JOCHEN HEINLOTH

LINEARE ALGEBRA

Eine Rohversion meiner Vorlesungsnotizen

Inhaltsverzeichnis

Vorwort

Im Wintersemester 2022/23 hatte ich einmal wieder die Gelegenheit die Anfängervorlesung zur linearen Algebra zu halten und dies sind meine Notizen, die ich als Skript parallel zur Vorlesung aufgeschrieben habe. Dies ist noch eine Rohversion mit Tippfehlern.

Die wiederkehrenden Gespräche mit Studierenden und Kolleg*innen über den Studieneinstieg und einige der Forschungsarbeiten hierzu, waren für mich ein Anlass, die Vorlesung gerade zu Beginn wieder etwas zu verändern und zudem die Übungen anders zu gestalten. Das möchte ich kurz erklären.

Für mich ist Mathematik insbesondere dazu da, komplizierte Probleme so einfach darzustellen, dass wir darüber nachdenken können.

Bei den Studierenden nehme ich zu Anfang des Studiums oft den umgekehrten Eindruck wahr, dass in der Mathematik zu Beginn einfache Dinge grundlos kompliziert gemacht werden. Ich erinnere mich aus meinem eigenen Studium zum Beispiel an die Wahrheitstabelle für die Schlussfolgerung „$\Rightarrow$“, bei der ich das merkwürdige Gefühl hatte, dass ich Sätze der Form „aus A folgt B“ eigentlich sicher verstehe, mich aber anstrengen muss, die Tabelleneinträge zu begreifen. Die Konstruktion der ganzen Zahlen mit Hilfe von Äquivalenzrelationen fand ich deutlich komplizierter, als einfach „$-$“ vor die positiven Zahlen zu schreiben, und die Aufgabe, darüber nachzudenken warum in jedem Körper die Gleichung $3 \cdot 0 = 0$ gilt, schien mir eher eine Fingerübung, denn ernsthafter Studieninhalt.

Im Nachhinein weiß ich, wieso diese Dinge aus gutem Grund in die Grundlagen der Mathematik Einzug gefunden haben, aber zu Anfang des Studiums habe ich diese selbst eher als Spielereien wahrgenommen, die leider für viele Studierende eine sehr ernstzunehmende Hürde darstellen.

Gespräche mit Studierenden bestätigen meine Sorge, dass die Auflösung, wieso das sinnvoll ist, vielfach auch bei guten Studierenden bis zum Ende des Studiums nicht eintritt.

In der Vorlesung habe ich darum versucht, die mathematischen Techniken und Begriffe in der Vorlesung nach Möglichkeit erst

dann zu entwickeln, wenn dazu ein Anlass geschaffen wurde. Ein positiver Nebeneffekt hiervon ist, dass sich die Begriffsbildungen damit etwas gleichmäßiger über das Semester verteilen.

Insbesondere habe ich versucht interessante Beispiele wo irgend möglich immer vor und nicht erst nach den zugehörigen Definitionen vorzustellen. Das ist kein neues Konzept, aber ich war überrascht, wie wenig es in Büchern für den Studieneinstieg verwendet wird.

Ich hatte etwas Sorge, dass ich mit diesem Vorgehen sehr viel Zeit mit Spaltenvektoren und dem Rechnen von Beispielen verbringen würde, ohne mit der Theorie wirklich voran zu kommen und das alles in dem Wissen, dass ich viele Aussagen dann später noch einmal in abstrakterer Sprache formulieren musste.

Am Ende war es umgekehrt so, dass die für Beispiele und Anwendungen genutzte Zeit, die allgemeinen Resultate erleichtert hat. Zum Beispiel war ich überrascht, dass mir die Studierenden in der Vorlesung in der wir nach vielen Beispielen den abstrakten Vektorraumbegriff einführen wollten, selbst eine insgesamt vollständige Liste von Axiomen vorschlagen konnten, in der nur eines – obgleich äquivalent – inhaltlich etwas von der Standardformulierung abwich.

DIE ÜBUNGSAUFGABEN habe ich im Vergleich zu den Vorjahren etwas angepasst. Mir hatten einige sehr gute Studierende in Seminaren zum Ende des Studiums berichtet, dass sie die Übungszettel zu Beginn ihres Studiums als fast ausschließlich frustrierend wahrgenommen hatten. Das hatte mich schockiert. Zusätzlich zur Übungsstruktur habe ich daher noch stärker versucht auf jedem Aufgabenblatt mehrere Aufgaben anzubieten, die leichtere Erfolgserlebnisse ermöglichen und die Studierenden explizit gebeten, mich zu warnen, wenn die Aufgaben als zu schwer wahrgenommen werden, da ich mich verschätzen kann. Das ist in Woche 6 dann leider tatsächlich passiert, aber so konnte ich immerhin versuchen, das abzumildern.

Bei der Formulierung der Aufgaben habe ich schließlich versucht, die Anregungen von Thomas Bauer und Lisa Hefendehl-Hebecker[1] wenigstens im Hinterkopf zu behalten. Leider habe ich nur in einigen wenigen Wochen Zeit für peer-instruction Modelle gefunden.

[1] Thomas Bauer and Lisa Hefendehl-Hebecker. *Mathematikstudium für das Lehramt an Gymnasien*. Springer Spektrum Wiesbaden, 2019. ISBN 978-3-658-26681-3. URL https://doi.org/10.1007/978-3-658-26682-0

DIE ÜBUNGSGRUPPEN zur Vorlesung haben wir zum niederländischem Modell umgestellt, d.h. statt einer Mischung aus Präsenz und Hausuaufgaben, sollten die Übungen vorwiegend dafür genutzt werden, an den Aufgaben zu arbeiten, die in der Woche abgegeben werden sollten. Dies, damit einerseits für die Frage „Ich weiß gar nicht wie ich anfangen soll?“ ein Ansprechpartner verfügbar ist und andererseits die Inhalte der Übungsgruppen mit denen der Vorlesung zeitlich besser übereinstimmen. Nach Abgabetermin wurden dann Lösungen für die Aufgaben bereit gestellt.

Zusätzlich zu den schriftlichen Aufgaben, gab es für die einzelnen Themenblöcke digitale Aufgaben zur Wiederholung, die von

Natascha Scheibke über die universitätseigene JACK-Umgebung in Moodle bereitgestellt wurden. Die Bearbeitung war freiwillig, zur Klausurvorbereitung wurden die Studierenden aber verpflichtet, sich ein eigenes unverbindliches Wochenziel für diese Aufgaben zu setzen.

Bei der nächsten Gelegenheit würde ich gerne noch am Ende der Kapitel jeweils eine strukturierte Übersicht über die jeweils im Kapitel enthaltenen Begriffe und Resultate hinzufügen. Zudem würde ich das Übungsformat gerne noch um einen Anlass erweitern, sich mit den Korrekturen der eigenen Lösungen zu beschäftigen.

Danksagung: An dieser Stelle möchte ich mich ganz herzliche bei Freunden, Kolleg*innen, Studierenden, Tutor*innen, kurz bei allen bedanken, die sich immer wieder Zeit genommen haben, um mit mir über Fragen, Inhalte, Formate und Stolpersteine zum Studieneinstieg zu sprechen, zu diskutieren und mir Anregungen zu geben. Vielen herzlichen Dank! Vielen Dank auch an alle Studierenden und Tutor:innen, die mir Kommentare, Fehler und Verbesserungsvorschläge geschickt haben: Fereshteh Fattahi, Kevin Kristen, Jonas Lensing, Florian Leptien, Natascha Scheibke und ebenfalls vielen Dank an die anonymen Hinweisgeber:innen auf Moodle.

Einleitung

Diese Notizen waren eigentlich nicht als Skript gedacht, sondern eher als Hilfestellung zur Erinnerung an die Vorlesung. Während des laufenden Semesters ist es schwierig, zusätzlich zur Vorlesung ein Buch zu schreiben – Sie können gerne einmal versuchen selbst eine Vorlesung im Computer aufzuschreiben, dann sehen Sie vielleicht, was ich meine – daher werden Sie hier auch mehr Tippfehler finden als mir lieb ist. Hinweise zu Fehlern und Tippfehlern nehme ich gerne entgegen.

Es gibt viele gute Quellen zur linearen Algebra und es ist nützlich, den Umgang mit anderen Quellen zu üben. Viele Bücher zur linearen Algebra sind über die Universitätsbibliothek auch online verfügbar, ein Beispiel ist das Buch von Gerd Fischer[2], es gibt auch viele Skripte wie zum Beispiel das von Ulrich Görtz[3] oder das von Wolfgang Soergel[4], englische Quellen wie das Buch von David Lay, Steven Lay und Judi McDonald[5] sind häufig sehr viel anwendungsorientierter und darum eine gute Ergänzung. Sie werden sehen, dass die Quellen fast den gleichen Inhalte umfassen, aber unterschiedlich erklären. Welche Darstellung Ihnen persönlich am leichtesten zugänglich ist, wird von Ihren Vorlieben und Vorkenntnissen abhängen.

[2] Gerd Fischer. *Lineare Algebra*, volume 17 of *Grundkurs Mathematik*. Friedr. Vieweg & Sohn, Braunschweig, fifth edition, 1979. ISBN 3-528-17217-7. URL https://link.springer.com/book/10.1007/978-3-8348-9365-9. In collaboration with Richard Schimpl

[3] Ulrich Görtz. Lineare algebra. Vorlesungsskript, 2020. URL https://math.ug/lecture-notes.html

[4] Wolfgang Soergel. Lineare algebra. Vorlesungsskript, 2022. URL http://home.mathematik.uni-freiburg.de/soergel/Skripten/XXLA1.pdf

[5] David Lay, Steven Lay, and Judi McDonald. *Linear Algebra and Its Applications*. Harlow: Pearson Education, Limited, 2015. ISBN 9781292092232. URL https://elibrary.pearson.de/book/99.150005/9781292092249

Die lineare Algebra ist ein erstaunlich universell einsetzbares Werkzeug, das Sie einerseits unbemerkt täglich im Alltag verwenden und das andererseits in fast jedem Stück moderner Mathematik eine Rolle spielt. Das ist für mich noch immer überraschend, denn vereinfacht gesagt geht es in der linearen Algebra nur darum, lineare Gleichungssysteme zu lösen und wir werden in der ersten Vorlesung ein Verfahren dafür kennenlernen — vielleicht kennen Sie das sogar schon aus der Schule.

Für mich ist das ein gutes Beispiel dafür, dass es sich oftmals lohnt, ein zunächst einfach erscheinendes Problem genauer anzuschauen. Haben wir den Kern des Problems so gut verstanden, dass wir den Kern kompakt zusammenfassen können, findet sich das gleiche Prinzip mit etwas Glück auch in ganz anderen Zusammenhängen wieder. Kurz: Abstraktion lohnt sich.

Das kompakte Zusammenfassen und genaue Verstehen ist nach den linearen Gleichungen (und der weitläufigen Verwandtschaft derselben) der zweite Kernpunkt der Vorlesung und vielleicht der schwierigere Teil, denn das Zusammenfassen ist mit der mathematische Sprache unverschämt gut möglich, aber wie mit jeder

Sprache braucht es Zeit und Übung, um sich an die Formulierungen und Formeln zu gewöhnen.

Mit etwas Gewöhnung ist es dann sehr erfreulich, dass wir damit komplexe Probleme so kompakt fassen können, dass wir einerseits leichter darüber nachdenken können — ich kann nicht ernsthaft über Lösungsmengen im 10 dimensionalen Raum nachdenken, aber viele natürliche Probleme haben viel mehr als 10 Parameter — und andererseits auch so sicher argumentieren können, dass wir verlässlich richtig und falsch unterscheiden können. Der zweite Punkt ist ebenso wichtig, denn in umständlicher Sprache werden Argumente leicht lang und unüberschaubar. Das genaue Argumentieren, das notwendig ist, um sicher zu sein, dass ein Argument wirklich auch in einem allgemeineren Kontext tragfähig ist, können Sie in der linearen Algebra gut lernen und üben.

Die moderne Form dieser algebraischen Sprache wurde wesentlich von Emmy Noether geprägt. Sie hat das selbst schön zusammengefasst:

> „Meine Methoden sind wirklich Methoden des Arbeitens und Denkens; deshalb haben sie sich überall anonym eingeschlichen."
> (*Emmy Noether, 1882-1935*)

Bildquelle: Wikipedia

„Entschuldigung, ich habe das nicht verstanden. Könnten Sie das bitte noch einmal wiederholen" sind wichtige Sätze, die Sie in jedem Sprachkurs ganz zu Anfang lernen, weil Sie sich ohne diese kaum unterhalten können. Das ist mit jeder Fachsprache genauso: Formulierungen, die Ihnen bei der ersten Begegnung unverständlich erscheinen, werden Ihnen mit etwas Nachfragen nach kurzer Zeit geläufig werden.

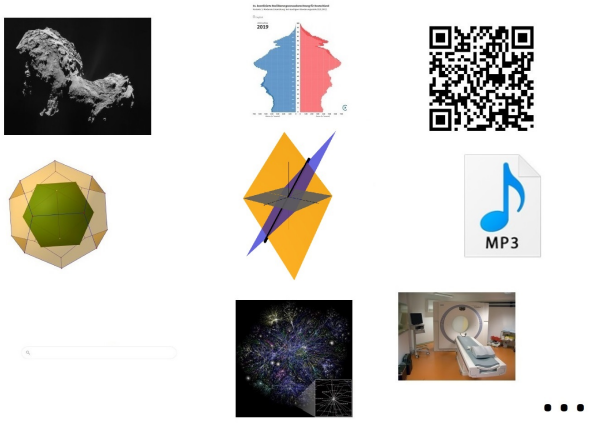
„Ich habe das nicht verstanden" bedeutet, dass Sie gerade versuchen, etwas zu verstehen. Haben Sie keine Angst vor diesem Satz. Das Nachfragen ist wichtig, denn es kann immer einmal passieren, dass ich versehentlich zu schnell mit der mathematischen Sprache umgehe und dabei Konstruktionen verwende, die für Sie ungewohnt oder gar unbekannt sind. Das ist keine Absicht. Zudem mache ich wie alle Menschen auch Fehler und dann ist es gut, wenn Sie das bemerken.

Sie werden niemals die einzige Person in der Vorlesung sein, die Ihre Frage hat, alle anderen werden sich heimlich bedanken, dass Sie sich getraut haben die Frage zu stellen.

Wissenschaft entsteht im Gespräch, nur selten in Monologen.

Lineare Gleichungssysteme

Lineare Algebra spielt in sehr unterschiedlichen Zusammenhängen eine wichtige Rolle, wahrscheinlich ohne dass Ihnen das bewusst ist. Hier eine kleine Auswahl:



Bildquellen: WikiCommons, Nasa, Statistisches Bundesamt, eigene Bilder

Die Liste könnte ich noch weiter verlängern, wenn Sie aus den Wirtschafts- oder Ingenieurwissenschaften kämen, hätte ich andere Bilder ausgewählt.

Am Beispiel der Internetsuche möchte ich Ihnen zeigen, wie der Zusammenhang zu Gleichungen in Problemen entstehen kann, die zunächst nicht mathematisch aussehen.

Wenn Sie eine Suchanfrage in eine der üblichen Suchmaschinen eintippen, bekommen Sie – nach werbefinanzierten Links – eine sortierte Liste, die meist in den ersten Einträgen eine für Sie relevante Seite enthält. Das Ursprungsproblem war hier:

Frage: Wie sortieren wir Seiten nach Relevanz, ohne den Inhalt zu analysieren?

Das hat zu Anfang ziemlich schlecht funktioniert, bis Google den PageRank-Algorithmus gefunden hat.

Die Idee war, nur auf das Netzwerk der Verknüpfungen von Seiten zu achten und dann anzunehmen, dass

1. eine Seite wichtig ist, wenn viele Seiten auf diese verlinken und

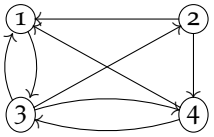
2. diese selbst nicht gleichzeitig auf sehr viele andere Seiten verweisen.

Ein mathematisches Modell für diese Grundidee könnte so aussehen: Die Relevanz einer Seite geben wir durch eine Zahl[6] w_{Seite} an, 0 sind völlig unwichtige Seiten, je größer die Zahl, desto wichtiger die Seite. Für die Verteilung stellen wir die folgenden Regeln auf:

[6] w für Wichtigkeit

1. Jede Seite gibt Ihre Wichtigkeit gleichmäßig an die Seiten ab, auf die sie verweist.
2. Jede Seite ist so wichtig, wie die Summe der Links auf die Seite.

Lassen Sie uns das an einem sehr kleinen Netzwerk aufschreiben:



Die Seiten sind im Bild nummeriert und die Links durch Pfeile gekennzeichnet. Die Anzahl der Links, die von den Seiten ausgehen sind:

Seite 1: 1 Link (nur nach 3),

Seite 2: 2 Links (nach 1 und 4)

Seite 3: 3 Links (nach 1,2,4)

Seite 4: 2 Links (nach 1 und 3).

Die Regeln können wir jetzt als Gleichungen lesen: Die Wichtigkeit w_1 von Seite 1 ist durch die Summe der Links gegeben. Die Seiten $2, 3, 4$ verweisen auf 1 und geben jeweils $1/(\#\text{ eigene Links})$ an w_1 weiter also:

Das Zeichen # möchte ich für „Anzahl“ verwenden.

$$w_1 = \frac{1}{2}w_2 + \frac{1}{3}w_3 + \frac{1}{2}w_4$$

Genauso erhalten wir

$$\begin{aligned} w_2 &= \frac{1}{3}w_3 \\ w_3 &= w_1 + \frac{1}{2}w_4 \\ w_4 &= \frac{1}{2}w_2 + \frac{1}{3}w_3 \end{aligned}$$

Bringen wir alle Terme auf eine Seite erhalten wir ein Gleichungssystem für die w_1, w_2, w_3, w_4:

$$\begin{array}{rcrcrcrcl} & w_1 & - & \frac{1}{2}w_2 & - & \frac{1}{3}w_3 & - & \frac{1}{2}w_4 & = & 0 \\ & & & w_2 & - & \frac{1}{3}w_3 & & & = & 0 \\ - & w_1 & & & & w_3 & - & \frac{1}{2}w_4 & = & 0 \\ & & - & \frac{1}{2}w_2 & - & \frac{1}{3}w_3 & + & w_4 & = & 0 \end{array}$$

Vielleicht sehen Sie hier eine für uns nutzlose Lösung sofort. Wenn Sie diese unnütze Lösung loswerden möchten könnten Sie die Summer der Relevanz aller Seiten fixieren, zum Beispiel als

$$w_1 + w_2 + w_3 + w_4 = 1$$

Dies ist ein lineares Gleichungssystem, d.h. es kommen hier nur Vielfache der Unbekannten w_1, w_2, w_3, w_4 vor, keine Produkte oder Potenzen.

Das Gleichungssystem für ein Netzwerk aus dem wirklichen Leben wird sehr groß sein, so groß, dass es nicht so klar ist, wie wir jemals eine Lösung finden.

FRAGE: Kennen Sie ein Lösungsverfahren, das für jedes derartige Gleichungssystem funktioniert?

In der Mathematik wird es vielfach um die Frage gehen, wie wir an ein Problem herangehen können, für das wir noch keine Lösung kennen.

Wie finden wir ein Lösungsverfahren?

EINE LÖSUNGSSTRATEGIE möchte ich an dieser Frage vorstellen. Selbst wenn Sie schon ein Lösungsverfahren kennengelernt haben, ist es nützlich, die Frage zu stellen, woher die Idee zu dem Verfahren vielleicht kam.

Wenn ich unser Beispiel anschaue, dann stören mich zunächst die Brüche, die noch schlimmer werden, wenn ich ein größeres Netzwerk mit mehr Links anschauen möchte.

SCHRITT EINS: Sammle Beispiele, die

- überschaubar genug sind, um leicht eine Lösung zu finden und
- zusammen eine Chance haben, typisch zu sein.[7]

[7] Ich würde das allgemeine Problem gern auf meine einfachen Beispiele zurückführen.

Insbesondere werde ich statt krummen Zahlen, zunächst einfache Koeffizienten ausprobieren, weil ich sonst die Übersicht verliere.

Genauso nenne ich meine Unbekannten $x_1, x_1, \ldots$, weil ich mich daran gewöhnt habe.

Für ein Gleichungssystem sehe ich zwei Möglichkeiten, meine Arbeit zu erleichtern:

1. Weniger Gleichungen.
2. Weniger Unbekannte.

Vielleicht fangen wir mit weniger Gleichungen an:

HABEN WIR NUR EINE GLEICHUNG zum Beispiel

$$x_1 + 2x_2 + 3x_3 + 4x_4 = 10$$

so können wir alle Lösungen sehen: Setzen wir nämlich für x_2, x_3, x_4 beliebige feste Werte a, b, c ein, so können wir den nötigen Wert für x_1 ausrechnen:

$$x_1 = 10 - 2a - 3b - 4c$$

Diese Notation ist nicht sehr nachhaltig, mit a, b, c werden uns die Buchstaben schnell ausgehen.

Damit haben wir die Lösungen der Gleichung bestimmt:

$$\text{Lösungsmenge} = \{(x_1, x_2, x_3, x_4) = (10 - 2a - 3b - 4c, a, b, c) | a, b, c \text{ beliebig}\}.$$

Die Lösungsmenge sähe also geometrisch aus, wie ein 3-dimensionaler Raum mit den Koordinaten a, b, c. Das passt ganz gut zu Beispielen mit weniger Unbekannten:

$$x_1 + 2x_2 = 10$$

beschreibt eine Gerade in der Ebene und

$$x_1 + 2x_2 + 3x_3 = 10$$

eine Ebene im Raum.

In Termen von Formeln wird das erfreulicherweise nicht schwieriger, wenn wir mehr Variablen haben als wir uns geometrisch vorstellen können.

Was machen wir nun mit mehereren Gleichungen? Ein einfaches Beispiel wäre vielleicht:

$$\begin{array}{rcrcrcl} x_1 & + & 2x_2 & + & 3x_3 & = & 10 \\ & & 4x_2 & + & 5x_3 & = & 10 \\ & & & & 6x_3 & = & 12. \end{array}$$

Aufgabe. Sehen Sie eine Lösung?

Wieso sehen wir die Lösung?

- In der letzten Zeile steht $x_3 = 2$ in komplizierterer Schreibweise.
- Setzen wir $x_3 = 2$ in der Zeile darüber ein, erhalten wir $4x_2 + 5 \cdot 2 = 10$ also $4x_2 = 0$ und damit $x_2 = 0$.
- Setzen wir die Werte $x_2 = 0$, $x_3 = 2$ in die erste Zeile ein erhalten wir $x_1 + 0 + 3 \cdot 2 = 10$ also $x_1 = 4$.

Fazit: Wenn das Gleichungssystem so aussieht, dass ab der 2. Gleichung nur noch die Unbekannten Nummer $2, 3, \ldots$ in der 3. Gleichung nur noch die Unbekannten Nummer $3, \ldots$ usw. vorkommen, so können wir die Lösung einfach bestimmen.

Lassen Sie uns die gerade verwendeten Argumente noch einmal systematisch durchgehen, um zu sehen, was wir daraus für kompliziertere System lernen können:

- In der letzten Zeile steht $x_3 = 2$ in komplizierterer Schreibweise.

Bedeutet: Wir teilen die letzte Zeile durch 6 (=multiplizieren mit $\frac{1}{6}$):

$$\begin{array}{rcrcrcl} x_1 & + & 2x_2 & + & 3x_3 & = & 10 \\ & & 4x_2 & + & 5x_3 & = & 10 \\ & & & & 6x_3 & = & 12. \mid \cdot \frac{1}{6} \end{array}$$

$$\begin{array}{rcrcrcl} x_1 & + & 2x_2 & + & 3x_3 & = & 10 \\ & & 4x_2 & + & 5x_3 & = & 10 \\ & & & & x_3 & = & 2. \end{array}$$

OPERATION 1: Die Lösungsmenge einer Gleichung ändert sich nicht, wenn wir die Gleichung mit einer Zahl $\neq 0$ multiplizieren.
Denn gilt $x = y$, so ist $ax = ay$. Wenn $a \neq 0$, so können wir umgekehrt $ax = ay$ mit $\frac{1}{a}$ multiplizieren und erhalten dann $x = y$ zurück.

Den zweiten Schritt:

- setzen wir $x_3 = 2$ in der Zeile darüber ein, erhalten wir $4x_2 + 5 \cdot 2 = 10$ also $4x_2 = 0$

können wir auch als: Subtrahiere das 5-fache der letzten Zeile von der zweiten:

$$\begin{array}{rcrcrcl} x_1 & + & 2x_2 & + & 3x_3 & = & 10 \\ & & 4x_2 & + & 5x_3 & = & 10 \longleftarrow + \\ & & & & x_3 & = & 2. \mid \cdot -5 \end{array}$$

$$\begin{array}{rcrcrcl} x_1 & + & 2x_2 & + & 3x_3 & = & 10 \\ & & 4x_2 & & & = & 0 \\ & & & & x_3 & = & 2. \end{array}$$

interpretieren.

OPERATION 2: Die Lösungsmenge einer Gleichung ändert sich nicht, wenn wir ein Vielfaches einer Gleichung zu einer anderen addieren.
Denn gilt $x = y$ und $z = w$, so gilt für jede Zahl a auch $x + az = y + aw$.

Umgekehrt erhalten wir aus $x + az = y + aw$ und $z = w$ die Gleichung $x = y$ zurück, indem wir das $-a$ fache der Gleichung $z = w$ zu $x + az = y + aw$ addieren.

Mit den beiden Operationen können wir jetzt auch den Rest unserer Lösung formulieren:

$$\begin{array}{rcrcrcl} x_1 & + & 2x_2 & + & 3x_3 & = & 10 \\ & & 4x_2 & + & & = & 0 \mid \cdot \frac{1}{4} \\ & & & & x_3 & = & 2. \end{array}$$

$$\begin{array}{rcrcrcl} x_1 & + & 2x_2 & + & 3x_3 & = & 10 \longleftarrow + \longleftarrow + \\ & & x_2 & + & & = & 0 \qquad \mid \cdot -2 \\ & & & & x_3 & = & 2. \mid \cdot -3 \end{array}$$

$$\begin{array}{rcrcrcl} & & x_1 & & & = & 4 \\ & & & & x_2 & = & 0 \\ & & & & x_3 & = & 2. \end{array}$$

Im einfachen Beispiel haben wir die Umformungen jetzt so formuliert, dass wir diese für allgemeine Gleichungssysteme verwenden können. Damit können wir jedes System so umschreiben, dass es wie im Fazit zu unserem einfachen Beispiel aussieht.

FAZIT: Wenn das Gleichungssystem so aussieht, dass ab der 2. Gleichung nur noch die Unbekannten Nummer $2, 3, \ldots$ in der 3. Gleichung nur noch die Unbekannten Nummer $3, \ldots$ usw. vorkommen, so können wir die Lösung einfach bestimmen.

Ein allgemeines Lösungsverfahren

2. Vorlesung 12.10.

DER GAUSS-ALGORITHMUS macht das wie folgt: Gegeben ein lineares Gleichungssystem für die Unbekannten $x_1, x_2, \ldots, x_n$.
Etappe 1: *Entferne von oben nach unten, schrittweise jeweils die vordere Variable aus den darunter stehenden Gleichungen.* Formal geht das so – keine Sorge wir machen gleich ein Beispiel in dem das hoffentlich einfach aussieht:

In der nächsten Vorlesung werde ich genauer formulieren, was hiermit gemeint ist

Schritt 1 Sortiere die Gleichungen so, dass die Unbekannte x_1 in der ersten Gleichung vorkommt. Das Kleingedruckte: Wenn x_1 in keiner Gleichung vorkommen sollte, kann der Wert für x_1 beliebig gewählt werden. Wir führen Schritt 1 dann für die nächste Variable x_2 durch.

Schritt 2 Teile die 1. Gleichung durch den Koeffizienten von x_1. Die Gleichung ist jetzt also $x_1 + \text{etwas} \cdot x_2 + \cdots = \text{Konstante}$.

Wir sollten noch eine bessere Notation für etwas finden.

Schritt 3 Subtrahiere dann von oben nach unten jeweils das Vielfache der 1. Gleichung von den darunter stehenden Gleichungen 2, 3, usw., das durch den Koeffizienten von x_1 gegeben ist. Damit wird x_1 aus den Gleichungen 2, 3, usw., entfernt.

Erstes Zwischenergebnis: x_1 kommt in den Gleichungen $2, 3, \ldots$ nicht mehr vor.

Schritt 4+ Wiederhole das Verfahren für die Gleichungen $2, 3 \ldots$ mit den Variablen x_2 und danach mit den nächsten Variablen.

ERGEBNIS DER ERSTEN ETAPPE: Wir erhalten durch Umformen von oben nach unten, ein Gleichungssystem in dem dass ab der 2. Gleichung nur noch die Unbekannten Nummer $2, 3, \ldots$ in der 3. Gleichung nur noch die Unbekannten Nummer $3, \ldots$ usw. vorkommen und der Koeffizient der führenden Variablen in jeder Gleichung jeweils 1 ist.

Wir sagen dazu: Das Gleichungssystem hat *Zeilenstufenform*.

$$\begin{aligned} x_1 + 2x_2 + 3x_3 + x_4 &= 5 \\ x_3 + 5x_4 &= 6 \\ x_4 &= 1. \end{aligned}$$

Etappe 2: *Entferne* - wie im einfachen Beispiel - *mit Umformungen von unten nach oben, die führenden Variablen der Gleichungen aus den darüberstehenden:*

Schritt 5 Subtrahiere jetzt das Vielfache der letzten Zeile von den darüberstehenden, das die führende Variable der letzten Gleichung aus den darüberstehenden entfernt.

Schritt 6+ Wiederhole das Verfahren danach von unten nach oben für alle Gleichungen, so dass die führenden Unbekannten jeweils nur in einer einzigen Gleichung vorkommen.

ERGEBNIS DER ZWEITEN ETAPPE: Im Gleichungssystem kommt die erste Variable nur in der ersten Gleichung vor, in den darunter stehenden fällt jeweils wenigstens eine weitere Variable weg. Die jeweils führende Variable in jeder Gleichung kommt in keiner anderen Gleichung mehr vor.

Wir sagen dazu: Das Gleichungssystem hat *reduzierte Zeilenstufenform*.

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

MERKSCHEMA:

$$\begin{array}{cccc|} \star & \star & \star & \star=\star \\ \star & \star & \star & \star=\star \\ \star & \star & \star & \star=\star \\ \downarrow & & & \end{array} \quad \rightsquigarrow \quad \begin{array}{cccc|} & x_1 & \star & \star & \star=\star \\ & 0 & x_2 & \star & \star=\star \\ & 0 & 0 & x_3 & \star=\star \\ & & & & \uparrow \end{array} \quad \rightsquigarrow \quad \begin{array}{cccc} & x_1 & 0 & 0 & \star=\star \\ & 0 & x_2 & 0 & \star=\star \\ & 0 & 0 & x_3 & \star=\star \end{array}$$

Jetzt können wir die Lösungen ablesen:

- Die Unbekannten, die nicht als führende Variable des vereinfachten Gleichungssystems vorkommen, können beliebig gewählt werden.
- Für die Unbekannten, die als führende Variable in einer Gleichung vorkommen, gibt diese Gleichung eine Formel für die Lösung.
- Ist eine der Gleichungen von der Form $0 =$ Konstante ungleich 0, so hat das System keine Lösung.

Das wird klarer, wenn wir ein Beispiel rechnen:

Ein Beispiel für das Gauß-Verfahren

Beispiel 1. Lösung des Gleichungssystems:

$$\begin{array}{rcrcrcl} & & x_1 & + & x_2 & + & x_3 & = & 1 \\ & & 3x_1 & + & 2x_2 & + & x_3 & = & 6 \\ & & 6x_1 & + & 5x_2 & + & 5x_3 & = & 6 \end{array}$$

$$\begin{array}{rcrcrcll} x_1 & + & x_2 & + & x_3 & = & 1 & \mid \cdot -3 \\ \mathbf{3}x_1 & + & 2x_2 & + & x_3 & = & 6 & \leftarrow + \\ \mathbf{6}x_1 & + & 5x_2 & + & 5x_3 & = & 6 & \leftarrow + \end{array} \qquad \mid \cdot -6$$

$$\begin{array}{rcrcrcll} x_1 & + & x_2 & + & x_3 & = & 1 & \\ & - & x_2 & - & 2x_3 & = & 3 & \mid \cdot -1 \\ & - & x_2 & - & x_3 & = & 0 & \end{array}$$

$$\begin{array}{rcrcrcll} x_1 & + & x_2 & + & x_3 & = & 1 & \\ & + & x_2 & + & 2x_3 & = & -3 & \\ & - & x_2 & - & x_3 & = & 0 & \leftarrow + \end{array}$$

$$\begin{array}{rcrcrcll} x_1 & + & x_2 & + & x_3 & = & 1 & \leftarrow + \\ & + & x_2 & + & 2x_3 & = & -3 & \leftarrow + \\ & & & + & x_3 & = & -3 & \mid \cdot -2 \end{array} \qquad \mid \cdot -1$$

Zeilenstufenform

$$\begin{array}{rcrcrcll} x_1 & + & x_2 & & & = & 4 & \leftarrow + \\ & + & x_2 & & & = & 3 & \mid \cdot -1 \\ & & & + & x_3 & = & -3 & \end{array}$$

$$\begin{array}{rcrcrcl} x_1 & & & & & = & 1 \\ & & x_2 & & & = & 3 \\ & & & & x_3 & = & -3 \end{array}$$

reduzierte Zeilenstufenform

Bemerkung. Es ist in fast jedem Beispiel verlockend, sich den Algorithmus zu erleichtern indem geschickte Umformungen eingeschoben werden. Oben hätte ich zum Beispiel erst einmal die letzte und die erste Zeile vertauscht, um mir Brüche zu ersparen.

Das erstaunliche ist aber, dass die Methode als Kochrezept immer funktioniert und ich selbst häufig weniger Rechenfehler mache, wenn ich mich schlicht an das Verfahren halte.

Aufgabe 1. Bestimmen Sie alle Lösungen des Gleichungssystems:

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

In der Vorlesung haben wir das gemeinsam gelöst:

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

$$\begin{array}{rcrcrcrcl} x_1 & + & x_2 & & & + & 2x_4 & = & 1 \\ & & x_2 & + & x_3 & + & x_4 & = & 1 \mid -1 \\ & & x_2 & + & x_3 & + & x_4 & = & 1 \end{array}$$

$$\begin{array}{rcrcrcrcl} x_1 & + & x_2 & & & + & 2x_4 & = & 1 \\ & & x_2 & + & x_3 & + & x_4 & = & 1 \mid -1 \\ & & & & & & 0 & = & 0 \end{array}$$

$$\begin{array}{rcrcrcrcl} x_1 & + & & & -x_3 & + & x_4 & = & 0 \\ & & x_2 & + & x_3 & & + & x_4 & = & 1 \end{array}$$

Also können x_3 und x_4 beliebig gewählt werden und jede Wahl bestimmt x_1, x_2 dann eindeutig:

$$\text{Lösungsmenge} = \left\{ \begin{array}{l} x_1 = s - t \\ x_2 = 1 - s - t \\ x_3 = s \\ x_4 = t \end{array} \;\middle|\; s, t \text{ beliebig} \right\}$$

Wie Sie die frei wählbaren Variablen nennen, bleibt Ihnen überlassen, ich nehme immer aufeinander folgende Buchstaben, aber wenn Sie statt s, t lieber „Emmy“ und „Oscar“ nehmen, ist das Ihre Entscheidung.

Aufgabe 2. Bestimmen Sie alle Lösungen des Gleichungssystems:

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

Matrizen - eine abkürzende Schreibweise für Gleichungssysteme

Wahrscheinlich ist Ihnen die Schreibarbeit in den Beispielen schon unangenehm vorgekommen. Das liegt an den Unbekannten x_i, die immerzu abgeschrieben werden, ohne wirklich an den Umformungen beteiligt zu sein. Um das loszuwerden sind Matrizen praktisch.

EIN LINEARES GLEICHUNGSSYSTEM mit n Unbekannten $x_1, x_2, \ldots, x_n$ und m Gleichungen ist ein Gleichungssystem der Form:

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

Der erste Index in a_{12} gibt die Nummer der Gleichung (die Zeile) an, der zweite die Nummer der Unbekannten (die Spalte). Also

$$a_{\text{Zeile,Spalte}}.$$

wobei $(a_{ij})_{\substack{i=1,\ldots m \\ j=1,\ldots m}}$ und b_i fest gewählte Zahlen sind.

Zur Abkürzung schreiben wir die Koeffizienten a_{ij} in eine Matrix

$$A := \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix}.$$

In unserem Beispiel

$$\begin{array}{rcrcrcl} x_1 & + & x_2 & + & x_3 & = & 1 \\ 3x_1 & + & 2x_2 & + & x_3 & = & 6 \\ 6x_1 & + & 5x_2 & + & 5x_3 & = & 6 \end{array}$$

Bekommen wir also

$$A := \begin{pmatrix} 1 & 1 & 1 \\ 3 & 1 & 1 \\ 6 & 5 & 5 \end{pmatrix}.$$

Wenn wir daraus das Gleichungssystem zurück bekommen möchten, schreiben wir die Variablen untereinander als Vektor

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Und definieren

$$A \cdot x = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}$$

ICH MERKE MIR DIE REGEL SO:

$$\overrightarrow{\begin{pmatrix} a_{1,1} & a_{1,2} & \ldots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ldots & a_{2,n} \\ \vdots & & & \vdots \\ a_{m,1} & a_{m,2} & \ldots & a_{m,n} \end{pmatrix}} \cdot \left.\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}\right\downarrow =$$

Laufen Sie mit zwei Fingern, an den beiden Pfeilen entlang, multiplizieren Sie jeweils die Einträge und addieren Sie die Ergebnisse auf. Das ist der Erste Eintrag im Ergebnisvektor. Dann machen Sie das für die 2. Zeile usw.

Der Vektor auf der rechten Seite der Gleichung für $A \cdot x$ hat m Einträge. Schreiben wir also

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

so wird unser Gleichungssystem zur Gleichung

Diese kompakte Schreibweise wird noch ihr Eigenleben entwickeln.

$$A \cdot x = b.$$

Fazit:
Mit

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ \vdots & & & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}, \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ und } \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

wird aus dem Gleichungssystem

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \cdots + a_{1,n}x_n &= b_1 \\ \vdots & \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + \cdots + a_{m,n}x_n &= b_m \end{aligned}$$

die kurze Gleichung

Diese einfache Gleichung wird bald noch ein Eigenleben entwickeln.

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

Im Gauss-Algorithmus mussten wir die rechte Seite des Gleichungssystems mitführen. In Matrixnotation schreiben wir dafür die rechte Seite hinter einem senkrechten Strich an die Matrix: Für das Beispiel

$$\begin{array}{rcrcrcrcl} & & & & x_3 & & & + \; x_5 & = & -4 \\ x_1 & - & x_2 & + & x_3 & - & 2x_4 & & = & 1 \\ 2x_1 & - & 2x_2 & + & 2x_3 & - & 2x_4 & & = & 0 \end{array}$$

schreiben wir also

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

Das nennen wir auch die *erweiterte Koeffizientenmatrix*.
Der senkrechte Strich markiert die Stelle des „=“ in den Gleichungen, also die Stelle an der die Unbekannten aufhören.

Den Gauß-Algorithmus hierfür können wir dann so aufschreiben:

$$
\begin{pmatrix} 0 & 0 & 1 & 0 & 1 & | & -4 \\ 1 & -1 & 1 & -2 & 0 & | & 1 \\ 2 & -2 & 2 & -2 & 0 & | & 0 \end{pmatrix} \leftarrow \\
\begin{pmatrix} 1 & -1 & 1 & -2 & 0 & | & 1 \\ 0 & 0 & 1 & 0 & 1 & | & -4 \\ 2 & -2 & 2 & -2 & 0 & | & 0 \end{pmatrix} \begin{matrix} | \cdot -2 \\ \\ \leftarrow \end{matrix} + \\
\begin{pmatrix} 1 & -1 & 1 & -2 & 0 & | & 1 \\ 0 & 0 & 1 & 0 & 1 & | & -4 \\ 0 & 0 & 0 & 2 & 0 & | & -2 \end{pmatrix} \begin{matrix} \leftarrow \\ | \cdot -1 \\ \\ \end{matrix} + \\
\begin{pmatrix} 1 & -1 & 0 & -2 & -1 & | & 5 \\ 0 & 0 & 1 & 0 & 1 & | & -4 \\ 0 & 0 & 0 & 2 & 0 & | & -2 \end{pmatrix} \begin{matrix} \leftarrow \\ \\ | \cdot 1 \end{matrix} + \\
\begin{pmatrix} 1 & -1 & 0 & 0 & -1 & | & 3 \\ 0 & 0 & 1 & 0 & 1 & | & -4 \\ 0 & 0 & 0 & 2 & 0 & | & -2 \end{pmatrix} | \cdot \tfrac{1}{2} \\
\begin{pmatrix} 1 & \mathbf{-1} & 0 & 0 & \mathbf{-1} & | & 3 \\ 0 & 0 & 1 & 0 & \mathbf{1} & | & -4 \\ 0 & 0 & 0 & 1 & \mathbf{0} & | & -1 \end{pmatrix}
$$

Daran können wir wie gehabt die Lösungsmenge ablesen: Die Variablen, die den fettgedruckten Spalten entsprechen können wir frei wählen und dann für die führenden Variablen eine Formel ablesen:

$$
\text{Lösungen}(A \cdot x = b) = \left\{ \begin{pmatrix} x_1 \\ \mathbf{x_2} \\ x_3 \\ x_4 \\ \mathbf{x_5} \end{pmatrix} = \begin{pmatrix} 3 + \mathbf{s} + \mathbf{t} \\ \mathbf{s} \\ -4 - \mathbf{t} \\ -1 \\ \mathbf{t} \end{pmatrix} \middle| \, \mathbf{s}, \mathbf{t} \text{ beliebig} \right\}
$$

Für den Moment, bleibt es Ihnen überlassen, ob Sie die längere Schreibweise bevorzugen.

Zahlbereiche

3. Vorlesung 17.10.

Wie bei der Suche nach einem Lösungsverfahren lohnt es sich, nachzuschauen, was wir im Verfahren wirklich benutzt haben, auch wenn das zunächst banal erscheint.

Zur Lösung von linearen Gleichungssystemen haben wir nur die Grundrechenarten $+, -, \cdot, :$ verwendet. Die kennen Sie natürlich lange aus unterschiedlichen Zahlbereichen:

Für die natürlichen Zahlen

$$\mathbb{N} = \{1, 2, 3, \dots\}$$

haben wir die Verknüpfungen $+$ und $\cdot$, aber schon die Subtraktion ist nicht immer möglich.

In den ganzen Zahlen

$$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$$

sind $+, -, \cdot$ definiert, nur dividieren macht Probleme.

In den rationalen Zahlen (den Brüchen)

$$\mathbb{Q} = \left\{ \text{Brüche } \frac{n}{m} \,\middle|\, n \in \mathbb{Z}, m \in \mathbb{N} \right\}$$

sind dann alle Verknüpfungen $+, -, \cdot, :$ definiert.

Ausserdem kennen Sie die reellen Zahlen $\mathbb{R}$, über die Sie wahrscheinlich als Dezimalzahlen, mit potentiell unendlich vielen Nachkommastellen nachdenken

$$\mathbb{R} = \left\{ n, d_1 d_2 d_3 \dots \mid n \in \mathbb{Z}, d_i \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, \} \right\};$$

zum Beispiel enthält $\mathbb{R}$ die Zahl $\pi = 3,141592653589\dots$.

Komplexe Zahlen

Es gibt weitere nützliche Zahlbereiche in denen wir genauso gut rechnen können, zum Beispiel die komplexen Zahlen $\mathbb{C}$.

Die komplexen Zahlen wurden ursprünglich erfunden, weil aufgefallen war, dass sich mit Zwischenergebnissen wie Termen aus

der p–q–Formel auch dann sinnvoll weiter rechnen lässt, wenn in der Wurzel eine negative Zahl steht. Dieser Zahlbereich hat sich dann als außerordentlich nützlich erwiesen.[8]

[8] Zum Beipiel kann ich mir selbst die Rechenregeln für die Winkelfunktionen sin, cos erst merken, seitdem ich weiß wie sich diese in den komplexen Zahlen zu einer sehr viel einfacheren Formel zusammensetzen.

Der Trick für die Definition der komplexen Zahlen ist, einfach der Algebra zu vertrauen: Da wir in $\mathbb{R}$ keine Zahl a mit $a^2 = -1$ finden, d.h. es gibt kein $a = \sqrt{-1}$ in $\mathbb{R}$, fügen wir Symbol i zu $\mathbb{R}$ hinzu, für das wir die Rechenregel $i^2 = -1$ festlegen (also $i = \sqrt{-1}$ definiert wird). Wir machen uns zunächst keine Gedanken über eine Bedeutung hiervon, sondern rechnen mit dem Symbol i.

Dazu werden wir mindestens alle Ausdrücke der Form $a + i \cdot b$ wobei a, b reelle Zahlen sind benötigen und es wird sich herausstellen, dass das schon ausreicht:

$$\mathbb{C} := \{a + ib \,|\, a, b \in \mathbb{R}\},$$

d.h. die komplexen Zahlen sind einfach formale Ausdrücke der Form $a + ib$ für reelle Zahlen a, b, also beispielsweise

$$1,\ i,\ 1+i,\ 1-i,\ 2+3i,\ \pi - \sqrt{2}i,\ \ldots$$

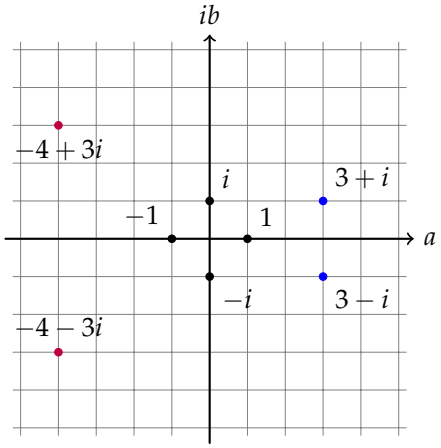
Sie sehen, dass ich statt $0 + i$ einfach i schreibe und für $1 + i \cdot 0$ einfach 1.

Für komplexe Zahlen können wir die Grundrechenarten $+, -, \cdot, :$ erweitern:

$$\begin{aligned}
(a+ib) + (c+id) &:= (a+c) + i(b+d) \\
&\text{denn } a + ib + c + id \overset{!}{=} a + c + ib + id \overset{!}{=} a + c + i(b+d) \\
(a+ib) \cdot (c+id) &:= ac - bd + i(ad+bc) \\
&\text{denn} \\
(a+ib) \cdot (c+id) &= ac + aid + ibc + ibid \text{ (Ausmultiplizieren)} \\
&= ac + iad + ibc + i^2 bd \\
&\overset{i^2=-1}{=} ac + i(ad+bc) - bd.
\end{aligned}$$

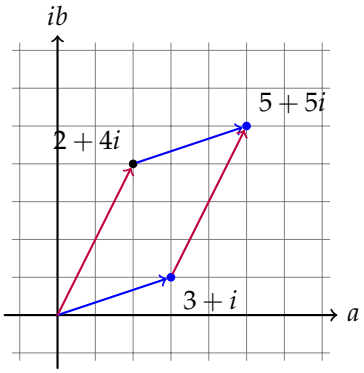
Die Rechenregeln ergeben sich also einfach aus den üblichen Regeln zum Ausmultiplizieren und Zusammenfassen für reelle Zahlen, wir müssen uns also keine neuen Formeln (außer $i^2 = -1$) merken.

Eine geometrische Interpretation dieser Regeln ist beruhigend und hilfreich. Wenn wir uns $\mathbb{C} := \{a + ib | a, b \in \mathbb{R}\}$ geometrisch als die Punkte in der Ebene vorstellen:



können wir die Addition und Multiplikation auch geometrisch interpretieren.

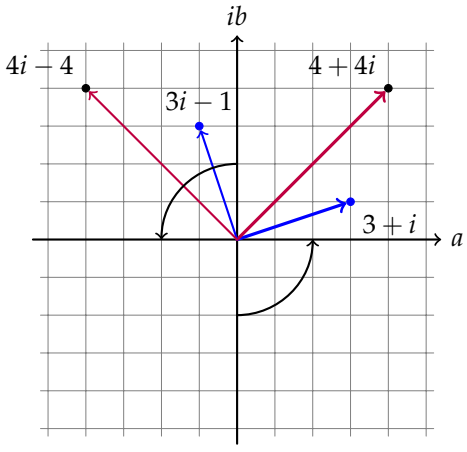
Die Addition ist durch aneinanderlegen von Ortsvektoren gegeben:



Die Multiplikation mit i ergibt

$$i \cdot (a + ib) = ia + i^2 b = -b + ia.$$

Die $a-$Achse wird also auf die ib-Achse gedreht und die $ib-$Achse auf die $-a$-Achse. Das entspricht also einer Drehung um 90°:



Lassen Sie uns testen, ob diese Vorstellung gut funktioniert. Wenn $i = \sqrt{-1}$ eine Drehung um 90° ist, sollte vielleicht $\sqrt{i}$ eine Drehung um 45° sein. Mein erster Versuch für die Drehung wäre

vielleicht, die Winkelhalbierende zwischen i und 1 zu nehmen. Das wäre

$$\frac{1+i}{|1+i|} = \frac{1+i}{\sqrt{2}} = \frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}},$$

weil die Diagonale im Einheitsquadrat die Länge $\sqrt{2}$ hat.

Probe:

$$\begin{aligned}(\frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}})^2 &= \frac{1}{\sqrt{2}}^2 + 2i(\frac{1}{\sqrt{2}})^2 + (i\frac{1}{\sqrt{2}})^2 \\ &= \frac{1}{2} + i + (i^2\frac{1}{2}) = \frac{1}{2} + i - \frac{1}{2} = i.\end{aligned}$$

Wir haben also tatsächlich geometrisch eine Wurzel gefunden. Sie können das gern für $\sqrt[4]{i}$ oder $\sqrt[3]{-1}$ auch selbst einmal probieren.

