

Lineare Algebra

Gertrud Desch

Zur Neufassung des Skriptums 2016/17:

Das ist die Neufassung eines Skriptums, das schon seit vielen Jahren meine Vorlesung über Lineare Algebra 1 und 2 an der Karl-Franzens-Universität Graz begleitet.

Der Inhalt hat sich nicht wesentlich geändert. Durch den Aufbau möchte ich meiner Erfahrung Rechnung tragen, dass zu viele abstrakte Begriffe zu Beginn die Anfängerinnen und Anfänger leicht verwirren können, und dass die Intuition und auch das klare Denken oft unter einem überwältigenden Aufwand an Formalismen ersticken. Daher entwickle ich in diesem Skriptum einen Teil der linearen Algebra zuerst nur für den aus der Schulmathematik für kleine Dimensionen bekannten, wenn nicht vertrauten, Raum \mathbb{R}^n der n -dimensionalen Spaltenvektoren. Der Kopf der Leserinnen und Leser soll zu diesem Zeitpunkt noch unbeschwert sein von abstrakten algebraischen Konzepten, und sich ganz auf die Funktionsweise der Vektor- und Matrizenrechnung konzentrieren dürfen. Es bietet auch dieses Umfeld genug Material, um Definieren und Beweisen zu üben und den Zusammenhang zwischen Lehrsätzen und Rechenmethoden zu erfassen.

Natürlich werden im weiteren Verlauf der Vorlesung im Sinne des Studienplanes auch allgemeine Körper und der abstrakte Vektorraumbegriff einbezogen.

Gertrud Desch, 30. Mai 2017
gertrud.desch@uni-graz.at
Telefon (0316) 380 5177
Karl-Franzens-Universität Graz
Institut für Mathematik und Wissenschaftliches Rechnen
Heinrichstrasse 36, oberster = 4. Stock, Zimmer 505,
durch die Glastür gleich beim Lift, dann vorletzte Tür
links.

Inhaltsverzeichnis

0. Einleitung	1
0.1. Mathematik verstehen	1
0.2. Der mathematische Lösungsprozess	3
0.3. Ein paar allgemeine Tipps für die Arbeit	5
Kapitel 1. Matrizenrechnung	9
1. Notation	9
1.1. Zahlen, Vektoren, Matrizen	9
1.2. Mengen	10
2. Lineare Gleichungssysteme	11
2.1. Grundbegriffe	11
2.2. Das Lösungsprinzip	14
2.3. Systeme mit keiner oder mehrfachen Lösungen	17
2.4. Der fertige Algorithmus	19
3. Analytische Geometrie	20
3.1. Endlichdimensionale affine Räume	20
3.2. Geometrische Aufgaben und lineare Gleichungssysteme	23
3.3. Unterräume und affine Unterräume	26
4. Basis und Dimension	30
4.1. Lineare Abhängigkeit	30
4.2. Basis	33
4.3. Basiserweiterung	35
5. Matrix mal Vektor	36
5.1. Funktionen	37
5.2. Produkt einer Matrix mit einem Vektor	40
5.3. Matrizen und lineare Abbildungen	42
5.4. Kern, Spaltenraum und Zeilenraum	45
5.5. Basen und Dimension von Kern, Spaltenraum und Zeilenraum	47
6. Matrix mal Matrix	52
6.1. Linearkombination von Matrizen	52
6.2. Hintereinanderausführung von Funktionen, Identität, Umkehrfunktion	53
6.3. Matrixprodukt	56
6.4. Reguläre Matrizen	60
6.5. Berechnung der inversen Matrix	63
Kapitel 2. Körper	69
1. Körperaxiome	69
1.1. Grundrechnungsarten	69
1.2. Spaltenvektoren über allgemeinen Körpern	73
2. Beispiele von Körpern	75
2.1. Komplexe Zahlen	75
2.2. Restklassen	83
Kapitel 3. Inneres Produkt und Orthogonalität	89

INHALTSVERZEICHNIS

1. Das euklidische innere Produkt	89
1.1. Motivation aus der Elementargeometrie	89
1.2. Definition von inneren Produkten und Normen	90
1.3. Adjungierte Matrix	94
2. Orthogonalität	94
2.1. Orthogonales Komplement	94
2.2. Orthonormalsysteme	98
Kapitel 4. Der abstrakte Vektorraum	103
1. Definitionen und Grundbegriffe	103
1.1. Definition von Vektorräumen, einfache Rechenregeln	103
1.2. Linearkombinationen und Unterräume	105
1.3. Lineare Abbildungen in allgemeinen Vektorräumen	106
1.4. Isomorphismen	109
1.5. Normen und Skalarprodukte in allgemeinen Vektorräumen	111
2. Funktionenräume	117
2.1. Definition von Funktionenräumen	117
2.2. Lineare Abbildungen in Funktionenräumen	119
2.3. Normen und Skalarprodukte in Funktionenräumen	120
3. Basis und Koordinaten	125
3.1. Lineare Unabhängigkeit und Basis	125
3.2. Koordinaten	128
3.3. Koordinatentransformation	131
4. Vektorräume aus Vektorräumen konstruieren	135
4.1. Durchschnitt und Summenraum	135
4.2. Direkte Summen	137
4.3. Faktorräume	141
Kapitel 5. Determinanten	147
1. Definition und Berechnung durch Zeilenumformungen	147
1.1. Definition der Determinante	147
1.2. Determinante und Zeilenumformungen	148
1.3. Dreiecksmatrizen und Blockmatrizen	150
1.4. Die Eindeutigkeit der Determinante	152
2. Multiplikationssatz für Determinanten	152
2.1. Determinanten von Elementarmatrizen	152
2.2. Der Multiplikationssatz	153
2.3. Determinante als Kriterium für Regularität	153
3. Existenz der Determinante	154
3.1. Permutationen	154
3.2. Die Formel von Leibniz	157
4. Entwicklungssatz von Laplace	159
4.1. Entwicklungssatz von Laplace	159
4.2. Die Komplementärmatrix und die Inverse	161
4.3. Cramersche Regel	162
4.4. Das äußere Produkt im dreidimensionalen Raum	163
Kapitel 6. Eigenwerte	165
1. Polynome	165
1.1. Polynom und Polynomfunktion	165
1.2. Division mit Rest	168
1.3. Nullstellen	170
2. Eigenvektoren und Eigenwerte	172

INHALTSVERZEICHNIS

2.1. Beispiele	172
2.2. Eigenwerte, Eigenvektoren, Eigenraum	173
2.3. Berechnung von Eigenwerten	174
2.4. Charakteristisches Polynom	175
3. Diagonalisierung	178
3.1. Lineare Unabhängigkeit von Eigenvektoren	178
3.2. Diagonalisierung einer Matrix	180
3.3. Anwendungsbeispiele	182
4. Hauptraumzerlegung	184
4.1. Invariante Unterräume	184
4.2. Lemma von Fitting	185
4.3. Haupträume	187
4.4. Jordansche Normalform	192
5. Matrizen in Polynome eingesetzt	195
5.1. Potenzen einer Matrix	195
5.2. Minimalpolynom	196
5.3. Satz von Cayley-Hamilton	197
Kapitel 7. Normale Matrizen	199
1. Orthogonale und unitäre Matrizen	199
1.1. Unitäre Matrizen	199
1.2. Orthogonale Diagonalisierung	201
2. Normale Matrizen	202
2.1. Definition und Beispiele	202
2.2. Orthogonale Diagonalisierbarkeit normaler Matrizen	202
3. Selbstadjungierte Matrizen und Definitheit	203
3.1. Diagonalisierung selbstadjungierter Matrizen	203
3.2. Definitheit	203
3.3. Sesquilinearformen und Matrizen	205

INHALTSVERZEICHNIS

0. Einleitung

Willkommen in der Linearen Algebra, einem mathematischen Kerngebiet, das seine Finger in fast alle Teilgebiete der Mathematik ausstreckt. Mit dieser Vorlesung möchte ich Ihnen Wissen und Verständnis für Vektoren und lineare Gleichungssysteme vermitteln. Vor allem möchte ich Ihnen aber auch ein Bild mitgeben, wie man Mathematik betreibt, wie eine mathematische Theorie gebaut ist, wie man sie beweist, und wie man sie zu Papier bringt.

0.1. Mathematik verstehen. Mathematik kann und muss auf vielen verschiedenen Ebenen verstanden werden, und daraus folgt auch, wie man sie lernen muss: Mit einem wachen und flexiblen Geist, der an den Text aus vielen verschiedenen Blickwinkeln herangeht. Die schlechte Nachricht zuerst: Und mit viel, viel Arbeit und Mühe. Um eine halbe Seite Mathematik zu verstehen, kann man manchmal einen Tag lang sitzen. Bitte lassen Sie sich nicht entmutigen! Gerade dann, wenn man den Widerstand eines Faches spürt, findet ein intensiver Lernprozess statt.

0.1.1. *Mathematik als Werkzeug.* Sie kennen Mathematik wahrscheinlich vor allem als ein Kompendium von Standardregeln, die man auf Standardprobleme anwendet, um eine richtige Lösung zu erhalten. Mathematik muss für den Benutzer, die Benutzerin ein zuverlässiges Werkzeug sein. Zu lernen gibt es verschiedene Methoden, zum Beispiel, wie man ein Gleichungssystem löst, wie man Matrizen miteinander multipliziert, wie man eine Determinante ausrechnet und was sie bedeutet. Diese Dinge trainiert man durch wiederholtes Üben, nicht zuletzt in den Proseminaren zur Vorlesung.

0.1.2. *Fachsprache.* Mathematische Fachliteratur, auch dieses Skriptum, hat ihre eigene, sphinxenartig steinerne Sprache: Texte sind in Definitionen, Sätze und Beweise aufgebrochen. Die Sprache ist monoton, es wiederholen sich immer wieder die Standardphrasen, an die Sie sich gewöhnen werden. Alles ist auf Unmissverständlichkeit ausgerichtet. Die kleinen Worte (dann, dann und nur dann, es gibt, es gibt genau ein, für alle, ...) sind sehr wichtig und dürfen nicht überlesen oder beliebig gegen andere ausgetauscht werden. Diese Sprache lernen Sie gleichzeitig mit dem Nachvollziehen der Beweise, wenn Sie mathematische Texte, zum Beispiel dieses Skriptum, genau durcharbeiten und sich immer wieder fragen: Wie würde ich denselben Text formulieren? Warum verwendet die Autorin genau diese umständliche Formulierung? Hätte ich wirklich mit meiner Formulierung dasselbe gesagt, was im Skriptum steht, oder besteht ein feiner Unterschied? Enthält meine Formulierung alles, was ein Leser, eine Leserin wissen muss, wenn er/sie nicht bei mir rückfragen können? Versuchen Sie gleich von Anfang an, alles was Sie an mathematischen Inhalten ausdrücken wollen, korrekt auszudrücken. Jedes Zeichen, das Sie setzen, jedes kleine Wort muss seine Bedeutung haben und bewusst gesetzt sein. Vermeiden Sie langatmigiges Herumreden. Schreiben Sie Ihre mathematischen Gedanken auf, dann werden Sie ganz von selbst unnötige Phrasen weglassen.

0.1.3. *Nachprüfbarkeit.* Wir wollen Sie in dieser Lehrveranstaltung davon überzeugen, dass Sie selbst nachprüfen können, ob unsere Theorien stimmen.¹ Bitte glauben Sie mir nichts außer den Definitionen. Wenn ich eine Behauptung aufstelle, schulde ich Ihnen einen Beweis. Das Wesentliche eines Beweises ist, dass er von Behauptungen ausgeht, die bereits auf sicherem Boden stehen, und in kleinen Schritten immer mehr Einsichten gewinnt, bis man sieht, dass die gewünschte Aussage unvermeidlich gelten muss. Wenn Sie Beweise verstehen wollen, brauchen Sie mehrere Durchgänge.

¹Schon deshalb, weil in diesem Skriptum immer wieder Tippfehler gefunden werden, und wenn ich einen korrigiert habe, wachsen zehn neue nach, wie die Köpfe der Hydra selig.

- (1) Zunächst fragen Sie sich: Was soll bewiesen werden? Kenne ich die Definitionen der Objekte, von denen die Rede ist? (Bei Bedarf nachschlagen.) Welche Voraussetzungen werden gemacht?
- (2) Dann ein sorgfältiges Durcharbeiten von vorne nach hinten. Fragen Sie sich bei jedem Schritt: Warum darf denn die Autorin das behaupten? Wissen wir das schon, und woher?
- (3) Wenn Sie sich überzeugt haben, dass der Beweis korrekt ist, sollten Sie ihn nach folgenden Gesichtspunkten wiederholt durchlesen: Was sind die entscheidenden Ideen, die Wendepunkte im Beweis, und was ergibt sich einfach automatisch durch Nachrechnen.
- (4) An welchen Stellen werden die Voraussetzungen gebraucht? Werden wirklich alle Voraussetzungen benötigt?

Lesen Sie die Definitionen sehr genau und lernen Sie sie auswendig, dass Sie nicht immer nachschlagen müssen, worüber wir reden. Merken Sie sich aber die Beweise nicht durch Auswendiglernen, sondern indem Sie sich die großen Linien des Gedankengangs, die grundlegenden Ideen, merken, und üben Sie, bei geschlossenem Skriptum den Beweis in eigenen Worten aufzufüllen und nachzuvollziehen. Nehmen Sie sich die Mühe und schreiben Sie Ihren Beweis exakt aus, oder erzählen Sie ihn einer/m KollegIn. In der Mathematik hat man nur das verstanden, was man auch exakt aussprechen kann. Fällt Ihnen ein anderer, kürzerer oder eleganterer Beweis ein? Hat die Verfasserin Fehler gemacht?² Arbeiten Sie bitte in kleinen Schritten, Beweis für Beweis. Es ist schwer, gerade am Anfang, Übersicht über ein ganzes Kapitel oder auch nur mehrere Seiten zu behalten.

0.1.4. *Theorie mit Höhepunkten.* Mathematik baut strukturierte Theorien. Es gibt zentrale Begriffe und Kernsätze, die Einsichten, um dererwillen die Theorie entwickelt wurde. Und es gibt viele Hilfssätze und Hilfsbegriffe, die helfen, sich zu den Beweisen der Kernsätze hochzuarbeiten. Wenn Sie einen Abschnitt durchgearbeitet haben, fragen Sie sich: Was sind die wesentlichen Begriffe und Aussagen des Kapitels, was ist die Pointe? Wie greifen die Aussagen des Kapitels ineinander? Wie baut das Kapitel auf vorigen Abschnitten auf?

0.1.5. *Abstraktion.* Mathematische Fachliteratur ist meist sehr abstrakt gehalten. Das heißt, es wird über die betrachteten Objekte so wenig wie möglich vorausgesetzt. Warum soll man für die reellen und komplexen Zahlen zweimal dasselbe beweisen, wenn man in Wirklichkeit nur ein paar Eigenschaften der Grundrechnungsarten braucht, die für beide Zahlenmengen gleich sind. Also reden wir von Körpern. Warum soll man eine Ebene oder den dreidimensionalen Raum getrennt betrachten, und gelten nicht auch dieselben Prinzipien im sechsdimensionalen Raum? Selbst Funktionen lassen sich als Vektoren auffassen, und auf diese Weise kann man Ideen der Geometrie auf Aufgaben übertragen, in denen die gesuchten Objekte Funktionen sind (zum Beispiel das Geschwindigkeitsfeld in einer strömenden Flüssigkeit). Auch wenn Sie zu Beginn eher den Eindruck haben werden, dass die Abstraktion eine Last ist, und sie Ihnen vielleicht sogar wie eine aufgeblasene akademische Seifenblase vorkommen mag: Wenn Sie im Studium fortschreiten, werden Sie immer mehr den Nutzen der Abstraktion schätzen lernen. Mit ihr kann man komplizierte Aufgaben behandeln, als wären sie einfach, weil sie einfach sind, wenn man sich auf das Wesentliche beschränkt und verwirrende Details weglassen kann. Für den Anfang ist wichtig, dass Sie sich daran gewöhnen, zwischen den Abstraktionsebenen hin und her zu schalten. Wenn Sie einen Satz über einen allgemeinen Vektorraum lesen, fragen Sie sich: Was bedeutet dieser Satz in einer Ebene oder im dreidimensionalen Raum? Versuchen Sie sich, den Sachverhalt durch eine Zeichnung klarer zu

²Das kommt leider vor. Bitte um ein e-Mail, wenn Sie einen Fehler finden!

machen. Verbirgt sich hinter dem mathematischen Formalismus vielleicht eine sehr einleuchtende Einsicht, die unsere Vorstellung von Raum und Ebene anspricht? Wenn Ihnen Ihre Intuition einen Sachverhalt nahelegt, der im dreidimensionalen Raum gilt, können Sie auch versuchen, ob Sie diesen Sachverhalt im allgemeinen Vektorraum beweisen können. Gilt er immer, oder brauchen Sie vielleicht Zusatzbedingungen, die nicht jeder Vektorraum erfüllt?

0.2. Der mathematische Lösungsprozess.

Ich bin manchmal gefragt worden, wie ich in einem so trockenen Beruf überleben kann. Die Antwort ist: Mathematik ist eine kreative Tätigkeit, weit entfernt von trocken. Mathematik ist für mich Wahrnehmen von Strukturen. Sehen ist keine maschinelle Tätigkeit, sondern ein Vorgang, der große Teile der Seele mit einbezieht, auch solche, die sich dem reinen Verstand entziehen.

Ich gehe jetzt bei meiner Beschreibung der schöpferischen Arbeit davon aus, dass Sie allein sind, und dass in der Ihnen zugänglichen Literatur die Lösung nicht zu finden ist, oder dass Sie darauf aus sind, das Beispiel allein zu lösen. Aber Sie wissen ja: Lesen und Diskussion mit KollegInnen machen Freude und sind unverzichtbare Quellen auf dem Weg der Wissenschaft.

Man darf sich den mathematischen Lösungsprozess nicht vorstellen, wie er leider oft in den Schulen eingepaukt wird, nämlich dass für eine oftmals geübte Musteraufgabe einfach das richtige Rechenverfahren eingeschaltet wird. Meist stehen wir ja vor Fragen, für die wir (zumindest selbst) noch kein Standardmuster zur Verfügung haben. Ich werde jetzt versuchen, den Lösungsprozess in Phasen aufzuteilen. Er läuft aber meistens nicht so linear ab. Immer wieder muss man von einer Phase in eine frühere zurückspringen. Einsicht ist ein Kreisprozess, eine Art Volleyballspiel zwischen den kreativen und den ordnenden Kräften im Kopf.

0.2.1. *Aufgabenstellung klären:* Oft ist dieser Abschnitt eine Routineaufgabe: Man überprüft, wie die Objekte definiert sind, und was die Aussage bedeutet, die man beweisen will. Manchmal hat man aber nur eine vage Vorstellung, dass ein bestimmter Sachverhalt gelten müsste, und muss selbst erst suchen, in welcher Form man ihn in einen mathematischen Satz gießen kann, wie man seine Objekte definieren muss, und welche Voraussetzungen passend sein könnten. In diesem Fall ist schon die Klärung der Aufgabenstellung ein Stück der Suchphase.

0.2.2. *Suchphase:* Das ist der eigentlich schöpferische Teil, für den es natürlich kein Patentrezept gibt. Wichtigste Spielregel ist, dass sich die Assoziationen frei entfalten dürfen. Alle Routinearbeiten werden für später aufgeschoben, jetzt sind nur die Aspekte interessant, die neue Ideen brauchen.

Man lässt alle Ideen an sich heran, und fragt sich dann: Hilft mir das? Hier sind einige Beispiele von Fragen, die man sich stellen kann, wenn man noch gar keine Vorstellung von einer möglichen Lösung hat:

- Was weiss ich über die Objekte, von denen die Rede ist? Welche Sätze kenne ich über sie?
- Käme ich mit zusätzlichen Voraussetzungen weiter, oder wenn ich in einer einfacheren Situation wäre, etwa in der Ebene statt im abstrakten Vektorraum?
- Kann ich eine Zeichnung machen? Was sagt mir meine Intuition?
- Kenne ich einen Satz, der ähnlich klingt wie der, den ich beweisen möchte? Wie wird er bewiesen? Wie unterscheidet er sich von meinem Problem?
- Kann ich mein Problem umformulieren, so dass es in eine mir bekannte Theorie fällt?

- Was muss noch gelten, wenn meine Behauptung wirklich stimmt? Was muss gelten, wenn sie nicht stimmt? Kann ich indirekt vorgehen?
- Stimmt meine Behauptung überhaupt? Gibt es ein Gegenbeispiel?
- Ich glaube, ich könnte die Aussage XXX beweisen. Würde mir die bei meinem Problem helfen?

Irgendwann fügen sich dann die vielen Gedanken zu einem Bild zusammen, in dem sich ein Weg abzeichnet. Versuchen Sie jetzt, diesen Weg zu gehen. Wo bleibt man hängen, welche Hürden müssen noch genommen werden? Am Ende schließt sich vielleicht ein zusammenhängender Weg von den Annahmen zur Aussage, oder vielleicht steht auch ein Gegenbeispiel da. Vielleicht kommt man auch zu keinem Ergebnis. Das ist in der Wissenschaft ganz normal. Wenn man alles beweisen könnte, was man will, wären Beweise ja sinnlos. Die vielen Irrwege sind ein wichtiger Teil der Erkenntnis.

Schmierpapier und Bleistifte sind die wertvollsten Ressourcen für schöpferische MathematikerInnen. Vielleicht müssen Sie sich, wie ich musste, erst daran gewöhnen, auf Blättern zu schreiben, auf denen schon viel durchgestrichen ist. Skizzieren Sie ihre Gedanken ruhig in allen Ecken und verbinden Sie sie mit Pfeilen quer über das Blatt. Radieren Sie nicht, streichen Sie einfach durch. Geniessen Sie es, auf einem Stück Papier ein Chaos zu inszenieren.

Wenn Sie sich festfahren und hängen bleiben, unterbrechen Sie und machen Sie eine Runde durch den Stadtpark oder sonst etwas Lockerndes, das Ihnen gut tut. In solchen Pausen tauchen oft schlagartig überraschende Einsichten auf. Schöpferische Arbeit ist ein Wechselspiel zwischen höchster Konzentration und Entspannung.

0.2.3. *Ordnungsphase*: Wenn die Lösung der Aufgabe gefunden ist, wird sie dokumentiert. Sie wird noch einmal systematisch ausformuliert. Ich habe (unpublizierte) Skizzen aus meiner mathematischen Sturm- und Drangzeit, die ich heute selbst nicht mehr verstehe, weil ich verabsäumt habe, sie nachvollziehbar hinzuschreiben. Oft merke ich auch beim sorgfältigen Hinschreiben, dass meine Gedanken noch Lücken haben, die zu schließen sind.

Um gut zu schreiben, müssen Sie nur einen Satz befolgen: Stellen Sie sich bitte immer, wenn Sie schreiben, einen Menschen vor, für den Sie schreiben.

Schreiben Sie alles, was diese Person zum Verständnis Ihrer Gedanken braucht. Verwirren Sie die LeserInnen nicht mit unnötigen Worten. Ordnen Sie den Text so, dass sich der Inhalt für die LeserInnen erschließt, ohne dass sie im Text ständig vor und rückwärts springen müssen.

Schreiben Sie einen Beweis möglichst so auf, dass Sie die Leser vom Bekannten zum Unbekannten führen. In der Suchphase geht es oft genau umgekehrt: "Ich möchte beweisen, dass XXX gilt, also muss ich zeigen, dass YYY gilt, also muss ich zeigen ...". Das ist zwar korrekt, aber anstrengend zu lesen, weil jede noch nicht bewiesene Aussage im Kopf des Lesers sozusagen mit einem roten Warnschild "Weiss ich aber noch nicht" gespeichert wird, und ein Gefühl von Unsicherheit verursacht. Kennzeichnen Sie deutlich Behauptungen, die noch zu zeigen sind, und orientieren Sie die LeserInnen über Ihre Strategie. (z.B., "Wir zeigen zunächst, dass die Behauptung XXX in endlichdimensionalen Vektorräumen gilt. Der unendlich dimensionale Fall wird dann daraus gefolgert." oder "Wir führen den Beweis indirekt. Sei also die Behauptung XXX falsch ...") Wenn Sie für Ihren Beweis Resultate brauchen, die schon bekannt sind, geben Sie Zitate, wo man diese Resultate mit Beweis nachschlagen kann. Ganz einfache Schritte können Sie auslassen, wenn Sie von der Leserschaft ausreichend Vorkenntnisse erwarten. (z.B., "Beweis als Übung.") Wenn Ihnen aber andere AutorInnen solche Stellen vorsetzen, überprüfen Sie bitte, ob der Beweis wirklich so leicht ist. Manchmal besteht die einfache Übung im Aufdecken eines fundamentalen Denkfehlers!

Am Ende der Ordnungsphase liegt das Ergebnis Ihrer Arbeit vor Ihnen, jederzeit nachvollziehbar für Sie und andere. In diesen schön aufgeschriebenen Resultaten sieht man oft nicht mehr die Mühe und auch nicht die Motivationen der Suchphase. Daher ist der Stil der mathematischen Literatur gut für die Dokumentation, als Lernende/r hätte man manchmal gern mehr Einblick auch in die Irrwege der schöpferischen Phase. Aber welche AutorInnen lassen sich schon gerne in die Karten gucken.

0.3. Ein paar allgemeine Tipps für die Arbeit.

0.3.1. *Offen sein.* Das ist vielleicht der beste und wichtigste Rat, den ich Ihnen mitgeben kann. Seien Sie offen und neugierig dem Studium gegenüber. Freuen Sie sich, dass Sie jetzt Mathematik betreiben können wie die “Großen”, und dass Sie in eine viel weitere, vielfältigere Gedankenwelt eingetreten sind, als Sie es sich wahrscheinlich erträumt hätten.

Sie müssen und sollen nichts wegwerfen, was Sie aus der Schule mitgebracht haben. Richtige Mathematik bleibt richtig. Aber klammern Sie sich auch nicht ängstlich an den Methoden fest, die sie in der Schule gelernt haben, probieren Sie aus, was wir hier anders machen, und machen Sie sich selbst ein Bild, wie Sie gerne arbeiten. Lassen Sie sich ruhig auf die ungewohnte Form der Mathematik ein! Sie sind gut in der Schulmathematik, sonst hätten Sie ja nicht dieses Fach gewählt. Sie werden auch gut in der Hochschulmathematik sein. Sie können nur dazugewinnen, und was Sie schon können, kann Ihnen niemand nehmen.

0.3.2. *Vorlesung.* Vorweg, der Besuch der Vorlesung ist nicht verpflichtend, und auch wenn es für mich angenehm ist, wenn ich pünktlich anfangen kann, wird Sie niemand fragen oder gar böse ansehen, wenn Sie einmal erst später dazukommen oder früher gehen müssen.

Die Vorlesung sollte eigentlich Vordenkung heißen. Ich möchte vor Ihnen den Stoff entstehen lassen, mit der richtigen Sprache, der richtigen Argumentation, und — unabsichtlich — den gelegentlichen Fehlern, die mit der Entwicklung einer Theorie einhergehen. Sie sollen Gelegenheit haben, simultan mitzudenken, und sofort zu fragen, wenn Ihnen ein Gedankengang unklar ist. Trauen Sie sich! Fragen sind willkommen und zeigen, dass Sie bei der Sache sind. Ich erkläre gern einen Sachverhalt ein zweites Mal mit anderen Worten, und wenn es sein muss, auch öfter.

Das Tempo der Vorlesung können Sie mitbestimmen. Bitte bremsen Sie mich ruhig, wenn ich für Sie zu schnell vorgehe, und treiben Sie mich an, wenn es langweilig wird. Und weisen Sie mich auch hin, wenn meine Stimme zu leise oder meine Handschrift, sagen wir wertfrei und politically correct, kreativ wird. In einer Vorlesung sind Sie im Hörsaal keinerlei Prüfungen ausgesetzt. Wenn ich Fragen in die Hörschaft richte, so nur, um Sie zum Mitdenken zu animieren, oder um ein Bild zu bekommen, wieviel Sie verstehen.

Niemand kann von Ihnen erwarten, über den Stoff in allen Details auf dem Laufenden zu sein, obwohl ein gewisses Maß an Mitlernen Ihnen später viel Arbeit spart und Sie von der Vorlesung mehr haben. Damit Sie verstehen können, was vorgetragen wird, sollten Sie die Definitionen geläufig haben, und eine Vorstellung haben, um welche Kerninhalte es zur Zeit geht. Die Beweise müssen Sie nicht von Stunde zu Stunde parat haben.

Es gibt zwei hervorragende Möglichkeiten, mit denen Sie für sich den Wert der Vorlesung zunichte machen können. Selbstverständlich verbietet Ihnen niemand, gelegentlich ein Wort mit Ihren NachbarInnen zu tauschen, solange das nicht permanent geschieht, und so leise, dass nicht andere ZuhörerInnen gestört werden. Aber immer wenn Sie Ihre Aufmerksamkeit von der Mathematik abziehen, geht Ihnen der Fluss der Argumentation verloren. Übrigens müssen in einer Mathematikvorlesung

die Tafeln sehr oft gelöscht werden, und die Pausen, die dabei entstehen, sind auch für Sie Entspannungs- und Gesprächspausen, und für mich eine Erholungsphase für die Stimme.

Die zweite Möglichkeit der Ressourcenverschwendung ist das kalligrafische Mitschreiben. Der ganze Stoff der Lehrveranstaltung liegt detailliert ausgearbeitet in diesem Skriptum vor Ihnen. Sie brauchen also als eigene Mitschrift nur einige wenige Notizen, die Sie vielleicht sogar in dieses Skriptum dazuschreiben können, wenn Sie es einseitig ausdrucken. Wenn Sie eine Mitschrift auf Schönschreiben, mit sauberen Unterstreichungen, eingerahmten Sätzen und mehreren Farben führen, verschenden Sie alle Konzentration und Zeit auf die äußere Form. Was Ihnen aber die Vorlesung über das Skriptum hinaus bieten kann und soll, ist die Erfahrung und die Übung im Mitdenken.

0.3.3. *Proseminar*. Die Proseminare dienen dazu, dass Sie selbst versuchen, eigenständig Aufgaben zu lösen, die mit dem Stoff der Vorlesung Hand in Hand gehen. Darunter finden sich Routineaufgaben, in denen die Anwendung der gelernten Rechenverfahren geübt wird, und leichte Beweisaufgaben, damit Sie den Umgang mit den Begriffen und das mathematische Argumentieren üben. Immer wieder dazwischen kommt auch einmal eine schwierigere Aufgabe. Wenn Sie zu denen gehören, denen das Fach leicht fällt, werden Sie Spaß daran haben, wenn man Sie ein wenig herausfordert.

Es besteht immanenter Prüfungscharakter, das heißt, Sie müssen bei der Lehrveranstaltung anwesend sein, denn die Leistungen, die Sie im Hörsaal erbringen, tragen zumindest teilweise zur Gesamtnote bei. Es ist durchaus nicht unerwünscht, wenn Sie die Hausübungsbeispiele gemeinsam mit KollegInnen vorbereiten, aber diskutieren Sie sie in diesem Fall so weit aus, bis alle TeilnehmerInnen das Beispiel auch verstanden haben. Wenn Sie an die Tafel kommen und etwas vortragen, das Sie gar nicht verstehen, findet das die/der Lehrende schnell heraus, und Sie müssen mit Punkteabzug rechnen. Sonst gilt: Natürlich haben Sie das Recht auf Fehler. Es kommt einfach vor, dass man ein Beispiel fehlerhaft vorbereitet, leider geht mir das auch in der Vorlesung manchmal so. Wenn der Übungsleiter, die Leiterin, in Ihrer Lösung einen Fehler aufdeckt, wird er/sie vielleicht versuchen, Sie auf den richtigen Gedanken hinzulenken. Denken Sie mit und lassen Sie sich ruhig führen. Sie lernen dabei gleich die unschätzbare Fähigkeit, ein bisschen zu improvisieren.

Ich möchte Ihnen empfehlen, Proseminar und Vorlesung möglichst zugleich zu besuchen. In allen Fragen, welche die Proseminare betreffen, wenden Sie sich bitte an den Leiter / die Leiterin der Übungsgruppe, der Sie zugeteilt sind. Ich selbst bin für das Proseminar nicht zuständig, und habe keinen Einfluss auf die ÜbungsleiterInnen. Sie arbeiten autonom und eigenverantwortlich.

0.3.4. *Tutorium*. Es wird auch ein Tutorium angeboten. Dieses ist formlos und wird, damit für Sie die Schwellenangst möglichst gering ist, von Studierenden gehalten. Es soll Ihnen helfen, Schwierigkeiten zu überbrücken, die in Vorlesung oder Proseminar für Sie aufgetreten sind. Sie können mit Ihren Fragen zu Vorlesung und Proseminar in das Tutorium kommen, oder auch einfach zuhören, was sonst im Tutorium geschieht.

0.3.5. *Fachbücher*. Diese Vorlesung folgt weitgehend dem Taschenbuch Lineare Algebra von Gerd Fischer, Springer-Vieweg Verlag 2014. Es ist ein gut ausgereiftes und sorgfältig geschriebenes Lehrbuch, das inzwischen seine achtzehnte Auflage erlebt. Ich empfehle Ihnen auch, in der Bibliothek des Institutes für Mathematik an der KFU am Standort für Lineare Algebra zu schmökern. Vieles wird jetzt noch weit über Ihrem Horizont liegen, aber es gibt auch sehr gut erklärte Anfängerlehrbücher. Mir selbst gefallen zum Beispiel die Bücher von Klaus Jänich und von Gilbert Strang. Blättern Sie auch ruhig in englischsprachiger Fachliteratur. Die

Schwierigkeit der Fremdsprache wird oft wett gemacht, indem englisch schreibende VerfasserInnen typischerweise mehr Wert auf Verständlichkeit und gute Erklärung legen.

0.3.6. *Fragen kommen.* Wenn Sie mit dem Verständnis anstehen, oder auch allgemeine Fragen oder Wünsche zur Vorlesung haben, scheuen Sie sich nicht, fangen Sie mich nach der Vorlesung ab, oder klopfen Sie bei mir an. Das gilt natürlich auch für Fragen, auf die Sie später beim Lernen auf die Prüfung stossen. Sie sind willkommen, ich gebe Ihnen gerne Auskunft. Das Schlimmste, was passieren kann, ist dass ich gerade nicht da bin oder keine Zeit habe. Wenn Sie mich ganz sicher erreichen wollen, machen Sie am besten einen Termin über e-Mail aus.

0.3.7. *Gemeinsam lernen.* Es ist ein guter Tipp, sich mit KollegInnen zusammen zu tun und gemeinsam zu lernen. Erstens kann man sich erzählen, was in den Lehrveranstaltungen geschehen ist, die jemand versäumt hat. Es ist auch gut, beim Lernen und Verstehen des Stoffes miteinander zu diskutieren. Vier oder sechs Augen finden schneller den springenden Punkt als zwei. Wenn Sie sich gegenseitig beim Lernen auf die Prüfung unterstützen, lassen Sie sich von der Kollegin oder dem Kollegen den Stoff genau und unmissverständlich erzählen, damit Unklarheiten aufgedeckt werden können, und auch die Fähigkeit zum exakten sprachlichen Ausdruck trainiert wird. Geben Sie bitte als Zuhörende/r nicht Ruhe, bis alles sauber und unmissverständlich ausformuliert ist. "Ich weiss eh schon, was Du meinst" hilft dem Lernpartner nicht.

0.3.8. *Prüfungen und Noten.* Ich werde die Vorlesungen Lineare Algebra 1 und 2 schriftlich prüfen, um mit der TU einheitlich vorzugehen und auch wegen der großen Studierendenzahlen. Es wird für jede der beiden Vorlesungen 6 Termine geben, der erste davon gleich nach Ende der Lehrveranstaltungen, die anderen verstreut über ein Jahr. Mustertests, an denen Sie sehen können, wie meine Prüfungen aufgebaut sind, werden rechtzeitig vor den ersten Prüfungsterminen auf meiner Homepage angeboten werden. Bitte unterschätzen Sie nicht die Schwierigkeit der Prüfungen an der Hochschule. Normalerweise reichen nicht bloss wenige Tage aus, sondern man braucht mehrere Wochen, um den umfangreichen und schwierigen Stoff zu bewältigen.

Gar nicht wenige von Ihnen sind anfangs ziemlich enttäuscht, wenn Sie auf der Höheren Schule stets gewohnt waren, in Mathematik zu glänzen, und nun auf der Universität vielleicht nur im Mittelfeld abschneiden, oder sich überhaupt sehr schwer tun. Lassen Sie sich nicht entmutigen! Waren Sie in Ihrer Klasse eine/r der wenigen, für die ernsthaft ein Mathematikstudium in Frage gekommen ist, so finden Sie sich jetzt unter lauter KollegInnen, die auch Mathematik gewählt haben. Das ist ein ganz anderer Konkurrenzdruck. Es ist absolut ok, nicht der/die Beste zu sein.

Die ersten Schritte des Mathematikstudiums sind oft von einer Art Kulturschock begleitet. Wenn man dann zu begreifen beginnt, wo der Kern des mathematischen Denkens liegt, öffnet sich das spröde Fach. Ich wünsche Ihnen, dass Sie nicht nur immer tiefer in das Verständnis der Mathematik vordringen, sondern dabei auch etwas von der Begeisterung und Faszination erleben, mit der Mathematik für mich und viele verbunden ist. Viel Erfolg und Freude im Studium!

Ein herzliches Dankeschön: Viele, viele Studierende haben die vorigen Versionen des Skriptums sorgfältig gelesen und mir Listen von Verbesserungsvorschlägen und Tippfehlern zukommen lassen. Ihnen allen möchte ich an dieser Stelle meinen herzlichen Dank aussprechen!

KAPITEL 1

Matrizenrechnung

In unserem ersten Kapitel befassen wir uns mit der linearen Algebra über den reellen Zahlen und in endlich dimensionalen Vektorräumen. Obwohl vieles in diesem Kapitel in viel allgemeineren Situationen anwendbar ist, wollen wir zu Beginn nicht allzu viele neue Begriffe einführen, damit Sie sich auf das Wesentliche konzentrieren können, und nicht unter einem Berg von Abstraktion den anschaulichen Zugang verlieren.

1. Notation

1.1. Zahlen, Vektoren, Matrizen.

Wir fixieren zunächst einige Schreibweisen, die wir im Laufe der Vorlesung immer wieder brauchen werden:

VEREINBARUNG 1.1.1.

- (1) Die Menge der natürlichen Zahlen ist $\mathbb{N} = \{1, 2, 3, \dots\}$.¹
- (2) Die Menge der ganzen Zahlen ist $\mathbb{Z} = \{0, 1, -1, 2, -2, \dots\}$.
- (3) Die reellen Zahlen sind alle Zahlen der Zahlengeraden, also 0 sowie alle positiven und alle negativen Zahlen, die eine Länge bezeichnen können. Die Menge der reellen Zahlen schreiben wir \mathbb{R} .

Beachten Sie bitte, dass die obige Vereinbarung viel zu vage für eine mathematische Definition wäre. Wir gehen davon aus, dass Sie mit den Zahlen und den Grundrechnungsarten vertraut sind. Natürlich gibt es einen mathematisch exakten Aufbau des Zahlensystems. Wenn man für die natürlichen Zahlen einige grundlegende Axiome voraussetzt, kann man die anderen Zahlenmengen, die Grundrechnungsarten darin und ihre Eigenschaften streng formal konstruieren. Speziell die Konstruktion der reellen Zahlen, wenn man einmal die Bruchzahlen hat, ist ziemlich genial und trickreich. Das würde uns aber hier zu weit von unserem Thema abbringen.

Matrizen und Vektoren werden uns viel Schreibarbeit ersparen und helfen, die lineare Algebra übersichtlich zu gestalten. Vorläufig führen wir sie einfach als Schreibweise ein. In Wirklichkeit sind sie mathematische Objekte, mit denen man rechnen kann, fast wie mit Zahlen. Sie bilden das Rückgrat der linearen Algebra.

DEFINITION 1.1.2. Ein n -dimensionaler Spaltenvektor über \mathbb{R} besteht aus n reellen Zahlen, welche in einer Spalte untereinander angeordnet sind:

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}.$$

Die Menge der n -dimensionalen Spaltenvektoren über \mathbb{R} bezeichnen wir mit \mathbb{R}^n .

¹Achtung, manche Autoren zählen auch 0 zu den natürlichen Zahlen.

DEFINITION 1.1.3. Eine $(m \times n)$ -Matrix über \mathbb{R} ist ein rechteckiges Schema, in dem in m Zeilen und n Spalten reelle Zahlen angeordnet sind:

$$(\alpha_{i,j})_{i=1 \dots m, j=1 \dots n} = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix}.$$

Die Menge aller $(m \times n)$ -Matrizen über \mathbb{R} bezeichnen wir mit $\mathbb{R}^{m \times n}$. Im Gegensatz zu einer Matrix oder einem Spaltenvektor bezeichnen wir eine einzelne reelle Zahl auch als einen Skalar.

Wir werden zur Bezeichnung von Skalaren griechische Kleinbuchstaben verwenden. Bitte machen Sie sich im Lauf der Vorlesung mit dem griechischen Alphabet vertraut:

alpha	A	α	ny	N	ν
beta	B	β	xi	Ξ	ξ
gamma	Γ	γ	omikron	O	o
delta	Δ	δ	pi	Π	π
epsilon	E	ϵ, ε	rho	P	ρ
zeta	Z	ζ	sigma	Σ	σ
eta	H	η	tau	T	τ
theta	Θ	θ, ϑ	ypsilon	Y	υ
iota	I	ι	phi	Φ	ϕ, φ
kappa	K	κ	chi	X	χ
lambda	Λ	λ	psi	Ψ	ψ
my	M	μ	omega	Ω	ω

1.2. Mengen.

Will man die Mengenlehre axiomatisch auf ein solides Fundament stellen, muss man sehr diffizile logische Probleme lösen. Mit unserer naiven Vorstellung: "Eine Menge ist eine Zusammenfassung irgendwelcher Element zu einem Ganzen" laufen wir sehr schnell in Widersprüche:

BEISPIEL 1.2.1 (Russelsche Antinomie). Sei M die Menge aller Mengen, die sich selbst nicht als Element enthalten.

Die harmlos scheinende Definition verfängt sich in Widersprüche, wenn wir entscheiden sollen, ob M selbst ein Element von M ist:

Sei $M \in M$. Die Menge M enthält nur Elemente, die sich selbst nicht als Element enthalten. Weil aber $M \in M$ vorausgesetzt wurde, kann also M kein Element von M sein.

Sei $M \notin M$. Nach Definition von M muss also M zu den Elementen von M gehören.

Wie man es dreht, es kommt immer ein Widerspruch heraus. Die Definition von M liefert keine sinnvolle Menge!

Trotz dieser Gefahren können wir ruhig für diese Vorlesung und für weite Bereiche der Mathematik mit unserer naiven Vorstellung fortfahren. Wir wenden keine so starken Mittel der Mengenlehre an, wir definieren keine so komplizierten Mengen, dass wir auf die Feinheiten der Axiomatik achten müssen. Dieses Kapitel soll nur die Schreibweisen festlegen, mit denen wir arbeiten. Wir gehen auf die Problematik der Mengentheorie nicht ein.

VEREINBARUNG 1.2.2. Seien a, b, c, \dots irgendwelche Objekte, und sei $A(x)$ eine Aussage, die man über ein Objekt x machen kann, und die dann wahr oder falsch sein kann. Wir führen folgende Mengenschreibweisen ein:

- (1) $\{a, b, c, \dots\}$ ist die Menge, die die Objekte a, b, c usw. enthält (aufzählende Mengendeklaration).
- (2) $\{x \mid A(x)\}$ ist die Menge aller Objekte x , für die die Aussage $A(x)$ wahr ist (beschreibende Mengendeklaration).

Die in einer Menge M enthaltenen Objekte nennen wir die Elemente von M . Ist x ein Element von M , schreiben wir $x \in M$. Die Menge, die gar kein Element enthält, heißt die leere Menge \emptyset .

Hier haben wir uns nicht getraut, von einer Definition zu reden. Mathematisch ist das alles sehr verschwommen. Was sind "irgendwelche Objekte"?

VEREINBARUNG 1.2.3. Seien A und B zwei Mengen. Wir sagen, A ist eine Teilmenge von B ($A \subseteq B$, B ist eine Obermenge von A , $B \supseteq A$), wenn jedes Element aus A auch in B enthalten ist.

A ist eine echte Teilmenge von B , wenn A eine Teilmenge von B ist und $A \neq B$ gilt. Wir schreiben auch $A \subsetneq B$.

Achtung: Die Notation $A \subset B$ ist mehrdeutig. Manche Autoren verwenden sie für alle Teilmengen, manche nur für echte Teilmengen.

VEREINBARUNG 1.2.4. Seien A, B, A_1, A_2, \dots Mengen. Wir verwenden folgende Bezeichnungen:

- $A \cap B$ Durchschnitt von A und B : Die Menge aller Elemente, die sowohl in A als auch in B liegen.
- $A \cup B$ Vereinigung von A und B : Die Menge aller Elemente, die in A oder B oder beiden liegen.
- $A \setminus B$ Differenzmenge. Die Menge aller Elemente von A , die nicht in B liegen.
- $\bigcap_{i=1}^{\infty} A_i$ Durchschnitt aller A_i . Die Menge aller Elemente, die in jedem A_i enthalten sind.
- $\bigcup_{i=1}^{\infty} A_i$ Vereinigung aller A_i . Die Menge aller Elemente, die mindestens in einem A_i enthalten sind.
- $A \times B$ Produktmenge. Menge aller geordneten Paare (a, b) mit $a \in A$ und $b \in B$.

2. Lineare Gleichungssysteme

2.1. Grundbegriffe.

BEISPIEL 2.1.1. Wir geben hier ein Beispiel für ein lineares Gleichungssystem von 3 Gleichungen in 3 Unbekannten:

$$\begin{array}{rcccc} 10\lambda & + & 12\mu & + & 7\nu & = & 35 \\ \lambda & + & 2\mu & + & \nu & = & 5 \\ 11\lambda & + & 30\mu & + & 18\nu & = & 58 \end{array}$$

Gesucht sind Zahlen λ, μ und ν , die gleichzeitig alle drei Gleichungen erfüllen².

Lösung:

²Lies: λ : lambda, μ : my, ν : ny.

- (1) Unser erstes Ziel ist, eine Unbekannte aus allen Gleichungen außer einer einzigen zu entfernen. Wir wählen dazu die Unbekannte ν und beschließen, dass sie in der zweiten Gleichung stehen bleiben darf. Wenn wir die zweite Gleichung sieben mal von der ersten Gleichung abziehen, und 18 mal von der dritten Gleichung abziehen, entstehen zwei neue Gleichungen, die die Unbekannte ν nicht mehr enthalten. Wenn die ursprünglichen 3 Gleichungen gelten, müssen auch die neuen Gleichungen gelten, denn wir haben ja jeweils links und rechts die selben Werte abgezogen. Wir lassen die alte erste und dritte Gleichung weg, und ersetzen sie durch die beiden neuen Gleichungen. Die zweite Gleichung behalten wir so, wie sie war. Wir dürfen sie ab jetzt zu keiner anderen Gleichung mehr addieren, sonst würden wir die Unbekannte ν wieder in die anderen Gleichungen hineinbringen.

$$\begin{array}{rclcl} 3\lambda & - & 2\mu & & = & 0 \\ \lambda & + & 2\mu & + & \nu & = & 5 \\ -7\lambda & - & 6\mu & & & = & -32 \end{array}$$

- (2) Wir haben noch die erste und dritte Gleichung zur Verfügung. Wieder wollen wir eine Unbekannte aus allen außer einer Gleichung ausschalten. Wir entscheiden, dass die Unbekannte μ eliminiert wird, und dass sie in der ersten Zeile stehen bleibt. Wir ziehen daher drei mal die erste Gleichung von der dritten Gleichung ab.

$$\begin{array}{rclcl} 3\lambda & - & 2\mu & & = & 0 \\ \lambda & + & 2\mu & + & \nu & = & 5 \\ -16\lambda & & & & & = & -32 \end{array}$$

Wenn wir die dritte Gleichung links und rechts durch -16 dividieren, bleibt sie gültig und wird übersichtlicher:

$$\begin{array}{rclcl} 3\lambda & - & 2\mu & & = & 0 \\ \lambda & + & 2\mu & + & \nu & = & 5 \\ \lambda & & & & & = & 2 \end{array}$$

- (3) Zwischenbilanz: Wir stehen am Ende der sogenannten *Eliminationsphase*: Wir haben erreicht, dass die dritte Gleichung nur mehr die eine Unbekannte λ enthält. Aus dieser Gleichung können wir λ direkt ablesen. Die erste Gleichung enthält zwei Unbekannte, wenn wir aber λ einmal kennen, können wir leicht μ berechnen. Die zweite Gleichung enthält alle Unbekannten, aber wenn wir λ und μ schon kennen, können wir nun ν berechnen. Wir rollen also das System von hinten nach vorne wieder auf, und werden die überflüssigen Unbekannten aus der ersten und zweiten Gleichung eliminieren. Diese Phase des Lösungsverfahrens nennt man *Rücksubstitutionsphase*.
- (4) Wir subtrahieren die dritte Gleichung 3 mal von der ersten und einmal von der zweiten. Nun ist die dritte Gleichung die einzige Gleichung im System, die noch λ enthält.

$$\begin{array}{rclcl} -2\mu & & & & = & -6 \\ 2\mu & + & \nu & & = & 3 \\ \lambda & & & & = & 2 \end{array}$$

Die erste Gleichung können wir durch -2 dividieren:

$$\begin{array}{rclcl} \mu & & & & = & 3 \\ 2\mu & + & \nu & & = & 3 \\ \lambda & & & & = & 2 \end{array}$$

- (5) Jetzt ziehen wir noch die erste Gleichung zweimal von der zweiten ab und haben die Unbekannten vor uns stehen:

$$\begin{array}{rclcl} \mu & & & & = & 3 \\ & & \nu & & = & -3 \\ \lambda & & & & = & 2 \end{array}$$

- (6) Wir können die Probe machen, indem wir die errechneten Unbekannten λ , μ und ν in die ursprünglichen Gleichungen einsetzen:

$$\begin{array}{rclcl} 10 \cdot 2 & + & 12 \cdot 3 & + & 7 \cdot (-3) & = & 35 \\ 2 & + & 2 \cdot 3 & + & \cdot (-3) & = & 5 \\ 11 \cdot 2 & + & 30 \cdot 3 & + & 18 \cdot (-3) & = & 58 \end{array}$$

Die drei Zahlen $\lambda = 2$, $\mu = 3$ und $\nu = -3$ erfüllen tatsächlich alle drei Gleichungen.

DEFINITION 2.1.2.

- (1) Ein lineares Gleichungssystem von m Gleichungen in n Unbekannten hat die Gestalt

$$(2.1.1) \quad \begin{array}{cccccc} \alpha_{1,1}\tau_1 & + & \alpha_{1,2}\tau_2 & + & \cdots & + & \alpha_{1,n}\tau_n & = & \beta_1 \\ \alpha_{2,1}\tau_1 & + & \alpha_{2,2}\tau_2 & + & \cdots & + & \alpha_{2,n}\tau_n & = & \beta_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ \alpha_{m,1}\tau_1 & + & \alpha_{m,2}\tau_2 & + & \cdots & + & \alpha_{m,n}\tau_n & = & \beta_m \end{array}$$

Dabei sind für $i = 1, \dots, m$ und $j = 1, \dots, n$ die $\alpha_{i,j}$ und β_i gegebene reelle Zahlen³. Gesucht ist ein n -Tupel (τ_1, \dots, τ_n) von reellen Zahlen, die zugleich alle m Gleichungen erfüllen.

- (2) Eine partikuläre Lösung des Systems (2.1.1) ist ein Spaltenvektor

$$\begin{pmatrix} \tau_1 \\ \vdots \\ \tau_n \end{pmatrix},$$

dessen Komponenten alle Gleichungen des Systems (2.1.1) gleichzeitig erfüllen.

- (3) Die allgemeine Lösung oder Lösungsmenge des Systems (2.1.1) ist die Menge aller partikulären Lösungen.

Ein lineares Gleichungssystem erkennt man also daran, dass die Unbekannten nur mit konstanten reellen Zahlen multipliziert werden und addiert werden dürfen. Es gibt also zum Beispiel keine Quadrate oder Wurzeln und keine Produkte von zwei Unbekannten miteinander.

BEMERKUNG 2.1.3. System (2.1.1) lässt sich kurz auch so anschreiben:

$$(\forall i = 1, \dots, m) \sum_{j=1}^n \alpha_{i,j}\tau_j = \beta_i.$$

DEFINITION 2.1.4. Die erweiterte Matrix des Gleichungssystems (2.1.1) ist die $(m \times (n + 1))$ -Matrix

$$\begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} & \beta_1 \\ \vdots & & \vdots & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} & \beta_m \end{pmatrix}.$$

BEMERKUNG 2.1.5. Die erweiterte Matrix eines Systems von m linearen Gleichungen in n Unbekannten hat m Zeilen und $(n + 1)$ Spalten. Jede Zeile beschreibt eine der Gleichungen, jede der ersten n Spalten gehört zu einer Variablen, nur die letzte Spalte gehört zu den Konstanten auf der rechten Seite der Gleichungen.

Es ist wichtig, die Koeffizienten zu den Variablen sorgfältig untereinander zu schreiben. Erscheint eine Variable in einer der Gleichungen nicht, muss an ihrem Platz eine Null geschrieben werden.

³Lies: α : alpha, β : beta, τ : tau. — Später werden wir sehen, dass es genausogut Elemente eines beliebigen Körpers sein können.

2.2. Das Lösungsprinzip.

DEFINITION 2.2.1. Sei A eine Matrix über \mathbb{R} . Die folgenden drei Umformungen heißen elementare Zeilentransformationen:

- (1) Zur Zeile Nummer i wird das γ -fache⁴ einer anderen Zeile (Nummer j) addiert. Dabei ist $\gamma \in \mathbb{R}$.
- (2) Die Zeile Nummer i wird mit einer reellen Zahl $\gamma \neq 0$ multipliziert.
- (3) Die Zeilen Nummer i und j tauschen Platz.

BEHAUPTUNG 2.2.2. Jede elementare Zeilentransformation lässt sich durch eine elementare Zeilentransformation rückgängig machen, und zwar

- (1) Addiere das γ -fache von Zeile j zu Zeile i — addiere das $(-\gamma)$ -fache von Zeile j zu Zeile i .
- (2) Multipliziere Zeile i mit $\gamma \neq 0$ — Multipliziere Zeile i mit $1/\gamma$.
- (3) Tausche Zeilen i und j — tausche nochmal Zeilen i und j .

SATZ 2.2.3. Sei

$$C = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} & \beta_1 \\ \vdots & & \vdots & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} & \beta_m \end{pmatrix}$$

die erweiterte Matrix eines linearen Gleichungssystems von m Gleichungen in n Unbekannten. Die Matrix \tilde{C} entstehe aus C durch eine elementare Zeilentransformation. Dann hat das Gleichungssystem mit der erweiterten Matrix \tilde{C} dieselbe Lösungsmenge wie das Gleichungssystem mit der erweiterten Matrix C .

BEWEIS. Wir zeigen zunächst für jede der drei Arten von elementaren Zeilentransformationen, dass jede partikuläre Lösung zur erweiterten Matrix C auch Lösung zur erweiterten Matrix \tilde{C} ist. Es sei also $\begin{pmatrix} \tau_1 \\ \vdots \\ \tau_n \end{pmatrix}$ eine partikuläre Lösung zur erweiterten Matrix C .

- (1) Wurde das γ -fache der Zeile j zu Zeile i addiert, so heißt die Gleichung zur neuen Zeile i :

$$(\alpha_{i,1} + \gamma\alpha_{j,1})\tau_1 + \cdots + (\alpha_{i,n} + \gamma\alpha_{j,n})\tau_n = (\beta_i + \gamma\beta_j).$$

Da die τ eine Lösung des ursprünglichen Systems ist, gilt aber

$$\begin{aligned} & (\alpha_{i,1} + \gamma\alpha_{j,1})\tau_1 + \cdots + (\alpha_{i,n} + \gamma\alpha_{j,n})\tau_n \\ &= [\alpha_{i,1}\tau_1 + \cdots + \alpha_{i,n}\tau_n] + [\gamma\alpha_{j,1}\tau_1 + \cdots + \gamma\alpha_{j,n}\tau_n] \\ &= [\alpha_{i,1}\tau_1 + \cdots + \alpha_{i,n}\tau_n] + \gamma[\alpha_{j,1}\tau_1 + \cdots + \alpha_{j,n}\tau_n] \\ &= \beta_i + \gamma\beta_j. \end{aligned}$$

Daher gilt auch die neue Gleichung Nummer i . Alle anderen Gleichungen gelten, da sie unverändert aus dem ursprünglichen System übernommen wurden.

⁴Lies: γ : gamma.

Das vorletzte Pivotelement war die (-2) an der Stelle (1,2). Wir räumen die zweite Spalte auf.

$$\begin{pmatrix} 0 & -2 & 0 & -6 \\ 0 & 0 & 1 & -3 \\ 1 & 0 & 0 & 2 \end{pmatrix}.$$

Wir kürzen die erste Zeile durch (-2):

$$\begin{pmatrix} 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -3 \\ 1 & 0 & 0 & 2 \end{pmatrix}.$$

Die Zeilen der obigen Matrix bedeuten

$$\mu = 2, \nu = -3, \lambda = 2.$$

Damit ist das Gleichungssystem gelöst. \square

BEISPIEL 2.2.5.

$$\begin{aligned} 2x + y + 3z + u &= 7 \\ 6x - 2y - 11z + 3u &= 1 \\ -6x + 6z - 2u &= 4 \\ 2x + 3y + 17z + 4u &= 45 \end{aligned}$$

Lösung: Zunächst stellen wir die erweiterte Matrix auf. Beachten Sie, dass in der dritten Gleichung y nicht vorkommt. Wir müssen aber an der Stelle für y in der Matrix eine Null schreiben, sodass alle Zeilen gleich lang sind.

$$\begin{pmatrix} 2 & 1 & 3 & 1 & 7 \\ 6 & -2 & -11 & 3 & 1 \\ -6 & 0 & 6 & -2 & 4 \\ 2 & 3 & 17 & 4 & 45 \end{pmatrix}$$

Wir formen nun die erweiterte Matrix mit Eliminationsschritten um. Pivotelemente kennzeichnen wir jeweils mit Stern *, die Nummer über dem Stern sagt, das wievielte Pivotelement hier steht.

$$\begin{aligned} & \begin{pmatrix} 2 & 1^{*1} & 3 & 1 & 7 \\ 6 & -2 & -11 & 3 & 1 \\ -6 & 0 & 6 & -2 & 4 \\ 2 & 3 & 17 & 4 & 45 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 1^{*1} & 3 & 1 & 7 \\ 10 & 0 & -5 & 5^{*2} & 15 \\ -6 & 0 & 6 & -2 & 4 \\ -4 & 0 & 8 & 1 & 24 \end{pmatrix} \\ \rightarrow & \begin{pmatrix} 2 & 1^{*1} & 3 & 1 & 7 \\ 2 & 0 & -1 & 1^{*2} & 3 \\ -2^{*3} & 0 & 4 & 0 & 10 \\ -6 & 0 & 9 & 0 & 21 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 1^{*1} & 3 & 1 & 7 \\ 2 & 0 & -1 & 1^{*2} & 3 \\ 1^{*3} & 0 & -2 & 0 & -5 \\ 0 & 0 & -3^{*4} & 0 & -9 \end{pmatrix} \\ \rightarrow & \begin{pmatrix} 2 & 1^{*1} & 3 & 1 & 7 \\ 2 & 0 & -1 & 1^{*2} & 3 \\ 1^{*3} & 0 & -2 & 0 & -5 \\ 0 & 0 & 1^{*4} & 0 & 3 \end{pmatrix} \end{aligned}$$

Für das letzte Pivotelement bestand der Pivotschritt nur in der Division der Pivotzeile durch das Pivotelement.

Mit dem vierten Pivotelement leiten wir nun die Rücksubstitutionsphase ein:

$$\begin{aligned} & \begin{pmatrix} 2 & 1^{*1} & 3 & 1 & 7 \\ 2 & 0 & -1 & 1^{*2} & 3 \\ 1^{*3} & 0 & -2 & 0 & -5 \\ 0 & 0 & 1^{*4} & 0 & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 1^{*1} & 0 & 1 & -2 \\ 2 & 0 & 0 & 1^{*2} & 6 \\ 1^{*3} & 0 & 0 & 0 & 1 \\ 0 & 0 & 1^{*4} & 0 & 3 \end{pmatrix} \\ \rightarrow & \begin{pmatrix} 0 & 1^{*1} & 0 & 1 & -4 \\ 0 & 0 & 0 & 1^{*2} & 4 \\ 1^{*3} & 0 & 0 & 0 & 1 \\ 0 & 0 & 1^{*4} & 0 & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 1^{*1} & 0 & 1 & -8 \\ 0 & 0 & 0 & 1^{*2} & 4 \\ 1^{*3} & 0 & 0 & 0 & 1 \\ 0 & 0 & 1^{*4} & 0 & 3 \end{pmatrix} \end{aligned}$$

Nun können wir die Lösung direkt ablesen:

$$x = 1, y = -8, z = 3, u = 4.$$

\square

2.3. Systeme mit keiner oder mehrfachen Lösungen.

BEISPIEL 2.3.1. Lösen Sie das System mit der erweiterten Matrix

$$\begin{pmatrix} 2 & 1 & 2 & 3 & 1 \\ 5 & 2 & -2 & 1 & 0 \\ 7 & 3 & 0 & 4 & 1 \\ 8 & 3 & -6 & -1 & 2 \end{pmatrix}.$$

Lösung:

$$\begin{pmatrix} 2 & 1 & 2^* & 3 & 1 \\ 5 & 2 & -2 & 1 & 0 \\ 7 & 3 & 0 & 4 & 1 \\ 8 & 3 & -6 & -1 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 1 & 2 & 3 & 1 \\ 7^* & 3 & 0 & 4 & 1 \\ 7 & 3 & 0 & 4 & 1 \\ 14 & 6 & 0 & 8 & 5 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 1 & 2 & 3 & 1 \\ 7 & 3 & 0 & 4 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

Nun können wir kein Pivotelement mehr finden, obwohl es Zeilen gibt, die noch nicht Pivotzeilen waren. Diese Zeilen enthalten aber in den ersten 4 Spalten nur Nullen, und ein Pivotelement muss ungleich Null sein.

Die dritte Zeile besteht nur aus Nullen und bedeutet

$$0\tau_1 + 0\tau_2 + 0\tau_3 + 0\tau_4 = 0.$$

Das gilt für jedes beliebige Quadrupel von τ_i . Daher ist jeder Vektor $\begin{pmatrix} \tau_1 \\ \vdots \\ \tau_4 \end{pmatrix}$, der die erste, zweite

und vierte Gleichung löst, automatisch eine Lösung der dritten Gleichung und löst damit das ganze System. Die dritte Zeile kann daher aus der Systemmatrix gestrichen werden.

Die vierte Zeile bedeutet aber

$$0\tau_1 + 0\tau_2 + 0\tau_3 + 0\tau_4 = 3,$$

und das kann niemals gelten. Daher gibt es keine τ_i , die die vierte Gleichung lösen, und daher gibt es erst recht keine τ_i , die das gesamte System lösen. Die Lösungsmenge des Systems ist leer. \square

BEHAUPTUNG 2.3.2. *An Hand des obigen Beispiels haben wir gesehen:*

- (1) *Enthält die erweiterte Matrix eines Gleichungssystems eine Zeile, die nur aus Nullen besteht, so kann diese Zeile aus dem System gestrichen werden, ohne dass sich die Lösungsmenge ändert.*
- (2) *Enthält die erweiterte Matrix eines Gleichungssystems eine Zeile*

$$(0 \ 0 \ \dots \ 0 \ \gamma) \quad \text{mit } \gamma \neq 0,$$

so hat das Gleichungssystem keine Lösung. (D.h., die allgemeine Lösung ist die leere Menge.)

BEISPIEL 2.3.3. Lösen Sie das System mit der erweiterten Matrix:

$$\begin{pmatrix} 10 & 2 & 14 & 4 & 6 \\ 1 & 3 & 19 & 0 & 5 \\ 11 & -2 & -11 & 5 & 0 \\ 14 & 0 & 2 & 6 & 4 \end{pmatrix}$$

Lösung: Wir beginnen mit den Eliminationsschritten.

$$\begin{pmatrix} 10 & 2^* & 14 & 4 & 6 \\ 1 & 3 & 19 & 0 & 5 \\ 11 & -2 & -11 & 5 & 0 \\ 14 & 0 & 2 & 6 & 4 \end{pmatrix} \rightarrow \begin{pmatrix} 5 & 1^* & 7 & 2 & 3 \\ -14 & 0 & -2 & -6 & -4 \\ 21 & 0 & 3 & 9 & 6 \\ 14 & 0 & 2^* & 6 & 4 \end{pmatrix} \\ \rightarrow \begin{pmatrix} 5 & 1^* & 7 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 7 & 0 & 1^* & 3 & 2 \end{pmatrix}$$

Nun finden wir kein Pivotelement mehr, denn die einzigen Zeilen, die noch nicht Pivotzeilen waren, sind Nullzeilen. Es bleibt ein Rücksubstitutionsschritt:

$$\begin{pmatrix} 5 & 1^{*1} & 7 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 7 & 0 & 1^{*2} & 3 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} -44 & 1^{*1} & 0 & -19 & -11 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 7 & 0 & 1^{*2} & 3 & 2 \end{pmatrix}$$

Die beiden Nullzeilen können wir streichen.

$$\begin{pmatrix} -44 & 1^{*1} & 0 & -19 & -11 \\ 7 & 0 & 1^{*2} & 3 & 2 \end{pmatrix}.$$

Das neue Problem ist, dass bei diesem System Spalten übrig geblieben sind, die keine Pivotspalten waren. Wir haben am Ende der Elimination, nach Streichung der überflüssigen Nullzeilen, weniger Gleichungen als Unbekannte übrig.

Wir unterscheiden nun zwei Arten von Spalten. Die zweite und dritte Spalte enthalten Pivotelemente, wir werden diese Spalten als basische⁵ Spalten bezeichnen. Die erste und die vierte Spalte sind nichtbasisch. Wir schreiben jetzt die Gleichungen aus, wobei wir die nichtbasischen Variablen auf die rechte Seite bringen:

$$\begin{aligned} \tau_2 &= -11 + 44\tau_1 + 19\tau_4 \\ \tau_3 &= 2 - 7\tau_1 - 3\tau_4 \end{aligned}$$

Jeder Vektor $\begin{pmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \end{pmatrix}$, der diese zwei Bedingungen erfüllt, ist eine partikuläre Lösung des Systems.

Dabei können τ_1 und τ_4 beliebige reelle Zahlen sein. Die allgemeine Lösung (die Menge aller Lösungsvektoren) ist also

$$\left\{ \begin{pmatrix} \tau_1 \\ -11 + 44\tau_1 + 19\tau_4 \\ 2 - 7\tau_1 - 3\tau_4 \\ \tau_4 \end{pmatrix} \mid \tau_1, \tau_4 \in \mathbb{R} \right\}.$$

Oft gibt man in solchen Fällen den nichtbasischen Variablen neue Namen, zum Beispiel $\tau_1 = \lambda$, $\tau_4 = \mu$ und zerlegt die Lösungsvektoren in eine Linearkombination von je einem Vektor mit den Konstanten, einen mit den Vielfachen von λ und einen mit den Vielfachen von μ . Weil wir bisher noch keine Vektoraddition eingeführt haben, bleibt das vorläufig einfach eine Schreibweise:

$$\left\{ \begin{pmatrix} 0 \\ -11 \\ 2 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 44 \\ -7 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 19 \\ -3 \\ 1 \end{pmatrix} \mid \lambda, \mu \in \mathbb{R} \right\}.$$

□

BEMERKUNG 2.3.4. Am Ende der Elimination, wenn keine weiteren Pivotelemente mehr gefunden werden können, sieht man normalerweise zwei Arten von Spalten und damit zwei Arten von Variablen:

- Basische Spalten enthalten ein Pivotelement.
- Nichtbasische Spalten wurden nie als Pivotspalten verwendet.

Nach der Rücksubstitution löst man das Gleichungssystem auf, indem man alle Terme, die nichtbasische Variablen enthalten, auf die rechte Seite des Systems verschiebt. Auf der linken Seite jeder Gleichung steht dann nur mehr eine Variable (außer Nullen), die Pivotvariable, zu der die jeweilige Gleichung Pivotzeile war. Diese basische Variable lässt sich nun durch die nichtbasischen Variablen ausdrücken. Die nichtbasischen Variablen sind dann freie Parameter, d.h., jeder beliebige Satz von Werten für die nichtbasischen Variablen liefert auf diese Art genau eine partikuläre Lösung des Systems.

⁵Der Begriff basische Spalte kommt aus der linearen Optimierung. Im Augenblick können wir ihn noch nicht erklären.

BEMERKUNG 2.3.5. Wenn ein lineares Gleichungssystem überhaupt Lösungen besitzt, hängt die Anzahl der freien Parameter in der allgemeinen Lösungen davon ab, wieviele Spalten nichtbasisch geblieben sind. Weil man in einem gewissen Rahmen die Pivotelemente frei wählen kann, können verschiedene Personen zu verschiedenen basischen Spalten kommen. Die Anzahl der nichtbasischen Spalten hängt aber, wie wir später beweisen werden, vom Verlauf des Lösungsweges nicht ab.

2.4. Der fertige Algorithmus.

Ein Algorithmus ist ein Rechenverfahren, das so detailliert in einzelne Rechenschritte aufgeschlüsselt ist, dass man es auf einem Computer programmieren könnte.

ALGORITHMUS 2.4.1 (Gaußsche Elimination). Zu lösen ist ein System von m linearen Gleichungen in n Unbekannten. Wir gehen von der erweiterten Matrix des Systems aus.

Eliminationsphase:

- 1) Gibt es eine Zeile i in der erweiterten Matrix, die noch nicht Pivotzeile war, und auf der linken Seite ein Element $\alpha_{i,j} \neq 0$ enthält? Wenn ja, wähle ein solches Element als Pivotelement. (Zeile i heißt dann Pivotzeile, Spalte j heißt Pivotspalte.) Wenn nein, dann gehe zu Schritt 4.
- 2) Von jeder anderen Zeile $k \neq i$, welche noch nicht Pivotzeile war, subtrahiere das $\alpha_{k,j}/\alpha_{i,j}$ -Fache der Pivotzeile. Es entsteht eine Null in der Pivotspalte.
- 3) Gehe zu Schritt 1.

Behandlung der Nullzeilen:

- 4) Alle Zeilen, die nur aus Nullen bestehen, werden gestrichen.
- 5) Gibt es eine Zeile, die links nur Nullen besitzt, aber rechts ein Element ungleich Null enthält? Wenn ja, dann hat das System überhaupt keine Lösung. Beende die Rechnung. Wenn nein, gehe weiter zu Schritt 6.

Rücksubstitutionsphase:

- 6) Gibt es noch Zeilen, welche nicht für die Rücksubstitution verwendet wurden? Wenn ja, wähle aus ihnen jene Zeile i , die als letzte Pivotzeile war und gehe zu Schritt 7. Wenn nein, gehe zu Schritt 10.
- 7) Wähle das ehemalige Pivotelement $\alpha_{i,j}$ wieder als Pivotelement.
- 8) Von allen Zeilen k , die vor der Zeile Nummer i Pivotzeilen waren, subtrahiere das $\alpha_{k,j}/\alpha_{i,j}$ -Fache der Pivotzeile. Es entstehen Nullen in der Pivotspalte.
- 9) Gehe zu Schritt 6.

Freie Parameter und Auflösung:

- 10) Gibt es Spalten, die nie Pivotspalten waren? Setze die entsprechenden Variablen ξ_j mit allgemeinen Zahlen (=Buchstaben) als freie Parameter an.
- 11) Aus jeder Zeile i , deren Pivotelement $\alpha_{i,j}$ war, läßt sich der Wert von τ_j jetzt ablesen. Ende des Verfahrens.

BEMERKUNG 2.4.2.

- (1) Von diesem Verfahren gibt es verschiedene Varianten. Eine ist, die Rücksubstitutionsphase nicht separat zu führen, sondern gleich in der Eliminationsphase auch in den Zeilen, die schon Pivotzeilen waren, zu eliminieren. Wenn man sich aber die Rücksubstitution für das Ende aufhebt, werden die Schritte wegen der vielen Nullen besonders einfach.

- (2) Oft wird bei der Erklärung des Gaußschen Verfahrens der Begriff der Zeilenstufenform eingeführt, die sich bei der Elimination ergibt, wenn man unter den möglichen Pivotelementen immer jenes aussucht, das am weitesten oben und dann am weitesten links steht. Das ist aber nicht notwendig.

3. Analytische Geometrie

3.1. Endlichdimensionale affine Räume.

DEFINITION 3.1.1.

Sei $n \in \mathbb{N}$.

- (1) Wir führen auf \mathbb{R}^n Grundrechnungsarten ein: Die Addition und die Subtraktion, und die Multiplikation mit einer reellen Zahl:

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} \pm \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} := \begin{pmatrix} \alpha_1 \pm \beta_1 \\ \alpha_2 \pm \beta_2 \\ \vdots \\ \alpha_n \pm \beta_n \end{pmatrix} \quad \gamma \cdot \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} := \begin{pmatrix} \gamma\alpha_1 \\ \gamma\alpha_2 \\ \vdots \\ \gamma\alpha_n \end{pmatrix}.$$

Wie bei der Multiplikation von reellen Zahlen lassen wir zumeist den Multiplikationspunkt weg: $\lambda a = \lambda \cdot a$. Wir verwenden auch die gewohnt Priorität der Rechnungsarten: Punktrechnung vor Strichrechnung.

- (2) Der Nullvektor ist der Spaltenvektor

$$\begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Weil normalerweise keine Gefahr der Verwechslung besteht, schreiben wir den Nullvektor mit demselben Zeichen 0 wie den Skalar Null.

BEHAUPTUNG 3.1.2. Die folgenden Rechenregeln lassen sich leicht nachprüfen. Sie gelten für alle $a, b, c \in \mathbb{R}^n$, $\lambda, \mu \in \mathbb{R}$:

- (1) $a + b = b + a$. Kommutativität.
- (2) $(a + b) + c = a + (b + c)$. Assoziativität.
- (3) $a + 0 = a$. Der Nullvektor ist ein neutrales Element für +.
- (4) $a + (-1) \cdot a = 0$. Der Vektor $(-1) \cdot a$ ist invers zu a für +.
- (5) $\lambda(a + b) = \lambda a + \lambda b$. Distributivität.
- (6) $(\lambda + \mu)a = \lambda a + \mu a$. Distributivität.
- (7) $\lambda(\mu a) = (\lambda\mu)a$.
- (8) $1 \cdot a = a$.

Wir werden später sagen: \mathbb{R}^n ist ein Vektorraum über \mathbb{R} .

BEWEIS. Die Beweise dieser Behauptungen folgen alle dem gleichen Schema. Wir beweisen als Beispiel das Distributivgesetz (5). Es seien

$$a = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \in \mathbb{R}^n, \quad \text{und} \quad \lambda \in \mathbb{R}.$$

Dann gilt

$$\begin{aligned} \lambda(a+b) &= \lambda\left[\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} + \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}\right] = \lambda\begin{pmatrix} \alpha_1 + \beta_1 \\ \vdots \\ \alpha_n + \beta_n \end{pmatrix} = \begin{pmatrix} \lambda(\alpha_1 + \beta_1) \\ \vdots \\ \lambda(\alpha_n + \beta_n) \end{pmatrix} \\ &= \begin{pmatrix} \lambda\alpha_1 + \lambda\beta_1 \\ \vdots \\ \lambda\alpha_n + \lambda\beta_n \end{pmatrix} = \begin{pmatrix} \lambda\alpha_1 \\ \vdots \\ \lambda\alpha_n \end{pmatrix} + \begin{pmatrix} \lambda\beta_1 \\ \vdots \\ \lambda\beta_n \end{pmatrix} = \lambda\begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} + \lambda\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \lambda a + \lambda b. \end{aligned}$$

□

Um die geometrischen Vorstellungen anschaulich zu machen, die wir mit \mathbb{R}^n verbinden, betrachten wir zunächst eine Ebene \mathbb{A} , die wir als Menge von Punkten auffassen. Die Punkte werden wir mit lateinischen Großbuchstaben bezeichnen.

In der Ebene fixieren wir einen Punkt O (den Koordinatenursprung) und ein Kreuz von Geraden durch O , die nicht zusammenfallen (die Koordinatenachsen). Auf den Geraden stecken wir jeweils einen zweiten, von O verschiedenen Punkt E_1, E_2 ab. Die Strecke von O zu E_1 bzw. E_2 bezeichnen wir als die Einheitslänge in Richtung 1 oder 2. (Es spielt für die jetzigen Überlegungen keine Rolle, ob die Geraden aufeinander senkrecht stehen, und ob die abgesteckten Längen in beide Richtungen gleich sind.)

Einen Vektor $\begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$ fassen wir nun als eine Parallelverschiebung auf, die jeden Punkt der Ebene um α_1 Längeneinheiten in Richtung der ersten Achse und um α_2 Längeneinheiten in Richtung der zweiten Achse verschiebt. Eine Verschiebung um eine negative "Länge" bedeutet in diesem Kontext eine Verschiebung entgegengesetzt zur Achsenrichtung. In der Zeichnung wird der Vektor durch einen Pfeil dargestellt. Wenn ein Punkt P um den Vektor v verschoben wird, bezeichnen wir das Ergebnis mit $P \boxplus v$.

Seien $a = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$, $b = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$ zwei Vektoren aus \mathbb{R}^2 und sei $\lambda \in \mathbb{R}$. Eine Verschiebung um den Vektor λa ist eine Verschiebung um $\lambda\alpha_1$ in die erste Achsenrichtung und um $\lambda\alpha_2$ in die zweite Achsenrichtung. Der Vektor λa entspricht also der Verschiebung a , verlängert oder verkürzt um den Faktor λ . Die Verschiebung um $a+b$ lässt sich bewerkstelligen, indem man einen Punkt erst um α_1 und α_2 in die beiden Achsenrichtungen verschiebt, und dann um β_1 und β_2 . Daher entspricht der Vektor $a+b$ der Hintereinanderausführung der beiden Verschiebungen um den Vektor a und um den Vektor b .

Für je zwei Punkte P, Q in der Ebene gibt es genau einen Vektor, der P nach Q verschiebt. Wir bezeichnen diesen Vektor als den Verbindungsvektor \vec{PQ} . Insbesondere hat jeder Punkt P einen eindeutig bestimmten Vektor, der den Koordinatenursprung nach P verschiebt. Wir nennen diesen Vektor den Ortsvektor von P . Der Ortsvektor von O ist der Nullvektor, und die Ortsvektoren von E_1 und E_2 sind die Einheitsvektoren $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Nun können wir jeden Punkt mit seinem Ortsvektor identifizieren. Rechentechnisch werden wir daher zwischen der Ebene \mathbb{A} und dem Raum \mathbb{R}^2 gar nicht unterscheiden müssen, obwohl nun Vektoren auf zwei verschiedene Weisen geometrisch gedeutet werden: Als Parallelverschiebungen in der Ebene, oder als Ortsvektoren zur Bezeichnung von Punkten der Ebene.

Beachten Sie bitte, dass das keine mathematischen Definitionen sind. Wir haben ja noch keinen festen Begriff von Länge und von Verschiebung. Die obigen Gedanken sollen Ihnen Bilder bereitstellen, die Sie sich zur folgenden exakten Definition eines affinen Raumes vorstellen können.

DEFINITION 3.1.3. Gegeben sei eine nichtleere Menge \mathbb{A} . Wir nehmen an, dass zu jedem $P \in \mathbb{A}$ und $v \in \mathbb{R}^n$ ein eindeutig bestimmtes Element $P \boxplus v \in \mathbb{A}$ (sprich: " P verschoben um v ")⁶ existiert, sodass die folgenden Regeln gelten:

- (1) Für alle $P \in \mathbb{A}$ und alle $v, w \in \mathbb{R}^n$ gilt $(P \boxplus v) \boxplus w = P \boxplus (v + w)$.

(Die Summe zweier Vektoren entspricht der Hintereinanderausführung von zwei Verschiebungen.)

⁶ \boxplus ist kein genormtes Symbol. In der Praxis kann man es ohne Gefahr von Verwechslungen durch ein gewöhnliches $+$ ersetzen.

- (2) Für alle $P, Q \in \mathbb{A}$ gibt es genau einen Vektor $v \in \mathbb{R}^n$, sodass $P \boxplus v = Q$.
(Von jedem Punkt zu jedem Punkt gibt es genau eine Verschiebung.)

Dann heißt \mathbb{A} ein n -dimensionaler affiner Raum über \mathbb{R} . Die Elemente von \mathbb{A} bezeichnen wir als Punkte. Für zwei Punkte $P, Q \in \mathbb{A}$ bezeichnen wir den Vektor mit der Eigenschaft $P \boxplus v = Q$ als Verbindungsvektor von P nach Q : $v = \vec{PQ}$.

DEFINITION 3.1.4. Wir fixieren einen beliebigen Punkt O eines affinen Raumes \mathbb{A} als Koordinatenursprung. Für jeden Punkt $P \in \mathbb{A}$ bezeichnen wir dann den Vektor \vec{OP} als den Ortsvektor von P .

Ergebnisse wie die folgende Bemerkung überraschen uns nicht, aber es gehört zur mathematischen Genauigkeit, dass auch sie aus den Definitionen hergeleitet werden.

BEHAUPTUNG 3.1.5. Sei \mathbb{A} ein n -dimensionaler affiner Raum über \mathbb{R} und O als Koordinatenursprung fixiert. Seien $P, Q \in \mathbb{A}$ und $a \in \mathbb{R}^n$.

- (1) Für alle $P \in \mathbb{A}$ gilt $P \boxplus 0 = P$.
Die "Verschiebung" um den Nullvektor läßt jeden Punkt stehen, wo er ist.
(2) Dann ist der Ortsvektor von $P \boxplus a$

$$[O, P \boxplus a]^\rightharpoonup = \vec{OP} + a.$$

- (3) Der Verbindungsvektor zweier Punkte in Gegenrichtung ist

$$\vec{QP} = -\vec{PQ}.$$

- (4) Der Verbindungsvektor zweier Punkte P und Q ist die Differenz ihrer Ortsvektoren:

$$\vec{PQ} = \vec{OQ} - \vec{OP}.$$

BEWEIS.

- (1) Sei $v = \vec{PP}$, der Verbindungsvektor von P nach P . Dann gilt

$$P \boxplus (v + v) = (P \boxplus v) \boxplus v = P \boxplus v = P.$$

Weil aber der Verbindungsvektor zwischen zwei Punkten eindeutig bestimmt ist, folgt daraus $v + v = v$ und daraus wieder $v = 0$. Also ist $P \boxplus 0 = P$.

- (2) Wir müssen zeigen: $O \boxplus (\vec{OP} + a) = P \boxplus a$:

$$O \boxplus (\vec{OP} + a) = (O \boxplus \vec{OP}) \boxplus a = P \boxplus a.$$

- (3) Wir müssen zeigen: $Q \boxplus (-\vec{PQ}) = P$:

$$Q \boxplus (-\vec{PQ}) = (P \boxplus \vec{PQ}) \boxplus (-\vec{PQ}) = P \boxplus (\vec{PQ} + (-\vec{PQ})) = P \boxplus 0 = P.$$

- (4) Wir müssen zeigen: $P \boxplus (\vec{OQ} - \vec{OP}) = Q$:

$$P \boxplus (\vec{OQ} - \vec{OP}) = P \boxplus (\vec{PO} + \vec{OQ}) = (P \boxplus \vec{PO}) \boxplus \vec{OQ} = O \boxplus \vec{OQ} = Q.$$

□

BEHAUPTUNG 3.1.6. Mit der Operation

$$P \boxplus v = P + v$$

für $P \in \mathbb{R}^n$, $v \in \mathbb{R}^n$ ist \mathbb{R}^n selbst ein n -dimensionaler affiner Raum über \mathbb{R} .
Mit dem Nullpunkt $O = 0$ wird jeder Punkt in \mathbb{R}^n zugleich sein eigener Ortsvektor.

BEWEIS. Wir müssen die zwei Bedingungen aus Definition 3.1.3 nachweisen.
Seien $P, Q, v, w \in \mathbb{R}^n$.

(1) Es gilt

$$(P \boxplus v) \boxplus w = (P + v) \boxplus w = (P + v) + w = P + (v + w) = P \boxplus (v + w).$$

(2) Der Vektor $v = Q - P$ ist der einzige Vektor sodass

$$Q = P \boxplus v = P + v.$$

□

3.2. Geometrische Aufgaben und lineare Gleichungssysteme.

DEFINITION 3.2.1. Sei \mathbb{A} ein n -dimensionaler affiner Raum über \mathbb{R} , sei $P \in \mathbb{A}$ ein Punkt und sei $v \neq 0 \in \mathbb{R}^n$. Die Gerade durch P in Richtung v ist definiert als

$$\{P \boxplus \lambda v \mid \lambda \in \mathbb{R}\}.$$

In Definition 3.2.1 war es wichtig, einen Richtungsvektor $v \neq 0$ zu haben. Für $v = 0$ würde das Gebilde auf den einen Punkt P zusammenfallen. Wie definiert man eine Ebene? Es reicht nicht, einfach 2 Richtungsvektoren zu nehmen, die nicht Null sind. Sie dürfen auch nicht Vielfache voneinander sein, sonst erhalten wir eine Gerade.

DEFINITION 3.2.2. Sei \mathbb{A} ein n -dimensionaler affiner Raum über \mathbb{R} , sei $P \in \mathbb{A}$ und seien $b_1, b_2 \in \mathbb{R}^n$ so, dass weder b_1 ein Vielfaches von b_2 ist, noch umgekehrt. Die Ebene durch P , deren Richtungen von b_1, b_2 aufgespannt werden, ist

$$\{P \boxplus (\lambda_1 b_1 + \lambda_2 b_2) \mid \lambda_1, \lambda_2 \in \mathbb{R}\}.$$

BEISPIEL 3.2.3. Gegeben seien 3 Punkte X, Y, Z eines dreidimensionalen affinen Raumes \mathbb{A} über \mathbb{R} mit den Ortsvektoren

$$\vec{OX} = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}, \quad \vec{OY} = \begin{pmatrix} 0 \\ 2 \\ 6 \end{pmatrix}, \quad \vec{OZ} = \begin{pmatrix} 4 \\ -3 \\ 0 \end{pmatrix}.$$

Legen Sie eine Ebene durch die drei Punkte.

Lösung: Wir können uns den Ausgangspunkt in der Ebene aussuchen, z.B. den Punkt X aus der Angabe, und als Richtungsvektoren kommen alle Vektoren in Frage, die zwei Punkte der Ebene verbinden, z.B. \vec{XY} und \vec{XZ} . Die Ortsvektoren der Ebene bilden dann die Menge

$$\{P \in \mathbb{A} \mid \vec{OP} = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix} + \lambda_1 \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 2 \\ -4 \\ -4 \end{pmatrix}\}.$$

Wir könnten genausogut einen anderen Punkt der Ebene als Ausgangspunkt nehmen und die Richtungsvektoren mit Konstanten multiplizieren, wir erhalten dieselbe Ebene:

$$\{P \in \mathbb{A} \mid \vec{OP} = \begin{pmatrix} 0 \\ 2 \\ 6 \end{pmatrix} + \lambda_1 \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix}\}.$$

□

Wir rechnen nun einige Beispiele, die den Zusammenhang zwischen analytischer Geometrie und linearen Gleichungssystemen deutlich machen, anders ausgedrückt, der affine Raum, den wir betrachten, ist der Raum \mathbb{R}^n selbst.

BEISPIEL 3.2.4. Bestimmen Sie den Durchschnitt der Geraden

$$G = \left\{ \begin{pmatrix} -1 \\ -3 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 3 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\},$$

$$H = \left\{ \begin{pmatrix} 1 \\ -2 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 2 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\}.$$

Lösung: Um nicht denselben Buchstaben λ in zwei verschiedenen Bedeutungen in der Rechnung zu haben, parametrisieren wir die Gerade H mit μ . Wie der Parameter heißt, ist schließlich gleichgültig. Wir suchen also einen Punkt x , der zugleich in zwei Mengen liegt:

$$x \in \left\{ \begin{pmatrix} -1 \\ -3 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 3 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\} \cap \left\{ \begin{pmatrix} 1 \\ -2 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 2 \end{pmatrix} \mid \mu \in \mathbb{R} \right\}.$$

Der Vektor x muss sich also auf zwei verschiedene Arten schreiben lassen, und wir erhalten

$$\begin{aligned} x &= \begin{pmatrix} -1 \\ -3 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \\ \lambda \begin{pmatrix} 2 \\ 3 \end{pmatrix} - \mu \begin{pmatrix} 1 \\ 2 \end{pmatrix} &= -\begin{pmatrix} -1 \\ -3 \end{pmatrix} + \begin{pmatrix} 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}. \end{aligned}$$

Schreiben wir die einzelnen Komponenten auf, erhalten wir zwei Gleichungen in zwei Unbekannten:

$$\begin{aligned} 2\lambda - \mu &= 2 \\ 3\lambda - 2\mu &= 1 \end{aligned}$$

Wir lösen das System:

$$\begin{pmatrix} 2 & -1 & 2 \\ 3 & -2 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & -1 & 2 \\ -1 & 0 & -3 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & -1 & -4 \\ -1 & 0 & -3 \end{pmatrix}$$

Nun können wir direkt ablesen:

$$\lambda = 3, \quad \mu = 4.$$

Wir setzen $\lambda = 3$ in die Gleichung von G ein und erhalten

$$x = \begin{pmatrix} -1 \\ -3 \end{pmatrix} + 3 \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}.$$

Als Probe können wir auch μ in die Gleichung von H einsetzen und müssen denselben Vektor x als Schnittpunkt erhalten:

$$x = \begin{pmatrix} 1 \\ -2 \end{pmatrix} + 4 \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}.$$

Die Probe zeigt also, dass wir richtig gerechnet haben. Der gesuchte Schnittpunkt ist

$$x = \begin{pmatrix} 5 \\ 6 \end{pmatrix}.$$

□

BEISPIEL 3.2.5. Gegeben seien die folgende Ebene E und die Punkte x, y . Liegen $x \in E, y \in E$?

$$x = \begin{pmatrix} 5 \\ -4 \\ 7 \end{pmatrix}, \quad y = \begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix}, \quad E = \left\{ \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} \mid \lambda, \mu \in \mathbb{R} \right\}.$$

Lösung: x liegt genau dann in E , wenn es sich in der Form

$$x = \begin{pmatrix} 5 \\ -4 \\ 7 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix}$$

schreiben lässt. Wenn wir diese Gleichung wieder in einzelne Komponenten auflösen, erhalten wir ein System von drei Gleichungen in zwei Unbekannten:

$$\begin{aligned} 2\lambda + \mu &= 4 \\ \lambda + 4\mu &= -5 \\ 3\lambda + 2\mu &= 5 \end{aligned}$$

Wenn wir überprüfen wollen, ob y in E liegt, erhalten wir ebenso ein Gleichungssystem

$$\begin{aligned} 2\lambda + \mu &= 1 \\ \lambda + 4\mu &= 4 \\ 3\lambda + 2\mu &= 6 \end{aligned}$$

Da die linken Seiten der Gleichungssysteme dieselben sind, können wir sie in einem Aufwaschen lösen: Wir führen die elementaren Zeilentransformationen mit beiden rechten Seiten zugleich durch.

$$\begin{aligned} & \begin{pmatrix} 2 & 1 & 4 & 1 \\ 1* & 4 & -5 & 4 \\ 3 & 2 & 5 & 6 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & -7* & 14 & -7 \\ 1 & 4 & -5 & 5 \\ 0 & -10 & 20 & -6 \end{pmatrix} \\ \rightarrow & \begin{pmatrix} 0 & -1 & 2 & -1 \\ 1 & 4 & -5 & 5 \\ 0 & 0 & 0 & -4(!) \end{pmatrix} \rightarrow \begin{pmatrix} 0 & -1 & 2 & -1 \\ 1 & 0 & 3 & 3 \\ 0 & 0 & 0 & -4(!) \end{pmatrix} \end{aligned}$$

Im ersten Fall ist $\lambda = 3$ und $\mu = -2$. Als Probe setzen wir λ und μ in die Gleichung der Ebene ein und erhalten tatsächlich x :

$$\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + 3 \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} - 2 \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ -4 \\ 7 \end{pmatrix}.$$

Der Vektor x liegt also in der Ebene E .

