Per Jensen

VORLESUNGSSKRIPT MATHEMATIK B FÜR CHEMIKER

BERGISCHE UNIVERSITÄT WUPPERTAL
THEORETISCHE CHEMIE
April 2004

Inhaltsverzeichnis

1 Komplexe Zahlen

1.1 Einführung

Betrachten wir das Polynom

$$f(z) = z^2 + 4z + 8. \tag{1.1}$$

Wir möchten die Nullstellen dieses Polynoms bestimmen. Es gilt

$$z^2 + 4z + 8 = (z+2)^2 + 4. \tag{1.2}$$

Damit haben wir an einer Nullstelle die Bedingung

$$(z+2)^2 + 4 = 0. \tag{1.3}$$

Formal existieren also die Nullstellen

$$z_- = -2 - \sqrt{-4} \;\text{ und }\; z_+ = -2 + \sqrt{-4}, \tag{1.4}$$

oder

$$z_- = -2 - 2\sqrt{-1} \;\text{ und }\; z_+ = -2 + 2\sqrt{-1}. \tag{1.5}$$

Die Zahl $\sqrt{-1}$ erfüllt die Gleichung

$$\left(\sqrt{-1}\right)^2 = -1. \tag{1.6}$$

Eine willkürliche reelle Zahl $x \in \mathbb{R}$ hat die Eigenschaft $x^2 \geq 0$. Also ist $\sqrt{-1}$ keine reelle Zahl.

Wir definieren die *imaginäre Einheit*

$$i = \sqrt{-1} \tag{1.7}$$

und es gilt

$$i^2 = -1. \tag{1.8}$$

Die Nullstellen des Polynoms in Gln. (1.2) lassen sich damit als

$$z_- = -2 - 2\,i \;\text{ und }\; z_+ = -2 + 2\,i \tag{1.9}$$

ausdrücken, und wir erhalten

$$z^2 + 4z + 8 = (z - z_-)\,(z - z_+) = (z + 2 + 2\,i)\,(z + 2 - 2\,i). \tag{1.10}$$

Die Größen z_- und z_+ sind Beispiele für *komplexe Zahlen.* Allgemein läßt sich eine komplexe Zahl z darstellen als

$$z = a + i\,b \in \mathbb{C}, \quad \text{wobei } \; a, b \in \mathbb{R} \tag{1.11}$$

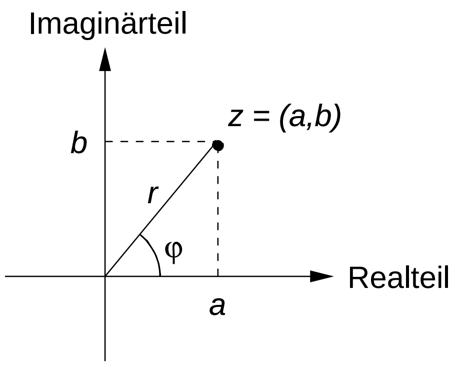

Abbildung 1.1: Die komplexe Zahlenebene.

und $\mathbb{C}$ die Menge aller komplexen Zahlen ist. Die Größe

$$a = \Re(z) = \Re(a + ib) \tag{1.12}$$

ist der *Realteil*, und

$$b = \Im(z) = \Im(a + ib) \tag{1.13}$$

der *Imaginärteil* von z.

Alternativ schreibt man komplexe Zahlen als geordnete 2-tupel der reellen Zahlen nach

$$z = (a, b). \tag{1.14}$$

Abbildung 1.1 zeigt die Darstellung der Zahl $z = (a, b)$ in der "komplexen Zahlenebene." Jeder Punkt in der komplexen Zahlenebene (das heißt, jede komplexe Zahl) ist auch durch den Abstand zum Ursprung r und den Winkel φ festgelegt (siehe Abb. 1.1). Es gilt

$$a = r \cdot \cos\varphi \quad \text{und} \quad b = r \cdot \sin\varphi. \tag{1.15}$$

Damit haben wir

$$z = a + i\,b = r\,\cos\varphi + i\,r\,\sin\varphi = r\,(\cos\varphi + i\,\sin\varphi) \tag{1.16}$$

mit $r = \sqrt{a^2 + b^2}$ und $\varphi = \arctan\left(\frac{b}{a}\right)$. Die Darstellung einer komplexen Zahl als $a+ib$ nennt man eine "kartesische Darstellung", die Darstellung als $r\,(\cos\varphi + i\,\sin\varphi)$ heißt "Polardarstellung."

Die zu $z = a + ib$ konjugiert komplexe Zahl bezeichnet man mit z^* und es gilt

$$z^* = a - ib. \tag{1.17}$$

In der Polardarstellung haben wir $z = r\,(\cos\varphi + i\,\sin\varphi)$ und damit

$$z^* = r\,(\cos\varphi - i\,\sin\varphi) = r\,(\cos(-\varphi) + i\,\sin(-\varphi)). \tag{1.18}$$

Abbildung 1.2 zeigt z und z^* in der komplexen Zahlenebene.

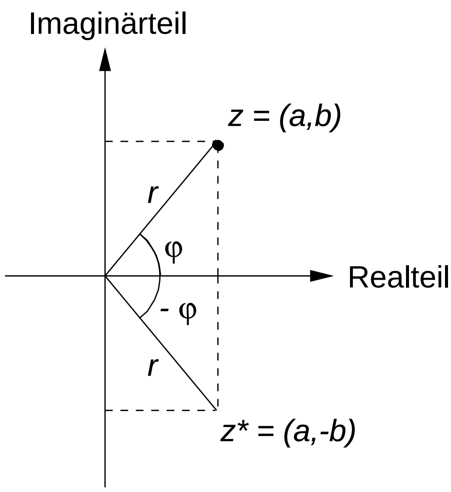

Abbildung 1.2: Konjugiert komplexe Zahlen.

1.2 Rechenregel

Wir definieren

$$z_1 = a_1 + i\, b_1 = r_1\, (\cos\varphi_1 + i\, \sin\varphi_1) \tag{1.19}$$

und

$$z_2 = a_2 + i\, b_2 = r_2\, (\cos\varphi_2 + i\, \sin\varphi_2). \tag{1.20}$$

Die zwei komplexen Zahlen z_1 und z_2 sind gleich, wenn

$$a_1 = a_2 \;\text{ und }\; b_1 = b_2 \tag{1.21}$$

sind, beziehungsweise wenn

$$r_1 = r_2 \;\text{ und }\; \varphi_1 = \varphi_2 + 2\, k\, \pi, \tag{1.22}$$

sind, wobei $k \in \mathbb{Z}$ eine beliebige ganze Zahl ist.

Komplexe Zahlen werden komponentenweise addiert, das heißt

$$z_1 + z_2 = (a_1 + a_2) + i\, (b_1 + b_2). \tag{1.23}$$

Es gilt $z_1 + z_2 = z_2 + z_1$.

Komplexe Zahlen werden multipliziert gemäß

$$\begin{aligned} z_1 \cdot z_2 &= (a_1 + i\, b_1) \cdot (a_2 + i\, b_2) \\ &= a_1 \cdot a_2 + a_1 \cdot i \cdot b_2 + i \cdot b_1 \cdot a_2 + i^2 \cdot b_1 \cdot b_2 \\ &= (a_1 \cdot a_2 - b_1 \cdot b_2) + i\, (a_1 \cdot b_2 + b_1 \cdot a_2). \end{aligned} \tag{1.24}$$

Es gilt $z_1 \cdot z_2 = z_2 \cdot z_1$ und $z_1 \cdot (z_2 + z_3) = z_1 \cdot z_2 + z_1 \cdot z_3$. Ein Sonderfall ist

$$z \cdot z^* = (a + i\,b) \cdot (a - i\,b) = a^2 + b^2 \in \mathbb{R}. \tag{1.25}$$

Der *Betrag* von z ist definiert als

$$|z| = \sqrt{z \cdot z^*} = \sqrt{a^2 + b^2}. \tag{1.26}$$

In der Polardarstellung haben wir

$$|z| = \sqrt{r^2\,\cos^2\varphi + r^2\,\sin^2\varphi} = r. \tag{1.27}$$

Zwei komplexe Zahlen werden dividiert, indem man mit dem konjugiert-komplexen des Nenners erweitert und dann die Multiplikationsregeln anwendet:

$$\begin{aligned} \frac{z_1}{z_2} &= \frac{a_1 + i\,b_1}{a_2 + i\,b_2} \\ &= \frac{(a_2 - i\,b_2) \cdot (a_1 + i\,b_1)}{(a_2 - i\,b_2) \cdot (a_2 + i\,b_2)} \\ &= \frac{a_1 \cdot a_2 + b_1 \cdot b_2}{a_2^2 + b_2^2} + i\,\frac{b_1 \cdot a_2 - a_1 \cdot b_2}{a_2^2 + b_2^2}. \end{aligned} \tag{1.28}$$

1.3 Exponentialdarstellung

Wir betrachten eine Taylorentwicklung der Exponentialfunktion

$$e^u = 1 + \frac{u}{1!} + \frac{u^2}{2!} + \frac{u^3}{3!} + \frac{u^4}{4!} + \frac{u^5}{5!} + \ldots. \tag{1.29}$$

Setzen wir hier $u = i\,\varphi$, erhalten wir

$$\begin{aligned} e^{i\,\varphi} &= 1 + i\,\frac{\varphi}{1!} - \frac{\varphi^2}{2!} - i\,\frac{\varphi^3}{3!} + \frac{\varphi^4}{4!} + i\,\frac{\varphi^5}{5!} - \ldots \\ &= \left(1 - \frac{\varphi^2}{2!} + \frac{\varphi^4}{4!} - \ldots\right) + i\,\left(\frac{\varphi}{1!} - \frac{\varphi^3}{3!} + \frac{\varphi^5}{5!} - \ldots\right) \end{aligned} \tag{1.30}$$

Wir erkennen den Realteil in Gln. (1.30) als die Taylorentwicklung von $\cos\varphi$ und den Imaginärteil als die Taylorentwicklung von $\sin\varphi$. Damit gilt

$$e^{i\,\varphi} = \cos\varphi + i\,\sin\varphi \tag{1.31}$$

und wir erhalten aus Gln. (1.16) die *Euler'sche Formel*

$$z = r\,(\cos\varphi + i\,\sin\varphi) = r\,e^{i\,\varphi} \tag{1.32}$$

oder, da $r = |z|$ ist,

$$z = |z|\, e^{i\,\varphi} = |z|\, \exp(i\,\varphi). \tag{1.33}$$

Der Betrag von $e^{i\,\varphi}$ ist

$$\begin{aligned} |e^{i\,\varphi}| &= \sqrt{e^{i\,\varphi}\,(e^{i\,\varphi})^*} \\ &= \sqrt{(\cos\varphi + i\,\sin\varphi)(\cos\varphi - i\,\sin\varphi)} = \sqrt{\cos^2\varphi + \sin^2\varphi} = 1. \end{aligned} \tag{1.34}$$

In der Euler'schen Darstellung ist es einfach, komplexe Zahlen zu multiplizieren und zu dividieren. Mit

$$z_1 = |z_1|\, e^{i\,\varphi_1} \;\; \text{und} \;\; z_2 = |z_2|\, e^{i\,\varphi_2} \tag{1.35}$$

berechnen wir das Produkt nach

$$z_1 \cdot z_2 = |z_1| \cdot |z_2| \cdot e^{i\,(\varphi_1+\varphi_2)}, \tag{1.36}$$

und die Division nach

$$\frac{z_1}{z_2} = \frac{|z_1|}{|z_2|} \cdot e^{i\,(\varphi_1-\varphi_2)}. \tag{1.37}$$

Potenzen können mittels der *Moivre-Formel* berechnet werden:

$$z^m = \left(|z|\, e^{i\,\varphi}\right)^m = |z|^m\, e^{i\,m\,\varphi}. \tag{1.38}$$

Für $k \in \mathbb{Z}$ gilt zum Beispiel

$$\begin{aligned} 1 &= e^{i\,0} = e^{i(0+2\,k\,\pi)} = \cos(2\,k\,\pi) + i\sin(2\,k\,\pi) \\ i &= e^{i\,\frac{\pi}{2}} = e^{i\left(\frac{\pi}{2}+2\,k\,\pi\right)} = \cos\left(\frac{\pi}{2} + 2\,k\,\pi\right) + i\sin\left(\frac{\pi}{2} + 2\,k\,\pi\right) \\ -1 &= e^{i\,\pi} = e^{i(\pi+2\,k\,\pi)} = \cos(\pi + 2\,k\,\pi) + i\sin(\pi + 2\,k\,\pi) \\ -i &= e^{i\,\frac{3\pi}{2}} = e^{i\left(\frac{3\pi}{2}+2\,k\,\pi\right)} = \cos\left(\frac{3\pi}{2} + 2\,k\,\pi\right) + i\sin\left(\frac{3\pi}{2} + 2\,k\,\pi\right). \end{aligned} \tag{1.39}$$

Die Zahl i^i berechnen wir als

$$i^i = \left(e^{i\,\frac{\pi}{2}}\right)^i = e^{i^2\,\frac{\pi}{2}} = e^{-\frac{\pi}{2}}, \tag{1.40}$$

das heißt, $i^i = e^{-\frac{\pi}{2}}$ ist eine reelle Zahl.

Weitere Beziehungen sind

$$e^{+i\varphi} = \cos\varphi + i\,\sin\varphi \tag{1.41}$$

$$e^{-i\varphi} = \cos\varphi - i\,\sin\varphi \tag{1.42}$$

$$\cos\varphi = \frac{1}{2}\left(e^{+i\varphi} + e^{-i\varphi}\right) \tag{1.43}$$

$$\sin\varphi = \frac{1}{2i}\left(e^{+i\varphi} - e^{-i\varphi}\right). \tag{1.44}$$

Die Gleichungen (1.43) und (1.44) können mit der Definitionsgleichungen der Funktionen cosh (hyperbolische Cosinusfunktion) und sinh (hyperbolische Sinusfunktion) verglichen werden:

$$\cosh x = \frac{1}{2}\left(e^{x} + e^{-x}\right) \tag{1.45}$$

und

$$\sinh x = \frac{1}{2}\left(e^{x} - e^{-x}\right). \tag{1.46}$$

Die Gleichung

$$z^m = 1 \tag{1.47}$$

besitzt für $m = 1, 2, 3, 4, \ldots$ und $z \in \mathbb{C}$ die m verschiedenen Lösungen

$$z_k = e^{i\,k\,\frac{2\pi}{m}}, \tag{1.48}$$

$k = 1, 2, 3, 4, \ldots, m$. Diese Zahlen liegen in der komplexen Zahlenebene auf dem Einheitskreis und bilden ein reguläres m-Eck.

Zum Beispiel haben wir für $m = 5$ (siehe Abb. 1.3)

$$\begin{aligned} z_1 &= e^{i\,\frac{2\pi}{5}}, \\ z_2 &= e^{i\,\frac{4\pi}{5}}, \\ z_3 &= e^{i\,\frac{6\pi}{5}} = e^{-i\,\frac{4\pi}{5}} = z_2^*, \\ z_4 &= e^{i\,\frac{8\pi}{5}} = e^{-i\,\frac{2\pi}{5}} = z_1^*, \\ z_5 &= e^{i\,\frac{10\pi}{5}} = e^{i\,2\,\pi} = 1. \end{aligned} \tag{1.49}$$

Die Lösungen der Gleichung

$$z^4 = 1 \tag{1.50}$$

sind

$$\begin{aligned} z_1 &= i \\ z_2 &= -1 \\ z_3 &= -i \\ z_4 &= 1. \end{aligned} \tag{1.51}$$

Wir definieren die Funktion

$$f(\varphi) = e^{i\,\varphi}. \tag{1.52}$$

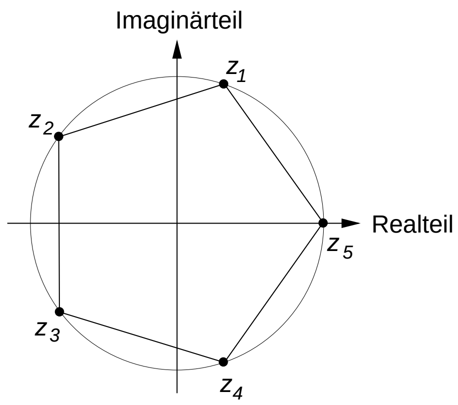

Abbildung 1.3: Die Lösungen z_1, z_2, z_3, z_4, z_5 (Gln. (1.49)) der Gleichung $z^5 = 1$, dargestellt in der komplexen Zahlenebene. Die Lösungen liegen alle auf dem Einheitskreis.

Wie berechnet man die Ableitung $f'(\varphi) = df/d\varphi$? Es gilt

$$\begin{aligned} \left(e^{i\,\varphi}\right)' &= (\cos\varphi + i\,\sin\varphi)' \\ &= -\sin\varphi + i\,\cos\varphi = i\,\left(\cos\varphi - \frac{1}{i}\,\sin\varphi\right) \\ &= i\,\left(\cos\varphi - \frac{i}{i^2}\,\sin\varphi\right) = i\,(\cos\varphi + i\,\sin\varphi) = i\,e^{i\,\varphi}. \end{aligned} \tag{1.53}$$

Später in der Vorlesung werden wir Differentialgleichungen behandeln, wie zum Beispiel

$$y'' + y = 0. \tag{1.54}$$

Hier ist $y(x)$ eine Funktion von x und $y'' = d^2y/dx^2$. Also erfüllt die Funktion die Gleichung

$$\frac{d^2y}{dx^2} = -y. \tag{1.55}$$

Wir können einfach nachprüfen, daß

$$y(x) = A\,e^{i\,x} + B\,e^{-i\,x} \tag{1.56}$$

eine Lösung ist, wobei A und B Konstanten sind.

2 Grundlagen der linearen Algebra

2.1 Problemstellungen

Als erstes Beispiel betrachten wir zwei Geraden in der (x, y)-Ebene:

$$\begin{aligned} y &= 2x + 1 \\ y &= -3x + 4 \end{aligned} \tag{2.1}$$

oder

$$2x - y = -1 \tag{2.2}$$
$$3x + y = 4. \tag{2.3}$$

Im Schnittpunkt der beiden Geraden sind die beiden Gleichungen (2.2) und (2.3) erfüllt. Wir suchen also ein Zahlenpaar (x, y), daß eine Lösung beider Gleichungen ist. Allgemein können wir solche Gleichungen als

$$ax + by = c \tag{2.4}$$
$$dx + ey = f \tag{2.5}$$

ausdrücken. Es kann sein, daß die beiden Geraden sich in einem Punkt schneiden. Dann haben die Gleichungen genau eine Lösung (x, y). Die beiden Geraden können aber auch parallel sein, dann existieren keine Lösungen. Schließlich können die beiden Geraden zusammenfallen und es gibt unendlich viele Lösungen. Wie bestimmen wir im Allgemeinen die Lösungen solcher *linearen Gleichungssysteme*? Wir schreiben sie als

$$\begin{pmatrix} a & b \\ d & e \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c \\ f \end{pmatrix}, \tag{2.6}$$

wobei die Konstruktionen

$$\begin{pmatrix} a & b \\ d & e \end{pmatrix}, \quad \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} c \\ f \end{pmatrix} \tag{2.7}$$

Matrizen genannt werden. Die Eigenschaften der beiden Matrizen

$$\begin{pmatrix} a & b \\ d & e \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} c \\ f \end{pmatrix} \tag{2.8}$$

entscheiden die Lösungsstruktur.

Als zweites Beispiel betrachten wir eine *Koordinatentransformation*. Wir haben zwei Koordinatensysteme xy und $x'y'$, wobei die Achsen x und x' einen Winkel von α bilden. Ein gegebener Punkt P hat die Koordinaten (x_P, y_P) im xy-System

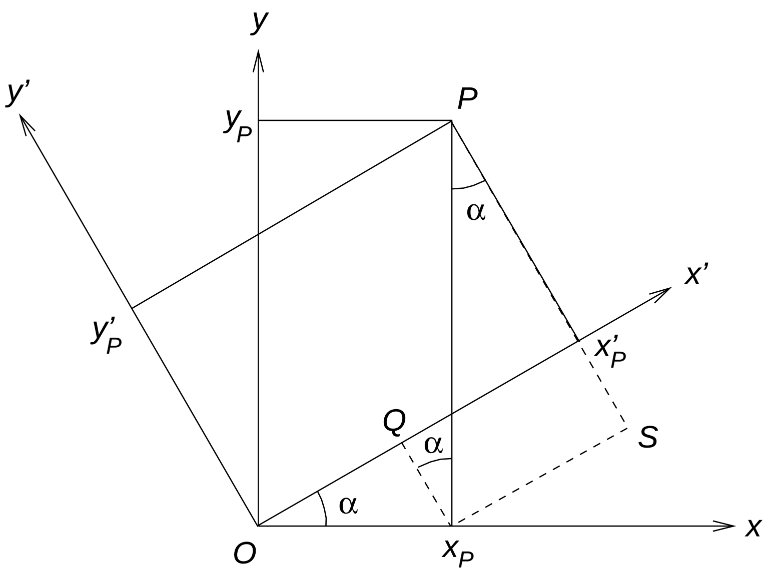

Abbildung 2.1: Die Koordinatentransformation.

und die Koordinaten (x'_P, y'_P) im $x'y'$-System. Abbildung 2.1 zeigt das Prinzip der Koordinatentransformation. Es gilt

$$|OQ| = x_P \cos\alpha. \tag{2.9}$$

Wir erhalten x'_P, indem wir $|OQ|$ zum Abstand zwischen Q und dem Punkt, wo PS die x'-Achse schneidet, addieren. Der letztere Abstand ist $y_P \sin\alpha$. Damit gilt

$$x'_P = x_P \cos\alpha + y_P \sin\alpha. \tag{2.10}$$

Ferner gilt

$$|PS| = y_P \cos\alpha. \tag{2.11}$$

Wir erhalten y'_P, indem wir von $|PS|$ den Abstand zwischen S und dem Punkt, wo PS die x'-Achse schneidet, subtrahieren. Der letztere Abstand ist $x_P \sin\alpha$. Damit gilt

$$y'_P = y_P \cos\alpha - x_P \sin\alpha. \tag{2.12}$$

Die beiden Gleichungen (2.10) und (2.12) können auch als eine Matrixgleichung ausgedruckt werden, als

$$\begin{pmatrix} x'_P \\ y'_P \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} x_P \\ y_P \end{pmatrix}. \tag{2.13}$$

2.2 Matrizen – Definition und Begriffe

Ein Rechteckschema mathematischer Größen (z. B. Zahlen, Funktionen oder Operatoren) aus n Zeilen und m Spalten nennt man eine *Matrix* der Dimension $n \times m$. Die $n \cdot m$ Einträge auf den einzelnen Plätzen dieses Schemas heißen *Matrixelemente*. Ein Beispiel für eine Matrix der Dimension $n \times m$ ist

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nm} \end{pmatrix}. \tag{2.14}$$

Das Element a_{ij} steht in der i'ten Zeile (i ist das *Zeilenindex*) und der j'ten Spalte (j ist das *Spaltenindex*). Die Elemente a_{ii} heißen *Diagonalelemente*, die Elemente a_{ij} mit $i \neq j$ heißen *Außerdiagonalelemente*.

A ist eine *reelle Matrix*, wenn alle $a_{ij} \in \mathbb{R}$ sind.

A ist eine *quadratische Matrix* der Dimension n, wenn $n = m$ ist, das heißt, wenn Zeilenzahl und Spaltenzahl übereinstimmen.

A ist eine *Nullmatrix*, wenn alle $a_{ij} = 0$ sind.

Die quadratische Matrix **A** ist *symmetrisch*, wenn gilt $a_{ij} = a_{ji}$. Ein Beispiel für eine symmetrische Matrix ist

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}. \tag{2.15}$$

Die quadratische Matrix **A** ist *antisymmetrisch* (oder *schiefsymmetrisch*), wenn gilt $a_{ij} = -a_{ji}$ $(i \neq j)$. Ein Beispiel für eine antisymmetrische Matrix ist

$$\begin{pmatrix} 1 & 2 & 3 \\ -2 & 4 & 5 \\ -3 & -5 & 6 \end{pmatrix}. \tag{2.16}$$

Die quadratische Matrix **A** ist eine *obere Dreiecksmatrix*, wenn gilt $a_{ij} = 0$ für $i > j$. Ein Beispiel für eine obere Dreiecksmatrix ist

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}. \tag{2.17}$$

Die quadratische Matrix **A** ist eine *untere Dreiecksmatrix*, wenn gilt $a_{ij} = 0$ für $i < j$. Ein Beispiel für eine untere Dreiecksmatrix ist

$$\begin{pmatrix} 1 & 0 & 0 \\ -2 & 4 & 0 \\ -3 & -5 & 6 \end{pmatrix}. \tag{2.18}$$

Die quadratische Matrix **A** ist eine *Diagonalmatrix*, wenn gilt $a_{ij} = 0$ für $i \neq j$. Ein Beispiel für eine Diagonalmatrix ist

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{pmatrix}. \tag{2.19}$$

Die quadratische Matrix **A** ist eine *Einheitsmatrix*, wenn gilt $a_{ij} = \delta_{ij}$. Das *Kroneckerdelta* ist definiert als

$$\delta_{ij} = \begin{cases} 1 & \text{wenn } \; i = j \; \text{ und} \\ 0 & \text{wenn } \; i \neq j \; \text{ ist.} \end{cases} \tag{2.20}$$

Ein Beispiel für eine Einheitsmatrix ist

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.21}$$

Eine Matrix der Dimension $n \times 1$ entspricht einen Spaltenvektor mit n Komponenten. Eine Matrix der Dimension $1 \times n$ entspricht einen Zeilenvektor mit n Komponenten.

