

Mathematik I

(für die Studiengänge Informatik, Elektrotechnik und Physik)

Peter Stollmann

Professur Analysis

Wintersemester 2019/20



Mathematik!
TU Chemnitz

Vorlesung

Prof. Peter Stollmann

Mo 9:15

Do 11:30

Übungen/Praktika

Clemens Bombach

Christian Rebs

Dr. Martin Tautenhahn

Die Übungen beginnen in der ersten Vorlesungswoche, das heißt (je nach Übungsgruppe) ab 15. Oktober 2019.

Das Praktikum beginnt in der zweiten Vorlesungswoche, das heißt (je nach Übungsgruppe) ab 21. Oktober 2019.

Webseite

<https://bildungsportal.sachsen.de/opal/auth/RepositoryEntry/21422997>

Stand 28.09.2018

	Studiengang	SWS	Klausurzeit	LP	AS	Studenten
B_In	Informatik	4V+2Ü+2P	120	9	270	28
B_AI	Angewandte Informatik	4V+2Ü+2P	120	9	270	29
B_BT	Biomedizinische Technik	4V+2Ü+2P	180	8	240	29
B_EM	Elektromobilität	4V+2Ü+2P	120	8	240	9
B_ET	Elektrotechnik und Informationstechnik	4V+2Ü+2P	120	8	240	31
B_RE	Regenerative Energietechnik	4V+2Ü+2P	120	8	240	13
B_Ph	Physik	4V+2Ü(+2P)	120	8	240	18
M_IG	Informatik f. Geistes- u. Soz.-Wiss.)	4V+2Ü+2P	120	9	270	23

180

(laut Modulbeschreibungen **Mathematik I** bzw. **Höhere Mathematik I**)

AS = Gesamtarbeitsaufwand in Stunden

LP = Leistungspunkte

Neben der Präsenzzeit von 4,5 h/Woche (= 67.5 gesamt) entfällt also ein erheblicher Anteil auf **Selbststudium**:

- Vor- und Nachbereitung von Lehrveranstaltungen,
- Literaturstudium,
- Lösen von Übungsaufgaben,
- Prüfungsvorbereitung.

- Klausurarbeit am Ende des Wintersemesters (Umfang laut Tabelle).
- Termin wird bekanntgegeben sobald von zentraler Prüfungsplanung verkündet.
- Zugelassene Hilfsmittel: Ausdruck dieser Folien, Randnotizen hierauf, Formelsammlung (kein Taschenrechner).

Inhalt:

- Grundlagen:
Aussagenlogik, Mengenlehre, Zahlen, Abbildungen, Relationen
- Lineare Algebra:
Vektorräume, Matrizen, lineare Gleichungssysteme, Determinanten, Skalarprodukt, analytische Geometrie, Eigenwerte und -vektoren

Ziele: Die Studierenden sollen

- Verständnis der 'mathematischen Sprache' entwickelt haben,
- das elementare technische Reservoir der Mathematik (Grundlagen, lineare Algebra und, im 2. Semester, Infinitesimalrechnung einer Variablen) kennen und beherrschen,
- einfache mathematische Modelle aus den Naturwissenschaften analysieren können.

- Arbeiten Sie **von Anfang an** intensiv mit.

- Arbeiten Sie **von Anfang an** intensiv mit.
- Arbeiten Sie die Vorlesungen nach.

- Arbeiten Sie **von Anfang an** intensiv mit.
- Arbeiten Sie die Vorlesungen nach.
Stellen Sie Fragen, wenn sich Unklarheiten ergeben. Werden Sie insbesondere sofort aktiv, wenn Sie den 'roten Faden' verloren haben.

- Arbeiten Sie **von Anfang an** intensiv mit.
- Arbeiten Sie die Vorlesungen nach.
Stellen Sie Fragen, wenn sich Unklarheiten ergeben. Werden Sie insbesondere sofort aktiv, wenn Sie den 'roten Faden' verloren haben.
- Bearbeiten Sie die Übungsaufgaben, Beispiele und Übungsklausuren **selbständig**.

- Arbeiten Sie **von Anfang an** intensiv mit.
- Arbeiten Sie die Vorlesungen nach.
Stellen Sie Fragen, wenn sich Unklarheiten ergeben. Werden Sie insbesondere sofort aktiv, wenn Sie den 'roten Faden' verloren haben.
- Bearbeiten Sie die Übungsaufgaben, Beispiele und Übungsklausuren **selbständig**.
- Nehmen Sie an den **Übungen** teil.

- Arbeiten Sie **von Anfang an** intensiv mit.
- Arbeiten Sie die Vorlesungen nach.
Stellen Sie Fragen, wenn sich Unklarheiten ergeben. Werden Sie insbesondere sofort aktiv, wenn Sie den 'roten Faden' verloren haben.
- Bearbeiten Sie die Übungsaufgaben, Beispiele und Übungsklausuren **selbständig**.
- Nehmen Sie an den **Übungen** teil.
- Nehmen Sie das **Praktikum** in Anspruch.

- Arbeiten Sie **von Anfang an** intensiv mit.
- Arbeiten Sie die Vorlesungen nach.
Stellen Sie Fragen, wenn sich Unklarheiten ergeben. Werden Sie insbesondere sofort aktiv, wenn Sie den 'roten Faden' verloren haben.
- Bearbeiten Sie die Übungsaufgaben, Beispiele und Übungsklausuren **selbständig**.
- Nehmen Sie an den **Übungen** teil.
- Nehmen Sie das **Praktikum** in Anspruch.
- Bilden Sie **Arbeitsgruppen**. Wer über Mathematik spricht, versteht diese besser und kann Probleme besser identifizieren.

- Arbeiten Sie **von Anfang an** intensiv mit.
- Arbeiten Sie die Vorlesungen nach.
Stellen Sie Fragen, wenn sich Unklarheiten ergeben. Werden Sie insbesondere sofort aktiv, wenn Sie den 'roten Faden' verloren haben.
- Bearbeiten Sie die Übungsaufgaben, Beispiele und Übungsklausuren **selbständig**.
- Nehmen Sie an den **Übungen** teil.
- Nehmen Sie das **Praktikum** in Anspruch.
- Bilden Sie **Arbeitsgruppen**. Wer über Mathematik spricht, versteht diese besser und kann Probleme besser identifizieren.
- Lesen Sie dem Stand der Vorlesung entsprechende Kapitel in der **Literatur**, um andere Aspekte kennenzulernen.

- Arbeiten Sie **von Anfang an** intensiv mit.
- Arbeiten Sie die Vorlesungen nach.
Stellen Sie Fragen, wenn sich Unklarheiten ergeben. Werden Sie insbesondere sofort aktiv, wenn Sie den 'roten Faden' verloren haben.
- Bearbeiten Sie die Übungsaufgaben, Beispiele und Übungsklausuren **selbständig**.
- Nehmen Sie an den **Übungen** teil.
- Nehmen Sie das **Praktikum** in Anspruch.
- Bilden Sie **Arbeitsgruppen**. Wer über Mathematik spricht, versteht diese besser und kann Probleme besser identifizieren.
- Lesen Sie dem Stand der Vorlesung entsprechende Kapitel in der **Literatur**, um andere Aspekte kennenzulernen.
- Versuchen Sie Freude, Ausdauer und sportlichen Ehrgeiz beim Lösen der Probleme zu entwickeln.

- Arbeiten Sie **von Anfang an** intensiv mit.
- Arbeiten Sie die Vorlesungen nach.
Stellen Sie Fragen, wenn sich Unklarheiten ergeben. Werden Sie insbesondere sofort aktiv, wenn Sie den 'roten Faden' verloren haben.
- Bearbeiten Sie die Übungsaufgaben, Beispiele und Übungsklausuren **selbständig**.
- Nehmen Sie an den **Übungen** teil.
- Nehmen Sie das **Praktikum** in Anspruch.
- Bilden Sie **Arbeitsgruppen**. Wer über Mathematik spricht, versteht diese besser und kann Probleme besser identifizieren.
- Lesen Sie dem Stand der Vorlesung entsprechende Kapitel in der **Literatur**, um andere Aspekte kennenzulernen.
- Versuchen Sie Freude, Ausdauer und sportlichen Ehrgeiz beim Lösen der Probleme zu entwickeln.
- Nehmen Sie sich von Anfang an **genügend Zeit**.

Beispiel zur Zeitplanung

Laut Modulbeschreibung ist der Selbststudiumsanteil im Modul auf 160–200 h ausgelegt. Diese könnte man exemplarisch wie folgt aufteilen:

- 120 h während des Semesters, d. h. durchschnittlich 8 h pro Woche (ggf. könnte man einen Teil dieser Zeit im Tutorium oder in einer Lerngruppe verbringen),
- 40–80 h Prüfungsvorbereitung in der vorlesungsfreien Zeit.

Wie Sie die Zeit aufteilen, ist natürlich Ihre Sache. Wir empfehlen Ihnen jedoch, von vornherein einen Zeitplan zu erstellen und auch konsequent einzuhalten.

Falls das noch nicht reicht.

hilft nur, sich mehr Zeit zu nehmen. Die Modulbeschreibung geht von durchschnittlichem Talent und Vorkenntnissen aus.

Angebote wie den *Lern | RAUM für Mathematik* nutzen.

In den seltensten Fällen ist jemand generell unfähig den Stoff zu verstehen – in der Regel braucht man einfach nur **mehr Zeit**.

Was sicher schiefgeht

... ist das Umsetzen verbreiteter 'pragmatischer' Einstellungen wie

- Ich kann in Vorlesung/Übung dem Dozenten recht gut folgen, also beherrsche ich den Stoff – 'Beschallenlassen' reicht vollkommen.

Was sicher schiefgeht

... ist das Umsetzen verbreiteter 'pragmatischer' Einstellungen wie

- Ich kann in Vorlesung/Übung dem Dozenten recht gut folgen, also beherrsche ich den Stoff – 'Beschallenlassen' reicht vollkommen.
- Die Vorlesung zeigt mir die theoretische Seite des Stoffes – ich brauche aber für die Klausur nur die Anwendung. Also reicht es, Übung bzw. Tutorium zu besuchen.

Was sicher schiefgeht

... ist das Umsetzen verbreiteter 'pragmatischer' Einstellungen wie

- Ich kann in Vorlesung/Übung dem Dozenten recht gut folgen, also beherrsche ich den Stoff – 'Beschallenlassen' reicht vollkommen.
- Die Vorlesung zeigt mir die theoretische Seite des Stoffes – ich brauche aber für die Klausur nur die Anwendung. Also reicht es, Übung bzw. Tutorium zu besuchen.
- Mathematische Techniken kann man wie Rezepte auswendig lernen; das reicht für Anwendung und Klausur. Ein tieferes Verständnis von Inhalten und Zusammenhängen benötigen nur Mathematiker.

Was sicher schiefgeht

... ist das Umsetzen verbreiteter 'pragmatischer' Einstellungen wie

- Ich kann in Vorlesung/Übung dem Dozenten recht gut folgen, also beherrsche ich den Stoff – 'Beschallenlassen' reicht vollkommen.
- Die Vorlesung zeigt mir die theoretische Seite des Stoffes – ich brauche aber für die Klausur nur die Anwendung. Also reicht es, Übung bzw. Tutorium zu besuchen.
- Mathematische Techniken kann man wie Rezepte auswendig lernen; das reicht für Anwendung und Klausur. Ein tieferes Verständnis von Inhalten und Zusammenhängen benötigen nur Mathematiker.
- Mein Stundenplan lässt mir nur wenig Zeit. Ich arbeite am Ende des Semesters einfach alles 'am Stück' nach. Ein, zwei Wochen (oder gar Tage) werden schon reichen.

Was sicher schiefgeht

... ist das Umsetzen verbreiteter 'pragmatischer' Einstellungen wie

- Ich kann in Vorlesung/Übung dem Dozenten recht gut folgen, also beherrsche ich den Stoff – 'Beschallenlassen' reicht vollkommen.
- Die Vorlesung zeigt mir die theoretische Seite des Stoffes – ich brauche aber für die Klausur nur die Anwendung. Also reicht es, Übung bzw. Tutorium zu besuchen.
- Mathematische Techniken kann man wie Rezepte auswendig lernen; das reicht für Anwendung und Klausur. Ein tieferes Verständnis von Inhalten und Zusammenhängen benötigen nur Mathematiker.
- Mein Stundenplan lässt mir nur wenig Zeit. Ich arbeite am Ende des Semesters einfach alles 'am Stück' nach. Ein, zwei Wochen (oder gar Tage) werden schon reichen.
- Bei der Klausur sind so viele Hilfsmittel zugelassen, dass ich auf Lernen und Üben verzichten kann.

Folien

- Die in der Vorlesung gezeigten Folien werden auf der Vorlesungswebseite im Voraus zum Herunterladen bereitgestellt.
- Diese stellen **kein** Vorlesungsskript dar, sondern sollen lediglich den Mitschreibaufwand minimieren. Eigene Ergänzungen bzw. Randnotizen sind unerlässlich.
- Der Download ersetzt natürlich nicht den Besuch von Vorlesung/Übung und auch nicht die Lektüre von Fachliteratur.

Einige Lehrbücher

- M. Schubert: Mathematik für Informatiker. Springer-Vieweg, 2012. **E**
- M. Drmota, B. Gittenberger, G. Karigl und A. Panholzer: Mathematik für Informatiker. Heldermann Verlag, 2008
- L. Papula: Mathematik für Ingenieure und Naturwissenschaftler. Vieweg+Teubner, Band 1–3, 2009. **E**

Danksagung: Diese Vorlesung entstand nach Vorlagen meines Kollegen Prof. Oliver Ernst.

Doch nun ...

lassen Sie uns endlich zur Mathematik kommen! Immerhin eine der ältesten wissenschaftlichen Beschäftigungen der Menschen überhaupt.



Papyrus Rhind

(ca. 1650 v. Chr.)

Approximation von π , Kreisfläche, Auflösen linearer Gleichungen

① Grundlagen

- 1.1 Elemente der Aussagenlogik
- 1.2 Elemente der Mengenlehre
- 1.3 Die reellen Zahlen
- 1.4 Natürliche Zahlen und Induktionsprinzip
- 1.5 Abbildungen und Funktionen
- 1.6 Komplexe Zahlen

② Lineare Algebra

- 2.1 Vektorräume
- 2.2 Matrizen und lineare Abbildungen
- 2.3 Lineare Gleichungssysteme
- 2.4 Determinanten
- 2.5 Invertierbare Matrizen
- 2.6 Orthogonalität, Skalarprodukt und Norm
- 2.7 Kreuz- und Spatprodukt
- 2.8 Elemente der analytischen Geometrie
- 2.9 Orthogonale Abbildungen
- 2.10 Eigenwerte und Eigenvektoren

2.11 Singulärwertzerlegung

- ① Grundlagen
- ② Lineare Algebra

Wie sag' ich's?

Werfen Sie einen kurzen Blick auf folgende lyrische Einlassung:

$$f : \mathbb{R} \rightarrow \mathbb{R} \text{ stetig in } x_0$$
$$:\Leftrightarrow (\forall \epsilon > 0 \exists \delta > 0 : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon)$$

Möglicherweise bewegen Sie jetzt Fragen wie

- Wie kann man jemals einen solchen Ausdruck durchschauen?
- Wozu muss man sich überhaupt derart kompliziert ausdrücken?

Wie sag' ich's?

Werfen Sie einen kurzen Blick auf folgende lyrische Einlassung:

$$f : \mathbb{R} \rightarrow \mathbb{R} \text{ stetig in } x_0 \\ :\Leftrightarrow (\forall \epsilon > 0 \exists \delta > 0 : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon)$$

Möglicherweise bewegen Sie jetzt Fragen wie

- Wie kann man jemals einen solchen Ausdruck durchschauen?
- Wozu muss man sich überhaupt derart kompliziert ausdrücken?

Lassen Sie uns dazu ein kleines Experiment durchführen:

Erklären Sie Ihrem Nachbarn in einer Minute, was eine reelle Zahl ist, und warum für diese das Kommutativgesetz der Addition gilt.

Wie sag' ich's?

Werfen Sie einen kurzen Blick auf folgende lyrische Einlassung:

$$f : \mathbb{R} \rightarrow \mathbb{R} \text{ stetig in } x_0 \\ \Leftrightarrow (\forall \epsilon > 0 \exists \delta > 0 : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon)$$

Möglicherweise bewegen Sie jetzt Fragen wie

- Wie kann man jemals einen solchen Ausdruck durchschauen?
- Wozu muss man sich überhaupt derart kompliziert ausdrücken?

Lassen Sie uns dazu ein kleines Experiment durchführen:

Erklären Sie Ihrem Nachbarn in einer Minute, was eine reelle Zahl ist, und warum für diese das Kommutativgesetz der Addition gilt.

Wahrscheinlich (hoffentlich) ist Ihnen aufgefallen, dass die Umgangssprache für eine präzise Beschreibung nicht ausreicht. Wir brauchen eine Fachsprache zur mathematischen Kommunikation. Diese muss erst erlernt werden.

① Grundlagen

- 1.1 Elemente der Aussagenlogik
- 1.2 Elemente der Mengenlehre
- 1.3 Die reellen Zahlen
- 1.4 Natürliche Zahlen und Induktionsprinzip
- 1.5 Abbildungen und Funktionen
- 1.6 Komplexe Zahlen

② Lineare Algebra

Elemente der Aussagenlogik

Aussagen

Eine **Aussage** p ist ein sprachliches oder formelmäßiges Gebilde, dem man entweder den **Wahrheitswert** 'wahr' oder 'falsch' zuordnen kann.

Sprechweisen

- ' p ist eine wahre (richtige) Aussage.', ' p gilt.'
- ' p ist eine falsche Aussage.', ' p gilt nicht.'

Insbesondere kann keine Aussage gleichzeitig wahr und falsch sein.

Sind folgende Konstrukte Aussagen? Was ist ggf. ihr Wahrheitswert?

'Wir sprechen gerade über Aussagenlogik.';

'Wie soll ich bloß die Prüfung schaffen?';

' $6^2 = 36$ ';

' $7 = \sqrt{48}$ ';

' $\sqrt{14}$ ';

'Rettet den Euro!';

'Wir befinden uns in Dresden oder Chemnitz.'

'Diese Aussage ist falsch.'

Elemente der Aussagenlogik

Verknüpfung von Aussagen

Eine **Aussagenverknüpfung** liefert eine neue Aussage, deren Wahrheitswert sich aus den Wahrheitswerten der verknüpften Aussagen ergibt.

Seien p und q Aussagen, dann heißt

- \bar{p} (auch $\neg p$) **Negation** von p (sprich: 'nicht p '),
- $p \vee q$ **Disjunktion** von p und q (sprich: ' p oder q '),
- $p \wedge q$ **Konjunktion** von p und q (sprich: ' p und q ').

Die Wahrheitswerte werden durch folgende Tabellen festgeschrieben:

p	\bar{p}
w	f
f	w

p	q	$p \vee q$	$p \wedge q$
w	w	w	w
w	f	w	f
f	w	w	f
f	f	f	f

Beachten Sie den rot markierten Wahrheitswert bei 'oder' – umgangssprachlich meint man mitunter nicht dasselbe!

Geben Sie den Wahrheitswert der Aussage

'Wir gehen heute abend ins Kino oder ins Theater'

in Abhängigkeit von der tatsächlich durchgeführten Aktivitäten an. Machen Sie sich den Unterschied zur Umgangssprache klar.

Geben Sie die Wahrheitswerte und Negationen zu folgenden Aussagen an:

- p : 'Unsere Vorlesung findet heute am Nachmittag statt.'
- q : 'Unsere Vorlesung findet heute am Abend statt.'

Warum ist q nicht die Negation von p ?

Elemente der Aussagenlogik

Implikation und Äquivalenz

Ein zentrales Element beim Aufbau mathematischer Theorie ist das logische Schließen. Dafür benötigen wir zwei weitere Verknüpfungen.

Seien p und q Aussagen; dann heißt

- $p \Rightarrow q$ **Implikation** ('Aus p folgt q '; 'Wenn p gilt, so gilt q .'),
- $p \Leftrightarrow q$ **Äquivalenz** (' p gilt **genau** dann, wenn q gilt.').

Eine Implikation ist genau dann falsch, wenn man aus **wahren** Aussagen die **falschen** Schlüsse zieht (p wahr, q falsch).

Eine Äquivalenz bedeutet, dass p und q immer den gleichen Wahrheitswert besitzen.

Die zugehörige Wahrheitstabelle spielt für den Praktiker kaum eine Rolle – die sprachliche Fassung der Begriffe ist handlicher:

p	q	$p \Rightarrow q$	$q \Rightarrow p$	$p \Leftrightarrow q$
w	w	w	w	w
w	f	f	w	f
f	w	w	f	f
f	f	w	w	w

Wichtig sind aber folgende alternative Sprechweisen:

- für $p \Rightarrow q$: ' p ist **hinreichend** für q .' oder ' q ist **notwendig** für p .'
- für $p \Leftrightarrow q$: ' p ist hinreichend und notwendig für q .'

und folgender Satz:

Satz 1.1

Seien p und q Aussagen. Dann gilt:

- $(p \Leftrightarrow q) \Leftrightarrow (p \Rightarrow q) \text{ und } (q \Rightarrow p)$
- $(p \Rightarrow q) \Leftrightarrow (\bar{q} \Rightarrow \bar{p})$

Die zweite Aussage bildet die Grundlage für das indirekte Beweisen ('Angenommen q gilt nicht, dann kann auch p nicht gelten.').

Interpretieren Sie die Aussage 'Wenn n^2 gerade ist, so ist auch n gerade.' vor dem Hintergrund der zweiten Beziehung in Satz 1.1.

Satz 1.2

Es seien p , q und r beliebige Aussagen. Dann gelten immer:

- $\overline{\overline{p}} \Leftrightarrow p$ (doppelte Verneinung),
- $p \vee \overline{p}$ (Satz vom ausgeschlossenen Dritten),
- $(p \vee q) \Leftrightarrow (q \vee p)$
 $(p \wedge q) \Leftrightarrow (q \wedge p)$ (Kommutativgesetze),
- $[(p \vee q) \vee r] \Leftrightarrow [p \vee (q \vee r)],$
 $[(p \wedge q) \wedge r] \Leftrightarrow [p \wedge (q \wedge r)]$ (Assoziativgesetze),
- $[p \wedge (q \vee r)] \Leftrightarrow [(p \wedge q) \vee (p \wedge r)],$
 $[p \vee (q \wedge r)] \Leftrightarrow [(p \vee q) \wedge (p \vee r)]$ (Distributivgesetze),
- $\overline{p \vee q} \Leftrightarrow (\overline{p} \wedge \overline{q}),$
 $\overline{p \wedge q} \Leftrightarrow (\overline{p} \vee \overline{q})$ (De Morgansche Regeln).

Punkte 1-4 scheinen sofort klar. Im Zweifel erfolgt der Nachweis solcher Äquivalenzen durch Aufstellen von Wahrheitstabellen.

Elemente der Aussagenlogik

Exkurs: Anwendung in der Digitaltechnik

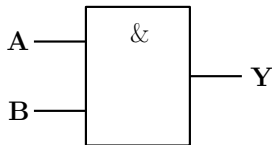
In der Digitaltechnik gibt es genau zwei Zustände (hohe/niedrige(keine) Spannung; 1/0), die man als logisches 'wahr' bzw. 'falsch' interpretiert.

Schaltungen bestehen häufig aus 'Logikgattern' (engl. *gates*).

Diese besitzen z. B. 2 Eingänge (A,B) und einen Ausgang (Y) und reproduzieren logische Verknüfungen.

Eine wichtige Anwendung von Satz 1.2 liegt in der Analyse und Vereinfachung logischer Schaltnetzwerke.

Beispiel: Schaltsymbol und Wahrheitstabelle für ein UND-Gatter



A	B	Y
1	1	1
1	0	0
0	1	0
0	0	0

Elemente der Aussagenlogik

Aussageformen, Quantoren

Eine **Aussageform** ist ein Gebilde, welches eine oder mehrere Variablen enthält, und das nach dem Ersetzen der Variablen durch konkrete Werte in eine Aussage übergeht.

Schreibweise: $p(x)$, falls x als Variable verwendet wird.

Mit $p(x)$ sei die Aussageform $x \leq 1$ bezeichnet, wobei x eine reelle Variable sei. Wie lauten die Aussagen $p(0)$ und $p(4)$ und was ist ihr Wahrheitswert?

Eine weitere Möglichkeit, Aussageformen in Aussagen zu überführen, ist die Bindung der Variablen an **Quantoren**. Die wichtigsten sind:

- \forall – ‘für alle ... gilt ...’
- \exists – ‘es existiert (mindestens) ein ..., so dass ...’

Beispiel: ‘Für alle reellen Zahlen x gilt $x > 1$.’

‘Es gibt eine reelle Zahl x mit $x^2 = -1$.’

(Beide Aussagen sind natürlich falsch.)

Definition 1.3

Sei $p(x)$ eine Aussageform; dann meinen wir mit

- der Negation von 'Für alle x gilt $p(x)$ ' die Aussage 'Es existiert ein x , für das $p(x)$ nicht gilt'; symbolisch:

$$\overline{\forall x : p(x)} := (\exists x : \overline{p(x)})$$

- der Negation von 'Es existiert ein x , für das $p(x)$ gilt' die Aussage 'Für alle x gilt $p(x)$ nicht'; symbolisch:

$$\overline{\exists x : p(x)} := (\forall x : \overline{p(x)})$$

Natürlich erfolgt diese Definition im Einklang mit unserer Intuition – die Intuition wird lediglich sinnvoll formalisiert.

Formulieren Sie die Negationen von 'Für alle reellen Zahlen x gilt $x > 1$.', 'Es gibt eine reelle Zahl x mit $x^2 = -1$.', 'Jeder Haushalt gibt ein Drittel seines Einkommens für Wohnen aus.' (SZ, 8.9.12) und 'Alle Politiker lügen gelegentlich.'

Elemente der Aussagenlogik

Axiome und Sätze

Gerüstet mit diesen Vorkenntnissen lassen sich die gebräuchlichen Sprachkonstrukte der Mathematik grob typisieren.

Ein **Axiom** ist eine Aussage, die innerhalb einer Theorie nicht begründet oder hergeleitet wird. Axiome werden beweislos vorausgesetzt. Wichtige Eigenschaften von Axiomensystemen sind **Widerspruchsfreiheit** und **Vollständigkeit**.¹

Beispiel: Kommutativgesetz für reelle Zahlen, siehe später.

Ein **Satz** ist eine Aussage, die mittels eines Beweises als wahr erkannt wurde. Den Beweis bildet dabei eine Kette von Schlussfolgerungen, die nur von Axiomen oder bereits bewiesenen Sätzen ausgeht.

Beispiel: ' $\sqrt{2}$ ist keine rationale Zahl.'

Gebräuchliche Ausprägungen von Sätzen sind auch

- **Lemma** (Hilfssatz) – ein Satz, der als Zwischenstufe in einem Beweis dient,
- **Korollar** (Folgerung) – ein Satz, der ohne großen Beweisaufwand aus einem anderen folgt.

Die Abgrenzung zum Satz ist jeweils fließend und subjektiv.

¹vgl. D. Hofstadter: *Gödel, Escher Bach*. Basic Books, New York NY 1979

Eine **Definition** dient zur Festlegung eines mathematischen Begriffs. Hier sind eher Wertungen wie 'sinnvoll/nicht sinnvoll' statt 'wahr/falsch' angebracht. Sinnvolle Definitionen erfüllen unter anderem folgende Anforderungen:

- Widerspruchsfreiheit,
- Zirkelfreiheit (das zu Definierende darf nicht selbst Bestandteil der Definition sein),
- Angemessenheit (nicht zu eng und nicht zu weit gefasst),
- Redundanzfreiheit (keine Bestandteile enthalten, die aus anderen logisch folgen).

Häufig schreibt man Definitionen in der Form von Gleichungen oder Äquivalenzen und verwendet die Symbole $:=$ und $:\Leftrightarrow$. Der zu definierende Ausdruck steht immer auf der Seite des Doppelpunkts.

Beispiel: $a^0 := 1$ für $a \in \mathbb{R}$.

Elemente der Aussagenlogik

Zum mathematischen Wahrheitsbegriff

Da in der Mathematik **alle** Aussagen per logischem Schluss bewiesen werden, entsteht ein sehr scharfer Wahrheitsbegriff.

In den (experimentellen) Naturwissenschaften ist der Wahrheitsbegriff stärker an Beobachtungen gekoppelt und daher weniger scharf (**deduktive** vs. **induktive** Schlussweisen).

Die Mathematik ist also im logischen Sinne präziser, dies erfordert aber eben auch eine kompliziertere Fachsprache.

Für unsere Vorlesung: Ein strenger Aufbau der Theorie kostet Zeit und ist nicht Aufgabe des Anwenders, daher folgender Kompromiss:

- Reihenfolge von Axiomen, Sätzen und Definitionen entspricht weitgehend dem strengen Aufbau der Theorie,
- Beweise und Beweisideen werden dort skizziert, wo sie dem Verständnis des Stoffes dienen,
- 'technische' und aufwändige Beweise werden weggelassen – verwenden Sie bei Interesse entsprechende Literatur.

① Grundlagen

- 1.1 Elemente der Aussagenlogik
- 1.2 Elemente der Mengenlehre**
- 1.3 Die reellen Zahlen
- 1.4 Natürliche Zahlen und Induktionsprinzip
- 1.5 Abbildungen und Funktionen
- 1.6 Komplexe Zahlen

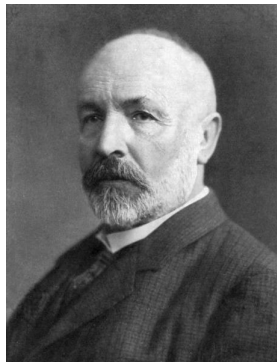
② Lineare Algebra

Definition 1.4 ('naive' Mengendefinition, Cantor*, 1895)

Eine **Menge** M ist eine Zusammenfassung von bestimmten, wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen.

Die erwähnten Objekte heißen **Elemente** dieser Menge M .

Schreibweise: $x \in M$ (bzw. $x \notin M$).



* Georg Cantor (1845-1918), deutscher Mathematiker, lieferte u.a. wichtige Beiträge zur Mengenlehre.

Elemente der Mengenlehre

Darstellungsmöglichkeiten für Mengen

- aufzählende Darstellung – Elemente werden explizit aufgelistet, z.B.:

$$A = \{\text{Chemnitz; Dresden; Freiberg; Leipzig}\},$$

$$B = \{2; 3; 5; 7\}.$$

- beschreibende Darstellung – Elemente werden durch eine Eigenschaft charakterisiert, z.B.:

$$A = \{x : x \text{ ist sächsische Universitätsstadt.}\},$$

$$B = \{p : p \text{ ist Primzahl und } p < 10\}.$$

Die Gesamtheit, die kein Element enthält heißt **leere Menge**.

Symbol: \emptyset oder $\{\}$.

Elemente der Mengenlehre

Mengenbeziehungen

Seien A und B Mengen.

- A heißt **Teilmenge** von B , wenn jedes Element von A auch Element von B ist. Schreibweise: $A \subseteq B$.
- A und B heißen **gleich**, wenn jedes Element von A auch Element von B ist und umgekehrt ($A \subseteq B$ und $B \subseteq A$).
Schreibweise: $A = B$.
- A heißt **echte Teilmenge** von B , wenn jedes Element von A auch Element von B ist, und ein Element von B existiert, das nicht zu A gehört. ($A \subseteq B$ und $A \neq B$). Schreibweise: $A \subset B^\dagger$.

Die Teilmengen einer Menge M können wiederum zur sogenannten **Potenzmenge** zusammengefasst werden:

$$\mathcal{P}(M) := \{A : A \subseteq M\}.$$

Man gebe die Potenzmenge von $A = \{1; 2; 4\}$ an.

[†]Aufgepasst: Manche Autoren benutzen das Symbol \subset auch für die allgemeine Teilmengenbeziehung (Gleichheit eingeschlossen) und für die echte Teilmengenbeziehung \subsetneq .

Elemente der Mengenlehre

Mengenoperationen

Seien A und B Teilmengen einer Grundmenge X . Dann heißt

- $A \cup B := \{x : x \in A \text{ oder } x \in B\}$ die **Vereinigung** von A und B ,
- $A \cap B := \{x : x \in A \text{ und } x \in B\}$ der **Durchschnitt** von A und B ,
- $A \setminus B := \{x : x \in A \text{ und } x \notin B\}$ die (mengentheoretische) **Differenz** von A und B (sprich 'A ohne B'),
- $\bar{A} := X \setminus A$ das Komplement von A in X .

Geben Sie zu $A = \{1; 2; 3; 4; 5; 6\}$ und $B = \{0; 4; 5\}$ die Mengen $A \cup B$, $A \cap B$, $A \setminus B$ und $B \setminus A$ an.

Wie lauten Vereinigung und Durchschnitt der Mengen

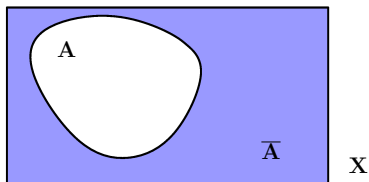
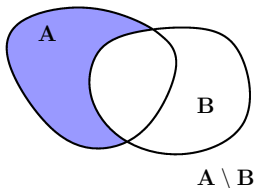
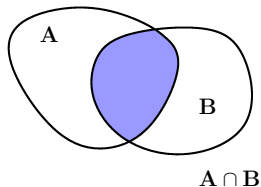
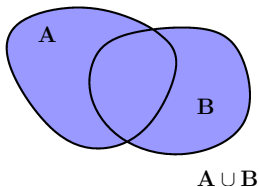
$$C = \{p : p \text{ ist Primzahl}\} \text{ und}$$

$$D = \{p : p \text{ ist Primzahl und } p < 10\}?$$

Elemente der Mengenlehre

Visualisierung

Visualisierung von Mengenoperationen und -beziehungen erfolgt häufig im sogenannten **Venn-Diagramm**:



Elemente der Mengenlehre

Rechenregeln für Mengenoperationen

Vereinigung (\cup), Schnitt (\cap) bzw. Komplement wurden letztlich über die logischen Junktoren 'oder' (\vee), 'und' (\wedge) bzw. Negation definiert. Daher gelten für Mengenoperationen analoge Regeln zu Satz 1.2:

Satz 1.5 (Regeln für das Operieren mit Mengen)

Es seien A, B, C Mengen. Dann gelten:

- $A \cap B = B \cap A, \quad A \cup B = B \cup A,$
- $A \cap (B \cap C) = (A \cap B) \cap C, \quad A \cup (B \cup C) = (A \cup B) \cup C,$
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C), \quad A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$
- $A \cap A = A, \quad A \cup A = A,$
- $A \cap \emptyset = \emptyset, \quad A \cup \emptyset = A,$
- $A \setminus B = A \Leftrightarrow A \cap B = \emptyset, \quad A \setminus B = \emptyset \Leftrightarrow A \subseteq B,$
- $A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C),$
 $A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C).$

Elemente der Mengenlehre

Kartesisches Produkt

Seien A und B Mengen. Dann heißt die Menge aller geordneten Paare (a, b) mit $a \in A$ und $b \in B$ **kartesisches Produkt** von A und B , kurz

$$A \times B := \{(a, b) : a \in A \text{ und } b \in B\}$$

Analog definiert man für Mengen A_1, A_2, \dots, A_n

$$A_1 \times A_2 \times \cdots \times A_n := \{(a_1, a_2, \dots, a_n) : a_1 \in A_1, a_2 \in A_2, \dots, a_n \in A_n\}$$

und schreibt im Falle der Gleichheit aller A_i kürzer

$$A^n := A \times A \times \cdots \times A \quad (n \text{ Faktoren}).$$

Als wichtigste Beispiele werden uns \mathbb{R}^2 und \mathbb{R}^3 begegnen – das sind die Entsprechungen von Anschauungsebene und -raum.

