von Vektoren ergibt dasselbe Resultat, wie wenn die lineare Abbildung zuerst auf jeden Vektor einzeln angewandt wird und die entsprechenden Resultatvektoren danach addiert werden.

(L2) Für alle Eingabevektoren v und jede beliebige Zahl α gilt

$$T(\alpha \cdot v) = \alpha \cdot T(v).$$

In Worten: Anwenden einer linearen Abbildung auf ein skalares Vielfaches eines Vektors ergibt dasselbe Resultat, wie wenn die lineare Abbildung zuerst auf den ursprünglichen Vektor angewandt wird und die Skalarmultiplikation danach ausgeführt wird.

Beispiel 3.2 Wir überprüfen, ob die Funktionen in Beispiel 3.1 linear sind oder nicht.

a) Die Funktion F ist nicht linear. Betrachten wir beispielsweise

$$F \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix},$$

dann müsste mit (L2) gelten, dass

$$F \left(\alpha \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right) = \alpha \cdot F \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \alpha \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 3\alpha \\ \alpha \end{pmatrix}.$$

Nun gilt aber

$$F \left(\alpha \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right) = F \begin{pmatrix} \alpha \\ \alpha \\ \alpha \end{pmatrix} = \begin{pmatrix} 3\alpha \\ \alpha^2 \end{pmatrix}.$$

Die Gleichung

$$\begin{pmatrix} 3\alpha \\ \alpha^2 \end{pmatrix} = \begin{pmatrix} 3\alpha \\ \alpha \end{pmatrix}$$

gilt im allgemeinen nicht (nur für $\alpha = 0$ oder $\alpha = 1$). Sie müsste jedoch für jedes α (und jeden Vektor) gelten. Somit ist (L2) nicht erfüllt.

b) Die Funktion G ist linear. Wir überprüfen die beiden obigen Bedingungen.

(L1) Wir nehmen zwei beliebige Vektoren

$$\mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Dann gilt:

$$\mathbf{G}(\mathbf{v} + \mathbf{w}) = \mathbf{G} \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \end{pmatrix} = \begin{pmatrix} x_2 + y_2 \\ x_1 + y_1 - (x_2 + y_2) \\ x_1 + y_1 + (x_2 + y_2) \\ 0 \\ 2(x_1 + y_1) \end{pmatrix} = \begin{pmatrix} x_2 + y_2 \\ x_1 - x_2 + y_1 - y_2 \\ x_1 + x_2 + y_1 + y_2 \\ 0 \\ 2(x_1 + y_1) \end{pmatrix}.$$

Ausserdem haben wir

$$\begin{aligned} \mathbf{G}(\mathbf{v}) + \mathbf{G}(\mathbf{w}) &= \mathbf{G} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \mathbf{G} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_1 - x_2 \\ x_1 + x_2 \\ 0 \\ 2x_1 \end{pmatrix} + \begin{pmatrix} y_2 \\ y_1 - y_2 \\ y_1 + y_2 \\ 0 \\ 2y_1 \end{pmatrix} \\ &= \begin{pmatrix} x_2 + y_2 \\ x_1 - x_2 + y_1 - y_2 \\ x_1 + x_2 + y_1 + y_2 \\ 0 \\ 2(x_1 + y_1) \end{pmatrix}. \end{aligned}$$

Wir sehen, dass (L1) immer erfüllt ist.

(L2) Ferner erhalten wir für eine beliebige Zahl α :

$$\begin{aligned} \mathbf{G}(\alpha \cdot \mathbf{v}) &= \mathbf{G}\left(\alpha \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = \mathbf{G}\begin{pmatrix} \alpha x_1 \\ \alpha x_2 \end{pmatrix} = \begin{pmatrix} \alpha x_2 \\ \alpha x_1 - \alpha x_2 \\ \alpha x_1 + \alpha x_2 \\ 0 \\ 2\alpha x_1 \end{pmatrix} = \alpha \cdot \begin{pmatrix} x_2 \\ x_1 - x_2 \\ x_1 + x_2 \\ 0 \\ 2x_1 \end{pmatrix} \\ &= \alpha \cdot \mathbf{G}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \alpha \cdot \mathbf{G}(\mathbf{v}). \end{aligned}$$

Also ist auch (L2) generell erfüllt.

c) Die Abbildung L ist nicht linear. Ein einfaches Beispiel zeigt dies:

$$L\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \sqrt{1^2 + 1^2 + 1^2 + 1^2} = 2.$$

Mit (L2) müsste dann gelten:

$$L\begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix} = L\left((-1) \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}\right) = (-1) \cdot L\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = -2.$$

Dies stimmt aber nicht, denn

$$L\begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix} = \sqrt{(-1)^2 + (-1)^2 + (-1)^2 + (-1)^2} = 2.$$

d) Drehungen sind lineare Abbildungen. Man kann dies überprüfen wie in b). Allerdings lässt sich in dieser geometrischen Situation auch intuitiv argumentieren:

(L1) Drehen einer Summe von Vektoren ergibt dasselbe, wie wenn die Vektoren separat gedreht und danach addiert werden.

(L2) Ändern der Länge eines Vektors (was der Skalarmultiplikation entspricht) vor oder nach einer Drehung ergibt dasselbe Ergebnis.

□

3.3 Lineare Abbildungen und Matrizen

Betrachten wir eine lineare Abbildung T von \mathbb{R}^m nach \mathbb{R}^n , wobei m und n gegebene Zahlen seien. Kurz,

$$T : \mathbb{R}^m \rightarrow \mathbb{R}^n.$$

Für einen Eingabevektor v in \mathbb{R}^m ist dann der Ausgabevektor $T(v)$ unter der Anwendung der linearen Abbildung T ein Vektor in \mathbb{R}^n . Nun lässt sich jeder Vektor v in \mathbb{R}^m als Linearkombination der Standardbasis (vgl. (2.31) und (2.32)) in \mathbb{R}^m schreiben:

$$v = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = x_1 \cdot e_1 + x_2 \cdot e_2 + \cdots + x_m \cdot e_m.$$

Dann gilt mit der Eigenschaft (L1):

$$\begin{aligned} T(v) &= T(x_1 \cdot e_1 + x_2 \cdot e_2 + \cdots + x_m \cdot e_m) \\ &= T(x_1 \cdot e_1) + T(x_2 \cdot e_2) + \cdots + T(x_m \cdot e_m). \end{aligned}$$

Ferner erhalten wir unter Anwendung von (L2):

$$T(x_1 \cdot e_1) + T(x_2 \cdot e_2) + \cdots + T(x_m \cdot e_m) = x_1 \cdot T(e_1) + x_2 \cdot T(e_2) + \cdots + x_m \cdot T(e_m),$$

und daher

$$\boxed{T(v) = x_1 \cdot T(e_1) + x_2 \cdot T(e_2) + \cdots + x_m \cdot T(e_m)}. \quad (3.3)$$

Es folgt also:

Der Resultatvektor $T(v)$ ist eine Linearkombination der Vektoren

$$T(e_1), \quad T(e_2), \quad \dots, \quad T(e_m),$$

und zwar wieder mit denselben Koeffizienten x_1, x_2, \dots, x_m von v .

Die obige Formel ist berechnungstechnisch wichtig. Sie zeigt, dass es beim Rechnen mit einer linearen Abbildung genügt, die Ausgabevektoren der Basisvektoren zu ermitteln. Dies ist besonders dann praktisch, wenn die lineare Abbildung auf eine ganze (möglicherweise grosse) Menge von Vektoren angewandt werden soll (zum Beispiel bei der Analyse von grossen Datenmengen). Wir zeigen dies am Beispiel 3.1 b).

Beispiel 3.4 Die Standardbasis in \mathbb{R}^2 besteht aus den zwei Vektoren

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

Anwenden der linearen Abbildung G auf die Standardbasis ergibt:

$$G(e_1) = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 2 \end{pmatrix}, \quad G(e_2) = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

Mit Formel (3.3) gilt dann für einen *beliebigen* Eingabevektor

$$v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

aus \mathbb{R}^2 , dass sich der entsprechende Ausgabevektor wie folgt schreiben lässt:

$$G(v) = x_1 \cdot G(e_1) + x_2 \cdot G(e_2). \quad (3.5)$$

Beispielsweise,

$$G \begin{pmatrix} 2 \\ -7 \end{pmatrix} = 2 \cdot G(e_1) + (-7) \cdot G(e_2) = 2 \cdot \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 2 \end{pmatrix} - 7 \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 9 \\ -5 \\ 0 \\ 4 \end{pmatrix}.$$

Das gleiche Resultat ergibt sich aus der ursprünglichen Definition von G . □

Im obigen Beispiel können wir die Vektoren $G(e_1)$ und $G(e_2)$ als Spalten in eine Matrix, die wir mit $[G]$ bezeichnen, schreiben¹:

$$[G] = \left(\begin{array}{c|c} G(e_1) & G(e_2) \end{array} \right) = \begin{pmatrix} 0 & 1 \\ 1 & -1 \\ 1 & 1 \\ 0 & 0 \\ 2 & 0 \end{pmatrix}.$$

Die Formel (3.5) schreiben wir dann mit Hilfe der Notation

$$G(v) = [G] \cdot v = [G] \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

wobei “ \cdot ” eine Multiplikation zwischen der Matrix $[G]$ und dem Vektor v bedeutet. Diese nennen wir deshalb auch **Matrix-Vektor-Multiplikation**.

¹Die vertikalen Striche “ $|$ ” haben in diesem Zusammenhang keine mathematische Bedeutung; sie sollen lediglich die Spaltenstruktur der einzelnen Vektoren innerhalb der Matrix unterstreichen.

Ganz allgemein definieren wir für eine lineare Abbildung $T : \mathbb{R}^m \rightarrow \mathbb{R}^n$ die zugehörige Matrix

$$[T] = (T(e_1) \mid T(e_2) \mid \cdots \mid T(e_m)),$$

wobei die Spalten von $[T]$ gerade die Vektoren $T(e_1), T(e_2), \dots, T(e_m)$ sind. Diese Matrix heisst **Matrix der linearen Abbildung T** . Sie hat n Zeilen und m Spalten. Die Formel (3.3) schreiben wir dann mit der Kurzschreibweise

$$\boxed{T(v) = [T] \cdot v = x_1 \cdot T(e_1) + x_2 \cdot T(e_2) + \cdots + x_m \cdot T(e_m),}$$

wobei

$$v = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}.$$

Das so definierte “Produkt” zwischen der Matrix $[T]$ und dem Vektor v nennen wir **Matrix-Vektor-Produkt**. In Worten berechnet es sich als

$$\begin{aligned} T(v) = & \text{(1. Komponente des Eingabevektors } v) \cdot \text{(1. Spalte der Matrix } [T]) \\ & + \text{(2. Komponente des Eingabevektors } v) \cdot \text{(2. Spalte der Matrix } [T]) \\ & + \text{(3. Komponente des Eingabevektors } v) \cdot \text{(3. Spalte der Matrix } [T]) \\ & \vdots \end{aligned}$$

Beispiel 3.6 Betrachten wir die Drehung aus Beispiel 3.1 d). Die Matrix der linearen Abbildung R ist gegeben durch

$$[R] = (R(e_1) \mid R(e_2)) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Mit Hilfe der Matrix-Vektor-Multiplikation können wir nun jeden beliebigen Vektor um 90° drehen. Beispielsweise den Vektor

$$v = \begin{pmatrix} -11 \\ 35 \end{pmatrix}.$$

Hier erhalten wir

$$R(v) = [R] \cdot v = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} -11 \\ 35 \end{pmatrix} = -11 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} + 35 \cdot \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} -35 \\ -11 \end{pmatrix}.$$

□

Beispiel 3.7 Sei S die lineare Abbildung, welche wie folgt auf Vektoren des \mathbb{R}^2 wirkt: Erste Komponente wird um einen Faktor 3 gestreckt, zweite Komponente wird um einen Faktor $\frac{1}{2}$ gestaucht, d.h.,

$$S \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 3x_1 \\ x_2/2 \end{pmatrix}.$$

Um die Matrix $[S]$ zu finden, wenden wir S auf die Standardbasis in \mathbb{R}^2 an:

$$[S] = (\mathbf{S}(e_1) \mid \mathbf{S}(e_2)) = \begin{pmatrix} 3 & 0 \\ 0 & 1/2 \end{pmatrix}.$$

Zum Beispiel:

$$S \begin{pmatrix} 11 \\ 32 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 1/2 \end{pmatrix} \cdot \begin{pmatrix} 11 \\ 32 \end{pmatrix} = \begin{pmatrix} 33 \\ 16 \end{pmatrix}.$$

□

Beispiel 3.8 Gegeben sei die Matrix einer linearen Abbildung L :

$$[L] = \begin{pmatrix} -2 & 3 & 0 \\ -3 & -1 & 5 \end{pmatrix}.$$

1. Wie sieht diese lineare Abbildung konkret aus?

Die Matrix $[L]$ hat 2 Zeilen und 3 Spalten, d.h., L ist eine lineare Abbildung mit Eingabevektoren in \mathbb{R}^3 und Ausgabevektoren in \mathbb{R}^2 . Für einen beliebigen Eingabevektor gilt

$$\begin{aligned} L \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= [L] \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -2 & 3 & 0 \\ -3 & -1 & 5 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= x_1 \cdot \begin{pmatrix} -2 \\ -3 \end{pmatrix} + x_2 \cdot \begin{pmatrix} 3 \\ -1 \end{pmatrix} + x_3 \cdot \begin{pmatrix} 0 \\ 5 \end{pmatrix} = \begin{pmatrix} -2x_1 + 3x_2 \\ -3x_1 - x_2 + 5x_3 \end{pmatrix}. \end{aligned}$$

2. Wie wird der Einheitswürfel in \mathbb{R}^3 durch die Abbildung L transformiert?

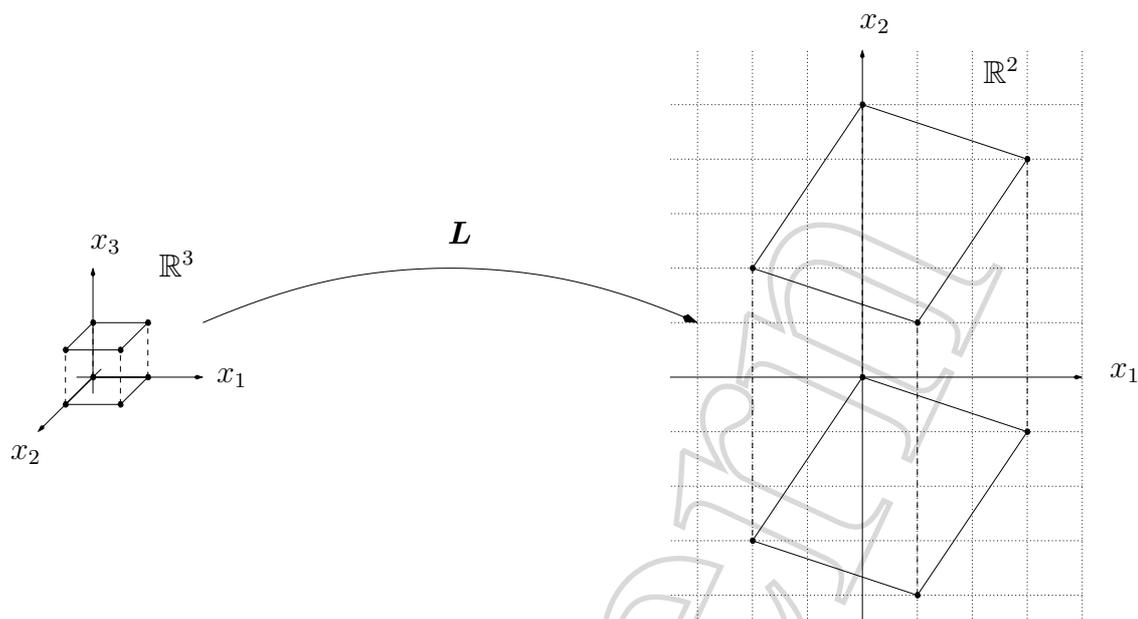
Die Eckpunkte des Einheitswürfels entsprechen den Ortsvektoren

$$\begin{aligned} \mathbf{p}_1 &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, & \mathbf{p}_2 &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, & \mathbf{p}_3 &= \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, & \mathbf{p}_4 &= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \\ \mathbf{p}_5 &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, & \mathbf{p}_6 &= \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, & \mathbf{p}_7 &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, & \mathbf{p}_8 &= \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}. \end{aligned}$$

Wir wenden die lineare Abbildung L auf diese Eckpunkte an:

$$\begin{aligned} L(\mathbf{p}_1) &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}, & L(\mathbf{p}_2) &= \begin{pmatrix} -2 \\ -3 \end{pmatrix}, & L(\mathbf{p}_3) &= \begin{pmatrix} 1 \\ -4 \end{pmatrix}, & L(\mathbf{p}_4) &= \begin{pmatrix} 3 \\ -1 \end{pmatrix}, \\ L(\mathbf{p}_5) &= \begin{pmatrix} 0 \\ 5 \end{pmatrix}, & L(\mathbf{p}_6) &= \begin{pmatrix} -2 \\ 2 \end{pmatrix}, & L(\mathbf{p}_7) &= \begin{pmatrix} 1 \\ 1 \end{pmatrix}, & L(\mathbf{p}_8) &= \begin{pmatrix} 3 \\ 4 \end{pmatrix}. \end{aligned}$$

Das entstehende Bild ist in Abbildung 3.2 gezeigt. Die lineare Abbildung L erlaubt also die Darstellung eines 3-dimensionalen Objektes in der Ebene \mathbb{R}^2 unter einer gewissen Perspektive. Die "Blickwinkel" können durch Anpassen der Matrixeinträge in $[L]$ beliebig geändert werden.

Abbildung 3.2: Lineare Abbildung von \mathbb{R}^3 nach \mathbb{R}^2 .

3. Welche Vektoren in \mathbb{R}^3 werden auf den Nullvektor in \mathbb{R}^2 abgebildet?

Es geht also darum alle Vektoren

$$\mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

zu finden, für welche gilt:

$$\mathbf{L}(\mathbf{v}) = \mathbf{0}.$$

Ausgeschrieben:

$$\mathbf{L} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -2x_1 + 3x_2 \\ -3x_1 - x_2 + 5x_3 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Dies führt auf ein lineares Gleichungssystem mit zwei Gleichungen und drei Unbekannten. Alle Lösungen sind gegeben in der Form

$$x_1 = \frac{15}{11}s, \quad x_2 = \frac{10}{11}s, \quad x_3 = s,$$

für beliebige Werte des Parameters s . Die Menge aller dieser Vektoren bilden einen linearen Unterraum von \mathbb{R}^3 , der **Kern von L** genannt und mit $\ker(L)$ bezeichnet wird. Hier

gilt:

$$\ker(\mathbf{L}) = \left\{ \begin{pmatrix} 15s/11 \\ 10s/11 \\ s \end{pmatrix} : s \text{ eine beliebige Zahl} \right\} = \text{span} \left(\begin{pmatrix} 15/11 \\ 10/11 \\ 1 \end{pmatrix} \right).$$

Es folgt, dass die Dimension des Kerns von \mathbf{L} gleich 1 ist.

□

Ganz allgemein definieren wir:

Der **Kern** einer linearen Abbildung $\mathbf{T} : \mathbb{R}^m \rightarrow \mathbb{R}^n$ ist die Menge aller Vektoren \mathbf{v} in \mathbb{R}^m für die gilt, dass $\mathbf{T}(\mathbf{v}) = \mathbf{0}$. Diese Menge ist ein linearer Unterraum von \mathbb{R}^m und wird mit $\ker(\mathbf{T})$ bezeichnet.

3.4 Verknüpfte lineare Abbildungen

Betrachten wir den Vektor

$$\mathbf{v} = \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix}$$

und die linearen Abbildungen \mathbf{R} und \mathbf{L} aus den obigen Beispielen 3.6 respektive 3.8. Dann gilt

$$\mathbf{L}(\mathbf{v}) = [\mathbf{L}] \cdot \mathbf{v} = \begin{pmatrix} -13 \\ 2 \end{pmatrix}.$$

Auf den Resultatvektor können wir nun weiter die lineare Abbildung \mathbf{R} anwenden:

$$\mathbf{R}(\mathbf{L}(\mathbf{v})) = [\mathbf{R}] \cdot ([\mathbf{L}] \cdot \mathbf{v}) = [\mathbf{R}] \cdot \begin{pmatrix} -13 \\ 2 \end{pmatrix} = \begin{pmatrix} -2 \\ -13 \end{pmatrix}. \quad (3.9)$$

Zusammengefasst haben wir also zuerst die Abbildung \mathbf{L} und dann die Abbildung \mathbf{R} auf den Eingabevektor \mathbf{v} angewandt. Dieses ‘‘Hintereinanderschalten’’ von Funktionen nennt man **Verknüpfung**. Man verwendet die Notation

$$(\mathbf{R} \circ \mathbf{L})(\mathbf{v}) = \mathbf{R}(\mathbf{L}(\mathbf{v})).$$

Mit $\mathbf{R} \circ \mathbf{L}$ ist also die Vektorfunktion gemeint, die aus einem Eingabevektor \mathbf{v} direkt den Ausgabevektor $\mathbf{R}(\mathbf{L}(\mathbf{v}))$ produziert; vgl. Abbildung 3.3. Man beachte, dass die Reihenfolge der Verknüpfung wichtig ist. So ist beispielsweise die Verknüpfung $\mathbf{L} \circ \mathbf{R}$ im obigen Zusammenhang gar nicht definiert (warum?).

