

# Ingenieurmathematik II

Olaf Ippisch  
Institut für Mathematik  
TU Clausthal  
Erzstr. 1  
D-38678 Clausthal-Zellerfeld  
E-mail: [olaf.ippisch@tu-clausthal.de](mailto:olaf.ippisch@tu-clausthal.de)

11. Juli 2017

# Inhaltsverzeichnis

<b>1. Einführung</b>	<b>1</b>
1.1. Ingenieurmathematik II . . . . .	3
1.1.1. Anwendungen . . . . .	3
1.1.2. Beispiel: Simulation von Temperaturverteilung in einem Werkstück . . . . .	5
1.1.3. Lineare Algebra . . . . .	7
1.1.4. Funktionen mehrerer Variablen . . . . .	8
1.2. Ausblick: Ingenieurmathematik III . . . . .	8
1.3. Ausblick: Ingenieurmathematik IV . . . . .	9
<b>I. Lineare Algebra</b>	<b>11</b>
<b>2. Matrizen und Vektoren</b>	<b>11</b>
2.1. Definitionen . . . . .	12
2.2. Rechenregeln für Matrizen . . . . .	14
2.3. Lineare Gleichungssysteme . . . . .	20
2.4. Gauß Algorithmus . . . . .	21
<b>3. Vektorraum</b>	<b>27</b>
<b>4. Determinanten</b>	<b>35</b>
<b>5. Lineare Gleichungssysteme revisited</b>	<b>41</b>
5.1. Spaltenraum . . . . .	41
5.2. Bestimmung einer Basis von Kern $\mathbf{A}$ (Nullraum) . . . . .	42
5.3. Bestimmung einer speziellen Lösung von $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ . . . . .	43
5.4. Vollständige Charakterisierung eines linearen Gleichungssystems . . . . .	43
5.5. Schlussbemerkung . . . . .	47
<b>6. Inverse Matrix</b>	<b>49</b>
<b>7. Längen und Winkel im <math>\mathbb{R}^n</math></b>	<b>53</b>
7.1. Skalarprodukt . . . . .	53
7.2. Normen . . . . .	54
7.3. Winkel im $\mathbb{R}^n$ , Orthogonalität . . . . .	57
7.4. Methode der kleinsten Quadrate . . . . .	62
7.5. Vektorprodukt . . . . .	65
7.6. Tensorprodukt . . . . .	66
<b>8. Lineare Abbildungen (Bonusmaterial)</b>	<b>67</b>
<b>9. Eigenwerte und Eigenvektoren</b>	<b>71</b>
9.1. Definition . . . . .	71
9.2. Berechnung von Eigenwerten . . . . .	71
9.3. Spezielle Matrizen . . . . .	74
9.4. Anwendung . . . . .	75

<b>II. Analysis von Funktionen in mehreren Variablen</b>	<b>83</b>
<b>10. Ableitung von Funktionen mehrerer Variablen</b>	<b>83</b>
10.1. Motivation . . . . .	83
10.2. Grundlegende Begriffe . . . . .	84
10.3. Grenzwert und Stetigkeit . . . . .	85
10.4. Partielle Ableitung . . . . .	85
10.5. Höhere Ableitungen . . . . .	86
10.6. Totale Ableitung . . . . .	87
10.7. Lokale Extremwerte . . . . .	88
10.8. Richtungsableitung . . . . .	90
10.9. Kurven im $\mathbb{R}^n$ . . . . .	90
10.10. Kettenregel . . . . .	92
10.11. Approximation höherer Ordnung (Taylor-Formel) . . . . .	95
10.12. Optimierung unter Nebenbedingungen . . . . .	96
10.13. Vektorwertige Funktionen . . . . .	100
10.13.1. Differentiation und Jacobi-Matrix . . . . .	100
10.13.2. Kettenregel . . . . .	103
10.13.3. Newtonverfahren im $\mathbb{R}^n$ . . . . .	103
<b>11. Integration von Funktionen mehrerer Variablen</b>	<b>105</b>
11.1. Kurvenintegrale . . . . .	105
11.2. Potential eines Gradientenfelds . . . . .	108
11.3. Flächen- und Volumenintegral . . . . .	109
11.3.1. Fläche und Volumen . . . . .	109
11.3.2. Flächenintegral . . . . .	110
11.3.3. Volumenintegral . . . . .	113
11.4. Variablentransformation für Integrale . . . . .	114
11.5. Oberflächenintegral . . . . .	117
11.6. Divergenz und Rotation . . . . .	121
11.7. Die Integralsätze von Green, Stokes und Gauß . . . . .	122
11.8. Partielle Differentialgleichungen . . . . .	128



# 1. Einführung

## Organisatorisches

### Übungen

- Große Übung (Dr. Franz Hanauska)
  - Dienstags, 13-15 Uhr, Hörsaal A (IfM)
  - Beginn: 11. April
- Tutorien
  - Beginn: 18. April
  - Anmeldung über Stud.IP ab heute, 10.04. um 14:00 Uhr
  - Abgabe der Lösungen der Übungsblätter in Gruppen von bis zu zwei Studierenden jeweils Montag zu Beginn der Vorlesung, für die Studierenden in den Tutorien am Montag von 8-10 Uhr direkt beim Tutor.
  - Korrektur einer Aufgabe pro Übungsblatt (wird erst nach der Abgabe festgelegt).
  - Rückgabe in den Tutorien

### Übungen

- Verpflichtende Hausübungen, erstes Übungsblatt diese Woche
  - Betrifft: Studierende der Studiengänge Verfahrenstechnik, Maschinenbau, Energie und Rohstoffe sowie Energietechnologien
  - Zum Bestehen der Hausübungen sind 40 % der erreichbaren Punkte über alle korrigierten Aufgaben notwendig. Außerdem muss der Studierende mindestens einmal eine Aufgabe im Tutorium vorgerechnet haben.
- Bonussystem
  - Betrifft: Studierende aller Fächer
  - Ein Bonus für die Klausur kann erzielt werden, wenn 60 % der erreichbaren Punkte über alle korrigierten Aufgaben erreicht werden. Außerdem muss der Studierende mindestens einmal eine Aufgabe im Tutorium vorgerechnet haben.
  - Sind die Voraussetzungen des Bonus erreicht, werden Punkte in der Klausur gutgeschrieben, die einer Notenstufe (z.B. von 1,7 auf 1,3) entsprechen.
  - Der Bonus wird nach allgemeiner Prüfungsordnung nur angewendet, wenn die Klausur auch ohne Bonus erreicht worden wäre.
  - Ist die Klausur nur sehr knapp bestanden, reichen die Bonuspunkte manchmal nicht zum Erreichen der 3,7.

*Intensive Beteiligung an den Übungen ist essentiell für den Lernerfolg (und für das Bestehen der Klausur)*

## Klausur

- Samstag, 22. Juli 2017, 9:00-11:00 Uhr
- Eine Beispielklausur wird rechtzeitig zusammen mit der Musterlösung verteilt.
- Erlaubte Hilfsmittel: 1 eigenhändig beschriebenes (nicht kopiertes) DIN A4 Blatt. Vorder- und Rückseite dürfen verwendet werden.
- Taschenrechner ist nicht erlaubt und auch nicht nötig
- Nachklausur Ingenieurmathematik I: 8. Juli 2017, 9:00-11:00 Uhr

## Vorlesungszeit

Beginn der Vorlesung jeweils um 10:15 Uhr und Ende um 11:55 Uhr. Dafür keine Vorlesungen in der letzten Semesterwoche.

## Vorlesungsmitschrift

- Version aus dem SS 2015 bereits im StudIP
- wird während der Vorlesung ergänzt
- enthält den Tafelanschrieb
- upgedatete Version ebenfalls erhältlich über Stud.IP

## Literatur

- Wikipedia
- [Wilhelm Merz, Peter Knabner: Mathematik für Ingenieure und Naturwissenschaftler - Lineare Algebra und Analysis in  \$\mathbb{R}\$ , Springer Spektrum](#)
- [Tilo Arens, Frank Hettlich, Christian Karpfinger, Ulrich Kockelkorn, Klaus Lichtenegger, Hellmuth Stachel: Mathematik, Springer Spektrum](#)
- [Lothar Papula: Mathematik für Ingenieure und Naturwissenschaftler, Band 2, Vieweg + Teubner](#)
- [Begründet von I.N. Bronstein und K.A. Semendjaew, weitergeführt von G. Grosche, V. Ziegler und D. Ziegler Herausgegeben von E. Zeidler: Springer-Taschenbuch der Mathematik, Springer Fachmedien Wiesbaden](#)
- Meyberg, Vachenauer: Höhere Mathematik 1: Differential- und Integralrechnung, Vektor- und Matrizenrechnung, Springer
- Meyberg, Vachenauer: Höhere Mathematik 2: Differentialgleichungen, Funktionentheorie, Fourier-Analyse, Variationsrechnung, Springer

[blaue Farbe: Erhältlich als Ebook in der Uni-Bibliothek](#)

## Programmierung mit Python

- Die Anwendung von Mathematik ist heute häufig mit der Anwendung oder Erstellung von Computerprogrammen verbunden
- Das Erlernen grundlegender Programmierfähigkeiten ist für die Anwendung vieler mathematischer Algorithmen unerlässlich
- Wir werden im Rahmen der Vorlesung Python einsetzen
- Dafür reichen maximal 3 Minuten pro Vorlesung völlig aus
- Python ermöglicht Ihnen teilweise auch die Ergebnisse Ihrer Hausübungen zu überprüfen
- Wir verwenden als Umgebung Pyzo
- Die Installationsanleitung für Pyzo unter Windows, Linux und OS X finden Sie unter <http://www.pyzo.org/start.html>
- Auf der Vorlesungshomepage finden Sie auch eine Kurzeinführung in Python (enthält im wesentlichen die im letzten Semester behandelte Funktionalität)

## Studierendenbeteiligung

Zur Vorlesung

- Rückmeldeblatt in jeder Vorlesung, nur zwei Fragen:
  - Was ist das Wichtigste, das Sie heute gelernt haben?
  - Was haben Sie am wenigsten verstanden?

Kann gerne für weitere Anmerkungen genutzt werden.

- Rückmeldung über Tutoren
- Mail an mich: [olaf.ippisch@tu-clausthal.de](mailto:olaf.ippisch@tu-clausthal.de)

Zu den Übungen

- Rückmeldung über Tutoren
- Mail an Herrn Dr. Hanauska: [franz.hanauska@tu-clausthal.de](mailto:franz.hanauska@tu-clausthal.de)

## 1.1. Ingenieurmathematik II

### 1.1.1. Anwendungen

Reale Probleme verknüpfen in der Regel mehrere Einflussvariablen

⇒ Eine oder mehrere Funktionen mehrerer Variablen involviert

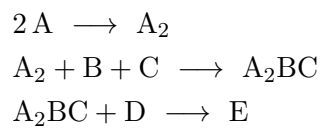
$$f(x_1, x_2, \dots, x_N)$$

oder

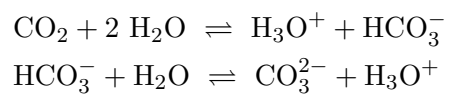
$$\begin{aligned} f_1(x_1, x_2, \dots, x_N) \\ f_2(x_1, x_2, \dots, x_N) \\ \vdots \\ f_N(x_1, x_2, \dots, x_N) \end{aligned}$$

## Chemische Reaktionen

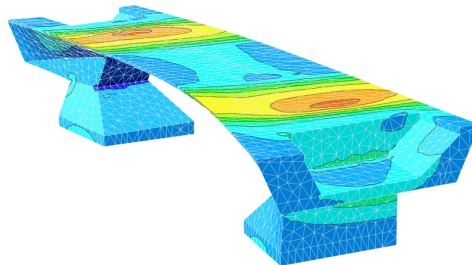
Kinetisch



Gleichgewicht

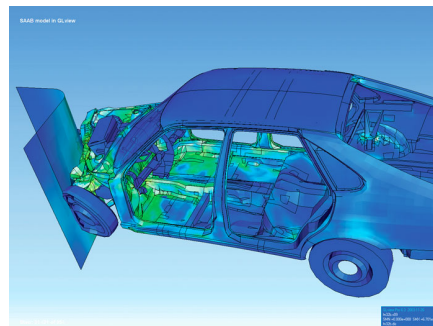


## Baustatik



Quelle: InfoGraph

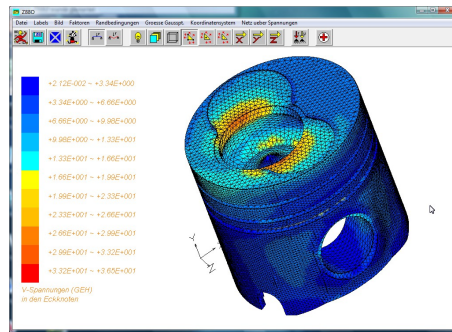
## Crash-Simulation



Quelle: WikiCommons

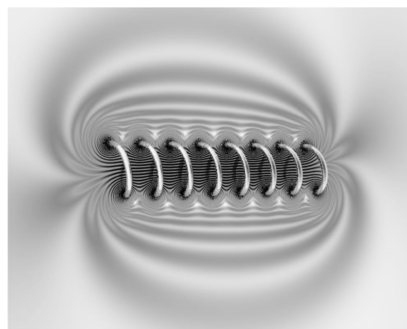


## Mechanische Belastung eines Kolbens



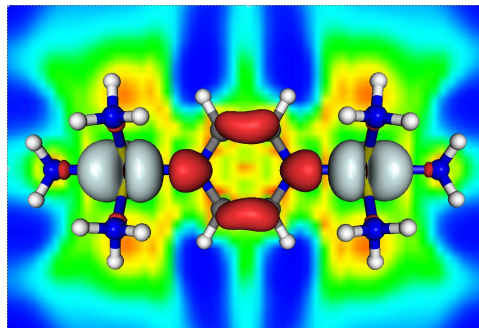
Quelle: Bal 79 auf WikiCommons (CC-BY-SA-3.0)

## Elektrodynamik



Quelle: Svjo auf WikiCommons (CC-BY-SA-3.0)

## Quantenchemie



Quelle: AG Kaupp, Institut für Chemie, TU Berlin

### Frage:

Wie behandelt man das mathematisch?

### 1.1.2. Beispiel: Simulation von Temperaturverteilung in einem Werkstück

#### Wärmeleitungsgleichung

Für die Temperatur  $T$  am Ort  $\mathbf{x}$  zur Zeit  $t$  gilt eine Energieerhaltungsgleichung:

$$\frac{\partial [C_v T(\mathbf{x}, t)]}{\partial t} - \nabla \cdot [\lambda \cdot \nabla T(\mathbf{x}, t)] = 0$$

- $C_v$ : Volumetrische Wärmekapazität
- $\lambda$ : Wärmeleitfähigkeit

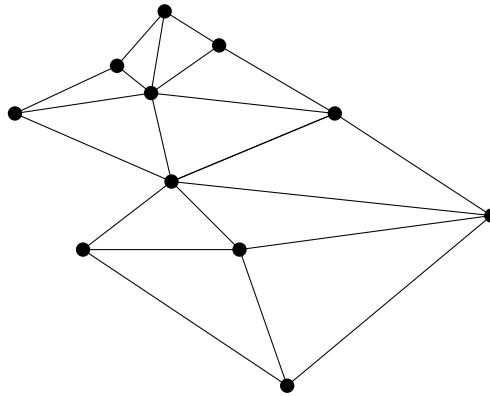
Eindimensional, mit konstanter Wärmeleitfähigkeit  $\lambda$ :

$$\frac{\partial [C_v T(x, t)]}{\partial t} - \lambda \cdot \frac{\partial^2 T(x, t)}{\partial x^2} = 0$$

Dies gilt an jedem Punkt im Raum für alle Zeiten.

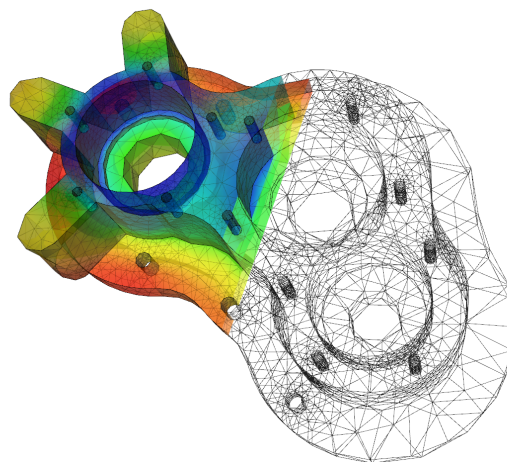
### Diskretisierung

Berechnung von Näherungslösungen durch Diskretisierung von Zeit und Raum:



Dies führt zu einem linearem Gleichungssystem mit  $N$  Gleichungen für  $N$  Unbekannte.

### Simulationsergebnis



Quelle: User A1 auf WikiCommons (CC-BY-SA-3.0)

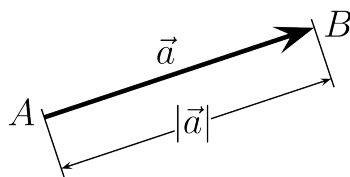
### 1.1.3. Lineare Algebra

#### Vektoren und Matrizen

Vektor

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

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



Quelle: Markus A. Hennig auf WikiCommons (CC-BY-SA-3.0)

Matrix

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

#### Lineare Gleichungssysteme

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

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$$

## Lineare Algebra

- Rechenregeln für Matrizen und Vektoren
- Skalarprodukt und Kreuzprodukt von Vektoren
- Lösung linearer Gleichungssysteme
- Vektorräume
- Eigenwerte und Eigenvektoren
- Inverse Matrizen
- Diagonalisierung
- Lineare Abbildungen

### 1.1.4. Funktionen mehrerer Variablen

$$\frac{\partial [C_v T(x, t)]}{\partial t} - \lambda \cdot \frac{\partial^2 T(x, t)}{\partial x^2} = 0$$

- Partielle Ableitungen
- Gradient, Divergenz und Rotation

$$\frac{\partial [C_v T(\mathbf{x}, t)]}{\partial t} - \nabla \cdot [\lambda \cdot \nabla T(\mathbf{x}, t)] = 0$$

- Ableitungen höherer Ordnung
- Taylorreihe im Mehrdimensionalen: Jacobi- und Hesse-Matrix
- Extremwerte im Mehrdimensionalen
- Integration im Mehrdimensionalen
- Partielle Differentialgleichungen
- Integralsätze von Gauß, Stokes und Green

## 1.2. Ausblick: Ingenieurmathematik III

### Computer (1949)



Quelle: Dryden Flight Research Center Photo Collection, NACA (NASA)

## Supercomputer (2012)



Quelle: Jülich Supercomputing Center (JSC)

- Fließkommazahlen
- Lösung linearer und nichtlinearer Gleichungssysteme
- Interpolation und Extrapolation mit Polynomen
- Trigonometrischen Interpolation, Diskrete Fouriertransformation
- Numerische Differentiation und Integration

### 1.3. Ausblick: Ingenieurmathematik IV

- Numerische Lösung gewöhnlicher Differentialgleichungen
  - Explizite Einschrittverfahren
  - Konvergenz
  - Stabilität
  - Implizite Verfahren
  - Mehrschrittverfahren
- Numerische Lösung partieller Differentialgleichungen
  - Finite-Differenzen
  - Finite-Elemente
  - Instationäre Probleme



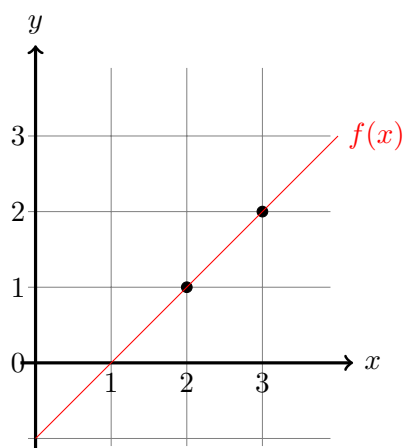
# Teil I.

# Lineare Algebra

## 2. Matrizen und Vektoren

Motivation: Interpolation von Punkten mit Polynomen.

$x_i$	$y_i$
2	1
3	2



Bedingungen für Gerade  $c_1x + c_2$  durch die Punkte  $(x_i, y_i)$ :

$$c_1x_i + c_2 = y_i$$

Also hier:

$$c_1 \cdot 2 + c_2 = 1$$

$$c_1 \cdot 3 + c_2 = 2$$

Kürzer:

$$\underbrace{\begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix}}_{\mathbf{A}} \cdot \underbrace{\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}}_{\mathbf{c}} = \underbrace{\begin{pmatrix} 1 \\ 2 \end{pmatrix}}_{\mathbf{y}}.$$

$\mathbf{A}$  ist eine  $2 \times 2$ -Matrix mit den Zeilen(vektoren)

$$\begin{pmatrix} 2 & 1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 3 & 1 \end{pmatrix}$$

und den Spalten(vektoren)

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

$\mathbf{A}$  hat die vier Komponenten

$$a_{11} = 2, \quad a_{12} = 1, \quad a_{21} = 3, \quad a_{22} = 1.$$

$\mathbf{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$  und  $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$  sind Vektoren (oder  $2 \times 1$  Matrizen).

**Rechenregeln:**

$$\mathbf{A} \cdot \mathbf{c} = \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 2 \cdot c_1 + c_2 \\ 3 \cdot c_1 + c_2 \end{pmatrix}$$

$$\mathbf{A} + \mathbf{A} = \begin{pmatrix} 2+2 & 1+1 \\ 3+3 & 1+1 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 6 & 2 \end{pmatrix},$$

$$2 \cdot \mathbf{A} = \begin{pmatrix} 2 \cdot 2 & 2 \cdot 1 \\ 2 \cdot 3 & 2 \cdot 1 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 6 & 2 \end{pmatrix}.$$

## 2.1. Definitionen

### Definition 2.1 (Matrix)

Ein rechteckiges Schema mit  $m$  Zeilen und  $n$  Spalten (reeller) Zahlen heißt (reelle)  $m \times n$  Matrix. Die Menge aller reellen  $m \times n$  Matrizen bezeichnet man mit  $\mathbb{R}^{m \times n}$ .

### Anmerkung 2.2

- Für Matrizen werden meist (lateinische Großbuchstaben verwendet, z. B.

$$\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$$

- Mit  $a_{ij}, b_{ij}, c_{ij}, \dots$  bezeichnen wir die Zahl (das Matricelement) in der  $i$ -ten Zeile und  $j$ -ten Spalte von  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$ . Man schreibt auch  $(\mathbf{A})_{ij}, (\mathbf{B})_{ij}, (\mathbf{C})_{ij}, \dots$ .
- Analog definiert man die Menge von Matrizen natürlicher Zahlen  $\mathbb{N}^{m \times n}$ , ganzer Zahlen  $\mathbb{Z}^{m \times n}$ , rationaler Zahlen  $\mathbb{Q}^{m \times n}$  und komplexer Zahlen  $\mathbb{C}^{m \times n}$  oder allgemein die Menge der Matrizen über einem Körper  $\mathbb{K}$  als  $\mathbb{K}^{m \times n}$ .

### Definition 2.3 (Vektoren)

Eine  $n \times 1$  Matrix bezeichnet man als Spaltenvektor, eine  $1 \times n$  Matrix als Zeilenvektor. Da Spaltenvektoren eine herausgehobene Rolle spielen, bezeichnet man sie oft einfach als Vektor. Die Menge aller Vektoren mit  $n$  Zeilen ist dann  $\mathbb{R}^n$ .

### Anmerkung 2.4

- Für Vektoren werden meist Kleinbuchstaben verwendet, z.B.

$$\mathbf{a}, \mathbf{b}, \mathbf{c}, \dots, \mathbf{x}, \mathbf{y}$$

- Die Zahl in der  $i$ -ten Zeile (die  $i$ -te Komponente) eines Vektors ist dann  $a_i, b_i, c_i, \dots, x_i, y_i$ .
- In gedruckten Werken werden Matrizen und Vektoren meist fett gedruckt.
- Teilweise wird für Vektoren auch ein Pfeil über den Buchstaben gesetzt:  $\vec{a}, \vec{b}, \vec{c}, \dots, \vec{x}, \vec{y}$



### Beispiel 2.5

$$\mathbf{A} = \begin{pmatrix} 1 & -2 & 1 & 4 & 0 \\ -2 & 5 & 2 & 1 & 2 \end{pmatrix} \in \mathbb{R}^{2 \times 5}$$
$$a_{24} = 1, \quad a_{13} = 1, \quad a_{22} = 5.$$

### Definition 2.6 (Transponierte)

Werden in einer Matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$  die Zeilen und Spalten miteinander vertauscht, erhält man die Transponierte  $\mathbf{A}^T \in \mathbb{R}^{n \times m}$  der Matrix  $\mathbf{A}$ . Damit gilt

$$(\mathbf{A}^T)_{ij} = (\mathbf{A})_{ji} \quad \forall 1 \leq i \leq n, 1 \leq j \leq m$$

### Anmerkung 2.7

- Die Transponierte eines (Spalten)-Vektors ist ein Zeilenvektor.
- $(\mathbf{A}^T)^T = \mathbf{A}$ .
- Sei  $\mathbf{x}$  ein Zeilenvektor aus  $\mathbb{R}^{1 \times n}$ , dann lässt sich das schreiben als  $\mathbf{x}^T \in \mathbb{R}^n$ .

### Beispiel 2.8

Für  $\mathbf{A}$  aus Beispiel 2.5:

$$\mathbf{A}^T = \begin{pmatrix} 1 & -2 \\ -2 & 5 \\ 1 & 2 \\ 4 & 1 \\ 0 & 2 \end{pmatrix},$$
$$\mathbf{x} = \begin{pmatrix} 1 \\ 3 \\ 4 \end{pmatrix} \Leftrightarrow \mathbf{x}^T = (1 \quad 3 \quad 4).$$

### 3-Minutes Python 1

In Python gibt es mit `numpy` ein sehr mächtiges Modul zum Rechnen mit Matrizen und Vektoren. Die grundlegende Datenstruktur von `numpy` ist ein multi-dimensionales Feld. Eine Matrix, einen Zeilenvektor und einen Spaltenvektor legt man damit z.B. folgendermaßen an:

```
import numpy as np
2 A = np.array([[2,4],[1,3]]) # Matrix
b = np.array([1,4])          # Zeilenvektor
4 x = np.array([[1],[4]])     # Spaltenvektor
```

Die Datenelemente eines Arrays können einen beliebigen Typ haben, der sich über das zusätzliche Argument `dtype=` mit folgendem Datentyp festlegen lässt, z.B.

```
>>> A=np.array([[2, 4], [1, 3]],dtype=float)
2 >>> print(A)
[[ 2.  4.]
4  [ 1.  3.]]
```

```

>>> A=np.array([[2, 4], [1, 3]],dtype=int)
6 >>> print(A)
  [[2 4]
 8  [1 3]]
>>> A=np.array([[2, 4], [1, 3]],dtype=complex)
10 >>> print(A)
  [[ 2.+0.j  4.+0.j]
12  [ 1.+0.j  3.+0.j]]

```

Der Zugriff auf Elemente eines Array funktioniert durch Angabe der durch Komma getrennten Indices in einer eckigen Klammer, wobei der erste Index 0 ist:

```

>>> A=np.array([[2, 4], [1, 3]])
2 >>> A[0,0]=7
>>> print(A)
4 [[7 4]
  [1 3]]

```

Die Transponierte eines Array erhält man schließlich durch Nachstellen von .T

```

>>> print(A)
2 [[7 4]
  [1 3]]
4 >>> print(A.T)
  [[7 1]
6  [4 3]]

```

## 2.2. Rechenregeln für Matrizen

### Definition 2.9

Seien  $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$ .

Die Summe der Matrizen  $\mathbf{A} + \mathbf{B}$  ist dann definiert als

$$(\mathbf{A} + \mathbf{B})_{ij} := a_{ij} + b_{ij} \quad \forall 1 \leq i \leq m, 1 \leq j \leq n.$$

Analog ist die Differenz  $\mathbf{A} - \mathbf{B}$  definiert

$$(\mathbf{A} - \mathbf{B})_{ij} := a_{ij} - b_{ij} \quad \forall 1 \leq i \leq m, 1 \leq j \leq n.$$

Die Matrizen sind gleich falls

$$a_{ij} = b_{ij} \quad \forall 1 \leq i \leq m, 1 \leq j \leq n.$$

Das Produkt einer Matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$  mit einem Skalar  $\lambda \in \mathbb{R}$  ist

$$(\lambda \cdot \mathbf{A})_{ij} := \lambda a_{ij} \quad \forall 1 \leq i \leq m, 1 \leq j \leq n.$$

### Beispiel 2.10

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 4 & 2 \end{pmatrix} + 3 \cdot \begin{pmatrix} 2 & 2 & 2 \\ 1 & 4 & 3 \end{pmatrix} \\ = \begin{pmatrix} 1+3 \cdot 2 & 2+3 \cdot 2 & 3+3 \cdot 2 \\ 4+3 \cdot 1 & 4+3 \cdot 4 & 2+3 \cdot 3 \end{pmatrix} = \begin{pmatrix} 7 & 8 & 9 \\ 7 & 16 & 11 \end{pmatrix}.$$

Wie leicht nachzuprüfen ist (komponentenweise Definition der Operationen, Rechenregeln für reelle Zahlen) gilt:

### Satz 2.11

Für  $\lambda, \omega \in \mathbb{R}$  und  $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^{m \times n}$  gilt:

$$\mathbf{A} + \mathbf{B} \in \mathbb{R}^{m \times n}, \quad (2.1)$$

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}, \quad (2.2)$$

$$\mathbf{A} + (\mathbf{B} + \mathbf{C}) = (\mathbf{A} + \mathbf{B}) + \mathbf{C}, \quad (2.3)$$

$$\exists \mathbf{0} \in \mathbb{R}^{m \times n} : \forall \mathbf{A} \in \mathbb{R}^{m \times n} : \mathbf{A} + \mathbf{0} = \mathbf{A}, \quad (2.4)$$

$$\forall \mathbf{A} \in \mathbb{R}^{m \times n} : \exists -\mathbf{A} : \mathbf{A} + (-\mathbf{A}) = \mathbf{0}, \quad (2.5)$$

$$\lambda \cdot \mathbf{A} \in \mathbb{R}^{m \times n}, \quad (2.6)$$

$$\lambda \cdot (\mathbf{A} + \mathbf{B}) = \lambda \cdot \mathbf{A} + \lambda \cdot \mathbf{B}, \quad (2.7)$$

$$(\lambda + \omega) \cdot \mathbf{A} = \lambda \cdot \mathbf{A} + \omega \cdot \mathbf{A}, \quad (2.8)$$

$$\lambda \cdot (\omega \cdot \mathbf{A}) = (\lambda \cdot \omega) \cdot \mathbf{A}, \quad (2.9)$$

$$1 \cdot \mathbf{A} = \mathbf{A}. \quad (2.10)$$

### Anmerkung 2.12

Da diese Bedingungen erfüllt sind, sagt man auch " $\mathbb{R}^{m \times n}$  ist ein Vektorraum" (genauer:  $\mathbb{R}^{m \times n}$  ist ein Vektorraum über  $\mathbb{R}$  bzw. ein reeller Vektorraum).

Analog bilden die komplexen Matrizen  $\mathbb{C}^{m \times n}$  einen Vektorraum über  $\mathbb{C}$ .

### Definition 2.13 (Matrix-Vektor Produkt)

Sei  $\mathbf{A} \in \mathbb{R}^{m \times n}$  und  $\mathbf{x} \in \mathbb{R}^n$ .

Dann ist  $\mathbf{A} \cdot \mathbf{x} \in \mathbb{R}^m$  definiert als

$$(\mathbf{A} \cdot \mathbf{x})_i := \sum_{1 \leq j \leq n} a_{ij} x_j \quad \forall 1 \leq i \leq m.$$

### Beispiel 2.14

1.

$$\begin{pmatrix} 1 & -2 \\ -2 & 5 \end{pmatrix} \cdot \begin{pmatrix} 9 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \cdot 9 - 2 \cdot 4 \\ -2 \cdot 9 + 5 \cdot 4 \end{pmatrix} = \begin{pmatrix} 9 - 8 \\ -18 + 20 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

2.

$$\begin{pmatrix} 2 & 1 \\ 3 & 1 \\ 4 & -2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 + x_2 \\ 3x_1 + x_2 \\ 4x_1 - 2x_2 \end{pmatrix}.$$

**Definition 2.15 (Einheitsmatrix)**

Die Matrix  $\mathbf{I} \in \mathbb{R}^{n \times n}$  mit

$$(\mathbf{I})_{ij} = \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{sonst} \end{cases}$$

wird Einheitsmatrix genannt.

**Beispiel 2.16**

Für  $n = 3$ :

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

**Satz 2.17**

Sei  $\mathbf{I} \in \mathbb{R}^{n \times n}$ ,  $\mathbf{x} \in \mathbb{R}^n$  dann gilt:

$$\mathbf{I} \cdot \mathbf{x} = \mathbf{x}$$

und  $\mathbf{I} \cdot \mathbf{x} = \mathbf{b}$  hat für jedes  $\mathbf{b} \in \mathbb{R}^n$  genau eine Lösung  $\mathbf{x} = \mathbf{b}$ .

**Beweis:**

$$(\mathbf{I} \cdot \mathbf{x})_i = \sum_{1 \leq j \leq n} (\mathbf{I})_{ij} \cdot x_j = x_i.$$

□

**Definition 2.18 (Besondere Matrizen)**

Eine Matrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$  (also mit gleich vielen Zeilen und Spalten) heißt quadratische Matrix.

Eine quadratische Matrix mit

- $a_{ij} = 0$  für  $i \neq j$  heißt Diagonalmatrix
- $a_{ij} = 0$  für  $i < j$  heißt untere Dreiecksmatrix
- $a_{ij} = 0$  für  $i > j$  heißt obere Dreiecksmatrix
- $a_{ij} = a_{ji} \quad \forall 1 \leq i, j \leq n$  heißt symmetrisch. Es gilt  $\mathbf{A}^T = \mathbf{A}$
- $a_{ij} = -a_{ji} \quad \forall 1 \leq i, j \leq n$  heißt schiefsymmetrisch. Dies erfordert, dass  $a_{ii} = 0$  für  $i = j$ .  
Es gilt  $\mathbf{A}^T = -\mathbf{A}$

**Beispiel 2.19**

- $\mathbf{I}$  ist eine Diagonalmatrix, bei der alle Diagonalelemente 1 sind.

- $\begin{pmatrix} 7 & 0 & 0 \\ 3 & 1 & 0 \\ 5 & 0 & 6 \end{pmatrix}$  ist eine untere Dreiecksmatrix.

- $\begin{pmatrix} 7 & 3 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 6 \end{pmatrix}$  ist eine obere Dreiecksmatrix.
- $\begin{pmatrix} 1 & 8 & -7 \\ 8 & 5 & 2 \\ -7 & 2 & 4 \end{pmatrix}$  ist symmetrisch.
- $\begin{pmatrix} 0 & 8 & -7 \\ -8 & 0 & 2 \\ 7 & -2 & 0 \end{pmatrix}$  ist schiefsymmetrisch.

### 3-Minutes Python 2

Es gibt eine Reihe von Funktionen in **numpy** um bestimmte Matrizen zu erzeugen. **zeros** erzeugt eine Matrix mit lauter Nullen, **ones** eine Matrix mit lauter Einsen, **empty** eine Matrix, bei der die Werte nicht initialisiert sind (sie haben einfach den Wert, der zur Zeit der Erzeugung zufällig im Speicher steht).

```
import numpy as np
2 >>> A=np.zeros( (3,4) )
>>> print(A)
4 [[ 0.  0.  0.  0.]
   [ 0.  0.  0.  0.]
   [ 0.  0.  0.  0.]]
6 B=np.ones( (2,3) )
8 >>> B=np.ones( (2,3) )
>>> print(B)
10 [[ 1.  1.  1.]
    [ 1.  1.  1.]]
12 >>> C=np.empty( (3,3) )
>>> print(C)
14 [[ -2.00000000e+000  -2.00000000e+000  -2.00000000e+000]
    [ -2.00000000e+000   2.81617418e-322   0.00000000e+000]
16 [  0.00000000e+000   0.00000000e+000   0.00000000e+000]]
```

Es gibt auch eine Funktion um die Einheitsmatrix zu erzeugen. Sie heißt **identity** und erwartet nur einen Parameter, da die Einheitsmatrix quadratisch ist.

```
>>> I=np.identity(3)
2 >>> print(I)
4 [[ 1.  0.  0.]
   [ 0.  1.  0.]
   [ 0.  0.  1.]]
```

Die Funktion **diag** erzeugt eine Diagonalmatrix mit den Elementen eines Zeilenvektors, der als Argument übergeben wird. Möchte man nicht die Hauptdiagonale setzen, dann kann man die Diagonale über einen optionalen zweiten Parameter angeben, der die Diagonale relativ zur Hauptdiagonale angibt.

```

>>> np.diag([1,3,4])
2 array([[1, 0, 0],
        [0, 3, 0],
        [0, 0, 4]])
>>> np.diag([1,3,4],-1)
6 array([[0, 0, 0, 0],
        [1, 0, 0, 0],
        [0, 3, 0, 0],
        [0, 0, 4, 0]])
10 >>> np.diag([1,3,4],+2)
    array([[0, 0, 1, 0, 0],
12         [0, 0, 0, 3, 0],
        [0, 0, 0, 0, 4],
14         [0, 0, 0, 0, 0],
        [0, 0, 0, 0, 0]])

```

### Definition 2.20 (Matrix-Produkt)

Sei  $\mathbf{A} \in \mathbb{R}^{m \times n}$ ,  $\mathbf{B} \in \mathbb{R}^{n \times k}$  dann ist  $\mathbf{A} \cdot \mathbf{B} \in \mathbb{R}^{m \times k}$  mit

$$(\mathbf{A} \cdot \mathbf{B})_{ij} = \sum_{1 \leq l \leq n} a_{il} \cdot b_{lj} \quad \forall 1 \leq i \leq m, 1 \leq j \leq k.$$

### Beispiel 2.21

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 2 \cdot 4 & 1 \cdot 2 + 2 \cdot 5 & 1 \cdot 3 + 2 \cdot 6 \\ 3 \cdot 1 + 4 \cdot 4 & 3 \cdot 2 + 4 \cdot 5 & 3 \cdot 3 + 4 \cdot 6 \\ 5 \cdot 1 + 6 \cdot 4 & 5 \cdot 2 + 6 \cdot 5 & 5 \cdot 3 + 6 \cdot 6 \end{pmatrix} = \begin{pmatrix} 9 & 12 & 15 \\ 19 & 26 & 33 \\ 29 & 40 & 51 \end{pmatrix}$$

### Anmerkung 2.22

- Auch für quadratische Matrizen  $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$  ist in der Regel  $\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A}$

$\Rightarrow$  Das Kommutativgesetz gilt für die Matrixmultiplikation nicht.

- Für zwei rechteckige Matrizen  $\mathbf{A} \in \mathbb{R}^{m \times n}$  und  $\mathbf{B} \in \mathbb{R}^{n \times m}$  gilt sogar:

$$\mathbf{A} \cdot \mathbf{B} \in \mathbb{R}^{m \times m}$$

$$\mathbf{B} \cdot \mathbf{A} \in \mathbb{R}^{n \times n}$$

### Beispiel 2.23

1.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 1 \cdot 5 + 2 \cdot 7 & 1 \cdot 6 + 2 \cdot 8 \\ 3 \cdot 5 + 4 \cdot 7 & 3 \cdot 6 + 4 \cdot 8 \end{pmatrix} = \begin{pmatrix} 5 + 14 & 6 + 16 \\ 15 + 28 & 18 + 32 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}$$

2.

$$\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 5 \cdot 1 + 6 \cdot 3 & 5 \cdot 2 + 6 \cdot 4 \\ 7 \cdot 1 + 8 \cdot 3 & 7 \cdot 2 + 8 \cdot 4 \end{pmatrix} = \begin{pmatrix} 5 + 18 & 10 + 24 \\ 7 + 24 & 14 + 32 \end{pmatrix} = \begin{pmatrix} 23 & 34 \\ 31 & 46 \end{pmatrix}$$

Ohne Beweis geben wir folgende Regeln für die Matrixmultiplikation:

**Satz 2.24**

Sei  $\mathbf{A} \in \mathbb{R}^{m \times n}$ .

Für  $\mathbf{B} \in \mathbb{R}^{n \times k}$ ,  $\mathbf{C} \in \mathbb{R}^{k \times l}$  und  $\lambda \in \mathbb{R}$  gilt:

Assoziativgesetz:

$$\mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{C}) = (\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{C} \quad (2.11)$$

$$\lambda \cdot (\mathbf{A} \cdot \mathbf{B}) = (\lambda \cdot \mathbf{A}) \cdot \mathbf{B} = \mathbf{A} \cdot (\lambda \cdot \mathbf{B}) \quad (2.12)$$

Multiplikation mit der Einheitsmatrix:

$$\mathbf{I} \cdot \mathbf{A} = \mathbf{A} \quad \text{für } \mathbf{I} \in \mathbb{R}^{m \times m} \quad (2.13)$$

$$\mathbf{A} \cdot \mathbf{I} = \mathbf{A} \quad \text{für } \mathbf{I} \in \mathbb{R}^{n \times n} \quad (2.14)$$

Distributivgesetze:

Seien  $\mathbf{B}, \mathbf{C} \in \mathbb{R}^{n \times k}$

$$\mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) = \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C} \quad (2.15)$$

Sei  $\mathbf{B} \in \mathbb{R}^{m \times n}$ ,  $\mathbf{C} \in \mathbb{R}^{n \times k}$

$$(\mathbf{A} + \mathbf{B}) \cdot \mathbf{C} = \mathbf{A} \cdot \mathbf{C} + \mathbf{B} \cdot \mathbf{C} \quad (2.16)$$

### 3-Minutes Python 3

Das Produkt zweier Matrizen oder einer Matrix mit einem Vektor berechnet sich durch die Methode `dot` der Matrix auf der linken Seite des Produktes. Python ist dabei etwas nachlässig mit der exakten Form des Vektors. Das Matrix-Vektor-Produkt lässt sich auch berechnen, wenn es sich um ein Produkt mit einem Zeilenvektor handelt. Man erhält das Ergebnis, das man für einen Spaltenvektor mit den gleichen Einträgen bekommen würde, nur als Zeilenvektor ausgegeben. Mathematisch ist das zwar falsch, aber der Bequemlichkeit geschuldet, da sich Zeilenvektoren einfacher schreiben lassen.

```
import numpy as np
2 >>> A=np.array([[2, 4], [1, 3]])
  >>> b=np.array([[1],[2]])
4 >>> A.dot(b)
  array([[10],
6         [ 7]])
  >>> A.dot(A)
8 array([[ 8, 20],
         [ 5, 13]])
10 >>> c=np.array([1,2])
  >>> A.dot(c)
12 array([10,  7])
```

Im Gegensatz dazu wird bei Verwendung des einfachen Multiplikationszeichens einfach komponentenweise multipliziert. Bei Multiplikation einer Matrix mit einem Vektor werden die Einträge einer Zeile jeweils mit dem entsprechenden Wert im Vektor multipliziert.

```
>>> A*A
```

```

2 array([[ 4, 16],
        [ 1,  9]])
4 >>> A*b
array([[2,  4],
6      [2,  6]])

```

Summe und Differenz von Matrizen und Vektoren, sowie das Produkt mit einem Skalar, werden ebenfalls komponentenweise berechnet.

```

>>> A=np.array([[2,  4], [1,  2]])
2 >>> A+A
array([[4,  8],
4      [2,  4]])
>>> A-A
6 array([[0,  0],
        [0,  0]])
8 >>> A-2*A
array([[ -2,  -4],
10      [-1,  -2]])

```

### 2.3. Lineare Gleichungssysteme

Ein lineares Gleichungssystem (LGS) mit  $m$  Gleichungen für  $n$  Unbekannte  $x_1, \dots, x_n$  hat die Form

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

in Matrixschreibweise:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$$

oder

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b} \quad \text{mit} \quad \mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{x} \in \mathbb{R}^n \text{ und } \mathbf{b} \in \mathbb{R}^m.$$

#### Definition 2.25

Ein lineares Gleichungssystem heißt homogen, falls die rechte Seite gleich  $\mathbf{0}$  ist, also  $\mathbf{b} = \mathbf{0}$ . Sonst heißt es inhomogen. Zur Lösung eines linearen Gleichungssystems kombiniert man  $\mathbf{A}$



und  $\mathbf{b}$  häufig zu einer erweiterten Koeffizientenmatrix

$$(\mathbf{A}|\mathbf{b}) := \left( \begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array} \right) \in \mathbb{R}^{m \times (n+1)}.$$

Folgende Operationen (elementare Zeilenumformungen) an der erweiterten Koeffizientenmatrix ändern die Lösung des LGS nicht:

1. das Vertauschen zweier Zeilen
2. die Multiplikation einer Zeile mit einer Zahl  $\lambda \neq 0$
3. die Addition (oder Subtraktion) des  $\lambda$ -fachen einer Zeile zu einer anderen.

### Definition 2.26

Entsteht eine Matrix  $\mathbf{B} \in \mathbb{R}^{m \times n}$  durch endlich viele elementare Zeilenumformungen aus  $\mathbf{A} \in \mathbb{R}^{m \times n}$ , dann sind  $\mathbf{A}$  und  $\mathbf{B}$  zueinander äquivalent.

## 2.4. Gauß Algorithmus

Zur Lösung eines linearen Gleichungssystems wird die erweiterte Koeffizientenmatrix durch elementare Zeilenumformungen so umgeformt, dass das äquivalente lineare Gleichungssystem einfacher lösbar ist.

