

Diese Notizen¹ sind während der Ausarbeitung der Anleitung zur Linearen Algebra I im Wintersemester 2010/2011 entstanden. Sie werden frei zur Verfügung gestellt, enthalten aber nicht alles, was in der Anleitung behandelt wurde, insbesondere nicht die veranschaulichenden Zeichnungen. Wer Fehler in diesem Dokument findet, schreibe bitte im Interesse seiner Kommilitonen eine Email an zemke@tu-harburg.de mit dem Betreff [Anleitung LAI]. Der Text wird dann bei Bedarf verbessert und/oder ergänzt.

Inhaltsverzeichnis

1	Anleitung vom 29.10.2010	2
1.1	Aussagenlogik	2
1.2	Zahlbegriffe	3
1.3	Komplexe Zahlen	4
1.4	Funktionen	8
2	Anleitung vom 12.11.2010	9
2.1	Orthogonalität	9
2.2	Geometrie im \mathbb{R}^n , $n = 2, 3$	12
2.3	Auswirkungen der Orthogonalität	18
3	Anleitung vom 26.11.2010	18
3.1	Vektorraum	19
3.2	Unterraum	20
3.3	Linearkombinationen; lineare Abhängigkeit	21
3.4	Basis; Dimension	23
4	Anleitung vom 10.12.2010	27
4.1	Skalarprodukte: Winkel & Längen	27
4.2	Orthogonalbasen im Polynomraum	31
4.3	Normen & Metriken	32
4.4	Orthogonal- und Orthonormalbasen	34
4.5	Metriken	36
4.6	Lineare Gleichungssysteme	37
4.7	Der Gauß-Algorithmus	40
5	Anleitung vom 07.01.2011	41
5.1	Lineare Mannigfaltigkeiten	42
5.2	Lineare Abbildungen	42
5.3	Matrixmultiplikation I: Vektoren	44
5.4	Matrixdarstellung linearer Abbildungen	44
5.5	Matrixmultiplikation II: Matrizen	46
6	Anleitung vom 21.01.2011	50
6.1	Kongruenztransformationen	50
6.2	Multiplikationstraining	53
6.3	Das Gaußsche Eliminationsverfahren	54
6.4	Die LR-Zerlegung	60
6.5	Regularität & Singularität	64

¹Version vom 6. März 2011, erstellt um 20:18 von Jens-Peter M. Zemke.

1 Anleitung vom 29.10.2010

Themen der Anleitung sind Aussagenlogik; Zahlen (natürliche, ganze, rationale, reelle und komplexe; besonderes Augenmerk ist dabei auf die zugrunde liegende Motivation und Veranschaulichung gerichtet), insbesondere der gravierende Unterschied zwischen den reellen und den komplexen Zahlen, namentlich der Begriff der Ordnung. Des weiteren werden die Begriffe der Surjektivität, Injektivität und Bijektivität von Funktionen eingeführt.

1.1 Aussagenlogik

In der Aussagenlogik beschäftigt man sich mit Aussagen und deren Wahrheitswert. Dabei geht man grundsätzlich davon aus, dass die behandelten Aussagen entweder wahr (exklusiv) oder falsch sind, das ist das sogenannte „Bivalenzprinzip“. Dabei bezeichnet man die beiden möglichen Werte „wahr“ und „falsch“ als Wahrheitswerte, meist abgekürzt zu den Buchstaben „w“ für „wahr“ und „f“ für „falsch“. Damit sind Aussagen der Form „Linux ist besser als Windows“ oder umgekehrt grundsätzlich nicht behandelbar und im Folgenden ausgeschlossen. Aussagen kann man (wie soeben geschehen) mittels „und“ und „oder“ (im Sinne von exklusiv oder) verknüpfen und mit „nicht“ negieren. Aussagen werden meist mit Großbuchstaben bezeichnet, also A, B, C, D, \dots ; die mathematischen Symbole zur Kombination und Negation sind dabei dann für „oder“ das \vee (soll an lat. „vel“ erinnern), für „und“ das \wedge und für „nicht“ das \neg . Zum besseren Verständnis werden Klammern gebraucht. Für Grundaussagen (atomare Aussagen; solche, die sich nicht in weitere Teilaussagen zerlegen lassen) A und B ist also $\neg(A \vee B)$ eine Aussage. Ein zweites Grundprinzip der Aussagenlogik ist es, dass der Wahrheitswert jede solcherart zusammengesetzte Aussage sich aus den Wahrheitswerten der atomaren Aussagen und den Verknüpfungen berechnen lässt. Seien die beiden Aussagen

$$A := \text{„Ich bin müde.“} \quad \text{und} \quad B := \text{„Ich sitze in der Anleitung.“} \quad (1)$$

gegeben. Dann kann jeder für sich entscheiden, welche Aussage wahr oder falsch ist. Nehmen wir mal an, dass A für die entsprechende Person falsch, und B wahr ist. Die kombinierte Aussage $\neg(A \vee B)$ lautet dann

$$\neg(A \vee B) = \text{„Nicht (ich bin müde oder ich sitze in der Anleitung).“} \quad (2)$$

Welchen Wahrheitswert hat diese Aussage im oben betrachteten Szenario? Dazu verwendet man sogenannte Wahrheitstabeln:

A	$\neg A$	A	B	$A \vee B$	$\neg(A \vee B)$
f	w	f	f	f	w
f	w	f	w	w	f
w	f	w	f	w	f
		w	w	w	f

(3)

Man sieht, unter der gegebenen „Belegung“ der Wahrheitswerte für die Aussagen A und B ist die Aussage $\neg(A \vee B)$ falsch. Nun würde (fast) niemand die Aussage in dieser Form (2) wiedergeben, eher würde man sagen

$$\text{„Ich bin nicht müde und ich sitze nicht in der Anleitung.“}, \quad (4)$$

also $\neg A \wedge \neg B$. Sind diese Aussagen für jede Belegung der Aussagen A und B äquivalent, d. h., sagen dasselbe aus? Dazu verwenden wir wieder eine Wahrheitstafel:

A	B	$A \vee B$	$\neg(A \vee B)$	$\neg A$	$\neg B$	$\neg A \wedge \neg B$
f	f	f	w	w	w	w
f	w	w	f	w	f	f
w	f	w	f	f	w	f
w	w	w	f	f	f	f

(5)

Man sieht, dass die beiden Spalten zu den Aussagen $\neg(A \vee B)$ und $\neg A \wedge \neg B$ die selben Wahrheitswerte ergeben. Damit sind diese Aussagen wirklich äquivalent, in Zeichen,

$$(\neg(A \vee B)) \Leftrightarrow (\neg A \wedge \neg B). \quad (6)$$

Dass der Wahrheitsgehalt jede Aussage sich aus den Teilaussagen und der Art der Zusammensetzung herleiten lässt nennt man das „Extensionalitätsprinzip“. Aussagen wie die in Gleichung (6), welche immer wahr sind, egal, wie die Teilaussagen belegt werden, nennt man Tautologien. Tautologien sind das, was in der Mathematik als „Theorem“ oder „Satz“ bekannt ist. Ein bekannter Satz ist der vom ausgeschlossenen Dritten (lat.: tertium non datur):

$$P \vee \neg P \quad (7)$$

ist immer wahr. Das bedeutet, jede Aussage ist *entweder* wahr *oder* falsch.

Von der Aussagenlogik geht man noch etwas weiter, nämlich zur Prädikatenlogik. Dort gibt es ausser Aussagen auch noch Prädikate (z. B. $S =$ „schwimmt auf dem See“) und Quantoren, nämlich den „Existenzquantor“ \exists und denn „Allquantor“ \forall . Damit kann man dann schon Aussagen bilden wie

$$\forall e \in M : S(e), \quad (8)$$

wobei hier die Variable e für beliebige Elemente aus der Menge M meiner Entchen steht.

1.2 Zahlbegriffe

Sie sollten aus der Schulmathematik und/oder dem Vorkurs der TUHH etwas mit den Zeichen \mathbb{N} , \mathbb{Z} , \mathbb{Q} und \mathbb{R} anfangen können. Wesentlich sind hier (in einem Mathematik-lastigen Universitätsstudium) einige Aspekte derselben. Die *natürlichen Zahlen* \mathbb{N} liefern historisch gesehen die natürlichste Vergleichsbasis von zählbaren Dingen, wie Schafen, Kühen oder Hühnern. Die Erfindung der Zahl Null ist ein zur Zeit der Babylonier erfolgter, nicht zu vernachlässigender Fortschritt.

Naheliegend ist die Idee einer „Verknüpfung“ von natürlichen Zahlen (2 Äpfel plus 3 Äpfel), welche als *Addition* bezeichnet wird und mit $+$ bezeichnet wird. Die *Multiplikation* \cdot ist schon eine gedanklich sehr viel weiter entfernte „Verknüpfung“, denn was wäre denn das Ergebnis von „2 Äpfel mal 3 Äpfel“? Hier liegt ein wesentlicher Abstraktionsschritt, nämlich die Abstraktion der Zahlen von der zugrunde liegenden Einheit. Diese Multiplikation über den natürlichen Zahlen ist eine Kurzschreibweise für die wiederholte Addition von gleichen Mengen.

Der nächste, in der Schule selten entsprechend gewürdigte Abstraktionsschritt ist die Frage nach der „Umkehrbarkeit“ (*Inversion*) der Verknüpfungen der Addition und Multiplikation. Der erste Schritt ist nicht immer möglich, daher wurden früh in der Menschheitsgeschichte Schulden, also negative Zahlen, und damit die *ganzen Zahlen* \mathbb{Z} „erfunden“. Der Gebrauch der negativen Zahlen hat sich zu den Zeiten der Medicis und Fuggers stark verbreitet. In der modernen Mathematik sind die ganzen Zahlen die *eindeutigen* Lösungen $x \in \mathbb{Z}$ von Gleichungen der Form $a + x = b$, für beliebig gegebene $a, b \in \mathbb{Z}$. Man sagt dazu auch, die Menge \mathbb{Z} , genauer: die (*additive*) *Gruppe* $(\mathbb{Z}, +)$, ist bzgl. der Addition und Inversenbildung ($-n$ ist das *additive Inverse* zu n) *abgeschlossen*. Die Umkehrung der Multiplikation in \mathbb{Z} führt analog auf die *rationalen Zahlen* \mathbb{Q} . Die (multiplikative) *Gruppe* (\mathbb{Q}^*, \cdot) , mit $\mathbb{Q}^* := \mathbb{Q} \setminus \{0\}$, ist wiederum bzgl. der Multiplikation und Inversenbildung ($1/r = q/p$ ist das *multiplikative Inverse* zu $r = q/p$) *abgeschlossen*.

Auf diesem Stand der Mathematik verharren die Griechen lange Zeit, sie glaubten, alle Strecken (bei ihnen überwog die geometrische Sicht der Dinge) seien „kommensurabel“, also mit einem gleichen Mass messbar. In heutige mathematische Sprache übersetzt heisst das, zu zwei gegebenen Zahlen gibt es immer eine Zahl, so dass die beiden Zahlen Vielfache dieser sind. Vom lateinischen „Ratio“ (Verhältnis) nennt man daher die in der Schule oft als Bruchzahlen bezeichneten Zahlen *rationale* Zahlen. Die Entdeckung, dass im Pentagon ein inkommensurables Streckenverhältnis, also eine *irrationale* Zahl, in Form des Goldenen Schnittes $\Phi = (1 + \sqrt{5})/2$ auftaucht, zerstörte ein Weltbild.

Die Erweiterung der rationalen Zahlen zu allen reell beobachtbaren (also in der Form von Streckenlängen) auftauchenden Zahlen, den sogenannten *reellen Zahlen* \mathbb{R} hängt mit einem Anspruch an die „Vollständigkeit“ der Zahlen zusammen. Man möchte sicher sein, dass bei beliebigen, auch ins Unendliche weitergehenden Vorgängen, keine neuen Zahlen mehr auftauchen. Die reellen Zahlen erhält man aus der „Vervollständigung“ der rationalen Zahlen. Die (später Ihnen eventuell in der Analysis begegnenden) bekanntesten Methoden der „Vervollständigung“ der rationalen Zahlen zu den reellen Zahlen sind die mittels Äquivalenzklassen von Cauchy-Folgen, Äquivalenzklassen von Intervallschachtelungen und Dedekindscher Schnitte. Solche Techniken werden selten akkurat in der Schulmathematik behandelt. In der Schule werden reelle Zahlen meist mit ihrer Veranschaulichung mittels der Zahlengeraden eingeführt und behandelt.

Den meisten Erstsemestern wahrscheinlich eher nicht so vertraut wie die reellen Zahlen sind die sogenannten *komplexen Zahlen* \mathbb{C} .

1.3 Komplexe Zahlen

Die komplexen Zahlen sind motiviert aus der Unzufriedenheit über die eventuelle Nichtlösbarkeit algebraischer (mittels Addition und Multiplikation aus einfachen Termen erzeugter) Gleichungen (Polynome), als einfachstes Beispiel beobachten wir

$$\forall x \in \mathbb{R} : x^2 + 1 \geq 1 \neq 0. \quad (9)$$

Demnach hat das quadratische Polynom $p(x) := x^2 + 1$ keine (reelle) Nullstelle. Das folgt auch ziemlich schnell aus der sogenannten „pq-Formel“ (aus binomischer Ergänzung), da der unter der Wurzel entstehende Ausdruck negativ ist. Im Allgemeinen ist es sehr schwierig, einem Polynom anzusehen, ob, und, allgemeiner, wieviele reelle Nullstellen es hat.

Als Gedankenspielerie begonnen, indem einfach „imaginäre Zahlen“ eingeführt wurden, zuerst die imaginäre Einheit

$$i := \sqrt{-1}, \quad (10)$$

zeigte sich, dass viele Gleichungen erst unter Verwendung von gemischt reellen Zahlen $x \in \mathbb{R}$ und imaginären Zahlen iy , $y \in \mathbb{R}$, sogenannten komplexen Zahlen \mathbb{C} , meist bezeichnet mit dem Buchstaben z ,

$$z := x + iy, \quad (11)$$

behandelbar wurden oder stark vereinfacht wurden. (Beispiele: Auflösbarkeit von linearen Differentialgleichungen, Eigenwertaufgaben, viele Vorgänge die auf Schwingungen basieren, Maxwell'sche Gleichungen in der reellen Form und komplexen Form.)

Die reellen Zahlen haben gewisse Eigenschaften, sie erfüllen sogenannte „Rechenregeln“, festgehalten in der Form von „Axiomen“, welche gerade das Rechnen mit ihnen ausmachen (Genauer und jetzt noch viel zu abstrakt: $(\mathbb{R}, +)$ und $(\mathbb{R}^* := \mathbb{R} \setminus \{0\}, \cdot)$ sind abelsche (kommutative) Gruppen, es gilt das Distributivgesetz, es gibt eine Ordnung \leq auf \mathbb{R} und bezüglich dieser ist \mathbb{R} ordnungsvollständig).

Solche Rechenregeln sollten auch für diese gemischten komplexen Zahlen erfüllt sein. Leider geht gerade die Ordnung verloren, welche die Existenz einer Vergleichbarkeit $<$ zwischen beliebigen Zahlen $a, b \in \mathbb{R}$ behandelt:

$$\forall a, b \in \mathbb{R} : \{a < b\} \vee \{a = b\} \vee \{a > b\}. \quad (12)$$

Einige dieser Eigenschaften haben weitreichende Konsequenzen, z. B. beim Beweis, dass das arithmetische Mittel von n reellen Zahlen immer zwischen dem Minimum und dem Maximum liegt, oder inwieweit die komplexen Zahlen doch noch eine „Ordnung“ aufweisen (Schranken für den Real- und Imaginärteil; Dreiecksungleichung). Obacht ist dabei insbesondere bei der Notation und mathematischen Strenge gewünscht: so sollte man nicht auf die Idee verfallen, die im reellen gültigen „Ordnungsaxiome“ für die komplexen Zahlen zu verwenden. Wie lassen sich denn die Zahlen 1 und i „vergleichen“? Welche Zahl der beiden ist größer oder kleiner?

Dank Gauß (Deutschland) und Argand (Frankreich) ist die Behandlung komplexer Zahlen heute alles andere als „komplex“ oder „imaginär“. Beide entdeckten, dass die komplexen Zahlen $z \in \mathbb{C}$ sich, wie $x \in \mathbb{R}$ als Zahlengerade, als *Zahlenebene* auffassen lassen.

Die Addition zweier komplexer Zahlen ist dabei komponentenweise definiert, also die übliche Addition zweier Vektoren in der Ebene. Die Subtraktion ist die Addition des negierten Wertes, also eine Umkehrung der „Richtung“ bei dem zweiten „Vektor“ und dann eine Aneinanderlegung der „Vektoren“.

Die Multiplikation ist am besten verständlich, wenn man die komplexe Exponentialfunktion (welche wir erstmal nur für rein reelle oder rein imaginäre Argumente kennenlernen werden) verwendet. Es gilt die nach Euler benannte Gleichung

$$e^{i\phi} = \cos(\phi) + i \sin(\phi), \quad \phi \in \mathbb{R}, \quad (13)$$

welche im Skript/in der Vorlesung als Definition gehandelt wird. Aus den Additionstheoremen folgert man, dass die fundamentale Eigenschaft der Exponentialfunktion

$$e^{i(\phi+\psi)} = e^{i\phi+i\psi} = e^{i\phi} e^{i\psi} \quad (14)$$

auch für rein imaginäre Argumente gilt. Das Wechselspiel zwischen einer (komplexen) Multiplikation (von komplexen Zahlen auf dem Einheitskreis) und der Addition der Winkel liefert eine einfache geometrische Interpretation der Multiplikation beliebiger komplexer Zahlen.

Nun liegen die wenigsten komplexen Zahlen auf dem Einheitskreis, also muss man zu dem (immer nur bis auf Vielfache von 2π , dem Umfang des Kreises; im Bogenmaß bestimmten) Winkel noch einen Radius angeben. Diese beiden (reellen) Größen charakterisieren wie der Real- und Imaginärteil die komplexe Zahl eindeutig, wenn man den Winkel einschränkt und den Radius Null gesondert behandelt. Es gilt für alle $z \in \mathbb{C}^* := \mathbb{C} \setminus \{0\}$

$$z = r e^{i\phi}, \quad \text{wobei} \quad \phi \in (-\pi, \pi]. \quad (15)$$

Das so eindeutig bestimmte ϕ (griechischer Buchstabe phi, gesprochen wie *vieh*, manchmal auch geschrieben als φ) nennt sich *Hauptwert* des Argumentes. *Das Argument* ist immer gefragt unter Angabe der Vielfachen von 2π . Ein Hinweis: An der Universität wird im Bogenmaß gerechnet. Ein Einheitskreis hat einen Umfang von 2π , also entspricht ein Vollkreis dem Winkel 2π , ein Halbkreis dem Winkel π und ein Viertelkreis dem Winkel $\pi/2$. Zwei aufeinander senkrecht stehende Geraden haben demnach einen Winkel von $\pi/2$ zueinander.

Die Multiplikation lässt sich nach den Regeln der Multiplikation reeller Zahlen, der fundamentalen Eigenschaft der Exponentialfunktion und der geometrischen Interpretation der Polarkoordinaten schreiben als

$$z_1 \cdot z_2 = r_1 e^{i\phi_1} \cdot r_2 e^{i\phi_2} = (r_1 \cdot r_2) e^{i\phi_1} \cdot e^{i\phi_2} = (r_1 \cdot r_2) e^{i\phi_1+i\phi_2} = (r_1 \cdot r_2) e^{i(\phi_1+\phi_2)}. \quad (16)$$

Damit erzeugt die Multiplikation zweier komplexer Zahlen eine neue komplexe Zahl, deren Winkel aus der Addition der beiden Winkel (modulo Vielfache von 2π) und deren Radius aus der Multiplikation der beiden Radien berechnet wird.

Um die umgekehrte Operation, also das Teilen zu verstehen, betrachtet man am besten die Operation z_1/z_2 als Multiplikation von z_1 mit $1/z_2$. Für diese Inversion gilt (dieses ist der rechnerische Weg):

$$\frac{1}{z} = \frac{1}{z} \frac{\bar{z}}{\bar{z}} = \frac{\bar{z}}{|z|^2}. \quad (17)$$

In Polarkoordinaten ergibt sich wegen $|z|^2 = r^2$

$$z = r e^{i\phi}, \quad \frac{1}{z} = \frac{1}{r} e^{-i\phi}. \quad (18)$$

Das bedeutet geometrisch, dass die neue Zahl durch Spiegelung an der reellen Achse (komplexe Konjugation) und „Inversion am Einheitskreis“ (Übergang vom Radius r zum Radius $1/r$) entsteht. Wie man diese Operationen an Beispielen durchführt, also Ergebnisse berechnet, ist Teil des aktuellen Übungsblattes.

Die Veranschaulichung der Operationen $+$, \cdot , $-$, $/$ als geometrische Operationen in der komplexen Ebene (die reelle Ebene mit der zusätzlichen Multiplikation der Tupel (x, y) aus Real- und Imaginärteil) ist für das Verständnis (und die Bewältigung der Klausuraufgaben zu komplexen Zahlen) unentbehrlich.

Was auch unentbehrlich ist, ist zu lernen, nicht alles auszurechnen, das ist zwar in der Schule angesagt, aber in der Realität weder gefragt noch teilweise möglich. Insbesondere sind die gegebenen Daten (reelle oder komplexe Zahlen) meist fehlerbehaftet, da sie aus Messungen resultieren. Nette ganze Zahlen und ganzzahlige Rechenergebnisse sind dort fast nie zu finden. Für viele Aussagen genügen grobe Abschätzungen.

Als Beispiele vier Aufgaben zu den komplexen Zahlen aus alten Klausuren. Die erste Aufgabe ist die Bewertung der Aussage, dass für alle $z \in \mathbb{C}$

$$\Re(z^2) \leq (\Re(z))^2. \quad (19)$$

Man probiert zuerst meist aus, das Gegenteil zu zeigen, dazu nimmt man ein paar „zufällige“ Werte, also z. B. $z = 1, -1, i, -i, 1 + i$ oder so. Man wird sehen, dass diese Aussage immer stimmt. Also sollte man sie allgemein beweisen können. Wir haben zwei allgemeine Darstellungen/Anschauungen komplexer Zahlen: Die als Vektoren in der Ebene oder mittels Polarkoordinaten.

Hier zeigt sich, trotz der Multiplikation (Quadrieren), dass die erste Darstellung die angepasste ist. Es gilt

$$z = a + ib, \quad a, b \in \mathbb{R}, \quad (20)$$

$$z^2 = (a + ib)(a + ib) = a^2 + (ib)^2 + 2iab = a^2 - b^2 + i2ab. \quad (21)$$

Der Vergleich der Realteile beider Seiten lautet

$$a^2 - b^2 \leq a^2, \quad (22)$$

eine Aussage, welche für beliebige reelle Zahlen a, b immer erfüllt ist.

Natürlich kann man den Beweis auch in der Polarkoordinatendarstellung führen:

$$z = re^{i\phi}, \quad z^2 = r^2 e^{i2\phi}, \quad (23)$$

der Vergleich der beiden Realteile läuft also auf den Vergleich zweier Kosinus-Ausdrücke hinaus,

$$r^2 \cos(2\phi) \leq (r \cos(\phi))^2 = r^2 \cos^2(\phi), \quad (24)$$

welches sich dank einer Division durch r^2 (für $r = 0$ ist die Aussage trivial) und Anwendung der Additionstheoreme umschreiben lässt in

$$\cos^2(\phi) - \sin^2(\phi) \leq \cos^2(\phi), \quad (25)$$

eine auch klarerweise wahre Aussage. Dazu muss man aber die Additionstheoreme können (oder sich aus der Euler-Formel schnell wieder herleiten).

Die zweite Aufgabe aus einer alten Klausur basierte auf den komplexen Zahlen

$$z_1 := 1 - i, \quad z_2 := \frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}}. \quad (26)$$

Zu bewerten war die Aussage, ob

$$|z_1^8 - z_2^8| > 10 \quad (27)$$

gelte. Natürlich kann man die Zahlen einfach ausrechnen und damit dann den Absolutbetrag berechnen. Man kann auch geometrisch vorgehen und die Lage der beiden Zahlen ungefähr

bestimmen. Am einfachsten ist die Aussage aber zu bewerten, wenn man sich nur die Radien anschaut, es gilt

$$r_1 = |z_1| = \sqrt{1^2 + 1^2} = \sqrt{2}, \quad (28)$$

$$r_2 = |z_2| = \sqrt{\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2} = 1. \quad (29)$$

Die Radien von $\{z_i^8\}_{i=1}^2$ sind demnach

$$r_1^8 = \left(2^{1/2}\right)^8 = 2^4 = 16, \quad r_2^8 = 1. \quad (30)$$

Damit muss man von dem Kreis mit Radius 16 um höchstens ± 1 abweichen, was als kleinsten Kreis einen mit Radius 15 ergibt, welcher deutlich ausserhalb einem mit Radius 10 liegt. Also ist die Behauptung eine wahre Aussage.

Eine dritte Klausuraufgabe war es, zu bewerten, ob aus

$$\Re(z) > 1 \quad (31)$$

die Implikation

$$\Re(z^2) > 1 \quad (32)$$

folgt.

Da

$$z = a + ib \quad \text{und} \quad z^2 = a^2 - b^2 + i2ab \quad (33)$$

gilt, muss man nur $|b|$ gross genug wählen, z. B. $a = 2, b = 2$ und erhält

$$\Re(z) = a = 2 > 1, \quad \Re(z^2) = a^2 - b^2 = 2^2 - 2^2 = 0 < 1. \quad (34)$$

Also ist die Behauptung eine falsche Aussage.

Die nächste Aufgabe war aufgrund von Schreibfehlern so in keiner Klausur enthalten, sie sollte aber wie folgt lauten. Seien die beiden komplexen Zahlen

$$z_1 := \sqrt{2} + i\sqrt{2}, \quad z_2 := \frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}} \quad (35)$$

definiert. Gilt dann

$$\Re(z_2^3 + \bar{z}_1) = 1? \quad (36)$$

Diese Frage bewertet man am einfachsten anhand einer Skizze, in der Anleitung geschehen. Hier folgt ein rechnerischer Versuch, die gezeigten/gezeichneten Schritte zu veranschaulichen.

Die Polardarstellung der beiden komplexen Zahlen lässt sich schnell an einer Skizze ablesen, es gilt (unter Auslassung des Termes $2\pi k, k \in \mathbb{Z}$)

$$z_1 = 2 \cdot e^{i\frac{\pi}{4}}, \quad z_2 = 1 \cdot e^{i\frac{\pi}{4}}. \quad (37)$$

Daraus folgt (Winkeladdition), dass

$$z_2^3 = \left(1 \cdot e^{i\frac{\pi}{4}}\right)^3 = 1 \cdot e^{i\frac{3\pi}{4}} = \frac{-1}{\sqrt{2}} + i\frac{1}{\sqrt{2}}. \quad (38)$$

Die Konjugation (Spiegelung an der reellen Geraden) ergibt

$$\bar{z}_1 = 2 \cdot e^{-i\frac{\pi}{4}} = \sqrt{2} - i\sqrt{2}. \quad (39)$$

Die Summation der Zahlen entspricht zeichnerisch also gerade der Vektoraddition eines Vektors der Länge Eins in Richtung $(-1, 1)^T$ und eines Vektors der Länge 2 in Richtung $(1, -1)^T$. Damit

ist das Ergebnis ein Vektor der Länge Eins in Richtung $(1, -1)^T$. Es handelt sich beim Ergebnis also um eine komplexe Zahl mit Länge Eins und einem Nichtnull-Imaginärteil, also

$$z_2^3 + z_1 = a + ib, \quad a^2 + b^2 = 1, \quad b \neq 0. \quad (40)$$

Damit ist die Behauptung, dass der Realteil des Ergebnisses, also a , gleich Eins sei, nicht mehr erfüllbar.

Man konnte natürlich alles explizit ausrechnen, aber hier sollte der einfachere geometrische Weg veranschaulicht werden, und daher wurden viele Größen nicht genau ausgerechnet.