Relationen

Äquivalenzrelationen

Mit **Relationen** werden Beziehungen zwischen Objekten ausgedrückt. Eine (binäre) Relation R auf einer Grundmenge M ist eine Teilmenge des kartesischen Produkts $M \times M$, also

$$R \subseteq M \times M.$$

Wichtige Eigenschaften von Relationen sind

- (a) **Reflexivität.** $\forall x \in M$ gilt $(x, x) \in R$.
- (b) **Symmetrie.** $\forall x, y \in M$ mit $(x, y) \in R \Rightarrow (y, x) \in R$.
- (c) **Transitivität.** $\forall x, y, z \in M$ mit $(x, y) \in R \wedge (y, z) \in R \Rightarrow (x, z) \in R$.

Eine Relation, die alle drei Eigenschaften erfüllt, heißt **Äquivalenzrelation**.

Für $(x, y) \in R$ schreibt man auch xRy (manchmal auch $x \sim y$).

Eine **Partition** \mathcal{P} einer Menge M bezeichnet eine Zerlegung in disjunkte Teilmengen, d.h.

$$M = \bigcup_{T \in \mathcal{P}} T \quad \text{und} \quad S \cap T = \emptyset \quad \text{für je zwei verschiedene } S, T \in \mathcal{P}.$$

Jede Partition einer Menge erzeugt eine Äquivalenzrelation auf ihr und umgekehrt.

a) $R = \{(1, 1), (2, 2), (3, 3), (1, 2), (3, 1)\}$

- **reflexiv:** $1, 2, 3 \in X$ und $(1, 1), (2, 2), (3, 3) \in R \implies$ reflexiv
- **transitiv:** $(3, 1), (1, 2) \in R$ aber $(3, 2) \notin R \implies$ nicht transitiv
- **symmetrisch:** $(1, 2) \in R$ aber $(2, 1) \notin R \implies$ nicht symmetrisch
- **antisymmetrisch:** für $(1, 2), (3, 1) \in R$ ist $(2, 1), (1, 3) \notin R$
und $(1, 1), (2, 2), (3, 3) \in R \implies 1 = 1, 2 = 2, 3 = 3 \implies$ antisymmetrisch

also ist R reflexiv und antisymmetrisch

b) $R = \{(1, 1), (2, 2), (1, 2), (2, 1)\}$

- **reflexiv:** $3 \in X$ aber $(3, 3) \notin R \implies$ nicht reflexiv

- **transitiv:**

- $(1, 2), (2, 1) \in R \implies (1, 1) \in R$
- $(2, 1), (1, 2) \in R \implies (2, 2) \in R$
- $(1, 1), (1, 1) \in R \implies (1, 1) \in R$
- $(2, 2), (2, 2) \in R \implies (2, 2) \in R$

\implies transitiv

- **symmetrisch:**

- $(1, 1) \in R \implies (1, 1) \in R$
- $(2, 2) \in R \implies (2, 2) \in R$
- $(1, 2) \in R \implies (2, 1) \in R$
- $(2, 1) \in R \implies (1, 2) \in R$

\implies symmetrisch

- **antisymmetrisch:** $(1, 2), (2, 1) \in R \not\Rightarrow 1 = 2 \implies$ nicht antisymmetrisch

also ist R transitiv und symmetrisch

b) $R = \{(1, 1)\}$

- **reflexiv:** $3 \in X$ aber $(3, 3) \notin R \implies$ nicht reflexiv
- **transitiv:** $(1, 1), (1, 1) \in R \implies (1, 1) \in R \implies$ transitiv
- **symmetrisch:** $(1, 1) \in R \implies (1, 1) \in R \implies$ symmetrisch
- **antisymmetrisch:** $(1, 1) \in R \wedge (1, 1) \in R \implies 1 = 1, \implies$ antisymmetrisch

also ist R transitiv, symmetrisch und antisymmetrisch

Betrachten die Menge $M = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ und dann die Zerlegung der Menge in

$$M_1 = \{3, 6, 9\}, M_2 = \{2, 5, 8\}, M_3 = \{1, 4, 7\}.$$

Gilt

$$M = \bigcup_{i=1,2,3} M_i?$$

Ist das eine Partition? Was für eine Relation könnten Sie sich vorstellen?

Weitere wichtige Eigenschaften von Relationen:

(d) **Antisymmetrie**. $\forall x, y \in M$ mit $xRy \wedge yRx \Rightarrow x = y$.

(e) **Asymmetrie**. $\forall x, y \in M$ mit $xRy \Rightarrow \neg(yRx)$.

Eine Relation auf M heißt **Ordnungsrelation**, wenn sie reflexiv, antisymmetrisch und transitiv ist.

Beispiel: \leq, \subseteq .

Eine Ordnungsrelation auf M heißt **vollständig** oder **linear**, falls für alle $x, y \in M$ gilt: $xRy \vee yRx$.

Eine Relation auf M heißt **strenge Ordnungsrelation**, wenn sie asymmetrisch und transitiv ist.

Beispiel: $<$.

Eine strenge Ordnungsrelation heißt **vollständig**, wenn für alle $x, y \in M$ mit $x \neq y$ gilt $xRy \vee yRx$.

① Grundlagen

- 1.1 Elemente der Aussagenlogik
- 1.2 Elemente der Mengenlehre
- 1.3 Die reellen Zahlen**
- 1.4 Natürliche Zahlen und Induktionsprinzip
- 1.5 Abbildungen und Funktionen
- 1.6 Komplexe Zahlen

② Lineare Algebra

Die reellen Zahlen

Erinnern Sie sich an unser Experiment in der allerersten Stunde?

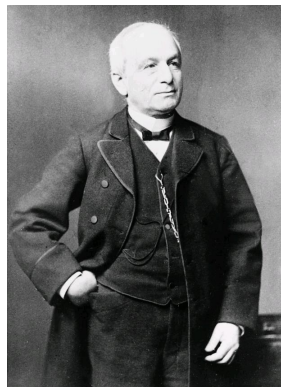
Obwohl **jeder** von uns schon seit Jahren reelle Zahlen verwendet, fällt es uns extrem schwer, das Wesen der Zahl in Worte zu fassen.

Es lohnt sich also eine nähere Betrachtung – gerade auch weil Messwerte für viele physikalische Größen als reelle Zahlen aufgefasst werden können.

Einen (eher historisch interessanten) Anfangspunkt liefert uns folgendes berühmte wie umstrittene Zitat:

'Die natürlichen Zahlen hat Gott gemacht, alles andere ist Menschenwerk.'

(Leopold Kronecker, deutscher Mathematiker, 1823–1891)



Skizzieren wir zunächst das klassische Vorgehen über Zahlbereichserweiterungen.

Wir starten mit

- $\mathbb{N} := \{1, 2, 3, \dots\}$ – **natürliche Zahlen**
(‘hat Gott gemacht’ oder werden über Peano-Axiome beschrieben)
- $\mathbb{N}_0 := \mathbb{N} \cup \{0\} = \{0, 1, 2, 3, \dots\}$

Nach Einführung der Addition stellt man fest, dass Gleichungen wie $x + 9 = 1$ in \mathbb{N} nicht lösbar sind, daher Erweiterung zu

- $\mathbb{Z} := \{\dots, -2, -1, 0, 1, 2, \dots\}$ – **ganze Zahlen**

Nimmt man die Multiplikation hinzu, sind wiederum Gleichungen wie $(-2) \cdot x = 1$ in \mathbb{Z} nicht lösbar, daher Erweiterung zu

- $\mathbb{Q} := \left\{ \frac{p}{q} \mid p, q \in \mathbb{Z}, q \neq 0 \right\}$ – **rationale Zahlen** (alle endlichen und periodischen Dezimalbrüche)

Wählt man die **Zahlengerade** als Modell, so kann man mit den rationalen Zahlen zumindest beliebig feine Einteilungen realisieren.

Die reellen Zahlen

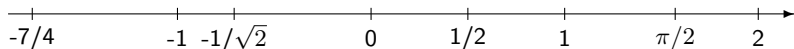
Lücken in \mathbb{Q}

Aber: Bereits Euklid (ca. 360-280 v. Chr.) überlieferte uns den Beweis zu

Satz 1.6

$\sqrt{2} \notin \mathbb{Q}$, d.h. $\sqrt{2}$ ist keine rationale Zahl.

Die rationalen Zahlen erfassen also nicht die gesamte Zahlengerade. Erst durch 'Vervollständigen' gelangt man zu den **reellen Zahlen** \mathbb{R} .



Jedem Punkt der Zahlengeraden entspricht genau eine reelle Zahl.

Die erste exakte Formulierung dieses Vervollständigungsprozesses stammt übrigens nicht aus der Antike. Dies gelang erst Karl Weierstraß (deutscher Mathematiker, 1815-1897).



Die reellen Zahlen

Moderner axiomatischer Ansatz

Klassische Zahlbereichserweiterungen sind anschaulich, erweisen sich aber im Detail als kompliziert, z. B. bei

- Einführung von Addition und Multiplikation sowie der damit verbundenen Rechengesetze,
- Einführung einer Ordnungsrelation ('größer/kleiner'),
- Ausformulieren des Vervollständigungsschritts.

Moderne Ansätze definieren daher \mathbb{R} direkt über Axiome, die unmittelbar die Rechengesetze liefern. Man benötigt:

- Körperaxiome (liefern Rechengesetze),
- Anordnungsgesetze (liefern Ordnungsrelation),
- Vollständigkeitsaxiom.

\mathbb{N} , \mathbb{Z} und \mathbb{Q} werden als entsprechende Teilmengen von \mathbb{R} aufgefasst.

Wir verzichten auf eine formale Angabe der Axiome, geben aber die unmittelbar resultierenden Gesetze an.

Die reellen Zahlen

Die arithmetischen Gesetze in \mathbb{R}

- | | |
|--|---|
| (1) $(a+b)+c = a+(b+c) \quad \forall a, b, c \in \mathbb{R}$ | (Assoziativgesetz der Addition), |
| (2) $a + 0 = a \quad \forall a \in \mathbb{R}$ | (neutrales Element der Addition), |
| (3) $a + (-a) = 0 \quad \forall a \in \mathbb{R}$ | (inverse Elemente der Addition), |
| (4) $a + b = b + a \quad \forall a, b \in \mathbb{R}$ | (Kommutativgesetz der Addition), |
| (5) $(ab)c = a(bc) \quad \forall a, b, c \in \mathbb{R}$ | (Assoziativgesetz der Multiplikation), |
| (6) $a \cdot 1 = a \quad \forall a \in \mathbb{R}$ | (neutrales Element der Multiplikation), |
| (7) $a \frac{1}{a} = 1 \quad \forall a \in \mathbb{R}, a \neq 0$ | (inverse Elemente der Multiplikation), |
| (8) $ab = ba \quad \forall a, b \in \mathbb{R}$ | (Kommutativgesetz der Multiplikation), |
| (9) $a(b + c) = ab + ac \quad \forall a, b, c \in \mathbb{R}$ | (Distributivgesetz). |

Diese Gesetze korrespondieren mit den **Körperaxiomen**.

Eine Menge die die Körperaxiome erfüllt, heißt **(Zahl-)Körper**.

Die reellen Zahlen

Die arithmetischen Gesetze in \mathbb{R}

Wir definieren für $a \in \mathbb{R}$ und $n \in \mathbb{N}$:

$$a^n := a \cdot a \cdot \dots \cdot a \quad (n \text{ Faktoren}), \quad a^0 := 1$$

und, falls $a \neq 0$:

$$a^{-n} := \frac{1}{a^n}.$$

Dann folgen aus den arithmetischen Gesetzen weitere Rechenregeln:

Satz 1.7

Für alle $a, b \in \mathbb{R}$ gelten

- $a \cdot 0 = 0$,
- $ab = 0 \Rightarrow a = 0$ oder $b = 0$,
- $a^2 = b^2 \Leftrightarrow a = b$ oder $a = -b$,
- $-(-a) = a$, $-(a + b) = -a - b$,
- $(-a)b = a(-b) = -ab$, $(-a)(-b) = ab$,
- $a^n a^m = a^{n+m} \quad \forall m, n \in \mathbb{Z}$ (falls $a \neq 0$),
- $a^n b^n = (ab)^n \quad \forall n \in \mathbb{Z}$ (falls $a, b \neq 0$).

Weiterhin lassen sich mit den arithmetischen Gesetzen folgende wohlbekannte Aussagen beweisen:

Satz 1.8 (Binomische Formeln)

Für alle $a, b \in \mathbb{R}$ gilt

- $(a + b)^2 = a^2 + 2ab + b^2$,
- $(a - b)^2 = a^2 - 2ab + b^2$,
- $(a + b)(a - b) = a^2 - b^2$.

Natürlich kann man auch Binome höheren Grades so auswerten. Z. B. gilt

$$(a \pm b)^3 = a^3 \pm 3a^2b + 3ab^2 \pm b^3.$$

Einen allgemeingültige Formel für $(a + b)^n$, $n \in \mathbb{N}$, finden Sie in der Literatur unter dem Namen 'Binomischer Lehrsatz'.

Die reellen Zahlen

Die Ordnungsrelation in \mathbb{R}

(10) Zwischen zwei reellen Zahlen a und b besteht immer genau eine der folgenden drei Größenbeziehungen:

$$a < b \text{ (} a \text{ kleiner } b\text{), } a = b \text{ (} a \text{ gleich } b\text{), } a > b \text{ (} a \text{ größer } b\text{)}.$$

(11) $(a < b \text{ und } b < c) \Rightarrow a < c \quad \forall a, b, c \in \mathbb{R}$ (Transitivität),

(12) $a < b \Leftrightarrow a + c < b + c \quad \forall a, b, c \in \mathbb{R}$ (Monotonie der Addition),

(13) $(a < b \text{ und } 0 < c) \Leftrightarrow ac < bc \quad \forall a, b, c \in \mathbb{R}$
(Monotonie der Multiplikation).

Definition:

$a \leq b$ (a kleiner oder gleich b) bedeutet $a < b$ oder $a = b$.

$a \geq b$ (a größer oder gleich b) bedeutet $a > b$ oder $a = b$.

Diese Gesetze korrespondieren mit den **Anordnungsaxiomen**.

Körper, die die Anordnungsaxiome erfüllen, heißen **geordnete Körper**. Beispiele sind \mathbb{R} , aber auch \mathbb{Q} .

Die reellen Zahlen

Die Ordnungsrelation in \mathbb{R}

Aus den Anordnungsaxiomen folgen weiterhin die Regeln für den Umgang mit Ungleichungen:

Satz 1.9

Für $a, b, c, d \in \mathbb{R}$ gelten:

- $a \leq b$ und $c \leq d \Rightarrow a + c \leq b + d$,
- $a \leq b$ und $c < 0 \Rightarrow ac \geq bc$ und $\frac{a}{c} \geq \frac{b}{c}$,
- $a \leq b \Rightarrow -b \leq -a$,
- $0 < a \leq b \Rightarrow 0 < \frac{1}{b} \leq \frac{1}{a}$,
- $a \leq b < 0 \Rightarrow \frac{1}{b} \leq \frac{1}{a} < 0$,
- $a < 0 < b \Rightarrow \frac{1}{a} < 0 < \frac{1}{b}$.

Finden Sie die Lösungen der Ungleichung $\frac{x-1}{x+3} < 2$ ($x \neq -3$).

Definition 1.10 (Obere und untere Schranken)

Sei $M \neq \emptyset$, eine Teilmenge der reellen Zahlen.

- M heißt **nach oben beschränkt**, wenn es eine Zahl $C \in \mathbb{R}$ gibt mit

$$x \leq C \quad \text{für alle } x \in M.$$

Jede solche Zahl C heißt **obere Schranke** von M . Die kleinste obere Schranke einer nach oben beschränkten Menge M heißt **Supremum** von M (Schreibweise: $\sup M$).

- M heißt **nach unten beschränkt**, wenn es eine Zahl $C \in \mathbb{R}$ gibt mit

$$x \geq C \quad \text{für alle } x \in M.$$

Jede solche Zahl C heißt **untere Schranke** von M . Die größte untere Schranke einer nach unten beschränkten Menge M heißt **Infimum** von M (Schreibweise: $\inf M$).

M heißt **beschränkt**, wenn M sowohl nach oben wie auch nach unten beschränkt ist. Gilt $\sup M \in M$ [$\inf M \in M$], so heißt $\sup M$ [$\inf M$] auch **Maximum** [**Minimum**] von M .

Beispiele:

Ist die Menge

$$M_1 := \left\{ \frac{1}{n} : n \in \mathbb{N} \right\} \subset \mathbb{R}$$

beschränkt? Geben Sie falls möglich obere und untere Schranken, Supremum und Infimum sowie Maximum und Minimum an.

Betrachten wir weiterhin

$$M_2 := \{q \in \mathbb{Q} : q^2 < 2\} \subset \mathbb{Q}.$$

Offenbar ist M_2 nach oben beschränkt, denn 42 ist obere Schranke. Ein Supremum besitzt die Menge in den **rationalen** Zahlen jedoch nicht, da $\sqrt{2} \notin \mathbb{Q}$.

In den **reellen** Zahlen ist die Angabe des Supremums jedoch unproblematisch: $\sup M_2 = \sqrt{2}$.

Das letzte Beispiel liefert den Schlüssel für den noch fehlenden Vervollständigungsschritt bei der Konstruktion von \mathbb{R} .

(14) Jede nach oben beschränkte Menge reeller Zahlen besitzt ein Supremum in \mathbb{R} .

Folgerungen:

- Jede nach unten beschränkte Menge reeller Zahlen besitzt ein Infimum in \mathbb{R} .
- Jede beschränkte Menge reeller Zahlen besitzt ein Supremum und ein Infimum in \mathbb{R} .

Damit ist die axiomatische Konstruktion von \mathbb{R} komplett. Kurz zusammenfassen kann man (1)-(14) in folgendem Satz:

\mathbb{R} ist ein ordnungsvollständiger geordneter Körper.

Intervalle sind 'zusammenhängende' Teilmengen von \mathbb{R} . Es wird dabei nach Inklusion der Randpunkte unterschieden.

Für $a, b \in \mathbb{R}$ mit $a \leq b$:

$$[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\} \quad (\text{abgeschlossenes Intervall}),$$

$$(a, b] := \{x \in \mathbb{R} : a < x \leq b\} \quad (\text{halboffenes Intervall}),$$

$$[a, b) := \{x \in \mathbb{R} : a \leq x < b\} \quad (\text{halboffenes Intervall}),$$

$$(a, b) := \{x \in \mathbb{R} : a < x < b\} \quad (\text{offenes Intervall}),$$

$$[a, \infty) := \{x \in \mathbb{R} : a \leq x\},$$

$$(a, \infty) := \{x \in \mathbb{R} : a < x\},$$

$$(-\infty, b] := \{x \in \mathbb{R} : x \leq b\},$$

$$(-\infty, b) := \{x \in \mathbb{R} : x < b\},$$

$$(-\infty, \infty) := \mathbb{R}.$$

Auf dem Zahlenstrahl lassen sich Intervalle als zusammenhängende Bereiche darstellen.

Der **Betrag** von $a \in \mathbb{R}$ ist durch

$$|a| := \begin{cases} a, & \text{für } a \geq 0, \\ -a, & \text{für } a < 0 \end{cases} \quad \text{definiert.}$$

Satz 1.11 (Rechnen mit Beträgen)

Für $a, b \in \mathbb{R}$, $c \geq 0$ gelten

- $|a| \geq 0$, wobei $|a| = 0 \Leftrightarrow a = 0$,
- $|a| = c \Rightarrow a = c$ oder $a = -c$,
- $|a| = |-a|$ und $|a - b| = |b - a|$,
- $|ab| = |a||b|$,
- $|a + b| \leq |a| + |b|$ (Dreiecksungleichung),
- $|a| \leq c \Leftrightarrow -c \leq a \leq c$.

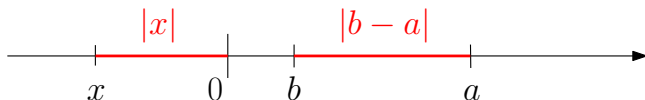
Die reellen Zahlen

Betrag und Abstand

Für den Betrag gibt es eine einfache aber wichtige Interpretation am Zahlenstrahl.

Für $x, a, b \in \mathbb{R}$

- gibt $|x|$ den Abstand von x zur Null an,
- gibt $|a - b| = |b - a|$ den Abstand zwischen a und b an.



Die reellen Zahlen

$\pi = 3.2$ per Gesetz? The 'Indiana Pi Bill'.

Zum Abschluss dieses Abschnitts ein Kuriosum aus der mathematischen Unterhaltungsliteratur. Es geht um eine Gesetzesvorlage vom Ende des 19. Jahrhunderts im Landesparlament des US-Bundesstaates Indiana..

- ab 1892 veröffentlicht der Arzt (und Hobby-Mathematiker) Edward J. Goodwin mehrere Arbeiten zur **Quadratur des Kreises**. Eine Behauptung darin: $\pi = 3.2$
- 1897 Gesetzesvorlage für Parlament von Indiana zur Festlegung von π . Angebot: Staat Indiana darf 'Resultat' gebührenfrei verwenden.
- Entwurf durch Unterhaus zunächst an Ausschuss für Sumpfgebiete³, dann an Bildungsausschuss verwiesen, von letzterem befürwortet.
- Entwurf passiert Unterhaus nach 3 Lesungen mit 67:0 Stimmen.
- Mathematik-Professor Clarence Abiathar Waldo (Purdue University) interveniert, 'berät' einige maßgeblichen Senatoren.
- Entscheidung auf unbestimmte Zeit vertagt (Wert der Vorlage nicht einschätzbar, Sache aber nicht für Gegenstand der Gesetzgebung gehalten)

³"where the bill could find a deserved grave."

① Grundlagen

- 1.1 Elemente der Aussagenlogik
- 1.2 Elemente der Mengenlehre
- 1.3 Die reellen Zahlen
- 1.4 Natürliche Zahlen und Induktionsprinzip**
- 1.5 Abbildungen und Funktionen
- 1.6 Komplexe Zahlen

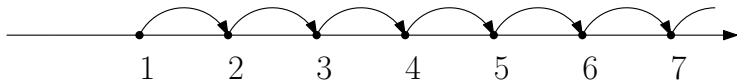
② Lineare Algebra

Natürliche Zahlen und Induktionsprinzip

Ausgehend von den reellen Zahlen findet man leicht folgende Charakterisierung der natürlichen Zahlen:

Die natürlichen Zahlen sind die kleinste Teilmenge von \mathbb{R} , die sowohl die Zahl 1 und mit jeder Zahl n auch deren **Nachfolger** $n + 1$ enthält.

Visualisierung am Zahlenstrahl:



Daraus können wir nun ein wichtiges Beweisprinzip ableiten.

Satz 1.12 (Prinzip der vollständigen Induktion)

Eine Aussageform $A(n)$ ist genau dann für alle $n \in \mathbb{N}$ wahr, wenn

- $A(1)$ wahr ist (*'Induktionsanfang'*),
- aus der Gültigkeit von $A(n)$ für ein beliebiges $n \in \mathbb{N}$ stets die Gültigkeit von $A(n+1)$ folgt (*'Induktionsschritt'*).

Der Induktionsschritt sichert also folgende 'Kette' von Implikationen:

$$A(1) \Rightarrow A(2) \Rightarrow A(3) \Rightarrow \dots \Rightarrow A(n) \Rightarrow A(n+1) \Rightarrow \dots$$

Man ist natürlich nicht zwingend auf den Startwert 1 festgelegt. Man kann im Induktionsanfang auch bei einem beliebigen $n_0 \in \mathbb{Z}$ starten, muss dann aber auch den Induktionsschritt für $n \geq n_0$ beweisen.

Beispiel 1.13 (Bernoullische Ungleichung)

Man beweise folgende Ungleichung mittels vollständiger Induktion:

$$(1+x)^n \geq 1+nx \quad \text{für alle } x \in [-1, \infty), n \in \mathbb{N}. \quad (1.1)$$

Wir führen eine vollständige Induktion über n aus.

Induktionsanfang:

Für $n = 1$ gilt

$$(1+x)^1 = 1+x = 1+n \cdot x,$$

d. h. die Aussage ist für alle $x \geq -1$ wahr.

Induktionsschritt:

Angenommen, (1.1) sei für ein $n \in \mathbb{N}$ wahr, d.h.

$$(1+x)^n \geq 1+nx \quad (x \geq -1) \quad (\text{Induktionsannahme, IA}).$$

Dann gilt

$$\begin{aligned}(1+x)^{n+1} &= (1+x)^n(1+x) \\ &\geq (1+nx)(1+x) \quad (\text{nach IA}) \\ &= 1+(n+1)x+nx^2 \\ &\geq 1+(n+1)x \quad (\text{da } nx^2 \geq 0).\end{aligned}$$

Aus der Gültigkeit von (1.1) für n folgt also die Gültigkeit für $n+1$.

① Grundlagen

- 1.1 Elemente der Aussagenlogik
- 1.2 Elemente der Mengenlehre
- 1.3 Die reellen Zahlen
- 1.4 Natürliche Zahlen und Induktionsprinzip
- 1.5 Abbildungen und Funktionen**
- 1.6 Komplexe Zahlen

② Lineare Algebra

Hier sollen nur die grundlegenden Begriffe bereitgestellt/wiederholt werden. Eine intensive Beschäftigung mit der Materie erfolgt später.

Definition 1.14 (Funktion)

Seien A und B Mengen. Eine **Abbildung** oder **Funktion** $f : A \rightarrow B$ ist eine Vorschrift, die jedem $x \in A$ **genau** ein $y = f(x) \in B$ zuordnet.

A heißt **Definitionsbereich** von f und $f(A) := \{f(x) : x \in A\} \subseteq B$ heißt **Wertebereich** oder **Bild** von f .

Für $x \in A$ heißt $y = f(x)$ **Bild von x unter f** oder **Funktionswert von f an der Stelle x** .

Zwei Funktionen $f : D_f \rightarrow B$, $x \mapsto f(x)$, und $g : D_g \rightarrow C$, $x \mapsto g(x)$, heißen gleich ($f = g$), wenn $D_f = D_g$, $B = C$ und $f(x) = g(x)$ für alle $x \in D_f = D_g$ gelten.

Schreibweisen:

$$f : A \rightarrow B \quad \text{oder} \quad f : A \rightarrow B$$
$$y = f(x) \quad \quad \quad x \mapsto f(x).$$

Beschreibung von Funktionen

Eine Funktion $f : A \rightarrow B$ kann man auf verschiedene Weisen beschreiben:

- analytisch, d.h. durch Angabe der Zuordnungsvorschrift,
- tabellarisch, d.h. durch eine Wertetabelle,
- graphisch, d.h. durch Visualisierung der Menge

$$\text{Graph}(f) := \{(x, f(x)) : x \in A\} \subseteq A \times B,$$

des sogenannten **Graphen von f** .

(Für $f : \mathbb{R} \rightarrow \mathbb{R}$ ist der Graph von f eine 'Kurve' im \mathbb{R}^2 .)

Definition 1.15

Sei $f : A \rightarrow B$ eine Funktion. Zu einer Menge $A' \subseteq A$ heißt

$$f(A') := \{f(x) : x \in A'\} \subseteq B$$

das **Bild** von A' unter f . Zu einer Menge $B' \subseteq B$ heißt

$$f^{-1}(B') := \{x \in A : f(x) \in B'\}$$

das **Urbild** von B' unter f .

Man bestimme $f([1, 2])$ und $f^{-1}([1, 4])$ für $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^2$.

Definition 1.16

Eine Funktion $f : A \rightarrow B$ heißt

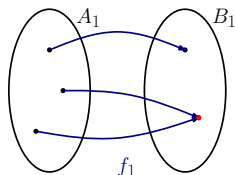
- **injektiv** (eindeutig), wenn für alle $x_1, x_2 \in A$ mit $x_1 \neq x_2$ stets $f(x_1) \neq f(x_2)$ gilt,
- **surjektiv**, wenn es zu jedem $y \in B$ ein $x \in A$ gibt mit $y = f(x)$,
- **bijektiv**, wenn f injektiv und surjektiv ist.

Ist f bijektiv, so existiert die **Umkehrfunktion**

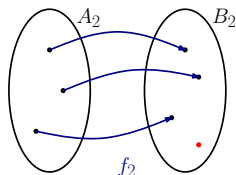
$$f^{-1} : B \rightarrow A, \quad f^{-1}(y) = x \Leftrightarrow y = f(x).$$

Abbildungen und Funktionen

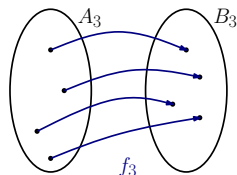
Umkehrbarkeit von Abbildungen



surjektiv, nicht injektiv



injektiv, nicht surjektiv



bijektiv

Ist die Funktion $f_1 : \mathbb{R} \rightarrow \mathbb{R}$, $f_1(x) = x^2$ injektiv/surjektiv/bijektiv?

Wie verhält es sich mit

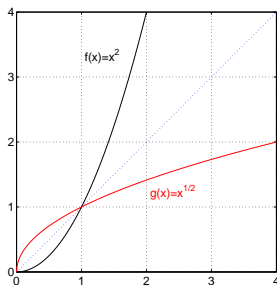
- $f_2 : \mathbb{N} \rightarrow \mathbb{N}$, $f_2(x) = x^2$,
- $f_3 : [0, \infty) \rightarrow \mathbb{R}$, $f_3(x) = x^2$,
- $f_4 : [0, \infty) \rightarrow [0, \infty)$, $f_4(x) = x^2$?

Beispiel 1.17 (n -te Wurzel)

Die Funktion $f : [0, \infty) \rightarrow [0, \infty)$, $f(x) = x^n$ ist für alle $n \in \mathbb{N}$ bijektiv. Die Umkehrfunktion ist die n -te Wurzel:

$$f^{-1} : [0, \infty) \rightarrow [0, \infty), \quad x \mapsto \sqrt[n]{x}.$$

Beachten Sie: die n -te Wurzel ist für negative Zahlen **nicht** definiert. Statt $\sqrt[n]{x}$ schreibt man kurz \sqrt{x} .



Situation für $n = 2$. Wie bei allen reellen Funktionen entsteht der Graph von f^{-1} durch Spiegeln des Graphen von f an der Geraden $y = x$.

① Grundlagen

- 1.1 Elemente der Aussagenlogik
- 1.2 Elemente der Mengenlehre
- 1.3 Die reellen Zahlen
- 1.4 Natürliche Zahlen und Induktionsprinzip
- 1.5 Abbildungen und Funktionen
- 1.6 Komplexe Zahlen

② Lineare Algebra

Erinnern Sie sich an die Zahlbereichserweiterungen? Unser Ziel war dabei, bestimmte Gleichungstypen uneingeschränkt lösen zu können.

Z. B. hatte $x + 5 = 2$ in \mathbb{N} keine Lösung. Erweitert man den Zahlbereich zu \mathbb{Z} , so lässt sich eine Lösung angeben ($x = -3$).

In den **reellen** Zahlen sind quadratische Gleichungen nicht immer lösbar: $x^2 = 1$ hat in \mathbb{R} zwei Lösungen ($x_{1/2} = \pm 1$), dagegen hat $x^2 = -1$ in \mathbb{R} keine Lösung.

Um diesen 'Mangel' zu beheben, kann man eine weitere Zahlbereichserweiterung durchführen. Der initiale Schritt ist dabei das Hinzufügen einer neuen Zahl i , die $i^2 = -1$ erfüllt.

Definition 1.18 (Komplexe Zahl)

Eine **komplexe Zahl** z ist ein Ausdruck der Form

$$z = a + ib \quad \text{mit } a, b \in \mathbb{R},$$

Die Zahl i (mit der Eigenschaft $i^2 = -1$) heißt **imaginäre Einheit**.

$a =: \operatorname{Re} z$ heißt **Realteil** von z ,

$b =: \operatorname{Im} z$ heißt **Imaginärteil** von z .

Zwei komplexe Zahlen sind genau dann gleich, wenn sowohl die Realteile als auch die Imaginärteile übereinstimmen.

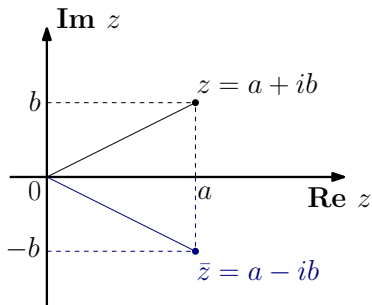
Die Menge aller komplexen Zahlen wird mit \mathbb{C} bezeichnet.

Statt $x+i0$ schreibt man x , d.h. reelle Zahlen sind komplexe Zahlen mit Imaginärteil 0. In diesem Sinne gilt also $\mathbb{R} \subseteq \mathbb{C}$.

Definition 1.19

Sei $z = a + ib$ ($a, b \in \mathbb{R}$) eine komplexe Zahl. Dann heißt

- $\bar{z} := a - ib$ die zu z **konjugiert komplexe Zahl**.
- $|z| := \sqrt{a^2 + b^2}$ der **Betrag** von $|z|$.



Eine komplexe Zahl z veranschaulicht man sich als Punkt der Gaußschen Zahlenebene mit den Koordinaten $(\text{Re } z, \text{Im } z)$.

Rechenoperationen in \mathbb{C}

Für $z = a + ib$, $w = c + id \in \mathbb{C}$ ($a, b, c, d \in \mathbb{R}$) definiert man

$$z + w := (a + c) + i(b + d),$$

$$z - w := (a - c) + i(b - d),$$

$$zw := (ac - bd) + i(ad + bc),$$

$$\frac{z}{w} := \frac{ac+bd}{c^2+d^2} + i\frac{bc-ad}{c^2+d^2} \quad (w \neq 0).$$

Diese Operationen sind so definiert, dass man die gewohnten Vorstellungen von den reellen Zahlen unter Beachtung von $i^2 = -1$ direkt übertragen kann.

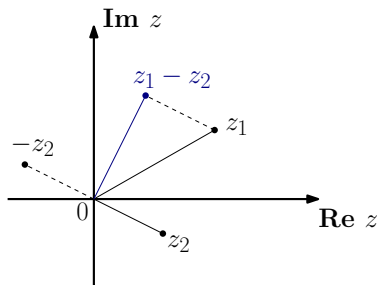
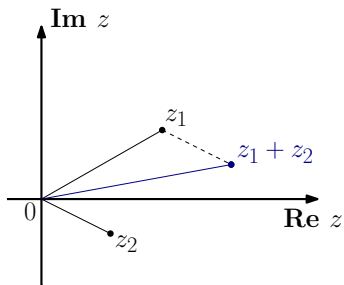
Die Division lässt sich als 'Erweitern' mit dem Konjugiert-Komplexen des Nenners auffassen:

$$\frac{z}{w} = \frac{(a + ib)(c - id)}{(c + id)(c - id)} = \frac{(a + ib)(c - id)}{c^2 + d^2} = \frac{z\bar{w}}{|w|^2}.$$

Komplexe Zahlen

Addition und Subtraktion in der Gaußschen Zahlenebene

Man gebe zu $z_1 = 1$, $z_2 = -i$, $z_3 = -3 + 2i$, $z_4 = 1 + i$ Real- und Imaginärteil an und zeichne die zugehörigen Punkte in die Gaußschen Zahlenebene. Man berechne $z_3 + z_4$, $z_3 - z_4$, $z_3 \cdot z_4$ sowie z_3/z_4 .
Wie lautet die kartesische Form der Zahl $\frac{1}{i}$?



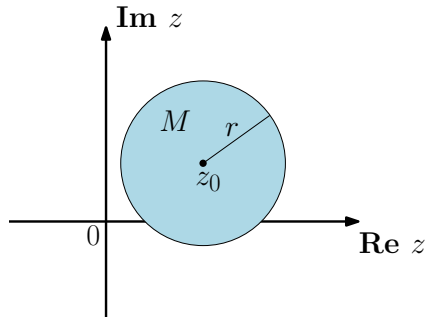
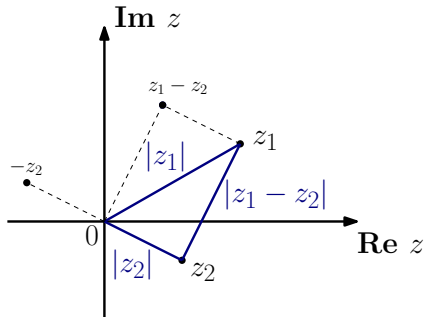
Beachten Sie die Analogie zur Vektorrechnung!