### Beispiel 2.27

1.

$$\mathbf{A} = \begin{pmatrix} 4 & -2 & -2 \\ -8 & 6 & -2 \\ 0 & 4 & -11 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 4 \\ 0 \\ 17 \end{pmatrix}$$

$$\left( \begin{array}{ccc|c} 4 & -2 & -2 & 4 \\ -8 & 6 & -2 & 0 \\ 0 & 4 & -11 & 17 \end{array} \right) + 2 \cdot 1. \text{ Zeile}$$

$$\left( \begin{array}{ccc|c} 4 & -2 & -2 & 4 \\ 0 & 2 & -6 & 8 \\ 0 & 4 & -11 & 17 \end{array} \right) - 2 \cdot 2. \text{ Zeile}$$

$$\left( \begin{array}{ccc|c} 4 & -2 & -2 & 4 \\ 0 & 2 & -6 & 8 \\ 0 & 0 & 1 & 1 \end{array} \right)$$

$$\begin{array}{rrcr} 4x_1 & -2x_2 & -2x_3 & = 4 \\ & 2x_2 & -6x_3 & = 8 \\ & & x_3 & = 1 \end{array}$$

$$\text{Auflösen nach } x_1, x_2, x_3 \quad \Rightarrow \quad \mathbf{x} = \begin{pmatrix} 5 \\ 7 \\ 1 \end{pmatrix}$$

2.

$$\mathbf{A} = \begin{pmatrix} 4 & -2 & -2 \\ -8 & 6 & -2 \\ 0 & 4 & -12 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 4 \\ 0 \\ 17 \end{pmatrix}$$
$$\Rightarrow \left( \begin{array}{ccc|c} 4 & -2 & -2 & 4 \\ 0 & 2 & -6 & 8 \\ 0 & 0 & 0 & 1 \end{array} \right)$$

$\Rightarrow$  keine Lösung, da

$$0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 = 1 \quad (2.17)$$

immer falsch ist.

3.

$$\mathbf{b} = \begin{pmatrix} 4 \\ 0 \\ 16 \end{pmatrix}$$
$$\left( \begin{array}{ccc|c} 4 & -2 & -2 & 4 \\ 0 & 2 & -6 & 8 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

$\Rightarrow$  Unendlich viele Lösungen, da letzte Zeile wahr unabhängig von  $\mathbf{x}$ .

Wähle  $x_3 = \lambda$ :

$$\begin{array}{rclcl} 4x_1 & -2x_2 & -2x_3 & = & 4 \\ & 2x_2 & -6x_3 & = & 8 \\ & & x_3 & = & \lambda \end{array}$$
$$\Rightarrow \begin{array}{rclcl} 4x_1 & -2x_2 & & = & 4 + 2\lambda \\ & 2x_2 & & = & 8 + 6\lambda \end{array}$$
$$\Rightarrow \mathbf{x} = \begin{pmatrix} (4 + 2\lambda + 8 + 6\lambda)/4 \\ (8 + 6\lambda)/2 \\ \lambda \end{pmatrix} = \begin{pmatrix} 3 + 2\lambda \\ 4 + 3\lambda \\ \lambda \end{pmatrix}$$
$$\mathbf{x} = \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix} + \lambda \cdot \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$$

Geradengleichung im dreidimensionalen Raum

Wir haben gesehen, dass es drei Fälle gibt,

- Es gibt genau eine Lösung
- Es gibt keine Lösung
- Es gibt unendlich viele Lösungen

Frage:

Wovon hängt es ab, welcher Fall eintritt?

Beobachtung:

Es scheint von

- **A** abzuhängen, ob es genau eine Lösung gibt.
- **b** abzuhängen, ob es keine oder unendlich viele Lösungen gibt.

### Definition 2.28

Sei  $\nu_i = \min \{j : a_{ij} \neq 0\}$  die Spalte mit dem ersten Element in Zeile  $i$ , das ungleich Null ist. Eine Matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$  hat Treppenform, falls gilt:

$$0 \leq \nu_0 < \nu_1 < \nu_2 < \dots < \nu_r$$

mit  $r \leq n$ . Am Anfang jeder Zeile gibt es also mindestens ein Nullelement mehr, als in der Zeile zuvor.

### Algorithmus 2.29 (Vorwärtselimination)

Bringe erweiterte Koeffizientenmatrix durch elementare Zeilenumformungen in Treppenform.

- 0) Setze  $k = 1$
- 1) Falls  $a_{kk} = 0$ 
  - a) Falls  $\nexists a_{ik} \neq 0$  mit  $i > k$ , dann ist auch die  $k$ -te Spalte schon in Treppenform, gehe zu 3)
  - b) Falls  $\exists a_{ik} \neq 0$  mit  $i > k$ , vertausche die Zeilen  $k$  und  $i$
- 2)
$$a_{ij} = a_{ij} - \frac{a_{ik}}{a_{kk}} \cdot a_{kj} \quad \forall k < i, j \leq n$$
$$a_{ik} = 0 \quad \forall k < i \leq n$$
- 3) Setze  $k = k + 1$ , falls  $k < n$  gehe zu 1)

### Definition 2.30 (Nullzeile)

Eine Matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$  hat eine Nullzeile falls  $a_{ij} = 0 \quad \forall 1 \leq j \leq n$ .

### Definition 2.31 (Rang einer Matrix)

Eine Matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$  hat den Rang  $k$ , falls die zu  $\mathbf{A}$  äquivalente Matrix in Treppenform genau  $k$  Zeilen hat, die keine Nullzeilen sind.

### Anmerkung 2.32

- Elementare Zeilenumformungen ändern den Rang einer Matrix nicht

### 3-Minutes Python 4

Mit der Funktion `linalg.matrix_rank` lässt sich der Rang einer Matrix berechnen:

```
>>> import numpy as np
2 >>> A=np.array([[2, 4], [1, 2]])
>>> np.linalg.matrix_rank(A)
4 1
>>> np.linalg.matrix_rank(np.array([[2, 4], [1, 3]]))
6 2
```

**Satz 2.33**

1. Ein lineares Gleichungssystem  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$  mit  $\mathbf{A} \in \mathbb{R}^{m \times n}$ ,  $\mathbf{x} \in \mathbb{R}^n$  und  $\mathbf{b} \in \mathbb{R}^m$  ist lösbar, falls der Rang der Matrix  $\mathbf{A}$  und der Rang der erweiterten Koeffizientenmatrix  $(\mathbf{A}|\mathbf{b})$  gleich sind, d.h.

$$\text{Rang}(\mathbf{A}) = \text{Rang}(\mathbf{A}|\mathbf{b}) = p.$$

2. Ist ein LGS lösbar, dann gibt es genau eine Lösung, falls  $p=n$  und unendlich viele Lösungen, falls  $p < n$ , wobei  $(n-p)$  Parameter frei gewählt werden können.

**Beweis:**

1. Ist der Rang der erweiterten Koeffizientenmatrix größer, als der Rang von  $\mathbf{A}$ , dann gibt es, wenn man beide Matrizen in Treppenform bringt, mindestens eine Zeile

$$0 \ 0 \ \dots \ 0 \ 0 \mid b_i$$

mit  $b_i \neq 0$ . Das entspricht der unlösbaren Gleichung

$$0 \cdot x_1 + 0 \cdot x_2 + \dots + 0 \cdot x_{n-1} + 0 \cdot x_n = b_i \neq 0.$$

2. Ist das LGS lösbar und der Rang  $(\mathbf{A}) = n$ , dann sind die ersten  $n$  Zeilen der zu  $\mathbf{A}$  äquivalenten Matrix in Treppenstufenform  $\mathbf{A}^*$  eine obere Dreiecksmatrix, für die das LGS eine eindeutige Lösung besitzt.

Für  $p < n$  gibt es nur  $p$  Gleichungen für  $n$ -Unbekannte, so dass  $(n-p)$  Parameter frei wählbar sind.

□

**Satz 2.34**

1. Das homogene LGS  $\mathbf{A} \cdot \mathbf{x} = \mathbf{0}$  mit  $\mathbf{A} \in \mathbb{R}^{m \times n}$ ,  $\mathbf{x} \in \mathbb{R}^n$  hat nur die "triviale" Lösung  $x_i = 0 \quad \forall 1 \leq i \leq n$ , falls  $\text{Rang}(\mathbf{A}) = n$ .
2. Ist  $\text{Rang}(\mathbf{A}) = p < n$ , dann besitzt  $\mathbf{A} \cdot \mathbf{x} = \mathbf{0}$  von  $\mathbf{0}$  verschiedene Lösungen mit  $n-p$  freien Parametern.
3. Ist das inhomogene LGS  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$  für  $\mathbf{b} \in \mathbb{R}^m$  lösbar, lässt sich jede Lösung  $\mathbf{v}$  darstellen als

$$\mathbf{v} = \mathbf{v}_0 + \mathbf{u}$$

wobei  $\mathbf{v}_0$  eine beliebige spezielle Lösung von  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$  ist und  $\mathbf{u}$  eine Lösung des homogenen Systems.

**Beweis:**

1. Nach Satz 2.33 gibt es genau eine Lösung, falls  $\text{Rang}(\mathbf{A}) = n$ , diese muss dann  $\mathbf{x} = \mathbf{0}$  sein.
2. folgt aus Satz 2.33

3. Da  $\mathbf{A} \cdot \mathbf{u} = \mathbf{0}$  und  $\mathbf{A} \cdot \mathbf{v}_0 = \mathbf{b}$  gilt  $\mathbf{A} \cdot (\mathbf{v}_0 + \mathbf{u}) = \mathbf{A} \cdot \mathbf{v}_0 + \mathbf{A} \cdot \mathbf{u} = \mathbf{b}$ . Andererseits folgt aus  $\mathbf{A} \cdot \mathbf{v} = \mathbf{b}$  und  $\mathbf{A} \cdot \mathbf{v}_0 = \mathbf{b}$ , dass  $\mathbf{u} := \mathbf{v} - \mathbf{v}_0$  eine Lösung des homogenen Systems sein muss.

□

### Anmerkung 2.35

Um alle möglichen Lösungen zu finden genügt es also eine (spezielle) Lösung von  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$  zu bestimmen (z.B. indem man alle freien Parameter auf Null setzt), sowie die Lösungsmenge des homogenen Systems (die nur von  $\mathbf{A}$  abhängt und für alle rechten Seiten  $\mathbf{b}$  gleich ist).

### Beispiel 2.36

Wir haben gezeigt (Beispiel 2.27, 3.), dass die Lösung von

$$\begin{pmatrix} 4 & -2 & -2 \\ -8 & 6 & -2 \\ 0 & 4 & -12 \end{pmatrix} \cdot \mathbf{x} = \begin{pmatrix} 4 \\ 0 \\ 16 \end{pmatrix}$$

die Form hat:

$$\mathbf{x} = \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix} + \lambda \cdot \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$$

Wir erhalten nach:

$$\begin{pmatrix} 4 & -2 & -2 \\ -8 & 6 & -2 \\ 0 & 4 & -12 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix} = \begin{pmatrix} 12 - 8 \\ -24 + 24 \\ 16 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ 16 \end{pmatrix}$$

$$\begin{pmatrix} 4 & -2 & -2 \\ -8 & 6 & -2 \\ 0 & 4 & -12 \end{pmatrix} \cdot \lambda \cdot \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} = \lambda \cdot \begin{pmatrix} 8 - 6 - 2 \\ -16 + 18 - 2 \\ 12 - 12 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix} \text{ ist eine spezielle Lösung des inhomogenen LGS und } \lambda \cdot \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} \text{ die Lösungsmenge des}$$

homogenen LGS.

### Algorithmus 2.37 (Rückwärtseinsetzen)

Ist das LGS  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$  eindeutig lösbar und ist die erweiterte Koeffizientenmatrix  $(\mathbf{A}|\mathbf{b})$  durch Vorwärtselimination in Treppenform  $(\mathbf{A}^*|\mathbf{b}^*)$  gebracht, dann erhält man die Lösung  $\mathbf{x}$  wie folgt:

1.  $x_n = b_n^*/a_{nn}^*$
2. Für  $i = (n-1), \dots, 1$ :

$$x_i = \frac{1}{a_{ii}^*} \left( b_i^* - \sum_{j>i} a_{ij}^* x_j \right)$$

### 3-Minutes Python 5

Die Lösung eines linearen Gleichungssystems lässt sich ganz einfach mit Hilfe der Methode `linalg.solve` von `numpy` berechnen. Dabei ist das erste Argument die Matrix und das zweite Argument die rechte Seite. `numpy` ist relativ lax, was die Form der rechten Seite angeht. Steht dort ein Zeilenvektor, dann berechnet es auch die richtige Lösung (was mathematisch nicht richtig ist) und liefert auch einen Zeilenvektor zurück. Ist die rechte Seite ein Spaltenvektor, dann liefert die Methode auch einen Spaltenvektor zurück.

```
>>> import numpy as np
2 >>> A=np.array([[2, 4], [1, 3]])
>>> b=np.array([1,2])
4 >>> np.linalg.solve(A,b)
array([-2.5,  1.5])
6 >>> b=np.array([[1],[2]])
>>> np.linalg.solve(A,b)
8 array([[ -2.5],
        [  1.5]])
```

Ist das lineare Gleichungssystem nicht eindeutig lösbar, dann bricht der Löser mit der Fehlermeldung ab, dass die Matrix singulär wäre.

```
1 >>> A=np.array([[2, 4], [1, 2]])
>>> np.linalg.solve(A,b)
3 Traceback (most recent call last):
  File "<stdin>", line 1, in <module>
5   File
    "/usr/local/lib/python3.6/site-packages/numpy/linalg/linalg.py",
    line 384, in solve
    r = gufunc(a, b, signature=signature, extobj=extobj)
7   File
    "/usr/local/lib/python3.6/site-packages/numpy/linalg/linalg.py",
    line 90, in _raise_linalgerror_singular
    raise LinAlgError("Singular matrix")
9 numpy.linalg.linalg.LinAlgError: Singular matrix
```

### 3. Vektorraum

Wir suchen Antworten auf die Fragen:

- Wieso gibt es für die gleiche Matrix keine oder unendlich viele Lösungen?
- Wie sieht die Lösungsmenge aus?

Zu deren Lösung lohnt sich eine alternative Betrachtungsweise eines linearen Gleichungssystems:

Welche Vektoren aus  $\mathbb{R}^m$  lassen sich als Summe von Vielfachen der Spaltenvektoren darstellen? Zum Beispiel als

$$\begin{pmatrix} 4 \\ -8 \\ 0 \end{pmatrix} \cdot x_1 + \begin{pmatrix} -2 \\ 6 \\ 4 \end{pmatrix} \cdot x_2 + \begin{pmatrix} -2 \\ -2 \\ -12 \end{pmatrix} \cdot x_3$$

Wir führen dazu zunächst ein neues Konzept ein:

#### Definition 3.1 (Vektorraum)

Eine nicht leere Menge  $\mathbf{V}$ , in der man zu je zwei Elementen  $\mathbf{a}, \mathbf{b} \in \mathbf{V}$  eine Summe  $\mathbf{a} + \mathbf{b} \in \mathbf{V}$  und zu jedem Element  $\mathbf{a} \in \mathbf{V}$  und jedem Skalar  $\lambda \in \mathbb{R}$  das Produkt  $\lambda \cdot \mathbf{a} \in \mathbf{V}$  bilden kann, heißt ein Vektorraum über  $\mathbb{R}$ , wenn folgende acht Rechengesetze (Vektorraum-Axiome) erfüllt sind:

Für  $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbf{V}$  und  $\lambda, \mu \in \mathbb{R}$  gilt:

$$\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a} \quad (3.1)$$

$$(\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}) \quad (3.2)$$

$$\exists \mathbf{0} \in \mathbf{V} : \forall \mathbf{a} \in \mathbf{V} : \mathbf{a} + \mathbf{0} = \mathbf{a} \quad (3.3)$$

$$\forall \mathbf{a} \in \mathbf{V} : \exists -\mathbf{a} \in \mathbf{V} : \mathbf{a} + (-\mathbf{a}) = \mathbf{0} \quad (3.4)$$

$$1 \cdot \mathbf{a} = \mathbf{a} \quad (3.5)$$

$$\lambda \cdot (\mu \cdot \mathbf{a}) = (\lambda \cdot \mu) \cdot \mathbf{a} \quad (3.6)$$

$$\lambda \cdot (\mathbf{a} + \mathbf{b}) = \lambda \cdot \mathbf{a} + \lambda \cdot \mathbf{b} \quad (3.7)$$

$$(\lambda + \mu) \cdot \mathbf{a} = \lambda \cdot \mathbf{a} + \mu \cdot \mathbf{a} \quad (3.8)$$

Die Elemente eines Vektorraums nennt man Vektoren. Statt  $\mathbf{a} + (-\mathbf{b})$  schreibt man  $\mathbf{a} - \mathbf{b}$ .

#### Beispiel 3.2

1. Die Menge aller Matrizen  $\mathbf{A} \in \mathbb{R}^{m \times n}$  bildet einen Vektorraum.

Dies gilt insbesondere für die Menge aller Zeilenvektoren

$$\mathbb{R}^{1 \times n} = \{(a_1, \dots, a_n) : a_i \in \mathbb{R}\}$$

und die Menge aller Spaltenvektoren

$$\mathbb{R}^n = \left\{ \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} : a_i \in \mathbb{R} \right\}.$$

2. Die Menge aller reellen Polynome vom Grad  $\leq n$ :

$$\mathbf{P}_n = \{\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_n x^n : \alpha_i \in \mathbb{R} \quad \forall 0 \leq i \leq n\}$$

Addition von Polynomen und Multiplikation mit einem Skalar sind so definiert wie bei Polynomen üblich.

$$\text{Nullvektor } p(x) := 0 \quad \forall x \in \mathbb{R}$$

$$\text{Zu } p(x) = \sum_{i=0}^n \alpha_i x^i \text{ negativer Vektor: } -p(x) := \sum_{i=0}^n (-\alpha_i) x^i.$$

3. Die Menge aller auf dem Intervall  $I \subseteq \mathbb{R}$  stetigen reellen Funktionen  $\mathcal{C}^0(I)$ <sup>1</sup> ist definiert als

$$\mathcal{C}^0(I) := \{f : I \rightarrow \mathbb{R}, f \text{ stetig}\}$$

$$(f + g)(x) := f(x) + g(x)$$

$$(\lambda f)(x) := \lambda \cdot f(x)$$

$$\text{Nullvektor } \mathbf{0}(x) := 0 \quad \forall x \in I$$

$$\text{Zu } f \in \mathcal{C}^0(I) \text{ negative Funktion } (-f)(x) := -f(x) \quad \forall x \in I$$

### Definition 3.3 (Unterraum)

Eine nicht leere Teilmenge  $\mathbf{U} \subseteq \mathbf{V}$  heißt Unterraum (oder linearer Teilraum) von  $\mathbf{V}$ , wenn für alle  $\mathbf{u}, \mathbf{v} \in \mathbf{U}$  auch  $(\mathbf{u} + \mathbf{v}) \in \mathbf{U}$  und für alle  $\mathbf{u} \in \mathbf{U}$  und  $\lambda \in \mathbb{R}$  auch  $\lambda \cdot \mathbf{u} \in \mathbf{U}$  gilt.

Damit ist jeder Unterraum auch wieder ein Vektorraum.

### Beispiel 3.4

1. Jeder Vektorraum  $\mathbf{V}$  hat die “trivialen” Unterräume  $\{\mathbf{0}\}$  und  $\mathbf{V}$ .
2. Für jedes  $\mathbf{v} \in \mathbf{V}$  ist

$$\{\lambda \cdot \mathbf{v} : \lambda \in \mathbb{R}\}$$

ein Unterraum von  $\mathbf{V}$ . Falls  $\mathbf{V} = \mathbb{R}^2$  oder  $\mathbf{V} = \mathbb{R}^3$ , ist dies die Menge aller zu  $\mathbf{v}$  parallelen Vektoren (inklusive Nullvektor).

3. Die Menge aller Lösungen  $\mathbf{x} \in \mathbb{R}^n$  des homogenen LGS  $\mathbf{A} \cdot \mathbf{x} = \mathbf{0}$  mit  $\mathbf{A} \in \mathbb{R}^{m \times n}$

$$\text{Kern } \mathbf{A} := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A} \cdot \mathbf{x} = \mathbf{0}\},$$

ist ein Unterraum von  $\mathbb{R}^n$  und heißt der Kern oder auch der Nullraum von  $\mathbf{A}$ .

### Beweis:

Da die triviale Lösung  $\mathbf{0} \in \text{Kern } \mathbf{A}$  existiert, ist die Menge nicht leer. Seien  $\mathbf{u}, \mathbf{v} \in \text{Kern } \mathbf{A}$  und  $\lambda \in \mathbb{R}$ :

$$\mathbf{A} \cdot (\mathbf{u} + \mathbf{v}) = \mathbf{A} \cdot \mathbf{u} + \mathbf{A} \cdot \mathbf{v} = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

$$\mathbf{A} \cdot (\lambda \mathbf{u}) = \lambda \cdot \mathbf{A} \cdot \mathbf{u} = \lambda \cdot \mathbf{0} = \mathbf{0}$$

$$\Rightarrow \mathbf{u} + \mathbf{v} \in \text{Kern } \mathbf{A} \text{ und } \lambda \cdot \mathbf{u} \in \text{Kern } \mathbf{A}$$

□

---

<sup>1</sup>wird oft auch nur als  $\mathcal{C}(I)$  bezeichnet



**Definition 3.5 (Linearkombination)**

Jede aus endlich vielen Vektoren  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \in \mathbf{V}$  gebildete Summe der Form

$$\sum_{i=1}^n \alpha_i \cdot \mathbf{v}_i$$

mit den Koeffizienten  $\alpha_i \in \mathbb{R}$  heißt eine Linearkombination der Vektoren  $\mathbf{v}_i$ .

Die Menge aller möglichen Linearkombinationen der  $\mathbf{v}_i$  heißt lineare Hülle der  $\mathbf{v}_i$  oder auch  $\text{Span}(\mathbf{v}_i)$ :

$$\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_n) := \left\{ \sum_{i=1}^n \alpha_i \mathbf{v}_i \quad : \quad \alpha_i \in \mathbb{R}, \quad \forall 1 \leq i \leq n \right\}$$

**Anmerkung 3.6**

Für  $\mathbf{v}_i \in \mathbf{V}$  ist  $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_n)$  ein Unterraum von  $\mathbf{V}$ , da:

$$\begin{aligned} \left( \sum_{i=1}^n \alpha_i \mathbf{v}_i \right) + \left( \sum_{i=1}^n \beta_i \mathbf{v}_i \right) &= \sum_{i=1}^n (\alpha_i + \beta_i) \mathbf{v}_i \\ \lambda \left( \sum_{i=1}^n \alpha_i \mathbf{v}_i \right) &= \sum_{i=1}^n (\lambda \alpha_i) \mathbf{v}_i \end{aligned}$$

**Beispiel 3.7**

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 3 \\ 2 \\ 0 \\ 1 \end{pmatrix}$$

$$\begin{aligned} \text{Span}(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) &= \{ \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \alpha_3 \mathbf{v}_3 : \alpha_i \in \mathbb{R} \} \\ &= \left\{ \begin{pmatrix} \alpha_1 + \alpha_2 + 3\alpha_3 \\ \alpha_1 + 2\alpha_3 \\ -\alpha_2 \\ \alpha_3 \end{pmatrix} : \alpha_i \in \mathbb{R} \right\} \end{aligned}$$

**Definition 3.8 (Erzeugendensystem)**

Ein Unterraum  $\mathbf{U}$  von  $\mathbf{V}$  wird von den Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbf{V}$  erzeugt/aufgespannt bzw.  $\mathbf{v}_1, \dots, \mathbf{v}_n$  ist ein Erzeugendensystem von  $\mathbf{U}$ , falls

$$\mathbf{U} = \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_n).$$

**Beispiel 3.9**

Die Vektoren  $\mathbf{e}_1 = (1, 0, 0)^T$ ,  $\mathbf{e}_2 = (0, 1, 0)^T$ ,  $\mathbf{e}_3 = (0, 0, 1)^T$  erzeugen den Vektorraum  $\mathbb{R}^3$ . Dies ist nicht eindeutig. Auch  $\mathbf{e}_1 + \mathbf{e}_2$ ,  $\mathbf{e}_1 - \mathbf{e}_2$ ,  $\mathbf{e}_2 + \mathbf{e}_3$  und  $\mathbf{e}_1 - \mathbf{e}_3$  erzeugen den  $\mathbb{R}^3$ .

**Frage:**

Sind alle Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_n$  notwendig um  $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_n)$  zu erzeugen (oder kann man einige Vektoren weglassen)?

**Definition 3.10 (Lineare Unabhängigkeit von Vektoren)**

Die Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbf{V}$  sind linear unabhängig, falls die Gleichung

$$\sum_{i=1}^n \alpha_i \mathbf{v}_i = \mathbf{0}$$

nur die Lösung  $\alpha_i = 0, \forall 1 \leq i \leq n$  hat. Sonst sind sie linear abhängig.

**Anmerkung 3.11**

Sind die Vektoren  $\mathbf{v}_i$  linear abhängig, dann lässt sich mindestens einer der Vektoren als Linearkombination der anderen Vektoren darstellen.

**Beweis:**

Falls die  $\mathbf{v}_i$  linear abhängig sind, existiert mindestens ein  $\alpha_j \neq 0$ , mit dem gilt  $\sum_{i=1}^n \alpha_i \mathbf{v}_i = \mathbf{0}$ .

Wir können daher nach  $\mathbf{v}_j$  auflösen:  $\mathbf{v}_j = -\frac{1}{\alpha_j} \cdot \sum_{i \neq j} \alpha_i \mathbf{v}_i$  □

**Beispiel 3.12**

- Zwei linear abhängige Vektoren  $\mathbf{v}_1, \mathbf{v}_2 \in \mathbf{V}$  sind parallel.

**Beweis:**

Linear abhängig  $\Rightarrow \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 = \mathbf{0}$  mit  $|\alpha_1| + |\alpha_2| > 0$  (d.h. mindestens einer der beiden Werte ist ungleich 0)

$$\begin{aligned} \Rightarrow \mathbf{v}_1 &= \frac{\alpha_2}{\alpha_1} \cdot \mathbf{v}_2 && \text{falls } \alpha_1 \neq 0 \\ \text{oder } \mathbf{v}_2 &= \frac{\alpha_1}{\alpha_2} \cdot \mathbf{v}_1 && \text{falls } \alpha_2 \neq 0 \end{aligned}$$

□

- Ist ein (einzelner) Vektor linear abhängig  $\Rightarrow \alpha_1 \mathbf{v}_1 = \mathbf{0}$  mit  $\alpha_1 \neq 0$   
 $\Rightarrow \mathbf{v}_1 = \mathbf{0}$

Ist bei den Vektoren der Nullvektor enthalten, dann sind sie daher immer linear abhängig, denn falls z.B.  $\mathbf{v}_k$  der Nullvektor ist, gilt:

$$\sum_{i=1}^n \alpha_i \mathbf{v}_i = \sum_{i \neq k} \alpha_i \mathbf{v}_i + \alpha_k \underbrace{\mathbf{v}_k}_{=\mathbf{0}} = \sum_{i \neq k} \alpha_i \mathbf{v}_i = \mathbf{0}$$

Dabei kann  $\mathbf{v}_k$  beliebig gewählt werden, also insbesondere ungleich Null.

**Definition 3.13 (Basis eines Vektorraums)**

Ein System von Vektoren  $(\mathbf{v}_1, \dots, \mathbf{v}_n)$  mit  $\mathbf{v}_i \in \mathbf{V}$  heißt eine Basis des Vektorraums  $\mathbf{V}$ , falls folgende Aussagen gelten:

1. Die Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_n$  sind linear unabhängig
2. Die Vektoren  $\mathbf{v}_i$  erzeugen  $\mathbf{V}$ , d.h.

$$\mathbf{V} = \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_n)$$

**Beispiel 3.14 (Kartesische Basis)**

Sei  $\mathbf{e}_i \in \mathbb{R}^n$  der  $i$ -te Einheitsvektor mit

$$(\mathbf{e}_i)_j = \begin{cases} 1 & j = i \\ 0 & \text{sonst} \end{cases}.$$

(also  $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0)^T$ ).

Dann ist  $\{\mathbf{e}_i \in \mathbb{R}^n : \forall 1 \leq i \leq n\}$  eine Basis von  $\mathbb{R}^n$ , da

$$\mathbf{x} = x_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \dots + x_n \cdot \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix}$$

und daher für  $\mathbf{x} = \mathbf{0}$  gelten muss, dass  $x_i = 0$  für  $1 \leq i \leq n$ .

**Satz 3.15**

Ist  $(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n)$  eine Basis von  $\mathbf{V}$ , dann gibt es zu jedem Vektor  $\mathbf{u} \in \mathbf{V}$  genau  $n$  reelle Zahlen (ein  $n$ -Tupel)  $\alpha_1, \alpha_2, \dots, \alpha_n$  mit

$$\mathbf{u} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots + \alpha_n \mathbf{v}_n$$

Ferner sind je  $m$  Vektoren aus  $\mathbf{V}$  linear abhängig, falls  $m > n$ .

**Beweis:**

Da eine Basis  $\mathbf{V}$  erzeugt, kann man jedes  $\mathbf{u} \in \mathbf{V}$  schreiben als

$$\mathbf{u} = \sum_{i=1}^n \alpha_i \mathbf{v}_i$$

- Eindeutigkeit:

Angenommen es gäbe von den  $\alpha_i$  verschiedene Zahlen  $\beta_i$ , so dass

$$\begin{aligned} \mathbf{u} &= \sum_{i=1}^n \alpha_i \mathbf{v}_i = \sum_{i=1}^n \beta_i \mathbf{v}_i \\ \Leftrightarrow \sum_{i=1}^n \alpha_i \mathbf{v}_i - \sum_{i=1}^n \beta_i \mathbf{v}_i &= \mathbf{0} \\ \Leftrightarrow \sum_{i=1}^n (\alpha_i - \beta_i) \mathbf{v}_i &= \mathbf{0} \\ \text{Basis, lin. unabh.} \quad \Leftrightarrow \quad \alpha_i - \beta_i &= 0 \quad \forall 1 \leq i \leq n \quad \nexists (\text{Widerspruch}) \end{aligned}$$

- Lineare Abhängigkeit:

Gegeben seine  $m$  Vektoren  $\mathbf{u}_1, \dots, \mathbf{u}_m \in \mathbf{V}$ , mit  $m > n$

Zu zeigen:

$$\sum_{j=1}^m \beta_j \mathbf{u}_j = 0 \quad \text{mit } \exists \beta_j \neq 0 \quad \text{mit } 1 \leq j \leq m$$

$$\begin{array}{l} \text{Basisdarst. von } \mathbf{u}_j \\ \Leftrightarrow \end{array} \quad \sum_{j=1}^m \left( \beta_j \cdot \sum_{i=1}^n \alpha_{ij} \mathbf{v}_i \right) = \sum_{i=1}^n \left[ \left( \sum_{j=1}^m \beta_j \alpha_{ij} \right) \cdot \mathbf{v}_i \right]$$

$$\begin{array}{l} \text{Basis, lin. unabh.} \\ \Rightarrow \end{array} \quad \text{zu erfüllen: } \sum_{j=1}^m \beta_j \alpha_{ij} = 0 \quad \forall 1 \leq i \leq n$$

Dieses lineare Gleichungssystem ist immer nicht-trivial lösbar (siehe Satz 2.34, 2.), da homogenes lineares Gleichungssystem  $\mathbf{A} \cdot \mathbf{b} = \mathbf{0}$  mit  $\mathbf{A} \in \mathbb{R}^{n \times m}$  und  $\mathbf{b} \in \mathbb{R}^m$ , also  $n < m$  Gleichungen für  $m$  Unbekannte.

□

### Beispiel 3.16

Eine Basis der Menge

$$\{f \in \mathcal{C}^2(\mathbb{R}) : f''(x) + f(x) = 0 \quad \forall x \in \mathbb{R}\}$$

ist  $f_1(x) = \sin(x)$  und  $f_2(x) = \cos(x)$ , d.h. jede Lösung der Gleichung  $f''(x) + f(x) = 0$  lässt sich darstellen als  $\alpha_1 \sin(x) + \alpha_2 \cos(x)$ . Dies ließe sich wieder schreiben als  $\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$  (wenn die Basis klar ist).

### Anmerkung 3.17

Die  $n$  reellen Koeffizienten für die Erzeugung eines Vektors durch die Basisvektoren sind das, was man gemeinhin unter einem Vektor versteht. Sie sind aber nur seine Darstellung in einer speziellen Basis. Bisher war dies meist die kartesische Basis. Dies muss aber nicht notwendigerweise die sinnvollste Wahl sein.

### Beispiel 3.18 (Basisdarstellung)

Der Vektor  $\mathbf{a} \in \mathbb{R}^2$  habe in der Basis  $\mathbf{e}_1, \mathbf{e}_2$  die Darstellung

$$\mathbf{a} = \begin{pmatrix} 2 \\ 3 \end{pmatrix} = 2 \cdot \mathbf{e}_1 + 3 \mathbf{e}_2$$

In der Basis  $\mathbf{e}_1 + \mathbf{e}_2, \mathbf{e}_1 - \mathbf{e}_2$  hat er die Darstellung:

$$\mathbf{a} = \begin{pmatrix} \frac{5}{2} \\ -\frac{1}{2} \end{pmatrix} = \frac{5}{2} \cdot (\mathbf{e}_1 + \mathbf{e}_2) - \frac{1}{2}(\mathbf{e}_1 - \mathbf{e}_2) = \left(\frac{5}{2} - \frac{1}{2}\right) \mathbf{e}_1 + \left(\frac{5}{2} + \frac{1}{2}\right) \mathbf{e}_2 = 2\mathbf{e}_1 + 3\mathbf{e}_2.$$

### Definition 3.19

Ein Vektorraum heißt endlich dimensional oder endlich erzeugt, wenn es endlich viele Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_n \in \mathbf{V}$  gibt, so dass  $\mathbf{V} = \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_n)$ .

**Satz 3.20 (Dimension eines Vektorraums)**

Jeder endlich erzeugte Vektorraum  $\mathbf{V} \neq \{\mathbf{0}\}$  besitzt eine (endliche) Basis  $(\mathbf{v}_1, \dots, \mathbf{v}_n)$ . Ist  $(\mathbf{u}_1, \dots, \mathbf{u}_m)$  ebenfalls eine Basis von  $\mathbf{V}$ , dann gilt  $m = n$ . Die gemeinsame Zahl  $n$  aller Basen von  $\mathbf{V}$  heißt die Dimension von  $\mathbf{V}$  abgekürzt  $\text{Dim } \mathbf{V}$ . Für  $\{\mathbf{0}\}$  setzt man  $\text{Dim}\{\mathbf{0}\} = 0$ .

**Beweis:**

Der Vektorraum ist endlich erzeugt, also muss es eine Basis  $(\mathbf{v}_1, \dots, \mathbf{v}_n)$  von  $\mathbf{V}$  geben. Nach Satz 3.15 sind  $m > n$  Vektoren aus  $\mathbf{V}$  immer linear abhängig. Ist  $(\mathbf{u}_1, \dots, \mathbf{u}_m)$  auch eine Basis von  $\mathbf{V}$ , dann muss deshalb  $m \leq n$  gelten (da die Vektoren einer Basis linear unabhängig sein müssen). Umgekehrt folgt aus  $(\mathbf{u}_1, \dots, \mathbf{u}_m)$  ist Basis von  $\mathbf{V}$ , dass  $n \leq m$ . Damit bleibt nur  $m = n$  übrig.  $\square$

**Satz 3.21 (Dimension der Unterräume)**

Jeder Unterraum  $\mathbf{U}$  eines endlich dimensionalen Vektorraums  $\mathbf{V}$  ist endlich dimensional. Wenn  $\mathbf{U} \neq \mathbf{V}$  (bzw.  $\mathbf{U} \subset \mathbf{V}$ ), gilt:

$$\text{Dim } \mathbf{U} < \text{Dim } \mathbf{V}. \quad (3.9)$$

**Beweis:**

Alle Vektoren in  $\mathbf{V}$  sind mit  $\text{Dim } \mathbf{V}$  Basisvektoren darstellbar. Da  $\mathbf{V}$  Vektoren enthält, die nicht in  $\mathbf{U}$  sind und  $\mathbf{U}$  ein Vektorraum ist folgt  $\text{Dim } \mathbf{U} < \text{Dim } \mathbf{V}$ .  $\square$

**Beispiel 3.22**

Sei  $\mathbf{V} = \mathbb{R}^3$ . Dann gilt  $\text{Dim } \mathbf{V} = 3$ , da sich mit drei Basisvektoren (z.B. der kartesischen Basis) jeder Vektor im  $\mathbb{R}^3$  darstellen lässt.

- Eine Gerade durch den Ursprung im  $\mathbb{R}^3$  bildet einen Unterraum der Dimension eins, z.B.

$$\left\{ \mathbf{u} : \mathbf{u} = \lambda \cdot (1, 2, 3)^T, \quad \lambda \in \mathbb{R} \right\}$$

- Eine ebene Fläche durch den Ursprung im  $\mathbb{R}^3$  bildet einen Unterraum der Dimension zwei, z.B.

$$\left\{ \mathbf{u} : \mathbf{u} = \lambda \cdot (1, 2, 3)^T + \mu \cdot (-1, 2, 3)^T, \quad \lambda, \mu \in \mathbb{R} \right\}$$

- Eine Gerade oder Fläche, die nicht durch den Ursprung geht, bildet hingegen keinen Unterraum, da sie den Nullvektor nicht enthält, z.B.

$$\left\{ \mathbf{u} : \mathbf{u} = \lambda \cdot (1, 2, 3)^T + (1, 0, 0)^T, \quad \lambda \in \mathbb{R} \right\}$$



## 4. Determinanten

Determinanten wurden erfunden um die eindeutige Lösbarkeit von linearen Gleichungssystemen zu bestimmen<sup>2</sup>.

### Definition 4.1 (Determinante)

Die Determinante einer quadratischen Matrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$  wird als  $\det \mathbf{A}$  oder auch  $|\mathbf{A}|$  bezeichnet und lässt sich rekursiv definieren als:

$$n = 1$$

$$\det \mathbf{A} = |\mathbf{A}| = |(a_{11})| := a_{11}$$

$$n \geq 2$$

Rekursive Entwicklung nach einer Zeile oder Spalte:

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

Dabei ist  $\mathbf{A}_{ij}$  die  $(n-1) \times (n-1)$  Matrix, die aus  $\mathbf{A} \in \mathbb{R}^{n \times n}$  durch Streichen der  $i$ -ten Zeile und  $j$ -ten Spalte entsteht. Dies wird rekursiv solange wiederholt, bis  $n = 1$ .

### Beispiel 4.2

$$\begin{aligned} \mathbf{A} &= \begin{pmatrix} 4 & -2 & -2 \\ -8 & 6 & -2 \\ 0 & 4 & -11 \end{pmatrix} \\ |\mathbf{A}| &= \begin{vmatrix} 4 & -2 & -2 \\ -8 & 6 & -2 \\ 0 & 4 & -11 \end{vmatrix} = (-1)^{1+1} \cdot 4 \cdot \begin{vmatrix} 6 & -2 \\ 4 & -11 \end{vmatrix} + (-1)^{2+1} \cdot (-8) \cdot \begin{vmatrix} -2 & -2 \\ 4 & -11 \end{vmatrix} + (-1)^{3+1} \cdot 0 \cdot \begin{vmatrix} -2 & -2 \\ 6 & -2 \end{vmatrix} \\ &= 4 \cdot [6 \cdot (-11) - 4 \cdot (-2)] + 8 \cdot [(-2) \cdot (-11) - 4 \cdot (-2)] = 4 \cdot (-58) + 8 \cdot 30 = 8 \end{aligned}$$

### Anmerkung 4.3

Es gibt verkürzte Rechenverfahren für Matrizen mit  $n = 2$  oder  $n = 3$  (Sarrus'sche Regel):

$$n = 2$$

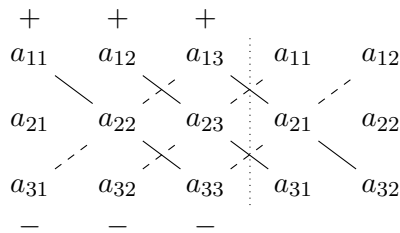
$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

$$n = 3$$

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

<sup>2</sup>China (3. Jh. A.D.), Cardano (1501-1576), Leibniz (1646-1716)

Merkhilfe *nur für*  $n = 3$ :



Schreibe erste und zweite Spalte noch einmal hinter die Matrix, addiere das Produkt der Werte in den drei Diagonalen von links oben nach rechts unten und subtrahiere das Produkt der Werte in den Diagonalen von links unten nach rechts oben.

**Wichtig:** Die Sarrus'sche Regel ist nur für  $n = 2$  und  $n = 3$  anwendbar. Für größere Matrizen liefert sie *nicht* das richtige Ergebnis.

#### Beispiel 4.4

$$\mathbf{A} = \begin{pmatrix} 4 & -2 & -2 \\ -8 & 6 & -2 \\ 0 & 4 & -12 \end{pmatrix}$$

$$|\mathbf{A}| = \begin{vmatrix} 4 & -2 & -2 \\ -8 & 6 & -2 \\ 0 & 4 & -12 \end{vmatrix}$$

$$= 4 \cdot 6 \cdot (-12) + (-2) \cdot (-2) \cdot 0 + (-2) \cdot (-8) \cdot 4$$

$$- 0 \cdot 6 \cdot (-2) - 4 \cdot (-2) \cdot 4 - (-12) \cdot (-8) \cdot (-2)$$

$$= 0$$

#### 3-Minutes Python 6

Zur Berechnung der Determinante gibt es die **numpy**-Funktion `linalg.det`, die als Argument das Array mit der Matrix erwartet, deren Determinante berechnet werden soll.

```
>>> import numpy as np
2 >>> A=np.array([[4,-2,-2],[-8,6,-2],[0,4,-12]])
>>> np.linalg.det(A)
4 0.0
>>> A=np.array([[4,-2,-2],[-8,6,-2],[0,4,-11]])
6 >>> np.linalg.det(A)
7.9999999999999947
```

Wie man sieht, erhält man ein Ergebnis, das aufgrund von Rundungsfehlern nur näherungsweise richtig ist. Ein besseres Ergebnis erhält man mit **sympy**. Dort gibt es den Datentyp **Matrix**. Dieser hat eine Methode `det`, die wie folgt benutzt wird:

```
>>> import sympy as sp
2 >>> A=sp.Matrix([[4,-2,-2],[-8,6,-2],[0,4,-11]])
```



```
>>> A.det()
4 8
```

Da die Berechnungen mit symbolischen Berechnungen durchgeführt werden, erhält man das exakte Ergebnis. Allerdings ist die Berechnung in **numpy** effizienter und daher schneller.

#### Satz 4.5

Für eine Dreiecksmatrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$  gilt:

$$\det \mathbf{A} = a_{11} a_{22} \cdots a_{nn} = \prod_{i=1}^n a_{ii}$$

#### Beweis:

Obere Dreiecksmatrix:

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

Da in der ersten Spalte außer  $a_{11}$  lauter Nullen stehen, bleibt bei Entwicklung nach dieser Spalte nur ein Term übrig, das geht dann rekursiv so weiter (analog ginge auch eine Entwicklung nach der letzten Zeile).

Untere Dreiecksmatrix: Entsprechende Entwicklung nach erster Zeile oder letzter Spalte.  $\square$

#### Satz 4.6 (Rechenregeln für Determinanten)

1.  $\det \mathbf{A}^T = \det \mathbf{A}$
2.  $\det (\mathbf{A} \cdot \mathbf{B}) = (\det \mathbf{A}) \cdot (\det \mathbf{B})$  (*Determinantenproduktsatz*)
3. *Elementare Zeilenumformungen:*  
Entsteht eine Matrix  $\tilde{\mathbf{A}} \in \mathbb{R}^{n \times n}$  aus an einer Matrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$  durch
  - Vertauschung von zwei Zeilen oder Spalten  $\Rightarrow \det \tilde{\mathbf{A}} = -\det \mathbf{A}$
  - Multiplizieren einer Zeile (oder Spalte) mit  $\lambda \in \mathbb{R} \Rightarrow \det \tilde{\mathbf{A}} = \lambda \det \mathbf{A}$
  - Addieren einer Zeile zu einer anderen  $\Rightarrow \det \tilde{\mathbf{A}} = \det \mathbf{A}$
  - Auch das Addieren des Vielfachen einer Zeile zu einer anderen ändert den Wert der Determinante nicht
4.  $\det(\lambda \mathbf{A}) = \lambda^n \det \mathbf{A}$

#### Beweis:

1. Große Übung
- 2+3. ohne Beweis.  
Addieren des Vielfachen einer Zeile entspricht Multiplizieren der Zeile mit  $\lambda$ , Addieren der Zeile zu einer anderen und anschließend Multiplizieren der Zeile mit  $\frac{1}{\lambda}$

$$\Rightarrow \det \tilde{\mathbf{A}} = \lambda \cdot \frac{1}{\lambda} \det \mathbf{A} = \det \mathbf{A}$$

4. folgt aus den Regel für das Multiplizieren einer Zeile mit  $\lambda$

□

#### Satz 4.7

Sei  $\mathbf{A} \in \mathbb{R}^{n \times n}$ , dann gilt

$$\det \mathbf{A} = 0 \Leftrightarrow \text{Rang } \mathbf{A} < n.$$

#### Beweis:

Wird die Matrix  $\mathbf{A}$  mit elementaren Zeilenumformungen in Treppenform  $\mathbf{A}^*$  gebracht, dann gibt es wegen Satz 4.6(3) ein  $\omega \in \mathbb{R} \setminus \{0\}$  so dass

$$\det \mathbf{A} = \omega \det \mathbf{A}^*.$$

Da  $\mathbf{A}^*$  eine obere Dreiecksmatrix ist, gilt  $\det \mathbf{A}^* = \prod_{i=1}^n a_{ii}^*$ .

$$\text{Rang } \mathbf{A} = n \Leftrightarrow a_{ii}^* \neq 0 \forall 1 \leq i \leq n \Leftrightarrow \det \mathbf{A}^* \neq 0 \Leftrightarrow \det \mathbf{A} \neq 0$$

Daraus folgt die Behauptung.

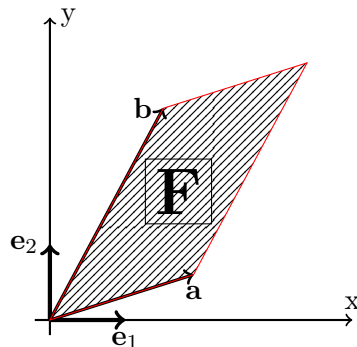
□

#### Beispiel 4.8

- Lineare Unabhängigkeit

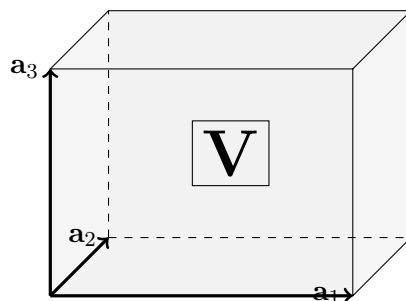
Vektoren  $\mathbf{v}_i \in \mathbb{R}^n$ ,  $1 \leq i \leq n$  sind genau dann linear abhängig, wenn die Determinante der Matrix, bei der die Vektoren  $\mathbf{v}_i$  die Spalten (oder Zeilen) bilden, gleich Null ist.

