

# Mathematik für Wirtschaftswissenschaftler

(Master Economics)

PROF. DR. ALFRED MÜLLER

Universität Siegen

Wintersemester 2008/09

# Vorwort

Dieses Manuskript enthält den wesentlichen Inhalt der Vorlesung *Mathematik für Wirtschaftswissenschaftler*, die im Wintersemester 2008/09 an der Universität Siegen für Studierende im Masterstudiengang Economics gehalten wurde.

# Inhaltsverzeichnis

<b>1</b>	<b>Integralrechnung</b>	<b>4</b>
1.1	Das unbestimmte Integral und Stammfunktionen . . . . .	5
1.2	Das bestimmte Integral und Flächeninhalte . . . . .	8
1.3	Spezielle Integrationstechniken . . . . .	13
1.4	Uneigentliche Integrale . . . . .	17
1.5	Integration von Funktionen mehrerer Veränderlicher . . . . .	19
<b>2</b>	<b>Lineare Algebra</b>	<b>21</b>
2.1	Definition von Eigenwerten . . . . .	22
2.2	Determinanten . . . . .	24
2.3	Lineare Unabhängigkeit, Basis und Dimension . . . . .	30
2.4	Rang einer Matrix . . . . .	33
2.5	Lösungsmengen linearer Gleichungssysteme . . . . .	35
2.6	Berechnung von Eigenwerten und Eigenvektoren . . . . .	38
<b>3</b>	<b>Differential- und Differenzgleichungen</b>	<b>42</b>
3.1	Differentialgleichungen . . . . .	42
3.2	Differenzgleichungen . . . . .	48

# 1 Integralrechnung

Unter *Integration* versteht man die Umkehrung der Differentiation. Man sucht also zu einer gegebenen Ableitung  $f'$  die ursprüngliche Funktion  $f$ . Ein ökonomisches Beispiel soll die Problemstellung verdeutlichen.

**Beispiel 1.1.** Ein Unternehmen, das nur ein Produkt herstellt, sehe sich folgender Grenzkostenfunktion  $K'$  gegenüber:

$$K'(x) = 0.3x^2 - 4x + 21. \quad (1.1)$$

Wie kann man daraus die Gesamtkostenfunktion  $K$  ermitteln? Gesucht ist also eine Funktion  $x \mapsto K(x)$  derart, dass die Ableitung  $K'$  genau mit der Grenzkostenfunktion (1.1) übereinstimmt. Mit Kenntnissen der Differentialrechnung kann man das Ergebnis erraten.

1. Die Ableitung von  $x^3$  ist  $3x^2$ , daher ist  $0.3x^2$  die Ableitung von  $0.1x^3$ . Analog:
2.  $-4x$  ist die Ableitung von  $-2x^2$ .
3.  $21$  ist die Ableitung von  $21x$ .

Also hat die Funktion

$$K(x) = 0.1x^3 - 2x^2 + 21x \quad (1.2)$$

die gewünschte Ableitung  $K'$ .

Man kann offensichtlich zur Kostenfunktion (1.2) noch einen beliebigen festen Fixkostenwert  $K_f$  addieren, ohne dass sich die Grenzkosten (1.1) ändern,

$$K(x) = 0.1x^3 - 2x^2 + 21x + K_f, \quad (1.3)$$

da beim Ableiten die Konstante  $K_f$  verschwindet. Die Kostenfunktion (1.3) ist also erst durch Vorgabe der Fixkosten eindeutig bestimmt.

Wir wollen diese Problemstellung nun genauer untersuchen.

## 1.1 Das unbestimmte Integral und Stammfunktionen

Eine der zwei Hauptaufgaben der Integralrechnung besteht darin, zu einer gegebenen Ableitung die ursprüngliche Funktion (auch Stammfunktion) zu finden. Es wird also Integration als Umkehrung der Differentiation betrachtet.

**Definition 1.2.** Sei  $f$  eine gegebene stetige Funktion im Intervall  $[a, b]$ . Eine differenzierbare Funktion  $F : [a, b] \rightarrow \mathbb{R}$  heißt **Stammfunktion** zu  $f$ , falls gilt:

$$F'(x) = f(x) \quad \text{für alle } x \in (a, b).$$

Wie wir an obigem Beispiel schon gesehen haben, sind Stammfunktionen nicht eindeutig. Wenn man zu einer Stammfunktion eine Konstante addiert, erhält man eine weitere Stammfunktion. Genauer gilt folgender Sachverhalt.

**Satz 1.3.** Sei  $f$  stetig auf  $[a, b]$ , und  $F_1$  eine Stammfunktion zu  $f$ . Dann erhält man sämtliche Stammfunktionen  $F$  zu  $f$  durch

$$F(x) = F_1(x) + c, \quad c \in \mathbb{R}.$$

Der Satz enthält also zwei Aussagen.

1. Wenn  $F(x)$  eine Stammfunktion zu  $f$  ist, so auch  $F(x) + c$  für jede beliebige Konstante  $c$ .
2. Wenn  $F_1$  und  $F_2$  Stammfunktionen zu  $f$  sind, so gilt stets  $F_2(x) = F_1(x) + c$  mit einer Konstanten  $c$ .

**Beispiel 1.4.** Sei  $f(x) = x^2$ . Dann rechnet man leicht nach, dass durch  $F_1(x) = \frac{1}{3}x^3$  eine Stammfunktion gegeben ist. Sämtliche Stammfunktionen lassen sich dann darstellen in der Form  $F(x) = \frac{1}{3}x^3 + c$ , also z.B.  $\frac{1}{3}x^3 + 5$  oder  $\frac{1}{3}x^3 + \ln(2)$  u.s.w.

Wir führen nun ein Symbol ein für die Menge aller (sich nur um Konstanten unterscheidenden) Stammfunktionen.

**Definition 1.5.** Die Menge aller Stammfunktionen zu  $f$  in  $[a, b]$  wird **unbestimmtes Integral** genannt und mit  $\int f(x)dx$  bezeichnet, d.h.

$$\int f(x)dx = \{F \mid F'(x) = f(x)\}$$

Die Schreibweise scheint im Moment unnötig kompliziert zu sein, weil man ja vielleicht auch kürzer  $\int f$  oder  $\int f(x)$  schreiben könnte. Der Grund für die scheinbar kompliziertere Schreibweise wird später klar werden, wenn wir die Hauptsätze der Differenzial- und Integralrechnung sowie Substitutionsregel u.ä. kennenlernen werden.

Wir werden häufig die (mathematisch nicht ganz korrekte) Schreibweise

$$\int f(x)dx = F(x) + c$$

verwenden. Dies ist mathematisch nicht ganz korrekt, weil links vom Gleichheitszeichen eine Menge steht und rechts eigentlich ein Repräsentant der Menge, aber dies führt eigentlich nie zu Missverständnissen und ist deshalb eine bequeme Schreibweise.

Wir erhalten folgende Tabelle für einige der wichtigsten Integrale:

$f(x)$	$\int f(x)dx$	Bemerkungen
0	$c$	$c = \text{const.}$
$x^n$	$\frac{1}{n+1}x^{n+1} + c$	$n \neq -1$
$(ax + b)^n$	$\frac{1}{a} \frac{(ax+b)^{n+1}}{n+1} + c$	$n \neq -1, a \neq 0$
$1/x$	$\ln(x) + c$	$x > 0$
$\frac{1}{ax+b}$	$\frac{1}{a} \ln(ax + b) + c$	$ax + b > 0, a \neq 0$
$e^x$	$e^x + c$	
$e^{ax+b}$	$\frac{1}{a}e^{ax+b} + c$	$a \neq 0$

**Beispiel 1.6.** (i)

$$\int 1dx = x + c;$$

(ii)

$$\int x^7 dx = \frac{1}{8}x^8 + c;$$

(iii)

$$\int \sqrt{x} dx = \int x^{1/2} dx = \frac{2}{3}x^{3/2} + c;$$

(iv)

$$\int (3z - 2)^2 dz = \frac{1}{3} \frac{(3z - 2)^3}{3} + c;$$

(v)

$$\int e^{-0.1t} dt = -10e^{-0.1t} + c.$$

**Bemerkung:** Obwohl man zeigen kann, dass jede stetige Funktion auch eine Stammfunktion besitzt, ist es nicht immer möglich, diese Stammfunktion in geschlossener Form (d.h. durch Kombination endlich vieler elementarer Funktionen) darzustellen. Dies gilt z.B. für folgende Integrale:

$$\int e^{-x^2} dx; \quad \int \frac{e^x}{x} dx; \quad \int \frac{dx}{\ln(x)}.$$

Einige elementare Rechenregeln für unbestimmte Integrale enthält der folgende Satz.

**Satz 1.7.** Es seien  $f, g$  stetige Funktionen. Dann gilt:

a)

$$\int af(x)dx = a \int f(x)dx,$$

b)

$$\int (f(x) + g(x))dx = \int f(x)dx + \int g(x)dx.$$

Der Beweis folgt sofort jeweils durch Ableiten der beiden Seiten unter Beachtung von

$$\frac{d}{dx} \int f(x)dx = f(x)$$

Die Eigenschaften aus Satz 1.7 bezeichnet man auch als *Linearität des Integrals*. Sie lässt sich auch in einer einzigen Gleichung zusammenfassen als

$$\int (af(x) + bg(x))dx = a \int f(x)dx + b \int g(x)dx.$$

**Beispiel 1.8.**

$$\int \left( 8x^3 + \frac{12}{\sqrt{4x+9}} \right) dx = 2x^4 + 6\sqrt{4x+9} + c.$$

Leider gibt es keine vergleichbare einfache Regel für die Integration von Produkten. Aus der Produktregel der Differenziation

$$(f(x) \cdot g(x))' = f'(x) \cdot g(x) + f(x) \cdot g'(x) \neq f'(x)g'(x)$$

kann man leider keine einfache Regel für die Integration von Produkten herleiten. Es gilt insbesondere

$$\int (f'(x)g'(x))dx \neq f(x) \cdot g(x).$$

## 1.2 Das bestimmte Integral und Flächeninhalte

Es sei  $f : [a, b] \rightarrow \mathbb{R}$  eine positive stetige Funktion. Dann besteht die zweite Hauptaufgabe der Integralrechnung darin, den Inhalt  $A$  der Fläche zu bestimmen, die vom Graphen der Funktion, der  $x$ -Achse sowie den Senkrechten  $x = a$  und  $x = b$  begrenzt wird.

Dazu muss zunächst formal definiert werden, was unter dem anschaulich klaren Begriff des Flächeninhalts im allgemeinen Fall verstanden werden soll. Dies geschieht mit Hilfe des Grenzwertbegriffes, indem die Fläche von unten und oben angenähert wird durch stückweise konstante Funktionen, für deren Flächenberechnung man nur wissen muss, dass der Inhalt eines Rechtecks mit Kantenlängen  $a$  und  $b$  gleich dem Produkt  $a \cdot b$  ist.

Wir zerlegen also das Intervall  $[a, b]$  in Teilintervalle  $[x_i, x_{i+1}]$  mit

$$a = x_0 < x_1 < \dots < x_n = b.$$

Auf dem Intervall  $[x_i, x_{i+1}]$  seien

$$u_i := \min\{f(x) : x_i \leq x \leq x_{i+1}\} \quad \text{und} \quad o_i := \max\{f(x) : x_i \leq x \leq x_{i+1}\}$$

der kleinste bzw. grösste Funktionswert von  $f$ . Das bedeutet, dass ein Rechteck mit Höhe  $u_i$  unter den Funktionsgraphen passt, und ein Rechteck der Höhe  $o_i$  den Funktionsgraphen einschliesst. Für eine derart gegebene Zerlegung  $Z$  definieren wir die zugehörige **Untersumme** und **Obersumme** durch

$$U(Z, f) = \sum_{i=0}^{n-1} u_i \cdot (x_{i+1} - x_i) \quad \text{und} \quad O(Z, f) = \sum_{i=0}^{n-1} o_i \cdot (x_{i+1} - x_i).$$

Offensichtlich ist  $U(Z, f) \leq O(Z, f)$  und bei einer vernünftiger Definition sollte der Flächeninhalt  $A$  zwischen diesen beiden Werten liegen. Je feiner die Zerlegung  $Z$ , desto näher sollten die oberen und unteren Schranken am gesuchten Flächeninhalt liegen.

Formal wird deshalb das Feinheitsmaß

$$|Z| = \max_{i=0, \dots, n-1} |x_{i+1} - x_i|$$

definiert als die Breite des längsten Intervalls der Zerlegung, und die Fläche  $A$  wird definiert durch Grenzübergang  $|Z| \rightarrow 0$ , d.h.

$$A := \lim_{|Z| \rightarrow 0} U(Z, f) = \lim_{|Z| \rightarrow 0} O(Z, f), \tag{1.4}$$

falls diese beiden Grenzwerte übereinstimmen. Man kann zeigen, dass diese beiden Grenzwerte immer existieren und übereinstimmen, falls  $f$  stetig ist. Man nennt diesen Grenzwert das **bestimmte Integral**.

**Definition 1.9.** Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine stetige Funktion. Dann nennt man den Grenzwert in (1.4) **bestimmtes Integral** von  $f$  über  $[a, b]$  und benutzt dafür folgende Schreibweise:

$$\int_a^b f(x) dx = \lim_{|Z| \rightarrow 0} U(Z, f) = \lim_{|Z| \rightarrow 0} O(Z, f).$$

Dabei heißt  $a$  **untere Integrationsgrenze** und  $b$  **obere Integrationsgrenze**. Die Funktion  $f$  heißt **Integrand** und  $x$  heißt **Integrationsvariable**.



Die Integrationsvariable kann offensichtlich beliebig umbenannt werden. Es gilt

$$\int_a^b f(x)dx = \int_a^b f(z)dz = \int_a^b f(t)dt.$$

Eine beliebige (evtl. nicht stetige) Funktion  $f$  heißt (Riemann-) **integrierbar**, falls die beiden Grenzwerte in (1.4) existieren und übereinstimmen. Es ist leicht einzusehen, dass auch stückweise stetige Funktionen mit endlich vielen Unstetigkeitsstellen integrierbar sind. Nur sehr ausgefallene Funktionen sind nicht Riemann-integrierbar wie z.B. die Funktion

$$f(x) = \begin{cases} 1, & \text{falls } x \in \mathbb{Q}, \\ 0, & \text{sonst.} \end{cases}$$

**Beispiel 1.10.** Wir wollen mit der Definition des Integrals die Trapezfläche  $\int_a^b x dx$  berechnen. Aus der Elementargeometrie sollte bekannt sein, dass für die zugehörige Fläche gilt

$$\int_a^b x dx = \frac{a+b}{2}(b-a) = \frac{1}{2}(b^2 - a^2).$$

Für den formalen Nachweis dieser Formel mit obiger Definition zerlegen wir das Intervall  $[a, b]$  in  $n$  äquidistante Teilintervalle der Länge  $(b-a)/n$ . Also ist  $x_i = a + i(b-a)/n$ . Wegen der Monotonie der Funktion  $f(x) = x$  ist deshalb

$$u_i = a + i \frac{b-a}{n} \quad \text{und} \quad o_i = a + (i+1) \frac{b-a}{n}.$$

Daraus ergibt sich

$$\begin{aligned} U(Z, f) &= \sum_{i=0}^{n-1} u_i \cdot (x_{i+1} - x_i) \\ &= \sum_{i=0}^{n-1} \left( a + i \frac{b-a}{n} \right) \cdot \frac{b-a}{n} \\ &= \frac{b-a}{n} \cdot \left( an + \frac{b-a}{n} \sum_{i=0}^{n-1} i \right). \end{aligned}$$

Wegen

$$\sum_{i=0}^{n-1} i = \frac{(n-1)n}{2}$$

ergibt sich damit

$$\begin{aligned} U(Z, f) &= \frac{b-a}{n} \cdot \left( an + \frac{b-a}{n} \frac{(n-1)n}{2} \right) \\ &= a(b-a) + \frac{n-1}{n} \frac{(b-a)^2}{2}. \end{aligned}$$

Durch Grenzübergang  $|Z| \rightarrow 0$  bzw.  $n \rightarrow \infty$  erhält man

$$\lim_{|Z| \rightarrow 0} U(Z, f) = a(b-a) + \frac{(b-a)^2}{2} = \frac{b^2 - a^2}{2}.$$

Analog erhält man für die Obersumme  $O(Z, f)$  wegen  $o_i = u_i + (b-a)/n$

$$O(Z, f) = U(Z, f) + \frac{b-a}{n} \cdot \frac{b-a}{n} \cdot n = U(Z, f) + \frac{(b-a)^2}{n}$$

den gleichen Grenzwert und somit

$$\int_a^b x dx = \lim_{|Z| \rightarrow 0} U(Z, f) = \lim_{|Z| \rightarrow 0} O(Z, f) = \frac{b^2 - a^2}{2}.$$

Folgende elementaren Eigenschaften des bestimmten Integrals lassen sich einfach aus der Definition herleiten und sind anschaulich klar.

**Satz 1.11.** *Es seien  $f, g$  stetige Funktionen. Dann gilt:*

a)

$$\int_a^b c f(x) dx = c \int_a^b f(x) dx,$$

b)

$$\int_a^b (f(x) + g(x)) dx = \int_a^b f(x) dx + \int_a^b g(x) dx.$$

Ausserdem gilt offensichtlich folgende **Intervalladditivität**.