Komplexe Zahlen

Geometrische Interpretation des Betrags

In der Gaußschen Zahlenebene charakterisiert

- $|z|$ den Abstand einer Zahl z zum Koordinatenursprung
- $|z_1 - z_2| = |z_2 - z_1|$ den Abstand zwischen z_1 und z_2 in der Gaußschen Zahlenebene.
- die Menge $M = \{z \in \mathbb{C} : |z - z_0| \leq r\}$ für $r > 0$ und $z_0 \in \mathbb{C}$ eine Kreisscheibe um z_0 mit Radius r .



Komplexe Zahlen

\mathbb{C} als Zahlkörper

Mit der eingeführten Addition und Multiplikation bildet \mathbb{C} einen Zahlkörper. Konsequenzen sind:

Die arithmetischen Gesetze der reellen Zahlen (vgl. S. 48) gelten auch für komplexe Zahlen.

Satz 1.20

Die Rechenregeln für reelle Zahlen (Satz 1.7) sowie die binomischen Formeln und deren Verallgemeinerungen (vgl. S. 50) gelten auch für komplexe Zahlen.

Wie gewohnt schreiben wir dabei für $z \in \mathbb{C}, n \in \mathbb{N}$

$$z^n := z \cdot z \cdot \dots \cdot z \quad (n \text{ Faktoren}), \quad z^0 := 1, \quad z^{-n} := \frac{1}{z^n} \quad (z \neq 0).$$

Finden Sie die komplexen Lösungen der Gleichung $z^2(z+1)(z-4+i) = 0$.

Satz 1.21

Für $z, w \in \mathbb{C}$ gelten

- $\overline{\bar{z}} = z$,
- $\overline{z \pm w} = \bar{z} \pm \bar{w}$, $\overline{z \cdot w} = \bar{z} \cdot \bar{w}$, $\overline{\left(\frac{z}{w}\right)} = \frac{\bar{z}}{\bar{w}}$,
- $z\bar{z} = |z|^2$, $|\bar{z}| = |z|$,
- $\operatorname{Re} z = \frac{1}{2}(z + \bar{z})$, $\operatorname{Im} z = \frac{1}{2i}(z - \bar{z})$,
- $z = \bar{z} \Leftrightarrow z \in \mathbb{R}$.

Beweisen Sie einige dieser Regeln.

Satz 1.22

Für $z, w \in \mathbb{C}$ gelten

- $|z| \geq 0$, wobei $|z| = 0 \Leftrightarrow z = 0$,
- $|z| = |-z|$,
- $|zw| = |z||w|$,
- $|z + w| \leq |z| + |w|$ (**Dreiecksungleichung**).

Warum lassen sich die Regeln

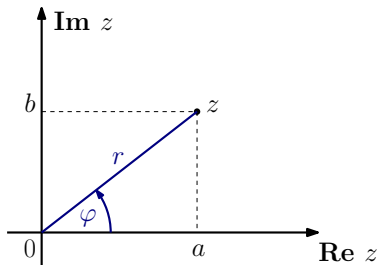
- $|a| = c \Rightarrow a = c$ oder $a = -c$,
- $|a| \leq c \Leftrightarrow -c \leq a \leq c$.

für $a \in \mathbb{R}$, $c \geq 0$ **nicht** auf komplexe Zahlen anwenden?

Komplexe Zahlen

Die Polarform einer komplexen Zahl

Eine komplexe Zahl $z \neq 0$ ist auch über ihre Polarkoordinaten in der Gaußschen Zahlenebene eindeutig charakterisiert:



Dabei ist

- $r = |z|$ der Abstand von z zu 0 ,
- φ der 'Drehwinkel' des Ortsvektors zu z , gemessen von der reellen Achse im Gegenuhrzeigersinn.

Komplexe Zahlen

Die Polarform einer komplexen Zahl

Mir den klassischen Definitionen von Sinus und Kosinus am Einheitskreis gilt also für $z = a + ib$:

$$z = r \cos \varphi + i \cdot r \sin \varphi = |z|(\cos \varphi + i \sin \varphi).$$

Definition 1.23 (Polarform)

Für $z = a + ib$ ($a, b \in \mathbb{R}$), $z \neq 0$, heißt $\varphi = \arg(z)$ **Argument** von z , falls

$$|z| \cos \varphi = a \quad \text{und} \quad |z| \sin \varphi = b. \quad (1.2)$$

Die Darstellung

$$z = |z|(\cos \varphi + i \sin \varphi).$$

heißt **Polarform** von z .

Das Argument ist nur bis auf ganzzahlige Vielfache von 2π festgelegt. Man wählt häufig $\varphi \in [0, 2\pi)$, um Eindeutigkeit zu erhalten, und spricht dann vom **Hauptwert** des Arguments.

- Von der Polarform $z = |z|(\cos \varphi + i \sin \varphi)$ zur kartesischen Form gelangt man durch Ausmultiplizieren.
- Von der kartesischen Form $z = a + ib$ zur Polarform gelangt man mit $|z| = \sqrt{a^2 + b^2}$ und (1.2).

Zur Bestimmung des Arguments kann man auch

$$\tan \varphi = \frac{b}{a} \quad (a \neq 0)$$

benutzen, muss dabei aber auf Quadrantenbeziehungen achten (der Fall $a = 0$ ist trivial).

Wie lautet die Polarform von $2i$, -1 , $-3 + 2i$?

Komplexe Zahlen

Die komplexe Exponentialfunktion

Die Zahl $e^{i\varphi}$ ($\varphi \in \mathbb{R}$) spielt im Zusammenhang mit der Polarform eine wichtige Rolle. Da uns noch keine Potenzreihen zur Verfügung stehen, definieren wir sie vorläufig mittels

Eulersche Formel

$$e^{i\varphi} := \cos \varphi + i \sin \varphi \quad (\varphi \in \mathbb{R})$$

Damit kann man jede komplexe Zahl $z \neq 0$ schreiben als

$$z = |z|(\cos \varphi + i \sin \varphi) = |z|e^{i\varphi} \quad (\varphi = \arg(z)).$$

Diese Darstellung nennt man **Eulersche, Exponential-** oder einfach wieder **Polar-****darstellung** von z .

Im Einklang mit den Potenzgesetzen schreiben wir ferner $e^z = e^a e^{ib}$ für eine beliebige komplexe Zahl $z = a + ib$ ($a, b \in \mathbb{R}$). Auch dies wollen wir vorläufig als Definition verstehen.

Es lässt sich zeigen, dass die komplexe Exponentialfunktion ähnlichen Gesetzen genügt, wie die reelle:

Satz 1.24

Für $z, w \in \mathbb{C}$, $n \in \mathbb{N}$ gelten:

$$e^{z+w} = e^z \cdot e^w, \quad e^{z-w} = \frac{e^z}{e^w}, \quad (e^z)^n = e^{nz},$$

insbesondere also für $\varphi, \psi \in \mathbb{R}$:

$$e^{i(\varphi+\psi)} = e^{i\varphi} \cdot e^{i\psi}, \quad e^{i(\varphi-\psi)} = \frac{e^{i\varphi}}{e^{i\psi}}, \quad (e^{i\varphi})^n = e^{in\varphi}.$$

Komplexe Zahlen

Die komplexe Exponentialfunktion

Die Eulersche Formel $e^{i\varphi} = \cos \varphi + i \sin \varphi$ ($\varphi \in \mathbb{R}$) liefert den Schlüssel für folgende Beobachtung:

- $e^{i\varphi}$ ist die komplexe Zahl auf dem Einheitskreis, deren Argument φ ist.

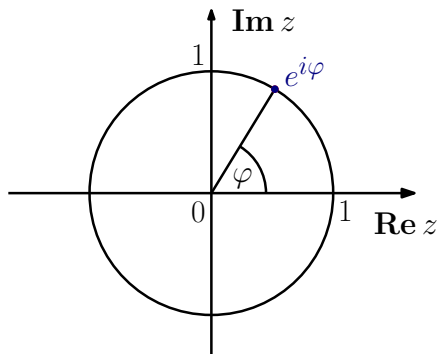


Bild rechts: Leonhard Euler (1707-1783), Schweizer Mathematiker.

Für zwei komplexe Zahlen $z, w \neq 0$ mit den Polardarstellungen

$$\begin{aligned}z &= |z| e^{i\varphi} = |z|(\cos \varphi + i \sin \varphi), \\w &= |w| e^{i\psi} = |w|(\cos \psi + i \sin \psi) \quad (\varphi, \psi \in \mathbb{R})\end{aligned}$$

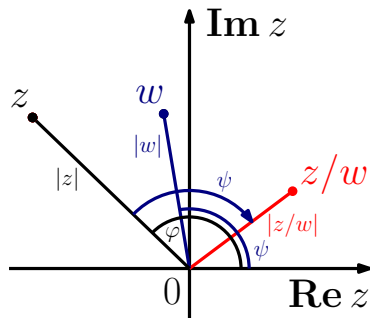
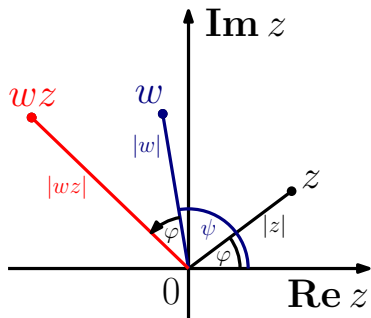
erhält man

$$\begin{aligned}z \cdot w &= |z||w| e^{i\varphi} e^{i\psi} \\&= |zw| e^{i(\varphi+\psi)} \\&= |zw|(\cos(\varphi + \psi) + i \sin(\varphi + \psi)), \\ \frac{z}{w} &= \frac{|z| e^{i\varphi}}{|w| e^{i\psi}} \\&= \left| \frac{z}{w} \right| e^{i(\varphi-\psi)} \\&= \left| \frac{z}{w} \right| (\cos(\varphi - \psi) + i \sin(\varphi - \psi)),\end{aligned}$$

d.h. $\arg(zw) = \arg(z) + \arg(w)$ und $\arg\left(\frac{z}{w}\right) = \arg(z) - \arg(w)$

Komplexe Zahlen

Geometrische Darstellung von Multiplikation und Division



Für eine komplexe Zahl $z \neq 0$ mit Polardarstellung $z = |z|e^{i\varphi}$ und $n \in \mathbb{N}$ gilt weiterhin

$$z^n = (|z|e^{i\varphi})^n = |z|^n (e^{i\varphi})^n = |z|^n e^{in\varphi} = |z|^n (\cos(n\varphi) + i \sin(n\varphi)).$$

Bei der Ermittlung der Lösungen von $w^n = z$ zu einer gegebenen komplexen Zahl $z = |z|e^{i\varphi}$ ist zu beachten, dass $\varphi \mapsto e^{i\varphi}$ im Gegensatz zum reellen Fall 2π -periodisch ist, d.h.

$$e^{i\varphi} = e^{i(\varphi+2k\pi)} \quad (k \in \mathbb{Z}).$$

Somit liefert die (formale) Anwendung der Potenzgesetze

$$w^n = |z|e^{i\varphi} \quad \Leftrightarrow \quad w = |z|^{\frac{1}{n}} e^{i\frac{\varphi+2k\pi}{n}} = \sqrt[n]{|z|} e^{i(\frac{\varphi}{n} + \frac{2k\pi}{n})}.$$

Auch die Periode 2π wird also durch n geteilt, wodurch n **verschiedene** 'n-te Wurzeln' von z entstehen.

Wir fassen unsere Vorüberlegung folgendermaßen zusammen:

Satz 1.25 (Potenzieren und Radizieren in \mathbb{C})

Sei $n \in \mathbb{N}$, dann gilt:

- Die n -te Potenz von $z = |z|e^{i\varphi} = |z|(\cos \varphi + i \sin \varphi)$ ($\varphi \in \mathbb{R}$) ergibt sich zu
$$z^n = |z|^n e^{in\varphi} = |z|^n (\cos(n\varphi) + i \sin(n\varphi)).$$

Insbesondere gilt die *de Moivresche Formel*

$$(\cos \varphi + i \sin \varphi)^n = \cos(n\varphi) + i \sin(n\varphi).$$

- Für jede Zahl $z = |z|e^{i\varphi}$ besitzt die Gleichung $w^n = z$ genau n verschiedene Lösungen

$$w_k = \sqrt[n]{|z|} e^{i\left(\frac{\varphi}{n} + \frac{2k\pi}{n}\right)} = \sqrt[n]{|z|} \left[\cos\left(\frac{\varphi}{n} + \frac{2k\pi}{n}\right) + i \sin\left(\frac{\varphi}{n} + \frac{2k\pi}{n}\right) \right]$$

mit $k = 0, \dots, n - 1$.

Komplexe Zahlen

Geometrische Darstellung der n -ten Wurzeln

Die n -ten Wurzeln von $|z|e^{i\varphi}$ liegen auf einem Kreis mit Radius $\sqrt[n]{|z|}$ um den Nullpunkt und bilden ein regelmäßiges n -Eck.

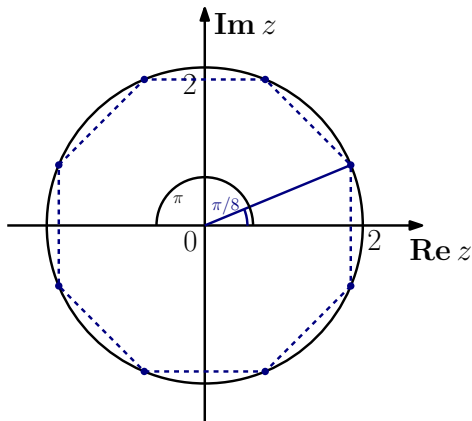


Illustration am Beispiel der 8-ten Wurzeln von $-256 = 256e^{i\pi}$.

Mit Hilfe der Eulerschen und der de Moivreschen Formel lassen sich Beziehungen für trigonometrische Funktionen zeigen, z. B.

- $\sin(x \pm y) = \sin x \cos y \pm \cos x \sin y,$
- $\cos(x \pm y) = \cos x \cos y \mp \sin x \sin y,$
- $\sin 2x = 2 \sin x \cos x,$
- $\cos 2x = \cos^2 x - \sin^2 x$ ($\sin^n x := (\sin x)^n$, analog für \cos)
- $\sin 3x = 3 \sin x - 4 \sin^3 x,$
- $\cos 3x = 4 \cos^3 x - 3 \cos x.$

Beweisen Sie einige dieser Beziehungen.

Bei der Bestimmung n -ter Wurzeln wurden Gleichungen der Form

$$z^n - a = 0$$

gelöst. Wir wollen die Frage der Lösbarkeit algebraischer Gleichungen allgemeiner untersuchen. Hierzu folgende Terminologie:

Eine Abbildung der Form

$$p : \mathbb{C} \rightarrow \mathbb{C}, \quad p(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0$$

mit festen Zahlen ('Koeffizienten') $a_0, a_1, \dots, a_n \in \mathbb{C}, a_n \neq 0$ heißt komplexes **Polynom** vom **Grad** $\deg(p) = n$.

Eine Zahl $w \in \mathbb{C}$ heißt **Nullstelle** von p , falls $p(w) = 0$.

Komplexe Zahlen

Abspalten von Linearfaktoren, Existenz von Nullstellen

Besitzt ein Polynom p vom Grad $n \geq 1$ eine Nullstelle w , so lässt sich p ohne Rest durch $(z - w)$ teilen, d.h. es gibt ein Polynom q von Grad $n - 1$ mit

$$p(z) = (z - w)q(z) \quad (z \in \mathbb{C}).$$

Anders als in \mathbb{R} gilt in den komplexen Zahlen folgende Aussage:

Satz 1.26 (Fundamentalsatz der Algebra)

*Jedes Polynom vom Grad $n \geq 1$ besitzt mindestens eine **komplexe** Nullstelle.*

Von jedem Polynom p mit Grad $n \geq 1$ kann man also einen Linearfaktor abspalten. Ist der verbleibende Rest mindestens vom Grad 1, kann man dies natürlich erneut tun usw. . . Wir halten fest:

Folgerung 1.27

Das Polynom $p : \mathbb{C} \rightarrow \mathbb{C}$ sei gegeben durch

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0.$$

Dann existieren komplexe Zahlen w_1, \dots, w_n mit

$$p(z) = a_n (z - w_1)(z - w_2) \dots (z - w_n). \quad (1.3)$$

Jedes komplexe Polynom hat also sogar n Nullstellen (Vielfachheiten mitgezählt) und zerfällt vollständig in Linearfaktoren.

Die Berechnung der Nullstellen gemäß Satz 1.26 und Folgerung 1.27 ist i. Allg. nur numerisch möglich. Ausnahmen bilden n -te Wurzeln oder quadratische Gleichungen mit **reellen** Koeffizienten:

Satz 1.28

Die komplexen Lösungen der quadratischen Gleichung

$$z^2 + pz + q = 0, \quad p, q \in \mathbb{R},$$

sind gegeben durch

$$z_{1,2} = \begin{cases} -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}, & \text{falls } \frac{p^2}{4} - q \geq 0, \\ -\frac{p}{2} \pm i\sqrt{q - \frac{p^2}{4}}, & \text{falls } \frac{p^2}{4} - q < 0. \end{cases}$$

Verifizieren Sie den zweiten Fall durch Modifikation der bekannten Herleitung der p - q -Formel. Geben Sie die komplexen Lösungen von $z^2 - 4z + 5 = 0$ an.

Das Auftreten konjugiert komplexer Zahlen in der p - q -Formel lässt sich verallgemeinern:

Satz 1.29

Sind die Koeffizienten des Polynoms

$$p: \mathbb{C} \rightarrow \mathbb{C}, \quad p(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0.$$

*allesamt **reell**, so sind die Nullstellen reell ($\lambda_j \in \mathbb{R}$) oder treten in konjugiert komplexen Paaren ($\alpha_j \pm i\beta_j$, $\alpha_j, \beta_j \in \mathbb{R}$) auf. Es existiert eine Zerlegung der Form*

$$p(z) = a_n \prod_{j=1}^k (z - \lambda_j) \prod_{j=1}^m ((z - \alpha_j)^2 + \beta_j^2), \quad (1.4)$$

*wobei $k + 2m = n$. Jedes **reelle Polynom** lässt sich also als Produkt von Linearfaktoren und quadratischen Polynomen schreiben.*

Geben Sie Zerlegungen von $p(z) = z^3 + z$ gemäß (1.3) und (1.4) an.

In der Wechselstromtechnik treten häufig sinus-/cosinusförmige Spannungs- bzw. Stromverläufe auf:

$$i(t) = I_m \cos(\omega t)$$

$$u(t) = U_m \cos(\omega t + \varphi)$$

(I_m, U_m Amplituden, t Zeit, $\omega = 2\pi f$ Kreisfrequenz, f Frequenz (Netz in D: $f = 50$ Hz))

Der Winkel φ charakterisiert die i.A. auftretende **Phasenverschiebung** zwischen Strom und Spannung.

Ursache für die Phasenverschiebung sind verschiedene 'Arten' von Widerständen im Wechselstromkreis:

- **Ohmscher Widerstand** R :

$$U_m = R I_m, \text{ keine Phasenverschiebung,}$$

- **Induktivität** L (Spule):

$$U_m = \omega L I_m, \text{ Phasenverschiebung } \frac{\pi}{2},$$

Hintergrund Selbstinduktion/Lenzsches Gesetz,

- **Kapazität** C (Kondensator):

$$U_m = \frac{1}{\omega C} I_m, \text{ Phasenverschiebung } -\frac{\pi}{2},$$

Hintergrund: fortwährende Ladungs-/Entladungsvorgänge.

Es müssen also nicht nur die reellen Proportionalitätsfaktoren, sondern auch die Phasenverschiebungen beachtet werden.

Trick: Verwende zur Berechnung komplexe Ströme und Spannungen und interpretiere dabei nur den Realteil:

$$i(t) = I_m e^{i\omega t}, \quad u(t) = U_m e^{i(\omega t + \varphi)}.$$

Mit folgenden **komplexen** Widerständen \tilde{R} gilt dann das Ohmsche Gesetz $u(t) = \tilde{R} i(t)$ auch hier:

- Ohmscher Widerstand $R \in \mathbb{R}$:

$$u(t) = U_m e^{i\omega t} = R I_m e^{i\omega t} = R i(t),$$

- Induktiver Widerstand $R_L = i\omega L$:

$$u(t) = U_m e^{i(\omega t + \frac{\pi}{2})} = i\omega L I_m e^{i\omega t} = R_L i(t) \quad (\text{beachte } i = e^{i\frac{\pi}{2}})$$

- Kapazitiver Widerstand $R_C = -i\frac{1}{\omega C}$:

$$u(t) = U_m e^{i(\omega t - \frac{\pi}{2})} = -\frac{i}{\omega C} I_m e^{i\omega t} = R_C i(t) \quad (\text{beachte } -i = e^{-i\frac{\pi}{2}})$$

Für einen komplexen Widerstand R bezeichnet man $\operatorname{Re} R$ als **Wirk-** und $\operatorname{Im} R$ als **Blindwiderstand**. Ohmsche Widerstände sind reine Wirkwiderstände, kapazitive und induktive reine Blindwiderstände.

Vorteile der komplexen Wechselstromrechnung:

- Das Ohmsche Gesetz gilt wie gewohnt (für die Realteile gilt es nicht!),
- Phasenverschiebung $\arg R$ und **Scheinwiderstand** $|R| = U_m/I_m$ werden gleichzeitig erfasst,
- Die Kirchhoffschen Gesetze und ihre Folgerungen gelten damit auch im Wechselstromkreis, z. B.

$$R = R_1 + R_2 \quad \text{bzw.} \quad \frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}$$

für Reihen-/Parallelschaltung zweier Widerstände.

Weitere Anwendungen der komplexen Zahlen finden sich an verschiedensten Stellen. Einige werden Sie im Laufe ihres Studiums sicher kennenlernen, z.B.

- Die mit bestimmten Differentialgleichungen verknüpften 'Eigenwerte' sind komplexe Zahlen. Standardanwendungen sind Federschwinger oder Fadenpendel.
- In der Signal- oder Bildanalyse verwendet man oft die komplexwertige Fouriertransformierte, um Frequenzanalysen oder -filterungen durchzuführen.
- In der Quantenmechanik werden Teilchen durch ihre komplexwertige Zustandsfunktion beschrieben. Interpretiert wird nur deren Quadrat ('Ortsauffindungswahrscheinlichkeitsdichte')

Ziele erreicht?

Sie sollten nun (bzw. nach Abschluss der Übungen/Nacharbeiten des Stoffes):

- sicher mit reellen Zahlen, Gleichungen und Ungleichungen umgehen können,
- sicher mit komplexen Zahlen umgehen können (Arithmetik in kartesischer und Polarform, Potenzen und Wurzeln, einfache Gleichungen, Veranschaulichungen in der Gaußschen Zahlenebene),
- einfache bis mäßig schwierige logische Zusammenhänge und Schlüsse erfassen können,
- Überblickswissen über den Aufbau mathematischer Theorie im allgemeinen besitzen,
- eine grobe Vorstellung vom axiomatischen Aufbau von Zahlenbereichen haben,
- erste Vorstellungen entwickelt haben, wozu man das bisher vermittelte Wissen in der Informatik/Elektrotechnik braucht.

Sie sind sich nicht sicher oder meinen 'nein'? Dann **werden Sie aktiv!**

① Grundlagen

② Lineare Algebra

Mit der linearen Algebra lernen wir nun ein großes Teilgebiet der Mathematik kennen.

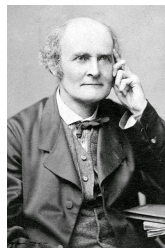
Dieses befasst sich unter anderem mit

- Vektorräumen,
- linearen Abbildungen zwischen Vektorräumen,
- linearen Gleichungssystemen,
- Determinanten und Matrizen,
- Eigenwerten und -vektoren.

Insbesondere werden wir hier auch die Grundlagen für die mehrdimensionale Differential- und Integralrechnung legen.

Die Entwicklung der modernen linearen Algebra erfolgte vor allem in der Mitte des 19. Jahrhunderts, wenngleich erste Grundlagen bereits wesentlich früher bekannt waren. Wichtige Beiträge lieferten die Mathematiker

- Gabriel Cramer (1704-1752, CH),
- Sir William Rowan Hamilton (1805–1865, IRL),
- Hermann Graßmann (1809–1877, D),
- Arthur Cayley (1821–1895, UK).



① Grundlagen

② Lineare Algebra

2.1 Vektorräume

2.2 Matrizen und lineare Abbildungen

2.3 Lineare Gleichungssysteme

2.4 Determinanten

2.5 Invertierbare Matrizen

2.6 Orthogonalität, Skalarprodukt und Norm

2.7 Kreuz- und Spatprodukt

2.8 Elemente der analytischen Geometrie

2.9 Orthogonale Abbildungen

2.10 Eigenwerte und Eigenvektoren

2.11 Singulärwertzerlegung

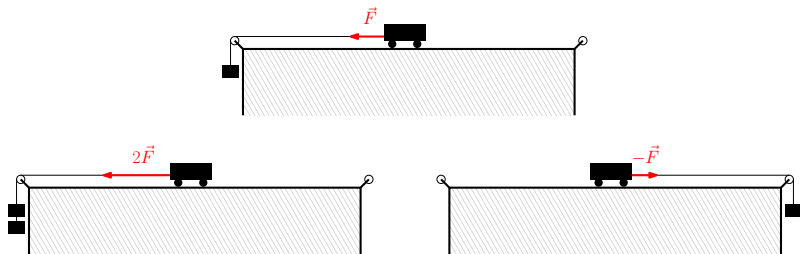
Vektorräume

Motivation: physikalische Kräfte

Physikalische Kräfte können nicht durch eine Zahl allein beschrieben werden: sie besitzen neben ihrem 'Betrag' auch eine Richtung.

Man beschreibt sie durch **Vektoren**. Wirken die Kräfte in einer Ebene, verwendet man 'zweidimensionale' Vektoren $\begin{bmatrix} \alpha \\ \beta \end{bmatrix}$.

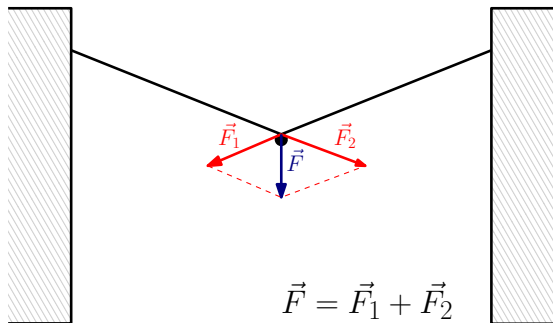
Man kann Kräfte (Vektoren) mit einer Zahl λ multiplizieren: dabei wird die Richtung beibehalten oder (bei negativem λ) umgekehrt und der Betrag mit $|\lambda|$ multipliziert.



Vektorräume

Motivation: physikalische Kräfte

Desweiteren kann man Kräfte (Vektoren) addieren. Dies visualisiert man am sogenannten Kräfteparallelogramm.



Man beachte, dass sich die Beträge der Kräfte **nicht** einfach addieren. Es gelten aber auch für die Vektoraddition viele gewohnte Gesetzmäßigkeiten, die in die Definition des Vektorraums einfließen.

Um die algebraische Struktur des Vektorraums und damit den Vektorbegriff mathematisch exakt zu fassen, benötigen wir zunächst einen Körper \mathbb{K} .

Dies ist eine Menge, auf der zwei Operationen ($+$ und \cdot) definiert sind, die den Gesetzen genügen, die in Abschnitt 1.3, Folie 48, aufgelistet wurden.

In den meisten Fällen werden wir als Körper die reellen Zahlen wählen ($\mathbb{K} = \mathbb{R}$), mitunter auch die komplexen Zahlen ($\mathbb{K} = \mathbb{C}$).

Grundsätzlich könnte man aber jeden Körper wählen (z. B. \mathbb{Q}).

Definition 2.1 (Vektorraum)

Ein \mathbb{K} -Vektorraum $\mathcal{V} := (V; +, \cdot)$ besteht aus einer Menge $V \neq \emptyset$, deren Elemente **Vektoren** genannt werden, sowie zwei Operationen: einer **(Vektor)addition** $+: V \times V \rightarrow V$ und einer **Skalarmultiplikation** $\cdot: \mathbb{K} \times V \rightarrow V$.

Dabei müssen folgende Regeln gelten:

- (1) $\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}$ für alle $\mathbf{a}, \mathbf{b}, \mathbf{c} \in V$,
- (2) es gibt einen Vektor $\mathbf{0}$ mit $\mathbf{a} + \mathbf{0} = \mathbf{a}$ für alle $\mathbf{a} \in V$,
- (3) zu jedem $\mathbf{a} \in V$ gibt es ein $-\mathbf{a} \in V$ mit $\mathbf{a} + (-\mathbf{a}) = \mathbf{0}$,
- (4) $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$ für alle $\mathbf{a}, \mathbf{b} \in V$,
- (5) $(\lambda\mu) \cdot \mathbf{a} = \lambda \cdot (\mu \cdot \mathbf{a})$ für alle $\lambda, \mu \in \mathbb{K}$ und alle $\mathbf{a} \in V$,
- (6) $(\lambda + \mu) \cdot \mathbf{a} = \lambda \cdot \mathbf{a} + \mu \cdot \mathbf{a}$ für alle $\lambda, \mu \in \mathbb{K}$ und alle $\mathbf{a} \in V$,
- (7) $\lambda \cdot (\mathbf{a} + \mathbf{b}) = \lambda \cdot \mathbf{a} + \lambda \cdot \mathbf{b}$ für alle $\lambda \in \mathbb{K}$ und alle $\mathbf{a}, \mathbf{b} \in V$,
- (8) $1 \cdot \mathbf{a} = \mathbf{a}$ für alle $\mathbf{a} \in V$.

Vektorräume

Vektorraumaxiome, Erläuterungen und Anmerkungen

- $+: V \times V \rightarrow V$ bedeutet, dass die Addition je zwei Vektoren \mathbf{a} und \mathbf{b} einen Vektor $\mathbf{a} + \mathbf{b}$ zuordnet
- $\cdot: \mathbb{K} \times V \rightarrow V$ bedeutet, dass die Skalarmultiplikation je einer Zahl λ und einem Vektor \mathbf{a} einen Vektor $\lambda \cdot \mathbf{a}$ zuordnet.

Dies entspricht genau dem Charakter der am Beispiel physikalischer Kräfte diskutierten Operationen.

Auch wenn sich die Menge V und die algebraische Struktur $\mathcal{V} := (V; +, \cdot)$ prinzipiell unterscheiden, verwendet man statt des umständlichen $(V; +, \cdot)$ fast immer nur V als Bezeichnung des Vektorraums.

Das (momentan) wichtigste Beispiel für einen Vektorraum ist

$$\mathbb{K}^n := \left\{ \vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} : a_j \in \mathbb{K}, j = 1, 2, \dots, n \right\}$$

mit

$$\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} := \begin{bmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{bmatrix} \quad \text{und} \quad \lambda \cdot \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} := \begin{bmatrix} \lambda a_1 \\ \lambda a_2 \\ \vdots \\ \lambda a_n \end{bmatrix}$$

($\lambda \in \mathbb{K}; a_j, b_j \in \mathbb{K}$). Addition und Skalarmultiplikation sind also **komponentenweise** definiert.

Vektorräume

Beispiele

Zwei Vektoren, $\vec{a} = [a_j]_{j=1}^n$ und $\vec{b} = [b_j]_{j=1}^n \in \mathbb{K}^n$, sind genau dann gleich, wenn $a_j = b_j$ für alle $j = 1, 2, \dots, n$ gilt.

Nullvektor und inverse Vektoren sind im \mathbb{K}^n gegeben durch

$$\vec{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{bzw.} \quad -\vec{a} = \begin{bmatrix} -a_1 \\ -a_2 \\ \vdots \\ -a_n \end{bmatrix}, \quad \text{falls } \vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}.$$

Zu $\vec{a}, \vec{b} \in \mathbb{R}^2$ mit $\vec{a} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ und $\vec{b} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ berechne man $\vec{a} + 3\vec{b}$ und $\vec{a} - \vec{b}$.

Zu $\vec{a}, \vec{b} \in \mathbb{C}^2$ mit $\vec{a} = \begin{bmatrix} 3 \\ 4i \end{bmatrix}$ und $\vec{b} = \begin{bmatrix} -1+i \\ 2 \end{bmatrix}$ berechne man $\vec{a} + (1+i)\vec{b}$ und $\vec{a} - i\vec{b}$.

Verifizieren Sie für $V = \mathbb{R}^n$ einige der in Definition 2.1 genannten Beziehungen.

In jedem der folgenden Beispiele sind Addition und Skalarmultiplikation punktweise zu verstehen.

- Die Menge der Polynome bildet einen Vektorraum.
- Die Menge der Polynome vom maximalen Grad n bildet einen Vektorraum.
- Die Menge der stetigen reellen Funktionen bildet einen Vektorraum (Bezeichnung: $C(\mathbb{R})$).
- Die Menge der k -mal stetig differenzierbaren reellen Funktionen bildet einen Vektorraum (Bezeichnung: $C^k(\mathbb{R})$).

- Koordinatenvektoren sind **Spaltenvektoren**. Weil das oft zuviel Platz beansprucht, schreiben wir auch

$$\vec{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} =: [a_1, a_2, \dots, a_n]^T.$$

a_j heißt j -te **Komponente** von \vec{a} .

- Solange wir allgemeine Vektorräume betrachten, verwenden wir für Vektoren fette kleine lateinische Buchstaben ($\mathbf{a}, \mathbf{b}, \dots$) und kleine griechische Buchstaben für Skalare.
- Für den Spezialfall \mathbb{K}^n , insbesondere für \mathbb{R}^n verwenden wir für Vektoren die Schreibweise mit dem Pfeil (\vec{a}, \vec{b}, \dots) und deren Komponenten die Schreibweise a_j, b_j, \dots ($j = 1, \dots, n$).
- Der Punkt \cdot , der für die Skalarmultiplikation steht, wird meistens unterdrückt.

In Vektorräumen gelten weiterhin folgende Rechenregeln:

Satz 2.2

Sei V ein \mathbb{K} -Vektorraum. Dann gelten:

- $0\mathbf{v} = \lambda\mathbf{0} = \mathbf{0}$ für alle $\lambda \in \mathbb{K}$ und alle $\mathbf{v} \in V$.
- $(-\lambda)\mathbf{v} = \lambda(-\mathbf{v}) = -(\lambda\mathbf{v})$ für alle $\lambda \in \mathbb{K}$ und alle $\mathbf{v} \in V$.
- $(-\lambda)(-\mathbf{v}) = \lambda\mathbf{v}$ für alle $\lambda \in \mathbb{K}$ und alle $\mathbf{v} \in V$.

Dem Anwender dürften diese Regeln intuitiv klar sein; aber genaugenommen müssen sie aus Definition 2.1 hergeleitet werden.

Machen Sie sich für mindestens einen Punkt klar, wie das geschehen könnte.

Definition 2.3 (Unterraum)

Ist U eine nichtleere Teilmenge eines \mathbb{K} -Vektorraums V mit

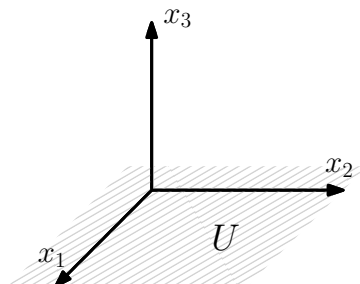
- $u + v \in U$ für alle $u, v \in U$ und
- $\lambda u \in U$ für alle $u \in U$ und alle $\lambda \in \mathbb{K}$,

dann nennt man U einen **Unterraum** von V .