Im zweiten Fall gibt es keine Lösung, also liegt y nicht in E . □

BEISPIEL 3.2.6. Bestimmen Sie den Durchschnitt der Ebene E mit der Geraden G :

$$\begin{aligned} E &= \left\{ \begin{pmatrix} 7 \\ 14 \\ -11 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 8 \\ 4 \end{pmatrix} + \mu \begin{pmatrix} 3 \\ 9 \\ 0 \end{pmatrix} \mid \lambda, \mu \in \mathbb{R} \right\}, \\ G &= \left\{ \begin{pmatrix} 11 \\ 2 \\ -19 \end{pmatrix} + \lambda \begin{pmatrix} 5 \\ 0 \\ -10 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\}. \end{aligned}$$

Lösung: Wir suchen wieder einen Punkt x , der beide Parameterformen erfüllt, also

$$x = \begin{pmatrix} 7 \\ 14 \\ -11 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 8 \\ 4 \end{pmatrix} + \mu \begin{pmatrix} 3 \\ 9 \\ 0 \end{pmatrix} = \begin{pmatrix} 11 \\ 2 \\ -19 \end{pmatrix} + \nu \begin{pmatrix} 5 \\ 0 \\ -10 \end{pmatrix}.$$

(Beachten Sie, dass wir wieder den Parameter aus G umbenannt haben, um nicht einen Buchstaben für zwei verschiedene Zahlen zu verwenden.) Wir sortieren die Terme und erhalten

$$\lambda \begin{pmatrix} 2 \\ 8 \\ 4 \end{pmatrix} + \mu \begin{pmatrix} 3 \\ 9 \\ 0 \end{pmatrix} - \nu \begin{pmatrix} 5 \\ 0 \\ -10 \end{pmatrix} = \begin{pmatrix} 11 \\ 2 \\ -19 \end{pmatrix} - \begin{pmatrix} 7 \\ 14 \\ -11 \end{pmatrix} = \begin{pmatrix} 4 \\ -12 \\ -8 \end{pmatrix}.$$

$$\begin{pmatrix} 2 & 3 & -5* & 4 \\ 8 & 9 & 0 & -12 \\ 4 & 0 & 10 & -8 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 3 & -5 & 4 \\ 8* & 9 & 0 & -12 \\ 8 & 6 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 3 & -5 & 4 \\ 8 & 9 & 0 & -12 \\ 0 & -3* & 0 & 12 \end{pmatrix}$$

Wir beginnen die Rücksubstitution

$$\begin{pmatrix} 2 & 0 & -5 & 16 \\ 8* & 0 & 0 & 24 \\ 0 & -3 & 0 & 12 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 0 & -5 & 10 \\ 8 & 0 & 0 & 24 \\ 0 & -3 & 0 & 12 \end{pmatrix}$$

Nun steht die Lösung vor uns:

$$\lambda = 3, \quad \mu = -4, \quad \nu = -2.$$

Setzen wir $\lambda = 3$, $\mu = -4$ in die Ebenengleichung ein, so erhalten wir den Schnittpunkt

$$x = \begin{pmatrix} 7 \\ 14 \\ -11 \end{pmatrix} + 3 \begin{pmatrix} 2 \\ 8 \\ 4 \end{pmatrix} - 4 \begin{pmatrix} 3 \\ 9 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}.$$

Wir können, als Probe, ebenso $\nu = -2$ in die Geradengleichung einsetzen, und müssen denselben Schnittpunkt erhalten:

$$x = \begin{pmatrix} 11 \\ 2 \\ -19 \end{pmatrix} - 2 \begin{pmatrix} 5 \\ 0 \\ -10 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}.$$

□

BEISPIEL 3.2.7. Bestimmen Sie den Durchschnitt der Ebenen

$$E = \left\{ \begin{pmatrix} 5 \\ -1 \\ 9 \end{pmatrix} + \lambda \begin{pmatrix} 4 \\ -4 \\ 5 \end{pmatrix} + \mu \begin{pmatrix} 4 \\ -2 \\ 3 \end{pmatrix} \mid \lambda, \mu \in \mathbb{R} \right\},$$

$$F = \left\{ \begin{pmatrix} 7 \\ 7 \\ 6 \end{pmatrix} + \lambda \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix} + \mu \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} \mid \lambda, \mu \in \mathbb{R} \right\}.$$

Lösung: Wir suchen wieder Punkte, die zugleich in beiden Parameterformen geschrieben werden können:

$$\begin{pmatrix} 5 \\ -1 \\ 9 \end{pmatrix} + \lambda \begin{pmatrix} 4 \\ -4 \\ 5 \end{pmatrix} + \mu \begin{pmatrix} 4 \\ -2 \\ 3 \end{pmatrix} = \begin{pmatrix} 7 \\ 7 \\ 6 \end{pmatrix} + \nu \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix} + \kappa \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix},$$

$$\lambda \begin{pmatrix} 4 \\ -4 \\ 5 \end{pmatrix} + \mu \begin{pmatrix} 4 \\ -2 \\ 3 \end{pmatrix} - \nu \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix} - \kappa \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 7 \\ 7 \\ 6 \end{pmatrix} - \begin{pmatrix} 5 \\ -1 \\ 9 \end{pmatrix} = \begin{pmatrix} 2 \\ 8 \\ -3 \end{pmatrix}.$$

Damit erhalten wir das folgende Schema, wobei jede Zeile wieder für eine Gleichung steht, und die Spalten nacheinander die Koeffizienten für $\lambda, \mu, \nu, \kappa$ und die rechte Seite beinhalten:

$$\begin{pmatrix} 4 & 4 & -2 & -3 & 2 \\ -4 & -2 & -2 & -1* & 8 \\ 5 & 3 & -3 & -2 & -3 \end{pmatrix} \rightarrow \begin{pmatrix} 16 & 10 & 4 & 0 & -22 \\ -4 & -2 & -2 & -1* & 8 \\ 13 & 7 & 1* & 0 & -19 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} -36 & -18 & 0 & 0 & 54 \\ -4 & -2 & -2 & -1* & 8 \\ 13 & 7 & 1* & 0 & -19 \end{pmatrix} \rightarrow \begin{pmatrix} -2 & -1* & 0 & 0 & 3 \\ -4 & -2 & -2 & -1* & 8 \\ 13 & 7 & 1* & 0 & -19 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} -2 & -1* & 0 & 0 & 3 \\ 0 & 0 & -2 & -1* & 2 \\ -1 & 0 & 1* & 0 & 2 \end{pmatrix} \rightarrow \begin{pmatrix} -2 & -1* & 0 & 0 & 3 \\ -2 & 0 & 0 & -1* & 6 \\ -1 & 0 & 1* & 0 & 2 \end{pmatrix}$$

Nun ist auch die Rücksubstitution fertig. Trotzdem können wir keine eindeutige Lösung ablesen, denn die erste Spalte war nie Pivotspalte. Unsere Gleichungen bedeuten:

$$\begin{aligned} -2\lambda & - \mu & & & & = & 3 \\ -2\lambda & & & & - \kappa & = & 6 \\ -\lambda & & + \nu & & & = & 2 \end{aligned}$$

Für jeden Wert von λ können wir Werte von μ, ν, κ berechnen, sodass sich eine Lösung ergibt:

$$\mu = -3 - 2\lambda, \quad \nu = 2 + \lambda, \quad \kappa = -6 - 2\lambda.$$

Wir setzen in die erste Ebenengleichung ein und erhalten

$$\begin{aligned} E \cap F &= \left\{ \begin{pmatrix} 5 \\ -1 \\ 9 \end{pmatrix} + \lambda \begin{pmatrix} 4 \\ -4 \\ 5 \end{pmatrix} + (-3 - 2\lambda) \begin{pmatrix} 4 \\ -2 \\ 3 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} 5 \\ -1 \\ 9 \end{pmatrix} - 3 \begin{pmatrix} 4 \\ -2 \\ 3 \end{pmatrix} + \lambda \begin{pmatrix} 4 \\ -4 \\ 5 \end{pmatrix} - 2\lambda \begin{pmatrix} 4 \\ -2 \\ 3 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} -7 \\ 5 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} -4 \\ 0 \\ -1 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\}. \end{aligned}$$

Das ist die Parameterform einer Geraden. Die beiden Ebenen schneiden einander also in einer Geraden. \square

3.3. Unterräume und affine Unterräume.

Einzelne Punkte, Geraden und Ebenen sind Sonderfälle von sogenannten affinen Unterräumen eines affinen Raumes. Obwohl sich unsere Anschauung jenseits der dritten Dimension schwer tut, kann es vom mathematischen Standpunkt aus auch höher dimensionale affine Räume geben.

Wir betrachten aber zunächst nur den Vektorraum \mathbb{R}^n und nicht die Punktmenge. In der geometrischen Deutung sind die Vektoren Parallelverschiebungen auf der Punktmenge. Wenn man zwei Parallelverschiebungen hintereinander ausführt, erhält man wieder eine Parallelverschiebung. Auch wenn eine Parallelverschiebung mit gleicher oder genau umgekehrter Richtung, aber veränderter Länge durchgeführt wird, ist das Ergebnis eine Parallelverschiebung. Man kann

also, ausgehend von einigen Parallelverschiebungen, eine große Anzahl weiterer Verschiebungen konstruieren.

DEFINITION 3.3.1. Seien v_1, \dots, v_k Vektoren in \mathbb{R}^n .

- (1) Sind $\lambda_1, \dots, \lambda_k \in \mathbb{R}$, so heißt der Vektor $\lambda_1 v_1 + \dots + \lambda_k v_k$ eine Linearkombination von v_1, \dots, v_k .
- (2) Die lineare Hülle des k -Tupels (v_1, \dots, v_k) ist die Menge aller Linearkombinationen dieser Vektoren:

$$\mathcal{L}(v_1, \dots, v_k) := \{\lambda_1 v_1 + \dots + \lambda_k v_k \mid \lambda_1, \dots, \lambda_k \in \mathbb{R}\}.$$

- (3) Die lineare Hülle der leeren Menge wird definiert als der Nullraum, das ist die Menge, die nur aus dem Nullvektor besteht:

$$\mathcal{L}(\emptyset) := \{0\}.$$

- (4) Ist U die lineare Hülle von (v_1, \dots, v_k) , so sagen wir auch, U wird von (v_1, \dots, v_k) aufgespannt, oder (v_1, \dots, v_k) ist ein Erzeugendensystem von U .

BEMERKUNG 3.3.2. Wir können die Definition der linearen Hülle natürlich auch mit dem Summenzeichen schreiben:

$$\mathcal{L}(v_1, \dots, v_k) = \left\{ \sum_{i=1}^k \lambda_i v_i \mid \lambda_i \in \mathbb{R} \right\}.$$

Bitte gewöhnen Sie sich möglichst schnell an den Gebrauch des Summenzeichens: Seien $p \leq q$ ganze Zahlen, und v_p, \dots, v_q Zahlen oder Vektoren, dann ist

$$\sum_{i=p}^q v_i := v_p + v_{p+1} + \dots + v_q.$$

D.h.: Zähle die Objekte v_i von $i = p$ bis $i = q$ auf, und addiere sie alle.

Der Summationsindex i kann jeden Namen haben, er ist nur ein Platzhalter für die Zahlen p, \dots, q , man könnte genausogut t oder jeden anderen Buchstaben schreiben.

Oben haben wir betrachtet, wie aus einer Menge von wenigen Vektoren viele weitere Vektoren gebildet werden können. Wie sieht nun eine Menge von Vektoren aus, aus der sich keine weiteren Vektoren mehr kombinieren lassen?

DEFINITION 3.3.3. Eine Teilmenge $U \subseteq \mathbb{R}^n$ heißt ein Unterraum von \mathbb{R}^n , wenn sie folgende Eigenschaften hat:

- (1) U ist nicht leer.
- (2) Für alle $x \in U$ und alle $\lambda \in \mathbb{R}$ ist $\lambda x \in U$.
- (3) Für alle $x, y \in U$ ist $x + y \in U$.

BEHAUPTUNG 3.3.4. Ist U ein Unterraum von \mathbb{R}^n so gilt:

- (1) Der Nullvektor ist Element von U .
- (2) Jede Linearkombination von Vektoren aus U liegt wieder in U .

BEWEIS.

- (1) Es gibt jedenfalls ein Element $x \in U$, weil U nicht leer ist. Dann ist aber auch $0 = 0x \in U$.
- (2) Sind v_1, \dots, v_k Vektoren aus U und $\lambda_1, \dots, \lambda_k \in \mathbb{R}$, so ist auch jedes $\lambda_i v_i \in U$, weil Multiplikation mit Skalaren nicht aus U herausführt. Weil aber auch Addition nicht aus U herausführt, ist $\sum_{i=1}^k \lambda_i v_i \in U$.

□

SATZ 3.3.5. Seien $v_1, \dots, v_k \in \mathbb{R}^n$. Dann ist die lineare Hülle $\mathcal{L}(v_1, \dots, v_k)$ der kleinste Unterraum des \mathbb{R}^n , der alle Vektoren v_1, \dots, v_k enthält. Das heißt:

- (1) $\mathcal{L}(v_1, \dots, v_k)$ enthält alle v_i .
- (2) $\mathcal{L}(v_1, \dots, v_k)$ ist ein Unterraum.
- (3) Ist U ein Unterraum, der alle v_i enthält, dann gilt $\mathcal{L}(v_1, \dots, v_k) \subseteq U$.

BEWEIS.

- (1) Sei $i \in \{1, \dots, k\}$. Dann ist

$$v_i = 0v_1 + \dots + 1v_i + \dots + 0v_k \in \mathcal{L}(v_1, \dots, v_k).$$

- (2) Wir brauchen die drei Unterraumeigenschaften:

$$0 = 0v_1 + \dots + 0v_k \in \mathcal{L}(v_1, \dots, v_k).$$

Sei $x = \sum \mu_i v_i \in \mathcal{L}(v_1, \dots, v_k)$ und $\lambda \in \mathbb{R}$. Dann ist $\lambda x = \sum_{i=1}^k (\lambda \mu_i) v_i \in \mathcal{L}(v_1, \dots, v_k)$.

Seien $x = \sum_{i=1}^k \mu_i v_i$ und $y = \sum_{i=1}^k \nu_i v_i \in \mathcal{L}(v_1, \dots, v_k)$. Dann ist $x + y = \sum_{i=1}^k (\mu_i + \nu_i) v_i \in \mathcal{L}(v_1, \dots, v_k)$.

- (3) Sei U ein Unterraum, der alle v_i enthält. Wegen Bemerkung 3.3.4 enthält U auch alle Linearkombinationen der v_i , also ist $\mathcal{L}(v_1, \dots, v_k) \subseteq U$.

□

BEHAUPTUNG 3.3.6. Der Durchschnitt zweier Unterräume des \mathbb{R}^n ist wieder ein Unterraum.

BEWEIS. Seien U und V zwei Unterräume des \mathbb{R}^n . Wir zeigen die Unterraumeigenschaften für die Menge $U \cap V$:

- (1) Weil jeder Unterraum den Nullvektor enthält, liegt 0 auch in U und in V , daher ist $0 \in U \cap V$. Insbesondere ist also $U \cap V$ nicht leer.
- (2) Sei $x \in U \cap V$ und $\lambda \in \mathbb{R}$. Weil $x \in U$ und U ein Unterraum ist, ist auch $\lambda x \in U$. Mit demselben Argument folgt auch $\lambda x \in V$. Daher ist $\lambda x \in U \cap V$.
- (3) Seien x und y beide Elemente von $U \cap V$. Dann ist insbesondere $x \in U$ und $y \in U$. Weil U ein Unterraum ist, ist auch $x + y \in U$. Ebenso zeigt man $x + y \in V$. Daher ist $x + y \in U \cap V$.

□

Für jetzt bleibt die Frage offen, ob wirklich jeder Unterraum des \mathbb{R}^n als lineare Hülle endlich vieler Vektoren dargestellt werden kann, und wenn das so ist, wieviele Vektoren mindestens gebraucht werden. Diese Frage wird uns später zum Begriff der Basis und der Dimension führen.

Wir kehren jetzt zur Betrachtung des affinen Raumes zurück.

DEFINITION 3.3.7. Sei \mathbb{A} ein n -dimensionaler affiner Raum über \mathbb{R} , seien U ein Unterraum des \mathbb{R}^n und $P \in \mathbb{A}$. Dann heißt die Menge⁷

$$(3.3.1) \quad W = P \boxplus U := \{P \boxplus u \mid u \in U\}$$

⁷Die Schreibweise $P + U$ für die Menge aller Punkte, die von P aus durch Anwendung einer Verschiebung aus U erreicht werden können, ist sehr bequem. Ebenso kann man für zwei Zahlen- oder Vektormengen A, B kurz definieren: $A \pm B = \{a \pm b \mid a \in A, b \in B\}$. Verwenden Sie diese Schreibweise aber mit Vorsicht. Zum Beispiel ist

$$\{0, 1\} + \{2, 3\} - \{2, 3\} = \{-1, 0, 1, 2\}.$$

ein affiner Unterraum von \mathbb{A} durch den Punkt P .

Der Unterraum U heißt der Tangentialraum des affinen Unterraumes.

BEMERKUNG 3.3.8. Der Sonderfall $U = \{0\}$, $W = \{P\}$ liefert einen einzelnen Punkt als affinen Unterraum.

BEMERKUNG 3.3.9. Wenn der Unterraum U die lineare Hülle der Vektoren v_1, \dots, v_k ist, kann der affine Unterraum in der Form

$$W = \{P \boxplus (\lambda_1 v_1 + \dots + \lambda_k v_k) \mid \lambda_i \in \mathbb{R}\}$$

geschrieben werden. Wir nennen das die Parameterform eines affinen Unterraumes.

Unterscheiden Sie die Begriffe Unterraum und affiner Unterraum: Ein affiner Unterraum ist eine Menge von Punkten, ein Unterraum ist eine Menge von Vektoren (also, in der geometrischen Deutung, von Parallelverschiebungen). Aber auch, wenn \mathbb{R}^n als affiner Raum aufgefasst wird, und jeder Punkt zugleich ein Vektor ist, ist nicht jeder affine Unterraum ein Unterraum, sondern nur jene affinen Unterräume, die den Nullpunkt enthalten. Die Sprache täuscht hier: Nicht jeder affine Unterraum ist ein Unterraum, aber jeder Unterraum ist ein affiner Unterraum!

Der Tangentialraum eines affinen Unterraumes W ist die Menge der Verschiebungen, die je zwei Punkte ineinander verschieben. Anders ausgedrückt, eine Verschiebung liegt genau dann im Tangentialraum, wenn sie nicht aus dem affinen Unterraum hinausführt.

Die Intuition sagt uns, dass wir jeden Punkt eines affinen Unterraumes W als Ausgangspunkt zur Erzeugung von ganz W heranziehen können. Die folgende Behauptung zeigt, dass unsere Anschauung hier recht hat:

BEHAUPTUNG 3.3.10. Sei \mathbb{A} ein n -dimensionaler affiner Raum über \mathbb{R} , sei $Q \in \mathbb{A}$ und W ein affiner Unterraum von \mathbb{A} mit Tangentialraum U .

$$W = \{Q \boxplus u \mid u \in U\}.$$

Dann sind folgende Aussagen äquivalent:

- (i) $Q \in W$.
- (ii) $W = \{Q \boxplus u \mid u \in U\}$.

BEWEIS. Der Kürze halber bezeichnen wir $Y = \{Q \boxplus u \mid u \in U\}$. Aussage (ii) sagt dann $W = Y$. Wir beweisen eine Äquivalenz zweier Aussagen. Daher zeigen wir, dass aus (i) die Aussage (ii) folgt, und aus (ii) die Aussage (i) folgt.

- (i) \Rightarrow (ii) Sei $Q \in W$. Dann gibt es also ein $u \in U$ sodass $Q = P \boxplus u$. Wir müssen zeigen: $Y = W$. Sei $R \in Y$, also $R = Q \boxplus v$ mit einem $v \in U$. Weil U ein Unterraum ist, gilt auch $u + v \in U$. Dann ist aber

$$R = Q \boxplus v = (P \boxplus u) \boxplus v = P \boxplus (u + v).$$

Nach Definition von W gilt also $R \in W$. Da dies für alle $R \in Y$ gilt, folgt $Y \subseteq W$.

Ist umgekehrt S ein Punkt aus W , so gibt es ein $w \in U$ mit

$$S = P \boxplus w = P \boxplus (u + (w - u)) = Q \boxplus (w - u) \in Y.$$

Also ist $W \subseteq Y$.

- (ii) \Rightarrow (i) Wegen (ii) ist für alle $u \in U$ der Punkt $Q \boxplus u \in W$. Dann ist insbesondere $Q = Q \boxplus 0 \in W$.

□

Im dreidimensionalen affinen Raum können sich 2 Ebenen gar nicht, oder als Gerade, oder als Ebene schneiden. Allgemeiner gilt die folgende Behauptung:

BEHAUPTUNG 3.3.11. Sei \mathbb{A} ein n -dimensionaler affiner Raum über \mathbb{R} , seien W, Y zwei affine Unterräume von \mathbb{A} mit Tangentialräumen U bzw. $V \subset \mathbb{R}^n$. Dann ist entweder $W \cap Y = \emptyset$ oder $W \cap Y$ ist ein affiner Unterraum von \mathbb{A} mit Tangentialraum $U \cap V$.

BEWEIS. Wenn $W \cap Y$ leer ist, ist die Behauptung bewiesen. Wir sehen noch den Fall an, dass $W \cap Y$ mindestens einen Punkt P enthält. Es ist dann

$$W = \{P \boxplus u \mid u \in U\} \text{ und } Y = \{P \boxplus v \mid v \in V\}.$$

Ist $Q \in W \cap Y$, so lässt sich Q also schreiben $Q = P \boxplus u$, wobei $u \in U$ und $u \in V$. Es ist also

$$Q \in \{P \boxplus u \mid u \in U \cap V\}.$$

Ist andererseits $Q = P \boxplus u$ mit $u \in U \cap V$, so ist $Q \in W$ weil $u \in U$ und $Q \in Y$ weil $u \in V$, also ist $u \in W \cap Y$. \square

4. Basis und Dimension

4.1. Lineare Abhängigkeit.

Wir zielen jetzt auf Erzeugendensysteme, die einen Unterraum U des \mathbb{R}^n aufspannen, und in denen kein Vektor überflüssig ist. Zunächst sehen wir uns überzählige Vektoren an:

BEHAUPTUNG 4.1.1. Sei (v_1, \dots, v_k) ein k -Tupel von Vektoren in \mathbb{R}^n , und i ein Index zwischen 1 und k . Die folgenden Aussagen sind äquivalent:

- (1) Die lineare Hülle des gesamten Vektorsystems ist zugleich die lineare Hülle des Teilsystems ohne v_i :

$$\mathcal{L}(v_1, \dots, v_i, \dots, v_k) = \mathcal{L}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k).$$

- (2) v_i lässt sich als Linearkombination der anderen Vektoren schreiben:

$$v_i \in \mathcal{L}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k).$$

- (3) Es gibt Skalare $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ sodass $\lambda_i \neq 0$ und $\lambda_1 v_1 + \dots + \lambda_k v_k = 0$.

BEWEIS.

- (1) \Rightarrow (2): Weil $v_i \in \mathcal{L}(v_1, \dots, v_i, \dots, v_k)$ folgt aus (1) dass

$$v_i \in \mathcal{L}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k).$$

Das heißt aber, dass v_i eine Linearkombination des übrigen Teilsystems ist.

- (2) \Rightarrow (3): Sei $v_i = \sum_{j \neq i} \mu_j v_j$. Dann ist

$$\mu_1 v_1 + \dots + \mu_{i-1} v_{i-1} - 1 v_i + \mu_{i+1} v_{i+1} + \dots + \mu_k v_k = 0.$$

Setze also

$$\lambda_j = \begin{cases} \mu_j & \text{für } j \neq i \\ -1 \neq 0 & \text{für } j = i \end{cases}.$$

- (3) \Rightarrow (2): Sei $\sum_{j=1}^k \lambda_j v_j = 0$ und $\lambda_i \neq 0$. Es ist dann

$$\lambda_i v_i = - \sum_{j \neq i} \lambda_j v_j, \text{ also } v_i = \sum_{j \neq i} \frac{-\lambda_j}{\lambda_i} v_j.$$

(2) \Rightarrow (1): Sei $v_i = \sum_{j \neq i} \lambda_j v_j$. Offensichtlich ist

$$\mathcal{L}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k) \subseteq \mathcal{L}(v_1, \dots, v_i, \dots, v_k),$$

weil jede Linearkombination $\sum_{j \neq i} \mu_j v_j$ sich schreiben lässt $\sum_{j=1}^k \mu_j v_j$ mit $\mu_i = 0$.

Sei umgekehrt $x \in \mathcal{L}(v_1, \dots, v_i, \dots, v_k)$. Dann ist für geeignete μ_j

$$\begin{aligned} x &= \sum_{j=1}^k \mu_j v_j = \sum_{j \neq i} \mu_j v_j + \mu_i v_i = \sum_{j \neq i} \mu_j v_j + \mu_i \sum_{j \neq i} \lambda_j v_j \\ &= \sum_{j \neq i} (\mu_j + \mu_i \lambda_j) v_j \in \mathcal{L}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k). \end{aligned}$$

□

Auf Grund von Behauptung 4.1.1 charakterisieren wir ein Vektorsystem ohne "überflüssigen" Vektor folgendermaßen:

DEFINITION 4.1.2. Seien v_1, \dots, v_k Vektoren aus \mathbb{R}^n . Das k -Tupel (v_1, \dots, v_k) heißt linear unabhängig, wenn es kein k -Tupel von reellen Zahlen $\lambda_1, \dots, \lambda_k$ außer $0, \dots, 0$ gibt, für welches gilt:

$$\lambda_1 v_1 + \dots + \lambda_k v_k = 0.$$

Andernfalls heißt das Vektoren- k -Tupel (v_1, \dots, v_k) linear abhängig. Die leere Menge betrachten wir immer als linear unabhängig.

Wir versuchen einige äquivalente Formulierungen:

(v_1, \dots, v_k) ist linear unabhängig, wenn gilt:

$$(\forall \lambda_1, \dots, \lambda_k \in \mathbb{K}) \left[\sum_{i=1}^k \lambda_i v_i = 0 \Rightarrow (\lambda_1, \dots, \lambda_k) = (0, \dots, 0) \right].$$

(v_1, \dots, v_k) ist linear abhängig, wenn gilt:

$$(\exists \lambda_1, \dots, \lambda_k \in \mathbb{K}) \left[[(\exists i \in \{1, \dots, k\}) \lambda_i \neq 0] \wedge \left[\sum_{i=1}^k \lambda_i v_i = 0 \right] \right].$$

BEHAUPTUNG 4.1.3.

- (1) Vektoren- k -Tupel, welche den Nullvektor enthalten, sind linear abhängig.
- (2) k -Tupel, in denen ein Vektor mehrmals vorkommt, sind linear abhängig.
- (3) Enthält ein k -Tupel (v_1, \dots, v_k) ein linear abhängiges System⁸ $(v_{i_1}, \dots, v_{i_r})$ mit paarweise verschiedenen⁹ Indizes zwischen 1 und k , dann ist auch (v_1, \dots, v_k) linear abhängig.

BEWEIS.

- (1) Im k -Tupel (v_1, \dots, v_k) sei $v_j = 0$ für einen Index j . Wir setzen

$$\lambda_i = \begin{cases} 0 & \text{falls } i \neq j, \\ 1 & \text{falls } i = j. \end{cases}$$

Dann ist

$$\sum_{i=1}^k \lambda_i v_i = 1 \cdot v_j = 0,$$

aber nicht alle λ_i sind Null. Also ist (v_1, \dots, v_k) linear abhängig.

⁸Die Doppelindex-Schreibweise ist halb so schlimm wie sie aussieht. Betrachten wir zum Beispiel das Quintupel (v_1, \dots, v_5) und das Untersystem (v_2, v_4, v_1) . Dann ist i_1 der erste Index, der im Tripel vorkommt: $i_1 = 2$, $v_{i_1} = v_2$. Es ist i_2 der zweite Index, der im Tripel vorkommt: $i_2 = 4$, $v_{i_2} = v_4$. Letztlich ist $i_3 = 1$, $v_{i_3} = v_1$. Insgesamt ist $(v_{i_1}, v_{i_2}, v_{i_3}) = (v_2, v_4, v_1)$.

⁹Das heißt: keine zwei Indizes i_p, i_q sind gleich.

- (2) Sei nun (v_1, \dots, v_k) ein n -Tupel, in dem für zwei Indices $p \neq q$ gilt: $v_p = v_q$. Wir haben

$$0v_1 + \dots + 1v_p + \dots + (-1)v_q + \dots + 0v_k = 0,$$

aber nicht alle Koeffizienten dieser Linearkombination sind Null. Also ist das k -Tupel linear abhängig.

- (3) Sei $\sum_{j=1}^l \lambda_j v_{i_j} = 0$ und nicht alle λ_j gleich Null. Wir setzen

$$\mu_i := \begin{cases} \lambda_{i_j} & \text{falls } i = i_j \\ 0 & \text{falls } i \notin \{i_1, \dots, i_r\} \end{cases}.$$

Dann ist $\mu_{i_j} \neq 0$ für alle jene j für die $\lambda_j \neq 0$ und

$$\sum_{i=1}^k \mu_i v_i = \sum_{j=1}^r \lambda_j v_{i_j} = 0.$$

□

SATZ 4.1.4. Seien $v_1, \dots, v_n \in \mathbb{R}^n$. Dann sind folgende Aussagen äquivalent:

- (1) Das k -Tupel (v_1, \dots, v_k) ist linear unabhängig.
- (2) Für jedes $x \in \mathcal{L}(\{v_1, \dots, v_k\})$ gibt es nur ein eindeutiges k -Tupel $(\lambda_1, \dots, \lambda_k)$ von reellen Zahlen, sodass $x = \sum_{i=1}^k \lambda_i v_i$.
- (3) Kein v_i aus dem k -Tupel lässt sich als Linearkombination der anderen Vektoren des k -Tupels schreiben.
- (4) Wird einer der Vektoren v_i aus dem System weggelassen, so verkleinert sich die lineare Hülle:

$$\mathcal{L}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k) \subsetneq \mathcal{L}(v_1, \dots, v_i, \dots, v_k).$$

BEWEIS.

- (1) \Rightarrow (2): Sei (v_1, \dots, v_k) linear unabhängig. Sei

$$x = \sum_{i=1}^k \lambda_i v_i = \sum_{i=1}^k \mu_i v_i.$$

Dann ist

$$\sum_{i=1}^k (\lambda_i - \mu_i) v_i = \sum_{i=1}^k \lambda_i v_i - \sum_{i=1}^k \mu_i v_i = x - x = 0.$$

Wegen der linearen Unabhängigkeit von (v_1, \dots, v_k) folgt für alle i dass $\lambda_i - \mu_i = 0$, also $\lambda_i = \mu_i$.

- (2) \Rightarrow (3): Angenommen, es gilt nicht (3) und für einen Index i gilt also

$$v_i = \lambda_1 v_1 + \dots + \lambda_{i-1} v_{i-1} + \lambda_{i+1} v_{i+1} + \dots + \lambda_k v_k.$$

Andererseits ist natürlich

$$v_i = 0v_1 + \dots + 1v_i + \dots + 0v_k.$$

Das sind zwei verschiedene Arten, v_i als Linearkombination der v_1, \dots, v_k anzusetzen, aber die kann es wegen (2) nicht geben.

- (3) \Rightarrow (4): Wegen (3) gilt für jeden Index i

$$v_i \notin \mathcal{L}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k).$$

Weil aber natürlich

$$v_i \in \mathcal{L}(v_1, \dots, v_i, \dots, v_k),$$

sind die beiden linearen Hüllen verschieden:

$$\mathcal{L}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k) \neq \mathcal{L}(v_1, \dots, v_i, \dots, v_k).$$

(4) \Rightarrow (1): Siehe Behauptung 4.1.1.

□

BEMERKUNG 4.1.5. Ob ein k -Tupel von Vektoren

$$(v_1, \dots, v_k) = \left(\begin{pmatrix} \alpha_{11} \\ \vdots \\ \alpha_{n1} \end{pmatrix}, \dots, \begin{pmatrix} \alpha_{1k} \\ \vdots \\ \alpha_{nk} \end{pmatrix} \right)$$

linear unabhängig ist, überprüft man durch den Ansatz

$$\lambda_1 v_1 + \dots + \lambda_k v_k = 0.$$

Ausgeschrieben führt dieser Ansatz auf das lineare Gleichungssystem

$$\begin{array}{ccccccc} \alpha_{11}\lambda_1 & + & \dots & + & \alpha_{1k}\lambda_k & = & 0 \\ \vdots & & & & \vdots & & \vdots \\ \alpha_{n1}\lambda_1 & + & \dots & + & \alpha_{nk}\lambda_k & = & 0 \end{array}$$

Wenn dieses Gleichungssystem neben der Null-Lösung noch andere Lösungen besitzt, ist das Vektorsystem (v_1, \dots, v_k) linear abhängig.

BEISPIEL 4.1.6. Ist das folgende System (v_1, v_2, v_3) linear abhängig?

$$v_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix}, \quad v_3 = \begin{pmatrix} -1 \\ 10 \\ 4 \end{pmatrix}.$$

Lösung: Wir setzen das Gleichungssystem

$$\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = 0$$

in Matrixform an und beginnen den Eliminationsalgorithmus:

$$\begin{pmatrix} 2 & 1 & -1 & 0 \\ 1 & 4 & 10 & 0 \\ 1 & 2 & 4 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & -3 & -9 & 0 \\ 0 & 2 & 6 & 0 \\ 1 & 2 & 4 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 2 & 4 & 0 \end{pmatrix}$$

Wir können hier zu rechnen aufhören!¹⁰ An dieser Stelle sehen wir, dass sich nur 2 Pivotelemente finden lassen, und die dritte Variable nichtbasisch bleibt. Daher kann sie frei gewählt werden, z.B. $\lambda_3 = 1$, und das System hat außer der Null-Lösung noch weitere Lösungen. Das System (v_1, v_2, v_3) ist linear abhängig. □

4.2. Basis.

Wir können jetzt definieren, was ein Erzeugendensystem ohne überflüssige Vektoren ist:

DEFINITION 4.2.1. Sei $U \neq \{0\}$ ein Unterraum von \mathbb{R}^n . Seien $v_1, \dots, v_k \in U$. Das System (v_1, \dots, v_k) heißt eine Basis von U , wenn folgende zwei Bedingungen erfüllt sind:

- 1) Das System (v_1, \dots, v_k) ist linear unabhängig.
- 2) Die Lineare Hülle des Systems ist $\mathcal{L}(v_1, \dots, v_k) = U$.

Die leere Menge ist per definitionem die einzige Basis des Nullraumes $\{0\}$.

SATZ 4.2.2. Sei $U \neq \{0\}$ ein Unterraum von \mathbb{R}^n . Seien $v_1, \dots, v_k \in U$. Folgende Aussagen sind äquivalent:

- 1) (v_1, \dots, v_k) ist eine Basis von U .

¹⁰Faulheit ist die Triebfeder der mathematischen Kreativität. Rechnen Sie nie weiter, als Sie müssen, um Ihre Frage zu beantworten.

- 2) (v_1, \dots, v_k) ist ein maximales linear unabhängiges System in U in folgendem Sinn: (v_1, \dots, v_k) ist linear unabhängig, und für jeden Vektor $w \in U$ ist das System (v_1, \dots, v_k, w) linear abhängig.
- 3) Für jeden Vektor $x \in U$ existieren eindeutig bestimmte $\kappa_1, \dots, \kappa_k \in \mathbb{R}$ so¹¹, dass

$$x = \kappa_1 v_1 + \dots + \kappa_k v_k.$$

- 4) (v_1, \dots, v_k) ist ein minimales Erzeugendensystem von U in folgendem Sinne: Das System (v_1, \dots, v_k) spannt den Vektorraum U auf, aber für keinen Index i erzeugt das System $(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k)$ ganz U .

BEWEIS. (1) \Rightarrow (2): Nach Definition einer Basis ist (v_1, \dots, v_k) linear unabhängig. Weil $\mathcal{L}(\{v_1, \dots, v_n\}) = U$, gibt es für jedes $w \in U$ Skalare $\lambda_1, \dots, \lambda_k$ sodass

$$w = \lambda_1 v_1 + \dots + \lambda_k v_k.$$

Wegen Satz 4.1.4 ist das System (v_1, \dots, v_n, w) linear abhängig.

(2) \Rightarrow (3): Es gelte (2). Sei $x \in U$. Dann ist das System (v_1, \dots, v_k, x) linear abhängig, also gibt es $\lambda_1, \dots, \lambda_n, \mu \in \mathbb{R}$ sodass

$$\lambda_1 v_1 + \dots + \lambda_k v_k + \mu x = 0,$$

wobei mindestens einer der Skalare $\lambda_1, \dots, \lambda_k, \mu$ ungleich null ist. Wäre $\mu = 0$, so wäre

$$\lambda_1 v_1 + \dots + \lambda_k v_k = 0$$

und mindestens eines der λ_i ungleich null. Das widerspricht aber der linearen Unabhängigkeit von (v_1, \dots, v_k) . Also ist $\mu \neq 0$. Wir erhalten

$$x = -\frac{\lambda_1}{\mu} v_1 - \dots - \frac{\lambda_k}{\mu} v_k.$$

Also lässt sich x als Linearkombination der v_i schreiben. Da (v_1, \dots, v_k) linear unabhängig sind, ist diese Linearkombination nach Satz 4.1.4 eindeutig bestimmt.

(3) \Rightarrow (4): Es gelte (3). Da sich jedes $x \in U$ als Linearkombination der v_i schreiben lässt, ist $U = \mathcal{L}(v_1, \dots, v_k)$. Weil nach (3) jedes $x \in U$ sich auf genau eine Weise als Linearkombination der v_i schreiben lässt, ist das System (v_1, \dots, v_k) nach Satz 4.1.4 linear unabhängig, und kein v_i kann weggelassen werden, ohne dass die lineare Hülle des restlichen Systems eine echte Teilmenge von U wird. Also ist (v_1, \dots, v_k) ein minimales Erzeugendensystem.

(4) \Rightarrow (1): Es gelte (4). Es ist also (v_1, \dots, v_k) ein Erzeugendensystem. Andererseits ist für jedes i

$$\mathcal{L}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n) \neq U = \mathcal{L}(v_1, \dots, v_n).$$

Nach Satz 4.1.4 folgt die lineare Unabhängigkeit. \square

SATZ 4.2.3 (Basisauswahlsatz). Sei $U = \mathcal{L}(v_1, \dots, v_k) \neq \{0\}$ ein Unterraum des \mathbb{R}^n . Dann lässt sich aus (v_1, \dots, v_k) eine Basis $(v_{i_1}, \dots, v_{i_r})$ von U auswählen.

BEWEIS. Wenn (v_1, \dots, v_k) nicht selbst eine Basis von U ist, dann ist das System linear abhängig, und es lässt sich ein v_i weglassen, sodass nach wie vor $\mathcal{L}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k) = U$ ist. Wenn nun $(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_k)$ linear unabhängig ist, haben wir eine Basis, sonst können wir wieder einen Vektor weglassen, solange bis das Verfahren bei einer Basis abbricht. Da es nur endlich viele Vektoren zum Weglassen gibt, muss das Verfahren einmal enden. \square

¹¹Lies: κ : kappa.

4.3. Basiserweiterung.

LEMMA 4.3.1. ¹² Sei $U = \mathcal{L}(v_1, \dots, v_k)$ ein Unterraum des \mathbb{R}^n , und (w_1, \dots, w_r) ein linear unabhängiges Vektorsystem in U . Dann ist $r \leq k$.

BEWEIS. Da $w_j \in U$, gibt es Skalare $\alpha_{1j}, \dots, \alpha_{kj}$ sodass

$$w_j = \sum_{i=1}^k \alpha_{ij} v_i.$$

Wir betrachten das Gleichungssystem

$$\begin{array}{ccccccc} \alpha_{11}\lambda_1 & + & \cdots & + & \alpha_{1r}\lambda_r & = & 0 \\ \vdots & & & & \vdots & & \vdots \\ \alpha_{k1}\lambda_1 & + & \cdots & + & \alpha_{kr}\lambda_r & = & 0 \end{array}$$

Für jede Lösung $\lambda_1, \dots, \lambda_r$ dieses Gleichungssystems gilt

$$\sum_{j=1}^r \lambda_j w_j = \sum_{j=1}^r \sum_{i=1}^k \lambda_j \alpha_{ij} v_i = \sum_{i=1}^k \left(\sum_{j=1}^r \alpha_{ij} \lambda_j \right) v_i = \sum_{i=1}^k 0 v_i = 0.$$

Weil (w_1, \dots, w_r) ein linear unabhängiges System ist, ist aber $\sum_{j=1}^r \lambda_j w_j = 0$ nur dann, wenn alle λ_j gleich null sind. Daher kann das Gleichungssystem nur die Null-Lösung haben.

Wenn nun $k < r$ ist, können bei der Lösung des Gleichungssystems höchstens k Pivotelemente vorkommen, und mindestens $r - k \geq 1$ Variablen bleiben nichtbasisch. Dann hat das Gleichungssystem aber auch andere Lösungen als Null, denn die nichtbasischen Variablen können ja frei gewählt werden. Also muss gelten $r \leq k$. \square

KOROLLAR 4.3.2. ¹³ Ist (v_1, \dots, v_r) ein linear unabhängiges Vektorsystem in \mathbb{R}^n , so ist $r \leq n$.

BEWEIS. \mathbb{R}^n wird von den n Einheitsvektoren

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

aufgespannt. \square

SATZ 4.3.3 (Basisergänzungssatz). Sei U ein Unterraum von \mathbb{R}^n und (v_1, \dots, v_k) ein linear unabhängiges Vektorsystem in U . Dann ist entweder (v_1, \dots, v_k) selbst eine Basis von U , oder es lässt sich mit Vektoren w_1, \dots, w_r zu einer Basis

$$(v_1, \dots, v_k, w_1, \dots, w_r)$$

von U erweitern.

¹²Ein Lemma (griechisch: Weg) ist eine Behauptung, die man aufstellt und beweist, um mit ihrer Hilfe den Beweis einer weiteren Aussage zu erreichen.

¹³Ein Korollar zu einer Aussage ist eine unmittelbare Folgerung aus dieser Aussage.

BEWEIS. Ist (v_1, \dots, v_k) keine Basis von U , so ist es nicht maximal linear unabhängig (nach Satz 4.2.2), und daher gibt es ein $w_1 \in U$, sodass (v_1, \dots, v_k, w_1) auch noch linear unabhängig ist. Nun ist entweder (v_1, \dots, v_k, w_1) maximal linear unabhängig und damit eine Basis, oder das System kann um einen weiteren Vektor w_2 ergänzt werden. So entstehen immer größere linear unabhängige Vektorsysteme. Weil aber ein linear unabhängiges System wegen Korollar 4.3.2 höchstens n Vektoren enthalten kann, muss das Verfahren abbrechen, und für ein geeignetes r ist $(v_1, \dots, v_k, w_1, \dots, w_r)$ eine Basis. \square

Jetzt können wir die Frage beantworten, ob wirklich jeder Unterraum von \mathbb{R}^n von endlich vielen Vektoren aufgespannt wird.

SATZ 4.3.4. *Jeder Unterraum des \mathbb{R}^n besitzt mindestens eine Basis.*

BEWEIS. Ist $U = \{0\}$ der Nullraum, so besitzt U die leere Menge als Basis. Andernfalls gibt es ein $v_1 \neq 0$ in U . Das Vektorsystem (v_1) ist linear unabhängig und lässt sich zu einer Basis von U ausdehnen. \square

Zwar hat ein Unterraum des \mathbb{R}^n , wenn er nicht der Nullraum ist, viele verschiedene Basen. Es enthalten aber alle Basen gleichviele Vektoren.

SATZ 4.3.5. *Sei U ein Unterraum von \mathbb{R}^n . Dann enthalten alle Basen von U dieselbe Anzahl von Elementen.*

BEWEIS. Seien (v_1, \dots, v_k) und (w_1, \dots, w_r) zwei Basen von U . Wir zeigen: $k = r$. Der Raum U wird von (v_1, \dots, v_k) erzeugt, und (w_1, \dots, w_r) ist ein linear unabhängiges Vektorsystem in U . Nach Lemma 4.3.1 ist also $r \leq k$. Nun nützen wir, dass (v_1, \dots, v_k) linear unabhängig und U von (w_1, \dots, w_r) aufgespannt wird, also ist wieder nach Lemma 4.3.1 $k \leq r$. Damit ist also $k = r$. \square

Diese Einsicht rechtfertigt die folgende Definition:

DEFINITION 4.3.6. Sei U ein Unterraum von \mathbb{R}^n und (v_1, \dots, v_k) eine Basis von U . Dann heißt k die Dimension von U ($k = \dim(U)$).

Zum Beispiel ist die Dimension des Nullraumes $\dim(\{0\}) = 0$, die Dimension des \mathbb{R}^n ist $\dim(\mathbb{R}^n) = n$. Der Tangentialraum einer Ebene hat die Dimension 2, der Tangentialraum einer Gerade nur die Dimension 1.

SATZ 4.3.7. *Seien U, W zwei Unterräume des \mathbb{R}^n , wobei $U \subseteq W$. Dann gilt*

- (1) $\dim(U) \leq \dim(W)$.
- (2) *Ist $\dim(U) = \dim(W)$, dann ist $U = W$.*

BEWEIS. Sei (v_1, \dots, v_k) eine Basis von U , also $\dim(U) = k$. Nach Satz 4.3.3 ist entweder (v_1, \dots, v_k) auch eine Basis von W , dann ist $\dim(W) = k$ und $W = \mathcal{L}(v_1, \dots, v_k) = U$, oder es gibt eine Basis $(v_1, \dots, v_k, w_1, \dots, w_r)$ von W mit $r \geq 1$. In diesem Fall ist

$$\dim(W) = k + r > k = \dim(U).$$

\square

5. Matrix mal Vektor

5.1. Funktionen.

Im kommenden Abschnitt wird viel von linearen Funktionen die Rede sein. Bevor wir uns auf diesen Spezialfall von Funktionen einschränken, brauchen wir ein wenig Wissen über Funktionen im Allgemeinen.

Wir beschreiben den Begriff der Funktion zunächst auf eine Weise, die der Anschauung gut entspricht. Wir werden gleich eine saubere Definition nachtragen.

VEREINBARUNG 5.1.1. Seien A und B zwei nichtleere Mengen. Eine Funktion (= Abbildung) von A nach B ist eine Vorschrift, die zu jedem $a \in A$ genau ein $b \in B$ bestimmt.

Mit der Notation $f : A \rightarrow B$ drücken wir aus, dass f eine Funktion von A nach B ist. Mit $f : a \mapsto b$ oder mit $f(a) = b$ (“ $f(a)$ ist b ”, “ f bildet a auf b ab”) drücken wir aus, dass f einem bestimmten Element $a \in A$ das Element $b \in B$ zuweist.

Die Menge A heißt der Definitionsbereich von f , die Menge B heißt Wertevorrat von f .

Das ist wieder eine sehr verwaschene Begriffserklärung: “Eine Vorschrift, die ...” — Für die Anschauung verständlich, aber mathematisch keine Definition. Wir wissen, dass der Begriff der Menge schon diffizil und gefährlich ist. Werden wir bei der Definition von Funktionen wieder Grundlagenprobleme bekommen, kann man “Vorschriften” definieren, die zwingend auf Widersprüche führen? Die gute Nachricht ist: Hat man einmal den Mengenbegriff auf die sichere Seite gerettet, kann man eine Funktion präzise definieren. Wir definieren die Funktion einfach als die Menge der Paare $(a, f(a))$:

DEFINITION 5.1.2. Seien A und B nichtleere Mengen. Eine Funktion f ist eine Teilmenge der Produktmenge $A \times B$ mit der Eigenschaft:

Für jedes $a \in A$ gibt es genau ein $b \in B$, sodass das Paar (a, b) in f liegt.

Sind $a \in A$, $b \in B$, mit $(a, b) \in f$, so schreiben wir auch $b = f(a)$.

Will man speziell auf die Definition der Funktion durch Paare hinweisen, nennt man die Menge $\{(a, f(a)) \mid a \in A\}$ auch Graph der Funktion.

Für unsere praktische Arbeit genügt aber die anschauliche Vorstellung aus Vereinbarung 5.1.1.

BEMERKUNG 5.1.3. Seien A und B Mengen, und $f : A \rightarrow B$ eine Funktion. Beachten Sie, dass f nicht den vollen Wertevorrat B ausschöpfen muss, es darf “überflüssige” Elemente in B geben. Dagegen muss sich f auf jedes Element in A anwenden lassen.

BEMERKUNG 5.1.4. Seien A und B Mengen, und $f : A \rightarrow B$ eine Funktion. Sei $a \in A$. Unterscheiden Sie genau:

f ist eine Funktion, also axiomatisch eine Menge geordneter Paare, anschaulich eine Abbildungsvorschrift.

$f(a)$ ist ein Element von B , jenes Element, auf das a abgebildet wird.

BEISPIEL 5.1.5. Hier finden Sie einige Beispiele von Funktionen:

- (1) Jeder/m Studierenden wird genau eine Matrikelnummer zugeordnet, niemand von ihnen hat keine oder mehrere Nummern. Wir haben daher eine Funktion von der Menge der Studierenden auf die 7-stelligen natürlichen Zahlen. Umgekehrt findet sich zu jeder Matrikelnummer ein/e Studierende/r. Trotzdem ist das keine Funktion von der Menge der 7-stelligen natürlichen Zahlen auf die Menge der Studierenden, weil es Zahlen gibt, die nicht als Matrikelnummer vorkommen.
- (2) Die Abbildung $x \mapsto \sqrt{x}$ ist eine Funktion von $[0, \infty)$ nach $[0, \infty)$.

- (3) Die Abbildung $x \mapsto x^2$ ist eine Abbildung von \mathbb{R} nach \mathbb{R} , denn jede reelle Zahl hat ihr eindeutiges Quadrat. Es macht nichts, dass der Wertevorrat nicht ausgeschöpft wird (negative Zahlen sind keine Quadrate), und auch nichts, dass jede positive Zahl zweimal als Wert (nämlich als Quadrat einer positiven und einer negativen Zahl) vorkommt.
- (4) Seien $\alpha_1, \dots, \alpha_n \in \mathbb{R}$. Die Formel

$$f : \begin{cases} \mathbb{R}^n & \rightarrow \mathbb{R} \\ \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} & \mapsto \sum_{i=1}^n \alpha_i \xi_i \end{cases}$$

definiert eine Funktion.

Funktionen von Intervallen reeller Zahlen in Intervalle lassen sich grafisch als Kurven darstellen. Jedes Paar $(x, f(x))$ des Graphen von f wird als Koordinatenpaar eines Punktes der Ebene dargestellt. Der gesamte Graph ergibt bei (eingermaßen "gutmütigen") Funktionen eine Kurve. Funktionen von \mathbb{R}^2 nach \mathbb{R} kann man als Flächen im dreidimensionalen Raum darstellen.

Eine Funktion, die auf einer endlichen Menge A definiert ist, lässt sich auch als Tabelle schreiben. Eine Funktion einer endlichen Menge A nach \mathbb{R} kann man als Stabdiagramm darstellen. Funktionen von einer endlichen Menge A in eine andere endliche Menge B lässt sich als Pfeildiagramm schreiben, indem man die Elemente von A und von B als Punkte darstellt, und jeweils einen Pfeil von a nach $f(a)$ zieht.

DEFINITION 5.1.6. Seien A und B Mengen und $f : A \rightarrow B$ eine Funktion. Seien $U \subseteq A$ und $V \subseteq B$ Teilmengen (die auch leer sein dürfen). Wir definieren

- (1) Das Bild von U besteht aus allen Elementen von B , die als Werte $f(u)$ mit einem $u \in U$ vorkommen: $f(U) := \{f(u) \mid u \in U\}$.
- (2) Das Urbild von V besteht aus allen Elementen von A , die durch f in Elemente von V abgebildet werden: $f^{-1}(V) = \{a \in A \mid f(a) \in V\}$.

BEMERKUNG 5.1.7. Bitte beachten Sie unbedingt, dass es zur Definition des Urbildes einer Menge unter einer Funktion f nicht notwendig ist, dass f eine Umkehrfunktion besitzt, auch wenn es durch die Schreibweise suggeriert wird. Wenn die Funktion den Wertevorrat nicht voll ausschöpft, dann gibt es auch nichtleere Mengen, deren Urbild die leere Menge ist. Dagegen gibt es nie eine nichtleere Menge, deren Bild leer ist.

BEISPIEL 5.1.8.

- (1) Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch $f(x) = x^2$. Sei $0 < a < b < \infty$. Dann sind Bild und Urbild des Intervalles (a, b) gegeben durch

$$f((a, b)) = (a^2, b^2), \quad f^{-1}((a, b)) = (-\sqrt{b}, -\sqrt{a}) \cup (\sqrt{a}, \sqrt{b}).$$

- (2) Seien $\alpha_{i,j} \in \mathbb{R}$ für $i = 1, \dots, m, j = 1, \dots, n$. Wir betrachten die Funktion

$$f : \begin{cases} \mathbb{R}^n & \rightarrow \mathbb{R}^m \\ \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} & \mapsto \begin{pmatrix} \alpha_{1,1}\xi_1 + \dots + \alpha_{1,n}\xi_n \\ \vdots \\ \alpha_{m,1}\xi_1 + \dots + \alpha_{m,n}\xi_n \end{pmatrix} \end{cases}.$$

Das Urbild des Nullraumes ist die Lösungsmenge des homogenen linearen Gleichungssystems

$$\begin{array}{ccccccc} \alpha_{1,1}\xi_1 & + & \cdots & + & \alpha_{1,n}\xi_n & = & 0 \\ \vdots & & & & \vdots & & \vdots \\ \alpha_{m,1}\xi_1 & + & \cdots & + & \alpha_{m,n}\xi_n & = & 0 \end{array}$$

Die Lösungsmenge kann aus einem Vektor (dem Nullvektor) oder einem mehrdimensionalen Unterraum des \mathbb{R}^n bestehen.

Injektivität und Surjektivität sind zentrale Begriffe zur Beschreibung von Funktionen. Sie spielen auch in der linearen Algebra eine große Rolle:

DEFINITION 5.1.9. Seien A und B Mengen, und $f : A \rightarrow B$ eine Funktion. Wir sagen

- (1) Die Funktion $f : A \rightarrow B$ heißt surjektiv, wenn $f(A) = B$, d.h., wenn alle Elemente von B als Werte $f(a)$ mit geeigneten $a \in A$ vorkommen.
- (2) Die Funktion $f : A \rightarrow B$ heißt injektiv, wenn für alle $x \neq y \in A$ gilt: $f(x) \neq f(y)$, wenn also verschiedene Elemente aus A auf verschiedene Elemente aus B abgebildet werden.
- (3) Die Funktion $f : A \rightarrow B$ heißt bijektiv, wenn sie sowohl injektiv als auch surjektiv ist.

Die folgende Bemerkung sagt genau dasselbe mit anderen Worten:

BEMERKUNG 5.1.10. Seien A und B Mengen, und $f : A \rightarrow B$ eine Funktion. Es gilt:

- (1) $f : A \rightarrow B$ ist genau dann surjektiv, wenn für jedes $y \in B$ die Gleichung $f(x) = y$ mindestens eine Lösung $x \in A$ besitzt.
- (2) $f : A \rightarrow B$ ist genau dann injektiv, wenn für jedes $y \in B$ die Gleichung $f(x) = y$ höchstens eine Lösung $x \in A$ besitzt.
- (3) $f : A \rightarrow B$ ist genau dann bijektiv, wenn für jedes $y \in B$ die Gleichung $f(x) = y$ genau eine Lösung $x \in A$ besitzt.

BEISPIEL 5.1.11. Ist die folgende Funktion injektiv, surjektiv, bijektiv?

$$f : \begin{cases} [0, \infty) & \rightarrow [0, \infty) \\ x & \mapsto (x^3 - 2)^2 \end{cases} .$$

Lösung: Wir untersuchen, ob die Gleichung $f(x) = y$ für alle $y \in [0, \infty)$ mindestens, höchstens, oder genau eine Lösung hat.

$$\begin{aligned} (x^3 - 2)^2 &= y \\ (x^3 - 2) &= \pm\sqrt{y} \quad \text{geht immer, weil } y \in [0, \infty) \\ x^3 &= \pm\sqrt{y} + 2 \\ x &= \begin{cases} \sqrt[3]{\sqrt{y} + 2} & \text{geht immer} \\ \sqrt[3]{-\sqrt{y} + 2} & \text{nur wenn } y \leq 4, \text{ sonst ist } x \notin [0, \infty) \end{cases} \end{aligned}$$

Daher gibt es für jedes $y \in [0, \infty)$ mindestens eine Lösung x von $f(x) = y$. Die Abbildung ist also surjektiv. Aber es gibt Werte $y \in [0, \infty)$, für die die Lösung nicht eindeutig bestimmt ist. Daher ist f nicht injektiv und nicht bijektiv. \square

BEISPIEL 5.1.12. Ist die folgende Abbildung $g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ injektiv, surjektiv, bijektiv?

$$g\left(\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}\right) = \begin{pmatrix} 2\xi_1 + \xi_2 \\ \xi_1 - 3\xi_2 \end{pmatrix} .$$

Lösung: Wir untersuchen die Lösbarkeit von $g(x) = y$, d.h.¹⁴

$$\begin{array}{rcl} 2\xi_1 & + & \xi_2 = \eta_1 \\ \xi_1 & - & 3\xi_2 = \eta_2 \end{array}$$

$$\begin{pmatrix} 2 & 1 & \eta_1 \\ 1 & -3 & \eta_2 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 7 & \eta_1 - 2\eta_2 \\ 1 & -3 & \eta_2 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 1 & \frac{1}{7}\eta_1 - \frac{2}{7}\eta_2 \\ 1 & 0 & \frac{3}{7}\eta_1 + \frac{1}{7}\eta_2 \end{pmatrix}$$

Das Gleichungssystem ist eindeutig lösbar, die Abbildung g ist bijektiv. \square

5.2. Produkt einer Matrix mit einem Vektor.

Wir beginnen nun Matrizen als Objekte aufzufassen, mit denen man rechnen kann. Die folgende Definition sieht auf den ersten Blick recht künstlich aus und scheint schwer zu merken. Sie werden sich ganz schnell an sie gewöhnen. Sie ist einer der genialen Schachzüge in der linearen Algebra.

DEFINITION 5.2.1. Das Produkt einer $(m \times n)$ -Matrix über \mathbb{R} mit einem n -dimensionalen reellen Spaltenvektor ist definiert durch

$$\begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} \cdot \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \alpha_{1,1}\beta_1 + \cdots + \alpha_{1,n}\beta_n \\ \vdots \\ \alpha_{m,1}\beta_1 + \cdots + \alpha_{m,n}\beta_n \end{pmatrix}.$$

Wichtig:

- Die Matrix steht immer vor dem Vektor. In der linearen Algebra kommt der Reihenfolge eine wichtige Bedeutung zu!
- Das Produkt $A \cdot b$ einer Matrix A mit einem Vektor b kann nur gebildet werden, wenn A so viele Spalten hat, wie b Einträge hat. Das Resultat $A \cdot b$ ist dann ein Spaltenvektor mit so vielen Einträgen wie A Zeilen hat.

Wie bei der Multiplikation von Skalaren lassen wir auch bei der Multiplikation einer Matrix mit einem Vektor den Multiplikationspunkt normalerweise weg.

BEISPIEL 5.2.2. Wir berechnen

$$\begin{pmatrix} 2 & 5 & -3 & 1 \\ 4 & 3 & 2 & -6 \\ 1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 7 \\ 8 \\ 9 \\ 10 \end{pmatrix}.$$

Lösung:

Unsere erste Frage ist, ob die Multiplikation überhaupt erlaubt ist. Die Matrix hat 4 Spalten, und der Vektor hat 4 Zeilen, also können wir multiplizieren. Wir wissen jetzt auch, dass das Produkt ein dreidimensionaler Spaltenvektor sein wird, weil die Matrix 3 Zeilen hat.

Wir berechnen das oberste Element. Achten Sie jetzt nur auf die erste Zeile der Matrix und den Spaltenvektor. Man tut sich leicht, wenn man mit dem Zeigefinger die erste Zeile, und mit dem Rechenstift den Vektor entlang fährt:

$$\begin{pmatrix} 2 & 5 & -3 & 1 \\ & & & \\ & & & \\ & & & \end{pmatrix} \begin{pmatrix} 7 \\ 8 \\ 9 \\ 10 \end{pmatrix} = \begin{pmatrix} 2 \cdot 7 + 5 \cdot 8 + -3 \cdot 9 + 1 \cdot 10 \\ & & & \\ & & & \\ & & & \end{pmatrix} = \begin{pmatrix} 37 \\ & & & \\ & & & \\ & & & \end{pmatrix}.$$

Nun berechnen wir den zweiten Eintrag. Wir konzentrieren uns auf die zweite Zeile der Matrix und den Vektor:

$$\begin{pmatrix} 2 & 5 & -3 & 1 \\ 4 & 3 & 2 & -6 \\ & & & \\ & & & \end{pmatrix} \begin{pmatrix} 7 \\ 8 \\ 9 \\ 10 \end{pmatrix} = \begin{pmatrix} & & & \\ 4 \cdot 7 + 3 \cdot 8 + 2 \cdot 9 - 6 \cdot 10 & & & \\ & & & \\ & & & \end{pmatrix} = \begin{pmatrix} & & & \\ 37 & & & \\ & & & \\ & & & \end{pmatrix}.$$

¹⁴Lies η : eta

Jetzt fehlt nur mehr der dritte Eintrag:

$$\begin{pmatrix} 2 & 5 & -3 & 1 \\ 4 & 3 & 2 & -6 \\ 1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 7 \\ 8 \\ 9 \\ 10 \end{pmatrix} = \begin{pmatrix} 37 \\ 10 \\ 1 \cdot 7 + 0 \cdot 8 + 1 \cdot 9 - 0 \cdot 10 \end{pmatrix} = \begin{pmatrix} 37 \\ 10 \\ 16 \end{pmatrix}.$$

Das Produkt ist fertig berechnet:

$$\begin{pmatrix} 2 & 5 & -3 & 1 \\ 4 & 3 & 2 & -6 \\ 1 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 7 \\ 8 \\ 9 \\ 10 \end{pmatrix} = \begin{pmatrix} 37 \\ 10 \\ 16 \end{pmatrix}.$$

Das Produkt einer Matrix mit einem Vektor ist maßgeschneidert für die Formulierung linearer Gleichungssysteme. Wir werden davon später noch Gebrauch machen.

BEHAUPTUNG 5.2.3. *Das lineare Gleichungssystem*¹⁵

$$\begin{array}{ccccccc} \alpha_{1,1}\xi_1 & + & \cdots & + & \alpha_{1,n}\xi_n & = & \beta_1 \\ \vdots & & & & \vdots & & \vdots \\ \alpha_{m,1}\xi_1 & + & \cdots & + & \alpha_{m,n}\xi_n & = & \beta_m \end{array}$$

lässt sich mit Hilfe des Matrixproduktes in der folgenden Form schreiben:

$$\begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix}.$$

BEWEIS. Man muss nur ausmultiplizieren:

$$\begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} \alpha_{1,1}\xi_1 + \cdots + \alpha_{1,n}\xi_n \\ \vdots \\ \alpha_{m,1}\xi_1 + \cdots + \alpha_{m,n}\xi_n \end{pmatrix}.$$

□

DEFINITION 5.2.4. Die Matrix A eines linearen Gleichungssystems $Ax = b$ heißt die Koeffizientenmatrix oder Systemmatrix des Gleichungssystems.

Das Produkt einer Matrix mit einem Vektor lässt sich auch als Linearkombination der Spalten der Matrix interpretieren.

DEFINITION 5.2.5. Sei $A \in \mathbb{R}^{m \times n}$ eine Matrix

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix}.$$

(1) Die Spalten von A sind die Spaltenvektoren

$$\begin{pmatrix} \alpha_{1,1} \\ \vdots \\ \alpha_{m,1} \end{pmatrix}, \begin{pmatrix} \alpha_{1,2} \\ \vdots \\ \alpha_{m,2} \end{pmatrix}, \dots, \begin{pmatrix} \alpha_{1,n} \\ \vdots \\ \alpha_{m,n} \end{pmatrix} \in \mathbb{R}^m.$$

¹⁵Lies: ξ : xi, nicht verwechseln mit ζ : zeta.

(2) Die Zeilen von A sind die $(1 \times n)$ -Matrizen

$$\begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \cdots & \alpha_{2,n} \\ \vdots & & \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix}.$$

(3) Eine $(1 \times n)$ -Matrix nennt man auch einen n -dimensionalen Zeilenvektor.

BEHAUPTUNG 5.2.6. *Das Produkt einer Matrix $A \in \mathbb{R}^{m \times n}$ mit einem Spaltenvektor $x \in \mathbb{R}^n$ ist eine Linearkombination der Spalten von A .*

BEWEIS.

$$\begin{aligned} & \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} \alpha_{1,1}\xi_1 + \cdots + \alpha_{1,n}\xi_n \\ \vdots \\ \alpha_{m,1}\xi_1 + \cdots + \alpha_{m,n}\xi_n \end{pmatrix} \\ &= \begin{pmatrix} \alpha_{1,1}\xi_1 \\ \vdots \\ \alpha_{m,1}\xi_1 \end{pmatrix} + \cdots + \begin{pmatrix} \alpha_{1,n}\xi_n \\ \vdots \\ \alpha_{m,n}\xi_n \end{pmatrix} = \xi_1 \begin{pmatrix} \alpha_{1,1} \\ \vdots \\ \alpha_{m,1} \end{pmatrix} + \cdots + \xi_n \begin{pmatrix} \alpha_{1,n} \\ \vdots \\ \alpha_{m,n} \end{pmatrix}. \end{aligned}$$

□

5.3. Matrizen und lineare Abbildungen.

DEFINITION 5.3.1. Eine Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ heißt linear, wenn die folgenden beiden Bedingungen gelten:

- (1) Für alle Vektoren $x \in \mathbb{R}^n$ und alle Skalare $\lambda \in \mathbb{R}$ ist $f(\lambda x) = \lambda f(x)$.
- (2) Für alle Vektoren $x, y \in \mathbb{R}^n$ ist $f(x + y) = f(x) + f(y)$.

BEHAUPTUNG 5.3.2.

- (1) *Eine Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ist genau dann linear, wenn für alle $x_1, \dots, x_k \in \mathbb{R}^n$ und alle $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ gilt:*

$$(5.3.1) \quad f(\lambda_1 x_1 + \cdots + \lambda_k x_k) = \lambda_1 f(x_1) + \cdots + \lambda_k f(x_k).$$

- (2) *Jede lineare Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ bildet den Nullvektor aus \mathbb{R}^n auf den Nullvektor aus \mathbb{R}^m ab.*

BEWEIS. (1) Diese Behauptung ist ein schönes Beispiel für einen Induktionsbeweis. Sei also $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine lineare Abbildung.

Wenn f die Bedingung (5.3.1) erfüllt, gilt sie insbesondere auch für $k = 1$, also

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

und $k = 2$ im Sonderfall dass $\lambda_1 = \lambda_2 = 1$, also

$$f(x + y) = f(1x + 1y) = 1f(x) + 1f(y) = f(x) + f(y).$$

Die Gegenrichtung beweisen wir durch vollständige Induktion nach k . Es sei also f linear.

Induktionsverankerung für $k = 1$: Wegen der Definition der Linearität gilt

$$f(\lambda_1 x_1) = \lambda_1 f(x_1).$$

Induktionsschritt von k auf $k+1$: Es gelte die Induktionsannahme

$$f(\lambda_1 x_1 + \cdots + \lambda_k x_k) = \lambda_1 f(x_1) + \cdots + \lambda_k f(x_k).$$

Dann gilt

$$\begin{aligned} & f(\lambda_1 x_1 + \cdots + \lambda_k x_k + \lambda_{k+1} x_{k+1}) \\ = & f([\lambda_1 x_1 + \cdots + \lambda_k x_k] + [\lambda_{k+1} x_{k+1}]) \quad (\text{Summe in zwei Teile zerlegt}) \\ = & f(\lambda_1 x_1 + \cdots + \lambda_k x_k) + f(\lambda_{k+1} x_{k+1}) \quad (\text{weil } f \text{ linear ist}) \\ = & [\lambda_1 f(x_1) + \cdots + \lambda_k f(x_k)] + [\lambda_{k+1} f(x_{k+1})] \quad (\text{Ind.annahme und } f \text{ linear}) \\ = & \lambda_1 f(x_1) + \cdots + \lambda_k f(x_k) + \lambda_{k+1} f(x_{k+1}). \end{aligned}$$

(2) Sei f linear und Θ_n der Nullvektor¹⁶ in \mathbb{R}^n . Es ist

$$f(\Theta_n) = f(0\Theta_n) = 0f(\Theta_n) = \Theta_m.$$

□

Vielleicht die wichtigste Eigenschaft der Matrizen ist, dass zu jeder Matrix eine lineare Abbildung, und zu jeder linearen Abbildung ein Matrix gehört, wie die beiden folgenden Sätze zeigen:

SATZ 5.3.3. *Sei A eine $(m \times n)$ -Matrix über \mathbb{R} . Wir definieren die Abbildung*

$$f : \begin{cases} \mathbb{R}^n & \rightarrow \mathbb{R}^m \\ x & \mapsto Ax \end{cases}.$$

(1) *Die Abbildung f ist linear.*

(2) *Die Spalten der Matrix A sind die Bilder der Einheitsvektoren.*

BEWEIS. Es sei¹⁷

$$A = \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix}, \quad x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}, \quad y = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix}, \quad \lambda \in \mathbb{R}.$$

Es gilt:

$$\begin{aligned} A(\lambda x) &= \begin{pmatrix} \sum_{j=1}^n \alpha_{1,j}(\lambda \xi_j) \\ \vdots \\ \sum_{j=1}^n \alpha_{m,j}(\lambda \xi_j) \end{pmatrix} = \begin{pmatrix} \lambda \sum_{j=1}^n \alpha_{1,j} \xi_j \\ \vdots \\ \lambda \sum_{j=1}^n \alpha_{m,j} \xi_j \end{pmatrix} = \lambda(Ax), \\ A(x+y) &= \begin{pmatrix} \sum_{j=1}^n \alpha_{1,j}(\xi_j + \eta_j) \\ \vdots \\ \sum_{j=1}^n \alpha_{m,j}(\xi_j + \eta_j) \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n \alpha_{1,j} \xi_j + \sum_{j=1}^n \alpha_{1,j} \eta_j \\ \vdots \\ \sum_{j=1}^n \alpha_{m,j} \xi_j + \sum_{j=1}^n \alpha_{m,j} \eta_j \end{pmatrix} \\ &= \begin{pmatrix} \sum_{j=1}^n \alpha_{1,j} \xi_j \\ \vdots \\ \sum_{j=1}^n \alpha_{m,j} \xi_j \end{pmatrix} + \begin{pmatrix} \sum_{j=1}^n \alpha_{1,j} \eta_j \\ \vdots \\ \sum_{j=1}^n \alpha_{m,j} \eta_j \end{pmatrix} = Ax + Ay. \end{aligned}$$

¹⁶Lies: Θ : Theta. Normalerweise bezeichnen wir den Nullvektor ja nur mit 0, aber hier wollen wir den Skalar 0 und den Nullvektor auseinanderhalten.

¹⁷Lies: η : eta.

Daher ist f linear.

Letztlich ist das Bild des k -ten Einheitsvektors die k -te Spalte von A :

$$\begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{m,1} & \cdots & \alpha_{m,n} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \text{ (} k\text{-te Stelle)} \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_{1,k} \\ \vdots \\ \alpha_{m,k} \end{pmatrix}.$$

□

BEMERKUNG 5.3.4. Sei $A \in \mathbb{R}^{m \times n}$. Die Linearität der Abbildung $x \mapsto Ax$ bedeutet wegen Behauptung 5.3.2 nichts anderes, als dass für alle $x_1, \dots, x_k \in \mathbb{R}^n$ und $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ gilt:

$$A\left(\sum_{j=1}^k \lambda_j x_j\right) = \sum_{j=1}^k \lambda_j Ax_j.$$

SATZ 5.3.5. Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine lineare Abbildung. Dann gibt es genau eine Matrix $A \in \mathbb{R}^{m \times n}$, so dass für alle $x \in \mathbb{R}^n$ gilt: $f(x) = Ax$.

BEWEIS. Wie wir bereits aus Satz 5.3.3 wissen, sind die Spalten einer solchen Matrix A jedenfalls die Bilder der Einheitsvektoren. Daher kann es höchstens eine Matrix geben, die f darstellt. Seien nun die Spalten a_1, \dots, a_n von A die Bilder der

Einheitsvektoren, also $a_k = f(e_k)$, und sei $x = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \in \mathbb{R}^n$. Wir zeigen: $Ax = f(x)$.

Wegen Behauptung 5.2.6 ist

$$Ax = \sum_{j=1}^n \xi_j a_j = \sum_{j=1}^n \xi_j f(e_j).$$

Wir verwenden die Linearität von f und rechnen weiter:

$$Ax = f\left(\sum_{j=1}^n \xi_j e_j\right) = f(x).$$

□

In der Sprache der linearen Gleichungssysteme drückt sich die Linearität der Abbildung $x \mapsto Ax$ durch das Superpositionsprinzip aus:

SATZ 5.3.6 (Superpositionsprinzip). Sei $A \in \mathbb{R}^{m \times n}$ eine $(m \times n)$ -Matrix, seien $b, b_1, b_2 \in \mathbb{R}^m$, $x, x_1, x_2 \in \mathbb{R}^n$, und¹⁸ $\chi_1, \chi_2 \in \mathbb{R}$.

- (1) Die Lösungsmenge des homogenen linearen Gleichungssystems $Ax = 0$ ist ein Unterraum.
- (2) Ist x_1 eine partikuläre Lösung von $Ax_1 = b_1$ und x_2 eine partikuläre Lösung von $Ax_2 = b_2$, so ist $y = \chi_1 x_1 + \chi_2 x_2$ eine partikuläre Lösung von $Ay = (\chi_1 b_1 + \chi_2 b_2)$.
- (3) Besitzt das Gleichungssystem $Ax = b$ überhaupt eine partikuläre Lösung x , so ist die Lösungsmenge dieses Gleichungssystems der affine Unterraum $\{x + u \mid Au = 0\}$.

¹⁸Lies: χ : chi.

BEWEIS.

- (1) Sei $U = \{u \in \mathbb{R}^n \mid Au = 0\}$. Es ist $A0 = 0$, also ist $U \neq \emptyset$. Ist $u \in U$ und $\lambda \in \mathbb{R}$, so ist

$$A(\lambda u) = \lambda Au = 0, \quad \text{also } (\lambda u) \in U.$$

Sind $u_1, u_2 \in U$, so ist

$$A(u_1 + u_2) = Au_1 + Au_2 = 0 + 0 = 0, \quad \text{also } u_1 + u_2 \in U.$$

- (2) Sei $Ax_1 = b_1$ und $Ax_2 = b_2$. Dann ist, wegen der Linearität der Abbildung $x \mapsto Ax$,

$$A(\chi_1 x_1 + \chi_2 x_2) = \chi_1 Ax_1 + \chi_2 Ax_2 = \chi_1 b_1 + \chi_2 b_2.$$

- (3) Sei $y \in \mathbb{R}^n$ beliebig. Wir setzen $u = y - x$. Dann gilt¹⁹

$$Au = 0 \iff A(y - x) = 0 \iff Ay - Ax = 0 \iff Ay = b.$$

Also ist $Ay = b$ genau dann, wenn $y = x + u$ mit einem u , so dass $Au = 0$.

□

5.4. Kern, Spaltenraum und Zeilenraum.

Wichtig! Erinnern Sie sich bitte an die Definition von Bildern und Urbildern von Mengen: Ist $f : A \rightarrow B$ eine Abbildung, ist $U \subseteq A$ und $W \subseteq B$, so ist

- $f(U) = \{f(u) \mid u \in U\}$, die Menge aller Punkte in B , die Bilder von Punkten aus U sind.
- $f^{-1}(W) = \{x \in A \mid f(x) \in W\}$, die Menge aller Punkte in A , die nach W abgebildet werden. Die Existenz einer Umkehrabbildung von f wird in dieser Definition nicht vorausgesetzt.

SATZ 5.4.1. Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine lineare Abbildung. Sei U ein Unterraum von \mathbb{R}^n und W ein Unterraum von \mathbb{R}^m . Dann gilt:

- (1) $f(U)$ ist ein Unterraum von \mathbb{R}^m .
- (2) $f^{-1}(W)$ ist ein Unterraum von \mathbb{R}^n .

BEWEIS.

- (1) Weil U ein Unterraum von \mathbb{R}^n ist, enthält U den Nullvektor. Dann ist aber $f(0) = 0$ in $f(U)$ enthalten, also ist $f(U) \neq \emptyset$. Sind $v_1, v_2 \in f(U)$ und $\lambda_1, \lambda_2 \in \mathbb{R}$, so gibt es $u_1, u_2 \in U$ sodass $v_1 = f(u_1)$, $v_2 = f(u_2)$. Weil U ein Unterraum ist, ist auch $\lambda_1 u_1 + \lambda_2 u_2 \in U$, und daher ist

$$\lambda_1 v_1 + \lambda_2 v_2 = \lambda_1 f(u_1) + \lambda_2 f(u_2) = f(\lambda_1 u_1 + \lambda_2 u_2) \in f(U).$$

- (2) Weil W ein Unterraum ist, enthält W den Nullvektor aus \mathbb{R}^m . Da aber der Nullvektor aus \mathbb{R}^n auf den Nullvektor aus \mathbb{R}^m abgebildet wird, enthält $f^{-1}(W)$ den Nullvektor aus \mathbb{R}^n und ist also nicht leer. Seien $u_1, u_2 \in f^{-1}(W)$, $\lambda_1, \lambda_2 \in \mathbb{R}$. Dann sind also $f(u_1) \in W$, $f(u_2) \in W$. Weil W ein Unterraum ist, ist auch $\lambda_1 f(u_1) + \lambda_2 f(u_2) \in W$. Dann ist

$$f(\lambda_1 u_1 + \lambda_2 u_2) = \lambda_1 f(u_1) + \lambda_2 f(u_2) \in W,$$

und damit ist $\lambda_1 u_1 + \lambda_2 u_2 \in f^{-1}(W)$.

□

¹⁹Lies \iff : genau dann, wenn

DEFINITION 5.4.2. Sei $A \in \mathbb{R}^{m \times n}$ mit den Zeilen $z_1, \dots, z_m \in \mathbb{R}^{1 \times n}$ und den Spalten $s_1, \dots, s_n \in \mathbb{R}^m$. Wir definieren folgende Mengen:

$$\begin{aligned} \ker(A) &= \{x \in \mathbb{R}^n \mid Ax = 0\} \quad \text{Kern von } A, \\ \text{rg}(A) &= \mathcal{L}(s_1, \dots, s_n) \subseteq \mathbb{R}^m \quad \text{Spaltenraum oder Range von } A, \\ \text{zr}(A) &= \mathcal{L}(z_1, \dots, z_m) \subseteq \mathbb{R}^{1 \times n} \quad \text{Zeilenraum von } A. \end{aligned}$$

Wir schreiben auch, wenn unser Augenmerk vor allem der linearen Abbildung $f : x \mapsto Ax$ gilt, $\ker(f)$ statt $\ker(A)$ und $\text{rg}(f)$ statt $\text{rg}(A)$.

Verwechseln Sie nicht das englische Wort range (Wertebereich, hier: Spaltenraum) mit dem deutschen Wort Rang, englisch rank. Der Rang einer Matrix ist eine Zahl und wird im nächsten Abschnitt eingeführt. Auch gibt es in der englischsprachigen Literatur den Begriff "nullspace of A ". Damit ist nicht unser Nullraum gemeint, sondern der Kern von A .

BEHAUPTUNG 5.4.3. Sei $A \in \mathbb{R}^{m \times n}$ und

$$f : \begin{cases} \mathbb{R}^n & \rightarrow \mathbb{R}^m \\ x & \mapsto Ax \end{cases}$$

die dazu gehörende lineare Abbildung.

- (1) Der Kern von A ist zugleich das Urbild des Nullraums unter f , d.h., $\ker(A) = f^{-1}(\{0\})$.
- (2) Der Kern von A ist die Lösungsmenge des homogenen linearen Gleichungssystems $Ax = 0$.
- (3) Der Kern von A ist ein Unterraum von \mathbb{R}^n .

BEWEIS.

- (1) $\ker(A) = \{x \in \mathbb{R}^n \mid Ax = 0\} = \{x \in \mathbb{R}^n \mid f(x) \in \{0\}\} = f^{-1}(\{0\})$.
- (2) Das ist genau die Definition des Kernes.
- (3) Das folgt aus der Linearität von f . Weil $\{0\} \subset \mathbb{R}^m$ ein Unterraum des \mathbb{R}^m ist, ist auch das Urbild davon ein Unterraum des \mathbb{R}^n .

□

SATZ 5.4.4. Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine lineare Abbildung. Es sind äquivalent:

- (i) f ist injektiv.
- (ii) $\ker(f) = \{0\}$.

BEWEIS. (i) \Rightarrow (ii): Sei f injektiv. Dann kann höchstens ein Vektor aus \mathbb{R}^n auf den Nullvektor abgebildet werden, und weil f linear ist, wissen wir, dass der Nullvektor auf Null abgebildet wird. Also ist $\ker(f) = \{0\}$.

(ii) \Rightarrow (i): Sei umgekehrt $\ker(f) = \{0\}$ und seien $x, y \in \mathbb{R}^n$ so, dass $f(x) = f(y)$. Dann ist $0 = f(x) - f(y) = f(x - y)$. Also ist $x - y \in \ker(f)$ und daher ist $x - y = 0$. Es können also keine zwei verschiedenen Vektoren dasselbe Bild haben, und daher ist f injektiv. □

BEHAUPTUNG 5.4.5. Sei $A \in \mathbb{R}^{m \times n}$ und

$$f : \begin{cases} \mathbb{R}^n & \rightarrow \mathbb{R}^m \\ x & \mapsto Ax \end{cases}$$

die dazu gehörende lineare Abbildung.

- (1) Der Spaltenraum von A ist zugleich das Bild von \mathbb{R}^n unter der Abbildung f , d.h., $\text{rg}(A) = f(\mathbb{R}^n)$.

- (2) Der Spaltenraum von A ist die Menge jener Vektoren b , für die das Gleichungssystem $Ax = b$ lösbar ist.
- (3) Der Spaltenraum von A ist ein Unterraum von \mathbb{R}^m .
- (4) Die Abbildung f ist genau dann surjektiv, wenn $\text{rg}(A) = \mathbb{R}^m$.

BEWEIS.

- (1) Seien $s_1, \dots, s_n \in \mathbb{R}^m$ die Spalten von A . Wegen Behauptung 5.2.6 gilt

$$f\left(\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}\right) = A \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \sum_{j=1}^n \xi_j s_j.$$

Daher ist das Bild jedes Vektors aus \mathbb{R}^n eine Linearkombination der Spalten von A , und jede Linearkombination der Spalten von A das Bild eines Vektors aus \mathbb{R}^n .

- (2) Das System $Ax = b$ ist genau dann lösbar, wenn $f(x) = b$ eine Lösung besitzt, und das ist genau dann der Fall, wenn $b \in f(\mathbb{R}^n)$.
- (3) Der Spaltenraum ist definiert als die lineare Hülle der Spalten, und die lineare Hülle ist immer ein Unterraum.
- (4) Nach Definition ist f genau dann surjektiv, wenn $f(\mathbb{R}^n)$ der ganze Raum \mathbb{R}^m ist.

□

Um den Zeilenraum im Kontext von linearen Abbildungen zu deuten, würden wir den Begriff der dualen Abbildung brauchen. Darauf lassen wir uns jetzt noch nicht ein.

5.5. Basen und Dimension von Kern, Spaltenraum und Zeilenraum.

Jetzt brauchen wir noch eine möglichst effiziente Methode, um bei gegebener Matrix A die Unterräume $\ker(A)$, $\text{rg}(A)$ und $\text{zr}(A)$ zu berechnen.

LEMMA 5.5.1. Sei $A \in \mathbb{R}^{m \times n}$ und $\tilde{A} \in \mathbb{R}^{\tilde{m} \times n}$. Die Matrix \tilde{A} entstehe aus A durch einen oder mehrere der folgenden Schritte:

- (1) Zur Zeile Nummer i wird das γ -fache einer anderen Zeile (Nummer j) addiert.
- (2) Zeile Nummer i wird mit einer reellen Zahl $\gamma \neq 0$ multipliziert.
- (3) Die Zeilen Nummer i und j tauschen Platz.
- (4) Eine Nullzeile wird hinzugefügt.
- (5) Eine Nullzeile wird weggelassen.

Es seien s_1, \dots, s_n die Spalten von A und $\tilde{s}_1, \dots, \tilde{s}_n$ die Spalten von \tilde{A} . Dann gilt:

- (a) A und \tilde{A} haben denselben Zeilenraum.
- (b) Seien $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Dann ist

$$\sum_{j=1}^n \lambda_j s_j = 0 \quad \text{genau dann, wenn} \quad \sum_{j=1}^n \lambda_j \tilde{s}_j = 0.$$

- (c) A und \tilde{A} haben denselben Kern.
- (d) Seien j_1, \dots, j_k Indizes zwischen 1 und n . Dann ist $(s_{j_1}, \dots, s_{j_k})$ genau dann ein Erzeugendensystem des Spaltenraumes von A , wenn $(\tilde{s}_{j_1}, \dots, \tilde{s}_{j_k})$ ein Erzeugendensystem des Spaltenraumes von \tilde{A} ist.
- (e) Seien j_1, \dots, j_k Indizes zwischen 1 und n . Dann ist $(s_{j_1}, \dots, s_{j_k})$ genau dann linear unabhängig, wenn $(\tilde{s}_{j_1}, \dots, \tilde{s}_{j_k})$ linear unabhängig ist.
- (f) Seien j_1, \dots, j_k Indizes zwischen 1 und n . Dann ist $(s_{j_1}, \dots, s_{j_k})$ genau dann eine Basis des Spaltenraumes von A , wenn $(\tilde{s}_{j_1}, \dots, \tilde{s}_{j_k})$ eine Basis des Spaltenraumes von \tilde{A} ist.

BEWEIS.

- (a) Der Beweis genügt natürlich für den Fall, dass \tilde{A} durch eine einzelne Umformung aus A entsteht. Wir zeigen zuerst, dass $\text{zr}(\tilde{A})$ eine Teilmenge von $\text{zr}(A)$ ist.

Bei Schritten (1) und (2) sind die neu entstandenen Zeilen Linearkombinationen der Zeilen von A , daher liegen sie im Zeilenraum von A . Schritt (3) (Vertauschung der Reihenfolge) ändert nichts am Gesamtbestand der Zeilen, und daher auch nichts am Zeilenraum. Und Schritte (4) und (5) verändern den Bestand der Zeilen nur um Nullzeilen, die zum Zeilenraum nichts beitragen. Also ist $\text{zr}(\tilde{A}) \subseteq \text{zr}(A)$.

Weil sich alle diese Schritte wieder durch Umformungen der Art (1)...(5) rückgängig machen lassen, kann man die Rollen von \tilde{A} und A im obigen Beweis vertauschen und es folgt auch $\text{zr}(A) \subseteq \text{zr}(\tilde{A})$.

- (b) Wir haben bereits in Satz 2.2.3 gezeigt, dass die elementaren Zeilenumformungen die Lösungsmenge des Gleichungssystems

$$A \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix} = 0, \quad \text{d.h.} \quad \sum_{j=1}^n \lambda_j s_j = 0$$

nicht verändern. Hinzunehmen oder Weglassen von Nullzeilen ändert wegen Behauptung 2.3.2 ebenfalls nichts an der Lösungsmenge. (Bei homogenen Gleichungssystemen steht rechts immer eine Null, daher kann es keine Zeilen geben, die auf Unlösbarkeit schließen lassen.)

- (c) Der Kern von A ist genau die Lösungsmenge des Gleichungssystems $Ax = 0$, und dies ist wieder die Lösungsmenge des Gleichungssystems $\tilde{A}x = 0$. Daher wird der Kern durch die Zeilenumformungen nicht verändert.
- (d) Das System $(s_{j_1}, \dots, s_{j_k})$ spannt genau dann den ganzen Spaltenraum von A auf, wenn sich jede Spalte s_q durch eine Linearkombination der Vektoren s_{j_1}, \dots, s_{j_k} bilden lässt, also

$$(\forall q \in 1, \dots, n) (\exists \lambda_1, \dots, \lambda_k \in \mathbb{R}) \quad s_q - \sum_{p=1}^k \lambda_p s_{j_p} = 0.$$

Wegen Punkt (b) ist das äquivalent zu

$$(\forall q \in 1, \dots, n) (\exists \lambda_1, \dots, \lambda_k \in \mathbb{R}) \quad \tilde{s}_q - \sum_{p=1}^k \lambda_p \tilde{s}_{j_p} = 0,$$

und das heißt wiederum, dass das System $(\tilde{s}_{j_1}, \dots, \tilde{s}_{j_k})$ den ganzen Spaltenraum von \tilde{A} aufspannt.

- (e) Das System $(s_{j_1}, \dots, s_{j_k})$ ist genau dann linear unabhängig, wenn die einzige Linearkombination, welche den Nullvektor ergibt, mit den Koeffizienten $\lambda_p = 0$ erfolgt, also

$$(\forall \lambda_1, \dots, \lambda_k \in \mathbb{R}) \quad \left[\sum_{p=1}^k \lambda_p s_{j_p} = 0 \Rightarrow \lambda_1 = \dots = \lambda_k = 0 \right].$$

Das ist wiederum wegen Punkt (b) äquivalent zu

$$(\forall \lambda_1, \dots, \lambda_k \in \mathbb{R}) \quad \left[\sum_{p=1}^k \lambda_p \tilde{s}_{j_p} = 0 \Rightarrow \lambda_1 = \dots = \lambda_k = 0 \right],$$

und das heißt, dass das System $(\tilde{s}_{j_1}, \dots, \tilde{s}_{j_k})$ linear unabhängig ist.

- (f) Wegen Punkt (d) und (e).

□

BEMERKUNG 5.5.2. Die Matrizen A und \tilde{A} in Lemma 5.5.1 haben zwar immer denselben Zeilenraum, aber normalerweise nicht denselben Spaltenraum. Nur die linearen Abhängigkeiten, die zwischen den Spalten bestehen, bleiben dieselben.

Wir werden ein Verfahren angeben, wie man Lemma 5.5.1 nützen kann, um Basen von Kern, Spaltenraum und Zeilenraum einer Matrix zu bestimmen. Statt durch einen formalen Beweis machen wir die Methode wahrscheinlich leichter verständlich, wenn wir sie an einem Beispiel durchführen und erklären, warum sie funktioniert.

BEISPIEL 5.5.3. Bestimmen Sie Basen des Kernes, des Spaltenraumes und des Zeilenraumes der Matrix

$$A = \begin{pmatrix} 8 & 1 & 15 & 2 & 6 & 2 \\ -5 & 2 & 3 & -2 & 0 & 1 \\ -8 & 2 & -6 & 1 & 15 & -2 \\ 0 & 1 & 3 & 1 & 7 & 0 \end{pmatrix}.$$

Lösung:

- *Zeilentransformationen:* Wir beginnen mit Zeilentransformationen, bis wir kein Pivotelement mehr finden, und führen dann die Rücksubstitution durch. Nullzeilen streichen wir aus der Matrix. Alle Pivotzeilen dividieren wir durch das Pivotelement, sodass an den Stellen des Pivotelementes je eine 1 zu stehen kommt.

$$\begin{aligned} A &= \begin{pmatrix} 8 & 1 & 15 & 2 & 6 & 2 \\ -5 & 2 & 3 & -2 & 0 & 1 \\ -8 & 2 & -6 & 1 & 15 & -2 \\ 0 & 1 & 3 & 1 & 7 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 8 & 0 & 12 & 1 & -1 & 2 \\ -5 & 0 & -3 & -4 & -14 & 1 \\ -8 & 0 & -12 & -1 & 1 & -2 \\ 0 & 1 & 3 & 1 & 7 & 0 \end{pmatrix} \\ &\rightarrow \begin{pmatrix} 8 & 0 & 12 & 1 & -1 & 2 \\ 27 & 0 & 45 & 0 & -18 & 9 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 3 & 1 & 7 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 8 & 0 & 12 & 1 & -1 & 2 \\ 3 & 0 & 5 & 0 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 3 & 1 & 7 & 0 \end{pmatrix} \\ &\rightarrow \begin{pmatrix} 2 & 0 & 2 & 1 & 3 & 0 \\ 3 & 0 & 5 & 0 & -2 & 1 \\ 0 & 1 & 3 & 1 & 7 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 0 & 2 & 1 & 3 & 0 \\ 3 & 0 & 5 & 0 & -2 & 1 \\ -2 & 1 & 1 & 0 & 4 & 0 \end{pmatrix} = \tilde{A}. \end{aligned}$$

Nun sind wir mit den Zeilentransformationen fertig. Jede Zeile war einmal Pivotzeile, alle anderen Zeilen wurden zu Nullzeilen und wurden gestrichen. Wir haben insgesamt drei Pivotelemente gefunden, daher hat die Matrix \tilde{A} drei Zeilen.

