Mathematik für Naturwissenschaftler II

Dr. Peter Bauer

Institut für Mathematik
Universität Frankfurt am Main

Sommersemester 2025

Lineare Algebra

1 Der mehrdimensionale Raum

1.1 Vektoren

Im Teil I dieser Vorlesung haben wir uns insbesondere mit Funktionen einer Variablen beschäftigt. Viele naturwissenschaftliche Beobachtungen hängen aber nicht nur von einer Größe, sondern von mehreren Größen ab.

Wir betrachten daher nun statt der (eindimensionalen) Zahlen mehrdimensionale Objekte:

1.1.1 Bezeichnung Sei $n \in \mathbb{N}$ und seien $v_1, v_2, \ldots, v_n$ reelle Zahlen. $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$ ist ein *n-dimensionaler Vektor*. Die Menge dieser Vektoren ist der *n-dimensionale reelle Raum* $\mathbb{R}^n$. Der Vektor $\vec{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ wird als *Nullvektor* bezeichnet.

Zur Unterscheidung werden reelle oder komplexe Zahlen im Gegensatz zu Vektoren oft als *Skalare* bezeichnet.

1.1.2 Bemerkung In der Literatur wird statt $\vec{v}$ oft $\mathbf{v}$ oder einfach v geschrieben. In letzterem Fall muss dann dem Kontext entnommen werden, ob v ein Vektor oder ein Skalar ist.

1.1.3 Bemerkung Verwenden wir in der obigen Definition komplexe statt reeller Zahlen, erhalten wir ganz analog den *n-dimensionalen komplexen Raum* $\mathbb{C}^n$.

Wir werden uns hier aber fast nur mit dem $\mathbb{R}^n$ befassen, die Behandlung des $\mathbb{C}^n$ ist völlig analog.

Natürlich kann mit Vektoren auch gerechnet werden:

1.1.4 Definition Seien $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ und $\vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$ Vektoren des $\mathbb{R}^n$ und sei $\lambda \in \mathbb{R}$. Es gilt

$$\vec{v} + \vec{w} = \begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{pmatrix} \qquad \text{und} \qquad \lambda\vec{v} = \begin{pmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{pmatrix} .$$

1.1.5 Beispiele

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

Für diese Rechenarten gelten die Regeln, die wir z.B. von den reellen Zahlen her kennen:

1.1.6 Satz (Rechenregeln) Seien $\lambda, \mu \in \mathbb{R}$ und $\vec{v}, \vec{w}, \vec{z} \in \mathbb{R}^n$. Es gilt

(1) $\vec{v}+\vec{w}=\vec{w}+\vec{v}$ (Kommutativität der Addition)

(2) $(\vec{v}+\vec{w})+\vec{z}=\vec{v}+(\vec{w}+\vec{z})$ (Assoziativität der Addition)

(3) $\lambda\cdot\vec{v}=\vec{v}\cdot\lambda$ (Kommutativität der Multiplikation)

(4) $\lambda(\mu\vec{v})=(\lambda\mu)\vec{v}$ (Assoziativität der Multiplikation)

(5) $\lambda(\vec{v}+\vec{w})=\lambda\vec{v}+\lambda\vec{w}$ und $(\lambda+\mu)\vec{v}=\lambda\vec{v}+\mu\vec{v}$ (Distributivität)

Die Addition von Vektoren und die Multiplikation eines Vektors mit einem Skalar lassen sich auch geometrisch interpretieren: Der Addition zweier Vektoren entspricht das Hintereinanderfügen der Vektoren, die Multiplikation mit einem Skalar streckt (bzw. staucht) den Vektor. Im Fall eines negativen Skalars wird die Orientierung des Vektors umgekehrt.

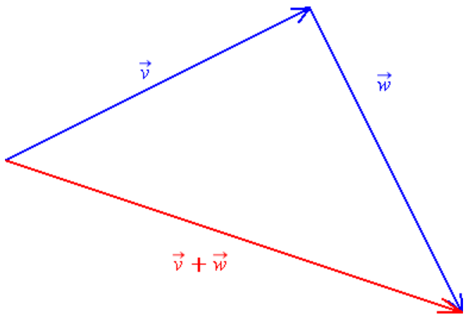


Abbildung 1: Addition von Vektoren.

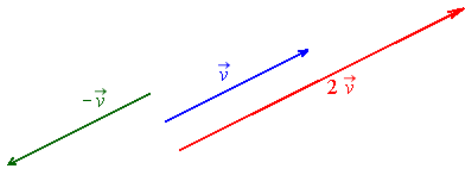


Abbildung 2: Multiplikation eines Vektors mit einem Skalar.

1.2 Norm

Die „Länge“ eines Vektors kann in einem zweidimensionalen Raum, also einer Ebene, über den Satz des Pythagoras bestimmt werden. Die Länge des Vektors $\binom{a}{b}$ ist damit $\sqrt{a^2+b^2}$.

Dies gilt analog auch in höheren Dimensionen. Statt „Länge“ verwendet man die Bezeichnung „Norm“.

1.2.1 Definition Sei $\vec{v}=\begin{pmatrix}v_1\\\vdots\\v_n\end{pmatrix}\in\mathbb{R}^n$. Die *Norm* von $\vec{v}$ ist

$$\|\vec{v}\|=\sqrt{v_1^2+\cdots+v_n^2}=\left(\sum_{i=1}^n v_i^2\right)^{1/2} .$$

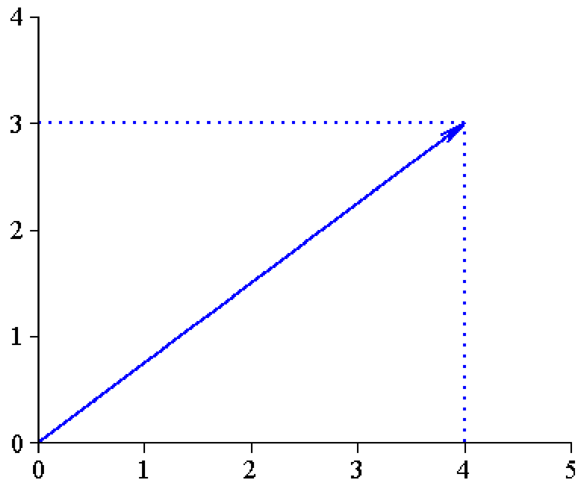


Abbildung 3: $\left\|\binom{4}{3}\right\| = \sqrt{4^2+3^2} = 5$

1.2.2 Beispiel

$$\left\|\begin{pmatrix} 1 \\ -1 \end{pmatrix}\right\| = \left(1^2 + (-1)^2\right)^{1/2} = \sqrt{2}$$

1.2.3 Bemerkung In einem *kartesischen Koordinatensystem*, bei dem die Koordinatenachsen senkrecht aufeinander stehen, entspricht die Norm eines Vektors gerade dem, was wir uns unter seiner Länge vorstellen.

Es gibt aber auch andere Koordinatensysteme, bei denen die Achsen nicht senkrecht aufeinander stehen. Die Norm ist auch dort definiert, sie entspricht jedoch nicht mehr unserer Vorstellung von Länge. Wir werden uns in Abschnitt 2.6 näher mit solchen Koordinatensystemen befassen.

1.2.4 Satz Seien $\vec{v}, \vec{w} \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$. Es gilt

(1) $\|\vec{v}\| \geq 0 \quad$ und $\quad \|\vec{v}\| = 0 \;\Leftrightarrow\; \vec{v} = \vec{0}$

(2) $\|\lambda\vec{v}\| = |\lambda| \cdot \|\vec{v}\|$

(3) $\|\vec{v} + \vec{w}\| \leq \|\vec{v}\| + \|\vec{w}\|$ (Dreiecksungleichung)

Die Bezeichnung „Dreiecksungleichung“ erklärt sich, wenn wir uns Abbildung 1 anschauen.

1.3 Skalarprodukt

Während die Addition von Vektoren in naheliegender Weise definiert werden kann, ist das bei der Multiplikation von Vektoren nicht so einfach. Wir kennen bereits die Multiplikation eines Vektors mit einem Skalar. Das ist aber eigentlich eine eher ungewöhnliche Multiplikation, da hier verschiedene Objekte, nämlich Vektoren und Skalare, miteinander verknüpft werden.

Tatsächlich gibt es mehrere Multiplikationen in Zusammenhang mit Vektoren, die alle in der ein oder anderen Hinsicht ungewöhnlich sind.

1.3.1 Definition Seien $\vec{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ und $\vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ Vektoren des $\mathbb{R}^n$. Dann heißt die reelle Zahl

$$\vec{a} \cdot \vec{b} = \sum_{i=1}^{n} a_i b_i = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

Skalarprodukt von $\vec{a}$ und $\vec{b}$.

1.3.2 Beispiele **(1)** $\begin{pmatrix} 1 \\ -3 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ -1 \\ -2 \end{pmatrix} = 1 \cdot 3 + (-3)(-1) + 2 \cdot (-2) = 3 + 3 - 4 = 2$

(2) $\vec{0} \cdot \vec{a} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = 0 \cdot a_1 + 0 \cdot a_2 + \cdots + 0 \cdot a_n = 0,$

also gilt $\vec{0} \cdot \vec{a} = 0$ für alle $\vec{a} \in \mathbb{R}^n$, was wir wohl nicht anders erwartet haben.

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

Es ist also möglich, dass das Skalarprodukt zweier Vektoren 0 ist, obwohl keiner der beiden Vektoren der Nullvektor ist.

1.3.3 Satz (Rechenregeln) Seien $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$. Dann gilt

(1) $\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}$ (Kommutativität)

(2) $\left(\vec{a} + \vec{b}\right) \cdot \vec{c} = \vec{a} \cdot \vec{c} + \vec{b} \cdot \vec{c}$ (Distributivität)

(3) $\lambda\left(\vec{a} \cdot \vec{b}\right) = \left(\lambda\vec{a}\right) \cdot \vec{b} = \vec{a} \cdot \left(\lambda\vec{b}\right)$

Das Skalarprodukt ist nicht nur ungewöhnlich, weil diese Multiplikation zweier Vektoren einen Skalar liefert, auch manche vertraute Rechenregeln gelten nicht. Das Skalarprodukt ist beispielsweise *nicht assoziativ*:

$$\left(\binom{1}{0} \cdot \binom{0}{2}\right) \cdot \binom{1}{1} = 0 \cdot \binom{1}{1} = \binom{0}{0}$$
$$\binom{1}{0} \cdot \left(\binom{0}{2} \cdot \binom{1}{1}\right) = \binom{1}{0} \cdot 2 = \binom{2}{0}$$

Unter Umständen ist also $\left(\vec{a} \cdot \vec{b}\right) \cdot \vec{c} \neq \vec{a} \cdot \left(\vec{b} \cdot \vec{c}\right)$.

1.3.4 Bemerkung Wir unterscheiden in der Schreibweise nicht zwischen den drei verschiedenen Multiplikationen (Multiplikation zweier Skalare, Multiplikation eines Vektors mit einem Skalar und das Skalarprodukt zweier Vektoren): in allen Fällen schreiben wir einen Punkt „·“, den wir meistens auch einfach weglassen.

Man muss also stets dem Kontext entnehmen, welches Produkt gemeint ist.

Es gibt einen oft recht nützlichen Zusammenhang zwischen dem Skalarprodukt und der Norm:

$$\vec{a} \cdot \vec{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = a_1^2 + a_2^2 + \cdots + a_n^2 = \|\vec{a}\|^2 \ .$$

Wir haben also

1.3.5 Satz Für $\vec{a} \in \mathbb{R}^n$ gilt $\vec{a} \cdot \vec{a} = \|\vec{a}\|^2$.

1.4 Winkel

In der Ebene gilt der *Cosinussatz*: Sind a, b, c die Längen der Seiten eines Dreiecks und liegt der Winkel γ der Seite c gegenüber, dann gilt

$$a^2 + b^2 - 2ab \cdot \cos\gamma = c^2 \ .$$

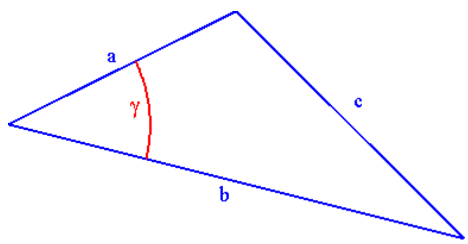


Abbildung 4: $a^2 + b^2 - 2ab \cdot \cos\gamma = c^2$ (Cosinussatz)

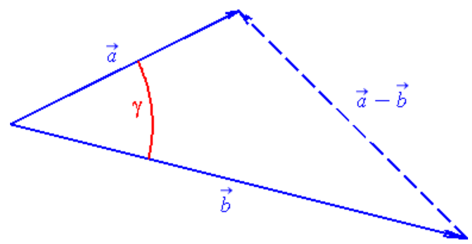


Abbildung 5: Winkeldefinition

Sind $\vec{a}$ und $\vec{b}$ Vektoren des $\mathbb{R}^n$, die den Winkel γ einschließen, dann bilden $\vec{a}$, $\vec{b}$ und $\vec{a} - \vec{b}$ ein Dreieck und der Cosinussatz ergibt

$$\begin{aligned}
\|\vec{a}\|^2 + \|\vec{b}\|^2 - 2\|\vec{a}\|\|\vec{b}\| \cos\gamma &= \|\vec{a} - \vec{b}\|^2 \\
&= \left(\vec{a} - \vec{b}\right) \cdot \left(\vec{a} - \vec{b}\right) \\
&= \vec{a} \cdot \left(\vec{a} - \vec{b}\right) - \vec{b} \cdot \left(\vec{a} - \vec{b}\right) \\
&= \vec{a} \cdot \vec{a} - \vec{a} \cdot \vec{b} - \vec{b} \cdot \vec{a} + \vec{b} \cdot \vec{b} \\
&= \|\vec{a}\|^2 - 2\,\vec{a} \cdot \vec{b} + \|\vec{b}\|^2 \\
\Rightarrow \qquad \|\vec{a}\|\|\vec{b}\| \cos\gamma &= \vec{a} \cdot \vec{b}
\end{aligned}$$

Dies können wir zur Definition des Winkels zwischen zwei Vektoren nutzen:

1.4.1 Definition Seien $\vec{a}, \vec{b} \in \mathbb{R}^n$ mit $\vec{a} \neq \vec{0}$ und $\vec{b} \neq \vec{0}$. Bezeichnet $\sphericalangle(\vec{a}, \vec{b})$ den *Winkel* zwischen $\vec{a}$ und $\vec{b}$, dann ist

$$\cos \sphericalangle(\vec{a}, \vec{b}) = \frac{\vec{a} \cdot \vec{b}}{\|\vec{a}\|\|\vec{b}\|} \ .$$

1.4.2 Bemerkung Wie schon bei der Norm entspricht der so definierte Winkel nur in kartesischen Koordinatensystemen unserer Vorstellung. Aber auch in nicht-kartesischen Koordinatensystemen kann der Winkel gemäß Definition 1.4.1 definiert werden.

1.4.3 Beispiele **(1)** $\vec{a} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ und $\vec{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

Es gilt

$$\cos \sphericalangle(\vec{a}, \vec{b}) = \frac{\binom{0}{1} \cdot \binom{1}{1}}{\left\| \binom{0}{1} \right\| \left\| \binom{1}{1} \right\|} = \frac{0+1}{1 \cdot \sqrt{2}} = \frac{1}{\sqrt{2}}$$

und damit ist $\sphericalangle(\vec{a}, \vec{b}) = \dfrac{\pi}{4}$.

Natürlich ist auch $\cos \frac{7\pi}{4} = \frac{1}{\sqrt{2}}$, wir wollen aber stets den kleineren Winkel zwischen den beiden Vektoren betrachten, also

$$0 \leq \sphericalangle(\vec{a}, \vec{b}) \leq \pi \ .$$

In diesem Intervall ist der Winkel auch eindeutig durch seinen Cosinus bestimmt.

(2) $\vec{a} = \begin{pmatrix} 2 \\ 2 \\ 0 \\ 0 \end{pmatrix}$ und $\vec{b} = \begin{pmatrix} 0 \\ 3 \\ 3 \\ 0 \end{pmatrix}$

Vektoren im $\mathbb{R}^4$ können wir uns zwar nicht vorstellen, den Winkel zwischen ihnen können wir aber leicht berechnen:

$$\cos \sphericalangle(\vec{a}, \vec{b}) = \frac{\begin{pmatrix} 2 \\ 2 \\ 0 \\ 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 3 \\ 3 \\ 0 \end{pmatrix}}{\left\| \begin{pmatrix} 2 \\ 2 \\ 0 \\ 0 \end{pmatrix} \right\| \left\| \begin{pmatrix} 0 \\ 3 \\ 3 \\ 0 \end{pmatrix} \right\|} = \frac{6}{\sqrt{4+4+0+0} \cdot \sqrt{0+9+9+0}} = \frac{6}{\sqrt{144}} = \frac{6}{12} = \frac{1}{2} \ ,$$

also ist $\sphericalangle(\vec{a}, \vec{b}) = \dfrac{2\pi}{3}$.

(3) $\vec{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ und $\vec{b} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$

$$\cos \sphericalangle(\vec{a}, \vec{b}) = \frac{\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}}{\left\| \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\| \left\| \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right\|} = \frac{1+2-3}{\left\| \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\| \left\| \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right\|} = 0 \ ,$$

also ist $\sphericalangle(\vec{a}, \vec{b}) = \dfrac{\pi}{2}$, die beiden Vektoren stehen also senkrecht aufeinander.

Das letzte Beispiel zeigt:

1.4.4 Satz Seien $\vec{a}, \vec{b} \in \mathbb{R}^n$, beide Vektoren seien $\neq \vec{0}$. Es gilt $\vec{a} \cdot \vec{b} = 0$ genau dann, wenn $\sphericalangle(\vec{a}, \vec{b}) = \frac{\pi}{2}$, wenn die Vektoren also senkrecht aufeinander stehen. $\vec{a}$ und $\vec{b}$ heißen dann *orthogonal* zueinander.

1.4.5 Beispiel Die Vektoren $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ sind jeweils paarweise orthogonal zueinander.

Die Vektoren des letzten Beispiels haben außerdem alle die Norm 1. Daher rührt folgende Bezeichnung:

1.4.6 Definition Die Vektoren des $\mathbb{R}^n$

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} , \ \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} , \ \dots \ , \ \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

heißen *Einheitsvektoren.*

Alle Einheitsvektoren sind jeweils paarweise orthogonal zueinander und haben die Norm 1.

Verwenden wir noch eine weitere Bezeichnung:

1.4.7 Bezeichnung Sind $\lambda_1, \dots, \lambda_k \in \mathbb{R}$, dann heißt ein Ausdruck der Form

$$\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + \dots + \lambda_k \vec{v}_k$$

Linearkombination der Vektoren $\vec{v}_1, \dots, \vec{v}_2 \in \mathbb{R}^n$[1]

Damit gilt dann:

1.4.8 Satz Jeder Vektor $\vec{v} \in \mathbb{R}^n$ ist in eindeutiger Weise als Linearkombination der Einheitsvektoren $\vec{e}_i \in \mathbb{R}^n$ darstellbar:

$$\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \sum_{i=1}^{n} v_i \vec{e}_i \ . \tag{1}$$

Die Komponenten v_i von $\vec{v}$ sind also gerade die Koeffizienten der Linearkombination (1).

1.5 Vektorprodukt

Wir kennen nun eine Multiplikation von Vektoren, nämlich das Skalarprodukt. Dieses Produkt hat als Ergebnis aber keinen Vektor, sondern einen Skalar. Man kennt auch eine Multiplikation von Vektoren, die als Ergebnis einen Vektor ergibt. Dieses Produkt gibt es allerdings *nur im dreidimesionalen Raum*:

1.5.1 Definition Seien $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} \in \mathbb{R}^3$. Dann ist das *Vektorprodukt* von $\vec{a}$ und $\vec{b}$

$$\vec{a} \times \vec{b} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix} \ .$$

1.5.2 Beispiele **(1)** Seien $\vec{a} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ und $\vec{b} = \begin{pmatrix} 0 \\ -4 \\ 6 \end{pmatrix}$.

$$\vec{a} \times \vec{b} = \begin{pmatrix} 2 \cdot 6 - 3 \cdot (-4) \\ 3 \cdot 0 - 1 \cdot 6 \\ 1 \cdot (-4) - 2 \cdot 0 \end{pmatrix} = \begin{pmatrix} 24 \\ -6 \\ -4 \end{pmatrix}$$

[1] Wenn keine Quadrate oder höheren Potenzen, keine trigonometrischen Funktionen o.a. auftreten, spricht man von *linear.*

(2) Seien $\vec{a}$ und $\vec{b}$ wie in (1).

$$\vec{b} \times \vec{a} = \begin{pmatrix} (-4) \cdot 3 - 6 \cdot 2 \\ 6 \cdot 1 - 0 \cdot 3 \\ 0 \cdot 2 - (-4) \cdot 1 \end{pmatrix} = \begin{pmatrix} -24 \\ 6 \\ 4 \end{pmatrix}$$

Es gilt hier also $\vec{a} \times \vec{b} \neq \vec{b} \times \vec{a}$, das Vektorprodukt ist folglich *nicht kommutativ* !

Es gilt aber:

1.5.3 Satz (Rechenregeln) Seien $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$ und $\lambda \in \mathbb{R}$. Dann gilt

(1) $\vec{a} \times \vec{b} = -(\vec{b} \times \vec{a})$ (Antikommutativität)

(2) $(\lambda\vec{a}) \times \vec{b} = \vec{a} \times (\lambda\vec{b}) = \lambda(\vec{a} \times \vec{b})$

(3) $\vec{a} \times (\vec{b} + \vec{c}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{c}$ und $(\vec{a} + \vec{b}) \times \vec{c} = \vec{a} \times \vec{c} + \vec{b} \times \vec{c}$ (Distributivität)

1.5.4 Bemerkung Da das Vektorprodukt ein Produkt ist, müssen nach der „Punkt vor Strich"-Regel die Produkte auf der rechten Seite der Gleichungen in (3) nicht in Klammern stehen.

Die Definition des Vektorprodukts ist recht abstrakt. Tatsächlich ist das Vektorprodukt aber eine Verknüpfung, die man sich sehr gut geometrisch vorstellen kann.

Betrachten wir die folgende Kombination von Vektor- und Skalarprodukt:

$$\begin{aligned} (\vec{a} \times \vec{b}) \cdot \vec{a} &= \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \\ &= a_2b_3a_1 - a_3b_2a_1 + a_3b_1a_2 - a_1b_3a_2 + a_1b_2a_3 - a_2b_1a_3 \\ &= 0 \end{aligned}$$

Ganz analog ist

$$(\vec{a} \times \vec{b}) \cdot \vec{b} = 0 \ .$$

Das heißt, der Vektor $\vec{a} \times \vec{b}$ steht senkrecht sowohl auf $\vec{a}$ als auch auf $\vec{b}$.

Stellen wir uns zwei Vektoren $\vec{a}$ und $\vec{b}$ im dreidimensionalen Raum vor, so sehen wir, dass es gerade zwei Richtungen gibt, die senkrecht auf diesen beiden Vektoren stehen. Die Orientierung des Produkts $\vec{a}\times\vec{b}$ lässt sich mit der *Rechte-Hand-Regel* (Abbildung 6) ermitteln: Zeigt der Daumen der rechten Hand in Richtung $\vec{a}$, der Zeigefinger in Richtung $\vec{b}$, dann deutet der Mittelfinger der rechten Hand in Richtung $\vec{a} \times \vec{b}$.

Wir haben damit eine Anschauung, in welche Richtung das Vektorprodukt zweier Vektoren zeigt. Auch für die Länge des Vektors $\vec{a} \times \vec{b}$ gibt es eine einfache geometrische Interpretation: es handelt sich gerade um die Fläche des Parallelogramms, das von $\vec{a}$ und $\vec{b}$ aufgespannt wird. Bezeichnet F diese Fläche, ist also

$$F = \|\vec{a} \times \vec{b}\| \ . \tag{2}$$

Die Fläche eines Parallelogramms ist das Produkt der Länge der Grundseite und der Höhe h, also

$$F = \|\vec{a}\| \cdot h \ .$$

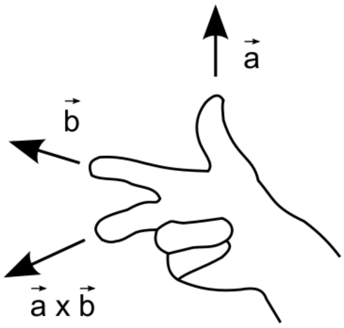


Abbildung 6: Rechte-Hand-Regel und Vektorprodukt

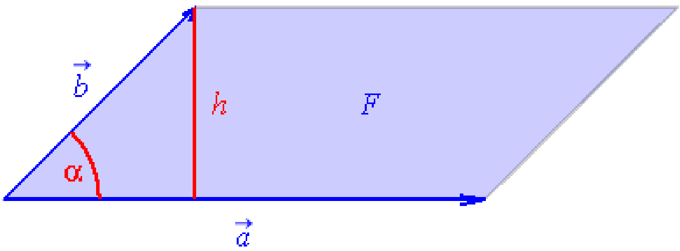


Abbildung 7: $\|\vec{a} \times \vec{b}\|$ entspricht der Fläche des aufgespannten Parallelogramms

Sei $\alpha = \sphericalangle(\vec{a}, \vec{b})$ der Winkel zwischen $\vec{a}$ und $\vec{b}$. Dann gilt

$$\sin\alpha = \frac{h}{\|\vec{b}\|} \qquad \Rightarrow \qquad h = \|\vec{b}\| \cdot \sin\alpha$$

und somit insgesamt

$$F = \|\vec{a}\| \cdot \|\vec{b}\| \cdot \sin\alpha \ .$$

Zusammen mit (2) haben wir also

1.5.5 Satz Seien $\vec{a}, \vec{b} \in \mathbb{R}^3$. Dann gilt

$$\|\vec{a} \times \vec{b}\| = \|\vec{a}\| \cdot \|\vec{b}\| \cdot \sin\sphericalangle(\vec{a}, \vec{b}) \ .$$

1.5.6 Bemerkung Es gilt also

$$\sin\sphericalangle(\vec{a}, \vec{b}) = \frac{\|\vec{a} \times \vec{b}\|}{\|\vec{a}\| \cdot \|\vec{b}\|} \ ,$$

aber dies ist zur Berechnung des Winkels weniger geeignet als Definition 1.4.1, da z.B. $\sin\frac{\pi}{4} = \frac{1}{\sqrt{2}} = \sin\frac{3\pi}{4}$. Der Winkel lässt sich aus dem Sinus also auch durch Beschränkung auf das Intervall $[0, \pi]$ nicht eindeutig bestimmen.

