Hochschule Konstanz
Technik, Wirtschaft und Gestaltung

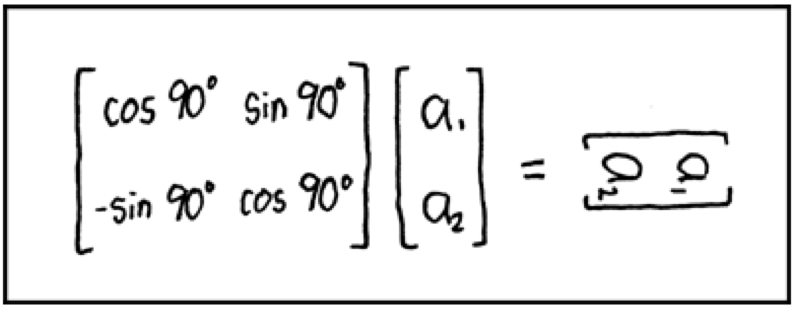
Fakultät Informatik

Mathematik

Teil I, Lineare Algebra

Skript

(R. Axthelm)



rebekka.axthelm@htwg-konstanz.de

Inhaltsverzeichnis

Elementare Algebra: zählen und sprechen 1

Wir behandeln:

- Mengen, Schreibweisen und Konventionen
- Aussagen treffen: Mathematik als Sprache

»Er hat ein paar LGS hingeschrieben, einige Theoreme eingestreut und eh ich mich's versah hab ich ihm den ganzen Beweis abgenommen ...«

In diesem Kapitel wollen wir uns dem Grundlegendsten der Mathematik widmen, nämlich den Zahlen. Zahlen als solche sind uns nichts Neues und es wird Sie überraschen, dass wir über das Rechnen mit Zahlen noch einmal nachdenken wollen. Lassen Sie es auf sich zukommen. Sie werden sicher etwas Neues lernen oder das bereits Erlernte aus einer anderen Perspektive betrachten. Des Weiteren wollen wir uns im "mathematisch" Sprechen bzw. Schreiben üben. Dabei geht es um die absolut korrekte Verwendung der mathematischen Symbole. Gerade in der Informatik ist es von besonderer Bedeutung, sich eineindeutig auszudrücken. Vieles kennen wir aus dem alltäglichen Leben aber Einiges sprechen wir im Alltag nicht korrekt.

Kind: "Bekomme ich ein Eis?"
Mutter: "Klar. Hier gebe ich dir Geld. Kaufe dir gerne ein Eis."

Nach einer Weile kommt das Kind zurück und hält zur Überraschung der Mutter zwei Eistüten in der Hand. Die Mutter schimpft, hat sie doch nur ein Eis erlaubt.

Nun. Aus Sicht der Mathematik hat das Kind keinen Fehler gemacht. "ein" ist gleichbedeutend mit "mindestens ein". Die Mutter hätte den Kauf von "genau einem" Eis erlauben müssen.

Im Alltag ist uns ja meist schon klar wie es gemeint ist aber so einem Computer eben nicht. Dem muss man es schon genau sagen. Und auch in einem mathematischen Beweis, muss jede Formulierung so gewählt werden, dass es keinen Raum für Diskussionen lässt. Es geht nicht um eine Meinung, die man vertritt. Darum ist es von elementarster Wichtigkeit, dass man erstens genau weiß was man sagen möchte und zweitens unmissverständlich genau das auch aussagt.

1.1 Schreibweisen & Operationen

Eine grundlegende Fähigkeit menschlichen Geistes ist es, Objekte zu einem Ganzen zusammenfassen zu können. So fassen wir die Einwohner von Konstanz zu einem Ganzen zusammen, indem wir sie Kontanzer nennen. Wir wollen Objekten, die wir aufgrund einer bestimmten Eigenschaft zusammenfassen einen Namen geben:

Definition 1.1 (Mengen)
Eine Menge *ist eine Zusammenfassung von bestimmten, wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen. Objekte einer Menge heißen* Elemente. *(Georg Cantor, 1845 - 1918)*

Ist a ein Element der Menge M so sagen wir "a ist Element von M" und schreiben

$$a \in M$$

und ist a hingegen kein Element von M so sagen wir "a ist nicht Element von M" und schreiben

$$a \notin M\,.$$

Wir unterscheiden zwischen Mengen in aufzählender und beschreibender Form. Zur beschreibenden Menge gehört zum Beispiel die folgende:

$$M := \{x \mid x \text{ hat die Eigenschaft } E\}$$

Wir sagen: "M ist definiert als (:=) die Menge aller Elemente x für die gilt (|), dass sie die Eigenschaft E haben."

Eine Menge aufzählender Form wäre etwa:

$$M := \{1, 2, 3, 4, 5\}$$

Wir sagen: "M ist definiert als die Menge bestehend aus den Zahlen $1, 2, 3, 4$ und 5."

Wir arbeiten auch mit Mengen, die keine Elemente enthalten. Wir sprechen dann von der *leeren Menge* und schreiben

$$\emptyset := \{\} \, .$$

N heißt *Teilmenge* von M, wenn jedes Element in N auch eine Element von M ist. Auch dafür gibt es eine Schreibweise:

$$N \subseteq M$$

Das bedeutet, dass jedes Element von N auch ein Element von M ist.

Man beachte, dass wenn N eine Teilmenge von M ist auch $N = M$ gelten kann. Soll die Gleichheit ausgeschlossen werden so sagen wir "N ist echte Teilmenge von M" und schreiben

$$N \subset M \, .$$

In anderen "Worten":

$$N \subseteq M \text{ und es gibt ein } x \in M \text{ mit } x \notin N$$

Spätestens jetzt sollten wir uns über die mathematische Bedeutung von "es gibt ein" klar werden. Im mathematischen Formalismus bedeutet "es gibt ein" stets "es gibt mindestens ein". Darf es nicht mehr geben so sagt man "es gibt genau ein".

Auch für diesen Formalismus existieren hilfreiche Ausdrucksformen, nämlich die sogenannten *Quantoren*. So heißt

$$\exists\, x \in M \mid E$$

"es existiert ein Element x der Menge M mit der Eigenschaft E". E ist hierbei nicht die Eigenschaft, die die Menge M auszeichnet. Die Menge M könnte beispielsweise die Menge aller männlichen Schweizer sein und die Eigenschaft E könnte schwarzhaarig bedeuten. Dann wäre die Aussage $\exists\, x \in M \mid E$ gleichbedeutend mit "Es gibt einen Schweizer, der schwarze Haare hat". Es kann hier, wir erinnern uns, durchaus auch bedeuten, dass mehrere oder gar alle schweizer Männer schwarze Haare haben.

Während hingegen

$$\exists!\, x \in M \mid E$$

bedeutet "es existiert genau ein Element x der Menge M mit der Eigenschaft E". Für unsere männlichen Schweizer würde es bedeuten, dass es im ganzen Land nur einen einzigen schwarzhaarigen Mann gibt. Eher unwahrscheinlich.

$$\forall x \in M \mid E$$

bezeichnet "alle Elemente x der Menge M, die die Eigenschaft E haben".

Bei der Beschreibung von Elementeigenschaften einer Menge sind die Zeichen $\vee$ für das logische ODER und $\wedge$ für das logische UND zur Steigerung der Übersicht und Darstellungseleganz äußerst nützlich.

Mengen können miteinander operieren. Man kann sie vereinen, gemeinsame Elemente bezeichnen, sie addieren und voneinander abziehen, um nur die grundlegendsten von allen möglichen Mengenoperationen zu nennen. Unter der *Vereinigung* von Mengen M und N

$$M \cup N := \{x \mid (x \in M) \vee (x \in N)\}$$

verstehen wir alle Elemente, die sowohl in M als auch in N enthalten sind, wohingegen der *Durchschnitt* oder auch kurz *Schnitt* der beiden Mengen

$$M \cap N := \{x \in M \mid x \in N\}$$

nur die Elemente beinhaltet, die in M und gleichzeitig in N enthalten sind. Ziehen wir die Menge N von der Menge M ab, sprich "M *ohne* N",

$$M \setminus N := \{x \in M \mid x \notin N\}$$

so erhalten wir alle Elemente, die in M und nicht in N sind. Unter dem Komplement der Menge M

$$M^{\complement} := \{x \mid x \notin M\}$$

verstehen wir die Menge aller Elemente, die nicht in M sind.

Beispiel 1 Äpfel und Birnen Vor Ihnen stehen zwei Obstkörbe. Der linke Korb enthält zwei Äpfel und drei Birnen. Der rechte Korb enthält zwei Äpfel und eine Orange. Was ist die Schnittmenge dieser beiden Mengen?

Die Schnittmenge ist die leere Menge. Warum?

Mengen M und N heißen *disjunkt*, wenn sie kein gemeinsames Element haben. Es gilt dann

$$M \cap N = \emptyset \,.$$

Zwei Mengen M und N heißen *gleich*, genau dann wenn

$$M \subseteq M \wedge N \subseteq M$$

gilt.

Mengen selbst können Elemente von Mengen sein. Zum Beispiel sind

$$\{\{1,2\},3\}\,, \quad \{\{\emptyset\}\}\,, \quad \{\emptyset,\{1,2\},a\}$$

Mengen im Sinne der Definition.

Definition 1.2 (Mengenschreibweisen)

Schreibweise	*Definition*	*Sprechweise*
$x \in M$		x *ist Element von* M
$x \notin M$		x *ist nicht Element von* M
	$\{x \mid x \text{ hat die Eigenschaft } E\}$	*Menge aller* x *mit der Eigenschaft* E.
$\emptyset$	$:= \{\}$	*leere Menge*
$N \subseteq M$		N *ist Teilmenge von* M
$N \subset M$		N *ist echte Teilmenge von* M
Durchschnitt:		
$M \cap N$	$:= \{x \in M \mid x \in N\}$	M *geschnitten* N
Vereinigung:		
$M \cup N$	$:= \{x \mid (x \in M) \vee (x \in N)\}$	M *vereinigt* N
Differenz:		
$M \setminus N$	$:= \{x \in M \mid x \notin N\}$	M *ohne* N
Komplement:		
$M^{\complement}$	$:= \{x \mid x \notin M\}$	*Komplement von* M

Ein wenig aufpassen müssen wir bei der Reihenfolge von Mengenoperationen. Wenn eine Klammerung vorgenommen wurde ist es klar, was gemeint ist:

$$(A \cup B) \cap C \neq A \cup (B \cap C)$$

Ohne Klammerung ist es auch klar, aber man muss wissen welcher Operator Vorfahrt hat:

$$A \cup B \cap C = A \cup (B \cap C)$$

Die Reihenfolge ist:

$$^{\complement} \quad \text{vor} \quad \cap \quad \text{vor} \quad \cup \quad \text{vor} \quad \setminus$$

Beispiel 2

(a) Es sei $A = \{1, 2, 3, 4, 5, 6\}$, $B = \{2, 4, 6\}$ und $C = \{3, 6\}$. Dann ist

$$\begin{aligned} A \cup (B \cap C) &= \{1, 2, 3, 4, 5, 6\} \cup (\{2, 4, 6\} \cap \{3, 6\}) \\ &= \{1, 2, 3, 4, 5, 6\} \cup \{6\} = \{1, 2, 3, 4, 5, 6\} = A \end{aligned}$$

und

$$\begin{aligned} (A \cup B) \cap C &= (\{1, 2, 3, 4, 5, 6\} \cup \{2, 4, 6\}) \cap \{3, 6\} \\ &= \{1, 2, 3, 4, 5, 6\} \cap \{3, 6\} = \{3, 6\} = C \end{aligned}$$

(b)
$$(A \cup B \setminus C \cap D \cup E \cap A) \setminus B = ((A \cup B) \setminus ((C \cap D) \cup (E \cap A))) \setminus B$$

Definition 1.3 (Quantoren)

	Schreibweise	Sprechweise
Existenz–		
quantor:	$\exists\, x$	*es existiert (mind.) ein x*
	$\nexists\, x$	*es existiert kein x*
	$\exists!\, x$	*es existieren genau ein x*
	$\exists_1\, x$	*es existieren genau ein x*
	$\exists_n\, x$	*es existieren genau n x*
Generalisierungs–		
quantor:	$\forall\, x$	*für alle x*
	$\vee$	*logisches ODER*
	$\wedge$	*logisches UND*

1.2 Zahlmengen

Wir wollen uns nun mit ganz speziellen Mengen befassen, nämlich den Zahlmengen. Wir haben in den voranstehenden Kapiteln schon mit ihnen gearbeitet, ganz so als wüssten wir worum es geht. Und in der Tat verhält es sich ja so, dass wir durch unsere Erfahrungen viel mit Zahlen zu tun hatten, woraufhin sich ein gewisser Gewöhnungseffekt eingestellt hat. Nichts desto trotz ist es so, dass Zahlen nicht etwa in der Natur zu beobachten sind – oder haben Sie schon einmal gesehen, wie sich eine 2 mit sich selbst multipliziert hat? – sondern ein geistiges Konstrukt des Menschen darstellen. Wir wollen in diesem Kapitel die für uns relevantesten Zahlmengen zusammenfassen:

Wir starten mit der Zahl 1. Dann nehmen wir noch eine weitere 1 hinzu und erhalten die 2. Noch eine 1 liefert die 3. Wir wollen diesen Prozess unendlich oft wiederholen und fassen aus zeitlichen Gründen alle Zahlen, die wir auf diese Weise erhalten zusammen zu einer Menge, versehen diese mit dem Symbol $\mathbb{N}$ und nennen sie die Menge der *natürlichen Zahlen*. Es gilt also
$$\mathbb{N} := \{1, 2, 3, \ldots\}\,.$$

Wollen wir auch die 0 einbeziehen so versehen wir unser Symbol mit einer entsprechenden Kennzeichnung $\mathbb{N}_0$. Es ist demnach
$$\mathbb{N}_0 := \{0, 1, 2, 3, \ldots\}\,.$$

Wir wollen auch mit negativen Zahlen rechnen, also fügen wir sinnvollerweise zu $\mathbb{N}_0$ alle natürlichen Zahlen, versehen mit einem negativen Vorzeichen hinzu und nennen die neue

Menge $\mathbb{Z}$. Das ist die Menge der *ganzen Zahlen* und hat die Darstellung

$$\mathbb{Z} := \{\ldots, -2, -1, 0, 1, 2, \ldots\}.$$

Bilden wir nun Brüche (*Zähler* durch *Nenner*) aus je zwei ganzen Zahlen, wobei wir peinlichst darauf achten, dass der Nenner immer ungleich Null ist, so erhalten wir die Menge der *rationalen Zahlen*, die wir mit $\mathbb{Q}$ bezeichnen und die mathematische Definition

$$\mathbb{Q} := \left\{ \frac{m}{n} \,\middle|\, m, n \in \mathbb{Z} \ \wedge \ n \neq 0 \right\}$$

besitzt. Für $m, n \in \mathbb{Z}$ mit $n \neq 0$ heißt der Bruch

$$\frac{m}{n} \in \mathbb{Q}$$

gekürzt oder auch *reduziert*, wenn m und n keinen gemeinsamen *nichttrivialen* Teiler haben. Ein trivialer Teiler wäre die 1.

Es gab also zunächst natürliche, ganze und rationale Zahlen. Bis man entdeckte, dass das nicht ausreicht[1]. Die Diagonale im Quadrat der Seitenlänge 1 zum Beispiel kann nicht als rationale Zahl geschrieben werden. Ich nehme die Spannung raus und verrate schon mal, dass diese Diagonale die Länge $\sqrt{2}$ hat. Aus der Schule wissen wir, dass $\sqrt{2}$ diejenige Zahl ist, die mit sich selbst multipliziert gerade die 2 ergibt. Aber können wir sie konkret darstellen? Wir kennen einen genäherten Wert $\sqrt{2} \approx 1.414213562\ldots$, aber vollständig beschrieben ist sie damit nicht. Wir werden uns in Kapitel 1.3 (Satz 1.10) davon überzeugen, dass es eine "vollständigere" Darstellung als die, die wir bereits verwendet haben, nämlich $\sqrt{2}$, nicht gibt.

Wir nennen alle Zahlen auf der Zahlengeraden die *reellen Zahlen* und bezeichnen diese mit dem Symbol $\mathbb{R}$. $\sqrt{2}$ zum Beispiel ist sicher eine Zahl, die sich in dieser Menge befindet, aber nicht in $\mathbb{Q}$ enthalten ist. Damit gilt schon mal, dass

$$\mathbb{R} \setminus \mathbb{Q} \neq \emptyset.$$

Wir gehen einmal davon aus, dass es noch mehr Zahlen gibt, die nicht rational sind, so dass es sich lohnt dieser speziellen Menge einen Namen zu geben. Sie hat kein eigenes Symbol,

[1] Die Definitionen der Zahlenmengen haben historische Gründe:

> *Die antiken Griechen, die Schöpfer der beweisenden Mathematik, vertraten unter dem Einfluß der pythagoreischen Schule lange die waghalsige Lehre, dass die Welt sich durch natürliche Zahlen und deren Verhältnis – also die rationalen Zahlen – beschreiben lasse. Die Entdeckung, dass man jedoch die Diagonale eines Quadrats so nicht beschreiben kann war für die Griechen ein tiefer Schock; den frevelhaften Entdecker dieser Ungeheuerlichkeit – pikanterweise ein Mitglied der pythagoreischen Schule selbst – sollen seine pythagoreischen Genossen denn auch zur Strafe während einer Seefahrt ins Meer geworfen haben.*

aber wir nennen sie die Menge der *irrationalen Zahlen*. Weitere Beispiele für irrationale Zahlen sind etwa

$\mathrm{e} = 2,71828...$	*Eulersche Zahl*
$\pi = 3,14159...$	*Kreiszahl* (sprich "pi")

Für einen Bereich der reellen Zahlen als Teilmenge, etwa

$$\{x \in \mathbb{R} \mid 0 \leq x \leq 1\} \subset \mathbb{R}\,,$$

verwenden wir einen speziellen Begriff, nämlich:

Definition 1.4 (Intervalle)

Schreibweise	*Definition*	*Beschreibung*
$[a,b]$	$:= \{x \in \mathbb{R} \mid a \leq x \leq b\} \subset \mathbb{R}$	*abgeschlossenes Intervall*
(a,b)	$:= \{x \in \mathbb{R} \mid a < x < b\} \subseteq \mathbb{R}$	*offenes Intervall*
$[a,b)$	$:= \{x \in \mathbb{R} \mid a \leq x < b\} \subset \mathbb{R}$	*(halb–)offenes Intervall*
$(a,b]$	$:= \{x \in \mathbb{R} \mid a < x \leq b\} \subset \mathbb{R}$	*(halb–)offenes Intervall*

Die Unterscheidung $\subset$ und $\subseteq$ beim offenen Intervall (a,b) hat damit zu tun, dass bei unbeschränkten Intervallgrenzen $a = -\infty$ und $b = \infty$

$$(a,b) = \mathbb{R}$$

gilt und $\mathbb{R}$ keine abgeschlossene Menge ist. Solcherlei mathematische Hintergründe interessieren uns aber nicht weiter. Wir vermerken es an dieser Stelle nur der Vollständigkeit wegen, da es immer mal wieder auftaucht und wir es einfach korrekt hinschreiben wollen.

Seien $a, b \in \mathbb{R}$ mit $-\infty < a, b < \infty$.

$\cancel{[a,\infty]}$ sondern immer $[a,\infty)$

genauso:

$\cancel{[-\infty,b]}$ sondern immer $(-\infty,b]$

Eine weitere übliche Schreibweise für offene Intervalle (a, b) ist durch

$$]a, b[$$

gegeben.

Wir wollen noch eine kleine Notation einführen, die es uns erlaubt zwischen positiven und negativen, sagen wir reellen, Zahlen zu unterscheiden. Sprechen wir von *positiven reellen Zahlen* so meinen wir Zahlen, die größer als Null sind und bezeichnen diese mit $\mathbb{R}^+$. Es ist also

$$\mathbb{R}^+ := \{x \in \mathbb{R} \mid x > 0\} .$$

Sprechen wir hingegen von *nicht negativen reellen Zahlen* so meinen wir Zahlen, die nicht kleiner als Null sind und bezeichnen sie mit $\mathbb{R}_0^+$. Sehen Sie den Unterschied? Es ist nämlich

$$\mathbb{R}_0^+ := \{x \in \mathbb{R} \mid x \geq 0\} .$$

$\mathbb{R}^-$ und $\mathbb{R}_0^-$ definieren wir analog.

Beispiel 3 weitere Zahlmengen

Name	Schreibweise	Definition
positive reelle Zahlen:	$\mathbb{R}^+$	$:= \{x \in \mathbb{R} \mid x > 0\}$
nicht negative reelle Zahlen:	$\mathbb{R}_0^+$	$:= \{x \in \mathbb{R} \mid x \geq 0\}$
negative reelle Zahlen:	$\mathbb{R}^-$	$:= \{x \in \mathbb{R} \mid x < 0\}$
nicht positive reelle Zahlen:	$\mathbb{R}_0^-$	$:= \{x \in \mathbb{R} \mid x \leq 0\}$

Wir können auch das Produkt von Mengen definieren. Sie dürfen sich allerdings darunter nicht vorstellen, dass je zwei Elemente aus zwei Mengen miteinander multipliziert werden! Wie soll das auch funktionieren, wenn die eine Menge aus Äpfeln und die andere aus Birnen besteht? Man sollte sich viel mehr vorstellen, dass das Produkt von Mengen eine Menge von Paaren, gebildet aus den Elementen der ursprünglichen Mengen, besteht. Wir definieren:

Definition 1.5 (Produktmenge) *Es seien M und N Mengen.*

Die Menge der geordneten Paare

$$M \times N := \{(a, b) \mid (a \in M) \wedge (b \in N)\}$$

heißt Produktmenge oder auch kartesisches Produkt .
Da eine Produktmenge ihrerseits eine Menge darstellt kann der Prozess sukkzessive fortgesetzt und Produktmengen von mehr als zwei Mengen gebildet werden. Seien M_1,..,M_n Mengen, dann ist

$$M_1 \times \cdots \times M_n := \{(a_1, \ldots, a_n) \mid a_i \in M_i, i = 1, \ldots, n\}$$

ebenfalls eine Produktmenge.

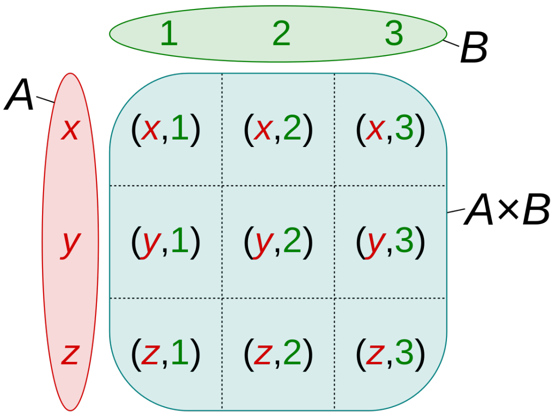
Notation & Sprechweisen:

$$\underbrace{M \times \cdots \times M}_{n \text{ mal}} =: M^n$$

Speziell heißen Elemente des

$$\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$$

n-Tupel . 2-Tupel des $\mathbb{R}^2$ nennen wir auch einfach Tupel.



Quelle: Wikipedia

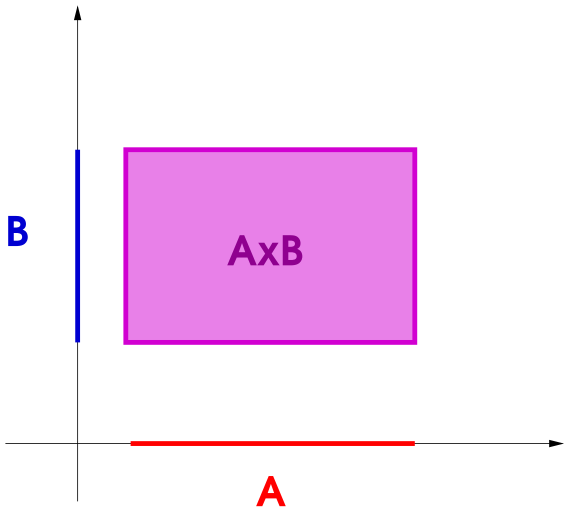
Zwei Elemente (c, d) und (e, f) aus $M \times N$ heißen *gleich* wenn $c = e$ und $d = f$ gilt. Wir schreiben das mal mathematisch:

Für (c, d), $(e, f) \in M \times N$ gilt: $(c, d) = (e, f) \;\Leftrightarrow\; c = e \wedge d = f$.

Beispiel 4 Produktmengen

Euklidsche Ebene

$$\begin{aligned}\mathbb{R}^2 &= \mathbb{R} \times \mathbb{R} \\ &= \{(x, y) \mid x \in \mathbb{R} \land y \in \mathbb{R}\}\end{aligned}$$



Der Raum unserer Anschauung

$$\begin{aligned}\mathbb{R}^3 &= \mathbb{R} \times \mathbb{R} \times \mathbb{R} \\ &= \{(x, y, z) \mid x, y, z \in \mathbb{R}\}\end{aligned}$$

Einheitswürfel im $\mathbb{R}^3$

$$[0, 1]^3 = \{(x, y, z) \in \mathbb{R}^3 \mid 0 \leq x \leq 1 \land 0 \leq y \leq 1 \land 0 \leq z \leq 1\}$$

Einheitswürfel im $\mathbb{R}^3$ ohne Rand

$$(0, 1)^3 = \{(x, y, z) \in \mathbb{R}^3 \mid 0 < x < 1 \land 0 < y < 1 \land 0 < z < 1\}$$

Die Menge aller Eckpunkte des Quadrats $[-1, 1]^2 \subseteq \mathbb{R}^2$

$$\{-1, 1\}^3 = \{(-1, -1), (-1, 1), (1, -1), (1, 1)\}$$

Bisher haben Sie nur Zahlen gesehen, die sie bereits kennen. Damit es Ihnen nicht langweilig wird sollen Sie an dieser Stelle schon mal eine neue Art von Zahlen kennenlernen dürfen. Es handelt sich dabei um die sogenannten Komplexen Zahlen. Keine Sorge. Der Name ist nicht Programm. Es ist ganz einfach. Zunächst stellen wir uns die Frage, wieso die Zahlen, die wir bisher gesehen haben nicht ausreichend sind. Ein einfaches Beispiel lässt die Dringlichkeit erahnen: Wir suchen eine Zahl x, die folgende Gleichung erfüllt:

$$x^2 = -1$$

Wir sind uns sicher einig, dass das jetzt erst mal nicht ganz klar ist. Die Idee der Komplexen Zahlen steckt im Grunde darin, eine solche Gleichung eben auch nach x auflösen zu können. Dazu braucht man eine Zahl, die genau diese Bedingung erfüllt. Wir nennen diese Zahl *imaginäre Zahl* und bezeichnen sie intuitiver Weise mit i. i kann nicht weiter benannt werden. Nur so, dass i^2 eben gerade -1 ist[2]. Das ist schon alles. Damit konstruiert man eine neue Zahlmenge:

Definition 1.6 (Komplexe Zahlen) *Die Zahlmenge* $\mathbb{C}$ *mit*

$$\mathbb{C} = \{x + i\,y \mid x, y \in \mathbb{R} \ \wedge \ i^2 = -1\}$$

heißt Menge der Komplexen Zahlen. Für eine komplexe Zahl $c = (x + iy) \in \mathbb{C}$ *heißt*

$$\begin{aligned} \mathrm{Re}(c) &:= x \ \textit{Realteil}, \\ \mathrm{Im}(c) &:= y \ \textit{Imaginärteil und} \\ \bar{c} &:= x - iy \ \textit{konjugiert Komplexes von } c. \end{aligned}$$

Wann immer man "Neues" erklärt muss man auch klären wie man mit diesem Objekt rechnen kann. Wir sollten also für diese neue Zahlmenge Rechenregeln erklären. Das ist nicht schwer. Die komplexen Zahlen basieren auf reelle Zahlen, deren Rechenregeln hier induzieren. Das ist auch gut so, denn im Grunde sind reelle Zahlen auch komplexe Zahlen - mit Imaginärteil gleich Null - und somit sollten die entsprechenden Regeln dann auch konsistent sein.

Rechenregeln für komplexe Zahlen:
Es seien $z, w \in \mathbb{C}$ mit $z = a + i\,b$ und $w = c + i\,d$.

Addition/Subtraktion:

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

Multiplikation:

$$z \cdot w = (a + i\,b)(c + i\,d) = (ac - bd) + i\,(bc + da)$$

konjugiert Komplexe:

$$\overline{z} = a - i\,b$$

Betrag:

$$|z| = \sqrt{z\overline{z}} = \sqrt{a^2 + b^2}$$

Division:

$$\frac{z}{w} = \frac{z\overline{w}}{|w|^2} = \frac{ac + bd}{|w|^2} + i\frac{bc - da}{|w^2|}$$

[2] Das ist vielleicht verwirrend aber im Grunde das Gleiche wie mit $\sqrt{2}$. $\sqrt{2}$ können wir auch nicht genauer benennen, außer, dass $\sqrt{2}^2 = 2$, oder? Daran haben wir uns halt schon gewöhnt.

Jede Operation mit komplexen Zahlen liefert wieder eine Zahl, die sich in Real- und Imaginärteil zerlegen läßt. Diese Darstellung nennen wir kartesische Darstellung . Wir werden später noch weitere Darstellungsformen von komplexen Zahlen kennenlernen.

Beispiel 5 Für $z = 1 + i$ und $w = -2 + 3\,i$ gilt

(a)

$$z + w = (1 + i) + (-2 + 3\,i) = -1 + 4\,i\,,$$

also

$$\mathrm{Re}(z + w) = -1\,,\ \mathrm{Im}(z + w) = 4\,.$$

(b)

$$z - w = (1 + i) - (-2 + 3\,i) = 3 - 2\,i\,,$$

also

$$\mathrm{Re}(z - w) = 3\,,\ \mathrm{Im}(z - w) = -2\,.$$

(c)

$$z \cdot w = (1 + i)(-2 + 3\,i) = -2 + 3\,i^2 + i = -5 + i\,,$$

also

$$\mathrm{Re}(z \cdot w) = -5\,,\ \mathrm{Im}(z \cdot w) = 1\,.$$

(d)

$$\frac{z}{w} = \frac{z \cdot \overline{w}}{w \cdot \overline{w}} = \frac{(1 + i)(-2 - 3\,i)}{(-2 + 3\,i)(-2 - 3\,i)} = \frac{1 - 5\,i}{13} = \frac{1}{13} + \frac{-5}{13}\,i\,,$$

also

$$\mathrm{Re}\left(\frac{z}{w}\right) = \frac{1}{13}\,,\ \mathrm{Im}\left(\frac{z}{w}\right) = \frac{-5}{13}\,.$$

(e)

$$\mathrm{Re}(z) = \frac{1}{2}(z + \overline{z})$$
$$\mathrm{Im}(z) = \frac{-i}{2}(z - \overline{z})$$

1.3 Rasiermesserscharfe Logik und Formulierung von Aussagen

Die Mathematik ist eine Sprache, mit der wir Aussagen treffen und Gegebenheiten beschreiben können und das auf - jaja - möglichst einfache Art. Die Mathematik hat den Anspruch einfach zu sein! Also eine Sprache, ja? Die Zahlen sind wie Wörter im Sprachgebrauch. Ganze Zahlmengen umfassen dann quasi einen Teil des Wörterbuchs. Mit Wörtern alleine lassen sich jedoch noch keine Aussagen treffen. Wir müssen Satzkonstrukte schaffen, Wörter miteinander kombinieren. Die Grammatik gibt uns dabei ein Regelwerk zur Hand, damit nicht sowas rauskommt wie "gehe trank Kino Bier gestern ich danach". Gut, an sich haben wir eine grobe Vorstellung wie der Abend verlaufen ist, aber das liegt an unserem Erfahrungsschatz, der dem Gehirn Interpolationen erlaubt. Der Ausdruck selbst liefert keinerlei Information. Beim Aneinanderreihen von Wörtern ergibt sich noch keine Aussage. Halten wir uns an die Regeln der Grammatik wird aus dem obigen Ausdruck sowas wie "Gestern ging ich in's Kino und trank danach ein Bier." Im Grunde genügt auch alleine die Grammatik nicht, denn der Satz "Heute ist es dunkler als kalt." ist grammatikalisch ein korrekter Satz, beinhaltet aber keinerlei Wahrheitsgehalt. In der Mathematik ist ein Ausdruck erst dann eine Aussage, wenn ihm eindeutig ein Wahrheitsgehalt "wahr" oder "falsch" zugeordnet werden kann.

Wir behandeln also das Verständnis logischer Schlussfolgerungen, die in der Aussagenlogik definiert ist. Wir benötigen dieses Werkzeug, um klar, unmissverständlich und eindeutig Dinge mathematisch beschreiben und mathematische Aussagen treffen zu können. Es wird Sie vielleicht beeindrucken, dass die komplette mathematische Grammatik auf einer kleinen überschaubaren Tabelle basiert.

Definition 1.7 (Aussage)

Eine Aussage *ist ein sinnvolles sprachliches Gebilde, das entweder wahr oder falsch ist.*

Etwas anderes als wahr oder falsch wird nicht zugelassen. Aussagen kann man miteinander verknüpfen. Die Aussage dieser Verknüpfungen werden in Wahrheitstabellen definiert. Bevor wir uns eine solche Tabelle anschauen führen wir noch ein paar sprachliche Elemente ein. Es seien Aussagen A und B gegeben.

Definition 1.8 (Verknüpfungen von Aussagen)

	Schreibweise	*Sprechweise*
Negation:	$\neg A$	*nicht* A
Konjunktion:	$A \wedge B$	A *und* B
Alternative:	$A \vee B$	A *oder* B
Implikation:	$A \Rightarrow B$	*aus* A *folgt* B
Äquivalenz:	$A \Leftrightarrow B$	A *ist äquivalent zu* B

Die *Negation* einer Aussage A beinhaltet die "minimalste" Störung, die dazu führt, die wahre Aussage A falsch, bzw die falsche Aussage A wahr zu machen. Nehmen wir ein Beispiel: Wir treffen die Aussage: "Alle Autos in Konstanz sind rot." Die Negation dieser Aussage ist dann "Es gibt ein Auto in Konstanz, das nicht rot ist."

Definition 1.9 (Wahrheitstafel/-tabelle)

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \Rightarrow B$	$A \Leftrightarrow B$
W	W	F	W	W	W	W
W	F	F	F	W	F	F
F	F	W	F	F	W	W
F	W	W	F	W	W	F

Aus der Wahrheitstabelle können wir logische Regeln erstellen. Hier sind ein paar Beispiele aufgelistet (Die ersten drei sollten Sie sich genauer anschauen. Die verwenden wir später für Satz 1.10 und Hilfssatz 1.11 noch.):

$$
\begin{aligned}
(A \Rightarrow B) &= (\neg A \vee B) = (\neg B \Rightarrow \neg A) && \text{Ersetzen der Implikation}\\
\neg(A \vee B) &= \neg A \wedge \neg B && \text{de Morgansche Regeln}\\
\neg(A \Rightarrow B) &= A \wedge \neg B && \text{Verneinung der Implikation}\\
\\
\neg\neg A &= A && \text{Doppelte Verneinung}\\
(A \Leftrightarrow B) &= ((A \Rightarrow B) \wedge (B \Rightarrow A)) &&\\
&= \neg(\neg A \Leftrightarrow B) && \text{Ersetzen der Äquivalenzrelation}\\
\neg(A \wedge B) &= \neg A \vee \neg B && \text{de Morgansche Regeln}\\
\neg(A \Leftrightarrow B) &= (A \wedge \neg B) \vee (\neg A \wedge B) && \text{Verneinung der Äquivalenzrelation}
\end{aligned}
$$

Die Aussagenlogik ist unentbehrlich bei der mathematischen Beweisführung. Eine Aussage der Form

A:="Es gibt (mindestens) ein Auto in Konstanz, das rot ist."

ist auf direktem Weg meist schwer zu beweisen. Gewöhnlich geht man dabei so vor, dass man die Negation davon annimmt und dies zu einem Widerspruch führt. Man nimmt also an es gelte $\neg A$:

$\neg A$ = "Alle Autos in Konstanz sind rot."

Wenn man diese Annahme zu einem Widerspruch führen kann, was meist dadurch erreicht wird, dass man einfach ein Gegenbeispiel findet, so folgt, dass $\neg A$ falsch ist, was direkt dazu führt, dass die Aussage A wahr ist. In diesem Fall wird sich ein Gegenbeispiel finden lassen, indem man einfach aus dem Fenster schaut.

Übrigens: Jede Aussage über die leere Menge ist wahr!

Mit dem Vermögen, Aussagen treffen zu können, wollen wir uns nun der Technik der Beweisführung widmen. Denn eine Aussage alleine ist nichts als Geschwätz, von dem wir nicht wissen, ob es wahr oder falsch ist. Beweisführungen sind im Grunde Methoden von Schlussfolgerungen, die uns basierend auf der Wahrheitstafel 1.9 sagen, ob die behauptete Aussage den Wahrheitsgehalt "wahr" hat. Nichts weiter. Ganz einfach. Wir werden das Thema der Beweisführung an dieser Stelle allerdings nicht überstrapazieren, da es im Laufe der Kapitel ja permanent auftaucht und in folge dessen auch immer wieder geübt wird.

Beweismethoden:

1. direkter Beweis
2. indirekter Beweis
3. vollständige Induktion

Beim *direkten Beweis* verfährt man "straight forward". Der *indirekte Beweis* funktioniert so, dass man zunächst die Negation der Aussage annimmt und diese dann zu einem Widerspruch führt, woraufhin die ursprüngliche Aussage Gültigkeit erhällt. Der dritten Beweismethode, dem sogenannte *Beweis durch Vollständige Induktion* widmen wir das folgende Unterkapitel, Kapitel 1.4 und werden auch erst dort genauer darauf eingehen.

Wir betrachten die Zahl $\sqrt{2}$ und wollen uns davon überzeugen, dass diese nicht rational ist, d.h. dass sie nicht durch eine rationale Zahl beschrieben werden kann. Das wollen wir mittels indirektem Beweis zeigen und formulieren den

Satz 1.10 *Für p, q mit $q \neq 0$ gilt:*

$$\left(\sqrt{2} = \frac{p}{q}\right) \Rightarrow (p \notin \mathbb{Z} \vee q \notin \mathbb{Z})$$

Um diesen Satz zu beweisen, benötigen wir folgende Aussagen, die wir dehalb vorweg formulieren, damit wir sie dann im Beweis verwenden können.

Hilfssatz 1.11 *Für $n \in \mathbb{Z}$ gilt:*

1. *Ist n gerade so ist n^2 ebenfalls gerade.*
2. *Ist n ungerade so ist n^2 ebenfalls ungerade.*

Wir wollen uns davon überzeugen:

Beweis Hilfssatz 1.11:

Zu 1.: Was wir zeigen wollen hat die folgende Aussagenstruktur. Sei $n \in \mathbb{Z}$*. Mit*

$$\begin{aligned} A &:= \text{“}n \textit{ ist gerade”} \textit{ und} \\ B &:= \text{“}n^2 \textit{ ist gerade”} \end{aligned}$$

gilt

$$A \Rightarrow B\,.$$

Das ist so einfach, dass wir den Beweis ohne Umstände direkt durchführen können. Wir nennen ihn deshalb “direkten Beweis”. Eine gerade Zahl n *kann dargestellt werden als* $n = 2k,\ k \in \mathbb{Z}$*. Dann ist*

$$(2k)^2 = 4k^2 = 2(2k^2)$$

selbstverständlich wieder eine gerade Zahl, da sie immer noch durch zwei teilbar ist.

Zu 2.: Diese Aussage beweisen wir analog zur ersten: Sei $k = 2n + 1$*. Dann ist*

$$(2k+1)^2 = 4k^2 + 4k + 1 = \underbrace{2(k^2 + 2k)}_{\textit{gerade Zahl}} + 1$$

eine ungerade Zahl, womit beide Aussagen bewiesen sind.

□

Da aus n gerade folgt, dass n^2 gerade ist, gilt in äquivalenter Weise, dass wenn n^2 ungerade folglich n auch ungerade ist. Das Analoge gilt für den zweiten, ungeraden Fall. So dass wir insgesamt die Aussagen

$$n \text{ gerade} \quad \Leftrightarrow \quad n^2 \text{ gerade}$$

und

$$n \text{ ungerade} \quad \Leftrightarrow \quad n^2 \text{ ungerade}$$

bewiesen haben.

Nun kommen wir zum Beweis des Satzes.

Beweis Satz 1.10:

Eine übliche Vorgehensweise bei der Beweisführung ist der sogenannte “indirekte Beweis”. Man nimmt einfach die Negation an und führt dies zu einem Widerspruch.

Wie sieht das formal aus? Wir haben zwei Aussagen:

$$\begin{aligned} A &:= \left(\sqrt{2} = \frac{p}{q}\right) \\ B &:= (p \notin \mathbb{Z} \vee q \notin \mathbb{Z}) \end{aligned}$$

unter der Voraussetzung (!), dass $q \neq 0$ *wollen wir foldende Behauptung zeigen:*

$$A \Rightarrow B \tag{1}$$

Wir nehmen die Negation an mit

$$\neg(A \Rightarrow B) \quad \Leftrightarrow \quad A \wedge \neg B\,.$$

Also Annahme: Es gilt

$$A \wedge \neg B. \tag{2}$$

Schauen wir zunächst mal was $\neg B$ *genau ist:*

$$\begin{aligned}\neg B &= \neg\,(p \notin \mathbb{Z} \vee q \notin \mathbb{Z}) \\ &= \neg\,(p \notin \mathbb{Z}) \wedge \neg\,(q \notin \mathbb{Z}) \\ &= (p \in \mathbb{Z}) \wedge (q \in \mathbb{Z}) \\ &= p, q \in \mathbb{Z}\end{aligned}$$

Gleichung (2) kann ersetzt werden durch

$$\left(\sqrt{2} = \frac{p}{q}\right) \wedge p, q \in \mathbb{Z} \tag{3}$$

Können wir die Aussage in (3) zu einem Widerspruch führen so gilt die ursprüngliche, zu beweisende Aussage (1).

Wir nehmen also an, dass es eine rationale Zahl gibt, die Wurzel aus zwei darstellen kann. Die rationale Zahl sei so dargestellt, dass $\frac{p}{q} \in \mathbb{Z}, q \neq 0$ *gekürzt ist. Das kann man ja immer machen.*

$$\begin{aligned} && \sqrt{2} &= \frac{p}{q} \\ \Leftrightarrow && 2 &= \frac{p^2}{q^2} \\ \Leftrightarrow && 2q^2 &= p^2 \end{aligned}$$

Dies bedeutet, dass p^2 *eine gerade Zahl ist. Mit der Folgerung aus Hilfssatz 1.11 gilt auch dass* p *eine gerade Zahl ist. Wir definieren* $r := \frac{p}{2}$*, was wieder eine ganze Zahl sein muss und setzen diese für* p *ein. Dann gilt weiter*

$$\begin{aligned} \Leftrightarrow && 2q^2 &= 4r^2 \\ \Leftrightarrow && q^2 &= 2r^2 \end{aligned}$$

was bedeutet, dass auch q *eine gerade Zahl ist. Dies steht im Widerspruch zur Annahme, dass* p *und* q *keinen gemeinsamen Teiler haben. Wir haben unsere Rechnungen also erfolgreich auf einen Widerspruch geführt, was die Schlußfolgerung erlaubt, dass die negierte Aussage falsch und demzufolge die eigentliche Aussage (1) wahr ist.*

$\square$

1.4 Vollständige Induktion

Wir haben bereits die Beweismethoden "direkter Beweis" und "indirekter Beweis" kennengelernt. Nun wollen wir uns der Beweismethode der *Vollständigen Induktion* widmen. Das Grundprinzip ist ganz einfach: Denken Sie an Dominosteine. Sie wissen, wenn ein Stein umgestoßen wird so wird dieser den nächsten Stein ebenfalls zu Fall bringen. Das ist zunächst ein reines Wissen, ohne, dass ein Stein wirklich fällt. Dies nennt man den *Indutionsschritt*. Wenn Sie nun einen Stein wirklich antippen und zu Fall bringen, so ist Ihnen aufgrund der Kettenreaktion, bzw. des Induktionsschrittes klar, dass alle Steine umfallen werden. Das Starten dieses Prozesses nennen wir *Induktionsanfang*.

In anderen Worten: Wenn die Aussage $A(1)$ gilt und wir gezeigt haben, dass $A(k) \Rightarrow A(k{+}1)$ gilt und zwar für alle $k \in \mathbb{N}$, so folgt, dass auch $A(2)$ gilt, denn wir haben gezeigt, dass $A(1) \Rightarrow A(2)$ gilt. Wenn $A(2)$ gilt, so folgt, dass auch $A(3)$ gilt und so weiter.

Zeile	$A(k)$	$A(k+1)$	$A(k) \Rightarrow A(k+1)$
1	W	W	W
2	W	F	F
3	F	W	W
4	F	F	W

Der Beweis besteht im Grunde aus zwei Teilen. Erinnern Sie sich zunächst mal an das profane Beispiel "Es regnet, folglich ist die Straße nass". Wenn Sie

Es regnet $\Rightarrow$ die Straße ist nass

zeigen können sind Sie bei der obigen Wahrheitstabelle in einer der Situation von Zeile 1, 3 oder 4. Wenn Sie also diese Folgerung bewiesen haben, haben Sie noch lange nicht bewiesen, dass es regnet, auch nicht dass die Straße nass ist. Sie haben lediglich gezeigt, dass für den Fall, dass es regnet die Straße dann nass sein wird. Wenn Sie also eigentlich zeigen wollen, dass die Straße nass ist, müssen Sie noch zusätzlich beweisen, dass es regnet. Dann sind Sie in der Wahrheitstabelle in der Zeile 1 gelandet! Et voilà!

Will man zeigen, dass eine Aussage $A(n)$ für alle natürlichen Zahlen $n \in \mathbb{N}$ richtig ist, muss man beweisen, dass die Aussage für einen Startwert, etwa $n = 1$ richtig ist und dass man von $n = k$ auf $n = k + 1$ schließen kann. (Die Aussage $A(n)$ könnte zum Beispiel sein: Der Dominostein mit Nummer n fällt.)

Wir formulieren das ordentlich:

Definition 1.12 (Vollständige Induktion) *$\forall n \in \mathbb{N}$ sei $A(n)$ ist eine Aussage über $n \in \mathbb{N}$. Wenn gilt:*

$$\begin{array}{rll} A(1) & \textit{ist wahr und} & \text{(Induktionsvoraussetzung)} \\ A(k) \Rightarrow A(k+1) & \textit{ist wahr,} & \text{(Induktionsschritt)} \end{array}$$

dann gilt: $\forall n \in \mathbb{N}\ :\ A(n)$ ist eine wahre Aussage.

Beispiel 6 durch 3 teilbar

Behauptung

$$\forall n \in \mathbb{N}\ :\ n^3 - n \quad \text{ist durch 3 teilbar}$$

Induktions-anfang

$$A(1)\ :\ 1^3 - 1 = 0 \quad \text{ist durch 3 teilbar.} \qquad \checkmark$$

Induktions-voraussetzung Es gelte $A(k)$ mit

$$A(k)\ :\ k^3 - k \quad \text{ist durch 3 teilbar}$$

Induktions-schritt $A(k) \Rightarrow A(k+1)$ mit

$$A(k+1)\ :\ (k+1)^3 - (k+1) \quad \text{ist durch 3 teilbar}$$

Induktions-beweis

$$\begin{aligned} (k+1)^3 - (k+1) &= k^3 + 3\,k^2 + 3\,k + 1 - k - 1 \\ &= \underbrace{(k^3 - k)}_{\text{dirch 3 teilb., (IV)}} + \underbrace{3\,(k^2 + k)}_{\text{durch 3 teilb.}} \end{aligned}$$

□

Beispiel 7 Vom "kleinen Gauß"

Wir wollen zeigen, dass

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$$

gilt; und das per Vollständiger Induktion (klar). Also:

Beweis durch Vollständige Induktion:

Behauptung $\forall n \in \mathbb{N}\ :$

$$1 + \cdots + n = \frac{n(n+1)}{2}$$

Induktions-anfang Die Aussage gilt für $n = 1$:

$$1 = \frac{1(1+1)}{2} \qquad \checkmark$$

Induktions–voraussetzung Die Behauptung gelte für $n = k$:

$$A(k): \qquad 1 + 2 + 3 + \cdots + k = \frac{k(k+1)}{2} \tag{4}$$

Induktions–schritt

$$A(k) \Rightarrow A(k+1)$$

mit

$$A(k+1): \qquad 1 + 2 + 3 + \cdots + k + (k+1) = \frac{(k+1)(k+2)}{2} \tag{5}$$

Induktions–beweis Hier geht man typischerweise so vor, dass man mit der linken Seite der Aussage $A(k+1)$ startet, so umformuliert, so dass man die Gleichung (4) einsetzen kann. Umformulierungen führen dann auf die rechte Seite in Gleichung (5), womit die Gleichheit in (5) unter der Annahme, dass $A(k)$ wahr ist, gilt.

$$\begin{aligned}
1 + 2 + 3 + \cdots + k + (k+1) &= 1 + 2 + 3 + \cdots + k + (k+1) \\
&= \frac{k(k+1)}{2} + (k+1) \\
&= \frac{k(k+1) + 2(k+1)}{2} \\
&= \frac{(k+1)(k+2))}{2}
\end{aligned}$$

□

Struktur der vollständigen Induktion:

Aussage:

Voraussetzungen

Behauptung: $A(n)$ gilt für alle $n \in \mathbb{N}$

Beweis: (vollständige Induktion)

Induktionsanfang: Prüfe $A(1)$!

Induktionsannahme: Gelte $A(k)$ für ein $k \in \mathbb{N}$.

Induktionsschritt: Zeige: $A(k) \Rightarrow A(k+1)$

Beispiel 8 Für die Aussage (die offensichtlich falsch ist)

$$1 + 2 + \cdots + n < \frac{n(n+1)}{2}$$

läßt sich der Induktionsschritt $A(k) \Rightarrow A(k+1)$ auch zeigen. Aber $A(1)$ ist nicht wahr.

Die Aussagenkette muss nicht bei $n = 1$ beginnen. Wenn $A(1)$ nicht wahr ist, aber $A(2)$ und der Induktionsschritt gezeigt werden kann, so gilt die Aussage eben für alle $n \geq 2$.

