Lineare Algebra und Analytische Geometrie I

W.P. Barth, P. Knabner

Wintersemester 08/09

Department Mathematik der Universität
Erlangen, Bismarckstraße $1\frac{1}{2}$
e-mail: barth@mi.uni-erlangen.de

Version vom 14. Januar 2009

Inhaltsverzeichnis

0 Einführung

An allen deutschen Universitäten müssen die Studenten der Mathematik (und auch der Physik) in den ersten Semestern die Vorlesungen „Analysis“ und „Lineare Algebra und Analytische Geometrie“ hören. Das ist sinnvoll, weil der Stoff beider Vorlesungen für jede Art von Mathematik - sei sie rein oder angewandt - fundamental ist. Der Stoff der Analysis-Vorlesungen ist Differentiation und Integration, das kennt man im Prinzip aus dem Gymnasium. Diese Art von Mathematik wurde um das Jahr 1700 herum, im wesentlichen von I. Newton und G.W. Leibniz erfunden. Die Entwicklung der Analysis verlief Hand in Hand mit der Entwicklung ihrer Anwendungen in Physik und Ingenieurwissenschaften. Um 1700 waren ja auch die Wissenschaften Mathematik und Physik bei weitem nicht so voneinander abgegrenzt wie heute. Die Analysis hat dann um 1900 ihren Einzug in den Mathematikstoff der Gymnasien gefunden. Sie ist eine natürlich gewachsene Disziplin, mit starker Beziehung zu Anwendungen.

Anders ist es mit der Linearen Algebra. Die gab es vor dem zweiten Weltkrieg nicht als eigenständige Vorlesung. Natürlich gab es ihre Rechentechniken (vor allem lineare Gleichungssysteme, Matrizen, Determinanten) schon viel früher. Aber man hat diese Rechentechniken damals nicht zu einem strukturierten System zusammengefasst. Das tat man erst, nachdem auch die Algebra, früher die Lehre von den Gleichungen, seit den zwanziger Jahren des letzten Jahrhunderts, zu einer Lehre von Rechenstrukturen wurde. Vor dem Krieg geschah dies vor allem unter dem Einfluss der in Erlangen geborenen Mathematikerin Emmy Noether, nach dem Krieg vor allem unter dem Einfluss einer Schule französischer Mathematiker, die sich hinter dem Pseudonym „N. Bourbaki“ verbargen.

Vorlesungen über Geometrie, sei es analytische oder andere, dagegen gibt es, seit es Mathematikvorlesungen gibt. Geometrie, schon den antiken Ägyptern und Griechen bekannt, ist zusammen mit der Algebra die älteste Teildisziplin der Mathematik. Die Analytische Geometrie verlor nach und nach ihren eigenständigen Charakter, seit Descartes die kartesischen Koordinaten erfand. Seitdem kann man geometrische Beweise auch durch Koordinatenrechnungen erbringen. Dies führte - vor allem im 19. Jahrhundert - zu einem erbitterten Gegensatz zwischen den reinen („synthetischen“) Geometern, und den Koordinatenrechnern, die die Geometrie als eine Anwendung der Algebra sahen.

Und es stellte sich eben heraus, dass genau das, was man unter „Analytischer Geometrie“ versteht, eine Anwendung der Linearen Algebra ist. So wurde dann aus der Vorlesung „Analytische Geometrie“, früher eine Standardvorlesung für Anfänger, die Vorlesung „Lineare Algebra und Analytische Geometrie“. Nach und nach traten dabei die Strukturen der Linearen Algebra in den Vordergrund, und die z.T. Jahrtausende alten geometrischen Inhalte in den Hintergrund. Das hat sich bis in die Gymnasien ausgewirkt, wo leider auch der Stellenwert der Geometrie sehr gelitten hat.

Die grundlegenden Rechenmethoden der Analysis (Grenzwertbildung, Differenzieren, Integrieren) und der Linearen Algebra (Lineare Gleichungssysteme, Vektoren, Matrizen) sind zumindest rudimentär vom Gymnasium her vertraut. Ich habe die Erfahrung gemacht, dass die Studenten im ersten Semester mit der Linearen Algebra ganz andersartige Probleme, als mit der Analysis haben. In der Linearen Algebra liegt nämlich sehr bald das Schwergewicht auf der Untersuchung mathematischer „Strukturen“. Das ist in der Geschichte der Mathematik etwas relativ Neues. Erst in den letzten 100 bis 150 Jahren schenkt man diesem Gesichtspunkt in der Mathematik Aufmerksamkeit. Und leider ist es genau das, was den Anfängern immer große Schwierigkeiten bereitet. Aber weil es eben eine der ganz wichtigen Methoden in der modernen Mathematik ist, müssen wir uns damit befassen.

Jetzt sollte eigentlich eine Liste von Lehrbüchern der Linearen Algebra kommen. Eine solche Liste zusammenzustellen ist nicht einfach: Bücher zur Linearen Algebra gibt es viele, aber keines, welches mir wirklich zusagt. Das ist auch der Grund dafür, dass ich im WS 91/92 ein Skriptum geschrieben habe, das ich damals an die Studenten ausgab. Das positive Echo auf jenes Skriptum hat mich bewogen, seitdem zu den meisten meiner Vorlesungen den Studenten ein eigenes Skriptum an die Hand zu geben. Das vorliegende Skriptum ist z.T. eine aufpolierte Version meines Skriptums von 91/92. Z.T. hat es aber durch Zusammenarbeit mit meinem Kollegen P. Knabner sehr an Inhalt gewonnen. Dabei folgen wir einer Tendenz, die in den letzten Jahrzehnten in der Mathematik stark an Bedeutung gwann. Früher gab es eine klare Unterscheidung zwischen reiner (zweckfreier) und angewandter (nützlicher) Mathematik. Diese Abgrenzung hat sich zu einem großen Teil aufgelöst. Im Jahr 2006/07 hielt P. Knabner diese Anfängervorlesung 'Lineare Algebra'. Dabei orientierte er sich wesentlich an meinem vorliegenden Skriptum, ergänzte es allerdings um Inhalte, die früher der angewandten Mathematik zugerechnet waren. Ich beabsichtige seine Ergänzungen auch in dieser Vorlesung zu besprechen. Ich kann nur hoffen, dass ich mit dem Schreiben des so modifizierten Skriptums zeitlich zurecht komme.

Trotzdem, noch eine Liste von Standardlehrbüchern:

E. Brieskorn: Lineare Algebra und analytische Geometrie 1, 2, Vieweg 1983, 85
G. Fischer: Lineare Algebra, Vieweg 1975
W. Graeub: Lineare Algebra, Springer 1958
H.J. Kowalsky: Lineare Algebra, de Gruyter 1963
F. Lorenz: Lineare Algebra I, II, BI 1981, 1982
G. Strang: Lineare Algebra, Springer 2003

Nach jedem Abschnitt finden Sie eine Anzahl von Übungsaufgaben. Einige habe ich mir selbst ausgedacht, andere aus Quellen übernommen, an die ich mich jetzt nicht mehr erinnere. Sehr viele aber stammen aus Staatsexamensklausuren: entweder aus dem früheren Vorexamen oder dem nicht-vertieften Examen. Aus diesen Aufgaben werde ich Ihre Übungsaufgaben auswählen.

Erlangen, 13.10.08

Wolf Barth

1 Der Zahlenraum $\mathbb{R}^n$

1.1 Lineare Gleichungssysteme

1.1.1 Beispiele und Spezialfälle

Lineare Gleichungssysteme sind die einzige Art von Gleichungen in der Mathematik, welche wirklich exakt lösbar sind. Wir beginnen mit einem Beispiel, wie es schon aus der Antike überliefert ist.

Beispiel 1.1 *In einem Käfig seien Hasen und Hühner. Die Anzahl der Köpfe sei insgesamt 4, die Anzahl der Beine sei insgesamt 10. Frage: Wieviele Hasen und wieviele Hühner sind es?*

Lösung: Es sei x die Anzahl der Hasen und y die Anzahl der Hühner. Dann gilt also

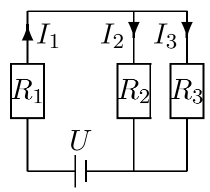
$$\begin{array}{rcrcr} x & + & y & = & 4 \\ 4x & + & 2y & = & 10 \end{array}$$

Dies ist ein System aus zwei linearen *Gleichungen in zwei Unbekannten x und y. Wir können aus der ersten Gleichung $x = 4 - y$ eliminieren und in die zweite einsetzen:*

$$\begin{aligned} 4(4-y)+2y &= 10 \\ 16-2y &= 10 \\ -2y &= -6 \\ y &= 3 \\ x &= 1 \end{aligned}$$

Antwort: Es sind drei Hühner und ein Hase.

Beispiel 1.2 *Gegeben sei ein elektrisches Netzwerk der Form*



Dabei seien U, R_1, R_2, R_3 gegeben und I_1, I_2 und I_3 gesucht. Weil es uns hier nur auf die Mathematik ankommt, vermeiden wir Fragen der Benennung. Die vorkommenden Größen $U, R_\nu, I_\nu, \nu = 1,2,3$, sind also reelle Zahlen. Und die Widerstände R_ν sind Zahlen > 0.

Lösung: Nach den sogenannten Kirchhoffschen Gesetzen der Physik hat man die Gleichungen $I_1 = I_2 + I_3$, sowie $R_2I_2 = R_3I_3$ und $R_1I_1 + R_2I_2 = U$. Das Aufstellen dieser Gleichungen heißt mathematische Modellierung *des Problems. Wir schreiben die Gleichungen als ein System aus drei linearen Gleichungen in den drei Unbekannten I_1, I_2 und I_3 :*

$$\begin{array}{rcrcrcr} I_1 & - & I_2 & - & I_3 & = & 0 \\ & & R_2I_2 & - & R_3I_3 & = & 0 \\ R_1I_1 & + & R_2I_2 & & & = & U \end{array}$$

Wir können hier etwa $I_1 = I_2 + I_3$ eliminieren, um folgendes System aus zwei linearen Gleichungen in den Unbekannten I_2 und I_3 zu erhalten:

$$\begin{array}{rcrcl} R_2 I_2 & - & R_3 I_3 & = & 0 \\ (R_1 + R_2) I_2 & + & R_1 I_3 & = & U \end{array}$$

Hier eliminieren wir $I_2 = \frac{R_3}{R_2} I_3$ (das geht so, weil $R_2 \neq 0$) und erhalten schließlich die Gleichung

$$\begin{array}{rcl} (R_1 + R_2)\dfrac{R_3}{R_2} I_3 + R_1 I_3 & = & U, \\ (R_1 R_2 + R_1 R_3 + R_2 R_3) I_3 & = & R_2 U. \end{array}$$

Weil $R_1 R_2 + R_1 R_3 + R_2 R_3 > 0$ ist, können wir hieraus

$$I_3 = \frac{R_2 U}{R_1 R_2 + R_1 R_3 + R_2 R_3}$$

berechnen. Aus den Eliminationsgleichungen für I_2 und I_1 erhalten wir

$$I_2 = \frac{R_3 U}{R_1 R_2 + R_1 R_3 + R_2 R_3}, \qquad I_1 = \frac{(R_2 + R_3) U}{R_1 R_2 + R_1 R_3 + R_2 R_3}.$$

Definition 1.1 *Eine lineare Gleichung ist eine Gleichung der Art*

$$a_1 x_1 + a_2 x_2 + \ldots + a_n x_n = b.$$

Dabei sind die Koeffizienten $a_1, \ldots, a_n$ und die rechte Seite b gegeben, und die Unbekannten $x_1, \ldots, x_n$ sind gesucht. Ein lineares Gleichungssystem *kurz LGS ist ein System*

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

aus mehreren linearen Gleichungen. Dabei sind die Koeffizienten $a_{\mu,\nu}$ und die rechten Seiten b_μ für $\mu = 1, \ldots, m, \nu = 1, \ldots, n$ gegeben. Die Unbekannten x_ν, $\nu = 1, \ldots, n$, sind gesucht.

Zunächst sollen alle Koeffizienten und auch die Unbekannten reelle Zahlen sein. Die Menge der reellen Zahlen bezeichnen wir wie üblich mit $\mathbb{R}$. Wir benutzen die Rechenstrukturen

$$\text{Addition: } a + b, \qquad \text{Multiplikation: } a \cdot b \text{ bzw. kurz } ab \qquad (a, b \in \mathbb{R})$$

wie in der Schule, ohne sie hier weiter zu analysieren oder zu problematisieren. Zahlen $x_1, \ldots, x_n \in \mathbb{R}$, wie sie etwa zu einer Lösung des obigen LGS gehören, fassen wir zusammen zu einem Spaltenvektor

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (x_\nu)_{\nu=1,\ldots,n}.$$

Die Menge aller solcher Spaltenvektoren nennt man den *Zahlenraum* $\mathbb{R}^n$.

Definition 1.2 *Die* Koeffizientenmatrix *des LGS aus Definition 1.1 ist das rechteckige Zahlenschema*

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

Der Zeilenvektor $\mathbf{a}_\mu = (a_{\mu,1}, ..., a_{\mu,n})$, $\mu = 1, ..., m$, *heißt die* μ*-te Zeile der Matrix* A. *Der Spaltenvektor*

$$\mathbf{a}^\nu = \begin{pmatrix} a_{1,\nu} \\ \vdots \\ a_{m,\nu} \end{pmatrix}, \quad \nu = 1, ..., n,$$

heißt die ν*-te Spalte der Matrix* A.

Wenn wir an die Matrix A *noch die rechten Seiten des LGS anfügen,*

$$A = \left(\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \dots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} & b_m \end{array} \right),$$

so nennen wir dies die erweiterte Koeffiizentenmatrix.

Beispiel 1.1 ist so ziemlich die einfachste Sorte eines linearen Gleichungssytems. Beispiel 1.2 gibt aber einen ersten Eindruck davon, wie lineare Gleichungssysteme Fragen aus Naturwissenschaften und Technik, aber auch aus der Ökonomie modellieren. Schon deswegen ist es wichtig, sie mathematisch zu untersuchen. Dabei stellen sich zwei wesentliche mathematische Fragen:

A) Das Existenzproblem: Hat ein vorgelegtes LGS (mindestens) eine Lösung? Diese Frage kann man positiv entscheiden durch:

a) Konkrete Angabe einer Lösung. Das geht allerdings nur bei einem konkreten Beispiel, nicht Diskussion einer allgemeinen Situation. Es bleibt dann auch die Frage, woher eine solche Lösung kommt.

b) Abstrakte Argumentation, z.B., durch einen Widerspruchsbeweis. Aus der Annahme, es gebe keine Lösung folgert man logisch einen Widerspruch. Eine Lösung wird dadurch aber nicht bekannt.

c) Angabe, bzw. Herleitung eines Algorithmus (= Rechenvorschrift) zur Bestimmung einer Lösung. Wenn dies nur endlich viele Rechenschritte erfordert, dann erhält man bei (exakter) Durchführung des Algorithmus eine (exakte) Lösung.

B) Das Eindeutigkeitsproblem: Ist die Lösung des vorgelegten LGS eindeutig bestimmt? Das heißt auf deutsch: Wenn x und y Lösungen sind, gilt dann $x = y$? Dies ist nur durch abstrakte Argumentation zu klären.

Die Fragen A) und B) sind i.A. unabhängig voneinander. Wenn beide positiv zu beantworten sind, dann sagt man: Es gibt *genau eine* Lösung.

Ein LGS, das aus den Anwendungen kommt, ist typischerweise sehr groß, es enthält etwa 10^3 bis 10^8 Unbekannte bzw. Gleichungen. Das Rechnen mit der Hand, wie in den obigen Beispielen, ist nicht mehr möglich. Die Frage nach (effizienten) Algorithmen ist deswegen besonders wichtig. Wir wollen diese Frage hier so weit wie möglich mitbehandeln. Vertieft wird sie in der Numerik-Vorlesung. Wir diskutieren also jetzt den Allgemeinfall eines LGS, wobei wir besonders darauf achten müssen, welche Spezialfälle und Ausnahmen auftreten können:

Spezialfall 1: Eine Gleichung. Bei der Behandlung einer linearen Gleichung

$$a_1x_1 + a_2x_2 + ... + a_nx_n = b,$$

müssen wir leider verschiedene Fälle unterscheiden:

A: Nicht alle Koeffizienten $a_1, ..., a_n$ sind $= 0$. Dann sei etwa $a_m, 1 \leq m \leq n$, der erste von 0 verschiedene Koeffizient. Die Gleichung sieht so aus:

$$0 \cdot x_1 + ... + 0 \cdot x_{m-1} + a_m \cdot x_m + a_{m+1} \cdot x_{m+1} + ... + a_n \cdot x_n = b.$$

Wir können also $x_1, ..., x_{m-1}$ beliebig wählen, auf die Gültigkeit der Gleichung hat dies keinen Einfluß. Ebenso können wir $x_{m+1}, ..., x_n$ beliebig wählen. Anschließend setzen wir

$$x_m := (b - a_{m+1}x_{m+1} - ... - a_nx_n)/a_m.$$

Damit haben wir für jede Wahl der $x_1, ..., x_{m-1}, x_{m+1}, ..., x_n$ die Gleichung gelöst. Dies ist auf diese Weise nur möglich, weil $a_m \neq 0$. Es fällt auf, dass hier $n-1$ Unbekannte frei wählbar sind. Die Zahl dieser 'Freiheitsgrade' zu präzisieren, ist Teil der zu entwickelnden Theorie.

B1: Alle Koeffizienten $a_1, ..., a_n$ sind 0, aber es ist $b \neq 0$. Das Gleichungssystem hat dann die merkwürdige Form

$$0 \cdot x_1 + ... + 0 \cdot x_n = b.$$

Egal, wie man auch die Unbekannten $x_1, ..., x_n$ wählt, diese Gleichung ist nie zu erfüllen. Sie ist *unlösbar.*

B2: Alle Koeffizienten $a_1, ..., a_n$ sind $= 0$ und auch $b = 0$. In diesem reichlich uninteressanten Fall ist die Gleichung stets erfüllt, sie stellt keinerlei Bedingungen an die Unbekannten.

Spezialfall 2: Diagonalsystem. Hier hat die Koeffizientenmatrix eine Form

$$A = \begin{pmatrix} a_{1,1} & 0 & ... & ... & ... & 0 \\ 0 & \ddots & & & & \vdots \\ \vdots & & a_{r,r} & & & \vdots \\ \vdots & & & 0 & & \vdots \\ \vdots & & & & \ddots & \vdots \\ 0 & ... & ... & ... & ... & \end{pmatrix}$$

mit $a_{\mu,\nu} = 0$ für $\mu \neq \nu$. Die Diagonalelemente $a_{\mu,\mu}$ sollen so angeordnet sein, dass $a_{\mu,\mu} \neq 0$ für $\mu \leq r$ und $a_{\mu,\mu} = 0$ für $\mu > r$.

Falls $r < m$ ist, haben wir hier Nullzeilen $\mathbf{a}_\mu = (0, ..., 0)$. Das ist z.B. immer dann der Fall, wenn $n < m$. Wenn zu einer solchen Nullzeile die rechte Seite $b_\mu \neq 0$ ist, dann ist die Gleichung in dieser Zeile unlösbar (B1). Dann ist auch das ganze System unlösbar. (Sowas gibt's!) Falls für alle Nullzeilen die rechten Seiten $b_\mu = 0$ sind, dann enthalten diese Zeilen keine Aussage, und wir brauchen sie nicht zu beachten (B2).

Die Zeilen $\mu = 1, ..., r$ legen die Unbekannten x_μ fest durch

$$x_\mu = \frac{b_\mu}{a_{\mu,\mu}}, \quad \mu = 1, ..., r.$$

Die weiteren Unbekannten $x_{r+1}, ..., x_n$ sind frei wählbar, falls nicht der unlösbare Fall vorliegt.

Eigentlich ist dieses System kein richtiges System, weil jede Unbekannte nur in einer Zeile mit einem Koeffizienten $\neq 0$ vorkommt. Die Unbekannten sind durch die Gleichungen nicht miteinander gekoppelt.

Spezialfall 3: Staffelsystem. Hier hat die Koeffizientenmatrix eine Form

$$\begin{pmatrix} a_{1,1} & ... & ... & ... & a_{1,n} \\ 0 & \ddots & & & \vdots \\ \vdots & \ddots & a_{r,r} & ... & a_{r,n} \\ \vdots & & 0 & ... & 0 \\ \vdots & & & & \vdots \\ 0 & ... & ... & ... & 0 \end{pmatrix},$$

wo

- das untere Dreieck der Matrix verschwindet, d.h., für $\mu > \nu$ ist $a_{\mu,\nu} = 0$,
- es für $\mu > r$ nur Nullzeilen gibt, d.h., für $\mu > r$ ist $a_{\mu,\nu} = 0$,
- die Diagonaleinträge $a_{\mu,\mu} \neq 0$ sind für $\mu = 1, ..., r$.

Wieder entscheiden die rechten Seiten b_μ, $\mu = r+1, ..., n$, darüber, ob das System lösbar ist oder nicht. Im lösbaren Fall $b_{r+1} = ... = b_m = 0$ sind die letzten $m - r$ Zeilen aussagelos und die Unbekannten $x_{r+1}, ..., x_n$ frei wählbar. Wegen $a_{r,r} \neq 0$ ist die r-te Gleichung nach x_r auflösbar vermöge

$$x_r = \frac{1}{a_{r,r}} \left(b_r - \sum_{\nu=r+1}^{n} a_{\mu,\nu} x_\nu \right).$$

Mit diesem bekannten x_r kann dann x_{r-1} aus der $r-1$-ten Zeile bestimmt werden usw. Dieses Verfahren nennt man **Rückwärtssubstitution**. Für $\mu = r, r-1, ..., 1$ erhält man also

$$\boxed{x_\mu = \frac{1}{a_{\mu,\mu}} \left(b_\mu - \sum_{\nu=\mu+1}^{n} a_{\mu,\nu} x_\nu \right).}$$

Für $r = n$, den Fall ohne frei wählbare Unbekannte ist hier bei $\mu = r$ die Regel

$$\sum_{n+1}^{n} ... = 0 \quad \text{(leerer Indexbereich)}$$

zu beachten.

Bei einigen Unterfällen lässt sich Genaueres über die Lösungsmenge sagen:

Spezialfall 3a: Wenn $r = n$ ist (und notwendigerweise $m \geq n$), sowie $b_\mu = 0$ ist für $\mu > n$, dann ist das System lösbar. Aber keine der Unbekannten ist frei wählbar. Die Lösung ist eindeutig bestimmt.

Spezialfall 3b: Wenn $m > r$ ist und ein $b_\mu \neq 0$ für $\mu > r$, dann ist das System unlösbar.

Spezialfall 4: Zeilenstufenform. Die Koeffizientenmatrix hat eine Art zerpflückter Staffelform

$$\begin{pmatrix} \overbrace{0...0}^{n_0} & \# & \overbrace{*...*}^{n_1} & * & ... & * & \overbrace{*...*}^{n_r} \\ . & 0 & 0...0 & \# & ... & * & *...* \\ & & & & \ddots & & \\ . & . & . & & 0...0 & \# & *...* \\ . & . & . & . & . & 0 & 0...0 \\ 0...0 & 0 & 0...0 & 0 & 0...0 & 0 & 0...0 \end{pmatrix}$$

Dabei bezeichnet # Koeffizienten $\neq 0$ und $*$ beliebige Koeffizienten. Die Stufenlängen $n_0, n_1, ..., n_r$ können eventuell $= 0$ sein, und r mit $1 \leq r \leq m, n$, die Anzahl der Stufen kann mit der Gesamtzahl m aller Zeilen übereinstimmen, sodass also keine Nullzeilen am unteren Ende der Matrix auftreten.