- Fläche eines Parallelogramms



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

- Volumen eines Spats



Der Spat mit den Kanten

$$\mathbf{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}, \quad \mathbf{a}_3 = \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix}$$

hat das Volumen

$$V = \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Die drei Vektoren  $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$  sind positiv orientiert/bilden ein Rechtssystem, wenn  $V$  positiv ist. Sie stehen dann so zueinander wie Daumen, Zeigefinger und Mittelfinger der rechten Hand.

- Kramer'sche Regel

Sei  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$  mit  $\mathbf{A} \in \mathbb{R}^{n \times n}$  und  $\det \mathbf{A} \neq 0$  sowie  $\mathbf{x}, \mathbf{b} \in \mathbb{R}^n$ . Dann gilt:

$$x_i = \frac{\det \mathbf{A}_i}{\det \mathbf{A}}$$

Dabei ist  $\mathbf{A}_i$  eine Matrix bei der die  $i$ -te Spalte von  $\mathbf{A}$  durch  $\mathbf{b}$  ersetzt wurde, also

$$\mathbf{A}_i = \begin{bmatrix} a_{11} & \dots & a_{1(i-1)} & b_1 & a_{1(i+1)} & \dots & a_{1n} \\ a_{21} & \dots & a_{2(i-1)} & b_2 & a_{2(i+1)} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{n(i-1)} & b_n & a_{n(i+1)} & \dots & a_{nn} \end{bmatrix}$$

Beispiel:

$$\begin{aligned} 2x + 3y &= 3 \\ 5x - 7y &= -1 \end{aligned}$$

$$x = \frac{\begin{vmatrix} 3 & 3 \\ -1 & -7 \end{vmatrix}}{\begin{vmatrix} 2 & 3 \\ 5 & -7 \end{vmatrix}} = \frac{-18}{-29} = \frac{18}{29}$$

$$y = \frac{\begin{vmatrix} 2 & 3 \\ 5 & -1 \end{vmatrix}}{\begin{vmatrix} 2 & 3 \\ 5 & -7 \end{vmatrix}} = \frac{-17}{-29} = \frac{17}{29}$$



## 5. Lineare Gleichungssysteme revisited

### 5.1. Spaltenraum

Zurück zur Frage:

Für welche rechten Seiten  $\mathbf{b} \in \mathbb{R}^m$  ist ein lineares Gleichungssystem mit einer bestimmten Matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$  lösbar?

Oder andersherum:

Welche  $\mathbf{b} \in \mathbb{R}^m$  lassen sich mit beliebigen  $\mathbf{x} \in \mathbb{R}^n$  als  $\mathbf{A} \cdot \mathbf{x}$  erzeugen/darstellen?

#### Definition 5.1 (Spaltenraum)

Sei  $\mathbf{A} \in \mathbb{R}^{m \times n}$  und  $\mathbf{a}^{(j)} \in \mathbb{R}^m$  die  $j$ -te Spalte von  $\mathbf{A}$ , also  $\mathbf{a}_i^{(j)} = a_{ij} \forall 1 \leq i \leq m, 1 \leq j \leq n$ .

Die Spalten von  $\mathbf{A}$  erzeugen einen Vektorraum, dieser heißt Spaltenraum oder auch Bild  $\mathbf{A}$  (weil er die Menge aller Vektoren ist, auf die sich  $\mathbf{x}$  mit Hilfe von  $\mathbf{A}$  abbilden lässt).

Der von den Zeilen erzeugte Vektorraum heißt Zeilenraum.

#### Satz 5.2

Die Dimension des Spaltenraums einer Matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$  ist gleich dem Rang von  $\mathbf{A}$  und auch gleich der Dimension des Zeilenraums.

ohne Beweis

#### Satz 5.3

Ein lineares Gleichungssystem  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$  mit  $\mathbf{A} \in \mathbb{R}^{m \times n}$ ,  $\mathbf{b} \in \mathbb{R}^m$ ,  $\mathbf{x} \in \mathbb{R}^n$  ist lösbar, falls  $\mathbf{b}$  im Spaltenraum von  $\mathbf{A}$  liegt. Es ist eindeutig lösbar, wenn die Spalten von  $\mathbf{A}$  eine Basis des Spaltenraums bilden.

#### Beweis:

Ein lineares Gleichungssystem ist gerade die Frage nach den Koeffizienten der Spaltenvektoren mit denen sich  $\mathbf{b}$  als Linearkombination der Spalten von  $\mathbf{A}$  ergibt. Wenn die Spaltenvektoren eine Basis des Spaltenraums bilden (also linear unabhängig sind) gibt es genau eine Darstellung jedes Vektors in der Basis (Satz 3.15)  $\Rightarrow$  das lineare Gleichungssystem ist eindeutig lösbar.  $\square$

**Frage:** Wie finden wir eine Basis des Spaltenraums von  $\mathbf{A}$ ?

Sei  $\mathbf{A}$  durch elementare Spaltenumformungen zu einer äquivalenten Matrix  $\mathbf{A}^*$  in Treppenform umgeformt, die  $p$  Nicht-Nullzeilen hat. Sie sieht dann in etwa wie folgt aus (wobei  $e_i$  für Werte ungleich Null steht und  $*$  für beliebige Werte):

$$\mathbf{A}^* = \begin{pmatrix} e_1 & \dots & * & * & * & * & \dots & * & * \\ 0 & \dots & 0 & e_2 & * & * & \dots & * & * \\ 0 & \dots & 0 & 0 & 0 & e_3 & \dots & * & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & e_p \\ 0 & \dots & 0 & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Wir nennen die Spalte  $j_i$  in Zeile  $i$  in der zum ersten Mal ein Matricelement nicht Null ist (also  $a_{ik}^* = 0 \quad \forall k < j_i$ ) eine Pivotspalte.

Eine Matrix mit Rang  $p$  hat  $p$  Pivotspalten. Da von links nach rechts jede Pivotspalte eine Zeile mehr enthält, in der Werte stehen, die nicht Null sind hat  $\sum_{i=1}^p \mathbf{a}^{*(j_i)} \cdot x_i = \mathbf{0}$  offensichtlich nur die Lösung  $x_i = 0 \quad \forall 1 \leq i \leq p$ . Daher sind die Pivotspalten linear unabhängig.

Weil  $\mathbf{A}^*$  zu  $\mathbf{A}$  äquivalent ist, gilt dies auch für  $\sum_{i=1}^p \mathbf{a}^{(j_i)} \cdot x_i = \mathbf{0}$ . Damit sind auch die entsprechenden Spalten von  $\mathbf{A}$  linear unabhängig.

Da die  $p$  Pivotspalten linear unabhängig sind und der Spaltenraum ebenfalls die Dimension  $p$  hat, müssen sie eine Basis des Spaltenraums bilden.

$$\Rightarrow \text{Span}(\mathbf{a}^{(1)}, \mathbf{a}^{(2)}, \dots, \mathbf{a}^{(n)}) = \text{Span}(\mathbf{a}^{(j_1)}, \mathbf{a}^{(j_2)}, \dots, \mathbf{a}^{(j_p)}) = \text{Bild } \mathbf{A}$$

Damit folgt der

#### Satz 5.4 (Rangsatz)

Sei  $\mathbf{A} \in \mathbb{R}^{m \times n}$ ,  $p = \text{Rang}(\mathbf{A})$  und  $\mathbf{A}^*$  eine zu  $\mathbf{A}$  äquivalente Matrix in Treppenform.

Für  $i = 1, \dots, p$  sei  $j_i := \min \{j : a_{ij}^* \neq 0\}$  (also eine Pivotspalte).

Dann ist  $\mathbf{a}^{(j_1)}, \mathbf{a}^{(j_2)}, \dots, \mathbf{a}^{(j_p)}$  eine Basis von Bild  $\mathbf{A}$ , d.h. des Raumes aller  $\mathbf{b}$ , für die  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$  eine Lösung hat.

Insbesondere gilt:

$\text{Rang}(\mathbf{A}) = \text{Dim Bild } \mathbf{A}$  (d.h. der Rang einer Matrix ist gleich der Dimension des Spaltenraums)

und

$$\text{Dim Bild } \mathbf{A} + \text{Dim Kern } \mathbf{A} = n.$$

(d.h. die Dimension der Matrix minus der Dimension des Spaltenraums ist die Dimension des Nullraums)

## 5.2. Bestimmung einer Basis von Kern $\mathbf{A}$ (Nullraum)

Sei  $\mathbf{A}^*$  eine zu  $\mathbf{A} \in \mathbb{R}^{m \times n}$  äquivalente Matrix in Treppenform.  $\mathbf{A}$  habe Rang  $p$ , es gibt also  $p$  Nicht-Nullzeilen in  $\mathbf{A}^*$  und  $n - p$  freie Parameter.

Wir schreiben eine  $n \times n$  Einheitsmatrix unter  $\mathbf{A}^*$ . Analog zur Gaußelimination verwenden wir nun elementare *Spaltenumformungen* um die Nicht-Pivotspalten von  $\mathbf{A}^*$  zu Nullspalten zu machen. Wir suchen also nach einer Lösung von  $\sum_{i=1}^p \mathbf{a}^{*(j_i)} \cdot x_i + \mathbf{a}^{*(k)} = 0$ , wobei  $k$  der Index der jeweiligen Nicht-Pivotspalte ist (für die anderen Nicht-Pivotspalten wird  $x_i$  dabei implizit auf Null gesetzt). Diese Spaltenumformungen wenden wir dabei auch auf die Einheitsmatrix an.

Jede Nicht-Pivotspalte enthält dann am Ende einen Basisvektor von Kern  $\mathbf{A}$ .

#### Beispiel 5.5

$$\mathbf{A} = \begin{pmatrix} 4 & -2 & -2 \\ -8 & 6 & -2 \\ 0 & 4 & -12 \end{pmatrix}, \quad \mathbf{A}^* = \begin{pmatrix} 4 & -2 & -2 \\ 0 & 2 & -6 \\ 0 & 0 & 0 \end{pmatrix}$$

Pivotspalten von  $\mathbf{A}^*$ :  $j = 1, 2$ , Forme Spalte 3 zu Nullspalte um:

$$\mathbf{A}^* = \begin{pmatrix} 4 & -2 & -2 \\ 0 & 2 & -6 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{+3 \cdot 2. \text{ Spalte}} \begin{pmatrix} 4 & -2 & -8 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{+2 \cdot 1. \text{ Spalte}} \begin{pmatrix} 4 & -2 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\Rightarrow \text{Kern } \mathbf{A} = \text{Span} \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$$

### 5.3. Bestimmung einer speziellen Lösung von $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$

Sei  $(\mathbf{A}^* | \mathbf{b}^*)$  eine zu  $(\mathbf{A} | \mathbf{b})$  äquivalente Matrix in Treppenform.  $\mathbf{A}$  habe, ebenso wie  $(\mathbf{A} | \mathbf{b})$ , den Rang  $p$ . Das System ist also lösbar.

Schreibt man die ersten  $p$  Zeilen der Pivotspalten von  $\mathbf{A}^*$  nebeneinander, erhält man eine obere Dreiecksmatrix. Zusammen mit den ersten  $p$  Zeilen von  $\mathbf{b}^*$  gibt dies ein (durch Rückwärtseinsetzen) eindeutig lösbares lineares Gleichungssystem für die  $p$  Unbekannten  $x_{j_i}$   $1 \leq i \leq p$ . Setzt man für die restlichen  $n-p$  Unbekannten Null ein, erhält man eine spezielle Lösung des linearen Gleichungssystems.

#### Beispiel 5.6

Mit den Pivotspalten der Matrix in Treppenform und der rechten Seite aus Beispiel 2.27, 3.:

$$\begin{pmatrix} 4 & -2 \\ 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 8 \end{pmatrix} \quad x_1 = 3, x_2 = 4$$

$$\Rightarrow \mathbf{x} = \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix}$$

### 5.4. Vollständige Charakterisierung eines linearen Gleichungssystems

#### Algorithmus 5.7

Sei  $\mathbf{A} \in \mathbb{R}^{m \times n}$ ,  $\mathbf{b} \in \mathbb{R}^m$  und  $\mathbf{x} \in \mathbb{R}^n$  die Lösung von  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ .

1. Bringe die erweiterte Koeffizientenmatrix  $(\mathbf{A} | \mathbf{b})$  durch elementare Zeilenumformungen in Treppenform  $(\mathbf{A}^* | \mathbf{b}^*)$  (Algorithmus 2.29).
2. Fallunterscheidung
  - a) Ist  $\text{Rang}(\mathbf{A}^*) = n$ :

$$\Rightarrow \text{Kern } \mathbf{A} = \{\mathbf{0}\}$$

Wenn  $m = n \Rightarrow \text{Bild } \mathbf{A} = \mathbb{R}^n$ , ansonsten ist das Bild von  $\mathbf{A}$  ein  $n$ -dimensionaler Unterraum des  $\mathbb{R}^m$ .

Fallunterscheidung:

- i.  $\text{Rang}(\mathbf{A}^* | \mathbf{b}^*) = n \Rightarrow$  genau eine Lösung, bestimmbar durch Rückwärtseinsetzen (Algorithmus 2.37).
  - ii.  $\text{Rang}(\mathbf{A}^* | \mathbf{b}^*) > n \Rightarrow$  keine Lösung, da  $\mathbf{b}$  nicht in Bild  $\mathbf{A}$  (überdeterminiertes System)
- b) Ist  $\text{Rang}(\mathbf{A}^*) = p < n$ :
- Die Spalten  $j_i$  ( $1 \leq i \leq p$ ) von  $\mathbf{A}$ , die Pivotspalten von  $\mathbf{A}^*$  sind, bilden eine Basis von Bild  $\mathbf{A}$  (siehe Abschnitt 5.1). Dies ist ein  $p$ -dimensionaler Unterraum des  $\mathbb{R}^m$ .
- Durch elementare Spaltenumformungen an  $\mathbf{A}^*$ , mit denen die Nicht-Pivotspalten zu Nullspalten gemacht werden, erhält man aus den entsprechenden Spalten einer unter die Matrix  $\mathbf{A}^*$  geschriebenen  $n \times n$  Einheitsmatrix eine Basis von Kern  $\mathbf{A}$  (siehe Abschnitt 5.2).
- Fallunterscheidung:
- i.  $\text{Rang}(\mathbf{A}^* | \mathbf{b}^*) = p \Rightarrow$  unendlich viele Lösungen. Bestimme eine spezielle Lösung  $\mathbf{v}$  mit den ersten  $p$  Zeilen der Pivotspalten und von  $\mathbf{b}^*$  (siehe Abschnitt 5.3). Alle möglichen Lösungen sind dann gegeben durch  $\mathbf{v} + \text{Kern } \mathbf{A}$ .
  - ii.  $\text{Rang}(\mathbf{A}^* | \mathbf{b}^*) > p \Rightarrow$  keine Lösung, da  $\mathbf{b}$  nicht in Bild  $\mathbf{A}$  (überdeterminiertes System)

### Beispiel 5.8

$$\mathbf{A} = \left( \begin{array}{cccc|c} 2 & 1 & 0 & 3 & 4 \\ -6 & -3 & 3 & -6 & -11 \\ 0 & 0 & 6 & 6 & 2 \end{array} \right) + 3 \cdot 1. \text{ Zeile}$$

$$\left( \begin{array}{cccc|c} 2 & 1 & 0 & 3 & 4 \\ 0 & 0 & 3 & 3 & 1 \\ 0 & 0 & 6 & 6 & 2 \end{array} \right) - 2 \cdot 2. \text{ Zeile}$$

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

Der Rang von  $\mathbf{A}$  ist also 2. Die 1. und die 3. Spalte sind Pivotspalten. Die Menge aller  $\mathbf{b}$ , die sich als Ergebnis von  $\mathbf{A} \cdot \mathbf{x}$  für beliebiges  $\mathbf{x}$  darstellen lassen (also das Bild von  $\mathbf{A}$ ), sind gegeben durch:

$$\text{Bild } \mathbf{A} = \text{Span} \left( \begin{pmatrix} 2 \\ -6 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 6 \end{pmatrix} \right)$$

Das entspricht einer Ebene im dreidimensionalen Raum.

Da  $\text{Rang}(\mathbf{A}^* | \mathbf{b}^*) = \text{Rang}(\mathbf{A}^*) = 2 < 4$  gibt es unendlich viele Lösungen mit 2 frei wählbaren Parametern.

Wir wollen nun alle möglichen Lösungen bestimmen. Zunächst eine spezielle Lösung. Dafür



nehmen wir die ersten 2 Zeilen der 2 Pivotspalten von  $\mathbf{A}^*$  und der rechten Seite  $\mathbf{b}^*$

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix} \Rightarrow x_3 = \frac{1}{3}, x_1 = 2 \Rightarrow \mathbf{x} = \begin{pmatrix} 2 \\ 0 \\ 1/3 \\ 0 \end{pmatrix}$$

Jetzt fehlt noch eine Basis für den Kern von  $\mathbf{A}$ , also die Menge aller Lösungen des homogenen Gleichungssystems. Dazu müssen wir die zweite und die vierte Spalte von  $\mathbf{A}^*$  zu Nullspalten machen:

$$\begin{aligned} \mathbf{A}^* &= \begin{pmatrix} 2 & 1 & 0 & 3 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{2.\text{Spalte} - 1/2 \cdot 1.\text{Spalte}} \begin{pmatrix} 2 & 0 & 0 & 3 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \\ 1 & -1/2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ &\xrightarrow{4.\text{Spalte} - 3/2 \cdot 1.\text{Spalte}} \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 \\ 1 & -1/2 & 0 & -3/2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{4.\text{Spalte} - 1 \cdot 3.\text{Spalte}} \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & -1/2 & 0 & -3/2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ \Rightarrow \text{Kern } \mathbf{A} &= \text{Span} \left( \begin{pmatrix} -1/2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -3/2 \\ 0 \\ -1 \\ 1 \end{pmatrix} \right) \end{aligned}$$

Alle möglichen Lösungen von  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$  sind dann gegeben durch

$$\mathbf{x} = \begin{pmatrix} 2 \\ 0 \\ 1/3 \\ 0 \end{pmatrix} + \lambda_1 \cdot \begin{pmatrix} -1/2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} -3/2 \\ 0 \\ -1 \\ 1 \end{pmatrix}$$

mit  $\lambda_1, \lambda_2 \in \mathbb{R}$ . Die Lösungen liegen in einer Ebene im  $\mathbb{R}^4$ .

### 3-Minutes Python 7

In Python können wir die vollständige Lösung mit Hilfe von `sympy`. In `sympy` lässt sich der Rang einer Matrix mit Hilfe der Methode `rank` bestimmen. Hat die Matrix vollen Rang, dann bestimmt man die Lösung mit der Methode `solve`:

```
>>> import sympy as sp
2 >>> A=sp.Matrix([[4,-2,-2],[-8,6,-2],[0,4,-11]])
>>> A.rank()
4 3
>>> b=sp.Matrix([4,0,17])
```

```

6 >>> A.solve(b)
  Matrix([
8 [5],
  [7],
10 [1]])

```

Ist das Gleichungssystem nicht eindeutig lösbar, dann müssen wir zunächst prüfen ob überhaupt eine Lösung existiert. Dabei spart uns die Methode `col_insert` Arbeit, deren erstes Argument die Spalte ist, nach der die Spalte eingefügt werden soll, die als zweites Argument gegeben wird.

```

>>> A=sp.Matrix([[2,1,0,3],[-6,-3,3,-6],[0,0,6,6]])
2 >>> A.rank()
  2
4 >>> b=sp.Matrix([4,-11,2])
>>> A.col_insert(4,b)
6 Matrix([
  [ 2,  1,  0,  3,  4],
8 [-6, -3,  3, -6, -11],
  [ 0,  0,  6,  6,  2]])
10 >>> A.col_insert(4,b).rank()
  2

```

In diesem Fall ist das Gleichungssystem also lösbar. Wir erhalten den Bildraum der Matrix mit der Methode `columnspace` und den Kern der Matrix mit der Methode `nullspace`

```

>>> A.columnspace()
2 [Matrix([
  [ 2],
4 [-6],
  [ 0]])], Matrix([
6 [0],
  [3],
8 [6]])]
>>> A.nullspace()
10 [Matrix([
  [-1/2],
12 [  1],
  [  0],
14 [  0]])], Matrix([
  [-3/2],
16 [  0],
  [-1],
18 [  1]])]

```

Mit Hilfe der Methode `gauss_jordan_solve` berechnet `sympy` auch direkt die allgemeine Lösung. Diese Methode liefert als ersten Rückgabewert die allgemeine Lösung und als zweiten Wert eine Matrix mit den Namen der freien Parameter.

```

>>> A.gauss_jordan_solve(b)
2 (Matrix([
3  [-_tau0/2 - 3*_tau1/2 + 2],
4  [          _tau0],
5  [          -_tau1 + 1/3],
6  [          _tau1]]) , Matrix([
7  [_tau0],
8  [_tau1]]))

```

Möchte man allerdings nur numerisch die Lösung eines großen eindeutig lösbaren linearen Gleichungssystems bestimmen, dann ist die Funktion `solve` in `numpy` eindeutig die schnellere Lösung.

## 5.5. Schlussbemerkung

### Anmerkung 5.9

Wir können aus den Überlegungen einige übliche Vorstellungen widerlegen:

- Gibt es bei einem linearen Gleichungssystem weniger Gleichungen als Unbekannte, kann dieses trotzdem unlösbar sein. Mathematisch heißt das, dass  $\mathbf{b}$  nicht in Bild  $\mathbf{A}$  liegt, anschaulich heißt es, dass sich Gleichungen widersprechen (z.B.  $x_1 + x_2 + x_3 = 1$  und  $x_1 + x_2 + x_3 = 0$ ).
- Gibt es bei einem linearen Gleichungssystem mehr Gleichungen als Unbekannte, kann dieses trotzdem lösbar sein. Mathematisch dann, wenn der Rang von  $(\mathbf{A} | \mathbf{b})$  nicht größer als  $n$  und gleich dem Rang von  $\mathbf{A}$  ist, weil Gleichungen linear abhängig sind. Anschaulich bedeutet das, dass mehrere Gleichungen die gleichen Informationen enthalten.



## 6. Inverse Matrix

### Definition 6.1

Eine quadratische Matrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$  heißt invertierbar, wenn es eine Matrix  $\mathbf{B} \in \mathbb{R}^{n \times n}$  gibt, so dass  $\mathbf{A} \cdot \mathbf{B} = \mathbf{B} \cdot \mathbf{A} = \mathbf{I}$ . Dann ist die Matrix  $\mathbf{B}$  eindeutig und wird mit  $\mathbf{A}^{-1}$  bezeichnet (d.h.  $\mathbf{A}^{-1} := \mathbf{B}$ ). Sie heißt zu  $\mathbf{A}$  inverse Matrix oder Inverse von  $\mathbf{A}$ .

### Satz 6.2

Gibt es zwei Matrizen  $\mathbf{B}, \mathbf{C} \in \mathbb{R}^{n \times n}$  zur Matrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$  mit  $\mathbf{B} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{C} = \mathbf{I}$ , dann ist  $\mathbf{A}$  invertierbar und  $\mathbf{B} = \mathbf{C} = \mathbf{A}^{-1}$ .

**Beweis:**

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

Also gilt  $\mathbf{B} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{B} = \mathbf{I}$  und  $\mathbf{B} = \mathbf{A}^{-1}$ . □

### Beispiel 6.3

$$\mathbf{A} = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}, \quad \mathbf{A}^{-1} = \begin{pmatrix} -2 & 3/2 \\ 1 & -1/2 \end{pmatrix}$$

$$\begin{aligned} \mathbf{A} \cdot \mathbf{A}^{-1} &= \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} -2 & 3/2 \\ 1 & -1/2 \end{pmatrix} \\ &= \begin{pmatrix} -2+3 & 1 \cdot 3/2 - 3 \cdot 1/2 \\ -4+4 & 2 \cdot 3/2 - 4/2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \mathbf{A}^{-1} \cdot \mathbf{A} &= \begin{pmatrix} -2 & 3/2 \\ 1 & -1/2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \\ &= \begin{pmatrix} -2+2 \cdot 3/2 & -2 \cdot 3 + 4 \cdot 3/2 \\ 1-2 \cdot 1/2 & 3-4/2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

### Satz 6.4

1. Die Inverse einer invertierbaren Matrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$  ist invertierbar mit

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

2. Das Produkt zweier invertierbarer  $n \times n$  Matrizen ist invertierbar mit

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

3. Die Transponierte  $\mathbf{A}^T$  einer Matrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$  ist genau dann invertierbar, wenn  $\mathbf{A}$  invertierbar ist. Es gilt:

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

**Beweis:**

$$1. \mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$$

2.

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

3.

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

□

Aus dem Satz folgt, dass  $(\mathbf{A}_1 \cdot \mathbf{A}_2 \cdot \dots \cdot \mathbf{A}_k)^{-1} = \mathbf{A}_k^{-1} \cdot \dots \cdot \mathbf{A}_2^{-1} \cdot \mathbf{A}_1^{-1}$

### Satz 6.5

Seien  $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^{n \times n}$ ,  $\mathbf{x} \in \mathbb{R}^n$ , dann sind folgende Aussagen äquivalent:

1.  $\mathbf{A}$  ist invertierbar.
2. Es gibt eine Matrix  $\mathbf{B}$  mit  $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}$ .
3. Es gibt eine Matrix  $\mathbf{C}$  mit  $\mathbf{C} \cdot \mathbf{A} = \mathbf{I}$ .
4.  $\mathbf{A} \cdot \mathbf{x} = \mathbf{0} \Rightarrow \mathbf{x} = \mathbf{0}$
5.  $\text{Rang}(\mathbf{A}) = n$ .

**Beweis:**

$$1) \mathbf{C} = \mathbf{A}^{-1} \Rightarrow \mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{I} \Rightarrow 3)$$

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{0} = \mathbf{C} \cdot \mathbf{0} = \mathbf{C} \cdot \mathbf{A} \cdot \mathbf{x} \stackrel{3)}{=} \mathbf{I} \cdot \mathbf{x} \Rightarrow 4)$$

$$4) \iff 5) \text{ Satz 2.34}$$

$$5) \Rightarrow \mathbf{A} \cdot \mathbf{x} = \mathbf{b} \text{ eindeutig lösbar für alle } \mathbf{b} \in \mathbb{R}^n.$$

Insbesondere gibt es dann  $\mathbf{b}^{(i)} \in \mathbb{R}^n$  mit  $\mathbf{A} \cdot \mathbf{b}^{(i)} = \mathbf{e}_i$  (i-ter Einheitsvektor)  
mit  $\mathbf{B} = (\mathbf{b}^{(1)}, \mathbf{b}^{(2)}, \dots, \mathbf{b}^{(n)})$

$$\Rightarrow \mathbf{A} \cdot \mathbf{B} = (\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n) = \mathbf{I} \Rightarrow 2)$$

$$2) + 3) \Rightarrow 1)$$

$$1) \Rightarrow 3) \Rightarrow 4) \Rightarrow 5) \Rightarrow 2) \Rightarrow 1)$$

Also sind alle Aussagen äquivalent

□

### Anmerkung 6.6

Die Inverse einer Diagonalmatrix  $\mathbf{D} \in \mathbb{R}^{n \times n}$  ist ebenfalls eine Diagonalmatrix mit  $(\mathbf{D}^{-1})_{ii} = 1/d_{ii}$

**Algorithmus 6.7 (Gauß-Jordan Verfahren)**

Wir suchen eine Matrix  $\mathbf{X} \in \mathbb{R}^{n \times n}$  mit

$$\mathbf{A} \cdot \mathbf{X} = \mathbf{I}$$

Dazu schreiben wir die Einheitsmatrix  $\mathbf{I} \in \mathbb{R}^{n \times n}$  neben die Matrix  $\mathbf{A}$  (entsprechend der erweiterten Koeffizientenmatrix, nur mit mehreren rechten Seiten):

$$(\mathbf{A} | \mathbf{I})$$

Dann formen wir  $\mathbf{A}$  durch elementare Zeilenumformungen erst in eine obere Dreiecksmatrix, anschließend in eine Diagonalmatrix und schließlich in die Identität um.

An der Stelle der ursprünglichen Einheitsmatrix steht dann  $\mathbf{A}^{-1}$

**Beispiel 6.8**

$$\left( \begin{array}{ccc|ccc} 4 & -2 & -2 & 1 & 0 & 0 \\ -8 & 6 & -2 & 0 & 1 & 0 \\ 0 & 4 & -11 & 0 & 0 & 1 \end{array} \right) + 2 \cdot 1. \text{ Zeile}$$

$$\left( \begin{array}{ccc|ccc} 4 & -2 & -2 & 1 & 0 & 0 \\ 0 & 2 & -6 & 2 & 1 & 0 \\ 0 & 4 & -11 & 0 & 0 & 1 \end{array} \right) - 2 \cdot 2. \text{ Zeile}$$

$$\left( \begin{array}{ccc|ccc} 4 & -2 & -2 & 1 & 0 & 0 \\ 0 & 2 & -6 & 2 & 1 & 0 \\ 0 & 0 & 1 & -4 & -2 & 1 \end{array} \right) + 2 \cdot 3. \text{ Zeile}$$

$$\left( \begin{array}{ccc|ccc} 4 & -2 & 0 & -7 & -4 & 2 \\ 0 & 2 & 0 & -22 & -11 & 6 \\ 0 & 0 & 1 & -4 & -2 & 1 \end{array} \right) + 2. \text{ Zeile}$$

$$\left( \begin{array}{ccc|ccc} 4 & 0 & 0 & -29 & -15 & 8 \\ 0 & 2 & 0 & -22 & -11 & 6 \\ 0 & 0 & 1 & -4 & -2 & 1 \end{array} \right)$$

$$\left( \begin{array}{ccc|ccc} 1 & 0 & 0 & -29/4 & -15/4 & 2 \\ 0 & 1 & 0 & -11 & -11/2 & 3 \\ 0 & 0 & 1 & -4 & -2 & 1 \end{array} \right)$$

$$\Rightarrow \mathbf{A}^{-1} = \begin{pmatrix} -29/4 & -15/4 & 2 \\ -11 & -11/2 & 3 \\ -4 & -2 & 1 \end{pmatrix}$$

Test:

$$\begin{pmatrix} -29/4 & -15/4 & 2 \\ -11 & -11/2 & 3 \\ -4 & -2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 4 & -2 & -2 \\ -8 & 6 & -2 \\ 0 & 4 & -11 \end{pmatrix}$$

$$= \begin{pmatrix} -29 + 30 + 0 & 29/2 - 45/2 + 8 & 29/2 + 15/2 - 22 \\ -44 + 44 + 0 & 22 - 33 + 12 & 22 + 11 - 33 \\ -16 + 16 + 0 & 8 - 12 + 4 & 8 + 4 - 11 \end{pmatrix}$$

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

### 3-Minutes Python 8

In `sympy` hat jede Matrix direkt eine Methode `inv`, die die Inverse zurückliefert, falls diese existiert. Der Multiplikationsoperator berechnet in `sympy` (im Gegensatz zu `numpy`) direkt das Matrixprodukt.

```
>>> import sympy as sp
2 >>> A=sp.Matrix([[4,-2,-2],[-8,6,-2],[0,4,-11]])
>>> print(A.inv())
4 Matrix([[-29/4, -15/4, 2], [-11, -11/2, 3], [-4, -2, 1]])
>>> print(A*A.inv())
6 Matrix([[1, 0, 0], [0, 1, 0], [0, 0, 1]])
>>> print(A.inv()*A)
8 Matrix([[1, 0, 0], [0, 1, 0], [0, 0, 1]])
```

In `numpy` gibt es eine entsprechende Funktion `np.linalg.inv`. Hier muss man für das Matrixprodukt zum Testen aber die Methode `dot` verwenden.

```
>>> A=np.array([[4,-2,-2],[-8,6,-2],[0,4,-11]])
2 >>> np.linalg.inv(A)
array([[ -7.25,  -3.75,   2.  ],
4        [-11.   ,  -5.5  ,   3.  ],
        [ -4.   ,  -2.   ,   1.  ]])
6 >>> A.dot(np.linalg.inv(A))
array([[ 1.,   0.,   0.],
8        [ 0.,   1.,   0.],
        [ 0.,   0.,   1.]])
10 >>> np.linalg.inv(A).dot(A)
array([[ 1.,   0.,   0.],
12        [ 0.,   1.,   0.],
        [ 0.,   0.,   1.]])
```



## 7. Längen und Winkel im $\mathbb{R}^n$

### 7.1. Skalarprodukt

#### Definition 7.1 (Skalarprodukt)

Ein Skalarprodukt oder inneres Produkt auf einem (reellen) Vektorraum ist eine positiv definite, symmetrische Bilinearform  $\langle \cdot, \cdot \rangle: \mathbf{V} \times \mathbf{V} \rightarrow \mathbb{R}$ , d.h. für  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$  und  $\lambda \in \mathbb{R}$  gilt:

1. bilinear

$$\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle \quad (7.1)$$

$$\langle \mathbf{u}, \mathbf{v} + \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle \quad (7.2)$$

$$\langle \mathbf{u}, \lambda \mathbf{v} \rangle = \lambda \langle \mathbf{u}, \mathbf{v} \rangle = \langle \lambda \mathbf{u}, \mathbf{v} \rangle \quad (7.3)$$

2. symmetrisch

$$\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle \quad (7.4)$$

3. positiv definit

$$\langle \mathbf{u}, \mathbf{u} \rangle \geq 0 \text{ und } \langle \mathbf{u}, \mathbf{u} \rangle = 0 \iff \mathbf{u} = \mathbf{0} \quad (7.5)$$

#### Satz 7.2 (Skalarprodukt im $\mathbb{R}^n$ )

Seien  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ , dann ist das Skalarprodukt gegeben durch

$$\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{x}^T \cdot \mathbf{y} = x_1 \cdot y_1 + x_2 \cdot y_2 + \cdots + x_n \cdot y_n = \sum_{i=1}^n x_i \cdot y_i$$

**Beweis:**

$$(7.1): \langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \sum_{i=1}^n (u_i + v_i) \cdot w_i = \sum_{i=1}^n (u_i \cdot w_i + v_i \cdot w_i) = \sum_{i=1}^n u_i \cdot w_i + \sum_{i=1}^n v_i \cdot w_i = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$$

analog für (7.2)

$$(7.3): \langle \mathbf{u}, \lambda \mathbf{v} \rangle = \sum_{i=1}^n u_i \cdot \lambda \cdot v_i = \lambda \cdot \sum_{i=1}^n u_i \cdot v_i = \lambda \cdot \langle \mathbf{u}, \mathbf{v} \rangle$$

(7.4): offensichtlich

$$(7.5): \langle \mathbf{u}, \mathbf{u} \rangle = \sum_{i=1}^n u_i^2 = \sum_{i=1}^n |u_i|^2 \geq \left( \max_{1 \leq i \leq n} |u_i| \right)^2 \geq 0$$
$$\implies \langle \mathbf{u}, \mathbf{u} \rangle = 0 \text{ falls } \max_{1 \leq i \leq n} |u_i| = 0 \iff \mathbf{u} = \mathbf{0}$$

□

**Anmerkung 7.3**

- Offensichtlich ist  $\langle \mathbf{u}, \sum_{i=1}^n (\alpha_i \cdot \mathbf{w}_i) \rangle = \sum_{i=1}^n \alpha_i \langle \mathbf{u}, \mathbf{w}_i \rangle$
- $\langle \mathbf{u}, \mathbf{0} \rangle = \langle \mathbf{0}, \mathbf{u} \rangle = 0$

**Beispiel 7.4**

$$\mathbf{x} = \begin{pmatrix} 3 \\ 4 \\ 1 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 7 \\ 0 \\ 2 \end{pmatrix}$$

$$\langle \mathbf{x}, \mathbf{y} \rangle = 3 \cdot 7 + 4 \cdot 0 + 1 \cdot 2 = 23$$

**7.2. Normen****Definition 7.5 (Norm)**

Eine Norm ist eine Abbildung, die jedem Element  $\mathbf{v}$  aus einem Vektorraum  $\mathbf{V}$  eine nicht-negative reelle Zahl zuweist.

$$\|\cdot\| : \mathbf{V} \rightarrow \mathbb{R}_0^+, \quad \mathbf{x} \mapsto \|\mathbf{x}\|$$

Diese muss für alle  $\mathbf{x}, \mathbf{y} \in \mathbf{V}$  und alle  $\lambda \in \mathbb{R}$  folgende Eigenschaften haben:

$$\|\mathbf{x}\| = 0 \Rightarrow \mathbf{x} = \mathbf{0} \quad (\text{sonst } \|\mathbf{x}\| > 0) \quad (7.6)$$

$$\|\lambda \cdot \mathbf{x}\| = |\lambda| \cdot \|\mathbf{x}\| \quad (7.7)$$

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \quad (7.8)$$

**Anmerkung 7.6**

- Aus (7.7) folgt  $\|-\mathbf{u}\| = \|\mathbf{u}\|$
- Eine Norm weist einem Vektor eine Länge zu.

**Definition 7.7**

Mögliche Normen für  $\mathbf{x} \in \mathbb{R}^n$  sind gegeben durch

$$\|\mathbf{x}\|_p := \sqrt[p]{\sum_{i=1}^n |x_i|^p} \quad 1 \leq p < \infty$$

Eine weitere Norm ist die sogenannte Maximumnorm

$$\|\mathbf{x}\|_\infty := \max_{1 \leq i \leq n} |x_i|$$

**Anmerkung 7.8**

- Auf dem  $\mathbb{R}^n$  (bzw. auf endlich dimensionalen Vektorräumen) sind alle Normen äquivalent, d.h. es gibt immer zwei positive Konstanten  $\lambda_1, \lambda_2 \in \mathbb{R}^+$ , so dass für alle  $\mathbf{x} \in \mathbf{V}$

$$\lambda_1 \cdot \|\mathbf{x}\|_a \leq \|\mathbf{x}\|_b \leq \lambda_2 \cdot \|\mathbf{x}\|_a$$

- Je größer  $p$  umso wichtiger werden die betragsgrößten Komponenten.

Es gilt  $\lim_{p \rightarrow \infty} \|\mathbf{x}\|_p = \|\mathbf{x}\|_\infty$ .

### Satz 7.9

Durch  $\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n x_i^2} = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$  ist eine Norm definiert ( $p = 2$  in Definition 7.7). Diese heißt die euklidische Norm oder auch vom Skalarprodukt induzierte Norm.

### Beweis:

Normeigenschaften:

(7.6): folgt aus (7.5)

(7.7):

$$\|\lambda \cdot \mathbf{x}\|_2 = \sqrt{\langle \lambda \mathbf{x}, \lambda \mathbf{x} \rangle} \stackrel{(7.3)}{=} \sqrt{\lambda^2 \cdot \langle \mathbf{x}, \mathbf{x} \rangle} = |\lambda| \cdot \|\mathbf{x}\|_2$$

(7.8):

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|_2^2 &= \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle = \\ &= \langle \mathbf{x}, \mathbf{x} \rangle + 2\langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle \\ &= \|\mathbf{x}\|_2^2 + 2\langle \mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|_2^2 \\ &\leq \|\mathbf{x}\|_2^2 + 2|\langle \mathbf{x}, \mathbf{y} \rangle| + \|\mathbf{y}\|_2^2 \end{aligned} \tag{7.9}$$

Wenn  $\mathbf{x}$  oder  $\mathbf{y} = \mathbf{0}$  ist, dann ist (7.8) erfüllt.

Wir wollen jetzt  $|\langle \mathbf{x}, \mathbf{y} \rangle|$  abschätzen für  $\mathbf{y} \neq \mathbf{0}$ :

Wegen (7.5) gilt für beliebiges  $\lambda \in \mathbb{R}$

$$0 \leq \langle \mathbf{x} - \lambda \mathbf{y}, \mathbf{x} - \lambda \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle - 2\lambda \langle \mathbf{x}, \mathbf{y} \rangle + \lambda^2 \langle \mathbf{y}, \mathbf{y} \rangle$$

Wir wählen jetzt  $\lambda = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle}$

$$\begin{aligned} \Rightarrow 0 &\leq \langle \mathbf{x}, \mathbf{x} \rangle - 2 \cdot \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle} \cdot \langle \mathbf{x}, \mathbf{y} \rangle + \frac{\langle \mathbf{x}, \mathbf{y} \rangle^2}{\langle \mathbf{y}, \mathbf{y} \rangle^2} \cdot \langle \mathbf{y}, \mathbf{y} \rangle \\ &= \langle \mathbf{x}, \mathbf{x} \rangle - 2 \cdot \frac{\langle \mathbf{x}, \mathbf{y} \rangle^2}{\langle \mathbf{y}, \mathbf{y} \rangle} + \frac{\langle \mathbf{x}, \mathbf{y} \rangle^2}{\langle \mathbf{y}, \mathbf{y} \rangle} = \langle \mathbf{x}, \mathbf{x} \rangle - \frac{\langle \mathbf{x}, \mathbf{y} \rangle^2}{\langle \mathbf{y}, \mathbf{y} \rangle} \end{aligned}$$

Da  $\langle \mathbf{y}, \mathbf{y} \rangle > 0$

$$\begin{aligned} \Rightarrow 0 &\leq \langle \mathbf{x}, \mathbf{x} \rangle \cdot \langle \mathbf{y}, \mathbf{y} \rangle - \langle \mathbf{x}, \mathbf{y} \rangle^2 \\ \Rightarrow \langle \mathbf{x}, \mathbf{y} \rangle^2 &\leq \langle \mathbf{x}, \mathbf{x} \rangle \cdot \langle \mathbf{y}, \mathbf{y} \rangle \\ \Rightarrow |\langle \mathbf{x}, \mathbf{y} \rangle|^2 &\leq \|\mathbf{x}\|_2^2 \cdot \|\mathbf{y}\|_2^2 \\ \Rightarrow |\langle \mathbf{x}, \mathbf{y} \rangle| &\leq \|\mathbf{x}\|_2 \cdot \|\mathbf{y}\|_2 \end{aligned} \tag{7.10}$$

Man nennt (7.10) die Cauchy-Schwarzsche Ungleichung.

Wir nehmen (7.9):

$$\|\mathbf{x} + \mathbf{y}\|_2^2 \leq \|\mathbf{x}\|_2^2 + 2|\langle \mathbf{x}, \mathbf{y} \rangle| + \|\mathbf{y}\|_2^2$$

und setzen (7.10) ein:

$$\begin{aligned}\|\mathbf{x} + \mathbf{y}\|_2^2 &\leq \|\mathbf{x}\|_2^2 + 2 \cdot \|\mathbf{x}\|_2 \cdot \|\mathbf{y}\|_2 + \|\mathbf{y}\|_2^2 = (\|\mathbf{x}\|_2 + \|\mathbf{y}\|_2)^2 \\ \Rightarrow \|\mathbf{x} + \mathbf{y}\|_2 &\leq \|\mathbf{x}\|_2 + \|\mathbf{y}\|_2\end{aligned}$$

□

### Beispiel 7.10

$$\mathbf{x} = \begin{pmatrix} 3 \\ -4 \\ 1 \end{pmatrix}$$

$$\|\mathbf{x}\|_1 = |3| + |-4| + |1| = 8$$

$$\|\mathbf{x}\|_2 = \sqrt{3^2 + (-4)^2 + 1^2} = \sqrt{26} \approx 5,1$$

$$\|\mathbf{x}\|_3 = \sqrt[3]{|3|^3 + |-4|^3 + |1|^3} = \sqrt[3]{27 + 64 + 1} = \sqrt[3]{92} \approx 4,5$$

$$\|\mathbf{x}\|_\infty = \max(|3|, |-4|, |1|) = 4$$

#### 3-Minutes Python 9

Bei numpy berechnet sich das Skalarprodukt mit der Funktion dot

```
>>> import numpy as np
2 >>> a=np.array([3,4,1])
>>> b=np.array([7,0,2])
4 >>> np.dot(a,b)
23
```

und bei sympy mit der Methode dot

```
>>> import sympy as sp
2 >>> a=sp.Matrix([3,4,1])
>>> b=sp.Matrix([7,0,2])
4 >>> a.dot(b)
23
```

Die Norm eines Vektors berechnet sich in **sympy** mit der Methode **norm**. Diese berechnet standardmäßig die euklidische Norm. Wird der Methode eine Zahl als Argument übergeben, dann wird die entsprechende Norm berechnet. Ist das Argument **oo** wird die Maximumnorm berechnet.

```
>>> a=sp.Matrix([3,4,1])
2 >>> a.norm(1)
8
4 >>> a.norm()
sqrt(26)
6 >>> a.norm(3)
92**(1/3)
8 >>> a.norm(sp.oo)
```

In `numpy` gibt es zur Normberechnung die Funktion `linalg.norm`. Das erste Argument ist der Vektor, dessen Norm berechnet werden soll, das zweite Argument die Ordnung der Norm  $p$  bzw. `inf`. Wird das Argument weggelassen, dann wird die Euklidische Norm berechnet.

```
>>> a=np.array([3,4,1])
2 >>> np.linalg.norm(a,1)
8.0
4 >>> np.linalg.norm(a)
5.0990195135927845
6 >>> np.linalg.norm(a,3)
4.5143574354740013
8 >>> np.linalg.norm(a,np.inf)
4.0
```

### 7.3. Winkel im $\mathbb{R}^n$ , Orthogonalität

Im Weiteren schreiben wir statt  $\|\cdot\|_2$  einfach nur  $\|\cdot\|$  für die euklidische Norm. Diese hat eine geometrische Interpretation:

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

Umgekehrt können wir dadurch den Winkel  $\varphi$  definieren, falls  $\|\mathbf{a}\| > 0$  und  $\|\mathbf{b}\| > 0$ :

#### Definition 7.11

Seien  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  und  $\mathbf{x}, \mathbf{y} \neq \mathbf{0}$ , dann nennt man  $\varphi := \arccos \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}$  den Winkel zwischen  $\mathbf{x}$  und  $\mathbf{y}$ . Man nennt  $\mathbf{x}$  und  $\mathbf{y}$  orthogonal, falls  $\langle \mathbf{x}, \mathbf{y} \rangle = 0$  (da  $\arccos 0 = 90^\circ$ ). Der Nullvektor ist orthogonal zu allen Vektoren des  $\mathbb{R}^n$ .