- *Einteilung in basische und nichtbasische Spaltenindizes:* Die basischen Spalten sind jene, die als Pivotspalten verwendet wurden. Weil wir drei Pivotelemente gefunden haben, gibt es drei basische Spalten. Jede Pivotspalte hat ihr Pivotelement, das zu 1 skaliert wurde. An allen anderen Stellen der Pivotspalte wurde eine Null erzeugt. Also steht in jeder basischen Spalte ein Einheitsvektor. Weil jede Zeile einmal als Pivotzeile verwendet wurde, kommt jeder Einheitsvektor in irgendeiner Spalte vor. In unserem Beispiel stehen in den basischen Spalten Nummer 4, 6 und 2 die drei Einheitsvektoren. Die Matrix A , und damit auch \tilde{A} , hat 6 Spalten. Drei davon sind basisch, also bleiben drei nichtbasische Spalten übrig. Es ist Zufall, dass wir in diesem Beispiel gleich viele basische wie nichtbasische Spalten haben, das ist nicht immer so.
- *Bestimmung einer Basis des Zeilenraumes von A :* Der Zeilenraum von A ist nach Lemma 5.5.1 zugleich der Zeilenraum von \tilde{A} . Er wird von den drei Zeilen

$$\begin{aligned} \tilde{z}_1 &= (2 \ 0 \ 2 \ 1 \ 3 \ 0), \\ \tilde{z}_2 &= (3 \ 0 \ 5 \ 0 \ -2 \ 1), \\ \tilde{z}_3 &= (-2 \ 1 \ 1 \ 0 \ 4 \ 0) \end{aligned}$$

aufgespannt, die in \tilde{A} verblieben sind. Wir behaupten, dass dieses System auch linear unabhängig ist. Damit ist $(\tilde{z}_1, \tilde{z}_2, \tilde{z}_3)$ eine Basis des Zeilenraumes von A .

Die lineare Unabhängigkeit rechtfertigen wir mit folgender Rechnung, bei der wir nur

auf die Koeffizienten mit den basischen Indices 4, 6, 2 achten

$$\begin{aligned} & \lambda_1 \tilde{z}_1 + \lambda_2 \tilde{z}_2 + \lambda_3 \tilde{z}_3 = \\ & \lambda_1 \begin{pmatrix} * & 0 & * & 1 & * & 0 \end{pmatrix} \\ & + \lambda_2 \begin{pmatrix} * & 0 & * & 0 & * & 1 \end{pmatrix} \\ & + \lambda_3 \begin{pmatrix} * & 1 & * & 0 & * & 0 \end{pmatrix} \\ & = \begin{pmatrix} * & \lambda_3 & * & \lambda_1 & * & \lambda_2 \end{pmatrix} \end{aligned}$$

Ist also $\lambda_1 \tilde{z}_1 + \lambda_2 \tilde{z}_2 + \lambda_3 \tilde{z}_3 = 0$, so sind $\lambda_1 = \lambda_2 = \lambda_3 = 0$. Daher ist das System $(\tilde{z}_1, \tilde{z}_2, \tilde{z}_3)$ linear unabhängig.

Weil der Zeilenraum von A eine Basis aus drei Zeilenvektoren hat, ist die Dimension des Zeilenraumes ebenfalls 3, und das ist zugleich die Anzahl der gefundenen Pivotelemente.

- *Berechnung einer Basis des Kerns von A* : Der Kern von A ist wegen Lemma 5.5.1 zugleich die Lösungsmenge des homogenen Gleichungssystems $\tilde{A}x = 0$. Wir bringen die nichtbasischen Variablen auf die rechte Seite des Gleichungssystems. Links bleibt für jede Zeile je eine basische Variable übrig:

$$\begin{aligned} \xi_4 &= -2\xi_1 - 2\xi_3 - 3\xi_5 \\ \xi_6 &= -3\xi_1 - 5\xi_3 + 2\xi_5 \\ \xi_2 &= 2\xi_1 - \xi_3 - 4\xi_5 \end{aligned}$$

Daher gibt es, wie wir schon lange wissen, für jeden beliebigen Satz von Werten ξ_1, ξ_3, ξ_5 , den man den nichtbasischen Variablen zuweist, genau eine Lösung des Gleichungssystems:

$$x = \xi_1 \begin{pmatrix} 1 \\ 2 \\ 0 \\ -2 \\ 0 \\ -3 \end{pmatrix} + \xi_3 \begin{pmatrix} 0 \\ -1 \\ 1 \\ -2 \\ 0 \\ -5 \end{pmatrix} + \xi_5 \begin{pmatrix} 0 \\ -4 \\ 0 \\ -3 \\ 1 \\ 2 \end{pmatrix} =: \xi_1 x_1 + \xi_3 x_3 + \xi_5 x_5$$

Wir wissen nun, dass jede Lösung $x \in \ker(A)$ sich als Linearkombination von diesen Vektoren x_1, x_3, x_5 schreiben lässt. Wir behaupten, dass diese drei Vektoren eine Basis von $\ker(A)$ bilden.

Dazu müssen wir zeigen, dass das System (x_1, x_3, x_5) linear unabhängig ist. In der folgenden Rechnung behalten wir nur die Koeffizienten im Auge, die an den nichtbasischen Indices 1,3,5 stehen. Es seien $\lambda_1, \lambda_3, \lambda_5 \in \mathbb{R}$ so, dass $\lambda_1 x_1 + \lambda_3 x_3 + \lambda_5 x_5 = 0$.

$$0 = \lambda_1 \begin{pmatrix} 1 \\ * \\ 0 \\ * \\ 0 \\ * \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ * \\ 1 \\ * \\ 0 \\ * \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ * \\ 0 \\ * \\ 1 \\ * \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ * \\ \lambda_2 \\ * \\ \lambda_3 \\ * \end{pmatrix}.$$

Daraus folgt $\lambda_1 = \lambda_3 = \lambda_5 = 0$. Also ist das System (x_1, x_3, x_5) linear unabhängig.

Der Kern von A hat eine Basis aus 3 Elementen, zu jeder nichtbasischen Variablen gehört ein x_i . Daher ist die Dimension des Kerns von A gleich drei.

- *Bestimmung einer Basis des Spaltenraumes von A* : Die Spalten mit Indices 6, 4, 2 von \tilde{A} sind die Einheitsvektoren und damit eine Basis des gesamten \mathbb{R}^3 . Damit sehen wir auch, dass der Spaltenraum von \tilde{A} der ganze \mathbb{R}^3 sein muss. Wegen Lemma 5.5.1 bilden also auch die Spalten (s_4, s_2, s_6) der Matrix A eine Basis des Spaltenraumes von A . Daher ist das System der Vektoren

$$s_4 = \begin{pmatrix} -2 \\ -2 \\ 1 \\ 1 \end{pmatrix}, \quad s_6 = \begin{pmatrix} 2 \\ 1 \\ -2 \\ 0 \end{pmatrix}, \quad s_2 = \begin{pmatrix} 1 \\ 2 \\ 2 \\ 1 \end{pmatrix}$$

eine Basis des Spaltenraumes von A . Der Spaltenraum hat eine Basis von 3 Vektoren, zu jeder basischen Spalte je einer. Daher ist die Dimension des Spaltenraumes von A gleich drei.

Damit ist das Beispiel 5.5.3 fertig analysiert.

ALGORITHMUS 5.5.4. Sei $A \in \mathbb{R}^{m \times n}$ eine $(m \times n)$ -Matrix über \mathbb{R} . Mit dem folgenden Verfahren kann man Basen von $\ker(A)$, $\text{rg}(A)$ und $\text{zr}(A)$ finden.

- (1) *Zeilentransformationen:* An der Matrix A werden Pivotschritte so lange ausgeführt, bis keine Zeile mehr als mögliche Pivotzeile übrigbleibt. Sollten Nullzeilen übrigbleiben, so werden diese gestrichen. Durch Rücksubstitution werden in den Pivotspalten alle Elemente außer dem Pivotelement auf Null gebracht. Jede Pivotzeile wird durch ihr Pivotelement dividiert. Es bleibt eine Matrix \tilde{A} mit $r \leq m$ Zeilen und n Spalten übrig. Unter den Spalten von \tilde{A} kommen alle r Einheitsvektoren des \mathbb{R}^r vor. Mit $\tilde{\alpha}_{i,j}$ bezeichnen wir die Einträge von \tilde{A} .
- (2) *Klassifizieren der Spalten:* Mit b_i bezeichnen wir die basische Spalte, die den i -ten Einheitsvektor enthält. $B = \{b_1, \dots, b_r\}$ bezeichnet die Menge der basischen Spaltenindices. Mit N bezeichnen wir die Menge der nicht-basischen Spaltenindices.
- (3) *Eine Basis des Kerns von A :* Für jeden nichtbasischen Spaltenindex $k \in N$ konstruieren wir den Vektor $x_k \in \mathbb{R}^n$ mit den Komponenten

$$\xi_j := \begin{cases} 1 & \text{wenn } j = k \\ 0 & \text{wenn } j \in N \setminus \{k\} \\ -\tilde{\alpha}_{i,j} & \text{wenn } j = b_i \in B \end{cases} .$$

Das System (x_1, \dots, x_{n-r}) ist dann eine Basis von $\ker(A)$.

- (4) *Eine Basis des Spaltenraumes von A :* Die Spalten mit den basischen Spaltenindices b_1, \dots, b_r der ursprünglichen Matrix A bilden eine Basis des Spaltenraumes von A .
- (5) *Eine Basis des Zeilenraumes von A :* Die Zeilen der Matrix \tilde{A} bilden eine Basis des Zeilenraumes von A .
- (6) *Dimensionen:* Die Dimensionen der drei Räume sind:

$$\dim(\ker(A)) = n - r, \quad \dim(\text{rg}(A)) = r, \quad \dim(\text{zr}(A)) = r .$$

DEFINITION 5.5.5. Sei $A \in \mathbb{R}^{m \times n}$. An A werden so lange Pivotschritte durchgeführt, bis sich keine möglichen Pivotzeilen mehr finden. Die Anzahl der gefundenen Pivotelemente heißt der Rang von A (geschrieben²⁰ $\rho(A)$ oder manchmal $\text{rk}(A)$).

Der Rang ist gleichzeitig die Dimension des Zeilenraumes und des Spaltenraumes von A .

Die Zahl $n - \rho(A)$ heißt auch Rangdefekt von A , und ist die Dimension des Kernes von A .

Mit der Formulierung "Zeilenrang = Spaltenrang" drückt man aus, dass die Dimension des Zeilenraumes einer Matrix dieselbe ist wie die Dimension des Spaltenraumes. Man könnte den Rang ebensogut mit elementaren Spaltenumformungen bestimmen.

SATZ 5.5.6. Sei $A \in \mathbb{R}^{m \times n}$, also eine Matrix mit m Zeilen und n Spalten. Zur Matrix A betrachten wir die lineare Abbildung

$$f : \begin{cases} \mathbb{R}^n & \rightarrow \mathbb{R}^m \\ x & \mapsto Ax \end{cases} .$$

Dann gilt:

- (1) Die Abbildung f ist genau dann surjektiv, wenn der Rang $\rho(A) = m$.
- (2) Die Abbildung f ist genau dann injektiv, wenn der Rang $\rho(A) = n$.

BEWEIS.

²⁰Lies: ρ : rho

- (1) f ist genau dann surjektiv, wenn das Bild von f der ganze \mathbb{R}^m ist, also $\text{rg}(A) = \mathbb{R}^m$, und das gilt wieder genau dann, wenn $\rho(A) = \dim(\text{rg}(A)) = m$.
- (2) f ist wegen Satz 5.4.4 genau dann injektiv, wenn $\ker(A) = \{0\}$, also $\dim(\ker(A)) = n - r = 0$.

□

6. Matrix mal Matrix

6.1. Linearkombination von Matrizen.

Bevor wir uns mit der wichtigsten Matrizenoperation befassen, der Matrixmultiplikation, führen wir noch ein, wie man mit Matrizen genauso wie mit Vektoren Linearkombinationen bilden kann.

DEFINITION 6.1.1. Seien $A = (\alpha_{i,j})_{i=1,\dots,m,j=1,\dots,n}$ und $B = (\beta_{i,j})$ zwei $(m \times n)$ -Matrizen in $\mathbb{R}^{m \times n}$. Sei $\lambda \in \mathbb{R}$. Wir definieren

$$A \pm B := \begin{pmatrix} \alpha_{1,1} \pm \beta_{1,1} & \alpha_{1,2} \pm \beta_{1,2} & \cdots & \alpha_{1,n} \pm \beta_{1,n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m,1} \pm \beta_{m,1} & \alpha_{m,2} \pm \beta_{m,2} & \cdots & \alpha_{m,n} \pm \beta_{m,n} \end{pmatrix},$$

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

Beachten Sie, dass man (ähnlich wie bei Spaltenvektoren) nur Matrizen derselben Größe miteinander addieren darf.

BEISPIEL 6.1.2.

$$2 \begin{pmatrix} 1 & 5 & 3 \\ -3 & 0 & 4 \end{pmatrix} + 5 \begin{pmatrix} 1 & 1 & 0 \\ -2 & 0 & 3 \end{pmatrix} - 3 \begin{pmatrix} 2 & -3 & 1 \\ -1 & 4 & -2 \end{pmatrix} = \begin{pmatrix} 1 & 24 & 3 \\ -13 & -12 & 29 \end{pmatrix}.$$

DEFINITION 6.1.3. Die $(m \times n)$ -Nullmatrix ist die $(m \times n)$ -Matrix

$$0_{m \times n} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}.$$

Aus dem Zusammenhang wird meistens klar, ob es sich um die reelle Zahl Null, den Nullvektor oder die Nullmatrix handelt, daher lassen wir normalerweise die Subscripts $m \times n$ weg und schreiben auch für die Nullmatrizen einfach 0.

BEHAUPTUNG 6.1.4. *Es gelten dieselben Rechenregeln wie für die Linearkombination von Spaltenvektoren (vgl. Behauptung 3.1.2)*

Seien $\lambda, \mu \in \mathbb{R}$ und $A, B, C \in \mathbb{R}^{m \times n}$ Matrizen derselben Größe. Es gilt

$$\begin{aligned} A + B &= B + A \\ A + (B + C) &= (A + B) + C \\ A + ((-1)A) &= 0_{m \times n} \\ A + 0_{m \times n} &= A \\ (\lambda + \mu)A &= \lambda A + \mu A \\ \lambda(A + B) &= \lambda A + \lambda B \\ \lambda(\mu A) &= (\lambda\mu)A \\ 1A &= A \end{aligned}$$

Versuchen Sie selbst, einige dieser Regeln zu beweisen. Es ist ganz leicht.

Den Wert der folgenden Umformung versteht man erst voll, wenn man sich mit dem inneren Produkt beschäftigt. Der Vollständigkeit halber führen wir sie aber jetzt gleich an.

DEFINITION 6.1.5. Sei A eine $(m \times n)$ -Matrix. Die Transponierte A^T von A ist die $(n \times m)$ -Matrix, die entsteht, wenn man die Zeilen von A als Spalten anschreibt, und umgekehrt:

$$\begin{pmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,n} \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & \alpha_{2,n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m,1} & \alpha_{m,2} & \cdots & \alpha_{m,n} \end{pmatrix}^T = \begin{pmatrix} \alpha_{1,1} & \alpha_{2,1} & \cdots & \alpha_{m,1} \\ \alpha_{1,2} & \alpha_{2,2} & \cdots & \alpha_{m,2} \\ \vdots & \vdots & & \vdots \\ \alpha_{1,n} & \alpha_{2,n} & \cdots & \alpha_{m,n} \end{pmatrix}.$$

Die folgenden Rechenregeln sind leicht zu beweisen:

BEHAUPTUNG 6.1.6. Seien $A, B \in \mathbb{R}^{m \times n}$, $\lambda \in \mathbb{R}$:

$$\begin{aligned} (A^T)^T &= A \\ (A + B)^T &= A^T + B^T \\ (\lambda A)^T &= \lambda A^T \end{aligned}$$

Trick: In einem eng gedruckten Text kann man einen Spaltenvektor $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$ manchmal leichter als transponierten Zeilenvektor $(1 \ 2 \ 3 \ 4)^T$ unterbringen.

6.2. Hintereinanderausführung von Funktionen, Identität, Umkehrfunktion.

DEFINITION 6.2.1. Seien A, B, C Mengen, und $f : A \rightarrow B$, $g : B \rightarrow C$ Funktionen. Die Hintereinanderausführung (Komposition) $g \circ f$ ("g nach f") ist die Funktion

$$g \circ f : \begin{cases} A & \rightarrow C \\ a & \mapsto g(f(a)) \end{cases}.$$

BEMERKUNG 6.2.2. Achten Sie darauf, dass die Hintereinanderausführung $g \circ f$ nur existiert, wenn der Wertevorrat von f im Definitionsbereich von g als Teilmenge enthalten ist.

SATZ 6.2.3. Die Hintereinanderausführung von Funktionen ist assoziativ im folgenden Sinn:

Seien A, B, C, D Mengen, und $f : A \rightarrow B$, $g : B \rightarrow C$, $h : C \rightarrow D$ Funktionen. Dann ist $(h \circ g) \circ f = h \circ (g \circ f)$.

BEWEIS. Es ist $g \circ f$ eine Funktion von A nach C und h eine Funktion von C nach D , daher existiert $h \circ (g \circ f)$ und ist eine Funktion von A nach D . Dabei ist für $a \in A$

$$[(h \circ (g \circ f))(a) = h((g \circ f)(a)) = h(g(f(a))).$$

Es ist f eine Funktion von A nach B , und $h \circ g$ eine Funktion von B nach D . Daher existiert $(h \circ g) \circ f$ und es ist

$$[(h \circ g) \circ f](a) = (h \circ g)(f(a)) = h(g(f(a))).$$

Beide Formeln geben dieselbe Funktion. □

BEMERKUNG 6.2.4. Die Reihenfolge der Hintereinanderausführung darf aber im Allgemeinen nicht vertauscht werden.

Beispiel 1: Seien $f, g : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch $f(x) = x^2$, $g(x) = x + 1$. Es ist

$$[g \circ f](x) = g(x^2) = x^2 + 1, \quad [f \circ g](x) = f(x + 1) = (x + 1)^2.$$

Beispiel 2: Seien $f : \mathbb{R} \rightarrow \mathbb{R}^2$ und $g : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ Funktionen. Es existiert dann $g \circ f : \mathbb{R} \rightarrow \mathbb{R}^3$, aber nicht $f \circ g$. Denn ist $x \in \mathbb{R}^2$, so ist $g(x) \in \mathbb{R}^3$, aber $f(g(x))$ ist nicht definiert, weil $f(y)$ nur für $y \in \mathbb{R}$ existiert. □

DEFINITION 6.2.5. Sei M eine beliebige nichtleere Menge, so ist die Identität auf M definiert durch

$$\text{id}_M : \begin{cases} M & \rightarrow M \\ x & \mapsto x \end{cases}.$$

BEMERKUNG 6.2.6. Seien A und B Mengen und $f : A \rightarrow B$ eine Funktion. Dann gilt

$$f = \text{id}_B \circ f = f \circ \text{id}_A.$$

DEFINITION 6.2.7. Seien A und B Mengen und $f : A \rightarrow B$ eine Funktion. Sei $g : B \rightarrow A$ eine Funktion mit der Eigenschaft

$$g \circ f = \text{id}_A, \quad f \circ g = \text{id}_B.$$

Dann heißt g Umkehrfunktion (inverse Funktion) zu f , und wir schreiben²¹ $g = f^{-1}$.

Das bedeutet: Die Funktion und die Umkehrfunktion heben einander auf. Führt man erst f , und anschließend f^{-1} aus, kommt man zum Ausgangswert zurück. Ebenso, wenn man erst f^{-1} und dann f ausführt.

Die Schreibweise $g = f^{-1}$ wird durch die folgende Bemerkung gerechtfertigt:

BEHAUPTUNG 6.2.8. Sei $f : A \rightarrow B$ eine Funktion. Dann besitzt f höchstens eine Umkehrfunktion.

BEWEIS. Seien $g : B \rightarrow A$ und $h : B \rightarrow A$ Umkehrfunktionen von f . Dann ist

$$g = g \circ \text{id}_B = g \circ (f \circ h) = (g \circ f) \circ h = \text{id}_A \circ h = h.$$

□

²¹Das Symbol f^{-1} hat also leider zwei verschiedene Bedeutungen: $f^{-1}(M) = \{x \mid f(x) \in M\}$ Urbild einer Menge $M \subset B$, gibt es für jede Funktion f . Aber: $f^{-1}(y)$ für $y \in B$ ist der Wert der Umkehrfunktion an der Stelle y . Den gibt es nur, wenn f eine Umkehrfunktion besitzt.

BEHAUPTUNG 6.2.9. Seien A und B Mengen, und $f : A \rightarrow B$ eine Funktion. Dann sind äquivalent:

- (i) Die Funktion f besitzt eine Umkehrfunktion.
- (ii) Für jedes $y \in B$ hat die Gleichung $f(x) = y$ genau eine Lösung $x \in A$. (Diese Lösung ist dann $f^{-1}(y)$.)
- (iii) f ist bijektiv.

BEWEIS.

(i) \Rightarrow (ii): Wir nehmen an, dass f eine Umkehrfunktion f^{-1} besitzt. Sei $y \in B$ beliebig. Es ist $x = f^{-1}(y)$ eine Lösung von $f(x) = y$, denn

$$f(x) = f(f^{-1}(y)) = [f \circ f^{-1}](y) = \text{id}_B y = y.$$

Ist z irgendeine Lösung von $f(z) = y$, so ist

$$z = \text{id}_A(z) = [f^{-1} \circ f](z) = f^{-1}(f(z)) = f^{-1}(y).$$

Daher ist $x = f^{-1}(y)$ die einzige Lösung der Gleichung $f(x) = y$.

(ii) \Rightarrow (i): Wir nehmen an, dass für jedes $y \in B$ genau ein x existiert, sodass $f(x) = y$. Wir definieren die Funktion

$$g : \begin{cases} B & \rightarrow A \\ y & \mapsto x \text{ mit } f(x) = y \end{cases}.$$

Wir zeigen, dass g die Umkehrfunktion von f ist. Sei $z \in A$. Setze $y = f(z)$. Dann ist $g(f(z)) = g(y)$ die Lösung x von $f(x) = y = f(z)$, davon gibt es nach Voraussetzung genau eine, und das ist natürlich z . Also ist $g \circ f = \text{id}_A$. Sei nun $y \in B$ beliebig. Es ist $g(y)$ die eindeutige Lösung x der Gleichung $f(x) = y$, und damit ist natürlich $f(g(y)) = f(x) = y$. Also ist $f \circ g = \text{id}_B$.

(ii) \Leftrightarrow (iii): Bemerkungen 5.1.10 und 6.2.9.

□

BEISPIEL 6.2.10. Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch $f(x) = x^3 + 1$. Gibt es eine Umkehrfunktion?

Lösung: Ist $y \in \mathbb{R}$, so muss $x = f^{-1}(y)$ die Gleichung $x^3 + 1 = y$ lösen. Wir formen diese Gleichung äquivalent um:

$$x^3 + 1 = y \iff x^3 = y - 1 \iff x = \sqrt[3]{y - 1}.$$

Wir haben also zu jedem $y \in \mathbb{R}$ genau eine Lösung gefunden. Also ist $f^{-1}(y) = \sqrt[3]{y - 1}$. □

BEISPIEL 6.2.11. Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch

$$f\left(\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}\right) := \xi_1 + \xi_2.$$

Gibt es eine Umkehrfunktion?

Lösung: Sei $y \in \mathbb{R}$. Wir untersuchen, ob die Gleichung $f(x) = y$ eindeutig lösbar ist. Nun sind sowohl der Vektor $\begin{pmatrix} y \\ 0 \end{pmatrix}$ als auch $\begin{pmatrix} 0 \\ y \end{pmatrix}$ Lösungen. Daher ist die Lösung nicht eindeutig, und es gibt keine Umkehrfunktion. □

BEHAUPTUNG 6.2.12. Seien A , B und C Mengen, und $f : A \rightarrow B$, $g : B \rightarrow C$ Funktionen. Es gilt:

- (1) Sind $f : A \rightarrow B$ und $g : B \rightarrow C$ surjektiv, so ist auch $g \circ f : A \rightarrow C$ surjektiv.
- (2) Sind $f : A \rightarrow B$ und $g : B \rightarrow C$ injektiv, so ist auch $g \circ f : A \rightarrow C$ injektiv.

(3) Sind $f : A \rightarrow B$ und $g : B \rightarrow C$ bijektiv, so ist auch $g \circ f : A \rightarrow C$ bijektiv. In diesem Fall ist $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

BEWEIS. Seien f, g surjektiv. Sei $z \in C$. Dann gibt es wegen der Surjektivität von g eine Lösung $y \in B$ von $g(y) = z$, und wegen der Surjektivität von f eine Lösung $x \in A$ von $f(x) = y$. Es ist dann $[g \circ f](x) = g(f(x)) = g(y) = z$, also hat $[g \circ f](x) = y$ eine Lösung.

Seien nun f und g injektiv, und seien $x_1 \neq x_2 \in A$. Wegen Injektivität von f ist $f(x_1) \neq f(x_2)$, und wegen der Injektivität von g ist $g(f(x_1)) \neq g(f(x_2))$. Also ist $[g \circ f](x_1) \neq [g \circ f](x_2)$.

Der dritte Punkt folgt unmittelbar aus den beiden ersten. Es gilt auch wegen der Assoziativität der Hintereinanderausführung (Satz 6.2.3)

$$\begin{aligned}(f^{-1} \circ g^{-1}) \circ (g \circ f) &= f^{-1} \circ (g^{-1} \circ g) \circ f = f^{-1} \circ \text{id}_B \circ f = f^{-1} \circ f = \text{id}_A, \\ (g \circ f) \circ (f^{-1} \circ g^{-1}) &= g \circ (f \circ f^{-1}) \circ g^{-1} = g \circ \text{id}_B \circ g^{-1} = g \circ g^{-1} = \text{id}_C.\end{aligned}$$

Also ist $f^{-1} \circ g^{-1}$ die Umkehrfunktion zu $g \circ f$. \square

BEMERKUNG 6.2.13. Beachten Sie die Umkehr der Reihenfolge: $(f \circ g)^{-1} = g^{-1} \circ f^{-1}$.

6.3. Matrixprodukt.

DEFINITION 6.3.1. Seien $k, m, n \in \mathbb{N}$. Sei $A = (\alpha_{i,j}) \in \mathbb{R}^{k \times m}$ und $B = (\beta_{i,j}) \in \mathbb{R}^{m \times n}$. Wir definieren die Produktmatrix $A \cdot B \in \mathbb{R}^{k \times n}$ durch

$$\begin{aligned}A \cdot B &:= (\gamma_{i,j})_{i=1, \dots, k, j=1, \dots, n} \text{ mit} \\ \gamma_{i,j} &:= \sum_{s=1}^m \alpha_{i,s} \beta_{s,j}.\end{aligned}$$

BEMERKUNG 6.3.2.

- (1) Um $A \cdot B$ zu bilden, muss also die Anzahl der Spalten von A gleich der Anzahl der Zeilen von B sein. Das Element $\gamma_{i,j}$ der Matrix $A \cdot B$ wird dann aus der i -ten Zeile von A und der j -ten Spalte von B berechnet.
- (2) Ist $A \in \mathbb{R}^{m \times n}$ und $x \in \mathbb{R}^n$, so ist das Produkt Ax nichts anderes als das Matrizenprodukt, wenn man x als eine Matrix mit n Zeilen und einer Spalte auffasst.

Sind b_1, \dots, b_n die Spalten von B , so sind Ab_1, \dots, Ab_n die Spalten von AB .

BEISPIEL 6.3.3. Wir berechnen

$$\begin{pmatrix} 2 & 1 & 3 & 1 \\ 1 & 0 & -4 & 0 \\ 3 & 5 & -5 & 2 \end{pmatrix} \cdot \begin{pmatrix} 3 & 1 \\ 4 & 0 \\ 2 & 3 \\ 5 & 1 \end{pmatrix}.$$

Lösung:

Die Multiplikation ist erlaubt, weil die erste Matrix 4 Spalten und die zweite Matrix gleichviel, ebenfalls 4, Zeilen hat. Das Ergebnis wird wie die erste Matrix 3 Zeilen, und wie die zweite Matrix zwei Spalten haben.

Wir berechnen jetzt das Element in der ersten Zeile und ersten Spalte. Achten Sie nur auf die erste Zeile der vorderen und die erste Spalte der hinteren Matrix:

$$\begin{pmatrix} 2 & 1 & 3 & 1 \\ & & & \\ & & & \\ & & & \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 4 \\ 2 \\ 5 \end{pmatrix} = \begin{pmatrix} 2 \cdot 3 + 1 \cdot 4 + 3 \cdot 2 + 1 \cdot 5 \\ & \\ & \\ & \end{pmatrix} = \begin{pmatrix} 24 \\ & \\ & \\ & \end{pmatrix}.$$

Das Element in der ersten Zeile und zweiten Spalte des Produktes entsteht aus der ersten Zeile der ersten Matrix und der zweiten Spalte der zweiten Matrix:

$$\begin{pmatrix} 2 & 1 & 3 & 1 \\ 1 & 0 & -4 & 0 \\ 3 & 5 & -5 & 2 \end{pmatrix} \cdot \begin{pmatrix} 3 & 1 \\ 4 & 0 \\ 2 & 3 \\ 5 & 1 \end{pmatrix} = \begin{pmatrix} 24 & 2 \cdot 1 + 1 \cdot 0 + 3 \cdot 3 + 1 \cdot 1 \\ \vdots & \vdots \end{pmatrix} = \begin{pmatrix} 24 & 12 \\ \vdots & \vdots \end{pmatrix}.$$

Die zweite Zeile des Produktes wird mit der zweiten Zeile der vorderen Matrix und den jeweiligen Spalten der hinteren Matrix gebildet:

$$\begin{pmatrix} 2 & 1 & 3 & 1 \\ 1 & 0 & -4 & 0 \\ 3 & 5 & -5 & 2 \end{pmatrix} \cdot \begin{pmatrix} 3 & 1 \\ 4 & 0 \\ 2 & 3 \\ 5 & 1 \end{pmatrix} = \begin{pmatrix} 24 & 12 \\ -5 & -11 \\ \vdots & \vdots \end{pmatrix}.$$

Nun berechnen wir noch die dritte Zeile:

$$\begin{pmatrix} 2 & 1 & 3 & 1 \\ 1 & 0 & -4 & 0 \\ 3 & 5 & -5 & 2 \end{pmatrix} \cdot \begin{pmatrix} 3 & 1 \\ 4 & 0 \\ 2 & 3 \\ 5 & 1 \end{pmatrix} = \begin{pmatrix} 24 & 12 \\ -5 & -11 \\ 29 & -10 \end{pmatrix}.$$

Jetzt ist das Produkt fertig berechnet:

$$\begin{pmatrix} 2 & 1 & 3 & 1 \\ 1 & 0 & -4 & 0 \\ 3 & 5 & -5 & 2 \end{pmatrix} \cdot \begin{pmatrix} 3 & 1 \\ 4 & 0 \\ 2 & 3 \\ 5 & 1 \end{pmatrix} = \begin{pmatrix} 24 & 12 \\ -5 & -11 \\ 29 & -10 \end{pmatrix}.$$

Noch ein paar Matrixprodukte:

BEISPIEL 6.3.4.

$$\begin{aligned} \begin{pmatrix} 2 & 1 & 3 \\ -1 & -2 & 0 \end{pmatrix} \cdot \begin{pmatrix} 4 & 5 \\ 3 & 1 \\ 2 & 2 \end{pmatrix} &= \begin{pmatrix} 17 & 17 \\ -10 & -7 \end{pmatrix}, \\ \begin{pmatrix} 4 & 5 \\ 3 & 1 \\ 2 & 2 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 & 3 \\ -1 & -2 & 0 \end{pmatrix} &= \begin{pmatrix} 3 & -6 & 12 \\ 5 & 1 & 9 \\ 2 & -2 & 6 \end{pmatrix}, \\ \begin{pmatrix} 2 & 1 & 3 \\ -1 & -2 & 0 \end{pmatrix} \cdot \begin{pmatrix} 4 & 3 & 2 \\ 5 & 1 & 2 \end{pmatrix} &\text{ gibt es nicht!} \\ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \text{aber} \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \\ \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \\ \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \cdot \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} &= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}. \end{aligned}$$

BEMERKUNG 6.3.5. An Beispiel 6.3.4 kann man sehen:

- (1) Das Matrixprodukt ist nicht kommutativ: Es gibt Matrizen A, B sodass $AB \neq BA$.
- (2) Das Matrixprodukt hat Nullteiler, das heißt, es gibt Matrizen A, B , beide nicht Null, sodass das Produkt AB trotzdem die Nullmatrix ergibt.
- (3) Idempotente Matrizen sind Matrizen $A \in \mathbb{R}^{n \times n}$ für die $AA = A$ ist. Das ist klar für die $n \times n$ -Nullmatrix und die Einheitsmatrix. Aber darüber hinaus gibt es auch andere idempotente Matrizen.
- (4) Nilpotente Matrizen sind $n \times n$ -Matrizen, für die es eine Potenz A^k gibt, welche die Nullmatrix ergibt. Das ist klar für die Nullmatrix, aber darüber hinaus gibt es auch andere nilpotente Matrizen.

Es folgt eine Litanei von Rechenregeln für das Matrixprodukt. Alle diese Regeln kann man durch direktes Nachrechnen beweisen: Beweis durch Förderung der Papierindustrie. Versuchen Sie selbst, die eine oder andere Regel zu beweisen.

BEHAUPTUNG 6.3.6. Seien $k, m, n, p \in \mathbb{N}$, sei²² $A \in \mathbb{R}^{k \times m}$, $B, C \in \mathbb{R}^{m \times n}$, $D \in \mathbb{R}^{n \times p}$, sei 0 die Nullmatrix, jeweils in geeigneten Dimensionen. Sei $\lambda \in \mathbb{R}$. Es gilt:

$$\begin{aligned} (A \cdot B) \cdot D &= A \cdot (B \cdot D) && \text{Assoziativit\u00e4t,} \\ A \cdot (B + C) &= A \cdot B + A \cdot C && \text{Distributivit\u00e4t,} \\ (B + C) \cdot D &= B \cdot D + C \cdot D && \text{Distributivit\u00e4t,} \\ (\lambda A) \cdot B &= \lambda(A \cdot B) = A \cdot (\lambda B), \\ B \cdot 0_{n \times p} &= 0_{m \times p}, \quad 0_{p \times m} \cdot B = 0_{p \times n}. \end{aligned}$$

Vorsicht ist geboten, wenn man Produkte transponiert: Die Reihenfolge der Faktoren kehrt sich um:

BEHAUPTUNG 6.3.7. Sei $A \in \mathbb{R}^{k \times m}$, $B \in \mathbb{R}^{m \times n}$. Dann ist

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

BEWEIS.

$$\begin{aligned} B^T A^T &= \begin{pmatrix} \beta_{11} & \cdots & \beta_{m1} \\ \vdots & & \vdots \\ \beta_{1n} & \cdots & \beta_{mn} \end{pmatrix} \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{k1} \\ \vdots & & \vdots \\ \alpha_{1m} & \cdots & \alpha_{km} \end{pmatrix} \\ &= \begin{pmatrix} \sum_{l=1}^m \alpha_{1l} \beta_{l1} & \cdots & \sum_{l=1}^m \alpha_{kl} \beta_{l1} \\ \vdots & & \vdots \\ \sum_{l=1}^m \alpha_{1l} \beta_{ln} & \cdots & \sum_{l=1}^m \alpha_{kl} \beta_{ln} \end{pmatrix} = (AB)^T. \end{aligned}$$

□

Das Matrixprodukt ist die \u00dcbersetzung der Hintereinanderausf\u00fchrung von Funktionen in die Sprache der Matrizen:

SATZ 6.3.8. Seien $F = (\phi_{ij}) \in \mathbb{R}^{m \times n}$, $G = (\gamma_{ij}) \in \mathbb{R}^{k \times m}$ und die dazugeh\u00f6rigen linearen Abbildungen

$$f : \begin{cases} \mathbb{R}^n & \rightarrow \mathbb{R}^m \\ x & \mapsto Fx \end{cases}, \quad g : \begin{cases} \mathbb{R}^m & \rightarrow \mathbb{R}^k \\ y & \mapsto Gy \end{cases}.$$

Dann ist die Hintereinanderausf\u00fchrung $g \circ f$ die Abbildung

$$g \circ f : \begin{cases} \mathbb{R}^n & \rightarrow \mathbb{R}^k \\ x & \mapsto (GF)x \end{cases}.$$

BEWEIS. Sei $H \in \mathbb{R}^{k \times n}$ die Matrix von $g \circ f$. Die Spalten von H sind die Bilder $[g \circ f](e_j)$ der Einheitsvektoren. Die Spalten von F sind die Bilder $f(e_j)$ der Einheitsvektoren. Wir betrachten die j -te Spalte von H .

$$\begin{aligned} [g \circ f](e_j) &= g(f(e_j)) = G(f(e_j)) \\ &= \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1m} \\ \vdots & & \vdots \\ \gamma_{k1} & \cdots & \gamma_{km} \end{pmatrix} \begin{pmatrix} \phi_{1j} \\ \vdots \\ \phi_{mj} \end{pmatrix} = \begin{pmatrix} \sum_{l=1:m} \gamma_{1l} \phi_{lj} \\ \vdots \\ \sum_{l=1:m} \gamma_{kl} \phi_{lj} \end{pmatrix} \end{aligned}$$

Das ist aber genau die j -te Spalte von GF . □

²²Die Dimensionen der Matrizen sind einfach so festgelegt, dass es die Produkte, die unten gebildet werden, wirklich gibt.

BEISPIEL 6.3.9. Wir bezeichnen mit $d : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ die Drehung um den Nullpunkt um 45° im Gegenuhrzeigersinn, und mit $s : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ die Spiegelung an der x -Achse. Was geschieht, wenn wir erst d und dann s ausführen?

Lösung:

Wir stellen erst die Matrizen von d und s auf. Die Matrix D von d hat in ihren Spalten die gedrehten Einheitsvektoren:

$$d\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{pmatrix}, \quad d\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} -\sqrt{2}/2 \\ \sqrt{2}/2 \end{pmatrix}, \quad D = \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix}.$$

Die Matrix S von s hat in den Spalten die an der x -Achse gespiegelten Einheitsvektoren:

$$s\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad s\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Wir suchen nun die Matrix von $s \circ d$ (Denken Sie daran: die Abbildung, die zuerst ausgeführt wird, steht sowohl beim Symbol \circ als auch bei der Matrixmultiplikation rechts im Produkt.) Die Matrix von $s \circ d$ ist

$$SD = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix} = \begin{pmatrix} \sqrt{2}/2 & -\sqrt{2}/2 \\ -\sqrt{2}/2 & -\sqrt{2}/2 \end{pmatrix}.$$

Wie man herausbekommt, dass diese Matrix eine Spiegelung beschreibt, werden wir erst im zweiten Semester im Rahmen der Eigenwerttheorie lernen.

BEHAUPTUNG 6.3.10. Seien $A = (\alpha_{ij}) \in \mathbb{R}^{k \times m}$ und $B = (\beta_{ij}) \in \mathbb{R}^{m \times n}$. Es gilt:

- (a) Der Spaltenraum von AB ist ein Unterraum des Spaltenraumes von A .
- (b) Der Zeilenraum von AB ist ein Unterraum des Zeilenraumes von B .
- (c) Für die Ränge gilt $\rho(AB) \leq \min(\rho(A), \rho(B))$.

BEWEIS.

- (a) Seien s_1, \dots, s_m die Spalten von A . Sei $j \in \{1 \dots n\}$. Die j -te Spalte von AB ist nach Definition des Matrixproduktes

$$\begin{pmatrix} \alpha_{11}\beta_{1j} + \dots + \alpha_{1m}\beta_{mj} \\ \vdots \\ \alpha_{k1}\beta_{1j} + \dots + \alpha_{km}\beta_{mj} \end{pmatrix} = \beta_{1j} \begin{pmatrix} \alpha_{11} \\ \vdots \\ \alpha_{k1} \end{pmatrix} + \dots + \beta_{mj} \begin{pmatrix} \alpha_{1m} \\ \vdots \\ \alpha_{km} \end{pmatrix} = \sum_{l=1}^m \beta_{lj} s_l.$$

Also liegen alle Spalten von AB im Spaltenraum von A .

- (b) Seien z_1, \dots, z_m die Zeilen von B . Sei $i \in \{1 \dots k\}$. Die i -te Zeile von AB ist nach Definition des Matrixproduktes

$$\begin{aligned} & (\alpha_{i1}\beta_{11} + \dots + \alpha_{im}\beta_{m1} \quad \dots \quad \alpha_{i1}\beta_{1n} + \dots + \alpha_{im}\beta_{mn}) \\ &= \alpha_{i1} (\beta_{11} \quad \dots \quad \beta_{1n}) + \dots + \alpha_{im} (\beta_{m1} \quad \dots \quad \beta_{mn}) \\ &= \sum_{l=1}^m \alpha_{il} z_l. \end{aligned}$$

Also liegen alle Zeilen von AB im Zeilenraum von B .

- (c) Der Rang ist die Dimension des Zeilenraumes. Weil $\text{zr}(AB) \subseteq \text{zr}(B)$ gilt auch $\rho(AB) \leq \rho(B)$. Andererseits ist der Rang auch die Dimension des Spaltenraumes. Weil $\text{rg}(AB) \subseteq \text{rg}(A)$ gilt auch $\rho(AB) \leq \rho(A)$.

□

6.4. Reguläre Matrizen.

DEFINITION 6.4.1. Eine Matrix heißt quadratisch wenn sie gleich viele Zeilen wie Spalten hat.

DEFINITION 6.4.2. Die $n \times n$ -Einheitsmatrix E_n ist die Matrix

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

SCHREIBWEISE 6.4.3. Den Index n lässt man oft weg und schreibt einfach E für die Einheitsmatrix, wenn ohnehin klar ist, wieviele Zeilen und Spalten sie hat. Andere gebräuchliche Symbole für die Einheitsmatrix sind I oder 1 . Ist $\lambda \in \mathbb{R}$ so schreibt man oft einfach λ anstatt λE — aus dem Zusammenhang muss dann geschlossen werden, ob man den Skalar meint oder eine Matrix, die auf der Diagonalen überall λ hat, und sonst Nullen.

DEFINITION 6.4.4. Das Kronecker-Delta $\delta : \mathbb{Z} \times \mathbb{Z} \rightarrow \{0, 1\}$ wird definiert durch

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

BEMERKUNG 6.4.5. Das Kronecker-Delta ist eine sehr praktische Schreibweise! Zum Beispiel kann man damit bequem die Einheitsmatrix definieren

$$E_n = (\delta_{i,j})_{i=1, \dots, n, j=1, \dots, n}$$

Auch die Einheitsvektoren in \mathbb{R}^n kann man schön hinschreiben:

$$e_k = (\delta_{i,k})_{i=1, \dots, n}.$$

SATZ 6.4.6. Sei $A \in \mathbb{R}^{m \times n}$, E_m und E_n die $m \times m$ bzw. $n \times n$ Einheitsmatrizen. Dann ist

$$E_m A = A E_n = A.$$

Insbesondere gilt für alle $x \in \mathbb{R}^n$: $E_n x = x$. Das heißt, die $n \times n$ -Einheitsmatrix ist die Matrix der Identitätsabbildung auf \mathbb{R}^n .

BEWEIS. Wir zeigen $E_m A = A$. Sei also $A = (\alpha_{i,j})_{i=1, \dots, m, j=1, \dots, n}$. Dann ist der Koeffizient an der Stelle i, j der Matrix $E_m A$ nach Definition des Matrixproduktes

$$\sum_{k=1}^m \delta_{i,k} \alpha_{k,j}.$$

Nun ist $\delta_{i,k}$ immer Null, ausser wenn $k = i$ ist, und daher bleibt von der ganzen Summe nur übrig

$$\sum_{k=1}^m \delta_{i,k} \alpha_{k,j} = \delta_{i,i} \alpha_{i,j} = \alpha_{i,j}.$$

Die Gleichung $A E_n = A$ geht ebenso. □

Manchmal ist der Beweis für einen sehr einfachen Sachverhalt schwierig formal aufzuschreiben. Wenn Ihnen nicht einleuchtet, was da oben steht, rechnen Sie einmal

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}.$$

Es ist grundsätzlich oft hilfreich für das Verständnis, wenn man abstrakte Beweise anhand einfacher Beispiele nachvollzieht.

DEFINITION 6.4.7. Sei $A \in \mathbb{R}^{n \times n}$ eine quadratische Matrix. Die Matrix $B \in \mathbb{R}^{n \times n}$ heißt inverse Matrix zu A wenn gilt:

$$AB = BA = E_n.$$

Wir schreiben in diesem Fall: $A^{-1} := B$.

Der folgende einfache Satz rechtfertigt die Schreibweise A^{-1} für die inverse Matrix:

SATZ 6.4.8.

- 1) Zu jeder Matrix kann es höchstens eine Inverse geben.
- 2) Besitzt eine quadratische Matrix A eine Inverse A^{-1} , so ist die Inverse zu A^{-1} wiederum A , also $(A^{-1})^{-1} = A$.

BEWEIS. Punkt (1): Seien B und C inverse Matrizen zu A . Dann ist

$$B = E_n B = (CA)B = C(AB) = CE_n = C.$$

Punkt (2): Trivial auf Grund der Definition der Inversen. □

BEMERKUNG 6.4.9. Die Beschränkung auf quadratische Matrizen ergibt sich ganz natürlich: Ist $A \in \mathbb{R}^{m \times n}$ und $B \in \mathbb{R}^{n \times m}$ so dass $AB = E_m$ und $BA = E_n$, dann ist $m = n$.

BEWEIS.

$$m = \rho(E_m) = \rho(AB) = \rho(AE_n B) \leq \rho(E_n) = n$$

und ebenso gilt $n \leq m$. □

BEISPIEL 6.4.10. Das folgende Beispiel zeigt, dass nicht jede quadratische Matrix eine Inverse besitzt: Sei $A = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$. Dann besitzt A keine inverse Matrix.

BEWEIS. Angenommen, A besitzt eine inverse Matrix A^{-1} . Man kann leicht nachrechnen: $A^2 := A \cdot A = 0$. Dann wäre aber

$$A = E_n A = (A^{-1} A) A = A^{-1} A^2 = A^{-1} 0 = 0,$$

aber $A \neq 0$, und wir haben einen Widerspruch. □

DEFINITION 6.4.11. Eine quadratische Matrix, heißt regulär, wenn sie eine Inverse besitzt. Andernfalls heißt die Matrix singular.

SATZ 6.4.12. Seien $A, B \in \mathbb{R}^{n \times n}$ zwei reguläre Matrizen. Dann ist auch AB regulär, und $(AB)^{-1} = B^{-1} A^{-1}$. (Achtung, die Reihenfolge dreht sich um!)

BEWEIS.

$$\begin{aligned} (AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} = AE_n A^{-1} = AA^{-1} = E_n, \\ (B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}A)B = B^{-1}E_n B = B^{-1}B = E_n. \end{aligned}$$

□

BEMERKUNG 6.4.13. Sei $A \in \mathbb{R}^{n \times n}$ eine reguläre Matrix. Dann ist auch die transponierte Matrix A^T regulär, und es gilt: $(A^T)^{-1} = (A^{-1})^T$.

BEWEIS.

$$A^T (A^{-1})^T = (A^{-1} A)^T = E_n^T = E_n, \quad (A^{-1})^T A^T = (A A^{-1})^T = E_n^T = E_n.$$

□

SATZ 6.4.14. Sei \mathbb{R} ein Körper, $A \in \mathbb{R}^{n \times n}$ eine quadratische Matrix, und

$$f : \begin{cases} \mathbb{R}^n & \rightarrow \mathbb{R}^n \\ x & \mapsto Ax \end{cases}$$

die dazugehörige lineare Abbildung

Dann ist A genau dann regulär, wenn f bijektiv ist. In diesem Fall gehört die Matrix A^{-1} zur Umkehrabbildung f^{-1} von f .

BEWEIS. Sei zunächst A regulär. Wir definieren die lineare Abbildung

$$g : \begin{cases} \mathbb{R}^n & \rightarrow \mathbb{R}^n, \\ x & \mapsto A^{-1}x. \end{cases}$$

Dann ist für $x \in \mathbb{R}^n$

$$\begin{aligned} [f \circ g](x) &= A(A^{-1}x) = (AA^{-1})x = E_n x = x, \\ [g \circ f](x) &= A^{-1}(Ax) = (A^{-1}A)x = E_n x = x. \end{aligned}$$

Also ist

$$f \circ g = g \circ f = \text{id}_{\mathbb{R}^n}$$

die Identität auf \mathbb{R}^n . Daher ist g die Umkehrabbildung zu f .

Sei nun f bijektiv und f^{-1} die Umkehrabbildung zu f . Wir wissen, dass zu f^{-1} eine Matrix B gehört, sodass für alle $x \in \mathbb{R}^n$ gilt $f^{-1}(x) = Bx$. Dann ist BA die Matrix zur Hintereinanderausführung $f^{-1} \circ f$, also zur Identität $\text{id}_{\mathbb{R}^n}$, also ist $BA = E_n$. Ebenso ist AB die Matrix zu $f \circ f^{-1} = \text{id}_{\mathbb{R}^n}$, also $AB = E_n$. Damit ist B die inverse Matrix zu A . \square

SATZ 6.4.15. Sei \mathbb{R} ein Körper und $A \in \mathbb{R}^{n \times n}$ eine $n \times n$ -Matrix. Dann sind folgende Aussagen äquivalent:

- 1) Der Rang von A ist n .
- 2) Der Spaltenraum von A ist der ganze Raum \mathbb{R}^n .
- 3) Die Spalten von A sind eine Basis des \mathbb{R}^n .
- 4) Die Spalten von A sind linear unabhängig.
- 5) Der Zeilenraum von A ist der ganze Raum $\mathbb{R}^{1 \times n}$.
- 6) Die Zeilen von A sind eine Basis des $\mathbb{R}^{1 \times n}$.
- 7) Die Zeilen von A sind linear unabhängig.
- 8) Der Kern von A ist der Nullraum.
- 9) A ist regulär.

BEWEIS. (1) \Rightarrow (2): Sei $\rho(A) = n$. Nach Definition des Ranges hat also der Spaltenraum von A die Dimension n . Weil aber $\text{rg}(A)$ also ein n -dimensionaler Unterraum des n -dimensionalen Raumes \mathbb{R}^n ist, gilt wegen Satz 4.3.7 Gleichheit: $\text{rg}(A) = \mathbb{R}^n$.

(2) \Rightarrow (3): Ist $\text{rg}(A) = \mathbb{R}^n$, so bilden die Spalten von A ein Erzeugendensystem des \mathbb{R}^n . Aus diesem läßt sich jedenfalls eine Basis auswählen (Satz 4.2.3). Da aber jede Basis des \mathbb{R}^n aus genau n Vektoren besteht, müssen alle Spalten bleiben. Sie bilden also ein minimales Erzeugendensystem, eine Basis, des \mathbb{R}^n .

(3) \Rightarrow (4): Jede Basis ist linear unabhängig.

(4) \Rightarrow (1): Seien die Spalten von A linear unabhängig. Wegen Satz 4.3.3 ist die Dimension des Spaltenraumes von A also mindestens n . Da der Spaltenraum von A aber ein Teilraum von \mathbb{R}^n ist, kann seine Dimension auch nicht größer als n sein. Wir haben also $\rho(A) = \dim(\text{rg}(A)) = n$.

(1) \Rightarrow (5) \Rightarrow (6) \Rightarrow (7) \Rightarrow (1): Genau wie die eben bewiesenen Äquivalenzen, unter Berücksichtigung, dass der Rang von A zugleich die Dimension des Zeilenraumes von A ist.

(1) \Leftrightarrow (8): Der Kern von A hat die Dimension $\dim(\ker(A)) = n - \rho(A)$, und diese ist genau dann 0, wenn $\rho(A) = n$ ist.

(1) \Leftrightarrow (9): Sei f die zu A gehörige lineare Abbildung

$$f : \begin{cases} \mathbb{R}^n & \rightarrow \mathbb{R}^n \\ x & \mapsto Ax. \end{cases}$$

Sei zunächst $\rho(A) = n$. Dann ist (wegen Punkt(2)) der Spaltenraum von A der ganze \mathbb{R}^n , also ist f surjektiv. Wegen Punkt (8) ist der Kern von f der Nullraum, also ist f injektiv (Satz 5.4.4). Also ist f bijektiv und nach Satz 6.4.14 ist A regulär. Ist umgekehrt A regulär, so ist f bijektiv, insbesondere surjektiv. Damit ist der Spaltenraum von A der ganze Raum \mathbb{R}^n . Das ist Punkt (2), und aus diesem folgt Punkt (1). \square

SATZ 6.4.16. Sei $A \in \mathbb{R}^{n \times n}$ eine quadratische Matrix. Dann sind äquivalent:

- (1) A ist regulär.
- (2) Für jedes $b \in \mathbb{R}^n$ hat das Gleichungssystem $Ax = b$ höchstens eine Lösung.
- (3) Für jedes $b \in \mathbb{R}^n$ hat das Gleichungssystem $Ax = b$ mindestens eine Lösung.
- (4) Für jedes $b \in \mathbb{R}^n$ ist das Gleichungssystem $Ax = b$ eindeutig lösbar.

In diesem Fall ist die Lösung des Gleichungssystems $Ax = b$ gegeben durch

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

BEWEIS. Sei f die zu A gehörende lineare Abbildung. Nach Satz 6.4.15 sind äquivalent

- (i) A ist regulär.
- (ii) $\text{rg}(A) = \mathbb{R}^n$
- (iii) $\ker(A) = \{0\}$.

Übersetzt in die Sprache der linearen Gleichungssysteme bedeutet aber $\text{rg}(A) = \mathbb{R}^n$, dass für jedes $b \in \mathbb{R}^n$ das Gleichungssystem $Ax = b$ mindestens eine Lösung hat. Die Aussage $\ker(A) = \{0\}$ bedeutet nach Satz 5.4.4, dass für jedes $b \in \mathbb{R}^n$ mindestens eine Lösung hat.

Sei nun A regulär. Wir lösen $Ax = b$.

$$Ax = b \Rightarrow A^{-1}(Ax) = A^{-1}b \Rightarrow E_n x = A^{-1}b \Rightarrow x = A^{-1}b.$$

Also ist $A^{-1}b$ der einzige Vektor, der als Lösung in Frage kommt. Weil es aber genau eine Lösung gibt, ist also $x = A^{-1}b$ die eindeutige Lösung. \square

Um ein einzelnes Gleichungssystem $Ax = b$ zu lösen, zahlt sich aber die Berechnung von A^{-1} nicht aus, sie macht mehr Arbeit als die Lösung des Gleichungssystems. Will man dagegen das System für viele verschiedene rechte Seiten lösen, ist es schon sinnvoll, die Inverse zu berechnen. Der Hauptnutzen des Satzes besteht aber in seiner Anwendung als Hilfsmittel für Beweise und zur Konstruktion weiterer Formeln und Rechenregeln.

6.5. Berechnung der inversen Matrix.

Wir werden jetzt einen Algorithmus zur Berechnung der inversen Matrix durch Pivotschritte entwickeln. Zuerst zeigen wir aber, dass die Zeilentransformationen selbst durch Matrixmultiplikationen beschrieben werden können.

DEFINITION 6.5.1. Sei $m \in \mathbb{N}$, $\lambda \in \mathbb{K} \setminus \{0\}$, $i \neq j \in \{1, \dots, m\}$. Wir definieren folgende Matrizen in $\mathbb{K}^{m \times m}$, die sogenannten Elementarmatrizen:

Wir haben zur besseren Orientierung die Zeilen- und Spaltennummern an den linken und oberen Rand der Matrizen geschrieben, und alle nicht eingefüllten Koeffizienten sind wie in der Einheitsmatrix. Unterstrichen sind die Koeffizienten, die sich von der Einheitsmatrix unterscheiden. Die Reihenfolge der Zeilen und Spalten von i und j kann auch umgekehrt sein, je nachdem ob $i < j$ oder $j < i$.

$$T_{\text{tausch}}(i, j) := \begin{pmatrix} & (1) & & (i) & & (j) & & (m) \\ (1) & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ & \vdots & \ddots & \vdots & & \vdots & & \vdots \\ (i) & 0 & \cdots & \underline{0} & \cdots & \underline{1} & \cdots & 0 \\ & \vdots & & \vdots & \ddots & \vdots & & \vdots \\ (j) & 0 & \cdots & \underline{1} & \cdots & \underline{0} & \cdots & 0 \\ & \vdots & & \vdots & & \vdots & \ddots & \vdots \\ (m) & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}$$

$$T_{\text{add}}(i, j, \lambda) := \begin{pmatrix} & (1) & & (i) & & (j) & & (m) \\ (1) & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ & \vdots & \ddots & \vdots & & \vdots & & \vdots \\ (i) & 0 & \cdots & 1 & \cdots & \underline{\lambda} & \cdots & 0 \\ & \vdots & & \vdots & \ddots & \vdots & & \vdots \\ (j) & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ & \vdots & & \vdots & & \vdots & \ddots & \vdots \\ (m) & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}$$

$$T_{\text{mult}}(i, \lambda) := \begin{pmatrix} & (1) & & (i) & & (j) & & (m) \\ (1) & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ & \vdots & \ddots & \vdots & & \vdots & & \vdots \\ (i) & 0 & \cdots & \underline{\lambda} & \cdots & 0 & \cdots & 0 \\ & \vdots & & \vdots & \ddots & \vdots & & \vdots \\ (j) & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ & \vdots & & \vdots & & \vdots & \ddots & \vdots \\ (m) & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}$$

SATZ 6.5.2. Die elementaren Zeilenumformungen (siehe Definition 2.2.1) können mit Hilfe der Matrixmultiplikation mit Elementarmatrizen geschrieben werden: Sind $m, n \in \mathbb{N}$, $A \in \mathbb{R}^{m \times n}$ und die $m \times m$ Elementarmatrizen wie in Definition 6.5.1 gegeben. Sei $\lambda \in \mathbb{R} \setminus \{0\}$, $i \neq j \in \mathbb{N}$. Dann gilt:

- 1) $T_{\text{tausch}}(i, j) \cdot A$ entsteht aus A indem man die i -te und die j -te Zeile von A vertauscht.
- 2) $T_{\text{add}}(i, j, \lambda) \cdot A$ entsteht aus A indem man das λ -fache der j -ten Zeile zur i -ten Zeile von A addiert.
- 3) $T_{\text{mult}}(i, \lambda) \cdot A$ entsteht aus A indem man die i -te Zeile von A mit λ multipliziert.

Ebenso erhält man durch Multiplikation mit Elementarmatrizen von rechts folgende Spaltentransformationen:

- 1) $A \cdot T_{\text{tausch}}(i, j)$ entsteht aus A indem man die i -te und die j -te Spalte von A vertauscht.

- 2) $A \cdot T_{\text{add}}(i, j, \lambda)$ entsteht aus A indem man das λ -fache der i -ten Spalte zur j -ten Spalte von A addiert.
- 3) $A \cdot T_{\text{mult}}(i, \lambda)$ entsteht aus A indem man die i -te Spalte von A mit λ multipliziert.

BEWEIS. Nachrechnen. □

BEMERKUNG 6.5.3. Die Matrix einer elementaren Zeilen- oder Spaltentransformation erhält man einfach, indem man diese Transformation auf die Einheitsmatrix anwendet.

BEHAUPTUNG 6.5.4. Die Elementarmatrizen sind regulär:

$$\begin{aligned} T_{\text{tausch}}(i, j)^{-1} &= T_{\text{tausch}}(i, j), \\ T_{\text{add}}(i, j, \lambda)^{-1} &= T_{\text{add}}(i, j, -\lambda), \\ T_{\text{mult}}(i, \lambda)^{-1} &= T_{\text{mult}}\left(i, \frac{1}{\lambda}\right). \end{aligned}$$

BEWEIS. Das sieht man, indem man überlegt, mit welchen Zeilentransformationen die jeweiligen Transformationen rückgängig gemacht werden können. □

Wir wenden uns jetzt der Berechnung der Inversen einer quadratischen Matrix zu.

LEMMA 6.5.5. Seien $A, B \in \mathbb{R}^{n \times n}$ quadratische Matrizen. Wenn gilt

$$BA = E_n,$$

dann sind A und B regulär und $B = A^{-1}$, $A = B^{-1}$.

Das heißt, es muss nicht erst überprüft werden, ob auch $BA = E_n$.

BEWEIS. Weil A und B nur n Zeilen haben, kann ihr Rang nicht größer als n sein. Andererseits gilt $n = \rho(E_n) = \rho(BA)$, sodass wegen Behauptung 6.3.10 die Ränge von A und B nicht kleiner als n sein können. Daher ist $\rho(A) = \rho(B) = n$, und daher sind die beiden Matrizen regulär. Insbesondere besitzt also A eine Inverse A^{-1} . Es ist

$$B = BE_n = B(AA^{-1}) = (BA)A^{-1} = E_n A^{-1} = A^{-1}.$$

Damit ist auch $B^{-1} = A$. □

ALGORITHMUS 6.5.6 (Berechnung der inversen Matrix). Sei A eine $n \times n$ -Matrix über \mathbb{R} . Der folgende Algorithmus entscheidet, ob A regulär ist, und berechnet gegebenenfalls die inverse Matrix A^{-1} :

- 1) Schreibe nebeneinander die Matrix A und die Einheitsmatrix E_n .
- 2) Forme die linke Seite A durch Pivotschritte um, solange man Pivotelemente findet. Führe gleichzeitig alle Zeilenumformungen auch auf der rechten Seite durch.
- 3) Wenn auf der linken Seite eine Nullzeile entsteht, ist A singulär, und es gibt keine Inverse. Stop.
- 4) Wenn auf der linken Seite keine Nullzeile entsteht, waren alle Zeilen Pivotelemente. Verwende nun Rücksubstitution, um überall ausser auf den Pivotelementen Nullen zu erzeugen. (Wieder werden alle Zeilentransformationen zugleich auf die rechte Seite ausgeführt.)
- 5) Dividiere jede Zeile durch ihr Pivotelement. Auf der linken Seite stehen nun nur mehr Einsen und Nullen. (Wieder werden die Divisionen auch auf der rechten Seite ausgeführt.)

- 6) Ordne die Zeilen so um, dass auf der linken Seite die Einheitsmatrix steht. (Wieder werden die Zeilen rechts ebenso umgeordnet.) Rechts steht dann die Inverse zu A .

BEWEIS. Lässt sich A durch Zeilentransformationen in eine Matrix \tilde{A} überführen, die eine Nullzeile enthält, so ist $\rho(A) = \rho(\tilde{A}) < n$ und daher ist A nicht regulär. Wenn man auf keine Nullzeile stößt, findet man n Pivotelemente und kann A in die Einheitsmatrix überführen. Die rechte Seite, die nun entstanden ist, nennen wir B . Man hat dann elementare Zeilentransformationen mit Elementarmatrizen T_1, \dots, T_k ausgeführt, sodass

$$T_k T_{k-1} \cdots T_1 A = E_n.$$

Zugleich wurde die Einheitsmatrix in die neue rechte Seite übergeführt:

$$B = (T_k T_{k-1} \cdots T_1 E_n).$$

Damit ist aber

$$BA = T_k \cdots T_1 A = E_n$$

und nach Lemma 6.5.5 ist A regulär und $B = A^{-1}$. \square

BEISPIEL 6.5.7. Ist die folgende Matrix A regulär? Wenn ja, berechne die Inverse:

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

Lösung:

1.) Pivotschritte:

$$\begin{pmatrix} 1 & -1 & 1 & -1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 3 & -1 & 0 & 1 & 0 & 0 \\ 3 & -3 & 4 & -2 & 0 & 0 & 1 & 0 \\ 1* & 0 & 2 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & -1 & -1 & -1 & 1 & 0 & 0 & -1 \\ 0 & 1* & 1 & -1 & 0 & 1 & 0 & -1 \\ 0 & -3 & -2 & -2 & 0 & 0 & 1 & -3 \\ 1 & 0 & 2 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 0 & 0 & 0 & -2* & 1 & 1 & 0 & -2 \\ 0 & 1 & 1 & -1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1* & -5 & 0 & 3 & 1 & -6 \\ 1 & 0 & 2 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Der Rang ist 4, es gibt eine Inverse!

2.) Rücksubstitutionsschritte und dividieren durch Pivotelemente:

$$\begin{pmatrix} 0 & 0 & 0 & -2* & 1 & 1 & 0 & -2 \\ 0 & 1 & 1 & -1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -5 & 0 & 3 & 1 & -6 \\ 1 & 0 & 2 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 0 & 0 & 1 & -0.5 & -0.5 & 0 & 1 \\ 0 & 1 & 1 & 0 & -0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 1* & 0 & -2.5 & 0.5 & 1 & -1 \\ 1 & 0 & 2 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} 0 & 0 & 0 & 1 & -0.5 & -0.5 & 0 & 1 \\ 0 & 1 & 0 & 0 & 2 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 & -2.5 & 0.5 & 1 & -1 \\ 1 & 0 & 0 & 0 & 5 & -1 & -2 & 3 \end{pmatrix}$$

Umordnen der Zeilen:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 5 & -1 & -2 & 3 \\ 0 & 1 & 0 & 0 & 2 & 0 & -1 & 1 \\ 0 & 0 & 1 & 0 & -2.5 & 0.5 & 1 & -1 \\ 0 & 0 & 0 & 1 & -0.5 & -0.5 & 0 & 1 \end{pmatrix}$$

Die Inverse von A ist

$$A^{-1} = \begin{pmatrix} 5 & -1 & -2 & 3 \\ 2 & 0 & -1 & 1 \\ -2.5 & 0.5 & 1 & -1 \\ -0.5 & -0.5 & 0 & 1 \end{pmatrix}.$$

□

KOROLLAR 6.5.8. *Jede reguläre Matrix lässt sich als Produkt von Elementarmatrizen schreiben.*

BEWEIS. Wie wir im Beweis zu Algorithmus 6.5.6 gesehen haben, lässt sich jede reguläre Matrix durch Elementarmatrizen in die Einheitsmatrix überführen:

$$T_k \cdots T_1 A = E_n.$$

Dann ist

$$A = T_1^{-1} \cdots T_k^{-1} E_n = T_1^{-1} \cdots T_k^{-1}.$$

Weil aber die Inversen der Elementarmatrizen wieder Elementarmatrizen sind, ist also A das Produkt von Elementarmatrizen. □

KAPITEL 2

Körper

1. Körperaxiome

1.1. Grundrechnungsarten.

Wir haben im vorigen Kapitel die Matrizenrechnung über den reellen Zahlen eingeführt. Über die reellen Zahlen weiß man sehr viel. Wir führen diese Eigenschaften nur an, um uns bewusst zu machen, dass \mathbb{R} eine äußerst reiche und komplizierte Struktur hat, die weit über die Grundrechnungsarten hinaus geht. Manches davon haben Sie vielleicht schon in Analysis genau diskutiert oder werden es diskutieren.

- Auf \mathbb{R} sind die Grundrechnungsarten $+, -, \cdot, /$ mit den bekannten Rechenregeln eingeführt.
- Auf \mathbb{R} gibt es eine Ordnungsstruktur \leq , die mit den Grundrechnungsarten verträglich ist.
- Wenn eine Teilmenge $A \subset \mathbb{R}$ eine obere Schranke besitzt, so besitzt A auch eine kleinste obere Schranke.
- Jede Cauchyfolge in \mathbb{R} besitzt einen Grenzwert.
- Die Menge der reellen Zahlen ist überabzählbar.
- usw.

Von allen diesen Eigenschaften der reellen Zahlen haben wir bei der Entwicklung der Matrizenrechnung nur die Grundrechnungsarten und ihre Rechenregeln verwendet. Das heißt, um Matrizen- und Vektorrechnung zu betreiben, brauchen wir nur eine Menge und darauf die vier Grundrechnungsarten. Im Begriff des Körpers sammeln wir alle Eigenschaften, die wir zur Einführung der Matrizenrechnung benötigen.

Die Menge, auf der wir arbeiten, muss nicht einmal aus dem bestehen, was wir als Zahlen kennen. Und eine Rechenvorschrift auf M ist nichts anderes als eine Abbildung, die zwei Elementen aus M ein neues Element auf M zuordnet:

DEFINITION 1.1.1. Sei M eine nichtleere Menge und \circ eine Abbildung

$$\begin{cases} M \times M & \rightarrow M \\ (a, b) & \mapsto a \circ b \end{cases} .$$

Dann heißt \circ eine innere Verknüpfung auf M .

BEMERKUNG 1.1.2.

- (1) Das Symbol \circ steht in Definition 1.1.1 für irgendeine Verknüpfung und muss nicht die Hintereinanderausführung von Funktionen bedeuten.
- (2) Wir reden von einer inneren Verknüpfung, weil alles innerhalb einer Menge passiert: Die Elemente $a, b \in M$ werden verknüpft, und es entsteht wieder ein Element $a \circ b$ aus M .

DEFINITION 1.1.3 (Körper). Sei \mathbb{K} eine Menge, welche mindestens zwei verschiedene Elemente 0 und 1 enthält¹. Auf \mathbb{K} seien zwei innere Verknüpfungen gegeben (die sogenannte Körperaddition und Körpermultiplikation):

$$+ : \begin{cases} \mathbb{K} \times \mathbb{K} & \rightarrow \mathbb{K} \\ (a, b) & \mapsto a + b \end{cases}, \quad \cdot : \begin{cases} \mathbb{K} \times \mathbb{K} & \rightarrow \mathbb{K} \\ (a, b) & \mapsto a \cdot b \end{cases}.$$

Das Quintupel $(\mathbb{K}, +, \cdot, 0, 1)$ heißt ein Körper, wenn folgende Eigenschaften gelten:

- Für $+$:
 - (plus 1) Assoziativgesetz: $(\forall x, y, z \in \mathbb{K}) (x + y) + z = x + (y + z)$.
 - (plus 2) 0 ist neutrales Element für $+$, das heißt: $(\forall x \in \mathbb{K}) x + 0 = x$.
 - (plus 3) Zu jedem $x \in \mathbb{K}$ gibt es ein inverses Element $y \in \mathbb{K}$ sodass $x + y = 0$.
 - (plus 4) Kommutativgesetz: $(\forall x, y \in \mathbb{K}) x + y = y + x$.
- Für \cdot :
 - (mal 1) Assoziativgesetz: $(\forall x, y, z \in \mathbb{K}) (x \cdot y) \cdot z = x \cdot (y \cdot z)$.
 - (mal 2) 1 ist neutrales Element für \cdot , das heißt: $(\forall x \in \mathbb{K}) x \cdot 1 = x$.
 - (mal 3) Zu jedem $x \in \mathbb{K} \setminus \{0\}$ gibt es ein inverses Element $y \in \mathbb{K}$ sodass $x \cdot y = 1$.
 - (mal 4) Kommutativgesetz: $(\forall x, y \in \mathbb{K}) x \cdot y = y \cdot x$.
- Für das Zusammenspiel von $+$ und \cdot :
 - (dist) Distributivgesetz²: $(\forall x, y, z \in \mathbb{K}) x \cdot (y + z) = x \cdot y + x \cdot z$.

In der Definition des Körpers haben wir nur soviel zusammengefasst, wie nötig ist, um die Existenz der Grundrechnungsarten und ihre Eigenschaften daraus herzuleiten. Zunächst führen wir eine Körpersubtraktion und eine Körperdivision ein.

BEHAUPTUNG 1.1.4. Sei $(\mathbb{K}, +, \cdot, 0, 1)$ ein Körper. Dann gilt

- (a) Zu jedem $a \in \mathbb{K}$ gibt es genau ein inverses Element bezüglich der Körperaddition.
- (b) Zu jedem $a \in \mathbb{K} \setminus \{0\}$ gibt es genau ein inverses Element bezüglich der Körpermultiplikation.

BEWEIS.

- (a) Wegen Körperaxiom (plus 3) besitzt a mindestens ein inverses Element bezüglich $+$. Seien b und c invers zu a bezüglich $+$. Dann ist

$$b = b + 0 = b + (a + c) = (b + a) + c = 0 + c = c.$$

Also gibt es genau ein inverses Element.

- (b) Sei $a \neq 0$. Wegen Körperaxiom (mal 3) besitzt a mindestens ein inverses Element bezüglich \cdot . Seien b, c invers zu a bezüglich \cdot . Dann ist

$$b = b \cdot 1 = b \cdot (a \cdot c) = (b \cdot a) \cdot c = 1 \cdot c = c.$$

□

Die Eindeutigkeit der inversen Elemente rechtfertigt die Verwendung der folgenden Schreibweisen $(-x)$ und x^{-1} . Beachten Sie, dass das nur Schreibweisen sind, die passende Assoziationen zu unseren Rechengewohnheiten aus \mathbb{R} wecken. Wir setzen nicht voraus, dass auf \mathbb{K} bereits eine Subtraktion und eine Division definiert sind. Im Gegenteil, durch die Existenz der inversen Elemente können wir anschließend Subtraktion und Division definieren.

DEFINITION 1.1.5. Sei $(\mathbb{K}, +, \cdot, 0, 1)$ ein Körper, und $a \in \mathbb{K}$.

¹Diese Elemente 0 und 1 müssen nicht die Zahlen 0 und 1 sein. Diese Symbole sind so gewählt, dass sie passende Assoziationen zu unseren gewohnten Rechenregeln in \mathbb{R} wecken. Es muss auch \mathbb{K} selbst nicht unbedingt eine Menge von Zahlen sein.

²Wir verwenden die Konvention "Punktrechnung vor Strichrechnung", die wir von den reellen Zahlen gewöhnt sind. Sonst müssten wir genauer schreiben: $(\forall x, y, z \in \mathbb{K}) x \cdot (y + z) = (x \cdot y) + (x \cdot z)$.

Das Symbol $(-a)$ bezeichnet das inverse Element zu a bezüglich der Körperaddition.

Ist $a \neq 0$, so bezeichnet a^{-1} das inverse Element von a bezüglich der Körpermultiplikation.

Weil also die inversen Elemente eindeutig bestimmt sind, können wir eindeutig eine Körpersubtraktion und eine Körperdivision durch die folgende Definition einführen:

DEFINITION 1.1.6 (Körpersubtraktion und -division). Sei $(\mathbb{K}, +, \cdot, 0, 1)$ ein Körper. Wir definieren Abbildungen (die sogenannte Körpersubtraktion und die Körperdivision) durch

$$\left\{ \begin{array}{l} \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K} \\ a - b \end{array} \right. := a + (-b) \quad , \quad \left\{ \begin{array}{l} \mathbb{K} \times (\mathbb{K} \setminus \{0\}) \rightarrow \mathbb{K} \\ \frac{a}{b} \end{array} \right. := a \cdot (b^{-1}) \quad .$$

Wir haben nun auf dem Körper alle vier Grundrechnungsarten eingeführt. Man kann nun zeigen, dass viele der Rechenregeln, die wir aus \mathbb{R} gewohnt sind, auch in jedem anderen Körper gelten. Wir werden das an verschiedenen Beispielen vorführen. Versuchen Sie selbst, andere bekannte Rechenregeln darauf zu überprüfen, ob sie in jedem Körper gelten.

BEHAUPTUNG 1.1.7 (Kürzungsregel). Sei $(\mathbb{K}, +, \cdot, 0, 1)$ ein Körper, seien $a, x, y \in \mathbb{K}$, $p \in \mathbb{K} \setminus \{0\}$. Dann gilt

- (a) $a + x = a + y \Rightarrow x = y$,
 (b) $p \cdot x = p \cdot y \Rightarrow x = y$.

BEWEIS.

- (a) Wegen (plus 3) existiert das inverse Element $(-a)$ zu a bezüglich der Addition. Es gilt

$$\begin{aligned} & a + x = a + y \\ \Rightarrow (-a) + (a + x) &= (-a) + (a + y) && \text{links und rechts dasselbe addiert} \\ \Rightarrow ((-a) + a) + x &= ((-a) + a) + y && \text{(Assoziativität von +)} \\ \Rightarrow 0 + x = 0 + y & && \text{weil } (-a) \text{ invers zu } a \text{ ist} \\ \Rightarrow x = y & && \text{(Weil 0 bezüglich + neutral ist.)} \end{aligned}$$

- (b) Weil nach Voraussetzung $p \neq 0$ gilt, gibt es das inverse Element p^{-1} zu p bezüglich der Multiplikation. Es gilt

$$\begin{aligned} & p \cdot x = p \cdot y \\ \Rightarrow p^{-1} \cdot (p \cdot x) &= p^{-1} \cdot (p \cdot y) && \text{links und rechts dasselbe multipliziert} \\ \Rightarrow (p^{-1} \cdot p) \cdot x &= (p^{-1} \cdot p) \cdot y && \text{Assoziativität von } \cdot \\ \Rightarrow 1 \cdot x = 1 \cdot y & && \text{weil } p^{-1} \text{ invers zu } p \text{ ist} \\ \Rightarrow x = y & && \text{weil 1 neutral bezüglich } \cdot \text{ ist} \end{aligned}$$

□

BEHAUPTUNG 1.1.8. Sei $(\mathbb{K}, +, \cdot, 0, 1)$ ein Körper, seien $x, y \in \mathbb{K}$. Dann gilt

$$x \cdot y = 0 \iff (x = 0 \text{ oder } y = 0).$$

BEWEIS.

\Leftarrow : Sei o.B.d.A. $x = 0$. Wir müssen zeigen: $0 \cdot y = 0$. (Der Fall $y = 0$ geht genauso.) Weil 0 das neutrale Element bezüglich der Körperaddition ist, gilt $0 + 0 = 0$. Mit dem Distributivgesetz erhalten wir

$$0 \cdot y + 0 \cdot y = (0 + 0) \cdot y = 0 \cdot y = 0 \cdot y + 0.$$

Mit Hilfe der Kürzungsregel für die Addition erhalten wir daraus

$$0 \cdot y = 0.$$

\Rightarrow : Sei $x \cdot y = 0$. Wir müssen zeigen: $x = 0$ oder $y = 0$. Wenn $x = 0$ ist, ist nichts mehr zu zeigen. Wir betrachten also nur noch den Fall $x \neq 0$. Wir wissen bereits von der anderen Beweisrichtung, dass $x \cdot 0 = 0$. Also haben wir die Gleichung

$$x \cdot y = 0 = x \cdot 0.$$

Weil $x \neq 0$, können wir die Kürzungsregel für die Multiplikation anwenden und erhalten

$$y = 0.$$

□

BEHAUPTUNG 1.1.9. Sei $(\mathbb{K}, +, \cdot, 0, 1)$ ein Körper. Seien $a, b, c \in \mathbb{K}$, $p, q \in \mathbb{K} \setminus \{0\}$. Dann gelten folgende Regeln

(a) $(-(-a)) = a$ und $(p^{-1})^{-1} = p$.

(b) $(-a) = (-1) \cdot a$.

(c) $(p \cdot q)^{-1} = p^{-1} \cdot q^{-1}$.

(d) $a \cdot (b - c) = a \cdot b - a \cdot c$.

(e) $\frac{a}{p} \cdot \frac{b}{q} = \frac{a \cdot b}{p \cdot q}$.

BEWEIS.

- (a) Es ist $(-(-a))$ das eindeutige Element r mit der Eigenschaft $(-a) + r = 0$. Nun gilt aber gerade für a , dass $(-a) + a = 0$, weil $(-a)$ invers zu a ist. Also ist $a = (-(-a))$. Die Aussage $(p^{-1})^{-1} = p$ beweist man ebenso.
- (b) Es ist $(-a)$ das eindeutige Element s mit der Eigenschaft $a + s = 0$. Nun ist aber $a + (-1) \cdot a = (1 + (-1)) \cdot a = 0 \cdot a = 0$.
- (c) $(p \cdot q)^{-1}$ ist das eindeutige Element t mit der Eigenschaft $(p \cdot q) \cdot t = 1$. Nun ist aber $(p \cdot q) \cdot (p^{-1} \cdot q^{-1}) = (p \cdot p^{-1}) \cdot (q \cdot q^{-1}) = 1 \cdot 1 = 1$.
- (d) $a \cdot (b - c) = a \cdot (b + (-1) \cdot c) = a \cdot b + a \cdot (-1) \cdot c = a \cdot b + (-1) \cdot (a \cdot c) = a \cdot b - a \cdot c$.
- (e) $\frac{a}{p} \cdot \frac{b}{q} = a \cdot p^{-1} \cdot b \cdot q^{-1} = (a \cdot b) \cdot (p^{-1} q^{-1}) = (a \cdot b) \cdot (p \cdot q)^{-1} = \frac{a \cdot b}{p \cdot q}$.

□

BEMERKUNG 1.1.10. Diese Sätzchen sind sehr einfach, aber sie enthalten für Anfängerinnen und Anfänger eine Schwierigkeit: "Warum gibt es überhaupt etwas zu beweisen? Wir kennen doch alle diese Rechenregeln schon aus der Schulzeit!"

Die Antwort lautet: Wir kennen sie eben nicht. Wir haben sie für die rationalen Zahlen (Bruchzahlen) und später für die reellen Zahlen kennen gelernt. Aber jetzt arbeiten wir in einem Körper, über den wir nichts wissen, als was in Definition 1.1.3 festgehalten ist. Und wir müssen zeigen, dass allein aus dieser Definition alle die Regeln, die uns für \mathbb{R} wohlbekannt sind, auch für alle anderen Körper gefolgert werden können.

Die Körperaxiome für die Addition und die Multiplikation sind dieselben, mit Ausnahme der Sonderstellung des Nullelementes bei der Multiplikation. Im Begriff der Gruppe werden diese vier Axiome zusammengefasst. Gruppen sind in der Mathematik allgegenwärtig, haben vielseitige Anwendungen, und eine sehr reiche Theorie. Leider haben wir nicht einmal für die Anfangsgründe Zeit. Ich muss Sie auf eine Einführungsvorlesung aus Algebra vertrösten. Aber wenigstens die Definition einer Gruppe sollen Sie gesehen haben.

DEFINITION 1.1.11. Sei G eine nichtleere Menge und \circ eine innere Verknüpfung auf G . Das Paar (G, \circ) heißt eine Gruppe, wenn gilt

- Die Verknüpfung \circ ist assoziativ, also

$$(\forall x, y, z \in G) (x \circ y) \circ z = x \circ (y \circ z).$$

- Es gibt ein neutrales Element e bezüglich \circ , also ein Element mit der Eigenschaft

$$(\forall x \in G) x \circ e = e \circ x = x.$$

- Zu jedem $x \in G$ gibt es ein inverses Element $x^{-1} \in G$, also ein Element mit der Eigenschaft

$$x \circ x^{-1} = x^{-1} \circ x = e.$$

Wenn zusätzlich \circ kommutativ ist, dann heißt (G, \circ) eine kommutative Gruppe oder eine abelsche³ Gruppe.

BEMERKUNG 1.1.12. Häufig wird die Gruppenoperation auch mit einem Multiplikationspunkt \cdot statt \circ geschrieben, und dieser Punkt kann auch, wie bei der Multiplikation von Zahlen, weggelassen werden: $ab := a \cdot b$.

In kommutativen Gruppen wird die Gruppenoperation oft mit einem $+$ geschrieben. Dann bezeichnet man das neutrale Element mit 0 und das inverse Element zu x mit $(-x)$.

1.2. Spaltenvektoren über allgemeinen Körpern.

Wir können nun über jedem Körper, wie über den reellen Zahlen, Spaltenvektoren und Matrizen definieren.

DEFINITION 1.2.1. Sei $(\mathbb{K}, +, \cdot, 0, 1)$ ein Körper und seien $m, n \in \mathbb{N}$. Ein rechteckiges Schema aus m Zeilen und m Spalten von Elementen aus \mathbb{K} heißt eine $(m \times n)$ -Matrix über \mathbb{K} :

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix}$$

Eine $(m \times 1)$ -Matrix über \mathbb{K} heißt ein m -dimensionaler Spaltenvektor.

Mit \mathbb{K}^m bezeichnen wir die Menge der m -dimensionalen Spaltenvektoren über \mathbb{K} , und mit $\mathbb{K}^{m \times n}$ die Menge der $(m \times n)$ -Matrizen über \mathbb{K} .

Zur Unterscheidung von Vektoren und Matrizen bezeichnen wir die Elemente des Körpers \mathbb{K} als Skalare.

DEFINITION 1.2.2. Sei $(\mathbb{K}, +, \cdot, 0, 1)$ ein Körper. Für Matrizen über \mathbb{K} führen wir folgende Rechenoperationen ein:

- Seien $A = (\alpha_{ij})_{i=1 \dots m, j=1 \dots n}$ und $B = (\beta_{ij})_{i=1 \dots m, j=1 \dots n}$ Matrizen aus $\mathbb{K}^{m \times n}$, und es sei $\lambda \in \mathbb{K}$. Wir definieren

$$A \pm B := \begin{pmatrix} \alpha_{11} \pm \beta_{11} & \cdots & \alpha_{1n} \pm \beta_{11} \\ \vdots & & \vdots \\ \alpha_{m1} \pm \beta_{m1} & \cdots & \alpha_{mn} \pm \beta_{mn} \end{pmatrix},$$

$$\lambda A := \begin{pmatrix} \lambda \alpha_{11} & \cdots & \lambda \alpha_{1n} \\ \vdots & & \vdots \\ \lambda \alpha_{m1} & \cdots & \lambda \alpha_{mn} \end{pmatrix}.$$

³Niels Henrik Abel, norwegischer Mathematiker. Bitte den Namen aussprechen wie den Bruder von Kain.

Beachten Sie, dass mit dieser Definition (im Sonderfall $n = 1$) auch die Linearkombinationen von Spaltenvektoren definiert sind.

Die $(m \times n)$ -Nullmatrix über \mathbb{K} ist die Matrix

$$0_{m \times n} := \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \in \mathbb{K}^{m \times n},$$

dabei ist 0 das Nullelement im Körper \mathbb{K} . Der Nullvektor in \mathbb{K}^n ist die $(n \times 1)$ -Nullmatrix.

- Seien $A = (\alpha_{ij})_{i=1 \dots k, j=1 \dots m} \in \mathbb{K}^{k \times m}$ und $B = (\beta_{ij})_{i=1 \dots m, j=1 \dots n} \in \mathbb{K}^{m \times n}$. Dann ist

$$A \cdot B = \begin{pmatrix} \sum_{p=1}^m \alpha_{1,p} \beta_{p,1} & \cdots & \sum_{p=1}^m \alpha_{1,p} \beta_{p,n} \\ \vdots & & \vdots \\ \sum_{p=1}^m \alpha_{k,p} \beta_{p,1} & \cdots & \sum_{p=1}^m \alpha_{k,p} \beta_{p,n} \end{pmatrix} \in \mathbb{K}^{k \times n}.$$

Die $(n \times n)$ -Einheitsmatrix über \mathbb{K} ist die Matrix

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

dabei sind 0 und 1 das Nullelement und Einselement des Körpers \mathbb{K} .

SATZ 1.2.3. *Die Theorie der linearen Gleichungssysteme, die Matrizen- und Vektorrechnung über \mathbb{R} , wie sie im Kapitel 1 entwickelt wurde, gilt wörtlich genauso über allgemeinen Körpern.*

BEWEIS. Zur Entwicklung der Theorie wurden von den reellen Zahlen lediglich die Grundrechnungsarten und ihre Rechenregeln herangezogen. \square

BEISPIEL 1.2.4. Wir adaptieren zum Beispiel die Theorie der Linearen Unabhängigkeit.

Diskussion von Beispiel 1.2.4:

In dieser Diskussion sei also $(\mathbb{K}, +, \cdot, 0, 1)$ ein Körper.

Man kann definieren:

Seien (v_1, \dots, v_k) ein System von Spaltenvektoren des \mathbb{K}^n . Das System (v_1, \dots, v_k) heißt linear unabhängig, wenn es kein k -Tupel von Skalaren $(\lambda_1, \dots, \lambda_k)$ von Skalaren aus \mathbb{K} außer $(0, \dots, 0)$ gibt, sodass

$$\lambda_1 v_1 + \cdots + \lambda_k v_k = 0.$$

Das ist nichts anderes als Definition 4.1.2, nur müssen wir nun allgemeiner denken. Statt reeller Zahlen stehen jetzt Elemente irgendeines Körpers. 0 bedeutet, je nach Zusammenhang, das Nullelement des Körpers oder den Nullvektor in \mathbb{K}^n . Im Text der Definition mussten wir aber nur \mathbb{R} an allen Stellen durch \mathbb{K} ersetzen.

Wir zeigen am Beispiel einer Behauptung (nämlich (1) \Rightarrow (2) aus Satz 4.1.4), dass sich auch die Behauptungen und Beweise so gut wie wörtlich aus der reellen Theorie übernehmen lassen:

Seien $v_1, \dots, v_k \in \mathbb{K}^n$ linear unabhängig und $x \in \mathcal{L}(v_1, \dots, v_k)$. Dann gibt es nur ein eindeutiges k -Tupel $(\lambda_1, \dots, \lambda_k)$ von Skalaren aus \mathbb{K} , sodass $x = \sum_{i=1}^k \lambda_i v_i$.

Beweis: Natürlich ist das so zu verstehen, dass schon vorher der Begriff der linearen Hülle für Vektoren aus \mathbb{K}^n angepasst wurde (vgl. Definition 3.3.1(2):

$$\mathcal{L}(v_1, \dots, v_k) = \left\{ \sum_{i=1}^k \lambda_i v_i \mid \lambda_1, \dots, \lambda_k \in \mathbb{K} \right\}.$$

Sei nun

$$x = \sum_{i=1}^k \lambda_i v_i = \sum_{i=1}^k \mu_i v_i.$$

Dann ist

$$\begin{aligned}
 & \sum_{i=1}^k (\lambda_i - \mu_i) v_i \\
 &= \sum_{i=1}^k (\lambda_i v_i - \mu_i v_i) && \text{Distributivgesetz} \\
 &= \sum_{i=1}^k \lambda_i v_i + \sum_{i=1}^k (-\mu_i v_i) && \text{Ass., Komm. von +} \\
 &= \sum_{i=1}^k \lambda_i v_i + \sum_{i=1}^k (-1) \mu_i v_i && \text{1.1.9(b)} \\
 &= \sum_{i=1}^k \lambda_i v_i + (-1) \sum_{i=1}^k \mu_i v_i && \text{Distributivgesetz für Vektoren} \\
 &= x + (-1)x && \text{nach Voraussetzung} \\
 &= 0.
 \end{aligned}$$

Weil die Vektoren (v_1, \dots, v_k) linear unabhängig ist, folgt nun für alle $i = 1, \dots, k$:

$$(\lambda_i - \mu_i) = 0 \quad \text{also} \quad \lambda_i = \mu_i.$$

Damit ist die Behauptung bewiesen.

Wiederum sind Behauptung und Beweis fast wörtlich wie in Satz 4.1.4. Wir müssen uns nur vor Augen halten, dass die durchgeführten Rechenoperationen nun Operationen im Körper beziehungsweise in \mathbb{K}^n sind und die Beweisschritte sich auch mit den Axiomen eines allgemeinen Körpers rechtfertigen lassen.

Versuchen Sie selbst, verschiedene Definitionen und Behauptungen aus Kapitel 1 von den reellen Zahlen auf allgemeine Körper zu übertragen.

2. Beispiele von Körpern

2.1. Komplexe Zahlen.

Die Zahlenmengen \mathbb{N} und \mathbb{Z} der natürlichen und ganzen Zahlen sind keine Körper. In \mathbb{N} gibt es keine inversen Elemente bezüglich der Addition. In \mathbb{Z} hat jedes Element sein Inverses für die Addition, das heißt, ganze Zahlen kann man subtrahieren und erhält wieder ganze Zahlen. Die einzigen ganzen Zahlen, die inverse Elemente bezüglich der Multiplikation in \mathbb{Z} besitzen, sind aber ± 1 . Dividieren können wir erst, wenn wir Bruchzahlen zulassen. Der Körper der rationalen Zahlen \mathbb{Q} besteht aus allen Zahlen, die sich als Brüche von ganzen Zahlen schreiben lassen. Auf \mathbb{Q} gibt es erstmals alle vier Grundrechnungsarten. Weil \mathbb{Q} ein Körper ist, lässt sich über \mathbb{Q} die gesamte Vektor- und Matrizenrechnung, wie wir sie bisher kennen, entwickeln.