Es lässt sich zeigen, dass die Verknüpfung von linearen Abbildungen wieder eine lineare Abbildung ist. Insbesondere gibt es also auch für sie eine zugehörige Matrix, die wir mit $[\mathbf{R} \circ \mathbf{L}]$

oder auch

$$\begin{aligned} (\mathbf{R} \circ \mathbf{L})(\mathbf{v}) &= \mathbf{R}(\mathbf{L}(\mathbf{v})) = \mathbf{R}([\mathbf{L}] \cdot \mathbf{v}) = \mathbf{R} \begin{pmatrix} -2x_1 + 3x_2 \\ -3x_1 - x_2 + 5x_3 \end{pmatrix} \\ &= [\mathbf{R}] \cdot \begin{pmatrix} -2x_1 + 3x_2 \\ -3x_1 - x_2 + 5x_3 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} -2x_1 + 3x_2 \\ -3x_1 - x_2 + 5x_3 \end{pmatrix} \\ &= \begin{pmatrix} 3x_1 + x_2 - 5x_3 \\ -2x_1 + 3x_2 \end{pmatrix}. \end{aligned}$$

Die Resultate sind in der Tat identisch.

Ganz allgemein gilt:

Für zwei lineare Abbildungen

$$\mathbf{S} : \mathbb{R}^m \rightarrow \mathbb{R}^n, \quad \mathbf{T} : \mathbb{R}^n \rightarrow \mathbb{R}^p$$

ist die Verknüpfung

$$\mathbf{T} \circ \mathbf{S} : \mathbb{R}^m \rightarrow \mathbb{R}^p$$

gegeben durch

$$(\mathbf{T} \circ \mathbf{S})(\mathbf{v}) = \mathbf{T}(\mathbf{S}(\mathbf{v}))$$

ebenfalls eine lineare Abbildung. Die zugehörige Matrix berechnet sich als

$$[\mathbf{T} \circ \mathbf{S}] = (([\mathbf{T}] \cdot \mathbf{s}_1) \mid ([\mathbf{T}] \cdot \mathbf{s}_2) \mid \cdots \mid ([\mathbf{T}] \cdot \mathbf{s}_m)),$$

wobei die Vektoren $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m$ die Spalten der Matrix $[\mathbf{S}]$ sind. Die Matrix $[\mathbf{T} \circ \mathbf{S}]$ heisst auch **Matrixprodukt** der Matrizen $[\mathbf{T}]$ und $[\mathbf{S}]$ und wird bezeichnet mit

$$[\mathbf{T}] \cdot [\mathbf{S}] = [\mathbf{T} \circ \mathbf{S}].$$

Wir halten folgende wichtige Bemerkungen fest:

1. Im Allgemeinen gilt *nicht* (!): $[\mathbf{T}] \cdot [\mathbf{S}] = [\mathbf{S}] \cdot [\mathbf{T}]$.
2. Im Allgemeinen ist die Matrixmultiplikation *nicht* die eintragsweise Multiplikation der einzelnen Matrizen.
3. Das Matrixprodukt kann nur dann berechnet werden, wenn die erste Matrix gleich viele Spalten wie die zweite Matrix Zeilen hat (wieso?).

Beispiel 3.10 In der Ebene \mathbb{R}^2 sei

$$\begin{aligned} \mathbf{S} &= \text{Spiegelung an der } x_1\text{-Achse,} \\ \mathbf{T} &= \text{Senkrechte Projektion auf die } x_2\text{-Achse,} \end{aligned}$$

d.h.,

$$\mathbf{S} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}, \quad \mathbf{T} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ x_2 \end{pmatrix}.$$

Die zugehörigen Matrizen sind:

$$[\mathbf{S}] = (\mathbf{S}(e_1) \mid \mathbf{S}(e_2)) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad [\mathbf{T}] = (\mathbf{T}(e_1) \mid \mathbf{T}(e_2)) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

In diesem Beispiel sind viele verschiedene Verknüpfungen möglich. Beispielsweise:

1. $\mathbf{T} \circ \mathbf{S}$: Dies bedeutet, dass zuerst eine Spiegelung an der x_1 -Achse durchgeführt und der resultierende Vektor dann auf die x_2 -Achse projiziert wird. Aus einem Vektor

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

wird somit der Vektor

$$[\mathbf{T} \circ \mathbf{S}] \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ -x_2 \end{pmatrix}.$$

Die zugehörige Matrix erhalten wir aus der Matrixmultiplikation

$$[\mathbf{T} \circ \mathbf{S}] = [\mathbf{T}] \cdot [\mathbf{S}] = \left(\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mid \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right) = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}.$$

2. In diesem Beispiel sind $[\mathbf{S} \circ \mathbf{T}]$ und $[\mathbf{T} \circ \mathbf{S}]$ identisch. Dies lässt sich sowohl geometrisch als auch durch Nachrechnen überprüfen. Dies ist aber typischerweise *nicht* wahr!
3. $\mathbf{T} \circ \mathbf{T}$: Wir projizieren zweimal auf die x_2 -Achse. Das ergibt dasselbe Resultat, wie wenn nur einmal projiziert wird. Tatsächlich sehen wir mit der Matrixmultiplikation:

$$[\mathbf{T} \circ \mathbf{T}] = [\mathbf{T}] \cdot [\mathbf{T}] = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = [\mathbf{T}].$$

□

Bemerkung 3.11 In MATLAB oder OCTAVE können Produkte von Matrizen und Vektoren mit der Operation "*" durchgeführt werden.

```
octave:1> L = [-2 3 0
              -3 -1 5]
```

L =

```
-2 3 0
-3 -1 5
```

```
octave:2> v=[1
            0
            0]
```

```
v =
```

```
1
0
0
```

```
octave:3> L*v
```

```
ans =
```

```
-2
-3
```

```
octave:4> R = [0 -1
                1  0]
```

```
R =
```

```
0 -1
1  0
```

```
octave:5> R*L
```

```
ans =
```

```
3  1 -5
-2 3  0
```

3.5 Umkehrung von linearen Abbildungen

Wir betrachten eine lineare Abbildung $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ (d.h., Eingabe- und Ausgaberaum sind gleich). Angenommen: wir wenden T auf einen *beliebigen* Eingabevektor v an und erhalten als Ausgabe den Vektor w , d.h.,

$$w = T(v).$$

Wir stellen uns die Frage: Gibt es eine lineare Abbildung S , die die lineare Abbildung T “umkehrt” und aus w wieder *genau eindeutig* den ursprünglichen Vektor v produziert? D.h.,

$$v = S(w).$$

Anders ausgedrückt soll

$$(S \circ T)(v) = S(\underbrace{T(v)}_{=w}) = v$$

gelten für jeden beliebigen Vektor aus \mathbb{R}^n . Wenn eine solche “umgekehrte” lineare Abbildung S existiert, dann nennen wir sie **inverse Abbildung von T** und schreiben:

$$S = T^{-1},$$

oder in Matrixform

$$[S] = [T^{-1}] = [T]^{-1}.$$

Die Matrix $[T]^{-1}$ heisst **inverse Matrix von $[T]$** .

Die Antwort, ob eine inverse Abbildung existiert, liegt im Kern $\ker(T)$ der linearen Abbildung T . Es gilt:

Zu einer linearen Abbildung $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ gibt es genau dann eine inverse Abbildung T^{-1} , falls der Kern $\ker(T)$ von T nur aus dem Nullvektor besteht. Äquivalent bedeutet dies, dass die Matrix $[T]$ von T *regulär* ist, d.h., $\det([T]) \neq 0$.

Beispiel 3.12

- a) Es bezeichne R die lineare Abbildung in \mathbb{R}^2 aus Beispiel 3.1 d), welche eine Drehung um 90° durchführt, d.h.,

$$R \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}, \quad [R] = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Gibt es eine inverse Abbildung? Ja, denn

$$\det([R]) = 1 \neq 0.$$

Wie lautet die inverse Abbildung? Geometrisch wird ein um 90° gedrehter Vektor durch eine Drehung um -90° auf den ursprünglichen Vektor zurückgedreht. Die inverse Abbildung ist also wie folgt gegeben:

$$R^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix}, \quad [R]^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

- b) Es sei T die Abbildung in \mathbb{R}^2 aus Beispiel 3.10, welche einen Vektor senkrecht auf die x_2 -Achse projiziert:

$$T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ x_2 \end{pmatrix}, \quad [T] = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Gibt es eine inverse lineare Abbildung? Nein, denn

$$\det([T]) = 0.$$

Dazu folgende Erklärung: Betrachten wir beispielsweise den Vektor

$$\mathbf{v} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}.$$

Dann gilt

$$\mathbf{T}(\mathbf{v}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Genau das gleiche Resultat würden wir aber zum Beispiel auch beim Anwenden von \mathbf{T} auf den Vektor

$$\mathbf{u} = \begin{pmatrix} -5 \\ 0 \end{pmatrix}$$

erhalten. Es gibt also mehr als nur einen Vektor, der durch Anwenden der linearen Abbildung \mathbf{T} den Nullvektor ergibt. Eine eindeutige Umkehrung kann daher nicht möglich sein (wir bemerken insbesondere, dass hier der Kern $\ker(\mathbf{T})$ nicht nur aus dem Nullvektor besteht).

□

Bemerkung 3.13 In MATLAB oder OCTAVE lassen sich inverse Matrizen einfach mit dem Befehl "inv" berechnen. Beim Versuch, eine singuläre Matrix zu invertieren, wird üblicherweise eine Fehlermeldung ausgegeben. Wir betrachten die Matrizen aus dem vorherigen Beispiel 3.12:

```
octave:1> R = [0 -1
              1  0]
```

```
R =
```

```
  0  -1
  1   0
```

```
octave:2> inv(R)
```

```
ans =
```

```
  0  1
 -1  0
```

```
octave:3> T = [0 0
               0 1]
```

```
T =
```

```
  0  0
  0  1
```

```
octave:4> inv(T)
```

```
warning: inverse: matrix singular to machine precision, rcond = 0
```

```
ans =
```

```
0 0
0 1
```

Die Berechnung von inversen Matrizen wird in der Praxis, wenn immer möglich, vermieden, da dies typischerweise ein sehr aufwendiger Prozess ist.

3.6 Übungsaufgaben

3.1. Sind die folgenden Abbildungen linear? Wenn ja, bestimmen Sie die zugehörige Matrix.

a) $F \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5x_1 \\ -x_2 + 6x_3 \end{pmatrix}$

b) $G \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \frac{1}{2}(x_1 + x_2)$

c) $H \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} x_1 \\ 1 \\ 1 \end{pmatrix}$

d) $J(x) = \begin{pmatrix} 2x \\ x \end{pmatrix}$

3.2. Gegeben sei der Vektor

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

Weiter sei P die lineare Abbildung, die einen beliebigen Vektor $\mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ senkrecht auf den Vektor \mathbf{w} projiziert.

- Finden Sie eine Formel für $P(\mathbf{v})$ mit Hilfe der Projektionseigenschaft.
- Finden Sie die zu P gehörige Matrix $[P]$.
- Zeigen Sie, dass $[P] \cdot [P] = [P]$ und geben Sie eine Interpretation dieser Tatsache.

3.3. Es seien

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} -1 & 2 \\ -3 & -4 \end{pmatrix}.$$

Berechnen Sie die Matrizen $(\mathbf{A} + \mathbf{B})^2$ und $\mathbf{A}^2 + 2\mathbf{A} \cdot \mathbf{B} + \mathbf{B}^2$; hier bedeutet $\mathbf{A}^2 = \mathbf{A} \cdot \mathbf{A}$, etc. Was fällt auf, und wie lässt sich die Beobachtung erklären?

3.4. Es sei $\mathbf{L} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ die lineare Abbildung mit der Matrix

$$[\mathbf{L}] = \begin{pmatrix} 3 & 4 \\ 2 & 1 \end{pmatrix}.$$

a) Skizzieren Sie, was mit dem Einheitsquadrat $Q = P_1P_2P_3P_4$ mit den Eckpunkten

$$P_1 = (0, 0), \quad P_2 = (1, 0), \quad P_3 = (1, 1), \quad P_4 = (0, 1)$$

unter Anwendung der linearen Abbildung \mathbf{L} geschieht.

b) Wie ist das Verhältnis der Flächen in a)? Vergleichen Sie dieses mit $|\det([\mathbf{L}])|$.

3.5. Wir betrachten folgende Rotationen in \mathbb{R}^3 .

\mathbf{R}_1 : Rotation um die x_1 -Achse um 90° (im Gegenuhrzeigersinn),

\mathbf{R}_2 : Rotation um die x_2 -Achse um 90° (im Gegenuhrzeigersinn).

a) Bestimmen Sie die zu \mathbf{R}_1 bzw. \mathbf{R}_2 gehörigen Matrizen $[\mathbf{R}_1]$ und $[\mathbf{R}_2]$.

b) Berechnen Sie die Matrizen $[\mathbf{R}_1 \circ \mathbf{R}_2]$ und $[\mathbf{R}_2 \circ \mathbf{R}_1]$ der entsprechenden Verknüpfungen mit Hilfe der *Matrixmultiplikation*.

c) Berechnen Sie $(\mathbf{R}_2 \circ \mathbf{R}_1)(\mathbf{v})$ für $\mathbf{v} = \begin{pmatrix} 3 \\ -2 \\ 15 \end{pmatrix}$.

d) Gibt es eine Matrix $[\mathbf{S}_1]$, so dass

$$(\mathbf{R}_1 \circ \mathbf{S}_1)(\mathbf{v}) = (\mathbf{S}_1 \circ \mathbf{R}_1)(\mathbf{v}) = \mathbf{v}$$

für alle Vektoren $\mathbf{v} \in \mathbb{R}^3$ gilt? Wie kann \mathbf{S}_1 geometrisch interpretiert werden?

3.6. Bestimmen Sie jeweils die zur linearen Abbildung \mathbf{T} gehörige Matrix $[\mathbf{T}]$:

$$\text{a) } \mathbf{T} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 - x_2 \\ x_1 + x_2 \end{pmatrix} \quad \text{b) } \mathbf{T} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 + x_3 \\ x_1 + 5x_2 \\ x_3 \end{pmatrix}$$

$$\text{c) } \mathbf{T} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 - x_1 \\ x_1 + 3x_2 \\ x_1 - x_2 \end{pmatrix} \quad \text{d) } \mathbf{T} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

3.7. Gegeben sind die Matrix $[\mathbf{T}]$ und der Vektor \mathbf{x} . Bestimmen Sie den Vektor $\mathbf{T}(\mathbf{x})$:

$$\text{a) } [\mathbf{T}] = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} 3 \\ -2 \end{pmatrix} \qquad \text{b) } [\mathbf{T}] = \begin{pmatrix} -1 & 1 \\ 2 & 4 \\ 7 & 8 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

3.8. Berechnen Sie im folgenden $\mathbf{T}(\mathbf{x})$ unter Verwendung der zu \mathbf{T} gehörigen Matrix $[\mathbf{T}]$. Kontrollieren Sie das Resultat anschliessend durch direkte Berechnung.

$$\text{a) } \mathbf{T} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -x_1 + x_2 \\ x_2 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} -1 \\ 4 \end{pmatrix}$$

$$\text{b) } \mathbf{T} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 - x_2 + x_3 \\ x_2 + x_3 \\ 0 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} 2 \\ 1 \\ -3 \end{pmatrix}.$$

3.9. Bestimmen Sie die zur linearen Abbildung $\mathbf{T} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ gehörige Matrix $[\mathbf{T}]$.

- a) \mathbf{T} = Spiegelung an der Geraden $y = x$.
- b) \mathbf{T} = Rotation am Ursprung um $\vartheta = 30^\circ$.
- c) \mathbf{T} = Rotation am Ursprung um $\vartheta = -60^\circ$.

3.10. In den folgenden Beispielen ist \mathbf{T} eine Verknüpfung von linearen Abbildungen $\mathbb{R}^3 \rightarrow \mathbb{R}^3$. Bestimmen Sie jeweils die zugehörige Matrix $[\mathbf{T}]$ als Matrixprodukt der Matrizen der einzelnen linearen Abbildungen.

- a) \mathbf{T} = Reflexion an der yz -Ebene, dann orthogonale Projektion auf die xz -Ebene.
- b) \mathbf{T} = Streckung am Ursprung mit Faktor 2, dann Rotation um 45° um die y -Achse, schliesslich Spiegelung an der y -Achse.
- c) \mathbf{T} = Rotation um 270° um die x -Achse, dann Rotation um 90° um die y -Achse, schliesslich Rotation um 180° um die z -Achse.

3.11. Die lineare Abbildung $\mathbf{F} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ ist durch die Matrix

$$[\mathbf{F}] = \begin{pmatrix} 2 & 2 & -4 \\ -5 & 3 & 2 \\ 1 & -3 & 2 \end{pmatrix}$$

gegeben. Lösen Sie die Gleichungen

$$\text{a) } \mathbf{F} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\text{b) } \mathbf{F} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}.$$

4

Kleinste Quadrate und diskrete Fouriertransformation

Angenommen, die Beschleunigung a eines Partikels, welches in der Atmosphäre aus der Ruhe zu fallen beginnt, soll durch Messungen bestimmt werden. In einem Laborversuch wird dies wie folgt simuliert: Es werden drei Strecken $s_1 < s_2 < s_3$ abgemessen. Dann wird das Partikel aus dem Ruhezustand über jede der Strecken fallen gelassen und die zugehörigen Fallzeiten $t_1 < t_2 < t_3$ werden gemessen; vgl. Abbildung 4.1. Wenn man davon ausgeht, dass die Bewegung gleichmässig beschleunigt verläuft, dann ist die Fallbeschleunigung a aus den Daten s_i und t_i abzulesen.

Zum Versuch werden einige Vorüberlegungen angestellt: Das Fallgesetz im Vakuum hat die Form

$$s = \frac{1}{2}at^2,$$

oder nach t aufgelöst

$$t = \sqrt{\frac{2s}{a}}.$$

Wir erwarten, dass die Grösse

$$c = \sqrt{\frac{2}{a}} \tag{4.1}$$

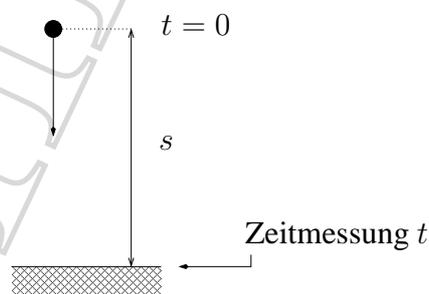


Abbildung 4.1: Freier Fall eines Partikels.

konstant ist (da die Beschleunigung a als gleichmässig angenommen wird). Folglich müssen die gemessenen Fallzeiten gemäss dieser Idealisierung eine Beziehung der Art

$$t_i = c\sqrt{s_i} \quad (4.2)$$

erfüllen. Um die Auswertung einfach zu halten, werden die Fallstrecken

$$s_1 = 0.64m, \quad s_2 = 1.00m, \quad s_3 = 1.96m$$

gewählt. Die Fallzeiten werden mit einer Lichtschranke auf schätzungsweise $0.02s$ genau bestimmt. Die im Versuch ermittelten Fallzeiten (Mittelwerte über je 5 Versuche) betragen

$$t_1 = 0.36s, \quad t_2 = 0.46s, \quad t_3 = 0.64s.$$

Benutzen wir die Formel (4.2), um aus den obigen Messungen die Konstante c zu ermitteln, so erhalten wir

$$c = \frac{t_1}{\sqrt{s_1}} = \frac{0.36s}{0.8m^{\frac{1}{2}}} = 0.45sm^{-\frac{1}{2}}.$$

Für die Beschleunigung ergibt sich damit aus (4.1):

$$a = \frac{2}{c^2} = \frac{2s_1}{t_1^2} \approx 9.88ms^{-2}. \quad (4.3)$$

Mit denselben Rechnungen, angewandt auf die zwei weiteren Datenpaare, erhalten wir:

$$a = \frac{2s_2}{t_2^2} \approx 9.45ms^{-2}, \quad a = \frac{2s_3}{t_3^2} \approx 9.57ms^{-2}.$$

Offensichtlich besteht ein gewisser Widerspruch zwischen den gemessenen Daten und dem angenommenen Gesetz (4.2). Die Beschleunigung müsste eigentlich für alle Messungen gleich sein. Dass dies nicht der Fall ist, ist aufgrund der Messfehler bei den Zeitmessungen nicht überraschend (die drei Werte $\sqrt{s_i}$ können als exakt angenommen werden).

Wir wollen nun einen "Kompromiss" eingehen, um eine möglichst gute Näherung der gesuchten Beschleunigung a zu finden. Hier bietet sich die sogenannte *Methode der kleinsten Quadrate* an. Wir besprechen die Idee am gegebenen Beispiel: Zunächst werden die Daten in den Vektoren

$$\mathbf{w} = \begin{pmatrix} \sqrt{s_1} \\ \sqrt{s_2} \\ \sqrt{s_3} \end{pmatrix} = \begin{pmatrix} 0.8 \\ 1.0 \\ 1.4 \end{pmatrix}$$

und

$$\mathbf{t} = \begin{pmatrix} 0.36 \\ 0.46 \\ 0.64 \end{pmatrix}$$

gespeichert. Da ein Gesetz der Art $t = c\sqrt{s}$ angenommen wird, sollte eine Darstellung für \mathbf{t} von der Art

$$\mathbf{t} = c \cdot \mathbf{w}$$

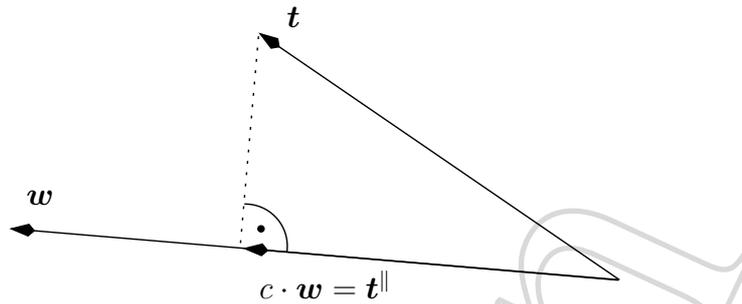


Abbildung 4.2: Kompromisslösung mittels kleinster Quadrate.

gelten. Wegen der Widersprüche in den Daten ist dies aber nicht exakt erfüllt. Wir stellen deshalb die folgende Frage: *Wie muss die Konstante c gewählt werden, damit die Differenz zwischen den Vektoren t und $c \cdot w$ möglichst klein wird?* Mathematisch ausgedrückt, wollen wir c finden, sodass

$$\|t - c \cdot w\| \longrightarrow \text{minimal.}$$

In Abbildung 4.2 ist die Situation grafisch dargestellt. Daraus sehen wir, dass die Differenz $\|t - c \cdot w\|$ genau dann minimal wird, wenn der Vektor $c \cdot w$ die Orthogonalprojektion des Vektors t auf den Vektor w ist.