**Satz 1.12.** *Es sei  $f$  in  $[a, b]$  und in  $[b, c]$  integrierbar. Dann gilt:*

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx.$$

Aus Satz 1.12 lässt sich sofort folgern, dass eine stückweise stetige Funktion mit endlich vielen Unstetigkeitsstellen

$$a < a_1 < a_2 < \dots < a_n < b$$

integrierbar ist, und dass man das Integral ausrechnen kann, indem man die Integrale

$$\int_{a_i}^{a_{i+1}} f(x) dx, \quad i = 0, \dots, n,$$

(mit  $a_0 := a$  und  $a_{n+1} := b$ ) ausrechnet als Integrale stetiger Funktionen auf den jeweiligen Intervallen, und anschliessend aufaddiert.

Wir untersuchen nun den Zusammenhang zwischen unbestimmtem und bestimmtem Integral. Dazu definieren wir die Integralfunktion (Flächeninhaltsfunktion)

$$F(x) = \int_a^x f(t) dt$$

als den Flächeninhalt unter dem Graphen zwischen den Punkten  $a$  und  $x$ .

Wir erhalten folgenden **1. Hauptsatz der Differenzial- und Integralrechnung**.

**Satz 1.13.** Sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig. Dann ist die Integralfunktion  $F(x)$  von  $f$  auf  $[a, b]$  differenzierbar und es gilt

$$F'(x) = f(x) \quad (1.5)$$

*Beweis.* Aus Satz 1.12 folgt sofort für ein beliebiges  $h > 0$  und  $x \in [a, b]$

$$\begin{aligned} F(x+h) - F(x) &= \int_a^{x+h} f(t)dt - \int_a^x f(t)dt \\ &= \int_x^{x+h} f(t)dt. \end{aligned}$$

Sei

$$m_h := \min\{f(t) : x \leq t \leq x+h\} \quad \text{und} \quad M_h := \max\{f(t) : x \leq t \leq x+h\}.$$

Dann gilt

$$hm_h \leq F(x+h) - F(x) \leq hM_h,$$

also

$$m_h \leq \frac{F(x+h) - F(x)}{h} \leq M_h.$$

Aus der Stetigkeit von  $f$  folgt  $\lim_{h \rightarrow 0} m_h = \lim_{h \rightarrow 0} M_h = f(x)$  und somit

$$f(x) \leq \lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} \leq f(x),$$

folglich  $F'(x) = f(x)$ . □

Wir wollen nun noch für spätere Zwecke bestimmte Integrale auch für den Fall  $a = b$  und  $a > b$  definieren, auch wenn diese keine Bedeutung als Flächeninhalte haben. Es ist naheliegend, das bestimmte Integral im Falle  $a = b$  auf Null zu setzen. Damit dann weiterhin Satz 1.12 gilt, muss man dann folgende Definition wählen.

**Definition 1.14.** Sei  $f : [a, b] \rightarrow \mathbb{R}$  integrierbar. Dann setzt man

$$a) \quad \int_a^a f(x)dx = 0, \quad b) \quad \int_a^b f(x)dx = - \int_b^a f(x)dx.$$

Damit gilt also wie gewünscht die Beziehung

$$\int_a^b f(x)dx + \int_b^a f(x)dx = \int_a^a f(x)dx = 0.$$

Der Satz 1.13 besagt gerade, dass die Integralfunktion  $F(x)$  eine Stammfunktion zu  $f(x)$  ist. Wir wissen aus Satz 1.3 dass sich jede andere Stammfunktion von  $F$  nur durch eine Konstante unterscheidet. Angenommen, wir können eine Stammfunktion  $F_1$  von  $f$  erraten. Dann wissen wir also, dass für die Integralfunktion gilt

$$F(x) = \int_a^x f(t) dt = F_1(x) + c.$$

Mit Hilfe von Definition 1.14 können wir nun aber  $c$  bestimmen! Es gilt nämlich

$$\int_a^a f(x) dx = F_1(a) + c = 0, \quad \text{also } c = -F_1(a).$$

Also gilt für jede beliebige Stammfunktion  $F_1$

$$\int_a^x f(t) dt = F_1(x) - F_1(a). \quad (1.6)$$

Diese Aussage ist als **2. Hauptsatz der Differenzial- und Integralrechnung** bekannt.

**Satz 1.15.** *Es sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig, und  $F$  eine beliebige Stammfunktion zu  $f$ . Dann gilt*

$$\int_a^b f(x) dx = F(b) - F(a).$$

Damit können wir uns die komplizierte Berechnung von bestimmten Integralen über die Berechnung des Grenzwertes von Ober- und Untersummen ersparen, und durch die Bestimmung von Stammfunktionen ersetzen. Ausserdem rechtfertigt dieser Satz im Nachhinein die Verwendung des gleichen Symbolik ( $\int$  bzw.  $\int_a^b$ ).

**Konvention:** Man benützt oft folgende Schreibweisen bei der Berechnung von bestimmten Integralen mittels Stammfunktionen:

$$\int_a^b f(x) dx = F(x) \Big|_a^b = \left[ F(x) \right]_a^b = F(b) - F(a).$$

**Beispiel 1.16.** a) Wir wiederholen die Berechnung von  $\int_a^b x dx$  mit Hilfe des 2. Hauptsatzes der Differenzial- und Integralrechnung. Eine Stammfunktion von  $f(x) = x$  ist gegeben durch  $F(x) = x^2/2 + c$  mit einer beliebigen Konstante  $c$ . Aus Satz 1.15 folgt damit

$$\int_a^b x dx = F(b) - F(a) = \frac{b^2}{2} + c - \left( \frac{a^2}{2} + c \right) = \frac{b^2 - a^2}{2}.$$

Aus der komplizierten Rechnung mit Hilfe der Grenzwerte von Ober- und Untersummen wird also ein einfacher Einzeiler, und wie man am Beispiel noch mal sieht, spielt die Wahl der Konstante  $c$  keine Rolle.

b) Wir berechnen den Flächeninhalt zwischen der Funktion  $f(x) = \sqrt{x}$ , der  $x$ -Achse, und den Geraden  $x = 1$  und  $x = 4$ :

$$\int_1^4 \sqrt{x} dx = \int_1^4 x^{1/2} dx = \frac{2}{3} x^{3/2} \Big|_1^4 = \frac{2}{3} 4^{3/2} - \frac{2}{3} 1^{3/2} = \frac{14}{3}.$$

Zur Berechnung von Flächeninhalten negativer Funktionen  $f < 0$  benutzt man die Tatsache, dass die Fläche zwischen der  $x$ -Achse und der Funktion  $f$  die gleiche ist wie die Fläche zwischen der positiven Funktion  $-f$  und der  $x$ -Achse. Deshalb ist der Flächeninhalt  $A$  zwischen der  $x$ -Achse, der negativen Funktion  $f$  und den Geraden  $x = a$  und  $x = b$  gegeben durch

$$A = \int_a^b -f(x) dx = - \int_a^b f(x) dx = \left| \int_a^b f(x) dx \right|.$$

Will man den Flächeninhalt der Fläche zwischen der  $x$ -Achse und einer Funktion  $f$  mit Vorzeichenwechseln auf  $[a, b]$  berechnen, so muss man beachten, dass dies nicht mehr als einfaches bestimmtes Integral berechnet werden kann. So gilt zum Beispiel

$$\int_0^3 (x^2 - 3) dx = \left. \frac{1}{3}x^3 - 3x \right|_0^3 = \frac{1}{3}27 - 9 = 0,$$

obwohl die Fläche offensichtlich von Null verschieden ist.

Bei der Bestimmung des bestimmten Integrals werden die Flächen oberhalb der  $x$ -Achse positiv und unterhalb der  $x$ -Achse negativ gezählt, und in diesem Beispiel sind die zufällig gerade gleich gross, weshalb das bestimmte Integral verschwindet.

Zur Bestimmung der Fläche muss man also erst die Nullstellen der Funktion im Intervall  $[a, b]$  bestimmen. Seien diese gegeben durch

$$a < x_1 < \dots < x_n < b.$$

Dann ist der gesuchte Flächeninhalt  $A$  gegeben durch

$$A = \left| \int_a^{x_1} f(x) dx \right| + \left| \int_{x_1}^{x_2} f(x) dx \right| + \dots + \left| \int_{x_n}^b f(x) dx \right|.$$

Im Falle der Funktion  $f(x) = x^2 - 3$  erhalten wir im Intervall  $[0, 3]$  die Nullstelle  $x_1 = \sqrt{3}$  und somit

$$\begin{aligned} A &= \left| \int_0^{\sqrt{3}} (x^2 - 3) dx \right| + \left| \int_{\sqrt{3}}^3 (x^2 - 3) dx \right| \\ &= \left| \left[ \frac{1}{3}x^3 - 3x \right]_0^{\sqrt{3}} \right| + \left| \left[ \frac{1}{3}x^3 - 3x \right]_{\sqrt{3}}^3 \right| \\ &= \left| \sqrt{3} - 3\sqrt{3} \right| + \left| 9 - 9 - (\sqrt{3} - 3\sqrt{3}) \right| \\ &= 4\sqrt{3}. \end{aligned}$$

Den Flächeninhalt zwischen zwei sich nicht schneidenden Funktionen  $f$  und  $g$  kann man als Differenz der beiden unter den Graphen liegenden Flächenstücke auffassen. Im Falle  $f(x) \geq g(x)$  auf  $[a, b]$  erhalten wir deshalb

$$\begin{aligned} A &= \int_a^b f(x) dx - \int_a^b g(x) dx \\ &= \int_a^b (f(x) - g(x)) dx. \end{aligned}$$

### 1.3 Spezielle Integrationstechniken

Für Integrale gibt es leider keine einfache Produktregel und auch keine einfache Kettenregel. So lassen sich zwar leicht Stammfunktionen für die Funktionen  $f(x) = e^x$  und  $g(x) = -x^2$

finden, für die verkettete Funktion  $f(g(x)) = e^{-x^2}$  ist es dagegen unmöglich, eine geschlossene Formel für das Integral anzugeben. Trotzdem lassen sich aus den entsprechenden Regeln für die Differenziation Tricks für die Integration von Produkten und verketteten Funktionen ableiten.

Wir betrachten zunächst den Fall von Produktintegration. Hier gilt folgende Regel, die man auch **partielle Integration** nennt.

**Satz 1.17.** *Es seien  $f, f', g, g'$  stetige Funktionen. Dann gilt*

$$\int f(x)g'(x) dx = f(x)g(x) - \int f'(x)g(x) dx. \quad (1.7)$$

*Beweis.* Wir zeigen, dass beide Seiten der Gleichung (1.7) die gleiche Ableitung besitzen. Für die Ableitung der linken Seite gilt nach der Definition des unbestimmten Integrals

$$\frac{d}{dx} \int f(x)g'(x) dx = f(x)g'(x)$$

und für die rechte Seite gilt nach der Produktregel der Differenziation

$$\frac{d}{dx} \left( f(x)g(x) - \int f'(x)g(x) dx \right) = (f'(x)g(x) + f(x)g'(x)) - f'(x)g(x) = f(x)g'(x).$$

Also stimmen die beiden Seiten der Gleichung überein. □

Dieses Verfahren der partiellen Integration empfiehlt sich, wenn folgende zwei Voraussetzungen erfüllt sind:

1. Der Integrand kann als Produkt aufgefasst werden, dessen einer Faktor (nämlich  $g'(x)$ ) leicht integriert werden kann (zu  $g(x)$ ).
2. Das auf der rechten Seite von (1.7) stehende Integral  $\int f'(x)g(x) dx$  ist einfacher zu lösen als das ursprüngliche Integral  $\int f(x)g'(x) dx$ .

**Beispiel 1.18.** Gesucht ist  $\int x \ln(x) dx$  (für  $x > 0$ ). Der zweite Faktor  $\ln(x)$  ist nicht ohne weiteres integrierbar, wohl aber der erste ( $= x$ ). Also setzen wir  $g'(x) = x$  und  $f(x) = \ln(x)$ . Dann ist  $g(x) = x^2/2$  eine Stammfunktion von  $g'(x)$  und  $f'(x) = 1/x$ . Somit erhalten wir

$$\begin{aligned} \int x \ln(x) dx &= \frac{x^2}{2} \ln(x) - \int \frac{x^2}{2} \frac{1}{x} dx \\ &= \frac{x^2}{2} \ln(x) - \frac{x^2}{4} + c \end{aligned}$$

Für das zugehörige bestimmte Integral gilt natürlich die analoge Regel

$$\boxed{\int_a^b f(x)g'(x) dx = \left[ f(x)g(x) \right]_a^b - \int_a^b f'(x)g(x) dx} \quad (1.8)$$

Also erhält man z.B.

$$\begin{aligned}\int_2^3 x \ln(x) dx &= \left[ \frac{x^2}{2} \ln(x) \right]_2^3 - \int_2^3 \frac{x^2}{2} \frac{1}{x} dx \\ &= \left[ \frac{x^2}{2} \ln(x) - \frac{x^2}{4} \right]_2^3 \approx 2.3075.\end{aligned}$$

Aus der Kettenregel der Differenziation kann man folgende **Substitutionsregel** herleiten, indem man im Integral  $\int f(x)dx$  die Variable  $x$  durch eine Funktion  $x = h(t)$  ersetzt.

**Satz 1.19.** *Es sei  $f$  stetig, und  $h$  stetig differenzierbar und **umkehrbar** mit Umkehrfunktion  $h^{-1}$ . Dann gilt*

$$\int f(x) dx = \int f(h(t)) \cdot h'(t) dt. \quad (1.9)$$

*Beweis.* Wir zeigen, dass die Ableitungen der beiden Seiten von Gleichung (1.9) nach  $t$  übereinstimmen. Wegen  $x = h(t)$  ist  $dx/dt = h'(t)$  und somit

$$\begin{aligned}\frac{d}{dt} \int f(x) dx &= \frac{d}{dx} \left( \int f(x) dx \right) \frac{dx}{dt} = f(x)h'(t) = f(h(t))h'(t) \\ &= \frac{d}{dt} \int f(h(t)) \cdot h'(t) dt.\end{aligned}$$

□

**Bemerkungen:** 1. Häufig liest man Satz 1.19 in folgender Weise:

$$\int f(h(x)) \cdot h'(x) dx = \int f(t) dt$$

mit  $h(x) = t$  bzw.  $x = h^{-1}(t)$ .

2. Bei diesem Satz wird spätestens deutlich, warum es sinnvoll ist, das Differential  $dx$  immer mit anzugeben bei der Schreibweise  $\int f(x)dx$ . Der Satz besagt nämlich, dass wenn man im Integranden  $f(x)$  die Variable  $x$  durch  $x = h(t)$  substituiert, dass man dann auch das Differential  $dx$  substituieren muss durch  $dx = h'(t)dt$ , weil eben für die Ableitung von  $x = h(t)$  gilt:  $dx/dt = h'(t)$ .

**Beispiel 1.20.** a) Gesucht ist  $\int x\sqrt{1-x^2}dx$ . Wir vereinfachen das Integral durch Substitution  $1-x^2 = t$ . Dann gilt  $dt = -2xdx$ . Deshalb erhält man

$$\begin{aligned}\int x\sqrt{1-x^2}dx &= -\frac{1}{2} \int \sqrt{1-x^2} \cdot (-2x)dx = -\frac{1}{2} \int \sqrt{t}dt \\ &= -\frac{1}{3}t^{3/2} + c = -\frac{1}{3}(1-x^2)^{3/2} + c = -\frac{1}{3}\sqrt{(1-x^2)^3} + c.\end{aligned}$$

b) Gesucht ist  $\int \frac{x^3+x}{x^4+2x^2} dx$ . Da der Zähler fast übereinstimmt mit der Ableitung vom Nenner, substituieren wir  $t = x^4 + 2x^2$ . Dann gilt  $dt = (4x^3 + 4x)dx$  und deshalb

$$\begin{aligned} \int \frac{x^3+x}{x^4+2x^2} dx &= \frac{1}{4} \int \frac{(4x^3+4x)dx}{x^4+2x^2} = \frac{1}{4} \int \frac{1}{t} dt \\ &= \frac{1}{4} \ln(t) + c = \frac{1}{4} \ln(x^4+2x^2) + c. \end{aligned}$$

Für das **bestimmte Integral** erhalten wir die folgende zu Satz 1.19 analoge Regel:

**Satz 1.21.** *Es sei  $f$  stetig, und  $h$  stetig differenzierbar und **umkehrbar** mit Umkehrfunktion  $h^{-1}$ . Dann gilt*

$$\int_a^b f(x) dx = \int_u^v f(h(t)) \cdot h'(t) dt \quad (1.10)$$

mit  $x = h(t)$ ,  $a = h(u)$ ,  $b = h(v)$ , also  $u = h^{-1}(a)$ ,  $v = h^{-1}(b)$ .

Die Gleichung (1.10) lässt sich wieder umschreiben in

$$\int_a^b f(h(x)) \cdot h'(x) dx = \int_{h(a)}^{h(b)} f(t) dt.$$

Hier ist also zu beachten, dass auch die Integrationsgrenzen entsprechend der Substitutionsfunktion zu transformieren sind.

**Beispiel 1.22.** Gesucht ist  $\int_1^2 x^3 \sqrt{x^4-1} dx$ . Wir substituieren

$$t = h(x) = x^4 - 1 \quad \Rightarrow \quad dt = 4x^3 dx.$$

Die Integrationsgrenzen  $x_1 = 1$  und  $x_2 = 2$  müssen entsprechend transformiert werden:

$$t_1 = h(x_1) = 1^4 - 1 = 0, \quad t_2 = h(x_2) = 2^4 - 1 = 15.$$

Also ist

$$\int_1^2 x^3 \sqrt{x^4-1} dx = \frac{1}{4} \int_0^{15} \sqrt{t} dt = \left[ \frac{1}{6} t^{3/2} \right]_0^{15} \approx 9.6825.$$

Alternativ könnte man die Stammfunktion durch "Resubstitution" gewinnen, und dann die ursprünglichen Integrationsgrenzen verwenden:

$$\int x^3 \sqrt{x^4-1} dx = \frac{1}{4} \int \sqrt{t} dt = \frac{1}{6} (x^4-1)^{3/2} + c.$$

Daraus folgt für das bestimmte Integral

$$\int_1^2 x^3 \sqrt{x^4-1} dx = \left[ \frac{1}{6} (x^4-1)^{3/2} \right]_1^2 \approx 9.6825.$$



## 1.4 Uneigentliche Integrale

Manchmal interessiert man sich für Integrale über unendliche Intervalle.