NUN FEHLEN UNS NOCH die Grundrechenarten „$-$“ und „:“. Die Subtraktion können wir als

$$(a+ib) - (c+id) = (a-c) + i(b-d)$$

definieren, denn $-$ ist die Operation die $+$ rückgängig macht.

DIE DIVISION ist vielleicht überraschender. Der Ausdruck $\frac{a+ib}{c+id}$ lässt sich mit einem Trick in die Form $e+if$ bringen. Die 3. binomische Formel sagt nämlich, dass

$$(c+id)\cdot(c-id) = c^2 - (id)^2 = c^2 - (-d^2) = c^2 + d^2$$

gilt und c^2+d^2 ist eine positive reelle Zahl wenn $c+id \neq 0$ ist. Also gilt:

$$\begin{aligned}\frac{a+ib}{c+id} &= \frac{(a+ib)(c-id)}{(c+id)(c-id)} && \text{erweitere den Bruch mit } c-id \\ &= \frac{(ac+bd+i(bc-ad)}{c^2+d^2} && \text{Ausmultiplizieren} \\ &= \underbrace{\frac{ac+bd}{c^2+d^2}}_{\text{in } \mathbb{R}} + i\,\underbrace{\frac{bc-ad}{c^2+d^2}}_{\text{in } \mathbb{R}} && \text{Bruchrechnen.}\end{aligned}$$

Beispiel 2.

- $\frac{1}{1+i}$ berechnen wir, indem wir den Bruch wie oben mit $1-i$ erweitern:

$$\frac{1}{1+i} = \frac{1(1-i)}{(1+i)(1-i)} = \frac{1-i}{1^2-i^2} = \frac{1-i}{1^2-(-1)} = \frac{1-i}{2} = \frac{1}{2} - i\frac{1}{2}.$$

- $\frac{1}{i} = \frac{1\cdot -i}{i\cdot -i} = \frac{-i}{-(-1)} = -i$. Also

$$\frac{1}{i} = -i.$$

Um das Rechen zu üben, haben wir in der Vorlesung noch den Gaußalgorithmus für das Gleichungssystem

$$\begin{aligned}ix_1 - 3x_2 &= 3i \\ 2ix_1 - (3+3i)x_2 &= 6\end{aligned}$$

durchgeführt.

$$\begin{array}{rcrcl} ix_1 & - & 3x_2 & = & 3i \mid \cdot \frac{1}{i} \\ 2ix_1 & - & (3+3i)x_2 & = & 6 \end{array}$$

$$\begin{array}{rcrcl} x_1 & + & 3ix_2 & = & 3 \mid -2i \\ 2ix_1 & - & (3+3i)x_2 & = & 6 \leftarrow + \end{array}$$

$$\begin{array}{rcrcl} x_1 & + & 3ix_2 & = & 3 \\ 0 & + & (3-3i)x_2 & = & 6-6i \mid \cdot \frac{1}{3-3i} \end{array}$$

$$\begin{array}{rcrcl} x_1 & + & 3ix_2 & = & 3 \leftarrow + \\ & & x_2 & = & 2 \mid \cdot -3i \end{array}$$

$$\begin{array}{rcl} x_1 & = & 3-6i \\ x_2 & = & 2. \end{array}$$

Also ist die einzige Lösung $x_1 = 3 - 6i$ und $x_2 = 2$.

Ich empfehle an dieser Stelle dringend, eine Probe zu machen, um das Ergebnis zu prüfen.

Es gibt jetzt ein ernstes Problem: Die Rechnungen sind schön und gut, aber wie können wir uns systematisch davon überzeugen, dass tatsächlich „alle üblichen Rechenregeln" auch für die komplexen Zahlen gelten? Einige davon haben wir gerade verwendet!

Bei der Gelegenheit wäre es auch schön zu klären, was genau hier mit „alle" gemeint ist.

Aufgabe 3. Sammeln Sie *bevor Sie weiterlesen* wenigstens 5 Eigenschaften und Rechenregeln für die 4 Grundrechenarten, die Sie beim Rechnen mit reellen Zahlen gewöhnt sind.

Das können einfache Termumformungen mit $+, -, \cdot, :$ sein, oder einfache Rückschlüsse, die Sie aus Gleichungen wie zum Beispiel $a \cdot b = 0$, ziehen können.

VIELLEICHT HABEN SIE einige der folgenden Regeln gesammelt:

- $a - (b - c) = a - b + c$
- Ein Produkt ist genau dann 0, wenn einer der Faktoren 0 ist.
- $a - b = a + (-1) \cdot b$
- Aus $a + b = a + c$ folgt $b = c$.
- Ist $a \neq 0$ und $a \cdot b = a \cdot c$, so gilt $b = c$.
- Wir teilen durch einen Bruch, indem wir mit dem Kehrwert multiplizieren, d.h. es gilt

  $$a : (b : c) = a \cdot (c : b)$$

  wenn $b \neq 0$ und $c \neq 0$.
- Gilt $a \cdot b = b$ so ist entweder $a = 1$ oder $b = 0$.

Für die komplexen Zahlen hatten wir zum Beispiel $\frac{1}{c+id}$ mit einem Trick berechnet, wenn Sie einen anderen Trick finden, könnte ein anderes Ergebnis herauskommen, d.h. es könnte doch sein, dass mehrere Zahlen als Inverse funktionieren, oder?

Es scheint ein mittlerer Alptraum alle diese Dinge nachzuprüfen, oder einen wirklichen Überblick über die Regeln zu bekommen. Schlimmer noch, wenn wir nächste Woche noch einen neuen Zahlbereich finden sollten, dann geht das wieder von vorne los.

DIE EINZIGE BRAUCHBARE LÖSUNG ist, eine möglichst minimale Liste von Regeln zu finden, aus der sich alle anderen ergeben. Wenn die Liste kurz genug ist, können wir das leicht nachprüfen und dann einmal überlegen, was sich alles aus der Liste ergibt.

Körperaxiome: Ein Minimum an Rechenregeln genügt

Vorlesung 4, 19.10.

Um nicht für die komplexen Zahlen $\mathbb{C}$ und jeden weiteren Zahlbereich keine lange Liste von Eigenschaften nachprüfen zu müssen, ist es hilfreich eine minimale Liste zu finden, aus der sich alle anderen automatisch ergeben. Die übliche Liste hierfür nennen wir die „Körperaxiome“, wobei „Körper“ der abstrakte Begriff für Zahlbereiche ist in denen die 4 Grundrechenarten verfügbar sind. Sie werden gleich merken, dass der erste Kniff ist, sich zu erinnern, dass sich schon bei den Brüchen $-$ aus $+$ und $:$ aus $\cdot$ ergeben hat.

KURZER EINSCHUB ZU ZWEI GRUNDBEGRIFFEN: Wir haben bereits den Begriff der *Menge* verwendet. Als umgangssprachliche Erklärung für diesen Begriff ist noch immer die Beschreibung von Cantor üblich, dass eine Menge „die Zusammenfassung von bestimmten, wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens ist“. Die Objekte die zu einer Menge M gehören, heißen die *Elemente* der Menge, die Notation $m \in M$ bedeutet, dass m ein Element von M ist und $m \notin M$, dass m kein Element von M ist.

Der Mengenbegriff lässt sich weiter formalisieren, aber das möchte ich hier nicht tun.

Die wesentliche Eigenschaften der Definition einer Menge M sind, dass

- prinzipiell geklärt ist, welche Objekte Elemente von M sind und welche nicht und
- das Wort „wohlunterschieden“ verlangt, dass wir von zwei Elementen sagen können, ob diese gleich, oder verschieden sind.

Zum Beispiel werden Sie in der Analysis klären, dass in den reellen Zahlen $0,99999\cdots = 1$ gilt. Die Anforderung zu klären, welche Elemente gleich sind, ist also nicht immer offensichtlich.

Auf Nachfrage hatten wir in der Vorlesung zwei Gründe gesehen:

1. Auch in $\mathbb{R}$ sollte $3 \cdot \frac{1}{3} = 1$ gelten. Dort schreiben wir aber $\frac{1}{3} = 0,333333\ldots$ und berechenen $1 = 3 \cdot (0,33333\ldots) = 0,99999\ldots$.
2. Für die endliche Dezimalzahlen gilt $1 - 0,999\ldots99 = 0,000\ldots01 = \frac{1}{10^n}$, die Differenz $1 - 0,9999999\ldots$ muss also kleiner als jede positive Zahl sein.

Sie haben schon viele Besipiele von Mengen gesehen, in dieser Vorlesung hatten wir schon Löungsmengen und die Zahlbereiche $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ gesehen.

Rechenoperationen wie $+$ sind Vorschriften, die aus zwei Zahlen eine neue Zahl machen. Das wird üblicherweise als Abbildung:

$$+ \colon M \times M \to M$$

aufgeschrieben.

Das ist schnell erklärt: Eine *Abbildung* $f \colon M \to N$ von einer Menge M in eine andere Menge N ist eine Vorschrift, die jedem Element $m \in M$ ein Element $f(m) \in N$ zuordnet.

In unserem Beispiel möchten wir zwei Zahlen ihre Summe zuordnen also aus 2 Elementen ein neues Element machen. Darum führen wir die Notation

$$M \times M := \{\, (a,b) \mid a \in M \text{ und } b \in M \,\}$$

für die Menge aller Paare von Elementen von M ein.

Die Verknüpfung $+$ ist damit eine Abbildung

$$\begin{aligned} + \colon \mathbb{R} \times \mathbb{R} &\to \mathbb{R} \\ (a,b) &\mapsto a + b. \end{aligned}$$

Der Pfeil „$\mapsto$“ bedeutet „wird abgebildet auf“.

Statt einen Begriff einer Funktion in mehreren Veränderlichen einzuführen, haben wir hier einfach Paare von Elementen als Elemente einer neuen Menge eingeführt. Das ist eine technische Kleinigkeit, aber wie bei den Grundrechenarten, ist es auch bei der Entwicklung der Grundlagen praktisch, mit den Begriffen sparsam umzugehen.

Mit diesen Grunbegriffen können wir jetzt die versprochene minimale Liste von Eigenschaften aufschreiben, die ein Zahlbereich haben sollte.

Definition 3. Ein *Körper* ist eine Menge K, zusammen mit zwei Verknüpfungen:

$$\begin{aligned} + \colon K \times K &\to K \\ (a,b) &\mapsto a + b \\ \cdot \colon K \times K &\to K \\ (a,b) &\mapsto a \cdot b, \end{aligned}$$

die folgende Eigenschaften erfüllen:

Ao (neutrales Element für +) Es gibt ein Element $0 \in K$ so dass für alle $a \in K$ gilt, dass $a + 0 = a$.

Jeder Zahlbereich sollte eine 0 enthalten.

AKom (Kommutativgesetz) Für alle $a, b \in K$ gilt $a + b = b + a$.

AAss (Assoziativgesetz) Für alle $a, b, c \in K$ gilt $(a + b) + c = a + (b + c)$.

Das war der Grund, warum wir in Rechnungen mit $+$ keine Klammern setzen müssen.

AInv (inverse Elemente) Zu jedem Element $a \in K$ existiert ein Element $-a \in K$ so dass $a + (-a) = 0$.

Hier ist $-$ versteckt.

M1 (neutrales Element für $\cdot$) Es gibt ein Element $1 \in K \setminus \{0\}$ so dass für alle $a \in K$ gilt, dass $1 \cdot a = a$.

Jeder Zahlbereich sollte eine 1 enthalten, die Bedingung $1 \neq 0$ hätten Sie wahrscheinlich zunächst vergessen.

MKom (Kommutativgesetz) Für alle $a, b \in K$ gilt $a \cdot b = b \cdot a$.

MAss (Assoziativgesetz) Für alle $a, b, c \in K$ gilt $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.

Das war der Grund, warum wir in Rechnungen mit $\cdot$ keine Klammern setzen müssen.

MInv (inverse Elemente) Zu jedem Element $a \in K$ mit $a \neq 0$ existiert ein Element $a^{-1} \in K$ so dass $a \cdot (a^{-1}) = 1$.

Hier ist : versteckt.

Dist (Distributivgesetz) Für alle $a, b, c \in K$ gilt $a \cdot (b + c) = a \cdot b + a \cdot c$.

FAST MÖCHTE ICH MICH für eine so lange Definition entschuldigen, aber ich hoffe, dass Sie umgekehrt eher überrascht sind, dass wir schon fertig sind.

DIE ERSTE GUTE NACHRICHT ist, dass es nicht so schwer ist, diese Liste von Eigenschaften für die komplexen Zahlen $\mathbb{C}$ zu prüfen, für die rationalen Zahlen $\mathbb{Q}$ und die reellen Zahlen $\mathbb{R}$ kennen Sie das ohnehin.

Behauptung 4. *Die komplexen Zahlen $\mathbb{C}$ mit den Verknüpfungen $+$ und $\cdot$ sind ein Körper.*

Beweis. Das ist recht schnell geprüft:Die 0 erfüllt

$$(a + ib) + 0 = (a + 0 + (b + 0)i) = a + ib,$$

genauso gilt $1 \cdot (a + ib) = a + ib$. Kommutativität folgt aus der Definition von $+$ und $\cdot$ weil diese Rechenregeln in $\mathbb{R}$ gelten. Das gleiche gilt für das AAss. Die Assoziativität der Multiplikation ist eine kleine Fleißaufgabe (von der wir uns noch überlegen werden, wieso wir die mit Nachdenken vermeiden können):

$$\begin{aligned}
\big((a+ib)(a'+ib')\big)(a''+ib'') &= (aa' - bb' + i(ab' + ba'))(a''+ib'') \\
&= (aa'a'' - bb'a'' - ab'b'' - ba'b'') \\
&\quad + i(aa'b'' - bb'b'' + ab'a'' + ba'a'') \\
(a+ib)\big((a'+ib')(a''+ib'')\big) &= (a+ib)\big(a'a'' - b'b'' + i(a'b'' + b'a'')\big) \\
&= (aa'a'' - ab'b'' - ba'b'' - bb'a'') \\
&\quad + i(aa'b'' + ab'a'' + ba'a'' - bb'b'')
\end{aligned}$$

und die beiden Ergebnisse stimmen tatsächlich überein.

Das Distributivgesetz rechnen Sie bitte selbst nach. □

Aus den zwei Verknüpfungen $+$ und $\cdot$ können wir wegen Bedingung AInv die Subtraktion $-$ und wegen MInv auch die Division : wie folgt erklären: Da es zu jedem $a \in K$ wie in AInv festgehalten ein Element $(-a) \in K$ gibt, können wir

$$\begin{aligned} - \colon K \times K &\to K \\ (b, a) &\mapsto b - a := b + (-a) \end{aligned}$$

„:=“ bedeutet „*wird definiert als*“, d.h. der Ausdruck auf der Seite des Doppelpunktes wird durch den Term auf der anderen Seite definiert.

und da es zu jedem $a \in K \setminus \{0\}$ ein a^{-1} wie in MInv gibt

$$\begin{aligned} : \colon K \times (K \setminus \{0\}) &\to K \\ (b, a) &\mapsto b : a := b \cdot (a^{-1}) \end{aligned}$$

$\setminus$ ist das Symbol mit dem wir eine Teilmenge aus einer Menge entfernen, also das Subtraktionssymbol für Mengen.

definieren.

Halt! Wenn Sie genau hinschauen, sehen Sie vielleicht, dass ich gerade geschummelt habe.

Sehen Sie wo?

In der Mathematik sollten Sie keinem Argument leichtfertig glauben.

Außerdem dürfen Sie sich daran gewöhnen, sehr genau hinzuschauen und nachzufragen.

Kleine Fehler in den Grundlagen, können später relativ große Resultate zum Einsturz bringen. Gerade wenn Sie Argumente später in Situationen anwenden möchten, an die Sie ursprünglich nicht gedacht hatten, ist es wichtig, ganz sicher zu sein, dass Sie wirklich alle Voraussetzungen für Ihre Schlüsse geprüft haben.

In der Definition hatten wir nur verlangt, dass es zu jedem Element a ein anderes Element $-a$ bzw a^{-1} gibt, das die gewünschte Eigenschaft hat. Das sagt aber nichts darüber, ob es davon vielleicht mehrere gibt.[9]

Das wäre schlecht, weil mit unterschiedlichen $(-a), (-a')$ mit $a + (-a) = 0 = a + (-a')$ nicht gleich klar ist, warum dann $b + (-a) = b + (-a')$ gelten sollte. Damit $-$ eine Abbildung wird, muss das geklärt sein.

[9] Das ist keine esoterische Frage: Bei den komplexen Zahlen hatten wir zum Beispiel $\frac{1}{c+id}$ mit einem Trick berechnet und es ist zunächst nicht so klar, ob Ihnen vielleicht noch ein anderer Trick mit einem anders aussehenden Ergebnis einfällt.

Behauptung 5. *In jedem Körper K sind die inversen Elemente $(-a)$ und a^{-1} aus den Axiomen AInv und MInv eindeutig bestimmt.*

Argumente dieser Art sind gewöhnungsbedürftig und brauchen manchmal einen Trick, Sie werden das aber schnell lernen. Mir scheint es in jedem Fall einfacher, mir das allgemein zu überlegen, als die Gleichungen in $\mathbb{C}$ nachzurechnen.

Beweis. Wir müssen zeigen: Ist $a \in K$ und sind $b, c \in K$ zwei[10] Elemente so dass $a + b = 0$ und $a + c = 0$ gilt, so ist $b = c$.

[10] Es ist sehr hilfreich, die zwei potenziellen inversen mit unterschiedlichen Buchstaben zu bezeichnen, damit die Frage weniger verwirrend aussieht.

Wie fange ich an? Eigentlich steht in $a + b = 0$, $a - a' = 0$, subtrahieren wir auf beiden Seiten a, bekommen wir $-a' = -a$. Das Problem ist nur, dass wir $-$ gerade definieren möchten und darum $-a'$ durch $+b$ ersetzen müssen. Dann sieht das so aus:

$$
\begin{aligned}
a + b = 0 \quad & | + c \text{ (brauche } c \text{ in der Gleichung)} \\
\Rightarrow (a + b) + c = 0 + c \quad & \text{benutze } c + 0 = c \text{ und kommutativ} \\
\Rightarrow (b + a) + c = c \quad & \text{benutze das Assoziativgesetz} \\
\Rightarrow b + (a + c) = c \quad & \text{gewonnen: Kann jetzt } a + c = 0 \text{ verwenden!} \\
\Rightarrow b + 0 = c & \\
\Rightarrow b = c. &
\end{aligned}
$$

„$\Rightarrow$“ ist die Kurzform für „daraus folgt“.

Also gilt $b = c$.

Führen Sie ein ähnliches Argument für (a^{-1}) jetzt selbst einmal durch. □

Eine andere Möglichkeit, auf einen Ansatz zu kommen, ist sich ganz formal zu überlegen, dass Sie beide Gleichungen zusammen benutzen möchten. Fangen Sie mit einer von beiden an und schauen Sie, welche Umformungen helfen könnten, die zweite einzusetzen.

Hier noch ein Beispiel für eine Rechenregel, die in jedem Körper gilt.

Behauptung 6. *In jedem Körper gilt „Minus von Minus ist Plus“, d.h. für alle $a \in K$ gilt*

$$-(-a) = a.$$

Beweis. (Lange Version mit Erklärung wie ich vorgehe.)
Für jedes $b \in K$ ist $-b$ das (nach dem vorigen Resultat eindeutig bestimmte) Element von K für das die Gleichung

Definition von $-$etwas nachgeschaut, weil die merkwürdig war.

$$b + (-b) = 0$$

gilt.

Um $-(-a) = a$ zu zeigen, müssen wir uns also überlegen, dass wenn wir $b = (-a)$ setzen, die Gleichung $b + a = 0$ gilt.

Setze die Definition von „$-$etwas“ in die Aussage der Behauptung ein.

Das stimmt aber, denn

$$(-a) + a \overset{AKomm}{=} a + (-a) \overset{AInv}{=} 0,$$

weil

- $(-a) + a = a + (-a)$ wegen des Kommutativgesetzes für $+$ gilt und
- $a + (-a) = 0$ gilt weil $-a$ das additive Inverse von a war.

□

Beweis. (Kurze Version) Wir wissen dass gilt

Das Argument ist das gleiche, nur ist es auf eine gewöhnungsbedürftige Art aufgeschrieben.

$$\begin{aligned} (-a) + a &= a + (-a) && + \text{ ist kommutativ} \\ &= 0 && \text{verwende Eigenschaft AInv}v \end{aligned}$$

Also erfüllt a die Bedingung AInv für das Element $-a$. □

In den Übungsaufgaben können Sie jetzt nachprüfen, dass die üblichen Rechenregeln tatsächlich in jedem Körper gelten und also insbesondere in $\mathbb{C}$. Ich hoffe, dass Sie sich nach dem Schreibaufwand für das Assoziativgesetz in $\mathbb{C}$ dabei auch davon überzeugen können, dass die allgemeinen Argumente sehr viel angenehmer sind, als wenn Sie diese Dinge direkt mit der Definition der Verknüpfungen in $\mathbb{C}$ nachrechnen würden.

Die rationalen Zahlen – Äquivalenzrelationen

Vorlesung 24.10.

Sie hatten relativ schnell zugestimmt, dass die rationalen Zahlen einen Körper bilden – andernfalls wäre die Definition auch schlecht. Ganz am Anfang der Definition „Ein Körper ist eine Menge K“ gibt es jedoch den kleinen Haken, dass wir die Menge der rationalen Zahlen angeben sollten. Ich hatte diese als:

$$\mathbb{Q} = \left\{ \text{Brüche } \frac{n}{m} \;\middle|\; n \in \mathbb{Z}, m \in \mathbb{N} \right\}$$

notiert.

Wenn wir die Definition als

$$\text{Menge aller Ausdrücke der Form } \frac{n}{m}$$

lesen, haben wir gerade bewiesen, dass diese Menge kein Körper sein wird, denn bei Bruchsymbolen sind multiplikativ inverse Elemente nicht eindeutig bestimmt, schließlich ist $\frac{1}{3} = \frac{2}{6}$, obwohl die Ausdrücke formal verschieden sind.

Eine Lösung hierfür wäre, nur gekürzte Brüche zuzulassen, was sofort unpraktisch ist, weil ich bei größeren Brüchen nicht sehe, ob der Bruch gekürzt ist oder nicht und schlimmer noch die Rechengesetze

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}$$

keine gekürzten Brüche liefern. Kurz, das wäre eine ganz unnatürliche Definition.

Um das in Ordnung zu bringen, sollten wir darum besser eine Methode finden, die erlaubt verschiedene Symbole als gleiche Elemente einer Menge aufzufassen.

In den Brüchen wissen wir, wie wir prüfen können, ob zwei komplizierte Brüche gleich sind. Wir berechnen einfach die Differenz und schauen ob 0 herauskommt:

$$\frac{a}{b} = \frac{c}{d} \Leftrightarrow \frac{ad-bc}{bd} = 0 \Leftrightarrow ad - bc = 0.$$

Brüche die diese Gleichung erfüllen sind also gleichwertig (=äquivalent).

$\Leftrightarrow$ bedeutet „genau dann wenn", also: Es gilt, sowohl, dass aus der linken Aussage die rechte folgt, als auch umgekehrt, dass aus der rechten Aussage, die linke folgt.

Wenn Sie eine genau-dann-wenn-Aussage begründen wollen, müssen Sie darum immer zwei Aussagen begründen: „$\Leftarrow$" und „$\Rightarrow$"!

Für die rationalen Zahlen $\mathbb{Q}$ sollten wir also einfach alle formalen Ausdrücke der Form $\frac{a}{b}$, die den gleichen Wert haben als unterschiedliche Repräsentanten des gleichen Elementes von $\mathbb{Q}$ auffassen.

Ganz ähnlich müssen wir bei den reellen Zahlen aufpassen, dass $0,9999\cdots = 1,0000\ldots$ gilt. Wenn Sie reelle Zahlen als unendliche Dezimalzahlen auffassen möchten, sollten Sie also darauf achten, dass Zahlen die auf $,\ldots\overline{9}$ (d.h eine unendliche Folge von 9) enden mit der Zahl identifiziert wird in der Sie die letzte Ziffer, die keine 9 war um eine erhöhen und dann $\overline{0}$ anfügen. Dann bekommen Sie tatsächlich einen Körper.

Das Zusammenfassen verschiedener Symbole zu einem Objekt formalisieren wir mit dem Begriff der Äquivalenzrelation (= Gleichwertigkeitsbegriff). Das ist wie bei Körpern eine minimale Liste an Anforderungen für einen Gleichwertigkeitsbegriff.

Definition 7. Eine *Relation* auf einer Menge M ist eine Teilmenge $R \subset M \times M$. Für zwei Elemente $a, b \in M$ schreiben wir

$$a \sim_R b :\Leftrightarrow (a,b) \in R.$$

Ich stelle mir $M \times M$ hier als Kreuztabelle vor, R gibt die angekreuzten Paare (a,b) an.

$\sim_R$ lese ich als „a ist in Relation zu b".

Eine *Äquivalenzrelation* auf einer Menge ist eine Relation $R \subset M \times M$ für die gilt, dass

- (reflexiv) Für alle $m \in M$ gilt $m \sim_R m$.
- (transitiv) Gilt $m \sim_R m'$ und $m' \sim_R m''$ so folgt dass $m \sim_R m''$.
- (symmetrisch) Gilt $m \sim_R m'$ so gilt auch $m' \sim_R m$.

m sollte gleichwertig mit sich selbst sein

Die Bedingungen sind hoffentlich einleuchtend, am Ende soll uns eine Äquivalenzrelation erlauben, m durch andere Symbole zu ersetzen, ohne die Bedeutung zu ändern. Die Symmetrie bedeutet einfach, dass Ersetzungen rückgängig gemacht werden dürfen, Transitivität formalisiert, dass wenn ich $\frac{2}{6}$ durch $\frac{1}{3}$ ersetzen darf und $\frac{1}{3}$ durch $\frac{5}{15}$, dann darf ich auch direkt $\frac{2}{6}$ durch $\frac{5}{15}$ ersetzen.

Die Reflexivität ist nur dafür da, festzuhalten, dass $m = m$ immer gelten soll.

Beispiel 8 (Rationale Zahlen)**.** Die Relation

$$\frac{a}{b} \sim \frac{c}{d} \Leftrightarrow ad - bc = 0$$

Wenn wir ein Konzept formalisiert haben, sollten wir testen, ob die Liste der Bedingungen auch wirklich im Beispiel funktioniert.

auf der Menge der Bruchsymbole:

$$\text{BruchSymb} = \{\frac{n}{m} \mid n \in \mathbb{Z}, m \in \mathbb{N}\}$$

ist eine Äquivalenzrelation.

ENTSCHULDIGUNG. Formal sollte ich die Relation eigentlich als Teilmenge

$$R = \left\{(\frac{a}{b}, \frac{c}{d}) \mid ad - bc = 0\right\} \subset \text{BruchSymb} \times \text{BruchSymb}$$

aufschreiben. Ich finde es aber einfacher, direkt an die Relation $\sim_R$ zu denken.

Beweis.

- (reflexiv) Für alle $\frac{n}{m} \in M$ gilt, $\frac{n}{m} \sim \frac{n}{m}$, weil in der Tat $mn - nm = 0$ gilt.
- (transitiv) Zu zeigen: Gilt $\frac{n}{m} \sim \frac{n'}{m'}$ und $\frac{n'}{m'} \sim \frac{n''}{m''}$ dann muss $\frac{n}{m} \sim \frac{n''}{m''}$ gelten, also $nm'' = n''m$:

  Zum Anfangen ist es immer nützlich, klar aufzuschreiben was wir eigentlich zeigen wollen.

  Wenn ich nicht weiß, wie ich anfangen soll, schreibe ich zunächst die Voraussetzungen aus.

  Die erste Voraussetzung $\frac{n}{m} \sim \frac{n'}{m'}$ bedeutet:

  $$nm' = n'm$$

  Die zweite Voraussetzung $\frac{n'}{m'} \sim \frac{n''}{m''}$ bedeutet:

  $$n'm'' = n''m'.$$

  So, jetzt weiß ich nicht weiter. Wie komme ich zu einem Ausdruck der Form nm''? In meinem Beispiel $\frac{2}{6} \sim \frac{1}{3} \sim \frac{5}{15}$ sehe ich das Problem auch schon, denn zum Vergleichen brauchen wir den Hauptnenner aller 3 Brüche. Die erste Gleichung weiß noch nichts vom 3. Nenner, aber es hilft, mit dem 3. Nenner zu multiplizieren.

  Das probiere ich einmal auf gut Glück im allgemeinen.

  Multiplizieren wir die erste Gleichung mit m'' (denn m''das sollte in der Gleichung die wir suchen vorkommen), dann folgt:

  $$\begin{aligned} nm'm'' &= n'mm'' && \text{Setze die 2. Gleichung ein} \\ &= n''m'm && \text{also} \\ nm'm'' &= n''m'm \end{aligned}$$

  Also gilt $m'(nm'') = m'(n''m)$ und da $m' \neq 0$ ist, folgt daraus die Behauptung:

  $$nm'' = n''m.$$

- (symmetrisch) Gilt $m \sim_R m'$ so gilt auch $m' \sim_R m$.

  Das ist einfacher als die Transitivität und ich überlasse es darum Ihnen, dies nachzuweisen.

□

MERKE:

1. Wenn ich etwas nachprüfen möchte und nicht weiß, wie ich anfangen soll, schreibe ich erst genau auf, was zu zeigen ist und dann, was die Voraussetzungen genau bedeuten.
2. Wenn ich zwischendurch nicht weiter weiß, hilft es oft auf einem Schmierzettel anzuschauen, wie wir in einem Beispiel weiterkommen.

Wie identifizieren wir jetzt gleichwertige Symbole? Die Idee ist so einfach wie genial: Statt nach einer tieferen Bedeutung der verschiedenen Ausdrücke zu suchen, fassen wir einfach die jeweils gleichwertigen Ausdrücke zu einem Objekt zusammen und betrachten die Elemente davon als die erlaubten Darstellungen für das Objekt.

Definition 9. Sei $\sim_R$ eine Äquivalenzrelation auf einer Menge M.

- Für ein Element $m \in M$ heißt die Menge

  $$[m] := \{m' \in M \mid m \sim_R m'\}$$

  die *Äquivalenzklasse* des Elementes m.

  Jedes Element $m' \in [m]$ heißt *Vertreter der Äquivalenzklasse* $[m]$.

  $[m]$ ist die Menge der gleichwertigen Symbole für m.

- Die Menge der Äquivalenzklassen:

  $$M/\sim_R := \{A \subseteq M \mid A = [m] \text{ für ein } m \in M\}$$

  heißt *Quotientenmenge* der Äquivalenzrelation.

Es gibt unterschiedliche Schrebweisen für Äquivalenzklassen, statt $[m]$ wird manchmal auch $\overline{m}$ verwendet.

Die Eigenschaften von Äquivalenzrelationen können wir in Eigenschaften der Äquivalenzklassen übersetzen, da leisten die das was wir uns wünschen:

- Reflexivität sagt $m \in [m]$, d.h. jedes Element taugt als Vertreter seiner Äquivalenzklasse.
- Symmetrie sagt $m' \in [m]$ gilt genau dann, wenn $m \in [m']$, d.h., wenn m' als Vertreter von m taugt, dann auch umgekehrt.
- Transitivität und Symmetrie zusammen besagen:

  $$\text{Aus } m' \in [m] \text{ folgt } [m'] = [m].$$

  Vielleicht ist Ihnen das bereits klar. Wenn nicht, dann sollten wir das noch nachweisen: Angenommen $m' \in [m]$. Dann bedeutet $a \in [m']$ per Definition $m' \sim_R a$ und aus $m \sim_R m' \sim_R a$ folgt wegen der Transitivität, dass $m \sim_R a$ also $a \in [m]$.

  Damit haben wir nachgewiesen, dass aus $m \sim_R m'$ folgt, dass[11] $[m'] \subseteq [m]$. Wegen der Symmetrie folgt aus $m \sim_R m'$ auch $m' \sim_R m$ und darum aus dem ersten Argument $[m] \subseteq [m']$. Zusammen folgt also $[m] = [m']$.

[11] „$\subseteq$“bedeutet „ist Teilmenge von“

Merke: Um eine Gleichheit von Mengen $M = N$ zu zeigen, müssen wir die beiden Aussagen

$M \subseteq N$ „M ist Teilmenge von N“ und

$M \supseteq N$ „N ist Teilmenge von M“

nachweisen.

Fazit: Eine Äquivalenzrelation $\sim_R$ zerlegt eine Menge in paarweise disjunkte Äquivalenzklassen. Die Quotientenmenge $M/\sim_R$ ist die Menge, mit der wir rechnen, wenn wir Elemente, die in M nur äquivalent waren, wirklich als gleich betrachten wollen.

Beispiel 10 (Rationale Zahlen)**.** Die Quotientenmenge der Menge der Bruchsymbole:

$$\text{BruchSymb} = \left\{ \frac{n}{m} \,\middle|\, n \in \mathbb{Z}, m \in \mathbb{N} \right\}$$

nach der Relation

$$\frac{a}{b} \sim \frac{c}{d} \Leftrightarrow ad - bc = 0$$

die Menge der rationalen Zahlen:

$$\mathbb{Q} = \text{BruchSymb}/\sim .$$

Bei Brüchen schreiben wir statt $[\frac{m}{n}]$ immer $\frac{m}{n}$ und denken uns den Bruchstrich als Erinnerung, dass es sich hier um Äquivalenzklassen handelt.

Bei jedem abstrakten Begriff, sollten wir uns Beispiele und Nicht-Beispiele suchen, um einerseits ein Gefühl für den Begriff zu bekommen und andererseits zu prüfen, ob das überhaupt ein nützliches Konzept ist. Wenn es nur ein Beispiel gibt, war die Begriffsbildung vielleicht eher übertriebener Aufwand.

Frage: Kennen Sie mehr Beispiele aus Mathematik oder Alltag, bei denen wir das Konzept als Äquivalenzrelation auffassen können und die Äquivalenzklassen vielleicht auch eine Bedeutung haben?

Beispiel 11 (Äquivalenzrelationen)**.** 1. Gerade/Ungerade

2. gleiche Endziffer

3. Gleiche Marke, Gleiches Modell

Frage: Was ist mit der Relation auf den reellen Zahlen

$$x \sim y \Leftrightarrow x \cdot y > 0,$$

ist das eine Äquivalenzrelation? Falls Nein, warum nicht? Falls Ja, was bedeuten die Äquivalenzklassen?

Frage: Kennen Sie Relationen, die keine Äquivalenzrelationen sind?

Beispiel 12 (Keine Äquivalenzrelationen)**.** Für Personen:

1. A kennt B oder

2. A hat B einmal die Hand gegeben oder

3. A mag B.

Vielleicht prüfen Sie einmal, welche der Axiome hier gelten, welche nicht und warum das für die Äquivalenzklassen jeweils problematisch wäre.

Vorlesung 26.10.

FÜR DIE MENGE der rationalen Zahlen:

$$\mathbb{Q} := \left\{ \frac{n}{m} \,\middle|\, n \in \mathbb{Z}, m \in \mathbb{N} \right\} / \left(\frac{a}{b} \sim \frac{c}{d} \Leftrightarrow ad - bc = 0 \right)$$

haben wir mit dem Konzept der Äquivalenzrelation jetzt eine gute Beschreibung gefunden. Um alle Körperaxiome ordentlich zu formulieren sollten wir der guten Form halber noch die Rechengesetze aufschreiben:

$$\begin{aligned} +\colon \mathbb{Q} \times \mathbb{Q} &\to \mathbb{Q} \\ \left(\frac{a}{b}, \frac{c}{d}\right) &\mapsto \frac{a}{b} + \frac{c}{d} := \frac{ad + cb}{bd} \\ \cdot\colon \mathbb{Q} \times \mathbb{Q} &\to \mathbb{Q} \\ \left(\frac{a}{b}, \frac{c}{d}\right) &\mapsto \frac{a}{b} \cdot \frac{c}{d} := \frac{ac}{bd}. \end{aligned}$$

SEHEN SIE, dass wir hier — wie bei der Definition von „$-$“ in Körpern — wieder ein kleines Problem haben?

Für die Bruchsymbole, war klar, dass das die Formeln Abbildungen definieren, aber für $\mathbb{Q}$, ist $[\frac{a}{b}] + [\frac{c}{d}] := [\frac{ad+cb}{bd}]$ eine Formel für die Äquivalenzklassen, d.h, wir dürfen die Etikette $\frac{a}{b}$ und $\frac{c}{d}$ jederzeit durch andere Elemente der jeweiligen Äquivalenzklassen ersetzen (also statt $\frac{2}{6}$ zum Beispiel $\frac{5}{15}$ schreiben). Dann ändert sich formal der Ausdruck auf der rechten Seite und wir sollten uns davon überzeugen, dass die beiden Ergebnisformeln die gleiche Äquivalenzklasse beschreiben.

Ist Ihnen das für die Bruchrechnung klar, oder sind Sie jetzt unsicher? „Lemma“ bedeutet Hilfssatz. Lemmata sind in der Regel kleine, nützliche Behauptungen, die wir gerne für etwas anderes verwenden möchten.

Hier ist es eine Hilfsbehauptung für die Aussage „$\mathbb{Q}$ ist ein Körper“.

Lemma 13. *Die Operationen $+\colon \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ und $\cdot\colon \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ sind wohldefiniert, d.h. die Restklasse der Ergebnisformel hängt nicht davon ab, welche Vertreter der Restklassen von $[\frac{a}{b}]$ und $[\frac{c}{d}]$ gewählt wurden.*

Beweis. Zu zeigen ist: Gilt $[\frac{a}{b}] = [\frac{a'}{b'}]$ und $[\frac{c}{d}] = [\frac{c'}{d'}]$, so gilt auch

$$\left[\frac{ad + cb}{bd}\right] = \left[\frac{a'd' + c'b'}{b'd'}\right] \text{ und } \left[\frac{ac}{bd}\right] = \left[\frac{a'c'}{b'd'}\right].$$

Schritt 1: Schreibe auf, was wir zeigen wollen.

Wir wissen:

Schritt 2: Schaue die Definitionen der Symbole nach.

1. $[\frac{a}{b}] = [\frac{a'}{b'}]$ bedeutet $ab' = a'b$ und
2. $[\frac{c}{d}] = [\frac{c'}{d'}]$ bedeutet $cd' = c'd$.

Wir wollen zeigen: $\left[\frac{ad+cb}{bd}\right] \stackrel{?}{=} \left[\frac{a'd'+c'b'}{b'd'}\right]$, das bedeutet

$$(ad + cb)(b'd') \stackrel{?}{=} (a'd' + c'b')(bd). \tag{1}$$

Um das zu zeigen, rechnen wir beide Seiten aus:

Schritt 3: Jetzt habe ich eine Aussage, die ich nachrechnen kann.

$$\begin{aligned} (ad + cb)(b'd') &= ab'dd' + bb'cd' && \text{(linke Seite von (1))} \\ (a'd' + c'b')(bd) &= a'bdd' + bb'c'd && \text{(rechte Seite von (1))} \end{aligned}$$

Wegen der Voraussetzung 1. und 2 sind also in der Tat beide Seiten gleich.

Schritt 4: Schreibe auf, dass wir erreicht haben, was wir zeigen wollten.

Die Behauptung für $\cdot$ ist einfacher, diese rechnen Sie als Übungsaufgabe bitte selber nach. □

Rechnen mit Restklassen

Die Beispiele von Äquivalenzrelationen „Gerade/Ungerade“ und „gleiche Endziffer“ (oder auch „3 letzten Stellen sind gleich“) beruhen auf dem selben Prinzip, denn wenn wir die ganzen Zahlen $\mathbb{Z}$ im Binärsystem aufschreiben würden, wäre die Endziffer die Zahl, die entscheidet ob eine Zahl gerade oder ungerade ist.[12]

Sie verwenden das gleiche Prinzip beim Rechnen mit Uhrzeiten, dort rechnen Sie für die Minuten bis auf Vielfache von 60, für die Stunden bis auf Vielfache von 12.

Lassen Sie uns das einmal allgemein formulieren:

[12] Mit der Wahl des Zahlsystems entscheiden Sie, welche Teilbarkeitsregeln einfach zu sehen sind, in Dezimalzahlen ist die Teilbarkeit durch $2, 5, 10$ sichtbar, aber Teilbarkeit durch 7 schwerer, im 7-er-System wäre es umgekehrt.

Beispiel 14 (Restklassen und $\mathbb{Z}/N\mathbb{Z}$)**.** Sei $N \in \mathbb{N}$ eine ganze Zahl, dann ist die Relation

$$a \sim_N b \text{ genau dann wenn } (a = b + k \cdot N \text{ für ein } k \in \mathbb{Z})$$

eine Äquivalenzrelation.

Die Bediungung „$a = b + k \cdot N$ für ein $k \in \mathbb{Z}$“ könne wir auch als „$a - b$ ist durch N teilbar)“ formulieren.

Statt $a \sim_N b$ sagen wir auch *„a ist kongruent zu b modulo N“*.

Die Äquivalenzklasse $[a]$ einer Zahl $a \in \mathbb{Z}$ ist die Menge

$$[a] = \{a + k \cdot N \,|\, k \in \mathbb{Z}\}.$$

Die Menge der Äquivalenzklassen bezeichnen wir als

$$\mathbb{Z}/N\mathbb{Z} := \mathbb{Z}/\sim_N,$$

und nennen diese die *Menge der Restklassen modulo N*.

Beweis. Den Nachweis der Eigenschaften überlasse ich Ihnen, fragen Sie nach, falls Sie mit einer der Eigenschaften Schwierigkeiten haben sollten. □

Die Menge $\mathbb{Z}/N\mathbb{Z}$ hat N Elemente:

$$\begin{aligned}\mathbb{Z}/N\mathbb{Z} &= \{[0], [1], \ldots, [N-1]\} \\ &= \{[1], [2], \ldots, [N]\} \\ &= \{[2], [3], \ldots, [N+1]\}\end{aligned}$$

Beispiel 15. Die IBAN verwendet eine Prüfziffer, die in $\mathbb{Z}/97\mathbb{Z}$ berechnet wird (Die Zahl 97 ist die größte zweistellige Primzahl). Das funktioniert so: Die IBAN ist aus Bankleitzahl und Kontonummer zusammengesetzt und sieht so aus:

$$DEXY567890123456789012$$

Das Länderkürzel wird in zwei 2stellige Zahlen übersetzt (A=10,B=11, usw., alsoDE = 12 13) die Ziffern XY werden dann so berechnet, dass

$$[5678901234567890121213\,XY] = [1] \in \mathbb{Z}/97\mathbb{Z}$$

gilt.

Hierbei wird XY in $\{02, 03, \ldots, 98\}$ gewählt.

Mit Restklassen können wir rechnen, genau wie Sie mit Endziffern rechnen können:

Behauptung 16. *Sei $N \geq 2$ eine ganz Zahl. Dann sind die Verknüpfungen*

$$\begin{aligned} +\colon \mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z} &\to \mathbb{Z}/N\mathbb{Z} \\ ([a],[b]) &\mapsto [a]+[b] := [a+b] \qquad \textit{und} \\ \cdot\colon \mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z} &\to \mathbb{Z}/N\mathbb{Z} \\ ([a],[b]) &\mapsto [a]\cdot[b] := [a\cdot b] \end{aligned}$$

wohldefiniert und erfüllen alle Rechenregeln, die für die ganzen Zahlen gelten, d.h. bis auf die Existenz von multiplikativen Inversen gelten die Körperaxiome für diese Verknüpfungen.

Beweis. Das ist einfach: Zu zeigen ist: Wenn $[a] = [a']$ und $[b] = [b']$ gilt, dann ist

$$[a+b] = [a'+b'] \text{ und } [a\cdot b] = [a'\cdot b'].$$

Die Bedingung $[a] = [a']$ bedeutet $a' = a + k\cdot N$ für ein $k \in \mathbb{Z}$ und $[b] = [b']$ bedeutet $b' = b + \ell\cdot N$ für ein $\ell \in \mathbb{Z}$.

Dann gilt aber

$$[a'+b'] = [a + k\cdot N + b + \ell\cdot N] = [a+b+(k+\ell)\cdot N] = [a+b]$$

und genauso:

$$\begin{aligned} [a'\cdot b'] &= [(a+k\cdot N)\cdot(b+\ell\cdot N)] \\ &= [a\cdot b + (a\ell + kb + k\ell N)\cdot N] \\ &= [a\cdot b]. \end{aligned}$$

Die Rechengesetze die für $+$ und $\cdot$ für ganze Zahlen gelten übertragen sich automatisch, denn die Formeln für die Verknüpfungen wurde einfach durch hinzufügen von $[\ \ ]$ erhalten, z.B.

$$a+b = b+a \Rightarrow [a+b] = [b+a].$$

□

Beispiel 17 (Rechnen mit Wochentagen)**.** Wenn wir die Wochentage von 1 bis 7 durchnummerieren können wir die mit den Elementen von $\mathbb{Z}/7\mathbb{Z} = \{[1],[2],\dots,[N]\}$. identifizieren.

Das erklärt vielleicht wie wir mit Wochentagen rechnen:

FRAGE: Heute ist Mittwoch der 25.10. welcher Wochentag ist der 25.11?

Der Oktober hat 31 Tage, $31 = 28+3$ also $[31] = [3] \in \mathbb{Z}/7\mathbb{Z}$. Der 25.11. ist also 3 Wochentage nach Mittwoch, also ein Samstag.

Bemerkung. Die Menge $\mathbb{Z}/7\mathbb{Z}$ ist mit den Verknüpfungen $+, \cdot$ ein Körper, $\mathbb{Z}/10\mathbb{Z}$ nicht.

Beweis. Für $\mathbb{Z}/7\mathbb{Z}$ müssen wir nur noch zeigen, dass jedes Element $[a] \neq [0]$ ein multiplikativ inverses Element besitzt. Das können wir entweder ausschreiben:

$$[1] \cdot [1] = [1], [2] \cdot [4] = [8] = [1]$$

$$[3] \cdot [5] = [15] = [1], [6] \cdot [6] = [-1] \cdot [-1] = [1],$$

oder überlegen uns, warum das funktioniert bzw. wie wir die Inversen berechnen können. (Das machen wir später, wir benötigen dafür nur, dass 7 eine Primzahl ist.)

Für $\mathbb{Z}/10\mathbb{Z}$ hatten Sie das Argument gegeben, dass $[5]$ kein inverses Element haben kann, da $n \cdot 5]$ immer auf 0 oder 5 endet, und darum $[n \cdot 5] \neq [1] \in \mathbb{Z}/10\mathbb{Z}$ für alle $n \in \mathbb{Z}$.

Mein Lieblingsargument ist, dass $[5] \cdot [2] = [10] = [0] \in \mathbb{Z}/10\mathbb{Z}$, aber in jedem Körper ein Produkt nur dann 0 ist, wenn einer der Faktoren 0 ist. (Das hatten Sie in den Übungen gezeigt.) □

Behauptung 18. *Die Menge $\mathbb{Z}/N\mathbb{Z}$ ist zusammen mit $+, \cdot$ genau dann ein Körper, wenn N eine Primzahl ist.*

Das überlegen wir uns am Montag.

WIR HABEN UNS SCHON ÜBERLEGT, dass wenn $N = a \cdot b$ keine Primzahl ist, die Gleichung

$$[a] \cdot [b] = [a \cdot b] = [N] = [0]$$

zeigt, dass $\mathbb{Z}/N\mathbb{Z}$ kein Körper sein kann, da in Körpern ein Produkt nur dann 0 ist, wenn einer der Faktoren 0 ist.

Damit haben wir eingesehen, dass wenn $\mathbb{Z}/N\mathbb{Z}$ ein Körper ist, N eine Primzahl sein muss.

FÜR DIE UMKEHRUNG („$N = p$ Primzahl $\Rightarrow \mathbb{Z}/N\mathbb{Z}$ Körper“) wäre es praktisch, wenn wir für eine Primzahl p zu jedem $[a] \neq [0] \in \mathbb{Z}/p\mathbb{Z}$ ein multiplikativ inverses $[a]^{-1}$ berechnen könnten.

Der euklidische Algorithmus

Haben Sie vielleicht in der Schule gelernt, wie wir den größten gemeinsamen Teiler $\mathrm{ggT}(a, b)$ von ganzen Zahlen a, b ausrechnen können? Das geht so:

Behauptung 19 (Teilen mit Rest)**.** *Sind $a, b \in \mathbb{Z}$ ganze Zahlen mit $b > 0$, so lässt sich a eindeutig mit Rest durch b teilen, d.h. es existieren eindeutige Zahlen $q, Rest \in \mathbb{Z}$ mit $0 \leq Rest < b$ so dass*

$$a = q \cdot b + Rest.$$

Für diese Zahlen gilt dann

- $\mathrm{ggT}(a, b) = \mathrm{ggT}(b, Rest)$ *und*
- $Rest = 1 \cdot a - q \cdot b$ *lässt sich als ganzzahligen linearen Ausdruck in a, b schreiben.*

Beweis. Der erste Teil ist Ihnen vielleicht klar: da $b > 0$ ist, gibt es eine größte[13] ganze Zahl q für die $q \cdot b \leq a$. Dann ist aber

$$0 \leq a - q \cdot b < b$$

denn sonst würde $q + 1$ ebenfalls $q + 1 \cdot b \leq a$ erfüllen und q wäre dann nicht maximal mit dieser Eigenschaft.

Wir können also $\text{Rest} := a - q \cdot b$ wählen. Die zweite angegebene Eigenschaft haben wir dann per Definition erfüllt.

Die Eindeutigkeit der Zahlen q, Rest ist vielleicht einsichtig, der Vollständigkeit halber hier ein Argument: Hätten wir eine zweite Lösung $a = q' \cdot b + \text{Rest}'$, so wäre

$$q' \cdot b + \text{Rest}' = a = q \cdot b + \text{Rest}$$

also nach umstellen:

$$(q' - q) \cdot b = \text{Rest} - \text{Rest}'.$$

Da $0 \leq |\text{Rest} - \text{Rest}'| < b$ gilt und die linke Seite durch b teilbar ist, muss also $0 = \text{Rest} - \text{Rest}'$ gelten und darum auch $(q' - q) \cdot b = 0$, also $(q' - q) = 0$.

