MATSE

Ausbildung

FH AACHEN

Lineare Algebra 1 & 2

Jacqueline Gottowik (gottowik@itc.rwth-aachen.de)

Matthias Grajewski (grajewski@fh-aachen.de)

Ruth Partzsch (r.partzsch@fz-juelich.de)

auf Grundlage eines Skripts von H. J. Pflug

01. September 2024

Inhaltsverzeichnis

Kapitel 1

Motivation und Vorbereitung

1.1 Inhalt und Anwendungen der Linearen Algebra

- Lineare Algebra stellt grundlegende Techniken und Begriffe bereit
- Lineare Algebra durchsetzt die gesamte Mathematik
- drei wesentliche Themenbereiche:

1. lineare Gleichungssysteme
2. analytische Geometrie
3. algebraischen Strukturen

- drei Beispiele zur Motivation

Beispiel 1.1:

- Einschmelzen von Schrott zur Eisenerzeugung
- kein chemisch reines Eisen: Metalle wie Kupfer, Zink u. a. enthalten
- Legierung mit vorgegebenen Anteilen an Fremdmetallen soll erschmolzen werden
- drei Schrottsorten S_1, S_2 und S_3:

	S_1	S_2	S_3
Eisen	$0,9$	$0,8$	$0,8$
Kupfer	0	$0,2$	$0,1$
Zink	$0,1$	0	$0,1$

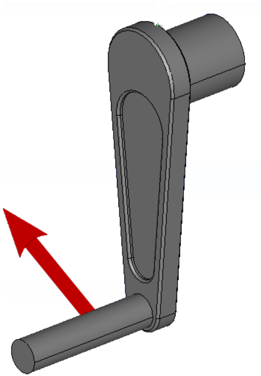
- Endprodukt: 80% Eisen, 12% Kupfer und 8% Zink
- Massebilanz für Eisen: $0,9a + 0,8b + 0,8c = 0,8$
- Massebilanz für Kupfer: $0a + 0,2b + 0,1c = 0,12$

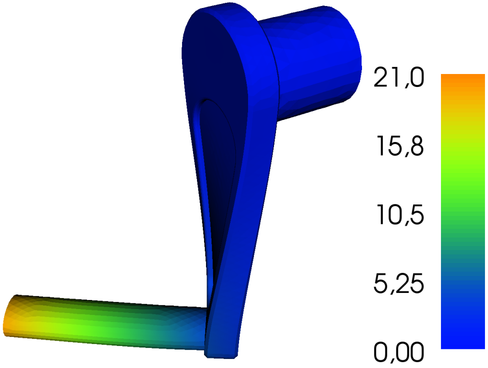
- Massebilanz für Zink: $0,1a + 0b + 0,1c = 0,08$
- lineares Gleichungssystem:

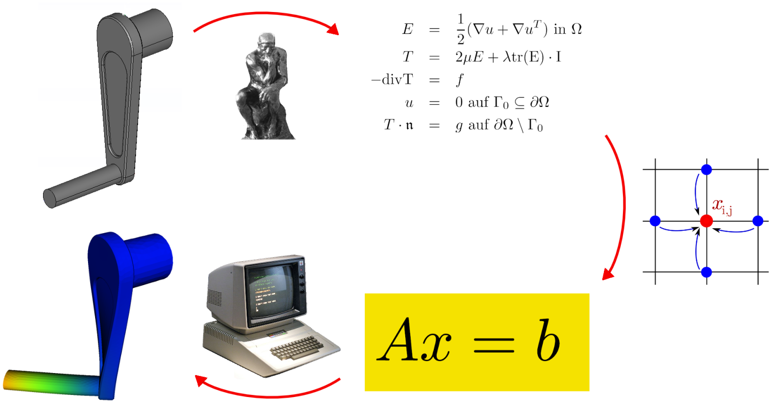
$$\begin{aligned} 0,9a + 0,8b \; &+ 0,8c = 0,8 \\ 0a + 0,2b \; &+ 0,1c = 0,12 \\ 0,1a + \quad 0b \; &+ 0,1c = 0,08 \end{aligned} \tag{1.1}$$

- Lösung ist $a = 0$, $b = 0,2$ und $c = 0,8$ (Nachweis durch Einsetzen)

Beispiel 1.2:







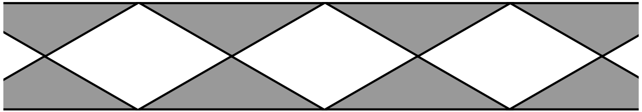
- Numerische Simulation: Verhaltensvorhersage eines physikalischen Systems mithilfe mathematischer Methoden

- Auto, Erde, Atom, mech. Bauteil , z. B. Kurbel
- herkömmlicher Entwicklungsprozess:
  - Entwurf, dann Bau und Test von Prototypen
  - Iteration bis zum Erfolg
  - Serienfertigung
- Numerische Simulation ersetzt Bau und Test von Prototypen durch Berechnungen
- Zeit- und Geldersparnis
- unverzichtbares Werkzeug im Maschinenbau

- Numerische Simulation z. B. mit der Methode der finiten Elemente
- Zusammenwirken vieler verschiedener mathematischer Disziplinen

- am Ende häufig: Lösen eines linearen Gleichungssystems
- 1000000 Gleichungen und Unbekannte nicht ungewöhnlich
- effiziente Lösung durch massiven Computereinsatz und hochentwickelte Algorithmen
- Numerische Lineare Algebra
- Voraussetzung dazu: Lineare Algebra
- ohne Lineare Algebra keine Numerische Simulation

Beispiel 1.3:



- bekannt: Gruppe $(M, \circ)$
- Fries: in eine Raumrichtung unendlich ausgedehnt, eben, periodisch
- Starrkörperbewegung: bijektive, längen- und winkelerhaltende Abbildung der Ebene in sich
- die Starrkörperbewegungen, die das Fries in sich überführen, bilden eine Gruppe bzgl. der Verkettung
- Automorphismengruppe („Friesgruppe”)

- zu jedem Fries-Typ gehört genau eine Automorphismengruppe und umgekehrt
- Untersuchung mithilfe der Gruppentheorie: es gibt im Wesentlichen nur 7 verschiedene Friesgruppen
- bis auf Farbe etc: es gibt nur 7 verschiedene Friese
- unendlich ausgedehnte periodische Objekte im Raum: Modell für Kristalle
- Fedorov: es existieren genau 230 Raumgruppen
- Durchbruch in der Kristallographie

1.2 Mathematisches Arbeiten und Problemlösen

- zentral: systematische Beschäftigung mit mathematischen Ideen, Problemen und ihrer Lösung
- Lösung durch strengen Beweis oder Gegenbeispiel
- in LA: Formulierung und Beweis mathematischer Aussagen sehr wichtig
- Unterschied z. B. zur Schule
- Lösungen mathematischer Probleme liegen oftmals nicht auf der Hand
- jetzt: grundlegende Lösungstechniken

- Situation: mathematische Aussage liegt vor. Beweisen oder Widerlegen Sie!

- Standardanforderung im Studium
- in der Praxis: (zielführende) Fragen sind oft schwieriger zu finden als Antworten

1. **Versuchen Sie, das Problem zu verstehen!**
    - Sind Ihnen alle verwendeten mathematischen Begriffe in der Formulierung klar? Könnten Sie einem Kommilitonen alle verwendeten mathematischen Begriffe erklären? Wenn nicht, wiederholen Sie zunächst die entsprechenden Inhalte der Vorlesung.
    - Sollte die Aussage eine Formel sein, versuchen Sie, dieselbe Aussage für Sie selbst als Text zu formulieren und umgekehrt.
    - Betrachten Sie Beispiele und Spezialfälle. Kann man an Beispielen bereits erkennen, warum die Aussage wahr oder falsch sein sollte?
    - Fertigen Sie, wann immer das Problem es ermöglicht, eine Skizze an!
2. **Kennen Sie ähnliche Probleme?**

- Kennen Sie bei vergleichbaren Aussagen sogar Beweis oder Gegenbeispiel?
- Kann man die gegebene Aussage auf Bekanntes zurückführen?
- Wenn nicht, was genau ist anders?

3. **Vorwärtsarbeiten**

   - Jeder Beweisversuch lebt von den Voraussetzungen. Was lässt sich mit den gegebenen Voraussetzungen anfangen? Welche Aussagen lassen sich damit zeigen?

4. **Rückwärtsarbeiten**

   - Unter welchen zusätzlichen Voraussetzungen könnten Sie denn die gewünschte Aussage beweisen?
   - Wie kann man sich im zweiten Schritt von diesen zusätzlichen Voraussetzungen befreien?

5. **Zwischenziele formulieren**

- Bei komplexeren Sachverhalten kann es helfen, dass man das Gesamtproblem in Teilprobleme zerlegt, die man dann getrennt bearbeitet.
- Für jede Etappe lassen sich die oben skizzierten Problemlösestrategien verwenden.

6. **Problemlösestrategien kombinieren und ausprobieren**

   - In vielen Fällen bringt erst eine Kombination der obigen Strategien den Erfolg.
   - Wenn eine Strategie nicht weiterführt, muss man eine andere ausprobieren. Mathematik bedeutet manchmal auch hartnäckiges Herumprobieren!

7. **Zum Schluss: richtig Aufschreiben!**

   - Schreiben Sie ihre Argumentation detailliert und nachvollziehbar auf. Oftmals wird umgekehrt die Argumentation erst bei ihrer Formulierung wirklich klar.
   - Vermeiden Sie umständliche Prosa, sondern bedienen Sie sich mathematischer Formeln und Formulierungen.

- Kontrollieren Sie zum Schluss: Ist Ihre Beweisführung lückenlos? Ist das Gegenbeispiel wirklich eins?
- Für weiterführende Literatur zur Formulierung mathematischer Gedanken verweisen wir auf Beutelspacher [1].

1.3 Vektoren im $\mathbb{R}^n$

- Durchbruch in der Geometrie: Einführung von Koordinaten durch Descartes
- dadurch: geometrische Aussagen können mit Methoden der Arithmetik untersucht werden und umgekehrt

Definition 1.4:

1. Für $n \in \mathbb{N}$ sei x das n-Tupel

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (x_i)_{i=1}^n$$

mit $x_i \in \mathbb{R}$ für $1 \leq i \leq n$. Dann heißt x *Vektor*, die Zahl x_i die i-te *Koordinate*.

2. Zu x wie oben sei

$$x^T := (x_1, \ldots, x_n)$$

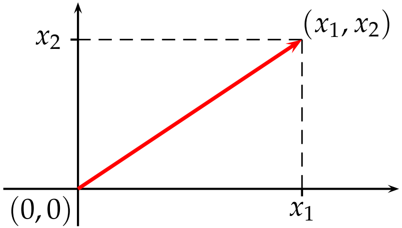
der *transponierte Vektor*. Weiter definieren wir $(x^T)^T := x$.

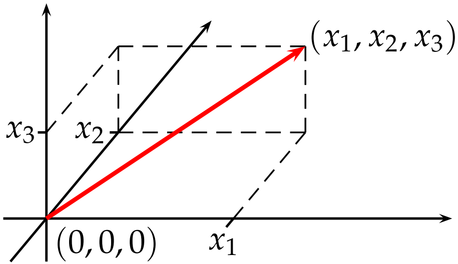
3. Zwei Vektoren x, y sind *gleich*, wenn alle ihre Koordinaten gleich sind.

4. Der Zahlraum
$$\mathbb{R}^n = \{(x_1, \dots, x_n)^T \,|\, x_1, \dots, x_n \in \mathbb{R}\}$$
sei die Menge aller geordneten n-Tupel reeller Zahlen.

- Fall $n = 2$
- Wahl eines Koordinatensystems
- dann: jeder Punkt der Ebene kann durch (x_1, x_2) parametrisiert werden
- umgekehrt: (x_1, x_2) lässt sich als Punkte der Ebene veranschaulichen
- man identifiziert die Ebene mit $\mathbb{R}^2$

- analog: $n = 3$





- Zahlraum dient auch zur Beschreibung nicht-geometrischer Zusammenhänge

Beispiel 1.5:

- Wirtschaftswissenschaften: Koordinaten geben die Gesamtproduktion von Waren wieder

- Maßstab: Preis (z. B. in 10^9 \$)
- 5 Wirtschaftszweige:

1.	Schwerindustrie	2.	Elektronik und Computer
3.	Automobilindustrie	4.	Landwirtschaft
5.	Finanzsektor		

- Modellierung der wirtschaftlichen Vorgänge im $\mathbb{R}^5$
- wir nehmen die obige Reihenfolge der Koordinaten an und betrachten

$$(44, 182, 404, 36, 230)^T,$$

- Bedeutung: z. B. Schwerindustrie produzierte Waren im Wert von 44 Milliarden \$
- Finanzsektor: Umsatz von 230 Milliarden \$

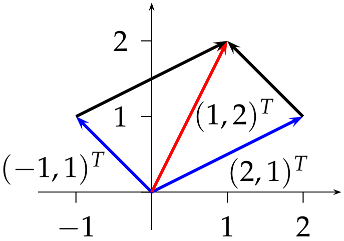
- jetzt: geometrische Deutung von Vektoren
- Interpretation als Beschreibung eines Punktes erfordert einen Koordinatenursprung
- Wo soll der sein?
- willkürliche Festlegung, da eigentlich alle Punkte gleichberechtigt
- daher: Deutung als Verschiebung
- eigentlich: eine Verschiebung ist eine Abbildung
- eindeutig festgelegt, wenn zu P sein Bildpunkt Q bekannt ist
- eindeutige Repräsentation einer Verschiebung durch gerichtete Strecke $\overrightarrow{PQ}$
- aber: ein anderer Punkt mit dessen Bildpunkt repräsentiert die Verschiebung genauso

- Alle gerichteten Strecken, die dieselbe Verschiebung repräsentieren, bilden eine Äquivalenzklasse
- in der euklidischen Geometrie: Das ist ein Vektor!
- ein (Verschiebungs)vektor hat Länge und Richtung, aber keine feste Lage
- Sei nun doch ein Koordinatenursprung 0 gegeben
- wähle $P = 0$
- damit wählt man aus der Äquivalenzklasse einen Vertreter aus
- der zeigt auf festen Punkt Q
- *Ortsvektor*; Schreibweise: $\overrightarrow{0Q}$ oder einfach q
- jetzt: seien bestimmte Richtungen ausgezeichnet als Koordinatenachsen

- dann: man kann Q über seine Koordinaten durch ein n-Tupel repräsentieren
- Das n-Tupel repräsentiert zugleich die Verschiebung und den Verschiebungsvektor
- Deutung eines n-Tupels abhängig vom Zusammenhang

1.4 Addition und Multiplikation im $\mathbb{R}^n$

- wir deuten $a = (2, 1)$ und $b = (-1, 1)$ als Verschiebungen in der Ebene
- Verschiebung längs a und dann längs b: Verschiebung um $c = (1, 2)$

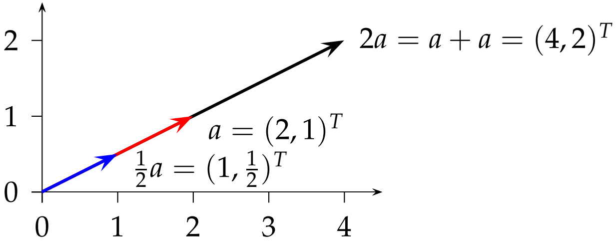


- analog im $\mathbb{R}^3$

Definition 1.6: Seien $a = (a_1, \ldots, a_n)^T, b = (b_1, \ldots, b_n)^T \in \mathbb{R}^n$. Dann ist

$$a + b := \begin{pmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{pmatrix}.$$

- verlängert man $a = (2, 1)$ in der Ebene um den Faktor 2, erhält man $(4, 2)$



Definition 1.7: Für $\lambda \in \mathbb{R}$ und $a = (a_1, \dots, a_n)^T \in \mathbb{R}^n$ sei

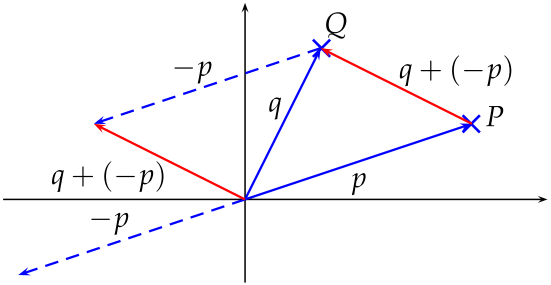
$$\lambda a := \begin{pmatrix} \lambda a_1 \\ \vdots \\ \lambda a_n \end{pmatrix}.$$

- Für $0 < \lambda \in \mathbb{R}$: λa entspricht Punkt mit derselben Richtung wie a zum Ursprung, aber λ-fachen Abstand
- für λ negativ: Richtungsumkehr

Definition 1.8: Für $a, b \in \mathbb{R}^n$ sei

$$a - b := a + (-1)b$$

der *Differenzvektor* von a und b.



- geometrische Deutung von $q-p$: Verschiebungsvektor von P nach Q
- $p-q$ ist der Verschiebungsvektor von Q nach P

Definition 1.9: Zwei Vektoren $a, b \neq 0$ heißen *parallel* (Schreibweise $a \parallel b$) $:\Longleftrightarrow \exists \alpha \in \mathbb{R} : a = \alpha b$. Gilt $\alpha > 0$, haben parallele Vektoren die *gleiche Richtung*, im Fall von $\alpha < 0$ *gegensätzliche Richtung*.
Der Nullvektor ist zu jedem Vektor parallel.

Beispiel 1.10: Die Vektoren $a = (1,2)^T$ und $b = (2,4)^T$ sind parallel und haben die gleiche Richtung; der Vektor $c = (1,3)^T$ ist weder zu a noch zu b parallel.

1.5 Lineare Gleichungssysteme und Matrizen

- lineares Gleichungssystem aus dem Beispiel der Eisenverhüttung:

$$\begin{aligned} 0,9a + 0,8b + 0,8c &= 0,8 \\ 0a + 0,2b + 0,1c &= 0,12 \\ 0,1a + \quad 0b + 0,1c &= 0,08 \end{aligned} \tag{1.2}$$

- Lösung hängt nur von den Koeffizienten ab
- zusammen mit der rechten Seite: alle Informationen über das LGS enthalten
- Kurzschreibweise: man notiert die Koeffizienten als *Koeffizientenmatrix*
- am Beispiel oben:

$$\begin{pmatrix} 0,9 & 0,8 & 0,8 \\ 0 & 0,2 & 0,1 \\ 0,1 & 0 & 0,1 \end{pmatrix}.$$

Definition 1.11: Eine $m \times n$- *Matrix* A ist ein rechteckiges Schema von reellen oder komplexen Zahlen a_{ij} (den *Elementen* der Matrix) mit m Zeilen und n Spalten

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad . \tag{1.3}$$

- i: Zeilenindex, j: Spaltenindex
- alternative Schreibweise von Matrizen: (a_{ij}) bzw. $(a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$
- Menge aller $m \times n$-Matrizen mit reellen bzw. komplexen Elementen: $\mathbb{R}^{m \times n}$ bzw. $\mathbb{C}^{m \times n}$
- $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

Definition 1.12:

1. Seien $A = (a_{ij}), B = (b_{ij}) \in \mathbb{K}^{m\times n}$. Man setzt $A = B :\Leftrightarrow a_{ij} = b_{ij}\, \forall 1 \le i \le m, 1 \le j \le n$.
2. Eine Matrix, deren Elemente alle den Wert 0 annehmen, heißt *Nullmatrix*.
3. Zu A wie oben sei
$$A^T := (a_{ji}) \in \mathbb{K}^{n\times m}$$
die *transponierte Matrix*. Man erhält also A^T, indem man die Spalten von A als Zeilen von A^T verwendet.

- allg. lineares Gleichungssystem mit m Gleichungen und n Unbekannten:

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + \cdots + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + \cdots + & a_{2n}x_n & = & b_2 \\ \cdots & + & \cdots & + \cdots + & \cdots & = & \cdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + \cdots + & a_{mn}x_n & = & b_m \end{array} \tag{1.4}$$

- Koeffizientenmatrix ist A (s. o.)

- keine vollständge Beschreibung des LGS, weil die rechte Seite fehlt
- man ergänzt A zu (A, b) (*erweiterte Koeffizientenmatrix*)
- Abtrennung der letzten Spalte durch senkrechten Strich üblich

Beispiel 1.13: Die erweiterte Koeffizientenmatrix des linearen Gleichungssystems aus Beispiel 1.1 lautet

$$\begin{pmatrix} 0,9 & 0,8 & 0,8 & 0,8 \\ 0 & 0,2 & 0,1 & 0,12 \\ 0,1 & 0 & 0,1 & 0,08 \end{pmatrix} \quad \text{bzw.} \quad \left(\begin{array}{ccc|c} 0,9 & 0,8 & 0,8 & 0,8 \\ 0 & 0,2 & 0,1 & 0,12 \\ 0,1 & 0 & 0,1 & 0,08 \end{array}\right).$$

Definition 1.14:

1. Eine $n \times n$-Matrix heißt *quadratisch*.
2. Eine quadratische Matrix A mit Elementen a_{ij} wird häufig in der Form $A = (a_{ij})_{i,j=1}^{n}$ geschrieben.

3. Sei $A = (a_{ij})_{i,j=1}^{n}$ quadratisch. Die Elemente von A mit $i = j$ bilden die *Hauptdiagonale* von A.

4. Eine quadratische Matrix, bei der alle Elemente oberhalb der Hauptdiagonalen gleich 0 sind, heißt *untere Dreiecksmatrix*.

5. Eine quadratische Matrix, bei der alle Elemente unterhalb der Hauptdiagonalen gleich 0 sind, heißt *obere Dreiecksmatrix*.

Beispiel 1.15: Die Matrix

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

ist eine untere Dreiecksmatrix.

Definition 1.16:

1. Sei $A \in \mathbb{K}^{n\times n}$. Gilt $a_{ij} = 0$ für alle $i \neq j$, heißt A *Diagonalmatrix*.

2. Die Diagonalmatrix

$$E_n := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \in \mathbb{K}^{n\times n}$$

heißt *Einheitsmatrix*.

Bemerkung 1.17:

- rechte Seite eines reellen linearen Gleichungssystems:

$$b = (b_1, \dots, b_m)^T \in \mathbb{R}^m$$

- Deutung als Punkt im Raum

- analog: Lösungsvektor $x \in \mathbb{R}^n$ entspricht Punkt im Raum

1.6 Lösung eines linearen Gleichungssystems

1.6.1 Das Gaußsche Eliminationsverfahren

- Gaußsches Eliminationsverfahren
- Idee: man überführt das gegebene LGS in ein anderes mit gleicher Lösung, dessen Lösung unmittelbar ablesbar ist

Definition 1.18:

1. Zwei Gleichungssysteme heißen *äquivalent*, falls sie die gleiche Lösungsmenge haben.

2. Zwei Matrizen A und B heißen äquivalent $(A \sim B)$, falls die entsprechenden Gleichungssysteme die gleiche Lösungsmenge besitzen.

Bemerkung 1.19: Äquivalenz im Sinne von Definition 1.18 ist eine *Äquivalenzrelation* (vgl. die Vorlesung „Mathematische Grundlagen”).

Beispiel 1.20: Das Gleichungssystem

$$\begin{array}{rrrl} 2x & +3y & -4z = & 8 \\ -4x & +2y & +3z = & -5 \\ 3x & +y & +2z = & 13 \end{array} \qquad (a)$$

ist äquivalent zu den Gleichungssystemen

$$\begin{array}{rrrl} 2x & +3y & -4z = & 8 \\ & 8y & -5z = & 11 \\ & & z = & 1 \end{array} \qquad (b)$$

und

$$\begin{array}{rl} x = & 3 \\ y = & 2 \\ z = & 1 \end{array} \qquad (c)$$

Auch die drei zugehörigen erweiterten Matrizen A, B, C sind dann äquivalent:

$$A = \begin{pmatrix} 2 & 3 & -4 & 8 \\ -4 & 2 & 3 & -5 \\ 3 & 1 & 2 & 13 \end{pmatrix}, B = \begin{pmatrix} 2 & 3 & -4 & 8 \\ 0 & 8 & -5 & 11 \\ 0 & 0 & 1 & 1 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

- (b): Stufenform
- (c): reduzierte Stufenform
- B: nur Nullen unterhalb der Hauptdiagonalen
- C: Eins aus der Hauptdiagonale, ansonsten Nulleinträge
- letzte Spalte: beliebige Werte

- Stufenform (Gauß-Algorithmus) übersichtlicher, Lösung einfacher zu bestimmen

- reduzierte Stufenform (Gauß-Jordan-Algorithmus): noch übersichtlicher; Lösung direkt ablesbar
- Ansatz: man bestimmt durch Umformungen ein äquivalentes Gleichungssystem in reduzierter Stufenform
- Äquivalenzumformungen der Zeilen:

Z1. Addition eines Vielfachen einer Zeile zu einer anderen.

Z2. Vertauschen zweier Zeilen.

Z3. Multiplikation einer Zeile mit einem Skalar $\lambda \neq 0, \lambda \in \mathbb{K}$.

- für die reduzierte Stufenform: zunächst Stufenform
- hauptsächlich durch Z1: Einträge unterhalb der Hauptdiagonalen zu Null

- Man beginnt in der ersten Spalte und arbeitet sich spaltenweise vor: *OHP*

- Stufenform nicht eindeutig
- nach Stufenform: zeilenweises Vorgehen
- Element aus der Hauptdiagonalen auf Eins bringen
- dann alle Elemente darüber auf Null
- im Beispiel: *OHP*

- Rücktransformation liefert die Lösung

$$x = 3$$
$$y = 2$$
$$z = 1$$

- Das Gaußsche Eliminationsverfahren führt in der Tat zu äquivalenten Gleichungssystemem (Beweis später)

- in der Praxis häufig: Berechnung mit Computern
- Programme hierzu: Matlab, python (http://www.python.org), Octave (http://www.gnu.org/software/octave/), ...
- interaktive Eingabe
- in Matlab
  - nach >> folgt die Eingabezeile
  - Ergebnis der Eingabezeile wird darunter ausgegeben
  - Ausgabe wird unterdrückt, indem man die Eingabezeile mit „;“ abschließt

Wir legen die Matrix zu Beispiel 1.20 in MATLAB an:

```
>> a=[1 2 3;4 5 6;7 8 9]
```

```
a =

1    2    3
4    5    6
7    8    9
```

- Eingabezeile ist nur die oberste Zeile
- Matrixzeilen werden durch Semikolon getrennt
- jetzt: Lösung eines linearen Gleichungssystems mit Matlab
- wir verwenden als eine Möglichkeit dazu den Befehl `rref`
- `format rat`: Ausgabe in Brüchen
- Matlab rechnet intern mit Fließkommazahlen

- Beispiel:

```
>> format rat
>> a=[2 3 -4 8;-4 2 3 -5;3 1 2 13]

a =

       2              3             -4              8
      -4              2              3             -5
       3              1              2             13

>> rref(a)

ans =

       1              0              0              3
```

```
0    1    0    2
0    0    1    1
```

- Ergebnis zurückübersetzt:

$$x = 3$$
$$y = 2$$
$$z = 1$$

1.6.2 Unter- und überbestimmte Gleichungssysteme

- Interpretation der reduzierten Stufenform bei über- oder unterbestimmten Gleichungssysteme
- *unterbestimmtes Gleichungssystem*: mehr als eine Lösung
- bei linearen Gleichungssystemen: dann sogar unendlich viele Lösungen
- *überbestimmtes Gleichungssystem*: keine Lösung
- Unterscheidung von drei Fällen:

1. Es existiert eine eindeutige Lösung.
2. Es existiert keine Lösung.
3. Es existieren unendlich viele Lösungen.

- die reduzierte Stufenform zeigt immer, welcher Fall vorliegt

Definition 1.21: Der erste Nicht-Null-Eintrag jeder Zeile einer Matrix heißt *Pivot-Element*. Eine Spalte, in der ein Pivot-Element vorkommt, heißt *Pivot-Spalte*.

- reduzierte Stufenform: Pivot-Element ist immer 1
- i. a. reduzierte Stufenform nicht eindeutig
- daher fordern wir zusätzlich:
- Kommen Nullzeilen vor, also Zeilen, die nur Nullen enthalten, dann müssen diese die untersten Zeilen der Matrix sein. Nullzeilen repräsentieren die Gleichung $0 = 0$ und können ignoriert werden.
- Die Pivot-Spalte der Zeile $i+1$ muss rechts der Pivot-Spalte der Zeile i liegen.
- Pivot-Spalten enthalten in der reduzierten Stufenform außer dem Pivot-Element nur Nullen. Diese Bedingung gilt nicht für die letzte Spalte der erweiterten Matrix.

- Forderungen lassen sich immer mit Z1 – Z3 erfüllen

Beispiel 1.22:

- Pivotelement: „P”; beliebiges Element: x

$$\begin{pmatrix} P & 0 & 0 & x \\ 0 & P & 0 & x \\ 0 & 0 & P & x \end{pmatrix}$$

- Matrix hat reduzierte Stufenform
- Alle Spalten bis auf die letzte sind Pivot-Spalten
- ebenso in reduzierter Stufenform:

$$\begin{pmatrix} P & 0 & x & 0 & x \\ 0 & P & x & 0 & x \\ 0 & 0 & 0 & P & x \\ 0 & 0 & 0 & 0 & P \end{pmatrix} \begin{pmatrix} P & 0 & x & x & 0 & x & x & x \\ 0 & P & x & x & 0 & x & x & x \\ 0 & 0 & 0 & 0 & P & x & x & x \end{pmatrix} \begin{pmatrix} P & 0 & 0 & x \\ 0 & P & 0 & x \\ 0 & 0 & P & x \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

- ab jetzt: die Spalten 1 bis $n-1$ der erweiteren Matrix heißen *Koeffizientenmatrix*
- die n-te Spalte: *letzte Spalte*
- im Fall einer eindeutigen Lösung erhält man immer die (reduzierte) Stufenform

$$\left(\begin{array}{ccc|c} P & x & x & x \\ 0 & P & x & x \\ 0 & 0 & P & x \end{array}\right) \quad \text{bzw.} \quad \left(\begin{array}{ccc|c} P & 0 & 0 & x \\ 0 & P & 0 & x \\ 0 & 0 & P & x \end{array}\right).$$

- letzte Spalte der Stufenform eine Pivot-Spalte $\Rightarrow$ Gleichungssystem *nicht lösbar*
- Beispiele:

$$\left(\begin{array}{cccc|c} P & x & x & x & x \\ 0 & P & x & x & x \\ 0 & 0 & 0 & 0 & P \end{array}\right) \quad \text{bzw.} \quad \left(\begin{array}{cccc|c} P & 0 & x & x & x \\ 0 & P & x & x & x \\ 0 & 0 & 0 & 0 & P \end{array}\right).$$

Beispiel 1.23: Die erweiterte Matrix sei

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

Die letzte Zeile lautet ausgeschrieben

$$0 \cdot x + 0 \cdot y + 0 \cdot z = 1\,.$$

- unerfüllbar für alle x, y, z
- lineares Gleichungssystem ist *unlösbar*

- in der (reduzierten) Stufenform weniger Pivotspalten als das Gleichungssystem Unbekannte hat $\Rightarrow$ mehrere Lösungen.
- Beispiele

$$\left(\begin{array}{cccc|c} P & x & x & x & x \\ 0 & P & x & x & x \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \quad \text{bzw.} \quad \left(\begin{array}{cccc|c} P & 0 & x & x & x \\ 0 & P & x & x & x \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

- *freie Parameter*: eine oder mehrere Variablen können beliebige Werte annehmen
- Lösungsregel: jede Nicht-Pivot-Spalte $\cong$ freie Variable

Beispiel 1.24:

$$\begin{aligned} 2\cdot x - 4\cdot y + 2\cdot z &= 8 \\ 1\cdot x - 1\cdot y - 7\cdot z &= 6 \end{aligned}$$

- Berechnung der äquivalenten reduzierten Stufenform mit Matlab:

```
>> format rat
>> a=[2 -4 2 8;1 -1 -7 6];
>> rref(a)

ans =

       1              0             -15                8
       0              1              -8                2
```

- Ergebnis als Gleichungssystem:

$$\begin{aligned} 1\cdot x + 0\cdot y - 15\cdot z &= 8 \\ 0\cdot x + 1\cdot y \quad - 8\cdot z &= 2 \end{aligned}$$

bzw.

$$\begin{aligned} x &= 8 + 15\cdot z \\ y &= 2 + 8\cdot z \end{aligned}$$

oder in Vektor-Schreibweise

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 8 \\ 2 \end{pmatrix} + z \begin{pmatrix} 15 \\ 8 \end{pmatrix}.$$

- Lösung noch nicht vollständig, weil sie drei Komponenten haben muss
- Wert für z ergänzen durch $z = 0 + 1\cdot z$
- Endergebnis:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 8 \\ 2 \\ 0 \end{pmatrix} + z \begin{pmatrix} 15 \\ 8 \\ 1 \end{pmatrix} .$$

- **Zusammenfassung:**
- Die Nicht-Pivot-Spalten werden mit der zugehörigen Variablen versehen auf die rechte Seite gebracht, wobei sich alle Vorzeichen umkehren.
- Zeilen für fehlende Variablen werden in den Lösungsvektor nach dem Schema $x = x$ eingefügt.

Beispiel 1.25: Man löse das lineare Gleichungssystem

$$\begin{array}{rrrrcr} 2a & +b & -2c & & = & -2 \\ -4a & -2b & +c & +2d & = & 2 \\ -2a & -b & & +d & = & 1 \end{array} .$$

Man erhält die reduzierte Stufenform

```
>> a=[2 1 -2 0 -2;-4 -2 1 2 2;-2 -1 0 1 1];
>> rref(a)

ans =

       1              1/2              0              0             -1
       0              0                1              0              0
       0              0                0              1             -1
```

Die zweite Spalte der Matrix ist keine Pivot-Spalte. Dies führt zu

$$\begin{pmatrix} a \\ c \\ d \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ -1 \end{pmatrix} + b \begin{pmatrix} -\frac{1}{2} \\ 0 \\ 0 \end{pmatrix}$$

und dann zum Endergebnis

$$\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \\ -1 \end{pmatrix} + b \begin{pmatrix} -\frac{1}{2} \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

Kapitel 2

Analytische Geometrie

2.1 Skalarprodukt und Norm

- Ziel: geometrische Konzepte wie Länge, Winkel auf den $\mathbb{R}^n$ übertragen
- für $n = 2$ und $n = 3$ müssen und werden sich die bekannten Längen- und Winkelbegriffe als Spezialfall ergeben
- entscheidend dazu: *Skalarprodukt*
- folgende Definition dient zur Vorbereitung

Definition 2.1: Seien $X_1, \ldots, X_n$ mit $n \in \mathbb{N}$ nichtleere Mengen. Dann versteht man unter dem *kartesischen Produkt* $X_1 \times \cdots \times X_n$ die Menge aller geordneten n-Tupel:

$$X_1 \times \cdots \times X_n := \{(x_1, \ldots, x_n)^T \,|\, x_i \in X_i\}\,.$$

Für $X_1 = \cdots = X_n = X$ schreibt man kurz X^n statt $X_1 \times \cdots \times X_n$.

Beispiel 2.2:

- $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$
- $\mathbb{R}^n = \mathbb{R} \times \ldots \times \mathbb{R}$ mit $X_1 = \ldots = X_n = \mathbb{R}$

Bemerkung 2.3:

1. - kartesisches Produkt verallgemeinert die Definition des Zahlraums
   - statt $\{1, \ldots, n\}$ als Indexmenge $\mathbb{N}$: unendliche kartesische Produkte
   - Elemente dieser Mengen: Folgen (siehe Analysis)

- ab jetzt: immer $a = (a_i)_{i=1}^n, b = (b_i)_{i=1}^n \in \mathbb{R}^n$.

Definition 2.4: Eine Abbildung $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ heißt *Skalarprodukt*, wenn folgende Bedingungen erfüllt sind:

SP1 (Symmetrie): $\forall a, b \in \mathbb{R}^n : \langle a, b \rangle = \langle b, a \rangle$.

SP2 $\forall a, b, c \in \mathbb{R}^n$:

$$\langle a, b + c \rangle = \langle a, b \rangle + \langle a, c \rangle$$

SP3 $\forall \alpha \in \mathbb{R}$ gilt

$$\langle \alpha a, b \rangle = \alpha \langle a, b \rangle = \langle a, \alpha b \rangle .$$

SP4 (positive Definitheit): $\forall a \in \mathbb{R}^n \setminus \{0\} : \langle a, a \rangle > 0$, und $\langle 0, 0 \rangle = 0$.

- Wichtigstes aller Skalarprodukte: *Standardskalarprodukt* oder *euklidisches Skalarprodukt*
- spricht man von „dem“ Skalarprodukt, ist das euklidische Skalarprodukt gemeint

Definition 2.5: Für a, b sei ihr *euklidisches Skalar-* bzw. *Punktprodukt* $\langle a, b \rangle$ definiert als

$$\langle a, b \rangle := a_1 b_1 + \cdots + a_n b_n = \sum_{i=1}^{n} a_i b_i \,.$$

Bemerkung 2.6:

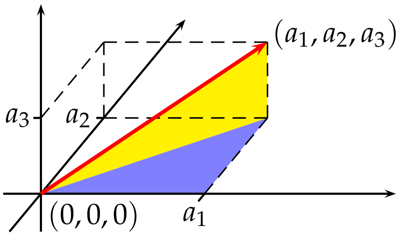
- alternative Schreibweise: $a \cdot b$ (daher die Bezeichnung „Punktprodukt“)
- (Skalar-)produkt von Vektoren entspricht in seinen Eigenschaften nicht dem Produkt zweier Zahlen
- höhere Potenzen von Vektoren sind nicht definiert

Beispiel 2.7: *OHP*

Satz 2.8: Das euklidische Skalarprodukt ist ein Skalarprodukt im Sinne der Definition 2.4.

Beweis *OHP*

- sei $a = (a_1, a_2)^T \in \mathbb{R}^2$
- Länge von a: $\sqrt{a_1^2 + a_2^2}$ nach Pythagoras
- sei weiter $a = (a_1, a_2, a_3)^T \in \mathbb{R}^3$
- doppelte Anwendung des Satzes von Pythagoras: Länge beträgt $\sqrt{a_1^2 + a_2^2 + a_3^2}$



- Verallgemeinerung auf den $\mathbb{R}^n$ liegt nahe

Definition 2.9: Einem Vektor a wird die *euklidische Norm* oder *Standardnorm* $\|a\|$ zugeordnet durch

$$\|a\| := \left(\sum_{i=1}^{n} a_i^2\right)^{1/2} . \tag{2.1}$$

Satz 2.10: Die in Gleichung (2.1) definierte Norm hat folgende Eigenschaften:

N0 : $\|a\| \in \mathbb{R}$.

N1 : $\|a\| \geq 0$.

N2 : $\|a\| = 0 \quad \Leftrightarrow \quad a = 0$.

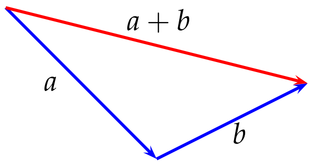
N3 : $\forall \lambda \in \mathbb{R} : \|\lambda a\| = |\lambda| \|a\|$

N4 : (Dreiecksungleichung) $\|a + b\| \leq \|a\| + \|b\|$

Beweis *OHP*

Bemerkung 2.11:

- Abbildung zeigt, warum **N4** *Dreiecksungleichung* genannt wird



- Mathematik strebt nach Abstraktion und Allgemeinheit
- bisher: Längenbegriff auf dem $\mathbb{R}^n$ durch Verallgemeinerung
- jetzt: weitere Verallgemeinerung

Definition 2.12: Eine Abbildung $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ heißt *Norm*, wenn sie die Eigenschaften **N0** bis **N4** aus Satz 2.10 besitzt.

- offensichtlich: die euklidische Norm entspricht der Wurzel des euklidischen Skalarprodukts eines Vektors mit sich selbst
- das euklidische Skalarprodukt *induziert* die euklidische Norm
- Zusammenhang gilt allgemeiner

Satz 2.13: Sei $\langle \cdot, \cdot \rangle$ ein beliebiges Skalarprodukt auf $\mathbb{R}^n$. Dann wird durch $\| \cdot \| := \sqrt{\langle \cdot, \cdot \rangle}$ eine Norm auf $\mathbb{R}^n$ induziert.