1.4 Funktionen

Eine Relation ist eine Teilmenge des kartesischen Produktes zweier Mengen, hier meist $\mathbb{R} \times \mathbb{R}$ oder $D \times \mathbb{R}$ mit $D \subset \mathbb{R}$, also eine Teilmenge der reellen Ebene. Beispiele: $x \leq y$ als Relation ist assoziiert mit der Teilmenge

$$\mathbb{T}_{\leq} = \{(x, y) : x, y \in \mathbb{R}, x \leq y\}. \quad (41)$$

Der Einheitskreis ist eine Teilmenge der Ebene, also eine Relation. Die Koordinaten (x, y) stehen in der Relation $x^2 + y^2 = 1$.

Eine Funktion ist eine spezielle Form der Relation, hier wird jedem Element x auf der linken Seite (Linkstotalität) ein eindeutiges Element y auf der rechten Seite zugeordnet (Rechtseindeutigkeit). Immer, wenn diese Eigenschaften erfüllt sind, kann man sich die Zuordnung als Funktionsvorschrift vorstellen, daher die Notation

$$f : D \subset \mathbb{R} \rightarrow \mathbb{R}, \quad (42)$$

$$f : x \mapsto y = f(x). \quad (43)$$

Die Eigenschaften, welche eine Funktion definieren, also, dass sie als Relation linkstotal und rechtseindeutig ist, klingen geradezu danach, dass auch die Eigenschaften „rechtstotal“ und „linkseindeutig“ interessant sein könnten, und sind klarerweise notwendig, um die Relation als „inverse“ Funktion $f^{-1} : y \mapsto x$ aufzufassen.

Dass eine Funktion linkseindeutig ist, bedeutet, dass

$$y = f(x_1) = f(x_2) \Rightarrow x_1 = x_2, \quad (44)$$

oder, äquivalent dazu, dass verschiedene Argumente immer verschiedene Funktionswerte haben,

$$x_1 \neq x_2 \Rightarrow y_1 = f(x_1) \neq f(x_2) = y_2. \quad (45)$$

Man nennt solche Funktionen meist *injektiv* anstatt linkseindeutig, den Begriff linkseindeutig verwendet man eher für Relationen. Den Begriff „injektiv“ sollte man sich merken, da dieses die international gebräuchliche Bezeichnung ist. (Blöde Merkregel: Injektionen bekommt man an einer eindeutigen Stelle des Körpers, nicht an mehreren.)

Dass eine Funktion rechtstotal ist, bedeutet, dass

$$\forall y \exists x : y = f(x). \quad (46)$$

Man nennt solche Funktionen meist *surjektiv* anstatt rechtstotal, den Begriff rechtstotal verwendet man eher für Relationen. Den Begriff „surjektiv“ sollte man sich merken, da dieses die international gebräuchliche Bezeichnung ist.

Wenn eine Funktion injektiv und surjektiv ist, ist sie gleichzeitig eine Funktion der Werte in die Argumente, also eine Funktion in inverser Richtung, damit eine Funktion in beide Richtungen. Diese „umkehrbaren“ Funktionen werden „bijektiv“ genannt. Diese eigentlich der Analysis entstammenden Begriffe korrespondieren im Bereich der linearen Algebra später direkt mit algebraischen Begrifflichkeiten, also sollte man sich diese merken.

Die Begriffe stammen wohl aus dem Französischen, auf Wikipedia wird vermutet, von dem Autorenkollektiv Nicolas Bourbaki. Beispiele dazu sind die Relationen zwischen Menschen und (rechtem) Daumenabdruck, Menschen und Personalausweisnummern. Man veranschauliche sich: Warum sind der Kosinus und der Sinus Funktionen? Woher kommen sie?

Apropos Begriffe: Später in der Vorlesung „Lineare Algebra“ werden Sie den Begriff der „Abbildung“ kennenlernen, welcher eigentlich nichts anderes als der Begriff der „Funktion“ ist. Diese synonym verwendeten Begriffe haben verschiedene historische Kontexte. In der Vorlesung „Lineare Algebra“ spricht man meist von (linearen) Abbildungen, in der Vorlesung „Analysis“ hingegen von Funktionen.

2 Anleitung vom 12.11.2010

Themen der Anleitung sind allgemein Vektoren im zwei- und dreidimensionalen Anschauungsraum. Eng verwandt mit Vektoren ist der Begriff der Linearkombination und der Frage nach der „Kombinierbarkeit“ gegebener Vektoren zu einem gesuchten Vektor. Es werden die Begriffe der orthogonalen Projektion, des Skalarproduktes und der Orthogonalität besprochen. Im dreidimensionalen Fall werden das Kreuzprodukt und das Spatprodukt definiert; im Zusammenhang mit diesen Produkten wird die Berechnung von Flächen und Volumina angesprochen. Desweiteren geben wir einige Details zu Punkten, Geraden und Ebenen im \mathbb{R}^2 und \mathbb{R}^3 sowie Schnitten derselben und Abständen zwischen ihnen.

2.1 Orthogonalität

Definition von Vektoren (Schule: Länge mit Richtung, modelliert als Pfeil; Uni: Anzahl von Komponenten, welche nach gewissen später sauber behandelten Gesetzen „verändert“, z. B. addiert, werden können). Anschauung von Vektoren im \mathbb{R}^n , $n = 2, 3$. Addition. Der Begriff der Linearkombination. Der Begriff der Länge (basierend auf dem Satz des Pythagoras). Der Begriff der Orthogonalität. Definition des Skalarproduktes (im \mathbb{R}^n für $n = 2, 3$ und später sogar allgemeines $n \in \mathbb{N}$). Definition des Kreuzproduktes (nur im \mathbb{R}^3) und des Spatproduktes (nur im \mathbb{R}^3 , da basierend auf dem Kreuzprodukt).

Das Skalarprodukt im \mathbb{R}^n , $n = 2, 3$ ist definiert als

$$\langle a, b \rangle := |a| \cdot |b| \cos(\angle(a, b)). \quad (47)$$

Diese Definition benötigt man später nur zur Berechnung von Winkeln. Besser ist die im Skript bewiesene Aussage

$$\langle a, b \rangle = \sum_{i=1}^n a_i b_i, \quad (48)$$

welche im allgemeinen \mathbb{R}^n als Definition des Standard-Skalarproduktes herangezogen wird und nur für $n = 3$ aus der obigen Definition hergeleitet wird. Was man aber an der Definition auch ohne die für die Berechnung schönere Darstellung sofort abliest, ist aufgrund von $\cos(0) = 1$ die auch für allgemeines $n \in \mathbb{N}$ und allgemeines $a \in \mathbb{R}^n$ gültige Gleichheit

$$\langle a, a \rangle = |a|^2 = \sum_{i=1}^n a_i^2. \quad (49)$$

Zwei Vektoren im \mathbb{R}^2 oder \mathbb{R}^3 sind genau dann senkrecht zueinander, wenn der Kosinus des Winkels zwischen ihnen gleich Null ist. Damit ist dann aber das Skalarprodukt gleich Null. Ein einfacher Test, ob zwei Vektoren senkrecht stehen, ist die Berechnung des Skalarproduktes mittels der umgeschriebenen Variante:

$$a = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad b = \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \quad \langle a, b \rangle = 1 \cdot 2 - 2 \cdot 1 = 0, \quad (50)$$

also stehen die beiden Vektoren a, b aus dem \mathbb{R}^2 aufeinander senkrecht. Anschaulich ist das natürlich sofort klar. Im \mathbb{R}^2 ist zu $(x, y)^T$ immer $(-y, x)^T$ senkrecht. Wenn Vektoren senkrecht aufeinander stehen, so nennt man sie auch orthogonal. Diese Orthogonalität hat viele Anwendungen und ist eine der Hauptbegriffe der Linearen Algebra. Zur eben verwendeten Notation: Ein nachgeführtes oben stehendes T nennet sich „Transponation“ und bedeutet, dass die durch Kommata getrennten in den Klammern stehenden Ausdrücke übereinander in den Klammern stehen sollen, also die Komponenten eines Vektors sind. Es gilt also z. B.

$$(x, y, z)^T = \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \quad (51)$$

Wie man leicht sieht, hat die Schreibweise mit dem T den Vorteil, besser in den Text zu passen.

Eine Anwendung des Skalarproduktes liegt in der Projektion (z. B. bei der orthogonalen Zerlegung von Kräften): Projiziere a auf die Richtung von b . Die Richtung von b kann man durch jedes Nichtnull-Vielfache von b beschreiben, um die Skalierung einfacher zu gestalten, wählt man einen Vektor in die Richtung von b mit der Länge Eins. Diesen berechnet man gemäß (siehe die in der Anleitung gegebene Zeichnung)

$$\frac{b}{|b|} = \frac{b}{\sqrt{\langle b, b \rangle}}. \quad (52)$$

Nun möchte man einen gegebenen Vektor a aus dem selben Raum in die Richtung von b projizieren. Zeichnerisch ist klar, dass die Länge des entstehenden Vektors durch $|a| \cos(\angle(a, b))$ gegeben ist. Damit haben wir die Länge und die Richtung und können den Gesamtvektor konstruieren:

$$P_b a = |a| \cos(\angle(a, b)) \frac{b}{\sqrt{\langle b, b \rangle}}. \quad (53)$$

Wie vorher wird man niemals einen Winkel ausrechnen, wenn man nicht muss. Daher drückt man die Länge durch das Skalarprodukt aus und erhält

$$P_b a = \frac{\langle a, b \rangle}{|b|} \frac{b}{\sqrt{\langle b, b \rangle}}. \quad (54)$$

Die Länge beschreibt man wieder als Wurzel des Skalarproduktes und fasst die Wurzelterme zusammen,

$$P_b a = \frac{\langle a, b \rangle}{\langle b, b \rangle} b = \frac{\sum_{i=1}^n a_i b_i}{\sum_{i=1}^n b_i^2} b. \quad (55)$$

Oft hat man zwei (linear unabhängige) Vektoren im \mathbb{R}^3 , welche z. B. die Richtungsvektoren einer Ebene sind, und möchte einen dazu senkrecht stehenden finden (z. B. um einen Einheitsnormalvektor zu finden). Um solch einen senkrecht auf zwei gegebenen Vektoren stehenden Vektor im \mathbb{R}^3 zu erhalten, kann man einfach das Kreuzprodukt verwenden.

Kreuzprodukt (nur im \mathbb{R}^3 definiert). Das Kreuzprodukt $a \times b \in \mathbb{R}^3$ zweier Vektoren $a, b \in \mathbb{R}^3$ ist definiert durch die folgenden drei Eigenschaften:

1. $|a \times b| = |a| \cdot |b| \cdot |\sin(\angle(a, b))|$,
2. $a \times b \perp a, b$,
3. $(a, b, a \times b)$ bilden ein Rechtssystem.

Es kann in Komponenten ausgedrückt werden mittels

$$a \times b = |a| \cdot |b| \cdot \sin(\angle(a, b)) \cdot e, \quad e \perp a, b, \quad |e| = 1, \quad (56)$$

$$= \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix} \quad (57)$$

$$= \begin{vmatrix} e_1 & e_2 & e_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \begin{vmatrix} e_1 & a_1 & b_1 \\ e_2 & a_2 & b_2 \\ e_3 & a_3 & b_3 \end{vmatrix}. \quad (58)$$

Hierbei ist der Vektor e ein auf Länge Eins normierter Vektor ($|e| = 1$), der mit den Vektoren a, b ein Rechtssystem (a, b, e) bildet. Die senkrechten Striche an dem quadratischen Zahlen-Tableau (der Matrix) bedeuten, dass die Determinante der Matrix zu berechnen ist (siehe Skript).

Das Kreuzprodukt ist eine nützliche Sache im \mathbb{R}^3 und kann nicht in natürlicher einfacher nutzbringender Weise auf den allgemeinen \mathbb{R}^n verallgemeinert werden. Das Kreuzprodukt sollte man nicht mit dem kartesischen Produkt verwechseln, beide haben das selbe Multiplikationszeichen, aber das Kreuzprodukt ordnet zwei Vektoren des \mathbb{R}^3 einen Vektor des \mathbb{R}^3 zu, wohingegen das kartesische Produkt zwei Mengen eine neue Menge zuordnet.

Auf dem Kreuzprodukt basiert das daher ebenfalls nur für den \mathbb{R}^3 definierte Spatprodukt,

$$\langle a \times b, c \rangle. \quad (59)$$

Das Spatprodukt bildet drei Vektoren des \mathbb{R}^3 auf eine reelle Zahl ab, deren Betrag die Größe des Spates (Parallelepipeds), welcher von diesen drei Vektoren erzeugt wird (Bild), darstellt, und deren Vorzeichen angibt, ob das System (a, b, c) der Vektoren ein Rechtssystem (Vorzeichen $+$) oder ein Linkssystem (Vorzeichen $-$) ist.

Zusammengefasst haben wir somit die folgenden „Produkte“ kennengelernt:

$$\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad (\text{Skalarprodukt})$$

$$\langle \cdot, \cdot \rangle : (a, b) \mapsto z = \langle a, b \rangle$$

$$= \sum_{i=1}^n a_i b_i,$$

$$\times : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3, \quad (\text{Kreuzprodukt})$$

$$\times : (a, b) \mapsto z = a \times b$$

$$= (a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1)^T,$$

$$\langle \cdot \times \cdot, \cdot \rangle : \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}, \quad (\text{Spatprodukt})$$

$$\langle \cdot \times \cdot, \cdot \rangle : (a, b, c) \mapsto z = \langle a \times b, c \rangle$$

$$= (a_2 b_3 - a_3 b_2) c_1 + (a_3 b_1 - a_1 b_3) c_2 + (a_1 b_2 - a_2 b_1) c_3$$

$$=: \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \pm \text{Vol}(a, b, c).$$

Genaugenommen haben wir vorher auch noch das „komplexe Produkt“ $\cdot_{\mathbb{C}}$ auf dem \mathbb{R}^2 kennengelernt:

$$\cdot_{\mathbb{C}} : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^2, \quad (\text{komplexes Produkt})$$

$$\cdot_{\mathbb{C}} : (a, b) \mapsto z = a \cdot_{\mathbb{C}} b$$

$$= (a_1 b_1 - a_2 b_2, a_1 b_2 + a_2 b_1)^T.$$

In dieser Darstellung verwenden wir die auf Gauß und Argand zurückgehende Identifikation des \mathbb{C} mit \mathbb{R}^2 , wobei die erste (horizontale) Achse dem Realteil und die zweite (vertikale) Achse dem Imaginärteil der komplexen Zahl entspricht, welche als Punkt in der Ebene dargestellt wird.

Normalerweise versteht man unter einer Addition oder Multiplikation (und damit unter einem Produkt) etwas, was die folgenden vier Rechenregeln erfüllt. Da „Dinge“ (ergo, *algebraische Strukturen*) mit diesen Eigenschaften häufig auftauchen, hier die Definition einer sogenannten (abelschen) Gruppe.

DEFINITION (ABELSCHE GRUPPE). Eine *abelsche Gruppe* ist eine Menge G mit einer Verknüpfung

$$\circ : \begin{cases} G \times G & \rightarrow G \\ (g, h) & \mapsto g \circ h = w \end{cases} \quad (60)$$

so dass

$$\begin{aligned} (g \circ h) \circ w &= g \circ (h \circ w) \quad \forall g, h, w \in G, & \text{(Assoziativität)} \\ \exists e \in G : h \circ e &= e \circ h = h \quad \forall h \in G, & \text{(neutrales Element)} \\ \forall h \in G \exists w \in G : h \circ w &= e, \quad (-h := w \quad \text{oder} \quad h^{-1} := w), & \text{(Inverse)} \\ g \circ h &= h \circ g. & \text{(Kommutativität)} \end{aligned}$$

Man spricht dann von der abelschen Gruppe (G, \circ) . □

Wenn das letzte Axiom, das der Kommutativität, *nicht* erfüllt ist, spricht man auch einfach von einer Gruppe. In abelschen, also kommutativen Gruppen wird die Verknüpfung meist als Addition mit $\circ = +$ gekennzeichnet und das neutrale Element als Nullelement $e = 0$ bezeichnet. Ist die Gruppe nicht notwendig abelsch, so wird häufig die Verknüpfung als Multiplikation mit $\circ = \cdot$ und das neutrale Element als Einselement $e = 1$ bezeichnet. Das inverse Element wird dann wie oben in der Definition angegeben additiv oder multiplikativ geschrieben. Das ist der Punkt, wo die „Operationen“ Subtraktion und Division ihren Ursprung haben.

Bekannte Beispiele abelscher Gruppen sind $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{R}, +)$, $(\mathbb{C}, +)$, (\mathbb{Q}^*, \cdot) , (\mathbb{R}^*, \cdot) , (\mathbb{C}^*, \cdot) , $(\mathbb{R}^2, +)$, $(\mathbb{R}^3, +)$ und zu guter Letzt $(\mathbb{R}^n, +)$, $n \in \mathbb{N}$. Man sieht also, ziemlich viele bekannte Dinge sind abelsche Gruppen. Im zweiten Semester werden wir uns dann auch mit nicht abelschen Gruppen herumschlagen, also Gruppen, in denen $a \cdot b \neq b \cdot a$ vorkommt. Ausserdem gibt es ja oft in einer algebraischen Struktur zwei Verknüpfungen, eine Addition und eine Multiplikation, und eventuell noch mehrere andere. Diese axiomatisch zu beschreiben ist ein Anliegen der (reinen) Mathematik und dient dem angewandten Mathematiker und dem Ingenieur als (reines) Hilfsmittel.

2.2 Geometrie im \mathbb{R}^n , $n = 2, 3$

Der Flächeninhalt eines Parallelogrammes lässt sich leicht berechnen mittels $F = |u \times v|$. Das liegt an der Definition des Kreuzproduktes zusammen mit der Herleitung einer Formel für den Flächeninhalt eines Parallelogrammes über den Satz von Cavalieri. Der Flächeninhalt berechnet sich aus Breite mal Höhe. Man nehme eine Seite (oBdA² die durch u erzeugte) als Grundseite, die Breite ist die Länge des Vektors (also $|u|$), und berechne die Höhe über den eingeschlossenen Winkel mittels des Sinus (Länge der Hypotenuse mal Sinus des Winkels), also z. B. $h = |v| \cdot |\sin(\angle(u, v))|$. Insgesamt gilt demnach $F = |u| \cdot |v| \cdot |\sin(\angle(u, v))|$, also erhält man dieselbe Zahl aus dem Absolutbetrag des Kreuzproduktes. Will man das Kreuzprodukt auch als Mittel zur Berechnung im \mathbb{R}^2 heranziehen, so muss man die Vektoren des \mathbb{R}^2 (in nicht eindeutiger Weise) in den \mathbb{R}^3 einbetten. Z. B. könnte man unten Nullen einfügen und erhält so ein Vielfaches des letzten Einheitsvektors e_3 .

Das Volumen eines Spates lässt sich leicht berechnen mittels $V = |\langle u \times v, w \rangle| = |\det(u, v, w)|$. Für das Volumen V eines Spates gilt (Prinzip von Cavalieri)

$$V = F \cdot h, \quad (61)$$

²Ohne Beschränkung der Allgemeinheit.

wobei F die Größe der Grundfläche bezeichnet und h die Höhe. Die Grundfläche ist das von v und w aufgespannte Parallelogramm und hat nach vorherigen Betrachtungen die Größe $|u \times v| = |u| \cdot |v| \cdot |\sin(\angle(u, v))|$. Um die Höhe zu erhalten, muss der dritte Vektor auf eine senkrecht zur Grundfläche stehenden Vektor projiziert werden. Dazu bietet sich der aus dem Kreuzprodukt erhaltene Vektor $u \times v$ an, der ja gerade senkrecht auf diesen beiden Vektoren, also auch auf der aufgespannten Ebene und damit auch der Grundfläche steht. Die Länge der Projektion, also die Höhe, berechnet sich zu

$$h = \left| \left\langle \frac{u \times v}{|u \times v|}, w \right\rangle \right|. \quad (62)$$

Für das Volumen gilt also

$$V = F \cdot h = |u \times v| \cdot \left| \left\langle \frac{u \times v}{|u \times v|}, w \right\rangle \right| = |\langle u \times v, w \rangle|. \quad (63)$$

Wer sich jetzt fragt, wozu man die orthogonale Projektion benötigt, der solle sich den folgenden Teil über Geraden und Ebenen bereits einmal anschauen. Er enthält geometrische Normalformen von Geraden und Ebenen (und Punkten) und die Berechnung von Abständen zwischen ihnen.

Berechnet werden können bei geeigneter Darstellung der Geraden, Ebenen usw. recht einfach die Abstände zwischen Punkten, Geraden und Ebenen und die Schnitte von Ebenen.

Eine Anmerkung zu den folgenden Beschreibungen von Geraden und Ebenen: Wir lassen oft die „äußeren“ Bedingungen an x weg, statt z. B. für die durch zwei Vektoren $a, r \in \mathbb{R}^2$ definierte Gerade g zu schreiben

$$g := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid \begin{pmatrix} x \\ y \end{pmatrix} = a + \lambda r, \quad \lambda \in \mathbb{R} \right\}, \quad (64)$$

schreiben wir der Einfachheit halber nur

$$\begin{pmatrix} x \\ y \end{pmatrix} = a + \lambda r, \quad \lambda \in \mathbb{R}. \quad (65)$$

Die Geradenbeschreibung erfolgt mittels Richtungs- und Normalenvektoren, bei Ebenen dito. Wir beginnen mit Geraden im \mathbb{R}^2 . Eine Gerade ist (in der Schulversion) beispielsweise gegeben durch die Menge aller Punkte $(x, y)^T$ mit

$$2x + 3y = 5. \quad (66)$$

Meist verwendet man in der Schule die explizite Form mit „Steigung“ und „y-Achsenabschnitt“,

$$y = -\frac{2}{3}x + \frac{5}{3}. \quad (67)$$

Diese Form schließt in unnatürlicher Weise (z. B.) die Gerade

$$1x + 0y = 5 \quad (68)$$

aus, welche eine horizontale Linie bei $x = 5$ darstellt. Wir verwenden nur die Form einer Gleichung, wobei die Unbekannten auf der einen, die Konstanten auf der anderen Seite stehen.

Mathematisch gesehen ist die Gleichungsbedingung eine Bedingung an die Punkte des \mathbb{R}^2 , die alle bis auf die Punkte auf einer Geraden, also einer zum $\mathbb{R}^1 = \mathbb{R}$ äquivalenten Menge, aussiebt. Die „Dimension“ dieser Menge ist im Allgemeinen später gegeben als die Dimension des Raumes, also n für den \mathbb{R}^n , minus die Anzahl der verschiedenen Bedingungen, also hier $2 - 1 = 1$.

Eine andere Formulierung ist gegeben durch einen Punkt auf der Geraden, im Beispiel liest man schnell den Punkt mit dem Ortsvektor $a = (1, 1)^T$ als zur Geraden gehörig ab, da

$$2 \cdot 1 + 3 \cdot 1 = 5, \quad (69)$$

und einen Vektor, der die Richtung der Geraden angibt, zu berechnen über zwei verschiedene Punkte auf der Geraden, als Beispiel neben dem Punkt A mit Ortsvektor $a = (1, 1)^T$ der Punkt B mit Ortsvektor

$$b = \begin{pmatrix} 4 \\ -1 \end{pmatrix}, \quad \text{oder} \quad b = \begin{pmatrix} 0 \\ \frac{5}{3} \end{pmatrix}, \quad \text{oder} \quad b = \begin{pmatrix} \frac{5}{2} \\ 0 \end{pmatrix}. \quad (70)$$

Die Differenz zweier Ortsvektoren ist ein Vektor in Richtung der Geraden, wir wählen das erste b , damit gilt

$$b - a = \begin{pmatrix} 4 \\ -1 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}. \quad (71)$$

Nun lässt sich jeder Punkt auf dieser Geraden erreichen durch den Ortsvektor, der entsteht, wenn wir erst auf die Gerade gehen und dann entlang der Richtung der Geraden, also mittels

$$x = a + \lambda(b - a) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} 3 \\ -2 \end{pmatrix}, \quad \lambda \in \mathbb{R}. \quad (72)$$

Der Vektor

$$r := \begin{pmatrix} 3 \\ -2 \end{pmatrix} \quad (73)$$

und seine Nichtnull-Vielfachen sind sogenannte Richtungsvektoren der Geraden. Alternativ kann man eine Richtung im \mathbb{R}^2 durch Angabe eines senkrecht dazu stehenden Vektors beschreiben, eines Normalenvektors. Laut vorherigem ist ein solcher Vektor (festgelegt nur bis auf Nichtnull-Vielfache) der Vektor

$$n := \begin{pmatrix} 2 \\ 3 \end{pmatrix}. \quad (74)$$

Dieser Normalenvektor lässt sich aber ohne weiteres direkt an der Gleichung (66) ablesen, denn es gilt mit dem eben definierten n offensichtlich

$$\left\langle n, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle = 2x + 3y = 5. \quad (75)$$

Der Beweis, dass der abgelesene Vektor n der Normalenvektor ist, erfolgt über die Verwendung der Ortsvektoren a, b zweier beliebiger verschiedener Punkte auf der Geraden und der Geradengleichung, da

$$\langle n, b \rangle = 5, \quad (76a)$$

$$\langle n, a \rangle = 5, \quad (76b)$$

$$\Rightarrow \langle n, b - a \rangle = \langle n, r \rangle = 0. \quad (76c)$$

Der Vektor $r := b - a$ ist von Null verschieden und ein Richtungsvektor, also steht der abgelesene Vektor n auf dem (jedem) so berechneten Richtungsvektor senkrecht, ist also ein Normalenvektor.

Die Form mit dem Skalarprodukt und dem normierten Normalenvektor

$$n^0 = \frac{n}{|n|} = \frac{1}{\sqrt{13}} \begin{pmatrix} 2 \\ 3 \end{pmatrix} \quad (77)$$

ist die sogenannte Hessesche Normalform, in LA meist geschrieben als

$$\langle n^0, x - a \rangle = 0, \quad (78)$$

in der Literatur meist in der Form

$$\langle n^0, x \rangle = d \quad (= \langle n^0, a \rangle). \quad (79)$$

Der Abstand zwischen einem Punkt und einer Geraden kann am Besten mittels der Hesseschen Normalform ermittelt werden. Sei dazu z. B. P ein Punkt mit dem Ortsvektor $p = (1, 2)^T$.

Dann ist nach dem Satz von Pythagoras der Abstand gegeben durch die Länge der Projektion von $p - a$ auf die Richtung von n , also durch

$$\langle n^0, p - a \rangle = \frac{1}{\sqrt{13}} \left\langle \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 1-1 \\ 2-1 \end{pmatrix} \right\rangle = \frac{3}{\sqrt{13}}. \quad (80)$$

Damit ist das oben gegebene $d = \langle n^0, a \rangle$ der Abstand der Geraden zum Nullpunkt.

Winkel zwischen Geraden im \mathbb{R}^2 können mittels der Richtungs- oder Normalenvektoren bestimmt werden,

$$\langle v, w \rangle = |v| \cdot |w| \cdot \cos(\angle(v, w)), \quad (81a)$$

$$\cos(\angle(v, w)) = \frac{\langle v, w \rangle}{|v| \cdot |w|}, \quad (81b)$$

$$\angle(v, w) = \arccos \left(\frac{\langle v, w \rangle}{|v| \cdot |w|} \right). \quad (81c)$$

Diese Gleichung gilt auch für sich schneidende Geraden im \mathbb{R}^n .

Jetzt die Ebenen im \mathbb{R}^3 . Diese lassen sich über die Angabe einer Gleichung

$$1x_1 + 2x_2 + 3x_3 = 6 \quad (82)$$

beschreiben. Hier gilt wieder: die Dimensionalität (Ebene, 2) des Objektes ist die Dimension des Raumes (3) minus die Anzahl der verschiedenen Bedingungen (hier 1) an die drei Punkte (x_1, x_2, x_3) , also an die Vektoren des \mathbb{R}^3 . Keine Bedingung ist z. B.

$$0x_1 + 0x_2 + 0x_3 = 0, \quad (83)$$

da alle Vektoren des \mathbb{R}^3 diese Gleichung erfüllen, und auch

$$0x_1 + 0x_2 + 0x_3 = 1 \quad (84)$$

ist keine Bedingung, da diese Gleichung nicht erfüllbar ist. Der Grund ist die „Häufung von Null“ auf der linken Seite.