Die Eigenschaft $\text{ggT}(a, b) = \text{ggT}(b, \text{Rest})$ folgt, weil die Gleichung

$$a = q \cdot b + \text{Rest}$$

besagt, dass der größte gemeinsame Teiler von $(b, \text{Rest}$ auch a teilt und umgekehrt besagt die Gleichung

$$\text{Rest} = 1 \cdot a - q \cdot b$$

dass der größte gemeinsame Teiler von a, b auch den Rest teilt. □

[13] Wir wissen sogar, dass $q \in \{-|a|, -|a| + 1, \ldots, 0, 1, \ldots, |a|\}$.

Mit Hilfe dieser Beobachtung können wir für Zahlen a, b den größten gemeinsamen Teiler bestimmen und diesen in der Form

$$\text{ggT}(a, b) = n \cdot a + m \cdot b \text{ für geeignete } n, m \in \mathbb{Z}$$

schreiben. Das würde unser Problem lösen, denn ist p eine Primzahl und $0 < a < p$ ein Vertreter einer Restklasse $\neq 0$ in $\mathbb{Z}/N\mathbb{Z}$, so ist $\text{ggT}(a, p) = 1$ (da p nur durch 1 und p teilbar ist, a aber nicht durch p) und wenn wir dann

$$1 = n \cdot a + m \cdot p$$

geschrieben haben, gilt

$$[1] = [n] \cdot [a] \text{ in } \mathbb{Z}/p\mathbb{Z}.$$

Den $\text{ggT}(a, b)$ können wir zusammen mit der Darstellung $\text{ggT}(a, b) = n \cdot a + m \cdot b$ wie folgt berechnen:

Wegen der Behauptung ist $\text{ggT}(a, b) = \text{ggT}(b, \text{Rest})$ und Rest ist $< b$. Wenn wir also induktiv:

$$\begin{aligned} a &= q_1 b + r_1 \text{ mit } 0 \leq r_1 < b & \text{ggT}(a, b) &= \text{ggT}(b, r_1) \\ b &= q_2 r_1 + r_2 \text{ mit } 0 \leq r_2 < r_1 & \text{ggT}(b, r_1) &= \text{ggT}(r_1, r_2) \\ &\ldots \end{aligned}$$

eine absteigende Folge $b = r_0 > r_1 > r_2 > \cdots \geq 0$ nicht-negativer ganzer Zahlen, d.h. nach endlich vielen Schritten muss $r_n = 0$ gelten und die letzte Zahl $r_i \neq 0$ ist der $\mathrm{ggT}(a, b)$, denn

$$\mathrm{ggT}(a, b) = \mathrm{ggT}(b, r_1) = \cdots = \mathrm{ggT}(r_i, 0) = r_i.$$

Durch rückwärts-Einsetzen können wir dann

$$r_i = r_{i-2} - q_i r_{i-1}$$

schreiben, in der rechten Seite können wir dann r_{i-1} mit der vorherigen Gleichung durch $r_{i-1} = r_{i-3} - q_{i-2} r_{i-2}$ ersetzen

$$r_i = r_{i-2} - q_i(r_{i-3} - q_{i-2} r_{i-2})$$

und erhalten damit induktiv die Darstellung von $\mathrm{ggT}(a, b)$.

Das ist an einem Beispiel leichter nachvollziehbar:

Beispiel 20. Bestimme das Inverse von 11 in $\mathbb{Z}/97\mathbb{Z}$:

$$\begin{aligned} 97 &= 8 \cdot 11 + \mathbf{9} \\ 11 &= 1 \cdot 9 + \mathbf{2} \\ 9 &= 4 \cdot 2 + \mathbf{1} \end{aligned}$$

Wenn der Rest 1 ist, haben wir den ggT gefunden.

Also

$$\begin{aligned} \mathbf{1} &= 9 - 4 \cdot \mathbf{2} \\ &= 9 - 4 \cdot (11 - 1 \cdot 9) \\ &= 5 \cdot \mathbf{9} - 4 \cdot 11 \\ &= 5 \cdot (97 - 8 \cdot 11) - 4 \cdot 11 \\ &= -44 \cdot 11 + 5 \cdot 97. \end{aligned}$$

Also ist $[11]^{-1} = [-44] = [53] \in \mathbb{Z}/97\mathbb{Z}$.

Probe: $[11] \cdot [53] = [530 + 53] = [583] = [582 + 1] = [97 \cdot 6 + 1] = [1]$.

Insbesondere haben wir jetzt gesehen, dass falls p eine Primzahl ist, für jede Zahl $a \in \mathbb{Z}$, die nicht durch p teilbar ist, wir eine Zahl $b \in \mathbb{Z}$ berechnen können für die

$$[a][b] = [1] \in \mathbb{Z}/p\mathbb{Z}.$$

Insbesondere ist $\mathbb{Z}/p\mathbb{Z}$ ein Körper.

Folgerung 21. *Teilt eine Primzahl p ein Produkt, so teilt p einen der Faktoren.*

Beweis. Nach Definition teilt p eine Zahl n genau dann, wenn $[n] = [0] \in \mathbb{Z}/p\mathbb{Z}$ gilt.

Also teilt p ein Produkt $a \cdot b$, genau dann wenn $[a][b] = [0] \in \mathbb{Z}/p\mathbb{Z}$. In einem Körper ist ein Produkt aber genau dann 0 wenn einer der Faktoren 0 ist. □

Hiermit haben Sie nun einige neue Zahlbereiche kennengelernt. Lassen Sie uns nun zu linearen Gleichungen zurückkehren.

Lineare Gleichungen und Matrizen als Abbildungen

Wir hatten gesehen, dass wir ein lineares Gleichungssystem

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

Mit der Schreibweise

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ \vdots & & & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}, \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ und } \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

kompakt als eine Gleichung

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

schreiben können.

Die Vektorschreibweise und der K^n

Für die Koeffizienten können wir nun Elemente eines beliebigen Körpers zulassen. In den reellen Zahlen hatten wir Lösungen **x** als Elemente des

$$\mathbb{R}^n = \{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mid x_1, \dots, x_n \in \mathbb{R} \}$$

aufgefasst. Hierbei stelle ich mir $\mathbb{R}^2$ als die Punkte in der Ebene und $\mathbb{R}^3$ als 3-dimensionalen Raum mit einem ausgezeichneten 0-Punkt vor.

Genauso werden wir für jeden Körper K die Lösungen **x** als Elemente der Menge

$$K^n = \{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mid x_1, \dots, x_n \in K \}$$

auffassen, die Elemente von K^n heißen *Vektoren*.

Zum Beispiel stelle ich mir die Elemente von $\mathbb{Z}/2\mathbb{Z}^3$ als die Ecken des Würfels mit Kantenlänge 1 vor. Mein Bild von $\mathbb{Q}^n$ ist von meiner Vorstellung von $\mathbb{R}^n$ kaum zu unterscheiden, bei $\mathbb{C}^n$ denke ich manchmal an einen $\mathbb{R}^{2n}$, in dem die Koordinaten paarweise zusammengefasst wurden.

Entsprechend zu $\mathbf{x} \in K^n$ ist $\mathbf{b} \in K^m$.

Bemerkung. Eine $m \times n$-Matrix ist eine 2-dimensionale Liste von Einträgen, das wird manchmal als $\mathbb{R}^{m\times n} = \mathbb{R}^m \otimes \mathbb{R}^n$ geschrieben. Wenn Sie Datenmengen beschreiben oder analysieren möchten, kommen öfters höherdimensionale Listen vor, Sie können sich sicher $(a_{i,j,k})_{i,j,k=0,\dots,n}$ als Würfel-Matrix vorstellen. Solche mehrdimensionale Listen heißen auch *Tensoren* – dies nur für den Fall, dass Sie einmal über das Wort stolpern sollten.

Vorlesung 2.11.

GENAUSO wie bei den komplexen Zahlen, können wir Elemente von K^n komponentenweise addieren

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix} = \begin{pmatrix} x_1 + x_1' \\ x_2 + x_2' \\ \vdots \\ x_n + x_n' \end{pmatrix}.$$

Das entspricht geometrisch wieder dem aneinanderlegen von Ortsvektoren im $\mathbb{R}^n$.

Die Streckung eines Ortsvektors um den Faktor $c \in K$ ist durch

$$c \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{pmatrix}$$

gegeben.

Als Verknüpfungen könnten wir das formal wieder als

$$\begin{aligned} +\colon K^n \times K^n &\to K^n \\ (\mathbf{x}, \mathbf{x}') &\mapsto \mathbf{x} + \mathbf{x}' \qquad \text{und} \\ \cdot\colon K \times K^n &\to K^n \\ (c, \mathbf{x}) &\mapsto c \cdot \mathbf{x} := \begin{pmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{pmatrix} \end{aligned}$$

schreiben. Hier verknüpft $\cdot$ Elemente aus unterschiedlichen Mengen (ein Element aus dem Körper und einen Vektor). Das Konzept der Verknüpfung als Abbildung ist so flexibel, dass wir damit auch solche Operationen formalisieren können.

Lösungsmengen und der Kern einer Matrix

Der Gauß-Algorithmus erlaubt uns, die Lösungsmenge von Gleichungssystemen $Ax = b$ zu bestimmen. Wenn wir für die gleiche Matrix die Lösungen für unterschiedliche b suchen, wird es lästig, das jedes Mal von neuem anzufangen.

FRAGE:

1. Können wir bestimmen für welche b das Gleichungssystem $Ax = b$ lösbar ist?
2. Wie hängen die Lösungsmengen von $Ax = b$ und $Ax = b'$ zusammen?

BEOBACHTUNG: Sind $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ und $\mathbf{x}' = \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}$ Lösungen von $A\mathbf{x} = \mathbf{b}$ und $A\mathbf{x} = \mathbf{b}'$, so gilt für $\mathbf{x} + \mathbf{x}'$, dass

$$A(\mathbf{x} + \mathbf{x}') = \mathbf{b} + \mathbf{b}', \text{ d.h.}$$

$$A \cdot (\mathbf{x} + \mathbf{x}') = A \cdot \mathbf{x} + A \cdot \mathbf{x}'. \tag{2}$$

DENN hierbei haben wir einfach nur die Gleichungssysteme

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

und

$$\begin{array}{rcl} a_{1,1}x'_1 + a_{1,2}x'_2 + \cdots + a_{1,n}x'_n & = & b'_1 \\ a_{2,1}x'_1 + a_{2,2}x'_2 + \cdots + a_{2,n}x'_n & = & b'_2 \\ \vdots & & \vdots \\ a_{m,1}x'_1 + a_{m,2}x'_2 + \cdots + a_{m,n}x'_n & = & b'_m \end{array}$$

addiert, addieren wir nämlich die jeweils i-te Gleichung erhalten wir:

$$\begin{array}{rccccccc} & a_{i,1}x_1 & + & a_{i,2}x_2 & +\cdots+ & a_{i,n}x_n & = & b_i \\ + & a_{i,1}x'_1 & + & a_{i,2}x'_2 & +\cdots+ & a_{i,n}x'_n & = & b'_i \\ \hline & a_{i,1}(x_1 + x'_1) & + & a_{i,2}(x_2 + x'_2) & +\cdots+ & a_{i,n}(x_n + x'_n) & = & b_i + b'_i \end{array}$$

und das ist die i-te Zeile von $A(\mathbf{x} + \mathbf{x}') = \mathbf{b} + \mathbf{b}'$.

GENAUSO können wir das erste Gleichungssystem mit einer Zahl c multiplizieren und erhalten:

$$A \cdot \mathbf{x} = \mathbf{b} \Rightarrow A \cdot (c\mathbf{x}) = c \cdot \mathbf{b}, \text{ d.h.}$$

$$A \cdot (c\mathbf{x}) = c \cdot (A \cdot \mathbf{x}), \tag{3}$$

da genauso

$$\begin{array}{rccccccl} & a_{i,1}x_1 & + & a_{i,2}x_2 & +\cdots+ & a_{i,n}x_n & = & b_i & |c\cdot \\ \Rightarrow & a_{i,1}(c\cdot x_1) & + & a_{i,2}(c\cdot x_2) & +\cdots+ & a_{i,n}(c\cdot x_n) & = & c\cdot b_i & \end{array}$$

Folgerung 22. *Je zwei Lösungen des Gleichungssystems $A\cdot\mathbf{x}=\mathbf{b}$ unterscheiden sich um eine Lösung von $A\cdot\mathbf{x}=0$.*

Beweis. Ist $A\cdot\mathbf{x}=\mathbf{b}$ und $A\cdot\mathbf{x}'=\mathbf{b}$ so gilt

$$A\cdot(\mathbf{x}-\mathbf{x}')=\mathbf{b}-\mathbf{b}=\mathbf{0}.$$

Ist umgekehrt $A\cdot\mathbf{x}=\mathbf{b}$ und $A\cdot\mathbf{z}=\mathbf{0}$, so gilt

$$A\cdot(\mathbf{x}+\mathbf{z})=\mathbf{b}+\mathbf{0}=\mathbf{b}.$$

□

Definition. Ist A eine $m\times n$-Matrix mit Koeffizienten in K, so heißt die Menge

$$\mathrm{Ker}(A):=\{\mathbf{z}\in K^n | A\cdot\mathbf{z}=0\}$$

Kern der Matrix A.

Damit haben wir die 2. Frage beantwortet:

Folgerung 23. *Ist für eine $n\times m$-Matrix A das Gleichungssystem $Ax=b$ lösbar und ist $v\in K^n$ eine Lösung dieser Gleichung, so ist die Lösungsmenge*

$$\mathit{Lösung}(Ax=b)=\{v+z\,|\,z\in\mathrm{Ker}(A)\}.$$

Insbesondere gibt es nur dann eindeutige Lösungen wenn $\mathrm{Ker}(A)=\{\mathbf{0}\}$.

Wir schreiben oft $\mathbf{0}=\begin{pmatrix}0\\0\\\vdots\\0\end{pmatrix}$. Die Notation 0 wird für sehr viele unterschiedliche 0-Elemente verwendet.

Matrizen als Abbildungen und das Bild

Für die erste Frage könnten wir einfach alle $\mathbf{x}$ in A einsetzen und schauen welche Ergebnisse wir erhalten, d.h. wir fassen die Matrix A als Abbildung

$$\begin{aligned} A\colon K^n &\to K^m \\ \mathbf{x} &\mapsto A\cdot\mathbf{x} \end{aligned}$$

auf, zum Beispiel definiert die Matrix $A=\begin{pmatrix}1&2&3\\4&5&6\end{pmatrix}$ die Abbildung

$$\begin{aligned} A\colon K^3 &\to K^2 \\ \mathbf{x} &\mapsto A\cdot\mathbf{x}, \end{aligned}$$

die zum Beispiel $\begin{pmatrix}1\\1\\1\end{pmatrix}$ auf

$$\begin{pmatrix}1&2&3\\4&5&6\end{pmatrix}\cdot\begin{pmatrix}1\\1\\1\end{pmatrix}=\begin{pmatrix}1+2+3\\4+5+6\end{pmatrix}=\begin{pmatrix}6\\15\end{pmatrix}$$

abbildet.

Die $\mathbf{b}$ zu finden, für die $A\mathbf{x} = \mathbf{b}$ lösbar ist, ist also das gleiche Problem, wie das Bild der Abbildung

$$\text{Bild}(A) := \{\mathbf{b} \in K^m \mid \text{es gibt ein } \mathbf{x} \in K^n \text{ mit } A \cdot \mathbf{x} = \mathbf{b}\}$$

zu bestimmen.

WIR SOLLTEN EIN BEISPIEL ausrechnen, um ein Gefühl für diese Menge zu bekommen. Sei

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 2 & 3 & 5 \\ 1 & 1 & 2 \end{pmatrix}$$

welche $\mathbf{b}$ sind im Bild von A?

$$\begin{aligned} A \cdot \mathbf{x} &= \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 2 & 3 & 5 \\ 1 & 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \\ &= \begin{pmatrix} 1 \cdot x_1 + & 0 \cdot x_2 + & 1 \cdot x_3 \\ 0 \cdot x_1 + & 1 \cdot x_2 + & 1 \cdot x_3 \\ 2 \cdot x_1 + & 3 \cdot x_2 + & 5 \cdot x_3 \\ 1 \cdot x_1 + & 1 \cdot x_2 + & 2 \cdot x_3 \end{pmatrix} \\ &= x_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \end{pmatrix} + x_2 \cdot \begin{pmatrix} 0 \\ 1 \\ 3 \\ 1 \end{pmatrix} + x_3 \cdot \begin{pmatrix} 1 \\ 1 \\ 5 \\ 2 \end{pmatrix}. \end{aligned}$$

FRAGE: Fällt Ihnen etwas auf?

IHRE ANTWORT WAR: Der letzte Vektor ist die Summe der beiden ersten:

$$\begin{pmatrix} 1 \\ 1 \\ 5 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \\ 3 \\ 1 \end{pmatrix}.$$

FRAGE: Können Sie damit entscheiden, ob ein zufällig gewählter Vektor, zum Beispiel $\begin{pmatrix} 7 \\ 5 \\ 3 \\ 2 \end{pmatrix}$ im Bild von A liegt?

IHRE ANTWORT WAR: Der Vektor liegt nicht im Bild. Da der letzte Vektor ist die Summe der beiden ersten ist, können wir den für die Beschreibung des Bildes vergessen, denn wir können eine Summe in der dieser Vektor vorkommt durch eine Summe in der nur die ersten beiden Vektoren vorkommen ersetzen.

Dann ist aber

$$x_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \end{pmatrix} + x_2 \cdot \begin{pmatrix} 0 \\ 1 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ 2x_1 + 3x_2 \\ x_1 + x2 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 7 \\ 5 \\ 3 \\ 2 \end{pmatrix}.$$

Das ist nicht lösbar, da die ersten Zeilen $x_1 = 7, x_2 = 5$ bedeuten, aber dann ist $2x_1 + 3x_2 \neq 3$.

DIESE ÜBERLEGUNGEN funktionieren genauso für allgemeine Matrizen. Lassen Sie uns das formal aufschreiben. Dazu ist ein neues Wort nützlich.

Definition. Sind $\mathbf{v}_1, \ldots, \mathbf{v}_r \in K^n$ und $a_1, \ldots, a_r \in K$, so heißen Ausdrücke der Form

$$a_1 \cdot \mathbf{v}_1 + a_2 \cdot \mathbf{v}_2 + \cdots + a_r \cdot \mathbf{v}_r$$

Linearkombinationen der Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_r$.

Die allgemeinen Aussagen zum Beispiel können wir jetzt so formulieren:

Bemerkung. 1. Das Bild der Matrix A besteht genau aus allen Linearkombinationen der Spaltenvektoren von A: Sind $\mathbf{v}_1, \ldots, \mathbf{v}_n$ die Spaltenvektoren von A, so ist

$$\text{Bild}(A) = \{b = c_1 \cdot v_1 + \cdots + c_n \cdot v_n | c_1, \ldots, c_n \in K\}.$$

2. Ist eine Spalte, eine Linearkombination der anderen, so können wir diese Spalte bei der Berechnung des Bildes weglassen:

   Ist $\mathbf{v}_n = a_1\mathbf{v}_1 + \ldots a_{n-1}\mathbf{v}_{n-1}$ so ist

   $$\{c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n \in K^m | c_1, \ldots, c_n \in K\} = \{d_1 \cdot \mathbf{v}_1 + \cdots + d_{n-1} \cdot \mathbf{v}_{n-1} \in K^m | c_1, \ldots, c_{n-1} \in K\}.$$

3. Das Bild von A ändert sich nicht, wenn wir ein Vielfaches einer Spalte zu einer anderen addieren: Ist $a \in K$ und $1 \leq i < n$ so gilt

   $$\{c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n \in K^m | c_1, \ldots, c_n \in K\} = \{d_1 \cdot \mathbf{v}_1 + \cdots + d_n \cdot (\mathbf{v}_n + a \cdot \mathbf{v}_i) \in K^m | d_1, \ldots, d_n \in K\}.$$

   Wir können das Bild also bestimmen, indem wir den Gauß-Algorithmus auf die Spalten der Matrix anwenden.

   Aber: Spaltenumformungen machen für das Gleichungssystem keinen Sinn, nur für das Bild.

BEWEIS: Wir schreiben jetzt einfach die Überlegung aus unserem Beispiel für eine allgemeine Matrix auf:

1. Wir hatten $A \cdot x$ definiert als

$$A \cdot x = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}$$

und

$$\begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} = x_1 \cdot \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \cdot \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n \cdot \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}.$$

Also ist

$$\begin{aligned} \text{Bild}(A) &= \{A \cdot \mathbf{x} | \mathbf{x} \in K^n\} \\ &= \{c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n | c_1, \ldots, c_n \in K\}. \end{aligned}$$

2. Die Inklusion „$\subseteq$“ folgt, indem wir wie im Beispiel den Ausdruck für v_n einsetzen: Ist $b = c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot \mathbf{v}_n$ und

$$\mathbf{v}_n = a_1\mathbf{v}_1 + \ldots a_{n-1}\mathbf{v}_{n-1}$$

so gilt

$$\begin{aligned} b &= c_1 \cdot \mathbf{v}_1 + \cdots + c_n \cdot (a_1\mathbf{v}_1 + \ldots a_{n-1}\mathbf{v}_{n-1}) \\ &= c_1 \cdot \mathbf{v}_1 + \cdots + c_n a_1 \mathbf{v}_1 + \cdots + c_n a_{n-1}\mathbf{v}_{n-1} \\ &= (c_1 + c_n a_1)\mathbf{v}_1 + \cdots + (c_{n-1} + c_n a_{n-1})\mathbf{v}_{n-1} \end{aligned}$$

und das ist eine Linearkombination der Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$.

Die umgekehrte Inklusion „$\supseteq$“ ist klar, da jede Linearkombination der Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ auch eine Linearkombination der Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_n$ ist, indem wir den Koeffizienten von $\mathbf{v}_n$ gleich 0 wählen.

3. Diesen Punkt können Sie ähnlich wie den vorigen Punkt zeigen, das möchte ich Ihnen als Übung überlassen. Fragen Sie nach, wenn Sie dabei auf Probleme stoßen.

Unterräume und lineare Unabhängigkeit

Für die Berechnung des Bildes, ist es demnach praktisch, einen Begriff zu haben, der erklärt wann wir in einer Liste $\mathbf{v}_1, \ldots, \mathbf{v}_r$ keinen Vektor als Linearkombination der anderen schreiben können.

Die Formulierung „Wir können keinen der Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_r$ als Linearkombination der anderen schreiben“ ist etwas sperrig, formal sähe das vielleicht so aus:

Für alle $i = 1, \ldots r$ und $a_1, \ldots, a_r \in K$ gilt

$$\mathbf{v}_i \neq a_1\mathbf{v}_1 + \cdots + a_{i-1}\mathbf{v}_{i-1} + a_{i+1}\mathbf{v}_{i+1} + \cdots + a_r\mathbf{v}_r.$$

Es wäre lästig, das nachzuprüfen, weil wir die Bedingung für alle r Vektoren einzeln prüfen müssten.

Mit einem Trick, können wir das Gleiche viel einfacher formulieren:

Definition. Eine Menge von Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_r \in K^n$ heißt *linear unabhängig*, wenn die Gleichung

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_r\mathbf{v}_r = 0$$

in K nur die triviale Lösung $a_1 = a_2 = \cdots = a_r = 0$ besitzt, d.h.

$$(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_r\mathbf{v}_r = 0) \Rightarrow a_1 = a_2 = \cdots = a_r = 0.$$

Entsprechend heißen $\mathbf{v}_1, \ldots, \mathbf{v}_r \in K^n$ *linear abhängig*, wenn die Gleichung

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_r\mathbf{v}_r = 0$$

eine Lösung besitzt, in der wenigstens ein $a_i \neq 0$ ist.

FRAGE: Sehen Sie, wieso diese Definition linear unabhängiger Vektoren auch eine korrekte Beschreibung der Bedingung „Wir können keinen der Vektoren $v_1, \ldots, v_r$ als Summe der anderen schreiben“ ist?

ANTWORT: Wenn wir $\mathbf{v}_1 = a_2\mathbf{v}_2 + \cdots + a_r\mathbf{v}_r$ schreiben können, so ist

$$(-1) \cdot \mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_r\mathbf{v}_r = 0$$

eine nicht-triviale Lösung der Gleichung $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_r\mathbf{v}_r = 0$, also sind die Vektoren dann nicht linear unabhängig.

Haben wir umgekehrt eine Lösung der Gleichung

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_r\mathbf{v}_r = 0$$

in der ein $a_i \neq 0$ ist, so können wir die Gleichung umstellen:

$$\begin{aligned} -a_i\mathbf{v}_i &= a_1\mathbf{v}_1 + \cdots + a_{i-1}\mathbf{v}_{i-1} + a_{i+1}\mathbf{v}_{i+1} + \ldots a_r\mathbf{v}_r \qquad | \cdot (-a_i)^{-1} \\ \Leftrightarrow v_i &= -\frac{a_1}{a_i}\mathbf{v}_1 - \cdots - \frac{a_{i-1}}{a_i}\mathbf{v}_{i-1} - \frac{a_{i+1}}{a_i}\mathbf{v}_{i+1} - \cdots - \frac{a_r}{a_i}\mathbf{v}_r \end{aligned}$$

und sehen, dass wir dann $\mathbf{v}_i$ als Linearkombination der anderen Vektoren schreiben können.

Notation 24. Für gegebene Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_r \in K^m$ bezeichnen wir mit

$$\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_r) := \{\mathbf{v} = a_1\mathbf{v}_1 + \ldots a_r\mathbf{v}_r \,|\, a_1, \ldots, a_r \in K\} \subseteq K^m$$

den Spann (oder die lineare Hülle) der Vektoren.

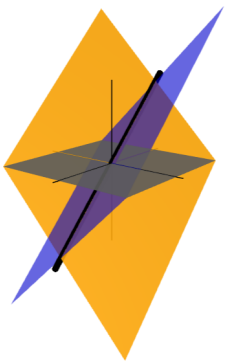
Per Konvention ist der Spann der leeren Menge $\mathrm{Span}() := \mathbf{0} \subset K^m$ der 0-Vektor.

Teilmengen dieser Art heißen auch Unterräume von K^m.

Definition. Eine *nicht leere* Teilmenge $U \subset K^m$ heißt *Unterraum* von K^m, wenn gilt

1. Für jeden Vektor $\mathbf{v} \in U$ sind alle Vielfachen ebenfalls in U, also:
$$(\mathbf{v} \in U, c \in K) \Rightarrow c \cdot \mathbf{v} \in U.$$
2. Für je zwei Vektoren $\mathbf{v}, \mathbf{v}' \in U$ ist auch ihre Summe $\mathbf{v} + \mathbf{v}' \in U$, also:
$$(\mathbf{v}, \mathbf{v}' \in U) \Rightarrow \mathbf{v} + \mathbf{v}' \in U.$$

Einige lineare Unterräume in $\mathbb{R}^3$:



Bemerkung. 1. Wenn Sie sich an die Formelschreibweise gewöhnen, können Sie irgendwann die Formelzeilen schneller lesen, als den gleichbedeutenden Text.

2. Statt zu fordern, dass U nicht leer ist, könnten wir genauso gut fordern, dass $\mathbf{0} \in U$ liegt, denn wenn ein $\mathbf{v} \in U$ ist, dann auch $0 \cdot \mathbf{v} = \mathbf{0}$.

   Wenn Sie von einer Teilmenge prüfen wollen, ob diese ein Unterraum ist, ist es oft einfacher $\mathbf{0} \in U$ zu zeigen, als ein anderes Element von U anzugeben.

Der Spann $\text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_r) \subseteq K^m$ von Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_r \in K^m$ ist ein Unterraum, insbesondere ist das Bild einer Matrix $\text{Bild}(A)$ ist ein Unterraum. Wir werden noch sehen, dass sich alle Unterräume als Spann geeigneter Vektoren schreiben lassen.

Wir waren aber eigentlich auf der Suche nach einem möglichst einfachen Erzeugendensystem für das Bild einer Matrix, zum Beispiel waren für das Bild der Matrix $A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 2 & 3 & 5 \\ 1 & 1 & 2 \end{pmatrix}$ die ersten beiden Spalten ausreichend und besonders nützlich, um zu sehen, ob ein Vektor $\mathbf{b}$ im Bild liegt oder nicht.

Das erste Kriterium für ein einfaches Erzeugendensystem sollte sein, dass wir keinen der Vektoren weglassen können. Diese Eigenschaft bekommt einen Namen:

Definition. Eine Teilmenge $\mathbf{v}_1, \ldots, \mathbf{v}_r \in U$ eines Unterraum $U \subseteq K^m$ heißt *Erzeugendensystem von* U, wenn sich jeder Vektor in U als Linearkombination der $\mathbf{v}_1, \ldots, \mathbf{v}_r$ schreiben lässt, d.h. wenn

$$U = \text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_r).$$

Eine *Basis* eines Unterraums ist ein linear unabhängiges Erzeugendensystem.

Behauptung 25. *Ist $\mathbf{v}_1, \ldots, \mathbf{v}_r \in U$ eine Basis von U, so lässt sich jeder Vektor $\mathbf{v} \in U$ eindeutig als Linearkombination $\mathbf{v} = a_1\mathbf{v}_1 + \ldots a_r\mathbf{v}_r$ schreiben, die Koeffizienten dieser Gleichung heißen* die Koordinaten *von $\mathbf{v}$ bezüglich der Basis $\mathbf{v}_1, \ldots, \mathbf{v}_r$.*

Beweis. Nach Definition ist eine Basis $\mathbf{v}_1, \ldots, \mathbf{v}_r \in U$ insbesondere ein Erzeugendensystem

$$\text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_r) = U$$

also lässt sich jedes Element $\mathbf{v} \in U$ als Linearkombination der $\mathbf{v}_i$ schreiben. Angenommen wir hätten zwei Darstellungen

$$a_1\mathbf{v}_1 + \ldots a_r\mathbf{v}_r = \mathbf{v} = b_1\mathbf{v}_1 + \ldots b_r\mathbf{v}_r,$$

dann gilt

$$(a_1 - b_1)\mathbf{v}_1 + \ldots (a_r - b_r)\mathbf{v}_r = 0.$$

Da die Vektoren linear unabhängig sind, muss dann $a_1 - b_1 = a_2 - b_2 = \cdots = a_r - b_r = 0$ gelten, d.h. $a_i = b_i$ für alle i. Die Darstellung war als eindeutig. □

Bemerkung. Das Argument des Beweises lässt sich umdrehen: Lässt sich jeder Vektor $\mathbf{v} \in U$ eindeutig als Linearkombination $\mathbf{v} = a_1\mathbf{v}_1 + \ldots a_r\mathbf{v}_r$ schreiben, so sind die Vektoren $\mathbf{v}_1, \ldots \mathbf{v}_r$ eine Basis von U.

Beweis. Hier wäre nur zu zeigen, dass die $\mathbf{v}_i$ linear unabhängig sind, d.h., dass das Gleichungssystem

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_r\mathbf{v}_r = 0$$

nur die triviale Lösung $a_i = 0$ besitzt. Das gilt aber, da sich der 0-Vektor nach Voraussetzung eindeutig als Linearkombination der $\mathbf{v}_i$ schreiben lässt. □

Mit diesen Begriffen können wir nun die Lösbarkeit von Gleichungssystemen neu beschreiben:

Folgerung 26. *Sei A eine $m \times n$ Matrix mit Koeffizienten in einem Körper K dann gilt:*

1. *Das Gleichungssystem $A\mathbf{x} = b$ ist genau dann für alle b lösbar, wenn die Spalten ein Erzeugendensystem von K^m bilden.*
2. *Das Gleichungssystem $A\mathbf{x} = b$ kann nur dann eindeutig lösbar sein, wenn die Spalten der Matrix A linear unabhängig sind.*
3. *Das Gleichungssystem $A\mathbf{x} = b$ ist genau dann für alle b eindeutig lösbar, wenn die Spalten der Matrix A eine Basis von K^m bilden.*

Woher kommt hier K^m? Berechnen wir für die $m \times n$-Matrix $A \cdot \mathbf{x}$ für einen Spaltenvektor $\mathbf{x}$ mit n Einträgen, so erhalten wir einen Spaltenvektor mit m Einträgen.

Spaltenvektoren mit m Einträgen sind genau die Elemente von K^m. Insbesondere sind die Spalten von A Elemente von K^m. Wir können uns also fragen, ob die Spaltenvekotren eine Basis oder ein Erzeugendensystem von K^m sind.

DAS IST ÜBERRASCHEND da der Gauß-Algorithmus nur die Zeilen der Matrix betrachtet hat, die Bedeutung der Spalten der Matrix war daraus nicht klar ersichtlich.

WIR HABEN SCHON GESEHEN, wie wir aus einem Erzeugendensystem eine Basis eines Unterraums $U = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_r)$ bestimmen können, denn wir wissen:

1. Für alle $c \in K$ und $i < r$ gilt

   $$\mathrm{Span}\left(\mathbf{v}_1, \ldots, \mathbf{v}_r\right) = \mathrm{Span}\left(\mathbf{v}_1, \ldots, (\mathbf{v}_r + c_i\mathbf{v}_i)\right),$$

   d.h. der Spann ändert sich nicht wenn wir ein Vielfaches eines Vektors zu einem anderen hinzuaddieren.

2. Für alle $c \in K$ mit $c \neq 0$ und $1 \leq i \leq r$ gilt

   $$\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_r) = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1}, c\mathbf{v}_i, \mathbf{v}_{i+1}, \ldots \mathbf{v}_r),$$

   d.h. der Spann ändert sich nicht wenn wir einen Vektor mit einer Zahl $\neq 0$ multiplizieren.

3. Ist $v_r = a_1\mathbf{v}_1 + \cdots + a_{r-1}v_{r-1}$ so gilt

$$\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_r) = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_{r-1}),$$

d.h. ist einer der Vektoren eine Linearkombination der anderen, z.B. der 0-Vektor, so können wir diesen Vektor weglassen.

Damit können wir den Gaußalgorithmus auf Spaltenvektoren anwenden, zum Beispiel für die folgenden Vektoren in $\mathbb{R}^3$:

$$\begin{aligned}
\mathrm{Span}\left(\begin{pmatrix}1\\4\\7\end{pmatrix}, \begin{pmatrix}2\\5\\8\end{pmatrix}, \begin{pmatrix}3\\6\\9\end{pmatrix}\right) &= \mathrm{Span}\left(\begin{pmatrix}\mathbf{1}\\4\\7\end{pmatrix}, \begin{pmatrix}2\\5\\8\end{pmatrix} - 2\cdot\begin{pmatrix}\mathbf{1}\\4\\7\end{pmatrix}, \begin{pmatrix}3\\6\\9\end{pmatrix} - 3\cdot\begin{pmatrix}1\\4\\7\end{pmatrix}\right) \\
&= \mathrm{Span}\left(\begin{pmatrix}1\\4\\7\end{pmatrix}, \begin{pmatrix}\mathbf{0}\\-3\\-6\end{pmatrix}, \begin{pmatrix}\mathbf{0}\\-6\\-12\end{pmatrix}\right) && \mid -1/3\cdot \text{(2. Vektor)} \\
&= \mathrm{Span}\left(\begin{pmatrix}\mathbf{1}\\4\\7\end{pmatrix}, \begin{pmatrix}\mathbf{0}\\1\\2\end{pmatrix}, \begin{pmatrix}\mathbf{0}\\-6\\-12\end{pmatrix}\right) && \mid \begin{array}{l}\text{addiere } 6\cdot 2.\\ \text{zum 3. Vektor}\end{array} \\
&= \mathrm{Span}\left(\begin{pmatrix}\mathbf{1}\\4\\7\end{pmatrix}, \begin{pmatrix}\mathbf{0}\\\mathbf{1}\\2\end{pmatrix}, \begin{pmatrix}\mathbf{0}\\\mathbf{0}\\0\end{pmatrix}\right) \\
&= \mathrm{Span}\left(\begin{pmatrix}\mathbf{1}\\4\\7\end{pmatrix}, \begin{pmatrix}\mathbf{0}\\\mathbf{1}\\2\end{pmatrix}\right) && \mid -4\cdot\text{(2.) zum (1.)} \\
&= \mathrm{Span}\left(\begin{pmatrix}\mathbf{1}\\\mathbf{0}\\-1\end{pmatrix}, \begin{pmatrix}\mathbf{0}\\\mathbf{1}\\2\end{pmatrix}\right).
\end{aligned}$$

Die letzten beiden Vektoren sind linear unabhängig, da die jeweils 1. Koordinate $\neq 0$ nur in einem der beiden Vektoren vorkommt.

Mit dieser Basis können wir von einem Vektor $\begin{pmatrix}a\\b\\c\end{pmatrix} \in \mathbb{R}^3$ sofort sehen, ob dieser ein Element des Spanns der Vektoren ist, da

$$x_1\begin{pmatrix}\mathbf{1}\\\mathbf{0}\\-1\end{pmatrix} + x_2\begin{pmatrix}\mathbf{0}\\\mathbf{1}\\2\end{pmatrix} = \begin{pmatrix}x_1\\x_2\\(-x_1+2x_2)\end{pmatrix},$$

also ist $\begin{pmatrix}a\\b\\c\end{pmatrix}$ genau dann im Spann wenn $c = -a + 2b$.

Der erste Blick auf diese Rechnung sagt Ihnen vielleicht, dass die letzte Koordinate für den Spann in gewisser Weise überflüssig ist. Überflüssige Koordinaten sind manchmal nützlich, Sie könnten die letzte Koordinate $-a + 2b$ hier als „Prüfziffer“ auffassen.

Exkurs: Lineare Codes

Codes dienen dazu, Fehler, die bei der Übermittlung von Daten immer einmal auftreten, zu erkennen und wenn möglich zu korrigieren. Mathematisch können wir das in der Sprache der linearen Algebra zum Beispiel so verstehen: Eine Nachricht ist eine Folge von Zeichen, zum Beispiel einfach eine Folge von 0en und 1en, die Buchstaben können wir in jedem Fall durchnummerieren und damit als Elemente eines Körpers K auffassen. Eine Nachricht ist dann ein Vektor $\mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in K^n$. Versenden wir eine Nachricht $\mathbf{v}$, so treten bei der Übertragung Fehler auf[14], d.h. einige Koordinaten von $\mathbf{v}$ kommen möglicherweise verkehrt an.

[14] z.B. $\mathbf{v} = \begin{pmatrix} x_1 \\ F \\ x_3 \\ \vdots \\ x_n \end{pmatrix}$ statt $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix}$

Codes versuchen geschickt Koordinaten so hinzu zu fügen, dass Fehler erkannt und bestenfalls korrigiert werden können. Das einfachste Beispiel ist ein Wiederholungscode, statt $\begin{pmatrix} a \\ b \end{pmatrix}$ könnten wir $\begin{pmatrix} a \\ a \\ b \\ b \end{pmatrix}$ verschicken, das entspricht in Matrixschreibweise der Abbildung

$$A \colon K^2 \to K^4 \tag{4}$$

$$\begin{pmatrix} a \\ b \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix}. \tag{5}$$

Erhalten wir dann zum Beispiel die Nachricht $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}$ so sehen wir, dass ein Fehler aufgetreten ist, wissen aber nicht, ob die erste oder die zweite Koordinate falsch übertragen wurde.

Das geht etwas überraschenderweise mit längeren Nachrichten besser. Zum Beispiel hat Hamming die folgende Abbildung

vorgeschlagen:

$$H\colon (\mathbb{Z}/2\mathbb{Z})^4 \to (\mathbb{Z}/2\mathbb{Z})^7 \tag{6}$$

$$\mathbf{v} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_2+x_3+x_4 \\ x_1+x_3+x_4 \\ x_1+x_2+x_4 \end{pmatrix} = \begin{pmatrix} \mathbf{1} & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 \\ 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & 0 & \mathbf{1} \\ 0 & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \mathbf{1} & 0 & \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} & 0 & \mathbf{1} \end{pmatrix} \cdot \mathbf{v}. \tag{7}$$

Behauptung 27. *Sind $\mathbf{v} \neq \mathbf{v}' \in (\mathbb{Z}/2\mathbb{Z})^4$ zwei unterschiedliche Nachrichten, so unterscheiden sich die codierten Nachrichten $H\mathbf{v}$ und $H\mathbf{v}'$ an mindestens 3 Koordinaten.*

Wird bei der Übertragung eine Koordinate falsch übertragen, so kann der Fehler erkannt und korrigiert werden.

Beweisidee Wenn sich $H\mathbf{v}$ und $H\mathbf{v}'$ an weniger als 3 Stellen unterscheiden, bedeutet das, dass im Vektor $H(\mathbf{v}-\mathbf{v}')$ mehr als $7-3=4$ Koordinaten $=0$ sind. Wir müssen also nur nachprüfen, dass für alle Matrizen H', die aus H entstehen indem wir 5 der Zeilen auswählen gilt, dass $\mathrm{Ker}(H') = \{0\}$. Wegen der Symmetrie der Definition ist müssen wir dafür nur recht wenige Matrizen einzeln prüfen, die vollständige Liste überlasse ich Ihnen als Übung.

Frage: Angenommen Sie empfangen die Nachricht $\begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$. Was war wahrscheinlich die Originalnachricht?

In der Vorlesung haben Sie die Auflösung: $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ gefunden: Für den Vektor $\mathbf{v} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$ wären die letzten Koordinaten von $H\mathbf{v}$

$$\begin{aligned} x_2+x_3+x_4 &= 1+1+0 = 0 \text{ in } \mathbb{Z}/2\mathbb{Z} \\ x_1+x_3+x_4 &= 1+1+0 = 0 \text{ in } \mathbb{Z}/2\mathbb{Z} \\ x_1+x_2+x_4 &= 1+1+0 = 0 \text{ in } \mathbb{Z}/2\mathbb{Z} \end{aligned}$$

damit sind 2 Prüfziffern falsch. Die einzige Koordinate die in der Berechnung der beiden falschen Ziffern vorkommt ist x_2, ändern

wir den Eintrag von 1 zu 0, so stimmen alle Prüfziffern, d.h. für

$$\mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \text{ gilt } H\mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

DER HAMMING CODE ist für Wörter der Länge 4 in den Zeichen $0, 1$ optimal und wird darum noch immer in Speicherbausteinen und Speichermedien verwendet[15]. Lassen Sie uns kurz überlegen, was mit „optimal“ gemeint ist.

Angenommen wir möchten Fehler im Vektor $\mathbf{w} = H\mathbf{v} \in (\mathbb{Z}/2\mathbb{Z})^7$ korrigieren. Es gibt 7 Stellen an denen der Fehler auftreten kann, also sollten $H\mathbf{v}$ und 7 weitere Vektoren als $\mathbf{w}$ erkannt werden. Das Bild von H enthält aber 2^4 Elemente, für jedes davon wollen wir 8 Vektoren als „richtig“ erkennen, das macht insgesamt $8 \cdot 2^4 = 2^7$ Vektoren. Genauso viele Elemente hat $(\mathbb{Z}/2\mathbb{Z})^7$.

[15] Auch dort werden zunehmend mit ähnlichen Verfahren längere Blöcke codiert, aber es wird fast nirgends einfach nur die Nachricht gespeichert, sondern immer ein Code dazu.

FÜR LÄNGERE NACHRICHTEN ist eine bessere Fehlerkorrektur möglich – das ist wichtig, wenn Sie zum Beispiel im Mobilfunk oder bei der Übertragung von Bildern von Satelliten oder dem James-Webb-Teleskop recht störungsanfällige Übertragungswege haben. Zum Beispiel verwenden QR-Codes einen linearen Code, der für ein Alphabet mit 256 Buchstaben Nachrichten so kodiert, dass bis zu 30% fehlerhaft übertragener Daten korrigiert werden können.

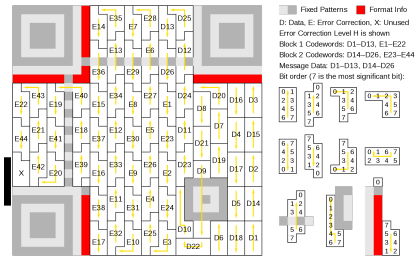


Abbildung 1: Grafik aus Wikipedia entnommen, der englische Artikel ist recht detailliert.

Die Dimension eines Unterraums

Zurück zur Bestimmung von Basen eines Unterraums. Erinnern Sie sich an die Definition, was eine Basis ist? Vielleicht zum Aufwärmen die Bemerkung, dass für Basen die gleichen Umfromungsregeln gelten, wie für den Spann:

Bemerkung. Sind $\mathbf{v}_1, \dots, \mathbf{v}_r$ eine Basis eines Unterraums U dann gilt:

1. Wir dürfen einen Vektor mit einer Zahl $\neq 0$ multiplizieren:

   Für alle $a \neq 0$ und $1 \leq i \leq r$ ist dann auch $\mathbf{v}_1, \dots, c \cdot \mathbf{v}_i, \mathbf{v}_{i+1}, \dots, \mathbf{v}_r$ eine Basis von U.

2. Wir können Vielfache eines (oder mehrerer) Basisvektors zu einem anderen addieren: Ist $\mathbf{v}_i' := \mathbf{v}_i + c_1\mathbf{v}_1 + \dots + c_{i-1}\mathbf{v}_{i-1} + c_{i+1}\mathbf{v}_{i+1} + \dots + c_r\mathbf{v}_r$ mit $c_1, \dots, c_r \in K$ dann ist auch $\mathbf{v}_1, \dots, \mathbf{v}_i', \mathbf{v}_{i+1}, \dots, \mathbf{v}_r$ eine Basis von U.

DENN: Wir hatten beide Aussagen für den Span schon gesehen, müssen also nur noch zeigen, dass die angegebenen Vektoren wieder linear unabhängig sind.

1. Zu zeigen ist: Das Gleichungsystem

$$a_1\mathbf{v}_1 + \cdots + a_i(c \cdot \mathbf{v}_i) + \cdots + a_r\mathbf{v}_r = 0$$

hat in K nur die triviale Lösung $a_1 = a_2 = \cdots = a_r = 0$.

Da aber $\mathbf{v}_1, \ldots, \mathbf{v}_n$ linear unabhängig sind folgt aus

$$a_1\mathbf{v}_1 + \cdots + (a_i \cdot c) \cdot \mathbf{v}_i + \cdots + a_r\mathbf{v}_r = 0,$$

dass $a_1 = \cdots = a_ic = \cdots = a_r = 0$ und weil $c \neq 0$ ist folgt auch $a_i = 0$.

2. Geht genauso: Angenommen

$$a_1\mathbf{v}_1 + \cdots + a_i\mathbf{v}_i' + \cdots + a_r\mathbf{v}_r = 0$$

Dann ist

$$\begin{aligned} 0 &= a_1\mathbf{v}_1 + \cdots + a_i\mathbf{v}_i' + \cdots + a_r\mathbf{v}_r \\ &= a_1\mathbf{v}_1 + \cdots + a_i \underbrace{(\mathbf{v}_i + c_1\mathbf{v}_1 + \cdots + c_{i-1}\mathbf{v}_{i-1} + c_{i+1}\mathbf{v}_{i+1} + \cdots + c_r\mathbf{v}_r)}_{=\mathbf{v}_i'} + \cdots + a_r\mathbf{v}_r \\ &= (a_1 + a_ic_1)\mathbf{v}_1 + \cdots + a_i\mathbf{v}_i + \cdots + (a_r + a_ic_r)\mathbf{v}_r. \end{aligned}$$

eine Linearkombination der $\mathbf{v}_i$, also ist $a_i = 0$ und alle $a_j + a_ic_1 = 0$, weil $a_i = 0$ ist, bedeutet das aber $a_j = 0$ für alle j. Daher sind auch $\mathbf{v}_1, \ldots, \mathbf{v}_i', \mathbf{v}_{i+1}, \ldots, \mathbf{v}_r$ linear unabhängig.

Wir haben Basen mit einem Algorithmus bestimmt. Für größere Gleichungssysteme ist leider nicht so klar, ob wir mit unterschiedlichen Verfahren unterschiedlich viele linear unabhängige Vektoren erhalten. Das sollte nicht sein, lassen Sie uns das formulieren.

Satz 28. *Je zwei Basen eines Unterraums haben gleich viele Elemente.*

ZU ZEIGEN IST: Sind $\mathbf{v}_1, \ldots, \mathbf{v}_n$ und $\mathbf{w}_1, \ldots, \mathbf{w}_d$ Basen eines Unterraums $U \subset K^m$, so gilt $n = d$.

Es gibt unterschiedliche Beweisideen mit denen Sie diese Aussage nachweisen können. Zum Beispiel könnten Sie im Fall, dass $d > n$ ist, alle $\mathbf{w}_i$ als Linearkombination der $\mathbf{v}_j$ schreiben, die Koeffizienten in eine Matrix eintragen und dann mit dem Gaußalgorithmus eine nichttriviale Linearkombination der 0 finden. Das ist leider ziemlich unübersichtlich und braucht viele Indices.

Ein besserer Trick ist, sich zu überlegen, dass wir nach und nach $\mathbf{v}_i$'s so durch $\mathbf{w}$'s ersetzen können dass wir wieder eine Basis erhalten.

Behauptung 29 (Austauschsatz)**.** *Sind die Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_n$ eine Basis des Unterraums U und $\mathbf{w}_1, \ldots, \mathbf{w}_j \in U$ linear unabhängig. Dann existiert eine Liste von j Indices $I = \{i_1, \ldots, i_j\} \subset \{1, \ldots, n\}$ so dass $(\mathbf{w}_1, \ldots, \mathbf{w}_j, (\mathbf{v}_\ell)_{\ell \in \{1,\ldots,n\} \setminus I})$ eine Basis von U ist. Insbesondere gilt dann $j \leq n$.*

DAS BEWEISPRINZIP für diese Aussage ist *vollständige Induktion*: Wir zeigen zunächst, dass die Aussage für $j = 1$ stimmt und überlegen uns dann, dass wenn die Aussage für $j = k$ gilt, dann auch für $j = k + 1$. Damit haben wir dann gezeigt, dass die Aussage für alle $j \in \mathbb{N}$ gelten muss.