#### Beispiel 7.12

1.

$$\begin{aligned} \mathbf{x} &= \begin{pmatrix} 3 \\ -4 \\ 1 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 7 \\ 0 \\ 2 \end{pmatrix} \\ \|\mathbf{x}\| &= \sqrt{26}, \quad \|\mathbf{y}\| = \sqrt{49 + 4} = \sqrt{53} \\ \langle \mathbf{x}, \mathbf{y} \rangle &= 23 \\ \varphi &= \arccos \left( \frac{23}{\sqrt{26 \cdot 53}} \right) = \arccos \left( \sqrt{\frac{529}{1378}} \right) \approx 51,7^\circ \end{aligned}$$

2.

$$\begin{aligned}\mathbf{x} &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \|\mathbf{x}\| &= 1, \quad \|\mathbf{y}\| = 1, \quad \langle \mathbf{x}, \mathbf{y} \rangle = 1 \cdot 0 + 0 \cdot 1 = 0 \\ &\Rightarrow \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|} = 0 \\ &\Rightarrow \mathbf{x} \text{ und } \mathbf{y} \text{ sind orthogonal.}\end{aligned}$$

**Satz 7.13 (Abstand eines Punktes von einer Geraden)**

Seien  $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$  und  $\mathbf{v} \neq \mathbf{0}$ . Sei  $\mathbf{U} = \text{Span}(\mathbf{v})$ . Dann gibt es genau einen eindeutig bestimmten Punkt  $\hat{\mathbf{w}} \in \mathbf{U}$ , der minimalen Abstand zu  $\mathbf{w}$  hat, d.h.

$$\|\mathbf{w} - \hat{\mathbf{w}}\| < \|\mathbf{w} - \mathbf{u}\| \quad \forall \mathbf{u} \in \mathbf{U} \setminus \{\hat{\mathbf{w}}\}. \quad (7.11)$$

Er ist gegeben durch

$$\hat{\mathbf{w}} = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \cdot \mathbf{v} \quad (7.12)$$

Es gilt

$$\langle \mathbf{w} - \hat{\mathbf{w}}, \mathbf{v} \rangle = 0 \quad (7.13)$$

$$\|\hat{\mathbf{w}}\| = \frac{|\langle \mathbf{v}, \mathbf{w} \rangle|}{\|\mathbf{v}\|} \leq \|\mathbf{w}\|, \quad (7.14)$$

und

$$\langle \mathbf{v}, \hat{\mathbf{w}} \rangle = \langle \mathbf{v}, \mathbf{w} \rangle \quad (7.15)$$

**Anmerkung 7.14**

Man kann sich (7.11) durch folgende Umformung anschaulich machen:

$$\hat{\mathbf{w}} = \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \cdot \mathbf{v} = \frac{\|\mathbf{v}\| \cdot \|\mathbf{w}\| \cdot \cos \varphi}{\|\mathbf{v}\|^2} \cdot \mathbf{v} = \|\mathbf{w}\| \cdot \cos \varphi \cdot \frac{\mathbf{v}}{\|\mathbf{v}\|}$$

Es handelt sich also um einen Vektor in Richtung von  $\mathbf{v}$ , dessen Länge dem Anteil von  $\mathbf{w}$  entspricht, der in Richtung der Gerade zeigt.

**Beweis:**

(7.13): folgt aus (7.12): Übung

$$(7.15): \langle \mathbf{v}, \mathbf{w} \rangle = \underbrace{\langle \mathbf{v}, \mathbf{w} - \hat{\mathbf{w}} \rangle}_{=0 \text{ (7.13)}} + \langle \mathbf{v}, \hat{\mathbf{w}} \rangle = \langle \mathbf{v}, \hat{\mathbf{w}} \rangle$$

(7.11):

$$\begin{aligned}\|\mathbf{w} - \mathbf{u}\|^2 &= \|\mathbf{w} - \hat{\mathbf{w}} + \hat{\mathbf{w}} - \mathbf{u}\|^2 \\ &= \langle \mathbf{w} - \hat{\mathbf{w}} + \hat{\mathbf{w}} - \mathbf{u}, \mathbf{w} - \hat{\mathbf{w}} + \hat{\mathbf{w}} - \mathbf{u} \rangle \\ &= \langle \mathbf{w} - \hat{\mathbf{w}}, \mathbf{w} - \hat{\mathbf{w}} \rangle + 2\langle \mathbf{w} - \hat{\mathbf{w}}, \hat{\mathbf{w}} - \mathbf{u} \rangle + \langle \hat{\mathbf{w}} - \mathbf{u}, \hat{\mathbf{w}} - \mathbf{u} \rangle \\ &= \|\mathbf{w} - \hat{\mathbf{w}}\|^2 + 2\langle \mathbf{w} - \hat{\mathbf{w}}, \hat{\mathbf{w}} - \mathbf{u} \rangle + \|\hat{\mathbf{w}} - \mathbf{u}\|^2\end{aligned}$$

da  $\hat{\mathbf{w}} - \mathbf{u} \in \mathbf{U}$ ,  $\mathbf{U} = \text{Span}(\mathbf{v})$  und  $\langle \mathbf{w} - \hat{\mathbf{w}}, \mathbf{v} \rangle \stackrel{(7.13)}{=} 0$   
 $\Rightarrow \|\mathbf{w} - \mathbf{u}\|^2 = \|\mathbf{w} - \hat{\mathbf{w}}\|^2 + \|\hat{\mathbf{w}} - \mathbf{u}\|^2 > \|\mathbf{w} - \hat{\mathbf{w}}\|^2 \quad \forall \mathbf{u} \in \mathbf{U} \setminus \{\hat{\mathbf{w}}\}$

$$(7.14): (7.12) \Rightarrow \|\hat{\mathbf{w}}\| = \left\| \frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \cdot \mathbf{v} \right\| = \frac{|\langle \mathbf{v}, \mathbf{w} \rangle|}{\|\mathbf{v}\|^2} \cdot \|\mathbf{v}\|$$

$$\Rightarrow \|\hat{\mathbf{w}}\| = \frac{|\langle \mathbf{v}, \mathbf{w} \rangle|}{\|\mathbf{v}\|}$$

$$\|\mathbf{w}\|^2 = \|\mathbf{w} - \hat{\mathbf{w}} + \hat{\mathbf{w}}\|^2 = \|\mathbf{w} - \hat{\mathbf{w}}\|^2 + 2 \underbrace{\langle \mathbf{w} - \hat{\mathbf{w}}, \hat{\mathbf{w}} \rangle}_{=0 \text{ 7.13}} + \|\hat{\mathbf{w}}\|^2$$

$$= \|\mathbf{w} - \hat{\mathbf{w}}\|^2 + \|\hat{\mathbf{w}}\|^2 \geq \|\hat{\mathbf{w}}\|^2$$

$$\Rightarrow \|\mathbf{w}\|^2 \geq \|\hat{\mathbf{w}}\|^2 \Rightarrow \|\mathbf{w}\| \geq \|\hat{\mathbf{w}}\|$$

□

### Anmerkung 7.15

Aus (7.14) ergibt sich direkt die Cauchy-Schwarzsche Ungleichung:

$$\|\mathbf{w}\| \geq \frac{|\langle \mathbf{v}, \mathbf{w} \rangle|}{\|\mathbf{v}\|} \iff \|\mathbf{v}\| \cdot \|\mathbf{w}\| \geq |\langle \mathbf{v}, \mathbf{w} \rangle|$$

### Definition 7.16

Die Vektoren  $\mathbf{u}_1, \dots, \mathbf{u}_p \in \mathbb{R}^n$  bilden ein Orthogonalsystem, falls die  $\mathbf{u}_i$  paarweise orthogonal zueinander sind.

Sind die  $\mathbf{u}_i$  normiert, d.h.  $\|\mathbf{u}_i\| = 1 \quad \forall 1 \leq i \leq p$ , bilden sie ein Orthonormalsystem. Eine Orthonormalbasis eines Vektorraums ist eine Basis, die ein Orthonormalsystem bildet.

### Satz 7.17 (Näherung eines Punktes in einem Unterraum)

Sei  $\mathbf{v}_1, \dots, \mathbf{v}_p \in \mathbb{R}^n$  ein Orthogonalsystem und  $\mathbf{v}_i \neq \mathbf{0} \quad \forall 1 \leq i \leq p$

Dann gilt:

1. Für jedes  $\mathbf{w} = \sum_{i=1}^p \alpha_i \mathbf{v}_i$   
gilt für  $i = 1, \dots, p$ :

$$\alpha_i = \frac{\langle \mathbf{v}_i, \mathbf{w} \rangle}{\langle \mathbf{v}_i, \mathbf{v}_i \rangle} \tag{7.16}$$

2. Die  $\mathbf{v}_i$  sind linear unabhängig.

3. Sei  $\mathbf{w} \in \mathbb{R}^n$  beliebig. Durch

$$\beta_i := \frac{\langle \mathbf{v}_i, \mathbf{w} \rangle}{\langle \mathbf{v}_i, \mathbf{v}_i \rangle} \quad \forall 1 \leq i \leq p$$

$$\hat{\mathbf{w}} := \sum_{i=1}^p \beta_i \mathbf{v}_i \tag{7.17}$$

ist der zu  $\mathbf{w}$  nächste Punkt aus  $\text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_p)$  eindeutig definiert, d.h.

$$\mathbf{w} - \hat{\mathbf{w}} \perp \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_p) \tag{7.18}$$

$$\|\mathbf{w} - \hat{\mathbf{w}}\| < \|\mathbf{w} - \mathbf{u}\| \quad \forall \mathbf{u} \in \text{Span}(\mathbf{v}_1, \dots, \mathbf{v}_p) \setminus \{\hat{\mathbf{w}}\} \tag{7.19}$$

**Beweis:**

1.

$$\begin{aligned}\mathbf{w} &= \sum_{i=1}^p \alpha_i \cdot \mathbf{v}_i \\ \langle \mathbf{v}_i, \mathbf{w} \rangle &= \sum_{j=1}^p \alpha_j \cdot \langle \mathbf{v}_i, \mathbf{v}_j \rangle = \alpha_i \langle \mathbf{v}_i, \mathbf{v}_i \rangle \\ \Rightarrow \alpha_i &= \frac{\langle \mathbf{v}_i, \mathbf{w} \rangle}{\langle \mathbf{v}_i, \mathbf{v}_i \rangle}\end{aligned}$$

2.

$$\begin{aligned}\mathbf{0} &= \sum_{i=1}^p \alpha_i \cdot \mathbf{v}_i \\ \Rightarrow \alpha_i &= \frac{\langle \mathbf{v}_i, \mathbf{0} \rangle}{\langle \mathbf{v}_i, \mathbf{v}_i \rangle} = \mathbf{0}\end{aligned}$$

$\Rightarrow \mathbf{v}_i$  sind linear unabhängig

3. Wegen (7.13) steht  $\mathbf{w} - \hat{\mathbf{w}}$  senkrecht auf jedem einzelnen Vektor  $\mathbf{v}_i$   
Da (7.11) für jeden Vektor  $\mathbf{v}_i$  gilt  $\Rightarrow$ (7.19)

□

### Anmerkung 7.18

- Aus einem Orthogonalsystem nach Satz 7.17 lässt sich durch Normierung

$$\mathbf{u}_i = \frac{\mathbf{v}_i}{\|\mathbf{v}_i\|}$$

eine Orthonormalbasis generieren

- Die Einheitsvektoren  $\mathbf{e}_i$  ( $1 \leq i \leq n$ ) bilden eine Orthonormalbasis des  $\mathbb{R}^n$

### Definition 7.19

Ein reeller Vektorraum  $\mathbf{V}$  mit Skalarprodukt, in dem die Länge von  $\mathbf{x} \in \mathbf{V}$  definiert ist durch  $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$ , der Abstand zwischen zwei Vektoren  $\mathbf{x}, \mathbf{y} \in \mathbf{V}$  durch  $\|\mathbf{x} - \mathbf{y}\|$  und der Winkel zwischen zwei Vektoren  $\mathbf{x}, \mathbf{y} \in \mathbf{V}$  durch

$$\varphi := \arccos \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}, \quad \mathbf{x} \perp \mathbf{y} \iff \langle \mathbf{x}, \mathbf{y} \rangle = 0$$

heißt euklidischer Vektorraum.

### Satz 7.20

Jeder  $n$ -dimensionale euklidische Vektorraum  $\mathbf{V}$  kann durch eine Orthonormalbasis  $\mathbf{v}_1, \dots, \mathbf{v}_n$  aufgespannt werden. Die Koeffizienten  $\alpha_i$  bezüglich dieser Basis ergeben sich als

$$\alpha_i = \langle \mathbf{x}, \mathbf{v}_i \rangle$$



**Beweis:**

Satz 7.17 + Normiertheit der Basis

□

**Anmerkung 7.21**

Ein kartesisches Koordinatensystem in einem euklidischen Vektorraum ist gegeben durch die Wahl einer bestimmten Orthonormalbasis und eines Koordinatenursprungs.

**Algorithmus 7.22 (Gram-Schmidtsches-Orthonormalisierungsverfahren)**

$\mathbf{v}_1, \dots, \mathbf{v}_p \in \mathbb{R}^n$  seien beliebige linear unabhängige Vektoren

1. Setze  $\mathbf{u}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|}$  und  $k = 1$ .
2. Angenommen  $\mathbf{u}_1, \dots, \mathbf{u}_k$  seien bereits orthonormale Vektoren  
 Berechne  $\hat{\mathbf{u}}_{k+1} = \sum_{i=1}^k \langle \mathbf{u}_i, \mathbf{v}_{k+1} \rangle \cdot \mathbf{u}_i$
3. Setze  $\mathbf{u}_{k+1} = \frac{\mathbf{v}_{k+1} - \hat{\mathbf{u}}_{k+1}}{\|\mathbf{v}_{k+1} - \hat{\mathbf{u}}_{k+1}\|}$
4. Wenn  $(k+1) < p$  wiederhole ab 2. mit  $k = k+1$ .
5.  $\mathbf{u}_1, \dots, \mathbf{u}_p$  bilden jetzt ein Orthonormalsystem

**Beispiel 7.23**

$$\begin{aligned} \mathbf{v}_1 &= \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \\ \|\mathbf{v}_1\| &= \sqrt{4} = 2 \\ \mathbf{u}_1 &= \frac{1}{2} \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ \hat{\mathbf{u}}_2 &= \left\langle \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix} \right\rangle \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 3 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix} \\ \mathbf{v}_2 - \hat{\mathbf{u}}_2 &= \begin{pmatrix} 3 \\ 2 \end{pmatrix} - \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} \\ \|\mathbf{v}_2 - \hat{\mathbf{u}}_2\| &= 2 \\ \mathbf{u}_2 &= \frac{1}{2} \cdot \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \Rightarrow \mathbf{u}_1 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{aligned}$$

**3-Minutes Python 10**

`sympy` stellt die Funktion `Gramschmidt` zur Gram-Schmidt Orthogonalisierung und Orthonormierung bereit. Dazu muss eine Liste von Vektoren übergeben werden. Wird als zweites Argument `True` übergeben, dann werden die orthogonalisierten Vektoren auch noch normiert.

```
>>> L = [sp.Matrix([2,3,5]), sp.Matrix([3,6,2]),
```

```

    sp.Matrix([8,3,6]])
2 >>> sp.GramSchmidt(L)
  [Matrix([
4 [2],
  [3],
6 [5]])], Matrix([
  [ 23/19],
8 [ 63/19],
  [-47/19]])], Matrix([
10 [ 1692/353],
  [-1551/706],
12 [ -423/706]])]
>>> W=sp.GramSchmidt(L)
14 >>> W[0].dot(W[0])
  38
16 >>> W[0].dot(W[1])
  0
18 >>> sp.GramSchmidt(L,True)
  [Matrix([
20 [ sqrt(38)/19],
  [3*sqrt(38)/38],
22 [5*sqrt(38)/38]])], Matrix([
  [ 23*sqrt(6707)/6707],
24 [ 63*sqrt(6707)/6707],
  [-47*sqrt(6707)/6707]])], Matrix([
26 [ 12*sqrt(706)/353],
  [-11*sqrt(706)/706],
28 [ -3*sqrt(706)/706]])]
>>> U=sp.GramSchmidt(L)
30 >>> U[0].dot(U[1])
  0
32 >>> U[0].dot(U[0])
  1

```

## 7.4. Methode der kleinsten Quadrate

Bestimmung von Systemparametern durch indirekte Messungen z.B. des Rollwiderstandskoeffizienten  $c_r$

### Beispiel 7.24

Rollreibung:  $F_R = c_r \cdot F_N$  ( $F_N$  : Normalenkomponente der Gewichtskraft)

Wir messen  $x(t)$  bei einer Anfangsgeschwindigkeit  $v_0$  für einen Körper der Masse  $m$ . Wegen des zweiten Newton'schen Gesetzes gilt:

$$m \cdot a = -F_R = -c_r \cdot F_N$$

$$a = -\frac{1}{m} (c_r \cdot F_N)$$

Befindet sich das Gefährt bei  $t_0 = 0$  an der Stelle  $x = 0$  mit Geschwindigkeit  $v = v_0$

$$\Rightarrow x = \frac{1}{2}at^2 + v_0 \cdot t = -\frac{1}{2m}c_r \cdot F_N \cdot t^2 + v_0 \cdot t$$

Dies ist eine lineare Gleichung in  $c_r$  und  $v_0$

Problem: Messwerte sind fehlerbehaftet

Lösung: Wir machen sehr viele Messungen

$$\begin{array}{cccccc} t & 2 & 4 & 6 & 8 & 10 \\ x & 60 & 120 & 175 & 235 & 290 \end{array}$$

$$m = 1600 \text{ kg}, \quad F_N = 16000 \text{ N}$$

$$\frac{F_N}{2m} = 5$$

$$x = -c_r \cdot 5 \cdot t^2 + v_0 \cdot t$$

$$\begin{pmatrix} -20 & 2 \\ -80 & 4 \\ -180 & 6 \\ -320 & 8 \\ -500 & 10 \end{pmatrix} \cdot \begin{pmatrix} c_r \\ v_0 \end{pmatrix} = \begin{pmatrix} 60 \\ 120 \\ 175 \\ 235 \\ 290 \end{pmatrix}$$

$\Rightarrow$  überbestimmtes Problem, wahrscheinlich keine exakte Lösung.

Wir suchen die Lösung für  $\mathbf{u} = (c_r, v_0)^T \in \mathbb{R}^2$ , für die  $\hat{\mathbf{x}} = \mathbf{A} \cdot \mathbf{u}$  den minimalen Abstand von  $\mathbf{x} = (60 \ 120 \ 175 \ 235 \ 290)^T$  hat.

Die Spalten von  $\mathbf{A}$  spannen einen zweidimensionalen Unterraum des  $\mathbb{R}^5$  auf (da sie linear unabhängig sind). Wir suchen also den Punkt in diesem Unterraum mit minimalem Abstand von  $\mathbf{x}$ . Für diesen gilt  $(\mathbf{x} - \hat{\mathbf{x}}) \perp$  für jede Spalten  $\mathbf{a}^{(i)}$  von  $\mathbf{A}$ :

$$\mathbf{a}^{(i)} \perp (\mathbf{x} - \hat{\mathbf{x}}) \iff \langle \mathbf{a}^{(i)}, (\mathbf{x} - \hat{\mathbf{x}}) \rangle = \mathbf{a}^{(i)T} \cdot (\mathbf{x} - \hat{\mathbf{x}}) = 0$$

$$\text{Gilt für alle Spalten} \Rightarrow \mathbf{A}^T \cdot (\mathbf{x} - \hat{\mathbf{x}}) = \mathbf{0} \iff \mathbf{A}^T \cdot (\mathbf{x} - \mathbf{A} \cdot \mathbf{u}) = \mathbf{0}$$

### Satz 7.25

Sei  $\mathbf{A} \in \mathbb{R}^{m \times n}$  und  $\mathbf{b} \in \mathbb{R}^m$ .  $\mathbf{u} \in \mathbb{R}^n$  ist genau dann eine "kleinste Quadrate-Lösung" der Gleichung  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$  (oder der Wert für  $\mathbf{x}$ , für den  $\mathbf{A} \cdot \mathbf{x}$  von  $\mathbf{b}$  den kleinsten Abstand in der euklidischen Norm hat), wenn  $\mathbf{u}$  eine Lösung der sog. Normalengleichung

$$\mathbf{A}^T \mathbf{A} \mathbf{u} = \mathbf{A}^T \mathbf{b}$$

ist.

### Anmerkung 7.26

Sind die Fehler der Messung normalverteilt, erhält man eine optimale Schätzung der Parameter.

### Beispiel 7.27 (Fortsetzung von Beispiel 7.24)

Mit

$$\mathbf{A}^T \mathbf{A} = \begin{pmatrix} -20 & -80 & -180 & -320 & -500 \\ 2 & 4 & 6 & 8 & 10 \end{pmatrix} \cdot \begin{pmatrix} -20 & 2 \\ -80 & 4 \\ -180 & 6 \\ -320 & 8 \\ -500 & 10 \end{pmatrix} = \begin{pmatrix} 391600 & -9000 \\ -9000 & 220 \end{pmatrix}$$

und

$$\mathbf{A}^T \mathbf{x} = \begin{pmatrix} -20 & -80 & -180 & -320 & -500 \\ 2 & 4 & 6 & 8 & 10 \end{pmatrix} \cdot \begin{pmatrix} 60 \\ 120 \\ 175 \\ 235 \\ 290 \end{pmatrix} = \begin{pmatrix} -262500 \\ 6430 \end{pmatrix}$$

erhalten wir

$$\begin{pmatrix} 391600 & -9000 \\ -9000 & 220 \end{pmatrix} \cdot \mathbf{u} = \begin{pmatrix} -262500 \\ 6430 \end{pmatrix}$$

$$\begin{pmatrix} 391600 & -9000 \\ 0 & 13,156 \end{pmatrix} \cdot \mathbf{u} = \begin{pmatrix} -262500 \\ 397,058 \end{pmatrix}$$

$$\Rightarrow \mathbf{u} = \begin{pmatrix} 0,023 \\ 30,18 \end{pmatrix}$$

$$\Rightarrow \hat{\mathbf{x}} = \mathbf{A} \cdot \mathbf{u} = \begin{pmatrix} 59,9 \\ 118,9 \\ 176,9 \\ 234,1 \\ 290,3 \end{pmatrix}$$

$$\|\mathbf{x} - \hat{\mathbf{x}}\| = 2,1$$

Zum Vergleich: die exakten Werte waren:  $c_r = 0.015$ ,  $v_0 = 30$  m/s, die Fehler waren nicht normalverteilt.

#### 3-Minutes Python 11

Die kleinste Quadrate-Lösung kann mit Hilfe der Funktion `linalg.lstsq` von `numpy` berechnet werden.

```
>>> import numpy as np
2 >>> A=np.array([[ -20,2], [ -80,4], [ -180,6], [ -320,8], [ -500,10]])
>>> b=np.array([60,120,175,235,290])
4 >>> np.linalg.lstsq(A,b)
(array([ 2.32919255e-02,  3.01801242e+01]), array([
    5.93167702]), 2, array([ 625.94476646,  3.62620134]))
```

Die Funktion liefert vier Werte zurück: Einen Vektor mit der Kleinste-Quadrate-Lösung, das Residuum (also die Summe der quadrierten Abweichungen zwischen Messwerten und geschätzten Werten), den Rang von  $\mathbf{A}$  und die sogenannten Singulärwerte von  $\mathbf{A}$ .

## 7.5. Vektorprodukt

### Definition 7.28

Seien  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$  und  $\mathbf{u} = (u_1, u_2, u_3)^T$ ,  $\mathbf{v} = (v_1, v_2, v_3)^T$  ihre Darstellung in kartesischen Koordinaten. Dann ist das Kreuzprodukt von  $\mathbf{u}$  und  $\mathbf{v}$  definiert als

$$\mathbf{u} \times \mathbf{v} := \begin{pmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{pmatrix}$$

### Satz 7.29

Das Kreuzprodukt von  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$  hat folgende Eigenschaften:

1.  $\mathbf{u} \times \mathbf{v} = \mathbf{0}$ , wenn  $\mathbf{u}$  und  $\mathbf{v}$  linear abhängig sind.
2. Seien  $\mathbf{u}$  und  $\mathbf{v}$  linear unabhängig. Dann gilt:
  - a)  $\mathbf{u} \times \mathbf{v} \perp \text{Span}(\mathbf{u}, \mathbf{v})$
  - b)  $\mathbf{u}$ ,  $\mathbf{v}$ , und  $\mathbf{u} \times \mathbf{v}$  sind positiv orientiert.
  - c)  $\|\mathbf{u} \times \mathbf{v}\|$  ist der Flächeninhalt des Parallelogramms mit den Ecken  $\mathbf{0}$ ,  $\mathbf{u}$ ,  $\mathbf{v}$  und  $\mathbf{u} + \mathbf{v}$ .

**Beweis:**

Übungen

□

### Anmerkung 7.30

Geometrisch bedeutet

- 7.29 2a: Der Vektor  $\mathbf{n} = \mathbf{u} \times \mathbf{v}$  steht senkrecht auf der Ebene, die von  $\mathbf{u}$  und  $\mathbf{v}$  aufgespannt wird (Normalenvektor).
- 7.29 2b:  $\mathbf{u}$ ,  $\mathbf{v}$  und  $\mathbf{n}$  zeigen wie Daumen, Zeigefinger und Mittelfinger der rechten Hand.

### 3-Minutes Python 12

Das Vektorprodukt berechnet sich in **numpy** mit der Funktion **cross** und in **sympy** mit der Methode **cross**:

```
>>> import numpy as np
2 >>> x = np.array([0,0,1])
>>> y = np.array([0,1,0])
4 >>>
>>> np.cross(x,y)
6 array([-1,  0,  0])
>>> np.cross(y,x)
8 array([1,  0,  0])
>>> a = sp.Matrix([0,1,0])
10 >>> b = sp.Matrix([1,0,0])
```

```

>>> a.cross(b)
12 Matrix([
    [ 0],
14    [ 0],
    [-1]])
16 >>> b.cross(a)
    Matrix([
18    [0],
    [0],
20    [1]])

```

## 7.6. Tensorprodukt

### Definition 7.31

Seien  $\mathbf{u} \in \mathbb{R}^n$  und  $\mathbf{v} \in \mathbb{R}^m$ . Dann ist das Tensorprodukt oder das dyadische Produkt (im Englischen *outer product*) von  $\mathbf{u}$  und  $\mathbf{v}$  definiert als

$$\mathbf{u} \otimes \mathbf{v} := \mathbf{u} \cdot \mathbf{v}^T = \begin{pmatrix} u_1 v_1 & u_1 v_2 & \dots & u_1 v_m \\ u_2 v_1 & u_2 v_2 & \dots & u_2 v_m \\ \vdots & \vdots & \ddots & \vdots \\ u_n v_1 & u_n v_2 & \dots & u_n v_m \end{pmatrix}$$

### Satz 7.32

1. Seien  $\mathbf{u} \in \mathbb{R}^n$  und  $\mathbf{v} \in \mathbb{R}^m$  mit  $\mathbf{u} \neq \mathbf{0}$  und  $\mathbf{v} \neq \mathbf{0}$ . Dann gilt:

$$\text{Rang}(\mathbf{u} \otimes \mathbf{v}) = 1$$

2. Seien  $\mathbf{u}, \mathbf{v}, \mathbf{x} \in \mathbb{R}^n$ . Dann gilt:

$$(\mathbf{u} \otimes \mathbf{v}) \cdot \mathbf{x} = \mathbf{u} \cdot \langle \mathbf{v}, \mathbf{x} \rangle$$

### Beweis:

1. Jede Zeile enthält jeweils die Elemente des Vektors  $\mathbf{v}$  multipliziert mit einem Element von  $\mathbf{u}$ . Die i-te Zeile des Produkts lässt sich also durch die j-te Zeile des Produkts durch Multiplikation mit  $\frac{u_i}{u_j}$  darstellen.
2. Es gilt:

$$(\mathbf{u} \otimes \mathbf{v}) \cdot \mathbf{x} = (\mathbf{u} \cdot \mathbf{v}^T) \cdot \mathbf{x} = \mathbf{u} \cdot (\mathbf{v}^T \cdot \mathbf{x}) = \mathbf{u} \cdot \langle \mathbf{v}, \mathbf{x} \rangle$$

□

## 8. Lineare Abbildungen (Bonusmaterial)

### Definition 8.1

Seien  $\mathbf{V}$  und  $\mathbf{W}$  Vektorräume über  $\mathbb{R}$ . Durch eine Abbildung  $f : \mathbf{V} \rightarrow \mathbf{W}$  (man sagt “eine Abbildung  $f$  von  $\mathbf{V}$  nach  $\mathbf{W}$ ”) wird jedem Vektor  $\mathbf{v} \in \mathbf{V}$  ein eindeutig bestimmter Vektor  $\mathbf{w} = f(\mathbf{v})$  zugeordnet.

Man nennt eine Abbildung

**injektiv** wenn es zu jedem  $\mathbf{w} \in \mathbf{W}$  höchstens ein  $\mathbf{v} \in \mathbf{V}$  gibt.

**surjektiv** wenn es zu jedem  $\mathbf{w} \in \mathbf{W}$  mindestens ein  $\mathbf{v} \in \mathbf{V}$  gibt.

**bijektiv** wenn die Abbildung injektiv und surjektiv ist, es also zu jedem  $\mathbf{w} \in \mathbf{W}$  genau ein  $\mathbf{v} \in \mathbf{V}$  gibt.

### Beispiel 8.2

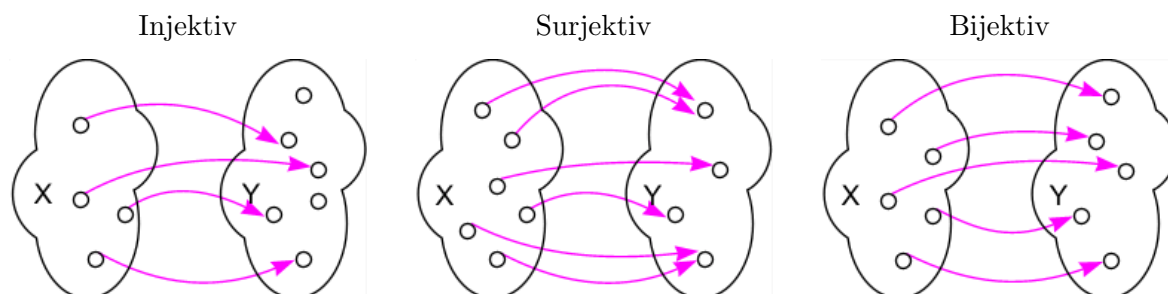
- injektiv

- $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x)$  streng monoton steigend oder fallend, z.B.  $x^3$  aber nicht  $x^2$
- $f : \mathbb{R}_0^+ \rightarrow \mathbb{R}$ ,  $f(x) = \sqrt{x}$

- surjektiv

- $f : \mathbb{R} \rightarrow \mathbb{R}$  stetig, Maxima und Minima erlaubt aber  $\lim_{x \rightarrow \pm\infty} f(x) = \pm\infty$  oder  $\lim_{x \rightarrow \pm\infty} f(x) = \mp\infty$ , z.B.  $ax^3 + bx^2 + cx + d$  aber nicht  $ax^2 + bx + c$ ,  $a \neq 0$ .
- $f : \mathbb{R} \rightarrow [-1 : 1]$ ,  $f(x) = \cos(x)$

Mengen:



### Definition 8.3

Eine Abbildung  $f : \mathbf{V} \rightarrow \mathbf{W}$  heißt linear, wenn gilt:

1.  $f$  ist homogen, d.h.  $f(\lambda \mathbf{v}) = \lambda f(\mathbf{v}) \quad \forall \lambda \in \mathbb{R}, \mathbf{v} \in \mathbf{V}$
2.  $f$  ist additiv, d.h.  $f(\mathbf{v} + \mathbf{u}) = f(\mathbf{v}) + f(\mathbf{u}) \quad \forall \mathbf{v}, \mathbf{u} \in \mathbf{V}$

### Anmerkung 8.4

Definition 8.3 ist äquivalent zu

$$f(\alpha_1 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n) = \alpha_1 f(\mathbf{v}_1) + \cdots + \alpha_n f(\mathbf{v}_n) \quad (8.1)$$

Damit reicht es zu zeigen, dass  $f(\lambda \cdot \mathbf{v} + \mathbf{u}) = \lambda \cdot f(\mathbf{v}) + f(\mathbf{u})$  für alle  $\lambda \in \mathbb{R}$  und  $\mathbf{v}, \mathbf{u} \in \mathbf{V}$  um zu beweisen, dass eine Abbildung linear ist.

**Beispiel 8.5**

1. Die Nullabbildung  $0 : \mathbf{V} \rightarrow \mathbf{V}$ ,  $0(\mathbf{v}) := \mathbf{0}$
2. Die identische Abbildung  $\text{Id} : \mathbf{V} \rightarrow \mathbf{V}$ ,  $\text{Id}(\mathbf{v}) := \mathbf{v}$  (bijektiv)
3. Die Projektion  $p_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $p_i(\mathbf{x}) := x_i$  ( $i$ -te Komponente von  $\mathbf{x}$ ) (surjektiv)
4. Multiplikation mit Matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$

$$f_A : \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad f_A(\mathbf{x}) := \mathbf{A} \cdot \mathbf{x}$$

5. Differentiationsoperator  $\frac{d}{dx}$  an einer festen Stelle  $x$ :

$$\mathcal{C}^1(\mathbf{I}) \rightarrow \mathcal{C}^0(\mathbf{I}), \quad \frac{d}{dx}f(x) := f'(x)$$

6. keine linearen Abbildungen sind:

$$g : \mathbf{V} \rightarrow \mathbf{V}, \quad g(\mathbf{v}) := \mathbf{v} + \mathbf{a} \quad (\text{Verschiebung um } \mathbf{a} \neq \mathbf{0})$$

$$f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} \times \mathbb{R}, \quad f(x, y) := (x^2, x + y) \quad (\text{da quadratisch in } x)$$

7. Skalarprodukt: Bilineare Abbildung, d.h. das Skalarprodukt ist eine lineare Abbildung für jedes einzelne Argument.

**Satz 8.6**

Sei  $f : \mathbf{V} \rightarrow \mathbf{W}$  eine lineare Abbildung, dann gilt:

1.  $f(\mathbf{0}) = \mathbf{0}$
2.  $f$  ist genau dann injektiv, wenn  $f(\mathbf{x}) \neq \mathbf{0} \quad \forall \mathbf{x} \neq \mathbf{0}$
3.  $\mathbf{U} := \{f(\mathbf{x}) : \mathbf{x} \in \mathbf{V}\}$  ist ein Unterraum von  $\mathbf{W}$
4. Sei  $f$  bijektiv, dann ist  $f$  invertierbar und die Umkehrabbildung  $f^{-1} : \mathbf{W} \rightarrow \mathbf{V}$  ist ebenfalls linear. Ist  $f$  injektiv, dann gibt es auch eine Umkehrabbildung, diese ist jedoch nur für die Bildmenge  $\mathbf{U}$  definiert.

ohne Beweis

**Satz 8.7**

Seien  $g : \mathbf{U} \rightarrow \mathbf{V}$  und  $f : \mathbf{V} \rightarrow \mathbf{W}$  linear. Dann ist auch  $f \circ g : \mathbf{U} \rightarrow \mathbf{W}$  linear.

$(f \circ g)(x)$  (sprich “ $f$  nach  $g$ ”) entspricht  $f(g(x))$

**Beweis:** Seien  $\mathbf{u}, \mathbf{v} \in \mathbf{U}$  und  $\lambda \in \mathbb{R}$

$$\begin{aligned} f \circ g(\mathbf{u} + \mathbf{v}) &= f(g(\mathbf{u} + \mathbf{v})) = f(g(\mathbf{u}) + g(\mathbf{v})) = f(g(\mathbf{u})) + f(g(\mathbf{v})) \\ &= f \circ g(\mathbf{u}) + f \circ g(\mathbf{v}) \\ f \circ g(\lambda \mathbf{v}) &= f(g(\lambda \mathbf{v})) = f(\lambda \cdot g(\mathbf{v})) = \lambda \cdot f(g(\mathbf{v})) = \lambda \cdot f \circ g(\mathbf{v}) \end{aligned}$$

□



**Beispiel 8.8**

Sei  $\mathbf{B} = (\mathbf{u}_1, \dots, \mathbf{u}_n)$  mit  $\mathbf{u}_i \in \mathbb{R}^n$  ( $1 \leq i \leq n$ ) eine Basis des Vektorraums  $\mathbf{V}$  und

$$\mathbf{v} = \alpha_1 \mathbf{u}_1 + \dots + \alpha_n \mathbf{u}_n$$

die eindeutige Darstellung des (beliebigen) Vektors  $\mathbf{v} \in \mathbf{V}$  in der Basis  $\mathbf{B}$ . Wir nennen

$$\mathbf{v}_{\mathbf{B}} := \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{R}^n$$

den Koordinatenvektor von  $\mathbf{v} \in \mathbf{V}$  bezüglich  $\mathbf{B}$ .

Die lineare Abbildung  $\mathbf{v} \rightarrow \mathbf{v}_{\mathbf{B}}$  heißt Koordinatenabbildung. Sie ist bijektiv.

**Satz 8.9**

Sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$  linear. Dann gibt es genau eine Matrix  $\mathbf{A}$ , sodass  $f = f_{\mathbf{A}}$ . Es gilt

$$\mathbf{A} = [f(\mathbf{e}_1) \quad f(\mathbf{e}_2) \quad \dots \quad f(\mathbf{e}_n)]$$

wobei  $\mathbf{e}_i \in \mathbb{R}^n$  die  $i$ -ten Einheitsvektoren sind.

Ganz wichtige Schlussfolgerung:

**Jede lineare Abbildung zwischen endlich-dimensionalen Vektorräumen lässt sich also als Matrixmultiplikation darstellen.**

**Beweis:** Mit  $\mathbf{x} = \sum_{i=1}^n x_i \cdot \mathbf{e}_i$  und Gleichung 8.1 folgt:

$$f(\mathbf{x}) = \sum_{i=1}^n x_i \cdot f(\mathbf{e}_i)$$

Wegen der Eigenschaften der Matrix-Vektormultiplikation folgt die Behauptung. □

**Satz 8.10**

Seien  $\mathbf{F}$  und  $\mathbf{G}$  die Abbildungsmatrizen zu den linearen Abbildungen  $f$  und  $g$ . Dann ist die Abbildungsmatrix zu  $h := f \circ g$  gegeben durch

$$\mathbf{H} = \mathbf{F} \cdot \mathbf{G}$$

**Beweis:**

$$f \circ g(x) = f(g(x)) = \mathbf{F} \cdot g(x) = \mathbf{F} \cdot \mathbf{G} \cdot \mathbf{x} = (\mathbf{F} \cdot \mathbf{G}) \cdot \mathbf{x}$$

□

**Beispiel 8.11**

Sei  $\alpha \in \mathbb{R}$ ,  $\mathbf{x} \in \mathbb{R}^2$ . Die Abbildung  $f_{\alpha} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  drehe den Vektor  $\mathbf{x}$  um den Winkel  $|\alpha|$

- gegen den Uhrzeigersinn, falls  $\alpha \geq 0$
- im Uhrzeigersinn, falls  $\alpha < 0$ .

$f_\alpha$  ist linear und die Drehmatrix  $A_\alpha$  ist gegeben durch:

$$A_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

Drehe erst um Winkel  $\alpha$ , dann um Winkel  $\beta$ , also insgesamt um Winkel  $\alpha + \beta$ :

$$f_\alpha \circ f_\beta = f_{\alpha+\beta}$$

$$\Rightarrow \mathbf{A}_{\alpha+\beta} = \mathbf{A}_\alpha \cdot \mathbf{A}_\beta$$

$$\begin{aligned} A_\alpha \cdot A_\beta &= \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \cdot \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix} \\ &= \begin{pmatrix} \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) & -\cos(\alpha)\sin(\beta) - \sin(\alpha)\cos(\beta) \\ \sin(\alpha)\cos(\beta) + \cos(\alpha)\sin(\beta) & -\sin(\alpha)\sin(\beta) + \cos(\alpha)\cos(\beta) \end{pmatrix} \\ &\stackrel{!}{=} \begin{pmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{pmatrix} \end{aligned}$$

$$\Rightarrow \cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$$

$$\sin(\alpha + \beta) = \sin(\alpha)\cos(\beta) + \cos(\beta)\sin(\alpha)$$

Wir erhalten auf diesem Weg also die Additionstheoreme für Sinus und Kosinus.

## 9. Eigenwerte und Eigenvektoren

### 9.1. Definition

#### Definition 9.1

Eine Zahl  $\lambda \in \mathbb{C}$  heißt Eigenwert der Matrix  $\mathbf{A} \in \mathbb{C}^{n \times n}$ , wenn es wenigstens einen Vektor  $\mathbf{x} \in \mathbb{C}^n$ ,  $\mathbf{x} \neq \mathbf{0}$  gibt mit

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

Jeder Vektor  $\mathbf{x} \neq \mathbf{0}$ , der diese Gleichung erfüllt, heißt Eigenvektor von  $\mathbf{A}$  zum Eigenwert  $\lambda$ .

#### Anmerkung 9.2

- Da Eigenwerte (EW) und Eigenvektoren (EV) häufig eine Rolle in Anwendungen spielen, die mit komplexen Zahlen rechnen, verwenden wir ab jetzt Matrizen und Vektoren mit komplexen Einträgen.

Alles bisher über reelle Matrizen und Vektoren Gelernte lässt sich direkt auf komplexe Matrizen und Vektoren übertragen.

#### Beispiel 9.3

Die Matrix

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

hat den Eigenwert  $\lambda = -2$  mit dem Eigenvektor  $\mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$ , da

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

- Geometrische Bedeutung:  
Ein Eigenvektor wird durch Multiplikation mit der Matrix gestreckt, falls  $|\lambda| > 1$  und gestaucht, falls  $|\lambda| < 1$ . Seine Richtung kehrt sich um, falls  $\lambda < 0$ , sonst bleibt sie gleich.

### 9.2. Berechnung von Eigenwerten

Für einen Eigenvektor gilt

$$\begin{aligned} \mathbf{A} \cdot \mathbf{x} &= \lambda \cdot \mathbf{x} = \lambda \cdot \mathbf{I} \cdot \mathbf{x} \\ \iff \mathbf{A} \cdot \mathbf{x} - \lambda \cdot \mathbf{I} \cdot \mathbf{x} &= \mathbf{0} \\ \iff (\mathbf{A} - \lambda \cdot \mathbf{I}) \cdot \mathbf{x} &= \mathbf{0} \end{aligned} \tag{9.1}$$

9.1 hat nicht triviale Lösungen nur, falls

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

(da sonst nur genau eine Lösung  $\mathbf{x} = \mathbf{0}$  existiert)

Für  $\mathbf{A} \in \mathbb{C}^{n \times n}$  gilt:

$$\det(\mathbf{A} - \lambda \mathbf{I}) = (-\lambda)^n + \left( \sum_{i=1}^n a_{ii} \right) \cdot (-\lambda)^{n-1} + \dots + \det \mathbf{A} = \sum_{i=0}^n \alpha_i \cdot (-\lambda)^i \quad (9.2)$$

#### Definition 9.4

$\chi_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I})$  heißt charakteristisches Polynom.

$\chi_{\mathbf{A}}(\lambda) = 0$  heißt charakteristische Gleichung.

Die Spur einer Matrix  $\mathbf{A} \in \mathbb{C}^{n \times n}$  ist gegeben durch

$$\text{Spur } \mathbf{A} = \sum_{i=1}^n a_{ii}$$

#### Satz 9.5

$\lambda \in \mathbb{C}$  ist genau dann ein Eigenwert der Matrix  $\mathbf{A} \in \mathbb{C}^{n \times n}$ , wenn  $\lambda$  eine Nullstelle des charakteristischen Polynoms  $\chi_{\mathbf{A}}(\lambda)$  ist, d.h.  $\det(\mathbf{A} - \lambda \mathbf{I}) = 0$ .

**Beweis:** s.o. □

#### Algorithmus 9.6 (Berechnung der Eigenwerte und Eigenvektoren)

1. Bestimme die (verschiedenen) Nullstellen  $\lambda_1, \dots, \lambda_r$  von  $\chi_{\mathbf{A}}$
2. zu jedem Eigenwert  $\lambda_i$  bestimmt man die Lösungsmenge des homogenen LGS

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

also Kern  $(\mathbf{A} - \lambda_i \cdot \mathbf{I})$ .

Jeder Vektor aus dieser Menge ist ein Eigenvektor zum Eigenwert  $\lambda_i$

#### Anmerkung 9.7

- Die Bestimmung der Nullstellen von  $\chi_{\mathbf{A}}$  kann anspruchsvoll sein  $\Rightarrow$  numerische Methoden
- Da  $\chi_{\mathbf{A}}$  ein Polynom vom Grad  $n$  ist, hat es  $n$  (komplexe) Nullstellen. Es gibt also immer  $n$  Eigenwerte. Eine Nullstelle  $\lambda_i$  kann mehrfach vorkommen. In der Regel gibt es daher  $r \leq n$  verschiedene Eigenwerte.
- $\chi_{\mathbf{A}}$  lässt sich darstellen als

$$\chi_{\mathbf{A}}(\lambda) = (\lambda_1 - \lambda)^{k_1} \cdot (\lambda_2 - \lambda)^{k_2} \cdot \dots \cdot (\lambda_r - \lambda)^{k_r} = \prod_{i=1}^r (\lambda_i - \lambda)^{k_i} \quad (9.3)$$

$k_i$  heißt algebraische Vielfachheit des Eigenwertes  $\lambda_i$

- Koeffizientenvergleich von (9.2) und (9.3) ergibt:

$$\begin{aligned} \text{Spur } \mathbf{A} &= k_1 \lambda_1 + \dots + k_r \cdot \lambda_r = \sum_{i=1}^r k_i \lambda_i \\ \det \mathbf{A} &= \lambda_1^{k_1} \cdot \lambda_2^{k_2} \cdot \dots \cdot \lambda_r^{k_r} = \prod_{i=1}^r \lambda_i^{k_i} \end{aligned}$$

- $\dim(\text{Kern}(\mathbf{A} - \lambda_i \mathbf{I}))$  heißt geometrische Vielfachheit des Eigenwertes  $\lambda_i$ . Die algebraische und geometrische Vielfachheit ist im allgemeinen nicht gleich, letztere ist aber niemals größer als erstere.
- Die Eigenvektoren zu *verschiedenen* Eigenwerten sind *immer* linear unabhängig.