Man sieht schnell, dass z. B. der Ortsvektor $a = (1, 1, 1)^T$ auf die Ebene führt, denn

$$1 \cdot 1 + 2 \cdot 1 + 3 \cdot 1 = 6. \quad (85)$$

Jetzt kann man zwei weitere Punkte auf der Ebene identifizieren (da ja bekanntlich drei Punkte eine Ebene eindeutig beschreiben, nicht umsonst haben Stühle mindestens drei und nicht zwei Beine), wir wählen

$$b = \begin{pmatrix} 0 \\ 3 \\ 0 \end{pmatrix} \quad \text{und} \quad c = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}. \quad (86)$$

Eine Punkt-Richtungsform erhält man nun durch Angabe eines Ortsvektors, z. B. a und zweier in der Ebene liegender Richtungen, wir wählen

$$r_1 := b - a = \begin{pmatrix} 0-1 \\ 3-1 \\ 0-1 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} \quad (87)$$

und

$$r_2 := c - a = \begin{pmatrix} 0-1 \\ 0-1 \\ 2-1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}. \quad (88)$$

Damit gilt dann

$$x = a + \lambda r_1 + \mu r_2, \quad \lambda, \mu \in \mathbb{R}, \quad (89)$$

oder in ausgeschriebener Form

$$x = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \lambda \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} + \mu \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}, \quad \lambda, \mu \in \mathbb{R}. \quad (90)$$

Alternativ kann man wieder einen senkrecht zur Ebene stehenden Vektor n verwenden, um die Ausrichtung der Ebene im Raum eindeutig zu charakterisieren. Für diese Zwecke bietet sich das Kreuzprodukt an,

$$n := r_1 \times r_2 = \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix} \times \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 - 1 \\ 1 - -1 \\ 1 - -2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}. \quad (91)$$

Erstaunlicherweise haben wir wieder den „ablesbaren“ Vektor berechnet ...

Allgemein ist es auch für Ebenen so, dass der ablesbare Vektor ein Normalenvektor ist, der Beweis ist analog zu dem für die Geraden. Es gilt mit dem eben berechneten/abgelesenen n ja offensichtlich für alle Punkte der Ebene mit Ortsvektor x

$$\langle n, x \rangle = 6. \quad (92)$$

Nun kann man ja drei beliebige (verschiedene) Punkte mit Ortsvektoren $a, b, c \in \mathbb{R}^3$ nehmen, welche die Geradengleichungen erfüllen und mittels Subtraktion Vektoren generieren, welche in der Ebene liegen und somit Richtungsvektoren sind,

$$\langle n, a \rangle = 6, \quad (93a)$$

$$\langle n, b \rangle = 6, \quad (93b)$$

$$\langle n, c \rangle = 6, \quad (93c)$$

$$\Rightarrow \langle n, r_1 \rangle = \langle n, b - a \rangle = 0, \quad (93d)$$

$$\Rightarrow \langle n, r_2 \rangle = \langle n, c - a \rangle = 0. \quad (93e)$$

Also steht der abgelesene Vektor n auf allen möglichen Richtungsvektoren senkrecht, ist also bereits ein Normalenvektor. Meist nimmt man die normalisierte Variante des Einheitsnormalenvektors

$$n^0 := \frac{n}{|n|} = \frac{1}{\sqrt{14}} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad (94)$$

und bastelt sich die Hessesche Normalform gemäß

$$\langle n^0, x - a \rangle = 0. \quad (95)$$

Der Abstand eines Punktes mit Ortsvektor p , z. B.

$$p = \begin{pmatrix} 0 \\ -1 \\ -2 \end{pmatrix} \quad (= a - n), \quad (96)$$

ist wieder durch den Betrag von

$$\langle n^0, p - a \rangle = \frac{1}{\sqrt{14}} \left\langle \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ -2 \\ -3 \end{pmatrix} \right\rangle = \frac{-14}{\sqrt{14}} = -\sqrt{14} = -|n|, \quad (97)$$

also wie erwartet, durch $|n|$ gegeben.

Geraden im \mathbb{R}^3 kann man gut mittels einer Punkt-Richtungs-Form darstellen,

$$g := \{x = a + \lambda r\}, \quad a, r \in \mathbb{R}^3. \quad (98)$$

Eine Aufgabe aus einer alten Klausur: Es geht um die Bewertung der Aussage, dass der Abstand der beiden Geraden

$$g_1 := \left\{ x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \lambda \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \lambda \in \mathbb{R} \right\} \quad (99)$$

und

$$g_2 := \left\{ x = \begin{pmatrix} 2 \\ 4 \\ 2 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \mu \in \mathbb{R} \right\} \quad (100)$$

größer als 3 ist. Hier kann man zwar die Punkte minimalen Abstandes berechnen, und danach die Länge der Verbindungsstrecke, aber man sieht sofort, dass die Gerade g_1 durch alle Vektoren x_1 der Gestalt

$$x_1 = \begin{pmatrix} 1 \\ 2 \\ \star_1 \end{pmatrix}, \star_1 \in \mathbb{R} \quad (101)$$

und die Gerade g_2 durch alle Vektoren x_2 der Gestalt

$$x_2 = \begin{pmatrix} \star_2 \\ 4 \\ 2 \end{pmatrix}, \star_2 \in \mathbb{R} \quad (102)$$

beschrieben wird. Die Differenz zweier solcher Vektoren ist gegeben durch

$$x_2 - x_1 = \begin{pmatrix} \star_2 - 1 \\ 2 \\ 2 - \star_1 \end{pmatrix}. \quad (103)$$

Um dessen Länge zu minimieren, kann man in der Definition der Länge,

$$|x_2 - x_1|^2 = (\star_2 - 1)^2 + 2^2 + (2 - \star_1)^2 \quad (104)$$

die beiden Größen unabhängig als $\star_2 = 1$ und $\star_1 = 2$ wählen und erhält somit als Differenzvektor und minimale Länge

$$\begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}, \quad \left| \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} \right| = 2 < 3. \quad (105)$$

Allgemein? Im Allgemeinen stellt man schnell fest, dass der Abstand minimal wird, wenn die Verbindungsgerade zwischen den Punkten mit behaupteten minimalem Abstand senkrecht steht auf den Richtungsvektoren der Geraden. Der Beweis erfolgt über den Satz des Pythagoras in der Ebene aufgespannt durch eine der Geraden und die Verbindungsgerade.

Also bildet man für zwei Punkte (i.e., durch feste Wahl von $\hat{\lambda}$ und $\hat{\mu}$) die inneren Produkte mit den beiden Richtungsvektoren und erhält ein Gleichungssystem mit zwei Unbekannten und zwei Gleichungen,

$$\langle r_1, a_1 - a_2 + \hat{\lambda}r_1 - \hat{\mu}r_2 \rangle = 0, \quad (106a)$$

$$\langle r_2, a_1 - a_2 + \hat{\lambda}r_1 - \hat{\mu}r_2 \rangle = 0. \quad (106b)$$

Zurück zu dem Beispiel. Im Beispiel ergibt sich für $\hat{\lambda}$ und $\hat{\mu}$ das Gleichungssystem

$$\left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 2 \\ 4 \\ 2 \end{pmatrix} + \hat{\lambda} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \hat{\mu} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle = 0, \quad (107a)$$

$$\left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 2 \\ 4 \\ 2 \end{pmatrix} + \hat{\lambda} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \hat{\mu} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle = 0, \quad (107b)$$

welches ausgeschrieben lautet

$$1 + \hat{\lambda} = 0, \quad (108a)$$

$$-1 - \hat{\mu} = 0. \quad (108b)$$

Die Lösung ist $\hat{\lambda} = -1$ und $\hat{\mu} = -1$. Die zugehörigen Punkte $\{x_i\}_{i=1}^2$ auf den jeweiligen Geraden $\{g_i\}_{i=1}^2$ haben die Ortsvektoren

$$x_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \hat{\lambda} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \quad (109)$$

und

$$x_2 = \begin{pmatrix} 2 \\ 4 \\ 2 \end{pmatrix} + \hat{\mu} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} \quad (110)$$

und (wie nicht anders zu erwarten war) den Differenzvektor

$$d = x_2 - x_1 = \begin{pmatrix} 1 \\ 4 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}. \quad (111)$$

Wieder sieht man, dass der Abstand gleich Zwei, damit kleiner als Drei und die Aussage eine falsche ist, ein.

2.3 Auswirkungen der Orthogonalität

Die Orthogonalität kann herangezogen werden, um Fälle der Gleichheit in der Dreiecksungleichung und der Cauchy-Schwarzschen Ungleichung (CSU)

$$|\langle v, w \rangle| \leq |v| \cdot |w| \quad \forall v, w \in \mathbb{R}^n \quad (112)$$

(ohne Absolutbeträge) zu charakterisieren, welche für den \mathbb{R}^2 und \mathbb{R}^3 schnell aus

$$\langle v, w \rangle = |v| \cdot |w| \cdot \cos(\angle(v, w)) \quad \text{und} \quad -1 \leq \cos(x) \leq 1, \quad x \in \mathbb{R} \quad (113)$$

folgt. Hier ist sehr schön die einfache geometrische Deutung der Länge („Norm“ im \mathbb{R}^n) und der Winkel (Skalarprodukt) zu erkennen. Zum besseren Verständniss der Orthogonalität gibt es in Aufgabe 4 vier Teilaufgaben mit Vektoren aus dem \mathbb{R}^3 , daher steht hier jetzt keine weitere Herleitung dieser Sachverhalte.

3 Anleitung vom 26.11.2010

Themen dieser Anleitung: Allgemeine Vektorräume; Unter(vektor)räume; Linearkombinationen, lineare Ab- und Unabhängigkeit; Basen, Dimension.

3.1 Vektorraum

In den meisten Mathematik-Büchern findet sich die folgende Definition eines allgemeinen Vektorraumes³: Eine nichtleere Menge V heißt (reeller) Vektorraum, wenn es eine Verknüpfung

$$+ : \begin{cases} V \times V & \rightarrow V \\ (u, v) & \mapsto u + v = w \end{cases} \quad (114a)$$

gibt, so dass $(V, +)$ eine sogenannte abelsche Gruppe ist, d.h., es gelten

$$\begin{aligned} (u + v) + w &= u + (v + w) \quad \forall u, v, w \in V, & (\text{Assoziativität}) \\ \exists 0 \in V : v + 0 &= 0 + v = v \quad \forall v \in V, & (\text{Nullelement}) \\ \forall v \in V \exists w \in V : v + w &= 0, \quad (-v := w), & (\text{Inverse}) \\ u + v &= v + u, & (\text{Kommutativität}) \end{aligned}$$

und eine Verknüpfung

$$\cdot : \begin{cases} \mathbb{R} \times V & \rightarrow V, \\ (\lambda, v) & \mapsto \lambda \cdot v = w, \end{cases} \quad (116a)$$

so dass (mit $\lambda, \mu \in \mathbb{R}$ und $u, v, w \in V$) die folgenden Verträglichkeiten sichergestellt sind,

$$\begin{aligned} (\lambda \cdot \mu) \cdot v &= \lambda \cdot (\mu \cdot v), & (\text{Assoziativität}) \\ (\lambda + \mu) \cdot v &= \lambda \cdot v + \mu \cdot v, & (\text{Distributivitätsgesetze}) \\ \lambda \cdot (v + w) &= \lambda \cdot v + \lambda \cdot w, & (\text{Distributivitätsgesetze}) \\ 1 \cdot v &= v. & (\text{Neutralität der Eins}) \end{aligned}$$

Die Assoziativität bewirkt, dass es für das Ergebnis egal ist, in welcher Reihenfolge gerechnet wird, ergo: welche Klammern man weglassen kann. Die Existenz des Nullelementes und der Inversen bewirken, dass man in diesem abstrakten „allgemeinen Vektorraum“ auch so rechnen kann wie bisher, man kann also jeden Weg z. B. wieder umkehren.

In der Vorlesung wird statt den beiden Axiomen der Existenz der Null und der Existenz von Inversen äquivalent die einzelne Forderung

$$\forall u, v \in V \exists! x \in V : u + x = v \quad (\text{Eindeutige Auflösbarkeit})$$

verwendet. Die Symbolfolge $\exists!$ bedeutet dabei „es existiert genau ein“. Dieses Axiom beschreibt die „eindeutige Auflösbarkeit“ von „Gleichungen“ in einem Vektorraum, anschaulicher gesprochen: Wo man vom Ursprung hin kommt, kommt man auch von jedem anderen Ort hin. Es gibt also immer eine Möglichkeit der Umkehr.

Da in den 7 Axiomen aus der Vorlesung nicht mehr klar ersichtlich ist, dass es wirklich eine Null und ein Inverses gibt, ist eine Aufgabe die Herleitung der Existenz und Eindeutigkeit eines Nullelementes aus den gegebenen sieben Axiomen.

Die skalare Multiplikation sichert insbesondere mit der Neutralität der Eins die Interpretation des Aneinanderlegens, Stauchens und Streckens von Vektoren.

Motiviert und abgeleitet sind diese einen Vektorraum definierenden Eigenschaften aus den „Modellen“ des \mathbb{R}^2 und des \mathbb{R}^3 . Der Schritt zum \mathbb{R}^n ist vom mathematischen her ein kleiner, statt nur zwei oder drei Einträge zuzulassen, lässt man jetzt einfach allgemein $n = 1, 2, 3, 4, 5, \dots$ Einträge zu.

³Es ist klar, dass man, will man möglichst viele (der in den Ingenieurwissenschaften auftretenden) Strukturen mit gemeinsamen Techniken behandeln, eben mit „linearer Algebra“, dass man nicht zu viele Forderungen stellen sollte, die erfüllt sein müssen. Dass es hier dann 8 oder 7 Axiome sind, die einen Vektorraum definieren, und nicht etwa deutlich weniger, ergibt sich automatisch, wenn man untersucht, was geschieht, wenn man Axiome wegzulassen versucht. Eine schöne Liste resultierender Gebilde findet sich (in englischer Sprache) unter http://www.math.niu.edu/~rusin/known-math/99/lin_alg. Leider ist dort aber auch fehlerhaftes: Statt „no-3: Can't happen!“ sollte dort „no-3: Set V to the empty set.“ stehen. Warum?

Die wichtigsten Änderungen scheinen auf den ersten Blick der Verlust jeglicher Interpretation in der Anschauung (Ebene, Anschauungsraum) und des Kreuzproduktes (und damit des Spatproduktes) zu sein.

Der Verlust der Anschauung ist gemindert, falls man nur mit einigen Vektoren zu tun hat. Zwei Vektoren lassen sich (fast) immer als eine Ebene aufspannend auffassen, drei Vektoren (fast) immer als einen Anschauungsraum aufspannend. Oft hilft diese Vorstellung.

Rechnerisch ist alles immer noch sehr einfach, man benötigt ja nur die komponentenweise Addition und Multiplikation,

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad (118a)$$

$$\lambda \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda \cdot x_1 \\ \lambda \cdot x_2 \\ \vdots \\ \lambda \cdot x_n \end{pmatrix}. \quad (118b)$$

Diese „endlich langen“ Vektoren kann man sogar zu (überabzählbar) „unendlich langen“ Vektoren machen. Damit erhält man die Funktionenräume (über \mathbb{R}), welche erfahrungsgemäß für Erstsemester eine große Hürde darstellen. In diesen wird die Addition und die skalare Multiplikation wieder „komponentenweise“ definiert. Dieses erfolgt nicht wie bei Vektoren, welche ja auch nur Funktionen

$$x : \{1, 2, \dots, n\} \rightarrow \mathbb{R}$$

sind, wobei die Funktionen eindeutig durch Angabe aller ihrer (endlich vielen) Werte charakterisiert werden und die Funktionswerte meist als Komponenten geschrieben werden,

$$x(i) = x_i, \quad i \in \{1, 2, \dots, n\}.$$

Im Gegensatz dazu wird einfach die folgende Notation verwendet. Sei \mathcal{F} die Menge aller Funktionen $f : \mathbb{R} \rightarrow \mathbb{R}$, dann sei die Summe zweier Funktionen und die skalare Multiplikation einer Funktion „komponentenweise“ definiert durch

$$(\mathbf{f} + \mathbf{g})(x) := f(x) + g(x), \quad f, g \in \mathcal{F}, \quad \forall x \in \mathbb{R}, \quad (119a)$$

$$(\lambda \cdot \mathbf{f})(x) := \lambda \cdot f(x), \quad f \in \mathcal{F}, \quad \forall x \in \mathbb{R}. \quad (119b)$$

Mann kann nachrechnen (und sollte das vielleicht auch mal tun), dass dieser „Funktionenraum“ ein Vektorraum ist.

3.2 Unterraum

Eine Teilmenge $W \subset V$ eines Vektorraumes V , die selber wieder ein Vektorraum ist, ist ein sogenannter Untervektorraum oder kurz Unterraum. Nun möchte man selten nachrechnen, ob auch alle Axiome für einen Vektorraum erfüllt sind. Glücklicherweise muss man das auch nicht, da man nur testen muss, ob die Auflösbarkeit von Gleichungen gegeben ist, oder nach Satz 2.23 der Vorlesung, ob

$$u, v \in W \Rightarrow u + v \in W, \quad (120a)$$

$$\lambda \in \mathbb{R}, \quad v \in W \Rightarrow \lambda \cdot v \in W, \quad (120b)$$

oder äquivalent dazu, ob sämtliche Linearkombinationen wieder in W liegen,

$$\lambda, \mu \in \mathbb{R}, \quad u, v \in W \Rightarrow \lambda \cdot u + \mu \cdot v \in W. \quad (121)$$

Die letzte Bedingung untersucht also, ob mit gegebenen Vektoren u, v, w, \dots alle „linearen“ Gebilde, also der Nullpunkt ($\lambda = 0 = \mu$), Geraden durch Null (z. B. $\lambda = 0$), Ebenen durch Null ($\lambda \cdot u + \mu \cdot v$), höhere Räume ($\lambda \cdot u + \mu \cdot v + \eta \cdot w$ etcpp.) in dem potentiellen Unterraum enthalten sind. Um zu testen, ob eine Untermenge eines „bekannten“ Vektorraumes ein Unterraum ist, verwendet man am besten die beiden Gleichungen (120).

3.3 Linearkombinationen; lineare Abhängigkeit

Wie eben gerade gesehen, sollten mit gewissen Vektoren v_1, \dots, v_r auch immer alle Linearkombinationen in einem Unterraum enthalten sein. Damit kann man aber einen Unterraum basteln, indem man eine Menge von Vektoren nimmt, und alle möglichen Linearkombinationen hinzunimmt. Abhängig von den gewählten Vektoren erhält man z. B. den gesamten Raum V oder auch nur einen Teilraum. Man sagt, dass die Vektoren v_1, \dots, v_r einen Raum aufspannen. Zu einem gegebenen (endlichdimensionalen) Raum gibt es auch immer ein Erzeugendensystem, d.h., eine Menge von Vektoren, die diesen Raum aufspannen. Verschiedene Mengen von Vektoren können ein und denselben Raum aufspannen, z. B. spannen sowohl

$$v_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad (122)$$

als auch

$$w_1 := \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad w_2 := \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \quad w_3 := \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix}, \quad w_4 := \begin{pmatrix} 4 \\ 1 \\ 0 \end{pmatrix}, \quad w_5 := \begin{pmatrix} 5 \\ 1 \\ 0 \end{pmatrix} \quad (123)$$

denselben Raum, nämlich den Unterraum des \mathbb{R}^3 gegeben durch alle Vektoren, welche eine Null in der letzten Komponente haben, alternativ, welche durch die Ebene $\langle e_3, x \rangle = 0$ definiert sind, auf.

Um etwas Ordnung einzubringen, bietet es sich an, minimale Erzeugendensysteme zu suchen. Dieses soll zuerst am Beispiel geschehen.

Man kann sich schnell überzeugen, dass es kein Erzeugendensystem aus nur einem Vektor bestehend geben kann, welches diese durch Null gehende Ebene aufspannt, da es ja (z. B. berechnet durch das Kreuzprodukt) immer noch einen weiteren auf dem gewählten Vektor und dem dritten Einheitsvektor e_3 senkrecht stehenden Vektor gibt. Nun ist aber die Auswahl der Vektoren anscheinend nicht eindeutig, denn mit dem Erzeugendensystem (122) aus zwei Vektoren ist auch die folgende Auswahl von zwei Vektoren ein Erzeugendensystem,

$$u_1 := \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad u_2 := \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}. \quad (124)$$

Invariant scheint nur die Anzahl der benötigten Vektoren zu sein. Dass dieses so ist, folgt mit einigen weiteren Definitionen (linear (un)abhängig, Basis) und Erkenntnissen aus dem Austauschsatz von Steinitz (siehe Skript).

Um diese Anzahl zu berechnen, schaut man, ob es möglich ist, einen der gegebenen Vektoren durch die anderen darzustellen, und damit herauszunehmen. Diese Darstellung ist immer genau dann möglich, wenn

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_r v_r = 0 \quad (125)$$

gilt, wobei mindestens ein (und damit im Fall $v_j \neq 0$, $j = 1, \dots, r$, mindestens zwei) α_j , $j = 1, \dots, r$, ungleich Null ist (sind). Sei oBdA (Umnummerierung) $j = r$. Dann gilt

$$v_r = -\frac{\alpha_1}{\alpha_r} v_1 - \dots - \frac{\alpha_{r-1}}{\alpha_r} v_{r-1}, \quad (126)$$

und v_r ist bereits in dem von v_1, \dots, v_{r-1} erzeugten Raum enthalten.

Man kann also das Erzeugendensystem immer um einen Vektor reduzieren, wenn es α_j , nicht alle gleich Null gibt, so dass die *Linearkombination*

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_r v_r = 0 \quad (127)$$

der Vektoren v_1, \dots, v_r das Nullelement ergibt. Das führt auf den Begriff der „linearen Abhängigkeit“. Umgekehrt definiert man die lineare *Unabhängigkeit* der Vektoren v_1, \dots, v_r über

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_r v_r = 0 \Rightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_r = 0. \quad (128)$$

Beispiele: Beliebige Vektoren aus dem \mathbb{R}^n für $n = 1, 2, 3, \dots$ basteln, gucken. Ist der Nullvektor linear abhängig? Die Funktionen

$$f_1 : \begin{cases} \mathbb{R} & \rightarrow \mathbb{R} \\ x & \mapsto x^3 \end{cases} \quad (129a)$$

und

$$f_2 : \begin{cases} \mathbb{R} & \rightarrow \mathbb{R} \\ x & \mapsto \cos(x) \end{cases} \quad (129b)$$

sind linear unabhängig. Dazu reicht es aus, die lineare Unabhängigkeit in endlich vielen Punkten nachzuweisen, wir wählen die Punkte $x_1 = 0$ und $x_2 = \pi$, damit gilt $0^3 = 0$, $\cos(0) = 1$, $\pi^3 = \pi^3$, $\cos(\pi) = -1$. Statt nachzuweisen, dass

$$\alpha_1 f_1 + \alpha_2 f_2 = 0 \Leftrightarrow (\alpha_1 f_1 + \alpha_2 f_2)(x) = 0 \quad \forall x \in \mathbb{R} \quad (130)$$

$$\Leftrightarrow \alpha_1 x^3 + \alpha_2 \cos(x) = 0 \quad \forall x \in \mathbb{R} \Rightarrow \alpha_1 = \alpha_2 = 0, \quad (131)$$

weist man nach, dass

$$\alpha_1 x^3 + \alpha_2 \cos(x) = 0 \Rightarrow \alpha_1 = \alpha_2 = 0 \quad x \in \{0, \pi\},$$

welches folgendermaßen geschieht,

$$\begin{aligned} \alpha_1 x_1^3 + \alpha_2 \cos(x_1) &= 0 \\ \alpha_1 x_2^3 + \alpha_2 \cos(x_2) &= 0 \end{aligned} \Rightarrow \alpha_1 = \alpha_2 = 0.$$

Diese Erkenntnis folgt sofort aus der Lösung des Gleichungssystems

$$\begin{aligned} \alpha_1 \cdot 0^3 + \alpha_2 \cdot (+1) &= 0, \\ \alpha_1 \cdot \pi^3 + \alpha_2 \cdot (-1) &= 0, \end{aligned}$$

denn man sieht in der ersten Zeile (der ersten Gleichung) vorne eine Null, also ist α_2 bestimmt durch 0 geteilt durch 1, also gleich Null. Damit ist durch Einsetzen nach der zweiten Zeile auch $\alpha_1 = 0$.

Wenn für gewisse Wahlen von endlich vielen Punkten die Funktionen linear abhängig sind, so muss keine lineare Abhängigkeit der Funktionen vorliegen, als Beispiel seien die Funktionen \cos und \sin ausgewählt und ausgewertet an beliebigen ganzzahligen Vielfachen von 2π . Der Kosinus ergibt also immer einen Vektor nur aus Einsen (oft in der Literatur mit dem kleinen Buchstaben e bezeichnet), der Sinus einen Vektor nur aus Nullen (oBdA mit 0 bezeichnet),

$$\begin{pmatrix} \cos(x_1) \\ \cos(x_2) \\ \vdots \\ \cos(x_n) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = e, \quad \begin{pmatrix} \sin(x_1) \\ \sin(x_2) \\ \vdots \\ \sin(x_n) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = 0, \quad x_i = 2\pi k_i, k_i \in \mathbb{Z}.$$

Trotzdem (Anschauung) ist der Sinus nicht aus dem Kosinus durch Skalierung mit der Null hervorgegangen, der Sinus ist ja nicht die Nullfunktion.

Wenn für alle Wahlen von endlich vielen Punkten die Funktionen linear abhängig sind, so gibt es eventuell so etwas wie eine Funktionalgleichung (Additionstheorem o.ä.), die man ausnutzen kann, um lineare Abhängigkeit der Funktionen zu zeigen. Ein Standardbeispiel wird konstruiert unter Verwendung der Funktionalgleichung der Exponentialfunktionen,

$$a^{x+y} = a^x \cdot a^y.$$

Wenn wir jetzt „die“ Exponentialfunktion verwenden und zwar zur Konstruktion der beiden Funktionen

$$g_1 : \begin{cases} \mathbb{R} & \rightarrow \mathbb{R} \\ x & \mapsto e^{42+x} \end{cases} \quad (132a)$$

und

$$g_2 : \begin{cases} \mathbb{R} & \rightarrow \mathbb{R} \\ x & \mapsto e^x \end{cases} \quad (132b)$$

so sieht man sofort, dass

$$e^{42+x} = e^{42} \cdot e^x =: \alpha_2 \cdot e^x$$

und mit $\alpha_1 = -1$ die beiden Funktionen linear abhängig sind,

$$\alpha_1 \cdot e^{42+x} + \alpha_2 \cdot e^x = 0 \quad \forall x \in \mathbb{R} \quad \Leftrightarrow \quad \alpha_1 \cdot g_1 + \alpha_2 \cdot g_2 = 0.$$

Weitere Beispiele für lineare Abhängigkeit von Funktionen finden sich in den kommenden Aufgaben.

Zwei miteinander verwandte Klausuraufgaben sind wie folgt gegeben. Es seien die vier Funktionen

$$\begin{aligned} m_1(x) &:= \min\{x, 0\}, & m_2(x) &:= \max\{x, 0\}, \\ m_3(x) &:= x & \text{und} & \quad m_4(x) := |x| \end{aligned}$$

definiert. Erste Aufgabe: gilt die Aussage

$$\text{span}\{m_1, m_2\} = \text{span}\{m_3, m_4\}?$$

Klarerweise gilt $m_3 = m_1 + m_2$ und auch $m_4 = m_2 - m_1$ (Skizze). Damit ist auch $2m_1 = m_3 - m_4$ und $2m_2 = m_3 + m_4$. Aus Division durch 2 folgt, dass die Aussage wahr ist.

Zweite Aufgabe: gilt die Aussage

$$\text{span}\{m_1, m_3\} = \text{span}\{m_2, m_4\}?$$

Nach eben Gezeigtem gilt $m_3 = 2m_2 - m_4$ und damit $m_1 = m_3 - m_2 = 2m_2 - m_4 - m_2 = m_2 - m_4$, was man aber auch gleich hätte sehen können. Umgekehrt gilt klarerweise $m_2 = m_3 - m_1$ und $m_4 = m_3 - 2m_1$.

3.4 Basis; Dimension

Offensichtlich ist ein minimales Erzeugendensystem aus linear unabhängigen Vektoren beschaffen, sonst wäre es nach vorherigen Betrachtungen nicht minimal. Die Anzahl dieser Vektoren ist nur vom Unterraum abhängig und heißt die Dimension des Raumes, im Beispiel (122) der Vektoren aus dem \mathbb{R}^3 ist die Dimension also 2.