Beweis. INDUKTIONS ANFANG: *Zeige, dass die Behauptung für $j = 1$ gilt*, d.h. zu zeigen: Ist $\mathbf{w}_1 \in U$ ein Vektor $\neq 0$, so existiert ein i, so dass $(\mathbf{w}_1, \mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_n)$ eine Basis von U ist.

Sie hatten in Aufgabe 3 (3) von Blatt 4 eine solche Ersetzung in einem Beispiel konstruiert.

Da $\mathbf{v}_1, \dots, \mathbf{v}_n$ eine Basis von U ist und $\mathbf{w}_1 \in U$, können wir $\mathbf{w}_1$ als Linearkombination

$$\mathbf{w}_1 = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n, \quad a_i \in K$$

schreiben. Da $\mathbf{w}_1 \neq 0$ ist, muss wenigstens ein $a_i \neq 0$ sein, dann gilt aber

$$\begin{aligned}
\operatorname{Span}(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_n) &= \operatorname{Span}(\mathbf{v}_1, \dots, a_i \cdot \mathbf{v}_i, \dots, v_n) && \text{mult.}\mathbf{v}_i \text{ mit } a_i \neq 0) \\
&= \operatorname{Span}(\mathbf{v}_1, \dots, \underbrace{a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n}_{=\mathbf{w}}, \dots, v_n) && \text{addiere Vielfache anderer Spalten} \\
&= \operatorname{Span}(\mathbf{w}_1, \mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_n).
\end{aligned}$$

Wir haben uns außerdem vorhin überlegt, dass diese Umformungen auch die lineare Unabhängigkeit der Vektoren erhalten, die Vektoren sind also wieder eine Basis von U.

INDUKTIONS SCHRITT: *Wenn wir die Aussage für eine Zahl k bewiesen haben, so folgt die Aussage auch für die Zahl $k + 1$.*

Sei also

$$\mathbf{w}_1, \dots, \mathbf{w}_k, (\mathbf{v}_\ell)_{\ell \in \{1,\dots,n\} \smallsetminus I}$$

eine Basis von U und die Vektoren $\mathbf{w}_1, \dots \mathbf{w}_{k+1}$ seien linear unabhängig. Dann müssen wir zeigen, dass wir einen der $\mathbf{v}_\ell$'s durch $\mathbf{w}_{k+1}$ ersetzen können.

Der Übersicht halber nummerieren wir die $\mathbf{v}_i$ so um, dass $I = \{1, \dots, k\}$, also $\{1, \dots, n\} \smallsetminus I = \{k + 1, \dots, n\}$

Dann können wir wie zuvor

$$\mathbf{w}_{k+1} = c_1\mathbf{w}_1 + \cdots + c_k\mathbf{w}_k + c_{k+1}\mathbf{v}_{k+1} + \dots c_n\mathbf{v}_n$$

als Linearkombination der Basis schreiben. Da die $\mathbf{w}_i$ linear unabhängig sind, lässt sich $\mathbf{w}_{k+1}$ nicht als Linearkombination der $\mathbf{w}_1, \dots, \mathbf{w}_k$ alleine schreiben. Darum muss eine der Konstanten $c_{k+1}, \dots, c_n$ ungleich 0 sein. Ist ein solches $c_\ell \neq 0$, können wir aber wie im ersten Schritt, $\mathbf{v}_i$ durch $\mathbf{w}_{k+1}$ austauschen, d.h. dann sind

$$\mathbf{w}_1, \dots, \mathbf{w}_{k+1}, (\mathbf{v}_i)_{i \in \{k+1,\dots n\} \smallsetminus \{\ell\}}$$

eine Basis.

Im letzten Schritt, haben wir auch gesehen, dass $j \leq n$ gelten muss. □

AUS DEM AUSTAUSCHSATZ folgt insbesondere, dass je zwei Basen $\mathbf{v}_1, \dots, \mathbf{v}_n$ und $\mathbf{w}_1, \dots, \mathbf{w}_d$ Basen eines Unterraums $U \subset K^m$ gleich

viele Elemente haben, da damit $d \leq n$ erhalten und wenn wir die Rollen der $\mathbf{v}_i$ und $\mathbf{w}_j$ vertauschen genauso $n \leq d$ gezeigt haben.

Damit ist die Anzahl der Elemente einer Basis ein sinnvoller Dimensionsbegriff.

Definition. Ist $U \subseteq K^m$ ein Unterraum, so heißt die Anzahl der Elemente einer Basis *Dimension* von U:

$$\dim_K U := \text{Anzahl der Elemente einer Basis von } U.$$

Wir geben manchmal den Körper K als Index des Symbols dim an, da zum Beispiel $\mathbb{C} = \mathbb{R}^2$ und damit $\dim_{\mathbb{R}} \mathbb{C} = 2$, aber $\dim_{\mathbb{C}} \mathbb{C} = 1$.

Beispiel 30. Es gilt für alle n, dass $\dim_K K^n = n$, denn die Vektoren

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

sind ein linear unabhängiges Erzeugendensystem, genannt die Standardbasis von K^n.

In den Übungsaufgaben werden Sie aus dem Austauschsatz das folgende Ergebnis ableiten.

Folgerung 31. *Sei $U \subseteq K^n$ ein Unterraum, dann gilt:*

1. *U besitzt eine Basis und es gilt $\dim_K U \leq n$.*
2. *Ist $\dim_K U = n$, so gilt sogar $U = K^n$.*
3. *Sind die Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_r \in U$ linear unabhängig, so können wir diese zu einer Basis $\mathbf{v}_1, \ldots, \mathbf{v}_r, \mathbf{v}_{r+1}, \ldots, \mathbf{v}_d$ von U ergänzen.*

Lasse Sie uns noch einige Beispiele berechnen:

Beispiel 32. 1. Sei $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ dann gilt

   (a) $\dim_K(\text{Bild}(A)) = 1$, denn $\text{Bild}(A) = \text{Span}(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix})$, also ist $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ eine Basis von $\text{Bild}(A)$.

   (b) $\dim_K(\text{Ker}(A)) = 2$

2. Sei $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$ dann gilt

   (a) $\dim_K(\text{Bild}(A)) = 2$, denn wir haben

   $$\text{Bild}(A) = \text{Span}(\begin{pmatrix} 1 \\ 4 \\ 7 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix}, \begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix}) = \text{Span}(\begin{pmatrix} \mathbf{1} \\ \mathbf{0} \\ -1 \end{pmatrix}, \begin{pmatrix} \mathbf{0} \\ \mathbf{1} \\ 2 \end{pmatrix})$$

   berechnet und damit eine 2-elementige Basis des Bildes gefunden.

(b) $\dim_K(\mathrm{Ker}(A)) = 1$, denn mit dem Gauß-Algorithmus finden wir

$$\mathrm{Ker}(A) = \text{Lösung}(Ax = 0) = \left\{ \begin{pmatrix} a \\ -2a \\ a \end{pmatrix} \mid a \in K \right\} = \mathrm{Span}(\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}).$$

3. Sei $A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$ dann gilt

(a) $\dim_K(\mathrm{Bild}(A)) = 3$, denn der Gauß-Algorithmus zeigt, dass $\text{Lösung}(Ax = 0) = \{0\}$ gilt, die Spalten der Matrix also linear unabhängig sind. Da $\dim_K K^3 = 3$ gilt, muss also $\mathrm{Bild}(A) = K^3$ gelten.

(b) $\dim_K(\mathrm{Ker}(A)) = 0$

4. Sei $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} : K^4 \to K^3$ dann gilt

(a) $\dim_K(\mathrm{Bild}(A)) = 1$, denn $\mathrm{Bild}(A) = \mathrm{Span}(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix})$, also ist $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ eine Basis von $\mathrm{Bild}(A)$.

(b) $\dim_K(\mathrm{Ker}(A)) = 3$, da

$$\text{Lösung}(Ax = 0) = \left\{ \begin{pmatrix} -(a+b+c) \\ a \\ b \\ c \end{pmatrix} \mid a, b, c \in K \right\}$$
$$= \mathrm{Span}(\begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}).$$

In der Vorlesung hatten Sie sich statt Beispiel 3 eine 4×3-Matrix gewünscht.

FRAGE: Fällt Ihnen an den Resultaten etwas auf?

Die Dimensionsformel

Sie hatten in der Vorlesung als Antwort die folgende Regelmäßigkeit vorgeschlagen:

Satz 33 (Dimensionsformel)**.** *Ist A eine $m \times n$ Matrix, mit Koeffizienten in einem Körper K so gilt für die Abbildung*

$$A \colon K^n \to K^m,$$

dass

$$n = \dim K^n = \dim(\mathrm{Bild}(A)) + \dim(\mathrm{Ker}(A)).$$

Beweis. Vorüberlegung: Die Aussage des Satzes scheint zu sagen, dass es möglich sein sollte, aus einer Basis von $\mathrm{Ker}(A)$ *und einer von* $\mathrm{Bild}(A)$ *eine Basis von* K^n *zu konstruieren. Lassen Sie uns das versuchen:* Sei $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathrm{Ker}(A)$ eine Basis von $\mathrm{Ker}(A) \subseteq K^n$ und $\mathbf{w}_1, \dots, \mathbf{w}_b \in \mathrm{Bild}(A)$ eine Basis von $\mathrm{Bild}(A) \subseteq K^m$.

Die Basis des Bildes besteht aus Elementen von K^m nicht K^n.

Da $\mathbf{w}_1, \dots, \mathbf{w}_b \in \mathrm{Bild}(A) = \{Ax \in K^m \mid x \in K^n\}$, existieren $\tilde{\mathbf{w}}_1, \dots, \tilde{\mathbf{w}}_b \in K^n$ mit $A\tilde{\mathbf{w}}_i = \mathbf{w}_i$.

BEHAUPTUNG: Die Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_k, \tilde{\mathbf{w}}_1, \dots, \tilde{\mathbf{w}}_b$ sind eine Basis von K^n, d.h die Vektoren sind linear unabhängig und erzeugen K^n.

Wir zeigen zunächst, dass die Vektoren linear unabhängig sind: Sei also

$$a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k + c_1\tilde{\mathbf{w}}_1 + \dots c_b\tilde{\mathbf{w}}_b = 0$$

Wir müssen irgendwie die Konstruktion der $\tilde{\mathbf{w}}_j$ verwenden, in dieser kam die Matrix A vor, also wenden wir A auf die Linearkombination an.

mit $a_i, c_j \in K$. Dann gilt

$$\begin{aligned}
0 = A0 &= A(a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k + c_1\tilde{\mathbf{w}}_1 + \dots c_b\tilde{\mathbf{w}}_b) && \text{da } A\cdot(\mathbf{v}+\mathbf{v}') = A\mathbf{v} + A\mathbf{v}' \\
&= A(\underbrace{a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k}_{\in \mathrm{Ker}(A)}) + A(c_1\tilde{\mathbf{w}}_1) + \cdots + A(c_b\tilde{\mathbf{w}}_b) && \text{da } A\cdot(c\mathbf{v}) = cA\mathbf{v} \\
&= 0 + c_1A\tilde{\mathbf{w}}_1 + \cdots + c_b \dots A\tilde{\mathbf{w}}_b && \text{da } A\tilde{w}_j = \mathbf{w}_j \\
&= c_1\mathbf{w}_1 + \cdots + c_b\mathbf{w}_b.
\end{aligned}$$

Die Vektoren $\mathbf{w}_1, \dots, \mathbf{w}_b$ sind aber linear unabhängig, also hat die Gleichung

$$c_1\mathbf{w}_1 + \cdots + c_b\mathbf{w}_b = 0$$

nur die Lösung $c_1 = \dots, c_b = 0$.

Also gilt

$$\begin{aligned}
0 &= a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k + \underbrace{c_1}_{=0}\tilde{\mathbf{w}}_1 + \dots \underbrace{c_b}_{=0}\tilde{\mathbf{w}}_b \\
&= a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k.
\end{aligned}$$

Da die Vektoren $\mathbf{v}_1, \dots, v_k$ linear unabhängig sind muss darum auch $a_1 = \cdots = a_k = 0$ gelten. Damit haben wir gezeigt, dass die Gleichung

$$a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k + c_1\tilde{\mathbf{w}}_1 + \dots c_b\tilde{\mathbf{w}}_b = 0$$

nur die triviale Lösung $a_1 = \cdots = a_k = c_1 = \cdots = c_b = 0$ hat, die Vektoren also linear unabhängig sind.

GANZ ÄHNLICH können wir auch zeigen, dass die Vektoren ein Erzeugendensystem bilden:

Sei $\mathbf{v} \in K^n$. Dann ist $A\mathbf{v} \in \mathrm{Bild}(A)$. Da die Vektoren $\mathbf{w}_j$ ein Erzeugendensystem von $\mathrm{Bild}(A)$ sind können wir also $A\mathbf{v} = c_1\mathbf{w}_1 + \cdots + c_b\mathbf{w}_b$ als Linearkombination der $\mathbf{w}_j$ schreiben.

Dann gilt

$$\begin{aligned}
A\mathbf{v} &= c_1\mathbf{w}_1 + \cdots + c_b\mathbf{w}_b \\
&= c_1A\tilde{\mathbf{w}}_1 + \cdots + c_b \dots A\tilde{\mathbf{w}}_b \\
&= A(c_1\tilde{\mathbf{w}}_1 + \dots c_b\tilde{\mathbf{w}}_b)
\end{aligned}$$

Also gilt

$$A(\mathbf{v} - (c_1\tilde{\mathbf{w}}_1 + \dots c_b\tilde{\mathbf{w}}_b)) = A\mathbf{v} - A\mathbf{v} = 0,$$

d.h. der Vektor $\mathbf{v} - (c_1\tilde{\mathbf{w}}_1 + \dots c_b\tilde{\mathbf{w}}_b) \in \mathrm{Ker}(A)$ ist ein Element des Kerns von A, lässt sich also als Linearkombination der $\mathbf{v}_i$ schreiben:

$$\begin{aligned}\mathbf{v} - (c_1\tilde{\mathbf{w}}_1 + \dots c_b\tilde{\mathbf{w}}_b) &= a_1\mathbf{v}_1 + \dots + a_k\mathbf{v}_k \\ \Leftrightarrow \quad \mathbf{v} &= a_1\mathbf{v}_1 + \dots + a_k\mathbf{v}_k + c_1\tilde{\mathbf{w}}_1 + \dots c_b\tilde{\mathbf{w}}_b.\end{aligned}$$

Also lässt sich jeder Vektor $\mathbf{v}$ als Linearkombination der $\mathbf{v}_1, \dots, \mathbf{v}_k, \tilde{\mathbf{w}}_1, \dots, \tilde{\mathbf{w}}_b$ schreiben, d.h. die Vektoren sind ein Erzeugendensystem von K^n. □

Die Dimensionsformel erklärt einige Muster, die Sie beim Lösen von Gleichungssystemen beobachtet haben:

Die Faustregel, dass für m Gleichungen in n Variablen $n - m$ freie Variable übrig bleiben werden, ist nur dann richtig, wenn die Spalten der Matrix linear unabhängig sind, für eine zufällig gewählte Matrix wird das meist der Fall sein. In speziellen Beispielen können mehr freie Variable übrig bleiben, jedoch niemals weniger.

Die Anzahl der freien Variablen am Ende des Gauß-Algorithmus können wir jetzt als Dimension des Kerns verstehen, diese hängt insbesondere nicht davon ab, mit welcher Methode wir Lösungen bestimmen.

Die Dimensionsformel ist außerdem praktisch, um einzusehen, ob wir eine Basis von Bild oder Kern gefunden haben, denn wenn wir die Dimension bereits kennen, und wir wissen von einer Liste von $\dim U$ Vektoren, dass diese entweder linear unabhängig sind oder ein Erzeugendensystem von U bilden, so folgt die andere Eigenschaft dann automatisch.

Der Rang einer Matrix

Eine weitere Merkwürdigkeit bei Gleichungssystemen war, dass wir für die Bestimmung von Lösungen Zeilenumformungen durchführen konnten, für die Bestimmung der Dimension des Bildes, aber Spaltenumformungen verwendet haben.

Behauptung 34 (Zeilenrang=Spaltenrang)**.** *Ist A eine $m \times n$ Matrix mit Koeffizienten in einem Körper K, so ist die maximale Anzahl linear unabhängiger Spalten der Matrix gleich der maximalen Anzahl linear unabhängiger Zeilen der Matrix.*

Beweis. Dieses Resultat folgt wieder aus der Dimensionsformel: Die maximale Anzahl linear unabhängiger Spalten der Matrix ist die Dimension des Bildes $\mathrm{Bild}(A) = n - \dim \mathrm{Ker}(A)$.

Der Gauß-Algorithmus bestimmt die Anzahl der linear unabhängigen Zeilen. Bleiben dort r Zeilen $\neq 0$ übrig, so haben wir $n - r$ freie Variable, also ist die Dimension des Kerns $\dim \mathrm{Ker}(A) = n - r$.

Damit ist dann aber

$$\dim \text{Bild}(A) = n - \dim \text{Ker}(A) = n - (n-r) = r.$$

Die Zahl r war genau die Anzahl linear unabhängiger Zeilen, $\dim \text{Bild}(A)$ die Anzahl linear unabhängiger Spalten. Das zeigt also die Behauptung. □

Die Anzahl der linear unabhängigen Zeilen/Spalten einer Matrix bekommt einen Namen.

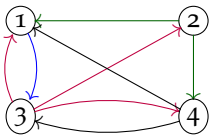
Definition 35. Ist A eine $m \times n$ Matrix mit Koeffizienten in einem Körper K, so heißt

$$\text{Rang}(A) := \dim(\text{Bild}(A))$$

der *Rang* der Matrix.

Anwendung von Zeilenrang=Spaltenrang

Mit diesem Resultat können wir jetzt eine Frage aus der ersten Vorlesung beantworten. Wir hatten für den PageRank-Algorithmus aus einem Netzwerk



eine Gleichungssystem gemacht, in dem die Variablen mit den Knoten des Netzwerks nummeriert waren, die Einträge der zugehörigen Matrix waren:

$$A_{Netzwerk} = (a_{i,j})_{i,j=1,\dots,n} \text{ mit } a_{i,j} = \frac{(\text{Anzahl Links von } j \to i)}{(\text{Anzahl Links von } j)}.$$

Im Beispiel ist

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

Falls von einem Knoten j gar keine Links ausgehen, hatten wir künstlich Links von j zu allen anderen Knoten hinzugefügt. Um die Relevanz der verschiedenen Knoten zu berechnen, hatten wir eine Lösung der Gleichung

$$A\mathbf{v} = \mathbf{v} \text{ mit } \mathbf{v} \in K^N \setminus \{0\}$$

gesucht. Das ist äquivalent dazu, eine Lösung von

$$A\mathbf{v} - \mathbf{v} = 0$$

zu finden. Wenn wir die $n \times n$ Matrix $\text{id}_n := (\delta_{i,j})$ als $\delta_{i,j} := \left\{ \begin{smallmatrix} 1 \text{ wenn } i=j \\ 0 \text{ wenn } i \neq 0 \end{smallmatrix} \right\}$ definieren, bedeutet

$$\text{id}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$A\mathbf{v} - \mathbf{v} = A\mathbf{v} - \text{id}_n\,\mathbf{v} = (A - \text{id}_n)\mathbf{v}.$$

FRAGE: Wieso sollte $(A - \text{id}_n)\mathbf{v} = 0$ immer eine Lösung $\mathbf{v} \neq 0$ haben, d.h. warum gilt $\text{Rang}(A - \text{id}_n) < n$?

ANTWORT Die Spaltensummen $s_j := a_{1,j} + a_{2,j} + \cdots + a_{n,j}$ von A ist immer 1, da

$$a_{1,j} + a_{2,j} + \cdots + a_{n,j} = \frac{(\text{Anzahl Links von } j \to 1)}{(\text{Anzahl Links von } j)} + \cdots + \frac{(\text{Anzahl Links von } j \to n)}{(\text{Anzahl Links von } j)} = 1.$$

Im Beispiel ist

$$B = A - \mathrm{id}_4 = \begin{pmatrix} -1 & \frac{1}{2} & \frac{1}{3} & 0 \\ 0 & -1 & \frac{1}{3} & 0 \\ 1 & 0 & -1 & 1 \\ 0 & \frac{1}{2} & \frac{1}{3} & -1 \end{pmatrix}$$

und

$$B^t = \begin{pmatrix} -1 & 0 & 1 & 0 \\ \frac{1}{2} & -1 & 0 & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & -1 & \frac{1}{3} \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

Also sind alle Spaltensummen der Matrix $B := A - \mathrm{id}_n$ gleich 0. Wir schreiben $B = (b_{i,j})_{i,j=1,\ldots,n}$.

Vertauschen wir in B die Zeilen und Spalten, d.h. wir definieren die Matrix $B^t = (c_{ij})$ mit Einträgen $c_{i,j} := b_{j,i}$ in denen wir Spalten und Zeilenindices vertauschen – B^t heißt *transponierte Matrix* – so hat B^t die Zeilensummen 0, d.h.

$$B^t \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = 0.$$

Also gilt $\mathrm{Rang}(B^t) < n$. Da Zeilenrang und Spaltenrang übereinstimmen gilt aber

$$\mathrm{Rang}(B^t) = \mathrm{Rang}(B).$$

Also ist auch $\mathrm{Rang}(B) < n$.

Wir haben hiermit eine Methode gefunden, von einer Matrix zu zeigen, dass diese einen nichttrivialen Kern hat, ohne dass wir eine konkrete Lösung angegeben hätten.

Lassen Sie uns das Argument formal zusammenfassen:

Folgerung 36 (Stochastische Matrizen)**.** *Ist A eine $n \times n$ Matrix mit Einträgen in einem Körper K, so dass die Summe der Einträge der Spalten von A gleich 1 ist, so existiert ein $\mathbf{v} \in K^n$ mit $\mathbf{v} \neq 0$ für das*

$$A\mathbf{v} = \mathbf{v}$$

gilt.

Bemerkung. Matrizen mit Spaltensumme 1 und Einträgen in $\mathbb{R}_{\geq 0}$ heißen *stochastische Matrizen.* Diese treten auf, wenn Sie Prozesse mit n möglichen Zuständen modellieren, die mit einer gewissen Wahrscheinlichkeit von einem Zustand i in einen anderen Zustand j wechseln, tragen Sie diese Übergangswahrscheinlichkeiten in eine Matrix A ein

$$a_{ij} = \text{Wahrscheinlichkeit dass Zustand } j \text{ in Zustand } i \text{ übergeht,}$$

so modelliert $\mathbf{v} \mapsto A \cdot \mathbf{v}$ die Entwicklung des Prozesses.

Eine Lösung der Gleichung $A \cdot \mathbf{v} = \mathbf{v}$ ist dann ein Gleichgewichtspunkt.

Fazit und Berechnung von Inversen

Wir haben jetzt gesehen, wie wir die Lösbarkeit von Gleichungssystemen in Termen des Rangs der zugehörigen Matrix erkennen können:

Folgerung 37. *Sei A eine $m \times n$-Matrix, dann gilt:*

1. *Das Gleichungssystem $A\mathbf{v} = \mathbf{b}$ ist genau dann für alle $\mathbf{b}$ lösbar, wenn* $\mathrm{Rang}(A) = m$.

2. *Das Gleichungssystem $A\mathbf{v} = \mathbf{0}$ hat genau dann eine eindeutige Lösung, wenn* $\mathrm{Rang}(A) = n$.

3. *Das Gleichungssystem $A\mathbf{v} = \mathbf{b}$ ist genau dann für alle $\mathbf{b}$ eindeutig lösbar, wenn $n = m = \mathrm{Rang}(A)$.*

Beweis. 1. Das Gleichungssystem $A\mathbf{v} = \mathbf{b}$ ist nach Definition des Bildes, genau dann für alle $\mathbf{b}$ lösbar, wenn $\mathrm{Bild}(A) = K^m$ gilt. Das ist nach Aufgabe 3 Blatt 6 gleichbedeutend mit $\dim(\mathrm{Bild}(A)) = m$. Da $\mathrm{Rang}(A) = \dim(\mathrm{Bild}(A))$ nach Definition gilt, folgt die Behauptung.

2. Das Gleichungssystem $A\mathbf{v} = \mathbf{0}$ hat genau dann eine eindeutige Lösung, wenn die Spalten der Matrix A linear unabhängig sind, d.h wenn der Spaltenrang von A gleich n ist.

3. Das Gleichungssystem $A\mathbf{v} = \mathbf{b}$ ist nach 1. genau dann für alle $\mathbf{b}$ lösbar, wenn $\mathrm{Rang}(A) = m$ gilt, je zwei Lösungen unterscheiden sich um ein Element von $\mathrm{Ker}(A)$. Nach 2. ist $\mathrm{Ker}(A) = \{0\}$ gleichbedeutend mit $\mathrm{Rang}(A) = n$. Also ist das Gleichungssystem genau dann für alle $\mathbf{b}$ eindeutig lösbar, wenn $m = \mathrm{Rang}(A) = n$ gilt.

□

BERECHNUNG DER INVERSEN. Ist A eine $n \times n$-Matrix vom Rang n, so können wir für das Gleichungssystem $A\mathbf{v} = \mathbf{b}$ eine Lösungsformel angeben.

Wenn wir nämlich für die Standardbasis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ von K^n Lösungen $\mathbf{v}_1, \ldots, \mathbf{v}_n$ der Gleichung $A\mathbf{v}_i = \mathbf{e}_i$ finden, dann ist für einen beliebigen Vektor $\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ der Vektor $\mathbf{v} := b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n$ eine Lösung von $A\mathbf{v} = \mathbf{b}$, da

$$\begin{aligned} A\mathbf{v} &= A(b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n) && \text{Definition von } \mathbf{v} \\ &= b_1A\mathbf{v}_1 + \cdots + b_nA\mathbf{v}_n && \text{Rechenregel für } A\mathbf{v} \\ &= b_1\mathbf{e}_1 + \cdots + b_n\mathbf{e}_n && \text{weil } A\mathbf{v}_i = \mathbf{e}_i \\ &= \mathbf{b}. \end{aligned}$$

Wenn wir also $A^{-1} := (\mathbf{v}_1 \ldots \mathbf{v}_n)$ als die Matrix mit den Spalten $\mathbf{v}_1, \ldots, \mathbf{v}_n$ definieren, gilt

$$A^{-1} \cdot \mathbf{b} = b_1\mathbf{v}_1 + \cdots + b_n\mathbf{v}_n = \mathbf{v}.$$

Die Matrix A^{-1} liefert uns also die gewünschte Lösungsformel.

DAS KÖNNEN WIR MIT DEM GAUSS-ALGORITHMUS BERECHNEN indem wir die Gleichungen $A\mathbf{v}_i = \mathbf{e}_i$ simultan lösen. Lassen Sie uns das an einem Beispiel durchführen:

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

Schreibe die erweiterte Koeffizientenmatrix:

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ 1 & 2 & 2 & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ 1 & 1 & 2 & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{array}\right)$$

auf und wende das Gauß-Verfahren an:

$$\begin{pmatrix} 1 & 1 & 1 & | & 1 & 0 & 0 \\ \mathbf{1} & 2 & 2 & | & 0 & 1 & 0 \\ \mathbf{1} & 1 & 2 & | & 0 & 0 & 1 \end{pmatrix} \begin{array}{l} \\ \leftarrow \\ \leftarrow \end{array} \begin{array}{l} \\ \cdot -1 \\ \cdot -1 \end{array} \begin{array}{l} \\ + \\ + \end{array}$$

$$\begin{pmatrix} 1 & 1 & \mathbf{1} & | & 1 & 0 & 0 \\ 0 & 1 & \mathbf{1} & | & -1 & 1 & 0 \\ 0 & 0 & 1 & | & -1 & 0 & 1 \end{pmatrix} \begin{array}{l} \leftarrow \\ \leftarrow \\ \\ \end{array} \begin{array}{l} \cdot -1 \\ \\ \cdot -1 \end{array} \begin{array}{l} + \\ + \\ \end{array}$$

$$\begin{pmatrix} 1 & \mathbf{1} & 0 & | & 2 & 0 & -1 \\ 0 & 1 & 0 & | & 0 & 1 & -1 \\ 0 & 0 & 1 & | & -1 & 0 & 1 \end{pmatrix} \begin{array}{l} \leftarrow \\ \\ \\ \end{array} \begin{array}{l} \cdot -1 \\ \\ \\ \end{array} \begin{array}{l} + \\ \\ \\ \end{array}$$

$$\begin{pmatrix} 1 & 0 & 0 & | & 2 & -1 & 0 \\ 0 & 1 & 0 & | & 0 & 1 & -1 \\ 0 & 0 & 1 & | & -1 & 0 & 1 \end{pmatrix} \begin{array}{l} \leftarrow \\ \\ \\ \end{array} \begin{array}{l} \cdot -1 \\ \\ \\ \end{array} \begin{array}{l} + \\ \\ \\ \end{array}$$

Also ist

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

BEISPIEL: Für $b = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ ist

$$A^{-1} \cdot b = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}$$

und dieser Vektor erfüllt in der Tat $A \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$.

PROBE: Wir können die Rechnung auch simultan für alle Vektoren prüfen, indem wir

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

berechnen – hierbei haben wir die Multiplikation von Matrizen einfach so definiert, dass wir $A \cdot$ Spalten von A^{-1} nebeneinander schreiben.

Bemerkung (Matrizenmultiplikation)**.** In der Probe haben wir gesehen, dass wir eine $m \times n$ Matrix A mit einer $n \times \ell$ Matrix B multiplizieren können, indem wir die ℓ Spalten von B mit A multiplizieren und die Ergebnisse als Spalten in eine $m \times \ell$-Matrix schreiben.

Lassen Sie uns nach diesem konkreten Verfahren noch einmal auf die Argumente zur Dimension und Unterräumen zurückkommen.

Allgemeine Vektorräume

Wenn wir die Argumente zur Dimension von Unterräumen, zu Basen und dem Beweis der Dimensionsformel anschauen, kommen Vektoren dort nur noch als Symbole $\mathbf{v}_i$ vor, nicht mehr als explizite Spaltenvektoren. Die Rechenoperationen die wir verwenden sind die Addition von Vektoren $+$ und in den Ausdrücken der Form $c \cdot \mathbf{v}$ die Multiplikation mit Elementen von K.

Diese Notation kennen Sie aus anderen Beispielen:

1. Sind $f, g\colon \mathbb{R} \to \mathbb{R}$ reelle Funktionen, so ist

$$\begin{aligned} f+g\colon \mathbb{R} &\to \mathbb{R} \\ x &\mapsto (f+g)(x) := f(x) + g(x) \end{aligned}$$

   wieder eine reelle Funktion und genauso ist für $c \in \mathbb{R}$ auch $(c \cdot f)(x) := c \cdot f(x)$ eine reelle Funktion.

   Das funktioniert für Funktionen auf einem Intervall $f\colon [a,b] \to \mathbb{R}$ ganz genauso.

2. Sind $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ zwei Folgen reeller Zahlen, so können wir diese ebenfalls addieren, oder mit einer Zahl multiplizieren.

Genau wie bei Körpern, lohnt es sich, die wesentlichen Grundeigenschaften der Operationen formal zu beschreiben, denn dann können wir unsere Resultate über Untervektorräume auch auf die obigen Beispiele anwenden.

Definition (Vektorraum)**.** Sei K ein Körper. Ein *K-Vektorraum* ist eine Menge V zusammen mit zwei Verknüpfungen:

$$\begin{aligned} +\colon V \times V &\to V \\ (\mathbf{v}, \mathbf{v}') &\mapsto \mathbf{v} + \mathbf{v}' \qquad\qquad \text{und} \\ \cdot\colon K \times V &\to V \\ (c, \mathbf{v}) &\mapsto c \cdot \mathbf{v} \end{aligned}$$

die die folgenden Eigenschaften erfüllen:

Vo (neutrales Element für +) Es gibt ein Element $0 \in V$ so dass für alle $\mathbf{v} \in V$ gilt, dass $\mathbf{v} + 0 = \mathbf{v}$.

Jeder Vektorraum sollte einen 0-Vektor enthalten.

VKom (Kommutativgesetz) Für alle $\mathbf{v}, \mathbf{v}' \in V$ gilt $\mathbf{v} + \mathbf{v}' = \mathbf{v}' + \mathbf{v}$.

VAss (Assoziativgesetz) Für alle $\mathbf{v}, \mathbf{v}', \mathbf{v}'' \in V$ gilt $(\mathbf{v} + \mathbf{v}') + \mathbf{v}'' = \mathbf{v} + (\mathbf{v}' + \mathbf{v}'')$.

VInv (inverse Elemente) Zu jedem Element $\mathbf{v} \in V$ existiert ein Element $-\mathbf{v} \in V$ so dass $\mathbf{v} + (-\mathbf{v}) = 0$.

S1 (neutrales Element für $\cdot$) Für alle $\mathbf{v} \in V$ gilt, dass $1 \cdot \mathbf{v} = \mathbf{v}$.

SAss (Assoziativgesetz) Für alle $a, b \in K$ und $\mathbf{v} \in V$ gilt

$$(a \cdot b) \cdot \mathbf{v} = a \cdot (b \cdot \mathbf{v}).$$

In S1 und SAss ist die Rechenregel $\mathbf{v} = (c^{-1} \cdot c) \cdot \mathbf{v} = c^{-1} \cdot (c \cdot \mathbf{v})$ versteckt.

Dist1 (Distributivgesetz) Für alle $a, b \in K$ und $\mathbf{v} \in V$ gilt

$$(a + b) \cdot \mathbf{v} = a \cdot \mathbf{v} + b \cdot \mathbf{v}.$$

Dist2 (Distributivgesetz) Für alle $a \in K$ und $\mathbf{v}, \mathbf{v}' \in V$ gilt

$$a \cdot (\mathbf{v} + \mathbf{v}') = a \cdot \mathbf{v} + a \cdot \mathbf{v}'.$$

Die Axiome für die Addition $+$ sind die gleichen, die wir auch für Körper haben, bei der Verknüpfung $\cdot$ müssen wir aufpassen, dass wir es nun mit zwei unterschiedlichen Argumenten (einer Zahl und einem Vektor) zu tun haben, daher müssen wir jetzt zwei Distributivgesetze formulieren, die sehr ähnlich aussehen, aber einmal das $+$ in K und einmal das $+$ in V verwenden.

BEISPIELE für Vektorräume haben wir schon gesehen:

1. Der K^n mit den üblichen Verknüpfungen $+, \cdot$ war das Beispiel, dass die Definition motiviert hat.
2. Vektorräume von Abbildungen

   $$\mathrm{Abb}([a, b], \mathbb{R}) := \{f \colon [a, b] \to \mathbb{R} \mid f \text{ Abbildung}\}$$

   sind mit den Verknüpfungen vom Anfang des Kaptiels ein Vektorraum.
3. Folgen $\mathrm{Folgen}_{\mathbb{R}} := \{(a_n)_{n \in \mathbb{N}} \mid a_n \in \mathbb{R} \text{ für alle } n \in \mathbb{N}\}$ bilden ebenso einen Vektorraum.

DIE VEKTORRÄUME von Folgen und Abbildungen, unterscheiden sich vom K^n dadurch, dass wir keine endliche Basis finden können – vielleicht finden Sie eine unendliche Liste linear unabhängiger Elemente? Für Unterräume spezieller Folgen oder spezieller Abbildungen kann das anders aussehen.

Übersetzung der Grundbegriffe im K^n für allgemeine Vektorräume

Die Definitionen von linear unabhängigen Vektoren, dem Spann einer Teilmenge, von Unterräumen des K^n, oder von Basen von Unterräumen können wir genauso für einen abstrakten Vektorraum V verwenden:

Definition. Eine Menge von Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_r \in K^n$ heißt *linear unabhängig*, wenn die Gleichung

$$a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_r\mathbf{v}_r = 0$$

in K nur die triviale Lösung $a_1 = a_2 = \cdots = a_r = 0$ besitzt.

Beispiele:

1. Die Funktionen $f(x) = x, g(x) = x + 1, h(x) = x^2 \in \mathrm{Abb}(\mathbb{R}, \mathbb{R})$ sind linear unabhängig, denn ist

   $$a_1 f + a_2 g + a_3 h = 0$$

   die Null-Abbildung, so gilt

   $0 = a_1 x + a_2(x+1) + a_3(x^2) = a_2 + (a_1 + a_2)x + a_3 x^2$ für alle $x \in \mathbb{R}$.

   Für $x = 0$ bedeutet das $a_2 = 0$ und damit

   $$0 = a_1 x + a_3 x^2 = x(a_1 + a_3 x)$$

   für alle x. Also gilt insbesondere (setze $x = \pm 1$), dass $a_1 + a_3 = 0 = a_1 - a_3$, aber daraus folgt $a_2 = a_3 = 0$.

2. Die Funktionen $f(x) = x, g(x) = x + 1, h(x) = x - 1$ sind in $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$ linear abhängig, da

   $$f(x) - g(x) = h(x) - f(x) \text{ also } 2f(x) - g(x) - h(x) = 0$$

   gilt.

3. Die Elemente $1, i \in \mathbb{C}$ sind im $\mathbb{R}$-Vektorraum $\mathbb{C}$ linear unabhängig, da wir jede komplexe Zahl eindeutig als $a + b \cdot i$ schreiben können, im $\mathbb{C}$-Vektorraum $\mathbb{C}$ sind die Elemente linear abhängig, da $1 + i \cdot i = 0$ gilt.

   Der Körper K spielt also in der Definition von K-Vektorräumen eine wichtige Rolle.

Notation 38. Für gegebene Vektoren $\mathbf{v}_1, \ldots, \mathbf{v}_r \in V$ bezeichnen wir mit

$$\mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_r) := \{\mathbf{v} = a_1\mathbf{v}_1 + \ldots a_r\mathbf{v}_r \,|\, a_1, \ldots, a_r \in K\} \subseteq V$$

den Spann (oder die lineare Hülle) der Vektoren.

Per Konvention ist der Spann der leeren Menge $\mathrm{Span}() := \mathbf{0} \subset V$ der 0-Vektor.

Definition. Eine Liste $\mathbf{v}_1, \ldots, \mathbf{v}_d \in V$ heißt Basis des K-Vektorraums V, wenn $\mathbf{v}_1, \ldots, \mathbf{v}_d$ ein linear unabhängiges Erzeugendensystem von V ist.

Die Dimension $\dim_K V$ von V ist die Anzahl der Elemente einer Basis.

Ein Vektorraum V heißt *endlich-dimensional* wenn $\dim V$ eine endliche Basis $\mathbf{v}_1, \ldots, \mathbf{v}_d \in V$ besitzt.

Die $\mathbb{R}$-Vekotrräume $\mathrm{Abb}(\mathbb{R},\mathbb{R})$ und $\mathrm{Folgen}_{\mathbb{R}}$ sind nicht endlich-dimensional, wir werden aber sehen, dass diese Räume viele interessante, endlich-dimensionale Unterräume haben.

Bemerkung. Wie zuvor sind $\mathbf{v}_1,\dots,\mathbf{v}_d \in V$ genau dann eine Basis von V, wenn sich jedes Element $\mathbf{v} \in V$ *eindeutig* als Linearkombination

$$\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_d\mathbf{v}_d$$

mit Koeffizienten $a_i \in K$ schreiben lässt, denn der Beweis dieser Aussage für Unterräume des K^n hatte nur die Rechenreglen für Vektoren verwendet, nicht die Eigenschaft dass die Vektoren Elemente von K^n sind.

Genauso hat der Beweis des Austauschsatzes und damit auch der Aussage, dass je zwei Basen die gleiche Anzahl von Elementen haben, nur die Rechenoperationen in Vektorräumen verwendet. Darum gelten diese Aussagen auch in allgemeinen Vektorräumen. Sie sollten sich die Beweise noch einmal anschauen, um zu prüfen, dass wir dort tatsächlich K^n durch V ersetzen können.

Definition. Sei V ein K-Vektorraum. Eine *nicht leere* Teilmenge $U \subset V$ heißt *Unterraum* von V, wenn gilt

1. Für jeden Vektor $\mathbf{v} \in U$ sind alle Vielfachen ebenfalls in U, also:
$$(\mathbf{v} \in U, c \in K) \Rightarrow c \cdot \mathbf{v} \in U.$$
2. Für je zwei Vektoren $\mathbf{v}, \mathbf{v}' \in U$ ist auch ihre Summe $\mathbf{v} + \mathbf{v}' \in U$, also:
$$(\mathbf{v}, \mathbf{v}' \in U) \Rightarrow \mathbf{v} + \mathbf{v}' \in U.$$

Beispiel 39. Sei $\mathrm{Fibo} \subset \mathrm{Folgen}_{\mathbb{R}} := \{(a_n)_{n\in\mathbb{N}} \in \mathbb{R}\}$ die Teilmenge der Fibonacci-Folgen, d.h.

$$\mathrm{Fibo} = \{(a_n)_{n\in\mathbb{N}} \in \mathbb{R} \,|\, a_{n+2} = a_{n+1} + a_n \text{ für alle } n\}$$

ist die Teilmenge der Folgen, für die jedes Folgenglied die Summe der beiden vorhergehenden Folgenglieder ist.

Zum Beispiel sind also die Folgen $(G_n) = (1,1,2,3,5,8,13,21,\dots)$ und $(F_n) = (0,1,1,2,3,5,8,13,\dots)$ Elemente von Fibo. Die Folge F_n heißt die Folge der Fibonacci-Zahlen. Diese wird Ihnen in der Natur, Musik und Kunst viel begegnen, wenn Sie die Augen offen halten.

Diese Teilmenge ist ein Unterraum, da

1. Wenn $\mathbf{v} = (a_n) \in \mathrm{Fibo}, c \in \mathbb{R}$, dann ist auch $c\mathbf{v} = (ca_n)_{n\in\mathbb{N}} \in \mathrm{Fibo}$, da
$$ca_{n+2} \overset{(a_n)\in\mathrm{Fibo}}{=} c(a_{n+1} + a_n) = ca_{n+1} + c_{a_n}$$
gilt.
2. Wenn $(a_n), (b_n) \in \mathrm{Fibo}$, so auch $(a_n + b_n)_{n\in\mathbb{N}}$, da wieder gilt
$$a_{n+2} + b_{n+2} \overset{(a_n),(b_n)\in\mathrm{Fibo}}{=} a_{n+1} + a_n + b_{n+1} + b_n = (a_{n+1} + b_{n+1}) + (a_n + b_n).$$

FRAGE: Was ist die Dimension $\dim_{\mathbb{R}}$ Fibo ? Können Sie eine Basis angeben?

ANTWORT: $\dim_{\mathbb{R}} \text{Fibo} = 2$, die beiden Folgen $(G_n) = (1,1,2,3,5,8,13,21,\dots)$ und $(F_n) = (0,1,1,2,3,5,8,13,\dots)$ sind linear unabhängig, also eine Basis.

Das gilt, da durch die Bedingung $a_{n+2} = a_{n+1} + a_n$ für alle n jedes Element von Fibo eindeutig durch die ersten beiden Folgenglieder a_1, a_2 bestimmt ist. Da die ersten Folgenglieder $(1,1)$ und $(0,1)$ von (G_n) und (F_n) in $\mathbb{R}^2$ linear unabhängig sind, lässt sich jede Folge $(a_n) \in \text{Fibo}$ eindeutig als Linearkombination der Folgen $(F_n), (G_n)$ schreiben.

Der Unterraum Fibo ist ein gutes Beispiel für einen Unterraum mit unterschiedlichen, interessanten Basen. Wenn Sie mit ähnlichen Bedingungen

$$a_{n+2} = ca_{n+1} + da_n$$

für unterschiedliche $c, d \in \mathbb{R}$ spielen, werden Sie feststellen, dass die Folgenglieder etwa exponentiell wachsen, d.h. etwa wie die Folge x^n für ein $x \in \mathbb{R}$ größer werden.

In Fibo könnten wir darum nach einer Folge (a_n) mit $a_n = x^{n-1}$ suchen. Wenn es so ein x gibt, so muss das die Gleichungen

Wir nehmen $a_n = x^{n-1}$ damit die Folge mit $a_1 = x^0 = 1$ anfängt. Wir könnten genauso gut $a_n = x^n$ versuchen, aber dann sind die Formeln am Ende etwas weniger schön.

$$\begin{aligned} x^2 &= x + 1 \\ x^3 &= x^2 + x = x(x+1) \\ x^4 &= x^3 + x^2 = x^2(x+1) \\ \vdots &= \vdots \end{aligned}$$

erfüllen, also muss entweder $x = 0$ sein oder x ist eine Nullstelle von $x^2 - x - 1$, d.h

Die Zahl $\phi = \frac{1+\sqrt{5}}{2}$ heißt *goldener Schnitt* und kommt in der Natur häufig vor. Sie hat die schönen Eigenschaften $\phi + 1 = \phi^2$ und $\phi - 1 = \frac{1}{\phi}$.

Wenn Sie ein Blatt mit Seitenverhältnis ϕ nehmen und davon ein Quadrat mit Kantenlänge der kürzeren Seite abschneiden, hat der Rest wieder Seitenverhältnis ϕ.

$$\phi_\pm = \pm\sqrt{1 + \frac{1}{4}} + \frac{1}{2} = \frac{\pm\sqrt{5}+1}{2} = \frac{1 \pm \sqrt{5}}{2}.$$

Damit haben wir zwei neue linear unabhängige Vektoren

$$\mathbf{v}_1 = (1, \phi_+, \phi_+^2, \dots), \mathbf{v}_2 = (1, \phi_-, \phi_-^2, \phi_-^3, \dots) \in \text{Fibo}$$

gefunden.

Diese beiden Vektoren sind wieder linear unabhängig, da die Vektoren $\begin{pmatrix} 1 \\ \phi_+ \end{pmatrix}$ und $\begin{pmatrix} 1 \\ \phi_- \end{pmatrix}$ der ersten beiden Folgenglieder in $\mathbb{R}^2$ linear unabhängig sind. Damit müssen $\mathbf{v}_1, \mathbf{v}_2$ aber eine neue Basis von Fibo sein, es gibt also $a, b \in \mathbb{R}$ so dass

$$\begin{aligned} (0,1,1,\dots) &= a\mathbf{v}_1 + b\mathbf{v}_2 \\ &= a \cdot (1, \phi_+, \phi_+^2, \dots) + b \cdot (1, \phi_-, \phi_-^2, \phi_-^3, \dots) \\ &= (a+b, a\phi_+ + b\phi_-, \dots). \end{aligned}$$

Damit muss $a + b = 0$, also $b = -a$ gelten und für das 2. Folgenglied

muss dann gelten:

$$\begin{aligned} 1 &= a\phi_+ + b\phi_- \\ &= a\phi_+ - a\phi_- \\ &= a(\phi_+ - \phi_) = a(\frac{1+\sqrt{5}}{2} - \frac{1-\sqrt{5}}{2}) \\ &= a\sqrt{5}. \end{aligned}$$

Also ist $a = \frac{1}{\sqrt{5}}$. Also gilt für die n-te Zahl F_n

$$\begin{aligned} F_n &= \frac{1}{\sqrt{5}}(\phi_+^{n-1} - \phi_-^{n-1}) \\ &= \frac{1}{\sqrt{5}}\left(\left(\frac{1+\sqrt{5}}{2}\right)^{n-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n-1}\right) \end{aligned}$$

Das ist eine bemerkenswerte Formel, die einerseits das Wachstum der Fibonacci-Zahlen gut erklärt und andererseits einen guten Blick für Formeln benötigt, um zu sehen, dass die Formel überhaupt ganze Zahlen liefert.

EIN WEITERES BEISPIEL eines K-Vektorraums, der interessante Unterräume enthält ist der Vektorraum der Polynome mit Koeffizienten in einem Körper K. Ein Polynom (in einer Variablen x) ist ein Ausdruck der Form

$$p(x) = a_0 + a_1 x + \cdots + a_n x^n$$

wobei $n \in \mathbb{Z}_{\geq 0}$ eine nicht-negative ganze Zahl ist und die Koeffizienten $a_i \in K$ Elemente unseres Körpers K sind. Die übliche Konvention ist, dass wir die Terme in denen die Koeffizienten 0 sind, weglassen dürfen, und $1 \cdot x^n$ als x^n lesen, d.h.

$$1 + 2 \cdot x + 1 \cdot x^2 + 0 \cdot x^3 = 1 + 2x + x^2.$$

Die Menge aller Polynome mit Koeffizienten in K bezeichnen wir mit

$$K[x] := \{p(x) = a_0 + a_1 x + \cdots + a_n x^n \mid n \in \mathbb{Z}_{\geq 0}, a_i \in K \text{ für } i = 1, \ldots, n\}.$$

Polynome können wir addieren, indem wir die Koeffizienten addieren

$$(a_0 + a_1 x + \cdots + a_n x^n) + (b_0 + b_1 x + \cdots + b_n x^n) := (a_0 + b_0) + (a_1 + b_1)x + \cdots + (a_n + b_n)x^n$$

und mit Elementen $c \in K$ multiplizieren:

$$c \cdot (a_0 + a_1 x + \cdots + a_n x^n) := (ca_0) + (ca_1)x + \cdots + (ca_n)x^n.$$

Damit wird $K[x]$ ein K-Vektorraum.

Bemerkung. Mir ist wichtig, dass wir Polynome tatsächlich als abstrakte Ausdrücke verstehen, d.h. das Symbol x ist einfach ein Symbol mit dem wir rechnen. Das ist ähnlich wie das Symbol $i \in \mathbb{C}$, nur dass für die Variable x alle Potenzen x^n verschieden sind, die Monome $1, x, x^2, x^3, \ldots$ sind in $K[x]$ linear unabhängig.

Es gibt auch den Begriff der Polynomfunktion $p\colon K \to K$, aber wenn der Körper sehr klein ist, wie zum Beispiel $K = \mathbb{Z}/2\mathbb{Z}$, dann gibt es viel weniger Abbildungen als Polynome, zum Beispiel definiert das Polynom $p(x) = x(x-1) = x^2 - x$, wenn wir für x die beiden Elemente $0, 1$ von $\mathbb{Z}/2\mathbb{Z}$ einsetzen die 0-Abbildung, obwohl das Polynom nicht das 0-Polynom ist.