Die Erweiterung von \mathbb{Q} zu \mathbb{R} bringt für die Grundrechnungsarten keine Neuerungen, dafür können aber nun Wurzeln aus allen nichtnegativen Zahlen gezogen werden. In \mathbb{R} stehen Mittel der Analysis zur Verfügung, mit denen man zum Beispiel spezielle Funktionen (Exponentialfunktion, Winkelfunktionen) einführen kann. Der Begriff der Stetigkeit und des Grenzwertes kann zwar bereits auf \mathbb{Q} definiert werden, ihre Stärke entfalten diese Begriffe aber erst, wenn mit \mathbb{R} die Vollständigkeit ins Spiel kommt.

Wir setzen uns nun in den Kopf, auch die Wurzel aus (-1) zu ziehen, und lassen uns zunächst auf eine Heuristik ein:

BEISPIEL 2.1.1. Wenn es einen Körper \mathbb{K} gäbe, der die reellen Zahlen als Teilmenge enthält, und außerdem ein Element i mit der Eigenschaft $i^2 = -1$ enthält, dann müssten die vier Grundrechnungsarten für Ausdrücke der Form $(a + bi)$ mit

reellen Zahlen a, b folgendermaßen aussehen:

$$\begin{aligned}
 & (a + bi) \pm (c + di) \\
 &= a \pm c + bi \pm di && \text{Assoziativität und Kommutativität von } + \\
 &= (a \pm c) + (b \pm d)i && \text{Distributivität} \\
 \\
 & \overline{(a + bi)(c + di)} \\
 &= ac + bci + adi + i^2bd && \text{Distributivität} \\
 &= ac + bci + adi - bd && \text{weil } i^2 = -1 \\
 &= (ac - bd) + (bc + ad)i && \text{Distributivität} \\
 \\
 & \frac{a + bi}{c + di} \\
 &= \frac{(a + bi)(c - di)}{(c + di)(c - di)} && \text{Multiplizieren mit } 1 = \frac{c - di}{c - di} \\
 &= \frac{(ac - bdi^2) + (bc - ad)i}{c^2 - d^2i^2} && \text{Distributivgesetze} \\
 &= \frac{(ac + bd) + (bc - ad)i}{c^2 + d^2} && \text{weil } i^2 = -1
 \end{aligned}$$

Noch wissen wir nicht, dass es wirklich einen solchen Körper gibt. Weil wir aber gefunden haben, wie die Rechenregeln lauten müssten, können wir das Pferd beim Schwanz aufzäumen. Man definiert sich die komplexen Zahlen $a + ib$ einfach als Paare (a, b) von reellen Zahlen, wobei man sich zum Paar (a, b) denkt, dass es am Ende $a + bi$ bedeuten wird. Auf der Menge $\mathbb{R} \times \mathbb{R}$ der Paare führt man die passenden Rechenregeln ein. Dann zeigt man, dass auf diese Weise ein Körper entsteht.

DEFINITION 2.1.2. Auf der Menge $\mathbb{C} := \mathbb{R} \times \mathbb{R}$ führen wir die folgenden inneren Verknüpfungen ein:

$$\begin{aligned}
 (a, b) + (c, d) &:= (a + c, b + d), \\
 (a, b) \cdot (c, d) &:= (ac - bd, ad + bc).
 \end{aligned}$$

Statt (a, b) schreibt⁴ man auch $a + ib$, insbesondere setzt man $i := (0, 1)$.

SATZ 2.1.3. Die komplexen Zahlen $(\mathbb{C}, +, \cdot, (0, 0), (1, 0))$ bilden einen Körper. Die inversen Elemente bezüglich Addition und Multiplikation sind gegeben durch

$$\begin{aligned}
 -(a, b) &= (-a, -b), \\
 (a, b)^{-1} &= \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right).
 \end{aligned}$$

Komplexe Zahlen werden nach der folgenden Regel dividiert:

$$\frac{(a, b)}{(c, d)} = \left(\frac{ac + bd}{c^2 + d^2}, \frac{bc - ad}{c^2 + d^2} \right).$$

Die komplexe Zahl $i := (0, 1)$ erfüllt $i^2 = (-1, 0)$.

BEWEIS. Der Beweis ist im Detail mühsam, aber nicht schwer, und erfolgt durch einfaches Nachrechnen aller erforderlichen Rechenregeln.

⁴Wir werden später auch diese Schreibweise verwenden. Aus didaktischen Gründen machen wir aber die einführenden Schritte mit der Paar-Schreibweise.

Wir beweisen zum Beispiel die Assoziativität der Multiplikation.

$$\begin{aligned}
 [(a, b) \cdot (c, d)] \cdot (e, f) &= (ac - bd, ad + bc) \cdot (e, f) \\
 &= ((ac - bd)e - (ad + bc)f, (ac - bd)f + (ad + bc)e) \\
 &= (ace - bde - adf - bcf, acf - bdf + ade + bce), \\
 (a, b) \cdot [(c, d) \cdot (e, f)] &= (a, b) \cdot (ce - df, cf + de) \\
 &= (a(ce - df) - b(cf + de), a(cf + de) + b(ce - df)) \\
 &= (ace - adf - bcf - bde, acf + ade + bce - bdf).
 \end{aligned}$$

In beiden Multiplikationen kommt dasselbe heraus, daher gilt das Assoziativgesetz in (\mathbb{C}, \cdot) . Ebenso sind die Kommutativität von Addition und Multiplikation, die Assoziativität der Addition und das Distributivgesetz zu beweisen.

Wir überzeugen uns, dass wir geeignete neutrale Elemente angegeben haben:

$$\begin{aligned}
 (0, 0) + (a, b) &= (0 + a, 0 + b) = (a, b), \\
 (1, 0) \cdot (a, b) &= (1 \cdot a - 0 \cdot b, 0 \cdot a + 1 \cdot b) = (a, b).
 \end{aligned}$$

Als nächstes überprüfen wir die inversen Elemente:

$$\begin{aligned}
 (a, b) + (-a, -b) &= (a + (-a), b + (-b)) = (0, 0), \\
 (a, b) \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right) &= \left(a \frac{a}{a^2 + b^2} - b \frac{-b}{a^2 + b^2}, b \frac{a}{a^2 + b^2} + a \frac{-b}{a^2 + b^2} \right) \\
 &= \left(\frac{a^2 + b^2}{a^2 + b^2}, \frac{ab - ab}{a^2 + b^2} \right) = (1, 0).
 \end{aligned}$$

Für die Division folgt nun

$$\frac{(a, b)}{(c, d)} = (a, b)(c, d)^{-1} = (a, b) \left(\frac{c}{c^2 + d^2}, \frac{-d}{c^2 + d^2} \right) = \left(\frac{ac + bd}{c^2 + d^2}, \frac{bc - ad}{c^2 + d^2} \right).$$

Letztlich ist

$$(0, 1)^2 = (0, 1)(0, 1) = (0 \cdot 0 - 1 \cdot 1, 0 \cdot 1 + 1 \cdot 0) = (-1, 0).$$

□

Wir wollen nun die reellen Zahlen als eine Teilmenge von \mathbb{C} auffassen. Dazu konstruieren wir uns eine sogenannte Einbettung:

SATZ 2.1.4.

(1) Die Abbildung

$$j : \begin{cases} \mathbb{R} & \rightarrow \mathbb{C} \\ a & \mapsto (a, 0) \end{cases}$$

ist injektiv.

(2) Die Abbildung j ist verträglich mit den Grundrechnungsarten im folgenden Sinn: Sind $x, y \in \mathbb{R}$, so gilt

$$\begin{aligned}
 j(x + y) &= j(x) + j(y), \\
 j(xy) &= j(x) \cdot j(y).
 \end{aligned}$$

(3) Die Menge $j(\mathbb{R}) = \{(a, 0) \mid a \in \mathbb{R}\} \subset \mathbb{C}$ ist ein bijektives Bild von \mathbb{R} .

BEWEIS.

(1)

$$j(x) = j(y) \Rightarrow (x, 0) = (y, 0) \Rightarrow x = y.$$

(2)

$$\begin{aligned} j(x) + j(y) &= (x, 0) + (y, 0) = (x + y, 0 + 0) = (x + y, 0) = j(x + y), \\ j(x) \cdot j(y) &= (x, 0) \cdot (y, 0) = (x \cdot y - 0 \cdot 0, x \cdot 0 + 0 \cdot y) = (xy, 0) = j(xy). \end{aligned}$$

(3) Offensichtlich ist $j : \mathbb{R} \rightarrow j(\mathbb{R})$ surjektiv. Weil j aber auch injektiv ist, ist j eine bijektive Abbildung von \mathbb{R} nach $j(\mathbb{R})$.

□

SCHREIBWEISE 2.1.5. Wir identifizieren die Teilmenge

$$j(\mathbb{R}) = \{(a, 0) \mid a \in \mathbb{R}\} \subset \mathbb{C}$$

mit der Menge der reellen Zahlen, indem wir jede komplexe Zahl $(a, 0)$ mit der reellen Zahl a gleichsetzen. Ist \circ eine der Verknüpfungen $+$ oder \cdot , so gilt

$$(a, 0) \circ (b, 0) = (a \circ b, 0).$$

Dadurch trägt $j(\mathbb{R})$ genau dieselbe Körperstruktur wie \mathbb{R} .

“Identifizieren” ist einleuchtend für die Vorstellung und bestens geeignet für den mathematischen Alltag, aber streng genommen unpräzise. Genauer müsste man sagen: Die komplexen Zahlen enthalten als Teilmenge einen Körper, der bijektiv auf \mathbb{R} abgebildet werden kann, wobei die Abbildung mit den Körperverknüpfungen verträglich ist.⁵

Künftig schreiben wir unbesorgt $a + ib$ statt (a, b) und können auf die formale Definition als Paare und den Einbettungshomomorphismus j vergessen: Sie wurden nur gebraucht, um \mathbb{C} auf eine solide formale Basis zu stellen. Jetzt wissen wir, dass es wirklich einen Körper gibt, in dem (-1) eine Wurzel hat. Matrizenrechnung über \mathbb{C} ist zwar etwas mühsam, bringt aber keine neuen Probleme, wenn man nur die Grundrechnungsarten in \mathbb{C} fließend beherrscht.

BEISPIEL 2.1.6.

$$\begin{aligned} (2 + 6i) - (4 - 3i) &= (2 - 4) + i(6 + 3) = -2 + 9i \\ (2 + 6i)(4 - 3i) &= (2 \cdot 4 + 6 \cdot 3) + i(-2 \cdot 3 + 6 \cdot 4) = 26 + 18i \\ \frac{2 + 6i}{4 - 3i} &= \frac{2 \cdot 4 - 6 \cdot 3}{4^2 + 3^2} + i \frac{2 \cdot 3 + 6 \cdot 4}{4^2 + 3^2} = -\frac{10}{25} + i \frac{30}{25} = 0,4 + 1,2i \end{aligned}$$

BEISPIEL 2.1.7. Überprüfen Sie, ob die folgende Matrix $A \in \mathbb{C}^{3 \times 3}$ regulär ist. Wenn A regulär ist, bestimmen Sie die Inverse von A .

$$A = \begin{pmatrix} i & 1 - i & 2 + 3i \\ -2i & 1 + 6i & -2 - 6i \\ i & 1 - i & 3 + 3i \end{pmatrix}.$$

Lösung:

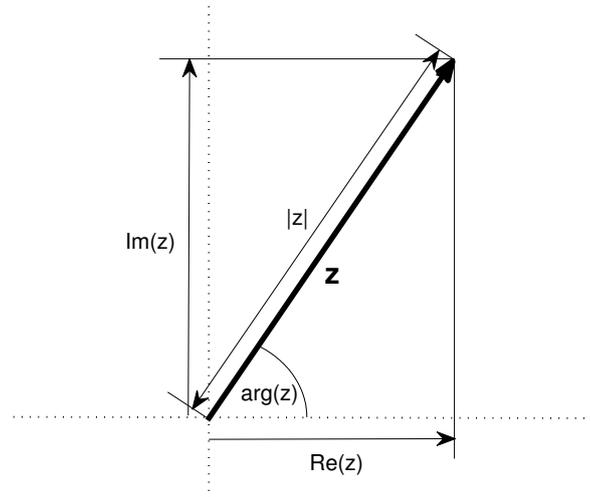
$$\begin{aligned} &\begin{pmatrix} i* & 1 - i & 2 + 3i & 1 & 0 & 0 \\ -2i & 1 + 6i & -2 - 6i & 0 & 1 & 0 \\ i & 1 - i & 3 + 3i & 0 & 0 & 1 \end{pmatrix} \\ \rightarrow &\begin{pmatrix} 1 & -1 - i & 3 - 2i & -i & 0 & 0 \\ 0 & 3 + 4i & 2 & 2 & 1 & 0 \\ 0 & 0 & 1* & -1 & 0 & 1 \end{pmatrix} \\ \rightarrow &\begin{pmatrix} 1 & -1 - i & 0 & 3 - 3i & 0 & -3 + 2i \\ 0 & 3 + 4i* & 0 & 4 & 1 & -2 \\ 0 & 0 & 1 & -1 & 0 & 1 \end{pmatrix} \\ \rightarrow &\begin{pmatrix} 1 & 0 & 0 & \frac{103 - 79i}{25} & \frac{7 - i}{25} & \frac{-89 + 52i}{25} \\ 0 & 1 & 0 & \frac{12 - 16i}{25} & \frac{3 - 4i}{25} & \frac{-6 + 8i}{25} \\ 0 & 0 & 1 & -1 & 0 & 1 \end{pmatrix} \end{aligned}$$

⁵Mit ein bisschen mehr Grundwissen aus Algebra könnte man sagen: \mathbb{C} enthält als Teilmenge einen Körper, der isomorph zu \mathbb{R} ist.

Also ist A regulär und

$$A^{-1} = \frac{1}{25} \begin{pmatrix} 103 - 79i & 7 - i & -89 + 52i \\ 12 - 16i & 3 - 4i & -6 + 8i \\ -25 & 0 & 25 \end{pmatrix}.$$

DEFINITION 2.1.8 (Gaussche Zahlenebene). In der Gaußschen Zahlenebene wird jede komplexe Zahl $z = a + bi$ als Vektor $(a, b)^T \in \mathbb{R}^2$ dargestellt:



Wir definieren die folgenden Größen für $z = a + bi$:

- den Realteil von z durch $\operatorname{Re}(z) := a$,
- den Imaginärteil von z durch $\operatorname{Im}(z) := b$,
- die konjugiert komplexe Zahl zu z durch $\bar{z} := a - bi$,
- den Betrag $|z| := r$ und das Argument $\arg(z) := \phi$ von z als jene nichtnegative reelle Zahl r und jenen Winkel ϕ für die gilt:

$$z = r(\cos(\phi) + i \sin(\phi)).$$

(In diesem Zusammenhang geben wir Winkel immer im Bogenmaß an: 2π entspricht einem Winkel von 360° . Beachten Sie dabei, dass Winkel stets nur auf Vielfache von 2π eindeutig bestimmt sind, der Winkel von $360^\circ = 2\pi$ ist derselbe Winkel wie $0^\circ = 0$.)

BEHAUPTUNG 2.1.9.

- (a) Der Betrag einer komplexen Zahl $a + bi$ errechnet sich durch die nichtnegative Quadratwurzel

$$|a + bi| = \sqrt{a^2 + b^2}.$$

- (b) Es gilt $|z|^2 = z\bar{z}$. Insbesondere ist $z\bar{z}$ immer reell und nichtnegativ, und Null nur dann, wenn $z = 0$ ist.
- (c) Die Addition von komplexen Zahlen stellt sich in der Zahlenebene als Vektoraddition dar.
- (d) Bei der Multiplikation multiplizieren sich die Beträge, und addieren sich die Argumente.
- (e) Bei der Division dividieren sich die Beträge und subtrahieren sich die Argumente, insbesondere ist $\arg\left(\frac{1}{z}\right) = -\arg(z)$.

BEWEIS.

- (a) Wir verwenden $\cos^2(\phi) + \sin^2(\phi) = 1$. Ist also $a = r \cos(\phi)$, $b = r \sin(\phi)$, dann gilt

$$a^2 + b^2 = r^2 \cos^2(\phi) + r^2 \sin^2(\phi) = r^2.$$

- (b) Offensichtlich ist

$$(a + bi) \cdot \overline{(a + bi)} = a^2 + b^2 = |a + bi|^2.$$

- (c) Die geometrische Deutung der Addition ist offensichtlich.
 (d) Zur geometrischen Deutung der Multiplikation verwenden wir die Additionstheoreme der Winkelfunktionen:

$$\begin{aligned} \cos(\phi + \psi) &= \cos(\phi) \cos(\psi) - \sin(\phi) \sin(\psi), \\ \sin(\phi + \psi) &= \cos(\phi) \sin(\psi) + \sin(\phi) \cos(\psi). \end{aligned}$$

Wir multiplizieren die komplexen Zahlen

$$u = r \cos(\phi) + r \sin(\phi)i, \quad v = s \cos(\psi) + s \sin(\psi)i$$

und erhalten

$$\begin{aligned} & [r \cos(\phi) + r \sin(\phi)i][s \cos(\psi) + s \sin(\psi)i] \\ &= [r \cos(\phi)s \cos(\psi) - r \sin(\phi)s \sin(\psi)] + [r \cos(\phi)s \sin(\psi) + r \sin(\phi)s \cos(\psi)]i \\ &= rs[\cos(\phi) \cos(\psi) - \sin(\phi) \sin(\psi)] + rs[\cos(\phi) \sin(\psi) + \sin(\phi) \cos(\psi)]i \\ &= rs \cos(\phi + \psi) + rs \sin(\phi + \psi)i. \end{aligned}$$

Aus dieser Darstellung von uv folgt $|uv| = rs$ und $\arg(uv) = \phi + \psi$.

- (e) Die Deutung der Division folgt aus Punkt (d).

□

BEMERKUNG 2.1.10. Der Betrag erfüllt die Dreiecksungleichung: gilt

$$(\forall u, v \in \mathbb{C}) \quad |u + v| \leq |u| + |v|.$$

Die Dreiecksungleichung ist die Grundlage dafür, dass man $|u - v|$ als Maß für den Abstand zweier komplexer Zahlen u, v verwenden kann, denn es gilt

$$|u - v| \leq |u - w| + |w - v|,$$

der Abstand der Direktverbindung von u nach v ist nie größer, als wenn man die Abstände des Umwegs über eine dritte Zahl w nimmt.

Wenn man erst einen Abstand definiert hat, kann man auch anfangen, Grenzwerte einzuführen. Der Absolutbetrag ist also die Grundlage für die Grenzwertrechnung, Integral- und Differentialrechnung in \mathbb{C} .

BEHAUPTUNG 2.1.11. Sei $u = r(\cos(\phi) + i \sin(\phi))$ (mit $r > 0$ und $\phi \in \mathbb{R}$) eine komplexe Zahl ungleich Null. Sei $p \in \mathbb{N}$ (hier wird \mathbb{N} ohne Null verstanden, also $p \neq 0$). Die Gleichung

$$z^p = u$$

besitzt genau p Lösungen, und zwar

$$z = r^{1/p} \left[\cos\left(\frac{\phi + 2k\pi}{p}\right) + i \sin\left(\frac{\phi + 2k\pi}{p}\right) \right],$$

mit $k \in \{0, \dots, p-1\}$.

BEWEIS. Ansatz: Setze $z = s(\cos(\psi) + i\sin(\psi))$. Nach den Regeln der Multiplikation (p -mal auf z angewendet) muss sich also der Betrag potenzieren, und das Argument multiplizieren:

$$r(\cos(\phi) + i\sin(\phi)) = s^p(\cos(p\psi) + i\sin(p\psi)).$$

Also $r = s^p$ oder $s = r^{1/p}$. Und $\psi = \phi/p$? Aber Vorsicht, das Argument ist ja nur eindeutig bis auf modulo 2π ! Also muss nur gelten

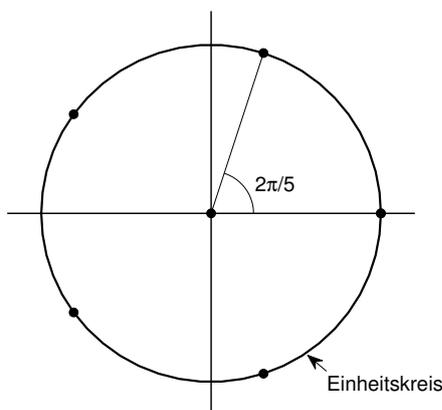
$$p\psi = \phi + 2k\pi \text{ für irgendein } k \in \mathbb{Z}.$$

Das ergibt p verschiedene Lösungen

$$\psi = \frac{\phi}{p} + \frac{k\pi}{p} \text{ mit } k \in \{0, \dots, p-1\}.$$

(Ab $k = p$ wiederholen sich dieselben Lösungen). □

Die folgende Graphik zeigt die fünften Einheitswurzeln, also die 5 Lösungen der Gleichung $z^5 = 1$, in der Gaußschen Zahlenebene. Alle liegen am Einheitskreis $\{z \in \mathbb{C} \mid |z| = 1\}$. Ihre Argumente sind die ganzzahligen Vielfachen von $\frac{2\pi}{5}$.



Die große Bedeutung der komplexen Zahlen steht auf drei Säulen. Was zuerst wie ein Spleen weltfremder Mathematiker aussieht, die unbedingt die Wurzel aus -1 ziehen wollen, hat sich als eines der fruchtbarsten und erstaunlichsten Konzepte der Mathematik bewährt. Sie sollen also eine Idee davon bekommen, worin die Kraft dieser Theorie besteht, auch wenn wir hier keine Beweise liefern können. Das kommt zum Teil in den Grundvorlesungen der Analysis, zum Teil in einer Vorlesung über Funktionentheorie.

Ein Polynom über einem Körper \mathbb{K} ist ein Ausdruck

$$\gamma_0 + \gamma_1 x + \gamma_2 x^2 + \dots + \gamma_n x^n$$

mit $\gamma_0, \dots, \gamma_n \in \mathbb{K}$. Eine Nullstelle des Polynomes ist ein Element $\xi \in \mathbb{K}$ sodass

$$\gamma_0 + \gamma_1 \xi + \gamma_2 \xi^2 + \dots + \gamma_n \xi^n$$

BEHAUPTUNG 2.1.12. *Jedes nicht konstante Polynom über \mathbb{C} hat mindestens eine komplexe Nullstelle.*

Das Polynom $x^2 + 1$ hätte zum Beispiel keine reelle Nullstelle. Braucht man also zur Behandlung eines Problems die Nullstellen eines Polynoms, so muss man reelle Polynome Fallunterscheidungen treffen zwischen Situationen, in denen das Polynom Nullstellen hat, und dem Fall, dass es keine Nullstelle gibt. Im Komplexen gibt es immer Nullstellen, und daher keine umständlichen Fallunterscheidungen. Dieses Thema wird im zweiten Semester eine Schlüsselrolle spielen, wenn wir die Theorie der Eigenwerte entwickeln.

BEHAUPTUNG 2.1.13. *Die Eulersche Zahl*

$$e := \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n \approx 2,7183$$

läßt sich auch zu komplexen Potenzen erheben. Sind r und ϕ reelle Zahlen, so gilt die Eulersche Formel

$$(2.1.1) \quad e^{r+i\phi} = e^r [\cos(\phi) + i \sin(\phi)].$$

In vielen Anwendungsgebieten der Mathematik, zum Beispiel in der Mechanik, treten lineare Differentialgleichungen auf. Sie handeln von vektorwertigen Funktionen $u : [0, T] \rightarrow \mathbb{R}^n$ mit

$$\frac{d}{dt} u(t) = Au(t),$$

dabei ist A eine quadratische Matrix. Für die Lösungen und ihre Komponenten sind drei Arten des Verlaufs charakteristisch:

- u kann exponentiell wachsen oder abklingen. Im einfachsten Fall lässt sich u durch die Formel

$$u(t) = Ce^{rt}$$

beschreiben. Dabei sind $C, r \in \mathbb{R}$. Das Vorzeichen von r entscheidet, ob die Funktion exponentiell steigt oder fällt.

- u kann eine periodische Schwingung ausführen:

$$u(t) = C \cos(\omega t) \quad \text{mit } C, \omega \in \mathbb{R}.$$

Nach (2.1.1) lässt sich ein solches u als der Realteil einer Exponentialfunktion schreiben:

$$u(t) = C \operatorname{Re}(e^{i\omega t}).$$

- Es gibt auch Schwingungen, die sich exponentiell aufschaukeln oder exponentiell abklingen:

$$u(t) = Ce^{rt} \cos(\omega t) = C \operatorname{Re}(e^{(t+i\omega)t}).$$

Durch die komplexen Zahlen lassen alle diese Funktion mit derselben Theorie anfassen, nämlich mit der komplexen Exponentialfunktion.

Die Fourieranalyse beschäftigt sich damit, Funktionen als Überlagerung von periodischen Schwingungen darzustellen, etwa, wie der komplexe Klang einer menschlichen Stimme oder eines Musikinstrumentes sich aus Grund- und Obertönen zusammensetzt. Auch diese Theorie wird durch den Gebrauch der komplexen Exponentialfunktion wesentlich übersichtlicher.

Vielleicht wissen Sie schon aus der Analysis oder noch aus der Schule, was es bedeutet, eine reelle Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ zu differenzieren, und auch, dass nicht jede Funktion differenziert werden kann. Die differenzierbaren Funktionen $f : \mathbb{C} \rightarrow \mathbb{C}$ haben eine besonders reiche Theorie mit überraschenden Ergebnissen. Einige davon finden Sie in der folgenden Behauptung. Was man alles mit dieser faszinierenden Theorie anfangen kann, kann ich Ihnen leider jetzt noch nicht erklären.

BEHAUPTUNG 2.1.14. *Sei $f : \mathbb{C} \rightarrow \mathbb{C}$ differenzierbar und nicht konstant Null.*

- f lässt sich immer durch eine Potenzreihe ausdrücken:

$$f(z) = \sum_{k=0}^{\infty} \gamma_k z^k.$$

- Die Nullstellen von f können sich nicht häufen, das heißt, um jede Nullstelle liegt ein Kreis, in der es keine weitere Nullstelle gibt.
- Der maximale Betrag, den die Funktion f auf einem Kreis (oder auch allgemeineren Gebieten) erreicht, wird immer am Rand des Gebietes angenommen.
- Die Werte der Funktion f am Rand eines Kreises bestimmen bereits eindeutig die Werte, die f im Inneren des Kreises annimmt.

2.2. Restklassen.

In diesem Kapitel werden wir mit der Kongruenz modulo m und den Restklassen ein Beispiel einer Äquivalenzrelation und der dazu gehörenden Klasseneinteilung kennenlernen. Eine allgemeine Abhandlung dieser wichtigen Begriffe müssen wir aus Zeitgründen auf das zweite Semester in das Kapitel über Faktorräume verschieben.

Die Menge der ganzen Zahlen ist bekanntlich kein Körper, weil es innerhalb der ganzen Zahlen keine Division gibt. In der Terminologie der Algebra ist \mathbb{Z} ein kommutativer Ring mit Eins:

DEFINITION 2.2.1. Sei R eine Menge mit mindestens einem Element $0 \in R$, seien $+$ und \cdot zwei innere Verknüpfungen in R . Das Quadrupel $(R, +, \cdot, 0)$ heißt ein Ring, wenn folgende Axiome gelten:

- Für $+$: $(R, +)$ ist eine abelsche Gruppe, das heißt
 - (plus 1) Assoziativgesetz: $(\forall x, y, z \in R) (x + y) + z = x + (y + z)$.
 - (plus 2) 0 ist neutrales Element für $+$, das heißt: $(\forall x \in R) x + 0 = x$.
 - (plus 3) Zu jedem $x \in R$ gibt es ein inverses Element $y \in R$ sodass $x + y = 0$.
 - (plus 4) Kommutativgesetz: $(\forall x, y \in R) x + y = y + x$.
- Für \cdot :
 - (mal 1) Assoziativgesetz: $(\forall x, y, z \in R) (x \cdot y) \cdot z = x \cdot (y \cdot z)$.
- Für das Zusammenspiel von $+$ und \cdot :
 - (dist 1) Distributivgesetz links: $(\forall x, y, z \in R) x \cdot (y + z) = x \cdot y + x \cdot z$.
 - (dist 2) Distributivgesetz rechts: $(\forall x, y, z \in R) (y + z) \cdot x = y \cdot x + z \cdot x$.

Wenn die Multiplikation zusätzlich kommutativ ist, d.h., $(\forall x, y \in R) x \cdot y = y \cdot x$, dann heißt $(R, \cdot, 0)$ ein kommutativer Ring.

Wenn die Multiplikation ein neutrales Element 1 besitzt, also $(\forall x \in R) 1 \cdot x = x \cdot 1 = x$, dann heißt $(R, +, \cdot, 0, 1)$ ein Ring mit Eins.

Wir werden nun aus \mathbb{Z} weitere kommutative Ringe mit Eins konstruieren, wobei wir uns vorerst um die Möglichkeit einer Division keine Gedanken machen.

DEFINITION 2.2.2 (Kongruenz modulo m). Sei $m \in \mathbb{N}$, seien $x, y \in \mathbb{Z}$:

- (1) x und y heißen zueinander kongruent modulo m , wenn die Differenz $x - y$ ein ganzzahliges Vielfaches von m ist:

$$x \equiv y \pmod{m} \text{ (oder kurz } x \equiv_m y) : \iff (\exists c \in \mathbb{Z}) (y - x) = cm.$$
- (2) Die Restklasse von x modulo m ist die Menge aller ganzen Zahlen, die zu x kongruent sind:

$$[x]_m := \{y \in \mathbb{Z} \mid x \equiv_m y\}.$$

- (3) Die Menge der Restklassen modulo m bezeichnen⁶ wir mit $\mathbb{Z}/m\mathbb{Z}$.

Wenn man die Restklassen modulo $-m$ ebenso definiert, erhält man wiederum dieselben Restklassen modulo m . Daher können wir uns wie oben auf $m \in \mathbb{N}$ beschränken.

BEISPIEL 2.2.3. Die vier Restklassen modulo 4 in \mathbb{Z} sind:

$$\begin{aligned} [0]_4 &= \{0, \pm 4, \pm 8, \dots\}, \\ [1]_4 &= \{1, -4, 5, -7, 9, \dots\}, \\ [2]_4 &= \{2, -2, 6, -6, 10, \dots\}, \\ [3]_4 &= \{3, -1, 7, -5, 11, \dots\}. \end{aligned}$$

Diese Restklassen wiederholen sich, z.B. ist

$$[4]_4 = [0]_4, [5]_4 = [1]_4, \dots, [-1]_4 = [3]_4, [-2]_4 = [2]_4, \dots.$$

⁶Das Zeichen / ist also ein Symbol für die Restklassenbildung, keine Division.

BEMERKUNG 2.2.4. Eine äquivalente Möglichkeit, die Restklasse $[x]_m$ anzuschreiben, ist

$$[x]_m = x + m\mathbb{Z} = \{x + mz \mid z \in \mathbb{Z}\}.$$

Insbesondere ist

$$[0]_m = m\mathbb{Z} = \{mz \mid z \in \mathbb{Z}\}.$$

BEHAUPTUNG 2.2.5. Sei $m \in \mathbb{N}$. Die Kongruenzrelation modulo m ist eine Äquivalenzrelation auf \mathbb{Z} , das heißt, es gelten folgende Eigenschaften:

- (1) $(\forall x \in \mathbb{Z}) x \equiv_m x$ (Reflexivität),
- (2) $(\forall x, y \in \mathbb{Z}) x \equiv_m y \iff y \equiv_m x$ (Symmetrie),
- (3) $(\forall x, y, z \in \mathbb{Z}) [x \equiv_m y \wedge y \equiv_m z] \Rightarrow [x \equiv_m z]$ (Transitivität).

BEWEIS.

- (1) Weil $x - x = 0 = 0m$, gilt $x \equiv_m x$.
- (2) Sei $x \equiv_m y$. Dann gibt es ein $c \in \mathbb{Z}$ sodass $y - x = cm$. Folglich ist $x - y = (-c)m$, und daraus folgt $y \equiv_m x$. Vertauscht man in diesem Beweis die Rollen von x und y , erhält man auch $y \equiv_m x \Rightarrow x \equiv_m y$.
- (3) Es sei $x \equiv_m y$ und $y \equiv_m z$, also $y - x = cm$ und $z - y = dm$ für geeignete $c, d \in \mathbb{Z}$. Es folgt $z - x = (z - y) + (y - x) = dm + cm = (d + c)m$, und daher $x \equiv_m z$.

□

BEHAUPTUNG 2.2.6. Sei $m \in \mathbb{N}$. Seien $x, y \in \mathbb{Z}$. Dann gilt

- (1) Es gilt immer $x \in [x]_m$.
- (2) Es sind äquivalent:
 - (i) $x \equiv_m y$,
 - (ii) $y \in [x]_m$,
 - (iii) $[x]_m = [y]_m$.
- (3) Es gilt entweder $[x]_m \cap [y]_m = \emptyset$ oder $[x]_m = [y]_m$.

BEWEIS.

- (1) $x = x + 0 = x + 0m$.
- (2) (i) \Rightarrow (iii): Sei $x \equiv_m y$. Für jedes $z \in [y]_m$ gilt nach Definition 2.2.2 (2) $y \equiv_m z$. Wegen der Transitivität gilt $x \equiv_m z$, also $z \in [x]_m$. Also gilt $[y]_m \subseteq [x]_m$. Wegen der Symmetrie gilt aber auch $y \equiv_m x$ und, mit vertauschten Rollen von x und y , folgt wie oben $[x]_m \subseteq [y]_m$.
(iii) \Rightarrow (ii): Sei $[y]_m = [x]_m$. Weil immer $y \in [y]_m$, folgt nun $y \in [x]_m$.
(ii) \Rightarrow (i): Definition 2.2.2 (2).
- (3) Sei $[x]_m \cap [y]_m \neq \emptyset$, also gibt es ein $z \in [x]_m \cap [y]_m$. Aus Punkt (2) folgt $[z]_m = [x]_m$ und $[z]_m = [y]_m$, also ist $[x]_m = [y]_m$.

□

BEHAUPTUNG 2.2.7. Sei $m \in \mathbb{N}$. Die Restklassen modulo m bilden eine Partition oder Klasseneinteilung auf \mathbb{Z} , das heißt:

- Keine Restklasse ist leer.
- Jede ganze Zahl x ist in genau einer Restklasse enthalten.

BEWEIS. Natürlich enthält jede Restklasse $[x]_m$ das Element x , also ist sie nicht leer, und wir haben auch eine Restklasse gefunden, die x enthält. Ist $x \in [z]_m$ für ein weiteres $z \in \mathbb{Z}$, so folgt $[x]_m = [z]_m$. Daher ist x nur in einer einzigen Restklasse enthalten. □

Wir werden nun auf der Menge $\mathbb{Z}/m\mathbb{Z}$ der Restklassen modulo m eine Addition und Multiplikation einführen. Dafür brauchen wir das folgende Lemma:

LEMMA 2.2.8. *Sei $m \in \mathbb{N}$ und $x_1, x_2, y_1, y_2 \in \mathbb{Z}$ mit $x_1 \equiv_m x_2$ und $y_1 \equiv_m y_2$. Dann gilt*

$$\begin{aligned}(x_1 + y_1) &\equiv_m (x_2 + y_2), \\ (x_1 y_1) &\equiv_m (x_2 y_2).\end{aligned}$$

BEWEIS. Es sei $x_2 = x_1 + cm$ und $y_2 = y_1 + dm$. Es ist

$$\begin{aligned}x_2 + y_2 &= x_1 + cm + y_1 + dm = (x_1 + y_1) + (c + d)m, \\ x_2 y_2 &= (x_1 + cm)(y_1 + dm) = x_1 y_1 + m(x_1 d + c y_1 + cdm).\end{aligned}$$

□

DEFINITION 2.2.9 (Rechnen mit Restklassen). Sei $m \in \mathbb{N}$. Wir definieren die Addition und Multiplikation⁷ von Restklassen aus $\mathbb{Z}/m\mathbb{Z}$ folgendermaßen:

$$\begin{aligned}[x]_m + [y]_m &:= [x + y]_m, \\ [x]_m [y]_m &:= [xy]_m.\end{aligned}$$

(Die Restklassen von x und y verknüpft man, indem man erst x und y verknüpft, und vom Ergebnis die Restklasse bildet.)

Erklärung: Diese Definition darf man nicht gedankenlos übernehmen, ohne sich vorher zu fragen, ob die Verknüpfungen dadurch wohldefiniert sind. Es kann ja dieselbe Restklasse auf verschiedene Weise angeschrieben sein (z.B. $[7]_4 = [3]_4$). Das Ergebnis darf aber nur von der Restklasse abhängen, nicht von der Art, wie sie angeschrieben wurde. Wenn also $[x_1]_m = [x_2]_m$ und $[y_1]_m = [y_2]_m$ jeweils dieselben Restklassen sind, dann muss auch in beiden Fällen dasselbe Ergebnis herauskommen: $[x_1 + y_1]_m = [x_2 + y_2]_m$ und $[x_1 y_1]_m = [x_2 y_2]_m$. Das haben wir durch Lemma 2.2.8 abgesichert.

SATZ 2.2.10. *Sei $m \in \mathbb{N}$. Dann ist $(\mathbb{Z}/m\mathbb{Z}, +, \cdot, [0]_m, [1]_m)$ ein kommutativer Ring mit Eins.*

BEWEIS. Alle Rechengesetze übertragen sich direkt von \mathbb{Z} auf $\mathbb{Z}/m\mathbb{Z}$. Wir zeigen zum Beispiel die Distributivität:

$$\begin{aligned}& [x]_m([y]_m + [z]_m) \\ &= [x]_m[y + z]_m = [x(y + z)]_m && \text{Definition 2.2.9} \\ &= [xy + xz]_m && \text{Distributivität auf } \mathbb{Z} \\ &= [xy]_m + [xz]_m = [x]_m[y]_m + [x]_m[z]_m && \text{Definition 2.2.9}\end{aligned}$$

Die Assoziativität und Kommutativität von $+$ und \cdot beweist man genauso.

Die Restklasse von Null ist neutral bezüglich der Addition, und die Restklasse von 1 ist neutral bezüglich der Multiplikation:

$$[x]_m + [0]_m = [x + 0]_m = [x]_m \text{ und } [x]_m [1]_m = [x1]_m = [x]_m.$$

Die Restklasse $[-x]_m$ ist invers zu $[x]_m$ bezüglich der Addition:

$$[x]_m + [-x]_m = [x + (-x)]_m = [0]_m.$$

□

⁷Das sind natürlich andere Verknüpfungen als $+$ und \cdot auf \mathbb{Z} . Aber weil aus dem Zusammenhang immer klar wird, ob Verknüpfungen von Zahlen oder von Restklassen gemeint sind, verwenden wir ohne Bedenken dieselben Symbole.

BEISPIEL 2.2.11. Der Restklassenring $\mathbb{Z}/6\mathbb{Z}$ hat folgende Verknüpfungstabellen:

+	[0]	[1]	[2]	[3]	[4]	[5]
[0]	[0]	[1]	[2]	[3]	[4]	[5]
[1]	[1]	[2]	[3]	[4]	[5]	[0]
[2]	[2]	[3]	[4]	[5]	[0]	[1]
[3]	[3]	[4]	[5]	[0]	[1]	[2]
[4]	[4]	[5]	[0]	[1]	[2]	[3]
[5]	[5]	[0]	[1]	[2]	[3]	[4]

·	[0]	[1]	[2]	[3]	[4]	[5]
[0]	[0]	[0]	[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]	[3]	[4]	[5]
[2]	[0]	[2]	[4]	[0]	[2]	[4]
[3]	[0]	[3]	[0]	[3]	[0]	[3]
[4]	[0]	[4]	[2]	[0]	[4]	[2]
[5]	[0]	[5]	[4]	[3]	[2]	[1]

Beachten Sie, dass [1] und [5] die einzigen Klassen sind, die inverse Elemente bezüglich der Multiplikation besitzen. Weil \mathbb{Z} kein Körper ist, gibt es keine Division, deren Rechenregeln sich direkt auf den Restklassenring modulo m übertragen könnte. Die Klassen [2], [3] und [4] sind Nullteiler: Obwohl sie selbst nicht die Nullklasse sind, können sie mit anderen Nicht-Nullklassen multipliziert werden, sodass [0] herauskommt. Zum Beispiel ist $[2][3]=[0]$.

Trotzdem kann $(\mathbb{Z}/m\mathbb{Z}, +, \cdot, [0]_m, [1]_m)$ ein Körper sein:

SATZ 2.2.12. Sei $m \in \mathbb{N}$. Der Restklassenring $(\mathbb{Z}/m\mathbb{Z}, +, \cdot, [0]_m, [1]_m)$ ist genau dann ein Körper, wenn m eine Primzahl ist.

BEWEIS. Sei m keine Primzahl. Dann gibt es $m = pq$ mit $1 < p < m$ und $1 < q < m$. Dann gilt

$$[p]_m [q]_m = [m]_m = [0]_m.$$

Wenn $(\mathbb{Z}/m\mathbb{Z}, \cdot, [0]_m, [1]_m)$ ein Körper wäre, müsste nach Behauptung 1.1.8 folgen,

$$[p]_m = [0]_m \text{ oder } [q]_m = [0]_m.$$

Aber weder p noch q ist ein ganzzahliges Vielfaches von m , daher ist $[p]_m \neq [0]_m$ und $[q]_m \neq [0]_m$. Also kann $(\mathbb{Z}/m\mathbb{Z}, +, \cdot, [0]_m, [1]_m)$ kein Körper sein.

Sei m eine Primzahl. Wir zeigen zunächst, dass für jede Restklasse $[x]_m \neq [0]_m$ die Abbildung

$$g_x : \begin{cases} \mathbb{Z}/m\mathbb{Z} & \rightarrow \mathbb{Z}/m\mathbb{Z} \\ [y]_m & \mapsto [x]_m [y]_m \end{cases}$$

injektiv ist. Seien also $[x]_m \neq [0]_m$ und $[y]_m, [z]_m \in \mathbb{Z}/m\mathbb{Z}$ so, dass

$$[x]_m [y]_m = [x]_m [z]_m.$$

Es ist dann

$$[x(z-y)]_m = [x]_m [z-y]_m = [x]_m ([z]_m - [y]_m) = [x]_m [z]_m - [x]_m [y]_m = [0]_m,$$

also ist $x(z-y)$ ein ganzzahliges Vielfaches von m . Die Primzahl m teilt aber nicht x , denn $[x]_m \neq [0]_m$, also muss m ein Teiler von $z-y$ sein. Also ist $[z-y]_m = [0]_m$, und $[z]_m = [y]_m$. Damit ist die Injektivität von g_x bewiesen.

Weil g_x injektiv ist und $\mathbb{Z}/m\mathbb{Z}$ aus m verschiedenen Restklassen besteht, enthält das Bild $g_x(\mathbb{Z}/m\mathbb{Z})$ ebenfalls m verschiedene Restklassen. Daher muss auch die Restklasse $[1]_m$ im Bild sein. Es gibt also eine Restklasse $[y]_m \in \mathbb{Z}/m\mathbb{Z}$, sodass

$$[x]_m [y]_m = g_x([y]_m) = [1]_m.$$

Daher besitzt die Restklasse $[x]_m$ ein inverses Element bezüglich der Multiplikation, und $(\mathbb{Z}/m\mathbb{Z}, +, \cdot, [0]_m, [1]_m)$ ist ein Körper. \square

Die Algebra hat zur Division in Restklassenkörpern ein Verfahren bereit, nämlich den Euklidischen Algorithmus. Für das Rechnen in kleinen Restklassenkörpern kann man sich aber leicht durch Probieren eine Tabelle der inversen Elemente bereitstellen und diese als Hilfsmittel zur Division heranziehen.

BEISPIEL 2.2.13. Wir bestimmen eine Tabelle der inversen Elemente im Restklassenkörper modulo 5.

Lösung: Das Nullelement $[0]_5$ hat natürlich kein inverses Element. Das Element $[1]_5$ ist das Einselement und muss zu sich invers sein. Ebenso muss $[-1]_5$ zu sich selbst invers sein also $[4]_5^{-1} = [4]_5$. Nun müssen wir nur mehr die Inversen von $[2]_5$ und $[3]_5$ bestimmen, in Frage kommen als Inverse kommen auch nur mehr $[2]_5$ und $[3]_5$. Durch Probieren stellt man fest $[2]_5[3]_5 = 1$, und die Tabelle ist fertig:

x	$[1]_5$	$[2]_5$	$[3]_5$	$[4]_5$
x^{-1}	$[1]_5$	$[3]_5$	$[2]_5$	$[4]_5$

Wie in allen Körpern, funktionieren auch in Restklassenkörpern die Methoden der Matrizenrechnung aus Kapitel 1:

BEISPIEL 2.2.14. Wir lösen in $\mathbb{Z}/5\mathbb{Z}$ das lineare Gleichungssystem

$$\begin{aligned} [2]_5x + [3]_5y + [2]_5z + [4]_5u &= [1]_5 \\ [3]_5x + [1]_5y + [1]_5z + [2]_5u &= [4]_5 \\ [3]_5x + [4]_5y + [1]_5z + [3]_5u &= [3]_5 \\ [4]_5x + [2]_5y + [0]_5z + [1]_5u &= [1]_5 \end{aligned}$$

Lösung: Um die Rechnung übersichtlicher anzuschreiben, ersparen wir uns während der Pivot-schritte das Schreiben der eckigen Klammern für die Restklassen. Wir dürfen aber nicht vergessen, dass wir nicht mit ganzen Zahlen, sondern mit Restklassen modulo 5 arbeiten.

Damit Sie gleich eine Division sehen, beginnen wir nicht mit der günstigsten Pivotzeile. Divisionen werden durch Multiplikationen mit den inversen Elementen durchgeführt. Additionen, Subtraktionen und Multiplikationen führt man durch, indem man wie mit ganzen Zahlen rechnet und am Ende der Rechnung die Restklasse bestimmt.

$$\begin{aligned} &\begin{pmatrix} 2* & 3 & 2 & 4 & 1 \\ 3 & 1 & 1 & 2 & 4 \\ 3 & 4 & 1 & 3 & 3 \\ 4 & 2 & 0 & 1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 4 & 1 & 2 & 3 \\ 0 & 4 & 3 & 1 & 0 \\ 0 & 2 & 3 & 2 & 4 \\ 0 & 1* & 1 & 3 & 4 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 4 & 1 & 2 & 3 \\ 0 & 0 & 4 & 4 & 4 \\ 0 & 0 & 1* & 1 & 1 \\ 0 & 1 & 1 & 3 & 4 \end{pmatrix} \\ &\rightarrow \begin{pmatrix} 1 & 4 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 1* & 0 & 2 & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 3 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 2 & 3 \end{pmatrix} \end{aligned}$$

Die vierte Variable bleibt nichtbasisch. Wir erhalten die Lösungsmenge

$$\left\{ \begin{pmatrix} [0]_5 \\ [3]_5 \\ [1]_5 \\ [0]_5 \end{pmatrix} + t \begin{pmatrix} [2]_5 \\ [3]_5 \\ [4]_5 \\ [1]_5 \end{pmatrix} \mid t \in \mathbb{Z}/5\mathbb{Z} \right\}.$$

Inneres Produkt und Orthogonalität

1. Das euklidische innere Produkt

1.1. Motivation aus der Elementargeometrie.

Die Theorie des inneren Produktes ist, grob gesprochen, die Übertragung des Satzes von Pythagoras in die lineare Algebra:

Seien O, A, B die Eckpunkte eines Dreiecks. (Einen davon haben wir o.B.d.A. in den Koordinatenursprung O verlegt.) Mit \overline{PQ} bezeichnen wir die Länge der Strecke von P nach Q . Der Winkel an der Ecke O ist genau dann ein rechter Winkel, wenn für die Seitenlängen gilt

$$\overline{OA}^2 + \overline{OB}^2 = \overline{AB}^2.$$

Gehen wir davon aus, dass die Einheitsvektoren des \mathbb{R}^n aufeinander im rechten Winkel stehen, so liefert der Satz von Pythagoras für die Länge $\|a\|$ eines Vektors

$$\left\| \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \right\|^2 = \alpha_1^2 + \cdots + \alpha_n^2.$$

Nun seien $a = (\alpha_1, \dots, \alpha_n)^T$ und $b = (\beta_1, \dots, \beta_n)^T$ die Ortsvektoren zweier Punkte A, B in einem n -dimensionalen affinen Raum. Der Verbindungsvektor von A nach B ist dann $\vec{AB} = b - a$. Es gilt

$$\begin{aligned} \overline{AB}^2 &= \sum_{j=1}^n (\beta_j - \alpha_j)^2 = \sum_{j=1}^n (\beta_j^2 + \alpha_j^2 - 2\alpha_j\beta_j) = \sum_{j=1}^n \beta_j^2 + \sum_{j=1}^n \alpha_j^2 - 2 \sum_{j=1}^n \alpha_j\beta_j \\ &= \overline{OB}^2 + \overline{OA}^2 - 2 \sum_{j=1}^n \alpha_j\beta_j. \end{aligned}$$

Nach Pythagoras gilt dann

$$\begin{aligned} &\text{Der Winkel in der Ecke } O \text{ ist ein rechter Winkel} \\ \Leftrightarrow &\overline{OA}^2 + \overline{OB}^2 = \overline{AB}^2 \\ \Leftrightarrow &\overline{OA}^2 + \overline{OB}^2 = \overline{OB}^2 + \overline{OA}^2 - 2 \sum_{j=1}^n \alpha_j\beta_j \\ \Leftrightarrow &0 = \sum_{j=1}^n \alpha_j\beta_j. \end{aligned}$$

Die Summe $\langle a, b \rangle := \sum_{j=1}^n \alpha_j\beta_j$ ist also der Schlüssel zur Behandlung des rechten Winkels in der linearen Algebra. Und auch die Länge von a ist nichts anderes als $\|a\| = \sqrt{\langle a, a \rangle}$.

1.2. Definition von inneren Produkten und Normen.

Damit wir uns Wiederholungen sparen, steht im ganzen Kapitel über innere Produkte \mathbb{K} durchweg für einen der Körper \mathbb{R} oder \mathbb{C} . Für reelle Zahlen gilt immer $\bar{a} = a$, die Konjugation kann also ignoriert werden, solange \mathbb{R} der Grundkörper ist.

Geleitet von der Heuristik, zäumen wir nun das Pferd vom Schwanz her auf und beginnen die axiomatische Arbeit mit der Definition des inneren Produktes:

DEFINITION 1.2.1 (Standard-Skalarprodukt). Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Das Standard-Skalarprodukt von zwei Vektoren aus \mathbb{K}^n ist definiert durch

$$\left\langle \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix} \right\rangle_2 := \sum_{i=1}^n \bar{\alpha}_i \beta_i.$$

Andere Schreibweisen für Skalarprodukte sind (a, b) , $a \cdot b$, $\langle a | b \rangle$.

Wenn Verwechslungen ausgeschlossen sind, lässt man auch gerne den Subskript 2 weg und schreibt für das Standard-Skalarprodukt einfach $\langle a, b \rangle$. Aber wie wir aber bald sehen werden, gibt es auch andere Skalarprodukte.

Das Standard-Skalarprodukt auf \mathbb{R}^n heißt auch das euklidische Skalarprodukt.

Die Komponenten von einem der beiden Vektoren fließen konjugiert ins Skalarprodukt ein. Wir werden sehen, dass dadurch $\langle a, a \rangle_2$ immer nichtnegativ ist. Das ist wiederum die Grundlage, dass man aus dem inneren Produkt die Norm von a durch $\|a\| = \sqrt{\langle a, a \rangle_2}$ definieren kann.

Viele Autoren setzen bei der Definition des Standard-Skalarproduktes die Konjugation auf die rechte Seite:

$$\langle a, b \rangle_2 := \sum_{j=1}^n \alpha_j \bar{\beta}_j.$$

BEHAUPTUNG 1.2.2. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Seien $a, b, c \in \mathbb{K}^n$, $\lambda, \mu \in \mathbb{K}$. Dann gilt

(a) $\langle a, a \rangle_2 \geq 0$. Es ist $\langle a, a \rangle_2 = 0$ genau dann, wenn $a = 0$.

(b) $\langle a, b \rangle_2 = \overline{\langle b, a \rangle_2}$.

(c) Linearkombinationen lassen sich aus dem Standard-Skalarprodukt herausziehen:

$$\begin{aligned} \langle \lambda a + \mu b, c \rangle_2 &= \bar{\lambda} \langle a, c \rangle_2 + \bar{\mu} \langle b, c \rangle_2, \\ \langle c, \lambda a + \mu b \rangle_2 &= \lambda \langle c, a \rangle_2 + \mu \langle c, b \rangle_2. \end{aligned}$$

BEWEIS. Seien

$$a = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}, \quad b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}, \quad c = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix}.$$

(a) Es gilt

$$\langle a, a \rangle_2 = \sum_{j=1}^n \bar{\alpha}_j \alpha_j = \sum_{j=1}^n |\alpha_j|^2.$$

Weil in dieser Summe alle Summanden nichtnegative reelle Zahlen sind, ist die Summe eine nichtnegative reelle Zahl. Wenn die Summe Null ist, sind alle Summanden gleich Null, und daher jedes $\alpha_j = 0$.

(b)

$$\overline{\langle a, b \rangle_2} = \overline{\sum_{j=1}^n \bar{\alpha}_j \beta_j} = \sum_{j=1}^n \alpha_j \bar{\beta}_j = \sum_{j=1}^n \bar{\beta}_j \alpha_j = \langle b, a \rangle_2.$$

(c)

$$\begin{aligned}
\langle \lambda a + \mu b, c \rangle_2 &= \sum_{j=1}^n \overline{(\lambda \alpha_j + \mu \beta_j)} \gamma_j = \sum_{j=1}^n (\overline{\lambda \alpha_j} + \overline{\mu \beta_j}) \gamma_j \\
&= \overline{\lambda} \sum_{j=1}^n \overline{\alpha_j} \gamma_j + \overline{\mu} \sum_{j=1}^n \overline{\beta_j} \gamma_j = \overline{\lambda} \langle a, c \rangle_2 + \overline{\mu} \langle b, c \rangle_2, \\
\langle c, \lambda a + \mu b \rangle_2 &= \sum_{j=1}^n \overline{\gamma_j} (\lambda \alpha_j + \mu \beta_j) = \lambda \sum_{j=1}^n \overline{\gamma_j} \alpha_j + \mu \sum_{j=1}^n \overline{\gamma_j} \beta_j \\
&= \lambda \langle c, a \rangle_2 + \mu \langle c, b \rangle_2.
\end{aligned}$$

□

Steht bei der Definition des inneren Produktes die Konjugation auf der rechten Seite

$$\langle a, b \rangle_2 := \sum_{j=1}^n \alpha_j \overline{\beta_j},$$

dann sind manche Details entsprechend zu modifizieren:

$$\begin{aligned}
\langle \lambda a + \mu b, c \rangle_2 &= \lambda \langle a, c \rangle_2 + \mu \langle b, c \rangle_2, \\
\langle c, \lambda a + \mu b \rangle_2 &= \overline{\lambda} \langle c, a \rangle_2 + \overline{\mu} \langle c, b \rangle_2.
\end{aligned}$$

Eine Abbildung $\langle \cdot, \cdot \rangle$, die so wie das Standard-Skalarprodukt alle Eigenschaften aus Behauptung 1.2.2 erfüllt, nennen wir ein Skalarprodukt. Das Standard-Skalarprodukt ist also ein Spezialfall eines Skalarproduktes:

DEFINITION 1.2.3 (Skalarprodukt). Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Gegeben sei eine Abbildung

$$\begin{cases} \mathbb{K}^n \times \mathbb{K}^n & \rightarrow \mathbb{K} \\ (a, b) & \mapsto \langle a, b \rangle \end{cases}.$$

Die Abbildung $\langle \cdot, \cdot \rangle$ heißt ein Skalarprodukt oder inneres Produkt auf \mathbb{K}^n wenn gilt:

- (1) $\langle \cdot, \cdot \rangle$ ist positiv definit, das heißt für alle $a \in \mathbb{K}^n$

$$\langle a, a \rangle \geq 0 \quad \text{und} \quad \langle a, a \rangle = 0 \Rightarrow a = 0.$$

- (2) $\langle \cdot, \cdot \rangle$ ist hermitesch, das heißt für alle $a, b \in \mathbb{K}^n$

$$\langle a, b \rangle = \overline{\langle b, a \rangle}.$$

(Ist $\mathbb{K} = \mathbb{R}$, dann ist das äquivalent zur Symmetrie: $\langle a, b \rangle = \langle b, a \rangle$.)

- (3) $\langle \cdot, \cdot \rangle$ ist sesquilinear, das heißt für alle $a, b, c \in \mathbb{K}^n$ und alle $\lambda, \mu \in \mathbb{K}$

$$\begin{aligned}
\langle c, \lambda a + \mu b \rangle &= \lambda \langle c, a \rangle + \mu \langle c, b \rangle, \\
\langle \lambda a + \mu b, c \rangle &= \overline{\lambda} \langle a, c \rangle + \overline{\mu} \langle b, c \rangle.
\end{aligned}$$

Wenn $\mathbb{K} = \mathbb{R}$, dann fallen die Konjugationen weg, und man sagt, die Abbildung $\langle \cdot, \cdot \rangle$ ist bilinear.

Wir kommen nun zur Definition der Norm (des Betrages) eines Vektors:

DEFINITION 1.2.4 (Norm zu einem Skalarprodukt). Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $\langle \cdot, \cdot \rangle$ ein inneres Produkt auf \mathbb{K}^n . Die von $\langle \cdot, \cdot \rangle$ induzierte Norm eines Vektors in \mathbb{K}^n ist definiert als die positive Quadratwurzel

$$\|a\| := \sqrt{\langle a, a \rangle}.$$

Insbesondere induziert das Standard-Skalarprodukt die sogenannte euklidische Norm

$$\left\| \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \right\|_2 := \sqrt{|\alpha_1|^2 + \cdots + |\alpha_n|^2}.$$

Üblicherweise schreibt man einfach $\|a\|$ ohne Subskript für die euklidische Norm, wenn keine Verwechslungen zu erwarten sind. Es gibt aber auch andere Normen auf \mathbb{R}^n , nur haben diese keine so anschauliche Interpretation als Länge.

Manchmal werden die Normen auf \mathbb{R}^n auch mit einfachen Strichen geschrieben: $|a|$. Statt des Wortes Norm wird auch oft das Wort Betrag oder Absolutbetrag verwendet.

SATZ 1.2.5. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, und $\langle \cdot, \cdot \rangle$ ein inneres Produkt auf \mathbb{K}^n mit der induzierten Norm $\|\cdot\|$. Seien $a \in \mathbb{K}^n$, $\lambda \in \mathbb{K}$. Dann gilt:

- (a) $\|a\| \geq 0$. Wenn $\|a\| = 0$, dann ist $a = 0$.
 (b) $\|\lambda a\| = |\lambda| \|a\|$.

BEWEIS.

- (a) Folgt direkt aus der positiven Definitheit des inneren Produktes.
 (b)

$$\|\lambda a\|^2 = \langle \lambda a, \lambda a \rangle = \bar{\lambda} \lambda \langle a, a \rangle = |\lambda|^2 \|a\|^2.$$

□

Für die dritte wichtige Eigenschaft der Norm, die Dreiecksungleichung, brauchen wir als Hilfsmittel die Cauchy-Schwarz'sche Ungleichung. Sie ist eine der wichtigsten Ungleichungen überhaupt.

SATZ 1.2.6 (Cauchy-Schwarz'sche Ungleichung). Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, sei $\langle \cdot, \cdot \rangle$ ein inneres Produkt auf \mathbb{K}^n und $\|\cdot\|$ die induzierte Norm. Seien $a, b \in \mathbb{K}^n$. Es gilt

$$|\langle a, b \rangle| \leq \|a\| \|b\|.$$

Gleichheit gilt genau dann, wenn a und b linear abhängig sind.

BEWEIS. Wir machen eine Fallunterscheidung:

Fall 1: a und b sind linear abhängig. Dann gibt es einen Vektor $c \in \mathbb{K}^n$ und Skalare $\alpha, \beta \in \mathbb{K}$ so, dass $a = \alpha c$ und $b = \beta c$. Dann ist

$$\langle a, b \rangle = \langle \alpha c, \beta c \rangle = \bar{\alpha} \beta \langle c, c \rangle = \bar{\alpha} \beta \|c\|^2$$

und daher

$$|\langle a, b \rangle| = |\alpha| |\beta| \|c\|^2 = (|\alpha| \|c\|) (|\beta| \|c\|) = \|a\| \|b\|.$$

Fall 2: a und b sind linear unabhängig. Für jedes beliebige $\varepsilon \in \mathbb{K}$ gilt dann $a + \varepsilon b \neq 0$, also

$$\begin{aligned} 0 &< \|a + \varepsilon b\|^2 = \langle a + \varepsilon b, a + \varepsilon b \rangle \\ &= \langle a, a \rangle + \varepsilon \langle a, b \rangle + \bar{\varepsilon} \langle b, a \rangle + \varepsilon \bar{\varepsilon} \langle b, b \rangle \\ &= \|a\|^2 + \varepsilon \langle a, b \rangle + \bar{\varepsilon} \overline{\langle a, b \rangle} + |\varepsilon|^2 \|b\|^2. \end{aligned}$$

Das gilt insbesondere auch für

$$\varepsilon = -\frac{\langle b, a \rangle}{\|b\|^2}.$$

In diesem Fall ist

$$\varepsilon \langle a, b \rangle = -\frac{1}{\|b\|^2} \langle b, a \rangle \langle a, b \rangle = -\frac{1}{\|b\|^2} |\langle a, b \rangle|^2,$$

und die obige Abschätzung ergibt

$$\begin{aligned} 0 &< \|a\|^2 - \frac{1}{\|b\|^2} |\langle a, b \rangle|^2 - \frac{1}{\|b\|^2} |\langle a, b \rangle|^2 + \frac{1}{\|b\|^4} |\langle a, b \rangle|^2 \|b\|^2 \\ &= \|a\|^2 - \frac{1}{\|b\|^2} |\langle a, b \rangle|^2. \end{aligned}$$

Es folgt

$$|\langle a, b \rangle|^2 < \|a\|^2 \|b\|^2.$$

□

Beachten Sie die elegante Beweistechnik: eine Abschätzung zunächst mit einer beliebigen Konstanten angesetzt, und gegen Ende des Beweises wird die Konstante so gewählt, dass das gesuchte Ergebnis herauskommt!

SATZ 1.2.7 (Dreiecksungleichung). Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, $\|\cdot\|$ die euklidische Norm auf \mathbb{K}^n und $a, b \in \mathbb{K}^n$. Dann gilt die Dreiecksungleichung

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

BEWEIS. In der letzten Zeile der folgenden Abschätzung verwenden wir die Cauchy-Schwarz'sche Ungleichung.

$$\begin{aligned} \|a + b\|^2 &= \langle a + b, a + b \rangle = \langle a, a \rangle + \langle a, b \rangle + \langle b, a \rangle + \langle b, b \rangle \\ &= \|a\|^2 + \langle a, b \rangle + \overline{\langle a, b \rangle} + \|b\|^2 = \|a\|^2 + 2 \operatorname{Re} \langle a, b \rangle + \|b\|^2 \\ &\leq \|a\|^2 + 2 \|a\| \|b\| + \|b\|^2 = (\|a\| + \|b\|)^2. \end{aligned}$$

□

DEFINITION 1.2.8 (Norm). Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $\|\cdot\|$ eine Abbildung

$$\begin{cases} \mathbb{K}^n & \rightarrow \mathbb{R} \\ a & \mapsto \|a\| \end{cases}$$

Diese Abbildung heißt eine Norm auf \mathbb{K}^n wenn für alle $a, b \in \mathbb{K}^n$, $\lambda \in \mathbb{K}$ gilt:

$$\begin{aligned} \|a\| &\geq 0 \quad \text{und} \quad \|a\| = 0 \Rightarrow a = 0, \\ \|\lambda a\| &= |\lambda| \|a\|, \\ \|a + b\| &\leq \|a\| + \|b\|. \end{aligned}$$

Wenn also ein Skalarprodukt eine Norm induziert, ist das ein Spezialfall von einer Norm auf \mathbb{K}^n . Es gibt aber auch andere Normen, zum Beispiel für $1 \leq p < \infty$

$$\begin{aligned} \left\| \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \right\|_p &:= \left[\sum_{j=1}^n |\alpha_j|^p \right]^{1/p}, \\ \left\| \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \right\|_\infty &:= \max_{j=1, \dots, n} |\alpha_j|. \end{aligned}$$

1.3. Adjungierte Matrix.

Es gibt einen sehr nützlichen Trick, um das Standardskalarprodukt als ein Matrixprodukt zu schreiben. Dazu brauchen wir aber zuerst den Begriff der adjungierten Matrix:

DEFINITION 1.3.1 (Adjungierte Matrix). Ist $A \in \mathbb{C}^{m \times n}$ eine Matrix, so entsteht die adjungierte Matrix A^* aus A , indem man A transponiert und zu jedem Koeffizienten die konjugiert komplexe Zahl nimmt.

$$\begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{pmatrix}^* = \begin{pmatrix} \bar{\alpha}_{11} & \cdots & \bar{\alpha}_{m1} \\ \vdots & & \vdots \\ \bar{\alpha}_{1n} & \cdots & \bar{\alpha}_{mn} \end{pmatrix}.$$

Ist $A \in \mathbb{R}^{m \times n}$, so ist $A^* = A^T$. Die Adjungierte einer komplexen Zahl ist einfach die konjugiert komplexe Zahl.

Die folgenden Regeln ergeben sich durch einfaches Nachrechnen:

BEHAUPTUNG 1.3.2. Seien A, B, C Matrizen mit Zeilen- und Spaltenzahl, sodass die folgenden Operationen erlaubt sind. Dann gilt:

$$\begin{aligned} (A \pm B)^* &= A^* \pm B^*, \\ (AB)^* &= B^* A^*, \quad \text{Achtung, die Reihenfolge kehrt sich um!} \\ (A^*)^* &= A, \\ (A^{-1})^* &= (A^*)^{-1}. \end{aligned}$$

Jetzt können wir den Trick formulieren:

BEHAUPTUNG 1.3.3. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und seien $a, b \in \mathbb{K}^n$. Das euklidische innere Produkt lässt sich durch Matrixmultiplikation schreiben:

$$\langle a, b \rangle = a^* b.$$

2. Orthogonalität

Obwohl sich ein Großteil dieses Abschnittes ebenso gut für andere Skalarprodukte als das Standardskalarprodukt durchziehen lassen, folgen wir unserem didaktischen Konzept und vermeiden allzu große Allgemeinheit und Abstraktion. In diesem Abschnitt ist also immer $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, es bezeichnet $\langle \cdot, \cdot \rangle$ das Standard-Skalarprodukt und $\|\cdot\|$ die Standardnorm auf \mathbb{K}^n .

2.1. Orthogonales Komplement.

Eine geometrisch befriedigende Definition des Winkels würde einen viel tieferen, axiomatischen Zugang zur Elementargeometrie bedeuten. Für die lineare Algebra genügt es aber, zur Definition des Winkels den Cosinussatz heranzuziehen: Gegeben sei ein Dreieck im \mathbb{R}^n mit Ecken O, A, B mit $a = \vec{OA}$, $b = \vec{OB}$. Es sei γ der Winkel an der Ecke O . Der Cosinussatz aus der Elementargeometrie sagt:

$$\begin{aligned} &\|a\|^2 + \|b\|^2 - 2\|a\| \|b\| \cos(\gamma) \\ &= \|b - a\|^2 = \langle b - a, b - a \rangle \\ &= \|b\|^2 - 2\langle a, b \rangle + \|a\|^2 \end{aligned}$$

sodass

$$\cos(\gamma) = \frac{\langle a, b \rangle}{\|a\| \|b\|}$$

Wieder verwenden wir diese Formel als die Definition des Winkels:

DEFINITION 2.1.1. Seien $a, b \in \mathbb{R}^n \setminus \{0\}$. Der Winkel zwischen a und b ist definiert durch

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

Diese Definition braucht eine Rechtfertigung: Der Arcuscosinus kann nur von Zahlen genommen werden, die im Intervall $[-1, 1]$ liegen. Nun ist aber wegen Satz 1.2.6

$$\left| \frac{\langle a, b \rangle}{\|a\| \|b\|} \right| \leq 1.$$

Damit der Arcuscosinus eindeutig bestimmt ist, lassen wir nur Winkel im Intervall $[0, \pi]$ zu.

Unser besonderes Interesse gilt aber dem rechten Winkel:

DEFINITION 2.1.2. Seien $a, b \in \mathbb{K}^n$, U, V zwei Teilmengen des \mathbb{K}^n . Dann sagt man

- 1) Die Vektoren a und b stehen aufeinander orthogonal, wenn gilt $\langle a, b \rangle = 0$.
- 2) Der Vektor a steht orthogonal auf die Menge U , wenn gilt: $(\forall u \in U) \langle a, u \rangle = 0$.
- 3) Die Mengen U und V stehen orthogonal aufeinander, wenn gilt: $(\forall u \in U) (\forall v \in V) \langle u, v \rangle = 0$.
- 4) Das orthogonale Komplement von U ist definiert durch

$$U^\perp := \{x \in \mathbb{R}^n \mid x \text{ steht orthogonal auf } U\}.$$

Obwohl diese Definition des orthogonalen Komplementes für alle Teilmengen $U \subset \mathbb{K}^n$ möglich ist, verwendet man sie zumeist nur für Unterräume.

BEMERKUNG 2.1.3. Besonders leicht ist das Auffinden von Normalvektoren in \mathbb{R}^2 : Ist $a = (\alpha_1, \alpha_2)^T \in \mathbb{R}^2 \setminus \{0\}$, so sind die beiden Vektoren $(\pm\alpha_2, \mp\alpha_1)^T$ die einzigen Vektoren des \mathbb{R}^2 , welche auf a orthogonal stehen, und die gleiche euklidische Norm wie a haben.

BEWEIS. Leichte Rechnung! □

BEHAUPTUNG 2.1.4. Sei U ein Unterraum von \mathbb{K}^n . Es gilt

- (a) Das orthogonale Komplement U^\perp ist ein Unterraum von \mathbb{K}^n .
- (b) $U \cap U^\perp = \{0\}$.
- (c) Sei $U = \mathcal{L}(b_1, \dots, b_k)$, $x \in \mathbb{K}^n$. Dann ist $x \in U^\perp$ genau wenn $\langle b_j, x \rangle = 0$ für alle $j = 1 \dots k$.
- (d) Sei $U = \mathcal{L}(b_1, \dots, b_k)$ und sei $B \in \mathbb{K}^{n \times k}$ die Matrix mit den Spalten b_1, \dots, b_k . Dann ist $U^\perp = \ker(B^*)$.
- (e) $\dim(U) + \dim(U^\perp) = n$.
- (f) $(U^\perp)^\perp = U$.

BEWEIS.

- (a) Für alle $u \in U$ gilt $\langle u, 0 \rangle = 0$, daher ist $0 \in U^\perp$. Sind $x, y \in U^\perp$ und $\lambda, \mu \in \mathbb{K}$, so gilt für alle $u \in U$

$$\langle u, \lambda x + \mu y \rangle = \lambda \langle u, x \rangle + \mu \langle u, y \rangle = \lambda 0 + \mu 0 = 0.$$

- (b) Weil U und U^\perp Unterräume sind, ist jedenfalls $0 \in U \cap U^\perp$. Ist $x \in U \cap U^\perp$, dann steht x als Element von U^\perp orthogonal auf alle Elemente von U , insbesondere auch auf x : $0 = \langle x, x \rangle$. Damit ist aber $x = 0$.

(c) Ist $x \in U^\perp$, dann ist $\langle b_j, x \rangle = 0$, weil $b_j \in U$ und x auf alle Vektoren aus U orthogonal steht.

Sei nun umgekehrt $\langle b_j, x \rangle = 0$ für alle $j = 1 \cdots k$. Jedes $u \in U$ lässt sich als Linearkombination $u = \sum_{j=1}^k \lambda_j b_j$ schreiben. Dann ist

$$\langle x, u \rangle = \langle x, \sum_{j=1}^k \lambda_j b_j \rangle = \sum_{j=1}^k \lambda_j \langle x, b_j \rangle = 0.$$

Also ist dann $x \in U^\perp$.

(d) Die Zeilen des Gleichungssystems $B^*x = 0$ sind, ausgeschrieben,

$$0 = b_j^*x = \langle b_j, x \rangle.$$

Wegen Punkt (c) ist also $B^*x = 0$ äquivalent zu $x \in U^\perp$.

(e) Wir verwenden in Punkt (d) eine Basis (b_1, \dots, b_k) von U . Dann ist also $k = \dim(U)$. Die Matrix B^* hat k linear unabhängige Zeilen, also ist der Rang $\rho(B^*) = k$, und folglich ist $\dim(\ker(B^*)) = n - k$. Wegen Punkt (d) ist aber $\ker(B^*) = U^\perp$.

(f) Zunächst zeigen wir: $U \subseteq (U^\perp)^\perp$. Ist nämlich $u \in U$ und $v \in U^\perp$, so ist $\langle u, v \rangle = 0$ wegen der Definition von U^\perp . Da das für alle $v \in U^\perp$ gilt, ist $U \subseteq (U^\perp)^\perp$. Nun ist

$$\dim((U^\perp)^\perp) = n - \dim(U^\perp) = n - (n - \dim(U)) = \dim(U).$$

Also sind U und $(U^\perp)^\perp$ ineinandergeschachtelte Unterräume des \mathbb{K}^n mit derselben Dimension, und deshalb sind sie gleich. □

Dimensionsgründe spielen eine große Rolle in den obigen Beweisen. Tatsächlich kann man auch unendlich dimensionale Vektorräume und dort innere Produkte und orthogonale Komplemente definieren. Es gilt aber nicht mehr unbedingt $(U^\perp)^\perp = U$, und auch die orthogonale Projektion, die wir gleich kennen lernen, muss es in unendlich dimensionalen Räumen nicht für alle Unterräume geben.

LEMMA 2.1.5 (Gramsche Matrix). *Seien $b_1, \dots, b_k \in \mathbb{K}^n$. Wir betrachten die Matrix $B \in \mathbb{K}^{n \times k}$ mit den Spalten b_1, \dots, b_k und die (sogenannte Gramsche) Matrix*

$$G = B^*B = \begin{pmatrix} \langle b_1, b_1 \rangle & \cdots & \langle b_1, b_k \rangle \\ \vdots & & \vdots \\ \langle b_k, b_1 \rangle & \cdots & \langle b_k, b_k \rangle \end{pmatrix} \in \mathbb{K}^{k \times k}.$$

Es gilt

$$(a) \ker(B) = \left\{ \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_k \end{pmatrix} \mid \sum_{j=1}^k \xi_j b_j = 0 \right\}.$$

$$(b) \ker(G) = \ker(B).$$

(c) G ist genau dann regulär wenn das System (b_1, \dots, b_k) linear unabhängig ist.

BEWEIS.

(a) Bekanntlich ist Bx die folgende Linearkombination der Spalten von B

$$B \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_k \end{pmatrix} = \sum_{j=1}^k \xi_j b_j.$$

(b) Sei $x \in \ker(B)$. Dann ist

$$Gx = (B^*B)x = B^*(Bx) = B^*0 = 0.$$

Sei umgekehrt $x \in \ker(G)$. Dann ist

$$\|Bx\|^2 = \langle Bx, Bx \rangle = (Bx)^*(Bx) = x^*(B^*B)x = x^*Gx = x^*0 = 0,$$

also ist dann $Bx = 0$.

(c) G ist genau dann regulär, wenn $\ker(G)$ nur aus dem Nullvektor besteht. Das ist wegen Punkten (a) und (b) äquivalent dazu, dass $\xi_1 = \dots = \xi_k = 0$ die einzige Lösung von $\sum_{j=1}^k \xi_j b_j = 0$ ist. Und das bedeutet wiederum, dass das System (b_1, \dots, b_k) linear unabhängig ist. □

SATZ 2.1.6 (orthogonale Projektion). Sei U ein Unterraum von \mathbb{K}^n mit Basis (b_1, \dots, b_k) . Sei $x \in \mathbb{K}^n$. Dann gilt:

- (a) Es gibt eindeutig bestimmte $u \in U$, $v \in U^\perp$ sodass $x = u + v$.
 (b) $u = B(B^*B)^{-1}B^*x$.

BEWEIS. Die kürzeste Beweismöglichkeit wäre wahrscheinlich, einfach anzusetzen $u = B(B^*B)^{-1}B^*x$, und dann zu zeigen: $u \in U$, $x - u \in U^\perp$. Wir führen den Beweis etwas langatmiger, aber anschaulicher. Wenn man sich den Beweisansatz merkt, muss man sich die Formel nicht merken.

Gesucht ist also $u \in U$ so, dass $x - u \in U^\perp$. Weil (b_1, \dots, b_k) den Raum U aufspannen, können wir u ansetzen

$$u = \sum_{j=1}^k \eta_j b_j = B \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_k \end{pmatrix}.$$

Die Bedingung $x - u \in U^\perp$ ist äquivalent zu:

$$(\forall i = 1 \dots k) \langle b_i, x - u \rangle = 0,$$

$$\text{d.h. } (\forall i = 1 \dots k) \langle b_i, u \rangle = \langle b_i, x \rangle,$$

$$\text{d.h. } (\forall i = 1 \dots k) \langle b_i, \sum_{j=1}^k \eta_j b_j \rangle = \langle b_i, x \rangle.$$

Ausgeschrieben erhält man

$$\begin{array}{ccccccc} \langle b_1, b_1 \rangle \eta_1 & + & \dots & + & \langle b_1, b_k \rangle \eta_k & = & \langle b_1, x \rangle \\ \vdots & & & & \vdots & & \vdots \\ \langle b_k, b_1 \rangle \eta_1 & + & \dots & + & \langle b_k, b_k \rangle \eta_k & = & \langle b_k, x \rangle \end{array}$$

Das ist ein lineares Gleichungssystem $Gy = h$ mit der Gramschen Matrix $G = B^*B$ als Systemmatrix und der rechten Seite $h = B^*x$. Weil das System (b_1, \dots, b_k) linear unabhängig ist, ist G nach Lemma 2.1.5 regulär, und das Gleichungssystem besitzt eine einzige Lösung y . Damit löst $u = By = BG^{-1}B^*x$ auch die Bedingung $(u - x) \in U^\perp$. □

DEFINITION 2.1.7 (orthogonale Projektion). Sei U ein Unterraum des \mathbb{K}^n und sei $x \in \mathbb{K}^n$.

Die orthogonale Projektion von x auf U ist jener Vektor $u \in \mathbb{K}^n$, für den gilt:

$$u \in U \quad \text{und} \quad (x - u) \in U^\perp.$$

Die Abbildung $\pi_U : \mathbb{K}^n \rightarrow \mathbb{K}^n$ ordnet jedem x in \mathbb{K}^n seine orthogonale Projektion auf U zu.

BEISPIEL 2.1.8. Seien $b_1, b_2, x \in \mathbb{R}^4$, und x wie folgt gegeben:

$$b_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \\ 1 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad x = \begin{pmatrix} 7 \\ -5 \\ 11 \\ 6 \end{pmatrix}.$$

Wir berechnen die orthogonale Projektion von x auf die lineare Hülle $\mathcal{L}(b_1, b_2)$.

Lösung: Es ist

$$\begin{aligned} \langle b_1, b_1 \rangle &= 10, \\ \langle b_1, b_2 \rangle &= 5, \\ \langle b_2, b_2 \rangle &= 3, \\ \langle b_1, x \rangle &= 25, \\ \langle b_2, x \rangle &= 13. \end{aligned}$$

Wir lösen das Gleichungssystem

$$\begin{pmatrix} 10 & 5 \\ 5 & 3 \end{pmatrix} \begin{pmatrix} \eta_1 \\ \eta_2 \end{pmatrix} = \begin{pmatrix} 25 \\ 13 \end{pmatrix}.$$

$$\begin{pmatrix} 10 & 5 & 25 \\ 5 & 3 & 13 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & -1 & -1 \\ 5 & 3 & 13 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 1 & 1 \\ 5 & 0 & 10 \end{pmatrix}$$

Also ist $\eta_1 = 2$, $\eta_2 = 1$, und

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

Eine besonders wichtige Eigenschaft der orthogonalen Projektion ist, dass sie den nächsten Punkt aus einem Unterraum an einen gegebenen Punkt liefert:

SATZ 2.1.9. Sei U ein Unterraum des \mathbb{K}^n und sei $x \in \mathbb{K}^n$. Sei u die orthogonale Projektion von x auf U . Dann gilt:

$$(\forall w \in U) \|x - u\| \leq \|x - w\|.$$

Dabei gilt Gleichheit nur für $w = u$ selbst.

BEWEIS. Sei $w \in U$. Dann ist $u - w \in U$ und daher orthogonal auf $x - u$. Es folgt

$$\begin{aligned} \|x - w\|^2 &= \|(x - u) + (u - w)\|^2 \\ &= \|x - u\|^2 + \langle x - u, u - w \rangle + \langle u - w, x - u \rangle + \|u - w\|^2 \\ &= \|x - u\|^2 + \|u - w\|^2. \end{aligned}$$

□

2.2. Orthonormalsysteme.

DEFINITION 2.2.1 (Orthonormalsystem). Ein Vektorsystem (a_1, \dots, a_k) in \mathbb{K}^n heißt ein Orthonormalsystem, wenn gilt

$$\langle a_i, a_j \rangle = \begin{cases} 0 & \text{wenn } i \neq j \\ 1 & \text{wenn } i = j \end{cases}.$$

Wenn das Orthonormalsystem zugleich Basis eines Unterraumes U von \mathbb{K}^n ist, dann heißt es eine Orthonormalbasis.

BEHAUPTUNG 2.2.2. Orthonormalsysteme sind immer linear unabhängig.

BEWEIS. Sei (a_1, \dots, a_k) ein Orthonormalsystem und $\sum_{i=1}^k \lambda_i a_i = 0$. Wir nehmen das innere Produkt mit a_j . In der folgenden Rechnung verschwinden alle inneren Produkte $\langle a_j, a_i \rangle$ außer $\langle a_j, a_j \rangle$:

$$0 = \langle a_j, \sum_{i=1}^k \lambda_i a_i \rangle = \sum_{i=1}^k \lambda_i \langle a_j, a_i \rangle = \lambda_j$$

Das geht für alle $j = 1, \dots, k$, daher sind alle $\lambda_j = 0$. \square

BEHAUPTUNG 2.2.3. Sei (a_1, \dots, a_k) ein Orthonormalsystem in \mathbb{K}^n und seien $\gamma_1, \dots, \gamma_k \in \mathbb{K}$. Dann ist

$$\left\| \sum_{i=1}^k \gamma_i a_i \right\|^2 = \sum_{i=1}^k |\gamma_i|^2.$$

BEWEIS.

$$\left\| \sum_{i=1}^k \gamma_i a_i \right\|^2 = \left\langle \sum_{i=1}^k \gamma_i a_i, \sum_{j=1}^k \gamma_j a_j \right\rangle = \sum_{i=1}^k \sum_{j=1}^k \overline{\gamma_i} \gamma_j \langle a_i, a_j \rangle = \sum_{i=1}^k \overline{\gamma_i} \gamma_i = \sum_{i=1}^k |\gamma_i|^2.$$

\square

Wenn man zu einem Unterraum U eine Orthonormalbasis kennt, ist die Bestimmung der orthogonalen Projektion auf U besonders einfach:

SATZ 2.2.4. Sei U ein Unterraum des \mathbb{K}^n mit einer Orthonormalbasis (a_1, \dots, a_k) . Sei $x \in \mathbb{K}^n$.

(1) Dann ist die orthogonale Projektion von x auf U

$$\pi_U(x) = \sum_{i=1}^k \langle a_i, x \rangle a_i.$$

(2) (Besselsche Ungleichung bzw., bei $x = u$, Parsevalsche Gleichung:)

$$\sum_{i=1}^k |\langle a_i, x \rangle|^2 \leq \|x\|^2,$$

dabei gilt Gleichheit genau dann, wenn $x \in U$.

BEWEIS.

(1) Sei

$$u := \sum_{i=1}^k \gamma_i a_i \text{ mit } \gamma_i = \langle a_i, x \rangle.$$

Weil u eine Linearkombination der a_i ist, ist $u \in U$. Wir zeigen noch, dass $x - u$ orthogonal auf ganz U steht, und dazu genügt, dass $x - u$ orthogonal auf jeden Basisvektor a_j steht. Es ist

$$\begin{aligned} \langle a_j, x - u \rangle &= \langle a_j, x - \sum_{i=1}^k \gamma_i a_i \rangle \\ &= \langle a_j, x \rangle - \sum_{i=1}^k \gamma_i \langle a_j, a_i \rangle = \langle a_j, x \rangle - \gamma_j = 0. \end{aligned}$$

(2) Es ist $\langle u, x - u \rangle = 0$ und daher ist

$$\|x\|^2 = \|x - u\|^2 + \|u\|^2 \geq \|u\|^2.$$

gleichheit gilt genau dann, wenn $\|x - u\|^2 = 0$, also wenn $x = u$. Wegen Behauptung 2.2.3 ist aber

$$\|u\|^2 = \left\| \sum_{i=1}^k \langle a_i, x \rangle a_i \right\|^2 = \sum_{i=1}^k |\langle a_i, x \rangle|^2.$$

□

SATZ 2.2.5 (Gram-Schmidtsches Orthonormalisierungsverfahren).

(1) Sei (b_1, \dots, b_k) ein linear unabhängiges Vektorsystem in \mathbb{K}^n . Dann gibt es ein Orthonormalsystem (a_1, \dots, a_k) , so dass für alle $m = 1, \dots, k$ die folgenden linearen Hüllen gleich sind:

$$\mathcal{L}(b_1, \dots, b_m) = \mathcal{L}(a_1, \dots, a_m).$$

(2) Insbesondere besitzt jeder Unterraum des \mathbb{K}^n eine Orthonormalbasis.

(3) Das Orthonormalsystem (a_1, \dots, a_k) kann folgendermaßen bestimmt werden:

$$\tilde{a}_1 = b_1,$$

$$\text{für } m = 2, \dots, k: \quad \tilde{a}_m = b_m - \sum_{j=1}^{m-1} \frac{1}{\|\tilde{a}_j\|^2} \langle \tilde{a}_j, b_m \rangle \tilde{a}_j,$$

$$\text{für } m = 1, \dots, k: \quad a_m = \frac{1}{\|\tilde{a}_m\|} \tilde{a}_m.$$

BEWEIS. Wir bestimmen (a_1, \dots, a_k) wie in Punkt (3) und zeigen durch Induktion nach m folgende Eigenschaften:

- (1) $\tilde{a}_m \neq 0$,
- (2) \tilde{a}_m steht orthogonal auf alle \tilde{a}_j mit $j < m$,
- (3) $\mathcal{L}(\tilde{a}_1, \dots, \tilde{a}_m) = \mathcal{L}(b_1, \dots, b_m)$.

Verankerung: Weil das System (b_1, \dots, b_k) linear unabhängig ist, ist $b_1 \neq 0$. Weil $\tilde{a}_1 = b_1$ ist also $\tilde{a}_1 \neq 0$ und $\mathcal{L}(\tilde{a}_1) = \mathcal{L}(b_1)$.

Induktionsschritt von m auf $m + 1$: Wir nehmen nun an, dass Punkte (1), (2), (3) für m gelten. Wegen der linearen Unabhängigkeit ist $b_{m+1} \notin \mathcal{L}(b_1, \dots, b_m)$ und wegen Punkt (1) der Induktionsannahme ist $b_{m+1} \notin \mathcal{L}(\tilde{a}_1, \dots, \tilde{a}_m)$. Daher muss gelten

$$b_{m+1} \neq \sum_{j=1}^m \frac{1}{\|\tilde{a}_j\|^2} \langle \tilde{a}_j, b_{m+1} \rangle \tilde{a}_j,$$

und daher ist $\tilde{a}_{m+1} \neq 0$.

Sei nun $i < m + 1$. Wir nützen in der folgenden Rechnung, dass \tilde{a}_i orthogonal auf alle \tilde{a}_j mit $j = 1, \dots, m$ steht:

$$\begin{aligned} & \langle \tilde{a}_i, \tilde{a}_{m+1} \rangle \\ &= \left\langle \tilde{a}_i, b_{m+1} - \sum_{j=1}^m \frac{1}{\|\tilde{a}_j\|^2} \langle \tilde{a}_j, b_{m+1} \rangle \tilde{a}_j \right\rangle \\ &= \langle \tilde{a}_i, b_{m+1} \rangle - \sum_{j=1}^m \frac{1}{\|\tilde{a}_j\|^2} \langle \tilde{a}_j, b_{m+1} \rangle \langle \tilde{a}_i, \tilde{a}_j \rangle \\ &= \langle \tilde{a}_i, b_{m+1} \rangle - \frac{1}{\|\tilde{a}_i\|^2} \langle \tilde{a}_i, b_{m+1} \rangle \langle \tilde{a}_i, \tilde{a}_i \rangle = 0. \end{aligned}$$

Es gilt

$\tilde{a}_j \in \mathcal{L}(\tilde{a}_1, \dots, \tilde{a}_m) = \mathcal{L}(b_1, \dots, b_m) \subseteq \mathcal{L}(b_1, \dots, b_{m+1})$ für $j = 1, \dots, m$,
und daher auch (mit geeigneten γ_j)

$$\tilde{a}_{m+1} = b_{m+1} - \sum_{j=1}^m \gamma_j \tilde{a}_j \in \mathcal{L}(b_1, \dots, b_{m+1}).$$

Daraus folgt

$$\mathcal{L}(\tilde{a}_1, \dots, \tilde{a}_{m+1}) \subseteq \mathcal{L}(b_1, \dots, b_{m+1}).$$

Umgekehrt ist

$$b_j \in \mathcal{L}(b_1, \dots, b_m) = \mathcal{L}(\tilde{a}_1, \dots, \tilde{a}_m) \subseteq \mathcal{L}(\tilde{a}_1, \dots, \tilde{a}_{m+1}) \quad \text{für } j = 1, \dots, m$$

und

$$b_{m+1} = \tilde{a}_{m+1} + \sum_{j=1}^m \gamma_j \tilde{a}_j \in \mathcal{L}(\tilde{a}_1, \dots, \tilde{a}_{m+1}).$$

Daher ist

$$\mathcal{L}(b_1, \dots, b_{m+1}) \subseteq \mathcal{L}(\tilde{a}_1, \dots, \tilde{a}_{m+1}).$$

□

BEISPIEL 2.2.6. Wir suchen eine Orthonormalbasis für die lineare Hülle von b_1, b_2, b_3 :

$$b_1 = \begin{pmatrix} 2 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \quad b_2 = \begin{pmatrix} 3 \\ 3 \\ 0 \\ 4 \end{pmatrix}, \quad b_3 = \begin{pmatrix} 2 \\ 9 \\ -2 \\ -6 \end{pmatrix}$$

Lösung:

$$\tilde{a}_1 = \begin{pmatrix} 2 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \quad \|\tilde{a}_1\| = 3.$$

$$\tilde{a}_2 = \begin{pmatrix} 3 \\ 3 \\ 0 \\ 4 \end{pmatrix} - \frac{1}{9} \left\langle \begin{pmatrix} 2 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 3 \\ 0 \\ 4 \end{pmatrix} \right\rangle \begin{pmatrix} 2 \\ 1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 0 \\ 4 \end{pmatrix} - \begin{pmatrix} 2 \\ 1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ -2 \\ 4 \end{pmatrix}$$

$$\|\tilde{a}_2\| = 5.$$

$$\tilde{a}_3 = \begin{pmatrix} 2 \\ 9 \\ -2 \\ -6 \end{pmatrix} - \frac{1}{9} \left\langle \begin{pmatrix} 2 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 9 \\ -2 \\ -6 \end{pmatrix} \right\rangle \begin{pmatrix} 2 \\ 1 \\ 2 \\ 0 \end{pmatrix} - \frac{1}{25} \left\langle \begin{pmatrix} 1 \\ 2 \\ -2 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 9 \\ -2 \\ -6 \end{pmatrix} \right\rangle \begin{pmatrix} 1 \\ 2 \\ -2 \\ 4 \end{pmatrix}$$

$$= \begin{pmatrix} 2 \\ 9 \\ -2 \\ -6 \end{pmatrix} - \begin{pmatrix} 2 \\ 1 \\ 2 \\ 0 \end{pmatrix} - 0 \begin{pmatrix} 1 \\ 2 \\ -2 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 8 \\ -4 \\ -6 \end{pmatrix},$$

$$\|\tilde{a}_3\| = \sqrt{116} = 2\sqrt{29}.$$

Die Normierung ergibt

$$a_1 = \frac{1}{3} \begin{pmatrix} 2 \\ 1 \\ 2 \\ 0 \end{pmatrix}, \quad a_2 = \frac{1}{5} \begin{pmatrix} 1 \\ 2 \\ -2 \\ 4 \end{pmatrix}, \quad a_3 = \frac{1}{\sqrt{29}} \begin{pmatrix} 0 \\ 8 \\ -4 \\ -6 \end{pmatrix}.$$

□

Der abstrakte Vektorraum

So, wie wir unsere Arbeit über dem Körper \mathbb{R} der reellen Zahlen begonnen haben, und später die Körperaxiome herausdestilliert haben, die ausreichen, um die ganze Matrizen- und Vektorrechnung im \mathbb{K}^n zu entwickeln, so haben wir auch mit einem ganz speziellen Vektorraum begonnen, dem Raum der n -dimensionalen Spaltenvektoren. Wir werden nun sehen, dass für viele Begriffe und Lehrsätze, die wir gefunden haben, ein viel allgemeinerer und abstrakterer Rahmen ausreicht: Der Vektorraum. Die hohe Abstraktion hat Vorteile: Mit der abstrakten Theorie lässt sich die geometrische Vorstellung, die wir aus \mathbb{R}^n mitbringen, in Analogie auf viel komplexere Objekte anwenden. Allerdings müssen Vektorräume nicht unbedingt endliche Basen haben. Auf die Bequemlichkeiten der Matrizenrechnung und die Folgerungen, die sich aus endlicher Dimension ergeben, müssen wir auf der höchsten Abstraktionsstufe verzichten.