Minimale Erzeugendensysteme werden kürzer als Basis bezeichnet (von dieser „Basis“ ausgehend, lässt sich ja alles in dem Raum durch Linearkombinationen eindeutig erstellen/darstellen).

Die Eindeutigkeit der Darstellung folgt (siehe Beweis in der Anleitung/Skript) über Beweis durch Widerspruch.

Welche Dimension hat der Raum

$$\text{span}\{m_1, m_2, m_3, m_4\}?$$

Auch dieser Raum hat die Dimension Zwei, da nach Vorherigem z. B. m_3 und m_4 sich mittels Linearkombinationen von m_1 und m_2 ausdrücken lassen, und die beiden Funktionen m_1, m_2 linear unabhängig sind, da die Auswertung an den Punkten ± 1 die offensichtlich linear unabhängigen Vektoren

$$M_1 := \begin{pmatrix} m_1(-1) \\ m_1(1) \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \text{und} \quad M_2 := \begin{pmatrix} m_2(-1) \\ m_2(1) \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

ergeben. Als Basen kommen nach den vorherigen Betrachtungen (und einer kurzen Überlegung, dass keine der Funktionen aus einer der anderen mittels einer Skalierung hervorgegangen sein kann) unter anderem die vier folgenden Basen in Frage:

$$(m_1, m_2), \quad (m_3, m_4), \quad (m_1, m_3), \quad (m_2, m_4).$$

Der Begriff der Basis zeigt auf, dass eigentlich alle „abstrakten“ endlichdimensionalen reellen Vektorräume (oder Betrachtungen „abstrakter“ reeller Vektorräume, die nur Untervektorräume, erzeugt mittels endlich vieler Vektoren, benötigen) sich durch den \mathbb{R}^n beschreiben lassen, wobei n die Dimension des Vektorraumes ist. Das ist der Fall, da man dann zu den Vektoren gebildet aus den Koeffizienten, die nötig sind, um die Vektoren in einer gewählten Basis darzustellen, übergehen kann.

Als Beispiel betrachte man die (sogenannten Legendre-)Polynome

$$P_0(x) := 1, \tag{133a}$$

$$P_1(x) := x, \tag{133b}$$

$$P_2(x) := \frac{1}{2}(3x^2 - 1), \tag{133c}$$

$$P_3(x) := \frac{1}{2}(5x^3 - 3x). \tag{133d}$$

Diese sind linear unabhängig und bilden eine Basis des Unterraumes Π_3 der Polynome vom Höchstgrad drei. Wir kennen bereits aus der Vorlesung die Basis der Monome,

$$\Pi_3 := \text{span}\{1, x, x^2, x^3\},$$

welche ja gerade in Form von Linearkombinationen in den *Polynomen* resultieren. Wir zeigen, dass beide Funktionssätze linear unabhängig sind, und dass beide eine Basis desselben Raumes bilden.

Um zu zeigen, dass die Monome bis zur Potenz $k = 3$ eine Basis eines Raumes der Dimension $k + 1 = 4$ bilden, wählt man am Besten $k + 1 = 4$ verschiedene Punkte, hier wählen wir einfach $x_j = j \in \{1, 2, 3, 4\}$ und erhalten durch Einsetzen aus der allgemeinen Forderung (beliebiges $x \in \mathbb{R}$ im Sinne von für alle $x \in \mathbb{R}$ gültig) der linearen Unabhängigkeit

$$\alpha_0 \cdot 1 + \alpha_1 \cdot x + \alpha_2 \cdot x^2 + \alpha_3 \cdot x^3 = 0 \quad \Rightarrow \quad \alpha_0 = \alpha_1 = \alpha_2 = \alpha_3 = 0$$

die spezielle Forderung ($j = 1, \dots, 4$)

$$\alpha_0 \cdot 1 + \alpha_1 \cdot x_j + \alpha_2 \cdot x_j^2 + \alpha_3 \cdot x_j^3 = 0 \quad \Rightarrow \quad \alpha_0 = \alpha_1 = \alpha_2 = \alpha_3 = 0,$$

welche auf ein Gleichungssystem

$$\begin{aligned} \alpha_0 \cdot 1 + \alpha_1 \cdot 1 + \alpha_2 \cdot 1 + \alpha_3 \cdot 1 &= 0, \\ \alpha_0 \cdot 1 + \alpha_1 \cdot 2 + \alpha_2 \cdot 4 + \alpha_3 \cdot 8 &= 0, \\ \alpha_0 \cdot 1 + \alpha_1 \cdot 3 + \alpha_2 \cdot 9 + \alpha_3 \cdot 27 &= 0, \\ \alpha_0 \cdot 1 + \alpha_1 \cdot 4 + \alpha_2 \cdot 16 + \alpha_3 \cdot 64 &= 0, \end{aligned}$$

mit der eindeutigen Lösung $\alpha_j = 0, j = 0, \dots, 3$ führt. Beweis: Subtraktion der ersten Gleichung von allen anderen führt auf das Gleichungssystem

$$\begin{aligned}\alpha_0 \cdot 1 + \alpha_1 \cdot 1 + \alpha_2 \cdot 1 + \alpha_3 \cdot 1 &= 0, \\ \alpha_0 \cdot 0 + \alpha_1 \cdot 1 + \alpha_2 \cdot 3 + \alpha_3 \cdot 7 &= 0, \\ \alpha_0 \cdot 0 + \alpha_1 \cdot 2 + \alpha_2 \cdot 8 + \alpha_3 \cdot 26 &= 0, \\ \alpha_0 \cdot 0 + \alpha_1 \cdot 3 + \alpha_2 \cdot 15 + \alpha_3 \cdot 63 &= 0.\end{aligned}$$

Erneute Subtraktion des Zweifachen der neuen zweiten Gleichung von der neuen dritten Gleichung und des Dreifachen der neuen zweiten Gleichung von der neuen vierten ergibt das Gleichungssystem

$$\begin{aligned}\alpha_0 \cdot 1 + \alpha_1 \cdot 1 + \alpha_2 \cdot 1 + \alpha_3 \cdot 1 &= 0, \\ \alpha_0 \cdot 0 + \alpha_1 \cdot 1 + \alpha_2 \cdot 3 + \alpha_3 \cdot 7 &= 0, \\ \alpha_0 \cdot 0 + \alpha_1 \cdot 0 + \alpha_2 \cdot 2 + \alpha_3 \cdot 12 &= 0, \\ \alpha_0 \cdot 0 + \alpha_1 \cdot 0 + \alpha_2 \cdot 6 + \alpha_3 \cdot 42 &= 0.\end{aligned}$$

Subtraktion der vorletzten Zeile multipliziert mit 3 von der letzten Zeile ergibt das „dreieckige“ Gleichungssystem

$$\begin{aligned}\alpha_0 \cdot 1 + \alpha_1 \cdot 1 + \alpha_2 \cdot 1 + \alpha_3 \cdot 1 &= 0, \\ \alpha_0 \cdot 0 + \alpha_1 \cdot 1 + \alpha_2 \cdot 3 + \alpha_3 \cdot 7 &= 0, \\ \alpha_0 \cdot 0 + \alpha_1 \cdot 0 + \alpha_2 \cdot 2 + \alpha_3 \cdot 12 &= 0, \\ \alpha_0 \cdot 0 + \alpha_1 \cdot 0 + \alpha_2 \cdot 0 + \alpha_3 \cdot 6 &= 0.\end{aligned}$$

Nun sollte jeder die eindeutige Lösung $\alpha_j = 0, j = 0, \dots, 3$ mittels „Rückwärtseinsetzen“ ablesen können. Anmerkung zu diesen Rechnungen: Das wird später alles nochmal sauber, etwas abstrakter und, vor allem, algorithmischer behandelt, die Schlagworte dazu sind dann „das Gaußsche Eliminationsverfahren“ und „die LR-Zerlegung“. Hier sollte nur einmal „auf die harte Tour“ gezeigt werden, wie man lineare Unabhängigkeit nachweist. Dieser Weg ist derjenige, der im Zweifelsfall (bei guter Wahl der Auswertungsstellen) immer funktioniert.

Genauso *könnte* man es im Falle der oben angegebenen Legendre-Polynome machen. Allerdings ist es schneller zu zeigen, dass alle Linearkombinationen, die man mit den Monomen bekommt, auch im Falle der Legendre-Polynome berechnet werden können. Dazu zeigt man, dass man aus den Legendre-Polynomen die Monome basteln kann. Die ersten beiden Basisfunktionen sind identisch, also ist dort nichts zu zeigen. Das zweite Monom (die dritte Basisfunktion x^2) lässt sich eindeutig mittels des nullten und zweiten Legendre-Polynomes schreiben,

$$x^2 = \frac{2}{3} \cdot P_2(x) + \frac{1}{3} \cdot P_0(x) = \frac{2}{3} \cdot \frac{1}{2}(3x^2 - 1) + \frac{1}{3} \cdot 1$$

das dritte Monom lässt sich eindeutig mittels des ersten und dritten Legendre-Polynomes schreiben,

$$x^3 = \frac{2}{5} \cdot P_3(x) + \frac{3}{5} \cdot P_1(x) = \frac{2}{5} \cdot \frac{1}{2}(5x^3 - 3x) + \frac{3}{5} \cdot x.$$

Damit lässt sich jede Linearkombination

$$p(x) = \alpha_0 x^0 + \alpha_1 x^1 + \alpha_2 x^2 + \alpha_3 x^3 = \sum_{j=0}^3 \alpha_j x^j$$

der Monome sofort auch als Linearkombination der Legendre-Polynome schreiben, es gilt

$$p(x) = \alpha_0 P_0(x) + \alpha_1 P_1(x) + \alpha_2 \left(\frac{2}{3} \cdot P_2(x) + \frac{1}{3} \cdot P_0(x) \right) + \alpha_3 \left(\frac{2}{5} \cdot P_3(x) + \frac{3}{5} \cdot P_1(x) \right) \quad (134)$$

oder

$$p(x) = \beta_0 P_0(x) + \beta_1 P_1(x) + \beta_2 P_2(x) + \beta_3 P_3(x) = \sum_{j=0}^3 \beta_j P_j(x)$$

mit

$$\beta_0 = \alpha_0 + \frac{1}{3}\alpha_2, \quad (135a)$$

$$\beta_1 = \alpha_1 + \frac{3}{5}\alpha_3, \quad (135b)$$

$$\beta_2 = \frac{2}{3}\alpha_2, \quad (135c)$$

$$\beta_3 = \frac{2}{5}\alpha_3. \quad (135d)$$

Dadurch, dass man eine bekannte Basis mit anderen Kandidaten (damit laut Invarianz der Dimension gleicher Anzahl an Vektoren) für eine Basis ausdrücken kann, gewinnt man die Information, dass es sich um eine andere Basis *desselben* Raumes handelt. Dieser Weg ist oft (insbesondere mit mehr Erfahrung) der schnellere.

Egal, ob man nun die vier Zahlen α_j , $j = 0, \dots, 3$, oder äquivalent dazu, β_j , $j = 0, \dots, 3$, verwendet, kann man sich ganz im \mathbb{R}^4 wäghen. Die Summe zweier Funktionen und das skalare Vielfache von Funktionen (die Funktionen werden hier als Vektoren gesehen) ist nichts anderes als die Addition zweier Vektoren und das skalare Vielfache von Vektoren im \mathbb{R}^4 . Sei ein Polynom beschrieben durch Wahl von $\{\alpha_j\}_{j=0}^3$, ein zweites durch Wahl von $\{\tilde{\alpha}_j\}_{j=0}^3$. Als Beispiel nehmen wir

$$p(x) = 0 + 1x + 2x^2 + 3x^3$$

und

$$\tilde{p}(x) = 1 + 1x + 1x^2 + 1x^3,$$

beschrieben in der Basis der Monome durch die Vektoren

$$\alpha = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix}, \quad \text{und} \quad \tilde{\alpha} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

des \mathbb{R}^4 . Die Summe der beiden Polynome ist

$$(p + \tilde{p})(x) = 1 + 2x + 3x^2 + 4x^3,$$

entspricht also in der Basis der Monome dem Vektor

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = \alpha + \tilde{\alpha} = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

des \mathbb{R}^4 . Das skalare Vielfache eines Polynomes, z. B. das Zweifache des ersten Polynomes ergibt das Polynom

$$(2p)(x) = 0 + 2x + 4x^2 + 6x^3$$

mit dem Vektor der Koeffizienten in der Basis der Monome

$$\begin{pmatrix} 0 \\ 2 \\ 4 \\ 6 \end{pmatrix} = 2 \cdot \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix}.$$

Genauso ergeben die Vektoren aus den Koeffizienten $\{\beta_j\}_{j=0}^3$ eine Darstellung, als \mathbb{R}^4 , der Summen und skalaren Vielfachen, mithin aller „Vektoren“ (Polynome vom Höchstgrad 3) aus dem Vektorraum.

4 Anleitung vom 10.12.2010

Themen der Anleitung sind: Allgemeine Skalarprodukte; Orthogonalbasen; Normen; Orthonormalbasen; Orthonormalisierung nach Gram und Schmidt; Metriken; der Gauß-Algorithmus zur Lösung linearer Gleichungssysteme;

4.1 Skalarprodukte: Winkel & Längen

Motiviert aus den letzten Beobachtungen hat man bei zwei Vektoren ja (fast) immer eine äquivalente Darstellung des aufgespannten Raumes als \mathbb{R}^2 , also als Ebene. Dort, wie auch im \mathbb{R}^3 , welcher aus drei (linear unabhängigen) Vektoren gebildet wird, hat man aber in natürlicher Weise ein Skalarprodukt, welches dort über die „Längen“ der „Vektoren“ und den „Winkel“ zwischen ihnen gegeben ist. Was die Interpretation als Vektorraum angeht, so gibt es keinen Unterschied zwischen dem „normalen“ \mathbb{R}^3 und einem Raum z. B. aus den linear unabhängigen Funktionen $1, x, x^2$ gebildet.

Hier geht man den umgekehrten Weg. Man kann ja in einfacher Weise „das“ Skalarprodukt im \mathbb{R}^2 und \mathbb{R}^3 analog auf den \mathbb{R}^n verallgemeinern, indem man die äquivalente Beschreibung in Komponenten verwendet, nämlich

$$\langle x, y \rangle := \sum_{j=1}^n x_j \cdot y_j. \quad (136)$$

Man überzeugt sich schnell, dass dabei die „wesentlichen“ Eigenschaften erhalten bleiben. Welche Eigenschaften Mathematiker nach langer Forschung als „wesentlich“ herausgearbeitet haben, findet sich weiter unten.

Man kann jetzt umgekehrt wieder eine Länge und einen Winkel zwischen allgemeinen Vektoren bezüglich dieses Skalarproduktes definieren. Im Anschauungsraum \mathbb{R}^3 galt ja, dass die Länge eines Vektors beschreibbar war (dank des Satzes des Pythagoras) mit dem Skalarprodukt wie folgt,

$$|x| = \sqrt{\langle x, x \rangle} = \sqrt{\sum_{j=1}^3 x_j^2}.$$

Nun definiert man eine Länge als die Wurzel dieses Skalarproduktes mit sich selbst,

$$|x| := \sqrt{\langle x, x \rangle} = \sqrt{\sum_{j=1}^n x_j^2}. \quad (137)$$

Nun kann man, wenn man den Raum der Polynome vom Höchstgrad 3 wieder als Beispiel heranzieht, ja anscheinend zwei verschiedene „Skalarprodukte“ auf diesem über das entsprechende Skalarprodukt des \mathbb{R}^4 in der „ α “-Form oder in der „ β “-Form definieren.

Das erste „Skalarprodukt“ wäre gegeben für die Verwendung des „Standard“-Skalarproduktes in der Monombasis,

$$\langle p, \tilde{p} \rangle_{\Pi_{3,\alpha}} = \langle \alpha, \tilde{\alpha} \rangle_{\mathbb{R}_4} = \sum_{j=0}^3 \alpha_j \tilde{\alpha}_j.$$

Das zweite „Skalarprodukt“ wäre gegeben für die Verwendung des „Standard“-Skalarproduktes in der Legendre-Polynom-Basis,

$$\langle p, \tilde{p} \rangle_{\Pi_{3,\beta}} = \langle \beta, \tilde{\beta} \rangle_{\mathbb{R}_4} = \sum_{j=0}^3 \beta_j \tilde{\beta}_j.$$

Durch die Beschreibung der β in Abhängigkeit von den α in Gleichung (135) sieht man schnell, dass im Allgemeinen für $\alpha_2 \neq 0$ oder $\alpha_3 \neq 0$ sofort

$$\langle p, \tilde{p} \rangle_{\Pi_{3,\alpha}} \neq \langle p, \tilde{p} \rangle_{\Pi_{3,\beta}}$$

folgt. Zum Beispiel liefern die beiden Skalarprodukte verschiedene Ergebnisse, wenn $q = p = \tilde{p}$ mit $\alpha_0 = \alpha_1 = \alpha_3 = 0, \alpha_2 = 1$ gilt, also das Skalarprodukt des Polynomes

$$q(x) = x^2$$

mit sich selber berechnet werden soll. Die Koeffizienten in der Monombasis sind gegeben, die in der Legendre-Polynom-Basis berechnen sich zu

$$\beta_0 = \frac{1}{3}, \quad \beta_1 = 0, \quad \beta_2 = \frac{2}{3}, \quad \beta_3 = 0.$$

Es ergibt sich

$$\langle q, q \rangle_{\Pi_{3,\alpha}} = 1 \neq \frac{5}{9} = \left(\frac{1}{3}\right)^2 + \left(\frac{2}{3}\right)^2 = \langle q, q \rangle_{\Pi_{3,\beta}}$$

Die Frage, die sich jetzt stellt, ist, was wollen wir denn nun für „Skalarprodukte“ als „echt“ zulassen?

Im Fall des \mathbb{R}^3 wurden ja aus der Definition über Länge und Winkel mehrere Eigenschaften hergeleitet, die gerade nötig waren, um die kürzere Berechnung und damit Verallgemeinerung auf den \mathbb{R}^n zu ermöglichen.

Allgemein fordert man von einer Abbildung

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

die folgenden vier Eigenschaften ($u, v, w \in V, \lambda \in \mathbb{R}$),

$$\begin{aligned} \langle u + v, w \rangle &= \langle u, w \rangle + \langle v, w \rangle, & (\text{Linearität}) \\ \langle \lambda \cdot u, v \rangle &= \lambda \cdot \langle u, v \rangle, & (\text{Homogenität}) \\ \langle u, v \rangle &= \langle v, u \rangle, & (\text{Symmetrie}) \\ \langle v, v \rangle &> 0 \quad \forall v \in V, \quad v \neq 0. & (\text{Definitheit}) \end{aligned}$$

Sind alle vier Forderungen erfüllt, so bezeichnet man die Abbildung als Skalarprodukt.

Am problematischsten ist die vierte Forderung. Um nachzuweisen, dass eine Abbildung (positiv) definit ist, muss gezeigt werden, dass sie für jedes Argument ungleich Null positiv ist. Dazu versucht man häufig, die Abbildung als Summe von Quadraten zu schreiben. Ein dafür nützlicher Trick sei am Beispiel der Funktion

$$g(x, y) = 4x^2 - 7xy + 4y^2$$

vorgeführt. Es wäre (binomische Formeln) kein Problem gewesen, den Ausdruck als einen quadratischen Term zu schreiben, wenn dort statt der -7 eine ± 8 gestanden hätte. Allerdings kann man versuchen, die beiden binomischen Formeln

$$b_1(x, y) := (2x + 2y)^2 = 4x^2 + 8xy + 4y^2 \geq 0$$

und

$$b_2(x, y) := (2x - 2y)^2 = 4x^2 - 8xy + 4y^2 \geq 0$$

herzunehmen und eine Mittelung zu finden, so dass

$$g(x, y) = \alpha_1 b_1(x, y) + \alpha_2 b_2(x, y)$$

gilt. Diese Mittelung kann man aus dem Gleichungssystem

$$\begin{aligned} \alpha_1 + \alpha_2 &= 1 \\ 8\alpha_1 - 8\alpha_2 &= -7 \end{aligned}$$

berechnen, es gilt

$$\alpha_1 = \frac{1}{16} > 0, \quad \alpha_2 = \frac{15}{16} > 0,$$

und damit

$$g(x, y) = \frac{1}{16}(2x + 2y)^2 + \frac{15}{16}(2x - 2y)^2 \geq 0.$$

Jetzt bleibt zu zeigen, dass der Fall $g(x, y) = 0$ nur für $x = y = 0$ eintreten kann. Klarerweise ist g nur dann Null, wenn beide Ausdrücke in den Klammern gleich Null sind, also wenn das Gleichungssystem

$$\begin{aligned} 2x + 2y &= 0 \\ 2x - 2y &= 0 \end{aligned}$$

erfüllt ist. Die eindeutige Lösung ist aber gerade $x = y = 0$.

Nun könn(ten) wir zeigen, dass die beiden vorher „Skalarprodukte“ genannten Varianten „ α “ und „ β “ echte Skalarprodukte gemäß der obigen Definition sind.

Jeder, der an etwas Übung interessiert ist, teste für beide Varianten alle vier definierenden Eigenschaften. Für die Neugierigen und Denkschwachen hier ein kleines Resultat, welches nicht Bestandteil der Anleitung ist, und diesen Aufwand vermeidet:

LEMMA 1: Seien V und W isomorphe Vektorräume, d.h. es gebe eine bijektive Abbildung (einen „Isomorphismus“) $T : V \rightarrow W$ mit

$$T(u + v) = T(u) + T(v) \quad \forall u, v \in V, \quad (\text{Erhalt der Linearität})$$

$$T(\lambda \cdot v) = \lambda \cdot T(v) \quad \forall \lambda \in \mathbb{R}, \quad \forall v \in V. \quad (\text{Erhalt der Homogenität})$$

Sei ein Skalarprodukt $\langle \cdot, \cdot \rangle_W$ in W definiert. Dann definiert

$$\langle u, v \rangle_V := \langle T(u), T(v) \rangle_W \quad (138)$$

ein Skalarprodukt in V .

BEWEIS. Die Linearität und die Homogenität folgen aus der Kombination der Linearität und Homogenität des Skalarproduktes im Vektorraum W und des Isomorphismus T . Zuerst die Linearität, es gilt

$$\begin{aligned} \langle u + v, w \rangle_V &= \langle T(u + v), T(w) \rangle_W = \langle T(u) + T(v), T(w) \rangle_W \\ &= \langle T(u), T(w) \rangle_W + \langle T(v), T(w) \rangle_W \\ &= \langle u, w \rangle_V + \langle v, w \rangle_V. \end{aligned}$$

Nun die Homogenität, es gilt

$$\begin{aligned} \langle \lambda \cdot u, v \rangle_V &= \langle T(\lambda \cdot u), T(v) \rangle_W = \langle \lambda \cdot T(u), T(v) \rangle_W \\ &= \lambda \cdot \langle T(u), T(v) \rangle_W = \lambda \cdot \langle u, v \rangle_V. \end{aligned}$$

Die Symmetrie folgt aus der Symmetrie des Skalarproduktes in W ,

$$\langle u, v \rangle_V = \langle T(u), T(v) \rangle_W = \langle T(v), T(u) \rangle_W = \langle v, u \rangle_V.$$

Die Definitheit folgt aus der Definitheit des Skalarproduktes in W und der Bijektivität von T , es gilt zumindest (zur besseren Unterscheidung bezeichnen wir hier die Nullelemente/Nullvektoren/Nullfunktionen mit $0_V \in V$ und $0_W \in W$)

$$\langle v, v \rangle_V = \langle T(v), T(v) \rangle_W > 0 \quad \forall T(v) \neq 0_W \in W.$$

Da aber $0_V = T^{-1}(0_W)$ aufgrund der Vektorraumisomorphie gilt, ist das Nullelement 0_V von V der einzige Vektor, der im Skalarprodukt mit sich selbst Null ergibt. \square

Ein Beispiel eines anderen (für Funktionenräume typischen) Skalarproduktes ist durch

$$\langle p, q \rangle := \int_{-1}^1 p(x)q(x) dx \quad (139)$$

gegeben. Seien die zugrunde liegenden Funktionen wieder die Polynome vom Höchstgrad drei (welches die Existenz sämtlicher Integrale sichert). Die Linearität im ersten Argument folgt aus der Linearität des Integrales,

$$\begin{aligned} \langle p + r, q \rangle &= \int_{-1}^1 (p(x) + r(x))q(x) dx \\ &= \int_{-1}^1 p(x)q(x) dx + \int_{-1}^1 r(x)q(x) dx = \langle p, q \rangle + \langle r, q \rangle. \end{aligned}$$

Die Homogenität im ersten Argument folgt aus der Homogenität des Integrales,

$$\begin{aligned} \langle \lambda \cdot p, q \rangle &= \int_{-1}^1 (\lambda \cdot p(x))q(x) dx \\ &= \lambda \cdot \int_{-1}^1 p(x)q(x) dx = \lambda \cdot \langle p, q \rangle. \end{aligned}$$

Die Symmetrie folgt sofort aus der (punktweisen; d.h. für alle $x \in \mathbb{R}$ gültigen) Symmetrie der Multiplikation zweier Funktionen.

Die Definitheit folgt aus der Beobachtung, dass das (elementweise) Produkt eines Polynomes vom Höchstgrad drei mit sich selber eine nichtnegative Funktion (sogar wieder ein Polynom) ist, die (Fundamentalsatz der Algebra) höchstens 6 Nullstellen hat oder die Nullfunktion (das Nullpolynom) ist. Damit ist das Integral für alle Nichtnull-Polynome echt größer als Null.

Hat man ein Skalarprodukt gefunden, so definiert man eine Pseudolänge, eine sogenannte Norm eines Vektors, als die Wurzel dieses allgemeinen Skalarproduktes mit sich selbst,

$$\|x\| := \sqrt{\langle x, x \rangle}. \quad (140)$$

Die Normen bzgl. des „ α “- „ β “- und Standard-Skalarproduktes (139) für Polynome vom Höchstgrad drei am Beispiel des Polynomes

$$q(x) = x^2$$

sind also

$$\begin{aligned} \|q\|_\alpha &:= \sqrt{\langle q, q \rangle_{\Pi_{3,\alpha}}} = \sqrt{1 \cdot 1} = 1, \\ \|q\|_\beta &:= \sqrt{\langle q, q \rangle_{\Pi_{3,\beta}}} = \sqrt{\left(\frac{1}{3}\right)^2 + \left(\frac{2}{3}\right)^2} = \frac{\sqrt{5}}{3}, \\ \|q\|_f &:= \sqrt{\langle q, q \rangle} = \sqrt{\int_{-1}^1 x^2 \cdot x^2 dx} = \sqrt{\left[\frac{x^5}{5}\right]_{-1}^1} = \sqrt{\frac{2}{5}}, \end{aligned}$$

also alle verschieden. Je nach Art der „Längenmessung“ ergibt sich eine andere nichtnegative Größe. Aus der Definitheit aller Skalarprodukte folgt aber bereits, dass alle Normen entweder gleichzeitig gleich Null oder ungleich Null sein müssen.

Aus den vier algebraischen Forderungen an ein Skalarprodukt und die Definition der „zugeordneten“ Norm als Wurzel des Skalarproduktes eines Vektors mit sich selbst folgt bereits die Gültigkeit der CSU (Cauchy-Schwarzsche Ungleichung). Damit lassen sich jetzt in strikter Analogie Winkel zwischen Vektoren definieren,

$$\cos(\angle(v, w)) := \frac{\langle v, w \rangle}{\|v\| \|w\|}. \quad (141)$$

Aus dem Beweis der CSU ist klar, dass man dabei nicht die Normen und Skalarprodukte mischen kann.

Damit definiert *ein* gegebenes Skalarprodukt bereits in eindeutiger Weise „Längen“ (Normen) von Vektoren und Winkel zwischen zwei Vektoren.

4.2 Orthogonalbasen im Polynomraum

Wir hatten eben festgestellt: „Damit definiert ein gegebenes Skalarprodukt bereits in eindeutiger Weise „Längen“ (Normen) von Vektoren und Winkel zwischen zwei Vektoren.“

Insbesondere kann man jetzt von orthogonalen/senkrechten Vektoren im allgemeinen Vektorraum sprechen, im Hinterkopf hat man dabei immer ein gegebenes Skalarprodukt. So sind die Monome vom Höchstgrad 3 z. B. im „ α “-Skalarprodukt paarweise orthogonal und die angegebenen ersten vier Legendre-Polynome im „ β “-Skalarprodukt paarweise orthogonal.