Wie wir wissen, ist das Skalarprodukt zweier Vektoren genau dann Null, wenn die Vektoren senkrecht aufeinander stehen. Stellen wir uns die analoge Frage: Wann ergibt das Vektorprodukt zweier Vektoren den Nullvektor $\vec{0}$?

Es ist

$$\vec{0} = \vec{a} \times \vec{b} \ \Leftrightarrow \ 0 = \|\vec{a} \times \vec{b}\| = \|\vec{a}\| \cdot \|\vec{b}\| \cdot \sin\sphericalangle(\vec{a}, \vec{b})$$

und dies ist der Fall, wenn $\vec{a} = \vec{0}$ oder $\vec{b} = \vec{0}$ oder $\sin \sphericalangle(\vec{a}, \vec{b}) = 0$ und dies gilt wiederum für

$$\sphericalangle(\vec{a}, \vec{b}) = 0 \text{ oder } = \pi ,$$

wenn also $\vec{a}$ und $\vec{b}$ „in einer Linie liegen“. Man bezeichnet solche Vektoren als *kollinear*.

Wir haben also:

1.5.7 Satz Seien $\vec{a}, \vec{b} \in \mathbb{R}^3$. Dann sind $\vec{a}$ und $\vec{b}$ genau dann kollinear, wenn $\vec{a} \times \vec{b} = \vec{0}$.

1.6 Spatprodukt

Das Spatprodukt ist ein Produkt von drei(!) Vektoren des $\mathbb{R}^3$. Es handelt sich um eine Kombination von Skalar- und Vektorprodukt:

1.6.1 Definition Seien $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$. Dann ist die reelle Zahl

$$\vec{a} \cdot (\vec{b} \times \vec{c})$$

das *Spatprodukt* von $\vec{a}$, $\vec{b}$ und $\vec{c}$.

Die wesentlichen Eigenschaften des Spatprodukts beschreibt der folgende Satz:

1.6.2 Satz Seien $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$.

(1) $\vec{a} \cdot (\vec{b} \times \vec{c}) = \vec{b} \cdot (\vec{c} \times \vec{a}) = \vec{c} \cdot (\vec{a} \times \vec{b})$ („zyklische Vertauschung“)

(2) $\vec{a} \cdot (\vec{b} \times \vec{c}) = 0$ gilt genau dann, wenn $\vec{a}, \vec{b}, \vec{c}$ *komplanar* sind, d.h. wenn diese Vektoren „in einer Ebene liegen“.

(3) $\left|\vec{a} \cdot (\vec{b} \times \vec{c})\right|$ ist das Volumen des Spats, der von $\vec{a}, \vec{b}, \vec{c}$ aufgespannt wird (Abbildung 8).

(4) Ist V das Volumen des von $\vec{a}, \vec{b}, \vec{c}$ aufgespannten Spats, dann gilt

$$V = \|\vec{a}\| \cdot \|\vec{b}\| \cdot \|\vec{c}\| \cdot \sin \sphericalangle(\vec{b}, \vec{c}) \cdot \cos \sphericalangle(\vec{a}, \vec{b} \times \vec{c}) .$$

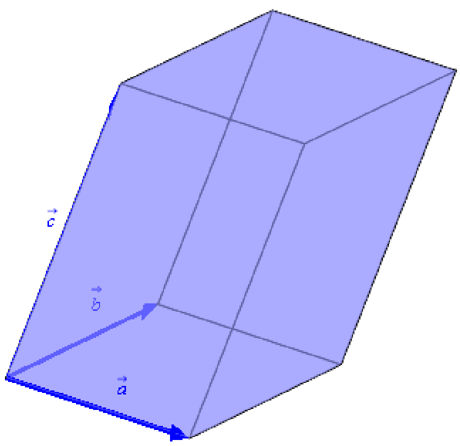


Abbildung 8: $|\vec{a} \cdot (\vec{b} \times \vec{c})|$ entspricht dem Volumen des aufgespannten Spats

Der Betrag des Spatprodukts gibt also das Volumen des aufgespannten Spats an. Auch das Vorzeichen hat eine Interpretation:

1.6.3 Bezeichnung Drei Vektoren $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$ bilden ein ...

(1) *Rechtssystem*, wenn $\vec{a} \cdot (\vec{b} \times \vec{c}) > 0$,

(2) *Linkssystem*, wenn $\vec{a} \cdot (\vec{b} \times \vec{c}) < 0$.

Offensichtlich ist hier die Reihenfolge der drei Vektoren entscheidend.

Übrigens bilden Daumen, Zeige- und Mittelfinger der rechten (linken) Hand ein Rechtssystem (Linkssystem).

1.7 Lineare Unabhängigkeit

Betrachten wir eine Menge von Vektoren im $\mathbb{R}^n$, beispielsweise drei Vektoren im $\mathbb{R}^2$, wie wir sie in Abbildung 9 sehen.

Es ist möglich, einen der Vektoren als Linearkombination der anderen Vektoren darzustellen, in Abbildung 9 ist zum Beispiel

$$\vec{v}_1 = \frac{5}{3}\vec{v}_2 + \frac{7}{3}\vec{v}_3 \ .$$

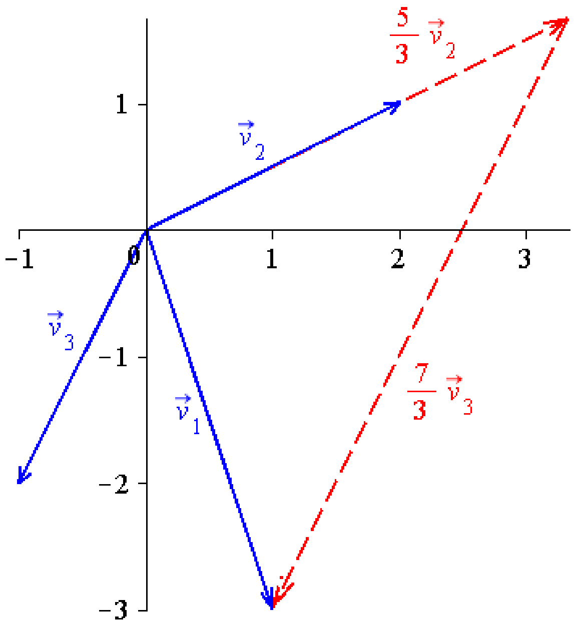


Abbildung 9: $\vec{v}_1 = \frac{5}{3}\vec{v}_2 + \frac{7}{3}\vec{v}_3$

1.7.1 Bezeichnung Kann in einer Menge von Vektoren im $\mathbb{R}^n$ ein Vektor als Linearkombination der anderen Vektoren dargestellt werden, nennt man diese Vektoren *linear abhängig*.

Ist dies nicht möglich, nennt man diese Vektoren *linear unabhängig*.

1.7.2 Beispiel Die Einheitsvektoren $\vec{e}_1, \ldots, \vec{e}_n \in \mathbb{R}^n$ sind linear unabhängig.

Anschaulich kann man sich dies im $\mathbb{R}^3$ (oder auch $\mathbb{R}^2$) klarmachen:

Betrachten wir nur zwei der drei Einheitsvektoren des $\mathbb{R}^3$, dann liegen diese beiden in einer Ebene. Jede Linearkombination der beiden Vektoren wird ebenfalls in dieser Ebene liegen.

Der dritte Einheitsvektor steht aber senkrecht auf der Ebene, kann also keine Linearkombination der anderen sein.

Eine Methode, auszurechnen, wann eine Menge von Vektoren linear unabhängig ist oder nicht, werden wir noch kennenlernen. Bis dahin können wir aber bereits einige allgemeine Gesetze notieren:

Haben wir z.B. Vektoren $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^n$, dann gilt stets $0 \cdot \vec{v}_1 + 0 \cdot \vec{v}_2 = \vec{0}$. Der Nullvektor kann also immer als Linearkombination anderer Vektoren dargestellt werden:

1.7.3 Satz Vektoren sind stets linear abhängig, wenn einer der Vektoren der Nullvektor ist.

Auch, wenn die Zahl der Vektoren größer ist, als die Dimension des Raumes, können wir immer eine Linearkombination finden:

1.7.4 Satz Seien $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k \in \mathbb{R}^n$. Ist $k > n$, dann sind die Vektoren linear abhängig.

Im $\mathbb{R}^n$ können also höchstens n Vektoren linear unabhängig sein und unter diesen darf nicht der Nullvektor vorkommen. Dieser Fall hat eine besondere Bedeutung:

1.7.5 Definition Seien $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n \in \mathbb{R}^n \setminus \{\vec{0}\}$ linear unabhängige Vektoren. Dann heißt die Menge $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ *Basis* des $\mathbb{R}^n$.

1.7.6 Beispiel Die Einheitsvektoren $\vec{e}_1, \ldots, \vec{e}_n \in \mathbb{R}^n$ sind (s.o.) linear unabhängig und bilden, da es sich um n Vektoren handelt, eine Basis des $\mathbb{R}^n$.

Die Bedeutung von Basen zeigt folgender Satz:

1.7.7 Satz Sei $B = \{\vec{b}_1, \ldots, \vec{b}_n\}$ eine Basis des $\mathbb{R}^n$. Dann ist jeder Vektor $\vec{v} \in \mathbb{R}^n$ in eindeutiger Weise als Linearkombination der Basisvektoren $\vec{b}_i \in B$ darstellbar:

$$\vec{v} = \sum_{i=1}^{n} \mu_i \vec{b}_i \qquad \text{mit } \mu_i \in \mathbb{R}. \tag{3}$$

Für die Basis aus Einheitsvektoren hatten wir dieses Ergebnis bereits in Satz 1.4.8. Dort konnten wir sogar die Koeffizienten μ_i angeben: Es handelte sich um die Komponenten v_i des Vektors

$$\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} .$$

1.7.8 Definition Sei $B = \{\vec{b}_1, \ldots, \vec{b}_n\}$ eine Basis des $\mathbb{R}^n$. Die Koeffizienten der Linearkombination (3) sind die *Koordinaten* des Vektors $\vec{v}$ bezüglich dieser Basis:

$$\vec{v} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}_B .$$

Bezüglich der Basis E aus Einheitsvektoren erhalten wir die gewohnte Darstellung

$$\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_E = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} .$$

1.7.9 Bemerkung Die Basis E aus Einheitsvektoren liefert also die Darstellung von Vektoren bezüglich des *kartesischen Koordinatensystems*. Andere Basen definieren andere, evtl. *schiefwinklige* Koordinatensysteme.

1.7.10 Beispiel Betrachten wir $\vec{v} = \binom{2}{1}$. Es ist offensichtlich $\vec{v} = 2\vec{e}_1 + 1\vec{e}_2$.

Die Vektoren $\vec{b}_1 = \binom{-1}{0}$ und $\vec{b}_2 = \binom{-1}{1}$ sind linear unabhängig, bilden also eine Basis $B = \{\vec{b}_1, \vec{b}_2\}$ des $\mathbb{R}^2$. Wegen

$$\vec{v} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} = -3\begin{pmatrix} -1 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \end{pmatrix} = -3\vec{b}_1 + 1\vec{b}_2$$

folgt

$$\vec{v} = \begin{pmatrix} -3 \\ 1 \end{pmatrix}_B$$

bezüglich der Basis B.

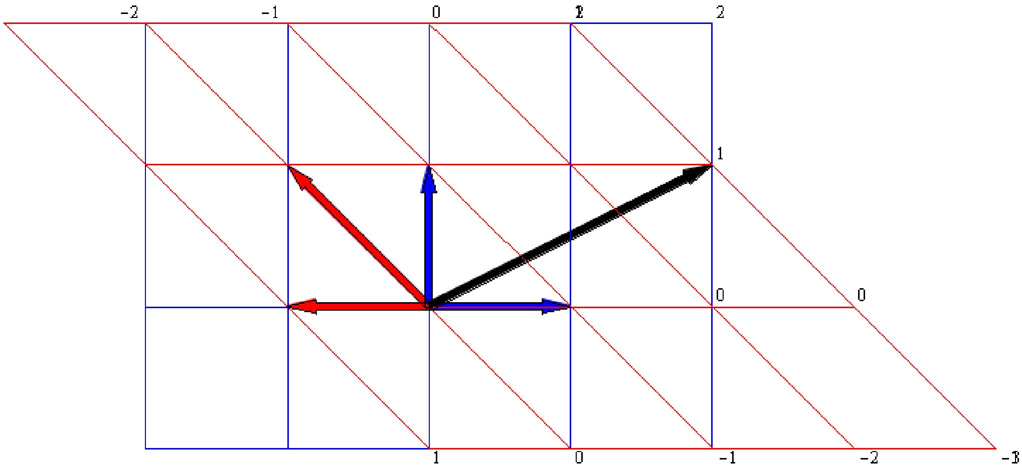


Abbildung 10: Der Vektor $\vec{v}$ (schwarz) bezüglich der Basis E (blau) und bezüglich der Basis B (rot). Es gilt $\vec{v} = \binom{2}{1}_E$ bzw. $\vec{v} = \binom{-3}{1}_B$.

Eine Frage bleibt: Wie findet man möglichst leicht die Darstellung eines Vektors bezüglich einer anderen Basis. Wir stellen diese Frage nun zurück, da sie in Abschnitt 2.6 viel leichter zu beantworten sein wird als mit unseren jetzigen Mitteln.

2 Lineare Gleichungssysteme

2.1 Matrizen

Gleichungen wie $3x = 4$ oder $x^2 - 2x + 1 = 0$ zu lösen, d.h. x zu bestimmen, ist eine der häufigsten Aufgaben, die sich uns stellen. Oftmals ist es sogar erforderlich, *Systeme* solcher Gleichungen zu lösen, d.h. mehrere Variablen aus mehreren Gleichungen zu bestimmen.

Im Allgemeinen ist es leider gar nicht möglich, eine solche Lösung zu finden. Handelt es sich aber um lineare Gleichungen, gibt es ein recht einfaches Verfahren, mit dem die Lösung(en) bestimmt werden können.

Ein *lineares Gleichungssystem* mit m Gleichungen und n *Unbekannten* $x_1, x_2, \dots, x_n$ ist das System

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= y_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= y_m \ . \end{aligned}$$

Gesucht sind die Variablen $x_1, x_2, \dots, x_n$, wenn die *Koeffizienten* $a_{11}, \dots, a_{mn}$ und die Konstanten $y_1, \dots, y_m$ gegeben sind.

2.1.1 Beispiel

$$\begin{aligned} x_1 - x_2 - 5x_3 &= 0 \\ 3x_1 + x_2 - 7x_3 &= 0 \\ x_1 + 3x_2 + 3x_3 &= 0 \end{aligned}$$

Der folgende Begriff dient zunächst der Vereinfachung der Schreibweise z.B. solcher Systeme:

2.1.2 Definition Das Koeffizientenschema

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = (a_{ik}) \qquad \text{mit } i = 1, \dots, m \text{ und } k = 1, \dots, n$$

heißt *Matrix* mit m Zeilen und n Spalten oder (m, n)*-Matrix* oder $m \times n$*-Matrix*.

2.1.3 Beispiel

$$\begin{pmatrix} 1 & -1 & -5 \\ 3 & 1 & -7 \\ 1 & 3 & 3 \end{pmatrix}$$

2.1.4 Bezeichnungen **(1)** Ist $m = n$, hat die Matrix also gleich viele Zeilen und Spalten, dann heißt die Matrix *quadratisch*.

(2) Die i-te Zeile $(a_{i1}, a_{i2}, \dots, a_{in})$ bildet einen *Zeilenvektor* des $\mathbb{R}^n$, die k-te Spalte $\begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{pmatrix}$ einen *Spaltenvektor* des $\mathbb{R}^m$.

(3) Die (m,n)-Matrix (a_{ik}) mit $a_{ik} = 0$ für alle i, k, also

$$\begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} = (0) = O$$

heißt *Nullmatrix*.

(4) Die (n,n)-Matrix

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} = (\delta_{ik}) = E \qquad \text{mit } \delta_{ik} = \begin{cases} 0 & \text{für } i \neq k \\ 1 & \text{für } i = k \end{cases}$$

heißt *Einheitsmatrix*.

2.2 Matrixoperationen

Auch wenn Matrizen zunächst als abkürzende Schreibweisen gedacht sind, ist auch möglich, mit ihnen zu rechnen:

2.2.1 Definition Sei $\lambda \in \mathbb{R}$ und seien $A = (a_{ik})$ und $B = (b_{ik})$ (m,n)-Matrizen, also

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \qquad \text{und} \qquad B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} .$$

Dann ist

(1) $A + B = (a_{ik} + b_{ik}) = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix}$

(2) $\lambda A = (\lambda a_{ik}) = \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix}$

(3) $-A = (-1) \cdot A$

2.2.2 Beispiel

$$2\begin{pmatrix} 2 & 0 \\ -1 & -2 \end{pmatrix} + \begin{pmatrix} 4 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ -2 & -4 \end{pmatrix} + \begin{pmatrix} 4 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 8 & -1 \\ -1 & -3 \end{pmatrix}$$

2.2.3 Bemerkung Die Addition von Matrizen, die unterschiedliche Zeilen- oder Spaltenzahl haben, ist nicht definiert.

Welche Rechenregeln haben ihre Gültigkeit ?

Die Definitionen der Addition und der Multiplikation mit einem Skalar ähneln den entsprechenden Operationen, die wir von Vektoren bereits kennen. Es ist also nicht überraschend, wenn auch die gleichen Gesetze gelten:

2.2.4 Satz (Rechenregeln „+“) Seien A, B, C (m, n)-Matrizen und seien $\lambda, \mu \in \mathbb{R}$. Es gilt:

(1) $A + B = B + A$ (Kommutativität)

(2) $(A + B) + C = A + (B + C)$ (Assoziativität)

(3) $A + O = A$ und $A - A = O$

(4) $\lambda(A + B) = \lambda A + \lambda B$ und $(\lambda + \mu)A = \lambda A + \mu A$ (Distributivität)

(5) $(\lambda\mu)A = \lambda(\mu A) = \mu(\lambda A)$ (Assoziativität)

Wie bei Vektoren ist auch bei Matrizen Multiplizieren deutlich komplizierter als z.B. die Addition:

2.2.5 Definition Sei $A = (a_{ij})$ eine (m, n)-Matrix und $B = (b_{jk})$ eine (n, p)-Matrix. Dann ist die (m, p)-Matrix $C = (c_{ik})$ mit

$$c_{ik} = \sum_{j=1}^{n} a_{ij} b_{jk} \qquad \text{für } i = 1, \ldots, m \text{ und } k = 1, \ldots, p$$

das *Produkt* von A und B:

$$C = A \cdot B \; .$$

2.2.6 Beispiel $\begin{pmatrix} 2 & 1 & -3 \\ -1 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 4 \\ 2 & 3 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 7 & 5 \\ -3 & 0 \end{pmatrix}$

2.2.7 Bemerkung Die Einträge c_{ik} von $C = A \cdot B$ sind das Skalarprodukt des i-ten Zeilenvektors von A und des k-ten Spaltenvektors von B.

Offenbar kann man nicht jede Matrix mit jeder anderen Matrix multiplizieren: wie aus der Definition hervorgeht, muss die Spaltenzahl des ersten Faktors der Zeilenzahl des zweiten Faktors entsprechen:

$$(m, n)\text{-Matrix} \cdot (n, p)\text{-Matrix} = (m, p)\text{-Matrix}$$

Dies bedeutet auch, dass das Matrixprodukt *nicht kommutativ* ist, d.h. im Allgemeinen gilt *nicht* $A \cdot B = B \cdot A$. Beispielsweise ist

$$\begin{pmatrix} 1 & 4 \\ 2 & 3 \\ -1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 & -3 \\ -1 & 0 & 2 \end{pmatrix} = \begin{pmatrix} -2 & 1 & 5 \\ 1 & 2 & 0 \\ -4 & -1 & 7 \end{pmatrix} .$$

Vergleichen wir dies mit Beispiel 2.2.6, sehen wir, dass hier $A \cdot B \neq B \cdot A$.

Bei der Multiplikation ist die Sache also nicht so einfach. Einige bekannte Rechenregeln haben aber auch beim Matrixprodukt ihre Gültigkeit:

2.2.8 Satz (Rechenregeln „·“) Seien A, B, C Matrizen und sei $\lambda \in \mathbb{R}$. Sofern die Produkte definiert sind, gilt

(1) $A \cdot (B \cdot C) = (A \cdot B) \cdot C$ (Assoziativität)

(2) $\lambda(A \cdot B) = (\lambda A) \cdot B = A \cdot (\lambda B)$

(3) $O \cdot A = O = A \cdot O$

(4) $A \cdot (B + C) = A \cdot B + A \cdot C$ und $(A + B) \cdot C = A \cdot C + B \cdot C$ (Distributivität)

(5) Ist A quadratisch und E die Einheitsmatrix gleicher Größe, dann gilt $A \cdot E = A = E \cdot A$.

2.3 Transposition

Im Zusammenhang mit Matrizen wird oft eine recht einfache, aber nützliche Operation benutzt:

2.3.1 Definition Sei $A = (a_{ik})$ eine (m,n)-Matrix. Die (n,m)-Matrix

$$A^T = (a_{ki})$$

heißt *transponierte* Matrix. A^T entsteht also durch Vertauschen von Zeilen und Spalten.

2.3.2 Beispiel $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$

2.3.3 Satz (Rechenregeln) Seien A, B, C Matrizen und $\lambda \in \mathbb{R}$. Sofern die folgenden Operationen definiert sind, gilt

(1) $\left(A^T\right)^T = A$

(2) $(\lambda A)^T = \lambda \left(A^T\right)$

(3) $(A+B)^T = A^T + B^T$

(4) $(A \cdot B)^T = B^T \cdot A^T$

2.3.4 Bemerkung Die Transposition erlaubt es, Vektoren wie Matrizen zu behandeln:

Betrachten wir $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ und $\vec{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$. Jeden Vektor können wir auch als $(n,1)$-Matrix auffassen. Wir können nun verschiedene Produkte bilden:

(1) $\vec{v} \cdot \vec{w}$ ist als Matrix-Produkt nicht definiert, dies ist also stets als Skalarprodukt zu verstehen.

(2) $\vec{v}^T \cdot \vec{w} = (v_1 \ \dots \ v_n) \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = (v_1 w_1 + \dots + v_n w_n) = \vec{v} \cdot \vec{w}$.

$\vec{v}^T \cdot \vec{w}$ ergibt als Matrixprodukt also gerade das Skalarprodukt.

(3) $\vec{v} \cdot \vec{w}^T = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \cdot (w_1 \ \dots \ w_n) = \begin{pmatrix} v_1 w_1 & v_1 w_2 & \cdots & v_1 w_n \\ v_2 w_1 & v_2 w_2 & \cdots & v_2 w_n \\ \vdots & \vdots & \ddots & \vdots \\ v_n w_1 & v_n w_2 & \cdots & v_n w_n \end{pmatrix}$

$\vec{v} \cdot \vec{w}^T$ ergibt eine (n,n)-Matrix. Dieses Produkt wird auch als *dyadisches* oder *tensorielles Produkt* bezeichnet.

2.4 Die inverse Matrix

2.4.1 Definition Eine (n,n)-Matrix A ist *invertierbar*, wenn eine Matrix A^{-1} existiert mit

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

A^{-1} heißt dann *inverse Matrix*.

2.4.2 Bemerkung Ist A invertierbar, dann gilt $A^{-1} \cdot A = A \cdot A^{-1} = E$, was wegen der Nicht-Kommutativität des Matrixprodukts ja nicht selbstverständlich ist.

2.4.3 Beispiel Betrachten wir $A = \begin{pmatrix} 1 & 1 \\ 3 & 4 \end{pmatrix}$. Es gilt

$$\begin{pmatrix} 1 & 1 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 4 & -1 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E ,$$

also ist A invertierbar und es ist $A^{-1} = \begin{pmatrix} 4 & -1 \\ -3 & 1 \end{pmatrix}$ die Inverse zu A.

2.4.4 Satz Seien A, B invertierbare Matrizen. Es gilt

(1) $\left(A^{-1}\right)^{-1} = A$

(2) $E^{-1} = E$

(3) $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$ (sofern die Produkte definiert sind)

(4) $\left(A^{-1}\right)^T = \left(A^T\right)^{-1}$

Nachzurechnen, ob eine Matrix invers zu einer anderen ist, ist recht leicht. Wie stellt man aber fest, ob eine Matrix überhaupt eine Inverse hat und wie findet man die inverse Matrix, falls sie existiert ?

Wir benötigen hierzu einige Matrixumformungen:

2.4.5 Definition Die folgenden Operationen bilden die *elementaren Zeilenoperationen*:

(1) Multiplikation einer Zeile mit einer Konstanten $\lambda \neq 0$.

(2) Vertauschung zweier Zeilen.

(3) Addition eines Vielfachen einer Zeile zu einer anderen Zeile.

2.4.6 Bemerkung Statt der elementaren Zeilenoperationen können jeweils auch entsprechende Spaltenoperationen benutzt werden. Zeilen- und Spaltenoperationen dürfen jedoch nicht „gemischt“ werden.

Mit Hilfe dieser Operationen ist es nun möglich, festzustellen, ob eine Matrix eine Inverse besitzt:

2.4.7 Satz Eine Matrix A ist genau dann invertierbar, wenn A durch elementare Zeilenoperationen in die Einheitsmatrix übergeführt werden kann.