Der Vektorraum $K[x]$ ist wiederum nicht endlich-dimensional, aber wenn wir den Unterraum der Polynome vom Grad $\leq n$

$$K[x]_{\leq n} := \{p(x) = a_0 + a_1 x + \cdots + a_n x^n \mid a_i \in K \text{ für } i = 1, \ldots, n\}$$

betrachten, so hat dieser die Basis $1, x, \ldots, x^n$, hat also Dimension

$$\dim_K K[x]_{\leq n} = n + 1.$$

Bemerkung. Wir können Polynome natürlich auch multiplizieren, indem wir das Produkt ausmultiplizieren und dann die Terme mit gleichen Exponenten zusammenfassen. Um das als Formel aufzuschreiben ist die Summenschreibweise

Summenzeichen $\sum$

$$\sum_{i=1}^{n} a_i := a_1 + \ldots a_n$$

hilfreich, insbesondere ist

$$a_0 + a_1 x + \cdots + a_n x^n = \sum_{i=0}^{n} a_i x^i.$$

Der Koeffizient von x^ℓ im Produkt

$$(a_0 + a_1 x + \cdots + a_n x^n) \cdot (b_0 + b_1 x + \cdots + b_n x^n)$$

ist in dieser Schreibweise

$$a_\ell b_0 + a_{\ell-1} b_1 + \cdots + a_1 b_{\ell-1} + a_0 b_\ell = \sum_{i=0}^{\ell} a_i b_{\ell-i},$$

wobei wir für $\ell > n$ heimlich $a_j = b_j = 0$ für $j > n$ gesetzt haben. Insgesamt können wir das Produkt dann so schreiben:

$$\begin{aligned}
&(a_0 + a_1 x + \cdots + a_n x^n) \cdot (b_0 + b_1 x + \cdots + b_n x^n) \\
&= a_0 b_0 + \left(\sum_{i=0}^{1} a_i b_{1-i}\right) x^1 + \left(\sum_{i=0}^{2} a_i b_{1-i}\right) x^2 + \cdots + \left(\sum_{i=0}^{2n} a_i b_{2n-i}\right) x^{2n} \\
&= \sum_{\ell=0}^{n} \left(\sum_{i=0}^{\ell} a_i b_{\ell-i}\right) x^\ell
\end{aligned}$$

Keine Sorge, die Formel ist nur eine kompakte Zusammenfassung der Regel, die Sie für die Multiplikation kennen. Die Summenschreibweise sieht zu Beginn etwas ungewohnt aus, ich werde diese darum nur hin und wieder verwenden.

Lineare Abbildungen

Für den K^n hatten $m \times n$-Matrizen A Abbildungen $A\colon K^n \to K^m$ definiert. Diese hatte die Eigenschaften

$$\begin{aligned} A(\mathbf{v}+\mathbf{v}') &= A\mathbf{v} + A\mathbf{v}' && \text{für alle } \mathbf{v}, \mathbf{v}' \in K^n \text{ und} \\ A(c\cdot\mathbf{v}) &= c\cdot A\mathbf{v} && \text{für alle } \mathbf{v}, \in K^n, c \in K. \end{aligned}$$

Diese Eigenschaft haben auch Abbildungen, die Sie auf unseren Beispielvekotrräumen $K[x]$ oder $\text{Folgen}_{\mathbb{R}}$ kennen, zum Beispiel kennen Sie die Ableitung von Polynomen

$$\begin{aligned} (\quad)'\colon K[x] &\to K[x] \\ p(x) &\mapsto p'(x) \\ a_0 + a_1 x + \cdots + a_n x^n &\mapsto a_1 + 2a_2 x + \cdots + n a_n x^{n-1} \\ &= \sum_{i=1}^{n} i \cdot a_{i+1} x^i. \end{aligned}$$

Diese erfüllt auch $(f(x)+g(x))' = f'(x) + g'(x)$ und $(cf(x))' = c(f'(x))$.

Es ist für Sie vielleicht ungewohnt, die Ableitung selbst als Abbildung aufzufassen, aber der Prozess aus einer Funktion eine andere Funktion zu machen ist aus unserer Perspektive ganz ähnlich zum Prozess aus einer Zahl, eine andere Zahl zu berechnen.

Definition (Lineare Abbildungen)**.** Eine Abbildung $F\colon V \to W$ zwischen K-Vektorräumen V, W heißt *linear* genau dann wenn

$$\begin{aligned} F(\mathbf{v}+\mathbf{v}') &= F(\mathbf{v}) + F(\mathbf{v}') && \text{für alle } \mathbf{v}, \mathbf{v}' \in V \text{ und} \\ F(c\cdot\mathbf{v}) &= c\cdot F(\mathbf{v}) && \text{für alle } \mathbf{v} \in V, c \in K. \end{aligned}$$

Bemerkung. Die Bedingung für die Linearität einer Abbildung F könnten wir auch so formulieren, dass wir sagen, dass F mit den Vektorraum-Verknüpfungen $+, \cdot$ verträglich ist, d.h. es ist für lineare Abbildung egal ob wir erst Vektoren in V addieren und dann abbilden, oder ob wir die Vektoren erst abbilden und dann in W addieren, das gleiche gilt für die Multiplikation mit Elementen aus K. Lineare Abbildungen heißen darum auch *Homomorphismen* (strukturerhaltende Abbildungen) von Vektorräumen.

Als Beispiele linearer Abbildungen haben wir gerade die Ableitung und für die Vektorräume K^n, K^m die Abbildungen die durch Matrizen definiert werden gesehen.

Wie für Matrizen können wir Kern und Bild einer linearen Abbildung $F\colon V \to W$ betrachten:

$$\text{Ker}(F) := \{\mathbf{v} \in V \mid F(\mathbf{v}) = 0 \in W\} \subseteq V$$

$$\text{Bild}(F) := \{\mathbf{b} \in W \mid \text{es gibt ein } \mathbf{v} \in V \text{ mit } F(\mathbf{v}) = \mathbf{b}\} \subseteq W.$$

Bemerkung. Für lineare Abbildungen $F\colon V \to W$ sind $\ker(F) \subseteq V$ und $\text{Bild}(F) \subseteq W$ Untervektorräume.

Bitte kontrollieren Sie selbst, dass diese Aussage wirklich stimmt, mit etwas Glück fällt Ihnen das schon nicht mehr so schwer.

Erstaunlicher ist jetzt vielleicht, dass wir im Beweis der Dimensionsformel für Matrizen wiederum nur mit Symbolen $\mathbf{v}_i$ gerechnet haben und dabei nur die einfachen Rechenregeln $A(\mathbf{v}+\mathbf{v}') = A\mathbf{v}+A\mathbf{v}'$ und $A(c\cdot\mathbf{v}) = c\cdot A\mathbf{v}$ verwendet haben. Ersetzen wir im Beweis K^n und K^m durch endlich-dimensionale Vektorräume, so erhalten wir darum eine Dimensionsformel für lineare Abbildungen.

Satz 40 (Dimensionsformel)**.** *Ist V ein endlich-dimensionaler K-Vektorraum und $F\colon V\to W$ eine lineare Abbildung von K-Vektorräumen, so gilt*

$$\dim V = \dim(\mathrm{Bild}(F)) + \dim(\mathrm{Ker}(F)).$$

Beispiel 41. Die Ableitung von Polynomen vom Grad $\leq n$ ist eine lineare Abbildung

$$(\quad)'\colon \mathbb{Q}[x]_{\leq n}\to\mathbb{Q}[x]_{\leq n}.$$

Sie wissen, dass die Ableitung eines Polynoms vom Grad n ein Polynom vom Grad $n-1$ liefert und weil $(x^\ell)' = \ell\cdot x^{\ell-1}$ kommt auch jedes Polynom vom Grad $n-1$ als Ableitung vor, d.h.

$$\mathrm{Bild}((\quad)'\colon \mathbb{Q}[x]_{\leq n}\to\mathbb{Q}[x]_{\leq n}) = \mathbb{Q}[x]_{\leq n-1}.$$

Also gilt $\dim(\mathrm{Bild}((\quad)')) = \dim\mathbb{Q}[x]_{\leq n-1} = n-1$.

Für den Kern überlegen wir uns, dass $p'(x)=0$ bedeutet, dass das Polynom konstant $p(x)$ war, also

$$\mathrm{Ker}((\quad)'\colon \mathbb{Q}[x]_{\leq n}\to\mathbb{Q}[x]_{\leq n}) = \{p(x)=a_0 | a_0\in\mathbb{Q}\} = \mathrm{Span}(1)$$

und darum $\dim(\mathrm{Ker}((\quad)') = 1$. Damit gilt tatsächlich

$$n+1 = \dim\mathbb{Q}[x]_{\leq n} = \dim(\mathrm{Bild}((\quad)')) + \dim(\mathrm{Ker}((\quad)')).$$

Die Dimensionsformel sagt uns, dass wir die Dimension des Kerns schon an der Dimension des Bildes hätten ablesen können. Es ist ganz erfreulich und für Differentialgleichungen nützlich, dass die lineare Algebra auf diese Art Aussagen über die Ableitung finden kann.

Matrizen zu linearen Abbildungen

Der Zusammenhang zwischen linearen Abbildungen und Matrizen ist sehr eng. Wahrscheinlich sind Ihnen im Moment noch Matrizen lieber, weil Ihnen diese konkreter erscheinen. Darum möchte ich kurz erklären, wie wir lineare Abbildungen tatsächlich immer durch Matrizen beschreiben können und damit Fragen zu linearen Abbildungen in Fragen zu Matrizen übersetzten können.

Das Prinzip ist einfach: Ist $F\colon V\to W$ linear und $\mathbf{v}_1,\dots,\mathbf{v}_n\in V$ eine Basis, so bestimmen die Werte $F(\mathbf{v}_1),\dots,F(\mathbf{v}_n)$ die Abbildung

F vollständig, denn wir können jeden Vektor $\mathbf{v} \in V$ eindeutig als Linearkombination

$$\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$$

unserer Basis schreiben. Ist F linear so können wir damit $F(\mathbf{v})$ aus den $F(\mathbf{v}_i)$ berechnen:

$$\begin{aligned} F(\mathbf{v}) &= F(a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) && \text{Formel für } \mathbf{v} \\ &= F(a_1\mathbf{v}_1) + \cdots + F(a_n\mathbf{v}_n) && \text{weil } F(\mathbf{v}+\mathbf{v}') = F(\mathbf{v}) + F(\mathbf{v}') \\ &= a_1F(\mathbf{v}_1) + \cdots + a_nF(\mathbf{v}_n) && \text{weil } F(c\mathbf{v}) = cF(\mathbf{v}). \end{aligned}$$

Dass die Bilder $F(\mathbf{v}_1), \ldots, F(\mathbf{v}_n)$ einer Basis die Abbildung F eindeutig bestimmen, entspricht für Matrizen $A\colon K^n \to K^m$ der Aussage, dass die Spalten der Matrix A die Bilder $A\mathbf{e}_1, \ldots, A\mathbf{e}_n$ der Standardbasis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ sind.

Basen liefern Koordinaten

Die Wahl einer Basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ eines Vektorraums V entspricht der Wahl eines Koordinatensystems, indem wir für jedes $\mathbf{v} \in V$ die Koeffizienten der eindeutigen Linearkombination

$$\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$$

als Koordinaten $\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ auffassen, d.h. wir bekommen eine Koordinatenabbildung

$$\begin{aligned} \mathrm{Koord}_{\underline{\mathbf{v}}}\colon V &\to K^n, \\ \mathbf{v} &\mapsto \mathrm{Koord}_{\underline{\mathbf{v}}}(\mathbf{v}) := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \end{aligned}$$

wobei die Einträge a_i des Spaltenvektors die Koeffizienten der eindeutigen Lösung der Gleichung

$$\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$$

sind.

Beispiel 42. 1. Ist $V = K[t]_{\leq 2}$ der Vektorraum der höchstens quadratischen Polynome und $\mathbf{v}_1 = 1, \mathbf{v}_2 = t, \mathbf{v}_3 = t^2$ die erste Basis, die uns für diesen Vektorraum einfällt, so sind die Koordinaten des Polynoms $p(t) = 3t^2 - 6t + 7$ leicht als

$$\mathrm{Koord}_{\underline{\mathbf{v}}}(3t^2 - 6t + 7) = \begin{pmatrix} 7 \\ (-6) \\ 3 \end{pmatrix}$$

abzulesen, denn

$$\begin{aligned} 3t^2 - 6t + 7 &= 7 \cdot 1 + (-6) \cdot t + 3 \cdot t^2 \\ &= 7 \cdot \mathbf{v}_1 + (-6) \cdot \mathbf{v}_2 + 3 \cdot \mathbf{v}_3. \end{aligned}$$

Hätten wir hingegen die Basis $\mathbf{w}_1 = 1, \mathbf{w}_2 = (t-1), \mathbf{w}_3 = (t-1)^2 = t^2 - 2t + 1$ gewählt, so würden wir, um die Koordinaten bezüglich dieser Basis zu bestimmen, $p(t)$ als Linearkombination dieser Vektoren schreiben, also

$$\begin{aligned} 3t^2 - 6t + 7 &= 3 \cdot (t^2 - 2t + 1) + 4 \cdot 1 \\ &= 4 \cdot \mathbf{w}_1 + 0 \cdot \mathbf{w}_2 + 3 \cdot \mathbf{w}_3 \end{aligned}$$

und erhalten damit den Koordinatenvektor

$$\mathrm{Koord}_{\underline{\mathbf{w}}}(3t^2 - 6t + 7) = \begin{pmatrix} 4 \\ 0 \\ 3 \end{pmatrix}.$$

2. Ist $V = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0 \right\} \subseteq \mathbb{R}^3$ die Ebene der Vektoren, deren Summe der Koordinaten 0 ist, und wählen wir zum Beispiel $\mathbf{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ als Basis, so können wir die Koordinaten des Vektors $\mathbf{v} = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$ berechnen, indem wir den Vektor als Linearkombination der Basis schreiben:

Geometrisch ist das die Ebene senkrecht zum Vektor $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

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

und erhalten dann

$$\mathrm{Koord}_{\underline{\mathbf{v}}}\left(\begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix} \right) = \begin{pmatrix} -2 \\ 3 \end{pmatrix}.$$

Das ist praktisch, weil wir damit für Rechnungen in diesem 2-dimensionalen Teilraum des $\mathbb{R}^3$ nur noch 2 Koordinaten benötigen.

Lassen Sie uns das jetzt allgemein formulieren.

Satz 43 (Basen liefern Koordinaten). *Sei $\underline{\mathbf{v}} := (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ eine Basis eines Vektorraums V. Die Koordinatenabbildungen sind die Abbildungen:*

Notation: Wir schreiben für eine Basis $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ oft kürzer

$$\underline{\mathbf{v}} := (\mathbf{v}_1, \ldots, \mathbf{v}_n).$$

$$\mathrm{Komb}_{\underline{\mathbf{v}}} \colon K^n \to V, \quad \mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \mapsto \mathrm{Komb}_{\underline{\mathbf{v}}}(\mathbf{a}) := a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$$

$$\mathrm{Koord}_{\underline{\mathbf{v}}} \colon V \to K^n, \quad \mathbf{v} \mapsto \mathrm{Koord}_{\underline{\mathbf{v}}}(\mathbf{v}) := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \textit{ wobei die } a_i \textit{ die eindeutige}$$

Lösung von $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$ sind.

Die Koordinatenabbildungen sind lineare Abbildungen, die zueinander invers sind, d.h. es gilt

$$\mathrm{Komb}_{\underline{\mathbf{v}}}(\mathrm{Koord}_{\underline{\mathbf{v}}}(\mathbf{v})) = \mathbf{v} \textit{ für alle } \mathbf{v} \in V \textit{und}$$
$$\mathrm{Koord}_{\underline{\mathbf{v}}}(\mathrm{Komb}_{\underline{\mathbf{v}}}(\mathbf{a})) = \mathbf{a} \textit{ für alle } \mathbf{a} \in K^n.$$

Nach Wahl einer Basis von V entspricht also jedem abstrakten Vektor $\mathbf{v} \in V$ genau ein Spaltenvektor und umgekehrt jedem Spaltenvektor genau ein Element $\mathbf{v} \in V$.

Beweis (Basen liefern Koordinaten): Die letzten Aussagen des Satzes $\mathrm{Komb}_{\underline{\mathbf{v}}}(\mathrm{Koord}_{\underline{\mathbf{v}}}(\mathbf{v})) = \mathbf{v}$ und $\mathrm{Koord}_{\underline{\mathbf{v}}}(\mathrm{Komb}_{\underline{\mathbf{v}}}(\mathbf{a})) = \mathbf{a}$ sind nach Konstruktion klar, denn $\mathrm{Koord}_{\underline{\mathbf{v}}}(\mathbf{v})$ ist genau der Spaltenvektor dessen Einträge die Koeffizienten der eindeutigen Linearkombination $\mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$ sind, also ist $\mathrm{Komb}_{\underline{\mathbf{v}}}(\mathrm{Koord}_{\underline{\mathbf{v}}}(\mathbf{v})) = \mathrm{Komb}_{\underline{\mathbf{v}}}(\mathbf{a}) = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = \mathbf{v}$. Genauso ist $\mathrm{Komb}_{\underline{\mathbf{v}}}(\mathbf{a}) = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n$ und darum nach Konstruktion $\mathrm{Koord}_{\underline{\mathbf{v}}}(a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) = \mathbf{a}$.

Es ist auch einsichtig, dass die Abbildungen linear sind, denn sind $\mathrm{Koord}_{\underline{\mathbf{v}}}(\mathbf{v}) = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \mathrm{Koord}_{\underline{\mathbf{v}}}(\mathbf{v}') = \begin{pmatrix} a_1' \\ \vdots \\ a_n' \end{pmatrix}$ so ist

$$\begin{aligned}\mathbf{v} + \mathbf{v}' &= (a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) + (a_1'\mathbf{v}_1 + \cdots + a_n'\mathbf{v}_n) \\ &= (a_1 + a_1')\mathbf{v}_1 + \cdots + (a_n + a_n')\mathbf{v}_n,\end{aligned}$$

und darum

$$\begin{aligned}\mathrm{Koord}_{\underline{\mathbf{v}}}(\mathbf{v} + \mathbf{v}') &= \begin{pmatrix} a_1 + a_1' \\ \vdots \\ a_n + a_n' \end{pmatrix} \\ &= \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} a_1' \\ \vdots \\ a_n' \end{pmatrix} \\ &= \mathrm{Koord}_{\underline{\mathbf{v}}}(\mathbf{v}) + \mathrm{Koord}_{\underline{\mathbf{v}}}(\mathbf{v}').\end{aligned}$$

Das gleiche Argument funktioniert für $c\mathbf{v}$. Dass die Abbildung $\mathrm{Komb}_{\underline{\mathbf{v}}}$ linear ist, ist genauso klar, denn

$$\begin{aligned}\mathrm{Komb}_{\underline{\mathbf{v}}}\left(\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} a_1' \\ \vdots \\ a_n' \end{pmatrix}\right) &= \mathrm{Komb}_{\underline{\mathbf{v}}}\left(\begin{pmatrix} a_1 + a_1' \\ \vdots \\ a_n + a_n' \end{pmatrix}\right) \\ &= (a_1 + a_1')\mathbf{v}_1 + \cdots + (a_n + a_n')\mathbf{v}_n \\ &= (a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) + (a_1'\mathbf{v}_1 + \cdots + a_n'\mathbf{v}_n) \\ &= \mathrm{Komb}_{\underline{\mathbf{v}}}(\mathbf{a}) + \mathrm{Komb}_{\underline{\mathbf{v}}}(\mathbf{a}').\end{aligned}$$

Wieder funktioniert das gleiche Argument für $c\mathbf{v}$. □

Notation 44. Sind $g\colon X \to Y$ und $f\colon Y \to Z$ Abbildungen, so schreiben wir $f \circ g$ für die Komposition (Hinternanderausführung) von f und g, d.h.

$$\begin{aligned}f \circ g\colon X &\to Z \\ x &\mapsto (f \circ g)(x) := f(g(x)).\end{aligned}$$

Die Verkettung $\circ$ von Abbildungen lesen wir *von rechts nach links*: $f \circ g \circ h$ bedeutet $f(g(h(\_\_)))$ also, wende erst h dann g und dann f an.

Definition. Eine lineare Abbildung $f\colon V \to W$ heißt *Isomorphismus* genau dann, wenn eine zu f inverse Abbildung $g\colon W \to V$ existiert, d.h. es existiert $g\colon W \to V$ mit $g \circ f = \mathrm{id}_V$ und $f \circ g = \mathrm{id}_W$.

$\mathrm{id}_V\colon V \to V$ ist die identische Abbildung d.h. $\mathrm{id}_V(v) := v$ für alle $v \in V$.

Bemerkung. Existiert für eine lineare Abbildung f eine inverse Abbildung g, so ist g automatisch ebenfalls linear. Darum hatten wir diese Bedingung nicht in die Definition aufgenommen. Wir werden das noch beweisen, das Argument ist aber ziemlich formal.

JEDE LINEARE ABBILDUNG $F\colon V \to W$ können wir nach Wahl von Basen $\underline{\mathbf{v}} := \mathbf{v}_1, \ldots, \mathbf{v}_n$ von V und $\underline{\mathbf{w}} := \mathbf{w}_1, \ldots, \mathbf{w}_m$ von W also konkret in Koordinaten durch eine $m \times n$-Matrix ${}_{\underline{\mathbf{w}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F)$ beschreiben, deren Spalten die Koordinaten der Bilder der Basisvektoren $F(\mathbf{v}_1), \ldots, F(\mathbf{v}_n)$ sind, d.h.

$$ {}_{\underline{\mathbf{w}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F) = (\mathrm{Koord}_{\underline{\mathbf{w}}}(F(\mathbf{v}_1)) \ldots \mathrm{Koord}_{\underline{\mathbf{w}}}(F(\mathbf{v}_n)))\,. $$

MERKE: In den Spalten der Matrix stehen die Koordinaten der Bilder der Basisvektoren!

Wenn wir dann das Bild $F(\mathbf{v}) = F(a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n)$ berechnen wollen, so können wir einfach

$$ {}_{\underline{\mathbf{w}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F) \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} $$

berechnen und wissen dann, dass

$$ F(\mathbf{v}) = b_1\mathbf{w}_1 + \ldots b_m\mathbf{w}_m, $$

denn:

$$
\begin{aligned}
F(\mathbf{v}) &= F(a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n) && \mathbf{v} = a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n \\
&= a_1F(\mathbf{v}_1) + \cdots + a_nF(\mathbf{v}_n) && F \text{ linear} \\
&= a_1\,\mathrm{Komb}_{\underline{\mathbf{w}}}(\mathrm{Koord}_{\underline{\mathbf{w}}}(F(\mathbf{v}_1))) + \cdots + a_n\,\mathrm{Komb}_{\underline{\mathbf{w}}}(\mathrm{Koord}_{\underline{\mathbf{w}}}(F(\mathbf{v}_n))) && \mathrm{Komb}_{\underline{\mathbf{w}}}\,\mathrm{Koord}_{\underline{\mathbf{w}}} = \mathrm{id}_W \\
&= \mathrm{Komb}_{\underline{\mathbf{w}}}(a_1\,\mathrm{Koord}_{\underline{\mathbf{w}}}(F(\mathbf{v}_1)) + \cdots + a_n\,\mathrm{Koord}_{\underline{\mathbf{w}}}(F(\mathbf{v}_n))) && \mathrm{Komb}_{\underline{\mathbf{w}}} \text{ linear} \\
&= \mathrm{Komb}_{\underline{\mathbf{w}}}({}_{\underline{\mathbf{w}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F) \cdot \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}) && \text{Def. } {}_{\underline{\mathbf{w}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F) \\
&= \mathrm{Komb}_{\underline{\mathbf{w}}}(\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}) && \mathbf{b} = {}_{\underline{\mathbf{w}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F)\mathbf{a} \\
&= b_1\mathbf{w}_1 + \ldots b_m\mathbf{w}_m. && \text{Def. } \mathrm{Komb}_{\underline{\mathbf{w}}}
\end{aligned}
$$

Insgesamt können wir $F(\mathbf{v})$ also bestimmen, indem wir den Koordinatenvektor von $\mathbf{v}$ mit der Matrix ${}_{\underline{\mathbf{w}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F)$ multiplizieren und das Ergebnis als Koordinaten in W auffassen. Diesen Satz können wir in der Formel:

$$ F(\mathbf{v}) = \mathrm{Komb}_{\underline{\mathbf{w}}} \circ {}_{\underline{\mathbf{w}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F) \circ \mathrm{Koord}_{\underline{\mathbf{v}}}(\mathbf{v}), $$

HALT! Bitte schauen Sie die Formel einmal genau und in Ruhe so lange an, bis Ihnen klar ist, warum die Formel und der Satz davor den gleichen Sachverhalt beschreiben.

Sie sollten am Ende der Vorlesung in der Lage sein, die Aussage einer Formel wie dieser für sich selbst zu entschlüsseln. Sie sehen an diesem Beispiel, dass es dafür unentbehrlich ist, sich ganz genau klar zu machen, was die Symbole bedeuten! Sobald Sie das gemacht haben, verändert sich Ihr Verständnis der Formel.

zusammenfassen.

Beispiel 45. Sei $D = (\quad)' \colon K[x]_{\leq 3} \to K[x]_{\leq 2}$ wieder die Abbildung, die durch die Ableitung von Polynomen gegeben ist. Eine Basis vom Raum der Polynome vom Grad ≤ 3 ist $\mathbf{v}_1 = 1, \mathbf{v}_2 = x, \mathbf{v}_3 = x^2, \mathbf{v}_4 = x^3$, Basis vom Raum der Polynome vom Grad ≤ 2 ist $\mathbf{w}_1 = 1, \mathbf{w}_2 = x, \mathbf{w}_3 = x^2$. Es gilt

$$\begin{aligned} D(\mathbf{v}_1) = D(1) = \quad 0 \quad &= 0 \cdot \mathbf{w}_1 + 0 \cdot \mathbf{w}_2 + 0 \cdot \mathbf{w}_3 \\ D(\mathbf{v}_2) = D(x) = \quad 1 \quad &= 1 \cdot \mathbf{w}_1 + 0 \cdot \mathbf{w}_2 + 0 \cdot \mathbf{w}_3 \\ D(\mathbf{v}_3) = D(x^2) = 2 \cdot x \ &= 0 \cdot \mathbf{w}_1 + 2 \cdot \mathbf{w}_2 + 0 \cdot \mathbf{w}_3 \\ D(\mathbf{v}_4) = D(x^3) = 3 \cdot x^2 &= 0 \cdot \mathbf{w}_1 + 0 \cdot \mathbf{w}_2 + 3 \cdot \mathbf{w}_3. \end{aligned}$$

Wir verwenden den Merksatz: In den Spalten der Matrix stehen die Koordinaten der Bilder der Basisvektoren! Also berechnen wir $D(\mathbf{v}_1) = 0$ (das Bild des ersten Basisvektores), schreiben das Ergebnis als Linearkombination von $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ und schreiben die Koeffizienten in die 1. Spalte.

Also ist

$${}_{\underline{\mathbf{w}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(D) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Und in der Tat können wir die Koeffizienten der Ableitung des Polynoms $p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$ auch bestimmen, indem wir

$${}_{\underline{\mathbf{w}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(D) \cdot \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ 2a_2 \\ 3a_3 \end{pmatrix}$$

ausrechnen. Die Einträge dieses Vektors sind in der Tat genau die Koeffizienten der Ableitung.

Bemerkung. Mir scheinen Formeln mit Kompositionen von Abbildungen, wie

$$F(\mathbf{v}) = \mathrm{Komb}_{\underline{\mathbf{w}}} \circ {}_{\underline{\mathbf{w}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F) \circ \mathrm{Koord}_{\underline{\mathbf{v}}}(\mathbf{v})$$

immer komplizierter als die Aussagen zu sein, die die Formeln ausdrücken. Für mich ist es einfacher, die Abbildungen dann in ein Diagramm zu schreiben, d.h. ich schreibe für jede Abbildung Definitions- und Wertebereiche auf und dekoriere den Pfeil dazwischen mit der Abbildung:

$$\begin{array}{ccc} V & \xrightarrow{F} & W \\ {\scriptstyle \mathrm{Koord}_{\underline{\mathbf{v}}}}\downarrow & & \uparrow{\scriptstyle \mathrm{Komb}_{\underline{\mathbf{v}}}} \\ K^n & \xrightarrow{{}_{\underline{\mathbf{w}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F)} & K^m \end{array}.$$

Die Aussage der Formel, ist dann, dass die beiden Wege entlang der Pfeile, die gleiche Abbildung definieren, wir sagen dazu „das Diagramm kommutiert“.

Im Beispiel ist die Verkettung der 3 Abbildungen dann sichtbar: Berechne für einen Vektor $\mathbf{v} \in F$ erst den Koordinatenvektor $\mathrm{Koord}_{\underline{\mathbf{v}}}(\mathbf{v})$, multipliziere diesen mit der Matrix und übersetze die Ergebniskoordinaten mittels $\mathrm{Komb}_{\underline{\mathbf{w}}}$ als Element von W.

Das kann ich mir ganz gut merken, die Formel dazu

$$F = \mathrm{Komb}_{\underline{\mathbf{w}}} \circ {}_{\underline{\mathbf{w}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F) \circ \mathrm{Koord}_{\underline{\mathbf{v}}},$$

sieht komplizierter aus, aber ich hoffe, dass wir uns auch daran gewöhnen können.

LASSEN SIE UNS die Beobachtung, dass wir lineare Abbildungen zwischen endlich-dimensionalen Vektorräumen nach Wahl von Basen als Matrizen auffassen können noch einmal formal zusammenfassen.

Satz 46. *1. Jede lineare Abbildung $L\colon K^n \to K^m$ ist durch die Multiplikation mit einer eindeutig bestimmten $m \times n$-Matrix gegeben.*

2. Ist $F\colon V \to W$ eine lineare Abbildung endlich dimensionaler K-Vektorräume und sind $\underline{\mathbf{v}} := \mathbf{v}_1, \dots, \mathbf{v}_n \in V$ und $\underline{\mathbf{w}} := \mathbf{w}_1, \dots, \mathbf{w}_m \in W$ Basen, so wird F in den Koordinaten bezüglich $\underline{\mathbf{v}}$ und $\underline{\mathbf{w}}$ durch die Matrix

Merke: In den Spalten der Matrix stehen die Koordinaten der Bilder der Basisvektoeren.

$$\underline{\mathbf{w}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F) = (\mathrm{Koord}_{\underline{\mathbf{w}}}(F(\mathbf{v}_1)) \dots \mathrm{Koord}_{\underline{\mathbf{w}}}(F(\mathbf{v}_n)))$$

beschrieben, d.h.

$$\mathrm{Koord}_{\underline{\mathbf{w}}} \circ F \circ \mathrm{Komb}_{\underline{\mathbf{v}}} = \underline{\mathbf{w}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F).$$

Beweis. 1. Die erste Aussage kennen wir schon, denn ist $\mathbf{e}_1, \dots, \mathbf{e}_n$ die Standardbasis von K^n so gilt für jede lineare Abbildung $L\colon K^n \to K^m$, dass

$$L\left(\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}\right) = a_1 L(\mathbf{e}_1) + \dots + a_n L(\mathbf{e}_n) = (L(\mathbf{e}_1) \dots L(\mathbf{e}_n)) \cdot \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix},$$

d.h. L ist durch die $n \times m$-Matrix gegeben, deren Einträge die Bilder der Standardbasisvektoren sind. Da umgekehrt für jede $n \times m$-Matrix A, der Vektor $A\mathbf{e}_i$ der i-te Spaltenvektor von A ist, ist A auch eindeutig bestimmt.

2. Die zweite Aussage hatten wir gerade bewiesen, denn wir hatten

$$F = \mathrm{Komb}_{\underline{\mathbf{w}}} \circ \underline{\mathbf{w}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F) \circ \mathrm{Koord}_{\underline{\mathbf{v}}}$$

nachgerechnet. Multiplizieren wir diese Gleichung von links mit $\mathrm{Koord}_{\underline{\mathbf{w}}}$ und von rechts mit $\mathrm{Komb}_{\underline{\mathbf{v}}}$, so erhalten wir die Formel

$$\mathrm{Koord}_{\underline{\mathbf{w}}} \circ F \circ \mathrm{Komb}_{\underline{\mathbf{v}}} = \underline{\mathbf{w}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F).$$

□

Bemerkung. Die Aussage „ Jede lineare Abbildung $L\colon K^n \to K^m$ ist durch die Multiplikation mit einer eindeutig bestimmten $m \times n$-Matrix gegeben.“ können wir noch anders lesen. Die Menge aller linearen Abbildungen zwischen zwei Vektorräumen wird mit

$$\mathrm{Hom}_K(V, W) := \{F\colon V \to W \mid F \text{ linear}\}$$

bezeichnet. Die Aussage ist also gleichbedeutend damit, dass die Abbildungen

$\underline{\mathbf{e}} = \mathbf{e}_1, \dots, \mathbf{e}_n$ ist wie immer die Standardbasis von K^n.

$$\begin{aligned} {}_{\underline{\mathbf{e}}}\mathrm{Mat}_{\underline{\mathbf{e}}}\colon \mathrm{Hom}_K(K^n, K^m) &\to \mathrm{Mat}_{m,n}(K) \\ F &\mapsto {}_{\underline{\mathbf{e}}}\mathrm{Mat}_{\underline{\mathbf{e}}}(F) \\ \mathrm{Hom}_K(K^n, K^m) &\leftarrow \mathrm{Mat}_{m,n}(K) : \mathrm{Mult} \\ A \cdot \_\_ &\leftarrow A \end{aligned}$$

zueinander inverse Abbildungen sind.

Der zweite Teil des Satzes ist entsprechend: Ist $\underline{\mathbf{v}}$ eine Basis von V und $\underline{\mathbf{w}}$ eine Basis von W, so sind die Abbildungen

$$\begin{aligned} {}_{\underline{\mathbf{w}}}\mathrm{Mat}_{\underline{\mathbf{v}}}\colon \mathrm{Hom}_K(V, W) &\to \mathrm{Mat}_{m,n}(K) \\ F &\mapsto {}_{\underline{\mathbf{w}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F) \\ \mathrm{Hom}_K(V, W) &\leftarrow \mathrm{Mat}_{m,n}(K) : \mathrm{Mult} \\ \mathrm{Komb}_{\underline{\mathbf{w}}} \circ A \circ \mathrm{Koord}_{\underline{\mathbf{v}}} &\leftarrow A \end{aligned}$$

zueinander invers.

DABEI LERNEN WIR gleich noch etwas hinzu:

1. Die Multiplikation von Matrizen entspricht der Komposition der zugehörigen linearen Abbildungen:

   $$A \cdot B = {}_{\underline{\mathbf{e}}}\mathrm{Mat}_{\underline{\mathbf{e}}}(A \circ B)$$

   denn die Spaltenvektoren von $A \cdot B$ sind nach Definition der Multiplikation von Matrizen genau die Vektoren $A \cdot$ Spalte von B, d.h. die i-te Spalte von $A \cdot B$ ist $A(B \cdot \mathbf{e}_i)$ und das ist nach Definition die i-te Spalte von ${}_{\underline{\mathbf{e}}}\mathrm{Mat}_{\underline{\mathbf{e}}}(A \circ B)$.

2. Die Menge $\mathrm{Mat}_{m,n}(K)$ der $m \times n$-Matrizen mit Koeffizienten in K bildet einen Vektorraum, denn wir können Matrizen addieren und mit Skalaren multiplizieren. Genauso ist für K-Vektorräume V, die Menge $\mathrm{Hom}_K(V, W)$ der linearen Abbildungen von V nach W ein Vektorraum. Die Operationen für Matrizen übersetzen sich einfach als:

   $$(F + G)(\mathbf{v}) := F(\mathbf{v}) + G(\mathbf{v}) \text{ und } (c \cdot F)(\mathbf{v}) := c \cdot F(\mathbf{v}).$$

   Die Abbildung ${}_{\underline{\mathbf{w}}}\mathrm{Mat}_{\underline{\mathbf{v}}}\colon \mathrm{Hom}_K(V, W) \to \mathrm{Mat}_{m,n}(K)$ ist ein Isomorphismus von Vektorräumen, denn die Abbildung ist per Konstruktion linear.

Gute Basen schlechte Basen: Basiswechsel

Ein Spezialfall der Beschreibung von linearen Abbildungen durch Matrizen ist natürlich, dass $V = K^n$ und $W = K^m$ unsere Standardbeispiele von Vektorräumen sind, wir aber statt der Standardbasis, andere Basen $\underline{\mathbf{v}} \in K^n$ und $\underline{\mathbf{w}} \in K^m$ gewählt haben.

In diesem Fall ist $\mathrm{Komb}_{\underline{\mathbf{v}}}\colon K^n \to K^n$ durch die $n \times n$-Matrix $\underline{\mathbf{v}} = (\mathbf{v}_1 \mathbf{v}_2 \dots \mathbf{v}_n)$ deren Spalten, die Basisvektoren $\mathbf{v}_1, \dots, v_n$ sind gegeben und entsprechend ist $\mathrm{Koord}_{\underline{\mathbf{w}}} = (\mathbf{w}_1 \dots \mathbf{w}_m)^{-1}$ die *inverse* der Matrix mit den Spaltenvektoren $\mathbf{w}_1, \dots, \mathbf{w}_m$, denn $\mathrm{Komb}_{\underline{\mathbf{w}}}$ ist die zu $\mathrm{Koord}_{\underline{\mathbf{w}}}$ inverse Abbildung.

In Formeln finden wir also die Formel für den *Basiswechsel* einer Matrix:

$$\underline{\mathbf{w}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A) = (\mathbf{w}_1 \dots \mathbf{w}_m)^{-1} \circ A \circ (\mathbf{v}_1 \dots \mathbf{v}_n)$$
$$(= \quad \mathrm{Koord}_{\underline{\mathbf{w}}} \circ \underline{\mathbf{e}}\mathrm{Mat}_{\underline{\mathbf{e}}}(A) \circ \mathrm{Komb}_{\underline{\mathbf{v}}})$$

Beispiel 47. Im Raum der Fibonacci-Folgen:

$$\mathrm{Fibo} = \{(a_n)_{n\in\mathbb{N}} \in \mathbb{R} \mid a_{n+2} = a_{n+1} + a_n \text{ für alle } n\}$$

war für uns die Basis

$$\mathbf{v}_1 = (1, \phi_+, \phi_+^2, \dots), \mathbf{v}_2 = (1, \phi_-, \phi_-^2, \phi_-^3, \dots) \in \mathrm{Fibo},$$

wobei $\phi_\pm = \frac{1\pm\sqrt{5}}{2}$ besonders nützlich.

Dieses Beispiel können wir noch auf eine andere Art anschauen. Die Formel $a_{n+2} = a_{n+1} + a_n$ können wir nämlich auch als lineare Vorschrift für die Berechnung $\begin{pmatrix} a_n \\ a_{n+1} \end{pmatrix} \mapsto \begin{pmatrix} a_{n+1} \\ a_{n+2} \end{pmatrix}$ auffassen:

$$A\colon \mathbb{R}^2 \to \mathbb{R}^2$$
$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \mapsto \begin{pmatrix} a_2 \\ a_1 + a_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}.$$

Damit ist also für eine Folge $(a_1, a_2, \dots) \in \mathrm{Fibo}$ für jedes $n \in \mathbb{N}$ der Vektor

$$\begin{pmatrix} a_{n+1} \\ a_{n+2} \end{pmatrix} = \underbrace{A(A(\cdots(A}_{n-\text{mal}} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix})\cdots))$$
$$= \underbrace{A \circ A \circ \cdots \circ A}_{n-\text{mal}}(\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}) = A^n \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}.$$

Eine Formel für die Matrix $A^n = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \cdot \dots \cdot \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$ anzugeben ist also genauso schwer, wie eine Formel für die Fibonacci-Zahlen anzugeben.

Wählen wir aber die Basis $\mathbf{v}_1 = \begin{pmatrix} 1 \\ \phi_+ \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ \phi_- \end{pmatrix} \in \mathbb{R}^2$, so ist

Erinnerung: $\phi_\pm^2 = \phi_\pm + 1$

$$A\mathbf{v}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \phi_+ \end{pmatrix} = \begin{pmatrix} \phi_+ \\ 1 + \phi_+ \end{pmatrix} = \begin{pmatrix} \phi_+ \\ \phi_+^2 \end{pmatrix}$$
$$= \phi_+ \cdot \begin{pmatrix} 1 \\ \phi_+ \end{pmatrix} = \phi_+ \cdot \mathbf{v}_1 = \phi_+ \cdot \mathbf{v}_1 + 0 \cdot \mathbf{v}_2$$
$$A\mathbf{v}_2 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ \phi_- \end{pmatrix} = \begin{pmatrix} \phi_- \\ 1 + \phi_- \end{pmatrix} = \begin{pmatrix} \phi_- \\ \phi_-^2 \end{pmatrix}$$
$$= \phi_- \cdot \begin{pmatrix} 1 \\ \phi_- \end{pmatrix} = \phi_- \cdot \mathbf{v}_2 = 0 \cdot \mathbf{v}_1 + \phi_- \cdot \mathbf{v}_2.$$

Damit ist also

$$ {}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A) = (\mathrm{Koord}_{\underline{\mathbf{v}}}(A\mathbf{v}_1)\mathrm{Koord}_{\underline{\mathbf{v}}}(A\mathbf{v}_2)) = \begin{pmatrix} \phi_+ & 0 \\ 0 & \phi_- \end{pmatrix} $$

und für diese Matrix lassen sich die Potenzen mühelos ausrechnen:

$$ \begin{pmatrix} \phi_+ & 0 \\ 0 & \phi_- \end{pmatrix}^2 = \begin{pmatrix} \phi_+ & 0 \\ 0 & \phi_- \end{pmatrix}\begin{pmatrix} \phi_+ & 0 \\ 0 & \phi_- \end{pmatrix} = \begin{pmatrix} \phi_+^2 & 0 \\ 0 & \phi_-^2 \end{pmatrix} $$
$$ \begin{pmatrix} \phi_+ & 0 \\ 0 & \phi_- \end{pmatrix}^3 = \begin{pmatrix} \phi_+^2 & 0 \\ 0 & \phi_-^2 \end{pmatrix}\begin{pmatrix} \phi_+ & 0 \\ 0 & \phi_- \end{pmatrix} = \begin{pmatrix} \phi_+^3 & 0 \\ 0 & \phi_-^3 \end{pmatrix}. $$

Damit können wir induktiv

$$ \begin{pmatrix} \phi_+ & 0 \\ 0 & \phi_- \end{pmatrix}^n = \begin{pmatrix} \phi_+^n & 0 \\ 0 & \phi_-^n \end{pmatrix} $$

beweisen, denn die Formel gilt für $n = 1, 2, 3$ und wenn die Formel für ein $k \in \mathbb{N}$ gilt, so auch für $k+1$ da

$$ \begin{aligned} \begin{pmatrix} \phi_+ & 0 \\ 0 & \phi_- \end{pmatrix}^{k+1} &= \begin{pmatrix} \phi_+ & 0 \\ 0 & \phi_- \end{pmatrix}^{k} \cdot \begin{pmatrix} \phi_+ & 0 \\ 0 & \phi_- \end{pmatrix} \\ &= \begin{pmatrix} \phi_+^k & 0 \\ 0 & \phi_-^k \end{pmatrix} \cdot \begin{pmatrix} \phi_+ & 0 \\ 0 & \phi_- \end{pmatrix} \quad \text{verwende Aussage für } k \\ &= \begin{pmatrix} \phi_+^{k+1} & 0 \\ 0 & \phi_-^{k+1} \end{pmatrix}. \end{aligned} $$

Damit ist also

$$ {}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A^n) = \begin{pmatrix} \phi_+^n & 0 \\ 0 & \phi_-^n \end{pmatrix}. $$

Mit der Basiswechselformel können wir daraus auch eine Formel für A^n ablesen:

$$ \begin{aligned} A^n = {}_{\underline{\mathbf{e}}}\mathrm{Mat}_{\underline{\mathbf{e}}}(A^n) &= \mathrm{Koord}_{\underline{\mathbf{e}}} \circ {}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A^n) \circ \mathrm{Komb}_{\underline{\mathbf{e}}} \\ &= (\mathbf{v}_1\mathbf{v}_2) \cdot {}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A^n) \cdot (\mathbf{v}_1\mathbf{v}_2)^{-1} \\ &= \begin{pmatrix} 1 & 1 \\ \phi_+ & \phi_- \end{pmatrix}\begin{pmatrix} \phi_+^n & 0 \\ 0 & \phi_-^n \end{pmatrix}\begin{pmatrix} 1 & 1 \\ \phi_+ & \phi_- \end{pmatrix}^{-1} \end{aligned} $$

Wir berechnen die inverse Matrix $\begin{pmatrix} 1 & 1 \\ \phi_+ & \phi_- \end{pmatrix}^{-1}$ mit dem Gauß-Verfahren:

$\phi_- - \phi_+ = \frac{1-\sqrt{5}}{2} - \frac{1+\sqrt{5}}{2} = -\sqrt{5}$

$$ \left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ \phi_+ & \phi_- & 0 & 1 \end{array}\right) \begin{array}{l} \mid \cdot -\phi_+ \\ \longleftarrow + \end{array} $$
$$ \left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & (\phi_- - \phi_+) & -\phi_+ & 1 \end{array}\right) = \left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & -\sqrt{5} & -\phi_+ & 1 \end{array}\right) \mid \tfrac{-1}{\sqrt{5}} $$
$$ \left(\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 1 & \frac{\phi_+}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \end{array}\right) \begin{array}{l} \longleftarrow + \\ \mid - \end{array} $$
$$ \left(\begin{array}{cc|cc} 1 & 0 & \frac{-\phi_-}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ 0 & 1 & \frac{\phi_+}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \end{array}\right) $$

Probe:

$$\begin{pmatrix} 1 & 1 \\ \phi_+ & \phi_- \end{pmatrix} \cdot \begin{pmatrix} \frac{-\phi_-}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{\phi_+}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \end{pmatrix} = \begin{pmatrix} \frac{-\phi_- + \phi_+}{\sqrt{5}} & \frac{1}{\sqrt{5}} - \frac{1}{\sqrt{5}} \\ \frac{-\phi_-\phi_+}{\sqrt{5}} + \frac{\phi_+\phi_-}{\sqrt{5}} & \frac{\phi_+ - \phi_-}{\sqrt{5}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \checkmark$$

Also gilt $\phi_+\phi_- = \frac{1+\sqrt{5}}{2}\frac{1-\sqrt{5}}{2} = -1$

$$\begin{aligned} A^n = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n &= \begin{pmatrix} 1 & 1 \\ \phi_+ & \phi_- \end{pmatrix} \begin{pmatrix} \phi_+^n & 0 \\ 0 & \phi_-^n \end{pmatrix} \begin{pmatrix} 1 & 1 \\ \phi_+ & \phi_- \end{pmatrix}^{-1} \\ &= \begin{pmatrix} \phi_+^n & \phi_-^n \\ \phi_+^{n+1} & \phi_-^{n+1} \end{pmatrix} \begin{pmatrix} \frac{-\phi_-}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{\phi_+}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{-\phi_+^n\phi_- + \phi_-^n\phi_+}{\sqrt{5}} & \frac{\phi_+^n - \phi_-^n}{\sqrt{5}} \\ \frac{-\phi_+^{n+1}\phi_- + \phi_-^{n+1}\phi_+}{\sqrt{5}} & \frac{\phi_+^{n+1} - \phi_-^{n+1}}{\sqrt{5}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{\phi_+^{n-1} - \phi_-^{n-1}}{\sqrt{5}} & \frac{\phi_+^n - \phi_-^n}{\sqrt{5}} \\ \frac{\phi_+^n - \phi_-^n}{\sqrt{5}} & \frac{\phi_+^{n+1} - \phi_-^{n+1}}{\sqrt{5}} \end{pmatrix}. \end{aligned}$$

Es lohnt sich also, geschickte Basen für lineare Abbildungen zu suchen, selbst wenn die Abbildung schon als Matrix gegeben ist.

Bei der Wahl zwischen

$${}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A^n) = \begin{pmatrix} \phi_+^n & 0 \\ 0 & \phi_-^n \end{pmatrix} \text{ und } {}_{\underline{\mathbf{e}}}\mathrm{Mat}_{\underline{\mathbf{e}}}(A^n) = \begin{pmatrix} \frac{\phi_+^{n-1} - \phi_-^{n-1}}{\sqrt{5}} & \frac{\phi_+^n - \phi_-^n}{\sqrt{5}} \\ \frac{\phi_+^n - \phi_-^n}{\sqrt{5}} & \frac{\phi_+^{n+1} - \phi_-^{n+1}}{\sqrt{5}} \end{pmatrix}$$

ist klar, welche Matrix die Abbildung übersichtlicher beschreibt.

Aufgabe 4. Bestimmen Sie die Matrix ${}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}$ der Abbildung $A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} : \mathbb{R}^2 \to \mathbb{R}^2$ bezüglich der Basis

$$\underline{\mathbf{v}} = \mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

SKIZZE: Wir wissen:

$$\begin{aligned} {}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A) &= (\mathbf{v}_1\mathbf{v}_2)^{-1} \cdot A \cdot (\mathbf{v}_1\mathbf{v}_2) \\ &= \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \end{aligned}$$

Die Inverse der ersten Matrix berechnen wir mit dem Gauß-Algorithmus und finden

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Damit berechnen wir:

$$\begin{aligned} {}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A) &= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}. \end{aligned}$$

Die Abbildung entspricht also der Streckung mit dem Faktor 2 entlang der Diagonalen $\mathbb{R}\mathbf{v}_1 = \mathbb{R}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ und der Streckung mit dem Faktor 4 entlang der Antidiagonalen $\mathbb{R}\mathbf{v}_2 = \mathbb{R}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Das erklärt, wieso die Abbildungsmatrix in den neuen Koordinaten viel einfacher aussieht.

SIE HATTEN IN DER VORLESUNG GEFRAGT wie wir zu einer gegebenen linearen Abbildung, eine geschickte Basen wie in den Beispielen finden. Das zu erklären, ist das Ziel der nächsten Wochen.