Für $\mu = 1, ..., m$ definieren wir den Index $j(\mu)$ durch

$$j(\mu) := \begin{cases} \min\limits_{\nu}\{a_{\mu,\nu} \neq 0\} & \text{falls} \quad \mu \leq r \\ n+1 & \text{falls} \quad \mu > r \end{cases}$$

Für $\mu = 1, ..., r$ ist also

$$a_{\mu,\nu} = 0 \text{ wenn } \nu \leq j(\mu) - 1, \quad a_{\mu,j(\mu)} \neq 0, \quad \text{sowie} \quad j(1) < j(2) < ... < j(r).$$

Die Stufenlängen sind

$$n_0 = j(1) - 1, \quad n_i = j(i+1) - j(i) - 1 \text{ für } i = 1, ..., r.$$

Falls $b_{r+1} = ... = b_m = 0$, ist das System lösbar, und auch hier lässt sich die Lösungsgesamtheit angeben:

Wir beginnen in der letzten Zeile mit der letzten Unbekannten. Entsprechend der Länge n_r der letzten Stufe sind die n_r Unbekannten $x_n, ..., x_{j(r)+1}$ frei wählbar. Zur Verdeutlichung nennen wir diese frei wählbaren Komponenten des Lösungsvektors *Parameter* und bezeichnen sie mit λ_ν:

$$\begin{array}{rcl} x_n & := & \lambda_n \\ & \vdots & \\ x_{j(r)+1} & := & \lambda_{n_r} \end{array} \qquad \lambda_\nu \in \mathbb{R}$$

Aber bei $x_{j(r)}$ steht ein Koeffizient # der $\neq 0$ ist. Deswegen ist diese Unbekannte durch die r-te Zeile des Gleichungssystems und durch die bereits gewählten Parameter eindeutig bestimmt. Weiter sind die Parameter

$$\begin{array}{rcl} x_{j(r)-1} & := & \lambda_{n_r+1} \\ & \vdots & \\ x_{j(r-1)+1} & := & \lambda_{n_r+n_{r-1}} \end{array} \qquad \lambda_\nu \in \mathbb{R}$$

frei wählbar. Und $x_{j(r-1)}$ ist wieder durch das Gleichungssystem und die bisher gewählten Parameter eindeutig bestimmt.

Dieses Verfahren kann man iterieren, so dass man also nach r Schritten eine Darstellung aller Lösungen mit Parametern

$$\lambda_1, ..., \lambda_{\bar{n}} \quad \text{mit} \quad \bar{n} = \sum_{\mu=1}^{r} n_i = n - r.$$

Der Spezialfall 3 ist hierin enthalten als

$$n_0 = ... = n_{r-1} = 0, \quad n_r = n - r = \bar{n}.$$

1.1.2 Das Eliminationsverfahren

Jedes LGS System kann man mit dem Eliminationsverfahren behandeln, so, wie wir es an den obigen einfachen Beispielen 1.1 und 1.2 gesehen haben. Wir beschreiben diese Elimination jetzt in einer etwas formaleren Weise, um die Übersicht nicht zu verlieren.

Wenn alle Koeffizienten $a_{1,1}, ..., a_{m,1}$ in der ersten Spalte 0 sind, stellt das System keine Bedingung an die Unbekannte x_1. Diese ist völlig frei wählbar und auf die erste Spalte des Systems kommt es nicht an. Ist aber einer der Koeffizienten $a_{1,1}, ..., a_{m,1}$ aus der ersten Spalte $\neq 0$, so sei etwa $a_{p,1}$ davon der erste. Wir *vertauschen* die erste und die $p-$te Zeile. Dabei ändern sich die Lösungen des Systems nicht. Aber danach haben wir $a_{1,1} \neq 0$. Deswegen können wir die erste Zeile *durch* $a_{1,1}$ *dividieren* und wieder ändern sich die Lösungen nicht. Dann sieht die erste Zeile so aus:

$$x_1 + \frac{a_{1,2}}{a_{1,1}} x_2 + ... + \frac{a_{1,n}}{a_{1,1}} = \frac{b_1}{a_{1,1}}.$$

Wir eliminieren nun x_1, allerdings ohne die Eliminationsgleichung explizit hinzuschreiben, aus den restlichen Gleichungen, indem wir von der zweiten, ..., m-ten Zeile $a_{2,1}$ mal, ..., $a_{m,1}$ mal die erste Zeile *subtrahieren.* Da wir dies, wenn wir wollen, auch wieder rückgängig machen können, ändern sich auch hier die Lösungen nicht, und unser Gleichungssystem nimmt die Form

$$\begin{array}{ccccccccc} x_1 & + & a'_{1,2}x_2 & + & ... & + & a'_{1,n}x_n & = & b'_1 \\ & & a'_{2,2}x_2 & + & ... & + & a'_{2,n}x_n & = & b'_2 \\ & & \vdots & & & & \vdots & & \vdots \\ & & a'_{m,2}x_2 & + & ... & + & a'_{m,n}x_n & = & b'_m \end{array}$$

an, mit neuen Koeffizienten $a'_{1,2}, ..., a'_{m,n}$ und neuen rechten Seiten $b'_1, ..., b'_m$. Jetzt kommt es nur noch darauf an, die letzten $m-1$ Gleichungen aufzulösen. Gelingt dies, so setzen wir deren Lösungen $x_2, ..., x_n$ in die erste Gleichung ein und berechnen daraus x_1.

Beispiel 1.3 *Zum Nachweis, dass es sich bei diesem Verfahren tatsächlich nur um Elimination handelt betrachten wir noch einmal das Hasen und Hühner-Beispiel 1.1 Mit der erweiterten Koeffizientenmatrix*

$$\left(\begin{array}{cc|c} 1 & 1 & 4 \\ 4 & 2 & 10 \end{array} \right) \quad II - 4 \cdot I$$

Wenn wir hier das 4*-fache der ersten Zeile von der zweiten subtrahieren, erhalten wir die umgeformte Matrix*

$$\left(\begin{array}{cc|c} 1 & 1 & 4 \\ 0 & -2 & -6 \end{array} \right).$$

Deren zweite Zeile kommt genauso in Beispiel 1.1 vor.

Indem wir dieses Verfahren sukzessive wiederholen, können wir das Gleichungssystem lösen, außer wir gelangen irgend wann zu einer unlösbaren linearen Gleichung (Fall B1 von oben).

Wegen seiner prinzipiellen Wichtigkeit müssen wir das Eliminationsverfahren noch etwas weiter formalisieren. Dazu vereinbaren wir folgenden Sprachgebrauch („Definitionen“):

Die Veränderungen, welche wir am Gleichungssystem vornahmen, ohne seine Lösungsmenge zu ändern, nennen wir *elementare Zeilenumformungen.* Es gibt drei Sorten davon:

Typ I:	Vertauschen zweier Zeilen,
Typ II:	Multiplikation einer Zeile mit einer Konstanten $c \neq 0$,
Typ III:	Addition des c-fachen einer Zeile, $c \in \mathbb{R}$ beliebig, zu einer *anderen.*

Mit diesen drei Sorten elementarer Zeilenumformungen können wir, links beginnend, eine Spalte nach der anderen leerfegen:

- Sind alle Koeffizienten in der Spalte $= 0$, so ändern wir nichts, sondern wenden uns sogleich der nächsten Spalte zu.
- Sind Koeffizienten in der Spalte $\neq 0$, davon der erste etwa in der p-ten Zeile, so vertauschen wir diese p-te Zeile mit der ersten (Umformung vom Typ I). Anschließend multiplizieren wir die erste Zeile mit dem Kehrwert dieses Koeffizienten durch (Typ II), um zu erreichen, dass in dieser ersten Zeile der erste Koeffizient 1 ist. Schließlich addieren wir ein geeignetes Vielfaches der ersten Zeile zu jeder der folgenden Zeilen (Typ III), um dort den Koeffizienten aus der ersten Spalte zu beseitigen.
- Haben wir erreicht, dass der erste Koeffizient einer Spalte $= 1$ und alle weiteren $= 0$ sind, lassen wir die erste Zeile beiseite und wenden uns der nächsten Spalte zu.

Beispiel 1.4

gegeben $\begin{pmatrix} 0 & 2 & 4 \\ 2 & 3 & 4 \end{pmatrix}$

vertauschen erste und zweite Zeile $\begin{pmatrix} 2 & 3 & 4 \\ 0 & 2 & 4 \end{pmatrix}$

Multiplikation beider Zeilen mit $\frac{1}{2}$ $\begin{pmatrix} 1 & \frac{3}{2} & 2 \\ 0 & 1 & 2 \end{pmatrix}$

Beispiel 1.5

gegeben $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \end{pmatrix}$

zweite Zeile minus zweimal erste $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & -1 \end{pmatrix}$

Multiplikation der zweiten Zeile mit -1 $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & 1 \end{pmatrix}$

Dieses Verfahren heißt *Gaußsches Eliminationsverfahren* oder kurz *Gauß-Verfahren.*

Wir fassen das Resultat unserer Matrizen-Manipulationen zusammen:

Satz 1.1 (Gauß-Elimination) *Jede Matrix lässt sich durch elementare Zeilenumformungen auf eine* Zeilenstufenform *(obiger Spezialfall 4) bringen. Dabei können die Stufenlängen $n_0, n_1, ..., n_r$ eventuell $= 0$ sein, und r, die Anzahl der Stufen kann mit der Gesamtzahl aller Zeilen übereinstimmen, sodass also keine Nullzeilen am unteren Ende der Matrix auftreten.*

Da sich bei elementaren Umformungen der Koeffizientenmatrix die Lösungen eines linearen Gleichungssystems nicht ändern, können wir diese Lösungen auch bestimmen, nachdem wir die erweiterte Koeffizientenmatrix in Zeilenstufenform vorliegen haben. Ob das System lösbar ist oder nicht, lesen wir an der letzten Stufe ab:

lösbar	unlösbar
$\left(\begin{array}{ccccccc\|c} 0...0 & 1 & ... & . & ... & . & ... & . \\ & 0 & 0...0 & 1 & ... & . & ... & . \\ & & & 0 & ... & . & ... & . \\ & & & & ... & 1 & ... & . \\ & & & & & 0 & 0...0 & 0 \end{array}\right)$	$\left(\begin{array}{ccccccc\|c} 0...0 & 1 & ... & . & ... & . & ... & . \\ & 0 & 0...0 & 1 & ... & . & ... & . \\ & & & 0 & 0...0 & . & ... & . \\ & & & & & 1 & ... & . \\ & & & & & 0 & 0...0 & 1 \\ & & & & & & ... & 0 \end{array}\right)$

In der Tat, die letzte Gleichung des rechten Systems lautet $0 \cdot x_n = 1$ und ist unlösbar. Und wenn eine Gleichung unlösbar ist, dann ist das ganze System auch nicht lösbar.

Aus einer Matrix in Zeilenstufenform kann man zu einer Matrix in Staffelform gelangen, indem man in der Matrix Spalten vertauscht. Dadurch kann man die Spalten, in denen eine Stufe steht, ganz nach links bringen. Wenn man dabei noch die zugehörigen Unbekannten entsprechend vertauscht, bleiben auch hier die Lösungen unverändert.

Satz 1.2 *Jede Matrix lässt sich durch elementare Zeilenumformungen und Spaltenvertauschungen auf Staffelform (Spezialfall 3) bringen. Wenn man dabei auch die Unbekannten entsprechend umnummeriert, ändern sich dabei die Lösungen nicht.*

Manche Mathematiker sind mit einer Zeilenstufenform noch nicht zufrieden. Wenn die Koeffizientenmatrix z.B. quadratisch ist, und die Zeilenstufenform so aussieht

$$Z = \left(\begin{array}{cccc|c} 1 & z_{1,2} & ... & z_{1,n} & b'_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & & \ddots & z_{n-1,n} & b'_{n-1} \\ 0 & & 0 & 1 & b'_n \end{array}\right),$$

kann man die Umformungen noch etwas weiter treiben:

Vorletzte Zeile	-	$z_{n-1,n}$-mal die letzte Zeile,
$(n-2)$-te Zeile	-	$z_{n-2,n}$-mal die letzte Zeile,
	$\vdots$	
erste Zeile	-	$z_{1,n}$-mal die letzte Zeile.

Damit hat man erreicht, dass in der letzten Spalte alle Einträge $= 0$ sind, bis auf den Eintrag $= 1$ ganz unten. Mit einem analogen Verfahren, kann man auch alle Einträge in der vorletzten Spalte auf 0 bringen, bis auf den vorletzten Eintrag $= 1$. Man muss dazu von jeder Zeile geeignete Vielfache der vorletzten Zeile abziehen. Wenn man dies von rechts nach links mit allen Spalten macht, sieht am Ende die erweiterte Koeffizientenmatrix so aus:

$$\begin{array}{cccc|c} 1 & 0 & ... & 0 & b''_1 \\ 0 & 1 & & \vdots & b''_2 \\ \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & ... & 0 & 1 & b''_n \end{array}.$$

Dazu gehört das besonders einfache Gleichungssystem

$$\begin{array}{rcl} 1 \cdot x_1 & = & b_1'' \\ & \vdots & \\ 1 \cdot x_n & = & b_n'' \end{array}$$

Diese Matrix hat Diagonalform (Spezialfall 2). Hier kann man die Lösung

$$x_1 = b_1'', \quad ..., \quad x_n = b_n''$$

sehr einfach ablesen. Bringt man die Koeffizientenmatrix auf diese Form, so nennt man dies das Gauß-Jordan-Verfahren.

Wir schreiben lineare Gleichungssysteme noch etwas formaler

$$\sum_{\nu=1}^{n} a_{\mu,\nu} \cdot x_\nu = b_\mu, \quad \mu = 1, ..., m,$$

und beschließen diesen ersten Abschnitt mit einigen allgemeinen Tatsachen.

Dabei nennen wir ein Gleichungssystem *homogen*, wenn alle Zahlen $b_1, ..., b_m$ auf seiner rechten Seite 0 sind. Andernfalls nennen wir das Gleichungssystem *inhomogen*. Ein homogenes Gleichungssystem hat immer die uninteressante *triviale Lösung* $x_1 = ... = x_n = 0$. Ist dabei das System vom Typ 3a (eindeutige Lösung), so gibt es nur diese triviale Lösung.

Satz 1.3 (Mehr Unbekannte als Gleichungen) *Das homogene lineare Gleichungssystem*

$$\sum_{\nu=1}^{n} a_{\mu,\nu} \cdot x_\nu = 0, \quad \mu = 1, ..., m,$$

habe n Unbekannte und $m < n$ Zeilen. Dann können in den Lösungen $(x_1, ..., x_n)$ mindestens $n - m$ Parameter frei gewählt werden.

Beweis. Die Anzahl der Stufen in einer Matrix mit n Spalten und m Zeilen ist höchstens m. Somit gibt es mindestens $n - m$ Spalten, in denen keine Stufe steht, und in denen die Unbekannte beliebig gewählt werden kann. □

Satz 1.4 (Struktursatz) *Ist eine* spezielle Lösung $(y_1, ..., y_n)$ *des inhomogenen Systems*

$$\sum_{\nu=1}^{n} a_{\mu,\nu} \cdot x_\nu = b_\mu, \quad \mu = 1, ..., m$$

bekannt, so erhält man daraus alle Lösungen des inhomogenen Systems durch Addition aller Lösungen des zugehörigen homogenen Systems.

Beweis. Nach Annahme ist für $\mu = 1, ..., m$

$$\sum_{\nu=1}^{n} a_{\mu,\nu} \cdot y_\nu = b_\mu.$$

Wenn $(x_1, ..., x_n)$ eine beliebige Lösung des inhomogenen Systems ist, also

$$\sum_{\nu=1}^{n} a_{\mu,\nu} \cdot x_\nu = b_\mu,$$

so folgt

$$\sum_{\nu=1}^{n} a_{\mu,\nu} \cdot (x_\nu - y_\nu) = \sum_{\nu=1}^{n} a_{\mu,\nu} \cdot x_\nu - \sum_{\nu=1}^{n} a_{\mu,\nu} \cdot y_\nu = b_\mu - b_\mu = 0.$$

Das heißt: $(h_1, ..., h_n) = (x_1 - y_1, ..., x_n - y_n)$ ist eine Lösung des homogenen Systems.

Ist umgekehrt $(h_1, ..., h_n)$ eine Lösung des homogenen Systems so gilt

$$\sum_{\nu=1}^{n} a_{\mu,\nu} h_\nu = 0 \text{ für } \mu = 1, ..., m.$$

Daraus folgt

$$\sum_{\nu=1}^{n} a_{\mu,\nu} \cdot (y_\nu + h_\nu) = \sum_{\nu=1}^{n} a_{\mu,\nu} \cdot y_\nu + \sum_{\nu=1}^{n} a_{\mu,\nu} \cdot h_\nu = b_\mu + 0 = b_\mu.$$

Also ist auch $(x_1, ..., x_n) = (y_1 + h_1, ..., y_n + h_n)$ eine Lösung des inhomogenen Systems. $\square$

Ist $(h_1, ..., h_n)$ eine Lösung eines homogenen Systems, so auch

$$c \cdot (h_1, ..., h_n) = (c \cdot h_1, ..., c \cdot h_n)$$

für alle $c \in \mathbb{R}$. Hat das homogene System eine *nicht-triviale* Lösung, so hat es deswegen unendlich viele Lösungen. Ist darüber hinaus ein zugehöriges inhomogenes System lösbar, so besitzt auch dieses unendlich viele Lösungen.

Aufgabe 1.1 *Wenn fünf Ochsen und zwei Schafe acht Taels Gold kosten, sowie zwei Ochsen und acht Schafe auch acht Taels, was ist dann der Preis eines Tieres? (Chiu-chang Suan-chu, $\sim$ 300 n.Chr.)*

Aufgabe 1.2 *Auf einem Markt gibt es Hühner zu kaufen. Ein Hahn kostet drei Geldstücke, eine Henne zwei, und Küken kann man drei für ein Geldstück haben. Wie muss man es einrichten, um für 100 Geldstücke 100 Hühner zu bekommen? (Hinweise: Es gibt mehrere Lösungen, alle sind zu bestimmen. Als Anzahlen von Hühnern sind dabei nur ganze Zahlen ≥ 0 zugelassen.)*

Aufgabe 1.3 *Bestimmen Sie alle Lösungen des folgenden Gleichungssystems*

$$\begin{array}{rcrcrcrcrcr}
2x_1 & - & x_2 & - & x_3 & + & 3x_4 & + & 2x_5 & = & 6 \\
-4x_1 & + & 2x_2 & + & 3x_3 & - & 3x_4 & - & 2x_5 & = & -5 \\
6x_1 & - & 2x_2 & + & 3x_3 & & & - & x_5 & = & -3 \\
2x_1 & & & + & 4x_3 & - & 7x_4 & - & 3x_5 & = & -8 \\
 & & x_2 & + & 8x_3 & - & 5x_4 & - & x_5 & = & -3
\end{array}$$

Aufgabe 1.4 *Lösen Sie das lineare Gleichungssystem*

$$\begin{array}{rcrcrcrcr}
x_1 & + & 2x_2 & + & 3x_3 & + & x_4 & = & 0, \\
x_1 & + & 3x_2 & + & 2x_3 & + & x_4 & = & 1, \\
2x_1 & + & x_2 & - & x_3 & + & 4x_4 & = & 1, \\
 & & 6x_2 & + & 4x_3 & + & 2x_2 & = & 0.
\end{array}$$

Aufgabe 1.5 *Entscheiden Sie, welches der folgenden beiden Gleichungssysteme lösbar ist, und lösen Sie dieses:*

$$\begin{array}{rcrcrcrcl} x_1 & + & 2x_2 & & & + & x_4 & = & 0, \\ 2x_1 & + & 3x_2 & - & 2x_3 & + & 3x_4 & = & 0, \\ x_1 & + & x_2 & - & 2x_3 & + & 2x_4 & = & 1. \end{array}$$

$$\begin{array}{rcrcrcrcl} x_1 & + & 2x_2 & + & x_3 & + & x_4 & = & 0, \\ 2x_1 & + & 3x_2 & - & 2x_3 & + & 3x_4 & = & 0, \\ x_1 & + & x_2 & - & 2x_3 & + & 2x_4 & = & 1. \end{array}$$

Aufgabe 1.6 *Es seien r, s, t drei verschiedene reelle Zahlen. Zeigen Sie, dass für alle $a, b, c \in \mathbb{R}$ das Gleichungssystem*

$$\begin{array}{rcrcrcl} x_1 & + & rx_2 & + & r^2x_3 & = & a \\ x_1 & + & sx_2 & + & s^2x_3 & = & b \\ x_1 & + & tx_2 & + & t^2x_3 & = & c \end{array}$$

genau eine reelle Lösung hat und bestimmen Sie diese.

Aufgabe 1.7 *Es sei n eine natürliche Zahl. Lösen Sie das Gleichungssystem*

$$\begin{array}{lcll} x_1 & + & x_2 & = & 0 \\ x_2 & + & x_3 & = & 0 \\ & \vdots & & \vdots & \\ x_{n-2} & + & x_{n-1} & = & 0 \\ x_{n-1} & + & x_n & = & 0 \\ x_n & + & x_0 & = & 0 \end{array}$$

Aufgabe 1.8 *Ein* 9*-tupel* $(x_1, ..., x_9) \in \mathbb{R}^9$ *heiße magisches Quadrat, wenn*

$$\begin{array}{ccccccccc} & x_1+x_2+x_3 & = & x_4+x_5+x_6 & = & x_7+x_8+x_9 & = & x_1+x_4+x_7 & = \\ = & x_2+x_5+x_8 & = & x_3+x_6+x_9 & = & x_1+x_5+x_9 & = & x_3+x_5+x_7 & \end{array}$$

gilt. Stellen Sie ein lineares Gleichungssystem auf, das diesen acht Bedingungen äquivalent ist, und lösen Sie dieses.

Aufgabe 1.9 *Bestimmen Sie $t \in \mathbb{R}$ so, dass das folgende System*

$$\begin{array}{rcrcrcl} 2x_1 & + & 3x_2 & + & tx_3 & = & 3 \\ x_1 & + & x_2 & - & x_3 & = & 1 \\ x_1 & + & tx_2 & + & 3x_3 & = & 2 \end{array}$$

keine Lösung, bzw. mehr als eine Lösung, bzw. genau eine Lösung hat.

Aufgabe 1.10 *Untersuchen Sie, ob die beiden folgenden Gleichungssysteme eine von Null verschiedene Lösung haben:*

$$a)\quad \begin{array}{rcrcrcl} x_1 & + & x_2 & - & x_3 & = & 0 \\ 2x_1 & - & 3x_2 & + & x_3 & = & 0 \\ x_1 & - & 4x_2 & + & 2x_3 & = & 0 \end{array} \qquad b)\quad \begin{array}{rcrcrcl} x_1 & + & x_2 & - & x_3 & = & 0 \\ 2x_1 & + & 4x_2 & - & x_3 & = & 0 \\ 3x_1 & + & 2x_2 & + & 2x_3 & = & 0 \end{array}$$

Aufgabe 1.11 *Bringen Sie die folgenden Matrizen durch elementare Zeilenumformungen auf Zeilenstufenform:*

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

Aufgabe 1.12 *a) Geben Sie alle möglichen Zeilenstufenformen einer Matrix mit zwei Zeilen und drei Spalten an.*
b) Geben Sie hinreichende und notwendige Bedingungen dafür an, dass die Matrix

$$\begin{pmatrix} a & b & r \\ c & d & s \end{pmatrix}$$

auf die Zeilenstufenform

$$\begin{pmatrix} 1 & * & * \\ 0 & 1 & * \end{pmatrix}, \quad \textit{bzw.} \begin{pmatrix} 0 & 1 & * \\ 0 & 0 & 0 \end{pmatrix}, \quad \textit{bzw.} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

gebracht werden kann.

Aufgabe 1.13 *Sei $m > n \geq 1$, und sei ein unlösbares lineares Gleichungssystem von m Gleichungen in n Unbekannten gegeben. Man begründe, dass es $n + 1$ dieser Gleichungen gibt, die bereits keine Lösung haben.*

Aufgabe 1.14 *Man bestimme alle $\lambda \in \mathbb{R}$, für die das lineare Gleichungssystem*

$$\begin{array}{rcrcrcl} 2x_1 & + & x_2 & & & = & 1 \\ 3x_1 & - & x_2 & + & 6x_3 & = & 5 \\ 4x_1 & + & 3x_2 & - & x_3 & = & 2 \\ & & 5x_2 & + & 2x_3 & = & \lambda \end{array}$$

lösbar ist.