1. Definitionen und Grundbegriffe

1.1. Definition von Vektorräumen, einfache Rechenregeln.

In der folgenden Definition verwenden wir, um den Unterschied zu unterstreichen, verschiedene Zeichen \oplus und $+$ für die Addition von Vektoren und Skalaren, und verschiedene Zeichen \odot und \cdot für die Multiplikation eines Vektors mit einem Skalar und zwischen zwei Skalaren. In der Praxis schreibt man für beides einfach $+$ beziehungsweise \cdot . Für den Nullvektor verwenden wir hier das Zeichen Θ , in der Praxis schreibt man einfach 0 .

DEFINITION 1.1.1. Sei $(\mathbb{K}, +, \cdot, 0, 1)$ ein Körper, und V eine nichtleere Menge. Gegeben seien zwei Verknüpfungen \oplus und \odot :

$$\oplus \begin{cases} V \times V & \rightarrow V \\ (a, b) & \mapsto a \oplus b \end{cases}, \quad \odot \begin{cases} \mathbb{K} \times V & \rightarrow V \\ (\alpha, a) & \mapsto \alpha \odot a \end{cases}$$

(V, \oplus, \odot) heißt ein Vektorraum über \mathbb{K} , wenn folgende Gesetze gelten:

- (1) (V, \oplus) ist eine abelsche Gruppe, das heißt
 - \oplus ist assoziativ, das heißt, für alle $x, y, z \in V$ gilt $(x \oplus y) \oplus z = x \oplus (y \oplus z)$.
 - Es gibt ein Element $\Theta \in V$, sodass für alle $x \in V$ gilt $x \oplus \Theta = x$.
 - Zu jedem $x \in V$ gibt es ein Element $(-x) \in V$, sodass $x \oplus (-x) = \Theta$.
 - \oplus ist kommutativ: $x \oplus y = y \oplus x$ für alle $x, y \in V$.

(2)

$$(\forall \alpha, \beta \in \mathbb{K}, v \in V) (\alpha\beta) \odot v = \alpha \odot (\beta \odot v).$$

- (3) Es gelten die Distributivgesetze (für alle $\alpha, \beta \in \mathbb{K}, v, w \in V$):

$$\begin{aligned} \alpha \odot (v \oplus w) &= \alpha \odot v \oplus \alpha \odot w, \\ (\alpha + \beta) \odot v &= \alpha \odot v \oplus \beta \odot v. \end{aligned}$$

- 4) Für das Einselement 1 des Körpers gilt

$$(\forall v \in V) 1 \odot v = v.$$

Die Elemente von V heißen dann Vektoren, die Elemente von \mathbb{K} heißen Skalare. Das neutrale Element $\Theta \in V$ heißt der Nullvektor. Für zwei Vektoren $u, v \in V$ definieren wir $u - v := u \oplus (-v)$.

Schreibweise: Bei der Multiplikation eines Skalars mit einem Vektor schreibt man den Skalar immer vor dem Vektor. Die Abbildung \odot ist ja auf $\mathbb{K} \times V$ definiert, nicht auf $V \times \mathbb{K}$.

BEISPIEL 1.1.2. Sei \mathbb{K} ein Körper. Der Raum \mathbb{K}^n der n -dimensionalen Spaltenvektoren, und der Raum $\mathbb{K}^{m \times n}$ der $(m \times n)$ -Matrizen über \mathbb{K} sind Vektorräume.

BEISPIEL 1.1.3. Ist L ein Körper, der einen Körper K als Unterkörper enthält, so ist L mit den Körperverknüpfungen $+$, \cdot , als Vektorraumaddition und Multiplikation mit Skalaren aufgefasst, ein Vektorraum über K . Beispielsweise ist \mathbb{C} ein zweidimensionaler Vektorraum über \mathbb{R} .

In dieser Situation nennt man L eine Körpererweiterung von K . Die Theorie der Körpererweiterungen und die Möglichkeit, den größeren Körper als Vektorraum über dem kleineren Körper aufzufassen, spielen eine große Rolle bei der Untersuchung der Lösbarkeit von Gleichungen. Beispielsweise führt diese Theorie zum bekannten Satz, dass man nicht jeden Winkel mit Zirkel und Lineal dritteln kann.

Dem besonders wichtigen Beispiel der Funktionenräume werden wir einen eigenen Abschnitt widmen.

Nachdem eine neue algebraische Struktur (hier: der Vektorraum) eingeführt ist, und durch Beispiele belegt ist, dass diese Struktur auf interessante und wichtige Objekte angewendet werden kann, werden in der Algebra typischerweise viele kleine Rechenregeln bewiesen, die in dieser Struktur gelten. Beachten Sie, dass wir nur die Axiome der Struktur dazu verwenden dürfen, kein Vorwissen über Beispiele dieser Struktur. Die folgende Behauptung ist ein Beispiel solcher Regeln, es gibt noch viele andere.

BEHAUPTUNG 1.1.4. Sei $(V, +, \cdot)$ ein Vektorraum über einem Körper \mathbb{K} . Ausnahmsweise verwenden wir zwei verschiedene Symbole für Null im Körper (0) und Nullvektor (Θ).

(1) Es gilt die Kürzungsregel: Für alle $u, v, w \in V$ gilt

$$u + w = v + w \Rightarrow u = v.$$

(2) $(\forall v \in V) 0v = \Theta$.

(3) $(\forall \alpha \in \mathbb{K}) \alpha\Theta = \Theta$.

(4) $(\forall \alpha \in \mathbb{K}, v \in V) \alpha v = \Theta \Rightarrow \alpha = 0 \vee v = \Theta$.

(5) $(\forall v \in V) (-1)v = -v$.

BEWEIS. (1)

$$\begin{aligned} u + w = v + w &\Rightarrow (u + w) + (-w) = (v + w) + (-w) \\ &\Rightarrow u + (w + (-w)) = v + (w + (-w)) \Rightarrow u + \Theta = v + \Theta \Rightarrow u = v. \end{aligned}$$

(2) Aus $0 + 0 = 0$ folgern wir mittels Distributivität:

$$0v + 0v = (0 + 0)v = 0v = 0v + \Theta.$$

Nun liefert die Kürzungsregel $0v = \Theta$.

(3) Wir beginnen mit $\Theta + \Theta = \Theta$ und verwenden Distributivität:

$$\alpha\Theta + \alpha\Theta = \alpha(\Theta + \Theta) = \alpha\Theta = \alpha\Theta + \Theta.$$

Die Kürzungsregel liefert $\alpha\Theta = \Theta$.

(4) Sei $\alpha v = \Theta$ aber $\alpha \neq 0$. Zeige: $v = \Theta$. Da $\alpha \neq 0$, gibt es das inverse Element α^{-1} , und es ist

$$\Theta = \alpha^{-1}\Theta = \alpha^{-1}(\alpha v) = (\alpha^{-1}\alpha)v = 1v = v.$$

(5) Es gilt wegen Distributivität:

$$v + (-1)v = 1v + (-1)v = (1 + (-1))v = 0v = \Theta.$$

Also ist $(-1)v$ invers zu v bezüglich der Vektorraumaddition. □

1.2. Linearkombinationen und Unterräume.

Der nächste typische Schritt der Theoriebildung für eine algebraische Struktur ist die Definition von Unterstrukturen. Linearkombination und Unterraum sind zentrale Begriffe, die wir in der Theorie des \mathbb{K}^n eingeführt haben. Dieser Abschnitt zeigt nun, dass diese Begriffe in jedem Vektorraum sinnvoll sind. Damit ist er, mit kleinen Modifikationen, eine Wiederholung dessen, was wir zu Beginn der Vorlesung Lineare Algebra 1 erarbeitet haben.

Beobachten Sie aber, dass in der Theorie, wie wir sie jetzt entwickeln, (vorläufig) nirgends Begriffe wie Koordinaten oder Dimension vorkommen.

DEFINITION 1.2.1 (Linearkombination, lineare Hülle, Erzeugendensystem). Sei V ein Vektorraum über einem Körper \mathbb{K} .

(1) Seien $v_1, \dots, v_k \in V$ und $\lambda_1, \dots, \lambda_k \in \mathbb{K}$. Der Vektor

$$\sum_{i=1}^k \lambda_i v_i$$

heißt eine Linearkombination der Vektoren v_1, \dots, v_k .

(2) Sei S eine nichtleere Teilmenge von V . Die lineare Hülle von S ist die Menge aller Linearkombinationen von Vektoren aus S :

$$\mathcal{L}(S) := \left\{ \sum_{i=1}^k \lambda_i v_i \mid k \in \mathbb{N}, \lambda_i \in \mathbb{K}, v_i \in S \right\}.$$

(3) Die lineare Hülle der leeren Menge ist der Nullraum:

$$\mathcal{L}(\emptyset) := \{0\}.$$

(4) Ist U die lineare Hülle einer Menge S , dann bezeichnen wir S als ein Erzeugendensystem von U .

Es gibt einen kleinen Unterschied zwischen der Art, wie wir im Abschnitt über Matrizenrechnung die lineare Hülle eingeführt haben, und wie wir sie eben wieder definiert haben: In Definition 3.3.1 gehen wir von einem geordneten k -Tupel (v_1, \dots, v_k) von Vektoren aus, in Definition 1.2.1 gehen wir von einer Menge von Vektoren aus. Das erlaubt, auch die linearen Hüllen von unendlichen Mengen zu definieren. Für die Praxis spielt der Unterschied keine wesentliche Rolle.

DEFINITION 1.2.2 (Unterraum). Sei V ein Vektorraum über einem Körper \mathbb{K} und U eine Teilmenge von V . Dann heißt U ein Unterraum von V , wenn gilt:

- (1) $U \neq \emptyset$,
- (2) Für alle $\alpha, \beta \in \mathbb{K}$ und alle $u, v \in U$ gilt $\alpha u + \beta v \in U$.

BEHAUPTUNG 1.2.3. Sei V ein Vektorraum über \mathbb{K} und U ein Unterraum von V . Dann gilt:

- (1) $0 \in U$.
- (2) Sind $v_1, \dots, v_k \in U$ und $\lambda_1, \dots, \lambda_k \in \mathbb{K}$, so ist $\sum_{i=1}^k \lambda_i v_i \in U$.

BEWEIS.

- (1) Nach Definition 1.2.2 ist $U \neq \emptyset$. Wir wählen ein beliebiges $u \in U$ und setzen $\lambda = 0$. Es ist dann $0 = \lambda u + \lambda u \in U$.
- (2) Induktion nach k . □

BEHAUPTUNG 1.2.4. Sei V ein Vektorraum über einem Körper \mathbb{K} . Dann gilt: Der Nullraum $\{0\}$ ist ein Unterraum von V .

BEWEIS. Natürlich ist $\{0\}$ nicht leer, denn es enthält ja den Nullvektor. Jede Linearkombination des Nullvektors ist aber wieder der Nullvektor, und damit ist $\{0\}$ ein Unterraum. \square

BEHAUPTUNG 1.2.5. Sei V ein Vektorraum über einem Körper \mathbb{K} und $S \subseteq V$. Dann gilt:

Die lineare Hülle $\mathcal{L}(S)$ ist der kleinste Unterraum, der S enthält, das heißt

- (1) $\mathcal{L}(S)$ ist ein Unterraum.
- (2) Wenn U ein Unterraum von V ist, der S enthält, dann gilt auch $\mathcal{L}(S) \subseteq U$.

BEWEIS. Wir betrachten zuerst den Sonderfall $S = \emptyset$. Dann ist also nach Definition 1.2.1 $\mathcal{L}(S) = \{0\}$. Und das ist, wie oben bewiesen, ein Unterraum. Weil jeder Unterraum den Nullvektor enthält, ist der Nullraum in jedem Unterraum als Teilmenge enthalten, und somit der kleinste Unterraum überhaupt.

Wir betrachten nun $S \neq \emptyset$.

- (1) Sei v ein beliebiger Vektor aus S . Dann ist $0 = 0v \in \mathcal{L}(S)$, und daher ist $\mathcal{L}(S)$ nicht leer. Seien $u, v \in \mathcal{L}(S)$ und $\alpha, \beta \in \mathbb{K}$. Nach Definition von $\mathcal{L}(S)$ lassen sich u und v als Linearkombinationen schreiben:

$$u = \sum_{i=1}^p \lambda_i v_i, \quad v = \sum_{j=1}^q \mu_j w_j \quad \text{mit } \lambda_i, \mu_j \in \mathbb{K} \text{ und } v_i, w_j \in S.$$

Dann ist

$$\alpha u + \beta v = \sum_{i=1}^p \alpha \lambda_i v_i + \sum_{j=1}^q \beta \mu_j w_j$$

eine Linearkombination von Vektoren aus S und daher ist

$$\alpha u + \beta v \in \mathcal{L}(S).$$

Also ist $\mathcal{L}(S)$ ein Unterraum von V .

- (2) Sei U ein Unterraum von V , welcher S enthält, und sei $u \in \mathcal{L}(S)$. Dann hat u die Gestalt

$$u = \sum_{i=1}^k \lambda_i v_i \quad \text{mit } \lambda_i \in \mathbb{K}, v_i \in S.$$

Weil $S \subseteq U$, liegt jedes $v_i \in U$. Weil U ein Unterraum ist, führt die Linearkombination von Vektoren aus U nicht aus U heraus, und es ist $u \in U$. Also ist $\mathcal{L}(S) \subseteq U$. \square

1.3. Lineare Abbildungen in allgemeinen Vektorräumen.

DEFINITION 1.3.1 (Lineare Abbildung). Sei \mathbb{K} ein Körper und V und W zwei Vektorräume über \mathbb{K} . Sei $f : V \rightarrow W$ eine Abbildung.

- (1) Die Abbildung f heißt eine lineare Abbildung oder ein (Vektorraum-) Homomorphismus, wenn gilt

$$(\forall \alpha, \beta \in \mathbb{K}) (\forall u, v \in V) f(\alpha u + \beta v) = \alpha f(u) + \beta f(v).$$

- (2) Die Menge der linearen Abbildungen $f : V \rightarrow W$ bezeichnet man mit $\text{Hom}(V, W)$.

BEHAUPTUNG 1.3.2. Sei \mathbb{K} ein Körper und seien V, W Vektorräume über \mathbb{K} . Für $f, g \in \text{Hom}(V, W)$, $\lambda \in \mathbb{K}$ definieren wir die Abbildungen λf und $f + g$ so, dass für alle $v \in V$ gilt:

$$\begin{aligned} [\lambda f](v) &:= \lambda f(v), \\ [f + g](v) &:= f(v) + g(v). \end{aligned}$$

Dann sind λf und $f + g$ wieder lineare Abbildungen von V nach W , und $\text{Hom}(V, W)$ ist mit diesen Operationen ein Vektorraum über \mathbb{K} .

Der Nullvektor in $\text{Hom}(V, W)$ und der inverse Vektor zu einer Abbildung f in $\text{Hom}(V, W)$ sind die Abbildungen

$$0: \begin{cases} V & \rightarrow W \\ v & \mapsto 0 \end{cases}, \quad (-f): \begin{cases} V & \rightarrow W \\ v & \mapsto -f(v) \end{cases}.$$

BEWEIS. Wir zeigen zunächst, dass Linearkombinationen von Homomorphismen wieder linear sind. Dazu seien $f, g \in \text{Hom}(V, W)$, $x, y \in V$, $\lambda, \alpha, \beta \in \mathbb{K}$. Es ist

$$\begin{aligned} &[\lambda f](\alpha x + \beta y) = \lambda f(\alpha x + \beta y) \quad \text{Definition von } \lambda f \\ &= \lambda(\alpha f(x) + \beta f(y)) \quad \text{weil } f \text{ linear ist} \\ &= \alpha(\lambda f(x)) + \beta(\lambda f(y)) \quad \text{Rechenregeln in } \mathbb{K} \\ &= \alpha[\lambda f](x) + \beta[\lambda f](y) \quad \text{Definition von } \lambda f. \end{aligned}$$

$$\begin{aligned} &[f + g](\alpha x + \beta y) = f(\alpha x + \beta y) + g(\alpha x + \beta y) \quad \text{Definition von } f + g \\ &= \lambda(\alpha f(x) + \beta f(y)) + (\alpha g(x) + \beta g(y)) \quad \text{weil } f, g \text{ linear sind} \\ &= \alpha(f(x) + g(x)) + \beta(f(y) + g(y)) \quad \text{Rechenregeln in } \mathbb{K} \\ &= \alpha[f + g](x) + \beta[f + g](y) \quad \text{Definition von } f + g. \end{aligned}$$

Als Beispiel für den Beweis der Rechenregeln beweisen wir eines der Distributivgesetze, nämlich $\alpha(f + g) = \alpha f + \alpha g$. Wir müssen also zeigen, dass für alle $v \in V$ gilt:

$$[\alpha(f + g)](v) = [\alpha f + \alpha g](v).$$

Es ist

$$\begin{aligned} &[\alpha(f + g)](v) = \alpha[f + g](v) \quad \text{Definition von } \alpha h \\ &= \alpha(f(v) + g(v)) \quad \text{Definition von } f + g \\ &= \alpha f(v) + \alpha g(v) \quad \text{Rechenregeln in } \mathbb{K} \\ &= [\alpha f](v) + [\alpha g](v) \quad \text{Definition von } \alpha h \\ &= [\alpha f + \alpha g](v) \quad \text{Definition von } f + g \end{aligned}$$

Die anderen Rechenregeln gehen ebenso zu zeigen¹. □

BEHAUPTUNG 1.3.3. Seien U, V, W Vektorräume über demselben Körper \mathbb{K} , und $f: U \rightarrow V$ und $g: V \rightarrow W$ lineare Abbildungen. Dann ist die Hintereinanderausführung $g \circ f$ ebenfalls linear.

BEWEIS. Seien $x, y \in U$, $\alpha, \beta \in \mathbb{K}$. Dann ist

$$\begin{aligned} &[g \circ f](\alpha x + \beta y) = g(f(\alpha x + \beta y)) = g(\alpha f(x) + \beta f(y)) \\ &= \alpha g(f(x)) + \beta g(f(y)) = \alpha[g \circ f](x) + \beta[g \circ f](y). \end{aligned}$$

□

¹Siehe auch den Beweis zur Behauptung 2.1.2

Wie im Fall des \mathbb{K}^n besteht auch in allgemeinen Vektorräumen ein enges Wechselspiel zwischen Unterräumen und linearen Abbildungen:

BEHAUPTUNG 1.3.4. *Seien V und W Vektorräume über demselben Körper \mathbb{K} , und sei $f : V \rightarrow W$ eine lineare Abbildung. Sei A ein Unterraum von V und B ein Unterraum von W .*

- (1) $f(0) = 0$. (Beachten Sie: Links steht der Nullvektor von V , rechts steht der Nullvektor von W .)
- (2) $f(A)$ ist ein Unterraum von W .
- (3) $f^{-1}(B)$ ist ein Unterraum von V .

BEWEIS. Zur Deutlichkeit unterscheiden wir hier den Nullvektor Θ_V in V , den Nullvektor Θ_W in W und den Skalar 0 .

(1)

$$f(\Theta_V) = f(0\Theta_V) = 0f(\Theta_V) = \Theta_W.$$

(2) Weil $f(\Theta_V) = \Theta_W$ und der Unterraum A den Nullvektor Θ_V enthält, ist $\Theta_W \in f(A)$ und damit ist $f(A) \neq \emptyset$.

Seien $p, q \in f(A)$ und $\lambda, \mu \in \mathbb{K}$. Es gibt also $x, y \in A$, so dass $p = f(x)$ und $q = f(y)$. Dann ist

$$\lambda p + \mu q = \lambda f(x) + \mu f(y) = f(\lambda x + \mu y).$$

Weil A ein Unterraum von V ist, ist $\lambda x + \mu y \in A$, und daher ist

$$\lambda p + \mu q = f(\lambda x + \mu y) \in f(A).$$

(3) Weil $\Theta_W \in B$ und $f(\Theta_V) = \Theta_W$, ist $\Theta_V \in f^{-1}(B)$.

Seien $x, y \in f^{-1}(B)$ und $\lambda, \mu \in \mathbb{K}$. Es ist $f(x) \in B$ und $f(y) \in B$, und weil B ein Unterraum ist, ist auch $\lambda f(x) + \mu f(y) \in B$. Nun ist

$$f(\lambda x + \mu y) = \lambda f(x) + \mu f(y) \in B,$$

also ist $\lambda x + \mu y \in f^{-1}(B)$.

□

DEFINITION 1.3.5 (Kern und Bildraum). Seien V und W Vektorräume über demselben Körper \mathbb{K} , und sei $f : V \rightarrow W$ eine lineare Abbildung. Wir definieren

(1) den Kern von f

$$\ker(f) = \{x \in V \mid f(x) = 0\} = f^{-1}(\{0\}),$$

(2) den Bildraum (Range) von f

$$\operatorname{rg}(f) = \{f(x) \mid x \in V\} = f(V).$$

BEHAUPTUNG 1.3.6. *Seien V und W Vektorräume über demselben Körper \mathbb{K} , und sei $f : V \rightarrow W$ eine lineare Abbildung.*

- (1) *Der Kern von f ist ein Unterraum von V . Die Abbildung f ist genau dann injektiv, wenn $\ker(f) = \{0\}$.*
- (2) *Der Bildraum von f ist ein Unterraum von W . Die Abbildung f ist genau dann surjektiv, wenn $\operatorname{rg}(f) = W$.*

BEWEIS. Wegen Behauptung 1.3.4 sind $f^{-1}(\{0\})$ und $f(V)$ Unterräume von V bzw. W . Die Definition der Surjektivität besagt, dass f genau dann surjektiv ist, wenn $f(V) = W$. Zu zeigen bleibt die Äquivalenz

$$f \text{ ist injektiv} \Leftrightarrow \ker(f) = \{0\}.$$

Sei zunächst f injektiv. Weil $f(0) = 0$, ist $0 \in \ker(f)$. Wenn $x \in \ker(f)$ für irgendein $x \in V$, so gilt $f(x) = 0 = f(0)$, und wegen der Injektivität ist $x = 0$. Also besteht der Kern von f genau aus dem Nullvektor.

Sei jetzt $\ker(f) = \{0\}$, und sei $f(x) = f(y)$. Dann ist

$$f(x - y) = f(x - y) = 0$$

und daher

$$x - y \in \ker(f) = \{0\}.$$

Also ist $x = y$. □

1.4. Isomorphismen.

DEFINITION 1.4.1. Seien V und W Vektorräume über demselben Körper \mathbb{K} .

- (1) Eine bijektive lineare Abbildung $f : V \rightarrow W$ heißt ein (Vektorraum-) Isomorphismus.
- (2) Wenn es einen Isomorphismus $f : V \rightarrow W$ gibt, dann heißen die Vektorräume V und W isomorph. Wir schreiben dann $V \simeq W$.

BEHAUPTUNG 1.4.2. Seien V und W Vektorräume über demselben Körper \mathbb{K} und $f : V \rightarrow W$ ein Isomorphismus.

Dann ist auch die Umkehrabbildung $f^{-1} : W \rightarrow V$ ein Isomorphismus.

BEWEIS. Es ist klar, dass die Umkehrabbildung einer bijektiven Abbildung wiederum bijektiv ist. Zu zeigen bleibt, dass f^{-1} linear ist.

Seien $x, y \in W$ und $\alpha, \beta \in \mathbb{K}$. Zu zeigen ist, dass

$$f^{-1}(\alpha x + \beta y) = \alpha f^{-1}(x) + \beta f^{-1}(y).$$

Wegen der Linearität von f gilt

$$f(\alpha f^{-1}(x) + \beta f^{-1}(y)) = \alpha f(f^{-1}(x)) + \beta f(f^{-1}(y)) = \alpha x + \beta y,$$

und weil f^{-1} die Umkehrabbildung von f ist, heißt die obige Zeile nichts anderes als

$$\alpha f^{-1}(x) + \beta f^{-1}(y) = f^{-1}(\alpha x + \beta y).$$

□

BEHAUPTUNG 1.4.3. Seien U, V, W Vektorräume über demselben Körper \mathbb{K} . Dann gilt:

- (1) U ist isomorph zu sich selbst.
- (2) Ist U isomorph zu V , so ist V isomorph zu U .
- (3) Ist U isomorph zu V und V isomorph zu W , dann ist U isomorph zu W .

BEWEIS.

- (1) Die Identität id_U ist eine lineare bijektive Abbildung $U \rightarrow U$.
- (2) Ist $f : U \rightarrow V$ ein Isomorphismus, so ist wegen Behauptung 1.4.2 auch $f^{-1} : V \rightarrow U$ ein Isomorphismus.
- (3) Seien $f : U \rightarrow V$ und $g : V \rightarrow W$ Isomorphismen. Die Hintereinanderausführung $g \circ f$ der beiden bijektiven Abbildungen g und f ist wiederum bijektiv, und weil g und f linear sind, ist wegen Behauptung 1.3.3 auch $g \circ f$ linear. Also ist $[g \circ f] : U \rightarrow W$ ein Isomorphismus.

□

BEHAUPTUNG 1.4.4. Wenn zwei Vektorräume V und W isomorph sind, dann gilt:

Jede Aussage, die sich mit Hilfe von Linearkombinationen ausdrücken lässt, und im Vektorraum V wahr ist, ist auch im Vektorraum W wahr.

Wenn wir also die Eigenschaften (im Kontext der linearen Algebra) eines Vektorraumes V kennen, kennen wir auch die Eigenschaften aller zu V isomorphen Vektorräume.

BEISPIEL 1.4.5. Seien V und W isomorphe Vektorräume über einem Körper \mathbb{K} .

Wenn V ein Erzeugendensystem aus n Vektoren $\{v_1, \dots, v_n\}$ hat, dann hat auch W ein Erzeugendensystem aus n Vektoren.

BEWEIS. Es sei nämlich $f : V \rightarrow W$ ein Isomorphismus. Wir zeigen, dass $\{f(v_1), \dots, f(v_n)\}$ ein Erzeugendensystem von W ist. Ist $x \in W$, so ist $f^{-1}(x) \in V$, und weil $\{v_1, \dots, v_n\}$ ein Erzeugendensystem von V ist, gibt es Skalare $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ so, dass

$$f^{-1}(x) = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

Wegen der Linearität von f gilt

$$x = f(f^{-1}(x)) = f(\lambda_1 v_1 + \dots + \lambda_n v_n) = \lambda_1 f(v_1) + \dots + \lambda_n f(v_n).$$

Also besteht ganz W aus Linearkombinationen von $f(v_1), \dots, f(v_n)$. \square

BEISPIEL 1.4.6. Mit der gängigen Addition und Multiplikation ist \mathbb{C} ein Vektorraum über \mathbb{R} . Es gilt: Als Vektorräume sind \mathbb{C} und \mathbb{R}^2 isomorph.

BEWEIS. Wir definieren

$$f : \begin{cases} \mathbb{C} & \rightarrow \mathbb{R}^2 \\ z & \mapsto \begin{pmatrix} \operatorname{Re}(z) \\ \operatorname{Im}(z) \end{pmatrix} \end{cases}.$$

Man kann leicht zeigen, dass f ein Isomorphismus ist. \square

BEISPIEL 1.4.7. Die Räume $\operatorname{Hom}(\mathbb{K}^n, \mathbb{K}^m)$ und $\mathbb{K}^{m \times n}$ sind isomorph.

BEWEIS. Wir definieren eine Abbildung $M : \operatorname{Hom}(\mathbb{K}^n, \mathbb{K}^m) \rightarrow \mathbb{K}^{m \times n}$ folgendermaßen: $M(f)$ ist jene Matrix A , für die gilt

$$(\forall v \in \mathbb{K}^n) f(v) = Av.$$

Wir wissen aus Satz 5.3.5, dass es genau eine solche Matrix gibt, und aus Satz 5.3.3, dass zu jeder Matrix genau eine lineare Abbildung gehört. Daher ist durch die obige Definition eine bijektive Abbildung gegeben.

Wir müssen noch zeigen, dass M linear ist. Seien $f, g \in \operatorname{Hom}(\mathbb{K}^n, \mathbb{K}^m)$, $A = M(f)$, $B = M(g)$ die dazu gehörigen Matrizen, und $\alpha, \beta \in \mathbb{K}$. Wir müssen zeigen

$$M(\alpha f + \beta g) = \alpha M(f) + \beta M(g),$$

das heißt, dass die Matrix $\alpha A + \beta B$ genau die Matrix zur linearen Abbildung $\alpha f + \beta g$ ist. Es ist für jedes $v \in \mathbb{K}^n$

$$\begin{aligned} & [\alpha f + \beta g](v) = \alpha f(v) + \beta g(v) \quad \text{Definition der Linearkombinationen in } \operatorname{Hom}(V, W) \\ & = \alpha Av + \beta Bv \quad A, B \text{ sind die Matrizen von } f, g \\ & = (\alpha A + \beta B)(v) \quad \text{Rechenregeln für das Matrixprodukt} \end{aligned}$$

\square

Den Begriff des Homomorphismus und des Isomorphismus gibt es nicht nur für Vektorräume, sondern für alle algebraischen Strukturen. Ein Homomorphismus ist, salopp gesprochen, eine Abbildung, die mit den Rechenoperationen verträglich ist. Es ist gleichgültig, ob zuerst die Rechenoperation ausgeführt wird und dann erst die Abbildung, oder umgekehrt. Beim Vektorraumhomomorphismus ist die betreffende Rechenoperation die Linearkombination.

Zum Beispiel ist die folgende Abbildung ein Gruppenisomorphismus von der Gruppe $((0, \infty), \cdot)$ auf die Gruppe $(\mathbb{R}, +)$:

$$\begin{cases} (0, \infty) & \rightarrow \mathbb{R} \\ x & \mapsto \log_{10}(x) \end{cases}.$$

Es ist nämlich

$$\log_{10}(x \cdot y) = \log_{10}(x) + \log_{10}(y).$$

Als es noch keine Taschenrechner gab, wurde diese Tatsache als Rechenhilfe ausgenutzt. Man hatte Tabellen, in denen man zu jeder fünfstelligen Zahl den dekadischen Logarithmus finden konnte. Sollte man etwa $5213 \cdot 214$ multiplizieren, addierte man zunächst die Logarithmen

$$\begin{array}{rcl} \log_{10}(5213) & \approx & 3.71709 \\ \log_{10}(214) & \approx & 2.33041 \\ \text{Summe} & & 6.04750 \end{array}$$

Weil die Summe mit $6. \dots$ beginnt, ist das Ergebnis eine Zahl zwischen einer und zehn Millionen. Aus der Logarithmentafel sieht man, dass

$$0.04750 \approx \log_{10}(1.11558).$$

Auf die Rechengenauigkeit der Tabellen genau ist also

$$5213 \cdot 214 \approx 1\,115\,580 \quad (\text{exakt ist } 1\,115\,582).$$

Weil Addieren viel bequemer ist als Multiplizieren (und Multiplizieren beziehungsweise Dividieren sehr viel bequemer sind als Potenzieren und Wurzeln ziehen), war das eine nennenswerte Erleichterung. Durch den Isomorphismus konnte man sich für die Rechenoperationen jene Gruppe aussuchen, in der die Rechnungen bequemer durchgeführt werden können. Mit Hilfe des Isomorphismus kehrte man dann bei Bedarf zur anderen Gruppe zurück.

Wir sehen, wie viel von den bekannten Eigenschaften des \mathbb{K}^n sich schon im ganz abstrakten Rahmen des Vektorraumes beweisen lassen. Ebenso wichtig ist, zu verstehen, dass der Raum \mathbb{K}^n Eigenschaften hat, die nur durch die endliche Dimension folgen:

WARNUNG 1.4.8. *Ist $f : \mathbb{K}^n \rightarrow \mathbb{K}^n$ eine lineare Abbildung, dann gilt bekanntlich f ist injektiv $\Leftrightarrow f$ ist surjektiv.*

Das folgt, weil sich Injektivität, Surjektivität und Bijektivität aus dem Rang der $n \times n$ -Matrix von f ablesen lassen.

Es gibt aber Vektorräume V und lineare Abbildungen $f : V \rightarrow V$, welche

- *injektiv, aber nicht surjektiv oder*
- *surjektiv, aber nicht injektiv sind.*

Es gibt viele und wichtige Beispiele dazu in Funktionenräumen, z.B. Beispiel 2.2.2.

1.5. Normen und Skalarprodukte in allgemeinen Vektorräumen.

DEFINITION 1.5.1. Sei \mathbb{K} einer der Körper \mathbb{R} oder \mathbb{C} und V ein Vektorraum über \mathbb{K} . Eine Abbildung

$$\begin{cases} V & \rightarrow \mathbb{R} \\ x & \mapsto \|x\| \end{cases}$$

heißt eine Norm auf V , wenn sie folgende Bedingungen erfüllt:

- (1) Für alle $x \in V$ ist $\|x\| \geq 0$, und $\|x\| = 0$ genau dann, wenn $x = 0$.
- (2) Für alle $x \in V$ und alle $\lambda \in \mathbb{K}$ gilt $\|\lambda x\| = |\lambda| \|x\|$.
- (3) Für alle $v, w \in V$ ist $\|v + w\| \leq \|v\| + \|w\|$.

Wenn auf einem Vektorraum V eine Norm $\|\cdot\|$ gegeben ist, dann heißt das Paar $(V, \|\cdot\|)$ ein normierter Vektorraum.

Die Bedeutung der Norm auf einem Vektorraum liegt darin, dass man mit Hilfe der Norm Grenzwerte definieren kann, und damit auf Vektorräumen Analysis betreiben kann. Die Theorie der Analysis auf allgemeinen Vektorräumen heißt Funktionalanalysis.

DEFINITION 1.5.2 (Grenzwert in normierten Vektorräumen). Es sei $(V, \|\cdot\|)$ ein normierter Vektorraum.

- (1) Ist $w \in V$ und $\varepsilon > 0$, so definieren wir die offene Kugel mit Radius ε und Mittelpunkt w :

$$K(w, \varepsilon) := \{u \in V \mid \|u - w\| < \varepsilon\}.$$

- (2) Sei $(v_n)_{n=1}^\infty$ eine Folge von Vektoren in V , und $w \in V$. Man sagt, die Folge (v_n) konvergiert gegen w , oder

$$\lim_{n \rightarrow \infty} v_n = w,$$

genau dann, wenn gilt:

$$(\forall \varepsilon > 0) (\exists n_0 \in \mathbb{N}) (\forall n \geq n_0) \|v_n - w\| \leq \varepsilon.$$

Offene Kugeln spielen in der Analysis im normierten Raum die Rolle, die offene Intervalle in der reellen Analysis spielen. Eine Folge (v_n) von Vektoren konvergiert genau dann gegen einen Vektor w , wenn jede offene Kugel $K(w, \varepsilon)$ die Folge (v_n) "einfängt".

Beispiele werden wir im Abschnitt über Funktionenräume sehen.

BEISPIEL 1.5.3 (Matrixnorm). Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Auf \mathbb{K}^m und \mathbb{K}^n betrachten wir Normen $\|\cdot\|_{\mathbb{K}^m}$ bzw. $\|\cdot\|_{\mathbb{K}^n}$.

- (1) Dann liefert die folgende Definition eine Norm auf dem Raum $\mathbb{K}^{m \times n}$

$$\|A\|_{\mathbb{K}^{m \times n}} := \sup_{x \in \mathbb{K}^n, \|x\|_{\mathbb{K}^n} = 1} \|Ax\|_{\mathbb{K}^m}.$$

- (2) Für alle $x \in \mathbb{K}^n$ gilt dann:

$$\|f(x)\|_{\mathbb{K}^m} \leq \|f\|_{\mathbb{K}^{m \times n}} \|x\|_{\mathbb{K}^n}.$$

BEWEIS. Man braucht Hilfsmittel aus der Analysis, nämlich den Begriff der Kompaktheit, um zu zeigen, dass das Supremum

$$\sup_{x \in \mathbb{K}^n, \|x\|_{\mathbb{K}^n} = 1} \|Ax\|_{\mathbb{K}^m}$$

wirklich endlich ist. Dazu haben wir im Rahmen dieser Vorlesung leider nicht die Zeit.

Gehen wir nun davon aus, dass das Supremum endlich ist. Wir zeigen zuerst Punkt (2). Das ist trivial für $x = 0$. Es sei $x \in \mathbb{K}^n \setminus \{0\}$. Wir definieren

$$\tilde{x} := \frac{1}{\|x\|_{\mathbb{K}^n}} x \quad \text{sodass} \quad \|\tilde{x}\|_{\mathbb{K}^n} = 1.$$

Daher ist

$$\|A\tilde{x}\|_{\mathbb{K}^m} \leq \|A\|_{\mathbb{K}^{m \times n}}$$

und

$$\|Ax\|_{\mathbb{K}^m} = \|A(\|x\|_{\mathbb{K}^n} \tilde{x})\|_{\mathbb{K}^m} = \|x\|_{\mathbb{K}^n} \|A\tilde{x}\|_{\mathbb{K}^m} \leq \|x\|_{\mathbb{K}^n} \|A\|_{\mathbb{K}^{m \times n}}.$$

Nun kann man die Normeigenschaften leicht überprüfen:

- (1) Weil $\|Ax\|_{\mathbb{K}^m} \geq 0$ für alle x , gilt auch für das Supremum $\|A\|_{\mathbb{K}^{m \times n}} \geq 0$. Ist $\|A\|_{\mathbb{K}^{m \times n}} = 0$, so folgt aus Punkt (2) für alle $x \in \mathbb{K}^n$, dass $\|Ax\|_{\mathbb{K}^m} = 0$, und daher $Ax = 0$. Weil das für alle $x \in \mathbb{K}^n$ gilt, ist dann A die Nullmatrix.

(2) Es gilt

$$\begin{aligned} \|\lambda A\|_{\mathbb{K}^{m \times n}} &= \sup_{x \in \mathbb{K}^n, \|x\|_{\mathbb{K}^n}=1} \|(\lambda A)x\|_{\mathbb{K}^m} = \sup_{x \in \mathbb{K}^n, \|x\|_{\mathbb{K}^n}=1} |\lambda| \|Ax\|_{\mathbb{K}^m} \\ &= |\lambda| \sup_{x \in \mathbb{K}^n, \|x\|_{\mathbb{K}^n}=1} \|Ax\|_{\mathbb{K}^m} = |\lambda| \|A\|_{\mathbb{K}^{m \times n}}. \end{aligned}$$

(3) Sei $x \in \mathbb{K}^n$ mit $\|x\|_{\mathbb{K}^n} = 1$. Dann gilt

$$\begin{aligned} \|(A+B)x\|_{\mathbb{K}^m} &= \|Ax + Bx\|_{\mathbb{K}^m} \leq \|Ax\|_{\mathbb{K}^m} + \|Bx\|_{\mathbb{K}^m} \\ &\leq \|A\|_{\mathbb{K}^{m \times n}} + \|B\|_{\mathbb{K}^{m \times n}}. \end{aligned}$$

Dass gilt für alle x mit $\|x\|_{\mathbb{K}^n} = 1$, daher auch für das Supremum:

$$\|A+B\|_{\mathbb{K}^{m \times n}} \leq \|A\|_{\mathbb{K}^{m \times n}} + \|B\|_{\mathbb{K}^{m \times n}}.$$

□

BEISPIEL 1.5.4. In \mathbb{R}^n betrachten wir als Norm $\|\cdot\|$ die euklidische Norm, und die dazu passende Matrixnorm $\|\cdot\|_{\mathbb{R}^{n \times n}}$.

Sei $A \in \mathbb{R}^{n \times n}$ regulär und $(b_i)_{i=1}^\infty$ eine Folge in \mathbb{R}^n , welche gegen einen Vektor b konvergiert. Dann konvergieren die (bekanntlich eindeutigen) Lösungen x_k des Gleichungssystems $Ax_k = b_k$ gegen die Lösung x des Gleichungssystems $Ax = b$.

BEWEIS. Es gilt

$$\|x_n - x\| = \|A^{-1}b_n - A^{-1}b\| = \|A^{-1}(b_n - b)\| \leq \|A^{-1}\|_{\mathbb{R}^{n \times n}} \|b_n - b\|.$$

Sei $\varepsilon > 0$. Weil (b_n) gegen b konvergiert, gibt es ein $n_0 \in \mathbb{N}$, sodass für alle $n \geq n_0$ gilt

$$\|b_n - b\| \leq \frac{\varepsilon}{\|A^{-1}\|_{\mathbb{R}^{n \times n}}}.$$

Dann gilt auch für alle $n \geq n_0$

$$\|x_n - x\| \leq \|A^{-1}\|_{\mathbb{R}^{n \times n}} \|b_n - b\| \leq \|A^{-1}\|_{\mathbb{R}^{n \times n}} \frac{\varepsilon}{\|A^{-1}\|_{\mathbb{R}^{n \times n}}} = \varepsilon.$$

Also konvergiert die Folge (x_n) gegen x .

□

WARNUNG 1.5.5. Für Homomorphismen $f \in \text{Hom}(V, W)$ zwischen allgemeinen Vektorräumen V, W über \mathbb{R} oder \mathbb{C} mit Normen $\|\cdot\|_V$ und $\|\cdot\|_W$ kann es sein, dass gilt

$$\sup_{v \in V, \|v\|_V=1} \|f(v)\|_W = \infty.$$

Man kann also normalerweise nicht auf $\text{Hom}(V, W)$ eine Norm nach Art der Matrixnorm einführen.

Man definiert dann den Raum der sogenannten beschränkten linearen Abbildungen von V nach W :

$$L(V, W) := \{f \in \text{Hom}(V, W) \mid \sup_{v \in V, \|v\|_V=1} \|f(v)\|_W < \infty\}.$$

Auf diesem Raum dient das obige Supremum wieder als Norm. Mehr darüber hören Sie in einer Vorlesung aus Funktionalanalysis.

Auch zur Definition eines Skalarproduktes auf einem allgemeinen reellen oder komplexen Vektorraum kann man die Definition eines Skalarproduktes auf \mathbb{K}^n kopieren:

DEFINITION 1.5.6 (Skalarprodukt im allgemeinen Vektorraum). Es sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und V ein Vektorraum über \mathbb{K} . Gegeben sei eine Abbildung

$$\begin{cases} V \times V & \rightarrow \mathbb{K} \\ (a, b) & \mapsto \langle a, b \rangle \end{cases}.$$

Die Abbildung $\langle \cdot, \cdot \rangle$ heißt ein Skalarprodukt oder inneres Produkt auf V wenn gilt:

- (1) $\langle \cdot, \cdot \rangle$ ist positiv definit, das heißt für alle $a \in V$

$$\langle a, a \rangle \geq 0 \quad \text{und} \quad \langle a, a \rangle = 0 \Rightarrow a = 0.$$

- (2) $\langle \cdot, \cdot \rangle$ ist hermitesch, das heißt für alle $a, b \in V$

$$\langle a, b \rangle = \overline{\langle b, a \rangle}.$$

(Ist $\mathbb{K} = \mathbb{R}$, dann ist das äquivalent zur Symmetrie: $\langle a, b \rangle = \langle b, a \rangle$.)

- (3) $\langle \cdot, \cdot \rangle$ ist sesquilinear, das heißt für alle $a, b, c \in V$ und alle $\lambda, \mu \in \mathbb{K}$

$$\begin{aligned} \langle c, \lambda a + \mu b \rangle &= \lambda \langle c, a \rangle + \mu \langle c, b \rangle, \\ \langle \lambda a + \mu b, c \rangle &= \overline{\lambda} \langle a, c \rangle + \overline{\mu} \langle b, c \rangle. \end{aligned}$$

Wenn $\mathbb{K} = \mathbb{R}$, dann fallen die Konjugationen weg, und man sagt, die Abbildung $\langle \cdot, \cdot \rangle$ ist bilinear.

Ein reeller Vektorraum mit Skalarprodukt heißt auch ein euklidischer Raum, ein komplexer Vektorraum mit Skalarprodukt heißt ein unitärer Raum.

DEFINITION 1.5.7. Sei V ein reeller oder komplexer Vektorraum mit innerem Produkt V . Wir definieren die von $\langle \cdot, \cdot \rangle$ induzierte Norm auf V durch

$$\|v\| := \sqrt{\langle v, v \rangle}.$$

Die wichtigste Zutat zum Beweis, dass die induzierte Norm wirklich eine Norm ist, ist die Cauchy-Schwarzsche Ungleichung. Obwohl man sie und ihren Beweis wörtlich von Satz 1.2.6 kopieren kann, wiederholen wir sie hier.

SATZ 1.5.8 (Cauchy-Schwarzsche Ungleichung). Sei V ein Vektorraum über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ mit einem inneren Produkt $\langle \cdot, \cdot \rangle$. Seien $x, y \in V$. Dann gilt

$$|\langle x, y \rangle| \leq \|x\| \|y\|.$$

Gleichheit gilt genau dann, wenn x, y Vielfache eines gemeinsamen Vektors z sind.

BEWEIS. Fall 1: $x = \alpha z, y = \beta z$. Dann ist

$$\begin{aligned} \|x\| &= \sqrt{\langle \alpha z, \alpha z \rangle} = \sqrt{\overline{\alpha} \alpha \langle z z \rangle} = \sqrt{|\alpha|^2 \langle z z \rangle} = |\alpha| \sqrt{\langle z, z \rangle}. \\ \|y\| &= |\beta| \sqrt{\langle z, z \rangle} \quad \text{ebenso.} \\ |\langle x, y \rangle| &= |\langle \alpha z, \beta z \rangle| = |\overline{\alpha} \beta \langle z, z \rangle| = |\alpha| |\beta| \langle z, z \rangle = \|x\| \|y\|. \end{aligned}$$

Fall 2: $x + \varepsilon y \neq 0$ für alle $\varepsilon \in \mathbb{K}$. Wegen der positiven Definitheit von $\langle \cdot, \cdot \rangle$ gilt für alle $\varepsilon \in \mathbb{K}$

$$\begin{aligned} 0 &< \langle x + \varepsilon y, x + \varepsilon y \rangle \\ &= \langle x \rangle + \varepsilon \langle x, y \rangle + \overline{\varepsilon} \langle y, x \rangle + \overline{\varepsilon} \varepsilon \langle y, y \rangle. \end{aligned}$$

Wir wählen jetzt ein bestimmtes ε so, dass die drei letzten Terme der letzten Zeile bis auf das Vorzeichen gleich sind, und sich zwei davon wegheben, also

$$\varepsilon := -\frac{\langle y, x \rangle}{\langle y, y \rangle}.$$

Wir erhalten

$$\begin{aligned} 0 &< \langle x, x \rangle - \frac{\langle y, x \rangle}{\langle y, y \rangle} \langle x, y \rangle - \frac{\langle x, y \rangle}{\langle y, y \rangle} \langle y, x \rangle + \frac{\langle y, x \rangle \langle x, y \rangle}{\langle y, y \rangle^2} \langle y, y \rangle \\ &= \langle x, x \rangle - \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle}. \end{aligned}$$

Umordnen der Ungleichung liefert

$$|\langle x, y \rangle|^2 < \langle x, x \rangle \langle y, y \rangle.$$

□

BEHAUPTUNG 1.5.9. Sei V ein Vektorraum über \mathbb{R} oder \mathbb{C} mit innerem Produkt $\{\cdot, \cdot\}$, und sei $\|\cdot\|$ die vom Skalarprodukt induzierte Norm. Dann ist $\|\cdot\|$ eine Norm auf V .

BEWEIS.

(1) Für alle $x \in V$ ist $\{x, x\} \geq 0$, daher existiert eine nichtnegative Wurzel $\sqrt{\langle x, x \rangle} \geq 0$. Wegen der positiven Definitheit des Skalarproduktes ist $\langle x, x \rangle = 0$ nur wenn $x = 0$.

(2)

$$\|\alpha x\| = \sqrt{\langle \alpha x, \alpha x \rangle} = \sqrt{|\alpha|^2 \langle x, x \rangle} = |\alpha| \sqrt{\langle x, x \rangle}.$$

(3)

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\ &= \|x\|^2 + 2 \operatorname{Re}(\langle x, y \rangle) + \|y\|^2 \leq \|x\|^2 + 2|\langle x, y \rangle| + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \end{aligned}$$

□

Wie in \mathbb{K}^n definieren wir nun den Begriff der Orthogonalität und des orthogonalen Komplementes:

DEFINITION 1.5.10. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und V ein Vektorraum über \mathbb{K} mit innerem Produkt $\{\cdot, \cdot\}$. Seien $u, v \in V$, S eine Teilmenge von V und U ein Unterraum von V . Wir definieren:

- (1) u und v heißen aufeinander orthogonal, wenn gilt $\langle u, v \rangle = 0$.
- (2) u heißt orthogonal auf S , wenn gilt $(\forall s \in S) \langle u, s \rangle = 0$.
- (3) Das orthogonale Komplement von U ist

$$U^\perp := \{v \in V \mid (\forall u \in U) \langle u, v \rangle = 0\}.$$

Viele Eigenschaften des orthogonalen Komplementes lassen sich im allgemeinen Vektorraum genauso beweisen wie in \mathbb{K}^n — siehe Behauptung 2.1.4 Punkte (a,b,c):

BEHAUPTUNG 1.5.11. Sei V ein Vektorraum über $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ mit einem inneren Produkt $\langle \cdot, \cdot \rangle$. Sei $x \in V$, $S \subseteq V$ und U ein Unterraum von V . Es gilt:

- (1) $u \in (\mathcal{L}(S))^\perp \iff (\forall s \in S) \langle u, s \rangle = 0$.
- (2) U^\perp ist ein Unterraum von V .
- (3) Ist U ein Unterraum von V , dann ist $U \cap U^\perp = \{0\}$.

WARNUNG 1.5.12. Für einen Unterraum $U \subseteq \mathbb{K}^n$ gilt $(U^\perp)^\perp = U$. In einem allgemeinen euklidischen oder unitären Vektorraum gilt zwar immer $U \subseteq (U^\perp)^\perp$, aber es kann sein, dass $(U^\perp)^\perp$ eine echte Obermenge von U ist.

BEWEIS. Sei $u \in U$ und $w \in U^\perp$. Dann ist natürlich $\langle u, w \rangle = 0$, weil w im Komplement von U und u in U liegt. Also ist $u \in (U^\perp)^\perp$.

Ein Beispiel, dass die Umkehrung nicht gehen muss, folgt in Beispiel 2.3.4. □

BEHAUPTUNG 1.5.13. Sei V ein euklidischer oder unitärer Vektorraum, U ein Unterraum von V , und $x \in V$. Dann gilt:

- (1) x lässt sich höchstens auf eine Weise in $x = u + w$ mit $u \in U$ und $w \in U^\perp$ zerlegen.
 (2) Sei $u \in U$. Es gilt $x - u \in U^\perp$ genau wenn

$$(\forall y \in U) \|x - u\| \leq \|x - y\|.$$

BEWEIS.

- (1) Die Eindeutigkeit der Zerlegung folgt wie im endlich-dimensionalen Fall: Ist $x = u_1 + w_1 = u_2 + w_2$ mit $u_i \in U$ und $w_i \in U^\perp$, dann ist

$$u_1 - u_2 = w_2 - w_1 \in U \cap U^\perp = \{0\}.$$

Also ist $u_1 = u_2$ und $w_1 = w_2$.

- (2) Sei $x = u + w$ mit $u \in U$ und $w \in U^\perp$. Dann gilt $\langle x - u, u - y \rangle = 0$ für alle $y \in U$, weil $x - u = w \in U^\perp$.

$$\begin{aligned} \|x - y\|^2 &= \|(x - u) + (u - y)\|^2 \\ &= \langle (x - u) + (u - y), (x - u) + (u - y) \rangle \\ &= \langle x - u, x - u \rangle + 2 \operatorname{Re}(\langle x - u, u - y \rangle) + \langle u - y, u - y \rangle \\ &= \|x - u\|^2 + 0 + \|u - y\|^2 \geq \|x - u\|^2. \end{aligned}$$

Sei umgekehrt $\|x - u\| \leq \|x - y\|$ für alle $y \in U$. Wir zeigen: $x - u \in U^\perp$. Sei $z \in U$ und $\alpha = \langle z, x - u \rangle$. Für ein zunächst beliebiges $\varepsilon > 0$ ist $u + \varepsilon\alpha z \in U$ und daher

$$\begin{aligned} 0 &\leq \|x - u - \varepsilon\alpha z\|^2 - \|x - u\|^2 \\ &= \|x - u\|^2 - 2 \operatorname{Re}(\langle x - u, \varepsilon\alpha z \rangle) + \|\varepsilon\alpha z\|^2 - \|x - u\|^2 \\ &= -2\varepsilon \operatorname{Re}(\alpha \langle x - u, z \rangle) + \varepsilon^2 |\alpha|^2 \|z\|^2 \\ &= -2\varepsilon |\alpha|^2 + \varepsilon^2 |\alpha|^2 \|z\|^2. \end{aligned}$$

Wir kürzen durch ε und erhalten

$$|\alpha|^2 \leq \varepsilon \|z\|^2 |\alpha|^2 \rightarrow 0 \quad \text{für } \varepsilon \rightarrow 0.$$

Also ist $\alpha = \langle z, x - u \rangle = 0$.

□

WARNUNG 1.5.14. Ist U ein Unterraum des \mathbb{K}^n und $x \in \mathbb{K}^n$, so gibt es immer eine Zerlegung

$$x = u + w \quad \text{mit } u \in U \text{ und } w \in U^\perp.$$

Ist aber U ein Unterraum eines allgemeinen euklidischen oder unitären Vektorraumes und ist $(U^\perp)^\perp \neq U$, dann gibt es ein $x \in V$, welches sich nicht in $u + w$ mit $u \in U$ und $w \in U^\perp$ zerlegen lässt.

BEWEIS. Sei $(x \in (U^\perp)^\perp \setminus U)$. Wenn es eine Zerlegung $x = u + w$ mit $u \in U$ und $w \in U^\perp$ gibt, dann ist auch $u \in (U^\perp)^\perp$, und daher

$$w = x - u \in (U^\perp)^\perp \cap U^\perp = \{0\}$$

und daher $x = u \in U$ im Widerspruch zur Wahl von x . Also kann x keine solche Zerlegung haben. □

DEFINITION 1.5.15. Sei V ein Vektorraum über \mathbb{R} oder \mathbb{C} mit einem inneren Produkt $\langle \cdot, \cdot \rangle$.

Eine nichtleere Teilmenge $A \subseteq V$ heißt ein Orthonormalsystem, wenn gilt:

$$(\forall a, b \in A) \langle a, b \rangle = \begin{cases} 0 & \text{wenn } a \neq b \\ 1 & \text{wenn } a = b \end{cases}.$$

BEHAUPTUNG 1.5.16. Sei (a_1, \dots, a_k) ein endliches Orthogonalsystem in einem Vektorraum V mit innerem Produkt $\langle \cdot, \cdot \rangle$. Sei $U = \mathcal{L}(a_1, \dots, a_k)$ und sei $x \in V$. Dann ist die orthogonale Projektion von x auf U ist

$$\sum_{i=1}^k \langle a_i, x \rangle a_i.$$

BEWEIS. Genau wie Satz 2.2.4. □

In allgemeinen euklidischen oder unitären Vektorräumen kann es unendliche Orthonormalsysteme geben. Sei A ein unendliches Orthogonalsystem. Die Summe

$$\sum_{a \in A} \langle a, x \rangle a$$

ist dann eine unendliche Reihe. Die Norm im Vektorraum eröffnet die Möglichkeit, konvergente Reihen in einem normierten Vektorraum zu definieren. Wenn der Vektorraum V vollständig² ist, kann man beweisen, dass die obige Summe tatsächlich konvergiert. Sie ist dann die orthogonale Projektion von x auf den kleinsten abgeschlossenen Unterraum, der A enthält.

2. Funktionenräume

Wir können im Rahmen dieser Vorlesung nicht einmal cursorisch eine Theorie der Funktionenräume vorstellen. Dieser Abschnitt dient eher dazu, an Hand von Beispielen zu zeigen, welche Objekte und Operationen in die Reichweite der linearen Algebra fallen, und warum wir ein so allgemeines Vektorraumkonzept eingeführt haben.

2.1. Definition von Funktionenräumen.

DEFINITION 2.1.1. [Funktionenräume] Sei $\Omega \neq \emptyset$ eine Menge und \mathbb{K} ein Körper. Mit \mathbb{K}^Ω oder $\mathcal{F}(\Omega, \mathbb{K})$ bezeichnen wir die Menge aller Funktionen von Ω nach \mathbb{K} . Für $f, g \in \mathcal{F}(\Omega, \mathbb{K})$, $\alpha, \beta \in \mathbb{K}$ definieren wir die Funktion $\alpha f + \beta g \in \mathcal{F}(\Omega, \mathbb{K})$ durch

$$[\alpha f + \beta g](x) := \alpha f(x) + \beta g(x).$$

BEHAUPTUNG 2.1.2. Mit den Operationen aus Definition 2.1.1 ist $\mathcal{F}(\Omega, \mathbb{K})$ ein Vektorraum über \mathbb{K} . Der Nullvektor ist die konstante Funktion $f(x) = 0$.

BEWEIS. Wir müssen die Axiome des Vektorraums beweisen. Der Deutlichkeit halber unterscheiden wir hier die Zeichen für die Addition $+$ im Körper und die Addition \oplus von Funktionen, und die Zeichen für die Multiplikation \cdot im Körper, und die Multiplikation \odot eines Skalars mit einer Funktion. Das macht man in der Praxis nicht.

- Assoziativität der Addition: Wir müssen zeigen

$$(f \oplus g) \oplus h = f \oplus (g \oplus h),$$

also: für alle $x \in \Omega$

$$[(f \oplus g) \oplus h](x) = [f \oplus (g \oplus h)](x).$$

²Abgeschlossen und vollständig sind Begriffe aus der Analysis, die sich mit Hilfe der Norm auch auf Vektorräume übertragen lassen.

Es gilt

$$\begin{aligned}
 & [(f \oplus g) \oplus h](x) = (f \oplus g)(x) + h(x) \quad \text{Definition der Summe von Funktionen} \\
 = & (f(x) + g(x)) + h(x) \quad \text{Definition der Summe von Funktionen} \\
 = & f(x) + (g(x) + h(x)) \quad \text{Assoziativität der Körperaddition} \\
 = & f(x) + (g \oplus h)(x) \quad \text{Definition der Summe von Funktionen} \\
 = & [f \oplus (g \oplus h)](x) \quad \text{Definition der Summe von Funktionen.}
 \end{aligned}$$

- Das neutrale Element für die Addition von Funktionen ist die konstante Funktion $\Theta(x) := 0$. Denn für alle $x \in \Omega$ gilt

$$\begin{aligned}
 & [f \oplus \Theta](x) = f(x) + \Theta(x) \quad \text{Definition der Summe von Funktionen} \\
 = & f(x) + 0 \quad \text{Definition von } \Theta \\
 = & f(x) \quad 0 \text{ ist Nullelement des Körpers.}
 \end{aligned}$$

also $f \oplus \Theta = f$.

- Das inverse Element zu einer Funktion f für die Funktionenaddition ist

$$(-f)(x) := -f(x).$$

Denn für alle $x \in \Omega$ gilt

$$\begin{aligned}
 & [f \oplus (-f)](x) = f(x) + (-f)(x) \quad \text{Definition der Summe von Funktionen} \\
 = & f(x) - f(x) \quad \text{Definition von } (-f) \\
 = & 0 \quad \text{Subtraktion im Körper} \\
 = & \Theta(x) \quad \text{Definition der Nullfunktion } \Theta.
 \end{aligned}$$

Daher ist $f \oplus (-f) = \Theta$.

- Die Addition von Funktionen ist kommutativ, denn für alle $x \in \Omega$ gilt

$$\begin{aligned}
 & [f \oplus g](x) = f(x) + g(x) \quad \text{Definition der Summe von Funktionen} \\
 = & g(x) + f(x) \quad \text{Kommutativität von } + \text{ im Körper} \\
 = & [g \oplus f](x) \quad \text{Definition der Summe von Funktionen.}
 \end{aligned}$$

- $\alpha \odot (f \oplus g) = (\alpha \odot f) \oplus (\alpha \odot g)$, denn für alle $x \in \Omega$ gilt:

$$\begin{aligned}
 & [\alpha \odot (f \oplus g)](x) = \alpha \cdot (f \oplus g)(x) \quad \text{Definition von } \odot \\
 = & \alpha(f(x) + g(x)) \quad \text{Definition von } \oplus \\
 = & \alpha f(x) + \alpha g(x) \quad \text{Distributivität im Körper} \\
 = & [\alpha \odot f](x) + [\alpha \odot g](x) \quad \text{Definition von } \odot \\
 = & [(\alpha \odot f) \oplus (\alpha \odot g)](x) \quad \text{Definition von } \oplus.
 \end{aligned}$$

- $(\alpha + \beta) \odot f = (\alpha \odot f) \oplus (\beta \odot f)$, denn für alle $x \in \Omega$ ist

$$\begin{aligned}
 & [(\alpha + \beta) \odot f](x) = (\alpha + \beta)f(x) \quad \text{Definition von } \odot \\
 = & \alpha f(x) + \beta f(x) \quad \text{Distributivität im Körper} \\
 = & [\alpha \odot f](x) + [\beta \odot f](x) \quad \text{Definition von } \odot \\
 = & [(\alpha \odot f) \oplus (\beta \odot f)](x) \quad \text{Definition von } \oplus.
 \end{aligned}$$

- $\alpha \odot (\beta \odot f) = (\alpha\beta) \odot f$, denn für alle $x \in \Omega$ ist

$$\begin{aligned}
 & [\alpha \odot (\beta \odot f)](x) = \alpha[\beta \odot f](x) \quad \text{Definition von } \odot \\
 = & \alpha(\beta f(x)) \quad \text{Definition von } \odot \\
 = & (\alpha\beta)f(x) \quad \text{Assoziativität der Körpermultiplikation} \\
 = & [(\alpha\beta) \odot f](x) \quad \text{Definition von } \odot.
 \end{aligned}$$

- $1 \odot f = f$, denn für alle $x \in \Omega$ ist

$$\begin{aligned} [1 \odot f](x) &= 1 \cdot f(x) && \text{Definition von } \odot \\ &= f(x) && 1 \text{ ist neutral bezüglich der Körpermultiplikation.} \end{aligned}$$

□

Normalerweise betrachtet man nicht den Raum aller Funktionen, sondern von Funktionen mit bestimmten Eigenschaften:

DEFINITION 2.1.3. Es sei \mathbb{K} einer der Körper \mathbb{R} oder \mathbb{C} , es sei J ein (beschränktes oder unbeschränktes) Intervall in \mathbb{R} . Wie in Behauptung 2.1.2 sei $\mathcal{F}(J, \mathbb{K})$ der Vektorraum der Funktionen von J nach \mathbb{K} . Wir definieren folgende Funktionenmengen

- $\mathcal{B}(J, \mathbb{K})$ die Menge der beschränkten³ Funktionen von J nach \mathbb{K} .
- $\mathcal{C}(J, \mathbb{K})$ die Menge der stetigen⁴ Funktionen $J \rightarrow \mathbb{K}$.
- $\mathcal{C}^k(J, \mathbb{K})$ die Menge aller stetigen Funktionen $J \rightarrow \mathbb{K}$, die stetige Ableitungen bis zur Ordnung k besitzen.

BEHAUPTUNG 2.1.4. Die in Definition 2.1.3 genannten Funktionenmengen sind Unterräume von $\mathcal{F}(J, \mathbb{K})$.

BEWEIS. In einer Einführungsveranstaltung zu Analysis lernt man, dass Summen und Produkte beschränkter, stetiger oder differenzierbarer Funktionen wieder beschränkt, stetig oder differenzierbar sind. Weil die konstante Funktion $f(t) = \lambda$ auch beschränkt, stetig, und beliebig oft differenzierbar ist, gilt auch, dass das Produkt eines Skalars λ mit einer beschränkten, stetigen oder differenzierbaren Funktion wieder beschränkt, stetig oder differenzierbar ist. □

2.2. Lineare Abbildungen in Funktionenräumen.

BEISPIEL 2.2.1. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, sei $t_0 \in [0, 1]$, $h \in \mathcal{C}([0, 1], \mathbb{K})$. Folgende Abbildungen sind linear:

$$\begin{aligned} T &: \begin{cases} \mathcal{F}([0, 1], \mathbb{K}) & \rightarrow \mathbb{K}, \\ [T(f)](x) & := f(t_0) \end{cases} \\ I &: \begin{cases} \mathcal{C}([0, 1], \mathbb{K}) & \rightarrow \mathbb{K} \\ I(f) & := \int_0^1 f(t) dt \end{cases} \\ D &: \begin{cases} \mathcal{C}^1([0, 1], \mathbb{K}) & \rightarrow \mathcal{C}([0, 1], \mathbb{K}) \\ [D(f)](t) & := \frac{d}{dt} f(t) \end{cases} \\ C &: \begin{cases} \mathcal{C}([0, 1], \mathbb{K}) & \rightarrow \mathcal{C}([0, 1], \mathbb{K}) \\ [Cf](t) & := \int_0^t h(t-s)f(s) ds \end{cases} \end{aligned}$$

BEWEIS. Alle Beweise folgen aus der Definition der Linearkombination von Funktionen, und den Rechenregeln der Differential- und Integralrechnung. Wir führen als Beispiel den Beweis für die Abbildung C .

Seien also $f, g \in \mathcal{C}([0, 1], \mathbb{K})$ und $\alpha, \beta \in \mathbb{K}$. Wir zeigen:

$$C(\alpha f + \beta g) = \alpha C(f) + \beta C(g).$$

³Eine Funktion $f : J \rightarrow \mathbb{K}$ heißt beschränkt, wenn es eine Konstante M gibt, sodass $(\forall t \in J) |f(t)| \leq M$.

⁴Die Begriffe Stetigkeit und Differenzierbarkeit lernen Sie in einer Einführungsverlesung über Analysis kennen.

Wir müssen also zeigen:

$$(\forall t \in [0, 1]) [C(\alpha f + \beta g)](t) = [\alpha C(f) + \beta C(g)](t).$$

Es ist

$$\begin{aligned} & [C(\alpha f + \beta g)](t) \\ &= \int_0^t h(t-s)[\alpha f + \beta g](s) ds \quad \text{Definition von } C \\ &= \int_0^t h(t-s)(\alpha f(s) + \beta g(s)) ds \quad \text{Definition der Lin.kombination von Funktionen} \\ &= \int_0^t (\alpha h(t-s)f(s) + \beta h(t-s)g(s)) ds \quad \text{Rechenregeln in } \mathbb{K} \\ &= \alpha \int_0^t h(t-s)f(s) ds + \beta \int_0^t h(t-s)g(s) ds \quad \text{Rechenregeln für das Integral} \\ &= \alpha [C(f)](t) + \beta [C(g)](t) \quad \text{Definition von } C \\ &= [\alpha C(f) + \beta C(g)](t) \quad \text{Definition der Lin.kombination von Funktionen} \end{aligned}$$

□

Funktionsräume eignen sich auch gut als Beispiele, um herauszuarbeiten, welche Aussagen nur in endlich dimensionalen Vektorräumen gelten:

BEISPIEL 2.2.2. [Ein surjektiver Homomorphismus, der nicht injektiv ist] Sei

$$T : \begin{cases} \mathcal{F}([0, \infty), \mathbb{R}) & \rightarrow \mathcal{F}([0, \infty), \mathbb{R}) \\ [T(f)](x) & := f(x+1) \text{ für alle } x \in [0, \infty) \end{cases}.$$

Dann ist T surjektiv, aber nicht injektiv.

BEWEIS. Sei $g : [0, \infty) \rightarrow \mathbb{R}$ eine Funktion. Wir setzen

$$f(x) := \begin{cases} 0 & \text{wenn } x < 1, \\ g(x-1) & \text{wenn } x \geq 1. \end{cases}$$

Dann ist $g = T(f)$, und daher ist T surjektiv.

Setzen wir nun $h(x) = \max(0, 1-x)$ und sei 0 die Nullfunktion, dann gilt $T(h) = T(0) = 0$. Also ist T nicht injektiv. □

2.3. Normen und Skalarprodukte in Funktionenräumen.

BEISPIEL 2.3.1. Sei $\mathcal{C}([0, 1], \mathbb{R})$ die Menge der stetigen Funktionen von $[0, 1]$ nach \mathbb{R} . Dann ist $\mathcal{C}([0, 1], \mathbb{R})$ ein Vektorraum über \mathbb{R} , und die folgenden Funktionen $\mathcal{C}([0, 1], \mathbb{R}) \rightarrow \mathbb{R}$ sind Normen:

$$\begin{aligned} \|f\|_\infty &:= \sup_{t \in [0, 1]} |f(t)| \quad \text{Maximumnorm,} \\ \|f\|_1 &= \int_0^1 |f(t)| dt \quad L^1\text{-Norm.} \end{aligned}$$

BEWEIS. Wir zeigen, dass $\|\cdot\|_\infty$ eine Norm auf $\mathcal{C}([0, 1], \mathbb{R})$ ist. Seien f und $g : [0, 1] \rightarrow \mathbb{R}$ stetig, und $\lambda \in \mathbb{R}$.

- (1) Weil für alle
- $t \in [0, 1]$
- gilt
- $|f(t)| \geq 0$
- , muss natürlich auch gelten

$$\|f\|_\infty = \sup_{t \in [0,1]} |f(t)| \geq 0.$$

Wenn $\sup_{t \in [0,1]} |f(t)| = 0$, muss für alle $t \in [0, 1]$ gelten $|f(t)| = 0$, und daher ist f dann die Nullfunktion.

- (2) Für den Fall
- $\lambda = 0$
- gilt trivialerweise

$$\|0f\|_\infty = \|0\|_\infty = 0 = 0\|f\|_\infty.$$

Wir betrachten nun den Fall $\lambda \neq 0$: Wegen der Definition von $\|f\|_\infty$ gilt $|f(t)| \leq \|f\|_\infty$ für alle $t \in [0, 1]$. Ist $\lambda \in \mathbb{R}$, so ist dann

$$|\lambda f(t)| = |\lambda| |f(t)| \leq |\lambda| \|f\|_\infty.$$

Weil das für alle $t \in [0, 1]$ gilt, ist auch

$$\|\lambda f\|_\infty = \sup_{t \in [0,1]} |\lambda f(t)| \leq |\lambda| \|f\|_\infty.$$

Wenn man dieses Resultat auf $1/\lambda$ statt λ und λf statt f anwendet, erhält man

$$\|f\|_\infty = \left\| \frac{1}{\lambda} (\lambda f) \right\|_\infty \leq \frac{1}{|\lambda|} \|\lambda f\|_\infty,$$

und es folgt

$$|\lambda| \|f\|_\infty \leq \|\lambda f\|_\infty.$$

- (3) Für alle
- $t \in [0, 1]$
- gilt
- $|f(t)| \leq \|f\|_\infty$
- und
- $|g(t)| \leq \|g\|_\infty$
- . Daher gilt auch

$$|[f + g](t)| = |f(t) + g(t)| \leq |f(t)| + |g(t)| \leq \|f\|_\infty + \|g\|_\infty.$$

Weil das für alle $t \in [0, 1]$ gilt, folgt

$$\|f + g\|_\infty = \sup_{t \in [0,1]} |[f + g](t)| \leq \|f\|_\infty + \|g\|_\infty.$$

Den Beweis, dass die L^1 -Norm eine Norm ist, überlassen wir als Übung. \square

Ob eine Folge in einem Funktionenraum konvergiert, hängt von der Norm ab, die man betrachtet.

BEISPIEL 2.3.2. Wir betrachten die Folge

$$f_n(x) = \begin{cases} n^{3/2}x & \text{für } x \in [0, \frac{1}{2n}] \\ n^{3/2}(\frac{n}{2n}x) & \text{für } x \in (\frac{1}{2n}, \frac{1}{n}) \\ 0 & \text{für } x \in [\frac{1}{n}, 1] \end{cases} \quad \text{im Raum } \mathcal{C}([0, 1], \mathbb{R}).$$

Dann gilt:

- (1) Für alle $x \in [0, 1]$ ist $\lim_{n \rightarrow \infty} f_n(x) = 0$ (punktweise Konvergenz).
- (2) $\lim_{n \rightarrow \infty} \|f_n\|_1 = 0$ (Konvergenz gegen die Nullfunktion hinsichtlich der L^1 -Norm).
- (3) $\|f_n\|_\infty \rightarrow \infty$ (keine Konvergenz, nicht einmal Beschränktheit bezüglich der Supremumsnorm).

BEWEIS.

- (1) Sei $x = 0$. Dann ist $f_n(x) = 0$ für alle $n \in \mathbb{N}$. Ist $x > 0$, so ist $f_n(x) = 0$ für alle $n \geq \frac{1}{x}$. In beiden Fällen ist $\lim_{n \rightarrow \infty} f_n(x) = 0$.
- (2)

$$\begin{aligned} \|f_n\|_1 &= \int_0^1 |f_n(x)| dx = \int_0^{1/(2n)} n^{3/2}x dx + \int_{1/(2n)}^{1/n} n^{3/2}(\frac{1}{n} - x) dx \\ &= n^{3/2} \frac{1}{8n^2} + n^{3/2} \frac{1}{8n^2} = \frac{1}{4\sqrt{n}} \rightarrow 0. \end{aligned}$$

(3)

$$\|f_n\|_\infty = \max_{x \in [0,1]} |f_n(x)| = n^{3/2} \frac{1}{2n} = \frac{1}{2} \sqrt{n} \rightarrow \infty.$$

□

BEHAUPTUNG 2.3.3. Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Auf $\mathcal{C}([0, 1], \mathbb{K})$ ergibt die folgende Definition ein Skalarprodukt:

$$\langle f, g \rangle := \int_0^1 \overline{f(x)} g(x) dx.$$

Es induziert die sogenannte L^2 -Norm

$$\|f\|_2 = \sqrt{\int_0^1 |f(x)|^2 dx}.$$

BEWEIS. Direktes Nachrechnen der Eigenschaften eines Skalarproduktes:

(1) Sesquilinearität:

$$\begin{aligned} \langle f, \alpha g + \beta h \rangle &= \int_0^1 \overline{f(x)} (\alpha g(x) + \beta h(x)) dx \\ &= \int_0^1 [\alpha \overline{f(x)} g(x) + \beta \overline{f(x)} h(x)] dx \\ &= \alpha \int_0^1 \overline{f(x)} g(x) dx + \beta \int_0^1 \overline{f(x)} h(x) dx \\ &= \alpha \langle f, g \rangle + \beta \langle f, h \rangle. \\ \langle \alpha g + \beta h, f \rangle &= \int_0^1 (\overline{\alpha g(x) + \beta h(x)}) f(x) dx \\ &= \int_0^1 [\overline{\alpha} \overline{g(x)} f(x) + \overline{\beta} \overline{h(x)} f(x)] dx \\ &= \overline{\alpha} \int_0^1 \overline{g(x)} f(x) dx + \overline{\beta} \int_0^1 \overline{h(x)} f(x) dx \\ &= \overline{\alpha} \langle g, f \rangle + \overline{\beta} \langle h, f \rangle. \end{aligned}$$

(2) Hermitizität:

$$\begin{aligned} \langle f, g \rangle &= \int_0^1 \overline{f(x)} g(x) dx \\ &= \int_0^1 \overline{f(x) g(x)} dx \\ &= \overline{\int_0^1 g(x) f(x) dx} \\ &= \overline{\langle g, f \rangle}. \end{aligned}$$

(3) Positive Definitheit:

$$\langle f, f \rangle = \int_0^1 \overline{f(x)} f(x) dx = \int_0^1 |f(x)|^2 dx.$$

Weil $|f(x)|^2$ immer nichtnegativ ist, muss das Integral nichtnegativ sein. Also ist $\langle f, f \rangle \geq 0$. In der Analysis gibt es den Satz: Ist g eine stetige Funktion $[0, 1] \rightarrow \mathbb{R}$ mit $g(x) \geq 0$ für alle x , und ist $\int_0^1 g(x) dx = 0$, dann

ist g die Nullfunktion. Daraus folgt: Ist $\langle f, f \rangle = 0$, so ist $f(x) = 0$ für alle $x \in [0, 1]$.

□

BEISPIEL 2.3.4 (Ein Unterraum mit $U^\perp \neq U$). Auf $\mathbb{C}([0, 1], \mathbb{R})$ betrachten wir das L^2 -innere Produkt und den Unterraum

$$U = \{f \in \mathcal{C}([0, 1], \mathbb{R}) \mid f(\frac{1}{2}) = 0\}.$$

Dann ist $U^\perp = \{0\}$ und daher $(U^\perp)^\perp = \mathcal{C}([0, 1], \mathbb{R})$.

BEWEIS. Sei $g : [0, 1] \rightarrow \mathbb{R}$ stetig mit $g \neq 0$. Sei $x \in [0, 1]$ so, dass $g(x) \neq 0$. Ohne Beschränkung der Allgemeinheit sei $x \geq \frac{1}{2}$ und $g(x) > 0$. Wegen der Stetigkeit gibt es ein Intervall $J = (x, x + 2\varepsilon) \subset [\frac{1}{2}, 1]$ sodass $g(x) > 0$ auf J . Wir setzen

$$f(y) = \begin{cases} 0 & \text{wenn } y \in [0, x] \\ y - x & \text{wenn } y \in (x, x + \varepsilon] \\ x + 2\varepsilon - y & \text{wenn } y \in (x + \varepsilon, x + 2\varepsilon] \\ 0 & \text{wenn } y > x + 2\varepsilon \end{cases}.$$

Insbesondere ist $f(\frac{1}{2}) = 0$, also $f \in U$. Dann ist $f(x)g(x) \geq 0$ für alle $x \in [0, 1]$ und $f(x)g(x) > 0$ für $x \in J$. Daher ist

$$\langle f, g \rangle = \int_0^1 f(x)g(x) dx > 0.$$

Also ist dann $g \neq U^\perp$. Es folgt $U^\perp = \{0\}$.

□

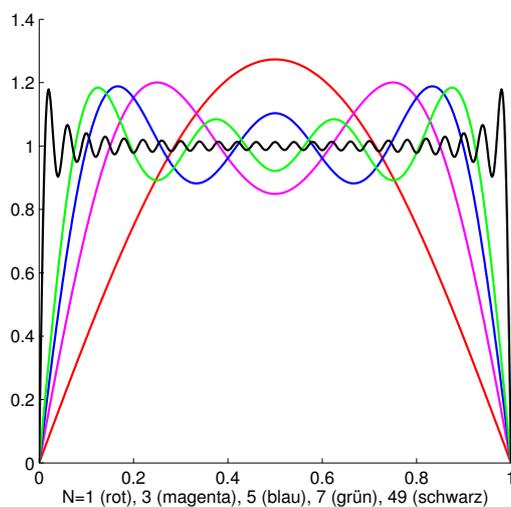
BEISPIEL 2.3.5. Wir definieren für $n \in \mathbb{N}$ die folgenden Funktionen $s_n \in \mathcal{C}([0, 1], \mathbb{R})$

$$s_n(x) = \sqrt{2} \sin(n\pi x).$$

- (1) Dann ist $\{s_n \mid n \in \mathbb{N}\}$ ein Orthonormalsystem bezüglich des L^2 -inneren Produktes.
- (2) Sei f die konstante Funktion $f(x) = 1$ und $N \in \mathbb{N}$. Die orthogonale Projektion von f auf die lineare Hülle von $\{s_1, \dots, s_N\}$ ist

$$\sum_{1 \leq n \leq N, n \text{ ungerade}} \frac{2\sqrt{2}}{n\pi} s_n.$$

Die folgende Grafik zeigt die orthogonalen Projektionen der konstanten Funktion $f(x) = 1$ auf die linearen Hüllen $\mathcal{L}(s_1, \dots, s_N)$ für einige N .



BEWEIS.

(1) Wir brauchen eine Formel über Winkelfunktionen

$$\sin(\alpha) \sin(\beta) = \frac{1}{2} [\cos(\alpha - \beta) - \cos(\alpha + \beta)].$$

Damit ergibt sich für $n \neq m$

$$\begin{aligned} \langle s_m, s_n \rangle &= 2 \int_0^1 \sin(n\pi x) \sin(m\pi x) dx \\ &= \int_0^1 \cos(n-m)\pi x dx - 2 \int_0^1 \cos((n+m)\pi x) dx \\ &= -\frac{1}{(n-m)\pi} \sin((n-m)\pi x) \Big|_{x=0}^1 + \frac{1}{(n+m)\pi} \sin((n+m)\pi x) \Big|_{x=0}^1 \\ &= 0. \end{aligned}$$

Für $m = n$ ergibt sich

$$\begin{aligned} \langle s_n, s_n \rangle &= 2 \int_0^1 \sin(n\pi x) \sin(m\pi x) dx \\ &= \int_0^1 \cos(0) dx - \int_0^1 \cos((n+m)\pi x) dx \\ &= 1 - \frac{2}{(n+m)\pi} \cos((n+m)\pi x) \Big|_{x=0}^1 = 0. \end{aligned}$$

(2)

$$\begin{aligned} \langle s_n, f \rangle &= \sqrt{2} \int_0^1 \sin(n\pi x) dx \\ &= -\frac{\sqrt{2}}{n\pi} \cos(n\pi x) \Big|_{x=0}^1 \\ &= \begin{cases} \frac{2\sqrt{2}}{n\pi} & \text{wenn } n \text{ ungerade} \\ 0 & \text{wenn } n \text{ gerade} \end{cases}. \end{aligned}$$

□

Wenn wir die Theorie der Reihen aus der Analysis adäquat für Vektorräume anpassen, können wir auf diese Weise Funktionen in Reihen von Sinusschwingungen oder auch anderen Orthonormalsystemen zerlegen. Man kommt zur Theorie der Fourierreihen und, wenn man die Reihen noch durch Integrale ersetzen kann, zur Fouriertransformation. Fourieranalysis spielt eine wichtige Rolle in Bereichen der Signalverarbeitung, Regeltechnik und auch Statistik.

3. Basis und Koordinaten

3.1. Lineare Unabhängigkeit und Basis.

DEFINITION 3.1.1. Sei V ein Vektorraum über einem Körper \mathbb{K} .

- (1) Ein System (v_1, \dots, v_k) von Vektoren in V heißt linear unabhängig, wenn die einzige Linearkombination, welche

$$\sum_{i=1}^k \lambda_i v_i = 0$$

ergibt, die Koeffizienten $\lambda_1 = \dots = \lambda_k = 0$ hat.

- (2) Eine Teilmenge $S \subseteq V$ heißt linear unabhängig, wenn jedes System (v_1, \dots, v_k) von endlich vielen, paarweise verschiedenen⁵ Vektoren aus S linear unabhängig ist.

Mit Hilfe dieser Definition können wir auch von unendlichen Mengen sagen, ob sie linear abhängig oder unabhängig sind.

Es besteht ein Unterschied zwischen Menge und System:

- Die Menge $\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$ ist linear unabhängig, denn es ist dieselbe Menge wie $\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$. ES macht für die Menge keinen Unterschied, wie oft ein Element aufgezählt wird, es kann nur entweder dazu gehören oder nicht.
- Das System $\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right)$ ist linear abhängig, weil ein Vektor doppelt darin vorkommt.

In der Praxis bereitet uns dieser Unterschied aber keine Probleme.

BEISPIEL 3.1.2. Im Raum $\mathcal{C}([0, \infty), \mathbb{R})$ der stetigen Funktionen von $[0, \infty)$ nach \mathbb{R} betrachten wir die Funktionen $f_k(x) = e^x$. Dabei ist $k \in \mathbb{R}$.

Dann ist die Menge $\{f_k \mid k \in \mathbb{R}\}$ linear unabhängig.

BEWEIS. Sei mit paarweise verschiedenen k_1, \dots, k_n

$$\sum_{i=1}^n \lambda_i f_{k_i} = 0.$$

Wir zeigen, dass alle λ_i gleich Null sind. Wenn das nicht so wäre, gäbe es einen Index j , sodass k_j der größte Exponent ist, für den $\lambda_j \neq 0$ gilt. Es ist dann auch für alle $x \in [0, \infty)$

$$0 = e^{-k_j x} \sum_{i=1}^k \lambda_i f_{k_i}(x) = \sum_{i=1}^k \lambda_i e^{(k_i - k_j)x}.$$

Für $i \neq j$ gilt dann entweder $\lambda_i = 0$ oder $k_i < k_j$. In jedem dieser Fälle ist

$$\lim_{x \rightarrow \infty} \lambda_i e^{(k_i - k_j)x} = 0.$$

Damit bleibt

$$0 = \lim_{x \rightarrow \infty} \sum_{i=1}^n \lambda_i e^{(k_i - k_j)x} = \lim_{x \rightarrow \infty} \lambda_j e^{(k_j - k_j)x} = \lambda_j$$

⁵das heißt: keine zwei Vektoren v_i, v_j sind gleich, außer $i = j$.

im Widerspruch zur Bedingung $\lambda_j \neq 0$ bei der Wahl des Index j . \square

DEFINITION 3.1.3. Sei V ein Vektorraum über einem Körper \mathbb{K} und S eine Teilmenge von V .

S heißt eine Basis von V , wenn S zugleich linear unabhängig und ein Erzeugendensystem von V ist.

Die folgende Charakterisierung einer Basis geht im allgemeinen Vektorraum genauso wie im \mathbb{K}^n :

BEHAUPTUNG 3.1.4. Sei V ein Vektorraum über einem Körper \mathbb{K} , und $S \subseteq V$. Es sind äquivalent:

- (1) S ist eine maximale linear unabhängige Teilmenge in V , das heißt, S ist linear unabhängig, und jede echte Obermenge von S ist linear abhängig.
- (2) S ist linear unabhängig und ein Erzeugendensystem von V .
- (3) Jedes $x \in V$ lässt sich auf genau eine Weise als Linearkombination

$$x = \sum_{i=1}^k \lambda_i v_i$$

mit $k \in \mathbb{N}$, paarweise verschiedenen $v_i \in S$ und $\lambda_i \in \mathbb{K}$ schreiben.

- (4) S ist ein minimales Erzeugendensystem von V , das heißt, die lineare Hülle von S ist V , aber keine echte Teilmenge von S spannt V auf.

BEWEIS. (i) \Rightarrow (ii): Sei S maximal linear unabhängig. Wir müssen zeigen, dass S ein Erzeugendensystem von V ist. Sei $x \in V$. Wenn $x \in S$, ist trivialerweise $x \in \mathcal{L}(S)$. Wenn $x \notin S$, dann ist $T = S \cup \{x\}$ eine echte Obermenge von S und daher linear abhängig. Dann gibt es eine Linearkombination

$$\sum_{i=1}^k \lambda_i v_i = 0$$

mit paarweise verschiedenen $v_i \in T$ und $\lambda_i \in \mathbb{K}$, wobei nicht alle λ_i gleich Null sind. Eines der v_i , für welches $\lambda_i \neq 0$ ist, muss x sein. Sonst hätten wir eine nichttriviale Linearkombination von Vektoren aus S , welche Null ergibt, und S wäre linear abhängig. Ohne Beschränkung der Allgemeinheit sei also $v_k = x$, $\lambda_k = 0$. Dann ist

$$x = v_k = -\frac{1}{\lambda_k} \sum_{i=1}^{k-1} \lambda_i v_i,$$

und daher liegt x in der linearen Hülle von S .

(ii) \Rightarrow (iii): Weil S ein Erzeugendensystem ist, lässt sich jedes $x \in V$ auf mindestens eine Weise als Linearkombination von Vektoren aus S schreiben. Angenommen, x lässt sich auf zwei Weisen schreiben:

$$x = \sum_{i=1}^k \lambda_i v_i = \sum_{j=1}^l \mu_j w_j$$

mit $v_i, w_j \in S$. Wir setzen

$$\{v_1 \cdots v_k\} \cup \{w_1 \cdots w_l\} = M.$$

Weil M endlich ist, können wir die Elemente benennen

$$M = \{u_1, \dots, u_m\}.$$

Alle Vektoren v_i und w_j kommen in M vor, daher kann man die beiden Linearkombinationen als Kombinationen von Vektoren aus M schreiben: (Für w_i , welche nicht in der ersten Linearkombination vorkommen, schreibt man einfach $\tilde{\lambda}_i = 0$.)

$$x = \sum_{i=1}^m \tilde{\lambda}_i w_i = \sum_{i=1}^m \tilde{\mu}_i w_i,$$

und es folgt

$$\sum_{i=1}^m (\tilde{\lambda}_i - \tilde{\mu}_i) w_i = 0.$$

Wegen der linearen Unabhängigkeit von S gilt für alle i

$$\tilde{\lambda}_i - \tilde{\mu}_i = 0,$$

also bestehen die beiden Linearkombinationen in Wirklichkeit aus denselben Vektoren mit denselben Skalaren.

(iii) \Rightarrow (iv): Weil sich jedes $x \in V$ als Linearkombination von Vektoren aus S schreiben lässt, ist S ein Erzeugendensystem. Wir müssen nur zeigen, dass S minimal ist. Sei T eine echte Teilmenge von S und $x \in S \setminus T$. Dann ist $x = 1x$ eine Möglichkeit, x als Linearkombination von Vektoren aus S zu schreiben. Wegen der Eindeutigkeit gibt es keine andere, und daher erst recht keine Möglichkeit, x als Linearkombination von Vektoren aus T zu schreiben. Also ist $x \notin \mathcal{L}(T)$, und T ist kein Erzeugendensystem von V .

(iv) \Rightarrow (i): Weil S ein Erzeugendensystem ist, kann keine echte Obermenge von S linear unabhängig sein. Denn ist $S \subsetneq T$ und $x \in T \setminus S$, so lässt sich x als Linearkombination von Vektoren aus S schreiben

$$x = \sum_{i=1}^k \lambda_i v_i \quad \text{mit } v_i \in S.$$

Dann ist aber

$$0 = (-1)x + \sum_{i=1}^k \lambda_i v_i$$

eine nichttriviale Linearkombination von Vektoren aus T , welche 0 ergibt.

S selbst ist linear unabhängig. Denn sonst gäbe es ein $y \in S$, welches sich durch die anderen Vektoren aus S linearkombinieren lässt. Dann wäre aber $S \setminus \{y\}$ auch noch ein Erzeugendensystem von V , im Widerspruch zur Minimalität von S . \square

Jetzt stellt sich die Frage, welche Vektorräume eine Basis besitzen. Eine einfache Variante des Basisauswahlsatzes ist das folgende Ergebnis:

BEHAUPTUNG 3.1.5. *Sei V ein Vektorraum über einem Körper \mathbb{K} mit einem endlichen Erzeugendensystem $\{v_1, \dots, v_n\}$. Dann lässt sich aus diesem Erzeugendensystem eine Basis von V auswählen.*

BEWEIS. Ist $\{v_1, \dots, v_n\}$ bereits ein minimales Erzeugendensystem, dann ist es selbst eine Basis von V . Andernfalls lässt man schrittweise Vektoren aus dem System weg, bis ein minimales Erzeugendensystem erreicht ist. \square

Man kann beweisen, dass jeder Vektorraum eine Basis besitzt. Allerdings benötigt man dafür ein schweres Geschütz der Mengenlehre, das Auswahlaxiom:

Sei M eine Menge, $\mathcal{P}(M)$ ihre Potenzmenge, das ist die Menge aller Teilmengen von M . Sei Ω eine weitere Menge und $g : \Omega \rightarrow \mathcal{P}(M)$ eine Abbildung. Dann gibt es eine Abbildung $h : \Omega \rightarrow M$, sodass für alle $x \in \Omega$ gilt $h(x) \in g(x)$.