### Beispiel 9.8

Berechne Eigenwerte und Eigenvektoren von

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

$$\begin{aligned} \det(\mathbf{A} - \lambda \mathbf{I}) &= \begin{vmatrix} -2 - \lambda & -5 \\ 1 & 4 - \lambda \end{vmatrix} \\ &= (-2 - \lambda) \cdot (4 - \lambda) - 1 \cdot (-5) \\ &= -8 - 4\lambda + 2\lambda + \lambda^2 + 5 \\ &= \lambda^2 - 2\lambda - 3 \end{aligned}$$

Nullstellen:

$$\begin{aligned} \lambda_{1/2} &= \frac{2 \pm \sqrt{4 + 12}}{2} = 1 \pm 2 \\ \lambda_1 &= 3, \quad \lambda_2 = -1 \end{aligned}$$

Eigenvektor zu  $\lambda_1$ :

$$\begin{aligned} (\mathbf{A} - 3 \cdot \mathbf{I}) \cdot \mathbf{x} &= \begin{pmatrix} -2 - 3 & -5 \\ 1 & 4 - 3 \end{pmatrix} \cdot \mathbf{x} = \begin{pmatrix} -5 & -5 \\ 1 & 1 \end{pmatrix} \cdot \mathbf{x} = \mathbf{0} \\ \Rightarrow x_1 &= -x_2, \quad \text{wähle } x_2 = \alpha \quad \Rightarrow \quad \mathbf{EV}_1 = \alpha \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} \end{aligned}$$

Eigenvektor zu  $\lambda_2$ :

$$\begin{aligned} (\mathbf{A} - (-1) \cdot \mathbf{I}) \cdot \mathbf{x} &= \begin{pmatrix} -2 + 1 & -5 \\ 1 & 4 + 1 \end{pmatrix} \cdot \mathbf{x} = \begin{pmatrix} -1 & -5 \\ 1 & 5 \end{pmatrix} \cdot \mathbf{x} = \mathbf{0} \\ \Rightarrow x_1 &= -5x_2, \quad \text{wähle } x_2 = \beta \quad \Rightarrow \quad \mathbf{EV}_2 = \beta \cdot \begin{pmatrix} -5 \\ 1 \end{pmatrix} \end{aligned}$$

Test:

$$\begin{aligned} \begin{pmatrix} -2 & -5 \\ 1 & 4 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} &= \begin{pmatrix} 2 - 5 \\ -1 + 4 \end{pmatrix} = \begin{pmatrix} -3 \\ 3 \end{pmatrix} = 3 \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} \\ \begin{pmatrix} -2 & -5 \\ 1 & 4 \end{pmatrix} \cdot \begin{pmatrix} -5 \\ 1 \end{pmatrix} &= \begin{pmatrix} 10 - 5 \\ -5 + 4 \end{pmatrix} = \begin{pmatrix} 5 \\ -1 \end{pmatrix} = -1 \cdot \begin{pmatrix} -5 \\ 1 \end{pmatrix} \end{aligned}$$

Überprüfe Anmerkung 9.7

$$\text{Spur } \mathbf{A} = -2 + 4 = 2 = 3 - 1$$

$$\begin{vmatrix} -2 & -5 \\ 1 & 4 \end{vmatrix} = -2 \cdot 4 - 1 \cdot (-5) = -8 + 5 = -3 = 3 \cdot (-1)$$

### 9.3. Spezielle Matrizen

#### Satz 9.9

Die Eigenwerte einer Dreiecksmatrix  $\mathbf{A} \in \mathbb{C}^{n \times n}$  sind gleich den Diagonalelementen von  $\mathbf{A}$

$$\lambda_i = a_{ii} \quad (1 \leq i \leq n)$$

#### Beweis:

Da  $\mathbf{A}$  Dreiecksmatrix  $\Rightarrow \det(\mathbf{A} - \lambda \mathbf{I})$  auch Dreiecksmatrix

$$\Rightarrow \det(\mathbf{A} - \lambda \mathbf{I}) \stackrel{(4.5)}{=} \prod_{i=1}^n (a_{ii} - \lambda)$$

$$\prod_{i=1}^n (a_{ii} - \lambda) = 0 \quad \Rightarrow \quad \text{Behauptung} \quad \square$$

#### Satz 9.10

Sei  $\mathbf{A} \in \mathbb{C}^{n \times n}$

1. Hat  $\mathbf{A}$  den Eigenvektor  $\mathbf{x} \in \mathbb{C}^n$  zum Eigenwert  $\lambda$ , dann ist dieser auch Eigenvektor von

$$\alpha \mathbf{A}, \mathbf{A}^m, \mathbf{A} + \beta \mathbf{I}, p(\mathbf{A}) := \alpha_p \mathbf{A}^p + \dots + \alpha_1 \mathbf{A} + \alpha_0 \mathbf{I}$$

allerdings zum Eigenwert

$$\alpha \lambda, \lambda^m, \lambda + \beta, p(\lambda) := \alpha_p \lambda^p + \dots + \alpha_1 \lambda + \alpha_0.$$

2.  $\mathbf{A}$  und  $\mathbf{A}^T$  haben die gleichen Eigenwerte (aber in der Regel andere Eigenvektoren).
3.  $\mathbf{A}$  ist genau dann invertierbar, wenn alle Eigenwerte  $\neq 0$  sind. Ist  $\lambda$  Eigenwert einer invertierbaren Matrix  $\mathbf{A}$  mit Eigenvektor  $\mathbf{x}$ , dann ist  $\frac{1}{\lambda}$  Eigenwert von  $\mathbf{A}^{-1}$  mit gleichem Eigenvektor  $\mathbf{x}$ .

#### Beweis:

1. Aus  $\mathbf{A} \cdot \mathbf{x} = \lambda \cdot \mathbf{x}$  folgt  $p(\mathbf{A}) \cdot \mathbf{x} = p(\lambda) \cdot \mathbf{x}$
2.  $\det(\mathbf{A}^T - \lambda \cdot \mathbf{I}) = \det(\mathbf{A} - \lambda \mathbf{I})^T = \det(\mathbf{A} - \lambda \mathbf{I})$
3.  $\mathbf{A}$  invertierbar  $\iff \det \mathbf{A} = \lambda_1^{k_1} \cdot \dots \cdot \lambda_r^{k_r} \neq 0 \iff$  alle  $\lambda_i \neq 0$

$$\mathbf{A} \cdot \mathbf{x} = \lambda \cdot \mathbf{x} \iff \frac{1}{\lambda} \mathbf{x} = \mathbf{A}^{-1} \mathbf{x}$$

$\square$

#### Definition 9.11

Sei  $\mathbf{A} \in \mathbb{C}^{n \times n}$  und  $\mathbf{A}^H = \overline{\mathbf{A}}^T$  ihre komplex-konjugierte und transponierte Matrix.  $\mathbf{A}$  heißt hermitesch, wenn gilt:

$$a_{ij} = \overline{a_{ji}} \quad \forall 1 \leq i, j \leq n \quad \text{bzw.} \quad \mathbf{A} = \mathbf{A}^H$$

Bei reellen Matrizen entspricht das einer symmetrischen Matrix ( $\mathbf{A} = \mathbf{A}^T$ ).

**Definition 9.12**

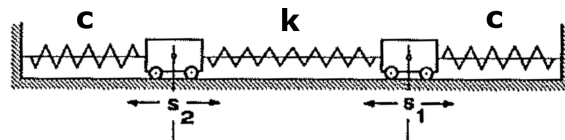
Eine Matrix  $\mathbf{U} \in \mathbb{C}^{n \times n}$  heißt unitär, wenn ihre Inverse gleich ihrer komplex-konjugierten und transponierten Matrix ist, d.h. falls  $\mathbf{U}\mathbf{U}^H = \mathbf{U}^H\mathbf{U} = \mathbf{I}$ . Für reelle Matrizen  $\mathbf{U} \in \mathbb{R}^{n \times n}$  heißt das, dass  $\mathbf{U}^{-1} = \mathbf{U}^T$ . Eine solche reelle Matrix heißt orthogonal.

**Satz 9.13 (Eigenwerte von hermiteschen Matrizen)**

Sei  $\mathbf{A} \in \mathbb{C}^{n \times n}$  hermitesch. Dann gilt

1. Alle  $n$  Eigenwerte sind reell.
2. Es gibt genau  $n$  linear unabhängige Eigenvektoren.
3. Zu jedem Eigenwert der algebraischen Vielfachheit  $k$  gibt es genau  $k$  linear unabhängige Eigenvektoren (d.h. die algebraische und geometrische Vielfachheit ist gleich).
4. Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.
5. Es lässt sich eine Orthonormalbasis des  $\mathbb{C}^n$  aus Eigenvektoren von  $\mathbf{A}$  aufbauen.
6. Es existiert eine Zerlegung  $\mathbf{W}^{-1}\mathbf{A}\mathbf{W} = \mathbf{D}$  bzw.  $\mathbf{A} = \mathbf{W}\mathbf{D}\mathbf{W}^{-1}$ . Die Spalten von  $\mathbf{W}$  sind dabei (linear unabhängige) Eigenvektoren  $\mathbf{E}\mathbf{V}_1, \dots, \mathbf{E}\mathbf{V}_n$  von  $\mathbf{A}$  und  $\mathbf{D}$  ist eine Diagonalmatrix, deren Diagonalelemente die zugehörigen Eigenwerte  $\lambda_1, \dots, \lambda_n$  sind.
7. Es existiert sogar eine entsprechende Zerlegung  $\mathbf{U}^H\mathbf{A}\mathbf{U} = \mathbf{D}$  bzw.  $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^H$ , mit einer unitären Matrix  $\mathbf{U}$ .

ohne Beweis

**9.4. Anwendung**

aus: Meyberg, Vachnauer: Höhere Mathematik 1, Springer

mit Federkonstanten  $c, k > 0$ , beide Wagen haben die gleiche Masse  $m > 0$ , die Masse der Federn sei vernachlässigbar

Betrachte Auslenkungen  $s_1, s_2$  aus der Ruhelage, Auslenkungen nach rechts seien positiv, in Ruhelage gelte  $s_1 = s_2 = 0$ .

Hookesches Gesetz:

$$\begin{aligned} \frac{d^2 s_1}{dt^2} &= -\frac{c}{m}s_1 + \frac{k}{m}(s_2 - s_1) = -\frac{c+k}{m}s_1 + \frac{k}{m}s_2 \\ \frac{d^2 s_2}{dt^2} &= \frac{k}{m}(s_1 - s_2) - \frac{c}{m}s_2 = \frac{k}{m}s_1 - \frac{c+k}{m}s_2 \\ \begin{pmatrix} \frac{d^2 s_1}{dt^2} \\ \frac{d^2 s_2}{dt^2} \end{pmatrix} + \underbrace{\begin{pmatrix} \frac{c+k}{m} & -\frac{k}{m} \\ -\frac{k}{m} & \frac{c+k}{m} \end{pmatrix}}_{\mathbf{A}} \cdot \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \end{aligned}$$

### Ansatz

$$\begin{pmatrix} s_1(t) \\ s_2(t) \end{pmatrix} = \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \cdot e^{i\omega t} = \mathbf{s} \cdot e^{i\omega t}$$

Einsetzen

$$\begin{aligned} \mathbf{s} \cdot (-\omega^2 e^{i\omega t}) + \mathbf{A} \cdot \mathbf{s} \cdot e^{i\omega t} &= (-\omega^2 \cdot \mathbf{s} + \mathbf{A} \cdot \mathbf{s}) \cdot e^{i\omega t} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \Leftrightarrow \mathbf{A} \cdot \mathbf{s} &= \omega^2 \cdot \mathbf{s}. \end{aligned}$$

Also muss  $\mathbf{s}$  ein Eigenvektor und  $\omega^2$  ein Eigenwert von  $\mathbf{A}$  sein.  
Zur Abkürzung schreiben wir

$$\begin{aligned} \alpha &= \frac{c+k}{m}, \quad \beta = \frac{k}{m} \\ \mathbf{A} &= \begin{pmatrix} \alpha & -\beta \\ -\beta & \alpha \end{pmatrix} \end{aligned}$$

### Eigenwerte

$$\begin{aligned} \det(\mathbf{A} - \lambda \mathbf{I}) &= \begin{vmatrix} \alpha - \lambda & -\beta \\ -\beta & \alpha - \lambda \end{vmatrix} = (\alpha - \lambda)^2 - \beta^2 = 0 \\ \Rightarrow (\alpha - \lambda)^2 &= \beta^2 \\ \Rightarrow \alpha - \lambda &= \pm \beta \\ \Rightarrow \lambda &= \alpha \pm \beta \\ \Rightarrow \lambda_1 &= \frac{c+k}{m} - \frac{k}{m} = \frac{c}{m} \\ \lambda_2 &= \frac{c+k}{m} + \frac{k}{m} = \frac{c+2k}{m} \\ \Rightarrow \omega_1 &= \sqrt{\frac{c}{m}}, \quad \omega_2 = -\sqrt{\frac{c}{m}} = -\omega_1, \quad \omega_3 = \sqrt{\frac{c+2k}{m}}, \quad \omega_4 = -\sqrt{\frac{c+2k}{m}} = -\omega_3 \end{aligned}$$

$\omega_1, \omega_2, \omega_3$  und  $\omega_4$  sind die Frequenzen von Eigenschwingungen.

### Eigenvektoren

$$\begin{aligned} (\mathbf{A} - (\alpha - \beta) \cdot \mathbf{I}) \cdot \mathbf{EV}_1 &= 0 \\ \begin{pmatrix} \alpha - (\alpha - \beta) & -\beta \\ -\beta & \alpha - (\alpha - \beta) \end{pmatrix} \cdot \mathbf{EV}_1 &= \begin{pmatrix} \beta & -\beta \\ -\beta & \beta \end{pmatrix} \cdot \mathbf{EV}_1 = 0 \\ \Rightarrow \mathbf{EV}_1 &= \gamma \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ \begin{pmatrix} \alpha - (\alpha + \beta) & -\beta \\ -\beta & \alpha - (\alpha + \beta) \end{pmatrix} \cdot \mathbf{EV}_2 &= \begin{pmatrix} -\beta & -\beta \\ -\beta & -\beta \end{pmatrix} \cdot \mathbf{EV}_2 = 0 \\ \Rightarrow \mathbf{EV}_2 &= \gamma \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} \end{aligned}$$



⇒ Bei Eigenwert  $\lambda_1$  schwingen die Massen in Phase

$$\overset{s_1}{\rightarrow} \quad \overset{s_2}{\rightarrow} \quad \text{oder} \quad \overset{s_1}{\leftarrow} \quad \overset{s_2}{\leftarrow}$$

Bei Eigenwert  $\lambda_2$  schwingen die Massen in Gegenphase

$$\overset{s_1}{\rightarrow} \quad \overset{s_2}{\leftarrow} \quad \text{oder} \quad \overset{s_1}{\leftarrow} \quad \overset{s_2}{\rightarrow}$$

Allgemeine Lösung:

$$\mathbf{s}(t) = \mu_1 \cdot \mathbf{EV}_1 \cdot e^{i\omega_1 t} + \mu_2 \cdot \mathbf{EV}_1 \cdot e^{-i\omega_1 t} + \mu_3 \cdot \mathbf{EV}_2 \cdot e^{i\omega_3 t} + \mu_4 \cdot \mathbf{EV}_2 \cdot e^{-i\omega_3 t}$$

Da die Eigenvektoren linear unabhängig sind, lässt sich jeder Anfangszustand als Linearkombination der Eigenvektoren darstellen.

$$\mathbf{s}_0 = \mathbf{s}(0) = (\mu_1 + \mu_2) \cdot \mathbf{EV}_1 + (\mu_3 + \mu_4) \cdot \mathbf{EV}_2$$

$\mu_1$  und  $\mu_2$  bestimmt man aus den Anfangsbedingungen. Zusätzlich brauchen wir noch Anfangsbedingungen für die Geschwindigkeit

$$\mathbf{v}(t) = \frac{d\mathbf{s}(t)}{dt} = \mu_1 \cdot \mathbf{EV}_1 \cdot (i\omega_1) \cdot e^{i\omega_1 t} + \mu_2 \cdot \mathbf{EV}_1 \cdot (-i\omega_1) \cdot e^{-i\omega_1 t} + \mu_3 \cdot \mathbf{EV}_2 \cdot (i\omega_3) \cdot e^{i\omega_3 t} + \mu_4 \cdot \mathbf{EV}_2 \cdot (-i\omega_3) \cdot e^{-i\omega_3 t}$$

$$\mathbf{v}_0 = \mathbf{v}(0) = (\mu_1 - \mu_2) \cdot (i\omega_1) \cdot \mathbf{EV}_1 + (\mu_3 - \mu_4) \cdot (i\omega_3) \cdot \mathbf{EV}_2 = [(\mu_1 - \mu_2) \cdot \omega_1 \cdot \mathbf{EV}_1 + (\mu_3 - \mu_4) \cdot \omega_3 \cdot \mathbf{EV}_2] i$$

Mögliche Anfangspositionen:

$$1. \quad s_1 = s_2 = s, \quad v_1 = v_2 = 0$$

$$\begin{aligned} \mathbf{s}_0 &= \begin{pmatrix} s \\ s \end{pmatrix} = (\mu_1 + \mu_2) \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + (\mu_3 + \mu_4) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} \mu_1 + \mu_2 + \mu_3 + \mu_4 \\ \mu_1 + \mu_2 - (\mu_3 + \mu_4) \end{pmatrix} \\ \mathbf{v}_0 &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} = (\mu_1 - \mu_2) \cdot (i\omega_1) \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} + (\mu_3 - \mu_4) \cdot (i\omega_2) \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ &= \begin{pmatrix} (\mu_1 - \mu_2) \cdot (i\omega_1) + (\mu_3 - \mu_4) \cdot (i\omega_3) \\ (\mu_1 - \mu_2) \cdot (i\omega_1) - (\mu_3 - \mu_4) \cdot (i\omega_3) \end{pmatrix} \end{aligned}$$

Wir haben also die vier Gleichungen

$$\begin{aligned} (1) : \quad & \mu_1 + \mu_2 + \mu_3 + \mu_4 = s \\ (2) : \quad & \mu_1 + \mu_2 - (\mu_3 + \mu_4) = s \\ (3) : \quad & (\mu_1 - \mu_2) \cdot (i\omega_1) + (\mu_3 - \mu_4) \cdot (i\omega_3) = 0 \\ (4) : \quad & (\mu_1 - \mu_2) \cdot (i\omega_1) - (\mu_3 - \mu_4) \cdot (i\omega_3) = 0 \\ (3) + (4) : \quad & \Rightarrow 2(\mu_1 - \mu_2) \cdot (i\omega_1) = 0 \quad \Rightarrow \mu_1 - \mu_2 = 0 \quad \Rightarrow \mu_1 = \mu_2 \\ (1) + (2) : \quad & \Rightarrow 2(\mu_1 + \mu_2) = 2s \quad \Rightarrow 4\mu_1 = 2s \quad \Rightarrow \mu_1 = \mu_2 = \frac{s}{2} \\ (1) - (2) : \quad & \Rightarrow 2(\mu_3 + \mu_4) = 0 \Rightarrow \mu_3 = -\mu_4 \\ (3) - (4) : \quad & \Rightarrow 2(\mu_3 - \mu_4) \cdot (i\omega_3) = 0 \Rightarrow 2\mu_3 \cdot (i\omega_3) = 0 \quad \Rightarrow \mu_3 = 0 \quad \Rightarrow \mu_4 = -\mu_3 = 0 \end{aligned}$$

Lösung:  $\mu_1 = \mu_2 = \frac{s}{2}$ ,  $\mu_3 = \mu_4 = 0 \Rightarrow$  Eigenschwingung 1

2.  $s_1 = -s_2 = s, v_1 = v_2 = 0$

Lösung:  $\mu_1 = \mu_2 = 0, \mu_3 = \mu_4 = \frac{s}{2} \Rightarrow$  Eigenschwingung 2

3.  $s_1 = s, s_2 = 0, v_1 = v_2 = 0$

$$\begin{pmatrix} s \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \mu_1 + \mu_2 + \mu_3 + \mu_4 \\ \mu_1 + \mu_2 - \mu_3 - \mu_4 \\ (\mu_1 - \mu_2) \cdot (i\omega_1) + (\mu_3 - \mu_4) \cdot (i\omega_3) \\ (\mu_1 - \mu_2) \cdot (i\omega_1) - (\mu_3 - \mu_4) \cdot (i\omega_3) \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ i\omega_1 & -i\omega_1 & i\omega_3 & -i\omega_3 \\ i\omega_1 & -i\omega_1 & -i\omega_3 & i\omega_3 \end{pmatrix} \cdot \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix} = \begin{pmatrix} s \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Löse lineares Gleichungssystem mit Gauß-Elimination:

```
>>> import sympy as sp
2 >>> om1=sp.Symbol("omega_1")
>>> om3=sp.Symbol("omega_3")
4 >>> A=sp.Matrix([[1,1,1,1],[1,1,-1,-1],
                    [om1*sp.I,-om1*sp.I,om3*sp.I,-om3*sp.I],
6                    [om1*sp.I,-om1*sp.I,-om3*sp.I,om3*sp.I]])
>>> s=sp.Symbol("s")
8 >>> b=sp.Matrix([s,0,0,0])
>>> A.solve(b)
10 Matrix([
[
s*(I*(-I*omega_1 - I*omega_3)/(8*omega_1) +
I*(-I*omega_1 + I*omega_3)/(8*omega_1))],
12 [s*(1/2 - I*(-I*omega_1 - I*omega_3)/(8*omega_1) -
I*(-I*omega_1 + I*omega_3)/(8*omega_1))],
[
s/4],
14 [
s/4]])
>>> sp.simplify(_)
16 Matrix([
[s/4],
18 [s/4],
[s/4],
20 [s/4]])
```

$$\Rightarrow \mu_1 = \mu_2 = \mu_3 = \mu_4 = \frac{s}{4}$$

$$\begin{aligned} \Rightarrow \mathbf{s}(t) &= \frac{s}{4} \left[ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \underbrace{(e^{i\omega_1 t} + e^{-i\omega_1 t})}_{=2 \cos(\omega_1 t)} + \begin{pmatrix} 1 \\ -1 \end{pmatrix} \underbrace{(e^{i\omega_3 t} + e^{-i\omega_3 t})}_{=2 \cos(\omega_3 t)} \right] \\ &= \frac{s}{2} \begin{pmatrix} \cos(\omega_1 t) + \cos(\omega_3 t) \\ \cos(\omega_1 t) - \cos(\omega_3 t) \end{pmatrix} \end{aligned}$$

mit den Parametern  $c = 1$ ,  $k = 1$  und  $m = 1$  erhalten wir

$$\omega_1 = \sqrt{1} = 1, \quad \omega_3 = \sqrt{3}$$

mit den Parameter  $c = 1$ ,  $k = 2$  und  $m = 1$

$$\omega_1 = 1, \quad \omega_3 = \sqrt{5}$$

Großes  $k \Rightarrow$  stärkerer Unterschied in den Eigenfrequenzen

Großes  $m \Rightarrow$  langsamere Schwingung

Großes  $c \Rightarrow$  schnellere Schwingung

4.  $s_1 = s_2 = 0$ ,  $v_1 = -1$ ,  $v_2 = 1$  Wir erhalten mit der Matrix von oben aber einer angepassten rechten Seite:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ i\omega_1 & -i\omega_1 & i\omega_3 & -i\omega_3 \\ i\omega_1 & -i\omega_1 & -i\omega_3 & i\omega_3 \end{pmatrix} \cdot \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}$$

Löse lineares Gleichungssystem mit Gauß-Elimination:

```
>>> import sympy as sp
2 >>> om1=sp.Symbol("omega_1")
>>> om3=sp.Symbol("omega_3")
4 >>> A=sp.Matrix([[1,1,1,1],[1,1,-1,-1],
                    [om1*sp.I,-om1*sp.I,om3*sp.I,-om3*sp.I],
                    [om1*sp.I,-om1*sp.I,-om3*sp.I,om3*sp.I]])
6 >>> s=sp.Symbol("s")
8 >>> b=sp.Matrix([0,0,-1,1])
>>> A.solve(b)
10 Matrix([
    [I*(-I*(-I*omega_1 - I*omega_3)/(2*omega_1) + I*(-I*omega_1
      + I*omega_3)/(2*omega_1))/(2*omega_3) + I/(2*omega_1)],
12 [ I*(I*(-I*omega_1 - I*omega_3)/(2*omega_1) - I*(-I*omega_1
      + I*omega_3)/(2*omega_1))/(2*omega_3) - I/(2*omega_1)],
    [
14 [
        I/(2*omega_3)],
        -I/(2*omega_3)]]])
>>> sp.simplify(_)
16 Matrix([
    [
18 [
        I/(2*omega_3)],
20 [-I/(2*omega_3)]]])
```

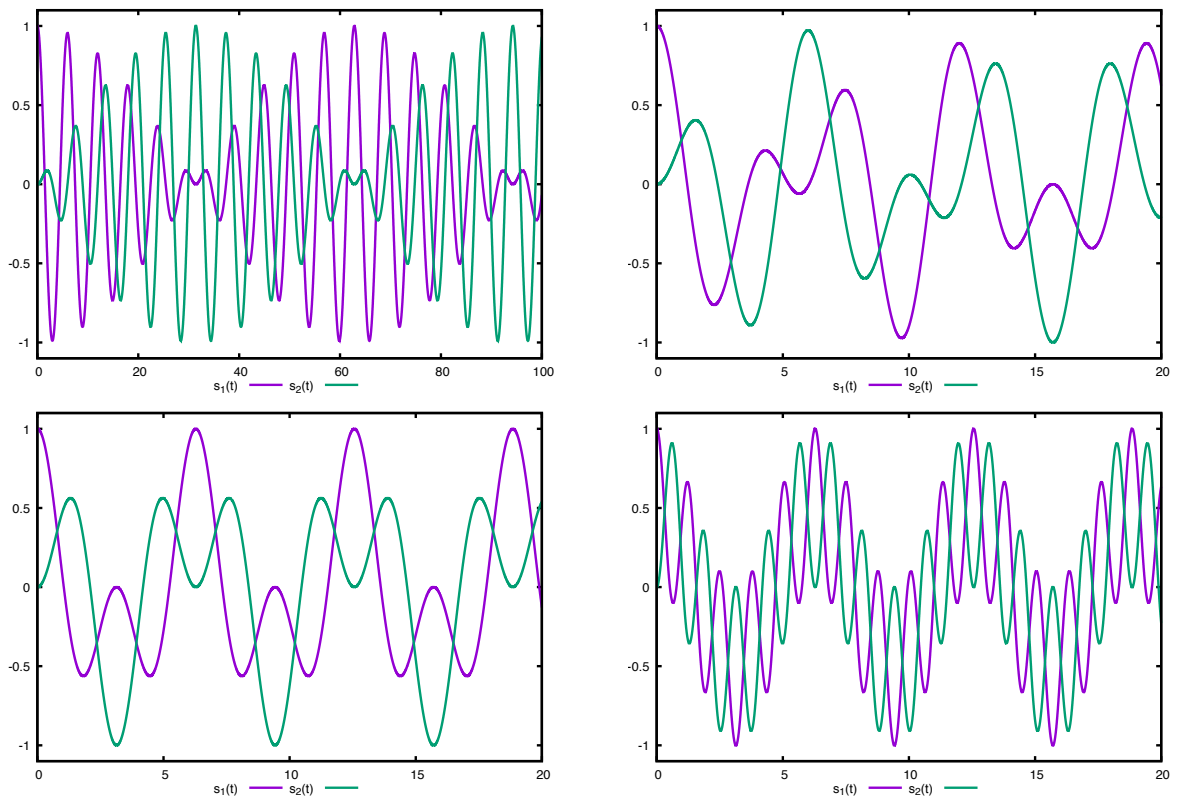


Abbildung 1: Koordinaten  $s_1(t)$  und  $s_2(t)$  für  $\alpha_1 = \alpha_2 = 0.5$  bei schwacher Kopplungsfeder ( $\omega_1 = 1$ ,  $\omega_3 = 1.1$ , oben links), etwas stärkerer Kopplung ( $\omega_1 = 1$ ,  $\omega_3 = 1.6$ , oben rechts), noch stärkerer Kopplung ( $\omega_1 = 1$ ,  $\omega_3 = 2$ , unten links) und starker Kopplung ( $\omega_1 = 1$ ,  $\omega_3 = 5$ , unten rechts)

$$\Rightarrow \mu_1 = \mu_2 = 0, \mu_3 = \frac{i}{2\omega_3}, \mu_4 = -\frac{i}{2\omega_3}$$

$$\Rightarrow \mathbf{s}(t) = \frac{i}{2\omega_3} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \underbrace{(e^{i\omega_3 t} - e^{-i\omega_3 t})}_{=2i \sin(\omega_3 t)} = \frac{1}{\omega_3} \begin{pmatrix} -\sin(\omega_3 t) \\ \sin(\omega_3 t) \end{pmatrix}$$

### 3-Minutes Python 13

In `sympy` hat jede Matrix die Methoden `eigenvals` und `eigenvecs` um die Eigenwerte und Eigenvektoren zu berechnen. `eigenvecs` gibt dabei jeweils den Eigenwert, seine algebraische Vielfachheit und den Eigenvektor zurück.

```
>>> A = sp.Matrix([[3, -2, 4, -2], [5, 3, -3, -2], [5, -2, 2,
-2], [5, -2, -3, 3]])
2 >>> A.eigenvals()
{3: 1, -2: 1, 5: 2}
4 >>> A.eigenvecs()
[(-2, 1, [Matrix([
```

```

6 [0],
  [1],
8 [1],
  [1]]]]), (3, 1, [Matrix([
10 [1],
   [1],
12 [1],
   [1]]]]), (5, 2, [Matrix([
14 [1],
   [1],
16 [1],
   [0]])], Matrix([
18 [ 0],
   [-1],
20 [ 0],
   [ 1]]]]))

```

Alternativ kann man sich auch das charakteristische Polynom mit `charpoly` ausgeben lassen, dieses mit `fact` faktorisieren und daraus die Eigenwerte ablesen.

```

>>> lamda=sp.symbols('lambda')
2 >>> p=A.charpoly(lamda)
>>> p
4 PurePoly(lambda**4 - 11*lambda**3 + 29*lambda**2 + 35*lambda
   - 150, lambda, domain='ZZ')
>>> sp.factor(p)
6 (lambda - 5)**2*(lambda - 3)*(lambda + 2)

```

Auch für die Diagonalisierung einer Matrix gibt es eine Methode:

```

>>> P,D=A.diagonalize()
2 >>> P
Matrix([
4 [0, 1, 1,  0],
  [1, 1, 1, -1],
6 [1, 1, 1,  0],
  [1, 1, 0,  1]])
8 >>> D
Matrix([
10 [-2, 0, 0, 0],
   [ 0, 3, 0, 0],
12 [ 0, 0, 5, 0],
   [ 0, 0, 0, 5]])
14 >>> P*D*P.inv()
Matrix([
16 [3, -2,  4, -2],
   [5,  3, -3, -2],
18 [5, -2,  2, -2],

```

```
[5, -2, -3, 3]])
```

In numpy gibt es die Funktionen `linalg.eigvals` und `linalg.eig`. Erstere liefert nur die Eigenwerte, letztere auch die Eigenvektoren zurück.

```
>>> A = np.array([[3, -2, 4, -2], [5, 3, -3, -2], [5, -2, 2,
-2], [5, -2, -3, 3]])
2 >>> np.linalg.eigvals(A)
array([ 3., -2., 5., 5.])
4 >>> np.linalg.eig(A)
(array([ 3., -2., 5., 5.]), array([[ 5.00000000e-01,
0.00000000e+00, -5.77350269e-01,
6 -6.07237672e-01],
[ 5.00000000e-01, -5.77350269e-01, -6.17589503e-16,
8 -5.01300427e-01],
[ 5.00000000e-01, -5.77350269e-01, -5.77350269e-01,
10 -6.07237672e-01],
[ 5.00000000e-01, -5.77350269e-01, -5.77350269e-01,
12 -1.05937245e-01]]))
```

Die Funktionen von `numpy` sind wesentlich schneller und lassen auch die Eigenwertberechnung für große Matrizen (mehrere tausend Einträge) in akzeptabler Zeit zu.

Zum Beispiel erzeugt das folgende Programm eine  $2000 \times 2000$  Matrix mit Zufallszahlen und berechnet die Eigenwerte in wenigen Sekunden:

```
A=np.random.rand(2000,2000)
2 np.linalg.eigvals(A)
```

## Teil II.

# Analysis von Funktionen in mehreren Variablen

## 10. Ableitung von Funktionen mehrerer Variablen

### 10.1. Motivation

In realen Anwendungen treten häufig Funktionen auf, die von mehreren Variablen abhängen.

#### Beispiel 10.1

- Temperaturverteilung  $T(x, y, z)$  in einem dreidimensionalen Körper (z.B. Motor) oder falls zeitabhängig  $T(x, y, z, t)$
- Energie eines bewegten Körpers

$$E = E_{\text{kin}} + E_{\text{pot}} = \frac{1}{2} \cdot m \cdot v^2 + m \cdot g \cdot h = \frac{1}{2} \cdot m \cdot \|\mathbf{v}\|^2 + m \cdot g \cdot h = \frac{1}{2} \cdot m \cdot \langle \mathbf{v}, \mathbf{v} \rangle + m \cdot g \cdot h = E(\mathbf{v}, h)$$

#### Anmerkung 10.2

- Funktionen mehrerer Variablen sind Abbildungen der Form  $f : D \rightarrow \mathbb{R}^m$  mit  $D \subseteq \mathbb{R}^n$ . Wir werden uns für's erste auf  $m = 1$  beschränken.

- Für  $f \left( \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right)$  schreiben wir der Einfachheit halber  $f(x_1, \dots, x_n)$ .

**Wiederholung:** Bei der Analyse von Funktionen einer Variablen haben wir die Begriffe der Stetigkeit und Differenzierbarkeit eingeführt. Diese waren wie folgt definiert:

#### Stetigkeit

- $f : I \rightarrow \mathbb{R}$  ist stetig in  $\xi \in I$ , wenn zu jedem  $\epsilon > 0$  ein  $\delta > 0$  existiert, so dass für alle  $x \in I$  mit  $|x - \xi| < \delta$  gilt:  $|f(x) - f(\xi)| < \epsilon$
- $f(x)$  ist stetig, wenn aus  $\lim_{k \rightarrow \infty} x_k = \xi$  immer folgt  $\lim_{k \rightarrow \infty} f(x_k) = f(\xi)$

#### Differenzierbarkeit

$f : I \rightarrow \mathbb{R}$ ,  $I \subseteq \mathbb{R}$  ist differenzierbar in  $x_0 \in I$ , wenn

1. Der Grenzwert

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

existiert.

2.  $\exists a \in \mathbb{R}$  und eine Funktion  $r$  mit  $f(x_0 + h) = f(x_0) + a \cdot h + r(h)$  mit  $\frac{r(h)}{h} \rightarrow 0$  für  $h \rightarrow 0$ .  
 $a$  ist dann die Ableitung von  $f$  im Punkt  $x_0$ .

*Wie übertragen wir diese Definitionen auf Funktionen mehrerer Variablen?*

## 10.2. Grundlegende Begriffe

Zunächst brauchen wir ein Konzept, um die Umgebung eines Punktes zu definieren:

### Definition 10.3

Sei  $\mathbf{x} \in \mathbb{R}^n$  und  $\delta > 0$ .

Wir nennen die Menge

$$U_\delta(\mathbf{x}) := \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{y}\| < \delta\}$$

die  $\delta$ -Umgebung von  $\mathbf{x}$ .

Dies ist eine offene Kugel (der Rand ist nicht enthalten) mit Radius  $\delta$  und Mittelpunkt  $\mathbf{x}$ .

Damit können wir Gebiete folgendermaßen charakterisieren:

### Definition 10.4

Sei  $D$  eine Teilmenge des  $\mathbb{R}^n$

1. Ein Punkt  $\mathbf{x} \in D$  heißt innerer Punkt von  $D$ , wenn es eine  $\delta$ -Umgebung von  $\mathbf{x}$  gibt, die ganz in  $D$  enthalten ist.
2.  $D$  heißt offen, wenn jeder Punkt von  $D$  ein innerer Punkt ist.
3. Ein Punkt  $\mathbf{x} \in \mathbb{R}^n$  heißt Randpunkt von  $D$ , wenn jede  $\delta$ -Umgebung von  $\mathbf{x}$  sowohl mindestens einen Punkt aus  $D$ , als auch mindestens einen nicht zu  $D$  gehörenden Punkt enthält.  
 Die Menge aller Randpunkte von  $D$  heißt Rand von  $D$  und wird mit  $\partial D$  bezeichnet.
4. Eine Menge heißt abgeschlossen, wenn sie alle ihre Randpunkte erhält.

### Beispiel 10.5

Offene Kreisscheibe:

$$K = \{(x, y) : (x - x_0)^2 + (y - y_0)^2 < r^2\}$$

Rand

$$\partial K = \{(x, y) : (x - x_0)^2 + (y - y_0)^2 = r^2\}$$

Abgeschlossene Kreisscheibe

$$K \cup \partial K = \{(x, y) : (x - x_0)^2 + (y - y_0)^2 \leq r^2\}$$



### 10.3. Grenzwert und Stetigkeit

#### Definition 10.6

Sei  $D \subseteq \mathbb{R}^n$ ,  $f : D \rightarrow \mathbb{R}$  und  $\mathbf{a} \in D$

1.  $f$  hat in  $\mathbf{a}$  den Grenzwert  $\lambda \in \mathbb{R}$

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = \lambda,$$

wenn es zu jedem  $\varepsilon > 0$  eine  $\delta$ -Umgebung  $\mathbf{U}_\delta$  von  $\mathbf{a}$  gibt, so dass

$$|f(\mathbf{x}) - \lambda| \leq \varepsilon \quad \forall \mathbf{x} \in D \cap \mathbf{U}_\delta(\mathbf{a})$$

2.  $f$  heißt in  $\mathbf{a} \in D$  stetig, wenn  $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = f(\mathbf{a})$
3.  $f$  heißt auf  $D$  stetig, wenn  $f$  in allen  $\mathbf{a} \in D$  stetig ist.

### 10.4. Partielle Ableitung

#### Definition 10.7

Sei  $D \subseteq \mathbb{R}^n$  offen,  $f : D \rightarrow \mathbb{R}$  und  $\mathbf{a} = (a_1, a_2, \dots, a_n)^T \in D$ . Existiert die Ableitung der Hilfsfunktion  $h_i : x \rightarrow f(a_1, \dots, a_{i-1}, x, a_{i+1}, \dots, a_n)$  bei der alle  $a_j$  mit  $j \neq i$  konstant angenommen werden, dann nennt man  $h'_i(a_i)$  die partielle Ableitung von  $f$  nach  $x_i$  an der Stelle  $\mathbf{a}$  und schreibt

$$\frac{\partial f}{\partial x_i}(\mathbf{a}) := h'_i(a_i)$$

#### Beispiel 10.8

$$f(x, y) = 100 \cdot x^2 + y^4 + x \cdot y$$

$$\frac{\partial f}{\partial x} = 200 \cdot x + y$$

$$\frac{\partial f}{\partial y} = 4y^3 + x$$

#### Anmerkung 10.9

Statt  $\frac{\partial f}{\partial x}(\mathbf{a})$  schreibt man auch  $\partial_x f(\mathbf{a})$  oder  $f_x(\mathbf{a})$

#### Definition 10.10

Sei  $D \subseteq \mathbb{R}^n$  offen,  $f : D \rightarrow \mathbb{R}$  und  $\mathbf{a} \in D$ .

$f$  heißt (stetig) partiell differenzierbar in  $\mathbf{a}$ , wenn alle partiellen Ableitungen existieren (und stetig sind).

Man nennt dann

$$\nabla f(\mathbf{a}) := \begin{pmatrix} \frac{\partial f}{\partial x_1}(\mathbf{a}) \\ \frac{\partial f}{\partial x_2}(\mathbf{a}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\mathbf{a}) \end{pmatrix} \in \mathbb{R}^n$$

den Gradienten von  $f$  an der Stelle  $\mathbf{a}$ .  $f$  ist (stetig) partiell differenzierbar, wenn  $f$  an jeder Stelle  $\mathbf{x} \in D$  (stetig) partiell differenzierbar ist.

## 10.5. Höhere Ableitungen

### Definition 10.11

Eine Funktion  $f : D \rightarrow \mathbb{R}$ ,  $D \subseteq \mathbb{R}^n$  offen heißt zweimal ( $k$ -mal) partiell differenzierbar, wenn alle zweiten partiellen Ableitungen

$$f_{x_i x_j} = \frac{\partial}{\partial x_j} \left( \frac{\partial f}{\partial x_i} \right) = \frac{\partial^2 f}{\partial x_j \partial x_i} = \partial_{x_j x_i} \quad 1 \leq i, j \leq n$$

(alle  $k$ -ten Ableitungen  $f_{x_{i_1}, \dots, x_{i_k}}$ ) existieren. Sind alle Ableitungen stetig, dann heißt  $f$  zweimal ( $k$ -mal) stetig partiell differenzierbar.

### Beispiel 10.12

$$f(x, y) = e^{x+2y} + 2x \sin(y)$$

$$\frac{\partial f}{\partial x} = e^{x+2y} + 2 \sin(y)$$

$$\frac{\partial f}{\partial y} = 2e^{x+2y} + 2x \cos(y)$$

$$\frac{\partial^2 f}{\partial x^2} = e^{x+2y}$$

$$\frac{\partial^2 f}{\partial x \partial y} = 2e^{x+2y} + 2 \cos(y)$$

$$\frac{\partial^2 f}{\partial y \partial x} = 2e^{x+2y} + 2 \cos(y)$$

$$\frac{\partial^2 f}{\partial y^2} = 4e^{x+2y} - 2x \sin(y)$$

### Definition 10.13

Wir bezeichnen die Menge aller auf einer offenen Menge  $D \subseteq \mathbb{R}^n$   $k$ -mal stetig partiell differenzierbaren Funktionen  $f : D \rightarrow \mathbb{R}$  mit  $\mathcal{C}^k(D, \mathbb{R})$  und nennen  $f \in \mathcal{C}^k(D, \mathbb{R})$  eine  $\mathcal{C}^k$ -Funktion. Eine  $\mathcal{C}^0$ -Funktion ist (nur) stetig.

### Satz 10.14

Für jede  $\mathcal{C}^2$ -Funktion  $f : D \rightarrow \mathbb{R}$ ,  $D \subseteq \mathbb{R}^n$  offen gilt:

$$\frac{\partial}{\partial x_i} \left( \frac{\partial f}{\partial x_j} \right) = \frac{\partial}{\partial x_j} \left( \frac{\partial f}{\partial x_i} \right) \quad 1 \leq i, j \leq n$$

ohne Beweis

### Definition 10.15

Für eine  $\mathcal{C}^k$ -Funktion mit  $k \geq 2$  heißt die (symmetrische) Matrix aller zweiten Ableitungen

$$\mathbf{H}_f(\mathbf{x}) := \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\mathbf{x}) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\mathbf{x}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_2 \partial x_2}(\mathbf{x}) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\mathbf{x}) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\mathbf{x}) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n}(\mathbf{x}) \end{pmatrix}$$

die Hesse-Matrix von  $f$  im Punkt  $\mathbf{x}$ .

### Beispiel 10.16

Für  $f(x, y) = e^{x+2y} + 2x \sin y$   
ist der Gradient gegeben durch

$$\nabla f(x, y) = \begin{pmatrix} e^{x+2y} + 2 \sin y \\ 2e^{x+2y} + 2x \cos y \end{pmatrix}$$

und die Hesse-Matrix durch

$$\mathbf{H}_f(x, y) = \begin{pmatrix} e^{x+2y} & 2e^{x+2y} + 2 \cos(y) \\ 2e^{x+2y} + 2 \cos(y) & 4e^{x+2y} - 2x \sin(y) \end{pmatrix}$$

## 10.6. Totale Ableitung

### Definition 10.17

Sei  $D \subseteq \mathbb{R}^n$  offen. Eine Funktion  $f : D \rightarrow \mathbb{R}$  heißt in  $\mathbf{x}_0 \in D$  total differenzierbar (oder linear approximierbar), wenn es einen Vektor  $\mathbf{a} \in \mathbb{R}^n$  gibt, so dass für  $\|\mathbf{x} - \mathbf{x}_0\| \rightarrow 0$  gilt:

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \mathbf{a}^T \cdot (\mathbf{x} - \mathbf{x}_0) + r(\mathbf{x} - \mathbf{x}_0)$$

mit  $\frac{r(\mathbf{x}-\mathbf{x}_0)}{\|\mathbf{x}-\mathbf{x}_0\|} \rightarrow 0$  für  $\|\mathbf{x} - \mathbf{x}_0\| \rightarrow 0$ .

### Satz 10.18

Ist  $f$  in  $\mathbf{x}_0 \in D$  total differenzierbar, d.h.  $f(\mathbf{x}) = f(\mathbf{x}_0) + \mathbf{a} \cdot (\mathbf{x} - \mathbf{x}_0) + r(\mathbf{x} - \mathbf{x}_0)$ , dann gilt:

1.  $f$  ist in  $\mathbf{x}_0$  stetig
2.  $\lim_{h \rightarrow 0} \frac{1}{h} [f(\mathbf{x}_0 + h \cdot \mathbf{v}) - f(\mathbf{x}_0)] = \mathbf{a}^T \mathbf{v}$  für alle  $\mathbf{v} \in \mathbb{R}^n$ ,  $\mathbf{v} \neq 0$
3.  $f$  ist partiell differenzierbar und  $\mathbf{a}$  ist eindeutig bestimmt als  $\mathbf{a} = \nabla f$

ohne Beweis

Anwendungen:

- Lineare Näherung komplizierter Funktionen
- Nullstellensuche (z.B. Gradienten-Verfahren)
- Fehlerabschätzung

### Satz 10.19

Jede  $\mathcal{C}^1$ -Funktion  $f : D \rightarrow \mathbb{R}$  ( $D \subseteq \mathbb{R}^n$  offen) ist auf  $D$  (d.h. in allen Punkten von  $D$ ) total differenzierbar.

ohne Beweis

## 10.7. Lokale Extremwerte

### Definition 10.20

1. Eine Menge  $D \subseteq \mathbb{R}^n$  heißt beschränkt, wenn es eine Konstante  $\lambda > 0$  gibt mit  $\|\mathbf{x}\| < \lambda$  für alle  $\mathbf{x} \in D$ .
2. Eine abgeschlossene und beschränkte Teilmenge des  $\mathbb{R}^n$  nennt man kompakt.