Beweis *OHP*

- es gibt viele Normen auf dem $\mathbb{R}^n$
- „die“ Norm: euklidische Norm
- „eine“ Norm: unbestimmte Norm
- sei stets $a = (a_i)_{i=1}^n \in \mathbb{R}^n$
- Auf $\mathbb{R}$ ist der Absolutbetrag eine Norm.
- Für $p \geq 1$ definiert man die ℓ_p-Norm durch

$$\|a\|_p := \left(\sum_{i=1}^{n} |a_i|^p \right)^{1/p} \tag{2.2}$$

  - Dreiecksungleichung: „diskrete Minkowskische Ungleichung“ (siehe Analysis)
  - für $0 < p < 1$ ist die Formel zwar definiert, die Dreiecksungleichung gilt i. A. aber nicht

- Der Spezialfall $p = 1$ wird als *Betragssummennorm* oder *Einernorm* bezeichnet und ist definiert als
$$\|a\|_1 = |a_1| + \cdots + |a_n|.$$
- Der Spezialfall $p = 2$ entspricht der euklidischen Norm.
- Die *Maximumnorm* oder ℓ_∞-Norm ist definiert als
$$\|a\|_\infty = \max\{|a_1|, \ldots, |a_n|\};$$
die Bezeichnung erklärt sich durch $\lim_{p\to\infty} \|a\|_p = \|a\|_\infty$.

Beispiel 2.14: *OHP*

Bemerkung 2.15:

- jedes Skalarprodukt induziert eine Norm (s. o.)
- Umkehrung falsch
- *keine* der ℓ_p-Normen ist von irgendeinem Skalarprodukt induziert bis auf den Spezialfall $p = 2$

2.1.1 Einheitsvektoren

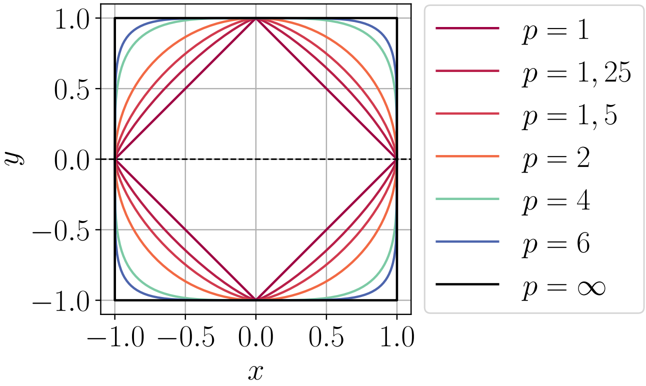
Definition 2.16: Sei $\|\cdot\|$ eine Norm. Ein Vektor $e \in \mathbb{R}^n$ heißt *Einheitsvektor* (zur Norm $\|\cdot\|$), wenn $\|e\| = 1$ ist.

Bemerkung 2.17:

- Eigenschaft „Einheitsvektor“ hängt von der Norm ab
- S^1 im $\mathbb{R}^2$:

$$S^1 := \left\{x \in \mathbb{R}^2 \mid \|x\| = 1\right\}.$$

- euklidische Norm: S^1 ist der Einheitskreis
- für andere ℓ_p-Normen:



- wichtige Einheitsvektoren: kanonische Einheitsvektoren in Richtung der Koordinatenachsen; $\|e_i\|_p = 1$

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix},$$

speziell in $\mathbb{R}^3$

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \tag{2.3}$$

die den Koordinatenrichtungen entsprechen und für die offensichtlich $\|e_i\|_p = 1$ für jede ℓ_p-Norm gilt.

- Normierung von $a \neq 0$
- finde Einheitsvektor e_a parallel und gleichgerichtet zu a durch

$$e_a = \frac{1}{\|a\|} a.$$

Beispiel 2.18: Sei $a = (2, -1, -2)^T$. Wegen $\|a\| = 3$ ist $e_a = {}^1\!/\!{}_3(2, -1, -2)^T$.

Bemerkung 2.19:

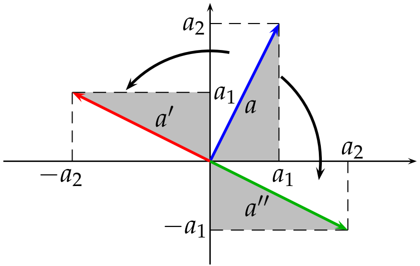
- v und λv $(\lambda > 0)$ zeigen in die gleiche Richtung
- nach Normierung gleich
- beim Normieren kann man zur Vereinfachung einen positiven Vorfaktor streichen

Beispiel 2.20: Möchte man den Vektor $b = \frac{11}{\sqrt{17}}(1, 0, 1)^T$ normieren, kann man statt dessen den Vektor $b' = (1, 0, 1)^T$ normieren und erhält bei einfacherer Rechnung das gleiche Ergebnis.

2.1.2 Orthogonale (senkrechte) Vektoren

Bemerkung 2.21:

- betrachte $a = (a_1, a_2)^T$



- a um 90° gegen den Uhrzeigersinn gedreht: $a' = (-a_2, a_1)^T$
- a um 90° im Uhrzeigersinn gedreht: $a'' = (a_2, -a_1)^T$

- euklidisches Skalarprodukt:

$$\langle a, a' \rangle = \langle a'', a \rangle = 0.$$

- umgekehrt: aus $\langle a, b \rangle = 0$ folgt $b = \lambda(-a_2, a_1)^T$
- $b \parallel a'$ und $b \parallel a''$, also $a \perp b$
- Verallgemeinerung auf den $\mathbb{R}^n$

Definition 2.22: Seien $a, b \in \mathbb{R}^n$ und $\langle \cdot, \cdot \rangle$ ein beliebiges Skalarprodukt. Die Vektoren a und b stehen *senkrecht* (oder auch *orthogonal*) zueinander bzgl. $\langle \cdot, \cdot \rangle$, Schreibweise $a \perp b$, wenn $\langle a, b \rangle = 0$ ist.

- Spezialfall euklidisches Skalarprodukt und $n = 2$: entspricht intuitiver Vorstellung von „senkrecht“

Bemerkung 2.23:

1. kanonische Einheitsvektoren stehen senkrecht aufeinander
2. Orthogonalität hängt vom gewählten Skalarprodukt ab:
   - $\langle a, b \rangle' := a_1 b_1 + 2 a_2 b_2$ ist Skalarprodukt
   - $(-1, 1)^T$ und $(1, 1)^T$ senkrecht bzgl. des euklidischen Skalarprodukts
   - aber: $\langle a, b \rangle' = 1$

Bemerkung 2.24:

- sei $a \perp b$
- dann: $\alpha a \perp \beta b$ wegen

$$\langle \alpha a, \beta b \rangle = \alpha \beta \langle a, b \rangle = 0.$$

Wir beweisen nun den Satz des Pythagoras.

Satz 2.25: Seien $a, b \in \mathbb{R}^n$ mit $a \perp b$, $\langle \cdot, \cdot \rangle$ ein beliebiges Skalarprodukt und $\| \cdot \|$ die dadurch induzierte Norm. Dann gilt

$$\|a + b\|^2 = \|a\|^2 + \|b\|^2.$$

Beweis *OHP*

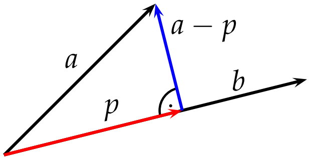
Bemerkung 2.26:

- Satz längst bekannt. Aber:

1. Satz gilt für *jedes* Skalarprodukt mit induzierter Norm
2. gilt im $\mathbb{R}^n$
3. Beweis elegant
4. rechtfertigt die vorgenommenen Abstraktionen

2.1.3 Winkel zwischen Vektoren

- sei Norm stets vom Skalarprodukt induziert
- a, b zwei Vektoren, $b \neq 0$
- Ziel: orthogonale Projektion p von a in Richtung von b



$$\begin{array}{lll} 1. \quad p \parallel b & \Rightarrow & p = \alpha b \\ 2. \quad (a-p) \perp b & \Rightarrow & \langle a-p, b\rangle = 0 \end{array} \tag{2.4}$$

OHP

Definition 2.27: Der für zwei Vektoren a und $b \neq 0$ durch (2.5) eindeutig bestimmte Vektor p heißt *orthogonale Projektion von a in Richtung b*, man schreibt auch $p_b(a)$.

OHP

Bemerkung 2.28: Ist b ein Einheitsvektor, so gilt:

$$p_b(a) = \langle a, b \rangle \cdot b.$$

Beispiel 2.29: *OHP*

Bemerkung 2.30:

- Konzept der orthogonalen Projektion wirkt unscheinbar
- eines der zentralsten Konzepte der Mathematik
- Finite Elemente basieren darauf
- allerdings: nochmals verallgemeinert
- ähnlich bedeutend: Cauchy-Schwarzsche Ungleichung

Satz 2.31 (Cauchy-Schwarzsche Ungleichung): Für $a, b \in \mathbb{R}^n$ gilt

$$|\langle a, b\rangle| \leq \|a\| \, \|b\|. \tag{2.8}$$

Beweis *OHP*

- jetzt: Beweis von **N4** in Satz 2.13

Satz 2.32 (Dreiecksungleichung): Für $a, b \in \mathbb{R}^n$ gilt

$$\|a + b\| \leq \|a\| + \|b\|.$$

Beweis *OHP*

Bemerkung 2.33:

- bei beliebiger Norm (evtl. nicht von Skalarprodukt induziert): **N4** folgt nicht aus obigem Satz
- **N4** muss von Fall zu Fall bewiesen werden

- jetzt: Definition von Winkeln zwischen 2 Vektoren
- $a, b \neq 0$ Vektoren in der Ebene; euklidsche Norm und euklidisches Skalarprodukt
- $0 \leq \theta \leq {}^{\pi}/_{2}$
- nutze $p_a(b)$

OHP

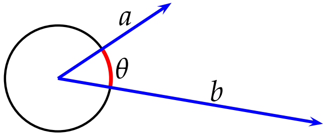
Definition 2.34: Seien $a, b \in \mathbb{R}^n \setminus \{0\}$. Der *Winkel* zwischen a und b, geschrieben $\angle(a, b)$, wird definiert als

$$\angle(a, b) := \arccos \frac{\langle a, b \rangle}{\|a\| \, \|b\|}. \tag{2.10}$$

- CSU: Winkel wohldefiniert
- $\angle(a, b) \in [0, \pi]$ wg. Wertebereich des Arcuscosinus
- bisheriger Winkelbegriff in der Ebene ist Spezialfall

Bemerkung 2.35:

- es gilt $0 \leq \theta \leq \pi$
- Winkel immer der kleinere der beiden möglichen



Beispiel 2.36: *OHP*

Beispiel 2.37: *OHP*

2.1.4 Das Vektorprodukt (Kreuzprodukt)

- Beschränkung auf $\mathbb{R}^3$ und das euklidische Skalarprodukt bzw. euklidische Norm
- gegeben $a = (a_1, a_2, a_3)^T$ und $b = (b_1, b_2, b_3)^T$ nicht parallel
- Ziel: Konstruktion eines auf a und b senkrechten Vektors
- Intuition: sollte es geben
- für einen strengen Beweis: Vektorprodukt als algebraische Operation
- dann: Nachweis der gewünschten geometrischen Eigenschaften

Definition 2.38: Seien a, b wie oben. Dann heißt

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

das *Vektorprodukt* oder *Kreuzprodukt* von a und b.

Beispiel 2.39: *OHP*

- 2 Schemata zur einfachen Berechnung des Kreuzprodukts
- Erstes Schema:
- Multiplikation über Kreuz
- dann: subtrahieren

$$\begin{pmatrix} 2 \\ -1 \\ -2 \end{pmatrix}\begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} \qquad \begin{pmatrix} 2 \\ -1 \\ -2 \end{pmatrix}\begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} \qquad \begin{pmatrix} 2 \\ -1 \\ -2 \end{pmatrix}\begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix}$$

1. Komponente: $1 - (-4) = 5$

2. Komponente: $-6 - (-2) = -4$

3. Komponente: $4 - (-3) = 7$

- Entlang der Pfeile wird überkreuz multipliziert
- dann wird entsprechend der Vorzeichen addiert bzw. subtrahiert

- Zweite Möglichkeit:

$$\begin{array}{rcrcrcr} & & 2 & & 3 & & \\ & & -1 & & 2 & & \\ {\color{red}-4} & & -2 & & -1 & & {\color{blue}1} \\ {\color{red}-2} & \longleftarrow & 2 & & 3 & \longrightarrow & {\color{blue}-6} \\ {\color{red}-3} & & -1 & & 2 & & {\color{blue}4} \end{array} \Rightarrow a \times b = \begin{pmatrix} {\color{blue}1} - {\color{red}(-4)} \\ {\color{blue}-6} - {\color{red}(-2)} \\ {\color{blue}4} - {\color{red}(-3)} \end{pmatrix} = \begin{pmatrix} 5 \\ -4 \\ 7 \end{pmatrix}$$

- Unter beide Vektoren werden die ersten beiden Komponenten des jeweiligen Vektors geschrieben

- Dann werden die Komponenten entlang der schrägen Pfeile multipliziert und die Produkte am rechten und linken Rand notiert.

- Schließlich werden die Produkte am linken Rand von denen am rechten Rand abgezogen

Es gelten folgende Rechenregeln:

Satz 2.40: Für a, b wie oben gilt

1. $a \times b = -b \times a$.
2. $a \times a = 0 \in \mathbb{R}^3$.
3. Sei $c \in \mathbb{R}^3$. Dann gilt $(a + b) \times c = a \times c + b \times c$.
4. Für $\alpha \in \mathbb{R}$ gilt $\alpha a \times b = \alpha(a \times b) = a \times (\alpha b)$.
5. $\langle a \times b, a\rangle = \langle a \times b, b\rangle = 0$.

Beweis *OHP*

- Aussage 5.) löst unser Problem nicht vollständig
- Was, wenn $a \times b$ der Nullvektor ist?
- Dann ist $a \times b$ nicht der gesuchte Vektor senkrecht zu a und b
- von jetzt an: $a, b \in \mathbb{R}^3 \setminus \{0\}$

Satz 2.41: Die Vektoren a, b sind genau dann nicht parallel, wenn $a \times b \neq 0$ gilt.

Beweis *OHP*

- jetzt: Länge von $a \times b$

Satz 2.42: Für a, b gilt

$$\|a \times b\|^2 = \|a\|^2 \|b\|^2 - \langle a, b \rangle^2 \, .$$

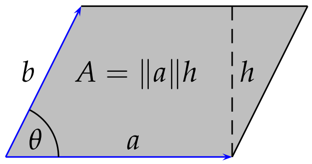
Beweis *OHP*

Folgerung 2.43: Sei θ der Winkel zwischen a und b. Dann gilt

$$\|a \times b\| = \|a\| \, \|b\| \, \sin\theta \, .$$

Beweis *OHP*

- geometrisch: Länge des Vektorprodukts als Fläche eines Parallelogramms



Bemerkung 2.44: *OHP*

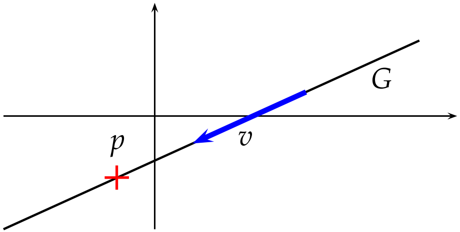
Bemerkung 2.45:

- alle geometrischen Eigenschaften von $a \times b$ gelten auch für $-(a \times b)$
- Kreuzprodukt ist so definiert, dass man $a, b, a \times b$ auf e_1, e_2, e_3 drehen kann
- Lage der Vektoren zueinander entsprechen der kanonischen Einheitsvektoren
- „Rechtssystem“
- rechte Hand: Daumen $\simeq e_1$, Zeigefinger $\simeq e_2$, Mittelfinger $\simeq e_3$
- falsch für $a, b, -a \times b$
- Bei jeder Drehung zeigt mindestens ein Vektor in Gegenrichtung eines kanonischen Einheitsvektors
- „Linkssystem“

2.2 Geraden und Ebenen

2.2.1 Definition und grundlegende Eigenschaften

- zur Definition einer Geraden in der Ebene: Anschauung
- sei immer $\alpha, \beta \in \mathbb{R}$ und $n, p, v \in \mathbb{R}^2$.



Definition 2.46:

1. Für einen *Ortsvektor* oder *Aufpunkt* p und einen *Richtungsvektor* $v \neq 0$ heiße

$$G := \{x \in \mathbb{R}^2 \,|\, \exists \alpha \in \mathbb{R} : x = p + \alpha v\} \tag{2.11}$$

   eine *Gerade*.

2. Die Gleichung

$$x = p + \alpha v$$

   aus (2.11) heißt *Punkt-Richtungsgleichung* von G, die reelle Zahl α aus Formel (2.11) nennt man *Parameter*.

Bemerkung 2.47:

- im Folgenden: Wir identifizieren eine Gerade mit ihrer Geradengleichung
- Aufpunkt und Richtungsvektor legen eine Gerade eindeutig fest

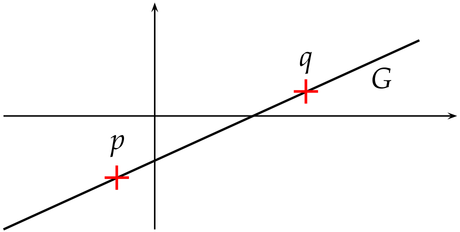
- aber: bei gegebener Geraden liegen Aufpunkt und Richtungsvektor nicht eindeutig fest

Satz 2.48: Sei G eine Gerade mit dem Richtungsvektor v und dem Aufpunkt p.

1. Jeder Vektor der Form $\tilde{v} = \beta v$ mit $\beta \neq 0$ ist ebenfalls Richtungsvektor.
2. Jeder Vektor der Form $\tilde{p} = p + \gamma v$ mit irgendeinem $\gamma \in \mathbb{R}$ ist ebenfalls Aufpunkt.

Beweis *OHP*

- jetzt: seien anstelle von Aufpunkt und Richtungsvektor zwei Punkte auf der Geraden bekannt



Satz 2.49: Seien $p, q \in \mathbb{R}^2$ mit $p \neq q$. Dann gibt es genau eine Gerade durch p und q.

Beweis *OHP*

Definition 2.50:

1. Seien p, q zwei verschiedene Punkte auf der Geraden G. Dann heißt die Geradengleichung

$$x = p + \alpha(q - p)$$

Zweipunktform von G. Man sagt, „die Gerade liegt in der Zwei-Punkte-Form vor“.

2. Alle Gleichungen, die einen reellen Parameter enthalten und eine Gerade beschreiben, nennt man *Parametergleichungen* oder *Parameterformen* einer Geraden.

Bemerkung 2.51:

1.
   - möglich: statt eines vorhandenen Richtungsvektors ein Vielfaches wählen
   - in der Praxis: Richtungsvektor mit „möglichst einfachen Zahlen“
   - Beispiel: $(\pi/3, \pi/6)^T$ eher ungeschickt, $(2,1)^T$ besser
   - Analog: Auswahl des Aufpunkts

2. • Umrechnung der Zwei-Punkte-Form in die Punkt-Richtungsform ist einfach:
   • Richtungsvektor z. B. durch $p-q$ oder $q-p$; Aufpunkt: z. B. p oder q
   • Umgekehrt: gegeben Richtungsvektor v. Dann sind p und z. B. $p+v$ oder $p-v$ zwei verschiedene Punkte auf der Geraden

3. • Dass man durch zwei Punkte genau eine Gerade legen kann, erscheint auch ohne Theorie einleuchtend
   • aber: strenger Beweis statt Appell an Einsicht
   • Indiz dafür, dass die algebraische Definition der geometrischen Objekts „Gerade“ sinnvoll ist

- jetzt: alternative Beschreibungen von Geraden ohne Parameter
- vorbereitend dazu folgendes Lemma

Lemma 2.52: Seien $a, b, c \in \mathbb{R}^2 \setminus \{0\}$ beliebig, $a \perp b$ und $b \perp c$. Dann gilt $a \| c$.

Beweis *OHP*

Satz 2.53: Sei $G \subset \mathbb{R}^2$ eine Gerade mit Richtungsvektor v und Aufpunkt p. Für jedes $n \neq 0$ mit $n \perp v$ gilt

$$x \in G \Leftrightarrow \langle x, n \rangle = \langle p, n \rangle \, .$$

Der Vektor n ist bis auf Skalierung eindeutig bestimmt.

Beweis *OHP*

Definition 2.54: Sei G eine Gerade in der Ebene und p der Aufpunkt von G.

1. Ein Vektor n wie in Satz 2.53 heißt *Normalenvektor* von G.
2. Die Gleichung
$$\langle x, n\rangle = \langle p, n\rangle$$
heißt *Normalform* von G.

- schon gesehen: jeder Punkt der Geraden eignet sich als Aufpunkt
- noch zu zeigen: die Normalform einer Geraden hängt von der Wahl des Aufpunkts nicht ab

Bemerkung 2.55: Seien p, q zwei Punkte auf der Geraden G mit Normalenvektor n. Dann gilt
$$\langle p, n\rangle = \langle q, n\rangle\ .$$

Beweis *OHP*

Bemerkung 2.56:

1.
- G, n und p wie zuvor
- $n = (a, b)^T \in \mathbb{R}^2$, $x = (x_1, x_2)^T$ und $c := \langle p, n \rangle \in \mathbb{R}$
- durch Ausrechnen folgt die allgemeine Geradengleichung

$$ax_1 + bx_2 = c\,.$$

2.
- sei $b \neq 0$
- setze $m = -a/b$ und $n = c/b$
- Geradengleichung nimmt folgende Form an:

$$y = mx + n,$$

- hier: $x := x_1$ und $y := x_2$

- Spezialfall der Normalform einer Geraden

Beispiel 2.57: *OHP*

- Normalform ist nur bis auf einen Faktor genau bestimmt
- Nomierung führt zur Hesseschen Normalform

Definition 2.58: Sei G eine Gerade mit Normalenvektor n. Gilt $\|n\| = 1$, so heißt die damit gebildete Normalform *Hessesche Normalform*.

Bemerkung 2.59: Man erhält die Hessesche Normalform aus einer beliebigen Normalform, indem man die Normalform durch $\|n\|$ teilt:

$$\frac{\langle x, n\rangle}{\|n\|} = \frac{\langle p, n\rangle}{\|n\|}.$$

Damit liegt der Normalvektor bis auf das Vorzeichen eindeutig fest.

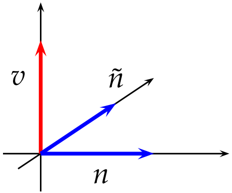
Beispiel 2.60: *OHP*

- jetzt: Geraden im Raum
- Definition analog zu Geraden in der Ebene
- die Sätze 2.48 und 2.49 gelten wortgleich mit wortgleichen Beweisen
- Zwei-Punkte-Form und Punkt-Richtungsform analog
- aber: es existiert keine parameterlose Beschreibung einer Geraden im Raum
- Grund: die Richtung eines Normalenvektors liegt nicht eindeutig fest

Beispiel 2.61:

- $v = (0, 0, 1)^T$ sei Richtungsvektor einer Geraden
- $n = (1, 0, 0)$ ist Normalenvektor, da $\langle v, n \rangle = 0$
- $\tilde{n} = (0, 1, 0)^T$ ist ebenso Normalenvekor

- Rechenregeln des Skalarprodukts: $\alpha n + \beta \tilde{n}$ ist Normalenvektor
- Konsequenz: jeder Vektor in der (x_1, x_2)-Ebene ist Normalenvektor



- jetzt: Ebenen im Raum
- $\mathbb{R}^2$ lässt sich als Ebene deuten
- wir identifizieren die Ebene mit

$$\tilde{E} := \left\{ x \in \mathbb{R}^3 \,\middle|\, x = (x_1, x_2, 0)^T;\ x_1, x_2 \in \mathbb{R} \right\}$$

- offenbar gilt

$$\tilde{E} := \left\{ x \in \mathbb{R}^3 \,\middle|\, x = (0,0,0)^T + x_1(1,0,0)^T + x_2(0,1,0)^T;\, x_1, x_2 \in \mathbb{R} \right\};$$

- Ebene: allen Vektoren der Form $0 + \alpha(1,0,0)^T + \beta(0,1,0)^T$ mit reellen Parametern α und β
- die Vektoren zeigen in unterschiedliche Richtungen
- die folgende Definition verallgemeinert diese Beobachtungen

Definition 2.62: Seien $p, v, w \in \mathbb{R}^3$, $v \neq 0$ und $w \neq 0$, und seien v und w nicht parallel. Dann heißt

$$E := \left\{ x \in \mathbb{R}^3 \,\middle|\, x = p + \alpha v + \beta w;\, \alpha, \beta \in \mathbb{R} \right\}$$

Ebene; die Vektoren v und w heißen *Richtungsvektoren*. Die Gleichung $x = p + \alpha v + \beta w$ wird *Punkt-Richtungsgleichung genannt*; man sagt, die „Ebene liegt in Punkt-Richtungsform vor“.

Bemerkung 2.63:

1. - ebenso wie bei Geraden in der Ebene: jeder Punkt eignet sich als Aufpunkt
   - beliebige zwei nichtparallele Vektoren ungleich Null in der Ebene sind Richungs- vektoren
2. Drei paarweise verschiedene Punkte, die nicht auf einer Geraden liegen, legen eine Ebene eindeutig fest.

Beweis

- Beweise ähnlich zu den Beweisen bei Geraden in der Ebene
- technisch komplizierter
- werden ausgelassen ■

Satz 2.64: Sei E eine Ebene mit Aufpunkt p und den Richtungsvektoren v und w. Dann existiert ein Vektor $n \neq 0$ mit $v \perp n$, $w \perp n$ und

$$x \in E \Leftrightarrow \langle x, n\rangle = \langle p, n\rangle \, .$$

Der Vektor n ist bis auf Skalierung eindeutig bestimmt.

Definition 2.65: Der Vektor n aus Satz 2.64 heißt *Normalenvektor* der Ebene, die Gleichung

$$\langle x, n\rangle = \langle p, n\rangle \tag{2.12}$$

heißt *Normalform* oder *Normalgleichung* der Ebene.

Bemerkung 2.66:

1. Die Normalform hängt nicht von der Wahl des Aufpunkts p ab.
2. • $n = (a, b, c)^T$ und $d = \langle p, n\rangle \in \mathbb{R}$

- dann: allgemeine Ebenengleichung

$$ax_1 + bx_2 + cx_3 = d\,.$$

3. - Hessesche Normalform wird analog definiert
   - bis auf Vorzeichen eindeutig

- Zusammenfassung

Raum	Objekt	Punkt-Richtungs-Gl.	Normalgleichung
$\mathbb{R}^2$	Gerade	$x = p + \alpha v$	$\langle x, n\rangle = \langle p, n\rangle$ $ax_1 + bx_2 = c$
$\mathbb{R}^3$	Gerade	$x = p + \alpha v$	Existiert nicht!
$\mathbb{R}^3$	Ebene	$x = p + \alpha v + \beta w$	$\langle x, n\rangle = \langle p, n\rangle$ $ax_1 + bx_2 + cx_3 = d$

2.2.2 Umrechnen zwischen verschiedenen Darstellungsformen

- Umrechnung zwischen Parameterform einer Geraden und Normalform
- es gibt keine Normalform von Geraden in $\mathbb{R}^3$
- deswegen: Geraden in $\mathbb{R}^2$
- Umrechnung zwischen Punkt-Richtungsform und Zwei-Punkte-Form einfach
- deswegen: nur Punkt-Richtungsform
- G gegeben durch $x = p + \alpha v$ mit $v = (v_1, v_2)^T$
- Normalenvektor n durch

$$n := \begin{pmatrix} v_2 \\ -v_1 \end{pmatrix}$$

- bis auf Skalierung einzige Möglichkeit
- rechte Seite durch Ausrechnen von $\langle p, n \rangle$

Beispiel 2.67: *OHP*

- geg. parameterlose Form $ax_1 + bx_2 = c$
- finde Richtungsvektor senkrecht zu $n = (a, b)^T$
- wähle $v = (b, -a)^T$
- p: wähle $x_1 = 0$ oder $x_2 = 0$ und berechne die andere Komponente aus der parameterlosen Form

Beispiel 2.68: *OHP*

- jetzt: Umrechnung zwischen Ebenendarstellungen
- man benötigt einen auf beiden Richtungsvektoren senkrechten Vektor
- in $\mathbb{R}^3$: Kreuzprodukt möglich
- Aufpunkt: jeder Punkt der Ebene, insbesondere p aus der Punkt-Richtungsform

Beispiel 2.69: *OHP*

- Umrechnung einer Normalform einer Ebene in die Punkt-Richtungsform
- gesucht: zwei nicht parallele Richtungsvektoren, beide senkrecht auf n
- Normalgleichung $ax_1 + bx_2 + cx_3 = d$
- lese Normalenvektor $n = (a, b, c)^T$ ab
- mindestens ein n_i von n ist ungleich 0
- vertausche n_i mit einer anderen Komponente n_j und verändere das Vorzeichen von n_i
- restliche Komponenten: 0
- es gibt zwei Möglichkeiten, $j \neq i$ zu wählen $\Rightarrow$ zwei Vektoren v und w
- $v \perp n$ und $w \perp n$, weder v noch w sind der Nullvektor
- fehlt: Aufpunkt

- zur Berechnung weist man zwei der drei Koordinaten von $x = (x_1, x_2, x_3)^T$ z. B. den Wert 0 zu
- restliche Koordinate durch Einsetzen in die Normalform
- wenn nicht möglich, wähle man ein anderes Koordinatenpaar
- es gibt immer zwei Koordinaten in x, mit denen das Verfahren funktioniert

Beispiel 2.70: *OHP*

2.2.3 Lagebeziehungen von Geraden und Ebenen

- zunächst: Orthogonalität und Winkel zwischen Geraden und Ebenen

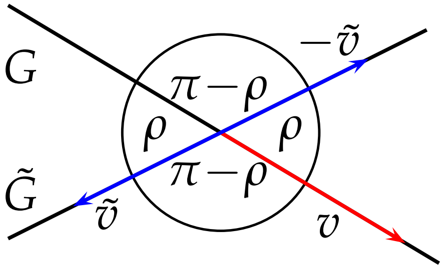
Definition 2.71:

1. Zwei Geraden in $\mathbb{R}^2$ oder $\mathbb{R}^3$ heißen *parallel*, wenn ihre Richtungsvektoren parallel sind.

2. Zwei sich schneidende Geraden heißen orthogonal, wenn ihre Richtungsvektoren orthogonal sind.

3. Seien G und $\tilde{G}$ zwei sich schneidende Geraden mit Richtungsvektoren v bzw. $\tilde{v}$. Der *Winkel* $\angle(G,\tilde{G})$ zwischen den Geraden wird definiert durch

$$\angle(G,\tilde{G}) := \min\{\angle(v,\tilde{v}), \angle(v,-\tilde{v})\}$$



Bemerkung 2.72:

1. - nicht möglich: Definition des Winkels zwischen zwei Geraden als Winkel zwischen den beiden Richtungsvektoren
   - Grund: siehe Abbildung
   - Skalierung des Richtungsvektors mit einer positiven Zahl unproblematisch
2. Abbildung zeigt: Winkel zwischen Geraden höchstens 90°

3. Für den Winkel zwischen zwei Vektoren $v, \tilde{v}$ gilt

$$\angle(v, \tilde{v}) = \arccos\left(\frac{\langle v, \tilde{v}\rangle}{\|v\|\,\|\tilde{v}\|}\right) \tag{2.13}$$

- Winkel größer als 90°: Skalarprodukt ist negativ
- Übergang von $\tilde{v}$ auf $-\tilde{v}$: Vorzeichenumkehr im obigen Bruch
- also gilt für den Winkel zwischen zwei sich schneidenden Geraden:

$$\angle(G, \tilde{G}) = \arccos\left(\frac{|\langle v, \tilde{v}\rangle|}{\|v\|\,\|\tilde{v}\|}\right) \tag{2.14}$$

Beispiel 2.73: *OHP*

Definition 2.74:

1. Zwei Ebenen heißen *parallel*, wenn ihre Normalenvektoren parallel sind. Sie heißen orthogonal, wenn ihre Normalenvektoren orthogonal sind.

2. Sei E eine Ebene mit Normalenvektor n. Eine Gerade mit Richtungsvektor v heißt *parallel* zur Ebene E, falls $v \perp n$.

3. Seien n und $\tilde{n}$ Normalenvektoren der beiden Ebenen E und $\tilde{E}$. Dann wird der Winkel $\angle(E,\tilde{E})$ zwischen den beiden Ebenen erklärt durch

$$\angle(E,\tilde{E}) := \min\{\angle(n,\tilde{n}), \angle(n,-\tilde{n})\}$$

Mit einer zum Fall der Geraden analogen Argumentation erkennt man

$$\angle(E,\tilde{E}) = \arccos\left(\frac{|\langle n,\tilde{n}\rangle|}{\|n\|\,\|\tilde{n}\|}\right). \tag{2.15}$$

Beispiel 2.75: *OHP*

2.2.4 Schnittmengen zwischen Geraden und Ebenen in $\mathbb{R}^2$ und $\mathbb{R}^3$

- Inhalt: Weg zur Bestimmung der Schnittmenge von Geraden und Ebenen
- gemeinsame Behandlung von Geraden in der Ebene und Ebenen im Raum
- „Hyperebene“ als gemeinsamer Begriff: „alles, was eine Normalform besitzt“
- im Folgenden mindestens eine Hyperebene beteiligt
- Gleichungen in Parameterform oder parameterlos

Rechenweg 2.76 (Bestimmung der Schnittmenge):

1. Man sorgt eventuell durch Umrechnung dafür, dass das eine Objekt durch eine parameterlose und das andere durch eine parameterbehaftete Gleichung beschrieben wird.

2. Man setzt die Parametergleichung in die parameterlose Gleichung ein und erhält Ausdrücke für den oder die Parameter.
3. Diese setzt man in die Parametergleichung ein und erhält eine Parametrisierung der Schnittmenge.

- Rechenweg nicht geeignet zur Bestimmung des Schnitts zweier Geraden in $\mathbb{R}^3$
- Grund: keine parameterfreie Darstellung verfügbar
- später

Beispiel 2.77: *OHP*

Satz 2.78: Sei E eine Hyperebene und G eine nichtparallele Gerade. Dann existiert ein eindeutiger Schnittpunkt von E und G, der sich mit obigem Verfahren berechnen lässt.

Beweis *OHP*

Analog, aber mit etwas mehr Aufwand zeigt man die Existenz einer Schnittgeraden von zwei nicht parallelen Ebenen.

- gegeben Geradengleichungen

$$G_1: \quad x = p_1 + \alpha v_1; \qquad G_2: \quad y = p_2 + \beta v_2$$

- Ansatz: Gleichsetzen und Auflösen

$$p_1 + \alpha v_1 = p_2 + \beta v_2$$

- 3 Fälle möglich:
  1. LGS hat genau eine Lösung: Die Geraden haben genau einen Schnittpunkt
  2. Geraden parallel: entweder mehrere Lösungen oder keine Lösung des LGS

3. Geraden nicht parallel, kein Schnittpunkt: Geraden sind *windschief*. Das LGS hat keine Lösung

Beispiel 2.79: *OHP*

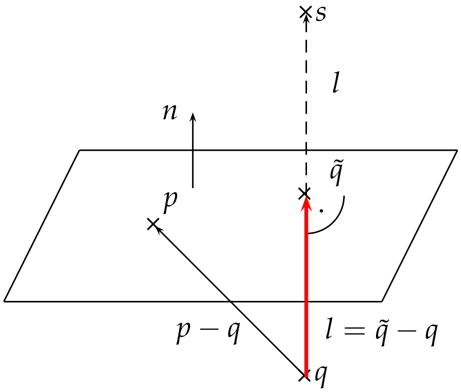
2.2.5 Abstandsbestimmung in $\mathbb{R}^2$ und $\mathbb{R}^3$

- seien E Hyperebene in $\mathbb{R}^2$ oder $\mathbb{R}^3$ mit Normalenvektor n, $q \notin E$ ein Punkt
- Ziel: Abstandsbegriff

Definition 2.80:

1. Seien E und q wie oben. Der Schnittpunkt $\tilde{q}$ der Geraden $q + \alpha n$ mit E heißt *Lotfußpunkt*, $l := \tilde{q} - q$ heißt *Lot*, $s := q + 2l$ heißt *Spiegelpunkt* von q an E.
2. Der Abstand d von q zu E wird definiert als die Länge des Lotes, also

$$d := \|l\|\,.$$



- Definition gerechtfertigt: Nach Satz 2.78 ex. eindeutiger Lotfußpunkt
- Warum gerade die Länge des Lotes als Abstandsbegriff?

Satz 2.81: Seien E und $q \notin E$ wie oben. Dann gilt

$$\|q - \tilde{q}\| = \min_{p \in E} \|q - p\|\,,$$

und der Lotfußpunkt ist der einzige Punkt in E mit minimalem Abstand zu q.

Beweis

- $\tilde{q} - p \perp n$
- Pythagoras:

$$\begin{aligned} \|p - q\|^2 &= \|(p - \tilde{q}) + (\tilde{q} - q)\|^2 \\ &= \|p - \tilde{q}\|^2 + \|\tilde{q} - q\|^2\,. \end{aligned}$$

- $\|p - q\|$ minimal, wenn $\|p - \tilde{q}\| = 0$

- also $p = \tilde{q}$ ■

- jetzt: Abstand Punkt-Gerade in $\mathbb{R}^3$
- Vorgehen analog

Definition 2.82: Sei G eine Gerade mit Richtungsvektor v und $q \notin G$. Ein Punkt $\tilde{q} \in G$ heißt *Lotfußpunkt*, wenn $l := q - \tilde{q} \perp G$ gilt, l heißt *Lot*, und der Abstand d zwischen einem Punkt und einer Geraden in $\mathbb{R}^3$ wird definiert durch $d := \|l\|$.

- es ex. immer ein eindeutiger Lotfußpunkt
- wie bei Hyperebene: zugleich der zu q nächste Punkt auf G
- jetzt: Abstand zweier nicht-paralleler Geraden in $\mathbb{R}^3$

Satz 2.83: Seien G_1 und G_2 zwei nicht parallele Geraden in $\mathbb{R}^3$. Dann steht der kürzeste Verbindungsvektor zwischen G_1 und G_2 senkrecht sowohl auf G_1 als auch auf G_2.

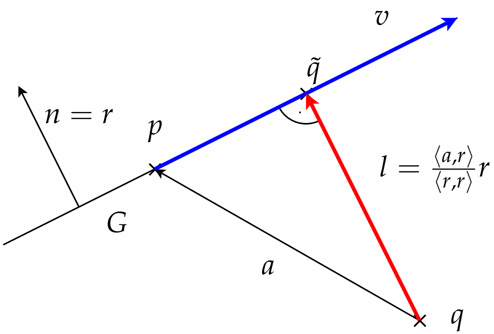
- Abstandsberechnungen mittels Berechnung des Lotfußpunktes
- dazu dient folgender Rechenweg

Rechenweg 2.84 (Abstandsberechnung):

1. Bestimme die Richtung r des Lots.
2. Bestimme jeweils einen Punkt auf den beiden Objekten und (durch Differenzbildung) den Abstandsvektor a zwischen den beiden Punkten.
3. Das Lot l ist die Projektion von a auf r. Der gesuchte Abstand ist $d = \|l\|$.

$$l = \frac{\langle a, r\rangle}{\langle r, r\rangle} r; \quad d = \frac{|\langle a, r\rangle|}{\langle r, r\rangle}\|r\| = \frac{|\langle a, r\rangle|}{\|r\|}$$

- Erläuterung am Abstand Punkt - Gerade:



- einziger Unterschied zwischen den Fällen: Berechnung von r

Beispiel 2.85: *OHP*

Beispiel 2.86: *OHP*

- Spezialfall Abstand einer Ebene vom Nullpunkt:

$$d = \frac{|\langle p, n\rangle|}{\|n\|},$$

- Länge der Projektion des Ortsvektors p zum Punkt P auf der Ebene mit dem Normalenvektor n.