Aufgabe 1.15 *Gegeben sei das lineare Gleichungssystem*

$$\begin{array}{rcrcrcrcrcl} 2x_1 & + & x_2 & + & ax_3 & + & x_4 & & & = & 0 \\ x_1 & & & & & + & ax_4 & - & ax_5 & = & 1 \\ & & 2x_2 & + & x_3 & & & + & 2x_5 & = & 2 \end{array}$$

a) Man bestimme die Lösungsmenge des Gleichungssystems für $a = 1$.
b) Gibt es ein $a \in \mathbb{R}$, für welches das Gleichungssytem keine Lösung hat?
c) Gibt es ein $a \in \mathbb{R}$, für welches das Gleichungssytem genau eine Lösung hat?

Aufgabe 1.16 *Für welche Werte des Parameters $s \in \mathbb{R}$ besitzt das lineare Gleichungssystem mit Koeffizientenmatrix A und rechter Seite $\mathbf{b}$, wo*

$$A = \begin{pmatrix} s-1 & -1 & 0 & -1 \\ 0 & s-2 & 1 & -1 \\ 1 & 0 & s & 0 \\ s & 1-s & 1 & 0 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} -1 \\ s \\ 1 \\ 1 \end{pmatrix},$$

a) genau eine Lösung?
b) keine Lösung ?
c) unendlich viele Lösungen?

Aufgabe 1.17 *a) Für welche Paare* $(a,b) \in \mathbb{R}^2$ *hat das Gleichungssystem*

$$\begin{array}{rcrcrcr} 2x_1 & + & 2x_2 & + & (a+1)x_3 & = & 2 \\ x_1 & + & 2x_2 & + & x_3 & = & 0 \\ ax_1 & + & & & bx_3 & = & -2 \end{array}$$

keine Lösung $(x_1, x_2, x_3) \in \mathbb{R}^3$*?*
b) Im Falle $b = 1$ *bestimme man alle Lösungen (in Abhängigkeit von* a*).*

Aufgabe 1.18 *Für welche reellen Zahlen* b *hat das Gleichungssystem*

$$\begin{array}{rcrcrcr} x_1 & + & x_2 & + & x_3 & = & 0 \\ x_1 & + & bx_2 & + & x_3 & = & 4 \\ bx_1 & + & 3x_2 & + & bx_3 & = & -2 \end{array}$$

keine, genau eine, unendlich viele Lösungen? Man gebe im letzten Fall die Lösungsmenge an.

Aufgabe 1.19 *Man bestimme die Lösungsgesamtheit (im* $\mathbb{R}^5$*) des Gleichungssystems*

$$\begin{array}{rcrcrcrcrcl} x_1 & - & x_2 & + & x_3 & - & x_4 & + & x_5 & = & 2 \\ x_1 & + & x_2 & + & x_3 & + & x_4 & + & x_5 & = & 1+\lambda \\ x_1 & & & + & \lambda x_3 & & & + & x_5 & = & 2 \end{array}$$

in Abhängigkeit von $\lambda \in \mathbb{R}$*.*

Aufgabe 1.20 *Betrachten Sie das lineare Gleichungssystem*

$$\begin{array}{rcrcrcl} \lambda x & + & y & & & = & \mu \\ x & + & \lambda y & + & z & = & \mu \\ & & y & + & \lambda z & = & \mu \end{array} \qquad \textit{in } x, y, z \in \mathbb{R}.$$

Für welche $\lambda, \mu \in \mathbb{R}$ *ist dieses Gleichungssystem lösbar?*

1.2 $\mathbb{R}$-Vektorräume

1.2.1 Vektorrechnung im $\mathbb{R}^n$

Unter einem *Vektor* verstehen wir vorerst ein n-tupel

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

reeller Zahlen $x_1, ..., x_n$. Es ist üblich, sich Vektoren als derartige Spaltenvektoren vorzustellen, während es aus schreibtechnischen Gründen besser wäre, Zeilenvektoren

$$\mathbf{x} = (x_1, ..., x_n)$$

zu benützen. Der Übergang von Zeile zu Spalte (und umgekehrt) wird durch das Symbol ^{t} (hochgestellt) gekennzeichnet. Es ist also

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}^t = (x_1, ..., x_n) \quad \text{und} \quad (x_1, ..., x_n)^t = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Allgemein gilt für Zeilen- wie Spaltenvektoren

$$\mathbf{x}^{tt} = \mathbf{x}.$$

Wir wollen Zahlenvektoren als Spalten auffassen, sie aus schreibtechnischen Gründen meist als Zeilenvektoren schreiben. Zur Verdeutlichung werden, so wie auch bisher schon, Elemente $\mathbf{x} \in \mathbb{R}^n$ durch **Fettdruck** von Zahlen $c \in \mathbb{R}$ unterschieden.

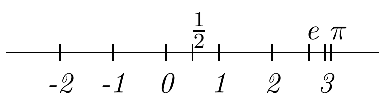
Das n-tupel $(x_1, ..., x_n)$ ist etwas anderes als die Menge $\{x_1, ..., x_n\}$, da es bei einem n-tupel auf die Reihenfolge der Einträge ankommt und bei einer Menge nicht.

Definition 1.3 *Der* n-dimensionale Zahlenraum *ist also die Menge*

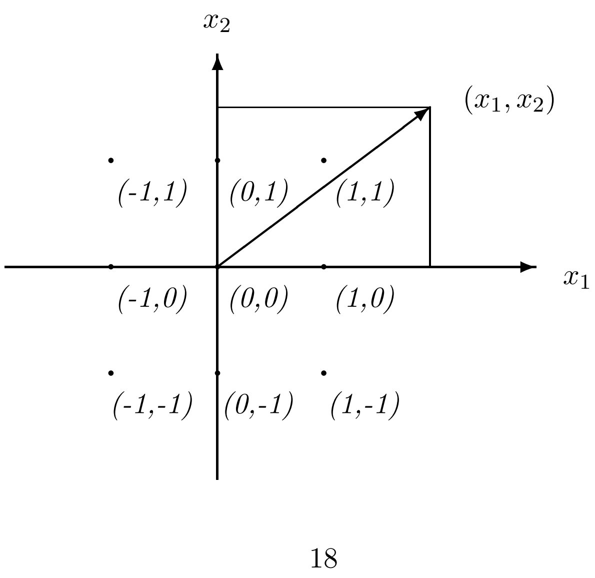
$$\mathbb{R}^n := \{(x_1, ..., x_n) \,:\, x_1 \in \mathbb{R}, ..., x_n \in \mathbb{R}\}$$

aller dieser Vektoren.

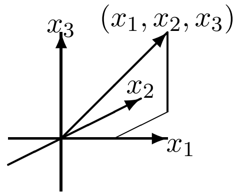
Beispiel 1.6 *Wir stellen die einfachsten Fälle vor:*
$n = 1.$ $\quad \mathbb{R}^1 = \mathbb{R}$ *ist die Zahlengerade*



$n = 2.$ *Seit Descartes ist es üblich, nach Wahl eines Koordinatensystems, die Punkte der Ebene durch Zahlenpaare* (x_1, x_2) *zu parametrisieren. Umgekehrt gibt die Ebene eine Veranschaulichung des Raums* $\mathbb{R}^2$ *der Zahlenpaare* (x_1, x_2). *Man „identifiziert“ den Zahlenraum* $\mathbb{R}^2$ *mit der Ebene.*



$n = 3$. *Ebenso, wie die Punkte der Ebene mit den Zahlenpaaren* $(x_1, x_2) \in \mathbb{R}^2$ *identifiziert werden können, so können nach Wahl eines Koordinatensystems die Punkte des Anschauungsraums mit Zahlentripeln* $(x_1, x_2, x_3) \in \mathbb{R}^3$ *identifiziert werden.*



$n = 4$. *Zu Beginn des 20. Jahrhunderts schlug A. Einstein den vierdimensionalen Zahlenraum* $\mathbb{R}^4$ *in seiner speziellen Relativitätstheorie als geometrisches Modell für den uns umgebenden Raum vor, wobei die Zeit als vierte Koordinate interpretiert wird. Erst wenige Jahre vorher war es in der Mathematik üblich geworden, geometrische Betrachtungen auch in mehr als drei Dimensionen durchzuführen. Die italienischen Geometer hatten diese Zahlenräume höherer Dimension, welche sie zunächst „Hyperräume" nannten, in die Mathematik eingeführt.*

Bei einem LGS mit n Unbekannten und m Zeilen treten n-tupel auf:

- der Vektor $(x_1, ..., x_n)^t$ der Unbekannten, bzw., wenn man diese ermittelt hat, der Lösungsvektor,
- die (Transponierten der) m Zeilenvektoren $\mathbf{a}_\mu = (a_{\mu,1}, ..., a_{\mu,n})$, $\mu = 1, ..., m$, der Koeffizientenmatrix

sowie auch m-tupel als

- die rechte Seite $(b_1, ..., b_m)^t$,
- die n Spalten $\mathbf{a}^\nu = (a_{1,\nu}, ..., a_{m,\nu})^t$, $\nu = 1, ..., n$, der Koeffizientenmatrix.

Mit den Vektoren des Zahlenraums $\mathbb{R}^n$ kann man die folgenden beiden Rechenoperationen durchführen:

Definition 1.4 *a) Die* Addition $'+' : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ *ist erklärt duch die Vorschrift*

$$\boxed{\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} \qquad \text{für alle } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.}$$

Der Vektor $\mathbf{x} + \mathbf{y}$ *heißt die* Summe *von* $\mathbf{x}$ *und* $\mathbf{y}$.

b) Die Operation $'\cdot' : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$, *die Multiplikation mit Skalaren ist erklärt durch*

$$\boxed{c \cdot \mathbf{x} = c \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} := \begin{pmatrix} c \cdot x_1 \\ \vdots \\ c \cdot x_n \end{pmatrix} \qquad \text{für alle } c \in \mathbb{R}, \mathbf{x} \in \mathbb{R}^n.}$$

Der Vektor $c \cdot \mathbf{x}$ *heißt* skalares Vielfaches *von* $\mathbf{x}$.

Beide Rechenoperationen sind komponentenweise nichts anderes als die übliche Addition und Multiplikation reeller Zahlen. Deswegen gelten auch hier die wohlbekannten Rechenregeln:

(A)	für die Addition	
(A.V1)	$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$,	(**Kommutativgesetz**)
(A.V2)	$\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z})$	(**Assoziativgesetz**)
(A.V3)	die Gleichung $\mathbf{a} + \mathbf{x} = \mathbf{b}$ hat bei jeder Vorgabe von $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ genau eine Lösung	
(M)	für die Multiplikation mit Skalaren	
(M.V1)	$(c + d) \cdot \mathbf{x} = c \cdot \mathbf{x} + d \cdot \mathbf{x}$	(**1. Distributivgesetz**)
(M.V2)	$c \cdot (\mathbf{x} + \mathbf{y}) = c \cdot \mathbf{x} + c \cdot \mathbf{y}$	(**2. Distributivgesetz**)
(M.V3)	$(c \cdot d) \cdot \mathbf{x} = c \cdot (d \cdot \mathbf{x})$	(**Assoziativgesetz**)
(M.V4)	$1 \cdot \mathbf{x} = \mathbf{x}$	(**neutrales Element**)

Die Eigenschaft (A.V3) ist äquivalent mit den folgenden beiden Eigenschaften zusammen.

(A.V3')	Es gibt genau ein $\mathbf{0} \in \mathbb{R}^n$ so, dass $\mathbf{x} + \mathbf{0} = \mathbf{x}$ für alle $\mathbf{x} \in \mathbb{R}^n$ nämlich $\mathbf{0} = (0, ..., 0)^t$.	(**neutrales Element**)
(A.V3")	Zu jedem $\mathbf{x} \in \mathbb{R}^n$ gibt es genau ein $-\mathbf{x} \in \mathbb{R}^n$ so, dass $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$, nämlich $-\mathbf{x} = (-x_1, ..., -x_n)^t$.	(**inverses Element**)

Wie üblich benutzen wir die Kurzschreibweise

$$\mathbf{x} - \mathbf{y} := \mathbf{x} + (-\mathbf{y}) = (x_1 - y_1, ..., x_n - y_n)^t.$$

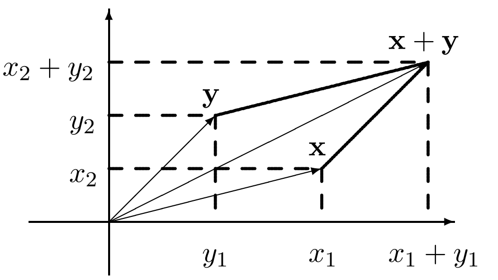
Damit wird $\mathbf{x} = \mathbf{b} - \mathbf{a}$ die Lösung der Gleichung $\mathbf{a} + \mathbf{x} = \mathbf{b}$.

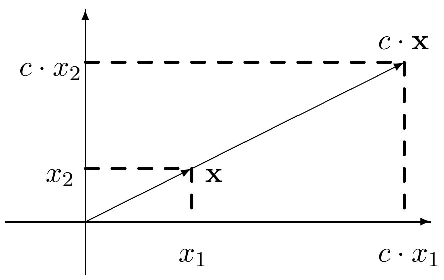
Durch komponentenweise Rechnung sieht man unmittelbar

$$(F.V) \qquad \left\{ \begin{array}{rcl} 0 \cdot \mathbf{x} & = & \mathbf{0}, \\ (-1) \cdot \mathbf{x} & = & -\mathbf{x}, \\ c \cdot \mathbf{0} & = & \mathbf{0}, \\ c \cdot \mathbf{x} = \mathbf{0} & \Leftrightarrow & c = 0 \text{ oder } \mathbf{x} = \mathbf{0}. \end{array} \right.$$

Definition 1.5 *Der Raum* $\mathbb{R}^n$ *mit den obigen Verknüpfungen* $'+'$ *und* $'\cdot'$ *heißt n-dimensionaler Skalarenvektorraum über* $\mathbb{R}$.

Viel mehr gibt es über das Rechnen mit Vektoren im $\mathbb{R}^n$ nicht zu sagen.



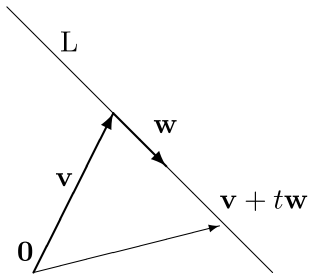


Den $\mathbb{R}^2$ bzw. $\mathbb{R}^3$ interpretiert man als Ebene bzw. den Anschauungsraum. Hier ist die komponentenweise Addition die Addition nach dem „Kräfteparallelogramm“ . Die Multiplikation für $|\lambda| > 1$ eine Streckung und für $|\lambda| < 1$ eine Stauchung dar, wobei für $\lambda < 0$ die Richtung umgekehrt wird.

Wir möchten noch an einem ganz einfachen Beispiel das Wesen der Linearen Algebra demonstrieren, das darin besteht, Algebra auf geometrische Sachverhalte anzuwenden, bzw. umgekehrt, intuitive Methoden aus der Geometrie für algebraische Anwendung zu abstrahieren. Als Beispiel diskutieren wir Geraden (in der Ebene und allgemeiner).

Definition 1.6 *Eine Gerade L im Zahlenraum $\mathbb{R}^n$ wird gegeben durch einen Anfangsvektor $\mathbf{v}$ und einen Richtungsvektor $\mathbf{0} \neq \mathbf{w} \in \mathbb{R}^n$. Sie ist die Menge*

$$L = \{\mathbf{v} + t\mathbf{w} \in \mathbb{R}^n \ : t \in \mathbb{R}\} =: \mathbf{v} + \mathbb{R} \cdot \mathbf{w}.$$



Satz 1.5 *Die Gerade L stimmt mit einer zweiten Geraden $L' = \{\mathbf{v}' + s\mathbf{w}' \ : s \in \mathbb{R}\}$ genau dann überein, wenn $\mathbf{v}' \in L$ und $\mathbf{w}' = c \cdot \mathbf{w}$ mit $0 \neq c \in \mathbb{R}$.*

Beweis. „$\Rightarrow$“ Wenn die Mengen $L = \{\mathbf{v}+t\mathbf{w} \ : t \in \mathbb{R}\}$ und $L' = \{\mathbf{v}'+s\mathbf{w}' \ : s \in \mathbb{R}\}$ übereinstimmen, dann ist insbesondere ($s = 0$) der Vektor $\mathbf{v}'$ ein Vektor auf L, also von der Form $\mathbf{v}' = \mathbf{v} + t_0\mathbf{w}$. Ebenso ist ($s = 1$) auch $\mathbf{v}' + \mathbf{w}' \in L$, also $\mathbf{v} + t_0\mathbf{w} + \mathbf{w}' = \mathbf{v}' + \mathbf{w}' = \mathbf{v} + t\mathbf{w}$. Daraus folgt $\mathbf{w}' = c\mathbf{w}$ mit $c = t - t_0$. Wegen $\mathbf{w}' \neq \mathbf{0}$ muss auch $c \neq 0$ sein.

„$\Leftarrow$“ Sei $\mathbf{v}' = \mathbf{v} + t_0\mathbf{w} \in L$ und $\mathbf{w}' = c\mathbf{w}$. Dann ist

$$L' = \{\mathbf{v}' + s\mathbf{w}' \ : s \in \mathbb{R}\} = \{\mathbf{v} + (t_0 + sc)\mathbf{w} \ : s \in \mathbb{R}\} = \{\mathbf{v} + t\mathbf{w} \ : t \in \mathbb{R}\},$$

denn wegen $c \neq 0$ durchläuft mit s auch $t = t_0 + sc$ alle reellen Zahlen. □

Satz 1.6 *Durch je zwei Vektoren $\mathbf{x} \neq \mathbf{y}$ des $\mathbb{R}^n$ gibt es genau eine Gerade L.*

Beweis. Existenz: wir wählen $\mathbf{v} := \mathbf{x}$ und $\mathbf{w} := \mathbf{y} - \mathbf{x}$. Dann enthält die Gerade $L = \{\mathbf{v} + t\mathbf{w} \ : t \in \mathbb{R}\} = \{\mathbf{x} + t(\mathbf{y} - \mathbf{x}) \ : t \in \mathbb{R}\}$ beide Vektoren $\mathbf{x}$ (für $t = 0$) und $\mathbf{y}$ (für $t = 1$).

Eindeutigkeit: Sei $L' = \{\mathbf{v}' + t\mathbf{w}' \ : t \in \mathbb{R}\}$ eine Gerade, welche die Vektoren $\mathbf{x}$ und $\mathbf{y}$ enthält. Wegen Satz 1.5 können wir diese Gerade auch schreiben als $L' = \{\mathbf{x} + t\mathbf{w}' \ : t \in \mathbb{R}\}$. Da $\mathbf{y} = \mathbf{x} + t_0\mathbf{w}'$ mit $t_0 \neq 0$ (wegen $\mathbf{x} \neq \mathbf{y}$), ist der Richtungsvektor $\mathbf{w}' = \frac{1}{t_0}(\mathbf{y} - \mathbf{x})$ ein Vielfaches des Richtungsvektors $\mathbf{y} - \mathbf{x}$ von L. Nach Satz 1.5 ist somit $L' = L$. □

Die Gerade durch $\mathbf{x}$ und $\mathbf{y}$ lässt sich etwas anders schreiben:

$$L = \{\mathbf{x} + t(\mathbf{y} - \mathbf{x}) \ : t \in \mathbb{R}\} = \{(1-t)\mathbf{x} + t\mathbf{y} \ : t \in \mathbb{R}\} = \{s\mathbf{x} + t\mathbf{y} \ : s, t \in \mathbb{R}, s + t = 1\}.$$

Die Gerade durch $\mathbf{x}$ und $\mathbf{y}$ ist nicht dasselbe, wie die *Strecke zwischen* $\mathbf{x}$ und $\mathbf{y}$

$$\{\mathbf{x} + t(\mathbf{y} - \mathbf{x}) \ : t \in \mathbb{R}\} \neq \{s\mathbf{x} + t\mathbf{y} \ : 0 \leq s, t \in \mathbb{R}, s + t = 1\}.$$

Für $s = t = \frac{1}{2}$ erhält man den *Mittelpunkt* $\frac{1}{2}(\mathbf{x} + \mathbf{y})$ dieser Strecke.

Nach diesen einfachen Tatsachen, welche in jedem Zahlenraum $\mathbb{R}^n$ richtig sind, betrachten wir jetzt den Zusammenhang von Geraden im $\mathbb{R}^2$ mit linearen Gleichungen in zwei Unbekannten.

Satz 1.7 *Für eine Teilmenge $L \subset \mathbb{R}^2$ sind folgende Eigenschaften äquivalent:*
(i): L ist eine Gerade **durch den Nullpunkt** *($\mathbf{0} \in L$).*
(ii): L ist Lösungsmenge einer **homogenen** *linearen Gleichung*

$$a_1x_1 + a_2x_2 = 0,$$

mit Koeffizienten a_1, a_2, die nicht beide 0 sind, d.h., wo $(a_1, a_2)^t \neq \mathbf{0}$.

Beweis. „(i) ⇒ (ii)“ : Als Anfangsvektor für L nehmen wir den Nullvektor und beschreiben unsere Gerade als

$$L = \{t\mathbf{w} \ : t \in \mathbb{R}\} = \{(tw_1, tw_2)^t \ : t \in \mathbb{R}\}$$

mit Koeffizienten w_1, w_2, die nicht beide 0 sind. Für unsere homogene Gleichung brauchen wir Koeffizienten a_1, a_2 mit der Eigenschaft $a_1w_1 + a_2w_2 = 0$. Die Zahlen

$$a_1 := w_2, \ a_2 := -w_1$$

bieten sich dafür an. Wir behaupten, dass L mit der Menge $\{(x_1, x_2)^t \in \mathbb{R}^2 \ : w_2x_1 - w_1x_2 = 0\}$ übereinstimmt. Wegen $w_2 \cdot tw_1 - w_1 \cdot tw_2 = 0$ ist klar, dass L in dieser Menge enthalten ist. Umgekehrt ist diese Menge aber, wie wir im nächsten Beweisschritt sehen werden, eine Gerade. Da sie $\mathbf{0}$ und $\mathbf{w}$ enthält, stimmt sie nach Satz 1.6 mit L überein.

„(ii) ⇒ (i)“ : Falls $a_1 \neq 0$, so erfüllt $\mathbf{x} = (x_1, x_2)^t$ die Gleichung $a_1x_1 + a_2x_2 = 0$ genau dann, wenn $x_1 = -\frac{a_2}{a_1}x_2$, das heißt, wenn $\mathbf{x} = x_2 \cdot (-\frac{a_2}{a_1}, 1)^t$ auf der Geraden durch $\mathbf{0}$ mit dem Richtungsvektor $\mathbf{w} = (-\frac{a_2}{a_1}, 1)^t$ liegt. Wenn aber $a_1 = 0$, so lautet die Gleichung $a_2x_2 = 0$. Da nun nach Voraussetzung $a_2 \neq 0$, ist dies äquivalent mit $x_2 = 0$. Diese Menge ist die Gerade durch den Nullpunkt mit Richtungsvektor $(1, 0)^t$. □

Satz 1.8 *Für eine Teilmenge $L \subset \mathbb{R}^2$ sind äquivalent:*
(i) L ist eine Gerade **nicht durch den Nullpunkt** *(nicht $\mathbf{0} \in L$).*
(ii) L ist Lösungsmenge einer **inhomogenen** *linearen Gleichung $a_1x_1 + a_2x_2 = b$, wobei $(a_1, a_2)^t \neq \mathbf{0}$ und $b \neq 0$.*

Beweis. „(i) $\Rightarrow$ (ii)“ : Wir schreiben $L = \{\mathbf{v} + t\mathbf{w} \, : t \in \mathbb{R}\}$ und betrachten die Gerade $L_0 := \{t\mathbf{w} \, : t \in \mathbb{R}\}$ mit demselben Richtungsvektor durch den Nullpunkt. Nach Satz 1.7 ist L_0 Lösungsmenge einer homogenen linearen Gleichung $a_1 x_1 + a_2 x_2 = 0$. Also ist

$$\begin{aligned} L &= \{\mathbf{v} + \mathbf{x} \, : \mathbf{x} \in L_0\} \\ &= \{\mathbf{v} + \mathbf{x} \, : a_1 x_1 + a_2 x_2 = 0\} \\ &= \{\mathbf{y} \in \mathbb{R}^2 \, : a_1 y_1 + a_2 y_2 = a_1 v_1 + a_2 v_2\}. \end{aligned}$$

Da L nicht durch den Nullpunkt geht, liegt $\mathbf{v}$ nicht auf L_0, und es ist $b := a_1 v_1 + a_2 v_2 \neq 0$.