Die Aufgabe lässt sich somit mit der Projektionseigenschaft (2.12) lösen. Es gilt

$$c \cdot w = t^{\parallel} = \frac{\langle t, w \rangle}{\|w\|^2} \cdot w,$$

und deshalb

$$c = \frac{\langle t, w \rangle}{\|w\|^2} = 0.456666 \dots \text{sm}^{-\frac{1}{2}}.$$

Die gesuchte Beschleunigung ergibt sich dann wie in (4.3), d.h.,

$$a \approx 9.59 \text{ms}^{-2}.$$

4.1 Orthogonalprojektionen und kleinste Quadrate

Im letzten Abschnitt haben wir eine Kompromisslösung für *eine* Variable c gefunden. Wir wollen nun ein Beispiel betrachten, welches einen Kompromiss für *zwei* Variablen erfordert.

Durch die drei Punkte

$$P_1 : (-3, 0), \quad P_2 : (1, 0), \quad P_3 : (2, 1)$$

in der Ebene soll eine Gerade gelegt werden. Die Aufgabe ist in dieser Form unlösbar, denn zwei Punkte liegen auf der x_1 -Achse, der dritte jedoch nicht. Wir gehen wiederum auf eine

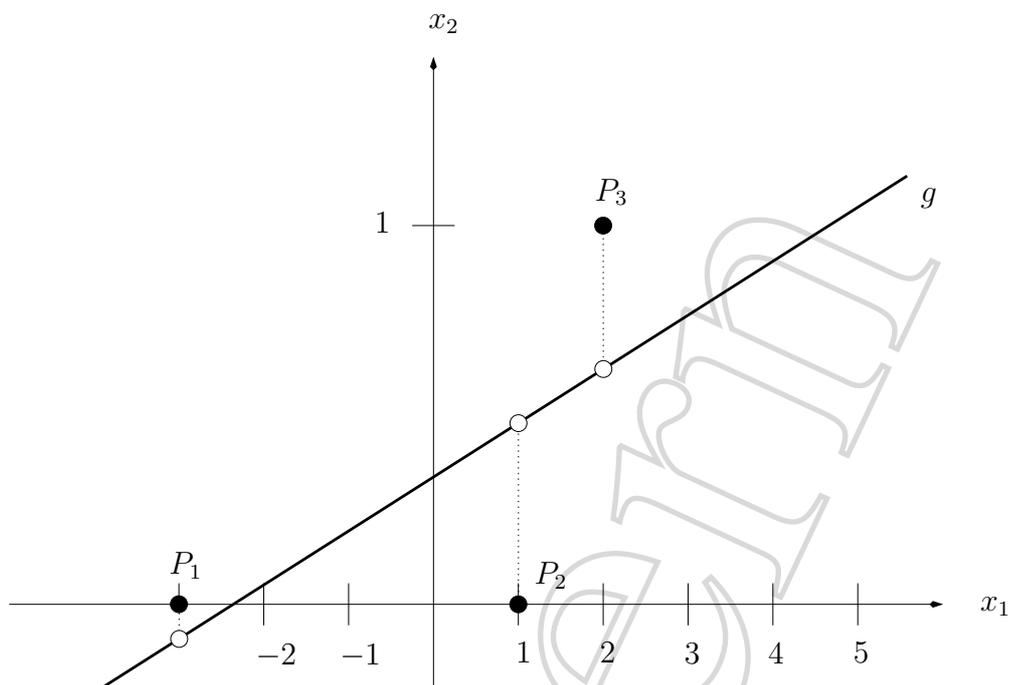


Abbildung 4.3: Gerade durch 3 Punkte.

Kompromisslösung ein, in dem wir eine Gerade finden, die möglichst “gut” durch die drei Punkte geht; siehe Abbildung 4.3.

Wir suchen eine Gerade g , welche gegeben ist durch

$$g : x_2 = ax_1 + b.$$

Hier sollen die Steigung a und der vertikale Achsenabschnitt b bestimmt werden. Setzen wir die Koordinaten der drei gegebenen Punkte in die Geradengleichung von g ein, so erhalten wir:

$$0 = -3a + b$$

$$0 = a + b$$

$$1 = 2a + b$$

Dies ist ein *überbestimmtes* lineares Gleichungssystem mit 2 Unbekannten (a und b) und drei Gleichungen. In Matrixform haben wir

$$\left(\begin{array}{cc|c} -3 & 1 & 0 \\ 1 & 1 & 0 \\ 2 & 1 & 1 \end{array} \right). \quad (4.4)$$

Dieses Gleichungssystem hat keine Lösung.

Messen wir die *Summe der Quadrate* der vertikalen Abstände der Punkte P_1 , P_2 und P_3 von der Geraden g (dies sind die “ $\cdot\cdot\cdot$ ”-Linien in Abbildung 4.3), so erhalten wir ein Mass für die *Gesamtabweichung* \mathcal{E} der Punkte von der Geraden:

$$\mathcal{E} = |-3a + b - 0|^2 + |a + b - 0|^2 + |2a + b - 1|^2. \quad (4.5)$$

Um einen möglichst optimalen Kompromiss zu finden, wollen wir a und b so wählen, dass die Abweichung \mathcal{E} minimal wird.

Dazu schreiben wir zunächst die Spaltenvektoren der erweiterten Matrix in (4.4) heraus:

$$\mathbf{w}_1 = \begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

und

$$\mathbf{v} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Dann gilt

$$\mathcal{E} = \left\| \begin{pmatrix} -3a + b - 0 \\ a + b - 0 \\ 2a + b - 1 \end{pmatrix} \right\|^2 = \left\| a \cdot \begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix} + b \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\|^2 = \|(a \cdot \mathbf{w}_1 + b \cdot \mathbf{w}_2) - \mathbf{v}\|^2. \quad (4.6)$$

Was lässt sich daraus schliessen? Der Vektor $a \cdot \mathbf{w}_1 + b \cdot \mathbf{w}_2$ liegt in der Ebene $E = \text{span}(\mathbf{w}_1, \mathbf{w}_2)$, welche aus allen Vektoren besteht, die als Linearkombinationen von \mathbf{w}_1 und \mathbf{w}_2 geschrieben werden können. Er soll minimalen Abstand zum Vektor \mathbf{v} haben, damit \mathcal{E} minimal wird. Dies ist dann der Fall, wenn der Vektor $a \cdot \mathbf{w}_1 + b \cdot \mathbf{w}_2$ die orthogonale Projektion des Vektors \mathbf{v} auf die Ebene E ist; vgl. Abbildung 4.4. Bezeichnen wir die orthogonale Projektion auf die Ebene E mit \mathbf{P}_E (dies ist eine lineare Abbildung), so soll also gelten, dass

$$\mathbf{P}_E(\mathbf{v}) = a \cdot \mathbf{w}_1 + b \cdot \mathbf{w}_2.$$

Aus Abbildung 4.4 sehen wir, dass der Vektor

$$\mathbf{v} - \mathbf{P}_E(\mathbf{v})$$

senkrecht auf der Ebene E und damit auf den beiden Vektoren \mathbf{w}_1 und \mathbf{w}_2 steht. Daraus folgt:

$$\begin{aligned} 0 &= \langle \mathbf{w}_1, \mathbf{v} - \mathbf{P}_E(\mathbf{v}) \rangle = \langle \mathbf{w}_1, \mathbf{v} - a \cdot \mathbf{w}_1 - b \cdot \mathbf{w}_2 \rangle = \langle \mathbf{w}_1, \mathbf{v} \rangle - a \langle \mathbf{w}_1, \mathbf{w}_1 \rangle - b \langle \mathbf{w}_1, \mathbf{w}_2 \rangle \\ 0 &= \langle \mathbf{w}_2, \mathbf{v} - \mathbf{P}_E(\mathbf{v}) \rangle = \langle \mathbf{w}_2, \mathbf{v} - a \cdot \mathbf{w}_1 - b \cdot \mathbf{w}_2 \rangle = \langle \mathbf{w}_2, \mathbf{v} \rangle - a \langle \mathbf{w}_2, \mathbf{w}_1 \rangle - b \langle \mathbf{w}_2, \mathbf{w}_2 \rangle. \end{aligned}$$

Daher

$$\begin{aligned} a \langle \mathbf{w}_1, \mathbf{w}_1 \rangle + b \langle \mathbf{w}_1, \mathbf{w}_2 \rangle &= \langle \mathbf{w}_1, \mathbf{v} \rangle \\ a \langle \mathbf{w}_2, \mathbf{w}_1 \rangle + b \langle \mathbf{w}_2, \mathbf{w}_2 \rangle &= \langle \mathbf{w}_2, \mathbf{v} \rangle. \end{aligned}$$

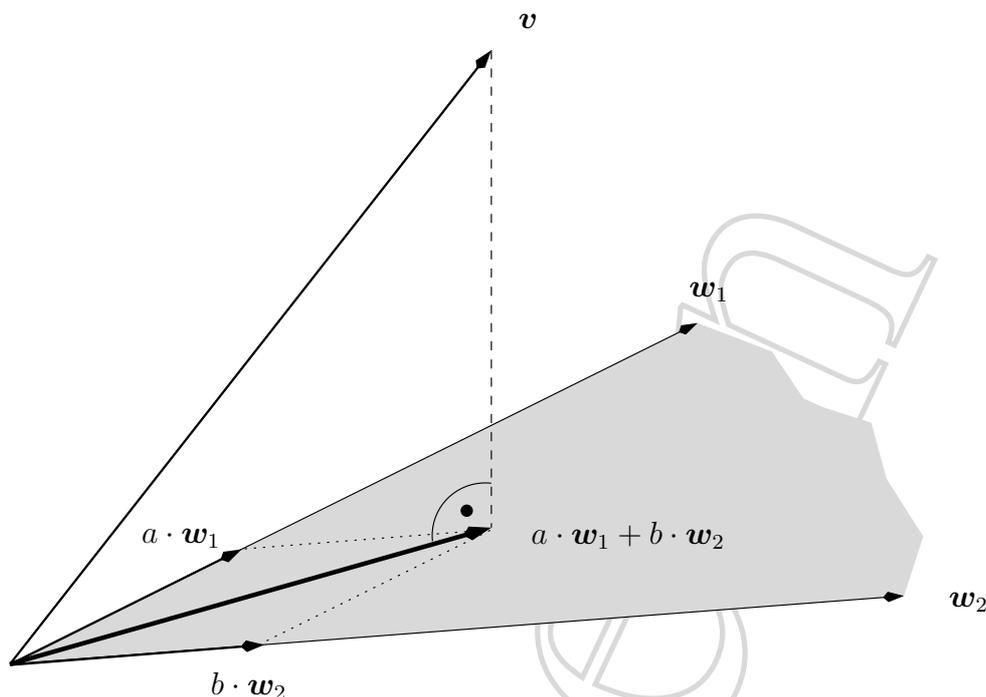


Abbildung 4.4: Orthogonalprojektion auf eine Ebene.

Dies ist ein Gleichungssystem mit zwei Gleichungen für die Unbekannten a und b . In Kurzschreibweise lautet es

$$\begin{pmatrix} \langle \mathbf{w}_1, \mathbf{w}_1 \rangle & \langle \mathbf{w}_1, \mathbf{w}_2 \rangle & | & \langle \mathbf{w}_1, \mathbf{v} \rangle \\ \langle \mathbf{w}_2, \mathbf{w}_1 \rangle & \langle \mathbf{w}_2, \mathbf{w}_2 \rangle & | & \langle \mathbf{w}_2, \mathbf{v} \rangle \end{pmatrix}. \quad (4.7)$$

Für das gegebene Beispiel berechnen wir die folgenden Matrixeinträge:

$$\begin{pmatrix} 14 & 0 & | & 2 \\ 0 & 3 & | & 1 \end{pmatrix}.$$

Auflösen ergibt die Lösung

$$a = \frac{1}{7}, \quad b = \frac{1}{3}.$$

Schliesslich erhalten wir die Gleichung der ‘‘Kompromissgeraden’’ g :

$$x_2 = \frac{1}{7}x_1 + \frac{1}{3}.$$

Sie ist so konstruiert worden, dass die Quadrate der vertikalen Abweichungen in (4.5) minimal sind.

Die hier diskutierte Technik heisst daher auch die **Methode der kleinsten Quadrate**. Das Verfahren lässt sich einfach verallgemeinern. Betrachten wir dazu ein überbestimmtes Gleichungs-

system mit m Gleichungen und n Unbekannten, wobei $m > n$:

$$\begin{array}{cccccc} a_{11}x_1 & + & a_{12}x_2 & + \cdots + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + \cdots + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \ddots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + \cdots + & a_{mn}x_n & = & b_m \end{array} \quad (4.8)$$

Gesucht sind die Unbekannten x_1, x_2, \dots, x_n , sodass das Gleichungssystem erfüllt ist. Da das System überbestimmt ist, wird es im Allgemeinen keine Lösung geben.

Wir suchen erneut einen Kompromiss mit Hilfe der Methode der kleinsten Quadrate. Dazu gehen wir gleich vor wie im obigen Beispiel und fassen zuerst die Spalten der erweiterten Matrix des Gleichungssystems als Vektoren auf:

$$\mathbf{w}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \quad \dots, \quad \mathbf{w}_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

und

$$\mathbf{v} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Dann können wir das Gleichungssystem schreiben als

$$x_1 \cdot \mathbf{w}_1 + x_2 \cdot \mathbf{w}_2 + \cdots + x_n \cdot \mathbf{w}_n = \mathbf{v}.$$

Eine bestmögliche Lösung durch kleinste Quadrate wird dann gefunden durch Minimierung der Gesamtabweichung

$$\mathcal{E} = \|(x_1 \cdot \mathbf{w}_1 + x_2 \cdot \mathbf{w}_2 + \cdots + x_n \cdot \mathbf{w}_n) - \mathbf{v}\|^2.$$

Die Kompromisslösung ist, genau wie zuvor, die Orthogonalprojektion $P_W(\mathbf{v})$ des Vektors \mathbf{v} auf den Unterraum

$$W = \text{span}(\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n).$$

Das entsprechende lineare Gleichungssystem für die Lösung x_1, x_2, \dots, x_n ist eine Verallgemeinerung von (4.7):

$$\begin{pmatrix} \langle \mathbf{w}_1, \mathbf{w}_1 \rangle & \langle \mathbf{w}_1, \mathbf{w}_2 \rangle & \langle \mathbf{w}_1, \mathbf{w}_3 \rangle & \cdots & \langle \mathbf{w}_1, \mathbf{w}_n \rangle & | & \langle \mathbf{w}_1, \mathbf{v} \rangle \\ \langle \mathbf{w}_2, \mathbf{w}_1 \rangle & \langle \mathbf{w}_2, \mathbf{w}_2 \rangle & \langle \mathbf{w}_2, \mathbf{w}_3 \rangle & \cdots & \langle \mathbf{w}_2, \mathbf{w}_n \rangle & | & \langle \mathbf{w}_2, \mathbf{v} \rangle \\ \langle \mathbf{w}_3, \mathbf{w}_1 \rangle & \langle \mathbf{w}_3, \mathbf{w}_2 \rangle & \langle \mathbf{w}_3, \mathbf{w}_3 \rangle & \cdots & \langle \mathbf{w}_3, \mathbf{w}_n \rangle & | & \langle \mathbf{w}_3, \mathbf{v} \rangle \\ \vdots & \vdots & \vdots & \ddots & \vdots & | & \vdots \\ \langle \mathbf{w}_n, \mathbf{w}_1 \rangle & \langle \mathbf{w}_n, \mathbf{w}_2 \rangle & \langle \mathbf{w}_n, \mathbf{w}_3 \rangle & \cdots & \langle \mathbf{w}_n, \mathbf{w}_n \rangle & | & \langle \mathbf{w}_n, \mathbf{v} \rangle \end{pmatrix}. \quad (4.9)$$

Diese Gleichungen nennt man die zum Problem (4.8) gehörigen **Normalgleichungen**.

Wir halten fest:

1. Die Koeffizientenmatrix des obigen Gleichungssystems (Normalgleichungen) wird wie folgt gebildet: der Eintrag in Zeile i und Spalte j ist das Skalarprodukt der Vektoren w_i und w_j . Da $\langle w_i, w_j \rangle = \langle w_j, w_i \rangle$, ist die Koeffizientenmatrix bezüglich ihrer Diagonalen *symmetrisch*.
2. Der i -te Eintrag im Rechte-Seite-Vektor in den Normalgleichungen berechnet sich als das Skalarprodukt der Vektoren w_i und v .

Anwendung 4.10 (Längenmessung)

Der Routenplaner von GOOGLE MAPS liefert folgende Distanzen (in km):

	Basel	Bern	Genf
Basel			
Bern	98		
Genf	251	157	
Zürich		126	278

Wir skizzieren die Lage der vier Städte in Abbildung 4.5 und führen die folgenden Distanzen ein:

$$\begin{aligned} d_1 &= \text{Bern-Basel} \\ d_2 &= \text{Bern-Genf} \\ d_3 &= \text{Bern-Zürich} \end{aligned}$$

Mit den Daten aus der obigen Tabelle muss gelten:

$$\begin{aligned} d_1 &= 98 \\ d_1 + d_2 &= 251 \\ d_2 &= 157 \\ d_3 &= 126 \\ d_2 + d_3 &= 278 \end{aligned}$$

Dies ist ein überbestimmtes lineares Gleichungssystem, welches keine Lösung hat (man sieht dies sofort durch Zusammenzählen der ersten und dritten Gleichung: $d_1 + d_2 = 98 + 157 = 255$; dies passt aber nicht mit der zweiten Gleichung zusammen). Um einen Kompromiss zu finden, benutzen wir die Methode der kleinsten Quadrate. Dazu schreiben wir das lineare Gleichungssystem in Matrixform,

$$\begin{pmatrix} 1 & 0 & 0 & | & 98 \\ 1 & 1 & 0 & | & 251 \\ 0 & 1 & 0 & | & 157 \\ 0 & 0 & 1 & | & 126 \\ 0 & 1 & 1 & | & 278 \end{pmatrix},$$

4.1 ORTHOGONALPROJEKTIONEN UND KLEINSTE QUADRATE

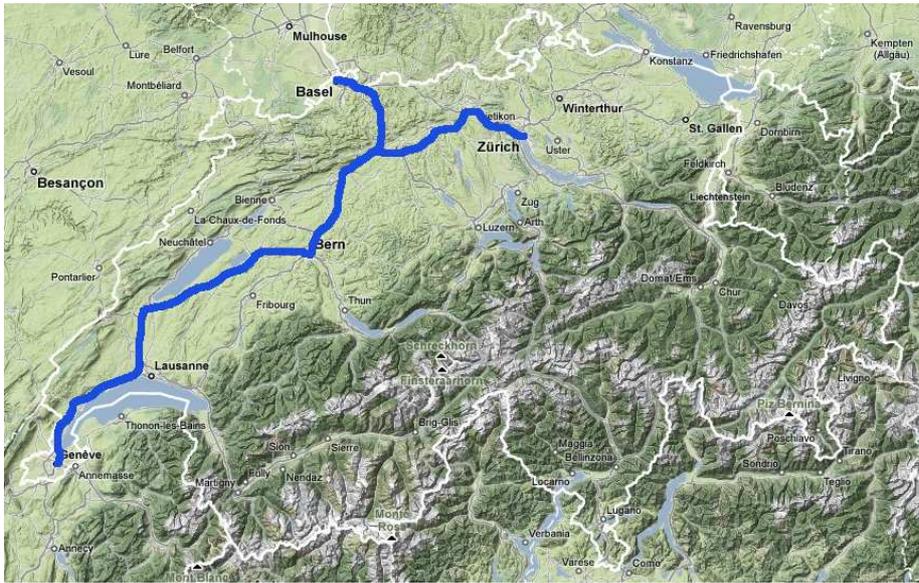


Abbildung 4.5: Verbindungen zwischen Basel, Bern, Genf und Zürich.

und schreiben die Spalten dieser Matrix heraus:

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{w}_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix},$$

und

$$\mathbf{v} = \begin{pmatrix} 98 \\ 251 \\ 157 \\ 126 \\ 278 \end{pmatrix}.$$

Nun stellen wir das Gleichungssystem (4.9) durch Berechnung der entsprechenden Skalarprodukte auf:

$$\begin{aligned} \langle \mathbf{w}_1, \mathbf{w}_1 \rangle &= 2 & \langle \mathbf{w}_1, \mathbf{v} \rangle &= 349 \\ \langle \mathbf{w}_1, \mathbf{w}_2 \rangle &= \langle \mathbf{w}_2, \mathbf{w}_1 \rangle = 1 & \langle \mathbf{w}_2, \mathbf{v} \rangle &= 686 \\ \langle \mathbf{w}_1, \mathbf{w}_3 \rangle &= \langle \mathbf{w}_3, \mathbf{w}_1 \rangle = 0 & \langle \mathbf{w}_3, \mathbf{v} \rangle &= 404 \\ \langle \mathbf{w}_2, \mathbf{w}_2 \rangle &= 3 \\ \langle \mathbf{w}_2, \mathbf{w}_3 \rangle &= \langle \mathbf{w}_3, \mathbf{w}_2 \rangle = 1 \\ \langle \mathbf{w}_3, \mathbf{w}_3 \rangle &= 2 \end{aligned}$$

Somit gilt es, das lineare Gleichungssystem

$$\left(\begin{array}{ccc|c} 2 & 1 & 0 & 349 \\ 1 & 3 & 1 & 686 \\ 0 & 1 & 2 & 404 \end{array} \right)$$

zu lösen. Die Lösung ist gegeben durch

$$d_1 = 97.125 \dots, \quad d_2 = 154.750 \dots, \quad d_3 = 124.625 \dots$$

◇

Bemerkung 4.11 In der Praxis wird die direkte Berechnung der Lösung des linearen Gleichungssystems (4.9) durch das Gaussverfahren aus Stabilitätsgründen (Rundungsfehler) oftmals vermieden. Alternativ werden beispielsweise sogenannte Matrixfaktorisierungsverfahren eingesetzt (Cholesky- und QR -Faktorisierungen), welche eine stabilere Lösung von (4.9) möglich machen.