- Hessesche Normalform

$$\frac{\langle x, n\rangle}{\|n\|} = \frac{\langle p, n\rangle}{\|n\|}$$

- Betrag der rechten Seite ist der Abstand der Ebene zum Nullpunkt

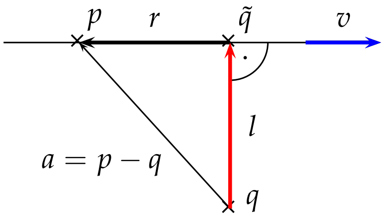
- bisheriger Ansatz schlägt fehl, weil eine Gerade im $\mathbb{R}^3$ keine Hyperebene ist

- Ansatz hierfür:

Rechenweg 2.87 (Abstand Punkt-Gerade in $\mathbb{R}^3$):

1. $a = l + r$, also $l = a - r$.
2. r ist die Projektion von a auf den Richtungsvektor der Geraden v, also
$$r = \frac{\langle a, v\rangle}{\langle v, v\rangle} v.$$
3. Zusammen ergibt sich:
$$l = a - \frac{\langle a, v\rangle}{\langle v, v\rangle} v$$

Die nachfolgende Grafik erläutert obigen Rechenweg.



Beispiel 2.88: *OHP*

2.3 Die Determinante in $\mathbb{R}^2$ und $\mathbb{R}^3$

2.3.1 Berechnung und geometrische Deutung

- Determinante für (1×1), (2×2)- und (3×3)-Matrizen
- nützliches Rechenwerkzeug
- Definition durch Angabe von Rechenvorschriften

Definition 2.89: Sei $A \in \mathbb{R}^{n \times n}$.

n=1: $A = (a_1)$. Dann gilt

$$\det(A) := a_1 \,.$$

n=2: Sei $A = (a, b)$ mit Spaltenvektoren $a, b \in \mathbb{R}^2$. Dann gilt

$$\det A = \det(a,b) = \det\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} := a_1b_2 - b_1a_2$$

$$- \qquad\qquad +$$

Entlang der Diagonalen wird das Produkt gebildet und entsprechend dem angezeigten Vorzeichen aufsummiert.

n=3: Für $A = (a,b,c)$ mit Spaltenvektoren $a,b,c \in \mathbb{R}^3$ setzen wir

$$\det(a,b,c) = \det\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$
$$:= a_1b_2c_3 + b_1c_2a_3 + c_1a_2b_3$$
$$-a_1c_2b_3 - b_1a_2c_3 - c_1b_2a_3\,.$$

- Veranschaulichung der Formel für $n = 3$ nach Sarrus:

$$\left|\begin{array}{ccc} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{array}\right|\begin{array}{cc} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{array}$$

$$- \quad - \quad - \qquad\qquad + \quad + \quad +$$

- Entlang der Diagonalen werden Produkte gebildet und diese mit den abgebildeten Vorzeichen versehen aufsummiert

- jetzt: geometrische Deutung der Determinanten für $n = 2$

$$\left|\det\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}\right| = |a_1 b_2 - a_2 b_1| = \left\|\begin{pmatrix} 0 \\ 0 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}\right\| = \left\|\begin{pmatrix} a_1 \\ a_2 \\ 0 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix}\right\|.$$

- euklidische Norm des Kreuzprodukts: Flächeninhalte des aufgespannten Parallelogramms

- also: Betrag der Determinante entspricht der Fläche des aufgespannten Parallelogramms

- geometrische Deutung für $n = 3$ benötigt folgende Definition

Definition 2.90: Für drei Vektoren $a, b, c \in \mathbb{R}^3$ nennt man

$$\langle a, b \times c \rangle \in \mathbb{R}$$

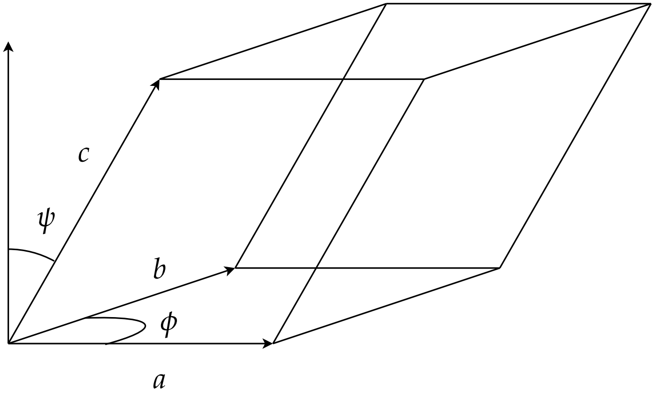
das *Spatprodukt* der drei Vektoren a, b, c.

Satz 2.91: Seien $a, b, c \in \mathbb{R}^3$, ϕ der Winkel zwischen a und b sowie ψ der Winkel zwischen den Vektoren $a \times b$ und c. Dann gilt:

$$\det(a, b, c) = \|a\| \cdot \|b\| \cdot \|c\| \sin\phi \cos\psi \tag{2.18}$$

Beweis *OHP*

- a, b, c : Kanten eines *Parallelepipeds* oder *Spats*
- Betrag der Determinante entspricht dem Volumen dieses Spats



Bemerkung 2.92:

- Vorzeichen der Determinante: Orientiertheit der Spaltenvektoren
- Vorzeichen positiv: Rechtssystem, sonst Linkssystem
- *orientiertes Volumen* des Spats

Satz 2.93: Seien $a, b, c \in \mathbb{R}^3$ Spaltenvektoren. Die Determinante hat folgende Eigenschaften.

D1 : $\det(a, b, c) = \det(c, a, b) = \det(b, c, a)$.

D2 : $\det(a, b, c) = -\det(b, a, c)$.

D3 : $\det(a, a, c) = 0$.

D4 : Für $\alpha \in \mathbb{R}$ gilt $\det(\alpha \cdot a, b, c) = \alpha \cdot \det(a, b, c)$.

D5 : $\det(a, b, c + d) = \det(a, b, c) + \det(a, b, d)$.

D6 : $\det(A) = \det(A^T)$

Für (2×2)-Matrizen $A = (a, b)$ mit $a, b \in \mathbb{R}^2$ gelten **D1** - **D6** sinngemäß.

Beweis Für $\mathbb{R}^2$: direktes Nachrechnen; für $\mathbb{R}^3$ verwendet man Satz 2.91 zusammen mit den Eigenschaften von Kreuz- und Skalarprodukt. ■

Erläuterungen, Beispiele und Folgerungen: *OHP*

Bemerkung 2.94: *OHP*

2.3.2 Lineare 3×3-Gleichungssysteme

- Anwendung von Determinanten: Untersuchung linearer Gleichungssysteme auf eindeutige Lösbarkeit

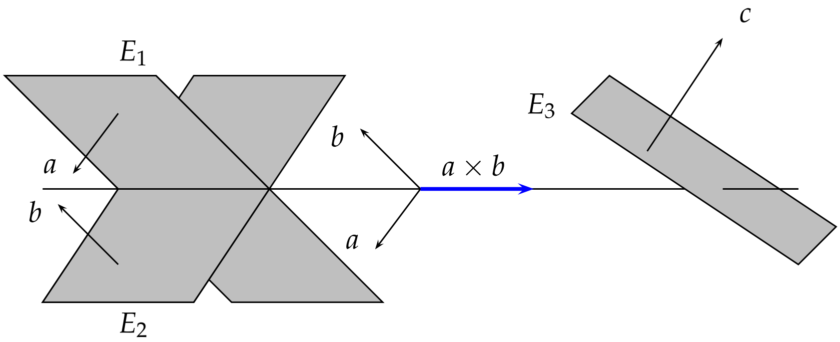
$$(LG) \left\{ \begin{array}{l} (E_1)\ a_1 \cdot x_1 + a_2 \cdot x_2 + a_3 \cdot x_3 = d_1 \\ (E_2)\ b_1 \cdot x_1 + b_2 \cdot x_2 + b_3 \cdot x_3 = d_2 \\ (E_3)\ c_1 \cdot x_1 + c_2 \cdot x_2 + c_3 \cdot x_3 = d_3 \end{array} \right\}$$

- Jede der drei Gleichungen beschreibt eine Ebene im Raum
- eindeutige Lösbarkeit: eindeutiger Schnittpunkt der Ebenen

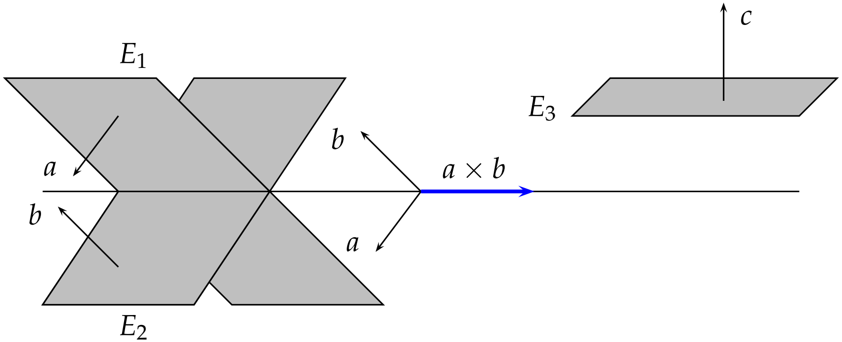
$$a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

- a Normalenvektor der Ebene E_1
- analog b, c für die Ebenen E_2 und E_3
- E_1 und E_2 dürfen nicht parallel sein
- also $a \times b \neq 0$
- $a \times b \neq 0$ zeigt in Richtung der Schnittgeraden $E_1 \cap E_2$
- diese darf nicht parallel zu E_3 sein
- also

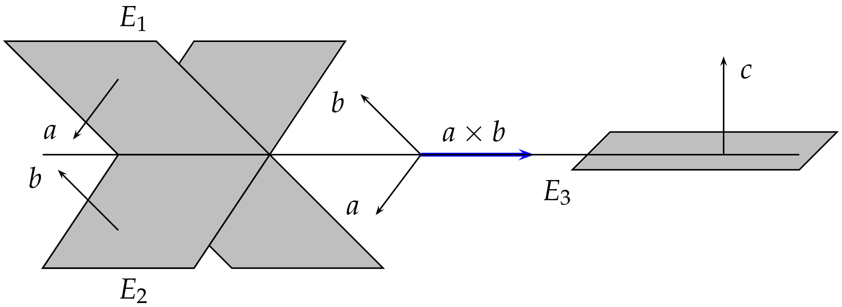
$$\langle (a \times b), c \rangle = \det(a, b, c) \neq 0.$$



Eindeutiger Schnittpunkt: $\langle a \times b, c \rangle \neq 0$



Kein Schnittpunkt: $\langle a \times b, c \rangle = 0$



Schnittgerade, kein eindeutiger Schnittpunkt: $\langle a \times b, c\rangle = 0$

Satz 2.95: Notwendig und hinreichend dafür, dass das lineare Gleichungssystem *LG* eine eindeutige Lösung besitzt, ist die Bedingung

$$\det(a, b, c) \neq 0.$$

Kapitel 3

Algebraische Strukturen

3.1 Gruppen

3.1.1 Grundlagen

- Mathematik: Strukturwissenschaft
- Untersuchung abstrakter Objekte
- Resultate der Algebra wichtig für angewandte Mathematik
- daher: wir behandeln kurz wesentliche algebraische Strukturen
- Beispiel hierfür: Gruppe

Definition 3.1:

1. Sei M eine Menge, und $\circ : M \times M \to M$ eine Abbildung, $(x, y) \mapsto \circ(x, y)$. Eine solche Abbildung heiße *Verknüpfung*; wir schreiben zur Abkürzung statt $\circ(x, y)$ einfacher $x \circ y$.

2. Für M und $\circ$ wie oben heißt das Paar $(M, \circ)$ eine *Gruppe*, wenn gilt:

   G1 (Assoziativität): Die Verknüpfung ist assoziativ, d. h. es gilt:

   $$\forall x, y, z \in M : \ (x \circ y) \circ z = x \circ (y \circ z)$$

   G2 (Neutralelement): Es existiert ein neutrales Element $e \in M$ so dass

   $$\forall x \in M : \ x \circ e = x$$

   G3 (Inverses Element): $\forall x \in M \, \exists x'$ mit

   $$x \circ x' = e.$$

3. Gilt für eine Gruppe $G = (M, \circ)$, dass $x \circ y = y \circ x \, \forall \, x, y \in M$, dann heißt G *abelsche Gruppe* oder *kommutative Gruppe*.

Bemerkung 3.2:

- Nachweis der Gruppeneigenschaft erfordert zu zeigen, dass $\circ$ wirklich eine Verknüpfung ist
- konkret: Man zeige, dass

1. $x \circ y$ für alle x, y existiert und eindeutig festgelegt ist.
2. das Ergebnis $x \circ y$ in M liegt.

- in diesem Fall: die Abbildung $\circ$ ist „wohldefiniert“

Beispiel 3.3:

1. $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$ und $(\mathbb{R}, +)$ Gruppen
2. - $M = \{-5, -4, \ldots, 0, \ldots, 4, 5\}$ zusammen mit der Addition bildet *keine* Gruppe

- **G1 – G3** sind erfüllt
- aber: wegen z. B. $4 + 4 = 8 \notin M$ ist „$+$“ *keine* Verknüpfung auf M
- also: keine Gruppe

Satz 3.4: Sei $G = (M, \circ)$ eine Gruppe. Dann gilt

1. $M \neq \emptyset$
2. Das inverse Element kommutiert, d. h. $\forall x \in M : x' \circ x = x \circ x'$.
3. Das neutrale Element kommutiert, also $\forall x \in M : x \circ e = e \circ x$.
4. Das inverse Element ist eindeutig, d. h. $\forall x$ existiert genau ein $x' \in M : x' \circ x = e$.
5. Das neutrale Element e ist eindeutig.

Beweis *OHP*

Bemerkung 3.5:

1. Man spricht von „dem“ inversen Element; Bezeichnung: a^{-1}
2. Bei einer Gruppe $G = (M, \circ)$ schreibt man oft kurz $a \in G$ (eigentlich richtig: $a \in M$)

Beispiel 3.6:

1. kleinste Gruppe: $(\{1\}, \cdot)$ (triviale Gruppe)
2.
   - $(\mathbb{N}, +)$ ist keine Gruppe
   - z. B. ist das additiv Inverse zu 1 keine natürliche Zahl (-1)
   - ein Inverses in $\mathbb{N}$ wäre auch Inverses in $\mathbb{Z}$
   - Widerspruch zur Eindeutigkeit des Inversen in $\mathbb{Z}$
3. $\mathbb{R}^n$ bildet mit der Vektoraddition eine Gruppe
4. $(\mathbb{R}, \cdot)$ keine Gruppe, da **G3** nicht erfüllt

5. $\mathbb{R}^3$ mit Kreuzprodukt keine Gruppe, da das Vektorprodukt nicht assoziativ ist

Bemerkung 3.7: $(\mathbb{R} \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe.

Beweis Wir zeigen **G1 – G3**. **G1** ist klar, ebenso **G2** mit $e = 1$. Mit $a^{-1} = 1/a$ gilt **G3**. Man beachte, dass $a \neq 0$ wegen $a \in \mathbb{R} \setminus \{0\}$ n. V. ∎

3.1.2 Vertiefung: Endliche Gruppen und Restklassen

- kurze Wiederholung von Inhalten aus „Mathematische Grundlagen"
- für $m \in \mathbb{N}$ definiere $\overline{k} = \{k + m\mathbb{Z}\}$ als Äquivalenzklasse von $a \sim b :\Leftrightarrow (a - b) \mod m = 0$
- disjunkte Zerlegung von $\mathbb{Z}$ in Äquivalenzklassen
- $\mathbb{Z}/m\mathbb{Z} := \{\overline{0}, \ldots, \overline{m-1}\}$

- definiere Verknüpfung ⊕ durch $\overline{k} \oplus \overline{l} := \overline{k+l}$.

Satz 3.8: Für $m \in \mathbb{N}$ ist $(\mathbb{Z}/m\mathbb{Z}, \oplus)$ eine abelsche Gruppe.

Beweis *OHP*

- endliche Gruppen werden häufig über Verknüpfungstafeln definiert
- Beispiel: $\mathbb{Z}/2\mathbb{Z}$ bzw. $\mathbb{Z}/4\mathbb{Z}$

$\oplus$	$\overline{0}$	$\overline{1}$
$\overline{0}$	$\overline{0}$	$\overline{1}$
$\overline{1}$	$\overline{1}$	$\overline{0}$

$\oplus$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{0}$	$\overline{1}$	$\overline{2}$

- Ist die Gruppe abelsch, dann ist die Verknüpfungstafel symmetrisch
- $m\mathbb{Z} = \{\ldots, -2m, -m, 0, m, 2m, \ldots\}$ Gruppe bzgl. $+$
- Gruppe in Gruppe

Definition 3.9: Sei $G = (M, \circ)$ eine Gruppe und $M' \subseteq M$. Bildet $U = (M', \circ)$ eine Gruppe, so heißt U *Untergruppe* von G, Schreibweise $U \leq G$.

Folgerung 3.10: Für G und U wie oben gilt $U \leq G$ genau dann, wenn gilt:

1. $M' \neq \emptyset$
2. $\forall a, b \in M' : a \circ b \in M'$
3. $\forall a \in M' : a^{-1} \in M'$

Beweis *OHP*

$\oplus$	$\overline{0}$	$\overline{2}$
$\overline{0}$	$\overline{0}$	$\overline{2}$
$\overline{2}$	$\overline{2}$	$\overline{0}$

Abbildung 3.1: Verküpfungstafel von $(\mathbb{Z}_2, \oplus)$

Beispiel 3.11: Offensichtlich gilt $(m\mathbb{Z}, +) \leq (\mathbb{Z}, +) \leq (\mathbb{Q}, +) \leq (\mathbb{R}, +)$.

- ein weiteres Mal die Verknüpfungstafel von $\mathbb{Z}/4\mathbb{Z}$
- $\mathbb{Z}_2 := (\{\overline{0}, \overline{2}\}, \oplus)$ bildet Untergruppe
- Verknüpfungstafel der Untergruppe gleicht der von $\mathbb{Z}/2\mathbb{Z}$
- beide Gruppen nicht gleich (verschiedene Restklassen)

- aber: gleiche Struktur

Definition 3.12: Zwei Gruppen $G_1 = (M_1, \circ)$ und $G_2 = (M_2, +)$ heißen *isomorph*, Schreibweise $G_1 \simeq G_2$, wenn es eine bijektive Abbildung $\phi : G_1 \to G_2$ gibt mit

$$\phi(a \circ b) = \phi(a) + \phi(b) \forall a, b \in G_1.$$

Bemerkung 3.13:

- „isomorph“ bedeutet “gleichgestaltig“
- ϕ heißt Gruppenisomorphismus
- ϕ „übersetzt“ zwischen den beiden Gruppen und respektiert dabei die Gruppenstruktur

Beispiel 3.14:

1. Isomorphie ist eine Äquivalenzrelation.

2.
- $\mathbb{Z}/2\mathbb{Z} \simeq \mathbb{Z}_2$
- Gruppenisomorphismus: $\phi(\overline{0}_2) = \overline{0}_4, \phi(\overline{1}_2) = \overline{2}_4$.
- Indizes bezeichnen jeweils den Modul m (wg. $\overline{0}_2 \neq \overline{0}_4$)
- $\mathbb{Z}/2\mathbb{Z} \simeq \mathbb{Z}_2 \leq \mathbb{Z}/4\mathbb{Z}$.

3.
- $(\mathbb{R}^4, +)$ und $(\mathbb{R}^2 \times \mathbb{R}^2, +)$ Gruppen
- formal nicht gleich
- aber: Gruppen isomorph durch
  $\varphi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^4$, $\varphi((x_1, x_2), (x_3, x_4)) := (x_1, x_2, x_3, x_4)$
- identifiziere $\mathbb{R}^4$ und $\mathbb{R}^2 \times \mathbb{R}^2$

4.
- $\mathbb{C} \neq \mathbb{R}^2$
- aber: $(\mathbb{C}, +) \simeq (\mathbb{R}^2, +)$ durch den Isomorphismus $\varphi : \mathbb{C} \to \mathbb{R}^2, \varphi(a + ib) = (a, b)^T$.
- Veranschaulichung der komplexen Zahlen als Ebene

Bemerkung 3.15 (Produktgruppen):

- $G_1 := (A, +)$ und $G_2 := (B, \circ)$
- definiere Verknüpfung auf $A \times B$

$$\begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \star \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} := \begin{pmatrix} a_1 + a_2 \\ b_1 \circ b_2 \end{pmatrix}, \quad a_1, a_2 \in A,\ b_1, b_2 \in B\,.$$

- $(A \times B, \star)$ ist Gruppe
- z. B.: $(\mathbb{R}^2, +)$ aus $(\mathbb{R}, +)$
- Enthält A m und B n Elemente, so enthält $A \times B$ mn Elemente
- also enthält $(\mathbb{Z}/2\mathbb{Z})^2$ genau 4 Elemente
- aber : $(\mathbb{Z}/2\mathbb{Z})^2 \not\simeq \mathbb{Z}/4\mathbb{Z}$

3.2 Körper

- jetzt: algebraische Strukturen ausgehend von den reellen Zahlen
- neben Addition auch Multiplikation definiert
- Eigenschaften von $(\mathbb{R}, +, \cdot)$:

1. $(\mathbb{R}, +)$ bildet eine abelsche Gruppe.
2. $(\mathbb{R} \setminus \{0\}, \cdot)$ bildet eine abelsche Gruppe (vgl. Bemerkung 3.7).
3. Es gelten die *Distributivgesetze*

$$\begin{aligned} \forall x, y, z \in \mathbb{R} : (x + y) \cdot z &= (x \cdot z) + (y \cdot z) \\ x \cdot (y + z) &= (x \cdot y) + (x \cdot z) \end{aligned}$$

Jedes Tripel $(M, \oplus, \odot)$, das die obigen drei Bedingungen erfüllt, wird *Körper* genannt:

Definition 3.16: Sei M eine nichtleere Menge. M zusammen mit zwei Verküpfungen $\oplus, \odot : M \times M \to M$ heißt *Körper*, wenn gilt :

1. $(M, \oplus)$ bildet eine kommutative Gruppe.
2. Nach Ausschluss des neutralen Elements der Verknüpfung $\oplus$ bilden die restlichen Elemente von M eine kommutative Gruppe bzgl. $\odot$.
3. Für $\oplus$ und $\odot$ gelten die *Distributivgesetze*:
   a) $\forall x, y, z \in M : \quad (x \oplus y) \odot z = (x \odot z) \oplus (y \odot z)$
   b) $\forall x, y, z \in M : \quad z \odot (x \oplus y) = (z \odot x) \oplus (z \odot y)$

- neutrales Element der Addition $\oplus$: „0“
- neutrales Element der Multiplikation $\odot$: „1“
- Schreibweise ⓪ und ①, um Verwechselungen mit den Zahlen 0 und 1 zu vermeiden

- in der Definition des Körpers: $⓪$ wird aus Eigenschaft 2 ausgeschlossen
- Grund ist der folgende Satz

Satz 3.17: Sei K ein Körper und $a \in K$. Dann gilt $a \odot ⓪ = ⓪$.

Beweis *OHP*

OHP

Beispiel 3.18:

1. Nach Konstruktion ist $(\mathbb{R}, +, \cdot)$ ein Körper.

2. $(\mathbb{Q}, +, \cdot)$ ist ein Körper, ebenso $(\mathbb{C}, +, \cdot)$ mit der komplexen Multiplikation.

3. $(\mathbb{Z}, +, \cdot)$ ist kein Körper, da $(\mathbb{Z} \setminus \{0\}, \cdot)$ keine Gruppe ist, weil z. B. $2^{-1} = \frac{1}{2} \notin \mathbb{Z}$, also **G3** nicht erfüllt ist.

- Es gibt endliche Gruppen.

- Gibt es auch endliche Körper?

Satz 3.19: Wir definieren auf $\mathbb{Z}/m\mathbb{Z}$ die Verknüpfung $\odot$ durch

$$\overline{a} \odot \overline{b} := \overline{a \cdot b} \, \forall \, \overline{a}, \overline{b} \in \mathbb{Z}/m\mathbb{Z}\,.$$

Ist m eine Primzahl, dann ist $(\mathbb{Z}/m\mathbb{Z}, \oplus, \odot)$ ein Körper.

- Beweis: siehe Literatur
- endliche Körper ohne größere Bedeutung für die Lineare Algebra

3.3 Vektorräume

- bisher: Zahlraum $\mathbb{R}^n$
- jetzt: Verallgemeinerung zum Vektorraum
- Vorgehen wie bei Gruppen und Körpern: Man sammelt alle wesentlichen algebraischen Eigenschaften des $\mathbb{R}^n$ und erhebt diese zur Definition

Definition 3.20: Sei K ein beliebiger Körper. Eine nicht-leere Menge V zusammen mit den beiden Abbildungen

$$\begin{aligned} \oplus : V \times V \to V, &\quad (x, y) \mapsto x \oplus y \in V \\ \text{und } \odot : K \times V \to V, &\quad (\lambda, x) \mapsto \lambda \odot x \in V \end{aligned}$$

heißt *Vektorraum über K* oder *K-Vektorraum*, wenn folgende Axiome gelten:

1. $(V, \oplus)$ ist eine kommutative Gruppe.

2. $\forall\, \lambda, \mu \in K$ und $x \in V$ gilt $\lambda \odot (\mu \odot x) = (\lambda\mu) \odot x$, wobei mit $\lambda\mu$ die Multiplikation aus K gemeint ist.

3. $\forall\, x \in V$ gilt ① $\odot\, x = x$ (① ist das neutrale Element der Multiplikation aus K).

4. $\forall\, \lambda \in K,\ x, y \in V$ gilt: $\lambda \odot (x \oplus y) = (\lambda \odot x) \oplus (\lambda \odot y)$.

5. $\forall\, \lambda, \mu \in K,\ x \in V$ gilt $(\lambda + \mu) \odot x = (\lambda \odot x) \oplus (\mu \odot x)$.

- üblicherweise: $K = \mathbb{R}$ oder $\mathbb{C}$
- Im Fall $K = \mathbb{R}$: V „reeller Vektorraum“
- Im Fall $K = \mathbb{C}$: V „komplexer Vektorraum“
- allgemein: Vektorraum über dem Körper K
- wir identifizieren häufig $(V, \oplus, \odot)$ mit V

Definition 3.21: Die Elemente der Menge V eines Vektorraums heißen *Vektoren*, die Elemente aus K *Skalare*.

Bemerkung 3.22:

- im Vektorraum: Multiplikation eines Skalars mit einem Vektor muss definiert sein
- die Multiplikation zweier Vektoren muss nicht definiert sein

Beispiel 3.23:

1. $\mathbb{R}^n$ ist nach Konstruktion ein reeller Vektorraum
2. $(\mathbb{C}^n, +, \cdot)$ ist ein komplexer Vektorraum
3. Jeder Körper ist ein Vektorraum über sich selbst mit den Verknüpfungen des Körpers $\oplus$ als Vektoraddition und $\odot$ als Multiplikation eines Vektors mit einem Skalar.

Beispiel 3.24:

1.

- $K = \mathbb{R}$ und $V = \mathcal{F}[a, b]$
- Summe und skalares Vielfaches von Funktionen werden punktweise definiert:

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

- Dann ist V ein reeller Vektorraum.
- Vektorraumaxiome durch Zurückführung auf reelle Zahlen

2.
- Dann ist V ein reeller Vektorraum.
- Vektorraumaxiome werden von V' „geerbt“
- zum Beweis: Summen und Vielfache stetiger Funktionen sind wieder stetig

Beispiel 3.25:

- Menge der $m \times n$-Matrizen mit Elementen im Körper K
- komponentenweise Addition: Für $A = (a_{ij})$ und $B = (b_{ij})$ sei $A + B := (a_{ij} + b_{ij})$.
- $\lambda A := (\lambda a_{ij})$
- Menge der $m \times n$-Matrizen ist K-Vektorraum
- Bezeichnung: $K^{m \times n}$

Beispiel 3.26: *OHP*

- analog zu Gruppen: es kann Vektorräume innerhalb von Vektorräumen geben
- Untervektorräume

Definition 3.27: Eine Teilmenge $U \subseteq V$ heißt *Untervektorraum* oder *Unterraum* von V, wenn $U \neq \emptyset$ (leere Menge) und $\forall\, x, y \in U$ und alle $\lambda \in K$ gilt:

$$x \oplus y \in U,\ \lambda \odot x \in U.$$

Bemerkung 3.28: $U \subseteq V$ ist genau dann ein Vektorraum, wenn U ein Untervektorraum von V im Sinne von Definition 3.27 ist.

Beweis *OHP*

Beispiel 3.29: *OHP*

Bemerkung 3.30:

- sei U ein Unterraum von V
- dann gilt $0 \in U$ wegen $0 \cdot x = 0$ für irgendein $x \in U$
- zudem ist $-x \in U$ wegen $(-1) \cdot x = -x$

Bemerkung 3.31: Es sind $\{0\}$ und V Unterräume von V.

Bemerkung 3.32: Sind U_1 und U_2 Unterräume von V, dann ist $U_1 \cap U_2$ ein Unterraum von V.

Beweis *OHP*

- Vereinigungen von Untervektorräumen sind i. A. keine UVR

Beispiel 3.33:

$$U_1 = \{(x, 0);\ x \in \mathbb{R}\}$$
$$U_2 = \{(0, y);\ y \in \mathbb{R}\}$$

- beides Untervektorräume
- $U_v = U_1 \cup U_2$
- in U_v: mindestens eine Komponente der Vektoren ist 0
- $x_1 = (1, 0)^T \in U_v$ und $x_2 = (0, 1)^T \in U_v$
- aber: $x_1 + x_2 = (1, 1)^T \notin U_v$
- Summen von Untervektorräumen sind Untervektorräume

Definition 3.34:

1. Seien U_1, U_2 Unterräume von V. Die *Summe* von U_1 und U_2 ist dann definiert durch

$$U_1 + U_2 := \{v \in V \mid \exists u_1 \in U_1, u_2 \in U_2 : v = u_1 + u_2\}.$$

2. Ein Vektorraum W heißt *direkte Summe* von U_1 und U_2, Schreibweise $W = U_1 \oplus U_2$, wenn gilt $W = U_1 + U_2$ und $U_1 \cap U_2 = \{0\}$.

Satz 3.35: Sind U_1 und U_2 Unterräume von V, ist $U_1 + U_2$ ein Unterraum von V.

Beweis Übungsaufgabe ∎

Beispiel 3.36:

$$U_1 = \{(x, 0);\ x \in \mathbb{R}\}$$
$$U_2 = \{(0, y);\ y \in \mathbb{R}\}$$

- Untervektorräume von $\mathbb{R}^2$

$$U_p = U_1 \oplus U_2 = \{(x+0, 0+y);\ x, y \in \mathbb{R}\} = \mathbb{R}^2$$

Bemerkung 3.37: Gilt $V = U_1 \oplus U_2$, ist die Darstellung $v = u_1 + u_2$ mit $u_1 \in U_1$ und $u_2 \in U_2$ eindeutig für alle $v \in V$.

Beweis *OHP*

3.4 Lineare Unabhängigkeit, Basis, Dimension

3.4.1 Lineare Unabhängigkeit

- intuitiv: Ebene besitzt zwei Dimensionen
- man möchte $\mathbb{R}^2$ die Dimension 2 zuordnen
- entsprechend möchte man $\mathbb{R}^3$ die Dimension 3 zuordnen
- Dimension eines allgemeinen Vektorraums?
- einige Vorarbeit erforderlich
- einfache Beobachtung: im $\mathbb{R}^3$ kann man jeden Vektor v als gewichtete Summe der kanonischen Einheitsvektoren schreiben:

$$v = x\begin{pmatrix}1\\0\\0\end{pmatrix} + y\begin{pmatrix}0\\1\\0\end{pmatrix} + z\begin{pmatrix}0\\0\\1\end{pmatrix}.$$

- Verallgemeinerung in folgender Definition:

Definition 3.38: Seien $v_1, \dots, v_r \in V$ Vektoren eines K-Vektorraums V:

(a) Lässt sich $v \in V$ als Summe dieser Vektoren mit Vorfaktoren darstellen,

$$v = \lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_r v_r \quad \text{mit} \quad \lambda_i \in K,$$

heißt v *Linearkombination* von $v_1, \dots, v_r$.

(b) Die Menge

$$L(v_1, \dots, v_r) := \{\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_r v_r \,|\, \lambda_i \in K\} \subseteq V$$

aller Linearkombinationen heißt *Lineare Hülle* von $v_1, \dots, v_r$.

Nach Definition ist jede Lineare Hülle ein Unterraum von V.

Beispiel 3.39: *OHP*

Beispiel 3.40: Beispiele in $\mathbb{R}^3$:

$$L(e_1, e_2, e_3) = \{(\lambda_1, \lambda_2, \lambda_3)^T \,|\, \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}\} = \mathbb{R}^3$$
$$L(e_1, e_2) = \{(\lambda_1, \lambda_2, 0)^T \,|\, \lambda_1, \lambda_2 \in \mathbb{R}\} \quad (x_1\text{-}x_2\text{-Ebene})$$
$$L(e_1, e_2, e_1 + e_2) = \{(\lambda_1 + \lambda_3, \lambda_2 + \lambda_3, 0)^T \,|\, \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}\}$$
$$(x_1\text{-}x_2\text{-Ebene})$$

- englischer Ausdruck für „Lineare Hülle“: *span*
- man schreibt oft $span(e_1, e_2)$ statt $L(e_1, e_2)$
- „Spann“ ungebräuchlich
- gebräuchliche Sprechweise: Die Vektoren e_1 und e_2 *spannen die* x_1-x_2-*Ebene auf*.
- Beispiel oben zeigt: $\mathbb{R}^3$ lässt sich als Menge von Linearkombinationen dreier Vektoren auffassen

Definition 3.41:

1. Sei V ein K-Vektorraum und $(v_1, \ldots, v_n)$ ein n-Tupel von Vektoren in V. Gilt $L(v_1, \ldots, v_n) = V$, nennt man $(v_1, \ldots, v_n)$ ein *Erzeugendensystem* von V.

2. V heißt *endlich erzeugt*, wenn es endlich viele Vektoren $v_1, \ldots, v_r$ gibt, so dass $L(v_1, \ldots, v_r) = V$, ansonsten nicht endlich erzeugt.

Beispiel 3.42:

- (e_1, e_2, e_3) ist ein Erzeugendensystem von $\mathbb{R}^3$
- $(0,1)^T$ und $(1,0)^T$ erzeugen den Zahlraum $\mathbb{R}^2$
- $\mathbb{R}^2$ und $\mathbb{R}^3$ endlich erzeugt
- $\mathbb{R}^n$ wird durch $e_1, \ldots, e_n$ (endlich) erzeugt.
- ab jetzt: der Vektorraum sei endlich erzeugt

- es gibt viele Erzeugendensysteme eines Vektorraums
- für $\mathbb{R}^3$: sowohl $E = (e_1, e_2, e_3)$ als auch $\tilde{E} = (e_1, e_2, e_3, e_1 + e_2)$ sind Erzeugendensysteme
- $\tilde{E}$ unnötig groß
- jeder Vektorraum enthält die 0
- schreibe 0 als Linearkombination der Vektoren des Erzeugendensystems
- in E: $0 = 0e_1 + 0e_2 + 0e_3$ (einzige Möglichkeit)
- in $\tilde{E}$: $0 = 0e_1 + 0e_2 + 0e_3 + 0(e_1 + e_2)$, aber auch $0 = (-1)e_1 + (-1)e_2 + 0e_3 + 1(e_1 + e_2)$
- Redundanz in $\tilde{E}$ hängt zusammen mit der Möglichkeit, die 0 auf verschiedene Weise als Linearkombination zu schreiben

Definition 3.43: Sei V ein K-Vektorraum. Ein r-Tupel $(v_1, \ldots, v_r)$ von Vektoren in V heißt *linear unabhängig*, wenn aus $\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_r v_r = 0$ stets folgt, dass $\lambda_1 = \lambda_2 = \cdots = \lambda_r = 0$ gilt.

- 0 eindeutig darstellbar $\Rightarrow$ *jeder* Vektor eindeutig darstellbar

Bemerkung 3.44: Alle Vektoren v aus $L(v_1, \ldots, v_n)$ sind genau dann eindeutig als Linearkombination der Vektoren $v_1, \ldots, v_n$ darstellbar, wenn das n-Tupel $(v_1, \ldots, v_n)$ linear unabhängig ist.

Beweis *OHP*

Bemerkung 3.45: Für linear unabhängige Vektoren gilt:

- Keiner der Vektoren ist eine Linearkombination der übrigen.
- Keiner der Vektoren ist der Nullvektor.

Beispiel 3.46 ($\mathbb{R}^n$): *OHP*

Beispiel 3.47:

- im $\mathbb{R}^n$: das n-Tupel der kanonischen Einheitsvektoren ist linear unabhängig
- $\tilde{E} = (e_1, e_2, e_3, e_1 + e_2)$ nicht linear unabhängig

3.4.2 Nachweis linearer Unabhängigkeit

- zunächst: $\mathbb{K}^n$
- Prototyp aller endlich erzeugten Vektorräume
- gegeben $v_1, \ldots, v_r$, $r \leq n$
- Test auf lineare Unabhängigkeit

$$\lambda_1 v_1 + \cdots + \lambda_r v_r = 0$$

- obige Gleichung aus der Definition ist ein LGS
- Vektoren genau dann linear unabhängig, wenn der Nullvektor die einzige Lösung ist

Beispiel 3.48: *OHP*

Beispiel 3.49: *OHP*

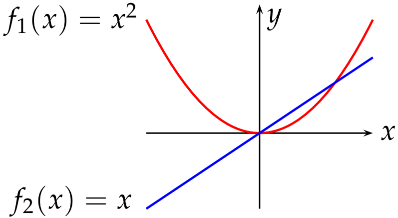
OHP

- jetzt: Vektorraum $\mathcal{C}[a,b]$
- Vektoren sind stetige Funktionen
- Überprüfung linearer Unabhängigkeit ausgehend von der Definition
- zu zeigen: Die Nullfunktion ist nur trivial darstellbar
- bzw. Eine Funktion ist Linearkombination der anderen

Beispiel 3.50: *OHP*

Beispiel 3.51:

$$f_1(x) = x^2; \quad f_2(x) = x$$



OHP

- im Gegensatz zum $\mathbb{K}^n$: es gibt kein festes Rechenschema zur Feststellung linearer (Un)abhängigkeit

- folgende Aussage kann nützlich sein

Bemerkung 3.52: Wir bilden zu den Funktionen $f_1, \ldots, f_n$ für paarweise verschiedene $x_1, \ldots, x_n$ die n Vektoren

$$(f_1(x_1), \ldots, f_1(x_n))^T, \ldots, (f_n(x_1), \ldots, f_n(x_n))^T \in \mathbb{K}^n \,.$$

Sind diese Vektoren linear unabhängig, dann sind die Funktionen $f_1, \ldots, f_n$ selbst linear unabhängig.

Beweis

- $0 = \sum_{i=1}^{n} \lambda_i f_i$

- gilt insbesondere für die x_i

- Vektoren $(f_i(x_j))_{i=1}^n$ linear unabhängig $\Rightarrow \lambda_j = 0$ ■

Beispiel 3.53:

- gegeben $f_1(x) = 1, f_2(x) = \sin x, f_3(x) = \cos x$ auf $[-\pi, \pi]$
- Funktionen linear unabhängig?
- Wähle $x_1 = 0$, $x_2 = \pi/2$ und $x_3 = \pi$

OHP

- Nachweis linearer (Un)abhängigkeit bei Funktionen:
- üblicherweise: Strukturinformationen ausnutzen
- Hilfsmittel aus der Analysis anwenden

3.4.3 Basis und Dimension

Definition 3.54: Sei $V \neq \{0\}$ ein endlich erzeugter K-Vektorraum und $v_1, \ldots, v_n \in V$. Das n-Tupel $\mathcal{B} = (v_1, \ldots, v_n)$ heißt *Basis* oder *minimales Erzeugendensystem* von V, wenn $\mathcal{B}$ linear unabhängig ist und wenn gilt $L(\mathcal{B}) = V$. Weiter sei $\emptyset$ die Basis des Nullvektorraums $\{0\}$.