**Definition 1.23.** Falls für eine integrierbare Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  der Grenzwert  $\lim_{b \rightarrow \infty} \int_a^b f(x) dx$  existiert, so schreibt man

$$\int_a^\infty f(x) dx = \lim_{b \rightarrow \infty} \int_a^b f(x) dx.$$

Man sagt dann, dass das **uneigentliche Integral**  $\int_a^\infty f(x) dx$  konvergiert. Falls der Grenzwert nicht existiert, sagt man, dass das Integral divergiert. Analog definiert man im Falle der Existenz der Grenzwerte

$$\int_{-\infty}^b f(x) dx = \lim_{a \rightarrow -\infty} \int_a^b f(x) dx$$

und

$$\int_{-\infty}^\infty f(x) dx = \lim_{a \rightarrow -\infty} \int_a^\infty f(x) dx = \int_{-\infty}^0 f(x) dx + \int_0^\infty f(x) dx.$$

Solche uneigentlichen Integrale spielen beispielsweise in der Statistik eine wichtige Rolle. Dort nennt man eine nichtnegative Funktion  $f : \mathbb{R} \rightarrow [0, \infty)$  Dichte, falls

$$\int_{-\infty}^\infty f(x) dx = 1.$$

Ein Beispiel für eine solche Dichte ist die Dichte der sogenannten Exponentialverteilung, welche gegeben ist durch

$$f(x) = \begin{cases} \alpha e^{-\alpha x}, & x > 0, \\ 0, & x < 0, \end{cases}$$

mit einem positiven Parameter  $\alpha > 0$ .

Eine wichtige Größe in der Wahrscheinlichkeitsrechnung ist die sogenannte Verteilungsfunktion

$$F(x) = \int_{-\infty}^x f(t) dt, \quad x \in \mathbb{R}.$$

Dies ist dann eine monoton wachsende Funktion mit

$$\lim_{x \rightarrow -\infty} F(x) = 0 \quad \text{und} \quad \lim_{x \rightarrow \infty} F(x) = 1.$$

Im Falle der obigen Exponentialverteilung erhalten wir für  $x > 0$

$$\begin{aligned} F(x) &= \int_{-\infty}^x f(t) dt = \int_{-\infty}^0 0 dt + \int_0^x \alpha e^{-\alpha t} dt \\ &= \left[ -e^{-\alpha t} \right]_0^x = -e^{-\alpha x} - (-e^{-\alpha \cdot 0}) \\ &= 1 - e^{-\alpha x}. \end{aligned}$$

Für  $x < 0$  gilt  $F(x) = 0$ , und wir sehen, dass tatsächlich gilt:

$$\int_{-\infty}^{\infty} f(t) dt = \lim_{x \rightarrow \infty} F(x) = \lim_{x \rightarrow \infty} (1 - e^{-\alpha x}) = 1.$$

**Beispiel 1.24.** Wir betrachten das Integral

$$\int_1^{\infty} \frac{1}{x^a} dx.$$

Im Falle  $a > 1$  erhalten wir für  $b > 1$

$$\int_1^b \frac{1}{x^a} dx = \int_1^b x^{-a} dx = \left[ \frac{x^{1-a}}{1-a} \right]_1^b = \frac{1}{1-a} (b^{1-a} - 1). \quad (1.11)$$

Damit ergibt sich für  $a > 1$

$$\int_1^{\infty} \frac{1}{x^a} dx = \lim_{b \rightarrow \infty} \frac{1}{1-a} (b^{1-a} - 1) = \frac{1}{a-1}.$$

Im Fall  $a = 1$  erhalten wir

$$\int_1^b \frac{1}{x} dx = \ln(b) - \ln(1) = \ln(b) \rightarrow \infty$$

für  $b \rightarrow \infty$ , und somit divergiert das Integral. Ebenso divergiert das Integral für  $a < 1$ , da die rechte Seite in Gleichung (1.11) gegen unendlich geht.

Analog zu Integralen über unbeschränkte Intervalle definiert man Integrale für unbeschränkte Funktionen, die Pole besitzen. Als Beispiel betrachte man die Funktion  $f(x) = 1/\sqrt{x}$ . Da kann man fragen, ob die Fläche unter dieser Funktion im Intervall  $(0, 2)$  endlich ist, und falls ja, welchen Wert sie hat. Da  $f(0)$  nicht definiert ist, und  $\lim_{h \rightarrow 0} f(h) = \infty$  gilt, behilft man sich wieder mit dem entsprechenden Grenzwert. Allgemein definiert man für den Fall, dass  $f(x)$  an der Stelle  $x = a$  nicht definiert ist, das uneigentliche Integral

$$\int_a^b f(x) dx = \lim_{h \rightarrow 0^+} \int_{a+h}^b f(x) dx,$$

falls der Grenzwert auf der rechten Seite existiert, und analog, falls  $f(x)$  an der Stelle  $x = b$  nicht definiert ist, das uneigentliche Integral

$$\int_a^b f(x) dx = \lim_{h \rightarrow 0^+} \int_a^{b-h} f(x) dx.$$

Mit dieser Definition erhalten wir

$$\begin{aligned} \int_0^2 \frac{1}{\sqrt{x}} dx &= \lim_{h \rightarrow 0^+} \int_h^2 \frac{1}{\sqrt{x}} dx \\ &= \lim_{h \rightarrow 0^+} \left[ 2\sqrt{x} \right]_h^2 = \lim_{h \rightarrow 0^+} (2\sqrt{2} - 2\sqrt{h}) \\ &= 2\sqrt{2}. \end{aligned}$$

## 1.5 Integration von Funktionen mehrerer Veränderlicher

Ähnlich wie man Flächen unter dem Graphen von Funktionen  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit Hilfe der Integralrechnung berechnen kann, so kann man auch das Volumen unter dem “Gebirge” eines Graphen einer Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  mit Hilfe der Integralrechnung bestimmen.

Ein Quader mit Kantenlängen  $x$ ,  $y$  und  $z$  hat bekanntlich das Volumen  $V = x \cdot y \cdot z$ . Betrachtet man eine konstante Funktion von zwei Veränderlichen  $f(x, y) = M > 0$  auf einem Rechteck  $[a, b] \times [c, d]$ , so erhält man als Volumen unter dem Graphen

$$V = M \cdot (b - a) \cdot (d - c).$$

Dies erhält man auch, indem man die Funktion  $f(x, y)$  erst bezüglich  $x$  integriert über  $[a, b]$  und dann bezüglich  $y$  über  $[c, d]$ .

$$\begin{aligned} V &= \int_c^d \left( \int_a^b f(x, y) dx \right) dy = \int_c^d \left( \int_a^b M dx \right) dy \\ &= \int_c^d (M \cdot (b - a)) dy = M \cdot (b - a) \cdot (d - c). \end{aligned}$$

Ähnlich wie bei der Bestimmung von Flächeninhalten bei Funktionen einer Veränderlichen kann man nun zeigen, dass allgemein gilt, dass für eine nichtnegative stetige Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  das Volumen zwischen dem Graphen und dem Rechteck  $[a, b] \times [c, d]$  gegeben ist durch das Doppelintegral

$$V = \int_c^d \left( \int_a^b f(x, y) dx \right) dy.$$

Man kann zeigen, dass bei diesem Doppelintegral die Reihenfolge der Integration keine Rolle spielt. Dies ist als **Satz von Fubini** bekannt.

**Satz 1.25.** *Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  stetig. Dann gilt*

$$\int_c^d \left( \int_a^b f(x, y) dx \right) dy = \int_a^b \left( \int_c^d f(x, y) dy \right) dx.$$

**Beispiel 1.26.** Wir berechnen

$$\int_0^1 \left( \int_{-1}^3 (x^2 y + xy^2 + 2x) dy \right) dx.$$

Dazu betrachten wir zunächst  $x$  als Konstante und berechnen das Integral bezüglich  $y$

$$\begin{aligned} \int_{-1}^3 (x^2 y + xy^2 + 2x) dy &= \left[ \frac{1}{2} x^2 y^2 + \frac{1}{3} xy^3 + 2xy \right]_{y=-1}^{y=3} \\ &= 4x^2 + \frac{52}{3}x. \end{aligned}$$

Nun integrieren wir ein zweites Mal bezüglich  $x$  und erhalten

$$\begin{aligned} \int_0^1 \left( \int_{-1}^3 (x^2y + xy^2 + 2x) dy \right) dx &= \int_0^1 \left( 4x^2 + \frac{52}{3}x \right) dx \\ &= \left[ \frac{4}{3}x^3 + \frac{26}{3}x^2 \right]_{x=0}^{x=1} = 10. \end{aligned}$$

Hätten wir das Integral in umgekehrter Reihenfolge berechnet, so hätten wir nach Satz 1.25 das gleiche Resultat erhalten müssen. Tatsächlich erhalten wir auch wenn wir erst bezüglich  $x$  integrieren zunächst

$$\begin{aligned} \int_0^1 (x^2y + xy^2 + 2x) dx &= \left[ \frac{1}{3}x^3y + \frac{1}{2}x^2y^2 + x^2 \right]_{x=0}^{x=1} \\ &= \frac{1}{3}y + \frac{1}{2}y^2 + 1, \end{aligned}$$

und damit

$$\int_{-1}^3 \left( \int_0^1 (x^2y + xy^2 + 2x) dx \right) dy = \int_{-1}^3 \left( \frac{1}{3}y + \frac{1}{2}y^2 + 1 \right) dy = 10.$$

Statt über ein Rechteck der Form  $A = [a, b] \times [c, d]$  zu integrieren, um das Volumen zwischen der Fläche  $A$  und dem Graphen der Funktion zu berechnen, kann man auch das Volumen über komplizierteren Grundflächen berechnen mit Hilfe von Integralen. Insbesondere kann man Grundflächen der Form

$$A = \{(x, y) : u(x) \leq y \leq v(x), a \leq x \leq b\}$$

mit integrierbaren Funktionen  $u(x) \leq v(x)$  auf  $[a, b]$  betrachten. Dann erhält man für das Volumen

$$V = \int_a^b \left( \int_{u(x)}^{v(x)} f(x, y) dy \right) dx.$$

Als Beispiel betrachten wir für  $A$  die Dreiecksfläche

$$A = \{(x, y) : 0 \leq y \leq 1 - x, 0 \leq x \leq 1\}.$$

Wir berechnen darüber das Integral der Funktion  $f(x, y) = x$  und erhalten

$$\begin{aligned} V &= \int_0^1 \left( \int_0^{1-x} x dy \right) dx \\ &= \int_0^1 [xy]_{y=0}^{y=1-x} dx = \int_0^1 x(1-x) dx \\ &= \left[ \frac{1}{2}x^2 - \frac{1}{3}x^3 \right]_0^1 = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}. \end{aligned}$$

Die geometrische Interpretation dieser Integralberechnung ist die Berechnung einer Pyramide mit einer Grundfläche  $1/2$  und einer Höhe  $1$ .

## 2 Lineare Algebra

In der linearen Algebra spielen Vektoren und Matrizen eine wichtige Rolle. Wir wiederholen deshalb zunächst die grundlegenden Begriffe und Notationen. Wir verwenden im Folgenden fette Buchstaben wie  $\mathbf{A}, \mathbf{x}, \dots$  für Matrizen und Vektoren, und zwar in der Regel grosse Fettsbuchstaben für Matrizen, und kleine Fettsbuchstaben für Vektoren. Ein rechteckiges Zahlenschema aus  $m$  Zeilen und  $n$  Spalten heisst  $m \times n$ -**Matrix**.

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

Die  $a_{ij} \in \mathbb{R}$  heissen **Elemente** der Matrix. Der erste Index  $i$  beschreibt die Zeile und der zweite Index  $j$  die Spalte.

Wir benutzen auch folgende Kurzschreibweisen:  $\mathbf{A} = (a_{ij})_{m \times n} = (a_{ij})$ . Durch Vertauschung von Zeilen und Spalten erhält man die **transponierte Matrix**  $\mathbf{A}^T = (a_{ji})$ .

Eine Matrix mit  $m = n$  heisst **quadratische Matrix**. Die Elemente  $a_{11}, \dots, a_{nn}$  heissen dann **Diagonalelemente**.

Die quadratische Matrix

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

mit

$$a_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases}$$

heißt **Einheitsmatrix (Identitätsmatrix)**.

Eine Matrix mit nur einer Zeile oder Spalte heisst **Vektor** und wird mit kleinen fetten Buchstaben beschrieben.

Eine  $n \times 1$ -Matrix mit nur einer Spalte heisst **Spaltenvektor**:

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

Eine  $1 \times n$ -Matrix mit nur einer Zeile heisst **Zeilenvektor**:

$$\mathbf{x} = (x_1 \ x_2 \ \dots \ x_n).$$

Bei Zeilenvektoren trennt man die Zahlen oft durch Kommata:

$$\mathbf{x} = (x_1, x_2, \dots, x_n).$$

Da man durch Transponieren eines Zeilenvektors einen Spaltenvektor erhält, schreibt man dann oft Spaltenvektoren platzsparend in der Form  $\mathbf{x}^T = (x_1, x_2, \dots, x_n)^T$ .

Wir setzen oft stillschweigend voraus, dass es sich bei einem Vektor um einen Spaltenvektor handelt, insbesondere wenn wir eine Matrix  $\mathbf{A}$  mit einem Vektor  $\mathbf{x}$  multiplizieren, so verstehen wir darunter das Produkt aus der Matrix  $\mathbf{A}$  mit dem Spaltenvektor  $\mathbf{x}$ :

$$\mathbf{Ax} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ a_{21}x_1 + \dots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix}.$$

## 2.1 Definition von Eigenwerten

Wir beginnen mit der Definition eines Eigenwertes.

**Definition 2.1.** Sei  $\mathbf{A}$  eine  $n \times n$ -Matrix. Gilt für einen Vektor  $\mathbf{x} \neq \mathbf{0}$

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

für eine reelle Zahl  $\lambda$ , so nennt man  $\mathbf{x}$  einen **Eigenvektor** von  $\mathbf{A}$  zum **Eigenwert**  $\lambda$ .

Wegen

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 10 \end{pmatrix} = 5 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

hat also zum Beispiel die Matrix

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

den Eigenvektor  $\mathbf{x}_1 = (1, 2)^T$  zum Eigenwert  $\lambda_1 = 5$ . Wegen

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0 \cdot \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

hat die gleiche Matrix  $\mathbf{A}$  auch den Eigenvektor  $\mathbf{x}_2 = (-2, 1)^T$  zum Eigenwert  $\lambda_2 = 0$ .

Bevor wir uns im Detail damit beschäftigen, wie man Eigenvektoren und Eigenwerte berechnen kann, wollen wir an einem Beispiel deutlich machen, warum diese sehr nützlich sind.

**Beispiel 2.2.** Wir betrachten die bekannte Fibonacci-Folge  $0, 1, 1, 2, 3, 5, 8, 13, 21, 34, \dots$ , bei welcher das nächste Folgenglied immer als die Summe der beiden davor liegenden gebildet wird, also  $x_{n+1} = x_n + x_{n-1}$  für alle  $n \geq 1$  mit  $y_0 = 0$ ,  $y_1 = 1$ . Mit Hilfe von Vektoren und Matrizen lässt sich das schreiben in der Form

$$\mathbf{x}_{n+1} = \begin{pmatrix} x_{n+1} \\ x_{n+2} \end{pmatrix} = \mathbf{Ax}_n = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x_n \\ x_{n+1} \end{pmatrix}.$$

Folglich erhalten wir durch Induktion

$$\mathbf{x}_n = \mathbf{A}^n \mathbf{x}_0$$

wobei  $\mathbf{A}^n$  das  $n$ -fache Matrizenprodukt von  $\mathbf{A}$  sein soll, das man erhält, wenn man  $\mathbf{A}$   $n$ -mal mit sich selbst multipliziert.