4.2 Diskrete Fouriertransformation und Analyse von periodischen Daten

In der Klimaforschung interessiert man sich unter anderem für Temperaturmittel über gewisse Zeitintervalle an verschiedenen Orten. Man würde intuitiv erwarten, dass die Mittelwerte umso “stabiler” sind, je länger das Zeitfenster der Messungen ist. Allerdings muss man natürlich auch bemerken, dass sich das Klima über die Jahrzehnte und Jahrhunderte wandeln kann.

In der folgenden Tabelle betrachten wir die Temperatur am Jungfraujoch im Zeitfenster von 1961 bis 1990:

Monat	Jan	Feb	Mär	Apr	Mai	Jun	Jul	Aug	Sep	Okt	Nov	Dez
Temp. [$^{\circ}C$]	-13.6	-14.2	-13.1	-10.8	-6.6	-3.7	-1.2	-1.2	-2.6	-5.2	-10.4	-12.3

Wir stellen diese Daten grafisch dar; siehe Abbildung 4.6 (hier steht Januar für $t = 0$, Februar für $t = 1$, März für $t = 2$, etc.). Wenn wir davon ausgehen, dass sich diese Kurve jährlich wiederholt, so können wir die Temperatur T (gemessen in $^{\circ}C$) als eine *periodische Funktion* der Zeit t (gemessen in Monaten) auffassen. Wir machen daher den Ansatz, dass die Funktion $T(t)$ beispielsweise wie folgt aussehen könnte:

$$T(t) = \hat{a}_0 + \hat{a}_1 \cos\left(\frac{2\pi}{12}t\right) + \hat{b}_1 \sin\left(\frac{2\pi}{12}t\right). \quad (4.12)$$

Hier sind \hat{a}_0 , \hat{a}_1 und \hat{b}_1 unbekannte Koeffizienten. Da die Funktionen

$$\cos\left(\frac{2\pi}{12}t\right), \quad \sin\left(\frac{2\pi}{12}t\right)$$

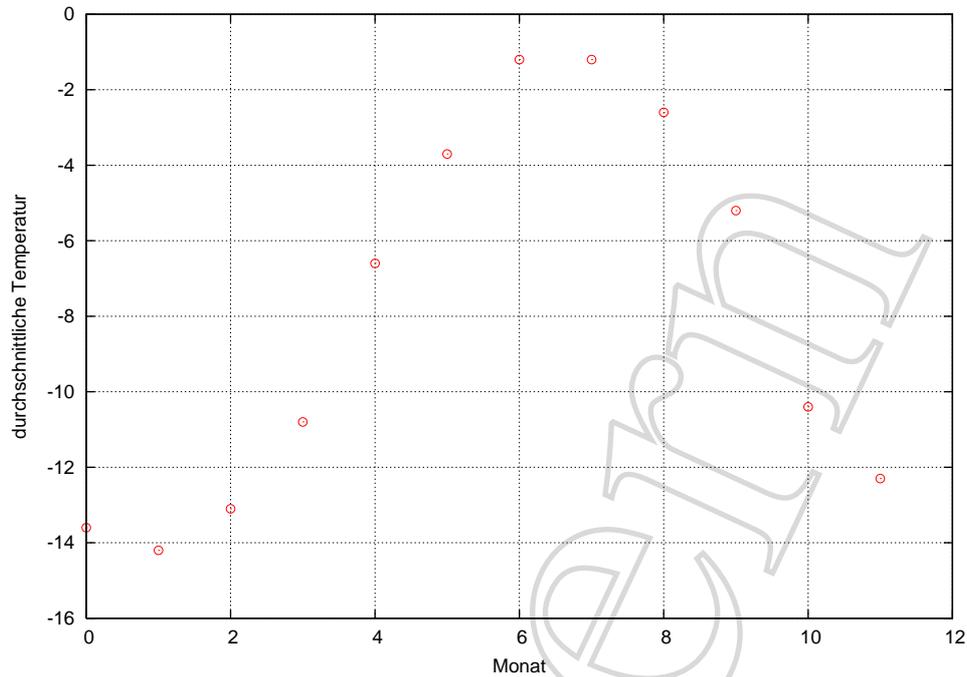


Abbildung 4.6: Monatliche Durchschnittstemperaturen am Jungfraujoch in den Jahren 1961 – 1990.

periodisch sind mit Periodenlänge 12, ist die angesetzte Funktion $T(t)$ “zusammengesetzt” aus Funktionen, welche dem erwarteten 12-monatigen Zyklus der Temperaturen genau entsprechen¹.

Unser Ziel ist es, die unbekanntenen Koeffizienten so zu bestimmen, dass die Daten möglichst gut mit der Funktion T in (4.12) übereinstimmen. Betrachten wir zunächst die Durchschnittstemperatur im Januar (d.h., $t = 0$), so muss gelten:

$$-13.6 = T(0) = \hat{a}_0 + \hat{a}_1 \cos\left(\frac{2\pi}{12} \cdot 0\right) + \hat{b}_1 \sin\left(\frac{2\pi}{12} \cdot 0\right).$$

Ähnlich erhalten wir dann für Februar ($t = 1$):

$$-14.2 = T(1) = \hat{a}_0 + \hat{a}_1 \cos\left(\frac{2\pi}{12}\right) + \hat{b}_1 \sin\left(\frac{2\pi}{12}\right).$$

Weiter für den Monat März ($t = 2$):

$$-13.1 = T(2) = \hat{a}_0 + \hat{a}_1 \cos\left(\frac{4\pi}{12}\right) + \hat{b}_1 \sin\left(\frac{4\pi}{12}\right).$$

¹In der Auswertung der Funktionen \cos und \sin verwenden wir hier die Winkelmessung in rad, d.h., $360 [\text{deg}] \hat{=} 2\pi [\text{rad}]$.

Für die restlichen 9 Monate ergeben sich dann die Gleichungen:

$$\begin{aligned}
 -10.8 &= T(3) = \hat{a}_0 + \hat{a}_1 \cos\left(\frac{6\pi}{12}\right) + \hat{b}_1 \sin\left(\frac{6\pi}{12}\right) \\
 -6.6 &= T(4) = \hat{a}_0 + \hat{a}_1 \cos\left(\frac{8\pi}{12}\right) + \hat{b}_1 \sin\left(\frac{8\pi}{12}\right) \\
 -3.7 &= T(5) = \hat{a}_0 + \hat{a}_1 \cos\left(\frac{10\pi}{12}\right) + \hat{b}_1 \sin\left(\frac{10\pi}{12}\right) \\
 -1.2 &= T(6) = \hat{a}_0 + \hat{a}_1 \cos\left(\frac{12\pi}{12}\right) + \hat{b}_1 \sin\left(\frac{12\pi}{12}\right) \\
 -1.2 &= T(7) = \hat{a}_0 + \hat{a}_1 \cos\left(\frac{14\pi}{12}\right) + \hat{b}_1 \sin\left(\frac{14\pi}{12}\right) \\
 -2.6 &= T(8) = \hat{a}_0 + \hat{a}_1 \cos\left(\frac{16\pi}{12}\right) + \hat{b}_1 \sin\left(\frac{16\pi}{12}\right) \\
 -5.2 &= T(9) = \hat{a}_0 + \hat{a}_1 \cos\left(\frac{18\pi}{12}\right) + \hat{b}_1 \sin\left(\frac{18\pi}{12}\right) \\
 -10.4 &= T(10) = \hat{a}_0 + \hat{a}_1 \cos\left(\frac{20\pi}{12}\right) + \hat{b}_1 \sin\left(\frac{20\pi}{12}\right) \\
 -12.3 &= T(11) = \hat{a}_0 + \hat{a}_1 \cos\left(\frac{22\pi}{12}\right) + \hat{b}_1 \sin\left(\frac{22\pi}{12}\right).
 \end{aligned}$$

Fassen wir die obigen Gleichungen zusammen, so ergibt sich ein überbestimmtes lineares Gleichungssystem mit 12 Gleichungen für die drei Unbekannten $\hat{a}_0, \hat{a}_1, \hat{b}_1$. In Matrixform:

$$\begin{pmatrix}
 1 & \cos(0) & \sin(0) & | & -13.6 \\
 1 & \cos(2\pi/12) & \sin(2\pi/12) & | & -14.2 \\
 1 & \cos(4\pi/12) & \sin(4\pi/12) & | & -13.1 \\
 1 & \cos(6\pi/12) & \sin(6\pi/12) & | & -10.8 \\
 1 & \cos(8\pi/12) & \sin(8\pi/12) & | & -6.6 \\
 1 & \cos(10\pi/12) & \sin(10\pi/12) & | & -3.7 \\
 1 & \cos(12\pi/12) & \sin(12\pi/12) & | & -1.2 \\
 1 & \cos(14\pi/12) & \sin(14\pi/12) & | & -1.2 \\
 1 & \cos(16\pi/12) & \sin(16\pi/12) & | & -2.6 \\
 1 & \cos(18\pi/12) & \sin(18\pi/12) & | & -5.2 \\
 1 & \cos(20\pi/12) & \sin(20\pi/12) & | & -10.4 \\
 1 & \cos(22\pi/12) & \sin(22\pi/12) & | & -12.3
 \end{pmatrix}$$

Bemerkenswert ist hier die Tatsache, dass die Vektoren \mathbf{c}_0 , \mathbf{c}_1 , \mathbf{s}_1 paarweise senkrecht aufeinander stehen. Insbesondere sind dann die Skalarprodukte jeweils gleich 0. Das lineare Gleichungssystem hat deshalb eine sehr einfache Form:

$$\left(\begin{array}{ccc|c} \langle \mathbf{c}_0, \mathbf{c}_0 \rangle & 0 & 0 & \langle \mathbf{c}_0, \mathbf{v} \rangle \\ 0 & \langle \mathbf{c}_1, \mathbf{c}_1 \rangle & 0 & \langle \mathbf{c}_1, \mathbf{v} \rangle \\ 0 & 0 & \langle \mathbf{s}_1, \mathbf{s}_1 \rangle & \langle \mathbf{s}_1, \mathbf{v} \rangle \end{array} \right).$$

Daraus lässt sich die Lösung direkt bestimmen:

$$\hat{a}_0 = \frac{\langle \mathbf{c}_0, \mathbf{v} \rangle}{\langle \mathbf{c}_0, \mathbf{c}_0 \rangle} = -7.91, \quad \hat{a}_1 = \frac{\langle \mathbf{c}_1, \mathbf{v} \rangle}{\langle \mathbf{c}_1, \mathbf{c}_1 \rangle} = -6.38, \quad \hat{b}_1 = \frac{\langle \mathbf{s}_1, \mathbf{v} \rangle}{\langle \mathbf{s}_1, \mathbf{s}_1 \rangle} = -2.27. \quad (4.14)$$

Diese Zahlen nennen wir **diskrete Fourierkoeffizienten**. Alle Koeffizienten zusammen heißen **diskrete Fouriertransformierte** der gegebenen Daten. Die Funktion T in (4.12) bezeichnen wir als **diskrete Fourierreihe**:

$$T(t) = -7.91 - 6.38 \cos\left(\frac{2\pi}{12}t\right) - 2.27 \sin\left(\frac{2\pi}{12}t\right).$$

Sie ist eine kontinuierliche Annäherung an die diskreten Daten. Zeichnen wir diese Funktion auf, so ergibt sich tatsächlich eine recht gute Approximation zu den gemessenen Daten; siehe Abbildung 4.7.

Verallgemeinerungen

Die Fourierreihe $T(t)$ lässt sich durch Hinzufügen weiterer Terme verbessern (d.h., die Approximation an die Daten wird genauer), indem weitere cos- und sin-Terme hinzugefügt werden:

$$\begin{aligned} T(t) = & \hat{a}_0 + \hat{a}_1 \cos\left(\frac{2\pi}{12}t\right) + \hat{a}_2 \cos\left(\frac{4\pi}{12}t\right) + \hat{a}_3 \cos\left(\frac{6\pi}{12}t\right) + \cdots + \hat{a}_m \cos\left(\frac{2m\pi}{12}t\right) \\ & + \hat{b}_1 \sin\left(\frac{2\pi}{12}t\right) + \hat{b}_2 \sin\left(\frac{4\pi}{12}t\right) + \hat{b}_3 \sin\left(\frac{6\pi}{12}t\right) + \cdots + \hat{b}_m \sin\left(\frac{2m\pi}{12}t\right), \end{aligned}$$

für eine gewisse Zahl m . Die kleinste Quadrate Methode führt dann, wie in (4.14), auf die diskreten Fourierkoeffizienten

$$\begin{aligned} \hat{a}_0 &= \frac{\langle \mathbf{c}_0, \mathbf{v} \rangle}{\langle \mathbf{c}_0, \mathbf{c}_0 \rangle}, & \hat{b}_1 &= \frac{\langle \mathbf{s}_1, \mathbf{v} \rangle}{\langle \mathbf{s}_1, \mathbf{s}_1 \rangle}, \\ \hat{a}_1 &= \frac{\langle \mathbf{c}_1, \mathbf{v} \rangle}{\langle \mathbf{c}_1, \mathbf{c}_1 \rangle}, & & \\ & \vdots & & \vdots \\ \hat{a}_m &= \frac{\langle \mathbf{c}_m, \mathbf{v} \rangle}{\langle \mathbf{c}_m, \mathbf{c}_m \rangle}, & \hat{b}_m &= \frac{\langle \mathbf{s}_m, \mathbf{v} \rangle}{\langle \mathbf{s}_m, \mathbf{s}_m \rangle}, \end{aligned} \quad (4.15)$$

4.2 DISKRETE FOURIERTRANSFORMATION UND ANALYSE VON PERIODISCHEN DATEN

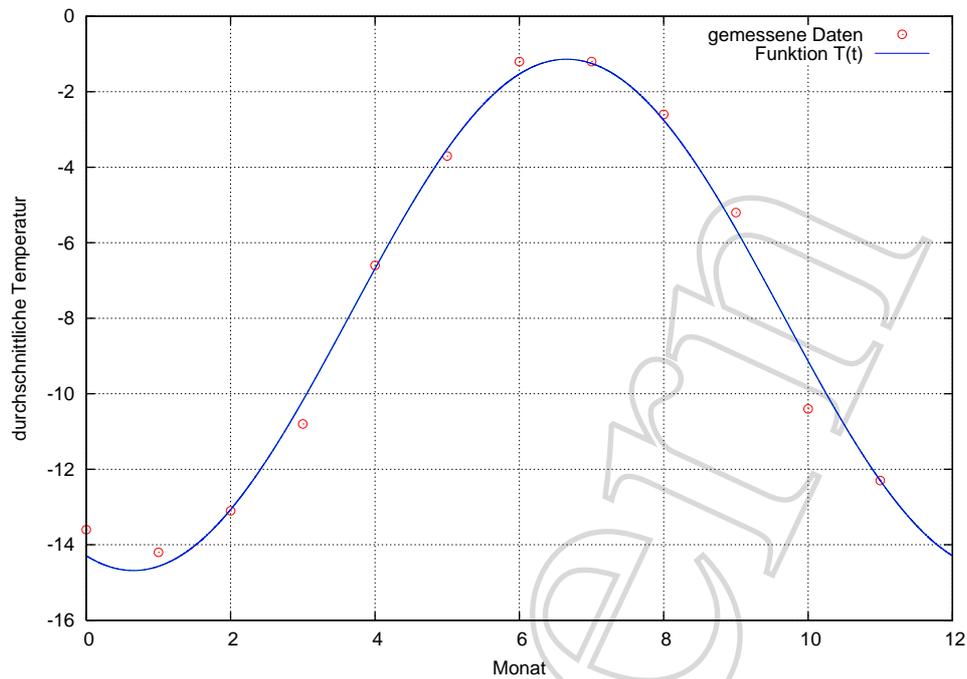


Abbildung 4.7: Monatliche Durchschnittstemperatur am Jungfrauoch (von 1961 bis 1990): Vergleich der Daten mit der Näherungsfunktion $T(t)$.

wobei c_0 wie in (4.13) ist, und

$$\mathbf{c}_1 = \begin{pmatrix} \cos(0) \\ \cos(2\pi/12) \\ \cos(4\pi/12) \\ \cos(6\pi/12) \\ \cos(8\pi/12) \\ \cos(10\pi/12) \\ \cos(12\pi/12) \\ \cos(14\pi/12) \\ \cos(16\pi/12) \\ \cos(18\pi/12) \\ \cos(20\pi/12) \\ \cos(22\pi/12) \end{pmatrix}, \quad \mathbf{c}_2 = \begin{pmatrix} \cos(0) \\ \cos(2 \cdot 2\pi/12) \\ \cos(2 \cdot 4\pi/12) \\ \cos(2 \cdot 6\pi/12) \\ \cos(2 \cdot 8\pi/12) \\ \cos(2 \cdot 10\pi/12) \\ \cos(2 \cdot 12\pi/12) \\ \cos(2 \cdot 14\pi/12) \\ \cos(2 \cdot 16\pi/12) \\ \cos(2 \cdot 18\pi/12) \\ \cos(2 \cdot 20\pi/12) \\ \cos(2 \cdot 22\pi/12) \end{pmatrix}, \quad \dots, \quad \mathbf{c}_m = \begin{pmatrix} \cos(0) \\ \cos(m \cdot 2\pi/12) \\ \cos(m \cdot 4\pi/12) \\ \cos(m \cdot 6\pi/12) \\ \cos(m \cdot 8\pi/12) \\ \cos(m \cdot 10\pi/12) \\ \cos(m \cdot 12\pi/12) \\ \cos(m \cdot 14\pi/12) \\ \cos(m \cdot 16\pi/12) \\ \cos(m \cdot 18\pi/12) \\ \cos(m \cdot 20\pi/12) \\ \cos(m \cdot 22\pi/12) \end{pmatrix},$$

sowie

$$\mathbf{s}_1 = \begin{pmatrix} \sin(0) \\ \sin(2\pi/12) \\ \sin(4\pi/12) \\ \sin(6\pi/12) \\ \sin(8\pi/12) \\ \sin(10\pi/12) \\ \sin(12\pi/12) \\ \sin(14\pi/12) \\ \sin(16\pi/12) \\ \sin(18\pi/12) \\ \sin(20\pi/12) \\ \sin(22\pi/12) \end{pmatrix}, \quad \mathbf{s}_2 = \begin{pmatrix} \sin(0) \\ \sin(2 \cdot 2\pi/12) \\ \sin(2 \cdot 4\pi/12) \\ \sin(2 \cdot 6\pi/12) \\ \sin(2 \cdot 8\pi/12) \\ \sin(2 \cdot 10\pi/12) \\ \sin(2 \cdot 12\pi/12) \\ \sin(2 \cdot 14\pi/12) \\ \sin(2 \cdot 16\pi/12) \\ \sin(2 \cdot 18\pi/12) \\ \sin(2 \cdot 20\pi/12) \\ \sin(2 \cdot 22\pi/12) \end{pmatrix}, \quad \dots, \quad \mathbf{s}_m = \begin{pmatrix} \sin(0) \\ \sin(m \cdot 2\pi/12) \\ \sin(m \cdot 4\pi/12) \\ \sin(m \cdot 6\pi/12) \\ \sin(m \cdot 8\pi/12) \\ \sin(m \cdot 10\pi/12) \\ \sin(m \cdot 12\pi/12) \\ \sin(m \cdot 14\pi/12) \\ \sin(m \cdot 16\pi/12) \\ \sin(m \cdot 18\pi/12) \\ \sin(m \cdot 20\pi/12) \\ \sin(m \cdot 22\pi/12) \end{pmatrix}.$$

Auch hier stehen alle Vektoren paarweise senkrecht aufeinander, wodurch die Koeffizientenmatrix in (4.9) “diagonal” wird und die explizite Lösung (4.15) möglich macht.