Das Auswahlaxiom ist weit weniger harmlos als es auf ersten Blick aussieht. Man kann damit zum Beispiel beweisen, dass man eine Kugel im \mathbb{R}^3 in drei disjunkte Bestandteile zerlegen kann,

diese Bestandteile im Raum drehen und verschieben und dann so zusammensetzen kann, dass eine echt größere Kugel daraus entsteht. Andererseits hat das Auswahlaxiom viele nützliche Konsequenzen, weshalb man es nicht missen möchte. Man verwendet es also, aber nur, um Resultate zu beweisen, die ohne Auswahlaxiom nicht beweisbar wären.

Für die angewandte Mathematik spielen Basen von unendlich dimensionalen Vektorräumen eine geringe Rolle. Lieber als durch endliche Linearkombinationen von Basisvektoren stellt man in solchen Räumen Vektoren als Reihensummen von unendlich vielen Vektoren, beispielsweise aus Orthonormalsystemen, dar.

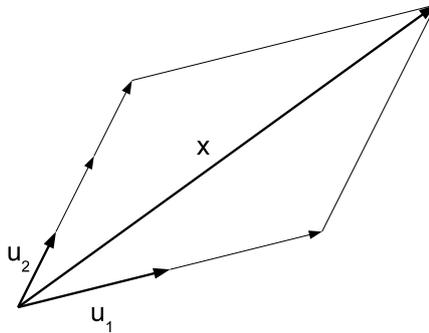
3.2. Koordinaten.

DEFINITION 3.2.1. Sei V ein Vektorraum über einem Körper \mathbb{K} mit einer endlichen Basis $B = (b_1, \dots, b_n)$. Sei $v \in V$.

Der Koordinatenvektor von v bezüglich der Basis B ist jener Vektor $(\xi_1, \dots, \xi_n)^T$ für den gilt:

$$v = \sum_{i=1}^n \xi_i b_i.$$

In der folgenden Grafik hat der Vektor x bezüglich der Basis (u_1, u_2) den Koordinatenvektor $(2, 3)^T$.



Wir werden ein Koordinatensystem nun als Isomorphismus zwischen V und \mathbb{K}^n interpretieren. Das sieht zunächst ein wenig umständlich aus, hilft uns aber später, die Übersicht zu behalten, wenn wir zwischen verschiedenen Koordinatensystemen hin und her wechseln.

BEHAUPTUNG 3.2.2. Sei V ein Vektorraum über einem Körper \mathbb{K} und sei $B = (b_1, \dots, b_n)$ eine endliche Basis von V .

Wir definieren eine Abbildung $\Phi_B : \mathbb{K}^n \rightarrow V$ durch

$$\Phi_B \left(\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \right) := \sum_{i=1}^n \xi_i b_i.$$

Dann ist Φ_B ein Isomorphismus.

BEWEIS. Weil B eine Basis von V ist, gibt es zu jedem $x \in V$ eine eindeutige Darstellung

$$x = \sum_{i=1}^n \xi_i b_i,$$

das heißt, es gibt einen eindeutigen Vektor $y \in \mathbb{K}^n$ mit $x = \Phi_B(y)$. Daher ist Φ_B bijektiv.

Die Linearität rechnet man leicht nach:

$$\begin{aligned} \Phi_B\left(\alpha \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} + \beta \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix}\right) &= \Phi_B\left(\begin{pmatrix} \alpha\xi_1 + \beta\eta_1 \\ \vdots \\ \alpha\xi_n + \beta\eta_n \end{pmatrix}\right) = \sum_{i=1}^n (\alpha\xi_i + \beta\eta_i)v_i \\ &= \alpha \sum_{i=1}^n \xi_i v_i + \beta \sum_{i=1}^n \eta_i v_i = \alpha \Phi_B\left(\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}\right) + \beta \Phi_B\left(\begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix}\right). \end{aligned}$$

□

BEMERKUNG 3.2.3. Es ist also Φ_B so definiert, dass $\Phi_B^{-1}(v)$ der Koordinatenvektor von v bezüglich der Basis B ist.

Wir erhalten also das folgende wichtige Resultat:

KOROLLAR 3.2.4. *Jeder endlich erzeugte Vektorraum über einem Körper \mathbb{K} ist isomorph zu \mathbb{K}^n mit einem geeigneten n .*

Dadurch können wir alles, was wir über den \mathbb{K}^n gelernt haben, auf jeden endlich erzeugten Vektorraum übertragen, darunter auch

BEHAUPTUNG 3.2.5. *Ist (b_1, \dots, b_n) eine Basis eines Vektorraumes V über einem Körper \mathbb{K} , dann bestehen auch alle anderen Basen von V aus genau n Elementen.*

DEFINITION 3.2.6. Ist (b_1, \dots, b_n) eine Basis eines Vektorraumes V , so nennen wir n die Dimension von V .

Alle weiteren Konsequenzen können Sie finden, indem Sie die Kapitel über Unterräume und Dimension des \mathbb{K}^n auf V übertragen.

Wir werden jetzt noch die Matrizenrechnung in allgemeinen, endlich dimensional Vektorräumen einführen.

BEHAUPTUNG 3.2.7. *Seien V und W zwei endlichdimensionale Vektorräume über demselben Körper \mathbb{K} . Sei $f : V \rightarrow W$ eine lineare Abbildung. Es sei $B = (b_1, \dots, b_n)$ eine Basis von V und $C = (c_1, \dots, c_m)$ eine Basis von W . Mit Φ_B und Φ_C bezeichnen wir wie üblich die Isomorphismen*

$$\Phi_B : \begin{cases} \mathbb{K}^n & \rightarrow V \\ \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} & \mapsto \sum_{j=1}^n \xi_j b_j \end{cases}, \quad \Phi_C : \begin{cases} \mathbb{K}^m & \rightarrow W \\ \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_m \end{pmatrix} & \mapsto \sum_{i=1}^m \eta_i c_i \end{cases}.$$

Dann ist

$$\Phi_C^{-1} \circ f \circ \Phi_B : \mathbb{K}^n \rightarrow \mathbb{K}^m$$

der Homomorphismus, der für jedes $x \in V$ den Koordinatenvektor von x (bezüglich B) in den Koordinatenvektor von $f(x)$ (bezüglich C) überführt.

BEWEIS. Sei $x \in V$. Dann ist der Koordinatenvektor von x bezüglich B der Vektor $\Phi_B^{-1}(x)$. Dann ist

$$[\Phi_C^{-1} \circ f \circ \Phi_B](\Phi_B^{-1}(x)) = \Phi_C^{-1}(f(x))$$

der Koordinatenvektor von $f(x)$ bezüglich der Basis C . \square

DEFINITION 3.2.8. Seien V und W zwei endlichdimensionale Vektorräume über demselben Körper \mathbb{K} . Sei $f : V \rightarrow W$ eine lineare Abbildung. Es sei $B = (b_1, \dots, b_n)$ eine Basis von V und $C = (c_1, \dots, c_m)$ eine Basis von W , und Φ_B, Φ_C die entsprechenden Koordinatenabbildungen.

Dann bezeichnet $M_C^B(f) \in \mathbb{K}^{m \times n}$ die Matrix des Homomorphismus $\Phi_C^{-1} \circ f \circ \Phi_B$. $M_C^B(f)$ heißt die Matrix von f bezüglich der Basen B und C .

Das folgende Diagramm macht diese Situation anschaulich:

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \uparrow \Phi_B & & \downarrow \Phi_C^{-1} \\ \mathbb{K}^n & \xrightarrow{M_C^B(f)} & \mathbb{K}^m \end{array}$$

Jeder Pfeil stellt einen Homomorphismus dar. Wird ein Pfeil durch einen weiteren fortgesetzt, drückt das die Hintereinanderausführung der beiden Homomorphismen aus. Führen mehrere Wege (also mehrere Hintereinanderausführungen von Homomorphismen) von einem Raum zum anderen, so ergeben die verschiedenen Hintereinanderausführungen dieselbe Abbildung: Wir nennen ein solches Diagramm ein kommutatives Diagramm. Zum Beispiel ist oben: $M_C^B(f) = \Phi_C^{-1} \circ f \circ \Phi_B$.

SATZ 3.2.9. Seien V und W zwei endlichdimensionale Vektorräume über demselben Körper \mathbb{K} . Sei $f : V \rightarrow W$ eine lineare Abbildung. Es sei $B = (b_1, \dots, b_n)$ eine Basis von V und $C = (c_1, \dots, c_m)$ eine Basis von W , und Φ_B, Φ_C die entsprechenden Koordinatenabbildungen. Sei $M_C^B(f)$ die Matrix von f bezüglich der Basen B und C . Dann gilt:

Die Spalten der Matrix $M_C^B(f)$ sind die Koordinaten der Vektoren $f(b_1), \dots, f(b_n)$ bezüglich der Basis C .

BEWEIS. Die Spalten von $M_C^B(f)$ sind die Bilder $M_C^B(f)e_1, \dots, M_C^B(f)e_n$ der Einheitsvektoren $e_1, \dots, e_n \in \mathbb{K}^n$. Nun ist $\Phi_B(e_j) = b_j$, und

$$M_C^B(f)e_j = \Phi_C^{-1} \circ f \circ \Phi_B(e_j) = \Phi_C^{-1}(f(b_j)).$$

Dies ist genau der Koordinatenvektor von $f(b_j)$. \square

SATZ 3.2.10. Seien U, V, W drei Vektorräume über \mathbb{K} mit den (endlichen) Dimensionen n, m, k . Seien $f : U \rightarrow V$ und $g : V \rightarrow W$ lineare Abbildungen. Seien B, C, D Basen von U, V, W und Φ_B, Φ_C, Φ_D die entsprechenden Koordinatenabbildungen. Es gilt:

$$M_D^B(g \circ f) = M_D^C(g)M_C^B(f).$$

BEWEIS. Die Situation wird durch das folgende kommutative Diagramm verdeutlicht:

$$\begin{array}{ccccc} U & \xrightarrow{f} & V & \xrightarrow{g} & W \\ \uparrow \Phi_B & & \uparrow \Phi_C & & \uparrow \Phi_D \\ \mathbb{K}^n & \xrightarrow{M_C^B(f)} & \mathbb{K}^m & \xrightarrow{M_D^C(g)} & \mathbb{K}^k \end{array}$$

Es ist

$$M_D^B(g \circ f) = \Phi_D^{-1} \circ (g \circ f) \circ \Phi_B = (\Phi_D^{-1} \circ g \circ \Phi_C) \circ (\Phi_C^{-1} \circ f \circ \Phi_B) = M_D^C(g)M_C^B(f).$$

□

SATZ 3.2.11. *Seien V und W zwei endlichdimensionale Vektorräume über demselben Körper \mathbb{K} mit den (endlichen) Dimensionen n und m . Sei $f : V \rightarrow W$ eine lineare Abbildung. Es seien B, C Basen von V, W , und Φ_B, Φ_C die entsprechenden Koordinatenabbildungen. Sei $M_C^B(f)$ die Matrix von f bezüglich der Basen B und C . Dann gilt:*

Die Abbildung f ist genau dann ein Isomorphismus, wenn $n = m$ und die Matrix $M_C^B(f)$ regulär ist. In diesem Fall ist

$$M_B^C(f^{-1}) = [M_C^B(f)]^{-1}.$$

BEWEIS. Wenn $n \neq m$, also wenn die Dimensionen von V und W verschieden sind, sind V und W bekanntlich nicht isomorph.

Sei nun $n = m$ und $M_C^B(f)$ regulär. Wir definieren den Homomorphismus $g : W \rightarrow V$ durch

$$g := \Phi_B \circ [M_C^B(f)]^{-1} \circ \Phi_C^{-1}.$$

Es gilt dann

$$\begin{aligned} g \circ f &= \Phi_B \circ [M_C^B(f)]^{-1} \circ \Phi_C^{-1} \circ f = \Phi_B \circ [M_C^B(f)]^{-1} \circ \Phi_C^{-1} \circ f \circ \Phi_B \circ \Phi_B^{-1} \\ &= \Phi_B \circ [M_C^B(f)]^{-1} \circ M_C^B(f) \circ \Phi_B^{-1} = \text{id}_V. \end{aligned}$$

Ebenso kann man zeigen:

$$f \circ g = \text{id}_W.$$

Also besitzt f eine Inverse, nämlich $f^{-1} = g$.

Sei nun f ein Isomorphismus und $M_B^C(f^{-1})$ die Matrix der Umkehrabbildung bezüglich der Basen C und B . Es ist

$$M_B^C(f^{-1})M_C^B(f) = \Phi_B^{-1} \circ f^{-1} \circ \Phi_C \circ \Phi_C^{-1} \circ f \circ \Phi_B = E_n.$$

Ebenso zeigt man

$$M_C^B(f)M_B^C(f^{-1}) = E_n.$$

Daher ist $M_B^C(f^{-1})$ die Inverse zu $M_C^B(f)$. □

3.3. Koordinatentransformation.

BEHAUPTUNG 3.3.1. *Sei V ein n -dimensionaler Vektorraum über \mathbb{K} . Seien $B = (b_1, \dots, b_n)$ und $C = (c_1, \dots, c_n)$ zwei Basen von V . Es seien Φ_B und Φ_C die Isomorphismen*

$$\Phi_B : \begin{cases} \mathbb{K}^n & \rightarrow V \\ \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} & \mapsto \sum_{j=1}^n \xi_j b_j \end{cases}, \quad \Phi_C : \begin{cases} \mathbb{K}^n & \rightarrow V \\ \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix} & \mapsto \sum_{i=1}^n \eta_i c_i \end{cases}.$$

Dann ist $\Phi_C^{-1}\Phi_B : \mathbb{K}^n \rightarrow \mathbb{K}^n$ jener Isomorphismus, der für jedes $v \in V$ den Koordinatenvektor $\Phi_B^{-1}(v)$ von v bezüglich B in den Koordinatenvektor $\Phi_C^{-1}(v)$ bezüglich C überführt.

Das folgende Diagramm macht die Situation anschaulich. Wir können $\Phi_C^{-1}\Phi_B$ auch als die Matrix $M_C^B(\text{id}_V)$ auffassen.

$$\begin{array}{ccc} V & \xrightarrow{\text{id}_V} & V \\ \uparrow \Phi_B & & \downarrow \Phi_C^{-1} \\ \mathbb{K}^n & \xrightarrow{T_C^B} & \mathbb{K}^n \end{array}$$

DEFINITION 3.3.2. Sei V ein n -dimensionaler Vektorraum über \mathbb{K} . Seien $B = (b_1, \dots, b_n)$ und $C = (c_1, \dots, c_n)$ zwei Basen von V und Φ_B und Φ_C die entsprechenden Koordinatenabbildungen. Wir definieren

$$T_C^B \text{ ist die Matrix von } \Phi_C^{-1} \circ \Phi_B.$$

Wir bezeichnen T_C^B auch als die Übergangsmatrix von der Basis B auf die Basis C .

SATZ 3.3.3. Sei V ein n -dimensionaler Vektorraum über \mathbb{K} . Seien $B = (b_1, \dots, b_n)$ und $C = (c_1, \dots, c_n)$ zwei Basen von V und Φ_B und Φ_C die entsprechenden Koordinatenabbildungen. Dann gilt:

- (1) Ist a der Koordinatenvektor eines Vektors $x \in V$ bezüglich der Basis B , dann ist $T_C^B a$ der Koordinatenvektor von x bezüglich der Basis C .
- (2) Die Matrix T_C^B ist zugleich $M_C^B(\text{id}_V)$.
- (3) Die Matrix T_C^B ist regulär, und ihre Inverse ist T_B^C .
- (4) Die Spalten von T_C^B sind die Basisvektoren b_1, \dots, b_n , geschrieben in Koordinaten bezüglich C .
- (5) Die Spalten von $[T_C^B]^{-1}$ sind die Basisvektoren c_1, \dots, c_n , geschrieben in Koordinaten bezüglich B .⁶

BEWEIS.

- (1) In dieser Situation ist $a = \Phi_B^{-1}(x)$ und

$$T_C^B a = (\Phi_C^{-1} \circ \Phi_B)(\Phi_B^{-1}(x)) = \Phi_C^{-1}(x).$$

- (2) Es ist

$$T_C^B = \Phi_C^{-1} \circ \Phi_B = \Phi_C^{-1} \circ \text{id}_V \circ \Phi_B = M_C^B(\text{id}_V).$$

- (3) Da id_V ein Isomorphismus ist, ist nach Satz 3.2.11 $T_C^B = M_C^B(\text{id}_V)$ regulär mit der Inversen

$$[T_C^B]^{-1} = M_B^C(\text{id}_V^{-1}) = M_B^C(\text{id}_V) = T_B^C.$$

- (4) Nach Satz 3.2.9 die Spalten von $T_C^B = M_C^B(\text{id}_V)$ die Koordinatenvektoren von

$$\text{id}_V(b_1) = b_1, \dots, \text{id}_V(b_n) = b_n$$

bezüglich C .

- (5) Punkt (3) und $[T_C^B]^{-1} = T_B^C$.

□

BEISPIEL 3.3.4. Sei C die Basis (c_1, c_2, c_3) in \mathbb{R}^3 .

- (1) Stellen Sie die Übergangsmatrix von der Basis C zur Einheitsbasis E auf.
- (2) Berechnen Sie die Übergangsmatrix von der Einheitsmatrix E zur Basis C .

⁶In manchen Aufgaben hat man die neuen Basisvektoren c_1, \dots, c_n in alten Koordinaten (bezüglich B) gegeben und kann so zuerst $[T_C^B]^{-1} = T_B^C$ bestimmen und anschließend die Inverse berechnen.

(3) Berechnen Sie die Koordinaten des Vektors x bezüglich der Basis C .

$$c_1 = \begin{pmatrix} 2 \\ 4 \\ 2 \end{pmatrix}, c_2 = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}, c_3 = \begin{pmatrix} 3 \\ 4 \\ 8 \end{pmatrix}, x = \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix}.$$

Loesung:

Wir haben die Koordinaten der Basisvektoren c_1, c_2, c_3 und die Koordinaten von x bezüglich der Einheitsbasis gegeben.

(1) Wir setzen die Basisvektoren c_1, c_2, c_3 zu einer Matrix zusammen, das ist die Übergangsmatrix von der Basis C zur Einheitsbasis E :

$$T_E^C = \begin{pmatrix} 2 & 1 & 3 \\ 4 & 3 & 4 \\ 2 & 1 & 8 \end{pmatrix}.$$

(2) Die Übergangsmatrix von der Einheitsbasis zur Basis C ist $T_C^E = (T_E^C)^{-1}$:

$$\begin{aligned} & \begin{pmatrix} 2 & 1 & 3 & 1 & 0 & 0 \\ 4 & 3 & 4 & 0 & 1 & 0 \\ 2 & 1 & 8 & 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 1 & 3 & 1 & 0 & 0 \\ -2 & 0 & -5 & -3 & 1 & 0 \\ 0 & 0 & 5 & -1 & 0 & 1 \end{pmatrix} \\ \rightarrow & \begin{pmatrix} 2 & 1 & 0 & 1.6 & 0 & -0.6 \\ -2 & 0 & 0 & -4 & 1 & 1 \\ 0 & 0 & 1 & -0.2 & 0 & 0.2 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 1 & 0 & -2.4 & 1 & 0.4 \\ 1 & 0 & 0 & 2 & -0.5 & -0.5 \\ 0 & 0 & 1 & -0.2 & 0 & 0.2 \end{pmatrix} \\ \rightarrow & \begin{pmatrix} 1 & 0 & 0 & 2 & -0.5 & -0.5 \\ 0 & 1 & 0 & -2.4 & 1 & 0.4 \\ 0 & 0 & 1 & -0.2 & 0 & 0.2 \end{pmatrix} \end{aligned}$$

Also ist

$$T_C^E = \begin{pmatrix} 2 & -0.5 & -0.5 \\ -2.4 & 1 & 0.4 \\ -0.2 & 0 & 0.2 \end{pmatrix}.$$

(3) Der Koordinatenvektor von x bezüglich C ist dann

$$T_C^E x = \begin{pmatrix} 2 & -0.5 & -0.5 \\ -2.4 & 1 & 0.4 \\ -0.2 & 0 & 0.2 \end{pmatrix} \begin{pmatrix} 4 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 6.5 \\ -7.2 \\ -0.6 \end{pmatrix}.$$

□

SATZ 3.3.5. Sei V ein n -dimensionaler Vektorraum mit Basen A und B und dazugehörigen Koordinatenabbildungen $\Phi_A, \Phi_B : \mathbb{K}^n \rightarrow V$. Sei W ein m -dimensionaler Vektorraum mit Basen C und D und dazugehörigen Koordinatenabbildungen $\Phi_C, \Phi_D : \mathbb{K}^m \rightarrow W$. Sei $f : V \rightarrow W$ eine lineare Abbildung.

Dann ist

$$M_D^A(f) = T_D^C M_C^B(f) T_B^A.$$

BEWEIS. Wir schreiben f als Hintereinanderausführung:

$$f = \text{id}_W \circ f \circ \text{id}_V,$$

wie das folgende kommutative Diagramm verdeutlicht:

$$\begin{array}{ccccccc} V & \xrightarrow{\text{id}_V} & V & \xrightarrow{f} & W & \xrightarrow{\text{id}_W} & W \\ \uparrow \Phi_A & & \uparrow \Phi_B & & \uparrow \Phi_C & & \uparrow \Phi_D \\ \mathbb{K}^n & \xrightarrow{T_B^A} & \mathbb{K}^n & \xrightarrow{M_C^B(f)} & \mathbb{K}^m & \xrightarrow{T_D^C} & \mathbb{K}^m \end{array}$$

Nach Satz 3.2.10 gilt

$$M_D^A(f) = M_D^C(\text{id}_W) M_C^B(f) M_B^A(\text{id}_V) = T_D^C M_C^B(f) T_B^A.$$

□

KOROLLAR 3.3.6. Sei V ein n -dimensionaler Vektorraum über \mathbb{K} , seien B und C zwei Basen von V mit den entsprechenden Koordinatenabbildungen $\Phi_B, \Phi_C : \mathbb{K}^n \rightarrow V$. Sei $f : V \rightarrow V$ ein Endomorphismus. Es gilt:

$$M_C^C(f) = T_C^B M_B^B(f) T_B^C = [T_B^C]^{-1} M_B^B(f) T_B^C.$$

BEWEIS. Satz 3.3.5 und Satz 3.3.3. □

DEFINITION 3.3.7. Seien $n \in \mathbb{N}$, $P, Q \in \mathbb{K}^{n \times n}$ Matrizen gleicher Dimension über einem Körper \mathbb{K} .

Die Matrizen P und Q heißen ähnliche Matrizen, wenn es eine reguläre Matrix $T \in \mathbb{K}^{n \times n}$ gibt, so dass $Q = T^{-1}PT$.

BEMERKUNG 3.3.8. Seien $n \in \mathbb{N}$, $P, Q \in \mathbb{K}^{n \times n}$ quadratische Matrizen gleicher Dimension über einem Körper \mathbb{K} .

Die Matrizen P und Q sind genau dann ähnlich, wenn es einen n -dimensionalen Vektorraum V mit zwei Basen A, B und einen Endomorphismus $f : V \rightarrow V$ gibt, so dass P und Q die Matrizen derselben Abbildung f bezüglich verschiedener Basen in V sind:

$$P = M_A^A(f), \quad Q = M_B^B(f).$$

BEWEIS. Übung! Können Sie das? Tipp: Für V und W kann man \mathbb{K}^n und \mathbb{K}^m nehmen. Für die Basen A und C kann man die Einheitsbasen heranziehen. □

BEISPIEL 3.3.9. Gegeben sei der Endomorphismus

$$f : \begin{cases} \mathbb{R}^3 & \rightarrow \mathbb{R}^3, \\ x & \mapsto Px, \end{cases} \quad \text{mit } P = \begin{pmatrix} 4 & 0 & -2 \\ 0 & 2 & 2 \\ -2 & 2 & 3 \end{pmatrix}.$$

(P ist also die Matrix von f bezüglich der Einheitsbasis E .)

Wir betrachten in \mathbb{R}^3 die neue Basis

$$C = \left(\begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix} \right).$$

Wie sieht die Matrix $M_C^C(f)$ von f bezüglich der Basis C aus?

Lösung: Wir berechnen zunächst die Matrix zur Koordinatentransformation: Die Spalten von T_E^C sind die Basisvektoren von C in Einheitskoordinaten, also ist

$$T_E^C = \begin{pmatrix} 2 & -2 & 1 \\ 2 & 1 & -2 \\ 1 & 2 & 2 \end{pmatrix}.$$

Als nächstes berechnen wir $T_C^E = [T_E^C]^{-1}$:

$$\begin{aligned} & \begin{pmatrix} 2 & -2 & 1 & 1 & 0 & 0 \\ 2 & 1 & -2 & 0 & 1 & 0 \\ 1* & 2 & 2 & 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & -6 & -3 & 1 & 0 & -2 \\ 0 & -3* & -6 & 0 & 1 & -2 \\ 1 & 2 & 2 & 0 & 0 & 1 \end{pmatrix} \\ \rightarrow & \begin{pmatrix} 0 & 0 & 9* & 1 & -2 & 2 \\ 0 & 1 & 2 & 0 & -1/3 & 2/3 \\ 1 & 0 & -2 & 0 & 2/3 & -1/3 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 0 & 1 & 1/9 & -2/9 & 2/9 \\ 0 & 1 & 0 & -2/9 & 1/9 & 2/9 \\ 1 & 0 & 0 & 2/9 & 2/9 & 1/9 \end{pmatrix} \\ \rightarrow & \begin{pmatrix} 1 & 0 & 0 & 2/9 & 2/9 & 1/9 \\ 0 & 1 & 0 & -2/9 & 1/9 & 2/9 \\ 0 & 0 & 1 & 1/9 & -2/9 & 2/9 \end{pmatrix} \end{aligned}$$

Also ist

$$T_C^E = \frac{1}{9} \begin{pmatrix} 2 & 2 & 1 \\ -2 & 1 & 2 \\ 1 & -2 & 2 \end{pmatrix}.$$

Nun ist

$$\begin{aligned} M_C^C(f) &= T_C^E P T_E^C \\ &= \frac{1}{9} \begin{pmatrix} 2 & 2 & 1 \\ -2 & 1 & 2 \\ 1 & -2 & 2 \end{pmatrix} \begin{pmatrix} 4 & 0 & -2 \\ 0 & 2 & 2 \\ -2 & 2 & 3 \end{pmatrix} \begin{pmatrix} 2 & -2 & 1 \\ 2 & 1 & -2 \\ 1 & 2 & 2 \end{pmatrix} \\ &= \frac{1}{9} \begin{pmatrix} 2 & 2 & 1 \\ -2 & 1 & 2 \\ 1 & -2 & 2 \end{pmatrix} \begin{pmatrix} 6 & -12 & 0 \\ 6 & 6 & 0 \\ 3 & 12 & 0 \end{pmatrix} \\ &= \frac{1}{9} \begin{pmatrix} 27 & 0 & 0 \\ 0 & 54 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Bezüglich der Basis C hat f die Diagonalmatrix

$$M_C^C(f) = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

□

Im Abschnitt über Eigenwerte werden wir lernen, wie man systematisch Basen finden kann, in denen sich Endomorphismen besonders einfach schreiben lassen.

4. Vektorräume aus Vektorräumen konstruieren

4.1. Durchschnitt und Summenraum.

4.1.1. Durchschnitt.

BEHAUPTUNG 4.1.1. Sei V ein Vektorraum über einem Körper \mathbb{K} , sei Λ eine Menge, und für jedes $\lambda \in \Lambda$ sei U_λ ein Unterraum von V . Dann ist der Durchschnitt

$$\bigcap_{\lambda \in \Lambda} U_\lambda$$

ein Unterraum von V .

BEWEIS. Weil jedes U_λ ein Unterraum ist, ist $0 \in U_\lambda$. Daher ist der Nullvektor auch im Durchschnitt aller U_λ enthalten, und der Durchschnitt ist nicht leer.

Seien $x, y \in \bigcap_{\lambda \in \Lambda} U_\lambda$ und $\alpha, \beta \in \mathbb{K}$. Weil jedes U_λ ein Unterraum ist, und x, y in jedem U_λ enthalten sind, ist auch

$$(\forall \lambda \in \Lambda) \alpha x + \beta y \in U_\lambda.$$

Daher ist $\alpha x + \beta y \in \bigcap_{\lambda \in \Lambda} U_\lambda$. \square

4.1.2. Summenraum.

Die Vereinigung von Unterräumen ist normalerweise kein Unterraum, wie man aus dem folgenden Beispiel sieht:

BEISPIEL 4.1.2. Die Mengen

$$U = \left\{ \lambda \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\} \quad \text{und} \quad W = \left\{ \lambda \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mid \lambda \in \mathbb{R} \right\}$$

sind Unterräume des \mathbb{R}^2 , aber $U \cup W$ ist kein Unterraum.

DEFINITION 4.1.3. Sei $(V, +, \cdot)$ ein Vektorraum über einem Körper \mathbb{K} . Für $i = 1, \dots, n$ sei W_i jeweils ein Unterraum von V . Wir definieren den Summenraum

$$W_1 + \dots + W_n := \{w_1 + \dots + w_n \mid (\forall i = 1 \dots n) w_i \in W_i\}.$$

BEHAUPTUNG 4.1.4. Sei $(V, +, \cdot)$ ein Vektorraum über einem Körper \mathbb{K} . Seien W_1, \dots, W_n Unterräume von V . Dann gilt:

- (1) Der Summenraum $W_1 + \dots + W_n$ ist ein Unterraum von V .
- (2) Für jedes $i = 1, \dots, n$ sei M_i jeweils ein Erzeugendensystem von W_i . Dann ist die Vereinigung $\bigcup_{i=1}^n M_i$ ein Erzeugendensystem von $W_1 + \dots + W_n$.

BEWEIS. *Punkt (1):* Natürlich ist $0 = 0 + \dots + 0 \in W_1 + \dots + W_n$. Seien $x, y \in W_1 + \dots + W_n$, $\alpha, \beta \in \mathbb{K}$. Wir zeigen: $\alpha x + \beta y \in W_1 + \dots + W_n$. Nach Definition des Summenraumes lassen sich x, y in folgender Weise schreiben:

$$x = \sum_{i=1}^n w_i, \quad y = \sum_{i=1}^n \tilde{w}_i, \quad \text{mit } w_i, \tilde{w}_i \in W_i.$$

Dann ist

$$\alpha x + \beta y = \sum_{i=1}^n (\alpha w_i + \beta \tilde{w}_i).$$

Da die W_i Unterräume von V sind, ist $\alpha w_i + \beta \tilde{w}_i \in W_i$, und somit ist $\alpha x + \beta y \in W_1 + \dots + W_n$.

Punkt (2): Wir zeigen nun, dass $W_1 + \dots + W_n$ der kleinste Unterraum ist, der alle M_i als Teilmengen enthält. Sei also U ein Unterraum von V , welcher alle M_i enthält. Zunächst enthält U jedes M_i , und damit auch $\mathcal{L}(M_i) = W_i$. Ist $x \in W_1 + \dots + W_n$, so lässt sich x schreiben als $x = \sum_{i=1}^n w_i$ mit $w_i \in W_i$. Da alle W_i in U enthalten sind, sind alle $w_i \in U$, und weil U ein Unterraum ist, ist damit auch $\sum_{i=1}^n w_i \in U$. Somit ist $W_1 + \dots + W_n \subseteq U$. \square

Achtung: Wenn $(w_{i,1}, \dots, w_{i,r_i})$ jeweils eine Basis von W_i ist, folgt daraus noch nicht, dass $(w_{1,1}, \dots, w_{n,r_n})$ eine Basis des Summenraumes ist. Zwar haben wir noch ein Erzeugendensystem des Summenraumes, aber dieses könnte linear abhängig sein. — Weiteres im Abschnitt über direkte Summen weiter unten.

4.1.3. Dimensionsformel.

SATZ 4.1.5 (Dimensionsformel). Sei $(V, +, \cdot)$ ein endlich erzeugter Vektorraum über einem Körper \mathbb{K} . Seien A, B zwei Unterräume von V . Dann gilt

$$\dim(A + B) + \dim(A \cap B) = \dim(A) + \dim(B).$$

BEWEIS. Da V endlich erzeugt ist, ist jeder Unterraum von V endlich erzeugt und besitzt eine Basis aus endlich vielen Vektoren. Wir beginnen mit einer Basis (x_1, \dots, x_r) von $A \cap B$. Weil $A \cap B$ sowohl in A als auch in B enthalten ist, ist (x_1, \dots, x_r) ein linear unabhängiges Vektorsystem sowohl in A als auch in B , und läßt sich nach dem Basisergänzungssatz zu je einer Basis von A und B erweitern. Wir haben also die folgenden Basen und Dimensionen:

$$\begin{array}{ll} (x_1, \dots, x_r) & \text{Basis von } A \cap B; \quad \dim(A \cap B) = r \\ (x_1, \dots, x_r, a_1, \dots, a_s) & \text{Basis von } A \quad ; \quad \dim(A) = r + s \\ (x_1, \dots, x_r, b_1, \dots, b_t) & \text{Basis von } B \quad ; \quad \dim(B) = r + t \end{array}$$

Wir zeigen nun, dass $(x_1, \dots, x_r, a_1, \dots, a_s, b_1, \dots, b_t)$ eine Basis von $A + B$ ist. Nach Behauptung 4.1.4 ist dieses System jedenfalls ein Erzeugendensystem von $A + B$, da es Vereinigung einer Basis von A mit einer Basis von B ist. Wir zeigen die lineare Unabhängigkeit. Es sei also

$$\sum_{i=1}^r \lambda_i x_i + \sum_{j=1}^s \mu_j a_j + \sum_{k=1}^t \nu_k b_k = 0.$$

Wir schreiben das in der Form

$$\sum_{i=1}^r \lambda_i x_i + \sum_{j=1}^s \mu_j a_j = - \sum_{k=1}^t \nu_k b_k.$$

Da alle $x_i \in A$, $a_j \in A$, ist $\sum_{i=1}^r \lambda_i x_i + \sum_{j=1}^s \mu_j a_j \in A$. Andererseits ist $-\sum_{k=1}^t \nu_k b_k \in B$. Also ist $-\sum_{k=1}^t \nu_k b_k \in A \cap B$. Weil aber (x_1, \dots, x_r) eine Basis von $A \cap B$ ist, muss es jedenfalls eine Darstellung der Form geben (mit geeigneten Skalaren ξ_i):

$$- \sum_{k=1}^t \nu_k b_k = \sum_{i=1}^r \xi_i x_i.$$

Das System $(x_1, \dots, x_r, b_1, \dots, b_t)$ ist aber eine Basis von B . Wir können also denselben Vektor auf zwei Arten als Linearkombination dieser Basis schreiben. Wegen der linearen Unabhängigkeit der Basis muss es sich in Wirklichkeit um dieselbe Linearkombination handeln. Die erste Darstellung enthält keinen der Basisvektoren x_i , die zweite enthält keinen der Basisvektoren b_k . Folglich sind alle Koeffizienten ν_k und ξ_i gleich Null.

$$\sum_{i=1}^r \lambda_i x_i + \sum_{j=1}^s \mu_j a_j = - \sum_{k=1}^t \nu_k b_k = 0,$$

und wegen der linearen Unabhängigkeit von $(x_1, \dots, x_r, a_1, \dots, a_s)$ sind auch alle λ_i und alle μ_j gleich Null. Also ist das System $(x_1, \dots, x_r, a_1, \dots, a_s, b_1, \dots, b_t)$ linear unabhängig, und somit ist es eine Basis von $A + B$. Insbesondere ist $\dim(A + B) = r + s + t$.

Nun können wir die Dimensionsformel zusammensetzen:

$$\dim(A + B) + \dim(A \cap B) = (r + s + t) + r = (r + s) + (r + t) = \dim(A) + \dim(B).$$

□

4.2. Direkte Summen.

4.2.1. *Direkte Summen innerhalb eines Vektorraums.*

DEFINITION 4.2.1. Sei $(V, +, \cdot)$ ein Vektorraum über \mathbb{K} . Seien W_1, \dots, W_n Unterräume von V . Der Summenraum $W_1 + \dots + W_n$ heißt direkte Summe von W_1, \dots, W_n wenn gilt:

Für jedes $x \in W_1 + \dots + W_n$ gibt es eindeutig bestimmte $w_i \in W_i$ sodass $x = \sum_{i=1}^n w_i$.

SCHREIBWEISE 4.2.2. Manche Autoren schreiben für die Summe von Unterräumen innerhalb eines Vektorraumes, wenn sie direkt ist, $W_1 \oplus \dots \oplus W_n$. In diesem Skriptum werde ich diese Schreibweise vermeiden, um nicht Konfusion mit der äußeren direkten Summe zu schaffen.

BEHAUPTUNG 4.2.3. Sei $(V, +, \cdot)$ ein Vektorraum über einem Körper \mathbb{K} . Sei $V = W_1 + \dots + W_n$ die direkte Summe von Unterräumen $W_i \subseteq V$.

(1) Dann ist für jedes $k = 1, \dots, n$ die folgende Abbildung wohldefiniert:

$$\pi_k : \begin{cases} V & \rightarrow W_k, \\ \sum_{i=1}^n w_i & \mapsto w_k. \end{cases}$$

(2) Die Abbildung π_k ist linear.

BEWEIS.

(1) Nach Definition des Summenraumes läßt sich jedes $v \in V$ als die Summe $\sum_{i=1}^n w_i$ schreiben. Weil die Summe direkt ist, ist diese Zerlegung von v eindeutig, sodass die obige Definition von π_k eine eindeutige Abbildung festlegt.

(2) Seien $x = w_1 + \dots + w_n \in V$, $y = \tilde{w}_1 + \dots + \tilde{w}_n \in V$ mit $w_i, \tilde{w}_i \in W_i$. Seien $\alpha, \beta \in \mathbb{K}$. Dann ist also

$$\pi_i(x) = w_i, \quad \pi_i(y) = \tilde{w}_i,$$

und

$$\alpha x + \beta y = \sum_{i=1}^n (\alpha w_i + \beta \tilde{w}_i),$$

sodass

$$\pi_i(\alpha x + \beta y) = \alpha w_i + \beta \tilde{w}_i = \alpha \pi_i(x) + \beta \pi_i(y).$$

□

DEFINITION 4.2.4. Die Abbildung π_k in Behauptung 4.2.3 heißt die Projektion des Summenraumes auf den Teilraum W_k entlang der Zerlegung $W_1 + \dots + W_n$.

Der Buchstabe π wird gerne für Projektionen verwendet. Aus dem Zusammenhang wird jeweils klar, ob π eine orthogonale Projektion, eine Projektion entlang einer Zerlegung, oder eine andere Art von Projektion bezeichnet.

BEHAUPTUNG 4.2.5. Sei $(V, +, \cdot)$ ein Vektorraum über einem Körper \mathbb{K} . Seien W_1, \dots, W_n Unterräume von V . Dann gilt:

(1) Die Summe $W_1 + \dots + W_n$ ist genau dann direkt, wenn gilt: Sind

$$w_1 \in W_{i_1} \setminus \{0\}, \dots, w_k \in W_{i_k} \setminus \{0\}$$

mit paarweise verschiedenen i_1, \dots, i_k , dann ist das System (w_1, \dots, w_k) linear unabhängig.

(2) Ist $n = 2$, so gilt: Die Summe $W_1 + W_2$ ist genau dann direkt, wenn gilt $W_1 \cap W_2 = \{0\}$.

- (3) Sind W_1, \dots, W_n endlich erzeugt, so gilt: Die Summe $W_1 + \dots + W_n$ ist genau dann direkt, wenn gilt: Sind $(w_{i,1}, \dots, w_{i,r_i})$ jeweils Basen von W_i , so ist das System $(w_{1,1}, \dots, w_{n,r_n})$ eine Basis von $W_1 + \dots + W_n$.
- (4) Sind W_1, \dots, W_n endlich erzeugt, so ist die Summe der W_i genau dann direkt, wenn gilt:

$$\dim(W_1 + \dots + W_n) = \sum_{i=1}^n \dim(W_i).$$

BEWEIS.

- (1) Sei die Summe direkt, seien $w_j \in W_{i_j}$, alle ungleich Null. Wir zeigen, dass (w_1, \dots, w_n) linear unabhängig ist. Sei also $\sum_{j=1}^n \lambda_j w_j = 0$. Da die Summe direkt ist, läßt sich der Nullvektor nur in einer eindeutigen Form $0 = \sum_{i=1}^n \tilde{w}_i$ mit $\tilde{w}_i \in W_i$ zerlegen, und natürlich geschieht das durch $\tilde{w}_i = 0$. Also muss für jedes j gelten: $\lambda_j w_j = 0$. Weil aber $w_j \neq 0$, folgt $\lambda_j = 0$ für alle j .

Umgekehrt sei nun jedes k -Tupel von Vektoren $w_j \in W_{i_j} \setminus \{0\}$ mit paarweise verschiedenen i_j linear unabhängig. Wir zeigen, dass die Summe direkt ist, also das jedes $w \in W_1 + \dots + W_n$ sich nur auf eine Weise in Vektoren aus W_i zerlegen lässt. Sei also $w = \sum_{i=1}^n u_i = \sum_{i=1}^n v_i$ mit $u_i, v_i \in W_i$. Dann ist $0 = \sum_{i=1}^n (u_i - v_i)$. Wir nehmen an, es wären nicht alle $u_i = v_i$. Wir setzen dann:

$$\begin{aligned} \{i_1, \dots, i_k\} &= \{i \mid u_i - v_i \neq 0\}, \\ w_j &= u_{i_j} - v_{i_j} \text{ für } j = 1, \dots, k. \end{aligned}$$

Dann ist $\sum_{j=1}^k w_j = 0$, im Widerspruch zur linearen Unabhängigkeit der w_j .

- (2) Sei $W_1 + W_2$ direkt, und $u \in W_1 \cap W_2$. Dann läßt sich u als Summe $u = w_1 + w_2$ mit $w_i \in W_i$ auf die folgenden Weisen zerlegen: $u = u + 0 = 0 + u$. Da die Zerlegung aber nach Voraussetzung eindeutig ist, folgt $u = 0$. Umgekehrt sei $W_1 \cap W_2 = \{0\}$ und $u = w_1 + w_2 = v_1 + v_2$ mit $v_i, w_i \in W_i$. Dann ist $w_1 - v_1 = v_2 - w_2$. Nun ist $w_1 - v_1 \in W_1$, derselbe Vektor $v_2 - w_2 \in W_2$. Daher ist $w_1 - v_1 \in W_1 \cap W_2$, also $w_1 - v_1 = 0$, und ebenso $v_2 - w_2 = 0$. Also ist die Zerlegung eindeutig: $v_1 = w_1, v_2 = w_2$.
- (3) Sei die Summe direkt. Wir zeigen: $(w_{1,1}, \dots, w_{n,r_n})$ ist linear unabhängig. Da dieses Vektorsystem den Summenraum erzeugt, ist es dann eine Basis. Sei also

$$\sum_{i=1}^n \sum_{j=1}^{r_i} \lambda_{i,j} w_{i,j} = 0.$$

Da die Summe direkt ist, läßt sich der Nullvektor nur auf eine Weise zerlegen, und es muss also gelten:

$$\sum_{j=1}^{r_i} \lambda_{i,j} w_{i,j} = 0 \text{ für alle } i.$$

Da aber $(w_{i,1}, \dots, w_{i,r_i})$ eine Basis von W_i , und daher linear unabhängig ist, folgt $\lambda_{i,j} = 0$ für alle i, j .

Sei umgekehrt $(w_{1,1}, \dots, w_{n,r_n})$ eine Basis des Summenraumes. Wir zeigen, dass die Summe direkt ist. Sei $x = \sum v_i = \sum u_i$ mit $u_i, v_i \in W_i$. Es lassen sich u_i, v_i mit Hilfe der Basen schreiben:

$$u_i = \sum_{j=1}^{r_i} \lambda_{i,j} w_{i,j}, \quad v_i = \sum_{j=1}^{r_i} \mu_{i,j} w_{i,j}.$$

Dann ist

$$x = \sum_{i=1}^n \sum_{j=1}^{r_i} \lambda_{i,j} w_{i,j} = \sum_{i=1}^n \sum_{j=1}^{r_i} \mu_{i,j} w_{i,j}.$$

Wegen der linearen Unabhängigkeit der $w_{i,j}$ folgt $\lambda_{i,j} = \mu_{i,j}$ für alle i, j , und folglich $u_i = v_i$ für alle i .

(4) Direkte Folge aus Punkt (3). □

BEHAUPTUNG 4.2.6. Sei \mathbb{K}^n die direkte Summe von Unterräumen W_1, \dots, W_k , wobei zu jedem W_i eine Basis $(w_{i,1}, \dots, w_{i,m_i})$ gegeben ist. Sei π_i die Projektionen von \mathbb{K}^n auf W_i entlang der Summenzerlegung (siehe Definition 4.2.4). Die Matrix P_i der Projektion π_i erhält man folgendermaßen:

- (1) Man erstellt eine Matrix T aus allen Basisvektoren, also eine Matrix mit den Spalten $w_{1,1} \dots w_{k,m_k}$.
- (2) Man berechnet T^{-1} .
- (3) Sind die Spalten von T mit Nummern $p \dots q$ die Basisvektoren $w_{i,1}, \dots, w_{i,m_i}$ des Unterraumes W_i , so setzt man aus den Spalten Nummer $p \dots q$ von T eine Matrix R_i , und aus den Zeilen Nummer $p \dots q$ der Matrix T^{-1} eine Matrix S_i zusammen.
- (4) $P_i = R_i S_i$.

BEWEIS. In diesem Beweis unterscheiden wir die Einheitsvektoren $e_1, \dots, e_n \in \mathbb{K}^n$ und die Einheitsvektoren $\tilde{e}_1, \dots, \tilde{e}_{m_i} \in \mathbb{K}^{m_i}$. Für jeden Einheitsvektor $e_\ell \in \mathbb{K}^n$ ist $T e_\ell$ einer der Basisvektoren $w_{j,r}$. Dabei ist $j = i$, wenn $1 \leq \ell \leq q$. Daher ist

$$R_i \tilde{e}_r = w_{i,r}.$$

Es folgt auch

$$T^{-1} w_{j,r} = e_\ell, \quad \text{wobei } p \leq \ell \leq q \iff j = i,$$

und daher

$$S_i w_{j,r} = \begin{cases} \tilde{e}_r & \text{wenn } j = i \\ 0 & \text{sonst} \end{cases}.$$

Daher ist

$$P_i w_{j,r} = \begin{cases} R_i \tilde{e}_r = w_{i,r} = \pi_i(w_{i,r}) & \text{wenn } j = i \\ R_i 0 = 0 = \pi_i(w_{j,r}) & \text{sonst} \end{cases}.$$

□

4.2.2. Äußere direkte Summe.

Manchmal will man aus zwei Vektorräumen, die nicht von vorneherein Teilräume eines gemeinsamen Oberraumes sind, eine Summe konstruieren.

DEFINITION 4.2.7. Seien V_1, V_2 zwei Vektorräume über demselben Körper \mathbb{K} . Die (äußere) direkte Summe von V_1 und V_2 wird definiert als Produktmenge:

$$V_1 \oplus V_2 := \{(v_1, v_2) \mid v_i \in V_i\},$$

mit folgenden Rechenoperationen für $(x_1, x_2), (y_1, y_2) \in V_1 \oplus V_2, \alpha \in \mathbb{K}$.

$$\begin{aligned} (x_1, x_2) + (y_1, y_2) &:= (x_1 + y_1, x_2 + y_2), \\ \alpha(x_1, x_2) &:= (\alpha x_1, \alpha x_2). \end{aligned}$$

Durch die injektiven Abbildungen

$$j_1 : \begin{cases} V_1 & \rightarrow V_1 \oplus V_2, \\ x_1 & \mapsto (x_1, 0), \end{cases} \quad j_2 : \begin{cases} V_2 & \rightarrow V_1 \oplus V_2, \\ x_2 & \mapsto (0, x_2), \end{cases}$$

betten wir V_1 und V_2 in den Summenraum ein. (d.h. wir identifizieren $x_1 \in V_1$ mit dem Vektor $j_1(x_1) \in V_1 \oplus V_2$, ebenso wird $x_2 \in V_2$ mit $j_2(x_2) \in V_1 \oplus V_2$ identifiziert.) Die Abbildungen

$$\pi_1 : \begin{cases} V_1 \oplus V_2 & \rightarrow V_1, \\ (x_1, x_2) & \mapsto x_1, \end{cases} \quad \pi_2 : \begin{cases} V_1 \oplus V_2 & \rightarrow V_2, \\ (x_1, x_2) & \mapsto x_2, \end{cases}$$

heißen die kanonischen Projektionen von $V_1 \oplus V_2$ auf V_1 bzw. V_2 .

BEMERKUNG 4.2.8. Diese Definition läßt sich analog ohne Schwierigkeiten auf die Summe von mehr als 2 Vektorräumen verallgemeinern.

BEHAUPTUNG 4.2.9. Seien V_1, V_2 zwei Vektorräume über demselben Körper \mathbb{K} , und sei $V_1 \oplus V_2$ deren äußere direkte Summe.

- (1) $V_1 \oplus V_2$ ist ein Vektorraum über \mathbb{K} .
- (2) Identifiziert man V_i mit $j_i(V_i)$, so sind V_1, V_2 Unterräume von $V_1 \oplus V_2$, und $V_1 \oplus V_2$ ist auch ihre direkte Summe im Sinne einer direkten Summe innerhalb eines Vektorraums.

BEWEIS.

- (1) Der Beweis erfolgt durch Nachrechnen aller Rechenregeln.
- (2) Jedes Paar $(x_1, x_2) \in V_1 \oplus V_2$ läßt sich in eindeutiger Weise in zwei Vektoren aus $j_1(V_1)$ und $j_2(V_2)$ zerlegen: $(x_1, x_2) = (x_1, 0) + (0, x_2)$. Weil es eine solche Zerlegung gibt, ist $V_1 \oplus V_2 = j_1(V_1) + j_2(V_2)$, und weil die Zerlegung eindeutig ist, ist diese Summe direkt. □

4.3. Faktorräume.

4.3.1. Äquivalenzrelationen.

In diesem Kapitel wird uns ein sehr wichtiges Stück Mengenlehre über den Weg laufen, nämlich der Begriff der Äquivalenzrelation. Im mathematischen Alltag stellt man sich eine Relation auf zwei Mengen A, B als eine Beziehung vor, in der je ein Element von A zu einem Element von B stehen kann oder nicht: z.B. sei A eine Menge von Studierenden, B eine Menge von Vorlesungen. Wir führen eine Relation \mathbf{r} ein: $a \mathbf{r} b$ bedeutet: Studierende(r) a besucht regelmäßig die Vorlesung b .

Was für den Alltag taugt, ist viel zu verwaschen für eine Definition, die allen Tücken der Mengenlehre Stand halten soll. Ähnlich wie eine Funktion definieren wir eine Relation formal als Teilmenge von $A \times B$:

DEFINITION 4.3.1 (Relation). Seien A und B nichtleere Mengen und $A \times B = \{(a, b) \mid a \in A, b \in B\}$ ihre Produktmenge. Eine Relation ist eine Teilmenge von $A \times B$.

Ist $\mathbf{r} \subseteq A \times B$ und sind $a \in A, b \in B$, so sagen wir, a steht zu b in der Relation \mathbf{r} , genau dann wenn $(a, b) \in \mathbf{r}$. In diesem Fall schreiben wir auch $a \mathbf{r} b$.

BEISPIEL 4.3.2 (Relationen).

- Die Ordnungsrelation auf \mathbb{N} : $a < b$.
- Die Gleichheit auf einer Menge M : $a = b$.
- Jede Funktion $f : A \rightarrow B$: $(a, b) \in f \iff f(a) = b$.
- Teilbarkeitsrelation: Seien $a, b \in \mathbb{Z}, a \neq 0$, dann definieren wir

$$a \mid b : \iff (\exists c \in \mathbb{Z}) ac = b \quad (\text{“}a \text{ teilt } b\text{“}).$$

DEFINITION 4.3.3 (Äquivalenzrelation). Sei M eine nichtleere Menge, und \sim eine Relation auf $M \times M$. Die Relation \simeq heißt eine Äquivalenzrelation, wenn sie folgende drei Bedingungen erfüllt:

$$\begin{array}{ll} (\forall x \in M) x \sim x & \text{Reflexivität} \\ (\forall x, y \in M) x \sim y \iff y \sim x & \text{Symmetrie} \\ (\forall x, y, z \in M) [(x \sim y) \wedge (y \sim z)] \Rightarrow [x \sim z] & \text{Transitivität} \end{array}$$

BEISPIEL 4.3.4. Folgende Relationen sind Äquivalenzrelationen:

- Gleichheit.
- Seien A, B Mengen und $f : A \rightarrow B$ eine Funktion. Auf A geben wir die Relation $x \sim y : \iff f(x) = f(y)$.
- In \mathbb{Z} Kongruenz modulo m .
- In $\mathbb{K}^{n \times n}$ Ähnlichkeit von Matrizen.

DEFINITION 4.3.5 (Klasseneinteilung). Sei M eine nichtleere Menge, und \mathcal{U} eine Teilmenge der Potenzmenge von M (d.h. die Elemente von \mathcal{U} sind Teilmengen von M). Man nennt \mathcal{U} eine Klasseneinteilung oder Partition auf M , wenn folgende Bedingungen gelten:

$$\begin{array}{l} \emptyset \notin \mathcal{U}, \\ (\forall x \in M) (\exists P \in \mathcal{U}) x \in P, \\ (\forall P, Q \in \mathcal{U}) P \neq Q \Rightarrow P \cap Q = \emptyset. \end{array}$$

BEMERKUNG 4.3.6. Man kann eine Partition auch so charakterisieren: Eine Partition auf einer nichtleeren Menge M ist eine Teilmenge \mathcal{U} der Potenzmenge von M , sodass gilt:

- Kein $P \in \mathcal{U}$ ist leer.
- Jedes $x \in M$ ist in genau einem $P \in \mathcal{U}$ enthalten.

Klasseneinteilungen und Äquivalenzrelationen gehen Hand in Hand:

SATZ 4.3.7. Sei M eine nichtleere Menge.

(a) Sei \sim eine Äquivalenzrelation auf M . Für alle $x \in M$ definieren wir die Äquivalenzklasse von x durch

$$[x]_{\sim} := \{y \in M \mid x \sim y\}.$$

Es sei $\mathcal{U} = \{[x]_{\sim} \mid x \in M\}$. Dann ist \mathcal{U} eine Klasseneinteilung auf M .

(b) Sei \mathcal{U} eine Klasseneinteilung auf M . Wir definieren die Relation

$$x \sim_{\mathcal{U}} y : \iff (\exists P \in \mathcal{U}) x \in P \wedge y \in P.$$

Dann ist $\sim_{\mathcal{U}}$ eine Äquivalenzrelation auf M .

BEWEIS. (a) Wir zeigen die Eigenschaften einer Partition.

- Jede Klasse $P \in \mathcal{U}$ ist nichtleer: Denn nach Definition von \mathcal{U} ist P von der Form $P = [x]_{\sim}$ für ein $x \in M$, und weil $x \sim x$, ist $x \in [x]_{\sim}$, und daher ist $[x]_{\sim} \neq \emptyset$.
- Jedes $x \in M$ liegt in einer Klasse: Ist $x \in M$, so ist $x \in [x]_{\sim}$, also gibt es ein $P \in \mathcal{U}$ (nämlich $P = [x]_{\sim}$) sodass $x \in P$.
- Klassen sind entweder disjunkt oder identisch: Seien nun $P = [x]_{\sim}$ und $Q = [y]_{\sim}$ so, dass $P \cap Q \neq \emptyset$. Wir zeigen $P = Q$. Es sei also $z \in [x]_{\sim} \cap [y]_{\sim}$. Das heißt, $x \sim z$ und $y \sim z$. Sei $w \in P = [x]_{\sim}$. Es gilt also $w \sim x$, und wegen der Transitivität mit $x \sim z$ folgt $w \sim z$. Noch einmal mit der Transitivität und $z \sim y$ folgt $w \sim y$, also $w \in [y]_{\sim}$. Es ist also $[x]_{\sim} \subseteq [y]_{\sim}$. Mit vertauschten Rollen von x und y folgt genauso: $[y]_{\sim} \subseteq [x]_{\sim}$.

(b) Wir zeigen die Eigenschaften einer Äquivalenzrelation.

- *Reflexivität:* Sei $x \in M$. Dann gibt es ein $P \in \mathcal{U}$ so, dass $x \in P$. Weil also $x \in P \wedge x \in P$, folgt $x \sim_{\mathcal{U}} x$.

- *Symmetrie:* Seien $x, y \in M$. Es gilt

$$\begin{aligned} x \sim_{\mathcal{U}} y &\iff (\exists P \in \mathcal{U}) [x \in P \wedge y \in P] \\ &\iff (\exists P \in \mathcal{U}) [y \in P \wedge x \in P] \iff y \sim_{\mathcal{U}} x. \end{aligned}$$

- *Transitivität:* Seien $x, y, z \in M$ sodass $x \sim_{\mathcal{U}} y$ und $y \sim_{\mathcal{U}} z$. Es gibt also $P, Q \in \mathcal{U}$ sodass

$$x \in P \wedge y \in P \quad \text{und} \quad y \in Q \wedge z \in Q.$$

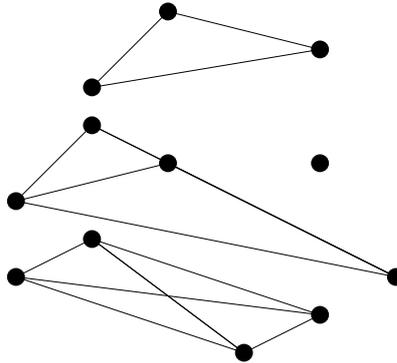
Weil nun $y \in P \cap Q$, ist $P = Q$ und daher ist

$$x \in P \wedge z \in P.$$

Daraus folgt aber $x \sim_{\mathcal{U}} z$.

□

BEISPIEL 4.3.8. Die Graphik zeigt eine Menge von 12 Punkten, auf der eine Äquivalenzrelation \sim definiert ist. Punkte, die in der Relation \sim stehen, wurden durch Linien verbunden. Wir sehen, dass die gesamte Menge in 4 Äquivalenzklassen zerfällt. Innerhalb jeder Klasse ist jeder Punkt mit jedem Punkt verbunden. Zwischen den Klassen gibt es keine Verbindungen. Eine der Klassen besteht nur aus einem Punkt.



4.3.2. Nebenklassen und Faktorräume.

Wir zeigen zunächst mit Zeichnungen einige Nebenklassen eines eindimensionalen Unterraumes von \mathbb{R}^2 , und Nebenklassen von Unterräumen des \mathbb{R}^3 .

SATZ 4.3.9. Sei V ein Vektorraum über \mathbb{K} , und W ein Unterraum von V . Wir definieren die folgende Relation auf V :

$$x \equiv_W y \iff y - x \in W.$$

Dann gilt:

- 1) Die Relation \equiv_W ist eine Äquivalenzrelation auf V .
- 2) Die Relation \equiv_W ist im folgenden Sinn verträglich mit den Vektorraumoperationen: Seien $x_1, x_2, y_1, y_2 \in V$, $\lambda \in \mathbb{K}$ so, dass $x_1 \equiv_W x_2$ und $y_1 \equiv_W y_2$.

Dann ist

$$x_1 + y_1 \equiv_W x_2 + y_2 \quad \text{und} \quad \lambda x_1 \equiv_W \lambda x_2.$$

3) Die Äquivalenzklasse $[x]$ von x bezüglich der Relation \equiv_W ist

$$[x] = x + W = \{x + w \mid w \in W\}.$$

BEWEIS. *Punkt (1):* Wir zeigen die drei Eigenschaften der Äquivalenzrelation. Wir werden immer wieder verwenden, dass W ein Unterraum ist. Zunächst gilt $x - x = 0 \in W$, also $x \equiv_W x$ für jedes $x \in V$. Sind $x, y \in V$ mit $x \equiv_W y$, also $v := y - x \in W$. Dann ist $x - y = -v \in W$, und damit $y \equiv_W x$. Sind $x, y, z \in V$ mit $x \equiv_W y$ und $y \equiv_W z$, so ist also $v := y - x \in W$, und $w = z - y \in W$. Damit ist $z - x = w + v \in W$ und folglich $x \equiv_W z$.

Punkt (2): Sei $x_1 \equiv_W x_2$, d.h., $v := x_2 - x_1 \in W$, und sei $y_1 \equiv_W y_2$, d.h. $w := y_2 - y_1 \in W$. Sei $\lambda \in \mathbb{K}$. Es ist dann $(x_2 + y_2) - (x_1 + y_1) = v + w \in W$, also $x_1 + y_1 \equiv_W x_2 + y_2$, und $\lambda x_2 - \lambda x_1 = \lambda v \in W$, also $\lambda x_1 \equiv_W \lambda x_2$.

Punkt (3): Es ist $y \in [x]$ genau dann, wenn $y \equiv_W x$, also, wenn es ein $w \in W$ gibt mit $y - x = w$, d.h. aber, $y = x + w$. \square

DEFINITION 4.3.10. Sei V ein Vektorraum über \mathbb{K} , und W ein Unterraum von V . Sei \equiv_W und $[x]$ definiert wie in Satz 4.3.9.

- 1) Zu jedem $x \in V$ heißt die Äquivalenzklasse $[x] = x + W$ die Nebenklasse oder Faser von x bezüglich W .
- 2) Die Partition $V/W := \{[x] \mid x \in V\}$ heißt der Faktorraum von V bezüglich W .

Die Nebenklassen von W sind einfach die affinen Unterräume von V , die W als Tangentialraum haben.

BEHAUPTUNG 4.3.11.

(1) Die folgenden Operationen auf V/W sind wohldefiniert: Seien $x, y \in V$ und $[x], [y] \in V/W$ ihre Nebenklassen. Sei $\lambda \in \mathbb{K}$.

$$[x] + [y] := [x + y], \quad \lambda[x] := [\lambda x].$$

(2) Mit diesen Operationen ist V/W ein Vektorraum über \mathbb{K} . Der Nullvektor in V/W ist $[0] = W$.

(3) Die kanonische Projektion

$$\pi : \begin{cases} V & \rightarrow V/W \\ x & \mapsto [x] \end{cases}$$

ist eine surjektive lineare Abbildung mit Kern W .

BEWEIS.

- (1) Jede Nebenklasse $[x] \in V/W$ lässt sich als Nebenklasse verschiedener Elemente aus V zeigen. Wir müssen zeigen, dass die Definition von der Wahl dieser Elemente unabhängig ist. Sei also $[x_1] = [x_2]$, $[y_1] = [y_2]$. Wir müssen zeigen $[x_1 + y_1] = [x_2 + y_2]$. Es gilt aber $x_1 \equiv_W x_2$ weil $x_2 \in [x_1] = [x_1]$ und $y_1 \equiv_W y_2$, also nach Satz 4.3.9 auch $x_1 + y_1 \equiv_W x_2 + y_2$ und $\lambda x_1 \equiv_W \lambda x_2$. Damit ist aber $[x_1 + y_1] = [x_2 + y_2]$ und $[\lambda x_1] = [\lambda x_2]$.

- (2) Hier muss man die Vektorraumaxiome Punkt für Punkt durchrechnen. Wir zeigen zum Beispiel eines der Distributivgesetze: Seien $[x], [y] \in V/W$, $\lambda \in \mathbb{K}$. Es ist dann (unter Benützung von Punkt (1) und des Distributivgesetzes in V):

$$\begin{aligned} \lambda([x] + [y]) &= \lambda[x + y] = [\lambda(x + y)] = [\lambda x + \lambda y] \\ &= [\lambda x] + [\lambda y] = \lambda[x] + \lambda[y]. \end{aligned}$$

Die anderen Axiome gehen ähnlich. Es ist $[-x] = -[x]$ und der Nullvektor in V/W ist $[0]$. Es ist aber $[0] = 0 + W = W$.

- (3) Als kanonische Projektion auf die Klassen einer Äquivalenzrelation ist π eine surjektive Abbildung. Es ist wegen Punkt (2):

$$\pi(\lambda x + \mu y) = [\lambda x + \mu y] = \lambda[x] + \mu[y],$$

also ist π linear. Letztlich ist

$$x \in \ker(\pi) \Leftrightarrow [x] = [0] \Leftrightarrow x \in W.$$

□

BEHAUPTUNG 4.3.12. Sei V ein Vektorraum über einem Körper \mathbb{K} und W ein Unterraum von V .

- 1) Dann gibt es einen Vektorraum Z und eine lineare Abbildung $f : V \rightarrow Z$ so, dass $\ker(f) = W$.
- 2) Ist zusätzlich $\dim(V) < \infty$, dann gilt

$$\dim(V/W) = \dim(V) - \dim(W).$$

BEWEIS. Punkt (1): Wähle $Z = V/W$ und $f = \pi$.

Punkt (2): Es ist $W = \ker(\pi)$ und $V/W = \text{rg}(\pi)$. Weil V und V/W endlich dimensionale Vektorräume sind, gilt für die lineare Abbildung $\pi : V \rightarrow V/W$:

$$\dim(\text{rg}(\pi)) + \dim(\ker(\pi)) = \dim(V).$$

□

Die Konstruktion von Faktorstrukturen wird in der Algebra oft verwendet: Analog kann man für Gruppen oder Ringe Faktorstrukturen definieren. Während aber jeder Unterraum eines Vektorraums die Bildung eines Faktorraumes erlaubt, und damit Kern einer linearen Abbildung sein kann, kann bei nicht-kommutativen Gruppen nur bezüglich ganz speziellen Untergruppen, den sogenannten Normalteilern, eine Faktorgruppe gebildet werden. Auch nicht jeder Unterring eines Ringes eignet sich zur Bildung eines Faktorringes, sondern nur die sogenannten Ideale.

4.3.3. Der Faktorisierungssatz.

SATZ 4.3.13. Seien U, V Vektorräume über \mathbb{K} , sei $f : V \rightarrow U$ eine lineare Abbildung. Es seien

$$\pi : \begin{cases} V & \rightarrow V/\ker(f) \\ x & \mapsto [x] \end{cases}, \quad j : \begin{cases} \text{rg}(f) & \rightarrow U \\ y & \mapsto y \end{cases}.$$

Dann gibt es einen Isomorphismus $\tilde{f} : V/\ker(f) \rightarrow \text{rg}(f)$ so, dass $f = j \circ \tilde{f} \circ \pi$.

Das folgende Diagramm wird also kommutativ:

$$\begin{array}{ccc} V & \xrightarrow{f} & U \\ \pi \downarrow \text{surjektiv} & & j \uparrow \text{injektiv} \\ V/\ker(f) & \xrightarrow[\text{bijektiv}]{\tilde{f}} & \text{rg}(f) \end{array}$$

BEWEIS. Wir definieren

$$\tilde{f} : \begin{cases} V/\ker(f) & \rightarrow \operatorname{rg}(f), \\ [x] & \mapsto f(x). \end{cases}$$

Wir müssen zunächst zeigen, dass \tilde{f} wohldefiniert ist. Jede Nebenklasse aus $V/\ker(f)$ lässt sich in der Form $[x]$ mit $x \in V$ schreiben. Ist $[x_1] = [x_2]$, so müssen wir zeigen, dass $f(x_1) = f(x_2)$. Es gilt aber

$$[x_1] = [x_2] \Leftrightarrow x_2 - x_1 \in \ker(f) \Leftrightarrow f(x_2) - f(x_1) = 0 \Leftrightarrow f(x_1) = f(x_2).$$

Wir haben damit gezeigt, dass \tilde{f} wohldefiniert ist, und zugleich erhalten, dass \tilde{f} injektiv ist. Da jedes $y \in \operatorname{rg}(f)$ sich in der Form $y = f(x)$ mit einem $x \in V$ schreiben lässt, ist

$$y = f(x) = \tilde{f}([x]),$$

also ist \tilde{f} surjektiv. Wir zeigen, dass \tilde{f} linear ist:

$$\tilde{f}(\lambda[x] + \mu[y]) = \tilde{f}([\lambda x + \mu y]) = f(\lambda x + \mu y) = \lambda f(x) + \mu f(y) = \lambda \tilde{f}([x]) + \mu \tilde{f}([y]).$$

□

KOROLLAR 4.3.14. *Seien U, V, W Vektorräume über \mathbb{K} , seien $f : V \rightarrow U$ und $g : V \rightarrow W$ lineare Abbildungen, und sei $\ker(f) = \ker(g)$. Dann sind die Bilder $\operatorname{rg}(f)$ und $\operatorname{rg}(g)$ isomorph.*

BEWEIS. Beide Bilder sind isomorph zu $V/\ker(f)$.

□

Determinanten

1. Definition und Berechnung durch Zeilenumformungen

1.1. Definition der Determinante.

DEFINITION 1.1.1. Sei \mathbb{K} ein Körper, $n \in \mathbb{N}$ und $\det : \mathbb{K}^{n \times n} \rightarrow \mathbb{K}$ eine Funktion. Die Funktion \det heißt eine Determinante, wenn die folgenden Eigenschaften erfüllt sind:

- 1) Die Funktion \det ist linear in den Zeilen, d.h., sind

$$A = \begin{pmatrix} - & a_1 & - \\ & \vdots & \\ - & a_i & - \\ & \vdots & \\ - & a_n & - \end{pmatrix}, \quad B = \begin{pmatrix} - & a_1 & - \\ & \vdots & \\ - & b_i & - \\ & \vdots & \\ - & a_n & - \end{pmatrix}$$

zwei Matrizen mit Zeilen a_k, b_k , die sich nur in der i -ten Zeile unterscheiden, und sind $\lambda, \mu \in \mathbb{K}$, so ist

$$\det \begin{pmatrix} - & a_1 & - \\ & \vdots & \\ - & \lambda a_i + \mu b_i & - \\ & \vdots & \\ - & a_n & - \end{pmatrix} = \lambda \det \begin{pmatrix} - & a_1 & - \\ & \vdots & \\ - & a_i & - \\ & \vdots & \\ - & a_n & - \end{pmatrix} + \mu \det \begin{pmatrix} - & a_1 & - \\ & \vdots & \\ - & b_i & - \\ & \vdots & \\ - & a_n & - \end{pmatrix}.$$

(Man sagt auch, die Determinante ist multilinear.)

- 2) Hat A zwei gleiche Zeilen, so ist $\det(A) = 0$:

$$\det \begin{pmatrix} & \vdots & \\ - & a_i & - \\ & \vdots & \\ - & a_i & - \\ & \vdots & \end{pmatrix} = 0.$$

- 3) Die Determinante der Einheitsmatrix ist $\det(E_n) = 1$. (Man sagt auch, \det ist normiert.)

BEMERKUNG 1.1.2. Wir werden schon im Zuge von Abschnitt 1 beweisen, dass es auf $\mathbb{K}^{n \times n}$ höchstens eine Determinante gibt. Wir zeigen nämlich, dass — wenn es überhaupt eine Determinante gibt — diese mit Hilfe von Pivottransformationen zu berechnen geht. Und damit ist das Ergebnis dieses Rechenverfahrens natürlich die einzige Determinante.

Erst in Abschnitt 3 werden wir, mit Hilfe einer komplizierten, für die praktische Rechnung wenig nützlichen Formel beweisen können, dass es eine Determinante gibt. Damit ist dann gezeigt, dass es genau eine Determinante gibt.

Erst damit wird die folgende Schreibweise endgültig gerechtfertigt, die wir bereits jetzt verwendet haben: $\det : \mathbb{K}^{n \times n} \rightarrow \mathbb{K}$ bezeichnet die Determinante. Statt $\det(A)$ schreibt man auch oft $|A|$, oder die Matrix A zwischen senkrechten Strichen statt Klammern.

BEMERKUNG 1.1.3. Die Determinante ist linear in den einzelnen Zeilen, aber im Allgemeinen gilt

$$\det(\lambda A) \neq \lambda \det(A), \quad \det(A + B) \neq \det(A) + \det(B).$$

Statt dessen gilt für $A \in \mathbb{K}^{n \times n}$:

$$\det(\lambda A) = \lambda^n \det(A).$$

1.2. Determinante und Zeilenumformungen.

Der folgende Satz zeigt, wie sich eine Determinante bei Anwendung von elementaren Zeilentransformationen verhält.

BEHAUPTUNG 1.2.1. Sei \mathbb{K} ein Körper und $\det : \mathbb{K}^{n \times n} \rightarrow \mathbb{K}$ eine Determinante. Dann gilt:

- 1) Entsteht B aus A , indem eine Zeile von A mit einem Skalar λ multipliziert wird, so ist $\det(B) = \lambda \det(A)$.
- 2) Entsteht B aus A , indem ein Vielfaches einer Zeile von A zu einer anderen Zeile von A addiert wird, so ist $\det(B) = \det(A)$.
- 3) Entsteht B aus A durch Vertauschung von zwei Zeilen, so ist $\det(B) = -\det(A)$. (Man sagt auch, die Determinante ist alternierend.)
- 4) Sind die Zeilen von A linear abhängig, so ist $\det(A) = 0$.

BEWEIS. *Punkt (1):* ist eine unmittelbare Folge von Definition 1.1.1 (1).

Punkt (2): Wir benützen erst Definition 1.1.1 (1), dann Definition 1.1.1 (2):

$$\begin{vmatrix} \vdots & & & & \\ - & a_i + \lambda a_j & - & & \\ \vdots & & & & \\ - & a_j & - & & \\ \vdots & & & & \end{vmatrix} = \begin{vmatrix} \vdots & & & & \\ - & a_i & - & & \\ \vdots & & & & \\ - & a_j & - & & \\ \vdots & & & & \end{vmatrix} + \lambda \begin{vmatrix} \vdots & & & & \\ - & a_j & - & & \\ \vdots & & & & \\ - & a_j & - & & \\ \vdots & & & & \end{vmatrix} = \begin{vmatrix} \vdots & & & & \\ - & a_i & - & & \\ \vdots & & & & \\ - & a_j & - & & \\ \vdots & & & & \end{vmatrix} + 0.$$

Punkt (3) Wir benützen mehrmals den eben bewiesenen Punkt (2), und ganz am Ende die Linearität in den Zeilen:

$$\begin{vmatrix} \vdots & & & & \\ - & a_j & - & & \\ \vdots & & & & \\ - & a_i & - & & \\ \vdots & & & & \end{vmatrix} = \begin{vmatrix} \vdots & & & & \\ - & a_j + a_i & - & & \\ \vdots & & & & \\ - & a_i & - & & \\ \vdots & & & & \end{vmatrix} = \begin{vmatrix} \vdots & & & & \\ - & a_j + a_i & - & & \\ \vdots & & & & \\ - & a_i - (a_j + a_i) & - & & \\ \vdots & & & & \end{vmatrix} \\ = \begin{vmatrix} \vdots & & & & \\ - & a_j + a_i & - & & \\ \vdots & & & & \\ - & -a_j & - & & \\ \vdots & & & & \end{vmatrix} = \begin{vmatrix} \vdots & & & & \\ - & a_i & - & & \\ \vdots & & & & \\ - & -a_j & - & & \\ \vdots & & & & \end{vmatrix} = - \begin{vmatrix} \vdots & & & & \\ - & a_i & - & & \\ \vdots & & & & \\ - & a_j & - & & \\ \vdots & & & & \end{vmatrix}$$

Punkt (4) Sei A eine Matrix mit den Zeilen a_1, \dots, a_n . Sind die Zeilen von A linear abhängig, so gibt es einen Index i und Koeffizienten $\lambda_k \in \mathbb{K}$ (für $k \neq i$), sodass

$$a_i = \sum_{k \neq i} \lambda_k a_k.$$

Wir verwenden nun die Multilinearität und anschließend Definition 1.1.1 (2):

$$\begin{vmatrix} - & a_1 & - \\ & \vdots & \\ - & \sum_{k \neq i} \lambda_k a_k & - \\ & \vdots & \\ - & a_n & - \end{vmatrix} = \sum_{k \neq i} \begin{vmatrix} & \vdots & \\ - & \lambda_k a_k & - \\ & \vdots & \\ - & a_k & - \\ & \vdots & \end{vmatrix} = \sum_{k \neq i} \lambda_k \begin{vmatrix} & \vdots & \\ - & a_k & - \\ & \vdots & \\ - & a_k & - \\ & \vdots & \end{vmatrix} = 0.$$

□

BEMERKUNG 1.2.2. Wir wissen jetzt: Ist \det eine Determinantenfunktion, und ist A eine singuläre Matrix, so ist $\det(A) = 0$. Später werden wir beweisen: A ist genau dann singulär, wenn $\det(A) = 0$.

ALGORITHMUS 1.2.3. Sei \mathbb{K} ein Körper, $\det : \mathbb{K}^{n \times n} \rightarrow \mathbb{K}$ eine Determinantenfunktion. Sei $A \in \mathbb{K}^{n \times n}$. Das folgende Verfahren liefert $\det(A)$:

- (1) Setze $D = 1$. (In D wird am Ende des Algorithmus die gesuchte Determinante stehen.)

Wenn in den folgenden Schritten Zeilentransformationen vorgenommen werden, ist D entsprechend zu ändern:

- Wird das Vielfache einer Zeile zu einer anderen addiert, wird D nicht verändert.
- Wird eine Zeile durch $\lambda \neq 0$ dividiert, so wird D mit λ multipliziert.
- Werden zwei Zeilen vertauscht, so wird das Vorzeichen von D umgekehrt.

- (2) Führe in A Pivotschritte durch, solange man Pivotelemente finden kann.
- (3) Wenn bei der Umformung eine Nullzeile entsteht, so ist $\det(A) = 0$. Stop.
- (4) Wenn bei der Umformung keine Nullzeile entsteht, wurden alle Zeilen als Pivotzeilen verwendet. Ordne die Zeilen durch Vertauschung so um, dass die Pivotelemente auf der Diagonalen stehen. (Hinweis: Die Nicht-Pivotelemente spielen keine weitere Rolle, man darf sie ohne Weiteres durch 0 ersetzen.)
- (5) Multipliziere D mit dem Produkt der Pivotelemente.
- (6) $\det(A) = D$.

BEWEIS. Durch Schritt (2) entsteht aus A eine neue Matrix B . Wegen Behauptung 1.2.1 ist $\det(A) = D \det(B)$.

Ist $\rho(A) < n$, also A singulär, so enthält B eine Nullzeile. In diesem Fall ist wegen Behauptung 1.2.1 (4) die Determinante $\det(A) = 0$.

Ist A regulär, so entsteht keine Nullzeile. Daher kommen alle Zeilen als Pivotzeilen und alle Spalten als Pivotspalten vor. Wir führen in Gedanken — nicht rechnerisch — die Rücksubstitutionsschritte aus, mit denen die Nicht-Pivotelemente auf 0 reduziert werden. Diese ändern die Determinante nicht. Daher dürfen wir alle Nicht-Pivotelemente durch Null ersetzen. Anschließend können wir die Zeilen vertauschen, bis alle Pivotelemente auf der Diagonalen stehen. Es entsteht eine Diagonalmatrix C . Mit jeder Vertauschung ändern wir das Vorzeichen von D , daher gilt nun nach Behauptung 1.2.1 (3) die Determinante $\det(A) = D \det(C)$. Letztlich dividieren

wir (in Gedanken) jede Zeile durch das Pivotelement, das in dieser Zeile steht, und multiplizieren dafür D mit dem Pivotelement. Es entsteht die Einheitsmatrix, und wegen Behauptung 1.2.1 und Definition 1.1.1 (3) ist $\det(A) = D \det(E_n) = D$. \square

BEISPIEL 1.2.4. Berechnen Sie die Determinante

$$\begin{vmatrix} -8 & 3 & 2 & -3 \\ 1 & -5 & -6 & -21 \\ 13 & 6 & -10 & -27 \\ 2 & 4 & 2 & 9 \end{vmatrix}$$

Lösung: Pivotschritte, welche die Determinante nicht verändern:

$$\begin{aligned} & \begin{vmatrix} -8 & 3 & 2 & -3 \\ 1 & -5 & -6 & -21 \\ 13 & 6 & -10 & -27 \\ 2 & 4 & 2 & 9 \end{vmatrix} = \begin{vmatrix} -10 & -1 & 0 & -12 \\ 7 & 7 & 0 & 6* \\ 23 & 26 & 0 & 18 \\ 2 & 4 & 2* & 9 \end{vmatrix} \\ & = \begin{vmatrix} 4 & 13 & 0 & 0 \\ 7 & 7 & 0 & 6* \\ 2* & 5 & 0 & 0 \\ 2 & 4 & 2* & 9 \end{vmatrix} = \begin{vmatrix} 0 & 3* & 0 & 0 \\ 7 & 7 & 0 & 6* \\ 2* & 5 & 0 & 0 \\ 2 & 4 & 2* & 9 \end{vmatrix} \end{aligned}$$

Wir wissen jetzt: Die Matrix hat vollen Rang, ist also regulär. Ab jetzt spielen die Nicht-Pivotelemente keine Rolle mehr, wir interessieren uns nur mehr für die Pivotelemente. Wir vertauschen Zeilen so lange paarweise, bis die Pivotelemente auf der Diagonalen liegen.

$$\begin{vmatrix} 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 6 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \end{vmatrix} = - \begin{vmatrix} 0 & 3 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 2 & 0 \end{vmatrix} = \begin{vmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 2 & 0 \end{vmatrix} = - \begin{vmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \end{vmatrix}$$

Also ist $\det(A) = -2 \cdot 3 \cdot 2 \cdot 6 = -72$. \square

1.3. Dreiecksmatrizen und Blockmatrizen.

DEFINITION 1.3.1. Sei $A = (\alpha_{i,j})_{i,j=1,\dots,n}$ eine quadratische Matrix.

- 1) A heißt obere Dreiecksmatrix, wenn unter der Diagonalen nur Nullen stehen, d.h., falls $i > j$, dann ist $\alpha_{i,j} = 0$.
- 2) A heißt untere Dreiecksmatrix, wenn über der Diagonalen nur Nullen stehen, d.h., falls $i < j$, dann ist $\alpha_{i,j} = 0$.
- 3) A heißt Diagonalmatrix, wenn außerhalb der Diagonalen nur Nullen stehen, d.h. falls $i \neq j$, dann ist $\alpha_{i,j} = 0$.

BEHAUPTUNG 1.3.2. Die Determinante einer Dreiecksmatrix ist das Produkt der Diagonalelemente.

BEWEIS. Wir betrachten die Diagonalelemente einer Dreiecksmatrix als Pivotelemente. Dann sind keine weiteren Pivotschritte nötig. \square

DEFINITION 1.3.3. Sei \mathbb{K} ein Körper, $m_1, m_2, n_1, n_2 \in \mathbb{N}$. Seien

$$\begin{aligned} A &= (\alpha_{i,j})_{i=1,\dots,m_1, j=1,\dots,n_1} \in \mathbb{K}^{m_1 \times n_1}, \\ B &= (\beta_{i,j})_{i=1,\dots,m_1, j=1,\dots,n_2} \in \mathbb{K}^{m_1 \times n_2}, \\ C &= (\gamma_{i,j})_{i=1,\dots,m_2, j=1,\dots,n_1} \in \mathbb{K}^{m_2 \times n_1}, \\ D &= (\delta_{i,j})_{i=1,\dots,m_2, j=1,\dots,n_2} \in \mathbb{K}^{m_2 \times n_2}. \end{aligned}$$

Wir definieren die Blockmatrix

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} := \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n_1} & \beta_{1,1} & \cdots & \beta_{1,n_2} \\ \vdots & & \vdots & \vdots & & \vdots \\ \alpha_{m_1,1} & \cdots & \alpha_{m_1,n_1} & \beta_{m_1,1} & \cdots & \beta_{m_1,n_2} \\ \gamma_{1,1} & \cdots & \gamma_{1,n_1} & \delta_{1,1} & \cdots & \delta_{1,n_2} \\ \vdots & & \vdots & \vdots & & \vdots \\ \gamma_{m_2,1} & \cdots & \gamma_{m_2,n_1} & \delta_{m_2,1} & \cdots & \delta_{m_2,n_2} \end{pmatrix}.$$

Ist $m_1 = n_1$ und $m_2 = n_2$ (also A und D quadratisch) und einer der Blöcke B oder C eine Nullmatrix, so heißt die Blockmatrix eine Blockdreiecksmatrix. Sind A und D quadratisch, und beide anderen Blöcke Nullmatrizen, so heißt die Blockmatrix eine Blockdiagonalmatrix.

BEMERKUNG 1.3.4. Definition 1.3.3 lässt sich analog auch für Matrizen mit mehr als 2×2 Blöcken verallgemeinern.

BEHAUPTUNG 1.3.5. Sei \mathbb{K} ein Körper, seien $A \in \mathbb{K}^{m \times m}$, $D \in \mathbb{K}^{n \times n}$ quadratisch und $B \in \mathbb{K}^{m \times n}$, $C \in \mathbb{K}^{n \times m}$. Seien $\det_m, \det_n, \det_{m+n}$ Determinanten auf $\mathbb{K}^{m \times m}$, $\mathbb{K}^{n \times n}$ und $\mathbb{K}^{(m+n) \times (m+n)}$. (Wir werden die Indizes $m, n, m+n$ weglassen, aus dem Zusammenhang ist klar, welche der drei Determinanten jeweils gemeint ist.) Dann ist

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det \begin{pmatrix} A & 0 \\ C & D \end{pmatrix} = \det(A) \det(D).$$

BEWEIS. Der Kürze halber schreiben wir

$$T := \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

Sei A singular. Dann sind die Spalten von A , und folglich die ersten m Spalten von T , linear abhängig, und damit ist die Blockmatrix singular. In diesem Fall ist $\det(T) = 0$ und auch $\det(A) \det(D) = 0 \cdot \det(D) = 0$.

Sei D singular. Dann sind die Zeilen von D , und folglich die letzten n Zeilen von T , linear abhängig, und daher ist T singular. Wieder folgt $0 = \det(T) = \det(A) \det(D)$. Nun seien sowohl A als auch D regulär. Durch Pivotschritte und Zeilenvertauschungen lässt sich A in eine Diagonalmatrix A_1 umwandeln. Wir führen dieselben Schritte mit den ersten m Zeilen von T durch und erhalten eine Matrix

$$\tilde{T} = \begin{pmatrix} A_1 & B_1 \\ 0 & D \end{pmatrix}.$$

Auch D lässt sich durch Zeilentransformationen in eine Diagonalmatrix D_1 umwandeln. Dieselben Transformationen, angewendet auf die letzten n Zeilen von \tilde{T} , ergeben eine Matrix

$$T_1 = \begin{pmatrix} A_1 & B_1 \\ 0 & D_1 \end{pmatrix}.$$

Sei nun a das Produkt der Diagonalelemente von A_1 und d das Produkt der Diagonalelemente D_1 , sei p die Anzahl der Zeilenvertauschungen unter den Transformationen, die A in A_1 umwandeln, und q die Anzahl der Zeilenvertauschungen bei der Umwandlung von D in D_1 . Es ist dann ad das Produkt der Diagonalelemente

von T_1 und $p+q$ die Anzahl der Zeilenvertauschungen bei der Umwandlung von T in T_1 . Dann ist

$$\det(A) = (-1)^p a, \quad \det(D) = (-1)^q d, \quad \det(T) = (-1)^{p+q} ad = \det(A) \det(D).$$

Der Beweis für die andere Blockdreiecksmatrix geht ebenso. \square

1.4. Die Eindeutigkeit der Determinante.

SATZ 1.4.1. *Auf $\mathbb{K}^{n \times n}$ gibt es höchstens eine Determinante.*

BEWEIS. Wenn es überhaupt eine Determinante gibt, ist sie das eindeutige Ergebnis der Berechnung durch Algorithmus 1.2.3. \square

2. Multiplikationssatz für Determinanten

2.1. Determinanten von Elementarmatrizen.

LEMMA 2.1.1. *Sei \mathbb{K} ein Körper, und seien $T_{\text{tausch}}(i, j)$, $T_{\text{add}}(i, j, \lambda)$, $T_{\text{mult}}(i, \lambda)$ die $n \times n$ -Elementarmatrizen aus Definition 6.5.1. Sei \det eine Determinante auf $\mathbb{K}^{n \times n}$. Dann gilt für alle Matrizen $A \in \mathbb{K}^{n \times n}$:*

$$\begin{aligned} \det(T_{\text{tausch}}(i, j)A) &= -\det(A), \\ \det(T_{\text{add}}(i, j, \lambda)A) &= \det(A), \\ \det(T_{\text{mult}}(i, \lambda)A) &= \lambda \det(A). \end{aligned}$$

Insbesondere ist

$$\det(T_{\text{tausch}}(i, j)) = -1, \quad \det(T_{\text{add}}(i, j, \lambda)) = 1, \quad \det(T_{\text{mult}}(i, \lambda)) = \lambda,$$

und damit gilt für jede Elementarmatrix T und jede Matrix A

$$\det(TA) = \det(T) \det(A).$$

BEWEIS. Die Linksmultiplikation mit Elementarmatrizen ist eine elementare Zeilenumformung. Daher ist das Lemma nichts anderes als eine Umformulierung von Behauptung 1.2.1. Die Determinanten der Elementarmatrizen ergeben sich, wenn man für A die Einheitsmatrix $A = E_n$ einsetzt. \square

LEMMA 2.1.2. *Jede reguläre Matrix in $\mathbb{K}^{n \times n}$ lässt sich als Produkt von Elementarmatrizen schreiben.*

BEWEIS. Ist A regulär, so haben wir im ersten Kapitel in Algorithmus 6.5.6 gesehen, dass sich A durch elementare Umformungen auf die Einheitsmatrix reduzieren lässt, also $E_n = T_m T_{m-1} \cdots T_1 A$, d.h., $A = T_1^{-1} \cdots T_m^{-1} E_n$. Da die Umkehrungen von elementaren Zeilentransformationen wieder elementare Zeilentransformationen sind, sind die Inversen von Elementarmatrizen wieder Elementarmatrizen. \square

2.2. Der Multiplikationssatz.

SATZ 2.2.1. Sei \mathbb{K} ein Körper und \det eine Determinantenfunktion auf $\mathbb{K}^{n \times n}$. Seien $A, B \in \mathbb{K}^{n \times n}$. Dann ist

$$\det(AB) = \det(A) \cdot \det(B).$$

BEWEIS. Ist A singular, so ist auch AB singular. Also gilt dann nach Behauptung 1.2.1 (4)

$$\det(A) = \det(AB) = 0.$$

Insbesondere ist also

$$\det(AB) = 0 = 0 \det(B) = \det(A) \det(B).$$

Ist A regulär, so lässt sich A nach Lemma 2.1.2 in der Form $A = T_1 \cdots T_m$ mit Elementarmatrizen T_k schreiben. Nach Lemma 2.1.1 gilt $\det(T_i C) = \det(T_i) \det(C)$ für jedes $C \in \mathbb{K}^{n \times n}$. Damit gilt (mittels vollständiger Induktion nach m):

$$\begin{aligned} \det(A) &= \det(T_1 \cdots T_m E_n) = \prod_{i=1}^m \det(T_i) \cdot 1, \\ \det(AB) &= \det(T_1 \cdots T_m B) = \prod_{i=1}^m \det(T_i) \cdot \det(B). \end{aligned}$$

Also ist $\det(AB) = \det(A) \det(B)$. \square

BEHAUPTUNG 2.2.2. Sei \det eine Determinante auf $\mathbb{K}^{n \times n}$ und $A \in \mathbb{K}^{n \times n}$. Dann hat die Transponierte A^T dieselbe Determinante wie A :

$$\det(A^T) = \det(A).$$

BEWEIS. Ist A singular, also $\rho(A) < n$, so ist auch $\rho(A^T) = \rho(A) < n$. Daher ist auch A^T singular, und beide Determinanten sind Null: $\det(A) = \det(A^T) = 0$. Ist A regulär, so ist nach Lemma 2.1.2 $A = T_1 \cdots T_m$ mit Elementarmatrizen T_i . Man sieht leicht, dass die Transponierten der Elementarmatrizen wieder Elementarmatrizen mit derselben Determinante sind. Also ist

$$\det(A^T) = \det(T_m^T \cdots T_1^T) = \prod_{i=1}^m \det(T_i^T) = \prod_{i=1}^m \det(T_i) = \det(A).$$

\square

KOROLLAR 2.2.3. Algorithmus 1.2.3 zur Berechnung der Determinante funktioniert ebenso, wenn man statt Zeilenumformungen Spaltenumformungen verwendet.