„(ii) $\Rightarrow$ (i)“ Jetzt ist

$$\begin{aligned} L &= \{\mathbf{x} \in \mathbb{R}^2 \, : a_1 x_1 + a_2 x_2 = b\} \\ &= \{\mathbf{v} + \mathbf{y} \in \mathbb{R}^2 \, : a_1 y_1 + a_2 y_2 = 0\} \end{aligned}$$

wo $\mathbf{v}$ eine spezielle Lösung der inhomogenen Gleichung $a_1 v_1 + a_2 v_2 = b$ ist (Satz 1.e). Nach Satz 1.7 beschreibt die homogene Gleichung $a_1 y_1 + a_2 y_2 = 0$ eine Gerade $L_0 = \{t\mathbf{w} \, : t \in \mathbb{R}\}$ durch den Nullpunkt. Somit ist $L = \{\mathbf{v} + t\mathbf{w} \, : t \in \mathbb{R}\}$ eine Gerade, die wegen $b \neq 0$ nicht durch den Nullpunkt geht. $\square$

Wir sahen, die Lösungsmenge einer linearen Gleichung in zwei Unbekannten, deren Koeffizienten nicht beide 0 sind, ist eine Gerade in der Zahlenebene $\mathbb{R}^2$. Die Lösungsmenge eines Systems von zwei derartigen linearen Gleichungen

$$\begin{array}{ll} a_{1,1}x_1 + a_{1,2}x_2 = b_1 & \text{(Lösungsmenge } L_1) \\ a_{2,1}x_1 + a_{2,2}x_2 = b_2 & \text{(Lösungsmenge } L_2) \end{array}$$

ist deswegen der Durchschnitt $L_1 \cap L_2$ der beiden Geraden. Für diesen Durchschnitt gibt es folgende Möglichkeiten:

1)	$L_1 = L_2$:		$L_1 \cap L_2$ ist die Gerade $L_1 = L_2$
2)	$L_1 \neq L_2$,	$L_1 \cap L_2 \neq \emptyset$	$L_1 \cap L_2$ ist ein Punkt
3)	$L_1 \neq L_2$,	L_1 und L_2 parallel	$L_1 \cap L_2$ ist leer

Zu diesen drei Möglichkeiten gehören die folgenden drei Stufenformen der Koeffizientenmatrix:

$$1) \quad \left(\begin{array}{cc|c} 1 & * & * \\ 0 & 0 & 0 \end{array}\right) \text{ oder } \left(\begin{array}{cc|c} 0 & 1 & * \\ 0 & 0 & 0 \end{array}\right)$$

$$2) \quad \left(\begin{array}{cc|c} 1 & * & * \\ 0 & 1 & * \end{array}\right)$$

$$3) \quad \left(\begin{array}{cc|c} 1 & * & * \\ 0 & 0 & 1 \end{array}\right) \text{ oder } \left(\begin{array}{cc|c} 0 & 1 & * \\ 0 & 0 & 1 \end{array}\right)$$

Eine analoge Situation ergibt sich im $\mathbb{R}^3$. Statt Geraden haben wir hier Ebenen.

Definition 1.7 *Es seien $\mathbf{v}, \mathbf{w}_1, \mathbf{w}_2 \in \mathbb{R}^n$ mit $\mathbf{w}_1, \mathbf{w}_2 \neq \mathbf{0}$. Weiter gebe es kein $c \in \mathbb{R}$ mit $\mathbf{w}_1 = c \cdot \mathbf{w}_2$. Dann heißt die Menge*

$$E := \{\mathbf{v} + s \cdot \mathbf{w}_1 + t \cdot \mathbf{w}_2 : \ s, t \in \mathbb{R}\} =: \mathbf{v} + \mathbb{R} \cdot \mathbf{w}_1 + \mathbb{R} \cdot \mathbf{w}_2$$

eine Ebene *im $\mathbb{R}^n$.*

Analog zu den Sätzen 1.7 und 1.8 gilt

Satz 1.9 *Die Lösungsmenge einer jeden linearen Gleichung*

$$a_1x_1 + a_2x_2 + a_3x_3 = b$$

mit Koeffizientenvektor $\mathbf{a} = (a_1, a_2, a_2)^t \neq \mathbf{0}$ *ist eine Ebene. Dabei ist* $b = 0$ *genau dann, wenn der Nullvektor* $\mathbf{0}$ *zur Ebene gehört.*

Beweis. Es sei $E \subset \mathbb{R}^3$ Lösungsmenge obiger Gleichung. Wegen Satz 1.4 genügt es, den homogenen Fall $b = 0$ zu betrachten. Wegen $\mathbf{a} \neq \mathbf{0}$ gibt es ein $a_i \neq 0$. Nach Vertauschung der Koordinaten können wir $a_1 \neq 0$ annehmen. Dann ist die allgemeine Lösung der Gleichung

$$\mathbf{x} = (-a_2x_2 - a_3x_3, x_2, x_3) = x_1\mathbf{v}_1 + x_2\mathbf{v}_2$$

mit $x_1, x_2 \in \mathbb{R}$ und

$$\mathbf{v}_1 = (-a_1, 1, 0), \quad \mathbf{v}_2 = (-a_2, 0, 1).$$

Offensichtlich sind $\mathbf{v}_1$ und $\mathbf{v}_2$ keine Vielfachen voneinander, also ist diese Menge eine Ebene. □

Der Durchschnitt $S = E_1 \cap E_2$ zweier Ebenen $E_i \subset \mathbb{R}^3$ wird also durch ein LGS mit drei Unbekannten und zwei Gleichungen beschrieben. Dabei gibt es die Möglichkeiten

	S
$E_1 = E_2$	Ebene
E_1 nicht parallel zu E_2	Gerade
$E_1 \neq E_2$ parallel	$\emptyset$

Dementsprechend wird der Durchschnitt von drei Ebenen durch ein LGS mit drei Unbekannten und drei Gleichungen beschrieben. Als weitere Möglichkeit haben wir hier

S besteht aus einem einzigen Punkt.

In diesem Fall ist das Gleichungssystem eindeutig lösbar.

Schließlich definieren wir im $\mathbb{R}^n$ allgemein

Definition 1.8 *Es seien* $\mathbf{a} \in \mathbb{R}^n$ *mit* $\mathbf{a} \neq \mathbf{0}$ *und* $b \in \mathbb{R}$. *Dann heißt*

$$H := \left\{ \mathbf{x} \in \mathbb{R}^n : \sum_{\nu=1}^{n} a_\nu x_\nu = b \right\}$$

eine Hyperebene *im* $\mathbb{R}^n$.

Eine Hyperebene im $\mathbb{R}^n$ ist also die Lösungsmenge einer **einzigen** linearen Gleichung in n Unbekannten. Im $\mathbb{R}^n$ mit $n = 2$ bzw. $= 3$ ist eine Hyperebene eine Gerade bzw. Ebene. Jede Zeile eines LGS beschreibt eine Hyperebene. Die Lösungsmenge des LGS ist der Durchschnitt all dieser Hyperebenen. Das ist die *zeilenweise Interpretation* eines LGS. Die Hyperebene H enthält genau dann den Nullvektor $\mathbf{0}$, wenn $b = 0$ ist. Deswegen enthält die Lösungsmenge eines LGS genau dann den Nullvektor, wenn das LGS homogen ist.

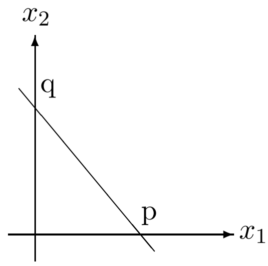
Schließlich noch ein Wort zur Nomenklatur: Die Beschreibung $L = \{\mathbf{v} + t\mathbf{w} : t \in \mathbb{R}\} = \mathbf{v} + \mathbb{R}\mathbf{w}$ heißt *Parametrisierung* oder *explizite Beschreibung* der Geraden L. Die Beschreibung $a_1x_1 + a_2x_2 = b$ einer Geraden in $\mathbb{R}^2$ heißt *implizit.*

Wenn $c \neq 0$, so ist $ca_1x_1 + ca_2x_2 = cb$ eine implizite Beschreibung der gleichen Geraden (Zeilenumformung vom Typ II). Wählt man - im Falle $b \neq 0$ - zum Beispiel $c = \frac{1}{b}$, so erhält man für

$$a_1 = \frac{1}{p} \neq 0, \quad a_2 = \frac{1}{q} \neq 0$$

die *Achsenabschnittsform*

$$\frac{1}{p}x_1 + \frac{1}{q}x_2 = 1.$$



Aufgabe 1.21 *Zeigen Sie:*
a) Die drei Geraden im $\mathbb{R}^2$

$$L_1 := \begin{pmatrix} -7 \\ 0 \end{pmatrix} + \mathbb{R} \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad L_2 := \begin{pmatrix} 5 \\ 0 \end{pmatrix} + \mathbb{R} \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad L_3 := \begin{pmatrix} 0 \\ 8 \end{pmatrix} + \mathbb{R} \cdot \begin{pmatrix} -1 \\ 4 \end{pmatrix}$$

schneiden sich in einem Punkt.
b) Die drei Punkte

$$\begin{pmatrix} 10 \\ -4 \end{pmatrix}, \quad \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} -5 \\ 6 \end{pmatrix}$$

liegen auf einer Geraden.

Aufgabe 1.22 *Es seien* $L \subset \mathbb{R}^2$ *die Gerade durch die Punkte* $(-1, 3)$ *und* $(5, -2)$, *sowie* M *die Gerade durch die Punkte* $(-2, -2)$ *und* $(1, 6)$. *Berechnen Sie den Schnittpunkt von* L *und* M.

Aufgabe 1.23 *Zeigen Sie, dass die drei Geraden im* $\mathbb{R}^2$ *mit den Gleichungen*

$$x + 2y - 1 = 0, \quad 3x + 2y + 2 = 0, \quad -x + 3y - 4 = 0$$

durch einen Punkt gehen und berechnen Sie diesen Punkt.

Aufgabe 1.24 *Untersuchen Sie, ob die Gerade*

$$L_1 := \begin{pmatrix} 5 \\ 2 \\ 0 \end{pmatrix} + \mathbb{R} \cdot \begin{pmatrix} 3 \\ -1 \\ 7 \end{pmatrix}$$

im $\mathbb{R}^3$ *die Gerade*

$$L_2 := \begin{pmatrix} 6 \\ 15 \\ 17 \end{pmatrix} + \mathbb{R} \cdot \begin{pmatrix} -2 \\ 4 \\ -1 \end{pmatrix}, \quad \textit{bzw.} \quad L_3 := \begin{pmatrix} -1 \\ 0 \\ 3 \end{pmatrix} + \mathbb{R} \cdot \begin{pmatrix} 4 \\ 2 \\ -1 \end{pmatrix}$$

schneidet und bestimmen Sie ggf. den Schnittpunkt.

Aufgabe 1.25 *Beweisen Sie, dass sich die drei Seitenhalbierenden eines Dreiecks in einem Punkt treffen.*

Aufgabe 1.26 *Beweisen Sie: Bei einem Tetraeder schneiden sich die Verbindungsgeraden der Mitten gegenüberliegender Kanten in einem Punkt.*

Aufgabe 1.27 *Es seien L_1, L_2, L_3 und L_4 vier verschiedene Geraden in der Ebene $\mathbb{R}^2$ derart, dass sich je zwei dieser Geraden in einem Punkt treffen. $S_{i,j}$ bezeichne den Schnittpunkt der Geraden S_i und S_j, ($1 \leq i < j \leq 4$). Die sechs Schnittpunkte $S_{i,j}$, $1 \leq i < j \leq 4$, seien alle verschieden. Bweisen Sie, dass die Mittelpunkte der drei Strecken $S_{1,2}S_{3,4}$, $S_{1,3}S_{2,4}$ sowie $S_{1,4}S_{2,3}$ auf einer Geraden liegen.*

1.2.2 Allgemeine $\mathbb{R}$-Vektorräume

In diesem Abschnitt verallgemeinern wir die Rechenstrukturen $'+'$ und $'\cdot'$ vom $\mathbb{R}^n$ auf allgemeinere Räume. Dies tun wir in zwei Schritten. Zunächst betrachten wir Räume, die sich vom Zahlenraum $\mathbb{R}^n$ nur unwesentlich unterscheiden, d.h., nur in der Art, wie wir ihre Elemente hinschreiben.

Beispiel 1.7 (Polynome vom Grad n) *Ein Polynom vom Grad $\leq n$ ist eine Funktion der Form*

$$f(x) = \sum_{\nu=0}^{n} a_\nu x^\nu, \quad a_0, ..., a_n \in \mathbb{R}.$$

Mit $\mathbb{R}_n[x]$ bezeichnen wir die Menge aller dieser Polynome vom Grad $\leq n$. Auch in diesem Raum sind Addition $'+'$ und Multiplikation $'\cdot'$ mit Skalaren definiert:

Addition: Sind

$$f(x) = \sum_{\nu=0}^{n} a_\nu x^\nu \quad \text{und} \quad g(x) = \sum_{\nu=0}^{n} b_\nu x^\nu \quad \in \mathbb{R}_n[x]$$

solche Polynome, so ist ihre Summe

$$(f+g)(x) = \sum_{\nu=0}^{n} a_\nu x^\nu + \sum_{\nu=0}^{n} b_\nu x^\nu = \sum_{\nu=0}^{n} (a_\nu + b_\nu) x^\nu.$$

Skalarenmultiplikation: Ist $f(x) \in \mathbb{R}_n[x]$ wie eben und $c \in \mathbb{R}$, so ist deren Produkt

$$(c \cdot f)(x) = \sum_{\nu=0}^{n} c \cdot a_\nu x^\nu.$$

Einem Polynom $f(x) \in \mathbb{R}_n[x]$ können wir seinen Koeffizientenvektor $(a_0, ..., a_n)^t \in \mathbb{R}^{n+1}$ zuordnen. Und umgekehrt ist das Polynom durch diesen Koeffizientenvektor eindeutig bestimmt. Die so definierte Abbildung

$$\mathbb{R}_n[x] \to \mathbb{R}^{n+1}$$

ist bijektiv. *Unter dieser Zuordung entspricht die Addition zweier Polynome der Addition ihrer Koeffizientenvektoren, die Multiplikation eines Polynoms mit einem Skalar der Multiplikation seines Koeffizientenvektors mit diesem Skalar. Deswegen gelten in $\mathbb{R}_n[x]$ genau die gleichen Rechenregeln wie im Zahlenraum $\mathbb{R}^{n+1}$.*

Beispiel 1.8 (Treppenfunktionen) *Wir betrachten ein abgeschlossenes Intervall $[a, b] \subset \mathbb{R}$ und eine Zerlegung dieses Intervalls in n gleichlange Intervalle der Länge*

$$h := \frac{b-a}{n}.$$

Diese Zerlegung wird definiert durch die Menge $\Delta = \{x_0, ..., x_n\}$ ihrer Zwischenpunkte mit

$$x_i := a + i \cdot h,\ 0 \leq i \leq n, \quad a = x_0 < x_1 < ... < x_{n-1} < x_n = b.$$

Eine Treppenfunktion, *bzw. ein* Histogramm *zu dieser Zerlegung Δ ist eine Funktion $t : [a, b] \to \mathbb{R}$, welche auf den Teilintervallen dieser Zerlegung konstant ist, also*

$$t(x) = a_i \text{ für } x \in [x_{i-1}, x_i[,\ i = 1, ..., n-1, \quad \text{und} \quad t(x) = a_n \text{ für } x \in [x_{n-1}, b].$$

Wir bezeichnen mit $T(\Delta)$ die Menge dieser Treppenfunktionen.

Einer solchen Treppenfunktion t kann man ihren Werte-Vektor $(a_1, ..., a_n) \in \mathbb{R}^n$ zuordnen, und sie ist umgekehrt durch diesen Vektor eindeutig bestimmt. Wieder ist die so definierte Abbildung $T(\Delta) \to \mathbb{R}^n$ bijektiv.

Auch in diesem Raum $T(\Delta)$ kann man zwei Treppenfunktionen addieren, indem man punktweise ihre Werte addiert. Und man multipliziert eine Treppenfunktion mit einem Skalar, indem man alle ihre Werte mit diesem Skalar multipliziert. Und wieder entspricht das unter der Abbildung $T(\Delta) \to \mathbb{R}^n$ den analogen Rechenoperationen im Zahlenraum $\mathbb{R}^n$

Beispiel 1.9 (Matrizen) *Wir betrachten hier Matrizen unabhängig von ihrer Interpretation als Koeffizientenmatrizen linearer Gleichungssysteme. Eine $m \times n$-*Matrix *ist also ein rechteckiges Zahlenschema*

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} = (a_{\mu,\nu})_{\substack{\mu=1,...,m \\ \nu=1,...,n}}$$

mit Zahlen $a_{\mu,\nu} \in \mathbb{R}$. Dabei heißt μ der Zeilenindex und ν der Spaltenindex. Die Menge aller dieser $m \times n$-Matrizen bezeichen wir mit $M(m \times n)$.

Auf diesem Raum $M(m \times n)$ wird Addition zweier Matrizen und Multiplikation einer Matrix mit einem Skalar komponentenweise definiert durch

$$A + B = (a_{\mu,\nu}) + (b_{\mu,\nu}) := (a_{\mu,\nu} + b_{\mu,\nu}), \quad c \cdot A = c \cdot (a_{\mu,\nu}) := (c \cdot a_{\mu,\nu}).$$

Eine $m \times n$-Matrix A kann man auffassen, als einen etwas seltsam hingeschriebenen Vektor

$$(a_{1,1}, ..., a_{1,n}, a_{2,1}, ..., a_{2,n}, ..., a_{m,n})^t \in \mathbb{R}^{m \cdot n}.$$

Und wieder entsprechen dabei die eben definierten Rechenoperationen der Matrizen den Rechenoperationen im Zahlenraum $R^{m \cdot n}$.

Nach diesen Vorbereitungen nun

Definition 1.9 *Es sei V eine nichtleere Menge mit zwei Verknüpfungen*

$$\textit{Addition: } V \times V \to V, \quad \textit{Skalarmultiplikation: } \mathbb{R} \times V \to V$$

so, dass die Regeln A.V1-3 und M.V1-4 gelten. Dann heißt $(V, +, \cdot)$ oder kurz V ein reeller Vektorraum *($\mathbb{R}$-Vektorraum). Die Elemente $\mathbf{x} \in V$ heißen Vektoren. Das neutrale Element wird mit $\mathbf{0}$ und das zu $\mathbf{x}$ inverse Element mit $-\mathbf{x}$ bezeichnet.*

Die Zahlenräume $\mathbb{R}^n$ und die in den Beispielen 1.7-1.9 definierten Räume sind $\mathbb{R}$-Vektorräume. Es gibt auch welche, die von diesen Vektorräumen wesentlich verschieden sind:

Beispiel 1.10 *Es sei $\mathbb{R}^\infty$ die Menge aller unendlichen Folgen $(a_\nu)_{\nu \in \mathbb{N}}$ reeller Zahlen mit den Rechenoperationen*

$$(a_\nu) + (b_\nu) := (a_\nu + b_\nu), \quad c \cdot (a_\nu) := (c \cdot a_\nu).$$

Dadurch wird ein $\mathbb{R}$-Vektorraum definiert. Es ist intuitiv klar, dass dieser Raum $\mathbb{R}^\infty$ wesentlich 'größer' ist als die Zahlenräume $\mathbb{R}^n$. Aber das präzise zu machen ist nicht ganz einfach (s. Definitionen 1.17 und 1.18)

Die Menge aller reellen Polynome (von beliebig großem, aber endlichen Grad) bezeichnet man mit $\mathbb{R}[x]$. Die Addition und Skalarenmultiplikation wird wie in Beispiel 1.7 definiert. Damit wird dieser Polynomraum $\mathbb{R}[x]$ ein $\mathbb{R}$-Vektorraum.

Es sei M eine beliebige unendliche Menge und $f : M \to \mathbb{R}$ eine Abbildung. Sowas nennt man auch reelle Funktion auf der Menge M. Die Menge all dieser Funktionen bezeichen wir mit $Abb(M, \mathbb{R})$. In dieser Allgemeinheit ist sie zu nichts gut, außer als Beispiel für einen i.A. echt großen $\mathbb{R}$-Vektorraum. Die beiden Rechenoperationen für Funktionen werden dabei wie in Beispiel 1.7 punktweise erklärt.

Die Rechenstrukturen in Beispiel 1.10 kann man nicht so mit den Vektorräumen $\mathbb{R}^n$ identifizieren, wie es in den Beispielen 1.7-1.9 gelungen ist. Deswegen sind die Regeln (F.V) keineswegs intuitiv klar. Aber man kann sie aus den Regeln $A.V1 - A.V3, M.V1 - M.V4$ herleiten.

Satz 1.10 *Die Formeln (F.V) gelten in jedem beliebigen $\mathbb{R}$-Vektorraum $(V, +, \cdot)$.*

Beweis. Es sei $\mathbf{x} \in V$ und $c \in \mathbb{R}$.

$0 \cdot \mathbf{x} = \mathbf{0}$: Wegen M.V1 (erstes Distributivgesetz) ist

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

Subtrahiert man den Vektor $0 \cdot \mathbf{x}$ auf beiden Seiten dieser Gleichung, so erhält man die zu zeigende Formel.

$-\mathbf{x} = (-1) \cdot \mathbf{x}$: Wieder mit dem Distributivgesetz folgt dies aus

$$\mathbf{x} + (-1) \cdot \mathbf{x} = (1 - 1) \cdot \mathbf{x} = 0 \cdot \mathbf{x} = \mathbf{0}.$$

$c \cdot \mathbf{0} = \mathbf{0}$: Mit dem zweiten Distributivgesetz berechnen wir

$$c \cdot \mathbf{0} = c \cdot (\mathbf{0} + \mathbf{0}) = c \cdot \mathbf{0} + c \cdot \mathbf{0}$$

und erhalten die behauptete Formel indem wir auf beiden Seiten dieser Gleichung den Vektor $c \cdot \mathbf{0}$ subtrahieren.