Verschiedene Anwendungen bringen verschiedene Periodenlängen mit sich. Insbesondere müssen diese natürlich nicht immer gleich 12 sein. Für eine allgemeine Periodenlänge L betrachten wir dann Funktionen der Form

$$f(t) = \hat{a}_0 + \hat{a}_1 \cos\left(\frac{2\pi}{L}t\right) + \hat{a}_2 \cos\left(\frac{4\pi}{L}t\right) + \dots + \hat{a}_m \cos\left(\frac{2m\pi}{L}t\right) \\ + \hat{b}_1 \sin\left(\frac{2\pi}{L}t\right) + \hat{b}_2 \sin\left(\frac{4\pi}{L}t\right) + \dots + \hat{b}_m \sin\left(\frac{2m\pi}{L}t\right)$$

zur Annäherung an periodische Daten. Hier gehen wir stets davon aus, dass das “Abtasten” der Daten, wie im vorherigen Beispiel, in *gleichmässigen* Abständen geschieht (also z.B. einmal pro Monat wie in der obigen Temperaturtabelle):

Zeitpunkte	0	1	2	...	$L-1$
Datenwerte	w_0	w_1	w_2	...	w_{L-1}

In dem Fall sind die entsprechenden Vektoren $\mathbf{c}_0, \mathbf{c}_1, \dots, \mathbf{c}_m$ und $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m$ wieder paarweise senkrecht zueinander. Die Lösungsformeln in (4.15) gelten dann nach wie vor.

In der Praxis gibt es sehr effiziente Methoden (beispielsweise die “Fast Fourier Transform (FFT)”), welche eine rasche Berechnung der diskreten Fourierkoeffizienten erlaubt.

4.3 Übungsaufgaben

- 4.1. Gemäss dem Hooke'schen Gesetz ist die Auslenkung einer Spiralfeder proportional zur angesetzten Kraft, d.h.,

$$F(s) = D \cdot s.$$

Dabei sind F [N] die angesetzte Kraft, s [m] die Auslenkung der Feder (beide positiv oder negativ) und D [$\frac{\text{N}}{\text{m}}$] die Federkonstante (welche die „Härte“ einer Feder angibt). Letztere soll für eine bestimmte Feder durch Messung von F in Abhängigkeit von s bestimmt werden; es ergibt sich

s	-0.3	-0.2	0.1	0.5
$F(s)$	-288	-194	98	479

Bestimmen Sie die Federkonstante D mit der Methode der kleinsten Quadrate.

- 4.2. In einem Stromkreis soll ein Widerstand mit Hilfe des Ohm'schen Gesetzes bestimmt werden. In einem Experiment wird für verschiedene Stromstärken der entsprechende Spannungsverlust gemessen:

Strom I	0.1	0.2	0.3	0.4	0.5
Spannung V	1.0	2.1	2.9	4.2	5.1

Stellen Sie ein überbestimmtes Gleichungssystem auf und lösen Sie es mit der Methode der kleinsten Quadrate.

- 4.3. Am 5. Oktober 1991 tauchte der Magellan-Orbiter in die Atmosphäre der Venus ein und übermittelte die Temperatur T (in Kelvin K) in Abhängigkeit vom Abstand h zur Planetenoberfläche. Das Signal brach auf der Höhe $h = 34$ km ab. In der Nähe der Oberfläche zeigen die übermittelten Daten einen Abfall der Temperatur in Abhängigkeit zur Höhe der (ungefähren) Form

$$T(h) = a \cdot h + b$$

für unbekannte Koeffizienten a und b . Verwenden Sie die folgende Datentabelle, um a und b mittels der Methode der kleinsten Quadrate zu bestimmen und berechnen Sie daraus die Temperatur an der Venusoberfläche.

Hohe h	60	55	50	45	40	35
Temperatur T (in K)	263	300	348	383	416	453

Bestimmen Sie mit der Methode der kleinsten Quadrate die Koeffizienten a , b und c so, dass die Ebene

$$E : ax_1 + bx_2 + cx_3 = 1$$

möglichst nahe „durch“ die Punkte $P_1(1, 0, -1)$, $P_2(2, 1, -2)$, $P_3(1, 1, 0)$ und $P_4(1, 1, -1)$ verläuft.

- 4.4. Bekanntlich gilt für den freien Fall eines Körpers in der Nähe der Erdoberfläche das Fallgesetz von Newton:

$$s(t) = s_0 + v_0 t + \frac{1}{2} g t^2,$$

wobei

- s = Vertikaldistanz gegenüber (willkürlichem) Nullpunkt P_0 ,
- s_0 = Vertikaldistanz zu P_0 zum Zeitpunkt $t = 0$,
- v_0 = Anfangsgeschwindigkeit (zum Zeitpunkt $t = 0$),
- g = Gravitationsbeschleunigung.

Durch Messung von $s(t)$ zu verschiedenen Zeiten t beim Fall eines Körpers soll g bestimmt werden. Es wird gemessen:

t [s]	0.2	0.3	0.4	0.5
$s(t)$ [cm]	11.7	42.1	82.4	132.7

Bestimmen Sie g mit der Methode der kleinsten Quadrate.

- 4.5. Die folgende Tabelle enthält die Monats-Mittelwerte der in Zürich in der 30-Jahre-Periode 1961–1990 gemessenen Temperaturen.

Monat	J	F	M	A	M	J	J	A	S	O	N	D
T [°C]	-0.5	0.9	4.2	7.9	12.2	15.4	17.7	16.8	13.9	9.2	3.9	0.6

- a) Bestimmen Sie die Koeffizienten \hat{a}_0 , \hat{a}_1 und \hat{b}_1 der zu $T(t)$ gehörigen diskrete Fourierreihe

$$T(t) = \hat{a}_0 + \hat{a}_1 \cos\left(\frac{2\pi}{12}t\right) + \hat{b}_1 \sin\left(\frac{2\pi}{12}t\right).$$

- b) Gleiche Aufgabe wie a) für die diskrete Fourierreihe 3. Ordnung

$$T(t) = \hat{a}_0 + \hat{a}_1 \cos\left(\frac{2\pi}{12}t\right) + \hat{a}_2 \cos\left(\frac{4\pi}{12}t\right) + \hat{a}_3 \cos\left(\frac{6\pi}{12}t\right) \\ + \hat{b}_1 \sin\left(\frac{2\pi}{12}t\right) + \hat{b}_2 \sin\left(\frac{4\pi}{12}t\right) + \hat{b}_3 \sin\left(\frac{6\pi}{12}t\right).$$

- 4.6. Von einer zunächst unbekanntem Funktion $f(t)$ sind folgende Paare von Argument und Funktionswert bekannt.

t	0	1	2	3	4	5
$f(t)$	0	$3\sqrt{3}/8$	$3\sqrt{3}/8$	0	$-3\sqrt{3}/8$	$-3\sqrt{3}/8$

Bestimmen Sie die diskrete Fourierreihe 3. Ordnung für $f(t)$, wie in Aufgabe 4.5 b) und interpretieren Sie das Resultat.

- 4.7. Der Brunstzyklus von Säugetieren wird hormonell gesteuert; beim Hausrind dauert ein Zyklus etwa 21 Tage. Die folgende Tabelle enthält die über 103 Rinder gemittelten Progesteronwerte $P(t)$ (in ng pro ml Blutplasma) in Abhängigkeit der Zeit t seit Beginn der Brunst.

t [d]	0	3	6	9	12	15	18
$P(t)$ [$\frac{\text{ng}}{\text{ml}}$]	0.17	0.94	4.08	6.35	7.31	7.92	1.34

Bestimmen Sie die diskrete Fourierreihe zweiter Ordnung,

$$\hat{P}(t) = \hat{a}_0 + \hat{a}_1 \cos(? \cdot t) + \hat{b}_1 \sin(? \cdot t) + \hat{a}_2 \cos(? \cdot 2t) + \hat{b}_2 \sin(? \cdot 2t),$$

welche die gemessenen Progesteronspiegel am besten approximiert (im Sinne der Methode der kleinsten Quadrate).

U
ni
B
erlin

5

Eigenwertprobleme

5.1 Iterative Prozesse

Wir leiten dieses Kapitel mit einem Anwendungsbeispiel ein.

Anwendung 5.1 (Markov-Prozesse)

Wer den Geysir Old Faithful (Abbildung 5.1) im Yellowstone Nationalpark beobachtet, wird feststellen, dass er seine Fontänen in unregelmässigen Abständen hochsteigen lässt. Wer aber die Wartezeiten zwischen zwei Ausbrüchen grob in kurz und lang einteilt und damit etwas Statistik treibt, wird bemerken, dass auf ein kurzes Intervall immer ein langes folgt, nicht aber umgekehrt. Genauer gilt:

- Die Chance (Wahrscheinlichkeit), dass auf eine kurze Wartezeit eine lange folgt, ist 100%.



Abbildung 5.1: Geysir Old Faithful.

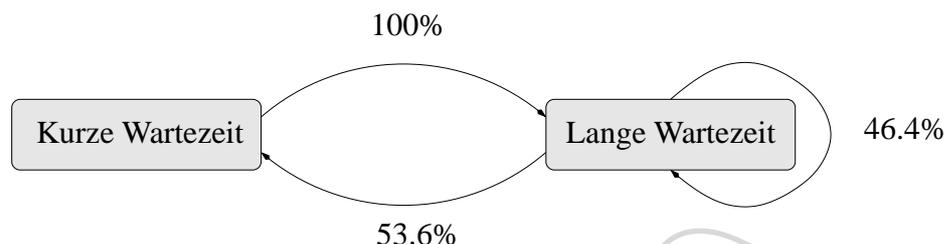


Abbildung 5.2: Gerichteter Graph für Geysir Old Faithful.

- Die Chance, dass auf eine lange Wartezeit eine kurze folgt, ist 53.6%, während sich mit einer Chance von 46.4% wiederholt ein langes Intervall anschliesst.

Diese Verhaltensweise lässt sich in Form eines *gerichteten Graphen* darstellen; siehe Abbildung 5.2. Hier bedeuten die rechteckigen Kästchen die möglichen Ereignisse, und die Pfeile stehen für die Übergänge zwischen diesen. Wir wollen die langzeitige Entwicklung der Wartezeiten studieren. Zu diesem Zweck bezeichnen wir die Chance für eine kurze respektive lange Wartezeit mit k und l . Soeben wurde ein langer Abstand zwischen zwei Ausbrüchen beobachtet. Dies entspricht den Wahrscheinlichkeiten

$$k_0 = 0\% = 0, \quad l_0 = 100\% = 1,$$

wobei wir den Index "0" benutzen, um den Startpunkt der Beobachtungen zu markieren. Die darauffolgenden Wartezeiten nummerieren wir dementsprechend mit 1, 2, 3, ...

Die Chance, dass als nächstes wieder eine lange Wartepause erfolgt, beträgt

$$l_1 = 46.4\% = 0.464.$$

Die Wahrscheinlichkeit für eine kurze Periode, beträgt

$$k_1 = 53.6\% = 0.536.$$

Wie gross sind nun die Chancen für eine kurze, respektive lange Wartezeit? Eine lange Wartezeit folgt einer kurzen mit Wahrscheinlichkeit 100% und einer langen mit Wahrscheinlichkeit 46.4%. Daher,

$$l_2 = 100\% \cdot k_1 + 46.4\% \cdot l_1 = k_1 + 0.464l_1 = 0.751296$$

Eine kurze Wartezeit kann nur nach einer vorherigen langen Periode auftreten und zwar mit einer Chance von 53.6%, d.h.,

$$k_2 = 0\% \cdot k_1 + 53.6\% \cdot l_1 = 0.536l_1 = 0.248704.$$

Genau gleich können wir nun weiterfahren, um die entsprechenden Wahrscheinlichkeiten für die dritte Wartezeit zu ermitteln:

$$l_3 = 100\% \cdot k_2 + 46.4\% \cdot l_2 = k_2 + 0.464l_2 = 0.597305344$$

$$k_3 = 0\% \cdot k_2 + 53.6\% \cdot l_2 = 0.536l_2 = 0.402694656.$$

n	0	1	2	3	4
\mathbf{v}_n	$\begin{pmatrix} 0.00000 \\ 1.00000 \end{pmatrix}$	$\begin{pmatrix} 0.53600 \\ 0.46400 \end{pmatrix}$	$\begin{pmatrix} 0.24870 \\ 0.75130 \end{pmatrix}$	$\begin{pmatrix} 0.40269 \\ 0.59731 \end{pmatrix}$	$\begin{pmatrix} 0.32016 \\ 0.67984 \end{pmatrix}$
n	5	6	7	8	9
\mathbf{v}_n	$\begin{pmatrix} 0.36440 \\ 0.63560 \end{pmatrix}$	$\begin{pmatrix} 0.34068 \\ 0.65932 \end{pmatrix}$	$\begin{pmatrix} 0.35339 \\ 0.64661 \end{pmatrix}$	$\begin{pmatrix} 0.34658 \\ 0.65342 \end{pmatrix}$	$\begin{pmatrix} 0.35023 \\ 0.64977 \end{pmatrix}$
n	10	11	12	13	14
\mathbf{v}_n	$\begin{pmatrix} 0.34828 \\ 0.65172 \end{pmatrix}$	$\begin{pmatrix} 0.34932 \\ 0.65068 \end{pmatrix}$	$\begin{pmatrix} 0.34876 \\ 0.65124 \end{pmatrix}$	$\begin{pmatrix} 0.34906 \\ 0.65094 \end{pmatrix}$	$\begin{pmatrix} 0.34890 \\ 0.65110 \end{pmatrix}$

Tabelle 5.1: Entwicklung der Iteration (5.2).

Allgemein erhalten wir für die Wahrscheinlichkeiten der Wartezeiten nach n Ausbrüchen:

$$\begin{aligned} l_{n+1} &= k_n + 0.464l_n \\ k_{n+1} &= 0 \cdot k_n + 0.536l_n. \end{aligned}$$

Speichern wir die Wahrscheinlichkeiten als Komponenten von Vektoren,

$$\mathbf{v}_0 = \begin{pmatrix} k_0 \\ l_0 \end{pmatrix}, \quad \mathbf{v}_1 = \begin{pmatrix} k_1 \\ l_1 \end{pmatrix}, \quad \dots, \quad \mathbf{v}_n = \begin{pmatrix} k_n \\ l_n \end{pmatrix},$$

so gilt mit der Matrix-Vektormultiplikation

$$\begin{pmatrix} k_{n+1} \\ l_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 0.536 \\ 1 & 0.464 \end{pmatrix} \cdot \begin{pmatrix} k_n \\ l_n \end{pmatrix},$$

oder äquivalent

$$\mathbf{v}_{n+1} = \mathbf{S} \cdot \mathbf{v}_n, \quad n = 0, 1, 2, \dots \quad (5.2)$$

wobei

$$\mathbf{S} = \begin{pmatrix} 0 & 0.536 \\ 1 & 0.464 \end{pmatrix}.$$

Ein Vorgang der Form (5.2), bei dem eine Grösse \mathbf{v}_{n+1} aus einer zuvor berechneten Grösse \mathbf{v}_n bestimmt wird, heisst **iterativer Prozess**. Im gegebenen Beispiel geschieht dies durch eine Matrix-Vektormultiplikation; wir nennen (5.2) deshalb auch **Matrix-Vektor-Iteration**, und die Matrix \mathbf{S} heisst **Iterations- oder Systemmatrix**. Falls, wie in dieser Anwendung, ein neuer Zustand aus einem vorherigen mit vorgegebenen Wahrscheinlichkeiten hervorgeht, heisst der iterative Prozess (**diskreter**) **Markovprozess oder Markovkette**.

Wie verhält sich der iterative Prozess (5.2) langfristig? Wir stellen dies in Tabelle 5.1 dar.

Offensichtlich strebt die Iteration gegen einen festen Vektor

$$\hat{\mathbf{v}} = \begin{pmatrix} 0.34\dots \\ 0.65\dots \end{pmatrix}.$$

Für fortgeschrittene Iterationszahlen n gilt also näherungsweise

$$\mathbf{v}_n \approx \hat{\mathbf{v}},$$

und diese Näherung wird umso besser, je grösser n ist. Erinnern wir uns an die Iterationsvorschrift (5.2), so bedeutet dies

$$\hat{\mathbf{v}} \approx \mathbf{v}_{n+1} = \mathbf{S} \cdot \mathbf{v}_n \approx \mathbf{S} \cdot \hat{\mathbf{v}}.$$

Im Grenzfall $n \rightarrow \infty$ wird die Approximation beliebig genau, und wir ersetzen " \approx " durch " $=$ ":

$$\mathbf{S} \cdot \hat{\mathbf{v}} = \hat{\mathbf{v}}.$$

Dies ist ein lineares Gleichungssystem für den Vektor $\hat{\mathbf{v}} = \begin{pmatrix} \hat{k} \\ \hat{l} \end{pmatrix}$, denn

$$\mathbf{S} \cdot \hat{\mathbf{v}} = \begin{pmatrix} 0.536\hat{l} \\ \hat{k} + 0.464\hat{l} \end{pmatrix},$$

und daher,

$$\mathbf{S} \cdot \hat{\mathbf{v}} = \begin{pmatrix} 0.536\hat{l} \\ \hat{k} + 0.464\hat{l} \end{pmatrix} = \begin{pmatrix} \hat{k} \\ \hat{l} \end{pmatrix} = \hat{\mathbf{v}}, \quad (5.3)$$

oder

$$\begin{aligned} -\hat{k} + 0.536\hat{l} &= 0 \\ \hat{k} - 0.536\hat{l} &= 0 \end{aligned}$$

Dieses homogene lineare Gleichungssystem hat die Lösungsmenge

$$\mathbb{L} = \left\{ \begin{pmatrix} 0.536s \\ s \end{pmatrix} : s \text{ eine beliebige Zahl} \right\}.$$

Nun erinnern wir uns, dass die Komponenten der Vektoren $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2, \dots$ die Bedeutung von Wahrscheinlichkeiten haben, die sich jeweils zu $100\% = 1$ addieren müssen. Dies muss auch für den Vektor $\hat{\mathbf{v}}$ gelten. Wir suchen also denjenigen Vektor in der obigen Lösungsmenge, für den gilt:

$$0.536s + s = 1,$$

d.h.,

$$s = \frac{1}{1.536} = 0.65104\dots$$

Somit erhalten wir

$$\hat{\mathbf{v}} = \begin{pmatrix} 0.536 \cdot 0.65104\dots \\ 0.65104\dots \end{pmatrix} = \begin{pmatrix} 0.34896\dots \\ 0.65104\dots \end{pmatrix}.$$

Dies ist der Vektor, entlang dessen sich die Matrix-Vektoriteration (5.2) stabilisiert. \diamond

n	0	1	2	3	4	5	6	7
\mathbf{v}_n	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$\begin{pmatrix} -1 \\ 4 \end{pmatrix}$	$\begin{pmatrix} -5 \\ 14 \end{pmatrix}$	$\begin{pmatrix} -19 \\ 46 \end{pmatrix}$	$\begin{pmatrix} -65 \\ 146 \end{pmatrix}$	$\begin{pmatrix} -211 \\ 454 \end{pmatrix}$	$\begin{pmatrix} -665 \\ 1394 \end{pmatrix}$	$\begin{pmatrix} -2059 \\ 4246 \end{pmatrix}$
$\ \mathbf{v}_n\ $	1.00	4.12	14.87	49.77	159.82	500.64	1544.49	
$\frac{1}{\ \mathbf{v}_n\ } \cdot \mathbf{v}_n$	$\begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$\begin{pmatrix} -0.24 \\ 0.97 \end{pmatrix}$	$\begin{pmatrix} -0.34 \\ 0.94 \end{pmatrix}$	$\begin{pmatrix} -0.38 \\ 0.92 \end{pmatrix}$	$\begin{pmatrix} -0.41 \\ 0.91 \end{pmatrix}$	$\begin{pmatrix} -0.42 \\ 0.91 \end{pmatrix}$	$\begin{pmatrix} -0.42 \\ 0.91 \end{pmatrix}$	$\begin{pmatrix} -0.43 \\ 0.90 \end{pmatrix}$

Tabelle 5.2: Entwicklung der Iteration (5.5) – (5.6).

Wir betrachten ein weiteres Beispiel.

Beispiel 5.4 Sei

$$\mathbf{S} = \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}, \quad (5.5)$$

und betrachte einen entsprechenden iterativen Prozess

$$\mathbf{v}_{n+1} = \mathbf{S} \cdot \mathbf{v}_n, \quad n = 0, 1, 2, \dots \quad (5.6)$$

Als Startvektor wählen wir $\mathbf{v}_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Auch hier studieren wir die langfristige Entwicklung der Iteration; siehe Tabelle 5.2 (oben). Auf den ersten Blick fällt folgendes auf:

1. Die erste Komponente der Vektorfolge wird immer kleiner (d.h., negativer), während die zweite immer grösser wird.
2. Das Verhältnis zwischen der ersten und zweiten Komponente entspricht mit wachsender Iterationszahl n grob dem Verhältnis $-1 : 2$.

Um diese Beobachtungen genauer zu untersuchen, tragen wir in Tabelle 5.2 (unten) sowohl die Längen der Vektoren \mathbf{v}_n als auch die entsprechenden auf Länge 1 skalierten Vektoren $\frac{1}{\|\mathbf{v}_n\|} \cdot \mathbf{v}_n$ ein (jeweils auf 2 Nachkommastellen gerundet). Wir erkennen, dass sich die Vektoren immer mehr einer stabilen Richtung

$$\hat{\mathbf{v}} \approx \begin{pmatrix} -0.43 \\ 0.90 \end{pmatrix}$$

angleichen. Ferner nehmen die Längen der Vektoren in jedem Schritt etwa um einen Faktor 3 zu. Wir formulieren diese Beobachtungen mathematisch:

1. Für grosse Iterationszahl gilt

$$\frac{1}{\|\mathbf{v}_n\|} \cdot \mathbf{v}_n \approx \hat{\mathbf{v}}, \quad (5.7)$$

wobei diese Approximation immer genauer und im Grenzfall $n \rightarrow \infty$ zur Gleichheit wird.