Wir wollen noch ein Beispiel rechnen

Beispiel 9 Bernoullische Ungleichung

Behauptung $\forall n \in \mathbb{N}\ \forall x \geq -1$:

$$(1+x)^n \geq 1 + nx$$

Induktions–anfang

$$(1+x) \geq 1 + x$$

Induktions–voraussetzung $k \in \mathbb{N} \ \wedge \ x \geq -1$:

$$A(k): \qquad (1+x)^k \geq 1 + kx$$

Induktions–schritt $A(k) \Rightarrow A(k+1)$ mit

$$A(k+1): \qquad (1+x)^{k+1} \geq 1 + (k+1)x$$

Induktions–beweis

$$\begin{aligned} (1+x)^{k+1} &= (1+x)^k(1+x) \\ &\geq (1+kx)(1+x) \\ &= 1 + (k+1)x + kx^2 \\ &\geq 1 + (k+1)x \end{aligned}$$

denn $kx^2 \geq 0$. (Und? warum muss jetzt $x \geq -1$ sein? Hm?) $\square$

Typischerweise sind Aussagen, die per Vollständiger Induktion bewiesen werden von der Form n–facher Summen oder Produkte. Die "Pünktchen"–Darstellung $1 + \cdots + n$ ist in der Mathematik aus gutem Grund nicht gerne gesehen. Deshalb führen wir nun zwei neue, sehr nützliche Zeichen ein, um Ausdrücke wie beim Beispiel 7 eleganter darstellen zu können. Das ist zum einen das Summenzeichen $\sum$ und zum anderen das Produktzeichen $\prod$.

Definition 1.13 (Summenzeichen) *Für $m, n, k \in \mathbb{Z}$, $a_k \in \mathbb{R}$*

$$\sum_{k=m}^{n} a_k := a_m + a_{m+1} + \cdots + a_n \qquad \text{für } m \leq n$$

Sprich: "Summe der a_k von k gleich m bis n". Ist $k > m$ so definieren wir

$$\sum_{k=m}^{n} a_k := 0 \qquad \text{für } m > n\,.$$

Rechenregeln für das Summenzeichen: Sei stets $m \leq n$. Dann gilt:
Summe/Differenz:

$$\sum_{k=m}^{n} a_k \pm \sum_{k=m}^{n} b_k = \sum_{k=m}^{n} (a_k \pm b_k)$$

Produkt mit einem Skalar ($c \in \mathbb{R}$):

$$\sum_{k=m}^{n} c\, a_k = c \sum_{k=m}^{n} a_k$$

Indexverschiebung :

$$\sum_{k=m}^{n} a_k = \sum_{k=m+p}^{n+p} a_{k-p} = \sum_{k=m-p}^{n-p} a_{k+p}$$

Teleskopsummen :

$$\sum_{k=m}^{n} (a_k - a_{k-1}) = a_n - a_{m-1}$$

Beispiel 10 Teleskopsumme Wir wollen die Summe

$$\sum_{k=1}^{999} \frac{1}{k(k+1)}$$

auswerten. Ohne Computer scheint das auf den ersten Blick ein wenig mühsam zu sein. Wir müssen ein wenig umformen. Zunächst machen wir aus dem komplizierten Term $\frac{1}{k(k+1)}$ zwei einfache:

$$\frac{1}{k(k+1)} = \frac{a}{k} + \frac{b}{k+1} = \frac{k(a+b)+a}{k(k+1)} \underset{a\,=\,1}{\overset{b\,=\,-a}{=}} \frac{1}{k} - \frac{1}{k+1}$$

Nun können wir die Summe umformen zu

$$\sum_{k=1}^{999} \frac{1}{k(k+1)} = \sum_{k=1}^{999} \Big(\underbrace{\frac{1}{k}}_{=:a_k} - \underbrace{\frac{1}{k+1}}_{=:a_{k+1}} \Big)$$

Es gilt für Teleskopsummen dieser Form

$$\begin{aligned} &= \sum_{k=1}^{999} (a_k - a_{k+1}) = a_1 - a_{1000} \\ &= 1 - \frac{1}{10^4} = \frac{999}{10^4} = 0.999 \end{aligned}$$

Insegesamt haben wir berechnet, dass

$$\boxed{\sum_{k=1}^{999} \frac{1}{k(k+1)} = 0.999}$$

Hätten Sie das Ergebnis vermutet?

Das Produktzeichen wird uns nicht gar so häufig begegnen wie das Summenzeichen, aber dennoch, der Vollständigkeit wegen und wo wir gerade dabei sind:

Definition 1.14 (Produktzeichen) *Für* $m, n \in \mathbb{Z}$, $a_k \in \mathbb{R}$

$$\prod_{k=m}^{n} a_k := a_m \cdot a_{m+1} \cdots a_n \qquad \text{für } m \leq n$$

Sprich: "Das Produkt über die a_k *von* k *gleich* m *bis* n*". Auch hier erklären wir die Situation*

$$\prod_{k=m}^{n} a_k := 1 \qquad \text{für } m > n\,.$$

Rechenregeln für das Produktzeichen:
Indexverschiebung:

$$\prod_{k=m}^{n} a_k = \prod_{k=m-p}^{n-p} a_{k+p}$$

Teleskopprodukt:

$$\prod_{k=m}^{n} \frac{a_k}{a_{k-1}} = \frac{a_n}{a_{m-1}}$$

Beispiel 11 mit Summenzeichen

Behauptung $\forall\, n \in \mathbb{N}$:

$$\sum_{k=1}^{n} (2\,k-1)^2 = \frac{n\,(2\,n-1)\,(2\,n+1)}{3}$$

Induktions–anfang

$$\begin{aligned} A(1)\ :\quad & \sum_{k=1}^{1} (2\,k-1)^2 = \frac{1\,(2\cdot 1-1)\,(2\cdot 1+1)}{3} \\ \Leftrightarrow\quad & (2\cdot 1-1)^2 = \frac{3}{3} \\ \Leftrightarrow\quad & 1 = 1 \qquad \checkmark \end{aligned}$$

Induktions–voraussetzung Es gelte $A(l)$ für ein $l \in \mathbb{N}$ mit

$$A(l)\ :\ \sum_{k=1}^{l} (2\,k-1)^2 = \frac{l\,(2\,l-1)\,(2\,l+1)}{3}$$

Induktions–schritt $A(l) \overset{!}{\Rightarrow} A(l+1)$ mit

$$A(l+1)\ :\ \sum_{k=1}^{l+1} (2\,k-1)^2 = \frac{(l+1)\,(2\,l+1)\,(2\,l+3)}{3}$$

**Induktions–
beweis**

$$\sum_{k=1}^{l+1}(2\,k-1)^2 = \sum_{k=1}^{l}(2\,k-1)^2 + (2\,(l+1)-1)^2$$

Hier setzen wir die Induktionsvoraussetzung (IV) ein

$$\begin{aligned}
&= \left(\frac{l\,(2\,l-1)\,(2\,l+1)}{3}\right) + (2\,(l+1)-1)^2 \\
&= \frac{l\,(2\,l-1)\,(2\,l+1) + 3\,(2\,l+1)^2}{3} \\
&= \frac{(l\,(2\,l-1) + 3\,(2\,l+1))\,(2\,l+1)}{3} \\
&= \frac{(2\,l^2 - l + 6\,l + 3)\,(2\,l+1)}{3} \\
&= \frac{(2\,l^2 + 5\,l + 3)\,(2\,l+1)}{3} \\
&= \frac{(2\,l+3)\,(l+1)\,(2\,l+1)}{3}
\end{aligned}$$

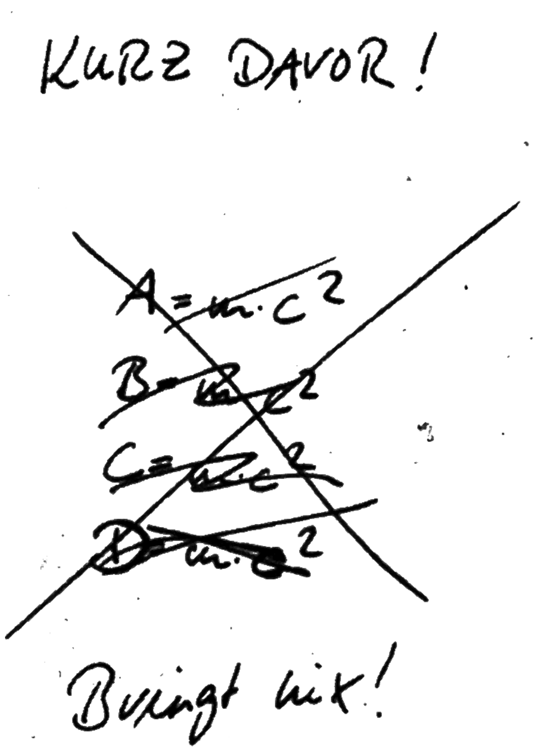
$\square$

Abstrakte Algebra: aufräumen und trennen

2

Wir behandeln:

- Relationen und Äquivalenzklassen
- Ordnung und Strukturen in Mengen



Wir behandeln in diesem Kapitel Teilgebiete der abstrakten Algebra. Grundsätzlich geht es darum, einer Menge eine Struktur zu geben. Das kann in zweierlei Hinsicht geschehen. Einmal zerlegt man eine Menge, die unendlich viele Elemente besitzt in weniger - wobei das ihrerseits unendlich viele sein können - disjunkte Teilmengen, die selbst wieder unendlich viele Elemente besitzt. Zum Beispiel können wir die natürlichen Zahlen in gerade und ungerade Zahlen zerlegen. Das waeren zwei Teilmengen, die zusammen wieder die gesamte Menge ergeben. Wenn es um die Eigenschaft gerade oder ungerade geht, bräuchte man dann nicht mehr jede Zahl betrachten sondern nur noch mit einem jeweiligen Repräsentanten, etwa der 2 und der 3, arbeiten.

Nehmen wir zu den Mengen noch verschiedene Operationen dazu, etwa Addition und Multiplikation, das dürfte jedem geläufig sein, lassen sich weitere "Ordnungsmöglichkeiten" erzeugen. Solche Strukturen erlauben uns etwa alle Rechnungen auf dem Compiuter alleine mit 0 und 1 durchzuführen. Das ist schon schick.

2.1 Relationen

Die Grundlage von Äquivalenzklassen sind Relationen. Sie stellen im Grunde die Bedingung dar, nach der Elemente einer Menge bestimmten Äquivalenzklassen zugeordnet werden, wenn es sich bei der Relation um eine Äquivalenrelation handelt. Wir wollen zunächst ein wenig über Relationen als solche sprechen.

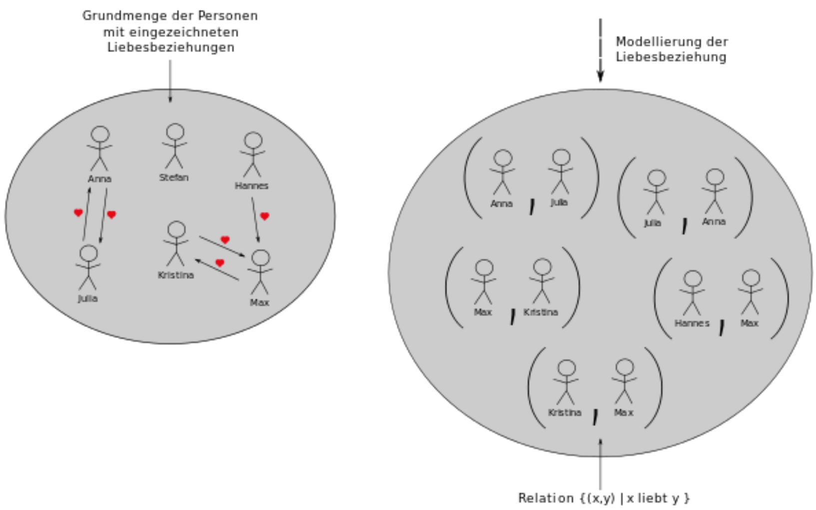
Die allgemeine Bedingung, die zwischen zwei Mengen bestehen kann nennen wir eine Relation.

Definition 2.1 (Relation) *Eine Teilmene $R \subseteq A_1 \times \cdots \times A_n$ heißt (n-stellige) Relation auf das kartesische Produkt $A_1 \times \cdots \times A_n$. Handelt es sich um eine 2-stellige Relation so sprechen wir von einer* binären Relation.
Schreibweisen für binäre Realtionen sind gegeben durch

$$(x, y) \in R \quad \text{oder} \quad xRy \quad \text{oder auch} \quad R(x, y) \quad \text{sowie} \quad x \sim y\,.$$

Relationen sind im Grunde zunächst nur Teilmengen von Produktmengen. Sie benennen lediglich Elemente, der beteiligten Mengen, die in irgendeiner bestimmten Beziehung zueinander stehen.

Beispiel 12 Relationen



(a)

$$A := \{Anna, Stefan, Hannes, Julia, Kristina, Max\}$$
$$R \subseteq A \times A = \{(x,y) \mid x\,,\ y \in A\}$$

mit xRy :="x liebt y" gilt dann

$$\begin{aligned} R &= \{(x,y) \in A^2 \mid xRy\} \\ &= \{(Anna, Julia)\,,\ (Julia, Anna)\,,\ (Kristina, Max)\,, \\ &\quad (Max, Kristina)\,,\ (Hannes, Max)\} \end{aligned}$$

(b) Eine Funktion $f\,:\ \mathbb{D}_f \subseteq \mathbb{R} \to f(\mathbb{D}_f) \subseteq \mathbb{R}$ ist ja gerde die Menge

$$f = \{(x,y) \in \mathbb{R}^2 \mid y = f(x)\}$$

und damit ist f eine Relation auf $\mathbb{R}^2$.

Definition 2.2 (Äquivalenzrelation) *Eine Äquivalenzrelation ist eine binäre Relation $R \subseteq A \times A$ mit folgenden Eigenschaften:*

Jedes Element steht zu sich selbst in Relation

(A1) $\forall\, x \in A\ :\ xRx$ *(reflexiv)*

Wenn ein Element zu einem anderen in Relation steht, so steht das andere auch zu dem ersteren in Relation

(A2) $\forall\, x,y \in A\ :\ xRy \Rightarrow yRx$ *(symmetrisch)*

Wenn ein Element zu einem anderen in Relation steht und dieses zu einem dritten, so steht auch das erste zum dritten in Relation

(A3) $\forall\, x,y,z \in A\ :\ xRy\ \wedge\ yRz \Rightarrow xRz$ *(transitiv)*

Beispiel 13 Äquivalenzrelation

Sei M die Menge der Menschen, a, b zwei (beliebige) Menschen aus M und "a ist gleich groß wie b" sei die Relation. Um zu zeigen, dass diese Relation eine Äquivalenzrelation ist, müssen wir überprüfen, ob die drei Regeln aus der Definition 2.2 erfüllt werden:

(A1) Sei a ein Mensch aus M. "a ist gleich groß wie a", d.h. ein Mensch ist so groß wie er selber. Da können wir mit gutem Gewissen zustimmen.

(A2) Seien $a, b \in M$. "Wenn a so groß ist wie b, so ist b so groß wie a." Wenn also ein Mensch so groß ist wie ein anderer, so ist der Andere so groß wie der Erste. Es braucht nicht allzuviel Lebenserfahrung, um auch diesem Sachverhalt zuzustimmen.

(A3) Seien $a, b, c \in M$. "Wenn a so groß ist wie b und b so groß ist wie c, so ist a so groß wie c." Wenn also eine beliebige 1. Person so groß wie eine 2. Person und diese 2. Person so groß wie eine Dritte ist, so ist die erste Person so groß wie die dritte Person. Ja, ist einleuchtend, stimmt.

Beispiel 14 nicht Äquivalenzrelationen

(a) Die Relationen im Beispiel 12 sind keine Äquivalenzrelationen. Beispiel (a) ist weder reflexiv, noch symmetrisch, noch transitiv.

(b) Beispiel (b) hängt von der genauen Definition der Abbildungsvorschrift von f ab. Nehmen wir einmal an: $f(x) = x^2$. Die Relation beinhaltet dann Punkte deren y-Wert das quadrat des x-Wertes ist, etwa $(2, 4)$, aber nicht $(4, 2)$. Es ist also $2R4$ aber nicht $4R2$, keine Symmetrie, also ist *(A2)* nicht erfüllt. Nur für $(1, 1)$ gilt Reflexivität *(A1)*. Transitivität *(A3)* gilt auch nur für $(1, 1)$: $xRy \ \wedge \ yRz$ bedeutet $y = x^2 \ \wedge \ z = y^2$. Daraus folgt $z = x^4$, was nicht das Gleiche ist wie xRz, bzw. $z = x^2$.

(c) Sei M die Menge der Menschen, a, b zwei (beliebige) Menschen aus M und "a mag b" sei die Relation. Sicher kann da jeder aus seinem reichhaltigen Erfahrungsschatz schöpfen und einsehen, dass die Symmetrie hier nicht gegeben ist.

(d) Etwas abstrakter: Seien $a, b \in \mathbb{Z}$ mit der Relation $aRb \ \Leftrightarrow \ a \cdot b \geq 0$. Wir prüfen das:

(A1) $a^2 \geq 0$ ist immer erfüllt. ✓

(A2) $a \cdot b \geq 0 \ \Leftrightarrow \ b \cdot a \geq 0$ gilt wegen der Kommutativität der üblichen Multiplikation in $\mathbb{R}$. ✓

(A3) $a \cdot b \geq 0 \ \wedge \ b \cdot c \geq 0$ führt nicht zwingend auf $a \cdot c \geq 0$. Zum Beispiel dann nicht, wenn $a > 0, b = 0$ und $c < 0$ ist.

Definition 2.3 (Ordnung) *Eine Relation $R \subseteq A \times A$ mit den Eigenschaften (A1), (A3) und*

(A4) $\forall\, x, y \in A \ : \ xRy \wedge yRx \quad \Rightarrow \quad x = y$

heißt Ordnung oder Halbordnung.

Beispiel 15 Ordnung

Die Relation "$a \geq b$" auf - sagen wir - $\mathbb{Z}^2$ ist eine Ordnung.

(A1) $a \geq a$ ✓

(A3) Wenn $a \geq b$ und $b \geq c$ ist, dann können wir getrost davon ausgehen, dass auch $a \geq c$ ist.

(A4) $a \geq b$ und $b \geq a$ kann nur dann gleichzeitig erfüllt sein, wenn $a = b$ ist.

Wohlbemerkt gilt die Symmetrie in diesem Fall nicht. Es handelt sich rein um eine Ordnung, nicht um eine Äquivalenzrelation.

Wenn man nun eine Äquivalenzrelation vorliegen hat dann handelt es sich ja um eine Teilmenge, bzw. um eine klar umrissene Menge beteiligter Elemente. Eines der Elemente greifen wir als Repräsentanten heraus.

Definition 2.4 (Äquivalenzklasse) *Ist $R \subseteq A \times A$ eine Äquivalenzrelation und $a \in A$, dann heißt die Menge*

$$[a]_R := \{b \in A \mid aRb\}$$

Äquivalenzklasse (ÄK) von R und a heißt Repräsentant der ÄK $[a]_R$.

Mit der Bildung von Äquivalenzklassen ist es und möglich, eine Menge vollständig in disjunkte Teilmengen zu zerlegen. Innerhalb jeder ÄK stehen die Elemente bzgl. der zugrundeliegenden Relation in Bezug zueinander.

Beispiel 16 Äquivalenzklasse

(a) Wir betrachten M, die Menge aller Studenten der HTWG und die auf $M \times M$ basierte Relation

$$R = \{(x, y) \in M \times M \mid x \text{ und } y \text{ besuchen die selbe Vorlesung}\}\ .$$

Zunächst einmal handelt es sich hierbei um eine Äquivalenzrelation. Wir wählen einen Studenten, der diese bestimmte Vorlesung besucht, nämlich der Student b, als Repräsentanten aller Studierenden in dieser Vorlesung:

$$[b]_R = \{a \in M \mid a \text{ besucht die selbe Vorlesung wie } b\}$$

Egal wen ich als b auswähle die Menge $[b]_R$ beinhaltet nun alle Studierenden dieser Vorlesung.

(b) Es sei $R \subseteq \mathbb{N}_0^2$ mit
$$R = \left\{ (x,y) \in \mathbb{N}_0^2 \;\middle|\; \frac{x-y}{3} \in \mathbb{Z} \right\}.$$
Überzeugen Sie sich davon, dass es sich hierbei um eine Äquivalenzrelation handelt. Wir bilden nun die Äquivalenzklassen:

$$\begin{aligned}
[0]_R &= \{x \in \mathbb{N}_0 \mid 0Rx\} = \left\{ x \in \mathbb{N}_0 \;\middle|\; \frac{-x}{3} \in \mathbb{Z} \right\} = \{\dots, -6, -3, 0, 3, 6, \dots\} \\
[1]_R &= \{x \in \mathbb{N}_0 \mid 1Rx\} = \left\{ x \in \mathbb{N}_0 \;\middle|\; \frac{1-x}{3} \in \mathbb{Z} \right\} = \{\dots, -5, -2, 1, 4, 7, \dots\} \\
[2]_R &= \{x \in \mathbb{N}_0 \mid 2Rx\} = \left\{ x \in \mathbb{N}_0 \;\middle|\; \frac{2-x}{3} \in \mathbb{Z} \right\} = \{\dots, -4, -1, 2, 5, 8, \dots\} \\
[3]_R &= [0]_R \ni 3 \\
[4]_R &= [1]_R \ni 4 \\
&\text{etc.}
\end{aligned}$$

Insgesamt gilt
$$[a]_R \cap [b]_R = \emptyset \quad \Leftrightarrow \quad a \neq b$$
und
$$[0]_R \cup [1]_R \cup [2]_R = \mathbb{N}_0 \,.$$

Diese beiden Eigenschaften sind die wesentlichen einer Äquivalenzklasse und gelten für diese Strukturen immer!!!

Satz 2.5 *Es sei $R \subseteq A^2$ eine ÄR und $A \neq \emptyset$. Dann gilt:*

(a) $\forall a \in A \;:\; [a]_R \neq \emptyset$

(b) $\forall b \in [a]_R \;:\; [b]_R = [a]_R$

(c) Wenn b nicht in Relation zu a steht, dann gilt
$$[b]_R \cap [a]_R = \emptyset \,.$$

(d)
$$A = \bigcup_{a \in A} [a]_R$$

Beweis auf Seite 153

2.2 Algebraische Strukturen

Wir wollen in diesem Kapitel verschiedene Strukturen auf Mengen zu studieren. Was eine Struktur auf einer Menge ganz ist, hängt von betrachteten Fragestellung ab. In dieser Vorlesung wird die Struktur stets aus einer oder mehreren zweistelligen Operationen auf der Menge bestehen, die dann bestimmten Gesetzmäßigkeiten, also Axiomen, genügen. Dabei ist eine zweistellige Operation auf einer Menge G eine Abbildung, die einem Paar (g, h) von Elementen aus G wieder ein Element in G zuordnet, also eine Abbildung $G \times G \to G$.

2.2.1 Gruppen

Definition 2.6 (Gruppe) *Eine Gruppe ist ein Paar $(G, \circ)$ bestehend aus einer nicht leeren Menge $G \neq \emptyset$ und einer zweistelligen Operation "$\circ$" mit*

$$g \circ h \in G\,,$$

für $g, h \in G$, die folgende Gruppenaxiome erfüllt:

(G1) $(g \circ h) \circ f = g \circ (h \circ f)\ \forall g, h, f \in G$ (Assoziativgesetz)

(G2) $\exists\, n \in G\ :\ \forall\, g \in G\ :\ n \circ g = g$ (neutrales Element)

(G3) $\forall\, g \in G\ \exists\, g^{-1} \in G\ :\ g^{-1} \circ g = n$ (inverses Element)

Gilt darüberhinaus

(G4) $\forall\ g\, h \in G\ :\ g \circ h = h \circ g$ (Kommutativität)

so heißt die Gruppe kommutativ *oder* abelsch.
Gilt hingegen nur das Assoziativgesetz so sprechen wir von einer Halbgruppe *und gelten nur die Axiome (G1) und (G2) so heißt G* Monoid.

Beispiel 17 Gruppe

(a) $(\mathbb{R}, +)$ mit der üblichen Addition als Gruppenoperation ist eine abelsche Gruppe. $n = 0$ erfüllt die Rolle des neutralen Elements, und zu jeder Zahl g existiert mit $-g$ ein inverses Element.

(b) $(\mathbb{R}^2, +)$ mit

$$(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 + b_2) \in \mathbb{R}^2$$

ist eine abelsche Gruppe.

(c) $(\mathbb{Z}, +)$ ist eine abelsche Gruppe.

(d) $(\{-1, 1\}, \cdot)$ mit

$\cdot$	-1	1
-1	1	-1
1	-1	1

beschreibt eine endliche, abelsche Gruppe. Jedes Element ist zu sich selbst das inverse Element und $n = 1$ ist das neutrale Element.

Beispiel 18 nicht Gruppe

(a) $(\mathbb{N}, +)$ ist keine Gruppe.

(b) $(\mathbb{N}_0, +)$ ist keine Gruppe aber ein abelsches Monoid.

(c) $(\mathbb{Z}, \cdot)$ ist keine Gruppe aber ein abelsches Monoid.

Satz 2.7 *Für eine Gruppe $(G, \circ)$ gelten die folgenden Eigenschaften:*

(a) Es gilt

$$g \circ g^{-1} = n\,.$$

(b) Es gilt

$$g \circ n = g\,.$$

(c) Das neutrale Element von G ist eindeutig bestimmt.

(d) Zu jedem Element von G gibt es genau ein Inverses.

(e) Das neutrale Element ist invers zu sich selbst.

(f) $\forall a \in G \;:\; (a^{-1})^{-1} = a$

(g) $\forall g, h \in G \;:\; (g \circ h)^{-1} = h^{-1} \circ g^{-1}$

(h) Kürzungsregeln: Es gilt

$$a \circ b = a \circ c \quad \Rightarrow \quad b = c$$

und

$$b \circ a = c \circ a \quad \Rightarrow \quad b = c\,.$$

(i) Lösbarkeit von Gleichungen: $\forall a, b \in G\; \exists! x, y \in G \;:\; a \circ x = b \;\wedge\; y \circ a = b$

Beweis auf Seite 153

Wir wollen hier schon mal eine neue Menge, nämlich die Menge von Matrizen einführen, die später an Bedeutung gewinnen wird.

Definition 2.8 (Matrix) *Eine Matrix A ist eine strukturierte Anordnung von $m \cdot n$ Zahlen aus $\mathbb{K}$ gemäß*

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

Die Matrix hat m Zeilen und n Spalten.
Notation:

$$A = (A_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \qquad \text{oder kurz} \qquad A = (A_{ij})_{ij}$$

Die Menge aller $m \times n$-Matrizen mit Komponenten aus der Menge $\mathbb{K}$ bezeichnen wir mit $\mathbb{K}^{m \times n}$.

A_{ij} bezeichnet die i–te Zeile und j–te Spalte. Immer im Sinne

"Zeilen zuerst, Spalten später."

Beispiel 19 Matrizen

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$$

A ist eine *quadratische Matrix*, da $m = n$ gilt.

Die Elemente A_{ii} einer quadratischen Matrix $A \in \mathbb{K}^{n\times n}$ heißen *Hauptdiagonale* und $A_{i,i+1}$, bzw $A_{i+1,i}$ heißen *Nebendiagonale*.

$$B = \begin{pmatrix} 1+i & 2i & 0 \\ 0 & -i & 4 \end{pmatrix} \in \mathbb{C}^{2\times 3}$$

$$C = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \in \mathbb{R}^{2\times 2}$$

C ist eine *symmetrische Matrix*, da $C_{ij} = C_{ji}$ gilt.

$$R = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix} \in \mathbb{R}^{3\times 3}$$

R ist eine *obere Dreiecksmatrix*, da $R_{ij} = 0$ für $i > j$ gilt.

$$U = \begin{pmatrix} 0 & 0 & 0 \\ -4 & 0 & 0 \\ -1 & 2 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 3}$$

U ist eine *echte untere Dreiecksmatrix*, da $U_{ij} = 0$ für $i \leq j$ gilt.

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{pmatrix} \in \mathbb{R}^{3\times 3}$$

D ist eine *Diagonalmatrix*, da $D_{ij} = 0$ für $i \neq j$ gilt.

$$E_n = (\delta_{ij})_{1\leq i,j\leq n} \in \mathbb{R}^{n\times n}$$

E heißt *Einheitsmatrix*. Dabei heißt

$$\delta_{ij} := \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j \end{cases} \quad \textit{Kronecker-Symbol}$$

$$E_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Wir schreiben auch einfach E für die Einheitsmatrix, wenn die Anzahl Zeilen, bzw Spalten durch dem Kontext klar ist. Manchmal verwendet man auch den Buchstaben I für Identität.

$$\mathbb{R}^2 \times \mathbb{R}^2 \neq \mathbb{R}^{2\times 2}$$

Beispiel 20 Matrizen als Gruppe mit $+$

$(\mathbb{R}^{2\times 2}, +)$ mit der *komponentenweisen Addition*

$$(A+B)_{ij} = A_{ij} + B_{ij}$$

ist eine abelsche Gruppe.

Beispiel 21 Matrizen als Monoid mit $\cdot$

$(\mathbb{R}^{2\times 2}, \cdot)$ mit der üblichen *Matrixmultiplikation*

$$C = (c_{ij})_{ij} = \left(\sum_{k=1}^{2} a_{ik}b_{kj}\right)_{ij} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

ist ein Monoid.

2.2.2 Ringe

Wenn uns nun zum Rechnen eine Operation nicht genügt, wir sind ja bereits gewohnt zu addieren und zu multiplizieren, dann erweitern wir einfach um eine weitere zweistellige Operation.

Definition 2.9 (Ringe) *Ein Ring ist ein Tripel $(R, +, \circ)$ bestehend aus einer nicht leeren Menge $R \neq \emptyset$ und zwei Operationen $\circ$ und $+$ mit*

$$g + h \in R \qquad \text{und} \qquad g \circ h \in R\,,$$

so dass folgende Ringaxiome erfüllt sind:

(R1) $(R, +)$ *ist eine abelsche Gruppe* (neutr. Elt. bzgl. Addition ist 0)

(R2) $\forall a, b, c \in R \ : \ a \circ (b \circ c) = (a \circ b) \circ c$ (Assoziativität)

(R3) $\forall a, b, c \in R \ :$ (Distributivität)

$$a \circ (b + c) = a \circ b + a \circ c$$

und

$$(b + c) \circ a = b \circ a + c \circ a$$

Ein Ring heißt kommutativ, *wenn Kommutativität bezügl. "$\circ$" ebenfalls gegeben ist. Existiert ein neutrales Element $n \in R$ bzgl. "$\circ$", so dass für alle $a \in R$ gilt $n \circ a = a$, so sprechen wir von einem* Ring mit Einselement.

Ein Ring mit Einselement ist im Grunde dann gegeben, wenn R bzgl. der Addition eine abelsche Gruppe, bzgl. der Multiplikation ein Monoid bildet und für beide Verknüpfungen das Distributivgesetz gilt.

Beispiel 22 Ring $(\mathbb{Z}, +, \cdot)$ ist ein kommutativer Ring mit Einselement.

Beispiel 23 Matrizen als Ring $(\mathbb{R}^{2\times 2}, +, \cdot)$ ist ein Ring mit Eins.

2.2.3 Körper

Definition 2.10 (Körper) *Eine Menge K mit zwei binären Operationen $+$ und $\circ$ heißt* Schiefkörper, *wenn K ein Ring ist und $K \setminus \{0\}$ bzgl. der Multiplikation eine Gruppe bildet.*
K heißt Körper, *wenn K Schiefkörper ist und $K \setminus \{0\}$ bzgl. der Multiplikation eine kommutative Gruppe bildet.*

Beispiel 24 Körper

(a) $(\mathbb{Q}, +, \cdot)$ und $(\mathbb{R}, +, \cdot)$ mit der üblichen Addition und Multiplikation sind Körper.

(b) $(\{0, 1\}, +, \cdot)$ mit

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

ist ein Körper.

(c) $(\mathbb{C}, +, \cdot)$ bildet einen Körper.

(d) $M \subset \mathbb{R}^{2\times 2}$ über $\mathbb{R}$ mit

$$M = \{A \in \mathbb{R}^{2\times 2} \mid a_{11}a_{22} \neq a_{12}a_{21}\}$$

ist ein Schiefkörper.

Beispiel 25 nicht Körper

(a) $(\mathbb{Z}, +, \cdot)$ mit der üblichen Addition und Multiplikation ist ein kommutativer Ring mit Eins aber kein Körper, denn es gibt keine multiplikativen Inversen.

(b) $(\{0, 1, 2, 3\}, +, \circ)$ mit

+	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

∘	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	2
3	0	3	2	1

ist ein kommutativer Ring mit Eins und kein Körper, denn die 2 hat kein multiplikatives Inverses.

2.2.4 Vektorräume

Von besonderer Bedeutung ist der Vektorraum , oder auch linearer Raum genannt, als algebraische Struktur. Im Gegensatz zu Gruppen und Ringen basiert ein Vektorraum nicht nur auf einer beliebigen Menge V sondern zusätzlich noch auf einem Körper $(\mathbb{K}, +, \cdot)$, kurz $\mathbb{K}$. Wir werden in der Regel $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ betrachten. Es gibt weitere Möglichkeiten, aber an denen sind in der Regel nur Mathematiker interessiert. Wir definieren eine Addition $\oplus$ zwischen Elementen eines Vektorraums, welche wir Vektoren nennen und eine Multiplikation $\odot$ zwischen Vektoren und Elementen des Körpers, welche wir Skalare nennen. Folglich sprechen wir von skalarer Multiplikation. Wirhaben hier also zwei Sorten von Mengen und zwei Sorten Addition, bzw. Multiplikaiton. Ansonsten ist nix wirklich Neues dazugekommen.

Definition 2.11 (Vektorraum (VR)) *Sei $V \neq \emptyset$ eine nicht leere Menge und $(\mathbb{K}, +, \cdot)$ ein Körper. Es bezeichne $\oplus$ die Addition zwischen Elementen in V und $\odot$ die (skalare) Multiplikation zwischen einem Skalar aus $\mathbb{K}$ und einem Element aus V. Für $\vec{a}, \vec{b} \in V$ und $\lambda \in \mathbb{K}$ gelte damit*

$$\vec{a} \oplus \vec{b} \in V \quad \wedge \quad \lambda \odot \vec{a} \in V\,.$$

$(V, (\mathbb{K}, +, \cdot), \oplus, \odot)$ heißt Vektorraum, *wenn folgende Bedingungen erfüllt sind:*

Addition:

(V1) $(V, \oplus)$ *ist eine abelsche Gruppe.*

Multiplikation:

(V2) $\exists 1 \in \mathbb{K}\, \forall \vec{a} \in V \;:\; 1 \odot \vec{a} = \vec{a}$ *(neutrales Element bezüglich der Multiplikation)*

(V3) $\forall \lambda, \mu \in \mathbb{K}\, \forall \vec{a} \in V \;:\; \lambda \odot (\mu \odot \vec{a}) = (\lambda \cdot \mu) \odot \vec{a}$ *(Assoziativgesetz)*

Addition und Multiplikation:

(V4) $\forall \lambda \in \mathbb{K}\, \forall \vec{a}, \vec{b} \in V \;:\; \lambda \odot (\vec{a} \oplus \vec{b}) = (\lambda \odot \vec{a}) \oplus (\lambda \odot \vec{b})$ *(Distributivgesetz)*

(V5) $\forall \lambda, \mu \in \mathbb{K}\, \forall \vec{a} \in V \;:\; (\lambda + \mu) \odot \vec{a} = (\lambda \odot \vec{a}) \oplus (\mu \odot \vec{a})$ *(Distributivgesetz)*

Elemente eines Vektorraums heißen Vektoren. *Das neutrale Element bezüglich der Addition heißt* Nullvektor.

Notation:

Wir schreiben kurz V über $\mathbb{K}$ oder auch nur V statt $(V, (\mathbb{K}, +, \cdot), \oplus, \odot)$.

Die kleinen Pfeile über den Vektoren $\vec{a}$ und $\vec{b}$ symbolisieren des besseren Verständnisses wegen, dass diese beiden Elemente aus V sind. Wir werden aber des Weiteren, der besseren Übersicht wegen, auf diese Zusatzbezeichnung verzichten:

$$a, b \qquad \text{statt} \qquad \vec{a}, \vec{b}$$

Wir werden später das Symbol $+$ für beide Additionen verwenden:

$$a + b \qquad \text{statt} \qquad a \oplus b$$

Die Symbole $\cdot$ und $\odot$ werden für gewöhnlich gar nicht geschrieben:

$$\lambda\, a\,,\ \lambda\, \mu \qquad \text{statt} \qquad \lambda \odot a\,,\ \lambda \cdot \mu$$

Ein Vektorraum ist auf jeden Fall eine abelsche Gruppe bzgl. der Addition. Wir werden auch Vektoren auf die ein oder andere Art miteinander multiplizieren aber es gibt keine definierte Multiplikation zwischen Vektoren, die sich in eine Struktur packen lässt.

Beispiel 26 Vektorräume

(a) $\mathbb{C}$ über $\mathbb{C}$ oder $\mathbb{C}$ über $\mathbb{R}$ oder $\mathbb{R}$ über $\mathbb{R}$, also $(\mathbb{C}, \mathbb{C})$, $(\mathbb{C}, \mathbb{R})$ oder $(\mathbb{R}, \mathbb{R})$ sind VRe.

(b) $\mathbb{R}$ über $\mathbb{C}$, bzw. $(\mathbb{R}, \mathbb{C})$ ist kein VR.

(c) $(\mathbb{R}^2, \mathbb{R})$ mit

$$(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 + b_2) \qquad \text{und} \qquad \lambda\,(a_1, a_2) = (\lambda\, a_1, \lambda\, a_2)$$

ist ein VR.

(d) $(\mathbb{C}^2, \mathbb{C})$ und $(\mathbb{C}^2, \mathbb{R})$ sind ebenfalls VRe, jedoch nicht $(\mathbb{R}^2, \mathbb{C})$.

(e) $(\mathbb{P}_3, \mathbb{R})$ die Menge aller Polynome vom Grad ≤ 3 mit reellen Koeffizienten der Form

$$p(x) = a_0 + a_1\, x + a_2\, x^2 + a_3\, x^3 = \sum_{k=0}^{3} a_k x^k\,,\ a_k \in \mathbb{R}$$

mit

$$(p + q)(x) = p(x) + q(x) = \sum_{k=0}^{3} a_k x^k + \sum_{k=0}^{3} b_k x^k = \sum_{k=0}^{3} (a_k + b_k) x^k \in \mathbb{P}_3$$

und

$$\lambda\, p(x) = \sum_{k=0}^{3} \lambda a_k x^k \in \mathbb{P}_3\,,\ \lambda \in \mathbb{R}$$

ist ein VR.

(f) $(\mathbb{P}_3, \mathbb{C})$ die Menge aller Polynome vom Grad ≤ 3 mit reellen Koeffizienten ist kein VR, denn

$$\lambda\, p(x) \notin \mathbb{P}_3$$

hat keine reellen Koeffizienten mehr.

(g) $(\mathbb{R}^{2\times 2}, (\mathbb{R}, +, \cdot), \oplus, \odot)$ - kurz $(\mathbb{R}^{2\times 2}, \mathbb{R})$ - mit der komponentenweisen Addition

$$A \oplus B := (A_{ij} + B_{ij})_{ij}$$

und skalaren Multiplikation

$$\alpha \odot A := (\alpha A_{ij})_{ij}$$

ist ein Vektorraum.

Lineare Algebra: Konzepte in Vektorräumen 3

Wir behandeln:

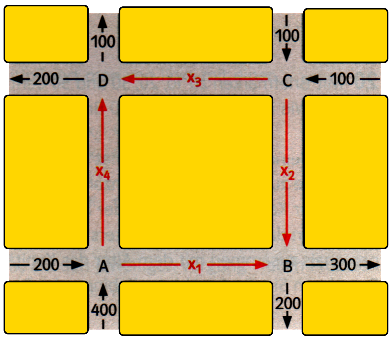
- Matrizen im LGS und als lineare oder affine Abbildungen
- Basis und Dimension sowie Eigenwerte und -vektoren als Basis
- Messen in Vektorräumen

Vektorräume sind also erst einmal Mengen - das können Zahlen, Tupel, Funktionen, Matrizen, etc. sein - deren Elemente, Vektoren genannt, in vorgegebener Weise verrechnet werden können. Das sind unsere jeweiligen Rechenregeln. Diese Rechenregeln unterliegen Gesetzen, müssen also bestimmte Eigenschaften erfüllen. Mit Vektorräumen lassen sich Konzepte aufbauen mit Hilfe derer wir dann etwa ein gutes Preis-Leistungs-Verhältnis bei Datenkompressionen (jpg, mpg, etc) erreichen oder Muster in Bildern erkennen können (waver-Produktion, Gesichtserkennung) und Vieles mehr. Aus diesen Konzepten resultieren Rechenmethoden, bzw. Formeln. Die Formeln vertrauen darauf, dass ihre Voraussetzungen erfüllt sind sonst liefern sie keine korrekten Ergebnisse; Ihr Gesicht würde nicht erkannt werden, Ihre Urlaubsfotos wären futsch und ihr Lieblingssong würde klingen als kratze jemand mit Fingenägeln über die Tafel. Darum ist es wichtig, das Basissetting (hier die Vektorraumregeln) streng einzuhalten.

Um unsere genannten Anwendungsbeispiele auch einmal rechnen zu können, benötigen wir ein wenig Methodenwerkzeug. Dazu machen wir einen kleinen Exkurs über das Lösen von linearen Gleichungssystemen im folgenden Unterkapitel.

3.1 Lösen linearer Gleichungssysteme

Beispiel 27 Geviert, Teil 1/3



Wir lassen uns von einem kleinen Beispiel begleiten: Bei einem Geviert aus Einbahnstraßen sind die Verkehrsdichten für die zu- und abfließenden Ströme bekannt.
Für eine Straßenbaufirma hat folgende Fragen:

(a) Ist eine Sperrung des Straßenstücks AD ohne Drosselung des Zuflusses möglich?

(b) Welches ist die minimale Verkehrsdichte auf dem Straßenstück AB?

(c) Welches ist die maximale Verkehrsdichte auf dem Straßenstück CD?

Für die Verkehrsdichten x_1 bis x_4 lässt sich die Situation in einem LGS zusammenfassen. Bei Kreuzung A kommt eine Verkehrsdichte von $200 + 400 = 600$ zusammen, die sich auf die Einbahnstraßen x_1 und x_4 verteilen. Wie im Detail wissen wir nicht. Wir wissen nur $x_1+x_4 = 600$. Damit haben wir die erste Gleichung des Systems. Sukzessiv fahren wir fort und erhalten insgesamt das LGS

$$
\begin{aligned}
A &: \quad x_1 + x_4 = 600 \\
B &: \quad x_1 + x_2 = 500 \\
C &: \quad x_2 + x_3 = 200 \\
D &: \quad x_3 + x_4 = 300\,.
\end{aligned}
$$

Das Aufstellen des LGS, also die mathematische Beschreibung einer gegebenen Situation, ist der erste Schritt. Der dritte Schritt, ich greife voraus, besteht darin, das System nach den Unbekannten aufzulösen und dazwischen werden wir das LGS ein wenig umformulieren, es in die sogenannte Matrix-Vektor-Darstellung bringen, um die Handhabung zu vereinfachen.

Wir betrachten nun noch einmal zusammenfassend grundlegende Matrizenoperationen auch für nicht quadratische Matrizen

Rechenregeln für Matrizen: Es seien $A, B \in \mathbb{K}^{M\times N}$. Die *Summe von Matrizen* ist definiert als

$$A + B = (A_{ij} + B_{ij})_{ij} \,.$$

Die *skalare Multiplikation von Matrizen* mit $\alpha \in \mathbb{R}$ ist definiert als

$$\alpha\, A = (\alpha\, A_{ij})_{ij} \,.$$

Das *Matrixprodukt* $A \cdot B$ ist definiert durch

$$C \cdot D = \left(\sum_{m=1}^{M} C_{nm} D_{ml} \right)_{1\leq n \leq N\,,\, 1 \leq l \leq L} \quad , \; C \in \mathbb{K}^{N\times M}\,,\; D \in \mathbb{K}^{M\times L}$$

Die Addiation und die skalare Multiplikation sorgen dafür, dass wir mit Matrizen einen Vektorraum bilden können. Die Matrixmultiplikation erlaubt weitere Operationen in diesem Vektorraum, sofern es sich um quadratische Matrizen handelt!

Vorsicht ist darin geboten, dass die Dimensionen im Sinne der Definition "passen". Bei der Addition müssen die beteiligten Matrizen gleich "groß" sein und bei einem Matrixprodukt muss die Anzahl der Spalten der "linken" Matrix der Anzahl der Zeilen der "rechten" Matrix entsprechen. Sonst passt es nicht.

Beispiel 28

(a)

$$\underbrace{\begin{pmatrix} 2 & 1 \\ 3 & -1 \\ 0 & 2 \end{pmatrix}}_{3\times 2} \underbrace{\begin{pmatrix} 2 \\ -1 \end{pmatrix}}_{2\times 1} = \underbrace{\begin{pmatrix} 3 \\ 7 \\ -2 \end{pmatrix}}_{3\times 1}$$

(b)

$$\underbrace{(3,4)}_{1\times 2} \underbrace{\begin{pmatrix} 2 & -1 \\ 8 & 9 \end{pmatrix}}_{2\times 2} = \underbrace{\begin{pmatrix} 38, 33 \end{pmatrix}}_{1\times 2}$$

(c)

$$\underbrace{(2,4)}_{1\times 2} \underbrace{\begin{pmatrix} 1 \\ 3 \end{pmatrix}}_{2\times 1} = \underbrace{14}_{1\times 1}$$

(d)

$$\underbrace{\begin{pmatrix} 2 \\ 4 \end{pmatrix}}_{2\times 1} \underbrace{(1,3)}_{1\times 2} = \underbrace{\begin{pmatrix} 2 & 6 \\ 4 & 12 \end{pmatrix}}_{2\times 2}$$

(e)

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ -2 & -1 \end{pmatrix} = \begin{pmatrix} -3 & 1 \\ -5 & 5 \end{pmatrix}$$
$$\begin{pmatrix} 1 & 3 \\ -2 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 10 & 14 \\ -5 & -8 \end{pmatrix}$$

Für die Matrixmultiplikation gilt nicht im Allgemeinen Kommutativität:

$$A \cdot B \neq B \cdot A$$

(Ausnahmen gibt es in speziellen Fällen.)

Noch eine kleine Besonderheit von Matrizen, die thematisch gerade gut an diese Stelle passt ist die

Definition 3.1 (Transponierte einer Matrix) *Die* Transponierte *einer Matrix A ist die Matrix*

$$\left(A^T\right)_{ij} = A_{ji}\,,$$

deren Spalten, bzw. Zeilen gerade die Zeilen, bzw Spalten der Matrix A sind.
A heitßt symmetrische Matrix *wenn sie gleich ihrer Transponierten ist, d.h. wenn*

$$A^T = A$$

gilt.

Beispiel 29

(a)

$$\mathbb{R}^{2\times 3} \ni \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix} \in \mathbb{R}^{3\times 2}$$

(b)

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}$$

ist symmetrisch.

Satz 3.2 *Seien A, B und C Matrizen mit passenden Dimensionen. Dann gilt:*

$$(AB)^T = B^T A^T$$

Beweis von Satz 3.2 auf Seite 157

Mit dem Symbol für transponierte Matrizen kann man nun die Matrixmultiplikation nochmal ein wenig modifiziert hinschreiben:

Beispiel 30 Auch Matrixmultiplikationen nur anders hingeschrieben:

$$A\,x = \begin{pmatrix} \sum_{j=1}^{N} A_{1j}x_j \\ \vdots \\ \sum_{j=1}^{N} A_{Nj}x_j \end{pmatrix}$$

$$x^T A = (x_1, \dots, x_N)\,A = \begin{pmatrix} \sum_{i=1}^{N} A_{i1}x_i & \cdots & \sum_{i=1}^{N} A_{iN}x_i \end{pmatrix}$$

Bemerkung 3.3

1. *Bei Diagonalmatrizen ist das Matrizenprodukt kommutativ.*
2. *Obere/untere Dreiecksmatrizen bzw. Diagonalmatrizen bleiben im Produkt obere/untere Dreiecksmatrizen bzw. Diagonalmatrizen.*

Die Erweiterung der Definition 3.1 von reelle auf komplexe Zahlen erhalten wir durch folgende

Definition 3.4 (hermitesch) *Es sei eine Matrix $A \in \mathbb{C}^{n\times n}$ gegeben. Die* komplex konjugierte Matrix $\bar{A}$ *erhalten wir, indem wir alle Einträge der Matrix komplex konjugieren. Die Transponierte und komplex konjugierte Matrix $A^\star = \bar{A}^T$ heißt* adjungierte Matrix. *A heißt* hermitesch, *wenn sie gleich ihrer Adjungierten ist, also wenn*

$$A = \bar{A}^T =: A^\star$$

gilt. Die Einträge einer komplex konjugierten Matrix ergeben sich als

$$(a_{ij})^\star_{ij} = (\bar{a}_{ji})_{ij}\ .$$

Beispiel 31

(a) komplex konjugiert und transponiert:

$$A = \begin{pmatrix} 2+i & -i \\ 3 & 3\,i-4 \end{pmatrix} \quad \Rightarrow \quad A^\star = \begin{pmatrix} 2-i & 3 \\ i & -3\,i-4 \end{pmatrix}$$

(b) hermitesch:

$$A = \begin{pmatrix} 2 & -i \\ i & 3 \end{pmatrix} \quad \Rightarrow \quad A^\star = \begin{pmatrix} 2 & -i \\ i & 3 \end{pmatrix} = A$$

Mit der neugewonnenen Virtuosität in Sachen Matrizenmultiplikation können wir uns nun leichten Fußes wieder unserem Straßengeviert zuwenden.