$c \cdot \mathbf{x} = \mathbf{0} \Leftrightarrow c = 0$ oder $\mathbf{x} = \mathbf{0}$: Es ist nur noch die Richtung „$\Rightarrow$“ zu zeigen. Entweder ist $c = 0$ oder wir können $c \neq 0$ annehmen. Dann schreiben wir

$$\mathbf{x} = 1 \cdot \mathbf{x} = (\frac{1}{c} \cdot c) \cdot \mathbf{x} = \frac{1}{c} \cdot (c \cdot \mathbf{x}) = \frac{1}{c} \cdot \mathbf{0} = \mathbf{0}.$$

□

1.3 Untervektorräume

Die Hyperebene $H \subset \mathbb{R}^n$ sei gegeben durch die homogene Gleichung

$$h(\mathbf{x}) := a_1 x_1 + ... + a_n x_n = 0.$$

Die Funktion h hat offensichtlich die Eigenschaften

$$h(\mathbf{x}+\mathbf{y}) = h(\mathbf{x}) + h(\mathbf{y}) \quad \text{für} \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \qquad h(c \cdot \mathbf{x}) = c \cdot h(\mathbf{x}) \quad \text{für} \quad c \in \mathbb{R}, \mathbf{x} \in \mathbb{R}^n.$$

Daraus folgt

$$\mathbf{x}, \mathbf{y} \in H,\ s, t \in \mathbb{R} \quad \Rightarrow s \cdot \mathbf{x} + t \cdot \mathbf{h} \in H.$$

Sei nun ein *homogenes* lineares Gleichungssystem

$$\sum_{\nu=1}^{n} a_{\mu,\nu} x_\nu = 0, \quad (\mu = 1, ..., m)$$

gegeben. Seine Lösungsmenge

$$U = \{\mathbf{u} \in \mathbb{R}^n : \sum_{\nu=1}^{n} a_{\mu,\nu} u_\nu = 0, \mu = 1, ..., m\} \subset \mathbb{R}^n$$

ist Durchschnitt von Hyperebenen H wie eben. Deswegen gilt also:

$$\boxed{\mathbf{x}, \mathbf{y} \in U,\ s, t \in \mathbb{R} \Rightarrow s\mathbf{x} + t\mathbf{y} \in U \qquad \text{(LIN)}}$$

Diese Eigenschaft (LIN) kann auch in zwei Teilen geschrieben werden:

$$\boxed{\begin{array}{lcll} \mathbf{x}, \mathbf{y} \in U & \Rightarrow & \mathbf{x} + \mathbf{y} \in U & \text{(LIN, add)} \\ \mathbf{x} \in U, c \in \mathbb{R} & \Rightarrow & c\mathbf{x} \in U & \text{(LIN, mul)} \end{array}}$$

Sie ist für die Lineare Algebra so wichtig, dass wir sie durch eine Definition hervorheben:

Definition 1.10 *Es sei V ein $\mathbb{R}$-Vektorraum. Eine Teilmenge $\emptyset \neq U \subset V$ heißt* linearer Unterraum *oder* Untervektorraum *wenn sie die Eigenschaft (LIN) besitzt.*

Bevor wir weitere Beispiele geben, notieren wir, dass jeder lineare Unterraum U den Nullvektor enthält: Denn weil U nicht leer ist, enthält U mindestens einen Vektor $\mathbf{x}$, und dann wegen (LIN) auch den Nullvektor $\mathbf{0} = 0 \cdot \mathbf{x}$. Die Rechenregeln für Addition $'+'$ und Multiplikation $'\cdot'$ gelten auch in U, weil sie in V gelten. Damit wird die Menge U mit diesen beiden Rechenoperationen selbst ein $\mathbb{R}$-Vektorraum.

Beispiel 1.11 *Sei V ein $\mathbb{R}$-Vektorraum. Eine Gerade $L \subset V$ durch den Nullpunkt ist (in Verallgemeinerung von Definition 1.6) eine Menge der Form*

$$L = \{w \cdot \mathbf{w},\ w \in \mathbb{R}\} \quad \textit{mit} \quad \mathbf{0} \neq \mathbf{w} \in V.$$

Sie ist ein linearer Unterraum: Seien $\mathbf{x} = w_1 \cdot \mathbf{w}$ und $\mathbf{y} = w_2 \cdot \mathbf{w}$ Vektoren in L. Dann ist auch

$$\mathbf{x} + \mathbf{y} = (w_1 + w_2) \cdot \mathbf{w} \in L, \quad c \cdot \mathbf{x} = (c \cdot w_1) \cdot \mathbf{w} \in L.$$

Beispiel 1.12 *Eine Ebene $E \subset V$ durch den Nullpunkt ist (in Verallgemeinerung von Definition 1.7) eine Menge der Form*

$$E = \{v \cdot \mathbf{v} + w \cdot \mathbf{w} : \, v, w \in \mathbb{R}\}$$

(wo $\mathbf{v}$ *und* $\mathbf{w}$ *keine Vielfache voneinander sind). Wie in Beispiel 1.11 zeigt man: $E \subset V$ ist ein Untervektorraum.*

Beispiel 1.13 *Aus ganz trivialen Gründen sind der* Nullraum $\{\mathbf{0}\}$, *der nur den Nullvektor enthält, und der* Totalraum V *der alle Vektoren enthält, lineare Unterräume.*

Beispiel 1.14 *Sind U_1 und $U_2 \subset V$ zwei lineare Unterräume, so ist auch ihr Durchschnitt $U_1 \cap U_2 \subset V$ ein linearer Unterraum. Ihre Vereinigung ist i.A. kein linearer Unterraum. Dagegen definiert man*

$$U_1 + U_2 = \{\mathbf{u}_1 + \mathbf{u}_2 \in V : \, \mathbf{u}_1 \in U_1, \mathbf{u}_2 \in U_2\}.$$

Diese Menge ist ein linearer Unterraum: Mit $\mathbf{u}_1 + \mathbf{u}_2 \in U_1 + U_2$ und $\mathbf{u}'_1 + \mathbf{u}'_2 \in U_1 + U_2$ gehört auch

$$(\mathbf{u}_1 + \mathbf{u}_2) + (\mathbf{u}'_1 + \mathbf{u}'_2) = (\mathbf{u}_1 + \mathbf{u}'_1) + (\mathbf{u}_2 + \mathbf{u}'_2) \quad zu \quad U_1 + U_2.$$

Und für $c \in \mathbb{R}$ ist

$$c \cdot (\mathbf{u}_1 + \mathbf{u}_2) = c \cdot \mathbf{u}_1 + c \cdot \mathbf{u}_2 \in U_1 + U_2$$

offensichtlich.

Definition 1.11 *Es sei $A \subset V$ eine beliebige (endliche oder unendliche) aber nicht leere Teilmenge. Jede endliche Summe*

$$\mathbf{x} = \sum_{\nu=1}^{k} c_\nu \mathbf{a}_\nu, \quad c_\nu \in \mathbb{R}, \mathbf{a}_\nu \in A, k \in \mathbb{N},$$

nennen wir eine Linearkombination *von Vektoren aus A. Und die Menge aller Linearkombinationen von Vektoren aus A*

$$span(A) := \left\{ \sum_{\nu=1}^{k} c_\nu \mathbf{a}_\nu \, : k \in \mathbb{N}, c_\nu \in \mathbb{R}, \mathbf{a}_\nu \in A \right\}$$

heißt der von A aufgespannte Unterraum. *Für endliche Mengen $A = \{\mathbf{a}_1, ..., \mathbf{a}_k\}$ benützen wir dabei immer die Abkürzung*

$$span(\mathbf{a}_1, ..., \mathbf{a}_k) := span(\{\mathbf{a}_1, ..., \mathbf{a}_k\}).$$

Satz 1.11 *1) Die Menge $span(A)$ ist der kleinste lineare Unterraum von V, der die Menge A enthält, d.h.:*

(i): $span(A)$ ist ein linearer Unterraum,

(ii): jeder lineare Unterraum $U \subset V$, der A enthält, enthält auch $span(A)$.

2) Sind $A_1 \subset A_2 \subset V$ zwei nicht-leere Teilmengen. Dann gilt auch $span(A_1) \subset span(A_2)$.

3) Sind $A_1, A_2 \subset V$ beliebige nichtleere Mengen, dann gilt

$$span(A_1 \cup A_2) = span(A_1) + span(A_2).$$

Insbesondere ist für lineare Unterräume $U_1, U_2 \subset V$

$$span(U_1 \cup U_2) = U_1 + U_2.$$

Beweis 1) (i): Seien $\mathbf{x}=\sum_1^k c_\mu \mathbf{a}_\mu$ und $\mathbf{y}=\sum_1^l d_\nu \mathbf{a}'_\nu$ Elemente in *span*(A). Dann ist auch $s\mathbf{x}+t\mathbf{y}=\sum_1^k sc_\mu \mathbf{a}_\mu+\sum_1^l td_\nu \mathbf{a}'_\nu$ eine Linearkombination von Vektoren $\mathbf{a}_\mu, \mathbf{a}'_\nu \in A$ und gehört zu *span*(A).

(ii): Enthält der lineare Unterraum $U \subset \mathbb{R}^n$ die Menge A, so wegen wiederholter Anwendung von (LIN) auch jede endliche Linearkombination von Vektoren aus A, und damit die Menge *span*(A).

2) Es ist $A_1 \subset A_2 \subset span(A_2)$. Weil $span(A_2)$ ein linearer Unterraum ist, folgt die Behauptung aus 2).

3) Weil $A_1 \cup A_2$ in dem linearen Unterraum $span(A_1)+span(A_2)$ enthalten ist, folgt die Inklusion $span(A_1 \cup A_2) \subset span(A_1)+span(A_2)$ aus 1). Wegen $A_1 \subset (A_1 \cup A_2)$ ist $span(A_1) \subset span(A_1 \cup A_2)$ nach 2). Analog gilt $span(A_2) \subset span(A_1) \cup span(A_2)$. Weil $span(A_1 \cup A_2) \subset V$ ein linearer Unterraum ist, ist dann auch jede Summe von Vektoren aus A_1 und A_2 diesem Unterraum enthalten. Insbesondere gilt auch die Inklusion $span(A_1)+span(A_2) \subset span(A_1 \cup A_2)$.

Sind $A_1=U_1$ und $A_2=U_2$ lineare Unterräume, so ist $span(U_1)=U_1$ und $span(U_2)=U_2$. Nach dem Bewiesenen ist also

$$span(U_1 \cup U_2)=U_1+U_2. \qquad \square$$

Wir betrachten Spezialfälle für derart aufgespannte lineare Unterräume.

Beispiel 1.15 *Eine Gerade* $\mathbb{R}\cdot\mathbf{w} \subset V$ *durch* $\mathbf{0}$ *ist die Menge* $span(\mathbf{w})$. *Eine Ebene* $\mathbb{R}\cdot\mathbf{v}+\mathbb{R}\cdot\mathbf{w} \subset V$ *ist* $span(\mathbf{v},\mathbf{w})$. *Sind* $\mathbf{v},\mathbf{w} \subset V$ *Vektoren mit* $\mathbf{w} \neq \mathbf{0}$ *und* $\mathbf{v}=c\cdot\mathbf{w}$ *für ein* $c \in \mathbb{R}$, *dann ist* $span(\mathbf{v},\mathbf{w})=span(\mathbf{w})$ *eine Gerade und keine Ebene.*

Beispiel 1.16 *Mit* $\mathbf{e}_\nu \in \mathbb{R}^n$ *werden wir stets den Vektor bezeichnen, der an der* ν*-ten Stelle den Eintrag 1 enthält und sonst lauter Nullen:*

$$\begin{array}{ccccccccccc} \mathbf{e}_\nu & = & (& 0 & ,..., & 0, & 1, & 0, & ..., & 0 &)^t \\ & & & \uparrow & & & \uparrow & & & \uparrow & \\ & & & 1 & & & \nu & & & n & \end{array}$$

Die Vektoren $\mathbf{e}_\nu$ *heißen* Einheitsvektoren *im* $\mathbb{R}^n$. *Für* $k=1,...,n$ *ist dann*

$$\begin{aligned} span(\mathbf{e}_1,...,\mathbf{e}_k) &= \left\{\mathbf{x}=\sum_1^k c_\nu \mathbf{e}_\nu\right\} \\ &= \{\mathbf{x}=(c_1,...,c_k,0,...,0)^t\} \\ &= \{\mathbf{x} \in \mathbb{R}^n : x_{k+1}=...=x_n=0\}. \end{aligned}$$

Beispiel 1.17 *Es sei* $U=U_1+U_2$. *Dann gibt es zu jedem* $\mathbf{u} \in U$ *eine Darstellung*

$$\mathbf{u}=\mathbf{u}_1+\mathbf{u}_2 \quad \textit{mit} \quad \mathbf{u}_1 \in U_1, \mathbf{u}_2 \in U_2.$$

Behauptung: Diese Darstellung ist für alle $\mathbf{u} \in U$ *eindeutig, genau dann, wenn* $U_1 \cap U_2=\{\mathbf{0}\}$.

Beweis. Sei die Darstellung eindeutig. Für jeden Vektor $\mathbf{u} \in U_1 \cap U_2$ *hat man dann aber zwei Darstellungen*

$$\begin{aligned} \mathbf{u} &= \mathbf{u}_1+\mathbf{0} \quad \textit{mit} \quad \mathbf{u}_1=\mathbf{u} \in U_1 \\ &= \mathbf{0}+\mathbf{u}_2 \quad \textit{mit} \quad \mathbf{u}_2=\mathbf{u} \in U_2 \end{aligned}$$

Aus der Eindeutigkeit der Darstellung folgt $\mathbf{u}=\mathbf{u}_1=\mathbf{0}$. *Also ist* $U_1 \cap U_2$ *der Nullraum.*

Sei umgekehrt $U_1 \cap U_2=\{\mathbf{0}\}$. *Es gebe einen Vektor* $\mathbf{u}$ *mit zwei Darstellungen*

$$\mathbf{u}=\mathbf{u}_1+\mathbf{u}_2=\mathbf{u}'_1+\mathbf{u}'_2 \;\textit{mit}\; \mathbf{u}_1,\mathbf{u}'_1 \in U, \mathbf{u}_2,\mathbf{u}'_2 \in U_2.$$

Daraus folgt

$$\mathbf{u}_1 - \mathbf{u}_1' = \mathbf{u}_2' - \mathbf{u}_2 \in U_1 \cap U_2 = \{\mathbf{0}\},$$

also $\mathbf{u}_1 = \mathbf{u}_1'$ *und* $\mathbf{u}_2 = \mathbf{u}_2'$. □

Ist die eben diskutierte Bedingung erfüllt, dann heißt die Summe $U_1 + U_2$ direkt*, bzw. diese 'Zerlegung von* U *in* U_1 *und* U_2*' eine direkte-Summen-Zerlegung. Man schreibt dann*

$$U = U_1 \oplus U_2.$$

Beispiel 1.18 *Im Vektorraum* $\mathbb{R}[x]$ *der Polynome betrachten wir die Potenzen* x^i*. (Man nennt sie auch* Monome.*) Dann ist*

$$span(1 = x^0, ..., x^n) = \mathbb{R}_n[x]$$

der Untervektorraum der Polynome vom Grad $\leq n$ *und*

$$span\{x^i,\, i \in \mathbb{N}\} = \mathbb{R}[x]$$

der Raum aller Polynome.

Beispiel 1.19 *Es sei* $g = (1-x)^2 \in \mathbb{R}[x]$*. Dann ist*

$$span(x^0, x^1, x^2, g) = span(x^0, x^1, x^2) = \mathbb{R}_2[x].$$

Beispiel 1.20 *Es sei* $T(\Delta)$ *der Vektorraum der Treppenfunktionen zur Zerlegung* $a = x_0 < x_1 < ... < x_{n-1} < x_n = b$ *(Beispiel 1.8). Wir definieren Treppenfunktionen* $t_1, ..., t_n \in T$ *durch*

$$t_i(x) := \left\{ \begin{array}{l} 1 \text{ falls } x \in [x_{i-1}, x_i[\\ 0 \text{ sonst} \end{array} \right\} (i = 1, ..., n), \qquad t_n(x) := \left\{ \begin{array}{l} 1 \text{ falls } x \in [x_{n-1}, x_n] \\ 0 \text{ sonst} \end{array} \right\}$$

Dann ist $T = span(t_1, ..., t_n)$.

Mit diesem Begriff des „aufgespannten Unterraums“ können wir die Lösbarkeitsbedingung für ein lineares Gleichungssystem

$$\sum_{\nu=1}^{n} a_{\mu,\nu} x_\nu = b_\mu, \quad \mu = 1, ..., m$$

anders formulieren: Wir bezeichnen mit $\mathbf{a}^\nu \in \mathbb{R}^m$ die Spaltenvektoren der Koeffizientenmatrix und mit $\mathbf{b}$ den Vektor auf der rechten Seite des Gleichungssystems:

$$\mathbf{a}^\nu = \begin{pmatrix} a_{1,\nu} \\ \vdots \\ a_{m,\nu} \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

Mit diesen Vektoren kann man das Gleichungssystem in Vektorschreibweise

$$\sum_{\nu=1}^{n} x_\nu \mathbf{a}^\nu = \mathbf{b}$$

schreiben. Man sieht:

Satz 1.12 *Das Gleichungssystem ist genau dann lösbar, wenn die rechte Seite* $\mathbf{b}$ *eine Linearkombination der Spaltenvektoren* $\mathbf{a}^1, ..., \mathbf{a}^n$ *ist, d.h.,wenn*

$$\mathbf{b} \in span\left(\mathbf{a}^1, ..., \mathbf{a}^n\right).$$

Die zeilenweise Sicht eines LGS mit m Gleichungen und n Unbekannten besteht in der Aufgabe:
- Beschreibe den Durchschnitt von m Hyperebenen im $\mathbb{R}^n$.

Demgegenüber ist die spaltenweise Sicht:
- Stelle die rechte Seite $\mathbf{b}$ als eine Linearkombination von n Spaltenvektoren im $\mathbb{R}^m$ dar.

Andersherum gesehen, haben wir ein Verfahren gefunden, um zu prüfen, ob ein Vektor $\mathbf{b} \in \mathbb{R}^m$ eine Linearkombination von gegebenen Vektoren $\mathbf{a}_1, ..., \mathbf{a}_n \in \mathbb{R}^m$ ist: Man definiere ein Koeffizientenmatrix $A \in M(m \times n)$ mit den gegebenen Vektoren $\mathbf{a}_\nu$ als Spalten und prüfe mit dem Gaußschen Eliminationsverfahren das durch $(A, \mathbf{b})$ gegebene LGS auf Lösbarkeit.

Die spaltenweise Sicht des LGS nehmen wir als Anlass, das Produkt einer Matrix $A \in M(m \times n)$ mit einem Vektor $\mathbf{x} \in \mathbb{R}^n$ zu definieren.

Definition 1.12 *Es sei* $A = (\mathbf{a}^1, ..., \mathbf{a}^n) \in M(m \times n)$ *eine Matrix mit den Spaltenvektoren* $\mathbf{a}^\nu$, $\nu = 1, ..., n$. *Ihr* Matrix-Vektor-Produkt *mit dem Vektor* $\mathbf{x} = (x_\nu) \in R^n$ *ist*

$$A \cdot \mathbf{x} := \sum_{\nu=1}^{n} x_\nu \mathbf{a}^\nu \in \mathbb{R}^m,$$

die Linearkombination der Vektoren $\mathbf{a}^\nu$ *mit den Koeffizienten* x_ν.

Der Sinn dieser Definition besteht für uns momentan darin, dass wir jetzt ein LGS mit Koeffizientenmatrix $A \in M(m \times n)$ und rechter Seite $\mathbf{b} \in \mathbb{R}^m$ als Vektorgleichung

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

für den gesuchten Vektor $\mathbf{x} \in \mathbb{R}^n$ schreiben können.

Satz 1.13 (Rechenregeln) *Für Matrizen* $A, B \in M(m \times n)$, *Vektoren* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *und Skalare* $c \in \mathbb{R}$ *gelten*

$$1)\, A \cdot (\mathbf{x} + \mathbf{y}) = A \cdot \mathbf{x} + A \cdot \mathbf{y}, \qquad 2)\, (A + B) \cdot \mathbf{x} = A \cdot \mathbf{x} + B \cdot \mathbf{x},$$
$$3)\, (c \cdot A) \cdot \mathbf{x} = c \cdot (A \cdot \mathbf{x}) = A \cdot (c \cdot \mathbf{x}).$$

Wir beweisen nur die Gleichung 1), die anderen ergeben sich ähnlich. Dazu seien $\mathbf{a}^1, ..., \mathbf{a}^n$ die Spaltenvektoren der Matrix A. Nach Definition ist

$$A \cdot (\mathbf{x} + \mathbf{y}) = \sum_{\nu=1}^{n} (x_\nu + y_\nu) \cdot \mathbf{a}^\nu = \sum_{\nu=1}^{n} x_\nu \cdot \mathbf{a}^\nu + y_\nu \cdot \mathbf{a}^\nu = \sum_{\nu=1}^{n} x_\nu \cdot \mathbf{a}^\nu + \sum_{\nu=1}^{n} y_\nu \cdot \mathbf{a}^\nu = A \cdot \mathbf{x} + B \cdot \mathbf{x}.$$

$\square$

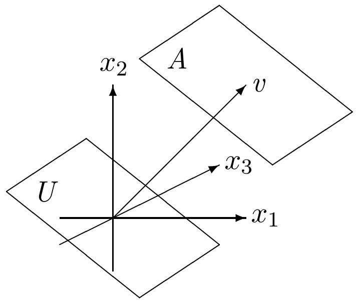
Schließlich treffen wir noch eine Vereinbarung, die an dieser Stelle überperfektionistisch erscheinen mag: Wenn die Menge A leer ist, so vereinbaren wir *span*(A) soll der Nullraum sein, d.h. der lineare Unterraum, welcher nur den Nullvektor enthält: *span*$(\emptyset) = \{\mathbf{0}\}$.

Manchmal bezeichnet man auch Lösungsmengen inhomogener Gleichungssysteme als Unterräume. Diese besitzen dann natürlich nicht die Eigenschaft (LIN). Wir werden solche Mengen *affine Unterräume* nennen.

Definition 1.13 *Es sei U ein Untervektorraum des $\mathbb{R}$-Vektorraums V und $\mathbf{v} \in V$. Dann heißt*

$$A = \{\mathbf{x} = \mathbf{v} + \mathbf{u} : \mathbf{u} \in U\} = \mathbf{v} + U$$

affiner Unterraum von V.



Es gilt also

Satz 1.14 *Die Lösungsmenge U eines LGS mit n Unbekannten ist ein affiner Unterraum des $\mathbb{R}^n$. U ist genau dann ein linearer Unterraum, wenn das LGS homogen ist.*

Im Wesentlichen ist das genau die Aussage des Struktursatzes 1.4.

Es seien $A_1 = \mathbf{v}_1 + U_1$ und $A_2 = \mathbf{v}_2 + U_2$ zwei affine Unterräume des $\mathbb{R}^n$, wobei

$$U_1 = span(\mathbf{a}_1, ..., \mathbf{a}_k),\, U_2 = span(\mathbf{a}_{k+1}, ..., a_m) \quad \text{mit} \quad \mathbf{a}_1, ..., \mathbf{a}_m \in \mathbb{R}^n.$$

Für den Durchschnitt $A = A_1 \cap A_2$ dieser affinen Unterräume gilt dann

$$\begin{aligned} \mathbf{v} \in A \quad &\Leftrightarrow \quad \text{es gibt } x_1, ..., x_m \in \mathbb{R} \text{ mit } \mathbf{v}_1 + \sum_{i=1}^{k} x_i \mathbf{a}_i = \mathbf{v}_2 + \sum_{i=k+1}^{m} x_i \mathbf{a}_i \\ &\Leftrightarrow \quad \sum_{i=1}^{k} x_i \mathbf{a}_i - \sum_{i=k+1}^{m} x_i \mathbf{a}_i = \mathbf{v}_2 - \mathbf{v}_1 \end{aligned}$$

Die Bestimmung aller Vektoren $\mathbf{v} \in A$ führt also darauf, das LGS mit Koeffizientenmatrix

$$(\mathbf{a}_1, ..., \mathbf{a}_k, -\mathbf{a}_{k+1}, ..., -\mathbf{a}_m)$$

und rechter Seite $\mathbf{b} = \mathbf{v}_2 - \mathbf{v}_1$ zu bestimmen.

Ist $A = \mathbf{v} + U$ ein affiner Unterraum und $\mathbf{w} \in A$ beliebig, dann ist auch $A = \mathbf{w} + U$. Sind $A_1 = \mathbf{v}_1 + U_1$ und $A_2 = \mathbf{v}_2 + U_2$ affine Unterräume, dann ist $A := A_1 \cap A_2$ entweder leer oder

$$A = \mathbf{a} + (U_1 \cap U_2) \quad \text{mit einem beliebigen Vektor } \mathbf{a} \in A.$$

Es gibt lineare Unterräume verschiedener Größe:

$\{\mathbf{0}\}$	Gerade	Ebene	...
0-dimensional	1-dimensional	2-dimensional	...