2. Das Längenverhältnis zweier aufeinanderfolgender Vektoren strebt immer mehr gegen 3:

$$\frac{\|\mathbf{v}_{n+1}\|}{\|\mathbf{v}_n\|} \approx 3. \quad (5.8)$$

Teilen wir die Gleichung (5.5) auf beiden Seiten durch $\|\mathbf{v}_n\|$, so folgt:

$$\frac{1}{\|\mathbf{v}_n\|} \cdot \mathbf{v}_{n+1} = \frac{1}{\|\mathbf{v}_n\|} \cdot (\mathbf{S} \cdot \mathbf{v}_n) = \mathbf{S} \cdot \left(\frac{1}{\|\mathbf{v}_n\|} \cdot \mathbf{v}_n \right).$$

Ausserdem gilt mit (5.8), dass

$$\frac{1}{\|\mathbf{v}_n\|} \cdot \mathbf{v}_{n+1} \approx \frac{3}{\|\mathbf{v}_{n+1}\|} \cdot \mathbf{v}_{n+1},$$

und deshalb

$$\mathbf{S} \cdot \left(\frac{1}{\|\mathbf{v}_n\|} \cdot \mathbf{v}_n \right) \approx 3 \cdot \left(\frac{1}{\|\mathbf{v}_{n+1}\|} \cdot \mathbf{v}_{n+1} \right).$$

Unter Benutzung von (5.7) wissen wir, dass

$$\hat{\mathbf{v}} \approx \frac{1}{\|\mathbf{v}_n\|} \cdot \mathbf{v}_n \approx \frac{1}{\|\mathbf{v}_{n+1}\|} \cdot \mathbf{v}_{n+1}.$$

Daher erhalten wir

$$\mathbf{S} \cdot \hat{\mathbf{v}} \approx 3 \cdot \hat{\mathbf{v}}.$$

Im Grenzfalle $n \rightarrow \infty$ dürfen wir davon ausgehen, dass die obige Näherung sogar zur Gleichheit wird, d.h., wir haben

$$\mathbf{S} \cdot \hat{\mathbf{v}} = 3 \cdot \hat{\mathbf{v}}. \quad (5.9)$$

Wie in (5.3) liegt auch hier ein lineares Gleichungssystem vor. Falls es eine Lösung $\hat{\mathbf{v}} \neq \mathbf{0}$ gibt, dann heisst ein solcher Vektor **Eigenvektor** von \mathbf{S} und die Zahl 3 der zugehörige **Eigenwert**. In der Tat gibt es auch hier eine nichttriviale (unendliche) Lösungsmenge

$$\mathbb{L} = \left\{ \begin{pmatrix} s \\ -2s \end{pmatrix} : s \text{ eine beliebige Zahl} \right\}.$$

Der auf Länge 1 skalierte Vektor $\hat{\mathbf{v}}$ muss dann

$$s^2 + (-2s)^2 = 1$$

erfüllen, d.h.,

$$s = \pm \frac{1}{\sqrt{5}}.$$

Er ist somit gegeben durch

$$\hat{\mathbf{v}} = \frac{1}{\sqrt{5}} \cdot \begin{pmatrix} 1 \\ -2 \end{pmatrix} \approx \begin{pmatrix} 0.44721 \\ -0.89443 \end{pmatrix} \quad \text{oder} \quad \hat{\mathbf{v}} = -\frac{1}{\sqrt{5}} \cdot \begin{pmatrix} 1 \\ -2 \end{pmatrix} \approx \begin{pmatrix} -0.44721 \\ 0.89443 \end{pmatrix}$$

Ein Blick auf Tabelle 5.2 (unten) zeigt, dass die Vektoren $\frac{1}{\|v_0\|} \cdot v_0, \frac{1}{\|v_1\|} \cdot v_1, \frac{1}{\|v_2\|} \cdot v_2, \dots$ gegen den zweiten der obigen zwei Lösungsvektoren strebt. Ferner bestätigt sich auch die “Gleichung” (5.8): Es gilt nämlich

$$\frac{\|v_{n+1}\|}{\|v_n\|} = \frac{\|S \cdot v_n\|}{\|v_n\|} \approx \frac{\|S \cdot \hat{v}\|}{\|\hat{v}\|} = \frac{\|3 \cdot \hat{v}\|}{\|\hat{v}\|} = \frac{3 \|\hat{v}\|}{\|\hat{v}\|} = 3.$$

Insbesondere folgt daraus, dass

$$\|v_{n+1}\| \approx 3 \|v_n\|,$$

d.h., die Längen der Vektoren nehmen immer um einen ungefähren Faktor 3 zu. □

5.2 Eigenwerte und Eigenvektoren

Im vorherigen Abschnitt haben wir bereits die Begriffe “Eigenwert” und “Eigenvektor” eingeführt. Wir betrachten nun die allgemeine Definition:

Für eine *quadratische* Matrix A (d.h., eine Matrix, die gleich viele Zeilen wie Spalten hat) betrachten wir das folgende Problem: *Finde Vektoren v (nicht gleich 0) und Zahlen λ , sodass die Gleichung*

$$A \cdot v = \lambda \cdot v \tag{5.10}$$

erfüllt ist.

Dann heisst v ein **Eigenvektor** und λ zugehöriger **Eigenwert** der Matrix A .

Beispiel 5.11 In Beispiel 5.4 haben wir gesehen, dass jedes skalare Vielfache des Vektors

$$\begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

(ausser der Nullvektor) die Gleichung (5.9) erfüllt und damit ein Eigenvektor der Matrix S zum Eigenwert $\lambda = 3$ ist. Insbesondere gibt es hier unendlich viele Eigenvektoren (dies ist übrigens immer so, da ja die Gleichung (5.10) mit einer beliebigen Zahl multipliziert und der Vektor v somit beliebig skaliert werden darf). □

Beispiel 5.12 Wir betrachten die Matrix

$$A = \begin{pmatrix} 2 & 3 \\ -1 & 6 \end{pmatrix}.$$

Um die Eigenvektoren und Eigenwerte von A zu bestimmen, betrachten wir die Gleichung (5.10) angewandt auf das aktuelle Beispiel:

$$\begin{pmatrix} 2 & 3 \\ -1 & 6 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \lambda \cdot \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

Ausmultiplizieren führt zu

$$\begin{aligned} 2v_1 + 3v_2 &= \lambda v_1 \\ -v_1 + 6v_2 &= \lambda v_2 \end{aligned}$$

und weiter

$$\begin{aligned} (2 - \lambda)v_1 + 3v_2 &= 0 \\ -v_1 + (6 - \lambda)v_2 &= 0 \end{aligned}$$

In Form der erweiterten Matrix lautet diese homogene lineare Gleichungssystem wie folgt:

$$\left(\begin{array}{cc|c} (2 - \lambda) & 3 & 0 \\ -1 & (6 - \lambda) & 0 \end{array} \right). \quad (5.13)$$

Wir suchen Lösungen v_1 und v_2 (die Koeffizienten der gesuchten Eigenvektoren), die nicht beide gleich 0 sind, da Eigenvektoren per Definition nicht der Nullvektor sein sollen. Dies bedeutet aber, dass das obige Gleichungssystem nicht nur die Nulllösung haben soll. Aus Abschnitt 1.4 wissen wir, dass die Koeffizientenmatrix

$$A_\lambda = \begin{pmatrix} (2 - \lambda) & 3 \\ -1 & (6 - \lambda) \end{pmatrix}$$

deshalb *singulär* sein soll, d.h.,

$$\det \begin{pmatrix} (2 - \lambda) & 3 \\ -1 & (6 - \lambda) \end{pmatrix} \stackrel{!}{=} 0.$$

Dies ist eine Gleichung für die Eigenwerte! Berechnen wir die Determinante, so gilt

$$\det \begin{pmatrix} (2 - \lambda) & 3 \\ -1 & (6 - \lambda) \end{pmatrix} = (2 - \lambda)(6 - \lambda) - (-1) \cdot 3 = \lambda^2 - 8\lambda + 15.$$

Dieser Ausdruck heisst **charakteristisches Polynom** der Matrix A . Wir wollen also

$$\lambda^2 - 8\lambda + 15 = 0$$

lösen, d.h., die Nullstellen des charakteristischen Polynoms finden. Wir erhalten

$$\lambda_1 = 3, \quad \lambda_2 = 5.$$

Diese Lösungen sind die beiden Eigenwerte der Matrix A .

Wir wollen nun zugehörige Eigenvektoren finden.

- Für $\lambda_1 = 3$: Wir setzen $\lambda = 3$ in (5.13) ein. Das Gleichungssystem hat dann die Form

$$\left(\begin{array}{cc|c} -1 & 3 & 0 \\ -1 & 3 & 0 \end{array} \right).$$

Es hat die Lösungsmenge

$$\left\{ \begin{pmatrix} 3s \\ s \end{pmatrix} : s \text{ eine beliebige Zahl} \right\}.$$

Alle Eigenvektoren zum Eigenwert $\lambda = 3$ sind somit skalare Vielfache des Vektors

$$\begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

- Für $\lambda_2 = 5$: Wir setzen $\lambda = 3$ in (5.13) ein. Das Gleichungssystem hat dann die Form

$$\left(\begin{array}{cc|c} -3 & 3 & 0 \\ -1 & 1 & 0 \end{array} \right).$$

Es hat die Lösungsmenge

$$\left\{ \begin{pmatrix} s \\ s \end{pmatrix} : s \text{ eine beliebige Zahl} \right\}.$$

Alle Eigenvektoren zum Eigenwert $\lambda = 5$ sind somit skalare Vielfache des Vektors

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

□

Der Vorgang zur Lösung des Eigenwertproblems (5.10) lässt sich in Form einer allgemein gültigen Methode formulieren.

1. Ausgehend von einer Matrix \mathbf{A} , definiere die Matrix \mathbf{A}_λ durch subtrahieren von λ (gedacht als Variable) entlang der Diagonalen der Matrix \mathbf{A} .

2. Finde die Nullstellen des charakteristischen Polynoms von \mathbf{A} , d.h., löse die Gleichung

$$\det(\mathbf{A}_\lambda) = 0 \tag{5.14}$$

nach λ auf. Die Lösungen sind die Eigenwerte der Matrix \mathbf{A} .

3. Für jeden Eigenwert λ löse das entsprechende homogene lineare Gleichungssystem

$$(\mathbf{A}_\lambda | \mathbf{0}).$$

Damit finden sich die zugehörigen Eigenvektoren der Matrix \mathbf{A} .

Wir halten folgendes fest:

1. Die Gleichung (5.14) ist ein Polynom vom Grad n in λ , wo n die Anzahl der Zeilen resp. Spalten der Matrix \mathbf{A} ist (dieses Polynom wird **charakteristisches Polynom** genannt). Die Gleichung hat maximal n verschiedene Lösungen, d.h., eine Matrix mit n Zeilen und Spalten *hat höchstens n Eigenwerte*.
2. Die Menge der Eigenvektoren, die zu einem Eigenwert gehören, bilden einen Unterraum (einen sogenannten **Eigenraum**). Dieser kann ein- oder auch mehrdimensional sein. Insbesondere kann es zu einem Eigenwert verschiedene linear unabhängige Eigenvektoren geben.

Beispiel 5.15 Sei

$$\mathbf{A} = \begin{pmatrix} 3 & 3 & 3 \\ 0 & -3 & -6 \\ 0 & -3 & 0 \end{pmatrix}.$$

Dann definieren wir:

$$\mathbf{A}_\lambda = \begin{pmatrix} 3 - \lambda & 3 & 3 \\ 0 & -3 - \lambda & -6 \\ 0 & -3 & -\lambda \end{pmatrix}.$$

Um die Eigenwerte zu finden, lösen wir die Gleichung

$$\det(\mathbf{A}_\lambda) = 0.$$

In MATLAB und OCTAVE lässt sich das charakteristische Polynom $\det(\mathbf{A}_\lambda)$ (hier vom Grad 3), mit dem Befehl `poly` bestimmen:

```
octave:1> A = [ 3  3  3
               0 -3 -6
               0 -3  0 ];
```

```
octave:2> poly(A)
```

```
ans =
```

```
1   0  -27  54
```

Die Antwort zeigt die Koeffizienten des charakteristischen Polynoms von \mathbf{A} , d.h., es gilt

$$\det(\mathbf{A}_\lambda) = 1 \cdot \lambda^3 - 0 \cdot \lambda^2 - 27 \cdot \lambda + 54.$$

Die Gleichung (5.14) lautet dann

$$\lambda^3 - 27\lambda + 54 = 0$$

und hat die Lösungen $\lambda_1 = 3$ und $\lambda_2 = -6$. Hier bemerken wir, dass das charakteristische Polynom in drei Faktoren zerfällt, $\lambda^3 - 27\lambda + 54 = (\lambda - 3)^2(\lambda + 6)$; da der Faktor $\lambda - 3$ zweimal

auftritt, heisst $\lambda_1 = 3$ *doppelte* Nullstelle. Die zugehörigen Eigenräume sind

$$E_{\lambda_1=3} = \left\{ \begin{pmatrix} s \\ t \\ -t \end{pmatrix} : s \text{ und } t \text{ beliebige Zahlen} \right\} = \text{span} \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \right)$$

und

$$E_{\lambda_2=-6} = \left\{ \begin{pmatrix} s \\ -2s \\ -s \end{pmatrix} : s \text{ eine beliebige Zahl} \right\} = \text{span} \left(\begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix} \right).$$

□

Bemerkung 5.16 Es kann durchaus vorkommen, dass eine Nullstelle des charakteristischen Polynoms (d.h., ein Eigenwert) einer Matrix *mehrfach* auftritt. Dies bedeutet, dass die Faktorisierung des charakteristischen Polynoms einen entsprechenden Faktor mehr als nur einmal enthält. Beispielsweise hat die Matrix

$$A = \begin{pmatrix} 1 & 3 & -4 \\ 0 & 1 & 15 \\ 0 & 0 & -2 \end{pmatrix}$$

das charakteristische Polynom

$$\det(A_\lambda) = \lambda^3 - 3\lambda + 2 = (\lambda - 1)^2(\lambda + 2).$$

Die Nullstelle und somit der Eigenwert $\lambda = 1$ tritt hier zweimal auf. Zu bemerken ist, dass die Häufigkeit des Auftretens der Nullstelle *nicht* unbedingt gleich der Dimension des entsprechenden Eigenraums sein muss.

5.3 Berechnung von Eigenwerten und Eigenvektoren in der Praxis

In praktischen Computeranwendungen ist die Berechnung der Eigenwerte mit Hilfe der Gleichung (5.14) **generell zu vermeiden**. Dies liegt daran, dass der Zusammenhang zwischen den Einträgen der Matrix A und den Lösungen von (5.14) “unscharf” ist; in anderen Worten: es besteht eine erhöhte Gefahr von Rundungsfehlern. Für eine Begründung verweisen wir auf die untenstehende Fussnote¹. Hinzu kommt, dass in praktischen Anwendungen oftmals nur die kleinsten oder grössten Eigenwerte (und zugehörige Eigenvektoren) von Interesse sind.

¹Wir betrachten die Matrix

$$A = \begin{pmatrix} 1 & 10^{-10} \\ 10^{-10} & 1 \end{pmatrix}.$$

Dann gilt:

$$\det(A_\lambda) = \lambda^2 - 2\lambda + (1 - 10^{-20}).$$

Die Eigenwerte von A sind die Lösungen von $\det(A_\lambda) = 0$, d.h.,

$$\lambda_1 = 1 + 10^{-10}, \quad \lambda_2 = 1 - 10^{-10}.$$

Numerische Verfahren zur Bestimmung von Eigenwerten und Eigenvektoren basieren auf anderen Lösungsansätzen wie die im vorherigen Abschnitt besprochene Methode. Hier werden oftmals Iterationsmethoden oder geeignete Matrixfaktorisierungen eingesetzt.

Im folgenden besprechen wir ein einfaches Verfahren zur Berechnung des grössten Eigenwerts und eines zugehörigen Eigenvektors. Zur Besprechung der Idee, erinnern wir uns an die Beispiel 5.4. Wir hatten dort die Matrix-Vektor-Iteration

$$\mathbf{v}_{n+1} = \mathbf{S} \cdot \mathbf{v}_n, \quad n = 0, 1, 2, \dots \quad (5.17)$$

betrachtet und festgestellt, dass die Folge der (auf Länge 1 skalierten) Vektoren

$$\frac{1}{\|\mathbf{v}_0\|} \cdot \mathbf{v}_0, \quad \frac{1}{\|\mathbf{v}_1\|} \cdot \mathbf{v}_1, \quad \frac{1}{\|\mathbf{v}_2\|} \cdot \mathbf{v}_2, \dots$$

gegen einen Vektor $\hat{\mathbf{v}}$ strebt, der die Gleichung

$$\mathbf{S} \cdot \hat{\mathbf{v}} = \lambda \cdot \hat{\mathbf{v}} \quad (5.18)$$

mit $\lambda = 3$ erfüllt.

Mathematisch lässt sich nun folgendes zeigen:

Unter gewissen (häufig erfüllten) Voraussetzungen² an die Matrix \mathbf{S} und den Startvektor \mathbf{v}_0 gilt für eine Matrix-Vektoriteration der Form (5.17):

- a) Das Längenverhältnis zweier aufeinander folgender Vektoren \mathbf{v}_n und \mathbf{v}_{n+1} strebt mit zunehmender Iterationszahl n immer mehr gegen den Betrag eines Eigenwert $\hat{\lambda}$ von \mathbf{S} , d.h.,

$$\frac{\|\mathbf{v}_{n+1}\|}{\|\mathbf{v}_n\|} \xrightarrow{n \rightarrow \infty} |\hat{\lambda}|.$$

Dieser Eigenwert $\hat{\lambda}$ ist derjenige Eigenwert von \mathbf{S} , der den grössten Betrag unter allen Eigenwerten hat. Wir nennen ihn deshalb auch **dominanten Eigenwert**.

- b) Die auf Länge 1 skalierten Vektoren

$$\frac{1}{\|\mathbf{v}_0\|} \cdot \mathbf{v}_0, \quad \frac{1}{\|\mathbf{v}_1\|} \cdot \mathbf{v}_1, \quad \frac{1}{\|\mathbf{v}_2\|} \cdot \mathbf{v}_2, \quad \frac{1}{\|\mathbf{v}_3\|} \cdot \mathbf{v}_3, \dots$$

streben mit wachsendem n immer mehr gegen einen Eigenvektor $\hat{\mathbf{v}}$ (mit der Länge 1) von \mathbf{S} , der zum dominanten Eigenwert $\hat{\lambda}$ gehört.

Numerische Software basiert oftmals auf etwa 15 Stellen exakter Rechnung. Das hat zur Folge, dass die Zahl $1 - 10^{-20}$ in der Regel nicht *exakt* abgespeichert wird, sondern auf 1 gerundet wird. Das charakteristische Polynom wird dann als

$$\det(\mathbf{A}_\lambda) = \lambda^2 - 2\lambda + 1$$

berechnet und führt zu den falschen Eigenwerten $\lambda_1 = \lambda_2 = 1$. Der bemerkenswerte Punkt besteht hier darin, dass die exakten Eigenwerte bei 15-stelliger Rechnung exakt dargestellt werden könnten. Die Berechnung mittels der Determinante respektive der Gleichung (5.14) führt aber zu Rundungsfehlern.

² Voraussetzungen, die die obigen Aussagen a) und b) sicherstellen, sind:

Bemerkung 5.19 Aus der obigen Aussage a) folgt, dass für genügend grosse Iterationszahlen n die Näherung

$$\|\mathbf{v}_{n+1}\| \approx |\hat{\lambda}| \|\mathbf{v}_n\|$$

gilt. Insbesondere sehen wir, dass die Längen der Vektoren \mathbf{v}_n in jedem Schritt (mindestens für grosse n)

- zunehmen um einen Faktor $|\hat{\lambda}|$, falls $|\hat{\lambda}| > 1$;
- etwa gleich bleiben, falls $|\hat{\lambda}| = 1$;
- abnehmen um einen Faktor $|\hat{\lambda}|$, falls $|\hat{\lambda}| < 1$.

Beispiel 5.20 Die Eigenwerte der Matrix \mathbf{A} aus Beispiel 5.15 sind $\lambda_1 = 3$ und $\lambda_2 = -6$. Die entsprechenden Beträge sind $|\lambda_1| = 3$ und $|\lambda_2| = 6$. Somit ist λ_2 der dominante Eigenwert. Die auf Länge 1 skalierten Vektoren der Vektoriteration

$$\mathbf{v}_{n+1} = \mathbf{A} \cdot \mathbf{v}_n, \quad n = 0, 1, 2, \dots$$

werden sich somit entlang des zugehörigen Eigenvektors

$$\pm \frac{1}{\sqrt{6}} \cdot \begin{pmatrix} 1 \\ -2 \\ -1 \end{pmatrix}$$

stabilisieren. Da $|\lambda_2| > 1$ werden die Längen der Vektoren in jedem Schritt um einen Faktor von etwa 6 zunehmen. \square

Die obigen Eigenschaften von Matrix-Vektoriterationen können nun für numerische Zwecke wie folgt benutzt werden: Um einen Eigenwert/Eigenvektor der Matrix \mathbf{S} zu berechnen, betrachte man die Iteration (5.17) und beobachte die resultierende (auf Länge 1 skalierte) Folge von Vektoren. Strebt diese Folge gegen einen festen Vektor $\hat{\mathbf{v}}$, so ist dies ein Eigenvektor der Matrix \mathbf{S} , und zwar derjenige, der zum dominanten Eigenwert gehört (immer vorausgesetzt, dass die nötigen Annahmen erfüllt sind). Der zugehörige dominante Eigenwert lässt sich ebenfalls ermitteln:

1. Der dominante Eigenwert $\hat{\lambda}$ ist ein einfacher Eigenwert von \mathbf{S} , d.h., $\hat{\lambda}$ ist eine Nullstelle des charakteristischen Polynoms, die einfach ist.
2. Angenommen, \mathbf{S} habe n Zeilen/Spalten. Dann soll es eine Basis von \mathbb{R}^n geben, die sich aus Eigenvektoren von \mathbf{S} zusammensetzt, d.h., es lassen sich n linear unabhängige Eigenvektoren finden. Eine solche Basis heisst **Eigenbasis** von \mathbf{S} und existiert nicht für jede Matrix.
3. Schreiben wir den Startvektor \mathbf{v}_0 in Koordinaten der Eigenbasis (vgl. Abschnitt 2.6), dann soll die dem dominanten Eigenvektor $\hat{\mathbf{v}}$ entsprechende Koordinate nicht gleich 0 sein.