Ist nun  $\mathbf{x}$  ein Eigenwert von  $\mathbf{A}$  zum Eigenwert  $\lambda$ , so gilt

$$\mathbf{A}^2 \mathbf{x} = \mathbf{A}(\mathbf{A}\mathbf{x}) = \mathbf{A}(\lambda \mathbf{x}) = \lambda(\mathbf{A}\mathbf{x}) = \lambda(\lambda \mathbf{x}) = \lambda^2 \mathbf{x},$$

und allgemeiner

$$\mathbf{A}^n \mathbf{x} = \lambda^n \mathbf{x}.$$

Im Fall der hier gegebenen  $2 \times 2$ -Matrix  $\mathbf{A}$  findet man zwei verschiedene Eigenvektoren  $\mathbf{z}_1$  und  $\mathbf{z}_2$  mit zugehörigen Eigenwerten  $\lambda_1$  und  $\lambda_2$ . Wir werden später sehen, wie man diese berechnet. Wenn wir nun noch den Startvektor  $\mathbf{x}_0 = (0, 1)^T$  darstellen als Linearkombination dieser Eigenvektoren in der Form

$$\mathbf{x}_0 = \alpha \mathbf{z}_1 + \beta \mathbf{z}_2,$$

was man leicht machen kann durch Lösung des zugehörigen linearen Gleichungssystems mit zwei Gleichungen und zwei Unbekannten, so kann man schliessen dass gilt

$$\begin{aligned} \mathbf{x}_n &= \mathbf{A}^n \mathbf{x}_0 = \mathbf{A}^n (\alpha \mathbf{z}_1 + \beta \mathbf{z}_2) \\ &= \alpha \mathbf{A}^n \mathbf{z}_1 + \beta \mathbf{A}^n \mathbf{z}_2 \\ &= \alpha \lambda_1^n \mathbf{z}_1 + \beta \lambda_2^n \mathbf{z}_2. \end{aligned}$$

Wir werden uns in nächster Zeit damit beschäftigen, wie man die konkreten Werte von  $\lambda_1, \lambda_2, \mathbf{z}_1, \mathbf{z}_2, \alpha$  und  $\beta$  berechnet. Hier sei aber schon mal das Ergebnis verraten, das man für die Fibonacci-Zahlen bekommt. Man erhält auf diese Weise für die  $n$ -te Fibonacci-Zahl  $x_n$  die Darstellung

$$x_n = \frac{1}{\sqrt{5}} \left( \frac{1 + \sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left( \frac{1 - \sqrt{5}}{2} \right)^n.$$

Da Fibonacci-Zahlen ganzzahlig sind, und für den zweiten Term

$$\left| \frac{1}{\sqrt{5}} \left( \frac{1 - \sqrt{5}}{2} \right)^n \right| < \frac{1}{2}$$

für alle  $n \in \mathbb{N}$  gilt, erhält man das interessante Ergebnis, dass man die  $n$ -te Fibonacci-Zahl dadurch erhält, dass man

$$\frac{1}{\sqrt{5}} \left( \frac{1 + \sqrt{5}}{2} \right)^n$$

auf die nächste ganze Zahl rundet. Der dabei vorkommende Bruch  $(1 + \sqrt{5})/2$  ist übrigens als der berühmte *goldene Schnitt* bekannt.

Durch Umformulierung der Definition eines Eigenwertes erhält man die Aussage, dass eine Zahl  $\lambda$  ein Eigenwert einer Matrix  $\mathbf{A}$  ist, wenn die Gleichung

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0} \quad (2.1)$$

eine nichttriviale Lösung  $\mathbf{x} \neq \mathbf{0}$  hat, wobei  $\mathbf{I}$  die Identitätsmatrix darstellt. Man beachte, dass ein homogenes Gleichungssystem der Form  $\mathbf{A}\mathbf{x} = \mathbf{0}$  immer die triviale Lösung  $\mathbf{x} = \mathbf{0}$  besitzt. Gesucht ist also eine weitere nichttriviale Lösung von Gleichung (2.6).

## 2.2 Determinanten

Die Frage, ob ein Gleichungssystem eine eindeutige Lösung hat, löst man oft mit Hilfe sogenannter **Determinanten**. Wir betrachten zunächst den Fall von Dimension  $n = 2$ . Gegeben sei also ein beliebiges Gleichungssystem mit zwei Variablen und zwei Unbekannten der Form  $\mathbf{A}\mathbf{x} = \mathbf{b}$ , also

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1, \\ a_{21}x_1 + a_{22}x_2 &= b_2. \end{aligned}$$

Löst man dieses in der üblichen Weise mit dem Gauß-Algorithmus, so erhält man als eindeutige Lösung

$$x_1 = \frac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{12}a_{21}}, \quad x_2 = \frac{b_2a_{11} - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}},$$

falls der gemeinsame Nenner  $a_{11}a_{22} - a_{12}a_{21}$  von Null verschieden ist. Dieser gemeinsame Nenner entscheidet also, ob das Gleichungssystem eine eindeutige Lösung hat, oder nicht. Man beachte, dass er durch die Einträge der Matrix  $\mathbf{A}$  bestimmt wird, und unabhängig vom Vektor  $\mathbf{b}$  ist. Man nennt ihn die **Determinante** der Matrix  $\mathbf{A}$ .

**Definition 2.3.** Sei  $\mathbf{A}$  eine  $2 \times 2$ -Matrix. Dann schreiben wir

$$\det(\mathbf{A}) = |\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

und nennen diesen Ausdruck die **Determinante** der Matrix  $\mathbf{A}$ .

Wir können damit folgenden Satz formulieren, mit Hilfe dessen sich prüfen lässt, ob ein Gleichungssystem eine eindeutige Lösung besitzt.

**Satz 2.4.** *Das Gleichungssystem  $\mathbf{A}\mathbf{x} = \mathbf{b}$  hat eine eindeutige Lösung genau dann, wenn*

$$|\mathbf{A}| \neq 0.$$

Wir wollen dies nun zunächst auf den Fall von Dimension  $n = 3$  übertragen, und dann auf den Fall einer beliebigen Dimension  $n$  verallgemeinern.



Löst man ein allgemeines lineares Gleichungssystem mit drei Gleichungen und drei Variablen der Form

$$\begin{aligned}a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1, \\a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2, \\a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3,\end{aligned}$$

so erhält man nach längerer Rechnung in der Lösung bei allen drei Variablen im Nenner den Ausdruck

$$a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}. \quad (2.2)$$

Den Ausdruck in (2.2) nennt man wieder die Determinante der zugehörigen Matrix  $\mathbf{A}$ . Der Ausdruck sieht zwar sehr chaotisch aus, aber es lässt sich bei genauerem Hinsehen ein Entwicklungsgesetz ableiten. Zunächst kann man feststellen, dass die drei Elemente  $a_{11}$ ,  $a_{12}$  und  $a_{13}$  der ersten Zeile der Matrix jeweils genau zweimal vorkommen. Ordnet man die Terme nach diesen Ausdrücken, so erhält man

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

Nun sieht man, dass die Ausdrücke in Klammern jeweils Determinanten der Ordnung 2 sind:

$$|\mathbf{A}| = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}. \quad (2.3)$$

Man erkennt, dass bei  $a_{1k}$  jeweils die Determinante der  $2 \times 2$ -Matrix steht, welche man erhält, wenn man in der  $3 \times 3$ -Matrix  $\mathbf{A}$  die erste Zeile und die  $k$ -te Spalte streicht ( $k = 1, 2, 3$ ).

Dieses Entwicklungsgesetz wollen wir nun benutzen, um allgemein eine Determinante  $n$ -ter Ordnung für eine  $n \times n$ -Matrix zu definieren. Dazu bezeichnen wir für eine beliebige  $n \times n$ -Matrix  $\mathbf{A}$  mit  $\mathbf{A}_{ij}$  die  $(n-1) \times (n-1)$ -Matrix, welche man erhält, wenn man in  $\mathbf{A}$  die  $i$ -te Zeile und die  $j$ -te Spalte streicht. Man nennt  $\mathbf{A}_{ij}$  einen **Co-Faktor**. Mit Hilfe solcher Co-Faktoren lässt sich die Determinante einer  $3 \times 3$ -Matrix in (2.3) schreiben als

$$|\mathbf{A}| = a_{11}|\mathbf{A}_{11}| - a_{12}|\mathbf{A}_{12}| + a_{13}|\mathbf{A}_{13}|. \quad (2.4)$$

Dies lässt sich nun verallgemeinern, um für eine Matrix beliebiger Dimension eine Determinante zu definieren.

**Definition 2.5.** Sei  $\mathbf{A}$  eine beliebige  $n \times n$ -Matrix. Für  $n = 1$  ist die Determinante von  $\mathbf{A}$  definiert durch

$$|\mathbf{A}| = a_{11}.$$

Für  $n > 1$  ist die Determinante rekursiv definiert durch

$$\begin{aligned}|\mathbf{A}| &= a_{11}|\mathbf{A}_{11}| - a_{12}|\mathbf{A}_{12}| + \dots + (-1)^{n+1}a_{1n}|\mathbf{A}_{1n}| \\ &= \sum_{j=1}^n (-1)^{j+1}a_{1j}|\mathbf{A}_{1j}|.\end{aligned}$$

Diese Definition ist konsistent mit den bereits angegebenen Formeln im Fall  $n = 2$  und  $n = 3$ :

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

und

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}).$$

Zur konkreten Berechnung benutzt man im Fall  $n = 3$  oft die sogenannte **Sarrus'sche Regel**. Dazu schreibt man rechts neben die Matrix zusätzlich noch einmal die ersten beiden Spalten,

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

und bildet für jede der insgesamt sechs Diagonalen (drei von links unten nach rechts oben, und drei von rechts oben nach links unten) das Produkt der Diagonalelemente. Dann addiert man die Produkte der von links oben nach rechts unten abwärts verlaufenden Diagonalen, und subtrahiert die Produkte der drei von links unten nach rechts oben aufwärts verlaufenden Diagonalen. Damit erhält man die Formel der Determinante wie oben

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

**Beispiel 2.6.** Wir suchen die Determinante

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

Wir erhalten für die Regel von Sarrus folgendes Schema:

$$\begin{array}{cccccc} 1 & 0 & 5 & 1 & 0 & \\ 0 & 2 & 7 & 0 & 2 & \\ -1 & 2 & 3 & -1 & 2 & \end{array}$$

und damit

$$\begin{aligned} |\mathbf{A}| &= 1 \cdot 2 \cdot 3 + 0 \cdot 7 \cdot (-1) + 5 \cdot 0 \cdot 2 - (-1) \cdot 2 \cdot 5 - 2 \cdot 7 \cdot 1 - 3 \cdot 0 \cdot 0 \\ &= 6 + 10 - 14 = 2. \end{aligned}$$

Eine vergleichbare Regel von Sarrus für Dimension  $n \geq 4$  gilt nicht!

Für Dimension  $n \geq 4$  benutzt man Rechenregeln für Determinanten, mit denen wir uns als nächstes beschäftigen werden.

Zunächst bemerken wir, dass durch Transponieren einer Matrix sich die Determinante nicht ändert.

**Satz 2.7.** Für jede quadratische Matrix  $\mathbf{A}$  gilt

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

Bei elementaren Umformungen von Matrizen gelten folgende Rechenregeln für Determinanten

**Satz 2.8.** a) Entsteht  $\tilde{\mathbf{A}}$  aus  $\mathbf{A}$  durch Vertauschen zweier Zeilen oder Spalten, so gilt

$$\det(\tilde{\mathbf{A}}) = -\det(\mathbf{A}),$$

b) Entsteht  $\tilde{\mathbf{A}}$  aus  $\mathbf{A}$  durch Addition des  $\lambda$ -fachen einer Zeile oder Spalte ( $\lambda \in \mathbb{R}$ ), so gilt

$$\det(\tilde{\mathbf{A}}) = \det(\mathbf{A}),$$

c) Entsteht  $\tilde{\mathbf{A}}$  aus  $\mathbf{A}$  durch Multiplikation einer Zeile oder Spalte mit  $\lambda \in \mathbb{R}$ , so gilt

$$\det(\tilde{\mathbf{A}}) = \lambda \cdot \det(\mathbf{A}).$$

Mit Hilfe solcher elementaren Umformungen kann man ähnlich wie beim Gauß-Algorithmus eine Matrix in eine obere Dreiecksmatrix umwandeln. Für letztere lässt sich aber die Determinante sehr einfach angeben.

**Satz 2.9.** Sei  $\mathbf{A}$  eine obere Dreiecksmatrix (also  $a_{ij} = 0$  für  $i > j$ ). Dann ist die Determinante das Produkt der Diagonalelemente

$$\det(\mathbf{A}) = \prod_{i=1}^n a_{ii} = a_{11} \cdot a_{22} \cdot \dots \cdot a_{nn}.$$

Mit Hilfe der Sätze 2.8 und 2.9 lassen sich Determinanten von Matrizen beliebiger Dimension ähnlich wie beim Gauß-Algorithmus berechnen.

**Beispiel 2.10.**

$$\begin{aligned} \begin{vmatrix} 2 & 2 & 4 & 6 \\ 1 & 1 & 3 & 7 \\ 4 & 6 & 2 & 9 \\ 2 & 0 & 6 & 1 \end{vmatrix} &= \begin{vmatrix} 2 & 2 & 4 & 6 \\ 0 & 0 & 1 & 4 \\ 0 & 2 & -6 & -3 \\ 0 & -2 & 2 & -5 \end{vmatrix} = - \begin{vmatrix} 2 & 2 & 4 & 6 \\ 0 & 2 & -6 & -3 \\ 0 & 0 & 1 & 4 \\ 0 & -2 & 2 & -5 \end{vmatrix} \\ &= - \begin{vmatrix} 2 & 2 & 4 & 6 \\ 0 & 2 & -6 & -3 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & -4 & -8 \end{vmatrix} = - \begin{vmatrix} 2 & 2 & 4 & 6 \\ 0 & 2 & -6 & -3 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 8 \end{vmatrix} = -32. \end{aligned}$$

Nach Satz 2.8 a) ändert sich beim Vertauschen von Zeilen höchstens das Vorzeichen einer Determinante. Da nach Satz 2.7 sich beim transponieren einer Matrix die Determinante nicht ändert, gilt das gleiche also auch für Spalten. Dies bedeutet aber, dass man in der Definition der Determinante die Entwicklung nach der ersten Zeile auch ersetzen kann durch die Entwicklung nach einer beliebigen Zeile oder Spalte.

**Satz 2.11.** a) Sei  $\mathbf{A}$  eine  $n \times n$ -Matrix und  $1 \leq i \leq n$ . Dann gilt

$$\det(\mathbf{A}) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(\mathbf{A}_{ij}).$$

b) Sei  $\mathbf{A}$  eine  $n \times n$ -Matrix und  $1 \leq j \leq n$ . Dann gilt

$$\det(\mathbf{A}) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(\mathbf{A}_{ij}).$$

Dieser Satz lässt sich oft vorteilhaft anwenden, indem man die Zeile oder Spalte, nach der man entwickelt, so wählt, dass sie möglichst viele Nullen enthält.

**Beispiel 2.12.** Für

$$\mathbf{A} = \begin{pmatrix} 4 & 0 & 1 & 0 \\ 2 & 1 & 2 & 5 \\ 0 & 2 & 0 & 1 \\ 5 & 0 & 2 & 0 \end{pmatrix}$$

entwickeln wir zuerst nach der ersten Zeile und erhalten

$$\det(\mathbf{A}) = \begin{vmatrix} 4 & 0 & 1 & 0 \\ 2 & 1 & 2 & 5 \\ 0 & 2 & 0 & 1 \\ 5 & 0 & 2 & 0 \end{vmatrix} = 4 \cdot \begin{vmatrix} 1 & 2 & 5 \\ 2 & 0 & 1 \\ 0 & 2 & 0 \end{vmatrix} + 1 \cdot \begin{vmatrix} 2 & 1 & 5 \\ 0 & 2 & 1 \\ 5 & 0 & 0 \end{vmatrix}.$$

Nun entwickeln wir die beiden  $3 \times 3$ -Matrizen jeweils nach der dritten Zeile.

$$\begin{vmatrix} 1 & 2 & 5 \\ 2 & 0 & 1 \\ 0 & 2 & 0 \end{vmatrix} = (-2) \cdot \begin{vmatrix} 1 & 5 \\ 2 & 1 \end{vmatrix} = (-2) \cdot (1 \cdot 1 - 2 \cdot 5) = 18,$$

bzw.

$$\begin{vmatrix} 2 & 1 & 5 \\ 0 & 2 & 1 \\ 5 & 0 & 0 \end{vmatrix} = 5 \cdot \begin{vmatrix} 1 & 5 \\ 2 & 1 \end{vmatrix} = 5 \cdot (1 \cdot 1 - 2 \cdot 5) = -45.$$

Damit erhält man

$$\det(\mathbf{A}) = 4 \cdot 18 - 45 = 27.$$

Des weiteren gelten folgende Rechenregeln.

**Satz 2.13.** *Seien  $\mathbf{A}, \mathbf{B}$  zwei  $n \times n$ -Matrizen und  $\lambda$  eine reelle Zahl. Dann gilt*

a)  $\det(\lambda \mathbf{A}) = \lambda^n \det(\mathbf{A}),$

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

Wir wollen nun noch eine geometrische Interpretation einer Determinante erwähnen. Gegeben seien Vektoren  $\mathbf{a} = (a_1, a_2)^T$  und  $\mathbf{b} = (b_1, b_2)^T$ . Wir betrachten das Parallelogramm, das von diesen beiden Vektoren aufgespannt wird. Formal ist das die Menge der Punkte

$$F = \{\alpha \mathbf{a} + \beta \mathbf{b} : 0 \leq \alpha, \beta \leq 1\}.$$

Dann ist die Fläche des Parallelogramms  $F$  gerade der Betrag der Determinante der Matrix

$$\mathbf{A} = (\mathbf{a}, \mathbf{b}) = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}.$$

Bild

Entsprechend gilt ein ähnlicher Sachverhalt im dreidimensionalen Raum. Gegeben seien Vektoren  $\mathbf{a} = (a_1, a_2, a_3)^T$ ,  $\mathbf{b} = (b_1, b_2, b_3)^T$  und  $\mathbf{c} = (c_1, c_2, c_3)^T$ . Wir betrachten das Parallelotop, das von diesen drei Vektoren aufgespannt wird. Dies ist die Menge der Punkte

$$V = \{\alpha \mathbf{a} + \beta \mathbf{b} + \gamma \mathbf{c} : 0 \leq \alpha, \beta, \gamma \leq 1\}.$$

Dann ist das Volumen des Parallelotops  $V$  gerade der Betrag der Determinante der Matrix

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

Bild

Der wichtigste Grund für die Einführung von Determinanten war aber Satz 2.4, der nun mit der allgemeinen Definition einer Determinante in beliebiger Dimension gilt.