2.3 Rechnen mit Matrizen

Mit **A** bezeichnen wir eine Matrix der Dimension $n_A \times m_A$ mit Elementen $a_{ij} \in \mathbb{R}$ und mit **B** bezeichnen wir eine Matrix der Dimension $n_B \times m_B$ mit Elementen $b_{ij} \in \mathbb{R}$. Die beiden Matrizen **A** und **B** sind *gleich*, wenn

1. $n_A = n_B$ und
2. $m_A = m_B$ und
3. $a_{ij} = b_{ij}$ für *alle* $i = 1, 2, 3, \ldots, n_A$ und $j = 1, 2, 3, \ldots, m_A$.

In diesem Fall schreiben wir

$$\mathbf{A} = \mathbf{B} \tag{2.22}$$

Zwei Matrizen **A** und **B** können addiert werden, wenn $n_A = n_B$ und $m_A = m_B$ sind. Wir schreiben

$$\mathbf{C} = \mathbf{A} + \mathbf{B}. \tag{2.23}$$

Das heißt, wir bezeichnen die Summe von **A** und **B** als **C**. **C** ist eine Matrix der Dimension $n_A \times m_A$ (oder $n_B \times m_B$) mit den Elementen

$$c_{ij} = a_{ij} + b_{ij}. \tag{2.24}$$

Es gilt
$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} \tag{2.25}$$
und
$$(\mathbf{A} + \mathbf{B}) + \mathbf{D} = \mathbf{A} + (\mathbf{B} + \mathbf{D}). \tag{2.26}$$
Ist $\mathbf{0}$ eine Nullmatrix der Dimension $n_A \times m_A$, so gilt
$$\mathbf{A} + \mathbf{0} = \mathbf{0} + \mathbf{A} = \mathbf{A}. \tag{2.27}$$

Mit $\widetilde{\mathbf{A}}$ bezeichen wir die Matrix
$$\widetilde{\mathbf{A}} = \begin{pmatrix} -a_{11} & -a_{12} & -a_{13} & \ldots & -a_{1m} \\ -a_{21} & -a_{22} & -a_{23} & \ldots & -a_{2m} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -a_{n1} & -a_{n2} & -a_{n3} & \ldots & -a_{nm} \end{pmatrix}. \tag{2.28}$$
Es gilt
$$\mathbf{A} + \widetilde{\mathbf{A}} = \widetilde{\mathbf{A}} + \mathbf{A} = \mathbf{0}. \tag{2.29}$$

Die Multiplikation einer Matrix $\mathbf{A}$ der Dimension $n_A \times m_A$ mit einem Skalar $\lambda \in \mathbb{R}$ ergibt eine Matrix
$$\mathbf{D} = \lambda\, \mathbf{A} \tag{2.30}$$
der Dimension $n_A \times m_A$; diese Matrix hat die Elemente
$$d_{ij} = \lambda\, a_{ij}. \tag{2.31}$$
Zum Beispiel gilt
$$3 \cdot \begin{pmatrix} 0 & 1 & 2 \\ -1 & 3 & 4 \end{pmatrix} = \begin{pmatrix} 0 & 3 & 6 \\ -3 & 9 & 12 \end{pmatrix}. \tag{2.32}$$
Die Matrix $\widetilde{\mathbf{A}}$ in Gln. (2.28) können wir als
$$\widetilde{\mathbf{A}} = (-1) \cdot \mathbf{A} = -\mathbf{A} \tag{2.33}$$
schreiben.

Zwei Matrizen $\mathbf{A}$ und $\mathbf{B}$ können multipliziert werden (das heißt, wir können die Produktmatrix $\mathbf{A} \cdot \mathbf{B}$ bilden), wenn gilt $m_A = n_B$. In diesem Falle ist $\mathbf{A} \cdot \mathbf{B}$ eine Matrix der Dimension $n_A \times m_B$. Für
$$\mathbf{C} = \mathbf{A} \cdot \mathbf{B} \tag{2.34}$$
gilt
$$c_{ij} = \sum_{k=1}^{m_A} a_{ik} \cdot b_{kj} \tag{2.35}$$

mit $i = 1, 2, 3, \ldots, n_A$ und $j = 1, 2, 3, \ldots, m_B$.

Wie betrachten als Beispiel die Matrizen

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \text{ und } \mathbf{B} = \begin{pmatrix} 5 & 0 \\ -1 & 3 \end{pmatrix}. \tag{2.36}$$

Hier gilt $n_A \times m_A = 3 \times 2$ und $n_B \times m_B = 2 \times 2$. Damit ist $m_A = n_B$ und die Matrix $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$ der Dimension $n_A \times m_B = 3 \times 2$ existiert. Ihre Elemente sind

$$\begin{aligned} c_{11} &= a_{11}\, b_{11} + a_{12}\, b_{21} = 1 \cdot 5 + 2 \cdot (-1) = 3 \\ c_{12} &= a_{11}\, b_{12} + a_{12}\, b_{22} = 1 \cdot 0 + 2 \cdot 3 = 6 \\ c_{21} &= a_{21}\, b_{11} + a_{22}\, b_{21} = 3 \cdot 5 + 4 \cdot (-1) = 11 \\ c_{22} &= a_{21}\, b_{12} + a_{22}\, b_{22} = 3 \cdot 0 + 4 \cdot 3 = 12 \\ c_{31} &= a_{31}\, b_{11} + a_{32}\, b_{21} = 5 \cdot 5 + 6 \cdot (-1) = 19 \\ c_{32} &= a_{31}\, b_{12} + a_{32}\, b_{22} = 5 \cdot 0 + 6 \cdot 3 = 18 \end{aligned} \tag{2.37}$$

und damit haben wir

$$\mathbf{C} = \begin{pmatrix} 3 & 6 \\ 11 & 12 \\ 19 & 18 \end{pmatrix} \tag{2.38}$$

Als weiteres Beispiel bilden wir das Produkt aus den folgenden Matrizen:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \text{ und } \mathbf{B} = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \tag{2.39}$$

Hier berechnen wir

$$\mathbf{A} \cdot \mathbf{B} = \begin{pmatrix} 2 & 1 \\ -2 & 0 \end{pmatrix} \text{ und } \mathbf{B} \cdot \mathbf{A} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}. \tag{2.40}$$

Das heißt,

$$\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A}. \tag{2.41}$$

Matrixmultiplikation ist also nicht kommutativ!

Für quadratische Matrizen $\mathbf{A}$, $\mathbf{B}$ und $\mathbf{C}$ derselben Dimension sind das Assoziativgesetz

$$\mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{C}) = (\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{C} \tag{2.42}$$

und das Distributivgesetz

$$\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C} \tag{2.43}$$

erfüllt.

Ist $\mathbf{A}$ eine Matrix der Dimension $n_A \times m_A$ und $\mathbf{E}$ eine Einheitsmatrix der Dimension $n_A \times n_A$, so gilt

$$\mathbf{E} \cdot \mathbf{A} = \mathbf{A}. \tag{2.44}$$

Zum Beispiel ist

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}. \tag{2.45}$$

Ist $\mathbf{E}$ eine Einheitsmatrix der Dimension $m_A \times m_A$, so gilt

$$\mathbf{A} \cdot \mathbf{E} = \mathbf{A}. \tag{2.46}$$

Beispielsweise ist

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}. \tag{2.47}$$

Gilt $m_A = n_A$, so daß $\mathbf{A}$ und $\mathbf{E}$ beide quadratisch sind, haben wir

$$\mathbf{A} \cdot \mathbf{E} = \mathbf{E} \cdot \mathbf{A} = \mathbf{A}. \tag{2.48}$$

Ein *Vektor* lässt sich als Zeilenmatrix

$$\vec{a} = \mathbf{a} = (a_1 \; a_2 \; a_3 \; \ldots a_n) \tag{2.49}$$

oder als Spaltenmatrix

$$\vec{b} = \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{pmatrix} \tag{2.50}$$

schreiben. Das Skalarprodukt $\vec{a} \cdot \vec{b}$ lässt sich dann als Matrixprodukt ausdrücken

$$\vec{a} \cdot \vec{b} = (a_1 \; a_2 \; a_3 \; \ldots a_n) \cdot \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{pmatrix} = \sum_{i=1}^{n} a_i \cdot b_i. \tag{2.51}$$

2.4 Quadratische Matrizen

Die Dimension

Die Dimension einer quadratischen Matrix $\mathbf{A}$ entspricht der Anzahl der Zeilen (die wiederum der Anzahl der Spalten gleicht):

$$\mathrm{Dim}\,(\mathbf{A}) = n. \tag{2.52}$$

Die Matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \tag{2.53}$$

hat die Dimension $\mathrm{Dim}\,(\mathbf{A}) = 3$.

Die Spur

Die Spur (Englisch: Trace) einer quadratischen Matrix der Dimension n ist die Summe der Diagonalelemente

$$\mathrm{Spur}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}. \tag{2.54}$$

Die Matrix $\mathbf{A}$ in Gln. (2.53) hat die Spur $1 + 5 + 9 = 15$.

Die Determinante

Einer quadratischen Matrix $\mathbf{A}$

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix} \tag{2.55}$$

aus n Zeilen und n Spalten kann man immer eine sogenannte Determinante n-ter Ordnung zuordnen. Man versteht darunter den mathematischen Ausdruck, der aus den Elementen der Matrix in folgender Weise gebildet wird: Man wählt zunächst aus jeder Zeile und aus jeder Spalte der Matrix jeweils genau ein Element, ordnet diese Elemente nach den Zeilen und bildet ihr Produkt:

$$a_{1j_1} a_{2j_2} \dots a_{nj_n}.$$

Die Größen $j_1 j_2 \dots j_n$ sind die jeweiligen Spaltenindizes. Da aus jeder Spalte genau ein Element gewählt wurde, bilden die Spaltenindizes $j_1 j_2 \dots j_n$ eine Permutation der Zahlen 1 bis n. Wenn es sich um eine gerade Permution handelt (das heißt, wenn man eine gerade Anzahl von Vertauschungen von je zwei Elementen braucht, um diese Permutation aus den Zahlen 1, 2, ..., n, das heißt, aus der *Referenzpermutation*, herzustellen), so fügen wir dem Produkt das positive Vorzeichen bei. Bei einer ungeraden Permutation (das heißt, wenn man eine ungerade Anzahl von Vertauschungen von je zwei Elementen braucht, um diese Permutation aus den Zahlen 1, 2, ..., n herzustellen) das negative Vorzeichen. Das kann man dadurch zum Ausdruck bringen, daß man vor das Produkt den Faktor $(-1)^{t(j_1 j_2 \ldots j_n)}$ setzt, wobei $t(j_1 j_2 \ldots j_n)$ die Zahl der Vertauschungen von je zwei Elementen angibt, die von der Zahlenfolge 1, 2, ..., n zur Folge $j_1 j_2 \ldots j_n$ führt. Eine Vertauschung von zwei Elementen wird auch eine *Transposition* genannt. Anschließend bilden wir die Summe über Produkte für alle $n!$ Permutationen der Spaltenindizes, was wir durch ein Summenzeichen ohne Angabe der Grenzen angeben wollen

$$\sum (-1)^{t(j_1 j_2 \ldots j_n)} a_{1j_1} a_{2j_2} \ldots a_{nj_n}.$$

Die so erhaltene Summe stellt die gesuchte Determinante D dar. Sie wird oft durch das Symbol $|\mathbf{A}|$ oder auch durch das gesamte Koffizientenschema, das zwischen zwei senkrechten Linien geschrieben wird, bezeichnet. Es gilt somit

$$D = |\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix} = \sum (-1)^{t(j_1 j_2 \ldots j_n)} a_{1j_1} a_{2j_2} \ldots a_{nj_n}. \tag{2.56}$$

Andere gebräuchliche Symbole für die Determinante einer Matrix $\mathbf{A}$ sind

$$D = \det(\mathbf{A}) = \det(a_{ij}) = |a_{ij}|\,. \tag{2.57}$$

Beispiele

a) Die Determinante einer Matrix der Dimension 1, also $\mathbf{A} = (a_{11})$ lautet $\det(\mathbf{A}) = a_{11}$.

b) Wir berechnen die zur Matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{2.58}$$

gehörende Determinante 2-ter Ordnung. Sie ist definiert mit Gln. (2.56) nach

$$D = (-1)^{t(12)}\, a_{11}\, a_{22} + (-1)^{t(21)}\, a_{12}\, a_{21}, \tag{2.59}$$

da die Spaltenindices 1 und 2 sich als 12 oder 21 permutieren lassen. Wir brauchen keine Transpositionen, um aus 12 die Permutation 12 herzustellen. Damit ist $t(12) = 0$. Wir brauchen eine Transposition, um aus 12 die Permutation 21 herzustellen: $t(21) = 1$. Folglich gilt

$$D = (-1)^0\, a_{11}\, a_{22} + (-1)^1\, a_{12}\, a_{21} = a_{11}\, a_{22} - a_{12}\, a_{21}. \tag{2.60}$$

c) Die $3! = 6$ möglichen Permutationen der drei Zahlen 1, 2, 3 sind 123, 132, 231, 213, 312 und 321. Damit ist die Determinante 3-ter Ordnung

$$D = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \tag{2.61}$$

als

$$\begin{aligned} D &= (-1)^{t(123)}\, a_{11}\, a_{22}\, a_{33} + (-1)^{t(132)}\, a_{11}\, a_{23}\, a_{32} \\ &+ (-1)^{t(231)}\, a_{12}\, a_{23}\, a_{31} + (-1)^{t(213)}\, a_{12}\, a_{21}\, a_{33} \\ &+ (-1)^{t(312)}\, a_{13}\, a_{21}\, a_{32} + (-1)^{t(321)}\, a_{13}\, a_{22}\, a_{31} \end{aligned} \tag{2.62}$$

definiert. In Tabelle 2.1 zählen wir für jede mögliche Permutation $j_1 j_2 j_3$ die Transpositionen, die erforderlich sind, um die fragliche Permutation aus den Zahlen 123 herzustellen, und bestimmen damit $t(j_1 j_2 j_3)$. Wir erhalten

$$\begin{aligned} D &= (-1)^0\, a_{11}\, a_{22}\, a_{33} + (-1)^1\, a_{11}\, a_{23}\, a_{32} \\ &+ (-1)^2\, a_{12}\, a_{23}\, a_{31} + (-1)^1\, a_{12}\, a_{21}\, a_{33} \\ &+ (-1)^2\, a_{13}\, a_{21}\, a_{32} + (-1)^1\, a_{13}\, a_{22}\, a_{31} \\ &= a_{11}\, a_{22}\, a_{33} - a_{11}\, a_{23}\, a_{32} + a_{12}\, a_{23}\, a_{31} \\ &- a_{12}\, a_{21}\, a_{33} + a_{13}\, a_{21}\, a_{32} - a_{13}\, a_{22}\, a_{31}. \end{aligned} \tag{2.63}$$

Die Unterdeterminante

Als *Unterdeterminante* D_{ij} $(n-1)$-ter Ordnung bezeichnet man die aus einer Determinante D n-ter Ordnung durch Entfernung der i-ten Zeile und j-ten Spalte hervorgehende Determinante. Haben wir zum Beispiel

$$D = |\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \tag{2.64}$$

Tabelle 2.1: Die Berechnung einer Determinante 3-ter Ordnung

$j_1 j_2 j_3$	Transpositionen	$t(j_1 j_2 j_3)$
123	123	0
132	123 $\rightarrow$ 132	1
231	123 $\rightarrow$ 213 $\rightarrow$ 231	2
213	123 $\rightarrow$ 213	1
312	123 $\rightarrow$ 321 $\rightarrow$ 312	2
321	123 $\rightarrow$ 321	1

so erhalten wir die Unterdeterminanten durch Streichen der 1. Zeile und 1. Spalte, bzw. 1. Zeile und 2. Spalte

$$D_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}, \quad D_{12} = \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}, \tag{2.65}$$

und so weiter.

Der Entwicklungsgesetz von Laplace

Jede Determinante läßt sich in der Form

$$D = \sum_{j=1}^{n} a_{ij} \, (-1)^{i+j} \, D_{ij} \tag{2.66}$$

nach den Elementen der i-ten Zeile entwickeln. Analogerweise kann nach den Elementen der k-ten Spalte entwickelt werden:

$$D = \sum_{i=1}^{n} a_{ik} \, (-1)^{i+k} \, D_{ik} \tag{2.67}$$

Die Wahl der Zeile oder Spalte ist beliebig. Es ist natürlich vorteilhaft, eine Zeile oder Spalte zu wählen, die viele Nullelemente enthält. Zum Beispiel sollte die Determinante

$$D = \begin{vmatrix} 1 & 4 & -2 \\ 0 & 1 & 3 \\ 4 & -5 & 1 \end{vmatrix} \tag{2.68}$$

nach der ersten Spalte oder nach der zweiten Zeile entwickelt werden. Entwicklung nach der ersten Spalte ergibt

$$\begin{aligned} D &= 1 \cdot (-1)^{1+1} \cdot \begin{vmatrix} 1 & 3 \\ -5 & 1 \end{vmatrix} + 0 \cdot (-1)^{2+1} \cdot \begin{vmatrix} 4 & -2 \\ -5 & 1 \end{vmatrix} + 4 \cdot (-1)^{3+1} \cdot \begin{vmatrix} 4 & -2 \\ 1 & 3 \end{vmatrix} \\ &= 1 \cdot (1 \cdot 1 - (-5) \cdot 3) + 4 \cdot (4 \cdot 3 - 1 \cdot (-2)) = 72. \end{aligned} \tag{2.69}$$

Entwicklung nach der zweiten Zeile ergibt

$$\begin{aligned} D &= 1 \cdot (-1)^{2+2} \cdot \begin{vmatrix} 1 & -2 \\ 4 & 1 \end{vmatrix} + 3 \cdot (-1)^{2+3} \cdot \begin{vmatrix} 1 & 4 \\ 4 & -5 \end{vmatrix} \\ &= 1 \cdot (1 \cdot 1 - 4 \cdot (-2)) - 3 \cdot (1 \cdot (-5) - 4 \cdot 4) = 72. \end{aligned} \tag{2.70}$$

Determinanten höherer Ordnung können durch Mehrfachanwendung des Entwicklungssatzes immer in eine Summe von Unterdeterminanten 2-ter Ordnung entwickelt werden.

Das Produkt

$$A_{ij} = (-1)^{i+j}\, D_{ij} \tag{2.71}$$

heißt das *algebraische Komplement* zum Matrixelement a_{ij}.

Die Eigenschaften der Determinante n-ter Ordnung

(a) Vertauschen wir zwei Zeilen (oder zwei Spalten) einer Determinante, so wechselt sie ihr Vorzeichen. Beispiel:

$$\begin{vmatrix} 1 & 3 \\ -4 & 5 \end{vmatrix} = 17 \tag{2.72}$$

aber

$$\begin{vmatrix} -4 & 5 \\ 1 & 3 \end{vmatrix} = \begin{vmatrix} 3 & 1 \\ 5 & -4 \end{vmatrix} = -17. \tag{2.73}$$

(b) Enthält eine Determinante eine Nullzeile (oder eine Nullspalte), so ist ihr Wert Null. Beispiel:

$$\begin{vmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 7 & 8 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 3 \\ 4 & 0 & 6 \\ 7 & 0 & 9 \end{vmatrix} = 0. \tag{2.74}$$

Der Beweis für diesen Satz kann über die Laplace'sche Entwicklung nach den Elementen der Nullzeile (oder Nullspalte) geführt werden.

(c) Der Wert einer Determinante ändert sich nicht, wenn wir zu einer Zeile ein Vielfaches einer *anderen* Zeile addieren, oder wenn wir zu einer Spalte ein Vielfaches einer *anderen* Spalte addieren. Beispiele:

$$\begin{vmatrix} 1 & 3 \\ -4 & 5 \end{vmatrix} = \begin{vmatrix} 1 & 3 \\ -3 & 8 \end{vmatrix} = \begin{vmatrix} 1 & 3 \\ -2 & 11 \end{vmatrix} = \begin{vmatrix} 1 & 4 \\ -4 & 1 \end{vmatrix} = 17. \tag{2.75}$$

(d) Enthält eine Determinante zwei gleiche Zeilen (oder zwei gleiche Spalten), so ist ihr Wert Null. Beispiel:

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 1 & 2 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 1 \\ 4 & 5 & 4 \\ 7 & 8 & 7 \end{vmatrix} = 0. \tag{2.76}$$

(e) Ist in einer Determinante eine Zeile als Linearkombination anderer Zeilen darstellbar, so ist der Wert der Determinante Null. Ist eine Spalte als Linearkombination anderer Spalten darstellbar, so ist der Wert der Determinante Null. Beispiel:

$$\begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 5 & 7 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 3 \\ 4 & 5 & 9 \\ 7 & 8 & 15 \end{vmatrix} = 0. \tag{2.77}$$

(f) Eine Determinante wird mit einer Zahl λ multipliziert, indem wir *alle* Elemente *einer* beliebigen Zeile oder Spalte mit λ multiplizieren. Beispiel:

$$3 \cdot \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = \begin{vmatrix} 3 & 6 \\ 3 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 2 \\ 9 & 12 \end{vmatrix} = \begin{vmatrix} 3 & 2 \\ 9 & 4 \end{vmatrix} = \begin{vmatrix} 1 & 6 \\ 3 & 12 \end{vmatrix} = -6. \tag{2.78}$$

(g) Die Determinante einer Dreiecksmatrix (oder eine Diagonalmatrix) ist gleich dem Produkt der Diagonalelemente mit

$$|\mathbf{A}| = a_{11}\, a_{22}\, a_{33} \ldots a_{nn} = \prod_{i=1}^{n} a_{ii}. \tag{2.79}$$

Durch Anwendung der Regel (c) können wir beliebige Determinanten in Determinanten einer Dreiecksmatrix umwandeln. In Gln. (2.68)-(2.70) berechneten wir durch Entwicklung nach Laplace

$$D = \begin{vmatrix} 1 & 4 & -2 \\ 0 & 1 & 3 \\ 4 & -5 & 1 \end{vmatrix} = 72. \tag{2.80}$$

Addieren wir das 4-fache der ersten Zeile zur dritten Zeile, so erhalten wir nach Regel (c):

$$D = \begin{vmatrix} 1 & 4 & -2 \\ 0 & 1 & 3 \\ 0 & -21 & 9 \end{vmatrix}. \tag{2.81}$$

In dieser neuen Determinante multiplizieren wir die zweite Zeile mit 21 und addieren das Ergebnis zur dritten Zeile:

$$D = \begin{vmatrix} 1 & 4 & -2 \\ 0 & 1 & 3 \\ 0 & 0 & 72 \end{vmatrix}. \tag{2.82}$$

Die Größe D kann jetzt als die Determinante einer Dreiecksmatrix berechnet werden und hat den Wert

$$D = 1 \cdot 1 \cdot 72 = 72. \tag{2.83}$$

(h) Die Determinante eines Produktes zweier quadratischen Matrizen $\mathbf{A}$ und $\mathbf{B}$ ist gleich dem Produkt der Determinanten der Einzelmatrizen:

$$\det(\mathbf{A} \cdot \mathbf{B}) = \det(\mathbf{A}) \cdot \det(\mathbf{B}). \tag{2.84}$$

(i) Die Einheitsmatrix $\mathbf{E}$ hat die Determinante

$$\det(\mathbf{E}) = 1. \tag{2.85}$$

(j) Die Transformationsmatrix

$$\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \tag{2.86}$$

besitzt die Determinante

$$\begin{vmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{vmatrix} = \cos^2\alpha + \sin^2\alpha = 1. \tag{2.87}$$

Rang

Der *Rang* einer Matrix $\mathbf{A}$, Rang$(\mathbf{A})$, ist gleich der Ordnung der größten darin enthaltenen und von Null verschiedenen Determinante. Gleichzeitig ist der Rang gleich der Maximalzahl linear unabhängiger Zeilen oder Spalten. Der Rang ist sowohl für quadratische als auch für nicht-quadratische Matrizen definiert.

Für eine quadratische Matrix $\mathbf{A}$ der Dimension $n \times n$ gilt offensichtlich

$$\text{Rang}\,(\mathbf{A}) \leq \text{Dim}\,(\mathbf{A}) = n. \tag{2.88}$$

Für eine nicht-quadratische Matrix $\mathbf{A}$ der Dimension $n_A \times m_A$ gilt

$$\text{Rang}\,(\mathbf{A}) \leq \min\,(n_A, m_A) \tag{2.89}$$

wobei $\min\,(n_A, m_A)$ gleich der kleinsten der beiden Zahlen n_A und m_A ist.