2.4.8 Beispiel

$$\begin{array}{c|l}
\begin{pmatrix} 2 & 7 \\ 1 & 4 \end{pmatrix} & (-2)\text{-fache 2. Zeile zu 1. Zeile addieren} \\
\begin{pmatrix} 0 & -1 \\ 1 & 4 \end{pmatrix} & 4\text{-fache 1. Zeile zu 2. Zeile addieren} \\
\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} & \text{1. und 2. Zeile vertauschen} \\
\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} & \text{2. Zeile mit } (-1) \text{ multiplizieren} \\
\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & = E
\end{array}$$

Man kann also $\begin{pmatrix} 2 & 7 \\ 1 & 4 \end{pmatrix}$ in die Einheitsmatrix überführen, die Matrix $\begin{pmatrix} 2 & 7 \\ 1 & 4 \end{pmatrix}$ ist demnach invertierbar.

Dieses Verfahren kann auch benutzt werden, um die Inverse zu berechnen: Dazu werden die gleichen elementaren Zeilenoperationen, die zur Einheitsmatrix führen, auf eine Einheitsmatrix angewandt.

2.4.9 Beispiele **(1)** Schauen wir uns die Umformungen aus Beispiel 2.4.8 noch einmal an und wenden die elementaren Zeilenoperationen auch auf eine Einheitsmatrix an:

$$\begin{array}{c|c|l}
\begin{pmatrix} 2 & 7 \\ 1 & 4 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & (-2)\text{-fache 2. Zeile zu 1. Zeile addieren} \\
\begin{pmatrix} 0 & -1 \\ 1 & 4 \end{pmatrix} & \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} & \text{4-fache 1. Zeile zu 2. Zeile addieren} \\
\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} & \begin{pmatrix} 1 & -2 \\ 4 & -7 \end{pmatrix} & \text{1. und 2. Zeile vertauschen} \\
\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} & \begin{pmatrix} 4 & -7 \\ 1 & -2 \end{pmatrix} & \text{2. Zeile mit } (-1) \text{ multiplizieren} \\
\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 4 & -7 \\ -1 & 2 \end{pmatrix} &
\end{array}$$

Es ist demnach $\begin{pmatrix} 2 & 7 \\ 1 & 4 \end{pmatrix}^{-1} = \begin{pmatrix} 4 & -7 \\ -1 & 2 \end{pmatrix}$, wie wir zur Sicherheit auch nachrechnen können:

$$\begin{pmatrix} 2 & 7 \\ 1 & 4 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 4 & -7 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} .$$

(2) Betrachten wir als zweites Beispiel eine $(3, 3)$-Matrix:

$$\begin{array}{c|c|l}
\begin{pmatrix} 3 & -1 & 1 \\ -15 & 6 & -5 \\ 5 & -2 & 2 \end{pmatrix} & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \text{5-fache 1. Zeile zu 2. Zeile addieren} \\
\begin{pmatrix} 3 & -1 & 1 \\ 0 & 1 & 0 \\ 5 & -2 & 2 \end{pmatrix} & \begin{pmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & (-2)\text{-fache 1. Zeile zu 3. Zeile addieren} \\
\begin{pmatrix} 3 & -1 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix} & \begin{pmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} & \text{1-fache 2. Zeile zu 1. Zeile addieren} \\
\begin{pmatrix} 3 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix} & \begin{pmatrix} 6 & 1 & 0 \\ 5 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} & \text{3-fache 3. Zeile zu 1. Zeile addieren} \\
\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 1 & 3 \\ 5 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} & \text{3. Zeile mit } (-1) \text{ multiplizieren} \\
\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 1 & 3 \\ 5 & 1 & 0 \\ 2 & 0 & -1 \end{pmatrix} & \text{1. und 3. Zeile vertauschen} \\
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & \begin{pmatrix} 2 & 0 & -1 \\ 5 & 1 & 0 \\ 0 & 1 & 3 \end{pmatrix} &
\end{array}$$

Folglich ist $\begin{pmatrix} 3 & -1 & 1 \\ -15 & 6 & -5 \\ 5 & -2 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & 0 & -1 \\ 5 & 1 & 0 \\ 0 & 1 & 3 \end{pmatrix}$.

Falls die Inverse einer Matrix existiert, können wir diese Inverse nun berechnen. Wir haben aber noch keine gute Möglichkeit, festzustellen, ob eine Matrix invertierbar ist. Ein recht einfaches Kriterium ist:

2.4.10 Satz Sei A eine (n, n)-Matrix. A ist genau dann invertierbar, wenn alle n Spaltenvektoren (oder alle n Zeilenvektoren) linear unabhängig sind.

2.4.11 Beispiel Betrachten wir

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

Sind $\vec{v}_1$, $\vec{v}_2$ und $\vec{v}_3$ die drei Spaltenvektoren dieser Matrix. Dann ist $\vec{v}_1 + \vec{v}_2 = \vec{v}_3$, die drei Vektoren sind also linear abhängig, A folglich nicht invertierbar.

2.5 Determinanten

Das einzige Kriterium, das wir bislang kennen, um zu entscheiden, ob eine Matrix überhaupt invertierbar ist, ist die lineare Unabhängigkeit der Spalten- oder Zeilenvektoren (Satz 2.4.10). Leider haben wir aber noch kein wirklich bequemes Verfahren kennengelernt, lineare Unabhängigkeit nachzuweisen.

Einen Schritt in diese Richtung werden wir nun unternehmen. Die *Determinante* einer Matrix ist eine reelle Zahl, die einer quadratischen Matrix zugeordnet wird und an der wir verschiedene Charakteristika der Matrix (z.B. deren Invertierbarkeit) erkennen können.

Wie berechnen wir diese Determinante ?

$(2, 2)$-Matrizen

Sei $A = (a_{ik})$ eine $(2, 2)$-Matrix. Die *Determinante* von A ist

$$\det A = |A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - cb .$$

2.5.1 Beispiel

$$\begin{vmatrix} 3 & 1 \\ -2 & 4 \end{vmatrix} = 12 - (-2) = 14$$

$(3, 3)$-Matrizen

Sei $A = (a_{ik})$ eine $(3, 3)$-Matrix. Die *Determinante* von A ist

$$\det A = |A| = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - gec - hfa - idb .$$

Man kann wohl nicht unbedingt erwarten, dass jemand sich dies merken möchte. Erfreulicherweise gibt es eine „Eselsbrücke“ für diese Formel, die *Sarrus-Regel*:

Danach werden die ersten beiden Spalten rechts der Matrix angefügt. Die drei Diagonalen von links oben nach rechts unten werden multipliziert und addiert, die drei Diagonalen von links unten nach rechts oben werden multipliziert und subtrahiert (Abbildung 11).

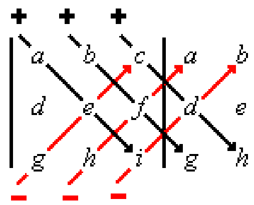


Abbildung 11: Regel von Sarrus

(n, n)-Matrizen

Bei größeren Matrizen gibt es leider keine Merkregel wie die Sarrus-Regel für die Berechnung der Determinanten. Statt dessen verwendet man einen anderen Weg:

Wir führen die Berechnung einer (n, n)-Determinante auf die Berechnung von $(n-1, n-1)$-Determinanten zurück. Diese werden dann auf $(n-2, n-2)$-Determinanten zurückgeführt und so weiter, bis wir bei $(3, 3)$- oder $(2, 2)$-Determinanten angelangt sind, die wir mit den obigen Formeln ausrechnen können.

Wir benötigen hierzu noch einen Begriff:

2.5.2 Definition Sei A eine (n, n)-Matrix. Die $(n-1, n-1)$-Matrix, die durch Weglassen der i-ten Zeile und der k-ten Spalte entsteht, heißt *Untermatrix* A_{ik}.

2.5.3 Beispiel Sei $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$. Dann ist $A_{11} = \begin{pmatrix} e & f \\ h & i \end{pmatrix}$, $A_{12} = \begin{pmatrix} d & f \\ g & i \end{pmatrix}$, usw.

Die Berechnung der Determinanten erfolgt nun über folgende Entwicklungen:

2.5.4 Satz Sei $A = (a_{ik})$ eine (n, n)-Matrix. Die *Determinante* von A ist mit einem $k \in \{1, \ldots, n\}$

$$\begin{aligned} \det A = |A| &= \sum_{i=1}^{n} (-1)^{i+k} a_{ik} |A_{ik}| \qquad (\textit{Entwicklung nach der } k\textit{-ten Spalte}) \\ &= (-1)^{1+k} \cdot a_{1k} \cdot |A_{1k}| + (-1)^{2+k} \cdot a_{2k} \cdot |A_{2k}| + \cdots + (-1)^{n+k} \cdot a_{nk} \cdot |A_{nk}| \end{aligned}$$

beziehungsweise mit einem $i \in \{1, \ldots, n\}$

$$\begin{aligned} \det A = |A| &= \sum_{k=1}^{n} (-1)^{i+k} a_{ik} |A_{ik}| \qquad (\textit{Entwicklung nach der } i\textit{-ten Zeile}) \\ &= (-1)^{i+1} \cdot a_{i1} \cdot |A_{i1}| + (-1)^{i+2} \cdot a_{i2} \cdot |A_{i2}| + \cdots + (-1)^{i+n} \cdot a_{in} \cdot |A_{in}| \,. \end{aligned}$$

In der Praxis geht man wie folgt vor:

(1) Wir versehen die Matrix mit einem „Vorzeichenschachbrett“ (links oben mit einem „+“ beginnend):

$$\begin{pmatrix} {}^{+}a_{11} & {}^{-}a_{12} & {}^{+}a_{13} & \cdots \\ {}^{-}a_{21} & {}^{+}a_{22} & {}^{-}a_{23} & \cdots \\ {}^{+}a_{31} & {}^{-}a_{32} & {}^{+}a_{33} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

(2) Wir suchen die Zeile oder Spalte mit den meisten Nullen.

(3) Wir entwickeln gem. Satz 2.5.4 nach dieser Zeile/Spalte. Das Vorzeichenschachbrett hilft uns dabei, das richtige Vorzeichen der jeweiligen Summanden zu finden.

(4) Wir wiederholen dieses Verfahren für jede im letzten Schritt aufgetretene (kleinere) Determinante, bis alle Determinanten (z.B. mit der Sarrus-Regel) berechnet werden können.

2.5.5 Beispiel Betrachten wir $A = \begin{pmatrix} 1 & 2 & 0 & 3 \\ -1 & 0 & 2 & 1 \\ 3 & 2 & 0 & 2 \\ 0 & 1 & 7 & 1 \end{pmatrix}$.

Mit Vorzeichenschachbrett versehen haben wir $\begin{pmatrix} {}^{+}1 & {}^{-}2 & {}^{+}0 & {}^{-}3 \\ {}^{-}-1 & {}^{+}0 & {}^{-}2 & {}^{+}1 \\ {}^{+}3 & {}^{-}2 & {}^{+}0 & {}^{-}2 \\ {}^{-}0 & {}^{+}1 & {}^{-}7 & {}^{+}1 \end{pmatrix}$.

Da die dritte Spalte zwei Nullen enthält, entwickeln wir nach dieser Spalte und erhalten (die Vorzeichen entnehmen wir dem Schachbrett)

$$
\begin{aligned}
\det A &= +0 \cdot |A_{13}| - 2 \cdot |A_{23}| + 0 \cdot |A_{33}| - 7 \cdot |A_{43}| \\
&= -2 \begin{vmatrix} 1 & 2 & 3 \\ 3 & 2 & 2 \\ 0 & 1 & 1 \end{vmatrix} - 7 \begin{vmatrix} 1 & 2 & 3 \\ -1 & 0 & 1 \\ 3 & 2 & 2 \end{vmatrix} \\
&= -2\,(2 + 0 + 9 - 0 - 2 - 6) - 7\,(0 + 6 + (-6) - 0 - 2 - (-4)) \\
&= -2 \cdot 3 - 7 \cdot 2 \\
&= -20\ .
\end{aligned}
$$

Wir könnten natürlich auch nach einer anderen Spalte oder Zeile entwicklen, dann würden aber weniger Summanden wegfallen und wir hätten mehr Arbeit. So brauchen wir nur zwei Determinanten mit der Sarrus-Regel berechnen.

2.5.6 Satz Seien A, B quadratische Matrizen und E eine Einheitsmatrix. Dann gilt:

(1) $\det(A^T) = \det A$

(2) $\det(A \cdot B) = \det A \cdot \det B$

(3) $\det E = 1$

(4) Ist A invertierbar, dann ist $\det(A^{-1}) = \dfrac{1}{\det A}$.

(5) Ist eine gesamte Zeile oder Spalte in A gleich Null, dann ist $\det A = 0$.

2.5.7 Bemerkung Im allgemeinen gilt *nicht* $\det(A + B) = \det A + \det B$.

Wir waren auf der Suche nach einer Technik, entscheiden zu können, ob eine Matrix invertierbar ist oder nicht. Tatsächlich erlauben Determinanten diese Unterscheidung:

2.5.8 Satz Eine quadratische Matrix A ist genau dann invertierbar, wenn $\det A \neq 0$.

Eine Folgerung aus diesem Ergebnis ist wegen Satz 2.4.10:

2.5.9 Satz Die Zeilenvektoren bzw. Spaltenvektoren einer Matrix A sind genau dann linear unabhängig, wenn $\det A \neq 0$.

Damit haben wir auch ein recht bequemes Verfahren gefunden, Vektoren auf ihre lineare Unabhängigkeit zu prüfen.

Bei einigen speziellen Matrizen lässt sich die Determinante sehr leicht und ohne Entwicklung berechnen:

2.5.10 Definition Eine quadratische Matrix $A = (a_{ik})$ ist eine ...

(1) *untere Dreiecksmatrix*, wenn $a_{ik} = 0$ für $i < k$,

(2) *obere Dreiecksmatrix*, wenn $a_{ik} = 0$ für $i > k$ und

(3) *Diagonalmatrix*, wenn $a_{ik} = 0$ für $i \neq k$.

2.5.11 Beispiele $\begin{pmatrix} 2 & 1 & 0 & 3 \\ 0 & -1 & 4 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ ist eine obere Dreiecksmatrix,

$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ ist eine Diagonalmatrix (und damit auch eine obere und eine untere Dreiecksmatrix).

2.5.12 Satz Sei $A = (a_{ik})$ eine (n, n)-Dreiecksmatrix. Dann gilt

$$\det A = a_{11}a_{22} \cdots a_{nn} \ .$$

Damit können wir nun Determinanten berechnen:

2.5.13 Beispiel

$$\begin{vmatrix} 2 & 1 & -4 & 0 \\ 0 & 3 & -1 & -1 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix} = 2 \cdot 3 \cdot 3 \cdot (-1) = -18$$

2.6 Basistransformation

Wie wir nach Definition 1.7.8 wissen, können Vektoren bezüglich verschiedener Basen dargestellt werden. Normalerweise verwenden wir die Basis aus Einheitsvektoren, die einem kartesischen Koordinatensystem entspricht. Dort entspricht die z.B. Norm eines Vektors auch dem, was wir uns unter seiner Länge vorstellen (vgl. Bemerkung 1.2.3).

In manchen Situationen ist es aber sinnvoll, auch andere Basen zu betrachten. Kristallgitter sind beispielsweise oft nicht rechtwinklig ausgerichtet. Hier ist dann ein entsprechendes schiefwinkliges Koordinatensystem zweckmäßiger.

Natürlich benötigen wir eine möglichst einfache Technik, zwischen den verschiedenen Basen zu wechseln, d.h. die Darstellung eines Vektors bezüglich einer Basis in die Darstellung bezüglich einer anderen Basis umzurechnen. Dies wird *Basistransformation* oder *Basiswechsel* genannt.

2.6.1 Definition Sei $E = \{\vec{e}_1, \ldots, \vec{e}_n\}$ die Basis des $\mathbb{R}^n$ aus Einheitsvektoren und sei $B = \{\vec{b}_1, \ldots, \vec{b}_n\}$ eine weitere Basis des $\mathbb{R}^n$.

Sei $\vec{b}_k = \begin{pmatrix} b_{1k} \\ b_{2k} \\ \vdots \\ b_{nk} \end{pmatrix}$ für $k = 1, \ldots, n$, dann ist

$$T_E^B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{pmatrix}$$

die *Transformationsmatrix* zum Basiswechsel von B nach E.

2.6.2 Bemerkung T_E^B ist die Matrix, deren Spalten gerade die B-Basisvektoren bilden:

$$T_E^B = \begin{pmatrix} \vec{b}_1 & \vec{b}_2 & \ldots & \vec{b}_n \end{pmatrix}$$

2.6.3 Beispiel Die folgenden beiden Vektoren sind linear unabhängig und bilden also eine Basis des $\mathbb{R}^2$: $B = \{\vec{b}_1, \vec{b}_2\}$ mit $\vec{b}_1 = \binom{-1}{0}$ und $\vec{b}_2 = \binom{-1}{1}$ Die Transformationsmatrix zum Wechsel von B nach E ist demnach

$$T_E^B = \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix} .$$

Eine solche Transformationsmatrix wird ihrem Namen tatsächlich gerecht:

2.6.4 Satz Sei $\vec{v} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}_B$ die Koordinatendarstellung von $\vec{v}$ bezüglich einer Basis B und sei $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$ die Darstellung dieses Vektor bezüglich der Basis aus Einheitsvektoren. Dann ist

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = T_E^B \cdot \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} .$$

2.6.5 Beispiel Betrachten wir wieder die Basis B aus Beispiel 2.6.3. Nehmen wir an, der Vektor $\vec{v}$ hat die Darstellung $\vec{v} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}_B$ bezüglich dieser Basis. Die Transformationsmatrix T_E^B haben wir bereits in Beispiel 2.6.3 berechnet. Damit hat $\vec{v}$ bezüglich der Matrix aus Einheitsvektoren die Darstellung

$$\vec{v} = T_E^B \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -3 \\ 2 \end{pmatrix} .$$

Kennen wir die Darstellung eines Vektors bezüglich einer Basis, ist es also sehr einfach, die Darstellung bezüglich der Basis aus Einheitsvektoren zu berechnen. Wie wechseln wir aber umgekehrt von der Basis aus Einheitsvektoren zu einer anderen Basis?

2.6.6 Satz Sei E die Basis des $\mathbb{R}^n$ aus Einheitsvektoren und sei B eine weitere Basis. Dann ist die Transformationsmatrix T_E^B invertierbar und die Transformationsmatrix zum Basiswechsel von E nach B ist

$$T_B^E = \left(T_E^B\right)^{-1} .$$

2.6.7 Beispiel Betrachten wir den Vektor $\vec{v} = \begin{pmatrix} -3 \\ 2 \end{pmatrix}$ bezüglich der Basis E und die Basis B aus den obigen Beispielen. Wir berechnen T_B^E durch Invertieren von $T_E^B = \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix}$:

$$\begin{array}{c|c|l}
\begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \text{1-fache 2. Zeile zu 1. Zeile addieren} \\
\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} & \text{1. Zeile mit } (-1) \text{ multiplizieren} \\
\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix} &
\end{array}$$

Also ist $T_B^E = \left(T_E^B\right)^{-1} = \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix}$ (die Matrix ist also zu sich selbst invers), wir können also die Darstellung von $\vec{v}$ bezüglich B bestimmen:

$$\vec{v} = \left(T_E^B\right)^{-1} \cdot \begin{pmatrix} -3 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} -3 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}_B .$$

Natürlich können wir nicht nur von B zu E oder von E nach B wechseln, auch Transformationen zwischen zwei schiefwinkligen Koordinatensystemen sind möglich: Wollen wir beispielsweise von einer Basis B zu einer Basis C wechseln, so wechseln wir zuerst von B nach E und anschließend von E nach C. Die entsprechende Transformationsmatrix erhalten wir durch Multiplikation:

$$T_C^B = T_C^E \cdot T_E^B = \left(T_E^C\right)^{-1} \cdot T_E^B .$$

2.6.8 Beispiel Betrachten wir die Basen $B = \{\vec{b}_1, \vec{b}_2, \vec{b}_3\}$ und $C = \{\vec{c}_1, \vec{c}_2, \vec{c}_3\}$ mit

$$\vec{b}_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} \quad , \quad \vec{b}_2 = \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} \quad , \quad \vec{b}_3 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$$

$$\vec{c}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad , \quad \vec{c}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \quad , \quad \vec{c}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} .$$

Wir haben also

$$T_E^B = \begin{pmatrix} 1 & 3 & 2 \\ 0 & 1 & 1 \\ 2 & 0 & 1 \end{pmatrix}$$

$$\text{und} \qquad T_C^E = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}^{-1} = \cdots = \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$$

(rechnen Sie das zur Übung einmal nach). Damit ist dann

$$T_C^B = \left(T_E^C\right)^{-1} \cdot T_E^B = \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & 1 \\ 1 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 3 & 2 \\ 0 & 1 & 1 \\ 2 & 0 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 3 & 2 & 2 \\ 1 & -2 & 0 \\ -1 & 4 & 2 \end{pmatrix} .$$

Ist $\vec{v} = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}_B$ bezüglich B, dann ist bezüglich C

$$\vec{v} = T_C^B \cdot \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 3 & 2 & 2 \\ 1 & -2 & 0 \\ -1 & 4 & 2 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 10 \\ 4 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \\ 0 \end{pmatrix}_C .$$

2.7 Der metrische Tensor

In der Praxis wird zur einfacheren Umrechnung bei Basistransformationen nicht nur die Transformationsmatrix benutzt.

Betrachten wir den $\mathbb{R}^3$ (da dies in Anwendungen die größte Rolle spielt) mit einer Basis $B = \{\vec{a}, \vec{b}, \vec{c}\}$. Die zugehörige Transformationsmatrix für den Wechsel von diesem Koordinatensystem in das kartesische nennen wir zur Abkürzung einfach M:

$$M = T_E^B = \begin{pmatrix} \vec{a} & \vec{b} & \vec{c} \end{pmatrix} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} .$$

Seien $\vec{v}_B$ und $\vec{w}_B$ die Darstellungen zweier Vektoren bezüglich B, dann erhalten wir die Darstellungen bezüglich E durch

$$\vec{v}_E = M \cdot \vec{v}_B \qquad \text{und} \qquad \vec{w}_E = M \cdot \vec{w}_B \ .$$

Schauen wir uns nun einmal das Skalarprodukt dieser beiden Vektoren an. Wie wir in Bemerkung 2.3.4 gesehen haben, ist das Skalarprodukt $\vec{v} \cdot \vec{w}$ nichts anderes als das Matrixprodukt $\vec{v}^T \cdot \vec{w}$.

Wir haben also unter Verwendung von Satz 2.3.3(2)

$$\vec{v}_E^T \cdot \vec{w}_E = (M\vec{v}_B)^T \cdot M\vec{w}_B = \vec{v}_B^T \cdot M^T M \cdot \vec{w}_B = \vec{v}_B^T \cdot G \cdot \vec{w}_B \ ,$$

wenn wir $G = M^T M$ abkürzen.

2.7.1 Satz Ist $M = T_E^B$ die Transformationsmatrix zum Basiswechsel von B nach E, dann heißt $G = M^T M$ *metrischer Tensor* und es gilt für das Skalarprodukt

$$\vec{v}_E \cdot \vec{w}_E = \vec{v}_B^T \cdot G \cdot \vec{w}_B \ .$$

Natürlich kann man G als Produkt $M^T M$ ausrechnen. Es geht aber auch einfacher:

$$\begin{aligned} G = M^T M &= \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \\ &= \begin{pmatrix} a_1a_1 + a_2a_2 + a_3a_3 & a_1b_1 + a_2b_2 + a_3b_3 & a_1c_1 + a_2c_2 + a_3c_3 \\ b_1a_1 + b_2a_2 + b_3a_3 & b_1b_1 + b_2b_2 + b_3b_3 & b_1c_1 + b_2c_2 + b_3c_3 \\ c_1a_1 + c_2a_2 + c_3a_3 & c_1b_1 + c_2b_2 + c_3b_3 & c_1c_1 + c_2c_2 + c_3c_3 \end{pmatrix} \\ &= \begin{pmatrix} \vec{a} \cdot \vec{a} & \vec{a} \cdot \vec{b} & \vec{a} \cdot \vec{c} \\ \vec{a} \cdot \vec{b} & \vec{b} \cdot \vec{b} & \vec{b} \cdot \vec{c} \\ \vec{a} \cdot \vec{c} & \vec{b} \cdot \vec{c} & \vec{c} \cdot \vec{c} \end{pmatrix} \end{aligned} \tag{4}$$

Der metrische Tensor G ist also also die Matrix der Skalarprodukte der Basisvektoren B.

Ist $\alpha = \sphericalangle(\vec{b}, \vec{c})$, $\beta = \sphericalangle(\vec{a}, \vec{c})$ und $\gamma = \sphericalangle(\vec{a}, \vec{b})$, können wir eine weitere Darstellung des metrischen Tensors finden.

Es ist z.B. $\vec{a} \cdot \vec{a} = \|\vec{a}\|^2$ (Satz 1.3.5) und nach der Winkeldefinition 1.4.1 ist beispielsweise $\vec{a} \cdot \vec{b} = \|\vec{a}\|\|\vec{b}\| \cos\gamma$. Aus (4) erhalten wir damit

$$G = \begin{pmatrix} \|a\|^2 & \|\vec{a}\|\|\vec{b}\| \cos\gamma & \|\vec{a}\|\|\vec{c}\| \cos\beta \\ \|\vec{a}\|\|\vec{b}\| \cos\gamma & \|b\|^2 & \|\vec{b}\|\|\vec{c}\| \cos\alpha \\ \|\vec{a}\|\|\vec{c}\| \cos\beta & \|\vec{b}\|\|\vec{c}\| \cos\alpha & \|c\|^2 \end{pmatrix}$$

Mit diesen Darstellungen des metrischen Tensors G können wir also das Skalarprodukt zweier Vektoren bezüglich der Basis aus Einheitsvektoren ausrechnen, ohne die Vektoren erst in diese Basis umrechnen zu müssen.