**Satz 2.14.** Das Gleichungssystem  $\mathbf{Ax} = \mathbf{b}$  hat eine eindeutige Lösung genau dann, wenn

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

## 2.3 Lineare Unabhängigkeit, Basis und Dimension

Insbesondere besitzt also ein homogenes lineares Gleichungssystem  $\mathbf{Ax} = \mathbf{0}$  nur die triviale Lösung  $\mathbf{x} = \mathbf{0}$ , falls  $\det(\mathbf{A}) \neq 0$ . Wir beschäftigen uns nun im allgemeinen Fall mit der Struktur der Lösungsmenge

$$L = \{\mathbf{x} : \mathbf{Ax} = \mathbf{0}\}.$$

Man sieht sofort, dass  $L$  die beiden folgenden Eigenschaften hat.

1. Sind  $\mathbf{x}$  und  $\mathbf{y}$  in  $L$ , so ist auch  $\mathbf{x} + \mathbf{y} \in L$ .
2. Ist  $\mathbf{x} \in L$ , so ist auch  $\lambda\mathbf{x} \in L$  für jedes  $\lambda \in \mathbb{R}$ .

Die erste Eigenschaft folgt, weil aus  $\mathbf{Ax} = \mathbf{0}$  und  $\mathbf{Ay} = \mathbf{0}$  auch

$$\mathbf{A}(\mathbf{x} + \mathbf{y}) = \mathbf{Ax} + \mathbf{Ay} = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

folgt, und die zweite Behauptung folgt, weil aus  $\mathbf{Ax} = \mathbf{0}$  auch

$$\mathbf{A}(\lambda\mathbf{x}) = \lambda(\mathbf{Ax}) = \lambda\mathbf{0} = \mathbf{0}$$

folgt.

Man nennt  $L$  deshalb einen linearen Unterraum des  $\mathbb{R}^n$ , weil er diese beiden Eigenschaften hat.

**Definition 2.15.** Eine Teilmenge  $U \subset \mathbb{R}^n$  heißt **linearer Unterraum**, falls folgende beiden Eigenschaften gelten:

1. Aus  $\mathbf{x} \in U$  und  $\mathbf{y} \in U$  folgt  $\mathbf{x} + \mathbf{y} \in U$ .
2. Ist  $\mathbf{x} \in U$ , so ist auch  $\lambda\mathbf{x} \in U$  für jedes  $\lambda \in \mathbb{R}$ .

**Definition 2.16.** a) Für Vektoren  $\mathbf{x}_1, \dots, \mathbf{x}_m$  nennen wir eine Summe

$$\lambda_1\mathbf{x}_1 + \dots + \lambda_m\mathbf{x}_m$$

mit  $\lambda_1, \dots, \lambda_m \in \mathbb{R}$  eine **Linearkombination** dieser Vektoren. Für eine beliebige endliche Menge  $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$  von Vektoren gleicher Dimension nennen wir die Menge aller Linearkombinationen

$$\begin{aligned} [\mathbf{x}_1, \dots, \mathbf{x}_m] &:= \{\lambda_1\mathbf{x}_1 + \dots + \lambda_m\mathbf{x}_m : \lambda_1, \dots, \lambda_m \in \mathbb{R}\} \\ &= \left\{ \sum_{i=1}^m \lambda_i\mathbf{x}_i : \lambda_1, \dots, \lambda_m \in \mathbb{R} \right\} \end{aligned}$$

die **lineare Hülle** von  $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ .

Es ist sofort einzusehen, dass die lineare Hülle einer beliebigen Menge von Vektoren ein Unterraum ist, denn sind  $\mathbf{x} = \lambda_1 \mathbf{x}_1 + \dots + \lambda_m \mathbf{x}_m$  und  $\mathbf{y} = \mu_1 \mathbf{x}_1 + \dots + \mu_m \mathbf{x}_m$  Linearkombinationen, so ist auch

$$\mathbf{x} + \mathbf{y} = (\lambda_1 + \mu_1) \mathbf{x}_1 + \dots + (\lambda_m + \mu_m) \mathbf{x}_m$$

eine Linearkombination, und  $\lambda \mathbf{x} = \lambda \lambda_1 \mathbf{x}_1 + \dots + \lambda \lambda_m \mathbf{x}_m$  ebenfalls.

Gibt es für einen linearen Unterraum  $U \subset \mathbb{R}^n$  eine Menge  $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ , deren lineare Hülle gerade  $U$  ist, also  $[\mathbf{x}_1, \dots, \mathbf{x}_m] = U$ , so nennt man die Menge  $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$  **Erzeugendensystem** von  $U$ . Man ist an möglichst kleinen Erzeugendensystemen interessiert. Betrachten wir zum Beispiel im  $\mathbb{R}^3$  den von den beiden ersten Einheitsvektoren  $\mathbf{e}_1 = (1, 0, 0)^T$  und  $\mathbf{e}_2 = (0, 1, 0)^T$  erzeugten Unterraum

$$U = [\mathbf{e}_1, \mathbf{e}_2] = \left\{ \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ 0 \end{pmatrix} : \lambda_1, \lambda_2 \in \mathbb{R} \right\},$$

so erhalten wir den Unterraum bestehend aus der Ebene der Punkte, deren dritten Koordinate verschwindet. Nehmen wir in das Erzeugendensystem zusätzlich den Vektor  $\mathbf{x} = (1, 1, 0)^T$  hinzu, so erhalten wir als lineare Hülle den gleichen Unterraum

$$U = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{x}] = \left\{ \begin{pmatrix} \lambda_1 + \lambda_3 \\ \lambda_2 + \lambda_3 \\ 0 \end{pmatrix} : \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ 0 \end{pmatrix} : \lambda_1, \lambda_2 \in \mathbb{R} \right\}.$$

Wir können also auf  $\mathbf{x}$  verzichten, wenn wir ein kleines Erzeugendensystem suchen. Der Grund dafür ist, dass sich  $\mathbf{x} = \mathbf{e}_1 + \mathbf{e}_2$  als Linearkombination der beiden anderen Vektoren schreiben lässt. Wir sprechen von linear unabhängigen Vektoren, wenn sie sich nicht als Linearkombination voneinander schreiben lassen.

**Definition 2.17.** Eine Menge von Vektoren  $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$  heißt **linear unabhängig**, falls die Gleichung

$$\lambda_1 \mathbf{x}_1 + \dots + \lambda_m \mathbf{x}_m = \mathbf{0}$$

nur für  $\lambda_1 = \dots = \lambda_m = 0$  erfüllt ist. Die Menge heißt **linear abhängig**, falls dies nicht gilt.

Die obige Menge  $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{x}\}$  ist linear abhängig, weil  $\mathbf{x} = \mathbf{e}_1 + \mathbf{e}_2$  gilt und somit die Gleichung  $\lambda_1 \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 + \lambda_3 \mathbf{x} = \mathbf{0}$  auch die Lösung  $\lambda_1 = 1, \lambda_2 = 1, \lambda_3 = -1$  besitzt.

Allgemein gilt: Sind  $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$  linear abhängig, so gibt es also Zahlen  $\lambda_1, \dots, \lambda_m$ , die nicht alle Null sind, so dass

$$\lambda_1 \mathbf{x}_1 + \dots + \lambda_m \mathbf{x}_m = \mathbf{0}$$

Ist also z.B.  $\lambda_1 \neq 0$ , so kann man die Gleichung nach  $\mathbf{x}_1$  auflösen, und erhält

$$\mathbf{x}_1 = -\frac{\lambda_2}{\lambda_1} \mathbf{x}_2 - \dots - \frac{\lambda_m}{\lambda_1} \mathbf{x}_m,$$

also ist  $\mathbf{x}_1$  eine Linearkombination von  $\mathbf{x}_2, \dots, \mathbf{x}_m$ . Genauso kann man  $\mathbf{x}_k$  als Linearkombination der anderen Vektoren darstellen, falls  $\lambda_k \neq 0$ . Bei linear abhängigen Vektoren ist also mindestens einer als Linearkombination der anderen darstellbar.

**Beispiel 2.18.** Gegeben seien die Vektoren

$$\mathbf{x} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad \mathbf{z} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \mathbf{u} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Die Menge  $\{\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{u}\}$  ist linear abhängig, da

$$3\mathbf{x} + (-1)\mathbf{y} + (-1)\mathbf{z} + 0\mathbf{u} = \mathbf{0}. \quad (2.5)$$

Die Menge  $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$  ist ebenfalls linear abhängig, da man in Gleichung (2.5) den Ausdruck  $+0\mathbf{u}$  natürlich auch weglassen kann.

Die Menge  $\{\mathbf{x}, \mathbf{y}, \mathbf{u}\}$  ist linear unabhängig. Es ist nämlich  $\lambda_1\mathbf{x} + \lambda_2\mathbf{y} + \lambda_3\mathbf{z} = \mathbf{0}$ , falls

$$\begin{aligned} \lambda_1 + \lambda_2 + \lambda_3 &= 0 \\ \lambda_1 + 2\lambda_2 + \lambda_3 &= 0 \\ \lambda_3 &= 0 \end{aligned}$$

Subtraktion der ersten Gleichung von der zweiten ergibt  $\lambda_2 = 0$ , und damit folgt dann aus der ersten Gleichung auch  $\lambda_1 = 0$ . Also ist  $\lambda_1 = \lambda_2 = \lambda_3 = 0$  die einzige Lösung der Gleichung.

Eine Menge von linear unabhängigen Vektoren heißt Basis des von ihr aufgespannten Unterraums.

**Definition 2.19.** Sei  $U$  ein Unterraum des  $\mathbb{R}^n$ . Eine Menge  $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$  von Vektoren heißt **Basis** von  $U$ , wenn folgende zwei Eigenschaften gelten:

1.  $\mathbf{x}_1, \dots, \mathbf{x}_m$  sind linear unabhängig;
2.  $[\mathbf{x}_1, \dots, \mathbf{x}_m] = U$ .

Für die Existenz von Basen gilt folgender Satz.

**Satz 2.20.** *Jeder Unterraum  $U$  des  $\mathbb{R}^n$  besitzt eine Basis. Ferner haben alle Basen von  $U$  die gleiche Anzahl an Vektoren.*

Dies gibt Anlass zu folgender Begriffsbildung.

**Definition 2.21.** Sei  $U$  ein Unterraum des  $\mathbb{R}^n$ . Dann heißt die Anzahl  $k$  der Elemente einer Basis **Dimension** des Unterraums. Man schreibt  $\dim(U) = k$ .

**Beispiel 2.22.** In der Regel hat ein Unterraum  $U$  viele verschiedene Basen. So hat z.B.  $U = \mathbb{R}^3$  offensichtlich die (sogenannte kanonische Basis) bestehend aus den drei Einheitsvektoren  $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ . Eine weitere Basis ist aber auch gegeben durch die Vektoren  $\{\mathbf{x}, \mathbf{y}, \mathbf{u}\}$  aus Beispiel 2.18. Um dies nachzuweisen, müssen wir zeigen, dass jeder beliebige Vektor  $\mathbf{a} = (a_1, a_2, a_3)^T$  sich als Linearkombination

$$\lambda_1\mathbf{x} + \lambda_2\mathbf{y} + \lambda_3\mathbf{u} = \mathbf{a}$$



darstellen lässt. Dies führt auf die Lösung eines linearen Gleichungssystems

$$\begin{aligned}\lambda_1 + \lambda_2 + \lambda_3 &= a_1 \\ \lambda_1 + 2\lambda_2 + \lambda_3 &= a_2 \\ \lambda_3 &= a_3\end{aligned}$$

Aus der dritten Gleichung folgt unmittelbar  $\lambda_3 = a_3$ . Subtraktion der ersten Gleichung von der zweiten liefert  $\lambda_2 = a_2 - a_1$ , und damit folgt dann  $\lambda_1 = 2a_1 - a_2 - a_3$ . Also ist

$$(2a_1 - a_2 - a_3)\mathbf{x} + (a_2 - a_1)\mathbf{y} + a_3\mathbf{u} = \mathbf{a},$$

d.h.  $[\mathbf{x}, \mathbf{y}, \mathbf{u}] = \mathbb{R}^3$ . Da wir schon gesehen haben, dass diese drei Vektoren linear unabhängig sind, bilden sie also eine Basis des  $\mathbb{R}^3$ .

Wir fassen die wichtigsten Aussagen über die Dimension von Unterräumen in einem Satz zusammen.

- Satz 2.23.** a) Sind  $U_1, U_2$  Unterräume mit  $U_1 \subseteq U_2$ , so ist  $\dim(U_1) \leq \dim(U_2)$ .  
 b) Sind  $U_1, U_2$  Unterräume mit  $U_1 \subseteq U_2$  und ist  $\dim(U_1) = \dim(U_2)$ , so ist  $U_1 = U_2$ .  
 c) Sind  $U_1, U_2$  Unterräume mit  $U_1 \subseteq U_2$  und ist  $\dim(U_1) < \dim(U_2)$ , so ist  $U_1 \neq U_2$ .  
 d) Ist  $U$  Unterraum von  $\mathbb{R}^n$ , so ist  $\dim(U) \leq n$ .  
 e) Für beliebige  $\mathbf{x}_1, \dots, \mathbf{x}_m \in \mathbb{R}^n$  gilt  $\dim([\mathbf{x}_1, \dots, \mathbf{x}_m]) \leq m$ .  
 f) Es gilt  $\dim([\mathbf{x}_1, \dots, \mathbf{x}_m]) = m$  genau dann, wenn  $\mathbf{x}_1, \dots, \mathbf{x}_m$  linear unabhängig sind.  
 g) Ist  $k > \dim(U)$ , so sind  $k$  Vektoren aus  $U$  immer linear abhängig.  
 h) Ist  $\dim(U) = k$ , so sind  $k$  linear unabhängige Vektoren aus  $U$  immer eine Basis von  $U$ .

## 2.4 Rang einer Matrix

Nach den Resultaten des vorigen Abschnittes entspricht die Dimension einer linearen Hülle  $[\mathbf{x}_1, \dots, \mathbf{x}_m]$  der maximalen Anzahl linear unabhängiger Vektoren in  $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$ . Ähnlich definiert man den Rang einer Matrix.

**Definition 2.24.** Sei  $A$  eine  $m \times n$ -Matrix. Wir betrachten  $A$  als Zusammenfassung von Spaltenvektoren

$$\mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_n),$$

wobei  $\mathbf{a}_i$  die  $i$ -te Spalte von  $\mathbf{A}$  bezeichnet. Der **Rang** von  $\mathbf{A}$  (als  $rg(\mathbf{A})$  bezeichnet) ist die maximale Anzahl linear unabhängiger Spalten von  $\mathbf{A}$ .

**Beispiel 2.25.** a) Gegeben sei die Matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 1 & 4 \\ 0 & 1 & 2 \end{pmatrix}$$

Dann sind die drei Spaltenvektoren von  $\mathbf{A}$  linear abhängig, da es sich um drei Vektoren in  $\mathbb{R}^2$  handelt. Es gilt

$$\begin{pmatrix} 2 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}.$$

Da die ersten beiden Spalten linear unabhängig sind, ist  $rg(\mathbf{A}) = 2$ .

b) Der Rang einer  $n \times n$ -Einheitsmatrix  $\mathbf{I}_n$  ist  $rg(\mathbf{I}_n) = n$ .

Offensichtlich gilt folgender Zusammenhang.

**Satz 2.26.** Sei  $\mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_n)$  eine  $m \times n$ -Matrix. Dann gilt

$$rg(\mathbf{A}) = \dim([\mathbf{a}_1, \dots, \mathbf{a}_n]).$$

Man kann sich natürlich auch nach der maximalen Anzahl linear unabhängiger Zeilen fragen. Da die Zeilen die Spalten der transponierten Matrix sind, ist dies äquivalent zur Frage nach dem Rang der transponierten Matrix. Dafür gilt folgender Satz.

**Satz 2.27.** Sei  $\mathbf{A}$  eine  $m \times n$ -Matrix. Dann gilt

$$rg(\mathbf{A}) = rg(\mathbf{A}^T).$$

Die maximale Anzahl linear unabhängiger Zeilen ist also gleich der maximalen Anzahl linear unabhängiger Spalten. Deshalb kann der Rang weder größer als die Spaltenzahl noch größer als die Zeilenzahl sein. Für eine  $m \times n$ -Matrix gilt also  $rg(\mathbf{A}) \leq \min\{m, n\}$ . Gilt  $rg(\mathbf{A}) = \min\{m, n\}$ , so sagt man, die Matrix habe **vollen Rang**.

Die Bestimmung des Ranges einer Matrix ist mit dem Gauß-Algorithmus möglich, da die dabei verwendeten Zeilenumformungen den Rang nicht ändern. Wenn man nach Durchführung des Gauß-Algorithmus aus der Matrix  $\mathbf{A}$  eine Matrix  $\tilde{\mathbf{A}}$  der Form

$$\tilde{\mathbf{A}} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1\ell} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2\ell} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & a_{\ell\ell} & \dots & a_{\ell n} \\ 0 & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

erhält, dann ist  $rg(\mathbf{A}) = \ell$ .

**Beispiel 2.28.** Gegeben sei die Matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 4 & 8 \\ 3 & 6 & 13 & 25 \\ 1 & 2 & 7 & 11 \end{pmatrix}.$$

Der Gauß-Algorithmus liefert folgende Umformungen. Wir ziehen die erste Zeile dreimal von der zweiten und einmal von der dritten ab und erhalten damit

$$\mathbf{A}_1 = \begin{pmatrix} 1 & 2 & 4 & 8 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 3 & 3 \end{pmatrix}.$$

Ziehen wir nun noch die zweite Zeile dreimal von der dritten ab, so erhalten wir

$$\mathbf{A}_2 = \begin{pmatrix} 1 & 2 & 4 & 8 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Damit sind die ersten beiden Zeilen linear unabhängig, und da die dritte Zeile eine Nullzeile ist, ergibt sich  $rg(\mathbf{A}_2) = rg(\mathbf{A}) = 2$ .