Der Rang einer Matrix bleibt unverändert, wenn wir

- zwei Zeilen (oder Spalten) vertauschen.
- zu einer Zeile ein Vielfaches einer *anderen* Zeile addieren, oder zu einer Spalte ein Vielfaches einer *anderen* Spalte addieren.

Betrachten wir zum Beispiel eine quadratische Matrix $\mathbf{A}$ der Dimension 4×4. Ist $\det(\mathbf{A}) \neq 0$, so haben wir $\text{Rang}\,(\mathbf{A}) = 4$. Ist dagegen $\det(\mathbf{A}) = 0$, müssen wir Unterdeterminanten 3-ter Ordnung untersuchen. Finden wir eine solche Unterdeterminante, die ungleich Null ist, gilt $\text{Rang}\,(\mathbf{A}) = 3$. Sind die Unterdeterminanten 3-ter Ordnung alle Null, müssen wir die Unterdeterminanten 2-ter Ordnung berechnen. Finden wir eine solche Unterdeterminante, die ungleich Null ist, gilt $\text{Rang}\,(\mathbf{A}) = 2$. Sind diese Unterdeterminanten auch alle Null, hat die Matrix, insofern sie nicht die Nullmatrix ist, den Rang 1.

Die Nullmatrix hat den Rang 0.

Eine Dreiecksmatrix (oder Diagonalmatrix) $\mathbf{A}$ der Dimension $n \times n$, für die alle Diagonalelemente $a_{ii} \neq 0$ sind, ist die Determinante mit $\prod_{i=1}^{n} a_{ii} \neq 0$ und die Matrix besitzt damit den Rang n.

Die Determinante der Matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 0 \\ -2 & 4 & 7 \\ -1 & 6 & 7 \end{pmatrix} \tag{2.90}$$

berechnet sich nach

$$\begin{aligned} |\mathbf{A}| &= 1 \cdot (-1)^{1+1} \cdot \begin{vmatrix} 4 & 7 \\ 6 & 7 \end{vmatrix} + 2 \cdot (-1)^{1+2} \cdot \begin{vmatrix} -2 & 7 \\ -1 & 7 \end{vmatrix} \\ &= 1 \cdot (-14) + 2 \cdot (-1) \cdot (-7) = 0, \end{aligned} \tag{2.91}$$

wenn wir nach der ersten Zeile entwickeln. Also ist $\text{Rang}\,(\mathbf{A}) < 3$. Die "obere rechte" Unterdeterminante

$$\begin{vmatrix} 2 & 0 \\ 4 & 7 \end{vmatrix} = 14 \tag{2.92}$$

ist aber ungleich Null, und damit ist Rang $(\mathbf{A}) = 2$.

Die Determinante

$$|\mathbf{A}| = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{vmatrix} = 0 \tag{2.93}$$

verschwindet. Mann kann einfach nachprüfen, daß alle neun Unterdeterminanten 2-ter Ordnung, zum Beispiel

$$\begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix} = 0, \tag{2.94}$$

auch gleich Null sind. Alle Unterdeterminanten erster Ordnung sind aber ungleich Null. Also ist Rang $(\mathbf{A}) = 1$. Multiplizieren wir in der Matrix $\mathbf{A}$ die erste Zeile mit 2, erhalten wir die zweite Zeile. Multiplikation der ersten Zeile mit 3 erzeugt die dritte Zeile, und durch Multiplikation der ersten Spalte mit 2 und 3 erzeugt man die zweite bzw. dritte Spalte. In der Matrix $\mathbf{A}$ ist die Maximalzahl linear unabhängiger Zeilen oder Spalten also gleich 1. In dieser Weise können wir auch erkennen, daß Rang $(\mathbf{A}) = 1$ ist.

2.5 Weitere Begriffe

Transposition

Unter der *transponierten Matrix* $\mathbf{A}^{\mathrm{T}}$ einer Matrix $\mathbf{A}$ der Dimension $n_A \times m_A$ verstehen wir eine Matrix der Dimension $m_A \times n_A$ mit den Elementen

$$(a^{\mathrm{T}})_{ij} = a_{ji}. \tag{2.95}$$

Eine quadratische Matrix ($m_A = n_A$) wird bei der Transposition an der Hauptdiagonalen gespiegelt; nur die Diagonalelemente a_{ii} bleiben auf ihren Plätzen.

Zum Beispiel erhalten wir aus

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} \tag{2.96}$$

die transponierte Matrix

$$\mathbf{A}^{\mathrm{T}} = \begin{pmatrix} 1 & 5 & 9 \\ 2 & 6 & 10 \\ 3 & 7 & 11 \\ 4 & 8 & 12 \end{pmatrix}. \tag{2.97}$$

Die Transposition eines Spaltenvektors

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix} \tag{2.98}$$

ergibt einen Zeilenvektor

$$\mathbf{a}^{\mathrm{T}} = (a_1\ a_2\ a_3\ \ldots a_n). \tag{2.99}$$

Die Transposition eines Zeilenvektors ergibt analogerweise einen Spaltenvektor.

Es gelten

$$(\mathbf{A}^{\mathrm{T}})^{\mathrm{T}} = \mathbf{A} \tag{2.100}$$

$$(\mathbf{A}+\mathbf{B})^{\mathrm{T}} = \mathbf{A}^{\mathrm{T}} + \mathbf{B}^{\mathrm{T}} \tag{2.101}$$

$$(\mathbf{A}\cdot\mathbf{B})^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}} \cdot \mathbf{A}^{\mathrm{T}}. \tag{2.102}$$

Für eine quadratische Matrix gilt

$$\det(\mathbf{A}^{\mathrm{T}}) = \det(\mathbf{A}). \tag{2.103}$$

Gilt

$$\mathbf{A}^{\mathrm{T}} = \mathbf{A} \tag{2.104}$$

sprechen wir von einer *symmetrischen* Matrix $\mathbf{A}$.

Für eine Matrix $\mathbf{A}$ mit komplexen Elementen definieren wir die *adjungierte Matrix* $\mathbf{A}^\dagger$ mit den Elementen

$$(a^\dagger)_{ij} = a^*_{ji}. \tag{2.105}$$

Das heißt

$$\mathbf{A}^\dagger = (\mathbf{A}^*)^{\mathrm{T}} = \left(\mathbf{A}^{\mathrm{T}}\right)^*. \tag{2.106}$$

Erfüllt $\mathbf{A}$ die Gleichung

$$\mathbf{A}^\dagger = \mathbf{A} \tag{2.107}$$

heißt sie *selbstadjungiert* oder *hermitesch.*

Die Matrix

$$\mathbf{A} = \begin{pmatrix} 1+2i & 3+4i \\ 5+6i & 7+8i \end{pmatrix} \tag{2.108}$$

hat die adjungierte Matrix

$$\mathbf{A}^\dagger = \begin{pmatrix} 1-2i & 5-6i \\ 3-4i & 7-8i \end{pmatrix}. \tag{2.109}$$

Die Matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 3+4i \\ 3-4i & 7 \end{pmatrix} \tag{2.110}$$

ist hermitesch. Alle symmetrischen Matrizen mit reellen Elementen sind hermitesch (warum?).

Inverse Matrizen

$\mathbf{A}$ ist eine Matrix der Dimension $n_A \times m_A$. Die Matrix $\mathbf{B}$ der Dimension $m_A \times n_A$ heißt *linksinvers* zu $\mathbf{A}$ wenn gilt

$$\mathbf{B} \cdot \mathbf{A} = \mathbf{E}, \tag{2.111}$$

wobei $\mathbf{E}$ eine Einheitsmatrix der Dimension $m_A \times m_A$ ist. Die Matrix $\mathbf{B}$ heißt *rechtsinvers* zu $\mathbf{A}$ wenn gilt

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{E}, \tag{2.112}$$

wobei $\mathbf{E}$ eine Einheitsmatrix der Dimension $n_A \times n_A$ ist.

Eine quadratische Matrix $\mathbf{A}$ der Dimension $n \times n$ heißt *regulär* (oder *nichtsingulär*), wenn es eine Matrix $\mathbf{A}^{-1}$ gibt, die sowohl linksinvers als auch rechtsinvers zu $\mathbf{A}$ ist. $\mathbf{A}^{-1}$ heißt die *Inverse* zu $\mathbf{A}$:

$$\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{E}. \tag{2.113}$$

Eine quadratische Matrix $\mathbf{A}$, die keine Inverse besitzt, heißt *singulär*. Nicht-quadratische Matrizen sind immer singulär.

Es gilt (Gln. (2.84) und (2.113))

$$\det(\mathbf{A}) \cdot \det\left(\mathbf{A}^{-1}\right) = \det(\mathbf{E}). \tag{2.114}$$

Ferner ist nach Gln. (2.79)

$$\det(\mathbf{E}) = 1 \tag{2.115}$$

und damit

$$\det\left(\mathbf{A}^{-1}\right) = \frac{1}{\det(\mathbf{A})}. \tag{2.116}$$

Hieraus folgt, daß $\mathbf{A}^{-1}$ nur dann existiert, wenn $\det(\mathbf{A}) \neq 0$ ist.

Für eine quadratische Matrix $\mathbf{A}$ mit $\det(\mathbf{A}) \neq 0$ können wir die Elemente der inversen Matrix durch die algebraischen Komplemente (Gln. (2.71)) ausdrücken. Es gilt

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{pmatrix} A_{11} & A_{21} & A_{31} & \dots & A_{n1} \\ A_{12} & A_{22} & A_{32} & \dots & A_{n2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ A_{1n} & A_{2n} & A_{3n} & \dots & A_{nn} \end{pmatrix}. \tag{2.117}$$

Hierbei ist auf die "transponierte" Indizierung zu achten. Berechnen wir mit dieser Definition ein Diagonalelement $(\mathbf{A}\mathbf{A}^{-1})_{ii}$ der Matrix $(\mathbf{A}\mathbf{A}^{-1})$, erhalten wir

$$\left(\mathbf{A}\mathbf{A}^{-1}\right)_{ii} = \frac{1}{\det(\mathbf{A})} \sum_{j=1}^{n} a_{ij} \, A_{ij} = 1, \tag{2.118}$$

wobei wir Gln. (2.66) und (2.71) benutzt haben. Ein Außerdiagonalelement dieser Matrix $(\mathbf{A}\mathbf{A}^{-1})_{ik}$ mit $k \neq i$ ist gegeben durch

$$\left(\mathbf{A}\mathbf{A}^{-1}\right)_{ik} = \frac{1}{\det(\mathbf{A})} \sum_{j=1}^{n} a_{ij} \, A_{kj}. \tag{2.119}$$

Nach Gln. (2.66) und (2.71) ist $(\mathbf{A}\mathbf{A}^{-1})_{ik}$ die Determinante einer Matrix, die aus der Matrix $\mathbf{A}$ dadurch hervorgegangen ist, daß wir die k-te Zeile durch die i-te Zeile ersetzt haben, ohne dabei die i-te Zeile zu verändern. Die resultierende Matrix hat zwei gleiche Zeilen und ihre Determinante verschwindet. Das heißt

$$\left(\mathbf{A}\mathbf{A}^{-1}\right)_{ik} = 0. \tag{2.120}$$

Wir haben jetzt gezeigt, daß

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{E} \tag{2.121}$$

ist. Mittels analoger Argumente können wir zeigen, daß

$$\mathbf{A}^{-1}\mathbf{A} = \mathbf{E} \tag{2.122}$$

ergibt.

Die Inverse der Matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{2.123}$$

mit der Determinante

$$D = a_{11} \, a_{22} - a_{12} \, a_{21} \neq 0 \tag{2.124}$$

berechnet sich nach Gln. (2.117) zu

$$\mathbf{A}^{-1} = \frac{1}{a_{11} \, a_{22} - a_{12} \, a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}. \tag{2.125}$$

Einige Rechenregeln:

$$\det\left(\mathbf{A}^{-1}\right) = \frac{1}{\det(\mathbf{A})}. \tag{2.126}$$

$$(\mathbf{A} \cdot \mathbf{B})^{-1} = \mathbf{B}^{-1} \cdot \mathbf{A}^{-1}. \tag{2.127}$$

$$\left(\mathbf{A}^{\mathrm{T}}\right)^{-1} = \left(\mathbf{A}^{-1}\right)^{\mathrm{T}}. \tag{2.128}$$

$$\mathbf{E}^{-1} = \mathbf{E}, \tag{2.129}$$

wobei $\mathbf{E}$ eine Einheitsmatrix ist.

Orthogonale Matrizen

$\mathbf{A}$ sei eine Matrix der Dimension $n_A \times m_A$. Eine reelle quadratische Matrix heißt dann *orthogonal*, wenn gilt

$$\mathbf{A}^{-1} = \mathbf{A}^{\mathrm{T}}. \tag{2.130}$$

Die Spalten einer orthogonalen Matrix sind orthogonal zueineinander. Daher gilt

$$\sum_{k=1}^{n} a_{ki}\, a_{kj} = \delta_{ij} \tag{2.131}$$

Entsprechend gilt für die Zeilen

$$\sum_{k=1}^{n} a_{ik}\, a_{jk} = \delta_{ij}. \tag{2.132}$$

Da für eine orthogonale Matrix $\mathbf{A}^{-1} = \mathbf{A}^{\mathrm{T}}$ ist, haben wir

$$\mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbf{E}, \tag{2.133}$$

wobei $\mathbf{E}$ eine Einheitsmatrix ist. Das heißt

$$\det\left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right) = \det\left(\mathbf{E}\right) = 1, \tag{2.134}$$

Die Gleichungen (2.84) und (2.103) ergeben

$$\det\left(\mathbf{A}^{\mathrm{T}}\right) \cdot \det\left(\mathbf{A}\right) = \left[\det\left(\mathbf{A}\right)\right]^2 = 1 \tag{2.135}$$

und wir folgern, daß

$$\det\left(\mathbf{A}\right) = \pm 1 \tag{2.136}$$

ist. Die Determinante einer orthogonalen Matrix ist also $+1$ oder -1.

Beispiele für orthogonale Matrizen sind die Transformationsmatrizen aus Gleichung. (2.13):

$$\mathbf{A} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}. \tag{2.137}$$

Für diese Matrizen erhalten wir

$$\mathbf{A}^{\mathrm{T}} \cdot \mathbf{A} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{2.138}$$

und

$$\mathbf{A} \cdot \mathbf{A}^{\mathrm{T}} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \cdot \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.139}$$

Folglich ist Gln. (2.130) erfüllt.

Ähnlichkeitstransformation

Zwischen den beiden n-komponentigen Vektoren

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} \quad \text{und} \quad \vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{pmatrix} \tag{2.140}$$

besteht der Zusammenhang

$$\vec{y} = \mathbf{A}\,\vec{x}, \tag{2.141}$$

wobei $\mathbf{A}$ eine quadratische Matrix der Dimension $n \times n$ ist.

Wir unterwerfen jetzt beide Vektoren $\vec{x}$ und $\vec{y}$ einer Transformation:

$$\vec{\mathcal{X}} = \mathbf{S}\,\vec{x}, \quad \text{und} \quad \vec{\mathcal{Y}} = \mathbf{S}\,\vec{y} \tag{2.142}$$

wobei $\mathbf{S}$ eine quadratische Transformationsmatrix der Dimension $n \times n$ ist.

Wie sieht nun der Zusammenhang zwischen $\vec{\mathcal{X}}$ und $\vec{\mathcal{Y}}$ aus? Gleichung (2.142) liefert

$$\vec{x} = \mathbf{S}^{-1}\,\vec{\mathcal{X}} \quad \text{und} \quad \vec{y} = \mathbf{S}^{-1}\,\vec{\mathcal{Y}}. \tag{2.143}$$

Eingesetzt in Gln. (2.141) ergibt

$$\mathbf{S}^{-1}\,\vec{\mathcal{Y}} = \mathbf{A}\,\mathbf{S}^{-1}\,\vec{\mathcal{X}}, \tag{2.144}$$

und damit

$$\vec{\mathcal{Y}} = \mathbf{S}\,\mathbf{A}\,\mathbf{S}^{-1}\,\vec{\mathcal{X}}. \tag{2.145}$$

Der Zusammenhang zwischen $\vec{\mathcal{X}}$ und $\vec{\mathcal{Y}}$ besteht also in

$$\vec{\mathcal{Y}} = \mathbf{A}' \, \vec{\mathcal{X}}, \tag{2.146}$$

wobei die neue Transformationsmatrix durch

$$\mathbf{A}' = \mathbf{S} \, \mathbf{A} \, \mathbf{S}^{-1} \tag{2.147}$$

gegeben ist.

Wir sagen, daß wir die Matrix $\mathbf{A}'$ durch eine *Ähnlichkeitstransformation* Zwischen den beiden n-komponentigen Vektoren der Matrix $\mathbf{A}$ erhalten haben.

Es gilt:

$$\begin{aligned} \det(\mathbf{A}') = \det\left(\mathbf{S}\,\mathbf{A}\,\mathbf{S}^{-1}\right) &= \det(\mathbf{S}) \, \det(\mathbf{A}) \, \det\left(\mathbf{S}^{-1}\right) \\ &= \det(\mathbf{S}) \, \det\left(\mathbf{S}^{-1}\right) \, \det(\mathbf{A}) \end{aligned} \tag{2.148}$$

Da wir mit

$$\mathbf{S}\,\mathbf{S}^{-1} = \mathbf{E} \tag{2.149}$$

eine Einheitsmatrix der Dimension $n \times n$ erhalten, gilt

$$\det\left(\mathbf{S}\,\mathbf{S}^{-1}\right) = \det(\mathbf{S}) \, \det\left(\mathbf{S}^{-1}\right) = \det(\mathbf{E}) = 1. \tag{2.150}$$

Setzen wir dieses Ergebnis in Gln. (2.148) ein, so können wir schreiben

$$\det(\mathbf{A}') = \det\left(\mathbf{S}\,\mathbf{A}\,\mathbf{S}^{-1}\right) = \det(\mathbf{A}) \tag{2.151}$$

Ferner gilt:

$$\mathrm{Spur}(\mathbf{A}') = \mathrm{Spur}\left(\mathbf{S}\,\mathbf{A}\,\mathbf{S}^{-1}\right) = \mathrm{Spur}(\mathbf{A}) \tag{2.152}$$

Diese Behauptung können wir wie folgt beweisen: Wir bezeichnen allgemein das Element in der i-ten Zeile und j-ten Spalte einer Matrix $\mathbf{X}$ (wobei im vorliegenden Fall $\mathbf{X} = \mathbf{A}$, $\mathbf{A}'$, $\mathbf{S}$, $\mathbf{S}^{-1}$, ... ist) als $(\mathbf{X})_{ij}$. In dieser Notation haben wir

$$\begin{aligned} \mathrm{Spur}\left(\mathbf{S}\,\mathbf{A}\,\mathbf{S}^{-1}\right) &= \sum_{i=1}^{n} \left(\mathbf{S}\,\mathbf{A}\,\mathbf{S}^{-1}\right)_{ii} \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} (\mathbf{S}\,\mathbf{A})_{ij} \, \left(\mathbf{S}^{-1}\right)_{ji} \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} (\mathbf{S})_{ik} \, (\mathbf{A})_{kj} \, \left(\mathbf{S}^{-1}\right)_{ji} \\ &= \sum_{j=1}^{n} \sum_{k=1}^{n} (\mathbf{A})_{kj} \left[\sum_{i=1}^{n} \left(\mathbf{S}^{-1}\right)_{ji} (\mathbf{S})_{ik} \right] \end{aligned} \tag{2.153}$$

Nach den Regeln der Matrixmultiplikation ist aber

$$\sum_{i=1}^{n} \left(\mathbf{S}^{-1}\right)_{ji} (\mathbf{S})_{ik} = \left(\mathbf{S}^{-1}\,\mathbf{S}\right)_{jk} = (\mathbf{E})_{jk} = \delta_{jk}, \tag{2.154}$$

wobei wir benutzt haben, daß

$$\mathbf{S}^{-1}\,\mathbf{S} = \mathbf{E} \tag{2.155}$$

ist. Damit erhalten wir

$$\mathrm{Spur}\left(\mathbf{S}\,\mathbf{A}\,\mathbf{S}^{-1}\right) = \sum_{j=1}^{n}\sum_{k=1}^{n} (\mathbf{A})_{kj}\ \delta_{jk} = \sum_{j=1}^{n} (\mathbf{A})_{jj} = \mathrm{Spur}\,(\mathbf{A})\,. \tag{2.156}$$

Betrachten wir als Beispiel für Gln. (2.141)

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}; \tag{2.157}$$

der Vektor $\vec{y}$ entsteht hier durch eine "Streckung" um einen Faktor 3 entlang der x_2-Richtung. Transformieren wir jetzt das Koordinatensystem, in dem die Vektoren $\vec{x}$ und $\vec{y}$ ausgedrückt werden, mit der Transformationsmatrix

$$\mathbf{S} = \begin{pmatrix} \cos\frac{\pi}{2} & \sin\frac{\pi}{2} \\ -\sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{2.158}$$

Diese Matrix beschreibt eine Drehung des Koordinatensystems um 90° (Gln. (2.13)). Da $\mathbf{S}$ orthogonal ist, gilt

$$\mathbf{S}^{-1} = \mathbf{S}^{\mathrm{T}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{2.159}$$

Wir berechnen daher nach Gln. (2.147)

$$\begin{aligned} \mathbf{A}' &= \mathbf{S}\,\mathbf{A}\,\mathbf{S}^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 3 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{2.160}$$

Im vorliegenden Fall wird also Gln. (2.146)

$$\begin{pmatrix} \mathcal{Y}_1 \\ \mathcal{Y}_2 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathcal{X}_1 \\ \mathcal{X}_2 \end{pmatrix}; \tag{2.161}$$

Der Vektor $\vec{y}$ entstand durch eine "Streckung" des Vektors $\vec{x}$ um einen Faktor 3 entlang der x_2-Richtung. Nachdem wir das Koordinatensystem um 90° gedreht haben, liegt die neue $\mathcal{X}_1$-Achse dort, wo die alte x_2-Achse früher lag und die neue $\mathcal{X}_2$-Achse liegt dort, wo die alte x_1-Achse früher lag. Folglich entsteht im neuen Koordinatensystem der Vektor $\vec{\mathcal{Y}}$ durch eine "Streckung" des Vektors $\vec{\mathcal{X}}$ um einen Faktor 3 entlang der $\mathcal{X}_1$-Richtung.

Prüfen Sie nach, daß die Matrizen $\mathbf{A}$ und $\mathbf{A}'$ des obigen Beispiels die Gleichungen (2.151) und (2.152) erfüllen.

2.6 Lineare Gleichungssysteme

Am Anfang dieses Kapitels diskutierten wir die zwei Gleichungen (2.2) und (2.3)

$$\begin{aligned} 2x - y &= -1 && (2.162)\\ 3x + y &= 4, && (2.163) \end{aligned}$$

mit den beiden Unbekannten x und y. Diese Gleichungen lassen sich in Matrixform schreiben:

$$\begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 \\ 4 \end{pmatrix} \tag{2.164}$$

Wir möchten jetzt solche *Gleichungssysteme* analysieren und betrachten n lineare Gleichungen mit den n Unbekannten x_1, x_2, x_3, $\ldots$, x_n:

$$\begin{aligned} a_{11}\,x_1 + a_{12}\,x_2 + \ldots + a_{1n}\,x_n &= b_1 \\ a_{21}\,x_1 + a_{22}\,x_2 + \ldots + a_{2n}\,x_n &= b_2 \\ &\vdots \\ a_{n1}\,x_1 + a_{n2}\,x_2 + \ldots + a_{nn}\,x_n &= b_n. \end{aligned} \tag{2.165}$$

Andere Schreibweisen für das Gleichungssystem sind

$$\sum_{k=1}^{n} a_{ik}\,x_k = b_i, \quad i = 1, 2, 3, \ldots, n, \tag{2.166}$$

oder

$$\mathbf{A}\,\vec{x} = \vec{b}, \tag{2.167}$$

wobei

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \ldots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \ldots & a_{nn} \end{pmatrix}, \quad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} \quad \text{und} \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{pmatrix} \tag{2.168}$$

ist.

Falls $b_1 = b_2 = b_3 = \ldots = 0$ ist, sprechen wir von einem *homogenen*, ansonsten von einem *inhomogenen* Gleichungssystem.

Die Eigenschaften linearer Gleichungssysteme

(a) Das lineare Gleichungssystem $\mathbf{A}\,\vec{x} = \vec{b}$ ist dann und nur dann lösbar, wenn der Rang der Koeffizientenmatrix $\mathbf{A}$ gleich dem Rang der erweiterten Koeffizientenmatrix

$$\mathbf{Ab} = \left(\begin{array}{ccccc|c} a_{11} & a_{12} & a_{13} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} & b_2 \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} & b_3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} & b_n \end{array}\right) \quad (2.169)$$

der Dimension $n \times (n+1)$ ist:

$$\mathrm{Rang}\,(\mathbf{A}) = \mathrm{Rang}\,(\mathbf{A}|\mathbf{b}). \quad (2.170)$$

(b) Ein homogenes lineares Gleichungssystem $\mathbf{A}\,\vec{x} = \vec{0}$ (wobei $\vec{0}$ der Nullvektor ist) ist somit immer lösbar; es hat die triviale Lösung $\vec{x} = \vec{0}$.