Für die in Definition 2.3 genannten Punkte verwendet man auch zusammenfassend die Sprechweise: U ist abgeschlossen unter Addition und Skalarmultiplikation.

Natürlich ist U damit selbst wieder ein Vektorraum; daher verwendet man auch die Bezeichnung Unter(vektor-)raum.

- Jeder Vektorraum V enthält als triviale Unterräume den gesamten Raum, also V , und den **Nullraum** $\{\mathbf{0}\}$, der nur aus dem Nullvektor besteht.
- Die Menge $\{\vec{x} \in \mathbb{R}^3 : x_3 = 0\}$ bildet einen Unterraum U des \mathbb{R}^3 .



Überzeugen Sie sich anhand der Definition, dass der zweite Punkt wahr ist.

Definition 2.4 (und Satz)

Sei V ein Vektorraum. Ein Vektor \mathbf{y} der Form

$$\mathbf{y} = \sum_{j=1}^k \lambda_j \mathbf{x}_j \quad (\lambda_j \in \mathbb{K}, \mathbf{x}_j \in V, k \in \mathbb{N})$$

heißt **Linearkombination** der Vektoren $\mathbf{x}_1, \dots, \mathbf{x}_k$.

Ist $\emptyset \neq X \subseteq V$, so ist

$$\text{span}(X) := \left\{ \sum_{j=1}^k \lambda_j \mathbf{x}_j : \lambda_j \in \mathbb{K}, \mathbf{x}_j \in X, k \in \mathbb{N} \right\}$$

ein Unterraum von V , genauer: der kleinste Unterraum von V , der X enthält. Man nennt $\text{span}(X)$ die **lineare Hülle** von X oder den von X **erzeugten Unterraum** von V .

Die lineare Hülle $\text{span}(X)$ ist also gerade die Menge aller Linearkombinationen von Vektoren aus X .

- $\text{span}\{\mathbf{v}\} = \{\lambda\mathbf{v} : \lambda \in \mathbb{K}\}$,
- $\text{span}\{\mathbf{v}, \mathbf{w}\} = \{\lambda\mathbf{v} + \mu\mathbf{w} : \lambda, \mu \in \mathbb{K}\}$
- Für $V = \mathbb{R}^3$ gilt

$$\text{span} \left(\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\} \right) = \{x \in \mathbb{R}^3 : x_3 = 0\},$$

aber auch

$$\text{span} \left(\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \right\} \right) = \{x \in \mathbb{R}^3 : x_3 = 0\}.$$

Definition 2.5 (Lineare Unabhängigkeit)

Seien V ein \mathbb{K} -Vektorraum und $X \subseteq V$ eine Teilmenge von V . Die Vektoren aus X heißen **linear unabhängig**, wenn der Nullvektor nur trivial als Linearkombination von Vektoren aus X dargestellt werden kann; d. h. wenn aus

$$\sum_{j=1}^k \lambda_j \mathbf{x}_j = \mathbf{0} \quad (\text{mit } \mathbf{x}_j \in X \text{ und } \lambda_j \in \mathbb{K})$$

stets

$$\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$$

folgt. Vektoren, die nicht linear unabhängig sind, nennt man **linear abhängig**.

Vektorräume

Lineare Unabhängigkeit (äquivalente Charakterisierung)

Seien V ein \mathbb{K} -Vektorraum und $X \subseteq V$ eine Teilmenge von V . Dann sind die Vektoren aus X genau dann linear unabhängig, wenn sich keiner der Vektoren aus X als Linearkombination der anderen schreiben lässt.

Dies ist wiederum äquivalent zur Forderung

$$\text{span}(X \setminus \{x\}) \subsetneq \text{span}(X) \quad \text{für alle } x \in X.$$

Machen Sie sich klar, dass es sich hierbei tatsächlich um eine äquivalente Charakterisierung handelt.

Wir werden später effiziente Möglichkeiten kennenlernen, Informationen über lineare Unabhängigkeit zu erhalten. Wir versuchen uns trotzdem bereits hier an folgender Aufgabe:

Für welche der folgenden Mengen $X_j \subset \mathbb{R}^2$ sind die Vektoren aus X_j linear unabhängig? Geben Sie jeweils eine schlüssige Begründung.

$$X_1 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}, \quad X_2 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

$$X_3 = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}, \quad X_4 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 0 \end{bmatrix} \right\}$$

Definition 2.6

Ist V ein Vektorraum, so heißt eine Teilmenge $X \subseteq V$ ein **Erzeugendensystem** von V , wenn man jeden Vektor $v \in V$ als Linearkombination von Vektoren aus X darstellen kann.

Ein Erzeugendensystem X von V , das aus linear unabhängigen Vektoren besteht, heißt **Basis** von V .

Erinnerung: 'Als Linearkombination darstellbar' bedeutet, dass zu jedem $v \in V$ Skalare $\lambda_1, \lambda_2, \dots, \lambda_k \in \mathbb{K}$ und Vektoren $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k \in X$ existieren mit

$$v = \lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \dots + \lambda_k \mathbf{x}_k = \sum_{j=1}^k \lambda_j \mathbf{x}_j,$$

oder kurz, dass $V = \text{span}(X)$ gilt.

Sei X eine Basis des \mathbb{K} -Vektorraums V . Dann gilt:

- Entfernt man aus X einen beliebigen Vektor x , dann ist $X \setminus \{x\}$ kein Erzeugendensystem von V .

Mit anderen Worten: Eine Basis von V ist ein **minimales Erzeugendensystem** von V .

- Fügt man zu X einen Vektor y ($y \notin X$) hinzu, dann sind die Vektoren aus $X \cup \{y\}$ nicht mehr linear unabhängig.

Mit anderen Worten: Eine Basis von V ist eine **maximale Menge linear unabhängiger Vektoren** aus V .

Ein Vektorraum V hat i.Allg. viele verschiedenen Basen, die aber alle dieselbe Anzahl von Elementen besitzen. Die Zahl der Vektoren, aus denen eine Basis von V besteht, heißt **Dimension** von V .

Schreibweise: $\dim(V)$.

- Für $V = \mathbb{R}^2$ gilt $\dim(V) = 2$, denn $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ ist eine Basis des \mathbb{R}^2 .
- Für den Vektorraum V_p der Polynome gilt $\dim(V_p) = \infty$. Eine Basis ist zum Beispiel gegeben durch $\{1, x, x^2, x^3, \dots\}$. (Warum?)

Wir befassen uns hier (fast) nur mit **endlich-dimensionalen** Vektorräumen V (d. h. $\dim(V) = n < \infty$).

Welche der folgenden Mengen $X_j \subset \mathbb{R}^2$ sind Erzeugendensysteme bzw. Basen des \mathbb{R}^2 ?

$$X_1 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}, \quad X_2 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

$$X_3 = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}, \quad X_4 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$$

Argumentieren Sie auch mit Hilfe der auf den letzten beiden Folien behandelten Eigenschaften von Basen.

Man charakterisiere die 1- und 2-dimensionalen Untervektorräume des \mathbb{R}^3 geometrisch.

Satz 2.7

Ist X eine Basis des \mathbb{K} -Vektorraums V , so lässt sich jeder Vektor $v \in V$ auf **eindeutige** Weise als Linearkombination der Basisvektoren darstellen.

- Für endlichdimensionale Vektorräume und $X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ existieren somit zu jedem $v \in V$ eindeutig bestimmte Skalare $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ mit
$$v = \sum_{j=1}^n \lambda_j \mathbf{x}_j.$$
- Im allgemeinen Fall mit $X = \{\mathbf{x}_j : j \in J\}$ ist diese Beziehung zu ersetzen durch $v = \sum_{j \in J} \lambda_j \mathbf{x}_j$, wobei jedoch nur endlich viele $\lambda_j \in \mathbb{K}$ ungleich 0 sein dürfen.

Man mache sich letzteres am Beispiel des Vektorraums der Polynome und der Exponentialfunktion (die natürlich kein Polynom ist) klar.

Der Vektor

$$\vec{e}_j := [0, \dots, 0, \underbrace{1}_{j\text{-te-Komponente}}, 0, \dots, 0]^T \in \mathbb{K}^n \quad (j = 1, 2, \dots, n)$$

heißt j -ter (n -dimensionaler) **Einheitsvektor** des \mathbb{K}^n . Es gilt:

Satz 2.8

Mehr als n Vektoren aus \mathbb{K}^n sind stets linear abhängig.

Im \mathbb{K}^n sind k paarweise verschiedene Einheitsvektoren ($k \leq n$) immer linear unabhängig.

*Insbesondere ist $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$ eine Basis des \mathbb{K}^n , die sogenannte **Standardbasis**, und $\dim(\mathbb{K}^n) = n$.*

Zwei Vektoren im \mathbb{R}^2 bilden genau dann eine Basis des \mathbb{R}^2 , wenn sie nicht auf einer Geraden liegen. Drei Vektoren im \mathbb{R}^3 bilden genau dann eine Basis des \mathbb{R}^3 , wenn sie nicht in einer Ebene liegen.

① Grundlagen

② Lineare Algebra

2.1 Vektorräume

2.2 Matrizen und lineare Abbildungen

2.3 Lineare Gleichungssysteme

2.4 Determinanten

2.5 Invertierbare Matrizen

2.6 Orthogonalität, Skalarprodukt und Norm

2.7 Kreuz- und Spatprodukt

2.8 Elemente der analytischen Geometrie

2.9 Orthogonale Abbildungen

2.10 Eigenwerte und Eigenvektoren

2.11 Singulärwertzerlegung

Matrizen und lineare Abbildungen

Matrizen

Eine $m \times n$ -**Matrix** A ist ein rechteckiges Zahlenschema, in dem $m \cdot n$ reelle oder komplexe Einträge in m Zeilen und n Spalten angeordnet sind:

$$A = [a_{i,j}]_{1 \leq i \leq m, 1 \leq j \leq n} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}.$$

Die Zahl $a_{i,j}$, die in der i -ten Zeile und j -ten Spalte von A steht, heißt der (i, j) -te Eintrag von A .

Die Menge der reellen (komplexen) $m \times n$ -Matrizen wird mit $\mathbb{R}^{m \times n}$ ($\mathbb{C}^{m \times n}$) bezeichnet. Um beide Fälle zu erfassen schreiben wir $\mathbb{K}^{m \times n}$.

Matrizen und lineare Abbildungen

Gleichheit von Matrizen, Vektoren als Spezialfall

Zwei Matrizen $A, B \in \mathbb{K}^{m \times n}$ mit $A = [a_{i,j}]_{1 \leq i \leq m, 1 \leq j \leq n}$ und $B = [b_{i,j}]_{1 \leq i \leq m, 1 \leq j \leq n}$ sind genau dann gleich, wenn

$$a_{i,j} = b_{i,j} \quad \text{für alle } i = 1, 2, \dots, m \text{ und alle } j = 1, 2, \dots, n.$$

Vektoren aus dem \mathbb{K}^n (Spaltenvektoren) kann man als Matrizen aus $\mathbb{K}^{n \times 1}$ auffassen.

Zeilenvektoren lassen sich analog als Elemente von $\mathbb{K}^{1 \times n}$ auffassen.

Dies wird insbesondere dann deutlich, wenn wir Addition und Skalarmultiplikation für Matrizen eingeführt haben.

Matrizen und lineare Abbildungen

Addition und Skalarmultiplikation für Matrizen

Wie bei Vektoren werden **Addition** und **Skalarmultiplikation** komponentenweise erklärt.

Für

$$A = [a_{i,j}]_{1 \leq i \leq m, 1 \leq j \leq n}, B = [b_{i,j}]_{1 \leq i \leq m, 1 \leq j \leq n} \in \mathbb{K}^{m \times n} \text{ und } \lambda \in \mathbb{K}$$

definiert man

$$A + B := [a_{i,j} + b_{i,j}]_{1 \leq i \leq m, 1 \leq j \leq n} \in \mathbb{K}^{m \times n}$$

und

$$\lambda \cdot A := [\lambda a_{i,j}]_{1 \leq i \leq m, 1 \leq j \leq n} \in \mathbb{K}^{m \times n}.$$

Man beachte, dass die Addition nur für Matrizen gleicher Größe (gleiche Anzahl Zeilen und Spalten) definiert ist.

Satz 2.9

- $A + (B + C) = (A + B) + C$ für alle $A, B, C \in \mathbb{K}^{m \times n}$,
- es gibt eine Matrix $O \in \mathbb{K}^{m \times n}$, die sogenannte **Nullmatrix**, mit $A + O = A$ für alle $A \in \mathbb{K}^{m \times n}$,
- zu jeder Matrix $A \in \mathbb{K}^{m \times n}$ gibt es eine Matrix $-A \in \mathbb{K}^{m \times n}$ mit $A + (-A) = O$,
- $A + B = B + A$ für alle $A, B \in \mathbb{K}^{m \times n}$,
- $(\lambda\mu)A = \lambda(\mu A)$ für alle $\lambda, \mu \in \mathbb{K}$ und alle $A \in \mathbb{K}^{m \times n}$,
- $(\lambda + \mu)A = \lambda A + \mu A$ für alle $\lambda, \mu \in \mathbb{K}$ und alle $A \in \mathbb{K}^{m \times n}$,
- $\lambda(A + B) = \lambda A + \lambda B$ für alle $\lambda \in \mathbb{K}$ und alle $A, B \in \mathbb{K}^{m \times n}$,
- $1 \cdot A = A$ für alle $A \in \mathbb{K}^{m \times n}$.

Zusammenfassend: $(\mathbb{K}^{m \times n}; +, \cdot)$ ist ein \mathbb{K} -Vektorraum.

Matrizen und lineare Abbildungen

Rechenregeln für Matrizen

Naheliegenderweise enthält die Nullmatrix $O \in \mathbb{K}^{m \times n}$ als Einträge nur Nullen, d. h.

$$O = [0]_{1 \leq i \leq m, 1 \leq j \leq n}.$$

Mit $A = [a_{i,j}] \in \mathbb{K}^{m \times n}$ gilt weiterhin $-A = [-a_{i,j}]$.

Für $A, B \in \mathbb{R}^{3 \times 2}$ mit

$$A = \begin{bmatrix} 1 & 3 \\ 0 & -1 \\ 0 & 5 \end{bmatrix} \quad \text{und} \quad B = \begin{bmatrix} -2 & 4 \\ 5 & 1 \\ 0 & -2 \end{bmatrix}$$

berechne man $A + B$ und $A - 3B$.

Ist $A = [a_{i,j}] \in \mathbb{K}^{m \times n}$, dann heißt

$$A^T := [a_{j,i}] \in \mathbb{K}^{n \times m}$$

die **Transponierte** von A . In A^T sind also die Rollen der Zeilen und Spalten von A vertauscht.

Transponiert man einen Vektor $\mathbf{a} = [a_j] \in \mathbb{K}^n$, so ergibt sich ein Zeilenvektor $\mathbf{a}^T = [a_1, a_2, \dots, a_n]$.

Für eine komplexe Matrix $A = [a_{i,j}] \in \mathbb{C}^{m \times n}$ definiert man weiterhin die **Konjugiert-Transponierte** von A :

$$A^H := \overline{A}^T = [\overline{a_{j,i}}] \in \mathbb{C}^{n \times m}.$$

Hat A nur reelle Einträge, so sind A^T und A^H identisch.

Satz 2.10

Für $A, B \in \mathbb{K}^{m \times n}$ und $\lambda \in \mathbb{K}$ gelten:

- $(A^T)^T = A$ und $(A^H)^H = A$,
- $(\lambda A)^T = \lambda A^T$ und $(\lambda A)^H = \bar{\lambda} A^H$,
- $(A + B)^T = A^T + B^T$ und $(A + B)^H = A^H + B^H$.

Man bestimme die Transponierten und Konjugiert-Transponierten von

$$A = \begin{bmatrix} 1 & 3 \\ 0 & -1 \\ 0 & 5 \end{bmatrix} \quad \text{und} \quad B = \begin{bmatrix} 1+i & 3i \\ 0 & -1 \end{bmatrix}$$

Definition 2.11

Für $A = [a_{i,j}] \in \mathbb{K}^{m \times n}$ und $B = [b_{i,j}] \in \mathbb{K}^{n \times p}$ ist das Produkt $C = A \cdot B = [c_{i,j}] \in \mathbb{K}^{m \times p}$ definiert durch

$$c_{i,j} = \sum_{k=1}^n a_{i,k} b_{k,j} \quad (i = 1, 2, \dots, m \text{ und } j = 1, 2, \dots, p).$$

Anmerkung: Das Produkt $C = AB$ ist nur dann erklärt, wenn A so viele Spalten wie B Zeilen hat. In diesem Fall übernimmt das Ergebnis C die Zeilenanzahl von A und die Spaltenanzahl von B , symbolisch:

$$\begin{array}{|c|} \hline A \\ \hline \end{array} \cdot \begin{array}{|c|} \hline B \\ \hline \end{array} = \begin{array}{|c|} \hline C \\ \hline \end{array}$$

Matrizen und lineare Abbildungen

Falk-Schema für Matrizenmultiplikation

					$b_{1,1}$	\cdots	$b_{1,j}$	\cdots	$b_{1,p}$
					$b_{2,1}$	\cdots	$b_{2,j}$	\cdots	$b_{2,p}$
					\vdots		\vdots		\vdots
					\vdots		\vdots		\vdots
					$b_{n,1}$	\cdots	$b_{n,j}$	\cdots	$b_{n,p}$
$a_{1,1}$	$a_{1,2}$	\cdots	\cdots	$a_{1,n}$	$c_{1,1}$	\cdots	$c_{1,j}$	\cdots	$c_{1,p}$
\vdots				\vdots	\vdots		\vdots		\vdots
$a_{i,1}$	$a_{i,2}$	\cdots	\cdots	$a_{i,n}$	$c_{i,1}$	\cdots	$c_{i,j}$	\cdots	$c_{i,p}$
\vdots				\vdots	\vdots		\vdots		\vdots
$a_{m,1}$	$a_{m,2}$	\cdots	\cdots	$a_{m,n}$	$c_{m,1}$	\cdots	$c_{m,j}$	\cdots	$c_{m,p}$

Matrizen und lineare Abbildungen

Anleitung zum Falk-Schema

Nur die grau markierte Zeile bzw. Spalte von A und B geht in die Berechnung von $c_{i,j}$ ein. Folgendes Vorgehen:

- Bilden Sie 'von außen kommend' Zahlenpärchen der Form $(a_{i,k}, b_{k,j})$.
- Multiplizieren Sie jeweils die beiden Zahlen und summieren Sie sämtliche Ergebnisse.

Berechnen Sie für

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 3 & -1 & 2 \end{bmatrix} \quad \text{und} \quad B = \begin{bmatrix} 1 & 2 \\ 0 & -2 \\ 2 & 2 \end{bmatrix}$$

die Produkte AB und BA und vergleichen Sie die Ergebnisse. Was können Sie aus dem Vergleich schließen?

Matrizen und lineare Abbildungen

Warnungen

- Auch wenn beide Produkte AB und BA definiert sind (was beispielsweise für $A, B \in \mathbb{K}^{n \times n}$ der Fall ist), gilt i. A. $AB \neq BA$.
- Aus $AB = O$ (Nullmatrix) folgt **keineswegs** $A = O$ oder $B = O$.
- Selbst aus $A^2 = AA = O$ folgt **nicht** $A = O$.

Berechnen Sie A^2 für

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Finden Sie eine Matrix $B \in \mathbb{R}^{2 \times 3}$, so dass $AB = O$.

Satz 2.12

Im Zusammenhang mit der Matrizenmultiplikation gelten folgende Rechenregeln:

- $(AB)C = A(BC)$ für alle $A \in \mathbb{K}^{m \times n}$, $B \in \mathbb{K}^{n \times p}$, $C \in \mathbb{K}^{p \times q}$,
- $A(B + C) = AB + AC$ für alle $A \in \mathbb{K}^{m \times n}$, $B, C \in \mathbb{K}^{n \times p}$,
- $(A + B)C = AC + BC$ für alle $A, B \in \mathbb{K}^{m \times n}$, $C \in \mathbb{K}^{n \times p}$,
- $\lambda(AB) = (\lambda A)B = A(\lambda B)$ für alle $\lambda \in \mathbb{K}$, $A \in \mathbb{K}^{m \times n}$, $B \in \mathbb{K}^{n \times p}$,
- $(AB)^T = B^T A^T$ und $(AB)^H = B^H A^H$ für alle $A \in \mathbb{K}^{m \times n}$, $B \in \mathbb{K}^{n \times p}$,

Satz 2.13

- Für die m -dimensionale **Einheitsmatrix**

$$I_m := \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \in \mathbb{K}^{m \times m},$$

gilt $I_m A = A$ für alle $A \in \mathbb{K}^{m \times n}$.

- Für die n -dimensionale Einheitsmatrix $I_n \in \mathbb{K}^{n \times n}$ gilt $AI_n = A$ für alle $A \in \mathbb{K}^{m \times n}$,
- Für die wie gekennzeichnet dimensionierten Nullmatrizen gilt $AO_{n \times p} = O_{m \times p}$ und $O_{q \times m}A = O_{q \times n}$ für alle $A \in \mathbb{K}^{m \times n}$.

Matrizen und lineare Abbildungen

Matrix-Vektor-Multiplikation

Ein Spezialfall der Matrizenmultiplikation ist die Multiplikation von Matrix und Vektor.

Für $A = [a_{i,j}] \in \mathbb{K}^{m \times n}$ und $\vec{x} \in \mathbb{K}^n$ ist $\vec{y} = A\vec{x}$ definiert durch

$$y_i = \sum_{j=1}^n a_{i,j} x_j \quad (i = 1, 2, \dots, m).$$

Somit ist $\vec{y} = A\vec{x} = \sum_{j=1}^n x_j \vec{a}_j$ eine Linearkombination der Spalten \vec{a}_j von A :

$$\begin{bmatrix} \vec{y} \end{bmatrix} = \begin{bmatrix} \vec{a}_1 & | & \vec{a}_2 & | & \dots & | & \vec{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} \vec{a}_1 \end{bmatrix} + x_2 \begin{bmatrix} \vec{a}_2 \end{bmatrix} + \dots + x_n \begin{bmatrix} \vec{a}_n \end{bmatrix}.$$

Matrizen und lineare Abbildungen

Lineare Abbildungen zwischen Vektorräumen

Jede Matrix $A \in \mathbb{K}^{m \times n}$ induziert nun über die Matrix-Vektor-Multiplikation eine Abbildung (die wir wieder mit A bezeichnen):

$$A : \mathbb{K}^n \rightarrow \mathbb{K}^m, \quad \vec{x} \mapsto \vec{y} = A\vec{x}.$$

Der Definitionsbereich von A ist \mathbb{K}^n und der Wertebereich von A ist in \mathbb{K}^m enthalten. Es werden Vektoren auf Vektoren abgebildet.

Die Abbildung A besitzt zwei bemerkenswerte Eigenschaften:

- $A(\vec{x}_1 + \vec{x}_2) = A\vec{x}_1 + A\vec{x}_2$ für alle $\vec{x}_1, \vec{x}_2 \in \mathbb{K}^n$,
- $A(\lambda\vec{x}) = \lambda(A\vec{x})$ für alle $\lambda \in \mathbb{K}, \vec{x} \in \mathbb{K}^n$.

Diese beiden Eigenschaften fasst man unter dem Begriff der Linearität zusammen.

Matrizen und lineare Abbildungen

Lineare Abbildungen zwischen Vektorräumen

Die Definition der Linearität für allgemeine Abbildungen zwischen Vektorräumen lautet wie folgt:

Definition 2.14 (Lineare Abbildung)

Seien V und W Vektorräume über \mathbb{K} . Eine Abbildung $f : V \rightarrow W$ heißt **linear**, wenn

- $f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y})$ für alle $\mathbf{x}, \mathbf{y} \in V$,
- $f(\lambda \mathbf{x}) = \lambda f(\mathbf{x})$ für alle $\lambda \in \mathbb{K}$, $\mathbf{x} \in V$.

Sind folgende Abbildungen linear im Sinne von Definition 2.14:

- $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = 3x$,
- $g : \mathbb{R} \rightarrow \mathbb{R}$, $g(x) = 3x + 42$,
- $h : \mathbb{R}^2 \rightarrow \mathbb{R}$, $h(\vec{x}) = [1, 0] \cdot \vec{x}$?

Matrizen und lineare Abbildungen

Lineare Abbildungen von \mathbb{K}^n nach \mathbb{K}^m

Auf Seite 147 hatten wir bereits bestätigt, dass Abbildungen der Form

$$A : \mathbb{K}^n \rightarrow \mathbb{K}^m, \quad \vec{x} \mapsto \vec{y} = A\vec{x},$$

linear sind.

Es stellt sich nun heraus, dass sich jede lineare Abbildung von \mathbb{K}^n nach \mathbb{K}^m in dieser Form schreiben lässt.

Satz 2.15

Sei $f : \mathbb{K}^n \rightarrow \mathbb{K}^m$ eine lineare Abbildung. Dann existiert eine eindeutig bestimmte Matrix $A \in \mathbb{K}^{m \times n}$, so dass

$$f(\vec{x}) = A\vec{x} \quad \text{für alle } \vec{x} \in \mathbb{K}^n.$$

In den Spalten \vec{a}_j von A stehen dabei gerade die Bilder der Einheitsvektoren \vec{e}_j , d. h.

$$\vec{a}_j = A\vec{e}_j \quad (j = 1, \dots, n).$$

Matrizen und lineare Abbildungen

Lineare Abbildungen von \mathbb{K}^n nach \mathbb{K}^m

Beweisidee:

Machen Sie sich die Aussage von Satz 2.15 klar, indem Sie

- verifizieren, dass eine lineare Abbildung $f : \mathbb{K}^n \rightarrow \mathbb{K}^m$ durch Angabe aller Werte $f(\vec{e}_j)$, $j = 1, \dots, n$, eindeutig bestimmt ist,
- die Beziehung $\vec{a}_j = A\vec{e}_j$ bestätigen,
- die beiden genannten Teilschritte zu einer Gesamtargumentation zusammenfügen.

Matrizen und lineare Abbildungen

Kern und Bild einer Matrix

Zu einer gegebenen Matrix $A \in \mathbb{K}^{m \times n}$ (bzw. der zugeordneten linearen Abbildung) heißt

- $\mathcal{N}(A) := \{\vec{x} \in \mathbb{K}^n : A\vec{x} = \vec{0}\} \subseteq \mathbb{K}^n$ **Nullraum** oder **Kern** von A .
- $\mathcal{R}(A) := \{\vec{y} = A\vec{x} \in \mathbb{K}^m : \vec{x} \in \mathbb{K}^n\} \subseteq \mathbb{K}^m$ das **Bild** von A (entspricht dem Wertebereich im bei uns gebrauchten Sinne).

Aufgrund der Linearität besitzen sowohl $\mathcal{N}(A)$ als auch $\mathcal{R}(A)$ eine spezielle Struktur:

Satz 2.16

Sei $A \in \mathbb{K}^{m \times n}$. $\mathcal{N}(A)$ ist ein Unterraum von \mathbb{K}^n . $\mathcal{R}(A)$ ist ein Unterraum von \mathbb{K}^m . Es gilt

$$\dim \mathcal{N}(A) + \dim \mathcal{R}(A) = n.$$

$\dim \mathcal{N}(A)$ heißt **Defekt** von A und $\text{Rang}(A) := \dim \mathcal{R}(A)$ **Rang** von A .

Matrizen und lineare Abbildungen

Kern und Bild einer Matrix

Illustrieren Sie die Aussagen von Satz 2.16 am Beispiel $A = [1, 1] \in \mathbb{R}^{1 \times 2}$.

Verifizieren Sie, dass $\mathcal{N}(A)$ ein Unterraum von \mathbb{K}^n ist.

Bezeichnen $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n \in \mathbb{K}^m$ die Spalten von $A \in \mathbb{K}^{m \times n}$, dann ist

$$\mathcal{R}(A) = \text{span}\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n\}.$$

(Warum?) Insbesondere ist die maximale Anzahl linear unabhängiger Spalten von A gleich dem Rang von A .

Satz 2.17

Für $A \in \mathbb{K}^{m \times n}$ gilt

$$\text{Rang}(A) = \dim \mathcal{R}(A) = \dim \mathcal{R}(A^T).$$

Die maximale Anzahl linear unabhängiger Spalten von A ist also gleich der maximalen Anzahl linear unabhängiger Zeilen, und beide stimmen mit dem Rang von A überein.

Insbesondere gilt

$$\text{Rang}(A) \leq \min\{m, n\}.$$

Bestimmen Sie den Rang der Matrizen $A = [1, 1]$ und $B = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$.

Anmerkung: Effiziente Verfahren zur Rangbestimmung werden wir im nächsten Abschnitt kennenlernen.

Matrizen und lineare Abbildungen

Matrizen in Trapezform

Es gibt aber Matrizen, bei denen man den Rang sofort ablesen kann. Ein Beispiel sind Matrizen in **Trapezform**

$$A = \left[\begin{array}{cccc|ccc} a_{1,1} & a_{1,2} & a_{1,3} & \dots & a_{1,r} & a_{1,r+1} & \dots & a_{1,n} \\ 0 & a_{2,2} & a_{2,3} & \dots & a_{2,r} & a_{2,r+1} & \dots & a_{2,n} \\ 0 & 0 & a_{3,3} & \dots & a_{3,r} & a_{3,r+1} & \dots & a_{3,n} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & a_{r,r} & a_{r,r+1} & \dots & a_{r,n} \\ \hline 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{array} \right] \in \mathbb{K}^{m \times n},$$

mit $a_{j,j} \neq 0$ ($j = 1, \dots, r$), wobei der untere, der rechte oder beide Teile entfallen können. Der Rang solcher Matrizen ist stets gleich r .

Machen Sie sich klar, warum das so ist.

① Grundlagen

② Lineare Algebra

2.1 Vektorräume

2.2 Matrizen und lineare Abbildungen

2.3 Lineare Gleichungssysteme

2.4 Determinanten

2.5 Invertierbare Matrizen

2.6 Orthogonalität, Skalarprodukt und Norm

2.7 Kreuz- und Spatprodukt

2.8 Elemente der analytischen Geometrie

2.9 Orthogonale Abbildungen

2.10 Eigenwerte und Eigenvektoren

2.11 Singulärwertzerlegung

Ein System aus Gleichungen der Form

$$\begin{array}{ccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \cdots & + & a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \cdots & + & a_{2,n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \cdots & + & a_{m,n}x_n & = & b_m \end{array} \quad (2.1)$$

heißt **lineares Gleichungssystem** (LGS), genauer: ein System von m linearen algebraischen Gleichungen in n Unbekannten.

Dabei sind die **Koeffizienten** $a_{i,j} \in \mathbb{K}$ sowie die Zahlen $b_i \in \mathbb{K}$ vorgegeben, während man die Zahlen $x_i \in \mathbb{K}$ zu bestimmen sucht.

Mit der **Koeffizientenmatrix**

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \in \mathbb{K}^{m \times n},$$

dem Vektor der **Unbekannten**

$$\vec{x} = [x_1, x_2, \dots, x_n]^T \in \mathbb{K}^n$$

sowie der **rechten Seite**

$$\vec{b} = [b_1, b_2, \dots, b_m]^T \in \mathbb{K}^m$$

schreibt man (2.1) kürzer als:

$$A\vec{x} = \vec{b}. \quad (2.2)$$

Ein lineares Gleichungssystem $A\vec{x} = \vec{b}$ heißt **homogen**, wenn $\vec{b} = \vec{0}$ und **inhomogen**, wenn $\vec{b} \neq \vec{0}$.

Lineare Gleichungssysteme

Lösbarkeit linearer Gleichungssysteme

Ein homogenes System besitzt immer (mindestens) eine Lösung, nämlich $\vec{x} = \vec{0}$. Die Lösungsmenge von $A\vec{x} = \vec{0}$ ist gerade der Kern $\mathcal{N}(A)$ von A .

Bei inhomogenen Systemen ist die Sache komplizierter. Wir beginnen mit folgendem Ergebnis:

Satz 2.18

Ist $A \in \mathbb{K}^{m \times n}$ und ist $\vec{b} \in \mathcal{R}(A)$, d. h. es gibt (mindestens) ein $\vec{x}_0 \in \mathbb{K}^n$ mit $A\vec{x}_0 = \vec{b}$, dann gilt

$$\begin{aligned}\{\vec{x} \in \mathbb{K}^n : A\vec{x} = \vec{b}\} &= \{\vec{x}_0\} + \mathcal{N}(A) \\ &= \{\vec{x}_0 + \vec{y} : \vec{y} \in \mathcal{N}(A)\}.\end{aligned}\tag{2.3}$$

Lineare Gleichungssysteme

Lösbarkeit linearer Gleichungssysteme

Im Falle der Existenz einer Lösung erhält man also die Lösungsmenge des inhomogenen Systems durch Verschieben der Lösungsmenge des homogenen Systems zur gleichen Koeffizientenmatrix.

Machen Sie sich dies anhand des relativ trivialen Gleichungssystems

$$x_1 + x_2 = 7$$

graphisch klar. Verwenden Sie im Ansatz (2.3) aus Satz 2.18 auch verschiedene \vec{x}_0 .

Verifizieren Sie die Beziehung (2.3).

Lineare Gleichungssysteme

Lösbarkeit linearer Gleichungssysteme

Zur weiteren Analyse der Lösbarkeit von LGS beleuchten wir die Lösungssuche noch unter einem weiteren Aspekt.

Sind $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ die Spalten der Matrix $A \in \mathbb{K}^{m \times n}$, dann lässt sich $A\vec{x} = \vec{b}$ auch schreiben als (vgl. S. 146):

$$x_1\vec{a}_1 + x_2\vec{a}_2 + \dots + x_n\vec{a}_n = \vec{b}$$

Die Suche nach einer Lösung von $A\vec{x} = \vec{b}$ ist also gleichbedeutend mit der Suche nach Linearkombinationen der Spalten \vec{a}_j , die die rechte Seite \vec{b} ergeben.

Wir halten somit fest:

Satz 2.19 (Lösbarkeit linearer Gleichungssysteme)

Es seien $A \in \mathbb{K}^{m \times n}$ mit den Spalten $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ und $\vec{b} \in \mathbb{K}^m$.

Das lineare Gleichungssystem $A\vec{x} = \vec{b}$ ist genau dann lösbar, wenn

$$\vec{b} \in \mathcal{R}(A) = \text{span}\{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n\}.$$

Das lineare Gleichungssystem $A\vec{x} = \vec{b}$ ist genau dann **eindeutig** lösbar, wenn $\vec{b} \in \mathcal{R}(A)$ gilt, und die Spalten $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ linear unabhängig sind:

$$A\vec{x} = \vec{b} \text{ ist eindeutig lösbar} \Leftrightarrow \vec{b} \in \mathcal{R}(A) \text{ und } \text{Rang}(A) = n.$$

Zusammenfassung: Es können somit genau drei Fälle eintreten:

- $A\vec{x} = \vec{b}$ besitzt **keine** Lösung, d.h. $\vec{b} \notin \mathcal{R}(A)$ (kann nur bei inhomogenen LGS passieren).
- $A\vec{x} = \vec{b}$ besitzt **genau eine** Lösung, d.h. $\vec{b} \in \mathcal{R}(A)$ und $\text{Rang}(A) = n$.
Dieser Fall kann nur für $m \geq n$ eintreten (mindestens so viele Gleichungen wie Unbekannte).
- $A\vec{x} = \vec{b}$ besitzt **unendlich viele** Lösungen, d.h. $\vec{b} \in \mathcal{R}(A)$ und $\text{Rang}(A) < n$.
Mit einer beliebigen Lösung \vec{x}_0 erhält man als Lösungsmenge

$$\vec{x}_0 + \mathcal{N}(A) = \vec{x}_0 + \{\text{alle Lösungen von } A\vec{x} = \vec{0}\}.$$

Lineare Gleichungssysteme mit mehr als einer, aber nur endlich vielen Lösungen existieren nicht.