Wir berechnen hier spaßeshalber die drei definierten inneren Produkte von $P_0(x) = 1$ und $P_2(x) = \frac{1}{2}(3x^2 - 1)$. Die Darstellungen in der Monom-Basis ist durch

$$\alpha(P_0) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \alpha(P_2) = \begin{pmatrix} -1/2 \\ 0 \\ 3/2 \\ 0 \end{pmatrix},$$

die Darstellung in der Legendre-Polynom-Basis durch (sic)

$$\beta(P_0) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \beta(P_2) = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

gegeben. Damit ergeben sich die drei Skalarprodukte zwischen P_0 und P_2 wie folgt,

$$\begin{aligned} \langle P_0, P_2 \rangle_{\Pi_{3,\alpha}} &= \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1/2 \\ 0 \\ 3/2 \\ 0 \end{pmatrix} \right\rangle = -\frac{1}{2}, \\ \langle P_0, P_2 \rangle_{\Pi_{3,\beta}} &= \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle = 0, \\ \langle P_0, P_2 \rangle &= \int_{-1}^1 1 \cdot \frac{1}{2}(3x^2 - 1) dx = \left[\frac{1}{2}(x^3 - x) \right]_{-1}^1 = 0. \end{aligned}$$

Also sind die beiden Legendre-Polynome P_0 und P_2 bezüglich der beiden letztgenannten Skalarprodukte orthogonal. Das die Legendre-Polynome bezüglich des zweiten Skalarproduktes paarweise orthogonal sind, ist trivial⁴. Wie sieht es mit dem über das Integral definierten Skalarprodukt aus? Sind dort auch alle Legendre-Polynome paarweise orthogonal zueinander? Ja. Der Beweis ist getrennt (Symmetrie-Eigenschaft des Skalarproduktes) in die Skalarprodukte mit

⁴trivial (adj., Mathematik): Sich meist nur durch langwierige und nicht sehr lehrreiche Rechnungen zeigen lassend. Diese werden dann oft dem geeigneten Leser überlassen. Niemand weiß heutzutage, warum der Leser dabei geneigt sein muß.

dem Legendre-Polynom P_0 ,

$$\langle P_0, P_1 \rangle = \int_{-1}^1 1 \cdot x \, dx = \left[\frac{1}{2} x^2 \right]_{-1}^1 = 0,$$

$$\langle P_0, P_2 \rangle = 0, \quad (\text{wie eben bereits gezeigt})$$

$$\langle P_0, P_3 \rangle = \int_{-1}^1 1 \cdot \frac{1}{2}(5x^3 - 3x) \, dx = \left[\frac{1}{2} \left(\frac{5}{4} x^4 - \frac{3}{2} x^2 \right) \right]_{-1}^1 = 0,$$

die noch nicht erfassten Skalarprodukte mit dem Legendre-Polynom P_1 ,

$$\langle P_1, P_2 \rangle = \int_{-1}^1 x \cdot \frac{1}{2}(3x^2 - 1) \, dx = \left[\frac{1}{2} \left(\frac{3}{4} x^4 - \frac{1}{2} x^2 \right) \right]_{-1}^1 = 0,$$

$$\langle P_1, P_3 \rangle = \int_{-1}^1 x \cdot \frac{1}{2}(5x^3 - 3x) \, dx = \left[\frac{1}{2} (x^5 - x^3) \right]_{-1}^1 = 0,$$

und zu guter Letzt das noch nicht erfasste Skalarprodukte mit dem Legendre-Polynom P_2 ,

$$\begin{aligned} \langle P_2, P_3 \rangle &= \int_{-1}^1 \frac{1}{2}(3x^2 - 1) \cdot \frac{1}{2}(5x^3 - 3x) \, dx \\ &= \int_{-1}^1 \frac{1}{4}(15x^5 - 14x^3 + 3x) \, dx \\ &= \left[\frac{1}{4} \left(\frac{5}{2} x^6 - \frac{7}{2} x^4 + \frac{3}{2} x^2 \right) \right]_{-1}^1 = 0. \end{aligned}$$

Damit ist gezeigt, dass die Legendre-Polynome P_0, \dots, P_3 eine Orthogonalbasis des Raumes der Polynome vom Höchstgrad 3 bezüglich des Integral-Skalarproduktes darstellen. Ausblick: Woher bekommt man (falls es solche gibt) so etwas wie Legendre-Polynome höheren Grades (also solche, die zu allen vorherigen Polynomen orthogonal sind)?

4.3 Normen & Metriken

Wir haben bereits gesehen, dass man in einem gegebenen Vektorraum V (Erinnerung: Was ist ein Vektorraum? \rightarrow 7 Axiome) mit einem gegebenen Skalarprodukt $\langle \cdot, \cdot \rangle$ (Erinnerung: Was ist ein Skalarprodukt? \rightarrow 4 Axiome) so etwas wie Längen definieren kann, nämlich die dem Skalarprodukt zugeordnete Norm

$$\|v\| := \sqrt{\langle v, v \rangle}. \quad (142)$$

Nun macht es auch Sinn, in einem Vektorraum ohne Skalarprodukt auch so etwas wie Längen zu haben, dann kann man nämlich anfangen, von Größe und Minimierung zu sprechen. Dazu muss man sich einigen, was man unter einer solchen (allgemeinen) Norm verstehen möchte. Es hat sich herausgestellt, dass drei Eigenschaften der von einem Skalarprodukt abgeleiteten Normen wichtig sind. Das führt auf die folgende Begriffsbildung.

Eine Abbildung eines Vektorraumes V nach \mathbb{R} ,

$$\|\cdot\| : \begin{cases} V \rightarrow \mathbb{R}, \\ v \mapsto \|v\|, \end{cases} \quad (143)$$

heißt eine Norm, wenn für $v, w \in V$, $\lambda \in \mathbb{R}$ gilt

$$\|v\| = 0 \Leftrightarrow v = 0, \quad (\text{Definitheit})$$

$$\|\lambda \cdot v\| = |\lambda| \cdot \|v\|, \quad (\text{Homogenität})$$

$$\|v + w\| \leq \|v\| + \|w\|. \quad (\text{Dreiecksungleichung})$$

Mit solch einer Norm hat man dann die Möglichkeiten, Dinge zu vergleichen und damit ein Instrument um den „optimalen“ Vektor für eine gegebene Problemstellung bezüglich der gewählten Norm herauszusuchen.

Bekannte Normen, die nicht direkt von einem Skalarprodukt stammen, sind die sogenannten Hölder-Normen oder p -Normen,

$$\|x\|_p := \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}, \quad p \geq 1, \quad x \in \mathbb{R}^n. \quad (144)$$

Im Limes $p \rightarrow \infty$ erhält man die Maximumnorm

$$\|x\|_\infty := \max_i |x_i|. \quad (145)$$

Diese wird häufig verwendet, ebenso die 1-Norm (Summennorm)

$$\|x\|_1 = \sum_{i=1}^n |x_i| \quad (146)$$

und die zu dem Euklidischen Skalarprodukt gehörige 2-Norm (Euklidische Norm)

$$\|x\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2} = \sqrt{\langle x, x \rangle}. \quad (147)$$

Normen werden verwendet, um Abstände zu messen, dazu misst man einfach die Länge der Differenz zweier Vektoren, berechnet also eine Distanz d gemäß

$$d(x, y) = \|x - y\|. \quad (148)$$

Solche „Distanzfunktionen“ $d(\cdot, \cdot)$ gibt es auch ohne Normen (ganz zu schweigen von Skalarprodukten). Wenn eine Abbildung

$$d : \begin{cases} V \times V \rightarrow \mathbb{R} \\ (x, y) \mapsto d(x, y) \end{cases} \quad (149)$$

die folgenden drei Eigenschaften (für alle $x, y, z \in V$)

$$d(x, y) = 0 \Leftrightarrow x = y, \quad (\text{Definitheit})$$

$$d(x, y) = d(y, x), \quad (\text{Symmetrie})$$

$$d(x, y) \leq d(x, z) + d(z, y), \quad (\text{Dreiecksungleichung})$$

hat, so spricht man von einer Metrik. Anschaulich kann man damit

- Punkte als gleich (Abstand Null) oder verschieden (Abstand ungleich Null) identifizieren, das entspricht der Definitheit,
- den Abstand zweier Punkte definieren, da der Abstand von x zu y gleich dem Abstand von y zu x ist, das entspricht der Symmetrie,
- Abstände durch Summen von Abständen beschränken oder zeigen, dass Umwege zu längeren Wegen führen, das folgt aus der Dreiecksungleichung.

Metriken sind Gegenstand von dem aktuellen Übungsblatt.

4.4 Orthogonal- und Orthonormalbasen

Sei V ein Vektorraum der Dimension k mit einem Skalarprodukt. Ein Satz von k paarweise aufeinander senkrecht stehenden Vektoren, von denen keiner der Nullvektor ist (ergo keiner die durch das Skalarprodukt induzierte Norm Null hat), nennt man eine Orthogonalbasis. Der „Orthogonal“-Teil ist klar, der „basis“-Teil folgt aus der folgenden Beobachtung. In einer Basis kann man jedes Element eines Raumes darstellen mittels einer (eindeutigen) Linearkombination der Basisvektoren. Die Eindeutigkeit ist zu der linearen Unabhängigkeit äquivalent, seien $v = \sum_{j=1}^k \alpha_j q^j$ und $v = \sum_{j=1}^k \beta_j q^j$ zwei Darstellungen von v in einer OGB $\{q^j\}_{j=1}^k$. Dann ist nach Subtraktion

$$0 = \sum_{j=1}^k (\alpha_j - \beta_j) q^j. \quad (150)$$

Wenn die Vektoren $\{q^j\}_{j=1}^k$ linear unabhängig sind, folgt, dass alle Skalare vor den Vektoren gleich Null sind, also die Eindeutigkeit aus

$$\alpha_j = \beta_j, \quad j = 1, \dots, k. \quad (151)$$

Ist umgekehrt die Eindeutigkeit gegeben, so sind die Vektoren $\{q^j\}_{j=1}^k$ linear unabhängig, was man schnell anhand eines Beweises durch Widerspruch einsieht. Wir nehmen also an, dass die $\{q^j\}_{j=1}^k$ linear abhängig sind. Dann gibt es eine nichttriviale Linearkombination der Null, also

$$0 = \sum_{j=1}^k \gamma_j q^j, \quad (152)$$

wobei mindestens ein γ_j ungleich Null ist. Dann ist mit einer gegebenen Darstellung $v = \sum_{j=1}^k \alpha_j q^j$ auch immer $v = \sum_{j=1}^k \beta_j q^j$, wobei $\beta_j = \alpha_j + \gamma_j$, denn

$$\sum_{j=1}^k \beta_j q^j = \sum_{j=1}^k (\alpha_j + \gamma_j) q^j = \sum_{j=1}^k \alpha_j q^j + \sum_{j=1}^k \gamma_j q^j = v + 0 = v. \quad (153)$$

Da mindestens ein β_j ungleich dem entsprechenden α_j ist, haben wir demnach zwei verschiedene Darstellungen gefunden. Sei jetzt E die Aussage der Eindeutigkeit, und LU die Aussage der linearen Unabhängigkeit. Dann haben wir eben gerade gezeigt, dass aus $\neg LU$ die Folgerung $\neg E$ gezogen werden kann, woraus wir mit $E = \neg(\neg E)$ schließen können, dass

$$\{\{\neg LU \Rightarrow \neg E\} \wedge \neg(\neg E)\} \Rightarrow \neg(\neg LU) = LU. \quad (154)$$

Dieser letzte Schluß ist in der Logik als *Modus (tollendo) tollens* bekannt.

Nun zeigen wir, dass sich in einer OGB alle Vektoren eindeutig darstellen lassen. Dazu nehmen wir eine gegebene Darstellung des generischen Vektors $v \in V$ an,

$$v = \sum_{j=1}^k \alpha_j q^j. \quad (155)$$

Jetzt können wir zeigen, dass die α_j sich bestimmen lassen als Anteile von Projektionen auf die (Richtung von) Basisvektoren. Es gilt mit dem gegebenen Skalarprodukt

$$\langle q^i, v \rangle = \sum_{j=1}^k \alpha_j \langle q^i, q^j \rangle = \sum_{j=1}^k \alpha_j \delta_{ij} \|q^i\|^2 = \alpha_i \|q^i\|^2. \quad (156)$$

Daraus (man beachte, dass die Normen der Vektoren als ungleich Null vorausgesetzt sind) kann man die Skalare $\{\alpha_j\}_{j=1}^k$ berechnen gemäß Projektion und Skalierung (Skalierung ist die Division durch oder Multiplikation mit einem Skalar) zu

$$\alpha_j = \frac{\langle q^j, v \rangle}{\|q^j\|^2} = \frac{\langle v, q^j \rangle}{\langle q^j, q^j \rangle}. \quad (157)$$

Insbesondere berechnen sich damit in eindeutiger Weise die Skalare $\{\gamma_j\}_{j=1}^k$ aus einer Linearkombination zu Null alle zu Null,

$$\gamma_j = \frac{\langle q^j, 0 \rangle}{\|q^j\|^2} = \frac{\langle 0, q^j \rangle}{\langle q^j, q^j \rangle} = 0. \quad (158)$$

Damit sind orthogonale Vektoren ungleich dem Nullvektor immer linear unabhängig. Die Anteile in der Darstellung

$$v = \sum_{j=1}^k \frac{\langle v, q^j \rangle}{\langle q^j, q^j \rangle} q^j \quad (159)$$

sind gerade die Projektionen von v in die Richtungen aller q^j . Der Vektor wird also gerade in seine Projektionen in orthogonale Richtungen zerlegt. Solche eine Zerlegung klappt nur, wenn alle Vektoren zueinander orthogonal sind.

Noch einfacher wird alles, wenn eine OGB vorliegt, in der jeder Vektor bezüglich der durch das gegebene Skalarprodukt gegebenen Norm die „Länge“ Eins hat. Solche eine OGB nennt man auch eine Orthonormalbasis („normal“ herrührend von normiert; „normiert“ herrührend von einer fest vorgegebene Norm, hier Eins, habend). Die Berechnung der Anteile vereinfacht sich dann zu

$$\alpha_j = \langle q^j, v \rangle, \quad (160)$$

d.h., dann gilt

$$v = \sum_{j=1}^k \alpha_j q^j = \sum_{j=1}^k \langle q^j, v \rangle q^j = \sum_{j=1}^k \langle v, q^j \rangle q^j. \quad (161)$$

Hat man eine Basis (diese kann man aus einem Erzeugendensystem konstruieren, das folgt aus dem Austauschsatz von Steinitz), so kann man versuchen, aus dieser eine OGB/ONB zu konstruieren. Die Idee dahinter geht zurück auf Gram und Schmidt.

Wie verwenden das resultierende Gram-Schmidt Orthonormalisierungsverfahren exemplarisch, um rekursiv eine ONB bezüglich des Skalarproduktes

$$\langle p, q \rangle := \int_{-1}^1 p(x)q(x) dx, \quad p, q \text{ Polynome} \quad (162)$$

für die Polynome zu erzeugen. Als gegebene Basis nehmen wir die Monome.

Das erste (oder, je nach Zählweise, das nullte) Monom ist die Funktion p_0 ,

$$p_0(x) = 1, \quad (163)$$

welches die Norm („Länge“)

$$\|p_0\| = \sqrt{\langle p_0, p_0 \rangle} = \sqrt{\int_{-1}^1 p_0(x)p_0(x) dx} = \sqrt{\int_{-1}^1 1 dx} = \sqrt{[x]_{-1}^1} = \sqrt{2} \quad (164)$$

hat. Teilt man dieses Monom durch seine Norm,

$$q_0 = \frac{p_0}{\sqrt{2}}, \quad (165)$$

so erhält man einen denselben Raum aufspannenden Vektor mit Norm Eins,

$$\|q_0\| = \sqrt{\langle q_0, q_0 \rangle} = \sqrt{\int_{-1}^1 \left(\frac{1}{\sqrt{2}}\right)^2 dx} = \sqrt{\left[\frac{1}{2}x\right]_{-1}^1} = 1. \quad (166)$$

Das zweite Monom ist die Funktion p_1 ,

$$p_1(x) = x. \quad (167)$$

Dieses ist bereits senkrecht zu q_0 , da

$$\langle q_0, p_1 \rangle = \int_{-1}^1 \frac{1}{\sqrt{2}} x dx = \left[\frac{1}{2\sqrt{2}} x^2 \right]_{-1}^1 = 0. \quad (168)$$

Das zweite Monom hat die Norm („Länge“)

$$\|p_1\| = \sqrt{\langle p_1, p_1 \rangle} = \sqrt{\int_{-1}^1 x^2 dx} = \sqrt{\left[\frac{x^3}{3} \right]_{-1}^1} = \sqrt{\frac{2}{3}}. \quad (169)$$

Damit hat die Funktion

$$q_1 : \left\{ x \mapsto \sqrt{\frac{3}{2}} x \right. \quad (170)$$

die Norm Eins und ist senkrecht zu q_0 . Damit haben wir bereits eine ONB für den Raum $\text{span}\{1, x\}$. Nun schauen wir, ob das dritte Monom p_2 ,

$$p_2(x) = x^2 \quad (171)$$

Anteile an den ersten beiden Basisvektoren q_0, q_1 hat, wenn ja, ziehen wir diese ab und normieren danach den Rest. Die Skalarprodukte sind gegeben als

$$\langle q_0, p_2 \rangle = \int_{-1}^1 \frac{1}{\sqrt{2}} x^2 dx = \left[\frac{1}{3\sqrt{2}} x^3 \right]_{-1}^1 = \frac{\sqrt{2}}{3}, \quad (172)$$

$$\langle q_1, p_2 \rangle = \int_{-1}^1 \sqrt{\frac{3}{2}} x^3 dx = \left[\sqrt{\frac{3}{2}} \frac{x^4}{4} \right]_{-1}^1 = 0. \quad (173)$$

Wir müssen also den Anteil in Richtung q_0 abziehen, der Anteil in Richtung q_1 ist sowieso gleich Null, also sind p_2 und q_1 schon orthogonal. Der bereinigte Vektor, der keine Anteil mehr an q_0 und q_1 hat, ist demnach gegeben als

$$\tilde{q}_2 = p_2 - \langle q_0, p_2 \rangle q_0 - \langle q_1, p_2 \rangle q_1, \quad \text{also gilt} \quad \tilde{q}_2(x) = x^2 - \frac{\sqrt{2}}{3} \frac{1}{\sqrt{2}} - 0 = x^2 - \frac{1}{3}, \quad (174)$$

also ist \tilde{q}_2 (bis auf eine Skalierung um $3/2$) durch das Legendre-Polynom P_2 gegeben (dieses wird definiert in Gleichung (133c)). Wir haben bereits gezeigt, dass die Legendre-Polynome eine Orthogonalbasis bilden, damit ist also \tilde{q}_2 orthogonal zu q_0 und q_1 . Die Norm von \tilde{q}_2 ist

$$\|\tilde{q}_2\| = \sqrt{\langle \tilde{q}_2, \tilde{q}_2 \rangle} = \sqrt{\int_{-1}^1 \frac{(3x^2 - 1)^2}{3^2} dx} = \sqrt{\frac{8}{45}} = \frac{2\sqrt{10}}{15}. \quad (175)$$

Der skalierte Vektor

$$q_2(x) = \frac{15}{2\sqrt{10}} \tilde{q}_2(x) = \frac{15}{2\sqrt{10}} \left(x^2 - \frac{1}{3} \right) = \frac{15}{2\sqrt{10}} x^2 - \frac{5}{2\sqrt{10}} \quad (176)$$

hat die Norm Eins und steht auf den ersten beiden Basisvektoren senkrecht, damit haben wir eine ONB für $\text{span}\{1, x, x^2\}$ erzeugt. So kann man weiter fortfahren und damit eine ONB für höhere Polynomräume konstruieren.

4.5 Metriken

Wie oben bereits im letzten Abschnitt erwähnt, gibt es „Distanzfunktionen“ $d(\cdot, \cdot)$ (auch ohne Normen; ganz zu schweigen von Skalarprodukten). Wenn eine Abbildung

$$d : \begin{cases} V \times V \rightarrow \mathbb{R} \\ (x, y) \mapsto d(x, y) \end{cases} \quad (177)$$

die folgenden drei Eigenschaften (für alle $x, y, z \in V$)

$$d(x, y) = 0 \Leftrightarrow x = y, \quad (\text{Definitheit})$$

$$d(x, y) = d(y, x), \quad (\text{Symmetrie})$$

$$d(x, y) \leq d(x, z) + d(z, y), \quad (\text{Dreiecksungleichung})$$

hat, so spricht man von einer Metrik. Anschaulich kann man damit

- Punkte als gleich (Abstand Null) oder verschieden (Abstand ungleich Null) identifizieren, das entspricht der Definitheit,
- den Abstand zweier Punkte definieren, da der Abstand von x zu y gleich dem Abstand von y zu x ist, das entspricht der Symmetrie,
- Abstände durch Summen von Abständen beschränken oder zeigen, dass Umwege zu längeren Wegen führen, das folgt aus der Dreiecksungleichung.

Hier ein auf den ersten Blick etwas ungewöhnliches Beispiel einer Metrik. Sei V ein Vektorraum, $\|\cdot\|$ eine auf diesem Vektorraum definierte Norm und

$$d(x, y) := \min\{\|x - y\|, 1\}, \quad x, y \in V.$$

Die drei Eigenschaften einer Metrik lassen sich schnell überprüfen:

$$d(x, y) = 0 \Leftrightarrow \min\{\|x - y\|, 1\} = 0$$

$$\Leftrightarrow \|x - y\| = 0 \Leftrightarrow x - y = 0 \Leftrightarrow x = y, \quad (\text{Definitheit})$$

$$d(x, y) = \min\{\|x - y\|, 1\} = \min\{\|y - x\|, 1\} = d(y, x), \quad (\text{Symmetrie})$$

$$\begin{aligned} d(x, y) &= \min\{\|x - y\|, 1\} \leq \min\{\|x - z\|, 1\} + \min\{\|z - y\|, 1\} \\ &= d(x, z) + d(z, y), \end{aligned} \quad (\text{Dreiecksungleichung})$$

wobei die letzten beiden Zeilen aus der Beobachtung folgen, dass die linke Seite in jedem Fall kleiner gleich Eins ist und falls auf der rechten Seite einer der beiden Summanden das Minimum bei Eins annimmt, die Ungleichung trivialerweise erfüllt ist. Der Fall, dass das Minimum jeweils bei dem Normenausdruck angenommen wird, läuft auf die Dreiecksungleichung der gegebenen Norm

$$\|x - y\| = \|x - z + z - y\| \leq \|x - z\| + \|z - y\|$$

hinaus. Bleibt theoretisch der Fall, dass $\max\{\|x - z\|, \|z - y\|\} \leq 1$ und $\|x - y\| > 1$, welcher aber auch durch die Dreiecksungleichung der Norm bereits erledigt ist. Man kann also eine Metrik konstruieren, indem man einen „Maximalabstand“ (hier als Eins gewählt) bei einer Metrik aus einer Norm einführt.

Ein noch ungewöhnlicheres Beispiel einer Metrik ist die *triviale Metrik*, welche definiert ist als

$$d(x, y) := \begin{cases} 0, & x = y, \\ 1, & x \neq y. \end{cases}$$

Dass diese Abbildung eine Metrik definiert weist jeder leicht selber nach.

4.6 Lineare Gleichungssysteme

Wir haben schon gesehen, dass Gleichungssysteme gebildet aus Forderungen an Linearkombinationen, also z. B.

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1,$$

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2,$$

...

$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m,$$

oft auftreten. Beispiele dazu sind Geraden- und Ebenengleichungen, auch die Untersuchung der linearen Unab- oder Abhängigkeit von Vektoren (siehe das Beispiel mit den Legendre-Polynomen) führen auf solche Gleichungssysteme.

Nun kann es (im Wesentlichen) drei verschiedene Möglichkeiten geben, wieviele Vektoren $x \in \mathbb{R}^n$ Lösungen sind. Diese drei Möglichkeiten zeigen wir hier einmal exemplarisch, dann in allgemeiner Darstellung.

Erste Möglichkeit: Es gibt keine Lösung. Beispiel ($m = n = 1$):

$$0 \cdot x_1 = 1.$$

Allgemeiner: Wenn der Vektor $b \in \mathbb{R}^m$ der rechten Seite nicht im von den Spaltenvektoren $a^i \in \mathbb{R}^m$,

$$a^j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix} \in \mathbb{R}^m, \quad j = 1, \dots, n \quad (179)$$

aufgespannten Raum

$$\mathcal{A} := \text{span}\{a^1, a^2, \dots, a^n\} \quad (180)$$

liegt, dann gibt es keine Lösung. Der Raum \mathcal{A} wird im Übrigen das *Bild von A* genannt. Das es keine Lösung geben kann, sieht man ein, da ein Gleichungssystem der obigen Gestalt ja eine Forderung an eine Linearkombination der Spalten ist, also

$$b = a^1 x_1 + a^2 x_2 + \dots + a^n x_n = \sum_{j=1}^n a^j x_j =: Ax. \quad (181)$$

Wenn b im Raum \mathcal{A} der Spaltenvektoren liegt, nennt man das Gleichungssystem „konsistent“.

Zweite Möglichkeit: Es gibt genau eine Lösung. Beispiel ($m = n = 1$):

$$1 \cdot x_1 = 1.$$

Allgemeiner: Wenn der Vektor b im Raum \mathcal{A} der Spaltenvektoren enthalten ist und die Spaltenvektoren linear unabhängig sind. Beweis: Seien zwei Lösungen x^1 und x^2 gegeben. Wir wollen zeigen, dass dann $x^1 = x^2$ gilt. In der Sicht als Linearkombinationen gilt

$$b = \sum_{j=1}^n a^j x_j^i = Ax^i, \quad i = 1, 2. \quad (182)$$

Subtraktion führt auf

$$0 = \sum_{j=1}^n a^j x_j^1 - \sum_{j=1}^n a^j x_j^2 = \sum_{j=1}^n (a^j x_j^1 - a^j x_j^2) = \sum_{j=1}^n a^j (x_j^1 - x_j^2) \quad (183a)$$

$$= Ax^1 - Ax^2 = A(x^1 - x^2). \quad (183b)$$

Damit haben wir einen Vektor $y := x^1 - x^2$ gefunden, der die linear unabhängigen Spaltenvektoren linear zur Null kombiniert,

$$Ay = A(x^1 - x^2) = Ax^1 - Ax^2 = b - b = 0 \quad (184)$$

$$= \sum_{j=1}^n a^j y_j = \sum_{j=1}^n a^j (x_j^1 - x_j^2) = 0. \quad (185)$$

Da aber die Spaltenvektoren als linear unabhängig vorausgesetzt waren, ist der Vektor y der Nullvektor, damit ist $x^1 = x^2$ bewiesen.

Dritte Möglichkeit: Es gibt unendlich viele Lösungen. Beispiel ($m = n = 2$):

$$\begin{aligned} 1x_1 + 2x_2 &= 3, \\ 1x_1 + 2x_2 &= 3. \end{aligned}$$

Das gegebene Gleichungssystem ist konsistent, denn das Bild \mathcal{A} von A besteht aus allen Vielfachen des Vektors aus lauter Einsen und b ist von diesem Typ. Eine Lösung ist klarerweise der Vektor

$$x^1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Aber auch der Vektor

$$x^2 = \begin{pmatrix} 5 \\ -1 \end{pmatrix}$$

ist eine Lösung, die beiden Lösungen sind überdies linear unabhängig. Wie vorher ist die Differenz

$$y = x^1 - x^2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 5 \\ -1 \end{pmatrix} = \begin{pmatrix} -4 \\ 2 \end{pmatrix}$$

der beiden Lösungen eine Lösung vom homogenen Gleichungssystem (d.h., das Gleichungssystem, wo die rechte Seite identisch gleich Null ist),

$$\begin{aligned} 1x_1 + 2x_2 &= 0, \\ 1x_1 + 2x_2 &= 0. \end{aligned}$$

Allgemein hat das Gleichungssystem unendlich viele Lösungen, wenn die rechte Seite b im Raum \mathcal{A} liegt und die Spaltenvektoren linear abhängig sind. Da b im Raum liegt, gibt es zumindest eine Linearkombination um b darzustellen, woraus man eine Lösung x gewinnt. Da die Spaltenvektoren linear abhängig sind, kann man immer einen Vektor aus Skalaren finden, so dass diese Skalare die Spaltenvektoren nichttrivial zur Null kombinieren. Sei dieser Vektor gegeben als y . Dann ist klar, dass

$$Ay = 0, \quad Ax = b. \quad (186)$$

Nun ist auch $\lambda \cdot y$ für alle $\lambda \in \mathbb{R}$ ein Vektor, dessen Komponenten die Spaltenvektoren linear zur Null kombinieren,

$$A(\lambda \cdot y) = \lambda \cdot Ay = \lambda \cdot 0 = 0. \quad (187)$$

Aus diesem Grund nennt man Gleichungssysteme, deren rechte Seite der Nullvektor ist, auch homogen.