(c) Die Lösungsmannigfaltigkeit eines lösbaren Systems ergibt sich aus

$$K = \mathrm{Dim}\,(\mathbf{A}) - \mathrm{Rang}\,(\mathbf{A}) = n - \mathrm{Rang}\,(\mathbf{A}) \quad (2.171)$$

wie folgt ($\vec{x}_0$, $\vec{x}_1$ und $\vec{x}_2$ sind n-Komponenten-Vektoren):

$K = 0$: Es gibt genau eine Lösung $\vec{x} = \vec{x}_0$,

$K = 1$: Die Lösungen bilden eine *einfach unendliche Menge* von Vektoren $\vec{x} = \vec{x}_0 + \lambda\,\vec{x}_1$, wobei $\lambda \in \mathbb{R}$ willkürlich ist,

$K = 2$: Die Lösungen bilden eine *zweifach unendliche Menge* von Vektoren $\vec{x} = \vec{x}_0 + \lambda_1\,\vec{x}_1 + \lambda_2\,\vec{x}_2$, wobei λ_1, $\lambda_2 \in \mathbb{R}$ willkürlich sind,

und so weiter.

(d) Ein homogenes lineares Gleichungssystem $\mathbf{A}\,\vec{x} = \vec{0}$ hat demnach außer der trivialen Lösung $\vec{x} = \vec{0}$ noch weitere Lösungen, wenn $\det(\mathbf{A}) = 0$ ist. In diesem Fall ist nämlich $\mathrm{Rang}\,(\mathbf{A}) < n$ und damit $K > 0$.

Das Gleichungssystem

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 1 \\ 1 & -2 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (2.172)$$

hat die Koeffizientenmatrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 1 \\ 1 & -2 & -3 \end{pmatrix}. \quad (2.173)$$

Wir entwickeln det(**A**) nach der ersten Spalte und erhalten

$$\det(\mathbf{A}) = \begin{vmatrix} 3 & 1 \\ -2 & -3 \end{vmatrix} + \begin{vmatrix} 0 & 2 \\ 3 & 1 \end{vmatrix} = -7 - 6 = -13. \qquad (2.174)$$

Das Gleichungssystem hat also nur die triviale Lösung $\vec{x} = \vec{0}$. Wir können dieses nachprüfen, indem wir das Gleichungssystem wie folgt aufschreiben:

$$x_1 + 2\,x_3 = 0 \qquad (2.175)$$
$$3x_2 + x_3 = 0 \qquad (2.176)$$
$$x_1 - 2\,x_2 - 3\,x_3 = 0 \qquad (2.177)$$

Wir subtrahieren Gln. (2.175) von Gln. (2.177) und erhalten damit ein äquivalentes Gleichungssystem

$$x_1 + 2\,x_3 = 0 \qquad (2.178)$$
$$3x_2 + x_3 = 0 \qquad (2.179)$$
$$-2\,x_2 - 5\,x_3 = 0. \qquad (2.180)$$

Multplikation der Gln. (2.179) mit 2/3 und Addition zu Gln. (2.180) ergeben

$$x_1 + 2\,x_3 = 0 \qquad (2.181)$$
$$3x_2 + x_3 = 0 \qquad (2.182)$$
$$-\frac{13}{3}\,x_3 = 0. \qquad (2.183)$$

Gleichung (2.183) hat die Lösung $x_3 = 0$ und Einsetzen in Gln. (2.181) und (2.182) liefert $x_1 = x_2 = 0$. Wir haben also bestätigt, daß das Gleichungssystem in Gln. (2.175)-(2.177) nur die triviale Lösung $\vec{x} = \vec{0}$ besitzt.

Als weiteres Beispiel betrachten wir das Gleichungssystem

$$\begin{pmatrix} 2 & -1 & 5 \\ -1 & 1 & -2 \\ 2 & 0 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \qquad (2.184)$$

mit der Koeffizientenmatrix

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 5 \\ -1 & 1 & -2 \\ 2 & 0 & 6 \end{pmatrix}. \qquad (2.185)$$

Wir entwickeln det(**A**) nach der dritten Zeile und erhalten

$$\det(\mathbf{A}) = 2 \cdot \begin{vmatrix} -1 & 5 \\ 1 & -2 \end{vmatrix} + 6 \cdot \begin{vmatrix} 2 & -1 \\ -1 & 1 \end{vmatrix} = 2 \cdot (-3) + 6 \cdot 1 = 0. \qquad (2.186)$$

Damit ist auch Rang $(\mathbf{A})$ kleiner Dim $(\mathbf{A})$. Wir finden eine Unterdeterminante zweiter Ordnung verschieden von Null, beispielsweise

$$\begin{vmatrix} -1 & 5 \\ 1 & -2 \end{vmatrix} = -3 \neq 0, \tag{2.187}$$

so daß Rang $(\mathbf{A}) = 2$ ist. Also ist $K = \text{Dim}\,(\mathbf{A}) - \text{Rang}\,(\mathbf{A}) = 3 - 2 = 1$, und die Lösungen des Gleichungssystems können als $\vec{x} = \vec{x}_0 + \lambda\,\vec{x}_1$ aufgeschrieben werden. Wir können dieses Ergebnis bestätigen durch Umformung des Gleichungssystems. Wir fangen an mit

$$\begin{aligned} 2\,x_1 - x_2 + 5\,x_3 &= 0 && (2.188)\\ -x_1 + x_2 - 2\,x_3 &= 0 && (2.189)\\ 2\,x_1 + 6\,x_3 &= 0. && (2.190) \end{aligned}$$

Multiplikation der Gln. (2.189) mit 2 und Addition zu den Gleichungen (2.188) und (2.189) erzeugen das Gleichungssystem

$$\begin{aligned} x_2 + x_3 &= 0 && (2.191)\\ -x_1 + x_2 - 2\,x_3 &= 0 && (2.192)\\ 2\,x_2 + 2\,x_3 &= 0. && (2.193) \end{aligned}$$

Wir multiplizieren Gln. (2.191) mit -2 und addieren das Ergebnis zu Gln. (2.193):

$$\begin{aligned} x_2 + x_3 &= 0 && (2.194)\\ -x_1 + x_2 - 2\,x_3 &= 0 && (2.195)\\ 0 &= 0. && (2.196) \end{aligned}$$

Gleichung (2.196) ist trivialerweise erfüllt. Setzen wir $x_3 = \lambda$, so erhalten wir durch Einsetzen in Gln. (2.194) $x_2 = -\lambda$ und durch weiteres Einsetzen in Gln. (2.195) $x_1 = -3\,\lambda$. Die Lösungen des Gleichungssystems lauten damit wie erwartet:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda \cdot \begin{pmatrix} -3 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \lambda \cdot \begin{pmatrix} -3 \\ -1 \\ 1 \end{pmatrix} \tag{2.197}$$

(e) Die Lösungsmenge eines linearen Gleichungssystems ändert sich nicht bei sogenannten *elementaren Zeilenoperationen.* In einer elementaren Zeilenoperation wird eine der n Gleichungen, zum Beispiel die mit dem Index k,

$$a_{k1}\,x_1 + a_{k2}\,x_2 + \ldots + a_{kn}\,x_n = b_k, \tag{2.198}$$

mit einer Zahl λ multpliziert,

$$\lambda\, a_{k1}\, x_1 + \lambda\, a_{k2}\, x_2 + \ldots + \lambda\, a_{kn}\, x_n = \lambda\, b_k, \tag{2.199}$$

und das Ergebnis wird zu einer anderen Gleichung (zum Beispiel mit Index p) addiert:

$$(a_{p1} + \lambda\, a_{k1})\; x_1 + (a_{p2} + \lambda\, a_{k2})\; x_2 + \ldots + (a_{pn} + \lambda\, a_{kn})\; x_n = b_p + \lambda\, b_k \tag{2.200}$$

Unter Punkt (d) haben wir bereits mehrere solcher Zeilenoperationen durchgeführt.

(f) Ist $\vec{x}_0$ *ein* spezieller Lösungsvektor (eine sogenannte *partikuläre Lösung*) des lösbaren inhomogenen Gleichungssystems $\mathbf{A}\,\vec{x} = \vec{b}$, so erhält man alle Lösungen $\vec{x}$ dieses Systems, wenn wir zu $\vec{x}_0$ die Lösungen des entsprechenden homogenen Systems $\mathbf{A}\,\vec{x}_\mathrm{h} = \vec{0}$ addieren:

$$\vec{x} = \vec{x}_0 + \vec{x}_\mathrm{h} \tag{2.201}$$

(g) Das inhomogene lineare Gleichungssystem $\mathbf{A}\,\vec{x} = \vec{b}$ besitzt genau dann eine eindeutige Lösung, wenn $\det \mathbf{A} \neq 0$ ist. In diesem Fall hat das entsprechende homogene Gleichungssystem $\mathbf{A}\,\vec{x} = \vec{0}$ die eindeutige Lösung $\vec{x} = \vec{0}$.

Das Gleichungssystem

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 1 \\ 1 & -2 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \tag{2.202}$$

hat die in Gln. (2.173) gegebene Koeffizientenmatrix $\mathbf{A}$ mit der Determinante $\det(\mathbf{A}) = -13$. Wir schreiben das Gleichungssystem als

$$\begin{aligned} x_1 + 2\,x_3 &= 1 && (2.203)\\ 3x_2 + x_3 &= 0 && (2.204)\\ x_1 - 2\,x_2 - 3\,x_3 &= 2 && (2.205) \end{aligned}$$

schreiben. Wir subtrahieren Gln. (2.203) von Gln. (2.205) und erhalten damit das äquivalente Gleichungssystem

$$\begin{aligned} x_1 + 2\,x_3 &= 1 && (2.206)\\ 3x_2 + x_3 &= 0 && (2.207)\\ -2\,x_2 - 5\,x_3 &= 1. && (2.208) \end{aligned}$$

Multplikation der Gln. (2.207) mit 2/3 und Addition zu Gln. (2.208) ergeben

$$\begin{aligned} x_1 + 2\,x_3 &= 1 \qquad (2.209)\\ 3x_2 + x_3 &= 0 \qquad (2.210)\\ -\frac{13}{3}\,x_3 &= 1. \qquad (2.211) \end{aligned}$$

Gleichung (2.211) hat die Lösung $x_3 = -3/13$. Einsetzen in Gln. (2.210) liefert nun $x_2 = 1/13$, und weiteres Einsetzen in Gln. (2.209) ergibt $x_1 = 19/13$. Die Lösung des Gleichungssystems ist damit

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \frac{1}{13} \begin{pmatrix} 19 \\ 1 \\ -3 \end{pmatrix}. \qquad (2.212)$$

Als eine bequeme Kurzschreibweise für ein Gleichungssystem, zum Beispiel

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \\ 2 \end{pmatrix}, \qquad (2.213)$$

genügt es, die erweiterte Koeffizientenmatrix nach Gln. (2.169), hier

$$\mathbf{Ab} = \begin{pmatrix} 1 & 1 & 1 & 4 \\ 1 & 2 & -1 & 5 \\ 1 & -1 & 5 & 2 \end{pmatrix}, \qquad (2.214)$$

aufzuschreiben. Wir können jetzt elementare Zeilenoperationen in dieser Matrix ausführen und erhalten zunächst

$$\begin{pmatrix} 1 & 1 & 1 & 4 \\ 0 & 1 & -2 & 1 \\ 0 & -2 & 4 & -2 \end{pmatrix}, \qquad (2.215)$$

und dann

$$\begin{pmatrix} 1 & 1 & 1 & 4 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \qquad (2.216)$$

Die dritte Gleichung des transformierten Systems ist trivialerweise erfüllt. Die beiden verbleibenden Gleichungen sind

$$\begin{aligned} x_1 + x_2 + x_3 &= 4 \qquad (2.217)\\ x_2 - 2\,x_3 &= 1. \qquad (2.218) \end{aligned}$$

Wählen wir hier $x_3 = \lambda$, so erhalten wir

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 & - & 3\lambda \\ 1 & + & 2\lambda \\ & & \lambda \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} + \lambda \cdot \begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix}. \tag{2.219}$$

Das definierte Gleichungssystem

$$\mathbf{Ab} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & -1 & 0 \\ 1 & -1 & 5 & 0 \end{pmatrix}, \tag{2.220}$$

hat dieselbe Koeffizientenmatrix wie das System in Gln. (2.214) aber einen anderen $\vec{b}$-Vektor. Die in Gln. (2.215) und (2.216) durchgeführten Zeilenoperationen ergeben in diesem Fall

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{2.221}$$

bzw.

$$x_1 + x_2 + x_3 = 0 \tag{2.222}$$

$$x_2 - 2\,x_3 = 0. \tag{2.223}$$

Mit $x_3 = \lambda$ ermitteln wir hier

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda \cdot \begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix}. \tag{2.224}$$

Wir sehen, daß die in Gln. (2.219) vollständige Lösung des inhomogenen Gleichungssystems in Gln. (2.213) sich als eine Summe zweier Terme schreiben läßt. Der erste Term

$$\vec{x}_0 = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}, \tag{2.225}$$

ist eine partikuläre Lösung des inhomogenen Gleichungssystems, und der zweite Term

$$\vec{x}_\mathrm{h} = \lambda \cdot \begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix}. \tag{2.226}$$

ist die vollständige Lösung des homogenen Gleichungssystems in Gln. (2.220).

Als Beispiel für ein drittes Gleichungssystem mit derselben Koeffizientenmatrix $\mathbf{A}$ betrachten wir

$$\mathbf{Ab} = \begin{pmatrix} 1 & 1 & 1 & 4 \\ 1 & 2 & -1 & 5 \\ 1 & -1 & 5 & 3 \end{pmatrix}, \tag{2.227}$$

so lässt sich durch die Zeilenoperationen der Gln. (2.215) und (2.216) in

$$\begin{pmatrix} 1 & 1 & 1 & 4 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{2.228}$$

umwandeln. Dieses Gleichungssystem ist **nicht** lösbar, da die dritte Gleichung einen Widerspruch beinhaltet. Die erweiterte Koeffizientenmatrix $\mathbf{Ab}$ in Gln. (2.227) hat den Rang 3, da die Unterdeterminante

$$\begin{vmatrix} 1 & 1 & 4 \\ 2 & -1 & 5 \\ -1 & 5 & 3 \end{vmatrix} = -3 \tag{2.229}$$

ist. Die Koeffizientenmatrix

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 5 \end{pmatrix} \tag{2.230}$$

hat $\det(\mathbf{A}) = 0$, aber Rang $(\mathbf{A}) = 2$, da die Unterdeterminante

$$\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1 \tag{2.231}$$

ist. Damit haben wir Rang $(\mathbf{A}) \neq$ Rang $(\mathbf{Ab})$, das heißt, ein nicht-lösbares Gleichungssystem.

Verallgemeinerung

Haben wir allgemein m lineare Gleichungen mit n Unbekannten, müssen wir drei Fälle unterscheiden:

- $m = n$: Diesen Fall haben wir hier in allen Einzelheiten behandelt.
- $m < n$: Das Gleichungssystem ist *unterbestimmt.* Ist Rang $(\mathbf{A}) \neq$ Rang $(\mathbf{Ab})$, so ist das System nicht lösbar. Für Rang $(\mathbf{A}) =$ Rang $(\mathbf{Ab})$ gibt es unendlich viele Lösungen in der Form von "Geraden", "Ebenen" etc.

- $m > n$: Das Gleichungssystem ist *überbestimmt.* Gibt es im System Gleichungen, die linear abhängig von anderen Gleichungen sind, können diese eliminiert werden und das System damit möglicherweise in ein $(m = n)$- oder $(m < n)$-Problem umgewandelt werden. Ansonsten ist keine exakte Lösung möglich. Man kann in diesem Fall durch *Ausgleichsrechnung* eine optimale Lösung (das heißt, eine Lösung, die die Gleichungen mit der kleinsten Abweichung erfüllt) ermitteln.

Berechnung der inversen Matrix

In Gln. (2.117) haben wir bereits einen allgemeinen Ausdruck für die Elemente der inversen Matrix $\mathbf{A}^{-1}$ gegeben. Diese Elemente können jedoch auch durch Lösung von linearen Gleichungssystemen bestimmt werden. Ist $\mathbf{A}$ eine quadratische Matrix der Dimension $n \times n$ und $\mathbf{E}$ eine Einheitsmatrix der Dimension $n \times n$, erfüllt die Matrix $\mathbf{X} = \mathbf{A}^{-1}$ die Gleichung

$$\mathbf{A}\,\mathbf{X} = \mathbf{E}. \tag{2.232}$$

Für $n = 2$ können wir diese Gleichung als

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.233}$$

aufschreiben. Diese Gleichung läßt sich auch als zwei Gleichungssysteme schreiben, nämlich

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{21} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{2.234}$$

und

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_{12} \\ x_{22} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.235}$$

Im Allgemeinfall können wir offensichtlich Gln. (2.232) als n lineare Gleichungssysteme ausdrücken. Jedes Gleichungssystem liefert als Lösung eine Spalte der inversen Matrix $\mathbf{A}^{-1}$. Wir können damit alle Elemente der inversen Matrix durch Lösung der Gleichungssysteme erhalten.

Benutzen wir eine Notation analog zur erweiterten Koeffizientenmatrix in Gln. (2.169), so können wir die n Gleichungssysteme, die $\mathbf{A}^{-1}$ als Lösung liefern, als

$$\left(\mathbf{A} \,\middle|\, \mathbf{E} \right) \tag{2.236}$$

aufschrieben. Für die zwei Gleichungssysteme in Gln (2.233) erhalten wir zum Beispiel die Kurzform

$$\left(\begin{array}{cc|cc} a_{11} & a_{12} & 1 & 0 \\ a_{21} & a_{22} & 0 & 1 \end{array} \right). \tag{2.237}$$

Wir können jetzt elementare Zeilenoperationen in dieser Matrix ausführen und damit die zwei Gleichungssysteme “in einem Aufwasch” lösen. Können wir durch Zeilenoperationen die Matrix in

$$\left(\begin{array}{cc|cc} 1 & 0 & r & s \\ 0 & 1 & t & u \end{array}\right) \tag{2.238}$$

überführen, so haben wir die Gleichungssysteme in

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} r & s \\ t & u \end{pmatrix} \tag{2.239}$$

umgewandelt. In diesem Fall gilt trivialerweise $x_{11} = r$, $x_{12} = s$, $x_{21} = t$ und $x_{22} = u$, das heißt

$$\mathbf{A}^{-1} = \begin{pmatrix} r & s \\ t & u \end{pmatrix}. \tag{2.240}$$

Im Allgemeinfall zielen wir also darauf, die Matrix in Gln. (2.236) durch Zeilenoperationen auf die Form

$$\left(\mathbf{E} \,\middle|\, \mathbf{B} \right) \tag{2.241}$$

zu bringen, da genau dann $\mathbf{A}^{-1} = \mathbf{B}$ ist.

Für

$$\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 2 & -1 \end{pmatrix} \tag{2.242}$$

liefert Gln. (2.236)

$$\left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 2 & -1 & 0 & 1 \end{array}\right). \tag{2.243}$$

Durch Zeilenoperationen erhalten wir zunächst

$$\left(\begin{array}{cc|cc} 1 & 3 & 1 & 0 \\ 0 & -7 & -2 & 1 \end{array}\right) \tag{2.244}$$

und dann

$$\left(\begin{array}{cc|cc} 1 & 0 & \frac{1}{7} & \frac{3}{7} \\ 0 & 1 & \frac{2}{7} & -\frac{1}{7} \end{array}\right). \tag{2.245}$$

Also gilt

$$\mathbf{A}^{-1} = \begin{pmatrix} \frac{1}{7} & \frac{3}{7} \\ \frac{2}{7} & -\frac{1}{7} \end{pmatrix}. \tag{2.246}$$

Prüfen Sie nach, ob Gln. (2.125) dieselbe Matrix $\mathbf{A}^{-1}$ liefert.

Für die Matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 7 & 0 \\ 2 & 0 & -1 \end{pmatrix} \tag{2.247}$$

lautet das Koeffizientenschema in Gln. (2.236)

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & 0 & 0 \\ 0 & 7 & 0 & 0 & 1 & 0 \\ 2 & 0 & -1 & 0 & 0 & 1 \end{array}\right). \tag{2.248}$$

Zeilenoperationen ergeben

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & 0 & 0 \\ 0 & 7 & 0 & 0 & 1 & 0 \\ 0 & 0 & -7 & -2 & 0 & 1 \end{array}\right), \tag{2.249}$$

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 3 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{7} & 0 \\ 0 & 0 & 1 & \frac{2}{7} & 0 & -\frac{1}{7} \end{array}\right) \tag{2.250}$$

und

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{7} & 0 & \frac{3}{7} \\ 0 & 1 & 0 & 0 & \frac{1}{7} & 0 \\ 0 & 0 & 1 & \frac{2}{7} & 0 & -\frac{1}{7} \end{array}\right), \tag{2.251}$$

so daß

$$\mathbf{A}^{-1} = \begin{pmatrix} \frac{1}{7} & 0 & \frac{3}{7} \\ 0 & \frac{1}{7} & 0 \\ \frac{2}{7} & 0 & -\frac{1}{7} \end{pmatrix} \tag{2.252}$$

ist. Prüfen Sie nach, ob die Gleichungen $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{E}$ erfüllt sind.

2.7 Das Matrixeigenwertproblem

Wir betrachten eine quadratische Matrix $\mathbf{A}$ der Dimension $n \times n$, n-komponentige Vektoren $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_n$ und komplexe Zahlen $\lambda \in \mathbb{C}$.

Gesucht sind zu einer solchen Matrix $\mathbf{A}$ diejenigen Vektoren $\vec{x}_i$ und die komplexen Zahlen λ_i, für die gilt:

$$\mathbf{A}\,\vec{x}_i = \lambda_i\,\vec{x}_i \tag{2.253}$$

Diese Gleichung kann auch als

$$\mathbf{A}\,\vec{x}_i = \lambda_i\,\mathbf{E}\,\vec{x}_i \tag{2.254}$$

ausgedrückt werden, wobei $\mathbf{E}$ eine Einheitsmatrix der Dimension $n \times n$ ist. Damit haben wir

$$(\mathbf{A} - \lambda_i\,\mathbf{E})\;\vec{x}_i = \vec{0}. \tag{2.255}$$

Die Matrix

$$\mathbf{A} - \lambda_i\,\mathbf{E} = \begin{pmatrix} a_{11}-\lambda_i & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda_i & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33}-\lambda_i & \dots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn}-\lambda_i \end{pmatrix} \tag{2.256}$$

heißt die *charakteristische Matrix.* Die Zahlen λ_i heißen *Eigenwerte* der Matrix $\mathbf{A}$, und die Vektoren $\vec{x}_i$ sind die zugehörigen *Eigenvektoren* der Matrix $\mathbf{A}$.

Mit $\vec{x}_i = \vec{0}$ sind die Gleichungen (2.253) und (2.255) für jeden Wert von λ_i erfüllt. Der Nullvektor $\vec{0}$ gilt aber definitionsgemäß nicht als Eigenvektor. Damit Gln. (2.255) Lösungen mit $\vec{x}_i \neq \vec{0}$ besitzt, muß die *charakteristische Gleichung*

$$\det\left(\mathbf{A} - \lambda_i\,\mathbf{E}\right) = 0 \tag{2.257}$$

erfüllt sein.

Wegen der unendlichen Lösungsmannigfaltigkeit erhalten wir für jeden Eigenwert λ_i unendlich viele Eigenvektoren, die wir als $\alpha\,\vec{x}_i$ schreiben können, wobei $\alpha \in \mathbb{C}$ beliebig ist. Erfüllen λ_i und $\vec{x}_i$ die Gln. (2.253) und multiplizieren wir beide Seiten dieser Gleichung mit α

$$\mathbf{A}\;(\alpha\,\vec{x}_i) = \lambda_i\;(\alpha\,\vec{x}_i)\,, \tag{2.258}$$

dann erfüllen auch λ_i und $\alpha\,\vec{x}_i$ die Eigenwertgleichung. Man wählt α oft so, daß $|\alpha\,\vec{x}_i|$ $= 1$ (mit $\alpha > 0$) ist. In diesem Fall spricht man von *normierten Eigenvektoren.*

Um das Eigenwertproblem einer gegebenen, quadratischen Matrix $\mathbf{A}$ der Dimension $n \times n$ zu lösen, gehen wir wie folgt vor:

1. Zunächst lösen wir die charakteristische Gleichung

$$\det\left(\mathbf{A} - \lambda\,\mathbf{E}\right) = 0. \tag{2.259}$$

Die linke Seite dieser Gleichung ist ein Polynom (das *charakteristische Polynom*) n-ter Ordnung in λ; die n Eigenwerte der Matrix $\mathbf{A}$ sind die n Nullstellen λ_i $(i = 1,\, 2,\, 3,\, \dots,\, n)$ dieses Polynoms.