Auch das Volumen V des Spats, der von den drei Basisvektoren $\vec{a}$, $\vec{b}$ und $\vec{c}$ aufgespannt wird, kann mittels des metrischen Tensors berechnet werden.

Unter Verwendung der Sarrus-Regel ist

$$\begin{aligned}\vec{a}\cdot(\vec{b}\times\vec{c}) &= \begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix}\cdot\begin{pmatrix}b_2c_3-b_3c_2\\b_3c_1-b_1c_3\\b_1c_2-b_2c_1\end{pmatrix}\\ &= a_1b_2c_3-a_1b_3c_2+a_2b_3c_1-a_2b_1c_3+a_3b_1c_2-a_3b_2c_1\\ &= a_1b_2c_3+b_1c_2a_3+c_1a_2b_3-a_3b_2c_1-b_3c_2a_1-c_3a_2b_1\\ &= \begin{vmatrix}a_1&b_1&c_1\\a_2&b_2&c_2\\a_3&b_3&c_3\end{vmatrix}\\ &= \det M\ ,\end{aligned} \tag{5}$$

also $V=|\vec{a}\cdot(\vec{b}\times\vec{c})|=|\det M|$. Daraus folgt

$$\det G=\det\left(M^TM\right)=\det M^T\cdot\det M=(\det M)^2=V^2\ ,$$

wir haben also

$$V=\sqrt{\det G}\ .$$

Die Beziehung in (5) erlaubt uns auch eine geometrische Anschauung davon zu bekommen, was die Determinante ist, zumindest im $\mathbb{R}^3$:

Betrachten wir die Spalten einer Matrix als Vektoren, ist die Determinante gerade das Spatprodukt, entspricht also dem Volumen des aufgespannten Spats, wobei sich das Vorzeichen danach richtet, ob es sich um ein Links- oder Rechtssystem handelt.

2.8 Lineare Gleichungssysteme lösen

Wir erinnern uns: eigentlich wollten wir lineare Gleichungssysteme lösen und haben Matrizen als abkürzende Schreibweisen für diese Gleichungssysteme kennen gelernt. Benutzen wir jetzt unsere Kenntnisse über Matrizen, um uns wieder dem ursprünglichen Ziel zu widmen.

Ein lineares Gleichungssystem

$$\begin{aligned}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n&=y_1\\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n&=y_2\\ &\vdots\\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n&=y_m\ .\end{aligned}$$

mit m Gleichungen und n Unbekannten können wir auch als Matrixprodukt schreiben:

$$\begin{pmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&\ddots&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{pmatrix}\cdot\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix}=\begin{pmatrix}y_1\\\vdots\\y_m\end{pmatrix}$$

Ist A eine (m,n)-Matrix, $\vec{x} \in \mathbb{R}^n$ und $\vec{y} \in \mathbb{R}^m$, dann ist also

$$A\vec{x} = \vec{y}$$

ein lineares Gleichungssystem mit m Gleichungen und n Unbekannten.

2.8.1 Beispiel

$$\begin{aligned} & \begin{pmatrix} 2 & -1 \\ 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \\ \Leftrightarrow \quad & \begin{pmatrix} 2x_1 - x_2 \\ x_1 + 3x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \\ \Leftrightarrow \quad & 2x_1 - x_2 = 1 \\ \text{und} \quad & x_1 + 3x_2 = 2 \end{aligned}$$

2.8.2 Bezeichnungen **(1)** Ist $m = n$, dann heißt das Gleichungssystem *quadratisch* (A ist dann auch quadratisch).

(2) Ist $\vec{y} = \vec{0}$, also $y_1 = y_2 = \cdots y_m = 0$, nennt man das Gleichungssystem *homogen*, ansonsten *inhomogen*.

2.8.3 Beispiele **(1)**

$$\begin{aligned} x_1 - x_2 &= 0 \\ 2x_1 - 2x_2 &= 1 \end{aligned} \quad \Leftrightarrow \quad \begin{pmatrix} 1 & -1 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

ist inhomogen und quadratisch.

Schauen wir uns die erste Gleichung näher an: $x_1 - x_2 = 0 \;\Leftrightarrow\; 2x_1 - 2x_2 = 0$ und das steht in Widerspruch zur zweiten Gleichung. Das Gleichungssystem kann also *keine Lösung* haben.

(2)

$$\begin{aligned} x_1 - x_2 &= 0 \\ 2x_1 - 2x_2 &= 0 \end{aligned} \quad \Leftrightarrow \quad \begin{pmatrix} 1 & -1 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

ist homogen und quadratisch und hat, wie wir durch Einsetzen leicht nachrechnen können, *mehrere Lösungen*: $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \ldots$

Tatsächlich ist jeder Vektor $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ mit $x_1 = x_2$ eine Lösung des Systems.

Ein lineares Gleichungssystem muss also nicht eindeutig lösbar sein, es kann auch keine oder mehrere Lösungen geben. Bei homogenen Systemen ist eine Lösung aber offensichtlich immer zu finden:

2.8.4 Satz Ein homogenes Gleichungssystem $A\vec{x} = \vec{0}$ hat stets die Lösung $\vec{x} = \vec{0}$.
$x_1 = x_2 = \ldots = x_n = 0$ ist also immer (eine) Lösung des Gleichungssystems.

Betrachten wir zuerst quadratische Gleichungssysteme. In diesem Fall können wir die Lösbarkeit, sogar die Eindeutigkeit der Lösung, recht leicht entscheiden:

2.8.5 Satz Sei $A\vec{x} = \vec{y}$ ein quadratisches lineares Gleichungssystem. Das System ist genau dann eindeutig lösbar, wenn

$$\det A \neq 0 \;.$$

Das Gleichungssystem ist also genau dann eindeutig lösbar, wenn A invertierbar ist. Diese Beobachtung können wir nutzen, um die Lösung auch zu berechnen:

Sei A invertierbar, dann gilt

$$A\vec{x} = \vec{y} \quad \Leftrightarrow \quad A^{-1}A\vec{x} = A^{-1}\vec{y} \quad \Leftrightarrow \quad E\vec{x} = A^{-1}\vec{y} \quad \Leftrightarrow \quad \vec{x} = A^{-1}\vec{y}$$

Wir haben also:

2.8.6 Satz Sei A eine quadratische Matrix mit $\det A \neq 0$. Dann ist $\vec{x} = A^{-1}\vec{y}$ die eindeutige Lösung des Gleichungssystems $A\vec{x} = \vec{y}$.

2.8.7 Beispiel

$$\begin{aligned} 3x_1 - x_2 + x_3 &= 2 \\ -15x_1 + 6x_2 - 5x_3 &= -10 \\ 5x_1 - 2x_2 + 2x_3 &= 3 \end{aligned} \qquad \Leftrightarrow \qquad \begin{pmatrix} 3 & -1 & 1 \\ -15 & 6 & -5 \\ 5 & -2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 \\ -10 \\ 3 \end{pmatrix}$$

Auf Seite 20 hatten wir diese Matrix bereits invertiert: $\begin{pmatrix} 3 & -1 & 1 \\ -15 & 6 & -5 \\ 5 & -2 & 2 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & 0 & -1 \\ 5 & 1 & 0 \\ 0 & 1 & 3 \end{pmatrix}$

Es folgt $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2 & 0 & -1 \\ 5 & 1 & 0 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} 2 \\ -10 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, die (einzige) Lösung des Gleichungssystems ist also

$$x_1 = 1 \;, \quad x_2 = 0 \;\text{ und }\; x_3 = -1 \;.$$

2.8.8 Bemerkung Die Sätze 2.8.4 und 2.8.5 können wir natürlich kombinieren. Betrachten wir ein homogenes quadratisches Gleichungssystem $A\vec{x} = \vec{0}$ mit $\det A \neq 0$. Dann ist $\vec{x} = \vec{0}$ die einzige Lösung.

Ist allerdings $\det A = 0$, dann muss es außer $\vec{x} = \vec{0}$ noch weitere Lösungen geben. Da A in diesem Fall aber nicht invertierbar ist, können wir diese Lösungen nicht durch Invertieren berechnen.

2.9 Der Gauß-Algorithmus

Wir haben schon gesehen, dass Dreiecksmatrizen z.B. bei der Berechnung von Determinanten große Vorteile bieten. Das ist auch der Fall, wenn lineare Gleichungssysteme durch Dreiecksmatrizen beschrieben werden.

Bevor wir uns dies genauer anschauen, überlegen wir uns, dass wir jedes lineare Gleichungssystem auf diese Gestalt bringen können.

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

1. Schritt

Wenn $a_{11} = 0$, vertausche zwei Zeilen, so dass anschließend $a_{11} \neq 0$.

2. Schritt

- Addiere die mit $-\dfrac{a_{21}}{a_{11}}$ multiplizierte 1. Zeile zur zweiten.
- Addiere die mit $-\dfrac{a_{31}}{a_{11}}$ multiplizierte 1. Zeile zur dritten.
- …
- Addiere die mit $-\dfrac{a_{n1}}{a_{11}}$ multiplizierte 1. Zeile zur n-ten.

Wir erhalten eine neue Matrix. In der neuen Matrix sieht jedoch die erste Spalte schon so aus, wie wir das von einer (oberen) Dreiecksmatrix erwarten:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & * & \cdots & * \end{pmatrix}$$

3. Schritt

Wiederhole das Verfahren mit der Untermatrix A_{11} statt A. Die erste Zeile und erste Spalte bleibt unverändert.

Als nächstes bleiben die ersten beiden Spalten und Zeilen unverändert und wir wiederholen das Verfahren für die verbleibende $(n-2, n-2)$-Matrix und so weiter…

Nach genügend Wiederholungen dieser Schritte erhalten wir eine Matrix der Form

$$\begin{pmatrix} * & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & * \end{pmatrix}$$

Schauen wir uns zwei Beispiele an:

2.9.1 Beispiele (1)

$$\begin{pmatrix} 1 & 0 & 2 \\ 2 & -1 & 3 \\ -3 & 2 & 7 \end{pmatrix} \qquad \begin{aligned} & a_{11} = 1 \neq 0 \\ & \left(-\tfrac{2}{1}\right) \cdot (1,0,2) = (-2,0,-4) \\ & \left(-\tfrac{-3}{1}\right) \cdot (1,0,2) = (3,0,6) \end{aligned}$$

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & -1 \\ 0 & 2 & 13 \end{pmatrix} \qquad \begin{aligned} & a_{22} = -1 \neq 0 \\ & \left(-\tfrac{2}{-1}\right) \cdot (0,-1,-1) = (0,-2,-2) \end{aligned}$$

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & -1 & -1 \\ 0 & 0 & 11 \end{pmatrix}$$

(2)

$$\begin{pmatrix} 1 & 0 & 2 \\ 2 & 0 & 3 \\ -3 & 2 & 7 \end{pmatrix} \qquad \begin{aligned} &a_{11} = 1 \neq 0 \\ &\left(-\tfrac{2}{1}\right) \cdot (1,0,2) = (-2,0,-4)) \\ &\left(-\tfrac{-3}{1}\right) \cdot (1,0,2) = (3,0,6) \end{aligned}$$

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 0 & -1 \\ 0 & 2 & 13 \end{pmatrix} \qquad a_{22} = 0, \text{ also 2. und 3. Zeile vertauschen}$$

$$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 13 \\ 0 & 0 & -1 \end{pmatrix}$$

2.10 Lineare Gleichungssysteme und Gauß-Algorithmus

Betrachten wir das Gleichungssystem

$$A\vec{x} = \vec{y}$$

mit einer (m, n)-Matrix A, also mit n Unbekannten und m Gleichungen. Um das Gleichungssystem zu lösen, fassen wir A und $\vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$ zu einer Matrix zusammen:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & y_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & y_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & y_m \end{pmatrix}$$

Als nächstes verwenden wir den Gauß-Algorithmus, um diese Matrix, soweit das natürlich möglich ist, auf „Dreiecksform“ zu bringen. Eventuell entstehen dabei Zeilen, die vollständig gleich 0 sind. Diese Zeilen streichen wir ersatzlos (m wird dann um 1 verringert).

Je nachdem, wie viele Zeilen und Spalten wir haben, erhalten wir Matrizen der Form[2]

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1n} & \beta_1 \\ & \alpha_{22} & \alpha_{23} & \cdots & \alpha_{2n} & \beta_2 \\ & & \alpha_{33} & \cdots & \alpha_{3n} & \beta_3 \\ & & & \ddots & \vdots & \vdots \\ & & & & \alpha_{nn} & \beta_n \end{pmatrix}$$

oder

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1n} & \beta_1 \\ & \alpha_{22} & \alpha_{23} & \cdots & \alpha_{2n} & \beta_2 \\ & & \alpha_{33} & \cdots & \alpha_{3n} & \beta_3 \\ & & & \ddots & \vdots & \vdots \\ & & & & \alpha_{in} & \beta_i \\ & & & & & \vdots \\ & & & & & \beta_m \end{pmatrix}$$

[2]Einträge, die durch den Gauß-Algorithmus 0 sind, werden aus Gründen der Übersichtlichkeit oft weggelassen.

oder aber

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \cdots & \cdots & \cdots & \alpha_{1n} & \beta_1 \\ & \alpha_{22} & \alpha_{23} & \cdots & \cdots & \cdots & \alpha_{2n} & \beta_2 \\ & & \alpha_{33} & \cdots & \cdots & \cdots & \alpha_{3n} & \beta_3 \\ & & & \ddots & \vdots & & \vdots & \vdots \\ & & & & \alpha_{mm} & \cdots & \alpha_{mn} & \beta_m \end{pmatrix}.$$

Die neue Matrix stellt natürlich auch ein lineares Gleichungssystem dar. Die Lösung dieses neuen Systems ist auch eine Lösung des ursprünglichen Gleichungssystems!

Schauen wir uns die drei Fälle näher an:

(1)

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1n} & \beta_1 \\ & \alpha_{22} & \alpha_{23} & \cdots & \alpha_{2n} & \beta_2 \\ & & \alpha_{33} & \cdots & \alpha_{3n} & \beta_3 \\ & & & \ddots & \vdots & \vdots \\ & & & & \alpha_{nn} & \beta_n \end{pmatrix}$$

Dies entspricht dem Gleichungssystem

$$\begin{aligned} \alpha_{11}x_1 + \alpha_{12}x_2 + \alpha_{13}x_3 + \cdots + \alpha_{1n}x_n &= \beta_1 \\ \alpha_{22}x_2 + \alpha_{23}x_3 + \cdots + \alpha_{2n}x_n &= \beta_2 \\ &\vdots \\ \alpha_{nn}x_n &= \beta_n \end{aligned}$$

Das Gleichungssystem ist durch „rückwärts Einsetzen“ recht leicht zu lösen: aus der letzten Zeile ergibt sich

$$x_n = \frac{\beta_n}{\alpha_{nn}}$$

und dies können wir in die vorletzte Zeile einsetzen und so weiter, bis wir alle x_k ausgerechnet haben.

2.10.1 Beispiel

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

Notieren wir dies als Matrix und führen den Gauß-Algorithmus durch:

$$\begin{pmatrix} 1 & 0 & 2 & 0 \\ 2 & -1 & 2 & 4 \\ -3 & 2 & 2 & 0 \end{pmatrix}$$

$$\rightsquigarrow \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & -1 & -2 & 4 \\ 0 & 2 & 8 & 0 \end{pmatrix}$$

$$\rightsquigarrow \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & -1 & -2 & 4 \\ 0 & 0 & 4 & 8 \end{pmatrix} \qquad \text{also} \qquad \begin{aligned} x_1 + 2x_3 &= 0 \\ -x_2 - 2x_3 &= 4 \\ 4x_3 &= 8 \end{aligned}$$

Aus der letzten Zeile folgt $x_3 = 2$. Setzen wir dies in die vorletzte Zeile ein, ergibt sich

$$-x_2 - 4 = 4 \quad \Leftrightarrow \quad x_2 = -8$$

und in der ersten Zeile erhalten wir

$$x_1 + 4 = 0 \quad \Leftrightarrow \quad x_1 = -4 \ .$$

(2)

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1n} & \beta_1 \\ & \alpha_{22} & \alpha_{23} & \cdots & \alpha_{2n} & \beta_2 \\ & & \alpha_{33} & \cdots & \alpha_{3n} & \beta_3 \\ & & & \ddots & \vdots & \vdots \\ & & & & \alpha_{nn} & \beta_n \\ & & & & & \vdots \\ & & & & & \beta_m \end{pmatrix}$$

Schauen wir uns z.B. die letzte Zeile an: Die zugehörige Gleichung lautet $0 = \beta_m$.

Da aber Zeilen, die nur aus Nullen bestehen, gestrichen wurden, muss $\beta_m \neq 0$ sein. Wir haben also einen Widerspruch; das lineare Gleichungssystem hat *keine Lösung*.

2.10.2 Beispiel

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

Der Gauß-Algorithmus ergibt hier

$$\begin{pmatrix} 1 & 2 & 1 & -1 & 1 \\ -2 & -3 & 2 & -3 & -2 \\ 3 & -1 & -2 & 3 & 4 \\ 2 & -2 & 1 & -1 & 1 \end{pmatrix}$$

$$\rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & -1 & 1 \\ 0 & 1 & 4 & -5 & 0 \\ 0 & -7 & -5 & 6 & 1 \\ 0 & -6 & -1 & 1 & -1 \end{pmatrix}$$

$$\rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & -1 & 1 \\ 0 & 1 & 4 & -5 & 0 \\ 0 & 0 & 23 & -29 & 1 \\ 0 & 0 & 23 & -29 & -1 \end{pmatrix}$$

$$\rightsquigarrow \begin{pmatrix} 1 & 2 & 1 & -1 & 1 \\ 0 & 1 & 4 & -5 & 0 \\ 0 & 0 & 23 & -29 & 1 \\ 0 & 0 & 0 & 0 & -2 \end{pmatrix}$$

Die letzte Zeile besagt $0 = -2$, das Gleichungssystem ist also nicht lösbar.

(3)

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} & \cdots & \cdots & \cdots & \alpha_{1n} & \beta_1 \\ & \alpha_{22} & \alpha_{23} & \cdots & \cdots & \cdots & \alpha_{2n} & \beta_2 \\ & & \alpha_{33} & \cdots & \cdots & \cdots & \alpha_{3n} & \beta_3 \\ & & & \ddots & \vdots & & \vdots & \vdots \\ & & & & \alpha_{mm} & \cdots & \alpha_{mn} & \beta_m \end{pmatrix}$$

mit dem zugehörigen linearen Gleichungssystem

$$\begin{aligned} \alpha_{11}x_1 + \alpha_{12}x_2 + \cdots + \alpha_{1n}x_n &= \beta_1 \\ \alpha_{22}x_2 + \cdots + \alpha_{2n}x_n &= \beta_2 \\ &\vdots \\ \alpha_{mm}x_m + \alpha_{m,m+1}x_{m+1} \cdots + \alpha_{mn}x_n &= \beta_m \end{aligned}$$

Die letzte Zeile ist äquivalent zu

$$x_m = \beta_m - \frac{\alpha_{m,m+1}x_{m+1} + \cdots + \alpha_{mn}x_n}{\alpha_{mm}} \ .$$

Dieses System hat also keine eindeutige Lösung, sondern mehrere:

Für jede Wahl von $x_{m+1}, x_{m+2}, \ldots, x_n \in \mathbb{R}$, können wir x_m und dann durch „rückwärts Einsetzen" $x_{m-1}, \ldots, x_1$ berechnen.

2.10.3 Beispiel

$$\begin{aligned} x_1 - 2x_2 + x_3 - x_4 &= 2 \\ -2x_1 + 3x_2 - x_3 + 3x_4 &= 1 \\ x_1 - x_2 - 2x_4 &= -3 \end{aligned}$$

Das Gauß-Verfahren liefert

$$\begin{aligned} &\begin{pmatrix} 1 & -2 & 1 & -1 & 2 \\ -2 & 3 & -1 & 3 & 1 \\ 1 & -1 & 0 & -2 & -3 \end{pmatrix} \\ \rightsquigarrow &\begin{pmatrix} 1 & -2 & 1 & -1 & 2 \\ 0 & -1 & 1 & 1 & 5 \\ 0 & 1 & -1 & -1 & -5 \end{pmatrix} \\ \rightsquigarrow &\begin{pmatrix} 1 & -2 & 1 & -1 & 2 \\ 0 & -1 & 1 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\ \rightsquigarrow &\begin{pmatrix} 1 & -2 & 1 & -1 & 2 \\ 0 & -1 & 1 & 1 & 5 \end{pmatrix} \qquad \text{also} \qquad \begin{aligned} x_1 - 2x_2 + x_3 - x_4 &= 2 \\ -x_2 + x_3 - x_4 &= 5 \end{aligned} \end{aligned}$$

Aus der letzten Zeile folgt für beliebige $x_3, x_4 \in \mathbb{R}$

$$x_2 = x_3 + x_4 - 5$$

und daraus und aus der ersten Zeile

$$\begin{aligned} x_1 &= 2 + 2x_2 - x_3 + x_4 \\ &= 2 + 2(x_3 + x_4 - 5) - x_3 + x_4 \\ &= x_3 + 3x_4 - 8 \ . \end{aligned}$$

Wir haben hier also unendlich viele Lösungen, nämlich jeweils eine für jede beliebige Wahl von x_3 und x_4. Beispielsweise erhalten wir für $x_3 = x_4 = 0$ die Lösung

$$x_1 = -8\ ,\ x_2 = -5\ ,\ x_3 = 0\ ,\ x_4 = 0$$

und für $x_3 = 1$ und $x_4 = 0$ erhalten wir

$$x_1 = -7\ ,\ x_2 = -4\ ,\ x_3 = 1\ ,\ x_4 = 0\ .$$

2.11 Eigenwerte und Eigenvektoren

Bei vielen naturwissenschaftlichen Effekten, z.B. bei der Untersuchung von Resonanzfrequenzen, in der Quantenphysik oder der Biologie tritt folgendes Problem auf:

Sei A eine quadratische Matrix. Für welchen Skalar $\lambda \in \mathbb{R}$ gilt

$$A\vec{x} = \lambda\vec{x} \tag{6}$$

für einen Vektor $\vec{x} \in \mathbb{R}^n$?

Natürlich gilt (6) immer für $\vec{x} = \vec{0}$ (die sog. *triviale Lösung*). Da das zu einfach ist, interessieren wir uns nur für Lösungen $\vec{x} \neq \vec{0}$:

2.11.1 Definition Sei A eine (n, n)-Matrix. Eine reelle Zahl λ, zu der ein Vektor $\vec{x} \in \mathbb{R}^n$ mit $\vec{x} \neq \vec{0}$ mit

$$A\vec{x} = \lambda\vec{x}$$

existiert, heißt *Eigenwert* von A. $\vec{x}$ nennt man dann den zugehörigen *Eigenvektor*.

2.11.2 Beispiel Sei $A = \begin{pmatrix} -1 & 1 \\ -4 & 4 \end{pmatrix}$. Dann gilt

$$\begin{pmatrix} -1 & 1 \\ -4 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 3 \\ 12 \end{pmatrix} \qquad \text{und} \qquad 3 \cdot \begin{pmatrix} 1 \\ 4 \end{pmatrix} = \begin{pmatrix} 3 \\ 12 \end{pmatrix}\ .$$

$\lambda = 3$ ist also Eigenwert zu A und $\binom{1}{4}$ der zu A und dem Eigenwert 3 gehörige Eigenvektor.

Es ist aber auch

$$\begin{pmatrix} -1 & 1 \\ -4 & 4 \end{pmatrix} \begin{pmatrix} 2 \\ 8 \end{pmatrix} = \begin{pmatrix} 6 \\ 24 \end{pmatrix} \qquad \text{und} \qquad 3 \cdot \begin{pmatrix} 2 \\ 8 \end{pmatrix} = \begin{pmatrix} 6 \\ 24 \end{pmatrix}\ .$$

Folglich ist auch $\binom{2}{8}$ Eigenvektor zum Eigenwert 3.

Wir sehen, dass zu einem Eigenwert einer Matrix mehrere Eigenvektoren existieren können. Tatsächlich gilt:

2.11.3 Satz Ist $\vec{x}$ Eigenvektor zum Eigenwert λ einer quadratischen Matrix A, dann sind auch alle Vielfachen $a\vec{x}$ mit $a \in \mathbb{R}$ und $a \neq 0$ Eigenvektoren zum Eigenwert λ.

Natürlich benötigen wir einen Weg, Eigenwerte und -vektoren zu finden. Es gilt

2.11.4 Satz Sei A eine quadratische Matrix. Die Eigenwerte von A sind genau die Nullstellen des *charakteristischen Polynoms*

$$P(\lambda) = \det(A - \lambda E)\ .$$

2.11.5 Beispiel Sei $A = \begin{pmatrix} 4 & -5 \\ 2 & -3 \end{pmatrix}$.

Das charakteristische Polynom hierzu ist

$$\begin{aligned} P(\lambda) &= \det\left(\begin{pmatrix} 4 & -5 \\ 2 & -3 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) \\ &= \det \begin{pmatrix} 4-\lambda & -5 \\ 2 & -3-\lambda \end{pmatrix} \\ &= (4-\lambda)(-3-\lambda) + 10 \\ &= \lambda^2 - \lambda - 2\ . \end{aligned}$$

Der p-q-Formel zufolge sind die Nullstellen dieses Polynoms

$$\lambda = \frac{1}{2} \pm \sqrt{\frac{1}{4} + 2} = \frac{1}{2} \pm \sqrt{\frac{9}{4}} = \frac{1}{2} \pm \frac{3}{2} = \begin{cases} 2 \\ -1 \end{cases} \quad ,$$

die Eigenwerte von A sind also 2 und -1.

Um auch die zugehörigen Eigenvektoren zu berechnen, betrachten wir Definition 2.11.1:

$$A\vec{x} = \lambda\vec{x} \qquad \Leftrightarrow \qquad A\vec{x} - \lambda\vec{x} = \vec{0} \qquad \Leftrightarrow \qquad (A - \lambda E) \cdot \vec{x} = \vec{0}$$

Da $(A - \lambda E)$ eine quadratische Matrix ist, ist dies ein homogenes, quadratisches, lineares Gleichungssystem, das wir mit den bekannten Methoden lösen können.