Sei ab jetzt stets $V \neq \{0\}$.

Definition 3.55: Sei $(v_1, \ldots, v_n)$ eine Basis von V und $v \in V$ mit

$$v = \lambda_1 v_1 + \cdots + \lambda_n v_n,$$

$\lambda_1, \ldots, \lambda_n \in K$. Das n-Tupel der obigen Vorfaktoren $(\lambda_1, \ldots, \lambda_n) \in K^n$ nennt man *Koordinaten* des Vektors v bzgl. der Basis $(v_1, \ldots, v_n)$.

Bemerkung 3.56: Die Koordinaten eines Vektors bzgl. einer gegebenen Basis sind eindeutig.

Beweis *OHP*

- Idee: Dimension als Anzahl der Basisvektoren in einer Basis
- zugleich: Anzahl der Koordinaten eines Vektors
- aber: Dimensionsbegriff nur wohldefiniert, wenn alle Basen gleichviele Elemente enthalten
- Hat überhaupt jeder Vektorraum eine Basis?
- jetzt: Vektorraum sei endlich erzeugt
- 2 Aussagen erforderlich:
- Jeder endlich erzeugte Vektorraum besitzt eine Basis.
- Alle Basen bestehen aus gleich vielen Vektoren (Eindeutigkeit der Basislänge).
- zuerst: Existenz von Basen

Satz 3.57 (Basisergänzungssatz): Sei V ein K-Vektorraum und seien

$$v_1, \dots, v_r, w_1, \dots, w_s$$

Vektoren in V. Ist $(v_1, \dots, v_r)$ linear unabhängig und ist

$$L(v_1, \dots, v_r, w_1, \dots, w_s) = V,$$

dann kann man $(v_1, \dots, v_r)$ durch evtl. Hinzunahme geeigneter Vektoren aus $\{w_1, \dots, w_s\}$ zu einer Basis von V ergänzen.

Beweis *OHP*

Folgerung 3.58: Jeder endlich erzeugte Vektorraum V hat eine Basis.

Beweis *OHP*

Beispiel 3.59:

- $V = \mathbb{R}^3$
-
$$v_1 = (1,0,0); \qquad v_2 = (0,1,0).$$
- v_1 und v_2 sind linear unabhängig
- sei weiter
$$w_1 = (1,1,0); \qquad w_2 = (0,0,1).$$
- $L(v_1, v_2, w_1, w_2) = \mathbb{R}^3$
- also Voraussetzungen des Basisergänzungssatzes erfüllt
- Aussage: (v_1, v_2, w_1), (v_1, v_2, w_2) oder (v_1, v_2, w_1, w_2) eine Basis des $\mathbb{R}^3$

- hier: (v_1, v_2, w_2)
- alle vier Vektoren gemeinsam bilden keine Basis, da linear abhängig

Satz 3.60 (Austauschlemma): Sind $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_m)$ Basen eines K-Vektorraums V, dann gibt es zu jedem v_i ein w_j, so dass aus $(v_1, \ldots, v_n)$ wieder eine Basis entsteht, wenn man in ihr v_i durch w_j ersetzt.

Beweis *OHP*

Beispiel 3.61:

- betrachte

$$v_1 = (1,0,0); \qquad v_2 = (0,1,0); \qquad v_3 = (0,0,1)$$
$$w_1 = (2,0,0); \qquad w_2 = (0,2,0); \qquad w_3 = (0,0,3)$$

- hierbei: (v_1, v_2, v_3) und (w_1, w_2, w_3) Basen des $\mathbb{R}^3$
- entferne v_3 aus der ersten Basis
- Austauschlemma: man kann einen der drei Vektoren w_i einsetzen, so dass eine Basis entsteht
- hier: w_3
- (v_1, v_2, w_3) ist eine Basis des $\mathbb{R}^3$.

- jetzt: Eindeutigkeit der Basislänge in endlich erzeugten Vektorräumen
- Konsequenz: Dimension kann als Anzahl der Basisvektoren definiert werden

Satz 3.62 (Eindeutigkeit der Basislänge): Sind $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_m)$ Basen eines K-Vektorraums V, dann gilt $n = m$.

Beweis *OHP*

Definition 3.63: Besitzt ein Vektorraum $V \neq \{0\}$ eine Basis $(v_1, \ldots, v_n)$, so definieren wir die *Dimension* von V als $\dim(V) := n$; besitzt V keine endliche Basis, dann setzt man $\dim(V) := \infty$. Weiter sei $\dim(\{0\}) := 0$.

- folgende Bemerkung liefert den entscheidenden Hinweis bei der Suche nach Basen

Bemerkung 3.64: Ist $\dim(V) = n$ und sind $v_1, \ldots, v_n$ n linear unabhängige Vektoren in V, so ist $(v_1, \ldots, v_n)$ eine Basis von V.

Beweis *OHP*

Beispiel 3.65: *OHP*

- Anwendung: Dimension von Summen von Untervektorräumen

Satz 3.66: Seien U, W Untervektorräume eines endlich erzeugten Vektorraums. Dann gilt

$$\dim(U+W) = \dim(U) + \dim(W) - \dim(U \cap W)\,.$$

Beweis *OHP*

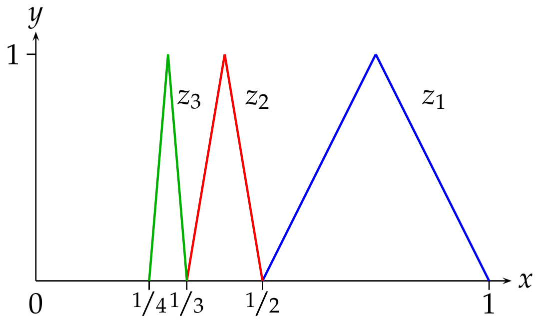
Folgerung 3.67: Sei V ein Vektorraum mit $\dim(V) = n$, U_1, U_2 Unterräume und $V = U_1 \oplus U_2$. Dann gilt $\dim(U_1) + \dim(U_2) = n$.

3.4.4 Exkurs: Nicht endlich erzeugte Vektorräume

- wir betrachten $\mathcal{C}[0,1]$
- definiere stetige Funktionen z_n durch

$$z_n(x) = \left\{ \begin{array}{c} 0,\ x \notin [\frac{1}{n+1}, \frac{1}{n}] \\ 1,\ \ x = (\frac{1}{n+1} + \frac{1}{n})/2 \\ \text{linear auf den Zwischenintervallen} \end{array} \right\}, n \in \mathbb{N}\,.$$

- $(z_1, \ldots, z_n)$ ist linear unabhängig für jedes $n \in \mathbb{N}$.
- Funktionsgraphen:



- Funktionswerte außerhalb sind Null

Satz 3.68: Der Vektorraum $\mathcal{C}[0,1]$ ist nicht endlich erzeugt.

Beweis

- sonst: es gibt $\mathcal{B}$ Basis mit m Vektoren

- Bemerkung 3.64: $\mathcal{B} = (z_1, \ldots, z_m)$ ist Basis, da linear unabhängig
- aber: für $0 \leq x < \frac{1}{m+1}$ gilt $z_i(x) = 0$, $1 \leq i \leq m$
- $f(x) \equiv 1 \notin L(z_1, \ldots, z_m)$
- Widerspruch ■

- nach diesem Ergebnis: Basen auf allgemeinen Vektorräumen?
- vorbereitend: Erweiterung der Begriffs des n-Tupels

Definition 3.69: Für eine beliebige geordnete Indexmenge I (z. B. die reellen Zahlen) und einer Menge $X \neq \emptyset$ bezeichnet man $(a_i)_{i \in I}$ mit $a_i \in X \; \forall i \in I$ als *Familie*.

- Indexmenge endlich: Familie ist n-Tupel

- Indexmenge abzählbar unendlich: Familie ist Folge
- folgende Definition mittels Zurückführung auf den endlich erzeugten Fall:

Definition 3.70: Sei V ein Vektorraum und $(v_i)_{i\in I}$ eine Familie von Vektoren.

1. **(Lineare Hülle)** Sei $L(v_i)_{i\in I}$ die Menge aller Vektoren, die sich als Linearkombination von endlich vielen Vektoren $v_1, \ldots, v_r \in (v_i)_{i\in I}$ darstellen lassen.
2. **(Erzeugendensystem)** Gilt $L(v_i)_{i\in I} = V$, so nennt man $(v_i)_{i\in I}$ ein *Erzeugendensystem*.
3. **(Lineare Unabhängigkeit)** Eine beliebige Familie heißt linear unabhängig, wenn jede endliche Teilfamilie linear unabhängig ist.

Bemerkung 3.71:

- eindeutige Darstellbarkeit eines Vektors als Linearkombination?
- gilt in beliebigen Vektorräumen durch Auswahl einer Teilfamilie

Beispiel 3.72:

- $(f_i)_{i\in\mathbb{N}}$ (s. o.) ist linear unabhängig, aber kein Erzeugendensystem
- z. B. $g(x) = 1$ kann nicht dargestellt werden:
- auf $[1/2, 1]$ ist nur die Funktion f_1 ungleich Null
- aber: kein Vielfaches von f_1 ist g
- Basisbegriff für beliebige Vektorräume

Definition 3.73: Sei V ein K-Vektorraum. Eine Familie $\mathcal{B} = (v_i)_{i\in I} \subseteq V$ heißt *Basis* oder *minimales Erzeugendensystem* von V, wenn $\mathcal{B}$ linear unabhängig ist und wenn gilt $L(\mathcal{B}) = V$.

Bemerkung 3.74:

- $f_1, \dots$ ist kein Erzeugendensystem

- keine Basis von $\mathcal{C}[0,1]$
- $\mathcal{C}^k[0,1]$ ist ein Vektorraum
- $\mathcal{C}^{k+1}[0,1]$ ist ein Unterraum von $\mathcal{C}^k[0,1]$ für alle k
- keiner dieser Räume endlich erzeugt
- Konstruktion von Funktionenfolgen analog zu $f_1, \ldots$ möglich
- die Funktionenfolgen sind alle keine Basen

Satz 3.75: Jeder Vektorraum hat eine Basis.

- Beweistechniken aus dem endlich erzeugten Fall nicht anwendbar
- Hilfsmittel aus der Mengenlehre
- Beweis nicht konstruktiv

Bemerkung 3.76:

- Folge statt Familie: nur abzählbare Basen möglich
- a priori nicht zu rechtfertigen
- $\mathcal{C}^k[a,b]$ besitzen tatsächlich überabzählbare Basen
- keine von ihnen lässt sich konkret angeben

3.4.5 Exkurs: Hyperebenen im $\mathbb{R}^n$

- Definition von Hyperebenen in allgemeinen endlich erzeugten Vektorräumen mit der jetzt zur Verfügung stehenden Theorie möglich
- Ausgangspunkt: deute Gerade in der Ebene als verschobenen Untervektorraum
- Dimension: $2-1=1$
- analog: deute Ebene im Raum als verschobenen Untervektorraum
- Dimension: $3-1=2$

Definition 3.77: Sei V ein K-Vektorraum, $\dim(V)=n$, U ein Untervektorraum mit $\dim(U)=n-1$ und $p\in V$ beliebig.

1. Dann heißt

$$H(U;p):=p+U:=\left\{x\in V \,\middle|\, x=p+u, u\in U\right\}$$

Hyperebene, p heißt *Aufpunkt* der Hyperebene.

2. Existieren linear unabhängige Vektoren $v_1, \ldots, v_{n-1}$, so dass

$$H = \left\{ x \in V \,\middle|\, x = p + \sum_{i=1}^{n-1} \alpha_i v_i, \alpha_i \in K \right\} \tag{3.1}$$

gilt, dann heißen $v_1, \ldots, v_{n-1}$ *Richtungsvektoren* von H, die Darstellung einer Hyperebene in Formel (3.1) heißt *Parameterform* der Hyperebene.

Bemerkung 3.78: Seien U, V, p wie oben.

1. Jede Hyperebene H besitzt eine Parameterdarstellung.
2. Richtungsvektoren sind nicht eindeutig.
3. Sei $H = p + U$ eine Hyperebene. Dann ist jeder Vektor $p' := p + u$ mit $u \in U$ ebenfalls Aufpunkt.

Beweis *OHP*

Beispiel 3.79:

1. - in $V = \mathbb{R}^2$: eine Hyperebene besitzt genau einen Richtungsvektor
   - deswegen:

$$H = \{x \in \mathbb{R}^2 \,|\, x = p + \alpha v, \alpha \in \mathbb{R}\}$$

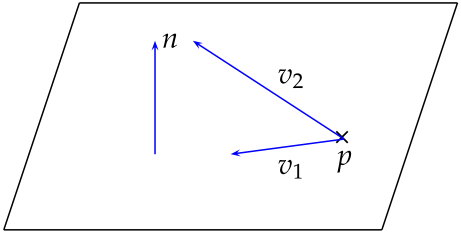
   - Ergebnis: Gerade in der Ebene

2. - Hyperebene in $\mathbb{R}^3$:

$$H = \left\{x \in \mathbb{R}^3 \,|\, x = p + \alpha_1 v_1 + \alpha_2 v_2,\ \alpha_1, \alpha_2 \in \mathbb{R}\right\}.$$

   - Ebenen in $\mathbb{R}^3$: Spezialfälle von Hyperebenen
   - lineare Unabhängigkeit: die beiden Richtungsvektoren der Ebene zeigen in verschiedene Richtungen

- anschaulich klar: ansonsten erhielte man eine Gerade



Bemerkung 3.80:

- Hyperebenen sind genau dann Untervektorräume, wenn sie 0 enthalten
- i. A. demnach keine Untervektorräume
- keine Dimension nach unserer Definition

- ab jetzt: $V = \mathbb{R}^n$

Satz 3.81: Sei $V = \mathbb{R}^n$ und H eine Hyperebene. Dann existiert ein *Normalenvektor* $w \in \mathbb{R}^n \setminus \{0\}$ mit

$$H = \{\langle x, w\rangle = \langle p, w\rangle\}$$

Der Vektor w ist bis auf seine Länge eindeutig festgelegt.

Beweis

- wird nachgeholt ■

Definition 3.82: Die Gleichung $\langle x, w\rangle = \langle p, w\rangle$ heißt *Normalgleichung* oder auch *Normalform* der Hyperebene. Eine Normalform mit $\|w\| = 1$ heißt *Hessesche Normalform*.

Bemerkung 3.83:

1. - Sei $w = (w_i)_{i=1}^n$ ein Normalenvektor von H, $p \in H$

- $\langle p, w \rangle = c \in \mathbb{R}$
- Normalfom von H:

$$\sum_{i=1}^{n} w_i x_i = c \,.$$

2.
   - gegeben lineares Gleichungssystem mit m Gleichungen und n Unbekannten
   - Zeilen der linearen Gleichungssystems: Normalformen von Hyperebenen
   - Lösungsvektor $x = (x_i)_{i=1}^{n}$ erfüllt alle Gleichungen simultan
   - x liegt also im Schnitt der Hyperebenen
   - schon bekannt: bei 3×3-Systemen besteht die Lösungsmenge aus dem Schnitt dreier Ebenen
   - hier: Verallgemeinerung der geometrischen Deutung auf den allgemeinen Fall

- zur Umrechnung einer Normalform aus einer Parameterform bei Hyperebenen

- Ansatz: Normalform ist parameterfrei
- also: deute die Parametergleichung als lineares Gleichungssystem und eliminiere alle Parameter
- Ergebnis: Normalform
- im Fall $\mathbb{R}^2$ und $\mathbb{R}^3$: vereinfachte Berechnungsweisen möglich

Beispiel 3.84: *OHP*

- zur Umrechnung einer Normalform in eine Parameterform: man finde $n-1$ linear unabhängige Vektoren v_i in H

- $v_i \perp w$

- Verfahren zur Konstruktion von v_i:

- geg. Normalenvektor $w = (w_i)_{i=1}^n$.

- wegen $w \neq 0$ existiert ein $w_i \neq 0$

- Für jeden auf w senkrechten Vektor führe man folgende Schritte durch

1. Man vertausche w_i mit einer anderen Komponente w_j; $j \neq i$.

2. Man ändere das Vorzeichen von w_i im so erzeugten Vektor.

3. Man setze alle Komponenten $w_k = 0$ für $k \neq i, j$.

- w hat Länge n
- genau $n-1$ verschiedene Möglichkeiten, $j \neq i$ zu wählen
- man erhält $n-1$ Vektoren
- nach Konstruktion senkrecht auf w

3.5 Polynome

- einfachste Anwendung der Vektorraumtheorie bei Funktionen: Vektorraum der Polynome
- einfache Funktionen
- Probleme können mit algebraischen Mitteln behandelt werden

Definition 3.85:

1. Ein *Polynom* oder ganzrationale Funktion $p : \mathbb{K} \to \mathbb{K}$ ist eine Funktion der Gestalt

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = \sum_{k=0}^{n} a_k x^k \tag{3.2}$$

mit den *Koeffizienten* $a_k \in \mathbb{K}, k = 0, 1, 2, \ldots, n$. Der Koeffizient a_n heißt *Leitkoeffizient*. Im Fall $a_n = 1$ heißt p *normiert*.

2. Die Funktion $p(x) \equiv 0$ heißt *Nullpolynom*.

3. Ist $a_n \neq 0$, so heißt n *Grad* des Polynoms, Schreibweise $deg(p) = n$. Weiter sei $deg(0) = -\infty$.

4. Sei P_n die Menge aller Polynome mit einem Grad von höchstens n.

Bemerkung 3.86:

1.
   - in der Algebra: man unterscheidet zwischen Polynom (endliche Koeffizientenfolge) und Polynomfunktion (siehe unsere Definition)
   - wichtig, wenn man endliche Körper zulässt
   - wir betrachten nur $\mathbb{K}$ und identifizieren deswegen Polynom und Polynomfunktion

2. Durch den Übergang des Laufindizes von k nach $n-k$ in Formel (3.2) erhält man sofort

$$p(x) = \sum_{k=0}^{n} a_k x^k = \sum_{k=0}^{n} a_{n-k} x^{n-k}$$

- Ziel der folgenden Betrachtungen: ein Polynom vom Grad n hat höchstens n reelle Nullstellen

- zur Vorbereitung:

Satz 3.87: Sei $x_0 \in \mathbb{K}$ beliebig und p ein Polynom mit $deg(p) = n \in \mathbb{N}$. Dann gilt für alle $x \in \mathbb{K}$:

$$p(x) = (x - x_0)p_{n-1}(x) + r$$

mit einem Polynom p_{n-1} vom Grad $n-1$ und $r \in \mathbb{K}$.

Beweis *OHP*

Bemerkung 3.88: Seien x_0, p, n wie in Satz 3.87.

1. Es gilt $r = p(x_0)$.

2. Ist x_0 Nullstelle von p, dann gilt

$$p(x) = (x - x_0)p_{n-1}(x).$$

Beweis

1. Es gilt $p(x_0) = (x_0 - x_0)p_{n-1}(x_0) + r = r$.

2. Man hat dann $0 = p(x_0) = r$ nach 1.) ■

- fortgesetzte Anwendung auf p_{n-1} liefert:

Satz 3.89: Hat ein Polynom p mit $deg(p) = n \in \mathbb{N}$ genau n Nullstellen, dann gilt die *Faktorzerlegung*

$$p_n(x) = (x - x_1)(x - x_2) \cdots (x - x_n)a_n.$$

- direkte Folgerung:

Satz 3.90: Ein Polynom vom Grad $n \in \mathbb{N}$ hat höchstens n Nullstellen.

Satz 3.91: Hat ein Polynom vom Grad $\leq n$ mehr als n Nullstellen, dann ist es das Nullpolynom, d. h. es gilt $a_0 = a_1 = \ldots = a_n = 0$.

Bemerkung 3.92:

- obiger Satz besagt nicht, dass ein reelles Polynom vom Grad n genau n reelle Nullstellen aufweist
- Beispiel: $p(x) = x^2 + 1$ besitzt keine reelle Nullstelle
- aber: Aussage gilt für $\mathbb{K} = \mathbb{C}$

Satz 3.93 (Fundamentalsatz der Algebra): Jedes nichtkonstante Polynom besitzt mindestens eine komplexe Nullstelle.

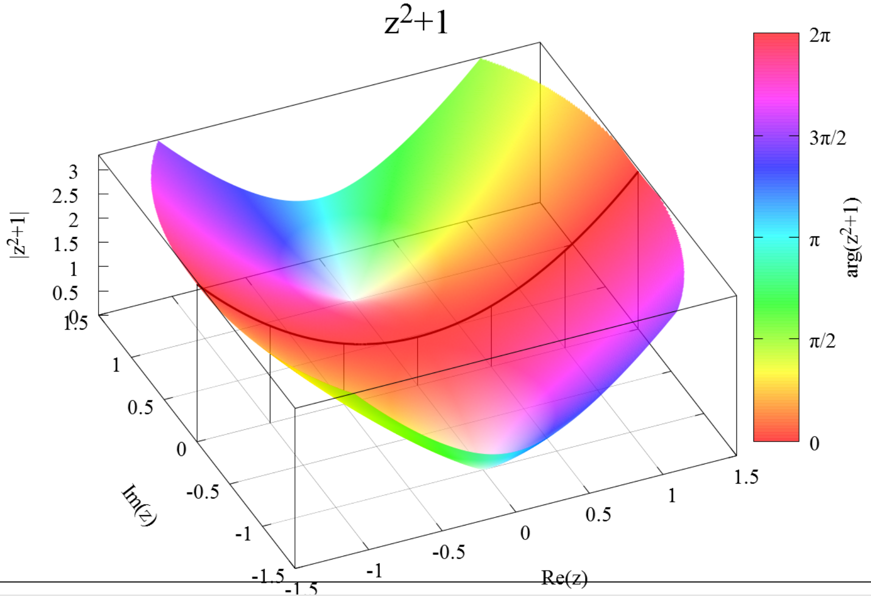
- erster vollständiger Beweis durch Gauß 1799 in seiner Dissertation
- Beweis sehr aufwändig; daher wird er ausgelassen

Folgerung 3.94: Jedes Polynom vom Grad $n \in \mathbb{N}$ weist genau n komplexe Nullstellen auf.

Beweis *OHP*

Beispiel 3.95:

- $p(z) = z^2 + 1$ besitzt die komplexen Nullstellen i und $-i$
- Visualisierung: Höhe der Fläche: Absolutbetrag von $p(z)$
- Färbung: Argument von $p(z)$
- konkret: Argument 0, π und 2π: Zahl reell
- Argument $\pi/2$ oder $3/2\pi$: Zahl imaginär



z²+1
|z²+1|
3
2.5
2
1.5
1
0.5
0
arg(z²+1)
2π
3π/2
π
π/2
0
Im(z)
1.5
1
0.5
0
-0.5
-1
-1.5
Re(z)
-1.5
-1
-0.5
0
0.5
1
1.5

Bemerkung 3.96:

- Existenz von Nullstellen bedeutet nicht unbedingt Berechenbarkeit
- Nullstellen von Polynomen vom Grad 2: *pq*-Formel
- Nullstellen von Polynomen vom Grad 3 und 4: Cardanische Formeln
- ziemlich kompliziert
- gefunden von Cardano (ca. 1500)
- Nullstellen von Polynomen vom Grad 5 und höher?
- Abel 1824: es gibt kein Analogon zur *pq*-Formel
- in der Praxis: Näherungslösungen
- jetzt: Polynome als Elemente von Vektorräumen

- Addition erfolgt punktweise
- skalare Vielfache von Polynomen werden ebenfalls punktweise erklärt

Satz 3.97: Seien $p, q \in P_n$. Dann sind für $\lambda \in \mathbb{K}$ sowohl λp als auch $(p+q)$ Polynome, und es gilt

$$deg(\lambda p) \leq n, \qquad deg(p+q) \leq n$$

Beweis *OHP*

Folgerung 3.98: Sei $+$ die Addition von Funktionen und $\cdot$ die Multiplikation einer Funktion mit einem Skalar. Dann bildet $(P_n, +, \cdot)$ einen $\mathbb{K}$-Vektorraum.

Beweis *OHP*

- jetzt: Basis und Dimension von P_n

Satz 3.99: Die Funktionen $1, x, x^2, \ldots, x^n$ bilden eine Basis des Vektorraums P_n, und es gilt

$$\dim(P_n) = n + 1\,.$$

Beweis *OHP*

Bemerkung 3.100:

1. Die Basis $(1, x, ...)$ aus Satz 3.99 wird *Monombasis* genannt.
2.
   - Koeffizientenvektor $(a_0, ..., a_n)$: Koordinaten von $p_n(x) = \sum_{k=0}^{n} a_k x^k$ bezüglich der Monombasis
   - Eindeutigkeit der Koordinatendarstellung: Koeffizienten eines Polynoms liegen eindeutig fest
3.
   - Es mögen $p, q \in P_n$ in mindestens $(n+1)$ Stellen übereinstimmen, also

$$p_n(x) = \sum_{k=0}^{n} a_k x^k \text{ und } q_n(x) = \sum_{k=0}^{n} b_k x^k$$

   - dann folgt:

$$a_k = b_k, k = 0, 1, ..., n.$$

Beweis *OHP*

- erste Anwendung von Polynomen
- gegeben seien Zeitpunkte $t_1 < \ldots < t_n$, zu denen Messungen durchgeführt wurden
- Anforderung: nachträglich einen Messwert zu einem anderen Messzeitpunkt generieren
- Ansatz: Finde f, die zu den Messzeitpunkten die gemessenen Werte annimmt
- dann: Werte f im Zwischenzeitpunkt aus
- der Wert von f ist dann (hoffentlich) ein geeigneter Ersatz für den realen Messwert, den es nicht gibt
- Ansatz: f sei Polynom
- Gibt es ein solches f immer? Ist es eindeutig?

Satz 3.101: Gegeben seien die $n+1$ Punkte $(x_k, y_k), 0 \leq k \leq n$ mit *paarweise verschiedenen* x_k. Dann existiert genau ein $p_n \in P_n$ mit $y_k = p_n(x_k)\ \forall 0 \leq k \leq n$. Dies ist das sogenannte *Interpolationspolynom*.

Beweis *OHP*

Beispiel 3.102: *OHP*

Bemerkung 3.103:

- oft: Polynome vom Grad n werden auf Teilmengen $M \subset \mathbb{K}$ eingeschränkt
- $P_n(M)$
- enthält M mehr als $n+1$ Elemente, gelten alle Aussagen des Kapitels, weil dann nicht alle Elemente Nullstellen sein können
- Folge: Aussagen gelten für sämtliche offenen Mengen, alle echten reellen Intervalle, Kreisscheiben, ...
- abschließend: Verfahren, mit dem man die lineare Abhängigkeit von Polynomen einfach bestimmen kann
- $p_1(x), \dots, p_n(x)$ Polynome vom Grad $\leq m$
- lineare Unabhängigkeit hier:

$$\sum_{i=1}^{n} \lambda_i p_i(x) = 0 \tag{3.3}$$

- Gleichung darf keine andere Lösung als $\lambda_1 = \lambda_2 = \ldots = 0$ besitzen
- $p_i(x)$ lässt sich immer schreiben als

$$p_i(x) = \sum_{k=0}^{m} a_{ki} x^k$$

- Einsetzen und Umordnen der Summanden:

$$0 = \sum_{i=1}^{n} \lambda_i \left(\sum_{k=0}^{m} a_{ki} x^k \right) = \sum_{k=0}^{m} \left(\sum_{i=1}^{n} \lambda_i a_{ki} \right) x^k.$$

- lineare Unabhängigkeit der Monome:

$$\sum_{i=0}^{n} a_{ki}\lambda_i = 0 \quad \forall 0 \leq k \leq m$$

- LGS in $(\lambda_1, \ldots, \lambda_n)$
- Gibt es andere Lösungen als $\lambda_1 = \ldots = \lambda_n = 0$, sind die Polynome linear abhängig, sonst linear unabhängig

Beispiel 3.104: Die Polynome

$$p_1(x) = (1-x)^2; \quad p_2(x) = (1-x)x; \quad p_3(x) = x^2$$

sollen auf lineare Unabhängigkeit geprüft werden. *OHP*

3.6 Skalarprodukt, euklidische und unitäre Räume

- bisher: Skalarprodukt auf $\mathbb{R}^n$
- jetzt: $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$
- Definition eines Skalarprodukts für allgemeine $\mathbb{K}$-Vektorräume

Definition 3.105: Sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$ heißt *Skalarprodukt*, wenn folgende Bedingungen erfüllt sind:

SP1: $\forall a, b \in V$:

$$\langle a, b \rangle = \begin{cases} \langle b, a \rangle & \text{für } \mathbb{K} = \mathbb{R} \\ \overline{\langle b, a \rangle} & \text{für } \mathbb{K} = \mathbb{C}. \end{cases}$$

SP2 $\forall a, b, c \in V$:

$$\langle a, (b+c)\rangle = \langle a, b\rangle + \langle a, c\rangle$$
$$\langle (a+b), c\rangle = \langle a, c\rangle + \langle b, c\rangle$$

SP3: $\forall \alpha \in \mathbb{K}$ gilt

$$\langle \alpha a, b\rangle = \alpha\langle a, b\rangle = \begin{cases} \langle a, \alpha b\rangle & \text{für } \mathbb{K} = \mathbb{R} \\ \langle a, \overline{\alpha} b\rangle & \text{für } \mathbb{K} = \mathbb{C}. \end{cases}$$

SP4: (positive Definitheit) $\forall a \in V \setminus \{0\} : \langle a, a\rangle > 0$, und $\langle 0, 0\rangle = 0$.

Bemerkung 3.106:

- positive Definitheit **SP4** nur sinnvoll, wenn $\langle a, a\rangle$ immer reell ist, auch für komplexes a
- Grund: nur $\mathbb{R}$ besitzt die Ordnungsrelation „>“

- Ausweg: Modifizierung der Symmetriebedingung wie in **SP1**
- dann: $\forall a \in V : \langle a, a\rangle = \overline{\langle a, a\rangle}$, also $Im(\langle a, a\rangle) = 0$
- daher ist **SP4** auch im Fall $\mathbb{K} = \mathbb{C}$ wohldefiniert

Beispiel 3.107:

- Ein Skalarprodukt auf $\mathbb{R}^n$ nach Definition 3.105 ist auch Skalarprodukt im Sinn von Definition 2.4
- also: beide Definitionen konsistent
- zunächst: Standardskalarprodukt auf $\mathbb{C}^n$ (analog zum euklidischen Skalarprodukt)

Definition 3.108: Auf $\mathbb{C}^n$ sei für $a = (a_i)_{i=1}^n, b = (b_i)_{i=1}^n$ das Standardskalarprodukt definiert durch

$$\langle a, b\rangle := \sum_{i=1}^{n} a_i \overline{b_i} \, . \tag{3.4}$$

Beispiel 3.109: *OHP*

Wir definieren nun ein abstraktes Skalarprodukt auf dem Vektorraum der stetigen Funktionen, der sich ja wesentlich von $\mathbb{K}^n$ unterscheidet.

Beispiel 3.110: Seien $f, g \in \mathcal{C}[a,b]$. Auf $\mathcal{C}[a,b]$ wird durch

$$\langle f, g \rangle := \int_a^b f(x) \cdot g(x)\, dx \tag{3.5}$$

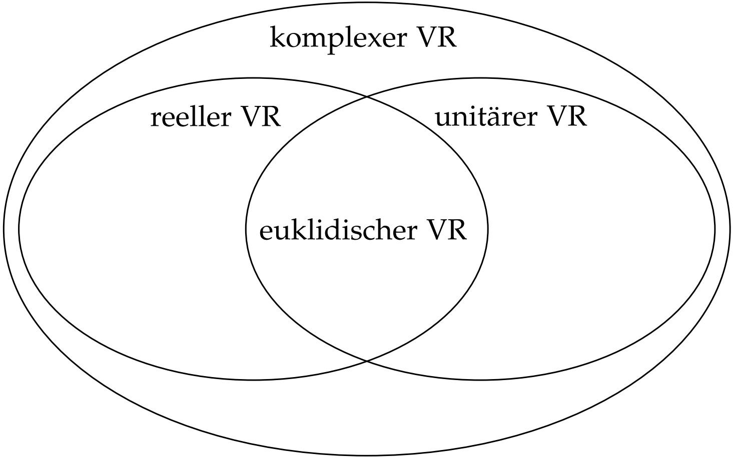
ein Skalarprodukt definiert.

Beweis *OHP*

Beispiel 3.111: *OHP*

Definition 3.112: Ein reeller Vektorraum gemeinsam mit einem Skalarprodukt heiße *Euklidischer Vektorraum*, ein komplexer Vektorraum mit einem Skalarprodukt heiße *Unitärer Vektorraum*.

- grafische Veranschaulichung:



Beispiel 3.113:

- $\mathbb{R}^n$ mit dem euklidischen Skalarprodukt und $\mathcal{C}[a,b]$ mit dem Integral-Skalarprodukt:

euklidische Vektorräume

- ebenso alle Unterräume wie z. B. $P_n([a, b])$
- zur Definition einer Norm auf allgemeinem V
- wir gehen von einer Norm auf $\mathbb{R}^n$ aus
- Idee: jede Abbildung $\|\cdot\| : V \to \mathbb{R}$ mit den Eigenschaften einer Norm auf $\mathbb{R}^n$ heißt Norm
- Konsistenz der Definitionen damit klar

Definition 3.114: Sei V ein $\mathbb{K}$-Vektorraum und $a, b \in V$. Eine Abbildung $\|\cdot\| : V \to \mathbb{R}$ heißt *Norm* genau dann, wenn

N0 : $\|a\| \in \mathbb{R}$.

N1 : $\|a\| \geq 0$.

N2 : $\|a\| = 0 \quad \Leftrightarrow \quad a = 0.$

N3 : $\forall \lambda \in \mathbb{K} : \|\lambda a\| = |\lambda| \|a\|$

N4 : (Dreiecksungleichung) $\|a + b\| \leq \|a\| + \|b\|$

- wie im Fall $V = \mathbb{R}^n$: Skalarprodukt induziert Norm

Satz 3.115: In einem unitären (bzw. euklidischen) Raum induziert das Skalarprodukt eine (Standard-)*Norm* durch

$$\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}\,.$$

Beweis *OHP*

In allen unitären Vektorräumen gilt die Cauchy-Schwarzsche Ungleichung, deren Beweis wir zunächst zurückstellen.

Satz 3.116: In allen unitären Vektorräumen V gilt die *Cauchy-Schwarzsche Ungleichung*

$$|\langle a,b\rangle| \le \|a\|\,\|b\| \quad \forall\, a,b \in V\,. \tag{3.6}$$

- Dreiecksungleichung **N4**: Cauchy-Schwarzsche Ungleichung und Symmetrie eines Skalarproduktes wurden verwendet
- gilt daher zunächst nur in euklidischen Räumen
- wir beweisen also die Dreiecksungleichung für unitäre Vektorräume

Beweis *OHP*

Beispiel 3.117:

1. Für $V = \mathbb{C}$ und $z = a + ib \in \mathbb{C}$ induziert das Standardskalarprodukt (3.4) die Norm

$$\|z\| = \sqrt{z \cdot \overline{z}} = \sqrt{(a+ib)(a-ib)} = \sqrt{a^2+b^2} = |z|,$$

   also den gewöhnlichen Betrag komplexer Zahlen.

2. Für $\mathbb{R}^n$ erhält man aus dem euklidischen Skalarprodukt die euklidische Norm (vgl. Kapitel 2.1).

3. Auf $V = \mathcal{C}[a,b]$ und $f \in V$ induziert das Skalarprodukt aus Beispiel 3.111 die sog. L_2-Norm

$$\|f\|_{L_2} := \left(\int_a^b f^2(x)\,dx \right)^{1/2} . \tag{3.7}$$

Sei z. B. $[a,b]=[-1,1]$ und $f(x)=x$. Dann ist

$$\begin{aligned}\|f\|_{L_2}^2=\|x\|_{L_2}^2 &= \int_{-1}^{1} x\cdot x\,dx\\ &= \frac{1}{3}x^3\Big|_{-1}^{1}=\frac{2}{3}.\end{aligned}$$

4. Nicht jede Norm auf einem Vektorraum wird durch ein Skalarprodukt induziert (vgl. Bemerkung 2.15). So wird die auf $\mathcal{C}[a,b]$ in der Analysis sehr gebräuchliche Maximumsnorm

$$\|f\|_\infty := \max_{x\in[a,b]} |f(x)|$$

von keinem Skalarprodukt induziert.

Bemerkung 3.118:

- Definition von Winkeln auf euklidischen Vektorräumen problemlos möglich, weil die Cauchy-Schwarzsche Ungleichung gilt
- aber: Bedeutung von Winkeln sehr gering
- viel problematischer: Winkel in unitären Räumen: Skalarprodukt kann komplexe Werte annehmen

3.7 Orthogonalität in unitären Vektorräumen

- in der analytischen Geometrie in $\mathbb{R}^n$: orthogonale Projektion fundamental
- daher: orthogonale Projektion wird auch in unitären Vektorräumen wichtig sein
- zunächst: genaue Definition der orthogonalen Projektion
- euklidische Vektorräume sind Spezialfälle von unitären Vektorräumen, daher nur unitäre Vektorräume
- ab jetzt: V sei ein unitärer Vektorraum

Definition 3.119: Seien $a, b \in V$. Es stehen a und b orthogonal zueinander, falls

$$\langle a, b \rangle = 0$$

gilt. Man schreibt $a \perp b$.

Bemerkung 3.120:

- Satz des Pythagoras $\|a+b\|^2 = \|a\|^2 + \|b\|^2$ für $a \perp b$ gilt in allen unitären Vektorräumen
- Beweis genau wie in Satz 2.25 für $\mathbb{R}^n$
- bei der Herleitung der orthogonalen Projektion in $\mathbb{R}^n$: nur **SP2** und **SP3** benutzt; Skalare wurden nur aus dem ersten Argument gezogen
- dies gilt in unitären Vektorräumen unverändert
- daher man erhält man

Satz 3.121: Für die orthogonale Projektion $p_b(a)$ eines Vektors a auf b, $b \neq 0$, gilt in jedem unitären Vektorraum

$$p_b(a) = \frac{\langle a, b\rangle}{\langle b, b\rangle} b \tag{3.8}$$

Wir holen nun den Beweis der Cauchy-Schwarzschen Ungleichung nach und schließen damit die Lücke in unserer Theorie.

Beweis *OHP*

- der Wert von $p_b(a)$ bleibt, wenn man b durch ein skalares Vielfaches ersetzt
- alle Vektoren in $L(b)$ führen zur gleichen orthogonalen Projektion
- deswegen: $p_b(a)$ als eine orthogonale Projektion von a auf den Untervektorraum $U = L(b)$
- man hat $a - p_b(a) \perp u \, \forall u \in U$

Definition 3.122: Sei U ein endlich erzeugter Untervektorraum von V und $a \in V$. Ein Vektor $p_U(a) \in U$ heißt *orthogonale Projektion* von a auf U, wenn

$$a - p_U(a) \perp u \quad \forall u \in U \tag{3.9}$$

gilt.

- Gibt es immer ein eindeutiges $p_U(a)$?

- folgender Begriff hilfreich

Definition 3.123: Für $M \subseteq V$ heißt

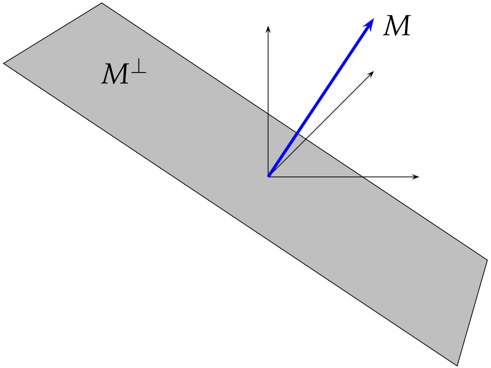
$$M^{\perp} = \{v \in V \,|\, v \perp u \;\forall u \in M\}$$

das *orthogonale Komplement* von M.

Bemerkung 3.124:

1. $M^{\perp}$ ist ein Untervektorraum von V.
2. Sei U ein Untervektorraum von V. Dann gilt $U \cap U^{\perp} = \{0\}$.