2. Für jeden Eigenwert λ_i bestimmen wir die zugehörigen Eigenvektoren $\vec{x}_i$ durch Lösung des Gleichungssystems

$$\left(\mathbf{A} - \lambda_i\,\mathbf{E}\right)\,\vec{x}_i = \vec{0}. \tag{2.260}$$

Als Beispiel betrachten wir die Matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}. \tag{2.261}$$

Die charakteristische Gleichung ist

$$\begin{vmatrix} 1-\lambda & 0 \\ -1 & 2-\lambda \end{vmatrix} = 0 \tag{2.262}$$

oder

$$(1-\lambda)\,(2-\lambda) = 0. \tag{2.263}$$

Die Eigenwerte der Matrix $\mathbf{A}$ erhalten wir als die Lösungen dieser Gleichung:

$$\lambda_1 = 1 \;\text{ und }\; \lambda_2 = 2 \tag{2.264}$$

Wir bestimmen jetzt die Eigenvektoren, die zum Eigenwert $\lambda_1 = 1$ gehören, indem wir das Gleichungssystem

$$\begin{pmatrix} 1-\lambda_1 & 0 \\ -1 & 2-\lambda_1 \end{pmatrix} \vec{x}_1 = \begin{pmatrix} 1-\lambda_1 & 0 \\ -1 & 2-\lambda_1 \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{2.265}$$

lösen. Einsetzen von $\lambda_1 = 1$ liefert die beiden Gleichungen

$$0 \cdot x_{11} + 0 \cdot x_{12} = 0 \tag{2.266}$$

$$(-1) \cdot x_{11} + 1 \cdot x_{12} = 0 \tag{2.267}$$

Gleichung (2.266) ist trivialerweise erfüllt für alle (x_{11}, x_{12}), aber Gln. (2.267) ergibt

$$x_{11} = x_{12}. \tag{2.268}$$

Mit $x_{11} = \alpha$ haben wir damit

$$\vec{x}_1 = \begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix} = \begin{pmatrix} \alpha \\ \alpha \end{pmatrix}. \tag{2.269}$$

Dieser Vektor soll, zusammen mit $\lambda_1 = 1$, die Eigenwertgleichung

$$\mathbf{A}\,\vec{x}_1 = \lambda_1\,\vec{x}_1 \tag{2.270}$$

erfüllen, das heißt

$$\begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} \alpha \\ \alpha \end{pmatrix} = \begin{pmatrix} \alpha \\ \alpha \end{pmatrix} = 1 \cdot \begin{pmatrix} \alpha \\ \alpha \end{pmatrix}. \tag{2.271}$$

Wir stellen fest, daß die Eigenwertgleichung erfüllt ist.

Der Eigenvektor $\vec{x}_1$ ist dann normiert, wenn

$$|\vec{x}_1| = \sqrt{\vec{x}_1 \cdot \vec{x}_1} = 1 \tag{2.272}$$

ist. Für den normierten Vektor gilt also

$$\vec{x}_1 \cdot \vec{x}_1 = \alpha^2 + \alpha^2 = 1, \tag{2.273}$$

und damit

$$\alpha = \frac{1}{\sqrt{2}}. \tag{2.274}$$

Der zum Eigenwert $\lambda_1 = 1$ gehörende, normierte Eigenvektor ist folglich

$$\vec{x}_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}. \tag{2.275}$$

Die zum Eigenwert $\lambda_2 = 2$ gehörenden Eigenvektoren berechnen sich aus

$$\begin{pmatrix} 1-\lambda_2 & 0 \\ -1 & 2-\lambda_2 \end{pmatrix} \vec{x}_2 = \begin{pmatrix} 1-\lambda_2 & 0 \\ -1 & 2-\lambda_2 \end{pmatrix} \begin{pmatrix} x_{21} \\ x_{22} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{2.276}$$

Einsetzen von $\lambda_2 = 2$ liefert die beiden identischen Gleichungen

$$(-1) \cdot x_{21} + 0 \cdot x_{22} = 0 \tag{2.277}$$

$$(-1) \cdot x_{21} + 0 \cdot x_{22} = 0 \tag{2.278}$$

mit der allgemeinen Lösung

$$\vec{x}_2 = \begin{pmatrix} x_{21} \\ x_{22} \end{pmatrix} = \begin{pmatrix} 0 \\ \alpha \end{pmatrix}, \tag{2.279}$$

wobei α beliebig ist. Einsetzen in die Eigenwertgleichung

$$\mathbf{A}\,\vec{x}_2 = \lambda_2\,\vec{x}_2 \tag{2.280}$$

zur Probe ergibt hier

$$\begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ \alpha \end{pmatrix} = \begin{pmatrix} 0 \\ 2\,\alpha \end{pmatrix} = 2 \cdot \begin{pmatrix} 0 \\ \alpha \end{pmatrix}. \tag{2.281}$$

Die Eigenwertgleichung ist erfüllt. Normierung des Eigenvektors nach

$$\vec{x}_2 \cdot \vec{x}_2 = 0^2 + \alpha^2 = 1 \tag{2.282}$$

liefert

$$\alpha = 1 \tag{2.283}$$

und damit lautet der zum Eigenwert $\lambda_2 = 2$ gehörende, normierte Eigenvektor

$$\vec{x}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.284}$$

Einige Sätze

(a) Hat die charakteristische Gleichung

$$\det(\mathbf{A} - \lambda\,\mathbf{E}) = 0 \tag{2.285}$$

n *verschiedene* Lösungen λ_i, so gilt

- zu jedem λ_i gibt es genau einen Eigenvektor $\vec{x}_i$ bestimmt bis auf einem Faktor α und
- die n verschiedenen Eigenvektoren $\vec{x}_i$ sind linear unabhängig.

Man spricht von *Entartung*, wenn mehrere Eigenwerte gleich sind.

(b) Ist die Matrix $\mathbf{A}$ reell und symmetrisch ($\mathbf{A}^{\mathrm{T}} = \mathbf{A}$), so gilt:

- alle n Eigenwerte sind reell.
- Eigenvektoren $\vec{x}_i$ und $\vec{x}_j$, die zu verschiedenen Eigenwerten gehören, sind zueinander orthogonal: $\vec{x}_i \cdot \vec{x}_j = 0$.
- Es läßt sich eine reelle orthogonale Matrix $\mathbf{C}$ finden, so daß

$$\mathbf{B} = \mathbf{C}^{-1}\,\mathbf{A}\,\mathbf{C} \tag{2.286}$$

eine Diagonalmatrix

$$\mathbf{B} = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix} \tag{2.287}$$

ist; ihre Diagonalelemente sind die Eigenwerte λ_i der Matrix $\mathbf{A}$. Die Spalten von $\mathbf{C}$ entsprechen den Eigenvektoren von $\mathbf{A}$:

$$\begin{pmatrix} C_{1i} \\ C_{2i} \\ C_{3i} \\ \vdots \\ C_{ni} \end{pmatrix} = \vec{x}_i. \tag{2.288}$$

Man sagt, daß $\mathbf{C}^{-1}\,\mathbf{A}\,\mathbf{C}$ eine *Hauptachsentransformation* von $\mathbf{A}$ ist. Gleichung (2.151) (mit $\mathbf{S} = \mathbf{C}^{-1}$) besagt, daß

$$\det(\mathbf{B}) = \det(\mathbf{A}) \tag{2.289}$$

ist. Aus Gln. (2.287) erhalten wir

$$\det(\mathbf{B}) = \lambda_1\,\lambda_2\,\lambda_3\ldots\lambda_n = \prod_{i=1}^{n}\lambda_i, \tag{2.290}$$

und damit

$$\prod_{i=1}^{n}\lambda_i = \det(\mathbf{A})\,. \tag{2.291}$$

Das Produkt der Eigenwerte von $\mathbf{A}$ ist gleich der Determinante von $\mathbf{A}$. Gleichung (2.152) (mit $\mathbf{S} = \mathbf{C}^{-1}$) besagt, daß

$$\mathrm{Spur}\,(\mathbf{B}) = \mathrm{Spur}\,(\mathbf{A}) \tag{2.292}$$

ist. Aus Gln. (2.287) erhalten wir

$$\mathrm{Spur}\,(\mathbf{B}) = \sum_{i=1}^{n}\lambda_i, \tag{2.293}$$

und damit

$$\sum_{i=1}^{n}\lambda_i = \mathrm{Spur}\,(\mathbf{A})\,. \tag{2.294}$$

Die Summe der Eigenwerte von $\mathbf{A}$ ist gleich der Spur von $\mathbf{A}$.

Beispiel: Wir bestimmen die Eigenwerte und Eigenvektoren der symmetrischen Matrix

$$\mathbf{A} = \begin{pmatrix} 7 & -3\sqrt{3} \\ -3\sqrt{3} & 13 \end{pmatrix}. \tag{2.295}$$

Die charakteristische Gleichung lautet

$$\begin{vmatrix} 7-\lambda & -3\sqrt{3} \\ -3\sqrt{3} & 13-\lambda \end{vmatrix} = 0 \tag{2.296}$$

oder

$$(7-\lambda)\,(13-\lambda) - 27 = \lambda^2 - 20\,\lambda + 64 = 0. \tag{2.297}$$

Die Eigenwerte der Matrix $\mathbf{A}$ erhalten wir als die Lösungen dieser Gleichung:

$$\lambda_1 = 4 \;\text{ und }\; \lambda_2 = 16. \tag{2.298}$$

Prüfen Sie nach, ob Gln. (2.291) und (2.294) erfüllt sind.

Der zum Eigenwert $\lambda_1 = 4$ zugehörige Eigenvektor

$$\vec{c}_1 = \begin{pmatrix} c_{11} \\ c_{21} \end{pmatrix} \tag{2.299}$$

erfüllt

$$\begin{pmatrix} 7-4 & -3\sqrt{3} \\ -3\sqrt{3} & 13-4 \end{pmatrix} \begin{pmatrix} c_{11} \\ c_{21} \end{pmatrix} = \begin{pmatrix} 3 & -3\sqrt{3} \\ -3\sqrt{3} & 9 \end{pmatrix} \begin{pmatrix} c_{11} \\ c_{21} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{2.300}$$

Die beiden Gleichungen im Gleichungssystem sind linear abhängig und liefern die Beziehung

$$c_{11} = \sqrt{3}\, c_{21}. \tag{2.301}$$

Das heißt

$$\vec{c}_1 = \begin{pmatrix} \sqrt{3}\,\alpha \\ \alpha \end{pmatrix}, \tag{2.302}$$

wobei α beliebig ist. Damit $\vec{c}_1$ normiert ist, muß gelten

$$3\,\alpha^2 + \alpha^2 = 1, \tag{2.303}$$

und damit

$$\alpha = \frac{1}{2}. \tag{2.304}$$

Der normierte Eigenvektor ist also

$$\vec{c}_1 = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix}. \tag{2.305}$$

Analogerweise findet man den zum Eigenwert $\lambda_2 = 16$ zugehörigen normierten Eigenvektor

$$\vec{c}_2 = \begin{pmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}. \tag{2.306}$$

Die beiden Eigenvektoren $\vec{c}_1$ und $\vec{c}_2$ gehören zu verschiedenen Eigenwerten einer reellen symmetrischen Matrix, und wir bestätigen, daß

$$\vec{c}_1 \cdot \vec{c}_2 = 0 \tag{2.307}$$

ist. Die beiden Eigenvektoren sind orthogonal.

Die Matrix $\mathbf{C}$ in Gln. (2.286) ergibt sich im vorliegenden Beispiel als (siehe Gln. (2.288))

$$\mathbf{C} = \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}. \tag{2.308}$$

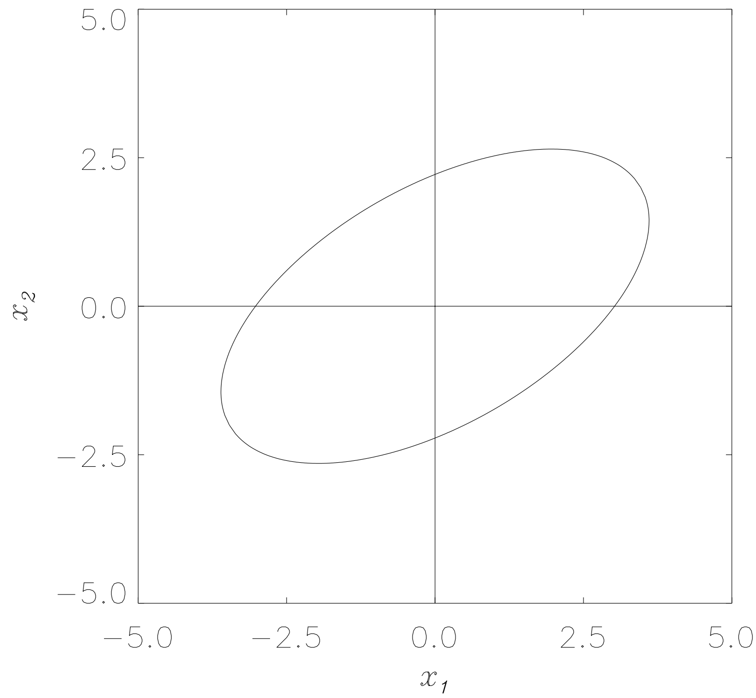

Abbildung 2.2: Die Ellipse ist definiert durch die Gleichung $7\,x_1^2 - 6\,\sqrt{3}\,x_1\,x_2 + 13\,x_2^2 = 64$.

Diese Matrix ist orthogonal, und damit gilt auch

$$\mathbf{C}^{-1} = \mathbf{C}^{\mathrm{T}} = \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix}. \tag{2.309}$$

Die Hauptachsentransformation der Matrix $\mathbf{A}$ ist folglich

$$\begin{aligned} \mathbf{C}^{-1}\,\mathbf{A}\,\mathbf{C} &= \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \begin{pmatrix} 7 & -3\sqrt{3} \\ -3\sqrt{3} & 13 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \\ &= \begin{pmatrix} 4 & 0 \\ 0 & 16 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \end{aligned} \tag{2.310}$$

Die Gleichung (2.286) (mit der Matrix $\mathbf{B}$ definiert in Gln. (2.287)) ist also erfüllt.

Die Bestimmung von Matrix-Eigenwerten spielt eine wichtige Rolle bei quantenmechanischen Berechnungen in der theoretischen Chemie. Wir werden hier eine andere Anwendung der Hauptachsentransformation, die Vereinfachung einer sogenannten *quadratischen Form*, exemplarisch darstellen. Betrachten wir die Funktion zweier Variablen

$$f(x_1, x_2) = a_{11}\, x_1^2 + 2\, a_{12}\, x_1\, x_2 + a_{22}\, x_2^2, \tag{2.311}$$

so können wir einfach nachvollziehen, daß wir diese Funktion als

$$f(x_1, x_2) = \vec{x}^{\mathrm{T}}\, \mathbf{A}\, \vec{x} \tag{2.312}$$

schreiben können, wobei

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{und} \quad \mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \tag{2.313}$$

sind; $\mathbf{A}$ ist also eine symmetrische Matrix. Die Funktion in Gln. (2.311) nennt man *eine quadratische Form.* Der Trick ist nun, durch eine Koordinatentransformation neue Koordinaten

$$\vec{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \mathbf{C}^{-1} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \mathbf{C}^{-1}\, \vec{x} \tag{2.314}$$

zu definieren. Dann gilt

$$\vec{x} = \mathbf{C}\, \vec{y} \;\; \text{und} \;\; \vec{x}^{\mathrm{T}} = \vec{y}^{\mathrm{T}}\, \mathbf{C}^{\mathrm{T}} = \vec{y}^{\mathrm{T}}\, \mathbf{C}^{-1}, \tag{2.315}$$

wobei wir die Tatsache benutzt haben, daß $\mathbf{C}$ orthogonal ist. Einsetzen in Gln. (2.312) ergibt

$$f(y_1, y_2) = \vec{y}^{\mathrm{T}}\, \mathbf{C}^{-1}\, \mathbf{A}\, \mathbf{C}\, \vec{y}. \tag{2.316}$$

Da

$$\mathbf{C}^{-1}\, \mathbf{A}\, \mathbf{C} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \tag{2.317}$$

ist, haben wir

$$\begin{aligned} f(y_1, y_2) &= \vec{y}^{\mathrm{T}} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \vec{y} \\ &= (y_1\; y_2) \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \lambda_1\, y_1^2 + \lambda_2\, y_2^2. \end{aligned} \tag{2.318}$$

Durch die Koordinatentransformation haben wir eine Vereinfachung der Funktion f erreicht; es gibt im neuen Ausdruck keinen "Kreuzterm" mehr, der proportional zu $y_1\, y_2$ ist.

Nehmen wir als Beispiel die Matrix $\mathbf{A}$ in Gln. (2.295) als Koeffizientenmatrix in Gln. (2.312). Die entsprechende quadratische Form ist

$$f(x_1, x_2) = 7\, x_1^2 - 6\,\sqrt{3}\, x_1\, x_2 + 13\, x_2^2. \tag{2.319}$$

Die Gleichung

$$f(x_1, x_2) = \lambda_1\, \lambda_2, \tag{2.320}$$

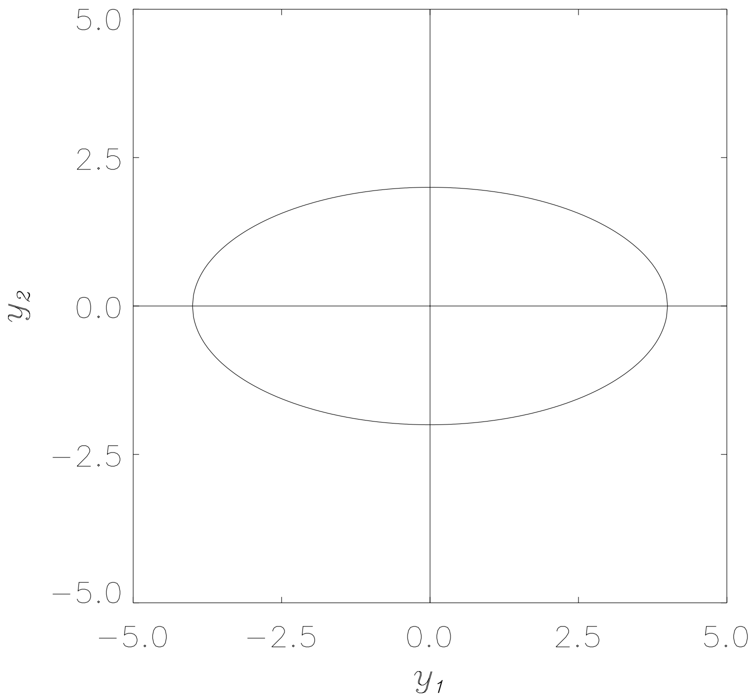

Abbildung 2.3: Die Ellipse ist definiert durch die Gleichung $(y_1/4)^2 + (y_2/2)^2 = 1$.

das heißt

$$7\,x_1^2 - 6\,\sqrt{3}\,x_1\,x_2 + 13\,x_2^2 = 64, \tag{2.321}$$

definiert eine Ellipse in der (x_1, x_2)-Ebene (Abb. 2.2). Die Ellipse steht "schräg" im Koordinatensystem. Durch die Koordinatentransformation

$$\vec{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \mathbf{C}^{-1}\,\vec{x} = \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} \end{pmatrix} \vec{x} \tag{2.322}$$

erhalten wir

$$f(y_1, y_2) = \lambda_1\,y_1^2 + \lambda_2\,y_2^2 = 4\,y_1^2 + 16\,y_2^2. \tag{2.323}$$

Die Gleichung

$$f(y_1, y_2) = \lambda_1\,\lambda_2 \tag{2.324}$$

läßt sich als

$$\lambda_1\,y_1^2 + \lambda_2\,y_2^2 = \lambda_1\,\lambda_2 \tag{2.325}$$

oder äquivalenterweise als

$$\frac{y_1^2}{\lambda_2} + \frac{y_2^2}{\lambda_1} = 1 \tag{2.326}$$

aufschreiben. Im vorliegenden Beispiel gilt also

$$\frac{y_1^2}{16} + \frac{y_2^2}{4} = \left(\frac{y_1}{4}\right)^2 + \left(\frac{y_2}{2}\right)^2 = 1. \tag{2.327}$$

Abbildung 2.3 zeigt die durch diese Gleichung definierte Ellipse in der (y_1, y_2)-Ebene. Die Koordinatentransformation in Gln. (2.322) bewirkt, daß die *Hauptachsen* (das heißt, die Symmetrieachsen) der Ellipse als Koordinatenachsen benutzt werden.

3 Differentialgleichungen

3.1 Grundbegriffe

Unter einer *gewöhnlichen Differentialgleichung n-ter Ordnung* versteht man eine Bestimmungsgleichung für eine unbekannte Funktion

$$y = \varphi(x) \tag{3.1}$$

einer unabhängigen Variablen x, die Ableitungen

$$\begin{aligned} y' &= \frac{d\varphi}{dx} \\ y'' &= \frac{d^2\varphi}{dx^2} \\ &\vdots \\ y^{(n)} &= \frac{d^n\varphi}{dx^n} \end{aligned} \tag{3.2}$$

bis zur n-ter Ordnung enthält. Ferner können in der Gleichung die unbekannte Funktion $\varphi(x)$ sowie die unabhängige Variable x auftreten. Allgemein läßt sich eine gewöhnliche Differentialgleichung n-ter Ordnung in der impliziten Form

$$F(x; y, y', y'', \ldots, y^{(n)}) = 0 \tag{3.3}$$

darstellen.

Die Ordnung einer gewöhnlichen Differentialgleichung wird also durch die Ordnung der *höchsten* in ihr enthaltenen Ableitung bestimmt.
Beispiel: $y' - 5x = 0$ ist eine Differentialgleichung 1. Ordnung. $xy'' + 2y' + y^2 - 2x = 0$ ist eine Differentialgleichung 2. Ordnung.

Eine gewöhnliche Differentialgleichung n-ter Ordnung heißt *linear*, wenn

(a) die unbekannte Funktion $\varphi(x)$ und alle ihre Ableitungen höchstens in 1. Potenz vorkommen, und

(b) keine Produkte dieser Größen auftreten.

Sie hat daher die allgemeine Form

$$a_n(x)\, y^{(n)} + a_{n-1}(x)\, y^{(n-1)} + \ldots + a_0(x)\, y = b(x) \tag{3.4}$$

wobei $a_n(x) \neq 0$ ist. Ist $b(x) = 0$, so heißt die lineare Differentialgleichung *homogen*, andernfalls *inhomogen*.

Beispiel: $y' - 5x = 0$ und $y^{(k)}0e^x$ sind lineare Differentialgleichung 1. bzw. k-ter Ordnung.
Beispiel: $y'' - a\,y' + b\,y = f(x)$ ist eine linear Differentialgleichung 2. Ordnung mit konstanten Koeffizienten (sog. *Schwingungsgleichung*). Differentialgleichungen von diesem Typ spielen in Physik und Chemie eine bedeutende Rolle (harmonischer Oszillator, ungedämpfte und gedämpfte mechanische und elektromagnetische Schwingungen; erzwungene Schwingungen; Molekülschwingungen).

Eine Funktion $y = \varphi(x)$, deren Funktionswert und deren Ableitungswerte eine Differentialgleichung identisch erfüllen, heißt eine Lösung oder ein Integral der Differentialgleichung. Die Gesamtheit aller Lösungen bildet die *Lösungsmenge* der Differentialgleichung.

Unsere Aufgabe besteht nun darin, *alle* Lösungen einer Differentialgleichung aufzusuchen. Man spricht in diesem Zusammenhang auch häufig von der *Integration* einer Differentialgleichung.

Die *allgemeine* Lösung (*allgemeines* Integral) einer gewöhnlichen Differentialgleichung n-ter Ordnung enthält noch *willkürliche Konstanten*, die man als *Parameter* oder *Integrationskonstanten* der Differentialgleichung bezeichnnet. Hierbei gilt, daß das allgemeine Integral einer gewöhnlichen Differentialgleichung n-ter Ordnung genau n willkürliche und unabhängige Integrationskonstanten enthält.

Weist man diesen n Integrationskonstanten im allgemeinen Integral einer Differentialgleichung einen festen Zahlenwert zu, so erhält man eine *spezielle* (oder *partikuläre*) Lösung.

3.2 Gewöhnliche Differentialgleichungen 1. Ordnung

Trennung der Variablen

Wir betrachten solche Differentialgleichungen, die sich als

$$y' = \frac{dy}{dx} = g(x)h(y) \tag{3.5}$$

schreiben lassen. Wir rechnen formal

$$\frac{dy}{h(y)} = g(x)dx; \quad h(y) \neq 0 \tag{3.6}$$

(wäre $h(y) = 0$, wäre die Lösung der Differentialgleichung trivial!), integrieren

$$\int \frac{dy}{h(y)} = \int g(x)dx \tag{3.7}$$

und lösen nach y auf.