EIN AUSBLICK: Das Besondere an den Basen $\mathbf{v}_1, \mathbf{v}_2$ in den Beispielen war, dass jeweils $A\mathbf{v}_1 = c \cdot \mathbf{v}_1$ und $A\mathbf{v}_2 = d \cdot \mathbf{v}_2$ galt und darum

$$ {}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A) = \begin{pmatrix} c & 0 \\ 0 & d \end{pmatrix}. $$

Als ersten Schritt sollten wir uns also überlegen für welche $c \in K$ es Vektoren $\mathbf{v}$ gibt, so dass $A\mathbf{v} = c \cdot \mathbf{v}$. Das ist gleichbedeutend mit $A\mathbf{v} - c \cdot \mathbf{v} = 0$ also $\mathbf{v} \in \mathrm{Ker}(A - c \cdot \mathrm{id}_n)$.

Dazu wäre es gut, eine Formel zu haben, die entscheidet, ob $\mathrm{Ker}(A - c \cdot \mathrm{id}_n) \neq \{0\}$ gilt, d.h. eine Formel, die entscheidet für welche c die Matrix $A - c\,\mathrm{id}_n$ invertierbar ist und für welche nicht. Die Determinante wird das für uns leisten.

Wörterbuch: Lösbarkeit von Gleichungen & Eigenschaften von Abbildungen.

Wir hatten Fragen zu linearen Gleichungssystemen in die Sprache der Matrizen übersetzt und diese jetzt zu linearen Abbildungen verallgemeinert.

Die Fragen zur Lösbarkeit von Gleichungen können wir entsprechend in Termen von Abbildungen umformulieren:

Sei $F\colon V \to W$ linear.

1. Die Gleichung $F(\mathbf{v}) =$ hat genau dann für alle $\in W$ eine Lösung $\mathbf{v} \in V$ wenn $\mathrm{Bild}(F) = W$.

2. Die Gleichung $F(\mathbf{v}) =$ hat genau dann für alle $\in W$ höchstens eine Lösung, wenn $\mathrm{Ker}(F) = \{0\}$.

3. Die Gleichung $F(\mathbf{v}) =$ hat genau dann für alle $\in W$ eine eindeutige Lösung, wenn F ein Isomorphismus ist, d.h. $\mathrm{Ker}(F) = \{0\}$ und $\mathrm{Bild}(F) = W$.

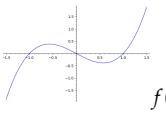
DIESE EIGENSCHAFTEN sind so nützlich, dass diese für allgemeine Abbildungen Namen bekommen haben.

Definition. Sei $f\colon X \to Y$ eine Abbildung von Mengen.

1. f heißt *surjektiv* genau dann, wenn die Gleichung $f(x) = b$ für alle $b \in Y$ wenigstens eine Lösung $x \in X$ besitzt.
2. f heißt *injektiv* genau dann, wenn es zu jedem $b \in Y$ höchstens eine Lösung der Gleichung $f(x) = b$ gibt, d.h. $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$.
3. $f\colon X \to Y$ heißt *bijektiv* genau dann, wenn die Gleichung $f(x) = b$ für alle $b \in Y$ eine eindeutige Lösung $x \in X$ besitzt.

Beispiel 48.

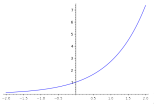
1. Die Abbildung $f\colon \mathbb{R} \to \mathbb{R}$, die durch $f(x) = (x-1)x(x+1) = x^3 - x$ gegeben ist, ist surjektiv, aber nicht injektiv.

   $f(x) = x^3 - x$

   Denn $x^3 - x = b$ ist für alle $b \in \mathbb{R}$ lösbar (surjektiv), aber $f(1) = 0 = f(0)$ (nicht injektiv).

2. Die Abbildung $f\colon \mathbb{R} \to \mathbb{R}$, die durch $f(x) = e^x$ gegeben ist, ist injektiv, aber nicht surjektiv.

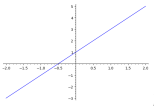
   

   $f(x) = e^x$

   $e^{x_1} = e^{x_2} \Leftrightarrow e^{x_1 - x_2} = 1 \Leftrightarrow x_1 = x_2$ (injektiv) in der Vorlesung hatten Sie vorgeschlagen stattdessen zu benutzen, dass $f(x) = ex$ streng monoton ist und darum aus $f(x_1) = f(x_2)$ folgt, dass $x_1 \not< x_2$ und $x_1 \not> x_2$ also $x_1 = x_2$.

   Da $e^x = -1$ in $\mathbb{R}$ keine Lösung besitzt ist f nicht surjektiv.

3. Die Abbildung $f\colon \mathbb{R} \to \mathbb{R}$, die durch $f(x) = 2x + 1$ gegeben ist, ist bijektiv.

   $f(x) = 2x + 1$

   Für alle $b \in \mathbb{R}$ hat die Gleichung $f(x) = 2x + 1 = b$ die eindeutige Lösung $x = \frac{b-1}{2}$. Die Abbildung $f\colon \mathbb{R} \to \mathbb{R}$ die durch $f(x) = x^3$ gegeben ist, ist auch bijektiv.

Aufgabe 5. Hotels schreiben im Gästebuch eine Abbildung

$$\text{Belegung}\colon \{\text{Hotelgäste}\} \to \{\text{Zimmer des Hotels}\}$$

auf.
Ein Paar kommt an die Rezeption und fragt nach einem Zimmer. Die Aushilfe an der Rezeption lernt gerade für die anstehende Mathematikprüfung und antwortet kryptisch mit einem der folgenden Sätze:

1. Es tut mir leid, aber die Belegung ist gerade surjektiv.
2. Es tut mir leid, aber die Belegung ist bei uns immer injektiv.
3. Es tut mir leid, aber die Belegung ist heute ausnahmsweise bijektiv.

FRAGE: Was will uns die Aushilfe damit sagen?

IHRE ANTWORTEN WAREN:

1. Alle Zimmer sind belegt.
2. Wir haben leider nur Einzelzimmer.
3. Alle Zimmer sind aktuell als Einzelzimmer belegt.

Die Determinante einer $(n \times n)$–Matrix

Wir hatten gelernt, wie wir von einer $(n \times n)$–Matrix $A\colon K^n \to K^n$ nachrechnen können, ob diese Matrix invertierbar ist. Wie am Ende des letzten Kapitels erwähnt, ist es nützlich, eine Formel in den Koeffizienten zu finden, die entscheidet, ob eine Matrix invertierbar ist oder nicht. Idealerweise sollte dies auch erlauben, eine explizite Formel für die inverse einer invertierbaren Matrix anzugeben.

Die Determinante ist eine Abbildung

$$\det\colon \mathrm{Mat}_{n,n}(K) \to K,$$

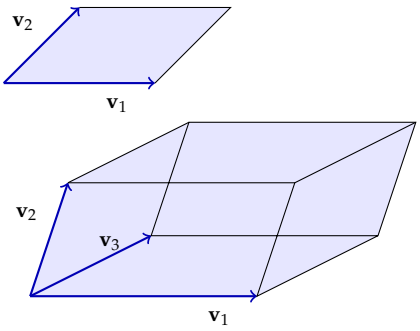
die das für uns leisten wird, d.h. $\det(A)$ wird genau dann $\neq 0$ sein, wenn A invertierbar ist.

Motivation: Die Suche nach einem Volumenbegriff

Da die Formel für die Determinante etwas sperrig ist, möchte ich Ihnen zunächst erklären, woher die Formel kommt. Für eine reelle $n \times n$-Matrix A wissen wir, dass die Matrix genau dann invertierbar ist, wenn die Spalten der Matrix eine Basis von $\mathbb{R}^n$ bilden.

Für $n = 2$ bedeutet das, dass die beiden Spalten nicht auf einer Geraden liegen. Das ist gleichbedeutend damit, dass das Parallelogramms mit den Seiten $\mathbf{v}_1, \mathbf{v}_2$ einen Flächeninhalt $\neq 0$ hat.

Für $n = 3$ bedeutet das, dass die 3 Spalten, nicht in einer Ebene liegen. Das bedeutet, das der von $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ definierte Spat, ein Volumen $\neq 0$ besitzt.



Die Frage, ob n Vektoren im $\mathbb{R}^n$ linear unabhängig sind, könnten wir also klären, wenn wir einen n-dimensionalen Volumenbegriff finden. Um dafür eine Formel zu suchen, könnten wir zunächst eine Wunschliste von Eigenschaften, die ein solches Volumen haben sollte, aufschreiben:

1. Die Formel sollte eine Abbildung definieren, die aus n Vektoren eine Zahl berechnet:

$$\begin{aligned}\det\colon \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n \text{ Faktoren}} &\to \mathbb{R}\\ (\mathbf{v}_1, \ldots, \mathbf{v}_n) &\mapsto \det(\mathbf{v}_1, \ldots, \mathbf{v}_n).\end{aligned}$$

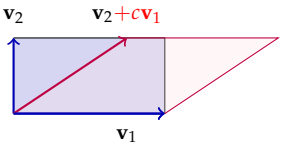
2. (Streckungen) Wenn wir einen der Vektoren mit einer Zahl $c \in \mathbb{R}$ multiplizieren, multipliziert sich das Volumen mit c, d.h. für alle $i \in \{1, \dots, n\}$ und $c \in \mathbb{R}$ gilt

$$\det(\mathbf{v}_1, \dots, c\mathbf{v}_i, \dots, \mathbf{v}_n) = c \det(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_n).$$

Eigentlich würden wir vielleicht mit $|c|$ multiplizieren, aber der Betrag ist nicht so schön, darum suchen wir lieber ein Volumen bis auf ein Vorzeichen.

3. (Scherungen) Wenn wir zu einem Vektor ein Vielfaches eines anderen addieren, sollte das Volumen unverändert bleiben:

$$\det(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_n) = \det(\mathbf{v}_1, \dots, \mathbf{v}_i + c\mathbf{v}_j, \dots, \mathbf{v}_n)$$



4. (Spate verkleben addiert Volumen) Für alle i gilt:

$$\det(\mathbf{v}_1, \dots, \mathbf{v}_i + \mathbf{v}_i', \dots, \mathbf{v}_n) = \det(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_n) + \det(\mathbf{v}_1, \dots, \mathbf{v}_i', \dots, \mathbf{v}_n)$$

5. (Doppelter Vektor) Wenn ein Vektor doppelt vorkommt, sollte das Volumen 0 sein: Für alle $i < j$ ist

$$\det(\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_{j-1}, \mathbf{v}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_n) = 0.$$

Bemerkung. 1. Einige der Bedingungen ergeben sich aus den anderen, vielleicht überlegen Sie sich einmal, welche.

2. Die Eigenschaften 2.-4. bedeuten einfach, dass die Abbildung det in jedem Eintrag $\mathbf{v}_i$ eine lineare Abbildung ist, d.h. für jede fest Wahl von Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots \mathbf{v}_n$ ist die Abbildung

$$K^n \ni x \mapsto \det(\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, x, \mathbf{v}_{i+1}, \dots \mathbf{v}_n) \in K$$

eine lineare Abbildung von K^n nach K.

3. Eine interessante Folgerung aus der 5. Eigenschaft und der Linearität ist, dass sich das Vorzeichen ändern muss, wenn wir zwei Vektoren vertauschen: Für alle $i < j$ gilt

$$\det(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_j, \dots, \mathbf{v}_n) = -\det(\mathbf{v}_1, \dots, \mathbf{v}_j, \dots, \mathbf{v}_i, \dots, \mathbf{v}_n)$$

Das folgt aus

$$\begin{aligned} 0 &= \det(\mathbf{v}_1, \dots, \mathbf{v}_i + \mathbf{v}_j, \dots, \mathbf{v}_i + \mathbf{v}_j, \dots, \mathbf{v}_n) \\ &= \det(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_i + \mathbf{v}_j, \dots, \mathbf{v}_n) + \det(\mathbf{v}_1, \dots, \mathbf{v}_j, \dots, \mathbf{v}_i + \mathbf{v}_j, \dots, \mathbf{v}_n) \\ &= \det(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_j, \dots, \mathbf{v}_n) + \det(\mathbf{v}_1, \dots, \mathbf{v}_j, \dots, \mathbf{v}_i, \dots, \mathbf{v}_n). \end{aligned}$$

Abbildungen, die die obigen Eigenschaften – die wir wieder in der Sprache von Vektorräumen formuliert haben – heißen alternierende Multilinearformen, das Anhängsel „form“ weist bei linearen Abbildungen darauf hin, dass das Ergebnis eine Zahl ist.

Definition (Alternierende Multilinearformen)**.** Sei V ein K-Vektorraum und $m \in \mathbb{N}$ und

$$\begin{aligned} f\colon \underbrace{V \times \cdots \times V}_{m \text{ Faktoren}} &\to K \\ (\mathbf{v}_1, \dots, \mathbf{v}_m) &\mapsto f(\mathbf{v}_1, \dots, \mathbf{v}_m) \end{aligned}$$

Die Symbole bedeuten: In f können wir m Vektoren einsetzen und erhalten dann eine Zahl als Ergebnis.

eine Abbildung.

1. Die Abbildung f heißt *Multilinearform*, wenn für alle $i \in \{1, \dots, m\}$ und alle $\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_m \in V$ die Abbildung
$$\mathbf{x} \mapsto f(\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{x}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_m)$$
linear ist.

2. Die Abbildung f heißt *alternierend*, wenn für alle $i < j$
$$f(\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_{j-1}, \mathbf{v}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_m) = 0$$
für alle $\mathbf{v}_1, \dots, \mathbf{v}_{i-1}, \mathbf{v}, \mathbf{v}_{i+1}, \dots, \mathbf{v}_{j-1}, \mathbf{v}, \mathbf{v}_{j+1}, \dots, \mathbf{v}_m \in V$.

Beispiel 49. 1. Wir wissen schon, dass lineare Abbildungen $f\colon K^n \to K$ durch $1 \times n$-Matrizen beschrieben werden, d.h. f ist immer von der Form
$$f\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = a_1 x_1 + \cdots + a_n x_n$$
für geeignete $a_1, \dots, a_n \in K$.

2. Ich behaupte, dass eine multilineare Abbildung
$$f\colon K^n \times K^n \to K$$
$$\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right) \mapsto f\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right)$$
ganz ähnlich durch eine Formel der Form
$$f\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i y_j$$
für geeignete $a_{ij} \in K$ gegeben ist.
Das folgt, da für festes $\mathbf{y} \in K^n$ die Abbildung $\mathbf{x} \mapsto f(\mathbf{x}, \mathbf{y})$ linear ist, also die Form $f\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right) = a_1(y) x_1 + \cdots + a_n(y) x_n$ hat. Wählen wir für $i \leq n$ den Vektor $\mathbf{x} = \mathbf{e}_i$, also $x_i = 1$ und $x_j = 0$ für $j \neq i$, dann ist $\mathbf{y} \mapsto f(e_i, \mathbf{y}) = a_i(\mathbf{y})$ und diese Abbildung ist ebenfalls linear. Also ist $a_i\left(\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right) = a_{i,1} y_1 + \cdots + a_{i,n} y_n$ und das war zu zeigen.

3. Die Abbildung
$$f\colon K^n \times K^n \to K$$
$$\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right) \mapsto f\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i y_j$$

ist genau dann alternierend, wenn für alle i gilt, dass $a_{i,i} = 0$ und für alle i, j gilt, dass $a_{i,j} = -a_{j,i}$.

Um das einzusehen, prüfen wir zunächst, dass eine Abbildung dieser Form alternierend ist, denn dann heben sich im Ausdruck $f(\mathbf{x}, \mathbf{x}) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$ jeweils die Terme $a_{i,j} x_i x_j$ und $a_{j,i} x_j x_i$ weg. Umgekehrt muss für jede alternierende Abbildung $0 = f(\mathbf{e}_i, \mathbf{e}_i) = a_{i,i})$ gelten und außerdem ist für alle i, j

$$a_{i,j} = f(e_i, e_j) = -f(e_j, e_i) = -a_{j,i}.$$

4. Alle alternierenden Multilinearformen $f\colon K^2 \times K^2 \to K$ sind also Vielfache der Abbildung

$$d(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}) = (x_1 y_2 - x_2 y_1),$$

d.h. $f(\mathbf{x}, \mathbf{y}) = c \cdot d(\mathbf{x}, \mathbf{y})$ für ein $c \in K$.

Aufgabe 6. Überlegen Sie sich analog, dass alle alternierenden Multilinearformen

$$f\colon K^3 \times K^3 \times K^3 \to K$$

Vielfache der Abbildung

$$\det(\begin{pmatrix} x_1 \\ x_2 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}, \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}) = x_1 y_2 z_3 - x_1 y_3 z_2 - x_2 y_1 z_3 + x_2 y_3 z_1 + x_3 y_1 z_2 - x_3 y_2 z_1$$

sind.

Wie zuvor können nur Terme $x_i y_j z_k$ mit 3 unterschiedlichen Indices vorkommen. Für diese finden wir Vorzeichen, denn fangen wir mit einem Term an, z.B. $x_1 y_2 z_3$, so änder sich bei Vertauschen von zwei Indices das Vorzeichen, also haben zum Beispiel $x_1 y_3 z_2$, $x_2 y_1 z_3$ und $x_3 y_2 z_1$ Vorzeichen $-$, wenn wir dann nochmal tauschen bekommen wir $-- = +$.

Die Beobachtung, dass es in den Beispielen bis auf Vielfache nur eine alternierende Möglichkeit gibt, lässt sich einfach erklären:

Behauptung 50. *Ist V ein n-dimensionaler K-Vektorraum, so gibt es bis auf Vielfache höchstens eine alternierende Multilinearform*

$$d\colon \underbrace{V \times \cdots \times V}_{n\text{Faktoren}} \to K.$$

Genauer gilt: Ist $\mathbf{v}_1, \ldots, \mathbf{v}_n$ eine Basis von V und $f\colon \underbrace{V \times \cdots \times V}_{n\text{Faktoren}} \to K$ eine alternierende Multilinearform, so ist f bereits durch den Wert $f(\mathbf{v}_1, \ldots, \mathbf{v}_n) = c \in K$ bestimmt, insbesondere ist f genau dann die 0-Abbildung, wenn $f(\mathbf{v}_1, \ldots, \mathbf{v}_n) = 0$ gilt.

Beweis. Das Argument funktioniert genau wie das Argument, warum lineare Abbildungen durch die Werte auf einer Basis bestimmt sind.

Ist $\mathbf{v}_1, \ldots, \mathbf{v}_n$ eine Basis von V und $\mathbf{v} = \mathbf{w}_1 \in V$ ein Vektor, so können wir $\mathbf{v}$ eindeutig als Linearkombination der Basis schreiben

$$\mathbf{v} = a_1 \mathbf{v}_1 + \cdots + a_n \mathbf{v}_n.$$

Dann ist für jede multilineare Abbildung $f\colon V^n \to K$

$$\begin{aligned} f(\mathbf{v}, \mathbf{w}_2, \dots, \mathbf{w}_n) &= f(a_1\mathbf{v}_1 + \dots + a_n\mathbf{v}_n, \mathbf{w}_2, \dots, \mathbf{w}_n) \\ &= a_1 f(\mathbf{v}_1, \mathbf{w}_2, \dots, \mathbf{w}_n) + \dots + a_n f(\mathbf{v}_n, \mathbf{w}_2, \dots, \mathbf{w}_n). \end{aligned}$$

Das Argument können wir induktiv auf die Vektoren $\mathbf{w}_2, \dots, \mathbf{w}_n$ anwenden, denn jedes der $\mathbf{w}_i$ lässt sich eindeutig als Linearkombination

$$\mathbf{w}_i = a_{i,1}\mathbf{v}_1 + \dots + a_{i,n}\mathbf{v}_n$$

schreiben.

Damit ist die Abbildung f durch die Funktionswerte $f(\mathbf{v}_{i_1}, \dots, \mathbf{v}_{i_n})_{i_1,\dots,i_n \in \{1,\dots,n\}}$ bestimmt.

Ist f zudem alternierend, so ist $f(\mathbf{v}_{i_1}, \dots, \mathbf{v}_{i_n}) = 0$ wann immer zwei der Indices übereinstimmen und sind alle Indices verschieden so gilt $\{i_1, \dots, i_n\} = \{1, \dots, n\}$, aber dann ist wegen

$$f(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_j, \dots, \mathbf{v}_n) = -f(\mathbf{v}_1, \dots, \mathbf{v}_j, \dots, \mathbf{v}_i, \dots, \mathbf{v}_n)$$

induktiv der Wert $f(\mathbf{v}_{i_1}, \dots, \mathbf{v}_{i_n})$ durch $f(\mathbf{v}_1, \dots, \mathbf{v}_n)$ bestimmt und stimmt mit diesem bis auf ein Vorzeichen überein. □

Folgerung 51. *Ist V ein $n-$dimensionaler K-Vektorraum und*

$$d\colon \underbrace{V \times \dots \times V}_{n\text{Faktoren}} \to K$$

eine alternierende Multilinearform, die nicht die 0-Abbildung ist, so sind n Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_n$ genau dann linear unabhängig, wenn

$$d(\mathbf{v}_1, \dots, \mathbf{v}_n) \neq 0 \in K.$$

Bemerkung. Die Folgerung können wir insbesondere auf die Abbildungen $\det\colon K^2 \times K^2 \to K$ die durch

$$\det\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = (x_1y_2 - x_2y_1)$$

gegeben ist und $\det\colon K^3 \times K^3 \times K^3 \to K$ mit

$$\det\left(\begin{pmatrix} x_1 \\ x_2 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}, \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}\right) = x_1y_2z_3 - x_1y_3z_2 - x_2y_1z_3 + x_2y_3z_1 + x_3y_1z_2 - x_3y_2z_1$$

aus den Beispielen anwenden. In diesen Beispielen gilt jeweils für die Standardbasis $\det(\mathbf{e}_1, \mathbf{e}_2) = 1 = \det(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$.

Beweis der Folgerung. Nach der Proposition ist die Abbildung eindeutig durch den Wert auf einer Basis bestimmt ist. Also gilt für alle Basen $d(\mathbf{v}_1, \dots, \mathbf{v}_n) \neq 0$. Sind die Vektoren umgekehrt linear abbhängig, so lässt sich einer der Vektoren als Linearkombination der anderen schreiben, d.h. es gibt einen Index i so dass

$$\mathbf{v}_i = \sum_{\substack{j=1\dots n \\ j\neq i}} a_j\mathbf{v}_j$$

Für $n = 3$ ist das Argument, dass wir $\mathbf{v}_i = a\mathbf{v}_j + b\mathbf{v}_k$ schreiben können, wobei $\{i, j, k\} = \{1, 2, 3\}$. Dann gilt:

$$\begin{aligned} \det(\mathbf{v}_i, \mathbf{v}_j, \mathbf{v}_k) &= \det(a\mathbf{v}_j + b\mathbf{v}_k, \mathbf{v}_j, \mathbf{v}_k) \\ &= a\det(\mathbf{v}_j, \mathbf{v}_j, \mathbf{v}_k) + b\det(\mathbf{v}_k, \mathbf{v}_j, \mathbf{v}_k) \\ &= 0. \end{aligned}$$

für geeignete $a_j \in K$. Dann ist aber

$$\begin{aligned} d(\mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_k) &= d(\mathbf{v}_1, \dots, \sum_{\substack{j=1\dots n \\ j\neq i}} a_j \mathbf{v}_j, \dots, \mathbf{v}_k) \\ &= \sum_{\substack{j=1\dots n \\ j\neq i}} a_j d(\mathbf{v}_1, \dots, \mathbf{v}_j, \dots, \mathbf{v}_k) \\ &= 0 \end{aligned}$$

denn im Ausdruck $d(\mathbf{v}_1, \dots, \mathbf{v}_j, \dots, \mathbf{v}_k)$ kommt der Vektor $\mathbf{v}_j$ doppelt vor.

□

Permutationen und Vorzeichen

Um analog zu den Beispielen für K^2 und K^3 für allgemeine n eine Formel für eine alternierende Multilinearform

$$\begin{aligned} \det\colon \underbrace{K^n \times \cdots \times K^n}_{n \text{ Faktoren}} &\to K \\ \left(\begin{pmatrix} x_{1,1} \\ \vdots \\ x_{n,1} \end{pmatrix}, \dots, \begin{pmatrix} x_{n,1} \\ \vdots \\ x_{n,n} \end{pmatrix} \right) &\mapsto \sum_{i_1,\dots,i_n} a_{i_1,\dots,i_n} x_{i_1} \cdots x_{i_n} \end{aligned}$$

angeben zu können, müssen wir wie in den Beispielen für jede Reihenfolge $i_1, \dots, i_n$ der Indices $1, \dots, n$ ein Vorzeichen für den Term $x_{i_1,1} \cdot \dots \cdot x_{i_n,n}$ so festlegen, dass sich das Vorzeichen ändert, wann immer wir zwei Indices vertauschen.

Um das aufzuschreiben, ist es hilfreich der Menge aller möglichen Vertauschungen einen Namen zu geben und die Struktur dieser Menge genauer anzuschauen.

Definition (Die symmetrische Gruppe)**.** Für jede natürliche Zahl n bezeichnen wir mit

$$S_n := \left\{ \sigma\colon \{1, \dots, n\} \to \{1, \dots, n\} \mid \sigma \text{ bijektiv} \right\}$$

Erinnern Sie sich, was bijektiv bedeutet? Wenn nicht, so schauen Sie das nach!

die Menge aller Permutationen (=Vertauschungen) der Zahlen $1, \dots, n$. Diese Menge heißt *symmetrische Gruppe* auf n Elementen.

Eine Möglichkeit, Permutationen anzugeben ist, eine Tabelle der Funktionswerte aufzuschreiben:

$$\begin{pmatrix} 1 & 2 & \dots & n-1 & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n-1) & \sigma(n) \end{pmatrix}$$

Zum Beispiel ist

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

die Abbildung, die 1 und 2 vertauscht (also $1 \to 2 \to 1$ abbildet) und die Zahlen $3, 4, 5$ zyklisch vertauscht (also $3 \to 4 \to 5 \to 3$ abbildet).

DIE ZYKELSCHREIBWEISE ist eine kompaktere Schreibweise für Permutationen. Dabei werden wie im Beispiel die Zahlen $1, \dots, n$ so hintereinander gruppiert, dass hinter jeder Zahl, die Zahl steht auf die diese abgebildet wird und die Gruppen einzuklammern, bei denen die letzte Zahl, wieder auf die erste abgebildet wird, also

$$(1,2)(3,4,5) \text{ für } 1 \to 2 \to 1, 3 \to 4 \to 5 \to 3.$$

Die eingeklammerten Gruppen heißen die Zykel der Permutation, die Abbildung σ stelle ich mir dabei so vor, dass die Gruppen jeweils in einem Kreis aufgeschrieben werden und σ alle Kreise um einen Schritt dreht.

Die Konvention ist, dass wir Zykel der Länge 1, d.h. die Zahlen, die auf sich selbst abgebildet werden weglassen, d.h. wir schreiben für die Abbildung in S_5, die $1, 2$ vertauscht und alle anderen Elemente fest lässt

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

kürzer

$$(1,2) \text{ statt } (1,2)(3)(4)(5).$$

Das ist praktisch, weil wir damit Elemente in S_n angeben können, ohne … zu verwenden. Für große n gibt es sehr viele Permutationen.

Behauptung 52. *Die Menge S_n hat $n! = n \cdot (n-1) \cdots 2 \cdot 1$ Elemente.*

Bemerkung. Die Anzahl der Elemente von S_n wächst sehr schnell:

n	2	3	4	5	6	10	20
n!	2	6	24	120	720	3.628.800	2.432.902.008.176.640.000.

Der Eindruck, dass sich mit recht wenigen Dingen schnell viel Unordnung herstellen lässt, trügt also nicht.

Beweis. Das Argument ist einfach: Für die möglichen Werte von $\sigma(1)$ gibt es die n-Möglichekeiten $1, \dots, n$. Ist $\sigma(1)$ gewählt, so bleiben für $\sigma(2)$ alle Zahlen außer $\sigma(1)$, also die $n-1$ Möglichkeiten $\{1, \dots, n\} \setminus \{\sigma(1)\}$, danach für $\sigma(3)$ die $n-2$-Möglichkeiten $\{1, \dots, n\} \setminus \{\sigma(1), \sigma(2)\}$ und so fort, insgesamt gibt es also

$$n \cdot (n-1) \cdots 2 \cdot 1 = n!$$

verschiedene Elemente in S_n. □

Beweis. Wenn wir das Argument formal ohne „und so fort" aufschreiben müssen, können wir das mit Induktion tun:

Wir beweisen mit Induktion, dass es für k-elementige Mengen I, J genau $k!$ bijektive Abbildungen $f\colon I \to J$ gibt.

Induktions Anfang: Für $n = 1$ gibt es nur eine Abbildung $\sigma\colon I \to J$ wenn I, J jeweils nur ein Element enthalten.

Angenommen wir wissen für ein k, dass es $k!$ bijektive Abbildungen zwischen K-elementigen Mengen gibt, so können wir die Elemente von S_{k+1} zählen, indem wir die Abbildungen

$\sigma\colon \{1,\dots,k+1\} \to \{1,\dots,k+1\}$ nach dem Funktionswert von $\sigma(1) =: i$ sortieren:

$$S_{k+1} = \bigcup_{i=1}^{k+1} \left\{ \sigma\colon \{1,\dots,k+1\} \to \{1,\dots,k+1\} \mid \begin{matrix} \sigma \text{ bijektiv} \\ \sigma(1) = i \end{matrix} \right\}$$

Für jedes i ist eine bijektive Abbildung $\sigma\colon \{1,\dots,k+1\} \to \{1,\dots,k+1\}$ mit $\sigma(k1) = i$ aber eindeutig durch die bijektive Abbildung

$$\begin{aligned} \sigma_k\colon \{2,\dots,k+1\} &\to \{1,\dots,k+1\} \smallsetminus \{i\} \\ j &\mapsto \sigma(j) \end{aligned}$$

bestimmt. Für k-elementige Mengen gibt es aber nach Induktion genau $k!$ bijektive Abbildungen. Da es $k+1$ Möglichkeiten für $i = \sigma(k+1)$ gibt, hat S_{k+1} also genau $(k+1)\cdot k! = (k+1)!$ Elemente. Und das gilt dann genauso für beliebige $k+1$-elementige Mengen. □

VERTAUSCHUNGEN der Zahlen $1,\dots,n$ können wir hintereinander ausführen — bei Kartenspielen nennen Sie das Mischen. Mathematisch bedeutet das, dass wir mit bijektiven Abbildungen wieder rechnen können:

- Sind $\sigma, \tau \in S_n$ so ist auch die Verkettung $\sigma \circ \tau \in S_n$.
- Ist $\sigma\colon \{1,\dots n\} \to \{1,\dots n\}$ bijektiv, so existiert eine Umkehrabbildung $\sigma^{-1}\colon \{1,\dots n\} \to \{1,\dots n\}$ mit $\sigma \circ \sigma^{-1}(i) = i$ und $\sigma^{-1} \circ \sigma(i) = i$ für alle $i = 1,\dots,n$.

Die Eigenschaften der Verknüpfung, bekommt wie bei Körpern und Vektorräumen einen Namen.

Definition. Eine *Gruppe* ist eine Menge G zusammen mit einer Verknüpfung $\cdot\colon G \times G \to G$ für die die folgenden Eigenschaften gelten:

1. (neutrales Element) Es gibt ein Element $e \in G$ so dass für alle $g \in G$ gilt
$$g \cdot e = g = e \cdot g.$$
2. (inverse Elemente) Für alle $g \in G$ existiert ein $g^{-1} \in G$ mit
$$g \cdot g^{-1} = e = g^{-1} \cdot g.$$
3. (Assoziativgesetz) Für alle $g, g', g'' \in G$ gilt
$$(g \cdot g') \cdot g'' = g \cdot (g' \cdot g'').$$

Beispiel 53. Gruppen beschreiben zumeist Symmetrien von Objekten:

1. Die symmetrische Gruppe S_n ist mit der Verknüpfung $\circ$ und dem neutralen Element $e = \mathrm{id}$ das durch die identische Abbildung $\mathrm{id}(x) = x$ für alle $x \in \{1,\dots n\}$ gegeben ist, eine Gruppe.

Das sind die möglichen Vertauschungen von n nummerierten Objekten. Die Gruppe S_3 beschreibt die Operationen im Hütchenspiel. Dass das verwirrend ist, liegt zu einem Teil daran, dass die Gruppe nicht kommutativ, d.h. oftmals ist $\sigma \circ \tau \neq \tau \circ \sigma$, zum Beispiel gilt für $\sigma = (12) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \tau = (23) = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$:

$$\sigma \circ \tau = \sigma(\tau) = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = (123)$$

$$\tau \circ \sigma = \tau(\sigma) = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = (132).$$

2. Die ganzen Zahlen $\mathbb{Z}$ sind mit $+$ eine Gruppe. Da $+n$ auf der Zahlengeraden der Verschiebung um n entspricht können wir $\mathbb{Z}$ auch als Teilmenge der Symmetrien der Zahlengeraden auffassen.

3. Wenn Sie die Ecken eines regelmäßigen n-Ecks von $1, \dots, n$ durchnummerieren, können Sie jede Symmetrie des n-Ecks eindeutig durch die Permutation der Ecken beschreiben, aber nicht alle Permutationen der Ecken liefern eine Symmetrie des $n-$Ecks. Die Teilmenge der Permuationen, die das tun bilden eine *Untergruppe* von S_n, d.h. eine Teilmenge, die mit der gleichen Verknüpfung $\circ$ selbst wieder eine Gruppe ist.

ZURÜCK ZUM VORZEICHEN einer Vertauschung. Für die Determinante suchen wir eine Möglichkeit, jeder Permutation ein Vorzeichen ± 1 so zuzuordnen, dass sich das Vorzeichen ändert, wann immer wir zwei Werte vertauschen, also eine Abbildung

$$\operatorname{sign}\colon S_n \to \{\pm 1\}.$$

Die Eigenschaft, dass sich das Vorzeichen bei Vertauschungen ändert, können wir erreichen, indem wir die Anzahl der Fehlstände zählen

$$\begin{aligned} \ell\colon S_n &\to \mathbb{Z}_{\geq 0} \\ \sigma &\mapsto \ell(\sigma) := \#\{(i,j) \in \{1,\dots,n\}^2 \mid i < j \text{ aber } \sigma(i) > \sigma(j)\} \end{aligned}$$

und beobachten, dass eine Vertauschung die Anzahl Fehlstände um eine ungerade Zahl verändert. Es ist vielleicht hilfreich, wenn Sie das an ein paar Beispielen selbst einmal ausprobieren, bevor wir versuchen, das formal zu beweisen.

Beispiel 54. 1. Für $1 \leq i < n$ ist für die Vertauschung $\tau_{i,i+1} = (i, i+1)$ von $i, i+1$, d.h.

$$\tau_{i,i+1}(k) = \begin{cases} i+1 & \text{wenn } k = i \\ i & \text{wenn } k = i+1 \\ k & \text{wenn } k \notin \{i, i+1\} \end{cases}$$

die Anzahl der Fehlstände $\ell(\tau_{i,i+1}) = 1$.

2. Seien $1 \leq i < j \leq n$ dann ist für die Vertauschung $\tau_{i,j} = (i,j)$ von i,j, d.h.
$$\tau(k) = \begin{cases} j & \text{wenn } k = i \\ i & \text{wenn } k = j \\ k & \text{wenn } k \notin \{i,j\} \end{cases}$$
die Anzahl der Fehlstände
$$\ell(\tau_{i,j}) = \ell((i,j)) = 2 \cdot (j-i-1) + 1$$
denn die Fehlstände sind genau durch die $j-i-1$ Zahlen k mit $i < k < j$ für die $i < k$ aber $\tau(i) = j > k = \tau(k)$ und $k < j$ aber $\tau(k) = k > i = \tau(j)$ gilt und das Paar i,j selbst gegeben. Das ist immer eine ungerade Zahl.

Satz 55 (sign ist multiplikativ)**.** *Die Abbildung*
$$\begin{aligned} \operatorname{sign}\colon S_n &\to \{\pm 1\} \\ \sigma &\mapsto (-1)^{\ell(\sigma)} = (-1)^{\textit{Anzahl der Fehlstände}(\sigma)} \end{aligned}$$
erfüllt $\operatorname{sign}(\sigma \circ \tau) = \operatorname{sign}(\sigma) \cdot \operatorname{sign}(\tau)$.

Die Formel $(-1)^N$ ist nur dazu da, die Parität (gerade/ungerade) von N zu bestimmen:
$$(-1)^N = \begin{cases} 1 & \text{wenn } N \text{ gerade} \\ -1 & \text{wenn } N \text{ ungerade.} \end{cases}$$
Wann immer Sie einen Ausdruck der Form $(-1)^{\text{Komplizierter Term}}$ sehen, sollten Sie das im Hinterkopf haben.

Beispiel 56. Seien $1 \leq i < j \leq n$ dann ist für die Vertauschung $\tau_{i,j}$ von i,j, dann ist die Anzahl der Fehlstände
$$\ell(\tau_{i,j}) = \ell((i,j)) = 2 \cdot (j-i-1) + 1$$
immer ungerade, also ist
$$\operatorname{sign}(\tau_{i,j}) = \operatorname{sign}((i,j)) = -1.$$

Es gibt recht viele verschiedene Methoden, die Multiplikativität von sign zu beweisen. Eine Methode ist, die Aussage in eine Formel zu übersetzen und dann zu üben, nicht vor komplizierten Formeln zu erschrecken.

Beweis des Satzes. Das Vorzeichen einer ganzen Zahl $a \in \mathbb{Z} \smallsetminus \{0\}$ können wir durch die Formel $Vz(a) = \frac{a}{|a|}$ angeben.

Um für eine Permutation σ die Zahl $(-1)^{\text{Anzahl der Fehlstände}(\sigma)}$ zu berechnen, können wir für alle $i < j$ jeweils entweder -1 wenn $\sigma(i) > \sigma(j)$, oder $+1$ wenn $\sigma(i) < \sigma(j)$ gilt aufmultiplizieren, also
$$\operatorname{sign}(\sigma) = \prod_{1 \leq i < j \leq n} Vz(\sigma(j) - \sigma(i)).$$

$$\begin{aligned} \operatorname{sign}(\sigma \circ \tau) &= \prod_{1 \leq i < j \leq n} Vz(\sigma(\tau(j)) - \sigma(\tau(i))) && \text{Formel ausgeschrieben} \\ &= \prod_{1 \leq i < j \leq n} \frac{Vz(\sigma(\tau(j)) - \sigma(\tau(i)))}{Vz(\tau(j) - \tau(i))} Vz(\tau(j) - \tau(i)) && \text{Füge } \frac{\operatorname{sign}\tau}{\operatorname{sign}\tau} \text{ hinzu.} \\ &= \left(\prod_{1 \leq i < j \leq n} \frac{Vz(\sigma(\tau(j)) - \sigma(\tau(i)))}{Vz(\tau(j) - \tau(i))} \right) \operatorname{sign}(\tau). \end{aligned}$$

Da τ eine bijektive Abbildung ist, kommen in der Liste $(\tau(i), \tau(j))_{i<j}$ alle Zahlenpaare $\{k,\ell\}_{k \neq \ell}$ genau einmal vor, nur eventuell in der falschen Reihenfolge.

Das erste Produkt in der obigen Formel, ist also wieder ein Produkt über alle Paare k, ℓ. Wenn $k = \tau(i) < \tau(j) = \ell$ gilt, dann ist

$$\frac{Vz(\sigma(\tau(j)) - \sigma(\tau(i)))}{Vz((\tau(j)) - (\tau(i)))} = \frac{Vz(\sigma(\ell) - \sigma(k))}{Vz(\ell - k)} = Vz(\sigma(\ell) - \sigma(k))$$

das Vorzeichen, das auch in der Formel für $\operatorname{sign}(\sigma)$ vorkommt und das gleiche gilt falls $k = \tau(i) > \tau(j) = \ell$, denn dann gilt

$$\frac{Vz(\sigma(\tau(j)) - \sigma(\tau(i)))}{Vz((\tau(j)) - (\tau(i)))}) = \frac{Vz(\sigma(\ell) - \sigma(k))}{-1} = Vz(\sigma(k) - \sigma(\ell)).$$

Damit ist also

$$\prod_{1 \le i < j \le n} \frac{Vz(\sigma(\tau(j)) - \sigma(\tau(i)))}{Vz((\tau(j)) - (\tau(i)))} = \prod_{1 \le k < \ell \le n} Vz(\sigma(\ell) - \sigma(k)) = \operatorname{sign}(\sigma).$$

Damit haben wir gezeigt, dass

$$\begin{aligned} \operatorname{sign}(\sigma \circ \tau) &= \left(\prod_{1 \le i < j \le n} \frac{Vz(\sigma(\tau(j)) - \sigma(\tau(i)))}{Vz(\tau(j) - \tau(i))} \right) \operatorname{sign}(\tau) \\ &= \operatorname{sign}(\sigma) \operatorname{sign}(\tau). \end{aligned}$$

□

Eine Formel für die Determinante

Wir haben nun eine Abbildung

$$\operatorname{sign}\colon S_n \to \{\pm 1\}$$

gefunden, die die Eigenschaft hat, dass für jede Vertauschung $\tau_{i,j}$ gilt, dass $\operatorname{sign}(\sigma \circ \tau_{i,j}) = -\operatorname{sign}(\sigma)$. Das ist, was wir benötigt haben um die Vorzeichen im Ausdruck, den wir für die Determinante erwarten festzulegen.

Satz 57 (Leibniz-Formel für die Determinante)**.** *Für jedes $n \in \mathbb{N}$ ist die Abbildung*

$$\det\colon \underbrace{K^n \times \cdots \times K^n}_{n \text{ Faktoren}} \to K$$

$$\det\left(\begin{pmatrix} x_{1,1} \\ \vdots \\ x_{n,1} \end{pmatrix}, \ldots, \begin{pmatrix} x_{1,n} \\ \vdots \\ x_{n,n} \end{pmatrix} \right) := \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) x_{\sigma(1),1} \cdot x_{\sigma(2),2} \cdots x_{\sigma(n),n}$$

eine alterniedende Multilinearform, so dass für die Standardbasis

$$\det(\mathbf{e}_1, \ldots, \mathbf{e}_n) = 1$$

gilt.

Beweis. Wir haben uns schon überlegt, dass Abbildungen der Form

$$\sum_{i_1, \ldots, i_n = 1}^{n} a_{i_1, \ldots, i_n} x_{i_1,1} \cdots x_{i_n,n}$$

multilinear sind.

Hier ist also

$$a_{i_1,\dots,i_n} = \left\{ \begin{array}{ll} \operatorname{sign}\left(\begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}\right) & \text{wenn alle Indices verschieden} \\ 0 & \text{sonst} \end{array} \right\}$$

Dass die Abbildung det alternierend ist, haben wir damit gerade so eingerichtet, denn wenn wir zwei Indices (i,j) vertauschen ist

$$\begin{aligned} a_{\dots,i_j,\dots,i_i,\dots} &= \operatorname{sign}\left(\begin{pmatrix} \dots & j & \dots & i & \dots \\ \dots & i_j & \dots & i_i & \dots \end{pmatrix}\right) \\ &= \operatorname{sign}\left(\begin{pmatrix} \dots & i & \dots & j & \dots \\ \dots & i_i & \dots & i_j & \dots \end{pmatrix} \circ (i,j)\right) \\ &= \operatorname{sign}\left(\begin{pmatrix} \dots & i & \dots & j & \dots \\ \dots & i_i & \dots & i_j & \dots \end{pmatrix}\right) \cdot \operatorname{sign}\big((i,j)\big) \\ &= -a_{\dots,i_i,\dots,i_j,\dots}. \end{aligned}$$

Haben wir also $\mathbf{v} = \begin{pmatrix} x_{1,i} \\ \vdots \\ x_{n,i} \end{pmatrix} = \begin{pmatrix} x_{1,j} \\ \vdots \\ x_{n,j} \end{pmatrix}$ für Indices $i \neq j$, so heben sich für alle $\sigma = \begin{pmatrix} \dots & i & \dots & j & \dots \\ \dots & \sigma(i) & \dots & \sigma(j) & \dots \end{pmatrix}$ die Terme für σ und $\sigma \circ (i,j)$ auf und wir erhalten

$$\det(\dots, \mathbf{v}, \dots, \mathbf{v}, \dots) = 0.$$

□

Bemerkung. Das Argument des Beweises, lässt sich in Termen von Gruppen etwas besser formulieren. Die Teilmenge

$$A_n := \{\sigma \in S_n \mid \operatorname{sign}(\sigma) = 1\} \subseteq S_n$$

heißt alternierende Gruppe. Für jede Transposition (i,j) gilt, dass

$$A_n \circ (i,j) := \{\sigma' = \sigma \circ (i,j) \mid \sigma \in A_n\} \quad = \{\sigma' \in S_n \mid \operatorname{sign}(\sigma') = -1\}.$$

Das Argument „Je zwei Urbilder unterscheiden Sich um ein Element des Kerns“ kommt hier in einer Variante für die Abbildung sign: $S_n \to \{\pm 1\}$ und „Ker(sign)“$= A_n$, zurück.

Damit gilt

$$\begin{aligned} \det(\mathbf{v}_1, \dots, \mathbf{v}_n) &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) x_{\sigma(1),1} \cdot x_{\sigma(2),2} \cdots x_{\sigma(n),n} \\ &= \sum_{\sigma \in A_n} \operatorname{sign}(\sigma) x_{\sigma(1),1} \cdot x_{\sigma(2),2} \cdots x_{\sigma(n),n} + \sum_{\sigma' \in A_n \circ (i,j)} \operatorname{sign}(\sigma') x_{\sigma'(1),1} \cdot x_{\sigma'(2),2} \cdots x_{\sigma'(n),n} \\ &= \sum_{\sigma \in A_n} 1 \cdot x_{\sigma(1),1} \cdot x_{\sigma(2),2} \cdots x_{\sigma(n),n} + \sum_{\sigma' = \sigma \circ (i,j) \in A_n \circ (i,j)} (-1) x_{(\sigma \circ (i,j)(1)),1} \cdot x_{\sigma \circ (i,j)(2),2} \cdots x_{\sigma \circ (i,j)(n),n} \\ &= \sum_{\sigma \in A_n} (\cdots x_{\sigma(i),i} \cdots x_{\sigma(j),j} \cdots) - (\cdots x_{\sigma(j),i} \cdots x_{\sigma(i),j} \cdots). \end{aligned}$$

Falls $\mathbf{v}_i = \mathbf{v}_j$ ist, sind in dieser Summe alle Differenzen 0.

Wie berechnen wir Determinanten?

Die Leibniz-Formel ist ein schönes Beispiel dafür, dass eine allgemeine Formel gleichzeitig hübsch aussehen und für die Berechnung (fast) völlig ungeeignet sein kann. Hier ist das so, weil die Summe $n!$ viele Summanden hat und das für größere n auch für schnelle Computer viel zu viele sind.

Die Formel ist aber auch ein gutes Beispiel dafür, dass es sich trotzdem lohnt, eine explizite Lösung zu haben, da wir die Eigenschaften der Formel verwenden können, um die Berechnung auf Spezialfälle zurückzuführen, in denen die Formel ganz einfach wird.

Behauptung 58. *Für Determinante von $n \times n$ Matrizen $A = (a_{i,j})_{i,j=1,\dots n}$ gelten die folgenden Rechenregeln:*

1. *Es gilt*

$$\det\left(\begin{pmatrix} a_{11} & a_{1,2} & \cdots & a_{1,n} \\ 0 & a_{22} & \cdots & a_{2,n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & a_{n,n} \end{pmatrix}\right) = a_{11} \cdot a_{2,2} \cdots a_{n,n},$$

   d.h. für obere Dreiecksmatrizen ist die Determinante das Produkt der Diagonaleinträge.

2. *Die transponierten Matrix A^t hat die gleiche Determinante wie A:*

$$\det(A) = \det(A^t).$$

   Insbesondere gilt die Formel in 1. auch für untere Dreiecksmatrizen.

3. *Die Determinante lässt sich mit dem Gauß-Algorithmus berechnen, denn es gilt:*

   (a) Die Determinante von A ändert sich nicht, wenn wir ein Vielfaches einer Zeile zu einer anderen addieren.

   (b) Multiplizieren wir eine Zeile von A mit $c \in K$, so multipliziert sich die Determinante ebenfalls mit c.

   (c) Vertauschen wir zwei Zeilen, so ändert sich die Determinante um ein Vorzeichen.

Beweis. 1. Die Leibnizformel besagt:

$$\det(A) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{\sigma(1),1} \cdot a_{\sigma(2),2} \cdots a_{\sigma(n),n}.$$

Für obere Dreiecksmatrizen gilt aber $a_{\sigma(i),i} = 0$ falls $\sigma(i) > i$ ist, damit sind in dieser Summe nur die Terme ungleich 0, in denen $\sigma(1) \leq 1, \sigma(2) \leq 2, \dots, \sigma(n) \leq n$ gilt. Da σ bijektiv ist, bedeutet das aber dass $\sigma(1) = 1, \sigma(2) = 2, \dots, \sigma(n) = n$ gelten muss, also $\det(A) = a_{1,1} \cdot a_{2,2} \cdots a_{n,n}$.

2. Die transponierte Matrix A^t ist die Matrix die entsteht, indem wir in A die Zeilen und Spalten-Indices vertauschen, also ist

$$\det(A^t) = \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdots a_{n,\sigma(n)},$$

d.h. in der Formel steht jetzt σ im zweiten Index, statt im ersten.