BEWEIS. Spaltenumformungen an A entsprechen Zeilenumformungen an A^T . \square

2.3. Determinante als Kriterium für Regularität.

SATZ 2.3.1. Sei \det eine Determinante auf $\mathbb{K}^{n \times n}$ und $A \in \mathbb{K}^{n \times n}$. Es gilt: A ist genau dann regulär, wenn $\det A \neq 0$. In diesem Fall ist

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

BEWEIS. Ist A singular, so sind die Zeilen von A linear abhängig, folglich ist nach Behauptung 1.2.1 die Determinante $\det(A) = 0$. Ist A regulär, so ist nach Behauptung 2.2.1

$$1 = \det(E_n) = \det(AA^{-1}) = \det(A) \det(A^{-1}).$$

\square

3. Existenz der Determinante

3.1. Permutationen.

DEFINITION 3.1.1. Sei $n \in \mathbb{N}$ und $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ eine bijektive Abbildung. Dann heißt σ eine Permutation auf der Menge $\{1, \dots, n\}$. Mit \mathcal{S}_n bezeichnen wir die Menge aller Permutationen auf $\{1, \dots, n\}$.

SCHREIBWEISE 3.1.2. Eine Permutation $\sigma \in \mathcal{S}_n$ können wir anschreiben, in dem wir in der ersten Zeile die Zahlen $1, \dots, n$, und darunter die Werte $\sigma(1), \dots, \sigma(n)$ anschreiben. Es ist üblich, diesen Ausdruck in Klammern zusammenzufassen. Aus dem Zusammenhang müssen Sie erkennen, dass dies keine Matrix, sondern die Beschreibung einer Permutation ist.

BEISPIEL 3.1.3. Die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 3 & 1 \end{pmatrix} \in \mathcal{S}_4$$

bildet folgendermaßen ab:

$$\sigma(1) = 2, \quad \sigma(2) = 4, \quad \sigma(3) = 3, \quad \sigma(4) = 1.$$

Sei

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

Berechnen Sie die Hintereinanderausführungen $\sigma \circ \tau$ und $\tau \circ \sigma$, und σ^{-1} .

Lösung: Wir berechnen nacheinander $[\sigma \circ \tau](i)$ für $i = 1, 2, 3, 4$ und tragen die Werte in die Beschreibung von $\sigma \circ \tau$ ein. Es ist $\tau(1) = 4$, also $[\sigma \circ \tau](1) = \sigma(4) = 1$. Es ist $\tau(2) = 3$, daher $[\sigma \circ \tau](2) = \sigma(3) = 3$, u.s.w.:

$$\begin{aligned} \sigma \circ \tau &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 2 \end{pmatrix}, \\ \tau \circ \sigma &= \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}. \end{aligned}$$

Zur Berechnung von σ^{-1} beachten wir zunächst: $\sigma(1) = 2$, daher ist $\sigma^{-1}(2) = 1$. Wir tragen in der Beschreibung von σ^{-1} den Wert 1 unter der Stelle 2 ein, u.s.w.:

$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix}.$$

□

SATZ 3.1.4. Sei $n \in \mathbb{N}$.

- 1) \mathcal{S}_n bildet mit der Hintereinanderausführung eine Gruppe, die sogenannte symmetrische Gruppe. Dabei ist das neutrale Element die Identität id , das inverse Element zu $\sigma \in \mathcal{S}_n$ die Umkehrabbildung σ^{-1} .
- 2) Ist $n \geq 3$, dann ist \mathcal{S}_n nicht kommutativ.
- 3) \mathcal{S}_n enthält $n! := n \cdot (n-1) \cdot \dots \cdot 2 \cdot 1$ Elemente. (Die Zahl $n!$ bezeichnet man als n faktorielle.)

BEWEIS. Punkt (1): Da die Hintereinanderausführung von bijektiven Abbildungen wieder bijektiv ist, ist \circ eine innere Verknüpfung auf \mathcal{S}_n . Bekanntlich ist die Hintereinanderausführung von Funktionen assoziativ. Es ist $\text{id} \circ \sigma = \sigma \circ \text{id} = \sigma$, daher ist id das neutrale Element. Für die Umkehrabbildung σ^{-1} gilt $\sigma \circ \sigma^{-1} = \sigma^{-1} \circ \sigma = \text{id}$, also ist σ^{-1} das inverse Element zu σ .

Punkt (2) Es ist

$$\begin{pmatrix} 1 & 2 & 3 & 4 \cdots \\ 2 & 1 & 3 & 4 \cdots \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \cdots \\ 1 & 3 & 2 & 4 \cdots \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \cdots \\ 2 & 3 & 1 & 4 \cdots \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 \cdots \\ 1 & 3 & 2 & 4 \cdots \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 & 4 \cdots \\ 2 & 1 & 3 & 4 \cdots \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \cdots \\ 3 & 1 & 2 & 4 \cdots \end{pmatrix}.$$

Punkt (3): Wenn wir eine Permutation σ auf $\{1, \dots, n\}$ zusammenbauen, haben wir zunächst n Möglichkeiten, das Bild $\sigma(1)$ festzulegen. Nun ist einer der Werte aus $1 \cdots n$ vergeben, und wir haben noch $n-1$ Möglichkeiten, das Bild $\sigma(2)$ festzulegen. Wir haben also bisher $n(n-1)$ verschiedene Möglichkeiten, $\sigma(1)$ und $\sigma(2)$ zu fixieren. Es bleiben $n-2$ Möglichkeiten für $\sigma(3)$, u.s.w. \square

DEFINITION 3.1.5.

- 1) Sei $\sigma \in \mathcal{S}_n$ eine Permutation. Ein Fehlstand von σ ist ein Paar (i, j) mit $i < j \in \{1, \dots, n\}$ so, dass $\sigma(j) < \sigma(i)$.
- 2) Sei m die Anzahl der Fehlstände von σ . Dann ist das Vorzeichen oder Signum von σ

$$\operatorname{sgn}(\sigma) := (-1)^m.$$

- 3) Ist $\operatorname{sgn}(\sigma) = 1$, so heißt σ eine gerade Permutation, sonst heißt σ eine ungerade Permutation.

BEISPIEL 3.1.6. Welche Fehlstände und welches Vorzeichen hat die Permutation

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

Lösung: Fehlstände sind überall da zu finden, wo in der zweiten Zeile eine größere Zahl vor einer kleineren steht: Es steht:

5 vor den kleineren Zahlen 3, 1, 4, 2. Daher sind die Paare $(1, 2)$, $(1, 3)$, $(1, 5)$ und $(1, 6)$ Fehlstände. Das sind einmal 4 Fehlstände.

3 vor 1 und 2, das gibt 2 weitere Fehlstände.

1 vor keiner kleineren Zahl, kein weiterer Fehlstand.

6 vor 4 und 2, gibt 2 weitere Fehlstände.

4 vor 2, noch ein Fehlstand.

Insgesamt haben wir 9 Fehlstände, eine ungerade Zahl. Also ist $\operatorname{sgn}(\tau) = -1$, und τ ist eine ungerade Permutation. \square

DEFINITION 3.1.7. Sei $1 \leq i < j \leq n$. Die Permutation

$$\tau(k) := \begin{cases} j & \text{wenn } k = i, \\ i & \text{wenn } k = j, \\ k & \text{wenn } k \notin \{i, j\}, \end{cases}$$

heißt eine Vertauschung. Ist insbesondere $j = i + 1$, so heißt τ eine Nachbarvertauschung oder Transposition.

LEMMA 3.1.8. Jede Vertauschung ist ungerade.

BEWEIS. Wir suchen die Fehlstände in der Vertauschung

$$\tau = \begin{pmatrix} \cdots & (i-1) & i & (i+1) & \cdots & (j-1) & j & (j+1) & \cdots \\ \cdots & (i-1) & j & (i+1) & \cdots & (j-1) & i & (j+1) & \cdots \end{pmatrix}.$$

Es steht in der zweiten Zeile j vor den kleineren Zahlen $(i+1), \dots, (j-1), i$, das gibt $j-i$ Fehlstände.

Es steht in der zweiten Zeile i nach den größeren Zahlen $j, (i+1), \dots, (j-1)$. Das

wären wieder $j - i$ Fehlstände, aber der Fehlstand (i, j) wurde schon oben gezählt. τ hat daher insgesamt $2(j - i) - 1$ Fehlstände und ist daher ungerade. \square

SATZ 3.1.9.

- 1) Sei $\tau \in \mathcal{S}_n$ eine Transposition, und $\sigma \in \mathcal{S}_n$. Dann hat $\sigma \circ \tau$ entweder genau einen Fehlstand mehr oder einen Fehlstand weniger als σ .
- 2) Sei $\sigma \in \mathcal{S}_n$ eine Permutation mit m Fehlständen. Dann gibt es (nicht unbedingt eindeutig bestimmte) Transpositionen τ_1, \dots, τ_m so, dass

$$\sigma = \tau_1 \circ \dots \circ \tau_m \circ \text{id}.$$

- 3) Lässt sich eine Permutation $\sigma \in \mathcal{S}_n$ als Hintereinanderausführung von m Transpositionen schreiben, so ist m ungerade, falls σ ungerade ist, und m ist gerade, falls σ gerade ist.
- 4) Sind $\sigma, \rho \in \mathcal{S}_n$, so ist

$$\text{sgn}(\sigma \circ \rho) = \text{sgn}(\sigma) \cdot \text{sgn}(\rho).$$

BEWEIS. *Punkt (1):* Sei τ die Nachbarvertauschung von i und $i + 1$. Es ist

$$\sigma = \begin{pmatrix} 1 & \dots & i & (i+1) & \dots & n \\ \sigma(1) & \dots & \sigma(i) & \sigma(i+1) & \dots & \sigma(n) \end{pmatrix},$$

$$\sigma \circ \tau = \begin{pmatrix} 1 & \dots & i & (i+1) & \dots & n \\ \sigma(1) & \dots & \sigma(i+1) & \sigma(i) & \dots & \sigma(n) \end{pmatrix}.$$

Im Vergleich zu σ haben in der unteren Zeile nur die zwei benachbarten Elemente $\sigma(i)$ und $\sigma(i+1)$ Platz getauscht. Daher ändert sich in der Stellung der Elemente der unteren Zeile zueinander nur das Paar $(\sigma(i), \sigma(i+1))$. War $(i, i+1)$ kein Fehlstand in σ , so ist es nun in $\sigma \circ \tau$ ein Fehlstand, und umgekehrt. Es gibt also einen Fehlstand mehr, oder einen Fehlstand weniger.

Punkt (2): Vollständige Induktion nach der Anzahl der Fehlstände m . Ist $m = 0$, so ist σ die Identität, und es braucht keine Transpositionen. Wir nehmen nun an, dass die Behauptung für Permutationen mit $m - 1$ Fehlständen gilt. Sei σ eine Permutation mit $m > 0$ Fehlständen. Dann gibt es ein i sodass $\sigma(i) > \sigma(i + 1)$. Sonst müsste für alle $i < j$ gelten: $\sigma(i) < \sigma(i + 1) < \dots < \sigma(j)$, und es könnte gar keinen Fehlstand geben. Sei nun τ_m die Vertauschung des Paares $(i, i + 1)$. Dann hat $\sigma \circ \tau_m$ einen Fehlstand weniger als σ , denn nun ist das Paar $(i, i + 1)$ kein Fehlstand mehr. Nach Induktionsannahme gibt es Transpositionen $\tau_1, \dots, \tau_{m-1}$ so, dass

$$\sigma \circ \tau_m = \tau_1 \circ \dots \circ \tau_{m-1}.$$

Beachten wir noch, dass für jede Transposition τ gilt $\tau \circ \tau = \text{id}$, dann ist

$$\sigma = \sigma \circ \tau_m \circ \tau_m = \tau_1 \circ \dots \circ \tau_m.$$

Punkt (3): Wegen Punkt (1) ist für jede Permutation σ und jede Transposition τ

$$\text{sgn}(\sigma \circ \tau) = -\text{sgn}(\sigma).$$

Daher ist

$$\text{sgn}(\tau_1 \circ \dots \circ \tau_m) = (-1)^m.$$

Punkt (4): Wegen Punkt (2) lassen sich σ und ρ als Hintereinanderausführung von Transpositionen schreiben:

$$\sigma = \tau_1 \circ \dots \circ \tau_m, \quad \rho = \tilde{\tau}_1 \circ \dots \circ \tilde{\tau}_k.$$

Dann ist

$$\sigma \circ \rho = \tau_1 \circ \dots \circ \tau_m \circ \tilde{\tau}_1 \circ \dots \circ \tilde{\tau}_k$$

eine Hintereinanderausführung von $m + k$ Transpositionen. Nach Punkt (3) ist

$$\operatorname{sgn}(\sigma) = (-1)^m, \quad \operatorname{sgn}(\rho) = (-1)^k, \quad \operatorname{sgn}(\sigma \circ \rho) = (-1)^{k+m} = (-1)^m \cdot (-1)^k.$$

□

3.2. Die Formel von Leibniz.

SATZ 3.2.1. Sei \mathbb{K} ein Körper und $n \in \mathbb{N}$. Die folgende Funktion $\det : \mathbb{K}^{n \times n} \rightarrow \mathbb{K}$ ist eine Determinante.

$$\det \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n} \end{pmatrix} := \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \alpha_{1,\sigma(1)} \alpha_{2,\sigma(2)} \cdots \alpha_{n,\sigma(n)}.$$

BEWEIS. *Linearität in den Zeilen:*

$$\begin{aligned} & \det \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \lambda \alpha_{i,1} + \mu \beta_{i,1} & \cdots & \lambda \alpha_{i,n} + \mu \beta_{i,n} \\ \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n} \end{pmatrix} \\ &= \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \alpha_{1,\sigma(1)} \cdots (\lambda \alpha_{i,\sigma(i)} + \mu \beta_{i,\sigma(i)}) \cdots \alpha_{n,\sigma(n)} \\ &= \lambda \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \alpha_{1,\sigma(1)} \cdots \alpha_{i,\sigma(i)} \cdots \alpha_{n,\sigma(n)} + \mu \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \alpha_{1,\sigma(1)} \cdots \beta_{i,\sigma(i)} \cdots \alpha_{n,\sigma(n)} \\ &= \lambda \det \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{i,1} & \cdots & \alpha_{i,n} \\ \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n} \end{pmatrix} + \mu \det \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \beta_{i,1} & \cdots & \beta_{i,n} \\ \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n} \end{pmatrix}. \end{aligned}$$

Matrizen mit zwei gleichen Zeilen haben Determinante 0: Seien die Zeilen mit Indices i und j der Matrix $(\alpha_{i,j})$ gleich. Sei τ die Vertauschung

$$\tau = \begin{pmatrix} 1 & \cdots & i & \cdots & j & \cdots & n \\ 1 & \cdots & j & \cdots & i & \cdots & n \end{pmatrix}$$

Nach Lemma 3.1.8 ist $\operatorname{sgn}(\tau) = -1$. Ist \mathcal{A}_n die Menge der geraden Permutationen, so erhält man jede ungerade Permutation ρ durch $\rho = \sigma \circ \tau$ mit einer eindeutigen

geraden Permutation $\sigma \in \mathcal{A}_n$.

$$\begin{aligned}
 & \det \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,n} \end{pmatrix} \\
 &= \sum_{\rho \in \mathcal{S}_n} \operatorname{sgn}(\rho) \alpha_{1,\rho(1)} \cdots \alpha_{i,\rho(i)} \cdots \alpha_{j,\rho(j)} \cdots \alpha_{n,\rho(n)} \\
 &= \sum_{\rho \in \mathcal{A}_n} \alpha_{1,\rho(1)} \cdots \alpha_{i,\rho(i)} \cdots \alpha_{j,\rho(j)} \cdots \alpha_{n,\rho(n)} \\
 &\quad - \sum_{\rho \in \mathcal{A}_n} \alpha_{1,\rho \circ \tau(1)} \cdots \alpha_{i,\rho \circ \tau(i)} \cdots \alpha_{j,\rho \circ \tau(j)} \cdots \alpha_{n,\rho \circ \tau(n)} \\
 &\quad \text{(Wir nützen jetzt die Gleichheit der Zeilen } i \text{ und } j \text{)} \\
 &= \sum_{\rho \in \mathcal{A}_n} \alpha_{1,\rho(1)} \cdots \alpha_{i,\rho(i)} \cdots \alpha_{j,\rho(j)} \cdots \alpha_{n,\rho(n)} \\
 &\quad - \sum_{\rho \in \mathcal{A}_n} \alpha_{1,\rho \circ \tau(1)} \cdots \alpha_{j,\rho \circ \tau(i)} \cdots \alpha_{i,\rho \circ \tau(i)} \cdots \alpha_{n,\rho \circ \tau(n)} \\
 &\quad \text{(Es ist } i = \tau(j), j = \tau(i) \text{ und } k = \tau(k) \text{ sonst:)} \\
 &= \sum_{\rho \in \mathcal{A}_n} \alpha_{1,\rho(1)} \cdots \alpha_{i,\rho(i)} \cdots \alpha_{j,\rho(j)} \cdots \alpha_{n,\rho(n)} \\
 &\quad - \sum_{\rho \in \mathcal{A}_n} \alpha_{1,\rho(1)} \cdots \alpha_{j,\rho(j)} \cdots \alpha_{i,\rho(i)} \cdots \alpha_{n,\rho(n)} \\
 &= 0.
 \end{aligned}$$

Determinante der Einheitsmatrix ist Eins: Die Koeffizienten der Einheitsmatrix sind $\delta_{i,j}$, dabei bedeutet δ das Kronecker-Delta:

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

Nun gibt es für jede Permutation $\sigma \neq \operatorname{id}$ einen Index i mit $\sigma(i) \neq i$. Damit ist

$$\operatorname{sgn}(\sigma) \delta_{1,\sigma(1)} \cdots \delta_{i,\sigma(i)} \cdots \delta_{n,\sigma(n)} = 0.$$

Nur für die Identität ist

$$\operatorname{sgn}(\operatorname{id}) \delta_{1,1} \cdots \delta_{n,n} = 1.$$

Daher ist

$$\det(E_n) = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) \delta_{1,\sigma(1)} \cdots \delta_{i,\sigma(i)} \cdots \delta_{n,\sigma(n)} = 1.$$

□

SATZ 3.2.2. *Es gibt genau eine Determinante $\mathbb{K}^{n \times n} \rightarrow \mathbb{K}$.*

BEWEIS. Durch die Leibnizsche Formel wird eine Determinante definiert, daher gibt es mindestens eine Determinante. Wenn \det eine Determinante ist, lässt sie sich mit Algorithmus 1.2.3 berechnen. Daher kann es höchstens eine Determinante geben. □

KOROLLAR 3.2.3. *Für 2×2 -Matrizen gilt:*

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

KOROLLAR 3.2.4 (Regel von Sarrus). *Um die Determinante einer 3×3 Matrix zu berechnen, schreibt man die ersten zwei Zeilen unter der Matrix noch einmal an. Man addiert dann die drei Produkte, die sich von links oben nach rechts unten ergeben, und subtrahiert die drei Produkte, die sich von rechts oben nach links unten ergeben:*

$$\begin{array}{cccccc}
 a_{1,1} & a_{1,2} & a_{1,3} & a_{1,1} & a_{1,2} & a_{1,3} \\
 & \searrow & & & & \swarrow \\
 a_{2,1} & a_{2,2} & a_{2,3} & a_{2,1} & a_{2,2} & a_{2,3} \\
 & \searrow & \searrow & & \swarrow & \swarrow \\
 a_{3,1} & a_{3,2} & a_{3,3} & a_{3,1} & a_{3,2} & a_{3,3} \\
 & \searrow & \searrow & & \swarrow & \swarrow \\
 a_{1,1} & a_{1,2} & a_{1,3} & a_{1,1} & a_{1,2} & a_{1,3} \\
 & & \searrow & & \swarrow & \\
 a_{2,1} & a_{2,2} & a_{2,3} & a_{2,1} & a_{2,2} & a_{2,3}
 \end{array}$$

$$[a_{1,1}a_{2,2}a_{3,3} + a_{2,1}a_{3,2}a_{1,3} + a_{3,1}a_{1,2}a_{2,3}] - [a_{1,3}a_{2,2}a_{3,1} + a_{2,3}a_{3,2}a_{1,1} + a_{3,3}a_{1,2}a_{2,1}].$$

4. Entwicklungssatz von Laplace

4.1. Entwicklungssatz von Laplace.

DEFINITION 4.1.1. Sei \mathbb{K} ein Körper, $A = (\alpha_{i,j}) \in \mathbb{K}^{n \times n}$ eine quadratische Matrix. Für jedes Indexpaar (i, j) entsteht die Matrix $A'_{i,j}$ durch Streichen der i -ten Zeile und der j -ten Spalte:

$$A'_{i,j} := \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,j-1} & \alpha_{1,j+1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \alpha_{i-1,1} & \cdots & \alpha_{i-1,j-1} & \alpha_{i-1,j+1} & \cdots & \alpha_{i-1,n} \\ \alpha_{i+1,1} & \cdots & \alpha_{i+1,j-1} & \alpha_{i+1,j+1} & \cdots & \alpha_{i+1,n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,j-1} & \alpha_{n,j+1} & \cdots & \alpha_{n,n} \end{pmatrix}$$

Wir definieren die komplementäre Matrix $A^\# \in \mathbb{K}^{n \times n}$ zu A :

$$A^\# = \begin{pmatrix} \alpha_{1,1}^\# & \cdots & \alpha_{1,n}^\# \\ \vdots & & \vdots \\ \alpha_{n,1}^\# & \cdots & \alpha_{n,n}^\# \end{pmatrix} \text{ mit}$$

$$\alpha_{i,j}^\# := (-1)^{i+j} \det(A'_{j,i}).$$

(Achtung auf die Umkehrung der Reihenfolge der Indizes in der obigen Formel bei $\det(A'_{j,i})$!)

BEMERKUNG 4.1.2 (Schachbrettregel). Die Vorzeichen $(-1)^{i+j}$ liegen wie ein Schachbrett auf der Matrix. Auf der Diagonalen liegt immer Plus:

$$\begin{pmatrix} + & - & + & \cdots & \pm \\ - & + & - & \cdots & \mp \\ + & - & + & \cdots & \pm \\ \vdots & \vdots & \vdots & & \vdots \\ \pm & \mp & \pm & \cdots & + \end{pmatrix}.$$

SATZ 4.1.3. Sei \mathbb{K} ein Körper und $A = (\alpha_{i,j}) \in \mathbb{K}^{n \times n}$.

1) ("Entwicklung nach einer Zeile":) Sei $i \in \{1, \dots, n\}$. Dann ist

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} \alpha_{i,j} \det(A'_{i,j}).$$

2) ("Entwicklung nach einer Spalte"): Sei $j \in \{1, \dots, n\}$. Dann ist

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} \alpha_{i,j} \det(A'_{i,j}).$$

BEWEIS. Wir zeigen die Entwicklung nach einer Zeile. Die Entwicklung nach einer Spalte folgt dann aus $\det(A) = \det(A^T)$.

Sei zunächst τ_i die Permutation

$$\tau_i = \begin{pmatrix} 1 & 2 & \cdots & i-1 & i & i+1 & \cdots & n-1 & n \\ i & 1 & \cdots & i-2 & i-1 & i+1 & \cdots & n-1 & n \end{pmatrix}$$

Offensichtlich sind die Fehlstände von τ_i die Paare $(1, k)$ mit $k = 2, \dots, i$. Das sind $i-1$ Fehlstände. Daher ist

$$\operatorname{sgn}(\tau_i) = (-1)^{i-1}.$$

Wir betrachten nun die Matrix

$$B_{i,j} := \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,j-1} & \alpha_{1,j} & \alpha_{1,j+1} & \cdots & \alpha_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \alpha_{i-1,1} & \cdots & \alpha_{i-1,j-1} & \alpha_{i-1,j} & \alpha_{i-1,j+1} & \cdots & \alpha_{i-1,n} \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \alpha_{i+1,1} & \cdots & \alpha_{i+1,j-1} & \alpha_{i+1,j} & \alpha_{i+1,j+1} & \cdots & \alpha_{i+1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \alpha_{n,1} & \cdots & \alpha_{n,j-1} & \alpha_{n,j} & \alpha_{n,j+1} & \cdots & \alpha_{n,n} \end{pmatrix}.$$

Wir tauschen die j -te Spalte ans linke Ende, und die i -te Zeile ans obere Ende der Matrix. Es ergibt sich eine Blockdreiecksmatrix, wenn man die erste Zeile und erste Spalte als einen 1×1 -Block, und die anderen Zeilen und Spalten als einen $(n-1) \times (n-1)$ -Block ansieht:

$$\det(B_{i,j}) = \operatorname{sgn}(\tau_j) \operatorname{sgn}(\tau_i) \begin{vmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \alpha_{1,j} & \alpha_{1,1} & \cdots & \alpha_{1,j-1} & \alpha_{1,j+1} & \cdots & \alpha_{1,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \alpha_{i-1,j} & \alpha_{i-1,1} & \cdots & \alpha_{i-1,j-1} & \alpha_{i-1,j+1} & \cdots & \alpha_{i-1,n} \\ \alpha_{i+1,j} & \alpha_{i+1,1} & \cdots & \alpha_{i+1,j-1} & \alpha_{i+1,j+1} & \cdots & \alpha_{i+1,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \alpha_{n,j} & \alpha_{n,1} & \cdots & \alpha_{n,j-1} & \alpha_{n,j+1} & \cdots & \alpha_{n,n} \end{vmatrix}.$$

Weil $\operatorname{sgn}(\tau_i) \operatorname{sgn}(\tau_j) = (-1)^{i-1} (-1)^{j-1} = (-1)^{i+j}$, und nach Behauptung 1.3.5, ergibt sich nun

$$\det(B_{i,j}) = (-1)^{i+j} \det(A'_{i,j}).$$

Nun sind wir in der Lage, den Entwicklungssatz fertig zu beweisen. Sind e_1^T, \dots, e_n^T die Einheits-Zeilenvektoren, so lässt sich die i -te Zeile von A schreiben:

$$\alpha_{i,1} e_1^T + \cdots + \alpha_{i,n} e_n^T.$$

Wegen der Linearität in der i -ten Zeile ist

$$\det(A) = \sum_{j=1}^n \alpha_{i,j} \det(B_{i,j}) = \sum_{j=1}^n (-1)^{i+j} \alpha_{i,j} \det(A'_{i,j}).$$

□

BEISPIEL 4.1.4. Der Laplacesche Entwicklungssatz eignet sich für Determinanten in Matrizen, die viele Nullen enthalten. Zum Beispiel berechnen wir die Determinante

$$\begin{vmatrix} \lambda - 4 & 0 & 1 & 2 \\ -2 & \lambda + 3 & 3 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 1 & 1 & \lambda - 2 \end{vmatrix}.$$

Lösung: Wir entwickeln zunächst nach der dritten Zeile, anschließend die beiden 3×3 -Determinanten nach der ersten Spalte. Die 2×2 -Determinanten lösen wir mittels Korollar 3.2.3 auf.

$$\begin{aligned} & \begin{vmatrix} \lambda - 4 & 0 & 1 & 2 \\ -2 & \lambda + 3 & 3 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 1 & 1 & \lambda - 2 \end{vmatrix} = \lambda \begin{vmatrix} \lambda - 4 & 0 & 2 \\ -2 & \lambda + 3 & 0 \\ 0 & 1 & \lambda - 2 \end{vmatrix} - 1 \begin{vmatrix} \lambda - 4 & 0 & 1 \\ -2 & \lambda + 3 & 3 \\ 0 & 1 & 1 \end{vmatrix} \\ &= \lambda(\lambda - 4) \begin{vmatrix} \lambda + 3 & 0 \\ 1 & \lambda - 2 \end{vmatrix} + 2\lambda \begin{vmatrix} 0 & 2 \\ 1 & \lambda - 2 \end{vmatrix} - (\lambda - 4) \begin{vmatrix} \lambda + 3 & 3 \\ 1 & 1 \end{vmatrix} - 2 \begin{vmatrix} 0 & 1 \\ 1 & 1 \end{vmatrix} \\ &= \lambda(\lambda - 4)(\lambda + 3)(\lambda - 2) - 4\lambda - (\lambda - 4)\lambda + 2 \\ &= \lambda^4 - 3\lambda^3 - 11\lambda^2 + 24\lambda + 2. \end{aligned}$$

□

4.2. Die Komplementärmatrix und die Inverse.

SATZ 4.2.1. Sei \mathbb{K} ein Körper, $A \in \mathbb{K}^{n \times n}$ und $A^\#$ die Komplementärmatrix zu A . Dann ist

$$AA^\# = A^\#A = \det(A) E_n.$$

Insbesondere gilt für jede reguläre Matrix A

$$A^{-1} = \frac{1}{\det(A)} A^\#.$$

BEWEIS. Wir zeigen $AA^\# = \det(A) E_n$. Das Produkt $A^\#A = \det(A) E_n$ geht genauso.

Sei also $A = (\alpha_{i,j})$, $A^\# = (\alpha_{i,j}^\#)$ und $AA^\# = (\gamma_{i,j})$. Wir bilden eine Matrix \tilde{A} , indem wir die j -te Zeile von A durch die i -te Zeile ersetzen. Für $i = j$ bleibt $\tilde{A} = A$, für $i \neq j$ kommt die i -te Zeile von A in der Matrix \tilde{A} zweimal vor. Daher ist

$$\det(\tilde{A}) = \begin{cases} \det(A) & \text{wenn } i = j \\ 0 & \text{wenn } i \neq j \end{cases} = \delta_{i,j} \det(A)$$

(mit dem Kronecker- δ). Wir entwickeln die Determinante von \tilde{A} nach der j -ten Zeile:

$$\begin{aligned} \delta_{i,j} \det(A) = \det(\tilde{A}) &= \begin{vmatrix} \text{(Zeile 1:)} & \alpha_{1,1} & \cdots & \alpha_{1,n} \\ & \vdots & & \vdots \\ \text{(Zeile } j\text{:)} & \alpha_{i,1} & \cdots & \alpha_{i,n} \\ & \vdots & & \vdots \\ \text{(Zeile } n\text{:)} & \alpha_{n,1} & \cdots & \alpha_{n,n} \end{vmatrix} \\ &= \sum_{k=1}^n (-1)^{j+k} \alpha_{i,k} \det(A'_{j,k}) = \sum_{k=1}^n \alpha_{i,k} \alpha_{k,j}^\# = \gamma_{i,j}. \end{aligned}$$

Damit ist der Satz bewiesen.

□

BEMERKUNG 4.2.2. Besonders einfach ist dieser Satz zur Invertierung von 2×2 -Matrizen anzuwenden:

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

4.3. Cramersche Regel.

SATZ 4.3.1 (Cramersche Regel). Sei $A \in \mathbb{K}^{n \times n}$ eine reguläre Matrix mit den Spalten a_1, \dots, a_n . Sei $b \in \mathbb{K}^n$ und $x = (\xi_1, \dots, \xi_n)^T$ die eindeutige Lösung des Gleichungssystems $Ax = b$. Dann gilt

$$\xi_i = \frac{1}{\det(A)} \begin{vmatrix} a_1 & \cdots & a_{i-1} & b & a_{i+1} & \cdots & a_n \end{vmatrix}$$

BEWEIS. Wir setzen

$$B_i := \begin{pmatrix} a_1 & \cdots & a_{i-1} & a_i & a_{i+1} & \cdots & a_n & b \\ 0 & & 0 & 1 & 0 & & 0 & \xi_i \end{pmatrix}$$

Dann ist

$$B_i \begin{pmatrix} x \\ -1 \end{pmatrix} = 0.$$

Daher ist B_i singulär, und $\det(B_i) = 0$. Wir entwickeln nach der letzten Zeile und vertauschen am Ende die Spalten:

$$\begin{aligned} 0 &= (-1)^{n+1+i} \det \begin{pmatrix} a_1 & \cdots & a_{i-1} & a_{i+1} & \cdots & a_n & b \\ a_1 & \cdots & a_{i-1} & a_i & a_{i+1} & \cdots & a_n \end{pmatrix} \\ &\quad + \xi_i \det \begin{pmatrix} a_1 & \cdots & a_{i-1} & a_i & a_{i+1} & \cdots & a_n \end{pmatrix} \\ &= - \det \begin{pmatrix} a_1 & \cdots & a_{i-1} & b & a_{i+1} & \cdots & a_n \end{pmatrix} \\ &\quad + \xi_i \det \begin{pmatrix} a_1 & \cdots & a_{i-1} & a_i & a_{i+1} & \cdots & a_n \end{pmatrix} \end{aligned}$$

Zur Bestimmung des Vorzeichens im letzten Schritt haben wir verwendet, dass die Spalte b , um an die Stelle i zu gelangen, die Spalten a_n, \dots, a_{i+1} überspringen muss. Das sind $(n-i)$ Transpositionen. \square

Die Cramersche Regel eignet sich gut zur Lösung von Systemen von 2 Gleichungen in 2 Unbekannten. Sonst ist die Bestimmung jeder einzelnen Determinante so aufwendig wie die Lösung des ganzen Systems, und wenn die Systemmatrix nicht regulär ist, scheitert die Cramersche Regel, während Pivotttransformationen die Lösungsmenge ergeben. Die Cramersche Regel ist aber von großem Wert für die Theorie, weil sie eine geschlossene Formel liefert, aus der sich Eigenschaften der Lösung ableiten lassen.

BEISPIEL 4.3.2. Wir lösen mit der Cramerschen Regel

$$\begin{aligned} 3x + 5y &= 7 \\ 2x - 7y &= 4 \end{aligned}$$

Lösung:

$$\begin{aligned} x &= \frac{\begin{vmatrix} 7 & 5 \\ 4 & -7 \end{vmatrix}}{\begin{vmatrix} 3 & 5 \\ 2 & -7 \end{vmatrix}} = \frac{-69}{-31} = \frac{69}{31}, \\ y &= \frac{\begin{vmatrix} 3 & 7 \\ 2 & 4 \end{vmatrix}}{\begin{vmatrix} 3 & 5 \\ 2 & -7 \end{vmatrix}} = \frac{-2}{-31} = \frac{2}{31}. \end{aligned}$$

4.4. Das äußere Produkt im dreidimensionalen Raum.

DEFINITION 4.4.1. Für zwei Vektoren des \mathbb{R}^3 definieren wir das äußere Produkt

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \times \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} := \begin{pmatrix} \begin{vmatrix} \alpha_2 & \beta_2 \\ \alpha_3 & \beta_3 \end{vmatrix} \\ -\begin{vmatrix} \alpha_1 & \beta_1 \\ \alpha_3 & \beta_3 \end{vmatrix} \\ \begin{vmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{vmatrix} \end{pmatrix}.$$

WARNUNG 4.4.2.

(1) Das äußere Produkt ist nicht kommutativ sondern

$$b \times a = -a \times b.$$

(2) Das äußere Produkt ist nicht assoziativ. Zum Beispiel gilt für die Einheitsvektoren $e_i \in \mathbb{R}^3$

$$(e_1 \times e_1) \times e_2 = 0 \times e_2 = 0,$$

$$\text{aber } e_1 \times (e_1 \times e_2) = e_1 \times e_3 = -e_2.$$

BEHAUPTUNG 4.4.3. In \mathbb{R}^3 betrachten wir das Standard-Skalarprodukt $\langle \cdot, \cdot \rangle$ und die davon induzierte euklidische Norm.

(1) Für alle Vektoren $a = (\alpha_i)_{i=1}^3$, $b = (\beta_i)_{i=1}^3$, $c = (\gamma_i)_{i=1}^3 \in \mathbb{R}^3$ gilt:

$$\langle a \times b, c \rangle = \begin{vmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{vmatrix}.$$

(2) $a \times b$ steht orthogonal auf a und b .

(3) $\det(a, b, a \times b) = \|a \times b\|^2$.

BEWEIS.

(1) Entwickeln von $\det(a, b, c)$ nach der letzten Spalte.

(2) $\langle a \times b, a \rangle = \det(a, b, a) = 0$, weil die Determinante zwei gleiche Spalten hat. Ebenso ist $\langle a \times b, b \rangle = \det(a, b, b) = 0$.

(3) $\det(a, b, a \times b) = \langle a \times b, a \times b \rangle = \|a \times b\|^2$.

□

Wir haben Fläche und Volumen nicht exakt definiert, doch ergibt sich als geometrische Interpretation für linear unabhängige $a, b \in \mathbb{R}^3$:

- Das System $(a, b, a \times b)$ ist positiv orientiert, weil die Determinante positiv ist.
- Das Volumen des Parallelepipeds, das von $a, b, a \times b$ aufgespannt ist, ist

$$|\det(a, b, a \times b)| = \|a \times b\|^2.$$

Zugleich ist es die Grundfläche des Parallelepipeds, also die Fläche des Parallelogramms, welches von a und b aufgespannt ist, multipliziert mit der Höhe $\|a \times b\|$ des Parallelepipeds. Daraus ergibt sich, dass $\|a \times b\|$ die Fläche des von a und b aufgespannten Parallelogramms ist.

Eigenwerte

1. Polynome

1.1. Polynom und Polynomfunktion.

DEFINITION 1.1.1. Sei \mathbb{K} ein Körper. Eine Polynomfunktion f über \mathbb{K} ist eine Funktion

$$f : \begin{cases} \mathbb{K} & \rightarrow \mathbb{K}, \\ x & \mapsto \alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n. \end{cases}$$

Dabei sind $\alpha_0, \dots, \alpha_n$ Elemente aus dem Körper \mathbb{K} und $n \in \mathbb{N} \cup \{0\}$.

SATZ 1.1.2. Seien $f(x) = \sum_{i=0}^m \alpha_i x^i$ und $g(x) = \sum_{j=0}^n \beta_j x^j$ zwei Polynomfunktionen über einem Körper \mathbb{K} . Wir definieren Addition und Multiplikation von Funktionen wie üblich: $[f + g](x) := f(x) + g(x)$, $[fg](x) := f(x)g(x)$. Dann sind auch $f + g$ und fg Polynomfunktionen, und es gilt:

$$\begin{aligned} [f + g](x) &= \sum_{k=0}^{\max(m,n)} (\alpha_k + \beta_k) x^k, \\ [fg](x) &= \sum_{k=0}^{m+n} \left(\sum_{j=0}^k \alpha_{k-j} \beta_j \right) x^k. \end{aligned}$$

In der Formel für die Summe setzen wir $\alpha_i = 0$ falls $i > m$ und $\beta_j = 0$ falls $j > n$.

BEWEIS. Um die Summenformel zu zeigen, setzen wir $N = \max(m, n)$, und $\alpha_i = 0$ für $i > m$, $\beta_j = 0$ für $j > n$, und können daher die Polynomfunktionen schreiben

$$f(x) = \sum_{k=0}^N \alpha_k x^k \quad g(x) = \sum_{k=0}^N \beta_k x^k.$$

Nun ist (unter Ausnützung der Kommutativität und Assoziativität der Körperaddition, und eines Distributivgesetzes)

$$f(x) + g(x) = \sum_{k=0}^N \alpha_k x^k + \sum_{k=0}^N \beta_k x^k = \sum_{k=0}^N (\alpha_k + \beta_k) x^k.$$

Für die Multiplikation erhalten wir

$$\begin{aligned}
 f(x)g(x) &= \left(\sum_{i=0}^m \alpha_i x^i \right) \left(\sum_{j=0}^n \beta_j x^j \right) \\
 &= \sum_{j=0}^n \sum_{i=0}^m \alpha_i \beta_j x^{i+j} \quad (\text{Klammern ausmultipliziert}) \\
 &= \sum_{j=0}^n \sum_{k=j}^{m+j} \alpha_{k-j} \beta_j x^k \quad (\text{Verschiebung des Summationsindex: } k := i + j) \\
 &= \sum_{k=0}^{m+n} \sum_{j=0}^k \alpha_{k-j} \beta_j x^k \quad (\text{Vertauschung der Summationsreihenfolge}) \\
 &= \sum_{k=0}^{m+n} \left(\sum_{j=0}^k \alpha_{k-j} \beta_j \right) x^k.
 \end{aligned}$$

□

BEISPIEL 1.1.3. Auf $\mathbb{K} = \{0, 1\}$ definieren wir die Rechenoperationen

$$\begin{array}{c|cc}
 + & 0 & 1 \\
 \hline
 0 & 0 & 1 \\
 1 & 1 & 0
 \end{array}
 \quad
 \begin{array}{c|cc}
 \cdot & 0 & 1 \\
 \hline
 0 & 0 & 0 \\
 1 & 0 & 1
 \end{array}$$

Man zeigt leicht, dass \mathbb{K} ein Körper ist. Wir berechnen die Polynomfunktion $f(x) = x^2 + x$ für $x \in \mathbb{K}$.

Lösung:

$$\begin{aligned}
 f(0) &= 0^2 + 0 = 0 + 0 = 0, \\
 f(1) &= 1^2 + 1 = 1 + 1 = 0.
 \end{aligned}$$

□

Dieses Beispiel zeigt, dass über diesem Körper die beiden Polynomfunktionen $f(x) = x^2 + x$ und $g(x) = 0$ dieselbe Funktion beschreiben. Trotzdem gibt es Gründe, zwischen den Ausdrücken $x^2 + x$ und 0 zu unterscheiden. Das ist der Grund, warum wir ein Polynom nicht einfach als Polynomfunktion definieren, sondern mit einem formalen Zugang. Natürlich definieren wir die Rechenvorschriften für Polynome genau so, dass sie zu den Rechenvorschriften für Polynomfunktionen passen.

DEFINITION 1.1.4. Sei \mathbb{K} ein Körper.

- (1) Ein Polynom über \mathbb{K} ist eine Folge $(\alpha_i)_{i=0}^{\infty}$ von Elementen $\alpha_i \in \mathbb{K}$, bei der nur endlich viele Glieder von 0 verschieden sind.
- (2) Auf der Menge der Polynome über \mathbb{K} führen wir die folgenden Rechenoperationen ein:

$$\begin{aligned}
 (\alpha_i)_{i=0}^{\infty} + (\beta_i)_{i=0}^{\infty} &= (\alpha_i + \beta_i)_{i=0}^{\infty}, \\
 (\alpha_i)_{i=0}^{\infty} \cdot (\beta_i)_{i=0}^{\infty} &= \left(\sum_{j=0}^k \alpha_{k-j} \beta_j \right)_{k=0}^{\infty}.
 \end{aligned}$$

- (3) Das Nullpolynom ist $0 = (0)_{i=0}^{\infty}$.
- (4) Ist $p = (\alpha_i)_{i=0}^{\infty}$ nicht das Nullpolynom, so definieren wir den Grad (engl.: degree) von p durch

$$\deg(p) := \max\{i \in \mathbb{N} \mid \alpha_i \neq 0\}.$$

Der Grad des Nullpolynoms wird mit $-\infty$ angesetzt.

- (5) Die Polynome von Grad 0 und das Nullpolynom heißen konstante Polynome.

(6) Die Polynome von Grad ≤ 1 heißen lineare Polynome.

Da die Schreibweise durch Folgen unhandlich und intuitiv wenig ansprechend ist, wählen wir ein beliebiges Symbol (z.B. x) und schreiben

$$(\alpha_i)_{i=0}^{\infty} = \alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n.$$

Wir nennen dann das Symbol x die Unbestimmte. Die Menge der Polynome über \mathbb{K} , mit der Unbestimmten x geschrieben, bezeichnen wir mit $\mathbb{K}[x]$.

DEFINITION 1.1.5. Sei R eine Menge, die mindestens ein Element 0 enthält. Seien $+$ und \cdot zwei innere Verknüpfungen auf R . Das Quadrupel $(R, +, \cdot, 0)$ heißt ein Ring, wenn gilt:

- (1) $(R, +)$ ist eine abelsche Gruppe mit neutralem Element 0 .
- (2) Die Multiplikation \cdot ist assoziativ.
- (3) Es gelten die Distributivgesetze (für alle $a, b, c \in R$):

$$(a + b) \cdot c = a \cdot c + b \cdot c,$$

$$(b + c) \cdot a = b \cdot a + c \cdot a.$$

Wenn zusätzlich \cdot kommutativ ist, dann heißt R ein kommutativer Ring.

Wenn ein Element $1 \in R$ existiert, sodass für alle $a \in R$ gilt

$$a \cdot 1 = 1 \cdot a = a,$$

dann heißt R ein Ring mit Eins.

BEHAUPTUNG 1.1.6. Sei \mathbb{K} ein Körper und $\mathbb{K}[x]$ die Menge der Polynome über \mathbb{K} .

- 1) $(\mathbb{K}[x], +, \cdot)$ ist ein kommutativer Ring mit Eins. Das Nullelement ist das Nullpolynom, das Einselement ist das konstante Polynom 1 .
- 2) Mit der Polynomaddition und dem Produkt

$$\lambda \cdot \sum_{i=0}^n \alpha_i x^i := \sum_{i=0}^n (\lambda \alpha_i) x^i$$

bildet $\mathbb{K}[x]$ einen Vektorraum über \mathbb{K} .

BEWEIS. Nachrechnen. □

BEHAUPTUNG 1.1.7. Seien $f \neq 0, g \neq 0$ zwei Polynome über einem Körper \mathbb{K} . Es gilt

$$\deg(f + g) = \max(\deg(f), \deg(g)) \quad \text{falls } \deg(f) \neq \deg(g),$$

$$\deg(f + g) \leq \deg(f) \quad \text{falls } \deg(f) = \deg(g),$$

$$\deg(fg) = \deg(f) + \deg(g).$$

BEWEIS. Sei $f = (\alpha_i)_{i=0}^{\infty}, g = (\beta_i)_{i=0}^{\infty}$. Sei $m = \deg(f)$ und $n = \deg(g)$. Ist $k > \max(m, n)$, so ist $\alpha_k = \beta_k = 0$ und folglich $\alpha_k + \beta_k = 0$. Daher kann $\deg(f + g)$ nicht größer als $\max(m, n)$ sein. Sei nun $m \neq n$, o.B.d.A. nehmen wir an $m < n$. Dann ist $\alpha_n + \beta_n = \beta_n \neq 0$, folglich ist $\deg(f + g) \geq n = \max(m, n)$. Ist $k > m + n$ und $j \leq k$, so ist entweder $j > n$ oder $k - j > m$ (oder beides), folglich muss gelten $\beta_j = 0$ oder $\alpha_{k-j} = 0$. Daher ist

$$\sum_{j=0}^k \alpha_{k-j} \beta_j = 0.$$

Es kann also der Grad von fg nicht größer als $m + n$ sein. Andererseits ist $\alpha_m \neq 0$ und $\beta_n \neq 0$. Daher ist

$$\sum_{j=0}^{m+n} \alpha_{k-j} \beta_j = \alpha_m \beta_n \neq 0.$$

Also ist $\deg(fg) \geq m + n$. \square

BEHAUPTUNG 1.1.8. *Sei \mathbb{K} ein Körper. Der Polynomring $\mathbb{K}[x]$ ist nullteilerfrei, d.h., sind $f, g \in \mathbb{K}[x] \setminus \{0\}$, so ist $fg \neq 0$.*

BEWEIS. Sind $f, g \neq 0$, so sind $\deg(f) \geq 0$, $\deg(g) \geq 0$. Dann ist nach Behauptung 1.1.7 auch $\deg(fg) \geq 0$, also ist $fg \neq 0$. \square

1.2. Division mit Rest.

SATZ 1.2.1 (Polynomdivision mit Rest). *Sei \mathbb{K} ein Körper, f, g zwei Polynome über \mathbb{K} , wobei $g \neq 0$. Dann gibt es eindeutig bestimmte Polynome p, r mit den Eigenschaften*

$$\begin{aligned} f &= pg + r, \\ \deg(r) &< \deg(g). \end{aligned}$$

BEWEIS. *Eindeutigkeit:* Sei $f = p_1g + r_1 = p_2g + r_2$. Dann ist also $g(p_1 - p_2) = r_2 - r_1$. Ist $p_1 \neq p_2$, so ist $\deg(p_1 - p_2) \geq 0$ und folglich $\deg(g(p_1 - p_2)) \geq 0 + \deg(g)$. Andererseits ist $\deg(r_2 - r_1) < \deg(g)$, und es entsteht ein Widerspruch. Also ist $p_1 = p_2$ und es folgt $r_1 = r_2$.

Existenz: Wir führen eine vollständige Induktion¹ nach $\deg(f)$:

Verankerung: Ist $\deg(f) < \deg(g)$, so ist $f = 0 \cdot g + f$. Wähle also $p = 0$, $r = f$.

Induktionsschritt: Sei nun $k \geq \deg(g)$ und jedes Polynom f mit $\deg(f) < k$ in der Form $f = pg + r$ darstellbar. Sei nun $f \in \mathbb{K}[x]$ mit $\deg(f) = k$. Wir zeigen, dass sich auch f zerlegen lässt. Es ist also $f = \sum_{i=0}^k \alpha_i x^i$, $g = \sum_{j=0}^m \beta_j x^j$, wobei $m = \deg(g) \leq k$, $\beta_m \neq 0$. Es ist dann

$$f_1 := f - \frac{\alpha_k}{\beta_m} x^{k-m} g$$

ein Polynom von Grad kleiner als k . Denn weil $\deg(f) = k$, $\deg(x^{k-m}g) = (k - m) + m = k$, ist $\deg(f_1) \leq k$. Aber der k -te Koeffizient von f_1 ist $\alpha_k - \frac{\alpha_k}{\beta_m} \beta_m = 0$. Nach Induktionsvoraussetzung kann man also zerlegen: $f_1 = p_1g + r$ mit $\deg(r) < \deg(g)$. Nun ist

$$f = f_1 + \frac{\alpha_k}{\beta_m} x^{k-m} g = \left(p_1 + \frac{\alpha_k}{\beta_m} x^{k-m} \right) g + r.$$

\square

DEFINITION 1.2.2. Seien f, g Polynome über \mathbb{K} mit $g \neq 0$, sei $f = pg + r$ mit $\deg(r) < \deg(g)$. Wir nennen das Polynom r den Rest bei der Division von f durch g .

BEMERKUNG 1.2.3. Der Existenzbeweis von Satz 1.2.1 liefert zugleich den Algorithmus für die Auffindung von p und r .

¹Das ist eine sogenannte starke Induktion. Angenommen eine Aussage A gilt für eine ganze Zahl m (Verankerung). Wenn man zeigen kann, dass die Aussage auch für $k + 1$ gilt, falls sie für alle ganzen Zahlen $m, m + 1, \dots, k$ gilt (Induktionsschritt), dann gilt die Aussage für alle ganzen Zahlen $\geq m$.

BEISPIEL 1.2.4. Dividiere mit Rest:

$$\underbrace{6x^4 - 7x^3 + 7x^2 + 5x - 5}_f : \underbrace{2x^2 - 3x + 2}_g.$$

Lösung:

$$\begin{array}{r r r r r r} 6x^4 & -7x^3 & +7x^2 & +5x & -5 & \frac{6x^4}{2x^2} = 3x^2 \\ 6x^4 & -9x^3 & +6x^2 & & & 3x^2g \\ \hline & 2x^3 & +x^2 & +5x & -5 & f_1 = f - 3x^2g \\ & 2x^3 & -3x^2 & +2x & & xg \\ \hline & & 4x^2 & +3x & -5 & f_2 = f_1 - xg \\ & & 4x^2 & -6x & +4 & 2g \\ \hline & & & 9x & -9 & r = f_2 - 2g \end{array}$$

Also ist

$$f = (3x^2 + x + 2)g + (9x - 9).$$

□

Einfach, aber von fundamentaler Bedeutung ist der folgende Satz. Er stellt die Verbindung zwischen der Theorie der Nullstellen und der Lehre von der Teilbarkeit der Polynome her:

SATZ 1.2.5. Sei $f = \sum_{i=0}^n \alpha_i x^i$ ein Polynom über einem Körper \mathbb{K} . Die dazugehörige Polynomfunktion sei

$$\tilde{f} : \begin{cases} \mathbb{K} & \rightarrow \mathbb{K}, \\ t & \mapsto \sum_{i=0}^n \alpha_i t^i. \end{cases}$$

Sei $\gamma \in \mathbb{K}$ und $f = p(x - \gamma) + r$ mit einem Polynom p (zugehörige Polynomfunktion \tilde{p}) und einem konstanten Polynom $r \in \mathbb{K}$.

Dann ist $r = \tilde{f}(\gamma)$.

BEWEIS. Wir setzen für x den Wert γ ein:

$$\tilde{f}(\gamma) = \tilde{p}(\gamma)(\gamma - \gamma) + r = r.$$

□

ALGORITHMUS 1.2.6 (Horner-Schema). Sei \mathbb{K} ein Körper, sei $f = \sum_{i=0}^n \alpha_i x^i \in \mathbb{K}[x]$, sei $\gamma \in \mathbb{K}$. Gesucht ist ein Polynom p und eine Konstante r so, dass $f = p(x - \gamma) + r$. (Insbesondere ist $r = \tilde{f}(\gamma)$, wobei \tilde{f} die zu f gehörende Polynomfunktion ist.)

Wir berechnen rekursiv die Koeffizienten von

$$p = \sum_{i=0}^{n-1} \beta_i x^i$$

und den Rest r , beginnend mit β_{n-1} :

$$\begin{aligned} \beta_{n-1} &= \alpha_n, \\ \beta_{k-1} &= \gamma \beta_k + \alpha_k \quad \text{für } k = n-2, \dots, 1, \\ r &= \gamma \beta_0 + \alpha_0. \end{aligned}$$

BEWEIS. Das ist nichts anderes als eine geschickte Art, die Division durch $(x - \gamma)$ mit Rest anzuschreiben. □

BEISPIEL 1.2.7. Sei $f = 4x^5 + 2x^4 - 6x^3 + 0x^2 - 3x + 2$, $\gamma = 2$. Schreiben Sie $f = p(x - 2) + r$. Was ist $\tilde{f}(2)$?

Lösung:

i	5	4	3	2	1	0
α_i	4	2	-6	0	-3	2
β_{i-1} bzw. r	4	10	14	28	53	108

Es ist also $p = 4x^4 + 10x^3 + 14x^2 + 28x + 53$, $r = \tilde{f}(2) = 108$. □

Viel schneller als die einzelnen Potenzen auszurechnen: $\tilde{f}(2) = 4 \cdot 2^5 + 2 \cdot 2^4 - 6 \cdot 2^3 - 3 \cdot 2 + 2!$

1.3. Nullstellen.

DEFINITION 1.3.1. Sei \mathbb{K} ein Körper und $f \in \mathbb{K}[x]$ ein Polynom mit zugehöriger Polynomfunktion \tilde{f} . Ein $\gamma \in \mathbb{K}$ heißt Nullstelle von f , wenn gilt $\tilde{f}(\gamma) = 0$.

SATZ 1.3.2. Sei \mathbb{K} ein Körper und $f \in \mathbb{K}[x]$ ein Polynom. Sei $\gamma \in \mathbb{K}$ und $f = p(x - \gamma) + r$. Dann ist γ genau dann eine Nullstelle von f , wenn $r = 0$, d.h., f ist durch $x - \gamma$ ohne Rest teilbar.

BEWEIS. Direkte Folge von Satz 1.2.5. □

DEFINITION 1.3.3. Sei \mathbb{K} ein Körper, $f \neq 0 \in \mathbb{K}[x]$, und $\gamma \in \mathbb{K}$. Die Vielfachheit von γ in f ist die größte Potenz m so, dass $(x - \gamma)^m$ ein Teiler von f ist.

Anders ausgedrückt: Ist m die Vielfachheit von γ in f , so ist $f = (x - \gamma)^m p$, wobei γ keine Nullstelle von p ist.

Ist γ keine Nullstelle von f , dann ist die Vielfachheit von γ in f gleich Null.

BEISPIEL 1.3.4. Überprüfen Sie, dass -3 eine Nullstelle von $f = x^4 + 7x^3 + 9x^2 - 27x - 54$ ist, und bestimmen Sie die Vielfachheit.

Lösung: Wie oft läßt sich $x + 3$ aus f ohne Rest abdividieren? Wir dividieren natürlich mit dem Horner-Schema:

1	7	9	-27	-54	f
1	4	-3	-18	0	Rest 0, $f = f_1(x + 3)$, $f_1 = x^3 + 4x^2 - 3x - 18$,
1	1	-6	0		Rest 0, $f_1 = f_2(x + 3)$, $f_2 = x^2 + x - 6$,
1	-2	0			Rest 0, $f_2 = f_3(x + 3)$, $f_3 = x - 2$,
1	-5				Rest $\tilde{f}_3(-3) = -5 \neq 0$.

Also ist $f = (x + 3)^3 f_3$ mit $\tilde{f}_3(-3) = -5 \neq 0$. Die Vielfachheit von -3 als Nullstelle von f ist 3. □

Wir erwähnen noch ohne Beweis eine nützliche Regel zur Bestimmung von Vielfachheiten mit Hilfe der Differentialrechnung.

BEMERKUNG 1.3.5. Sei \mathbb{K} ein Körper. Für ein Polynom $f = \sum_{i=0}^n \alpha_i x^i$ definieren wir die erste Ableitung $f' := \sum_{i=1}^n i \alpha_i x^{i-1}$, und die höheren Ableitungen f'' , $f^{(3)}$, $f^{(4)}$... wie üblich durch wiederholte Anwendung der ersten Ableitung. $f^{(0)}$ ist einfach f .

Sei nun $f \neq 0 \in \mathbb{K}[x]$ und $\gamma \in \mathbb{K}$ eine Nullstelle von f . Dann gilt:

Die Vielfachheit von γ in f ist m genau dann, wenn γ eine Nullstelle der Polynome $f, f', \dots, f^{(m-1)}$ ist, aber keine Nullstelle von $f^{(m)}$.

Unser Ziel ist nun, ein Polynom in Faktoren zu zerlegen

$$f = (x - \gamma_1)^{m_1} \cdots (x - \gamma_k)^{m_k} g$$

wobei $\gamma_1, \dots, \gamma_k$ die Nullstellen von f mit den Vielfachheiten m_1, \dots, m_k sind, und g keine Nullstellen hat.

LEMMA 1.3.6. Sei \mathbb{K} ein Körper. Sei $\gamma \in \mathbb{K}$ und $f = (x - \gamma)^m g \neq 0$ ein Polynom über \mathbb{K} . Sei $\lambda \neq \gamma$ und $n \in \mathbb{N}$.

Dann ist sind die Vielfachheiten von λ in f und in g gleich.

BEWEIS. Sei n die Vielfachheit von λ in f , und k die Vielfachheit von λ in g .
Es sei

$$f = (x - \lambda)^n p \quad \text{und} \quad g = (x - \lambda)^k q$$

wobei λ weder eine Nullstelle von p noch von q ist.

Dann ist also

$$f = (x - \gamma)^m g = (x - \lambda)^k (x - \gamma)^m q.$$

Also ist $(x - \lambda)^k$ ein Teiler von f , und die Vielfachheit von λ in f ist mindestens k .
Daher können wir schreiben

$$f = (x - \lambda)^n p = (x - \gamma)^m (x - \lambda)^k q, \\ \text{also} \quad (x - \lambda)^k [(x - \lambda)^{n-k} p - (x - \gamma)^m q] = 0,$$

und wegen der Nullteilerfreiheit des Polynomrings ist

$$(x - \lambda)^{n-k} p - (x - \gamma)^m q = 0.$$

Weil aber λ keine Nullstelle von $(x - \gamma)^m q$ ist, muss $n - k = 0$ sein. Daher sind die Vielfachheiten von λ in f und g gleich. \square

SATZ 1.3.7. Sei \mathbb{K} ein Körper und $f \neq 0 \in \mathbb{K}[x]$. Seien $\lambda_1, \dots, \lambda_k$ die Nullstellen von f mit den Vielfachheiten m_1, \dots, m_k .

- 1) Dann gibt es ein Polynom $p \in \mathbb{K}[x]$, welches keine Nullstellen besitzt, so dass

$$f = (x - \lambda_1)^{m_1} (x - \lambda_2)^{m_2} \cdots (x - \lambda_k)^{m_k} p.$$

- 2) Das Polynom f hat höchstens so viele Nullstellen (gezählt mit Vielfachheit) wie sein Grad ist: $\sum_{j=1}^k m_j \leq \deg(f)$.

BEWEIS. Teil (1): Vollständige Induktion nach k :

Induktionsverankerung $k = 1$: Weil λ_1 eine m_1 -fache Nullstelle von f ist, kann f zerlegt werden: $f = (x - \lambda_1)^{m_1} p$. Da jede Nullstelle von p auch eine weitere Nullstelle von f wäre, hat p keine Nullstellen.

Induktionsschritt: Der Satz sei richtig für ein Polynom mit k Nullstellen. Sei nun f ein Polynom mit $k + 1$ Nullstellen. Weil λ_{k+1} eine m_{k+1} -fache Nullstelle von f ist, gilt $f = (x - \lambda_{k+1})^{m_{k+1}} f_1$, dabei sind alle weiteren Nullstellen von f auch Nullstellen von f_1 mit derselben Vielfachheit. Daher hat f_1 genau k Nullstellen, und wir können die Induktionsvoraussetzung anwenden: $f_1 = p \prod_{j=1}^k (x - \lambda_j)^{m_j}$. Dabei hat p keine Nullstellen. Schließlich erhalten wir

$$f = (x - \lambda_{k+1})^{m_{k+1}} f_1 = p \prod_{j=1}^{k+1} (x - \lambda_j)^{m_j}$$

und die Induktion ist fertig.

Teil (2): In Hinblick auf Teil (1) ist

$$\deg(f) = \deg(p \prod_{j=1}^k (x - \lambda_j)^{m_j}) = \deg(p) + \sum_{j=1}^k m_j \geq \sum_{j=1}^k m_j.$$

\square

KOROLLAR 1.3.8. Sei \mathbb{K} ein Körper mit unendlich vielen Elementen, und seien f, g zwei verschiedene Polynome über \mathbb{K} . Dann sind auch die Polynomfunktionen \tilde{f}, \tilde{g} verschieden.

BEWEIS. Sei $\tilde{f} = \tilde{g}$, d.h., für alle $t \in \mathbb{K}$ ist $\tilde{f}(t) - \tilde{g}(t) = 0$. Es ist also jedes Element aus \mathbb{K} eine Nullstelle des Polynoms $f - g$. Wenn $f - g$ nicht das Nullpolynom wäre, könnte es aber nur endlich viele Nullstellen geben. Also ist $f = g$. \square

Überlegen Sie selbst, dass es über einem Körper mit endlich vielen Elementen immer Polynomfunktionen gibt, die durch mehrere verschiedene Polynome dargestellt werden können.

DEFINITION 1.3.9. Sei \mathbb{K} ein Körper. \mathbb{K} heißt algebraisch abgeschlossen, wenn jedes nicht konstante Polynom über \mathbb{K} mindestens eine Nullstelle besitzt.

BEMERKUNG 1.3.10. \mathbb{K} ist genau dann algebraisch abgeschlossen, wenn jedes Polynom $f \neq 0$ über \mathbb{K} in das Produkt von Linearfaktoren und einer Konstanten zerfällt, also wenn es $\lambda_1, \dots, \lambda_k, \alpha \in \mathbb{K}$ und $m_1, \dots, m_k \in \mathbb{N}$ gibt, so dass

$$f = \alpha \prod_{j=1}^k (x - \lambda_j)^{m_j}.$$

Anders ausgedrückt: Die Anzahl der Nullstellen von f , gezählt mit Vielfachheiten, ergibt genau den Grad von f .

BEWEIS. Nach Satz 1.3.7 lässt sich f zerlegen

$$f = p \prod_{j=1}^k (x - \lambda_j)^{m_j}.$$

Dabei hat p keine Nullstellen.

“ \Rightarrow ”: Sei \mathbb{K} algebraisch abgeschlossen, und hat p keine Nullstellen, so muss p ein konstantes Polynom sein. Es ist dann auch $\deg(p) = 0$ und folglich $\deg(f) = \sum_{j=1}^k m_j$.

“ \Leftarrow ”: Sei \mathbb{K} nicht algebraisch abgeschlossen. Dann gibt es ein nicht konstantes Polynom f ohne Nullstellen. Es ist dann in der obigen Zerlegung $f = p$ und p nicht konstant. \square

Für den Beweis des folgenden tiefen Satzes von fundamentaler Bedeutung fehlen uns hier die Grundlagen:

SATZ 1.3.11 (Hauptsatz der Algebra). *Der Körper \mathbb{C} der komplexen Zahlen ist algebraisch abgeschlossen.*

Überlegen Sie selbst, dass der Körper \mathbb{R} der reellen Zahlen nicht algebraisch abgeschlossen ist.

2. Eigenvektoren und Eigenwerte

2.1. Beispiele.

BEISPIEL 2.1.1. Wir betrachten die folgende Matrix $A \in \mathbb{R}^{2 \times 2}$ und die Vektoren $x, y \in \mathbb{R}^2$:

$$A = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}, \quad x = \begin{pmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}, \quad y = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix}.$$

Wir berechnen Ax und Ay . Welche Folgerungen können wir über die lineare Abbildung f schließen, die durch A gegeben wird?

Lösung: Einfache Rechnung zeigt: $Ax = -x$, $Ay = y$. Überdies ist $B = (x, y)$ eine Orthonormalbasis des \mathbb{R}^2 . Die Matrix von f bezüglich der Basis B erhalten wir nach der Regel: Die Spalten sind die Bilder der Basisvektoren (natürlich auch in neuen Koordinaten geschrieben):

$$M_B^B(f) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Also gibt es ein Koordinatensystem in \mathbb{R}^2 , sodass die Abbildung f das Vorzeichen der ersten Koordinate umkehrt, und die zweite Koordinate unverändert lässt. Die Abbildung f ist also eine Spiegelung an der Geraden

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{pmatrix}.$$

\square

BEISPIEL 2.1.2. Wir betrachten die lineare Abbildung mit der folgenden Matrix $A \in \mathbb{R}^{3 \times 3}$ und die Vektoren $x, y, z \in \mathbb{R}^3$.

$$A = \begin{pmatrix} 4 & 2 & -6 \\ 8 & 19 & -42 \\ 4 & 8 & -18 \end{pmatrix}, \quad x = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \quad y = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \quad z = \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix}.$$

Wieder berechnen wir Ax , Ay und Az .

Lösung: Wir erhalten: $Ax = 2x$, $Ay = 3y$, $Az = 0$. Verwenden wir $B = (x, y, z)$ als Basis für ein neues Koordinatensystem, so ist

$$M_B^B(f) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

In diesem Koordinatensystem lässt sich f also leicht beschreiben: Die erste Koordinate wird verdoppelt, die zweite Koordinate auf das Dreifache gestreckt, und die dritte Koordinate auf 0 gesetzt. Beachten Sie aber, dass diese Basis kein Orthonormalsystem ist. \square

In beiden Beispielen gewinnen wir Einsicht in die Struktur der linearen Abbildung, indem wir das Koordinatensystem wechseln. Unsere Basis bestand immer aus Vektoren der Form $Ax = \lambda x$ mit $\lambda \in \mathbb{K}$.

BEISPIEL 2.1.3. Sei $A \in \mathbb{R}^{3 \times 3}$ die Matrix einer Drehung um die Achse $0 + tv$. Dann lässt A die Punkte auf der Achse fest, also $Av = v$.

2.2. Eigenwerte, Eigenvektoren, Eigenraum.

DEFINITION 2.2.1. Sei \mathbb{K} ein Körper und $A \in \mathbb{K}^{n \times n}$ eine quadratische Matrix. Ein Vektor $v \in \mathbb{K}^n \setminus \{0\}$ heißt Eigenvektor von A zum Eigenwert $\lambda \in \mathbb{K}$, wenn gilt $Av = \lambda v$. Insbesondere heißt also $\lambda \in \mathbb{K}$ ein Eigenwert von A , wenn es dazu einen Eigenvektor $v \in \mathbb{K}^n \setminus \{0\}$ gibt.

SATZ 2.2.2. Sei \mathbb{K} ein Körper und $A \in \mathbb{K}^{n \times n}$ eine quadratische Matrix. Sei $\lambda \in \mathbb{K}$, und sei $W = \{v \in \mathbb{K}^n \mid Av = \lambda v\}$.

- 1) $W \neq \{0\}$ genau dann wenn λ ein Eigenwert von A ist.
- 2) $W = \ker(\lambda E_n - A)$.
- 3) W ist ein Unterraum von \mathbb{K}^n .

BEWEIS. *Teil (1):* Für jedes $\lambda \in \mathbb{K}$ gilt $A0 = \lambda 0$. Es ist λ genau dann ein Eigenwert, wenn es zusätzlich ein $v \neq 0$ gibt mit $Av = \lambda v$.

Teil (2): Es gilt

$$Av = \lambda v \Leftrightarrow 0 = \lambda v - Av \Leftrightarrow 0 = (\lambda E_n - A)v.$$

Teil (3): Der Kern einer Matrix ist immer ein Unterraum. \square

DEFINITION 2.2.3. Sei \mathbb{K} ein Körper und $A \in \mathbb{K}^{n \times n}$ eine quadratische Matrix. Sei $\lambda \in \mathbb{K}$ ein Eigenwert von A . Der Raum $\ker(\lambda E_n - A)$ heißt Eigenraum von A zum Eigenwert λ . Die Dimension von $\ker(\lambda E_n - A)$ heißt geometrische Vielfachheit von λ als Eigenwert von A .

SCHREIBWEISE 2.2.4. Statt $\lambda E_n - A$ schreibt man auch einfach $\lambda - A$.

BEMERKUNG 2.2.5. Der Eigenraum von A zu $\lambda = 0$ ist einfach der Kern von A .

2.3. Berechnung von Eigenwerten.

SATZ 2.3.1. Sei \mathbb{K} ein Körper und $A \in \mathbb{K}^{n \times n}$ eine quadratische Matrix. Sei $\lambda \in \mathbb{K}$. Dann gilt:
 λ ist genau dann ein Eigenwert von A wenn $\det(\lambda - A) = 0$ ist.

BEWEIS. Wegen Satz 2.2.2 ist λ genau dann ein Eigenwert, wenn $\ker(\lambda - A) \neq \{0\}$ ist, und das gilt genau dann, wenn $(\lambda - A)$ singulär ist. Eine Matrix ist wiederum genau dann singulär, wenn ihre Determinante null ist. \square

Damit ergibt sich folgendes Verfahren zur Berechnung von Eigenwerten:

ALGORITHMUS 2.3.2. Die Eigenräume und Eigenwerte von $A \in \mathbb{K}^{n \times n}$ kann man folgendermaßen berechnen:

- 1) Setze λ als Unbekannte an und berechne $\det(\lambda - A)$. Es ergibt sich ein Polynom vom Grad n .
- 2) Berechne (bei Bedarf mit einem numerischen Gleichungslöser) die Nullstellen dieses Polynoms. Sie sind die Eigenwerte von A .
- 3) Zu jedem Eigenwert λ ist $\ker(\lambda - A)$ der zugehörige Eigenraum von A .

Für spezielle Matrizen (z.B. hermitesche) gibt es ganz andere Zugänge zur Berechnung der Eigenwerte. Sie hören davon mehr in einer Lehrveranstaltung über numerische Mathematik.

BEISPIEL 2.3.3. Berechnen Sie alle Eigenwerte und Eigenräume der Matrix A als Matrix über \mathbb{R} und als Matrix über \mathbb{C} .

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

Lösung: Schritt 1: Berechnung von $\det(\lambda - A)$ mittels Entwicklung nach der ersten Zeile.

$$\begin{aligned} & \det \left(\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} - \begin{pmatrix} -1 & 1 & 2 \\ -1 & -5 & 2 \\ 2 & -2 & -4 \end{pmatrix} \right) \\ &= \begin{vmatrix} \lambda + 1 & -1 & -2 \\ 1 & \lambda + 5 & -2 \\ -2 & 2 & \lambda + 4 \end{vmatrix} \\ &= (\lambda + 1) \begin{vmatrix} \lambda + 5 & -2 \\ 2 & \lambda + 4 \end{vmatrix} - (-1) \begin{vmatrix} 1 & -2 \\ -2 & \lambda + 4 \end{vmatrix} + (-2) \begin{vmatrix} 1 & \lambda + 5 \\ -2 & 2 \end{vmatrix} \\ &= (\lambda + 1)[(\lambda + 5)(\lambda + 4) + 4] + [\lambda + 4 - 4] - 2[2 + 2(\lambda + 5)] \\ &= \lambda^3 + 10\lambda^2 + 30\lambda. \end{aligned}$$

Schritt 2: Nullstellen suchen: Offensichtlich ist $\lambda_1 = 0$ eine Nullstelle, damit haben wir einen Eigenwert. Die anderen Eigenwerte, falls welche existieren, sind Nullstellen des quadratischen Polynoms $\lambda^2 + 10\lambda + 30$. Diese gibt es nur im Körper der komplexen Zahlen: $\lambda_{2,3} = -5 \pm i\sqrt{5}$. Es hat also A über \mathbb{R} nur einen Eigenwert, nämlich 0, über \mathbb{C} besitzt A zusätzlich 2 konjugiert komplexe Eigenwerte.

Schritt 3: Berechnung der Eigenräume. Wir arbeiten mit Pivotschritten.

$W_1 := \ker(0 - A)$ über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$:

$$\begin{pmatrix} 1* & -1 & -2 \\ 1 & 5 & -2 \\ -2 & 2 & 4 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & -1 & -2 \\ 0 & 6* & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Die dritte Spalte bleibt als nichtbasische Spalte, wir setzen $x_3 = t$ und erhalten

$$\ker(-A) = \left\{ t \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \mid t \in \mathbb{K} \right\}.$$

$W_2 = \ker(-5 + i\sqrt{5} - A)$ über \mathbb{C} :

$$\begin{pmatrix} -4 + i\sqrt{5} & -1 & -2 \\ 1 & i\sqrt{5} & -2 \\ -2 & 2 & -1 + i\sqrt{5} \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 4 + 4i\sqrt{5} & -10 + 2i\sqrt{5} \\ 1 & i\sqrt{5} & -2 \\ 0 & 2 + 2i\sqrt{5} & -5 + i\sqrt{5} \end{pmatrix} \\ \rightarrow \begin{pmatrix} 0 & 0 & 0 \\ 1 & i\sqrt{5} & -2 \\ 0 & 2 + 2i\sqrt{5} & -5 + i\sqrt{5} \end{pmatrix}$$

Wir setzen $x_3 = t$ und erhalten

$$x_2 = -t \frac{-5 + i\sqrt{5}}{2 + 2i\sqrt{5}} = -t \frac{i\sqrt{5}(i\sqrt{5} + 1)}{2(i\sqrt{5} + 1)} = -t \frac{i\sqrt{5}}{2}.$$

Letztlich ist

$$x_1 = -i\sqrt{5}x_2 + 2t = -\frac{t}{2}.$$

Damit ist der Eigenraum zu $\lambda_2 = -5 + i\sqrt{5}$

$$\ker(5 - i\sqrt{5} - A) = \left\{ t \begin{pmatrix} -\frac{1}{2} \\ -\frac{\sqrt{5}}{2}i \\ 1 \end{pmatrix} \mid t \in \mathbb{C} \right\}.$$

Für $\lambda_3 = -5 - i\sqrt{5}$ ist die Rechnung ebenso, nur mit konjugiert komplexen Zahlen, und es ist

$$\ker(-5 + i\sqrt{5} - A) = \left\{ t \begin{pmatrix} -\frac{1}{2} \\ \frac{\sqrt{5}}{2}i \\ 1 \end{pmatrix} \mid t \in \mathbb{C} \right\}.$$

Alle drei Eigenräume sind eindimensional. Damit besitzt A über \mathbb{R} einen Eigenwert mit geometrischer Vielfachheit 1. Über \mathbb{C} besitzt A drei Eigenwerte mit geometrischer Vielfachheit jeweils 1. \square

2.4. Charakteristisches Polynom.

SATZ 2.4.1. Sei \mathbb{K} ein Körper, $A \in \mathbb{K}^{n \times n}$ eine quadratische Matrix. Dann ist $\det(\lambda - A)$ ein Polynom vom Grad n mit höchstem Koeffizienten 1:

$$\det(\lambda - A) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0.$$

BEWEIS. Sei $A = (\alpha_{ij})$. Nach der Leibnizformel ist

$$\det(\lambda - A) \\ = \sum_{\sigma \in \mathcal{S}_n} \operatorname{sgn}(\sigma) (\lambda\delta_{1,\sigma(1)} - \alpha_{1,\sigma(1)}) \cdots (\lambda\delta_{n,\sigma(n)} - \alpha_{n,\sigma(n)}).$$

Jeder der Summanden in dieser Summe ist ein Polynom in λ , wobei der Grad höchstens n beträgt, und nur in einem Fall tatsächlich n erreicht, nämlich für die Permutation $\sigma = \operatorname{id}$. Dieser Summand ist

$$(\lambda - \alpha_{1,1}) \cdots (\lambda - \alpha_{n,n}) = \lambda^n + \sum_{j=0}^{n-1} g_j \lambda^j$$

mit geeigneten Koeffizienten g_j . \square

DEFINITION 2.4.2. Sei \mathbb{K} ein Körper und $A \in \mathbb{K}^{n \times n}$ eine quadratische Matrix.

- 1) Dann heißt das Polynom $\det(\lambda - A)$ das charakteristische Polynom von A .
- 2) Ist γ ein Eigenwert von A , dann ist die algebraische Vielfachheit von γ als Eigenwert von A die Vielfachheit von γ als Nullstelle des charakteristischen Polynoms.

Ohne Beweis führen wir eine Formel für die Koeffizienten des charakteristischen Polynoms an.

BEHAUPTUNG 2.4.3. Sei \mathbb{K} ein Körper, $A \in \mathbb{K}^{n \times n}$ eine quadratische Matrix. Für $J \subseteq \{1, \dots, n\}$ sei $\#J$ die Anzahl der Elemente von J , und A_J die quadratische Matrix, die man erhält, wenn man alle Zeilen und Spalten mit den Indices aus J aus der Matrix A streicht. Dann gilt

$$\begin{aligned} \det(\lambda - A) &= \sum_{k=0}^n c_k \lambda^k \quad \text{mit} \\ c_k &= (-1)^{n-k} \sum_{J \subseteq \{1, \dots, n\}, \#J=k} \det(A_J) \quad \text{für } 0 \leq k < n, \\ c_n &= 1. \end{aligned}$$

BEISPIEL 2.4.4. Wir berechnen $\det(\lambda - A)$ für

$$A = \begin{pmatrix} 2 & 1 & 3 \\ 4 & 4 & -1 \\ 0 & -1 & 5 \end{pmatrix}.$$

Lösung: Die Determinante von A bekommen wir mittels Regel von Sarrus:

$$\begin{aligned} c_0 &= (-1)^3 \begin{vmatrix} 2 & 1 & 3 \\ 4 & 4 & -1 \\ 0 & -1 & 5 \end{vmatrix} = -(40 - 12 - 2 - 20) = -6, \\ c_1 &= (-1)^2 \left(\begin{vmatrix} 4 & -1 \\ -1 & 5 \end{vmatrix} + \begin{vmatrix} 2 & 3 \\ 0 & 5 \end{vmatrix} + \begin{vmatrix} 2 & 1 \\ 4 & 4 \end{vmatrix} \right) = 19 + 10 + 4 = 33, \\ c_2 &= (-1)^1 (2 + 4 + 5) = -11. \end{aligned}$$

Daher ist

$$\det(\lambda - A) = \lambda^3 - 11\lambda^2 + 33\lambda - 6.$$

□

DEFINITION 2.4.5. Sei \mathbb{K} ein Körper und $A = (\alpha_{i,j})_{i,j=1 \dots n} \in \mathbb{K}^{n \times n}$ eine quadratische Matrix. Die Spur (engl.: trace) von A ist die Summe der Diagonalelemente:

$$\text{trace}(A) = \sum_{i=1}^n \alpha_{i,i}.$$

BEHAUPTUNG 2.4.6. Sei \mathbb{K} ein Körper und $A \in \mathbb{K}^{n \times n}$ eine quadratische Matrix mit charakteristischem Polynom

$$\lambda^n + c_{n-1}\lambda^{n-1} + \dots + c_0.$$

- (1) Dann ist $c_{n-1} = -\text{trace}(A)$ und $c_0 = (-1)^n \det(A)$.

(2) Hat das charakteristische Polynom (mit Vielfachheiten gezählt) n Eigenwerte $\gamma_1, \dots, \gamma_n$, dann ist

$$\text{trace}(A) = \sum_{i=1}^n \gamma_i, \quad \det(A) = \prod_{i=1}^n \gamma_i.$$

(Eigenwerte mit algebraischer Vielfachheit ≥ 2 werden in der Liste mehrfach aufgezählt.)

BEWEIS. Teil (1) ist die direkte Folge von Behauptung 2.4.3.

Teil (2): Es sei also das charakteristische Polynom von A

$$\det(\lambda - A) = (\lambda - \gamma_1)(\lambda - \gamma_2) \cdots (\lambda - \gamma_n).$$

Ausmultiplizieren und sortieren nach Potenzen von λ gibt für

$$\begin{aligned} -\text{trace}(A) &= c_{n-1} = -\gamma_1 - \gamma_2 - \cdots - \gamma_n, \\ (-1)^n \det(A) &= c_0 = (-\gamma_1)(-\gamma_2) \cdots (-\gamma_n). \end{aligned}$$

□

BEISPIEL 2.4.7. Bestimmen Sie die Eigenwerte, Eigenräume, und die algebraischen und geometrischen Vielfachheiten der Eigenwerte der Matrix A :

$$A = \begin{pmatrix} -1 & -1 & -3 & 0 \\ 0 & -1 & -6 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & 6 & 1 \end{pmatrix}$$

Lösung: Das charakteristische Polynom ermittelt man in diesem Fall am bequemsten durch den Laplaceschen Entwicklungssatz nach der vierten Spalte, dann hat man eine Dreiecksmatrix:

$$\begin{aligned} \begin{vmatrix} \lambda + 1 & 1 & 3 & 0 \\ 0 & \lambda + 1 & 6 & 0 \\ 0 & 0 & \lambda - 1 & 0 \\ 0 & -2 & -6 & \lambda - 1 \end{vmatrix} &= (\lambda - 1) \begin{vmatrix} \lambda + 1 & 1 & 3 \\ 0 & \lambda + 1 & 6 \\ 0 & 0 & \lambda - 1 \end{vmatrix} \\ &= (\lambda - 1)^2 (\lambda + 1)^2. \end{aligned}$$

Daher besitzt A die Eigenwerte ± 1 , jeweils mit der algebraischen Vielfachheit 2.

Wir berechnen jetzt den Eigenraum für $\lambda = 1$, also den Kern von $(1 - A)$, durch Pivotschritte:

$$\begin{pmatrix} 2 & 1 & 3 & 0 \\ 0 & 2 & 6 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & -6 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 2^* & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -2^* & -6 & 0 \end{pmatrix}$$

Die dritte und vierte Spalte bleiben nichtbasisch. Wir setzen $x_3 = s$, $x_4 = t$ und erhalten

$$\ker(1 - A) = \left\{ s \begin{pmatrix} 0 \\ -3 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \mid s, t \in \mathbb{R} \right\}.$$

Der Eigenraum ist zweidimensional, der Eigenwert 1 hat auch die geometrische Vielfachheit 2.

Letztlich berechnen wir den Eigenraum zu $\lambda = -1$, also den Kern von $(-1 - A)$:

$$\begin{pmatrix} 0 & 1* & 3 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & -2 & -6 & -2 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 1 & 3 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & -2* & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2* & 0 \\ 0 & 0 & 0 & -2* \end{pmatrix}$$

Nichtbasisch bleibt nur die erste Spalte, der Kern hat die Dimension 1:

$$\ker(-1 - A) = \left\{ t \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \mid t \in \mathbb{R} \right\}.$$

Der Eigenwert -1 hat die algebraische Vielfachheit 2 und die geometrische Vielfachheit 1. \square

3. Diagonalisierung

3.1. Lineare Unabhängigkeit von Eigenvektoren.

Das Matrizenprodukt ist nicht kommutativ, aber es gilt immer:

LEMMA 3.1.1. *Sei \mathbb{K} ein Körper, $A \in \mathbb{K}^{n \times n}$ eine quadratische Matrix, und $\lambda, \mu \in \mathbb{K}$. Dann gilt:*

$$(\lambda - A)(\mu - A) = (\mu - A)(\lambda - A).$$

BEWEIS. Wir müssen nur die Klammern ausmultiplizieren:

$$\begin{aligned} (\lambda - A)(\mu - A) &= \lambda\mu E - (\lambda E)A - A(\mu E) + A^2 \\ &= \mu\lambda E - A(\lambda E) - (\mu E)A + A^2 = (\mu - A)(\lambda - A). \end{aligned}$$

\square

LEMMA 3.1.2. *Sei \mathbb{K} ein Körper, $A \in \mathbb{K}^{n \times n}$ eine quadratische Matrix.*

(1) *Sind $\lambda \neq \mu \in \mathbb{K}$, $r, s \in \mathbb{N} \cup \{0\}$. Dann ist*

$$\ker(\lambda - A)^r \cap \ker(\mu - A)^s = \{0\}.$$

(2) *Sind $\lambda \neq \mu \in \mathbb{K}$, $r, s \in \mathbb{N} \cup \{0\}$. Dann ist*

$$\ker(\lambda - A)^r \subseteq \text{rg}(\mu - A)^s.$$

(3) *Sind $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ paarweise verschieden, $r_1, \dots, r_k \in \mathbb{N}$, sind $v_1, \dots, v_k \in \mathbb{K}^n \setminus \{0\}$ so, dass für alle $i = 1, \dots, k$ gilt*

$$(\lambda_i - A)^{r_i} v_i = 0.$$

Dann ist das System (v_1, \dots, v_k) linear unabhängig.

(4) *Sind $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ paarweise verschieden, $r_1, \dots, r_k \in \mathbb{N}$. Dann ist der Summenraum*

$$\ker(\lambda_1 - A)^{r_1} + \dots + \ker(\lambda_k - A)^{r_k}$$

eine direkte Summe.

BEWEIS.

- (1) Vollständige Induktion nach
- $r + s$
- .

Induktionsverankerung:

Für $r = 0$ ist $\ker(\lambda - A)^0 = \ker(E_n) = \{0\}$, für $s = 0$ ist $\ker(\mu - A)^0 = \{0\}$.In beiden Fällen ist die Behauptung trivial. Wir beweisen die Behauptung auch für $r = s = 1$. Sei $(\lambda - A)x = (\mu - A)x = 0$. Dann ist

$$x = \frac{1}{\lambda - \mu} [(\lambda - A)x - (\mu - A)x] = 0.$$

Induktionsschritt: Sei der Satz richtig, falls $r + s \leq m$, und sei nun

$$(\lambda - A)^r v = (\mu - A)^s v = 0$$

mit $r + s = m + 1$. OBdA sei $r > 1$. Wir setzen $w = (\lambda - A)v$. Es ist dann

$$(\lambda - A)^{r-1} w = (\lambda - A)^r v = 0,$$

$$(\mu - A)^s w = (\mu - A)^s (\lambda - A)v = (\lambda - A)(\mu - A)^s v = 0.$$

Da $(r-1) + s = m$ folgt aus der Induktionsannahme, dass $w = (\lambda - A)v = 0$. Nun ist $(\lambda - A)v = (\mu - A)^s v = 0$, und weil $1 + s \leq m$, gilt wieder nach der Induktionsannahme $v = 0$.

- (2) Vollständige Induktion nach
- $m = r + s$
- .

Induktionsverankerung: Für $r = 0$ ist $\ker(\lambda - A)^r = \{0\}$. Für $s = 0$ ist $\operatorname{rg}(\mu - A)^s = \mathbb{K}^n$. In beiden Fällen ist die Behauptung trivial. Für $r = s = 1$ und $x \in \ker(\lambda - A)$ ist

$$x = \frac{1}{\mu - \lambda} [(\mu - A)x - (\lambda - A)x] = \frac{1}{\mu - \lambda} (\mu - A)x \in \operatorname{rg}(\mu - A).$$

Induktionsschritt von m auf $m + 1$. Nun sei $\ker(\lambda - A)^r \subseteq \operatorname{rg}(\mu - A)^s$ für alle $r + s \leq m$. Sei nun $r + s = m + 1$ und $x \in \ker(\lambda - A)^r$. Es ist $r + s - 1 = m$. Nach Induktionsvoraussetzung ist also $x \in \operatorname{rg}(\mu - A)^{s-1}$ und daher $(\mu - A)x \in \operatorname{rg}(\mu - A)^s$. Andererseits ist $(\lambda - A)x \in \ker(\lambda - A)^{r-1}$, und nach Induktionsvoraussetzung ist $(\lambda - A)x \in \operatorname{rg}(\mu - A)^s$. Daher ist

$$x = \frac{1}{\mu - \lambda} [(\mu - A)x - (\lambda - A)x] \in \operatorname{rg}(\mu - A)^s.$$

- (3) Vollständige Induktion nach
- k
- . Für
- $k = 1$
- ist die Aussage trivial, weil
- $v_1 \neq 0$
- .

Induktionsschritt: Sei der Satz richtig für k , und seien für $i = 1, \dots, k + 1$ die Vektoren $v_i \in \ker((\lambda_i - A)^{r_i}) \setminus \{0\}$. Für $i = 1, \dots, k$ setzen wir $w_i = (\lambda_{k+1} - A)^{r_{k+1}} v_i$. Wäre eines der $w_i = 0$, so wäre

$$(\lambda_{k+1} - A)^{r_{k+1}} v_i = w_i = 0 = (\lambda_i - A)^{r_i} v_i,$$

und wegen Teil (1) wäre dann $v_i = 0$ im Widerspruch zur Annahme. Daher sind alle $w_i \neq 0$. Es ist dann für $i = 1 \dots k$

$$\begin{aligned} (\lambda_i - A)^{r_i} w_i &= (\lambda_i - A)^{r_i} (\lambda_{k+1} - A)^{r_{k+1}} v_i \\ &= (\lambda_{k+1} - A)^{r_{k+1}} (\lambda_i - A)^{r_i} v_i = 0, \end{aligned}$$

also ist $w_i \in \ker((\lambda_i - A)^{r_i})$.

Sei nun

$$\sum_{i=1}^{k+1} \alpha_i v_i = 0.$$

Multiplikation mit $(\lambda_{r+1} - A)^{r_{k+1}}$ ergibt

$$\sum_{i=1}^k \alpha_i w_i + 0 = 0,$$

und weil (w_1, \dots, w_k) wegen der Induktionsannahme linear unabhängig sind, folgt $\alpha_1 = \dots = \alpha_k = 0$. Weil $v_{k+1} \neq 0$, muss nun auch $\alpha_{k+1} = 0$ sein, und das System (v_1, \dots, v_{k+1}) ist linear unabhängig.

- (4) Wegen Teil (2) und Behauptung 4.2.5 ist der Summenraum

$$\ker((\lambda_1 - A)^{r_1}) + \dots + \ker((\lambda_k - A)^{r_k})$$

eine direkte Summe. □

KOROLLAR 3.1.3. Sei \mathbb{K} ein Körper und $A \in \mathbb{K}^{n \times n}$ eine quadratische Matrix. Dann gilt:

- 1) Jedes System (v_1, \dots, v_k) von Eigenvektoren von A zu paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_k$ ist linear unabhängig.
- 2) Seien $\lambda_1, \dots, \lambda_k$ die Eigenwerte von A mit den geometrischen Vielfachheiten m_1, \dots, m_k . Der Raum \mathbb{K}^n besitzt genau dann eine Basis aus Eigenvektoren von A , wenn $m_1 + \dots + m_k = n$.
- 3) Wenn es n verschiedene Eigenwerte $\lambda_1, \dots, \lambda_n$ von A gibt, und v_1, \dots, v_n Eigenvektoren zu $\lambda_1, \dots, \lambda_n$ sind, dann ist (v_1, \dots, v_n) eine Basis von \mathbb{K}^n .

BEWEIS. Teil (1): Lemma 3.1.2) mit $r_1 = \dots = r_k = 1$.

Teil (2): Seien $W_i = \ker((\lambda_i - A))$ die Eigenräume von A . Wegen Lemma 3.1.2 (3) ist die Summe $W_1 + \dots + W_k$ direkt. Wegen Behauptung 4.2.5 ist die Dimension des Summenraums gleich $m_1 + \dots + m_k$. Insbesondere ist $W_1 + \dots + W_k = \mathbb{K}^n$ genau dann wenn $m_1 + \dots + m_k = n$.

Besitzt \mathbb{K}^n eine Basis aus Eigenvektoren, so ist offensichtlich $\mathbb{K}^n = W_1 + \dots + W_k$. Sei umgekehrt $\mathbb{K}^n = W_1 + \dots + W_k$. Für jeden Eigenraum W_i wählen wir eine Basis $(v_{i,1}, \dots, v_{i,m_i})$, diese besteht natürlich aus Eigenvektoren zum Eigenwert λ_i . Wegen Behauptung 4.2.5 ist die Vereinigung der Basen, also $(v_{1,1}, \dots, v_{k,m_k})$, eine Basis des Summenraums, folglich von \mathbb{K}^n .