Zusammengefasst gilt dann für alle $\lambda \in \mathbb{R}$

$$A(x + \lambda \cdot y) = Ax + \lambda \cdot Ay = Ax = b. \quad (188)$$

Damit haben wir aber unendlich viele Lösungen gefunden.

Eine Lösung findet man immer, indem man eine Basis von \mathcal{A} aus Spaltenvektoren berechnet und in dieser dann die eindeutige Lösung berechnet. Man kann dann zu dieser Lösung jede Lösung des homogenen Systems hinzuaddieren und erhält wieder eine Lösung.

Die Menge aller Lösungen des homogenen Systems ist trivialerweise eine Untermenge des \mathbb{R}^n , man kann schnell sehen, dass sie sogar ein Unterraum des \mathbb{R}^n ist. Zu überprüfen ist die Abgeschlossenheit bzgl. der Vektoraddition und der Skalarmultiplikation. Seien dazu $y, z \in \mathbb{R}^n$ Lösungen des homogenen Systems und $\lambda \in \mathbb{R}$. Dann gilt

$$A(y + z) = Ay + Az = 0 + 0 = 0 \quad (189)$$

$$A(\lambda \cdot y) = \lambda \cdot Ay = \lambda \cdot 0 = 0. \quad (190)$$

Damit bildet die Gesamtheit der Lösungen eines homogenen Gleichungssystems immer einen Untervektorraum des \mathbb{R}^n . Dieser Untervektorraum wird der *Kern von A* genannt.

Falls das inhomogene Gleichungssystem eine Lösung hat, gilt nach Satz 3.8 für die Menge aller Lösungen \hat{x} eine Darstellung der Gestalt

$$\hat{x} \in x_{\text{inhomogen}} + \mathcal{L}_{\text{homogen}}, \quad (191)$$

wobei der Vektorraum $\mathcal{L}_{\text{homogen}}$ der Vektorraum der Lösungen der homogenen Gleichung, also der Kern von A ist.

Beispiel: Eine homogene Gleichung in drei Variablen (also meist eine Ebene durch Null),

$$x_1 + 2x_2 + 3x_3 = 0.$$

Der Lösungsraum hat die Dimension $2 = 3 - 1$. Aufgespannt wird der Lösungsraum durch die beiden Vektoren

$$y^1 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \quad y^2 = \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}.$$

Allgemeiner gilt nach Satz 3.9

$$\dim(\mathcal{L}_{\text{homogen}}) = n - \dim(\mathcal{A}) \quad (192)$$

also ist die Summe der Dimensionen von Bild und Kern von A gleich der Anzahl der Spalten von A . Wenn zwei dieser Größen gegeben sind, ist die dritte trivialerweise berechenbar.

Eine Warnung, da diese Begriffe oft durcheinander gebracht werden: Die Menge der Lösungen eines homogenen Gleichungssystems bilden einen Untervektorraum, nämlich den Kern der Matrix, welcher also ein **Lösungsraum** ist, die Menge der Lösungen eines inhomogenen Gleichungssystems ist aber leider „nur“ eine Menge, also eine **Lösungsmenge**. Diesen speziellen Mengen gibt man aber auch einen Namen, der die Struktur dieser Mengen widerspiegelt, nämlich den der „**linearen Mannigfaltigkeit**“. Dazu später mehr. Zuerst *das* Standardtool zum Lösen linearer Gleichungssysteme: Der Gauß-Algorithmus.

4.7 Der Gauß-Algorithmus

Die Lösung von linearen Gleichungssystemen wurde schon vor ca. 2000 Jahren von den Chinesen beherrscht, die axiomatische Herangehensweise wurde hierzulande von Carl Friedrich Gauß (1777-1855) etabliert und der resultierende Algorithmus trägt noch heute diesen Namen.

Der Gauß-Algorithmus bietet eine einfache Methode, die obigen Betrachtungen in einen simplen Algorithmus zu packen. Was jeden Ingenieur interessieren sollte, ist die Durchführung des Algorithmus, da dieser immer wieder benötigt wird. In der Anleitung verzichten wir auf die Erwähnung der dahinter liegenden Theorie.

Die Grundidee hinter dem Algorithmus ist die Transformation der Matrix des Gleichungssystems in eine einfachere Gestalt, hier: Dreiecksgestalt. Beispiel zur Erklärung, Zahlen in Dreiecksmatrix schreiben, Diagonale ungleich Null, dann auflösen (vorwärts und rückwärts).

Bringen auf Dreiecksgestalt mittels elementaren Umformungen:

- Multiplikation einer Zeile mit Zahlen ungleich Null $v_i \rightarrow \lambda v_i$,
- Addition des Vielfachen einer Zeile zu einer anderen Zeile $v_j \rightarrow v_j + \lambda v_i$,
- **Vertauschen zweier Zeilen** $v_i \leftrightarrow v_j$,
- **Vertauschen zweier Spalten** $v^i \leftrightarrow v^j$.

Diese elementaren Umformungen ändern den Rang (die maximale Anzahl linear unabhängiger Zeilen, entspricht laut Beweis in der Vorlesung der maximalen Anzahl linear unabhängiger Spalten) der Matrix nicht.

Die Verwendung der ersten **drei** Umformungen führt auf Gauß-Elimination mit Spalten-tausch (und später mit partieller (oder Spalten-)Pivotisierung), die Verwendung aller **vier** Umformungen führt auf die Gauß-Elimination mit Spalten- und Zeilentausch (und später totaler Pivotisierung).

Beispiele ohne, mit Spalten- und mit vollständiger Pivotsuche an Tafel (insbesondere wahrscheinlich in der nächsten Anleitung).

Klausuraufgaben: Es sind die folgenden drei Vektoren gegeben,

$$v_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \quad v_2 := \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 := \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \quad (193)$$

Eine Aufgabe war es, die Aussage

„Ist das System $v_1x_1 + v_2x_2 + v_3x_3 = b$ für ein $b \in \mathbb{R}^4$ lösbar, so ist es für alle $b \in \mathbb{R}^4$ lösbar.“

zu bewerten. Nun löst

$$x = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad (194)$$

sicherlich das Gleichungssystem

$$v_1x_1 + v_2x_2 + v_3x_3 = v_1, \quad (195)$$

also haben wir eine rechte Seite gefunden (v_1), für die das System sicherlich lösbar ist (wir haben sogar die Lösung explizit angegeben). Andererseits ist aber der Vektor

$$b = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \quad (196)$$

sicher nicht im Spann der drei Vektoren v_1, v_2 und v_3 enthalten, da dieser Vektor auf allen drei Vektoren senkrecht steht, also existiert keine Lösung für dieses b . Damit ist die zu bewertende Aussage falsch.

Eine andere zu bewertende Aussage war

„Die Menge aller $b \in \mathbb{R}^4$, für die das System $v_1x_1 + v_2x_2 + v_3x_3 = b$ lösbar ist, bildet einen Unterraum des \mathbb{R}^4 .“

Nun ist ein lineares Gleichungssystem genau dann lösbar, wenn b im Untervektorraum aufgespannt von den Spaltenvektoren liegt. Damit ist die Menge aller b , für die das System lösbar ist, genau der *Untervektorraum*, welcher von v_1, v_2 und v_3 aufgespannt wird. Damit ist die zu bewertende Aussage wahr.

5 Anleitung vom 07.01.2011

Themen der Anleitung sind: Der Gauß-Algorithmus; lineare Mannigfaltigkeiten; lineare Abbildungen; Matrixdarstellung linearer Abbildungen.

5.1 Lineare Mannigfaltigkeiten

In der Beschreibung verschiedener geometrischer Dinge (Punkte, Geraden, Ebenen etcpp.) und bei der Beschreibung der Lösungsgesamtheit von linearen Gleichungssystemen tauchen immer wieder „verschobene“ Untervektorräume der Gestalt $W = v + U \subset V$, wobei $U \subset V$ ein Untervektorraum des Vektorraumes V ist, auf. Dabei ist die Notation

$$v + U := \{v + u, \quad u \in U\} \quad (197)$$

zu lesen als die Menge der Elemente, die sich aus der Addition von $v \in V$ und Elementen von $u \in U \subset V$ ergeben. Wenn $v \notin U$ ist, so ist diese Menge kein Vektorraum. Da man aber einen festen Ankerpunkt (wie bei den Geraden und Ebenen immer ausgenutzt) hat, und daran angeknüpft einen Vektorraum (linearen Raum), so nennt man diese „verschobenen“ Unterräume auch *affine Räume* oder *lineare Mannigfaltigkeiten*.

Beispiele: Funktionenraum, $5 \sin(x) + \alpha_1 \cos(x) + \alpha_2 \tan(x)$, wobei die Definitionsmenge eingeschränkt ist auf Punkte, für die die Funktionen auch definiert sind.

Erkennen: Eine lineare Mannigfaltigkeit muss mindestens ein Element eines Vektorraumes enthalten. Wenn es nur ein Element gibt, hat man als lineare Mannigfaltigkeit dieses Element als Ankerpunkt und angefügt den trivialen Vektorraum. Damit sind alle $v \in V$ bereits einfache lineare Mannigfaltigkeiten.

Unterräume $U \subset V$ von V sind auch lineare Mannigfaltigkeiten, denn als Ankerpunkt kann man ja den Nullvektor wählen (aber auch jeden anderen Vektor aus dem Unterraum).

Wie man gerade gesehen hat, ist dieser „Ankerpunkt“ im Allgemeinen nicht eindeutig. Dieses war aber auch schon bei den Geraden und den Ebenen im \mathbb{R}^3 bekannt. Z. B. ist eine lineare Mannigfaltigkeit durch

$$W = \left\{ g \in \mathbb{R}^2 \mid g = \begin{pmatrix} -3 \\ 1 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \end{pmatrix} \lambda, \quad \lambda \in \mathbb{R} \right\} \quad (198)$$

$$= \left\{ g \in \mathbb{R}^2 \mid g = \begin{pmatrix} -1 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \end{pmatrix} \lambda, \quad \lambda \in \mathbb{R} \right\} \quad (199)$$

parametrisiert. Wenn man erst einmal lineare Mannigfaltigkeiten hat, so ist jeder *nichtleere* Schnitt von zwei linearen Mannigfaltigkeiten (in dem selben Vektorraum V) eine lineare Mannigfaltigkeit (siehe dazu Satz 2.77). Das ist auch von den Geraden (ein eventueller Schnittpunkt ist lineare Mannigfaltigkeit) und Ebenen (eine eventuelle Schnittgerade ist lineare Mannigfaltigkeit) bekannt.

Ansonsten muss man überprüfen, ob alle Differenzen zweier Elemente aus einer linearen Mannigfaltigkeit W einen Vektorraum U bilden.

Die Entscheidung findet sinnvollerweise folgendermaßen statt:

1. Teste, ob $W \subset V$ (wenn V angegeben ist).
2. Teste, ob ein Element $w \in W$ existiert.
3. Wähle ein beliebiges Element $v \in W$ aus (potentieller „Ankerpunkt“).
4. Beweise/Widerlege, dass $U = \{w - v, w \in W\}$ ein Unterraum von V ist.

5.2 Lineare Abbildungen

In den Ingenieurwissenschaften treten aufgrund der einfachen geometrischen Bedeutung häufig Abbildungen $T : V \rightarrow W$ zwischen Vektorräumen V und W auf, welche die Eigenschaften

$$T(v + w) = T(v) + T(w), \quad \forall v, w \in V, \quad (\text{Erhalt der Addition})$$

$$T(\lambda \cdot v) = \lambda \cdot T(v), \quad \forall v \in V, \forall \lambda \in \mathbb{R} \quad (\text{Erhalt der Multiplikation})$$

haben, oder, äquivalent dazu, die Eigenschaft

$$T(\alpha v + \beta w) = \alpha T(v) + \beta T(w) \quad (\text{Linearität})$$

für alle $v, w \in V$ und für alle $\alpha, \beta \in \mathbb{R}$ haben.

Diese Abbildungen bilden lineare Gebilde, wie Geraden, Ebenen, Räume, lineare Mannigfaltigkeiten etcpp., auf ebensolche ab. Meist interessiert hierbei nur der Spezialfall von Geraden, Ebenen usw. durch die Null des Vektorraumes. Die genannten lineare Gebilde werden in diesem Falle mathematisch durch Untervektorräume beschrieben. Untervektorräume haben eine Basis, von denen sie aufgespannt werden; der Spann ist durch sämtliche Linearkombinationen der Basisvektoren gegeben.

Als Beispiel sei eine Ebene aufgespannt von $v_1, v_2 \in V$ gegeben, also eine Menge von Vektoren der Gestalt

$$\mathfrak{E} = \{\alpha_1 v_1 + \alpha_2 v_2, \quad \alpha_1, \alpha_2 \in \mathbb{R}\}. \quad (200)$$

Dann ist die Menge aller Bilder von Elementen von \mathfrak{E} unter einer linearen Abbildung $T : V \rightarrow W$ gegeben als

$$T(\mathfrak{E}) = \{T(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 T(v_1) + \alpha_2 T(v_2) = \alpha_1 w_1 + \alpha_2 w_2, \quad \alpha_1, \alpha_2 \in \mathbb{R}\}, \quad (201)$$

wobei $w_1 := T(v_1) \in W$ und $w_2 := T(v_2) \in W$. Wie man sieht, ist auch das Bild einer Ebene unter einer linearen Abbildung wieder eine Ebene (sofern w_1 und w_2 linear unabhängig sind). Gleiches gilt für andere Unterräume und auch für lineare Mannigfaltigkeiten und ist der Grund, diese Art von Abbildungen als „linear“ zu bezeichnen. Das Wort „linear“ taucht hier so oft auf, weil es sich bei dieser Vorlesung um die Vorlesung „Lineare Algebra“ handelt, und jetzt sollte Ihnen klar sein, warum diese so heisst und womit sie sich befasst (ergo: lineare Räume = Vektorräume, lineare Mannigfaltigkeiten, lineare Abbildungen \approx Matrizen, Linearkombinationen usw. usw.; was Algebra (von *al-jabr/al-ğabr* = Ergänzen) ist, und was man damit heute unter anderem bezeichnet, kommt im Abschnitt 5.5 und nie in der Vorlesung).

Lineare Abbildungen haben die schöne Eigenschaft, dass man nur die Wirkung der Abbildung auf eine (dann jede) Basis kennen muss, da alles anderen Werte sich automatisch aus der Eigenschaft der Linearität ergeben.

Sei dazu eine lineare Abbildung $T : V \rightarrow W$ zwischen zwei Vektorräumen V und W gegeben, und Basis des Raumes V sei v_1, \dots, v_n . Dann lässt sich jedes $v \in V$ aufgrund der Basiseigenschaft schreiben als

$$v = \sum_{i=1}^n \alpha_i v_i \quad (202)$$

für gewisse eindeutig gegebene $\{\alpha_i\}_{i=1}^n$. Die Bilder der $\{v_i\}_{i=1}^n$ unter T , also die

$$t_i = T(v_i) \in W, \quad i = 1, \dots, n, \quad (203)$$

seien bekannt. Dann gilt aufgrund der Linearität

$$T(v) = T\left(\sum_{i=1}^n \alpha_i v_i\right) = \sum_{i=1}^n \alpha_i T(v_i) = \sum_{i=1}^n \alpha_i t_i. \quad (204)$$

Die Bilder eines Vektors sind also eine Linearkombination der Bilder der Basisvektoren. Genauer: die Bilder eines Vektors, erzeugt als Linearkombination der Basisvektoren $\{v_i\}_{i=1}^n$ mittels gegebener $\{\alpha_i\}_{i=1}^n$ sind die Linearkombination der Bilder $\{t_i = T(v_i)\}_{i=1}^n$ der Basisvektoren gebildet mit denselben Skalaren $\{\alpha_i\}_{i=1}^n$.

5.3 Matrixmultiplikation I: Vektoren

Eine Matrix ist eine Sammlung von „zusammengehörigen“ Zahlen der Gestalt

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \in \mathbb{R}^{m \times n}, \quad m, n \in \mathbb{N}. \quad (205)$$

Matrizen tauchen primär bei Gleichungssystemen auf. Eine neue Operation Matrix (hier aus $\mathbb{R}^{m \times n}$) mal Vektor (hier aus \mathbb{R}^n) wird als Linearkombination (aus GLS) oder als „Zeile mal Spalte“ eingeführt,

$$A \cdot x = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad (206a)$$

$$= \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} x_2 + \cdots + \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n \in \mathbb{R}^m, \quad (206b)$$

$$A \cdot x = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \quad (207a)$$

$$= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} \in \mathbb{R}^m. \quad (207b)$$

Diese beiden Interpretationen ein und derselben Sache, nämlich der Matrix-Vektor-Multiplikation (MVM) sollten für alle Zeiten fest im Gehirn verankert werden.

Die erste Deutung, die als Linearkombination, zeigt, dass Matrizen lineare Abbildungen sind, respektive, darstellen, wenn man die Spalten der Matrix als Bilder von Basisvektoren der Abbildung begreift. Die Matrixmultiplikation (mit Matrizen der oben angegebenen Gestalt) bildet ja den \mathbb{R}^n in den \mathbb{R}^m ab, die Basisvektoren sind dabei die der Standardbasis, die Bilder sind die Spaltenvektoren der Matrix. Das Bild einer Linearkombination

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} x_1 + \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \end{pmatrix} x_2 + \cdots + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} x_n \quad (208)$$

ist gerade die entsprechende Linearkombination der Spalten der Matrix A , vergleiche mit Gleichung (206b).

Also stellen Matrizen lineare Abbildungen, besser gesagt, den *Prototyp der linearen Abbildung*, dar.

5.4 Matrixdarstellung linearer Abbildungen

Eine lineare Abbildung ist immer durch ihre Wirkung auf einen Satz von Basisvektoren festgelegt. Damit legen zwei Basen, eine im Urbildraum und eine im Bildraum eine Reihe von Zahlen

(eine Matrix) fest, die die lineare Abbildung beschreibt. Diese Matrix heisst eine der linearen Abbildung zugeordnete Matrix (bezüglich der gegebenen Basen).

Sei $T : V \rightarrow W$ linear, die Basis von V durch v_1, \dots, v_n gegeben, die Basis von W durch w_1, \dots, w_m gegeben. Nun berechnet man die Bilder $t_i = T(v_i)$ der Basisvektoren und stellt diese mittels der Basis w_j dar, berechnet also die Koeffizienten der Linearkombination der w_j , die als Ergebnis t_i ergibt,

$$t_i = w_1 a_{1i} + \dots + w_m a_{mi}. \quad (209)$$

Die Koeffizienten ergeben dann die Matrixdarstellung M_T der linearen Abbildung T ,

$$M_T = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \in \mathbb{R}^{m \times n}, \quad m, n \in \mathbb{N}. \quad (210)$$

Als Beispiel sei die Basis von $V = \mathbb{R}^3$ die Standardbasis,

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad (211)$$

und eine lineare Abbildung $\text{diff} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ gegeben als

$$\text{diff} \left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \right) = \begin{pmatrix} x_2 - x_1 \\ x_3 - x_2 \end{pmatrix}. \quad (212)$$

Diese lineare Abbildung liefert die diskrete Variante der Ableitung, nämlich die Differenzen (Differenzenquotienten) zu der Funktion $x : i \rightarrow x_i, i = 1, 2, 3$. Um eine Matrixdarstellung dieser „Ableitung“ zu haben, muss jetzt noch eine Basis im \mathbb{R}^2 gegeben sein, aus Langeweile und der Einfachheit halber wählen wir auch hier die Standardbasis,

$$e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{und} \quad e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (213)$$

Jetzt müssen wir erst die Bilder der drei Basisvektoren $e_i \in \mathbb{R}^3$ berechnen, welche als

$$\text{diff}(e_1) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad \text{diff}(e_2) = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad \text{und} \quad \text{diff}(e_3) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (214)$$

gegeben sind. Die Darstellung dieser Vektoren in der Standardbasis (sic) steht ja bereits da, also müssen wir nur noch die Matrix spaltenweise mit diesen Vektoren füllen,

$$M_{\text{diff}} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}. \quad (215)$$

Damit gilt dann

$$M_{\text{diff}} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_2 - x_1 \\ x_3 - x_2 \end{pmatrix} = \text{diff}(x). \quad (216)$$

Standardbasen sind zwar schön, aber selten. Also wählen wir im selben Beispiel als zweite Basis die Orthogonalbasis

$$q_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{und} \quad q_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (217)$$

Nun müssen wir die Ergebnisse aus Gleichung (214) noch in dieser Basis darstellen:

$$\text{diff}(e_1) = \begin{pmatrix} -1 \\ 0 \end{pmatrix} = -\frac{1}{2}(q_1 + q_2), \quad (218)$$

$$\text{diff}(e_2) = \begin{pmatrix} 1 \\ -1 \end{pmatrix} = q_2, \quad (219)$$

$$\text{diff}(e_3) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{2}(q_1 - q_2). \quad (220)$$

Die Koeffizienten lassen sich ablesen und werden spaltenweise in eine Matrix geschrieben:

$$\tilde{M}_{\text{diff}} = \frac{1}{2} \begin{pmatrix} -1 & 0 & 1 \\ -1 & 2 & -1 \end{pmatrix}. \quad (221)$$

Auch dieses ist eine (äquivalente) Matrixdarstellung, dieses Mal aber zu anderen Basen (eigentlich nur zu *einer* anderen Basis, da beide Male die Standardbasis im Urbildraum verwendet wurde). Dieses Beispiel wird gleich wieder aufgegriffen.

Die Differentiation von Elementen aus dem Polynomraum Π_n ergibt Polynome, deren Grad um Eins niedriger ist, bildet also einen $n + 1$ -dimensionalen Raum in einen n -dimensionalen ab, nämlich in der Raum der Polynome vom Höchstgrad kleiner n , Π_{n-1} . Wir wählen die Monombasis in beiden Räumen, und erhalten als Bilder der Basisvektoren skalierte Monome,

$$\frac{d}{dx}(x^k) = k \cdot x^{k-1}, \quad k = 0, 1, \dots, n. \quad (222)$$

Die Matrixdarstellung der Differentiation von Polynomen bezüglich der Monomenbasis ist also durch die Matrix

$$M_{\frac{d}{dx}} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 0 & n \end{pmatrix} \quad (223)$$

gegeben. Dieses mache man sich selbst anhand von Aufgabe 3 des Übungsblattes 5 klar.

5.5 Matrixmultiplikation II: Matrizen

Abbildungen kann man „verketteten“ / „kombinieren“ / hintereinander ausführen. Man kann sich schnell überzeugen, dass die Verkettung $S \circ T$ zweier linearer Abbildungen T und S ,

$$T : V \rightarrow W, \quad S : W \rightarrow U, \quad \Rightarrow \quad (S \circ T) : V \rightarrow U, \quad (224)$$

wieder linear ist, also wieder eine lineare Abbildung ist.

Da Matrizen den Prototyp der linearen Abbildung darstellen, kann man diese Verkettung von linearen Abbildungen auch für Matrizen hinschreiben. Diese Verkettung ist nur definiert, wenn es *einen* mittleren Raum gibt, der selbstverständlich eine *eindeutige* Dimension hat. Seien also zwei lineare Abbildungen \mathcal{A}, \mathcal{B} und ihre Verkettung \mathcal{C} ,

$$\mathcal{B} : \mathbb{R}^n \rightarrow \mathbb{R}^k, \quad \mathcal{A} : \mathbb{R}^k \rightarrow \mathbb{R}^m, \quad \Rightarrow \quad \underbrace{(\mathcal{A} \circ \mathcal{B})}_{=\mathcal{C}} : \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad (225)$$

durch Matrizen

$$A \in \mathbb{R}^{m \times k}, \quad B \in \mathbb{R}^{k \times n} \quad \text{und} \quad C \in \mathbb{R}^{m \times n} \quad (226)$$

gegeben. Dann kann man schnell aus der Identität

$$C \cdot x = A \cdot (B \cdot x) \quad (227)$$

durch Einsetzen der Standardbasisvektoren für $x \in \mathbb{R}^n$,

$$c_i = C \cdot e_i = A \cdot (B \cdot e_i) = A \cdot (b_i) = \sum_{j=1}^k a_j b_{ji} \quad (228)$$

$$\begin{pmatrix} c_{1i} \\ \vdots \\ c_{mi} \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mk} \end{pmatrix} \begin{pmatrix} b_{1i} \\ \vdots \\ b_{ki} \end{pmatrix} \quad (229)$$

$$= \sum_{j=1}^k \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix} b_{ji} = \begin{pmatrix} a_{11}b_{1i} + \cdots + a_{1k}b_{ki} \\ \vdots \\ a_{m1}b_{1i} + \cdots + a_{mk}b_{ki} \end{pmatrix} \quad (230)$$

$$= \begin{pmatrix} \sum_{j=1}^k a_{1j}b_{ji} \\ \vdots \\ \sum_{j=1}^k a_{mj}b_{ji} \end{pmatrix} \quad (231)$$

die Verkettung oder *Multiplikation* zweier Matrizen

$$C := A \cdot B \quad (232)$$

eingeführt als Matrix gebildet aus der Linearkombination der Spalten von A mit den Koeffizienten in den Vektoren der Spalten von B . Die Elemente c_{ji} der Matrix C sind also durch die Ausdrücke

$$c_{ji} = \sum_{\ell=1}^k a_{j\ell} b_{\ell i} \quad (233)$$

gegeben, graphisch kann man sich diesen Ausdruck als das Skalarprodukt von der j ten Zeile von A mit der i ten Spalte von B veranschaulichen (Merkphrase: „Zeile mal Spalte“).

Damit ist die Operation „Matrix mal Matrix“ definiert als die Verkettung zweier linearer Abbildungen, und wurde bereits interpretiert als Matrix mal viele Spaltenvektoren (zwei Deutungen, siehe MVM), oder andersherum als Linearkombinationen der Zeilen der zweiten Matrix, und als Hauptdeutung als $n \cdot m$ Skalarprodukte („Zeile mal Spalte“).

Was als Interpretation noch möglich und manchmal recht nützlich ist, ist die Interpretation als Summe von k dyadischen Produkten, und als absolute Verallgemeinerung, die Blockmatrixmultiplikation. Zur blockweisen Multiplikation begnügen wir uns mit ein paar Beispielen, eines davon ist dann die Multiplikation interpretiert als Summe von dyadischen Produkten.

Aus der Deutung der Hintereinanderschaltung zweier linearer Abbildungen ist klar, dass die Dimension des „mittleren“ Raumes dabei gleich bleiben sollte, mit anderen Worten, dass die Anzahl der Spalten der ersten Matrix gleich der Anzahl der Zeilen der zweiten Matrix sein muss.

Matrizen bilden neben den Vektorräumen das wichtigste Konstrukt der Linearen Algebra. Die Menge aller $n \times n$ Matrizen zusammen mit der Matrizenmultiplikation bilden eine sogenannte (assoziative) *Algebra*. Eine (assoziative) Algebra ist ein Vektorraum (hier der Vektorraum der Matrizen, der $\mathbb{R}^{n \times n}$ (Warum: Was ist die Addition von Matrizen? Wie definiert man die Skalarmultiplikation mit Skalaren aus \mathbb{R} ? Zum Abschluß die 7 Axiome eines Vektorraumes überprüfen.), in dem eine Multiplikation definiert ist, welche assoziativ ist, also, für die

$$A \cdot (B \cdot C) = (A \cdot B) \cdot C \quad (234)$$

gilt, und in der die Addition und Multiplikation verträglich sind, also, wo (da keine Kommutativität vorausgesetzt wird) zwei Distributivitätsgesetze

$$A \cdot (B + C) = A \cdot B + A \cdot C, \quad (A + B) \cdot C = A \cdot C + B \cdot C \quad (235)$$

gelten. Dass diese zwei Gesetze anstatt nur einem nötig sind, folgt daraus, dass für $n \geq 2$ die Matrixmultiplikation in aller Regel nicht kommutativ ist.