In jedem Ausdruck $a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdots a_{n,\sigma(n)} = \prod_{i=1}^n a_{i,\sigma(i)}$ können wir das Produkt aber nach dem zweiten Index umsortieren:

$$\prod_{i=1}^n a_{i,\sigma(i)} = \prod_{j=1}^n a_{\sigma^{-1}(j),j},$$

denn $\sigma(i) = j \Leftrightarrow i = \sigma^{-1}(j)$. Damit ist

$$\begin{aligned}\det(A^t) &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdots a_{n,\sigma(n)} \\ &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{\sigma^{-1}(1),1} \cdot a_{\sigma^{-1}(2),2} \cdots a_{\sigma^{-1}(n),n}.\end{aligned}$$

Außerdem ist $1 = \operatorname{sign}(\sigma \circ \sigma^{-1}) = \operatorname{sign}(\sigma)\operatorname{sign}(\sigma^{-1})$ und darum $\operatorname{sign}(\sigma) = \operatorname{sign}(\sigma^{-1})$. Darum können wir die Gleichung weiter umformen:

$$\begin{aligned}\det(A^t) &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma) a_{\sigma^{-1}(1),1} \cdot a_{\sigma^{-1}(2),2} \cdots a_{\sigma^{-1}(n),n} \\ &= \sum_{\sigma \in S_n} \operatorname{sign}(\sigma^{-1}) a_{\sigma^{-1}(1),1} \cdot a_{\sigma^{-1}(2),2} \cdots a_{\sigma^{-1}(n),n} \\ &= \sum_{\sigma' \in S_n} \operatorname{sign}(\sigma') a_{\sigma'(1),1} \cdot a_{\sigma'(2),2} \cdots a_{\sigma'(n),n} && \text{denn Inverse in } S_n \text{ sind eindeutig} \\ &= \det(A).\end{aligned}$$

3. Die Eigenschaften hatten wir in der Definition von alternierenden Multilinearformen für die Spaltenvektoren gefordert. Spaltenoperationen entsprechen Zeilenoperationen der transponierten Matrix, also gelten die Eigenschaften wegen 2. genauso für Zeilenoperationen.

□

FAZIT: Die Determinante berechnet sich am schnellsten mit dem Gaußalgorithmus, wobei wir diesmal sogar aufhören können sobald wir die Matrix in eine obere Dreiecksmatrix umgeformt haben.

Beispiel 59. Die Determinante der Matrix $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix}$ können wir mit Zeilenumformungen wie folgt berechnen:

$$\begin{aligned}\det\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \end{pmatrix} &= \det\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 2 & 8 \end{pmatrix} && \text{Umformung: 2. und 3. Zeile - 1.} \\ &= \det\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{pmatrix} && \text{Umformung: 3. Zeile - 2x 2.} \\ &= 1 \cdot 1 \cdot 2 = 2. && \text{Formel für obere Dreiecksmatrizen}\end{aligned}$$

Die Spaltenvektoren sind also in $\mathbb{R}^3$ linear unabhängig.

In $\mathbb{Z}/2\mathbb{Z}$ wäre $[2] = [0]$, die Spalten also linear abhängig, das ist sichtbar, denn in $\mathbb{Z}/2\mathbb{Z}^3$ sind die letzten beiden Spalten der Matrix gleich.

Beispiel 60. Das vorige Beispiel gehört zu einer Familie von Matrizen, die Ihnen hin und wieder begegnen werden.

$$\begin{aligned}\det \overset{\overset{-x \quad +}{\ulcorner\!\!\searrow}}{\begin{pmatrix} 1 & x \\ 1 & y \end{pmatrix}} &= \det \begin{pmatrix} 1 & 0 \\ 1 & y-x \end{pmatrix} \\ &= (y-x)\end{aligned}$$

$$\begin{aligned}\det \overset{\overset{-x \quad +}{\ulcorner\!\!\searrow}}{\begin{pmatrix} 1 & x & x^2 \\ 1 & y & y^2 \\ 1 & z & z^2 \end{pmatrix}} &= \det \overset{\overset{-x \quad +}{\ulcorner\!\!\searrow}}{\begin{pmatrix} 1 & x & 0 \\ 1 & y & y^2-xy \\ 1 & z & z^2-zy \end{pmatrix}} \\ &= \det \begin{pmatrix} 1 & 0 & 0 \\ 1 & y-x & y^2-xy \\ 1 & z-x & z^2-x^2 \end{pmatrix} \begin{matrix} \\ \ulcorner\!\!\leftarrow -1 \\ \leftarrow + \end{matrix} \begin{matrix} \\ \urcorner -1 \\ \urcorner + \end{matrix} \\ &= \det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1\cdot(y-x) & y\cdot(y-x) \\ 0 & 1\cdot(z-x) & z\cdot(z-x) \end{pmatrix} \\ &= (y-x)(z-x)\det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & y \\ 0 & 1 & z \end{pmatrix} \\ &= (y-x)(z-x)(z-y)\end{aligned}$$

Sie fragen sich hier wahrscheinlich, warum ich Spalten und nicht Zeilenumformungen anwende.

Eine Antwort ist, dass ich bei Zeilenumformungen, für den zweiten Schritt in allen Einträgen Variablen und keine 1en mehr finde. Schlimmer noch, wenn ich Zeilenumformungen anwende, erhalte ich nach dem ersten Schritt eine Matrix, die ich nicht gut vereinfachen kann. Versuchen Sie das ruhig selbst einmal.

Darum versuche ich dann einmal mein Glück mit Spaltenumformungen.

Um das Beispiel mit vollständiger Induktion allgemein ausrechnen zu können, ist es hilfreich uns noch eine weitere Rechenregel für Determinanten zu überlegen.

Behauptung 61. *Ist $A \in \mathrm{Mat}_{n,n}$ eine $n \times n$-Matrix und $B \in \mathrm{Mat}_{m,m}$ eine $m \times m$-Matrix, so gilt für die Determinante der $(n+m) \times (n+m)$-Matrix* $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$

$$\det(\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}) = \det(A) \cdot \det(B).$$

Beweis. Einen Beweis dieser Behauptung möchte ich Ihnen etwas später als Übungsaufgabe geben, wenn wir einen Trick gelernt haben, wie wir derartige Aussagen elegant, ohne große Rechnung beweisen können.

Eine Möglichkeit, wie wir das jetzt schon einsehen könnten, wäre allerdings auf den Gauß-Algorithmus zu verweisen: Die Aussage stimmt sicher, wenn A und B beide obere Dreiecksmatrizen sind,

denn dann ist auch $\det(\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix})$ eine obere Dreiecksmatrix und die Determinante dann also das Produkt der Diagonaleinträge. Mit dem Gauß-Verfahren können wir A und B durch Zeilenumformungen in obere Dreiecksmatrizen umformen. Die gleichen Umformungen funktionieren dann auch für $\det(\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix})$ und das reduziert die Behauptung auf den Fall von oberen Dreiecksmatrizen. □

Beispiel 62 (Vandermonde-Determinante). [16] Sind $x_1, \dots, x_n \in K$, so gilt

$$\det \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & & & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{pmatrix} = \prod_{1 \le i < j \le n} (x_j - x_i)$$

[16] Dies ist ein Beispiel einer wohl unbegründeten Namensgebung, Vandermonde (1735-1796) war Chemiker, Mathematiker und Musiker die Formel für die Determinante findet sich in seinen gesammelten Werken nicht.

Für $n = 2$ und $n = 3$ haben wir die Aussage in den Beispielen bereits gezeigt. Die Rechnung für $n = 3$ können wir induktiv verwenden. Subtrahieren wir jeweils das x_1-fache der vorletzten Spalte zur letzten, dann das x_1-fach der $n-2$-ten Spalte von der $n-1$-ten uund

so fort, erhalten wir:

$$\det \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-2} & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-2} & x_2^{n-1} \\ \vdots & & & & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-2} & x_n^{n-1} \end{pmatrix}$$

$$= \det \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-2} & 0 \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-2} & x_2^{n-1} - x_1(x_2^{n-2}) \\ \vdots & & & & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-2} & x_n^{n-1} - x_1(x_n^{n-2}) \end{pmatrix}$$

$$= \det \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 1 & x_2 - x_1 & x_2^2 - x_1x_2 & \dots & x_2^{n-1} - x_1(x_2^{n-2}) \\ \vdots & & & & \vdots \\ 1 & x_n - x_1 & x_n^2 - x_1x_n & \dots & x_n^{n-1} - x_1(x_n^{n-2}) \end{pmatrix}$$

$$= \det \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & (x_2 - x_1) \cdot 1 & (x_2 - x_1)x_2 & \dots & (x_2 - x_1)x_2^{n-2} \\ \vdots & & & & \vdots \\ 0 & (x_n - x_1) \cdot 1 & (x_n - x_1)x_n & \dots & (x_n - x_1)x_n^{n-2} \end{pmatrix}$$

$$= (x_2 - x_1) \cdots (x_n - x_1) \cdot \det \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & x_2 & \dots & x_2^{n-2} \\ \vdots & & & & \vdots \\ 0 & 1 & x_n & \dots & x_n^{n-2} \end{pmatrix}$$

$$= \Big(\prod_{j=2}^{n}(x_j - x_1)\Big) \det \begin{pmatrix} 1 & x_2 & \dots & x_2^{n-2} \\ \vdots & & & \vdots \\ 1 & x_n & \dots & x_n^{n-2} \end{pmatrix}$$

$$= \Big(\prod_{j=2}^{n}(x_j - x_1)\Big) \prod_{2 \le i < j \le n} (x_j - x_i) = \prod_{1 \le i < j \le n} (x_j - x_i).$$

Das ist die behauptete Formel.

Insbesondere ist die Vandermonde-Matrix genau dann invertierbar, wenn die Zahlen $x_1, \dots, x_n$ paarweise verschieden sind.

Der Multiplikationssatz und die Determinante eines Endomorphismus

Wir hatten die Determinante durch die Suche nach einem Volumenbegriff gefunden, d.h. die Determinante einer reellen Matrix können

wir als Volumen des Spats (=mehrdimensionales Parallelogramm) auffassen, dass durch die Spaltenvektoren definiert wird. Fassen wir die Matrix als die lineare Abbildung auf, die die Standardbasis von $\mathbb{R}^n$ auf die Spaltenvektoren abbildet, so bildet diese den Einheitswürfel auf den Spat der Spaltenvektoren ab. Damit sollten wir die Determinante also genauso als Volumenänderung, die die lineare Abbildung verursacht verstehen können. Diese Veränderung sollte unter Komposition von Abbildungen multiplikativ sein. Lassen Sie uns diese Aussage genau formulieren und beweisen.

Satz 63 (Determinanten–Multiplikationssatz)**.** *Sind $A, B \in \mathrm{Mat}_{n,n}(K)$ zwei $n \times n$-Matrizen mit Koeffizienten in einem Körper K, so gilt:*

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

Beweis. Es gibt sehr viele verschiedene Wege, diese Aussage zu beweisen. Mein Favorit ist das folgende Argument, das ohne Rechnung auskommt. Die Abbildung

$$d_A \colon \underbrace{K^n \times \cdots \times K^n}_{n \text{ Faktoren}} \to K$$

$$d_A\left(\begin{pmatrix} x_{1,1} \\ \vdots \\ x_{n,1} \end{pmatrix}, \ldots, \begin{pmatrix} x_{1,n} \\ \vdots \\ x_{n,n} \end{pmatrix}\right) := \det\left(A \begin{pmatrix} x_{1,1} \\ \vdots \\ x_{n,1} \end{pmatrix}, \ldots, A \begin{pmatrix} x_{1,n} \\ \vdots \\ x_{n,n} \end{pmatrix}\right)$$

ist genau wie det eine multilineare, alternierende Abbildung. Eine solche multilineare alternierende Abbildung, ist aber eindeutig durch den Wert auf einer beliebigen Basis, zum Beispiel der Standardbasis $\mathbf{e}_1, \ldots, \mathbf{e}_n$ bestimmt (Behauptung 50). Nach Definition ist $d_A(\mathbf{e}_1, \ldots, \mathbf{e}_n) = \det(A)$ und darum stimmen die Abbildungen $d_A(\_\_)$ und $\det(A) \cdot \det(\_\_)$ überein:

$$d_A(\quad) = \det(A) \cdot \det(\quad).$$

Setzen wir auf beiden Seiten die Spaltenvektoren von B ein, so erhalten wir

$$\det(A \cdot B) = d_A(B) = \det(A) \cdot \det(B).$$

□

In der Vorlesung hatten Sie darum gebeten, das Argument aufzuschreiben: d_A ist multilinear, da

$$\begin{aligned}
&d_A(v_1, \ldots, c \cdot v_i, \ldots) && \text{Def } d_A \\
&= \det(Av_1, \ldots, A(cv_i), \ldots) && A \text{ linear} \\
&= \det(Av_1, \ldots, cAv_i, \ldots) && \det \text{multi lin.} \\
&= c \det(Av_1, \ldots, Av_i, \ldots) && \text{Def } d_A \\
&= c d_A(v_1, \ldots, v_i, \ldots)
\end{aligned}$$

und

$$\begin{aligned}
&d_A(v_1, \ldots, v_i + v_i', \ldots) \\
&= \det(Av_1, \ldots, A(v_i + v_i'), \ldots) \\
&= \det(Av_1, \ldots, Av_i + Av_i', \ldots) \\
&= \det(Av_1, \ldots, Av_i, \ldots) + \det(Av_1, \ldots, Av_i', \ldots) \\
&= d_A(v_1, \ldots, v_i, \ldots) + d_A(v_1, \ldots, v_i', \ldots).
\end{aligned}$$

Die Abbildug ist alternierend, da

$$\begin{aligned}
&d_A(\ldots, v, \ldots, v, \ldots) && \text{Def } d_A \\
&= \det(\ldots, Av, \ldots, Av, \ldots) && \det \text{ alternierend} \\
&= 0
\end{aligned}$$

Aufgabe. Verwenden Sie den gleichen Trick, um einen neuen Beweis der Aussage

$$\det\left(\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\right) = \det(A) \cdot \det(B)$$

aus Behauptung 61 zu geben.

Insbesondere kennen wir durch die Multiplikativität der Determinante, die Determinante der inversen Matrix

Folgerung 64. *Ist $A \in \mathrm{Mat}_{n,n}(K)$ invertierbar, so gilt*

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Beweis. Für die Einheitsmatrix $E_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$ gilt

$$1 = \det(E_n) = \det(A \cdot A^{-1}) = \det(A) \cdot \det(A^{-1}).$$

□

Beispiel 65. Lassen Sie uns eine Probe machen: Ist $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, so ist

$$\det(A) = \det(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = ad - bc.$$

Sie hatten in der Übung berechnet, dass falls $ad - bc \neq 0$ gilt, dass

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

Damit gilt in der Tat

$$\det(A^{-1}) = \det \begin{pmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix} = \frac{da - (-b)(-c)}{(ad-bc)^2} = \frac{1}{ad-bc}.$$

Folgerung 66. *Die Determinante einer Matrix bleibt bei Basiswechsel unverändert, d.h. ist $A \in \mathrm{Mat}_{n,n}(K)$ eine $n \times n$-Matrix, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in K^n$ eine Basis und $B := (\mathbf{v}_1 \ldots \mathbf{v}_n)$ die Matrix von $\mathrm{Komb}_{\underline{\mathbf{v}}} \colon K^n \to K^n$, so gilt*

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

und ${}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A) = B^{-1}AB$.

Es ist bemerkenswert, dass es überhaupt Abbildungen $\mathrm{Mat}_{n,n}(K) \to K$ gibt, die sich bei Basiswechsel nicht verändern.

Das erlaubt insbesondere, für endlich dimensionale K-Vektorräume V die Determinante einer linearen Abbildung $F \colon V \to V$ definieren. Ich möchte das zunächst konkret mit Hilfe der Wahl einer Basis erklären und danach zurückblicken, um eine alternative Definition, die ohne Wahlen auskommt zu geben.

Zunächst können wir für $F \colon V \to V$ eine Basis $\mathbf{v}_1 \ldots, \mathbf{v}_n$ von V wählen und

$$\det(F) := \det({}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F))$$

definieren. Das Resultat hängt nicht von der Wahl der Basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ ab, denn für eine andere Basis $\mathbf{w}_1, \ldots, \mathbf{w}_n$ können wir die Matrix $B = {}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{w}}}(\mathrm{id}_V)$ verwenden, um

$${}_{\underline{\mathbf{w}}}\mathrm{Mat}_{\underline{\mathbf{w}}}(F) = {}_{\underline{\mathbf{w}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(\mathrm{id}_V){}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F){}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{w}}}(\mathrm{id}_V) \quad = B^{-1}{}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F)B$$

zu berechnen. Damit gilt

$$\det({}_{\underline{\mathbf{w}}}\mathrm{Mat}_{\underline{\mathbf{w}}}(F)) = \det(B^{-1}{}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F)B) = \det({}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F)).$$

Es gilt damit wie zuvor, dass $\det(F) \neq 0$ genau dann wenn F invertierbar ist.

Etwas abstrakter könnten wir die Determinante eines Endomorphismus $F\colon V \to V$ auch definieren, ohne dazu eine Basis zu wählen. Dazu können wir uns den Beweis der Multiplikativität der Determinante noch einmal anschauen und alle auftretenden Objekte für F und V formulieren.

Zunächst haben wir dort den Vektorraum

$$D_V := \{d\colon \underbrace{V \times \cdots \times V}_{n \text{ Faktoren}} \to K \mid d \text{ multilinear und alternierend}\}$$

der multilinearen alternierenden Abbildungen betrachtet. Da solche Abbildungen eindeutig durch den Wert auf einer Basis bestimmt sind (Behauptung 50) ist dieser Vektorraum höchstens eindimensional und weil wir für K^n ein Element $\det \neq 0$ gefunden haben, ist der Vektorraum tatsächlich für alle endlichdimensionalen Vektorräume $1-$dimensional.

Genau wie A eine d_A definiert hat, definiert F eine Abbildung

$$\begin{aligned} F^*\colon D_V &\to D_V \\ d &\mapsto F^*(d)(v_1,\dots,v_n) := d(F(v_1),\dots,F(v_n)). \end{aligned}$$

Die Abbildung F^* ist aber eine lineare Abbildung eines 1-dimensionalen Vektorraums und darum durch die Multiplikation mit einer Zahl gegeben: $F^*(d) = c_F \cdot d$. Die Determinante von F ist dann $\det(F) := c_F$.

Das hilft zwar nicht, um die Determinante auszurechnen, aber so müssen wir nicht mehr nachrechnen, dass die Definition unabhängig von der Wahl einer Basis war und das Argument wird viel kürzer. Es ist gut, sich an die abstrakteren Argumente zu gewöhnen, weil kompaktere Argumente kompliziertere Resultate erst möglich machen. Vergessen Sie dabei aber nicht, dass die Arbeit eine Determinante zu konstruieren noch immer wesentlich in das Argument eingeht, um zu zeigen, dass D_V nicht der 0-Vektorraum ist.

Laplace-Entwicklung und die Cramersche Regel

Die Berechnungsformel für die Determinante ist zum Beweisen oftmals praktischer als das Gauß-Verfahren. Dafür ist es aber wichtig, die Formel übersichtlich zu machen, indem wir die Abzählung aller Permutationen verwenden, um eine induktive Berechnung der Determinante zu erhalten.

Für die Determinante einer $n \times n$-Matrix $A = (a_{i,j})_{i,j=1,\dots,n} \in \mathrm{Mat}_{n,n}(K)$ können wir das wie folgt verwenden: Da det in jeder

In der Vorlesung hatten wir das Verfahren zunächst für das Beispiel

$$\det \begin{pmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

durchgeführt. Diese Rechnung fehlt hier noch.

Spalte linear ist, gilt

$$\det(A) = \sum_{i=1}^{n} a_{i,1} \det\left(\begin{pmatrix} 0 & a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & & \vdots \\ 1 & a_{i,2} & \cdots & a_{i,n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} i\text{-te Zeile} \right)$$

$$= \sum_{i=1}^{n} a_{i,1} \det\left(\begin{pmatrix} 0 & a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & & \vdots \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} i\text{-te Zeile} \right)$$

In diesem Ergebnis können wir die i-te Zeile an die erste Zeile verschieben, um eine Blockmatrix zu erhalten. Für das Verschieben müssen wir nacheinander die i-te mit der (i-1)-sten Zeile, dann mit der (i-2)ten bis zur 1. Zeile vertauschen, also $i-1$ Vertauschungen vornehmen, also ändert sich das Vorzeichen dabei $i-1$ mal und wir erhalten

$$\det(A) = \sum_{i=1}^{n} (-1)^{i-1} a_{i,1} \det\left(\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{i-1,2} & \cdots & a_{i-1,n} \\ 0 & a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} \right)$$

$$= \sum_{i=1}^{n} (-1)^{i-1} a_{i,1} \det\left(\begin{pmatrix} a_{1,2} & \cdots & a_{1,n} \\ \vdots & \vdots & \vdots \\ a_{i-1,2} & \cdots & a_{i-1,n} \\ a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & \vdots & \vdots \\ a_{n,2} & \cdots & a_{n,n} \end{pmatrix} \right)$$

Wir können die Determinante der $n \times n$ Matrix also induktiv als alternierende Summe über Eintrag der Spalte mal die Determinanten der $n-1 \times -n-1$ Matrizen die entstehen wenn wir 1. Spalte und i-te Zeile streichen berechnen.

alternierende Summe=abwechselnde Vorzeichen

Diese Formel heißt darum auch die Entwicklung der Determinante nach der ersten Spalte. Natürlich können wir das Argument genauso für die $j-$te statt der ersten Spalte durchführen. Das können wir auf die erste Spalte zurückführen indem wir die $j-$te Spalte nach vorne verschieben, dazu benötigen wir $j-1$ Vertauschungen und erhalten damit die Leibnizformel für die Entwicklung der

Determinante nach der $j-$ten Spalte:

$$\det(A) = \sum_{i=1}^{n} (-1)^{i-1}(-1)^{j-1} a_{i,j} \det\begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}$$

$$= \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \det\begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}.$$

Bemerkung. 1. Das Vorzeichen $(-1)^{i+1}$ für den (i,j)ten Eintrag der Matrix lässt sich als Schachbrettmuster merken:

$$((-1)^{i+j})_{i,j} = \begin{pmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

2. Da die Determinante einer Matrix gleich der Determinante der transponierten Matrix ist ($\det(A) = \det(A^t)$) gilt die entsprechende Formel auch für die Entwicklung nach der i-ten Zeile.

Lassen Sie uns das Ergebnis noch formal als Satz festhalten.

Satz 67 (Entwicklungssatz für die Determinante)**.** *Sei A eine $n \times n$-Matrix mit Koeffizienten in einem Körper K. Für $1 \leq i, j \leq j$ bezeichne*

$$A_{i,j} := \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}$$

die $(n-1) \times (n-1)$ Matrix, die entsteht indem in A die $i-$te Zeile und $j-$te Spalte gestrichen wird. Dann gilt für alle $i_0, j_0 \in \{1, \ldots n\}$, dass

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j_0} \cdot a_{i,j_0} \cdot \det(A_{i,j_0}) \quad \textit{(Entwicklung nach der } j_0\textit{-ten Spalte)}$$

$$= \sum_{j=1}^{n} (-1)^{i_0+j} \cdot a_{i_0,j} \cdot \det(A_{i_0,j}) \quad \textit{(Entwicklung nach der } i_0\textit{-ten Zeile).}$$

Beweis. Für die Entwicklung nach der j_0-ten Spalte, hatten wir das Ergebnis gerade berechnet. Die Formel für die Entwicklung nach der i_0-ten Zeile folgt daraus, indem wir das Resultat auf die transponierte Matrix A^t anwenden und uns erinnern, dass $\det(A) = \det(A^t)$ gilt. □

Diese Formel ist für die Berechnung von Determinanten nur nützlich, wenn die Matrix sehr viele 0-en enthält. Allerdings ist die Formel zum Beweisen sehr häufig nützlich, zum Beispiel um für spezielle Typen von Matrizen mittels vollständiger Induktion eine allgemeine Formel herzuleiten. Ein weiteres Beispiel dafür, dass die Formel zum Beweisen nützlich ist, ist die Cramersche Regel, die eine explizite Formel für die Inverse einer Matrix angibt. Um diese Formel zu finden, ist es hilfreich, die Formel für die Multiplikation von Matrizen mit der Formel aus dem Entwicklungssatz zu vergleichen. Seien also $A = (a_{i,j}), B = (b_{i,j}) \in \mathrm{Mat}_{n,n}(K)$ zwei $n \times n$ Matrizen. Dann gilt für die Einträge $c_{i,j}$ der Produktmatrix $C := A \cdot B$

In der Vorlesung hatten wir einige Beispiele berechnet, die im Skript noch fehlen.

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}.$$

Setzen wir also $b_{i,j} := (-1)^{i+j} \det(A_{j,i})$, d.h. der (i,j)-te Eintrag ist bis auf ein Vorzeichen die Determinante der Untermatrix on A, die entsteht indem wir die i-te Spalte und die j-te Zeile streichen, so sind die Diagonaleinträge von $A \cdot B$ nach dem Entwicklungssatz alle gleich $\det(A)$.

Beachten Sie die Vertauschung von Zeile und Spalte! Das kennen Sie aus der Merkregel für die Inverse einer 2×2-Matrix.

Satz 68 (Cramersche Regel)**.** *Sei A eine $n \times n$-Matrix mit Koeffizienten in einem Körper K. Für $1 \leq i, j \leq j$ bezeichne wieder $A_{i,j}$ die $(n-1) \times (n-1)$ Matrix, die entsteht indem in A die $i-$te Zeile und $j-$te Spalte gestrichen wird. Dann gilt für die Komplementärmatrix $A^\# := (a^\#_{i,j})$ mit Einträgen*

$$a^\#_{i,j} := (-1)^{i+j} \det(A_{j,i})$$

dass

$$A \cdot A^\# = A^\# A = \det(A) \cdot E_n$$

insbesondere gilt falls $\det(A) \neq 0$*, dass* $A^{-1} = \frac{1}{\det(A)} A^\#$.

Beweis. Der (i,j)-te Eintrag der Produktmatrix $(c_{i,j})_{i,j=1,\dots n} := A^\# A$ ist

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

Für $i = j$ ist dieser Ausdruck für $c_{i,i}$ nach dem Entwicklungssatz genau $\det(A)$, für $i \neq j$ ist der Ausdruck nach dem Entwicklungssatz die Determinante der Matrix, die aus A entsteht, indem wir die i-te Spalte von A durch die $j-$te ersetzen. Die Determinante einer Matrix mit zwei gleichen Spalten ist aber $= 0$, also gilt $A^\# \cdot A = \det(A) E_n$. Das Argument für $A \cdot A^\#$ ist analog, indem wir Zeilen statt Spalten verwenden. □

In der Vorlesung hatten wir das Ergebnis für den Fall einer 2×2-Matrix noch mit der Formel, die wir schon kennen verglichen.

Remark 69. Dieses Ergebnis ist zum Beispiel praktisch, weil es zeigt, dass die Abbildung

$$\begin{aligned}(\quad)^{-1}\colon \mathrm{GL}_n(\mathbb{R}) = \{A \in \mathrm{Mat}_{n,n}(\mathbb{R}) | \det(A) \neq 0\} &\to \mathrm{GL}_n(\mathbb{R}) \\ A &\mapsto A^{-1}\end{aligned}$$

eine stetige Abbildung ist.

Die Abbildung ist sogar differenzierbar und die Cramersche Regel erlaubt es, die Ableitung zu berechnen, aber das ist ein Thema für die Analysis 2.

Mit dem Gauß-Verfahren wäre es nicht so leicht, das zu zeigen, weil wir dabei Fallunterscheidungen wie $a_{1,1} = 0 / a_{1,1} \neq 0$ verwenden.

Die Formel ist auch nützlich, wenn Sie wissen möchten, welche Nenner in der Inversen einer Matrix auftreten.

Diagonalisierbarkeit und Eigenvektoren

Mit der Determinante können wir nun das bisher Gelernte zusammenfügen, um eine erste Antwort auf die Frage, wie wir geschickte Basen für eine gegebene lineare Abbildung finden können zu erhalten. Für die Matrizen die, die wichtigen Knoten in einem Netzwerk bestimmen wird das eine unerwartete neue Methode liefern, für diese sehr großen Matrizen Lösungen zu finden.

Wir hatten in Beispielen schon gesehen, dass wir für manche $n \times n$ Matrizen A Basen $\mathbf{v}_1, \ldots, \mathbf{v}_n \in K^n$ finden können, so dass die Matrix bezüglich dieser Basis nur noch auf der Diagonalen von 0 verschiedene Einträge hat:

$$\underline{\mathbf{v}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A) = \begin{pmatrix} t_1 & 0 & \ldots & 0 \\ 0 & t_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & t_n \end{pmatrix}.$$

Zum Beispiel hatten wir für $A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$ und die Basis $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \in \mathbb{R}^2$ gesehen, dass

$$\begin{aligned} &\underline{\mathbf{v}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A) \\ &= \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}. \end{aligned}$$

Damit war es einerseits leicht, eine geometrische Vorstellung der Abbildung zu erhalten (die Koordinaten werden jeweils um den Faktor d_i gestreckt) und andererseits konnten wir so auch eine Formel für A^n bestimmen, da die Potenzen von Diagonalmatrizen einfach zu berechnen sind.

Lassen Sie uns den auftretenden Objekten Namen geben.

Definition (Diagonalisierbarkeit)**.** 1. Eine $n \times n$-Matrix $A \in \mathrm{Mat}_{n,n}(K)$ heißt *diagonalisierbar* wenn es eine Basis $\mathbf{v}_1, \ldots, \mathbf{v}_n \in K^n$ gibt, so dass $\underline{\mathbf{v}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A) = \begin{pmatrix} t_1 & 0 & \ldots & 0 \\ 0 & t_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & t_n \end{pmatrix}$ eine Diagonalmatrix ist.

2. Allgemeiner heißt eine lineare Abbildung $F\colon V \to V$ eines n-dimensionalen K-Vektorraums *diagonalisierbar* wenn es eine Basis $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ gibt, so dass $\underline{\mathbf{v}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F) = \begin{pmatrix} t_1 & 0 & \ldots & 0 \\ 0 & t_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & t_n \end{pmatrix}$ eine Diagonalmatrix ist.

Wenn es Ihnen leichter fällt, können Sie „lineare Abbildung F“ in diesem Kapitel zunächst auch immer als „$n \times n$-Matrix A“ lesen.

DIE BEDINGUNG, dass die i-te Spalte der Matrix ${}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F)$ der Spaltenvektor $\begin{pmatrix} 0 \\ \vdots \\ t_i \\ 0 \\ \vdots \end{pmatrix}$ ist, bedeutet nach Definition der Abbildungsmatrix, dass

$$F(\mathbf{v}_i) = t_i \cdot \mathbf{v}_i.$$

Definition (Eigenvektoren und Eigenwerte)**.** Ist $F\colon V \to V$ eine lineare Abbildung, so heißt ein Vektor $\mathbf{v} \in V \smallsetminus \{0\}$ *Eigenvektor von* F, falls $F(\mathbf{v}) = c \cdot \mathbf{v}$ für ein $c \in K$. Die Zahl c heißt dann *Eigenwert* des Eigenvektors $\mathbf{v}$.

Existiert zu $c \in K$ ein Eigenvektor von F, so heißt c *Eigenwert* der Abbildung F.

Beispiel 70. In unserem Beispiel $A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$ gilt

$$\begin{aligned} A\mathbf{v}_1 &= \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 \\ 2 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \end{aligned}$$

also ist $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ein Eigenvektor zum Eigenwert 2 und

$$\begin{aligned} A\mathbf{v}_1 &= \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ &= \begin{pmatrix} 4 \\ -4 \end{pmatrix} = 4 \cdot \begin{pmatrix} 1 \\ -1 \end{pmatrix} \end{aligned}$$

und damit ist ist $\mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ein Eigenvektor zum Eigenwert 4.

In der Vorlesung hatten Sie sich noch ein Beispiel für eine linearen Abbildung gewünscht: Für Abbildung

$$\begin{aligned} F\colon K[t]_{\leq 2} &\to K[t]_{\leq 2} \\ p(t) &\mapsto F(p(t)) := t \cdot p'(t) \end{aligned}$$

Gilt $F(t^2) = t \cdot (2t) = 2 \cdot t^2$, also ist t^2 ein Eigenvektor von F zum Eigenwert 2.
IN DER ANALYSIS ist die Ableitung $f \mapsto f'$ auch eine lineare Abbildung und die Exponantialfunktion e^x erfüllt $(e^x)' = e^x$, die Funktion e^x ist also ein Eigenvektor für die Ableitung mit dem Eigenwert 1.
Aus dieser Eigenschaft lassen sich im übrigen alle anderen Eigenschaften der Exponentialfunktion relativ leicht herleiten.

EINE WICHTIGE BEOBACHTUNG ist nun, dass wir mit Hilfe der Determinante berechnen können, welche Zahlen als Eigenwerte einer Matrix vorkommen. Dazu schreiben wir die Gleichung $A\mathbf{v} = c\dot{\mathbf{v}}$ um

$$A\mathbf{v} = c \cdot \mathbf{v} \Leftrightarrow A\mathbf{v} - c \cdot \mathbf{v} = 0.$$

Um nun Matrizen-Rechnung zu verwenden schreiben wir $\mathbf{v} = E_n\mathbf{v}$ wobei $E_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}$ wieder die Einheitsmatrix, also die Matrix der identischen Abbildung $\mathrm{id}\colon K^n \to K^n, \mathrm{id}(\mathbf{v}) = \mathbf{v}$ ist. Dann gilt

$$\begin{aligned} A\mathbf{v} - c \cdot \mathbf{v} &= A\mathbf{v} - c \cdot E_n\mathbf{v} \\ &= (A - c \cdot E_n)\mathbf{v}. \end{aligned}$$

Also ist $\mathbf{v}$ genau dann ein Eigenvektor zum Eigenwert c von A wenn $\mathbf{v}$ im Kern von $A - cE_n$ liegt. Aber für eine $n \times n$-Matrix B gilt

$$\begin{aligned}\operatorname{Ker}(B) \neq \{0\} &\Leftrightarrow \text{Spalten von } B \text{ linear abhängig} \\ &\Leftrightarrow \det(B) = 0.\end{aligned}$$

Wenn wir c jetzt als Variable auffassen und darum t nennen, so wird für $A = (a_{ij})_{i,j=1,\dots n}$ aus $\det((A - t \cdot E_n))$ ein Polynom in $K[t]$:

$$\begin{aligned}\det(A - t \cdot E_n) &= \det(\begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ a_{n,1} & \cdots & a_{n,n-1} & a_{n,n} \end{pmatrix} - \begin{pmatrix} t & 0 & \dots & 0 \\ 0 & t & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & t \end{pmatrix}) \\ &= \det(\begin{pmatrix} a_{1,1} - t & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} - t & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ a_{n,1} & \cdots & a_{n,n-1} & a_{n,n} - t \end{pmatrix}).\end{aligned}$$

In unserem Beispiel erhalten wir:

$$\begin{aligned}&\det(\begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix} - \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix}) \\ &= \det \begin{pmatrix} 3-t & -1 \\ -1 & 3-t \end{pmatrix} \\ &= (3-t)(3-t) - (-1)(-1) \\ &= t^2 - 6t + 8 = (t-2)(t-4)\end{aligned}$$

Das Polynom hat also die Nullstellen 2 und 4.

Die Gleichung $A\mathbf{v} = c\mathbf{v}$ hat also genau dann eine Lösung $\mathbf{v} \in V \setminus \{0\}$ wenn c eine Nullstelle dieses Polynoms ist.

Definition. Für eine $n \times n$-Matrix A heißt das Polynom

$$\det(A - tE_n) \in K[t]$$

charakteristisches Polynom von A.

Satz 71. *Die Eigenwerte einer linearen Abbildung $F\colon V \to V$ sind die Nullstellen des charakteristischen Polynoms* $\det(F - t\,\mathrm{id}_V)$ *von F.*

Beweis. Das haben wir gerade gesehen: Ist c eine Nullstelle des charakteristischen Polynoms $\det(A - tE_n)$, so ist $\det(A - cE_n) = 0$. Das bedeutet, dass $\operatorname{Ker}(A - cE_n) \neq \{0\}$ und jeder Vektor $\mathbf{v} \in \operatorname{Ker}(A - cE_n) \neq \{0\}$ erfüllt $A\mathbf{v} = c\mathbf{v}$, d.h. c ist dann Eigenwert von A.

Ist umgekehrt c ein Eigenwert von A, so existiert $\mathbf{v} \in V \setminus \{0\}$ mit $A\mathbf{v} = c\mathbf{v}$. Also ist dann $\mathbf{v} \in \ker(A - cE_N) \setminus \{0\}$. Das bedeutet aber, dass $\det(A - cE_n) = 0$, d.h. c ist dann Nullstelle des charakteristischen Polynoms. □

In unserem Beispiel $A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$ hatten wir $\det(A - tE_2) = (t-2)(t-4)$ berechnet. A hat also die Eigenwerte $c_1 = 2, c_2 = 4$.

Eigenvektoren können wir jetzt als Elemente der Kerne von $A - c_1E_2$ und $A - c_2E_2$ bestimmen:

$$\operatorname{Ker}(A - 2E_2) = \operatorname{Ker}(\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}) = \operatorname{Ker}(\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}) = \operatorname{Span}(\begin{pmatrix} 1 \\ 1 \end{pmatrix})$$

$$\operatorname{Ker}(A - 4E_2) = \operatorname{Ker}(\begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}) = \operatorname{Ker}(\begin{pmatrix} -1 & -1 \\ 0 & 0 \end{pmatrix}) = \operatorname{Span}(\begin{pmatrix} 1 \\ -1 \end{pmatrix}).$$

Die Vektoren $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ sind linear unabhängig und darum eine Basis von $\mathbb{R}^2$. Weil v_1 ein Eigenvektor zum Eigenwert 2 und v_2 ein Eigenvektor zum Eigenwert 4 ist, ergibt sich jetzt auch ohne Rechnung, dass

$$ {}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A) = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}. $$

Beispiel 72. In der Vorlesung haben Sie das nun selbst für die Matrix

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

berechnet:

$$ \begin{aligned} \det(A - tE_3) &= \det \begin{pmatrix} -t & 1 & 0 \\ 0 & -t & 1 \\ -2 & 1 & 2-t \end{pmatrix} \\ &= -t \cdot \det \begin{pmatrix} -t & 1 \\ 1 & 2-t \end{pmatrix} + (-2) \det \begin{pmatrix} 1 & 0 \\ -t & 1 \end{pmatrix} \\ &= -t(-t(2-t)-1) - 2 \\ &= -t(t^2 - 2t - 1) - 2 = -t^3 + 2t^2 + t - 2 \end{aligned} $$

Dieses Polynom hat die Nullstelle 1, die können wir ausklammern

$$ -t^3 + 2t^2 + t - 2 = (t-1)(-t^2 - t + 2) \quad = -(t-1)(t+1)(t-2). $$

Die Eigenwerte sind also $t_1 = 1, t_2 = 2, t_3 = -1$. Für die Eigenvektoren finden wir mit dem Gauß-Verfahren

$$ \begin{aligned} \ker(A - 1E_3) &= \ker(\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ -2 & 1 & 1 \end{pmatrix}) \\ &= \ker(\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix}) = \mathrm{Span}(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}) \\ \ker(A - 2E_3) &= \ker(\begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ -2 & 1 & 0 \end{pmatrix}) \\ &= \ker(\begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & 0 \end{pmatrix}) = \mathrm{Span}(\begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}) \\ \ker(A - (-1)E_3) &= \ker(\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ -2 & 1 & 3 \end{pmatrix}) \\ &= \ker(\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}) = \mathrm{Span}(\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}) \end{aligned} $$

In der Vorlesung haben wir die Polynomdivison hierfür ausgeführt:

$$ \begin{array}{rl} -t^3 + 2t^2 & +t-2 = (t-1)\,(-t^2+t+2). \\ \underline{t^3 - t^2} & \\ t^2 & +t \\ \underline{-t^2} & \underline{+t} \\ & 2t-2 \\ & \underline{-2t+2} \\ & 0 \end{array} $$

Das Prinzip ist wie bei der schriftlichen Division ganzer Zahlen das Polynom, durch das Sie teilen möchten (hier $(t-1)$) jeweils mit der Potenz von $ct^?$ zu multiplizieren, dass der führende Term mit dem des großen Polynoms übereinstimmt. Das Vielfache subtrahieren wir und fahren dann mit dem Rest fort.

Wenn hier ein konstanter Rest r übrig bliebe, so hätten wir das Polynom als $q(t) \cdot (t-1) + r$ geschrieben, aber dieses Polynom würde für $t = 1$ den Wert r annehmen, d.h. 1 wäre keine Nullstelle.

Und also für $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$

$$ {}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}. $$

Hier habe ich versäumt nachzurechnen, dass $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ linear unabhängig sind. Wir überlegen uns gleich, wieso das hier automatisch der Fall ist.

Einschub: Polynomdivision und Ausklammern von Nullstellen

In der Randbemerkung zur Polynomdivision habe ich recht knapp erklärt, wieso wir Nullstellen von Polynomen immer als Linearfaktoren ausklammern können. Da das eine sehr nützliche Bemerkung ist, habe ich das in der Vorlesung noch einmal etwas ausführlicher erklärt.

Das Grundprinzip der Polynomdivision ist das gleiche Prinzip wie bei der schriftliche Division ganzer Zahlen: Sind $a(x), b(x) \in K[x]$ Polynome, dann können wir versuchen $b(x)$ als Vielfaches von $a(x)$ zu schreiben. Ist $a(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ und $b(x) = b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0$ mit $a_m \neq 0$ und $b_m \neq 0$ und $n = \mathrm{Grad}(a(x)) \leq \mathrm{Grad}(b(x)) = m$, so fällt in der Differenz

$$ b(x) - \frac{b_m}{a_n} x^{m-n} \cdot a(x) =: \tilde{b}(x) $$

der führende Term weg und wir erhalten ein Polynom vom Grad $< m$, d.h. $b(x) = \frac{b_m}{a_n} x^{m-n} \cdot a(x) + \tilde{b}(x)$. Das gleiche können wir mit $\tilde{b}(x)$ so lange fortführen, bis der Grad des übriggebliebenen Polynoms kleiner als der Grad von $a(x)$ ist. Das führt zur folgende Aussage:

Satz 73 (Teilen mit Rest für Polynome)**.** *Sind $a(x), b(x) \in K[x]$ Polynome mit Koeffizienten in einem Körper K, so existieren eindeutige Polynome $q(x), r(x) \in K[x]$ mit $\mathrm{Grad}(r(x)) < \mathrm{Grad}(a(x))$ so dass*

$$ b(x) = q(x) \cdot a(x) + r(x). $$

Beweis. Die Existenz haben wir uns gerade überlegt, bitte versuchen Sie einmal selbst, das Argument formal als Induktionsbeweis aufzuschreiben. Die Eindeutigkeit können wir ähnlich wie bei den ganzen Zahlen zeigen: Seien $b(x) = q(x) \cdot a(x) + r(x)$ und $b(x) = \tilde{q}(x) \cdot a(x) + \tilde{r}(x)$ zwei Lösungen des Problems mit $\mathrm{Grad}(r(x)) < \mathrm{Grad}(a(x))$ und $\mathrm{Grad}(\tilde{r}(x)) < \mathrm{Grad}(a(x))$. Dann gilt

$$ (q(x) - \tilde{q}(x))a(x) = \underbrace{\tilde{r}(x) - r(x)}_{\mathrm{Grad} < \mathrm{Grad}(a)}. $$

Da auf der linken Seite ein Vielfaches des Polynoms $a(x)$ steht, kann der Grad von $(q(x) - \tilde{q}(x))a(x)$ nur dann $< \mathrm{Grad}(a(x))$ sein, wenn $(q(x) - \tilde{q}(x)) = 0$ ist, also $q(x) = \tilde{q}(x)$, aber dann muss auch die rechte Seite 0 sein $0 = \tilde{r}(x) - r(x)$. Das zeigt die Eindeutigkeit. □

Beispiel: Für $b(x) = 3x^4 + 2x + 1$ und $a(x) = x^2 - 2$, ist

$$ \begin{aligned} &(3x^4 + 2x + 1) - 3x^2 \cdot (x^2 - 2) \\ &= 3x^4 + 2x + 1 - (3x^4 - 6x^2) \\ &= 6x^2 + 2x + 1 \end{aligned} $$

Das Ergebnis hat Grad $= 2 \geq$ Grad$(a(x)$ also weiter:

$$ \begin{aligned} &(6x^2 + 2x + 1) - 6 \cdot (x^2 - 2) \\ &= 6x^2 + 2x + 1 - 6x^2 + 12 \\ &= 2x + 13. \end{aligned} $$

Also ist

$$ b(x) = (3x^2 + 6)a(x) + 2x + 13. $$

Das können wir als Polynomdivision auch so schreiben:

$$ \begin{array}{rl} 3x^4 & +2x \ +1 = (x^2 - 2)\,(3x^2 + 6) \ +2x + 13 \\ \underline{-3x^4 + 6x^2} & \\ 6x^2 + 2x & +1 \\ \underline{-6x^2} & \underline{+12} \\ 2x + 13 & \end{array} $$

Bemerkung. Mit diesem Resultat könnten wir jetzt genau wie für ganze Zahlen auch einen euklidischen Algorithmus für Polynome durchführen und damit den größten gemeinsamen Teiler der Polynome $a(x), b(x)$ bestimmen und diesen als Linearkombination von $a(x)$ und $b(x)$ schreiben.

Auch das Rechnen in $\mathbb{Z}/n\mathbb{Z}$ könnten Sie damit auf $K[x]/p(x)K[x]$, die Menge in der wir Polynome bis auf Vielfache eines Polynoms $p(x)$ betrachten übertragen.

Wenn Sie das zum Beispiel für $\mathbb{R}[x]/(x^2+1)\mathbb{R}[x]$ betrachten, erfüllt die Äquivalenzklasse $[x]$ von x dort die Gleichung $[x]^2 + [1] = [x^2+1] = [0]$ und also $[x]^2 = [-1]$. Das liefert eine neue Konstruktion der komplexen Zahlen in der wir das Assoziativgesetz ohne Rechnung aus der entsprechenden Rechenregel für Polynome erhalten.

Folgerung 74 (Nullstellen können wir ausklammern)**.** *Ist $b(x) \in K[x]$ ein Polynom und $c \in K$ eine Nullstelle von $b(x)$, so können wir $b(x)$ als*

$$b(x) = (x-c) \cdot q(x)$$

für ein $q(x) \in K[x]$ schreiben.

Beweis. Wenden wir Teilen mit Rest für $a(x) = (x-c)$ auf $b(x)$ an, so erhalten wir

$$b(x) = (x-c) \cdot q(x) + r(x)$$

mit $\mathrm{Grad}(r(x)) < 1$, d.h. $r(x) = r_0$ ist konstant. Wir setzen wir $x = c$ in diese Gleichung ein und finden:

$$0 = b(c) = (c-c) \cdot q(c) + r_0 = 0 + r_0 = r_0.$$

Also ist $r_0 = 0$ und also $b(x) = (x-c)a(x)$. □

Folgerung 75. *Ein nicht konstantes Polynom $b(x) \in K[x]$ vom Grad n hat in einem Körper K höchstens n Nullstellen.*

Beweis. Ist $c_1 \in K$ eine Nullstelle von $b(x)$, so können wir nach dem vorangehenden Resultat $b(x) = (x-c_1)q(x)$ schreiben, wobei $\mathrm{Grad}(q(x)) = n-1$ gilt. Ist $c_2 \neq c_1$ eine weitere Nullstelle, so gilt $0 = b(c_2) = \underbrace{(c_2 - c_1)}_{\neq 0} q(c_2)$ also ist c_2 dann eine Nullstelle von $q(x)$. Hat b also n verschiedene Nullstellen, so können wir $b(x)$ induktiv als

$$b(x) = (x-c_1) \cdot \cdots \cdot (x-c_n) \cdot q_0$$

schreiben.

Da in jedem Körper ein Produkt genau dann 0 ist, wenn einer der Faktoren 0 ist, hat das Polynom

$$(x-c_1) \cdot \cdots \cdot (x-c_n) \cdot q_0$$

genau die Nullstellen $c_1, \ldots, c_n$ und keine weitere. Also hat auch $b(x)$ keine weitere Nullstelle. □

In der Vorlesung hatten Sie gefragt, wann wir n Nullstellen finden. Das hängt vom Körper ab. In den rationalen Zahlen $\mathbb{Q}$ hatte $x^2 - 2$ keine Nullstelle, aber das Problem konnten wir in den reellen Zahlen $\mathbb{R}$ beheben. Dort hatte $x^2 + 1$ keine Nullstelle, weshalb wir die komplexen Zahlen $\mathbb{C}$ eingeführt hatten. Diese reichen dann tatsächlich aus:

Fakt (Fundamentalsatz der Algebra)**.** Jedes nicht konstante Polynom $p(x) \in \mathbb{C}[x]$ hat in $\mathbb{C}$ eine Nullstelle.

Für dieses Resultat werden Sie später im Studium hoffentlich verschiedene schöne Beweise kennenlernen, in der linearen Algebra ist das leider noch nicht möglich.

Bemerkung. Für Polynome vom Grad 2 haben Sie mir in der Globalübung erzählt, dass Sie nicht alle wissen, wie Sie Nullstellen sofort ablesen können. Das müssen wir dringend ändern: Für mich kommen die Lösungsformeln aus der binomischen Formel

$$(x+a)^2 = x^2 + 2ax + a^2,$$

für Ausdrücke dieser Form können wir also eine Wurzel hinschreiben. Wollen wir allgemein für Zahlen p, q die Gleichung

$$x^2 + px + q = 0$$

lösen, so können wir das schnell in die einfache Form bringen und dann lösen:

$$\begin{aligned} x^2 + px + q = 0 \qquad & \Leftrightarrow \\ x^2 + 2\left(\frac{p}{2}\right)x + \left(\frac{p}{2}\right)^2 = -q + \left(\frac{p}{2}\right)^2 \qquad & \Leftrightarrow \\ (x + \left(\frac{p}{2}\right))^2 = \left(\frac{p^2}{4} - q\right) \qquad & \text{also} \\ x = \pm\sqrt{\left(\frac{p^2}{4} - q\right)} - \frac{p}{2}. & \end{aligned}$$

Beispiel: Die Nullstellen von

$$x^2 - x - 1 = 0$$

sind

$$\begin{aligned} x_\pm &= \pm\sqrt{\frac{1}{4} + 1} + \frac{1}{2} \\ &= \frac{\pm\sqrt{5}+1}{2} \end{aligned}$$

Für Polynome vom Grad 3 und 4 lassen sich mit etwas mehr Mühe, aber einem ähnlichen Grundprinzip auch Lösungsformeln finden. Für Gleichungen höheren Grades werden wir in der Algebra sehen, dass es keine allgemeine Lösungsformel gibt, die mir Wurzelzeichen auskommt.