Diese Größe nennt man „Dimension“ eines linearen Unterraums. Der folgende Abschnitt dient u. a. der präzisen Definition des Dimensionsbegriffs.

Aufgabe 1.28 *Betrachten Sie die acht Mengen von Vektoren* $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$ *definiert durch die Bedingungen*

1) $x_1 + x_2 = 0$,
2) $(x_1)^2 + (x_2)^2 = 0$,
3) $(x_1)^2 - (x_2)^2 = 0$,
4) $x_1 - x_2 = 1$,
5) $(x_1)^2 + (x_2)^2 = 1$,
6) *Es gibt ein* $t \in \mathbb{R}$ *mit* $x_1 = t$ *und* $x_2 = t^2$,
7) *Es gibt ein* $t \in \mathbb{R}$ *mit* $x_1 = t^3$ *und* $x_2 = t^3$,
8) $x_1 \in \mathbb{Z}$.

Welche dieser Mengen sind lineare Unterräume?

Aufgabe 1.29 *Liegt der Vektor* $(3, -1, 0, -1) \in \mathbb{R}^4$ *im Unterraum, der von den Vektoren* $(2, -1, 3, 2)$, $(-1, 1, 1, -3)$ *und* $(1, 1, 9, -5)$ *aufgespannt wird?*

Aufgabe 1.30 *Es seien* U_1 *und* $U_2 \subset V$ *Untervektorräume des* $\mathbb{R}$*-Vektorraums* V. *Zeigen Sie:* $U_1 \cup U_2 \subset V$ *ist genau dann ein Untervektorraum, wenn entweder* $U_1 \subset U_2$ *oder* $U_2 \subset U_1$.

Aufgabe 1.31 *Es sei* V *ein* $\mathbb{R}$*-Vektorraum und* $U \subset V$ *ein Untervektorraum. Zeigen Sie:*

a) Für $\mathbf{v}, \mathbf{v}' \in V$ *ist* $\mathbf{v} + U = \mathbf{v}' + U$ *genau dann, wenn* $\mathbf{v} - \mathbf{v}' \in U$. *Insbesondere ist* $\mathbf{v} + U = U$ *genau dann, wenn* $\mathbf{v} \in U$.

b) Auf V *wird durch*

$$\mathbf{v} \sim \mathbf{v}' \Leftrightarrow \mathbf{v} - \mathbf{v}' \in U \Leftrightarrow \mathbf{v} + U = \mathbf{v}' + U \quad \textit{für } \mathbf{v}, \mathbf{v}' \in V$$

eine Äquivalenzrelation *eingeführt, d.h., es gelten*

Reflexivität: $\mathbf{v} \sim \mathbf{v}$ *für alle* $\mathbf{v} \in V$,
Symmetrie: $\mathbf{v} \sim \mathbf{v}' \quad \Rightarrow \quad \mathbf{v}' \sim \mathbf{v}$,
Transitivität: $\mathbf{v} \sim \mathbf{v}'$ *und* $\mathbf{v}' \sim \mathbf{v}'' \quad \Rightarrow \quad \mathbf{v} \sim \mathbf{v}''$.

Aufgabe 1.32 *Zeigen Sie: Durch die folgenden Bedingungen an Polynome* $p \in \mathbb{R}[x]$ *werden Untervektorräume von* $\mathbb{R}[x]$ *definiert:*

$$1)\, p(0) = 0, \quad 2)\, p(2x) = 4p(x), \quad 3)\, p(-x) = p(x), \quad 4)\, p(-x) = -p(x).$$

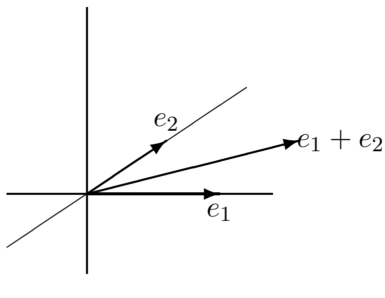
1.4 Lineare (Un-) Abhängigkeit

1.4.1 Dimension

Beispiel 1.21 *Die beiden Vektoren* $\mathbf{e}_1 = (1, 0, 0)$ *und* $\mathbf{e}_2 = (0, 1, 0) \in \mathbb{R}^3$ *spannen die Ebene* $\{\mathbf{x} \in \mathbb{R}^3 : x_3 = 0\}$ *auf. Dieselbe Ebene wird aber auch von den drei Vektoren*

$$\mathbf{e}_1,\, \mathbf{e}_2,\, \mathbf{e}_1 + \mathbf{e}_2 = (1, 1, 0)$$

aufgespannt. Jeden dieser drei Vektoren könnte man weglassen, die restlichen beiden spannen diese Ebene immer noch auf. Wir sagen: Diese drei Vektoren sind linear abhängig.



Im Folgenden sei V ein beliebiger $\mathbb{R}$-Vektorraum.

Definition 1.14 *Eine Menge $A \subset \mathbb{R}^n$ heißt* linear abhängig, *wenn es eine echte Teilmenge $A' \subset A$, $A' \neq A$ gibt mit $span(A') = span(A)$. Sonst heißt A* linear unabhängig.

Beispiel 1.22 *1) Die oben betrachtete Menge $A = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_1 + \mathbf{e}_2\} \subset \mathbb{R}^3$ ist linear abhängig, denn für $A' = \{\mathbf{e}_1, \mathbf{e}_2\} \subset A$ gilt $A' \neq A$ und $span(A') = span(A)$.*

2) Die Menge $A = \{\mathbf{e}_1, \mathbf{e}_2\}$ enthält die folgenden echten Teilmengen:
$A' = \{\mathbf{e}_1\}$ mit $span(\mathbf{e}_1) =$ Gerade $\mathbb{R} \cdot \mathbf{e}_1$,
$A' = \{\mathbf{e}_2\}$ mit $span(\mathbf{e}_2) =$ Gerade $\mathbb{R} \cdot \mathbf{e}_2$,
$A' = \emptyset$ mit $span(\emptyset) =$ Nullraum.
Für keine davon gilt $span(A') = span(A) =$ Ebene $\{x_3 = 0\}$. Also ist A linear unabhängig.

Beispiel 1.23 *Jede Menge, welche den Nullvektor enthält, ist linear abhängig, denn wenn $\mathbf{0} \in A$ und $A' = A \setminus \{\mathbf{0}\}$, dann ist $A' \neq A$, aber $span(A') = span(A)$.*

Beispiel 1.24 *Enthält A einen Vektor $\mathbf{a}$ mit $\mathbf{a} \in span(A \setminus \{\mathbf{a}\})$, dann ist A linear abhängig. Denn für $A' := A \setminus \{\mathbf{a}\}$ gilt $A \neq A'$, aber wegen $\mathbf{a} = \sum_1^l d_j \mathbf{a}_j, \mathbf{a}_j \in A'$,*

$$\begin{aligned} span(A) &= \{c_0\mathbf{a} + \sum_1^k c_m\mathbf{b}_m \ : k \in \mathbb{N}, c_0, c_1, ..., c_k \in \mathbb{R}, \mathbf{b}_m \in A'\} \\ &= \{c_0(\sum_1^l d_j\mathbf{a}_j) + \sum_1^k c_m\mathbf{b}_m \ : \ \mathbf{a}_j, \mathbf{b}_m \in A'\} \\ &\subset span(A'). \end{aligned}$$

Beispiel 1.25 *Jede Obermenge einer linear abhängigen Menge ist linear abhängig. Jede Teilmenge einer linear unabhängigen Menge ist linear unabhängig.*

Beispiel 1.26 *Wenn (voneinander verschiedene) Vektoren $\mathbf{v}_1, ..., \mathbf{v}_k \in A$ existieren und Zahlen $c_1, ..., c_k \in \mathbb{R}$, nicht $c_1 = ... = c_k = 0$, mit*

$$\boxed{\sum_{m=1}^k c_m\mathbf{v}_m = \mathbf{0}, \qquad \textit{(nicht-triviale lineare Relation)}}$$

dann ist A linear abhängig. Denn weil nicht alle $c_m = 0$ sind, können wir nach Vertauschen der Indizes annehmen $c_1 \neq 0$ und dann schreiben

$$c_1\mathbf{v}_1 = -\sum_2^k c_m\mathbf{v}_m, \qquad \mathbf{v}_1 = \sum_2^k -\frac{c_m}{c_1}\mathbf{v}_m \in span(A'),$$

wo $A' := A \setminus \{\mathbf{v}_1\}$.

Diese Beispiele sollten zunächst den Sachverhalt der linearen Abhängigkeit verdeutlichen. Das letzte Beispiel ist aber bereits typisch dafür, wie wir künftig lineare Un-/Abhängigkeit überprüfen werden:

Satz 1.15 (Test auf lineare Abhängigkeit) *Eine Teilmenge $A \subset \mathbb{R}^n$ ist genau dann linear abhängig, wenn es eine nichttriviale lineare Relation zwischen (voneinander verschiedenen) Vektoren aus A gibt.*

Satz 1.16 (Test auf lineare Unabhängigkeit) *Eine Teilmenge $A \subset \mathbb{R}^n$ ist genau dann linear unabhängig, wenn sie folgende Eigenschaft besitzt:*

> *Sind $\mathbf{v}_1, ..., \mathbf{v}_k$ endlich viele (voneinander paarweise verschiedene) Vektoren in A und $c_1, ..., c_k$ Zahlen in $\mathbb{R}$ mit*
>
> $$\boxed{\sum_{m=1}^{k} c_m \mathbf{v}_m = \mathbf{0}}$$
>
> *dann ist $c_1 = ... = c_k = 0$.*

Satz 1.16 ist nur eine Umformulierung von Satz 1.15. Deswegen genügt es, Satz 1.15 zu beweisen.

Beweis von Satz 1.15. „$\Leftarrow$“ Diese Beweisrichtung wurde als Beispiel 1.26 oben schon behandelt.

„$\Rightarrow$“ Sei A linear abhängig, d.h., es gibt eine Teilmenge $A' \subset A$ mit $span(A') = span(A)$ und $A' \neq A$. Dann gibt es also einen Vektor $\mathbf{v} \in A$, der nicht zu A' gehört. Wegen $\mathbf{v} \in A \subset span(A) = span(A')$ ist $\mathbf{v}$ eine Linearkombination $\mathbf{v} = \sum_1^k c_\nu \mathbf{v}_\nu$ von Vektoren $\mathbf{v}_\nu \in A'$. Dann ist

$$1 \cdot \mathbf{v} - \sum_{1}^{k} c_\nu \mathbf{v}_\nu = \mathbf{0}$$

eine nichttriviale (da $\mathbf{v}$ einen Koeffizienten $\neq 0$ hat) lineare Relation zwischen Vektoren aus A. $\square$

Noch zwei weitere Beispiele:

Beispiel 1.27 *Es sei $A \subset \mathbb{R}^n$ eine Teilmenge, die mehr als n Vektoren enthält. Dann ist A linear abhängig.*

Beweis. A enthält mindestens $n+1$ Vektoren $\mathbf{v}_1, ..., \mathbf{v}_{n+1}$. Das homogenen lineare Gleichungssystem

$$\begin{array}{ccccccc} c_1 \cdot v_{1,1} & + & ... & + & c_{n+1} \cdot v_{n+1,1} & = & 0 \\ \vdots & & & & \vdots & & \vdots \\ c_1 \cdot v_{1,n} & + & ... & + & c_{n+1} \cdot v_{n+1,n} & = & 0 \end{array}$$

aus n Gleichungen in den $n+1$ Unbekannten $c_1, ..., c_{n+1}$ hat nach Satz 1.3 eine Lösung $(c_1, ..., c_{n+1}) \neq (0, ..., 0)$. Damit haben wir eine nichttriviale lineare Relation $\sum_1^{n+1} c_\nu \mathbf{v}_\nu = \mathbf{0}$ zwischen $\mathbf{v}_1, ..., \mathbf{v}_{n+1}$. Nach Satz 1.15 ist A linear abhängig. $\square$

Beispiel 1.28 *Es seien*

$$\begin{array}{rcllll} \mathbf{z}_1 & = & (0, ..., 0, & 1, ... & .... & ...) \\ \mathbf{z}_2 & = & (0, ..., 0, & 0, ..., 0, & 1, ... & ...) \\ \vdots & & & & & \\ \mathbf{z}_r & = & (0, ..., 0, & 0, ..., 0, & 0, ..., 0, & 1, ..) \end{array}$$

die ersten r Zeilenvektoren aus einer Matrix in Zeilenstufenform, wobei r die Anzahl der Stufen der Matrix ist. Diese Vektoren sind linear unabhängig.

Beweis. Der Vektor $\mathbf{z}_k$ habe seinen ersten Eintrag $\neq 0$ in der n_k-ten Spalte, $k = 1, ..., r$. Da die Matrix Zeilenstufenform hat, ist

$$1 \leq n_1 < n_2 < ... < n_r \leq n.$$

Wir testen auf lineare Unabhängigkeit: sei eine Linearkombination $\sum_1^r c_k \mathbf{z}_k = \mathbf{0}$ gegeben. Da nur der erste Vektor $\mathbf{z}_1$ in der n_1-ten Spalte einen Eintrag $\neq 0$ besitzt, folgt hieraus $c_1 = 0$. Von den übrigen Vektoren hat nur $\mathbf{z}_2$ einen Eintrag $\neq 0$ in der n_2-ten Spalte, was $c_2 = 0$ zur Folge hat, usw. □

Beispiel 1.27 lässt sich auf beliebige endlich-erzeugte IR-Vektorräume V übertragen:

Satz 1.17 *Es sei V ein IR-Vektorraum, der von $\mathbf{v}_1, ..., \mathbf{v}_n \in V$ aufgespannt wird. Sind $\mathbf{w}_1, ..., \mathbf{w}_{n+k}$ mit $k > 0$ Vektoren in V, so sind diese linear abhängig.*

Beweis. Die Vektoren $\mathbf{w}_i$ sind Linearkombinationen

$$\mathbf{w}_i = \sum_{j=1}^{n} a_{i,j} \mathbf{v}_j \quad (i = 1, ..., n+k)$$

der Vektoren $\mathbf{v}_j$. Wir betrachten die $n+k$ Vektoren

$$\mathbf{a}_i := (a_{i,1}, ..., a_{i,n})^t \in \mathbb{R}^n, \quad i = 1, ..., n+k.$$

Nach Beispiel 1.27 sind diese linear abhängig, es gibt also Koeffizienten $c_1, ..., c_{n+k} \in \mathbb{R}$, nicht alle $= 0$, mit

$$\sum_{i=1}^{n+k} c_i \mathbf{a}_i = \mathbf{0}, \quad \text{d.h.} \quad \sum_{i=1}^{n+k} c_i a_{i,j} = 0 \text{ für } j = 1, ..., n.$$

Mit diesen Koeffizienten erhalten wir

$$\sum_{i=1}^{n+k} c_i \mathbf{w}_i = \sum_{i=1}^{n+k} c_i \sum_{j=1}^{n} a_{i,j} \mathbf{v}_j = \sum_{j=1}^{n} \left(\sum_{i=1}^{n+k} c_i a_{i,j} \right) \mathbf{v}_j = \mathbf{0}.$$

□

Satz 1.17 ist das erste Auftreten eines allgemeinen Prinzips: Eine Aussage über beliebige IR-Vektorräume wir durch Betrachtung geeigneter Koeffizienten-Vektoren auf eine Aussage im $\mathbb{R}^n$ zurückgeführt.

Beispiel 1.29 *Wir betrachten im Vektorraum $T(\Delta)$ aus Beispiel 1.8 die in Beispiel 1.20 definierten Treppenfunktionen $t_1, ..., t_n$. Wir testen sie auf lineare Unabhängigkeit:*

$$\sum_{i=1}^{n} c_i t_i = 0, \quad \textit{d.h.,} \quad \sum_{i=1}^{n} c_i t_i(x) = 0 \textit{ für alle } x \in [a, b].$$

Im Punkt $x = a$ haben alle Treppenfunktionen t_i, $i > 1$, den Wert $= 0$. Setzen wir $x = a$ in obige Gleichung ein, so sehen wir

$$0 = \sum_{i=1}^{n} c_i t_i(a) = c_1.$$

Durch Einsetzen der anderen Zwischenpunkte $x_1, ..., x_{n-1}$ finden wir ähnlich $c_2 = ... = c_n = 0$.

Beispiel 1.30 *Die Monome $x^0 = 1, x, ..., x^n$ im Polynom-Vektorraum $\mathbb{R}_n[x]$ sind linear unabhängig. Dazu setzen wir unseren Test an:*

$$p(x) = \sum_{i=0}^{n} c_i x_i \equiv 0$$

sei das Nullpolynom. Wegen $p(0) = c_0$ finden wir $c_0 = 0$. Wir können p also durch x dividieren um

$$\frac{p(x)}{x} = \sum_{i=0}^{n-1} c_{i+1} x^i$$

zu erhalten. Das Polynom $p(x)/x$ hat den Wert 0 für alle $x \neq 0$. Jetzt müssen wir eine Anleihe bei der Analysis oder der Algebra aufnehmen. (In einem isolierten Beispiel ist das erlaubt.) Z.B. aus der Stetigkeit des Polynoms $p(x)/x$ folgt, dass es auch in $x = 0$ den Wert 0 hat. Damit folgt $c_1 = 0$ usw.

Beispiel 1.31 *Aber auch die unendlich vielen Monome x^i, $i \in \mathbb{N}$ im Vektorraum $\mathbb{R}[x]$ aller Polynome sind linear unabhängig. Nach Satz 1.16 brauchen wir nämlich immer nur endlich viele dieser Monome auf Lineare Unabhängigkeit zu testen. Und das haben wir in Beipiel 1.30 getan.*

Gelegentlich haben wir es nicht mit einer *Menge* $\{\mathbf{v}_1, \mathbf{v}_2, ...\}$ von Vektoren zu tun, sondern mit einer *Folge* $\mathbf{v}_1, \mathbf{v}_2, ...$, in der etwa Vektoren auch mehrmals vorkommen können. Eine solche (endliche oder unendliche) Folge werden wir auch *System* von Vektoren nennen und $[\mathbf{v}_1, \mathbf{v}_2, ...]$ schreiben. Genauer:

$$\begin{array}{ll} [\mathbf{v}_1, ..., \mathbf{v}_n] & \text{für ein endliches} \\ [\mathbf{v}_i]_{i\in\mathbb{N}} & \text{für ein unendliches} \end{array}$$

System, oder zusammenfassend $[\mathbf{v}_i]_{i\in I}$ für endliche oder abzählbare Indexmengen I. Die Zeilenvektoren einer Matrix sind z.B. so ein System.

Definition 1.14 kann wörtlich auf Systeme übertragen werden.

Definition 1.15 *Ein System $[\mathbf{v}_i]_{i\in I}$ von Vektoren in V heißt* linear abhängig, *wenn es ein $k \in I$ gibt mit*

$$\mathbf{v}_k \in span\{\mathbf{v}_i\}_{i\in I, i\neq k}.$$

Alle obigen Überlegungen übertragen sich auf Systeme. Insbesondere ist der Test auf lineare Unabhängigkeit für ein System:

$$\sum_{1}^{k} c_\nu \mathbf{v}_\nu = \mathbf{0} \quad \Rightarrow \quad c_1 = ... = c_k = 0 \quad ?$$

für alle $k \in \mathbb{N}$. Ein System, in dem derselbe Vektor mehrmals vorkommt, ist somit stets linear abhängig.

Definition 1.16 *Es sei V ein $\mathbb{R}$-Vektorraum. Eine* Basis *von V ist eine Menge $\{\mathbf{v}_i\}_{i\in I}$ von Vektoren aus V mit*

(i) $V = span\{\mathbf{v}_i\}_{i\in I}$,

(ii) die Menge $\{\mathbf{v}_i\}_{i\in I}$ ist linear unabhängig.

Die Zahl r heißt Länge der Basis.

Beispiel 1.32 *Für eine Gerade $\mathbb{R} \cdot \mathbf{v} \subset V$ durch $\mathbf{0}$ bildet der Vektor $\mathbf{v}$ eine Basis.*

Eine Basis für eine Ebene $\mathbb{R} \cdot \mathbf{v} + \mathbb{R} \cdot \mathbf{w}$ durch $\mathbf{0}$ wie in Beispiel 1.12 bilden die Vektoren $\mathbf{v}$ und $\mathbf{w}$ eine Basis. Wenn die Vektoren $\mathbf{v}, \mathbf{w}$ linear abhängig wären, wäre einer ein Vielfaches des anderen.

Sind die Vektoren $\mathbf{v}$ und $\mathbf{w}$ linear abhängig, dann ist $\mathbb{R} \cdot \mathbf{v} = \mathbb{R} \cdot \mathbf{w}$ eine Gerade.

Beispiel 1.33 *Die Vektoren* $\mathbf{e}_1, ..., \mathbf{e}_n$*, bilden eine Basis des* $\mathbb{R}^n$*. Wir nennen sie die* Standardbasis*, die Vektoren nennen wir* Koordinatenvektoren.

Beispiel 1.34 *Der Nullvektorraum* $\{\mathbf{0}\}$ *hat die leere Menge* $\emptyset$ *als Basis.*

Definition 1.17 *Ein* $\mathbb{R}$*-Vektorraum* V *heißt* endlich erzeugt*, wenn es endlich viele Vektoren* $\mathbf{v}_1, ..., \mathbf{v}_n \in V$ *gibt mit* $V = span(\mathbf{v}_1, ..., \mathbf{v}_m)$*.*

Satz 1.18 (Basis-Satz) *Jeder endlich erzeugte* $\mathbb{R}$*-Vektorraum hat eine endliche Basis.*

Dies ist der Spezialfall $W = \{\mathbf{0}\}$ des folgenden Satzes 1.19 so dass wir nur diesen Satz 1.19 zu beweisen brauchen.

Satz 1.19 (Basis-Ergänzungs-Satz) *Der* $\mathbb{R}$*-Vektorraum* V *sei endlich erzeugt. Weiter sei* $W \subset V$ *ein Untervektorraum mit einer Basis* $\mathbf{w}_1, ..., \mathbf{w}_r$*. Dann gibt es Vektoren* $\mathbf{v}_1, ..., \mathbf{v}_s \in V$ *so, dass das System* $\mathbf{w}_1, ..., \mathbf{w}_r, \mathbf{v}_1, ..., \mathbf{v}_s$ *eine Basis von* V *ist.*

Beweis. Der Vektorraum V werde durch n Vektoren aufgespannt. Wenn $W = V$ ist, dann ist nichts zu beweisen ($s = 0$). Wenn $W \neq V$ ist, dann existiert ein $\mathbf{v} \in V$, das nicht $\in W$ ist. Wir behaupten, das System $\mathbf{w}_1, ..., \mathbf{w}_r, \mathbf{v}$ ist linear unabhängig und verwenden den Test aus Satz 1.16. Sei also

$$\sum_1^r c_\nu \mathbf{w}_\nu + c\mathbf{v} = \mathbf{0}$$

eine lineare Relation. Dann muss $c = 0$ gelten, denn sonst würde $\mathbf{v} = -\frac{1}{c}\sum c_\nu \mathbf{w}_\nu$ zu W gehören. Weil nun $c = 0$, so lautet die lineare Relation nur noch

$$\sum_1^r c_\nu \mathbf{w}_\nu = \mathbf{0}.$$

Weil die $\mathbf{w}_1, ..., \mathbf{w}_r$ eine Basis von W bilden, sind sie insbesondere linear unabhängig. Deswegen folgt jetzt auch $c_1 = ... = c_r = 0$ und $\mathbf{w}_1, ..., \mathbf{w}_r, \mathbf{v}$ sind linear unabhängig.