Beispiel 32 Geviert, Teil 2/3

Das LGS des Gevierts aus Beispiel 27 in Matrix-Darstellung sieht dann so aus:

$$\begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 600 \\ 500 \\ 200 \\ 300 \end{pmatrix}$$

Wir können nun ein LGS in Matrix-Form, d.h.

Zu gegebenem $y \in \mathbb{R}^m$ und $A \in \mathbb{R}^{m\times n}$ ist ein $x \in \mathbb{R}^n$ gesucht mit

$$A\,x = y\,,$$

darstellen. Wir fragen uns, ob es zu dieser Art Problemstellung eine Lösung gibt und wenn ja wieviele und wie können wir diese berechnen; wenn nein, warum nicht?

Schauen wir zunächst nach der Lösungsmethode. Wir werden schon merken wo wir hängenbleiben...

Definition 3.5 (Stufenform und Zeilenumformungen) *Zeilenumformungen beinhalten zwei Operationen:*

1. *Multiplikation einer Zeile mit einer Zahl aus* $\mathbb{R} \setminus \{0\}$.
2. *Addition einer Zeile zu einer anderen.*

Eine obere Dreiecksmatrix, die aus $A \in \mathbb{R}^{m\times n}$ *mittels Zeilenumformungen aus* A *entstanden ist heißt* Stufenform *von* A.

Zeilenumformungen von (A, y) ändern die Lösungesmenge des LGS $A\,x = y$ nicht.

Beispiel 33 irgendein Beispiel

Das Gleichungssystem in Matrix-Vektor-Darstellung:

$$\begin{aligned} x_1 - x_2 + 2\,x_3 &= 6 \\ 2\,x_1 + 3\,x_2 + 3\,x_3 &= 5 \\ 2\,x_1 + 8\,x_2 + 2\,x_3 &= 2 \end{aligned} \quad \Leftrightarrow \quad \begin{pmatrix} 1 & -1 & 2 \\ 2 & 3 & 3 \\ 2 & 8 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ 5 \\ 2 \end{pmatrix}$$

Ziel ist, die erweiterte Matrix (A, y) mittels Zeilenumformungen in Stufenform zu bringen:

$$\left(\begin{array}{ccc|c} 1 & -1 & 2 & 6 \\ 2 & 3 & 3 & 5 \\ 2 & 8 & 2 & 2 \end{array}\right) \underset{(\mathrm{III}-2\,\mathrm{I})/2}{\overset{\mathrm{II}-2\,\mathrm{I}}{\longrightarrow}} \left(\begin{array}{ccc|c} 1 & -1 & 2 & 6 \\ 0 & 5 & -1 & -7 \\ 0 & 5 & -1 & -5 \end{array}\right)$$

$$\overset{\mathrm{III}-\mathrm{II}}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & -1 & 2 & 6 \\ 0 & 5 & -1 & -7 \\ 0 & 0 & 0 & 2 \end{array}\right)$$

In der 3-ten Zeile steht jetzt $0 \cdot x_1 + 0 \cdot x_2 + 0 \cdot x_3 = 2$, was natürlich direkt auf einen Wiederspruch führt. Es gibt also kein x, welches auch nur die dritte Zeile erfüllt, insbesondere auch nicht das ganze LGS. Damit hat dieses LGS keine Lösung. Die Lösungsmenge $\mathbb{L}$ ist die Menge aller x, die das LGS erfüllen. In diesem Fall gilt dann:

$$\mathbb{L} = \emptyset$$

Wir sagen:

Das LGS hat keine Lösung.

Tja.

Nehmen wir ein anderes Beispiel.

Beispiel 34 ein anderes Beispiel

$$\left(\begin{array}{ccc|c} 1 & -1 & 2 & 6 \\ 2 & 3 & 3 & 5 \\ 2 & 8 & 0 & 2 \end{array}\right) \underset{(\mathrm{III}-2\,\mathrm{I})/2}{\overset{\mathrm{II}-2\,\mathrm{I}}{\longrightarrow}} \left(\begin{array}{ccc|c} 1 & -1 & 2 & 6 \\ 0 & 5 & -1 & -7 \\ 0 & 5 & -2 & -5 \end{array}\right)$$

$$\overset{\mathrm{III}-\mathrm{II}}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & -1 & 2 & 6 \\ 0 & 5 & -1 & -7 \\ 0 & 0 & -1 & 2 \end{array}\right)$$

Das kann man nun wieder als LGS umformulieren (in Gedanken eigentlich nur) zu

$$\begin{pmatrix} 1 & -1 & 2 \\ 0 & 5 & -1 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ -7 \\ 2 \end{pmatrix} \quad \Leftrightarrow \quad \begin{aligned} x_1 - x_2 + 2\,x_3 &= 6 \\ 5\,x_2 - x_3 &= -7 \\ -x_3 &= 2 \end{aligned}$$

Rekursives Auflösen führt auf

$$\left.\begin{aligned} x_3 &= -2 \\ x_2 &= \tfrac{1}{5}(-7 + x_3) = -\tfrac{9}{5} \\ x_1 &= 6 + x_2 - 2\,x_3 = 6 - \tfrac{9}{5} + 4 = \tfrac{41}{5} \end{aligned}\right\} \Rightarrow x = \begin{pmatrix} \frac{41}{5} \\ -\frac{9}{5} \\ -2 \end{pmatrix}$$

Also lautet die Lösungsmenge:

$$\mathbb{L} = \left\{ \frac{-1}{5} \begin{pmatrix} -41 \\ 9 \\ 10 \end{pmatrix} \right\}$$

Die Lösungsmenge beinhaltet genau einen Vektor, bzw. genau einen Punkt im $\mathbb{R}^3$. Wir sagen:

Das LGS ist eindeutig lösbar.

Beispiel 35 noch ein anderes Beispiel

$$\left(\begin{array}{ccc|c} 1 & -1 & 2 & 6 \\ 2 & 3 & 3 & 5 \\ 2 & 8 & 2 & -2 \end{array}\right) \underset{(\mathrm{III}-2\,\mathrm{I})/2}{\overset{\mathrm{II}-2\,\mathrm{I}}{\longrightarrow}} \left(\begin{array}{ccc|c} 1 & -1 & 2 & 6 \\ 0 & 5 & -1 & -7 \\ 0 & 5 & -1 & -7 \end{array}\right)$$

$$\overset{\mathrm{III}-\mathrm{II}}{\longrightarrow} \left(\begin{array}{ccc|c} 1 & -1 & 2 & 6 \\ 0 & 5 & -1 & -7 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$$\begin{pmatrix} 1 & -1 & 2 \\ 0 & 5 & -1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 6 \\ -7 \\ 0 \end{pmatrix} \quad \Leftrightarrow \quad \begin{aligned} x_1 - x_2 + 2\,x_3 &= 6 \\ 5\,x_2 - x_3 &= -7 \\ 0 &= 0 \end{aligned}$$

In der dritten Zeile können wir keinen Wert festlegen. Im Grunde steht da: "Welches $x_3 \in \mathbb{R}$ erfüllt $0 \cdot x_3 = 0$?" Die Antwort ist: "Alle Zahlen x_3 in $\mathbb{R}$ tun das." Also wählen wir für das rekursive Auflösen für x_3 einen Parameter t. Wir könnten im Grunde als Parametername x_3 stehen lassen. Es ist nur ein Name. Also weiter:

Rekursives Auflösen:

$$\left.\begin{aligned} x_3 &= t \\ x_2 &= \tfrac{1}{5}(-7+t) \\ x_1 &= 6 + x_2 - 2\,x_3 = 6 + \tfrac{1}{5}(-7+t) - 2\,t \end{aligned}\right\} \Rightarrow x = \begin{pmatrix} \frac{37}{5} - \frac{9}{5}\,t \\ -\frac{7}{5} + \frac{1}{5}\,t \\ t \end{pmatrix}$$

Also lautet die Lösungsmenge:

$$\mathbb{L} = \left\{ x \in \mathbb{R}^3 \;\middle|\; x = \frac{1}{5} \begin{pmatrix} 37 \\ -7 \\ 0 \end{pmatrix} + \begin{pmatrix} -9 \\ 1 \\ 5 \end{pmatrix} t\,,\; t \in \mathbb{R} \right\}$$

Da wir für t beliebig viele Werte einsetzen dürfen hat diese Lösungsmenge unendlich viele Elemente. Wir sagen:

Das LGS hat unendlich viele Lösungen.

Es gibt immer eine dieser drei Möglichkeiten:
Ein LGS hat keine, genau eine oder unendlich viele Lösungen!

Und wie sieht's denn nun mit unserem Straßengeviert aus? Das können wir ja auch mal durchrechnen:

Beispiel 36 Geviert, Teil 3/3 Wir können nun das LGS aus Beispiel 27 und 32 nach seinen Unbekannten auflösen.

$$\begin{aligned} x_2 + x_3 &= 200 \\ x_1 + x_2 &= 500 \\ x_3 + x_4 &= 300 \\ x_1 + x_4 &= 600 \end{aligned} \Leftrightarrow \left(\begin{array}{cccc|c} 1 & 1 & 0 & 0 & 500 \\ 1 & 0 & 0 & 1 & 600 \\ 0 & 1 & 1 & 0 & 200 \\ 0 & 0 & 1 & 1 & 300 \end{array}\right) \longrightarrow \left(\begin{array}{cccc|c} 1 & 1 & 0 & 0 & 500 \\ 0 & 1 & 0 & -1 & -100 \\ 0 & 0 & 1 & 1 & 300 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

$$\Rightarrow \quad \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 600 \\ -100 \\ 300 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}$$

Das LGS für sich hat also unendlich viele Lösungen. Soweit verhält es sich auf der abstrakten, mathematischen Ebene. Wenn man eine Problemstellung dann aber wieder in seinen realen Kontext überführt werden oft Lösungen als irrelevant ausgeschlossen. Zum Beispiel handelt es sich bei dem Geviert um Einbahnstraßen. Das bedeutet, dass alle Komponenten des Lösungsvektors nicht negativ (also ≥ 0) sind. Damit gilt

$$\begin{aligned} x_4 &\in [100, 300] \qquad & x_1 &\in [300, 500] \\ x_2 &\in [0, 200] \qquad & x_3 &\in [0, 200] \end{aligned}$$

Wir können dann etwa die folgenden, praktikablen Aussagen aus der Lösungsmenge ableiten:

(a) Eine Sperrung des Straßenstücks AD ist nicht möglich.

(b) Die minimale Verkehrsdichte auf dem Straßenstück AB ist 300.

(c) Die maximale Verkehrsdichte auf dem Straßenstück CD ist 200.

Aus praktischen Gründen erweitern wir unser Fachvokabular an dieser Stelle noch um ein paar Begriffe.

Definition 3.6 (Pivot, Rang, Regularität & Singularität)

- *Das jeweils erste nicht Nullelement in einer Stufenform einer Matrix* $A \in \mathbb{R}^{m\times n}$ *heißt* Pivotelement*.*
- *Die Anzahl der Pivotelemente in der Stufenform der Matrix* A *heißt* Rang *der Matrix* A *und wird mit* $\operatorname{rang}(A)$ *bezeichnet.*
- *Gilt bei einer quadratischen* $n \times n$*-Matrix* $\operatorname{rang}(A) = n$ *so heißt die Matrix* regulär*, andernfalls* singulär*.*

Wir sagen auch ein LGS hat den Rang $\operatorname{rang}(A)$. Die rechte Seite des LGS spielt dabei aber keine Rolle, auch wenn diese zum LGS an sich dazugehört.

Die Vorgehensweise:

<table>
<tr><th colspan="4">Vorgehen beim Lösen eines LGS</th></tr>
<tr><td colspan="4">1. System auf Stufenform bringen.</td></tr>
<tr><td colspan="4">2. Rangbestimmen $r = \operatorname{rang}(A)$, $A \in \mathbb{K}^{m\times n}$</td></tr>
<tr><td colspan="4">3. Gibt es Nullzeilen?</td></tr>
<tr><td colspan="2">ja: $r < m$</td><td colspan="2">nein: $r = m$</td></tr>
<tr><td colspan="2">Ist ein Element in $\{y_{r+1}, \ldots, y_m\}$ ungl. Null?</td><td colspan="2">$r < n$?</td></tr>
<tr><td>ja</td><td>nein</td><td>ja</td><td>nein</td></tr>
<tr>
<td>$\left(\begin{array}{ccccc|c} \star & & & & & \\ & \ddots & & & & \star \\ 0 & & \star & & \cdots & \cdots \\ 0 & \cdots & \cdots & & 0 & 0 \\ \vdots & & & & \vdots & \star \\ 0 & \cdots & \cdots & & 0 & 0 \end{array}\right)$
$\mathbb{L} = \emptyset$</td>
<td>$\left(\begin{array}{ccccc|c} \star & & & & & \\ & \ddots & & & & \star \\ 0 & & \star & & \cdots & \cdots \\ 0 & \cdots & \cdots & & 0 & 0 \\ \vdots & & & & \vdots & \vdots \\ 0 & \cdots & \cdots & & 0 & 0 \end{array}\right)$
$\operatorname{Dim} \mathbb{L} = n - r$</td>
<td>$\left(\begin{array}{cccccc|c} 0 & \cdots & 0 & \star & & \star & \star \\ \vdots & & \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & 0 & & \star & \star \end{array}\right)$
$\operatorname{Dim} \mathbb{L} = n - r$</td>
<td>$\left(\begin{array}{ccc|c} \star & & \star & \star \\ & \ddots & & \vdots \\ 0 & & \star & \star \end{array}\right)$
$\operatorname{Dim} \mathbb{L} = 0$</td>
</tr>
<tr><td>"Es gibt keine Lösung"</td><td>"Es gibt ∞ viele Lösungen"</td><td>"Es gibt ∞ viele Lösungen"</td><td>" Es gibt genau eine Lösung"</td></tr>
<tr><td colspan="4">4. Im Fall, dass es Lösungen gibt: Rekursives Auflösen.</td></tr>
</table>

Sie haben sicher bereits eine Idee was die Dimension eines Raumes beschreibt. Mathematisch werden wir das später im Kapitel 3.2 klären. Was Sie für's Erste wissen sollten ist

Die Anzahl der Matrixspalten minus die Anzahl der Pivotelemente der Stufenform der Matrix ergibt die Dimension der Lösungsmenge des lösbaren (!!!) LGS.

Durch das Anwenden von elementaren Zeilenumformungen kann eine reguläre Matrix nicht nur auf obere Dreiecksmatrix, sondern darüberhinaus bis auf Diagonalgestalt gebracht werden. Das LGS kann dann ohne Rekursion direkt aufgelöst werden. Schauen Sie:

Beispiel 37

$$\begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

$$\left(\begin{array}{cc|c} 2 & 1 & 1 \\ 3 & 4 & 3 \end{array}\right) \overset{\text{2II-3I}}{\longrightarrow} \left(\begin{array}{cc|c} 2 & 1 & 1 \\ 0 & 5 & 3 \end{array}\right) \overset{\text{5I-II}}{\longrightarrow} \left(\begin{array}{cc|c} 10 & 0 & 2 \\ 0 & 5 & 3 \end{array}\right) \underset{\text{II/5}}{\overset{\text{I/10}}{\longrightarrow}} \left(\begin{array}{cc|c} 1 & 0 & \frac{1}{5} \\ 0 & 1 & \frac{3}{5} \end{array}\right)$$

DIe erweiterte Matrix in Stufenform ist ja äquivalent zum LGS

$$\begin{aligned} 1 \cdot x_1 + 0 \cdot x_2 &= \frac{1}{5} \\ 0 \cdot x_1 + 1 \cdot x_2 &= \frac{3}{5}\,. \end{aligned}$$

Wir sehen, dass wir die Lösung nun direkt ablesen können. Sie lautet demnach:

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \frac{1}{5}$$

Jetzt kommt ein kleiner Trick! Nehmen wir einmal an wie wollten mehrere Gelichungssysteme lösen, die alle durch ein und dieselbe Matrix beschrieben werden. Etw sowas: Gegeben ist die Matrix $A \in \mathbb{R}^{3\times 3}$ sowie die drei rechten Seiten $y^1, y^2, y^3 \in \mathbb{R}^3$. Gesucht sind $x^1, x^2, x^3 \in \mathbb{R}^3$ mit

$$A\,x^1 = y^1\,,\ A\,x^2 = y^2\,,\ A\,x^3 = y^3\,.$$

Bei naivem Vorgehen würden wir jetzt für drei erweiterte Matrizen, nämlich

$$(A, y^1)\,,\ (A, y^2)\,,\ (A, y^3)$$

jeweils die Stufenform berechnen und im Anschluss feststellen, dass wir drei Mal die jeweils gleichen Zeilenumformungen durchgeführt haben. Diesen Prozess kann man leicht verküzen, indem man alle drei rechten Seiten in die erweiterte Matrix packt

$$(A, y^1, y^2, y^3)$$

und alle Rechnungen auf einen Schlag durchführt.

Beispiel 38

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} x_1^1 \\ x_2^1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} x_1^2 \\ x_2^2 \end{pmatrix} = \begin{pmatrix} 3 \\ 7 \end{pmatrix}$$

(Hier ist x_n^m die n-te Komponente von x^m.) Wir gehen dann so vor:

$$\left(\begin{array}{cc|cc} 1 & 2 & -1 & 3 \\ 3 & 4 & -1 & 7 \end{array}\right) \xrightarrow[\text{(I+II)}]{\text{II-3 I}} \left(\begin{array}{cc|cc} 1 & 0 & 1 & 1 \\ 0 & -2 & 2 & -2 \end{array}\right)$$
$$\xrightarrow{\text{II/(-2)}} \left(\begin{array}{cc|cc} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 1 \end{array}\right)$$

Die erste Spalte in der Erweiterung beinhaltet nun x^1, die Lösung von $A\,x^1 = y^1$ und die zweite Spalte der Erweiterung enthält x^2, die Lösung des zweiten Gleichungssystems $A\,x^2 = y^2$. Wir machen den Test:

$$A\,x^1 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix} = y^1$$

und

$$A\,x^1 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 7 \end{pmatrix} = y^2$$

✓

Warum wir das jetzt gemacht haben hat einen bestimmten Grund. Es geht weniger darum, wie man mehrere LGS mit gleicher Matrix auf einen Schlag lösen kann, sondern vielmehr geht es darum ein multiplikatives Inverses einer regulären Matrix, denn nur diese besitzt überhaupt eins, zu berechnen. Wie das?

Wir zäumen das Pferd von hinten auf:

Das neutrale Element bezüglich der Multiplikation einer quadratischen Matrix $A \in \mathbb{R}^{n\times n}$ ist ja E_n, denn

$$A\,E_n = A\,.$$

Angenommen es gäbe ein multiplikatives Inverses A^{-1}, dann wäre ja

$$A^{-1}\,A = E_n \quad \text{bzw.} \quad A\,A^{-1} = E_n\,.$$

Klar, oder?

$$A^{-1} \neq \frac{1}{A} \qquad \text{und} \qquad A\,x = y \quad \cancel{\Leftrightarrow \quad x = \frac{y}{A}}$$

"Durch eine Matrix teilen" bedeutet immer "mit ihrer (multiplikativen) Inversen multiplizieren", wenn es diese gibt. Im Grunde ist das im Reellen auch so.

Also angenommern wir hätten zur Matrix $A \in \mathbb{R}^{n\times n}$ unseres LGS $A\,x = y$ ein multiplikatives Inverses (wir sagen auch *inverse Matrix*) A^{-1} dann gilt doch

$$A\,x = y \quad \Leftrightarrow \quad A^{-1}A\,x = A^{-1}\,y \quad \Leftrightarrow \quad E_n\,x = A^{-1}\,y \quad \Leftrightarrow \quad x = A^{-1}\,y$$

(Beachten Sie, dass wir auf beiden Seiten des LGS A^{-1} jeweils von links multipliziert haben. Da die Matrixmultiplikation nicht kommutativ ist müssen wir peinlichst darauf achten.)

Wir erhalten nun - Achtung! Das ist ein Trick - als Lösung x^1 die erste Spalte von A^{-1}, wenn wir $y^1 = e_1 := (\delta_{1j}))_j$, den sogenannten *Einheitsvektor*, in unser LGS $A\,x^1 = y^1$ einsetzen. Für y^2 wählen wir e_2, usw. Sukkessives Fortführen liefert uns n LGSe

$$A\,x^1 = e_1\,,\ \ldots\,,\ A\,x^n = e_n$$

und jede Lösung x^i stellt die i-te Spalte von A^{-1} dar. Die gesamte Matrix A^{-1} können wir dann aus der Umformung

$$(A, E_n) \underset{\text{umformungen}}{\overset{\text{Zeilen-}}{\longrightarrow}} (E_n, A^{-1})$$

ablesen.

Beispiel 39 inverse Matrix

Wir suchen zu $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ das multiplikative Inverse A^{-1}.

$$(A, E_2) = \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right) \underset{\text{umformungen}}{\overset{\text{Zeilen-}}{\longrightarrow}} \left(\begin{array}{cc|cc} 1 & 0 & -2 & 1 \\ 0 & 1 & \frac{3}{2} & \frac{-1}{2} \end{array}\right) = (E_2, A^{-1})$$

Demnach gilt

$$A^{-1} = \frac{1}{2}\begin{pmatrix} -4 & 2 \\ 3 & -1 \end{pmatrix}$$

Ist $A \in \mathbb{R}^{n\times n}$ eine reguläre Matrix, so kann man das LGS $Ax = y$ auch mit Hilfe der Inversen A^{-1} berechnen:

$$Ax = y \qquad \Leftrightarrow \qquad x = A^{-1}y$$

Beispiel 40 Matrizen als Gruppen $(\mathbb{R}^{2\times 2}, \cdot)$ bildet mit der üblichen Matrixmultiplikation keine Gruppe, $(A, \cot)$ mit

$$A = \left\{A \in \mathbb{R}^{n\times n} \;\middle|\; \operatorname{rang}(A) = n\right\}$$

aber schon, d.h. die Menge der quadratischen, regulären Matrizen bilden mit der üblichen Matrixmultiplikation eine Gruppe (nicht abelsch!).

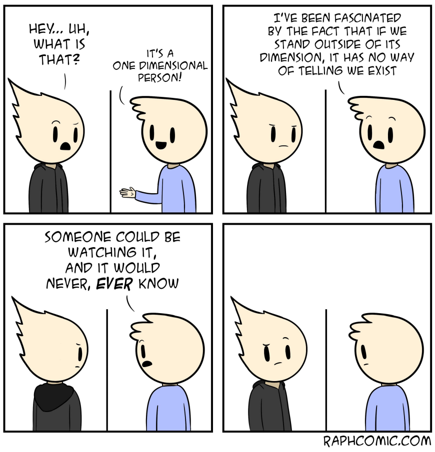
Hier noch eine kleine aber sehr nützliche Formel für $A \in \mathbb{R}^{2\times 2}$:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \Rightarrow \quad A^{-1} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

(Machen Sie den Test!)

Natürlich gibt es noch jede Menge weitere Matrizeneigenschaften aber für das, was wir als nächstes Lernen wollen genügt das mal. Wir werden uns tiefergreifend mit Matrizen zu gegebener Zeit beschäftigen.

3.2 Basis und Dimension von Vektorräumen



Vektorräume zeichnen sich durch die Eigenschaft der Dimension aus. Unter der Dimension versteht man die Anzahl sogenennter Basisvektoren, die mindestens erforderlich sind, um den ganzen Raum zu erzeugen.

Stellen wir uns ein Zelt vor. Mit zwei Stangen lässt es sich kaum zu einem komfortablen Raum aufspannen, mit drei schon. Eine vierte Stange kann man hinzunehmen, ist aber nicht nötig. Die vier Stangen sind ein Erzeugendensystem des Raums, drei Stangen (die in verschiedene Richtungen zeigen) bilden ein minimales Erzeugendensystem, wären damit eine Basis und zwei genügen nicht. Die Dimension unseres "Zeltraums" ist dann drei.

Wir werden sehen, dass die Wahl einer Basis nicht eindeutig ist und je nach Anwendung ist eine Wahl einer Basis "schlauer" als eine andere Wahl.

Sie haben vielleicht gerade schon bemerkt, dass es von besonderer Bedeutung ist in welche Richtungen unsere Zeltstangen zeigen. In der Mathematik verwenden wir den Begriff der linear (Un-)Abhängigkeit. Da zum Beispiel auch Polynome Vektoren sein können ist der Begriff "Richtung" eigentlich nicht der naheliegendste.

(Am Ende des Kapitels sollten Sie verstehen was "outside of its dimensions" bedeutet und das Comic diskutieren können.)

Definition 3.7 (lineare (Un-) Abhängigkeit) *Die Vektoren $v_1, \ldots, v_n \in V$ heißen* linear abhängig (la) *falls es Skalare $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ gibt, die nicht alle gleich Null sind mit*

$$\sum_{l=1}^{n} \lambda_l \, v_l = 0 \, .$$

Die Vektoren heißen linear unabhängig (lu) *wenn sie nicht linear abhängig sind.*

Zur besseren Lesbarkeit schreiben wir Skalare meistens mit griechischen Buchstaben; das aber nicht zwingend.

Da $\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$ für $\lambda_1 = \cdots = \lambda_n = 0$ immer erfüllt ist, sucht man bei der Frage nach linear Abhängigkeit nach einem $\lambda_i \neq 0$.

Beispiel 41 linear un-/abhängig

(a) Gegeben seien die Vektoren

$$a = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \; b = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}, \; c = \begin{pmatrix} -1 \\ 0 \\ 4 \end{pmatrix} \quad \text{und} \quad d = \begin{pmatrix} 0 \\ 2 \\ 7 \end{pmatrix}.$$

Es sind

$\{a, b\}$ sind linear abhängig (la), denn es gilt

$$2 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 6 \end{pmatrix}, \text{d.h. } 2\,a = b\,.$$

$\{a, c\}$ sind linear unabhängig (lu), denn es gilt: $\nexists t \in \mathbb{R}$:

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

$\{a, c, d\}$ sind paarweise (pw) linear unabhängig zusammen aber linear abhängig. a und c sind lu, a und d sowie c und d auch, aber

$$1 \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + 1 \cdot \begin{pmatrix} -1 \\ 0 \\ 4 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 0 \\ 2 \\ 7 \end{pmatrix} = 0$$

(b) Gegeben seien die Vektoren

$$p_1(x) = 2\,x + 1\,,\; p_2(x) = x + x^2\,,\; p_3(x) = -x^2 + 3\,x + 1 \in \mathbb{P}_2\,.$$

Wir prüfen p_1, p_2 und p_3 auf linear Abhängigkeit:

$$
\begin{aligned}
& & \lambda_1\, p_1(x) + \lambda_2\, p_2(x) + \lambda_3\, p_3(x) &= 0 \\
\Leftrightarrow & & \lambda_1\,(2\,x+1) + \lambda_2\,(x+x^2) + \lambda_3\,(-x^2+3\,x+1) &= 0 \\
\Leftrightarrow & & x^2\,\underbrace{(\lambda_2-\lambda_3)}_{\stackrel{!}{=}\,0} + x\,\underbrace{(2\,\lambda_1+\lambda_2+3\,\lambda_3)}_{\stackrel{!}{=}\,0} + \underbrace{(\lambda_1+\lambda_3)}_{\stackrel{!}{=}\,0} &= 0 \\
\Leftrightarrow & & \begin{pmatrix} 0 & 1 & -1 \\ 2 & 1 & 3 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
\Leftrightarrow & & \begin{pmatrix} 0 & 1 & -1 \\ 2 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}
\end{aligned}
$$

Hier ist die Lösungsmenge

$$\mathbb{L} = \{0\}$$

also gilt $\lambda_1 = \lambda_2 = \lambda_3 = 0$ und keine anderen Lambdas kommen in Frage. Damit sind p_1, p_2, p_3 lu.

(c) Für welches a sind die Vektoren

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}, \begin{pmatrix} 9 & 10 \\ 11 & a \end{pmatrix}$$

linear abhängig? (Jaja, Matrizen als Elemente eines Vektorraums $\mathbb{R}^{n\times n}$ sind auch Vektoren!) Wir prüfen für welche γ_1 , γ_2 γ_3 folgendes LGS erfüllt ist:

$$
\begin{aligned}
& & \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\gamma_1 + \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}\gamma_2 + \begin{pmatrix} 9 & 10 \\ 11 & a \end{pmatrix}\gamma_3 &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\
\Leftrightarrow & & \begin{pmatrix} \gamma_1+5\gamma_2+9\gamma_3 & 2\gamma_1+6\gamma_2+10\gamma_3 \\ 3\gamma_1+7\gamma_2+11\gamma_3 & 4\gamma_1+8\gamma_2+a\gamma_3 \end{pmatrix} &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\
\Leftrightarrow & & \begin{array}{rl} \gamma_1+5\gamma_2+9\gamma_3 & =0 \\ 2\gamma_1+6\gamma_2+10\gamma_3 & =0 \\ 3\gamma_1+7\gamma_2+11\gamma_3 & =0 \\ 4\gamma_1+8\gamma_2+a\gamma_3 & =0 \end{array} & \\
\Leftrightarrow & & \begin{pmatrix} 1 & 5 & 9 \\ 2 & 6 & 10 \\ 3 & 7 & 11 \\ 4 & 8 & a \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}
\end{aligned}
$$

Aus

$$\left(\begin{array}{ccc|c} 1 & 5 & 9 & 0 \\ 2 & 6 & 10 & 0 \\ 3 & 7 & 11 & 0 \\ 4 & 8 & a & 0 \end{array}\right) \xrightarrow[\text{umformungen}]{\text{Zeilen-}} \left(\begin{array}{ccc|c} 1 & 5 & 9 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & \frac{a-36}{-12} - 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

folgern wir dann die la für

$$a = 12$$

Die Vorgehensweise:

1. Linearkombination der zu untersuchenden Vektoren (Tupel, Polynome, Matrizen, whatever) gleich Null setzen.

2. Das entstandene homogene LGS lösen.

3. Beinhaltet die Lösungsmenge mehr als die Null so sind die Vektoren la, andernfalls lu.

Bei zwei Vektoren im $\mathbb{R}^2$ bedeutet linear Unabhängigkeit, dass die Vektoren in verschiedene Richtungen zeigen. Drei Vektoren im $\mathbb{R}^2$ können *paarweise (pw)* lu aber zusammen dennoch la sein...... jaja.
Linear Unabhängigkeit erkennt man bei Vektoren im $\mathbb{R}^n$ direkt, wenn es bei verschiedenen Komponenten im Tupel Nullen gibt. Haben Sie bereits ein Bild dazu?

Satz 3.8 *Eine Menge von Vektoren ist genau dann linear abhängig, wenn sich einer von ihnen als Linearkombination der jeweils anderen darstellen lässt.*

Beweis auf Seite 158

Es ist immer die Menge aller Vektoren, die einen Raum aufspannen das Erzeugendensystem dieses Raums. Die Frage, die sich stellt, ist die, ob alle angegebenen Vektoren auch nötig sind, um diesen Raum aufzuspannen? Gibt es also welche darunter, die überflüssig sind und die man folglich ignorieren kann?

Definition 3.9 (Erzeugendensystem & Basis) *Seien $v_1, \ldots, v_m \in V$. Dann bezeichnet die Menge*

$$U = \operatorname{Span}(v_1, \ldots, v_m) \subseteq V$$

die Menge aller Linearkombinationen aus v_1 bis v_m. Wir nennen U Spann *oder auch* lineare Hülle*. Die Vektoren $v_1, \ldots, v_m$ heißen* Erzeugendensystem *von U.*

Ist das Erzeugendensystem $v_1, \ldots, v_n$ minimal, das heißt es sind alle v_i linear unabhängig, so nennen wir $v_1, \ldots, v_n$ eine Basis *von U. Die Anzahl der Basisvektoren bestimmt die* Dimension *des Raumes U. Wir schreiben*

$$\operatorname{Dim}(U) = n\,.$$

Beispiel 42

Im Beispiel 41 (a) stellt $\{a, c, d\}$ ein Erzeugendensystem des Raums

$$S = \operatorname{Span}(a, b, c, d)\ ,\ \operatorname{Dim}(S) = ?$$

dar aber keine Basis, da die Vektoren a, b, c, d la sind. Wir wissen deshalb nicht welche Dimenson der Raum S hat. Hingegen sind $\{a, c\}$, $\{a, d\}$, $\{b, c\}$ und $\{d, c\}$ jeweils Basen und die entsprechend aufgespannten Räume haben Dimension zwei. Es gilt:

$$\operatorname{Span}(a, c, d) = \operatorname{Span}(a, c) = \operatorname{Span}(a, d) = \operatorname{Span}(d, c)$$

und

$$\operatorname{Span}(a, b, c) = \operatorname{Span}(a, c) = \operatorname{Span}(b, c) \neq \operatorname{Span}(a, b)\ .$$

Eine Basis ist immer auch ein Erzeugendensystem; ein Erzeugendensystem, aber nicht zwingend eine Basis. Wenn wir bei der Basis ein Element entfernen, so ändert sich der erzeugte Raum. Bei einem Erzeugendensystem hingegen ist es so, sofern linear abhängige Vektoren vorhanden sind, dass das Entfernen bestimmter Vektoren, nämlich der linear abhängigen, sich auf den aufgepannten Raum nicht auswirkt.

Beispiel 43 Basen

$$\mathbb{R}^2 = \operatorname{Span}\left(\begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\right)$$

und

$$\mathbb{R}^2 = \operatorname{Span}\left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right)$$

Die Basisdarstellung eines Vektorraums ist nicht eindeutig.

Nun können wir einem Vektorraum eine Dimension zuordnen. Wir wissen was Basen sind und das wir diese verschiedentlich wählen können. Wir wissen auch, dass etwa zwei Vektoren

$$v_1, v_2 \in \mathbb{R}^3\,,$$

die lu sind einen zweidimensionalen Raum aufspannen, also

$$\operatorname{Dim}\operatorname{Span}(v_1, v_2) = 2$$

aber wir wissen zunächst nicht ob das dann auch der gesamte $\mathbb{R}^2$ ist, bzw. wie $\operatorname{Span}(v_1, v_2)$ genau aussieht. Machen wir dazu am besten ein Beispiel:

Beispiel 44

1. Es sei
$$P = \operatorname{Span}(p_1, p_2, p_3) \subseteq \mathbb{P}_3$$
mit
$$p_1(x) = x^3 + 1\,,\ p_2(x) = x + x^2\,,\ p_3(x) = 1 + x - x^2\,.$$
Die Frage ist nun: "Wie sieht P aus?" oder anders formuliert: "enthält P alle Elemente von $\mathbb{P}_3$?" Nochmal anders formuliert: "Kann man jedes kubische Polynom durch die Vektoren p_1, p_2 und p_3 darstellen?" Um diese Frage zu beantworten probieren wir es einfach aus. Es sei
$$q(x) = a_0 + a_1\,x + a_2\,x^2 + a_3\,x^3 \in \mathbb{P}_3$$
ein beliebiges Polynom dritten Grades. Wir versuchen es mit den p_i, $i = 1, 2, 3$ darzustellen und schauen was passiert:
$$\begin{aligned} & p_1\,\lambda_1 + p_2\,\lambda_2 + p_3\,\lambda_3 = a_0 + a_1\,x + a_2\,x^2 + a_3\,x^3 \\ \Leftrightarrow\quad & (x^3+1)\,\lambda_1 + (x+x^2)\,\lambda_2 + (1+x-x^2)\,\lambda_3 = a_0 + a_1\,x + a_2\,x^2 + a_3\,x^3 \end{aligned}$$
Koeffizientenvergleich
$$\Leftrightarrow\quad x^3\,\underbrace{(\lambda_3 - a_3)}_{=0} + x^2\,\underbrace{(\lambda_2 - \lambda_3 - a_2)}_{=0} + x\,\underbrace{(\lambda_2 + \lambda_3 - a_1)}_{=0} + \underbrace{(\lambda_1 + \lambda_3 - a_0)}_{=0} = 0$$
$$\Rightarrow\quad \begin{aligned} \lambda_3 &= a_3 \\ \lambda_2 - \lambda_3 &= a_2 \\ \lambda_2 + \lambda_3 &= a_1 \\ \lambda_1 + \lambda_3 &= a_0 \end{aligned}$$

Nun haben wir wieder ein LGS zu lösen:

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & a_3 \\ 0 & 1 & -1 & a_2 \\ 0 & 1 & 1 & a_1 \\ 1 & 0 & 1 & a_0 \end{array}\right) \underset{\text{umformungen}}{\overset{\text{Zeilen-}}{\longrightarrow}} \left(\begin{array}{ccc|c} 1 & 0 & 0 & a_3 \\ 0 & 1 & -1 & a_2 \\ 0 & 0 & 1 & a_0 - a_3 \\ 0 & 0 & 0 & 2\,a_0 - 2\,a_3 - a_1 + a_2 \end{array}\right)$$

Was bedeutet das jetzt? Es bedeutet, dass Polynome mit der Eigenschaft $2\,a_0 - 2\,a_3 - a_1 + a_2 \neq 0$ nicht durch die Basis von P dargestellt werden können. Etwa $q(x) = x^3 + x^2 + x + 1 \notin P$, aber $q \in \mathbb{P}_3$. Es ist also P echt in $\mathbb{P}_3$ enthalten:

$$P \subset \mathbb{P}_3$$

Es ist also

$$\operatorname{Dim}(P) = 3$$

wärend

$$\operatorname{Dim}(\mathbb{P}_3) = \operatorname{Dim}\left(\operatorname{Span}\left(1, x, x^2, x^3\right)\right) = 4$$

gilt. Letzteres dürfen Sie sich gerne selbst überlegen.

2. Kommen wir zu einer sehr spannenden Frage: Welche Dimension hat der Raum der komplexen Zahlen $\mathbb{C}$? Ich schätze, dass Sie jetzt an die 2 denken. Ist auch nicht ganz falsch. Falsch ist erst mal die Frage, denn es fehlt eine Information. Die Dimension von $\mathbb{C}$ als VR hängt entscheidend von der Wahl des Körpers, über dem der VR definiert ist, ab.

   Denken wir zunächst an den VR $(\mathbb{C}, \mathbb{R})$, d.h. die komplexen Zahlen über den reellen Zahlen. Darin ist

   $$\mathcal{B} = \{1, i\}$$

   eine Basis. 1 und i sind lu, denn

   $$1 = \alpha\, i$$

   würde auf

   $$\alpha = -i \notin \mathbb{R}$$

   führen was aber keine reelle Zahl ist. Damit ist

   $$\operatorname{Dim}(\mathbb{C}) = 2\,.$$

   .

   Betrachten wir aber $(\mathbb{C}, \mathbb{C})$, d.h. komplexe Zahlen über $\mathbb{C}$, so sind mit obiger Rechnung 1 und i sehr wohl la. Wir können 1 oder i als Basis das ganzen Raums nehmen, etwa

   $$\mathcal{B} = \{1\}\,.$$

   Und dann ist eben

   $$\operatorname{Dim}(\mathbb{C}) = 1\,.$$

   Lustig, oder?

Die Vorgehensweise:

1. Linearkombination der angegebenen Basis/Erzeugendensystem gleich einem beliebigen Vektor des Raumes setzen.
2. Das entstandene inhomogene LGS lösen.
3. Prüfen ob die Parameter des beliebigen Vektors frei wählbar sind oder einer Einschränkung unterliegen.

Die beiden Vorgehensweisen bei der Untersuchung von lu/la und des Raums der aufgespannt wird können in einem Aufwasch erledigt werden.

Definition 3.10 (Unter(vektor)raum (UVR)) *Sei V ein Vektorraum. Eine Teilmenge $U \subseteq V$ heißt* Untervektorraum *oder kurz* Unterraum, *falls U selbst wieder ein Vektorraum ist.*

Es genügt zu prüfen, dass $\forall\, a, b \in U \;:\; a + b \in U$ und $\alpha\, a \in U$ gilt. Alle anderen Vektorraumeigenschaften übertragen sich automatisch von V auf U. Wir erhalten etwa automatisch einen UVR wenn wir als Basis eine Teilmenge der Raumbasis verwenden.

$$V = \mathrm{Span}\,(v_1, \ldots, v_n) \supset U = \mathrm{Span}\,(v_1, \ldots, v_{m \leq n})$$

Definition 3.11 (spezielle UVRe) *Seien $v, w \in \mathbb{R}^n$.*
Die Parameterdarstellung der Geraden *durch den Ursprung in Richtung v lautet*

$$g := \{x \in \mathbb{R}^n \mid x = t\, v\,,\; t \in \mathbb{R}\}\,.$$

g heißt Ursprungsgerade. *Wir schreiben kurz:*

$$g \,:\, t\, v$$

Seien v, w lu. Die Parameterdarstellung der Ebenen *durch den Ursprung in Richtung v und w wird beschrieben durch*

$$E := \{x \in \mathbb{R}^3 \mid x = t\, v + s\, w\,,\; t, s \in \mathbb{R}\}\,.$$

E heißt Ursprungsebene. *Wir schreiben kurz:*

$$E \,:\, v\, t + w\, s$$

Beispiel 45 UVRe

(a)

$$g\,:\,t\binom{2}{1}$$

ist die Menge bestehend aus allen Punkten im $\mathbb{R}^2$, die zu $\binom{1}{2}$ linear abhängig sind. Die Lösungsmenge eines homogenen LGS mit einer Nullzeile hat zum Beispiel die geometrische Form einer Geraden. g ist eine Ursprungsgerade, denn $0 \in g$ (für $t = 0$).

(b)

$$E\,:\,\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} t + \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} s$$

(c)

$$M = \left\{ A \in \mathbb{R}^{2\times 2} \,\middle|\, A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + t \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},\ t \in \mathbb{R} \right\}$$

enthält nicht die Nullmatrix. Sie können auch gerne mal die Abgeschlossenheit prüfen:

$$\left.\begin{matrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \in M \quad (t=0) \\ \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \in M \quad (t=1) \end{matrix}\right\} \quad \text{aber} \quad \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 0 & 0 \end{pmatrix} \notin M\,,$$

denn

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + t \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 0 & 0 \end{pmatrix} \quad \Leftrightarrow \quad \begin{pmatrix} 0 & t \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

und das ist ein Widerspruch.

Definition 3.12 (spezielle affine Räume) *Seien $a, v, w \in \mathbb{R}^n$.*
Die Parameterdarstellung der Geraden *durch den Punkt a in Richtung v lautet*

$$g := \{x \in \mathbb{R}^n \mid x = a + t\,v\,,\ t \in \mathbb{R}\}\,.$$

a heißt Stützvektor *der Geraden. Wir schreiben kurz:*

$$g\ :\ a + t\,v$$

Seien v, w lu. Die Parameterdarstellung der Ebenen *durch den Punkt a in Richtung v und w wird beschrieben durch*

$$E := \{x \in \mathbb{R}^3 \mid x = a + t\,v + s\,w\,,\ t, s \in \mathbb{R}\}\,.$$

a heißt Stützvektor *der Ebene. Wir schreiben kurz:*

$$E\ :\ a + v\,t + w\,s$$

Beispiel 46 affine Räume

(a)

$$g\ :\ \begin{pmatrix}0\\1\end{pmatrix} + t\begin{pmatrix}2\\1\end{pmatrix}$$

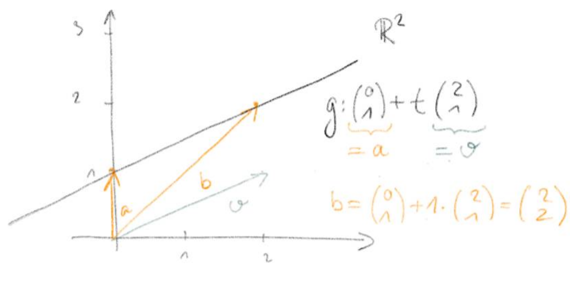
ist eine Gerade im $\mathbb{R}^2$, die durch den Punkt $(0,1)$ und $(2,2)$ ($t \in \{0,1\}$) veräuft. g ist eine Teilmenge des $\mathbb{R}^2$ aber selbst kein VR, denn die Abgeschlossenheit gilt nicht: Für die Punkte $(0,1)\,,\ (2,2) \in g$

gilt

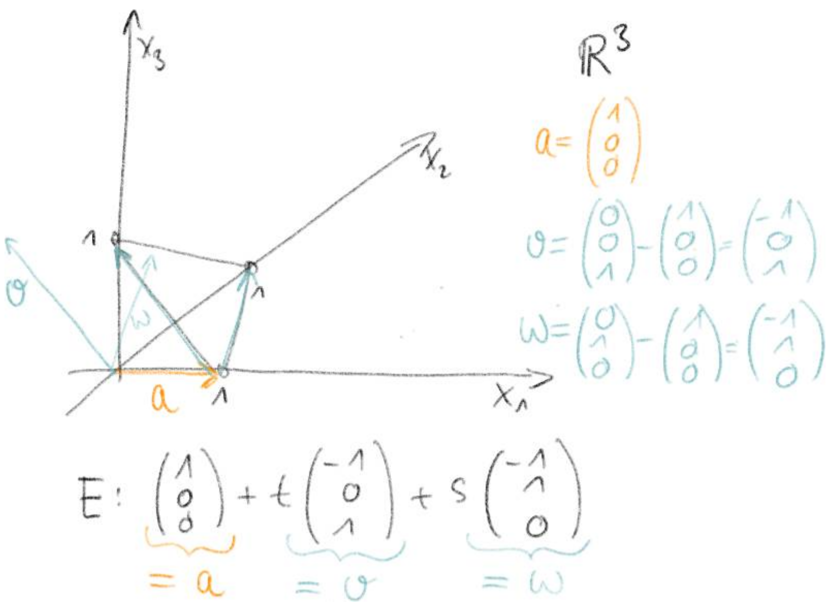
$$\begin{pmatrix}0\\1\end{pmatrix} + \begin{pmatrix}2\\2\end{pmatrix} = \begin{pmatrix}2\\3\end{pmatrix} \notin g\,,$$

denn

$$\begin{aligned}
& \begin{pmatrix}0\\1\end{pmatrix} + t\begin{pmatrix}2\\1\end{pmatrix} = \begin{pmatrix}2\\3\end{pmatrix} \\
\Leftrightarrow\quad & t\begin{pmatrix}2\\1\end{pmatrix} = \begin{pmatrix}2\\2\end{pmatrix} \\
\Leftrightarrow\quad & t = 1\ \wedge\ t = 2\ \Rightarrow \text{Wds}
\end{aligned}$$



(b)



(c)

$$M = \{f \in \mathbb{P}_1 \mid f(x) = 1 + t\,x\,,\ t \in \mathbb{R}\}$$

ist kein VR, denn es enthält die Null nicht:

$$1 + t\,x = 0 \quad \Leftrightarrow \quad t = \frac{-1}{x} \quad \Rightarrow \quad \text{Wds}\,,$$

denn $t \neq t(x)$, d.h. t hängt nicht von x ab. Die Aussagen müssen jeweils für alle x gelten.

Affine Räume sind Teilmengen von VRe. Sie können erst dann selbst VRe darstellen, wenn Sie die Null enthalten. Viele Formeln gelten nur auf VRen und liefern falsche Ergebnisse, wenn man sie unbedacht auf affine Räume anwendet.

Wir wollen zum Abschluss des Unterkapitels noch ein paar Zusammenhänge zwischen UVRe vom $\mathbb{R}^n$ und Lösungsmengen von LGSen herstellen. Bleiben wir dazu der Einfachheit halber im $\mathbb{R}^3$, weil es dort anschaulich ist. Das Phänomen selbst lässt sich problemlos auf höhere Dimensionen skalieren.

Zunächst betrachten wir eine Eben im $\mathbb{R}^3$ in einer anderen Darstellung, nämlich

Definition 3.13 (Koordinatendarstellung einer Ebene) *Seien $\alpha, \beta, \gamma \in \mathbb{R}$, dann ist durch*

$$E := \{x \in \mathbb{R}^3 \mid \alpha\,x_1 + \beta\,x_2 + \gamma\,x_3 + \delta = 0\}$$

eine Koordinatendarstellung der Ebene E in $\mathbb{R}^3$ gegeben.

Zur Ebene gehören alle Punkte des $\mathbb{R}^3$, die diese Gleichung erfüllen. Sieht etwas ander aus als die Parameterdarstellung aber wir werden gleich sehen, dass sich die eine Darstellung jeweils in die andere umformulieren lässt. Die Koordinatendarstellung beinhaltet eine Zeile eines LGS. Wir könnten es ja auch so schreiben

$$\begin{pmatrix} \alpha & \beta & \gamma \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -\delta \\ 0 \\ 0 \end{pmatrix}$$

und haben direkt schon die Stufenform. Und die Lösungsmenge des LGS hat die Form

$$\mathbb{L} \, : \, a + t\,v + s\,w \, ,$$

worin wir direkt die Parameterform der Ebene erkennen. Den Stützvektor a erhalten wir weil wir es mit einem inhomogenen LGS zu tun haben (0 ist keine Lösung, auch nicht in E enhalten).

Lösungsmengen von homogenen LGSen sind UVRe des $\mathbb{R}^n$ und solche von inhomogenen LGSen sind affine Räume.

Beispiel 47 $\quad \mathbb{L}$ homogener LGS ($Ax = 0$) Lösungen von homogenen LGS, d.h.

$$A\,x = 0$$

mit $A \in \mathbb{R}^{m\times n}$ sind UVRe des $\mathbb{R}^n$ und haben die Form:

$$\mathbb{L} = \left\{ x \in \mathbb{R}^n \,\middle|\, x = \sum_{l=1}^{L} v_l\,t_l = \begin{pmatrix} v_1^1 \\ \vdots \\ v_1^n \end{pmatrix} t_1 + \cdots + \begin{pmatrix} v_L^1 \\ \vdots \\ v_L^n \end{pmatrix} t_L \right\} ,$$

wobei $v_1, \ldots, v_L \in \mathbb{R}^n$ lu sind, d.h. eine Basis der Lösungsmenge bilden.

$$\mathrm{Dim}\,(\mathbb{L}) = L < n$$

L entspricht dabei der Differenz von Anzahl Spalten von A mit Anzahl Pivotelemente der Stufendorm von A. Ist A eine quadratische Matrix, so entspricht L gerade der Anzahl der Nullzeieln in der Stufenform. Gehen Sie gerne nochmal durch die LGS-Beispiele und überprüfen Sie das.