Dazu multiplizieren wir die Gleichung (5.18) mit dem Vektor \hat{v} selbst (mit dem Skalarprodukt). Dies führt zu

$$\langle \hat{v}, \mathbf{S} \cdot \hat{v} \rangle = \langle \hat{v}, \lambda \cdot \hat{v} \rangle = \lambda \cdot \langle \hat{v}, \hat{v} \rangle = \lambda \|\hat{v}\|^2.$$

Nun gilt $\|\hat{v}\| = 1$ und somit

$$\lambda = \langle \hat{v}, \mathbf{S} \cdot \hat{v} \rangle.$$

Das hier beschriebene Verfahren heisst **Potenzmethode** (“power method” in Englisch). Wir fassen die Methode im folgenden Algorithmus zusammen:

Algorithmus 5.21 (Potenzmethode)

1. Wähle einen Startvektor v_0 .
2. Berechne

$$\begin{array}{ll} \tilde{v}_1 = \mathbf{A} \cdot v_0, & v_1 = \frac{1}{\|\tilde{v}_1\|} \cdot \tilde{v}_1 \\ \tilde{v}_2 = \mathbf{A} \cdot v_1, & v_2 = \frac{1}{\|\tilde{v}_2\|} \cdot \tilde{v}_2 \\ \tilde{v}_3 = \mathbf{A} \cdot v_2, & v_3 = \frac{1}{\|\tilde{v}_3\|} \cdot \tilde{v}_3 \\ \vdots & \vdots \end{array}$$

Fahre solange fort, bis die Iteration einen genügend stabilen Vektor \hat{v} berechnet hat (in der Praxis wird dies oft solange gemacht, bis sich in den Iterationen keine grossen Änderungen mehr ergeben, also beispielsweise bis $\|v_{n+1} - v_n\| \leq \text{vorgegebene Toleranz}$). Der Vektor \hat{v} ist dann eine Näherung eines Eigenvektors zum betragsmässig grössten Eigenwert von \mathbf{A} .

3. Der Eigenwert mit dem grössten Betrag ergibt sich dann als

$$\lambda = \langle \hat{v}, \mathbf{S} \cdot \hat{v} \rangle.$$

Beispiel 5.22 Wir wenden die Potenzmethode auf die Matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & 1 & 5 \\ 2 & 0 & 1 & 3 \\ 3 & 4 & 7 & 1 \\ 4 & 3 & 1 & 1 \end{pmatrix}$$

an. Als Startvektor wählen wir

$$v_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Wir berechnen die (auf Länge 1 skalierten Vektoren) v_1, v_2, v_3, \dots in Algorithmus 5.21 mit OCTAVE. Die 4 Komponenten der Vektoren $v_0, v_1, v_2, \dots, v_9$ schreiben wir jeweils waagrecht in einer Zeile auf:

```

1.00000  1.00000  1.00000  1.00000
0.47565  0.28539  0.71348  0.42809
0.40793  0.28741  0.77878  0.38012
0.38776  0.26852  0.80556  0.35863
0.37612  0.26353  0.81678  0.34921
0.37150  0.26064  0.82170  0.34476
0.36924  0.25944  0.82391  0.34283
0.36826  0.25888  0.82490  0.34192
0.36780  0.25862  0.82535  0.34152
0.36760  0.25851  0.82555  0.34134

```

Nach 9 Iterationen gilt $\|v_9 - v_8\| = 0.0003588$. Wir wollen uns in diesem Beispiel mit dieser Genauigkeit begnügen und akzeptieren den Vektor v_9 als Eigenvektor, d.h.,

$$\hat{v} = \begin{pmatrix} 0.36760 \\ 0.25851 \\ 0.82555 \\ 0.34134 \end{pmatrix}$$

Als zugehörigen Eigenwert erhalten wir

$$\lambda = \hat{v} \cdot (S \cdot \hat{v}) = 10.001.$$

In der Tat ist der grösste Eigenwert von A gleich 10. Die Potenzmethode hat also bereits nach 9 Iterationen ein vernünftiges Resultat geliefert. \square

Bemerkung 5.23 In MATLAB und OCTAVE lassen sich Eigenwerte und Eigenvektoren mit dem Befehl `eig` berechnen. Wir zeigen dies anhand des vorherigen Beispiels:

```

octave:1> A = [ 1 3 1 5
                2 0 1 3
                3 4 7 1
                4 3 1 1 ];

```

```

octave:2> [EV,EW] = eig(A)

```

EV =

```

0.36743  0.36021  0.61221  0.42640
0.25841  0.18011  0.30611 -0.85280
0.82572 -0.85831 -0.22482  0.21320

```

0.34119 0.31799 -0.69351 0.21320

EW =

10.00000	0.00000	0.00000	0.00000
0.00000	4.53113	0.00000	0.00000
0.00000	0.00000	-3.53113	0.00000
0.00000	0.00000	0.00000	-2.00000

Als Ausgabe liefert eig die beiden Matrizen EV und EW. Hier sind die Diagonaleinträge von EW die Eigenwerte der Matrix A , und die Spalten von EV sind die zugehörigen (in der gleichen Reihenfolge) Eigenvektoren (skaliert auf Länge 1).

5.4 Weitere Anwendungen

Mathematische Modelle in den Naturwissenschaften und der Wirtschaft können – nach geeigneten (teilweise starken) Vereinfachungen – oftmals durch Matrix-Vektoriterationen beschrieben werden. So lassen sich beispielsweise gewisse Prozesse in der Wirtschaft oder Ökologie durch Markov-Ketten beschreiben. Weiter kennt man sogenannte Leslie-Modelle aus der Biologie, die entwickelt wurden, um Populationsmodelle mit Generationenstruktur zu studieren. Auch sie basieren auf Matrix-Vektoriterationen. Im folgenden wollen wir einen kleinen Einblick in diese Anwendungsbereiche geben und die entsprechenden iterativen Modelle mit Hilfe von Eigenwerten und Eigenvektoren untersuchen.

5.4.1 Markovmodelle in Ökologie und Wirtschaft

Wir betrachten ein grosses Waldgebiet, in dem es drei verschiedene Klassen 1, 2, 3 von Bäumen gibt (hier kann eine Klasse auch mehrere Baumarten beinhalten). Wir nehmen an, dass alle Bäume etwa gleich alt werden. Ausserdem erneuert sich der Wald stetig, d.h., an die Stelle eines alten Baums tritt ohne Unterbruch ein neuer Baum. Wir gehen davon aus, dass der Wechsel der Baumgenerationen mit *festen* Wahrscheinlichkeiten erfolgt. Wird also beispielsweise ein alter Baum der Klasse 2 ersetzt durch einen neuen Baum der Klasse 3, so geschieht dies immer mit derselben Wahrscheinlichkeit. Wir bezeichnen diese Zahl mit $p_{3,2}$. Ganz allgemein verwenden wir die folgende Schreibweise: wird ein alter Baum einer Klasse i (wo i gleich 1, 2 oder 3 sein kann) ersetzt durch einen jungen Baum einer Klasse j (hier darf j ebenfalls gleich 1, 2 oder 3 sein), so geschieht dies immer mit genau der gleichen Wahrscheinlichkeit, die wir mit $p_{j,i}$ bezeichnen (zu beachten ist, dass ein Wechsel von Klasse i nach Klasse j mit Wahrscheinlichkeit $p_{j,i}$ geschieht, d.h., die Indizes von p sind genau umgekehrt zum Wechsel $i \rightarrow j$). Die verschiedenen Wahrscheinlichkeiten speichern wir in einer Matrix (Systemmatrix):

$$P = \begin{pmatrix} p_{1,1} & p_{1,2} & p_{1,3} \\ p_{2,1} & p_{2,2} & p_{2,3} \\ p_{3,1} & p_{3,2} & p_{3,3} \end{pmatrix}.$$

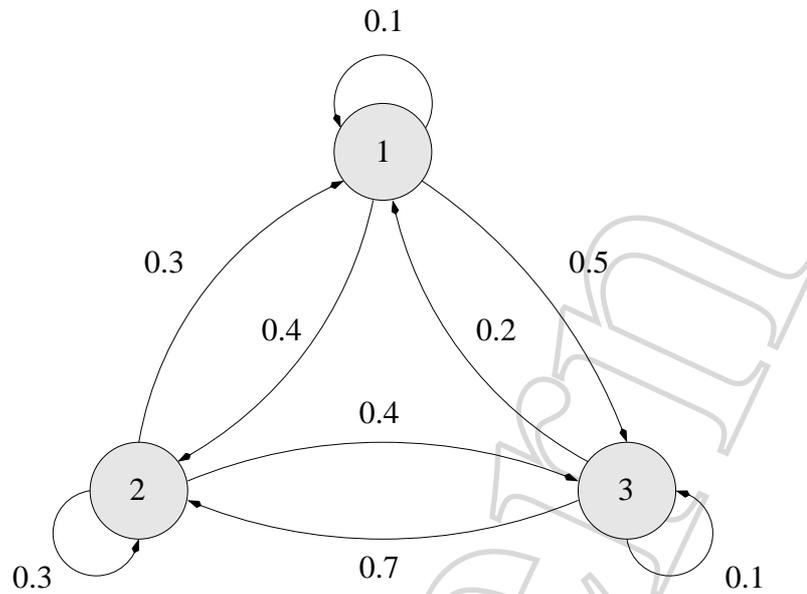


Abbildung 5.3: Gerichteter Graph für das Baummodell.

Betrachten wir ein konkretes Beispiel:

$$P = \begin{pmatrix} 0.1 & 0.3 & 0.2 \\ 0.4 & 0.3 & 0.7 \\ 0.5 & 0.4 & 0.1 \end{pmatrix}. \quad (5.24)$$

Dies bedeutet beispielsweise, dass auf einen alten Baum aus der Klasse 2 mit den Wahrscheinlichkeiten $p_{1,2} = 0.3, p_{2,2} = 0.3, p_{3,2} = 0.4$ ein Baum aus der Klasse 1, 2 bzw. 3 folgt. Wir können dies auch, wie beim Geysir Old Faithful (Anwendung 5.1), in Form eines gerichteten Graphen darstellen; siehe Abbildung 5.3.

Nun wollen wir die langfristige Entwicklung eines Waldes mit der obigen Systemmatrix betrachten. Dazu bezeichnen wir mit a_n, b_n und c_n die (prozentualen) Anteile der Baumklassen 1, 2, und 3 zum Zeitpunkt n . Hier deutet der Index \cdot_n den Iterationsschritt an, d.h., $n = 0$ entspricht der anfänglichen Verteilung der Klassen, $n = 1$ der Verteilung nach einem Wechsel, und so weiter. Es gilt:

$$\begin{aligned} a_n = & \text{(Anteil } a_{n-1} \text{ der Klasse 1)} \cdot \text{(Wahrscheinlichkeit, dass ein Baum der Klasse 1} \\ & \text{durch einen Baum der Klasse 1 ersetzt wird)} \\ & + \text{(Anteil } b_{n-1} \text{ der Klasse 2)} \cdot \text{(Wahrscheinlichkeit, dass ein Baum der Klasse 2} \\ & \text{durch einen Baum der Klasse 1 ersetzt wird)} \\ & + \text{(Anteil } c_{n-1} \text{ der Klasse 3)} \cdot \text{(Wahrscheinlichkeit, dass ein Baum der Klasse 3} \\ & \text{durch einen Baum der Klasse 1 ersetzt wird)}. \end{aligned}$$

Daher

$$a_n = a_{n-1}p_{1,1} + b_{n-1}p_{1,2} + c_{n-1}p_{1,3}.$$

Genau gleich erhalten wir

$$b_n = a_{n-1}p_{2,1} + b_{n-1}p_{2,2} + c_{n-1}p_{2,3}$$

$$c_n = a_{n-1}p_{3,1} + b_{n-1}p_{3,2} + c_{n-1}p_{3,3}.$$

Die drei Gleichungen lassen sich schreiben in Form einer Matrix-Vektoriteration,

$$\begin{pmatrix} a_n \\ b_n \\ c_n \end{pmatrix} = \mathbf{P} \cdot \begin{pmatrix} a_{n-1} \\ b_{n-1} \\ c_{n-1} \end{pmatrix}, \quad (5.25)$$

wobei \mathbf{P} die Matrix in (5.24) ist.

Am Anfang muss $a_0 + b_0 + c_0 = 100\% = 1$ gelten, denn die einzelnen prozentualen Anteile von Baumklassen ergeben zusammen 1. Es lässt sich einfach nachrechnen, dass dies für alle Iterationen so bleibt, d.h., wir haben immer $a_n + b_n + c_n = 1$. Weiter lässt sich zeigen, dass der betragsmässig grösste Eigenwert von \mathbf{P} gleich 1 ist, und dass es genau einen zugehörigen Eigenvektor gibt, dessen Komponenten alle positiv sind und die Summe 1 haben (dies ist typisch für Markovprozesse). Unserer Ausführungen (vgl. beispielsweise Anwendung 5.1) in den vorherigen Abschnitten zu Folge, ist dies auch der Vektor entlang dem sich die Vektoriteration (5.25) stabilisiert.

Als Beispiel berechnen wir die Eigenwerte und Eigenvektoren der Matrix \mathbf{P} in (5.24) mit OCTAVE:

```
octave:1> P = [ 0.1 0.3 0.2
               0.4 0.3 0.7
               0.5 0.4 0.1 ];

octave:2> [EV, EW] = eig(P)

ev =

    0.3717    0.3873 + 0.2236i    0.3873 - 0.2236i
    0.7540   -0.7746             -0.7746
    0.5416    0.3873 - 0.2236i    0.3873 + 0.2236i

ew =

    1.0000    0                0
           0   -0.2500 + 0.0866i    0
           0    0                -0.2500 - 0.0866i
```

Wir erhalten einen Eigenwert $\lambda_1 = 1$ mit zugehörigem Eigenvektor

$$\mathbf{v}_1 = \begin{pmatrix} 0.3717 \\ 0.7540 \\ 0.5416 \end{pmatrix};$$

die anderen beiden Eigenwerte sind hier komplexwertig, mit Betrag kleiner als 1. Um die langfristige Verteilung der Baumklassen zu berechnen, dividieren wir den Vektor \mathbf{v}_1 durch die Summe seiner Komponenten, sodass die Summe der Komponenten des neuen Vektors genau gleich 1 ist:

```
octave:3> v1 = ev(:,1)
```

```
v1 =
```

```
0.3717
0.7540
0.5416
```

```
octave:4> w1 = v1/sum(v1)
```

```
w1 =
```

```
0.2229
0.4522
0.3248
```

Der berechnete Vektor \mathbf{w}_1 zeigt uns nun die langfristige Verteilung der Bäume: etwa 22.3% Bäumen der Klasse 1, 45.2% Bäume der Klasse 2 und 32.5% der Klasse 3. Allerdings sei hier die wichtige Bemerkung gemacht, dass die einzelnen Wahrscheinlichkeitszahlen in der Matrix \mathbf{P} typischerweise mit der Zeit variieren und das obige Modell deshalb eine klare Vereinfachung der Wirklichkeit darstellt und – wenn überhaupt – nur innerhalb eines beschränkten zeitlichen Rahmens gültig sein kann.

Ähnliche Modelle kennt man auch aus der Wirtschaft. Betrachten wir als fiktives Beispiel einen Wirtschaftsraum, der aus vier verschiedenen Wirtschaftszweigen besteht: (A) Landwirtschaft und Nahrungsmittelproduktion, (B) Industrie, (C) Energieversorgung, Transporte und Entsorgung, (D) Dienstleistungen und Verwaltung. Im Laufe eines gewissen Zeitraums (d.h., einer Iteration), finden Geldtransaktionen zwischen den verschiedenen Branchen statt: von einem Euro, der sich im Wirtschaftszweig (X) befindet gelangt ein prozentualer Anteil p_{YX} in die Branche (Y). Der Geldfluss ist dann eine Markovkette, die durch eine 4×4 -Systemmatrix \mathbf{P} beschrieben wird. Zum Beispiel:

$$\mathbf{P} = \begin{pmatrix} 0.30 & 0.20 & 0.15 & 0.15 \\ 0.30 & 0.30 & 0.30 & 0.15 \\ 0.20 & 0.40 & 0.25 & 0.15 \\ 0.20 & 0.10 & 0.30 & 0.55 \end{pmatrix}.$$

Auch hier hat die Matrix P einen dominanten Eigenwert 1. Die langfristige Entwicklung der einzelnen Wirtschaftszweige kann nun, wie oben im Waldmodell, mit Hilfe des entsprechenden Eigenvektors untersucht werden.

5.4.2 Leslie-Modelle in der Biologie

In diesem Abschnitt wollen wir ein *stark vereinfachtes Modell* zur Beschreibung der Entwicklung einer Population von Menschen, Tieren oder Pflanzen untersuchen. Das Modell soll auf einer diskreten Zeitentwicklung basieren, d.h., auf Zeitpunkten $n = 0, 1, 2, \dots$. Der Abstand dieser Zeitpunkte soll einer festen Zeiteinheit (ZE) Δt entsprechen; bei einer menschlichen Population könnte Δt beispielsweise gleich 1 Jahr sein, d.h., $\Delta t \cdot n$ steht hier für die Anzahl Jahre seit dem Anfang. Weiter wollen wir berücksichtigen, dass es in einer Bevölkerung typischerweise verschiedene Generationen und somit eine Altersstruktur gibt. Nehmen wir dazu an, dass es in der betrachteten Population m verschiedene Generationen gibt. Hierbei besteht die erste Generation aus allen Individuen, die zwischen 0 und 1 Zeiteinheiten alt sind, die zweite Generation aus allen Individuen, die zwischen 1 und 2 Zeiteinheiten alt sind, etc. Die m -te Generation bildet sich dann aus allen Individuen, die ein Alter zwischen $m - 1$ und m Zeiteinheiten haben. Wir gehen davon aus, dass kein Individuum in der Population älter als m Zeiteinheiten wird.

Wir bezeichnen mit $N_{n,1}$ die Anzahl der Individuen, die sich zum Zeitpunkt n in der ersten Generation befinden (also gerade geboren wurden). Weiter sei $N_{n,2}$ die Anzahl der Individuen, die zum Zeitpunkt n in der zweiten Generation (also zwischen 1 und 2 ZE alt) sind, usw. Die Altersstruktur der Population zum Zeitpunkt n (also nach einer Zeit von $\Delta t \cdot n$) lässt sich als Vektor darstellen:

$$v_n = \begin{pmatrix} N_{n,1} \\ N_{n,2} \\ N_{n,3} \\ \vdots \\ N_{n,m} \end{pmatrix} = \begin{pmatrix} \# \text{ Individuen im Alter von 0 bis 1 ZE zum } n\text{-ten Zeitschritt} \\ \# \text{ Individuen im Alter von 1 bis 2 ZE zum } n\text{-ten Zeitschritt} \\ \# \text{ Individuen im Alter von 2 bis 3 ZE zum } n\text{-ten Zeitschritt} \\ \vdots \\ \# \text{ Individuen im Alter von } m-1 \text{ bis } m \text{ ZE zum } n\text{-ten Zeitschritt} \end{pmatrix}.$$

Nun beziehen wir die Vermehrung in das Modell mit ein. Dazu nehmen wir an, dass jede Generation eine gewisse Anzahl Junge erzeugt: Generation 1 erzeugt durchschnittlich J_1 Junge pro ZE und Individuum, Generation 2 erzeugt durchschnittlich J_2 Junge pro ZE und Individuum, ..., Generation m erzeugt J_m Junge pro ZE und Individuum. Die Anzahl der neugeborenen Jungtiere im aktuellen Zeitabschnitt macht die neue 1. Generation des nächsten Zeitabschnitts aus und besteht aus

$$N_{n+1,1} = J_1 N_{n,1} + J_2 N_{n,2} + \dots + J_m N_{n,m} \quad (5.26)$$

Individuen.

Es fehlt noch die Beschreibung der Überlebenschancen von Generation zu Generation. Wir bezeichnen die Überlebenschancen der 1. Generation mit p_1 (gemessen in %, d.h. $100 \cdot p_1$ von 100 Individuen überleben), die Überlebenschancen der 2. Generation mit p_2 , etc. Die neue zweite Generation wird nun von den Überlebenden der 1. Generation gebildet:

$$N_{n+1,2} = p_1 \cdot N_{n,1}. \quad (5.27)$$

Genau gleich gilt

$$\begin{aligned}
 N_{n+1,3} &= p_2 \cdot N_{n,2} \\
 N_{n+1,4} &= p_3 \cdot N_{n,3} \\
 &\vdots \\
 N_{n+1,m} &= p_{m-1} \cdot N_{n,m-1}.
 \end{aligned} \tag{5.28}$$

Die m -te Generation überlebt nicht.