Wir setzen $\mathbf{v}_1 := \mathbf{v}$ und $V_1 := span(\mathbf{w}_1, ..., \mathbf{w}_r, \mathbf{v}_1)$. Dann bilden die Vektoren $\mathbf{w}_1 ..., \mathbf{w}_r, \mathbf{v}_1$ eine Basis von V_1. Wenn $V_1 = V$ ist, dann sind wir fertig. Andernfalls wiederholen wir diese Konstruktion immer wieder. So erhalten wir für alle $k \geq 1$ Untervektorräume $V_k \subset V$ mit einer Basis $\mathbf{w}_1, ..., \mathbf{w}_r, \mathbf{v}_1, ..., \mathbf{v}_k$. Spätestens wenn $r + k = n + 1$ ist, können die $n + 1$ Vektoren $\mathbf{w}_1, ..., \mathbf{w}_r, \mathbf{v}_1, ..., \mathbf{v}_k$ nicht mehr linear unabhängig sein (Satz 1.17). Es muss also vorher schon einmal ein $k = s$ gegeben haben mit $V_s = V$. $\square$

Satz 1.20 (Korollar zu Satz 1.19) *In der Situation von Satz 1.19 gibt es zu* W *einen Untervektorraum* $W' \subset V$ *mit*

$$V = W \oplus W'.$$

Beweis. Wir setzen $W' = span(\mathbf{v}_1, ..., \mathbf{v}_s)$. $\square$

Satz 1.21 (Basis-Auswahl-Satz) *Es sei* $V = span(\mathbf{v}_1, ..., \mathbf{v}_k)$ *ein endlich-erzeugter* $\mathbb{R}$*-Vektorraum. Dann gibt es unter den Vektoren* $\mathbf{v}_1, ..., \mathbf{v}_k$ *eine Basis* $\mathbf{v}_{i_1}, ..., \mathbf{v}_{i_r}$ *für* V*.*

Beweis. Wenn $\mathbf{v}_1, ..., \mathbf{v}_k$ linear unabhängig sind, dann bilden sie eine Basis von V und wir sind fertig. Andernfalls gibt es unter ihnen einen Vektor $\mathbf{v}_j$ der eine Linearkombination $\sum_{i \neq j} c_i \mathbf{v}_i$ der anderen Vektoren ist. Dann wird U auch schon von den $k-1$ Vektoren $\mathbf{v}_1, ..., \mathbf{v}_{j-1}, \mathbf{v}_{j+1}, ..., \mathbf{v}_k$ aufgespannt. Spätestens nachdem wir diesen Schritt $k-1$-mal wiederholt haben, gelangen wir zu einem linear unabhängigen Teilsystem der $\mathbf{v}_1, ..., \mathbf{v}_k$, welches V aufspannt. □

Satz 1.22 (Invarianz der Basis-Länge) *Die Länge einer Basis für einen* $\mathbb{R}$*-Vektorraum* V *hängt nur von* V *ab und nicht von der gewählten Basis.*

Beweis. Seien $\mathbf{v}_1, ..., \mathbf{v}_r$ und $\mathbf{w}_1, ..., \mathbf{w}_s$ zwei Basen für V. Wir haben $s \leq r$ zu zeigen. Wenn aber $s > r$ sein sollte, dann wären die Vektoren $\mathbf{w}_1, ..., \mathbf{w}_s$ nach Satz 1.17 linear abhängig. Weil sie auch eine Basis von V bilden, geht das nicht. □

Die Sätze 1.18 und Satz 1.22 ermöglichen folgende Definition:

Definition 1.18 *Die* Dimension *eines endlich-erzeugten* $\mathbb{R}$*-Vektorraums* V *- in Zeichen* $dimV$ *- ist die Länge einer Basis für* V*. Für* $V = \{\mathbf{0}\}$ *setzt man* $dimV = 0$*.*

Beispiel 1.35 *Da* $\mathbf{e}_1, ..., \mathbf{e}_n \in \mathbb{R}^n$ *eine Basis bilden ist*

$$dim(\mathbb{R}^n) = n.$$

Eine Gerade hat die Dimension 1, eine Ebene Dimension 2.

Beispiel 1.36 *Im Raum* $M(m \times n)$ *der reellen* $m \times n$*-Matrizen fixieren wir die* $m \cdot n$ *Matrizen* $E_{i,j}$*, welche in der* i*-ten Zeile an der* j*-ten Stelle einen Eintrag* $= 1$ *haben, und wo alle anderen Einträge* $= 0$ *sind. Diese Matrizen bilden eine Basis von* $M(m \times n)$*, folglich ist* $dimM(m \times n) = m \cdot n$*.*

Beispiel 1.37 *Für den Raum* $T(\Delta)$ *der Treppenfunktionen aus Beispiel 1.8 bilden die Funktionen* $t_1, ..., t_n$ *aus Beispiel 1.20 eine Basis. Also ist* $dim\, T(\Delta) = n$*.*

Für den Raum $\mathbb{R}_n[x]$ *der Polynome vom Grad* $\leq n$ *bilden die* $n+1$ *Monome* x^i, $0 \leq i \leq n$, *eine Basis. Also ist* $dim\, \mathbb{R}_n[x] = n + 1$*.*

Der Vektorraum $\mathbb{R}[x]$ *aller Polynome ist nicht endlich erzeugbar. Endlich viele Polynome können nämlich als Linearkombinationen nur Polynome mit einem festen, endlichen Maximalgrad ergeben. Also hat der Vektorraum* $\mathbb{R}[x]$ *keine endliche Basis. Andererseits bildet die Menge* x^i, $i \in \mathbb{N}$, *aller Monome eine Basis des Raums* $\mathbb{R}[x]$*.*

Die Frage nach der Existenz einer Basis in allgemeinen, nicht endlich erzeugten Vektorräumen, berühren wir hier nicht. Wenn man das *Auswahlaxiom*, bzw. äquivalent dazu das *Zornsche Lemma* akzeptiert - wogegen nichts spricht (P.K.), wofür allerdings auch nichts (W.B.) - kann man für jeden Vektorraum die Existenz einer Basis beweisen. Dieser Beweis ist allerdings nicht konstruktiv. Den Beweis kann man z.B. finden in T.J. Jech: The Axiom of Choice, North Holland 1973, auf p. 12.

Die in Beispiel 1.37 angegebene Basis für $\mathbb{R}[x]$ ist abzählbar. Das liegt daran, dass Polynome noch sehr spezielle Funktionen sind. Für den $\mathbb{R}$-Vektorraum aller stetigen Funktionen $f : \mathbb{R} \to \mathbb{R}$ kann es keine abzählbare Basis geben (ohne Beweis). Hier beschränken wir uns auf die Redeweise

Definition 1.19 *Der* $\mathbb{R}$*-Vektorraum sei nicht endlich-erzeugt. Dann heißt* V unendlich-dimensional*, in Zeichen* $dim\, V = \infty$*.*

Satz 1.23 *Wir betrachten* $\mathbb{R}$*-Vektorräume* $U \subset V$*. Dann gelten*
1) $dim\,U \le dim\,V$*,*
2) $dim\,U = dim\,V = n$ *endlich* $\quad \Rightarrow \quad U = V$*.*

Beweis. 1) Für $dim\,V = \infty$ ist nichts zu zeigen. Wenn V endlich-dimensional ist, folgt die Aussage aus dem Basis-Ergänzungs-Satz 1.19.

2) Wegen 1) ist U endlich-dimensional und besitzt eine endliche Basis. Nach dem Basis-Ergänzungs-Satz können wir diese Basis durch $dim\,V - dim\,U = 0$ Vektoren zu einer Basis von V ergänzen. Alle Vektoren aus V sind Linearkombinationen dieser Basisvektoren. Es folgt $V \subset U$. □

Für unendlich-dimensionale Vektorräume ist Aussage 2) von Satz 1.23 i.A. falsch. Sei etwa V der Vektorraum aller stetigen Funktionen $f : \mathbb{R} \to \mathbb{R}$ und U sein Unterraum $\mathbb{R}[x]$. Beide Vektorräume sind unendlich-dimensional, stimmen aber nicht überein.

Satz 1.24 (Dimensionsformel) *Für je zwei lineare Unterräume* U_1, U_2 *eines* $\mathbb{R}$*-Vektorraums* V *gilt*

$$dim(U_1 \cap U_2) + dim(U_1 + U_2) = dim(U_1) + dim(U_2).$$

Beweis. Sei $\mathbf{u}_1, ..., \mathbf{u}_d$ eine Basis von $U_1 \cap U_2$. Wir ergänzen sie zu einer Basis $\mathbf{u}_1, ..., \mathbf{u}_d, \mathbf{v}_1, ..., \mathbf{v}_r$ von U_1 und einer Basis $\mathbf{u}_1, ..., \mathbf{u}_d, \mathbf{w}_1, ..., \mathbf{w}_s$ von U_2. Wir testen das System $\mathbf{u}_1, ..., \mathbf{u}_d, \mathbf{v}_1, ..., \mathbf{v}_r, \mathbf{w}_1, ..., \mathbf{w}_s$ auf lineare Unabhängigkeit: sei etwa die lineare Relation

$$\underbrace{a_1\mathbf{u}_1 + ... + a_d\mathbf{u}_d + b_1\mathbf{v}_1 + ... + b_r\mathbf{v}_r}_{\in U_1} + \underbrace{c_1\mathbf{w}_1 + ... + c_s\mathbf{w}_s}_{\in U_2} = \mathbf{0}$$

zwischen diesen Vektoren vorgelegt. Dann ist

$$c_1\mathbf{w}_1 + ... + c_s\mathbf{w}_s = -(a_1\mathbf{u}_1 + ... + a_d\mathbf{u}_d + b_1\mathbf{v}_1 + ... + b_r\mathbf{v}_r) \in (U_1 \cap U_2),$$

also

$$c_1\mathbf{w}_1 + ... + c_s\mathbf{w}_s = \alpha_1\mathbf{u}_1 + ... + \alpha_d\mathbf{u}_d \quad \text{mit} \quad \alpha_1, ..., \alpha_d \in \mathbb{R}.$$

Da aber $\mathbf{u}_1, ..., \mathbf{u}_d, \mathbf{w}_1, ..., \mathbf{w}_s$ als Basis von U_2 linear unabhängig waren, folgt hieraus $c_1 = ... = c_s = 0$. Ganz analog folgt $b_1 = ... = b_r = 0$, sodass die lineare Relation schließlich $a_1\mathbf{u}_1 + ... + a_d\mathbf{u}_d = 0$ lautet. Hieraus folgt dann endlich noch $a_1 = ... = a_d = 0$.

Da $\mathbf{u}_1, ..., \mathbf{u}_d, \mathbf{v}_1, ..., \mathbf{v}_r, \mathbf{w}_1, ..., \mathbf{w}_s$ den Unterraum $U_1 + U_2$ aufspannen, haben wir bewiesen, dass sie eine Basis von $U_1 + U_2$ bilden. Somit ist

$$\begin{array}{ll} dim(U_1) = d + r & dim(U_2) = d + s \\ dim(U_1 \cap U_2) = d & dim(U_1 + U_2) = d + r + s \\ dim(U_1) + dim(U_2) = 2d + r + s & dim(U_1 \cap U_2) + dim(U_1 + U_2) = 2d + r + s. \end{array}$$

Damit ist Formel (ii) bewiesen □

Im Spezialfall $U = U_1 \cap U_2 = \{\mathbf{0}\}$ wird die Dimensionsformel

$$dim\,U = dim(U_1 + U_2) = dim\,U_1 + dim\,U_2.$$

In diesem Fall ist aber $U = U_1 \oplus U_2$ eine direkte Summe. Jeder Untervektorraum $U_2 \subset U$ mit $U = U_1 \oplus U_2$ heißt ein *Komplement* von U_1 in U. Das Beispiel $U = \mathbb{R}^2$ und $U_1 = \mathbb{R} \cdot (1, 0)^t$ zeigt, dass es i.A. unendlich viele solche Komplemente U_2 gibt. Aber jedes dieser Komplemente hat die gleiche Dimension $= dim(U) - dim(U_1)$.

Aufgabe 1.33 *Sind die vier Vektoren*

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} 3 \\ 1 \\ 3 \\ 1 \end{pmatrix}, \quad \mathbf{w}_3 = \begin{pmatrix} 0 \\ 2 \\ 2 \\ 1 \end{pmatrix}, \quad \mathbf{w}_4 = \begin{pmatrix} 8 \\ 4 \\ 12 \\ 5 \end{pmatrix}$$

im $\mathbb{R}^4$ *linear abhängig oder linear unabhängig?*

Aufgabe 1.34 *Im* $\mathbb{R}^3$ *seien die Vektoren*

$$\mathbf{w}_1 = \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix}, \quad \mathbf{w}_3 = \begin{pmatrix} -4 \\ 0 \\ 5 \end{pmatrix}, \quad \mathbf{w}_4 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

gegeben. Welche der folgenden Mengen sind linear unabhängig?

$$a)\,\{\mathbf{w}_1, \mathbf{w}_2\}, \quad b)\,\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}, \quad c)\,\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \mathbf{w}_4\}.$$

Aufgabe 1.35 *Welche der folgenden vier Tripel von Vektoren im* $\mathbb{R}^3$ *sind linear unabhängig, welche sind linear abhängig?*
a) $(1,1,1), (1,1,0), (1,0,0)$;
b) $(1,1,1), (1,2,1), (0,2,0)$;
c) $(1,1,0), (0,1,1), 1,0,-1)$;
d) $(1,2,3), (4,5,6), (7,8,9)$.

Aufgabe 1.36 *Aus den Zeilenvektoren der Matrix*

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

lassen sich 15 verschiedene, nichtleere Mengen von Vektoren bilden, und ebenso aus den Spaltenvektoren. Welche dieser 30 Mengen sind linear abhängig?

Aufgabe 1.37 *Es sei* $U \subset \mathbb{R}^n$ *ein* k*-dimensionaler Untervektorraum. Zeigen Sie, dass für jede Teilmenge* $M \subset U$ *die folgenden Eigenschaften äquivalent sind:*
1) M *ist eine Basis von* U,
2) M *ist linear unabhängig und besteht aus* k *Vektoren,*
3) M *spannt* U *auf und besteht aus* k *Vektoren.*

Aufgabe 1.38 *Es seien*

$$\begin{aligned} U &:= \{\mathbf{x} \in \mathbb{R}^4 : x_1 + 2x_2 = x_3 + 2x_4\}, \\ V &:= \{\mathbf{x} \in \mathbb{R}^4 : x_1 = x_2 + x_3 + x_4\}. \end{aligned}$$

Bestimmen Sie Basen von $U, V, U \cap V$ *und* $U + V$.

Aufgabe 1.39 *Gegeben sei die reelle* 4×4*-Matrix*

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

Bestimmen Sie eine Basis für den Lösungsraum des homogenen linearen Gleichungssystems mit Koeffizientenmatrix A.

Aufgabe 1.40 *Es sei* U *der von* $(1,2,3)$ *und* $(4,5,6)$, *sowie* V *der von* $(2,1,2)$ *und* $(1,2,0)$ *aufgespannte Unterraum des* $\mathbb{R}^3$. *Man bestimme eine Basis von* $U \cap V$.

Aufgabe 1.41 *Es seien* $U, V, W \subset \mathbb{R}^n$ *lineare Unterräume. Beweisen oder widerlegen Sie die Formeln*

$$\begin{aligned} dim(U \cap (V+W)) &= dim(U \cap V) + dim(U \cap W) - dim(U \cap V \cap W), \\ dim(U+V+W) &= dim(U+V) + dim(U+W) - dim(U + (V \cap W)). \end{aligned}$$

Aufgabe 1.42 *Im* $\mathbb{R}^4$ *seien die Unterräume*

$$U_1 = \mathbb{R}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \mathbb{R}\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} \quad \textit{und} \quad U_2 = \mathbb{R}\begin{pmatrix} 6 \\ 5 \\ 1 \\ 2 \end{pmatrix} + \mathbb{R}\begin{pmatrix} 3 \\ 4 \\ 5 \\ 0 \end{pmatrix} + \mathbb{R}\begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \end{pmatrix}$$

gegeben. Man berechne eine Basis von $U_1 \cap U_2$.

Aufgabe 1.43 *Im* $\mathbb{R}^4$ *seien die folgenden Punktmengen gegeben:*

$$\begin{aligned} E_1 &= \{\mathbf{x} : x_1 + x_2 + x_3 + x_4 = x_1 + 2x_2 + 3x_3 + 4x_4 = 0\}, \\ E_2 &= \{\mathbf{x} : 5x_1 + 8x_2 + 11x_3 + 14x_4 = 2x_1 + 5x_2 + 8x_3 + 11x_4 = 0\}. \end{aligned}$$

Man zeige:
a) E_1 *und* E_2 *sind zweidimensionale Unterräume des* $\mathbb{R}^4$*;*
b) $E_1 = E_2$.

Aufgabe 1.44 *Im* $\mathbb{R}^4$ *seien die Vektoren*

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{v}_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 3 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$$

gegeben. Man bestimme ein Gleichungssystem, dessen Lösungsgesamtheit $\mathbb{R}\mathbf{v}_1 + \mathbb{R}\mathbf{v}_2 + \mathbb{R}\mathbf{v}_3$ *ist.*

Aufgabe 1.45 *Gegeben sei das lineare Gleichungssystem*

$$\begin{array}{ccccccc} 6x & + & 15y & + & 4z & = & a \\ 3x & + & 5y & + & 2z & = & b \\ 9x & & & + & 6z & = & c \end{array}$$

für $(x,y,z) \in \mathbb{R}^3$.
a) Man gebe die Menge E *aller rechten Seiten* $(a,b,c) \in \mathbb{R}^3$ *an, für die es lösbar ist. Welche Struktur hat diese Menge* E*?*
b) Welche Dimension hat der affine Lösungsraum für festes $(a,b,c) \in E$*?*
c) Man gebe denselben in Abhängigkeit von $(a,b,c) \in E$ *an.*

Aufgabe 1.46 *Geben Sie mit Begründung die Dimension der linearen Hülle (= span) der folgenden Teilmengen des* $\mathbb{R}^3$ *an:*
a) $S_1 := \{(1+n, 1+2n, 1+3n);\ n = 0, 1, ...\}$,
b) $S_2 := \{(n, n^2, n^3); n = 1, 2, ...\}$.

Aufgabe 1.47 *Seien* $n, k \in \mathbb{N}$, *seien* $\mathbf{v}_1, , \mathbf{v}_2, ..., \mathbf{v}_n \in \mathbb{R}^k$ *Vektoren, und sei* $\mathbf{w}_i := \sum_{j=1}^{i} \mathbf{v}_j$ *für* $i = 1, ..., n$. *Man zeige, dass die Folge* $(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n)$ *genau dann linear unabhängig ist, wenn die Folge* $(\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_n)$ *linear unabhängig ist.*

Aufgabe 1.48 U_1, *bzw.* U_2 *seien die durch die Vektoren*

$$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 4 \\ 3 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \qquad \textit{bzw.} \qquad \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

aufgespannten Unterräume des $\mathbb{R}^4$. *Man bestimme die Dimension und eine Basis von* $U_1 \cap U_2$.

Aufgabe 1.49 *Im reellen Vektorraum* $\mathbb{R}^5$ *seien folgende Vektoren gegeben:*

$$\mathbf{u}_1 = (-1, 4, -3, 0, 3);\ \mathbf{u}_2 = (2, -6, 5, 0, -2);\ \mathbf{u}_3 = (-2, 2, -3, 0, 6).$$

Sei U *der von* $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ *aufgespannte Unterraum im* $\mathbb{R}^5$. *Bestimmen Sie ein reelles lineares Gleichungssystem, dessen Lösungsraum genau* U *ist.*

1.4.2 Lineare Gleichungssysteme und zugehörige Unterräume

Wir wenden unseren Dimensionsbegriff jetzt noch auf lineare Gleichungssysteme an. Eine Matrix $A \subset M(m \times n)$ definiert zwei Untervektorräume: den von ihren Spalten $\mathbf{a}^\nu$ aufgespannten Unterraum

$$\text{Spaltenraum } S(A) = \mathit{span}(\mathbf{a}^1, ..., \mathbf{a}^n) \subset \mathbb{R}^m$$

und den von ihren Zeilen $\mathbf{a}_\mu$ aufgespannten Unterraum

$$\text{Zeilenraum } Z(A) = \mathit{span}(\mathbf{a}_1, ..., \mathbf{a}_m) \subset \mathbb{R}^n.$$

Definition 1.20 *Es sei* A *eine* $m \times n$*-Matrix.*

Der Spaltenrang von A *ist die Dimension ihres Spaltenraums im* $\mathbb{R}^m$. *Damit ist der Spaltenrang* $\leq m$. *Man sagt: die Matrix* A *hat vollen (bzw. maximalen) Spaltenrang, wenn dieser* $= n$, *die Anzahl aller Spalten ist.*

Der Zeilenrang von A *ist die Dimension ihres Zeilenraums im* $\mathbb{R}^n$. *Damit ist der Zeilenrang* $\leq n$. *Man sagt: die Matrix* A *hat vollen (bzw. maximalen) Zeilenrang, wenn dieser* $= m$, *die Anzahl aller Zeilen ist.*

Satz 1.25 *Es sei* A *eine* $m \times n$*-Matrix.*
*1) Der Zeilen*__raum__ *von* A *ändert sich nicht bei elementaren Zeilenumformungen und damit ändert sich auch der Zeilen*__rang__ *nicht.*
*2) Auch der Spalten*__rang__ *von* A *ändert sich nicht bei elementaren Zeilenumformungen.*
3) Es gilt

$$\boxed{\textit{Zeilenrang} = \textit{Spaltenrang.}}$$

Beweis. 1) Bei Umformungen vom Typ I wird nur die Reihenfolge der Zeilenvektoren geändert, der von ihnen aufgespannte Unterraum im $\mathbb{R}^n$ bleibt gleich. Wird die Matrix A durch eine Umformung vom Typ II oder III in eine Matrix A' umgeformt, so sind alle Zeilen der Matrix A' Linearkombinationen von Zeilen der Matrix A. Der Zeilenraum Z' der umgeformten Matrix A' ist also im Zeilenraum Z der Matrix A enthalten. Weil man die Umformung durch eine Umformung vom gleichen Typ wieder rückgängig machen kann, gilt auch $Z \subset Z'$.

2) Der Spaltenrang von A sei r. Nach Satz 1.21 (Basis-Auswahl-Satz) gibt es r linear unabhängige Spalten $\mathbf{b}_1 := \mathbf{a}_{\nu_1}, ..., \mathbf{b}_r = \mathbf{a}_{\nu_r}$ der Matrix A. Weil die Spalten der damit gebildeten $m \times r$-Matrix $B := (\mathbf{b}_1, ..., \mathbf{b}_r)$ linear unabhängig sind, hat das LGS mit dieser Matrix $B \cdot \mathbf{x} = \mathbf{0}$ nur die Null-Lösung.

Die Matrix A werde durch eine elementare Zeilenumformung in die Matrix A' übergeführt. Dabei wird auch die Teilmatrix B von A in eine Matrix B' übergeführt. Bei dieser Zeilenumformung der Matrix B ändert sich der Lösungsraum des Gleichungssytems $B \cdot \mathbf{x} = \mathbf{0}$ nicht. Also hat auch das LGS $B' \cdot \mathbf{x} = \mathbf{0}$ nur die Null-Lösung. Deswegen sind die r Spalten der Matrix B' linear unabhängig. Diese sind auch Spalten der Matrix A'. Also gilt für der Zeilenrang r' von A' dass $r' \geq r$. Weil man die durchgeführte Zeilenumformung durch eine Umformung vom gleichen Typ wieder rückgängig machen kann, gilt auch $r \geq r'$.