Mit dem Gauß-Algorithmus kann man auch leicht überprüfen, ob Vektoren linear unabhängig sind. Es gilt offensichtlich folgender Zusammenhang.

**Satz 2.29.** *Die Vektoren  $\mathbf{a}_1, \dots, \mathbf{a}_n$  sind linear unabhängig, falls*

$$rg((\mathbf{a}_1, \dots, \mathbf{a}_n)) = n.$$

## 2.5 Lösungsmengen linearer Gleichungssysteme

Wir haben bereits gesehen, dass die Menge  $L = L(\mathbf{A}, \mathbf{0}) = \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{0}\}$  ein Unterraum des  $\mathbb{R}^n$  ist. Für die Dimension dieses Unterraums gilt folgende Dimensionsformel.

**Satz 2.30.** *Sei  $\mathbf{A}$  eine  $m \times n$ -Matrix. Dann gilt*

$$rg(\mathbf{A}) + \dim(L(\mathbf{A}, \mathbf{0})) = n.$$

Die Dimension der Lösungsmenge ist also  $\dim(L(\mathbf{A}, \mathbf{0})) = n - rg(\mathbf{A})$ . Die Lösung ist eindeutig (gleich dem Nullvektor), falls die Dimension des Lösungsraums Null ist, also falls  $rg(\mathbf{A}) = n$ . Wir erhalten somit folgende äquivalente Bedingungen für die eindeutige Lösbarkeit.

**Satz 2.31.** *Für eine  $m \times n$ -Matrix  $\mathbf{A}$  sind folgende Aussagen äquivalent:*

- (i) *Die Lösung des homogenen linearen Gleichungssystems  $\mathbf{A}\mathbf{x} = \mathbf{0}$  ist eindeutig;*
- (ii)  *$rg(\mathbf{A}) = n$ ;*
- (iii) *die Spalten von  $\mathbf{A}$  sind linear unabhängig.*

Im Fall  $m = n$  ist dies nach Satz 2.14 außerdem äquivalent zur Aussage, dass  $\det(\mathbf{A}) \neq 0$  ist.

Erhält man durch die Umformungen beim Gauß-Algorithmus eine obere Dreiecksmatrix der Form

$$\tilde{\mathbf{A}} = \begin{pmatrix} 1 & 0 & \dots & 0 & a_{1,\ell+1} & \dots & a_{1n} \\ 0 & 1 & \dots & 0 & a_{2,\ell+1} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & a_{\ell,\ell+1} & \dots & a_{\ell n} \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

so ist also  $rg(\mathbf{A}) = \ell$  und damit die Dimension des Lösungsraums  $\dim(L(\mathbf{A}, \mathbf{0})) = n - \ell$ . Eine Basis von  $L(\mathbf{A}, \mathbf{0})$  ist gegeben durch

$$\mathbf{x}_1 = \begin{pmatrix} a_{1,\ell+1} \\ a_{2,\ell+1} \\ \vdots \\ a_{\ell,\ell+1} \\ -1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} a_{1,\ell+2} \\ a_{2,\ell+2} \\ \vdots \\ a_{\ell,\ell+2} \\ 0 \\ -1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \mathbf{x}_{n-\ell} = \begin{pmatrix} a_{1,n} \\ a_{2,n} \\ \vdots \\ a_{\ell,n} \\ 0 \\ 0 \\ 0 \\ \vdots \\ -1 \end{pmatrix}.$$

Man erhält also die Basis, indem man im Endtableau rechts unten das Negative einer Einheitsmatrix anfügt.

**Beispiel 2.32.** Sei die obere Dreiecksdarstellung nach Durchführung des Gauß-Algorithmus gegeben durch

$$\tilde{\mathbf{A}} = \begin{pmatrix} 1 & 0 & 0 & 2 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 2 \end{pmatrix}.$$

Dann erhält man also als Basis des Lösungsraums die Vektoren

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

Also ist

$$L(\mathbf{A}, \mathbf{0}) = [\mathbf{x}_1, \mathbf{x}_2] = \left\{ \lambda \begin{pmatrix} 2 \\ 1 \\ 1 \\ -1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \\ -1 \end{pmatrix} : \lambda, \mu \in \mathbb{R} \right\}.$$

Wir betrachten nun die Lösungsmenge  $L(\mathbf{A}, \mathbf{b})$  eines inhomogenen linearen Gleichungssystems  $\mathbf{A}\mathbf{x} = \mathbf{b}$  mit  $\mathbf{b} \neq \mathbf{0}$ . Angenommen, wir kennen eine spezielle Lösung  $\mathbf{x}^*$  mit  $\mathbf{A}\mathbf{x}^* = \mathbf{b}$ , und es sei  $\mathbf{y}$  eine beliebige weitere Lösung. Dann gilt für  $\mathbf{z} = \mathbf{y} - \mathbf{x}^*$  dass

$$\mathbf{A}\mathbf{z} = \mathbf{A}\mathbf{y} - \mathbf{A}\mathbf{x}^* = \mathbf{b} - \mathbf{b} = \mathbf{0}.$$

Also ist  $\mathbf{z} \in L(\mathbf{A}, \mathbf{0})$ . Ist umgekehrt  $\mathbf{x}^*$  eine spezielle Lösung von  $\mathbf{A}\mathbf{x}^* = \mathbf{b}$  und  $\mathbf{z} \in L(\mathbf{A}, \mathbf{0})$  eine beliebige Lösung von  $\mathbf{A}\mathbf{x} = \mathbf{0}$ , dann gilt für  $\mathbf{y} = \mathbf{x}^* + \mathbf{z}$  die Gleichung

$$\mathbf{A}\mathbf{y} = \mathbf{A}\mathbf{x}^* + \mathbf{A}\mathbf{z} = \mathbf{b} + \mathbf{0} = \mathbf{b}.$$

Also ist  $\mathbf{y} \in L(\mathbf{A}, \mathbf{b})$ . Wir haben damit das folgende Resultat gezeigt.

**Satz 2.33.** Sei  $\mathbf{A}$  eine  $m \times n$ -Matrix,  $\mathbf{b} \neq \mathbf{0}$ , und  $\mathbf{x}^* \in L(\mathbf{A}, \mathbf{b})$  eine spezielle Lösung des inhomogenen Gleichungssystems  $\mathbf{Ax} = \mathbf{b}$ . Dann ist die Lösungsmenge des inhomogenen Systems gegeben durch

$$L(\mathbf{A}, \mathbf{b}) = \{\mathbf{x}^* + \mathbf{z} : \mathbf{z} \in L(\mathbf{A}, \mathbf{0})\}.$$

Wir wollen uns nun noch mit der Frage beschäftigen, unter welchen Bedingungen das lineare Gleichungssystem  $\mathbf{Ax} = \mathbf{b}$  überhaupt eine Lösung  $\mathbf{x} \in \mathbb{R}^n$  besitzt, und unter welchen Bedingungen diese Lösung eindeutig ist.

Betrachtet man die Matrix  $\mathbf{A}$  als Aneinanderreihung von Spaltenvektoren  $\mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_n)$ , so kann man die linke Seite der Gleichung schreiben als

$$\mathbf{Ax} = x_1\mathbf{a}_1 + \dots + x_n\mathbf{a}_n.$$

Somit existiert eine Lösung genau dann, wenn reelle Zahlen  $x_1, \dots, x_n$  existieren mit

$$x_1\mathbf{a}_1 + \dots + x_n\mathbf{a}_n = \mathbf{b}.$$

Dies bedeutet, dass sich  $\mathbf{b}$  als Linearkombination von  $\mathbf{a}_1, \dots, \mathbf{a}_n$  schreiben lässt. Das bedeutet aber nichts anderes als dass die lineare Hülle  $[\mathbf{a}_1, \dots, \mathbf{a}_n]$  den Vektor  $\mathbf{b}$  enthält, also dass

$$[\mathbf{a}_1, \dots, \mathbf{a}_n] = [\mathbf{a}_1, \dots, \mathbf{a}_n, \mathbf{b}]$$

Aus Satz 2.26 und Satz 2.23 b) und c) folgt aber, dass dies genau dann gilt, falls

$$rg(\mathbf{A}) = rg(\mathbf{A}_\mathbf{b}),$$

wobei mit  $\mathbf{A}_\mathbf{b}$  die um den Vektor  $\mathbf{b}$  ergänzte Matrix  $\mathbf{A}$  bezeichnet werde:

$$\mathbf{A}_\mathbf{b} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}.$$

Wir erhalten also folgendes Resultat über die Existenz von Lösungen.

**Satz 2.34.** Sei  $\mathbf{A}$  eine  $m \times n$ -Matrix und  $\mathbf{b} \in \mathbb{R}^m$ . Das lineare Gleichungssystem  $\mathbf{Ax} = \mathbf{b}$  besitzt genau dann mindestens eine Lösung, falls  $rg(\mathbf{A}) = rg(\mathbf{A}_\mathbf{b})$ .

Die Anzahl der Lösungen von  $\mathbf{Ax} = \mathbf{b}$  ist nach Satz 2.33 gleich der Anzahl der Lösungen von  $\mathbf{Ax} = \mathbf{0}$ . Letztere ist nach Satz 2.30 eindeutig, falls  $rg(\mathbf{A}) = n$ . Also gilt folgendes Resultat.

**Satz 2.35.** Sei  $\mathbf{A}$  eine  $m \times n$ -Matrix und  $\mathbf{b} \in \mathbb{R}^m$ . Das lineare Gleichungssystem  $\mathbf{Ax} = \mathbf{b}$  besitzt genau dann eine eindeutige Lösung, falls  $rg(\mathbf{A}) = rg(\mathbf{A}_\mathbf{b}) = n$ .

## 2.6 Berechnung von Eigenwerten und Eigenvektoren

Wir kommen nun zurück zum Problem der Berechnung von Eigenwerten und Eigenvektoren. Dazu sei daran erinnert, dass ein **Eigenwert** einer Matrix  $\mathbf{A}$  eine reelle Zahl  $\lambda$  ist mit der Eigenschaft, dass es einen Vektor  $\mathbf{x} \neq \mathbf{0}$  gibt mit

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

Ein solcher zugehöriger Vektor  $\mathbf{x}$  heißt dann **Eigenvektor**.

Durch Umformulierung der Definition eines Eigenwertes erhält man die Aussage, dass eine Zahl  $\lambda$  ein Eigenwert einer Matrix  $\mathbf{A}$  ist, wenn die Gleichung

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0} \tag{2.6}$$

eine nichttriviale Lösung  $\mathbf{x} \neq \mathbf{0}$  hat, wobei  $\mathbf{I}$  die Identitätsmatrix darstellt.

Aus Satz 2.14 folgt somit, dass eine reelle Zahl  $\lambda$  genau dann ein Eigenwert der Matrix  $\mathbf{A}$  ist, falls  $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ . Wir fassen das zentrale Ergebnis im folgenden Satz zusammen.

**Satz 2.36.** *Sei  $\mathbf{A}$  eine  $n \times n$ -Matrix. Dann ist  $\det(\mathbf{A} - \lambda\mathbf{I})$  ein Polynom  $n$ -ten Grades in  $\lambda \in \mathbb{R}$ . Es heißt **charakteristisches Polynom** von  $\mathbf{A}$ . Die Lösungen  $\lambda_i$  der so genannten charakteristischen Gleichung*

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0$$

*sind die Eigenwerte von  $\mathbf{A}$ . Die Vektoren aus der Lösungsmenge  $L(\mathbf{A} - \lambda_i\mathbf{I}, \mathbf{0})$  des zugehörigen homogenen Gleichungssystems*

$$(\mathbf{A} - \lambda_i\mathbf{I})\mathbf{x} = \mathbf{0}$$

*mit  $\mathbf{x} \neq \mathbf{0}$  sind die zugehörigen Eigenvektoren.*

Den Unterraum  $L(\mathbf{A} - \lambda_i\mathbf{I}, \mathbf{0})$ , bestehend aus den Eigenvektoren zum Eigenwert  $\lambda_i$  und dem Nullvektor, nennt man auch den **Eigenraum** von  $\lambda_i$ .

**Beispiel 2.37.** Es seien die Eigenwerte und Eigenvektoren der Matrix

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

gesucht. Dann ist

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} 1 - \lambda & -2 \\ 1 & 4 - \lambda \end{vmatrix} = (1 - \lambda) \cdot (4 - \lambda) - 1 \cdot (-2) = \lambda^2 - 5\lambda + 6.$$

Die charakteristische Gleichung ist also  $\lambda^2 - 5\lambda + 6 = 0$ . Sie hat die Lösungen

$$\lambda_{1,2} = \frac{5}{2} \pm \sqrt{\frac{25}{4} - 6} = \frac{5}{2} \pm \sqrt{\frac{1}{4}} = \frac{5}{2} \pm \frac{1}{2},$$

also gibt es die beiden Eigenwerte  $\lambda_1 = 3$  und  $\lambda_2 = 2$ .

Um die Eigenvektoren zum Eigenwert  $\lambda_1 = 3$  zu bestimmen, suchen wir die Lösungen der Gleichung  $(\mathbf{A} - 3\mathbf{I})\mathbf{x} = \mathbf{0}$ . Dies führt auf das Gleichungssystem

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

Offensichtlich sind die beiden Gleichungen identisch. Setzt man  $x_2 = \alpha$  mit  $\alpha \in \mathbb{R}$  beliebig, so erhält man  $x_1 = -\alpha$ . Der zugehörige Lösungsraum und damit der Eigenraum zum Eigenwert  $\lambda_1 = 3$  ist also der 1-dimensionale Unterraum

$$L(\mathbf{A} - 3\mathbf{I}, \mathbf{0}) = \left\{ \alpha \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} : \alpha \in \mathbb{R} \right\}.$$

Die Menge der Eigenvektoren ist also

$$L(\mathbf{A} - 3\mathbf{I}, \mathbf{0})/\{\mathbf{0}\} = \left\{ \alpha \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} : \alpha \in \mathbb{R}/\{0\} \right\}.$$

Analog erhält man die Eigenvektoren zum Eigenwert  $\lambda_2 = 2$  als Lösungen der Gleichung  $(\mathbf{A} - 2\mathbf{I})\mathbf{x} = \mathbf{0}$ . Dies führt auf das Gleichungssystem

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

Offensichtlich sind auch diese beiden Gleichungen identisch. Setzt man  $x_2 = \alpha$  mit  $\alpha \in \mathbb{R}$  beliebig, so erhält man  $x_1 = -2\alpha$ . Der zugehörige Eigenraum zum Eigenwert  $\lambda_2 = 2$  ist also der 1-dimensionale Unterraum

$$L(\mathbf{A} - 2\mathbf{I}, \mathbf{0}) = \left\{ \alpha \cdot \begin{pmatrix} -2 \\ 1 \end{pmatrix} : \alpha \in \mathbb{R} \right\}.$$

und die Menge der Eigenwerte ist

$$L(\mathbf{A} - 2\mathbf{I}, \mathbf{0})/\{\mathbf{0}\} = \left\{ \alpha \cdot \begin{pmatrix} -2 \\ 1 \end{pmatrix} : \alpha \in \mathbb{R}/\{0\} \right\}.$$

Die beiden die Eigenräume aufspannenden Vektoren  $(-1, 1)^T$  und  $(-2, 1)^T$  sind linear unabhängig und bilden deshalb zusammen eine Basis des  $\mathbb{R}^2$ . Es lässt sich also jeder Vektor im  $\mathbb{R}^2$  als Linearkombination dieser beiden Eigenvektoren darstellen.

Man interessiert sich natürlich für die Anzahl der Eigenwerte und die Dimensionen der Eigenräume einer Matrix. Hierüber gibt folgender Satz über Polynome Auskunft.

**Satz 2.38.** *Jedes Polynom  $n$ -ten Grades hat höchstens  $n$  verschiedene Nullstellen. Ist  $n$  ungerade, so gibt es mindestens eine Nullstelle.*

Eine Rolle spielt auch noch die so genannte *Vielfachheit* einer Nullstelle. Ist  $\lambda$  eine Nullstelle des Polynoms  $p_n(x)$ , so erhält man durch Polynom-Division durch  $(x - \lambda)$  ein Polynom  $p_{n-1}(x)$  mit

$$p_n(x) = (x - \lambda)p_{n-1}(x).$$

Ist  $\lambda$  nun auch noch eine Nullstelle von  $p_{n-1}(x)$ , so spricht man von einer mehrfachen Nullstelle, da dann

$$p_n(x) = (x - \lambda)^2 p_{n-2}(x).$$

Genauer spricht man von einer  $k$ -fachen Nullstelle, falls

$$p_n(x) = (x - \lambda)^k p_{n-k}(x)$$

mit  $p_{n-k}(\lambda) \neq 0$ .

Für Eigenwerte und Eigenvektoren erhält man damit folgendes Resultat.

**Satz 2.39.** *Sei  $\mathbf{A}$  eine  $n \times n$ -Matrix. Dann gelten folgende Aussagen.*

- a)  $\mathbf{A}$  besitzt höchstens  $n$  verschiedene Eigenwerte.
- b) Ist der Eigenwert  $\lambda$  eine  $k$ -fache Nullstelle des charakteristischen Polynoms, so ist die Dimension des zugehörigen Eigenraumes höchstens  $k$ .
- c) Sind  $\lambda_1, \dots, \lambda_m$  die verschiedenen Eigenwerte von  $\mathbf{A}$  und  $\mathbf{v}_1, \dots, \mathbf{v}_m$  zugehörige Eigenvektoren, so sind  $\mathbf{v}_1, \dots, \mathbf{v}_m$  linear unabhängig.