In unserem Beispiel lautet dieses Gleichungssystem

$$\begin{pmatrix} 4-\lambda & -5 \\ 2 & -3-\lambda \end{pmatrix} \cdot \vec{x} = \vec{0}\ .$$

(1) Für $\lambda = 2$ haben wir also[3]

$$\begin{pmatrix} 2 & -5 \\ 2 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \text{also} \qquad \begin{array}{rcl} 2x_1 - 5x_2 & = & 0 \\ 2x_1 - 5x_2 & = & 0 \end{array} \qquad \Rightarrow \qquad x_1 = \frac{5}{2}x_2$$

Wir wissen ja, dass Eigenvektoren nicht eindeutig sind, es kann uns also nicht überraschen, dass auch dieses Gleichungssystem mehrere Lösungen hat. Alle Vektoren der Form

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{5}{2}x_2 \\ x_2 \end{pmatrix} = x_2 \cdot \begin{pmatrix} \frac{5}{2} \\ 1 \end{pmatrix}$$

mit $x_2 \in \mathbb{R} \setminus \{0\}$ sind Eigenvektoren zum Eigenwert 2, z.B. also $\binom{5/2}{1}$ oder $\binom{5}{2}$.

(2) Für den zweiten Eigenwert $\lambda = -1$ ergibt sich

$$\begin{pmatrix} 5 & -5 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \text{also} \qquad \begin{array}{rcl} 5x_1 - 5x_2 & = & 0 \\ 2x_1 - 2x_2 & = & 0 \end{array} \qquad \Rightarrow \qquad x_1 = x_2$$

Alle Vektoren $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_2 \end{pmatrix} = x_2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ mit $x_2 \in \mathbb{R} \setminus \{0\}$ sind also Eigenvektoren zu $\lambda = -1$.

Besonders einfach ist die Ermittlung der Eigenwerte bei Dreiecksmatrizen. Es gilt nämlich:

2.11.6 Satz Sei $A = (a_{ik})$ eine Dreiecksmatrix. Dann sind die Einträge a_{ii} auf der Diagonalen von links oben nach rechts unten die Eigenwerte von A.

[3] der Gauß-Algorithmus würde natürlich das gleiche Ergebniss liefern, wäre aber aufwändiger

Funktionen mehrerer Variablen

3 Funktionen im $\mathbb{R}^n$

3.1 Bezeichnungen und Beispiele

Naturwissenschaftliche Beobachtungen werden oft mathematisch durch Funktionen beschrieben. Ein Beispiel mag das bekannte Newtonsche Gravitationsgesetz sein: Sind m_1 und m_2 zwei Massen und ist r der Abstand dieser Massen, so bezeichnet $F(m_1, m_2, r) = G \cdot \frac{m_1 m_2}{r^2}$ die Kraft, die zwischen diesen Massen wirkt (G ist eine geeignete Konstante). $F(m_1, m_2, r)$ ist also eine Funktion, die drei reelle Werte auf einen reellen Wert abbildet. Dies kann auch als Abbildung eines dreidimensionalen Vektor aufgefasst werden.

Solche Funktionen müssen keine reelle Zahl ergeben, das Ergebnis könnte auch ein mehrdimensionaler Vektor sein.

3.1.1 Definition Sei $D \subset \mathbb{R}^n$ eine Teilmenge des $\mathbb{R}^n$. Eine *Funktion* $f : D \to \mathbb{R}^m$ ordnet jedem Vektor $\vec{x} \in D$ einen Vektor $\vec{y} = f(\vec{x}) \in \mathbb{R}^m$ zu.

D heißt *Definitionsbereich* von f.

Die Menge $W(f) = \{\vec{y} \in \mathbb{R}^m : \vec{y} = f(\vec{x})$ für ein $\vec{x} \in D\}$ ist der *Wertebereich* (das *Bild*) von f.

Vor allem aus Gründen der Bequemlichkeit schreibt man beim Umgang mit Funktionen Vektoren auch als Zeilenvektoren, also z.B. (x, y) statt $\binom{x}{y}$.

3.1.2 Beispiele **(1)** $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x, y) = x^2 - y^2$

Diese Funktion wird wegen ihres Graphen auch als *Sattelfläche* (Abbildung 12) bezeichnet.

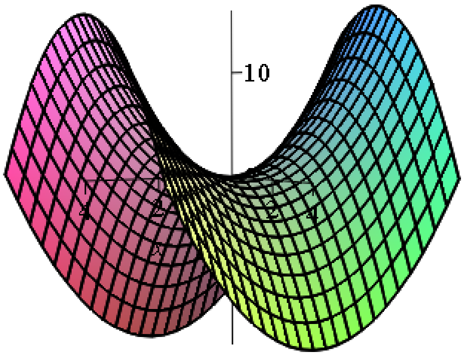


Abbildung 12: $f(x, y) = x^2 - y^2$

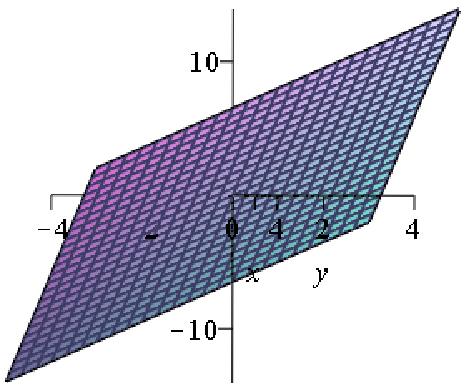


Abbildung 13: $f(x, y) = 2x + \frac{3}{2}y$

(2) $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x, y) = 2x + \dfrac{3}{2}y$

Der Graph einer linearen Funktion wie dieser stellt eine Ebene dar (Abbildung 13).

(3) $f : D \to \mathbb{R}$ mit $D = \left\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\right\}$ und $f(x, y) = \sqrt{1 - x^2 - y^2}$

Diese Funktion ist offenbar für $x^2 + y^2 > 1$ nicht definiert. Innerhalb des Definitionsbereichs stellt der Graph eine Halbkugel dar (Abbildung 14).

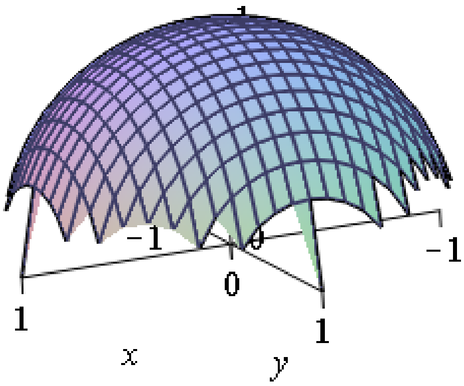


Abbildung 14: $f(x, y) = \sqrt{1 - x^2 - y^2}$

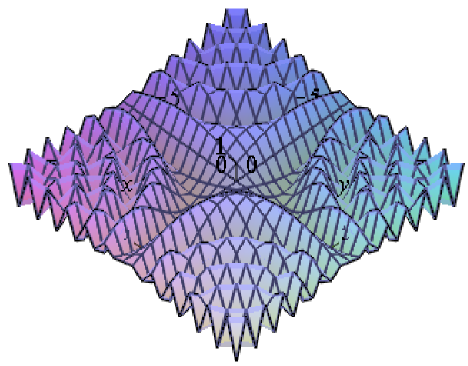


Abbildung 15: $f(x, y) = \sin(xy)$

(4) $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x, y) = \sin(xy)$ (Abbildung 15)

(5) $f : \mathbb{R}^3 \to \mathbb{R}^2$ mit $f(x, y, z) = (4x - 2y^2, z^2 - xy)$

Der Graph dieser Funktion ist 5-dimensional, also nicht darstellbar.

(6) $f : \mathbb{R} \to \mathbb{R}^2$ mit $f(t) = (\sin t, t)$

Während die Graphen von Funktionen $\mathbb{R}^2 \to \mathbb{R}$ „Berg- und Tallandschaften“ ähneln, sind Graphen von Funktionen $\mathbb{R} \to \mathbb{R}^2$ Kurven im Raum. Der Graph dieser Funktion ist eine schräg im Raum liegende Sinuskurve (Abbildung 16).

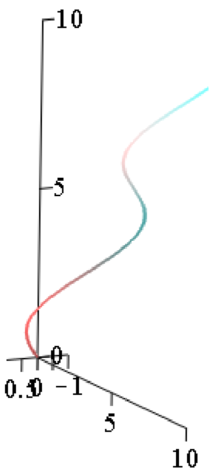


Abbildung 16: $f(t) = (\sin t, t)$

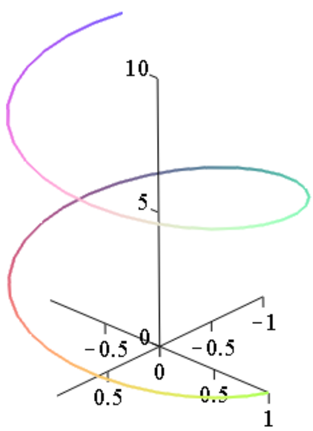


Abbildung 17: $f(t) = (\sin t, \cos t)$

(7) $f : \mathbb{R} \to \mathbb{R}^2$ mit $f(t) = (\sin t, \cos t)$ Hier ist der Graph eine Schraubenlinie (Abbildung 17).

3.2 Stetigkeit

Die Stetigkeit von Funktionen $f : \mathbb{R} \to \mathbb{R}$ kennen wir bereits aus Teil I. Dort hatten wir zuerst Folgen betrachtet, um den Begriff der Stetigkeit definieren zu können. So gehen wir auch bei mehrdimensionalen Funktionen vor:

Seien $(x_k)_k = x_1, x_2, \ldots$ und $(y_k)_k = y_1, y_2, \ldots$ zwei Folgen in den reellen Zahlen.

$\binom{x_k}{y_k} = \binom{x_1}{y_1}, \binom{x_2}{y_2}, \ldots$ bildet dann eine Folge im $\mathbb{R}^2$ und dies lässt sich natürlich auch in höheren Dimensionen fortführen: Im $\mathbb{R}^n$ betrachten wir Folgen

$$(\vec{x}_k)_k = \vec{x}_1, \vec{x}_2, \vec{x}_3, \ldots = \begin{pmatrix} x_{11} \\ \vdots \\ x_{1n} \end{pmatrix}, \begin{pmatrix} x_{21} \\ \vdots \\ x_{2n} \end{pmatrix}, \begin{pmatrix} x_{31} \\ \vdots \\ x_{3n} \end{pmatrix}, \ldots .$$

Für $i = 1, \ldots, n$ stellt also jedes $(x_{ki})_k = x_{1i}, x_{2i}, \ldots$ eine Folge in $\mathbb{R}$ dar.[4]

3.2.1 Beispiel

$$(\vec{x})_k = \begin{pmatrix} \frac{1}{k} \\ \frac{1}{k^2} \end{pmatrix}_k = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{1}{2} \\ \frac{1}{4} \end{pmatrix}, \begin{pmatrix} \frac{1}{3} \\ \frac{1}{9} \end{pmatrix}, \ldots$$

ist eine Folge im $\mathbb{R}^2$ mit $(x_{k1})_k = \left(\frac{1}{k}\right)_k$ und $(x_{k2})_k = \left(\frac{1}{k^2}\right)_k$.

Wenn wir eine Folge im $\mathbb{R}^n$ als n Folgen in $\mathbb{R}$ betrachten, fällt es uns leicht, die Konvergenz mehrdimensionaler Folgen zu definieren:

3.2.2 Definition Eine Folge $(\vec{x}_k)_k = \left(\begin{pmatrix} x_{k1} \\ \vdots \\ x_{kn} \end{pmatrix}\right)_k$ des $\mathbb{R}^n$ ist *konvergent*, wenn alle Komponentenfolgen $(x_{ki})_k$ für alle $i = 1, \ldots, n$ konvergieren.

3.2.3 Beispiele **(1)** $\lim\limits_{k\to\infty} \begin{pmatrix} \frac{1}{k} \\ \frac{1}{k^2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \vec{0}$

(2) $\begin{pmatrix} 1 - \frac{1}{k} \\ k^2 \end{pmatrix}_k$ konvergiert nicht: $(1 - \frac{1}{k})_k$ konvergiert zwar, aber $(k^2)_k$ ist divergent.

Nun können wir uns überlegen, wie man Stetigkeit auch bei Funktionen $f : \mathbb{R}^n \to \mathbb{R}^m$ definieren kann.

Erinnern wir uns zuerst an die Situation in den reellen Zahlen: Eine Funktion f *konvergiert* in $a \in \mathbb{R}$ gegen einen Grenzwert $g \in \mathbb{R}$ (d.h. $\lim\limits_{x\to a} f(x) = g$), wenn für jede Folge (a_k) mit $\lim\limits_{k\to\infty} a_k = a$ gilt: $\lim\limits_{k\to\infty} f(a_k) = g$.

f ist *stetig* in a, wenn $\lim\limits_{x\to a} f(x) = f(a)$.

Ganz analog definieren wir nun:

3.2.4 Definition Sei $D \subset \mathbb{R}^n$. Eine Funktion $f : D \to \mathbb{R}^m$ *konvergiert* in $\vec{a} \in D$ gegen $\vec{g} \in \mathbb{R}^m$, wenn für jede Folge $(\vec{a}_k)_k$ mit $\lim\limits_{k\to\infty} \vec{a}_k = \vec{a}$ gilt: $\lim\limits_{k\to\infty} f(\vec{a}_k) = \vec{g}$.

f ist *stetig* in $\vec{a}$, wenn f in $\vec{a}$ gegen $f(\vec{a})$ konvergiert, also, wenn

$$\lim_{k\to\infty} f(\vec{a}_k) = f(\vec{a}) \qquad \text{für jede Folge } (\vec{a}_k)_k \text{ mit } \lim_{k\to\infty} \vec{a}_k = \vec{a} \,.$$

[4] Auch, wenn die Schreibweise x_{ki} an Komponenten von Matrizen erinnert, handelt es sich hier um keine Matrix: $k = 1, 2, 3, \ldots$ läuft in Folgen „bis ins Unendliche“.

Wie wir es schon von Funktionen $f : \mathbb{R} \to \mathbb{R}$ kennen, sind aus stetigen Funktionen zusammengesetzte Funktionen wiederum stetig:

3.2.5 Satz Sind $f, g : \mathbb{R}^n \to \mathbb{R}^m$ und $h : \mathbb{R}^m \to \mathbb{R}^l$ stetig, so sind auch $h(f(\vec{x}))$ und $f(\vec{x}) \pm g(\vec{x})$ stetig. Sind $f, g : \mathbb{R}^n \to \mathbb{R}$ stetig, dann ist auch $f(\vec{x}) \cdot g(\vec{x})$ und, falls $g(\vec{x}) \neq 0$, $\dfrac{f(\vec{x})}{g(\vec{x})}$ stetig.

3.2.6 Beispiele

$$f(x,y,z) = \frac{x}{z^2} - x^2yz^3 + 2 \qquad \text{und} \qquad f(x,y) = e^{x-y}$$

sind (für $z \neq 0$) stetig, da sich diese Funktionen aus stetigen Funktionen zusammensetzen.

Für $f : \mathbb{R}^n \to \mathbb{R}^m$ ist $f(\vec{x}) \in \mathbb{R}^m$ und daher können wir den Funktionswert auch als Vektor schreiben, dessen Komponenten Funktionen $f_i : \mathbb{R}^n \to \mathbb{R}$ sind:

$$f(\vec{x}) = \begin{pmatrix} f_1(\vec{x}) \\ f_2(\vec{x}) \\ \vdots \\ f_m(\vec{x}) \end{pmatrix} .$$

3.2.7 Satz Sei $D \subset \mathbb{R}^n$. $f : D \to \mathbb{R}^m$ ist genau dann stetig, wenn alle *Komponentenfunktionen* $f_i(\vec{x}) : D \to \mathbb{R}$ für $i = 1, \ldots, n$ stetig sind.

3.2.8 Beispiele **(1)** $f(t) = (\sin t, \cos t)$ ist stetig, da $f_1(t) = \sin t$ und $f_2(t) = \cos t$ stetig sind.

(2) $f(x,y) = \left(2x^2 - y, \sqrt{y}, \dfrac{1}{x-y}\right)$ ist stetig für $x \neq y$ und $y \geq 0$, da $f_1(x,y) = 2x^2 - y$, $f_2(x,y) = \sqrt{y}$ und für $x \neq y$ auch $f_3(x,y) = \frac{1}{x-y}$ stetig sind. Für $x = y$ oder $y < 0$ ist die Funktion nicht definiert, also weder stetig noch unstetig.

(3)

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{für } \binom{x}{y} \neq \vec{0} \\ 0 & \text{für } \binom{x}{y} = \vec{0} \end{cases}$$

ist stetig für $\binom{x}{y} \neq \vec{0}$, da sich die Funktion dort aus stetigen Funktionen zusammensetzt.

Im Punkt $\binom{x}{y} = \vec{0}$ mag man vermuten (vgl. Abbildung 18), dass die Funktion dort unstetig ist. Um dies zu zeigen, müssen wir wegen Definition 3.2.4 nachweisen, dass es eine Folge $\begin{pmatrix} x_k \\ y_k \end{pmatrix}$ mit $\lim\limits_{k\to\infty} \begin{pmatrix} x_k \\ y_k \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ gibt, für die gilt:

$$\lim_{k\to\infty} f(x_k, y_k) \neq f(0,0) .$$

Betrachten wir dazu $\begin{pmatrix} x_k \\ y_k \end{pmatrix} = \begin{pmatrix} \frac{1}{k} \\ \frac{1}{k} \end{pmatrix} \overset{k\to\infty}{\longrightarrow} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Es gilt

$$f(x_k, y_k) = f\left(\frac{1}{k}, \frac{1}{k}\right) = \frac{\frac{1}{k} \cdot \frac{1}{k}}{\frac{1}{k^2} + \frac{1}{k^2}} = \frac{1}{2} \overset{k\to\infty}{\longrightarrow} \frac{1}{2} \neq 0 = f(0,0) ,$$

f ist also unstetig in $\vec{0}$.

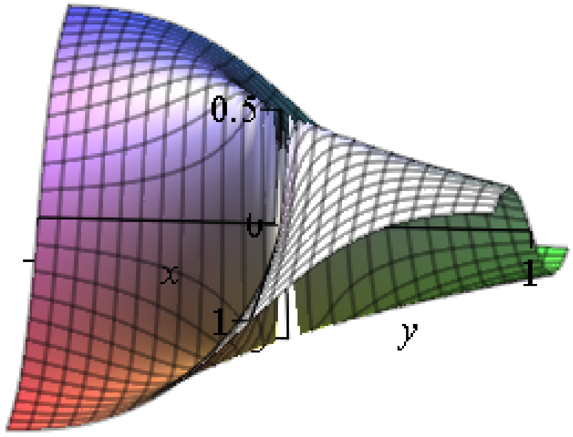


Abbildung 18: $f(x,y) = \frac{xy}{x^2+y^2}$ für $\binom{x}{y} \neq \vec{0}$

3.3 Differenzierbarkeit

Zur Erinnerung: Die Ableitung einer reellen Funktion $f : \mathbb{R} \to \mathbb{R}$ an einer Stelle $x_0 \in \mathbb{R}$ ist

$$\begin{aligned}\frac{d}{dx} f(x_0) = f'(x_0) &= \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \\ &= \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} \qquad \text{mit } h = x - x_0. \end{aligned} \tag{7}$$

Analog können wir bei Funktionen mehrerer Variablen verfahren: Betrachten wir beispielsweise eine Funktion $f : \mathbb{R}^2 \to \mathbb{R}$, dann können wir diese Funktion sowohl nach der ersten als auch nach der zweiten Komponente von $\binom{x}{y}$ ableiten. Wir erhalten

$$\begin{aligned}\frac{\partial}{\partial x} f(x_0, y_0) &= \lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} \ , \\ \frac{\partial}{\partial y} f(x_0, y_0) &= \lim_{h \to 0} \frac{f(x_0, y_0 + h) - f(x_0, y_0)}{h} \ .\end{aligned}$$

Außer der etwas anderen Schreibweise „$\frac{\partial}{\partial x}$“ statt „$\frac{d}{dx}$“, entspricht dies gerade der eindimensionalen Differenzierbarkeit in (7).

Allgemein:

3.3.1 Definition Eine Funktion $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$, heißt in $\vec{x} \in D$ *partiell differenzierbar nach* x_k, wenn der Grenzwert

$$\frac{\partial}{\partial x_k} f(\vec{x}) = f_{x_k}(\vec{x}) = \lim_{h \to 0} \frac{f(x_1, \ldots, x_{k-1}, x_k + h, x_{k+1}, \ldots, x_n) - f(x_1, \ldots, x_n)}{h}$$

existiert.

Partielle Ableitungen höherer Ordnung erhält man durch wiederholte partielle Differenziation. Ist f nach allen $x_1, \ldots, x_n$ partiell differenzierbar, dann heißt f *partiell differenzierbar* in $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.

3.3.2 Beispiele **(1)** $f(x,y) = 2x^2 + xy - 2y$

Es ist $\frac{\partial}{\partial x} f(x,y) = f_x(x,y) = 4x + y$ und $\frac{\partial}{\partial y} f(x,y) = f_y(x,y) = x - 2$.

Die höheren partiellen Ableitungen sind

$$\begin{aligned}
\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x} f(x,y)\right) &= \frac{\partial^2}{\partial x \partial x} f(x,y) = f_{xx}(x,y) = 4 \\
\frac{\partial}{\partial y}\left(\frac{\partial}{\partial x} f(x,y)\right) &= \frac{\partial^2}{\partial y \partial x} f(x,y) = f_{xy}(x,y) = 1 \\
\frac{\partial}{\partial x}\left(\frac{\partial}{\partial y} f(x,y)\right) &= \frac{\partial^2}{\partial x \partial y} f(x,y) = f_{yx}(x,y) = 1 \\
\frac{\partial}{\partial y}\left(\frac{\partial}{\partial y} f(x,y)\right) &= \frac{\partial^2}{\partial y \partial y} f(x,y) = f_{yy}(x,y) = 0
\end{aligned}$$

(2) $f(x,y) = e^{x-y}$
Es folgt

$$\begin{aligned}
f_x(x,y) &= e^{x-y} & f_y(x,y) &= -e^{x-y} \\
f_{xx}(x,y) &= e^{x-y} & f_{yx}(x,y) &= -e^{x-y} \\
f_{xy}(x,y) &= -e^{x-y} & f_{yy}(x,y) &= e^{x-y} \qquad \text{usw.}
\end{aligned}$$

(3) $f(x,y,z) = x^3 + xy^2 + xyz$
Die partiellen Ableitungen (erster Ordnung) sind

$$f_x(x,y,z) = 3x^2 + y^2 + yz \qquad f_y(x,y,z) = 2xy + xz \qquad f_z(x,y,z) = xy$$

und die partiellen Ableitungen zweiter Ordnung lauten

$$\begin{aligned}
f_{xx}(x,y,z) &= 6x & f_{yx}(x,y,z) &= 2y + z & f_{zx}(x,y,z) &= y \\
f_{xy}(x,y,z) &= 2y + z & f_{yy}(x,y,z) &= 2x & f_{zy}(x,y,z) &= x \\
f_{xz}(x,y,z) &= y & f_{yz}(x,y,z) &= x & f_{zz}(x,y,z) &= 0\ .
\end{aligned}$$

Unter den partiellen Ableitungen dritter Ordnung haben wir beispielsweise $f_{xxx}(x,y,z) = 6$ oder $f_{zyx}(x,y,z) = 1$ usw.

Obwohl es „eigentlich“ nicht gleichgültig ist, in welcher Reihenfolge man nach den verschiedenen Komponenten partiell ableitet, fällt auf, dass das Ergebnis unabhängig von dieser Reihenfolge zu sein scheint: es ist z.B. im letzten Beispiel $f_{xyz}(x,y,z) = f_{zyx}(x,y,z) = 1$. Tatsächlich gilt

3.3.3 Satz Existieren alle partiellen Ableitungen k-ter Ordnung und sind diese alle stetig, dann kann die Reihenfolge der partiellen Differenziationen vertauscht werden.

Bei einer Funktion $f : \mathbb{R}^n \to \mathbb{R}$ erhalten wir also (wenn die Funktion partiell differenzierbar ist) n partielle Ableitungen. Wie sieht es bei Funktionen $f : \mathbb{R}^n \to \mathbb{R}^m$ mit $m \geq 2$ aus?

In dieser Situation können wir die Funktion wieder in m Komponentenfunktionen zerlegen und diese wie gehabt partiell ableiten:

3.3.4 Definition Eine Funktion $f : D \to \mathbb{R}^m$, $D \subset \mathbb{R}^n$, $f(\vec{x}) = \begin{pmatrix} f_1(\vec{x}) \\ \vdots \\ f_m(\vec{x}) \end{pmatrix}$, mit $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ ist *partiell differenzierbar nach* x_k, wenn alle Komponentenfunktionen $f_1(\vec{x}), \ldots, f_m(\vec{x})$ nach x_k partiell differenzierbar sind.

f heißt *partiell differenzierbar*, wenn f nach allen $x_1, \ldots, x_n$ partiell differenzierbar ist.

3.3.5 Beispiel Betrachten wir $f(x,y) = (\sin x, \cos y, x-y)$.

Hier ist $f_1(x,y) = \sin x$, $f_2(x,y) = \cos y$ und $f_3(x,y) = x - y$ und wir erhalten

$$\begin{aligned}
\frac{\partial}{\partial x} f_1(x,y) &= \cos x & \qquad \frac{\partial}{\partial y} f_1(x,y) &= 0 \\
\frac{\partial}{\partial x} f_2(x,y) &= 0 & \qquad \frac{\partial}{\partial y} f_2(x,y) &= -\sin y \\
\frac{\partial}{\partial x} f_3(x,y) &= 1 & \qquad \frac{\partial}{\partial y} f_3(x,y) &= -1 \ .
\end{aligned}$$

3.3.6 Definition Sei $f : D \to \mathbb{R}^m$, $D \subset \mathbb{R}^n$, eine partiell differenzierbare Funktion mit stetigen partiellen Ableitungen $\frac{\partial}{\partial x_k} f_i(\vec{x})$.