Beweis *OHP*



Beispiel 3.125: *OHP*

Folgerung 3.126: Seien U, V und a wie in Definition 3.122. Die orthogonale Projektion von a auf U ist eindeutig.

Beweis *OHP*

Lemma 3.127: Sei U wie zuvor und $(u_1, \ldots, u_m)$ eine Basis von U. Für $v \in V$ gilt:

$$v \in U^\perp \Leftrightarrow \langle v, u_i \rangle = 0 \quad \forall\, 1 \leq i \leq m$$

Beweis *OHP*

- sei $(u_1, \dots, u_m)$ eine Basis von U. Existiert $p_U(a) \in U$, dann gilt

$$p_U(a) = \sum_{i=1}^{m} \alpha_i u_i \tag{3.10}$$

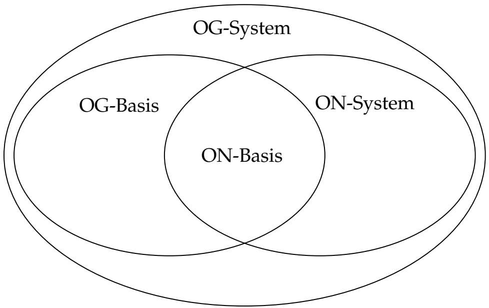
- α_i sind noch zu bestimmen
- mit obigen Lemma und Orthogonalitätsbedingung:

$$\begin{aligned} & \left\langle a - \sum_{i=1}^{m} \alpha_i u_i, u_j \right\rangle = 0 \quad \forall\, 1 \leq j \leq m \\ \Leftrightarrow \quad & \sum_{i=1}^{m} \langle u_i, u_j \rangle \alpha_i = \langle a, u_j \rangle \quad \forall\, 1 \leq j \leq m \\ \Leftrightarrow \quad & G\alpha = b \end{aligned} \tag{3.11}$$

- Koeffizientenvektor $\alpha = (\alpha_1, \ldots, \alpha_m)^T$
- Matrix $G = (\langle u_i, u_j \rangle)_{i,j=1}^m$ (Gramsche Matrix)
- LGS in den Koeffizienten
- die orthogonale Projektion existiert genau dann, wenn dieses LGS eine Lösung besitzt
- Frage nach der Lösbarkeit wird zunächst zurückgestellt
- die Matrix G hängt von der Wahl der Basis von U ab
- beste Wahl für die Basis: $\langle u_i, u_j \rangle = 0$ für $i \neq j$ und $\langle u_i, u_i \rangle = 1$ für $i = j$
- Basisvektoren paarweise orthogonal
- dann: G ist die Einheitsmatrix, und man müsste kein Gleichungssystem mehr lösen

Definition 3.128: Sei $\mathcal{B} = (v_1, \ldots, v_m)$ ein m-Tupel mit Vektoren in $V \setminus \{0\}$.

1. $\mathcal{B}$ heißt *Orthogonalsystem* in V, falls alle v_i paarweise orthogonal sind.
2. Ein Orthogonalsystem, für das zusätzlich $\|v_i\| = 1 \forall i = 1, \ldots, m$ gilt, heißt *Orthonormalsystem*.
3. Ein Orthogonalsystem, das eine Basis von V bildet, heißt *Orthogonalbasis* von V.
4. Ein Orthonormalsystem, das eine Basis von V bildet, heißt *Orthonormalbasis* von V.



Bemerkung 3.129: Mit dem *Kronecker-Symbol*[1]

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

gilt in jedem Orthonormalsystem $\langle v_i, v_j \rangle = \delta_{i,j}$.

Beispiel 3.130: *OHP*

[1] Leopold Kronecker (1823–1891), dt. Mathematiker; Beiträge vornehmlich zur Algebra und Zahlentheorie

Beispiel 3.131: *OHP*

Beispiel 3.132: *OHP*

Beispiel 3.133: Die Vektoren

$$a_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad a_2 = \begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \quad \text{und} \quad a_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

sind eine Basis, orthogonal und normiert. Sie bilden somit eine Orthonormalbasis.

Satz 3.134: Ein Orthogonalsystem $(v_1, \ldots, v_n)$ ist linear unabhängig.

Beweis *OHP*

Satz 3.135: Ist $\mathcal{B} = (v_1, \ldots, v_n)$ eine Orthogonalbasis von V, dann gilt für jedes $v \in V$:

$$v = \sum_{k=1}^{n} \frac{\langle v, v_k \rangle}{\langle v_k, v_k \rangle} v_k,$$

d. h. v hat bzgl. $\mathcal{B}$ die Koordinaten $\left(\frac{1}{\|v_k\|^2}\langle v, v_k\rangle, 1 \leq k \leq n\right)^T$.

Beweis *OHP*

Bemerkung 3.136:

- k-te Koordinate eines Vektors bzgl. einer Orthogonalbasis: Länge der Projektion auf den k-ten Basisvektor
- im Fall einer Orthonormalbasis: v besitzt die Koordinaten $(\langle v, v_k\rangle, 1 \leq k \leq n)^T$

Beispiel 3.137:

$$a_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}, \quad a_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix} \quad \text{und} \quad a_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

- ON-Basis des $\mathbb{R}^3$
- also $v = (5, 3, 7)^T = \lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3$

- Vorfaktoren λ_i aus:

$$\lambda_1 = \langle v, a_1 \rangle = \frac{5}{\sqrt{2}} + \frac{3}{\sqrt{2}} = 4 \cdot \sqrt{2}$$
$$\lambda_2 = \langle v, a_2 \rangle = \frac{5}{\sqrt{2}} - \frac{3}{\sqrt{2}} = \sqrt{2}$$
$$\lambda_3 = \langle v, a_3 \rangle = 7$$

- also $v = 4\sqrt{2} \cdot a_1 + \sqrt{2} \cdot a_2 + 7 \cdot a_3$

Satz 3.138: Sei $\mathcal{B} = (v_1, \ldots, v_m)$ ein Orthogonalsystem in V und $U = L(\mathcal{B})$ der von $\mathcal{B}$ aufgespannte Unterraum.

1. Für jedes $v \in V$ gilt

$$p_U(v) = \sum_{i=1}^{m} \frac{\langle v, v_i \rangle}{\langle v_i, v_i \rangle} v_i.$$

2. Jedes $v \in V$ lässt sich eindeutig als Summe $v = p_U(v) + w$ mit $w \in U^\perp$ schreiben. Dabei gilt $w = v - p_U(v)$.

3. $V = U \oplus U^\perp$.

4. Sei $\dim(V) = n$. Dann gilt $\dim(U) + \dim(U^\perp) = n$ für jeden Untervektorraum U.

Beweis *OHP*

Beispiel 3.139:

$$\mathcal{B} = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right)$$

- $\mathcal{B}$ Orthonormalsystem
- es gilt

$$U = L(\mathcal{B}) = \left\{ \begin{pmatrix} a \\ b \\ 0 \end{pmatrix}; \quad a, b \in \mathbb{R} \right\}.$$

- orthogonales Komplement:

$$U^\perp = \left\{ \begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix}; \quad c \in \mathbb{R} \right\}.$$

3.8 Das Verfahren von Gram-Schmidt und Anwendungen

- wir haben gesehen: orthogonale Projektion leicht möglich, wenn eine OG- oder sogar ON-Basis vorliegt
- Besitzt jeder endlich erzeugte unitäre Vektorraum eine ON-Basis?
- Wenn ja, wie kann man die konstruieren?
- Antwort: Orthogonalisierungsverfahren von Gram-Schmidt
- unabhängige Veröffentlichung durch Erhard Schmidt (1907) und Jørgen Pederson Gram (1879)
- Verwendung bereits um 1836 durch Cauchy
- gegeben: m linear unabhängige Vektoren aus VR mit Dimension n
- $m \leq n$

- Ziel: erzeuge einen Satz orthonormaler Vektoren $w_1, \ldots, w_m$ mit

$$L(v_1, \ldots, v_m) = L(w_1, \ldots, w_m),$$

- zur Orthogonalisierung: verwende orthogonale Projektion
- zur Veranschaulichung: Betrachte (v_1, v_2, v_3) beliebige Basis von $\mathbb{R}^3$
- Vorgehen:

1.
   - setze $w_1 = \frac{1}{\|v_1\|} v_1$
   - dann: $\|w_1\| = 1$
2.
   - konstruiere Vektor r_2 senkrecht zu w_1
   - dazu: projiziere v_2 auf den von w_1 erzeugten Unterraum $L(w_1)$
   - setze $r_2 := v_2 - p_{L(v_1)}(v_2)$

- $v_1 \perp r_2$ nach Definition der orthogonalen Projektion
- konkrete Berechnung nach Satz 3.138:

$$r_2 = v_2 - \langle v_2, w_1 \rangle w_1 \,.$$

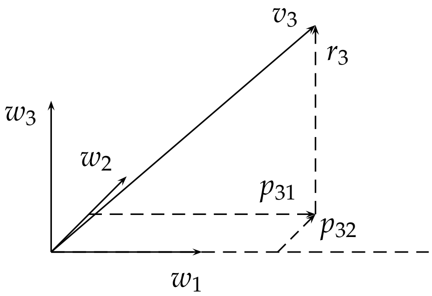
- man hat $r_2 \neq 0$, da $v_2 \notin L(w_1)$
- Normierung von r_2 liefert w_2
- r_2 und damit w_2 ist Linearkombination von v_1 und v_2
- also $L(w_1, w_2) \subseteq L(v_1, v_2)$
- weiter: w_1 und w_2 orthonormal, also linear unabhängig
- es folgt $\dim(L(w_1, w_2)) = \dim(L(v_1, v_2))$
- also:

$$L(w_1, w_2) = L(v_1, v_2) \,.$$

3. - setze $r_3 := v_3 - p_{L(v_1,v_2)}(v_3)$
   - nach Konstruktion: senkrecht auf $L(v_1, v_2)$
   - es folgt $r_3 \perp v_1$ und $r_3 \perp v_2$
   - v_i linear unabhängig: v_3 nicht in $L(v_1, v_2) = L(w_1, w_2)$
   - also $r_3 \neq 0$
   - w_3 durch Normierung von r_3
   - bekannt: (w_1, w_2) ON-Basis von $L(v_1, v_2)$
   - deswegen einfache Berechnung durch

$$r_3 = v_3 - \langle v_3, w_1 \rangle w_1 - \langle v_3, w_2 \rangle w_2$$

Das folgende Bild veranschaulicht die Konstruktion von w_3.



- einfache Verallgemeinerung möglich

Satz 3.140: (Gram-Schmidtsches Orthogonalisierungsverfahren) Sei V ein unitärer Vektorraum und $v_1, \ldots, v_m$ linear unabhängig. Seien

$$w_1 := \frac{v_1}{\|v_1\|}$$

$$r_{k+1} := v_{k+1} - \sum_{i=1}^{k} \langle v_{k+1}, w_i \rangle w_i$$

$$w_{k+1} = \frac{r_{k+1}}{\|r_{k+1}\|}$$

Dann bilden $(w_1, \ldots, w_m)$ eine Orthonormalbasis von $L(v_1, \ldots, v_m)$.

Beweis

- Beweis per vollständiger Induktion
- Induktionsanfang: Schritt 2.
- Induktionsschluss: wie bei Schritt 3 (nur allgemeinere Indizes) ■

Beispiel 3.141: *OHP*

Folgerung 3.142: Jeder endlich erzeugte unitäre Vektorraum besitzt eine Orthonormalbasis.

Bemerkung 3.143: Folgerung 3.142 ist in nicht endlich erzeugten Vektorräumen i. A. falsch.

- jetzt: Existenzbeweis der orthogonalen Projektion

Folgerung 3.144: Sei V ein unitärer Vektorraum und U ein endlich erzeugter Untervektorraum. Dann existiert für jedes $v \in V$ die orthogonale Projektion $p_U(v)$ von v auf U.

Beweis

- Sei $U = L(v_1, \dots, v_m)$
- finde ON-Basis $(w_1, \dots, w_m)$ von U (ex. nach Gram-Schmidt)

- wende Satz 3.138 an ■

Folgerung 3.145: Sei V ein endlich erzeugter unitärer Vektorraum und U irgend ein Untervektorraum. Dann gilt $V = U \oplus U^\perp$, und $\dim(V) = \dim(U) + \dim(U^\perp)$.

Beweis

- Aussage schon bewiesen, wenn U durch ein Orthogonalsystem erzeugt wird (Satz 3.138)
- Gram-Schmidt-Verfahren liefert für jedes U ein ON-System
- also ist die Voraussetzung unnötig ■

- wir zeigen jetzt, dass jede Hyperebene eine Normaldarstellung besitzt (Beweis stand noch aus)

Folgerung 3.146: Jede Hyperebene in $\mathbb{R}^n$ besitzt eine Normaldarstellung; der Normalenvektor ist bis auf Skalierung eindeutig.

Beweis *OHP*

Bemerkung 3.147:

- Ergebnis des Gram-Schmidt-Verfahrens hängt von der Reihenfolge der $v_1, \dots, v_m$ ab
- vertauscht man die Reihenfolge, erhält man i. A. *nicht* die $w_1, \dots, w_m$ in vertauschter Reihenfolge
- aber: Satz 3.140 unverändert gültig

Satz 3.148: Sei V wie oben und $v_1, \dots, v_m \in V$. Gelingt es, aus diesen mithilfe des Gram-Schmidt-Verfahrens orthonormale Vektoren $w_1, \dots, w_m$ zu erzeugen, dann sind $(v_1, \dots, v_m)$ linear unabhängig.

Beweis

- nehmen wir an, $(v_1, \dots, v_m)$ seien linear abhängig
- dann ex. v_k mit $v_k \in L(v_1, \dots, v_{k-1})$

- dann: $v_k = p_{L(v_1,\ldots,v_{k-1})}(v_k)$, also $r_k = 0$
- r_k nicht normierbar, das Gram-Schmidt-Verfahren bricht ab ■

Es ist also nicht erforderlich, die lineare Unabhängigkeit der Ausgangsvektoren bei der Orthonormalisierung nach Gram-Schmidt zu prüfen, weil das im Laufe der Rechnung ohnehin klar wird.

Beispiel 3.149: *OHP*

- orthonormale Funktionensysteme sehr wichtig in der angewandten Mathematik
- in der Praxis häufig: man möchte komplizierte Funktionen durch einfache Funktionen ersetzen
- dabei: Fehler durch Ersetzung möglichst klein
- „Approximation“
- Approximationstheorie: Konstruktion approximierender Funktionen und Fehlerschranken
- naheliegender Ansatz: Polynome
- Testfall: Approximation auf $[-5, 5]$ durch Polynome vom Höchstgrad n von

$$f(x) = \frac{1}{1+x^2}$$

- drei Ansätze

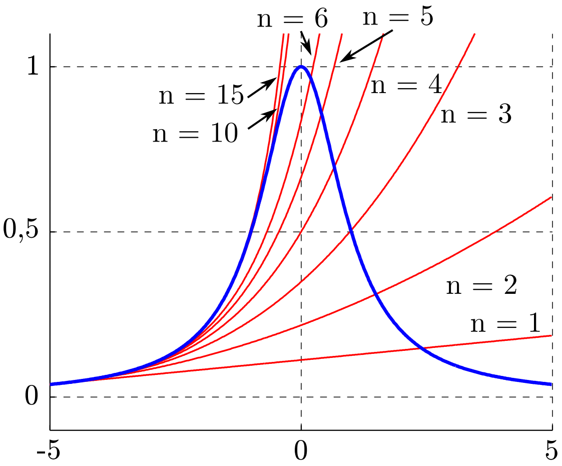
1. Approximation durch Taylorpolynome mit Grad n (Polynome t_n)
2. Approximation durch Interpolationspolynome mit Grad n (Polynome i_n)
3. Approximation durch orthogonale Projektion auf P_n (Polynome p_n)

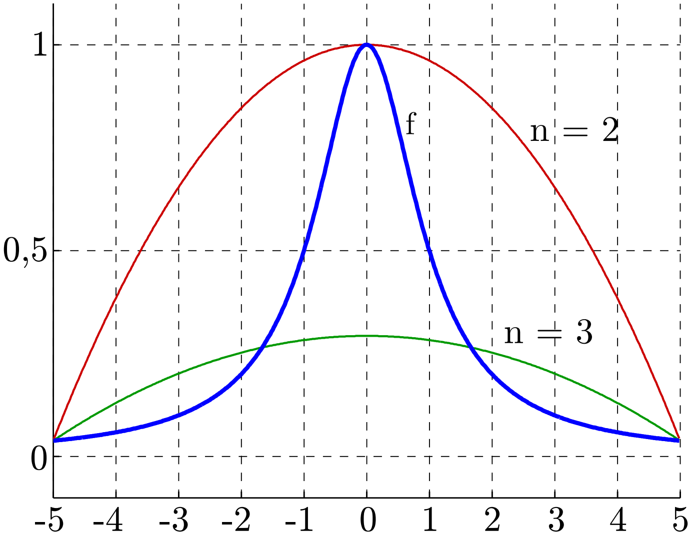
- Berechnungen in MATLAB
- $n+1$ Interpolationspunkte werden gleichabständig verteilt
- Entwicklungspunkt für Taylorpolynome: $x_0 = -5$.

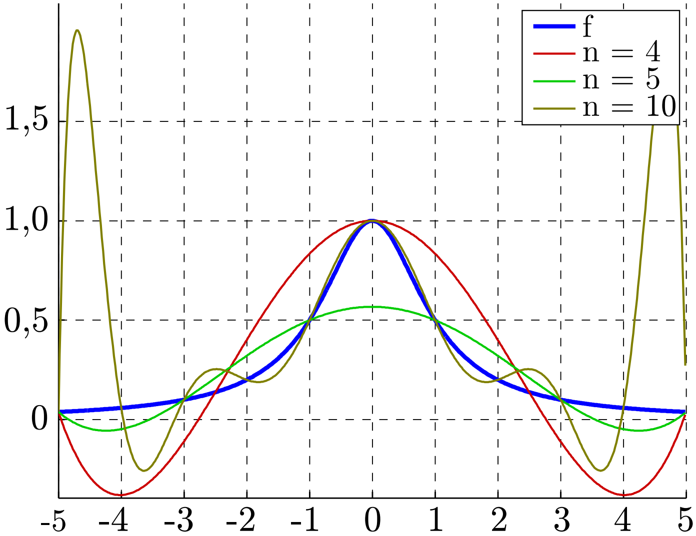
n	$\|f - t_n\|_{L_2}$	$\|f - i_n\|_{L_2}$	$\|f - p_n\|_{L_2}$
1	1,045	1,170	0,9007
2	1,108	1,438	0,9007
3	2,037	0,7712	0,6092
4	4,483	0,8829	0,6092
5	9,463	0,462	0,4103
6	19,02	0,8663	0,4103
10	218,4	1,835	0,1857
15	1074	-	0,0564

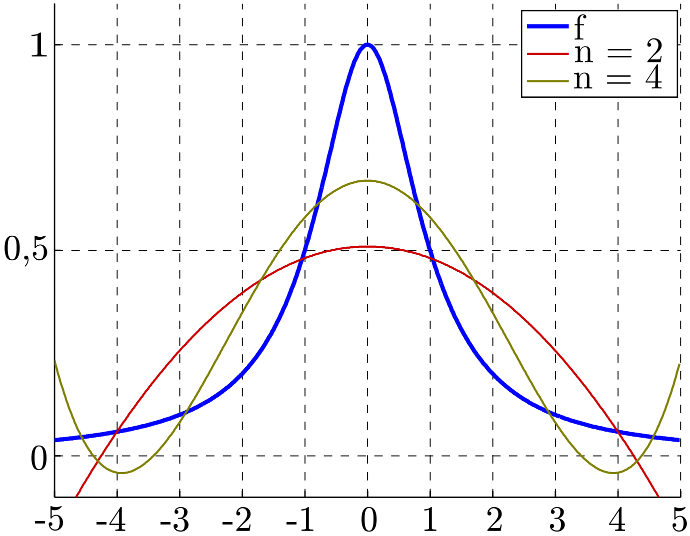
- Vergleich der drei Ansätze:
- allein die orthogonale Projektion von f auf P_n ist brauchbar
- man hat $\|f - p_n\|_{L_2} \to 0, n \to \infty$ (ohne Beweis)
- dann: Approximationsfehler kann (im Prinzip) beliebig klein gemacht werden

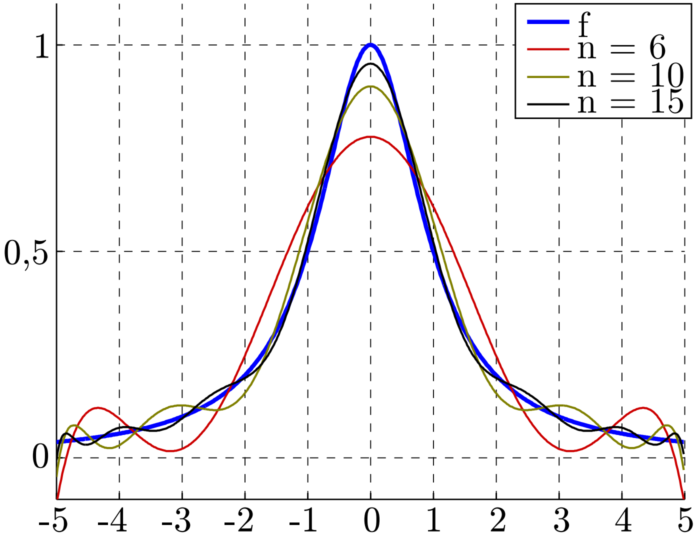
- bei den anderen Ansätzen: dies gilt nicht
- Vergleich der Polynome:




1
0,5
0
f
n = 2
n = 3
-5 -4 -3 -2 -1 0 1 2 3 4 5


f
n = 4
n = 5
n = 10
1,5
1,0
0,5
0
-5 -4 -3 -2 -1 0 1 2 3 4 5


f
n = 2
n = 4
1
0,5
0
-5
-4
-3
-2
-1
0
1
2
3
4
5


f
n = 6
n = 10
n = 15
1
0,5
0
-5 -4 -3 -2 -1 0 1 2 3 4 5

- Taylorpolynome: hervorragende Approximation nahe x_0, aber nicht auf dem ganzen Intervall
- Interpolationspolynome: starke Oszillation
- lässt sich aber durch bessere Wahl von Stützstellen reduzieren
- Warum ist die orthogonale Projektion so überlegen?

Definition 3.150: Sei V ein unitärer Vektorraum und $v \in V$ sowie $M \subset V$ eine beliebige nichtleere Menge. Ein Vektor $v^* \in M$ heißt *Bestapproximation* in M an v, falls

$$\|v^* - v\| = \inf_{x \in M} \|x - v\| \,.$$

- bei allgemeinen Mengen: es muss keine Bestapproximation geben
- Bestapproximation muss nicht eindeutig sein

- aber: bei Vektorräumen schon

Satz 3.151 (Bestapproximation): Sei V wie oben, $v \in V$ und U ein endlich erzeugter Untervektorraum von V. Dann gilt

$$\|v - p_U(v)\| = \min_{u \in U} \|u - v\|,$$

die orthogonale Projektion von v auf U ist also die einzige Bestapproximation an v in U.

Beweis *OHP*

- Satz 3.151: Approximationsfehler gemessen in der L_2-Norm der p_n an f bestmöglich
- p_n sind also die bestmögliche Wahl

Folgerung 3.152:

- es gilt $P_n \subset P_{n+1}$
- also $\|f - p_{P_{n+1}}(f)\|_{L_2} \leq \|f - p_{P_n}(f)\|_{L_2}$
- durch Erhöhung des Polynomgrades wird der Fehler keinesfalls größer
- aber: eventuell auch nicht kleiner (siehe Tabelle)

Bemerkung 3.153:

- Bestapproximation bzgl. anderer Normen auch denkbar
- Beispiel: Maximumsnorm auf $\mathcal{C}[a, b]$

- aber: diese wird durch kein Skalarprodukt induziert
- Theorie oben nicht anwendbar
- Theorie und Berechnungsverfahren existieren, sind aber viel komplizierter

Kapitel 4

Lineare Abbildungen

4.1 Vorbereitung

- „Menge“ und „Abbildung“ sind fundamentale Begriffe
- kurze Wiederholung als Vorbereitung
- Abbildung (intuitiv): eindeutige Zuordnung von Elementen einer Menge zu Elementen einer anderen Menge

Definition 4.1:

1. Unter einer *Abbildung* $f : X \to Y$ der *Definitionsmenge* X auf die *Zielmenge* Y versteht man eine Vorschrift, die jedem Element aus X ein eindeutiges Element aus Y zuordnet.

2. Für $x \in X$ heißt $f(x) \in Y$ der *Wert von f in x* oder auch *das Bild von x bezüglich f*.

- statt „Definitionsmenge“ auch „Definitionsbereich“

- statt „Zielmenge“ auch „Wertebereich“
- jetzt: Abbildungen ganzer Teilmengen

Definition 4.2: Seien f, X, Y wie in Definition 4.1 und $\tilde{X} \subseteq X$.

1. Das *Bild* $f(\tilde{X})$ *von* $\tilde{X}$ *bezüglich* f wird erklärt durch
$$f(\tilde{X}) := \{f(x) \,|\, x \in \tilde{X}\} \subseteq Y\,.$$
2. Für $\tilde{Y} \subseteq Y$ sei
$$f^{-1}(\tilde{Y}) := \{x \in X \,|\, f(x) \in \tilde{Y}\}$$
das *Urbild* von $\tilde{Y}$ bezüglich f.
3. Im Spezialfall einer einelementigen Menge $\tilde{Y} = \{a\}$ sei
$$f^{-1}(a) := f^{-1}(\{a\})$$

4. Man definiert $\text{Bild}(f) := f(X)$.

Bemerkung 4.3:

- $f^{-1}(a)$: *Menge* !!
- $f : \mathbb{R} \to \mathbb{R}, f(x) = x^2$
- $f^{-1}(-1) = \emptyset$
- $f^{-1}(4) = \{-2, 2\}$

- folgende Eigenschaften sind fundamental in Bezug auf Abbildungen

Definition 4.4: Seien f, X, Y wie in Definition 4.1. Die Abbildung f heißt

1. *injektiv*, wenn

$$f(x) = f(\tilde{x}) \Rightarrow x = \tilde{x} \quad \forall x, \tilde{x} \in X$$

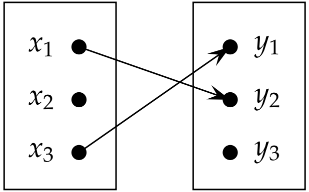
   gilt.

2. *surjektiv*, falls $\mathrm{Bild}(f) = Y$ gilt.

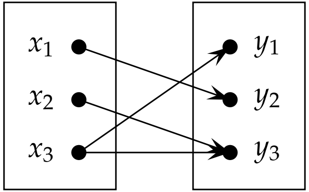
3. *bijektiv*, falls f injektiv und surjektiv ist.

4. *invertierbar*, falls es eine *Umkehrabbildung* oder *inverse Abbildung* $g : Y \to X$ gibt mit $f(g(y)) = y \, \forall y \in Y$ und $g(f(x)) = x \, \forall x \in X$.

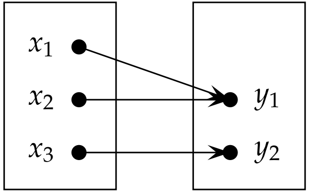
In den folgenden Diagrammen sind verschiedene Abbildungen zwischen Mengen skizziert, an denen die Begriffe aus Definition 4.4 illustriert werden.

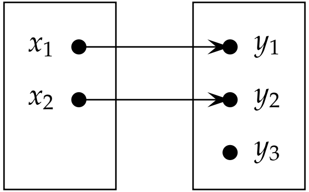


Keine Abbildung, kein Element ist $x_2 \in X$ zugeordnet.



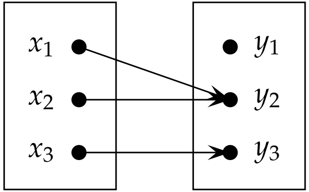
Keine Abbildung, zwei Elemente sind $x_3 \in X$ zugeordnet.

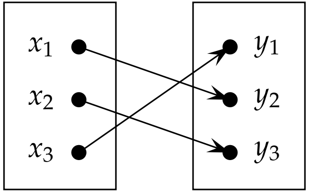




Surjektiv, aber nicht injektiv.

Injektiv, aber nicht surjektiv.





Weder injektiv noch surjektiv.

Injektiv + surjektiv = bijektiv.

Bemerkung 4.5: Eine Abbildung ist genau dann invertierbar, wenn sie bijektiv ist.

Beweis *OHP*

Bemerkung 4.6: Die zu f inverse Abbildung ist eindeutig.

Beweis *OHP*

- Bezeichnung der eindeutigen Umkehrabbildung von f: f^{-1}
- eigentlich Missbrauch der Notation, weil f^{-1} eine Menge beschreibt
- aber: Ist f bijektiv, dann identifiziert man $f^{-1}(y)$ mit ihrem einzigen Element
- die Zuordnung $y \to f^{-1}(y)$ erklärt dann die Umkehrabbildung

Beispiel 4.7: *OHP*

4.2 Grundlegende Eigenschaften linearer Abbildungen

- jetzt: Betrachtung von Vektorräumen (zentral in LA)
- *lineare* Abbildungen zwischen Vektorräumen
- zunächst: elementare Eigenschaften
- ab jetzt: V und W zwei K-Vektorräume über demselben Körper K

Definition 4.8:

1. Eine Abbildung $f : V \to W$ heißt *linear* oder ein *Homomorphismus*, falls gilt:

$$\begin{array}{cr} f(x+y) = f(x) + f(y) & \text{(Additivität)} \\ f(\lambda x) = \lambda f(x) & \text{(Homogenität)} \end{array} \quad \forall\, x, y \in V, \forall\, \lambda \in K\,.$$

2. Es sei $Hom(V, W)$ die Menge aller Homomorphismen von V nach W.

Beispiel 4.9: *OHP*

Beispiel 4.10:

- $f : \mathbb{R} \to \mathbb{R}, f(x) = x^2$ nicht linear
- denn: z. B. gilt $f(1+1) = 4 \neq 2 = f(1) + f(1)$

- Warum untersucht man lineare Abbildungen?
- Versuch einer Motivation

Bemerkung 4.11:

1. gegeben sei eine Funktion $f \in \mathcal{C}^2(\mathbb{R})$
   - Satz von Taylor
   - *OHP*

- jede glatte Funktion ist Summe einer Konstanten, eines linearen Anteils und eines Restglieds (klein in der Nähe des Entwicklungspunktes)
- f lokal beliebig gut durch eine lineare Funktion approximierbar
- analoge Resultate im $\mathbb{R}^n$

2. Wieviele wesentlich verschiedene Vektorräume gibt es eigentlich?

- analoge Situation bei Gruppen schon dagewesen
- Lösung: Isomorphiebegriff
- Isomorphismus: strukturerhaltende bijektive Abbildung
- will man die algebraische Struktur von Vektorräumen erhalten, kommt man zu linearen Abbildungen:
- Additivität wegen $(V, \oplus)$ Gruppe
- Homogenität aus der Multiplikation eines Vektors mit einem Skalar

Bemerkung 4.12: Seien $f, g : V \to W$ linear.

1. Für eine lineare Funktion f gilt $f(0) = 0$.
2. Die Funktion f ist genau dann linear, wenn

$$f(x + \lambda y) = f(x) + \lambda f(y) \; \forall x, y \in V, \lambda \in K \tag{4.1}$$

   gilt.

3. Summen, Vielfache linearer Abbildungen und vektorwertige Abbildungen, deren Komponenten aus linearen Abbildungen bestehen, sind wiederum linear.

Beweis *OHP*

Beispiel 4.13:

- $f(x) = x + 3$ ist *nicht* linear
- denn: dann müsste $f(0) = 0$ gelten
- aber: $f(0) = 3$
- allgemeiner: Funktionen mit konstantem Anteil *niemals linear*

Folgerung 4.14: $(Hom(V, W), +, \cdot)$ ist ein K-Vektorraum.

Beweis

- $Abb(V, W)$ ist Vektorraum
- Nachweis: Nachrechnen der Vektorraumaxiome durch Zurückführen auf die von V und W

- zeige also: $Hom(V,W)$ ist Untervektorraum
- konkret: Abgeschlossenheit von $Hom(V,W)$ gegenüber der Addition und der Multiplikation mit einem Skalar
- folgt aus direkt aus Bemerkung 4.12.3. ■

- Der Begriff der linearen Abbildung ist sehr umfassend
- zur Erläuterung: folgendes Beispiel

Beispiel 4.15:

1.
   - V euklidisch, $\langle \cdot , \cdot \rangle$ das Skalarprodukt und $a \in V$
   - $f_a : V \to \mathbb{R}, f_a(x) := \langle a, x \rangle$ ist linear

- Eigenschaften des Skalarprodukts:

$$f_a(x+\lambda y) = \langle a, x+\lambda y\rangle = \langle a, x\rangle + \langle a, \lambda y\rangle = f_a(x) + \lambda f_a(y)\,.$$

2. - $D : \mathcal{C}^1(\mathbb{R}) \to \mathcal{C}(\mathbb{R}), D(f) := f'$
   - Summen- und Faktorregel der Differentiation:

$$D(f+\lambda g) = (f+\lambda g)' = f' + (\lambda g)' = f' + \lambda g' = D(f) + \lambda D(g)$$

   - „Ableiten an sich“: lineare Abbildung

3. - T: Transpositionsabbildung

   $T : \mathbb{R}^{n\times m} \to \mathbb{R}^{m\times n}, T(A) := A^T.$

   - $A, B \in \mathbb{R}^{n\times m}$

$$\begin{aligned}(A+B)^T &= (a_{ij}+b_{ij})^T_{i=1\ldots n,j=1\ldots m}\\ &= (a_{ji}+b_{ji})_{i=1\ldots n,j=1\ldots m}\\ &= (a_{ji})_{i=1\ldots n,j=1\ldots m} + (b_{ji})_{i=1\ldots n,j=1\ldots m} = A^T + B^T\,.\end{aligned}$$

- $(\lambda A)^T = (\lambda a_{ji})_{i=1\ldots n,j=1\ldots m} = \lambda A^T$
- Folgerung:

$$T(A+\lambda B) = (A+\lambda B)^T = A^T + \lambda B^T = T(A) + \lambda T(B)\,.$$

- Fazit: Transponieren einer Matrix ist lineare Abbildung

- bisher gesehen: Injektivität, Surjektivität und Bijektivität herausragend
- jetzt: Untersuchung linearer Abbildungen auf diese Eigenschaften
- Hilfsmittel: *Kern*

Definition 4.16: Der *Kern* einer linearen Abbildung $f : V \to W$ wird definiert durch

$$\ker(f) := f^{-1}(0)\,.$$

Beispiel 4.17:

1. $f(x) = \alpha x$
   - reellwertige Funktion
   - $\ker(f) = \{0\}$ für $\alpha \neq 0$
   - $\ker(f) = \mathbb{R}$ für $\alpha = 0$
2. $f_a : V \to \mathbb{R}$
   $$\ker(f_a) = \{x \in V \,|\, f_a(x) = 0\} = \{x \in V \,|\, \langle a, x\rangle = 0\} = \{a\}^\perp$$
   - orthgonales Komplement von a
   - bekannt: es handelt sich um einen Untervektorraum

Satz 4.18: Sei $f : V \to W$ linear. Dann gilt:

1. $\mathrm{Bild}(f)$ ist ein Untervektorraum von W.
2. $\ker(f)$ ist ein Untervektorraum von V.

Beweis *OHP*

- einfache Charakterisierung von Injektivität möglich

Satz 4.19: Eine lineare Abbildung f ist genau dann injektiv, wenn $\ker(f) = \{0\}$ gilt.

Beweis *OHP*

Beispiel 4.20:

- nach Aussagen oben: $f(x) = \alpha x$ für $\alpha \neq 0$ ist injektiv

- ab jetzt: V und W endlich erzeugt
- Bild und Kern sind Unterräume mit endlicher Dimension
- Zusammenhang zwischen den Dimensionen der vorkommenden Räume?
- folgende Definition vereinfacht die Formulierungen

Definition 4.21: Für $f : V \to W$ linear definiert man den *Rang* von f durch

$$\mathrm{rg}(f) := \dim(\mathrm{Bild}(f)) .$$

Beispiel 4.22: *OHP*

Satz 4.23 (Dimensionsformel für lineare Abbildungen): Es sei $f : V \to W$ linear und $\dim(V) = n$. Dann gilt

$$\dim(\ker(f)) + \operatorname{rg}(f) = n.$$

Beweis *OHP*

OHP

- erste Anwendung: lineare bijektive Abbildungen

Definition 4.24: Sei $f : V \to W$ linear.

1. Ist f bijektiv, dann heißt f *Isomorphismus*.
2. Ist f bijektiv und gilt $V = W$, dann heißt f *Automorphismus*.

Bemerkung 4.25: Eine lineare Abbildung $f : V \to W$ ist genau dann ein Isomorphismus, wenn $\ker(f) = \{0\}$ und $\text{Bild}(f) = W$ gilt.

Folgerung 4.26:

1. Gelte $\dim(V) = \dim(W) = n$ und sei $f : V \to W$ linear. Dann gilt: f ist injektiv $\Leftrightarrow f$ ist surjektiv $\Leftrightarrow f$ ist bijektiv.

2. Sei $f : V \to W$ ein Isomorphismus. Dann gilt $\dim(V) = \dim(W)$.

Beweis *OHP*

Beispiel 4.27:

1. $f(x) = \alpha x$ ist Automorphismus für $\alpha \neq 0$

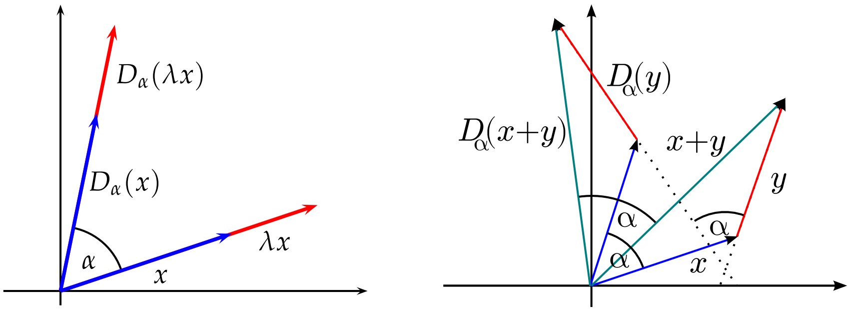
2.

   - zur Prüfung, ob f Isomorphismus: Dimensionen von V und W vergleichen!
   - wenn ungleich, dann kann f kein Isomorphismus sein
   - $\dim(V) < \dim(W) \Rightarrow f$ nicht surjektiv
   - $\dim(V) > \dim(W) \Rightarrow f$ nicht injektiv
   - Beispiel: für V mit $\dim(V) \geq 2$ ist f_α kein Isomorphismus (Aussage klar *ohne Rechnung!*)

Satz 4.28: Sei $f : V \to W$ ein Isomorphismus. Dann ist $f^{-1} : W \to V$ ebenfalls ein Isomorphismus.

Beweis *OHP*

Beispiel 4.29: *OHP*



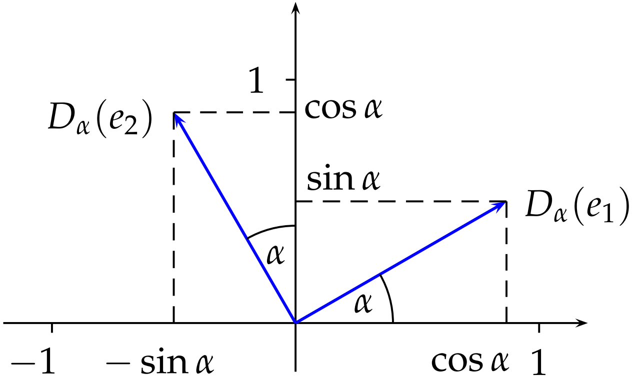
- einfaches Kriterium, dass f ein Isomorphismus ist, im folgenden Satz

Satz 4.30: Sei $\dim(V) = \dim(W) = n$, $(v_1, \ldots, v_n)$ eine Basis von V und $f : V \to W$ linear. Die Abbildung f ist genau dann ein Isomorphismus, wenn die Bilder $f(v_1), \ldots, f(v_n)$ der Basisvektoren von V eine Basis von W bilden.