Die Operation „Matrix mal Matrix“ ist auch umgekehrt ausführbar, wenn beide Matrizen quadratisch sind (also die Anzahl der Spalten gleich der Anzahl der Zeilen ist) und beide gleich groß. In diesem Fall gilt trotzdem meist

$$A \cdot B \neq B \cdot A. \quad (236)$$

Als Beispiel berechnen wir mal kurz die beiden Produkte der Matrizen

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix}. \quad (237)$$

Die Ergebnisse sind

$$A \cdot B = \begin{pmatrix} 10 & 13 \\ 22 & 29 \end{pmatrix} \neq \begin{pmatrix} 11 & 16 \\ 19 & 28 \end{pmatrix} = B \cdot A \quad (238)$$

Selbstverständlich gilt diese Ungleichheit immer, wenn beide Produkte definiert sind, aber die Ergebnisse ungleich groß interpretiert als Matrix.

Nun ein kleiner Exkurs zur blockweisen Multiplikation. Man kann Matrizen in Blöcke zerlegen, z. B.

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \quad (239)$$

und dann (bei passender Blockgrösse) „blockweise“ multiplizieren,

$$\mathbf{A} \cdot \mathbf{B} =: \mathbf{C} \quad (240)$$

$$= \begin{pmatrix} A_{11} \cdot B_{11} + A_{12} \cdot B_{21} & A_{11} \cdot B_{12} + A_{12} \cdot B_{22} \\ A_{21} \cdot B_{11} + A_{22} \cdot B_{21} & A_{21} \cdot B_{12} + A_{22} \cdot B_{22} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}. \quad (241)$$

Die Blockzerlegung muss dabei nicht in quadratische Blöcke oder mit einem quadratischen Schema von Blöcken erfolgen, nur alle auftretenden Produkte müssen definiert sein. Als Beispiel seien hier die Matrizen

$$F = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} F_{11} \\ F_{21} \end{pmatrix} \quad \text{und} \quad G = \begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix} = \begin{pmatrix} G_{11} & G_{12} \end{pmatrix} \quad (242)$$

mit

$$F_{11} = \begin{pmatrix} 1 & 2 \end{pmatrix}, \quad F_{21} = \begin{pmatrix} 3 & 4 \end{pmatrix}, \quad G_{11} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, \quad G_{12} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}, \quad (243)$$

gegeben. Da $F = A$ und $G = B$ ist das Ergebnis für $F \cdot G$ ja bereits bekannt. Die blockweise Multiplikation ist möglich, da jedes Mal die Anzahlen der Spalten und Zeilen der beiden zu multiplizierenden Matrizen übereinstimmen und ergibt, wie nicht anders zu erwarten,

$$F \cdot G = \begin{pmatrix} F_{11} \cdot G_{11} & F_{11} \cdot G_{12} \\ F_{21} \cdot G_{11} & F_{21} \cdot G_{12} \end{pmatrix} \quad (244)$$

$$= \begin{pmatrix} \left\langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \end{pmatrix} \right\rangle & \left\langle \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \end{pmatrix} \right\rangle \\ \left\langle \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \end{pmatrix} \right\rangle & \left\langle \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \end{pmatrix} \right\rangle \end{pmatrix} \quad (245)$$

$$= \begin{pmatrix} 2 + 8 & 3 + 10 \\ 6 + 16 & 9 + 20 \end{pmatrix} = \begin{pmatrix} 10 & 13 \\ 22 & 29 \end{pmatrix}. \quad (246)$$

Dieses ist die meistgebrauchte Art der Matrixmultiplikation und sollte von jedem beherrscht werden. Um sich diese Art der Matrixmultiplikation zu merken, muss man nur die Phrase „Zeilen mal Spalten“ behalten. Man sieht hierdurch, dass die Matrixmultiplikation aus n^2 Skalarprodukten (oder inneren Produkten) besteht.

Die gewählten Blöcke von F und G lassen aber auch die umgekehrte Multiplikation zu, nämlich als Summe von n dyadischen (oder äusseren) Produkten,

$$G \cdot F = (G_{11} \cdot F_{11} + G_{21} \cdot F_{21}) \quad (247)$$

$$= \begin{pmatrix} 2 \\ 4 \end{pmatrix} \cdot (1 \ 2) + \begin{pmatrix} 3 \\ 5 \end{pmatrix} \cdot (3 \ 4) \quad (248)$$

$$= \begin{pmatrix} 2 & 4 \\ 4 & 8 \end{pmatrix} + \begin{pmatrix} 9 & 12 \\ 15 & 20 \end{pmatrix} = \begin{pmatrix} 11 & 16 \\ 19 & 28 \end{pmatrix}, \quad (249)$$

wie nicht anders zu erwarten war. Für die Langsamen wird hier jetzt ein dyadisches Produkt zuerst in der Form jedes Element mit jedem Element berechnet:

$$\begin{pmatrix} 2 \\ 4 \end{pmatrix} \cdot (1 \ 2) = \begin{pmatrix} 2 \cdot 1 & 2 \cdot 2 \\ 4 \cdot 1 & 4 \cdot 2 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 4 & 8 \end{pmatrix}. \quad (250)$$

Zeilenweise umgedeutet kann man auch so rechnen:

$$\begin{pmatrix} 2 \\ 4 \end{pmatrix} \cdot (1 \ 2) = \begin{pmatrix} 2 \cdot (1 \ 2) \\ 4 \cdot (1 \ 2) \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 4 & 8 \end{pmatrix}. \quad (251)$$

Spaltenweise umgedeutet kann man auch so rechnen:

$$\begin{pmatrix} 2 \\ 4 \end{pmatrix} \cdot (1 \ 2) = \left(\begin{pmatrix} 2 \\ 4 \end{pmatrix} \cdot 1 \quad \begin{pmatrix} 2 \\ 4 \end{pmatrix} \cdot 2 \right) = \begin{pmatrix} 2 & 4 \\ 4 & 8 \end{pmatrix}. \quad (252)$$

Der *Rang einer Matrix* A ist die Anzahl linear unabhängiger Spalten von A . Welchen Rang hat eine Matrix berechnet aus *einem* dyadischen Produkt zweier Nicht-Null-Vektoren?

Ein kleiner Exkurs zu dem Beispiel der diskreten Ableitung als linearer Abbildung vom \mathbb{R}^3 in den \mathbb{R}^2 . Wir hatten die beiden Darstellungen

$$M_{\text{diff}} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}, \quad \tilde{M}_{\text{diff}} = \frac{1}{2} \begin{pmatrix} -1 & 0 & 1 \\ -1 & 2 & -1 \end{pmatrix} \quad (253)$$

dieser Abbildung zu der Standardbasis im \mathbb{R}^3 und der Standardbasis (M_{diff}) respektive der Orthogonalbasis bestehend aus den Basisvektoren

$$q_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{und} \quad q_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad (254)$$

zur Gewinnung der Matrixdarstellungen verwendet. Schreibt man die beiden Basisvektoren der Orthogonalbasis in eine Matrix Q ,

$$Q = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad (255)$$

so lässt sich M_{diff} aus \tilde{M}_{diff} wieder berechnen, indem man von links mit Q multipliziert,

$$M_{\text{diff}} = Q \cdot \tilde{M}_{\text{diff}}, \quad (256)$$

$$\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \frac{1}{2} \begin{pmatrix} -1 & 0 & 1 \\ -1 & 2 & -1 \end{pmatrix} \quad (257)$$

Die Matrix \tilde{Q} definiert als

$$\tilde{Q} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{2} Q \quad (258)$$

leistet Ungewöhnliches, indem sie die „Abbildung“ Q „aufhebt“ (invertiert), da

$$\tilde{Q} \cdot Q = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2 \quad (259)$$

und

$$\tilde{Q} \cdot M_{\text{diff}} = \tilde{M}_{\text{diff}}, \quad (260)$$

$$\frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} -1 & 0 & 1 \\ -1 & 2 & -1 \end{pmatrix}. \quad (261)$$

Matrizen mit solchen oder ähnlichen Eigenschaften werden uns demnächst mehr interessieren und sind danach dann nicht mehr wegzudenken. Frage: Welche Matrix leistet eine solche „Aufhebung“ für die Matrix \hat{Q} definiert als

$$\hat{Q} := \frac{1}{\sqrt{2}} Q, \quad (262)$$

welche in den Spalten eine ONB (also eine OGB mit zusätzlicher Normierung) des \mathbb{R}^2 enthält?

Wir haben gesehen, dass man auch „Spaltenvektoren“ (also $n \times 1$ -Matrizen) und „Zeilenvektoren“ (also $1 \times n$ -Matrizen) miteinander multiplizieren kann, um damit eine $n \times n$ -Matrix zu erhalten. Da dabei etwas „eindimensionales“, nämlich ein Vektor, auf etwas „zweidimensionales“, nämlich eine Matrix abgebildet wird, nennt man dieses Matrixprodukt auch das äussere Produkt. Wählt man als Zeilenvektor den zum Spaltenvektor transponierten Vektor v und teilt dann durch das Skalarprodukt (manchmal auch das „innere Produkt“ genannt) von v mit sich selbst, so erhält man die Projektion in Richtung von v , da für jeden Vektor $a \in \mathbb{R}^n$

$$P_v a = \frac{v v^T}{v^T v} a = v \frac{v^T a}{v^T v} = v \frac{\langle v, a \rangle}{\langle v, v \rangle}.$$

6 Anleitung vom 21.01.2011

Themen der Anleitung sind Kongruenztransformationen, orthogonale Matrizen; etwas Training und besseres Verständnis von Matrix-Matrix-Multiplikation (insbesondere mit Zeilen- und Spaltenvektoren); das Gaußsche Eliminationsverfahren; „die“ LR-Zerlegung mit partieller Pivotisierung; Regularität und Singularität von Matrizen.

6.1 Kongruenztransformationen

Wir haben bereits Abbildungen kennengelernt, welche Vektorraum-Eigenschaften (Vektoraddition, skalare Multiplikation) erhalten, nämlich die linearen Abbildungen. Die linearen Abbildungen sind gerade die, die lineare Mannigfaltigkeiten auf lineare Mannigfaltigkeiten abbilden, indem die Vektoraddition und die skalare Multiplikation erhalten bleiben,

$$T(v_1 + v_2) = T(v_1) + T(v_2), \quad T(\lambda \cdot v) = \lambda \cdot T(v).$$

Wir schränken diese linearen Abbildungen auf solche ein, die einen Vektorraum in sich selber abbilden, in der zugeordneten Matrixdarstellung ergeben sich dann quadratische Matrizen. Jetzt hat ein Vektorraum oft mehr Struktur, oft ist nämlich auch ein Skalarprodukt $\langle \cdot, \cdot \rangle$ in dem Vektorraum V definiert. Wie wir bereits in Abschnitt 4.1 gesehen haben, definieren Skalarprodukte Winkel zwischen Vektoren und über die zugehörige Norm so etwas wie Längen.

Jetzt schränken wir die wichtige Klasse der linearen Abbildungen (auch Transformationen genannt) weiter ein, nämlich auf lineare Abbildungen eines Vektorraumes $(V, +, \cdot, \langle \cdot, \cdot \rangle)$ mit Skalarprodukt in sich selbst (sogenannte Endomorphismen)

$$Q : V \rightarrow V, \quad Q : v \mapsto w = Q \cdot v \quad (263)$$

die die Winkel zwischen zwei beliebigen Vektoren und die Länge erhalten, also welche

$$\langle Qv_1, Qv_2 \rangle = \langle v_1, v_2 \rangle \quad \forall v_1, v_2 \in V \quad (264)$$

erfüllen. Diese linearen Abbildungen nennt man Kongruenztransformationen. Die zur gleichen Orthonormalbasis in Urbildraum (z. B. gegeben als V) und Bildraum (auch gegeben als V) gehörigen Matrixdarstellungen von Kongruenztransformationen werden *orthogonale Matrizen* genannt. Wir zeigen im folgenden, dass diese orthogonalen Matrizen $Q \in \mathbb{R}^{n \times n}$ sich durch die Gleichung $Q^T Q = E_n$ charakterisieren lassen.

Sind die in den Spalten einer zugeordneten Matrix stehenden Vektoren des \mathbb{R}^n linear abhängig, so gibt es einen Nicht-Nullvektor $x \in \mathbb{R}^n$, den man auch auf Länge $\|x\| = \sqrt{\langle x, x \rangle} = 1$ skalieren kann, so dass $Qx = o_n$ gilt, wobei o_n der Nullvektor des \mathbb{R}^n ist. Es gilt mit diesem x also

$$0 = \langle o_n, o_n \rangle = \langle Qx, Qx \rangle \neq \langle x, x \rangle = 1,$$

also bildet diese lineare Abbildung nie *alle* Vektoren der Länge Eins auf Vektoren der Länge Eins ab, kann also keine Kongruenztransformation sein. Damit ist klar, dass Kongruenztransformationen regulär sein müssen. Es gibt Kongruenztransformationen, da z. B. die Identität trivialerweise eine solche ist, auch Drehungen und Spiegelungen gehören dazu. Um die Kongruenztransformationen insgesamt zu charakterisieren, benötigen wir die Definition der adjungierten Abbildung zu einer linearen Abbildung.

Man definiert die Adjungierte A^* einer linearen Abbildung A durch die Bedingung

$$\langle x, A^* y \rangle = \langle Ax, y \rangle, \quad \forall x, y \in V. \quad (265)$$

Im Falle des Standard-Skalarproduktes über \mathbb{R} kann man schnell über die Verwendung der Standardbasis $\{e_i\}_{i=1}^n$ für x und y nachrechnen (und danach gleich wieder den Rechenweg vergessen), dass

$$A^* = A^T \quad (266)$$

gilt, wobei das nachgestellte T eine Spiegelung der Position entlang der Diagonalen der Matrix bedeutet, also z. B.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}, \quad (0 \quad 1 \quad 2)^T = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \quad (267)$$

wohingegen im Falle des Standard-Skalarproduktes über \mathbb{C}

$$A^* = A^H := A^* := \overline{A}^T = \overline{A^T} \quad (268)$$

gilt, wobei die Überstreichung die komplexe Konjugation bedeutet. Allgemein gilt für Kongruenztransformationen für alle $x, y \in V$

$$0 = \langle Qx, Qy \rangle - \langle x, y \rangle \quad (269a)$$

$$= \langle x, Q^* Qy \rangle - \langle x, y \rangle \quad (269b)$$

$$= \langle x, Q^* Qy - y \rangle \quad (269c)$$

$$= \langle x, (Q^* Q - E)y \rangle. \quad (269d)$$

Nun wählt man $x = (Q^* Q - E)y$ und erhält, dass für alle $y \in V$

$$0 = \langle x, (Q^* Q - E)y \rangle \quad (270a)$$

$$= \langle (Q^* Q - E)y, (Q^* Q - E)y \rangle \quad (270b)$$

$$= \|(Q^* Q - E)y\|^2. \quad (270c)$$

Durch die Wahl des Vektors y als einen der Basisvektoren sieht man, dass die Matrixdarstellung von $Q^*Q - E$ in der entsprechenden Spalte (damit in allen Spalten) eine Nullspalte hat. Damit gilt dann

$$Q^*Q = E. \quad (271)$$

Analog sieht man, dass auch

$$QQ^* = E \quad (272)$$

gilt. Wir betrachten meist nur den Fall des Standardskalarproduktes im \mathbb{R}^n oder \mathbb{C}^n . Im ersten Fall werden Kongruenztransformationen durch orthogonale Matrizen

$$Q^T Q = E = Q Q^T, \quad \text{also} \quad Q^{-1} = Q^T, \quad Q \in \mathbb{R}^{n \times n} \quad (273)$$

im zweiten Fall durch unitäre Matrizen

$$U^H U = E = U U^H, \quad \text{also} \quad U^{-1} = U^H, \quad U \in \mathbb{C}^{n \times n} \quad (274)$$

beschrieben.

Diese Matrixgleichungen sind eine Kurzschreibweise für eigentlich n^2 Gleichungen, hier am Beispiel der orthogonalen Matrizen, die Aussagen für unitäre Matrizen folgen analog. Die Diagonalelemente entsprechen der Normalisierung der Spaltenvektoren (ergo, $q_i \in \mathbb{R}^n = \mathbb{R}^{n \times 1}$) respektive, Zeilenvektoren (ergo, $q^i \in \mathbb{R}^{1 \times n}$),

$$q_i^T q_i = \langle q_i, q_i \rangle = \|q_i\|^2 = 1, \quad q^i (q^i)^T = \langle q^i, q^i \rangle = \|q^i\|^2 = 1, \quad (275)$$

wobei q_i hier die Spalten, q^i die Zeilen von Q bezeichnen sollen. Wenn eine Matrix also Zeilen oder Spalten aufweist, die nicht die Länge (Norm) Eins haben, so kann es sich nicht um eine Kongruenztransformation handeln.

Die Elemente der Gleichung ausserhalb der Diagonalen geben die Orthogonalität der Vektoren wieder,

$$q_i^T q_j = \langle q_i, q_j \rangle = 0, \quad q^i (q^j)^T = \langle q^i, q^j \rangle = 0, \quad (276)$$

wobei $i \neq j$ gilt. Also gibt jede Matrix, welche spalten- oder zeilenweise aus den Vektoren einer ONB des gesamten Vektorraumes aufgebaut ist, eine orthogonale Matrix.

Z. B. ist so die Matrix

$$G = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \quad (277)$$

eine orthogonale Matrix, da die Zeilen und Spalten senkrecht aufeinander stehen und auf Eins normiert sind, m.a.W.⁵, eine ONB des Raumes \mathbb{R}^2 darstellen. Genauso ist die Matrix

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad (278)$$

eine orthogonale Matrix. Die Matrix G (der Buchstabe G wurde wegen der Verwendung solcher Matrizen durch Wallace Givens gewählt) ist eine Drehung (manchmal auch eine Givens-Rotation genannt), die Matrix H (der Buchstabe H wurde wegen der Verwendung solcher Matrizen durch Alston Scott Householder gewählt) ist eine Spiegelung (oft auch Householder-Reflektion genannt). Allgemeiner sind Matrizen der Form

$$G(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}, \quad \theta \in [0, 2\pi) \quad (279)$$

Drehungen (Givens-Rotationen) des \mathbb{R}^2 um einen Winkel θ . Es gilt $G = G(\pi/4)$. Matrizen der Form

$$H(x) = E - 2 \frac{xx^T}{x^T x}, \quad x \in \mathbb{R}^n, x \neq 0, \quad (280)$$

⁵mit anderen Worten

sind Spiegelungen an der Hyperebene senkrecht zu x , da für alle $y \perp x$ gilt

$$H(x)x = \left(E - 2\frac{xx^T}{x^Tx}\right)x = x - 2\frac{xx^T}{x^Tx}x = x - 2x\frac{x^Tx}{x^Tx} = x - 2x = -x, \quad (281a)$$

$$H(x)y = \left(E - 2\frac{xx^T}{x^Tx}\right)y = y - 2\frac{xx^T}{x^Tx}y = y - 2x\frac{x^Ty}{x^Tx} = y - 0x = y, \quad (281b)$$

und somit für jede Orthonormalbasis mit $x/\|x\|$ als Element die lineare Abbildung für die Basis die Eigenschaften einer Spiegelung erfüllt. Es gilt

$$H = H(x), \quad \text{wobei} \quad x = \begin{pmatrix} -1 \\ 1 + \sqrt{2} \end{pmatrix}. \quad (282)$$

(Der Beweis sei Interessierten als Rechenübung nahegelegt, soll heissen, er ist lang und -wierig und vielleicht auch etwas -weilig.)

Man sieht schnell, dass die Transponation von Produkten die Reihenfolge der Einzelfaktoren vertauscht (sonst wären die Produkte bei rechteckigen Matrizen ja auch gar nicht definiert), also

$$(AB)^T = B^T A^T. \quad (283)$$

Damit ist aber auch sofort klar, dass das Produkt orthogonaler Matrizen wieder eine orthogonale Matrix ist, da

$$(Q_1 Q_2)^T Q_1 Q_2 = Q_2^T \underbrace{Q_1^T Q_1}_{=E} Q_2 = Q_2^T Q_2 = E. \quad (284)$$

Allgemeiner kann man sogar zeigen, dass die Menge der orthogonalen Matrizen $Q \in \mathbb{R}^{n \times n}$ bzgl. der Matrixmultiplikation eine (nichtkommutative) Gruppe $O(n)$ bildet. Gleiches gilt für die Gruppe der unitären Matrizen $U(n)$.

6.2 Multiplikationstraining

Aus der Definition der Matrix und der Matrix-Vektor-Multiplikation ist ersichtlich, dass die k te Spalte einer Matrix $A \in \mathbb{R}^{m \times n}$ durch die Multiplikation der Matrix und dem k ten Einheitsvektor des \mathbb{R}^n entsteht. Eine Präsenzaufgabe des letzten Übungszettels trainierte diese und andere Möglichkeiten, Teile der Matrix A mittels geeigneter Multiplikationen der Matrix mit (Zeilen- und Spalten-) Vektoren herauszuziehen. Dieser Teil wird für die Erkenntnis benötigt, dass die Adjungierte im Standardskalarprodukt die Transponierte ist, was man dort durch Einsetzen der Standardbasisvektoren $\{e_i\}_{i=1}^n$ einsieht.

Es wurde in der Vorlesung erzählt, wie man erkennen kann, wie die z. B. bei dem Gaußschen Eliminationsverfahren (ab hier unter Speicherung der „Multiplikatoren“, ergo, der Zahlen, welche angeben, wie oft die dem Spaltenindex entsprechende Zeile im jeweiligen Schritt von der dem Zeilenindex entsprechenden Zeile abgezogen wurde in einer Matrix L), also der LR-Zerlegung auftretenden Transformationen sich als Multiplikationen mit Matrizen realisieren lassen.

Als zusätzliches Beispiel betrachten wir hier die Multiplikation einer Matrix von links oder rechts mit einer Diagonalmatrix,

$$DA = (d_{ii}a^i)_{i=1}^n, \quad (285)$$

wobei a^i die Zeilen von A darstellen, und

$$AD = (a_i d_{ii})_{i=1}^n, \quad (286)$$

wobei a_i die Spalten von A darstellen. Beispiel:

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 2 & 3 \end{pmatrix}. \quad (287)$$

Analog permutiert die Multiplikation mit einer Permutationsmatrix von links die Zeilen, während eine Multiplikation mit einer Permutationsmatrix von rechts die Spalten permutiert. Beispiel:

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 3 & 4 & 5 \\ 0 & 1 & 2 \\ 7 & 8 & 9 \end{pmatrix}, \quad (288a)$$

$$\begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 \\ 4 & 3 & 5 \\ 8 & 7 & 9 \end{pmatrix}. \quad (288b)$$

Auch interessant sind Multiplikationen einer quadratischen Matrix $A \in \mathbb{R}^{n \times n}$ mit einer Permutationsmatrix P und ihrer Transponierten P^T der Form

$$B = P^T \cdot A \cdot P.$$

Was ist z. B. die Transformation $J^T A J$ einer 3×3 -Matrix A mit der Permutationsmatrix J definiert als

$$J = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} ? \quad (289)$$

Kann man diese als eine Art „Spiegelung“ der Matrixelemente an einer Matrixposition deuten?

6.3 Das Gaußsche Eliminationsverfahren

Wir betrachten hier kurz (in der Anleitung noch kürzer, das war unser Schwerpunktthema einer vorangehenden Anleitung) noch einmal das Gaußsche Eliminationsverfahren.

Wir zeigen zuerst exemplarisch, dass das Gaußsche Eliminationsverfahren (mit einem eventuellen Zeilen- und Spaltentausch) eine gegebene Matrix $A \in \mathbb{R}^{m \times n}$ immer auf die Gestalt

$$A_{\text{Gauß}} = \begin{pmatrix} R_k & X_{k,n-k} \\ O_{m-k,k} & O_{m-k,n-k} \end{pmatrix}, \quad (290)$$

wobei $R_k \in \mathbb{R}^{k \times k}$, $0 \leq k \leq \min\{m, n\}$, $X_{k,n-k} \in \mathbb{R}^{k \times (n-k)}$,

bringen kann. Hierbei sei die Matrix R_k eine obere Dreiecksmatrix mit Elementen ungleich Null auf der Diagonalen. Die Notation $O_{ij} \in \mathbb{R}^{i \times j}$ bezeichne einen Block aus lauter Nullen der durch die Indizes angegebenen Dimensionen.

Die Dimensionen der Matrix A sind, außer der natürlichen Voraussetzung $n, m \in \mathbb{N}$, beliebig, also gilt z. B. $m = n$, $m < n$ oder $n < m$. Die Fälle $m = k$ und/oder $n = k$ sind hierbei eingeschlossen. Im ersten Fall entfallen die beiden Null-Blöcke, im zweiten Fall entfallen die Matrix $X_{k,n-k}$ und der darunter stehende Null-Block.

Als erstes Beispiel betrachten wir eine Matrix, die bereits eine obere Dreiecksmatrix ist,

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 3 \\ 0 & 0 & 3 \end{pmatrix} = R_3 \in \mathbb{R}^3.$$

Diese hat $m = n = 3$ und die Blöcke aus Nullen und die Matrix $X_{k,n-k}$ entfallen.

Als zweites Beispiel betrachten wir die Matrix

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

Der erste Schritt der Gaußschen Elimination basiert darauf, Vielfache der ersten Zeile von den anderen Zeilen abzuziehen, so dass die Elemente unterhalb des ersten Elements in der ersten

Spalte alle gleich Null werden. Die Vielfachen speichern wir in dieser ersten Spalte, da dort nach Konstruktion Nullen stehen, welche wir uns nicht extra merken müssen,

$$\left(\begin{array}{c|cc} 1 & 1 & 1 \\ \hline 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{array}\right) \Rightarrow \tilde{A} = \left(\begin{array}{c|cc} 1 & 1 & 1 \\ \hline 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right) = \begin{pmatrix} R_1 & X_{1,2} \\ O_{3,1} & O_{3,2} \end{pmatrix}.$$

Als nächstes Beispiel betrachten wir

$$A = \begin{pmatrix} 2 & 2 & 2 \\ 4 & 4 & 4 \\ 8 & 8 & 8 \\ 6 & 6 & 6 \end{pmatrix}.$$

Der erste Schritt der Gaußschen Elimination basiert wieder darauf, Vielfache der ersten Zeile von den anderen Zeilen abzuziehen, so dass die Elemente unterhalb des ersten Elements in der ersten Spalte alle gleich Null werden,

$$\left(\begin{array}{c|cc} 2 & 2 & 2 \\ \hline 2 & 0 & 0 \\ 4 & 0 & 0 \\ 3 & 0 & 0 \end{array}\right) \Rightarrow \tilde{A} = \left(\begin{array}{c|cc} 2 & 2 & 2 \\ \hline 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right) = \begin{pmatrix} R_1 & X_{1,2} \\ O_{3,1} & O_{3,2} \end{pmatrix}.$$

Dieses Mal sieht man einen Unterschied in der ersten Spalte, es werden die Elemente in der ersten Spalte unterhalb des ersten Elementes durch das erste Element in dieser Spalte geteilt.