Die (m,n)-Matrix

$$f'(\vec{x}) = Df(\vec{x}) = \begin{pmatrix} \frac{\partial}{\partial x_1} f_1(\vec{x}) & \dots & \frac{\partial}{\partial x_n} f_1(\vec{x}) \\ \vdots & & \vdots \\ \frac{\partial}{\partial x_1} f_m(\vec{x}) & \dots & \frac{\partial}{\partial x_n} f_m(\vec{x}) \end{pmatrix}$$

ist die *Ableitung* von f. Weitere übliche Bezeichnungen sind *Funktionalmatrix* oder *Jacobi-Matrix*.

3.3.7 Beispiel $f(x,y) = (\sin x, \cos y, x-y)$ (wie in Beispiel 3.3.5).

Dann ist die Ableitung von f die Matrix $f'(x,y) = \begin{pmatrix} \cos x & 0 \\ 0 & -\sin y \\ 1 & -1 \end{pmatrix}$.

3.3.8 Bemerkung Im Fall eindimensionaler Funktionen $f : \mathbb{R} \to \mathbb{R}$ gibt die Ableitung in einem Punkt a bekanntlich gerade die Steigung der Tangenten in diesem Punkt an. Die Tangente ist eine Gerade $t(x) = A + B \cdot (x - a)$ (Geradengleichung), wobei $A = f(a)$ den Funktionswert an der Stelle a und $B = f'(a)$ die Steigung bezeichnet:

$$t(x) = f(a) + f'(a) \cdot (x - a) \ .$$

Im mehrdimensionalen Fall, $f : \mathbb{R}^n \to \mathbb{R}^m$, ist dies vergleichbar mit der *Tangentialebene* in einem Punkt $\vec{a} \in \mathbb{R}^n$:

$$T(\vec{x}) = f(\vec{a}) + f'(\vec{a}) \cdot (\vec{x} - \vec{a}) \ . \tag{8}$$

3.3.9 Beispiel Betrachten wir $f(x,y) = x^2 - y^3$ im Punkt $\vec{a} = \binom{1}{-1}$ (vgl. Abbildung 19).

f' ist hier die $(1,2)$-Matrix $f'(x,y) = \begin{pmatrix} 2x & -3y^2 \end{pmatrix}$, also haben wir in $\vec{a}$ die Tangentialebene

$$\begin{aligned}
T(x,y) &= f(1,-1) + f'(1,-1) \cdot \left(\begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right) \\
&= 1^2 - (-1)^3 + \begin{pmatrix} 2 & -3 \end{pmatrix} \cdot \begin{pmatrix} x-1 \\ y+1 \end{pmatrix} \\
&= 2 + 2(x-1) - 3(y+1) \\
&= 2x - 3y - 3 \ .
\end{aligned}$$

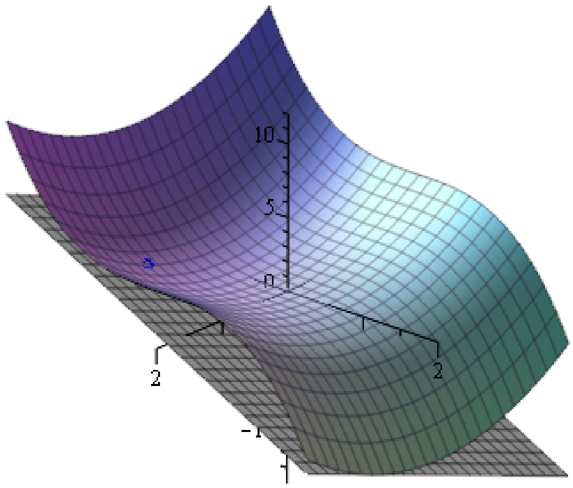


Abbildung 19: $f(x,y) = x^2 - y^3$ mit Tangentialebene in $\binom{1}{-1}$

3.4 Extremwerte

Bei Funktionen $f : \mathbb{R} \to \mathbb{R}$ stellt man sich die Ableitung in einem Punkt gerne als die Steigung des Funktionsgraphen in diesem Punkt vor. Bei Funktionen mit mehrdimensionalen Bildern macht der Begriff der Steigung keinen Sinn mehr, da dort keine „größer“- oder „kleiner“-Relation existiert.

Auch bei Funktionen $f : \mathbb{R}^n \to \mathbb{R}$ (deren Graph eine Art „Berg- und Tallandschaft“ darstellt) können wir nicht von „der“ Steigung in einem Punkt sprechen, da die Steigung in einem Punkt von der Richtung abhängt, in der sie gemessen wird. Statt dessen interessiert man sich eher für die maximale Steigung in einem Punkt.

3.4.1 Definition Ist $f : D \to \mathbb{R}$ mit $D \subset \mathbb{R}^n$, in $\vec{a} \in D$ partiell differenzierbar, dann ist der Vektor der partiellen Ableitungen an der Stelle $\vec{a} \in D$ der *Gradient* von f in $\vec{a}$:

$$\operatorname{grad} f(\vec{a}) = \left(\frac{\partial}{\partial x_1} f(\vec{a}), \frac{\partial}{\partial x_2} f(\vec{a}), \ldots, \frac{\partial}{\partial x_n} f(\vec{a}) \right)$$

3.4.2 Bemerkung Der Gradient entspricht der Ableitung von f bei Skalar-wertigen Funktionen, die in diesem Fall eine $(1,n)$-Matrix ist. Es ist also $\operatorname{grad} f(\vec{a}) = f'(\vec{a})$.

3.4.3 Satz Der Gradient von f zeigt in die Richtung des stärksten Anstiegs von f.

3.4.4 Beispiele **(1)** $f(x,y) = x^2 - y^3$ in $\vec{a} = \binom{1}{-1}$ (vgl. Abbildung 19).

Wir haben $\operatorname{grad} f(x,y) = \big(2x, -3y^2\big)$, also $\operatorname{grad} f(1,-1) = (2,-3)$. Im Punkt $\binom{1}{-1}$ ist der stärkste Anstieg also in Richtung $\binom{2}{-3}$ zu finden.

(2) $f(x,y) = x^2 + y^2 + 1$.

Der Gradient dieser Funktion ist $\operatorname{grad} f(x,y) = (2x, 2y)$, also z.B.

$$\operatorname{grad} f(0,0) = \vec{0}\ .$$

In $\binom{0}{0}$ haben hier also keine Richtung des stärksten Anstiegs!

Grund hierfür ist, dass die Funktion an der Stelle $\vec{0}$ ein Minimum hat und die Tangentialebene in diesem Punkt daher waagerecht ist (vgl. Abbildung 20). Dies entspricht bei eindimensionalen Funktionen einer waagerechten Tangente mit Steigung 0.

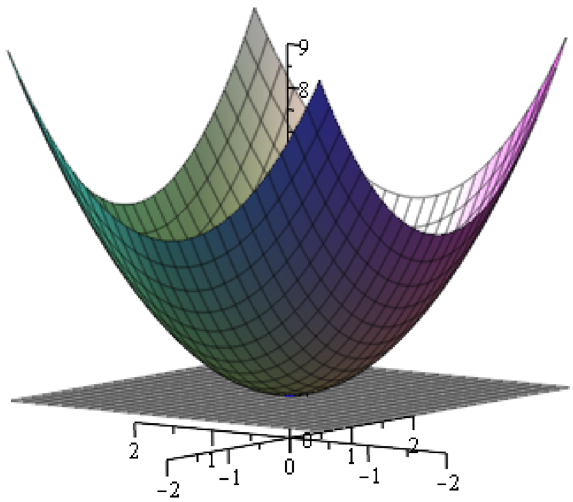


Abbildung 20: $f(x, y) = x^2 + y^2 + 1$ mit Tangentialebene in $\vec{0}$

Bei Funktionen $f : \mathbb{R} \to \mathbb{R}$ wissen wir, dass in einem lokalen Extremum stets $f'(x) = 0$ gilt (das *notwendige Kriterium* für Extrema). Im Mehrdimensionalen ist die Situation ganz ähnlich:

3.4.5 Satz (Notwendiges Kriterium für Extremwerte) Eine Funktion $f : D \to \mathbb{R}$ mit $D \subset \mathbb{R}^n$ sei in $\vec{a} \in D$ partiell differenzierbar. Hat f in $\vec{a}$ ein lokales Extremum, dann gilt

$$\operatorname{grad} f(\vec{a}) = \vec{0} \,.$$

3.4.6 Bemerkung Wie im Eindimensionalen können wir dies nutzen, um *mögliche* Stellen für Extrema zu finden. Es ist aber auch hier unklar, ob an diesen Stellen wirklich Extremwerte vorliegen. Betrachten wir beispielsweise die Sattelfläche $f(x, y) = x^2 - y^2$ (Abbildung 12). Der Gradient ist $\operatorname{grad} f(x, y) = (2x, -2y)$.

In $\vec{0}$ ist also $\operatorname{grad} f(\vec{0}) = \vec{0}$, das notwendige Kriterium ist somit erfüllt. Tatsächlich gibt es in $\vec{0}$ aber keinen Extremwert!

Im Eindimensionalen konnten wir die Entscheidung, ob ein Extremum vorliegt, meistens treffen, indem wir das Vorzeichen der zweiten Ableitung prüften: ist $f'(a) = 0$ und $f''(a) > 0$, haben wir ein Minimum in a, ist $f'(a) = 0$ und $f''(a) < 0$, haben wir ein Maximum.

Im Mehrdimensionalen ist dies leider nicht mehr so einfach, da wir nicht *eine* zweite Ableitung haben, sondern mehrere partielle Ableitungen zweiter Ordnung:

3.4.7 Definition Sei $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$, eine zweimal partiell differenzierbare Funktion. Die (n, n)-Matrix der partiellen Ableitungen zweiter Ordnung in $\vec{a} \in D$ heißt *Hesse-Matrix* von f in $\vec{a}$

$$H_f(\vec{a}) = \begin{pmatrix} f_{x_1,x_1}(\vec{a}) & f_{x_1,x_2}(\vec{a}) & \cdots & f_{x_1,x_n}(\vec{a}) \\ f_{x_2,x_1}(\vec{a}) & f_{x_2,x_2}(\vec{a}) & \cdots & f_{x_2,x_n}(\vec{a}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_n,x_1}(\vec{a}) & f_{x_n,x_2}(\vec{a}) & \cdots & f_{x_n,x_n}(\vec{a}) \end{pmatrix} .$$

3.4.8 Bemerkung Gibt es alle partiellen Ableitungen zweiter Ordnung und sind diese stetig, dann kann nach Satz 3.3.3 die Reihenfolge der Differenziation vertauscht werden: $f_{x_i,x_k}(\vec{a}) = f_{x_k,x_i}(\vec{a})$. $H_f(\vec{a})$ ist in diesem Fall „symmetrisch“.

3.4.9 Beispiel Betrachten wir $f(x,y) = x^3 + y^3$.

Es ist $f_x(x,y) = 3x^2$ und $f_y(x,y) = 3y^2$ und

$$\begin{aligned} f_{xx}(x,y) &= 6x & \qquad f_{xy}(x,y) &= 0 \\ f_{yx}(x,y) &= 0 & \qquad f_{yy}(x,y) &= 6y \end{aligned}$$

und die Hesse-Matrix lautet daher

$$H_f(x,y) = \begin{pmatrix} 6x & 0 \\ 0 & 6y \end{pmatrix} .$$

Beispielsweise ist also $H_f(1,2) = \begin{pmatrix} 6 & 0 \\ 0 & 12 \end{pmatrix}$.

Die Hesse-Matrix wird nun also die Rolle der zweiten Ableitung übernehmen, wir benötigen aber noch ein Kriterium, das statt $f''(x) > 0$ bzw. $f''(x) < 0$ verwendet werden kann, da wir bei einer Matrix natürlich nicht einfach nach dem Vorzeichen fragen können:

3.4.10 Definition Eine symmetrische Matrix A habe die Eigenwerte $\lambda_1, \ldots, \lambda_n$.

A ist *positiv definit*, wenn alle $\lambda_i > 0$ sind,

A ist *negativ definit*, wenn alle $\lambda_i < 0$ sind,

A ist *positiv semidefinit*, wenn alle $\lambda_i \geq 0$ sind,

A ist *negativ semidefinit*, wenn alle $\lambda_i \leq 0$ sind und

A ist *indefinit*, wenn es positive und negative Eigenwerte gibt.

Damit haben wir nun alle Werkzeuge zusammen, um die Extrema einer Funktion finden zu können:

3.4.11 Satz (Hinreichendes Kriterium für Extremwerte) Sei $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$, eine Funktion mit stetigen partiellen Ableitungen erster und zweiter Ordnung.

In $\vec{a} \in D$ gelte $\operatorname{grad} f(\vec{a}) = \vec{0}$. f hat dann in $\vec{a}$

...ein Maximum, wenn $H_f(\vec{a})$ negativ definit ist,

...ein Minimum, wenn $H_f(\vec{a})$ positiv definit ist und

...kein lokales Extremum, wenn $H_f(\vec{a})$ indefinit ist.

Ist $H_f(\vec{a})$ semidefinit, können wir keine Aussage über ein Extremum in diesem Punkt machen.

3.4.12 Beispiele **(1)** $f(x,y) = x^2 + y^2 + 1$ (vgl. Abbildung 20)
Es ist $f_x(x,y) = 2x$ und $f_y(x,y) = 2y$, es ist also

$$\operatorname{grad} f(x,y) = (2x, 2y) = (0,0) \quad \Leftrightarrow \quad \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} .$$

An der Stelle $\binom{0}{0}$ ist also ein Extremum möglich.
Die Hesse-Matrix von f lautet

$$H_f(\vec{a}) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \qquad \text{für alle } \vec{a}.$$

Bei einer solchen Dreiecksmatrix sind die Eigenwerte nach Satz 2.11.6 besonders leicht zu ermitteln: sie stehen auf der Diagonalen von links oben nach rechts unten. In diesem Fall gibt es also nur einen Eigenwert, nämlich 2. $H_f(\vec{0})$ ist folglich positiv definit und f hat daher ein Minimum in $\vec{0}$.

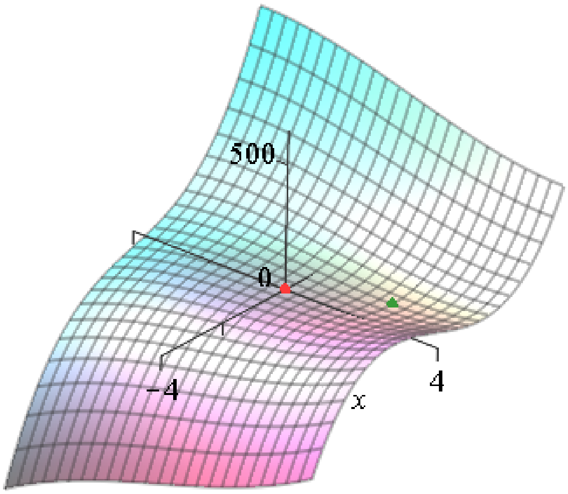


Abbildung 21: $f(x, y) = x^3 - 12xy + 8y^3$ mit möglichen Extrema in $\binom{0}{0}$ und $\binom{2}{1}$

(2) $f(x, y) = x^3 - 12xy + 8y^3$ (vgl. Abbildung 21)
Es gilt

$$\begin{aligned} & \operatorname{grad} f(x,y) = \left(3x^2 - 12y, -12x + 24y^2\right) = \vec{0} \\ \Leftrightarrow \quad & 3x^2 - 12y = 0 \qquad (9) \\ & \text{und} \quad -12x + 24y^2 = 0 \qquad (10) \end{aligned}$$

Dies ist zwar ein Gleichungssystem, es ist aber offensichtlich nicht linear. Unsere üblichen Verfahren zur Lösung linearer Gleichungssysteme können wir also nicht verwenden.

Gleichung (9) ergibt

$$y = \frac{1}{4}x^2 \tag{11}$$

und damit erhalten wir aus Gleichung (10)

$$\begin{aligned} & -12x + \frac{3}{2}x^4 = 0 \\ \Leftrightarrow \quad & x\left(-12 + \frac{3}{2}x^3\right) = 0 \\ \Leftrightarrow \quad & x = 0 \quad \text{oder} \quad \frac{3}{2}x^3 = 12 \\ \Leftrightarrow \quad & x = 0 \quad \text{oder} \quad x = 2 \end{aligned}$$

Unter Verwendung von (11) ist

$$\begin{aligned} x = 0 \quad & \Rightarrow \quad y = 0 \\ x = 2 \quad & \Rightarrow \quad y = 1\ , \end{aligned}$$

also können mögliche lokale Extrema in $\binom{0}{0}$ und in $\binom{2}{1}$ liegen.

Um zu entscheiden, ob es dort tatsächlich Extremwerte gibt, müssen wir die Definitheit der Hesse-Matrix prüfen. Es ist

$$H_f(x, y) = \begin{pmatrix} 6x & -12 \\ -12 & 48y \end{pmatrix} .$$

Im Punkt $\binom{0}{0}$ ist $H_f(0,0) = \begin{pmatrix} 0 & -12 \\ -12 & 0 \end{pmatrix}$. Die Eigenwerte dieser Matrix sind die Nullstellen des charakteristischen Polynoms

$$P(\lambda) = \det(H_f(0,0) - \lambda E) = \begin{vmatrix} -\lambda & -12 \\ -12 & -\lambda \end{vmatrix} = \lambda^2 - 144 = 0 \qquad \Rightarrow \qquad \lambda = \pm 12 \ .$$

$H_f(0,0)$ hat also sowohl einen positiven als auch einen negativen Eigenwert, ist also indefinit und somit gibt es an der Stelle $\binom{0}{0}$ kein Extremum.

Im Punkt $\binom{2}{1}$ ist $H_f(2,1) = \begin{pmatrix} 12 & -12 \\ -12 & 48 \end{pmatrix}$ und die Eigenwerte sind

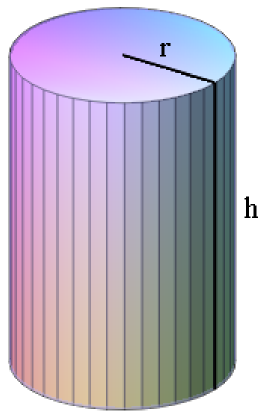
$$\begin{aligned} P(\lambda) &= \det(H_f(2,1) - \lambda E) \\ &= \begin{vmatrix} 12-\lambda & -12 \\ -12 & 48-\lambda \end{vmatrix} \\ &= (12-\lambda)(48-\lambda) - 144 \\ &= \lambda^2 - 60\lambda + 432 = 0 \\ \Rightarrow \qquad \lambda &= 30 \pm \sqrt{468} \approx \begin{cases} 51.63 & > 0 \\ 8.36 & > 0 \end{cases} \end{aligned}$$

$H_f(2,1)$ ist demnach positiv definit, wir haben in $\binom{2}{1}$ also ein Minimum.

3.5 Extrema mit Nebenbedingungen

Betrachten wir folgendes Problem:

Welche Maße muss eine Dose haben, damit möglichst wenig Blech verbraucht wird ?



Die Oberfläche einer Dose (Zylinder) der Höhe h und mit Radius r ist

$$\underbrace{2\pi r}_{\text{Umfang}} \cdot \underbrace{h}_{\text{Höhe}} + 2 \cdot \underbrace{\pi r^2}_{\text{Fläche der Ober-/Unterseite}} \ .$$

Wir müssen also nach einem Minimum von $F(r,h) = 2\pi rh + 2\pi r^2$ suchen:

$$\operatorname{grad} F(r,h) = (2\pi h + 4\pi r, 2\pi r) = \vec{0}$$

also

$$2\pi h + 4\pi r = 0 \quad \text{und} \quad 2\pi r = 0 \ .$$

Aus der zweiten Gleichung ergibt sich $r = 0$ und wenn wir dies in die erste Gleichung einsetzen, erhalten wir $h = 0$. Ein mögliches Extremum des Materialverbrauchs haben wir also bei einer Dose mit Höhe und Radius 0. Wie man das bei mathematischen Resultaten erwarten kann, ist dieses Ergebnis natürlich völlig richtig — es ist aber nicht wirklich nützlich.

Sinnvoll wird unsere Überlegung erst, wenn wir zusätzlich verlangen, dass die Dose ein gegebenes Volumen haben soll:

$$V(r,h) = \pi r^2 \cdot h = C \ .$$

Dies ist eine Extremwertaufgabe mit einer *Nebenbedingung*:

Wann ist die Oberfläche $F(r,h)$ minimal unter der Nebenbedingung, dass $V(r,h) = C$?

Es gilt:

3.5.1 Satz Wenn es ein Extremum einer Funktion $f : D \to \mathbb{R}$, $D \subset \mathbb{R}^n$, unter der Nebenbedingung $g(\vec{x}) = 0$ für $g : \mathbb{R}^n \to \mathbb{R}$ mit $\operatorname{grad} g(\vec{x}) \neq \vec{0}$ gibt, dann in einem Punkt $\vec{x}$ mit

$$g(\vec{x}) = 0 \qquad \text{und} \qquad \operatorname{grad} f(\vec{x}) = \lambda \cdot \operatorname{grad} g(\vec{x}) \qquad \text{für ein } \lambda \in \mathbb{R}.$$

3.5.2 Bemerkung **(1)** So gut wie jede Nebenbedingung lässt sich in der Form $g(\vec{x}) = 0$ ausdrücken. Unsere obige Bedingung $V(r,h) = \pi r^2 \cdot h = C$ können wir beispielsweise in der Form $g(r,h) = V(r,h) - C = 0$ schreiben.

(2) λ heißt *Lagrange-Multiplikator*. Der Wert von λ hat keine weitere Bedeutung.

(3) Der Satz sagt nicht, wo Extrema liegen, es handelt sich nur um ein notwendiges Kriterium. Wir können aber hiermit die Stellen errechnen, an denen Extrema *möglich* sind. Ob es dort dann tatsächlich Extremwerte gibt, muss auf anderem Weg entschieden werden — in der Praxis genügt hierzu meistens die Anschauung.

3.5.3 Beispiele **(1)** Unser Dosen-Beispiel: Wo liegt ein mögliches Minimum von

$$F(r,h) = 2\pi r h + 2\pi r^2$$

unter der Nebenbedingung

$$g(r,h) = V(r,h) - C = \pi r^2 \cdot h - C = 0 \quad ? \tag{12}$$

Wir müssen zuerst prüfen, ob die Voraussetzungen von Satz 3.5.1 erfüllt sind: Es gilt

$$\operatorname{grad} g(r,h) = \left(2\pi r h, \pi r^2\right) \neq (0,0) \qquad \text{für } r \neq 0.$$

Für $r \neq 0$ können wir unseren Satz also benutzen (der Fall $r = 0$ ist auch offensichtlich uninteressant). Nach Satz 3.5.1 gilt also an der Stelle eines Extremums unter der Nebenbedingung (12)

$$\begin{array}{rcr@{\,=\,}l}
 & & g(r,h) & 0 \\
 & \text{und} & \operatorname{grad} F(r,h) & \lambda \operatorname{grad} g(r,h) \\[1ex]
\Leftrightarrow & & \pi r^2 h - C & 0 \\
 & \text{und} & (2\pi h + 4\pi r, 2\pi r) & \lambda \left(2\pi r h, \pi r^2\right) \\[1ex]
\Leftrightarrow & & \pi r^2 h & C \\
 & \text{und} & 2\pi h + 4\pi r & 2\lambda \pi r h \\
 & \text{und} & 2\pi r & \lambda \pi r^2 \\[1ex]
\Leftrightarrow & & \pi r^2 h & C \\
 & \text{und} & h + 2r & \lambda r h \\
 & \text{und} & 2 & \lambda r
\end{array}$$

Aus der letzten Zeile folgt $r = \frac{2}{\lambda}$. Setzen wir dies in die vorletzte Zeile ein, erhalten wir $h + \frac{4}{\lambda} = 2h \;\Leftrightarrow\; h = \frac{4}{\lambda}$.

Insgesamt gilt also $h = 2r$, d.h. wenn es ein Minimum gibt, dann bei Dosen, deren Höhe doppelt so groß ist wie ihr Radius.

Auch die Werte für h und r können wir berechnen: Ist C das gewünschte Volumen, dann gilt wegen $h = 2r$

$$\pi r^2 h = C \qquad \Leftrightarrow \qquad \pi r^2 \cdot 2r = C \qquad \Leftrightarrow \qquad 2\pi r^3 = C \qquad \Leftrightarrow \qquad r = \sqrt[3]{\frac{C}{2\pi}}\,.$$

Soll die Dose beispielsweise ein Volumen von $C = 500\,\text{ml}$ haben, so erhalten wir $r \approx 4.3\,\text{cm}$ und damit $h \approx 8.6\,\text{cm}$.