Beweis *OHP*

OHP

Beispiel 4.31: *OHP*



Bemerkung 4.32:

- Zum Nachweis, dass f Isomorphismus ist: Ausrechnen der Umkehrabbildung möglich
- aber: sehr aufwändig, daher wenn möglich zu vermeiden
- Alternative: Bild und Kern ausrechnen und mit Dimensionsformel argumentieren
- Verfahren hierzu im nächsten Kapitel
- wenn Bild und Kern nicht gesucht sind: Berechnungen vermeiden
- stattdessen: Satz 4.30 anwenden
- damit: Problem ist auf Prüfung linearer Unabhängigkeit reduziert

4.3 Matrizen und lineare Abbildungen

- in diesem Kapitel: Zusammenhänge zwischen Matrizen und linearen Abbildungen
- Nebenergebnis: konkrete Gestalt für lineare Abbildungen

- Matrizen bieten die Möglichkeit, ein lineares Gleichungssystem effizient zu notieren

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + \cdots + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + \cdots + & a_{2n}x_n & = & b_2 \\ \cdots & + & \cdots & + \cdots + & \cdots & = & \cdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + \cdots + & a_{mn}x_n & = & b_m \end{array} \tag{4.2}$$

- Beschreibung durch Koeffizientenmatrix $A = (a_{ij})_{i=1\ldots m, j=1\ldots n}$ und Vektor $b = (b_1, \ldots, b_m)^T$
- man möchte auch den Lösungsvektor $x = (x_1, \ldots, x_n)^T$ in die Notation integrieren

Definition 4.33: Seien $A = (a_{ij})_{i=1\ldots m, j=1\ldots n} \in K^{m\times n}$ und $x = (x_i)_{i=1}^n \in K^n$. Dann sei

$$Ax := \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \cdots + \cdots + \cdots + \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} \in K^m .$$

Beispiel 4.34:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \quad , \quad x = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$$

- gesucht: Ax
- *OHP*

Bemerkung 4.35: Ein Vektor $x \in K^n$ beinhaltet genau dann eine Lösung des linearen Gleichungssystems (4.2), wenn

$$Ax = b$$

gilt.

- jetzt: Matrix-Vektor-Produkt als Abbildung

Satz 4.36: Sei $A \in K^{m \times n}$. Die Abbildung $f_A : K^n \to K^m, f_A(x) := Ax$ ist linear.

Beweis

- $a_1, \ldots, a_m$ die Zeilenvektoren von A
- sei $\langle \cdot, \cdot \rangle$ das euklidische Skalarprodukt

- dann

$$f_A(x) = \begin{pmatrix} a_1 \\ \vdots \\ a_m \end{pmatrix} \cdot x = \begin{pmatrix} \langle a_1^T, x \rangle \\ \vdots \\ \langle a_m^T, x \rangle \end{pmatrix}.$$

- f_A ist vektorwertige Funktion
- jede Komponente ist eine Abbildung der Gestalt $x \to \langle a_i^T, x \rangle$ und ist deswegen linear
- also ist f_A linear ■

Bemerkung 4.37:

- es gilt $(A+B)x = Ax + Bx$ und $(\lambda A)x = \lambda Ax$ mit $A, B \in K^{m \times n}$ und $\lambda \in K$
- Beweis: Nachrechnen!

- derselbe Sachverhalt: Für festes $x \in K^n$ ist $F_x : K^{m \times n} \to K^m, F_x(A) = Ax$ linear.
- Die Umkehrung des obigen Satzes gilt auch.

Satz 4.38: Sei $f : K^n \to K^m$ linear. Dann gibt es genau ein $A \in K^{m\times n}$ mit $f(x) = Ax \; \forall x \in K^n$, und es gilt

$$A = (f(e_1), \dots, f(e_n)).$$

Beweis *OHP*

Beispiel 4.39:

- sei $a \in \mathbb{R}^3$
- dann: $f : \mathbb{R}^3 \to \mathbb{R}^3$,
$$f(x) = a \times x$$
  ist linear
- es existiert A mit $f(x) = Ax$
- „In den Spalten von A stehen die Bilder der kanonischen Basisvektoren“

OHP

Beispiel 4.40:

- gesucht: Matrixdarstellung der Drehung D_α
- Ansatz: Bilder der kanonischen Basisvektoren berechnen
- bekannt: $D_\alpha(e_1) = (\cos\alpha, \sin\alpha)^T$ sowie $D_\alpha(e_2) = (-\sin\alpha, \cos\alpha)^T$
- $$D_\alpha(x) = \underbrace{\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}}_{:=A_\alpha} x\,.$$
- Matrizen der Form A_α: *Drehmatrizen*

Bemerkung 4.41: *OHP*

- Abbildungsmatrix durch die Bilder der kanonischen Einheitsvektoren als Spalten
- A legt f eindeutig fest
- also legen die Bilder der Basisvektoren eine lineare Abbildung eindeutig fest
- jetzt: beliebige lineare Abbildung f , beliebige Vektorräume (statt wie bisher K^n), Bilder von beliebigen Vektoren
- Wird f immer noch eindeutig durch die Bilder dieser Vektoren festgelegt?

Satz 4.42: Gegeben seien Vektoren $v_1, \ldots, v_n \in V$ und $w_1, \ldots, w_n \in W$. Bildet $(v_1, \ldots, v_n)$ eine Basis von V, dann gibt es genau ein $f \in Hom(V, W)$ mit $f(v_i) = w_i, 1 \le i \le n$. Die Abbildung f hat folgende Eigenschaften:

1. $\text{Bild}(f) = L(f(v_1), \ldots, f(v_n))$.
2. f ist injektiv $\Leftrightarrow w_1, \ldots, w_n$ sind linear unabhängig.

Beweis *OHP*

- bisher gesehen: für $V = K^n$ und $W = K^m$ lässt sich jede lineare Abbildung in der Form $f(x) = Ax$ schreiben
- jetzt: allgemeine (endlich) erzeugte K-Vektorräume

Satz 4.43: Seien V und W zwei K-Vektorräume, $\mathcal{B}_V = (v_1, \ldots, v_n)$ eine Basis von V und $\mathcal{B}_W = (w_1, \ldots, w_m)$ eine Basis von W und weiter $f : V \to W$ linear. Dann existiert genau eine Matrix $M^{\mathcal{B}_V}_{\mathcal{B}_W}(f) = (a_{ij}) \in K^{m \times n}$ mit

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i \quad \forall j = 1, \ldots, n \tag{4.5}$$

Beweis *OHP*

Bemerkung 4.44:

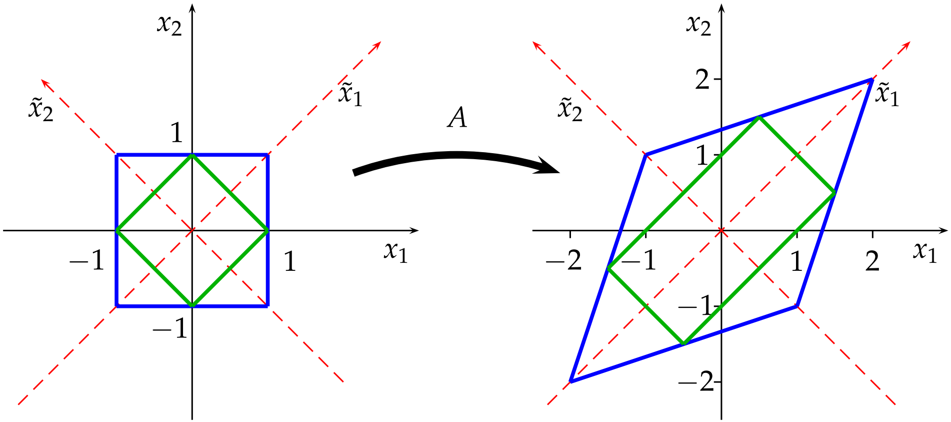
- $M_{\mathcal{B}_W}^{\mathcal{B}_V}(f)$ kodiert die Bilder der Basisvektoren
- also legt $M_{\mathcal{B}_W}^{\mathcal{B}_V}(f)$ f eindeutig fest
- Sprechweise: $M_{\mathcal{B}_W}^{\mathcal{B}_V}(f)$ ist die *Darstellungsmatrix* von f bezüglich der Basen $\mathcal{B}_V$ und $\mathcal{B}_W$.
- offensichtlich: andere Basen führen bei derselben lineare Abbildung zu unterschiedlichen Darstellungsmatrizen
- genauere Untersuchung später
- Merksatz:

  In den Spalten von A stehen die Koordinaten der Bilder der Basisvektoren von V bzgl. der gewählten Basis von W.

Beispiel 4.45: *OHP*

Beispiel 4.46:

$$A = \begin{pmatrix} 1,5 & 0,5 \\ 0,5 & 1,5 \end{pmatrix},$$



OHP

- Die Frage, wieviel Vektorräume es gibt, ist jetzt leicht zu beantworten.

Definition 4.47: Seien V und W zwei Vektorräume über demselben Körper K. Dann heißen V und W *isomorph*, Schreibweise $V \simeq W$, falls ein Isomorphismus von V nach W existiert.

Satz 4.48: Seien V und W zwei K-Vektorräume mit $\dim(V) = \dim(W) = n$. Dann gilt $K^n \simeq V \simeq W$.

Beweis *OHP*

- Fazit: „Es gibt bis auf Isomorphie nur K^n“

Beispiel 4.49: *OHP*

- jetzt wieder $V = K^n$
- Ziel: konkrete Berechnungen
- große Vereinfachung durch folgende Tatsache.

Folgerung 4.50: Sei $A \in K^{m \times n}$, und $f(x) = Ax$. Dann ist das Bild von f gleich der Linearen Hülle der Spaltenvektoren von A.

Beweis

- $e_i \in K^n$ der i-te Vektor der Standardbasis
- $f(e_i)$: i-te Spalte von A
- lineare Hülle der Spaltenvektoren von A: genau die lineare Hülle der Bilder der Basis-vektoren unter f
- also gleich $\mathrm{Bild}(f)$ ∎

- Berechnung von Bild und Kern
- viele Rechenwege möglich
- besonders effizient: folgende Reihenfolge der Berechnungen

1. Bestimmung des Kerns
2. Bestimmung der Dimension des Kerns
3. Bestimmung des Rangs ($\Rightarrow$ Dimensionsformel)
4. Bestimmung des Bildes

Beispiel 4.51: Gegeben ist

$$f\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 + x_2 \\ x_1 - x_2 + x_3 \\ 4x_1 - x_2 + 2x_3 \end{pmatrix}.$$

- zu zeigen: f linear
- man bestimme Bild(f), ker(f) und deren Dimensionen

OHP

Beispiel 4.52: *OHP*

- wesentlich: Sind die Räume endlich erzeugt, lassen sich alle relevanten Berechnungen auf Berechnungen an Matrizen zurückführen

4.4 Abbildungsverkettung und Matrizenmultiplikation

- in diesem Kapitel: Hintereinanderausführung linearer Abbildungen

Satz 4.53: Seien U, V, W K-Vektorräume und $f : V \to W$ sowie $g : U \to V$ linear. Dann ist auch $f \circ g : U \to W$ linear.

Beweis *OHP*

Folgerung 4.54: Führt man Isomorphismen hintereinander aus, erhält man wiederum einen Isomorphismus.

- etwas allgemeiner: Ranggleichung

Satz 4.55: Seien in der Situation von Satz 4.53 $\dim(V) = \dim(W) = n$ und f ein Isomorphismus. Dann gilt:

$$\operatorname{rg}(f \circ g) = \operatorname{rg}(g)\,.$$

Beweis *OHP*

- seien $f : K^n \to K^m$ und $g : K^\ell \to K^n$
- $f \circ g$ linear
- es existiert $C \in K^{m \times \ell}$ mit $(f \circ g)(x) = Cx$
- jetzt: Berechnung von C

OHP

Definition 4.56: Es sei $A \in K^{m\times n}$ und $B \in K^{n\times \ell}$. Dann heißt $C = AB \in K^{m\times \ell}$ mit

$$c_{ik} = a_{i1}b_{1k} + \cdots + a_{in}b_{nk} = \sum_{j=1}^{n} a_{ij}b_{jk} \qquad \forall i = 1,\ldots,m; k = 1,\ldots,\ell$$

das *Produkt* der Matrizen A und B.

$$\begin{pmatrix} b_{11} & \cdots & b_{1k} & \cdots & b_{1\ell} \\ b_{21} & \cdots & b_{2k} & \cdots & b_{2\ell} \\ \vdots & & \vdots & & \vdots \\ b_{n1} & \cdots & b_{nk} & \cdots & b_{n\ell} \end{pmatrix} = B$$

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \qquad \begin{pmatrix} & & \\ & c_{ik} & \\ & & \end{pmatrix} = C$$

Beispiel 4.57: Gegeben sind die Matrizen

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad B = \begin{pmatrix} 4 & 3 & 2 & 1 \\ 0 & -1 & -2 & -3 \\ 2 & 5 & 4 & 3 \end{pmatrix}.$$

OHP

Bemerkung 4.58:

- Produktbildung nur möglich, wenn die Spaltenzahl von A mit der Zeilenzahl von B übereinstimmt.
- sonst: Produkt von A und B nicht definiert
- wichtige Eigenschaften der Matrixmultiplikation

Satz 4.59: Seien A, B, C so, dass die nachfolgend vorkommenden Matrixmultiplikationen definiert sind. Dann gilt:

1. $A(BC) = (AB)C$ (Assoziativgesetz)
2. $A(B + C) = AB + AC$ und $(A + B)C = AC + BC$ (Distributivgesetz)
3. $(AB)^T = B^T A^T$
4. Sei $A \in K^{m\times n}$ und $E_k \in K^{k\times k}$ die $(k \times k)$-Einheitsmatrix. Dann gilt $AE_n = E_m A = A$.

5. Es sei 0 eine Nullmatrix. Falls sich die Produkte $0 \cdot A$ und $A \cdot 0$ bilden lassen, sind sie ebenfalls eine Nullmatrix.

Beweis *OHP*

Bemerkung 4.60: Das Matrixprodukt ist im Allgemeinen *nicht* kommutativ: Für

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

gilt

$$AB = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{aber} \quad BA = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Beispiel 4.61: *OHP*

Bemerkung 4.62:

- Deutung eines Vektors als einspaltige Matrix mit n Zeilen
- dann: euklidisches Skalarprodukt als Matrixmultiplikation

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i = x^T \cdot y,$$

- dabei: Man identifiziert die 1×1-Matrix $x^T \cdot y$ mit ihrem einzigen Eintrag $\langle x, y \rangle$
- analog: Umwandlung eines Spalten- in einen Zeilenvektor als Spezialfall der Transposition von Matrizen
- analog: Matrix-Vektor-Multiplikation als Spezialfall der Matrixmultiplikation
- $f : V \to W$ ein Isomorphismus mit Darstellungsmatrix A

- Darstellungsmatrix des Isomorphismus f^{-1}: B
- $f \circ f^{-1} = f^{-1} \circ f = id$
- also $AB = BA = E$
- B verhält sich wie das inverse Element zu A.

Definition 4.63: Sei A eine quadratische Matrix. Gibt es eine Matrix A^{-1} mit

$$AA^{-1} = A^{-1}A = E,$$

so heißt A *invertierbar* oder auch *regulär*. A^{-1} wird als *Inverse* von A bezeichnet.

Folgerung 4.64: Eine lineare Abbildung $f : K^n \to K^m$ ist genau dann invertierbar, wenn ihre Darstellungsmatrix invertierbar ist.

Folgerung 4.65: Jede invertierbare Matrix ist quadratisch.

Beweis

- $f : K^n \to K^m, f(x) = Ax$ ist invertierbar, also Isomorphismus
- es folgt $\dim(K^n) = \dim(K^m)$, also $n = m$ ■

Beispiel 4.66:

- Nicht jede quadratische Matrix ist invertierbar
- betrachte Nullmatrix
- $f : K^n \to K^n, f(x) = 0$ nicht injektiv und damit nicht invertierbar
- also Nullmatrix nicht invertierbar

Wir stellen wesentliche Eigenschaften im Hinblick auf Inverse in folgendem Satz zusammen.

Satz 4.67: Seien $A, B \in K^{n \times n}$ invertierbar. Dann gilt:

1. $$AB = E \Leftrightarrow BA = E \Leftrightarrow B = A^{-1}$$
2. AB ist invertierbar, und es gilt $(AB)^{-1} = B^{-1}A^{-1}$.
3. A^{-1} ist invertierbar, und es gilt $(A^{-1})^{-1} = A$.
4. A^T ist invertierbar, und es gilt $(A^T)^{-1} = (A^{-1})^T$.
5. Für $\lambda \in K \setminus \{0\}$ ist λA invertierbar, und es gilt $(\lambda A)^{-1} = \frac{1}{\lambda}A^{-1}$.

Beweis *OHP*

- Konsequenz der Eigenschaften der inversen Matrix:

Satz 4.68: Die Menge aller invertierbaren Matrizen in $K^{n\times n}$ bilden zusammen mit der Matrixmultiplikation eine Gruppe.

Beweis

- Menge aller invertierbaren $n \times n$-Matrizen: $GL(n; K)$
- oben bewiesen: Das Produkt von invertierbaren Matrizen und die Inverse einer invertierbaren Matrix sind invertierbar
- Multiplikation zweier Matrizen ist eine Verknüpfung auf $GL(n; K)$
- Assoziativität: In Satz 4.59 gezeigt, ebenso die Existenz des neutralen Elements E
- Existenz der Inversen klar ∎

Bemerkung 4.69:

1.
   - $GL(n;K)$ „general linear group“
   - eines der wichtigsten Beispiele nicht-abelscher Gruppen

2.
   - $GL(n;K)$ kein Körper
   - ein Grund: Summe zweier invertierbarer Matrizen muss nicht invertierbar sein
   - Beispiel:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

   - offenbar invertierbar

- aber:

$$A + B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

- Bild von $A + B$: Lineare Hülle der Spaltenvektoren
- offenbar $\mathrm{rg}(f) = 1$
- Abbildung nicht surjektiv, also nicht bijektiv
- damit $A + B$ nicht invertierbar

- jetzt: *Berechnung* der Inversen einer geg. Matrix $A \in GL(n; K)$
- $A^{-1} = B = (b_1, \ldots, b_n)$
- Satz 4.67: Inverse B ist Lösung von $AB = E$

- sei $E = (e_1, \dots, e_n)$
- spaltenweise: $Ab_i = e_i$
- n lineare Gleichungssystem mit derselben Matrix
- Fortgang des Gauß-Verfahrens hängt von A, nicht von b_i ab
- simultanes Lösen möglich

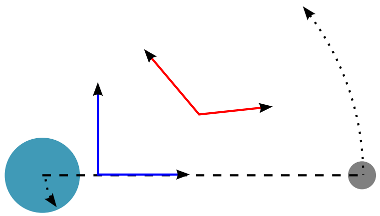
Beispiel 4.70: Gegeben sei

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

OHP

4.5 Koordinatentransformationen

- wesentlich bei mathematischer Modellbildung: Problem in Gleichungen fassen
- in vielen Fällen: geeignetes Koordinatensystem finden
- Beispiel: Bewegung von Erde und Mond



- rotes Koordinatensystem ungünstig, blaues wohl eher geeignet

- Koordinatentransformation
- allgemein in der Analysis: Koordinatentransformation ist ein *Diffeomorphismus*
- Diffeomorphismus: invertierbare stetig differenzierbare Abbildung, deren Inverse ebenfalls stetig differenzierbar ist
- hier: *lineare* Koordinatentransformationen
- wesentliche Vereinfachung
- Folge: Lage des Nullpunkts ändert sich wegen $f(0) = 0$ nicht

Beispiel 4.71:

- bisher: Koordinaten eines Vektors in K^n immer bezüglich der kanonischen Einheitsvektoren $e_1, \ldots, e_n$
- z. B. $x = (3, 1, 4)^T$: $x = 3e_1 + e_2 + 4e_3$

- Man konnte x mit seinen Koordinaten identifizieren
- jetzt: neue Basis bestehend aus

$$b_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, b_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, b_3 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

- x hat bezüglich der Basis $\mathcal{B} = (b_1, b_2, b_3)$ die Koordinaten $(2, 1, 3)^T$
- Schreibweise:

$$K_{\mathcal{B}}(x) = (2, 1, 3)^T \,.$$

- keine Basis angegeben: Standardbasis aus den kanonischen Einheitsvektoren
- jetzt: beliebige endlich erzeugte Vektorräume

Satz 4.72: Sei V ein K-Vektorraum mit einer Basis $\mathcal{B} = (v_1, \ldots, v_n)$. Dann existiert genau ein Isomorphismus $\varphi_{\mathcal{B}} : K^n \to V$ mit $\varphi_{\mathcal{B}}(e_i) = v_i, 1 \le i \le n$.

Beweis Man wähle in Satz 4.42 $V = K^n$ und $W = V$. ∎

- damit: eindeutige Zuordnung eines abstrakten Vektors v zu einem Tupel von Elementen aus K und umgekehrt immer vorhanden
- Tupel von Elementen in K^n: Koordinaten von v

Definition 4.73: Der Isomorphismus $\varphi_{\mathcal{B}}$ aus Satz 4.72 heißt *Koordinatenabbildung*, und für $v \in V$ heißen $K_{\mathcal{B}}(v) := \varphi_{\mathcal{B}}^{-1}(v) \in K^n$ die *Koordinaten von v bezüglich $\mathcal{B}$*.

- es wurden früher schon Koordinaten definiert
- Sind beide Definitionen konsistent?
- früher: Koordinaten als die Vorfaktoren der Linearkombination der Basisvektoren

- bekannt: Diese Koordinaten sind eindeutig
- sei $v \in V$ und $\mathcal{B} = (v_1, \dots, v_n)$ beliebige Basis
- dann: $v = \sum_{i=1}^{n} \alpha_i v_i$ mit eindeutigen $\alpha_i \in K$
- andererseits: $\varphi_{\mathcal{B}}$ ist linear. Deswegen gilt

$$\varphi_{\mathcal{B}}((\lambda_1, \dots, \lambda_n)^T) = \varphi_{\mathcal{B}}\Big(\sum_{i=1}^{n} \lambda_i e_i\Big) = \sum_{i=1}^{n} \lambda_i \varphi_{\mathcal{B}}(e_i) \,.$$

- Nach Konstruktion: $\varphi_{\mathcal{B}}(e_i) = v_i$
- deswegen $\varphi_{\mathcal{B}}((\lambda_1, \dots, \lambda_n)^T) = \sum_{i=1}^{n} \lambda_i v_i$
- es soll aber auch $\varphi_{\mathcal{B}}((\lambda_1, \dots, \lambda_n)^T) = v$

- nur möglich, wenn $\alpha_i = \lambda_i, 1 \leq i \leq n$
- beide Definitionen von Koordinaten liefern dasselbe
- beide Definitionen von Koordinaten sind konsistent

Beispiel 4.74:

- gegeben $V = K^n$ mit der Standardbasis
- $\varphi_{\mathcal{E}} = id$
- der Vektor ist gleich seinen Koordinaten in der Standardbasis
- obiges Beispiel: $\varphi_{\mathcal{B}}(x) = Bx$ mit der Matrix $B = (b_1, b_2, b_3)$
- denn: $Be_j = b_j$ für $1 \leq j \leq 3$

Bemerkung 4.75:

1.
- ohne Angabe einer Basis: Koordinaten wertlos
- betrachte Vektorraum P_2
- Monombasis $(1, x, x^2)$: $(1, 0, -1)$ entspricht $p(x) = 1 - x^2$
- Basis $(1, 1-x, (1-x)^2)$: $(1, 0, -1)$ entspricht $q(x) = 2x - x^2$
- bisher: Koordinaten immer implizit bzgl. der Standardbasis $\mathcal{E}$ angenommen
- aber: in abstrakten Vektorräumen gibt es kein Analogon zur Standardbasis

2.
- alternativ zur Basis: Angabe der Koordinatenabbildung $\varphi_{\mathcal{B}}$
- dann: jeder Basisvektor bestimmt durch $\varphi_{\mathcal{B}}(e_j)$
- umgekehrt: Basis $\mathcal{B}$ legt die Koordinatenabbildung eindeutig fest
- beide Angaben sind äquivalent

- jetzt: Vektorraum der Dimension n mit Basen $\mathcal{A} = (a_1, \ldots, a_n)$ und $\mathcal{B} = (b_1, \ldots, b_n)$
- für $v \in V$: $K_{\mathcal{A}}(v)$ und $K_{\mathcal{B}}(v)$

- die Basen $\mathcal{A}$ und $\mathcal{B}$ induzieren zwei Koordinatenabbildungen $\varphi_\mathcal{A}$ und $\varphi_\mathcal{B}$
- Darstellung als Diagramm

$$\begin{tikzcd} & V & \\ K^n \arrow[ur, "\varphi_\mathcal{A}"] \arrow[rr, "T_\mathcal{B}^\mathcal{A} = \varphi_\mathcal{B}^{-1} \circ \varphi_\mathcal{A}"'] & & K^n \arrow[ul, "\varphi_\mathcal{B}"'] \end{tikzcd}$$

- im Diagramm: nur Isomorphismen
- Diagramm kann in beliebiger Richtung durchlaufen werden; das Ergebnis bleibt gleich
- „kommutatives Diagramm“
- $\varphi_\mathcal{B}^{-1} \circ \varphi_\mathcal{A} : K^n \to K^n$ ist ein Automorphismus von K^n

- es existiert eine Matrixdarstellung
- $T_{\mathcal{B}}^{\mathcal{A}}$: *Transformationsmatrix des Basiswechsels von* $\mathcal{A}$ *nach* $\mathcal{B}$

Bemerkung 4.76:

- geg. $v \in V$ beliebig, $K_{\mathcal{A}}(v) = (x_1, \ldots, x_n)^T$ und $K_{\mathcal{B}}(v) = (y_1, \ldots, y_n)^T$

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = T_{\mathcal{B}}^{\mathcal{A}} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

- Sind die Koordinaten von v bezüglich $\mathcal{A}$ bekannt, kann man mithilfe der Matrix $T_{\mathcal{B}}^{\mathcal{A}}$ die Koordinaten von v bezüglich $\mathcal{B}$ berechnen.

Bemerkung 4.77:

- $T^{\mathcal{B}}_{\mathcal{A}} = (T^{\mathcal{A}}_{\mathcal{B}})^{-1}$ (vgl. kommutatives Diagramm)
- Koordinatentransformation in „umgekehrte Richtung“: Lösen eines LGS bzw. Invertierung einer Matrix
- sei jetzt $V = K^n$
- bei beliebiger Basis: quadratische Matrix aus den Basisvektoren
- die Vektoren von $\mathcal{A}$ mögen die Matrix A bilden
- die Vektoren von $\mathcal{B}$ mögen die Matrix B bilden
- kommutatives Diagramm:

$$\begin{tikzcd} x \arrow[r, "E"] & x \\ K_{\mathcal{A}}(x) \arrow[u, "A"] \arrow[r, "T^{\mathcal{A}}_{\mathcal{B}}"'] & K_{\mathcal{B}}(x) \arrow[u, "B"'] \end{tikzcd}$$

Beispiel 4.78: *OHP*

- jetzt: Änderung der Darstellungsmatrix bei Basiswechsel
- gegeben: $f : V \to W$ linear
- $\mathcal{A}$ Basis von V, $\mathcal{B}$ Basis von W
- Koordinatenabbildungen: $\varphi_{\mathcal{A}} : K^n \to V$ und $\varphi_{\mathcal{B}} : K^m \to W$
- kommutatives Diagramm:

$$\begin{tikzcd} K^n \arrow[r, "M_{\mathcal{B}}^{\mathcal{A}}(f)"] \arrow[d, "\varphi_{\mathcal{A}}"'] & K^m \arrow[d, "\varphi_{\mathcal{B}}"] \\ V \arrow[r, "f"'] & W \end{tikzcd}$$

- Das kommutative Diagramm ist wirklich kommutativ.

Satz 4.79: Seien V, W endlich erzeugte K-Vektorräume mit Basen $\mathcal{A}$ und $\mathcal{B}$ sowie $f \in Hom(V; W)$. Dann gilt

$$M_{\mathcal{B}}^{\mathcal{A}}(f) = \varphi_{\mathcal{B}}^{-1} \circ f \circ \varphi_{\mathcal{A}} \,.$$

Beweis *OHP*

- 1. Weg: direkte Anwendung von f auf v
- 2. Weg: Man errechnet die Koordinaten von v, bestimmt dann mit der Darstellungsmatrix die Koordinaten von $f(v)$, und ermittelt daraus $f(v)$
- Beide Wege sind gleichwertig!

Eine lineare Abbildung bildet einen Vektor v auf sein Bild $f(v) = w$ ab, ihre Darstellungsmatrix bildet die Koordinaten von v auf die Koordinaten von w ab.

Folgerung 4.80:

- Spezialfall $V = W$, verschiedene Basen $\mathcal{A}$ und $\mathcal{B}$
- sei $f = id$

$$M_{\mathcal{B}}^{\mathcal{A}}(id) = \varphi_{\mathcal{B}}^{-1} \circ id \circ \varphi_{\mathcal{A}} = T_{\mathcal{B}}^{\mathcal{A}} \,.$$

- Darstellungsmatrizen verallgemeinern Transformationsmatrizen des Basiswechsels.

- Hauptergebnis dieses Kapitels:

Satz 4.81: Seien V und W endlich erzeugt mit Basen $\mathcal{A}$ und $\mathcal{A}'$ bzw. $\mathcal{B}$ und $\mathcal{B}'$. Sei weiter $f : V \to W$ linear. Dann gilt

$$M_{\mathcal{B}'}^{\mathcal{A}'}(f) = T_{\mathcal{B}'}^{\mathcal{B}} \cdot M_{\mathcal{B}}^{\mathcal{A}}(f) \cdot \left(T_{\mathcal{A}'}^{\mathcal{A}}\right)^{-1} .$$

$$\begin{tikzcd}
K^n \arrow[rrr, "M_{\mathcal{B}}^{\mathcal{A}}(f)"] \arrow[dr, "\varphi_{\mathcal{A}}"] \arrow[dd, "T_{\mathcal{A}'}^{\mathcal{A}}"] & & & K^m \arrow[dl, "\varphi_{\mathcal{B}}"'] \arrow[dd, "T_{\mathcal{B}'}^{\mathcal{B}}"] \\
& V \arrow[r, "f"] & W & \\
K^n \arrow[ur, "\varphi_{\mathcal{A}'}"] \arrow[rrr, "M_{\mathcal{B}'}^{\mathcal{A}'}(f)"] & & & K^m \arrow[ul, "\varphi_{\mathcal{B}'}"']
\end{tikzcd}$$

Beweis

- Überlegungen oben: alle vier Teildiagramme sind kommutativ
- also auch das gesamte Diagramm ■

Folgerung 4.82:

- Sei $V = K^n = W$, $\mathcal{A}$ und $\mathcal{B}$ Basen von K^n
- S: Matrix des Basiswechsels von $\mathcal{A}$ nach $\mathcal{B}$
- gegeben f linear mit Darstellungsmatrix A bzgl. $\mathcal{A}$
- Darstellungsmatrix B von f bzgl. auf $\mathcal{B}$:

$$B = SAS^{-1}\,.$$

Beispiel 4.83: *OHP*

Beispiel 4.84: *OHP*

Kapitel 5

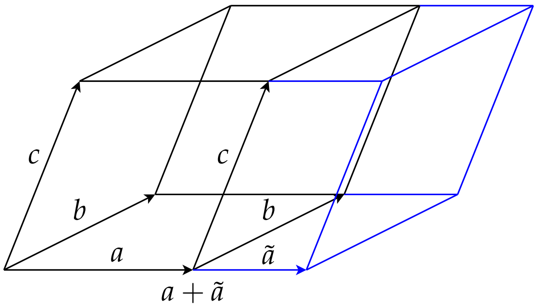
Determinanten

5.1 Motivation und Einführung

- in LA1: Determinante für 2×2- und 3×3-Matrizen betrachtet
- Rechenvorschrift
- Anwendung: $Ax = b$ hat genau dann eine eindeutige Lösung, wenn $\det(A) \neq 0$
- zudem: det liefert ein einfache Formel für das orientierte Volumen eines Spats
- jetzt: Determinante für beliebiges n
- Vorgehen bei Verallgemeinerungen: wesentliche Eigenschaften wählen, alles mit diesen Eigenschaften mit demselben Namen bezeichnen wie den vorherigen Spezialfall
- Welche Eigenschaften wählen?
- dazu: Eigenschaften aus der Charakterisierung als Volumenformel

1.
- Reihenfolge zweier Matrixspalten vertauschen: Rechtssystem wird Linkssystem oder umgekehrt
- Vorzeichen der Determinante ändert sich

2.
- Volumen des Spats ist die Summe der Volumina der Teilspate



3.
- Verlängern einer Seite um λ: Volumen vervielfacht sich um λ

4. Das Volumen des Einheitswürfels beträgt eins.

- ab jetzt immer: K Körper und $n \in \mathbb{N}$

Definition 5.1: Eine Abbildung $\det : K^{n \times n} \to K$ heißt *Determinante*, wenn gilt

1. det ist *alternierend*: Vertauscht man in $A = (a_1, \dots, a_n) \in K^{n \times n}$ zwei benachbarte Spalten, dann ändert $\det(A)$ ihr Vorzeichen:
$$\det(\dots, a_i, a_{i+1}, \dots) = -\det(\dots, a_{i+1}, a_i, \dots)$$
2. det ist *homogen*: Multipliziert man eine Spalte von A mit $\lambda \in K$, so gilt
$$\det(a_1, \dots, \lambda a_i, \dots, a_n) = \lambda \det(a_1, \dots, a_i, \dots, a_n)$$
3. det ist *additiv*: Stimmen zwei Matrizen überall bis auf die i-te Spalte überein, so addiert man ihre Determinanten, indem man die beiden i-ten Spalten addiert und dann die

Determinante der Summe bildet:

$$\begin{aligned} &\det(a_1,\ldots,a_{i-1}, \quad a_i, \quad a_{i+1},\ldots,a_n) \\ +&\det(a_1,\ldots,a_{i-1}, \quad a_i^*, \quad a_{i+1},\ldots,a_n) \\ =&\det(a_1,\ldots,a_{i-1}, a_i+a_i^*, a_{i+1},\ldots,a_n) \end{aligned}$$

4. det ist *normiert*: Für die Einheitsmatrix E gilt $\det(E)=1$.

Bemerkung 5.2:

- Definition über abstrakte Eigenschaften: *axiomatische Definition*
- dann: (einige) Eigenschaften sind klar
- Existenz und Eindeutigkeit sind nicht klar
- Konsequenz: Man muss die Wohldefiniertheit beweisen

- Alternative: Definition z. B. über eine Rechenvorschrift
- Ist diese immer durchführbar, ist Existenz und Eindeutigkeit klar.
- aber: Eigenschaften unklar
- Determinanten wurden historisch als Rechenvorschriften definiert
- axiomatische Definition zuerst bei Weierstraß
- „summierter Beweisaufwand“ bei beiden Zugängen in etwa gleich

Bemerkung 5.3:

- bisweilen Kurzschreibweise $|A|$ statt $\det(A)$.

$$\begin{vmatrix} 1 & -1 \\ 2 & 5 \end{vmatrix} = \det\begin{pmatrix} 1 & -1 \\ 2 & 5 \end{pmatrix}$$

5.2 Vorbereitung: Elementarmatrizen

- Untersuchung der Eigenschaften der Determinante wird durch *Elementarmatrizen* stark vereinfacht
- ebenso bei der Analyse des Gauß-Verfahrens
- daher: wir betrachten Elementarmatrizen

Definition 5.4: Seien $1 \le i, j \le n$ mit $i \neq j$ und $\lambda \in K \setminus \{0\}$ gegeben. Dann sei

$$C1 := \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & \lambda & \ddots & \\ & & & 1 \end{pmatrix} \in K^{n \times n} \text{ bzw. } C1 := \begin{pmatrix} 1 & & & \\ & \ddots & \lambda & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \in K^{n \times n},$$

wobei der (i,j)-te Eintrag den Wert λ annehmen soll und alle anderen Einträge außerhalb der Hauptdiagonalen 0 sein sollen. Sei $C2$ die Matrix, die man aus der Einheitsmatrix gewinnt, indem man die i-te und j-te Spalte vertauscht, also

$$C2 := \begin{pmatrix}
1 & & & \vert & & & & \vert & & & \\
& \ddots & & \vert & & & & \vert & & & \\
& & 1 & \vert & & & & \vert & & & \\
- & - & - & 0 & - & - & - & 1 & - & - & - \\
& & & \vert & 1 & & & \vert & & & \\
& & & \vert & & \ddots & & \vert & & & \\
& & & \vert & & & 1 & \vert & & & \\
- & - & - & 1 & - & - & - & 0 & - & - & - \\
& & & \vert & & & & \vert & 1 & & \\
& & & \vert & & & & \vert & & \ddots & \\
& & & \vert & & & & \vert & & & 1
\end{pmatrix}.$$

Zuletzt definieren wir

$$C3 := \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & \lambda & \\ & & & 1 \end{pmatrix} \in K^{n\times n}.$$

Matrizen der Gestalt $C1, C2$ oder $C3$ nennt man *Elementarmatrizen*.

Bemerkung 5.5: Die Multiplikation einer Matrix A von links mit einer Elementarmatrix Ci entspricht der Anwendung einer elementaren Zeilenoperation Zi des Gauß-Verfahrens auf A (vgl. Kapitel 1.6.1).

Beispiel 5.6:

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

OHP

- Beispiel zeigt: eine Folge von elementaren Zeilenoperationen entspricht der Multiplikation mit Elementarmatrizen von links

- allgemeiner gilt:

Folgerung 5.7: Eine Matrix A lässt sich genau dann als Produkt von Elementarmatrizen darstellen, wenn sie sich mit dem Gauß-Verfahren in die reduzierte Stufenform überführen lässt.

Bemerkung 5.8: Elementarmatrizen nach Definition 5.4 sind invertierbar, ihre Inversen sind wiederum Elementarmatrizen.

Beweis

$$C1^{-1} = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & -\lambda & \ddots & \\ & & & 1 \end{pmatrix} \quad , \quad C3^{-1} = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1/\lambda & \\ & & & 1 \end{pmatrix} ,$$

- es gilt $\lambda \neq 0$
- $C2^{-1} = C2$ ∎

- jetzt: Multiplikation mit Elementarmatrizen von *rechts*
- dies entspricht *Spalten*- statt *Zeilen*operationen
- Bezeichnung: S1 – S3 analog zu Z1 – Z3
- Man erhält folgenden Satz durch elementares Ausrechnen

Satz 5.9: Die Spaltenoperation Si auf A entspricht genau der Multiplikation mit der Elementarmatrix Ci von rechts, $i \in \{1, 2, 3\}$.

5.3 Eigenschaften der Determinante

- Wohldefiniertheit der Determinante: später
- jetzt: weitere Eigenschaften der Determinante?
- zur Vereinfachung: Körper $K \in \{\mathbb{R}, \mathbb{C}\}$
- Der Beweis des folgenden Satzes wird nachgeholt

Satz 5.10: Für $A \in K^{n \times n}$ gilt

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

wobei A^T die zu A transponierte Matrix ist.

- Folgerung: Die definierenden Eigenschaften der Determinante gelten sowohl für Zeilen als auch für Spalten

- daher: wir zeigen die folgenden Aussagen für Spalten
- für dieselben Aussagen für Zeilen: betrachte A^T

Bemerkung 5.11:

1. Vertauscht man zwei beliebige Spalten/Zeilen, so ändert die Determinante ihr Vorzeichen.
2. Besitzt eine Matrix A zwei gleiche Spalten/Zeilen, so gilt $\det(A) = 0$.