### Beispiel 10.21

Die Vollkugel um  $\mathbf{a}$  mit Radius  $\delta$

$$\mathbf{V}_\delta(\mathbf{a}) := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{a}\| \leq \delta\}$$

ist kompakt.

### Satz 10.22

1. *Satz vom Minimum und Maximum:*  
Jede auf einer kompakten Menge  $D \subseteq \mathbb{R}^n$  stetige Funktion nimmt auf  $D$  ein Minimum und ein Maximum an, d.h. es gibt  $\mathbf{a}, \mathbf{b} \in D$  mit  $f(\mathbf{b}) \leq f(\mathbf{x}) \leq f(\mathbf{a}) \quad \forall \mathbf{x} \in D$
2. *Satz von der gleichmäßigen Stetigkeit:*  
Jede auf einer kompakten Menge  $D \subseteq \mathbb{R}^n$  stetige Funktion ist dort gleichmäßig stetig, d.h. zu  $\varepsilon > 0$  gibt es ein  $\delta > 0$ , so dass für alle  $\mathbf{x}, \mathbf{y} \in D$  gilt

$$\|\mathbf{x} - \mathbf{y}\| < \delta \Rightarrow \|f(\mathbf{x}) - f(\mathbf{y})\| < \varepsilon$$

### Anmerkung 10.23

10.22 gilt nur, wenn  $D$  kompakt ist, da Extremwerte auf dem Rand (erfordert Abgeschlossenheit) oder im Unendlichen (bei fehlender Beschränktheit) angenommen werden können.

### Definition 10.24

Sei  $f : D \rightarrow \mathbb{R}$ ,  $D \subseteq \mathbb{R}^n$ . Ein Punkt  $\mathbf{a} \in D$  heißt lokales Maximum (bzw. Minimum) von  $f$ , wenn es eine  $\delta$ -Umgebung  $\mathbf{U}_\delta(\mathbf{a})$  gibt, so dass für alle  $\mathbf{x} \in \mathbf{U}_\delta(\mathbf{a}) \cap D$  gilt:

$$f(\mathbf{x}) \leq f(\mathbf{a}) \quad (\text{bzw. } f(\mathbf{x}) \geq f(\mathbf{a}))$$

Gelten die Ungleichungen für alle  $\mathbf{x} \in D$ , heißt  $\mathbf{a}$  globales Maximum (bzw. Minimum).

### Definition 10.25

Man nennt  $\mathbf{a} \in \mathbb{R}^n$  einen stationären (oder kritischen) Punkt einer differenzierbaren Funktion  $f$ , wenn  $\nabla f(\mathbf{a}) = 0$ .

Ein stationärer Punkt ist entweder eine Extremalstelle oder ein sogenannter Sattelpunkt.

### Definition 10.26 (Definitheit)

Sei  $\mathbf{A} \in \mathbb{R}^{n \times n}$ , dann heißt eine hermitesche Matrix  $\mathbf{A}$  (d.h. eine symmetrische Matrix im Reellen)

**positiv definit** falls  $\mathbf{x}^T \mathbf{A} \mathbf{x} > 0$  (alle Eigenwerte sind positiv)

**negativ definit** falls  $\mathbf{x}^T \mathbf{A} \mathbf{x} < 0$  (alle Eigenwerte sind negativ)

**positiv semidefinit** falls  $\mathbf{x}^T \mathbf{A} \mathbf{x} \geq 0$  (alle Eigenwerte sind positiv oder Null)

**negativ semidefinit** falls  $\mathbf{x}^T \mathbf{A} \mathbf{x} \leq 0$  (alle Eigenwerte sind negativ oder Null)

**indefinit** wenn keiner der anderen Fälle zutrifft (es gibt positive und negative Eigenwerte)

für alle  $\mathbf{x} \in \mathbb{R}^n$ ,  $\mathbf{x} \neq 0$ .

**Satz 10.27**

$D \subseteq \mathbb{R}^n$  sei offen,  $f : D \rightarrow \mathbb{R}$  eine  $\mathcal{C}^2$ -Funktion.

Für einen stationären Punkt  $\mathbf{x} \in D$  gilt:

Ist die Hesse-Matrix  $\mathbf{H}_f(\mathbf{x})$

1. positiv definit, dann hat  $f$  in  $\mathbf{x}$  ein lokales Minimum.
2. negativ definit, dann hat  $f$  in  $\mathbf{x}$  ein lokales Maximum.
3. indefinit, dann hat  $f$  in  $\mathbf{x}$  einen Sattelpunkt.
4. semidefinit, dann lässt sich nur mit der Hesse-Matrix keine Aussage treffen.

**Beispiel 10.28**

Untersuche

$$f(x, y) = x^3 - 12x + 5y^2$$

auf Extremstellen.

Stationäre Punkte:

$$\nabla f(x, y) = \begin{pmatrix} 3x^2 - 12 \\ 10y \end{pmatrix}$$

$f$  ist stationär, wenn  $y = 0$  und  $3x^2 - 12 = 0$

$\Rightarrow$  an den Stellen  $(-2, 0)$  und  $(2, 0)$

Hesse-Matrix:

$$\mathbf{H}_f(x, y) = \begin{pmatrix} 6x & 0 \\ 0 & 10 \end{pmatrix}$$

am ersten stationären Punkt:

$$\mathbf{H}_f(2, 0) = \begin{pmatrix} 12 & 0 \\ 0 & 10 \end{pmatrix}$$

$\Rightarrow$  beide Eigenwerte  $\lambda_1 = 12$ ,  $\lambda_2 = 10$  positiv

$\Rightarrow$  Minimum bei  $(2, 0)$

am zweiten stationären Punkt:

$$\mathbf{H}_f(-2, 0) = \begin{pmatrix} -12 & 0 \\ 0 & 10 \end{pmatrix}$$

$\Rightarrow$  Eigenwerte  $\lambda_1 = -12$  und  $\lambda_2 = 10$  haben unterschiedliches Vorzeichen

$\Rightarrow$  Sattelpunkt bei  $(-2, 0)$

## 10.8. Richtungsableitung

Die partiellen Ableitungen geben die Änderung bezüglich der Koordinatenachsen an. Wie finden wir Ableitungen in andere Richtungen?

### Definition 10.29

Existiert für eine Funktion  $f : D \rightarrow \mathbb{R}$ ,  $D \subseteq \mathbb{R}^n$  offen und einen Vektor  $\mathbf{v} \in \mathbb{R}^n$  an der Stelle  $\mathbf{x} \in \mathbb{R}^n$  der Grenzwert

$$\partial_{\mathbf{v}} f(\mathbf{x}) := \lim_{h \rightarrow 0} \frac{1}{h} [f(\mathbf{x} + h \cdot \mathbf{v}) - f(\mathbf{x})]$$

so nennen wir diesen die Ableitung von  $f$  an der Stelle  $\mathbf{x}$  längs  $\mathbf{v}$ . Ist  $\|\mathbf{v}\| = 1$ , dann heißt  $\partial_{\mathbf{v}} f(\mathbf{x})$  die Richtungsableitung (oder der Anstieg) von  $f$  an Stelle  $\mathbf{x}$  in Richtung  $\mathbf{v}$ .

### Satz 10.30

Sei  $D \subseteq \mathbb{R}^n$  offen,  $f$  auf  $D$  total differenzierbar und  $\mathbf{v} \in \mathbb{R}^n$  beliebig,  $\mathbf{v} \neq 0$  sowie  $\mathbf{x} \in D$ . Dann ist

$$\partial_{\mathbf{v}} f(\mathbf{x}) = \langle \nabla f(\mathbf{x}), \mathbf{v} \rangle = \nabla f(\mathbf{x})^T \cdot \mathbf{v} = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}) \cdot v_i.$$

ohne Beweis

### Anmerkung 10.31

Sei  $f : D \rightarrow \mathbb{R}$ ,  $D \subseteq \mathbb{R}^n$  eine  $\mathcal{C}^1$ -Funktion,  $\mathbf{v} \in \mathbb{R}^n$ ,  $\|\mathbf{v}\| = 1$  und  $\mathbf{x} \in D$ . Dann ist

$$\partial_{\mathbf{v}} f(\mathbf{x}) = \langle \nabla f(\mathbf{x}), \mathbf{v} \rangle = \|\nabla f(\mathbf{x})\| \cdot \|\mathbf{v}\| \cdot \cos \alpha = \|\nabla f(\mathbf{x})\| \cdot \cos \alpha$$

wobei  $\alpha$  der Winkel zwischen  $\nabla f$  und  $\mathbf{v}$  ist.

$\partial_{\mathbf{v}} f(\mathbf{x})$  ist maximal für  $\cos \alpha = 1 \iff \alpha = 0$

$\Rightarrow$  Gradient ist die Richtung des steilsten Anstiegs (und  $-\nabla f$  die Richtung des steilsten Abstiegs).

### Beispiel 10.32

Straße auf den Zuckerhut in Rio de Janeiro. Höhenprofil gegeben durch  $h(x, y) = -x^2 \cdot y^2 + 100$ . An der Stelle  $(-2, -3)^T$  ist die Richtung der Straße  $\mathbf{v} = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})^T$ . Wie groß ist die Steigung?

$$\nabla h(x, y) = \begin{pmatrix} -2xy^2 \\ -2x^2y \end{pmatrix}$$

$$\partial_{\mathbf{v}} f(-2, -3) = (36, \quad 24) \cdot \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{36 + 24}{\sqrt{2}} = \frac{60}{\sqrt{2}}$$

## 10.9. Kurven im $\mathbb{R}^n$

Oft ist es nötig z.B. Kurven im Raum in Abhängigkeit von der Zeit zu beschreiben.

**Definition 10.33**

Sei  $I \subseteq \mathbb{R}$  und  $x_i : I \rightarrow \mathbb{R} \quad (1 \leq i \leq n)$  eine stetig differenzierbare Funktion. Dann ist

$$\mathbf{x} : I \rightarrow \mathbb{R}^n \text{ mit } \mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix}, \quad t \in I$$

eine sogenannte  $\mathcal{C}^1$ -Kurve in Abhängigkeit vom Parameter  $t$ .

Grenzwert, Stetigkeit und Ableitung können leicht komponentenweise definiert werden:

$$\lim_{t \rightarrow t_0} \mathbf{x}(t) = \lim_{t \rightarrow t_0} \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \iff \lim_{t \rightarrow t_0} x_i(t) = c_i \quad (1 \leq i \leq n)$$

$\mathbf{x}$  ist auf  $I$  stetig, bzw. differenzierbar, wenn alle  $x_i$  auf  $I$  stetig bzw. differenzierbar sind.

$$\dot{\mathbf{x}}(t) = \frac{d}{dt} \mathbf{x}(t) = \lim_{h \rightarrow 0} \frac{1}{h} [\mathbf{x}(t+h) - \mathbf{x}(t)] = \begin{pmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \vdots \\ \dot{x}_n(t) \end{pmatrix}$$

$\dot{\mathbf{x}}(t)$  heißt Tangentialvektor von  $\mathbf{x}$  an der Stelle  $t \in I$ .

**Definition 10.34**

Sei  $t \in [a, b]$  und  $\mathbf{x}$  eine  $\mathcal{C}^1$ -Kurve, dann heißt

$$s(t) = \int_a^t \|\dot{\mathbf{x}}(t)\| dt$$

die Bogenlänge der Kurve.

**Anmerkung 10.35 (Anwendung)**

Bewegung eines Körpers im  $\mathbb{R}^n$  (meistens gilt  $n = 3$ ):

$\mathbf{x}(t)$  Position zur Zeit  $t$

$\dot{\mathbf{x}}(t)$  Geschwindigkeit des Körpers zur Zeit  $t$

$\ddot{\mathbf{x}}(t)$  Zur Zeit  $t$  auf den Körper wirkende Beschleunigung

$s(t)$  zurückgelegte Strecke

**Beispiel 10.36**

Sei  $\mathbf{x}(t) := (\cos(t), \sin(t))^T \quad a < t < b$

$$\|\mathbf{x}(t)\| = \sqrt{\cos^2(t) + \sin^2(t)} = 1$$

$\mathbf{x}(t)$  liegt auf der Einheitskreislinie  $C := \{(x_1, x_2)^T \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\}$ .

Geschwindigkeit:

$$\dot{\mathbf{x}}(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix}$$

$\Rightarrow$  Tangente an den Kreis  
Beschleunigungsvektor:

$$\ddot{\mathbf{x}}(t) = \begin{pmatrix} -\cos(t) \\ -\sin(t) \end{pmatrix} = - \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix} = -\mathbf{x}(t)$$

$\Rightarrow$  zeigt nach innen zum Mittelpunkt des Kreises  
Bahnlänge:

$$\begin{aligned} \|\dot{\mathbf{x}}(t)\| &= \sqrt{\sin^2(t) + \cos^2(t)} = 1 \\ s(t) &= \int_a^t \|\dot{\mathbf{x}}(t')\| dt' = \int_a^t 1 dt' = t - a \end{aligned}$$

Es gilt:

- $t < a$  Pfad entlang des Einheitskreises im Uhrzeigersinn
- $t > a$  Pfad entlang des Einheitskreises gegen den Uhrzeigersinn

## 10.10. Kettenregel

### Kettenregel für gewöhnliche Funktionen:

Seien  $I, J \subseteq \mathbb{R}$  offen,  $f : I \rightarrow \mathbb{R}$  und  $g : J \rightarrow \mathbb{R}$  mit  $g(x) \in I \quad \forall x \in J$  und  $f$  und  $g$  differenzierbar.

Dann ist die zusammengesetzte Funktion  $f \circ g : J \rightarrow \mathbb{R}$ ,  $(f \circ g)(x) = f(g(x))$  differenzierbar und es gilt:

$$(f \circ g)'(x) = f'(g(x)) \cdot g'(x)$$

### Satz 10.37

Für jede  $\mathcal{C}^1$ -Funktion  $f : D \rightarrow \mathbb{R}$ ,  $D \subseteq \mathbb{R}^n$  offen und für jedes Kurvenstück  $\mathbf{x} : ]a, b[ \rightarrow D$ ,  $]a, b[ \subseteq \mathbb{R}$  gilt:

$$\begin{aligned} \frac{d}{dt} f(\mathbf{x}(t)) &= \frac{d}{dt} f(x_1(t), x_2(t), \dots, x_n(t)) = f_{x_1}(\mathbf{x}(t)) \cdot \dot{x}_1(t) + \dots + f_{x_n}(\mathbf{x}(t)) \cdot \dot{x}_n(t) \\ &= \nabla f(\mathbf{x}(t))^T \cdot \dot{\mathbf{x}}(t) \end{aligned}$$

### Beweis:

Lineare Approximation in Kurvenpunkt  $\mathbf{x}_0 := \mathbf{x}(t_0)$

$$\begin{aligned} f(\mathbf{x}(t)) &= f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^T \cdot (\mathbf{x}(t) - \mathbf{x}_0) + r(\mathbf{x}(t), \mathbf{x}_0) \\ \Rightarrow \frac{f(\mathbf{x}(t)) - f(\mathbf{x}_0)}{t - t_0} &= \nabla f(\mathbf{x}_0)^T \cdot \frac{(\mathbf{x}(t) - \mathbf{x}_0)}{t - t_0} + \frac{r(\mathbf{x}(t), \mathbf{x}_0)}{t - t_0} \\ \Rightarrow \lim_{t \rightarrow t_0} \frac{r(\mathbf{x}(t), \mathbf{x}_0)}{t - t_0} &= \lim_{t \rightarrow t_0} \frac{\|\mathbf{x}(t) - \mathbf{x}_0\|}{t - t_0} \cdot \frac{r(\mathbf{x}(t), \mathbf{x}_0)}{\|\mathbf{x}(t) - \mathbf{x}_0\|} = 0 \end{aligned}$$

□



### Beispiel 10.38 (Polarkoordinaten im $\mathbb{R}^2$ )

$$x = r \cos \varphi, \quad y = r \sin \varphi$$

Sei  $f$  eine  $\mathcal{C}^2$ -Funktion auf  $D \in \mathbb{R}^2$  und  $D$  offen. Transformiere

$$f(x, y) : (x, y) \in D \rightarrow \mathbb{R}$$

und die ersten beiden Ableitungen nach

$$F(r, \varphi) := f(r \cos \varphi, r \sin \varphi)$$

Wir kürzen ab  $f_x = \frac{\partial f}{\partial x}(r \cos \varphi, r \sin \varphi)$ ,  $F_r = \frac{\partial F}{\partial r}(r, \varphi)$ ,  $F_\varphi = \frac{\partial F}{\partial \varphi}(r, \varphi)$ ,  $\dots$   
Für die ersten Ableitungen erhalten wir:

$$\begin{aligned} F_r &= \nabla f^T \cdot \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} = (f_x, f_y) \cdot \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} = f_x \cos \varphi + f_y \sin \varphi \\ F_\varphi &= \nabla f^T \cdot \begin{pmatrix} -r \sin \varphi \\ r \cos \varphi \end{pmatrix} = (f_x, f_y) \cdot \begin{pmatrix} -r \sin \varphi \\ r \cos \varphi \end{pmatrix} = -f_x r \sin \varphi + f_y r \cos \varphi = r \cdot (f_y \cos \varphi - f_x \sin \varphi) \end{aligned}$$

und für die zweiten Ableitungen:

$$\begin{aligned} F_{rr} &= \cos \varphi \cdot \nabla f_x^T \cdot \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} + \sin \varphi \cdot \nabla f_y^T \cdot \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \\ &= (\cos \varphi \cdot \nabla f_x + \sin \varphi \cdot \nabla f_y)^T \cdot \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \\ &= \begin{pmatrix} \cos \varphi \cdot f_{xx} + \sin \varphi \cdot f_{yx} \\ \cos \varphi \cdot f_{xy} + \sin \varphi \cdot f_{yy} \end{pmatrix}^T \cdot \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \\ &= f_{xx} \cos^2 \varphi + f_{yx} \sin \varphi \cos \varphi + f_{xy} \cos \varphi \sin \varphi + f_{yy} \sin^2 \varphi \\ &= f_{xx} \cos^2 \varphi + f_{xy} \cdot 2 \sin \varphi \cos \varphi + f_{yy} \sin^2 \varphi \\ &= f_{xx} \cos^2 \varphi + f_{xy} \cdot \sin(2\varphi) + f_{yy} \sin^2 \varphi \\ F_{r\varphi} &= \cos \varphi \cdot \nabla f_x^T \cdot \begin{pmatrix} -r \sin \varphi \\ r \cos \varphi \end{pmatrix} - f_x \sin \varphi + \sin \varphi \cdot \nabla f_y^T \cdot \begin{pmatrix} -r \sin \varphi \\ r \cos \varphi \end{pmatrix} + f_y \cos \varphi \\ &= (\cos \varphi \cdot \nabla f_x + \sin \varphi \cdot \nabla f_y)^T \cdot \begin{pmatrix} -r \sin \varphi \\ r \cos \varphi \end{pmatrix} - f_x \sin \varphi + f_y \cos \varphi \\ &= \begin{pmatrix} f_{xx} \cos \varphi + f_{yx} \sin \varphi \\ f_{xy} \cos \varphi + f_{yy} \sin \varphi \end{pmatrix}^T \cdot \begin{pmatrix} -r \sin \varphi \\ r \cos \varphi \end{pmatrix} - f_x \sin \varphi + f_y \cos \varphi \\ &= -r \cdot f_{xx} \sin \varphi \cos \varphi + r \cdot f_{xy} \cdot (\cos^2 \varphi - \sin^2 \varphi) + r \cdot f_{yy} \sin \varphi \cos \varphi - f_x \sin \varphi + f_y \cos \varphi \\ &= r \cdot \left[ \frac{1}{2} (f_{yy} - f_{xx}) \sin(2\varphi) + f_{xy} \cos(2\varphi) \right] - f_x \sin \varphi + f_y \cos \varphi \end{aligned}$$

$$\begin{aligned}
F_{\varphi\varphi}(r, \varphi) &= (-r \sin \varphi \cdot \nabla f_x + r \cos \varphi \cdot \nabla f_y)^T \cdot \begin{pmatrix} -r \sin \varphi \\ r \cos \varphi \end{pmatrix} - r \cos \varphi \cdot f_x - r \sin \varphi \cdot f_y \\
&= \begin{pmatrix} -r \cdot f_{xx} \sin \varphi + r \cdot f_{yx} \cos \varphi \\ -r \cdot f_{xy} \sin \varphi + r \cdot f_{yy} \cos \varphi \end{pmatrix}^T \cdot \begin{pmatrix} -r \sin \varphi \\ r \cos \varphi \end{pmatrix} - r \cdot f_x \cos \varphi - r \cdot f_y \sin \varphi \\
&= r^2 \cdot f_{xx} \sin^2 \varphi - r^2 \cdot f_{xy} \cdot 2 \cos \varphi \sin \varphi + r^2 \cdot f_{yy} \cos^2 \varphi - f_x r \cos \varphi - f_y r \sin \varphi \\
&= r^2 \cdot [f_{xx} \sin^2 \varphi - f_{xy} \sin(2\varphi) + f_{yy} \cos^2 \varphi] - r \cdot [f_x \cos \varphi + f_y \sin \varphi]
\end{aligned}$$

Alternativer Lösungsweg für zweite Ableitungen mit Hesse-Matrix  $\mathbf{H}_f$  als Ableitung des Gradienten (zweite Ableitung von  $f$ ):

$$\begin{aligned}
F_{rr}(r, \varphi) &= \left[ \mathbf{H}_f^T \cdot \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \right]^T \cdot \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} = \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}^T \cdot \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} \cdot \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \\
&= \begin{pmatrix} f_{xx} \cos \varphi + f_{xy} \sin \varphi \\ f_{yx} \cos \varphi + f_{yy} \sin \varphi \end{pmatrix}^T \cdot \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \\
&= f_{xx} \cos^2 \varphi + f_{xy} \sin \varphi \cos \varphi + f_{xy} \sin \varphi \cos \varphi + f_{yy} \sin^2 \varphi \\
&= f_{xx} \cos^2 \varphi + f_{xy} \sin 2\varphi + f_{yy} \sin^2 \varphi \\
F_{r\varphi}(r, \varphi) &= \left[ \mathbf{H}_f^T \cdot \begin{pmatrix} -r \sin \varphi \\ r \cos \varphi \end{pmatrix} \right]^T \cdot \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} + \nabla f^T \cdot \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix} \\
&= \begin{pmatrix} -r \sin \varphi \\ r \cos \varphi \end{pmatrix}^T \cdot \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} \cdot \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} + (f_x, f_y) \cdot \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix} \\
&= \begin{pmatrix} -r f_{xx} \sin \varphi + r f_{xy} \cos \varphi \\ -r f_{yx} \sin \varphi + r f_{yy} \cos \varphi \end{pmatrix}^T \cdot \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} - f_x \sin \varphi + f_y \cos \varphi \\
&= r \cdot [-f_{xx} \sin \varphi \cos \varphi + f_{xy} \cos^2 \varphi - f_{xy} \sin^2 \varphi + f_{yy} \sin \varphi \cos \varphi] - f_x \sin \varphi + f_y \cos \varphi \\
&= r \cdot \left[ \frac{1}{2} (f_{yy} - f_{xx}) \sin 2\varphi + f_{xy} \cos 2\varphi \right] - f_x \sin \varphi + f_y \cos \varphi \\
F_{\varphi\varphi}(r, \varphi) &= \left[ \mathbf{H}_f^T \cdot \begin{pmatrix} -r \sin \varphi \\ r \cos \varphi \end{pmatrix} \right]^T \cdot \begin{pmatrix} -r \sin \varphi \\ r \cos \varphi \end{pmatrix} + \nabla f^T \cdot \begin{pmatrix} -r \cos \varphi \\ -r \sin \varphi \end{pmatrix} \\
&= \begin{pmatrix} -r \sin \varphi \\ r \cos \varphi \end{pmatrix}^T \cdot \begin{pmatrix} f_{xx} & f_{yx} \\ f_{xy} & f_{yy} \end{pmatrix} \cdot \begin{pmatrix} -r \sin \varphi \\ r \cos \varphi \end{pmatrix} + (f_x, f_y) \cdot \begin{pmatrix} -r \cos \varphi \\ -r \sin \varphi \end{pmatrix} \\
&= \begin{pmatrix} -r f_{xx} \sin \varphi + r f_{xy} \cos \varphi \\ -r f_{yx} \sin \varphi + r f_{yy} \cos \varphi \end{pmatrix}^T \cdot \begin{pmatrix} -r \sin \varphi \\ r \cos \varphi \end{pmatrix} - r \cdot f_x \cos \varphi - r \cdot f_y \sin \varphi \\
&= r^2 \cdot [f_{xx} \sin^2 \varphi - f_{xy} \sin \varphi \cos \varphi - f_{xy} \sin \varphi \cos \varphi + f_{yy} \cos^2 \varphi] - r \cdot [f_x \sin \varphi + f_y \cos \varphi] \\
&= r^2 \cdot [f_{xx} \sin^2 \varphi - f_{xy} \sin 2\varphi + f_{yy} \cos^2 \varphi] - r \cdot [f_x \sin \varphi + f_y \cos \varphi]
\end{aligned}$$

Nach Auflösen z.B. nach  $f_x$  und  $f_y$  lassen sich auch die partiellen Ableitungen nach  $x$  und  $y$  durch die partiellen Ableitungen nach  $r$  und  $\varphi$  darstellen:

$$\begin{aligned}
f_x &= F_r \cos \varphi - \frac{1}{r} F_\varphi \sin \varphi \\
f_y &= F_r \sin \varphi + \frac{1}{r} F_\varphi \cos \varphi
\end{aligned}$$

### 10.11. Approximation höherer Ordnung (Taylor-Formel)

Für  $(n+1)$ -mal stetig differenzierbare Funktionen reeller Zahlen  $f: I \rightarrow \mathbb{R}$ ,  $I \subseteq \mathbb{R}$  gilt mit  $x, x_0 \in I$ :

$$f(x) = \sum_{j=0}^n \frac{f^{(j)}(x_0)}{j!} (x - x_0)^j + R_n(x)$$

mit dem Restglied

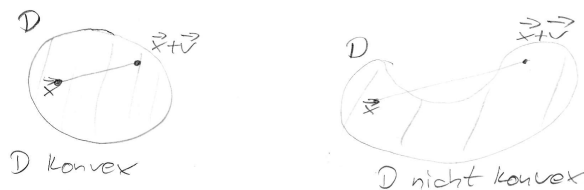
$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \cdot (x - x_0)^{n+1} \text{ für ein } \xi \in [x_0, x]$$

Wir möchten nun auch für eine Funktion  $f(\mathbf{x})$  mit  $\mathbf{x} \in \mathbb{R}^n$  eine besser als lineare Approximation finden.

#### Definition 10.39

$D \subseteq \mathbb{R}^n$  heißt konvex, wenn gilt:

$$\mathbf{x}, \mathbf{y} \in D \Rightarrow \mathbf{x} + t \cdot (\mathbf{y} - \mathbf{x}) \in D \quad \forall t \in [0, 1]$$



Führen wir die Hilfsfunktionen  $g(t) := \mathbf{x} + t \cdot \mathbf{v}$  und  $h(t) := f(g(t))$  ein, können wir die Taylorformel von oben verwenden

$$h(t) = \sum_{j=0}^n \frac{h^{(j)}(0)}{j!} \cdot t^j + R_n(t)$$

mit  $R_n(t) = \frac{h^{(n+1)}(\xi)}{(n+1)!} \cdot t^{n+1}$  für  $\xi \in [0, 1]$

$$\begin{aligned} h'(t) &= \frac{d}{dt} f(g(t)) = \nabla f(g(t))^T \cdot \frac{d}{dt} g(t) \\ &= \nabla f(\mathbf{x} + t \cdot \mathbf{v})^T \cdot \mathbf{v} \\ &= \langle \nabla f(\mathbf{x} + t \cdot \mathbf{v}), \mathbf{v} \rangle \\ \Rightarrow h'(0) &= \langle \nabla f(\mathbf{x}), \mathbf{v} \rangle \end{aligned}$$

$$\begin{aligned}
h''(t) &= \frac{d}{dt} h'(t) = \frac{d}{dt} \left[ \nabla f(g(t))^T \cdot \mathbf{v} \right] \\
&= \left[ \mathbf{H}_f^T(g(t)) \cdot \mathbf{v} \right]^T \cdot \frac{d}{dt} g(t) = \\
&= \left[ \mathbf{H}_f^T(\mathbf{x} + t \cdot \mathbf{v}) \cdot \mathbf{v} \right]^T \cdot \mathbf{v} = \\
&= \mathbf{v}^T \mathbf{H}_f(\mathbf{x} + t \cdot \mathbf{v}) \cdot \mathbf{v} \\
&= \langle \mathbf{v}, \mathbf{H}_f(\mathbf{x} + t \cdot \mathbf{v}) \cdot \mathbf{v} \rangle \\
\Rightarrow h''(0) &= \langle \mathbf{v}, \mathbf{H}_f(\mathbf{x}) \cdot \mathbf{v} \rangle
\end{aligned}$$

**Satz 10.40**

Sei  $D \subseteq \mathbb{R}^n$  offen und konvex,  $f \in \mathcal{C}^{n+1}(D, \mathbb{R})$  und  $\mathbf{x} \in D$ , dann gilt mit  $\mathbf{x} + \mathbf{v} \in D$ :

$$f(\mathbf{x} + \mathbf{v}) = f(\mathbf{x}) + \sum_{j=1}^n \frac{1}{j!} \partial_{\mathbf{v}}^j f(\mathbf{x}) + R_n(\mathbf{x}, \mathbf{v})$$

mit  $R_n(\mathbf{x}, \mathbf{v}) = \frac{1}{(n+1)!} \partial_{\mathbf{v}}^{n+1} f(\mathbf{x} + \xi_n \mathbf{v})$  und  $\xi_n \in [0, 1]$ .

Insbesondere gilt für  $n = 2$

$$f(\mathbf{x}) \approx f(\mathbf{x}_0) + \langle \nabla f(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle + \frac{1}{2} \langle \mathbf{x} - \mathbf{x}_0, \mathbf{H}_f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \rangle$$

**Beweis:**

folgt aus obiger Herleitung mit  $t = 1$  und  $\mathbf{v} = \mathbf{x} - \mathbf{x}_0$ . □

*Wichtig:* Wenn (und wirklich nur wenn) die dritte Ableitung von  $f$  beschränkt ist, dann hat die quadratische Approximation einen Fehler, der mindestens mit  $\|\mathbf{x} - \mathbf{x}_0\|^3$  gegen Null geht für  $\mathbf{x} \rightarrow \mathbf{x}_0$ .

## 10.12. Optimierung unter Nebenbedingungen

- Normalfall im Alltag:
  - Finde besten Fußballer, der für 2 Millionen Euro verfügbar ist.
  - Baue sparsamsten Motor, der 100'000 km läuft und 10'000 Euro kostet
  - Finde Körper mit Volumen  $V = V_0$  und minimaler Oberfläche
- Nebenbedingungen mathematisch:

$$g(\mathbf{x}) = 0, \quad \mathbf{x} \in \mathbb{R}^n$$

$$\text{z.B. Kosten} - 2 \text{ Millionen Euro} = 0$$

$$\text{Leistung} - 100'000 \text{ km} = 0$$

$$V - V_0 = 0$$

- Problem:  
Finde  $\mathbf{x} \in D$  mit  $f(\mathbf{x})$  maximal/minimal unter Bedingung  $g(\mathbf{x}) = 0$

**Beispiel 10.41**

Schneide aus einem Stamm mit Radius  $r$  einen Balken mit rechteckiger Querschnittsfläche mit Dicke  $h$  und Breite  $b$ , der maximales Widerstandsmoment hat. Dieses ist

$$W(b, h) = \frac{1}{6} \cdot b \cdot h^2$$

Nebenbedingung (Satz des Pythagoras und Satz des Thales)

$$b^2 + h^2 = (2r)^2$$

$$g(b, h) = b^2 + h^2 - 4r^2 = 0$$

**Direkter Lösungsansatz: Auflösen der Nebenbedingungen und einsetzen**

$$b^2 + h^2 - 4r^2 = 0 \iff h^2 = 4r^2 - b^2$$

$$\Rightarrow W(b) = \frac{1}{6} \cdot (4r^2 \cdot b - b^3)$$

$$\frac{\partial W}{\partial b} = \frac{1}{6} \cdot (4r^2 - 3b^2) = 0$$

$$\Rightarrow 3b^2 = 4r^2$$

$$\Rightarrow b = \pm \sqrt{\frac{4}{3}} \cdot r$$

$$\frac{\partial^2 W}{\partial b^2} = -b \quad \text{negativ} \Rightarrow \text{Maximum bei positivem } b$$

$$\Rightarrow b = \frac{2}{\sqrt{3}} \cdot r$$

$$\Rightarrow h^2 = 4r^2 - \frac{4}{3}r^2 = 4 \cdot \frac{2}{3}r^2 = \frac{8}{3}r^2$$

$$\Rightarrow h = \sqrt{\frac{8}{3}} \cdot r$$

**Lagrange'sches Multiplikatorverfahren**

Der Gradient von  $f(\mathbf{x})$  muss im Optimum senkrecht auf der Linie (Hyperfläche)  $g(\mathbf{x}) = 0$  stehen, auf der die Nebenbedingung erfüllt ist, da sonst  $f(\mathbf{x})$  entlang der Linie noch weiter zu- oder abnimmt.

**Satz 10.42**

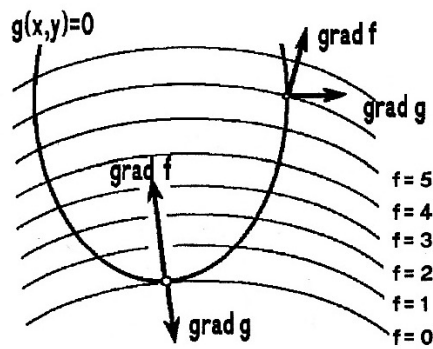
Sei  $D \subseteq \mathbb{R}^n$  offen und  $g : D \rightarrow \mathbb{R}$  eine  $\mathcal{C}^1$ -Funktion, dann definiert  $g(\mathbf{x}) = c = \text{const}$  ( $c \in \mathbb{R}$ ) eine Hyperfläche in  $D$ . Der Gradientenvektor  $\nabla g$  steht in jedem Punkt  $\mathbf{x} \in D$  mit  $g(\mathbf{x}) = c$  senkrecht auf dieser Hyperfläche

**Beweis:**

Für jede parametrisierte Kurve  $\mathbf{x} : I \rightarrow \mathbb{R}^n$ ,  $I \subseteq \mathbb{R}$  auf dieser Hyperfläche gilt:  $g(\mathbf{x}(t)) = c$ . Mit Kettenregel:

$$\nabla g(\mathbf{x}(t))^T \cdot \dot{\mathbf{x}}(t) = 0$$

Da  $\dot{\mathbf{x}}(t)$  Tangentialvektor  $\Rightarrow$  Behauptung □



aus Meyberg, Vachena: Höhere Mathematik 1

Für unser Optimierungsproblem heißt das, dass  $\nabla f(\mathbf{x})$  und  $\nabla g(\mathbf{x})$  parallel sein müssen.

$$\Rightarrow \nabla f(\mathbf{x}) = -\lambda \cdot \nabla g(\mathbf{x}), \quad \lambda \in \mathbb{R}$$

$$\Rightarrow \nabla f(\mathbf{x}) + \lambda \nabla g(\mathbf{x}) = \mathbf{0}.$$

Idee: Definiere Hilfsfunktion  $h(\mathbf{x}, \lambda)$ ,  $\mathbf{x} \in \mathbb{R}^n$ ,  $\lambda \in \mathbb{R}$  durch

$$h(\mathbf{x}, \lambda) := f(\mathbf{x}) + \lambda g(\mathbf{x})$$

$\lambda$  nennt man Lagrangeschen Multiplikator.

Setze  $\nabla h(\mathbf{x}, \lambda) = \mathbf{0}$

$$\begin{aligned} \Rightarrow \begin{pmatrix} \frac{\partial h}{\partial x_1}(\mathbf{x}, \lambda) \\ \vdots \\ \frac{\partial h}{\partial x_n}(\mathbf{x}, \lambda) \\ \frac{\partial h}{\partial \lambda}(\mathbf{x}, \lambda) \end{pmatrix} &= \mathbf{0} \\ \Leftrightarrow \begin{pmatrix} \frac{\partial f}{\partial x_1}(\mathbf{x}) + \lambda \cdot \frac{\partial g}{\partial x_1}(\mathbf{x}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\mathbf{x}) + \lambda \cdot \frac{\partial g}{\partial x_n}(\mathbf{x}) \\ g(\mathbf{x}) \end{pmatrix} &= \mathbf{0} \\ \Leftrightarrow \nabla f(\mathbf{x}) + \lambda \cdot \nabla g(\mathbf{x}) &= \mathbf{0} \\ g(\mathbf{x}) &= 0 \end{aligned}$$

An dieser Stelle ist also ein stationärer Punkt unter der Nebenbedingung  $g(\mathbf{x}) = 0$ . Das geht auch für mehrere Nebenbedingungen:

$$g_i(\mathbf{x}) = 0 \quad (1 \leq i \leq m)$$

Die Hilfsfunktion ist dann

$$h(\mathbf{x}, \lambda_1, \dots, \lambda_m) = f(\mathbf{x}) + \sum_{i=1}^m \lambda_i g_i(\mathbf{x})$$

am stationären Punkt gilt dann

$$\nabla f(\mathbf{x}) + \sum_{i=1}^m \lambda_i \nabla g_i(\mathbf{x}) = \mathbf{0}$$

$$g_1(\mathbf{x}) = 0$$

$\vdots$

$$g_m(\mathbf{x}) = 0$$

$\Rightarrow n + m$  Gleichungen für  $n + m$  Unbekannte

### Beispiel 10.43

Hilfsfunktion:

$$f(h, b, \lambda) = \frac{1}{6} \cdot b \cdot h^2 + \lambda \cdot (b^2 + h^2 - 4r^2)$$

$$\nabla f = \begin{pmatrix} \frac{1}{3} \cdot b \cdot h + \lambda \cdot 2h \\ \frac{1}{6} h^2 + \lambda \cdot 2b \\ b^2 + h^2 - 4r^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$1. \text{ Gleichung: } \Rightarrow \lambda = -\frac{1}{6} \cdot b$$

in 2. Gleichung:

$$\frac{1}{6} h^2 - \frac{2}{6} b^2 = 0$$
$$\Rightarrow h^2 = 2b^2$$

in 3. Gleichung:

$$b^2 + 2b^2 = 4r^2$$

$$\Rightarrow b = \sqrt{\frac{4}{3}} \cdot r$$

$$h^2 = \frac{2 \cdot 4}{3} r^2$$

$$h = \sqrt{\frac{8}{3}} \cdot r$$

$$W_{max} = \frac{1}{6} \cdot \frac{8}{3} \cdot r^2 \cdot \frac{2}{\sqrt{3}} \cdot r = \frac{8}{9 \cdot \sqrt{3}} \cdot r^3$$

### Anmerkung 10.44

- Vorteile des Lagrange'schen Multiplikatorverfahrens:
  - Funktioniert auch, wenn sich die Nebenbedingung nicht nach einer Variablen auflösen lässt.
  - Funktioniert genauso für viele Nebenbedingungen.
- Die konkreten Werte der  $\lambda_i$  sind nicht von Interesse. Es ist deshalb von Vorteil, sie gleich beim Lösen zu eliminieren.

- Für jeden stationären Punkt muss noch geprüft werden, ob es sich um das gewünschte Maximum/Minimum handelt.