Fassen wir die Gleichungen (5.26) – (5.28) in Vektorform zusammen, so gilt:

$$\begin{pmatrix} N_{n+1,1} \\ N_{n+1,2} \\ N_{n+1,3} \\ N_{n+1,4} \\ \vdots \\ N_{n+1,m} \end{pmatrix} = \begin{pmatrix} J_1 N_{n,1} + J_2 N_{n,2} + \dots + J_m N_{n,m} \\ p_1 \cdot N_{n,1} \\ p_2 \cdot N_{n,2} \\ p_3 \cdot N_{n,3} \\ \vdots \\ p_{m-1} \cdot N_{n,m-1} \end{pmatrix},$$

oder äquivalent,

$$\mathbf{v}_{n+1} = \mathbf{L} \cdot \mathbf{v}_n, \tag{5.29}$$

mit der Matrix

$$\mathbf{L} = \begin{pmatrix} J_1 & J_2 & J_3 & \dots & J_{m-1} & J_m \\ p_1 & 0 & 0 & \dots & 0 & 0 \\ 0 & p_2 & 0 & \dots & 0 & 0 \\ 0 & 0 & p_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & p_{m-1} & 0 \end{pmatrix}. \tag{5.30}$$

Hier wird \mathbf{L} **Leslie-Matrix** und die Matrix-Vektoriteration (5.29) **diskretes Leslie-Modell** genannt. Startend von einer anfänglichen Altersverteilung, gegeben durch einen Vektor \mathbf{v}_0 , entwickelt sich eine Population im Leslie-Modell nun nach der Vorschrift (5.29). Es lässt sich einfach mit Hilfe der Matrix-Vektormultiplikation zeigen: wenn alle Komponenten von \mathbf{v}_0 positiv sind (was in der Praxis immer der Fall ist, da es ja keine negative Anfangspopulation geben kann), so haben auch alle weiteren Vektoren $\mathbf{v}_1, \mathbf{v}_2, \dots$ nur positive Komponenten.

Aussagen über Wachstum, Stillstand, oder Rückgang der Population, sowie über das Konvergieren gegen eine stabile Verteilung lassen sich nun ähnlich wie zuvor mit Hilfe der Eigenvektoren und Eigenwerte der Leslie-Matrix \mathbf{L} diskutieren. Es lässt sich mathematisch beweisen, dass es bei Leslie-Matrizen immer einen positiven Eigenwert gibt (dominanter Eigenwert), der grösser als alle anderen Eigenwerte ist (falls \mathbf{L} komplexwertige Eigenwerte hat, was vorkommen kann, sind die Realteile dieser Eigenwerte ebenfalls kleiner als der dominante Eigenwert). Der zugehörige Eigenvektor hat Komponenten, die *alle das gleiche Vorzeichen haben*. Er gibt die Richtung an, entlang derer sich die Iteration (5.29) (fast immer) stabilisiert.

Wir betrachten zwei fiktive menschliche Bevölkerungen. Als Zeiteinheit wählen wir 10 Jahre. Hier betrachten wir lediglich die Alterstufen zwischen 0 und 50 Jahren, da die Generationen

5 EIGENWERTPROBLEME

der über 50-jährigen nur noch sehr wenig zur Vermehrung der Bevölkerung beitragen. Gegeben seien

$$\mathbf{L}_1 = \begin{pmatrix} 0 & 0.05 & 0.75 & 0.25 & 0.1 \\ 0.98 & 0 & 0 & 0 & 0 \\ 0 & 0.99 & 0 & 0 & 0 \\ 0 & 0 & 0.99 & 0 & 0 \\ 0 & 0 & 0 & 0.95 & 0 \end{pmatrix}, \quad \mathbf{L}_2 = \begin{pmatrix} 0 & 0.53 & 1.7 & 1.33 & 0.3 \\ 0.52 & 0 & 0 & 0 & 0 \\ 0 & 0.59 & 0 & 0 & 0 \\ 0 & 0 & 0.67 & 0 & 0 \\ 0 & 0 & 0 & 0.44 & 0 \end{pmatrix}.$$

Bei \mathbf{L}_1 könnte es sich um eine Leslie-Matrix einer Population im technologisch hochentwickelten Zustand handeln (grosse Überlebenswahrscheinlichkeiten, kleine Kindersterblichkeit, wenig Nachkommen), während \mathbf{L}_2 eine vorindustrielle Bevölkerung beschreiben könnte (kleinere Überlebenschancen und grosse Kindersterblichkeit, grössere Vermehrung).

Wir berechnen die Eigenwerte und Eigenvektoren dieser Matrizen mit Hilfe von OCTAVE:

```
octave:1> L1 = [0      0.05   0.75   0.25   0.1
               0.98   0       0       0       0
               0      0.99   0       0       0
               0      0       0.99  0       0
               0      0       0       0.95  0];

octave:2> [EV, EW] = eig(L1)

EV =

 0.4907 -0.0507+0.2674i -0.0507-0.2674i  0.0076-0.0180i  0.0076+0.0180i
 0.4663  0.3238-0.0834i  0.3238+0.0834i -0.0511+0.0048i -0.0511-0.0048i
 0.4476 -0.2647-0.3197i -0.2647+0.3197i  0.0753+0.1134i  0.0753-0.1134i
 0.4297 -0.2192+0.4662i -0.2192-0.4662i  0.1708-0.3185i  0.1708+0.3185i
 0.3959  0.6135          0.6135          -0.9208          -0.9208

EW =

 1.0313  0          0          0          0
 0 -0.3394+0.7218i  0          0          0
 0 0          -0.3394-0.7218i  0          0
 0 0          0          -0.1762+0.3287i  0
 0 0          0          0          -0.1762-0.3287i
```

Die Eigenwerte von \mathbf{L}_1 sehen wir auf der Diagonalen der Matrix EW. Die entsprechenden (auf Länge 1 normierten) Eigenvektoren sind die Spalten der Matrix EV. Wir sehen, dass es einen reellen positiven Eigenwert (dominanter Eigenwert) $\lambda_1 = 1.0313$ und vier komplexwertige Eigenwerte (mit negativen Realteilen) gibt. Das Modell wird sich somit entlang des ersten

Eigenvektors

$$\hat{v}_1 = \begin{pmatrix} 0.4907 \\ 0.4663 \\ 0.4476 \\ 0.4297 \\ 0.395 \end{pmatrix}$$

stabilisieren, der zum Eigenwert λ_1 gehört. Da λ_1 leicht grösser als 1 ist, erwarten wir einen sanften Zuwachs der Bevölkerung.

Betrachten wir die Matrix L_2 :

```
octave:1> L2= [0      0.53   1.7    1.33   0.3
               0.52   0       0      0      0
               0      0.59   0       0      0
               0      0       0.67   0      0
               0      0       0       0.44  0 ];
```

```
octave:2> [EV, EW]=eig(L2)
```

EV =

```
-0.8515  -0.6214  -0.6214  -0.2395  0.0029
-0.4294   0.1609+0.4440i  0.1609-0.4440i  0.2856  -0.0118
-0.2457   0.3127-0.2609i  0.3127+0.2609i  -0.3865  0.0540
-0.1596  -0.3446-0.2009i  -0.3446+0.2009i  0.5938  -0.2808
-0.0681  -0.0460+0.2524i  -0.0460-0.2524i  -0.5992  0.9582
```

EW =

```
1.0312   0           0           0           0
      0  -0.2331 + 0.6432i  0           0           0
      0   0           -0.2331 - 0.6432i  0           0
      0   0           0           -0.4360  0
      0   0           0           0          -0.1289
```

Hier gibt es drei reelle und zwei komplexe Eigenwerte. Nur der erste Eigenwert $\lambda_1 = 1.0312$ hat einen Realteil grösser als 0. Er ist dominant. Die Bevölkerungsstruktur wird sich folglich entlang des ersten Eigenvektors (multipliziert mit -1) stabilisieren:

$$\hat{v}_1 = \begin{pmatrix} 0.8515 \\ 0.4294 \\ 0.2457 \\ 0.1596 \\ 0.0681 \end{pmatrix}$$

Auch hier ist der dominante Eigenwert leicht grösser als 1, sodass ein leichter Anstieg der Bevölkerung erwartet wird.

Wenn wir die beiden Modelle für L_1 und L_2 vergleichen, so scheint deren Analysis zunächst sehr ähnlich. Trotzdem macht sich ein Unterschied sehr deutlich: Mit dem Eigenvektor \hat{v}_1 der ersten Bevölkerung sehen wir, dass sich die Population auf einen Zustand zubewegt, bei dem alle Altersgruppen etwa gleich stark vertreten sind (ausser die fünfte, die einen etwas kleineren Anteil hat). Beim zweiten Modell ist die Situation ganz anders. Betrachten wir dort die stabile Richtung, so erkennen wir deutlich, dass die erste Altersschicht sehr stark vertreten ist, während die Grösse der folgenden Altersgruppen jeweils um einen Faktor von etwa 2 kleiner sind. In den beiden Modellen stellt sich also langfristig eine ganz andere Alterstruktur ein.

Leslie-Modelle mit konstanten Matrizen sind üblicherweise ungenügend, um eine langzeitige Entwicklung einer Population zu beschreiben. Viele zeitabhängige Faktoren, wie medizinische, umweltbezogene oder wirtschaftliche Veränderungen, oder Zu- und Wegwanderungen, spielen bei der Entwicklung einer Bevölkerung eine wichtige Rolle. Eine mögliche Verbesserung des Leslie-Modells besteht darin, zeitabhängige Koeffizienten in die Leslie-Matrizen einzubauen. So wurden beispielsweise in vielen Teilen Westeuropas seit Beginn des letzten Jahrhunderts hohe Geburtenraten und niedrigere Lebenserwartung durch weniger Nachkommenszahlen und höhere Lebenserwartung abgelöst.

5.5 Übungsaufgaben

5.1. Betrachten Sie die Matrix

$$A = \begin{pmatrix} 5 & 3 \\ -1 & 1 \end{pmatrix}.$$

- Bestimmen Sie alle Eigenwerte und die zugehörigen Eigenvektoren von A "von Hand".
- Es sei $v_0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Berechnen Sie die ersten drei Iterierten v_1 , v_2 und v_3 der Matrix-Vektor-Iteration

$$v_{n+1} = A \cdot v_n \quad n \geq 0$$

und skalieren Sie sie auf Länge 1. Berechnen Sie dann den Vektor \hat{v} , gegen welchen die Folge der Vektoren $\frac{v_0}{\|v_0\|}$, $\frac{v_1}{\|v_1\|}$, $\frac{v_2}{\|v_2\|}$, ... (für $n \rightarrow \infty$) strebt, *exakt*.

5.2. Wiederholen Sie die obige Aufgabe für die Matrix

$$A = \begin{pmatrix} -3 & 1 \\ 0 & 1 \end{pmatrix}.$$

- Ein Fotokopierer ist in einem der beiden Betriebszustände [b] (betriebsbereit) oder [d] (defekt). Jedes defekte Gerät kommt innerhalb von 24 Stunden zur Servicestelle und wird im Lauf eines Tages mit Wahrscheinlichkeit 50% repariert. Mit der gleichen Wahrscheinlichkeit lässt sich der Defekt nicht innerhalb eines Tages beheben. Ein betriebsbereites

Gerät wird mit Wahrscheinlichkeit 1% im Laufe eines Tages einen Defekt erleiden und mit Wahrscheinlichkeit 99% betriebsbereit bleiben.

- Stellen Sie den gerichteten Graphen auf, welcher dieses Modell beschreibt.
- Beschreiben Sie das Verhalten des Systems durch eine Matrix-Vektor-Iteration.
- Welches ist der stabile Vektor der Matrix-Vektor-Iteration, und was bedeutet er in diesem Beispiel?

- 5.4. Wir betrachten das folgende (stark vereinfachte) Modell für Produktion und Abbau von roten Blutkörperchen (rBK). Wir setzen voraus, dass jeden Tag ein gewisser Anteil k ($0 < k < 1$) von rBK abgebaut wird und dass am gleichen Tag so viele neue rBK produziert werden, wie am vorherigen Tag abgebaut wurden. Dies führt zu

$$\begin{aligned} A_n &= (1 - k)A_{n-1} + N_n \\ N_n &= \gamma k A_{n-1}. \end{aligned}$$

Hier sind A_n die Anzahl rBK im Kreislauf und N_n die neuproduzierten rBK am Tag n .

- Erklären Sie diese Gleichungen.
 - Leiten Sie eine Matrix-Vektor-Formulierung her und zeigen Sie, dass die entsprechende Matrix die Struktur einer Leslie-Matrix hat.
 - Man kann Homöostase (Gleichgewicht im Körper) wie folgt definieren: Es gibt eine konstante Grösse G , sodass $A_n \rightarrow G$ für grosse n . Wie müssen die Parameter k und γ in diesem Modell sein, damit dies erreicht wird?
- 5.5. Wir betrachten eine Population von weiblichen Blauwalen. Die Population wird in die folgenden Altersgruppen eingeteilt:

- I: 0 bis 3 Jahre
- II: 4 bis 7 Jahre
- III: 8 bis 11 Jahre
- IV: mindestens 12 Jahre.

Die Populationsdynamik wird durch den gerichteten Graphen in Abbildung 5.4 beschrieben.

- Begründen oder widerlegen Sie die Aussage: Das langfristige Überleben der Population ist sichergestellt.
- Angenommen, die Population hat schon lange unter gleichen Bedingungen gelebt. Welche prozentuale Verteilung der Tiere auf die Altersklassen ist zu erwarten?
- Angenommen, das mittlere Alter in den Altersklassen werde durch folgende Tabelle beschrieben:

Altersklasse	I	II	III	IV
mittleres Alter	1	5	9	14

Welches ist das mittlere Alter der ganzen Population, falls die Altersverteilung stabil ist?

5 EIGENWERTPROBLEME

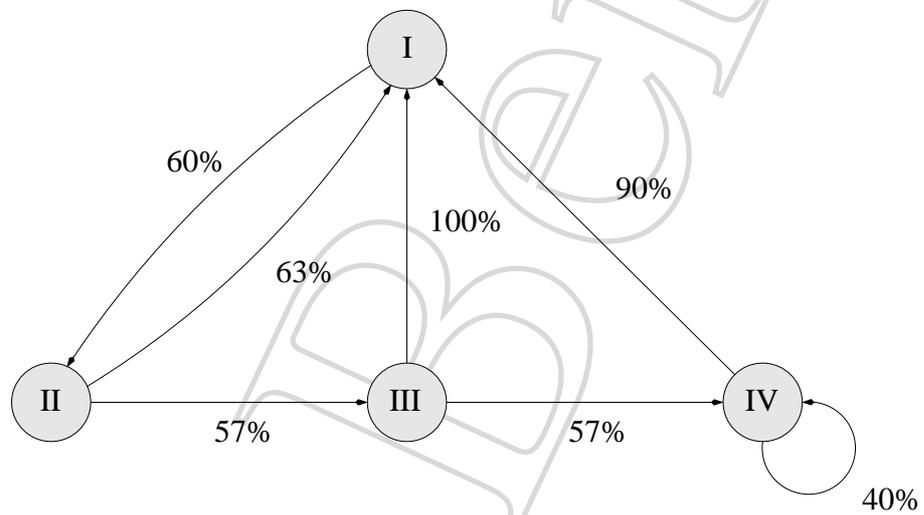


Abbildung 5.4: Dynamik einer Blauwalpopulation.

Ergänzende Literatur

H. Anton und C. Rorres, Elementary Linear Algebra (Applications Version), 9. Ausgabe, Verlag John Wiley & Sons, 2005.

P. Bützer, Mathematisches Ausgleichen chemischer Reaktionsgleichungen, frei verfügbar auf www.swisseduc.ch/ (Chemie), 2006.

A. Quarteroni und F. Saleri, Scientific Computing with MATLAB and Octave, 2. Auflage, Springer Verlag, 2006.

H. R. Schneebeli, Von Kormoranen, Eiern und Matrizen, frei verfügbar auf www.swisseduc.ch (Mathematik), 2006.

H. R. Schneebeli und H. R. Vollmer, Skalarprodukte - Schwingungen - Signale, frei verfügbar auf www.swisseduc.ch (Mathematik), 2005.

U
ni
B
erlin

Abbildungsverzeichnis

1.1	Elektrischer Stromkreis	8
1.2	Vereinfachtes Atmosphärenmodell	18
2.1	Zahlengerade.	39
2.2	Linearkombination zweier Vektoren.	43
2.3	Punkt P und Ortsvektor v in \mathbb{R}^2	45
2.4	Zwischenwinkel zweier Vektoren.	49
2.5	Projektion von v auf z	50
3.1	Drehung um 90°	79
3.2	Lineare Abbildung von \mathbb{R}^3 nach \mathbb{R}^2	86
3.3	Verknüpfung zweier Funktionen.	88
4.1	Freier Fall eines Partikels.	97
4.2	Kompromisslösung mittels kleinster Quadrate.	99
4.3	Gerade durch 3 Punkte.	100
4.4	Orthogonalprojektion auf eine Ebene.	102
4.5	Verbindungen zwischen Basel, Bern, Genf und Zürich.	105
4.6	Durchschnittstemperaturen am Jungfraujoch 1	107
4.7	Durchschnittstemperaturen am Jungfraujoch 2	111
5.1	Geysir Old Faithful. Quelle: wikipedia.org	117
5.2	Gerichteter Graph für Geysir Old Faithful.	118
5.3	Gerichteter Graph für das Baummodell.	133
5.4	Dynamik einer Blauwalpopulation.	142

U
ni
B
erlin

Index

- span (), 57
- \mathbb{R} , 39
- \mathbb{R}^2 , 39
- \mathbb{R}^n , 40
 - linearer Unterraum von, 53
 - Standardbasis, 67
- rang(), 24
- \, 30
- det, 29

- Absolutbetrag, 44
- Anwendungen
 - chemische Reaktionen, 31
 - Composition Space, 70
 - Datenübermittlung, 47, 66
 - Geometrie, 19
 - Interpolation von Daten, 17
 - Längenmessung, 104
 - Mischungen und Konzentrationen, 18

- Basis, 64
- Betrag, 44

- charakteristisches Polynom, 124
- chemische Reaktionen, 31

- Dateninterpolation, 17
- Determinante, 27, 28
 - in Matlab und Octave, 29
- Dimension, 64
- Dreiecksungleichung, 46

- Eigenraum, 126
- Eigenvektor, 122, 123
- Eigenwert, 122, 123
 - dominanter, 128
 - mehrfacher, 127

- Eigenwertproblem
 - in Matlab und Octave, 131
 - Lösung, 125
 - Potenzmethode, 130

- Fouriertransformierte, 110

- Gauss-Verfahren, 7, 13
- gerichteter Graph, 118
- Geysir Old Faithful, 117

- Interpolation von Daten, 17
- inverse lineare Abbildung, 92
- ISBN-Nummer, 47
- Iterationsmatrix, 119
- iterativer Prozess, 119

- Kern einer linearen Abbildung, 86
- Koeffizientenmatrix, 16
- Konzentrationen, 18
- Koordinaten, 69

- Lösungsmenge, 22
- Lesliematrix, 137
- Lesliemodelle, 136
- linear abhängig, 62
- linear unabhängig, 62
- lineare Abbildungen
 - Definition, 79
 - inverse, 92
 - Kern, 86
 - Matrizen von, 84
 - Verknüpfungen, 87
- lineare Gleichungssysteme, 13
 - homogene, 23
 - inhomogene, 23
 - Lösung in Octave, 30

INDEX

- Lösung mit Gauss-Verfahren, 13
 - überbestimmt, 30
 - unterbestimmt, 31
 - Unterräume, 55
- lineare Hülle, 57
- lineare Unterräume
 - Basis, 64
 - Dimension, 64
 - von \mathbb{R}^n , 53
- Linearkombination, 42
- Markovmodelle, 132
- Markovprozess, 119
- Matlab, *siehe* Octave
- Matrix, 10
 - Determinanten, 27, 28
 - Eingabe in Octave, 29
 - erweiterte, 16
 - inverse, 92
 - Koeffizientenmatrix, 16
 - quadratische, 28
 - reguläre, 23
 - singuläre, 23
- Matrix-Vektor-Iteration, 119
- Matrix-Vektor-Multiplikation, 83
- Matrixmultiplikation, 89
- Mengen, 39
- Mischungen, 18
- Norm, 45
- Normalgleichungen, 103
- Nullzeile, 24
- Octave
 - \, 30
 - det, 29
 - Determinante, 29
 - Eigenwertproblem, 131
 - Eingabe von Matrizen, 29
 - lineare Gleichungssysteme, 30
- Orthogonalprojektion, 101
- Ortsvektoren, 39
- Parameter, 21
- Potenzmethode, 130
- Projektionseigenschaft, 50
- Rückwärtseinsetzen, 10
- Rang, 24, 63
- reelle Zahlen, 39
- Skalarmultiplikation, 41
- Skalarprodukt, 47
 - Eigenschaften, 48
- stöchiometrische Koeffizienten, 31
- Standardbasis, 67
- Stromkreise, 7
- Systemmatrix, 119
- Unterraum, *siehe* linearer Unterraum
- Vektor
 - Länge, 44
 - linear abhängig, 62
 - linear unabhängig, 62
 - lineare Hülle, 57
 - Norm, 45
 - Operationen, 41
 - Rechteseite-, 16
 - Skalierung, 46
 - Unbekannten-, 16
- Vektorfunktionen, 77
- Zahlengerade, 39