Beweis *OHP*

- jetzt: Wirkung von Z1 – Z3 bzw. S1 – S3 auf die Determinante

Bemerkung 5.12:

1. S1 und Z1 ändern die Determinante einer Matrix nicht.
2. S2 und Z2 kehren das Vorzeichen der Determinante um.
3. S3 und Z3: Multiplikation einer Zeile oder Spalte mit einem Faktor λ vervielfacht den Wert der Determinante um den Faktor λ.

Beweis *OHP*

Satz 5.13:

1. Für die Elementarmatrizen aus Definition 5.4 gilt $\det(C1) = 1$, $\det(C2) = -1$ und $\det(C3) = \lambda$.

2. Für eine Elementarmatrix C und eine beliebige quadratische Matrix A gilt $\det(CA) = \det(C)\det(A) = \det(AC)$.

Beweis *OHP*

Folgerung 5.14: Für eine obere Dreiecksmatrix

$$A = \begin{pmatrix} \lambda_1 & \cdots & \\ & \ddots & \vdots \\ & & \lambda_n \end{pmatrix}$$

gilt $\det(A) = \lambda_1 \cdot \ldots \cdot \lambda_n$.

Beweis *OHP*

- zentrale Anwendung der Determinante: Kriterium zur eindeutigen Lösbarkeit eines linearen Gleichungssystems

Satz 5.15: Sei $A \in K^{n \times n}$. Dann gilt

$$A \text{ invertierbar } \Leftrightarrow \det(A) \neq 0\,.$$

Beweis *OHP*

Satz 5.16: Für $A, B \in K^{n \times n}$ gilt $\det(AB) = \det(A)\det(B)$.

Beweis *OHP*

Folgerung 5.17: Ist A invertierbar, dann gilt

$$\det(A^{-1}) = (\det(A))^{-1}\,.$$

Beweis Es gilt $\det(A)\det(A^{-1}) = \det(AA^{-1}) = \det(E) = 1.$ ■

5.4 Verfahren zur Berechnung der Determinante

5.4.1 Leibniz-Formel

- Ziel: explizite Formel zur Berechnung der Determinante
- Ausgangspunkt: 2 × 2-Matrix $A = (a, b)$

$$|A| = \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}.$$

Mit den kanonischen Einheitsvektoren e_1 und e_2 gilt

$$a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = a_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = a_1 e_1 + a_2 e_2$$

und daher

$$\det(a, b) = \det(a_1 e_1 + a_2 e_2, b) = a_1 \det(e_1, b) + a_2 \det(e_2, b)$$

- analoge Rechnung für b:

OHP

- analoge Rechnung für $n = 3$: Sarrus-Regel (siehe LA1)
- bei allgemeinem n:

1. Man schreibt a_1 bis a_n als $a_i = \sum_{k=1}^{n} a_{ki} e_k$.
2. - Additivität der Determinante für alle Spalten nutzen
   - Ergebnis: Summe aus n^n Summanden
   - jeder Summand: n Vorfaktoren und Determinante aus Einheitsvektoren
3. - Taucht derselbe Einheitsvektor mehrmals auf, fällt der Summand weg.
   - Nur Summanden mit Permutationen von $(e_1, \ldots, e_n)$ in der Determinanten bleiben.
4. - Determinanten: 1 oder -1
   - abhängig davon, ob eine gerade oder ungerade Anzahl von Vertauschungen zur Überführung in $\det(E)$ nötig ist

Definition 5.18: Eine bijektive Abbildung $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ heißt *Permutation*.

Beispiel 5.19:

1. Die identische Abbildung $\sigma(i) = i \, \forall \, 1 \leq i \leq n$ ist eine Permutation.
2. Die Inverse einer Permutation σ ist wiederum eine Permutation wegen $(\sigma^{-1})^{-1} = \sigma$.
3. Sind σ und ν Permutationen, dann auch $\sigma \circ \nu$, weil Verkettungen bijektiver Funktionen wieder bijektiv sind.

Bemerkung 5.20:

- Menge aller Permutationen auf $\{1, \ldots, n\}$: Gruppe bzgl. der Abbildungsverkettung
- *symmetrische Gruppe* S_n
- S_n ist für $n > 2$ nicht abelsch

- S_n besitzt $n!$ Elemente

Beweis *OHP*

- Ergebnis: Leibniz-Formel

Satz 5.21 (Leibniz-Formel): Für $A \in K^{n\times n}$ gilt

$$\det(A) = \sum_{\sigma\in S_n} \delta(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}$$

mit

$$\delta(\sigma) = \begin{cases} 1, & \text{falls } \sigma(1), \ldots, \sigma(n) \text{ durch eine gerade Zahl von Vertau-} \\ & \text{schungen zu } 1, 2, \ldots, n \text{ umgeordnet werden kann} \\ -1, & \text{sonst} \end{cases}$$

- Existenz der Determinante ist jetzt klar, weil eine konkrete Berechnungsformel gefunden ist

- es fehlt: Nachweis der Eindeutigkeit (siehe Literatur)

- es fehlt: Das Ergebnis der Leibnizschen Formel besitzt die definierenden Eigenschaften der Determinante (siehe Literatur)

- jetzt endlich: Beweis von $|A| = |A^T|$

Beweis *OHP*

- Zum Beweis wurden keine Eigenschaften der Determinante verwendet, die mit $|A| = |A^T|$ gezeigt wurden
- also: kein Zirkelschluss

5.4.2 Laplacescher Entwicklungssatz

- jetzt: alternative Berechnungsmethode
- ohne Beweis und Herleitung (siehe Literatur)

Definition 5.22: Für $A \in K^{n\times n}$ bezeichne A_{ij} die Matrix in $K^{(n-1)\times(n-1)}$, die durch Streichen der i-ten Zeile und der j-ten Spalte aus A hervorgeht.

Satz 5.23 (Laplacescher Entwicklungssatz): Es sei $A = (a_{ij}) \in K^{n\times n}$ und j mit $1 \leq j \leq n$. Dann gilt

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

Man spricht von der *Entwicklung der Determinante nach der j-ten Spalte*. Ebenso ist eine

Entwicklung der Determinante nach einer Zeile möglich:

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

- Berechnungsformel verwendbar als rekursive Definition der Determinante

Beispiel 5.24: *OHP*

5.4.3 Gauß-Algorithmus

- Vorgehen: Man bringt A mit Z1 – Z3 bzw. S1 – S3 auf Stufenform B
- dann: Determinante von B als Produkt der Hauptdiagonalelemente
- aber: Z2 bzw. S2 verändern das Vorzeichen der Determinante
- bei Anwendung von Z3 oder S3: Determinante ändert sich mit dem Faktor der Skalierung
- Dies muss man in der Rechnung ausgleichen.

Beispiel 5.25: *OHP*

Beispiel 5.26: *OHP*

Beispiel 5.27:

- Kombination von Zeilen- und Spaltenoperationen ist möglich
- Rechnung kann so sehr vereinfacht werden

OHP

5.4.4 Determinantenberechnung in der Praxis

- In der Praxis: Welches Verfahren wählen?
- Gegenfrage: Berechnung per Computer oder per Hand?
- üblicherweise: Determinanten für $n > 5$ sind der Hand kaum sinnvoll zu berechnen
- zur Handrechnung:
- nur Erfahrungswerte angebbar
- Grund: persönliche Vorlieben und Gewohnheit entscheiden bei der Wahl des Verfahrens mit

1. Wir raten von der Leibnizschen Formel als konkretes Rechenverfahren ab.
2. - Ausnahme: $n = 2$ und $n = 3$

- in diesen Fällen: Verwendung der expliziten Formeln (siehe LA 1) vorteilhaft
- Spezialfälle der Leibnizschen Formel

3.
- sowohl Gauß-Verfahren als auch Laplace-Entwicklung als Rechenverfahren geeignet
- Laplace-Entwicklung überlegen, wenn Spalten oder Zeilen mit wenigen Nicht-Null-Einträgen vorliegen
- dann: Entwicklung nach dieser Spalte/Zeile
- es müssen viele Summanden erst gar nicht berechnet werden, da ohnehin 0

4.
- Kombination von Verfahren
- z. B.: gegeben sei eine 4×4-Matrix
- Start mit Laplace-Entwicklung, Berechnung der entstehenden 3×3-Matrizen mit der Sarrus-Regel

- oder: Man schafft durch Gauß-Elimination Zeilen oder Spalten mit wenigen Nicht-Null-Einträgen
- dann: Laplace-Entwicklung

5. unerlässlich: Übung und Routine

Beispiel 5.28: *OHP*

- jetzt: Berechnung mit dem Computer
- (grobes) Komplexitätsmaß: Anzahl der arithmetischen Operationen
- Gauß-Verfahren: etwa $\frac{n^3-n}{3}$ Operationen zum Erzeugen der Nullen unterhalb der Diagonalen
- Laplace-Entwicklung: Aufwand etwa $n!$: Man erhält n Determinanten von $(n-1)\times(n-1)$-Matrizen
- pro $(n-1)\times(n-1)$-Matrix: Summe von $(n-1)$ Determinanten von $(n-2)\times(n-2)$-Matrizen
- usw.
- Leibnizsche Formel: $n!$ Summanden mit jeweils n Faktoren, also $n-1$ Multiplikationen pro Summen

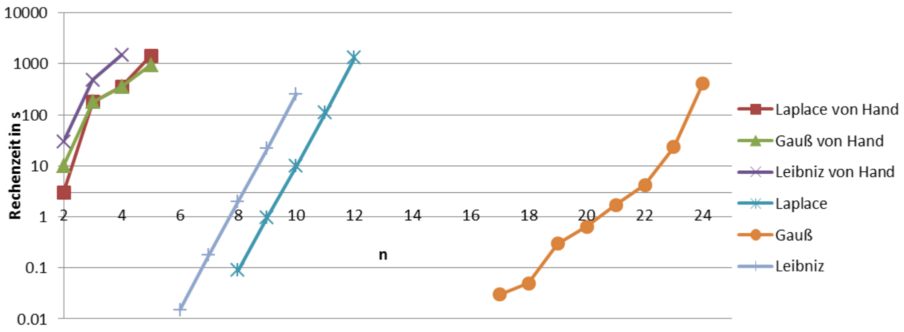
- also insgesamt $(n-1)n!$ arithmetische Operationen nötig.
- betrachten das Wachstumsverhalten der drei Verfahren abhängig von n
- Rechenzeit t
- Anzahl von arithmetischen Operationen: Ops
- Annahme: Computer schafft 10^9 arithmetische Operationen pro Sekunde

n	5	10	15	20	50	100
Ops t	40 < 1 ms	330 < 1 ms	$1,1$E3 < 1 ms	$2,7$E3 < 1 ms	$4,2$E4 < 1 ms	$3,3$E5 < 1 ms
Ops t	120 < 1 ms	$3,6$E6 $3,6$ ms	$1,3$E12 22 m	$2,4$E18 77 a	$3,0$E64 $> 10^{12}$ a	$9,3$E157 $> 10^{12}$ a
Ops t	480 < 1 ms	$3,3$E7 33 ms	$1,8$E13 $5,1$ h	$4,6$E19 $1,4$E3 a	$1,5$E66 $> 10^{12}$ a	$9,2$E159 $> 10^{12}$ a

- für Laplace-Entwicklung und Leibniz-Formel: Anstieg der Rechenzeit verheerend
- nur das Gauß-Verfahren geeignet zur Berechnung von Determinanten größerer Matrizen

Beispiel 5.29:

- realistischer Vergleich: Implementierung der drei Verfahren in python
- absolute Rechenzeiten nicht repräsentativ
- aber: Vergleich der Rechenzeiten ist repräsentativ
- zusätzlich: Vergleich mit Zeiten für Handrechnung
- MATSE-Studentin musste rechnen
- logarithmische Skalierung im Diagramm


10000
1000
100
10
1
0.1
0.01
Rechenzeit in s
2
4
6
8
10
12
14
16
18
20
22
24
n
Laplace von Hand
Gauß von Hand
Leibniz von Hand
Laplace
Gauß
Leibniz

5.5 Exkurs: Invertierung von Matrizen mittels Unterdeterminanten

- jetzt: Methode zu Invertierung eine Matrix
- Basis: Unterdeterminanten

Definition 5.30: Unter der zu A *komplementären Matrix* $\tilde{A}$ versteht man die Matrix mit den Elementen

$$\tilde{a}_{ij} = (-1)^{i+j} \cdot \det(A_{ji}) \, .$$

Satz 5.31: Sei $A \in GL(n;K)$. Dann gilt

$$A^{-1} = \frac{1}{\det(A)} \tilde{A} \, .$$

Für den Beweis verweisen wir auf z. B. Fischer [2, Kap.3.3].

Beispiel 5.32: *OHP*

- Anwendung: explizite Formel der Inversen einer 2 × 2-Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{5.1}$$

- obiges Rechenschema:

$$B = \begin{pmatrix} d & c \\ b & a \end{pmatrix} \Rightarrow C = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = \tilde{A}^T \Rightarrow \tilde{A} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Satz 5.33: Sei A wie in Formel (5.1) und invertierbar. Dann gilt

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Beispiel 5.34: *OHP*

- In der Praxis: Gauß-Verfahren ist mindestens ebenbürtig
- Das neue Verfahren hat interessante Konsequenzen.

1. Ist man in der Lage, eine Determinante zu berechnen, ist man auch fähig, ein lineares Gleichungssystem zu lösen.
2. Das Gauß-Verfahren ist mit allen seinen Varianten nicht die einzige Verfahrensklasse zum Lösen von linearen Gleichungssystemen, es existieren auch grundsätzlich andersartige Verfahren.

- dritte Konsequenz
- in der Praxis: A manchmal nicht genau bekannt
- z. B.: Koeffizienten von A wurden mithilfe von Messwerten berechnet
- Messungen sind immer ungenau.

- stattdessen verfügbar: gestörte Matrix $\hat{A}$
- Ist es überhaupt sinnvoll, ein gestörtes Gleichungssystem zu lösen?
- Grund: Man wird eine ebenfalls gestörte Lösung $\hat{x}$ erhalten.
- Antwort: Ja, wenn man $\|x-\hat{x}\| < \varepsilon$ garantieren könnte.
- Anwender gibt Fehlerschranke ε je nach erforderlicher Genauigkeit vor
- ε wird implizit festlegen, wie genau A bekannt sein muss
- Problem: Winzigste Störungen in A könnten zu erheblichen Abweichungen in x führten.
- in diesem Fall: seriöse Berechnung unmöglich
- Annahme: b bekannt
- Es gilt $x = A^{-1}b$

- $x - \hat{x} = (A^{-1} - \hat{A}^{-1})b$
- prinzipelle Berechenbarkeit dann gegeben, wenn $\|A^{-1} - \hat{A}^{-1}\|$ beliebig klein wird, wenn man man nur $\|A - \hat{A}\|$ klein genug wählt.

Satz 5.35:

1. Die Determinante $\det : \mathbb{C}^{n \times n} \to \mathbb{C}$ ist stetig.
2. Sei $A \in GL(n, \mathbb{C})$ und seien alle Matrizen in einer Umgebung von A ebenfalls regulär. Dann ist $\Psi : GL(n; \mathbb{C}) \to GL(n; \mathbb{C}), \Psi(B) = B^{-1}$ stetig im Punkt A, und es gilt

$$\forall \varepsilon > 0 \, \exists \delta > 0 : \|A - \hat{A}\| < \delta \Rightarrow \|A^{-1} - \hat{A}^{-1}\| < \varepsilon \, . \tag{5.2}$$

Beweis *OHP*

Kapitel 6

Lineare Gleichungssysteme

6.1 Lösbarkeit eines linearen Gleichungssystems

- zuerst: Definition eines allgemeinen linearen Gleichungssystems

Definition 6.1: Seien $A = (a_{ij}) \in K^{m\times n}$ und $b = (b_1, \ldots, b_m)^T \in K^m$. Dann heißt

$$\begin{array}{lcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ \vdots & & \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b_m \end{array}$$

lineares Gleichungssystem bzgl. $(x_1, \ldots, x_n)$ mit Koeffizienten a_{ij} in K. Hierbei sind $x_1, \ldots, x_n$ die *Unbekannten* des Systems. Für $b = 0$ nennt man das lineare Gleichungssystem *homogen*, sonst *inhomogen*.

Bemerkung 6.2: Jedes lineare Gleichungssystem kann in der Form $Ax = b$ geschrieben werden.

Definition 6.3: In der Situation von Definition 6.1 ist die *Lösungsmenge* $L(A,b)$ des zu (A,b) gehörigen Gleichungssystems festgelegt durch

$$L(A,b) := \{x \in K^n | Ax = b\} .$$

- jetzt: einfaches Kriterium zur Lösbarkeit eines beliebigen LGS
- als Vorbereitung dazu: Definition des Ranges einer Matrix

Definition 6.4: Die lineare Abbildung $L_A : K^n \to K^m$ sei gegeben durch $L_A(x) := Ax$. Dann sei $\mathrm{rg}(A) := \mathrm{rg}(L_A)$. Der *Spaltenrang* $\mathrm{rg}_S(A)$ sei die maximale Anzahl linear unabhängiger Spaltenvektoren von A.

Satz 6.5: Es gilt $\mathrm{rg}(A) = \mathrm{rg}_S(A)$.

Beweis *OHP*

Satz 6.6: Das lineare Gleichungssystem $Ax = b$ ist genau dann lösbar, wenn $\mathrm{rg}(a_1, \ldots, a_n) = \mathrm{rg}(a_1, \ldots, a_n, b)$ gilt. Kürzer schreibt man

$$\mathrm{rg}(A) = \mathrm{rg}(A, b)\,.$$

Beweis *OHP*

Satz 6.7: Sei $x_s \in K^n$ Lösung von $Ax = b$. Dann gilt

$$L(A,b) = x_s + \ker(A) = \{x_s + x \mid x \in \ker(A)\}\,.$$

Beweis *OHP*

Bemerkung 6.8: Sei x_s eine Lösung von $Ax = b$ und $(v_1, \ldots, v_r)$ eine Basis von $\ker(A)$. Dann gilt

$$L(A,b) = \{x_s + \lambda_1 v_1 + \ldots + \lambda_r v_r |\, \lambda_i \in K\}\,.$$

Wegen der Dimensionsformel Satz 4.23 gilt $r = \dim(\ker(A)) = n - \mathrm{rg}(A)$.

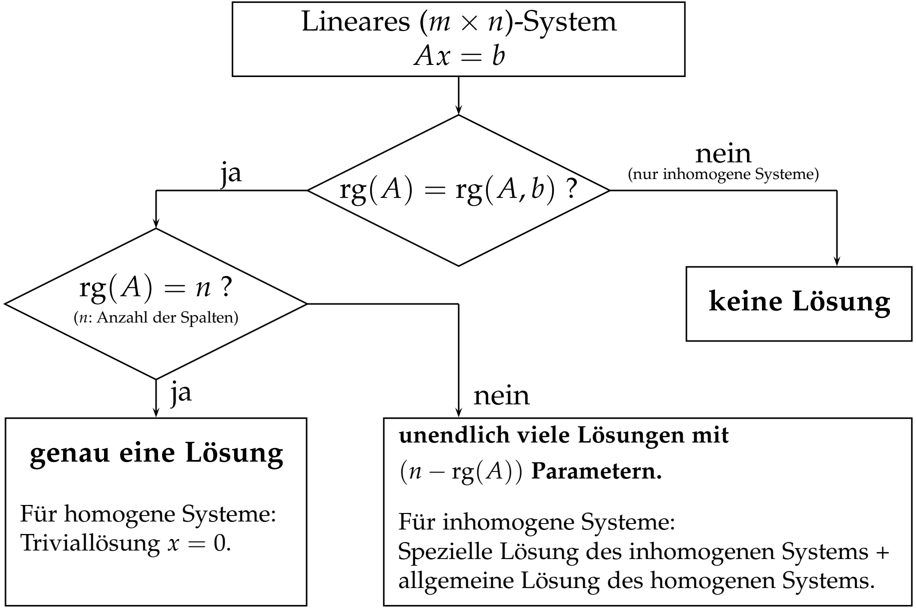
- gelegentlich: konkreter Lösungsvektor: *spezielle Lösung*, Lösungsvektor mit frei wählbaren Parametern: allgemeine Lösung

- Damit gilt folgender Merksatz:

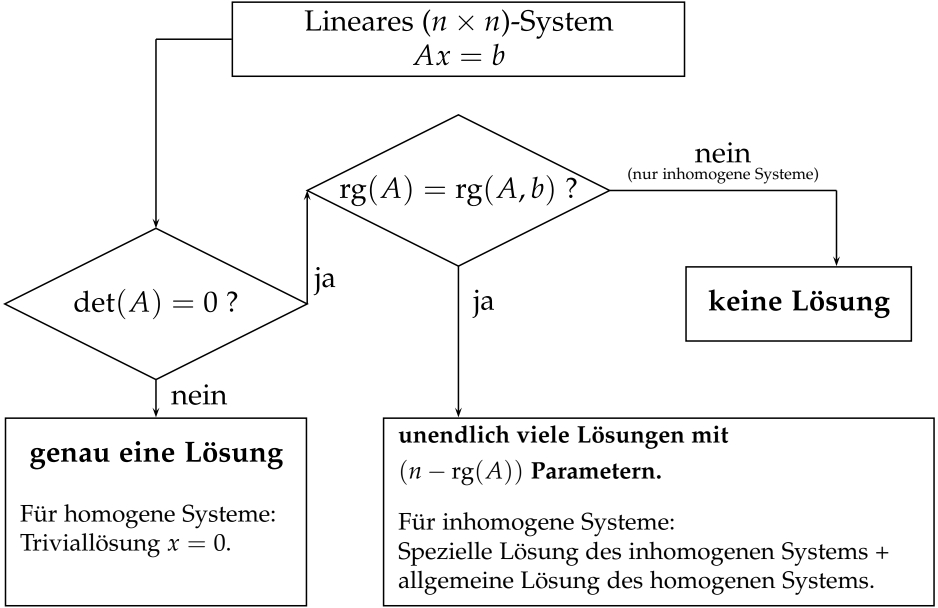
- Ist x_s eine spezielle Lösung von $Ax = b$ und ist x_h die *allgemeine Lösung* von $Ax = 0$, dann ist $x_a = x_h + x_s$ die *allgemeine* Lösung von $Ax = b$.

Folgerung 6.9: Das lineare Gleichungsssystem $Ax = b$ sei lösbar. Es ist genau dann eindeutig lösbar, falls $\ker(A) = \{0\}$, d. h. $\mathrm{rg}(A) = n$ gilt, also genau dann, wenn $\mathrm{rg}(A) = \mathrm{rg}(A,b) = n$ ist.

Folgerung 6.10: Sei die Abbildung L_A gegeben durch $L_A(x) = Ax$. Dann ist $Ax = b$ genau dann lösbar, wenn $b \in \text{Bild}(L_A)$ gilt.


Lineares $(m \times n)$-System
$Ax = b$
$\mathrm{rg}(A) = \mathrm{rg}(A, b)$?
ja
nein
(nur inhomogene Systeme)
keine Lösung
$\mathrm{rg}(A) = n$?
(n: Anzahl der Spalten)
ja
nein
genau eine Lösung
Für homogene Systeme:
Triviallösung $x = 0$.
unendlich viele Lösungen mit
$(n - \mathrm{rg}(A))$ Parametern.
Für inhomogene Systeme:
Spezielle Lösung des inhomogenen Systems +
allgemeine Lösung des homogenen Systems.

Bemerkung 6.11: Für quadratisches A ist das Gleichungssystem $Ax = b$ genau dann eindeutig lösbar, wenn $|A| \neq 0$ gilt.


Lineares $(n \times n)$-System
$Ax = b$
$\det(A) = 0$?
ja
nein
$\mathrm{rg}(A) = \mathrm{rg}(A, b)$?
nein
(nur inhomogene Systeme)
ja
keine Lösung
genau eine Lösung
Für homogene Systeme:
Triviallösung $x = 0$.
unendlich viele Lösungen mit
$(n - \mathrm{rg}(A))$ Parametern.
Für inhomogene Systeme:
Spezielle Lösung des inhomogenen Systems +
allgemeine Lösung des homogenen Systems.

Beispiel 6.12:

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}; \qquad b = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

- $n = 2$
- $\det(A) = 5$
- $\mathrm{rg}(A) = \mathrm{rg}(A, b) = 2$
- es gibt es eine eindeutige Lösung

Beispiel 6.13:

$$A = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 3 & 1 \end{pmatrix}; \quad b = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

- $n = 3$

- $\mathrm{rg}(A) = \mathrm{rg}(A, b) = 2$
- es gibt unendlich viele Lösungen

Beispiel 6.14:

$$A = \begin{pmatrix} 1 & -1 & 2 \\ 2 & -2 & 4 \end{pmatrix}; \quad b = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

- $n = 3$
- $\mathrm{rg}(A) = 1$
- $\mathrm{rg}(A, b) = 2$
- es existiert keine Lösung

Beispiel 6.15:

$$A = \begin{pmatrix} 1 & 3 \\ -1 & 1 \\ 0 & 1 \end{pmatrix}; \quad b = \begin{pmatrix} -2 \\ -2 \\ -1 \end{pmatrix}.$$

OHP

Bemerkung 6.16: Man muss zwei verschiedene Situationen klar trennen:

1. Ein Gleichungssystem $Ax = b$ mit $A \in K^{n\times n}$ ist genau dann eindeutig lösbar, wenn $\det(A) \neq 0$.

2. Sei $Ax = b$ gegeben mit

   (a) $A \in K^{(n+1)\times n}$ (eine Zeile mehr als Spalten) und

   (b) die Spalten von A sind linear unabhängig.

   Dann ist $Ax = b$ genau dann eindeutig lösbar, wenn $\det(A, b) = 0$ ist.

Beispiel 6.17: *OHP*

- jetzt: (beinahe) sämtliche Bedingungen für die eindeutige Lösbarkeit eines linearen Gleichungssystems

Satz 6.18: Sei $K \in \{\mathbb{R}, \mathbb{C}\}$. Für $A \in K^{n\times n}$ und die dadurch gegebene lineare Abbildung L_A sind folgende Bedingungen äquivalent:

1. A ist invertierbar.
2. $Ax = 0$ hat nur die triviale Lösung $x = 0$.
3. Durch Zeilen- und Spaltenumformungen kann A auf die Einheitsmatrix transformiert werden.
4. A ist darstellbar als Produkt von Elementarmatrizen.
5. $Ax = b$ besitzt für jedes $b \in K^n$ mindestens eine Lösung.
6. $Ax = b$ hat genau eine Lösung für jedes $b \in K^n$.

7. $\det(A) \neq 0$.
8. $\text{Bild}(L_A) = K^n$.
9. L_A ist bijektiv.
10. Die Spaltenvektoren von A sind linear unabhängig.
11. Die Zeilenvektoren von A sind linear unabhängig.
12. Die Spaltenvektoren von A bilden eine Basis von K^n.
13. Die Zeilenvektoren von A bilden eine Basis von K^n.
14. $\text{rg}(A) = n$.
15. $\ker(L_A) = \{0\}$.
16. $(\ker(L_A))^\perp = K^n$.

17. Das orthogonale Komplement des von den Zeilen von A aufgespannten Raums ist $\{0\}$.

18. $A^T A$ ist invertierbar.

6.2 Das Gaußsche Eliminationsverfahren reloaded

- bisher: Gauß-Verfahren wurde in vielfältiger Weise angewendet
- aber: Beweis der Korrektheit des Verfahrens fehlt
- Das wird jetzt nachgeholt.
- Sei $A \in K^{n \times n}$ invertierbar
- Der Gauß-Algorithmus besteht aus zwei Schritten

1. - forme (A, b) mit Z1 und Z2 um
   - Ziel: $(\tilde{A}, \tilde{b})$ mit oberer Dreicksmatrix $\tilde{A}$
2. - Man löst das System $\tilde{A}x = \tilde{b}$ durch Rückwärtseinsetzen
   - Lösung x erfüllt zugleich $Ax = b$

- nächster Satz: Das Gauß-Verfahren „erfindet oder unterschlägt“ keine Lösungen
- gilt für beliebige lineare Gleichungssysteme

Satz 6.19: Die Zeilenumformungen Z1 – Z3 des Gauß-Verfahrens erhalten die Lösungsmenge.

Beweis *OHP*

- z. z. bleibt, dass das Gauß-Verfahren für reguläre Matrizen immer durchführbar ist
- zunächst:

Satz 6.20: Die Zeilen- bzw. Spaltenoperationen des Gauß-Verfahrens erhalten den Rang einer Matrix.

Beweis *OHP*

- der Einfachheit halber sei A im folgenden regulär

Satz 6.21: Sei $A \in K^{n\times n}$ invertierbar. Dann lässt sich A durch eine Folge von Operationen vom Typ Z1 oder Z2 in eine obere Dreiecksmatrix umformen, wobei kein Hauptdiagonalelement den Wert 0 annimmt.

Beweis *OHP*

- allgemeiner gilt folgender Satz
- Ansatz ähnlich, nur technisch deutlich komplizierter

Satz 6.22: Sei $A \in K^{m \times n}$ mit $\mathrm{rg}(A) = r$. Durch eine Folge von Zeilen- und Spaltenoperationen der Form Z1 – Z3 bzw. S1 – S3 kann man die Matrix A in eine Matrix der Form

$$E_r = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

umformen. Es gilt $a_{ii} = 1$ für $i \leq r$.

- neben *Spaltenrang* einer Matrix: *Zeilenrang* gebräuchlich

Definition 6.23: Für $A \in K^{m \times n}$ sei die maximale Anzahl linear unabhängiger Zeilenvektoren von A der *Zeilenrang* $\mathrm{rg}_Z(A)$ von A.

- Anwendung der Theorie zum Gauß-Verfahren:

Satz 6.24 (Zeilenrang = Spaltenrang): Für $A \in K^{m \times n}$ gilt $\mathrm{rg}(A) = \mathrm{rg}_S(A) = \mathrm{rg}_Z(A)$.

Beweis *OHP*

- obiger Satz liefert ein Rechenverfahren zur Rangbestimmung
- Man bringt die Matrix mittels Zeilen- und Spaltenumformungen auf die Gestalt E_r
- dann: Rang als Anzahl der Nicht-Null-Einträge ablesen

Beispiel 6.25: *OHP*

- alternatives Rechenverfahren zur Rangbestimmung von Matrizen

Bemerkung 6.26: Es sei $A \in K^{m \times n}$ und r die Zeilen/Spaltenanzahl der größten $(r \times r)$– Untermatrix von A mit Determinante ungleich Null. Dann ist $\mathrm{rg}(A) = r$.

Beispiel 6.27:

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

OHP

6.3 Die Cramersche Regel

- in diesem Kapitel: Formel zur Lösung eines linearen Gleichungssystems mit Determinanten

Satz 6.28 (Cramersche Regel): Es seien $A = (a_1, \ldots, a_n) \in K^{n \times n}$ und $x, b \in K^n$ sowie $Ax = b$ ein lineares Gleichungssystem, und es gelte $\det(A) \neq 0$. Seien

$$A_i := (a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n), i = 1, \ldots, n\,.$$

Dann gilt

$$x_i = \frac{\det(A_i)}{\det(A)}, i = 1, \ldots, n\,.$$

Beweis *OHP*

Beispiel 6.29:

$$\begin{array}{cc|c} 1 & 2 & 5 \\ 3 & 2 & 11 \end{array}$$

OHP

Beispiel 6.30:

- Cramersche Regel bei Handrechnung manchmal vorteilhaft
- Grund: keine aufwändigen Divisionen innerhalb der Berechnung wie bei Gauß

$$\begin{array}{rrr|r} 1 & 2 & 7 & -2 \\ -3 & 1 & -4 & 1 \\ 5 & 11 & 5 & -1 \end{array} .$$

OHP

6.4 Überbestimmte lineare Gleichungssysteme

Allen recht getan ist eine Kunst, die niemand kann.

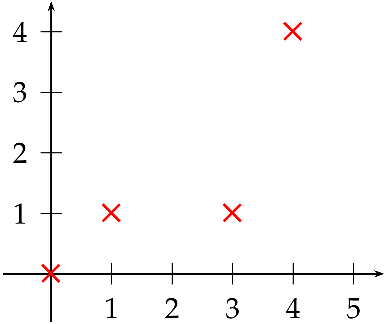
(alte Volksweisheit)

- in der Praxis häufig: Datenpunkte gegeben
- Datenpunkte repräsentieren Messwerte, wirtschaftliche Größen etc.
- Man nimmt einen linearen Zusammenhang an
- aber: aufgrund z. B. von Messfehlern liegen die Punkte nicht auf einer Geraden
- daher: versuche, die „am besten passende" Gerade zu finden
- „passen": Summe der Quadrate der Abweichungen der Kurve von den gegebenen Punkten wird minimiert (Fitting)

- eventuell trifft die Ausgleichsgerade keinen der Punkte exakt
- analoge Überlegungen gelten auch für beliebige funktionale Zusammenhänge
- im Fall der Geraden: *lineare Regression*

Beispiel 6.31:

- gegeben $(0;0),(1;1),(3;1),(4;4)$
- Ziel: Ausgleichsgerade berechnen



- Ansatz: $g(x) = \alpha x + \beta$

- Einsetzen der Punkte führt zu

$$\begin{aligned}
\alpha \cdot 0 + \beta &= 0 \\
\alpha \cdot 1 + \beta &= 1 \\
\alpha \cdot 3 + \beta &= 1 \\
\alpha \cdot 4 + \beta &= 4
\end{aligned}$$

- in Matrixform:

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 3 & 1 \\ 4 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 4 \end{pmatrix}.$$

- Die Punkte liegen nicht auf einer Geraden
- Konsequenz: Das lineare Gleichungssystem ist nicht lösbar
- in der Praxis: Eine Lösung muss her!

- Konsequenz: Wir müssen unseren Lösungsbegriff überdenken!
- Ziel: neuer Lösungsbegriff + Rechenverfahren für obigen Fall
- sei $Ax = b$ überbestimmt
- nur lösbar, wenn $b \in \mathrm{Bild}(L_A)$
- i. A. nicht gegeben
- Idee: Generiere (Ersatz)lösung, indem man b durch eine rechte Seite aus $\mathrm{Bild}(L_A)$ ersetzt
- Ersatz muss b möglichst nahe kommen
- Bekannt: $\mathrm{Bild}(L_A)$ ist ein Untervektorraum
- dann: die *orthogonale Projektion* $p_A(b)$ von b auf $\mathrm{Bild}(L_A)$ die bestmögliche Wahl

- bekannt: Liegt eine ON-Basis eines Unterraums vor, kann man die orthogonale Projektion darauf leicht berechnen

- jetzt: nur irgendeine Basis bekannt

Bemerkung 6.32: *OHP*

Satz 6.33: Die Normalgleichungen sind für jede reelle Matrix $A \in \mathbb{R}^{m \times n}$ lösbar. Im Falle $\operatorname{rg}(A) = n$ existiert mit

$$x = (A^T A)^{-1} A^T b \tag{6.4}$$

sogar eine eindeutige Lösung. In diesem Fall heißt $(A^T A)^{-1} A^T$ *verallgemeinerte Inverse* von A.

Beweis *OHP*

Der folgende Satz rechtfertigt den Ansatz, bei einem überbestimmten linearen Gleichungssystem $Ax = b$ das b durch $p_A(b)$ zu ersetzen.

Satz 6.34 (Methode der kleinsten Quadrate): Gegeben ist das Gleichungssystem

$$Ax = b, \quad A \in \mathbb{R}^{m \times n}, \, b \in \mathbb{R}^m, \, m \geq n\,.$$

Im Fall $\mathrm{rg}(A) = n$ gilt für

$$x_s = (A^T A)^{-1} A^T b,$$

dass

$$\|b - Ax_s\| = \min_{z \in \mathbb{R}^n} \|b - Az\|\,.$$

Der Vektor x_s heißt *Näherungslösung nach der Methode der kleinsten Quadrate*.

Beweis *OHP*

Gelegentlich spricht man von einer klassischen Lösung eines linearen Gleichungssystems, wenn $Ax = b$ erfüllt ist in Abgrenzung von einer Lösung nach der Methode der kleinsten Quadrate.

Beispiel 6.35: Das Gleichungssystem

$$\begin{aligned} x &= 1 \\ x &= -1 \end{aligned}$$

ist überbestimmt und besitzt keine klassische Lösung. *OHP*

Beispiel 6.36: Das Gleichungssystem

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

ist überbestimmt und im klassischen Sinn unlösbar.

OHP

- Wir kommen zur Ausgleichsrechnung zurück
- Spezialfall: euklidische Norm
- dann: $\|Ax-b\|^2$ genau der Summe der Quadrate der Abweichungen der Ausgleichsgerade zu den Datenpunkten
- minimal, wenn α und β mithilfe der Normalengleichungen bestimmt werden
- deswegen: „Methode der kleinsten Fehlerquadrate“
- Berechnung der Ausgleichsgeraden durch $A^TA(\alpha,\beta)^T = A^Tb$

OHP

- altenativ: Ausgleichsparabel statt Ausgleichsgerade („quadratischer Fit“)
- Ansatz: $q(x) = \alpha x^2 + \beta x + \gamma$
- Man erhält dieses lineare Gleichungssystem

$$\begin{aligned}
\alpha \cdot 0^2 + \beta \cdot 0 + \gamma &= 0 \\
\alpha \cdot 1^2 + \beta \cdot 1 + \gamma &= 1 \\
\alpha \cdot 3^2 + \beta \cdot 3 + \gamma &= 1, \quad \text{bzw.} \\
\alpha \cdot 4^2 + \beta \cdot 4 + \gamma &= 4
\end{aligned}$$

- in Matrixform:

$$\begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 9 & 3 & 1 \\ 16 & 4 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 4 \end{pmatrix},$$

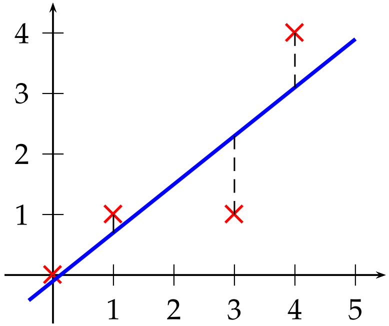


Abbildung 6.1: Ausgleichsgerade durch gegebene Datenpunkte

- Ergebnis: $p(x) = {}^1\!/_3 x^2 - {}^8\!/_{15} x + {}^2\!/_5$

Bemerkung 6.37: *OHP*

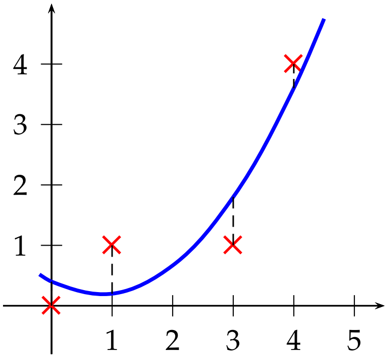


Abbildung 6.2: Ausgleichsparabel durch gegebene Datenpunkte

6.5 Unterbestimmte lineare Gleichungssysteme

- Motivation: Ausgleichsfunktion mit Polynom p
- jetzt: $\deg(p) \geq n$
- gegeben: $(-1,1)$, $(0,1)$ und $(1,1)$, $p(x) = a + bx + cx^2 + dx^3$
- keine eindeutige Lösung, viele Polynome erfüllen die Interpolationsaufgabe
- u. a. $p_1(x) = 1$, aber auch $p_2(x) = x^3 - x + 1$
- in der Praxis: man wählt wohl p_1
- genauere Analyse: lineares Gleichungssystem der Interpolationsaufgabe untersuchen

$$\begin{pmatrix} 1 & -1 & 1 & -1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} .$$

- Koeffizientenvektor von p_1: $v_1 = (1, 0, 0, 0)^T$
- Koeffizientenvektor von p_2: $v_2 = (1, -1, 0, 1)^T$
- beide Vektoren erfüllen das LGS
- aber: v_1 besitzt von allen Lösungsvektoren die kleinstmögliche Norm
- Grund: wegen des zweiten Punktes muss $a = 1$ sein
- die restlichen Komponenten sind bereits 0

Satz 6.38: Gegeben ist ein lineares Gleichungssystem $Ax = b$ mit $A \in \mathbb{R}^{m \times n}, x \in \mathbb{R}^n, b \in \mathbb{R}^m$ und $m \leq n$. Für $\operatorname{rg}(A) = m$ löst

$$x_s = A^T(AA^T)^{-1}b$$

das Gleichungssystem $Ax = b$, und es gilt:

1. x_s ist die Lösung mit der kleinsten Norm aller Lösungen x, d. h.

$$\|x_s\| = \min_{Ax=b}\{\|x\|\}.$$

2. x_s steht orthogonal zu sämtlichen Lösungen z von $A\,z = 0$, d. h.

$$x_s \perp \ker(A)$$

Für $\operatorname{rg}(A) = m$ heißt $A^T(AA^T)^{-1}$ die *verallgemeinerte Inverse* von A.