Beispiel 48 $\quad \mathbb{L}$ inhomogener LGS ($Ax \neq 0$) Lösungen von inhomogenen LGS, die lösbar sind (!!!) d.h.

$$A\,x = y$$

mit $A \in \mathbb{R}^{m\times n}$ sind UVRe des $\mathbb{R}^n$ und haben die Form:

$$\mathbb{L} = \left\{ x \in \mathbb{R}^n \;\middle|\; x = a + \sum_{l=1}^{L} v_l \, t_l = \underbrace{\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}}_{\text{Stützvektor}} + \underbrace{\begin{pmatrix} v_1^1 \\ \vdots \\ v_1^n \end{pmatrix} t_1 + \cdots + \begin{pmatrix} v_L^1 \\ \vdots \\ v_L^n \end{pmatrix} t_L}_{\text{lineare Hülle}} \right\},$$

wobei $a, v_1, \ldots, v_L \in \mathbb{R}^n$ lu sind, d.h. eine Basis der Lösungsmenge bilden.

$$\mathrm{Dim}\,(\mathbb{L}) = L < n$$

Suchen wir die Schnittmenge zweier Ebenen (hier in Koordinatendarstellung), d.h. wir suchen alle $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, die die folgenden beiden Gleichungen erfüllen

$$\begin{aligned} E_1 &: \; a\,x_1 + b\,x_2 + c\,x_3 + d \\ E_2 &: \; e\,x_1 + f\,x_2 + g\,x_3 + d \end{aligned}$$

so haben wir es mit einem inhomogenen LGS im $\mathbb{R}^3$ zu tun, das entweder keine Lösung hat (parallele Ebenen) oder unendlich viele, nämlich die Schnittgerade, die sie als Lösungsmenge erhalten bei einer Nullzeile in der Stufenform.

Hübsch nicht wahr?

3.3 Darstellung von Vektoren

Ein Koordinatensystem, welches durch die Einheitsvektoren (hier eine Basis des $\mathbb{R}^n$), auch *kanonische Basis* genannt,

$$\mathcal{E} = \{\mathrm{e}_1, \ldots, \mathrm{e}_n\}$$

mit $\mathrm{e}_i = (\delta_{ij})_{1 \le j \le n} \in \mathbb{R}^n$ (das ist ein Vektor, dessen i-te Komponente eine Eins ist und alle anderen Nullen sind) erzeugt wird heißt *kartesisches Koordinatensystem*. Es ist im Grunde genau das, mit welchem wir bisher gearbeitet haben. Jeder Vektor a eines kartesischen Koordinatensystems wird dargestellt als Linearkombination der Basisvektoren (hier e_i) und den zugehörigen Koordinaten a_i ($i = 1, \ldots, n$):

$$a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = a_1 \underbrace{\begin{pmatrix} 1 \\ \vdots \\ 0 \end{pmatrix}}_{=e_1} + \cdots + a_n \underbrace{\begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix}}_{=e_n} = a_1\,e_1 + \cdots + a_n\,e_n = \sum_{i=1}^{n} a_i \mathrm{e}_i$$

Der Vektor

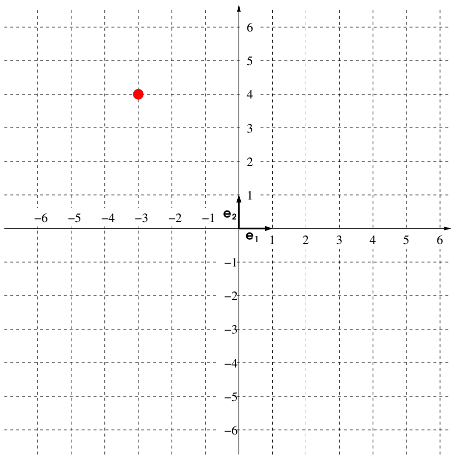
$$a_{\mathcal{E}} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

beschreibt dann den zu a gehörenden Koordinatenvektor bezüglich der Basis $\mathcal{E}$. Bei der Wahl der kaninonischen Basis entält der Koordinatenvektor zu a die gleichen Komponenten wie a selbst. Da fragt man sich natürlich wozu es einen neuen Begriff zur gleichen Darstellung braucht! Wir werden es erkennen, wenn wir mit anderen Basen oder auch in anderen Räumen als der $\mathbb{R}^n$ unterwegs sind.

Beispiel 49

Es sei

$$\mathcal{A} = \{e_1, e_2\} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$$



eine Basis des $\mathbb{R}^2$, bestehend aus den Einheitsvektoren e_1 und e_2. Dann ist der Koordinatenvektor $a_{\mathcal{A}}$ zum Vektor

$$a = \begin{pmatrix} -3 \\ 4 \end{pmatrix} = -3\,e_1 + 4\,e_2$$

gegeben durch

$$a_{\mathcal{A}} = \begin{pmatrix} -3 \\ 4 \end{pmatrix}.$$

Nun ist es aber so, dass die Basis eines VRs nicht eindeutig ist. Wir können den Vektorraum mit den unterschiedlichsten Basen darstellen. Wählen wir als Basis für den $\mathbb{R}^2$ etwa eine andere Basis, so ändert sich auch die Darstellung eines Vektors a. Damit ist gemeint, dass die Linearkombination eine andere ist. Der dargestellte Vektor a bleibt der gleiche.

Definition 3.14 (Koordinaten & Koordinatenvektor) *Sei $\mathcal{V} = \{v_1, \ldots, v_n\}$ eine Basis von V und $v \in V$ mit der eindeutigen Darstellung*

$$v = \sum_{k=1}^{n} \lambda_k v_k\,.$$

Die eindeutigen Skalare $\lambda_1, \ldots, \lambda_n$ heißen Koordinaten des Vektors v bezüglich der Basis $\mathcal{V}$ und

$$v_{\mathcal{V}} = (\lambda_1\,,\ \ldots\,,\ \lambda_n)^T \in \mathbb{R}^n$$

heißt Koordinatenvektor von v bezüglich $\mathcal{V}$.

Beispiel 50 affines Koordinatensystem

(a)

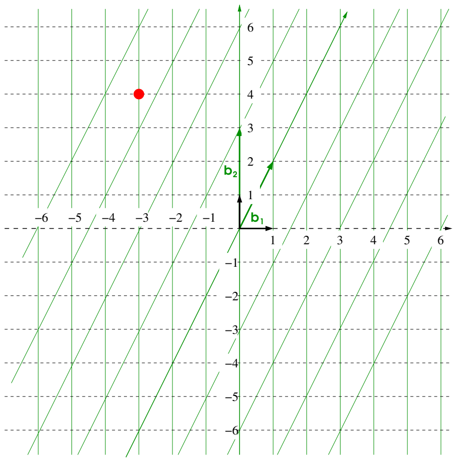
Es sei

$$\mathcal{B} = \{b_1, b_2\} = \left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \end{pmatrix} \right\}$$

eine Basis des $\mathbb{R}^2$. Dann lässt sich der Vektor

$$a = \begin{pmatrix} -3 \\ 4 \end{pmatrix} \in \mathbb{R}^2$$

als Linearkombination der Basisvektoren $\mathcal{B}$ wie folgt darstellen:



$$a = \begin{pmatrix} -3 \\ 4 \end{pmatrix} = \lambda_1 \, b_1 + \lambda_2 \, b_2 = \lambda_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}$$

Es ist also kurzum: $a = B\,\lambda \;\Leftrightarrow\; \lambda = B^{-1}a$, also[3]

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}}_{=:B}^{-1} \begin{pmatrix} -3 \\ 4 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 3 & 0 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} -3 \\ 4 \end{pmatrix} = \begin{pmatrix} -3 \\ \frac{10}{3} \end{pmatrix}$$

Das bedeutet, dass der Koordinatenvektor $a_{\mathcal{B}}$ von a bezüglich der Basis $\mathcal{B}$ gegeben ist durch

$$a_{\mathcal{B}} = \begin{pmatrix} -3 \\ \frac{10}{3} \end{pmatrix}.$$

Anschaulich bedeutet das: Wenn wir vom Ursprung aus -3 Mal in Richtung b_1 und dann $\frac{10}{3}$ Mal in Richtung b_2 "laufen" so erreichen wir genau den Punkt $a = (-3, 4)$.

(b) Es sei

$$\mathcal{P} = \{x^2 + 1\,,\; 2\,x - 1\} = \{p_1, p_2\}$$

eine Basis[4] von

$$P = \mathrm{Span}\,(p_1, p_2)\;.$$

Wie lautet jetzt etwa der Koordinatenvektor $q_{\mathcal{P}} \in \mathbb{R}^2$ von

$$q(x) = -x^2 + 4\,x - 3$$

[3] Wenn Sie hier vor lauter Wald die Bäume nicht mehr sehen dann nehmen Sie drei Farbstifte (rot, grün und gelb) und unterstreichen alle Terme wie zum Kapiteleingang: grün ist Basis, rot ist Koordinate und gelb Vektor

[4] Prüfen Sie zur Übung, dass die Elemente lu sind.

bezüglich der Basis[5] $\mathcal{P}$

$$\begin{aligned}
& & \mu_1\, p_1 + \mu_2\, p_2 &= q \\
&\Leftrightarrow & \mu_1\,(x^2+1) + \mu_2\,(2\,x-1) &= -x^2 + 4\,x - 3 \\
&\Leftrightarrow & x^2(\mu_1) + x(2\,\mu_2) + (\mu_1 - \mu_2) &= -x^2 + 4\,x - 3 \\
&\Leftrightarrow & x^2\,\underbrace{(\mu_1+1)}_{=0} + x\,\underbrace{(2\,\mu_2-4)}_{=0} + \underbrace{(\mu_1-\mu_2+3)}_{=0} &= 0 \\
&\Rightarrow & \mu_1 = -1 \;\wedge\; \mu_2 &= 2 \\
&\Rightarrow & &
\end{aligned}$$

$$q_{\mathcal{P}} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

(c) $(\mathbb{C}, \mathbb{R})$ ist der VR $\mathbb{C}$ über $\mathbb{R}$. Es sei

$$\mathcal{B} = \{1, i\}$$

eine Basis[6]. Dann lautet der Koordinatenvektor von $z = 2 - 3\,i$

$$z_{\mathcal{B}} = \begin{pmatrix} 2 \\ -3 \end{pmatrix}.$$

Bezüglich der Basis

$$\mathcal{C} = \{1 + i, 2\}$$

lautet der Koordinatenvektor

$$z_{\mathcal{C}} = \begin{pmatrix} 6 \\ -1 \end{pmatrix} \frac{1}{2}.$$

Satz 3.15 *Sei $v_1, \ldots, v_n \in V$ eine Basis und $a \in V$. Dann gibt es genau n Skalare $\lambda_i\,,\ i = 1, \ldots, n$ mit*

$$a = \sum_{k=1}^{n} \lambda_k v_k\,.$$

Beweis von Satz 3.15 auf Seite 159

Das bedeutet, dass es zu jeder Basis genau eine Darstellung des Vektors gibt, was daran liegt, dass eine Matrix, deren Spalten lu sind (ist ja hier jeweils eine Basis) immer regulär ist!

[5] Greifen Sie auch hier gerne zum Buntstift. Es hilft! Ich habe schon mal angefangen.
[6] Prüfen Sie das mal.

Satz 3.16 (Basiswechsel) *Es sei V ein n-dimensionaler VR mit den Basen*

$$\mathcal{V} = \{v_1, \dots, v_n\} \quad \text{und} \quad \mathcal{W} = \{w_1, \dots, w_n\}\,.$$

Des weiteren beschreiben $B_{\mathcal{V}}$ und $B_{\mathcal{W}}$ die Matrizen, deren Spalten die Basisvektoren sind:

$$B_{\mathcal{V}} = (v_1, \dots, v_n) = \begin{pmatrix} v_1^1 & \cdots & v_n^1 \\ \vdots & \ddots & \vdots \\ v_1^n & \cdots & v_n^n \end{pmatrix}, \; B_{\mathcal{W}} = (w_j^i)_{ij}$$

Des weiteren beschreiben $a_{\mathcal{V}}$ und $a_{\mathcal{W}}$ die Koordinatenvektoren von $a \in V$ bzgl. der Basen $\mathcal{V}$ und $\mathcal{W}$. Dann gilt:

$$a_{\mathcal{V}} = B_{\mathcal{V}}^{-1} B_{\mathcal{W}} a_{\mathcal{W}} \quad \text{und} \quad a_{\mathcal{W}} = B_{\mathcal{W}}^{-1} B_{\mathcal{V}} a_{\mathcal{V}}\,.$$

Die Matrizen $B_{\mathcal{W}}^{-1} B_{\mathcal{V}}$, bzw. $B_{\mathcal{V}}^{-1} B_{\mathcal{W}}$ heišen Transformationsmatrizen *zum* Basiswechsel *von $\mathcal{W}$ nach $\mathcal{V}$, bzw. umgekehrt.*

Ein kleiner, auch anderweitig gut brauchbarer Hilfssatz erleichtert den Beweis dieses Satzes.

Hilfssatz 3.17 *Es seien $A, B \in \mathbb{K}^{n \times n}$ reguläre Matrizen. Dann gilt*

$$(A \cdot B)^{-1} = B^{-1} A^{-1}\,.$$

Beweis von Hilfssatz 3.17 auf Seite 158

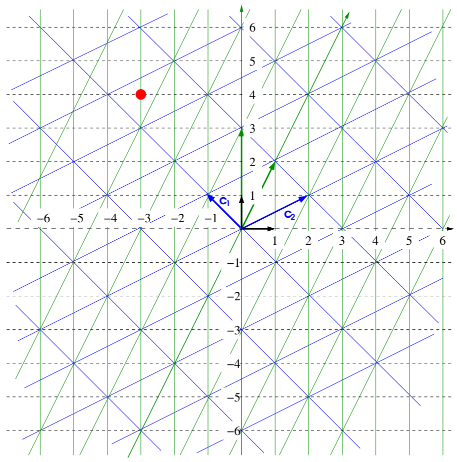
Beweis von Satz 3.16 auf Seite 158

Beispiel 51 Basiswechsel

(a) Wir gehen vom Setting in Beispiel 50, Teil (a) aus.

Sei nun

$$\mathcal{C} = \{c_1, c_2\} = \left\{ \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\}$$



eine weitere Basis des $\mathbb{R}^2$. Der Koordinatenvektor von a (aus Teil (a)) bezüglich der Basis $\mathcal{C}$ lautet nun:

$$a_{\mathcal{C}} = \underbrace{\begin{pmatrix} -1 & 2 \\ 1 & 1 \end{pmatrix}}_{=:C}^{-1} \begin{pmatrix} -3 \\ 4 \end{pmatrix}$$
$$= \frac{1}{3} \begin{pmatrix} -1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -3 \\ 4 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 11 \\ 1 \end{pmatrix}$$

Wollen wir nun den Koordinatenvektor von a bezüglich der Basis $\mathcal{B}$ direkt in den bezüglich der Basis $\mathcal{C}$ umrechnen so können wir das tun, ohne den Vektor a im kartesischen System zu berechnen. Nämlich so: Es ist ja

$$a = B\, a_{\mathcal{B}} \qquad \text{und} \qquad a = C\, a_{\mathcal{C}}\,.$$

Daraus folgt dann direkt

$$B\, a_{\mathcal{B}} = C\, a_{\mathcal{C}}\,.$$

Multiplizieren wir nun beide Seiten von links mir der Inversen von C so erhalten wir direkt

$$a_{\mathcal{C}} = C^{-1} B a_{\mathcal{B}} = \frac{1}{3} \begin{pmatrix} -1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix} a_{\mathcal{B}} = \underbrace{\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}}_{Transformationsmatrix} a_{\mathcal{B}}\,.$$

Wir setzen zum Test mal die Werte für $a_{\mathcal{B}}$ ein:

$$a_{\mathcal{C}} = C^{-1} B a_{\mathcal{B}} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} -3 \\ \frac{10}{3} \end{pmatrix} = \begin{pmatrix} \frac{11}{3} \\ \frac{1}{3} \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 11 \\ 1 \end{pmatrix}$$

Voilà!

Beispiel 52 Anwendungsbeispiele

(a) Matrixmultiplikation statt Ableitung: Mit der Koordinatendarstellung können wir zum Beispiel mit unseren Rechnungen von einem Polynomraum $\mathbb{P}_{n-1}$ in den $\mathbb{R}^n$ wechseln. Statt eine Ableitung berechnen zu müssen, was ja auf dem Computer immer etwas mühsam ist, kann man auch einfach eine Matrix-Vektor-Multiplikation ausführen. Das geht so:

Wir wählen folgendes Setting:

$$\text{Basis: } \mathcal{P} = \{1, x, x^2\} \subseteq \mathbb{P}_2$$
$$\text{Vektor: } q(x) = \{-x^2 + 4\,x - 3\} \in \mathbb{P}_2$$
$$\text{Koordinatenvektor: } q_{\mathcal{P}} = \begin{pmatrix} -1 \\ 4 \\ -3 \end{pmatrix} \in \mathbb{R}^3$$

Die Ableitungsfunktion von q ist

$$q'(x) = -2\,x + 4$$

und der zugehörige Koordinatenvektor

$$q'_{\mathcal{P}} = \begin{pmatrix} 0 \\ -2 \\ 4 \end{pmatrix} .$$

Die Matrix $A \in \mathbb{R}^{3\times 3}$, mit

$$q'_{\mathcal{P}} = A\,q_{\mathcal{P}}\,, \quad \begin{pmatrix} 0 \\ -2 \\ 4 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} -1 \\ 4 \\ -3 \end{pmatrix}$$

lautet

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

Wie wir so eine Matrix berechnen können besprechen wir in Kapitel 3.5. Im Moment soll das nur ein Beispiel sein, was wir so mit unseren neuen Erkenntnissen machen können.

Machen Sie den Test mit $p(x) = 4\,x^2 - 5\,x + 100$.

$$p_{\mathcal{P}} = \begin{pmatrix} 4 \\ -5 \\ 100 \end{pmatrix} , \; Ap_{\mathcal{P}} = \begin{pmatrix} 0 \\ 8 \\ -5 \end{pmatrix}$$

zurück in den Polynomraum ergibt:

$$8\,x - 5 = p'$$

✓

(b) Datenkompression:

Wir wollen eine Menge von Daten mit geringerem Speicherbedarf verwalten und nehmen dafür auch einen gewissen Informationsverlaust in Kauf. Es ist leicht einzusehen, dass man mehr verliert, wenn man mehr spart. Wir wollen am folgenden Beispiel auf einfachster Ebene begreiflich machen, dass man durch geschickte Wahl einer "guten" Basis bei gleichbleibendem Speicheraufwand weniger Information verlieren kann.

Wir betachten einmal den Punkt

$$P = (2,3) \in \mathbb{R}^2$$

und die Basis

$$\mathcal{B} = \{b_1, b_2\} = \left\{ \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}, \frac{1}{\sqrt{5}}\begin{pmatrix}1\\2\end{pmatrix} \right\}$$

des $\mathbb{R}^2$. Der Koordinatenvektor α_P zum Punkt P bezüglich der Basis $\mathcal{B}$ lautet

$$\begin{aligned}\alpha_P = (b_1, b_2)^{-1} P &= \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{5}} \\ -\frac{1}{\sqrt{2}} & \frac{2}{\sqrt{5}} \end{pmatrix}^{-1} \begin{pmatrix}2\\3\end{pmatrix} \\ &= \frac{\sqrt{10}}{3} \begin{pmatrix} \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix}2\\3\end{pmatrix} = \frac{1}{3}\begin{pmatrix}\sqrt{2}\\5\sqrt{5}\end{pmatrix} \approx \begin{pmatrix}0.47\\3.78\end{pmatrix}\end{aligned}$$

Es ist demnach

$$P = \alpha_{P,1} b_1 + \alpha_{P,2} b_2 = \frac{\sqrt{2}}{3}\, b_1 + \frac{5\sqrt{5}}{3}\, b_2\,.$$

Der Basisvektor b_1 hat ein Gewicht von ca. 0.47 und b_2 eins von ca. 3.78. Da b_1 eine geringere Bedeutung bei der Darstellung des Punktes P hat und wir Speicher sparen wollen werfen wir diesen Anteil an Information einfach weg. Der rekonstruierte Wert lautet dann:

$$\tilde{P} = \underbrace{\cancel{\alpha_1 b_1}}_{\text{Speicher sparen}} + \alpha_2\, b_2 = \frac{5\sqrt{5}}{3} \cdot \frac{1}{\sqrt{5}} \cdot \begin{pmatrix}1\\2\end{pmatrix} = \frac{5}{3}\begin{pmatrix}1\\2\end{pmatrix} \approx \begin{pmatrix}1.67\\3.33\end{pmatrix}$$

Wir erhalten natürlich nicht mehr unsere Originaldaten aber einen Wert, der dem ursprünglichen recht ähnlich ist. Gar nicht schlecht, dafür, dass wir ihn mit halbem Speicheraufwand erhalten können. Wir wollen dieses Konzept an einer ganzen Datensammlung, die allerdings immer noch überschaubar ist, durchführen. Wir nehmen dazu die Punktemenge bestehend aus 8 Punkten im $\mathbb{R}^2$:

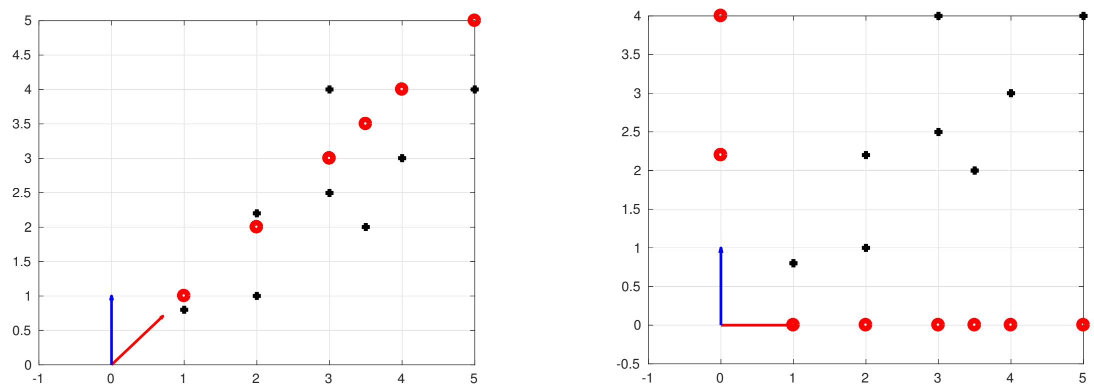
$$P = (P^1, P^2, \ldots, P^8) = \begin{pmatrix} 1 & 2 & 2 & 3.5 & 3 & 3 & 4 & 5 \\ 0.8 & 1 & 2.2 & 2 & 2.5 & 4 & 3 & 4 \end{pmatrix}$$

Von jedem dieser Punkte P^j berechnen wir den Koeffizientenvektor α_{P^j} bezüglich einer Basis

$$\mathcal{B} = \{b_1, b_2\}\,.$$

Wir entledigen uns jeweils des Anteils im Koordinatenvektor α_{Pj}, der betragsmäßig den kleineren Anteil trägt, rekonstruieren den Datensatz und betrachten das Ergebnis mit einem plot. Abbildung 1 zeigt, wie unterschiedlich die Ergebnisse unserer einfachen Datenkompression ausfallen, je nachdem welche Basis wir wählen. "Passt" die Basis zur geometrischen Ausrichtung der "Punktwolke" so erhalten wir eine qualitativ bessere Datenkompression. Bei Punkten in der Ebene kann man das noch gut erkennen. Wenn das aber tausende von Punkten im $\mathbb{R}^{17}$ sind wird es schwierig. Da braucht man dann eine Methode, die sich, ohne "Hingucken" zu müssen, automatisieren lässt. Wir werden im Kapitel 3.6 mehr darüber erfahren.

"gute Basis" "schlechte Basis"



`Basis:` $b_1 = \frac{1}{\sqrt{2}}\binom{1}{1}$; $b_2 = \binom{0}{1}$;
`Ausgabe: Error = 2.57e+00`

`Basis:` $b_1 = \binom{1}{0}$; $b_2 = \binom{0}{1}$;
`Ausgabe: Error = 6.07e+00`

Abbildung 1: Variation der Basis bei Datenkompression

(c) Mustererkennung:

Es sei

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

eine Basis[7] des $\mathbb{R}^2$ Wir wollen uns die Koordinatenvektoren der folgenden drei Matrizen anschauen:

$$A = \begin{pmatrix} 80 & 17 \\ 10 & 69 \end{pmatrix}, C = \begin{pmatrix} 1 & 18 \\ 16 & 2 \end{pmatrix}, F = \begin{pmatrix} 20 & 20 \\ -1 & 3 \end{pmatrix}$$

Die zugehörigen Koordiantenvektoren lauten[8]

[7] Vergewissen Sie sich, dass das eine Basis ist und, dass diese Basis den ganzen $\mathbb{R}^{2\times 2}$ aufspannt.
[8] Das Nachrechnen dürfen Sie als kleine Übungsaufgabe verstehen. :)

$$A_{\mathcal{B}} = \begin{pmatrix} 73 \\ 10 \\ 7 \\ -4 \end{pmatrix}, \; C_{\mathcal{B}} = \begin{pmatrix} -1 \\ 16 \\ 2 \\ 3 \end{pmatrix}, \; F_{\mathcal{B}} = \begin{pmatrix} -1 \\ -1 \\ 21 \\ 4 \end{pmatrix}.$$

Wir können jetzt an den Koordinaten ablesen welche von den Mustern, die wir für die Basisvektoren gewählt haben am stärksten in den Vektoren (hier Matrizen) repräsentiert werden. Die Matrix A etwa ist diagonaldominant[9] und wird am stärksten vom ersten Basisvektor, was eine Diagonalmatrix ist dargestellt.

3.4 Messen in Vektorräumen

Wir haben nun Räume in verschiedenen Strukturen, zuletzt den Vektorraum, kennengelernt und Sie haben bereits erfahren, dass es für die Anwendung von Formeln dringend nötig ist, dass die geforderten Voraussetzungen erfüllt sind. Nun wollen wir aber auch ein paar Formeln kennenlernen. Viele Formeln beinhalten Messungen. Was wir in VRen messen könnten - nicht immer - sind Längen und manchmal auch Winkel. Sicher können Sie sich bereits darunter was vorstellen. Ihre Idee von "senkrecht" hat wahrscheinlich etwas mit einem Winkel von "$90°$" zu tun. Diese Idee soll Ihnen nicht zerstört werden, sie ist schon richtig, aber sie soll abstrahiert und dann um diverse Erkenntnisse erweitert werden. Es kann etwas senkrecht (wir sprechen dann von orthogonal) sein und gleichzeitig in gewissem Sinne einen Winkel von $35°$ haben. In vielen Fällen ist man aber auch gar nicht an einer konkreten Winkelzahl interessiert sondern nur daran, ob etwas orthogonal ist oder nicht. Eine Winkelzahl für den Winkel zwischen zwei Polynomen (als Vektoren) kann man nirgends gebrauchen, aber eine Orthogonalitätseigenschaft schon.

Konzeptionell wird das so aussehen, dass wir unseren VRen ein Maß für Länge oder Winkel mitgeben. Wir erweitern damit seine Struktur. Starten wir mit der Messung von Längen. Das ist am einfachsten.

[9] Es dürfte aus dem Bauch heraus klar sein was dieser Begriff beschreibt.

3.4.1 Längen und Winkel

Definition 3.18 (Norm) *Eine Abbildung* $\|\cdot\|_V : V \to \mathbb{R}$ *heißt* Norm, *wenn für beliebige* $a, b \in V$ *und* $\lambda \in \mathbb{K}$ *gilt:*

(N1) $\|a\|_V \geq 0 \wedge \|a\|_V = 0 \;\Leftrightarrow\; a = 0$

(N2) $\|\lambda\, a\|_V = |\lambda|\, \|a\|_V$

(N3) $\|a + b\|_V \leq \|a\|_V + \|b\|_V$ *(Dreieckungleichung)*

$(V, (\mathbb{K}, +, \cdot), \oplus, \odot, \|\cdot\|_V)$ *oder kurz* $(V, \|\cdot\|_V)$ *heißt normierter Raum*
Wir sagen ein Vektor a *habe die Länge* $\|a\|_V$ *(bezüglich der Norm* $\|\cdot\|_V$*).*

Beispiel 53 Normen im $\mathbb{R}^n$

(a) Eine Norm, die Sie sicher bereits kennen ist der Betrag im $\mathbb{R}^n$:

$$\|a\|_2 = \sqrt{\sum_{i=1}^{n} a_i^2}$$

Für

$$a = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \in \mathbb{R}^3$$

gilt dann

$$\|a\|_2 = \sqrt{a_1^2 + a_2^2 + a_3^2} = \sqrt{1 + 4 + 9} = \sqrt{14}$$

und beschreibt den Abstand vom Punkt a zum Ursprung[10]

(b) *Manhattan-Norm*:

$$\|a\|_1 = \sum_{i=1}^{n} |a_i|$$

Für

$$a = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} \in \mathbb{R}^3$$

gilt dann

$$\|a\|_1 = |1| + |-2| + |3| = 6\,.$$

Es beschreibt die Länge des Weges von 0 nach a, wenn man auf den Linien eines Rasters entlangläuft. Daher der Name. :)

[10] Überlegen Sie sich ein Beispiel im $\mathbb{R}^2$, machen Sie sich eine Skizze dazu und erinnern Sie sich gerne bei dieser Gelegenheit an den Pythagoras.

(c) allgemein für $0 < p < \infty$ gilt:

$$\|a\|_p = \sqrt[p]{\sum_{i=1}^{n} |a_i|^p}$$

Für

$$a = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} \in \mathbb{R}^3$$

gilt dann

$$\|a\|_5 = \sqrt[5]{|1|^5 + |-2|^5 + |3|^5} = (1 + 32 + 243)^{\frac{1}{5}} \approx 3.08\,.$$

(d) und für $p = \infty$:

$$\|a\|_\infty = \max_{i=1,..,n} |a_i|$$

Zum Beispiel ist dann

$$\left\| \begin{pmatrix} -5 \\ 2 \end{pmatrix} \right\|_\infty = 5\,.$$

Für die 2-Norm gelten folgende Notationen gleichermaßen:

$$\|\cdot\|_2 = \|\cdot\| = |\cdot|$$

Dass all diese beispielhaften Vorschriften tatsächlich Normen sind muss man im Grunde prüfen; streng genommen. Das heißt, sobald man von einer Abbildung behauptet sie sei eine Norm, müssen die Axiome $(N1)$, $(N2)$ und $(N3)$ nachgewiesen werden.

Zum Beispiel gilt für die 2-Norm wegen

$$\|0\| = \sqrt{\sum_i 0^2} = \sqrt{0} = 0$$

und mit

$$\underbrace{a_1^2}_{\geq 0} + \cdots \underbrace{a_n^2}_{\geq 0} = 0 \quad \Rightarrow \quad a_1 = \cdots = a_n = 0$$

schon einmal $(N1)$. Und $(N2)$ lässt sich auch leicht nachprüfen, denn es gilt

$$\|\alpha\, a\| = \sqrt{\sum_i |\alpha\, a_i|^2} = \sqrt{|\alpha|^2 \sum_i |a_i|^2} = \sqrt{|\alpha|^2}\sqrt{\sum_i |a_i|^2} = |\alpha|\sqrt{\sum_i |a_i|^2} = |\alpha|\|a\|$$

Die Dreiecksungleichung $(N3)$ ist der schwierigste Kandidat. Der Nachweis dieser Gültigkeit ist einfach, wenn wir erst noch ein paar mathematische Begriffe lernen. Der Nachweis steckt dann im Satz 3.20 weiter unten.

Schauen wir uns zunächst aber noch ein paar weitere Beispiele an:

Beispiel 54 Normen im $\mathbb{P}_n(I)$

(a) quadratintegrierbare Funktionen auf $I = [0, 1]$:

$$\|q\|_2 = \sqrt{\int_0^1 q^2(x)\, dx}$$

Etwa ist

$$\|x^2\|_2 = \sqrt{\int_0^1 x^4\, dx} = \frac{1}{\sqrt{5}} \approx 0.45$$

(b) allgemein für $0 < p < \infty$:

$$\|q\|_p = \left(\int_I |q(x)|^p\, dx \right)^{\frac{1}{p}}$$

(c) und für $p = \infty$:

$$\|q\|_\infty = \sup_{x \in I} |q(x)|$$

Die Norm gibt uns einen abstrakten Begriff für Längen von Vektoren. Wir erhalten damit auch einen Längenbegriff für Funktionen und meinen dabei nicht die Länge der Kurve. Es kann ja auch etwas Höherdimensionales sein.

Wenn wir von Länge sprechen müssen wir im Grunde immer auch sagen bzgl. welcher Norm das gilt. Da die Norm aber an den VR geknüpft ist, ist es an sich aus dem Kontext heraus klar.

Machen wir uns weiter auf den Weg in Richtung eines Winkelbegriffs. Dafür definieren wir zunächst folgende Abbildung.

Definition 3.19 (Skalarprodukt) *Eine Abbildung $s : V \times V \to \mathbb{R}$ mit*

(S1) $s(\alpha_1\, a_1 + \alpha_2\, a_2, b) = \alpha_1\, s(a_1, b) + \alpha_2\, s(a_2, b)$ *(Bilinearität)*

(S2) $s(a, b) = s(b, a)$ *(Kommutativität)*

(S3) $s(a, a) \geq 0$ *(positiv definit)*

(S3) $s(a, a) = 0 \;\Rightarrow\; a = 0$

heißt Skalarprodukt. $(V, (\mathbb{K}, +, \cdot), \oplus, \odot, s)$ kurz (V, s) heißt Prähilbertraum. Prähilbertraum ist ein Überbegriff. Wählt man $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ so heißt $(V, \mathbb{C}, s)$ unitärer VR und $(V, \mathbb{R}, s)$ Euklidischer VR . Im speziellen Fall, dass $V = \mathbb{R}^2$ ist nennen wir den Vektorraum Euklidische Ebene.

Folgerung: (S1) und (S2) liefern Linearität auch im zweiten Argument:

$$s(a, \beta_1\, b_1 + \beta_2\, b_2) = \beta_1\, s(a, b_1) + \beta_2\, s(a, b_2)\,.$$

Aus (S1) folgt

$$a = 0 \;\Rightarrow\; s(a, a) = 0\,.$$

darum ist $(S3)$ nicht als Äquivalenz beschrieben. Die $\Leftarrow$-Richtung wäre überflüssig im axiomatischen Ansatz.

Beispiel 55 Skalarprodukte

(a) *Standardskalarprodukt* im $\mathbb{R}^n$:

$$< a, b >:= s_{\mathbb{R}^n}(a, b) = \sum_{i=1}^{n} a_i b_i$$

wie etwa im $\mathbb{R}^2$ mit den Vektoren

$$a = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \; b = \begin{pmatrix} -5 \\ 3 \end{pmatrix}.$$

Da gilt dann

$$\left\langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -5 \\ 3 \end{pmatrix} \right\rangle = 1 \cdot (-5) + 2 \cdot 3 = -5 + 6 = 1\,.$$

(b) Im Polynomraum gibt es zum Beispiel dieses Skalarprodukt:

$$s(p, q) = \int_I pq\, dx$$

Für die Funktionen $p(x) = x$ und $q(x) = 1 - x^2$ aus $\mathbb{P}_2(I)$, mit $I = [0, 1]$ gilt dann

$$\int_I p(x)q(x)\, dx = \int_0^1 x\,(1 - x^2)\, dx = \int_0^1 x - x^3\, dx = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}\,.$$

Wir schreiben statt $s(a, b)$ oder $< a, b >$ auch gerne kurz

$$a \cdot b$$

und setzen dabei einen deutlich sichtbaren Punkt zwischen a und b.

Auch bei der Aufstellung eines Skalarproduktes muss man im Grunde die Axiome $(S1)$ bis $(S4)$ nachweisen. Für das Standardskalarprodukt können Sie das als kleine Übungsaufgabe gerne mal versuchen.

Auch wenn es jede Menge verschiedener Skalarprodukte gibt, ist das Standardskalarprodukt am gebräuchlisten; genaus wie die 2-Norm. Warum das so ist sehen wir gleich.

Satz 3.20 *Eine durch ein Skalarprodukt induzierte Abbildung gemäß*

$$\|a\|_s := \sqrt{s(a,a)}$$

stellt eine Norm dar und heißt Euklidische Norm.

Um zu zeigen, dass eine durch ein Skalarprodukt induzierte Abbildung gemäß $\|\cdot\|_s = \sqrt{s(\cdot,\cdot)}$ tatsächlich eine Norm darstellt ist folgender Satz hilfreich.

Satz 3.21 *(Cauchy–Schwarzsche Ungleichung)*

$$|s(x,y)| \leq \|x\|_s \cdot \|y\|_s \tag{6}$$

Beweis von Satz 3.21 auf Seite 160, Beweis von Satz 3.20 auf Seite 159

Die 2-Norm $\|\cdot\|_2$ ist die Euklidische Norm zum Standardskalarprodukt, denn

$$\sqrt{< a, a >} = \sqrt{\sum_i a_i \cdot a_i} = \sqrt{\sum_i a_i^2} = \|a\|_2 \,.$$

Deshalb und wegen Satz 3.20 gelten dann auch alle Normaxiome für diese Norm. Bei einer Euklidischen Norm muss man demnach die Normaxiome gar nicht mehr nachweisen. Schön, nicht wahr?

Da jedes Skalarprodukt eine Euklidische Norm induziert, ist jeder Prähilbertraum automatisch auch ein normierter Raum. Man kann aber nicht zu jeder Norm ein "passendes" Skalarprodukt finden, so dass nicht jeder normierte Raum auch ein Prähilbertraum ist.

Es ist tatsächlich so. Wenn Sie in Ihrem Raum mit der Manhattannorm - bei DeepLearning-Methoden ist das zum Beispiel so - rechnen wollen, dann arbeiten Sie in einem normierten Raum. Sie haben kein Skalarprodukt. Kann man sich fragen warum das komisch sein sollte. Mit den beiden Begriffen "Norm" und "Skalarprodukt" können wir nun einen "Winkelbegriff" erklären. Das bedeutet, wenn Sie kein Skalarprodukt haben dann gibt es auch keine Winkel. An so einen Gedanken muss man sich erst mal gewöhnen. Schauen wir zunächst nach der etwas abstrakteren Version der Winkeldefinition.

Definition 3.22 (Winkel) *Der Winkel φ zwischen zwei Vektoren a und b im Prähilbertraum (V, s) ist definiert durch*

$$\cos\varphi = \frac{s(a,b)}{\|a\|_s \|b\|_s}.$$

Dabei bezeichnet $\|\cdot\|_s$ die durch das Skalarprodukt $s(\cdot,\cdot) : V \times V \to \mathbb{K}$ induzierte Euklidische Norm.

Damit das nicht nur so eine dahingeplätscherte Definition ist wollen wir uns zunächst davon überzeugen, dass dieser Winkel genau der ist, den wir aus der Geometrie in der Euklidischen Ebene her kennen. Es wäre ratsam an dieser Stelle, wenn Sie sich Kapitel A.1 über Trigonometrie im Anhang durchlesen würden. Es sind grundlegende trigonometrische Eigenschaften, die wir jetzt benötigen.

Es sei nun $V = \mathbb{R}^2$ ein Prä-HR mit dem Standardskalarprodukt

$$s(\cdot,\cdot) = <\cdot,\cdot>$$

und

$$\|\cdot\|_s = \|\cdot\|_2 = \sqrt{<\cdot,\cdot>}$$

die zugehörige Euklidische Norm.

Sie dürfen im Moment gerne an den altbekannten Winkel von 0-360 Grad denken. Wir wollen uns vegewissen, dass in diesem Kontext, also in $(\mathbb{R}^2, <\cdot,\cdot>)$ die neue Definition eines Winkels genau das Gleiche ist. Also weiter:

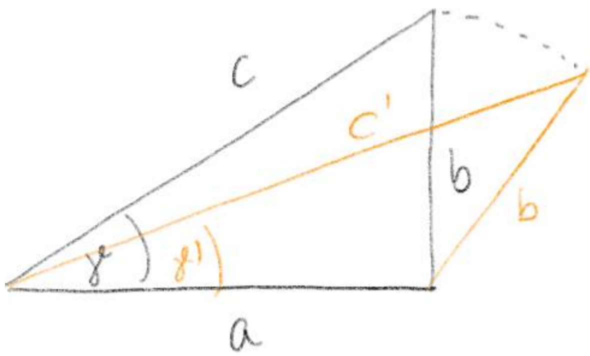


Abbildung 2: Pythagoras und Kosinussatz

Sie kennen alle den Pythagoras, der besagt, dass für ein Dreieck mit Seitenlängen a, b, c, bei dem die Seiten zu den Längen a und b senkrecht, also im 90° Winkel, aufeinanderstoßen die Beziehung

$$a^2 + b^2 = c^2$$

gilt. Haben Sie sich schon einmal überlegt, ob die Umkehrung der Aussage auch gilt? Also wenn für ein Dreieck die Beziehung $a^2 + b^2 = c^2$ erfüllt ist, handelt es sich dann auch um ein rechtwinkliges Dreieck? Die Antwort ist ja nicht selbstverständlich. Der *Konsinussatz* liefert diese Beziehung, denn er stellt eine Erweiterung des Pythagoras für beliebige, nicht zwingend rechtwinklige, Dreiecke dar (s. Abb. 2).

Satz 3.23 (Kosinussatz) *Es sei ein Dreieck mit Seitenlängen $a, b, c \in \mathbb{R}$ gegeben, wobei $\gamma \in \mathbb{R}$ den Winkel beschreibt gegenüber der Kante mit Seitenlänge c. Dann gilt*

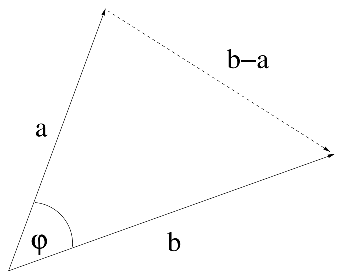
$$c^2 = a^2 + b^2 - 2\,a\,b\,\cos\gamma\,.$$

Beweis von Satz 3.23 auf Seite 154

Aus dem Kosinussatz[11] lässt sich dann sofort der Zusammenhang zwischen dem Winkel aus Definition 3.22 und dem Herkömmlichen herstellen.

Es werde ein Dreieck durch die beiden Vektoren $a, b \in \mathbb{R}^2$ beschrieben. Die dritte Seite erhalten wir aus $b - a$. Dann liefert der Kosinussatz direkt die Beziehung

$$\begin{aligned} & \|b-a\|^2 = \|a\|^2 + \|b\|^2 - 2\,\|a\|\,\|b\|\,\cos\varphi \\ \Leftrightarrow\quad & \cos\varphi = \frac{\|b-a\|^2 - \|a\|^2 - \|b\|^2}{-2\,\|a\|\,\|b\|} \end{aligned}$$



Betrachten wir nur den Zähler

$$< b-a, b-a > - < a, a > - < b, b > = -2 < a, b >$$

so erhalten wir insgesamt

$$\cos\varphi = \frac{< a, b >}{\|a\|\,\|b\|}$$

Das ist doch schön. Wir halten hier immerhin eine unglaublich simple Methode zur Berechnung des Winkels in Händen!

Zur Berechnung des Winkels müssen wir den Arkuskosinus heranziehen. Auf Grund seiner Definition erhalten wir als Ergebnis nur Winkel aus $[0°, 180°]$. Wir erhalten also bei der Berechnung des Winkels zwischen zwei Vektoren immer den inneren Winkel. Wollen wir den äußeren wissen, so müssen wir ihn nachträglich umrechnen.

[11] Sie können sich direkt überlegen ob der Satz plausibel ist, wenn Sie die Beziehung "größer/kleiner werden con c" mit "γ ist größer/kleiner als 90°" reflektieren. Machen Sie sich auch hier eine kleine Skizze dazu.

Beispiel 56 Winkelberechnung

(a) Seien $a, b, c \in \mathbb{R}^2$ gegeben durch

$$a = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \; b = \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \; c = \begin{pmatrix} 2 \\ -2 \end{pmatrix}.$$

Dann ist

$$\angle(a,b) = \arccos\left(\frac{< a, b >}{\|a\|\|b\|}\right) = \arccos\left(\frac{5}{\sqrt{26}}\right) \approx \arccos 0.98 \approx 0.2 (\approx 11.47^\circ)$$

$$\angle(a,c) = \arccos\left(\frac{< a, c >}{\|a\|\|c\|}\right) = \arccos 0 = \frac{\pi}{2}$$

(b) Die Polynome aus Beispiel 55, Teil (b) haben mit

$$s(p,q) = \int_0^1 p(x)q(x)\,dx \quad \text{und zug. Eukl. Norm} \quad \|p\|_s = \left(\int_0^2 p(x)^2\,dx\right)^{\frac{1}{2}}$$

den Winkel

$$\cos\varphi = \frac{s(p,q)}{\|p\|_s\|q\|_s} = \frac{\sqrt{45}}{8\sqrt{2}}$$

Denn

$$\int_0^1 p(x)q(x)\,dx = \int_0^1 x(1-x^2)\,dx = \frac{1}{4}$$

$$\left(\int_0^1 p^2(x)\,dx\right)^{\frac{1}{2}} = \left(\int_0^1 x^2\,dx\right)^{\frac{1}{2}} = \frac{1}{\sqrt{3}}$$

$$\left(\int_0^1 q^2(x)\,dx\right)^{\frac{1}{2}} = \left(\int_0^1 1 - 2x^2 + x^4\,dx\right)^{\frac{1}{2}} = 2\sqrt{\frac{2}{15}}$$

Abbildung 3 zeigt die geometrische Interpretation des Skalarprodukts. Wir können schon hier erahnen welche Aufgabe das Skalarprodukt bei der orthogonalen Projektion haben wird, welche wir im anschließenden Kapitel 3.4.2 besprechen werden.

Es sind zwei Eigenschaften, die wir über Skalarprodukte und Normen erhalten, die von besonderer Bedeutung sind: Die Länge Eins und der "rechte" Winkel. Und weil die so besonders sind gibt es zu diesen Eigenschaften auch besondere Nomenklaturen. Wir erweitern also nochmal ein wenig unser Mathe-Vokabelheft:

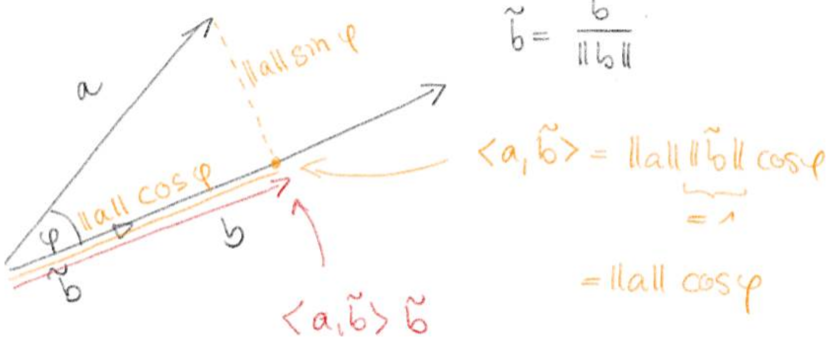


Abbildung 3: Geometrische Interpretation des Standardskalarprodukts im $\mathbb{R}^2$

Definition 3.24 (orthogonale und parallele Vektoren) *Zwei Vektoren $a, b \in V$ heißen* orthogonal *oder* senkrecht *(bezüglich des Skalarprodukts $s(\cdot,\cdot)$) wenn*

$$s(a,b) = 0$$

gilt. Sie heißen parallel *(bezüglich des Skalarprodukts $s(\cdot,\cdot)$) wenn*

$$|s(a,b)| = \|a\|_s \|b\|_s$$

gilt.

Beispiel 57

(a) Für die Vektoren

$$a = \begin{pmatrix} 2 \\ 4 \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

gilt

$$< a, b >= \left\langle \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right\rangle = 2 \cdot (-2) + 4 \cdot 1 = 0\,.$$

a und b sind demnach orthogonal.

(b) Die Vektoren

$$a = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

sind Paralle, denn es gilt

$$10 =< a, b >= \|a\| \|b\| = \sqrt{5}\sqrt{20} = 10\,.$$

Dass diese beiden Vektoren in die gleiche Richtung zeigen sieht man sofort ein, wenn man eine kleine Skizze macht; vlt auch vor dem geistigen Auge...

(c) Bezüglich des Skalarprodukts

$$s(p, g) = \int_0^\pi p(x)g(x)\, dx$$

sind die Vektoren

$$p(x) = \cos x \quad \text{und} \quad g(x) = \sin x$$

orthogonal, denn es gilt

$$s(p, g) = \int_0^\pi \cos x \, \sin x \, dx = 0\,.$$

Definition 3.25 (Normalvektor) *Sei $a \in V$. Dann heißt $a^\perp \in V$* Normalvektor *zu a, wenn*

$$s(a, a^\perp) = 0$$

erfüllt ist.

Es ist also $\sin x^\perp = \cos x$. Wichtig ist, dass die Orthogonalität sich auf das speziell gewählte Skalarprodukt, hier aus Beispiel 57, Teil (c), bezieht und bei der Wahl des Skalaprodukts über ein Integral auch der Integrationsbereich relevant ist, denn es gilt etwa

$$\int_0^{\frac{\pi}{2}} \cos x \, \sin x \, dx \neq 0\,!!!$$

Bezüglich dieses Skalarprodukts, mit anderen Integrationsgranzen, sind die Vektoren nicht orthogonal.

Im $\mathbb{R}^2$ erhält man generelle den Normalvektor über diese Vorschrift:

$$\begin{pmatrix} a \\ b \end{pmatrix}^\perp = \begin{pmatrix} -b \\ a \end{pmatrix}$$

Es handelt sich dabei um eine Drehung gegen den Uhrzeigersinn. Man kann genauso mit dem Uhrzeigersinn drehen. Es muss nur einmal klargestellt sein wie das Symbol definiert ist und in der Mathe "geht" man ja vorzugsweise gegen den Uhrzeigersinn.

Definition 3.26 (normierter Vektor) $a \in (V, \|\cdot\|_V)$ *mit*

$$\|a\|_V = 1$$

heißt normiert bezüglich $\|\cdot\|_V$ *oder kurz normiert.*

Ein Normalvektor ist nicht zwingend normiert!