Das nächste Beispiel

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 3 & 3 \\ 3 & 5 & 6 \\ 4 & 7 & 9 \end{pmatrix}$$

erfordert auch noch Eliminationsschritte in der zweiten und dritten Spalte,

$$\left(\begin{array}{ccc} 1 & 1 & 1 \\ 2 & 3 & 3 \\ 3 & 5 & 6 \\ 4 & 7 & 9 \end{array}\right) \Rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & \\ \hline 2 & 1 & 1 & \\ 3 & 2 & 3 & \\ 4 & 3 & 5 & \end{array}\right) \Rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & \\ \hline 2 & 1 & 1 & \\ 3 & 2 & 1 & \\ 4 & 3 & 2 & \end{array}\right) \Rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & \\ \hline 2 & 1 & 1 & \\ 3 & 2 & 1 & \\ 4 & 3 & 2 & \end{array}\right).$$

Hier gilt jetzt

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & \\ \hline 2 & 1 & 1 & \\ 3 & 2 & 1 & \\ 4 & 3 & 2 & \end{array}\right) \Rightarrow \tilde{A} = \left(\begin{array}{ccc|c} 1 & 1 & 1 & \\ \hline 0 & 1 & 1 & \\ 0 & 0 & 1 & \\ 0 & 0 & 0 & \end{array}\right) = \begin{pmatrix} R_3 \\ O_{1,3} \end{pmatrix}.$$

Die vorangegangenen Spalten werden dabei beim Abziehen nicht mehr verändert. In der zweiten Spalte zieht man von der nach dem ersten Eliminationsschritt in der ersten Spalte entstandene neue zweite Zeile so oft von den Zeilen drei bis vier ab, dass dort Nullen entstehen. In den neuen freien Plätzen merkt man sich, wie oft man abgezogen hat. Analog verfährt man mit der dritten Spalte.

Manchmal kann es sein, dass man auch noch Zeilen- und/oder Spalten tauschen muss, um fortfahren zu können. Als Beispiel dazu

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 8 \\ 3 & 0 & 5 \\ 4 & 2 & 4 \end{pmatrix}.$$

Der erste Schritt ist analog zu dem Vorhergehenden,

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & \\ 2 & 4 & 8 & \\ 3 & 0 & 5 & \\ 4 & 2 & 4 & \end{array}\right) \Rightarrow \left(\begin{array}{ccc|c} 1 & 2 & 3 & \\ \underline{2} & 0 & 2 & \\ 3 & -6 & -4 & \\ 4 & -6 & -8 & \end{array}\right).$$

Jetzt müssen wir Zeilen tauschen, da wir ansonsten ein Vielfaches von Null suchen, welches -6 ergibt. Wir tauschen die zweite und die dritte Zeile, inklusive der gespeicherten Multiplikatoren in der ersten Spalte, damit haben wir eine weiter zerlegbare Matrix

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & \\ \underline{3} & -6 & -4 & \\ 2 & 0 & 2 & \\ 4 & -6 & -8 & \end{array}\right) \Rightarrow \left(\begin{array}{ccc|c} 1 & 2 & 3 & \\ \underline{3} & -6 & -4 & \\ 2 & 0 & 2 & \\ 4 & 1 & -4 & \end{array}\right) \Rightarrow \left(\begin{array}{ccc|c} 1 & 2 & 3 & \\ \underline{3} & -6 & -4 & \\ 2 & 0 & 2 & \\ 4 & 1 & -2 & \end{array}\right)$$

Damit gilt dann

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & \\ \underline{3} & -6 & -4 & \\ 2 & 0 & 2 & \\ 4 & 1 & -2 & \end{array}\right) \Rightarrow \tilde{A} = \left(\begin{array}{ccc|c} 1 & 2 & 3 & \\ 0 & -6 & -4 & \\ 0 & 0 & 2 & \\ 0 & 0 & 0 & \end{array}\right) = \begin{pmatrix} R_3 \\ O_{1,3} \end{pmatrix}.$$

Hat man eine rechte Seite b aus einem Gleichungssystem dabei, so wird diese in einer Extra-Spalte am Ende mitgeführt und auch mitbehandelt. Wenn mehrere rechte Seiten von Gleichungssystemen mit der selben Systemmatrix A vorhanden sind, schreibt man alle rechten Seiten spaltenweise in einen Block, welcher hinten angefügt wird. Seien also z. B. die zwei Gleichungssysteme

$$Ax_1 = b_1 := \begin{pmatrix} 6 \\ 14 \\ 8 \\ 10 \end{pmatrix} \quad \text{und} \quad Ax_2 = b_2 := \begin{pmatrix} -5 \\ 2 \\ -1 \\ 1 \end{pmatrix}$$

mit dem eben verwendeten A gegeben. Eine kleine Randbemerkung zu den beiden rechten Seiten: Die erste rechte Seite ist der „Mackens“-Vektor $A \cdot e$, wobei $e \in \mathbb{R}^3$ der Vektor aus lauter Einsen ist, die zweite rechte Seite steht auf allen Spalten von A senkrecht. Die erweiterte Matrix sieht dann wie folgt aus:

$$\left(\begin{array}{ccc|cc} 1 & 2 & 3 & 6 & -5 \\ 2 & 4 & 8 & 14 & 2 \\ 3 & 0 & 5 & 8 & -1 \\ 4 & 2 & 4 & 10 & 1 \end{array}\right).$$

Der erste Eliminationsschritt liefert dann

$$\left(\begin{array}{ccc|cc} 1 & 2 & 3 & 6 & -5 \\ 2 & 4 & 8 & 14 & 2 \\ 3 & 0 & 5 & 8 & -1 \\ 4 & 2 & 4 & 10 & 1 \end{array}\right) \Rightarrow \left(\begin{array}{ccc|cc} 1 & 2 & 3 & 6 & -5 \\ \underline{2} & 0 & 2 & 2 & 12 \\ 3 & -6 & -4 & -10 & 14 \\ 4 & -6 & -8 & -14 & 21 \end{array}\right).$$

Der Tausch inklusive weiterer Eliminationen liefert dann

$$\begin{aligned} \left(\begin{array}{ccc|cc} 1 & 2 & 3 & 6 & -5 \\ \underline{3} & -6 & -4 & -10 & 14 \\ 2 & 0 & 2 & 2 & 12 \\ 4 & -6 & -8 & -14 & 21 \end{array}\right) &\Rightarrow \left(\begin{array}{ccc|cc} 1 & 2 & 3 & 6 & -5 \\ \underline{3} & -6 & -4 & -10 & 14 \\ 2 & 0 & 2 & 2 & 12 \\ 4 & 1 & -4 & -4 & 7 \end{array}\right) \\ &\Rightarrow \left(\begin{array}{ccc|cc} 1 & 2 & 3 & 6 & -5 \\ \underline{3} & -6 & -4 & -10 & 14 \\ 2 & 0 & 2 & 2 & 12 \\ 4 & 1 & -2 & 0 & 31 \end{array}\right). \end{aligned}$$

Nach der erfolgreichen Gauß-Elimination haben wir in kompakterer Schreibweise die beiden äquivalenten Gleichungssysteme

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -6 & -4 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} (x_1 \ x_2) = \begin{pmatrix} 6 & -5 \\ -10 & 14 \\ 2 & 12 \\ 0 & 31 \end{pmatrix}.$$

Das zweite Gleichungssystem

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -6 & -4 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} x_2 = \begin{pmatrix} -5 \\ 14 \\ 12 \\ 31 \end{pmatrix}$$

hat aufgrund der unerfüllbaren Forderung

$$0 = (0 \ 0 \ 0) x_2 = 31$$

in der letzten Zeile natürlich keine Lösung.

Die letzte Zeile des ersten Gleichungssystems

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -6 & -4 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} x_1 = \begin{pmatrix} 6 \\ -10 \\ 2 \\ 0 \end{pmatrix}$$

is hingegen trivialerweise erfüllt. Die Lösung des Gleichungssystems bestimmt sich aus dem „Rückwärtsauflösen“ des vereinfachten äquivalenten Gleichungssystems

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -6 & -4 \\ 0 & 0 & 2 \end{pmatrix} x_1 = \begin{pmatrix} 6 \\ -10 \\ 2 \end{pmatrix},$$

indem von unten nach oben die Zeilen verwendet werden, um die Komponenten von

$$x_1 = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix}$$

der Reihe nach zu finden. Die letzte Zeile liefert

$$2\xi_3 = 2 \quad \Leftrightarrow \quad \xi_3 = 1.$$

Einsetzen des soeben erzielten Ergebnisses in die zweite Zeile liefert

$$-6\xi_2 - 4\xi_3 = -6\xi_2 - 4 = -10 \quad \Leftrightarrow \quad \xi_2 = 1.$$

Unter Verwendung aller bisher erzielten Resultate gilt

$$1\xi_1 + 2\xi_2 + 3\xi_3 = \xi_1 + 2 + 3 = 6 \quad \Leftrightarrow \quad \xi_1 = 1,$$

also haben wir die eindeutige Lösung unseres Gleichungssystems gefunden.

Ein Beispiel, in dem auch ein Tausch von Spalten nötig ist, begegnet uns in den Übungsaufgaben.

Bisher haben wir nur Gleichungssysteme betrachtet, in denen die Matrix $X_{k,n-k}$ nicht weiter verwendet wurde. Als letztes Beispiel daher die Lösung eines lösbaren Gleichungssystems,

welches aber nicht eindeutig zu lösen ist: Eine Ebenengleichung im \mathbb{R}^3 . Sei das Gleichungssystem gegeben als

$$\langle n, x \rangle = 3x_1 - 6x_2 - 9x_3 = 3.$$

Die Matrix hat die Gestalt

$$n^T = (3 \mid -6 \quad -9) = (R_1 \quad X_{1,2}).$$

Eine Lösung des Gleichungssystems, die gesuchte spezielle Lösung, bestimmt man nur unter Verwendung von R_k , hier also $R_1 = 3$. Die anderen Komponenten des Lösungsvektors werden auf Null gesetzt, die entsprechenden Spalten der Linearkombination Ax (die Operation Matrix-mal-Vektor ist definiert als die Linearkombination der Spalten von A mit den in x enthaltenen Zahlen als Koeffizienten der Linearkombination) werden also nicht berücksichtigt. Das solcherart „eingeschränkte“ Gleichungssystem lautet

$$(3 \quad -6 \quad -9) \begin{pmatrix} x_1 \\ 0 \\ 0 \end{pmatrix} = 3x_1 = 3$$

und hat die eindeutige Lösung

$$x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Um jetzt linear unabhängige Lösungen des homogenen Gleichungssystems

$$(3 \quad -6 \quad -9) \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = 0$$

zu finden, welche dann den Unterraum zur linearen Mannigfaltigkeit liefert, welcher die Lösungsgesamtheit des inhomogenen Gleichungssystems beschreibt, setzt man die hinteren beiden Komponenten, also die, die zu den bis dato weggelassenen Spalten gehören, auf Vektoren einer Basis des $\mathbb{R}^{n-k} = \mathbb{R}^2$. Meist verwendet man (und wir verwenden hier) die Standard-Basis-Vektoren e_1, \dots, e_{n-k} des \mathbb{R}^{n-k} . Wir suchen also $y_a, y_b \in \mathbb{R}^3$, welche gegeben sind in der Form

$$y_a = \begin{pmatrix} a \\ 1 \\ 0 \end{pmatrix}, \quad y_b = \begin{pmatrix} b \\ 0 \\ 1 \end{pmatrix}.$$

Einsetzen in das homogene Gleichungssystem und Umformen liefert die eindeutigen Lösungen

$$(3 \quad -6 \quad -9) \begin{pmatrix} a \\ 1 \\ 0 \end{pmatrix} = 0 \Leftrightarrow 3a = 6 \Leftrightarrow a = 2$$

und

$$(3 \quad -6 \quad -9) \begin{pmatrix} b \\ 0 \\ 1 \end{pmatrix} = 0 \Leftrightarrow 3b = 9 \Leftrightarrow b = 3,$$

also die linear unabhängigen Vektoren

$$y_a = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad y_b = \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix}.$$

Beide lösen das homogene lineare Gleichungssystem. Die allgemeine Lösung des inhomogenen Gleichungssystems lautet damit

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \text{span} \left\{ \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} \right\},$$

so ist z. B.

$$x_{\text{anders}} := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 1 \\ 1 \end{pmatrix}$$

eine Lösung, wovon man sich schnell anhand von

$$(3 \quad -6 \quad -9) \begin{pmatrix} 6 \\ 1 \\ 1 \end{pmatrix} = 18 - 6 - 9 = 18 - 15 = 3$$

selber überzeugen kann.

Das nächste Beispiel ist ein Standardbeispiel, welches zeigt, dass bei einer Gaußschen Elimination die Elemente exponentiell wachsen können. Dieses gilt sogar, falls Zeilen getauscht werden, wenn ein größeres Element in der jeweils behandelten Spalte unterhalb der Diagonalen existiert. Es sei $A_n \in \mathbb{R}^{n \times n}$ gegeben als eine quadratische Matrix mit Einsen auf der Diagonalen und in der letzten Spalte, -1 sen im strikten unteren Dreieck und Nullen sonst, also

$$A_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 1 \\ -1 & 1 & 0 & \cdots & 0 & 1 \\ -1 & -1 & 1 & \cdots & 0 & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -1 & -1 & -1 & \cdots & 1 & 1 \\ -1 & -1 & -1 & \cdots & -1 & 1 \end{pmatrix}.$$

Als kleines Beispiel hier einmal A_3 ,

$$A_3 = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix}.$$

Zum besseren Verständnis sind die Einsen in der Diagonalen in blau oder blaugrün dargestellt, die Einsen in der letzten Spalte in grün oder blaugrün, die Elemente gleich -1 im strikten unteren Dreieck in rot und die übrig bleibende Null in schwarz.

Die durch das Gaußsche Eliminationsverfahren erzeugte obere Dreiecksmatrix (inklusive Speicherung der Multiplikatoren) lautet, wie man sich schnell überzeugt,

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & 1 \\ -1 & 1 & 0 & \cdots & 0 & 2 \\ -1 & -1 & 1 & \cdots & 0 & 4 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -1 & -1 & -1 & \cdots & 1 & 2^{n-2} \\ -1 & -1 & -1 & \cdots & -1 & 2^{n-1} \end{pmatrix}.$$

So hat die Matrix A_3 nach der Elimination (inklusive der Speicherung der Multiplikatoren) die Gestalt

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

Bei einer Matrix der Dimensionen $n \times n$ können die Elemente also um einen Faktor 2^{n-1} verstärkt werden. Im gegebenen Beispiel sind die Matrixelemente $-1, 0, 1$ alle vom Betrag kleiner gleich Eins und das größte Element unter der erzeugten oberen Dreiecksmatrix R_n und den Multiplikatoren ist 2^{n-1} . Dieses Verhalten nennt man exponentielles Wachstum. Solcherartes exponentielles

Wachstum ist unerwünscht, denn meist wird die Durchführung des Gaußschen Eliminationsverfahrens dem Rechner überlassen, welcher die Zahlen in ein aus Speicherplatzgründen beschränktes Format übertragen muss. Dadurch ist der *relative* Fehler meist in der Größenordnung von 10^{-16} , also ist der Fehler in der Lösung eines solchen Gleichungssystems, skaliert, so dass die Elemente der Matrix, und der rechten Seite ungefähr vom Betrage Eins sind, mit einem exponentiellen „Wachstumsfaktor“ in der absoluten Größenordnung von $10^{-16} \cdot 2^{n-1}$ behaftet und daher eventuell unbrauchbar.

6.4 Die LR-Zerlegung

Die LR-Zerlegung einer Matrix gewinnt man, indem man die bei der Elimination nach Gauß die Multiplikatoren, mit denen man die erste Zeile malnimmt, in einer unteren Dreiecksmatrix L mit Einsen auf der Diagonalen speichert. Man unterscheidet die (einfache) LR-Zerlegung, die LR-Zerlegung mit partieller Pivotisierung (suche in der aktuellen verbleibenden Spalte das betragsgrößte Element; auch Spaltenpivotsuche genannt) und die LR-Zerlegung mit vollständiger Pivotisierung (suche in der verbleibenden quadratischen Restmatrix das betragsgrößte Element).

Ein Beispiel zur einfachen LR-Zerlegung ohne Pivotisierung: Gegeben sei die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 6 & 8 \\ 1 & 6 & 14 \end{pmatrix}. \quad (291)$$

Die Elimination nach Gauß nimmt als erste Operationen die zweite minus der ersten Zeile und die dritte minus der ersten Zeile, die Multiplikatoren sind beide Male gleich Eins. Damit hat die Matrix L_1 die Gestalt

$$L_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \Rightarrow L_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \quad (292)$$

und die entstehende Matrix A_1 , in der bereits die erste Spalte auf obere Dreiecksgestalt gebracht ist, die Gestalt

$$A_1 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 4 & 11 \end{pmatrix} = L_1^{-1} A. \quad (293)$$

Als nächstes wird von der entstandenen dritten Zeile die zweite Zeile abgezogen, der Multiplikator ist wiederum gleich Eins, also ist die Matrix L_2 gegeben als

$$L_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \Rightarrow L_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}, \quad (294)$$

und die entstehende Matrix $A_2 = R$ hat die Gestalt

$$A_2 = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix} = R = L_2^{-1} A_1 = L_2^{-1} L_1^{-1} A. \quad (295)$$

Die Matrix L ist gegeben als

$$L = L_1 L_2 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \Rightarrow L^{-1} = L_2^{-1} L_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}, \quad (296)$$

und wir haben A zerlegt in ein Produkt zweier Dreiecksmatrizen L und R ,

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 6 & 8 \\ 1 & 6 & 14 \end{pmatrix} = LR = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}, \quad (297)$$

also die LR-Zerlegung von A berechnet.

Um mal eine andere LR-Zerlegung zu sehen, die noch keine Pivotisierung benötigt, aber wo die Multiplikatoren nicht alle gleich Eins sind, werden wir jetzt die Matrix

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 4 \\ 1 & 5 & 8 \end{pmatrix} \quad (298)$$

LR-zerlegen. Als erstes wird die erste Zeile zweimal von der zweiten und einmal von der dritten Zeile abgezogen, die entstehende Matrix L_1 hat die Gestalt

$$L_1 = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \Rightarrow L_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \quad (299)$$

und die Matrix A_1 , bei der bereits die erste Spalte eliminiert wurde, hat die Gestalt

$$A_1 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 4 & 7 \end{pmatrix} = L_1^{-1}A. \quad (300)$$

Elimination der zweiten Spalte mittels einer zweifachen Subtraktion der zweiten Zeile von der dritten liefert als nächste „L“-Matrix L_2 die Matrix

$$L_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}, \Rightarrow L_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}, \quad (301)$$

und die vollständig auf obere Dreiecksgestalt gebrachte Matrix $A_2 = R$ ist gegeben als

$$A_2 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix} = R = L_2^{-1}A_1 = L_2^{-1}L_1^{-1}A. \quad (302)$$

Mit

$$L = L_1L_2 = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}, \Rightarrow L^{-1} = L_2^{-1}L_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 3 & -2 & 1 \end{pmatrix}, \quad (303)$$

lautet die LR-Zerlegung der Matrix

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 4 \\ 1 & 5 & 8 \end{pmatrix} = LR = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}. \quad (304)$$

Jetzt zu einem Beispiel einer Matrix A , die nicht mehr ohne Pivotisierung LR-zerlegbar ist, nämlich der Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (305)$$

Die LR-Zerlegung würde existieren, wenn es ein reelles Vielfaches der ersten Zeile gäbe, welches von der zweiten Zeile abgezogen, im ersten Element eine Null erzeugt. Da $1 - x \cdot 0 = 1$ für alle $x \in \mathbb{R}$ gilt, ist es also nicht möglich, diese Matrix direkt in ein Produkt $L \cdot R$ zu zerlegen.

Solche Probleme treten immer dann auf, wenn bei der Berechnung der LR-Zerlegung in einer Matrix A_j , $j = 0, 1, 2, \dots$, wobei $A_0 = A$ und die anderen A_j die im j ten Schritt entstehenden Matrizen sind, ein Nullelement im $j + 1$ ten Diagonalelement entsteht. Als weiteres Beispiel betrachten wir die Matrix

$$A = \begin{pmatrix} 2 & 2 & 2 \\ 4 & 4 & 8 \\ 6 & 9 & 9 \end{pmatrix}, \quad (306)$$

in welcher keine offensichtlichen Nullelemente in der Diagonalen stehen. Die Matrix L_1 ist schnell berechnet als

$$L_1 = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}, \Rightarrow L_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix}, \quad (307)$$

und die Matrix A_1 hat die Gestalt

$$A_1 = \begin{pmatrix} 2 & 2 & 2 \\ 0 & 0 & 4 \\ 0 & 3 & 3 \end{pmatrix} = L_1^{-1}A, \quad (308)$$

welche eine Null im zweiten Diagonalelement aufweist und daher nicht mehr direkt LR-zerlegbar ist.

Da beide Matrizen jeweils den vollen Rang (2, respektive, 3) haben, führt man die partielle Pivotsuche/Spaltenpivotsuche (*pivot* (franz.) Dreh-, Angelpunkt) ein. In dieser werden die Gleichungen j bis n (ergo, die Zeilen) getauscht, bis in der Diagonale zu dieser Spalte (daher auch der Name *Spaltenpivotsuche*) ein Element ungleich Null steht.

Jetzt wird einfach die zweite Zeile gegen die dritte Zeile getauscht, also, mit der Permutationsmatrix

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad (309)$$

von links malgenommen. Jetzt ist

$$PA_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 2 & 2 \\ 0 & 0 & 4 \\ 0 & 3 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 2 \\ 0 & 3 & 3 \\ 0 & 0 & 4 \end{pmatrix} = R \quad (310)$$

eine rechte obere Dreiecksmatrix. Wie sieht jetzt die LR-Zerlegung mit Pivotisierung aus? Wir sehen, da $A_1 = L_1^{-1}A$ gilt, dass

$$R = PA_1 = PL_1^{-1}A. \quad (311)$$

Nun ist P eine Permutationsmatrix, also eine orthogonale Matrix, und die Inverse von P ist P^T . Wir hätten gerne ein P vor dem A , aber Matrixmultiplikation ist bekanntlich nicht kommutativ. Aber aufgrund der Orthogonalität von P gilt $E = P^T P$, und demnach können wir schreiben

$$R = PA_1 = PL_1^{-1}A = PL_1^{-1}EA = PL_1^{-1}P^T PA, \quad (312a)$$

$$\Rightarrow PA = (PL_1^{-1}P^T)^{-1}R = (PL_1P^T)R. \quad (312b)$$

Wie sieht die Matrix PL_1P^T aus? Durch Nachrechnen stellt man fest, dass

$$PL_1P^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad (313a)$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 3 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \quad (313b)$$

wieder eine untere Dreiecksmatrix mit Einsen auf der Diagonalen ist, die LR-Zerlegung mit partieller Pivotisierung demnach durch

$$PA = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 2 & 2 \\ 4 & 4 & 8 \\ 6 & 9 & 9 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 2 & 2 \\ 0 & 3 & 3 \\ 0 & 0 & 4 \end{pmatrix} = LR \quad (314)$$

gegeben. Allgemein vertauscht die Pivotisierung bei der Operation $P_i L_i P_i^T$ nur die Elemente im Vektor der Multiplikatoren, also hier, die Zahlen 2 und 3 in der ersten Spalte, vergleiche mit den farbigen Hervorhebungen in den Gleichungen (313).

Wenn solch ein Element ungleich Null nicht in der im j ten Schritt bearbeiteten Spalte existiert, hat die Matrix A_j des j ten Schrittes die folgende Gestalt,

$$A_j = \begin{pmatrix} r_{11} & \cdots & r_{1j} & r_{1,j+1} & r_{1,j+2} & \cdots & r_{1n} \\ 0 & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \ddots & r_{jj} & r_{j,j+1} & r_{j,j+2} & \cdots & r_{jn} \\ 0 & \cdots & 0 & \mathbf{0} & \tilde{r}_{j+1,j+2} & \cdots & \tilde{r}_{j+1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \mathbf{0} & \tilde{r}_{n,j+2} & \cdots & \tilde{r}_{nn} \end{pmatrix}. \quad (315)$$

Damit ist bestimmt die $j+1$ te Spalte der Matrix A_j im Spann der ersten j Spalten von A_j , also sind die ersten $j+1$ Spalten linear abhängig, also ist A_j singulär. Da

$$A_j = L_j^{-1} L_{j-1}^{-1} \cdots L_1^{-1} A \Rightarrow A = L_1 L_2 \cdots L_j A_j, \quad (316)$$

ist demnach auch bereits A singulär. Wir nehmen also an, dass unsere Matrix A quadratisch ist und vollen Rang hat, also regulär ist.

Das Vertauschen von Gleichungen/Zeilen wird durch eine Multiplikation mit einer Permutationsmatrix von links realisiert, für das erste Beispiel ist klar, dass mit

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = A \quad (317)$$

die Matrix PA ,

$$PA = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2, \quad (318)$$

LR-zerlegbar ist (die LR-Zerlegung mit Pivotisierung ist $PA = E_2 \cdot E_2$).

Eine Matrix $A \in \mathbb{R}^{n \times n}$ ist eindeutig LR-zerlegbar ohne Pivotisierung, wenn alle ersten $n-1$ Hauptunterabschnittsmatrizen

$$A_1 = (a_{11}), \quad A_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad A_3 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad \dots \quad (319)$$

regulär sind. Wenn eine Hauptunterabschnittsmatrix von A nicht regulär ist, so gibt es möglicherweise noch eine LR-Zerlegung, wenn alle folgenden Hauptunterabschnittsmatrizen ebenfalls singulär sind, aber die Zerlegung als solche ist nicht mehr eindeutig, ein Beispiel sind die beiden folgenden LR-Zerlegungen von A definiert als

$$A = \begin{pmatrix} 5 & 3 & 0 \\ 5 & 3 & 3 \\ 5 & 3 & 6 \end{pmatrix} = L_1 R_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 5 & 3 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 3 \end{pmatrix} \quad (320a)$$

$$= L_2 R_2 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 5 & 3 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 6 \end{pmatrix}. \quad (320b)$$

In diesem Beispiel ist die erste Hauptunterabschnittsmatrix regulär, die beiden folgenden sind es nicht, da die ersten beiden Spalten linear abhängig sind, und damit der Rang (mindestens, hier genau) um Eins niedriger als die Anzahl der Spalten ist.

6.5 Regularität & Singularität

Laut Skript heisst eine Matrix $A \in \mathbb{R}^{m \times n}$ regulär, wenn $m = n$, also die Matrix A quadratisch ist, und wenn der (Zeilen- und Spalten-) Rang gleich n ist. Damit ist die zugehörige Abbildung bijektiv, also invertierbar. Die Umkehrabbildung ist auch linear und entspricht auch einer Matrix, welche mit A^{-1} bezeichnet wird und inverse Matrix von A heisst.

Wenn $A \in \mathbb{R}^{n \times n}$ nicht regulär ist, wird A auch singular genannt. Es gelten für reguläre Matrizen A die wichtigen Gleichungen

$$A \cdot A^{-1} = A^{-1} \cdot A = E_n.$$

Ob eine Matrix regulär ist, läßt sich allerdings auch ohne explizite Berechnung der Inversen sehen. Ein wichtiges Kriterium ist, dass Gleichungssysteme

$$Ax = b, \quad A \in \mathbb{R}^{n \times n}, \quad b \in \mathbb{R}^n \quad (321)$$

immer lösbar sind, egal, welches b auf der rechten Seite steht.

Im Gauß-Algorithmus, wobei immer eine partielle Pivotsuche stattfinden sollte, wird eine LR-Zerlegung der Matrix PA berechnet, wobei die Permutationsmatrix P im Laufe der LR-Zerlegung erst konstruiert wird. Im Laufe der Berechnung der LR-Zerlegung mit partieller Pivotsuche stellt sich automatisch heraus, ob die gegebene quadratische Matrix A regulär ist. Wenn nämlich die LR-Zerlegung existiert, und sämtliche Diagonalelemente der oberen Dreiecksmatrix R ungleich Null sind (die der unteren Dreiecksmatrix L sind alle gleich Eins, also ungleich Null, und die Permutationsmatrix ist immer invertierbar, da sie orthogonal ist und $PP^T = P^TP = E$ gilt), dann ist die Ausgangsmatrix regulär. Wenn die LR-Zerlegung bei/mit einer Nullspalte in der verbleibenden Restmatrix abbricht, so ist A singular.

Merken sollte man sich, dass das Produkt regulärer Matrizen immer regulär und das Produkt von mehreren Matrizen bereits bei einer auftretenden singulären Matrix immer singular ist. Damit läßt sich das oben gesagte an der Gleichung

$$A = P^T L R, \quad (322)$$

ablesen, da das Produkt von P^T (regulär), L (regulär) und R den selben Rang wie A hat, also bei Existenz der LR-Zerlegung mit partieller Pivotisierung genau dann regulär ist, wenn R es ist. Wenn keine LR-Zerlegung mit regulärem R bei partieller Pivotisierung existiert, ist nach obigen Betrachtungen A singular.