Die Dimension des Eigenraumes muss nicht immer gleich der Vielfachheit des zugehörigen Eigenwertes sein. Wir illustrieren dies an einem einfachen Beispiel, und es gibt Matrizen, die gar keine reellen Eigenwerte haben.

**Beispiel 2.40.** a) Die Matrix

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

hat das charakteristische Polynom  $\lambda^2 + 1$ , welches keine Nullstelle hat. Damit besitzt die Matrix keine Eigenwerte.

b) Die Matrix

$$\mathbf{A} = \begin{pmatrix} 4 & 1 \\ 0 & 4 \end{pmatrix}$$

hat das charakteristische Polynom  $(4 - \lambda)^2$ , also ist  $\lambda = 4$  zweifache Nullstelle. Für  $\lambda = 4$  ist

$$\mathbf{A} - \lambda \mathbf{I} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Damit  $(\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = \mathbf{0}$  gilt, kann also  $x_1$  beliebig sein, während  $x_2 = 0$  sein muss. Also ist der Eigenraum

$$L(\mathbf{A} - 4\mathbf{I}, \mathbf{0}) = \left\{ \alpha \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} : \alpha \in \mathbb{R} \right\},$$

welcher offensichtlich nur Dimension 1 hat.



Wir betrachten nun noch spezielle Typen von Matrizen. Aus Satz 2.9 folgt sofort das folgende Resultat.

**Satz 2.41.** *Sei  $\mathbf{A}$  eine obere Dreiecksmatrix (also  $a_{ij} = 0$  für  $i > j$ ). Dann sind die Eigenwerte von  $\mathbf{A}$  gerade die Diagonalelemente von  $\mathbf{A}$ .*

Wir untersuchen nun noch Definitheitseigenschaften von symmetrischen Matrizen. Eine  $n \times n$ -Matrix  $\mathbf{A}$  heißt **symmetrische Matrix**, falls  $\mathbf{A} = \mathbf{A}^T$ , d.h. falls

$$a_{ij} = a_{ji} \quad \text{für alle } i, j = 1, \dots, n.$$

Wichtige Beispiele von symmetrischen Matrizen sind Hesse-Matrizen von zweimal differenzierbaren Funktionen  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ . Man benützt folgende Begriffe von Definitheit.

**Definition 2.42.** Sei  $\mathbf{A}$  eine symmetrische  $n \times n$ -Matrix. Dann heißt  $\mathbf{A}$

- **positiv definit**, falls  $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$  für alle  $\mathbf{x} \neq \mathbf{0}$ ;
- **positiv semidefinit**, falls  $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0$  für alle  $\mathbf{x} \neq \mathbf{0}$ ;
- **negativ definit**, falls  $\mathbf{x}^T \mathbf{A} \mathbf{x} < 0$  für alle  $\mathbf{x} \neq \mathbf{0}$ ;
- **negativ semidefinit**, falls  $\mathbf{x}^T \mathbf{A} \mathbf{x} \leq 0$  für alle  $\mathbf{x} \neq \mathbf{0}$ ;
- **indefinit**, falls  $\mathbf{x}^T \mathbf{A} \mathbf{x}$  sowohl positive als auch negative Werte annimmt.

Für die Bestimmung von Extremwerten mehrdimensionaler Funktionen gilt folgender Satz.

**Satz 2.43.** *Sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  zweimal stetig differenzierbar. Ist der Gradient  $\nabla f(\mathbf{x}) = \mathbf{0}$ , so gilt:*

- *ist die Hesse-Matrix  $H_f(\mathbf{x})$  an der Stelle  $\mathbf{x}$  **positiv definit**, so hat  $f$  an der Stelle  $\mathbf{x}$  ein lokales Minimum;*
- *ist die Hesse-Matrix  $H_f(\mathbf{x})$  an der Stelle  $\mathbf{x}$  **negativ definit**, so hat  $f$  an der Stelle  $\mathbf{x}$  ein lokales Maximum;*
- *ist die Hesse-Matrix  $H_f(\mathbf{x})$  an der Stelle  $\mathbf{x}$  **indefinit**, so hat  $f$  an der Stelle  $\mathbf{x}$  kein Extremum.*

Der wesentliche Grund für diese Tatsache besteht darin, dass nach der Taylor-Entwicklung zweiter Ordnung um einen Punkt  $\mathbf{x}$  für eine solche Funktion gilt:

$$f(\mathbf{y}) \approx f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{y} - \mathbf{x}) + \frac{1}{2} (\mathbf{y} - \mathbf{x})^T H_f(\mathbf{x}) (\mathbf{y} - \mathbf{x}).$$

Ist also  $\nabla f(\mathbf{x}) = \mathbf{0}$ , so bestimmt die Definitheit von  $H_f(\mathbf{x})$ , ob die Funktionswerte in einer Umgebung von  $\mathbf{x}$  grösser oder kleiner als im Punkt  $\mathbf{x}$  selbst sind.

Es ist schwierig, Definitheit von Matrizen über die Definition zu überprüfen. Deshalb helfen hier Aussagen, welche diese mit Hilfe von Eigenwerten charakterisieren.

**Satz 2.44.** Sei  $\mathbf{A}$  eine symmetrische  $n \times n$ -Matrix. Dann gilt:

a)  $\mathbf{A}$  besitzt genau  $n$  reelle Eigenwerte, wenn man mehrfache Eigenwerte mit ihrer Vielfachheit zählt.

b)  $\mathbf{A}$  ist genau dann

- **positiv definit**, wenn alle Eigenwerte von  $\mathbf{A}$  größer als Null sind;
- **positiv semidefinit**, wenn alle Eigenwerte von  $\mathbf{A}$  größer oder gleich Null sind;
- **negativ definit**, wenn alle Eigenwerte von  $\mathbf{A}$  kleiner als Null sind;
- **negativ semidefinit**, wenn alle Eigenwerte von  $\mathbf{A}$  kleiner oder gleich Null sind;
- **indefinit**, wenn  $\mathbf{A}$  sowohl positive als auch negative Eigenwerte hat.

Die Aussage von Teil b) ist intuitiv klar und kann man sich leicht merken, wenn man bedenkt dass für  $\mathbf{x} \neq \mathbf{0}$

$$\mathbf{x}^T \mathbf{x} = \sum_{i=1}^n x_i^2 > 0$$

und deshalb im Falle  $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$  wegen

$$\mathbf{x}^T \mathbf{A}\mathbf{x} = \mathbf{x}^T \lambda\mathbf{x} = \lambda \cdot \mathbf{x}^T \mathbf{x}$$

das Vorzeichen von  $\mathbf{x}^T \mathbf{A}\mathbf{x}$  durch das Vorzeichen von  $\lambda$  bestimmt wird.

## 3 Differential- und Differenzgleichungen

### 3.1 Differentialgleichungen

Wir beginnen mit einem Beispiel einer Differentialgleichung, welche sich aus einer volkswirtschaftlichen Fragestellung ergibt.

**Beispiel 3.1.** Gesucht ist eine Nachfragefunktion  $q = f(p)$ , abhängig vom Preis  $p$ , so dass die Preiselastizität

$$\text{El}_p f(p) = f'(p) \cdot \frac{p}{f(p)}$$

konstant sei mit einem Wert  $-\lambda$ , wobei wir davon ausgehen, dass  $\lambda > 0$ , dass also ein steigender Preis eine verminderte Nachfrage zur Folge hat. Gesucht ist also eine Nachfragefunktion mit

$$f'(p) \cdot \frac{p}{f(p)} = -\lambda$$

bzw. aufgelöst nach  $f'(p)$

$$f'(p) = -\lambda \cdot \frac{f(p)}{p}.$$

Es ist also eine Gleichung zu lösen, bei der die gesuchte Funktion sowie ihre erste Ableitung vorkommen.

Allgemein spricht man von **Differentialgleichungen** wenn in einer Gleichung eine Funktion gesucht wird, und in der Gleichung eine unabhängige Variable, eine Funktion dieser Variablen und/oder Ableitungen von ihr vorkommen. Wenn die höchste vorkommende Ableitung die Ordnung  $m$  hat, so spricht man von einer **Differentialgleichung  $m$ -ter Ordnung**. Formal ist das also eine Gleichung

$$F(x, y(x), y'(x), \dots, y^{(m)}(x)) = 0,$$

wobei  $F$  eine Funktion mehrerer Variablen ist. Eine Funktion  $f$  heißt Lösung einer Differentialgleichung, wenn

$$F(x, f(x), f'(x), \dots, f^{(m)}(x)) = 0$$

ist für alle  $x$  im Definitionsbereich.

Wir beschäftigen uns hier nur mit Differentialgleichungen erster Ordnung, in denen nur  $x$ ,  $y(x)$  und  $y'(x)$  vorkommen. Oft kann man eine solche Differentialgleichung nach  $y'$  auflösen und erhält dann die Form

$$y'(x) = g(x, y(x))$$

für eine Funktion  $g$ .

Differentialgleichungen sind oft nicht eindeutig lösbar. So hat die einfache Differentialgleichung  $y' = y$  offenbar die Lösung  $f(x) = e^x$ . Jede andere Funktion der Form  $f(x) = ce^x$  mit  $c \in \mathbb{R}$  ist aber auch eine Lösung. Deshalb benötigt man oft für die Eindeutigkeit noch eine so genannte **Anfangsbedingung**  $y_0 = f(x_0)$ , um die Lösung eindeutig zu machen. Fordert man im obigen Fall also z.B.  $f(0) = 3$ , so erhält man die eindeutige Lösung  $f(x) = 3e^x$ .

Der einfachste Fall einer Differentialgleichung ist gegeben, wenn die rechte Seite überhaupt nicht von  $y$  abhängt, wenn also die Differentialgleichung die Form  $y' = g(x)$  hat. Diese ist natürlich durch einfache Integration zu lösen.

**Satz 3.2.** *Gegeben sei eine Differentialgleichung der Form  $y' = g(x)$ .*

*Ist  $G$  eine Stammfunktion von  $g$ , so erhält man alle Lösungen dieser Differentialgleichung gemäß*

$$y = \int g(x)dx = G(x) + c$$

*mit  $c \in \mathbb{R}$ . Die Konstante  $c$  ergibt sich gegebenenfalls aus einer Anfangsbedingung.*

Wir führen nun lineare Differentialgleichungen ein.

**Definition 3.3.** Eine Differentialgleichung der Form

$$y' + a(x)y = b(x)$$

bezeichnet man als **lineare Differentialgleichung** erster Ordnung. Hier sind  $a(x)$  und  $b(x)$  gegebene Funktionen von  $x$ .

Ist  $b(x) = 0$  für alle  $x$ , so spricht man von einer **homogenen** linearen Differentialgleichung, sonst von einer **inhomogenen** linearen Differentialgleichung.

Die Lösung der homogenen linearen Differentialgleichung ist relativ einfach.

**Satz 3.4.** Eine homogene lineare Differentialgleichung  $y' + a(x)y = 0$  hat die Lösung

$$y = \gamma \exp\left(-\int a(x)dx\right),$$

wobei  $\gamma \in \mathbb{R}$  eine reelle Konstante ist.

Die eindeutige Lösung zur Anfangsbedingung  $y_0 = f(x_0)$  ist gegeben durch

$$y = y_0 \exp\left(-\int_{x_0}^x a(t)dt\right).$$

**Beispiel 3.5.** Das eingangs beschriebene Problem, eine Funktion zu finden mit konstanter Preiselastizität

$$q' = -\lambda \frac{q}{p} \Leftrightarrow q' + \frac{\lambda}{p}q = 0$$

stellt eine homogene lineare Differentialgleichung dar, wobei hier  $q$  die Rolle des  $y$  und  $p$  die Rolle des  $x$  übernimmt. Als allgemeine Lösung erhält man somit

$$q = \gamma \exp\left(-\int \frac{\lambda}{p}dp\right) = \gamma \exp(-\lambda \ln(p)) = \gamma p^{-\lambda} = \frac{\gamma}{p^\lambda}.$$

Sucht man eine Lösung zur Anfangsbedingung  $q(p_0) = q_0$ , so ergibt sich

$$\begin{aligned} q &= q_0 \exp\left(-\int_{p_0}^p \frac{\lambda}{t}dt\right) = q_0 \exp(-\lambda \ln(p) + \lambda \ln(p_0)) \\ &= q_0 p^{-\lambda} p_0^\lambda = q_0 \left(\frac{p_0}{p}\right)^\lambda. \end{aligned}$$

Die Lösung der allgemeinen inhomogenen linearen Differentialgleichung  $y' + a(x)y = b(x)$  ist leider sehr schwierig. Wir betrachten deshalb hier zunächst nur den einfachen Spezialfall von konstanten Koeffizienten, also der Differentialgleichung  $y' + ay = b$  mit reellen Zahlen  $a, b \neq 0$ . Wählt man für  $y$  die konstante Funktion  $y(x) = b/a$ , so sieht man durch Einsetzen leicht, dass dies eine spezielle Lösung der Differentialgleichung ist. Um die allgemeine Lösung zu erhalten, kann man dann das Wissen über die Lösung der zugehörigen homogenen Differentialgleichung ausnützen. Sind nämlich  $y_1$  und  $y_2$  zwei spezielle Lösungen der inhomogenen Gleichung, so gilt also

$$\begin{aligned} y_1' + ay_1 &= b \\ y_2' + ay_2 &= b. \end{aligned}$$

Zieht man diese beiden Gleichungen voneinander ab, so erhält man

$$(y_1' - y_2') + a(y_1 - y_2) = b - b = 0, \text{ also } (y_1 - y_2)' + a(y_1 - y_2) = 0.$$

Somit ist die Differenz  $y_1 - y_2$  eine Lösung der homogenen Differentialgleichung, welche nach Satz 3.4 aber durch  $y = \gamma e^{-ax}$  gegeben ist mit  $\gamma \in \mathbb{R}$ . Deshalb gilt folgendes Resultat.

**Satz 3.6.** Eine inhomogene lineare Differentialgleichung mit konstanten Koeffizienten der Form  $y' + ay = b$  hat für  $a \neq 0$  die Lösung

$$y = \frac{b}{a} + \gamma e^{-ax},$$

wobei  $\gamma \in \mathbb{R}$  eine reelle Konstante ist.

Die eindeutige Lösung zur Anfangsbedingung  $y_0 = f(x_0)$  ist gegeben durch

$$y = \frac{b}{a} + \left( y_0 - \frac{b}{a} \right) e^{-a(x-x_0)}.$$

**Beispiel 3.7.** Wir betrachten das Modell der Preisfestsetzung nach Evans. Hierzu bezeichne  $d(t)$  die Nachfrage nach einem Gut zur Zeit  $t$  sowie  $s(t)$  das Angebot an diesem Gut. Ferner sei  $p(t)$  der Preis des Gutes zur Zeit  $t$ . Es wird angenommen, dass Angebot und Nachfrage linear vom Preis abhängen, d.h. es sei

$$\begin{aligned} d(t) &= a - bp(t) \\ s(t) &= c + dp(t). \end{aligned}$$

Ferner wird angenommen, dass die Änderung des Preises proportional zum Nachfrageüberschuss ist, d.h.

$$p'(t) = \lambda(d(t) - s(t)).$$

Dabei wird angenommen, dass  $b, d, \lambda$  positive Konstanten sind. Einsetzen der ersten beiden Gleichungen in die dritte liefert

$$p'(t) = \lambda(a - bp(t) - (c + dp(t))) = \lambda(a - c) - \lambda(b + d)p(t).$$

Wir erhalten also eine lineare inhomogene Differentialgleichung mit konstanten Koeffizienten

$$p'(t) + \lambda(b + d)p(t) = \lambda(a - c).$$

Aus Satz 3.6 erhalten wir also die allgemeine Lösung

$$p(t) = \frac{\lambda(a - c)}{\lambda(b + d)} + \gamma e^{-\lambda(b+d)t}.$$

Man nennt den Preis  $\bar{p}$ , bei welchem Angebot und Nachfrage übereinstimmen, den **Gleichgewichtspreis**. Da  $d(t) = s(t)$  gilt, wenn

$$a - bp(t) = c + dp(t)$$

ist, erhält man also durch Lösung dieser Gleichung für den Gleichgewichtspreis  $\bar{p} = (a - c)/(b + d)$ . Unter der gemachten Annahme  $b, d, \lambda > 0$  gilt

$$\lim_{t \rightarrow \infty} p(t) = \lim_{t \rightarrow \infty} \left( \frac{a - c}{b + d} + \gamma e^{-\lambda(b+d)t} \right) = \frac{a - c}{b + d} = \bar{p}.$$

Mit zunehmender Zeit konvergiert also in diesem Modell der Preis gegen den Gleichgewichtspreis.

Eine weitere wichtige Klasse von Differentialgleichungen sind **Differentialgleichungen mit getrennten Veränderlichen**. Darunter versteht man Differentialgleichungen der Form

$$y' = g(x)h(y)$$

bei welchen man die rechte Seite als ein Produkt von zwei Funktionen schreiben kann, von denen die eine Funktion nur von  $x$  und die andere nur von  $y$  abhängt. Diese haben folgende allgemeine Lösung.