Beispiel 58 normierter Vektor

$$\left\| \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \right\|_2 = \sqrt{\left(\frac{1}{\sqrt{2}}\right)^2 + \left(-\frac{1}{\sqrt{2}}\right)^2} = \sqrt{\frac{1}{2} + \frac{1}{2}} = 1$$

Jeder Vektor $a \in (V, \|\cdot\|_V)$ kann gemäß

$$\tilde{a} = \frac{a}{\|a\|_V} \qquad \Rightarrow \qquad \|\tilde{a}\|_V = 1$$

normiert werden.

Beispiel 59 Es sei $V = (\mathbb{R}^2, \|\cdot\|_\infty)$ ein normierter Raum und

$$a = \begin{pmatrix} -4 \\ 2 \end{pmatrix}.$$

Dann ist

$$\|a\|_\infty = 4 \quad \text{und} \quad \tilde{a} = \frac{a}{\|a\|_\infty} = \begin{pmatrix} -1 \\ \frac{1}{2} \end{pmatrix}.$$

$\tilde{a}$ ist bezüglich $\|\cdot\|_\infty$ ein normierter Vektor, nicht aber bezüglich $\|\cdot\|_2$, denn $\|\tilde{a}\|_2 = \sqrt{5} \neq 1$.

Bisher haben wir uns mit zwei Arten von Produkten im VR beschäftigt. Das ist zum Einen die skalare Mutliplikation - als eine Art Produkt aus Vektor und Skalar mit einem Vektor als Ergebnis - und zum Anderen das Skalarprodukt - als eine Art Produkt von zwei Vektoren mit einem Skalar als Ergebnis. Es wäre enttäuschend wenn wir damit schon am Ende wären. Im $\mathbb{R}^2$ und $\mathbb{R}^3$ gibt es noch eine weitere Art von Produkt, nämlich das Kreuzprodukt. Das Kreuzprodukt hat ganz besondere Eigenschaften, die wir uns anschauen wollen. Man fragt sich warum es dieses Produkt nicht in $\mathbb{R}^{n,\, n>3}$ gibt. Im Grunde stimmt das nicht. Es gibt das Kreuzprodukt in einem anderen Kontext und unter anderem Namen auch in höherdimensionalen Räumen.

Man spricht dann von Determinante aber darüber sprechen wir erst in Kapitel 3.5. Kreuzprodukte und Determinanten haben eines gemeinsam: Sie beschreiben etwas was mit der Fläche/Volumen zu tun hat, die von zwei ($n \in \{2,3\}$) oder mehr ($n > 3$) Vektoren im $\mathbb{R}^n$ aufgespannt werden. Jetzt schauen wir erst mal nach der

Definition 3.27 (Kreuzprodukt im $\mathbb{R}^2$) *Es seien $a, b \in \mathbb{R}^2$ zwei Vektoren mit*

$$a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Dann heißt die Zahl

$$a \times b := a_1 b_2 - a_2 b_1 \in \mathbb{R}$$

das Kreuzprodukt von a und b.

Satz 3.28 *Der Betrag des Kreuzprodukts im $\mathbb{R}^2$ liefert als Wert den Flächeninhalt des Parallelogramms, das von den beteiligten Vektoren aufgespannt wird.*

Beweis von Satz 3.28 auf Seite 155

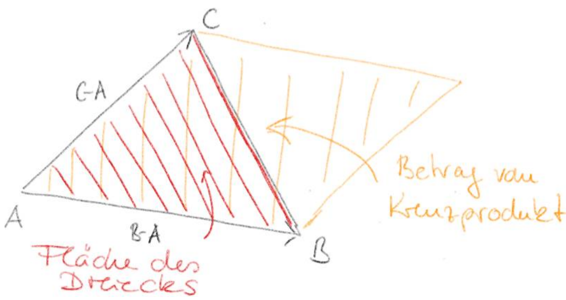
Beispiel 60

Es sei das Dreieck T durch die drei Ecken

$$A = (1,1)\,,\ B = (3,2) \quad \text{und} \quad C = (2,3)$$

gegeben. Dann erhalten wir den Flächeninhalt $|T|$ des Dreiecks durch

$$|(C-A) \times (B-A)| = \frac{1}{2}\left|\begin{pmatrix} 1 \\ 2 \end{pmatrix} \times \begin{pmatrix} 2 \\ 1 \end{pmatrix}\right| = \frac{3}{2}\,.$$



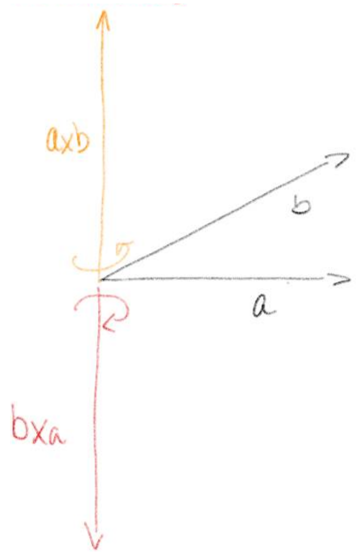
Definition 3.29 (Kreuzprodukt im $\mathbb{R}^3$) *Seien $a, b \in \mathbb{R}^3$. Dann heißt der Vektor*

$$a \times b = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} := \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix} \in \mathbb{R}^3$$

Kreuzprodukt oder auch äußeres Produkt oder auch Vektorprodukt.

Beispiel 61

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



1. Das Kreuzprodukt ist orthogonal zu den beteiligten Vektoren, d.h. es gilt
$$< a, a \times b >=< b, a \times b >= 0\,.$$
2. Der Betrag des Kreuzproduktes $a \times b$ entspricht genau dem Flächeninhalt des Parallelogramms, das durch a und b aufgespannt wird.

Denn

1. Diesen Sachverhalt können Sie leicht nachrechnen, wenn Sie die Definitionen des Kreuz- und des Skalarprodukts und das Ergebnis "aus-x-en".
2. Mit
$$\|a \times b\|_2^2 = \|a\|_2^2 \|b\|_2^2 - < a, b >^2\,,$$
worüber Sie sich in einer Übungsaufgabe klar werden dürfen, gilt dann auch

$$\begin{aligned} \|a \times b\|^2 &= \|a\|^2\|b\|^2 - < a, b >^2 \\ &= \|a\|^2\|b\|^2 - \|a\|^2\|b\|^2 \cos^2\varphi \\ &= \|a\|^2\|b\|^2(1 - \cos^2\varphi) \\ &= \|a\|^2\|b\|^2 \sin^2\varphi \end{aligned}$$

$$\Rightarrow \qquad |a \times b| = \underbrace{|b|}_{\text{Grundseite}} \cdot \underbrace{|a||\sin\varphi|}_{\text{Höhe}}$$

□

Rechenregeln für das Kreuzprodukt:

$$
\begin{array}{lll}
(1) & a \times b = 0 \;\Leftrightarrow\; a || b & \text{(Richtungsgleichheit)} \\
(2) & (\lambda\, a) \times b = \lambda\,(a \times b) = a \times (\lambda\, b) & \\
(3) & (a + b) \times c = a \times c + b \times c & \\
 & a \times (b + c) = a \times b + a \times c & \text{(Distributivität)} \\
(4) & a \times b = -b \times a & \text{(Antikommutativität)}
\end{array}
$$

Es gilt nicht die Assoziativität:

$$a \times (b \times c) \neq (a \times b) \times c\,,!$$

Beispiel 62

$$\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right) \times \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$$

aber

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \times \left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Beispiel 63 Kreuzprodukt im $\mathbb{R}^3$

(a) Das Kreuzprodukt der beiden Richtungsvektoren einer Ebene E in Parameterdarstellung

$$E \;:\; v\,t + w\,s$$

liefert eine Normale (es gibt ja immer 2!) an die Ebene:

$$n = v \times w$$

(b) Es sei $n \perp v$ und $n \perp w$, dann gilt

$$E \;:\; n_1\,x_1 + n_2\,x_2 + n_3\,x_3 = 0\,,$$

denn

$$
\begin{array}{lll}
 & x = v\,t + w\,s & | \cdot n \\
\Rightarrow & < x, n > = < v\,t, n > + < w\,s, n > & \\
\Leftrightarrow & < x, n > = \underbrace{< v, n >}_{=0} t + \underbrace{< w, n >}_{=0} s & \\
\Leftrightarrow & n_1\,x_1 + n_2\,x_2 + n_3\,x_3 = 0 &
\end{array}
$$

(c) Über den Normalvektor kann man also ganz leicht von einer Ebenendarstellung in die andere wechseln. Von Parameterform in Koordinatenform haben wir in Teil (2) gesehen. Umgekehrt ist auch ganz einfach: Es sei

$$n = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

wegen $E\ :\ x_1 + 2\,x_2 + 3\,x_3 = 0$ der Normalvektor von E. Beide Basisvektoren von E sind orthogonal zu n. Einer könnte etwa so aussehen:

$$v = \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix}$$

Und den zweiten erhalten wir über das Kreuzprodukt von n mit v (die Länge ist egal):

$$w = n \times v = \begin{pmatrix} 1 \\ -5 \\ 3 \end{pmatrix}$$

Es ist dann die Parameterdarstellung von E:

$$E\ :\ \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix} t + \begin{pmatrix} 1 \\ -5 \\ 3 \end{pmatrix} s$$

Den Normalvektor an eine Ebene kann man direkt aus deren Koordinatendarstellung ablesen. Auch wenn es keine Ursprungseben ist. (Überzeugen Sie sich gerne selbst davon.)

3.4.2 orthogonale Projektion

Wenn Sie einen Stein in die Hand nehmen und ihn fallen lassen, so ist der Aufprall des Steins auf den Boden eine orthogonale Projektion. Im Grunde war's das schon. In der Schule spricht man von "Lot fällen". So gesehen nichts Neues. Wir wollen das Thema in zweierlei Hinsicht erweitern: Wir betrachten Projektionen von Punkten auf Ebenen. Das ist die kleine Schwester von Texturprojektionen auf geometrische Objekte, wie man es in der Computergraphik praktiziert. Des Weiteren wollen wir den Projektionsbegriff dahingehend abstrahieren, dass wir ihn auch auf Funktionenräume anwenden können. Es wird Ihnen bei der Frequenzanalyse im dritten Semester ein tieferes Verständnis ermöglichen.

Definition 3.30 (Orthogonal- und Orthonormalbasis) *Sei (V, s) ein Prähilbertraum mit Skalarprodukt $s(\cdot,\cdot)$ und Euklidischer Norm $\|\cdot\|_s := \sqrt{s(\cdot,\cdot)}$. Eine Basis $b_1, \dots, b_n$ von V heißt* Orthogonalbasis*, falls gilt:*

$$\forall\, b_i, b_j \; i \neq j \;\; : \;\; s(b_i, b_j) = 0$$

Die Basis heißt Orthonormalbasis *falls gilt:*

$$\forall\, b_i, b_j \; i \neq j \;\; : \;\; s(b_i, b_j) = \delta_{ij}$$

Insbesondere gilt für Orthonormalbasen, dass

$$\|b_i\|_s = 1 \qquad 1 \leq i \leq n\,.$$

Wenn wir von Orthogonalität sprechen denken wir zunächst daran, wie zwei Vektoren zueinander stehen. Es gibt aber auch den Begriff der Orthogonalität von einem Vektor v zu einer Menge U. Das bedeutet, dass v zu jedem Element $u \in U$ orthogonal steht. Bei einerm VR U genügt es zu wissen, dass der v orthogonal zu jedem Basisvektor b_i steht, denn dann gilt für alle $u \in U$, dass $u = \sum_{i=1}^{n} \alpha_i b_i$ und damit:

$$s(v, u) = s(v, \sum_{i=1}^{n} \alpha_i b_i) = \sum_{i=1}^{n} \alpha_i s(v, b_i) = 0$$

Definition 3.31 (orthogonales Komplement) *Sei (V, s) ein Prähilbertraum und $U \subset V$ ein* UVR *von V. Das* orthogonale Komplement *von U in V ist eine Teilmenge $W \subset V$ mit*

$$w \perp U\,, \;\; \forall w \in W\,.$$

Ein Vektor w ist orthogonal zu einem Vektorraum bedeutet, dass w zu jedem beliebigen Element in V orthogonal ist.

Die Basen eines UVR und die seines orthogonalen Komplements sind immer paarweise linear unabhängig.

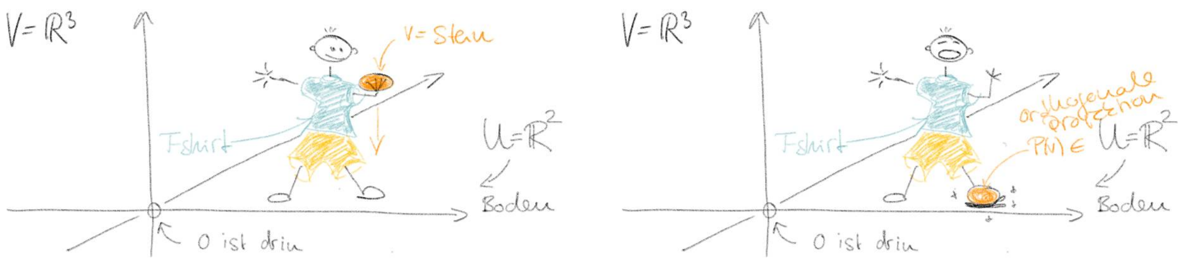
Definition 3.32 (orthogonale Projektion) *Sei V ein VR und U ein UVR von V. Eine Abbildung $P : V \to U$ heißt* orthogonale Projektion *auf U, falls*

$$P(v) \in U \qquad \textit{und} \qquad v - P(v) \perp U$$

für alle $v \in V$.

Dieser Satz ist nun in der Tat ein wenig abstrakt und wir wollen ihn direkt ein wenig entmystifizieren.

Sagen wir etwa V ist der $\mathbb{R}^3$, also der Raum um uns herum (sehr lokal nur) und U sei der Boden auf dem wir stehen; auf Höhe 0. U ist dann also ein Untervektorraum. $v \in V$ ist der Stein in unserer Hand und $P(v) \in U$ beschreibt die Position des Steins auf dem Boden, nachdem wir ihn haben fallen lassen. Der Stein auf dem Boden ist dann die orthogonale Projektion des Steins in unserer Hand. Der Weg des Steins als Gerade gedacht, in Richtung "Stein auf dem Boden also $P(v)$ minus Stein in der Hand, also v, ist orthogonal zum Boden. Jetzt klar, oder?



Sowas im $\mathbb{R}^3$ gedacht ist einfach gerechnet. Einfach die dritte Komponente Null setzen. Wir wollen das aber etwas allgemeiner kennenlernen. Wie wir eine solche Projektion rechnen können für beliebige VRe und UVRe ist in Satz 3.33 beschrieben. Er zeigt uns zunächst wie so eine oprthogonale Projektion $P(v)$ überhaupt aussieht, unter der Voraussetzung, dass der UVR eine Orthogonalbasis besitzt. Es muss allerdings für den UVR eine Orthogonalbasis angegeben sein. Wenn diese noch nicht gegeben ist können wir sie aber mit dem sogenannten Gram-Schmidtschen Orthogonalisierungsverfahren (siehe Satz 3.34) aufstellen. Das geht immer!

Satz 3.33 *Sei* (V, s) *ein Prähilbertraum und* $U \subset V$ *ein* UVR *mit Orthgonalbasis* $\mathcal{B} = \{b_1, \dots, b_n\}$. *Dann gilt für alle* $v \in V$

$$P(v) = \sum_{k=1}^{n} \frac{s(v, b_k)}{\|b_k\|_s^2} b_k \,.$$

Ist $\mathcal{B}$ *sogar eine Orthonormalbasis so verkürzt sich die Formel auf*

$$P(v) = \sum_{k=1}^{n} s(v, b_k)\, b_k \,.$$

Beweis von Satz 3.33 auf Seite 160

Schauen wir uns das zunächst einmal für Ebenen als UVR an. Das kann man sich noch gut vorstellen. Die Ebene muss natürlich eine Ursprungsebene sein und gehen wir für's Erste mal davon aus, dass sie durch zwei orthogonale Vektoren aufgespannt wird.

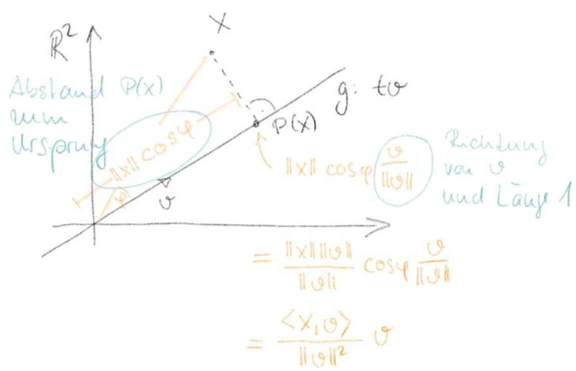
Speziell für (Ursprungs-)Geraden im $\mathbb{R}^2$: Sei $v \in \mathbb{R}^2$, $t \in \mathbb{R}$ und die Gerade

$$g \,:\, v\,t$$

gegeben. Dann lautet die orthogonale Projektion $P(X) \in g$ des Punktes $X \in \mathbb{R}^2$ auf g:

$$P(X) = \frac{< X, v >}{\|v\|^2}\, v$$

Die Graphik rechts veranschaulicht die Formel zur Projektion eines Punktes auf eine Gerade. Sie gilt im Übrigen auch in höheren Raumdimensionen.
Die Erweiterung zur Projektion eines Punktes auf eine Ebene ist die Summe der orthogonalen Projektionen auf die orthogonalen (!) Geraden durch die Basisvektoren der Ebene.



Wenn man das dann sukzessive fortführt erhält man für den $\mathbb{R}^n$ als VR die Formel in Satz 3.33.
Jetzt erst einmal ein Beispiel.

Beispiel 64 orthogonale Projektion eines Punktes auf eine Gerade im $\mathbb{R}^3$ Es sei $X = (1, 2, 3)$ und die

Gerade g gegeben durch

$$g\,:\,t\begin{pmatrix}1\\1\\1\end{pmatrix}.$$

Dann ist

$$P(X)=\frac{<X,v>}{\|v\|^2}\,v=\begin{pmatrix}1\\2\\3\end{pmatrix}\cdot\begin{pmatrix}1\\1\\1\end{pmatrix}\frac{1}{3}\begin{pmatrix}1\\1\\1\end{pmatrix}=2\begin{pmatrix}1\\1\\1\end{pmatrix}.$$

Dass $P(X)\in g$ gilt ist sofort ersichtlich.

$$\langle X-P(X),v\rangle=\left\langle\begin{pmatrix}1\\2\\3\end{pmatrix}-2\begin{pmatrix}1\\1\\1\end{pmatrix},\begin{pmatrix}1\\1\\1\end{pmatrix}\right\rangle=\left\langle\begin{pmatrix}-1\\0\\1\end{pmatrix},\begin{pmatrix}1\\1\\1\end{pmatrix}\right\rangle=0$$

beweist, dass es sich bei $P(X)$ tatsächlich um eine orthogonale Projektion handelt.

Speziell für (Ursprungs-)Ebenen im $\mathbb{R}^3$: Sei $v,w\in\mathbb{R}^3$, $t,s\in\mathbb{R}$ und die Ebene

$$E\,:\,v\,t+w\,s$$

gegeben mit $v\perp w$. Dann lautet die orthogonale Projektion $P(X)\in E$ des Punktes $X\in\mathbb{R}^3$ auf E:

$$P(X)=\frac{<X,v>}{\|v\|^2}\,v+\frac{<X,w>}{\|w\|^2}\,w$$

Achten Sie immer darauf, dass bei der orthogonalen Projektion auf eine Ebene, diese durch orthogonale Vektoren aufgespannt wird, also eine Orthogonalbasis besitzt.

Beispiel 65 orthogonale Projektion eines Punktes auf eine Ebene im $\mathbb{R}^3$ Sei $V=(\mathbb{R}^3,<\cdot,\cdot>)$ und $U\subset V$ ein UVR mit

$$U=\mathrm{Span}\,(a_1,a_2)\ ,\ a_1=\begin{pmatrix}1\\2\\0\end{pmatrix}\ ,\ a_2=\begin{pmatrix}2\\-1\\-2\end{pmatrix}.$$

$\{a_1,a_2\}$ ist eine Orthogonalbasis (überzeugen Sie sich davon!) mit

$$\|a_1\|=\sqrt{5}\quad\text{und}\quad\|a_2\|=3\,.$$

Es sei

$$v=\begin{pmatrix}3\\-1\\1\end{pmatrix}\in V\,.$$

Es ist $v \notin U$ (überzeugen Sie sich davon!) und wir erhalten als orthogonaler Projektion

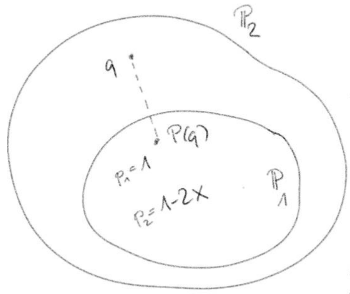
$$P(v) = \frac{< v, a_1 >}{\|a_1\|^2} a_1 + \frac{< v, a_2 >}{\|a_2\|^2} a_2 = \frac{1}{5}\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + \frac{5}{9}\begin{pmatrix} 2 \\ -1 \\ -2 \end{pmatrix} = \frac{1}{45}\begin{pmatrix} 59 \\ -7 \\ -50 \end{pmatrix}$$

Es ist dann

$$v - P(v) = \frac{1}{45}\begin{pmatrix} 76 \\ -38 \\ 95 \end{pmatrix} .$$

überzeugen Sie sich davon, dass $P(v) \in U$ und $v - P(v) \in U^\perp$ gilt.

Wir wollen nun noch ein etwas abstrakteres Beispiel rechnen. Wir versuchen es so klein als möglich zu halten. Bei Polynomen als VR können wir nicht mehr so gut Skizzen zur Veranschuung erzeugen. Solche Skizzen sind dann am Ende doch wieder auf eine Art sehr abstrakt. Siehe Graphik rechts.



$s(q - P(q), p_1) = 0$
$\wedge$
$s(q - P(q), p_2) = 0$
bzw.
$\int_0^1 (q - P(q)) p_1 \, dx = 0$
$\wedge$
$\int_0^1 (q - P(q)) p_2 \, dx = 0$

Beispiel 66 orthogonale Projektion eines kubischen Polynomes auf $\mathbb{P}_1$ Gesucht ist die orthogonale Projektion von $q(x) = x^2 \in \mathbb{P}_2$ auf den UVR

$$U = \mathrm{Span}\,(1, 1 - 2\,x) = \mathbb{P}_1 .$$

U besitze die Orthogonalbasis

$$\mathcal{P} = \{1, 1 - 2\,x\} ,$$

bzgl. des Skalarprodukts

$$s(p, q) = \int_0^1 p(x)q(x)\, dx .$$

Wir prüfen eben die Orthogonalität:

$$\int_0^1 1 \cdot (1 - 2\,x)\, dx = x - x^2 \Big|_0^1 = 1 - 1^2 = 0$$

Bingo!

Wir wollen die nun die orthogonale Projektion von $q(x) = x^2$ auf U berechnen und gehen dabei so vor:

1. orthogonale Projektion $P_1(q)$ von q auf $\mathrm{Span}\,(1)$.

2. orthogonale Projektion $P_2(q)$ von q auf $\operatorname{Span}(1 - 2\,x)$.

3. dann ist
$$P(q) = P_1(q) + P_2(q)$$
dir orthogonale Projektion von q auf U.

Berechnen wir zunächst alle Einzelteile und setzen diese in der Formel am Ende zusammen:

$$\begin{aligned}
s(q, p_1) &= \int\limits_0^1 x^2 \cdot 1\, dx = \frac{1}{3} \\
s(q, p_2) &= \int\limits_0^1 x^2 \cdot (1 - 2\,x)\, dx = \int\limits_0^1 x^2 - 2\,x^3\, dx = \frac{1}{3} - \frac{1}{2} = -\frac{1}{6} \\
\|p_1\|_s^2 &= \int\limits_0^1 1^2\, dx = 1 \\
\|p_2\|_s^2 &= \int\limits_0^1 (1 - 2\,x)^2\, dx = \int\limits_0^1 1 - 4\,x + 4\,x^2\, dx = 1 - 2 + \frac{4}{3} = \frac{1}{3}
\end{aligned}$$

Das ist alles, was wir brauchen und damit gilt dann

$$P(q) = \frac{s(q, p_1)}{\|p_1\|_s^2}\, p_1 + \frac{s(q, p_2)}{\|p_2\|_s^2}\, p_2 = \frac{1}{3} - \frac{1}{2}\,(1 - 2\,x) = x - \frac{1}{6}$$

Prüfen Sie selbst, ob $P(q) \in \mathbb{P}_1$ und $q - P(q) \perp U$ erfüllt ist.

Sowohl in den Beispielen 65 und 66 als auch im Satz 3.33 haben wir davon Gebrauch gemacht, dass eine Orthogonalbasis vorhanden ist. Was aber wenn die angegebene Basis gar nicht orthogonal ist? Ein einfacher Ausweg liefert uns folgender Satz, der besagt, dass man immer eine Orthognal- bzw. Orthonormalbasis konstruieren kann. Lustigerweise basiert diese Konstruktion wieder mithilfe der orthogonalen Projektion und - nein! - hier beißt sich nicht die Katze in den Schwanz. Ehrlich.

Wir sprechen hier immer von endlichdimensionalen VRe! Es lässt sich alles auf VRe mit unendlich aber abzählbar vielen Basen erweitern nicht jedoch auf überabzählbar unendlichdimensionale VRe! Aber keine Sorge. Letzteres ist ohnehin nicht unser Thema.

Satz 3.34 Gram-Schmidtsches Orthogonalisierungsverfahren
Sei (V, s) ein Prähilbertraum mit Basis $\mathcal{A} = \{a_1, \ldots, a_n\}$. Dann gibt es eine eindeutig bestimmte Orthonormalbasis $\mathcal{B} = \{b_1, \ldots, b_n\}$ mit

$$\mathrm{Span}\,(b_1, \ldots, b_n) = \mathrm{Span}\,(a_1, \ldots, a_n)\ .$$

Die Orthonormalbasis ist gegeben durch

$$b_1 = \frac{1}{\|a_1\|_s}\, a_1$$

$$b_k = \frac{1}{\|\tilde{b}_k\|_s}\, \tilde{b}_k \quad \tilde{b}_k = a_k - \sum_{i=1}^{k-1} s(a_k, b_i)\, b_i\ , \ \text{für } k = 2, \ldots, n$$

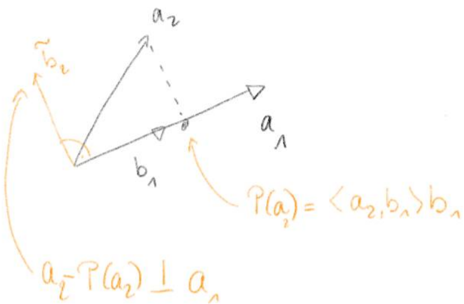
ohne Beweis

Beweisen lässt sich der Satz elegant mit der Methode der Vollständigen Induktion. Wir lassen das und begnügen uns mit anschaulichen Beispielen.

Beispiel 67 Gram-Schmidtsches Orthogonalisierungsverfahren

(a)
Wir suchen die Orthonormalbasis zur Basis

$$\mathcal{A} = \left\{ \begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 0 \\ 4 \\ -1 \end{pmatrix} \right\}$$



$$b_1 = \frac{1}{\|a_1\|}\, a_1 = \frac{1}{3} \begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}$$

$$\tilde{b}_2 = a_2 - < a_2, b_1 > \ b_1 = \frac{-1}{3} \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix}$$

$\tilde{b}_2 \in \mathrm{Span}\,(a_1, a_2) \wedge \tilde{b}_2 \perp a_1$

So fährt man sukzessive mit allen weiteren Dimensionen fort.

$$b_2 = \frac{1}{3\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix} \quad \text{und} \quad b_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

Überzeugen Sie sich davon, dass

$$\mathrm{Span}\,(a_1, a_2) = \mathrm{Span}\,(b_1, b_2)$$

oder Variationen davon gelten.

(b) Sei $V = (C^\circ([0,1]), (\cdot,\cdot)_{L^2})$ der Raum der quadratintegrierbaren, stetigen Funktionen gegeben mit dem Skalarprodukt

$$(f,g)_{L^2} = \int\limits_0^1 f(x)\, g(x)\, dx\,.$$

Für den Unterraum, aufgespannt durch die *Monome* x^k, $k \in \mathbb{N}_0$, also mit der Basis

$$\mathcal{A} = \{a_0, a_1, a_2, a_3, \ldots\} = \{x^0, x^1, x^2, x^3, \ldots\}$$

ist bezüglich dem angegebenen Skalarprodukt eine Orthonormalbasis $\mathcal{B} = \{b_0, b_1, b_2, b_3, \ldots\}$ gesucht.

$$\begin{aligned}
& & (a_0,a_0)_{L^2} &= \int\limits_0^1 1\, dx = 1 \\
&\Rightarrow & b_0 &= 1 \\
& & (a_1,b_0)_{L^2} &= \int\limits_0^1 x \cdot 1\, dx = \frac{1}{2} \\
&\Rightarrow & \tilde{b}_1 &= a_1 - (a_1,b_0)_{L^2} b_0 = x - \frac{1}{2} \\
& & (\tilde{b}_1,\tilde{b}_1)_{L^2} &= \int\limits_0^1 (x - \frac{1}{2})^2\, dx = \frac{1}{12} \\
&\Rightarrow & b_1 &= 2\sqrt{3}\left(x - \frac{1}{2}\right) \\
&\overset{\text{ÜA}}{\Rightarrow} & b_2 &= 6\sqrt{5}\left(x^2 - x + \frac{1}{6}\right)
\end{aligned}$$

Wir erhalten also die entsprechende Orthonormalbasis

$$\mathcal{B} = \left\{1, 2\sqrt{3}\left(x - \frac{1}{2}\right), 6\sqrt{5}\left(x^2 - x + \frac{1}{6}\right), \ldots\right\}.$$

Wenn man aus den Monomen

$$\mathcal{P} = \{1, x, x^2, x^3, \ldots, x^n\} \subset \mathbb{P}_n([-1,1])\,,$$

die eine Basis des $\mathbb{P}_n$ darstellen, mit dem Gram-Schmidtschen Orthogonalisierungsverfahren und dem Skalarprodukt

$$s(p,q) = \int\limits_{-1}^{1} p(x)q(x)\, dx$$

eine Orthogonalbasis erzeugt so erhält man die sogenannten Legendre-Polynome. Diese finden nennenswerte Anwendung in der Numerik und der theoretischen Physik. Darauf wollen wir hier nicht eingehen. Was wir aber an dieser Stelle gut verstehen können ist folgende Überlegung:

Es sei

$$\mathcal{P} = \{p_0, \ldots, p_n\}$$

eine Orthonormalbasis von $\mathbb{P}_n(I)$, $I \subset \mathbb{R}$, d.h. es gilt

$$s(p_i, p_j) = \delta_{ij}\,.$$

Dann lassen sich die Koordinatenvektoren von beliebigen Polynomen $p \in \mathbb{P}_n(I)$ ganz einfach berechnen. Es ist nämlich

$$\begin{aligned}
& & p(x) &= \sum_{i=0}^{n} \alpha_i p_i(x) \\
\Rightarrow & & p(x)p_j(x) &= \sum_{i=0}^{n} \alpha_i p_i(x) p_j(x) \\
\Rightarrow & & \int_I p(x)p_j(x)\,dx &= \int_I \sum_{i=0}^{n} \alpha_i p_i(x) p_j(x)\,dx \\
& & &= \alpha_i \sum_{i=0}^{n} \int_I p_i(x) p_j(x)\,dx \\
& & &= \alpha_i \sum_{i=0}^{n} \underbrace{s(p_i, p_j)}_{=\delta_{ij}} = \alpha_j
\end{aligned}$$

Ohne die Orhtonormalität der Basis würden wir, wie wir es in Kapitel 3.3 gelernt haben, mit der Linearkombination

$$p(x) = \sum_{i=0}^{n} \alpha_i p_i(x)$$

und einem Koeffizientenvergleich auf ein $n \times n$ LGS stoßen und das bei jedem Polynom, dessen Koordinatenvektor wir ermitteln wollen. Das ist sehr rechenaufwendig.

Im dritten Semester werden Sie in der Vorlesung "Signale, Systeme und Sensoren" auf diesen Trick stoßen. Eine Frequenzanalyse ist nichts anderes als die Berechnung von Koordinatenvektoren von Funktionen (Signalen) mit bestimmten Eigenschaften in einem VR, dessen Basis aus Sinusfunktionen unterschiedlicher Frequenz bestehen und die sind orthonormal:

$$\frac{1}{\pi} \int_0^{2\pi} \sin(k\,x)\sin(j\,x)\,dx = \delta_{kj}$$

Mal sehen ob Sie sich in einem Jahr wieder daran erinnern ;-)

Ein letzte Beispiel zur orthogonalen Projektion wollen wir uns noch ansehen.

Beispiel 68 **orthogonale Projektion auf eine Ebene** Gesucht ist die orthogonale Projektion $P(X)$ des Punktes $X = 0 \in \mathbb{R}^3 \setminus E$ mit

$$E \,:\, x_1 + x_2 + x_3 = 1\,.$$

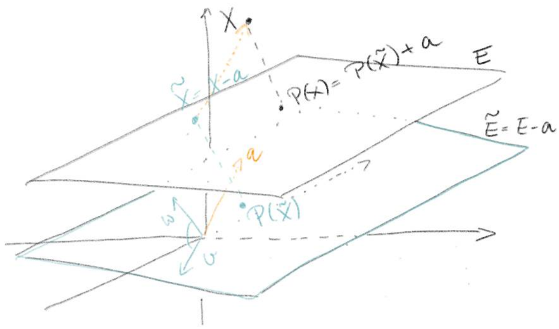
Da gibt es natürlich Varianten. Eine geht so:
Wir wissen sofort, dass

$$\tilde{n} = (1, 1, 1)^T$$

ein Normalvektor und

$$a = (1, 0, 0)^T$$

Stützvektor zu E ist. Diese Ebene ist keine Ursprungsebene, denn $0 \notin E$. Wir müssen nun irgend zwei Normalvektoren zu $\tilde{n}$ finden, denn die sind gleichzeitig Basisvektoren von E.



Es könnte etwa

$$v = (1, 1, -2)^T$$

sein. Damit die Basis von E direkt orthogonal ist[12], wählen wir als zweiten Basisvektor w das Kreuzpordukt von $\tilde{n}$ und v:

$$w = \tilde{n} \times v = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} -3 \\ 3 \\ 0 \end{pmatrix}$$

Das liefert uns also die Parameterdarstellung von E:

$$E \,:\, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} t + \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} s$$

Die orthogonale Projektion kann man so wie wir sie kennengelernt haben nur auf einem VR durchführen. Dazu verschieben wir die ganze Situation, d.h. Ebene und zu prjizierenden Punkt X um $-a$, so dass die Ebene den Ursprung enthält. Es ist dann

$$\tilde{E} = E - a \,:\, \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} t + \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} s \quad \text{und} \quad \tilde{X} = X - a = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$$

[12] Andernfalls wählen wir ähnlich frei ein $\tilde{w}$, das orthogonal zu $\tilde{n}$ und lu zu v ist und berechnen mit dem Gram-Schmidtschen Orthogonalisierungsverfahren aus $\tilde{w}$ ein w mit $w \perp v$ und $w \perp \tilde{n}$.

Die orthogonale Projektion von $\tilde{X}$ auf $\tilde{E}$ lautet dann:

$$P(\tilde{X}) = \frac{< \tilde{X}, v >}{\|v\|^2} v + \frac{< \tilde{X}, w >}{\|w\|^2} w = \frac{-1}{6} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}$$

Das Ergebnis schieben wir um a zurück auf E und erhalten

$$P(x) = P(\tilde{X}) + a = \frac{1}{3} \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Die Orthogonalität sehen wir direkt (ja, oder?) und auch, dass $P(X) \in E$ erfüllt ist. Wenn Sie sich unsicher sind dann prüfen Sie das direkt nach.

Definition 3.35 ((direkte) Summe von Unterräumen) *Seien U_1, U_2 UVR von V. Dann ist die* Summe *von U_1 und U_2 erklärt durch*

$$U_1 + U_2 = \{a + b \mid a \in U_1 \wedge b \in U_2\} .$$

Ist $U_1 \cap U_2 = \{0\}$, so heißt die Summe von U_1 und U_2 direkte Summe *und wir schreiben*

$$U_1 \oplus U_2 .$$

Wenn wir zwei Unterräume addieren so werfen wir alle beteiligten Basisvektoren in einen Topf und erhalten dann zunächst ein Erzeugendensystem. Ist das Erzeugendensystem wieder linear unabhängig so stellt es eine Basis des neuen Raums dar, die Dimensionen der beteiligten Unterräume addieren sich und die Summe ist eine direkte Summe.
Schauen Sie sich dazu nochmal das Comic ganz zu Beginn von Kapitel 3.2 an. Verstehen Sie jetzt?

Beispiel 69 direkte Summe

Es seien die Räume $U_1 = \mathrm{Span}\,(e_1, e_2)$ und $U_2 = \mathrm{Span}\,(e_3)$ gegeben. Dann ist

$$\mathbb{R}^3 = U_1 \oplus U_2 .$$

Beachten Sie an dieser Stelle den Unterschied bei der mengentheoretischen “Zusammenführung” zweier Mengen durch Vereinigung. Es gilt hier

$$\mathbb{R}^3 \neq U_1 \cup U_2 .$$

Es ist etwa der Vektor $a \in \mathbb{R}^3$ mit

$$a = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

weder in U_1 noch in U_2 enthalten und somit auch nicht in $U_1 \cup U_2$. Wohl ist er aber in $U_1 \oplus U_2$ enthalten, da er sich als Linearkombination der durch die Vereinigung der Mengen der Basisvektoren darstellen lässt:

$$a = \underbrace{e_1 + 2\,e_2}_{\in U_1} + \underbrace{3\,e_3}_{\in U_2} \;.$$

3.5 lineare und affine Abbildungen

Definition 3.36 (lineare Abbildung) *Eine Abbildung $\mathcal{A} : V \to W$ zwischen den beiden $\mathbb{K}$–Vektorräumen V und W heißt* lineare Abbildung *oder* Homomorphismus, *wenn für alle Vektoren $x, y \in V$ und alle Skalare $\kappa, \lambda \in \mathbb{K}$*

$$\mathcal{A}(\kappa\,x + \lambda\,y) = \kappa\,\mathcal{A}(x) + \lambda\,\mathcal{A}(y)$$

gilt. Die Menge der Homomorphismen von V nach W wird mit $\mathrm{Hom}\,(V, W)$ bezeichnet.

Eine lineare Abbildung ist eine Abbildung, die Vektoren aus V Vektoren aus W zuordnet. Dabei ist es egal, ob man zunächst eine Lineakombination von Vektoren aus V bildet und dann $\mathcal{A}$ anwendet oder ob man die Linearkombination in W bildet.

(siehe Abb. 4):

$$\underbrace{\mathcal{A}(\overbrace{\kappa\,x + \lambda\,y}^{\in V})}_{\in W} \quad = \quad \underbrace{\kappa\;\underbrace{\mathcal{A}(\overbrace{\;x\;}^{\in V})}_{\in W} + \lambda\;\underbrace{\mathcal{A}(\overbrace{\;y\;}^{\in V})}_{\in W}}_{\in W}$$

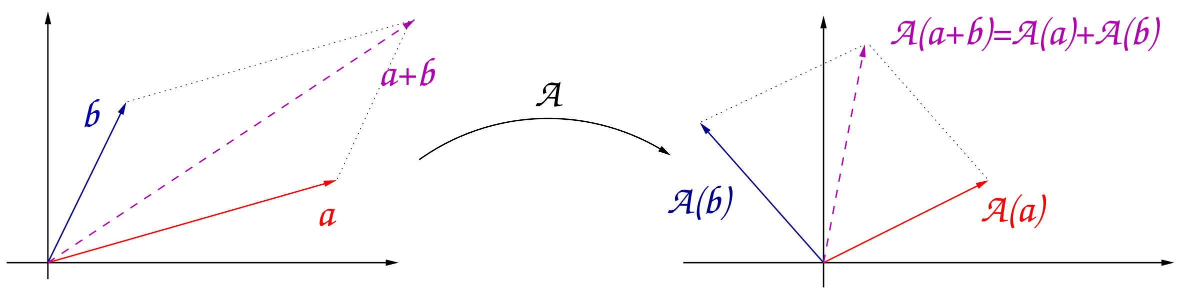


Abbildung 4: Eigenschaft einer linearen Abbildung

Beispiel 70 lineare und nichtlineare Abbildungen

(a)

$$g:\ \mathbb{R} \to \mathbb{R}$$
$$x \mapsto 2\,x$$

ist eine lineare Abbildung.

(b)

$$g:\ \mathbb{R} \to \mathbb{R}$$
$$x \mapsto x^2$$

ist keine lineare Abbildung.

(c)

$$g:\ \mathbb{R} \to \mathbb{R}$$
$$x \mapsto 2$$

ist keine lineare Abbildung.

(d)

$$g:\ \mathbb{R} \to \mathbb{R}$$
$$x \mapsto x + 2$$

ist keine lineare Abbildung.

(e)

$$\int_0^1 \cdot\, dx:\ C^0(\mathbb{R}) \to \mathbb{R}$$
$$f(x) \mapsto \int_0^1 f(x)\, dx$$

ist eine lineare Abbildung.

(f) Na? Ist e^x ein Homomorphismus? Wenn Satz 3.37 nicht erfüllt ist so können wir getrost diese Frage mit NEIN beantworten.

Eine lineare Funktion ist nicht zwingend eine lineare Abbildung.

Satz 3.37 *Für eine lineare Abbildung $\mathcal{A}$ gilt immer*

$$\mathcal{A}(0) = 0\,.$$

Beweis von Satz 3.37 auf Seite 161

Satz 3.37 besagt, dass $\mathcal{A} \in \mathrm{Hom}\,(V, W) \;\Rightarrow\; \mathcal{A}(0) = 0$. Es gilt aber noch lange nicht die Umkehrung, d.h. aus $\mathcal{A}(0) = 0$ folgt nicht, dass $\mathcal{A}$ eine lineare Abbildung ist. Betrachten Sie etwa $\mathcal{A}(x) = x^2$.

Beispiel 71 Matrizen als lineare Abbildung

Zu jeder Matrix $A \in \mathbb{K}^{m\times n}$ erhalten wir mit Hilfe der Matrix–Vektor–Multiplikation durch die Vorschrift

$$\mathcal{A}(x) := A\,x = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^n a_{1k}x_k \\ \sum_{k=1}^n a_{2k}x_k \\ \vdots \\ \sum_{k=1}^n a_{mk}x_k \end{pmatrix}$$

eine Abbildung $\mathcal{A} \in \mathrm{Hom}\,(\mathbb{K}^n, \mathbb{K}^m)$. Diese Abbildung ist linear.

Wärend hingegen eine Abbildung der Form $\mathcal{B} : \mathbb{K}^n \to \mathbb{K}^m$ mit

$$\mathcal{B}(x) := A\,x + b = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^n a_{1k}x_k + b_1 \\ \vdots \\ \sum_{k=1}^n a_{mk}x_k + b_m \end{pmatrix}$$

nicht linear ist. Wir sehen direkt ein, dass $\mathcal{B}(0) \neq 0$ ist. Wir nennen $\mathcal{B}$ eine *affine Abbildung*.

Beispiel 72 mehr lineare Abbildungen

1. Spiegelung an der y-Achse

$$\mathcal{A}(x) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} x$$

2. Spiegelung an der x-Achse

$$\mathcal{A}(x) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} x$$

3. Drehung am Ursprung um den Winkel φ

$$\mathcal{A}(x) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} x$$

4. Drehung am Punkt $M \neq 0$ um den Winkel φ ist eine affine, keine lineare Abbildung

$$\mathcal{B}(x) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} (x - M) + M \notin \mathrm{Hom}\,(\mathbb{R}^2, \mathbb{R}^2)$$

5. Die Projektion $\mathcal{P} : \mathbb{R}^3 \to E_{12}$:

$$\mathcal{P}(x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} x = \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix}$$

Definition 3.38 (Bild & Kern) *Für $\mathcal{A} \in \operatorname{Hom}(V, W)$ heißt ein Vektor $w \in W$* Bildvektor *von $\mathcal{A}$, falls gilt:*

$$\exists\, v \in V \mid \mathcal{A}(v) = w$$

Der UVR bestehend aus der Menge aller Bildvektoren

$$\operatorname{Bild}(\mathcal{A}) := \{\mathcal{A}(v) \in W \mid v \in V\}$$

heißt Bild *von $\mathcal{A}$.*
Die Menge aller Vektoren, deren Bild der Nullvektor ist heißt Kern *von $\mathcal{A}$ und wird mit*

$$\operatorname{Kern}(\mathcal{A}) := \{v \in V \mid \mathcal{A}(v) = 0\}$$

bezeichnet.

Nicht jeder Vektor in W ist ein Bildvektor. Es kann aber Vektoren in W geben, die Bildvektoren von mehreren Vektoren in V sind. Besonders interessant sind Vektoren deren Bildvektor der Nullvektor ist.

Beispiel 73 Bild einer Abbildung Die orthogonale Projektion aus Beispiel 71 ist im Grunde eine Abbildung von $\mathbb{R}^3$ nach $\mathbb{R}^3$. Es ist $\mathcal{P}: \mathbb{R}^3 \to \mathbb{R}^3$ mit

$$\mathcal{P}(x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} x = \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix}$$

gegeben. Das Bild von $\mathcal{P}$ ist aber nur die E_{12}-Ebene. Es ist etwa

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \notin \operatorname{Bild}(\mathcal{P}).$$

Für alle $x = (0, 0, x_3)^T$ ist $\mathcal{P}(x) = 0$, d.h. das der Kern der Abbildung gerade die ganze x_3-Achse ist.

Das Bild einer linearen Abbildung in Matrix-Mal-Vektor-Form $A\,x$ ist gerade die lineare Hülle der Spaltenvektoren der Matrix A.
Der Kern einer linearen Abbildung, definiert über die Matrix $A \in \mathbb{K}^{n \times m}$ gemäß $\mathcal{A}(x) = A\,x$ entspricht gerade der Lösungsmenge $\mathbb{L}$ des homogenen LGS $A\,x = 0$:

$$x \in \operatorname{Kern}(\mathcal{A}) \quad \Leftrightarrow \quad A\,x = 0$$

Beispiel 74 Kern einer Abbildung Es sei $\mathcal{A}(x) = A\,x$ mit

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

Dann gilt

$$\operatorname{Kern} \mathcal{A} = \{x \in \mathbb{R}^3 \mid A\,x = 0\} = \left\{ \begin{pmatrix} 4 \\ -2 \\ 1 \end{pmatrix} t \mid t \in \mathbb{R} \right\}.$$

Beispiel 75 Kern einer Abbildung

Es sei die Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$ mit $f(x, y) = (x - y, y - x)$ gegeben.

(a) f ist linear

(b) f kann auch als Matrix–Vektor Produkt geschrieben werden

$$F : \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - y \\ -x + y \end{pmatrix}$$

(c)

$$\operatorname{Kern}(f) = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 = x_2\}$$
$$\operatorname{Bild}(f) = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 = -x_2\}$$

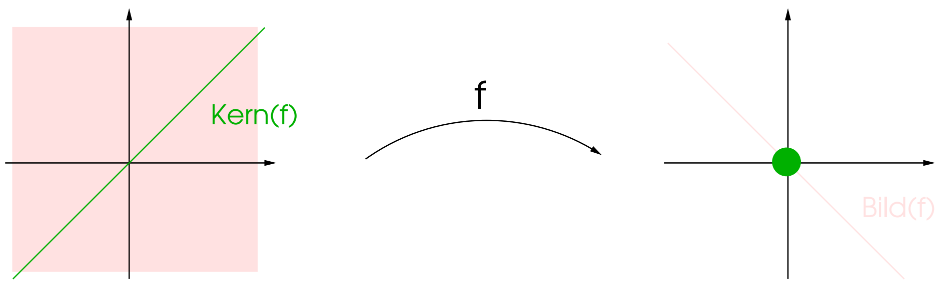


Abbildung 5: Kern und Bild einer linearen Abbildung

Ist $\mathcal{A} \in \operatorname{Hom}(V, W)$, so bilden die Spaltenvektoren das Erzeugendensystem des Bildraumes. Sind alle Spaltenvektoren linear unabhängig, der Rang der Matrix also maximal, so ist das Erzeugendensystem (siehe Definition 3.9, S. 59) eine Basis des Bildraumes.

Es sei etwa $\mathcal{A} \in \operatorname{Hom}(\mathbb{R}^n, \mathbb{R}^n)$ mit $\mathcal{A}(x) = A\,x$ und $\mathcal{K} = \{e_1, \ldots, e_n\}$ die Kanonische Basis des $\mathbb{R}^n$. Dann ist $\mathcal{A}(e_i)$ die i-te Spalte von A, denn

$$\mathcal{A}(e_i) = A\,e_i = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix} =: a_i$$

Damit ist

$$\mathcal{B} = \{a_1, \dots, a_n\}$$

ein Erzeugendensystem von $\mathrm{Bild}\,(\mathcal{A})$. Sind nun die Spalten von A linear unabhängig, d.h. $\mathrm{rang}(A) = n$ oder auch $\det A \neq 0$ so ist $\mathcal{B}$ sogar eine Basis von $\mathrm{Bild}\,(\mathcal{A})$.

Schön, nicht? Wir werden später in Satz 3.43, S. 117 noch einmal darauf zu sprechen kommen.

Doch zunächst noch eine wertvolle Eigenschaft linearer Abbildungen. Diese bilden nämlich linear unabhängige Vektoren, die nicht im Kern enthalten sind auf wiederum linear unabhängige Vektoren ab.

Satz 3.39 *Sei $\mathcal{A} \in \mathrm{Hom}\,(V, W)$ und $v_1, v_2 \in V$ lu. Dann gilt:*

$$w_1 = \mathcal{A}(v_1)\,, \;\; w_2 = \mathcal{A}(v_2) \text{ lu}\,,$$

falls $v_1, v_2 \notin \mathrm{Kern}\,\mathcal{A}$.

Beweis von Satz 3.39 auf Seite 161

Zunächst aber können wir diese Gegebenheit dahingehend gewinnbringend einsetzen, dass wir eine Matrix-Darstellung einer linearen Abbildung einfach und in wenigen Schritten berechnen können. Nicht immer ist das so einfach und direkt ersichtlich wie im Beispiel 75 Teil (c).