## 10.13. Vektorwertige Funktionen

### 10.13.1. Differentiation und Jacobi-Matrix

Wir betrachten jetzt Funktionen, bei denen sowohl Argument als auch Ergebnis Vektoren sind.

$$\mathbf{f} : D \rightarrow \mathbb{R}^m, \quad D \subseteq \mathbb{R}^n, \quad \mathbf{x} \in D$$

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} f_1(x_1, x_2, \dots, x_n) \\ f_2(x_1, x_2, \dots, x_n) \\ \vdots \\ f_m(x_1, x_2, \dots, x_n) \end{pmatrix}$$

#### Definition 10.45

Seien  $\mathbf{f} : D \rightarrow \mathbb{R}^m$ ,  $D \subseteq \mathbb{R}^n$  offen und  $\mathbf{x}, \mathbf{x}_0 \in D$  dann gilt:

1. Der Grenzwert und die partielle Differentiation sind komponentenweise definiert:

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{f}(\mathbf{x}) := \begin{pmatrix} \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f_1(\mathbf{x}) \\ \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f_2(\mathbf{x}) \\ \vdots \\ \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f_m(\mathbf{x}) \end{pmatrix}$$

$$\frac{\partial \mathbf{f}}{\partial x_i}(\mathbf{x}) := \begin{pmatrix} \frac{\partial f_1}{\partial x_i}(\mathbf{x}) \\ \frac{\partial f_2}{\partial x_i}(\mathbf{x}) \\ \vdots \\ \frac{\partial f_m}{\partial x_i}(\mathbf{x}) \end{pmatrix}$$

2.  $\mathbf{f}$  ist genau dann stetig, partiell differenzierbar oder eine  $\mathcal{C}^r$ -Funktion, wenn sämtliche Komponentenfunktionen  $f_i$  ( $1 \leq i \leq m$ ) stetig, partiell differenzierbar bzw.  $\mathcal{C}^r$ -Funktionen sind.
3.  $\mathbf{f}$  heißt in  $\mathbf{x}_0 \in D$  total differenzierbar oder linear approximierbar, wenn es eine  $m \times n$ -Matrix  $\mathbf{A}$  und eine  $\delta$ -Umgebung  $U_\delta(\mathbf{x}_0) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{x}_0\| < \delta\}$  von  $\mathbf{x}_0$  gibt, so dass für alle  $\mathbf{x} \in U_\delta(\mathbf{x}_0)$  gilt:

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}_0) + \mathbf{A} \cdot (\mathbf{x} - \mathbf{x}_0) + \mathbf{r}(\|\mathbf{x} - \mathbf{x}_0\|)$$

mit

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{\mathbf{r}(\|\mathbf{x} - \mathbf{x}_0\|)}{\|\mathbf{x} - \mathbf{x}_0\|} = \mathbf{0}$$



**Definition 10.46**

Die Matrix

$$\mathbf{J}_f(\mathbf{x}) := \begin{pmatrix} \nabla f_1(\mathbf{x})^T \\ \nabla f_2(\mathbf{x})^T \\ \vdots \\ \nabla f_m(\mathbf{x})^T \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}) & \dots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}) \end{pmatrix}$$

heißt Jacobi-Matrix

**Satz 10.47**

Sei  $f : D \rightarrow \mathbb{R}^m$ ,  $D \subseteq \mathbb{R}^n$  eine  $\mathcal{C}^1$ -Abbildung, dann gilt:

1.  $f$  ist für alle  $\mathbf{x} \in D$  linear approximierbar. Die Matrix  $\mathbf{A}$  ist eindeutig bestimmt durch  $\mathbf{A} = \mathbf{J}_f(\mathbf{x}_0)$
2. (Schränkensatz): Ist  $D$  konvex und  $\|\nabla f_i(\mathbf{x})\| \leq M \in \mathbb{R}$  für alle  $\mathbf{x} \in D$  und  $1 \leq i \leq m$ , dann gilt  $\|f(\mathbf{x}) - f(\mathbf{x}_0)\| \leq M\|\mathbf{x} - \mathbf{x}_0\| \quad \forall \mathbf{x}, \mathbf{x}_0 \in D$

**Beweis:**

1. Jede Komponente  $\mathcal{C}^1$ -Funktion  $\Rightarrow$  jede Komponente linear approximierbar mit

$$\begin{aligned} \mathbf{a}_i &= \left( \frac{\partial f_i}{\partial x_1} \quad \dots \quad \frac{\partial f_i}{\partial x_n} \right)^T \\ \Rightarrow x_i &= (\mathbf{x}_0)_i + \mathbf{a}_i^T \cdot (\mathbf{x} - \mathbf{x}_0) = (\mathbf{x}_0)_i + \langle \mathbf{a}_i, \mathbf{x} - \mathbf{x}_0 \rangle \\ \Rightarrow \mathbf{A} &= \begin{pmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_m^T \end{pmatrix} = \mathbf{J}_f(\mathbf{x}_0) \end{aligned}$$

2. ohne Beweis

□

**Beispiel 10.48**

- $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $f(\mathbf{x}) = \mathbf{A} \cdot \mathbf{x}$ ,  $\mathbf{A} \in \mathbb{R}^{m \times n}$ , ist total differenzierbar

$$\mathbf{J}_f(\mathbf{x}) = \mathbf{A} \quad \forall \mathbf{x} \in \mathbb{R}^n$$

- Polarkoordinaten

$$\begin{aligned} \mathbf{x}(r, \varphi) &= \begin{pmatrix} r \cos \varphi \\ r \sin \varphi \end{pmatrix} \\ \mathbf{J}_x(r, \varphi) &= \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix} \end{aligned}$$

**Satz 10.49**

Für  $\mathbf{x} \in \mathbb{R}^n$ ,  $f(\mathbf{x}), g(\mathbf{x}) \in \mathbb{R}^m$ ,  $\lambda(\mathbf{x}) \in \mathbb{R}$  und  $\alpha, \beta \in \mathbb{R}$  gilt, wenn  $f(\mathbf{x})$ ,  $g(\mathbf{x})$  und  $\lambda(\mathbf{x})$   $\mathcal{C}^1$ -Funktionen sind:

1.  $\mathbf{J}_{\alpha \mathbf{f} + \beta \mathbf{g}}(\mathbf{x}) = \alpha \cdot \mathbf{J}_{\mathbf{f}}(\mathbf{x}) + \beta \cdot \mathbf{J}_{\mathbf{g}}(\mathbf{x})$  (Linearität)
2.  $\mathbf{J}_{\lambda \cdot \mathbf{f}}(\mathbf{x}) = \lambda(\mathbf{x}) \cdot \mathbf{J}_{\mathbf{f}}(\mathbf{x}) + \mathbf{f}(\mathbf{x}) \cdot (\nabla \lambda(\mathbf{x}))^T$  (Produktregel für Multiplikation mit skalarer Funktion)
3.  $\mathbf{J}_{\langle \mathbf{f}, \mathbf{g} \rangle}(\mathbf{x}) = \mathbf{J}_{\mathbf{f}^T \mathbf{g}}(\mathbf{x}) = \mathbf{f}(\mathbf{x})^T \cdot \mathbf{J}_{\mathbf{g}}(\mathbf{x}) + \mathbf{g}(\mathbf{x})^T \cdot \mathbf{J}_{\mathbf{f}}(\mathbf{x})$  (Produktregel für Skalarprodukt)
4. für  $m = 3$ :  
 $\mathbf{J}_{\mathbf{f} \times \mathbf{g}}(\mathbf{x}) = \mathbf{f}(\mathbf{x}) \times \mathbf{J}_{\mathbf{g}}(\mathbf{x}) - \mathbf{g}(\mathbf{x}) \times \mathbf{J}_{\mathbf{f}}(\mathbf{x})$  (Produktregel für Kreuzprodukt)

dabei ist für  $\mathbf{v} \in \mathbb{R}^3$ ,  $\mathbf{A} \in \mathbb{R}^{3 \times n}$

$$\mathbf{v} \times \mathbf{A} = \begin{pmatrix} v_2 \cdot a_{31} - v_3 \cdot a_{21} & \dots & v_2 \cdot a_{3n} - v_3 \cdot a_{2n} \\ v_3 \cdot a_{11} - v_1 \cdot a_{31} & \dots & v_3 \cdot a_{1n} - v_1 \cdot a_{3n} \\ v_1 \cdot a_{21} - v_2 \cdot a_{11} & \dots & v_1 \cdot a_{2n} - v_2 \cdot a_{1n} \end{pmatrix} \quad \text{bzw.} \quad (\mathbf{v} \times \mathbf{A})^{(i)} = \mathbf{v} \times \mathbf{A}^{(i)}$$

(also spaltenweise definiert)

ohne Beweis

### Beispiel 10.50

1. Quasilineare Abbildung:  $\mathbf{f}(\mathbf{x}) = \lambda(\mathbf{x}) \cdot \mathbf{A} \cdot \mathbf{x}$  mit  $\lambda : D \rightarrow \mathbb{R}$ ,  $D \subseteq \mathbb{R}^n$ ,  $\mathcal{C}^1$ -Funktion und  $\mathbf{A} \in \mathbb{R}^{m \times n}$

$$\mathbf{J}_{\mathbf{f}}(\mathbf{x}) = \lambda(\mathbf{x}) \cdot \mathbf{A} + (\mathbf{A} \cdot \mathbf{x}) \cdot (\nabla \lambda(\mathbf{x}))^T$$

2. Konkretes Beispiel für quasilineare Abbildung:  $\mathbf{f}(\mathbf{x}) = \mathbf{A} \cdot \frac{\mathbf{x}}{\|\mathbf{x}\|}$  für  $\mathbf{A} \in \mathbb{R}^{m \times n}$ ,  $\mathbf{x} \in \mathbb{R}^n$ ,  $\|\mathbf{x}\| \neq 0$

$$\mathbf{f}(\mathbf{x}) = \mathbf{A} \cdot \frac{\mathbf{x}}{\|\mathbf{x}\|} = \frac{1}{\|\mathbf{x}\|} \cdot \mathbf{A} \cdot \mathbf{x}$$

$$\mathbf{J}_{\mathbf{f}}(\mathbf{x}) = \frac{1}{\|\mathbf{x}\|} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{x} \cdot \nabla \left( \frac{1}{\|\mathbf{x}\|} \right)$$

Nebenrechnung:

$$\frac{\partial}{\partial x_i} \frac{1}{\|\mathbf{x}\|} = -\frac{1}{\|\mathbf{x}\|^2} \cdot \frac{1}{2} \frac{1}{\sqrt{\sum_i x_i^2}} \cdot 2x_i = -\frac{x_i}{\|\mathbf{x}\|^3}$$

$$\begin{aligned} \mathbf{J}_{\mathbf{f}}(\mathbf{x}) &= \frac{1}{\|\mathbf{x}\|} \cdot \mathbf{A} + \mathbf{A} \cdot \mathbf{x} \cdot \left( -\frac{x_1}{\|\mathbf{x}\|^3} \quad \dots \quad -\frac{x_n}{\|\mathbf{x}\|^3} \right) = \frac{1}{\|\mathbf{x}\|} \cdot \mathbf{A} - \frac{1}{\|\mathbf{x}\|^3} \mathbf{A} \cdot \mathbf{x} \cdot \mathbf{x}^T \\ &= \frac{\mathbf{A}}{\|\mathbf{x}\|} \cdot \left( \mathbf{I}_n - \frac{\mathbf{x}}{\|\mathbf{x}\|} \cdot \frac{\mathbf{x}^T}{\|\mathbf{x}\|} \right) \end{aligned}$$

a)  $\mathbf{A} = \mathbf{I}_2 \in \mathbb{R}^{2 \times 2}$ :

$$\begin{aligned} \mathbf{J}_{\mathbf{f}}(\mathbf{x}) &= \frac{\mathbf{I}_2}{\|\mathbf{x}\|} \left[ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{1}{\|\mathbf{x}\|^2} \cdot \begin{pmatrix} x_1 x_1 & x_1 x_2 \\ x_2 x_1 & x_2 x_2 \end{pmatrix} \right] \\ &= \frac{1}{\|\mathbf{x}\|} \begin{pmatrix} 1 - \frac{x_1^2}{\|\mathbf{x}\|^2} & -\frac{x_1 x_2}{\|\mathbf{x}\|^2} \\ -\frac{x_1 x_2}{\|\mathbf{x}\|^2} & 1 - \frac{x_2^2}{\|\mathbf{x}\|^2} \end{pmatrix} \end{aligned}$$

$$\text{b) } \mathbf{A} = \begin{pmatrix} 1 & 4 \\ 3 & 0 \\ 2 & 1 \end{pmatrix}:$$

$$\begin{aligned} \mathbf{J}_{\mathbf{f}}(\mathbf{x}) &= \frac{1}{\|\mathbf{x}\|} \begin{pmatrix} 1 & 4 \\ 3 & 0 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 - \frac{x_1^2}{\|\mathbf{x}\|^2} & -\frac{x_1 x_2}{\|\mathbf{x}\|^2} \\ -\frac{x_1 x_2}{\|\mathbf{x}\|^2} & 1 - \frac{x_2^2}{\|\mathbf{x}\|^2} \end{pmatrix} \\ &= \frac{1}{\|\mathbf{x}\|^3} \begin{pmatrix} \|\mathbf{x}\|^2 - x_1^2 - 4x_1 x_2 & -x_1 x_2 + 4\|\mathbf{x}\|^2 + 4x_2^2 \\ 3\|\mathbf{x}\|^2 - 3x_1^2 & -3x_1 x_2 \\ 2\|\mathbf{x}\|^2 - 2x_1^2 - x_1 x_2 & -2x_1 x_2 + \|\mathbf{x}\|^2 - x_2^2 \end{pmatrix} \end{aligned}$$

### 10.13.2. Kettenregel

#### Satz 10.51

Seien  $\mathbf{g} : D \rightarrow \mathbb{R}^m$ ,  $D \subseteq \mathbb{R}^n$  offen und  $\mathbf{f} : G \rightarrow \mathbb{R}^q$ ,  $G \subseteq \mathbb{R}^m$  in  $\mathbf{x}_0$  bzw.  $\mathbf{g}(\mathbf{x}_0)$  linear approximierbar, dann ist  $\mathbf{f} \circ \mathbf{g}$  in  $\mathbf{x}_0$  auch linear approximierbar und es gilt:

$$\mathbf{J}_{\mathbf{f} \circ \mathbf{g}}(\mathbf{x}_0) = \mathbf{J}_{\mathbf{f}}(\mathbf{g}(\mathbf{x}_0)) \cdot \mathbf{J}_{\mathbf{g}}(\mathbf{x}_0)$$

Das heißt die Jacobi-Matrix der verketteten Abbildung ergibt sich als Produkt der Jacobi-Matrix der äußeren Abbildung am Punkt  $\mathbf{g}(\mathbf{x}_0)$  und der Jacobi-Matrix der inneren Abbildung.

**Beweis:**

$$\begin{aligned} \mathbf{f}(\mathbf{g}(\mathbf{x})) &= \mathbf{f}(\mathbf{g}(\mathbf{x}_0)) + \mathbf{J}_{\mathbf{f}}(\mathbf{g}(\mathbf{x}_0)) \cdot (\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{x}_0)) + \mathbf{r}_{\mathbf{f}}(\|\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{x}_0)\|) \\ &= \mathbf{f}(\mathbf{g}(\mathbf{x}_0)) + \mathbf{J}_{\mathbf{f}}(\mathbf{g}(\mathbf{x}_0)) \cdot [\mathbf{g}(\mathbf{x}_0) + \mathbf{J}_{\mathbf{g}}(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \mathbf{r}_{\mathbf{g}}(\|\mathbf{x} - \mathbf{x}_0\|) - \mathbf{g}(\mathbf{x}_0)] \\ &\quad + \mathbf{r}_{\mathbf{f}}(\|\mathbf{g}(\mathbf{x}_0) + \mathbf{J}_{\mathbf{g}} \cdot (\mathbf{x} - \mathbf{x}_0) + \mathbf{r}_{\mathbf{g}}(\|\mathbf{x} - \mathbf{x}_0\|) - \mathbf{g}(\mathbf{x}_0)\|) \\ &= \mathbf{f}(\mathbf{g}(\mathbf{x}_0)) + \mathbf{J}_{\mathbf{f}}(\mathbf{g}(\mathbf{x}_0)) \cdot [\mathbf{J}_{\mathbf{g}}(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) + \mathbf{r}_{\mathbf{g}}(\|\mathbf{x} - \mathbf{x}_0\|)] \\ &\quad + \mathbf{r}_{\mathbf{f}}(\|\mathbf{J}_{\mathbf{g}} \cdot (\mathbf{x} - \mathbf{x}_0) + \mathbf{r}_{\mathbf{g}}(\|\mathbf{x} - \mathbf{x}_0\|)\|) \end{aligned}$$

□

#### Beispiel 10.52

Gegeben seien die beiden Funktionen  $\mathbf{x}(r, \varphi) = (r \cos \varphi \quad r \sin \varphi)^T$  und  $\mathbf{f}(\mathbf{x}) = (x_1 + x_2 \quad x_1 - x_2)^T$ . Wir wollen die direkt berechnete Jacobi-Matrix und die mit der Kettenregel berechnete vergleichen:

$$\begin{aligned} \mathbf{f}(r, \varphi) &= \begin{pmatrix} r \cos \varphi + r \sin \varphi \\ r \cos \varphi - r \sin \varphi \end{pmatrix} \\ J_{\mathbf{f}}(r, \varphi) &= \begin{pmatrix} \cos \varphi + \sin \varphi & -r \sin \varphi + r \cos \varphi \\ \cos \varphi - \sin \varphi & -r \sin \varphi - r \cos \varphi \end{pmatrix} \\ J_{\mathbf{f} \circ \mathbf{x}} &= \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix} = \begin{pmatrix} \cos \varphi + \sin \varphi & -r \sin \varphi + r \cos \varphi \\ \cos \varphi - \sin \varphi & -r \sin \varphi - r \cos \varphi \end{pmatrix} \end{aligned}$$

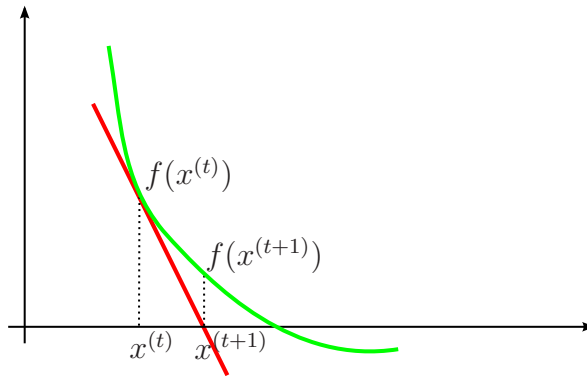
### 10.13.3. Newtonverfahren im $\mathbb{R}^n$

**Problem:** Sei  $\mathbf{f} : D \rightarrow \mathbb{R}^m$ ,  $D \subseteq \mathbb{R}^n$  offen, eine  $\mathcal{C}^1$ -Funktion

Gesucht ist  $\mathbf{x}_0 \in \mathbb{R}^n$  mit  $\mathbf{f}(\mathbf{x}_0) = \mathbf{0}$

**Idee:**

1. Wähle Startwert  $\mathbf{x}^{(0)}$
2. Approximiere  $\mathbf{f}(\mathbf{x})$  in  $\mathbf{x}^{(0)}$  linear.
3. Suche Punkt  $\mathbf{x}^{(1)}$  für den die lineare Approximation Null ist.
4. Wiederhole und hoffe, dass die Folge  $\mathbf{x}^{(i)}$  gegen  $\mathbf{x}_0$  konvergiert.



Lineare Approximation:

$$\mathbf{f}(\mathbf{x}) \approx \mathbf{f}(\mathbf{x}^{(i)}) + \mathbf{J}_{\mathbf{f}}(\mathbf{x}^{(i)}) \cdot (\mathbf{x} - \mathbf{x}^{(i)})$$

Für einen Iterationsschritt setze

$$\Delta \mathbf{x}^{(i)} = \mathbf{x}^{(i+1)} - \mathbf{x}^{(i)} \quad \Rightarrow \quad \mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} + \Delta \mathbf{x}^{(i)}$$

Zu lösen:

$$\begin{aligned} \mathbf{f}(\mathbf{x}^{(i)}) + \mathbf{J}_{\mathbf{f}}(\mathbf{x}^{(i)}) \cdot \Delta \mathbf{x}^{(i)} &= 0 \\ \Leftrightarrow \mathbf{J}_{\mathbf{f}}(\mathbf{x}^{(i)}) \cdot \Delta \mathbf{x}^{(i)} &= -\mathbf{f}(\mathbf{x}^{(i)}) \end{aligned}$$

$\Rightarrow$  Erfordert Lösung eines linearen Gleichungssystems in jedem Iterationsschritt.

Es lässt sich zeigen, dass das Newton-Verfahren quadratisch konvergiert, wenn der Startwert  $\mathbf{x}^{(0)}$  nahe genug bei  $\mathbf{x}_0$  liegt, d.h es gilt dann:

$$\|\mathbf{x}^{(i+1)} - \mathbf{x}_0\| \leq C \cdot \|\mathbf{x}^{(i+1)} - \mathbf{x}^{(i)}\|^2 \quad \text{mit } C \in \mathbb{R} \text{ konstant}$$

oder

$$\|\mathbf{x}^{(i)} - \mathbf{x}_0\| \leq Cq^{(2^i)} \quad 0 < q < 1$$

# 11. Integration von Funktionen mehrerer Variablen

## 11.1. Kurvenintegrale

Wir möchten eine Funktion  $f(\mathbf{x}(t))$  mit  $f : D \rightarrow \mathbb{R}$ ,  $D \subseteq \mathbb{R}^n$  entlang des Kurvenstücks  $\mathbf{x} : [a, b] \rightarrow \mathbb{R}^n$ ,  $[a, b] \subseteq \mathbb{R}$  integrieren.

Dazu wählen wir Punkte  $\mathbf{p}_i = (x_1(t_i), \dots, x_n(t_i))^T$ ,  $0 \leq i \leq k$  entlang von  $\mathbf{x}$  mit  $\mathbf{p}_0 = \mathbf{x}(a)$  und  $\mathbf{p}_k = \mathbf{x}(b)$ . Die Länge der Kurve zwischen  $\mathbf{p}_{i-1}$  und  $\mathbf{p}_i$  approximieren wir durch  $\|\mathbf{p}_i - \mathbf{p}_{i-1}\|$ . Analog zur Integration gewöhnlicher Funktionen bilden wir die Summe

$$\begin{aligned} & \sum_{i=1}^k [f(x_1(t_i), \dots, x_n(t_i)) \cdot \|\mathbf{p}_i - \mathbf{p}_{i-1}\|] \\ &= \sum_{i=1}^k \left[ f(x_1(t_i), \dots, x_n(t_i)) \cdot \sqrt{\sum_{j=1}^n (x_j(t_i) - x_j(t_{i-1}))^2} \right] \\ &= \sum_{i=1}^k \left[ f(\mathbf{x}(t_i)) \cdot (t_i - t_{i-1}) \cdot \sqrt{\sum_{j=1}^n \left( \frac{x_j(t_i) - x_j(t_{i-1})}{t_i - t_{i-1}} \right)^2} \right] \end{aligned}$$

Für  $k \rightarrow \infty$  mit  $\max_{1 \leq k \leq n} (t_i - t_{i-1}) \rightarrow 0$  und  $\sum_{i=1}^n (t_i - t_{i-1}) = b - a$  konstant konvergiert die Summe gegen

$$\int_a^b f(\mathbf{x}(t)) \cdot \sqrt{\sum_{j=1}^n (\dot{x}_j(t))^2} dt = \int_a^b f(\mathbf{x}(t)) \cdot \|\dot{\mathbf{x}}(t)\| dt$$

### Definition 11.1

Sei  $D \subseteq \mathbb{R}^n$  offen,  $\mathbf{x} : [a, b] \rightarrow D$  ein Kurvenstück und  $f : D \rightarrow \mathbb{R}$ , sowie  $f \circ \mathbf{x} : [a, b] \rightarrow \mathbb{R}$  stetig. Dann heißt

$$\int_{\mathbf{x}} f ds := \int_a^b f(\mathbf{x}(t)) \cdot \|\dot{\mathbf{x}}(t)\| dt$$

das Kurvenintegral von  $f$  längs  $\mathbf{x}$ .

### Beispiel 11.2

Bestimme die Masse einer Feder.

Die Dichte der Feder sei  $\rho(\mathbf{x}) = x_1^2 \cdot x_2^2 + x_3^2$

Parametrisierung der Feder (Form):

$$\begin{aligned}
 \mathbf{x}(t) &= \begin{pmatrix} \cos 2t \\ \sin 2t \\ \frac{1}{2}t \end{pmatrix} & 0 \leq t \leq 2\pi \\
 \dot{\mathbf{x}}(t) &= \begin{pmatrix} -2 \sin 2t \\ 2 \cos 2t \\ \frac{1}{2} \end{pmatrix} \\
 \|\dot{\mathbf{x}}(t)\| &= \sqrt{4 \sin^2 2t + 4 \cos^2 2t + \frac{1}{4}} = \sqrt{4(\cos^2 2t + \sin^2 2t) + \frac{1}{4}} \\
 &= \sqrt{\frac{17}{4}} = \frac{1}{2}\sqrt{17} \\
 \Rightarrow ds &= \frac{1}{2}\sqrt{17} dt \\
 M &= \int_{\mathbf{x}} \rho(\mathbf{x}) ds = \int_0^{2\pi} (x_1^2 x_2^2 + x_3^2) \cdot \frac{1}{2}\sqrt{17} dt \\
 &= \frac{1}{2}\sqrt{17} \int_0^{2\pi} \left( \cos^2 2t \cdot \sin^2 2t + \frac{1}{4}t^2 \right) dt \\
 &= \frac{1}{2}\sqrt{17} \left( \int_0^{2\pi} \cos^2 2t \cdot \sin^2 2t dt + \frac{1}{4} \left[ \frac{t^3}{3} \right]_0^{2\pi} \right)
 \end{aligned}$$

Nebenrechnung:

$$\begin{aligned}
 \cos^2 2t \cdot \sin^2 2t &= \frac{1}{4} \sin^2 4t && \text{wegen } 2 \cdot \cos \alpha \cdot \sin \alpha = \sin 2\alpha \\
 &= \frac{1}{8} (1 - \cos 8t) && \text{wegen } \sin^2 \frac{\alpha}{2} = \frac{1}{2} (1 - \cos \alpha)
 \end{aligned}$$

$$\begin{aligned}
 \Rightarrow M &= \frac{1}{2}\sqrt{17} \left( \frac{1}{8} \int_0^{2\pi} (1 - \cos 8t) dt + \frac{1}{4} \cdot \frac{8\pi^3}{3} \right) \\
 &= \frac{1}{2}\sqrt{17} \left( \frac{1}{8} \left[ t - \frac{1}{8} \sin 8t \right]_0^{2\pi} + \frac{2}{3}\pi^3 \right) \\
 &= \frac{1}{2}\sqrt{17} \left( \frac{1}{8} \cdot 2\pi + \frac{2}{3}\pi^3 \right) \\
 &= \frac{1}{8}\sqrt{17} \left( \pi + \frac{8}{3}\pi^3 \right)
 \end{aligned}$$

Erweiterung auf stückweise stetig differenzierbare Kurvenstücke (z.B. Kurven mit Ecken):

**Definition 11.3**

1. Unter einer Kurve  $\mathbf{x}$  in  $D \subseteq \mathbb{R}^n$  verstehen wir eine Folge  $\mathbf{x}_1, \dots, \mathbf{x}_k$  von (stetig differenzierbaren) Kurvenstücken  $\mathbf{x}_i : [a_i, b_i] \rightarrow D$  ( $1 \leq i \leq k$ ) mit  $\mathbf{x}_i(b_i) = \mathbf{x}_{i+1}(a_{i+1})$  ( $1 \leq i \leq k-1$ )
2. Für eine aus den Kurvenstücken  $\mathbf{x}_1, \dots, \mathbf{x}_k$  bestehende Kurve ist das Kurvenintegral einer Funktion  $f : D \rightarrow \mathbb{R}$  längs  $\mathbf{x}$  definiert als

$$\int_{\mathbf{x}} f \, ds := \sum_{i=1}^k \int_{\mathbf{x}_i} f \, ds$$

3. Ist die Kurve  $\mathbf{x}$  geschlossen, schreibt man auch

$$\oint_{\mathbf{x}} f \, ds.$$

**Definition 11.4**

1. Eine Funktion  $f : D \rightarrow \mathbb{R}$ ,  $D \subseteq \mathbb{R}^n$  die jedem Punkt  $\mathbf{x} \in D$  im Raum einen Wert  $f(\mathbf{x})$  zuordnet nennt man auch Skalarfeld.
2. Ordnet  $\mathbf{v} : D \rightarrow \mathbb{R}^n$ ,  $D \subseteq \mathbb{R}^n$  jedem Punkt  $\mathbf{x} \in D$  einen Vektor  $\mathbf{v}(\mathbf{x})$  zu, nennt man  $\mathbf{v}$  auch Vektorfeld.

**Definition 11.5**

Sei  $D \subseteq \mathbb{R}^n$  offen,  $\mathbf{x} : [a, b] \rightarrow D$  eine stetig differenzierbare Kurve und  $\mathbf{v} : D \rightarrow \mathbb{R}^n$ ,  $D \subseteq \mathbb{R}^n$  ein stetiges Vektorfeld. Man nennt

$$\int_{\mathbf{x}} \mathbf{v} \cdot d\mathbf{s} := \int_a^b \mathbf{v}(\mathbf{x}(t)) \cdot \dot{\mathbf{x}}(t) \, dt$$

das Kurvenintegral von  $\mathbf{v}$  längs  $\mathbf{x}$ .

**Anmerkung 11.6**

- Im weiteren sei mit dem Produkt  $\mathbf{x} \cdot \mathbf{y}$  zweier Vektoren  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  immer das Skalarprodukt gemeint.
- Sei  $\mathbf{y}(t) = \frac{\dot{\mathbf{x}}(t)}{\|\dot{\mathbf{x}}(t)\|}$  der normierte Tangentialvektor. Wegen  $\dot{\mathbf{x}}(t) = \mathbf{y}(t) \cdot \|\dot{\mathbf{x}}(t)\|$  ist dann

$$\int_{\mathbf{x}} \mathbf{v}(\mathbf{x}(t)) \cdot \dot{\mathbf{x}}(t) \, dt = \int_{\mathbf{x}} \mathbf{v}(t) \cdot \mathbf{y}(t) \cdot \|\dot{\mathbf{x}}(t)\| \, dt = \int_{\mathbf{x}} \mathbf{v}(\mathbf{x}(t)) \cdot \mathbf{y}(t) \, ds$$

Es wird also die Tangentialkomponente des Vektorfelds integriert.

- Sowohl für skalare Funktionen  $f, g$  als auch für Vektorfelder  $\mathbf{u}, \mathbf{v}$  gilt mit  $\alpha, \beta \in \mathbb{R}$

$$\begin{aligned} \int_{\mathbf{x}} (\alpha f + \beta g) \, ds &= \alpha \int_{\mathbf{x}} f \, ds + \beta \int_{\mathbf{x}} g \, ds \\ \int_{\mathbf{x}} (\alpha \mathbf{u} + \beta \mathbf{v}) \cdot d\mathbf{s} &= \alpha \int_{\mathbf{x}} \mathbf{u} \cdot d\mathbf{s} + \beta \int_{\mathbf{x}} \mathbf{v} \cdot d\mathbf{s} \end{aligned}$$

- Aufgrund der Eigenschaften des Skalarprodukts gilt für  $\mathbf{v} : D \rightarrow \mathbb{R}^n$ ,  $D \subseteq \mathbb{R}^n$

$$\begin{aligned} \int_{\mathbf{x}} \mathbf{v} \cdot \dot{\mathbf{x}}(t) dt &= \int_{\mathbf{x}} \mathbf{v}^T \cdot \begin{pmatrix} \frac{dx_1}{dt} \\ \vdots \\ \frac{dx_n}{dt} \end{pmatrix} dt = \int_{\mathbf{x}} \left( v_1 \frac{dx_1}{dt} + \cdots + v_n \frac{dx_n}{dt} \right) dt \\ &= \int_{\mathbf{x}} (v_1 dx_1 + v_2 dx_2 + \cdots + v_n dx_n) = \sum_{i=1}^n \int_{\mathbf{x}} v_i dx_i \end{aligned}$$

### Beispiel 11.7

$$F = \int_{\mathbf{x}} \begin{pmatrix} x^2 y \\ x - z \\ xyz \end{pmatrix} \cdot d\mathbf{s} = \int_{\mathbf{x}} x^2 y dx + (x - z) dy + (xyz) dz$$

längs der Kurve mit  $y = x^3$  ( $0 \leq x \leq 2$ ),  $z = 2$

$$\mathbf{x}(t) = \begin{pmatrix} t \\ t^3 \\ 2 \end{pmatrix} \quad 0 \leq t \leq 2$$

$$dx = dt, \quad dy = 3t^2 dt, \quad dz = 0$$

$$F = \int_0^2 (t^2 t^3 + (t - 2) 3t^2) dt = \left[ \frac{t^6}{6} + \frac{3}{4} t^4 - 2t^3 \right]_0^2 = \frac{32}{3} + 12 - 16 = \frac{20}{3}$$

Alternativ:

$$F = \int_0^2 \begin{pmatrix} t^2 \cdot t^3 \\ t - 2 \\ 2t \cdot t^3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 3t^2 \\ 0 \end{pmatrix} dt = \int_0^2 (t^5 + (t - 2) 3t^2) dt = \frac{20}{3}$$

## 11.2. Potential eines Gradientenfelds

### Definition 11.8

Eine Teilmenge  $\Omega \subseteq \mathbb{R}^n$  heißt Gebiet, wenn gilt:

1.  $\Omega$  ist offen
2.  $\Omega$  ist zusammenhängend, d.h. zu je zwei Punkten  $\mathbf{x}_0, \mathbf{y}_0 \in \Omega$  gibt es eine reguläre Kurve  $\mathbf{x} : [a, b] \rightarrow \Omega$  mit  $\mathbf{x}(a) = \mathbf{x}_0$ ,  $\mathbf{x}(b) = \mathbf{y}_0$ , die komplett in  $\Omega$  verläuft.

Eine Kurve  $\mathbf{x}$  ist regulär, wenn  $\|\dot{\mathbf{x}}(t)\| \neq 0 \quad \forall t \in [a, b]$

### Definition 11.9

Sei  $\Omega \subseteq \mathbb{R}^n$  ein Gebiet. Man nennt ein Vektorfeld  $\mathbf{v} \in \mathcal{C}^0(\Omega, \mathbb{R}^n)$  konservativ oder ein Potential- bzw. ein Gradientenfeld, wenn es eine Funktion  $f \in \mathcal{C}^1(\Omega, \mathbb{R})$  gibt mit  $\mathbf{v}(\mathbf{x}) = \nabla f(\mathbf{x})$  ( $\forall \mathbf{x} \in \Omega$ ).  $f$  heißt dann Stammfunktion und  $U := -f$  Potentialfunktion (oder ein Potential) von  $\mathbf{v}$ .

### Beispiel 11.10

- Gewichtskraft  $\rho \cdot \mathbf{g}$  und potentielle Energie  $E_{pot} = - \int_{\mathbf{x}} \rho \mathbf{g} \cdot d\mathbf{s}$



- Elektrische Feldstärke  $\mathbf{E}$  und Coulomb-Potential  $\Phi = - \int_{\mathbf{x}} \mathbf{E} \cdot d\mathbf{s}$

**Satz 11.11 (1. Hauptsatz für Kurvenintegrale)**

Ist  $\mathbf{v} : \Omega \rightarrow \mathbb{R}^n$  ein stetiges Gradientenfeld auf dem Gebiet  $\Omega \subseteq \mathbb{R}^n$  mit Stammfunktion  $f$ , dann gilt für jede stückweise reguläre Kurve  $\mathbf{x}$  in  $\Omega$  mit Anfang  $\mathbf{x}(a)$  und Ende  $\mathbf{x}(b)$

$$\int_{\mathbf{x}} \mathbf{v} \cdot d\mathbf{s} = f(\mathbf{x}(b)) - f(\mathbf{x}(a))$$

**Beweis:**

$$\int_{\mathbf{x}} \mathbf{v} \cdot d\mathbf{s} = \int_{\mathbf{x}} \nabla f \cdot d\mathbf{s} = \int_a^b \nabla f(\mathbf{x}(t)) \cdot \dot{\mathbf{x}}(t) dt = \int_a^b \left( \frac{d}{dt} f(\mathbf{x}(t)) \right) dt = f(\mathbf{x}(b)) - f(\mathbf{x}(a))$$

□

**Satz 11.12**

Für ein stetiges Vektorfeld  $\mathbf{v} \in \mathcal{C}^0(\Omega, \mathbb{R}^n)$  auf dem Gebiet  $\Omega \subseteq \mathbb{R}^n$  sind die folgenden Aussagen äquivalent:

1.  $\mathbf{v}$  ist ein Potentialfeld.
2. Für alle regulären Kurven  $\mathbf{x}$  in  $\Omega$  hängt  $\int_{\mathbf{x}} \mathbf{v} \cdot d\mathbf{s}$  nur vom Anfangs- und Endpunkt ab (Integral ist wegunabhängig).
3. Für alle geschlossenen regulären Kurven in  $\Omega$  gilt:

$$\oint_{\mathbf{x}} \mathbf{v} \cdot d\mathbf{s} = 0$$

ohne Beweis

**Definition 11.13**

Ein Gebiet  $\Omega \subseteq \mathbb{R}^n$  heißt einfach zusammenhängend, wenn jede geschlossene, doppelunktfreie Kurve in  $\Omega$  (d.h. eine Kurve ohne Überschneidungen) stetig auf einen Punkt in  $\Omega$  zusammengezogen werden kann ohne  $\Omega$  zu verlassen (d.h.  $\Omega$  hat keine Löcher).

**Satz 11.14 (2. Hauptsatz für Kurvenintegrale)**

Ein  $\mathcal{C}^1$ -Vektorfeld  $\mathbf{v} : \Omega \rightarrow \mathbb{R}^n$  auf einem einfach zusammenhängenden Gebiet  $\Omega \subseteq \mathbb{R}^n$  ist genau dann ein Potentialfeld, wenn die "Integrabilitätsbedingung"  $\mathbf{J}_{\mathbf{v}}(\mathbf{x}) = \mathbf{J}_{\mathbf{v}}(\mathbf{x})^T$  für alle  $\mathbf{x} \in \Omega$  erfüllt ist (d.h. wenn die Jacobi-Matrix von  $\mathbf{v}$  überall in  $\Omega$  symmetrisch ist).

## 11.3. Flächen- und Volumenintegral

### 11.3.1. Fläche und Volumen

Zunächst wollen wir in der Lage sein, die Fläche und das Volumen von beliebig geformten Gebieten zu bestimmen.

Sei  $M \subseteq \mathbb{R}^2$  eine Menge von Punkten in der  $(x, y)$ -Ebene. Wir zerlegen die  $(x, y)$ -Ebene durch die Koordinatenlinien  $x = n \cdot 2^{-k}$  und  $y = n \cdot 2^{-k}$ , ( $n = 0, \pm 1, \pm 2, \dots$ ) in Quadrate mit dem Flächeninhalt  $2^{-2k}$  und erhalten damit eine Folge von Gittern, die um so feiner sind, je größer  $k$  wird.

Sei  $s_k(M)$  der Flächeninhalt der Quadrate, die komplett innerhalb von  $M$  liegen und  $S_k(M)$  der Flächeninhalt der Quadrate, die mindestens einen Punkt von  $M$  enthalten. Dann gilt:

$$s_k(M) \leq S_k(M) \quad s_k(M) \leq s_{k+1}(M) \quad S_{k+1}(M) \leq S_k(M)$$

Daher müssen die Grenzwerte

$$A_i(M) := \lim_{k \rightarrow \infty} s_k(M) \quad \text{innerer Flächeninhalt}$$

$$A_a(M) := \lim_{k \rightarrow \infty} S_k(M) \quad \text{äußerer Flächeninhalt}$$

existieren.

Wenn  $A_a(M) = A_i(M) =: A(M)$ , heißt  $M$  Riemann-messbar und  $A(M)$  heißt der Flächeninhalt von  $M$ .

### Satz 11.15

Ein beschränktes Gebiet  $\Omega \subset \mathbb{R}^2$  mit stückweise regulärem Rand besitzt einen Flächeninhalt

$$A = \lim_{k \rightarrow \infty} s_k(\Omega) = \lim_{k \rightarrow \infty} S_k(\Omega).$$

Für Volumina verfährt man analog unter Einschluss der dritten Koordinatenachse. Satz 11.15 gilt dann entsprechend bei Gebieten, deren Rand aus stückweise regulären Flächen besteht.

### 11.3.2. Flächenintegral

Wir möchten nun eine Funktion über einen Bereich  $B$  integrieren. Für eine Funktion  $f(x, y)$  wäre  $B$  eine Fläche, für  $f(x, y, z)$  ein Volumen.

Damit so ein Integral existiert müssen Bedingungen an  $f$  und  $B$  erfüllt sein:

#### Definition 11.16

Ein Bereich  $B \subset \mathbb{R}^2$  ( $B \subset \mathbb{R}^3$ ) heißt regulär, wenn

1. der Rand  $\partial B$  aus endlich vielen regulären Kurvenstücken (stückweise regulären Flächen, siehe Definition 11.32) besteht.
2. das Innere  $B \setminus \partial B$  ein nicht leeres beschränktes Gebiet im  $\mathbb{R}^2$  ( $\mathbb{R}^3$ ) ist (d.h. offen und zusammenhängend).
3.  $B$  abgeschlossen ist, d.h.  $\partial B \subset B$

Sei  $f : B \rightarrow \mathbb{R}$  eine auf  $B$  stetige und beschränkte Funktion. Wir zerlegen  $B$  durch ein Netz regulärer Kurven in  $n$  Teilbereiche  $B_1, \dots, B_n$ . Der Durchmesser eines Teilbereichs  $\delta(B_i)$  ist definiert als das Infimum der Durchmesser aller Kreise, die  $B_i$  vollständig überdecken. Nach Satz 11.15 hat  $B_i$  einen Flächeninhalt  $\Delta A_i$ . Wir wählen für jedes  $B_i$  einen beliebigen Punkt  $(x_i, y_i) \in B_i$  (z.B. den Schwerpunkt) und bilden die Riemann-Summe

$$Z_n := \sum_{i=1}^n f(x_i, y_i) \Delta A_i.$$

Unter den gegebenen Voraussetzungen konvergieren die Riemann-Summen  $Z_n$  bei Verfeinerung des Netzes gegen einen Grenzwert, der unabhängig von der spezifischen Zerlegung ist, falls der maximale Durchmesser

$$\delta_{\max} := \max_{1 \leq i \leq n} \delta(B_i)$$

gegen Null geht.

### Definition 11.17

Der Grenzwert

$$\iint_B f(x, y) \, dA = \iint_B f \, dA := \lim_{\substack{\delta_{\max} \rightarrow 0 \\ n \rightarrow \infty}} \sum_{i=1}^n f(x_i, y_i) \Delta A_i$$

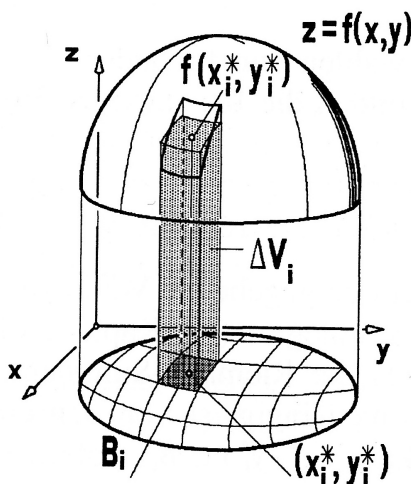
wird (wenn er existiert) Doppelintegral oder Gebietsintegral von  $f$  über  $B$  genannt. Das Symbol  $dA$  heißt Flächenelement. In kartesischen Koordinaten setzt man  $dA = dx \, dy$ .

### Anmerkung 11.18

- Das Integral über 1 liefert den Flächeninhalt von  $B$ :  $A_B = \iint_B dA$
- Damit lässt sich ein Mittelwert definieren als

$$\bar{f} = \frac{\iint_B f \, dA}{\iint_B dA}$$

- Falls  $f(x, y) > 0$  in  $B$  ( $< 0$  in  $B$ ) dann entspricht  $V = \iint_B f \, dA$  ( $-V$ ) dem Volumen eines Körpers über der Fläche  $B$ , der jeweils die Höhe  $f(x, y)$  hat.



Quelle: Meyberg, Vachenauer: Höhere Mathematik 1, Springer

**Definition 11.19**

Wir nennen  $B_1 \subseteq \mathbb{R}^2$  einen Normalbereich vom Typ I, wenn es  $a, b \in \mathbb{R}$  und  $\mathcal{C}^1$ -Funktionen  $g, h : [a, b] \rightarrow \mathbb{R}$  gibt mit  $g(x) \leq h(x)$  und

$$B_1 = \{(x, y) : a \leq x \leq b, \quad g(x) \leq y \leq h(x)\}$$

und einen Normalbereich vom Typ II, falls

$$B_2 = \{(x, y) : \ell(y) \leq x \leq r(y), \quad c \leq y \leq d\}$$

mit  $c, d \in \mathbb{R}$  und  $\mathcal{C}^1$ -Funktionen  $\ell, r : [c, d] \rightarrow \mathbb{R}$  mit  $\ell(y) \leq r(y)$ .

**Satz 11.20**

1. Für jede stetige Funktion  $f : B_1 \rightarrow \mathbb{R}$  mit  $B_1 \subseteq \mathbb{R}^2$  Normalbereich vom Typ I gilt:

$$\iint_{B_1} f(x, y) \, dx \, dy = \int_a^b \left( \int_{g(x)}^{h(x)} f(x, y) \, dy \right) dx$$

2. Für jede stetige Funktion  $f : B_2 \rightarrow \mathbb{R}$  mit  $B_2 \subseteq \mathbb{R}^2$  Normalbereich vom Typ II gilt:

$$\iint_{B_2} f(x, y) \, dx \, dy = \int_c^d \left( \int_{\ell(y)}^{r(y)} f(x, y) \, dx \right) dy$$

ohne Beweis

Komplizierte Gebiete zerlegt man in Normalbereiche und summiert die Ergebnisse.

**Satz 11.21**

Für  $B = \{(x, y) : a \leq x \leq b, \quad c \leq y \leq d\}$  gilt:

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

**Beispiel 11.22**

1.

$$\begin{aligned} \int_0^2 \int_0^3 x^2 y^3 \, dx \, dy &= \int_0^2 \left( \int_0^3 x^2 \cdot y^3 \, dx \right) dy = \\ &= \int_0^2 y^3 \cdot \left[ \frac{1}{3} x^3 \right]_0^3 dy = \int_0^2 9 \cdot y^3 \, dy = 9 \cdot \left[ \frac{1}{4} y^4 \right]_0^2 = 36 \end{aligned}$$

2.

$$\int_0^{2\pi} \int_0^1 e^{x^2} \sin(y) \, dx \, dy =$$

$e^{x^2}$  nicht einfach integrierbar

$$\begin{aligned} &= \int_0^1 \int_0^{2\pi} e^{x^2} \sin y \, dy \, dx = \int_0^1 e^{x^2} [-\cos y]_0^{2\pi} \, dx \\ &= \int_0^1 e^{-x^2} \cdot (-1 + 1) \, dx = 0 \end{aligned}$$

### 11.3.3. Volumenintegral

Zerlegt man ein reguläres Volumen  $B$  analog zu den Flächenintegralen in Teilbereiche  $B_i$  mit Volumen  $\Delta V_i$ , definiert deren maximalen Durchmesser  $\delta_{\max}$  durch das Infimum des Durchmessers einer Kugel, die  $B_i$  vollständig enthält, und bildet die Riemann-Summen  $Z_n := \sum_{i=1}^n f(x_i, y_i, z_i) \Delta V_i$  mit beliebig gewählten Punkten  $(x_i, y_i, z_i) \in B_i$  für eine stetige und beschränkte Funktion  $f : B \rightarrow \mathbb{R}$ , dann kann man das Volumenintegral wie folgt definieren:

#### Definition 11.23

Der Grenzwert

$$\iiint_B f(x, y, z) \, dV = \iiint_B f \, dV := \lim_{\substack{\delta_{\max} \rightarrow 0 \\ n \rightarrow \infty}} \sum_{i=1}^n f(x_i, y_i, z_i) \Delta V_i$$

wird (wenn er existiert) Dreifachintegral (oder Volumenintegral) von  $f$  über  $B$  genannt. Das Symbol  $dV$  heißt Volumenelement. In kartesischen Koordinaten ist  $dV = dx \, dy \, dz$ .

#### Anmerkung 11.24

- $\iiint_B dV = V_B$  liefert das Volumen von  $B$
- Wie im Zweidimensionalen lässt sich ein Gebiet in Normalbereiche aufteilen. Man erhält dann, z.B.

$$\begin{aligned} \iiint_B f \, dx \, dy \, dz &= \iint_D \left( \int_{g(x,y)}^{h(x,y)} f(x, y, z) \, dz \right) dx \, dy \\ &= \int_a^b \left( \int_{u(x)}^{v(x)} \left( \int_{g(x,y)}^{h(x,y)} f(x, y, z) \, dz \right) dy \right) dx \end{aligned}$$

#### Satz 11.25

Ist  $B = \{(x, y, z) : x_0 \leq x \leq x_1, y_0 \leq y \leq y_1, z_0 \leq z \leq z_1\}$  gilt: Die Integrationsreihenfolge ist beliebig, z.B.

$$\begin{aligned} \iiint_B f \, dx \, dy \, dz &= \int_{x_0}^{x_1} \left( \int_{y_0}^{y_1} \left( \int_{z_0}^{z_1} f(x, y, z) \, dz \right) dy \right) dx \\ &= \int_{z_0}^{z_1} \left( \int_{x_0}^{x_1} \left( \int_{y_0}^{y_1} f(x, y, z) \, dy \right) dx \right) dz = \dots \end{aligned}$$

### Beispiel 11.26

$$\begin{aligned}\int_{-1}^1 \int_{-1}^1 \int_{-1}^1 (x^2 + y^2) \, dx \, dy \, dz &= \int_{-1}^1 \left( \int_{-1}^1 \left[ \frac{1}{3} x^3 + y^2 \cdot x \right]_{-1}^1 dy \right) dz \\&= \int_{-1}^1 \left( \int_{-1}^1 \frac{2}{3} + 2y^2 \, dy \right) dz = \int_{-1}^1 \left[ \frac{2}{3} y + \frac{2}{3} y^3 \right]_{-1}^1 dz \\&= \int_{-1}^1 \frac{4}{3} + \frac{4}{3} \, dz = \frac{8}{3} [z]_{-1}^1 = \frac{16}{3}\end{aligned}$$

### Anmerkung 11.27

Wenn die Dichte eines Körpers durch die Funktion  $\rho(x, y, z)$  gegeben ist, dann lassen sich die folgenden Größen berechnen

- Gesamtmasse

$$M = \iiint_B \rho \, dV$$

- $k$ -tes Moment in Richtung  $i$  ( $i = 1, \dots, 3$ )

$$M_{k,x} = \iiint_B x^k \rho \, dV, \quad M_{k,y} = \iiint_B y^k \rho \, dV, \quad M_{k,z} = \iiint_B z^k \rho \, dV$$

- Schwerpunkt: Die Koordinate des Schwerpunkts in Richtung  $i$

$$x_S = \frac{M_{1,x}}{M}, \quad y_S = \frac{M_{1,y}}{M}, \quad z_S = \frac{M_{1,z}}{M}$$

- geometrischer Schwerpunkt:

$$\bar{x} = \frac{1}{V_B} \iiint_B x \, dV, \quad \bar{y} = \frac{1}{V_B} \iiint_B y \, dV, \quad \bar{z} = \frac{1}{V_B} \iiint_B z \, dV$$

Entsprechendes lässt sich im Zweidimensionalen definieren.

## 11.4. Variablentransformation für Integrale

Eine Koordinatentransformation (z.B. von kartesischen Koordinaten in Kugelkoordinaten) lässt sich durch eine (in der Regel nichtlineare) Abbildung darstellen.

### Satz 11.28

Sei  $B \subseteq \mathbb{R}^n$  ein regulärer Bereich,  $f : B \rightarrow \mathbb{R}$  stetig und  $\phi : U \rightarrow B$ ,  $U \subseteq \mathbb{R}^n$  eine bijektive  $\mathcal{C}^1$ -Abbildung. Dann gilt

$$\int_B f(\mathbf{x}) \, d\mathbf{x} = \int_U f(\phi(\mathbf{y})) \cdot |\det \mathbf{J}_\phi(\mathbf{y})| \, d\mathbf{y}$$

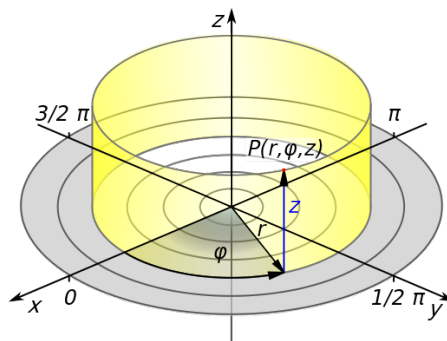
Dabei ist  $\det \mathbf{J}_\phi(\mathbf{y})$  die Determinante der Jacobi-Matrix von  $\phi$  an der Stelle  $\mathbf{y}$  und  $d\mathbf{x}$  und  $d\mathbf{y}$  stehen für  $dx_1 \dots dx_n$  und  $dy_1 \dots dy_n$ .