Lineare Gleichungssysteme

Berechnung der Lösung

Bislang wissen wir zwar über die Lösbarkeit eines LGS $A\vec{x} = \vec{b}$ Bescheid; können aber die Lösung selbst noch nicht konkret berechnen.

Für diese Aufgabe stehen verschiedenste Algorithmen zur Verfügung, so zum Beispiel die hier behandelten:

- Gauß-Algorithmus,
- Gauß-Jordan-Algorithmus – eine Erweiterung des Gauß-Algorithmus.

Desweiteren gibt es eine ganze Reihe numerischer Algorithmen, die Sie mittels Literatur oder in einer Numerik-Vorlesung erlernen können.

Die hier vorgestellten exakten Verfahren beruhen auf Umformungen in ein äquivalentes Gleichungssystem $\tilde{A}\vec{x} = \tilde{\vec{b}}$, für das man die Lösung relativ mühelos angeben kann.

Zunächst erzeugen wir aus der Koeffizientenmatrix $A \in \mathbb{K}^{m \times n}$ und der rechten Seite $b \in \mathbb{K}^m$ die **erweiterte Koeffizientenmatrix** $[A | \vec{b}]$. Diese enthält die vollständige Information über das betrachtete LGS.

Satz 2.20

Die folgenden **elementaren Umformungen** der erweiterten Koeffizientenmatrix verändern die Lösungsmenge eines linearen Gleichungssystems nicht:

- Multiplikation einer Gleichung (Zeile) mit $\lambda \in \mathbb{K}$, $\lambda \neq 0$,
- Addition des Vielfachen einer Gleichung (Zeile) zu einer anderen Gleichung (Zeile),
- Vertauschung von zwei Gleichungen (Zeilen),
- Umnummerierung von zwei Unbekannten (Vertauschung von zwei Spalten).
Hier muss man sich allerdings merken, welche Unbekannten man umnummeriert hat.

Lineare Gleichungssysteme

Gauß-Algorithmus

Beim Gauß-Algorithmus wird das LGS $A\vec{x} = \vec{b}$ in ein äquivalentes System $\tilde{A}\vec{x} = \tilde{b}$ mit einer Matrix $[\tilde{A} | \tilde{b}]$ in Trapezform überführt:

$$\left[\begin{array}{cccccccc|c} \tilde{a}_{1,1} & \tilde{a}_{1,2} & \tilde{a}_{1,3} & \dots & \tilde{a}_{1,r} & \tilde{a}_{1,r+1} & \dots & \tilde{a}_{1,n} & \tilde{b}_1 \\ 0 & \tilde{a}_{2,2} & \tilde{a}_{2,3} & \dots & \tilde{a}_{2,r} & \tilde{a}_{2,r+1} & \dots & \tilde{a}_{2,n} & \tilde{b}_2 \\ 0 & 0 & \tilde{a}_{3,3} & \dots & \tilde{a}_{3,r} & \tilde{a}_{3,r+1} & \dots & \tilde{a}_{3,n} & \tilde{b}_3 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \tilde{a}_{r,r} & \tilde{a}_{r,r+1} & \dots & \tilde{a}_{r,n} & \tilde{b}_r \end{array} \right]$$

mit $\tilde{a}_{j,j} \neq 0$ ($j = 1, \dots, r$).

Die Lösung solcher 'gestaffelter' Gleichungssysteme ist durch Rückwärtsauflösen leicht zu bestimmen.

Lineare Gleichungssysteme

Symbolischer Fortgang des Algorithmus

$$\left[\begin{array}{cccccccc|c} * & * & * & \dots & * & * & * & \dots & * & * \\ 0 & * & * & \dots & * & * & * & \dots & * & * \\ 0 & * & * & \dots & * & * & * & \dots & * & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & * & * & \dots & * & * & * & \dots & * & * \\ 0 & * & * & \dots & * & * & * & \dots & * & * \end{array} \right]$$

$$\left[\begin{array}{cccccccc|c} * & * & * & \dots & * & * & * & \dots & * & * \\ 0 & * & * & \dots & * & * & * & \dots & * & * \\ 0 & 0 & * & \dots & * & * & * & \dots & * & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & * & \dots & * & * & * & \dots & * & * \\ 0 & 0 & * & \dots & * & * & * & \dots & * & * \end{array} \right]$$

⋮

$$\left[\begin{array}{cccccccc|c} * & * & * & \dots & * & * & * & \dots & * & * \\ 0 & * & * & \dots & * & * & * & \dots & * & * \\ 0 & 0 & * & \dots & * & * & * & \dots & * & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & * & * & * & \dots & * & * \\ 0 & 0 & 0 & \dots & 0 & * & * & \dots & * & * \end{array} \right]$$

roter Stern: Eintrag zwingend ungleich Null; schwarzer Stern: beliebiger Eintrag

Lineare Gleichungssysteme

Exakte Beschreibung des Verfahrens

Vor jedem Einzelschritt werden zunächst alle Zeilen der Form

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

ersatzlos gestrichen.

Gibt es Zeilen der Form

$$0 \ 0 \ \dots \ 0 \mid \tilde{b}_j,$$

(mit $\tilde{b}_j \neq 0$) so wird der Algorithmus abgebrochen; das LGS hat dann keine Lösung.

Warum sind diese Schritte sinnvoll?

Schritt 1:

- Ist $a_{1,1} = 0$, so suche in der ersten Spalte nach einem Eintrag $a_{j,1} \neq 0$. Existiert ein solcher, tausche Zeilen 1 und j .
- Ist $a_{1,1} = 0$ und der erste Punkt war nicht erfolgreich, so existiert in der ersten Zeile ein Eintrag $a_{1,k} \neq 0$
Tausche in diesem Fall Spalten 1 und k ; merke wie die Unbekannten unnummeriert wurden,
- Ist jetzt $\tilde{a}_{1,1} \neq 0$, so erzeuge unterhalb von $\tilde{a}_{1,1}$ lauter Nullen durch Addition des jeweils $(-\frac{\tilde{a}_{i,1}}{\tilde{a}_{1,1}})$ -fachen der ersten Zeile zur Zeile i .

Die erste Zeile und Spalte von $[\tilde{A} | \tilde{b}]$ ist damit ermittelt und wird nicht mehr angefasst.

Schritt 2 etc.:

Nach Ausführung der Vorarbeiten (S. 167) wird nun mit der um die erste Zeile und Spalte reduzierten Matrix analog verfahren usw.

Im Falle der Existenz einer Lösung steht am Ende des Gauß-Algorithmus ein äquivalentes System der Form

$$\begin{bmatrix} \tilde{a}_{1,1} & \tilde{a}_{1,2} & \tilde{a}_{1,3} & \cdots & \tilde{a}_{1,r} & \tilde{a}_{1,r+1} & \cdots & \tilde{a}_{1,n} \\ 0 & \tilde{a}_{2,2} & \tilde{a}_{2,3} & \cdots & \tilde{a}_{2,r} & \tilde{a}_{2,r+1} & \cdots & \tilde{a}_{2,n} \\ 0 & 0 & \tilde{a}_{3,3} & \cdots & \tilde{a}_{3,r} & \tilde{a}_{3,r+1} & \cdots & \tilde{a}_{3,n} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \tilde{a}_{r,r} & \tilde{a}_{r,r+1} & \cdots & \tilde{a}_{r,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_r \end{bmatrix}.$$

Die Variablen x_{r+1}, \dots, x_n können als Parameter frei gewählt werden ($x_{r+1} = \lambda_{r+1}, \dots, x_n = \lambda_n$).

Die anderen Variablen x_r, x_{r-1}, \dots, x_1 lassen sich dann bestimmen, indem man die Gleichungen von unten her sukzessive auflöst.

Achtung: Umnummerierungen muss man natürlich berücksichtigen!

Trainieren Sie den Gauß-Algorithmus:

1. Bestimmen Sie die Lösungen von $A\vec{x} = \vec{b}$ und $A\vec{x} = \vec{0}$ für

$$A = \begin{bmatrix} 2 & -1 & -1 & 2 \\ 6 & -2 & 3 & -1 \\ -4 & 2 & 3 & -2 \\ 2 & 0 & 4 & -3 \end{bmatrix} \quad \text{und} \quad \vec{b} = \begin{bmatrix} 3 \\ -3 \\ -2 \\ -1 \end{bmatrix}$$

2. Bestimmen Sie die Lösungen von

$$6x_1 + 4x_2 + 8x_3 + 17x_4 = -20$$

$$3x_1 + 2x_2 + 5x_3 + 8x_4 = -8$$

$$3x_1 + 2x_2 + 7x_3 + 7x_4 = -4$$

$$2x_3 - x_4 = 4$$

3. Wieviele Lösungen besitzt das LGS $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$?

Lineare Gleichungssysteme

Rangbestimmung mittels Gauß-Verfahren

Die elementaren Umformungen von S. 164 lassen nicht nur die Lösung eines LGS unverändert.

Sie erhalten desweiteren natürlich auch den Rang der Koeffizientenmatrix. Es gilt also mit den bisherigen Bezeichnungen und den Erkenntnissen von S. 154:

$$\text{Rang}(A) = \text{Rang}(\tilde{A}) = r.$$

Somit steht uns mit dem Gauß-Verfahren auch ein effizientes Verfahren zur Rangbestimmung von Matrizen zur Verfügung.

Geben Sie die Ränge der Koeffizientenmatrizen aus den Beispielen von S. 170 an.

Beim **Gauß-Jordan-Algorithmus** wird das LGS $A\vec{x} = \vec{b}$, $A \in \mathbb{K}^{m \times n}$, ebenfalls in ein äquivalentes System mit Trapezform überführt.

Allerdings führt man die Umformungen noch weiter und erzeugt schließlich eine Matrix $[\tilde{A} | \tilde{\vec{b}}]$, in der \tilde{A} folgende Blockstruktur besitzt:

$$\tilde{A} = [I_r \ R] = \begin{bmatrix} 1 & 0 & \dots & 0 & * & \dots & * \\ 0 & 1 & \dots & 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \\ 0 & 0 & \dots & 1 & * & \dots & * \end{bmatrix}. \quad (2.4)$$

Dabei ist I_r die r -dimensionale Einheitsmatrix und $R \in \mathbb{K}^{r \times (n-r)}$ eine beliebige Restmatrix, die auch entfallen kann

Wir wollen hier nur den Fall betrachten, dass $A\vec{x} = \vec{b}$ eindeutig lösbar ist ($r = n$). Das Gleichungssystem wird dann in ein äquivalentes System $I_n\vec{x} = \tilde{\vec{b}}$ überführt, dessen Lösung $\vec{x} = \tilde{\vec{b}}$ direkt abgelesen werden kann.

Lineare Gleichungssysteme

Symbolischer Fortgang des Gauß-Jordan-Algorithmus

(A quadratisch mit vollem Rang)

$$\left[\begin{array}{cccccc|c} 1 & * & * & \dots & * & * & * \\ 0 & * & * & \dots & * & * & * \\ 0 & * & * & \dots & * & * & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & * & * & \dots & * & * & * \\ 0 & * & * & \dots & * & * & * \end{array} \right] \rightarrow \left[\begin{array}{cccccc|c} 1 & 0 & * & \dots & * & * & * \\ 0 & 1 & * & \dots & * & * & * \\ 0 & 0 & * & \dots & * & * & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & * & \dots & * & * & * \\ 0 & 0 & * & \dots & * & * & * \end{array} \right]$$

$$\rightarrow \dots \rightarrow \left[\begin{array}{cccccc|c} 1 & 0 & 0 & \dots & 0 & 0 & * \\ 0 & 1 & 0 & \dots & 0 & 0 & * \\ 0 & 0 & 1 & \dots & 0 & 0 & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & * \\ 0 & 0 & 0 & \dots & 0 & 1 & * \end{array} \right]$$

Bei Bedarf können auch mehrere rechte Seiten parallel mitgeführt werden.

Lineare Gleichungssysteme

Vorgehen im k -ten Schritt des Gauß-Jordan-Algorithmus

- Ist $a_{k,k} = 0$, dann tausche die Zeilen (oder Spalten) wie beim Gauß-Algorithmus, so dass $\tilde{a}_{k,k} \neq 0$.
- Teile nun die k -te Zeile durch $\tilde{a}_{k,k}$, d. h. erzeuge eine 1 an Position (k, k) .
- Erzeuge unter- und oberhalb der Position (k, k) in **allen** Zeilen Nullen durch Addition des jeweils $(-\tilde{a}_{i,k})$ -fachen der k -ten Zeile zur Zeile i .

Wie beim Gauß-Algorithmus muss dabei über unterwegs vorgenommene Umnummerierungen von Variablen (Spaltenvertauschungen) sorgfältig buchgeführt werden.

Lineare Gleichungssysteme

Bestimmung der Lösung im Gauß-Jordan-Algorithmus

Fehlt in (2.4) der Block R , so ist das Gleichungssystem $A\vec{x} = \vec{b}$ äquivalent zu $I_n\vec{x} = \vec{b}$. Es gibt eine eindeutige Lösung, die man direkt ablesen kann:

$$\vec{x} = \vec{b}.$$

Ist das Gleichungssystem äquivalent zu $[I_r \ R]\vec{x} = \vec{b}$ mit $r < n$, so teilt man \vec{x} in zwei Teile $\vec{x}^1 = [x_1, \dots, x_r]^T$ und $\vec{x}^2 = [x_{r+1}, \dots, x_n]^T$.

Die Variablen x_{r+1}, \dots, x_n können als Parameter frei gewählt werden. Wegen

$$\vec{b} = \begin{bmatrix} I_r & R \end{bmatrix} \begin{bmatrix} \vec{x}^1 \\ \vec{x}^2 \end{bmatrix} = \vec{x}^1 + R\vec{x}^2$$

ist die Lösung von $A\vec{x} = \vec{b}$ gegeben durch

$$\mathcal{L} = \left\{ \vec{x} = \begin{bmatrix} \vec{x}^1 \\ \vec{x}^2 \end{bmatrix} : \vec{x}^2 \in \mathbb{R}^{n-r} \text{ beliebig und } \vec{x}^1 = \vec{b} - R\vec{x}^2 \right\}.$$

Lineare Gleichungssysteme

Gauß- vs. Gauß-Jordan-Verfahren

Ob man ein LGS mittels Gauß- oder Gauß-Jordan-Verfahren löst, ist letztlich Geschmackssache.

Beim Gauß-Jordan-Verfahren spart man das Rückwärtseinsetzen, muss aber eine höhere Anzahl elementarer Umformungen in Kauf nehmen.

Besonders vorteilhaft ist Gauß-Jordan jedoch dann, wenn es um die Lösung mehrerer LGS mit gleicher Koeffizientenmatrix $A \in \mathbb{R}^{n \times n}$, $\text{rang}(A) = n$, und verschiedenen rechten Seiten geht:

$$A\vec{x} = \vec{b}_1, \quad A\vec{x} = \vec{b}_2, \quad \dots, \quad A\vec{x} = \vec{b}_\ell$$

Diese Aufgabe ist gleichbedeutend mit der Lösung der Matrixgleichung

$$AX = B,$$

wobei $X \in \mathbb{R}^{n \times \ell}$ gesucht ist, und $B \in \mathbb{R}^{n \times \ell}$ als Spalten gerade die rechten Seiten \vec{b}_i enthält.

Lineare Gleichungssysteme

Beispiel Gauß-Jordan-Verfahren

Mit Hilfe des Gauß-Jordan-Algorithmus bestimme man die Lösung der Matrixgleichung $AX = B$, wobei

$$A = \begin{bmatrix} 3 & 5 & 1 \\ 2 & 4 & 5 \\ 1 & 2 & 2 \end{bmatrix} \quad \text{und} \quad B = \begin{bmatrix} 6 & 13 \\ 17 & 10 \\ 7 & 5 \end{bmatrix}.$$

① Grundlagen

② Lineare Algebra

2.1 Vektorräume

2.2 Matrizen und lineare Abbildungen

2.3 Lineare Gleichungssysteme

2.4 Determinanten

2.5 Invertierbare Matrizen

2.6 Orthogonalität, Skalarprodukt und Norm

2.7 Kreuz- und Spatprodukt

2.8 Elemente der analytischen Geometrie

2.9 Orthogonale Abbildungen

2.10 Eigenwerte und Eigenvektoren

2.11 Singulärwertzerlegung

Die **Determinante** ist eine Abbildung, die einer quadratischen Matrix $A \in \mathbb{K}^{n \times n}$ eine Zahl aus \mathbb{K} zuordnet:

$$\begin{aligned}\det : \mathbb{K}^{n \times n} &\rightarrow \mathbb{K} \\ A &\mapsto \det A\end{aligned}$$

Sie wird u. A. benutzt zur Beschreibung

- der eindeutigen Lösbarkeit von LGS,
- der linearen Unabhängigkeit von n Vektoren,
- der Veränderung von Flächen/Volumina bei Anwendung der Abbildung A .

Historisches zum Determinantenbegriff

Determinanten mit $n = 2$ wurden erstmals am Ende des 16. Jh. von Cardano, größere ca. 100 Jahre später von Leibniz behandelt. Ein moderner axiomatischer Ansatz (1864) geht auf Weierstraß zurück.

Determinanten

Determinanten für $n \leq 3$

Wir werden den axiomatischen Aufbau hier umgehen und für die Fälle mit $n \leq 3$ explizite Formeln angeben. Für $n > 3$ definieren wir die Determinante induktiv.

$$n = 1 : \quad A = \begin{bmatrix} a_{1,1} \end{bmatrix}, \quad \det A := a_{1,1}.$$

$$n = 2 : \quad A = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}, \quad \det A := a_{1,1}a_{2,2} - a_{1,2}a_{2,1}.$$

$$n = 3 : \quad A = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix},$$

$$\det A := +a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} \\ - a_{1,3}a_{2,2}a_{3,1} - a_{1,1}a_{2,3}a_{3,2} - a_{1,2}a_{2,1}a_{3,3}.$$

Determinanten

Regel von Sarrus*

Im Fall $n = 3$ verwendet man zur Berechnung einer Determinante gerne folgendes Schema:

$$\begin{array}{ccccc} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,1} & a_{3,2} \\ & - & - & -, + & + & + \end{array}$$

- Man schreibt 1. und 2. Spalte nochmals neben die Determinante.
- Entlang der Diagonalen ermittelt man die Produkte der Einträge und versieht die Ergebnisse mit den dargestellten Vorzeichen.
- Man summiert die 6 vorzeichenbehafteten Produkte.

* Pierre Frédéric Sarrus, 1798-1861, französischer Mathematiker

Determinanten

Determinanten für $n \geq 4$

Für Determinanten mit $n \geq 4$ gehen wir induktiv vor, greifen also auf Determinanten geringerer Größe zurück.

Für $i, j \in \{1, 2, \dots, n\}$ bezeichne $A_{i,j} \in \mathbb{K}^{(n-1) \times (n-1)}$ diejenige Matrix, die durch Streichen der Zeile i und der Spalte j aus A entsteht.

Sei j nun ein beliebiger Spaltenindex. Dann definieren wir

$$\det A := \sum_{i=1}^n a_{i,j} (-1)^{i+j} \det A_{i,j} \quad (2.5)$$

Diese Darstellung wird **Laplace-Entwicklung** nach der j -ten Spalte genannt.

- Natürlich müsste man zunächst zeigen, dass die Darstellung in (2.5) von der gewählten Spalte unabhängig ist.
- Bei Berechnungen mit Hilfe der Laplace-Entwicklung ist es häufig zweckmäßig, Spalten (oder Zeilen, siehe später) mit möglichst vielen Nulleinträgen zu wählen.
- Die Laplace-Entwicklung (2.5) gilt auch für Determinanten mit $n = 2$ oder 3 .
- In der Literatur wird die Determinante häufig anders aufgebaut. Formel (2.5) ist dann ein Bestandteil des **Laplaceschen Entwicklungssatzes**.

Determinanten

Notation, Beispiele

Man verwendet für Determinanten mit $n \geq 2$ auch folgende verkürzende Schreibweise:

$$\begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{vmatrix} := \det \left(\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix} \right).$$

Berechnen Sie folgende Determinanten:

$$\begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix}, \begin{vmatrix} 1 & 0 & 2 \\ 3 & -1 & 1 \\ 3 & 1 & 5 \end{vmatrix}, \begin{vmatrix} 1 & 42 & 23 \\ 0 & 17 & -110 \\ 0 & 0 & 3 \end{vmatrix}, \begin{vmatrix} 1 & 0 & 1 & 3 \\ -1 & 2 & 0 & 0 \\ 0 & 2 & -1 & 0 \\ 1 & -2 & 0 & -1 \end{vmatrix}.$$

Satz 2.21

Für $A = [a_{i,j}] \in \mathbb{K}^{n \times n}$ und $\lambda \in \mathbb{K}$ gelten:

- *Multipliziert man eine Spalte von A mit λ , so multipliziert sich auch die Determinante mit λ :*

$$\begin{vmatrix} a_{1,1} & \cdots & \lambda a_{1,j} & \cdots & a_{1,n} \\ a_{2,1} & \cdots & \lambda a_{2,j} & \cdots & a_{2,n} \\ \vdots & & \vdots & & \vdots \\ a_{n,1} & \cdots & \lambda a_{n,j} & \cdots & a_{n,n} \end{vmatrix} = \lambda \cdot \begin{vmatrix} a_{1,1} & \cdots & a_{1,j} & \cdots & a_{1,n} \\ a_{2,1} & \cdots & a_{2,j} & \cdots & a_{2,n} \\ \vdots & & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j} & \cdots & a_{n,n} \end{vmatrix}$$

- *Die Determinante einer Matrix A ändert sich nicht, wenn man zu einer Spalte das Vielfache einer **anderen** Spalte addiert.*
- *Vertauscht man in einer quadratischen Matrix zwei verschiedene Spalten, so multipliziert sich die Determinante mit -1 .*

Satz 2.22

Seien $A, B \in \mathbb{K}^{n \times n}$, $A = [a_{i,j}]$ und $\lambda \in \mathbb{K}$. Dann gelten:

- $\det(\lambda A) = \lambda^n \det A$.
- $\det(A^T) = \det A$.
- $\det(AB) = \det(A) \det(B)$.
- Ist A eine Dreiecksmatrix (d.h. $a_{i,j} = 0$ für alle $i < j$ oder alle $i > j$), dann gilt

$$\det A = \prod_{i=1}^n a_{i,i},$$

d.h. die Determinante ist das Produkt der (Haupt-) Diagonalelemente.

- Wegen (2.5) und Satz 2.22, Punkt 2, lässt sich eine Determinante auch nach der i -ten Zeile entwickeln:

$$\det(A) = \sum_{j=1}^n a_{i,j} (-1)^{i+j} \det(A_{i,j})$$

- Elementare Zeilen- und Spaltenumformungen wie im Gauß- Algorithmus können die Determinante der Koeffizientenmatrix ändern.
Die Eigenschaft, gleich oder ungleich Null zu sein, bleibt dabei aber immer erhalten.
- Die Determinante einer Dreiecksmatrix ist genau dann Null, wenn für eines der Diagonalelemente $a_{i,i} = 0$ gilt.

Determinanten

Determinante und Rang

Kombiniert man die letzten beiden Punkte mit den Kenntnissen über den Endzustand der Matrix \tilde{A} beim Gauß-Algorithmus, so ergibt sich weiterhin:

- Ist $\text{rang}(A) = n$, so ist $\det A \neq 0$.
- Ist $\text{rang}(A) < n$, so ist $\det A = 0$.

Dabei ist natürlich immer $A \in \mathbb{K}^{n \times n}$ anzusetzen.

Das bedeutet wiederum, dass man die eindeutige Lösbarkeit eines linearen Gleichungssystems leicht mit der Determinante verifizieren kann.

Bestimmen Sie alle $\alpha \in \mathbb{R}$, für die das LGS $A\vec{x} = \vec{b}$ mit

$$A = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & 2\alpha & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

und beliebigem $\vec{b} \in \mathbb{R}^3$ eine eindeutige Lösung besitzt

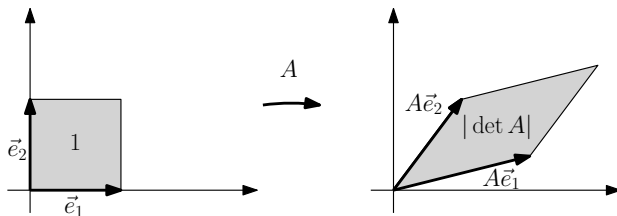
Determinanten

Geometrische Interpretation

Sei $A \in \mathbb{R}^{n \times n}$ und $\text{vol}(S)$ das n -dimensionale Volumen* einer geeigneten** Punktmenge S . Dann ist das Volumen des Bildes $f(S)$ unter $f(\vec{x}) = A\vec{x}$ gegeben durch

$$\text{vol}(f(S)) = |\det(A)| \cdot \text{vol}(S).$$

Determinanten können also als Flächen- oder Volumenverzerrungsfaktor interpretiert werden:



* in 2D ist das eine Fläche, in 3D das gewohnte Volumen

**'Geeignet' lässt sich im Rahmen der Maßtheorie sauber definieren. Gängige geometrische Objekte machen i.d.R. kein Problem.

① Grundlagen

② Lineare Algebra

2.1 Vektorräume

2.2 Matrizen und lineare Abbildungen

2.3 Lineare Gleichungssysteme

2.4 Determinanten

2.5 Invertierbare Matrizen

2.6 Orthogonalität, Skalarprodukt und Norm

2.7 Kreuz- und Spatprodukt

2.8 Elemente der analytischen Geometrie

2.9 Orthogonale Abbildungen

2.10 Eigenwerte und Eigenvektoren

2.11 Singulärwertzerlegung

Definition 2.23

Eine quadratische Matrix $A \in \mathbb{K}^{n \times n}$ heißt **invertierbar** oder **regulär**, wenn es eine Matrix $B \in \mathbb{K}^{n \times n}$ gibt mit

$$AB = BA = I_n. \quad (2.6)$$

Die Matrix B ist dann eindeutig durch A bestimmt, wird die **Inverse** von A genannt und mit A^{-1} bezeichnet.

Warum ist $B = A^{-1}$ eindeutig bestimmt? Sie finden die Antwort leicht, wenn Sie (2.6) als Menge von LGS der Form $A\vec{x} = \vec{e}_i$ lesen.

Invertierbare Matrizen

Berechnung der Inversen

Die Inverse einer invertierbaren 2×2 -Matrix

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}$$

kann man explizit angeben:

$$A^{-1} = \frac{1}{\det(A)} \begin{bmatrix} a_{2,2} & -a_{1,2} \\ -a_{2,1} & a_{1,1} \end{bmatrix} = \frac{1}{a_{1,1}a_{2,2} - a_{1,2}a_{2,1}} \begin{bmatrix} a_{2,2} & -a_{1,2} \\ -a_{2,1} & a_{1,1} \end{bmatrix}.$$

Für $n \geq 3$ existieren zwar auch Formeln, diese sind jedoch sperrig und kaum in Gebrauch.

Bestätigen Sie obige Formel durch Nachrechnen.

Invertierbare Matrizen

Berechnung der Inversen

Im allgemeinen Fall muss man die Matrixgleichung $AX = I_n$ lösen, oder eben sämtliche LGS

$$A\vec{x} = \vec{e}_i, \quad i = 1, \dots, n.$$

Für diese Aufgabe bietet sich der auf Folie 172 ff behandelte Gauß-Jordan-Algorithmus an.

$$\begin{array}{c|c} A & I_n \\ \hline \vdots & \vdots \\ \hline I_n & A^{-1} \end{array}$$

Man berechne die Inversen von

$$A = \begin{bmatrix} 1 & 23 \\ 0 & 42 \end{bmatrix} \quad \text{und} \quad B = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 3 & 0 \\ 3 & 4 & 1 \end{bmatrix}.$$

Satz 2.24

Sei $A \in \mathbb{K}^{n \times n}$. Dann sind folgende Aussagen äquivalent:

- A ist invertierbar.
- $\det A \neq 0$.
- Es gilt $\text{rang}(A) = n$, d.h. die Spalten (Zeilen) von A bilden eine Basis des \mathbb{K}^n .
- Das homogene System $A\vec{x} = \vec{0}$ besitzt nur die triviale Lösung $\vec{x} = \vec{0}$.
- Für jede rechte Seite $\vec{b} \in \mathbb{K}^n$ besitzt das System $A\vec{x} = \vec{b}$ genau eine Lösung, nämlich

$$\vec{x} = A^{-1}\vec{b}.$$

Satz 2.25

Seien $A, B \in \mathbb{K}^{n \times n}$ invertierbar. Dann gelten:

- A^T und A^H sind invertierbar, und es gilt

$$(A^T)^{-1} = (A^{-1})^T \quad \text{bzw.} \quad (A^H)^{-1} = (A^{-1})^H.$$

- A^{-1} ist invertierbar, und es gilt

$$(A^{-1})^{-1} = A.$$

- AB ist invertierbar, und es gilt

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

Notation: Da es nach dem ersten Anstrich egal ist, welche der beiden Operationen zuerst ausgeführt wird, schreibt man auch, etwas weniger umständlich, A^{-T} bzw. A^{-H} .

① Grundlagen

② Lineare Algebra

2.1 Vektorräume

2.2 Matrizen und lineare Abbildungen

2.3 Lineare Gleichungssysteme

2.4 Determinanten

2.5 Invertierbare Matrizen

2.6 Orthogonalität, Skalarprodukt und Norm

2.7 Kreuz- und Spatprodukt

2.8 Elemente der analytischen Geometrie

2.9 Orthogonale Abbildungen

2.10 Eigenwerte und Eigenvektoren

2.11 Singulärwertzerlegung

- Visualisieren wir zwei Vektoren \vec{x}, \vec{y} aus \mathbb{R}^2 oder \mathbb{R}^3 als Pfeile mit Startpunkt in $\vec{0}$, so haben wir eine recht konkrete Vorstellung, was mit 'senkrecht' oder dem 'Winkel zwischen \vec{x} und \vec{y} ' gemeint ist.
- Die bisherigen Überlegungen blenden diese Anschauung aber noch völlig aus – es fehlen mathematische Begriffe zur Beschreibung solcher Eigenschaften.
- Um diesen Mangel zu beheben, benötigen wir den Begriff des **Innen-** oder **Skalarprodukts**.
- Wir geben zunächst die allgemeine Form, ziehen uns aber dann im wesentlichen auf \mathbb{R}^n (manchmal \mathbb{C}^n) zurück.

Definition 2.26

Sei V ein \mathbb{K} -Vektorraum, wobei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Dann heißt eine Abbildung

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{K}$$

Innenprodukt oder **Skalarprodukt**, wenn für alle $\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{y} \in V$ und $\lambda \in \mathbb{K}$ gilt:

- $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$, wobei $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$ ('positiv definit'),
- $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$, falls $\mathbb{K} = \mathbb{R}$ ('symmetrisch'),
 $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$, falls $\mathbb{K} = \mathbb{C}$ ('hermitesch'),
- $\langle \lambda \mathbf{x}_1 + \mathbf{x}_2, \mathbf{y} \rangle = \lambda \langle \mathbf{x}_1, \mathbf{y} \rangle + \langle \mathbf{x}_2, \mathbf{y} \rangle$ ('linear im ersten Argument').

Wir nennen $\mathbf{x}, \mathbf{y} \in V$ **orthogonal** ($\mathbf{x} \perp \mathbf{y}$), wenn $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ gilt.

Ein Vektorraum, auf dem ein Skalarprodukt definiert ist, heißt **Innenproduktraum** oder **Prä-Hilbert-Raum** (Ist $\dim V < \infty$ so ist V dann bereits ein Hilbert-Raum.)

Hinweis: Denken Sie am besten bereits hier an das Standard-Skalarprodukt $\langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^n x_i y_i$ im \mathbb{R}^n (auch als **Euklidisches Skalarprodukt** bekannt)

Orthogonalität, Skalarprodukt und Norm

Zum Skalarprodukt gehörende Norm

Hat man einmal ein Skalarprodukt festgelegt, so ist damit immer eine 'Längenmessung' für Vektoren verbunden:

Satz 2.27 (und Definition)

Sei V ein \mathbb{K} -Vektorraum mit $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ und $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf V .
Dann gilt **Cauchy-Schwarzsche-Ungleichung**

$$|\langle \mathbf{x}, \mathbf{y} \rangle|^2 \leq \langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle, \quad (\mathbf{x}, \mathbf{y} \in V) \quad (\text{CSU})$$

mit Gleichheit genau dann, wenn \mathbf{x} und \mathbf{y} linear abhängig sind. Die durch

$$\|\mathbf{x}\| := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \quad (\mathbf{x} \in V)$$

definierte Abbildung heißt die von diesem Skalarprodukt erzeugte **Norm**. Jede Norm erfüllt für alle $\mathbf{x}, \mathbf{y} \in V, \lambda \in \mathbb{K}$ folgende Beziehungen:

- $\|\mathbf{x}\| \geq 0$, wobei $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$ ('positiv definit'),
- $\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|$ ('homogen'),
- $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ ('Dreiecksungleichung').

Orthogonalität, Skalarprodukt und Norm

Beispiele

Das (für uns) bei weitem wichtigste Skalarprodukt ist das **Euklidische Skalarprodukt**. Auf \mathbb{R}^n ist es definiert durch

$$\langle \vec{x}, \vec{y} \rangle := \vec{y}^T \vec{x} = \sum_{i=1}^n x_i y_i \quad (\vec{x}, \vec{y} \in \mathbb{R}^n).$$

Die Entsprechung in \mathbb{C}^n lautet

$$\langle \vec{x}, \vec{y} \rangle := \vec{y}^H \vec{x} = \sum_{i=1}^n x_i \bar{y}_i \quad (\vec{x}, \vec{y} \in \mathbb{C}^n),$$

so dass wir nun z. B. auch über Orthogonalität von Vektoren aus \mathbb{C}^n entscheiden können.

Anmerkung: Man kann auch andere Skalarprodukte auf \mathbb{K}^n definieren. Diese modellieren dann aber nicht unbedingt unsere klassischen Vorstellungen von Winkelmessung und Orthogonalität.

Prüfen Sie folgende Vektoren des \mathbb{R}^3 auf paarweise Orthogonalität:

$$\vec{a} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad \vec{c} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Bestimmen Sie das Skalarprodukt der Vektoren $\vec{x} = (1, i)^T$ und $\vec{y} = (1 + i, i)^T$ in \mathbb{C}^2 . Lassen sich diese beiden Vektoren mit unseren gewohnten Vorstellungen visualisieren?

Bestätigen Sie die in Definition 2.26 genannten Punkte für einen der Fälle $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.

Das Euklidische Skalarprodukt erzeugt die **Euklidische Norm**

$$\|\vec{x}\|_2 := \|\vec{x}\| := \sqrt{\sum_{i=1}^n |x_i|^2} \quad (\vec{x} \in \mathbb{R} \text{ bzw. } \vec{x} \in \mathbb{C}).$$

In \mathbb{R}^n kann man die Betragsstriche weglassen, d. h. $\|\vec{x}\| = \sqrt{\sum_{i=1}^n x_i^2}$.

Nach dem Satz des Pythagoras ist dies gerade die Länge des \vec{x} zugeordneten Vektorpfeils.

Machen Sie sich das am Beispiel des \mathbb{R}^2 anhand einer Skizze klar.

Berechnen Sie die Norm der Vektoren \vec{a} , \vec{b} und \vec{c} von Seite 201.

Anmerkung: In der Literatur findet man auch das Symbol $|\vec{x}|$ und die Bezeichnung 'Betrag' für die Euklidische Norm auf \mathbb{R}^n .