Beispiel 76 Matrix-Darstellung berechnen Es sei die Abbildung $f : \mathbb{R}^2 \to \mathbb{R}^2$ mit $f(x, y) = (x - y, y - x)$ gegeben. Gesucht ist die Matrix-Darstellung

$$f(x_1, x_2) = \mathcal{F}(x) = A\,x\,.$$

Die Spaltenvektoren der Matrix A erhalten wir durch die Bildvektoren zur Kanonischen Basis $\{e_1, e_2)\}$ von $\mathbb{R}^2$:

$$\mathcal{F}(e_1) = A\,e_1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} = f(1,0)^T = (1,-1)^T$$
$$\mathcal{F}(e_2) = A\,e_2 = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} = f(0,1)^T = (-1,1)^T$$

Damit ergibt sich insgesamt

$$\mathcal{F}(x) = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} x$$

Wir wollen das an einem etwas komplexeren Beispiel nochmal betrachten.

Beispiel 77 Spiegelung an einer Geraden Es sei die Gerade

$$g = \{x \in \mathbb{R}^2 \mid x = v\,t\,,\; t \in \mathbb{R}\,,\; v \in \mathbb{R}^2\} = \mathrm{Span}\,(v)$$

in der Gaußschen Ebene gegeben. Wir wollen eine lineare Abbildung in Matrixdarstellung angeben, die eine Spiegelung eines Punktes $x \in' R^2$ an der Geraden g vornimmt; wie in Abbildung 6 dargestellt. Zunächst projezieren wir x orthogonal auf die Gerade und erhalten

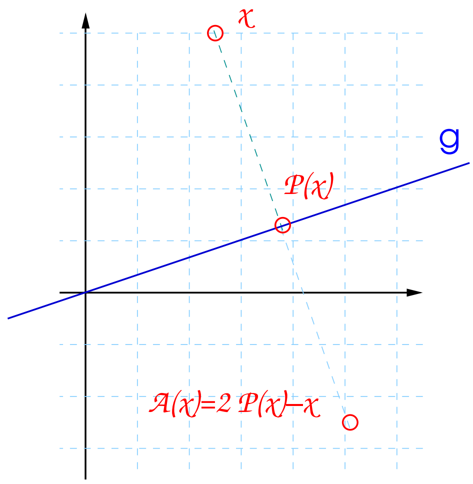


Abbildung 6: Spiegelung an einer Geraden

den Punkt $P(x) \in g$ mit $x - P(x) \perp g$:

$$P(x) = \frac{< x, v >}{\|v\|^2}\, v$$

Es ist dann die Abbildung gegeben durch

$$\mathcal{A}(x) = 2\,P(x) - x = 2\,\frac{< x, v >}{\|v\|^2}\, v - x\,.$$

Was wir aber wollen ist eine Matrixdarstellung, also $\mathcal{A}(x) = S\,x$. Wir erhalten die Spalten der Matrix S durch die Bildvektoren zu den Basisvektoren der Kanonischen Basis:

$$\mathcal{A}(e_1) = \begin{pmatrix} s_{11} \\ s_{21} \end{pmatrix} = 2\,\frac{< e_1, v >}{\|v\|^2}\, v - e_1 = 2\,\frac{v_1}{\|v\|^2}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2\,\frac{v_1^2}{\|v\|^2} - 1 \\ 2\,\frac{v_1\,v_2}{\|v\|^2} \end{pmatrix}$$

$$\mathcal{A}(e_2) = \begin{pmatrix} s_{12} \\ s_{22} \end{pmatrix} = 2\,\frac{< e_2, v >}{\|v\|^2}\, v - e_2 = 2\,\frac{v_2}{\|v\|^2}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2\,\frac{v_2\,v_1}{\|v\|^2} \\ 2\,\frac{v_2^2}{\|v\|^2} - 1 \end{pmatrix}$$

$$\Rightarrow \quad \mathcal{A}(x) = \begin{pmatrix} 2\,\frac{v_1^2}{\|v\|^2} - 1 & 2\,\frac{v_2\,v_1}{\|v\|^2} \\ 2\,\frac{v_1\,v_2}{\|v\|^2} & 2\,\frac{v_2^2}{\|v\|^2} - 1 \end{pmatrix} x$$

Lassen wir es mal konkret werden und wählen wie in Abbildung 6 $v = (3,1)^T$. Dann erhalten wir als Abbildung

$$\mathcal{A}(x) = \frac{1}{5}\begin{pmatrix} 4 & 3 \\ 3 & -4 \end{pmatrix} x\,.$$

Es ist dann der an g gespiegelte Punkt $x = (2.5, 5)^T$ gegeben durch

$$\mathcal{A}\left(\begin{pmatrix} 2.5 \\ 5 \end{pmatrix}\right) = \frac{1}{5}\begin{pmatrix} 4 & 3 \\ 3 & -4 \end{pmatrix}\begin{pmatrix} 5 \\ 10 \end{pmatrix}\frac{1}{2} = \begin{pmatrix} 5 \\ -2.5 \end{pmatrix}$$

Wir können auch eine Abbildung angeben zur Spiegelung an Geraden, die nicht durch den Ursprung verlaufen. Dies sind aber keine UVRe und wir erhalten dann auch keine lineare Abbildung sondern eine affine Abbildung. Eine Spiegelung an $g = a + t\,v$ ist dann in Matrix-Darstellung gegeben durch

$$\mathcal{B}(x) = S\,(x-a) + a = S\,x + (E_2 - S)\,a\,.$$

Beispiel 78 Spiegelung an einer Ebene Es sei die Ebene $E = v\,t + w\,s$ gegeben, die als Spiegel diene. Ist $\{v, w\}$ eine Orthogonalbasis von E so kann man die Projektion eines Punktes $x \in \mathbb{R}^3$ auf E direkt hinschreiben. Gehen wir aber nicht davon aus, so projezieren wir x auf die Einheitsnormale $n = \frac{v \times w}{\|v \times w\|}$ gemäß

$$D(x) =< x, n > \;n\,.$$

Wir erhalten dann als projezierten Punkte auf E

$$P(x) = x - D(x) = x- < x, n > \;n$$

und es ist dadurch der gespiegelte Punkt gegeben durch die Abbildung

$$\mathcal{A}(x) = 2\,P(x) - x = x - 2\,\frac{< x, v \times w >}{\|v \times w\|^2}\,v \times w\,.$$

Das wollen wir jetzt in Matrix-Darstellung haben: Für $n = \frac{v \times w}{\|v \times w\|}$ gilt

$$\mathcal{A}(x) = A\,x = (\mathcal{A}(e_1), \mathcal{A}(e_2), \mathcal{A}(e_3)) = \begin{pmatrix} 1 - 2\,n_1^2 & -2\,n_2\,n_1 & -2\,n_3\,n_1 \\ -2\,n_1\,n_2 & 1 - 2\,n_2^2 & -2\,n_3\,n_2 \\ -2\,n_1\,n_3 & -2\,n_2\,n_3 & 1 - 2\,n_3^2 \end{pmatrix} x\,.$$

Schauen wir uns etwas Konkretes an: Wir wollen den Punkt $x = (1,2,2)^T$ an der Ebene $x_1 - x_2 = 0$ spiegeln. Es ist $n = (1,-1,0)^T/\sqrt{2}$ und damit die Abbildung folgendermaßen gegeben:

$$\mathcal{A}(x) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} x\,.$$

Spiegeln wir $x = (1,2,2)^T$ an der Ebene so erhalten wir den Punkt

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

Fertigen Sie sich gerne eine Skizze dazu an.

Beispiel 79 Abbildung von Strecken Eine Strecke $\overline{PQ}$ zwischen den Punkten P und Q besteht aus allen Punkten auf der Geraden, die durch P und Q verläuft, zwischen P und Q:

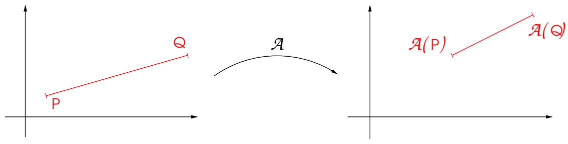
$$\overline{PQ} = \{x \in \mathbb{R}^n \mid x = P + t\,(Q-P)\,,\ t \in [0,1]\}$$

Wird eine ganze Strecke durch $\mathcal{A} \in \mathrm{Hom}\,(\mathbb{R}^n, \mathbb{R}^n)$ abgebildet, so erhalten wir wieder eine Strecke, nämlich

$$\mathcal{A}(\overline{PQ}) = \overline{\mathcal{A}(P)\,\mathcal{A}(Q)}\,,$$

denn für $x \in \overline{PQ}$ gilt

$$\begin{aligned} \mathcal{A}(x) &= \mathcal{A}(P + t\,(Q-P)) \\ &= \mathcal{A}(P) + t\,(\mathcal{A}(Q) - \mathcal{A}(P))\,. \end{aligned}$$

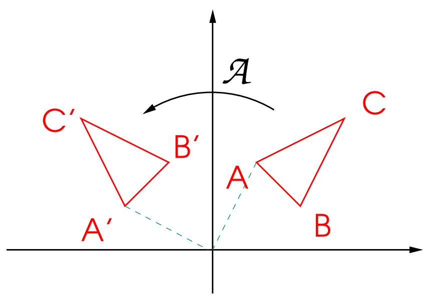


Beispiel 80 Dreieck drehen Es sei das Dreieck T durch die Eckpunkte $A = (1,2)$, $B = (2,1)$ und $C = (3,3)$ gegeben. Wir drehen das Dreieck um den Winkel $\varphi = \frac{\pi}{4}$, indem wir jede Ecke mit

$$\mathcal{C}(x) = \begin{pmatrix} \cos\frac{\pi}{4} & -\sin\frac{\pi}{4} \\ \sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{pmatrix} x$$

abbilden und erhalten das Bilddreieck T':

$$\begin{aligned} A' &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} A = \begin{pmatrix} -2 \\ 1 \end{pmatrix} \\ B' &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} B = \begin{pmatrix} -1 \\ 2 \end{pmatrix} \\ C' &= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} C = \begin{pmatrix} -3 \\ 3 \end{pmatrix} \end{aligned}$$



Beispiel 81 Viereck skalieren Es sei das Viereck Q durch die Punkte

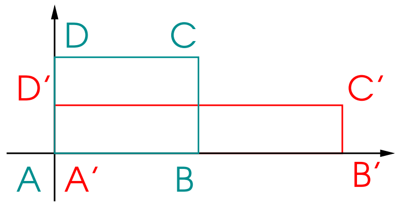
$$A = (0,0)\,,\ B = (3,0)\,,\ C = (3,2) \quad \text{und} \quad D = (0,2)$$

gegeben. Es wird in x_1-Richtung um den Wert $\alpha = 2$ und in x_2-Richtung um den Wert $\beta = \frac{1}{2}$ mittels der Abbildung

$$\mathcal{S}(x) = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} x$$

skaliert. Wir erhalten das neue Viereck mit den Ecken

$$A' = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$
$$B' = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \end{pmatrix},$$
$$C' = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ 1 \end{pmatrix},$$
$$D' = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

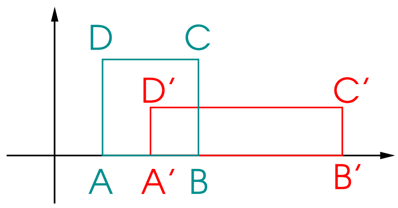


Die linke untere Ecke bleibt wo sie ist, da es zufällig der Nullpunkt ist und die lineare Abbildung diesen unverändert lässt. Wählt man zum Abbilden etwa das Viereck R durch die Punkte

$$A = (1,0)\,,\ B = (3,0)\,,\ C = (3,2) \quad \text{und} \quad D = (1,2)$$

so erhalten wir nach der Abbildung

$$A' = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix},$$
$$B' = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 3 \\ 0 \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \end{pmatrix},$$
$$C' = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 6 \\ 1 \end{pmatrix},$$
$$D' = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$



Definition 3.40 (Komposition) *Es seien $f : U \to V$ und $g : V \to W$ Abbildungen. Dann heißt die Abbildung $g \circ f : U \to W$ mit*

$$(g \circ f)(x) := g(f(x))$$

Komposition oder zusammengesetzte Abbildung von f und g.

Satz 3.41 *Es seien $\mathcal{A} \in \mathrm{Hom}\,(\mathbb{R}^n, \mathbb{R}^m)$ und $\mathcal{B} \in \mathrm{Hom}\,(\mathbb{R}^m, \mathbb{R}^l)$ mit*

$$\mathcal{A}(x) = A\,x \qquad \text{und} \qquad \mathcal{B}(y) = B\,y\,,$$

wobei $A \in \mathbb{R}^{m\times n}$, $x \in \mathbb{R}^n$, $B \in \mathbb{R}^{l\times m}$ und $y \in \mathbb{R}^m$. Dann ist

$$\mathcal{C}(x) := (\mathcal{B} \circ \mathcal{A})\,(x) = C\,x$$

mit

$$C_{ij} = \sum_{k=1}^{m} B_{ik} A_{kj}$$

und $\mathcal{C} \in \mathrm{Hom}\,(\mathbb{R}^n, \mathbb{R}^l)$.

ohne Beweis

Beispiel 82 Verkettung von Abbildungen

Es seien die Abbildungen

$$\mathcal{A}(x) = \begin{pmatrix} 1 & 2 \\ 0 & -4 \end{pmatrix} x \quad \text{und} \quad \mathcal{B}(x) = \begin{pmatrix} 0 & -1 \\ -2 & 0 \end{pmatrix} x$$

gegeben. Dann lautet die Verkettung

$$\begin{aligned} \mathcal{C}(x) = (\mathcal{B} \circ \mathcal{A})(x) = \begin{pmatrix} 0 & -1 \\ -2 & 0 \end{pmatrix} \mathcal{A}(x)\,x \\ = \begin{pmatrix} 0 & -1 \\ -2 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & -4 \end{pmatrix} x = \begin{pmatrix} 0 & 4 \\ -2 & -8 \end{pmatrix} x \end{aligned}$$

und

$$\begin{aligned} \mathcal{D}(x) = (\mathcal{A} \circ \mathcal{B})(x) = \begin{pmatrix} 1 & 2 \\ 0 & -4 \end{pmatrix} \mathcal{B}(x)\,x \\ = \begin{pmatrix} 1 & 2 \\ 0 & -4 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ -2 & 0 \end{pmatrix} x = \begin{pmatrix} -4 & -1 \\ 8 & 0 \end{pmatrix} x \end{aligned}$$

Ganz nebenbei bemerken wir, dass die Verkettung nicht kommutativ ist, was an der Nicht-kommutativität der Matrixmultiplikation liegt.

Beispiel 83 Verkettung von Abbildungen

Bei der Spiegelung eines Punktes im $\mathbb{R}^2$ an einer Geraden $g\ :\ a + t\,v$, $a, v \in \mathbb{R}^2$, $t \in \mathbb{R}$, die nicht durch den Ursprung verläuft muss zunächst das ganze System in den Urspung verschoben werden

$$\mathcal{A}(x) = x - a\,,$$

dann kann an der dadurch erhaltenen Geraden durch den Urspung gespiegelt werden

$$\mathcal{B}(x) = B\,x$$

und schlussendlich muss alles wieder in die ursprüngliche Position zurückgeschoben werden

$$\mathcal{C}(x) = x + a\,.$$

Das sind insgesamt zwei affine und eine lineare Abbildung, die man durch eine Verkettung zu einer Abbildung zusammenfassen kann:

$$\mathcal{D}(x) = (\mathcal{C} \circ \mathcal{B} \circ \mathcal{A})(x) = (\mathcal{B} \circ \mathcal{A})(x) + a = B\,\mathcal{A}(x) + a = B\,(x - a) + a = B\,x + (E_2 - B)\,a$$

Die Verkettung von beliebig vielen linearen Abbildungen ist assoziativ und wieder eine lineare Abbildung. Das bedeutet, dass bei der Hintereinanderschaltung von linearen Abbildungen die Reihenfolge keine Rolle spielt.

$$f \circ (g \circ u) = (f \circ g) \circ u$$

$f \in \mathrm{Hom}\,(V, W)$ sind ungerade, d.h. $f(-v) = -f(v)$. Das liegt ebenfalls an der Linearität.

Definition 3.42 (Bijektivität) *Es sei $\mathcal{A} : V \to W$. Die Abbildung $\mathcal{A}$ heißt*

- *injektiv, falls gilt: $\forall\ x_1, x_2 \in V\ \ x_1 \neq x_2\ :\ \mathcal{A}(x_1) \neq \mathcal{A}(x_2)$,*
- *surjektiv, falls gitl: $\forall\ y \in W\ \exists\ x \in V\ :\ \mathcal{A}(x) = y$ und*
- *bijektiv, falls gilt: $\mathcal{A}$ ist surjektiv und injektiv.*

Gilt überdies noch, dass $\mathcal{A} \in \mathrm{Hom}\,(V, W)$ ist, so heißt $\mathcal{A}$

Monomorphismus	*wenn $\mathcal{A}$ injektiv ist,*
Epimorphismus	*wenn $\mathcal{A}$ surjektiv ist,*
Isomorphismus	*wenn $\mathcal{A}$ bijektiv ist,*
Endomorphismus	*wenn $V = W$ ist und*
Automorphismus	*wenn $V = W$ und $\mathcal{A}$ bijektiv ist.*

Injektivität bedeutet also, dass je zwei verschiedenen Vektoren auch zwei verschiedene Bildvektoren zugeordnet werden. Surjektivität bedeutet, dass jeder Vektor in W erreicht wird. Gilt beides so ist durch $\mathcal{A}$ eine eineindeutige Abbildung gegeben, die jedem (!) Punkt in V genau einen (!) Punkt in W und umgekehrt jedem (!) Punkt in W genau einen (!) in V zuordnet.

Beispiel 84 Kneipenprügelei

Stellen wir uns nun eine Kneipe vor, in dem sich einige Herren (...) feucht-fröhlich amüsieren. Es kommt eine Gruppe weiblicher Amazonen (...) daher, die ziemlich auf Krawall gebürstet ist. Nach einigen frechen Sprüchen seitens der Herren eskaliert die Situation, und die Amazonen verprügeln die Männer.

Wat hat dat denn mit Mathe zu tun? Naja, die Amazonen sind unsere Menge X, die Kerle die Menge Y. Und wann immer eine Amazone x einem Kerl y eine reinhaut, haben wir eine Funktion von X nach Y. Damit das ganze auch wirklich eine Funktion ist (wir erinnern uns, bei einer Funktion wird einem x höchstens ein y zugewiesen), nehmen wir an, dass jede Amazone jeweils maximal einem Mann eine reinhaut.

Es gibt nun drei Möglichkeiten, wie es ablaufen könnte:

Injektiv

Injektiv heißt anschaulich: Jedem y-Wert wird maximal ein x-Wert zugeordnet! Bei unserer Prügelei heißt das, dass niemals zwei Amazonen auf einen Mann gehen würden. Dann gäbe es nämlich zwei x-Werte, die auf einen y-Wert abbilden. Das wäre nicht injektiv! Das heißt aber nicht, dass auch alle Männer etwas abkriegen. Helmut zum Beispiel hat direkt Angst bekommen und sich hinter der Theke versteckt, er kriegt keine rein, es gibt also kein x, welches auf den speziellen y-Wert "Helmut" abgebildet wird. Ist aber egal, trotzdem injektiv!

Surjektiv

Surjektiv ist die Schlägerei, wenn wirklich jeder Mann eine reinkriegt. Das heißt nämlich, dass für jeden y-Wert mindestens ein x da ist, um auf diesen y-Wert abzubilden. Wenn wie vorher tatsächlich Helmut hinter der Bar versteckt bleibt, ist unsere Schlägerei nicht surjektiv, denn dann wäre Helmut genau der y-Wert, der von gar keinem x erreicht wird. Hier lässt sich auch eine spezielle Sache der Surjektivität erklären: Angenommen, Helmut wäre gar nicht da, unsere -Menge also um 1 kleiner, ansonsten aber keinerlei Änderung beim Kampf: Dann wäre die Schlägerei surjektiv, denn außer Helmut kriegen alle etwas ab. Das heißt: Surjektivität hängt ganz besonders von der Y-Menge ab: Wir können die Funktion gleich lassen (an der Schlägerei an sich ändert sich nichts), machen unsere Menge Y aber etwas kleiner und alle y-Werte werden getroffen: Die Funktion, die vorher nicht surjektiv war, ist es nun.

Bijektiv

Bijektiv ist die Rauferei, wenn wirklich jede Amazone genau einen Mann verkloppt und auch kein Mann seinem Schicksal entgeht, also jeder etwas abbekommt. Es gibt also eine 1-zu-1-Beziehung, wiklich jede Amazone haut genau einen Mann und niemand bleibt übrig. Die Abbildung ist damit injektiv(jeder Mann bekommt von maximal einer Amazone etwas ab) und surjektiv(ohne Ausnahme bekommen jeder Mann etwas ab). Das heißt insbesondere: Die Funktion ist umkehrbar! Die Männer, die nun ordentlich eingesteckt haben, drehen den Spieß um, und verprügeln die Amazonen, und zwar ein jeder genau die, von der er eine abbekommen hat! X bildet nun auf Y ab und natürlich wieder 1-zu-1.

Nachdem die Kerle nun also die Frauen zurückgehauen haben, einigte man sich darauf, dass das ganze mit dem Prügeln doch irgendwie Schwachsinn ist und entschloss sich, lieber ein paar Biere zu trinken. Danach ging man gemeinsam nach Hause um Liebe zu machen. Ob das immernoch bijektiv war, oder vielleicht doch mehrere x auf ein y kamen, weiß ich nicht mehr. :)

Zitiert aus: http://matheistkeinarschloch.de/die-injektive-surjektive-bijektive-kneipenpruegelei/

Beispiel 85 injektive, surjektive und bijektive Abbildungen

(a) $\mathcal{P} : \mathbb{R}^3 \to E_{12}$ mit

$$\mathcal{P}(x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} x$$

ist eine surjektive Abbildung. Weil $\mathcal{B} \in \mathrm{Hom}\left(\mathbb{R}^3, E_{12}\right)$ ist $\mathcal{P}$ ein Epimorphismus. $\mathcal{P}$ ist nicht injektiv.

(b) $\mathcal{P} : \mathbb{R}^3 \to \mathbb{R}^3$ mit

$$\mathcal{P}(x) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} x$$

weder injektiv noch surjektiv. Da aber $V = \mathbb{R}^3 = W = \mathbb{R}^3$ ist nennen wir $\mathcal{P}$ einen Endomorphismus.

(c) Die Spiegelung einer Geraden $h : \binom{1}{1} t$ an der Geraden $g : \binom{3}{1} t$ gegeben durch die lineare Abbildung $\mathcal{A} : h \to \mathbb{R}^2$ mit

$$\mathcal{A}(x) = \frac{1}{5} \begin{pmatrix} 4 & 3 \\ 3 & -4 \end{pmatrix} x$$

ist eine injektive Abbildung und somit ein Monomorphismus. Das Bild ist die Gerade

$$h' : \mathcal{A}(h)\, \tilde{t} = \frac{1}{5} \begin{pmatrix} 4 & 3 \\ 3 & -4 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \tilde{t} = \begin{pmatrix} 7 \\ -1 \end{pmatrix} t\,.$$

Somit handelt es sich bei der Abbildung $\mathcal{A}$ nicht um eine surjektive Abbildung.

(d) Schřanken wir den Wertebereich von $\mathcal{A}$ in (c) ein und definieren die Abbildung so: $\mathcal{A} : h \to h'$, so erhalten wir eine injektive und surjektive also eine bijektive Abbildung, bzw. einen Isomorphismus.

Beispiel 85.(d) zeigt, dass die Frage nach Bijektivität im Wesentlichen die Frage nach Injektivität ist. Fehlt die Surjektivität, so schränkt man einfach den Wertebereich der Abbildung auf ihr Bild ein. Das kann man ja immer machen.

Satz 3.43 (Dimensionsformel für lineare Abbildungen) *Es seien* V, W VRe *endlicher Dimension und* $f \in \operatorname{Hom}(V, W)$. *Dann gilt*

$$\operatorname{Dim} V = \operatorname{Dim} \operatorname{Kern} f + \operatorname{Dim} \operatorname{Bild} f\,. \tag{7}$$

Wir sagen auch $\operatorname{rang} f = \operatorname{Dim} \operatorname{Bild} f$ *und schreiben*

$$\operatorname{Dim} V = \operatorname{Dim} \operatorname{Kern} f + \operatorname{rang} f\,.$$

ohne Beweis

Ist die lineare Abbildung über eine Matrix erklärt, sagen wir $\mathcal{A}(x) = Ax$ so bezeichnet $\operatorname{rang} \mathcal{A}$ gerade den Rang der Matrix A: $\operatorname{rang} \mathcal{A} = \operatorname{rang} A$. Die Dimensionsformel (7) lautet dann auch

$$\operatorname{Dim} V = \operatorname{Dim} \operatorname{Kern} \mathcal{A} + \operatorname{rang} A\,.$$

Nun können wir auch einsehen, dass der Kern einer linearen Abbildung aus verschiedenen Gründen interessant ist: Wir können etwa an ihm ablesen, ob die Abbildung injektiv ist. Das ist nämlich genau dann der Fall, wenn der Kern nur aus dem Nullvektor besteht. Dieser liegt immer im Kern, denn es gilt nach Satz 3.37 stets für lineare Abbildungen, dass $\mathcal{A}(0) = 0$.

Definition 3.44 (Umkehrabbildung) *Ist*

$$\begin{aligned} f:\ X &\to Y \\ x &\mapsto f(x) \end{aligned}$$

bijektiv also ein Isomorphismus, so heißt

$$\begin{aligned} f^{-1}:\ Y &\to X \\ f(x) &\mapsto x \end{aligned}$$

die Umkehrabbildung *von* f.
Es ist

$$(f \circ f^{-1})(x) = (f^{-1} \circ f)(x) = x\,.$$

Die Umkehrabbildung eines Isomorphismus ist selbst wieder ein Isomorphismus. Wir schreiben dann auch

$$f:\ X \xrightarrow{\cong} Y$$

Wir sagen die Räume X und Y sind *isomorphe Räume*.

Ist $\mathcal{A} \in \mathrm{Hom}\,(V, W)$ mit $\mathcal{A}(x) = Ax$ ein Isomorphismus erhalten wir die Umkehrabbildung $\mathcal{A}^{-1} \in \mathrm{Hom}\,(W, V)$ über die Inverse der Matrix A^{-1}:

Sei zunächst $\mathcal{A}^{-1}(x) = Bx$, dann gilt

$$(\mathcal{A}^{-1} \circ \mathcal{A})(x) = BAx = x \quad \Leftrightarrow \quad B = A^{-1}$$

Also ist $\mathcal{A}^{-1}(x) = A^{-1}x$.

Beispiel 86 Es sei die Abbildung $\mathcal{A} \in \mathrm{Hom}\,(\mathbb{R}^3, \mathbb{R}^3)$ gegeben mit

$$\mathcal{A}(x) = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix} x\,.$$

Die Matrix A aus $\mathcal{A}(x) = A\,x$ ist regulär und somit ist $\mathcal{A}$ ein Isomorphismus und besitzt eine Umkehrabbildung $\mathcal{A}^{-1}$. Sie errechnet sich aus der Inversen A^{-1} von A und lautet demnach

$$\mathcal{A}^{-1}(x) = \frac{1}{2} \begin{pmatrix} 1 & -2 & 1 \\ 2 & -2 & -2 \\ -1 & 2 & 1 \end{pmatrix} x\,.$$

Die Komposition von $\mathcal{A}$ und $\mathcal{A}^{-1}$ liefert die *Identität* $\mathcal{I}(x) = E_3\,x = x$:

$$\begin{aligned}(\mathcal{A} \circ \mathcal{A}^{-1})(x) &= \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix} \mathcal{A}^{-1}(x) \\ &= \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 1 & 0 & 1 \end{pmatrix} \frac{1}{2} \begin{pmatrix} 1 & -2 & 1 \\ 2 & -2 & -2 \\ -1 & 2 & 1 \end{pmatrix} x \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} x = x\end{aligned}$$

Eine inverse Matrix gibt es, wenn die Matrix A quadratisch ist, also $\mathrm{Dim}\,V = \mathrm{Dim}\,W =: n$ und $\mathrm{rang}A = n$. Das wiederum heißt, dass nur der Nullvektor die Gleichung $Ax = 0$ refüllt, was gleichbedeutend ist mit $\mathrm{Kern}\,\mathcal{A} = \{0\}$. Damit ergibt sich die Dimensionsformel

$$n = \mathrm{Dim}\,V = \mathrm{Dim}\,\mathrm{Kern}\,\mathcal{A} + \mathrm{rang}A = 0 + n \qquad \checkmark$$

Satz 3.45 *Es sei* $f \in \mathrm{Hom}(V, W)$. *Dann gilt*

(i) $\mathrm{Kern}\, f$ *ist ein* UVR *von* V.

(ii) $\mathrm{Kern}\, f = \{0\} \quad \Leftrightarrow \quad f$ *ist injektiv*

(iii) Ist $U \subset V$ *ein* UVR *von* V, *dann ist* $f(U) \subset W$ *ein* UVR *von* W.

Beweis von Satz 3.45 auf Seite 162

Ist $\mathcal{A} \in \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ über eine Matrix $A \in \mathbb{R}^{m \times n}$, mit $A = (a_1, \ldots, a_n)$, $a_i \in \mathbb{R}^m$, $i = 1, \ldots, n$ gemäß

$$\begin{aligned} \mathcal{A}:\ \mathbb{R}^n &\to \mathbb{R}^m \\ x &\mapsto Ax \end{aligned}$$

definiert, so gilt folgendes

- Das Bild von $\mathcal{A}$ wird durch die Spaltenvektoren der Matrix A erzeugt:

$$\mathrm{Bild}\,\mathcal{A} = \mathrm{Span}\,(a_1, \ldots, a_n)$$

- Die Einheitsvektoren e_k, $k = 1, \ldots, n$ von $\mathbb{R}^n$ werden auf die Spaltenvektoren abgebildet:

$$a_k = \mathcal{A}(e_k)$$

- $$\mathrm{Kern}\,\mathcal{A} = \{x \in \mathbb{R}^n \mid Ax = 0\}$$

- Die Dimension des Bildes ist gleich der Anzahl der linear unabhängigen Spalten von $\mathcal{A}$.

Beispiel 87 Basis des Bildraums Es sei $\mathcal{A} \in \mathrm{Hom}(\mathbb{R}^3, \mathbb{R}^3)$ gegeben durch

$$\mathcal{A}(x) = A\,x = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 3 \\ 0 & 1 & -1 \end{pmatrix} x\,.$$

Bei der Berechnung des Kerns lösen wir das homogenen LGS

$$A\,x = 0$$

und bringen dazu die Matrix A in Stufenform:

$$\begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 3 \\ 0 & 1 & -1 \end{pmatrix} \xrightarrow[\text{III+II}]{\text{(II-2 I)/3}} \begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Aus der Stufenform können wir dann zum einen die Lösung des LGS ablesen, was uns den Kern der Abbildung liefert, nämlich

$$\mathrm{Kern}\,\mathcal{A} = \left\{ x \in \mathbb{R}^3 \,\middle|\, x = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} t\,,\ t \in \mathbb{R} \right\}$$

und zum andern sehen wir, dass die ersten beiden Spaltenvektoren linear unabhängig sind und somit die Basis des Bildraums darstellen:

$$\mathrm{Bild}\,\mathcal{A} = \mathrm{Span}\left(\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \right)$$

Es gilt

$$\mathrm{Dim}\,\mathrm{Kern}\,\mathcal{A} + \mathrm{Dim}\,\mathrm{Bild}\,\mathcal{A} = 1 + 2 = 3 = \mathrm{Dim}\,\mathbb{R}^3\,.$$

Kern und Bild eines Endomorphismus $\mathcal{A} \in \mathrm{Hom}\,(V, V)$ bilden eine direkte Summe, welche den Gesamtraum V darstellt.

Wir wollen uns einer besonderen Eigenschaft von Abbildungen widmen, nämlich ihrer Skalierungseigenschaft. Diese erhalten wir über die sogenannte *Determinante* der quadratischen Matrix A, die die lineare oder affin lineare Abbidlung $\mathcal{A}$ beschreibt. Die mathematische Definition der Determinante ist uns zu abstrakt. Wir begnügen uns mit der Berechnungsmethode und einer geometrischen Interpretation. Die ist nämlich in diesem Fall auch - wie so oft - sehr hübsch.

Die Determinante einer 2×2-Matrix ist gerade das Kreuzprodukt der beiden Spaltenvektoren

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{1\star} \times a_{2\star} = a_{11}a_{22} - a_{12}a_{21}\,.$$

Ihr Betrag beschreibt den Flächeninhalt des Parallelogramms, das von den Spaltenvektoren aufgespannt wird.

Beispiel 88

$$\det \begin{pmatrix} 2 & 1 \\ 3 & 8 \end{pmatrix} = \begin{vmatrix} 2 & 1 \\ 3 & 8 \end{vmatrix} = 2 \cdot 8 - 3 \cdot 1 = 13$$

Das Einheitsquadrat $[0, 1]^2$ mit Flcheninhalt 1 wird auf ein Parallelogramm mit Flächeninhalt 13 abgebildet.

Wenn wir uns einmal auf das Gedankenexperiment einlassen wollen, was in einer Raumdimension weiter also in 3d das Analogon wäre, dann würde das bedeuten, dass man etwas sucht, womit man das Volumen eines Spats berechnen kann. Dazu dient zunächst das sogenannte

Definition 3.46 (Spatprodukt) *Seien $u, v, w \in \mathbb{R}^3$, dann heißt die Zahl, gegeben durch*

$$< u, v, w > := < u, v \times w >$$

Spatprodukt der Vektoren u, v und w.

Satz 3.47 *Der betragsmäßige Wert des Spatprodukts $< u, v, w >$ beschreibt das Volumen des Spats, der durch u, v und w aufgespannt wird.*

Beweis von Satz 3.47 auf Seite 161

Nehmen wir einmal an, dass es sich bei den drei Vektoren um die Spalten einer Matrix $A \in \mathbb{R}^{3\times 3}$ handelt:

$$\begin{aligned}
&\left\langle \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} \times \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix} \right\rangle \\
&\qquad = \left\langle \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix}, \begin{pmatrix} a_{22}a_{33} - a_{23}a_{32} \\ a_{13}a_{32} - a_{12}a_{33} \\ a_{12}a_{23} - a_{13}a_{22} \end{pmatrix} \right\rangle \\
&\qquad = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31} \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}
\end{aligned}$$

Beispiel 89 Entwicklung nach der zweiten Spalte

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 3 & -1 & 0 \\ 0 & 1 & 4 \end{pmatrix} = -2 \begin{vmatrix} 3 & 0 \\ 0 & 4 \end{vmatrix} + (-1) \begin{vmatrix} 1 & 3 \\ 0 & 4 \end{vmatrix} - 1 \begin{vmatrix} 1 & 3 \\ 3 & 0 \end{vmatrix}$$
$$= -2\,(3\cdot 4 - 0\cdot 0) - 1\,(1\cdot 4 - 0\cdot 3) - 1\,(1\cdot 0 - 3\cdot 3) = -24 - 4 + 9 = -19$$

Entwicklung nach der dritten Zeile

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 3 & -1 & 0 \\ 0 & 1 & 4 \end{pmatrix} = 0 \begin{vmatrix} 2 & 3 \\ -1 & 0 \end{vmatrix} - 1 \begin{vmatrix} 1 & 3 \\ 3 & 0 \end{vmatrix} + 4 \begin{vmatrix} 1 & 2 \\ 3 & -1 \end{vmatrix}$$
$$= -1\,(1\cdot 0 - 3\cdot 3) + 4\,(1\cdot(-1) - 3\cdot 2) = 9 - 28 = -19$$

Definition 3.48 (Streichmatrix) $A_{\{ij\}} \in \mathbb{K}^{(n-1)\times(n-1)}$ *sei die Matrix, die entsteht, wenn man in* A *die* i*–te Zeile und* j*–te Spalte entfernt. Wir nennen diese Matrix* *Streichmatrix zu* A.

Beispiel 90

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 3 & 1 \\ 4 & 5 & 6 \end{pmatrix} \qquad \Rightarrow \qquad A_{\{32\}} = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}$$

Entwicklung der Determinante von A nach der i–ten Zeile:

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

Entwicklung der Determinante von A nach der j–ten Spalte:

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

Folgerung: *Ist* A eine obere (untere) Dreiecksmatrix, so gilt

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

Man beachte, dass eine Diagonalematrix sowohl eine obere als auch eine untere Dreiecksmatrix ist. Insbesondere gilt

$$\det E = 1$$

Beispiel 91

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{pmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ 0 & a_{33} \end{vmatrix} = a_{11}\, a_{22}\, a_{33}$$

Satz 3.49 *(Eigenschaften der Determinante): Für die Determinante einer Matrix gilt:*

1. *Ist A' aus A durch Vertauschen zweier Zeilen entstanden, so gilt*
$$\det(A') = -\det(A)\,.$$

2. *Ist A' aus A durch Multiplikation einer Zeile von A mit $\kappa \in \mathbb{K}$ entstanden, so gilt*
$$\det(A') = \kappa\det(A)\,.$$
Insbesondere gilt dann auch
$$\det(\kappa\, A) = \kappa^n \det(A)\,.$$

3. *Ist A' durch Addition einer Zeile mit dem Vielfachen einer anderen Zeile entstanden so gilt*
$$\det(A') = \det(A)\,.$$

4. $$\det(A^T) = \det(A)$$

5. $$\det(AB) = \det(A)\det(B)$$

6. $$\det(A) \neq 0 \quad \Leftrightarrow \quad \mathrm{rang}(A) = n$$

ohne Beweis

Folgerung: Es gilt
$$\det A^{-1} = \frac{1}{\det A} \tag{8}$$

Die Tatsache, dass $\det A^T = \det A$ gilt, erlaubt es uns, alle Aussagen, Matrixzeilen betreffend, auch über Matrixspalten zu treffen.

Mit Hilfe der Determinante lassen sich inverse Matrizen berechnen, sofern es eine gibt, versteht sich.

Für $A \in \mathbb{R}^{3\times 3}$ gilt
$$\det A = < a_1, a_2, a_3 >\,,$$
wobei die a_i die Spaltenvektoren von A sind und $< \cdot, \cdot, \cdot >$ das Spatprodukt im $\mathbb{R}^3$ beschreibt. Ist V_e der Einheitswürfel mit Volumen $|V_e| = 1$, der von e_1, e_2, e_3 aufgespannt wird,

und V_a der Spat, den wir durch $\mathcal{A} : V_e \to V_a$ erhalten, so beschreibt $|\mathrm{det} A|$ das Volumen, des von a_1, a_2, a_3 aufgespannten Spates, also $|V_a| = |\mathrm{det} A|$.

> Die Determinante einer zu einer affinen, insbesondere linearen Abbildung gehörenden Matrix ist ein Maß für Streckungs– und Stauchungseigenschaften der Abbildung.

Wir wollen uns das an einigen Beispielen vergegenwärtigen. Es ist in dieser Situation das Beste was man tun kann.

Beispiel 92 Skalierungseigenschaften von Abbildungen

Es sei immer $\Phi_A(x) = A\,x$ und $\Phi_{A,b}(x) = A\,x + b$.

(a) Es seien

$\hat{T}$	gegeben durch die Punkte	$\hat{a} = (0,0)$	$\hat{b} = (1,0)$	$\hat{c} = (0,1)$
T	gegeben durch die Punkte	$a = (2,2)$	$b = (3,4)$	$c = (5,2)$

Die Abbildung

$$\Phi_{A,a}(\hat{x}) = \begin{pmatrix} 1 & 3 \\ 2 & 0 \end{pmatrix} x + \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

bildet $\hat{T}$ auf T ab. Die Determinante der Matrix in $\Phi_{A,a}$ im Betrag beschreibt die Änderung des Flächeninhalts der Dreiecke durch die Abbildung:

$$3 = |T| = |\hat{T}|\,|\det A| = \frac{1}{2}\,6$$

(b) Bei der Streckung in x_1- und x_2-Richtung

$$\Phi_B(x) = \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} x$$

wird das Standarddreieck $\hat{T}$ mit Flächeninhalt $\frac{1}{2}$ auf das Dreieck durch die Punkte $(0,0)$, $(\alpha,0)$ und $(0,\beta)$ mit dem Flächeninhalt

$$\frac{1}{2}\,\alpha\,\beta = |\hat{T}|\,|\det B| = |T|$$

abgebildet.

(c) Die Scherung

$$\Phi_C(x) = \begin{pmatrix} 1 & \alpha \\ 0 & 1 \end{pmatrix} x$$

erhält den Flächeninhalt des abgebildetetn Objekts, denn es gilt

$$\det C = 1\,.$$

(d) Die Drehung

$$\Phi_D(x) = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} x$$

ändert am Flächeninhalt ebenfalls nichts. Es gilt

$$\det D = \cos^2 x + \sin^2 x = 1$$

Besitzt eine affine Abbildung $\Phi_{F,f}(x)$ eine Umkehrabbildung $\Phi_{F,f}^{-1} = F^{-1}(x - f)$ so ist der Skalierungsfaktor der inversen Abbildung gerade der Kehrwert des Skalierungsfaktors von $\Phi_{F,f}$, denn

$$\det F^{-1} = \frac{1}{\det F}\,.$$

Siehe dazu nochmal Folgerung 8.

Beispiel 93 Die Umkehrabbildung von $\Phi_B(x)$ aus Beispiel 92, (b) lautet

$$\Phi_B^{-1}(x) = \begin{pmatrix} \frac{1}{\alpha} & 0 \\ 0 & \frac{1}{\beta} \end{pmatrix} x$$

und der Skalierungsfaktor lautet

$$|\det B^{-1}| = \frac{1}{\alpha\,\beta} = \frac{1}{|\det B|}\,.$$

Bei der Verkettung von Abbildungen multiplizieren sich die Skalierungsfaktoren, da (siehe Satz 3.49, 5.)

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

gilt.

Beispiel 94 Es sei $\Phi_{G,g}$ die Abbildung des Dreiecks S, gegeben durch die Ecken $(0,0)$, $(2,0)$ und $(0,4)$ auf das Dreieck T aus Beispiel 92 (a). Wir bilden $\Phi_{G,g}$ aus der Komposition von $\Phi_{A,a}$ und Φ_B^{-1} mit $\alpha = 2$ und $\beta = 4$:

$$\begin{aligned}
\Phi_{G,g}(x) = \left(\Phi_{A,a} \circ \Phi_B^{-1}\right)(x) &= \begin{pmatrix} 1 & 3 \\ 2 & 0 \end{pmatrix} \Phi_B^{-1}(x) + \binom{2}{2} \\
&= \begin{pmatrix} 1 & 3 \\ 2 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{4} \end{pmatrix} x + \binom{2}{2} = \begin{pmatrix} \frac{1}{2} & \frac{3}{4} \\ 1 & 0 \end{pmatrix} x + \binom{2}{2}
\end{aligned}$$

Es ist dann

$$\frac{3}{4} = |\det G| = |\det A| \cdot |\det B^{-1}| = |\det A| \cdot \frac{1}{|\det B|} = 6 \cdot \frac{1}{8} = \frac{3}{4}\,.$$

3.6 Eigenwerte und Eigenvektoren

Eigenvektoren gehören immer zu einer Abbildung und Eigenwerte immer zu einem Eigenvektor. Ein Eigenvektor einer Abbildung ist ein Vektor, der unter der Abbildung seine Richtung nicht ändert, etwa die Drehachse bei einer Drehung. Er kann aber durchaus seine Länge ändern und der entsprechende Skalierungsfaktor ist dann der Eigenwert. Eigenwert 1 bedeutet demnach, dass ein Vektor, hier der Eigenvektor, nicht nur seine Richtung sondern auch seine Länge beibehält. Er bleibt also unangetastet. Jetzt mathematisch:

Definition 3.50 (Eigenwert und Eigenvektor) *Sei* V *ein* VR *und* $f \in \mathrm{Hom}\,(V, V)$. *Unter einem Eigenvektor* v *von* f *zum Eigenwert* $\lambda \in \mathbb{K}$ *verstehen wir den Vektor* $v \in V \setminus \{0\}$ *mit der Eigenschaft*

$$f(v) = \lambda\, v\,.$$

Eigenvektoren behalten unter der Abbildung f die Richtung bei. Der Eigenwert beschreibt dann gerade die Längenskalierung des Eigenvektors.

Da stellt sich gleich die spannende Frage: "Wie berechnen wir EVen und EWe? Eigenvektoren sind offensichtlich Elemente des Kerns der Abbildung

$$(f - \lambda\,\mathrm{Id}\,)(v)\,.$$

Wir müssen also den Kern einer Abbildung berechnen. Das ist gut, denn das haben wir ja schon kräftig geübt. Allerdings können wir das nur dann konkret, wenn ein $\mathrm{EW}\,\lambda$ schon vorhanden ist. Wir brauchen also erst mal den EW.

Betrachten wir einmal die Abbildung $\mathcal{A} \in \mathrm{Hom}\,(\mathbb{R}^n, \mathbb{R}^m)$ der Gestalt $\mathcal{A}(x) = A\,x$. Wir suchen Vektoren $v \in \mathbb{R}^n$ und Skalare $\lambda \in \mathbb{R}$ mit

$$A\,v = \lambda\,v\,.$$

Oder anders geschrieben

$$(A - \lambda\,E_n)\,v = 0\,.$$

Wir sehen ein homogenes Gleichungssystem und suchen Lösungen v zu verschiedenen Skalaren λ. Den trivialen Fall $v = 0$ lassen wir aus, denn die Null ist kein EV. Damit haben wir unser Problem dahingehend umformuliert, dass wir Elemente des Kerns der Abbildung $(A - \lambda\,E_n)\,v$ berechnen wollen. Und zwar solche, die nicht Null sind. Wann enthält denn der Kern mehr als die Null? Nun, wenn die Matrix singulär ist, das heißt wenn ihre Determinante verschwindet. Im ersten Schritt ermitteln wir diejenigen λ, für die

$$\det\,(A - \lambda\,E_n) = 0$$

gilt.

Beispiel 95 Eigenwert einer Matrix Wir bestimmen die Eigenwerte λ der Matrix

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

Wir suchen λ, so dass

$$\det(A - \lambda E_2) = 0$$

erfüllt ist.

$$\begin{aligned} \det(A - \lambda E_2) &= \det\left(\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\right) \\ &= \det\begin{pmatrix} 1-\lambda & 2 \\ 0 & 1-\lambda \end{pmatrix} \\ &= (1-\lambda)^2 = 0 \end{aligned}$$

Diese Matrix hat den $\mathrm{EW}\,\lambda = 1$. Diejenigen Vektoren, welche auch immer das sind, die unter der Abbildung $\mathcal{A} = A\,x$ ihre Richtung beibehalten, werden auch nicht in ihrer Länge verändert.

Satz 3.51 (charakteristisches Polynom) *Sei V ein VR mit $\mathrm{Dim}\,(V) = n$ und $f \in \mathrm{Hom}\,(V,V)$. Dann gibt es Elemente $a_0, \ldots, a_{n-1} \in \mathbb{K}$ mit*

$$p_f(\lambda) := \det(f - \lambda\,\mathrm{Id}\,) = (-1)^n \lambda^n + \sum_{i=0}^{n-1} a_i \lambda^i \qquad \forall \lambda \in \mathbb{K}\,.$$

p_f heißt charakteristisches Polynom von f.

ohne Beweis

Beispiel 96 Eigevektor einer Matrix zum Eigenwert Wir bestimmen die EVen der Matrix A aus Beispiel 95 zum $\mathrm{EW}\,\lambda = 1$.

$$\mathrm{Kern}\,(A - E_2) = \mathrm{Kern}\begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} = \left\{ v \in \mathbb{R}^2 \,\middle|\, v = \begin{pmatrix} 1 \\ 0 \end{pmatrix} t\,,\ t \in \mathbb{R} \right\}$$

Also behalten alle Vektoren auf der x_1-Achse sowohl ihre Richtung als auch ihre Länge bei. Wir überzeugen und eben noch davon:

$$A\,e_1\,t = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} t = \begin{pmatrix} 1 \\ 0 \end{pmatrix} t$$

Anschaulich klar, denn die Abbildung beschreibt eine Scherung an der x_1-Achse.

Die Eigenwerte einer Matrix A sind die Nullstellen des charakterischtischen Polynoms von $\mathcal{A}$ mit $\mathcal{A}(x) = Ax$. Da nach dem Fundamentalsatz der Algebra ein Polynom über $\mathbb{C}$ immer mindestens eine Nullstelle in $\mathbb{C}$ besitzt gibt es für eine Matrix in $\mathbb{K}^{n\times n}$ mit $\mathbb{K} = \mathbb{C}$ auch immer mindestens einen Eigenwert. Für $\mathbb{K} = \mathbb{R}$ gilt das mit Sicherheit nur dann, wenn $n = 2\,k + 1$ für ein $k \in \mathbb{N}$, das heißt wenn n ungerade ist.

Beispiel 97 Eigenwert einer Drehung im $\mathbb{R}^2$ Wir berechnen die Eigenwerte zur Drehung.

$$\det \begin{pmatrix} \cos\phi - \lambda & \sin\phi \\ -\sin\phi & \cos\phi - \lambda \end{pmatrix} = \lambda^2 - 2\,\cos\phi\,\lambda + 1 = 0$$

$\Leftrightarrow$

$$\lambda_{1,2} = \cos\phi \pm \sqrt{-\sin\phi} = \cos\phi \pm i\,\sin\phi$$

Wir erhalten in diesem Fall nur dann einen reellen EW , wenn $\sin\phi = 0$ ist also wenn $\phi = k\,\pi$, $k \in \mathbb{Z}$. Für den EW gilt dann

$$\lambda = \begin{cases} 1 & \text{für } \phi = 2\,n\,\pi \\ -1 & \text{für } \phi = (2\,n+1)\,\pi \end{cases} \quad , n \in \mathbb{Z}$$

Anschaulich bedeutet das, dass es nur reelle EWe und zugehörige EVen gibt bei Drehungen um 180 bzw. 360 Grad. Wie sieht es in diesem Fall mit den EVen aus?