Ein erstes Hindernis hierfür können Sie an der obigen Formel schon erkennen: Wenn in einem Körper K, wie zum Beispiel $\mathbb{Z}/2\mathbb{Z}$ die merkwürdige Gleichung $1 + 1 = 0$ gilt, können wir nicht durch $2 = 0$ teilen und darum auch die obige Formel nicht verwenden.

Das passt dazu, dass $p(x) = x^2 + x + 1$ in $\mathbb{Z}/2\mathbb{Z}$ keine Nullstelle hat.

Zurück zu Eigenwerten

Es gibt ein nützliches Resultat, das uns oft erspart auszurechnen, dass Eigenvektoren eine Basis von K^n sind.

Satz 76. *Eigenvektoren zu verschiedenen Eigenwerten einer Matrix A sind linear unabhängig. Ist $A \in \mathrm{Mat}_{n,n}(K)$ und sind $\mathbf{v}_1, \dots, \mathbf{v}_r$ Eigenvektoren zu paarweise verschiedene Eigenwerten $c_1, \dots, c_r \in K$ von A, so sind $\mathbf{v}_1, \dots, \mathbf{v}_r$ linear unabhängig.*

Beweis. Diese Aussage können wir mit Induktion über die Anzahl r beweisen. Für $r = 1$ ist die Aussage richtig, denn ein Vektor $\mathbf{v}_1 \in V \smallsetminus \{0\}$ ist linear unabhängig.

Angenommen die Aussage stimmt für eine Zahl k von Eigenvektoren und seien $\mathbf{v}_1, \dots, \mathbf{v}_{k+1}$ Eigenvektoren zu paarweise verschiedenen Eigenwerten $c_1, \dots, c_{k+1}$. Dann ist zu zeigen, dass die Gleichung

$$a_1\mathbf{v}_1 + \dots a_{k+1}\mathbf{v}_{k+1} = 0 \tag{8}$$

nur die triviale Lösung $a_1 = \dots = a_{k+1} = 0$ hat.

Sei also $a_1, \dots, a_{k+1}$ eine Lösung. Der Trick ist nun A auf den Vektor $a_1\mathbf{v}_1 + \dots a_{k+1}\mathbf{v}_{k+1}$ anzuwenden:

$$\begin{aligned} A(a_1\mathbf{v}_1 + \dots a_{k+1}\mathbf{v}_{k+1}) &= a_1A\mathbf{v}_1 + \dots + a_{k+1}A\mathbf{v}_{k+1} \qquad A \text{ linear} \\ &= a_1c_1\mathbf{v}_1 + \dots + a_{k+1}c_{k+1}\mathbf{v}_{k+1} \end{aligned}$$

Da $A0 = 0$, gilt also auch

$$a_1c_1\mathbf{v}_1 + \dots + a_{k+1}c_{k+1}\mathbf{v}_{k+1} = A0 = 0 \tag{9}$$

In der Differenz von c_{k+1} mal der ersten Gleichung

$$c_{k+1}a_1\mathbf{v}_1 + \dots c_{k+1}a_{k+1}\mathbf{v}_{k+1} = 0$$

und

$$a_1c_1\mathbf{v}_1 + \dots + a_{k+1}c_{k+1}\mathbf{v}_{k+1} = 0$$

fällt der Term $a_{k+1}c_{k+1}\mathbf{v}_{k+1}$ weg und wir erhalten

$$(c_{k+1} - c_1)a_1\mathbf{v}_1 + \dots + (c_{k+1} - c_k)a_k\mathbf{v}_k = 0.$$

Nach Induktionsvoraussetzung sind dann für alle i die Koeffizienten $(c_{k+1} - c_i)a_i = 0$ und da $(c_{k+1} - c_i) \neq 0$ folgt daraus $a_1 = \dots = a_k = 0$. Aus der ersten Gleichung folgt dann $a_{k+1}v_{k+1} = 0$, also auch $a_{k+1} = 0$. □

Wie komme ich auf die Idee? Ich schreibe auf, was wir wissen: Wir haben die Gleichung

$$a_1\mathbf{v}_1 + \dots a_{k+1}\mathbf{v}_{k+1} = 0$$

und die Voraussetzung, dass $\mathbf{v}_i$ Eignvektoren sind, d.h.

$$A\mathbf{v}_i = c_i\mathbf{v}_i.$$

Das müssen wir beides verwenden. Mir fällt nicht viel anderes ein, als A auf die Gleichung $a_1\mathbf{v}_1 + \dots a_{k+1}\mathbf{v}_{k+1} = 0$ anzuwenden und zu schauen, was herauskommt.

Folgerung 77. *Hat eine $n \times n$-Matrix n verschiedene Eigenwerte in K, so ist A diagonalisierbar.*

Beweis. Wir haben schon gezeigt, dass die Eigenwerte von A genau die Nullstellen des charakteristischen Polynoms $\det(A - tE_n)$ sind (Satz 71). Sind $c_1, \dots, c_n$ nun n Nullstellen des charakteristischen Polynoms, so existieren für $i = 1, \dots, n$ Eigenvektoren $\mathbf{v}_i \in \mathrm{Ker}(A - c_iE_n) \smallsetminus \{0\}$. Sind die c_i paarweise verschieden, so sind diese Eigenvektoren nach dem vorigen Satz linear unabhängig (Satz 76). Da $\dim K^n = n$ gilt, sind diese Vektoren darum eine Basis von K^n für die ${}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F) = \begin{pmatrix} c_1 & 0 & \dots & 0 \\ 0 & c_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & c_n \end{pmatrix}$ nach Konstruktion eine Diagonalmatrix ist. □

Bemerkung (Charakteristisches Polynom für lineare Abbildungen)**.** Wie $\det(A)$ ändert sich das charakteristische Polynom nicht, wenn wir eine andere Basis von K^n wählen, denn

$$\begin{aligned}\det(B^{-1}AB - tE_n) &= \det(B^{-1}AB - tB^{-1}E_nB) && \text{da } BE_nB^{-1} = BB^{-1} = E_n\\ &= \det(B^{-1}(A - tE_n)B) && \text{Distributivgesetz}\\ &= \det(B^{-1})\det(A - tE_n)\det(B)\\ &= \det(A - tE_n).\end{aligned}$$

Darum können wir das charakteristische Polynom auch für lineare Abbildung $F\colon V \to V$ eines n-dimensionalen K-Vektorraums V definieren, indem wir für eine beliebige Basis $\underline{v}$ von V

$$\det(F - t\,\mathrm{id}_V) := \det({}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F) - t\,{}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(\mathrm{id}_V)) = \det({}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F) - tE_n)$$

setzen.

Bemerkung (Die Spur einer Matrix)**.** Die vorangehende Bemerkung sagt, dass alle Koeffizienten des charakteristischen Polynoms bei Basiswechsel unverändert bleiben. Einige dieser Koeffizienten können wir mit der Leibnizformel explizit berechnen denn diese berechnet

$$\det\begin{pmatrix} a_{1,1}-t & a_{1,2} & \dots & a_{1,n}\\ a_{2,1} & a_{2,2}-t & \ddots & \vdots\\ \vdots & \ddots & \ddots & a_{n-1,n}\\ a_{n,1} & \dots & a_{n,n-1} & a_{n,n}-t\end{pmatrix}$$

als Summe über alle Produkte der Einträge, die entstehen, wenn wir aus jeder Zeile und jeder Spalte einen Koeffizienten auswählen. Für die Diagonale erhalten wir zum Beispiel den Summanden

$$\begin{aligned}&(a_{1,1}-t)\cdot(a_{2,2}-t)\cdots(a_{n,n}-t)\\ &= (-t)^n + (a_{1,1}+a_{2,2}+\cdots+a_{n,n})(-t)^{n-1} + \text{Terme mit kleineren Potenzen von } t.\end{aligned}$$

Bei den anderen Summanden der Leibnizformel kann nach ausmultiplizieren aber nirgends t^n oder t^{n-1} vorkommen, denn dafür müssten wir wenigstens $n-1$ Spalten, den Eintrag auf der Diagonalen auswählen. Wenn wir das tun, bleibt uns in der verbleibenden i-ten Spalte aber auch nur noch der Diagonaleintrag übrig, da wir noch einen Koeffizienten aus der in der i-ten Zeile auswählen müssen.

Also gilt

$$\det(A - tE_n) = (-t)^n + (\sum_{i=1}^{n} a_{i,i})(-t)^{n-1} + \text{Terme mit kleineren Potenzen von } t.$$

Insbesondere bleibt die Summe der Diagonaleinträge

$$\mathrm{Spur}(A) := a_{1,1} + a_{2,2} + \cdots + a_{n,n}$$

bei Basiswechsel unverändert.

Beispiel: Für $A = \begin{pmatrix} 0 & 1 & 0\\ 0 & 0 & 1\\ -2 & 1 & 2\end{pmatrix}$ ist

$$\mathrm{Spur}(A) = 0+0+2 = 2.$$

In unserer Basis aus Eigenvektoren hatten wir ${}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A) = \begin{pmatrix} 1 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & -1\end{pmatrix}$ berechnet. Hier gilt in der Tat wieder

$$\mathrm{Spur}({}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A)) = 1+2+(-1) = 2.$$

Das ist eine erste Probe, wenn Sie ein charakteristisches Polynom und die Eigenwerte berechnen.

Anwendung: Ein neuer Blick auf lineare Rekursionen

Lassen Sie uns das Beispiel $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 1 & 2 \end{pmatrix}$ noch einmal anschauen. Für diese Matrix gilt

$$A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 1 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_3 \\ -2x_1 + x_2 + 2x_3 \end{pmatrix}.$$

Wenn wir also eine Folge x_n rekursiv definieren durch Startwerte x_{-2}, x_{-1}, x_0 und dann

$$x_{n+1} := 2x_n + x_{n-1} - 2x_{n-2}$$

setzen, ist

$$A \begin{pmatrix} x_{n-2} \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} x_{n-1} \\ x_n \\ x_{n+1} \end{pmatrix}.$$

Also

$$A^n \begin{pmatrix} x_{-2} \\ x_{-1} \\ x_0 \end{pmatrix} = \begin{pmatrix} x_{n-2} \\ x_{n-1} \\ x_n \end{pmatrix}.$$

Eine allgemeine Formel für die x_n anzugeben ist also gleichbedeutend damit die Potenzen A^n von A zu berechnen.

Bemerkung (Allgemeine lineare Rekursionen)**.** Ist allgemeiner eine Folge (x_n) reeller Zahlen durch Startwerte $x_{-r}, \dots, x_0$ und die Formel $x_{n+1} := a_0 x_n + a_2 x_{n-1} + \cdots + a_r x_{n-r}$ für feste Zahlen $a_0, \dots, a_r \in K$ gegeben, so gilt für die $r+1 \times r+1$-Matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & & 0 \\ 0 & \cdots & 0 & 0 & 1 \\ a_r & a_{r-1} & \cdots & a_1 & a_0 \end{pmatrix}$$

genauso

$$A \begin{pmatrix} x_{n-r} \\ x_{(n-r)+1} \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_{(n-r)+1} \\ x_{(n-r)+2} \\ \vdots \\ x_{n+1} \end{pmatrix} \text{ und } A^n \begin{pmatrix} x_{-r} \\ x_{-r+1} \\ \vdots \\ x_0 \end{pmatrix} = \begin{pmatrix} x_{n-r} \\ x_{(n-r)+1} \\ \vdots \\ x_n \end{pmatrix}$$

Auf Blatt 11 haben Sie das charakteristische Polynom dieser Matrix bestimmt, es ist das Polynom vom Grad $r+1$ dessen Koeffizienten bis auf Vorzeichen genau die a_i sind.

Eine Formel für A^n kennen wir aber, wenn wir A diagonalisieren können, denn ist $\mathbf{v}_1, \dots, \mathbf{v}_{r+1}$ eine Basis aus Eigenvektoren zu den Eigenwerten $c_1 \dots, c_{r+1}$ so gilt $A^n \mathbf{v}_i = c_i^n \mathbf{v}_i$ also

$${}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A^n) = \begin{pmatrix} c_1^n & 0 & \cdots & 0 \\ 0 & c_2^n & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & c_{r+1}^n \end{pmatrix}.$$

Wenn wir A diagonalisieren könnnen, so können also für x_n eine explizite Formel finden.

IN UNSEREM BEISPIEL $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 1 & 2 \end{pmatrix}$ gilt für unsere Basis aus Eigenvektoren $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$, dass ${}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ und also

$$ {}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A)^n = \begin{pmatrix} 1^n & 0 & 0 \\ 0 & 2^n & 0 \\ 0 & 0 & (-1)^n \end{pmatrix}. $$

WAS BEDEUTET DAS KONKRET? Wenn wir den Vektor $\mathbf{v} = \begin{pmatrix} x_{-2} \\ x_{-1} \\ x_0 \end{pmatrix}$ als Linearkombination

$$ \mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + a_3\mathbf{v}_3 $$

der Eigenvektoren schreiben können wir

$$ \begin{aligned} A^n\mathbf{v} &= A^n(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + a_3\mathbf{v}_3) \\ &= a_1 \cdot A^n\mathbf{v}_1 + a_2 \cdot A^n\mathbf{v}_2 + a_3 \cdot A^n\mathbf{v}_3 \\ &= a_1 \cdot (1^n \cdot \mathbf{v}_1) + a_2(2^n \cdot \mathbf{v}_2) + a_3((-1)^n \cdot \mathbf{v}_3) \end{aligned} $$

leicht berechnen.

Die Gleichung $\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + a_3\mathbf{v}_3$ bedeutet, dass der Koordinatenvektor von $\mathbf{v}$ bezüglich der Basis $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ der Vektor der Koeffizienten $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ ist, d.h. $\mathrm{Koord}_{\underline{\mathbf{v}}}(\mathbf{v}) = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$.

Die Rechnung: Berechne erst die $\underline{\mathbf{v}}$-Koordinaten von $\mathbf{v}$, wende dann ${}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A^n)$ an und interpretiere die Koeffizienten als Linearkombination der $\mathbf{v}_i$ entpricht genau der komplizierten Formel

$$ A^n\mathbf{v} = \mathrm{Komb}_{\underline{\mathbf{v}}} \circ {}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(A^n) \circ \mathrm{Koord}_{\underline{\mathbf{v}}}(\mathbf{v}). $$

WÄHLEN WIR ALS STARTWERTE $x_{-2} = x_{-1} = 0$ und $x_0 = 1$ und wollen $A^n \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ berechnen müssen wir $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ als Linearkombination der $\mathbf{v}_i$ schreiben, d.h. wir lösen das Gleichungssystem $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + a_3\mathbf{v}_3$, das wir als

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

schreiben mit dem Gauß-Verfahren. Wir finden

Sie sollten das selbst einmal tun.

$$ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = -\frac{1}{2} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \frac{1}{3} \cdot \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} + \frac{1}{6} \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}. $$

Also ist

$$\begin{pmatrix} x_{n-2} \\ x_{n-1} \\ x_n \end{pmatrix} = A^n \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = -\frac{1}{2} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \frac{1}{3} \cdot 2^n \cdot \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} + \frac{1}{6} \cdot (-1)^n \cdot \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

und wir finden

$$x_n = -\frac{1}{2} + \frac{1}{3} \cdot 2^n \cdot 4 + (-1)^n \cdot \frac{1}{6}.$$

Bemerkung. 1. Das Verfahren können wir für jede lineare Rekursion verwenden, für die das charakteristische Polynom der zugehörigen Matrix $r+1$ verschiedene Nullstellen besitzt. Die Lösungsformel für x_n ist dann immer eine Linearkombination der n-ten Potenzen der Eigenwerte von A.

2. Das Prinzip erklärt insbesondere, wieso wir für die Fibonacci-Zahlen eine Formel erhalten haben die eine Linearkombination der Potenzen der Nullstellen von

$$t^2 - t - 1 = \det(\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} - tE_2)$$

war.

Aufgabe 7 (Knobelaufgabe)**.** Vielleicht ist Ihnen aufgefallen, dass im konkreten Beispiel der Rekursion $x_{n+1} = 2x_n + x_{n-1} - 2x_{n-2}$, die Koordinaten der Eigenvektoren $(1,1,1)$, $(1,2,4)$, $(1,-1,1)$ gerade die ersten Potenzen der zugehörigen Eigenwerte $1, 2$ und -1 waren.

1. Ist das Zufall? Falls nein, woran liegt das für Matrizen dieser speziellen Form?

2. Falls nicht, könnten Sie jetzt einen Beweis ohne Rechnung für die Aussage finden, dass die Vandermonde-Determinante genau dann $\neq 0$ ist, wenn die auftretenden Zahlen paarweise verschieden sind! Sehen Sie wie?

Vandermonde-Determinante:
$$\det \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & & & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{n-1} \end{pmatrix}$$

Eigenraumzerlegung und mehrfache Eigenwerte

Sie hatten mir in der Vorlesung Beispiele für 5×5-Matrizen vorgeschlagen, die ein charakteristisches Polynom mit einer dreifachen und einer doppelten Nullstelle, zum Beispiel $(3-t)^3(5-t)^2$ haben, nämlich:

$$A = \begin{pmatrix} 3 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix} \text{ und } B = \begin{pmatrix} 3 & 1 & 1 & 1 & 1 \\ 0 & 3 & 1 & 1 & 1 \\ 0 & 0 & 3 & 1 & 1 \\ 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix}$$

Die erste Matrix A ist diagonalisierbar, denn A ist schon selbst eine Diagonalmatrix. Das zweite Beispiel B besitzt keine Basis aus

diagonalisierbar=besitzt Basis aus Eigenvektoren

Eigenvektoren, denn die Eigenwerte von B sind die Nullstellen des charakteristischen Polynoms, also 3 und 5, aber

$$B - 3E_5 = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

hat schon fast Zeilenstufenform, so dass wir sehen können, dass diese Matrix Rang 4 hat und damit $\dim(\operatorname{Ker}(A - 3E_5)) = 1$ gilt, da der erste Standarbasisvektor e_1 in diesem Kern liegt ist

$$\operatorname{Ker}(B - 3E_5) = \operatorname{Span}(e_1).$$

Genauso ist $\dim(\operatorname{Ker}(B - 5E_5) = 1$, wir finden also auch zum Eigenwert 5 bis auf Vielfache nur einen Eigenvektor. Insbesondere gibt es für B also keine Basis aus Eigenvektoren.

Sie hatten in Aufgabe 1 von Blatt 2 beim Lösen der Rekursion $x_{n+1} = 2x_n - x_{n-1}$ schon gesehen, dass Sie, obwohl es für $A = \begin{pmatrix} 0 & 1 \\ -1 & 2 \end{pmatrix}$ keine Basis aus Eigenvektoren gibt, wir den Eigenvektor $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ durch $\mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ zu einer Basis ergänzen können für die $A\mathbf{v}_2 = v_2 + \mathbf{v}_1$ gilt und wir immer noch A^n bestimmen können. Wir werden noch sehen, dass wir für allgemeine Matrizen einen ähnlichen Ersatz für eine Basis aus Eigenvektoren finden können.

Wir hatten uns überlegt, dass für eine lineare Abbildung $F\colon V \to V$, die Eigenvektoren zum Eigenwert c genau die Elemente von $\operatorname{Ker}(F - c \cdot \mathrm{id}_V)$ sind.

Notation 78. Ist c ein Eigenwert der lineare Abbildung $F\colon V \to V$, so heißt

$$\begin{aligned} \operatorname{Eig}(F, c) &:= \{v \in V \mid F(v) = c \cdot v\} \\ &= \operatorname{Ker}(F - c \cdot \mathrm{id}_V) \end{aligned}$$

der *Eigenraum* von F zum Eigenwert c.

Bemerkung. 1. Da wir uns schon überlegt haben, dass Kerne von linearen Abbildungen immer Unterräume sind, müssen wir uns jetzt nicht neu überlegen, dass Eigenräume immer Untervektorräume sind.

2. Ist c ein Eigenwert von $F\colon V \to V$ so heißt die Dimension des Eigenraums $\dim \operatorname{Eig}(F, c)$ *geometrische Vielfachheit des Eigenwertes*.

   Die Vielfachheit der Nullstelle c des charakteristischen Polynoms $\det(F - t\,\mathrm{id}_V)$ heißt *algebraische Vielfachheit des Eigenwertes*.

Die Aussage, dass Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig sind, zeigt auch, dass wenn wir Basen aller Eigenräume einer Abbildung wählen, die erhaltenen Vektoren zusammen wieder linear unabhängig sind. Um diese Aussage kompakt zu formulieren führen wir das Symbol $\oplus$ ein:

Definition. Sind $U_1, \dots, U_r \subset V$ Untervektorräume eines K-Vektorraums V so bezeichnen wir mit

$$+_{i=1}^{r} U_i := \{\mathbf{v} \in V \mid \text{es gibt } \mathbf{u}_i \in U_i \text{ mit } \mathbf{v} = \mathbf{u}_1 + \dots + \mathbf{u}_r\}$$

die Summe der Unterräume.

Die Summe, heißt *direkte Summe*,

$$\bigoplus_{i=1}^{r} U_i = +_{i=1}^{r} U_i$$

wenn für alle $\mathbf{v} \in +_{i=1}^{r} U_i$ die Darstellung $\mathbf{v} = \mathbf{u}_1 + \dots + \mathbf{u}_r$ mit $\mathbf{u}_i \in U_i$ *eindeutig* ist.

Beispiel:

$$\mathrm{Span}(e_1, e_2) \oplus \mathrm{Span}(e_3, e_4) = K^4$$

aber für die Unterräume $U_1 = \mathrm{Span}(\begin{pmatrix} 1 \\ 1 \end{pmatrix}), U_2 = \mathrm{Span}(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, U_3 = \mathrm{Span}(\begin{pmatrix} 0 \\ 1 \end{pmatrix}) \subset \mathbb{R}^2$ ist $+_{i=1}^{n} U_i$ keine direkte Summe, da zum Beispiel

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

auch als

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

geschrieben werden kann.

Aufgabe 8. Zeigen Sie, dass für eine direkte Summe von Unterräumen

$$\dim \bigoplus_{i=1}^{r} U_i = \sum_{i=1}^{r} \dim U_i$$

gilt.

Folgerung 79. *Sind $c_1, \dots, c_r$ Eigenwerte der Abbildung $F\colon V \to V$ so enthält V die direkte Summe der zugehörigen Eigenräume:*

$$\oplus_{i=1}^{r} \mathrm{Eig}(F, c_i) \subseteq V.$$

Beweis. Wir müssen nur zeigen, dass für Vektoren $\mathbf{v} = u_1 + \dots + u_r$ mit $u_i \in \mathrm{Eig}(F, c_i)$ die Darstellung als Summe eindeutig ist.

Angenommen $\mathbf{v} = u_1' + \dots + u_r'$ mit $u_i' \in \mathrm{Eig}(F, c_i)$ ist eine weitere Darstellung, dann ist

$$\mathbf{v} - \mathbf{v} = \underbrace{(u_1 - u_1')}_{\in \mathrm{Eig}(F, c_1)} + \dots + \underbrace{(u_r - u_r')}_{\in \mathrm{Eig}(F, c_r)} = 0.$$

Da $u_i - u_i'$ jeweils im Eigenraum zum Eigenwert c_i liegen und Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig sind, kann diese Linearkombination nur 0 sein, wenn alle $u_i - u_i' = 0$ sind, d.h. $u_i = u_i'$ für alle $i = 1, \dots, r$. □

Im Beispiel der nicht diagonalisierbaren Matrix B können wir einen Ersatz für den Eigenraum zum Eigenwert 3 sehen, der die erwartete Dimension 3 hat: Auf dem Unterraum $\mathrm{Span}(e_1, e_2, e_3)$ erfüllt die Abbildung $B - 3E_5$ nämlich $(B - 3E_5)^3(ae_1 + be_2 + ce_3) = 0$, denn

$$B = \begin{pmatrix} 3 & 1 & 1 & 1 & 1 \\ 0 & 3 & 1 & 1 & 1 \\ 0 & 0 & 3 & 1 & 1 \\ 0 & 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix}$$

$$\begin{aligned} (B - 3E_5)(ae_1 + be_2 + ce_3) &= a \cdot 0 + be_2 + c(e_1 + e_2) \\ (B - 3E_5)(be_2 + c(e_1 + e_2)) &= (b + c)e_1 \\ (B - 3E_5)((b + c)e_1) &= 0. \end{aligned}$$

Allgemein müssen wir $B - 3E_5$ nur durch $F - c\,\mathrm{id}_V$ ersetzen.

Satz 80 (Hauptraumzerlegung)**.** *Sei $F\colon V \to V$ eine lineare Abbildung mit charakteristischem Polynom*

$$\det(F - t\,\mathrm{id}_V) = \prod_{i=1}^{r}(c_i - t)^{n_i}$$

wobei $c_1, \ldots, c_r \in K$ paarweise verschieden sind. Dann gilt

$$\oplus_{i=1}^{r} \mathrm{Ker}((F - c_i\,\mathrm{id}_V)^n) = V$$

Bemerkung. Die Unterräume $\mathrm{Ker}((F - c_i\,\mathrm{id}_V)^n)$ heißen auch *Haupträume zum Eigenwert c_i* (oder verallgemeinerte Eigenräume) der Abbildung F.

Der Satz sagt also, dass wir K^n immer in verallgemeinerte Eigenräume zerlegen können, wenn wir nur das charakteristische Polynom in Linearfaktoren zerlegen können. Für $K = \mathbb{C}$ ist das zum Beispiel immer möglich.

Als Vorbereitung mit einfacheren Formeln, schauen wir zunächst den Eigenwert 0 an.

Behauptung 81. *Sie $F\colon V \to V$ eine lineare Abbildung eines endlichdimensionalen K-Vektorraums V. Dann gilt:*

1. *Die Unterräume*

   $$\mathrm{Ker}(F) \subseteq \mathrm{Ker}(F^2) \subseteq \cdots \subseteq \mathrm{Ker}(F^n)$$

   sind ineinander enthalten und es gibt ein $m \leq n$ so dass $\mathrm{Ker}(F^m) = \mathrm{Ker}(F^{m+1}) = \mathrm{Ker}(F^{m+k})$ für alle $k \in \mathbb{N}$. Genauso gilt

   $$\mathrm{Bild}(F) \supseteq \mathrm{Bild}(F^2) \supseteq \cdots \supseteq \mathrm{Bild}(F^n)$$

   und für das gleiche m wie zuvor gilt

   $$\mathrm{Bild}(F^m) = \mathrm{Bild}(F^{m+1}) = \mathrm{Bild}(F^{m+k}).$$

2. *Es gilt $\mathrm{Ker}(F^m) \oplus \mathrm{Bild}(F^m) = V$ und die Abbildung F bildet die Unterräume $\mathrm{Ker}(F^m)$ und $\mathrm{Bild}(F^m)$ auf sich selbst ab, auf $\mathrm{Bild}(F^m)$ ist F sogar ein Isomorphismus.*

Beweis. 1. Die Inklusionen folgen aus der Definition, denn ist $\mathbf{v} \in \mathrm{Ker}(F^m)$ so gilt $F^m(\mathbf{v}) = 0$ also ist auch $F^{m+1}(\mathbf{v}) = F(F^m(\mathbf{v})) = F(0) = 0$.

Genauso ist jeder Vektor $\mathbf{v} \in \mathrm{Bild}(F^{m+1})$ nach Definition von der Form $\mathbf{v} = F^{m+1}(\mathbf{w}) = F^m(F(\mathbf{w}))$ also ist $\mathbf{v}$ dann insbesondere auch in $\mathrm{Bild}(F^m)$.

Wenn $\mathrm{Ker}(F^m) \subsetneq \mathrm{Ker}(F^{m+1})$ gilt, so ist $\dim(\mathrm{Ker}(F^m)) < \dim(\mathrm{Ker}(F^{m+1}))$, die Folge der Dimensionen ist also aufsteigend aber durch $\dim V$ beschränkt, also muss die Folge spätestens nach n Schritten unverändert bleiben.

Wegen der Dimensionsformel gilt für alle m:

$$\dim(\mathrm{Ker}(F^m)) + \dim(\mathrm{Bild}(F^m)) = \dim V$$

also muss die Folge der Bilder für das gleiche m konstant werden. Das bedeutet aber insbesondere, dass die Abbildung F auf $\mathrm{Bild}(F^m) \twoheadrightarrow \mathrm{Bild}(F^{m+1}) = \mathrm{Bild}(F^m)$ ein Isomorphismus ist.

2. Wir müssen nur zeigen, dass $\mathrm{Ker}(F^m) \oplus \mathrm{Bild}(F^m)$ eine direkte Summe ist, denn dann muss dieser Unterraum wegen der Dimensionsformel schon ganz V sein.

   Angenommen $\mathbf{v} = u_K + u_B = u'_K + u'_B$ mit $u_K, u'_K \in \mathrm{Ker}(F^m), u_B, u'_B \in \mathrm{Bild}(F^m)$. Dann ist

   $$u_K - u'_K = u'_B - u_B \in \mathrm{Ker}(F^m) \cap \mathrm{Bild}(F^m).$$

   Es gilt aber $\mathrm{Ker}(F^m) \cap \mathrm{Bild}(F^m) = \{0\}$, denn für Vektoren $\mathbf{w} \in \mathrm{Bild}(F^m)$ existiert nach Definition ein $u \in V$ so dass $\mathbf{w} = F^m(u)$. Wäre $\mathbf{w} \in \mathrm{Ker}(F^m)$, d.h $0 = F^m(\mathbf{w}) = F^m(F^m(u)) = F^{2m}(u)$, ist $u \in \mathrm{Ker}(F^{2m})$. Aber m ist so gewählt, dass $\mathrm{Ker}(F^{2m}) = \mathrm{Ker}(F^m)$, also ist $u \in \mathrm{Ker}(F^m)$ und darum $0 = F^m(u) = \mathbf{w}$. Also ist $\mathrm{Bild}(F^m) \cap \mathrm{Ker}(F^m) = \{0\}$.

□

Bemerkung. Wählen wir in der Situation der Behauptung Basen $\mathbf{v}_1, \ldots, \mathbf{v}_k$ von $\mathrm{Ker}(F^m)$ und $\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n$ von $\mathrm{Bild}(F^m)$ so ist die Matrix von F bezüglich dieser Basis

$${}_{\underline{\mathbf{v}}}\mathrm{Mat}_{\underline{\mathbf{v}}}(F) = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

eine Block-Matrix denn F bildet den $\mathrm{Ker}(F^m) = \mathrm{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ und $\mathrm{Bild}(F^m) = \mathrm{Span}(\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n)$ auf sich selbst ab. Zudem ist $\det(B) \neq 0$, da F auf $\mathrm{Bild}(F^m)$ ein Isomorphismus ist.

Da die Determinante einer Block-Matrix das Produkt der Determinanten der Blöcke ist, sehen wir daraus dass charakteristische Polynom von F das Produkt der charakteristischen Polynome der Abbildungen auf den Unterräumen $\mathrm{Ker}(F^m), \mathrm{Bild}(F^m)$ ist.

Auf dem Unterraum $\mathrm{Ker}(F^m)$ wissen wir außerdem, dass 0 der einzige mögliche Eigenwert ist, denn ist $\mathbf{v} \in \mathrm{Ker}(F^m)$ ein Eigenvektor zum Eigenwert c so gilt einerseits $F^m(\mathbf{v}) = c^m\mathbf{v}$ und andererseits $F^m(\mathbf{v}) = 0$ weil $\mathbf{v}$ im Kern der Abbildung liegt.

Bemerkung. Im Beweis haben wir außerdem gezeigt, dass für zwei Unterräume $U_1, U_2 \subseteq V$ die Summe genau dann eine direkte Summe ist wenn $U_1 \cap U_2 = \{0\}$ gilt, d.h.

$$U_1 \oplus U_2 \subset V \Leftrightarrow U_1 \cap U_2 = \{0\}.$$

Für mehr als zwei Unterräume, ist die Bedingung komplizierter.

Um aus der Behauptung das Ergebnis des Satzes abzuleiten, zahlt es sich jetzt aus, dass wir unsere Resultate für allgemeine Vektorräume formuliert haben, denn so können wir unsere Resultate auch auf Teilräume (wie $\mathrm{Bild}(F^m) \subset V$) anwenden, in denen es keine „Standardbasis“ gibt. Damit können wir einen Induktionsbeweis angeben.

Beweis von Satz 80 (Hauptraumzerlegung). Wir beweisen die Aussage mit Induktion über die Anzahl der Eigenwerte von F.

Vorbemerkung: Das charakteristische Polynom der Abbildung $F - c\,\mathrm{id}$ ist $\det(F - c\,\mathrm{id} - t\,\mathrm{id}_V) = \det(F - (c+t)\,\mathrm{id}_V)$. Ist $p(t) = \det(F - t\,\mathrm{id}_V)$ das charakteristische Polynom von F, so ist das charakteristische Polynom von $F - c\,\mathrm{id}_V$ also $p(t+c)$, insbesondere sind die Eigenwerte von $F - c\,\mathrm{id}_V$ die Zahlen $c_1 - c, c_2 - c, \ldots, c_r - c$.
Induktionsstart: Ist $r = 1$, so ist

$$\det(F - t\,\mathrm{id}_V) = (c_1 - t)^n$$

mit $n = \dim_K(V)$. Wenden wir die gerade bewiesene Behauptung auf die Abbildung $G := F - c_1\,\mathrm{id}_V$ an, so finden wir ein m, so dass:

$$V = \mathrm{Ker}(G^m) \oplus \mathrm{Bild}(G^m) = \mathrm{Ker}((F - c_1\,\mathrm{id}_V)^m) \oplus \mathrm{Bild}((F - c_1\,\mathrm{id}_V)^m)$$

und die Abbildung G induziert einen Isomorphismus auf dem Unterraum $\mathrm{Bild}(G^m)$.

Für den Satz müssen wir zeigen, dass $\mathrm{Bild}(G^m) = 0$ gilt. Dafür wählen wir wie in der letzten Bemerkung Basen $\mathbf{v}_1, \ldots, \mathbf{v}_k$ von $\mathrm{Ker}(G^m)$ und $\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n$ von $\mathrm{Bild}(G^m)$. Dann ist

$$\underline{\mathbf{v}}\mathrm{Mat}_{\underline{\mathbf{v}}}(G) = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

und also gilt für das charakteristische Polynom $(-t)^n = \det(G - t\,\mathrm{id}_V) = \det(A - t\,\mathrm{id}_{\mathrm{Ker}(G^m)}) \cdot \det(B - t\,\mathrm{id}_{\mathrm{Bild}(G^m)})$. Da 0 eine n-fache Nullstelle von $(-t)^n$, aber keine Nullstelle von $\det(B - t\,\mathrm{id}_V)$ ist, kann der Faktor $\det(B - t\,\mathrm{id}_V)$ nicht vorkommen, d.h. es gilt $\mathrm{Bild}(G^m) = \{0\}$.

Damit haben wir gezeigt, dass im Fall $r = 1$ gilt, dass $V = \mathrm{Ker}((F - c_1\,\mathrm{id}_V)^n)$.

Der Induktionsstart $r = 1$ ist selbst eine interessante Aussage: Wir haben gerade gezeigt, dass wir am charakteristischen Polynom ablesen können, ob $F^n = 0$ ist.

Induktionsschritt: Ist der Satz für eine Zahl $r = k$ richtig, so auch für die nächste $k+1$: Sei also

$$\det(F - t\,\mathrm{id}_V) = \prod_{i=1}^{k+1}(c_i - t)^{n_i}.$$

Wir wenden die Behauptung auf die Abbildung $G := F - c_{k+1}\,\mathrm{id}_V$ an und finden ein $m \leq \dim V$ so dass

$$V = \mathrm{Ker}(G^m) \oplus \mathrm{Bild}(G^m) = \mathrm{Ker}((F - c_{k+1}\,\mathrm{id}_V)^m) \oplus \mathrm{Bild}((F - c_{k+1}\,\mathrm{id}_V)^m)$$

und G ist auf dem Unterraum $\mathrm{Bild}(G^m)$ ein Isomorphismus.

Wir haben uns schon überlegt, dass das charakteristische Polynom von G das Produkt der charakteristischen Polynome der Einschränkungen $G|_{\mathrm{Ker}(G^m)}$ und $G|_{\mathrm{Bild}(G^m)}$ der Abbildung auf die Unterräume $\mathrm{Ker}(G^m)$ und $\mathrm{Bild}(G^m)$ ist.

Das Symbol

$$G|_{\mathrm{Bild}(G^m)}\colon \mathrm{Bild}(G^m) \to \mathrm{Bild}(G^m)$$

notiert nur, dass wir vorübergehend in G nur Elemente aus dem Unterraum $\mathrm{Bild}(G^m)$ einsetzen.

$$\prod_{i=1}^{k+1}(c_i - c_{k+1} - t)^{n_i} = \det(G|_{\mathrm{Ker}(G^m)} - t\,\mathrm{id}_{\mathrm{Ker}(G^m)}) \cdot \det(G|_{\mathrm{Bild}(G^m)} - t\,\mathrm{id}_{\mathrm{Bild}(G^m)})$$

Das Polynom $\det(G|_{\mathrm{Bild}(G^m)} - t\,\mathrm{id}_{\mathrm{Bild}(G^m)})$ hat bei 0 keine Nullstelle, da B invertierbar ist und $\det(A - t\,\mathrm{id}_{\mathrm{Ker}(G^m)})$ hat nur die Nullstelle 0, da G auf $\mathrm{Ker}(G^m)$ nur den Eigenwert 0 hat. Also ist

$$\det(B - t\,\mathrm{id}_{\mathrm{Bild}(G^m)}) = \prod_{i=1}^{k}(c_i - c_{k+1} - t)^{n_i}.$$

Nach Induktionsannahme können wir den Satz also auf die Einschränkung von $G|_{\mathrm{Bild}(G^m)}$ von G auf den Unterraum $\mathrm{Bild}(G^m)$ anwenden und erhalten:

$$\mathrm{Bild}(G^m) = \oplus_{i=1}^{k} \mathrm{Ker}((G|_{\mathrm{Bild}(G^m)} - (c_i - c_{k+1})\,\mathrm{id}_{\mathrm{Bild}(G^m)})^n).$$

und auf da $F - c_i\,\mathrm{id}_V$ auf $\mathrm{Ker}(G^m) = \mathrm{Ker}((F - c_{k+1}\,\mathrm{id}_V)^n)$ invertierbar ist, ändert sich der Kern nicht, wenn wir wieder ganz V als Definitionsbereich zulassen:

$$\mathrm{Ker}((G|_{\mathrm{Bild}(G^m)} - (c_i - c_{k+1})\,\mathrm{id}_{\mathrm{Bild}(G^m)})^n) = \mathrm{Ker}((F - c_i\,\mathrm{id}_V)^n).$$

Insgesamt ist also:

$$\begin{aligned} V &= \mathrm{Ker}(G^m) \oplus \mathrm{Bild}(G^m) \\ &= \mathrm{Ker}((F - c_{k+1}\,\mathrm{id}_V)^m) \oplus \left(\oplus_{i=1}^{k} \mathrm{Ker}((F - c_i\,\mathrm{id}_V)^n)\right). \end{aligned}$$

Das war zu zeigen. □

Ausblick: Potenzen von stochastischen Matrizen und PageRank

Ganz am Anfang des Semesters hatten wir gesehen, wie wir mit Hilfe von linearen Gleichungssystemen versuchen können, die Wichtigkeit von Knoten in einem Netzwerk zu bewerten. Dazu hatten wir dem Netzwerk eine $N \times N$-Matrix A zugeordnet – N war hier die Anzahl der Knoten im Netzwerk – die die besondere Eigenschaft hatte, dass die Einträge alle nicht-negativ waren und die Summe der Einträge in jeder Spalte jeweils 1 ergab. Für die Bewertung haben wir nun einen Vektor $\mathbf{v} \in \mathbb{R}^N \smallsetminus \{0\}$ gesucht, für den $A\mathbf{v} = \mathbf{v}$ galt, d.h. $\mathbf{v}$ ist ein Eigenvektor zum Eigenwert 1.

Nachdem wir „Zeilenrang=Spaltenrang“ gezeigt hatten, konnten wir zeigen, dass es für diese Matrizen tatsächlich immer einen Eigenvektor zum Eigenwert 1 gibt.

Leider ist N für große Netzwerke so groß, dass selbst das Gauß-Verfahren nicht praktikabel ist. Mit Hilfe von Eigenwerten können wir ein ganz anderes Verfahren erklären.

Satz 82 (Satz von Perron für stochastische Matrizen)**.** *Ist $A = (a_{i,j})_{i,j=1,\dots n} \in \mathrm{Mat}_{n,n}(\mathbb{R})$ eine reelle $n \times n$-Matrix mit nicht-negativen Einträgen ($a_{i,j} \geq 0$ für alle i, j) und Spaltensumme 1 (d.h. $\sum_{i=1}^{n} a_{i}j = 1$ für alle j) dann gilt:*

1. *Die Zahl 1 ist ein Eigenwert von A und alle anderen (möglicherweise komplexen) Eigenwerte c von A erfüllen $|c| \leq 1$.*

2. *Gilt zusätzlich dass die Einträge von A positiv sind, d.h. $a_{i,j} > 0$ für alle i, j so ist 1 eine einfache Nullstelle des charakteristischen Polynoms und alle anderen Eigenwerte erfüllen $|c| < 1$.*

DIESEN SATZ KÖNNTEN wir jetzt ohne weitere Hilfsmittel beweisen, das ist – wenn wir uns an den Trick, die transponierten Matrix anzuschauen erinnern – überraschend einfach und wir holen das im kommenden Semester nach.

Nach unserem Satz zur Hauptraumzerlegung können wir für eine gegebene Matrix $A\colon \mathbb{C}^N \to \mathbb{C}^N$ jeden Vektor $\mathbf{v} \in \mathbb{C}^N$ eindeutig als Summe:

$$\mathbf{v} = \mathbf{v}_1 + \cdots + \mathbf{v}_r$$

mit $\mathbf{v}_i \in \operatorname{Ker}((A - c_i \operatorname{id}_{K^N})^n)$ schreiben, wobei $c_1, \ldots, c_r$ die Eigenwerte von A sind.

Für einen Eigenvektor $\mathbf{v}_i$ zum Eigenwert c_i ist $A^n\mathbf{v}_i = c_i^n\mathbf{v}_i$. Wenn $|c_i| < 1$ ist, werden die Potenzen $|c_i^n|$ immer kleiner, d.h. $A^n v_i \to_{n\to\infty} 0$ ist eine Folge von Vektoren, die gegen 0 konvergiert.

Die gleiche Eigenschaft gilt auch, wenn $\mathbf{v}_i$ nur im Hauptraum zum Eigenwert c_i mit $|c_i| < 1$ liegt. (Das ist nicht so schwer nachzurechnen, denn für Eigenvektoren stimmt das und die Differenz von F und c_i id (das operiert wie auf Eigenvektoren) ist auf dem Hauptraum so, dass eine Potenz 0 ist – ein genaues Argument wäre eine kurze Rechnung, die nicht mehr in die letzte Vorlesung passt.)

Sei nun A eine Matrix wie im Satz von Perron, d.h. A hat die den Eigenwert 1 als einfache Nullstelle des charakteristischen Polynoms mit Eigenvektor $\mathbf{v}_1$ und alle anderen Eigenwerte von A haben Betrag < 1.

Dann wird für jeden Vektor

$$\mathbf{v} = a_1\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_r$$

mit $\mathbf{v}_i \in \operatorname{Ker}((A - c_i \operatorname{id}_{K^N})^n)$

$$A^n\mathbf{v} \to_{n\to\infty} a_1\mathbf{v}_1$$

gegen $a_1\mathbf{v}_1$ konvergieren, denn die Vektoren $A^n\mathbf{v}_2, \ldots, A^n\mathbf{v}_r$ konvergieren alle gegen den 0-Vektor. Wir können damit einen Eigenvektor näherungsweise bestimmen, indem wir $A^n\mathbf{v}$ für einen zufällig gewählten Vektor $\mathbf{v}$ ausrechnen.

LASSEN SIE UNS das an einem Beispiel ausprobieren. (In der Vorlesung haben wir ein Beispie in SAGE ausprobiert.)

SIE KÖNNTEN EINWENDEN, dass im Beispiel einige Einträge von A gleich 0 waren und darum der Satz nicht anwendbar ist. Bei PageRank wurde das einfach so korrigiert, dass A für eine kleine Zahl ϵ durch die Matrix

$$(1-\epsilon)A + \epsilon \cdot (\text{Matrix aus lauter 1sen})$$

ersetzt wurde. Darauf können wir den Satz anwenden und das Ergebnis funktioniert in der Praxis gut.

MEHR DAZU im nächsten Semester, wenn Sie neugierig sind können Sie zum Beispiel im Buch„Google's PageRank and beyond: the science of search engine rankings“[17], das über die Universitätsbibliothek online verfügbar ist mehr dazu lesen.

Es gibt für die veränderte Matrix ein schönes Modell: Die Matrix A beschreibt das Verhalten von Nutzern, die im Netz mit gleicher Wahrscheinlichkeit auf einen der Links auf der gegebenen Seite klicken, die neue Matrix beschreibt das Verhalten, wenn außerdem mit der Wahrscheinlichkeit ϵ stattdessen eine zufällige neue Seite gewählt wird.

[17] Amy N. Langville and Carl D. Meyer. *Google's PageRank and beyond: the science of search engine rankings*. Princeton University Press, Princeton, NJ, 2012. Paperback edition of the 2006 original, online verfügbar

Das ist eine weitere überraschende Anwendung von unseren Ergebnissen über Matrizen und lineare Abbildungen. Wenn wir gleich mit dem Problem angefangen hätten, hätten wir angesichts der Größe der Netzwerke, die uns interessieren wahrscheinlich den Überblick verloren. Die kompakte Sprache der Mathematik hat uns ermöglicht, das Problem erst an einfachen Beispielen — wir kamen immer mit kleinen Matrizen aus — so gut zu verstehen, dass wir das dann auf viel kompliziertere Probleme anwenden konnten. Die Mathematik hat also ein kompliziertes Problem sehr viel einfacher gemacht, weil wir das Problem damit kompakt formulieren konnten (Suche einen Eigenvektor zum größten Eigenwert). Das Problem konnten wir dann im kleinen leicht verstehen und damit allgemein lösen. Abstraktion lohnt sich.

Die verschiedenen Bewertungsalgorithmen von Netzwerken können Sie für gute und schlechte Zwecke verwenden. Im 2022 erschienenen Buch[18] der Friedensnobelpreisträgerin Maria Ressa werden Sie am Ende Bewertungsprofile im Facebook-Netzwerk finden, die erklären wieso sich manche Falschinformationen über spezielle Konten stark verbreiten.

[18] Maria Ressa. *How to stand up to a dictator*. WH Allen, 2022. ISBN 978-0753559192

Glossar mathematischer Symbole

Mathematische Symbole haben meist eine genauere Bedeutung, als die entsprechenden Wörter, beim Lesen ersetzen wir die Symbole im Kopf durch Wörter. Damit können Sie in Ihren Lösungen prüfen, ob das was Sie aufschreiben lesbar sein könnte.

$=$ „(ist) gleich“

$\neq$ „(ist) nicht gleich“

$\Rightarrow$ „daraus folgt“ oder „impliziert, dass“

$\Leftrightarrow$ „(gilt) genau dann, wenn“

\# Anzahl

$\{?\,|\,\text{etwas}\}$ „Die Menge der ? für die etwas gilt“

$\in$ „aus“ oder „Element von“

$\ni$ Ist das umgekehrte „Element von“, d.h. „$x \in M$“ und „$M \ni x$“ bedeuten beide „x ist Element von M“

$\subseteq$ „enthalten in“ oder „Teilmenge von“

$\supseteq$ „enthält“ oder „Obermenge von“

$\cup$ „vereinigt mit“ $\cup$ ist das Symbol für die Vereinigung zweier Mengen (d.h. für die Menge die Entsteht, wenn wir die Elemente der beiden Mengen zu einer Menge zusammenfassen).

$\cap$ „geschnitten mit“ $\cap$ ist das Symbol für den Durchschnitt zweier Mengen (d.h. die Menge, die aus den Elementen besteht, die in beiden Mengen enthalten sind).

$M \setminus N$ „Die Menge M ohne die Elemente aus N“, bzw. „Entferne N aus der Menge M“. Das Symbol „$\setminus$“ ist das Minuszeichen für Mengen, also „entferne“.

$\forall$ „Für alle“

\# „Anzahl“

$\circ$ „verknüpft mit“. Das Symbol „$\circ$“ bezeichnet die Komposition von Abbildungen, d.h. $f \circ g(x) := f(g(x))$.

Literaturverzeichnis

Thomas Bauer and Lisa Hefendehl-Hebecker. *Mathematikstudium für das Lehramt an Gymnasien*. Springer Spektrum Wiesbaden, 2019. ISBN 978-3-658-26681-3. URL https://doi.org/10.1007/978-3-658-26682-0.

Gerd Fischer. *Lineare Algebra*, volume 17 of *Grundkurs Mathematik*. Friedr. Vieweg & Sohn, Braunschweig, fifth edition, 1979. ISBN 3-528-17217-7. URL https://link.springer.com/book/10.1007/978-3-8348-9365-9. In collaboration with Richard Schimpl.

Ulrich Görtz. Lineare algebra. Vorlesungsskript, 2020. URL https://math.ug/lecture-notes.html.

Amy N. Langville and Carl D. Meyer. *Google's PageRank and beyond: the science of search engine rankings*. Princeton University Press, Princeton, NJ, 2012. Paperback edition of the 2006 original, online verfügbar.

David Lay, Steven Lay, and Judi McDonald. *Linear Algebra and Its Applications*. Harlow: Pearson Education, Limited, 2015. ISBN 9781292092232. URL https://elibrary.pearson.de/book/99.150005/9781292092249.

Maria Ressa. *How to stand up to a dictator*. WH Allen, 2022. ISBN 978-0753559192.

Wolfgang Soergel. Lineare algebra. Vorlesungsskript, 2022. URL http://home.mathematik.uni-freiburg.de/soergel/Skripten/XXLA1.pdf.