Beispiel: Für die Differentialgleichung

$$y' = \frac{dy}{dx} = -ky \tag{3.8}$$

ist eine Trennung der Variablen möglich. Wir wählen

$$g(x) = -k \;\text{ und }\; h(y) = y. \tag{3.9}$$

Gleichung (3.6) ergibt

$$\frac{dy}{y} = -kdx. \tag{3.10}$$

Nach Integration (Gln. (3.7)) erhalten wir

$$\ln|y| + C_y = -kx + C_x, \tag{3.11}$$

wobei C_y und C_x Integrationskonstanten sind. Also

$$\ln|y| = -kx + (C_x - C_y) = -kx + C \tag{3.12}$$

mit $C = C_x - C_y$ und

$$y = \pm\exp(-kx + C) = \pm e^C e^{-kx} = y_0\, e^{-kx}; \tag{3.13}$$

$y_0 = \pm e^C$. Wir haben $y(x) \neq 0$ für alle x. Wählen wir eine Lösung mit $y_0 = +e^C$ ist $y(x) > 0$ für alle x. Wählen wir $y_0 = -e^C$ ist $y(x) < 0$ für alle x.

Beispiel: Für die Differentialgleichung

$$y' = \frac{dy}{dx} = -\frac{x}{y} \tag{3.14}$$

haben wir

$$g(x) = x \;\text{ und }\; h(y) = -\frac{1}{y}. \tag{3.15}$$

Gleichung (3.6) ergibt

$$-ydy = xdx. \tag{3.16}$$

Nach Integration (Gln. (3.7)) erhalten wir

$$-\frac{1}{2}y^2 + C_y = \frac{1}{2}x^2 + C_x, \tag{3.17}$$

wobei C_y und C_x Integrationskonstanten sind, und damit

$$x^2 + y^2 = 2(C_y - C_x) = c^2. \tag{3.18}$$

Die Lösungen der Differentialgleichung sind also "kreisförmige" Funktionen

$$y = \pm\sqrt{c^2 - x^2}. \tag{3.19}$$

Damit y reell ist, müssen die Bedingungen $c^2 = 2(C_y - C_x) > 0$ und $-|c| \leq x \leq |c|$ erfüllt sein; $|c|$ ist Radius des Kreises.

“Variation der Konstanten”

Für die Differentialgleichung

$$\frac{dy}{dx} + P(x)\,y = Q(x) \tag{3.20}$$

ist eine Trennung der Variablen nicht möglich für $Q(x) \neq 0$. Hier benutzen wir einen Trick, die “Variation der Konstanten” nach Lagrange.

Der erste Schritt besteht darin, die *homogene* Differentialgleichung

$$\frac{dy}{dx} + P(x)\,y = 0 \tag{3.21}$$

zu lösen. Hier können wir die Variablen trennen:

$$\frac{dy}{y} = -P(x)dx \tag{3.22}$$

und

$$\int \frac{dy}{y} = -\int P(x)dx \tag{3.23}$$

so daß

$$y = C\,e^{-\int P(x)dx} \tag{3.24}$$

wobei C eine Integrationskonstante ist.

Um die ursprüngliche, inhomogene Differentialgleichung in Gln. (3.20) lösen zu können, “variieren wir jetzt die Konstante”, indem wir C in Gln. (3.24) durch eine Funktion von x, $C(x)$, ersetzen. Der Ansatz für y ist dann

$$y = C(x)e^{-\int P(x)dx} = C(x)e^{-R(x)}, \tag{3.25}$$

wobei wir

$$R(x) = \int P(x)dx \tag{3.26}$$

definiert haben. Das heißt

$$\frac{dy}{dx} = \frac{dC}{dx}\,e^{-R(x)} - C(x)\,e^{-R(x)}\,\frac{dR}{dx}. \tag{3.27}$$

Da ja $dR/dx = P(x)$, haben wir

$$\frac{dy}{dx} = \frac{dC}{dx}\,e^{-R(x)} - C(x)e^{-R(x)}\,P(x). \tag{3.28}$$

Diesen Ausdruck setzen wir zusammen mit Gln. (3.25) in die inhomogene Differentialgleichung, Gln. (3.20), ein und erhalten

$$\frac{dC}{dx}\,e^{-R(x)} - C(x)e^{-R(x)}\,P(x) + P(x)\,C(x)e^{-R(x)} = Q(x) \tag{3.29}$$

oder

$$\frac{dC}{dx} = Q(x)\, e^{R(x)}. \tag{3.30}$$

Es folgt daß

$$C(x) = \int Q(x)\, e^{R(x)} dx + K, \tag{3.31}$$

wobei K eine Integrationskonstante ist, und wir erhalten von Gln. (3.25) das Endergebnis

$$y = \left[\int Q(x)\, e^{R(x)} dx + K\right] e^{-R(x)} = \left[\int Q(x)\, e^{\int P(x)dx} dx + K\right] e^{-\int P(x)dx} \tag{3.32}$$

Man sollte nicht Gln. (3.32) auswendig lernen, sondern jedesmal die Schritte zur Lösung von Gln. (3.20) einzeln durchführen.

Beispiel: Die Differentialgleichung

$$y' - y = \frac{dy}{dx} - y = \epsilon\, x \tag{3.33}$$

(ϵ ist eine Konstante) hat die Form der Gln. (3.20) mit

$$P(x) = -1 \;\text{ und }\; Q(x) = \epsilon\, x. \tag{3.34}$$

Erst lösen wir

$$\frac{dy}{dx} - y = 0 \tag{3.35}$$

mittels Trennung der Variablen:

$$\frac{dy}{y} = dx \tag{3.36}$$

ergibt

$$\ln y + C_y = x + C_x \tag{3.37}$$

und

$$y = \exp\left(x + C_x - C_y\right) = Ce^x \tag{3.38}$$

wobei $C = \exp(C_x - C_y)$ ist.

Wir variieren jetzt die Konstante und setzen y als

$$y = C(x)\, e^x \tag{3.39}$$

an. Also ist

$$\frac{dy}{dx} = \frac{dC}{dx}\, e^x + C(x)\, e^x. \tag{3.40}$$

Einsetzen von Gln. (3.39) und (3.40) in Gln. (3.33) liefert

$$\frac{dC}{dx}\,e^x + C(x)\,e^x - C(x)\,e^x = \epsilon\,x, \tag{3.41}$$

und wir haben folglich

$$\frac{dC}{dx} = \epsilon\,x\,e^{-x}. \tag{3.42}$$

Wir bestimmen $C(x)$ mittels *partieller Integration.* Für willkürliche Funktionen $f(x)$ und $g(x)$ gilt

$$\frac{d}{dx}(f\,g) = \frac{df}{dx}g + f\frac{dg}{dx} \tag{3.43}$$

und damit

$$\int f\frac{dg}{dx}dx = f\,g - \int \frac{df}{dx}g dx. \tag{3.44}$$

Mit $f(x) = \epsilon x$ ist $df/dx = \epsilon$. Mit $dg/dx = e^{-x}$ ist $g(x) = -e^{-x} + K_g$. Gleichung (3.44) liefert in diesem Fall

$$\begin{aligned}
C(x) &= \int \epsilon x e^{-x} dx \\
&= \epsilon\,x\,(-e^{-x} + K_g) - \int \epsilon\,(-e^{-x} + K_g)\,dx + K \\
&= -\epsilon\,x\,e^{-x} + \epsilon\,x\,K_g + \int \epsilon\,e^{-x}dx - \int \epsilon\,K_g\,dx + K \\
&= -\epsilon\,x\,e^{-x} + \epsilon\,x\,K_g - \epsilon\,e^{-x} - \epsilon\,x\,K_g + K \\
&= -\epsilon\,x\,e^{-x} - \epsilon\,e^{-x} + K = -\epsilon\,(x+1)\,e^{-x} + K,
\end{aligned} \tag{3.45}$$

wobei K_g und K Integrationskonstanten sind. Von Gln. (3.39) erhalten wir damit

$$y = -\epsilon\,(x+1) + K\,e^x. \tag{3.46}$$

Prüfen Sie nach, daß diese Funktion wirklich eine Lösung von Gln. (3.33) ist.
Hinweis: Allgemein gilt, daß die Lösung der Differentialgleichung in Gln. (3.20)

$$\frac{dy}{dx} + P(x)\,y = Q(x) \tag{3.47}$$

in der Form

$$y = y_{\mathrm{H}} + y_{\mathrm{P}} \tag{3.48}$$

geschrieben werden kann. Hier ist y_{H} die Lösung der *homogenen* Differentialgleichung

$$\frac{dy_{\mathrm{H}}}{dx} + P(x)\,y_{\mathrm{H}} = 0, \tag{3.49}$$

und y_P ist eine *partikuläre* Lösung der gesamten, inhomogenen Differentialgleichung in Gln. (3.47). Können wir also eine Lösung y_P raten, können wir sämtliche Lösungen der inhomogenen Differentialgleichung dadurch bestimmen, daß wir zu y_P sämtliche Lösungen der homogenen Differentialgleichung addieren. Im obigen Beispiel ist $y_\mathrm{P} = -\epsilon\,(x+1)$ und $y_\mathrm{H} = K\,e^x$ (Prüfen Sie nach!).

Exakte Differentialgleichungen

Eine Differentialgleichung

$$g(x,y) + h(x,y)\frac{dy}{dx} = 0 \tag{3.50}$$

heißt *exakt*, wenn es eine Funktion $u(x,y)$ gibt, so daß

$$g(x,y) = \frac{\partial u}{\partial x} \;\text{ und }\; h(x,y) = \frac{\partial u}{\partial y}. \tag{3.51}$$

In diesem Fall ist die Differentialgleichung Gln. (3.50) erfüllt, wenn das sogenannte *totale Differential*

$$du = \frac{\partial u}{\partial x}\,dx + \frac{\partial u}{\partial y}\,dy = g(x,y)\,dx + h(x,y)\,dy = 0. \tag{3.52}$$

Ist die Differentialgleichung exakt, haben wir

$$\frac{\partial g}{\partial y} = \frac{\partial h}{\partial x} = \frac{\partial^2 u}{\partial x \partial y}; \tag{3.53}$$

Die Bedingung, die erfüllt sein muß, damit Gln. (3.50) exakt ist, ist also

$$\frac{\partial g}{\partial y} = \frac{\partial h}{\partial x}. \tag{3.54}$$

Ist das totale Differential (Gln. (3.52)) $du = 0$, dann ist die Funktion u konstant. Das heißt, der Zusammenhang zwischen x und y ist bestimmt durch

$$u(x,y) = K, \tag{3.55}$$

wobei K eine Integrationskonstante ist, und wir können $y(x)$ bestimmen, wenn wir diese Gleichung nach y auflösen.
Beispiel: Die Differentialgleichung

$$\left(\frac{1}{x} - y\right) + (y-x)\frac{dy}{dx} = 0 \tag{3.56}$$

hat

$$g(x,y) = \frac{1}{x} - y \quad \text{und} \quad h(x,y) = y - x. \tag{3.57}$$

Es gilt

$$\frac{\partial g}{\partial y} = -1 \quad \text{und} \quad \frac{\partial h}{\partial x} = -1; \tag{3.58}$$

Gln. (3.54) ist erfüllt und die Differentialgleichung ist exakt.

Wir möchten die Funktion $u(x,y)$ bestimmen. Wir haben

$$u(x,y) = \int \frac{\partial u}{\partial x}\, dx = \int g(x,y)\, dx = \int \left(\frac{1}{x} - y\right) dx = \ln|x| - xy + C_1(y). \tag{3.59}$$

Die Funktion $C_1(y)$ hängt von y ab; sie ist unabhängig von x. Weiter wissen wir, daß

$$\frac{\partial u}{\partial y} = h(x,y) = y - x. \tag{3.60}$$

Vergleichen wir Gln. (3.59) und (3.60), sehen wir

$$-x + \frac{\partial C_1}{\partial y} = y - x \tag{3.61}$$

so daß

$$\frac{\partial C_1}{\partial y} = y \tag{3.62}$$

und

$$C_1(y) = \frac{1}{2} y^2 + c. \tag{3.63}$$

c ist hier eine Integrationskonstante. Wir haben also

$$u(x,y) = \ln|x| - xy + \frac{1}{2} y^2 + c, \tag{3.64}$$

und die Bedingung in Gln. (3.55) können wir als

$$\ln|x| - xy + \frac{1}{2} y^2 + C = 0 \tag{3.65}$$

schreiben, wobei $C = c - K$ ist. Durch Lösung dieser Gleichung können wir $y(x)$ erhalten. Das Ergebnis ist

$$y = x \pm \sqrt{x^2 - 2(\ln|x| + C)}. \tag{3.66}$$

Es kann sein, daß eine Differentialgleichung in der Form der Gln. (3.50) geschrieben werden kann, aber daß die Bedingung in Gln. (3.54) nicht erfüllt ist. Man sucht dann oft einen "integrierenden Faktor" $\lambda(x,y)$ so daß

$$\lambda(x,y)\, g(x,y) + \lambda(x,y)\, h(x,y) \frac{dy}{dx} = 0 \tag{3.67}$$

eine exakte Differentialgleichung ist. Das heißt, mit den Definitionen

$$A(x,y) = \lambda(x,y)\, g(x,y) \quad \text{und} \quad B(x,y) = \lambda(x,y)\, h(x,y) \tag{3.68}$$

bestimmen wir $\lambda(x,y)$ so daß

$$\frac{\partial A}{\partial y} = \frac{\partial B}{\partial x}. \tag{3.69}$$

Beispiel: Die Differentialgleichung

$$(1 - xy) + (xy - x^2)\frac{dy}{dx} = 0 \tag{3.70}$$

hat

$$g(x,y) = 1 - xy \quad \text{und} \quad h(x,y) = xy - x^2. \tag{3.71}$$

Damit sind

$$\frac{\partial g}{\partial y} = -x \quad \text{und} \quad \frac{\partial h}{\partial x} = y - 2x; \tag{3.72}$$

Gln. (3.54) ist *nicht* erfüllt. Multiplikation der Differentialgleichung mit $\lambda(x,y) = \lambda(x) = 1/x$ erzeugt jedoch

$$\left(\frac{1}{x} - y\right) + (y - x)\frac{dy}{dx} = 0, \tag{3.73}$$

das heißt die exakte Differentialgleichung in Gln. (3.56).

Falls der integrierende Faktor λ nur von einer der beiden Koordinaten x und y abhängt (das heißt, er kann als $\lambda(x)$ oder $\lambda(y)$ geschrieben werden), dann kann er elementar gefunden werden.

Nehmen wir für die Differentialgleichung in Gln. (3.70) an, daß wir einen integrierenden Faktor $\lambda(x)$ finden können. Dann soll nach Gln. (3.69) gelten

$$\frac{\partial}{\partial y}\left[\lambda(x)\,(1 - xy)\right] = \frac{\partial}{\partial x}\left[\lambda(x)\,(xy - x^2)\right]. \tag{3.74}$$

Also haben wir

$$\lambda(x)\,(-x) = \frac{\partial \lambda}{\partial x}\,(xy - x^2) + \lambda(x)\,(y - 2x), \tag{3.75}$$

$$\frac{\partial \lambda}{\partial x}\, x\,(y - x) + \lambda(x)\,(y - x) = 0, \tag{3.76}$$

$$\frac{\partial \lambda}{\partial x}\, x + \lambda(x) = 0, \tag{3.77}$$

$$\frac{d\lambda}{\lambda} = -\frac{dx}{x}, \tag{3.78}$$

$$\ln|\lambda| = -\ln|x| + C \tag{3.79}$$

und damit

$$\lambda = \frac{K}{x}; \quad K = \pm\exp(C). \tag{3.80}$$

3.3 Gewöhnliche Differentialgleichungen 2. Ordnung

Wir betrachten nur lineare Differentialgleichungen 2. Ordnung mit konstanten Koeffizienten

$$a\,\frac{d^2y}{dx^2} + b\,\frac{dy}{dx} + c\,y = s(x), \tag{3.81}$$

wobei $a \neq 0$, b und c konstante, komplexe Zahlen sind: a, b, $c \in \mathbb{C}$.

Differentialgleichungen 2. Ordnung kommen oft in der klassischen Mechanik vor. Betrachten wir ein Teilchen mit Masse m, das harmonische Schwingungen auf einer y-Achse ausführt. Die Potentialfunktion ist

$$V(y) = \frac{1}{2}\,k\,y^2, \tag{3.82}$$

wobei die Konstante $k \in \mathbb{R}$. Die Kraft (entlang der y-Achse), die auf das Teilchen wirkt, ist damit (das Hook'sche Gesetz)

$$F(y) = -\frac{dV}{dy} = -k\,y, \tag{3.83}$$

und die Bewegungsgleichung ist (Newton's 2. Gesetz)

$$m\frac{d^2y}{dt^2} = -k\,y \tag{3.84}$$

oder

$$m\frac{d^2y}{dt^2} + k\,y = 0, \tag{3.85}$$

wobei t die Zeit bezeichnet.

Die homogene lineare Differentialgleichung 2. Ordnung mit konstanten Koeffizienten

Wir kehren zur allgemeinen Differentialgleichung in Gln. (3.81) zurück. Zuerst betrachten wir das entsprechende *homogene Problem:*

$$a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + c\,y = 0. \tag{3.86}$$

Wir raten, daß die Funktion

$$y(x) = \exp(rx) \tag{3.87}$$

(wobei r eine komplexe Zahl ist) Gln. (3.86) erfüllt. Offensichtlich gilt

$$\frac{dy}{dx} = r\,\exp(rx) \;\text{ und }\; \frac{d^2y}{dx^2} = r^2\,\exp(rx). \tag{3.88}$$

Durch Einsetzen in Gln. (3.86) erhalten wir

$$ar^2\,\exp(rx) + br\,\exp(rx) + c\exp(rx) = \big(a\,r^2 + b\,r + c\big)\exp(rx) = 0. \tag{3.89}$$

Wir sehen, daß die Funktion $y(x) = \exp(rx)$ eine Lösung von Gln. (3.86) ist, wenn r die *charakteristische Gleichung*

$$a\,r^2 + b\,r + c = 0 \tag{3.90}$$

erfüllt. Die möglichen r-Werte sind folglich

$$r_1 = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \quad \text{und} \quad r_2 = \frac{-b + \sqrt{b^2 - 4ac}}{2a}. \tag{3.91}$$

Man kann einfach nachrechnen, daß

$$a\,r^2 + b\,r + c = a\,(r - r_1)\,(r - r_2) \tag{3.92}$$

ist, wobei r_1 und r_2 in Gln. (3.91) gegeben sind. Die homogene Differentialgleichung in Gln. (3.86) kann in analoger Weise 'faktorisiert' werden. Wir haben (nachrechnen!)

$$b = -a\,(r_1 + r_2) \quad \text{und} \quad c = a\,r_1\,r_2; \tag{3.93}$$

diese Ausdrücke werden in Gln. (3.86) eingesetzt:

$$\begin{aligned} & a\frac{d^2y}{dx^2} - a\,(r_1 + r_2)\,\frac{dy}{dx} + a\,r_1\,r_2\,y \\ = \;& a\left(\frac{d^2y}{dx^2} - r_1\,\frac{dy}{dx} - r_2\,\frac{dy}{dx} + r_1\,r_2\,y\right) \\ = \;& a\left(\frac{d}{dx}\left[\frac{dy}{dx} - r_1\,y\right] - r_2\left[\frac{dy}{dx} - r_1\,y\right]\right) = 0. \end{aligned} \tag{3.94}$$

Mit der Definition

$$z(x) = \frac{dy}{dx} - r_1\,y \tag{3.95}$$

erhalten wir aus Gln. (3.94)

$$\frac{dz}{dx} - r_2\,z(x) = 0. \tag{3.96}$$

Diese Differentialgleichung kann mittels Trennung der Variablen gelöst werden; das Ergebnis ist

$$z(x) = C_2\exp\left(r_2\,x\right), \tag{3.97}$$

wobei C_2 eine komplexe Integrationskonstante ist. Die Gleichungen (3.95) und (3.97) liefern jetzt

$$\frac{dy}{dx} - r_1\,y = C_2\exp\left(r_2\,x\right). \tag{3.98}$$

Diese Gleichung kann durch Variation der Konstanten gelöst werden. Sie hat die Form von Gln. (3.20) mit

$$P(x) = -r_1 \text{ und } Q(x) = C_2 \exp(r_2\, x)\,. \tag{3.99}$$

Das heißt (Gln. (3.26))

$$R(x) = \int P(x)dx = -r_1\, x. \tag{3.100}$$

Gleichung (3.32) ergibt damit

$$\begin{aligned} y &= \left[\int C_2 e^{r_2\, x}\, e^{-r_1\, x} dx + C_1\right] e^{r_1 x} \\ &= C_2\, e^{r_1 x} \int e^{(r_2 - r_1)\, x} dx + C_1\, e^{r_1 x} \end{aligned} \tag{3.101}$$

wobei wir die komplexe Integrationskonstante K von Gln. (3.32) in C_1 umbenannt haben. Für $r_1 = r_2$ erhalten wir

$$y = C_2\, e^{r_1 x} \int dx + C_1\, e^{r_1 x} = (C_1 + C_2\, x)\; e^{r_1 x}, \tag{3.102}$$

während für $r_1 \neq r_2$

$$y = C_1\, e^{r_1 x} + \frac{C_2}{r_2 - r_1} e^{r_2 x}. \tag{3.103}$$

Wenn die Zahlen a, b, c in Gln. (3.86) alle reell sind, möchten wir normalerweise alle reelle Lösungen dieser Differentialgleichung ermitteln. Abhängig vom Vorzeichen der Größe (siehe Gln. (3.91))

$$D = b^2 - 4ac \tag{3.104}$$

müssen wir drei Fälle unterscheiden:

$D = 0$ Wir haben (Gln. (3.91)) $r_1 = r_2 = -b/(2a)$; das heißt r_1 und r_2 sind reell. Die reellen Lösungen der Differentialgleichung sind nach Eq. (3.102)

$$y = (C_1 + C_2\, x)\; e^{r_1 x}, \tag{3.105}$$

wobei C_1, $C_2 \in \mathbb{R}$.

$D > 0$ Nach Gln. (3.91) haben wir $r_1 \neq r_2$; r_1 und r_2 sind jedoch reell. Die reellen Lösungen der Differentialgleichung sind nach Eq. (3.103)

$$y = C_1'\, e^{r_1 x} + C_2'\, e^{r_2 x}, \tag{3.106}$$

wobei C_1', $C_2' \in \mathbb{R}$.

$D < 0$ In diesem Fall sind r_1 und r_2 komplex:

$$r_1 = \frac{-b - i\sqrt{|D|}}{2a} = p - iq \;\text{ und }\; r_2 = \frac{-b + i\sqrt{|D|}}{2a} = p + iq, \tag{3.107}$$

wobei wir die Definitionen

$$p = -\frac{b}{2a} \;\text{ und }\; q = \frac{\sqrt{|D|}}{2a} \tag{3.108}$$

eingeführt haben. Die Lösungen der Differentialgleichung sind nach Eq. (3.103)

$$\begin{aligned} y &= C_1'\, e^{(p-iq)x} + C_2'\, e^{(p+iq)x} \\ &= e^{px}\left[C_1'\, (\cos(-qx) + i\sin(-qx)) + C_2'\, (\cos(qx) + i\sin(qx))\right] \\ &= e^{px}\left[C_1'\, (\cos(qx) - i\sin(qx)) + C_2'\, (\cos(qx) + i\sin(qx))\right] \\ &= e^{px}\left[(C_1' + C_2')\cos(qx) + i\,(-C_1' + C_2')\sin(qx)\right]. \end{aligned} \tag{3.109}$$

Diese Lösungen sind reell, wenn $C_1'' = C_1' + C_2'$ und $C_2'' = i\,(C_2' - C_1')$ reell sind. Dieses wird erreicht für $C_2' = (C_1')^*$. Wir können also die reellen Lösungen als

$$y = e^{px}\left[C_1''\, \cos(qx) + C_2''\, \sin(qx)\right] \tag{3.110}$$

schreiben, wobei C_1'', $C_2'' \in \mathbb{R}$.

Beispiel: Die Differentialgleichung

$$\frac{d^2y}{dx^2} - 3\,\frac{dy}{dx} + 2\,y = 0 \tag{3.111}$$

hat die charakteristische Gleichung

$$r^2 - 3\,r + 2 = 0 \tag{3.112}$$

mit den Lösungen $r_1 = 1$, $r_2 = 2$. Nach Gln. (3.103) hat die Differentialgleichung die Lösungen

$$y = C_1\, e^x + C_2\, e^{2x}. \tag{3.113}$$

wobei C_1 und C_2 willkürliche komplexe Konstanten sind.