Teil (3): Wegen Teil (1) sind (v_1, \dots, v_n) linear unabhängig, und da es n Vektoren sind, bilden sie eine Basis des \mathbb{K}^n . □

Versuchen Sie selbst, Korollar 3.1.3 direkt zu beweisen, ohne auf Lemma 3.1.2 zurückzugreifen. Das ist etwas einfacher, weil keine Potenzen r_i ins Spiel kommen, und es würde auch für diesen Abschnitt völlig ausreichen. Später werden wir aber brauchen können, dass wir mit Kernen von $(\lambda - A)^r$ statt nur von $(\lambda - A)$ arbeiten können.

3.2. Diagonalisierung einer Matrix.

DEFINITION 3.2.1. Sei \mathbb{K} ein Körper und $A \in \mathbb{K}^{n \times n}$ eine quadratische Matrix. A heißt genau dann diagonalisierbar, wenn es eine reguläre Matrix T und eine Diagonalmatrix D gibt, sodass $T^{-1}AT = D$ (d.h., wenn A ähnlich zu einer Diagonalmatrix ist).

SATZ 3.2.2. Sei \mathbb{K} ein Körper und $A \in \mathbb{K}^{n \times n}$. Sei T eine reguläre Matrix mit den Spalten v_1, \dots, v_n , und sei D eine Diagonalmatrix, auf deren Diagonalen die Elemente $\lambda_1, \dots, \lambda_n$ (nicht unbedingt paarweise verschieden) stehen. Dann gilt: $D = T^{-1}AT$ genau dann, wenn jedes v_i ein Eigenvektor von A zum Eigenwert λ_i ist.

Es ist also A genau dann diagonalisierbar, wenn \mathbb{K}^n eine Basis aus Eigenvektoren von A hat.

BEWEIS. Weil T regulär ist, ist $T^{-1}AT = D$ genau dann, wenn $AT = TD$. (Wir müssen nur von links mit T multiplizieren.) Die Spalten von AT sind Av_1, \dots, Av_n , die Spalten von TD sind $\lambda_1 v_1, \dots, \lambda_n v_n$. Also ist $T^{-1}AT = D$ genau dann, wenn $Av_i = \lambda_i v_i$ für alle i gilt. □

KOROLLAR 3.2.3. Sei \mathbb{K} ein Körper und $A \in \mathbb{K}^{n \times n}$. Sei $f : \mathbb{K}^n \rightarrow \mathbb{K}^n$ der Endomorphismus $f(x) = Ax$. Sei $B = (v_1, \dots, v_n)$ eine Basis des \mathbb{K}^n aus Eigenvektoren von A zu den Eigenwerten $\lambda_1, \dots, \lambda_n$. Mit $E = (e_1, \dots, e_n)$ bezeichnen wir die Einheitsbasis des \mathbb{K}^n . Dann gilt:

- 1) Die Spalten der Übergangsmatrix T_E^B sind die Eigenvektoren v_1, \dots, v_n .
- 2) Die Matrix des Endomorphismus f bezüglich der Basis B ist

$$M_B^B(f) = (T_E^B)^{-1} A T_E^B = D.$$

BEWEIS. Die Spalten von T_E^B sind die Basisvektoren von B , in Einheitskoordinaten geschrieben, also v_1, \dots, v_n . Weil diese Eigenvektoren zu den Eigenwerten $\lambda_1, \dots, \lambda_n$ sind, gilt wegen Satz 3.2.2: $(T_E^B)^{-1} A T_E^B = D$. \square

BEISPIEL 3.2.4. Diagonalisieren Sie die Matrix

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

Lösung: Berechnung des charakteristischen Polynoms und der Eigenwerte.

$$\begin{aligned} c_2 &= -\text{trace}(A) = -(-1 - 3 + 5) = -1, \\ c_1 &= \begin{vmatrix} -1 & 2 \\ 4 & -3 \end{vmatrix} + \begin{vmatrix} -1 & 4 \\ -2 & 5 \end{vmatrix} + \begin{vmatrix} -3 & -8 \\ 2 & 5 \end{vmatrix} = -5 + 3 + 1 = -1, \\ c_0 &= -\det(A) = 1. \end{aligned}$$

Damit ist das charakteristische Polynom

$$\lambda^3 - \lambda^2 - \lambda + 1 = (\lambda - 1)(\lambda^2 - 1) = (\lambda - 1)^2(\lambda + 1).$$

Daher hat A die Eigenwerte -1 mit algebraischer Vielfachheit 1 und $+1$ mit algebraischer Vielfachheit 2.

Eigenräume: Berechnung von $\ker(-1 - A)$.

$$\begin{aligned} (-1 - A) &= \begin{pmatrix} -1 + 1 & -2 & -4 \\ -4 & -1 + 3 & 8 \\ 2 & -2 & -1 - 5 \end{pmatrix} = \begin{pmatrix} 0 & -2 & -4 \\ -4 & 2 & 8 \\ 2 & -2 & -6 \end{pmatrix}. \\ \begin{pmatrix} 0 & -2 & -4 \\ -4 & 2 & 8 \\ 2 & -2 & -6 \end{pmatrix} &\rightarrow \begin{pmatrix} 0 & -2 & -4 \\ 0 & -2 & -4 \\ 2 & -2 & -6 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 0 & 0 \\ 0 & -2 & -4 \\ 2 & 0 & -2 \end{pmatrix}. \end{aligned}$$

Nichtbasisch bleibt nur die dritte Variable, die wir nun mit 1 ansetzen. Wir erhalten

$$\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

Berechnung von $\ker(1 - A)$.

$$\begin{aligned} (1 - A) &= \begin{pmatrix} 1 + 1 & -2 & -4 \\ -4 & 1 + 3 & 8 \\ 2 & -2 & 1 - 5 \end{pmatrix} = \begin{pmatrix} 2 & -2 & -4 \\ -4 & 4 & 8 \\ 2 & -2 & -4 \end{pmatrix}. \\ \begin{pmatrix} 2 & -2 & -4 \\ -4 & 4 & 8 \\ 2 & -2 & -4 \end{pmatrix} &\rightarrow \begin{pmatrix} 2 & -2 & -4 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Nichtbasisch bleiben die zweite und dritte Variable. Wir setzen jeweils eine davon mit 1, die andere mit 0 an, und erhalten eine Basis des Kerns:

$$\left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \right).$$

Diagonalisierung: Damit besitzt \mathbb{R}^3 eine Basis aus 3 linear unabhängigen Eigenvektoren von A :

$$B = \left(\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \right).$$

Die Übergangsmatrix T_E^B ist dann die Matrix, deren Spalten die Basisvektoren sind:

$$T_E^B = \begin{pmatrix} 1 & 1 & 2 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Aus den Eigenwerten ergibt sich die Diagonalmatrix

$$D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Diese Matrix stellt den Endomorphismus $f(x) = Ax$ bezüglich der Basis B dar. Man kann sich leicht durch direktes Nachrechnen (MATLAB?) überzeugen, dass tatsächlich gilt:

$$M_B^B(f) = (T_E^B)^{-1} A T_E^B = D.$$

□

SATZ 3.2.5. *Ähnliche Matrizen haben dasselbe charakteristische Polynom.*

BEWEIS. Weil $\det(T) \det(T^{-1}) = 1$, gilt

$$\begin{aligned} \det(\lambda - T^{-1}AT) &= \det(\lambda T^{-1}T - T^{-1}AT) = \det(T^{-1}(\lambda - A)T) \\ &= \det(T^{-1}) \det(\lambda - A) \det(T) = \det(\lambda - A). \end{aligned}$$

□

3.3. Anwendungsbeispiele.

BEISPIEL 3.3.1. Berechnen Sie die Potenzen A^m für $m \in \mathbb{N}$ und

$$A = \begin{pmatrix} 5 & -2 \\ 7 & -4 \end{pmatrix}.$$

Lösung: Wenn A diagonalisierbar sein sollte, also $D = T^{-1}AT$, so wäre $A = TDT^{-1}$ und

$$\begin{aligned} A^m &= (TDT^{-1})(TDT^{-1}) \cdots (TDT^{-1}) \\ &= TD(T^{-1}T)D(T^{-1}T) \cdots (T^{-1}T)DT^{-1} = TD^mT^{-1}. \end{aligned}$$

Die Potenzen einer Diagonalmatrix sind aber leicht zu berechnen:

$$\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}^m = \begin{pmatrix} \lambda_1^m & 0 \\ 0 & \lambda_2^m \end{pmatrix}.$$

Wir versuchen also, A zu diagonalisieren.

Das charakteristische Polynom ist

$$\lambda^2 - (5 - 4)\lambda + \begin{vmatrix} 5 & -2 \\ 7 & -4 \end{vmatrix} = \lambda^2 - \lambda - 6.$$

Die Eigenwerte sind $\lambda_1 = -2$ und $\lambda_2 = 3$. Weil A also zwei verschiedene Eigenwerte besitzt, hat \mathbb{R}^2 eine Basis aus Eigenvektoren, und A ist diagonalisierbar. Wir berechnen die Eigenräume:

$$\begin{aligned}\ker(-2 - A) &= \ker \begin{pmatrix} -7 & 2 \\ -7 & 2 \end{pmatrix} = \mathcal{L}\left(\begin{pmatrix} 2 \\ 7 \end{pmatrix}\right), \\ \ker(3 - A) &= \ker \begin{pmatrix} -2 & 2 \\ -7 & 7 \end{pmatrix} = \mathcal{L}\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right).\end{aligned}$$

Wir haben also $D = T^{-1}AT$ mit

$$D = \begin{pmatrix} -2 & 0 \\ 0 & 3 \end{pmatrix}, \quad T = \begin{pmatrix} 2 & 1 \\ 7 & 1 \end{pmatrix}, \quad T^{-1} = \begin{pmatrix} -0,2 & 0,2 \\ 1,4 & -0,4 \end{pmatrix}.$$

Dann ist

$$\begin{aligned}A^m &= TD^mT^{-1} = \begin{pmatrix} 2 & 1 \\ 7 & 1 \end{pmatrix} \begin{pmatrix} (-2)^m & 0 \\ 0 & 3^m \end{pmatrix} \begin{pmatrix} -0,2 & 0,2 \\ 1,4 & -0,4 \end{pmatrix} \\ &= (-2)^m \begin{pmatrix} 2 & 1 \\ 7 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} -0,2 & 0,2 \\ 1,4 & -0,4 \end{pmatrix} + 3^m \begin{pmatrix} 2 & 1 \\ 7 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -0,2 & 0,2 \\ 1,4 & -0,4 \end{pmatrix} \\ &= (-2)^m \begin{pmatrix} -0,4 & 0,4 \\ -1,4 & 1,4 \end{pmatrix} + 3^m \begin{pmatrix} 1,4 & -0,4 \\ 1,4 & -0,4 \end{pmatrix}.\end{aligned}$$

□

Aus dem obigen Beispiel sehen wir

BEMERKUNG 3.3.2. Ist eine $(n \times n)$ -Matrix A diagonalisierbar mit Eigenwerten $\lambda_1, \dots, \lambda_k$, so ist $A^m = \lambda_1^m C_1 + \dots + \lambda_k^m C_k$ mit Matrizen C_1, \dots, C_k , welche von m nicht abhängen.

BEISPIEL 3.3.3. Lösen Sie das folgende System von zwei Differentialgleichungen in zwei unbekanntenen Funktionen:

$$\begin{aligned}\frac{d}{dt}x_1(t) &= 5x_1(t) - 2x_2(t), & x_1(0) &= 1, \\ \frac{d}{dt}x_2(t) &= 7x_1(t) - 4x_2(t), & x_2(0) &= 2.\end{aligned}$$

Lösung: In Matrixform geschrieben gibt das System eine Differentialgleichung für eine Funktion $x: \mathbb{R} \rightarrow \mathbb{R}^2$:

$$\frac{d}{dt}x(t) = \begin{pmatrix} 5 & -2 \\ 7 & -4 \end{pmatrix} x(t), \quad x(0) = x_0 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Die Matrix aus dem Gleichungssystem ist die Matrix A aus Beispiel 3.3.1. Wir wissen aus diesem Beispiel, dass A diagonalisierbar ist: $D = T^{-1}AT$ mit

$$D = \begin{pmatrix} -2 & 0 \\ 0 & 3 \end{pmatrix}, \quad T = \begin{pmatrix} 2 & 1 \\ 7 & 1 \end{pmatrix}, \quad T^{-1} = \begin{pmatrix} -0,2 & 0,2 \\ 1,4 & -0,4 \end{pmatrix}.$$

Wir setzen nun

$$y(t) := T^{-1}x(t), \quad \text{also } x(t) = Ty(t).$$

Nun ist

$$\frac{d}{dt}y(t) = \frac{d}{dt}T^{-1}x(t) = T^{-1}Ax(t) = T^{-1}ATy(t) = Dy(t).$$

Wir erhalten also das System

$$\frac{d}{dt}y(t) = \begin{pmatrix} -2 & 0 \\ 0 & 3 \end{pmatrix} y(t), \quad y(0) = T^{-1}x(0) = \begin{pmatrix} 0,2 \\ 0,6 \end{pmatrix}.$$

Ausgeschrieben ergeben sich zwei entkoppelte (das heißt, voneinander unabhängige) Differentialgleichungen:

$$\begin{aligned}\frac{d}{dt}y_1(t) &= -2y_1(t), & y_1(0) &= 0, 2, \\ \frac{d}{dt}y_2(t) &= 3y_2(t), & y_2(0) &= 0, 6.\end{aligned}$$

Die Lösungen sind

$$y_1(t) = 0, 2 e^{-2t}, \quad y_2(t) = 0, 6 e^{3t}.$$

Dann ist

$$x(t) = Ty(t) = \begin{pmatrix} 2 & 1 \\ 7 & 1 \end{pmatrix} \begin{pmatrix} 0, 2 e^{-2t} \\ 0, 6 e^{3t} \end{pmatrix} = e^{-2t} \begin{pmatrix} 0, 4 \\ 1, 4 \end{pmatrix} + e^{3t} \begin{pmatrix} 0, 6 \\ 0, 6 \end{pmatrix}.$$

□

Aus dem obigen Beispiel sehen wir:

BEMERKUNG 3.3.4. Ist A diagonalisierbar mit Eigenwerten $\lambda_1, \dots, \lambda_k$, so schreibt sich jede Lösung der Differentialgleichung $x'(t) = Ax(t)$ in der Form

$$x(t) = e^{\lambda_1 t} v_1 + \dots + e^{\lambda_k t} v_k$$

mit Vektoren $v_i \in \ker(\lambda_i - A)$, welche von t nicht abhängen.

4. Hauptraumzerlegung

4.1. Invariante Unterräume.

DEFINITION 4.1.1.

- 1) Sei V ein Vektorraum, W ein Unterraum von V , und $f : V \rightarrow V$ ein Endomorphismus. Dann heißt W invariant bezüglich f , wenn gilt: $f(W) \subseteq W$.
- 2) Sei \mathbb{K} ein Körper, W ein Unterraum von \mathbb{K}^n , und $A \in \mathbb{K}^{n \times n}$. Dann heißt W invariant bezüglich A , wenn gilt: $(\forall w \in W) Aw \in W$.

BEMERKUNG 4.1.2. Sei $A \in \mathbb{K}^{n \times n}$.

- (1) Die folgenden Räume sind invariant bezüglich A :

$$\{0\}, \mathbb{K}^n, \ker(A), \operatorname{rg}(A).$$

- (2) Ist W invariant bezüglich A , so ist W auch invariant bezüglich jeder Potenz A^m und bezüglich $(\lambda - A)$ für jeden Skalar λ .
- (3) Sind W_1, \dots, W_k invariante Unterräume bezüglich A . Dann sind auch Summenraum $W_1 + \dots + W_k$ und Durchschnitt $\bigcap_{i=1}^k W_i$ invariante Unterräume bezüglich A .

LEMMA 4.1.3. *Ist V ein Vektorraum, $f : V \rightarrow V$ ein Endomorphismus, und W ein f -invarianter Unterraum von V , dann ist die Einschränkung*

$$f|_W: \begin{cases} W & \rightarrow W \\ x & \mapsto f(x) \end{cases}$$

ein Endomorphismus auf W .

BEWEIS. Weil f den Raum W nach Voraussetzung in W abbildet, ist $f|_W$ als Abbildung von W nach W wohldefiniert. Die Linearität erbt die Einschränkung natürlich von f . □

SATZ 4.1.4. Sei V ein endlich dimensionaler Vektorraum, $f : V \rightarrow V$ ein Endomorphismus. Seien W_1, \dots, W_k f -invariante Unterräume von V , sodass V die direkte Summe ist: $V = W_1 + \dots + W_k$. Mit f_i bezeichnen wir die Einschränkung von f auf den Unterraum W_i :

$$f_i : \begin{cases} W_i & \rightarrow W_i \\ x & \mapsto f(x) \end{cases}.$$

Für $i = 1, \dots, k$ sei $B_i = (b_{i,1}, \dots, b_{i,m_i})$ eine Basis von W_i , und es sei C_i die Matrix des Homomorphismus f_i bezüglich der Basis B_i . Es sei B die Basis $(b_{1,1}, b_{1,2}, \dots, b_{k,m_k})$ von V , die sich als Vereinigung der Basen der W_i ergibt.

Dann hat die Matrix $M_B^B(f)$ von f bezüglich B die Blockdiagonalform

$$\begin{pmatrix} C_1 & 0 & \cdots & 0 \\ 0 & C_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & C_k \end{pmatrix}.$$

BEWEIS. Der Übersicht halber indizieren wir die Zeilen und Spalten der Matrix C so wie die Basisvektoren von B mit Doppelindices.

Die Spalten von $M_B^B(f)$ sind die Bilder der Basisvektoren $b_{1,1}, \dots, b_{k,m_k}$, geschrieben in Koordinaten bezüglich B . Jeder Basisvektor $b_{i,j}$ liegt in W_i . Daher ist $f(b_{i,j}) = f_i(b_{i,j})$. Es sei

$$\begin{pmatrix} \gamma_{i,j,1} \\ \vdots \\ \gamma_{i,j,m_i} \end{pmatrix}$$

die j -te Spalte der Matrix C_i . Dann ist

$$f_i(b_{i,j}) = \sum_{q=1}^{m_i} \gamma_{i,j,q} b_{i,q} = \sum_{q=1}^{m_i} \gamma_{i,j,q} b_{i,q} + \sum_{p \neq i} \sum_{q=1}^{m_p} 0 b_{p,q}.$$

Daher ist der Eintrag der (i, j) -ten Spalte von C in der Zeile (p, q) gerade

$$\begin{cases} \gamma_{i,j,q} & \text{wenn } i = p \\ 0 & \text{wenn } i \neq p \end{cases}.$$

□

4.2. Lemma von Fitting.

SATZ 4.2.1 (Lemma von Fitting). Sei \mathbb{K} ein Körper, $F \in \mathbb{K}^{n \times n}$ eine quadratische Matrix. Dann gilt:

(1)

$$\begin{aligned} \{0\} &\subseteq \ker(F) \subseteq \ker(F^2) \subseteq \cdots, \\ \mathbb{K}^n &\supseteq \operatorname{rg}(F) \supseteq \operatorname{rg}(F^2) \supseteq \cdots. \end{aligned}$$

(2) Es gibt ein $r \leq n$ sodass $\ker(F^r) = \ker(F^{r+1})$.

(3) Sei $r \in \mathbb{N}$. Folgende Aussagen sind äquivalent:

- (i) $\ker(F^r) = \ker(F^{r+1})$,
- (ii) $(\forall s > r) \ker(F^r) = \ker(F^s)$,
- (iii) $\ker(F^r) \cap \operatorname{rg}(F^r) = \{0\}$,
- (iv) $\operatorname{rg}(F^r) = \operatorname{rg}(F^{r+1})$,
- (v) $(\forall s > r) \operatorname{rg}(F^r) = \operatorname{rg}(F^s)$,
- (vi) $\mathbb{K}^n = \ker(F^r) + \operatorname{rg}(F^r)$.

- (4) Wenn die Aussagen aus Punkt (3) zutreffen, dann ist \mathbb{K}^n die direkte Summe $\mathbb{K}^n = \ker(F^r) + \operatorname{rg}(F^r)$.

BEWEIS.

- (1) Ist $x \in \ker(F^i)$, so ist $F^{i+1}x = FF^ix = F0 = 0$, also $x \in \ker(F^{i+1})$. Ist $x \in \operatorname{rg}(F^{i+1})$, so gibt es ein $y \in \mathbb{K}^n$ sodass $x = F^{i+1}y$. Dann ist aber $x = F^iFy \in \operatorname{rg}(F^i)$.
- (2) Wäre $\ker(F^r) \neq \ker(F^{r+1})$ für $r = 0, \dots, n$, so wäre

$$0 < \dim(\ker(F)) < \dim(\ker(F^2)) < \dots < \dim(\ker(F^{n+1})),$$

und es würde folgen $\dim(\ker(F^{n+1})) > n$, was für einen Unterraum des \mathbb{K}^n nicht möglich ist.

(3)

- (i) \Rightarrow (ii): Wir nehmen an, dass (ii) nicht gilt. Sei s die kleinste Zahl für die gilt: $s > r$ und $\ker(F^s) \neq \ker(F^r)$. Sei $x \in \ker(F^s) \setminus \ker(F^r)$. Dann ist $F(x) \in \ker(F^{s-1}) = \ker(F^r)$, und daher ist $x \in \ker(F^{r+1}) = \ker(F^r)$ im Widerspruch zur Wahl von x .
- (ii) \Rightarrow (iii): Angenommen, (ii) gilt, aber (iii) gilt nicht. Dann gibt es ein $y \neq 0$ mit $y \in \ker(F^r) \cap \operatorname{rg}(F^r)$. Weil $y \in \operatorname{rg}(F^r)$, gibt es ein $x \in \mathbb{K}^n$ mit $F^r x = y$. Nun ist $F^{2r}x = F^r y = 0$ weil $y \in \ker(F^r)$. Weil aber $\ker(F^{2r}) = \ker(F^r)$, folgt $y = F^r x = 0$ im Widerspruch zur Wahl von y .
- (iii) \Rightarrow (i): Angenommen, (i) gilt nicht. Dann gibt es also ein $x \in \mathbb{K}^n$ mit $F^{r+1}x = 0$, aber $F^r x \neq 0$. Wir setzen $y = F^r x \neq 0$. Offensichtlich ist $y \in \operatorname{rg}(F^r)$, und weil $Fy = F^{r+1}x = 0$, ist erst recht $F^r y = 0$, und $y \in \ker(F^r)$. Daher ist $\ker(F^r) \cap \operatorname{rg}(F^r) \neq \{0\}$.
- (i) \Leftrightarrow (iv): Da $\ker(F^r) \subseteq \ker(F^{r+1})$ und $\operatorname{rg}(F^{r+1}) \subseteq \operatorname{rg}(F^r)$, gilt

$$\ker(F^r) = \ker(F^{r+1}) \iff \dim(\ker(F^r)) = \dim(\ker(F^{r+1})),$$

$$\operatorname{rg}(F^r) = \operatorname{rg}(F^{r+1}) \iff \dim(\operatorname{rg}(F^r)) = \dim(\operatorname{rg}(F^{r+1})).$$

Nun gilt aber $\dim(\operatorname{rg}(F^r)) = n - \dim(\ker(F^r))$, sodass

$$\begin{aligned} \dim(\ker(F^r)) &= \dim(\ker(F^{r+1})) \\ \Leftrightarrow \dim(\operatorname{rg}(F^r)) &= \dim(\operatorname{rg}(F^{r+1})). \end{aligned}$$

(ii) \Leftrightarrow (v): ebenso.

(iii) \Leftrightarrow (vi): Nach der Dimensionsformel ist

$$\begin{aligned} &\dim(\ker(F^r) + \operatorname{rg}(F^r)) + \dim(\ker(F^r) \cap \operatorname{rg}(F^r)) \\ &= \dim(\ker(F^r)) + \dim(\operatorname{rg}(F^r)) = n. \end{aligned}$$

Daher ist

$$\begin{aligned} \ker(F^r) + \operatorname{rg}(F^r) &= \mathbb{K}^n \iff \dim(\ker(F^r) + \operatorname{rg}(F^r)) = n \\ \iff \dim(\ker(F^r) \cap \operatorname{rg}(F^r)) &= 0 \iff \ker(F^r) \cap \operatorname{rg}(F^r) = \{0\}. \end{aligned}$$

- (4) Wenn die Aussagen aus Teil (3) zutreffen, ist $\ker(F^r) \cap \operatorname{rg}(F^r) = \{0\}$ und daher ist die Summe $\mathbb{K}^n = \ker(F^r) + \operatorname{rg}(F^r)$ direkt.

□

4.3. Haupträume.

DEFINITION 4.3.1. Sei \mathbb{K} ein Körper, $A \in \mathbb{K}^{n \times n}$, $\lambda \in \mathbb{K}$ ein Eigenwert von A . Der Raum $\ker(\lambda - A)^n$ heißt der verallgemeinerte Eigenraum oder Hauptraum von A zum Eigenwert λ .

BEMERKUNG 4.3.2. Wegen Satz 4.2.1 ist der Hauptraum von A zum Eigenwert λ zugleich der größtmögliche Raum, der sich als $\ker(\lambda - A)^r$ mit irgendeiner Potenz r schreiben läßt. Wenn gilt

$$0 \subsetneq \ker(\lambda - A) \subsetneq \ker(\lambda - A)^2 \subsetneq \cdots \subsetneq \ker(\lambda - A)^r = \ker(\lambda - A)^{r+1},$$

dann ist $\ker(\lambda - A)^r$ der Hauptraum.

SATZ 4.3.3. Sei \mathbb{K} ein Körper, $A \in \mathbb{K}^{n \times n}$ und f der zugehörige Endomorphismus. Seien $\lambda_1, \dots, \lambda_k$ paarweise verschiedene Eigenwerte von A , und $\ker(\lambda_i - A)^{r_i}$ der Hauptraum zu λ_i .

Dann ist \mathbb{K}^n die direkte Summe

$$\mathbb{K}^n = \ker(\lambda_1 - A)^{r_1} + \cdots + \ker(\lambda_k - A)^{r_k} + \bigcap_{i=1}^k \operatorname{rg}(\lambda_i - A)^{r_i}.$$

BEWEIS. Sei $x \in \mathbb{K}^n$. Nach dem Lemma von Fitting ist \mathbb{K}^n die direkte Summe von $\ker(\lambda_i - A)^{r_i}$ und $\operatorname{rg}(\lambda_i - A)^{r_i}$, also gibt es $x_i \in \ker(\lambda_i - A)^{r_i}$ so, dass $x - x_i \in \operatorname{rg}(\lambda_i - A)^{r_i}$. Wir betrachten nun

$$y = x - x_1 - x_2 - \cdots - x_k = x - x_i - \sum_{j \neq i} x_j.$$

Nun ist

$$\begin{aligned} (x - x_i) &\in \operatorname{rg}(\lambda_i - A)^{r_i}, \\ x_j &\in \ker(\lambda_j - A)^{r_j} \subseteq \operatorname{rg}(\lambda_i - A)^{r_i} \end{aligned}$$

wegen Lemma 3.1.2. Also ist $y \in \operatorname{rg}(\lambda_i - A)^n$ für alle i . Es ist daher

$$\mathbb{K}^n = \ker(\lambda_1 - A)^{r_1} + \cdots + \ker(\lambda_k - A)^{r_k} + \bigcap_{i=1}^k \operatorname{rg}(\lambda_i - A)^n.$$

Wir zeigen nun, dass die Summe direkt ist. Es sei also

$$x = \sum_{i=1}^k x_i + y$$

eine Zerlegung von x in $x_i \in \ker(\lambda_i - A)^{r_i}$ und $y \in \bigcap_{i=1}^k \operatorname{rg}(\lambda_i - A)^{r_i}$. Wir zeigen, dass die x_i eindeutig bestimmt sind. Es folgt dann daraus, dass auch y eindeutig bestimmt ist. Wegen Lemma 3.1.2 ist für $j \neq i$ der Vektor $x_j \in \ker(\lambda_j - A)^{r_j} \subseteq \operatorname{rg}(\lambda_i - A)^{r_i}$, und daher ist

$$x = x_i + (y + \sum_{j \neq i} x_j)$$

eine Zerlegung von x in einen Vektor in $\ker(\lambda_i - A)^{r_i}$ und einen Vektor in $\operatorname{rg}(\lambda_i - A)^{r_i}$. Diese Zerlegung ist aber nach dem Lemma von Fitting eindeutig, und daher ist x_i eindeutig bestimmt. \square

SATZ 4.3.4 (Hauptraumzerlegung). Sei \mathbb{K} ein algebraisch abgeschlossener Körper und $A \in \mathbb{K}^{n \times n}$. Seien $\lambda_1, \dots, \lambda_k$ die Eigenwerte von A und $\ker(\lambda_i - A)^{r_i}$ der Hauptraum zu λ_i . Dann ist \mathbb{K}^n die direkte Summe

$$\mathbb{K}^n = \ker(\lambda_1 - A)^n + \cdots + \ker(\lambda_k - A)^n.$$

BEWEIS. Nach Satz 4.3.3 ist

$$\mathbb{K}^n = \ker(\lambda_1 - A)^{r_1} + \cdots + \ker(\lambda_k - A)^{r_k} + W$$

mit

$$W = \bigcap_{i=1}^k \operatorname{rg}(\lambda_i - A)^{r_k}.$$

Wir nehmen nun an, dass $W \neq \{0\}$ und zielen auf einen Widerspruch.

Jede Matrix hat ihr charakteristisches Polynom, und über einem algebraisch abgeschlossenen Körper besitzt dieses eine Nullstelle. Es kann also in einem endlichdimensionalen Vektorraum über \mathbb{K} (außer dem Nullraum) keinen Endomorphismus ohne Eigenwert geben. Die lineare Abbildung

$$f : \begin{cases} W & \rightarrow W \\ x & \mapsto Ax \end{cases}$$

muss dann einen Eigenwert μ besitzen. Sei $y \in W$ ein dazugehöriger Eigenvektor, also insbesondere $y \neq 0$. Dann ist aber $(\mu - A)y = \mu y - f(y) = 0$, also ist y auch ein Eigenvektor von A , und μ ist einer der Eigenwerte λ_i von A . Dann ist aber $y \in \ker(\lambda_i - A)^n$, und wegen der Direktheit der Summe

$$\mathbb{K}^n = \ker(\lambda_1 - A)^n + \cdots + \ker(\lambda_k - A)^n + W.$$

ist $y = 0$ im Widerspruch zur Wahl von y . Also ist W der Nullraum. \square

BEHAUPTUNG 4.3.5. Sei \mathbb{K} algebraisch abgeschlossen, $A \in \mathbb{K}^{n \times n}$ eine Matrix und $f : x \mapsto Ax$ die zugehörige lineare Abbildung. Die Eigenwerte von A seien $\lambda_1, \dots, \lambda_k$ mit den Haupträumen $H_i = \ker(\lambda_i - A)^{r_i}$. Für $i = 1, \dots, k$ sei $(b_{i,1}, \dots, b_{i,m_i})$ eine Basis von H_i .

(a) Bezüglich der Basis

$$B = (b_{1,1}, \dots, b_{1,m_1}, \dots, b_{k,1}, \dots, b_{k,m_k})$$

hat f die Blockdiagonalmatrix

$$M_B^B(f) = \begin{pmatrix} C_1 & 0 & \cdots & 0 \\ 0 & C_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & C_k \end{pmatrix}.$$

Die einzelnen Blöcke C_i sind die $(m_i \times m_i)$ -Matrizen der Einschränkungen von f auf die Haupträume H_i .

(b) Für die Blöcke gilt $(\lambda_i - C_i)^{r_i} = 0$, und für $j \neq i$ ist $(\lambda_i - C_j)$ regulär.

(c) Die Dimensionen m_i der Haupträume sind die Vielfachheiten der Eigenwerte λ_i im charakteristischen Polynom von A .

BEWEIS.

(a) Weil die Haupträume H_i invariante Unterräume für A sind, ergibt sich die Blockdiagonalform nach Satz 4.1.4.

(b) C_i ist die Matrix des Endomorphismus

$$f_i : \begin{cases} H_i & \rightarrow H_i \\ x & \mapsto Ax \end{cases}.$$

Für $x \in H_i$ gilt $(\lambda_i - A)^{r_i}x = 0$, daher ist $(\lambda_i - C_i)^{r_i} = 0$. Andererseits gilt für $i \neq j$ und $x \in H_j \setminus \{0\}$ immer $(\lambda_i - A)x \neq 0$, denn $\ker(\lambda_i - A) \cap \ker(\lambda_j - A)^{r_j} = \{0\}$ nach Lemma 3.1.2. Daher ist $\ker(\lambda_i - C_j) = \{0\}$, und $(\lambda_i - C_j)$ ist regulär.

- (c) Weil ähnliche Matrizen dasselbe charakteristische Polynom haben, untersuchen wir die Vielfachheiten von λ_i in $\det(\lambda - M_B^B(f))$. Die Matrix C_i besitzt nur den Eigenwert λ_i , daher ist

$$\det(\lambda - C_i) = (\lambda - \lambda_i)^{m_i}.$$

Für $j \neq i$ ist $\lambda_i - C_j$ regulär, also ist $\det(\lambda_i - C_j) \neq 0$, und daher ist $(\lambda - \lambda_i)$ kein Teiler von $\det(\lambda - C_j)$. Wegen der Blockdiagonalform ist

$$\det(\lambda - M_B^B(f)) = \det(\lambda - C_1) \cdots \det(\lambda - C_k).$$

In diesem Produkt kommt $(\lambda - \lambda_i)$ als Nullstelle mit der Vielfachheit m_i vor. □

BEISPIEL 4.3.6. Die Matrix A hat die Eigenwerte 2 und -1, jeweils mit der algebraischen Vielfachheit 3. Bestimmen Sie die Haupträume von A und die Matrizen der Projektionen von \mathbb{R}^n auf die Haupträume.

$$A = \begin{pmatrix} 0 & 2 & 1 & -1 & 0 & 0 \\ -2 & 3 & 1 & -1 & 0 & 0 \\ -4 & 7 & 4 & -5 & 0 & 0 \\ -2 & 4 & 1 & -2 & 0 & 0 \\ -6 & 6 & 3 & -6 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}$$

Wir beginnen mit dem Eigenwert $\lambda = 2$ und untersuchen zuerst den Kern von $2 - A$:

$$\begin{aligned} & \begin{pmatrix} 2 & -2 & -1 & 1 & 0 & 0 \\ 2 & -1 & -1 & 1 & 0 & 0 \\ 4 & -7 & -2 & 5 & 0 & 0 \\ 2 & -4 & -1 & 4 & 0 & 0 \\ 6 & -6 & -3 & 6 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & -2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -3 & 0 & 3 & 0 & 0 \\ 0 & -2 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix} \\ & \rightarrow \begin{pmatrix} 2 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix} \\ & \rightarrow \begin{pmatrix} 2 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 2 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \\ & =: S_1. \end{aligned}$$

Im letzten Schritt haben wir nur noch eine Nullzeile gestrichen und eine Zeile durch 3 gekürzt. Die Pivotschritte haben auf eine Matrix S_1 mit Rang 5 geführt. Der Kern von $2 - A$, zugleich der Kern von S_1 , ist also eindimensional. Weil der Hauptraum aber die Dimension 3 haben muss, schreiten wir jetzt zur Bestimmung von $\ker(2 - A)^2$ fort. Dabei nützen wir einen Rechenvorteil: Es ist

$$x \in \ker(2 - A)^2 \iff (2 - A)x \in \ker(2 - A) \iff (2 - A)x \in \ker(S_1) \iff x \in \ker(S_1(2 - A)).$$

Auf Grund der einfachen Gestalt von S_1 ist $S_1(2-A)$ leichter zu berechnen als $(2-A)^2$, außerdem hat $S_1(2-A)$ nur mehr 5 Zeilen.

$$\begin{aligned}
 S_1(2-A) &= \begin{pmatrix} 2 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & -2 & -1* & 1 & 0 & 0 \\ 2 & -1 & -1 & 1 & 0 & 0 \\ 4 & -7 & -2 & 5 & 0 & 0 \\ 2 & -4 & -1 & 4 & 0 & 0 \\ 6 & -6 & -3 & 6 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix} \\
 &= \begin{pmatrix} 0 & 3 & 0 & -3 & 0 & 0 \\ 2 & -1 & -1 & 1 & 0 & 0 \\ 2 & -4 & -1 & 4 & 0 & 0 \\ 6 & -6 & -3 & 6 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix}.
 \end{aligned}$$

Wir berechnen den Kern:

$$\begin{aligned}
 &\begin{pmatrix} 0 & 3* & 0 & -3 & 0 & 0 \\ 2 & -1 & -1 & 1 & 0 & 0 \\ 2 & -4 & -1 & 4 & 0 & 0 \\ 6 & -6 & -3 & 6 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 1 & 0 & -1 & 0 & 0 \\ 2* & 0 & -1 & 0 & 0 & 0 \\ 2 & 0 & -1 & 0 & 0 & 0 \\ 6 & 0 & -3 & 0 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix} \\
 \rightarrow &\begin{pmatrix} 0 & 1 & 0 & -1 & 0 & 0 \\ 2 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 & *3 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 1 & 0 & -1 & 0 & 0 \\ 2 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \\
 \rightarrow &\begin{pmatrix} 0 & 1 & 0 & -1 & 0 & 0 \\ 2 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} =: S_2.
 \end{aligned}$$

Es entsteht eine Matrix S_2 mit Rang 4. Daher hat $\ker(2-A)^2 = \ker(S_2)$ die Dimension 2. Wir müssen noch weiter, um $\ker(2-A)^3$ zu berechnen. Wieder verwenden wir, dass $\ker(2-A)^3 = \ker(S_2(2-A))$.

$$\begin{aligned}
 S_2(2-A) &= \begin{pmatrix} 0 & 1 & 0 & -1 & 0 & 0 \\ 2 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & -2 & -1* & 1 & 0 & 0 \\ 2 & -1 & -1 & 1 & 0 & 0 \\ 4 & -7 & -2 & 5 & 0 & 0 \\ 2 & -4 & -1 & 4 & 0 & 0 \\ 6 & -6 & -3 & 6 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix} \\
 &= \begin{pmatrix} 0 & 3 & 0 & -3 & 0 & 0 \\ 0 & -3 & 0 & -3 & 0 & 0 \\ 6 & -6 & -3 & 6 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix}.
 \end{aligned}$$

Wir berechnen den Kern:

$$\begin{aligned}
 &\begin{pmatrix} 0 & 3* & 0 & -3 & 0 & 0 \\ 0 & -3 & 0 & -3 & 0 & 0 \\ 6 & -6 & -3 & 6 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 6 & 0 & -3 & 0 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 & 3* \end{pmatrix} \\
 \rightarrow &\begin{pmatrix} 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 6 & 0 & -3 & 0 & 3* & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \\
 \rightarrow &\begin{pmatrix} 0 & 1 & 0 & -1 & 0 & 0 \\ 2 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}
 \end{aligned}$$

Der Kern von $(2-A)^3$ ist dreidimensional. Wir haben nun den Hauptraum erreicht und lesen aus S_3 eine Basis des Hauptraumes ab:

$$\text{Basis von } \ker(2-A)^3 : \left(\begin{pmatrix} 1 \\ 0 \\ 0 \\ -2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right).$$

Wir untersuchen nun den Hauptraum zum Eigenwert -1. Wir beginnen mit $\ker(-1-A)$:

$$\begin{aligned} & \begin{pmatrix} -1 & -2 & -1 & 1* & 0 & 0 \\ 2 & -4 & -1 & 1 & 0 & 0 \\ 4 & -7 & -5 & 5 & 0 & 0 \\ 2 & -4 & -1 & 1 & 0 & 0 \\ 6 & -6 & -3 & 6 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} -1 & -2 & -1 & 1 & 0 & 0 \\ 3 & -2 & 0 & 0 & 0 & 0 \\ 9 & 3* & 0 & 0 & 0 & 0 \\ 3 & -2 & 0 & 0 & 0 & 0 \\ 12 & 6 & 3 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\ & \rightarrow \begin{pmatrix} 5 & 0 & -1 & 1 & 0 & 0 \\ 9* & 0 & 0 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 & 0 \\ 9 & 0 & 0 & 0 & 0 & 0 \\ -6 & 0 & 3 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 0 & 0 & -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\ & \rightarrow \begin{pmatrix} 0 & 0 & -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & -1 \end{pmatrix} =: S_4. \end{aligned}$$

Der Hauptraum muss die Dimension 3 haben, der Kern von S_4 ist aber erst zweidimensional. Wir schreiten fort zur Berechnung von $\ker(-1-A)^2 = \ker(S_4(-1-A))$.

$$\begin{aligned} S_4(-1-A) &= \begin{pmatrix} 0 & 0 & -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} -1 & -2 & -1 & 1* & 0 & 0 \\ 2 & -4 & -1 & 1 & 0 & 0 \\ 4 & -7 & -5 & 5 & 0 & 0 \\ 2 & -4 & -1 & 1 & 0 & 0 \\ 6 & -6 & -3 & 6 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -2 & 3 & 4 & -4 & 0 & 0 \\ -1 & -2 & -1 & 1 & 0 & 0 \\ 2 & -4 & -1 & 1 & 0 & 0 \\ 12 & -21 & -15 & 15 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Wir berechnen den Kern:

$$\begin{aligned} & \begin{pmatrix} -2 & 3 & 4 & -4 & 0 & 0 \\ -1 & -2 & -1 & 1 & 0 & 0 \\ 2 & -4 & -1 & 1 & 0 & 0 \\ 12 & -21 & -15 & 15 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} -2 & 3 & 4 & -4 & 0 & 0 \\ -1 & -2 & -1 & 1* & 0 & 0 \\ 2 & -4 & -1 & 1 & 0 & 0 \\ 4 & -7 & -5 & 5 & 0 & 0 \end{pmatrix} \\ & \rightarrow \begin{pmatrix} -6 & -5 & 0 & 0 & 0 & 0 \\ -1 & -2 & -1 & 1 & 0 & 0 \\ 3 & -2 & 0 & 0 & 0 & 0 \\ 9 & 3* & 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 9* & 0 & 0 & 0 & 0 & 0 \\ 5 & 0 & -1 & 1 & 0 & 0 \\ 9 & 0 & 0 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \\ & \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \\ & =: S_5. \end{aligned}$$

Wir sehen, dass der Kern von $(-1 - A)^2$ Dimension 3 hat, damit ist der Hauptraum bereits erreicht.

$$\text{Basis von } \ker(-1 - A)^2 : \left(\begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right).$$

Um die Projektionsmatrizen entlang der Hauptraumzerlegung zu bestimmen, setzen wir die Basisvektoren der Haupträume zu einer Matrix T zusammen:

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

Eine einfache Rechnung² ergibt die Inverse:

$$T^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 \\ 2 & -1 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Die ersten drei Spalten von T gehören zum Hauptraum von $\lambda = 2$. Nach Behauptung 4.2.6 ist die Projektion P_2 aus den ersten drei Spalten von T und den ersten drei Zeilen von T^{-1} zu berechnen:

$$\begin{aligned} P_2 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ -2 & 1 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Die Projektionsmatrix P_{-1} auf den Hauptraum zu $\lambda = -1$ könnten wir ebenso aus den letzten drei Spalten von T und den letzten drei Zeilen von T^{-1} berechnen. Weil aber die Summe aller Projektionsmatrizen die Einheitsmatrix sein muss, erhalten wir sie einfacher:

$$P_{-1} = E_6 - P_2 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 \\ 2 & -1 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

4.4. Jordansche Normalform.

Auf den einzelnen Haupträumen kann man die Struktur von A noch weiter untersuchen und Basen suchen, bezüglich derer sich A besonders einfach darstellen lässt. Auf diese Weise erhält man die Jordansche Normalform.

²besonders einfach mit MATLAB

DEFINITION 4.4.1. Die $r \times r$ -Matrix

$$\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}$$

heißt Jordanblock der Länge r zum Eigenwert λ .

Ein Jordanblock der Länge 1 ist einfach eine (1×1) -Matrix (λ) .

Eine Blockdiagonalmatrix, deren Blöcke Jordanblöcke sind, heißt Matrix in Jordanscher Normalform.

SATZ 4.4.2 (Jordansche Normalform). Sei \mathbb{K} ein algebraisch abgeschlossener Körper und $A \in \mathbb{K}^{n \times n}$. Dann ist A ähnlich zu einer Matrix in Jordanscher Normalform:

$$T^{-1}AT = \begin{pmatrix} J_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & J_q \end{pmatrix}.$$

Dabei sind die J_i Jordanblöcke zu (nicht unbedingt paarweise verschiedenen) Eigenwerten λ_i von A .

Der Beweis dieses Satzes und die rechnerische Bestimmung einer Basis, die zur Jordanschen Normalform führt, sind ziemlich aufwändig. Aus Zeitgründen lassen wir uns in dieser Vorlesung nicht darauf ein. Auch ist für praktische Zwecke die Bestimmung einer solchen Basis normalerweise nicht notwendig. Allerdings kann man, ohne die Basis zu kennen, die Jordansche Normalform einer Matrix auffinden.

SATZ 4.4.3. Sei $A \in \mathbb{K}^{n \times n}$ ähnlich zu einer Matrix in Jordanscher Normalform

$$B = T^{-1}AT = \begin{pmatrix} J_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & J_q \end{pmatrix}.$$

Dabei seien J_1, \dots, J_q Jordanblöcke der Längen t_1, \dots, t_q zu den (nicht unbedingt paarweise verschiedenen) Eigenwerten $\lambda_1, \dots, \lambda_q$. Dann gilt:

- (1) Die algebraische Vielfachheit eines Eigenwertes λ von A ist die Summe der Längen der Jordanblöcke zum Eigenwert λ .
- (2) Die geometrische Vielfachheit eines Eigenwertes λ von A ist die Anzahl der Jordanblöcke zum Eigenwert λ .
- (3) Die kleinste Potenz r , für die $\ker(\lambda - A)^r$ der Hauptraum zum Eigenwert λ ist, ist zugleich die größte Länge der Jordanblöcke zum Eigenwert λ .
- (4) Die Anzahl der Jordanblöcke mit Länge größer als i zu einem Eigenwert λ ist

$$\rho((\lambda - A)^i) - \rho((\lambda - A)^{i+1}).$$

- (5) Die Jordanblöcke J_i sind, bis auf ihre Reihenfolge auf der Diagonalen, eindeutig bestimmt.

BEWEIS. Es ist

$$\det(\lambda - B) = \det(T^{-1}(\lambda - A)T) = \det(T^{-1}) \det(\lambda - A) \det(T) = \det(\lambda - A),$$

daher haben A und B dasselbe charakteristische Polynom und dieselben Eigenwerte, mit denselben algebraischen Vielfachheiten. Es ist

$$(\lambda - B)^i = (\lambda - T^{-1}AT)^i = [T^{-1}(\lambda - A)T]^i = T^{-1}(\lambda - A)^i T,$$

also sind $(\lambda - A)^i$ und $(\lambda - B)^i$ ähnliche Matrizen und haben daher dieselben Ränge. Wir können also o.B.d.A. annehmen, dass $A = B$ selbst in Jordanscher Normalform vorliegt.

Dann ist $\lambda - A$ eine obere Dreiecksmatrix, und ihre Determinante ist

$$\det(\lambda - A) = \prod_{m=1}^q (\lambda - \lambda_m)^{t_m},$$

damit ist die algebraische Vielfachheit eines Eigenwertes λ gegeben durch die Summe der Längen der Jordanblöcke zum Eigenwert λ .

Für jeden Jordanblock J_m ist der Rang

$$\rho((\lambda_m - J_m)^i) = \begin{cases} t_m - i & \text{wenn } i = 0 \cdots t_m - 1 \\ 0 & \text{wenn } i \geq t_m. \end{cases}$$

Damit ist

$$\rho((\lambda_m - J_m)^i) - \rho((\lambda_m - J_m)^{i+1}) = \begin{cases} 1 & \text{wenn } i = 0 \cdots t_m - 1 \\ 0 & \text{wenn } i \geq t_m. \end{cases}$$

Daher ist $\rho((\lambda_m - A)^i) - \rho((\lambda_m - A)^{i+1})$ die Anzahl aller Jordanblöcke zum Eigenwert λ_m mit Länge größer als i . Insbesondere, für $i = 0$, erhalten wir die geometrische Vielfachheit von λ_m als Anzahl aller Jordanblöcke zum Eigenwert λ_m .

Aus diesen Regeln läßt sich eindeutig die Anzahl der Jordanblöcke jeder bestimmten Länge zu den einzelnen Eigenwerten mit Hilfe der Ränge von $(\lambda_m - A)^i$ feststellen. \square

BEISPIEL 4.4.4. Wir erklären in diesem Beispiel, wie man die Jordansche Normalform einer Matrix bestimmen kann.

Vorgangsweise: Wir nehmen an, wir haben eine Matrix $A \in \mathbb{C}^{10 \times 10}$ gegeben. Bekanntlich ist \mathbb{C} algebraisch abgeschlossen, sodass der Satz von der Jordannormalform gilt.

- 1) Wir bestimmen das charakteristische Polynom und die Eigenwerte von A . Nehmen wir an, wir erhalten $\lambda_1 = 3$ mit algebraischer Vielfachheit 2, und $\lambda_2 = 5$ mit algebraischer Vielfachheit 8. (Probe: Die algebraischen Vielfachheiten müssen sich auf 10 summieren.) Wir wissen nun, dass der Hauptraum zu $\lambda = 3$ die Dimension 2 besitzt, und der Hauptraum zu $\lambda = 5$ die Dimension 8 besitzt.
- 2) Wir verfolgen nun den Eigenwert $\lambda = 3$: Bisher wissen wir, dass entweder zwei Jordanblöcke der Länge 1, oder ein Jordanblock der Länge 2 vorliegen, da sich die Längen der Jordanblöcke auf die Dimension des Hauptraumes summieren. Wir berechnen nun $\ker(3 - A)$. Nehmen wir an, wir erhalten einen eindimensionalen Kern. Wir wissen nun: Die geometrische Vielfachheit von $\lambda = 3$ beträgt 1. Es gibt einen Jordanblock, und daher muss das ein Block der Länge 2 sein. Der Hauptraum zum Eigenwert 3 ist $\ker(3 - A)^2$, weil 2 die Länge des größten (einzigen) Jordanblocks zu $\lambda = 3$ ist.
- 3) Wir verfolgen nun den Eigenwert $\lambda = 5$. Bisher wissen wir, dass sich die Längen der Jordanblöcke auf 8 summieren. Wir berechnen $\ker(5 - A)$. Nehmen wir an, wir erhalten einen vierdimensionalen Kern. Dann ist die geometrische Vielfachheit von $\lambda = 5$ also 4, und es gibt 4 Jordanblöcke. Es gibt aber immer noch viele Möglichkeiten, die Gesamtlänge 8 auf vier Blöcke aufzuteilen.

Für $i \neq 0$ gilt nach dem Binomischen Lehrsatz, und weil $(A - \lambda_i)^{r_i} P_i = 0$,

$$\begin{aligned} A^m P_i &= [(A - \lambda_i) + \lambda_i E_n]^m P_i \\ &= \sum_{j=0}^m \binom{m}{j} (A - \lambda_i)^j \lambda_i^{m-j} P_i = \lambda_i^m \sum_{j=0}^{r_i-1} \binom{m}{j} \frac{1}{\lambda_i^j} (A - \lambda_i)^j P_i \\ &= \lambda_i^m \sum_{j=0}^{r_i-1} m(m-1) \cdots (m-j+1) \frac{1}{j! \lambda_i^j} (A - \lambda_i)^j P_i \\ &= \lambda_i^m \sum_{j=0}^{r_i-1} \sum_{l=0}^j m^l Q_{i,j,l} \end{aligned}$$

mit konstanten Matrizen $Q_{i,j,l}$. Mit $Q_{i,l} = \sum_{j=0}^{r_i-1} Q_{i,j,l}$ ergibt sich

$$A^m P_i = \lambda_i^m \sum_{l=0}^{r_i-1} m^l Q_{i,l}.$$

□

5.2. Minimalpolynom.

DEFINITION 5.2.1. Sei \mathbb{K} ein Körper, $p(x) = \gamma_0 + \gamma_1 x + \gamma_2 x^2 + \cdots + \gamma_r x^r$ ein Polynom in $\mathbb{K}[x]$. Sei $A \in \mathbb{K}^{n \times n}$ eine quadratische Matrix. Wir definieren die Matrix

$$p(A) := \gamma_0 E_n + \gamma_1 A + \gamma_2 A^2 + \cdots + \gamma_r A^r.$$

Weil der Raum $\mathbb{K}^{n \times n}$ die Dimension n^2 hat, muss das System von Matrizen (E_n, A, \dots, A^{n^2}) linear abhängig sein, daher muss es ein Polynom vom Grad höchstens n^2 geben, sodass $p(A) = 0$. Der folgende Satz zeigt, dass in Wirklichkeit ein viel kleinerer Grad ausreicht.

LEMMA 5.2.2. Sei \mathbb{K} ein Körper, $p \in \mathbb{K}[x]$ ein Polynom, $A \in \mathbb{K}^{n \times n}$ eine quadratische Matrix, und $T \in \mathbb{K}^{n \times n}$ eine reguläre Matrix. Dann gilt: $p(T^{-1}AT) = T^{-1}p(A)T$.

BEWEIS. Zunächst berechnen wir

$$\begin{aligned} [T^{-1}AT]^n &= [T^{-1}AT][T^{-1}AT] \cdots [T^{-1}AT] \\ &= T^{-1}A[TT^{-1}]A[TT^{-1}] \cdots [TT^{-1}]AT = T^{-1}A^n T. \end{aligned}$$

Damit ist aber

$$p(T^{-1}AT) = \sum_{i=0}^r \gamma_i [T^{-1}AT]^i = \sum_{i=0}^r \gamma_i T^{-1} A^i T = T^{-1} \left[\sum_{i=0}^r \gamma_i A^i \right] T.$$

□

SATZ 5.2.3. Sei \mathbb{K} ein algebraisch abgeschlossener Körper, sei $A \in \mathbb{K}^{n \times n}$. Es seien $\lambda_1, \dots, \lambda_k$ die Eigenwerte von A , und es seien r_1, \dots, r_k die kleinsten Potenzen, sodass $\ker(\lambda_i - A)^{r_i} = \ker(\lambda_i - A)^{r_i+1}$. (Es ist also $\ker(\lambda_i - A)^{r_i}$ der Hauptraum zum Eigenwert λ_i .) Wir setzen

$$p(x) = (x - \lambda_1)^{r_1} (x - \lambda_2)^{r_2} \cdots (x - \lambda_k)^{r_k}.$$

Dann gilt:

- 1) $p(A) = 0$.
- 2) Ist $q \in \mathbb{K}[x]$ ein Polynom, so dass $q(A) = 0$, so ist p ein Teiler von q .

BEWEIS. Nach Satz 4.3.4 gibt es eine reguläre Matrix T , sodass $C = T^{-1}AT$ Blockdiagonalform hat:

$$C = \begin{pmatrix} C_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & C_k \end{pmatrix}.$$

Dabei sind die $C_i \in \mathbb{K}^{m_i \times m_i}$, und die m_i sind die algebraischen Vielfachheiten von λ_i . Die C_i repräsentieren die Einschränkung des Homomorphismus $x \mapsto Ax$ auf die Haupträume $\ker(\lambda_i - A)^{r_i}$. Wegen Lemma 5.2.2 ist $p(A) = 0$ genau dann, wenn $p(C) = 0$. Direktes Nachrechnen ergibt

$$p(C) = \begin{pmatrix} p(C_1) & 0 & \cdots & 0 \\ 0 & p(C_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p(C_k) \end{pmatrix}.$$

Auf Grund der Definition des Hauptraumes und der r_i ist r_i die kleinste Potenz sodass $(\lambda_i - C_i)^{r_i} = 0$. Für $\mu \neq \lambda_i$ gilt $\ker(\mu - A) \cap \ker(\lambda_i - A)^{r_i} = \{0\}$, anders gesagt, der Hauptraum zu λ_i enthält von $\ker(\mu - A)$ nur den Nullvektor. Also gehört die Matrix $(\mu - C_i)$ zu einem injektiven Endomorphismus, deshalb ist sie regulär.

Sei nun $q \in \mathbb{K}[x] \setminus \{0\}$. Weil \mathbb{K} algebraisch abgeschlossen ist, zerfällt q in Linearfaktoren:

$$q(x) = \gamma \prod_{u=1}^k (\lambda_u - x)^{s_u} \prod_{j=1}^l (\mu_j - x)^{t_j}.$$

Dabei sind $s_u \in \mathbb{N} \cup \{0\}$, $t_j \in \mathbb{N}$, $\mu_j \notin \{\lambda_1, \dots, \lambda_k\}$, $\gamma \in \mathbb{K} \setminus \{0\}$. Es ist $q(A) = 0$ genau dann, wenn $q(C) = 0$, und das gilt genau dann, wenn für alle i gilt: $q(C_i) = 0$. Nun ist

$$q(C_i) = \gamma \prod_{u \neq i} (\lambda_u - C_i)^{s_u} \prod_{j=1}^l (\mu_j - C_i)^{t_j} (\lambda_i - C_i)^{s_i}.$$

Da alle Matrizen $(\lambda_u - C_i)$ für $u \neq i$ und alle Matrizen $(\mu_j - C_i)$ regulär sind, ist

$$q(C_i) = 0 \Leftrightarrow (\lambda_i - C_i)^{s_i} = 0 \Leftrightarrow s_i \geq r_i,$$

und das gilt genau dann, wenn p ein Teiler von q ist. \square

DEFINITION 5.2.4. Sei \mathbb{K} ein Körper, $p \in \mathbb{K}[x]$ ein Polynom, $A \in \mathbb{K}^{n \times n}$ eine quadratische Matrix. Sei $p(x) = x^r + \gamma_{r-1}x^{r-1} + \dots + \gamma_0$ in $\mathbb{K}[x]$ so, dass für jedes $q \in \mathbb{K}[x]$ gilt:

$$q(A) = 0 \Leftrightarrow p \text{ teilt } q.$$

Dann heißt p das Minimalpolynom von A .

Diese Definition funktioniert auch über nicht algebraisch abgeschlossenen Körpern, nur die Beschreibung des Minimalpolynoms mittels Satz 5.2.3 funktioniert nicht mehr immer, weil sich nicht für jede Matrix der Raum \mathbb{K}^n in Haupträume zerlegen lässt.

5.3. Satz von Cayley-Hamilton.

SATZ 5.3.1 (Satz von Cayley-Hamilton). Sei \mathbb{K} ein algebraisch abgeschlossener Körper, $A \in \mathbb{K}^{n \times n}$ eine quadratische Matrix, und $q \in \mathbb{K}[x]$ ihr charakteristisches Polynom. Dann ist $q(A) = 0$.

BEWEIS. Wir wissen bereits aus Satz 5.2.3, dass das Minimalpolynom von A sich in der folgenden Form schreiben lässt:

$$p(x) = \prod_{i=1}^k (x - \lambda_i)^{r_i}.$$

Dabei sind $\lambda_1, \dots, \lambda_k$ die Eigenwerte von A , und r_i jeweils die kleinste Potenz sodass $\ker(\lambda_i - A)^{r_i} = \ker(\lambda_i - A)^{r_i+1}$. Dagegen ist das charakteristische Polynom von A

$$q(x) = \prod_{i=1}^k (x - \lambda_i)^{m_i},$$

wobei m_i die algebraische Vielfachheit von λ_i ist. Zugleich wissen wir aus Satz 4.3.4, dass m_i die Dimension des Hauptraumes $\ker(\lambda_i - A)^{r_i}$ ist. Nun ist

$$1 \leq \dim(\ker(\lambda_i - A)) < \dim(\ker(\lambda_i - A)^2) < \dots < \dim(\ker(\lambda_i - A)^{r_i}) = m_i.$$

Daher ist $r_i \leq m_i$, und das Minimalpolynom teilt das charakteristische Polynom. Folglich ist $q(A) = 0$. \square

BEISPIEL 5.3.2. Bestimmen Sie das charakteristische Polynom und das Minimalpolynom der folgenden Matrix A über \mathbb{C} . Überprüfen Sie, dass A , in beide Polynome eingesetzt, die Nullmatrix ergibt.

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

Lösung: Das charakteristische Polynom erhalten wir durch Entwicklung nach der zweiten Zeile:

$$\begin{vmatrix} \lambda - 3 & -1 & -3 \\ 0 & \lambda - 2 & 0 \\ 1 & 1 & \lambda + 1 \end{vmatrix} = (\lambda - 2) \begin{vmatrix} \lambda - 3 & -3 \\ 1 & \lambda + 1 \end{vmatrix} \\ = (\lambda - 2)[(\lambda - 3)(\lambda + 1) + 3] = (\lambda - 2)(\lambda^2 - 2\lambda) = (\lambda - 2)^2 \lambda = \lambda^3 - 4\lambda^2 + 4\lambda.$$

Es gibt zwei Eigenwerte $\lambda = 2$ und $\lambda = 0$. Wir bestimmen die geometrische Vielfachheit von $\lambda = 2$, also die Dimension von $\ker(2 - A)$:

$$\begin{pmatrix} -1 & -1 & -3 \\ 0 & 0 & 0 \\ 1 & 1 & 3 \end{pmatrix} \rightarrow \begin{pmatrix} -1 & -1 & -3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Daher ist der Eigenraum zu $\lambda = 2$ zweidimensional. Weil aber die algebraische Vielfachheit von $\lambda = 2$ auch nur 2 ist, ist der Eigenraum bereits der Hauptraum. Damit ist das Minimalpolynom

$$(\lambda - 2)^1 \lambda = \lambda^2 - 2\lambda.$$

Um die Probe zu machen, berechnen wir erst die Potenzen von A :

$$A^2 = \begin{pmatrix} 6 & 2 & 6 \\ 0 & 4 & 0 \\ -2 & -2 & -2 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 12 & 4 & 12 \\ 0 & 8 & 0 \\ -4 & -4 & -4 \end{pmatrix}.$$

Wir setzen in das charakteristische Polynom ein:

$$A^3 - 4A^2 + 4A = \begin{pmatrix} 12 & 4 & 12 \\ 0 & 8 & 0 \\ -4 & -4 & -4 \end{pmatrix} - \begin{pmatrix} 24 & 8 & 24 \\ 0 & 16 & 0 \\ -8 & -8 & -8 \end{pmatrix} + \begin{pmatrix} 12 & 4 & 12 \\ 0 & 8 & 0 \\ -4 & -4 & -4 \end{pmatrix} = 0.$$

Wir setzen in das Minimalpolynom ein:

$$A^2 - 2A = \begin{pmatrix} 6 & 2 & 6 \\ 0 & 4 & 0 \\ -2 & -2 & -2 \end{pmatrix} - \begin{pmatrix} 6 & 2 & 6 \\ 0 & 4 & 0 \\ -2 & -2 & -2 \end{pmatrix} = 0.$$

\square

Normale Matrizen

Wir werden in diesem Kapitel stets den Raum \mathbb{R}^n bzw. \mathbb{C}^n mit dem euklidischen inneren Produkt und der dazugehörigen Norm betrachten.

1. Orthogonale und unitäre Matrizen

1.1. Unitäre Matrizen.

DEFINITION 1.1.1.

- 1) Eine Matrix $T \in \mathbb{R}^{n \times n}$ heißt orthogonal, wenn sie regulär ist und ihre Transponierte zugleich die Inverse ist: $T^{-1} = T^T$.
- 2) Eine Matrix $T \in \mathbb{C}^{n \times n}$ heißt unitär, wenn sie regulär ist und ihre Adjungierte zugleich die Inverse ist: $T^{-1} = T^*$.

Eine orthogonale Matrix ist also nichts anderes als eine reelle unitäre Matrix. Wenn wir in den folgenden Sätzen immer wieder orthogonale und unitäre Matrizen betrachten, werden wir normalerweise nur den Beweis für unitäre Matrizen über \mathbb{C} führen. Der reelle, orthogonale Fall geht dann genauso.

SATZ 1.1.2.

- 1) Sei $T \in \mathbb{R}^{n \times n}$. Dann sind äquivalent:
 - (i) T ist orthogonal.
 - (ii) Für alle $x \in \mathbb{R}^n$ gilt $\|Tx\| = \|x\|$.
 - (iii) Für alle $x, y \in \mathbb{R}^n$ gilt $\langle Tx, Ty \rangle = \langle x, y \rangle$.
- 2) Sei $T \in \mathbb{C}^{n \times n}$. Dann sind äquivalent:
 - (i) T ist unitär.
 - (ii) Für alle $x \in \mathbb{C}^n$ gilt $\|Tx\| = \|x\|$.
 - (iii) Für alle $x, y \in \mathbb{C}^n$ gilt $\operatorname{Re}(\langle Tx, Ty \rangle) = \operatorname{Re}(\langle x, y \rangle)$.
 - (iv) Für alle $x, y \in \mathbb{C}^n$ gilt $\langle Tx, Ty \rangle = \langle x, y \rangle$.

BEWEIS. Wir beweisen Teil (2), dann geht Teil (1) ebenso.

(i) \Rightarrow (ii): Sei $T^* = T^{-1}$. Es ist

$$\|Tx\|^2 = \langle Tx, Tx \rangle = (Tx)^*(Tx) = x^*T^*Tx = x^*x = \|x\|^2.$$

(ii) \Rightarrow (iii): Es ist

$$\begin{aligned} \|x+y\|^2 &= \|x\|^2 + \|y\|^2 + \langle x, y \rangle + \langle y, x \rangle = \|x\|^2 + \|y\|^2 + 2\operatorname{Re}(\langle x, y \rangle), \\ \|x-y\|^2 &= \|x\|^2 + \|y\|^2 - \langle x, y \rangle - \langle y, x \rangle = \|x\|^2 + \|y\|^2 - 2\operatorname{Re}(\langle x, y \rangle). \end{aligned}$$

Wir subtrahieren diese Gleichungen und erhalten

$$4\operatorname{Re}(\langle x, y \rangle) = \|x+y\|^2 - \|x-y\|^2.$$

Man kann in (ii) den Vektor x einmal durch $x+y$, und dann durch $x-y$ ersetzen, und erhält

$$4\operatorname{Re}(\langle Tx, Ty \rangle) = \|T(x+y)\|^2 - \|T(x-y)\|^2 = \|x+y\|^2 - \|x-y\|^2 = 4\operatorname{Re}(\langle x, y \rangle).$$

(iii) \Rightarrow (iv): Es gilt:

$$\operatorname{Re}(\langle x, iy \rangle) = \operatorname{Re}(i\langle x, y \rangle) = -\operatorname{Im}(\langle x, y \rangle).$$

Aus (iii) folgt daher

$$\operatorname{Im}(\langle Tx, Ty \rangle) = -\operatorname{Re}(\langle Tx, T(iy) \rangle) = -\operatorname{Re}(\langle x, iy \rangle) = \operatorname{Im}(\langle x, y \rangle).$$

Nochmals wegen (iii) folgt also $\langle Tx, Ty \rangle = \langle x, y \rangle$.

(iv) \Rightarrow (i): Es gelte (iv). Seien $x, y \in \mathbb{C}^n$. Dann gilt

$$\langle y, T^*Tx - x \rangle = \langle y, T^*Tx \rangle - \langle y, x \rangle = \langle Ty, Tx \rangle - \langle y, x \rangle = 0.$$

Setzen wir insbesondere $y = T^*Tx - x$, so folgt $\|T^*Tx - x\|^2 = 0$, also $T^*Tx = x$. Damit ist T regulär, und T^* ist die Inverse von T . \square

DEFINITION 1.1.3. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, sei $f : \mathbb{K}^n \rightarrow \mathbb{K}^n$ eine (nicht notwendig lineare) Abbildung mit der Eigenschaft:

$$(\forall x, y \in \mathbb{K}^n) \|f(x) - f(y)\| = \|x - y\|.$$

Dann heißt f eine isometrische Abbildung oder Isometrie.

Eine lineare Abbildung $\mathbb{K}^n \rightarrow \mathbb{K}^n$ ist also genau dann eine Isometrie, wenn sie durch eine unitäre (bzw. im reellen Fall orthogonale) Matrix beschrieben wird.

BEISPIEL 1.1.4. Die folgende Matrix beschreibt eine Drehung im \mathbb{R}^2 um den Nullpunkt mit einem Drehwinkel von ϕ :

$$D_\phi = \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix}.$$

Zeigen Sie, dass diese Matrix orthogonal ist.

Lösung: Aus der geometrischen Anschauung wissen wir, dass eine Drehung eine Isometrie ist, also erwarten wir dass D_ϕ orthogonal ist. Wir können das aber auch nachrechnen:

$$\begin{aligned} D_\phi^T D_\phi &= \begin{pmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix} \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix} \\ &= \begin{pmatrix} \cos^2(\phi) + \sin^2(\phi) & 0 \\ 0 & \sin^2(\phi) + \cos^2(\phi) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

\square

BEISPIEL 1.1.5. Bewegungen des \mathbb{R}^2 oder \mathbb{R}^3 , welche den Nullpunkt festlassen, werden durch orthogonale Matrizen beschrieben. Also z.B. Drehungen um Achsen durch den Nullpunkt, Spiegelungen an Geraden oder Ebenen durch den Nullpunkt, und Hintereinanderausführung solcher Bewegungen. Dagegen ist die Parallelverschiebung entlang eines Vektors a , also $f(x) = x + a$ zwar eine Isometrie, aber keine lineare Abbildung.

SATZ 1.1.6.

- 1) Das Produkt orthogonaler (bzw. unitärer) Matrizen ist orthogonal (bzw. unitär).
- 2) Eine reguläre Matrix A ist genau dann orthogonal (unitär), wenn A^{-1} orthogonal (unitär) ist.

Trauen Sie sich und beweisen Sie es selbst! Entweder direkt mit der Definition von Orthogonalität, oder mit Hilfe von Satz 1.1.2.

1.2. Orthogonale Diagonalisierung.

SATZ 1.2.1. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $T \in \mathbb{K}^{n \times n}$. Seien (s_1, \dots, s_n) die Spalten von T und z_1, \dots, z_n die Zeilen von T . Dann sind folgende Aussagen äquivalent:

- 1) T ist orthogonal bzw. unitär.
- 2) (s_1, \dots, s_n) ist ein Orthonormalsystem.
- 3) (z_1^*, \dots, z_n^*) ist ein Orthonormalsystem.

BEWEIS. (1) \Rightarrow (2): Seien s_1, \dots, s_n die Spalten von T , und e_1, \dots, e_n die Einheitsvektoren in \mathbb{K}^n . Bekanntlich ist $s_i = T(e_i)$. Wegen Satz 1.1.2 gilt für orthogonale (unitäre) Matrizen

$$\langle s_i, s_j \rangle = \langle Te_i, Te_j \rangle = \langle e_i, e_j \rangle = \delta_{i,j}.$$

(Dabei ist $\delta_{i,j}$ das Kronecker-Delta.)

(2) \Rightarrow (1): Sei nun (s_1, \dots, s_n) ein Orthonormalsystem. Die Zeilen von T^* sind die Adjungierten der Spalten von T . Damit ist

$$T^*T = \begin{pmatrix} - & s_1^* & - \\ & \vdots & \\ - & s_n^* & - \end{pmatrix} \begin{pmatrix} | & & | \\ s_1 & \cdots & s_n \\ | & & | \end{pmatrix} = \begin{pmatrix} s_1^*s_1 & \cdots & s_1^*s_n \\ \vdots & & \vdots \\ s_n^*s_1 & \cdots & s_n^*s_n \end{pmatrix} = E_n.$$

Also ist T^* die Inverse zu T .

(1) \Leftrightarrow (3): Es ist T unitär genau dann, wenn T^* unitär ist. Dies gilt aber wegen Punkt (2) genau dann, wenn die Spalten von T^* ein Orthonormalsystem sind. Die Spalten von T^* sind aber die Adjungierten der Zeilen von T . \square

KOROLLAR 1.2.2. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, sei $B = (b_1, \dots, b_n)$ eine Basis von \mathbb{K}^n und $E = (e_1, \dots, e_n)$ die Einheitsbasis. Dann gilt: Die Übergangsmatrix T_E^B (und damit auch ihre Inverse T_E^B) ist genau dann orthogonal bzw. unitär, wenn B eine Orthonormalbasis ist.

BEWEIS. Die Spalten von T_E^B sind die Basisvektoren b_1, \dots, b_n , in Einheitskoordinaten geschrieben. Nach Satz 1.2.1 sind die Spalten von T_E^B genau dann ein Orthonormalsystem, wenn T_E^B orthogonal bzw. unitär ist. \square

DEFINITION 1.2.3. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $A \in \mathbb{K}^{n \times n}$. Die Matrix A heißt orthogonal (unitär) diagonalisierbar, wenn \mathbb{K}^n eine Orthonormalbasis aus Eigenvektoren der Matrix A besitzt.

SATZ 1.2.4. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $A \in \mathbb{K}^{n \times n}$. Die Matrix A ist genau dann orthogonal (unitär) diagonalisierbar, wenn es eine orthogonale (unitäre) Matrix T gibt, sodass $T^{-1}AT$ eine Diagonalmatrix ist.

BEWEIS. Ist $T \in \mathbb{K}^{n \times n}$ regulär, so ist $T^{-1}AT$ genau dann eine Diagonalmatrix, wenn die Spalten von T Eigenvektoren von A sind. Die Spalten bilden genau dann zusätzlich ein Orthonormalsystem, wenn T orthogonal (unitär) ist. \square

SATZ 1.2.5. Eine Matrix $A \in \mathbb{C}^{n \times n}$ ist genau dann unitär diagonalisierbar, wenn die folgenden beiden Aussagen gelten:

- 1) Sind λ, μ zwei verschiedene Eigenwerte von A und x, y dazugehörige Eigenvektoren, so ist $\langle x, y \rangle = 0$. (Anders ausgedrückt, die Eigenräume stehen aufeinander orthogonal.)
- 2) Ist λ ein Eigenwert von A , so ist $\ker(\lambda - A)^2 = \ker(\lambda - A)$. (Anders ausgedrückt, die Vielfachheit des Eigenwertes im Minimalpolynom ist 1, oder wieder anders, der Hauptraum zu λ ist der Eigenraum.)

BEWEIS. \Rightarrow : Wenn A diagonalisierbar ist, ist \mathbb{C}^n die direkte Summe der Eigenräume. Daher müssen in diesem Fall die Haupträume gleich den Eigenräumen sein. Damit gilt (2). Sei nun A unitär diagonalisierbar mit paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_k$, seien $b_{i,1}, \dots, b_{i,m_i}$ Eigenvektoren zu λ_i sodass sich mit $(b_{1,1}, \dots, b_{k,m_k})$ eine Orthonormalbasis von \mathbb{C}^n ergibt. Sei x ein beliebiger Eigenvektor zu λ_i und y ein Eigenvektor zu λ_j mit $i \neq j$. Man sieht leicht, dass $x \in \mathcal{L}(\{b_{i,1}, \dots, b_{i,m_i}\})$ und $y \in \mathcal{L}(\{b_{j,1}, \dots, b_{j,m_j}\})$. Weil alle $b_{i,s}$ auf alle $b_{j,t}$ orthogonal stehen, folgt $\langle x, y \rangle = 0$. Also gilt (1).

\Leftarrow : Wir nehmen nun an, dass (1) und (2) gelten. Wegen der algebraischen Abgeschlossenheit von \mathbb{C} und Satz 4.3.4 ist \mathbb{C}^n die direkte Summe der Haupträume von A . Wegen (2) sind aber die Haupträume zugleich die Eigenräume. Seien wieder $\lambda_1, \dots, \lambda_k$ die Eigenwerte von A . Für jedes i wählen wir eine Orthonormalbasis $(b_{i,1}, \dots, b_{i,m_i})$ des Eigenraumes $\ker(\lambda_i - A)$. Wegen (1) stehen für $i \neq j$ die Vektoren $b_{i,s}$ und $b_{j,t}$ orthogonal. Daraus ergibt sich, dass $(b_{1,1}, \dots, b_{k,m_k})$ eine Orthonormalbasis des \mathbb{C}^n ist. \square

BEMERKUNG 1.2.6. Beachten Sie aber, dass eine reelle $n \times n$ -Matrix, für deren reelle Eigenvektoren und Eigenwerte die Bedingungen von Satz 1.2.5 gelten, nicht diagonalisierbar sein muss. Sie kann ja komplexe Eigenwerte haben.

2. Normale Matrizen

2.1. Definition und Beispiele.

DEFINITION 2.1.1. Sei $A \in \mathbb{C}^{n \times n}$ eine quadratische Matrix.

- 1) A heißt normal, wenn gilt $AA^* = A^*A$.
- 2) A heißt selbstadjungiert oder hermitesch, wenn gilt: $A = A^*$. Eine selbstadjungierte Matrix $A \in \mathbb{R}^{n \times n}$ heißt auch symmetrisch.
- 3) A heißt schiefadjungiert, wenn gilt: $A = -A^*$.

BEMERKUNG 2.1.2. Unitäre, selbstadjungierte und schiefadjungierte Matrizen sind Spezialfälle von normalen Matrizen.

2.2. Orthogonale Diagonalisierbarkeit normaler Matrizen.

SATZ 2.2.1. Sei $A \in \mathbb{C}^{n \times n}$ eine normale Matrix. Dann gilt:

- 1) $\ker(\lambda - A) = \ker(\bar{\lambda} - A^*)$
- 2) $\ker(\lambda - A) \perp \ker(\mu - A)$ falls $\lambda \neq \mu$.
- 3) $\ker(\lambda - A)^2 = \ker(\lambda - A)$

BEWEIS. Teil (1): Sei $(\lambda - A)x = 0$. Dann ist

$$\begin{aligned} \|(\bar{\lambda} - A^*)x\|^2 &= \langle (\bar{\lambda} - A^*)x, (\bar{\lambda} - A^*)x \rangle = [(\bar{\lambda} - A^*)x]^* [(\bar{\lambda} - A^*)x] \\ &= x^*(\lambda - A^*)(\bar{\lambda} - A^*)x = x^*(\lambda - A^*)(\lambda - A)x \\ &= \langle (\lambda - A)x, (\lambda - A)x \rangle = \|(\lambda - A)x\|^2 = 0. \end{aligned}$$

Also ist $\ker(\lambda - A) \subseteq \ker(\bar{\lambda} - A^*)$. Weil auch A^* normal ist, folgt nun auch

$$\ker(\bar{\lambda} - A^*) \subseteq \ker(\lambda - A).$$

Teil (2): Sei $(\lambda - A)x = 0$, $(\mu - A)y = 0$ für $\lambda \neq \mu$. Wegen Teil (1) gilt auch $(\bar{\lambda} - A^*)x = 0$. Daher ist

$$\mu \langle x, y \rangle = \langle x, \mu y \rangle = \langle x, Ay \rangle = \langle A^*x, y \rangle = \langle \bar{\lambda}x, y \rangle = \lambda \langle x, y \rangle.$$

Da $\lambda \neq \mu$, folgt $\langle x, y \rangle = 0$.

Teil (3): Sei $(\lambda - A)^2x = 0$. Dann ist also $(\lambda - A)x \in \ker(\lambda - A)$. Wegen Teil (1) ist auch $(\lambda - A)x \in \ker(\bar{\lambda} - A^*)$. Es ist

$$\|(\lambda - A)x\|^2 = \langle (\lambda - A)x, (\lambda - A)x \rangle = \langle x, (\bar{\lambda} - A^*)(\lambda - A)x \rangle = \langle x, 0 \rangle = 0.$$

□

SATZ 2.2.2. Sei $A \in \mathbb{C}^{n \times n}$. Dann gilt: Die Matrix A ist genau dann normal, wenn A unitär diagonalisierbar ist.

BEWEIS. Ist A normal, so folgt aus Satz 2.2.1 und Satz 1.2.5, dass A unitär diagonalisierbar ist.

Sei nun $D = T^{-1}AT$ mit einer unitären Matrix T und einer Diagonalmatrix D . Man überprüft leicht, dass Diagonalmatrizen miteinander kommutieren, sodass $D^*D = DD^*$. Es ist dann $A = TDT^{-1}$ und $T^{-1} = T^*$. Daher ist $A^* = (T^{-1})^*D^*T^* = TD^*T^{-1}$ und daher

$$AA^* = TDT^{-1}TD^*T^{-1} = TDD^*T^{-1} = TD^*DT^{-1} = TD^*T^{-1}TDT^{-1} = A^*A.$$

□

3. Selbstadjungierte Matrizen und Definitheit

3.1. Diagonalisierung selbstadjungierter Matrizen.

SATZ 3.1.1. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, $A \in \mathbb{K}^{n \times n}$ selbstadjungiert.

- 1) Alle Eigenwerte von A sind reell.
- 2) A ist über \mathbb{K} unitär (orthogonal) diagonalisierbar.

BEWEIS. Sei $\lambda x = Ax$ für einen Vektor $x \neq 0$. Dann ist wegen Satz 2.2.1 auch $\bar{\lambda}x = A^*x = Ax = \lambda x$. Da $x \neq 0$ folgt $\lambda = \bar{\lambda}$.

Die Matrix A ist jedenfalls diagonalisierbar über \mathbb{C} nach Satz 2.2.2. Ist $A \in \mathbb{R}^{n \times n}$, so können die Eigenvektoren reell gewählt werden, weil auch die Eigenwerte reell sind. Es folgt, dass A eine Orthonormalbasis aus reellen Eigenvektoren hat. □

3.2. Definitheit.

DEFINITION 3.2.1. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $A \in \mathbb{K}^{n \times n}$ selbstadjungiert. Die Matrix A heißt

- 1) positiv (negativ) definit, wenn für alle $x \in \mathbb{K}^n \setminus \{0\}$ gilt $x^*Ax > 0$ (bzw., $x^*Ax < 0$).
- 1) positiv (negativ) semidefinit, wenn für alle $x \in \mathbb{K}^n$ gilt $x^*Ax \geq 0$ (bzw., $x^*Ax \leq 0$).

SATZ 3.2.2. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $A \in \mathbb{K}^{n \times n}$ eine selbstadjungierte Matrix. Dann gilt:

- 1) A ist genau dann positiv definit, wenn alle Eigenwerte von A größer als 0 sind.
- 2) A ist genau dann positiv semidefinit, wenn alle Eigenwerte von A größer oder gleich 0 sind.

(Ein analoges Kriterium gilt für negativ definit und negativ semidefinit.)

BEWEIS. Sei A positiv definit und λ ein Eigenwert von A . Sei $x \neq 0$ ein Eigenvektor zum Eigenwert λ . Dann ist

$$\lambda \|x\|^2 = \lambda \langle x, x \rangle = \langle x, \lambda x \rangle = \langle x, Ax \rangle > 0.$$

Seien nun umgekehrt alle Eigenwerte von A positiv. Sei $x \neq 0$. Weil A unitär diagonalisierbar ist, gilt $x = \sum_{i=1}^k x_i$ mit $x_i \in \ker(\lambda_i - A)$, wobei die λ_i paarweise verschiedene Eigenwerte von A sind. Dabei sind nicht alle $x_i = 0$. Wir nützen nun aus, dass die Eigenräume aufeinander orthogonal stehen und erhalten

$$\begin{aligned} \langle x, Ax \rangle &= \sum_{i=1}^k \sum_{j=1}^k \langle x_i, Ax_j \rangle = \sum_{i=1}^k \sum_{j=1}^k \langle x_i, \lambda_j x_j \rangle \\ &= \sum_{i=1}^k \sum_{j=1}^k \lambda_j \langle x_i, x_j \rangle = \sum_{i=1}^k \lambda_i \|x_i\|^2 > 0. \end{aligned}$$

Der gleiche Beweis funktioniert auch für Semidefinitheit. \square

LEMMA 3.2.3. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $A \in \mathbb{K}^{n \times n}$ eine hermitesche, positiv definite Matrix. Sei $b \in \mathbb{K}^n$ und $\gamma \in \mathbb{R}$ (reell!). Wir betrachten die hermitesche $(n+1) \times (n+1)$ -Blockmatrix

$$M := \begin{pmatrix} A & b \\ b^* & \gamma \end{pmatrix}.$$

Dann ist M genau dann positiv definit, wenn $\det(M) > 0$ gilt.

BEWEIS. Weil A positiv definit ist, ist A regulär. Wir betrachten die reguläre Blockmatrix

$$T = \begin{pmatrix} E_n & -A^{-1}b \\ 0 & 1 \end{pmatrix}$$

und

$$\begin{aligned} \tilde{M} &:= T^* M T = \begin{pmatrix} E_n & 0 \\ -b^* A^{-1} & 1 \end{pmatrix} \begin{pmatrix} A & b \\ b^* & \gamma \end{pmatrix} \begin{pmatrix} E_n & -A^{-1}b \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} E_n & 0 \\ -b^* A^{-1} & 1 \end{pmatrix} \begin{pmatrix} A & 0 \\ b^* & -b^* A^{-1}b + \gamma \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & \gamma - b^* A^{-1}b \end{pmatrix}. \end{aligned}$$

Die Matrix T ist eine obere Dreiecksmatrix mit Einsen auf der Diagonale, daher ist $\det(T) = 1$ und auch $\det(T^*) = \det(T) = 1$, und

$$\det(\tilde{M}) = \det(T^*) \det(M) \det(T) = \det(M).$$

Damit ist

$$\det(M) = \det(\tilde{M}) = (\gamma - b^* A^{-1}b) \det(A) > 0$$

genau dann wenn $\gamma - b^* A^{-1}b > 0$.

Andererseits gilt für alle $x \in \mathbb{K}^{n+1}$

$$x^* M x = [T^{-1}x]^* T^* M T [T^{-1}x] = [T^{-1}x]^* \tilde{M} [T^{-1}x].$$

Daher ist M genau dann positiv definit, wenn \tilde{M} positiv definit ist. Die Blockdiagonalmatrix \tilde{M} ist aber genau dann positiv definit, wenn ihre Diagonallöcke A und $\gamma - b^* A^{-1}b$ positiv definit sind. Nun ist A nach Voraussetzung positiv definit, und es bleibt die Bedingung dass $\gamma - b^* A^{-1}b > 0$, d.h., $\det(M) > 0$. \square

SATZ 3.2.4 (Hauptminorenkriterium für Definitheit). Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $A = (\alpha_{i,j})_{i,j=1,\dots,n} \in \mathbb{K}^{n \times n}$ eine hermitesche Matrix. Wir definieren für $k = 1, \dots, n$

$$A_k := \begin{pmatrix} \alpha_{1,1} & \cdots & \alpha_{1,k} \\ \vdots & & \vdots \\ \alpha_{k,1} & \cdots & \alpha_{k,k} \end{pmatrix}.$$

Dann ist A genau dann positiv definit, wenn für alle $k = 1, \dots, n$ gilt $\det(A_k) > 0$.

BEWEIS. Sei zunächst A positiv definit. Für $y = (\eta_1, \dots, \eta_k)^T \in \mathbb{K}^k \setminus \{0\}$ setzen wir $\tilde{y} = (\eta_1, \dots, \eta_k, 0, \dots, 0)^T \in \mathbb{K}^n$. Dann ist $y^* A_k y = \tilde{y}^* A \tilde{y} > 0$. Also ist A_k positiv definit und hat nach Satz 3.2.2 lauter positive Eigenwerte. Die Determinante von A_k ist aber das Produkt der Eigenwerte und muß folglich auch positiv sein.

Es sei nun $\det(A_k) > 0$ für alle k . Wir zeigen durch Induktion nach k : Die Matrix A_k ist positiv definit. Für $k = n$ ergibt sich dann: $A = A_n$ ist positiv definit.

Verankerung $k = 1$. Es ist $A_1 = (\alpha_{1,1})$. Offensichtlich ist diese 1×1 -Matrix genau dann positiv definit, wenn $\alpha_{1,1} > 0$. Es ist aber $\det(A_1) = \alpha_{1,1}$.

Induktionsschritt: Sei A_k positiv definit. Wir können A_{k+1} als Blockmatrix mit einem geeigneten Vektor $b \in \mathbb{K}^n$ und $\gamma = \alpha_{k+1,k+1} \in \mathbb{R}$ schreiben:

$$A_{k+1} = \begin{pmatrix} A_k & b \\ b^* & \gamma \end{pmatrix}.$$

Nach Lemma 3.2.3 ist A_{k+1} genau dann positiv definit, wenn $\det(A_{k+1}) > 0$. \square

BEMERKUNG 3.2.5. Beachten Sie aber, dass sich das Hauptminorenkriterium nicht auf semidefinite Matrizen verallgemeinern lässt. Es sind alle Hauptminoren der folgenden Matrix nichtnegativ, aber die Matrix ist weder positiv noch negativ semidefinit:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

3.3. Sesquilinearformen und Matrizen.

DEFINITION 3.3.1. Sei V ein Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Eine Abbildung

$$a : \begin{cases} V \times V & \rightarrow \mathbb{K}, \\ (x, y) & \mapsto a(x, y) \end{cases}$$

heißt

Bilinearform, wenn gilt:	Sesquilinearform, wenn gilt:
$a(z, \lambda x + \mu y) = \lambda a(z, x) + \mu a(z, y)$	$a(z, \lambda x + \mu y) = \lambda a(z, x) + \mu a(z, y)$
$a(\lambda x + \mu y, z) = \lambda a(x, z) + \mu a(y, z)$	$a(\lambda x + \mu y, z) = \overline{\lambda} a(x, z) + \overline{\mu} a(y, z)$

Eine Bilinearform bzw. Sesquilinearform a heißt

symmetrisch, wenn gilt:	hermitesch, wenn gilt:
$a(y, x) = a(x, y),$	$a(y, x) = \overline{a(x, y)}.$

Eine Bilinearform bzw. Sesquilinearform a heißt

positiv semidefinit, wenn für alle $x \in V$ gilt: $a(x, x) \in [0, \infty)$,
 positiv definit, wenn für alle $x \in V \setminus \{0\}$ gilt $a(x, x) \in (0, \infty)$.

BEMERKUNG 3.3.2. In einem Vektorraum über \mathbb{R} sind also Bilinearform und Sesquilinearform dasselbe.

Ein inneres Produkt ist in dieser Terminologie eine hermitesche, positiv definite Sesquilinearform.

BEHAUPTUNG 3.3.3. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

- 1) Sei $a : \mathbb{K}^n \times \mathbb{K}^n \rightarrow \mathbb{K}$ eine Sesquilinearform. Dann gibt es eine eindeutig bestimmte Matrix $A \in \mathbb{K}^{n \times n}$ sodass für alle $x, y \in \mathbb{K}^n$ gilt $a(x, y) = x^*Ay$.
- 2) Ist $A \in \mathbb{K}^{n \times n}$, so ist die Abbildung

$$a : \begin{cases} \mathbb{K}^n \times \mathbb{K}^n & \rightarrow \mathbb{K} \\ (x, y) & \mapsto x^*Ay \end{cases}$$

eine Sesquilinearform.

- 3) Sind e_i die Einheitsvektoren in \mathbb{K}^n , so ist $A = (\alpha_{i,j})_{i,j=1,\dots,n}$ gegeben durch $\alpha_{i,j} = a(e_i, e_j)$.
- 4) Die Matrix A ist genau dann selbstadjungiert, wenn a hermitesch ist.
- 5) Die Matrix A ist genau dann positiv (semi)-definit, wenn a positiv (semi)-definit ist.

BEWEIS. Beweis als Übung. □

BEHAUPTUNG 3.3.4. Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und $B = (b_1, \dots, b_n)$ eine Basis des \mathbb{K}^n . Mit T_E^B bezeichnen wir wie immer die Übergangsmatrix von der Basis B zur Einheitsbasis, d.i. die Matrix, deren Spalten die Basisvektoren b_i sind. Sei $A \in \mathbb{K}^{n \times n}$ und a die Sesquilinearform $a(x, y) = x^*Ay$.

Sind \tilde{x}, \tilde{y} die Koordinaten bezüglich B von Vektoren $x, y \in \mathbb{K}^n$, so ist

$$a(x, y) = \tilde{x}^*C\tilde{y} \quad \text{mit } C = (T_E^B)^*AT_E^B.$$

BEWEIS. Weil $x = T_E^B\tilde{x}$ und $y = T_E^B\tilde{y}$, gilt

$$x^*Ay = (T_E^B\tilde{x})^*A(T_E^B\tilde{y}) = \tilde{x}^*[T_E^B]^*AT_E^B\tilde{y}.$$

□

Beachten Sie: Wenn A als die Matrix einer linearen Abbildung $f : x \mapsto Ax$ aufgefasst wird, geschieht der Basiswechsel durch Ähnlichkeit: $[T_E^B]^{-1}AT_E^B$. Wenn aber A als die Matrix einer Sesquilinearform $a(x, y) = x^*Ay$ aufgefasst wird, geschieht der Basiswechsel mit Hilfe der Adjungierten: $[T_E^B]^*AT_E^B$. Wenn B eine Orthonormalbasis ist, ist $[T_E^B]^* = [T_E^B]^{-1}$, und die Basiswechsel für lineare Abbildungen und für Sesquilinearformen sind gleich.