#### Anmerkung 11.29

- **Affine Koordinaten** (lineare Abbildung):  
 $\mathbf{x} = \mathbf{A}\mathbf{y} + \mathbf{x}_0$ ,  $\mathbf{A} \in \mathbb{R}^{n \times n}$  mit  $\det \mathbf{A} \neq 0$

$$\iiint_B f(\mathbf{x}) dx_1 dx_2 dx_3 = \iiint_U f(\mathbf{A}\mathbf{y} + \mathbf{x}_0) \cdot |\det \mathbf{A}| dy_1 dy_2 dy_3$$

- **Zylinderkoordinaten:**

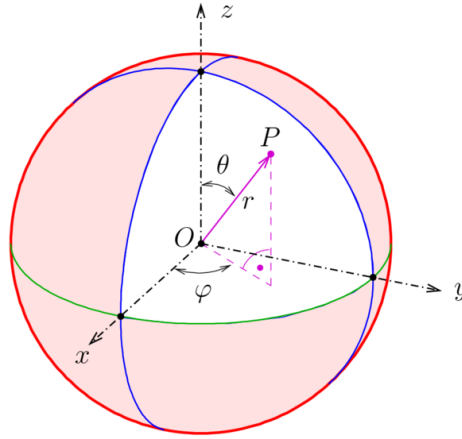


Zylinderkoordinaten eines Punktes  $P$  (hier wird  $\varphi$  für  $\phi$  verwendet)

Quelle: Wiki Commons, CC BY-SA 3.0, von Honina

$$\begin{aligned} \mathbf{x} &= \begin{pmatrix} r \cos \phi \\ r \sin \phi \\ z \end{pmatrix} \\ |\det \mathbf{J}| &= \left| \det \begin{pmatrix} \cos \phi & -r \sin \phi & 0 \\ \sin \phi & r \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \right| \underbrace{=}_{\text{Entwicklung nach 3. Spalte}} \left| \det \begin{pmatrix} \cos \phi & -r \sin \phi \\ \sin \phi & r \cos \phi \end{pmatrix} \right| \\ &= |r \cos^2 \phi + r \sin^2 \phi| = r \\ \iiint_B f(\mathbf{x}) dx dy dz &= \iiint_U f(r \cos \phi, r \sin \phi, z) \cdot r dr d\phi dz \end{aligned}$$

- **Kugelkoordinaten:**



Kugelkoordinaten eines Punktes  $P$  (hier wird  $\Theta$  für  $\psi$  und  $\varphi$  für  $\phi$  verwendet)

Quelle: Wiki Commons, CC BY-SA 3.0, von Ag2gaeh

$$\begin{aligned}
 \mathbf{x} &= \begin{pmatrix} r \sin \psi \cos \phi \\ r \sin \psi \sin \phi \\ r \cos \psi \end{pmatrix} \\
 |\det \mathbf{J}| &= \left| \det \begin{pmatrix} \sin \psi \cos \phi & r \cos \psi \cos \phi & -r \sin \psi \sin \phi \\ \sin \psi \sin \phi & r \cos \psi \sin \phi & r \sin \psi \cos \phi \\ \cos \psi & -r \sin \psi & 0 \end{pmatrix} \right| \\
 &= r^2 \sin \psi \cos^2 \psi \cos^2 \phi + r^2 \sin^3 \psi \sin^2 \phi + r^2 \sin \psi \cos^2 \psi \sin^2 \phi + r^2 \sin^3 \psi \cos^2 \phi \\
 &= r^2 [\sin^3 \psi (\sin^2 \phi + \cos^2 \phi) + \sin \psi \cos^2 \psi (\sin^2 \phi + \cos^2 \phi)] \\
 &= r^2 \sin \psi (\sin^2 \psi + \cos^2 \psi) \\
 &= r^2 \sin \psi \\
 \iiint_B f(\mathbf{x}) \, dx \, dy \, dz &= \iiint_U f(r \sin \psi \cos \phi, r \sin \psi \sin \phi, r \cos \psi) \cdot r^2 \sin \psi \, dr \, d\psi \, d\phi
 \end{aligned}$$

### Beispiel 11.30

Axiales Trägheitsmoment der Kugel  $K$  mit Radius  $R$  und homogener Dichte  $\rho = 1$  um die  $z$ -Achse:

$$\begin{aligned}
 I &= \iiint_K (x^2 + y^2) \rho \, dx \, dy \, dz = \int_0^{2\pi} \int_0^\pi \int_0^R (r^2 \sin^2 \psi \cos^2 \phi + r^2 \sin^2 \psi \sin^2 \phi) \cdot r^2 \sin \psi \, dr \, d\psi \, d\phi \\
 &= \int_0^{2\pi} \int_0^\pi \int_0^R r^4 \sin^3 \psi (\cos^2 \phi + \sin^2 \phi) \, dr \, d\psi \, d\phi = 2\pi \cdot \frac{R^5}{5} \cdot \int_0^\pi \sin^3 \psi \, d\psi \\
 &= 2\pi \cdot \frac{R^5}{5} \cdot \left[ \frac{1}{3} \cos^3 \psi - \cos \psi \right]_0^\pi = 2\pi \cdot \frac{R^5}{5} \cdot \left( -\frac{1}{3} + 1 - \frac{1}{3} + 1 \right) \\
 &= \frac{8}{15} \pi R^5
 \end{aligned}$$



## 11.5. Oberflächenintegral

Ziel: Integration einer Funktion über eine Oberfläche im Raum, z.B. die Erdoberfläche

### Definition 11.31

1. Unter einem regulären Flächenstück  $S$  verstehen wir das Ergebnis einer Abbildung  $\mathbf{x} : \Omega \rightarrow \mathbb{R}^3$ ,  $\mathbf{x}(u, v) = (x_1(u, v), x_2(u, v), x_3(u, v))^T$  mit  $\Omega \subseteq \mathbb{R}^2$  regulär und einer  $\mathcal{C}^1$ -Funktion  $\mathbf{x}$  unter den Bedingungen:
  - a) für  $(u, v), (u', v') \in \Omega$  beliebig und  $(u, v) \neq (u', v')$   
 $\Rightarrow \mathbf{x}(u, v) \neq \mathbf{x}(u', v')$
  - b) Seien

$$\mathbf{x}_u(u, v) = \begin{pmatrix} \frac{\partial x}{\partial u}(u, v) \\ \frac{\partial y}{\partial u}(u, v) \\ \frac{\partial z}{\partial u}(u, v) \end{pmatrix} \text{ und } \mathbf{x}_v(u, v) = \begin{pmatrix} \frac{\partial x}{\partial v}(u, v) \\ \frac{\partial y}{\partial v}(u, v) \\ \frac{\partial z}{\partial v}(u, v) \end{pmatrix}$$

die Tangentialvektoren an das durch  $\mathbf{x}$  definierte Flächenstück  $S$ , dann muss gelten

$$\mathbf{x}_u(u, v) \times \mathbf{x}_v(u, v) \neq \mathbf{0} \quad \forall (u, v) \in \Omega,$$

d.h. die Tangentialvektoren sind nicht parallel, da sie sonst keine Fläche aufspannen.

2. Der auf dem Flächenstück senkrecht stehende Einheitsvektor  $\mathbf{n}$  ist gegeben durch

$$\mathbf{n} := \frac{1}{\|\mathbf{x}_u \times \mathbf{x}_v\|} \cdot (\mathbf{x}_u \times \mathbf{x}_v)$$

3. Durch

$$\mathbf{v}(\lambda, \mu) = \mathbf{x}(u, v) + \lambda \mathbf{x}_u(u, v) + \mu \mathbf{x}_v(u, v)$$

mit  $\lambda, \mu \in \mathbb{R}$  wird die Tangentialebene an  $S$  im Punkt  $\mathbf{x}(u, v)$  aufgespannt.

### Definition 11.32

Eine stückweise reguläre Fläche  $S$  entsteht durch Vereinigung von  $k$  regulären Flächenstücken  $S_i$  bei denen gilt

$$S = \bigcup_{1 \leq i \leq k} S_i$$

$$S_i \cap S_j = \partial S_i \cap \partial S_j \text{ (Teilflächen hängen nur an Rändern zusammen)}$$

$S$  ist zusammenhängend

$\partial S$  besteht aus den Rändern, die nur zu einem Flächenstück gehören. Falls  $\partial S$  leer ist, heißt  $S$  geschlossen

### Beispiel 11.33

1. Graph einer zweidimensionalen Funktion  
 Mit  $z = h(u, v)$ ,  $h \in \mathcal{C}^1(\Omega, \mathbb{R})$

$$\mathbf{x}(u, v) = \begin{pmatrix} u \\ v \\ h(u, v) \end{pmatrix}, \quad (u, v) \in \Omega$$

Bei kartesischem Koordinatensystem verlaufen die Tangentialvektoren parallel zu den Koordinatenachsen.

$$\mathbf{x}_u = \begin{pmatrix} 1 \\ 0 \\ h_u \end{pmatrix}, \quad \mathbf{x}_v = \begin{pmatrix} 0 \\ 1 \\ h_v \end{pmatrix}$$

$$\mathbf{x}_u \times \mathbf{x}_v = \begin{pmatrix} -h_u \\ -h_v \\ 1 \end{pmatrix}$$

z.B. Paraboloidoberfläche

$$\mathbf{x}(u, v) = \begin{pmatrix} u \\ v \\ u^2 + v^2 \end{pmatrix}, \quad \mathbf{x}_u \times \mathbf{x}_v = \begin{pmatrix} -2u \\ -2v \\ 1 \end{pmatrix}$$

2. Zylindermantel mit z-Achse als Zylinderachse und Radius  $r$

$$\mathbf{x}(z, \phi) = \begin{pmatrix} r \cos \phi \\ r \sin \phi \\ z \end{pmatrix}, \quad z_0 \leq z \leq z_1, \quad 0 \leq \phi \leq 2\pi$$

$$\mathbf{x}_z \times \mathbf{x}_\phi = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \times \begin{pmatrix} -r \sin \phi \\ r \cos \phi \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \cdot 0 - r \cos \phi \\ 1 \cdot (-r \sin \phi) - 0 \cdot 0 \\ 0 \cdot (-r \sin \phi) - 0 \cdot (r \cos \phi) \end{pmatrix} = \begin{pmatrix} -r \cos \phi \\ -r \sin \phi \\ 0 \end{pmatrix}$$

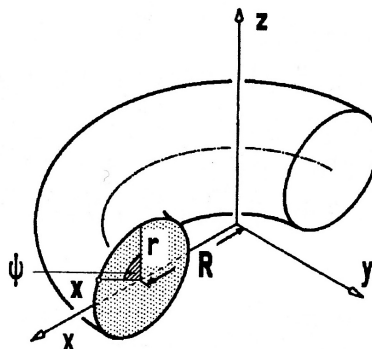
3. Kugeloberfläche um Ursprung mit Radius  $r$

$$\mathbf{x}(\psi, \phi) = \begin{pmatrix} r \sin \psi \cos \phi \\ r \sin \psi \sin \phi \\ r \cos \psi \end{pmatrix}, \quad 0 \leq \psi \leq \pi, \quad 0 \leq \phi \leq 2\pi$$

$$\mathbf{x}_\psi \times \mathbf{x}_\phi = \begin{pmatrix} r^2 \sin^2 \psi \cos \phi \\ r^2 \sin^2 \psi \sin \phi \\ r^2 \sin \psi \cos \psi \end{pmatrix} = r^2 \sin \psi \begin{pmatrix} \sin \psi \cos \phi \\ \sin \psi \sin \phi \\ \cos \psi \end{pmatrix}$$

An den Polen ist  $\mathbf{x}_\psi \times \mathbf{x}_\phi$  bei dieser Darstellung gleich 0. Am Meridian ist  $\mathbf{x}$  nicht eindeutig, da  $\mathbf{x}(\psi, 0) = \mathbf{x}(\psi, 2\pi)$

4. Torus (Ringoberfläche) mit Ringradius  $R$  und Ringquerschnittsradius  $r$

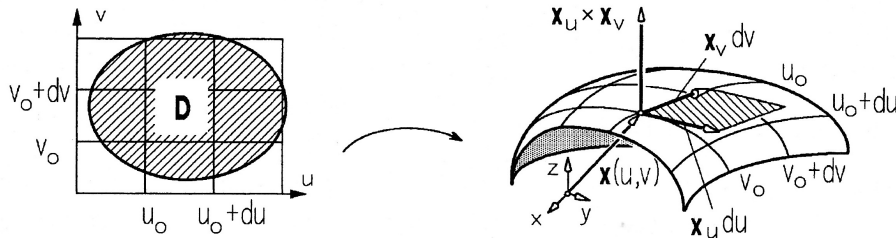


Halber Torus.

Quelle: Meyberg, Vachener: Höhere Mathematik 1, Springer

$$\mathbf{x}(\psi, \phi) = \begin{pmatrix} (R + r \cdot \sin \psi) \cos \phi \\ (R + r \cdot \sin \psi) \sin \phi \\ r \cos \psi \end{pmatrix}, \quad 0 \leq \phi \leq 2\pi, \quad 0 \leq \psi \leq 2\pi$$

$$\mathbf{x}_\psi \times \mathbf{x}_\phi = r \cdot (R + r \cdot \sin \psi) \begin{pmatrix} \sin \psi \cos \phi \\ \sin \psi \sin \phi \\ \cos \psi \end{pmatrix}$$



Integrationsgebiet  $\Omega$  (links) Oberfläche und Integrationselement (rechts).

Quelle: Meyberg, Vachnauer: Höhere Mathematik 1, Springer

### Definition 11.34

Sei  $\mathbf{x} : \Omega \rightarrow \mathbb{R}^3$ ,  $\Omega \subseteq \mathbb{R}^2$  eine Parameterdarstellung eines regulären Flächenstücks  $S$

1. Ist  $f$  ein auf  $S$  stetiges Skalarfeld, dann nennt man

$$\iint_S f \, dO := \iint_\Omega f(x(u, v), y(u, v), z(u, v)) \cdot \|\mathbf{x}_u(u, v) \times \mathbf{x}_v(u, v)\| \, du \, dv$$

das Oberflächenintegral von  $f$  über  $S$ .

2. Ist  $\mathbf{v}$  ein auf  $S$  stetiges Vektorfeld, dann nennt man das Flächenintegral der skalaren Normalkomponente  $\mathbf{v} \cdot \mathbf{n}$

$$\iint_S \mathbf{v} \cdot d\mathbf{O} = \iint_S (\mathbf{v} \cdot \mathbf{n}) \, dO = \iint_\Omega \mathbf{v} \cdot \underbrace{\frac{\mathbf{x}_u \times \mathbf{x}_v}{\|\mathbf{x}_u \times \mathbf{x}_v\|}}_{\mathbf{n}} \cdot \|\mathbf{x}_u \times \mathbf{x}_v\| \, du \, dv = \iint_\Omega \mathbf{v} \cdot (\mathbf{x}_u \times \mathbf{x}_v) \, du \, dv$$

den Fluss von  $\mathbf{v}$  durch  $S$ .

3. Die Oberfläche  $O_S$  ist gegeben durch

$$O_S := \iint_S dO = \iint_\Omega \|\mathbf{x}_u(u, v) \times \mathbf{x}_v(u, v)\| \, du \, dv$$

4. Das Oberflächenintegral über eine stückweise reguläre Fläche  $S = \bigcup_{1 \leq i \leq k} S_i$  berechnet man als Summe der Integrale über die regulären Flächenstücke:

$$\iint_S f \, dO = \sum_{i=1}^k \iint_{S_i} f \, dO$$

**Beispiel 11.35**

## 1. Torusoberfläche

$$\begin{aligned}
\|\mathbf{x}_\psi \times \mathbf{x}_\phi\| &= \sqrt{r^2(R + r \sin \psi)^2} \cdot \underbrace{\left\| \begin{pmatrix} \sin \psi \cos \phi \\ \sin \psi \sin \phi \\ \cos \psi \end{pmatrix} \right\|}_{= 1 \text{ (Einheitskugel)}} \\
&= r(R + r \sin \psi) \\
O &= \int_0^{2\pi} \int_0^{2\pi} r(R + r \sin \psi) \, d\psi \, d\phi = \\
&= 2\pi r \int_0^{2\pi} (R + r \sin \psi) \, d\psi = 2\pi r \left[ R\psi - r \cos \psi \right]_0^{2\pi} \\
&= 4\pi^2 r R
\end{aligned}$$

2. Axiales Trägheitsmoment der Kugelschale mit Radius  $R$  um z-Achse mit homogener Dichte  $\rho = \frac{m}{4\pi R^2}$

$$\begin{aligned}
I &= \iint_S (x^2 + y^2) \cdot \frac{m}{4\pi R^2} \, dO \\
&= \frac{m}{4\pi R^2} \cdot \int_0^{2\pi} \int_0^\pi (R^2 \sin^2 \psi \cos^2 \phi + R^2 \sin^2 \psi \sin^2 \phi) \cdot R^2 \sin \psi \, d\psi \, d\phi \\
&= \frac{m}{4\pi R^2} \int_0^{2\pi} \int_0^\pi R^2 \sin^2 \psi \cdot R^2 \sin \psi \, d\psi \, d\phi \\
&= \frac{m}{2} R^2 \int_0^\pi \sin^3 \psi \, d\psi = \\
&= \frac{m}{2} R^2 \left[ \frac{\cos^3 \psi}{3} - \cos \psi \right]_0^\pi = \frac{m}{2} R^2 \left( -\frac{1}{3} + 1 - \frac{1}{3} + 1 \right) \\
&= \frac{2}{3} m R^2
\end{aligned}$$

3. Elektrisches Feld einer Punktladung:

$$\mathbf{E} = \frac{c \cdot Q}{\|\mathbf{r}\|^2} \cdot \frac{\mathbf{r}}{\|\mathbf{r}\|} \text{ mit } c = \frac{1}{4\pi\epsilon_0\epsilon_r}$$

(quadratischer Abfall, Richtung bei positiver Ladung nach außen)

Integral über Kugeloberfläche:  $\mathbf{n} = \frac{\mathbf{r}}{\|\mathbf{r}\|}$

$$\begin{aligned}\iint_S (\mathbf{E} \cdot \mathbf{n}) \, dO &= \iint_S \frac{c \cdot Q}{\|\mathbf{r}\|^3} \cdot \mathbf{r} \cdot \frac{\mathbf{r}}{\|\mathbf{r}\|} \, dO \\ \int_0^{2\pi} \int_0^\pi \frac{c \cdot Q}{r^2} r^2 \sin \psi \, d\psi \, d\phi &= 2\pi c \cdot Q \cdot \int_0^\pi \sin \psi \, d\psi = 4\pi c \cdot Q \\ &= 4\pi \cdot \frac{1}{4\pi \varepsilon_0 \varepsilon_r} Q = \frac{1}{\varepsilon_0 \varepsilon_r} \cdot Q\end{aligned}$$

Gilt für jede beliebige geschlossene Fläche um Punktladung, da elektrisches Feld Potentialfeld.

## 11.6. Divergenz und Rotation

Für die weiteren Betrachtungen wollen wir den  $\nabla$  Operator einführen als

$$\nabla := \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}$$

Der Gradient von  $f$  ist dann einfach der  $\nabla$ -Operator angewendet auf  $f$  also

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

Weitere wichtige Vektoroperatoren sind Rotation und Divergenz.

### Definition 11.36

Sei  $\mathbf{v} : D \rightarrow \mathbb{R}^n$ ,  $D \subseteq \mathbb{R}^n$  ein  $\mathcal{C}^1$ -Vektorfeld. Dann ist die Divergenz  $\operatorname{div} \mathbf{v} : D \rightarrow \mathbb{R}$  definiert als

$$\operatorname{div} \mathbf{v}(\mathbf{x}) = \nabla \cdot \mathbf{v}(\mathbf{x}) = \frac{\partial v_1}{\partial x_1}(\mathbf{x}) + \dots + \frac{\partial v_n}{\partial x_n}(\mathbf{x}) = \sum_{i=1}^n \frac{\partial v_i}{\partial x_i}(\mathbf{x})$$

### Definition 11.37

Sei  $\mathbf{v} : D \rightarrow \mathbb{R}^3$ ,  $D \subseteq \mathbb{R}^3$  ein  $\mathcal{C}^1$ -Vektorfeld. Dann ist die Rotation  $\operatorname{rot} \mathbf{v} : D \rightarrow \mathbb{R}^3$  definiert als

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

### Definition 11.38

Sei  $f : D \rightarrow \mathbb{R}$ ,  $D \subseteq \mathbb{R}^n$  ein  $\mathcal{C}^2$ -Skalarfeld. Dann ist der Laplace-Operator  $\Delta f : D \rightarrow \mathbb{R}^n$  definiert als

$$\Delta f := \operatorname{div}(\nabla f) = \nabla \cdot \nabla f = \sum_{i=1}^n \frac{\partial^2 f_i}{\partial x_i^2}(\mathbf{x}).$$

### Anmerkung 11.39

- Im Rahmen einer Erhaltungsgleichung lässt sich die Divergenz als Quelldichte eines Vektorfeldes interpretieren, wobei Quellen positive und Senken negative Divergenz haben. Ist die Divergenz überall Null, dann ist das Feld quellenfrei.
- Die Rotation eines Vektorfeldes lässt sich als Wirbeldichte eines Vektorfeldes interpretieren. Ist  $\mathbf{v}(\mathbf{x})$  ein Geschwindigkeitsfeld, dann gibt  $\nabla \times \mathbf{v}(\mathbf{x})$  das doppelte der Winkelgeschwindigkeit an, mit der sich ein mitschwimmender Körper am Ort  $\mathbf{x}$  dreht.
- Der Laplace-Operator taucht in der Regel dann auf, wenn die Divergenz eines Gradientenfeldes berechnet wird.

### Satz 11.40

Seien  $f : D \rightarrow \mathbb{R}$ ,  $D \subseteq \mathbb{R}^3$  ein  $\mathcal{C}^2$ -Skalarfeld und  $\mathbf{v} : D \rightarrow \mathbb{R}^3$  ein  $\mathcal{C}^1$ -Vektorfeld. Für Divergenz und Rotation gilt dann:

- $\operatorname{rot}(\nabla f) = \nabla \times \nabla f = 0$       “Gradientenfeld ist wirbelfrei”
- $\operatorname{div}(\operatorname{rot} \mathbf{v}) = \nabla \cdot (\nabla \times \mathbf{v}) = 0$       “Feld der Rotation ist quellenfrei”

ohne Beweis

### Beispiel 11.41

Die Maxwell’schen Gesetze (Grundlage der klassischen Elektrodynamik) lassen sich als System von vier sogenannten partiellen Differentialgleichungen formulieren, in denen Divergenz und Rotation vorkommen:

$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$	“Ladungen sind die Quellen des elektrischen Feldes”
$\nabla \cdot \mathbf{B} = 0$	“Das Feld der magnetischen Flussdichte ist quellenfrei, es gibt keine magnetischen Monopole.”
$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$	“Eine Änderung der magnetischen Flussdichte führt zu einem elektrischen Wirbelfeld”
$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$	“Elektrische Ströme und Änderungen des elektrischen Feldes führen zu einem magnetischen Wirbelfeld”

mit:

$\mathbf{E}$	elektrische Feldstärke
$\rho$	Ladungsdichte
$\epsilon_0$	Permittivität des Vakuums
$\mathbf{B}$	magnetische Flussdichte
$\mu_0$	magnetische Permeabilität
$\mathbf{j}$	elektrische Stromdichte

## 11.7. Die Integralsätze von Green, Stokes und Gauß

Es gibt drei wichtige Sätze über Integrale von Rotation und Divergenz, die insbesondere bei der Lösung partieller Differentialgleichungen von Bedeutung sind.

Zur Motivation des Satzes von Green, integrieren wir ein Vektorfeld entlang der Kanten eines Rechtecks. Wir teilen das Integral in eine Summe über Integrale für jede Kante und nähern das

Integral über eine Kante durch den Werte in der Mitte der Kante mal die Länge der Kante.  
Den Wert in der Mitte der Kante berechnen wir als lineare Näherung.

$$\oint_{\partial S} \mathbf{v} \cdot d\mathbf{s} = \sum_{i=1}^4 \left[ \int_{\text{Kante } i} \mathbf{v} \cdot d\mathbf{s} \right]$$

$$\approx \underbrace{\left( \mathbf{v}(\mathbf{x}) - \mathbf{J}_{\mathbf{v}} \frac{\Delta x_2}{2} \mathbf{e}_2 \right) \cdot \underbrace{\mathbf{e}_1}_{\text{Tangentialvektor}} \cdot \underbrace{\Delta x_1}_{\text{Kantenlänge}}}_{\text{unten}} + \underbrace{\left( \mathbf{v}(\mathbf{x}) + \mathbf{J}_{\mathbf{v}} \frac{\Delta x_1}{2} \mathbf{e}_1 \right) \cdot \mathbf{e}_2 \cdot \Delta x_2}_{\text{rechts}}$$

$$+ \underbrace{\left( \mathbf{v}(\mathbf{x}) + \mathbf{J}_{\mathbf{v}} \frac{\Delta x_2}{2} \mathbf{e}_2 \right) \cdot (-\mathbf{e}_1) \cdot \Delta x_1}_{\text{oben}} + \underbrace{\left( \mathbf{v}(\mathbf{x}) - \mathbf{J}_{\mathbf{v}} \frac{\Delta x_1}{2} \mathbf{e}_1 \right) \cdot (-\mathbf{e}_2) \cdot \Delta x_2}_{\text{links}}$$

$$\text{mit } (\mathbf{J}_{\mathbf{v}} \cdot \mathbf{e}_i) \cdot \mathbf{e}_j = \left( \begin{pmatrix} \frac{\partial v_1}{\partial x_1} & \frac{\partial v_1}{\partial x_2} & \frac{\partial v_1}{\partial x_3} \\ \frac{\partial v_2}{\partial x_1} & \frac{\partial v_2}{\partial x_2} & \frac{\partial v_2}{\partial x_3} \\ \frac{\partial v_3}{\partial x_1} & \frac{\partial v_3}{\partial x_2} & \frac{\partial v_3}{\partial x_3} \end{pmatrix} \cdot \mathbf{e}_i \right) \cdot \mathbf{e}_j = \begin{pmatrix} \frac{\partial v_1}{\partial x_i} \\ \frac{\partial v_2}{\partial x_i} \\ \frac{\partial v_3}{\partial x_i} \end{pmatrix} \cdot \mathbf{e}_j = \frac{\partial v_j}{\partial x_i}$$

$$= \left( -\frac{1}{2} \frac{\partial v_1}{\partial x_2} + \frac{1}{2} \frac{\partial v_2}{\partial x_1} - \frac{1}{2} \frac{\partial v_1}{\partial x_2} + \frac{1}{2} \frac{\partial v_2}{\partial x_1} \right) \Delta x_1 \Delta x_2$$

$$= \left( \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \right) \Delta A$$

### Satz 11.42 (Satz von Green)

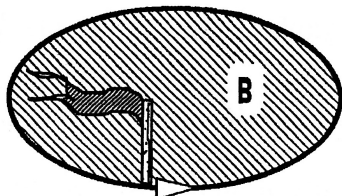
Sei  $B$  ein regulärer Bereich, dessen Rand  $\partial B$  aus endlich vielen geschlossenen, stückweise glatten Kurven besteht. Die Parametrisierung sei so, daß  $B$  stets links zur Durchlaufung liegt (positiver Umlauf). Sei  $D \subseteq \mathbb{R}^2$  eine offene Menge mit  $B \subset D$ , dann gilt

1. für zwei  $\mathcal{C}^1$ -Skalarfelder  $f : D \rightarrow \mathbb{R}$  und  $g : D \rightarrow \mathbb{R}$

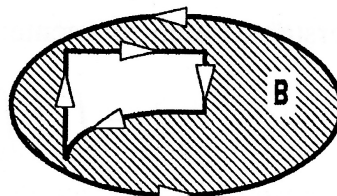
$$\oint_{\partial B} f dx_1 + g dx_2 = \iint_B \left( \frac{\partial g}{\partial x_1} - \frac{\partial f}{\partial x_2} \right) dx_1 dx_2$$

2. insbesondere gilt für jedes ebene  $\mathcal{C}^1$ -Vektorfeld  $\mathbf{v} : D \rightarrow \mathbb{R}^2$

$$\oint_{\partial B} \mathbf{v} \cdot d\mathbf{s} = \iint_B \left( \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \right) dx_1 dx_2$$



$\partial B = \mathbf{w}_1$



$\partial B = \mathbf{w}_1 \cup \mathbf{w}_2$

Flächen mit positiv umlaufenden Rand.

Quelle: Meyberg, Vachenauer, Höhere Mathematik 1, Springer

**Beispiel 11.43**

Mit  $\mathbf{v} = (0, x_1)^T$  bzw.  $\mathbf{v} = (-x_2, 0)^T$  ergibt sich die Fläche  $A_B$  von  $B$  als

$$\begin{aligned} A_B &= \iint_B dx_1 dx_2 = \iint_B (1 - 0) dx_1 dx_2 = \int_{\partial B} x_1 dx_2 \\ &= \iint_B (0 - (-1)) dx_1 dx_2 = - \int_{\partial B} x_2 dx_1 \\ &= \frac{1}{2} \int_{\partial B} -x_2 dx_1 + x_1 dx_2 \end{aligned}$$

Die Fläche lässt sich also rein durch Integration über den Rand bestimmen.

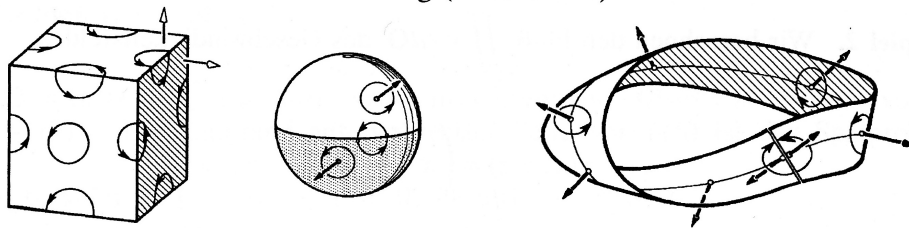
Oberflächen im  $\mathbb{R}^3$  sind zweiseitig. Wir brauchen deshalb für eine Verallgemeinerung des Satzes von Green eine Definition der Richtung von Flächen.

**Definition 11.44**

Eine stückweise reguläre Fläche  $S = \bigcup_{1 \leq i \leq k} S_i$  heißt zweiseitig oder orientierbar, wenn sich die Oberseite aller Flächenstücke  $S_i$  so wählen lassen, dass sich der Umlaufsinn über die “Kanten”  $S_i \cap S_j$  hinweg “stetig fortsetzt”. Die Oberseite wird durch einheitliche Wahl des Vorzeichens von

$$\mathbf{n} = \pm \frac{\mathbf{x}_u \times \mathbf{x}_v}{\|\mathbf{x}_u \times \mathbf{x}_v\|}$$

auf jedem Flächenstück  $S_i$  bestimmt und ist die Seite aus der die Normalen heraus zeigen.



Zweiseitige Flächen (links), ein nicht-zweiseitiges bzw. nicht orientierbares Möbius-Band (rechts).

Quelle: Meyberg, Vachenauer, Höhere Mathematik 1, Springer

**Satz 11.45 (Satz von Stokes)**

Sei  $S$  eine zweiseitige, stückweise reguläre Fläche mit überschneidungsfreier und geschlossener Randkurve  $\partial S$ , die so durchlaufen wird, dass  $S$  “links” liegt und der Umlaufsinn zusammen mit der Normalenrichtung  $\mathbf{n}$  von  $S$  eine Rechtsschraubung ergibt. Ferner sei  $U \subseteq \mathbb{R}^3$  eine offene Menge die  $S$  enthält. Dann gilt für alle  $\mathcal{C}^1$ -Vektorfelder  $\mathbf{v} : U \rightarrow \mathbb{R}^3$

$$\oint_{\partial S} \mathbf{v} \cdot d\mathbf{s} = \iint_S \text{rot } \mathbf{v} \cdot \mathbf{n} dO = \iint_S (\nabla \times \mathbf{v}) \cdot \mathbf{n} dO$$



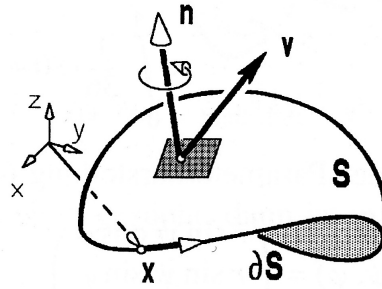


Illustration des Satz von Stokes.

Quelle: Meyberg, Vachenauer, Höhere Mathematik 1, Springer

#### Anmerkung 11.46

Bei einer Rechtsschraubung liegt die Fläche immer links von der Umlaufrichtung, wenn die Normale nach oben zeigt. Dies entspricht der sogenannten Rechte-Hand-Regel (Wenn der Daumen in Richtung der Normale weist, zeigen die Finger in Drehrichtung).

Das Linienintegral ist die Netto-Komponente der Tangentialkomponente von  $\mathbf{v}$  entlang einer geschlossenen Kurve und quantifiziert daher die Wirbelstärke. Bei positivem Integral ergibt die Drehrichtung des Wirbels mit der Normalenrichtung auch eine Rechtsschraubung.

Beachte: Das Ergebnis des Linienintegrals ist völlig unabhängig von der Form der Fläche, solange der Rand der Fläche gleich bleibt.

#### Beispiel 11.47

Gegeben sei die Paraboloidfläche

$$z = 4 - x^2 - y^2, \quad 0 \leq z \leq 4, \quad \mathbf{e}_z \cdot \mathbf{n} > 0$$

und das Vektorfeld

$$\mathbf{v}(x, y, z) = \begin{pmatrix} y \\ yz \\ y^2 - x^2 \end{pmatrix}$$

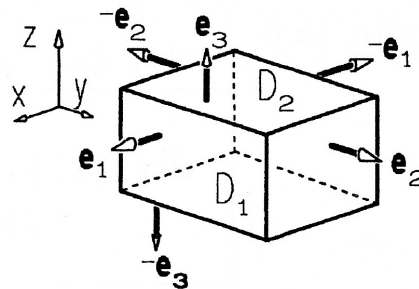
Berechne das Integral über die Paraboloidfläche von  $\nabla \times \mathbf{v}$ :

$$\iint_F \nabla \times \mathbf{v} \, d\mathbf{O} \stackrel{\text{Satz von Stokes}}{=} \int_{\partial F} \mathbf{v} \cdot d\mathbf{s}$$

Die Randkurve ist gegeben durch  $\mathbf{x} = (2 \cos t, 2 \sin t, 0)^T$ ,  $0 \leq t \leq 2\pi$ .

$$\begin{aligned} \int_{\partial F} \mathbf{v} \cdot d\mathbf{s} &= \int_0^{2\pi} \dot{\mathbf{x}}(t) \cdot \mathbf{v} \, dt = \int_0^{2\pi} \begin{pmatrix} -2 \sin t & 2 \cos t & 0 \end{pmatrix} \cdot \begin{pmatrix} 2 \sin t \\ 0 \\ 4 (\sin^2 t - \cos^2 t) \end{pmatrix} dt \\ &= \int_0^{2\pi} -4 \sin^2 t \, dt \\ &\quad \text{Nebenrechnung:} \\ &\quad \int \sin^2 t \, dt = t \sin^2 t - \int t \cdot 2 \sin t \cos t \, dt = t \sin^2 t - \int t \sin 2t \, dt \\ &\quad = t \sin^2 t - t \left( -\frac{1}{2} \cos 2t \right) - \frac{1}{2} \int \cos 2t \, dt \\ &\quad = t \cdot \left[ \sin^2 t + \frac{1}{2} (1 - 2 \sin^2 t) \right] - \frac{1}{4} \sin 2t + C \\ &\quad = \frac{t}{2} - \frac{1}{4} \sin 2t + C \\ &= \left[ -2t + \sin(2t) \right]_0^{2\pi} = -4\pi \end{aligned}$$

Zur Motivation des Gauß'schen Satzes bilanzieren wir Normalenflüsse über die Seiten eines Quaders. Wir summieren die Integrale über die Seitenflächen und nähern das Integral für eine Seite durch das Produkt aus dem Wert im Mittelpunkt der Seite mit der Fläche. Den Wert im Mittelpunkt der Seite nähern wir linear ausgehend vom Mittelpunkt des Quaders  $\mathbf{x}$ .



Bilanzierungsvolumen zur Ableitung der Divergenz.  
 Quelle: Meyberg, Vachenauer, Höhere Mathematik 1, Springer

$$\begin{aligned}
\iint_S \mathbf{v} \cdot d\mathbf{O} &= \sum_{i=1}^6 \left[ \iint_{\text{Seite } i} \mathbf{v} \cdot \mathbf{n} dA \right] \\
&\approx \underbrace{\left( \mathbf{v}(\mathbf{x}) + \mathbf{J}_v \frac{\Delta x_1}{2} \mathbf{e}_1 \right)}_{\text{rechts}} \cdot \underbrace{\mathbf{e}_1}_{\text{Normalenvektor}} \cdot \underbrace{\Delta x_2 \Delta x_3}_{\text{Seitenfläche}} \\
&\quad + \underbrace{\left( \mathbf{v}(\mathbf{x}) - \mathbf{J}_v \frac{\Delta x_1}{2} \mathbf{e}_1 \right)}_{\text{links}} \cdot (-\mathbf{e}_1) \cdot \Delta x_2 \Delta x_3 \\
&\quad + \underbrace{\left( \mathbf{v}(\mathbf{x}) + \mathbf{J}_v \frac{\Delta x_2}{2} \mathbf{e}_2 \right)}_{\text{hinten}} \cdot \mathbf{e}_2 \cdot \Delta x_1 \Delta x_3 + \underbrace{\left( \mathbf{v}(\mathbf{x}) - \mathbf{J}_v \frac{\Delta x_2}{2} \mathbf{e}_2 \right)}_{\text{vorne}} \cdot (-\mathbf{e}_2) \cdot \Delta x_1 \Delta x_3 \\
&\quad + \underbrace{\left( \mathbf{v}(\mathbf{x}) + \mathbf{J}_v \frac{\Delta x_3}{2} \mathbf{e}_3 \right)}_{\text{oben}} \cdot \mathbf{e}_3 \cdot \Delta x_1 \Delta x_2 + \underbrace{\left( \mathbf{v}(\mathbf{x}) - \mathbf{J}_v \frac{\Delta x_3}{2} \mathbf{e}_3 \right)}_{\text{unten}} \cdot (-\mathbf{e}_3) \cdot \Delta x_1 \Delta x_2 \\
&\quad \text{mit } (\mathbf{J}_v \cdot \mathbf{e}_i) \cdot \mathbf{e}_i = \left( \begin{pmatrix} \frac{\partial v_1}{\partial x_1} & \frac{\partial v_1}{\partial x_2} & \frac{\partial v_1}{\partial x_3} \\ \frac{\partial v_2}{\partial x_1} & \frac{\partial v_2}{\partial x_2} & \frac{\partial v_2}{\partial x_3} \\ \frac{\partial v_3}{\partial x_1} & \frac{\partial v_3}{\partial x_2} & \frac{\partial v_3}{\partial x_3} \end{pmatrix} \cdot \mathbf{e}_i \right) \cdot \mathbf{e}_i = \begin{pmatrix} \frac{\partial v_1}{\partial x_1} \\ \frac{\partial v_2}{\partial x_2} \\ \frac{\partial v_3}{\partial x_3} \end{pmatrix} \cdot \mathbf{e}_i = \frac{\partial v_i}{\partial x_i} \\
&= \left( \frac{1}{2} \frac{\partial v_1}{\partial x_1} + \frac{1}{2} \frac{\partial v_1}{\partial x_1} + \frac{1}{2} \frac{\partial v_2}{\partial x_2} + \frac{1}{2} \frac{\partial v_2}{\partial x_2} + \frac{1}{2} \frac{\partial v_3}{\partial x_3} + \frac{1}{2} \frac{\partial v_3}{\partial x_3} \right) \Delta x_1 \Delta x_2 \Delta x_3 \\
&= \left( \frac{\partial v_1}{\partial x_1} + \frac{\partial v_2}{\partial x_2} + \frac{\partial v_3}{\partial x_3} \right) \Delta x_1 \Delta x_2 \Delta x_3 \\
&= (\nabla \cdot \mathbf{v}) \cdot \Delta V
\end{aligned}$$

### Satz 11.48 (Satz von Gauß)

Sei  $\Omega \subseteq \mathbb{R}^n$  eine kompakte Menge mit abschnittsweise glattem Rand  $\partial\Omega$ . Der Rand sei orientiert, wobei  $\mathbf{n}$  die aus  $\Omega$  nach außen weisende Einheitsnormale ist. Sei  $\mathbf{v} : U \rightarrow \mathbb{R}^n$  auf einer offenen Menge  $U \in \mathbb{R}^n$  mit  $\Omega \subseteq U$  ein  $\mathcal{C}^1$ -Vektorfeld. Dann gilt

$$\iiint_{\Omega} \nabla \cdot \mathbf{v} dV = \iint_{\partial\Omega} (\mathbf{v} \cdot \mathbf{n}) dO$$

Der ‘Gesamtfluss’ durch alle Seiten eines geschlossenen Volumens entspricht der Summe der Quellen (oder bei negativem Vorzeichen Senken) in dem Volumen.

### Beispiel 11.49

- Aus Maxwell’schen Gesetzen:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$$

Bilde Oberflächenintegral der Normalenkomponente des Feldes über eine beliebige geschlossene Fläche  $\partial\Omega$ , die ein Gebiet  $\Omega$  umschließt:

$$\iint_{\partial\Omega} \mathbf{E} \cdot \mathbf{n} dO = \iiint_{\Omega} \nabla \cdot \mathbf{E} dV = \iiint_{\Omega} \frac{\rho}{\epsilon_0} dV = \frac{1}{\epsilon_0} \iiint_{\Omega} \rho dV$$

Das Integral der elektrischen Feldstärke über eine geschlossene Fläche entspricht also immer der eingeschlossenen Gesamtladung.

Entsprechend erhält man durch Integration über das entsprechende Gesetz für die magnetische Flussdichte:

$$\iint_{\partial\Omega} \mathbf{B} \cdot \mathbf{n} dO = \iiint_{\Omega} \nabla \cdot \mathbf{B} dV = 0$$

- Energieerhaltung bei der Wärmeleitung

$$\frac{\partial (C_v T)}{\partial t} = -\nabla \cdot (-\lambda \nabla T)$$

mit  $C_v$  volumetrische Wärmekapazität,  $\lambda$  Wärmeleitfähigkeit.

Dies ist eine partielle Differentialgleichung 2. Ordnung.

Integration über ein Gebiet  $\Omega$  liefert:

$$\iiint_{\Omega} \frac{\partial (C_v T)}{\partial t} dV = \iiint_{\Omega} \nabla \cdot (\lambda \nabla T) dV = \iint_{\partial\Omega} \lambda \nabla T \cdot \mathbf{n} dO$$

## 11.8. Partielle Differentialgleichungen

### Definition 11.50

Eine Gleichung zur Bestimmung einer Funktion  $u : \Omega \rightarrow \mathbb{R}$ ,  $\Omega \subseteq \mathbb{R}^n$ ,  $u \in \mathcal{C}^m$  in der neben der Funktion  $u(\mathbf{x})$  und Ausdrücken in  $\mathbf{x}$  auch partielle Ableitungen von  $u$  bis zur  $m$ -ten Ordnung vorkommen, heißt partielle Differentialgleichung (PDGL oder PDE)  $m$ -ter Ordnung. Allgemein kann man diese in der Form

$$F\left(\frac{\partial^m u}{\partial x_1^m}(\mathbf{x}), \frac{\partial^m u}{\partial x_1^{m-1} \partial x_2}(\mathbf{x}), \dots, \frac{\partial^m u}{\partial x_n^m}(\mathbf{x}), \dots, \frac{\partial^{m-1} u}{\partial x_1^{m-1}}(\mathbf{x}), \dots, u(\mathbf{x})\right) = 0 \quad \forall \mathbf{x} \in \Omega \quad (11.1)$$

schreiben.

Eine  $m$ -mal partiell differenzierbare Funktion  $u : \Omega \rightarrow \mathbb{R}$ ,  $\Omega \subseteq \mathbb{R}^n$  heißt Lösung der PDGL, wenn sie die PDGL in allen Punkten  $\mathbf{x} \in D$  erfüllt. Zur eindeutigen Festlegung von  $u$  sind noch „zusätzliche Bedingungen“ erforderlich, sogenannte Randbedingungen (bei Randbedingungen in der Zeit spricht man von Anfangsbedingungen).

Lässt sich die PDGL als Linearkombination partieller Ableitungen schreiben, bei der die Koeffizienten vor den partiellen Ableitungen nicht von  $u$  abhängen, sprechen wir von einer linearen partiellen Differentialgleichung.

Ist  $\mathbf{u}$  eine vektorwertige Funktion bei der eine partielle Differentialgleichung für jede Komponente des Vektors angegeben ist, dann sprechen wir von einem System partieller Differentialgleichungen.

**Beispiel 11.51**

1. Für Raumdimension  $n = 2$  und Ordnung  $m = 2$  lautet die allgemeine lineare PDGL:

$$\underbrace{a(x, y) \frac{\partial^2 u}{\partial x^2}(x, y) + 2b(x, y) \frac{\partial^2 u}{\partial x \partial y}(x, y) + c(x, y) \frac{\partial^2 u}{\partial y^2}(x, y)}_{\text{Hauptteil}} + d(x, y) \frac{\partial u}{\partial x}(x, y) + e(x, y) \frac{\partial u}{\partial y}(x, y) + f(x, y)u(x, y) + g(x, y) = 0 \quad \text{in } \Omega. \quad (11.2)$$

Die ersten drei Terme stellen den sog. „Hauptteil“ der Gleichung dar.

2. Die Maxwell'schen Gesetze bilden ein System von vier gekoppelten partiellen Differentialgleichungen.

Die Lösung von partiellen Differentialgleichungen erfolgt meist numerisch. Näheres hierzu in Ingenieurmathematik IV.