Beispiel: Die Differentialgleichung

$$\frac{d^2y}{dx^2} + 8\,\frac{dy}{dx} + 16\,y = 0 \tag{3.114}$$

hat die charakteristische Gleichung

$$r^2 + 8\,r + 16 = 0 \tag{3.115}$$

mit den Lösungen $r_1 = r_2 = -4$. Nach Gln. (3.102) hat die Differentialgleichung die Lösungen

$$y = (C_1 + C_2\,x)\;e^{-4x}, \tag{3.116}$$

wobei C_1 und C_2 willkürliche komplexe Konstanten sind.
Beispiel: Die Differentialgleichung

$$a\,\frac{d^2y}{dt^2} + c\,y = 0 \tag{3.117}$$

hat die charakteristische Gleichung

$$a\,r^2 + c = 0 \tag{3.118}$$

mit den Lösungen

$$r_1 = +\sqrt{-\frac{c}{a}} \;\text{ und }\; r_2 = -\sqrt{-\frac{c}{a}}. \tag{3.119}$$

Nehmen wir an, daß a und c reell sind. Wir haben dann die Möglichkeiten

a) $c = 0$. Dann haben wir $r_1 = r_2 = 0$ und damit nach Gln. (3.102)

$$y(t) = C_1 + C_2\,t, \tag{3.120}$$

wobei C_1 und C_2 willkürliche Konstanten sind.

b) $-c/a > 0$. Dann ist $\omega = \sqrt{-c/a}$ reell, $r_1 = +\omega$ und $r_2 = -\omega$. Nach Gln. (3.103) hat die Differentialgleichung die Lösungen

$$y(t) = C_1\,e^{+\omega t} + C_2\,e^{-\omega t}. \tag{3.121}$$

wobei C_1 und C_2 willkürliche Konstanten sind.

c) $-c/a < 0$. Dann definieren wir $\omega = \sqrt{c/a}$. Die Größe ω ist reell, $r_1 = +i\omega$ und $r_2 = -i\omega$. Nach Gln. (3.103) hat die Differentialgleichung die Lösungen

$$y(t) = C_1\,e^{+i\omega t} + C_2\,e^{-i\omega t}, \tag{3.122}$$

wobei C_1 und C_2 willkürliche Konstanten sind. Wir schreiben

$$\begin{aligned} y(t) &= C_1\,(\cos(\omega t) + i\,\sin(\omega t)) + C_2\,(\cos(-\omega t) + i\,\sin(-\omega t)) \\ &= C_1\,(\cos(\omega t) + i\,\sin(\omega t)) + C_2\,(\cos(\omega t) - i\,\sin(\omega t)) \\ &= (C_1 + C_2)\,\cos(\omega t) + i\,(C_1 - C_2)\,\sin(\omega t). \end{aligned} \tag{3.123}$$

Wir wählen $C_1 = C_2^*$, so daß

$$C_a = C_1 + C_2 \;\text{ und }\; C_b = i\,(C_1 - C_2) \tag{3.124}$$

reel sind. Wir definieren dann zwei weitere reelle Konstanten C und δ, die so bestimmt sind, daß

$$(C_a, C_b) = (C \, \cos\delta, C \, \sin\delta)\,. \tag{3.125}$$

Mit diesen Definitionen gilt

$$y(t) = C \, (\cos\delta \, \cos(\omega t) + \sin\delta \, \sin(\omega t)) = C \, \cos(\omega t - \delta). \tag{3.126}$$

Für $-c/a < 0$ ist Gln. (3.117) die *Schwingungsgleichung*. Die beiden freien Parameter C und δ sind bzw. Schwingungsamplitude und Phasenverschiebung.

Die inhomogene lineare Differentialgleichung 2. Ordnung mit konstanten Koeffizienten

Jede inhomogene lineare Differentialgleichung 2. Ordnung mit konstanten Koeffizienten kann in der Form

$$\frac{d^2y}{dx^2} + b\frac{dy}{dx} + c\,y = s(x), \tag{3.127}$$

geschrieben werden. Diese Gleichung kann dadurch erhalten werden, daß wir in Gln (3.81) auf der rechten und linken Seite durch a dividieren; $a \neq 0$ (wäre $a = 0$, hätten wir eine Differentialgleichung 1. Ordnung). Wir definieren für den späteren Gebrauch

$$L_2[y] = \frac{d^2y}{dx^2} + b\frac{dy}{dx} + c\,y, \tag{3.128}$$

damit die Differentialgleichung als

$$L_2[y] = s(x) \tag{3.129}$$

geschrieben werden kann.

Man kann zeigen, daß sich die allgemeine Lösung einer linearen Differentialgleichung 2. Ordnung aus der allgemeinen Lösung (y_{H}) der zugehörigen homogenen Differentialgleichung und einer partikulären Lösung (y_{P}) der inhomogenen Differentialgleichung zusammensetzt:

$$y = y_{\mathrm{H}} + y_{\mathrm{P}} \tag{3.130}$$

Wir wissen bereits, wie man eine lineare homogene Differentialgleichung 2. Ordnung mit konstanten Koeffizienten löst und damit können wir alle benötigte y_{H} in Gln. (3.130) bestimmen. Für das Lösen einer inhomogenen linearen Differentialgleichung 2. Ordnung mit konstanten Koeffizienten ist es deshalb nur noch erforderlich, ein Verfahren zur Bestimmung einer partikulären Lösung y_{P} bereitzustellen.

Eine allgemeine Lösungsmethode: Es gibt eine relativ allgemeine, aber auch relativ komplizierte Methode, eine partikuläre Lösung der Gln. (3.127) zu bestimmen. Betrachten wir zunächst das homogene Problem

$$\frac{d^2y_{\mathrm{H}}}{dx^2} + b\frac{dy_{\mathrm{H}}}{dx} + c\,y_{\mathrm{H}} = 0. \tag{3.131}$$

Die Lösungen $y_{\rm H}$ sind in Gln. (3.102) und (3.103) gegeben. In beiden dieser Fälle enthält die Lösungsfunktion zwei frei wählbare Konstanten, C_1 und C_2. Wir bestimmen die Werte von C_1 und C_2, so daß wir eine Lösung y_0 der homogenen Gln. (3.131) haben, für die

$$y_0(0) = 0 \;\; \text{und} \;\; y_0'(0) = 1 \tag{3.132}$$

wobei $y_0' = dy_0/dx$ ist. Dann hat die inhomogene Gln. (3.127) die partikuläre Lösung

$$y = \int_0^x y_0(x-t)\, s(t)\, dt. \tag{3.133}$$

Wir werden nachprüfen, daß die Funktion y in Gln. (3.133) eine Lösung der inhomogenen Differentialgleichung ist. Wegen der Tatsache, daß die Variable x in in Gln. (3.133) sowohl explizit im Integranden als auch als obere Integrationsgrenze vorkommt, gestaltet sich die Differentation von y nach x etwas kompliziert. Allgemein gilt für

$$f(x) = \int_0^x g(x,t)\, dt \tag{3.134}$$

daß

$$\frac{df}{dx} = g(x,x) + \int_0^x \frac{\partial g}{\partial x}(x,t)\, dt. \tag{3.135}$$

Für die Funktion y in Gln. (3.133) ergibt dies

$$\frac{dy}{dx} = y_0(0)\, s(x) + \int_0^x y_0'(x-t)\, s(t)\, dt = \int_0^x y_0'(x-t)\, s(t)\, dt \tag{3.136}$$

und

$$\frac{d^2y}{dx^2} = y_0'(0)\, s(x) + \int_0^x y_0''(x-t)\, s(t)\, dt = s(x) + \int_0^x y_0''(x-t)\, s(t)\, dt, \tag{3.137}$$

wobei $y_0'' = d^2y_0/dx^2$ ist. Setzen wir Gln. (3.133), (3.136) and (3.137) in Gln. (3.127) ein, erhalten wir

$$\begin{aligned}
&\frac{d^2y}{dx^2} + b\frac{dy}{dx} + c\,y \\
&= \;\; s(x) + \int_0^x y_0''(x-t)\, s(t)\, dt + b \int_0^x y_0'(x-t)\, s(t)\, dt + c \int_0^x y_0(x-t)\, s(t)\, dt \\
&= \;\; s(x) + \int_0^x \left[y_0''(x-t) + b\, y_0'(x-t) + c y_0(x-t)\right]\, s(t) dt = s(x).
\end{aligned} \tag{3.138}$$

Die Funktion y ist also eine Lösung der inhomogenen Differentialgleichung. In Gln. (3.138) haben wir

$$\left[y_0''(x-t) + b\, y_0'(x-t) + c y_0(x-t)\right] = 0, \tag{3.139}$$

da y_0 eine Lösung der homogenen Differentialgleichung ist.
Lösungen intelligent raten: Um Gln. (3.133) einsetzen zu können, müssen wir die Stammfunktion der Funktion $y_0(x-t)\,s(t)$ (aufgefaßt als eine Funktion von t) bestimmen. Dieses kann unmöglich oder mindestens sehr mühsam sein. Für einige Funktionen $s(x)$ in Gln. (3.127) ist die Form der partikulären Lösungen y_P bekannt. Wir geben hier einige Beispiele für das intelligente Raten von Lösungen.

- Hat eine inhomogene lineare Differentialgleichung die Form

$$L_2[y] = s_1(x) + s_2(x) \tag{3.140}$$

  und ist $y_{\mathrm{p}1}$ eine partikuläre Lösung der Differentialgleichung $L_2[y] = s_1(x)$, $y_{\mathrm{p}2}$ eine partikuläre Lösung der Differentialgleichung $L_2[y] = s_2(x)$, so ist

$$y_\mathrm{p} = y_{\mathrm{p}1} + y_{\mathrm{p}2} \tag{3.141}$$

  eine partikuläre Lösung der Differentialgleichung (3.140).

- Wir betrachten die Differentialgleichung

$$L_2[y] = P_m(x)\,e^{\alpha x} \tag{3.142}$$

  mit

$$P_m(x) = p_0 + p_1\,x + p_2\,x^2 + \ldots + p_m\,x^m. \tag{3.143}$$

  Der Lösungsansatz ist

$$y_\mathrm{P} = B_m(x)\,e^{\alpha x} \tag{3.144}$$

  mit

$$B_m(x) = b_0 + b_1\,x + b_2\,x^2 + \ldots + b_m\,x^m \tag{3.145}$$

  wenn α nicht Lösungswert der charakteristischen Gleichung der zugehörigen homogenen Differentialgleichung ist. Der Grad des Polynoms $B_m(x)$ ist der gleiche wie der des Polynoms $P_m(x)$ in $s(x)$. Der Lösungsansatz ist

$$y_\mathrm{P} = x^q\,B_m(x)\,e^{\alpha x}, \tag{3.146}$$

  wenn α der Wert einer q-fachen Lösung der charakteristischen Gleichung ist.

  Die Koeffizienten b_0, b_1, b_2, $\ldots$, b_m sind durch Koeffizientenvergleich nach einsetzen von y_P, y_P' und y_P'' in die zu lösende inhomogene Differentialgleichung $L_2[y] = P_m(x)\,e^{\alpha x}$ zu bestimmen.

- Wir betrachten die Differentialgleichung

$$L_2[y] = P_{m1}(x)\,\cos(\beta x) + P_{m2}(x)\,\sin(\beta x). \tag{3.147}$$

Der Lösungsansatz ist

$$y_\mathrm{P} = B_m(x)\,\cos(\beta x) + C_m(x)\,\sin(\beta x) \tag{3.148}$$

mit

$$\begin{aligned} B_m(x) &= b_0 + b_1\,x + b_2\,x^2 + \ldots + b_m\,x^m \\ C_m(x) &= c_0 + c_1\,x + c_2\,x^2 + \ldots + c_m\,x^m, \end{aligned} \tag{3.149}$$

wenn $\pm i\beta$ nicht Lösung der charakteristischen Gleichung der zugehörigen homogenen Differentialgleichung ist. Sind $P_{m1}(x)$ und $P_{m2}(x)$ Polynome von höchstens m-tem Grade, so sind für $B_m(x)$ und $C_m(x)$ Polynome m-ten Grades anzusetzen. Der Lösungsansatz ist

$$y_\mathrm{P} = x^q\,[B_m(x)\,\cos(\beta x) + C_m(x)\,\sin(\beta x)] \tag{3.150}$$

wenn $i\beta$ und $-i\beta$ q-fache Lösungen der charakteristischen Gleichung ist.

Die Koeffizienten b_0, b_1, b_2, ..., b_m und c_0, c_1, c_2, ..., c_m sind durch Koeffizientenvergleich nach einsetzen von y_P, y'_P und y''_P in die zu lösende inhomogene Differentialgleichung $L_2[y] = B_m(x)\,\cos(\beta x) + C_m(x)\,\sin(\beta x)$ zu bestimmen.

- Wir betrachten die Differentialgleichung

$$L_2[y] = e^{\alpha x}\,[P_{m1}(x)\,\cos(\beta x) + P_{m2}(x)\,\sin(\beta x)]\,. \tag{3.151}$$

Der Lösungsansatz ist

$$y_\mathrm{P} = e^{ax}\,[B_m(x)\,\cos(\beta x) + C_m(x)\,\sin(\beta x)] \tag{3.152}$$

wenn $\alpha \pm i\beta$ nicht Lösung der charakteristischen Gleichung der zugehörigen homogenen Differentialgleichung ist. Der Lösungsansatz ist

$$y_\mathrm{P} = x^q e^{ax}\,[B_m(x)\,\cos(\beta x) + C_m(x)\,\sin(\beta x)] \tag{3.153}$$

wenn $\alpha \pm i\beta$ q-fache Lösung der charakteristischen Gleichung ist.

Beispiel: Wir lösen die Differentialgleichung

$$\frac{d^2y}{dx^2} + y = \sin x \tag{3.154}$$

mit $s(x) = \sin x$. Die entsprechende homogene Differentialgleichung

$$\frac{d^2 y_\mathrm{H}}{dx^2} + y_\mathrm{H} = 0 \tag{3.155}$$

hat die charakteristische Gleichung

$$r^2 + 1 = 0 \tag{3.156}$$

mit den Lösungen $r_1 = -i$, $r_2 = +i$. Nach Gln. (3.103) hat die homogene Differentialgleichung die Lösungen

$$y_\mathrm{H} = C_1 \, e^{-ix} + C_2 \, e^{+ix}. \tag{3.157}$$

wobei C_1 und C_2 willkürliche komplexe Konstanten sind.

a) Wir lösen zunächst die inhomogene Differentialgleichung mit Hilfe von Gln. (3.133). Die besondere Lösung y_0 der homogenen Differentialgleichung, die wir für die Anwendung von Gln (3.133) benötigen, erfüllt Gln. (3.132). Mit

$$y_\mathrm{H}' = \frac{dy_\mathrm{H}}{dx} = -i \, C_1 \, e^{-ix} + i \, C_2 \, e^{+ix}. \tag{3.158}$$

haben wir also

$$y_\mathrm{H}(0) = C_1 \, e^{-i0} + C_2 \, e^{+i0} = C_1 + C_2 = 0 \tag{3.159}$$

und

$$y_\mathrm{H}'(0) = -i \, C_1 \, e^{-i0} + i \, C_2 \, e^{+i0} = i \, (C_2 - C_1) = 1. \tag{3.160}$$

Es folgt daß

$$C_1 = -C_2 = -\frac{1}{2i} \tag{3.161}$$

und damit haben wir von Gln. (3.157)

$$y_0 = -\frac{1}{2i} \left(e^{-ix} - e^{+ix} \right) = \sin x. \tag{3.162}$$

Wir können jetzt in Gln (3.133) einsetzen. Wir haben $s(x) = \sin x$ [Gln. (3.154)] und $y_0(x) = \sin x$ [Gln. (3.162)]. Damit hat die inhomogene Differentialgleichung (3.154) die partikuläre Lösung

$$\begin{aligned} y_\mathrm{P} &= \int_0^x \sin(x-t) \, \sin t \, dt \\ &= \frac{1}{2} \int_0^x \left(\cos(x-2t) - \cos x \right) dt \\ &= \left[-\frac{1}{4} \sin(x-2t) - \frac{1}{2} \, t \, \cos x \right]_0^x \\ &= \frac{1}{2} \sin x - \frac{1}{2} \, x \, \cos x. \end{aligned} \tag{3.163}$$

Wir haben hier die allgemeine Relation

$$\sin\alpha\,\sin\beta = \frac{1}{2}\,\cos(\alpha-\beta) - \frac{1}{2}\,\cos(\alpha+\beta) \qquad (3.164)$$

benutzt. Nach Gln. (3.130) sind sämtliche Lösungen der inhomogenen Differentialgleichung also

$$\begin{aligned} y &= y_\mathrm{H} + y_\mathrm{P} \\ &= C_1\,e^{-ix} + C_2\,e^{+ix} + \frac{1}{2}\sin x - \frac{1}{2}\,x\,\cos x \\ &= C_1'\,e^{-ix} + C_2'\,e^{+ix} - \frac{1}{2}\,x\,\cos x, \end{aligned} \qquad (3.165)$$

wobei $C_1' = C_1 - 1/(4i)$ und $C_2' = C_2 + 1/(4i)$ willkürliche komplexe Konstanten sind. Wir benutzten

$$\sin x = \frac{1}{2i}\left(e^{ix} - e^{-ix}\right). \qquad (3.166)$$

b) Die Funktion $s(x) = \sin x$ hat die Form von Gln. (3.147) mit $\beta = 1$. Das heißt, $i\beta$ ist eine 1-fache Lösung der charakteristischen Gleichung der zugehörigen homogenen Differentialgleichung, und wir erhalten den Lösungsansatz von Gln. (3.150)

$$y_\mathrm{P} = x\,(b_0\,\sin x + c_0\,\cos x)\,. \qquad (3.167)$$

Dann sind

$$\begin{aligned} \frac{dy_\mathrm{P}}{dx} &= b_0\,\sin x + c_0\,\cos x + x\,(b_0\,\cos x - c_0\,\sin x)\,, \\ \frac{d^2y_\mathrm{P}}{dx^2} &= 2\,(b_0\,\cos x - c_0\,\sin x) - x\,(b_0\,\sin x + c_0\,\cos x) \end{aligned} \qquad (3.168)$$

und damit

$$\frac{d^2y_\mathrm{P}}{dx^2} + y_\mathrm{P} = 2\,(b_0\,\cos x - c_0\,\sin x)\,. \qquad (3.169)$$

Die Differentialgleichung (3.154) ist erfüllt für $b_0 = 0$, $c_0 = -1/2$, und wir haben die partikuläre Lösung

$$y_\mathrm{P} = -\frac{1}{2}\,x\,\cos x \qquad (3.170)$$

gegeben in Gln. (3.165) für $C_1' = C_2' = 0$.

Beispiel: Wir lösen die Differentialgleichung

$$\frac{d^2y}{dx^2} - \frac{dy}{dx} = x \tag{3.171}$$

mit $s(x) = x$. Die entsprechende homogene Differentialgleichung

$$\frac{d^2y_\mathrm{H}}{dx^2} - \frac{dy_\mathrm{H}}{dx} = 0 \tag{3.172}$$

hat die charakteristische Gleichung

$$r^2 - r = 0 \tag{3.173}$$

mit den Lösungen $r_1 = 0$, $r_2 = 1$. Nach Gln. (3.103) hat die homogene Differentialgleichung die Lösungen

$$y_\mathrm{H} = C_1\, e^x + C_2. \tag{3.174}$$

wobei C_1 und C_2 willkürliche komplexe Konstanten sind.

a) Wir lösen zunächst die inhomogene Differentialgleichung mit Hilfe von Gln. (3.133). Mit

$$y'_\mathrm{H} = \frac{dy_\mathrm{H}}{dx} = C_1\, e^x \tag{3.175}$$

haben wir für die besondere Lösung y_0

$$y_0(0) = C_1\, e^0 + C_2 = C_1 + C_2 = 0 \tag{3.176}$$

und

$$y'_0(0) = C_1\, e^0 = C_1 = 1. \tag{3.177}$$

Es folgt daß

$$C_1 = -C_2 = 1 \tag{3.178}$$

und damit

$$y_0 = e^x - 1. \tag{3.179}$$

Wir können jetzt in Gln (3.133) einsetzen. Wir haben $s(x) = x$ und $y_0(x) = e^x - 1$. Das heißt, die inhomogene Differentialgleichung (3.171) hat die partikuläre Lösung

$$\begin{aligned} y_\mathrm{P} &= \int_0^x \left(e^{x-t} - 1\right) t\, dt \\ &= e^x \int_0^x t\, e^{-t}\, dt - \int_0^x t\, dt \\ &= e^x \left[-(t+1)\, e^{-t}\right]_0^x - \left[\frac{1}{2}\, t^2\right]_0^x \\ &= e^x \left(-(x+1)\, e^{-x} + 1\right) - \frac{1}{2}\, x^2 \\ &= -\frac{1}{2}\, x^2 - x - 1 + e^x. \end{aligned} \tag{3.180}$$

Nach Gln. (3.130) sind sämtliche Lösungen der inhomogenen Differentialgleichung also

$$y = y_{\mathrm{H}} + y_{\mathrm{P}} = C_1\, e^x + C_2 - \frac{1}{2}\, x^2 - x - 1 + e^x = -\frac{1}{2}\, x^2 - x + c_1\, e^x + c_2, \quad (3.181)$$

wobei $c_1 = C_1 + 1$ und $c_2 = C_2 - 1$ willkürliche komplexe Konstanten sind.

b) Wir haben hier $s(x) = x = P_m(x)\, e^{\alpha x}$, wobei $P_m(x) = x$ ($m = 1$) und $\alpha = 0$ sind. Damit ist α der Wert einer 1-fachen Lösung der charakteristischen Gleichung, und wir benutzen den Lösungsansatz von Gln. (3.146)

$$y_{\mathrm{P}} = x\ (b_0 + b_1\, x) = b_0\, x + b_1\, x^2. \quad (3.182)$$

Dann haben wir

$$\frac{dy_{\mathrm{P}}}{dx} = b_0 + 2\, b_1\, x \quad \text{und} \quad \frac{d^2 y_{\mathrm{P}}}{dx^2} = 2\, b_1 \quad (3.183)$$

und damit

$$\frac{d^2 y_{\mathrm{P}}}{dx^2} - \frac{dy_{\mathrm{P}}}{dx} = 2\, b_1 - b_0 - 2\, b_1\, x. \quad (3.184)$$

Folglich ist Gln. (3.171) erfüllt für $b_1 = -1/2$, $b_0 = -1$, und wir erhalten die partikuläre Lösung

$$y_{\mathrm{P}} = -\frac{1}{2}\, x^2 - x \quad (3.185)$$

gegeben in Gln. (3.181) für $c_1 = c_2 = 0$.