Orthogonalität, Skalarprodukt und Norm

Geometrische Interpretation des Skalarprodukts

In \mathbb{R}^2 oder \mathbb{R}^3 lässt sich mittels elementarer Geometrie zeigen, dass für den von zwei Vektoren \vec{x} und \vec{y} eingeschlossenen Winkel $\phi \in [0, \pi)$ gilt:

$$\cos \phi = \frac{\vec{x}^T \vec{y}}{\|\vec{x}\| \|\vec{y}\|}. \quad (2.7)$$

Anmerkung: Mittels

$$\cos \phi = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|}$$

lassen sich auch Winkel in beliebigen Vektorräumen mit Skalarprodukt definieren*, dies nutzt man praktisch aber selten.

Die Wohldefiniertheit wird dabei durch die allgemeingültige Cauchy-Schwarz-Ungleichung

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\| \quad (\mathbf{x}, \mathbf{y} \in V)$$

gesichert. Machen Sie sich klar, welche Bedingung es eigentlich zu sichern gilt.

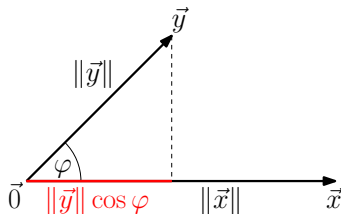
Orthogonalität, Skalarprodukt und Norm

Geometrische Interpretation des Skalarprodukts

Aus Gleichung (2.7) erhält man in \mathbb{R}^n desweiteren folgende Darstellung des Euklidischen Skalarprodukts:

$$\vec{x}^T \vec{y} = \|\vec{x}\| \|\vec{y}\| \cos \phi$$

Interpretiert man diese Formel an folgender Skizze,



so sieht man, dass das Skalarprodukt gleich der (vorzeichenbehafteten) Länge der orthogonalen Projektion von \vec{y} auf \vec{x} , multipliziert mit der Länge von \vec{x} ist.

Definition 2.28

Eine Basis $B = \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n\}$ des \mathbb{R}^n heißt **Orthonormalbasis** (ONB), wenn

$$\vec{b}_j^T \vec{b}_i = \begin{cases} 1, & \text{für } i = j, \\ 0, & \text{für } i \neq j. \end{cases}$$

Bei einer Orthonormalbasis stehen die Basisvektoren also paarweise aufeinander senkrecht und haben allesamt die Länge 1.

Ein prominentes Beispiel für eine Orthonormalbasis des \mathbb{R}^n ist die Standardbasis $\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n\}$.

Aus den Vektoren $\vec{a}, \vec{b}, \vec{c}$ auf S. 201 kann man durch Normieren eine ONB des \mathbb{R}^3 erhalten. Geben Sie diese an.

Orthogonalität, Skalarprodukt und Norm

Orthonormalbasen

Die Komponenten eines Vektors bezüglich einer Orthonormalbasis lassen sich besonders leicht über Orthogonalprojektionen berechnen:

Satz 2.29

Ist $B = \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_n\}$ eine Orthonormalbasis des \mathbb{R}^n , dann besitzt jeder Vektor $\vec{x} \in \mathbb{R}^n$ die eindeutige Darstellung

$$\vec{x} = \sum_{j=1}^n \left(\vec{b}_j^T \vec{x} \right) \vec{b}_j.$$

Machen Sie sich die Gültigkeit der Aussage für \mathbb{R}^2 anhand der Skizze auf S. 204 klar. Bestätigen Sie sie für den Fall der Standardbasis auf \mathbb{R}^2 auch rechnerisch.

Machen Sie sich klar, warum zwei von Null verschiedene orthogonale Vektoren immer linear unabhängig sind.

Orthogonalität, Skalarprodukt und Norm

Gram-Schmidtsches Orthogonalisierungsverfahren

Aus einem Satz Vektoren $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ kann man durch folgendes Verfahren eine Orthonormalbasis des Unterraumes $\text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ konstruieren.

$$\mathbf{b}_1 = \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|},$$

$$\tilde{\mathbf{b}}_2 = \mathbf{a}_2 - \langle \mathbf{a}_2, \mathbf{b}_1 \rangle \mathbf{b}_1,$$

$$\mathbf{b}_2 = \frac{\tilde{\mathbf{b}}_2}{\|\tilde{\mathbf{b}}_2\|},$$

$$\tilde{\mathbf{b}}_3 = \mathbf{a}_3 - \langle \mathbf{a}_3, \mathbf{b}_2 \rangle \mathbf{b}_2 - \langle \mathbf{a}_3, \mathbf{b}_1 \rangle \mathbf{b}_1,$$

$$\mathbf{b}_3 = \frac{\tilde{\mathbf{b}}_3}{\|\tilde{\mathbf{b}}_3\|},$$

⋮

$$\tilde{\mathbf{b}}_n = \mathbf{a}_n - \sum_{k=1}^{n-1} \langle \mathbf{a}_n, \mathbf{b}_k \rangle \mathbf{b}_k,$$

$$\mathbf{b}_n = \frac{\tilde{\mathbf{b}}_n}{\|\tilde{\mathbf{b}}_n\|}.$$

Im Fall dass die Vektoren $\mathbf{a}_1, \dots, \mathbf{a}_n$ nur einen Unterraum der Dimension $m < n$ aufspannen gilt $\mathbf{b}_{m+1} = \dots = \mathbf{b}_n = \mathbf{0}$.

Orthogonalität, Skalarprodukt und Norm

Exkurs: Skalarprodukte und Normen in der Quantenmechanik

In der Quantenmechanik beschreibt man den Zustand von Teilchen (in einer Raumdimension) mittels einer komplexwertigen **Wellen- oder Zustandsfunktion**

$$\phi : \mathbb{R} \rightarrow \mathbb{C}.$$

Diese Funktion ist selbst schwer interpretierbar, allerdings liefert

$$\int_a^b |\phi(x)|^2 dx$$

die Wahrscheinlichkeit, dass sich das Teilchen zwischen a und b aufhält.

Da sich das Teilchen an irgendeiner Stelle auf der Zahlengeraden befinden muss, ist es sinnvoll,

$$\int_{-\infty}^{\infty} |\phi(x)|^2 dx = 1$$

zu fordern.

Orthogonalität, Skalarprodukt und Norm

Exkurs: Skalarprodukte und Normen in der Quantenmechanik

Als zugrundeliegende Räume verwendet man daher sogenannte L^2 -Räume*. Grob gesprochen sind das Räume von Funktionen mit

$$\|f\|_2 := \left(\int_{-\infty}^{\infty} |f(x)|^2 dx \right)^{\frac{1}{2}} < \infty.$$

Die so definierte L^2 -Norm lässt sich mit folgendem (komplexen) Skalarprodukt erzeugen:

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x) \overline{g(x)} dx.$$

Eine reelle Version verwendet man häufig bei der theoretischen und numerischen Behandlung partieller Differentialgleichungen.

* Das 'L' steht für den hier benötigten erweiterten Integralbegriff – das 'Lebesgue-Integral' (Henri Lebesgue, 1875-1941).

Mit Hilfe einer Norm lässt sich schließlich ein Abstandsbegriff (**Metrik**) in V einführen. Der Abstand zweier Vektoren $\mathbf{x}, \mathbf{y} \in V$ wird durch

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| \quad (2.8)$$

erklärt.

Für alle $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ gelten:

- $d(\mathbf{x}, \mathbf{x}) = 0$, $d(\mathbf{x}, \mathbf{y}) > 0$, falls $\mathbf{x} \neq \mathbf{y}$ (Definitheit),
- $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ (Symmetrie),
- $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$ (Dreiecksungleichung),
- $d(\mathbf{x} + \mathbf{z}, \mathbf{y} + \mathbf{z}) = d(\mathbf{x}, \mathbf{y})$ (Translationsinvarianz).

Orthogonalität, Skalarprodukt und Norm

Metrik: Geometrische Interpretation

Setzt man wieder die Euklid-Norm auf \mathbb{R}^n an, so ergibt sich aus (2.8) die gewohnte Abstandsformel für zwei Punkte $\vec{x}, \vec{y} \in \mathbb{R}^n$:

$$d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}.$$

Mit welchem klassischen Satz deckt sich diese Formel?

Berechnen Sie den Abstand der Punkte $(5, 2)^T$ und $(1, -1)^T$ in \mathbb{R}^2 .

① Grundlagen

② Lineare Algebra

2.1 Vektorräume

2.2 Matrizen und lineare Abbildungen

2.3 Lineare Gleichungssysteme

2.4 Determinanten

2.5 Invertierbare Matrizen

2.6 Orthogonalität, Skalarprodukt und Norm

2.7 Kreuz- und Spatprodukt

2.8 Elemente der analytischen Geometrie

2.9 Orthogonale Abbildungen

2.10 Eigenwerte und Eigenvektoren

2.11 Singulärwertzerlegung

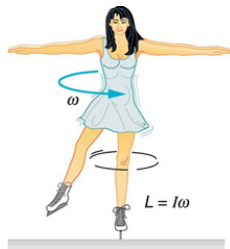
Kreuz- und Spatprodukt

Vektorprodukte

Bei vielen physikalischen Anwendungen benötigt man eine weitere Operation, das sogenannte **Vektor- oder Kreuzprodukt**.

Beispiele sind

- das **Drehmoment** $\vec{M} = \vec{r} \times \vec{F}$ zum Abstandsvektor \vec{r} und zur Kraft \vec{F} ,

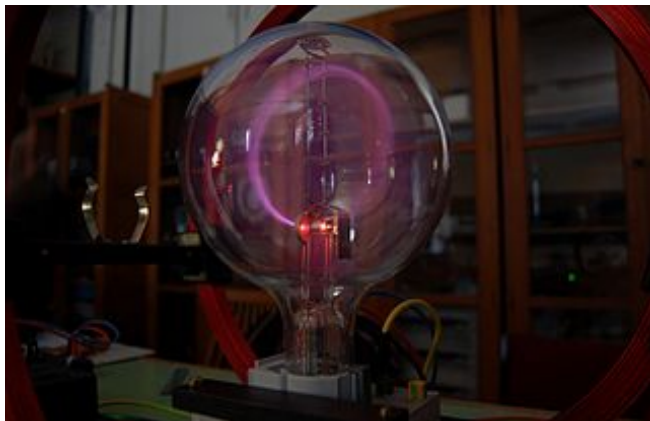


Verwandt: **Drehimpuls** $\vec{L} = \vec{r} \times \vec{p}$ bei Impuls $m\vec{v}$. Bei Drehbewegung mit Winkelgeschwindigkeit ω ist $|\vec{L}| = mr^2\omega$.

Kreuz- und Spatprodukt

Vektorprodukte

- die **Lorentzkraft** $\vec{F}_L = q(\vec{v} \times \vec{B})$, welche auf ein Teilchen mit Ladung q und Geschwindigkeit \vec{v} im Magnetfeld \vec{B} wirkt;



Fadenstrahlrohr

Kreuz- und Spatprodukt

Vektorprodukte

- die **Corioliskraft** $\vec{F}_C = 2m(\vec{v} \times \vec{\omega})$, welche auf einen Körper der Masse m wirkt, welcher sich mit Geschwindigkeit \vec{v} relativ zu einem mit Winkelgeschwindigkeit $\vec{\omega}$ rotierenden Bezugssystem bewegt;

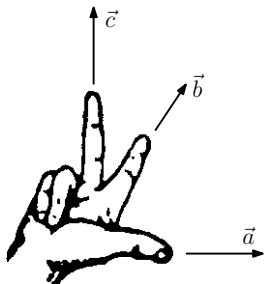


Kreuz- und Spatprodukt

Vektorprodukte

Wir gehen bei der Definition des Kreuzprodukts von den gewünschten geometrischen Eigenschaften aus und leiten dann die Rechenregeln sowie die Formel zur Berechnung her. Sämtliche Vektoren in diesem Kapitel sind als Elemente des \mathbb{R}^3 aufzufassen.

Drei linear unabhängige Vektoren $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$ bilden (in dieser Reihenfolge) ein **Rechtssystem**, wenn sich — von der Spitze von \vec{c} aus gesehen — \vec{a} durch Drehung um einen Winkel $\phi \in [0, \pi)$ im Gegenuhrzeigersinn in die gleiche Richtung wie \vec{b} bringen lässt.



Ob ein Rechtssystem vorliegt, kann mit der Rechten-Hand-Regel (Korkenzieherregel) entschieden werden:

Analog kann man ein **Linkssystem** definieren.

Kreuz- und Spatprodukt

Geometrische Definition des Kreuzprodukts

Definition 2.30

Seien $\vec{a}, \vec{b} \in \mathbb{R}^3$ linear unabhängig. Der Vektor

$$\vec{c} = \vec{a} \times \vec{b} \in \mathbb{R}^3$$

heißt **Kreuzprodukt** (auch **Vektorprodukt** oder **äußeres Produkt**) von \vec{a} und \vec{b} , wenn

(1) $\vec{c} \perp \vec{a}$ und $\vec{c} \perp \vec{b}$ (d.h. $\vec{c} \perp \text{span}\{\vec{a}, \vec{b}\}$),

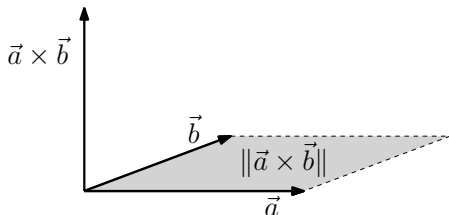
(2) $\|\vec{c}\| = \|\vec{a}\| \|\vec{b}\| |\sin \angle(\vec{a}, \vec{b})|$,

(3) $\vec{a}, \vec{b}, \vec{c}$ bilden in dieser Reihenfolge ein Rechtssystem.

Sind \vec{a} und \vec{b} linear abhängig (was die Fälle $\vec{a} = \vec{0}$ oder $\vec{b} = \vec{0}$ mit einschließt), so definieren wir $\vec{a} \times \vec{b} := \vec{0}$.

Kreuz- und Spatprodukt

Interpretation von Definition 2.30



- Punkt (1) legt die Gerade fest, auf der $\vec{a} \times \vec{b}$ liegt. Die Richtung von $\vec{a} \times \vec{b}$ ist damit bis auf das Vorzeichen eindeutig bestimmt.
- Punkt (2) legt die Länge von $\vec{a} \times \vec{b}$ fest. Diese stimmt mit der Fläche des von \vec{a} und \vec{b} aufgespannten Parallelogramms überein.
- Punkt (3) liefert die Entscheidung, welcher der verbliebenen zwei Vektoren (Richtungen) zu verwenden ist.

Kreuz- und Spatprodukt

Rechenregeln

Alein aufgrund der geometrischen Definition ergeben sich folgende Rechenregeln für Kreuzprodukte:

Satz 2.31

Seien $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$ und $\lambda \in \mathbb{R}$. Dann gelten:

- $\vec{b} \times \vec{a} = -(\vec{a} \times \vec{b})$,
- $(\vec{a} + \vec{b}) \times \vec{c} = \vec{a} \times \vec{c} + \vec{b} \times \vec{c}$,
- $\lambda(\vec{a} \times \vec{b}) = (\lambda\vec{a}) \times \vec{b} = \vec{a} \times (\lambda\vec{b})$.

Illustrieren Sie zumindest den ersten und dritten Punkt anhand geeigneter Skizzen.

Kreuz- und Spatprodukt

Explizite Berechnung des Kreuzprodukts

Satz 2.32

Für $\vec{a}, \vec{b} \in \mathbb{R}^3$ gilt

$$\vec{a} \times \vec{b} = \begin{bmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{bmatrix}. \quad (2.9)$$

Beweisidee:

- Für die Einheitsvektoren des \mathbb{R}^3 gilt offenbar $\vec{e}_1 \times \vec{e}_2 = \vec{e}_3$, $\vec{e}_2 \times \vec{e}_3 = \vec{e}_1$ und $\vec{e}_3 \times \vec{e}_1 = \vec{e}_2$.
- Schreibe $\vec{a} = a_1 \vec{e}_1 + a_2 \vec{e}_2 + a_3 \vec{e}_3$ bzw. $\vec{b} = b_1 \vec{e}_1 + b_2 \vec{e}_2 + b_3 \vec{e}_3$ und wende Satz 2.31 an.

Man führe die Rechnung zum zweitgenannten Punkt aus.

Kreuz- und Spatprodukt

Tipp zum praktischen Rechnen

Formel (2.9) merkt man sich am besten mit Hilfe der formalen 3×3 -Determinanten

$$\vec{a} \times \vec{b} = \det \left(\begin{bmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} \right),$$

die man mit der Regel von Sarrus auswertet.

Berechnen Sie auf diese Weise das Kreuzprodukt

$$\begin{bmatrix} 1 \\ 2 \\ -4 \end{bmatrix} \times \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}.$$

Kreuz- und Spatprodukt

Spatprodukt

Mitunter ist noch eine Kombination von Kreuz- und Skalarprodukt in Gebrauch. Unter dem **Spatprodukt** (auch **gemischten Produkt**) dreier Vektoren $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$ versteht man die reelle Zahl

$$[\vec{a}, \vec{b}, \vec{c}] := \vec{a}^T (\vec{b} \times \vec{c}).$$

Eine einfache Rechnung zeigt, dass sich das Spatprodukt als gewöhnliche Determinante interpretieren lässt:

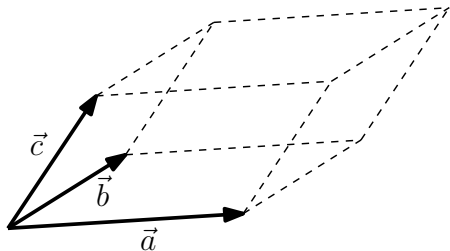
$$[\vec{a}, \vec{b}, \vec{c}] = \det \left(\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \right).$$

Beim Spatprodukt handelt es sich also eigentlich um nichts Neues.

Kreuz- und Spatprodukt

Geometrische Interpretation

Wie im Abschnitt Determinanten erörtert lässt sich damit der Betrag des Spatprodukts als Volumen des von \vec{a} , \vec{b} und \vec{c} aufgespannten Parallelepipeds interpretieren.



Der Begriff **Spat** steht synonym für Parallelepipeds; dies begründet die Namensgebung.

Satz 2.33

Seien $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$. Dann gelten:

- Sind $\vec{a}, \vec{b}, \vec{c}$ linear abhängig, so ist $[\vec{a}, \vec{b}, \vec{c}] = 0$.
- Sind $\vec{a}, \vec{b}, \vec{c}$ linear unabhängig, und bilden sie in dieser Reihenfolge ein Rechtssystem, so ist $[\vec{a}, \vec{b}, \vec{c}] > 0$.
- Sind $\vec{a}, \vec{b}, \vec{c}$ linear unabhängig, und bilden sie in dieser Reihenfolge ein Linkssystem, so ist $[\vec{a}, \vec{b}, \vec{c}] < 0$.
- $[\vec{b}, \vec{a}, \vec{c}] = [\vec{c}, \vec{b}, \vec{a}] = [\vec{a}, \vec{c}, \vec{b}] = -[\vec{a}, \vec{b}, \vec{c}]$.
- $[\vec{a}, \vec{b}, \vec{c}] = [\vec{c}, \vec{a}, \vec{b}] = [\vec{b}, \vec{c}, \vec{a}]$.

① Grundlagen

② Lineare Algebra

2.1 Vektorräume

2.2 Matrizen und lineare Abbildungen

2.3 Lineare Gleichungssysteme

2.4 Determinanten

2.5 Invertierbare Matrizen

2.6 Orthogonalität, Skalarprodukt und Norm

2.7 Kreuz- und Spatprodukt

2.8 Elemente der analytischen Geometrie

2.9 Orthogonale Abbildungen

2.10 Eigenwerte und Eigenvektoren

2.11 Singulärwertzerlegung

Elemente der analytischen Geometrie

Geometrische Interpretation von Vektoren im \mathbb{R}^n

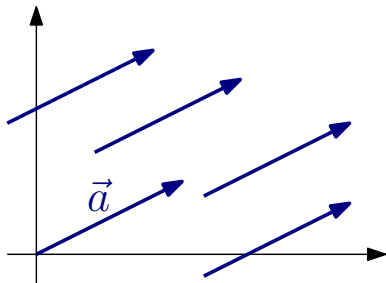
Vektoren aus \mathbb{R}^2 bzw. \mathbb{R}^3 lassen sich anschaulich auf verschiedene Weisen interpretieren:

- als Punkt der Anschauungsebene bzw. des Anschauungsraums,
- als Pfeil vom Koordinatenursprung zu eben diesem Punkt ('Ortsvektor'), also einem Objekt mit Richtung und Länge.

Manchmal ist es zudem anschaulicher, die so entstandenen Pfeile an bestimmte Stellen in der Ebene bzw. im Raum zu verschieben (z. B. wenn eine Kraft \vec{F} an einem bestimmten Punkt 'angreift').

Elemente der analytischen Geometrie

Bild zur Pfeilinterpretation



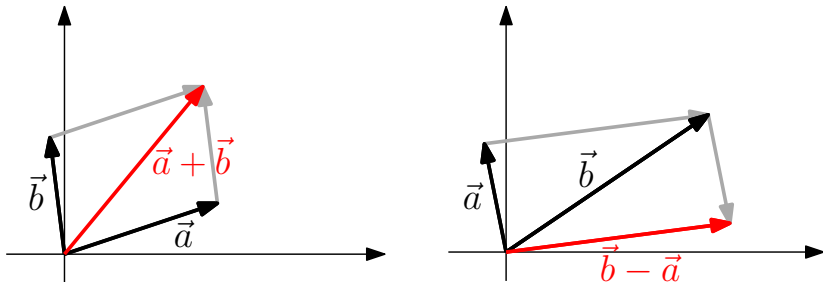
verschiedene Repräsentanten
des Vektors \vec{a} .

Vektoren des \mathbb{R}^n lassen sich also mit sogenannten Pfeilklassen assoziieren. Jede dieser Pfeilklassen korrespondiert wiederum mit einer speziellen Parallelverschiebung.

Elemente der analytischen Geometrie

Bild zur Pfeilinterpretation

Wie zweckmäßig solche Verschiebungen des Ortsvektors sind, wird an der geometrischen Interpretation der Addition und Subtraktion von Vektoren in deutlich:



In welchem Zusammenhang haben Sie diese Skizzen schon einmal gesehen? Was ist der Grund für diese Analogie?

In der analytischen Geometrie werden Punkte und Vektoren mitunter auch als verschiedene Objekte betrachtet.

Punkte werden dann meist durch Großbuchstaben gekennzeichnet (z. B. O für den Ursprung) und Vektoren häufig in der Form \overrightarrow{AB} über ihre Anfangs- und Endpunkte. Es gilt dann:

- Jeder Punkt P korrespondiert mit seinem Ortsvektor $\vec{p} = \overrightarrow{OP}$.
- Sind $\vec{a} = \overrightarrow{OA}$ und $\vec{b} = \overrightarrow{OB}$ die Ortsvektoren zu zwei Punkten A und B , so gilt

$$\overrightarrow{AB} = \vec{b} - \vec{a}$$

Wir werden bei unserer bisherigen Notation bleiben und \vec{p} statt \overrightarrow{OP} bzw. $\vec{b} - \vec{a}$ statt \overrightarrow{AB} schreiben.

Elemente der analytischen Geometrie

Punkte, Geraden und Ebenen im Raum

In diesem Abschnitt betrachten wir nur noch Objekte im \mathbb{R}^3 . Einige Ergebnisse haben ihre Entsprechungen im \mathbb{R}^2 , die klar ersichtlich sind.

Konkret wollen wir uns mit den Lagebeziehungen von Punkt, Gerade und Ebene befassen.

Der Stoff gehört zum Standardrepertoire an Gymnasien und wird dort auch in der nötigen Tiefe behandelt. Daher sollen diese Inhalte nur überblicksartig dargestellt werden.

Elemente der analytischen Geometrie

Geraden

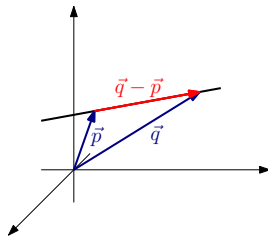
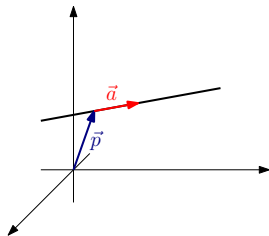
Eine Gerade, die durch den Punkt \vec{p} und parallel zur Richtung $\vec{a} \neq \vec{0}$ verläuft, besitzt die Parameterdarstellung

$$\vec{x} = \vec{p} + \lambda \vec{a}, \quad \lambda \in \mathbb{R}.$$

Diese Form wird auch **Punkt-Richtungs-Form** genannt.

Alternativ ist eine Gerade durch zwei verschiedene Punkte \vec{p} und \vec{q} eindeutig festgelegt. Man erhält die Parameterdarstellung über die **Zwei-Punkte-Form**

$$\vec{x} = \vec{p} + \lambda (\vec{q} - \vec{p}), \quad \lambda \in \mathbb{R}.$$



Analog ist eine Ebene durch einen Punkt \vec{p} und zwei linear unabhängige Richtungsvektoren \vec{a} und \vec{b} eindeutig festgelegt. Sie besitzt also die Parameterform

$$\vec{x} = \vec{p} + \lambda \vec{a} + \mu \vec{b}, \quad \lambda, \mu \in \mathbb{R}.$$

Diese Form wird auch **Punkt-Richtungs-Form der Ebene** genannt.

Alternativ kann man zur eindeutigen Festlegung drei Punkte \vec{p} , \vec{q} und \vec{r} verwenden. Man erhält die Parameterdarstellung über die **Drei-Punkte-Form**

$$\vec{x} = \vec{p} + \lambda (\vec{q} - \vec{p}) + \mu (\vec{r} - \vec{p}), \quad \lambda, \mu \in \mathbb{R}.$$

Elemente der analytischen Geometrie

Ebenen

Jeder Vektor $\vec{n} \neq 0$, der senkrecht auf der Ebene steht, heißt **Normalenvektor** der Ebene.

Auch durch Vorgabe eines Punktes \vec{p} und eines Normalektors \vec{n} ist eine Ebene eindeutig festgelegt. Es entsteht die **Normalenform**

$$\begin{aligned}\vec{n}^T(\vec{x} - \vec{p}) &= 0 \quad \text{bzw.} & (2.10) \\ n_1x_1 + n_2x_2 + n_3x_3 &= c \quad (\text{mit } c = \vec{n}^T\vec{p}).\end{aligned}$$

Ist \vec{n} auf Länge Eins normiert, so ist der Abstand des Punktes $\vec{x} = (x_1, x_2, x_3)$ von der Ebene gegeben durch (**Hessesche Normalform**)

$$d = |n_1x_1 + n_2x_2 + n_3x_3 - c|.$$

Wie kann man einen Normalenvektor zur Ebene $\vec{x} = \vec{p} + \lambda\vec{a} + \mu\vec{b}$ berechnen?

Interpretieren Sie Formel (2.10) im Kontext orthogonaler Projektionen auf den Normalenvektor.

Elemente der analytischen Geometrie

Graphische Darstellung

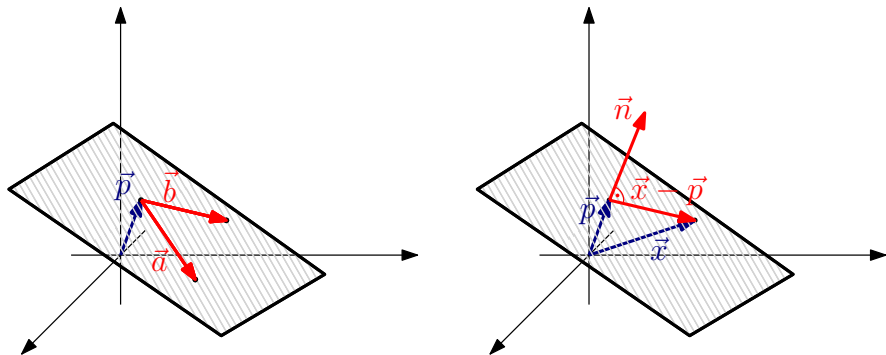


Illustration von Punkt-Richtungsform (links) und Normalenform (rechts).

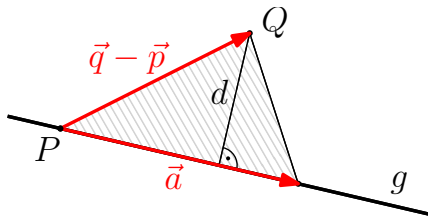
Elemente der analytischen Geometrie

Lagebeziehungen Punkt-Gerade

Ob ein Punkt \vec{q} auf einer Geraden $g : \vec{x} = \vec{p} + \lambda \vec{a}$ liegt oder nicht, kann man direkt mit der Parameterdarstellung prüfen. Alternativ berechnet man einfach den Abstand

$$d(\vec{q}, g) = \frac{\|\vec{a} \times (\vec{q} - \vec{p})\|}{\|\vec{a}\|}. \quad (2.11)$$

Bestätigen Sie die Formel. Berechnen Sie dazu den Flächeninhalt des von \vec{a} und $\vec{q} - \vec{p}$ aufgespannten Dreiecks auf zwei verschiedene Weisen.



Für zwei Geraden

$$g_1 : \vec{x} = \vec{p} + \lambda \vec{a} \quad \text{und} \quad g_2 : \vec{x} = \vec{q} + \mu \vec{b}$$

im \mathbb{R}^3 liegt genau eine der drei folgenden Situationen vor:

- sie sind **parallel** (Spezialfall der Gleichheit einbezogen),
- sie schneiden sich in genau einem Punkt,
- sie sind **windschief**, d.h. sie sind weder parallel noch schneiden sie sich.

Parallelität lässt sich am leichtesten erkennen: In diesem Fall sind die beiden Vektoren \vec{a} und \vec{b} linear abhängig, unterscheiden sich also nur um einen skalaren Faktor.

Im Falle der Parallelität kann man Formel (2.11) zur Abstandsbestimmung nutzen (warum?) und erhält

$$d(g_1, g_2) = \frac{\|\vec{a} \times (\vec{q} - \vec{p})\|}{\|\vec{a}\|} = \frac{\|\vec{b} \times (\vec{p} - \vec{q})\|}{\|\vec{b}\|}.$$

Für die restlichen Fälle erzeugt man durch Gleichsetzen das lineare Gleichungssystem

$$\vec{p} + \lambda \vec{a} = \vec{q} + \mu \vec{b} \quad (2.12)$$

(drei Gleichungen für die zwei Unbekannten λ und μ).

Hat (2.12) genau eine Lösung (λ^*, μ^*) , so schneiden sich die Geraden im Punkt $\vec{p} + \lambda^* \vec{a}$ (identisch mit $\vec{q} + \mu^* \vec{b}$). Der Schnittwinkel ist gegeben durch

$$\cos \phi = \frac{\vec{a}^T \vec{b}}{\|\vec{a}\| \|\vec{b}\|}.$$

Hat (2.12) unendlich viele Lösungen, so sind die Geraden g_1 und g_2 identisch. (Das kann man aber im Rahmen des Tests auf Parallelität bereits entscheiden.)

Hat (2.12) keine Lösung, so sind die Geraden g_1 und g_2 windschief oder aber parallel und verschieden (s. o.).

Liegt der windschiefe Fall vor, erhält man den Abstand der Geraden mittels

$$d(g_1, g_2) = \frac{|[\vec{a}, \vec{b}, \vec{q} - \vec{p}]|}{\|\vec{a} \times \vec{b}\|}. \quad (2.13)$$

Formel (2.13) lässt sich in Analogie zu (2.11) bestätigen, indem man das Volumen eines geeigneten Körpers auf zwei verschiedene Weisen berechnet. Um welchen Körper handelt es sich?

Ob ein Punkt \vec{q} zur Ebene

$$E : \vec{n}^T (\vec{x} - \vec{p}) = 0$$

gehört, lässt sich durch Einsetzen prüfen. (Gilt $\vec{n}^T (\vec{q} - \vec{p}) = 0$?)

Allgemein berechnet sich der Abstand von \vec{q} zu E gemäß

$$d(\vec{q}, E) = \frac{|\vec{n}^T (\vec{q} - \vec{p})|}{\|\vec{n}\|}. \quad (2.14)$$

Welcher Ansatz liegt Formel (2.14) zugrunde?

Eine Gerade und eine Ebene sind im \mathbb{R}^3 entweder

- parallel (beinhaltet den Fall, dass die Gerade in der Ebene liegt),
- oder sie schneiden sich in genau einem Punkt.

Die Gerade $g : \vec{x} = \vec{q} + \lambda \vec{a}$ und die Ebene $E : \vec{n}^T (\vec{x} - \vec{p}) = 0$ sind genau dann parallel, wenn \vec{a} und \vec{n} orthogonal sind, d. h. wenn

$$\vec{n}^T \vec{a} = 0.$$

In diesem Fall ist der Abstand Gerade-Ebene nach (2.14) gegeben durch

$$d = \frac{|\vec{n}^T (\vec{q} - \vec{p})|}{\|\vec{n}\|}.$$

Sind die Ebene und die Gerade nicht parallel ($\vec{n}^T \vec{a} \neq 0$), dann ist

$$\vec{q} + \frac{\vec{n}^T (\vec{p} - \vec{q})}{\vec{n}^T \vec{a}} \vec{a}$$

ihr Schnittpunkt. Der Schnittwinkel ϕ erfüllt

$$\sin \phi = \frac{|\vec{n}^T \vec{a}|}{\|\vec{n}\| \|\vec{a}\|}.$$

Leiten Sie beide Formeln her.

Zwei Ebenen im \mathbb{R}^3 sind

- entweder parallel (Spezialfall: identisch)
- oder schneiden sich entlang einer Geraden.

Zwei Ebenen

$$E_1 : \vec{n}_1^T (\vec{x} - \vec{p}_1) = 0 \quad \text{und} \quad E_2 : \vec{n}_2^T (\vec{x} - \vec{p}_2) = 0,$$

sind genau dann parallel, wenn \vec{n}_1 und \vec{n}_2 linear abhängig (also bis auf Vielfache gleich) sind. In diesem Fall ist der Abstand der Ebenen

$$d(E_1, E_2) = \frac{|\vec{n}_1^T (\vec{p}_2 - \vec{p}_1)|}{\|\vec{n}_1\|} = \frac{|\vec{n}_2^T (\vec{p}_1 - \vec{p}_2)|}{\|\vec{n}_2\|}.$$

Schneiden sich die Ebenen entlang einer Geraden $\vec{x} = \vec{p} + \lambda \vec{a}$ (äquivalent zu $\vec{n}_1 \times \vec{n}_2 \neq \vec{0}$), dann ist

$$\vec{a} = \vec{n}_1 \times \vec{n}_2,$$

und \vec{p} ist (jede) Lösung des linearen Gleichungssystems

$$\vec{n}_1^T (\vec{p} - \vec{p}_1) = 0,$$

$$\vec{n}_2^T (\vec{p} - \vec{p}_2) = 0$$

(zwei Gleichungen für drei Unbekannte, nämlich die drei Komponenten von \vec{p}). Der Schnittwinkel ϕ erfüllt

$$\cos \phi = \frac{\vec{n}_1^T \vec{n}_2}{\|\vec{n}_1\| \|\vec{n}_2\|}.$$

① Grundlagen

② Lineare Algebra

2.1 Vektorräume

2.2 Matrizen und lineare Abbildungen

2.3 Lineare Gleichungssysteme

2.4 Determinanten

2.5 Invertierbare Matrizen

2.6 Orthogonalität, Skalarprodukt und Norm

2.7 Kreuz- und Spatprodukt

2.8 Elemente der analytischen Geometrie

2.9 Orthogonale Abbildungen

2.10 Eigenwerte und Eigenvektoren

2.11 Singulärwertzerlegung

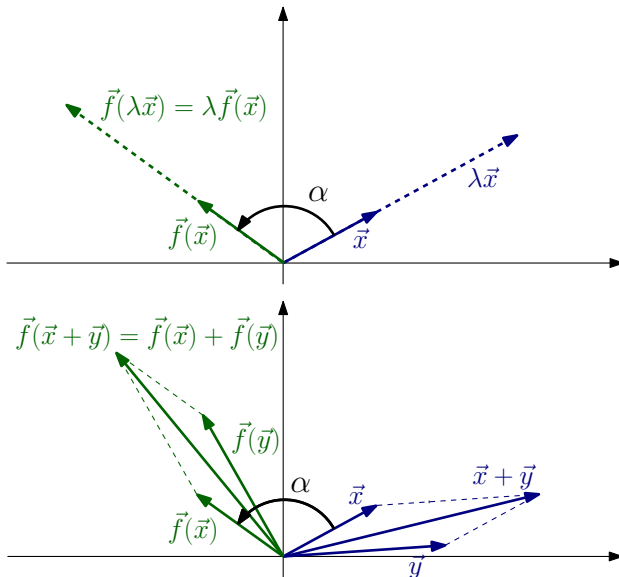
Orthogonale Abbildungen

- Eine große Klasse linearer Abbildungen sind die **orthogonalen Abbildungen**.
- Sie spielen insbesondere bei geometrischen Transformationen eine Rolle und zeichnen sich durch die Eigenschaft aus, **Längen und Winkel unverändert zu lassen**.
- So lassen sich zum Beispiel Drehungen und Spiegelungen mit orthogonalen Abbildungen mathematisch beschreiben.