**Satz 3.8.** Die allgemeine Lösung der Differentialgleichung mit getrennten Veränderlichen  $y' = g(x)h(y)$  erhält man aus der Gleichung

$$\int \frac{1}{h(y)} dy = \int g(x) dx.$$

Die eindeutige Lösung zur Anfangsbedingung  $y_0 = f(x_0)$  ist gegeben durch

$$\int_{y_0}^y \frac{1}{h(s)} ds = \int_{x_0}^x g(s) ds.$$

Man kann sich die Lösungsmethode gut merken, indem man  $dy/dx$  schreibt für  $y'$  und dann umformt, so dass die Variablen getrennt sind, d.h. auf verschiedenen Seiten der Gleichung stehen. Dann fügt man noch auf beiden Seiten ein Integralzeichen hinzu, um die Lösung zu erhalten.

$$y' = \frac{dy}{dx} = g(x)h(y) \quad \Leftrightarrow \quad \frac{dy}{h(y)} = g(x)dx \quad \Leftrightarrow \quad \int \frac{dy}{h(y)} = \int g(x)dx.$$

**Beispiel 3.9.** Wir betrachten einfache Wachstumsmodelle für ökonomische Größen abhängig von der Zeit  $t$ . Das einfachste denkbare Wachstumsmodell nimmt an, dass die absolute Änderung der Größe  $y$  proportional zu  $y$  ist. Dies führt zur Differentialgleichung  $y' = \lambda y$  mit der allgemeinen Lösung  $y(t) = \gamma e^{\lambda t}$  und der speziellen Lösung  $y(t) = y_0 e^{\lambda t}$ , wenn  $y_0$  der Startwert ist. Dies nennt man naheliegenderweise **exponentielles Wachstum**. Diese Annahme ist aber oft unrealistisch. Etwas realistischer ist deshalb ein Modell mit **logistischem Wachstum**. Hier nimmt man eine Sättigungsgrenze  $G$  an, und nimmt an, dass das relative Wachstum proportional ist zum Abstand von  $y$  zur Sättigungsgrenze  $G$ . Dies führt auf die Differentialgleichung

$$\frac{y'}{y} = \lambda(G - y) \quad \text{bzw.} \quad y' = \lambda y(G - y)$$

Dies ist eine Differentialgleichung mit getrennten Veränderlichen (mit  $g(t)$  konstant). Für die Lösung gilt also

$$\frac{dy}{y(G - y)} = \lambda dt$$

Unter der Anfangsbedingung  $y(0) = y_0$  führt dies auf die Gleichung

$$\int_{y_0}^y \frac{ds}{s(G - s)} = \int_0^t \lambda ds$$

Zur Lösung des linken Integrals schreiben wir den Integranden als

$$\frac{1}{s(G-s)} = \frac{1}{sG} + \frac{1}{G(G-s)}.$$

Damit erhält man

$$\begin{aligned} \int_{y_0}^y \frac{ds}{s(G-s)} &= \int_{y_0}^y \frac{1}{sG} ds + \int_{y_0}^y \frac{1}{G(G-s)} ds \\ &= \frac{1}{G} (\ln(y) - \ln(y_0)) + \frac{1}{G} (-\ln(G-y) + \ln(G-y_0)) \\ &= \frac{1}{G} \ln \left( \frac{y}{y_0} \cdot \frac{G-y_0}{G-y} \right) \end{aligned}$$

Man erhält also die Lösung der Differentialgleichung durch Auflösung der Gleichung

$$\frac{1}{G} \ln \left( \frac{y}{y_0} \cdot \frac{G-y_0}{G-y} \right) = \lambda t$$

nach  $y$ . Anwendung der Exponentialfunktion und Umformen liefert

$$\frac{y}{G-y} = \frac{y_0}{G-y_0} e^{\lambda G t}$$

und damit durch Auflösen nach  $y$

$$y = G \frac{\frac{y_0}{G-y_0} e^{\lambda G t}}{1 + \frac{y_0}{G-y_0} e^{\lambda G t}} = G \frac{1}{1 + \left(\frac{G}{y_0} - 1\right) e^{\lambda G t}}.$$

Diese Funktion nennt man **logistische Funktion**.

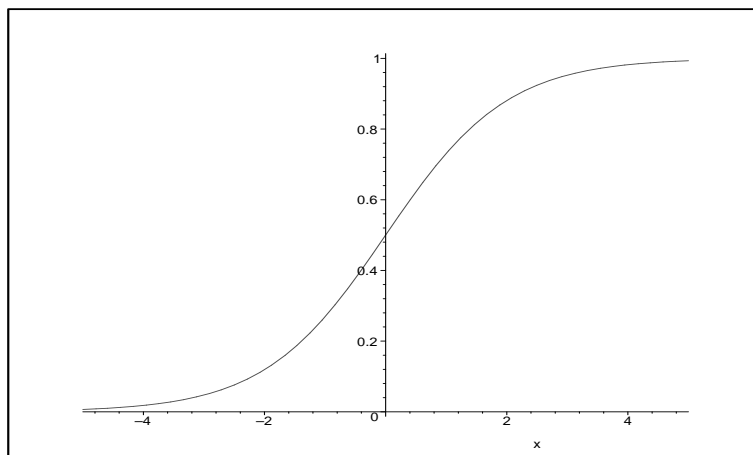


Abbildung 1: logistische Funktion

## 3.2 Differenzgleichungen

Im vorigen Abschnitt haben wir Modelle wie das Preisfestsetzungsmodell von Evans betrachtet, bei welchen der Zeitparameter ein Intervall reeller Zahlen durchlief. Dies führte auf Differentialgleichungen. Nur betrachten wir sogenannten *Periodenmodelle* bzw. *Modelle in diskreter Zeit*. Das bedeutet, dass wir die relevanten ökonomischen Größen nur zu bestimmten Zeitpunkten oder Zeitperioden betrachten. Wir nehmen deshalb an, dass die Zeitvariable  $t$  nur einen zusammenhängenden Abschnitt der ganzen Zahlen durchläuft.

**Beispiel 3.10.** Wir betrachten das sogenannte **Spinnweb-Modell** für Preisfestsetzung. Ähnlich wie in Beispiel 3.7 bezeichnen wir mit

$d_t$  die Nachfrage nach einem Gut in Periode  $t$ ,  
 $s_t$  das Angebot an dem Gut in Periode  $t$ ,  
 $p_t$  den Preis des Gutes in Periode  $t$ .

Es wird wieder angenommen, dass Angebot und Nachfrage linear vom Preis abhängen. Die Nachfrage in einer Periode hängt linear vom Preis des Gutes in der Periode ab,

$$d_t = a - bp_t.$$

Das Angebot reagiert um eine Periode verzögert auf Preisänderungen

$$s_t = c + dp_{t-1}.$$

In jeder Periode bestimmt sich der Preis so, dass Angebot und Nachfrage sich ausgleichen, d.h.  $d_t = s_t$ . Einsetzen in die ersten beiden Gleichungen liefert

$$a - bp_t = c + dp_{t-1}$$

bzw.

$$p_t + \frac{d}{b}p_{t-1} = \frac{a-c}{b}.$$

Dies ist ein typisches Beispiel einer so genannten **Differenzgleichung**.

**Definition 3.11.** Eine **lineare Differenzgleichung  $m$ -ter Ordnung mit konstanten Koeffizienten** ist eine Gleichung der Form

$$y_t + a_1y_{t-1} + \dots + a_my_{t-m} = b_t.$$

Hierbei ist  $y_t = f(t)$ ,  $t \in D \subset \mathbb{Z}$ , eine unbekannte Funktion. Die Koeffizienten  $a_1, \dots, a_m$  sind reelle Zahlen und  $b : D \rightarrow \mathbb{R}$  ist eine gegebene Funktion. Ist  $b_t = 0$  für alle  $t$ , so heißt die Gleichung **homogen**, sonst **inhomogen**.

Es ist klar, dass die Lösung einer Differenzgleichung  $m$ -ter Ordnung eindeutig wird, wenn man  $m$  Anfangsbedingungen vorgibt.



**Satz 3.12.** *Werden bei einer linearen Differenzgleichung  $m$ -ter Ordnung  $m$  aufeinanderfolgende Werte (z.B.  $y_0, y_1, \dots, y_{m-1}$ ) vorgegeben, so existiert genau eine Lösung  $y_t$ , die diese vorgegebenen Werte annimmt.*

Eine homogene lineare Differenzgleichung erster Ordnung hat die allgemeine Form  $y_t + ay_{t-1} = 0$ , d.h.  $y_t = (-a)y_{t-1}$ . Ist also  $y_0 = \gamma$  vorgegeben, so erhält man  $y_1 = (-a)\gamma$ ,  $y_2 = (-a)y_1 = (-a)^2\gamma$ ,  $y_3 = (-a)^3\gamma$ , ... Es gilt also folgendes Resultat.

**Satz 3.13.** *Die homogene lineare Differenzgleichung erster Ordnung  $y_t + ay_{t-1} = 0$  hat die allgemeine Lösung*

$$y_t = \gamma(-a)^t,$$

wobei  $\gamma$  eine reelle Konstante ist. Die eindeutige Lösung bei gegebenem Startwert  $y_0$  ist

$$y_t = y_0(-a)^t.$$

Im Fall einer inhomogenen linearen Differenzgleichung erster Ordnung kann man bei konstanter rechter Seite ähnlich wie beim Fall der Differentialgleichung vorgehen um die allgemeine Lösung zu finden. Die Differenzgleichung hat die Gestalt

$$y_t + ay_{t-1} = b.$$

Man sieht leicht, dass es eine konstante Lösung gibt, falls  $-a \neq 1$ :

$$y_t = \frac{b}{1+a} \quad \text{für alle } t.$$

Dies ist die sogenannte **Gleichgewichtslösung**.

Die allgemeine Lösung ergibt sich nun wieder als Summe dieser speziellen Lösung und einer Lösung der zugehörigen homogenen Gleichung. Sind nämlich  $y_t$  und  $z_t$  zwei Lösungen der inhomogenen Gleichung, also

$$\begin{aligned} y_t + ay_{t-1} &= b, \\ z_t + az_{t-1} &= b, \end{aligned}$$

so ist die Differenz dieser beiden Lösungen eine Lösung der zugehörigen homogenen Gleichung.

$$(y_t - z_t) + a(y_{t-1} - z_{t-1}) = b - b = 0.$$

Deshalb erhalten wir folgendes Resultat.

**Satz 3.14.** *Die inhomogene lineare Differenzgleichung erster Ordnung  $y_t + ay_{t-1} = b$  hat die allgemeine Lösung*

$$y_t = \begin{cases} \frac{b}{1+a} + \gamma(-a)^t, & \text{falls } -a \neq 1, \\ bt + \gamma, & \text{falls } -a = 1, \end{cases}$$

wobei  $\gamma$  eine reelle Konstante ist. Die eindeutige Lösung bei gegebenem Startwert  $y_0$  ist

$$y_t = \begin{cases} \frac{b}{1+a} + \left(y_0 - \frac{b}{1+a}\right)(-a)^t, & \text{falls } -a \neq 1, \\ bt + y_0, & \text{falls } -a = 1. \end{cases}$$

**Beispiel 3.15.** Wir kommen zurück zum Beispiel des Spinnweb-Modells aus Beispiel 3.10. Nach Satz 3.14 hat die dort auftretende Differenzgleichung

$$p_t + \frac{d}{b}p_{t-1} = \frac{a-c}{b}$$

die allgemeine Lösung

$$p_t = \frac{a-c}{b+d} + \gamma \left(-\frac{d}{b}\right)^t = \bar{p} + \gamma \left(-\frac{d}{b}\right)^t.$$

Der Ausdruck  $\gamma(-d/b)^t$  beschreibt also die Abweichung des Preises in Periode  $t$  vom Gleichgewichtspreis. Da  $b, d > 0$ , ist  $-d/b < 0$  und somit liegt der Preis abwechselnd über und unter dem Gleichgewichtspreis. Wir schauen uns das langfristige Verhalten an.

Fall 1: Im Fall  $d < b$  ist  $|d/b| < 1$  und es gilt deshalb  $\lim_{t \rightarrow \infty} (-d/b)^t = 0$ . Also konvergiert der Preis gegen den Gleichgewichtspreis.

Fall 2:  $d = b$ . Dann ist  $p_t = \bar{p} + \gamma(-1)^t$ . Es findet also eine konstante Oszillation um den Gleichgewichtspreis  $\bar{p}$  statt.

Fall 3: Für  $d > b$  erhält man  $|d/b| > 1$  und dies bedeutet, dass sich der Preis in jeder Periode weiter vom Gleichgewichtspreis entfernt. Es liegt eine explodierende Oszillation um den Gleichgewichtspreis vor.

Wir betrachten nun noch die allgemeine Lösung einer linearen Differenzgleichung  $m$ -ter Ordnung. Wir haben gesehen, dass bei der Gleichung erster Ordnung die Lösung die Form  $y_t = \gamma(-a)^t$  hat. Wir verwenden deshalb für die Lösung der homogenen linearen Differenzgleichung  $m$ -ter Ordnung ebenfalls den Ansatz  $y_t = \gamma z^t$ . Setzt man dies in die Definition der Differenzgleichung

$$y_t + a_1 y_{t-1} + \dots + a_m y_{t-m} = 0$$

ein, so erhält man

$$\gamma z^t + a_1 \gamma z^{t-1} + \dots + a_m \gamma z^{t-m} = 0.$$

Division durch  $\gamma$  und  $z^{t-m}$  führt auf die Gleichung

$$z^m + a_1 z^{m-1} + \dots + a_{m-1} z + a_m = 0.$$

Diese Gleichung nennt man die **charakteristische Gleichung** der Differenzgleichung. Die linke Seite nennt man naheliegenderweise **charakteristisches Polynom**. Die Lösung der Differenzgleichung reduziert sich auf die Bestimmung der Nullstellen des charakteristischen Polynoms.

**Satz 3.16.** *Besitzt die charakteristische Gleichung*

$$z^m + a_1 z^{m-1} + \dots + a_{m-1} z + a_m = 0$$

*der homogenen linearen Differenzgleichung  $m$ -ter Ordnung genau  $m$  verschiedene reelle Lösungen  $z_1, \dots, z_m$ , so ist die allgemeine Lösung durch*

$$y_t = \gamma_1 z_1^t + \gamma_2 z_2^t + \dots + \gamma_m z_m^t$$

*gegeben, wobei  $\gamma_1, \dots, \gamma_m \in \mathbb{R}$  beliebige Konstanten sind.*

Die allgemeine Lösung für den Fall, dass keine  $m$  verschiedene reelle Lösungen der charakteristischen Gleichung existieren, ist leider deutlich komplizierter, und kann hier nicht behandelt werden.

**Beispiel 3.17.** Wir kommen zurück zum Beispiel der Fibonacci-Folge aus Beispiel 2.2. Da man die Rekursionsgleichung

$$y_n = y_{n-1} + y_{n-2}$$

umformen kann zu  $y_n - y_{n-1} - y_{n-2} = 0$ , liegt also eine homogene lineare Differenzgleichung zweiter Ordnung vor mit charakteristischem Polynom  $z^2 - z - 1$ . Dies hat die Nullstellen

$$z_{1,2} = \frac{1}{2} + \sqrt{\frac{1}{4} + 1} = \frac{1 \pm \sqrt{5}}{2}.$$

Die allgemeine Lösung ist somit

$$y_n = \gamma_1 \left( \frac{1 + \sqrt{5}}{2} \right)^n + \gamma_2 \left( \frac{1 - \sqrt{5}}{2} \right)^n$$

Um die Anfangsbedingung  $y_0 = 0$  und  $y_1 = 1$  zu erfüllen, muss gelten

$$\begin{aligned} 0 &= \gamma_1 \left( \frac{1 + \sqrt{5}}{2} \right)^0 + \gamma_2 \left( \frac{1 - \sqrt{5}}{2} \right)^0, \\ 1 &= \gamma_1 \left( \frac{1 + \sqrt{5}}{2} \right)^1 + \gamma_2 \left( \frac{1 - \sqrt{5}}{2} \right)^1, \end{aligned}$$

also wegen der ersten Gleichung  $\gamma_1 + \gamma_2 = 0$ , d.h.  $\gamma_2 = -\gamma_1$ . Dies in die zweite Gleichung eingesetzt, liefert

$$\gamma_1 \left( \frac{1 + \sqrt{5}}{2} \right)^1 - \gamma_1 \left( \frac{1 - \sqrt{5}}{2} \right)^1 = \gamma_1 \sqrt{5} = 1.$$

Also ist  $\gamma_1 = 1/\sqrt{5}$  und  $\gamma_2 = -1/\sqrt{5}$  und somit

$$y_n = \frac{1}{\sqrt{5}} \left( \frac{1 + \sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \left( \frac{1 - \sqrt{5}}{2} \right)^n.$$

Für die inhomogene Gleichung mit konstanter rechter Seite

$$y_t + a_1 y_{t-1} + \dots + a_m y_{t-m} = b$$

errät man wieder die konstante Lösung

$$y_t = \frac{b}{1 + a_1 + \dots + a_m}$$

(falls  $1 + a_1 + \dots + a_m \neq 0$ ), und wie oben zeigt man auch hier durch Differenzenbildung, dass die Differenz zwischen zwei Lösungen der inhomogenen Gleichung eine Lösung der homogenen Gleichung ist. Wir erhalten also folgendes Resultat.

**Satz 3.18.** Ist  $1 + a_1 + \dots + a_m \neq 0$ , so besitzt die inhomogene lineare Differenzgleichung

$$y_t + a_1 y_{t-1} + \dots + a_m y_{t-m} = b$$

die allgemeine Lösung

$$y_t = \frac{b}{1 + a_1 + \dots + a_m} + y_t^{hom},$$

wobei  $y_t^{hom}$  die allgemeine Lösung der zugehörigen homogenen linearen Differenzgleichung bezeichnet.