Beispiel: Wir lösen die Differentialgleichung

$$\frac{d^2 y}{dx^2} - 2\, y = \frac{1}{4}\, e^{2x} \quad (3.186)$$

mit $s(x) = e^{2x}/4$. Die entsprechende homogene Differentialgleichung

$$\frac{d^2 y_{\mathrm{H}}}{dx^2} - 2\, y_{\mathrm{H}} = 0 \quad (3.187)$$

hat die charakteristische Gleichung

$$r^2 - 2 = 0 \quad (3.188)$$

mit den Lösungen $r_1 = -\sqrt{2}$, $r_2 = +\sqrt{2}$. Nach Gln. (3.103) hat die homogene Differentialgleichung die Lösungen

$$y_{\mathrm{H}} = C_1\, e^{-\sqrt{2}x} + C_2\, e^{+\sqrt{2}x}. \quad (3.189)$$

wobei C_1 und C_2 willkürliche komplexe Konstanten sind.

a) Wir lösen zunächst die inhomogene Differentialgleichung mit Hilfe von Gln. (3.133). Mit

$$y_{\mathrm{H}}' = \frac{dy_{\mathrm{H}}}{dx} = -\sqrt{2}\, C_1\, e^{-\sqrt{2}x} + \sqrt{2}\, C_2\, e^{+\sqrt{2}x} \tag{3.190}$$

haben wir für die besondere Lösung y_0

$$y_0(0) = C_1\, e^{-\sqrt{2}0} + C_2\, e^{+\sqrt{2}0} = C_1 + C_2 = 0 \tag{3.191}$$

und

$$y_0'(0) = -\sqrt{2}\, C_1\, e^{-\sqrt{2}0} + \sqrt{2}\, C_2\, e^{+\sqrt{2}0} = \sqrt{2}\,(C_2 - C_1) = 1. \tag{3.192}$$

Es folgt daß

$$C_1 = -C_2 = -\frac{1}{2\sqrt{2}} \tag{3.193}$$

und damit

$$y_0 = -\frac{1}{2\sqrt{2}} \left(e^{-\sqrt{2}x} - e^{+\sqrt{2}x} \right). \tag{3.194}$$

Wir können jetzt in Gln (3.133) einsetzen. Wir haben $s(x) = e^{2x}/4$ und $y_0(x) = -(e^{-\sqrt{2}x} - e^{+\sqrt{2}x})/(2\sqrt{2})$. Das heißt, die inhomogene Differentialgleichung (3.186) hat die partikuläre Lösung

$$\begin{aligned}
y_{\mathrm{P}} &= -\frac{1}{8\sqrt{2}} \int_0^x \left(e^{-\sqrt{2}(x-t)} - e^{+\sqrt{2}(x-t)} \right) e^{2t}\, dt \\
&= -\frac{1}{8\sqrt{2}} \left\{ e^{-\sqrt{2}x} \int_0^x e^{(2+\sqrt{2})t}\, dt - e^{+\sqrt{2}x} \int_0^x e^{(2-\sqrt{2})t}\, dt \right\} \\
&= -\frac{1}{8\sqrt{2}} \left\{ e^{-\sqrt{2}x} \left[\frac{1}{2+\sqrt{2}}\, e^{(2+\sqrt{2})t} \right]_0^x - e^{+\sqrt{2}x} \left[\frac{1}{2-\sqrt{2}}\, e^{(2-\sqrt{2})t} \right]_0^x \right\} \\
&= -\frac{1}{8\sqrt{2}} \left\{ e^{-\sqrt{2}x}\, \frac{1}{2+\sqrt{2}} \left(e^{(2+\sqrt{2})x} - 1 \right) - e^{+\sqrt{2}x}\, \frac{1}{2-\sqrt{2}} \left(e^{(2-\sqrt{2})x} - 1 \right) \right\} \\
&= -\frac{1}{8\sqrt{2}} \left\{ \left(\frac{1}{2+\sqrt{2}} - \frac{1}{2-\sqrt{2}} \right) e^{2x} - \frac{1}{2+\sqrt{2}}\, e^{-\sqrt{2}x} + \frac{1}{2-\sqrt{2}}\, e^{+\sqrt{2}x} \right\} \\
&= \frac{1}{8}\, e^{2x} + \frac{1}{8\sqrt{2}} \left\{ \frac{1}{2+\sqrt{2}}\, e^{-\sqrt{2}x} - \frac{1}{2-\sqrt{2}}\, e^{+\sqrt{2}x} \right\}.
\end{aligned} \tag{3.195}$$

Nach Gln. (3.130) sind sämtliche Lösungen der inhomogenen Differentialgleichung also

$$\begin{aligned}
y &= y_{\mathrm{H}} + y_{\mathrm{P}} \\
&= C_1\, e^{-\sqrt{2}x} + C_2\, e^{+\sqrt{2}x} + \frac{1}{8}\, e^{2x} + \frac{1}{8\sqrt{2}} \left\{ \frac{1}{2+\sqrt{2}}\, e^{-\sqrt{2}x} - \frac{1}{2-\sqrt{2}}\, e^{+\sqrt{2}x} \right\} \\
&= c_1\, e^{-\sqrt{2}x} + c_2\, e^{+\sqrt{2}x} + \frac{1}{8}\, e^{2x}
\end{aligned} \tag{3.196}$$

wobei

$$c_1 = C_1 + \frac{1}{8\sqrt{2}(2+\sqrt{2})} \text{ und } c_2 = C_2 - \frac{1}{8\sqrt{2}(2-\sqrt{2})} \tag{3.197}$$

willkürliche komplexe Konstanten sind.

b) Mit $s(x) = e^{2x}/4$ erhalten wir von Gln. (3.144) ($\alpha = 2$ ist *nicht* Lösungswert der charakteristischen Gleichung) den Lösungsansatz

$$y_\mathrm{P} = b_0\, e^{2x}. \tag{3.198}$$

Dann haben wir

$$\frac{dy_\mathrm{P}}{dx} = 2\, b_0\, e^{2x} \text{ und } \frac{d^2 y_\mathrm{P}}{dx^2} = 4\, b_0\, e^{2x}. \tag{3.199}$$

Durch Einsetzen in Gln. (3.186) erhalten wir

$$4\, b_0\, e^{2x} - 2\, b_0\, e^{2x} = \frac{1}{4}\, e^{2x}. \tag{3.200}$$

Diese Gleichung ist erfüllt für $b_0 = 1/8$ und wir erhalten in relativ einfacher Weise die partikuläre Lösung

$$y_\mathrm{P} = \frac{1}{8}\, e^{2x}, \tag{3.201}$$

die in Gln. (3.196) für $c_1 = c_2 = 0$ gegeben ist.

4 Ausgleichsrechnung

4.1 Grundbegriffe

Im Abschnitt 6.2 des Vorlesungsskriptes zur "Mathematik A für Chemiker und Lebensmittelchemiker" haben wir bereits die Situation diskutiert, bei der wir n Messungen der Größe y durchgeführt und die Meßwerte $(a_1, a_2, \ldots, a_n)$ erhalten haben. Wir nehmen an, daß jede Messung einer Gauss'schen Normalverteilung unterliegt. Ferner wird angenommen, daß alle Messungen dieselbe Präzision haben. Diese Präzision wird durch die Standardabweichung σ_i der Normalverteilung von der i-ten Messung ausgedrückt. Es gilt damit $\sigma_1 = \sigma_2 = \ldots = \sigma_n = \sigma$. Nach Gauß bestimmen wir den *wahrscheinlichsten Wert* von y, $\hat{y}$, durch Minimierung der Summe der Fehlerquadrate

$$S(\hat{y}) = \sum_{i=1}^{n} (\hat{y} - a_i)^2 . \tag{4.1}$$

Wir zeigten, daß $S(\hat{y})$ für

$$\hat{y} = \frac{1}{n} \sum_{i=1}^{n} a_i \tag{4.2}$$

den kleinstmöglichen Wert erhält. Das arithmetische Mittel

$$\bar{a} = \frac{1}{n} \sum_{i=1}^{n} a_i \tag{4.3}$$

der einzelnen Meßpunkte a_i ist derjenige Wert, der mit größter Wahrscheinlichkeit dem wahren Wert y entspricht.

Die Bestimmung von optimalen Parameterwerten durch Minimierung der Größe $S(\hat{y})$ wird oft als eine *Ausgleichsrechnung* oder als eine *Anpassung nach der Methode der kleinsten Fehlerquadrate* bezeichnet.

Es kann sein, daß die n Messungen von unterschiedlicher Präzision sind, indem die n σ-Werte σ_1, σ_2, $\ldots$, σ_n nicht alle gleich sind. In diesem Fall bestimmen wir $\hat{y}$ durch Minimierung der Größe

$$S(\hat{y}) = \sum_{i=1}^{n} \left(\frac{\hat{y} - a_i}{\sigma_i} \right)^2 = \sum_{i=1}^{n} p_i \, (\hat{y} - a_i)^2 , \tag{4.4}$$

wobei der *Gewichtfaktor*

$$p_i = \frac{1}{\sigma_i^2} \tag{4.5}$$

ist. Wir berechnen

$$\frac{dS}{d\hat{y}} = 2 \cdot \sum_{i=1}^{n} p_i \, (\hat{y} - a_i) \tag{4.6}$$

und

$$\frac{d^2S}{d\hat{y}^2} = 2 \cdot \sum_{i=1}^{n} p_i > 0. \tag{4.7}$$

Folglich besitzt $S(\hat{y})$ ein Minimum für

$$\frac{dS}{d\hat{y}} = 2 \cdot \sum_{i=1}^{n} p_i \, (\hat{y} - a_i) = 0 \tag{4.8}$$

und Auflösung nach $\hat{y}$ ergibt

$$\hat{y} = \frac{\sum_{i=1}^{n} p_i \cdot a_i}{\sum_{i=1}^{n} p_i}. \tag{4.9}$$

In einer von Gauß eingeführten Notation schreiben wir

$$[pa] = \sum_{i=1}^{n} p_i \cdot a_i \quad \text{und} \quad [p] = \sum_{i=1}^{n} p_i \tag{4.10}$$

und haben damit

$$\hat{y} = \frac{[pa]}{[p]}. \tag{4.11}$$

4.2 Ausgleich bedingter Beobachtungen

Als Beispiel für bedingte Beobachtungen betrachten wir drei Winkel α, β, γ. Wir haben n_α Messungen von α durchgeführt und die Ergebnisse $a_1, a_2, \ldots, a_{n_\alpha}$ erhalten. Analogerweise haben wir n_β Messungen von β (mit den Ergebnissen b_1, b_2, $\ldots$, b_{n_β}) und n_γ Messungen von β (mit den Ergebnissen c_1, c_2, $\ldots$, c_{n_γ}) durchgeführt. Wir nehmen an, daß alle Messungen von gleicher Präzision sind. Sind nun α, β und γ unabhängig voneinander, bestimmen wir ihre optimalen Werte durch Minimierung der Funktion

$$S(\alpha, \beta, \gamma) = \sum_{i=1}^{n_\alpha} (a_i - \alpha)^2 + \sum_{i=1}^{n_\beta} (b_i - \beta)^2 + \sum_{i=1}^{n_\gamma} (c_i - \gamma)^2 . \tag{4.12}$$

Es kann sein, daßα, β und γ nicht unabhängig voneinander sind, sondern eine Bedingung erfüllen. Als Beispiel betrachten wir die Bedingung

$$\phi(\alpha, \beta, \gamma) = \alpha + \beta + \gamma - 180^\circ = 0. \tag{4.13}$$

Das heißt, die Summe der drei Winkel ist immer 180°.

Die optimalen Winkelwerte $\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}$ können am einfachsten dadurch bestimmt werden, daß wir Gln. (4.13) in Gln. (4.12) einsetzen, zum Beispiel durch Eliminierung von γ:

$$S(\alpha,\beta) = \sum_{i=1}^{n_\alpha} (a_i - \alpha)^2 + \sum_{i=1}^{n_\beta} (b_i - \beta)^2 + \sum_{i=1}^{n_\gamma} (c_i - 180^\circ + \alpha + \beta)^2 \tag{4.14}$$

Wir bestimmen dann $\hat{\alpha}$ und $\hat{\beta}$ aus den Gleichungen

$$\frac{\partial S}{\partial \alpha} = -\sum_{i=1}^{n_\alpha} (a_i - \hat{\alpha}) + \sum_{i=1}^{n_\gamma} \left(c_i - 180^\circ + \hat{\alpha} + \hat{\beta}\right) = 0 \tag{4.15}$$

und

$$\frac{\partial S}{\partial \beta} = -\sum_{i=1}^{n_\beta} \left(b_i - \hat{\beta}\right) + \sum_{i=1}^{n_\gamma} \left(c_i - 180^\circ + \hat{\alpha} + \hat{\beta}\right) = 0 \tag{4.16}$$

zusammen mit

$$\hat{\gamma} = 180^\circ - \hat{\alpha} - \hat{\beta}. \tag{4.17}$$

Der Vollständigkeit halber soll hier auch erwähnt werden, daß es eine weitere Möglichkeit gibt, den Ausgleich bedingter Beobachtungen durchzuführen. Das mathematische Problem, bei welchem man eine Funktion $f(x_1, x_2, \ldots, x_n)$ (in unserem Beispiel die Funktion $S(\alpha,\beta,\gamma)$ in Gln. (4.12)) unter der Bedingung minimiert, daß die Variablen $x_1, x_2, \ldots, x_n$, von denen die Funktion abhängt (in unserem Beispiel die Winkel α, β, γ), einen Satz von m zusätzlichen Gleichungen $g_i(x_1, x_2, \ldots, x_n) = 0$ ($i = 1, 2, \ldots, m$), erfüllen müssen (in unserem Beispiel die eine Gleichung (4.13)), ist in der mathematischen Literatur ausgiebig behandelt worden. Oft löst man das hier aufgestellte Problem durch Anwendung sogenannter *Lagrangescher Multiplikatoren*[1]. Lagrange hat gezeigt, daß wir durch Minimierung der Funktion

$$\begin{aligned} T(x_1, x_2, \ldots, x_n) &= f(x_1, x_2, \ldots, x_n) + \lambda_1 \cdot g_1(x_1, x_2, \ldots, x_n) \\ &+ \lambda_2 \cdot g_2(x_1, x_2, \ldots, x_n) + \ldots + \lambda_m \cdot g_m(x_1, x_2, \ldots, x_n), \end{aligned} \tag{4.18}$$

wobei wir $x_1, x_2, \ldots, x_n$ als unabhängige Variable auffassen können, das Minimum der Funktion f unter Berücksichtigung der Bedingungsgleichungen $g_i(x_1, x_2, \ldots, x_n) = 0$ ($i = 1, 2, \ldots, m$) ermitteln können. Die anfänglich unbekannten Grössen λ_1, λ_2, ..., λ_m sind die Lagrange'schen Multiplikatoren.

[1] Diese Technik ist im Buch von Zachmann [H.G. Zachmann, *Mathematik für Chemiker*, Verlag Chemie GmbH, Weinheim/Bergstr., 1974] auf den Seiten 344-347 behandelt.

Im vorliegenden Beispiel haben wir einen einzigen Lagrange'schen Multiplikator λ und wir sollen die Funktion

$$\begin{aligned} T(\alpha, \beta, \gamma) &= S(\alpha, \beta, \gamma) + \lambda \cdot \phi(\alpha, \beta, \gamma) \\ &= \sum_{i=1}^{n_\alpha} (a_i - \alpha)^2 + \sum_{i=1}^{n_\beta} (b_i - \beta)^2 \\ &+ \sum_{i=1}^{n_\gamma} (c_i - \gamma)^2 + \lambda \cdot (\alpha + \beta + \gamma - 180^\circ) \end{aligned} \tag{4.19}$$

minimieren. Im Minimumspunkt gelten die Beziehungen

$$\frac{\partial S}{\partial \alpha} = -\sum_{i=1}^{n_\alpha} (a_i - \hat{\alpha}) + \lambda = 0, \tag{4.20}$$

$$\frac{\partial S}{\partial \beta} = -\sum_{i=1}^{n_\beta} \left(b_i - \hat{\beta}\right) + \lambda = 0 \tag{4.21}$$

und

$$\frac{\partial S}{\partial \gamma} = -\sum_{i=1}^{n_\gamma} (c_i - \hat{\gamma}) + \lambda = 0. \tag{4.22}$$

Gleichung (4.22) liefert zusammen mit Gln. (4.13)

$$\lambda = \sum_{i=1}^{n_\gamma} \left(c_i - 180^\circ + \hat{\alpha} + \hat{\beta}\right) \tag{4.23}$$

und durch Einsetzen in Gln. (4.20) und (4.21) erhalten wir wieder Gln. (4.15) und (4.16).

4.3 Ausgleich linearer Zusammenhänge

Zwischen x und y besteht der Zusammenhang

$$y = \alpha + \beta \cdot x. \tag{4.24}$$

Wir führen n Messungen von Wertepaaren (x_i, a_i) von (x, y) durch, wobei die i-te Messung den Gewichtfaktor p_i hat, und möchten nun die optimalen Werte der Parameter α und β bestimmen. Der Parameter β wird manchmal *Regressionskoeffizient* benannt.

Wir ermitteln die optimalen Werte von α und β (genannt $\hat{\alpha}$ und $\hat{\beta}$) durch Minimierung der Summe der Fehlerquadrate

$$S(\hat{\alpha},\hat{\beta}) = \sum_{i=1}^{n} p_i\,(\hat{y}_i - a_i)^2 = \sum_{i=1}^{n} p_i \left(\hat{\alpha} + \hat{\beta}\cdot x_i - a_i\right)^2 \tag{4.25}$$

wobei

$$\hat{y}_i = \hat{\alpha} + \hat{\beta}\cdot x_i \tag{4.26}$$

ist. Die Ableitungen der Fehlerquadrat-Summe nach $\hat{\alpha}$ und $\hat{\beta}$ sind

$$\frac{dS}{d\hat{\alpha}} = 2\cdot\sum_{i=1}^{n} p_i \left(\hat{\alpha} + \hat{\beta}\cdot x_i - a_i\right) \tag{4.27}$$

$$\frac{dS}{d\hat{\beta}} = 2\cdot\sum_{i=1}^{n} p_i\cdot x_i\cdot \left(\hat{\alpha} + \hat{\beta}\cdot x_i - a_i\right). \tag{4.28}$$

Im gesuchten Minimum der Funktion $S(\hat{\alpha},\hat{\beta})$ verschwinden beide Ableitungen. Damit haben wir

$$\frac{1}{2}\frac{dS}{d\hat{\alpha}} = \hat{\alpha}\cdot\sum_{i=1}^{n} p_i + \hat{\beta}\cdot\sum_{i=1}^{n} p_i\,x_i - \sum_{i=1}^{n} p_i\,a_i = 0 \tag{4.29}$$

$$\frac{1}{2}\frac{dS}{d\hat{\beta}} = \hat{\alpha}\cdot\sum_{i=1}^{n} p_i\cdot x_i + \hat{\beta}\cdot\sum_{i=1}^{n} p_i\cdot x_i^2 - \sum_{i=1}^{n} p_i\,a_i\,x_i = 0. \tag{4.30}$$

Nach Gauß schreiben wir diese Gleichungen als

$$\hat{\alpha}\cdot[p] + \hat{\beta}\cdot[px] = [pa] \tag{4.31}$$

$$\hat{\alpha}\cdot[px] + \hat{\beta}\cdot[px^2] = [pax] \tag{4.32}$$

wobei, für die allgemeine Größe z,

$$[z] = \sum_{i=1}^{n} z_i \tag{4.33}$$

ist. Die Gleichungen (4.31)-(4.32) können als eine Matrixgleichung geschrieben werden:

$$\begin{pmatrix} [p] & [px] \\ [px] & [px^2] \end{pmatrix} \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} [pa] \\ [pax] \end{pmatrix}. \tag{4.34}$$

und die Koeffizientenmatrix

$$\mathbf{A} = \begin{pmatrix} [p] & [px] \\ [px] & [px^2] \end{pmatrix} \tag{4.35}$$

kann nach Gln. (2.125) invertiert werden. Das Ergebnis ist

$$\mathbf{A}^{-1} = \frac{1}{[p]\,[px^2] - [px]^2} \begin{pmatrix} [px^2] & -[px] \\ -[px] & [p] \end{pmatrix}. \tag{4.36}$$

Also gilt

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \frac{1}{[p]\,[px^2] - [px]^2} \begin{pmatrix} [px^2] & -[px] \\ -[px] & [p] \end{pmatrix} \begin{pmatrix} [pa] \\ [pax] \end{pmatrix} \tag{4.37}$$

oder

$$\hat{\alpha} = \frac{[px]\,[pax] - [px^2]\,[pa]}{[px]^2 - [p]\,[px^2]} \tag{4.38}$$

$$\hat{\beta} = \frac{[px]\,[pa] - [p]\,[pax]}{[px]^2 - [p]\,[px^2]}. \tag{4.39}$$

Haben wir $p_1 = p_2 = p_3 = \ldots = p_n = p_0$, gilt $[p] = n \cdot p_0$. Wir erhalten von Gln. (4.38) und (4.39)

$$\hat{\alpha} = \frac{[x]\,[ax] - [x^2]\,[a]}{[x]^2 - n\,[x^2]} \tag{4.40}$$

$$\hat{\beta} = \frac{[x]\,[a] - n\,[ax]}{[x]^2 - n\,[x^2]}. \tag{4.41}$$

Der mittlere Fehler für eine Einzelmessung vom Gewicht p_i ist

$$m_i = \frac{m}{\sqrt{p_i}} \tag{4.42}$$

wobei

$$\begin{aligned} m &= \sqrt{\frac{\sum_{k=1}^{n} p_k \,(a_k - \bar{y}_k)^2}{n-2}} \\ &= \sqrt{\frac{\sum_{k=1}^{n} p_k \left(a_k - \hat{\alpha} - \hat{\beta} \cdot x_k\right)^2}{n-2}} \end{aligned} \tag{4.43}$$

ist. Die mittleren Fehler der beiden Parameter $\hat{\alpha}$ und $\hat{\beta}$ sind gegeben als

$$m_{\hat{\alpha}} = m \sqrt{\frac{[px^2]}{[p][px^2] - [px]^2}}, \tag{4.44}$$

$$m_{\hat{\beta}} = m \sqrt{\frac{[p]}{[p][px^2] - [px]^2}}. \tag{4.45}$$

4.4 Ausgleich nichtlinearer Zusammenhänge

Der meßbare Größe y ist gegeben als

$$y = \beta_0\, z_0 + \beta_1\, z_1 + \ldots + \beta_k\, z_k \tag{4.46}$$

wobei z_0, z_1, …, z_k andere meßbare Größen sind. Wir führen n Einzelmessungen von jeweils einem Datensatz $z_{0,i}$, $z_{1,i}$, …, $z_{k,i}$, a_i ($i = 1, 2, \ldots, n$) durch, wobei die i-te Messung den Gewichtfaktor p_i besitzt. Die optimalen Parameterwerte $\hat{\beta}_0$, $\hat{\beta}_1$, …, $\hat{\beta}_k$ werden durch Minimierung der Größe

$$S = \sum_{i=1}^{n} p_i\, \left(a_i - \beta_0\, z_{0,i} - \beta_1\, z_{1,i} - \ldots - \beta_k\, z_{k,i}\right)^2 \tag{4.47}$$

ermittelt. Am Minimumspunkt gilt

$$\frac{\partial S}{\partial \beta_0} = \frac{\partial S}{\partial \beta_1} = \ldots = \frac{\partial S}{\partial \beta_k} = 0. \tag{4.48}$$

Berechnen wir die partiellen Ableitungen, erhalten wir die k *Normalgleichungen*

$$\begin{aligned}
\beta_0\, [pz_0^2] + \beta_1\, [pz_0 z_1] + \beta_2\, [pz_0 z_2] + \ldots + \beta_k\, [pz_0 z_k] &= [paz_0] \\
\beta_0\, [pz_1 z_0] + \beta_1\, [pz_1^2] + \beta_2\, [pz_1 z_2] + \ldots + \beta_k\, [pz_1 z_k] &= [paz_1] \\
&\vdots \\
\beta_0\, [pz_k z_0] + \beta_1\, [pz_k z_1] + \beta_2\, [pz_k z_2] + \ldots + \beta_k\, [pz_k^2] &= [paz_k].
\end{aligned} \tag{4.49}$$

Lösung dieses linearen Gleichungssystem liefert die optimalen Parameterwerte.

Ein oft vorkommender Fall hat $z_0 = 1$, $z_1 = x$, $z_2 = x^2$, …, $z_k = x^k$. Die entsprechenden Normalgleichungen sind

$$\begin{aligned}
\beta_0\, [p] + \beta_1\, [px] + \beta_2\, [px^2] + \ldots + \beta_k\, [px^k] &= [pa] \\
\beta_0\, [px] + \beta_1\, [px^2] + \beta_2\, [px^3] + \ldots + \beta_k\, [px^{k+1}] &= [pax] \\
&\vdots \\
\beta_0\, [px^k] + \beta_1\, [px^{k+1}] + \beta_2\, [px^{k+2}] + \ldots + \beta_k\, [px^{2k}] &= [pax^k].
\end{aligned} \tag{4.50}$$

4.5 Darstellung einer Funktion mit Hilfe “einfacherer” Funktionen

Betrachten wir den Zusammenhang

$$y(x) = \beta_0\, \phi_0(x) + \beta_1\, \phi_1(x) + \ldots + \beta_k\, \phi_k(x), \tag{4.51}$$

wobei $\phi_0(x)$, $\phi_1(x)$, ..., $\phi_k(x)$ bekannte Funktionen von x sind. Es kann sein, daß $f(x)$ an diskreten Stellen x_i bekannt ist ($i = 1, 2, \ldots, n$; $n > k$). Dann definieren wir

$$S = \sum_{i=1}^{n} p_i \left(y(x_i) - \beta_0\,\phi_0(x_i) - \beta_1\,\phi_1(x_i) - \ldots - \beta_k\,\phi_k(x_i)\right)^2. \quad (4.52)$$

Ist dagegen $f(x)$ in einem Intervall $a \leq x \leq b$ bekannt, dann definieren wir

$$S = \int_a^b \left(y(x) - \beta_0\,\phi_0(x) - \beta_1\,\phi_1(x) - \ldots - \beta_k\,\phi_k(x)\right)^2 dx. \quad (4.53)$$

Definieren wir jetzt die Hilfsgrößen

$$[\phi_j \phi_k] = \begin{cases} \sum_{i=1}^{n} \phi_j(x_i)\phi_k(x_i) & \text{(diskret)} \\ \int_a^b \phi_j(x)\phi_k(x)dx & \text{(kontinuierlich)} \end{cases} \quad (4.54)$$

und

$$[y\phi_k] = \begin{cases} \sum_{i=1}^{n} y(x_i)\phi_k(x_i) & \text{(diskret)} \\ \int_a^b y(x)\phi_k(x)dx & \text{(kontinuierlich)} \end{cases}, \quad (4.55)$$

finden wir in beiden Fällen, daß die optimalen Werte der Parameter β_0, β_1, ..., β_k durch Lösung der Normalgleichungen

$$\begin{array}{rcl} \beta_0\,[\phi_0^2] + \beta_1\,[\phi_0\phi_1] + \beta_2\,[\phi_0\phi_2] + \ldots + \beta_k\,[\phi_0\phi_k] & = & [y\phi_0] \\ \beta_0\,[\phi_1\phi_0] + \beta_1\,[\phi_1^2] + \beta_2\,[\phi_1\phi_2] + \ldots + \beta_k\,[\phi_1\phi_k] & = & [y\phi_1] \\ & \vdots & \\ \beta_0\,[\phi_k\phi_0] + \beta_1\,[\phi_k\phi_1] + \beta_2\,[\phi_k\phi_2] + \ldots + \beta_k\,[\phi_k^2] & = & [y\phi_k]. \end{array} \quad (4.56)$$

ermittelt werden können.

Index