Drehungen des \mathbb{R}^2 um einen Winkel α und Mittelpunkt in $\vec{0}$ sind zunächst lineare Abbildungen, wie an folgenden Skizzen deutlich wird:

Orthogonale Abbildungen

Drehung als lineare Abbildung



Orthogonale Abbildungen

Drehung und Spiegelung als lineare Abbildung

Zeichnen Sie eine analoge Skizze für Spiegelung des \mathbb{R}^2 an einer Geraden durch $\vec{0}$.

Natürlich lässt sich dieses geometrische Argument analog auf Drehungen und auf Spiegelungen an Ebenen im \mathbb{R}^3 anwenden.

Sowohl Drehungen um den Ursprung als auch Spiegelungen an einer Geraden (Ebene) durch den Ursprung können also als Matrix-Vektor-Multiplikationen beschrieben werden.

Wir begeben uns auf die Suche nach den Abbildungsmatrizen zu diesen Abbildungen.

Orthogonale Abbildungen

Drehungen um den Ursprung im \mathbb{R}^2

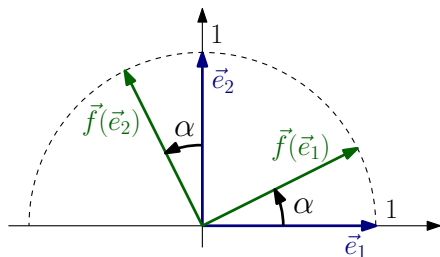
Die Spalten der gesuchten Abbildungsmatrix sind gerade die Bilder der Einheitsvektoren unter der Drehung.

Es gilt

$$\vec{f}(\vec{e}_1) = \begin{bmatrix} \cos \alpha \\ \sin \alpha \end{bmatrix}$$

und

$$\vec{f}(\vec{e}_2) = \begin{bmatrix} -\sin \alpha \\ \cos \alpha \end{bmatrix}.$$



Damit ist die gesuchte Abbildungsmatrix

$$D_\alpha = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}.$$

Orthogonale Abbildungen

Drehungen um den Ursprung im \mathbb{R}^2

Die Matrix D_α ist invertierbar, denn

$$\det D_\alpha = \cos^2 \alpha + \sin^2 \alpha = 1.$$

Ihre Inverse realisiert gerade die Drehung um $-\alpha$, d. h.

$$D_\alpha^{-1} = D_{-\alpha} = \begin{bmatrix} \cos(-\alpha) & -\sin(-\alpha) \\ \sin(-\alpha) & \cos(-\alpha) \end{bmatrix} = \begin{bmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{bmatrix}$$

Für D_α gilt also die bemerkenswerte Beziehung

$$D_\alpha^{-1} = D_\alpha^T.$$

Berechnen Sie Drehmatrix für eine Drehung um den Ursprung mit $\alpha = 30^\circ$.
Geben Sie das Bild des Vektors $[2, 3]^T$ an.

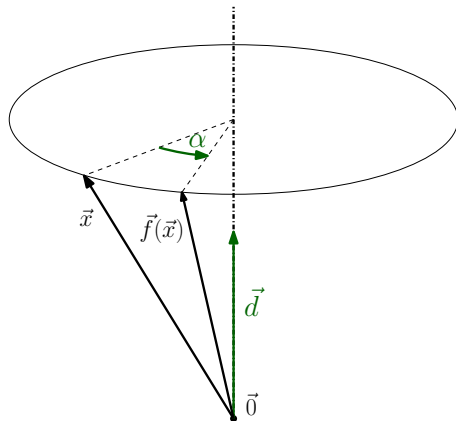
Wiederholen Sie bei Bedarf die Begriffe Inverse, Invertierbarkeit und Determinante aus Abschnitte 2.4 und 2.5.

Orthogonale Abbildungen

Drehungen im Raum

Drehungen im \mathbb{R}^3 werden durch eine Drehachse (die den Ursprung enthält) und einen Drehwinkel α festgelegt.

Die Drehachse wird dabei durch einen Vektor \vec{d} festgelegt, der in die positive Achsenrichtung zeigt.



Orthogonale Abbildungen

Drehungen im Raum

Besonders einfach wird die Angabe der Drehmatrizen, wenn man die Einheitsvektoren (also die Koordinatenachsen) als Drehachsen verwendet:

$$D_{x,\alpha} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix}, \quad D_{y,\alpha} = \begin{bmatrix} \cos \alpha & 0 & \sin \alpha \\ 0 & 1 & 0 \\ -\sin \alpha & 0 & \cos \alpha \end{bmatrix},$$

$$D_{z,\alpha} = \begin{bmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Jede beliebige Drehung im \mathbb{R}^3 lässt sich als Komposition von Drehungen um die Koordinatenachsen schreiben.

Physiker sprechen daher häufig von drei möglichen Freiheitsgraden der Rotation.

Machen Sie sich an einem der obigen Beispiele klar, dass die angegebene Matrix die gewünschte Transformation realisiert. Berechnen Sie dazu das Produkt $D\vec{x}$ für die gewählte Drehmatrix D .

Orthogonale Abbildungen

Exkurs: allgemeine Drehmatrix

Natürlich kann man auch im allgemeinen Fall die Drehmatrix angeben. Bei vorgegebenem Achsvektor \vec{d} (mit $\|\vec{d}\| = 1$) und Winkel α lautet diese

$$\begin{bmatrix} \cos \alpha + (1 - \cos \alpha)d_1^2 & (1 - \cos \alpha)d_1d_2 - d_3 \sin \alpha & (1 - \cos \alpha)d_1d_3 + d_2 \sin \alpha \\ (1 - \cos \alpha)d_1d_2 + d_3 \sin \alpha & \cos \alpha + (1 - \cos \alpha)d_2^2 & (1 - \cos \alpha)d_2d_3 - d_1 \sin \alpha \\ (1 - \cos \alpha)d_1d_3 - d_2 \sin \alpha & (1 - \cos \alpha)d_2d_3 + d_1 \sin \alpha & \cos \alpha + (1 - \cos \alpha)d_3^2 \end{bmatrix}$$

Eine explizite Abbildungsvorschrift ist gegeben durch

$$\vec{f}(\vec{x}) = \cos \alpha \vec{x} + (1 - \cos \alpha)(\vec{x}^T \vec{d}) \vec{d} + \sin \alpha(\vec{d} \times \vec{x}).$$

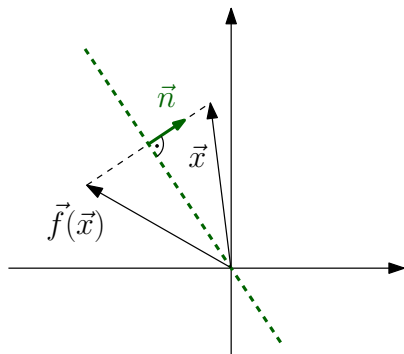
Dies alles schaut man bei Bedarf aber am besten in der Literatur nach.

Orthogonale Abbildungen

Spiegelungen in der Ebene

Wir betrachten zunächst Spiegelungen im \mathbb{R}^2 an einer Geraden durch $\vec{0}$ senkrecht zum Vektor \vec{n} . Dabei sei \vec{n} auf Länge 1 normiert ($\|\vec{n}\| = 1$).

Wir lesen das Spiegelbild $\vec{f}(x)$ von \vec{x} aus folgender Skizze ab:



$$\vec{f}(x) = \vec{x} - 2(\vec{n}^T \vec{x})\vec{n}$$

Beachten Sie dabei, dass wegen der Normierung von \vec{n} die Länge der Projektion von \vec{x} auf \vec{n} gerade $\vec{n}^T \vec{x}$ ist.

Orthogonale Abbildungen

Spiegelungen in der Ebene

Es gilt also

$$\begin{aligned}\vec{f}(x) &= \vec{x} - 2(\vec{n}^T \vec{x})\vec{n} = \vec{x} - 2\vec{n}(\vec{n}^T \vec{x}) \\ &= \vec{x} - 2(\vec{n}\vec{n}^T)\vec{x} = (I - 2\vec{n}\vec{n}^T)\vec{x},\end{aligned}$$

d. h. die Spiegelung wird durch Multiplikation mit der Matrix

$$S_{\vec{n}} = I - 2\vec{n}\vec{n}^T = \begin{bmatrix} 1 - 2n_1^2 & -2n_1n_2 \\ -2n_1n_2 & 1 - 2n_2^2 \end{bmatrix} = \begin{bmatrix} n_2^2 - n_1^2 & -2n_1n_2 \\ -2n_1n_2 & n_1^2 - n_2^2 \end{bmatrix}$$

beschrieben (beachte $n_1^2 + n_2^2 = 1$).

Geben Sie die Spiegelungsmatrix für die Spiegelung an der Geraden $x_2 = -x_1$ an. Multiplizieren Sie diese Matrix mit einem Vektor \vec{x} . Kann man das Ergebnis auch rein geometrisch begründen?

Orthogonale Abbildungen

Spiegelungen im Raum

Bei Spiegelungen im \mathbb{R}^3 verwendet man statt der Spiegelachse eine Spiegelebene durch den Ursprung, welche ganz analog durch den Normalenvektor \vec{n} festgelegt ist ($\|\vec{n}\| = 1$).

Die Spiegelungsmatrix besitzt jetzt drei Zeilen und Spalten, allerdings die gleiche Struktur:

$$S_{\vec{n}} = I - 2\vec{n}\vec{n}^T.$$

Zeichnen Sie eine geeignete Skizze, in welcher die Analogie sichtbar wird. Wie lautet die Matrix $S_{\vec{n}}$ in ausgeschriebener Form?

Orthogonale Abbildungen

Allgemein

Drehungen und Spiegelungen gehören zur Klasse der **orthogonalen linearen Abbildungen**, die sich auch für höhere Dimensionen erklären lassen:

Definition 2.34

Eine Matrix $U \in \mathbb{R}^{n \times n}$ heißt **orthogonal**, wenn sie das Innenprodukt nicht verändert, d. h. wenn

$$(U\mathbf{y})^T(U\mathbf{x}) = \mathbf{y}^T\mathbf{x} \quad \text{für alle } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

Insbesondere ist eine orthogonale Matrix

- **längenerhaltend**, d.h. $\|U\mathbf{x}\| = \|\mathbf{x}\|$ für alle \mathbf{x} und
- **winkeltreu**, d.h. $\angle(U\mathbf{x}, U\mathbf{y}) = \angle(\mathbf{x}, \mathbf{y})$ für alle \mathbf{x}, \mathbf{y} .

Satz 2.35 (Charakterisierung orthogonaler Matrizen)

Sei $U \in \mathbb{R}^{n \times n}$. Die folgenden Aussagen sind äquivalent:

- U ist orthogonal.
- Die Spalten (Zeilen) von U bilden eine Orthonormalbasis des \mathbb{R}^n .
- U ist invertierbar mit $U^{-1} = U^T$.

Machen Sie sich klar, warum aus Punkt 1 Punkt 2 und daraus wiederum Punkt 3 folgt.

Bestätigen Sie mit Punkt 3, dass die 2D-Spiegelungsmatrix $S = I - 2\vec{n}\vec{n}^T$ orthogonal ist.

Orthogonale Abbildungen

Exkurs: Unitäre Abbildungen

Die Entsprechung zu orthogonalen Matrizen im \mathbb{C}^n sind **unitäre** Matrizen. Auch hier verwendet man das unveränderte Skalarprodukt zur Definition.

Dabei muss natürlich statt $\mathbf{y}^T \mathbf{x}$ immer das komplexe Skalarprodukt $\mathbf{y}^H \mathbf{x}$ verwendet werden.

Die Aussagen von Satz 2.35 gelten dann analog – die Beziehung im letzten Punkt lautet dabei

$$U^{-1} = U^H.$$

Unitäre Abbildungen werden Ihnen möglicherweise in der Quantenmechanik begegnen.

① Grundlagen

② Lineare Algebra

2.1 Vektorräume

2.2 Matrizen und lineare Abbildungen

2.3 Lineare Gleichungssysteme

2.4 Determinanten

2.5 Invertierbare Matrizen

2.6 Orthogonalität, Skalarprodukt und Norm

2.7 Kreuz- und Spatprodukt

2.8 Elemente der analytischen Geometrie

2.9 Orthogonale Abbildungen

2.10 Eigenwerte und Eigenvektoren

2.11 Singulärwertzerlegung

Eigenwerte und Eigenvektoren

Motivation: Resonanzphänomene

Wir betrachten ein Flugzeug, das auf einer holperigen Piste landet.

Modell des Flugzeugs:

- drei Massen: m_1 (Rumpf und Motor); m_2, m_3 (Flügel),
- drei Steifigkeiten (k_1, k_2, k_3) für die 'federnde' Verbindung der Teile.

Die Piste wird durch eine Sinuskurve

$$r(t) = r_0 \sin(\omega_0 t)$$

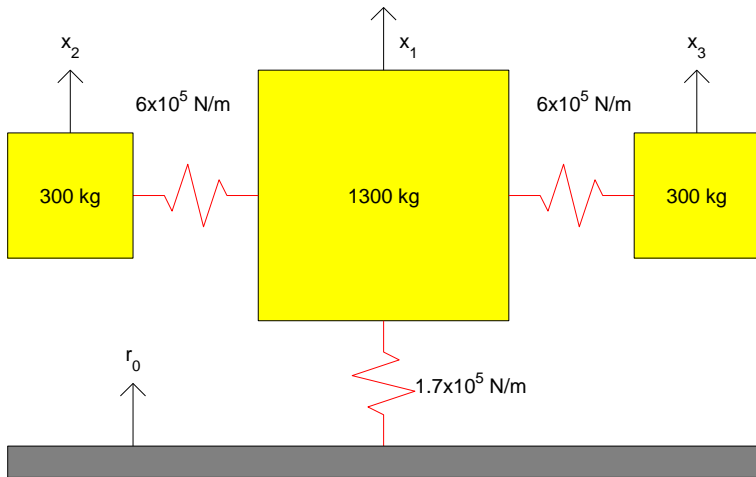
modelliert. Der Rumpf ist dann einer externen Kraft

$$f_1(t) = k_1 r_0 \sin(\omega_0 t)$$

ausgesetzt. Die Frequenz ω_0 hängt von der Landungsgeschwindigkeit v ab.

Eigenwerte und Eigenvektoren

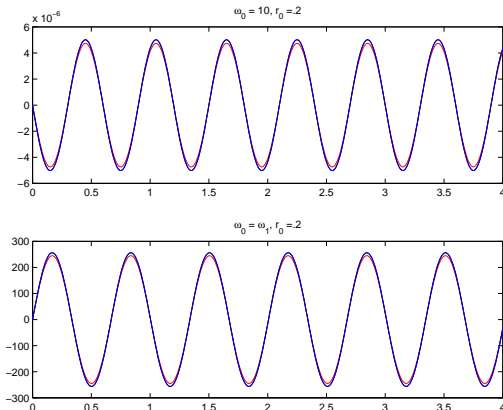
Motivation: Resonanzphänomene



Eigenwerte und Eigenvektoren

Motivation: Resonanzphänomene

Die mathematische Analyse des Beispiels wird erst am Ende des nächsten Semesters gelingen. Wir zeigen hier aber schon die Lösungen $x_{1,2,3}(t)$ [m] über t [s] für $v = 120$ km/h und $v = 108$ km/h:



Beachten Sie, dass sich die Amplituden bei den verschiedenen Landungsgeschwindigkeiten um 7 Größenordnungen(!!!) unterscheiden.

Eigenwerte und Eigenvektoren

Motivation: Resonanzphänomene

Im zweiten Fall tritt ein **Resonanzphänomen** auf: Wenn die Anregungsfrequenz ω_0 (nahezu) mit einer der **Eigenfrequenzen** $\omega_{1,2,3}$ des Flugzeugs übereinstimmt, kommt es zu gefährlich großen Oszillationen.

Die Nichtbeachtung von Eigenfrequenzen und Resonanz kann z. B. bei Brücken oder Hochhäusern katastrophale Auswirkungen haben:



Tacoma Narrows Bridge (WA, 1940).
Bild: Prelinger Archives

Weitere Beispiele:

Broughton suspension bridge, Manchester 1831;
Millennium footbridge, London 2000.

Eigenwerte und Eigenvektoren

Ranking relevanter Webseiten

Betrachte System aus n Webseiten (n groß), teilweise untereinander verlinkt.

Sei $\vec{x} \in \mathbb{R}^n$, $x_i \geq 0, i = 1, \dots, n$ sei ein Maß für die Relevanz der i -ten Webseite, etwa für eine Suchanfrage.

Prinzip: x_i hängt von der Wichtigkeit der auf x_i verweisenden Webseiten $\{x_j\}_{j=1}^n$ ab.

Sei $A \in \{0, 1\}^{n \times n}$ mit Einträgen

$$a_{i,j} = \begin{cases} 1 & \text{Webseite } j \text{ verweist auf Webseite } i \\ 0 & \text{sonst.} \end{cases}$$

Obigem Prinzip folgend könnte man dann folgendes Modell verwenden:

$$x_i = \sum_{j=1}^n a_{i,j} x_j, \quad \text{oder } A\vec{x} = \vec{x}$$

Diese (grundlegende) Ansatz liegt Google's PageRank-Verfahren [Brin & Page, 1998] zugrunde.

Eigenwerte und Eigenvektoren

Definition

Um z.B. Resonanzphänomene zu analysieren, benötigt man die folgenden Begriffe:

Definition 2.36

Sei $A \in \mathbb{R}^{n \times n}$ (oder $\mathbb{C}^{n \times n}$). Eine Zahl $\lambda \in \mathbb{C}$ heißt **Eigenwert** von A , wenn es einen Vektor $v \in \mathbb{C}^n$, $v \neq \mathbf{0}$, gibt, so dass

$$Av = \lambda v. \quad (2.15)$$

Jeder Vektor $v \in \mathbb{C}^n \setminus \{\mathbf{0}\}$, der (2.15) erfüllt, heißt **Eigenvektor** von A zum Eigenwert λ .

Achtung: Auch wenn wir nur reelle Matrizen betrachten: bei Eigenwerten und Eigenvektoren lässt man immer auch komplexe Zahlen zu!

Eigenwerte und Eigenvektoren

Definition

Die Richtung eines Eigenvektors wird durch die lineare Abbildung A nicht verändert – der Eigenvektor wird lediglich gestreckt.

Das rechte Bild entsteht z. B. aus dem linken durch eine ‘Scherung’ der Leinwand.

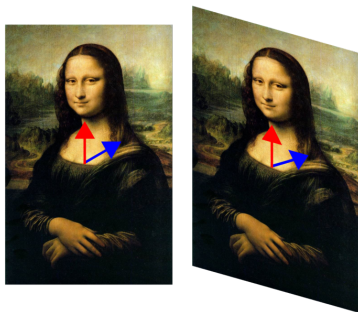


Bild: Wikimedia Commons

Der rote Vektor bleibt unverändert und ist damit ein Eigenvektor zum Eigenwert 1. Der blaue Vektor ändert hingegen seine Richtung und ist daher kein Eigenvektor der Scherungsabbildung.

Eigenwerte und Eigenvektoren

Welche reellen Eigenwerte besitzt eine Drehung im Raum um die x_1 -Achse, und welcher reelle Eigenvektor kommt in Frage?

Wie verhält es sich mit einer Spiegelung in der Ebene an einer Geraden durch 0 und senkrecht zu \vec{n} ?

Argumentieren Sie rein geometrisch!

Bestätigen Sie, dass $\lambda = 1$ Eigenwert von $A = \begin{bmatrix} 5 & -8 \\ -1 & 3 \end{bmatrix}$ mit zugehörigem Eigenvektor $\mathbf{v} = [2, 1]^T$ ist.

Zeigen Sie, dass jedes komplexe Vielfache des Vektors $[1, i]^T$ ein Eigenvektor von $B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ zum Eigenwert $\lambda = i$ ist.

Formulieren Sie eine allgemeingültige Aussage und bestätigen Sie diese durch Einsetzen in (2.15).

Anmerkung: Bislang wissen wir weder, wie man EW und EV berechnet, noch ob in den Beispielen alle EW und EV erfasst wurden.

Eigenwerte und Eigenvektoren

Berechnung von Eigenwerten

Wir benutzen zur Herleitung der Formel für die Eigenwertberechnung folgende Äquivalenzkette:

$\lambda \in \mathbb{C}$ ist Eigenwert von $A \in \mathbb{R}^{n \times n}$.

\Leftrightarrow Es gibt einen Vektor $v \neq \mathbf{0}$ mit $Av = \lambda v$.

\Leftrightarrow Ein Vektor $v \neq \mathbf{0}$ löst das homogene LGS $(A - \lambda I)v = \mathbf{0}$.

$\Leftrightarrow \det(A - \lambda I) = 0$.

Die Eigenwerte sind also gerade die Nullstellen der Funktion

$$c_A(\lambda) := \det(A - \lambda I).$$

Diese Funktion ist ein Polynom vom Grad n und wird das **charakteristische Polynom** von A genannt.

Eigenwerte und Eigenvektoren

Berechnung von Eigenwerten

Satz 2.37

Für $A \in \mathbb{R}^{n \times n}$ ist das charakteristische Polynom $c_A(\lambda) = \det(A - \lambda I)$ ein Polynom vom exakten Grad n mit reellen Koeffizienten und Höchstkoeffizienten 1 oder -1 .

Die Eigenwerte von A sind genau die Nullstellen von c_A , also die Lösungen der Gleichung

$$c_A(\lambda) = \det(A - \lambda I) \stackrel{!}{=} 0. \quad (2.16)$$

Berechnen Sie sämtliche Eigenwerte der Matrizen

$$A = \begin{bmatrix} 5 & -8 \\ -1 & 3 \end{bmatrix} \quad \text{und} \quad B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

(vgl. 2. und 3. Kasten auf S. 267).

Eigenwerte und Eigenvektoren

Berechnung von Eigenvektoren

Sei λ Eigenwert der Matrix A . Die zugehörigen Eigenvektoren von A sind die (nicht-trivialen) Lösungen des homogenen Gleichungssystems

$$(A - \lambda I)v = \mathbf{0}. \quad (2.17)$$

Zusammen mit $\mathbf{0}$ bilden sie einen Unterraum des \mathbb{C}^n , den sogenannten **Eigenraum** von A zum Eigenwert λ ,

$$\begin{aligned} \text{Eig}(A, \lambda) &:= \{\mathbf{x} \in \mathbb{C}^n : A\mathbf{x} = \lambda\mathbf{x}\} \\ &= \{\mathbf{x} \in \mathbb{C}^n : (A - \lambda I)\mathbf{x} = \mathbf{0}\} \\ &= \mathcal{N}(A - \lambda I) \end{aligned}$$

Berechnen Sie sämtliche Eigenvektoren zu den Eigenwerten der Matrizen A und B von Seite 269.

Definition 2.38

Zwei Matrizen $A, B \in \mathbb{C}^{n \times n}$ heißen **ähnlich**, wenn es eine invertierbare Matrix $V \in \mathbb{C}^{n \times n}$ gibt, so dass

$$A = V^{-1}BV. \quad (2.18)$$

Bemerkungen:

1. Gleichung (2.18) ist äquivalent zu

$$VA = BV \quad \text{und} \quad B = VAV^{-1}.$$

2. Ähnlichkeit ist eine Äquivalenzrelation.

Eigenwerte und Eigenvektoren

Ähnliche Matrizen

Ähnliche Matrizen haben das gleiche charakteristische Polynom, wie folgende Rechnung zeigt: aus $A = V^{-1}BV$ erhalten wir

$$\begin{aligned}c_A(\lambda) &= c_{V^{-1}BV}(\lambda) = \det(V^{-1}BV - \lambda I) \\ &= \det(V^{-1}(B - \lambda I)V) \\ &= \det(V^{-1}) \det(B - \lambda I) \det V \\ &= (\det V)^{-1} \det(B - \lambda I) \det V \\ &= c_B(\lambda).\end{aligned}$$

Wir fassen zusammen:

Satz 2.39

Ähnliche Matrizen haben das gleiche charakteristische Polynom und damit die gleichen Eigenwerte.

Die Eigenvektoren zu einem Eigenwert λ sind dabei allerdings verschieden: \vec{x} ist genau dann Eigenvektor von $A = V^{-1}BV$, wenn $V\vec{x}$ Eigenvektor von B ist.

Eigenwerte und Eigenvektoren

Hintergrund: Ähnlichkeit und Basistransformation

Um den Ähnlichkeitsbegriff vollständig zu verstehen, muss man sich mit Basistransformationen auseinandersetzen. Wir gehen dabei von der Darstellung $B = VAV^{-1}$ von S. 271 aus.

Zu einem gegebenen Vektor x lässt sich Vx als Linearkombination der Spalten v_1, \dots, v_n von V schreiben:

$$Vx = x_1 v_1 + x_2 v_2 + \dots + x_n v_n.$$

Fasst man \vec{x} als Koordinatenvektor bezüglich der Basis $\mathcal{B}_V = \{v_1, \dots, v_n\}$ auf, so ist Vx gerade der zugehörige Koordinatenvektor bezüglich der Standardbasis.

Man sagt daher auch, V stellt eine **Basistransformation** von der Basis \mathcal{B}_V in die Standardbasis dar.

Eigenwerte und Eigenvektoren

Hintergrund: Ähnlichkeit und Basistransformation

Die Abbildung V^{-1} macht die Transformation rückgängig und stellt somit eine Basistransformation von der Standardbasis nach \mathcal{B}_V dar.

Verifizieren Sie beide Aussagen am Beispiel der Basis $\mathcal{B} = \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$ und des Vektors $\begin{bmatrix} \sqrt{2} \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}_{\mathcal{B}}$.

In $Bx = VAV^{-1}x$ kann man die rechte Seite nun von rechts nach links wie folgt lesen:

- Stelle x als Koordinatenvektor bezüglich der Basis $\mathcal{B}_V = \{v_1, \dots, v_n\}$ dar (d. h. Multiplikation mit V^{-1}).
- Führe die durch B beschriebene lineare Abbildung in den Koordinaten bezüglich \mathcal{B}_V aus (d. h. Multiplikation mit A).
- Transformiere das Ergebnis wieder zurück in die Standardkoordinaten (Multiplikation mit V).

Eigenwerte und Eigenvektoren

Hintergrund: Ähnlichkeit und Basistransformation

Wir haben somit erkannt:

Satz 2.40

Ähnliche Matrizen stellen die gleiche lineare Abbildung bezüglich verschiedener Basen dar.

Folglich besitzen ähnliche Matrizen

- dieselbe Determinante,
- denselben Rang und
- denselben Defekt.

Auch die Aussage von Satz 2.39 wird mit dieser Erkenntnis noch ein Stück verständlicher.

Eigenwerte und Eigenvektoren

Diagonalisierbare Matrizen

Definition 2.41

Eine Matrix $A \in \mathbb{C}^{n \times n}$ heißt **diagonalisierbar**, wenn A ähnlich zu einer Diagonalmatrix $D \in \mathbb{C}^{n \times n}$ ist, d. h. wenn es eine invertierbare Matrix $V \in \mathbb{C}^{n \times n}$ gibt mit

$$D = V^{-1}AV.$$

Satz 2.42

Eine Matrix $A \in \mathbb{C}^{n \times n}$ ist genau dann diagonalisierbar, wenn es n linear unabhängige Eigenvektoren von A – also eine Basis aus Eigenvektoren von A – gibt.

Eigenwerte und Eigenvektoren

Diagonalisierbare Matrizen

Sei $A \in \mathbb{C}^{n \times n}$ eine Matrix, für die eine Basis des \mathbb{C}^n aus Eigenvektoren von A existiert (vgl. Satz 2.42).

Dann kann die Abbildung A bezüglich dieser Basis nur durch Streckungen der Basisvektoren ausgedrückt werden (Multiplikation mit einer Diagonalmatrix).

Beispiel. Die Matrix $A = \begin{bmatrix} 5 & -8 \\ -1 & 3 \end{bmatrix}$ besitzt die Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = 7$ mit den Eigenvektoren $\mathbf{v}_1 = [2, 1]^T$ und $\mathbf{v}_2 = [-4, 1]^T$ (vgl. S. 269). Es gilt die Darstellung

$$A = VDV^{-1} = \begin{bmatrix} 2 & -4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 7 \end{bmatrix} \begin{bmatrix} 2 & -4 \\ 1 & 1 \end{bmatrix}^{-1}.$$

In der Diagonalmatrix D stehen die Eigenwerte und in den Spalten von V die Eigenvektoren.

Eigenwerte und Eigenvektoren

Diagonalisierbare Matrizen

Wir widmen uns nun der Frage, wie groß die Dimension des von den Eigenvektoren einer Matrix $A \in \mathbb{C}^{n \times n}$ aufgespannten Unterraums ist.

Insbesondere wollen wir wissen, **wann** es eine Basis des \mathbb{C}^n aus Eigenvektoren von A gibt.

Wir beginnen mit folgendem Satz:

Satz 2.43

*Gehören die Eigenvektoren v_1, \dots, v_r zu **verschiedenen** Eigenwerten $\lambda_1, \dots, \lambda_r$ der $n \times n$ -Matrix A , dann sind sie linear unabhängig.*

Machen Sie sich dies zumindest für den Fall $r = 2$ klar.

Wenn es n **verschiedene** Eigenwerte zur Matrix $A \in \mathbb{C}^{n \times n}$ gibt, existiert also eine Basis des \mathbb{C}^n , die nur aus Eigenvektoren von A besteht.

Definition 2.44

Die **algebraische Vielfachheit** $\nu_{\text{alg}}(\lambda)$ eines Eigenwerts λ von A ist seine Vielfachheit als Nullstelle des charakteristischen Polynoms c_A von A .

Die **geometrische Vielfachheit** $\nu_{\text{geom}}(\lambda)$ eines Eigenwerts λ von A ist die Dimension des zugehörigen Eigenraums $\dim(\text{Eig}(A, \lambda))$.

Anmerkung zur Berechnung Um die algebraischen Vielfachheiten zu bestimmen, muss man nur das charakteristische Polynom c_A kennen. Um die geometrische Vielfachheiten zu berechnen, reicht die Kenntnis von c_A allein nicht aus.

Eigenwerte und Eigenvektoren

Vielfachheiten von Eigenwerten

Nach dem Fundamentalsatz der Algebra summieren sich die algebraischen Vielfachheiten der Eigenwerte einer Matrix $A \in \mathbb{C}^{n \times n}$ zu n .

Die Summe der geometrischen Vielfachheiten kann dagegen kleiner sein. Es gilt folgender Satz:

Satz 2.45

Sei λ Eigenwert der Matrix $A \in \mathbb{C}^{n \times n}$ mit algebraischer Vielfachheit $\nu_{\text{alg}}(\lambda)$ und geometrischer Vielfachheit $\nu_{\text{geom}}(\lambda)$. Dann gilt:

$$1 \leq \nu_{\text{geom}}(\lambda) \leq \nu_{\text{alg}}(\lambda) \leq n.$$

Machen Sie sich am Beispiel der Matrix $\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$ klar, dass die geometrische Vielfachheit eines Eigenwerts tatsächlich kleiner sein kann als die algebraische.

Eigenwerte und Eigenvektoren

Vielfachheiten von Eigenwerten

Beim Beispiel im Kasten auf S. 280 handelt es sich um einen sogenannten **Jordan-Block**, d. h. eine Matrix der Form

$$\begin{bmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

Dieser besitzt nur den Eigenwert λ mit geometrischer Vielfachheit 1 und algebraischer Vielfachheit n .

Bestätigen Sie diese Aussagen. Für die geometrische Vielfachheit nutzen Sie am besten die Beziehung

$$\nu_{\text{geom}}(\lambda) = \dim(\text{Eig}(A, \lambda)) = n - \text{Rang}(A - \lambda I).$$

Eigenwerte und Eigenvektoren

Diagonalisierbare Matrizen

Diagonalisierbarkeit lässt sich nun mittels Vielfachheiten auch folgendermaßen charakterisieren:

Satz 2.46

Eine Matrix $A \in \mathbb{C}^{n \times n}$ ist genau dann diagonalisierbar, wenn die algebraische und geometrische Vielfachheit für jeden Eigenwert übereinstimmen.

Sind $\lambda_1, \dots, \lambda_n$ die Eigenwerte und v_1, \dots, v_n zugeordnete Eigenvektoren, die eine Basis des \mathbb{C}^n bilden, so gilt die Darstellung

$$A = VDV^{-1}.$$

Dabei gilt $D = \text{diag}(\lambda_1, \dots, \lambda_n)$, und v_1, \dots, v_n sind in dieser Reihenfolge die Spalten von V .

Die Matrix $A = \begin{bmatrix} 5 & -8 \\ -1 & 3 \end{bmatrix}$ besitzt die Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = 7$ mit den Eigenvektoren $v_1 = [2, 1]^T$ und $v_2 = [-4, 1]^T$ (vgl. S. 269 f.). Sie bilden eine Basis des \mathbb{C}^2 (und in diesem Falle auch eine Basis des \mathbb{R}^2).

Eigenwerte und Eigenvektoren

Mögliche Konstellationen für $A \in \mathbb{R}^{2 \times 2}$

Das charakteristische Polynom einer reellen 2×2 -Matrix A besitzt die Struktur

$$c_A(\lambda) = \lambda^2 + p\lambda + q \quad (\text{mit } p, q \in \mathbb{R}).$$

In Kombination mit Satz 2.45 ergeben sich verschiedene Möglichkeiten für die Eigenwerte. Es existieren

- entweder zwei verschiedene reelle Eigenwerte λ_1 und λ_2 , die geometrische und algebraische Vielfachheit 1 haben. Zu jedem Eigenwert gibt es reelle Eigenvektoren.
- oder zwei verschiedene konjugiert-komplexe Eigenwerte λ_1 und λ_2 , die geometrische und algebraische Vielfachheit 1 haben. Zu jedem Eigenwert gibt es komplexe Eigenvektoren.

(b. w.)

Eigenwerte und Eigenvektoren

Mögliche Konstellationen für $A \in \mathbb{R}^{2 \times 2}$

- oder nur einen Eigenwert λ , der dann zwangsläufig reell (konjugiert-komplex zu sich selbst) ist.

Er besitzt die algebraische Vielfachheit 2 und

- entweder die geometrische Vielfachheit 2, d. h. jeder Vektor aus \mathbb{C}^2 ist Eigenvektor. A ist dann ähnlich zu λI , d.h. es existiert eine invertierbare Matrix V mit zwei linear unabhängigen Eigenvektoren zu λ als Spalten sodass

$$V^{-1}AV = \lambda I = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$$

- oder die geometrische Vielfachheit 1.

Ordnen Sie die Beispiele von S. 269 und S. 280 den entsprechenden Fällen zu.