<u>EVen zum EW $\lambda = 1$:</u>

$$\text{Kern }(D - 1\,E_2) = \text{Kern }\begin{pmatrix} 1-1 & 0 \\ 0 & 1-1 \end{pmatrix} = \mathbb{R}^2$$

<u>EVen zum EW $\lambda = -1$:</u>

$$\text{Kern }(D - (-1)\,E_2) = \text{Kern }\begin{pmatrix} -1+1 & 0 \\ 0 & -1+1 \end{pmatrix} = \mathbb{R}^2$$

Ist der Drehwinkel $\phi \neq k\,\pi$ so erhalten wir nur komplexe EWe . Das macht aber nix. Damit können wir auch rechnen:

<u>EVen zum EW $\lambda = \cos\phi + i\,\sin\phi$:</u>

$$\operatorname{Kern}\begin{pmatrix} \cos\phi-\lambda & \sin\phi \\ -\sin\phi & \cos\phi-\lambda \end{pmatrix} = \operatorname{Kern}\begin{pmatrix} -i\,\sin\phi & \sin\phi \\ -\sin\phi & -i\,\sin\phi \end{pmatrix}$$
$$= \operatorname{Kern}\begin{pmatrix} -i\,\sin\phi & \sin\phi \\ i\,\sin\phi & -\sin\phi \end{pmatrix}$$
$$= \operatorname{Kern}\begin{pmatrix} -i\,\sin\phi & \sin\phi \\ 0 & 0 \end{pmatrix}$$
$$= \left\{ v \in \mathbb{C}^2 \;\middle|\; v = \begin{pmatrix} -i \\ 1 \end{pmatrix} t \right\}$$

EVen zum EW $\lambda = \cos\phi - i\,\sin\phi$:

$$\operatorname{Kern}\begin{pmatrix} \cos\phi-\lambda & \sin\phi \\ -\sin\phi & \cos\phi-\lambda \end{pmatrix} = \operatorname{Kern}\begin{pmatrix} i\,\sin\phi & \sin\phi \\ -\sin\phi & i\,\sin\phi \end{pmatrix}$$
$$= \operatorname{Kern}\begin{pmatrix} i\,\sin\phi & \sin\phi \\ -i\,\sin\phi & -\sin\phi \end{pmatrix}$$
$$= \operatorname{Kern}\begin{pmatrix} i\,\sin\phi & \sin\phi \\ 0 & 0 \end{pmatrix}$$
$$= \left\{ v \in \mathbb{C}^2 \;\middle|\; v = \begin{pmatrix} i \\ 1 \end{pmatrix} t \right\}$$

Beispiel 98 Berechnung der Drehachse Folgende Abbildung beschreibt eine Drehung im Raum um den Winkel $\frac{\pi}{4}$, bzw 45°. Welches ist die Drehachse?

$$\Phi(x) = \frac{1}{3}\begin{pmatrix} 1+\sqrt{2} & 1-\sqrt{2} & 1 \\ 1 & 1+\sqrt{2} & 1-\sqrt{2} \\ 1-\sqrt{2} & 1 & 1+\sqrt{2} \end{pmatrix} x$$

Da es sich um eine Drehung handelt können wir davon ausgehen, dass die Matrix einen Eigenwert $\lambda = 1$ hat. Wir wollen uns dennoch davon überzeugen. Es sollte hier $\det(A - E_3) = 0$ erfüllt sein.

$$\det(A - E_3) = \frac{1}{27}\det\begin{pmatrix} \sqrt{2}-2 & 1-\sqrt{2} & 1 \\ 1 & \sqrt{2}-2 & 1-\sqrt{2} \\ 1-\sqrt{2} & 1 & \sqrt{2}-2 \end{pmatrix}$$
$$= \left((\sqrt{2}-2)^3 + (1-\sqrt{2})^3 + 1 - 3\,(1-\sqrt{2})(\sqrt{2}-2)\right)\frac{1}{27} = 0$$

Gut. Dann suchen wir jetzt einen zugehörigen EV . Diesen finden wir im Kern von $(A - E_3)$. Dazu bringen wir $(A - E_3)$ zunächst auf Stufenform und führen, um die Situation etwas übersichtlicher zu gestalten, Hilfsvariablen $a := \sqrt{2} - 2$ und $b := 1 - \sqrt{2}$ ein. Dann erhalten wir die Umformungen

$$\begin{pmatrix} a & b & 1 \\ 1 & a & b \\ b & 1 & a \end{pmatrix} \xrightarrow[\text{III-b II}]{\text{a II-I}} \begin{pmatrix} a & b & 1 \\ 0 & a^2 - b & ab - 1 \\ 0 & 1 - ab & a - b^2 \end{pmatrix} \xrightarrow{\text{III/(1-ab)-II/(a}^2\text{-b)}} \begin{pmatrix} 1 & b & 1 \\ 0 & a^2 - b & ab - 1 \\ 0 & 0 & 0 \end{pmatrix} .$$

Daraus ergibt sich die Lösungsmenge $\mathbb{L}$ des homogenen LGS $(A - E_3)\, v = 0$ zu

$$\mathbb{L} = \left\{ v \in \mathbb{R}^3 \mid v = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} t\,,\ t \in \mathbb{R} \right\} .$$

Bevor es weiter geht fassen wir das Bisherige zusammen:

Die Vorgehensweise:
Sei $E \in \mathbb{R}^{n \times n}$ die Einheitsmatrix und $A \in \mathbb{K}^{n \times n}$.

$\lambda \in \mathbb{K}$ ist EW von $\mathcal{A}(x) = Ax$, wenn

$$\det(A - \lambda\, E) = 0\,.$$

$v \in \mathbb{K}^n$ ist EV von A zum Eigenwert λ, falls

$$v \in \mathrm{Kern}\,(A - \lambda\, E) \setminus \{0\}\,.$$

Ein Eigenwert kann den Wert 0 annehmen, aber der Nullvektor ist nicht als Eigenvektor zugelassen. Eigenvektoren gehören immer zu einem Eigenwert. Der Nullvektor erfüllt aber die Gleichung $f(0) = \lambda\, 0$ immer, ganz gleich welchen Wert λ annimmt.

Die EVen , die wir erhalten sind immer Elemente der Lösungsmenge eines homogenen LGS. Das sind ja unendlich viele und jeder davon ist ein - und deshalb sagen wir ein und nicht der - EV . In den vorangegangenen Beispielen war die Lösungsmenge jeweils eine Gerade, so dass alle möglichen EVen zueinander linear abhängig sind. Es ist aber möglich, dass wir zu einem einzigen EW EVen erhalten, die Elemente einer Ebene oder eines noch höher dimensionalen Vektorraumes sind.

Beispiel 99

Die Matrix

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

hat die charakteristische Gleichung

$$\det \begin{pmatrix} 2 - \lambda & 0 \\ 3 & 2 - \lambda \end{pmatrix} = (2 - \lambda)^2 = 0\,.$$

Das charakteristische Polynom $p(\lambda) = (2-\lambda)^2$ hat die Nullstelle $\lambda = 2$ mit algebraischer Vielfachheit 2. Nun berechnen wir EVen :

$$(A - 2\,E_2)\,v = \begin{pmatrix} 0 & 0 \\ 3 & 0 \end{pmatrix} v = 0 \quad \Leftrightarrow \quad v = \begin{pmatrix} 0 \\ 1 \end{pmatrix} t\,,\ t \in \mathbb{R}$$

Beispiel 100

Die Matrix

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

hat die charakteristische Gleichung

$$\det \begin{pmatrix} 2-\lambda & 0 \\ 0 & 2-\lambda \end{pmatrix} = (2-\lambda)^2 = 0\,.$$

Das charakteristische Polynom $p(\lambda) = (2-\lambda)^2$ hat die Nullstelle $\lambda = 2$ mit algebr. Vfh 2. Nun berechnen wir EVen :

$$(A - 2\,E_2)\,v = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} v = 0 \quad \Leftrightarrow \quad v = \begin{pmatrix} 1 \\ 0 \end{pmatrix} t + \begin{pmatrix} 0 \\ 1 \end{pmatrix} s\,,\ t, s \in \mathbb{R}$$

In beiden Beispielen erhalten wir einen jeweils den $\mathrm{EW}\,2$ mit algebr. Vfh 2. In Beispiel 99 sind die zugehörigen EVen Elemente einer Geraden, wir sagen er hat die geometrische Vielfachheit 1, und in Beispiel 100 sind die zugehörigen EVen Elemente einer Ebene und wir sagen er habe die geom. Vfh [13] 2. Die algebr. Vfh bezeichnet die Vfh des entsprechenden Linearfaktors im charakteristischen Polynom und die geom. Vfh entspricht der Dimension der Lösungsmenge $\mathbb{L}$ des homogenen LGS $(A - \lambda\,E)\,v = 0$. Diese beiden Vfhen werden noch eine besondere Bedeutung erhalten und sind es Wert weitere Begrifflichkeiten zu definieren:

Definition 3.52 (Eigenraum) *Sei $f \in \mathrm{Hom}\,(V,V)$ und $\lambda \in \mathbb{K}$ ein EW von f. Dann heißt*

$$U_\lambda = \{v \in V \mid f(v) = \lambda\,v\}$$

Eigenraum von f zum $\mathrm{EW}\ \lambda$.

$$0 \in U_\lambda \text{ !!!}$$

[13] Vielfachheit

Satz 3.53 *Sei* $\operatorname{Dim} V = n$ *und* $f \in \operatorname{Hom}(V, V)$. *Sei* $\mu \in \mathbb{K}$ *ein* EW *von* f *der Vfh* r, *das heißt*

$$p_f(\lambda) = (\mu - \lambda)^r \cdot q(\lambda)\,.$$

Dann gilt

$$\operatorname{Dim} U_\mu \leq r\,.$$

ohne Beweis

Satz 3.54 *Sei* $\operatorname{Dim} V = n$ *und* $f \in \operatorname{Hom}(V, V)$ *mit* $\lambda_1, \ldots, \lambda_r$ *pwv* [14] EWe *von* f *und zugehörigen* EVen $v_1, \ldots, v_r \in V$. *Dann sind die* v_i, $i = 1, \ldots, r$ *linear unabhängig.*

Beweis von Satz 3.54 auf Seite 163

Wenn Sie also zu $f \in \operatorname{Hom}(V, V)$ EWe und zugehörige EVen berechnet haben so erzeugen diese Eigenräume, die jeweils UVRe zu V sind. Die Basisvektoren der jeweiligen Eigenräume sind linear unabhängig, was bedeutet, dass die Summe der Eigenräume direkt ist (siehe Definition 3.35, S. 102)

Beispiel 101

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

EWe:

$$\begin{aligned} \det(A - \lambda E_3) &= \det \begin{pmatrix} -\lambda & 0 & -2 \\ 1 & 2-\lambda & 1 \\ 1 & 0 & 3-\lambda \end{pmatrix} \\ &= (2-\lambda) \det \begin{pmatrix} -\lambda & -2 \\ 1 & 3-\lambda \end{pmatrix} = (2-\lambda)^2(1-\lambda) = 0 \end{aligned}$$

EVen zu $\lambda = 2$:

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

$\Rightarrow$

$$U_2 = \left\{ v \in \mathbb{R}^3 \,\middle|\, v = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} t + \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} s \,,\ t, s \in \mathbb{R} \right\}$$

EVen zu $\lambda = 1$:

$$\begin{pmatrix} -1 & 0 & -2 \\ 1 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix} \xrightarrow[\text{III+I, -I}]{\text{II+I}} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

$\Rightarrow$

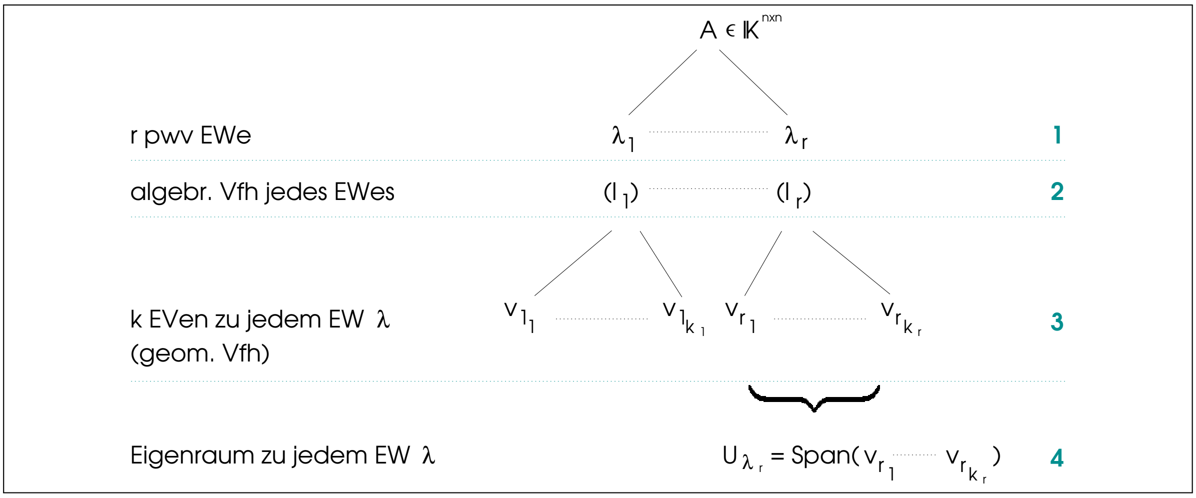
$$U_1 = \left\{ v \in \mathbb{R}^3 \,\middle|\, v = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} t \,,\ t \in \mathbb{R} \right\}$$

Es ist

$$\mathbb{R}^3 = U_1 \oplus U_2 \,.$$

In Beispiel 99 galt $U_2 \subset \mathbb{R}^2$ und in Beispiel 100 dagegen $U_2 = \mathbb{R}^2$.

In Abbildung 7 ist noch einmal alles, was wir in diesem Kapitel besprochen haben in einer Übersicht zusammengefasst.



1 Eine $n \times n$-Matrix hat maximal n EWe und maximal $r \leq n$ pwv EWe. In $\mathbb{C}$ lässt sich das charakt. Pololynom

$$p(\lambda) = \det\left(A - \lambda\, E_n\right) = (\lambda - \lambda_1) \cdots (\lambda - \lambda_n) = (\lambda - \lambda_1)^{l_1} \cdots (\lambda - \lambda_r)^{l_r}$$

immer (Fundamentalsatz der Algebra) in n-Linearfaktoren zerlegen. Dabei sind die λ_1 bis λ_r pwv. In $\mathbb{R}$ wissen wir nur, dass wir für ungerade n wenigstens einen reellen EW erhalten.

2 λ_1 bis λ_r seien also die pwv EWe von A, dann besitzt jeder EW λ_i das Attribut algeb. Vfh. l_i. Es gilt

$$r \leq \sum_{i=1}^{r} l_i = n \quad \text{für } \mathbb{K} = \mathbb{C} \quad \text{und} \quad r \leq \sum_{i=1}^{r} l_i \leq n \quad \text{für } \mathbb{K} = \mathbb{R}\,.$$

3 Jeder EW λ_i besitzt $1 \leq k_i \leq l_i$ (k_i =geom. Vfh.) lu EVen $v_{i_1}, \ldots, v_{i_{k_i}}$.

4 Es ist $U_{\lambda_i} = \text{Span}\,()$ der Eigenraum zum EW λ_i mit Dim $U_{\lambda_i} = k_i$. Es gilt

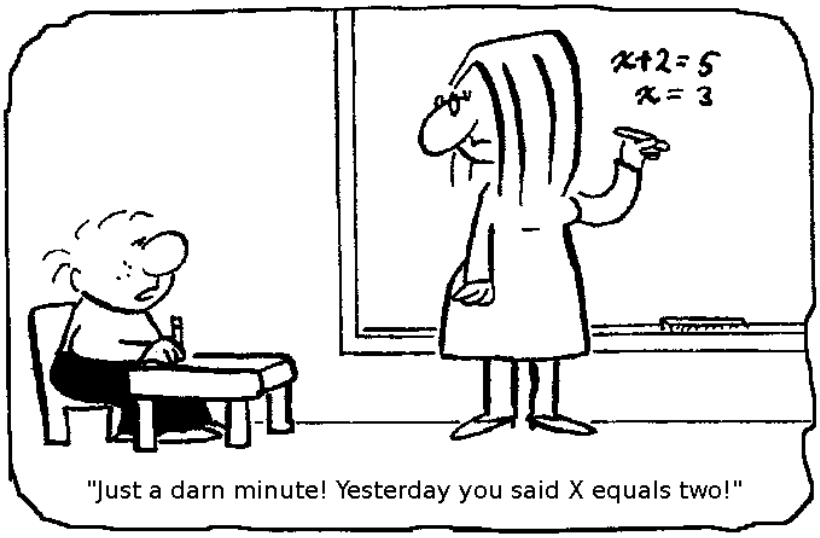
$$\sum_{i=1}^{r} k_i \leq \sum_{i=1}^{r} l_i \leq n \quad \wedge \quad \bigoplus_{i=1}^{r} U_{\lambda_i} = \begin{cases} \subsetneq \mathbb{K}^{n\times n} & \text{für } \sum k_i < n \\ = \mathbb{K}^{n\times n} & \text{für } \sum k_i = n \end{cases} .$$

Abbildung 7: Übersichtsdarstellung zu Eigenwerten, -vektoren und -räumen

Komplexe Zahlen und der schönste Satz der Mathematik 4

Wir behandeln:

- Polar- und Exponentialform von Komplexen Zahlen
- Potenzen und Wurzeln von Komplexen Zahlen



Wir haben bereits Komplexe Zahlen kennengelernt und auch Rechenoperationen wie Addition/Subtraktion und Multiplikation/Division. Die Form, in der Komplexen Zahlen definiert sind, nämlich

$$a + i\,b\,,$$

heißt kartesische Form. Wir werden in den beiden folgenden Unterkapiteln zwei weitere Formen kennenlernen, nämlich die Polarform

$$r\,(\cos\varphi + i\,\sin\varphi\,,$$

was ein vorbereitender Schritt ist zur Exponentialform

$$r\,\mathrm{e}^{i\,\varphi}\,.$$

Die Exponentialform ist von besonders interessanter Struktur und erlaubt ein sehr einfaches Berechnen von Potenzen und Wurzeln von Komplexen Zahlen.

4.1 Komplexe Zahlen in Polarkoordinaten

Komplexe Zahlen kann man als Punkte im $\mathbb{R}^2$ auffassen und entsprechend graphisch darstellen. Es werden dabei der Realteil auf die x-Achse und der Imaginärteil auf der y-Achse aufgetragen. So gesehen können wir komplexe Zahlen auch als Vektoren interpretieren. Sie haben demzufolge einen Abstand zum Ursprung und einen Winkel zur x-Achse. Diese beiden Eigenschaften definieren eine komplexe Zahl ebenso eindeutig, wie ihr Real- und Imaginärteil.

Definition 4.1 (Polarkoordinaten) *Es sei $z = a + i\,b$ eine komplexe Zahl. Dann heißt*

$$r = \|z\| = \sqrt{a^2 + b^2}$$

Betrag von z und

$$\varphi = \operatorname{Arg}(z)$$

Argument von $z \neq 0$. Wir sagen, dass

$$z = a + i\,b$$

in Kartesicher Form und

$$z = r(\cos\varphi + i\,\sin\varphi)$$

in Polarform gegeben ist.

Beispiel 102 Es sei $z = 2 + i\,3$ in kartesischer Form gegeben. Dann ist

$$|z| = \sqrt{13}$$
$$\operatorname{Arg}(\varphi) = \operatorname{atan2}\frac{3}{2} \approx 0.983(56.3°)$$

z in Polarform lautet dann

$$z = \sqrt{13}(\cos 0.983 + i\,\sin 0.983)\,.$$

Das Argument: Winkel eines Vektors im $\mathbb{R}^2$ bezüglich der x–Achse erhalten wir durch die folgende Funktion

$$\operatorname{Arg}(z) := \begin{cases} \arctan\left(\frac{y}{x}\right) & \text{für} \quad x > 0 & \text{(1. und 4. Quadrant)} \\ \arctan\left(\frac{y}{x}\right) + \pi & \text{für} \quad x < 0 \wedge y \geq 0 & \text{(2. Quadrant)} \\ \arctan\left(\frac{y}{x}\right) - \pi & \text{für} \quad x < 0 \wedge y < 0 & \text{(3. Quadrant)} \\ \frac{\pi}{2} & \text{für} \quad x = 0 \wedge y > 0 & \text{(positive } y\text{–Achse)} \\ -\frac{\pi}{2} & \text{für} \quad x = 0 \wedge y < 0 & \text{(negative } y\text{–Achse)} \\ \text{nicht definiert} & \text{für} \quad x = 0 \wedge y = 0 & \text{(im Ursprung)} \end{cases}$$

Der Definitionsbereich ist

$$\mathbb{D}_{\operatorname{Arg}} = [-\pi, \pi] \,.$$

Wenn wir also das Argument einer komplexen Zahl berechnen müssen reicht der Arkustangens mit seinem Wertebereich $\left(\frac{-\pi}{2}, \frac{\pi}{2}\right)$ nicht aus. Wir müssen zu Fuß eine Korrektur vornehmen, je nachdem in welchem Quadranten sich die entsprechende komplexe Zahl in der Gaußschen Ebene befindet.

Wir können nach Belieben zwischen den Darstellungen der kartesischen und der Polarform wechseln. Operationen in kartesischer Form bewirken das Gleiche, wie die entsprechenden in Polarform. Betrachten wir dies anhand der Multiplikation:

Satz 4.2 (Additionstheoreme) *Es gelten folgende Zusammenhänge*

$$\begin{aligned} \sin(x+y) &= \sin x \cos y + \sin y \cos x && (9) \\ \sin(x-y) &= \sin x \cos y - \cos x \sin y && (10) \\ \cos(x+y) &= \cos x \cos y - \sin y \sin x && (11) \\ \cos(x-y) &= \cos x \cos y + \sin x \sin y && (12) \end{aligned}$$

Beweis auf Seite 155

Satz 4.3 (Produkt in Polarform) *Für $z, w \in \mathbb{C}$ gilt*

$$|z \cdot w| = |z||w| \quad \textit{und} \quad \operatorname{Arg}(z \cdot w) = \operatorname{Arg}(z) + \operatorname{Arg}(w) \,.$$

Beweis auf Seite 156

Satz 4.4 (Quotient in Polarform) *Es seien* $z, w \in \mathbb{C}$ *mit*

$$z = r\ (\cos\varphi + i\ \sin\varphi)$$
$$w = s\ (\cos\psi + i\ \sin\psi)$$

gegeben. Dann gilt

$$\frac{z}{w} = \frac{r}{s}\,(\cos(\varphi - \psi) + i\ \sin(\varphi - \psi))\ .$$

Beweis auf Seite 157

Multiplizieren wir zwei komplexe Zahlen so erhalten wir eine neue komplexe Zahl, deren Betrag sich aus dem Produkt der Beträge der Faktoren und deren Argument sich aus der Summe der Argumente der Faktoren ergibt. Anschaulich ist $z_1 \cdot z_2$ eine Zahl, die aus z_1 durch Streckung um $|z_2|$ und Drehung um den Ursprung mit dem Winkel φ_2 entsteht. Die Multiplikation irgendeiner Zahl z_1 mit z_2 entspricht also einer Drehstreckung.

4.2 Komplexe Zahlen in Exponentialform

Das Bonbon kommt immer zum Schluss und hier ist sie, die schönste Formel der Mathematik[15]:

$$\mathrm{e}^{i\,\pi} + 1 = 0 \tag{13}$$

Dahinter stecken im Wesentlichen die sogenannten *Eulergleichungen*:

Satz 4.5 Eulergleichung *Für* $r,\ \varphi \in \mathbb{R}$ *gilt*

$$r\,\mathrm{e}^{i\varphi} = r\,(\cos\varphi + i\ \sin\varphi)\ .$$

Beweis auf Seite 163

Geometrisch interpretiert handelt es sich bei der Menge

$$\left\{\mathrm{e}^{i\varphi} \in \mathbb{C}\ |\ \varphi \in \mathbb{R}\right\}$$

um die Bahnkurve auf dem Kreisrand mit Radius 1.

[15] "Vor einigen Jahren gab es eine Umfrage unter Mathematikern: Welche Formel ist die schönste? Zur Auswahl standen Beispiele aus verschiedenen Bereichen der Mathematik. Am Ende siegte eine Formel aus dem 18. Jahrhundert, die auf den Mathematiker Euler zurückgeht. Er war damals Mathematiker am Hofe Friedrichs des Großen in Berlin." (Quelle: www.welt.de, 2004)

Mit Satz 4.5 ist lässt sich dann die Beziehung (13) direkt einsehen, denn es ist ja

$$r\,\mathrm{e}^{i\varphi} = r\,(\cos\varphi + i\,\sin\varphi)$$

und mit $r = 1$ und $\varphi = \pi$ gilt

$$\begin{aligned} & 1\,\mathrm{e}^{i\,\pi} = 1\,(\cos\pi + i\,\sin\pi) \\ \Leftrightarrow\quad & \mathrm{e}^{i\,\pi} = -1 + i\cdot 0 \\ \Leftrightarrow\quad & \mathrm{e}^{i\,\pi} + 1 = 0\,. \end{aligned}$$

□

Insgesamt haben wir jetzt drei Darstellungsmöglichkeiten von komplexen Zahlen kennengelernt; hier nochmal im Überblick:

Darstellungsmöglichkeiten von komplexen Zahlen: Sei $z \in \mathbb{C}$ mit $z = a + i\,b$. Dann gilt:

$z = a + i\,b$	kartesische Form
$z = \lvert z\rvert\,(\cos\mathrm{Arg}(z) + i\,\sin\mathrm{Arg}(z))$	Polarform
$z = \lvert z\rvert\,\mathrm{e}^{i\mathrm{Arg}(z)}$	Exponentialform

Beispiel 103

Für $z = 2 - 3\,i$ gilt

$$\begin{aligned} |z| &= \sqrt{2^2 + 3^2} = \sqrt{13} \\ \mathrm{Arg}(z) &= \arctan\frac{3}{2} \approx 0.98 \end{aligned}$$

Damit erhalten wir die Polarform (näherungsweise!)

$$z = \sqrt{13}\,\left(\cos 0.98 + i\,\sin 0.98\right)$$

und die Exponentialform (auch näherungsweise)

$$z = \sqrt{13}\,\mathrm{e}^{0.98\,i}$$

Der unschlagbare Vorteil der Exponentialform ist die leichte Handhabbarkeit bei der Berechnung von Potenzen. Lassen Sie sich überraschen mit dem folgenden

Satz 4.6 Potenz von komplexen Zahlen *Es sei $z \in \mathbb{C}$ mit*

$$z = r\,\mathrm{e}^{i\,\varphi}\,.$$

Dann gilt

$$z^n = r^n\,\mathrm{e}^{i\,n\,\varphi}\,.$$

Beweis ist klar

Die n–te Potenz einer komplexen Zahl in Exponentialform ergibt sich aus der n–ten Potenz des Betrages und das n–fache des Argumentes.

Beispiel 104

$$(2-3\,i)^3 = \left(\sqrt{13}\,\mathrm{e}^{0.98\,i}\right)^3 = \sqrt{13}^3\,\mathrm{e}^{2.95\,i}$$

Im Gegensatz zur e–Funtkion im Reellen ist die e–Funktion im Komplexen eine periodische Funktion, weil sich die Periodizität der Sinus– und Kosinusfunktion überträgt:

$$\mathrm{e}^{i\,\varphi} = \cos\varphi + i\,\sin\varphi = \cos(\varphi + 2\,\pi\,k) + i\,\sin(\varphi + 2\,\pi\,k) = \mathrm{e}^{i\,(\varphi+2\,\pi\,k)} \quad \text{für } k \in \mathbb{Z}$$

Beispiel 105

$$(2-3\,i)^3 = \sqrt{13}^3\,\mathrm{e}^{2.95\,i} = \sqrt{13}^3\,\mathrm{e}^{9.23\,i} = \sqrt{13}^3\,\mathrm{e}^{15.52\,i} = \ldots \text{etc.}$$

Definition 4.7 (Wurzel einer komplexen Zahl) *Für eine komplexe Zahl $a \in \mathbb{C}$ und eine natürliche Zahl $n \in \mathbb{N}$ heißt jede Lösung $z \in \mathbb{C}$ der Gleichung*

$$z^n = a$$

n–te Wurzel von a.

Wir sagen nicht z sei die n-te Wurzel von a, also $a^{\frac{1}{n}}$ sondern z sind alle Lösungen der Gleichung $z^n = a$. Der Unterschied besteht streng genommen darin, dass es nicht **die** n-te Wurzel gibt, sondern immer eine ganze, im Grunde "unendlich große" Lösungsmenge vorhanden ist.

Satz 4.8 Wurzel einer komplexen Zahl *Jede Zahl*

$$z_k = \sqrt[n]{|z|}\, \mathrm{e}^{i\left(\frac{\operatorname{Arg}(z)}{n} + \frac{2\pi k}{n}\right)}$$

ist n–te Wurzel von $z \in \mathbb{C}$.
Für $z \neq 0$ und $n \in \mathbb{N}$ hat die Gleichung $z^n = a$ genau n verschiedene Lösungen.

Beweis auf Seite 164

Beispiel 106

Wir suchen alle Lösungen von

$$z^3 = -i$$

. Für $a = -i$ gilt $|a| = 1$ und $\operatorname{Arg}(a) = \frac{3}{2}\pi$. Dann ist a in der Exponentialform gegeben als

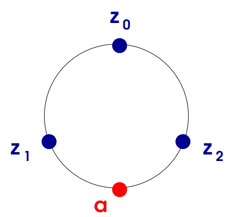
$$a = \mathrm{e}^{\frac{3}{2}\pi i}\,.$$

Die k=ten Wurzeln für $k = 0, 1, 2$ ergeben sich zu

$$z_k = \mathrm{e}^{i\left(\frac{1}{2}\pi + \frac{2}{3}\pi k\right)}\,.$$

Im Einzelnen haben wir

$$z_0 = \mathrm{e}^{i\left(\frac{1}{2}\pi\right)}$$
$$z_1 = \mathrm{e}^{i\left(\frac{7}{6}\pi\right)}$$
$$z_2 = \mathrm{e}^{i\left(\frac{11}{6}\pi\right)}$$
$$z_3 = \mathrm{e}^{i\left(\frac{15}{6}\pi\right)} = \mathrm{e}^{i\left(2\pi + \frac{1}{2}\pi\right)} = \mathrm{e}^{i\,\frac{1}{2}\pi} = z_0$$



a und z liegen hier nur deshalb auf dem selben Kreisrand, weil der Radius 1 ist!

Wurzelziehen ist Potenzieren mit einem Bruch.
Die Lösungen von $z^n = a$ bilden in der komplexen Zahlenebene die Ecken eines regelmäßigen n–Ecks auf dem Kreis mit Radius $\sqrt[n]{|a|}$.

Beispiel 107

(i) Wir lösen die Gleichung

$$z^2 = 4\,.$$

Es ist $a = 4 + i \cdot 0 = 4\,\mathrm{e}^{i \cdot 0}$ und wir erhalten damit die Lösungen

$$z_0 = \sqrt[2]{4}\, e^{i \cdot \frac{0}{2}} = 2$$

und

$$z_1 = \sqrt[2]{4}\, e^{i\cdot\left(\frac{0}{2}+\frac{2\,\pi}{2}\cdot 1\right)} = 2\,(\cos\pi + i\cdot\sin\pi) = 2\,(-1 + i\cdot 0) = -2\,,$$

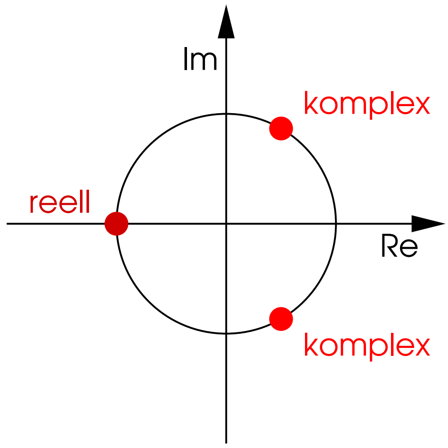
was nun nicht gerade überraschend ist. Es ist aber dennoch schön zu sehen, dass unser Wissen aus dem Reellen auch hier noch Gültigkeit hat.

(ii) Für die Lösungen der Gleichung

$$z^3 = -1$$

gilt mit $-1 = \mathrm{e}^{i\,\pi}$

$$\begin{aligned}
z_0 &= e^{i\cdot\frac{\pi}{3}} = \cos\frac{\pi}{3} + i\,\sin\frac{\pi}{3} = \frac{1}{2} + i\,\frac{\sqrt{3}}{2}\\
z_1 &= e^{i\cdot\left(\frac{\pi}{3}+\frac{2\,\pi}{3}\right)} = \cos\pi + i\,\sin\pi = -1 + i\cdot 0\\
z_2 &= e^{i\cdot\left(\frac{\pi}{3}+\frac{4\,\pi}{3}\right)} = \cos\frac{5\,\pi}{3} + i\,\sin\frac{5\,\pi}{3} = \frac{1}{2} - i\,\frac{\sqrt{3}}{2}
\end{aligned}$$



Wir sehen, dass diese Gleichung drei Lösungen hat, klar, denn es ist ja nach den dritten Wurzeln gefragt, aber nur eine davon ist wieder eine reelle Zahl.

Im Komplexen gibt es immer n verschiedene n-te Wurzeln. Im Reellen ist das nicht im Allgemeinen so. Die Anzahl der Wurzeln hängt davon ab wieviele Wurzeln auf der Realteil-Achse im Koordinatensystem liegen. Es gibt dann nur drei Möglichkeiten: keine Wurzel (Bsp: $\sqrt{-1}$), eine Lösung (Bsp: $\sqrt[3]{-1}$) oder zwei Lösungen (Bsp: $\sqrt{4}$).

A Vorwissen

A.1 Trigonometrie

Definition 1.1 (Sinus und Kosinus)

Zu einem rechtwinkligen Dreieck mit den Kathetenlängen a und b und der Hypotenusenlänge c, bezeichnen wir den Winkel gegenüber der Kante a mit α und definieren folgende Funktionen:

$$\sin\alpha = \frac{\text{Gegenkathete}}{\text{Hypothenuse}} = \frac{a}{c}$$

$$\cos\alpha = \frac{\text{Ankathete}}{\text{Hypothenuse}} = \frac{b}{c}$$

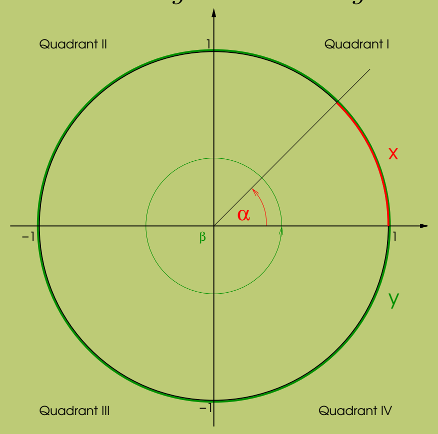
Sinus und Kosinus sind sogenannte trigonometrische Funktionen. Der Begriff Trigonometrie steht für die Beziehung zwischen den Winkeln und den Seitenlängen eines Dreiecks.

tri (drei) gon (ecke) meter (messen)[a]

[a] www.canoonet.de

Bei der Wahl des Argumentes der Sinusfunktion haben wir die Freiheit, zwischen dem Winkel in Grad mit der Kennzeichnung $^\circ$ oder dem sogenannten Bogenmaß zu wählen. Das Bogenmaß beschreibt die Länge eines Kreisbogens am Einheitskreis, wobei Einheitskreis den Kreis mit Radius 1 meint.

Definition 1.2 (Bogenmaß) *Das Bogenmaß x eines Winkels α ist die Länge des Kreisbogens, der dem Winkel α gegenüber liegt, wenn man ihn im Einheitskreis gegen den Uhrzeigersinn abträgt.*



Umfang des Einheitskreises:

Winkel $\beta = 360°$

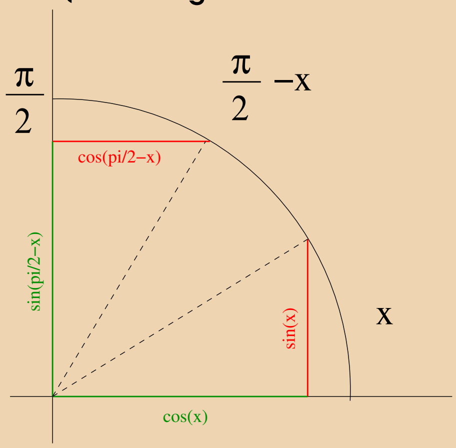
Bogenmaß $y = 2\,\pi$

Kreisabschnitt:

Winkel α

Bogenmaß $x = \frac{2\,\pi}{360°}\alpha$

Satz 1.3 (Beziehungen zwischen Sinus und Kosinus)



$$\cos x = \sin\left(\frac{\pi}{2} - x\right) = \sin\left(\frac{\pi}{2} + x\right)$$
$$\sin x = \cos\left(\frac{\pi}{2} - x\right) = -\cos\left(\frac{\pi}{2} + x\right)$$
$$1 = \sin^2 x + \cos^2 x$$

A.2 Polynome

A.2.1 Fundamentalsatz der Algebra

Hier geht es darum wie man Polynome in Linearfaktoren zerlegt. Das sieht dann so aus:

$$x^3 - x^2 + x - 1 = (x-1)(x+i)(x-i)$$

Links steht irgendein Polynom vom Grad n und rechts ein Produkt aus n Linearfaktoren (lineare Polynome). Das kann man, wenn man komplexe Zahlen zulässt, immer machen. Das sagt der "Fundamentalsatz der Algebra".

Die Frage ist wie. In den Linearfaktoren stehen die Zahlen $-1, i, -i$, was genau die drei Nullstellen vom Polynom auf der linken Seite sind. Wenn man die Nullstellen kennt kann man so eine Zerlegung einfach hinschreiben.

Da stellt sich natürlich gleich die nächste: Wie berechnet man die Nullstellen eines Polynoms? Das ist in der Tat nicht einfach und muss in der Regel näherungsweise gemacht werden. Ist dann aber Thema der Numerik. Analytisch kann man was machen, wenn die Koeffizienten des Polynoms ganzzahlig sind und die gesuchten Nullstellen rational. Dann gibt es Tricks.

Ein Trick besagt, welche Menge von rationalen Zahlen in Frage kommen. Die muss man dann probieren. Das ist wieder sehr aufwendig. Mit dem Horner-Schema kann man das ganz fix durchprobieren. Und wenn eine Zahl passt hat man auch gleich eine Zerlegung miterledigt.

Jetzt also Schritt-für-Schritt im Detail:

Definition 1.4 (Polynom) *Für $i \in \mathbb{N}_0$ mit $0 \leq i \leq n$ seien $a_i \in \mathbb{R}$ mit $a_n \neq 0$. Dann heißt*

$$p(x) := \sum_{j=0}^{n} a_j x^j = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

(reelles) Polynom n–ten Grades.

Schreibweise	*Sprechweise*
$a_0, \ldots, a_n$	*heißen Koeffizienten,*
a_0	*heißt absolutes Glied und*
a_n	*heißt Hauptkoeffizient des Polynoms*

Definition 1.5 (Nullstelle von Polynomen) *Es sei p ein Polynom vom Grad n. x_0 heißt* Nullstelle von p*, falls*

$$p(x_0) = 0$$

gilt. Es gibt dann ein Polynom g vom Grad $n-1$, so dass

$$p(x) = (x - x_0) g(x) \,.$$

Wir nennen den Ausdruck $(x - x_0)$ Linearfaktor *von p.*

Wir können diesen Prozess sukzessive fortführen, so lange das neu entstandene Polynom (hier g) eine weitere Nullstelle besitzt. Wir nennen ein Polynom *reduzibel*, wenn es sich vollständig in Linearfaktoren zerlegen läßt, ansonsten heißt es *irreduzibel*. Wir sagen ein Polynom ist *in* $\mathbb{R}$ *reduzibel*, wenn es sich vollständig in Linearfaktoren aus $\mathbb{R}$ zerlegen läßt, andernfalls nennen wir es *in* $\mathbb{R}$ *irreduzibel*. (Das gilt analog für $\mathbb{N}$, $\mathbb{Z}$ oder $\mathbb{C}$ anstatt $\mathbb{R}$.)

Da sich durch Abspalten des Linearfaktors der Grad des Polynoms um eins erniedrigt kann ein Polynom vom Grad n maximal n Nullstellen haben.

Satz 1.6 Fundamentalsatz der Algebra *Jedes Polynom $p(x)$ mit komplexen Koeffizienten a_i, $i = 0, \ldots, n$ läßt sich als ein Produkt von n Linearfaktoren schreiben:*

$$p(x) = a_n \prod_{j=1}^{n} (x - b_j)$$

Die komplexen Zahlen b_j, $j = 1, \ldots, n$ sind Nullstellen von p, die nicht paarweise verschieden sein müssen. Gibt es l Nullstellen mit $b_{j_1} = \ldots = b_{j_l}$ so sagen wir die Nullstelle b_j habe die algebraische Vielfachheit (Vfh) *l.*

Ein Polynom von ungeradem Grad hat mindestens eine reelle Nullstelle.

Also fassen wir unseren aktuellen Wissensstand kurz in anderen Worten zusammen: Ein Polynom läßt sich in so viele Linearfaktoren zerlegen, wie Nullstellen vorhanden sind. In $\mathbb{C}$ finden wir auf jeden Fall so viele Nullstellen, dass wirklich nur noch lineare Faktoren übrig sind. Beschränken wir uns auf die reellen Zahlen, so kann es passieren, dass quadratische Ausdrücke übrig bleiben.

Beispiel 108 Linearfaktoren

$$p(x) = x^3 - x^2 + x - 1 = (x - 1)(x^2 + 1) = (x - 1)(x + i)(x - i)$$

hat die Nullstellen $1, i, -i$, also eine Nullstelle in $\mathbb{R}$ und zwei in $\mathbb{C}$.

Wie aber finden wir nun Nullstellen bei einem Polynom? Gehen wir mal von dem einfachsten Fall aus (wir überspringen die Polynome vom Grad 1!) und denken an ein Polynom vom Grad 2:

Lösen der quadratischen Gleichung:

Die Lösungen x_1 und x_2 von
$$ax^2 + bx + c = 0$$
sind gegeben durch
$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},$$
sofern der Ausdruck unter der Summe nicht negativ ist. Es gilt folgendes:

1. Ist $b^2 - 4ac < 0$, so gibt es in $\mathbb{R}$ keine Lösung
2. Ist $b^2 - 4ac = 0$, so gibt es genau eine Lösung, nämlich $x_1 = \frac{-b}{2a}$
3. Ist $b^2 - 4ac > 0$, so gibt es genau zwei Lösungen, nämlich
   $x_1 = \frac{-b+\sqrt{b^2-4ac}}{2a}$ und $x_2 = \frac{-b-\sqrt{b^2-4ac}}{2a}$.

Sonderfälle sind solche bei denen der Grad reduziert werden kann. Hat ein Polynom etwa die Darstellung
$$p(x) = a_4x^4 + a_2x^2 + a_0$$
so setzen wir zunächst $y = x^2$ ein und suchen die Nullstelle des quadratischen Polynoms
$$p(y) = a_4y^2 + a_2y + a_0$$
und erhalten dann für die Nullstellen y_1, y_2 von $p(y)$ die Nullstellen
$$x_{1,2} = \pm\sqrt{y_1}\,,\ x_{3,4} = \pm\sqrt{y_2}$$
von $p(x)$.

In allen anderen Fällen müssen Nullstellen geraten werden. Sie haben richtig gelesen: GERATEN! Aber die Mathematik wäre nicht was sie ist, wenn es nicht wenigstens eine kleine Hilfestellung dazu gäbe. Zunächst definieren wir ein neues Zeichen:

Teiler: Seien $a, b \in \mathbb{Z}$ mit $a \neq 0$, dann definieren wir
$$a|b :\Leftrightarrow \frac{b}{a} \in \mathbb{Z}$$
Wir sagen a ist ein *Teiler* von b.

Nullstellen von Polynomen mit ganzen Koeffizienten:
Es sei $p(x)$ ein Polynom vom Grad n mit ganzen Koeffizienten, das heißt $a_i \in \mathbb{Z}$ für $i = 0, \ldots, n$. Dann gilt für eine Nullstelle x_0 von p

1. $$x_0 \in \mathbb{Z} \quad \Rightarrow \quad x_0 | a_0$$
2. $$x_0 \in \mathbb{Q} \quad \Rightarrow \quad x_0 = \frac{r}{s} \ : \ r|a_0 \wedge s|a_n$$

Das liefert uns nun nicht eine alleserschlagende Formel für eine Nullstelle, aber es kann die Suche sehr erleichtern. Wir überzeugen uns wieder anhand eines Beispiels:

Beispiel 109 Polynomdivision

Wir suchen eine Nullstelle des Polynoms

$$p(x) = x^3 - 2x^2 - x + 2$$

Wenn es dazu eine Nullstelle $x_0 \in \mathbb{Z}$ gibt, so ist diese ein Teiler von a_0. Es kommen also die Zahlen

$$\{\pm 1, \pm 2\}$$

in Frage. Probieren führt dazu, dass $x_1 = 1$ eine Nullstelle ist. Wir können nun den Linearfaktor $(x - 1)$ abspalten, um

$$p(x) = (x - 1) \cdot g(x)$$

zu erhalten. $g(x)$ berechnen wir mittels Polynomdivision:

$$\begin{array}{rrrrrrrrl}
 & x^3 & - & 2x^2 & - & x & + & 2 & = (x-1)(x^2 - x - 2) \\
- & x^3 & + & x^2 & & & & & \\
\hline
 & & - & x^2 & - & x & & & \\
 & & + & x^2 & - & x & & & \\
\hline
 & & & & - & 2x & + & 2 & \\
 & & & & + & 2x & - & 2 & \\
\hline
 & & & & & & & 0 &
\end{array}$$

Es ist nun

$$g(x) = (x^2 - x - 2)\,,$$

dessen Nullstellen wir über die "Mitternachtsformel" berechnen können:

$$x_{1,2} = \frac{1 \pm \sqrt{1+8}}{2} = \frac{1 \pm 3}{2}$$

Insgesamt können wir nun mit den drei berechneten Nullstellen $-1, 1, 2$ unser Polynom in drei Linearfaktoren zerlegen:

$$p(x) = x^3 - 2x^2 - x + 2 = (x-1)(x+1)(x-2)$$

Beispiel 110

$$p(x) = x^3 + x - x^2 - 1 = (x-1)\underbrace{(x^2+1)}_{=g(x)}$$

Das dürfen Sie gerne mal selbst rechnen. g hat in diesem Fall keine weiteren reellen Nullstellen. In $\mathbb{C}$ zerfällt es zu

$$(x^2+1) = (x-i)(x+i)\,.$$

Beispiel 111

$$p(x) = 6x^4 + 7x^3 - 13x^2 - 4x + 4$$

hat, wenn es rationale Nullstellen gibt welche von der Form

$$\{\pm\frac{2}{3}\,,\ \pm\frac{4}{3}\,,\ \pm 2\,,\ \pm 4\,,\ \pm 1\,,\ \pm\frac{1}{6}\,,\ \pm\frac{1}{3}\,,\ \pm\frac{1}{2}\}$$

Nun ist das ja trotz Hilfestellung doch noch mal eine ganz schöne Rechnerei. Da wird es doch Zeit eine weitere Hilfestellung hinzuzunehmen; das sogenannte *Horner–Schema*. Es beinhaltet die Idee, dass Polynom so umzuformen, dass man mit weniger Rechenaufwand Werte berechnen kann und funktioniert so:

$$\begin{aligned}
p(x) &= a_n x^n + a_{n-1}x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0 && (14)\\
&= x\left(a_n x^{n-1} + a_{n-1}x^{n-2} + \cdots + a_2 x + a_1\right) + a_0\\
&= x\left(x\left(a_n x^{n-2} + a_{n-1}x^{n-3} + \cdots + a_3 x + a_2\right) + a_1\right) + a_0\\
&= \underbrace{x\,(x\,(\cdots x}_{(n-1)\text{–mal}}\ (a_n x + a_{n-1}) + a_{n-2}) + \cdots + a_1) + a_0 && (15)
\end{aligned}$$

Das ist eine rechenfreundlichere Darstellung, wenn man die benötigten Additionen und Multiplikationen zählt.

Bei der Form in (14) benötigen wir

$$\sum_{j=1}^{n} j = \frac{n(n+1)}{2} = \mathcal{O}(n^2) \qquad \text{Multiplikationen,}$$

während in (15) nur

$$n = \mathcal{O}(n) \qquad \text{Multiplikationen}$$

von Nöten sind. Additionen hat man in beiden Fällen n Stück. Wenn man also viele Auswertungen vornehmen muss, lohnt es sich, das Polynom im Horner–Schema darzustellen.

Horner–Schema:

	-1	5	-3	-9	+
x= 2		-2	6	6	
	-1	3	3	-3	=p(x)

$$\begin{aligned} p(x) &= -x^3 + 5x^2 - 3x - 9 \\ &= x(x(x(-1) + 5) - 3) - 9 \\ p(2) &= 2(2(2(-1) + 5) - 3) - 9 \end{aligned}$$

Beispiel 111 dürfen Sie gerne selbst zu Ende rechnen.

Sie können sich bei der Zerlegung in Linearfaktoren auch die Polynomdivision sparen, wenn Sie bei der Nullstellensuche das Horner–Schema verwenden. Die Koeffizienten des Polynoms $q(x)$ aus $p(x) = (x - x_0)q(x)$ mit $p(x_0) = 0$ lassen sich aus dem Horner–Schema direkt ablesen.

Betrachten wir dazu ein Beispiel:

Beispiel 112

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

$$\begin{aligned} p(x) &= 1x^3 - 2x^2 - 1x + 2 \\ &= (x-2)(1x^2 + 0x - 1) \\ &= (x-2)(x-1)(1x+1) \end{aligned}$$

Die Werte in der untersten Zeile des Schemas entsprechen gerade den Koeffizienten von $q(x)$. Sie können sich also, nachdem Sie eine Nullstelle gefunden haben (!!) durch dieses Schema die nachfolgende Polynomdivision schenken.