Beweis *OHP*

Beispiel 6.39:

$$2x_1 + 2x_2 - x_3 = 6$$

OHP

Kapitel 7

Geometrie linearer Abbildungen

- „Matrix“ und „Lineare Abbildung“ zentral in der Linearen Algebra
- bekannt: Matrix A induziert eine lineare Abbildung f_A und umgekehrt
- aber: Sei A gegeben. Was genau „macht f_A mit den Vektoren“?
- Ziel: Herleitung geometrischer Eigenschaften von f_A aus algebraischen Eigenschaften von A und umgekehrt
- ab jetzt immer: euklidische Norm, euklidisches Skalarprodukt

7.1 Orthogonale Abbildungen und Matrizen

- Sei D eine Drehung in der Ebene um 0
- Es gilt $\|D(x)\| = \|x\| \; \forall x \in \mathbb{R}^2$
- allgemeiner: Ein f mit $\|f(x) - f(y)\| = \|x - y\| \; \forall x, y$ heißt *Isometrie*
- Dreht man zwei Vektoren auf die gleiche Weise, ändert sich ihre Lage zueinander nicht
- Folge: rechte Winkel bleiben erhalten
- weitere Folge: Die Standardbasis als Orthonormalbasis wird auf eine Orthonormalbasis abgebildet
- bekannt: Drehungen sind lineare Abbildungen mit Darstellungsmatrix A
- Diese Überlegungen gelten allgemein für lineare winkelerhaltende Isometrien (z. B. auch Spiegelungen)

Definition 7.1:

1. Eine Matrix $A \in \mathbb{R}^{n\times n}$ heißt *orthogonal*, wenn ihre Spaltenvektoren eine Orthonormalbasis bilden.
2. Die Menge aller orthogonalen Matrizen in $\mathbb{R}^{n\times n}$ heiße $O(n)$.

- Achtung: Eine Ortho*gonal*matrix liegt vor bei einer Ortho*normal*basis in den Spalten
- Ortho*gonal*basis genügt nicht

Beispiel 7.2:

1. Die Einheitsmatrix E liegt in $O(n)$.
2. - In $\mathbb{R}^2$: Drehung um α hat bzgl. der Standardbasis die Darstellung

$$A_\alpha = (a_1\, a_2) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

- Wg. $\cos^2\alpha + \sin^2\alpha = 1$: Matrixspalten haben Länge 1
- weiter: $\langle a_1, a_2 \rangle = 0$
- somit $A_\alpha \in O(2)$

3.
- Eine Spiegelung eines Vektors in $\mathbb{R}^2$ an der x-Achse ist eine lineare Abbildung
- Darstellungsmatrix:

$$S_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

- $S_x \in O(2)$

Satz 7.3: Für $A \in \mathbb{R}^{n \times n}$ sind äquivalent:

1. $A \in O(n)$
2. A ist invertierbar, und es gilt $A^{-1} = A^T$.

3. $A^T \in O(n)$

Beweis *OHP*

Satz 7.4: $O(n)$ bildet mit der Matrizenmultiplikation die sog. *orthogonale Gruppe*.

Beweis *OHP*

- zurück zu den Betrachtungen am Anfang
- zusammengefasst gilt: Jede winkelerhaltende lineare Isometrie wird bezogen auf die Standardbasis durch eine orthogonale Matrix dargestellt.
- jetzt: Umkehrung?

Bemerkung 7.5: Sei $f : \mathbb{R}^n \to \mathbb{R}^n$ beliebig. Dann sind äquivalent:

1. $$\langle f(x), f(y)\rangle = \langle x, y\rangle \; \forall x, y \in \mathbb{R}^n \tag{7.1}$$
2. f ist eine winkel- und längenerhaltende Isometrie.

Beweis *OHP*

- offenbar: Eigenschaft (7.1) maßgeblich
- wir definieren daher:

Definition 7.6: Eine Abbildung $f : \mathbb{R}^n \to \mathbb{R}^n$ heißt *orthogonal*, wenn (7.1) gilt.

- Zu Anfang: Wir haben angenommen, dass alle Drehungen und Spiegelungen linear seien.
- Das stimmt auch.

Satz 7.7: Sei $f : \mathbb{R}^n \to \mathbb{R}^n$ orthogonal. Dann ist f linear.

Beweis *OHP*

- allgemeine Isometrien sind nicht immer linear
- z. B.: $r : \mathbb{R} \to \mathbb{R}, r(x) = x + 1$.

Satz 7.8: Die Abbildung $f : \mathbb{R}^n \to \mathbb{R}^n$ ist genau dann orthogonal, wenn die zugehörige Darstellungsmatrix Q (bzgl. der Standardbasis) eine Orthogonalmatrix ist, also $Q \in O(n)$.

Beweis *OHP*

- A als Darstellungsmatrix von f hängt von der Wahl der Basis ab
- geometrische Eigenschaften von f natürlich nicht
- Folge: Algebraische Eigenschaften von A dürfen, wenn sie geometrische Bedeutung haben sollen, nicht von der Wahl der Basis abhängen.
- Orthogonalität einer Matrix bei Basiswechsel?

Bemerkung 7.9: Sei f orthogonal. Dann ist die Darstellungsmatrix von f bezogen auf *jede* Orthonormalbasis orthogonal.

Beweis *OHP*

Bemerkung 7.10:

- Bemerkung 7.9 ist für allgemeine Basen falsch
- betrachte

$$A = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

- Sei $\mathcal{B}$ die Basis, die aus den Spalten von B besteht
- sei $A = M^{\mathcal{E}}_{\mathcal{E}}(f)$
- dann:

$$M^{\mathcal{B}}_{\mathcal{B}} = BAB^{-1} = \frac{1}{\sqrt{2}}\begin{pmatrix} 3 & -7 \\ 1 & -3 \end{pmatrix},$$

- $M_{\mathcal{B}}^{\mathcal{B}}$ ist offenbar nicht orthogonal
- jetzt: Volumina unter orthogonalen Abbildungen

Bemerkung 7.11: Für $A \in O(n)$ gilt $|\det(A)| = 1$.

Beweis *OHP*

Bemerkung 7.12:

- aber: Nicht jede Matrix A mit $\det(A) = 1$ ist eine Orthogonalmatrix

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

- Verallgemeinerung der Substitutionsregel für eindimensionale Integrale: Transformationsformel (siehe Analysis)
- $\Omega \in \mathbb{R}^n$ glatt berandetes beschränktes Gebiet
- $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ stetig differenzierbar mit stetig differenzierbarer Umkehrfunktion
- weiter $f : \Omega \to \varphi(\Omega)$
- dann:

$$\int_{\Omega} (f \circ \varphi)(x) |\det(J\varphi(x))| \, dx = \int_{\varphi(\Omega)} f(z) \, dz$$

- $J\varphi$: Jacobi-Matrix von φ
- es gilt $m(\Omega) = \int_{\Omega} 1 dx$, analog $m(Q(\Omega))$
- setze also $\varphi(x) = Qx$ mit einer orthogonalen Matrix und $f = 1$
- man errechnet $J\varphi(x) = Q$
- daher:

$$m(Q(\Omega)) = \int_{Q(\Omega)} 1 \, dz = \int_{\Omega} 1 |\det(Q)| \, dx = m(\Omega).$$

- Winkelerhaltende Isometrien erhalten Volumina.
- winkelerhaltende Isometrien erhalten i. A. die Orientierung nicht

- Beispiel: Spiegelung S (orthogonal!) transformiert Links- in Rechtssystem und umgekehrt
- Veranschaulichung: Spiegelschrift
- $\det(S) = -1$ (vgl. Bemerkung 7.11)
- Drehungen D erhalten die Orientierung, also $\det(D) = 1$
- es gilt $\det(E) = 1$ und $\det(AB) = \det(A)\det(B)$
- Konsequenz:

Bemerkung 7.13: Die *spezielle orthogonale Gruppe*

$$SO(n) := \{Q \in O(n) \mid \det(Q) = 1\}$$

bildet eine Untergruppe von $O(n)$. Die Darstellungsmatrix einer beliebigen Drehung um 0 (bzgl. einer Orthonormalbasis) liegt in $SO(n)$, die einer Spiegelung in $O(n) \setminus SO(n)$.

- Zusammenfassung:
- Eine orthogonale Abbildung ist linear und erhält Abstände, Winkel und Volumina
- Ihre Darstellungsmatrix bezogen auf eine beliebige Orthonormalbasis ist orthogonal.
- umgekehrt: Jede Orthogonalmatrix induziert bezogen auf eine Orthonormalbasis eine orthogonale Abbildung.

7.2 Exkurs: QR-Zerlegung und Anwendungen

- orthogonale Matrizen haben weitergehende Anwendungen
- betrachte $Ax = b$ mit $A \in \mathbb{R}^{m \times n}$, $m \geq n$ und $\mathrm{rg}(A) = n$
- Annahme: es gebe eine QR-Zerlegung

$$A = QR$$

- dabei: $R \in \mathbb{R}^{n \times n}$ obere Dreicksmatrix
- $Q = (q_1, \ldots, q_n) \in \mathbb{R}^{m \times n}$ mit $\langle q_i, q_j \rangle = \delta_{ij}$
- sei $x_S = (A^T A)^{-1} A^T b$ die Lösung im Sinne der kleinsten Quadrate
- man hat $\mathrm{rg}(Q) = n = \mathrm{rg}(A)$
- aus Satz 4.55 (Rangerhaltung): $\mathrm{rg}(R) = n$

- R ist somit invertierbar
- Einsetzen von $A = QR$ in die Formel für x_S:

OHP

- x_S ist also zugleich Lösung von

$$Rx_S = Q^T b \tag{7.2}$$

- Berechnung des Ausgleichsproblems wurde zurückgeführt auf die (klassische) Lösung eines linearen Gleichungssystems mit einer Dreiecksmatrix
- statt Gauß-Elimination: Rückwärtseinsetzen
- der „klassische“ Fall $m = n$ ist ein Spezialfall
- Konsequenz: QR-Zerlegung ist eine Alternative zur Gauß-Elimination bei der Lösung linearer Gleichungssysteme
- Wie erhält man eine solche QR-Zerlegung?

Satz 7.14: Sei $A = (a_1, \dots, a_n) \in \mathbb{R}^{m\times n}$ mit $m \geq n$ und $\text{rg}(A) = n$. Dann gibt es eine in den Spalten orthogonale Matrix $Q \in \mathbb{R}^{m\times n}$ und eine obere Dreiecksmatrix $R \in \mathbb{R}^{n\times n}$ mit

$A = QR$. Hierbei können die Spalten von Q mithilfe des Verfahrens von Gram-Schmidt aus den Spalten von A erzeugt werden, und es gilt $\operatorname{rg}(R) = n$.

Beweis *OHP*

Beispiel 7.15:

- gesucht: *QR*-Zerlegung von

$$A = (a_1, a_2) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & -1 \end{pmatrix}$$

OHP

Bemerkung 7.16:

1. Q ist für $m \neq n$ *keine* Orthogonalmatrix, da nicht quadratisch
2. - wenn A quadratisch und Q mittels Gram-Schmidt errechnet:
   - $A = QR \Leftrightarrow R = Q^\top A$
   - R ist Dreiecksmatrix, da $L(a_1, \dots, a_i) = L(q_1, \dots, q_i)$ beim Verfahren von Gram-Schmidt gilt
3. - QR-Zerlegung einer Matrix $A \in \mathbb{R}^{m \times n}, m \geq n$ und $rg(A) = n$ ist nur bis auf Vorzeichen eindeutig
   - betrachte A aus Beispiel 7.15
   - es gilt $A = \tilde{Q}\tilde{R}$ mit $\tilde{Q} = (-q_1\, q_2)$ und

$$\tilde{R} = \begin{pmatrix} -\varrho_{11} & -\varrho_{12} \\ 0 & \varrho_{22} \end{pmatrix}.$$

4.
- Einschränkungen an A in Satz 7.14 nur wegen Beweistechnik (Gram-Schmidt-Verfahren)
- jede reelle rechteckige Matrix besitzt eine QR-Zerlegung mit $\mathrm{rg}(R) = \mathrm{rg}(A)$

5.
- QR-Zerlegung mit dem Verfahren von Gram-Schmidt ungeeignet für Berechnungen mit dem Computer
- Grund: unvermeidliche Rundungsfehler können sich im Laufe der Berechnung verheerend auswirken
- Gram-Schmidt-Verfahren *numerisch instabil*
- effiziente und numerisch stabile Verfahren: numerische Mathematik

6.
- in der Praxis: Lösung bei Handrechnung eher mit Gauß-Elimination als mit QR-Zerlegung
- für A invertierbar: dasselbe gilt für Computerberechnungen
- aber: numerische Behandlung von überbestimmten linearen Gleichungssystemen üblicherweise auf Basis einer QR-Zerlegung

7.3 Eigenwerte und Eigenvektoren

- wiederum: Was „macht eine lineare Abbildung anschaulich mit den Vektoren“?
- betrachten $\mathbb{C}$-Vektorräume V; der reelle Fall ist als Spezialfall enthalten
- Sprechweise:

Definition 7.17: Eine lineare Abbildung $f : V \to V$ heißt *Endomorphismus*.

- Sei ab jetzt f ein Endomorphismus
- Wunschvorstellung: zu f gebe es Vektoren v_i mit

$$f(v_i) = \lambda_i v_i, \quad \lambda_i \in \mathbb{C} \tag{7.6}$$

- weiter : $\mathcal{B} = (v_1, \ldots, v_n)$ sei Basis von V

- sei $M_{\mathcal{B}}^{\mathcal{B}}(f)$ die Darstellungsmatrix von f bzgl. $\mathcal{B}$
- in den Spalten von $M_{\mathcal{B}}^{\mathcal{B}}(f)$ stehen die Koordinaten der Bilder der Basisvektoren v_i
- $K_{\mathcal{B}}(v_i) = e_i$
- weiter: $K_{\mathcal{B}}(f(v_i)) = \lambda_i K_{\mathcal{B}}(v_i) = \lambda_i e_i$
- es folgt

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

- „Wirkung“ von f ist leicht zu verstehen:

- f skaliert jede Koordinate bezogen auf $\mathcal{B}$ eines Vektors mit λ_i
- Darstellungsmatrix so einfach wie nur möglich
- nur: gibt es derartige v_i und wie kann man sie berechnen?

Definition 7.18: Existiert für f ein $\lambda \in \mathbb{C}$ und $v \in V \setminus \{0\}$ mit

$$f(v) = \lambda v\,, \tag{7.7}$$

dann heißt v *Eigenvektor* von f zum *Eigenwert* λ.

Bemerkung 7.19:

1. - Eigenvektoren als Verallgemeinerung von Fixpunkten ($f(x) = x$)
   - Fixpunkte bleiben unter f invariant
   - Eigenvektoren: Richtung bleibt invariant, Länge eventuell nicht

2.
- Nullvektor wird als Eigenvektor ausgeschlossen
- sonst: $0 = f(0) = \lambda 0$, und jedes λ wäre Eigenwert
- Begriff „Eigenwert“ in diesem Fall sinnlos

- elementare Eigenschaften von Eigenvektoren:

Bemerkung 7.20:

1. Sei λ ein Eigenwert von f und $v_1, \ldots, v_k$ Eigenvektoren von f zu λ. Dann ist auch $v \in L(v_1, \ldots, v_k) \setminus \{0\}$ ein Eigenvektor von f zu λ.

2. Für $\lambda \in \mathbb{C}$ ist $Eig(f;\lambda) := \{v \in V \,|\, f(v) = \lambda v\}$, der *Eigenraum* von f zu λ, ein Untervektorraum von V.

3. Für $\lambda \neq \gamma$ gilt $Eig(f;\lambda) \cap Eig(f;\gamma) = \{0\}$.

4. Eigenvektoren zu unterschiedlichen Eigenwerten sind linear unabhängig.

Beweis *OHP*

- zur Berechnung von Eigenwerten und Eigenvektoren: $V = \mathbb{C}^n$; Standardbasis
- dann: f lässt sich mit seiner Darstellungsmatrix A identifizieren
- aus der Definition folgt:

$$Ax = \lambda x \Leftrightarrow (A - \lambda E)x = 0\,. \tag{7.8}$$

- (7.8) hat genau dann nichttriviale Lösungen, wenn $\det(A - \lambda E) = 0$ gilt
- entscheidende Bedingung an die Eigenwerte gefunden

Satz 7.21: Sei $A \in \mathbb{C}^{n\times n}$.

1. Die Funktion

$$\chi_A(\lambda) := \det(A - \lambda E) \tag{7.9}$$

ist ein Polynom mit $\deg(\chi_A) = n$ und heißt *charakteristisches Polynom* von A.

2. $\lambda \in \mathbb{C}$ ist Eigenwert von $A \Leftrightarrow \chi_A(\lambda) = 0$.
3. A hat (mit Vielfachheit) genau n Eigenwerte $\lambda_i \in \mathbb{C}$.
4. $Eig(f;\lambda) = \ker(A - \lambda E)$.

Beweis *OHP*

- gegeben sei eine konkrete Matrix $A \in \mathbb{C}^{n \times n}$
- Ziel: Berechnung der Eigenwerte und -vektoren
- Vorgehen bei Handrechnung:

1. - Man finde alle Nullstellen $\lambda_1, \lambda_2, ..., \lambda_n$ von $\chi_A(\lambda) \in \mathbb{C}$
   - Nullstellen: Eigenwerte von A
   - mehrfache Eigenwerte möglich
2. - zu λ_i: Man bestimmte die Eigenvektoren als Lösung von $(A - \lambda_i E)x = 0$
   - unendlich viele Lösungen!
   - insgesamt: Man löst so viele verschiedene lineare Gleichungssysteme wie es verschiedene Eigenwerte gibt
3. - effiziente Beschreibung von $Eig(A; \lambda)$ durch Angabe einer Basis

- löst man $(A-\lambda_i E)x = 0$, ergibt sich diese implizit durch die Parametrisierung der Lösungsmenge

Beispiel 7.22: Gesucht sind alle Eigenwerte und Eigenvektoren von

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

OHP

Beispiel 7.23: Wir bestimmen die Eigenwerte und Eigenvektoren der Drehmatrix

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

aus Beispiel 7.2.

- setze $c := \cos\alpha$, $s := \sin\alpha$

OHP

- Weglassen des Betrags: Änderung der (willkürlichen Reihenfolge) der Eigenwerte
- Eulerscher Satz: $\lambda_{1,2} = e^{\pm i\alpha}$
- $\lambda_1 = \overline{\lambda_2}$ und $|\lambda_i| = 1$
- Eigenwerte liegen auf dem komplexen Einheitskreis
- reelle Eigenwerte, wenn $\sin\alpha = 0$, also für $\alpha = 0$ oder $\alpha = 180°$
- erster Fall: identische Abbildung
- zweiter Fall: $f(x) = -x$, also Punktspiegelung um 0
- allgemein: eine Drehung ändert die Richtung eines Vektors, es können daher keine reellen Eigenwerte existieren
- bei reellen Eigenwerten: reelle Eigenvektoren behalten Richtung

- jetzt: Berechnung der Eigenvektoren für $\sin \alpha \neq 0$

OHP

Bemerkung 7.24:

1.
   - Eigenwerte der Drehung aus obigen Beispiel besitzen Betrag eins
   - allgemeiner: sei f lineare Isometrie mit Eigenwert λ und Eigenvektor x
   - dann: $\|x\| = \|f(x)\| = \|\lambda x\| = |\lambda| \, \|x\|$
   - also $|\lambda| = 1$

2.
   - „kritischer“ Teil der Berechnung von Eigenwerten/Eigenvektoren: Finden der Nullstellen des charakteristischen Polynoms
   - Polynomgrad 3 und 4: Cardanosche Formeln oder Raten einer Nullstelle mit anschließender Polynomdivision
   - Polynomgrad ≥ 5: Es existiert kein Rechenverfahren, mit dem sich mit endlich vielen Rechenschritten alle Nullstellen eines beliebigen Polynoms bestimmen ließen
   - „Satz von Abel“

  - bei praktischen Eigenwertproblemen: häufig Grad ≥ 5
  - dann: nur Näherungslösungen möglich

- jetzt wieder: allgemeiner Endomorphismus f
- sei A Darstellungsmatrix von f bezogen auf irgendeine Basis
- B Darstellungsmatrix bzgl. irgendeiner anderen Basis
- dann: $B = SAS^{-1}$
- es folgt
- Darstellungsmatrizen von f zu verschiedenen Basen haben alle die gleichen Eigenwerte
- daher: Man kann *irgendeine* Darstellungsmatrix A von f wählen
- man bestimmt die Eigenwerte von A

- dies sind zugleich die Eigenwerte von f
- Eigenvektoren von A: Koordinaten der Eigenvektoren von f in der Basis, auf die sich die Darstellungsmatrix bezieht
- Beweis und Details: Literatur
- jetzt: Zusammenhang von Determinante und Eigenwerten

Satz 7.25: Für $A = (a_1, \ldots, a_n) \in \mathbb{C}^{n \times n}$ mit Eigenwerten $\lambda_i, 1 \le i \le n$ gilt

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

Beweis *OHP*

- bei allgemeinen Matrizen: Eigenwerte können schwierig zu finden sein
- bei Dreiecksmatrizen: Eigenwerte sehr leicht zu finden

Bemerkung 7.26: Die Eigenwerte einer Dreiecksmatrix sind die Werte auf der Hauptdiagonalen.

Beweis *OHP*

Bemerkung 7.27: Sei x ein Eigenvektor zum Eigenwert λ von A. Dann gilt

$$A^k x = \lambda^k x\,.$$

Beweis

$$A^k x = A^{k-1} A x = A^{k-1}(\lambda x) = \lambda A^{k-1} x = \ldots = \lambda^k x \qquad \blacksquare$$

7.4 Diagonalisierung linearer Abbildungen

- Überlegung zu Beginn des letzten Kapitels: eine Basis $\mathcal{B}$ aus Eigenvektoren von f führt zu einer diagonalen Darstellungsmatrix $D = M_{\mathcal{B}}^{\mathcal{B}}(f)$
- auf der Hauptdiagonalen von D: Eigenwerte von f
- D ist einfach zu interpretieren
- sei A die Darstellungsmatrix von f bezogen auf irgendeine Basis $\tilde{\mathcal{B}}$
- sei B die Matrix des Basiswechsels von $\mathcal{B}$ nach $\tilde{\mathcal{B}}$
- dann: $D = B^{-1}AB$ bzw. $A = BDB^{-1}$
- nur: Gibt es immer eine Basis aus Eigenvektoren?

Definition 7.28: Ein $A \in \mathbb{C}^{n \times n}$ heißt *diagonalisierbar* $:\Longleftrightarrow \exists B \in GL(n;\mathbb{C}), D$ Diagonalmatrix mit

$$A = BDB^{-1}\,. \tag{7.10}$$

Bemerkung 7.29: Es ist A genau dann diagonalisierbar, wenn eine Basis aus Eigenvektoren existiert. In diesem Fall bilden die Spalten von B aus Formel (7.10) eine solche Basis aus Eigenvektoren.

Beweis *OHP*

Beispiel 7.30:

- betrachte $f(x) = Ax$ mit

$$A = \begin{pmatrix} 1,5 & 0,5 \\ 0,5 & 1,5 \end{pmatrix}$$

- wurde schon früher diskutiert
- geometrische Überlegungen: Basis $\tilde{\mathcal{E}} = (\tilde{e}_1, \tilde{e}_2)$
- dies entspricht einer um 45° gedrehten Standardbasis
- Darstellungsmatrix:

$$\tilde{A} = T^{\mathcal{E}}_{\tilde{\mathcal{E}}} A T^{\tilde{\mathcal{E}}}_{\mathcal{E}} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \tag{7.11}$$

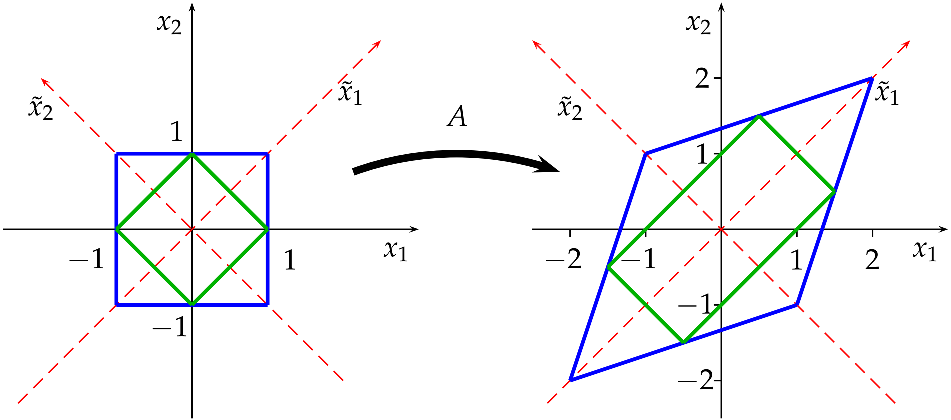


Abbildung 7.1: Quadrate unter der linearen Abbildung f aus den Beispielen 4.46, 4.84 und 7.30

- $T^{\mathcal{E}}_{\tilde{\mathcal{E}}}$ bezeichnet die Matrix des Basiswechsels von $\mathcal{E}$ nach $\tilde{\mathcal{E}}$
- A ist damit diagonalisierbar, weil $A = B^{-1}DB$ mit $B = T^{\mathcal{E}}_{\tilde{\mathcal{E}}}$ und $\tilde{A} = D$
- jetzt: Anwendung der neu entwickelten Theorie auf dieses Beispiel

OHP

- das ist das schon bekannte Ergebnis
- aber: Eigenwerttheorie zeigt einen systematischen Weg zu einer Basis wie $\tilde{\mathcal{E}}$
- Zerlegung $A = BDB^T$ wurde algebraisch gefunden
- geometrische Deutung möglich
- Abbildung f wirkt in drei Schritten:

1.
   - durch B^T: Drehung um 45° gegen den Uhrzeigersinn
   - Grund: wegen $|B| = 1$ gilt $B \in SO(2)$, also auch $B^T \in SO(2)$
   - Drehmatrix mit Drehwinkel α; $\cos\alpha = 1/\sqrt{2}$, also $\alpha = 45°$
2.
   - durch D Skalierung längs der gedrehten x_1-Achse um 2
   - gedrehte x_2-Koordinate bleibt unverändert

3. • B: Rückdrehung um 45° im Uhrzeigersinn.

Beispiel 7.31: • B und D nicht eindeutig

$$A = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$$

Nicht jede Matrix lässt sich diagonalisieren.

Beispiel 7.32: Sei $s(x) = Sx$ mit

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

- Abbildung: s ist eine Scherung
- x_1-Achse invariant
- alle anderen Vektoren: Richtungsänderung durch S, können also keine Eigenvektoren sein

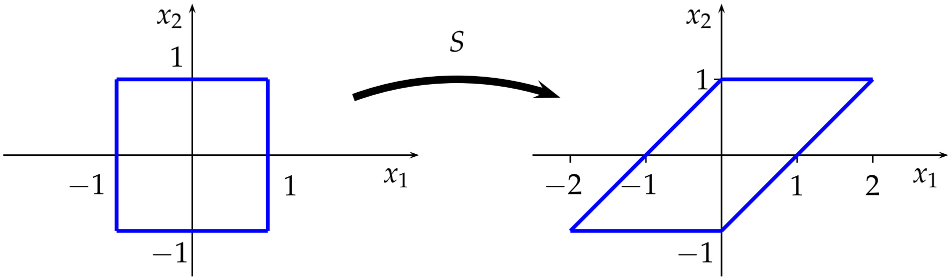


Abbildung 7.2: Einheitsquadrat unter der linearen Abbildung s aus Beispiel 7.32

- $\chi_S(\lambda) = (1 - \lambda)^2$: 1 ist doppelte Nullstelle
- 1 einziger Eigenwert von S
- Eigenvektoren:

OHP

- Ergebnis entspricht der geometrische Anschauung
- S ist nicht diagonalisierbar, weil alle Eigenvektoren auf der x-Achse liegen
- Basis von $\mathbb{C}^2$ aus Eigenvektoren existiert nicht
- jetzt: Kriterien zur Diagonalisierbarkeit
- Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig
- damit erhält man

Satz 7.33: Sei $A \in \mathbb{C}^{n\times n}$, und es mögen n verschiedene Eigenwerte existieren. Dann ist A diagonalisierbar.

Bemerkung 7.34:

- Satz 7.33 garantiert nicht die Diagonalisierbarkeit in $\mathbb{R}$

- betrachte Drehmatrix A_α: Eigenwerte sind $e^{\pm i\alpha}$
- Eigenvektoren komplexwertig
- Nicht-Diagonalisierbarkeit nach obigem Ergebnis nur möglich bei mehrfachen Eigenwerten
- Wir betrachten diesen Fall näher.

Definition 7.35: Sei $A \in \mathbb{C}^{n\times n}$ und λ ein Eigenwert. Die Vielfachheit der Nullstelle λ von χ_A heiße *algebraische Vielfachheit* $a(\lambda)$, weiter sei $g(\lambda) := \dim(Eig(A;\lambda))$ die *geometrische Vielfachheit* von λ.

Satz 7.36: Sei A wie oben. Existiert ein Eigenwert $\tilde{\lambda}$ mit $a(\tilde{\lambda}) > g(\tilde{\lambda})$, dann ist A nicht diagonalisierbar.

Beweis *OHP*

Beispiel 7.37:

- Beispiel der Scherung S: einziger Eigenwert $\lambda = 1$
- $a(1) = 2 > g(1) = 1$
- S ist nicht diagonalisierbar.
- viel übersichtlicher: Situation bei symmetrischen Matrizen

Satz 7.38: Eigenwerte reeller symmetrischer Matrizen sind immer reell. Zu jedem Eigenwert einer reellen symmetrischen Matrix existieren reelle Eigenvektoren.

Beweis *OHP*

Satz 7.39: Sei $A \in \mathbb{R}^{n \times n}$ symmetrisch und $\lambda \neq \mu$ zwei Eigenwerte von A mit Eigenvektoren v bzw. w. Dann gilt $v \perp w$.

Beweis *OHP*

Wir kommen nun zum Hauptergebnis dieses Kapitels.

Satz 7.40: Zu jeder reellsymmetrischen Matrix A gibt es eine Orthogonalmatrix Q und eine Diagonalmatrix D mit

$$A = QDQ^T \, .$$

Auf der Hauptdiagonalen von D stehen die Eigenwerte von A, in den Spalten von Q Eigenvektoren von A.

Beweis *OHP*

Beispiel 7.41:

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

- A symmetrisch
- also: es existieren v_1, v_2, v_3 orthonormal zu den drei reellen Eigenwerten $\lambda_1, \lambda_2, \lambda_3$
- sei $Q = (v_1, v_2, v_3)$
- dann gilt:

$$Q^T A Q = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}.$$

OHP

Es gilt die Umkehrung von Satz 7.40.

Bemerkung 7.42: Gibt es zu $A \in \mathbb{R}^{n\times n}$ ein $Q \in O(n)$ und eine Diagonalmatrix D mit $A = QDQ^T$, dann ist A symmetrisch.

Beweis

$$A^T = (QDQ^T)^T = (Q^T)^T D^T Q^T = QDQ^T = A$$

- hier noch ein weiteres Kriterium zur Diagonalisierbarkeit

Definition 7.43: $A \in \mathbb{C}^{n\times n}$ heißt *normal*, wenn gilt:

$$A\overline{A}^T = \overline{A}^T A\,. \tag{7.12}$$

Satz 7.44: Sei $A \in \mathbb{C}^{n\times n}$ normal. Dann existiert eine bzgl. des euklidischen Skalarprodukts in $\mathbb{C}^n$ orthonormale Basis aus Eigenvektoren, d. h. A ist diagonalisierbar.

- Beweis: siehe Literatur

Folgerung 7.45:

1. Jede reelle symmetrische Matrix ist diagonalisierbar. Die Eigenwerte sind reell.
2. Jede reelle antisymmetrische Matrix (d. h. $A^T = -A$) ist diagonalisierbar. Die Eigenwerte sind rein imaginär oder 0.
3. Jede reelle orthogonale Matrix ist diagonalisierbar.

Beweis *OHP*

7.5 Definitheit und Skalarprodukte

- in der Vorlesung: Skalarprodukt spielt eine zentrale Rolle
- bekannt: es existiert auf $\mathbb{R}^n$ eine Fülle von Skalarprodukten
- systematische Methode zur Konstruktion von Skalarprodukten?
- ein Skalarprodukt $(\cdot,\cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ muss drei Bedingungen erfüllen:

1. $(\cdot,\cdot)$ ist linear in den Spalten.
2. $(\cdot,\cdot)$ ist symmetrisch.
3. $(x,x) > 0$ für $x \neq 0$.

- wähle beliebige Matrix $A \in \mathbb{R}^{n\times n}$
- betrachte $(\cdot,\cdot)_A : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, (x,y)_A := \langle x, Ay\rangle$

- dabei: $\langle\cdot,\cdot\rangle$ ist das euklidische Skalarprodukt
- Bedingung 1. immer erfüllt
- sei $A = (a_1, \ldots, a_n)$, e_i und e_j Standard-Einheitsvektoren
- dann:

$$(e_i, e_j)_A = e_i^T A e_j = e_i^T a_j = a_{ij},$$

- $(e_i, e_j)_A = (e_j, e_i)_A \Leftrightarrow a_{ij} = a_{ji}$
- $(\cdot,\cdot)_A$ genau dann symmetrisch, wenn A symmetrisch

Bemerkung 7.46:

- Einschränkung der Struktur auf $(\cdot,\cdot)_A$ ist keine
- Sei $[\cdot,\cdot]$ irgendein Skalarprodukt und A definiert durch $a_{ij} = [e_i, e_j]$.

- dann: $[\cdot,\cdot] = \langle x, Ay \rangle$
- Jedes Skalarprodukt wird durch eine Matrix vollständig beschrieben.

Beweis *OHP*

- nicht jede symmetrische Matrix definiert ein Skalarprodukt
- die Nullmatrix sicher nicht
- Welche Matrizen erfüllen Kriterium 3. von oben?

Definition 7.47: Sei A ein symmetrische Matrix $A \in \mathbb{R}^{n\times n}$.

1. Die Abbildung $x \to \langle x, Ax\rangle$ wird *quadratische Form* genannt.
2. A heißt *positiv definit*, wenn $\langle x, Ax\rangle > 0\,\forall x \in \mathbb{R}^n \setminus \{0\}$.
3. A heißt *negativ definit*, wenn $\langle x, Ax\rangle < 0\,\forall\, x \in \mathbb{R}^n \setminus \{0\}$.
4. A heißt *positiv semidefinit*, wenn $\langle x, Ax\rangle \geq 0\,\forall\, x \in \mathbb{R}^n \setminus \{0\}$.
5. A heißt *negativ semidefinit*, wenn $\langle x, Ax\rangle \leq 0\,\forall\, x \in \mathbb{R}^n \setminus \{0\}$.

6. A heißt *indefinit*, falls sie weder positiv noch negativ (semi-)definit ist, d. h.

$$\exists x, y \in \mathbb{R}^n \setminus \{0\} : \langle x, Ax \rangle > 0 \quad \wedge \quad \langle y, Ay \rangle < 0.$$

7. Häufig kürzt man „**s**ymmetrisch **p**ositiv **d**efinit“ mit „spd“ ab.

Beispiel 7.48: Die quadratische Form trägt ihren Namen zu Recht. Es sei

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

OHP

Folgerung 7.49: Die Abbildung $(\cdot,\cdot)_A$ ist genau dann ein Skalarprodukt, wenn A spd ist.

Definition 7.50: Für $A \in \mathbb{R}^{n\times n}$ seien A_k die links oben beginnenden $k \times k$-Untermatrizen $A_k = (a_{ij})_{i,j=1}^{k}$ von A. Dann heißen $D_k = \det(A_k)$ die *Hauptunterdeterminanten* oder *Hauptminoren* von A.

Beispiel 7.51:

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

OHP

Satz 7.52: Für eine reelle symmetrische Matrix A sind äquivalent:

1. A ist positiv definit.
2. A besitzt nur positive Eigenwerte.
3. Alle Hauptminoren von A sind positiv.

Weiter ist A genau dann positiv semidefinit, wenn alle Eigenwerte von A nicht negativ sind. In diesem Fall sind alle Hauptminoren von A nicht negativ.

- vollständiger Beweis in der Literatur (z. B. Fischer, Kap. 5.7)

Beispiel 7.53: *OHP*

Bemerkung 7.54: Eine spd-Matrix $A \in \mathbb{R}^{n \times n}$ ist invertierbar.

Beweis Nach Satz 7.52.3 ist $\det(A) = D_n \neq 0$ und daher A invertierbar. ∎

Bemerkung 7.55:

- Sei A negativ definit
- dann: $-\langle x, Ax\rangle = \langle x, -Ax\rangle > 0$
- also $-A$ positiv definit
- analog: A genau dann negativ semidefinit, wenn $-A$ positiv semidefinit

Bemerkung 7.56:

- gegeben Hauptminoren D_k einer negativ definiten Matrix
- dann:

$$(-1)^k D_k = (-1)^k \det(A_k) = \det(-A_k) > 0, \tag{7.13}$$

- Die Hauptminoren einer negativ definiten Matrix haben wechselndes Vorzeichen
- beginnend mit negativem Vorzeichen
- Sei $f : \mathbb{R}^n \to \mathbb{R}$ hinreichend oft differenzierbar
- lokales Maximum in x_0 sicher dann, wenn $\nabla f(x_0) = 0$ und $\nabla^2 f$ in x_0 negativ definit
- Prüfung üblicherweise mit Kriterium (7.13)
- bekannt als *Hurwitz-Kriterium*

	D_1	D_2	D_3	D_4	$\ldots$	Kriterium
positiv definit	> 0	> 0	> 0	> 0	$\ldots$	hinreichend
negativ definit	< 0	> 0	< 0	> 0	$\ldots$	hinreichend
positiv semidefinit	≥ 0	≥ 0	≥ 0	≥ 0	$\ldots$	*notwendig*
negativ semidefinit	≤ 0	≥ 0	≤ 0	≥ 0	$\ldots$	*notwendig*
indefinit	sonst					hinreichend

- letzte Spalte: Vorzeichenfolge nur notwendig oder aber hinreichend

Bemerkung 7.57:

1. - A aus Beispiel 7.51: indefinit (siehe Vorzeichenfolge)
   - Eigenwerte: $\lambda_1 \approx -4,37, \lambda_2 \approx -1,12$ sowie $\lambda_3 \approx 3,10, \lambda_4 \approx 5,39$

2. - Kriterien zur Semidefinitheit nur notwendig:

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

- offenbar indefinit (Eigenwerte!)
- aber: $D_1 = D_2 = D_3 = 0 \geq 0$

- jetzt: spezielles Beispiel einer spd-Matrix

Bemerkung 7.58: Sei $m \geq n$ und $A \in \mathbb{R}^{m \times n}$ mit $\operatorname{rg}(A) = n$. Dann ist $A^T A$ positiv definit. Für $\operatorname{rg}(A) < n$ ist $A^T A$ positiv semidefinit.

Beweis *OHP*

Bemerkung 7.59: Für $A \in \mathbb{R}^{m \times n}$ mit $m \leq n$ ist AA^T positiv semidefinit und sogar positiv definit, falls $\mathrm{rg}(A) = m$. Dies zeigt man analog zu Bemerkung 7.58.

Index

Literaturverzeichnis

Literaturverzeichnis

[1] A. Beutelspacher. *Das ist o. B. d. A. trivial!* Vieweg, 8 edition, 2006.

[2] G. Fischer. *Lineare Algebra*. Springer, 18 edition, 2014.