(2) Wo liegen mögliche Extrema von $f(x,y) = xy$ unter der Nebenbedingung $x^2 + y^2 = 2$?
Hier formulieren wir die Nebenbedingung in der Form $g(x,y) = x^2 + y^2 - 2 = 0$ und es gilt

$$\operatorname{grad} g(x,y) = (2x, 2y) \neq (0,0) \qquad \text{für} \qquad (x,y) \neq \vec{0} \ .$$

Sei also $(x,y) \neq \vec{0}$. Laut Satz 3.5.1 erfüllen eventuelle Extremalstellen

$$\begin{aligned} && g(x,y) &= 0 \\ & \text{und} & \operatorname{grad} f(x,y) &= \lambda \operatorname{grad} g(x,y) \\ \Leftrightarrow && x^2 + y^2 - 2 &= 0 \\ & \text{und} & (y,x) &= \lambda(2x, 2y) \\ \Leftrightarrow && x^2 + y^2 &= 2 \\ & \text{und} & y &= 2\lambda x \\ & \text{und} & x &= 2\lambda y \ . \end{aligned}$$

Setzen wir die letzte Gleichung in die vorletzte ein, so erhalten wir

$$y = 2\lambda \cdot 2\lambda y \qquad \Leftrightarrow \qquad \lambda^2 = \frac{1}{4} \qquad \Leftrightarrow \qquad \lambda = \pm\frac{1}{2} \ .$$

Wir können nun x und y berechnen:
Für $\lambda = \frac{1}{2}$ folgt $y = 2\lambda x \;\Leftrightarrow\; y = x$ und

$$x^2 + y^2 = 2 \qquad \Leftrightarrow \qquad 2x^2 = 2 \qquad \Leftrightarrow \qquad x^2 = 1 \qquad \Leftrightarrow \qquad x = \pm 1 \ ,$$

mögliche Extrema liegen also in $\binom{1}{1}$ und $\binom{-1}{-1}$.
Für $\lambda = -\frac{1}{2}$ folgt $y = 2\lambda x \;\Leftrightarrow\; y = -x$ und wie oben

$$x^2 + y^2 = 2 \qquad \Leftrightarrow \qquad 2x^2 = 2 \qquad \Leftrightarrow \qquad x = \pm 1 \ ,$$

mögliche Extrema liegen hier also in $\binom{1}{-1}$ und $\binom{-1}{1}$.

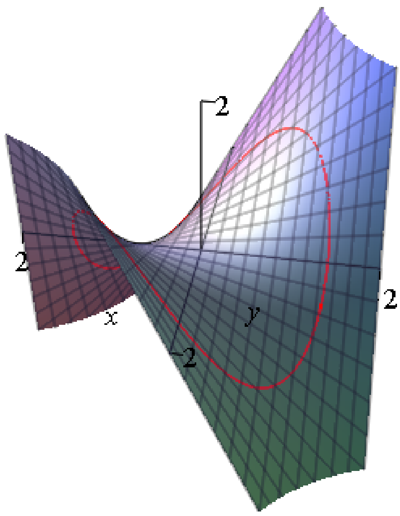


Abbildung 22: $f(x,y) = xy$ und die Punkte mit $x^2 + y^2 = 2$

Insgesamt haben wir also vier Stellen, an denen eventuell Extrema unter der Nebenbedingung existieren. Am Graphen (Abbildung 22, dort sind die Punkte, für die die Nebenbedingung gilt, mit einer roten Linie markiert) können wir erkennen, dass es Maxima in $\binom{1}{1}$ und $\binom{-1}{-1}$ ($f(x,y) = 1$) sowie Minima in $\binom{-1}{1}$ und $\binom{1}{-1}$ ($f(x,y) = -1$) gibt.

3.6 Vektorfelder

Kommen wir noch einmal auf den Gradienten zurück. Nach Def. 3.4.1 ist

$$\operatorname{grad} f(\vec{a}) = \left(\frac{\partial}{\partial x_1} f(\vec{a}), \frac{\partial}{\partial x_2} f(\vec{a}), \ldots, \frac{\partial}{\partial x_n} f(\vec{a})\right)$$

für eine Funktion $f : D \to \mathbb{R}$ mit $D \subset \mathbb{R}^n$.

f ist also eine Funktion, die jedem Punkt eines mehrdimensionalen Raumes einen Skalar zuweist. Eine solche Funktion wird deshalb auch als *Skalarfeld* bezeichnet.

Analog bezeichnet man eine Funktion $v : D \to \mathbb{R}^n$ mit $D \subset \mathbb{R}^n$, die jedem Punkt eines Raumes einen Vektor (gleicher Dimension) zuweist, als *Vektorfeld*.

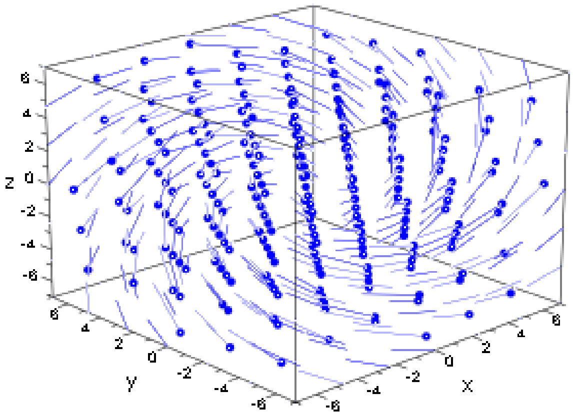


Abbildung 23: 3-dimensionales Vektorfeld $v : \mathbb{R}^3 \to \mathbb{R}^3$

Ein Beispiel für ein Vektorfeld ist der Gradient einer Funktion (Gradient eines Skalarfeldes): er weist einem Punkt $\vec{a} \in \mathbb{R}^n$ den Vektor $\left(\frac{\partial}{\partial x_1} f(\vec{a}), \frac{\partial}{\partial x_2} f(\vec{a}), \ldots, \frac{\partial}{\partial x_n} f(\vec{a})\right) \in \mathbb{R}^n$ zu. Man spricht deshalb auch von einem *Gradientenfeld*. In diesem Zusammenhang verwendet man oft eine andere Schreibweise:

Mit dem *Nabla-Operator* $\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_n}\right)$ ist[5]

$$\nabla f = \operatorname{grad} f \ .$$

Betrachten wir ein *partiell differenzierbares Vektorfeld*, d.h. alle Komponenten-Funktionen sind partiell differenzierbar, dann lassen sich mit dem Nabla-Operator zwei weitere Funktionen beschreiben.

Die *Divergenz* beschreibt die Tendenz eines Vektorfeldes „auseinanderzufließen“, die *lokale Quellendichte*. Es handelt sich dabei um das Skalarprodukt des Nabla-Operators mit einem Vektorfeld $v : \mathbb{R}^n \to \mathbb{R}^n$:

$$\operatorname{div} v = \nabla \cdot v = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \frac{\partial}{\partial x_1} v_1 + \frac{\partial}{\partial x_2} v_2 + \cdots + \frac{\partial}{\partial x_n} v_n \ .$$

[5] Das Zeichen ∇ ähnelt der Form einer Harfe, alt-griechisch „Nabla“.

Analog bezeichnet die *Rotation* die Tendenz eines dreidimensionalen Vektorfeldes, um Punkte zu rotieren, die *lokale Wirbeldichte*. Dies ist das Vektorprodukt des Nabla-Operators mit dem Vektorfeld:

$$\operatorname{rot} v = \nabla \times v = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \frac{\partial}{\partial x_3} \end{pmatrix} \times \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial x_2} v_3 - \frac{\partial}{\partial x_3} v_2 \\ \frac{\partial}{\partial x_3} v_1 - \frac{\partial}{\partial x_1} v_3 \\ \frac{\partial}{\partial x_1} v_2 - \frac{\partial}{\partial x_2} v_1 \end{pmatrix} .$$

Differenzialgleichungen

4 Gewöhnliche Differenzialgleichungen

4.1 Bezeichnungen

Betrachten wir folgende Situation: $n(t)$ sei die Anzahl der Atome in einer radioaktiven Probe zum Zeitpunkt t. Durch den Zerfall ändert sich diese Zahl, die Ableitung $n'(t)$ gibt bekanntlich diese Veränderung an.

Diese Veränderung ist proportional zur jeweiligen Zahl der Atome $n(t)$, mit einer Substanz-spezifischen Konstanten $k \geq 0$ gilt also

$$n'(t) = -k \cdot n(t) \ . \tag{13}$$

Unsere Aufgabe besteht nun darin, $n(t)$ zu bestimmen, also eine Funktion zu finden, die die Gleichung (13) erfüllt.

4.1.1 Bezeichnungen Eine Gleichung, wie beispielsweise (13), in der neben einer unbekannten Funktion auch deren Ableitung auftritt, heißt *Differenzialgleichung*. Wenn es sich wie in (13) um die erste Ableitung handelt, spricht man von einer Differenzialgleichung *erster Ordnung*, allgemein handelt es sich um eine Differenzialgleichung *n-ter Ordnung*, wenn höhere Ableitungen bis zur n-ten Ordnung auftreten.

Wenn die gesuchte Funktion von mehreren Variablen abhängt, können auch partielle Ableitungen auftreten, dann handelt es sich um eine *partielle Differenzialgleichung*, sonst um eine *gewöhnliche*.

Die gesuchte Funktion, die die Differenzialgleichung erfüllt, heißt *Lösung* der Differenzialgleichung.

Wir werden uns zunächst nur mit gewöhnlichen Differenzialgleichungen erster Ordnung befassen.

4.1.2 Notation Die gesuchte Funktion wird meist $y(x)$ geschrieben und zur Abkürzung wird „(x)“ hinter y, y',... einfach weggelassen. Aus (13) wird also

$$y' = -ky \ .$$

Allgemein kann man eine gewöhnliche Differenzialgleichung erster Ordnung in der Form

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

und eine Differenzialgleichung n-ter Ordnung in der Form $y^{(n)} = f\left(x, y, y', \ldots, y^{(n-1)}\right)$ schreiben.

4.1.3 Beispiel Eine Lösung der Differenzialgleichung $y' = -2y$ ist $y(x) = e^{-2x}$, denn es folgt $y'(x) = -2e^{-2x} = -2y(x)$.

Auch $y(x) = 2e^{-2x}$ oder $y(x) = -e^{-2x}$ sind Lösungen der Differenzialgleichung, wie man leicht nachrechnet.

Wie wir sehen, gibt es im Allgemeinen mehrere Lösungen einer Differenzialgleichung, man spricht dann von einer *Lösungsschar*.

4.2 Richtungsfelder

Die verschiedenen Lösungen einer Differenzialgleichung können graphisch dargestellt werden, sogar ohne, dass man diese Lösungen explizit kennen muss:

Betrachten wir eine Differenzialgleichung

$$y' = f(x, y) \ ,$$

(also $y'(x) = f(x, y(x))$) und einen Punkt $\binom{a}{b} \in \mathbb{R}^2$. Nehmen wir an, wir hätten eine Lösung $y(x)$ der Differenzialgleichung, deren Graph durch den Punkt $\binom{a}{b}$ läuft, d.h., für die $y(a) = b$ gilt.

Wir kennen dann die Steigung von $y(x)$ in diesem Punkt, nämlich

$$y'(a) = f(a, y(a)) = f(a, b) \ .$$

4.2.1 Bezeichnungen Da uns also in jedem Punkt $\binom{a}{b}$ die Steigung von y bekannt ist, können wir diese Steigung für jeden Punkt der Ebene als kleinen Pfeil einzeichnen. Dies bezeichnet man als *Richtungsfeld*.

Zeichnen wir eine Kurve, die diesen Pfeilen folgt und durch einen bestimmten Punkt $\binom{a}{b}$ läuft, so erhalten wir den Graphen einer einzelnen Lösung, die (im Gegensatz zur Lösungsschar) als *partikuläre Lösung* bezeichnet wird. Für diese Lösung gilt dann $y(a) = b$.

Die Aufgabe, zu einer Differenzialgleichung eine partikuläre Lösung mit $y(a) = b$ zu finden, heißt *Anfangswertproblem*.

4.2.2 Beispiele Betrachten wir $y' = -\dfrac{x}{y}$

Tragen wir $-\frac{x}{y}$ für verschiedene x und y in eine Tabelle ein, so erhalten wir z.B.

	$x = -2$	-1	0	1	2
$y = 1$	2	1	0	-1	-2
2	1	$\frac{1}{2}$	0	$-\frac{1}{2}$	-1

Das Richtungsfeld erhalten wir, indem wir diese Werte als Steigungen in ein Koordinatensystem einzeichnen (Abbildung 24).

Folgen wir von einem Punkt $\binom{x_0}{y_0}$ den Pfeilen, erhalten wir die partikuläre Lösung des Anfangswertproblems $y(x_0) = y_0$.

Weitere Beispiele sind in den Abbildungen 25, 26 und 27 zu finden.

4.3 Elementare Lösungsverfahren

Es gibt leider kein allgemeines Verfahren, eine beliebige Differenzialgleichung zu lösen. Bei vielen bestimmten Typen von Differenzialgleichungen lassen sich aber Lösungen bestimmen.

Wir werden nun einige der gewöhnliche Differenzialgleichungen erster Ordnung klassifizieren und jeweils zugehörige Lösungsansätze ermitteln.

Typ (A) $y' = f(x)$

Bei diesem Typ hängt die „rechte Seite“ nicht von y ab.

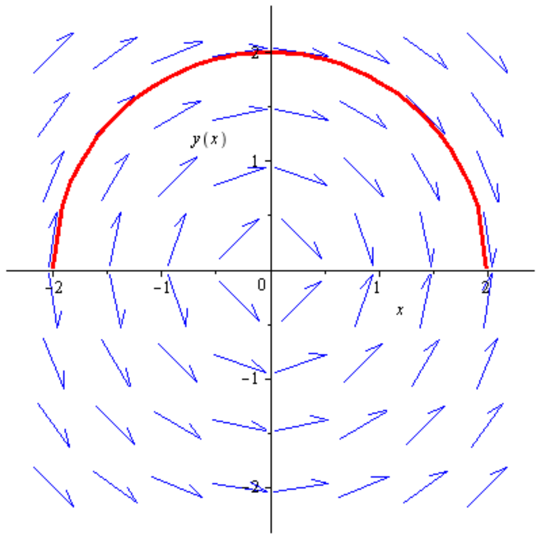


Abbildung 24: Richtungsfeld von $y' = -\frac{x}{y}$ mit partikulärer Lösung zu $y(0) = 2$

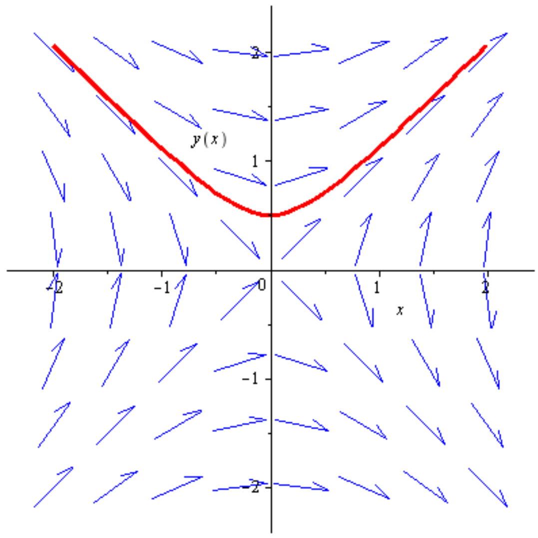


Abbildung 25: Richtungsfeld von $y' = \frac{x}{y}$ mit partikulärer Lösung zu $y(0) = \frac{1}{2}$

4.3.1 Beispiel $y' = 2x$

Eine Lösung ist hier recht leicht zu finden: Wir integrieren beide Seiten der Differenzialgleichung und erhalten

$$y'(x) = 2x \qquad \Rightarrow \qquad \int y'(x)\,dx = \int 2x\,dx \qquad \Rightarrow \qquad y(x) = x^2 + C\ .$$

Die Integrationskonstanten beider Seiten haben wir hier zu einer Konstanten C zusammengefasst. $y(x) = x^2 + C$ ist also unsere Lösungsschar.

Für jedes $C \in \mathbb{R}$ erhalten wir eine partikuläre Lösung: Ist zum Beispiel das Anfangswertproblem $y(0) = 0$ gegeben, so ergibt sich

$$y(0) = 0^2 + C = 0 \qquad \Rightarrow \qquad C = 0\ .$$

$y(x) = x^2$ ist folglich in diesem Fall die partikuläre Lösung.

Im Allgemeinen lösen wir eine Differenzialgleichung diesen Typs also durch Integration:

$$y' = f(x) \qquad \Rightarrow \qquad \int y'(x)\,dx = \int f(x)\,dx \qquad \Rightarrow \qquad \boxed{y(x) = \int f(x)\,dx} \tag{14}$$

Auch einige wichtige Anwendungen von Differenzialgleichungen gehören zu diesem Typ:

4.3.2 Beispiel Manche chemische Reaktionen gehorchen einem sog. „Geschwindigkeitsgesetz nullter Ordnung“ — ein Beispiel hierfür ist z.B. der Alkoholabbau im Blut, der näherungsweise diesem Gesetz folgt. Dabei ist der Abbau der Konzentration stets konstant, also unabhängig von der Konzentration. Wenn $c_A(t)$ die Konzentration zum Zeitpunkt t beschreibt, gilt also

$$c_A' = -k \qquad \text{mit einem Geschwindigkeitskoeffizienten } k\ .$$

In (14) setzen wir $y = c_A$ und $f(t) = -k$ und erhalten

$$c_A(t) = \int -k\,dt = -k \cdot t\ .$$

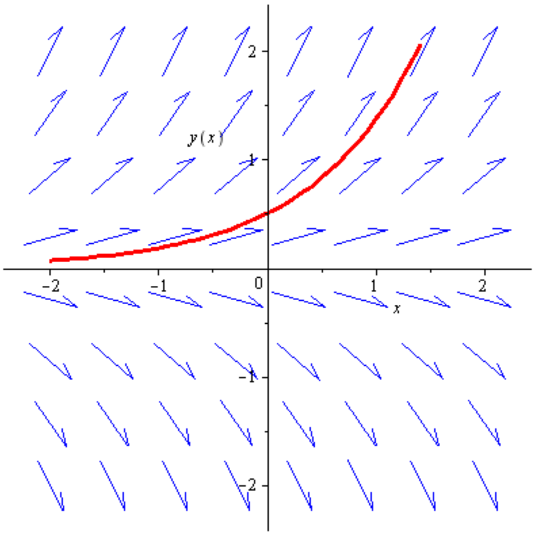


Abbildung 26: Richtungsfeld von $y' = y$ mit partikulärer Lösung zu $y(0) = \frac{1}{2}$

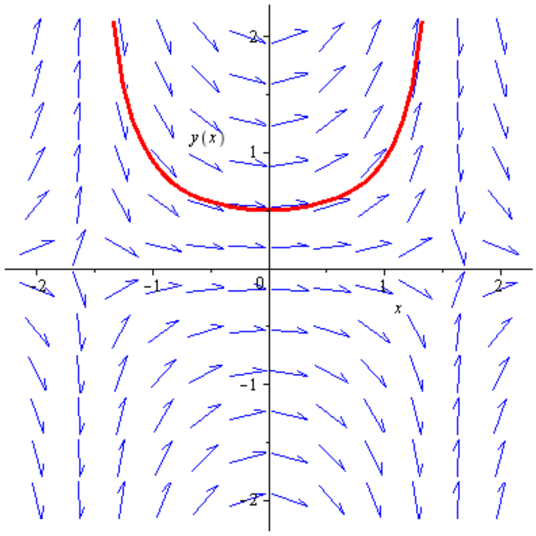


Abbildung 27: Richtungsfeld von $y' = y \cdot \tan x$ mit partikulärer Lösung zu $y(0) = \frac{1}{2}$

4.4 Autonome Differenzialgleichungen

Typ (B) Autonome Differenzialgleichung $y' = f(y)$

Bei einer *autonomen Differenzialgleichung* taucht x nicht auf der rechten Seite auf. Um eine Lösung herzuleiten, verwenden wir eine „*formale Rechnung*“: dies ist keine mathematisch exakte Überlegung, sondern eher eine Art „Merkregel“, um eine Lösung zu finden. Überraschend oft funktioniert dies und führt zu einem Ergebnis, das sich auch mit exakten Mitteln beweisen ließe:

Sei nun $f(x) \neq 0$. Unsere Differenzialgleichung $y' = f(y)$ kann auch geschrieben werden als

$$\frac{dy}{dx} = f(y) \qquad \Leftrightarrow \qquad \frac{dx}{dy} = \frac{1}{f(y)} \qquad \Leftrightarrow \qquad x' = \frac{1}{f(y)} \ .$$

Durch diese formale Rechnung haben wir also x und y vertauscht und so eine Differenzialgleichung des Typs (A) erhalten. Nach (14) finden wir unsere Lösung also auch hier durch Integration und unser Lösungsansatz lautet nun

$$\boxed{x(y) = \int \frac{1}{f(y)}\, dy} \tag{15}$$

Wenn wir dieses Integral berechnen und die Gleichung dann nach y auflösen, erhalten wir eine Lösung von $y' = f(y)$, wenn $f(y) \neq 0$ ist.

Ist dagegen $f(y) = 0$, dann sei y_0 diese Nullstelle, d.h. $f(y_0) = 0$. Die Funktion $y(x) = y_0$ ist dann eine Lösung von $y' = f(y)$. Solche Lösungen bezeichnet man als *Sonderlösungen*.

4.4.1 Beispiel $y' = y$

Hier ist also $f(y) = y$. Überlegen wir uns zuerst die Sonderlösungen, also die Fälle, in denen $f(y) = 0$. Dies gilt nur für $y = 0$,

$$y(x) = 0$$

ist also die einzige Sonderlösung der untersuchten Differenzialgleichung.

Sei nun $f(y) \neq 0$, d.h. $y \neq 0$. Nach unserem Lösungsansatz (15) erhalten wir

$$x(y) = \int \frac{1}{f(y)}\, dy = \int \frac{1}{y}\, dy = \ln|y| + C_1 \ .$$

Dr. Peter Bauer
SS 25 — 4. Juli 2025
Rev: 14158

Dies müssen wir nach y auflösen:

$$x = \ln|y| + C_1 \quad \Rightarrow \quad e^x = e^{\ln|y|+C_1} = |y| \cdot e^{C_1} \quad \Rightarrow \quad |y| = e^{-C_1} \cdot e^x \quad \Rightarrow \quad y = \pm e^{-C_1} \cdot e^x \; .$$

$\pm e^{-C_1}$ ist dabei auch nur eine reelle Konstante $C \neq 0$, als Lösungsschar erhalten wir also, wenn wir die obige Sonderlösung hinzunehmen,

$$y(x) = 0 \qquad \text{oder} \qquad y(x) = Ce^x \quad \text{mit } C \neq 0.$$

Eine partikuläre Lösung, zum Beispiel für den Anfangswert $y(1) = 1$, können wir wieder durch Einsetzen in die Lösungsschar finden:

$$y(1) = Ce^1 = 1 \qquad \Rightarrow \qquad C = \frac{1}{e} \; ,$$

die Lösung des Anfangswertproblems ist also

$$y(x) = \frac{1}{e} \cdot e^x = e^{x-1} \; .$$

Wäre andererseits eine Lösung mit $y(1) = 0$ gesucht, dann wäre die Sonderlösung $y(x) = 0$ für alle x die partikuläre Lösung.

4.4.2 Beispiel In Beispiel 4.3.2 haben wir bereits die Reaktionsgeschwindigkeit nullter Ordnung betrachtet. Zerfallsprozesse (also Reaktionen der Art A $\longrightarrow$ B + C) gehorchen dagegen einem Geschwindigkeitsgesetz 1. Ordnung.

Hier ist die Reaktionsgeschwindigkeit proportional von der Konzentration des zerfallenden Stoffs abhängig. Ist $c_A(t)$ die Konzentration von A, gilt also

$$c_A' = k \cdot c_A \; .$$

Dies ist eine autonome Differenzialgleichung. Es ist natürlich $kc_A \neq 0$, sonst würde nichts zerfallen. Daher gibt es keine Sonderlösungen.

Unser Ansatz lautet also

$$t = \int \frac{1}{kc_A} dc_A = \frac{1}{k} \ln|c_A| + C_1 \; .$$

Die Konzentration ist natürlich nicht negativ, also ist $|c_A| = c_A$ und lösen wir dies nach c_A auf, erhalten wir

$$t = \frac{1}{k} \ln c_A + C_1 \quad \Leftrightarrow \quad k(t - C_1) = \ln c_A \quad \Leftrightarrow \quad c_A = e^{k(t-C_1)} = e^{-kC_1} \cdot e^{kt} = C \cdot e^{kt}$$

mit einer Konstanten $C = e^{-kC_1}$. Diese Konstante lässt sich recht leicht bestimmen:

$$c_A(0) = C \cdot e^0 = C \; ,$$

es handelt sich also gerade um die Anfangskonzentration $c_A(0)$ und insgesamt wird die Reaktionsgeschwindigkeit beschrieben durch

$$c_A(t) = c_A(0)e^{kt} \; .$$

4.5 Differenzialgleichungen mit getrennten Variablen

Typ (C) Differenzialgleichung mit „getrennten Variablen“ $y' = h(x) \cdot g(y)$

Auch hier rechnen wir wieder „formal“: Für $g(y) \neq 0$ gilt

$$\frac{dy}{dx} = h(x) \cdot g(y) \quad \Leftrightarrow \quad \frac{dy}{g(y)} = h(x) \cdot dx \quad \Leftrightarrow \quad \boxed{\int \frac{1}{g(y)}\, dy = \int h(x)\, dx} \; . \tag{16}$$

Da hier die Variablen x und y getrennt auf den beiden Seiten des „=“ stehen, so dass man die Gleichung integrieren kann, spricht man bei diesem Verfahren von einer *Trennung der Variablen.*

An den Nullstellen $g(y_0) = 0$ haben wir dagegen die Sonderlösungen

$$y(x) = y_0 \ .$$

4.5.1 Beispiel $y' = -\frac{x}{y}$

Das Richtungsfeld dieser Differenzialgleichung kennen wir ja schon aus Abbildung 24.

Wir haben hier eine Differenzialgleichung mit getrennten Variablen, denn es ist $y' = -x \cdot \frac{1}{y}$, also $h(x) = -x$ und $g(y) = \frac{1}{y}$.

Da stets $g(y) \neq 0$, gibt es hier keine Sonderlösungen.