Zusammenfassung zu Nullstellen von Polynomen:

- $p \in \mathbb{P}_1$: Wir lösen die lineare Gleichung nach x auf.
- $p \in \mathbb{P}_2$: Mit der Mitternachtsformel bestimmen wir mögliche Nullstellen. Ist der Wert unter der Wurzel

  < 0 : so hat das Polynom keine reellen Nullstellen,

  $= 0$: so hat das Polynom eine reelle Nullstelle mit der Vielfachheit 2 und

  > 0 : so hat das Polynom zwei reelle Nullstellen.
- $p \in \mathbb{P}_n$ mit $n > 2$:
  - Hat das Polynom eine spezielle Form, die es erlaubt auf eine einfachere zurückzuführen, so tun wir das. Etwa $p(x) = x^4 + x^2 - 1$ ersetzen wir zunächst durch $p(y) = y^2 + y - 1$ gemäss $y = x^2$. Wir berechnen zunächst die Nullstellen (sofern in $\mathbb{R}$ vorhanden) von $p(y)$ und ermitteln dann (sofern in $\mathbb{R}$ möglich) aus $y = x^2$ die entsprechenden Werte für x.
  - Hat das Polynom keine spezielle Form, so gehen wir folgendermaßen vor:
    1. Wir setzen $k = 1$ und $p_k(x) = p(x)$
    2. Wir "raten" eine Nullstelle x_k von $p_k(x)$ unter Zuhilfenahme von Regel **??** und dem Horner–Schema zur Arbeitserleichterung.
    3. Mittels Polynomdivision (siehe Beispiel 109) bestimmen wir das Polynom $p_{k+1}(x) = \frac{p_k(x)}{x-x_k}$.
    4. Solange wir Nullstellen finden können setzen wir an dieser Stelle $k = k+1$ und gehen zu 2.

A.2.2 Integration von Polynomen

Für manche Beispiele berechnen wir bestimmte Integrale über Polynome. Integration ist erst Thema im zweiten Semester. Diese Integrationsbeispiele zeigen aber warum die Betrachtung auf einer abstrakten Ebene sinnvoll ist. Die Beispiel sind extra ganz klein gehalten und benutzen nur die folgenden Regeln, die Sie bereits aus der Schule kennen. Hier also eine kleine Wiederholung:

Integration von Polynomen:

$$\int_a^b c\,x^n + d\,x^m\,dx = \frac{c}{n+1}\,x^{n+1} + \frac{d}{m+1}\,x^{m+1}\bigg|_a^b$$

$$= \left(\frac{c}{n+1}\,b^{n+1} + \frac{d}{m+1}\,b^{m+1}\right) - \left(\frac{c}{n+1}\,a^{n+1} + \frac{d}{m+1}\,a^{m+1}\right)$$

In unseren Beispielen wird dann $a = 0$ und $b = 1$ sein und dann zerfällt das ganze zu diesem überschaubaren Ausdruck:

$$\int_0^1 c\,x^n + d\,x^m\,dx = \frac{c}{n+1} + \frac{d}{m+1}$$

Beispiel 113 Integration von Polyonmen

$$\int_0^1 2\,x^3 - \frac{1}{2}\,x\,dx = \frac{1}{2}\,x^4 - \frac{1}{4}\,x^2\bigg|_0^1 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

Mehr passiert hier nicht.

B Beweise

Beweis Satz 2.5:

(a) *Wegen der Reflexivität der ÄR gilt immer $a \in [a]_R$. Somit ist die ÄK nie leer.*

(b) *$\forall b \in [a]_R$ gilt ja, dass aRb ist und wegen der Symmetrie der ÄR ist dann auch bRa und somit $a \in [b]_R$. Das gilt für alle beteiligten Elemente und demnach sind die Mengen gleich. Oder kurz:*

$$b \in [a]_R \quad \Rightarrow \quad aRb \quad \Rightarrow \quad bRa \quad \Rightarrow \quad a \in [b]_R$$

Kurzum: Es ist egal welchen Repräsentanten man wählt.

(c) *Die Schnittmenge muss leer sein, denn sonst würde wegen der Transitivität wieder jedes Element in beiden Klassen enthalten sein.*

$$\left.\begin{array}{lllll} c \in [a]_R & \Rightarrow & aRc & \Rightarrow & cRa \\ c \in [b]_R & \Rightarrow & bRc & & \end{array}\right\} \Rightarrow \quad bRa \quad \Rightarrow a \in [b]_R$$

(d) *Da man für die ÄK nur Elemente in A wählt wird es nicht mehr und da immer $a \in [a]_R$ gilt wegen der Reflexivität fehlt am Ende auch nix.*

$\square$

Beweis Satz 2.7:

Sei $(G, \circ)$ eine Gruppe.

(a) *Es sei $g \in G$, $g' \in G$ Inverses zu g und $g'' \in G$ Inverses zu g'. Dann gilt*

$$\begin{aligned} g \circ g' = n \circ (g \circ g') = (g'' \circ g') \circ (g \circ g') = g'' \circ (g' \circ (g \circ g')) \\ = g'' \circ ((g' \circ g) \circ g') = g'' \circ (n \circ g') = g'' \circ g' = n\,. \end{aligned}$$

(b)

$$g \circ n = g \circ (g' \circ g) = (g \circ g') \circ g = n \circ g = g$$

(c) *Seien n und $\tilde{n}$ neutrale Elemente bzgl. $\circ$. Dann gilt*

$$\tilde{n} = \tilde{n} \circ n = n\,.$$

(d) *Sei $g \in G$ und $g', \tilde{g}' \in G$ zwei inverse Elemente mit*

$$g \circ g' = n \quad \text{und} \quad g \circ \tilde{g}' = n\,.$$

Dann gilt

$$\tilde{g}' = \tilde{g}' \circ n = \tilde{g}' \circ (g \circ g') = (\tilde{g}' \circ g) \circ g' = n \circ g' = g'\,.$$

(e) Das ist klar, denn es gilt

$$n = n \circ n\,.$$

(f) Für $a, a', a'' \in G$ mit $a \circ a' = n$ und $a' \circ a'' = n$ gilt

$$a = a \circ n = a \circ (a' \circ a'') = (a \circ a') \circ a'' = n \circ a'' = a''\,.$$

(g)

$$\begin{aligned} h^{-1} \circ g^{-1} &= n \circ (h^{-1} \circ g^{-1}) = \big((g \circ h)^{-1} \circ (g \circ h)\big) \circ (h^{-1} \circ g^{-1}) \\ &= (g \circ h)^{-1} \circ \big((g \circ h) \circ (h^{-1} \circ g^{-1})\big) = (g \circ h)^{-1} \circ \big((g \circ (h \circ h^{-1}) \circ g^{-1})\big) \\ &= (g \circ h)^{-1} \circ \big((g \circ n) \circ g^{-1})\big) = (g \circ h)^{-1} \circ (g \circ g^{-1}) = (g \circ h)^{-1} \circ n = (g \circ h)^{-1} \end{aligned}$$

(h) Die erste Gleichung ist eine Übungsaufgabe. Wir machen die zweite. Es gilt

$$b = b \circ n = b \circ (a \circ a') = (b \circ a) \circ a' = (c \circ a) \circ a' = c \circ (a \circ a') = c \circ n = c\,.$$

(i) Das ist klar, da ja jedes $a \in G$ genau ein Inverses a' besitzt und dann gilt:

$$\forall a, b \in G\ \exists!\, x \in G\ :\ x = a' \circ b$$

Für dieses x gilt dann

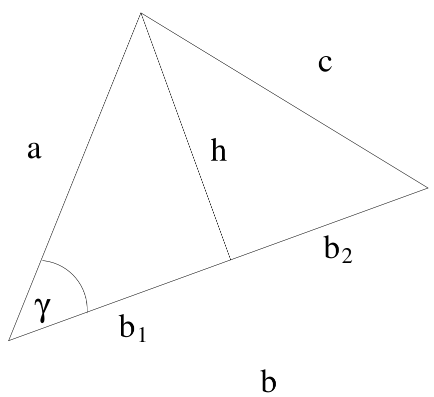
$$a \circ x = a \circ (a' \circ b) = (a \circ a') \circ b = n \circ b = b\,.$$

$\square$

Beweis Satz 3.23:

Es gelten die Beziehungen

$$\begin{aligned} & c^2 = h^2 + b_2^2 \\ \text{und} \quad & a^2 = h^2 + b_1^2 \\ & \quad = h^2 + (a \cos\gamma)^2 \end{aligned}$$



Dann ist

$$\begin{aligned} & c^2 = a^2 + b^2 - 2\,a\,b\,\cos\gamma \\ \Leftrightarrow\quad & h^2 + b_2^2 = h^2 + a^2\cos^2\gamma + b^2 - 2\,a\,b\,\cos\gamma \\ \Leftrightarrow\quad & b_2^2 = (b - \underbrace{a\,\cos\gamma}_{=b_1})^2 \end{aligned}$$

$\square$

Beweis Satz ??:

Das kann man leicht einsehen, denn es gilt wegen der Definition des Kosinus

$$\|P-a\| = \|X-a\| \cos\varphi$$

mit der Definition des Winkels gilt weiter

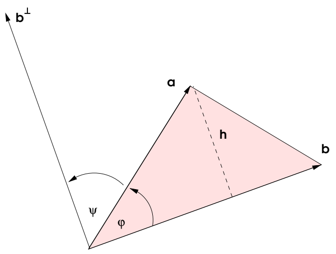
$$\begin{aligned}
&= \frac{< X-a, P-a >}{\|P-a\|} = \left\langle X-a, \frac{P-a}{\|P-a\|} \right\rangle \\
&= \left\langle X-a, \frac{v}{\|v\|} \right\rangle = \frac{\langle X-a, v\rangle}{\|v\|}
\end{aligned}$$

Und damit gilt insgesamt

$$P = a + \frac{\langle X-a, v\rangle}{\|v\|}\frac{v}{\|v\|} = a + \frac{\langle X-a, v\rangle}{\|v\|^2}\, v$$

$\square$

Beweis Satz 3.28:



$$\begin{aligned}
a \times b = a_1 b_2 - b_1 a_2 &= \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \cdot \begin{pmatrix} -b_2 \\ b_1 \end{pmatrix} =< a, b^\perp > \\
&= \|a\|\|b\| \cos\psi\,, \quad \psi = \angle(a, b^\perp) \\
&= \|a\|\|b\| \sin\left(\frac{\pi}{2} - \psi\right) = \|a\|\|b\| \sin\varphi\,, \quad \varphi = \angle(a,b) \\
&= h\,\|b\| \quad \text{also Höhe mal Grundseite des Dreiecks.}
\end{aligned}$$

$\square$

Beweis Satz 4.2:

(9)

$$\begin{aligned}
\sin x \cos y + \cos x \sin y &= \begin{pmatrix} \sin x \\ \cos x \end{pmatrix} \begin{pmatrix} \cos y \\ \sin y \end{pmatrix} \\
&=< P', Q > \qquad \text{mit} \qquad P' := \begin{pmatrix} \sin x \\ \cos x \end{pmatrix},\ Q := \begin{pmatrix} \cos y \\ \sin y \end{pmatrix} \\
&= \underbrace{\|P'\|}_{=1}\underbrace{\|Q\|}_{=1} \cos\psi \qquad \text{mit} \qquad \psi = \angle(P', Q) \\
&= \cos\psi
\end{aligned}$$

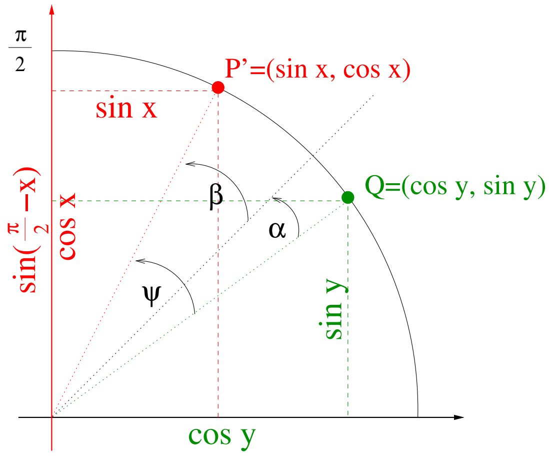
Wir wissen, dass

$$\beta + x = \frac{\pi}{4}$$
$$\alpha + y = \frac{\pi}{4}$$

gilt. Daraus folgt dann

$$\psi = \alpha + \beta = \frac{\pi}{2} - (x+y)\,.$$

Mit



$$\cos\psi = \cos(\alpha+\beta) = \cos\left(\frac{\pi}{2} - (x+y)\right) = \sin(x+y)$$

folgt dann die Behauptung!

(10) *Wir setzen in* $(AT1)$ *statt* y *einfach* $-y$ *und erhalten mit*

$$sin(x + (-y)) = \sin x \cos(-y) + \cos x \sin(-y) = \sin x \cos y - \cos x \sin y$$

die Behauptung.

(11) *Wir verschieben, so dass wir den Kosinus in Sinus ausdrücken können, verwenden dann* $(AT1)$ *und schieben dann das ganze wieder zurück:*

$$\begin{aligned}\cos(x+y) &= \sin\left(\left(\frac{\pi}{2} - x\right) + (-y)\right)\\ &= \sin\left(\frac{\pi}{2} - x\right)\cos y - \cos\left(\frac{\pi}{2} - x\right)\sin y\\ &= \cos x \cos y - \sin x \sin y\end{aligned}$$

(12) *Hier verwenden wir den gleichen Trick wie bei* $(AT1)$*: Wir setzen* $-y$ *statt* y *in die Gleichung* $(AT3)$*. Dann gilt*

$$\cos(x + (-y)) = \cos x \cos y + \sin x \sin y\,.$$

□

Beweis Satz 4.3:

Unabhängig von der Art der Darstellung ergibt sich aus der Definition des Betrages

$$|z \cdot w| = \sqrt{z\,w\,\overline{z}\,\overline{w}} = \sqrt{z\,\overline{z}}\sqrt{w\,\overline{w}} = |z||w|\,.$$

Es seien $z = a + i\,b$ *und* $w = c + i\,d$ *gegeben mit der Darstellung in Polarform:*

$$\begin{aligned} z &= r\,(\cos\varphi + i\,\sin\varphi) \\ w &= s\,(\cos\psi + i\,\sin\psi) \end{aligned}$$

Das Produkt in Polarform ergibt

$$\begin{aligned} z \cdot w &= r\,(\cos\varphi + i\,\sin\varphi) \cdot s\,(\cos\psi + i\,\sin\psi) \\ &= r\,s\,(\cos\varphi\,\cos\psi - \sin\varphi\sin\psi + i\,(\cos\varphi\sin\psi + \cos\psi\sin\varphi)) \\ &= r\,s\,(\cos(\varphi + \psi) + i\,\sin(\varphi + \psi)) \\ &=: t\,(\cos\varrho + i\,\sin\varrho) \\ &=: v \end{aligned}$$

□

Beweis Satz 4.4:

$$\begin{aligned} \frac{z}{w} &= \frac{r\,(\cos\varphi + i\,\sin\varphi)}{s\,(\cos\psi + i\,\sin\psi)} \\ &= \frac{r}{s}\frac{(\cos\varphi + i\,\sin\varphi)\,(\cos\psi - i\,\sin\psi)}{(\cos\psi + i\,\sin\psi)\,(\cos\psi - i\,\sin\psi)} \\ &= \frac{r}{s}\,(\cos\varphi + i\,\sin\varphi)\,(\cos\psi - i\,\sin\psi) \\ &= \frac{r}{s}\,(\cos\varphi\cos\psi + \sin\varphi\sin\psi + i(\sin\varphi\cos\psi - \cos\varphi\sin\psi)) \\ &= \frac{r}{s}(\cos(\varphi - \psi) + i\sin(\varphi - \psi)) \end{aligned}$$

□

Beweis Satz 3.2:

$$\begin{aligned} (A \cdot B)^T_{ki} &= \sum_{j=1}^{n} A_{ij}B_{jk} = \sum_{j=1}^{n} B_{jk}A_{ij} \\ &= \sum_{j=1}^{n} B^T_{kj}A^T_{ji} = (B^T \cdot A^T)_{ki} \end{aligned}$$

$\square$

Beweis Hilfssatz 3.17:

$$
\begin{aligned}
& (A \cdot B)^{-1}(A \cdot B) = E \\
\Leftrightarrow \quad & (A \cdot B)^{-1} A = B^{-1} \\
\Leftrightarrow \quad & (A \cdot B)^{-1} = B^{-1} A^{-1} \qquad \checkmark
\end{aligned}
$$

$\square$

Beweis Satz 3.16:

$$
\begin{aligned}
& a = B_{\mathcal{V}} a_{\mathcal{V}} \;\wedge\; a = B_{\mathcal{W}} a_{\mathcal{W}} \\
\Rightarrow \quad & B_{\mathcal{V}} a_{\mathcal{V}} = B_{\mathcal{W}} a_{\mathcal{W}} \\
\Leftrightarrow \quad & a_{\mathcal{V}} = B_{\mathcal{V}}^{-1} B_{\mathcal{W}} a_{\mathcal{W}}
\end{aligned}
$$

Mit dem Hilfssatz 3.17 gilt dann auch

$$
\begin{aligned}
& a_{\mathcal{V}} = B_{\mathcal{V}}^{-1} B_{\mathcal{W}} a_{\mathcal{W}} \\
\Leftrightarrow \quad & \left(B_{\mathcal{V}}^{-1} B_{\mathcal{W}}\right)^{-1} a_{\mathcal{V}} = \left(B_{\mathcal{V}}^{-1} B_{\mathcal{W}}\right)^{-1} B_{\mathcal{V}}^{-1} B_{\mathcal{W}} a_{\mathcal{W}} \\
\Leftrightarrow \quad & B_{\mathcal{W}}^{-1} B_{\mathcal{V}} a_{\mathcal{V}} = a_{\mathcal{W}}
\end{aligned}
$$

$\square$

$\square$

Beweis Satz ??:

1. $P^T = P^{-1}$
2. $||Px||_2 = ||x||_2$ *und* $< Px, Py >=< x, y >$.
3. $P \cdot Q$ *ist orthogonal:* $(P \cdot Q)^T = Q^T \cdot P^T = Q^{-1} \cdot P^{-1} = (P \cdot Q)^{-1}$.

$\square$

Beweis Satz 3.8:

Es sei V *ein VR und* $a_1, \ldots, a_n \in V$, *die alle nicht gleich Null sind.*

$\Leftarrow$: Es lasse sich ein, Œ: a_1, als Linearkombination der anderen Vektoren a_i, $i = 2, \dots, n$ darstellen:

$$\begin{aligned} & a_1 = \alpha_2\, a_2 + \cdots + \alpha_n\, a_n \\ \Leftrightarrow \quad & 0 = -a_1 + \alpha_2\, a_2 + \cdots + \alpha_n\, a_n \\ & \quad = \alpha_1\, a_1 + \alpha_2\, a_2 + \cdots + \alpha_n\, a_n\,, \quad \alpha_1 = -1 \end{aligned} \tag{16}$$

Damit ist die lineare Unabhängigkeit gezeigt.

$\Rightarrow$

Die erste und die letzte Zeile in den Gleichungen aus (16) sind äquivalent. Daher gilt die Rückrichtung bereits.

$\square$

Beweis Satz 3.15:

$$\begin{aligned} & a = \sum_{i=1}^{n} \lambda_i\, v_i \\ \Leftrightarrow \quad & \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \lambda_1 \begin{pmatrix} v_1^1 \\ \vdots \\ v_1^n \end{pmatrix} + \cdots \lambda_n \begin{pmatrix} v_n^1 \\ \vdots \\ v_n^n \end{pmatrix} = \underbrace{\begin{pmatrix} v_1^1 & \cdots & v_n^1 \\ \vdots & & \vdots \\ v_1^n & \cdots & v_n^n \end{pmatrix}}_{=:A} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} \end{aligned}$$

Da die Spalten in der Matrix A Basisvektoren und somit linear unabhängig sind, ist die Matrix regulär und somit ist das LGS in Form $A\,\lambda = a$ eindeutig lösbar. Bei gegebener Basis gibt es also für jedes $a \in V$ eindeutig bestimmte Koordinaten λ, was ja genau die Behauptung war.

$\square$

Beweis Satz 3.20:

Es sei (V, s) ein Prähilbertraum und die Abbildung $\|\cdot\|_s : V \to \mathbb{R}$ definiert durch $\|\cdot\|_s = \sqrt{s(\cdot,\cdot)}$. Zu zeigen ist, dass für die Abbildung $\|\cdot\|_s$ die Normaxiome $(N1)$ bid $(N3)$ aus Definition 3.18 erfüllt sind.

(N1) $\|\alpha\, v\|_s = |\alpha|\, \|v\|_s$:

$$\|\alpha\, v\|_s = \sqrt{s(\alpha\, v, \alpha\, v)} = \sqrt{\alpha^2\, s(v,v)} = |\alpha|\, \sqrt{s(v,v)} = |\alpha|\, \|v\|_s$$

(N2) $\|v\|_s >= 0$ *und* $\|v\|_s = 0 \quad \Leftrightarrow \quad v = 0$ *folgt direkt aus der positiv Definitheit des Skalarprodukts (siehe (S3) und (S4) in Def. 3.19).*

(N3) $\|v+w\|_s \leq \|v\|_s + \|w\|_s$:

$$\|v+w\|_s^2 = s(v+w, v+w) \leq s(v,v) + 2|s(v,w)| + s(w,w)$$
$$\overset{\text{Satz 3.21}}{\leq} \|v\|_s^2 + 2\,\|v\|_s\|w\|_s + \|w\|_s^2 = (\|v\|_s + \|w\|_s)^2$$

$\square$

Beweis Satz 3.21:

Es sei (V,s) *ein Prähilbertraum mit induzierter Norm* $\|\cdot\|_V$*. Dann gilt* $\forall x,y \in V \setminus \{0\}$ *und* $\forall \lambda \in \mathbb{K}$:

$$\begin{aligned} 0 \leq \|x - \lambda\, y\|_V^2 &= s(x - \lambda\, y, x - \lambda\, y) \\ &= s(x,x) - 2\,\lambda\, s(x,y) + \lambda^2\, s(y,y) \end{aligned}$$

Wir wählen $\lambda = \frac{s(x,y)}{s(y,y)}$, λ *ist ja beliebig, und dann erhalten wir*

$$\begin{aligned} &= s(x,x) - 2\,\frac{s(x,y)}{s(y,y)}\, s(x,y) + \frac{s(x,y)^2}{s(y,y)^2}\, s(y,y) \\ &= s(x,x) - 2\,\frac{s(x,y)^2}{s(y,y)} + \frac{s(x,y)^2}{s(y,y)} \\ &= s(x,x) - \frac{s(x,y)^2}{s(y,y)} \end{aligned}$$

Da $s(y,y) > 0$ *ist dieser Ausdruck äquivalent zu*

$$s(x,y)^2 \leq \|x\|_V^2 \|y\|_V^2$$

$\square$

Beweis Satz 3.33:

Wir zeigen, dass für jedes $v \in V$

$$v - P(v) \perp b_j\,, \quad \forall b_j \in \mathcal{B}\,.$$

Für jedes beliebige $b_j \in \mathcal{B}$ *gilt*

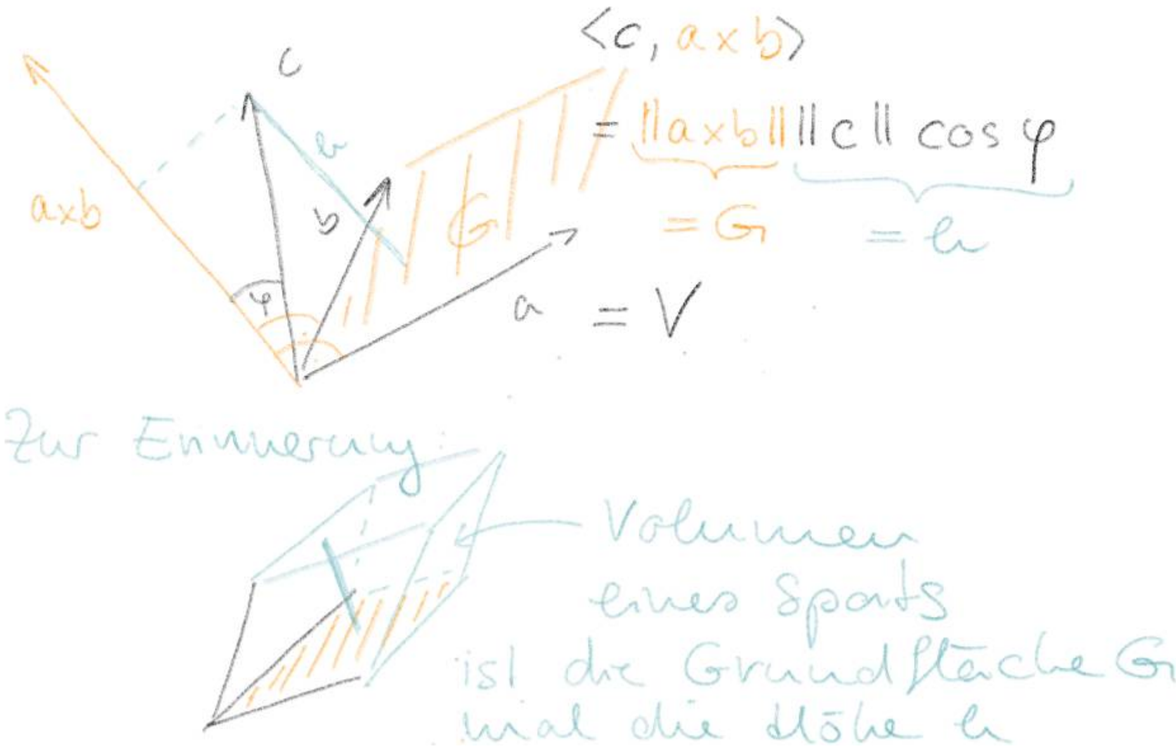
$$s(v - P(v), b_j) = s(v,b_j) - s(P(v), b_j) = s(v,b_j) - \sum_{k=1}^{n} \frac{s(v,b_k)}{\|b_k\|_s^2} \underbrace{s(b_k, b_j)}_{\delta_{kj}\,\|b_j\|_s^2}$$
$$= s(v,b_j) - \frac{s(v,b_j)}{\cancel{\|b_j\|_s^2}} \cancel{\|b_j\|_s^2} = 0$$

Damit ist auch klar, dass $s(v - P(v), w) = 0 \ \forall w \in U$*, denn mit* $w = \sum_{j=1}^{n} \beta_j \, b_j$ *folgt*

$$s(v - P(v), w) = s(v - P(v), \sum_{j=1}^{n} \beta_j \, b_j) = \sum_{j=1}^{n} \beta_j \underbrace{s(v - P(v), b_j)}_{=0} = 0$$

$\square$

Beweis Satz 3.47:



$\square$

Beweis Satz 3.37:

$$\mathcal{A}(0) = \mathcal{A}(x - x) = \mathcal{A}(x) - \mathcal{A}(x) = 0$$

$\square$

Beweis Satz 3.39:

Da die v_1, v_2 *lu sind folgt aus*

$$\lambda_1 v_1 + \lambda_2 v_2 = 0$$

direkt

$$\lambda_1 = \lambda_2 = 0\,.$$

Für w_1, w_2 gilt nun

$$0 = \lambda_1 w_1 + \lambda_2 w_2 = \lambda_1 \mathcal{A}(v_1) + \lambda_2 \mathcal{A}(v_2) = \mathcal{A}(\lambda_1 v_1 + \lambda_2 v_2)$$

Damit sind auch w_1 und w_2 lu.

□

Beweis Satz 3.45:

(i) *Zu zeigen ist hier, dass für $\alpha, \beta \in \mathbb{K}$ und $a, b \in \operatorname{Kern} f$ immer $\alpha\, a + \beta\, b \in \operatorname{Kern} f$ gilt. Das erhalten wir aber gerade über die Linearität von f, denn*

$$f(\alpha\, a + \beta\, b) = \alpha\, f(a) + \beta\, f(b) = \alpha\, 0 + \beta\, 0 = 0\,.$$

(ii) $\operatorname{Kern} f = \{0\} \quad \Leftrightarrow \quad f$ *ist injektiv*

"$\Leftarrow$" *Angenommen es gebe ein $a \in \operatorname{Kern} f$ mit $a \neq 0$, dann folgte daraus ein Widerspruch zur Injektivität von f, denn es wäre dann*

$$f(a) = 0 = f(0)\,.$$

Es würden also zwei verschiedene Vektoren auf ein und denselben, nämlich 0 abgebildet werden.

"$\Rightarrow$" *Die Aussage*

$$\text{“}(\operatorname{Kern} f = \{0\}) \quad \Rightarrow \quad (f \text{ ist injektiv})\text{”}$$

ist äquivalent zur Aussage

$$\text{“} \neg\ (f \text{ ist injektiv}) \Rightarrow \neg\ (\operatorname{Kern} f = \{0\})\text{”}$$

bzw

$$\text{“ } (f \text{ ist nicht injektiv}) \Rightarrow (\exists a \in V\ a \neq 0\ :\ f(a) = 0)\text{”}$$

Die neue Voraussetzung ist, dass f nicht injektiv ist. Es existieren also $a, b \in V$ mit $a \neq b$ und $f(a) = f(b)$.

$$\begin{aligned} &\Rightarrow & f(a) - f(b) &= 0 \\ &\Rightarrow & f(a-b) &= 0 \\ &\Rightarrow & \underbrace{a-b}_{\neq 0} &\in \operatorname{Kern} f \end{aligned}$$

Damit besteht dann auch der Kern aus mehr als einem Vektor.

(iii) Es seien $\alpha, \beta \in \mathbb{K}$ und $a, b \in U$. Da U ein UVR *von V ist, gilt $\alpha\, a + \beta\, b \in U$. Dann ist $f(a), f(b), f(\alpha\, a + \beta\, b \in) \in W$ und wegen der Linearität von f gilt auch $f(\alpha\, a + \beta\, b \in) = \alpha\, f(a) + \beta\, f(b)$ und damit ist dann automatisch*

$$\alpha\, f(a) + \beta\, f(b) \in W \;\forall f(a), f(b) \in W\,.$$

□

Beweis Satz 3.54:

Es gebe OE[16] zwei EWe *λ_1 und λ_2 und zugehörige* EVen *v_1 und v_2. Wären diese linear abhängig, gelte also $v_1 = \alpha\, v_2$ dann wäre v_2 ebenfalls ein* EV *zum* EW *λ_1, denn es gelte*

$$A\, v_1 = \lambda_1\, v_1 \quad \Leftrightarrow \quad A\, \alpha\, v_2 = \lambda_1\, \alpha\, v_2 \quad \Leftrightarrow \quad A\, v_2 = \lambda_1\, v_2\,.$$

Dieses Grundprinzip lässt sich sukzessive auf höhere Raumdimensionen übertragen.

□

Beweis Satz C.2:

Für $v_i \geq 0$ ist $\|v\|_1 = \sum_{i=1}^{n} v_i$. Dann gilt für eine spaltenstochastische Matrix A

$$\|A\, v\|_1 = \sum_i \sum_j a_{ij} v_j = \sum_j \underbrace{\sum_i a_{ij}}_{=1} v_j = \|v\|_1\,.$$

□

Beweis Satz 4.5:

Die Eulergleichungen erhält man über die Potenzreihendarstellungen von sin, cos *und* e*:*

$$\sin x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2\,k+1)!} x^{2\,k+1}$$
$$\cos x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2\,k)!} x^{2\,k}$$

[16] Ohne Einschränkung der Allgemeinheit

$$\begin{aligned}
\cos x + i\,\sin x &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2\,k)!} x^{2\,k} + i\,\frac{(-1)^k}{(2\,k+1)!} x^{2\,k+1} \\
&= \sum_{k=0}^{\infty} \frac{(i^2)^k}{(2\,k)!} x^{2\,k} + i\,\frac{(i^2)^k}{(2\,k+1)!} x^{2\,k+1} \\
&= \sum_{k=0}^{\infty} \frac{(i\,x)^{2\,k}}{(2\,k)!} + \frac{(i\,x)^{2\,k+1}}{(2\,k+1)!} = \sum_{k=0}^{\infty} \frac{(i\,x)^k}{k!} = \mathrm{e}^{i\,x}
\end{aligned}$$

□

Beweis Satz 4.8:

Sei $a \in \mathbb{C}$ mit $a = \alpha\,\mathrm{e}^{i\,\psi}$, so muss für die Polardarstellung $z = r\,\mathrm{e}^{i\,\varphi}$ jeder n–ten Wurzel z von a

$$r^n = \alpha \quad \textit{und} \quad n\,\varphi = \psi + 2\,\pi\,k$$

mit einem $k \in \mathbb{Z}$ gelten. Also ist

$$r = \sqrt[n]{\alpha} \quad \textit{und} \quad \varphi = \frac{\psi}{n} + \frac{2\,\pi\,k}{n}\,.$$

Jede Lösung der Gleichung $z^n = a$ hat daher die Form

$$z_k = \sqrt[n]{\alpha}\,\mathrm{e}^{i\left(\frac{\psi}{n}+\frac{k}{n}2\,\pi\right)}$$

mit $\alpha = |a|$ und $k \in \mathbb{Z}$.

Unter diesen unendlich vielen Zahlen sind nur $z_0, z_1, \ldots, z_{n-1}$ voneinander verschieden, denn es ist

$$\begin{aligned}
z_n &= \sqrt[n]{\alpha}\,\mathrm{e}^{i\left(\frac{\psi}{n}+\frac{n}{n}2\,\pi\right)} = \sqrt[n]{\alpha}\,\mathrm{e}^{i\left(\frac{\psi}{n}+2\,\pi\right)} = \sqrt[n]{\alpha}\,\mathrm{e}^{i\left(\frac{\psi}{n}\right)} && = z_0 \\
z_{n+1} &= \sqrt[n]{\alpha}\,\mathrm{e}^{i\left(\frac{\psi}{n}+\frac{n+1}{n}2\,\pi\right)} = \sqrt[n]{\alpha}\,\mathrm{e}^{i\left(\frac{\psi}{n}+\frac{n}{n}2\,\pi+\frac{1}{n}2\,\pi\right)} = \sqrt[n]{\alpha}\,\mathrm{e}^{i\left(\frac{\psi}{n}+\frac{1}{n}2\,\pi\right)} = \sqrt[n]{\alpha}\,\mathrm{e}^{i\left(\frac{\psi}{n}\right)} && = z_1 \\
z_{n+2} &= \sqrt[n]{\alpha}\,\mathrm{e}^{i\left(\frac{\psi}{n}+\frac{n+2}{n}2\,\pi\right)} = \sqrt[n]{\alpha}\,\mathrm{e}^{i\left(\frac{\psi}{n}+\frac{n}{n}2\,\pi+\frac{2}{n}2\,\pi\right)} = \sqrt[n]{\alpha}\,\mathrm{e}^{i\left(\frac{\psi}{n}+\frac{2}{n}2\,\pi\right)} = \sqrt[n]{\alpha}\,\mathrm{e}^{i\left(\frac{\psi}{n}\right)} && = z_2 \\
&\vdots
\end{aligned}$$

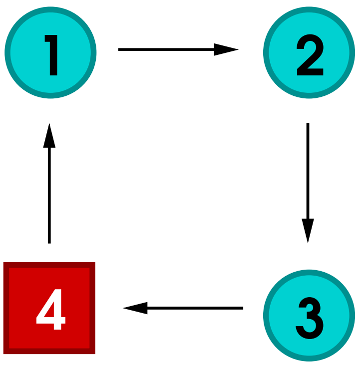
□

C Anwendungsbeispiele

C.1 Anwendungsbeispiele zu Eigenwerten und Eigenvektoren

Beispiel 114 MiniPoly

Wir entwickeln ein Miniaturmonopoly und berechnen, ob der Spielverlauf interessant oder langweilig sein wird.



Spielregeln:

- Gestartet wird auf einem beliebigen Feld.
- Auf 1-3 wird 1 Mal gewürfelt.
- Auf 4 kommt man mit einer 6 auf die 1, sonst bleibt man sitzen.

Frage: Mit welcher Wahrscheinlichkeit steht die Spielfigur auf den jeweiligen Feldern?

Wir erinnern uns an ein paar Grundelemente aus der Wahrscheinlichkeitsrechnung: Die Wahrscheinlichkeit $P(A)$ eines Ereignisses A ergibt sich aus

$$P(A) = \frac{\text{Anzahl günstiger Fälle}}{\text{Anzahl möglicher Fälle}}\,.$$

Die Zahl 4 würfeln hat demnach die Wahrscheinlichkeit $P(4) = \frac{1}{6}$. Es gilt generell:

$$\begin{array}{ll} P(A) = 0\text{:} & A \text{ trifft nie zu} \\ P(A) = 1\text{:} & A \text{ trifft sicher zu} \\ P(A) \in (0,1)\text{:} & \text{sonst} \end{array}$$

und

$$P(A \vee B) = P(A) + P(B)\text{:} \quad A \text{ oder } B \text{ trifft zu}$$

Damit ergibt sich die Wahrscheinlichkeit 3 oder 5 zu würfeln aus $P(3 \vee 5) = P(3) + P(5) = \frac{1}{3}$. Mehr brauchen wir nicht. Zurück zum Spiel:

Wir überlegen uns die Wahrscheinlichkeiten, um vom Feld j auf das Feld i zu gelangen und tragen die Ergenbisse in ij-te Komponente der Matrix A ein:

$$A = \begin{pmatrix} \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{6} & 0 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{6} & 0 \\ \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{5}{6} \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 1 & 1 & 2 & 1 \\ 2 & 1 & 1 & 0 \\ 2 & 2 & 1 & 0 \\ 1 & 2 & 2 & 5 \end{pmatrix}$$

Wir starten nun einmal mit einer Figur auf Feld 1. Es beinhalte der Vektor $x_0 \in \mathbb{R}^4$ die Wahrscheinlichkeiten, auf einer bestimmten Position zu stehen, dann ist

$$x_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

. Nach einem Wurf ergeben sich die neuen Wahrscheinlichkeiten zu

$$x_1 = A\, x_0 = \frac{1}{6} \begin{pmatrix} 1 & 1 & 2 & 1 \\ 2 & 1 & 1 & 0 \\ 2 & 2 & 1 & 0 \\ 1 & 2 & 2 & 5 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} 1 \\ 2 \\ 2 \\ 1 \end{pmatrix}$$

Nach zwei Würfen:

$$x_2 = A\, x_0 = \frac{1}{6} \begin{pmatrix} 1 & 1 & 2 & 1 \\ 2 & 1 & 1 & 0 \\ 2 & 2 & 1 & 0 \\ 1 & 2 & 2 & 5 \end{pmatrix} \frac{1}{6} \begin{pmatrix} 1 \\ 2 \\ 2 \\ 1 \end{pmatrix} = \frac{1}{36} \begin{pmatrix} 8 \\ 6 \\ 8 \\ 14 \end{pmatrix}$$

Daraus ergibt sich eine Folge x_3, x_4, ...

$$\begin{pmatrix} 0.2037 \\ 0.1389 \\ 0.1667 \\ 0.4907 \end{pmatrix}, \begin{pmatrix} 0.1944 \\ 0.1188 \\ 0.1420 \\ 0.5448 \end{pmatrix}, \ldots$$

Die entscheidende Frage ist jetzt, wie sich diese Wahrscheinlichkeiten im Laufe eines längeren Spieleabends entwickeln werden und ob das Ergebnis von der Startposition abhängig ist.

$$\begin{aligned} x_0 &= \text{Startwert} \\ x_1 &= A\, x_0 \\ &\vdots \\ x_k &= A\, x_{k-1} \end{aligned}$$

Bei Konvergenz $\lim_{k\to\infty} x_k = x$ gilt für diesen Grenzwert

$$x = A\, x$$

Wenn die Matrix A einen EW 1 hat dann ist der Grenzwert dieser Folge gerade der EV zu diesem EW . Nicht wahr?

Das Matlab-Programm

```
A=[[1 1 2 1];[2 1 1 0];[2 2 1 0];[1 2 2 5]]/6;
[V,D]=eig(A);
[val i]=min(abs(diag(D)-1));
lambda = D(i,i);
Ev = V(:,i)/sum(V(:,i));
fprintf('Der EV zum EW lambda=%f lautet\nv=[%f %f %f %f]\n',lambda,Ev);
```

liefert das Ergebnis

```
Der EV zum EW lambda=1.000000 lautet
v=[0.185484 0.096774 0.112903 0.604839]
```

Das Ergebnis zeigt, dass man am häufigsten im Gefängnis sitzt, am wenigsten häufig auf Feld 2, dann relativ ausgeglichen auf 1 und 3. Könnte also ein eher langweiliger Spieleabend werden..... :-(

Definition 3.1 (spaltenstochastische Matrix) *Eine Matrix* $A \in \mathbb{R}^{n\times n}$ *heißt spaltenstochastisch, falls gilt*

$$\sum_{i=1}^{n} a_{ij} = 1 \quad \forall j$$

Satz 3.2 *Für eine spaltenstochastische Matrix* A *und den* EV v, *beide mit nichtnegativen Komponenten, gilt*

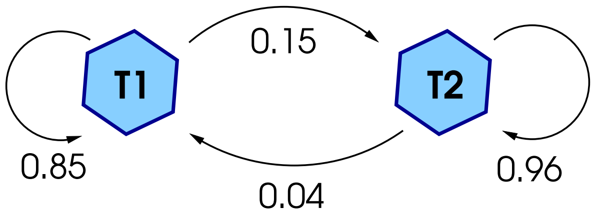
$$\|A\,v\|_1 = \|v\|_1\,.$$

Beweis von Satz C.2 auf Seite 163

Satz C.2 besagt, dass EVen zum EW 1 nur positive oder negative Einträge haben können.

Beispiel 115 Versicherungspolice

Eine Kfz-Versicherung wird zu zwei Tarifgruppen angeboten. Der Anbieter beobachtet, dass im Laufe eines Jahres 15% der Versicherungsnehmer von Tarifgruppe 1 zu Tarifgruppe 2 wechseln. Bei der Tarifgruppe 2 wechseln 4% zur Tarifgruppe 1. Ein Übergangsdiagramm stellt diese Zusammenhänge graphisch dar:



Die jährliche Änderung der Tarifgruppengröße wird durch die Matrix

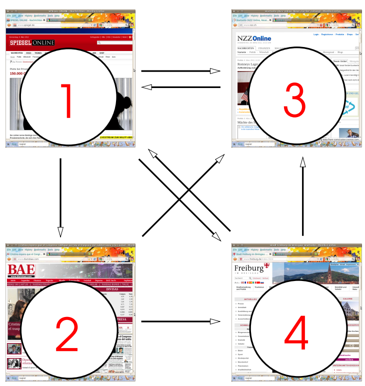
$$A = \begin{pmatrix} 0.85 & 0.04 \\ 0.15 & 0.96 \end{pmatrix}$$

dargestellt. Sie sehen, dass es sich um eine spaltenstochastische Matrix handelt. Ist auch klar, wenn man mal kurz drarüber nachdenkt.

Die Versicherung ist nun daran interessiert, zu wissen, wohin sich die Anzahl der Versicherten der jeweiligen Tarifgruppen im Laufe der Zeit entwickelt. Die entsprechende Verteilung der Versicherungsnehmer können wir dem EV zum EW entnehmen:

$$v = \frac{\text{Anzahl Gesamtpersonen}}{19} \begin{pmatrix} 4 \\ 15 \end{pmatrix}$$

Beispiel 116 Googlematrix



Sie geben einen Suchbegriff ein, drücken die Eingabetaste und mit kindlicher Vorfreude studieren Sie die ersten Treffer aus einer Liste von Tausenden. Schier unermässlich scheint der Wille des Internets Ihnen Ihre Frage beantworten zu wollen. Schier unmöglich ist es hingegen, alle angebotenen Seiten nach der besten Information zu durchforsten. Die Suchmaschine hat bereits eine Vorsortierung für Sie vorgenommen. Nun weiss kein Computer, welche Seite für Ihre Bedürfnisse die beste Wahl ist, aber nach ganz speziellen Kriterien, und wir wollen hier von den erkauften Plätzen in vorderster Reihe absehen, wurden bestimmte Seiten, die zu Ihrer Trefferliste gehören, als "wichtig" erklärt und nicht nur das sondern auch bezüglich einer Wichtigkeitsfunktion angeordnet. Wichtigste Seite zuerst, zweitwichtigste als zweite und so weiter.

Und was macht eine Seite nun wichtig? Wir könnten annehmen, dass eine Seite dann wichtig ist, wenn viele andere Seiten auf sie verweisen. Gehen wir einmal von einem äußerst einfachen Fall aus, nämlich, dass das Internet genau aus vier Seiten besteht und alle vier Seiten auch Ihren Suchbegriff enthalten. Diese vier Seiten seien untereinander verlinkt, so wie es im Bild links dargestellt ist. In dieser Situation wäre dann die Seite mit der Nummer 3 die wichtigste. Auf sie zeigen drei links. Die Seiten 1 und 4 kämen auf Position zwei mit jweils zwei Zuweisungen und die Seite 2 wäre dann die unwichtigste Seite. Die Ergebnisliste sähe also so aus:

$$\begin{bmatrix} 3 \\ 1,4 \\ 2 \end{bmatrix}$$

Wie auch immer mit der Rangfolge zwischen Seite 1 und 4 verfahren würde. Wir wollen folgende Überlegung anstellen: Es sollten nicht nur quantitative sondern auch qualitative Kriterien einfließen. Das heißt, dass es nicht nur auf die Anzahl der Zuweisungen ankommt sondern zusätzlich noch darauf ob diese Zuweisungen von wichtigen oder weniger wichtigen Seiten kommen. Was nutzen zehn Verweise von völlig nichtssagenden Seiten? Darauf sollte unmöglich die Wertschätzung einer Internetseite beruhen.

Die Idee ist folgende: Jede Seite erhält eine Stimme, die sie auf die Seiten auf die sie verweist gleichermaßen verteilt. Dazu sei W die Menge der Webseiten, die wir mit w_i, $i = 1, \ldots, 4$ bezeichnen und

$$k_i := \text{Anzahl der Links, die von der Seite } w_i \text{ ausgehen}$$

In unserem Fall also

$$k_1 = 3\,,\ k_2 = 2\,,\ k_3 = 1\,,\ k_4 = 2\,.$$

Jede Seite w_i verteilt seine Stimme gleichmäßig zu Teilen $\frac{1}{k_i}$ an die Seiten, auf die sie verweist. Die Stimme ist um so gewichtiger, je wichtiger die Seite w_i ist. Dazu definieren wir die sogenannte Wichtigkeitsfunktion (PageRank$^{\text{TM}}$):

$$\begin{aligned} r:\ W &\to \mathbb{R}_0^+ \\ w_i &\mapsto r_i \end{aligned}$$

Jeder Stimmanteil der Seite w_i wird nun mit seiner Wichtigkeit gewichtet und an die entsprechend verlinkte Seite abgegeben. Unsere Seite 1 erhält dann die Wichtigkeit

$$r_1 = \frac{r_3}{k_3} + \frac{r_4}{k_4} = \frac{r_3}{1} + \frac{r_4}{2}$$

Seite 1 erhält von Seite 2 gar keine Stimme. Es gibt keinen Verweis von Seite 2 auf sie, so dass man annhemen kann, dass Seite 2 sie für absolut irrelevant hält. Von Seite 3 bekommt sie den vollen Stimmanteil. Seite 3 hält nur sie für erwähnenswert also besonders wichtig. Das ergibt einen hohen Stimmanteil, wenn Seite 3 selbst sehr wichtig ist, das heißt r_3 einen großen Wert enthält. Ansonsten eben nicht. Seite 4 hält sowohl Seite 1 als auch Seite 3 für erwähnenswert und verteilt ihre Stimmen damit zu gleichen Teilen. Auch hier fließt die Wichtigkeit der Seite 4 gewinnbringend für Seite 1 ein. Nach diesem Schema verfahren wir mit allen Seiten und erhalten insgesamt

$$\begin{aligned} r_1 &= \frac{r_3}{1} + \frac{r_4}{2} \\ r_2 &= \frac{r_1}{3} \\ r_3 &= \frac{r_1}{3} + \frac{r_2}{2} + \frac{r_4}{2} \\ r_4 &= \frac{r_1}{3} + \frac{r_2}{2} \end{aligned}$$

Matrix-Vektor Schreibweise

$$\begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & 0 \end{pmatrix}}_{=:A} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{pmatrix}$$

Wir suchenalso ein $r \in \mathbb{R}^n$ mit

$$r = A\,r\,,$$

das heißt wir suchen den EV zum EW 1, sofern es diesen gibt. Wir sehen direkt, dass es sich um eine spaltenstochastische Matrix handelt mit positiven Einträgen. Den EW 1 haben wir demnach sicher.

Damit ist klar, dass wir Vektoren r aus dem Kern der Abbildung $(A - E)x$ suchen. Für welche r gilt also

$$\begin{pmatrix} -1 & 0 & 1 & \frac{1}{2} \\ \frac{1}{3} & -1 & 0 & 0 \\ \frac{1}{3} & \frac{1}{2} & -1 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{2} & 0 & -1 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{pmatrix} = 0\,?$$

Elementare Zeilenumformungen führen auf

$$\begin{pmatrix} -1 & 0 & 1 & \frac{1}{2} \\ 0 & -1 & \frac{1}{3} & \frac{1}{6} \\ 0 & 0 & -\frac{1}{2} & \frac{3}{4} \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ r_4 \end{pmatrix} = 0 \quad \Rightarrow \quad \mathrm{Kern}\,\mathcal{A} = \left\{ r \in \mathbb{R}^4 \;\middle|\; r = \begin{pmatrix} 12 \\ 4 \\ 9 \\ 6 \end{pmatrix} \lambda\,,\ \lambda \in \mathbb{R} \right\},$$

was uns direkt die Rangfolge der Internetseiten bezüglich ihrer Wichtigkeit liefert:

$$\begin{bmatrix} 1 \\ 3 \\ 4 \\ 2 \end{bmatrix}$$

Mehr dazu in Bryan und Leise (2006) oder auch auf http://www.rose-hulman.edu/ bryan/google.html.

Literatur

Bryan, K. und Leise, T. The \$ 25,000,000,000 Eigenvector: The Linear Algebra behind Google. *SIAM Review*, **48**(3): (2006), 569–581.

Index