Führen Sie eine ähnliche Analyse für den 3×3 -Fall durch. (Auf die Unterscheidung zwischen reellen und komplexen Eigenwerten können Sie dabei verzichten.)

Gilt $\nu_{\text{geom}}(\lambda) < \nu_{\text{alg}}(\lambda)$ für ein Eigenwert λ einer Matrix $A \in \mathbb{C}^{n \times n}$, so ist A nicht diagonalisierbar. Folgende 'fast diagonale' Form existiert aber im allgemeinen Fall:

Eigenwerte und Eigenvektoren

Jordansche Normalform

Satz 2.47 (Jordansche Normalform)

Jede Matrix $A \in \mathbb{C}^{n \times n}$ ist ähnlich zu einer Blockdiagonalmatrix $J = \text{diag}(J_1, \dots, J_s)$ mit Diagonalblöcken der Form

$$J_k = \begin{bmatrix} \lambda_k & 1 & & \\ & \lambda_k & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_k \end{bmatrix}, \quad \text{d.h.} \quad A = VJV^{-1} \quad (2.19)$$

mit Eigenwerten λ_k von A . Zu jedem Eigenwert von A existieren ein oder mehrere solcher **Jordan-Blöcke**. Ist n_j die Dimension des größten zum Eigenwert λ_j von A gehörenden Jordan-Blocks, so heißt $m_A(\lambda) = \prod_j (\lambda - \lambda_j)^{n_j}$ das **Minimalpolynom** von A , welches das charakteristische Polynom von A teilt. Für die Spaltenvektoren v von V im Indexbereich eines Jordanblocks zum Eigenwert λ_j gilt

$$(A - \lambda_j I)^\ell v = \mathbf{0} \quad \text{für ein } \ell \text{ mit } 1 \leq \ell \leq n_j.$$

Bemerkungen:

- (1) Die Spalten der Ähnlichkeitstransformation V in (2.19) heißen **Hauptvektoren**. Zu jeden Eigenwert λ von A existieren $\nu_{\text{alg}}(\lambda)$ linear unabhängige Hauptvektoren. Da Hauptvektoren zu verschiedenen Eigenwerten linear unabhängig sind existiert somit eine Basis des \mathbb{C}^n aus Hauptvektoren von A .
- (2) Im Fall $\nu_{\text{geom}}(\lambda) = \nu_{\text{alg}}(\lambda)$ sind alle Hauptvektoren von A zum Eigenwert λ auch Eigenvektoren.
- (3) Es gilt $m_A(A) = O$ und m_A ist das Polynom niedrigsten Grades mit dieser Eigenschaft. Somit gilt auch $c_A(A) = O$, was auch als Satz von Cayley-Hamilton bekannt ist.

Eigenwerte und Eigenvektoren

Eigenwerte spezieller Matrizen

Satz 2.48

Sei $A \in \mathbb{R}^{n \times n}$. Dann gelten:

- A und A^T besitzen dasselbe charakteristische Polynom, also dieselben Eigenwerte (mit i. A. verschiedenen Eigenräumen).
- Besitzt A den Eigenvektor x zum Eigenwert λ , dann besitzen

$$\alpha A, A^m, A + \beta I_n, p(A) = \alpha_m A^m + \cdots + \alpha_1 A + \alpha_0 I_n$$

denselben Eigenvektor x , allerdings zum Eigenwert

$$\alpha\lambda, \lambda^m, \lambda + \beta, p(\lambda) = \alpha_m \lambda^m + \cdots + \alpha_1 \lambda + \alpha_0.$$

- A ist genau dann invertierbar, wenn alle Eigenwerte von A von 0 verschieden sind. Ist dann λ ein Eigenwert von A mit Eigenvektor x , so ist λ^{-1} ein Eigenwert von A^{-1} mit demselben Eigenvektor x .

Verifizieren Sie einige dieser Aussagen.

Eigenwerte und Eigenvektoren

Eigenwerte spezieller Matrizen

Weiterhin kann man Eigenwerte sehr einfach bestimmen, wenn A bestimmte strukturelle Eigenschaften besitzt:

Satz 2.49

Ist A eine (untere oder obere) Dreiecksmatrix, so sind die Hauptdiagonaleinträge von A genau die Eigenwerte von A .

Dies trifft insbesondere dann zu, wenn A eine Diagonalmatrix ist. In diesem Fall sind die Einheitsvektoren zugehörige Eigenvektoren.

Erinnerung/Bemerkung: Eine Diagonalmatrix enthält nur auf der Hauptdiagonalen Einträge ungleich Null. Sind $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ die Einträge entlang der Diagonalen, so schreibt man $D = \text{diag}(\lambda_1, \dots, \lambda_n)$.

Bei Dreiecksmatrizen sind alle Elemente ober- bzw. unterhalb der Hauptdiagonalen gleich Null.

Eigenwerte und Eigenvektoren

Eigenwerte symmetrischer Matrizen

Eine Matrix $A \in \mathbb{R}^{n \times n}$ heißt **symmetrisch**, wenn $A = A^T$ gilt, d. h. die Einträge symmetrisch zur Hauptdiagonalen liegen.

Symmetrische Matrizen haben bemerkenswerte Eigenschaften:

Satz 2.50

Ist $A \in \mathbb{R}^{n \times n}$ symmetrisch, so besitzt A nur reelle Eigenwerte. Es gibt eine Orthonormalbasis des \mathbb{R}^n , die aus Eigenvektoren von A besteht.

Es gibt also eine Diagonalmatrix $D \in \mathbb{R}^{n \times n}$ und eine orthogonale Matrix $U \in \mathbb{R}^{n \times n}$ (d. h. $U^T = U^{-1}$), so dass

$$A = UDU^T.$$

Insbesondere ist A diagonalisierbar.

Eigenwerte und Eigenvektoren

Eigenwerte orthogonaler Matrizen

Für orthogonale Matrizen gibt es ein ähnliches Ergebnis. Beachten Sie aber, dass die Eigenwerte hier i. A. komplex sind!

Satz 2.51

Ist $A \in \mathbb{R}^{n \times n}$ orthogonal, so besitzt A nur Eigenwerte mit Betrag 1. Es gibt eine Orthonormalbasis des \mathbb{C}^n , die aus Eigenvektoren von A besteht.

Es gibt also eine Diagonalmatrix $D \in \mathbb{C}^{n \times n}$ und eine unitäre Matrix $U \in \mathbb{C}^{n \times n}$ (d. h. $U^H = U^{-1}$), so dass

$$A = UDU^H.$$

Insbesondere ist A diagonalisierbar.

Bemerkung: Die Klasse der Matrizen, welche eine Basis aus **orthonormalen** Eigenvektoren besitzen, ist größer als die der symmetrischen (Hermiteschen) Matrizen. Genau trifft dies zu für **normale** Matrizen. Diese sind charakterisiert durch die Eigenschaft

$$AA^T = A^T A, \quad \text{bzw.} \quad AA^H = A^H A.$$

Eigenwerte und Eigenvektoren

Eigenwerte orthogonaler Matrizen

Bestimmen Sie die Eigenwerte und Eigenvektoren der Matrizen

$$A = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & -3 & 1 \end{bmatrix} \quad \text{und} \quad D_\alpha = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}$$

Geben Sie in beiden Fällen eine Darstellung der Form VDV^{-1} an. Erkennen Sie die Matrix B von S. 269 f. wieder?

Eigenwerte und Eigenvektoren

Anwendung: Hauptachsentransformation

Die durch die Gleichungen

$$\frac{x^2}{4} + \frac{y^2}{9} = 1 \quad \text{bzw.} \quad y = 4(x + 1)^2 + 2$$

festgelegten Punktmenge erkennt man leicht als **Ellipse** symmetrisch zu den Koordinatenachsen mit Halbachsen 2 und 3 bzw. nach oben offene **Parabel** mit Scheitelpunkt $(x_S, y_S) = (-1, 2)$.

Dies ist weniger offensichtlich bei

$$\frac{x^2}{9} + 4xy - \frac{y^2}{4} + x - 2y = 4.$$

Die **Hauptachsentransformation** bietet eine Variablensubstitution, mit Hilfe derer eine Klassifikation quadratischer Funktionen einfach möglich ist.

Eigenwerte und Eigenvektoren

Anwendung: Hauptachsentransformation

Jede quadratische Funktion $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ besitzt die allgemeine Form

$$\phi(\mathbf{x}) = \mathbf{x}^T A \mathbf{x} + \mathbf{b}^T \mathbf{x} + c \quad (2.20)$$

mit einer symmetrischen Matrix $A \in \mathbb{R}^{n \times n}$, einem Vektor \mathbf{b} und einer Zahl $c \in \mathbb{R}$.

Beispiel: $n = 2$

$$\phi(x_1, x_2) = a_{1,1}x_1^2 + a_{1,2}x_1x_2 + a_{2,1}x_2x_1 + a_{2,2}x_2^2 + b_1x_1 + b_2x_2 + c$$

Die Aufgabe der **Hauptachsentransformation** besteht darin, durch einen geeigneten Basiswechsel im \mathbb{R}^n die quadratische Funktion in eine einfache Form zu bringen, an der man leicht ihren Typ ablesen kann.

Dies ist auch hilfreich bei der Klassifikation von **Quadriken** (auch **Kegelschnitte** genannt), also den Punktmengen

$$Q = \{\mathbf{x} \in \mathbb{R}^n : \phi(\mathbf{x}) = 0\}.$$

Eigenwerte und Eigenvektoren

Anwendung: Hauptachsentransformation

Da die Matrix A ohne Beschränkung der Allgemeinheit als symmetrisch angenommen werden kann besitzt diese eine Basis aus orthonormalen Eigenvektoren, d.h. eine orthogonale Matrix Q mit

$$Q^T A Q = D = \text{diag}(\lambda_1, \dots, \lambda_n)$$

mit reellen Eigenwerten $\lambda_1, \dots, \lambda_n$. Der Hauptteil $x^T A x$ einer quadratischen Funktion (2.20) heißt auch **quadratische Form**. Die orthogonalen Eigenrichtungen der Matrix A heißen **Hauptachsen** der quadratischen Funktion.

Für die neue Variable $\mathbf{y} := Q^T \mathbf{x}$ gilt dann $\mathbf{x} = Q \mathbf{y}$ und, mit dem neuen Vektor $\mathbf{d} := Q^T \mathbf{b}$ erhalten wir

$$\begin{aligned} \phi(\mathbf{x}) &= \mathbf{x}^T A \mathbf{x} + \mathbf{b}^T \mathbf{x} + c = (Q \mathbf{y})^T A (Q \mathbf{y}) + \mathbf{b}^T (Q \mathbf{y}) + c \\ &= \mathbf{y}^T Q^T A Q \mathbf{y} + \mathbf{b}^T Q \mathbf{y} + c = \mathbf{y}^T D \mathbf{y} + \mathbf{d}^T \mathbf{y} + c \\ &= \sum_{j=1}^n \lambda_j y_j^2 + \sum_{j=1}^n d_j y_j + c. \end{aligned}$$

Eigenwerte und Eigenvektoren

Anwendung: Hauptachsentransformation

Als letzter Schritt folgt nun quadratisches Ergänzen in jeder Variablen y_j :

$$\lambda_j y_j^2 + d_j y_j = \lambda_j \left(y_j + \frac{d_j}{2\lambda_j} \right)^2 - \frac{d_j^2}{4\lambda_j}.$$

Mit einer weiteren Substitution $z_j = y_j + \frac{d_j}{2\lambda_j}$, $j = 1, \dots, n$ erhalten wir

$$\phi(\mathbf{x}) = \sum_{j=1}^n \lambda_j z_j^2 + d = \lambda_1 z_1^2 + \dots + \lambda_n z_n^2 + d.$$

mit

$$d := c - \sum_{j=1}^n \frac{d_j^2}{4\lambda_j}.$$

Führen Sie die Hauptachsentransformation durch für die Quadrik

$$2x^2 - y^2 + 4xy - 2x + y - 6 = 0.$$

Eigenwerte und Eigenvektoren

Quadrikeneinteilung im \mathbb{R}^2

Alle Eigenwerte von A ungleich Null

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0 \quad \text{Ellipse mit Halbachsen } a, b$$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + 1 = 0 \quad \text{leere Menge}$$

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 = 0 \quad \text{Hyperbel}$$

$$x^2 + a^2y^2 = 0 \quad \text{Punkt } \{(0, 0)\}$$

$$x^2 - a^2y^2 = 0 \quad \text{Geradenpaar } y = \pm|a|x$$

Ein Eigenwert von A gleich Null

$$x^2 - 2py = 0 \quad \text{Parabel}$$

$$x^2 - a^2 = 0 \quad \text{paralleles Geradenpaar}$$

$$x^2 + a^2 = 0 \quad \text{leere Menge}$$

$$x^2 = 0 \quad \text{Gerade } x = 0 \text{ (} y\text{-Achse)}$$

① Grundlagen

② Lineare Algebra

2.1 Vektorräume

2.2 Matrizen und lineare Abbildungen

2.3 Lineare Gleichungssysteme

2.4 Determinanten

2.5 Invertierbare Matrizen

2.6 Orthogonalität, Skalarprodukt und Norm

2.7 Kreuz- und Spatprodukt

2.8 Elemente der analytischen Geometrie

2.9 Orthogonale Abbildungen

2.10 Eigenwerte und Eigenvektoren

2.11 Singulärwertzerlegung

Singulärwertzerlegung

Auch wenn sie zum Verständnis wichtig sind weisen die bisher betrachteten Zerlegungen bzw. Normalformen für Matrizen einige Schönheitsfehler auf:

- Nicht alle Matrizen sind diagonalisierbar, und die Jordansche Normalform lässt sich bei Rundungsfehlern nicht stabil berechnen.
- Auch bei diagonalisierbaren Matrizen können Eigenvektoren zu verschiedenen Eigenwerten nahezu linear abhängig sein, sodass auch hier sich Rundungsfehler stark bemerkbar machen können.
- In vielen Anwendungen beschreiben Matrizen Daten oder Signale und sind nicht notwendig quadratisch, d.h. es sind auch Zerlegungen für rechteckige Matrizen nötig.

Die **Singulärwertzerlegung** stellt die für die Praxis wichtigste Matrixzerlegung dar, mit Anwendungen u.a. in der Datenanalyse, Maschinenlernen, Computer-Vision, digitalen Signalverarbeitung, insbesondere Signal-Rauschentrennung.

Satz 2.52 (Singulärwertzerlegung)

Sei $A \in \mathbb{R}^{m \times n}$ eine Matrix vom Rang r . Dann gibt es orthogonale Matrizen $U \in \mathbb{R}^{m \times m}$ und $V \in \mathbb{R}^{n \times n}$ sowie eine 'Diagonalmatrix'

$$\Sigma = \begin{bmatrix} \Sigma_r & O \\ O & O \end{bmatrix} \in \mathbb{R}^{m \times n} \quad \text{mit} \quad \Sigma_r = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r) \in \mathbb{R}^{r \times r}$$

und $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$, so dass A die Zerlegung

$$A = U \Sigma V^T \quad (\text{SVD})$$

besitzt.

Die Darstellung (SVD) heißt **Singulärwertzerlegung** von A (engl. singular value decomposition).

Die positiven Zahlen $\sigma_1, \dots, \sigma_r$ nennt man die **Singulärwerte** von A .

Schreibt man $U = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m]$ und $V = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]$, so heißen $\mathbf{u}_i \in \mathbb{R}^m$ bzw. $\mathbf{v}_i \in \mathbb{R}^n$ zugehörige **linke** bzw. **rechte Singulärvektoren**.

(1) Darstellung von A als Summe von r Rang-1-Matrizen:

$$A = U\Sigma V^\top = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r] \Sigma_r [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r]^\top = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^\top$$

(2) Es gelten:

$$A\mathbf{v}_i = \begin{cases} \sigma_i \mathbf{u}_i & i = 1, 2, \dots, r, \\ \mathbf{0} & i = r + 1, \dots, n \end{cases}$$

und

$$A^\top \mathbf{u}_i = \begin{cases} \sigma_i \mathbf{v}_i & i = 1, 2, \dots, r, \\ \mathbf{0} & i = r + 1, \dots, m. \end{cases}$$

(3)

$\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$	ist eine ON-Basis von	$\mathcal{R}(A)$.
$\{\mathbf{u}_{r+1}, \dots, \mathbf{u}_m\}$	ist eine ON-Basis von	$\mathcal{N}(A^\top) = \mathcal{R}(A)^\perp$.
$\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$	ist eine ON-Basis von	$\mathcal{R}(A^\top) = \mathcal{N}(A)^\perp$.
$\{\mathbf{v}_{r+1}, \dots, \mathbf{v}_n\}$	ist eine ON-Basis von	$\mathcal{N}(A)$.

(4) $\sigma_1^2, \dots, \sigma_r^2$ sind die von Null verschiedenen Eigenwerte von $A^\top A$ bzw. AA^\top :

$$A^\top A = V \Sigma^\top \Sigma V^\top = V \begin{bmatrix} \Sigma_r^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} V^\top,$$

$$AA^\top = U \Sigma \Sigma^\top U^\top = U \begin{bmatrix} \Sigma_r^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} U^\top.$$

Insbesondere sind die Singulärwerte $\sigma_1, \dots, \sigma_r$ durch A eindeutig festgelegt. Die rechten Singulärvektoren $\mathbf{v}_1, \dots, \mathbf{v}_n$ bilden eine ON-Basis des \mathbb{R}^n aus Eigenvektoren von $A^\top A$:

$$A^\top A \mathbf{v}_i = \begin{cases} \sigma_i^2 \mathbf{v}_i & i = 1, 2, \dots, r, \\ \mathbf{0} & i = r + 1, \dots, n. \end{cases}$$

Die linken Singulärvektoren $\mathbf{u}_1, \dots, \mathbf{u}_m$ bilden eine ON-Basis des \mathbb{R}^m aus Eigenvektoren von AA^\top :

$$AA^\top \mathbf{u}_i = \begin{cases} \sigma_i^2 \mathbf{u}_i & i = 1, 2, \dots, r, \\ \mathbf{0} & i = r + 1, \dots, m. \end{cases}$$

Singulärwertzerlegung

Eigenschaften

- (5) Ist $A = A^T \in \mathbb{R}^{n \times n}$ mit von Null verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_r$, $|\lambda_1| \geq \dots \geq |\lambda_r| > 0$, dann sind $\sigma_i = |\lambda_i|$ die Singulärwerte von A .
- (6) Das Bild der (n -dimensionalen) Einheitskugel unter A ist ein Ellipsoid (im \mathbb{R}^m) mit Mittelpunkt $\mathbf{0}$ und Halbachsen $\sigma_i \mathbf{u}_i$ ($\sigma_i := 0$ für $i > r$).
- (7) Für $A \in \mathbb{R}^{m \times n}$ gilt $\|A\|_2 = \sigma_1$.
Ist $A \in \mathbb{R}^{n \times n}$ invertierbar, gilt außerdem $\|A^{-1}\|_2 = \sigma_n^{-1}$.
- (8) Analoge Aussagen gelten für komplexe Matrizen $A = U\Sigma V^H$ (U, V unitär).
(In (5) ist 'symmetrisch' durch 'normal' zu ersetzen.)

Singulärwertzerlegung

Anwendung: lineare Ausgleichsrechnung

Die **lineare Ausgleichsrechnung** führt auf Minimierungsprobleme der Form

$$\|\mathbf{b} - A\mathbf{x}\|_2 \rightarrow \min_{\mathbf{x} \in \mathbb{R}^n} \quad (2.21)$$

zu gegebener Matrix $A \in \mathbb{R}^{m \times n}$ ($m \geq n$) und Vektor $\mathbf{b} \in \mathbb{R}^m$.

Satz 2.53

Es sei $A \in \mathbb{R}^{m \times n}$ eine Matrix vom Rang r mit Singulärwertzerlegung $A = U\Sigma V^\top = U \begin{bmatrix} \Sigma_r & O \\ O & O \end{bmatrix} V^\top$. Dann löst

$$\mathbf{x}^* = V \begin{bmatrix} \Sigma_r^{-1} & O \\ O & O \end{bmatrix} U^\top \mathbf{b}$$

die lineare Ausgleichsaufgabe (2.21). Darüberhinaus ist \mathbf{x}^* die eindeutig bestimmte Lösung von (2.21) mit minimaler Euklid-Norm.

Mit Hilfe der Singulärwertzerlegung einer Matrix A lassen sich optimale Approximationen an A durch Matrizen niedrigen Rangs konstruieren. Hierzu ist ein Abstands begriff für Matrizen erforderlich. Die in Satz 2.27 eingeführten Normen für Vektoren lassen sich auch auf Matrizen verallgemeinern:

Definition 2.54

Eine **Matrixnorm** ist eine Abbildung

$$\|\cdot\| : \mathbb{K}^{n \times n} \rightarrow \mathbb{R}_0^+, \quad A \mapsto \|A\|, \quad \text{die}$$

- **positiv definit**, d.h. $\|A\| > 0 \quad \forall A \in \mathbb{K}^{n \times n}, A \neq O$, sowie
- **homogen**, d.h. $\|\alpha A\| = |\alpha| \|A\| \quad \forall \alpha \in \mathbb{K}, A \in \mathbb{K}^{n \times n}$, ist und der
- **Dreiecksungleichung**, d.h. $\|A + B\| \leq \|A\| + \|B\| \quad \forall A, B \in \mathbb{K}^{n \times n}$, genügt. Zusätzlich soll sie
- **submultiplikativ** sein, d.h. $\|AB\| \leq \|A\| \|B\| \quad \forall A, B \in \mathbb{K}^{n \times n}$.

Beispiel: **Frobenius** oder **Schur-Norm** ($A = [a_{i,j}]_{1 \leq i,j \leq n} \in \mathbb{K}^{n \times n}$)

$$\|A\|_F := \left(\sum_{i,j=1}^n |a_{i,j}|^2 \right)^{1/2}.$$

Definition 2.55

Jede Vektornorm $\|\cdot\|_V$ im \mathbb{K}^n induziert durch

$$\|A\|_M := \max_{\|x\|_V=1} \|Ax\|_V = \max_{x \neq 0} \frac{\|Ax\|_V}{\|x\|_V}$$

eine Matrixnorm in $\mathbb{K}^{n \times n}$, die von $\|\cdot\|_V$ **induzierte Matrixnorm**.

Es ist üblich, für $\|\cdot\|_V$ und $\|\cdot\|_M$ das gleiche Symbol zu verwenden:

- $\|\cdot\|_1$ induziert die **Spaltensummennorm** $\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{i,j}|$.
- $\|\cdot\|_2$ induziert die **Spektralnorm** $\|A\|_2 = \sqrt{\lambda_{\max}(A^H A)}$.
- $\|\cdot\|_\infty$ induziert die **Zeilensummennorm** $\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{i,j}|$.

Definition 2.56

Eine Vektornorm $\|\cdot\|_V$ und eine Matrixnorm $\|\cdot\|_M$ sind miteinander **verträglich** (oder passen zueinander), wenn

$$\|A\mathbf{x}\|_V \leq \|A\|_M \|\mathbf{x}\|_V \quad \forall \mathbf{x} \in \mathbb{K}^n, A \in \mathbb{K}^{n \times n}$$

gilt.

- Bezeichnet $\|\cdot\|_M$ die von der Vektornorm $\|\cdot\|_V$ induzierte Matrixnorm, so sind $\|\cdot\|_V$ und $\|\cdot\|_M$ miteinander verträglich.
- $\|\cdot\|_M$ ist die kleinste Matrixnorm, die mit $\|\cdot\|_V$ verträglich ist: Für alle $A \in \mathbb{K}^{n \times n}$ gilt

$$\|A\|_M = \min\{\|A\| : \|\cdot\| \text{ ist mit } \|\cdot\|_V \text{ verträglich}\}.$$

- Die Euklidische Vektornorm $\|\cdot\|_2$ ist mit der Frobenius-Norm $\|\cdot\|_F$ verträglich, was $\|A\|_2 \leq \|A\|_F$ für alle $A \in \mathbb{K}^{n \times n}$ impliziert.

Singulärwertzerlegung

Singulärwertzerlegung: Niedrigrangapproximation

Satz 2.57 (Schmidt, 1907; Eckart & Young, 1936; Mirsky, 1960)

Für eine Matrix $A \in \mathbb{R}^{m \times n}$ vom Rang r mit Singulärwertzerlegung $A = U\Sigma V^\top$ besitzt die Approximationsaufgabe

$$\min\{\|A - B\|_2 : B \in \mathbb{R}^{m \times n} \text{ und } \text{Rang}(B) \leq k\}$$

für $k < r$ die Lösung

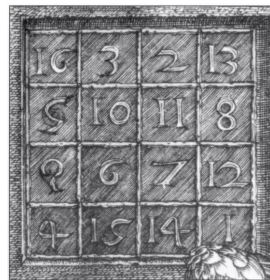
$$A_k := \sum_{i=1}^k \sigma_i \mathbf{u}_i \mathbf{v}_i^\top \quad \text{mit} \quad \|A - A_k\|_2 = \sigma_{k+1}.$$

Singulärwertzerlegung

Anwendung Datenkompression



Digitale Fotografien kann man als Matrizen aus Helligkeits- und Farbwerten auffassen. Wir betrachten als Beispiel einen Ausschnitt aus Albrecht Dürers Kupferstich Melancholie I (1514), der ein magisches Quadrat enthält.



Singulärwertzerlegung

Anwendung Datenkompression

- Die Bildinformation ist in einer Pixelmatrix A der Dimension 359×371 gespeichert, deren Einträge – ganze Zahlen zwischen 1 und 64 – verschiedene Graustufen repräsentieren.
- Wir approximieren A durch Matrizen niedrigen Rangs k (vgl. Satz 2.57). Im technischen Berechnungssystem MATLAB geschieht dies bspw. durch die Anweisungen

```
load detail.mat;  
[U,S,V] = svd(A);  
A_k = U(:,1:k)*S(1:k,1:k)*V(:,1:k)';  
image(X_k), colormap('gray'), axis('image'), axis('off')
```

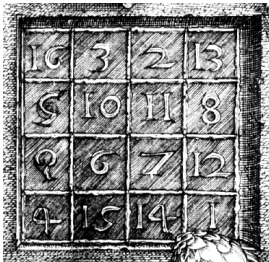
- Zur Speicherung von A_k sind im Wesentlichen

$$k(m+n) = 730k$$

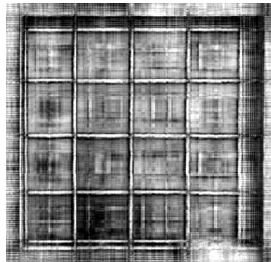
Zahlen erforderlich; die volle Matrix A erfordert

$$mn = 359 \cdot 371 = 133\,189$$

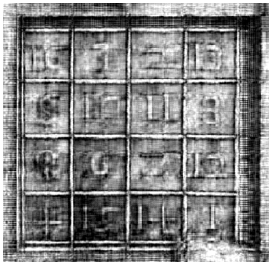
Original



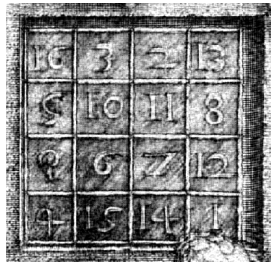
k=10



k=20



k=40



Als Maße für die Qualität der Kompression kann man den relativen Fehler

$$\frac{\|A - A_k\|_2}{\|A\|_2} = \frac{\sigma_{k+1}}{\sigma_1}$$

sowie die Kompressionsrate

$$\frac{\text{Speicherbedarf von } A_k}{\text{Speicherbedarf von } A} = \frac{k(m+n)}{mn}$$

heranziehen. In diesem Beispiel:

k	Relativer Fehler σ_{k+1}/σ_1	Kompressionsrate
10	0.0666	0.055
20	0.0528	0.110
40	0.0382	0.219

Wünschenswerte Eigenschaften von Suchmaschinen:

- Geschwindigkeit
- Möglichst viele relevante Dokumente (Webseiten). Wird als Präzision (**precision**) bezeichnet.
- Möglichst wenige relevante Dokumente fehlen. Wird als hoher Rücklauf (**recall**) bezeichnet.

Die Leistungsfähigkeit moderner Suchmaschinen beruht auch auf effizienten Algorithmen der (numerischen) linearen Algebra.

Beispiel: Google's PageRank Algorithmus beruht auf der Berechnung des Eigenvektors zum größten Eigenwert einer Matrix, deren Dimension der Anzahl der indizierten Webseiten entspricht, d.h. mehrere Milliarden. Effiziente Implementierung erfordert Techniken der Informatik (Datenhaltung, schneller Zugriff), Mathematik (Aufdatierungstechniken bei kleinen Änderungen der Matrix) und viele Heuristiken.

Wir betrachten hier eine weitere Suchtechnik: **latent semantic analysis (LSA)**.

Singulärwertzerlegung

Anwendung LSA

- Die ersten Suchmaschinen: Prüfe Dokumente auf Enthaltensein des Suchbegriffs, z.B. *Mark Twain*.
Nachteile: langsam, findet Synonyme nicht, z.B. *Samuel Clemens*.
- **Vektorraummodell** aus Termen und Dokumenten.
- Grundgesamtheit von m Begriffen (ein oder mehrere Worte) in einem Wörterbuch (**dictionary**).
- Jedes Dokument wird als Vektor im \mathbb{R}^m repräsentiert; i -te Komponente kodiert die Wichtigkeit des i -ten Termes in diesem Dokument.
- n Dokumente, als Spalten einer **Term-Dokument-Matrix** $A \in \mathbb{R}^{m \times n}$ angeordnet.
- **Beispiel:** Testdatensatz *Cranfield*, bestehend aus ca. 1400 Kurzfassungen von Texten über Luftfahrt.
<http://www.cs.utk.edu/~lsi> unter "Corpora",
<http://www.cs.umd.edu/users/oleary/SCCS/supp/lsi1>

Suchanfrage nun deutlich schneller:

1. Bestimme Index i des Suchbegriffs im Wörterbuch
2. Für jedes Dokument (1 bis n): prüfe anhand von a_{ij} ob Dokument j für Suchbegriff i relevant (relativ zu anderen Einträgen in Zeile i von A).

Schritt 1 kann effizient implementiert werden durch binäre Suche, wenn das Wörterbuch alphabetisch sortiert ist.

Schritt 2 kann auf einen Schlag geschehen durch die Matrixoperation

$$\mathbf{q}^\top A, \quad \mathbf{q} = \mathbf{e}_i.$$

Interessiert man sich für Begriff i und Begriff j , so kann man $\mathbf{q} = \mathbf{e}_i + \mathbf{e}_j$ verwenden.

- Angenommen wir stufen für Anfragevektor \mathbf{q} die ℓ Dokumente mit den größten Werten im Zeilenvektor $\mathbf{q}^\top A$ als relevant ein.
- Sei r die Anzahl der hiervon für Anfrage \mathbf{q} **tatsächlich** relevanter Dokumente.
- **precision**: $P(\ell) = r/\ell$
(niedrig, falls viele irrelevante Dokumente geliefert).
- **recall**: $R(\ell) = r/(\#\text{tatsächlich relevanter Dokumente})$
(niedrig, falls viele relevante Dokumente übersehen werden)
- Ein Graph von $P(\ell)$ und $R(\ell)$ als Funktion von ℓ illustriert wie erfolgreich relevante Dokumente gefunden werden.

Beispiel: Bei einer Dokumentensammlung aus 5 Dokumenten seien Dokumente 2,3 und 5 für eine Anfrage relevant, die Anfragebewertung liefere

$$\mathbf{q}^\top A = [6, 3, 1, 2, 10].$$

Wir erhalten

$$P(1) = 1,$$

$$R(1) = 1/3,$$

$$P(2) = 1/2,$$

$$R(2) = 1/3.$$

Singulärwertzerlegung

Anwendung LSA

- Problem: verschiedene Begriffe fügen gleiche/verwandte Konzepte. Diese werden durch Prüfen auf Auftreten nicht gefunden.
- Verschiedene Sprachen; unterschiedlicher Fachjargon (Herzanfall, Myokardinfarkt)
- Idee: Enthalten viele Dokumente die vier Begriffe *Mark*, *Twain*, *Samuel*, *Clemens*, so kann man auf eine Verwandtschaft schließen.
- **Beispiel:**

Begriff	D1	D2	D3	D4	D5
mark	20	31	0	5	2
twain	53	65	0	30	1
samuel	5	4	6	10	0
clemens	10	20	40	43	0
europe	30	10	25	52	70

Dokumente D1, D2, D3 und D4 sind relevant für die Anfrage *Mark Twain Europe*, aber auch für die Anfrage *Samuel Clemens Europe*. Letztere hätten wir mit dem reinen Vektorraummodell nicht entdeckt.

LSA:

- Ersetze A durch $A_k \in \mathbb{R}^{m \times n}$, $\text{Rang}(A_k) = k \ll \min\{m, n\}$.
- Bestapproximation an A durch Matrizen vom Rang k durch **abgeschnittene Singulärwertzerlegung**:

$$A = A_k + E, \quad \|E\|_2 = \|A - A_k\|_2 = \sigma_{k+1}$$

falls σ_j die Singulärwerte von A bezeichnen.

- Bewertungen für Anfrage q : $q^\top A_k$ anstelle von $q^\top A$, kleiner Fehler $q^\top E$.
- Idee: Approximation entfernt 'Rauschen' aus der Matrix, etwa die Variation der Bezeichnungen für dieselben Inhalte.
- k nicht zu klein, da sonst Spalten (Dokumente) zu ähnlich (zuviel Rauschen).
- k nicht zu groß, da Effekt sonst nicht eintritt (und Speicher-/Rechenaufwand zu groß).

D. P. O'Leary: What's the Score? Matrices, Documents, and Queries

Singulärwertzerlegung

Weitere Anwendungen

- Bildentzerrung (image deblurring)
- Gesichtserkennung
- Maschinenlernen
- Datenanalyse
- Signaltrennung (Cocktail Party Problem)
- Rauschunterdrückung

Ziele erreicht?

Sie sollten nun (bzw. nach Abschluss der Übungen/Selbststudium):

- wissen, was die Begriffe Vektorraum, Unterraum, Basis und lineare Unabhängigkeit bedeuten,
- diese Vorstellungen insbesondere im Fall des \mathbb{K}^n sicher anwenden können,
- sicher mit Matrizen rechnen können und den Zusammenhang zwischen Matrizen und linearen Abbildungen kennen,
- Dimension von Kern und Bild (Rang) einer Matrix (linearen Abbildung) sicher bestimmen können,
- lineare Gleichungssysteme sicher von Hand lösen können und sicher über die Anzahl der Lösungen entscheiden können,
- Determinanten moderater Größe sicher berechnen können und anhand der Ergebnisse über die eindeutige Lösbarkeit von LGS entscheiden können,
- wissen, was man unter Invertierbarkeit von Matrizen versteht und welche Beziehungen zur eindeutigen Lösbarkeit von LGS bestehen,
- die Inverse für kleinere Matrizen sicher berechnen können,

Ziele erreicht?

(Fortsetzung)

- über Skalarprodukte, Orthogonalität, Längen- und Winkelmessung bescheidwissen (vor allem im Fall \mathbb{K}^n),
- Kreuz- und Spatprodukt sicher berechnen können und über deren geometrische Interpretation bescheidwissen,
- Lagebeziehungen von Punkten, Geraden und Ebenen im Raum sicher analysieren können.
- über die Zusammenhänge zwischen Matrizen und linearen Abbildungen bescheidwissen,
- einfache geometrische Transformationen wie Drehungen und Spiegelungen mit Hilfe orthogonaler Matrizen beschreiben können,
- wissen, was orthogonale Matrizen charakterisiert,
- die Begriffe Eigenwert, Eigenvektor und Vielfachheit tiefgreifend verstanden haben,
- Eigenwerte und Eigenvektoren sicher berechnen können, ggf. auch unter Beachtung der Matrixstruktur,
- wissen, was es mit Basen von Eigenvektoren und Diagonalisierbarkeit auf sich hat.