Der Lösungsansatz (16) liefert uns

$$\int y\,dy = \int -x\,dx \quad \Rightarrow \quad \frac{y^2}{2} + C_1 = -\frac{x^2}{2} + C_2 \quad \Rightarrow \quad y^2 = -x^2 + C \ ,$$

wobei wir die beiden Integrationskonstanten C_1 und C_2 mit $C = 2(C_2 - C_1)$ zu einer Konstanten zusammengefasst haben. Es folgt also

$$y(x) = \pm\sqrt{C - x^2} \ .$$

Suchen wir die partikuläre Lösung zu $y(0) = 2$ (vgl. Abbildung 24), finden wir

$$y(0) = \pm\sqrt{C - 0} = 2 \quad \Rightarrow \quad C = 4 \ ,$$

und so

$$y(x) = \sqrt{4 - x^2} \ .$$

4.6 Lineare Differenzialgleichungen

Typ (D) Lineare Differenzialgleichung $y' = g(x) \cdot y + r(x)$

4.6.1 Bezeichnungen Eine lineare Differenzialgleichung $y' = g(x) \cdot y + r(x)$ nennt man *homogen*, wenn $r(x) = 0$ für alle x, ansonsten *inhomogen*.

Der Term $r(x)$ wird *Inhomogenität* oder *Störung* genannt.

1. Fall: homogene lineare Differenzialgleichung $y' = g(x) \cdot y$

Dies ist eine Differenzialgleichung mit getrennten Variablen vom Typ (C) mit $h(y) = y$. Sonderlösungen erhalten wir für $h(y_0) = y_0 = 0$, also ist

$$y(x) = 0 \qquad \text{für alle } x$$

die einzige Sonderlösung. Die allgemeine Lösung ergibt sich aus (16): Sei $G(x)$ eine Stammfunktion von $g(x)$, dann ist

$$\begin{aligned} & \int \frac{1}{y}\,dy = \int g(x)\,dx \\ \Rightarrow \quad & \ln|y| + C_1 = G(x) + C_2 \\ \Rightarrow \quad & |y| = e^{G(x)+C_2-C_1} = e^{G(x)} \cdot e^{C_2-C_1} \\ \Rightarrow \quad & y = C \cdot e^{G(x)} \qquad \text{mit } C = \pm e^{C_2-C_1} \neq 0. \end{aligned}$$

Die gefundene Sonderlösung $y(x) = 0$ entspricht $C = 0$ in der allgemeinen Lösung, die Lösungsschar einer homogenen linearen Differenzialgleichung ist also

$$\boxed{y(x) = C \cdot e^{G(x)}} \tag{17}$$

mit $C \in \mathbb{R}$ und wobei $G(x)$ eine Stammfunktion von $g(x)$ ist.

4.6.2 Beispiel Wir schauen uns ein Beispiel für eine homogene lineare Differenzialgleichung an:

$$y' = 3x^2 y$$

Hier ist $g(x) = 3x^2$. Nach (17) ist die Lösungsschar $y(x) = Ce^{G(x)}$. Als Stammfunktion von $g(x)$ wählen wir $G(x) = x^3$, die Lösungsschar ist also

$$y(x) = C \cdot e^{x^3} \ .$$

Haben wir beispielsweise als Anfangswert $y(-1) = 1$ gegeben, gilt

$$y(-1) = C \cdot e^{(-1)^3} = 1 \quad \Rightarrow \quad C = e$$

und erhalten deshalb als partikuläre Lösung $y(x) = e \cdot e^{x^3} = e^{x^3+1}$.

2. Fall: inhomogene lineare Differenzialgleichung $y' = g(x) \cdot y + r(x)$

Die Lösungsschar einer inhomogenen Differenzialgleichung setzt sich aus der Lösungsschar der zugehörigen homogenen Differenzialgleichung und einer partikulären Lösung der inhomogenen Differenzialgleichung zusammen:

4.6.3 Satz Die Lösungsschar einer inhomogenen linearen Differenzialgleichung $y' = g(x) \cdot y + r(x)$ ist

$$y(x) = y_{\text{hom}}(x) + y_{\text{part}}(x) \ ,$$

wobei

- $y_{\text{hom}}(x)$ die Lösungsschar der homogenen linearen Differenzialgleichung $y' = g(x) \cdot y$
- $y_{\text{part}}(x)$ eine (beliebige) partikuläre Lösung der linearen Differenzialgleichung

darstellt. Als $y_{\text{part}}(x)$ kann jede Funktion der Schar

$$\int \frac{r(x)}{y_{\text{hom}}(x)} \, dx \cdot y_{\text{hom}}(x)$$

gewählt werden.

4.6.4 Beispiel Ein Beispiel für eine inhomogene lineare Differenzialgleichung wäre:

$$y' = 3x - \frac{y}{x} = \underbrace{-\frac{1}{x}}_{=g(x)} \cdot y + \underbrace{3x}_{=r(x)}$$

Die Lösung finden wir in zwei Schritten: zuerst lösen wir die zugehörige homogene Differenzialgleichung $y' = -\frac{1}{x} \cdot y$.

$-\ln x$ ist eine Stammfunktion von $-\frac{1}{x}$ und daher ist

$$y_{\text{hom}}(x) = C \cdot e^{-\ln x} = C \cdot x^{-1} \ .$$

Im zweiten Schritt berechnen wir eine partikuläre Lösung:

$$\begin{aligned} \int \frac{r(x)}{y_{\text{hom}}(x)} \, dx \cdot y_{\text{hom}}(x) &= \int \frac{3x}{C \cdot x^{-1}} \, dx \cdot Cx^{-1} \\ &= 3 \int x^2 \, dx \cdot x^{-1} \\ &= \left(x^3 + C_1\right) \cdot x^{-1} \\ &= x^2 + \frac{C_1}{x} \end{aligned}$$

Mit $C_1 = 0$ erhalten wir zum Beispiel $y_{\text{part}}(x) = x^2$ als eine partikuläre Lösung, die Lösungsschar unserer inhomogenen linearen Differenzialgleichung ist also

$$y(x) = C \cdot x^{-1} + x^2 \ .$$

4.6.5 Beispiel Überlegen wir uns die Situation beim Sprung mit einem Fallschirm: Bekanntlich genügt die Kraft, die auf einen beschleunigten Körper wirkt, dem Gesetz „Masse mal Beschleunigung“ $m \cdot a$.

Andererseits wirkt auf einen Fallschirmspringer die Erdanziehungskraft $m \cdot g$ durch die Erdbeschleunigung g, die reduziert wird um die Bremswirkung durch den Luftwiderstand des Fallschirms. Diese bremsende Kraft ist proportional zur Geschwindigkeit v, mit einem Reibkoeffizient $k > 0$ als Proportionalitätskonstante, die Bremskraft ist also $k \cdot v$.

Insgesamt haben wir

$$m \cdot a = m \cdot g - k \cdot v \ .$$

Die Beschleunigung a ist ja gerade die Veränderung der Geschwindigkeit v, also $v' = a$. Damit erhalten wir

$$mv' = mg - kv \qquad \Leftrightarrow \qquad v' = -\frac{k}{m}\, v + g \tag{18}$$

und dies ist eine inhomogene lineare Differenzialgleichung.

Wir bestimmen wieder zuerst die Lösungsschar der zugehörigen homogenen Differenzialgleichung $v' = -\frac{k}{m}\, v$. $G(t) = -\frac{k}{m}\, t$ ist eine Stammfunktion von $-\frac{k}{m}$, also ist die Lösungsschar der homogenen Differenzialgleichung

$$v_{\text{hom}}(t) = C \cdot e^{G(t)} = Ce^{-kt/m} \ .$$

Um eine partikuläre Lösung der inhomogenen Differenzialgleichung zu finden, berechnen wir

$$\begin{aligned}
\int \frac{g}{v_{\text{hom}}(t)}\, dt \cdot v_{\text{hom}}(t) &= \int \frac{g}{e^{-kt/m}}\, dt \cdot e^{-kt/m} \\
&= g \int e^{kt/m}\, dt \cdot e^{-kt/m} \\
&= g \left(\frac{m}{k} e^{kt/m} + C_1 \right) \cdot e^{-kt/m} \\
&= \frac{gm}{k} + C_1 g e^{-kt/m}
\end{aligned}$$

Mit $C_1 = 0$ erhalten wir $v_{\text{part}}(t) = \frac{gm}{k}$ und so ist die Lösungsschar von (18)

$$v(t) = Ce^{-kt/m} + \frac{gm}{k} \ .$$

Eine partikuläre Lösung erhalten wir, wenn wir annehmen, dass der Springer beim Öffnen des Fallschirms (Zeitpunkt $t = 0$) die Geschwindigkeit v_0 hatte. Dann ist

$$v(0) = Ce^0 + \frac{gm}{k} = v_0 \qquad \Rightarrow \qquad C = v_0 - \frac{gm}{k} \ ,$$

also ist

$$v(t) = \left(v_0 - \frac{gm}{k} \right) e^{-kt/m} + \frac{gm}{k} \ .$$

Je länger der Springer am Schirm hängt, desto kleiner wird der erste Summand; es gilt $\lim\limits_{t \to \infty} v(t) = \frac{gm}{k}$. Egal wie schnell der Springer also beim Öffnen des Fallschirms war, hängt er nur lange genug am Schirm, nähert sich seine Geschwindigkeit dem Wert $\frac{gm}{k}$. Ist der Reibkoeffizient k groß genug, ist diese Geschwindigkeit hinreichend klein, die Landung unbeschadet zu überstehen.

5 Exkurs: Differenzialgleichungen und Schwingungen

5.1 Lineare Differenzialgleichungen zweiter Ordnung

Bei einer *Differenzialgleichung zweiter Ordnung* tritt neben y und y' auch die zweite Ableitung y'' auf. Ähnlich wie die bereits untersuchten Differenzialgleichungen erster Ordnung haben auch Differenzialgleichungen zweiter Ordnung oft unendlich viele Lösungen, also eine *Lösungsschar*. *Partikuläre Lösungen* gibt es auch hier zu *Anfangswertproblemen*, die nun aber zwei Gleichungen benötigen: Entweder werden sie durch

$$y(a_1) = b_1 \qquad \text{und} \qquad y(a_2) = b_2$$

für $a_1 \neq a_2$ angegeben oder durch

$$y(a_1) = b_1 \qquad \text{und} \qquad y'(a_2) = b_2 \ .$$

Differenzialgleichungen zweiter (oder höherer) Ordnung können im allgemeinen nur noch numerisch gelöst werden. Eine Ausnahme bilden dabei lineare Differenzialgleichungen zweiter Ordnung, für die es Lösungsansätze gibt:

5.1.1 Bezeichnungen Eine *lineare Differenzialgleichung zweiter Ordnung* hat die allgemeine Form

$$a_2(x)y'' + a_1(x)y' + a_0(x)y = b(x) \ .$$

Ist $b(x) = 0$ für alle x aus dem Definitionsbereich der Differenzialgleichung, dann heißt die Differenzialgleichung *homogen*, sonst *inhomogen*.

Homogene lineare Differenzialgleichungen zweiter Ordnung

Wichtig ist hier ein Begriff, den wir in ähnlicher Form bereits bei Vektoren kennen gelernt haben:

5.1.2 Definition Seien $f, g : D \to \mathbb{R}$ mit $D \subset \mathbb{R}$ zwei Funktionen. f und g heißen *linear unabhängig*, wenn für $\alpha_1, \alpha_2 \in \mathbb{R}$ die Gleichung

$$\alpha_1 f(x) + \alpha_2 g(x) = 0 \qquad \text{für alle } x \in D$$

nur für $\alpha_1 = \alpha_2 = 0$ erfüllt werden kann.

Damit können wir nun die Lösungsschar einer homogenen linearen Differenzialgleichung zweiter Ordnung konstruieren:

5.1.3 Satz Zu einer homogenen linearen Differenzialgleichung zweiter Ordnung mit stetigen, auf einem gemeinsamen Intervall definierten Koeffizientenfunktionen $a_0(x)$, $a_1(x)$ und $a_2(x)$ lassen sich stets zwei linear unabhängige Lösungen $f(x)$ und $g(x)$ finden, so dass die allgemeine Lösungsschar durch

$$y(x) = c_1 f(x) + c_2 g(x) \qquad \text{mit } c_1, c_2 \in \mathbb{R}$$

gegeben ist. f und g bezeichnet man dann als *Fundamentalsystem*.

Zur Prüfung der linearen Unabhängigkeit zweier Funktionen dient folgendes Kriterium:

5.1.4 Satz Seien f und g Lösungen einer homogenen linearen Differenzialgleichung. Dann ist die *Wronski-Determinante* von f und g

$$W(x) = \det \begin{pmatrix} f(x) & g(x) \\ f'(x) & g'(x) \end{pmatrix}$$

entweder gleich 0 für alle x oder ungleich 0 für alle x. Ist $W(x) \neq 0$, dann sind $f(x)$ und $g(x)$ linear unabhängig, bilden also ein Fundamentalsystem.

5.1.5 Beispiel Betrachten wir die Differenzialgleichung $y'' - 2y' - 3y = 0$.

Man kann leicht nachrechnen, dass $y_1(x) = e^{3x}$ und $y_2(x) = e^{-x}$ Lösungen dieser Differenzialgleichung sind. (Wie wir diese finden, sehen wir später.) Es gilt

$$W(x) = \det \begin{pmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{pmatrix} = \det \begin{pmatrix} e^{3x} & e^{-x} \\ 3e^{3x} & -e^{-x} \end{pmatrix} .$$

Da $W(x)$ entweder immer $= 0$ oder immer $\neq 0$ ist, genügt es, $W(x)$ an einer beliebigen Stelle zu berechnen, z.B. in $x = 0$:

$$W(0) = \det \begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix} = -1 - 3 = -4 \neq 0 ,$$

y_1 und y_2 bilden also ein Fundamentalsystem und die Lösungsschar ist daher

$$y(x) = c_1 y_1(x) + c_2 y_2(x) = c_1 e^{3x} + c_2 e^{-x} \qquad \text{mit } c_1, c_2 \in \mathbb{R}.$$

Homogene lineare Differenzialgleichungen zweiter Ordnung mit konstanten Koeffizienten

Im Spezialfall, dass die Koeffizientenfunktionen $a_0(x)$, $a_1(x)$ und $a_2(x)$ konstant sind, also gar nicht von x abhängen, können wir das Fundamentalsystem, das die Differenzialgleichung löst, recht leicht angeben. (Das obige Beispiel gehört in diesen Fall, denn da ist $a_1(x) = -2$ und $a_0(x) = -3$.)

5.1.6 Satz Sei $a_2 y'' + a_1 y' + a_0 y = 0$ eine homogene lineare Differenzialgleichung mit konstanten Koeffizienten. Dann ist $D = a_1^2 - 4a_2 a_0$ die *Diskriminante* der Differenzialgleichung und es gilt:

1. Fall: D > 0 Mit $\lambda_1 = -\frac{a_1 + \sqrt{D}}{2a_2}$ und $\lambda_2 = -\frac{a_1 - \sqrt{D}}{2a_2}$ bilden

$$f(x) = e^{\lambda_1 x} \qquad \text{und} \qquad g(x) = e^{\lambda_2 x}$$

ein Fundamentalsystem.

2. Fall: D < 0 Mit $\lambda = -\frac{a_1}{2a_2}$ und $\beta = \frac{\sqrt{-D}}{2a_2}$ bilden

$$f(x) = e^{\lambda x} \cos \beta x \qquad \text{und} \qquad g(x) = e^{\lambda x} \sin \beta x$$

ein Fundamentalsystem.

3. Fall: D = 0 Mit $\lambda = -\frac{a_1}{2a_2}$ bilden

$$f(x) = e^{\lambda x} \qquad \text{und} \qquad g(x) = x e^{\lambda x}$$

ein Fundamentalsystem.

In allen drei Fällen ist $y(x) = c_1 f(x) + c_2 g(x)$ mit $c_1, c_2 \in \mathbb{R}$ die allgemeine Lösungsschar.

5.1.7 Beispiel Betrachten wir die Differenzialgleichung aus Beispiel 5.1.5:

$$y'' - 2y' - 3y = 0$$

also $a_2 = 1$, $a_1 = -2$ und $a_0 = -3$. Es folgt

$$D = (-2)^2 - 4 \cdot 1 \cdot (-3) = 16 \quad \Rightarrow \quad \lambda_{1,2} = -\frac{-2 \pm \sqrt{16}}{2 \cdot 1} = \begin{cases} -1 \\ 3 \end{cases} \quad .$$

e^{-x} und e^{3x} bilden nach unserem Satz also ein Fundamentalsystem und es ist klar, wie die Lösungen aus Beispiel 5.1.5 gefunden wurden.

Inhomogene lineare Differenzialgleichungen zweiter Ordnung

Die Lösungsschar einer inhomogenen linearen Differenzialgleichung zweiter Ordnung

$$y'' + a_1(x)y' + a_0(x)y = b(x)$$

kann ähnlich zu inhomogenen linearen Differenzialgleichung erster Ordnung bestimmt werden:

(1) Wir ermitteln die allgemeine Lösung y_{hom} der homogenen Differenzialgleichung

$$y'' + a_1(x)y' + a_0(x)y = 0 \ .$$

(2) Wir ermitteln (dies ist leider oft schwierig) eine partikuläre Lösung y_{part} der inhomogenen Differenzialgleichung.

(3) Die allgemeine Lösung der inhomogenen Differenzialgleichung ist dann gegeben durch

$$y(x) = y_{\text{hom}}(x) + y_{\text{part}}(x) \ .$$

Klassische Beispiele für solche Differenzialgleichungen sind Schwingungen, mit denen wir uns im folgenden Abschnitt befassen.

5.2 Schwingungen

Schwingungen liegen vielen physikalischen und anderen naturwissenschaftlichen Betrachtungen zugrunde. Wir betrachten hier als Beispiel ein *Federpendel*, bei dem eine Masse an einer Feder hängt. Andere Schwingungen werden aber genauso behandelt, beispielsweise hätte ein elektrischer Schwingkreis aus Kondensator und Spule die gleiche mathematische Beschreibung (und Lösung).

Im folgenden wollen wir uns die mathematische Modellierung anschauen.

Ungedämpfte Schwingung

Ein einfacher Fall ist die *ungedämpfte* Schwingung:

Eine Masse m hängt an einer Feder und wird aus der Ruhelage um die Strecke $s_0 = s(0)$ ausgelenkt. $s(t)$ ist jeweils die Auslenkung zur Zeit t.

Zu jedem Zeitpunkt t wirkt die Kraft $F(t)$ der Feder proportional zur Auslenkung $s(t)$ mit einer *Federkonstanten* $k > 0$:

$$F(t) = -k \cdot s(t)$$

($F(t)$ wirkt entgegen der Auslenkung, daher das negative Vorzeichen).

Nach dem Newtonschen Gesetz “Kraft ist Masse mal Beschleunigung” ist andererseits

$$F(t) = m \cdot s''(t)$$

(die Beschleunigung ist die Ableitung der Geschwindigkeit, die Geschwindigkeit $s'(t)$ ist die Ableitung des Weges). Insgesamt ist also $ms''(t) = -ks(t)$ und damit

$$ms''(t) + ks(t) = 0 \ . \tag{19}$$

Dies ist die *Schwingungsdifferentialgleichung einer ungedämpften Schwingung.*

Hier ist also $a_2 = m$, $a_1 = 0$ und $a_0 = k$. Damit ist $D = -4mk < 0$ (2. Fall aus Satz 5.1.6) und wir können das Fundamentalsystem berechnen:

Es ist $\lambda = -\dfrac{a_1}{2a_2} = 0$ und $\beta = \dfrac{\sqrt{-D}}{2a_2} = \dfrac{\sqrt{4mk}}{2m} = \sqrt{\dfrac{k}{m}}$.

β wird auch *Winkelgeschwindigkeit* oder *Kreisfrequenz* genannt und mit ω_0 bezeichnet. Die *Periodendauer* ist $T = \frac{2\pi}{\omega_0}$ und die *Eigenfrequenz* der Schwingung ist $f = \frac{\omega_0}{2\pi}$.

$f(t) = \cos \omega_0 t$ und $g(t) = \sin \omega_0 t$ bilden also ein Fundamentalsystem und die allgemeine Lösung der Schwingungsdifferentialgleichung ist

$$s(t) = c_1 \cos \omega_0 t + c_2 \sin \omega_0 t \quad \text{mit } \omega_0 = \sqrt{\frac{k}{m}} \ .$$

5.2.1 Beispiel Betrachten wir das Pendel ohne zusätzlichen Anfangsimpuls, d.h. $s'(0) = 0$, und mit einer Anfangsauslenkung $s(0) = s_0$. Das Anfangswertproblem ist hier also

$$s(0) = s_0 \quad \text{und} \quad s'(0) = 0 \ .$$

Es ist $s'(t) = -c_1\omega_0 \sin \omega_0 t + c_2\omega_0 \cos \omega_0 t$, also folgt aus $s'(0) = c_2\omega_0 = 0$, dass $c_2 = 0$. Es folgt weiter $s(t) = c_1 \cos \omega_0 t$, insbesondere $s(0) = c_1 = s_0$ und insgesamt die partikuläre Lösung

$$s(t) = s_0 \cos \omega_0 t \ .$$

5.3 Gedämpfte Schwingungen

In der Praxis werden Schwingungen meist durch Reibung, Luftwiderstand o.ä. gedämpft. Diese Dämpfung ist proportional zur Geschwindigkeit, mit der die Masse sich jeweils bewegt. Beschrieben wird dies durch

$$ms''(t) + \rho s'(t) + ks(t) = 0 \qquad \text{mit einer } \textit{Dämpfungskonstanten } \rho > 0.$$

Dies entspricht der Differenzialgleichung einer ungedämpften Schwingung (19), es kommt jedoch noch eine *Dämpfung* hinzu, die proportional zur Geschwindigkeit ist. Diese Dämpfung ist also $\rho s'(t)$.

Den Wert $\delta = \dfrac{\rho}{2m}$ nennt man *Abklingkonstante*.

Die Diskriminante dieser Differenzialgleichung ist $D = a_1^2 - 4a_2a_0 = \rho^2 - 4mk$. Hier können nun alle drei Fälle auftreten: $D > 0$, $D < 0$ und $D = 0$.

1. Fall $D > 0 \Leftrightarrow \rho^2 - 4mk > 0$. Mit $\delta = \frac{\rho}{2m} \Leftrightarrow \rho = 2\delta m$ folgt daraus

$$4\delta^2 m^2 - 4mk > 0 \quad \Leftrightarrow \quad \delta^2 m^2 > mk \quad \Leftrightarrow \quad \delta > \sqrt{\frac{k}{m}} = \omega_0 \ .$$

Es ist also $D > 0 \Leftrightarrow \delta > \omega_0$, die Abklingkonstante ist größer als die Winkelgeschwindigkeit. Man spricht in diesem Fall von einer *starken Dämpfung.*

Die Lösung der Schwingungsdifferentialgleichung ist nach Satz 5.1.6

$$s(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t} ,$$

wobei

$$\lambda_{1,2} = -\frac{a_1 \pm \sqrt{D}}{2a_2} = -\frac{\rho \pm \sqrt{\rho^2 - 4mk}}{2m} = -\frac{\rho}{2m} \pm \sqrt{\left(\frac{\rho}{2m}\right)^2 - \frac{k}{m}} = -\delta \pm \sqrt{\delta^2 - \omega_0^2}$$

Ein Beispiel für eine solche Funktion ist in Abbildung 28 dargestellt. Die starke Dämpfung sorgt dafür, dass das Pendel gar nicht hin- und her, sondern in die Ruhelage zurück-„schwingt".

2. Fall $D < 0 \Leftrightarrow \delta < \omega_0$ (Diese Äquivalenz errechnet sich analog zum ersten Fall). Hier ist die Abklingkonstante kleiner als die Winkelgeschwindigkeit, dies ist eine *schwach gedämpfte* Schwingung.

Hier ist $\lambda = -\frac{a_1}{2a_2} = -\frac{\rho}{2m} = -\delta$ und $\beta = \frac{\sqrt{-D}}{2a_2} = \sqrt{\frac{k}{m} - \frac{\rho^2}{4m^2}} = \sqrt{\omega_0^2 - \delta^2}$.

Mit $\omega_\mathrm{d} = \sqrt{\omega_0^2 - \delta^2}$ lautet die allgemeine Lösungsschar also

$$s(t) = e^{-\delta t} \left(c_1 \cos \omega_\mathrm{d} t + c_2 \sin \omega_\mathrm{d} t\right) .$$

Eine solche Funktion ist in Abbildung 29 dargestellt.

3. Fall Der Fall $D = 0 \Leftrightarrow \delta = \omega_0$ ist ein Grenzfall zwischen schwacher und starker Dämpfung. Wie im zweiten Fall ist $\lambda = -\frac{a_1}{2a_2} = -\frac{\rho}{2m} = -\delta$ und die Lösungen sind

$$s(t) = c_1 e^{-\delta t} + c_2 t e^{-\delta t} = (c_1 + c_2 t) e^{-\delta t} .$$

Abbildung 30 zeigt eine derartige Funktion.

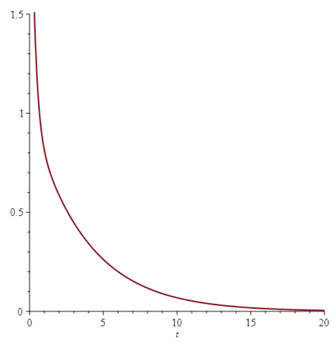


Abbildung 28: Schwingung mit starker Dämpfung

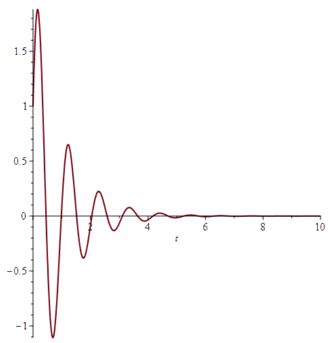


Abbildung 29: Schwingung mit schwacher Dämpfung

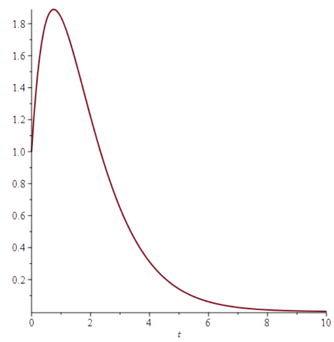


Abbildung 30: Grenzfall zwischen schwacher und starker Dämpfung

Inhaltsverzeichnis

Die Abbildungen 6, 11 und 23 sind Wikimedia Commons entnommen (Public Domain).