3) Wir überführen die Matrix A durch elementare Zeilenumformungen in eine Matrix Z in Zeilenstufenform. Nach 1) und 2) ändern sich dabei weder Zeilenrang, noch Spaltenrang. Wir brauchen die Behauptung also nur für die Matrix Z zu beweisen. Die Anzahl ihrer Stufen = Anzahl ihrer Zeilen $\neq \mathbf{0}$ sei s. Nach Beispiel 1.28 hat Z den Zeilenrang s. Wendet man die Argumentation aus Beispiel 1.28 auf die Spalten von Z, anstatt auf Zeilen an, so sieht man: die s Spalten von Z, die eine Stufe enthalten, sind linear unabhängig. Also hat die Matrix Z auch den Spaltenrang s. □

Definition 1.21 *Der* Rang *einer Matrix ist der gemeinsame Wert ihres Zeilen- und Spaltenrangs.*

Satz 1.25 liefert ein allgemeines Bestimmungsverfahren für den Rang einer Matrix. Mit elementaren Zeilenumformungen bringe man die Matrix auf Zeilenstufenform. Der Rang der Matrix ist dann die Anzahl der Stufen bzw. die Anzahl der Zeilen $\neq \mathbf{0}$ in dieser Zeilenstufenform.

Analog kann man auch die Dimension eines Untervektorraums $U = span(\mathbf{v}_1, ..., \mathbf{v}_k) \subset \mathbb{R}^n$ bestimmen, der von k Vektoren $\mathbf{v}_1, ..., \mathbf{v}_k \in \mathbb{R}^n$ aufgespannt wird. Man schreibe diese Vektoren als Zeilen $\mathbf{v}_1^t, ..., \mathbf{v}_k^t$ in eine $k \times n$-Matrix A. Der Zeilenraum dieser Matrix ist U. Bringen wir die Matrix A durch elementare Zeilenumformungen auf Zeilenstufenform A', so ändert sich der Zeilenraum U nicht. Eine Basis von U bilden also die Zeilenvektoren $\neq \mathbf{0}$ dieser Matrix A'.

Seien weiter $j_1, ..., j_r$ die Spalten der Matrix A', in denen ihre Stufen stehen und

$$\{i_1, ..., i_{n-r}\} = \{1, ..., n\} \setminus \{j_1, ..., j_r\}.$$

Die $n \times n$-Matrix A'' entstehe aus A', indem wir die Einheitsvektoren $\mathbf{e}_{i_1}^t, ..., \mathbf{e}_{i_{n-r}}^t$ als Zeilen hinzunehmen. Durch elementare Zeilenumformungen kann man diese Matrix A'' so verändern, dass in jeder ihrer Zeilen oder Spalten genau eine 1 und sonst lauter Nuller stehen. Die Matrix A'' hat also den Rang n. Somit hat die Matrix mit den Zeilen $\mathbf{e}_{i_1}^t, ..., \mathbf{e}_{i_{n-r}}^t$ den Rang $n - r$ und ihre Zeilenvektoren spannen einen Untervektorraum

$$W = span(\mathbf{e}_{i_1}, ...\mathbf{e}_{i_{n-r}}) \subset \mathbb{R}^n$$

auf mit

$$U \oplus W = \mathbb{R}^n.$$

Hat eine $m \times n$-Matrix A vollen Zeilenrang $r = m$, so gibt es in ihrer Zeilenstufenform keine Nullzeilen. Das LGS $A \cdot \mathbf{x} = \mathbf{b}$ ist also für jede Wahl $\mathbf{b}$ der rechten Seite lösbar.

Hat A vollen Spaltenrang $r = n$, dann sind die n Spalten von A linear unabhängig und das LGS $A \cdot \mathbf{x} = \mathbf{0}$ hat nur die Null-Lösung. Ist ein LGS $A \cdot \mathbf{x} = \mathbf{b}$ lösbar, so ist die Lösung $\mathbf{x}$ nach Satz 1.4 eindeutig bestimmt.

Satz 1.26 *Es sei A eine $m \times n$-Matrix. Wir betrachten das homogene LGS*

$$A \cdot \mathbf{x} = \mathbf{0} \quad \textit{mit dem Lösungsraum} \quad U \subset \mathbb{R}^n.$$

Für die Zahlen

$$d := \textit{Dimension des Lösungsraums } U \qquad\qquad r := \textit{Rang der Koeffizientenmatrix } A$$

gilt dann die Beziehung

$$\boxed{\quad d + r = n \quad}$$

Beweis. Bei elementaren Zeilenumformungen der Koeffizientenmatrix ändern sich weder d noch r (Satz 1.25). Wir können daher o.B.d.A. annehmen, die Koeffizientenmatrix habe Zeilenstufenform. Die Zahl der Stufen ist dann r. Es gibt also $n - r$ Spalten ohne Stufe in der Koeffizientenmatrix. Bei elementaren Spaltenumformungen ändert sich zwar der Lösungsraum, aber nicht seine Dimension. Wir können deswegen sogar annehmen, die Matrix A habe Staffelform. Stufen stehen also in den ersten r Spalten der Matrix.

Jeder Lösungsvektor $\mathbf{x} \in U$ hat deswegen eine Form

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}' \\ \mathbf{x}'' \end{pmatrix} \quad \mathbf{x}' \in \mathbb{R}^r,\ \mathbf{x}'' \in \mathbb{R}^{n-r},$$

wo $\mathbf{x}''$ beliebig gewählt werden kann. Der Vektor $\mathbf{x}'$ ist dann durch $\mathbf{x}''$ festgelegt. Wählen wir hier $\mathbf{x}'' = \mathbf{e}_{r+1}, .., \mathbf{e}_n$, so erhalten wir $n - r$ linear unabhängige Lösungsvektoren

$$\mathbf{x}_i = \begin{pmatrix} \mathbf{x}'_i \\ \mathbf{e}_i \end{pmatrix}, \quad i = r+1, ..., n.$$

Sei nun $\mathbf{x} \in U$ eine beliebige Lösung. Der zugehörige Vektor $\mathbf{x}''$ hat die Form

$$\mathbf{x}'' = \sum_{i=r+1}^{n} x_i \mathbf{e}_i, \quad x_i \in \mathbb{R}.$$

Auch der Vektor

$$\sum_{i=r+1}^{n} x_i \mathbf{x}_i = \sum_{i=r+1}^{n} x_i \begin{pmatrix} \mathbf{x}'_i \\ \mathbf{e}_i \end{pmatrix} = \begin{pmatrix} \sum_{i=r+1}^{n} x_i \mathbf{x}'_i \\ \mathbf{x}'' \end{pmatrix}$$

ist ein Lösungsvektor zum gleichen Vektor $\mathbf{x}''$. Weil $\mathbf{x}$ durch $\mathbf{x}''$ eindeutig bestimmt ist, muss

$$\mathbf{x} = \sum_{i=r+1}^{n} x_i \mathbf{x}_i$$

gelten.

Die $n - r$ Lösungsvektoren $\mathbf{x}_i$ bilden also eine Basis des Lösungsraums U. □

Satz 1.27 (Korollar) *Jeder lineare Unterraum $U \subset \mathbb{R}^n$ ist der Lösungsraum eines homogenen LGS.*

Beweis. Sei $dim(U) = k$ und $\mathbf{u}_1, ..., \mathbf{u}_k \in U$ eine Basis. Die $k \times n$-Matrix

$$B := \begin{pmatrix} \mathbf{u}_1^t \\ \vdots \\ \mathbf{u}_k^t \end{pmatrix}$$

hat also den Rang k. Diese Matrix ist die Koeffizientenmatrix eines homogenen LGS

$$B \cdot \mathbf{y} = \mathbf{0}$$

für die Unbekannten $y_1, ..., y_n$. Nach der Dimensionsformel von Satz 1.26 hat sein Lösungsraum $W \subset \mathbb{R}^n$ die Dimension $n - k$.

Wir wählen eine Basis $\mathbf{a}_1, ..., \mathbf{a}_{n-k} \in W$. Dann gilt also

$$\sum_{\nu=1}^{n} u_{\mu,\nu} a_{\lambda,\nu} = 0 \text{ für } \mu = 1, ..., k, \ \lambda = 1, ..., n - k.$$

Jetzt kehren wir die Rolle von Koeffizienten und Unbekannten in diesem Gleichungssystem um. Dann sind die Vektoren $\mathbf{u}_1, ..., \mathbf{u}_k$ Lösungsvektoren des homogenen linearen Gleichungssystems

$$\sum_{\nu=1}^{n} a_{\lambda,\nu} u_\nu = 0, \quad \lambda = 1, ..., n - k.$$

Die Koeffizientenmatrix dieses Systems hat die Zeilenvektoren $\mathbf{a}_1^t, ..., \mathbf{a}_{n-k}^t$ und damit den Zeilenrang $n - k$. Alle Vektoren $\mathbf{u}_1, ..., \mathbf{u}_k$ sind Lösungen dieses neuen Systems. Damit enthält sein Lösungsraum $V \subset \mathbb{R}^n$ auch $U = span(\mathbf{u}_1, ..., \mathbf{u}_k)$. Wieder verwenden wir die Dimensionsformel und folgern

$$dim(V) = n - (n - k) = k = dim(U).$$

Wegen $U \subset V$ folgt daraus $U = V$, d.h., U ist der Lösungsraum des Gleichungssystems mit den Zeilenvektoren $\mathbf{a}_1, ..., \mathbf{a}_{n-k}$. □

Satz 1.28 *Es sei A eine $n \times n$-Matrix. Wir betrachten das quadratische LGS*

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

(genauso viele Gleichungen wie Unbekannte!) Für dieses LGS sind die folgenden Aussagen äquivalent:

(i) Bei jeder Wahl der $b_1, ..., b_n$ auf der rechten Seite ist das Gleichungssystem lösbar. (Existenz)

(ii) Bei jeder Wahl der $b_1, ..., b_n$ auf der rechten Seite gibt es höchstens eine *Lösung des Systems. (Eindeutigkeit)*

(iii) Das zugehörige homogene System

$$A \cdot \mathbf{x} = \mathbf{0}$$

hat nur die Null-Lösung.

(iv) Der Rang der Koeffizientenmatrix ist n.

Beweis. Eigenschaft (i) ist damit äquivalent, dass die Spaltenvektoren der Koeffizientenmatrix A den ganzen Vektorraum $\mathbb{R}^n$ aufspannen. Dies ist damit äquivalent, dass die Koeffizientenmatrix den (Spalten-) Rang n besitzt (iv). Daraus folgt, dass die n Spaltenvektoren linear unabhängig sind. Denn wären sie linear abhängig, würden auch schon $n-1$ dieser Vektoren den $\mathbb{R}^n$ aufspannen. Aus dem Basis-Auswahlsatz würde der Widerspruch $dim(\mathbb{R}^n) < n$ folgen. Eigenschaft (iii) ist aber nichts anderes als der Test auf lineare Unabhängigkeit für die Spaltenvektoren und folgt somit aus (i). Ist umgekehrt (iii) erfüllt, so spannen die Spaltenvektoren von A einen n-dimensionalen Teilraum von $\mathbb{R}^n$ auf, der wegen Satz 1.23 mit $\mathbb{R}^n$ übereinstimmen muss. Deswegen folgt auch (i) aus (iii).

Eigenschaft (iii) ist der Spezialfall $\mathbf{b} = \mathbf{0}$ von (ii) und folgt somit aus (ii). Umgekehrt folgt (ii) aus (iii) mit dem Struktursatz 1.4. □

Aufgabe 1.50 *Berechnen Sie den Zeilenrang der Matrizen*

$$A = \begin{pmatrix} 1 & 3 & 6 & 10 \\ 3 & 6 & 10 & 15 \\ 6 & 10 & 15 & 21 \\ 10 & 15 & 21 & 28 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 3 & 6 & 10 \\ 3 & 6 & 10 & 1 \\ 6 & 10 & 1 & 3 \\ 10 & 1 & 3 & 6 \end{pmatrix}.$$

Aufgabe 1.51 *Es seien a, b, c reelle Zahlen. Bestimmen Sie den Zeilenrang der Matrix*

$$A = \begin{pmatrix} b & 0 & c & 0 \\ 0 & b & 0 & c \\ a & 0 & b & 0 \\ 0 & a & 0 & b \end{pmatrix}.$$

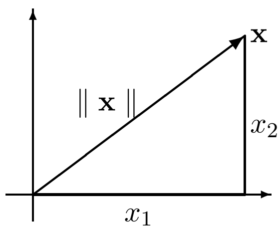
1.5 Das Skalarprodukt

1.5.1 Skalarprodukt im $\mathbb{R}^n$

In diesem Abschnitt 1.5.1 sollen zwei eng zusammenhängende Begriffe betrachtet werden, welche über die Vektorraumstruktur hinausgehen: Längen und Winkel.

Wir erinnern zunächst an den elementargeometrischen Begriff der Länge in $n = 1, 2$ und 3 Dimensionen:

n=1: Für $x \in \mathbb{R}$ ist $|x| := \sqrt{x^2}$ der *Betrag* der Zahl x.



n=2: Die Länge eines Vektors $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$ ist

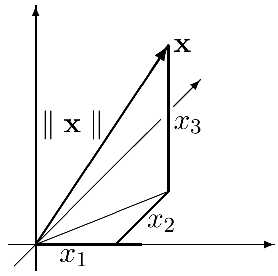
$$\| \mathbf{x} \| := \sqrt{x_1^2 + x_2^2}.$$

Dies ist der Inhalt des elementargeometrischen Satzes von Pythagoras.

n=3: Die Länge eines Vektors $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3$ ist

$$\| \mathbf{x} \| := \sqrt{x_1^2 + x_2^2 + x_3^2}.$$

Dies ergibt sich nach zweimaligem Anwenden des Pythagoras.



Es liegt nahe, wie dieser Längenbegriff für beliebige Dimension zu verallgemeinern ist:

Definition 1.22 *Sei* $\mathbf{x} = (x_1, ..., x_n)^t \in \mathbb{R}^n$. *Dann heißt*

$$\| \mathbf{x} \| := \sqrt{x_1^2 + x_2^2 + ... + x_n^2}$$

die Länge *oder die* Norm *von* $\mathbf{x}$.

Auch für den quadratischen Ausdruck unter der Wurzel führen wir eine eigene Notation ein:

Definition 1.23 *Seien* $\mathbf{x} = (x_1, ..., x_n)^t$ *und* $\mathbf{y} = (y_1, ..., y_n)^t$ *Vektoren im Zahlenraum* $\mathbb{R}^n$. *Dann heißt*

$$(\mathbf{x}.\mathbf{y}) := \sum_{\nu=1}^{n} x_\nu \cdot y_\nu$$

das Skalarprodukt *von* $\mathbf{x}$ *mit* $\mathbf{y}$.

Mit dieser Definition des Skalarprodukts ist also

$$\| \mathbf{x} \| = \sqrt{(\mathbf{x}.\mathbf{x})}.$$

Das Skalarprodukt $(\mathbf{x}.\mathbf{y})$ hat folgende Eigenschaften:

(i) *Bi-Linearität:* $((c_1\mathbf{x}_1 + c_2\mathbf{x}_2).\mathbf{y}) = c_1(\mathbf{x}_1.\mathbf{y}) + c_2(\mathbf{x}_2.\mathbf{y}), \quad \mathbf{x}_1, \mathbf{x}_2, \mathbf{y} \in \mathbb{R}^n, c_1, c_2 \in \mathbb{R},$
$(\mathbf{x}.(c_1\mathbf{y}_1 + c_2\mathbf{y}_2)) = c_1(\mathbf{x}.\mathbf{y}_1) + c_2(\mathbf{x}.\mathbf{y}_2), \quad \mathbf{x}, \mathbf{y}_1, \mathbf{y}_2 \in \mathbb{R}^n, c_1, c_2 \in \mathbb{R}.$

(ii) *Symmetrie:* $(\mathbf{x}.\mathbf{y}) = (\mathbf{y}.\mathbf{x}), \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$

(iii) *Definitheit:* Für alle $\mathbf{x} \in \mathbb{R}^n$ ist $(\mathbf{x}.\mathbf{x}) \geq 0$ und es gilt $(\mathbf{x}.\mathbf{x}) = 0$ genau dann, wenn $\mathbf{x} = \mathbf{0}$.

Eigenschaften der Norm sind:

(iv) *Definitheit:* Für alle $\mathbf{x} \in \mathbb{R}^n$ ist $\| \mathbf{x} \| \geq 0$ und es gilt $\| \mathbf{x} \| = 0$ genau dann, wenn $\mathbf{x} = \mathbf{0}$.

(v) *Homogenität:* $\| c \cdot \mathbf{x} \| = |c| \cdot \| \mathbf{x} \|$ für $c \in \mathbb{R}$ und $\mathbf{x} \in \mathbb{R}^n$.

Die Beweise für diese Eigenschaften sind offensichtlich. Einen Zusammenhang zwischen Skalarprodukt und Norm beschreibt

(vi) *Cauchy-Schwarz-Ungleichung:* $|(\mathbf{x}.\mathbf{y})| \leq \| \mathbf{x} \| \cdot \| \mathbf{y} \|$ (C.S.U.)

Beweis: Für alle $a, b \in \mathbb{R}$ ist

$$0 \leq \parallel a\mathbf{x} - b\mathbf{y} \parallel^2 = (a\mathbf{x} - b\mathbf{y} \,.\, a\mathbf{x} - b\mathbf{y}) = a^2 \parallel \mathbf{x} \parallel^2 -2ab(\mathbf{x}.\mathbf{y}) + b^2 \parallel \mathbf{y} \parallel^2,$$

oder äquivalent damit

$$2ab(\mathbf{x}.\mathbf{y}) \leq a^2 \parallel \mathbf{x} \parallel^2 + b^2 \parallel \mathbf{y} \parallel^2 .$$

Setzen wir z.B. $a = \parallel \mathbf{y} \parallel$ und $b = \parallel \mathbf{x} \parallel$, so erhalten wir

$$2 \parallel \mathbf{x} \parallel \cdot \parallel \mathbf{y} \parallel (\mathbf{x}.\mathbf{y}) \leq 2 \parallel \mathbf{x} \parallel^2 \cdot \parallel \mathbf{y} \parallel^2 .$$

Da die Behauptung für $\mathbf{x} = \mathbf{0}$ oder $\mathbf{y} = \mathbf{0}$ richtig ist, können wir o.B.d.A. $\mathbf{x} \neq \mathbf{0} \neq \mathbf{y}$ annehmen. Dann dürfen wir in der letzten Gleichung kürzen und erhalten

$$(\mathbf{x}.\mathbf{y}) \leq \parallel \mathbf{x} \parallel \cdot \parallel \mathbf{y} \parallel .$$

Für $-\mathbf{x}$ statt $\mathbf{x}$ gilt dieselbe Ungleichung, sodass also auch

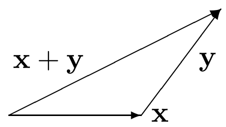
$$-(\mathbf{x}.\mathbf{y}) = (-\mathbf{x}.\mathbf{y}) \leq \parallel \mathbf{x} \parallel \cdot \parallel \mathbf{y} \parallel$$

gilt. Daraus folgt schließlich

$$|(\mathbf{x}.\mathbf{y})| = max\{(\mathbf{x}.\mathbf{y}), \, -(\mathbf{x}.\mathbf{y})\} \leq \parallel \mathbf{x} \parallel \cdot \parallel \mathbf{y} \parallel . \qquad \square$$

Mit der Cauchy-Schwarz-Ungleichung beweist man die

(vii) *Dreiecksungleichung:* $\parallel \mathbf{x} + \mathbf{y} \parallel \leq \parallel \mathbf{x} \parallel + \parallel \mathbf{y} \parallel$.



Beweis:

$$\begin{aligned} \parallel \mathbf{x} + \mathbf{y} \parallel^2 &= (\mathbf{x} + \mathbf{y}.\mathbf{x} + \mathbf{y}) \\ &= \parallel \mathbf{x} \parallel^2 + 2(\mathbf{x}.\mathbf{y}) + \parallel \mathbf{y} \parallel^2 \\ &\leq \parallel \mathbf{x} \parallel^2 + 2 \parallel \mathbf{x} \parallel \cdot \parallel \mathbf{y} \parallel + \parallel \mathbf{y} \parallel^2 \\ &= (\parallel \mathbf{x} \parallel + \parallel \mathbf{y} \parallel)^2. \end{aligned} \qquad \square$$

Nicht nur die Norm eines Vektors, auch das Skalarprodukt zweier Vektoren hat eine geometrische Bedeutung. Dazu betrachten wir zunächst zwei Einheitsvektoren (= Vektoren der Länge 1) im $\mathbb{R}^2$:

$$\begin{aligned} \mathbf{x} &= (cos(\alpha), sin(\alpha)) \\ \mathbf{y} &= (cos(\beta), sin(\beta)) \\ (\mathbf{x}.\mathbf{y}) &= cos(\alpha)cos(\beta) + sin(\alpha)sin(\beta) \\ &= cos(\alpha - \beta). \end{aligned}$$

Aus dem Additionstheorem für die *cos*-Funktion folgt also, dass das Skalarprodukt $(\mathbf{x}.\mathbf{y})$ zweier Einheitsvektoren der Cosinus des Winkels zwischen beiden Vektoren ist. Für zwei beliebige Vektoren $\mathbf{x} \neq \mathbf{0} \neq \mathbf{y}$ definieren wir zunächst die Einheitsvektoren

$$\hat{\mathbf{x}} := \frac{1}{\parallel \mathbf{x} \parallel}\mathbf{x}, \qquad \hat{\mathbf{y}} := \frac{1}{\parallel \mathbf{y} \parallel}\mathbf{y}$$

und erhalten dann für den Cosinus des Winkels zwischen $\mathbf{x}$ und $\mathbf{y}$

$$(\hat{\mathbf{x}}.\hat{\mathbf{y}}) = \frac{(\mathbf{x}.\mathbf{y})}{\parallel \mathbf{x} \parallel \cdot \parallel \mathbf{y} \parallel}.$$

Dies nehmen wir zum Anlass für die entsprechende Definition im $\mathbb{R}^n$:

Definition 1.24 *Seien* $\mathbf{x} \neq \mathbf{0} \neq \mathbf{y}$ *Vektoren im* $\mathbb{R}^n$. *Aus der Cauchy-Schwarz-Ungleichung folgt*

$$-1 \leq \frac{(\mathbf{x}.\mathbf{y})}{\parallel \mathbf{x} \parallel \cdot \parallel \mathbf{y} \parallel} \leq 1.$$

Da die Cosinus-Funktion das Intervall $[0, \pi]$ *bijektiv auf das Intervall* $[-1, 1]$ *abbildet, gibt es genau einen Winkel* $\alpha \in \mathbb{R}$ *mit*

$$cos(\alpha) = \frac{(\mathbf{x}.\mathbf{y})}{\parallel \mathbf{x} \parallel \cdot \parallel \mathbf{y} \parallel}, \qquad 0 \leq \alpha \leq \pi.$$

Wir nennen diesen Winkel α *den* Winkel zwischen den Vektoren $\mathbf{x}$ *und* $\mathbf{y}$. *(Dieser Winkel hat kein Vorzeichen! Er hängt nicht von der Reihenfolge der Vektoren* $\mathbf{x}$ *und* $\mathbf{y}$ *ab.)*

Hier haben wir ziemlich großzügig Gebrauch von den Eigenschaften der Cosinus-Funktion aus Analysis I gemacht. Die Beziehung zwischen Skalarprodukt und Cosinus des Zwischenwinkels ist für das Verständnis und die Anwendungen (z.B. in der analytischen Geometrie) von großer Bedeutung. Im weiteren Aufbau der Linearen Algebra selbst werden wir aber von dieser Tatsache keinen Gebrauch machen, sondern nur, um den Bezug zur Anschauung aufrecht zu erhalten. In diesem Sinn sollte uns deswegen die Anleihe bei der Vorlesung Analysis erlaubt sein.

Definition 1.25 *Zwei Vektoren* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *heißen* orthogonal *oder* senkrecht aufeinander, *in Zeichen* $\mathbf{x} \perp \mathbf{y}$, *wenn sie den Winkel* $\frac{\pi}{2}$ *einschließen, also wenn* $(\mathbf{x}.\mathbf{y}) = 0$ *ist. (Hier ist auch* $\mathbf{x} = \mathbf{0}$ *oder* $\mathbf{y} = \mathbf{0}$ *zugelassen.)*

Satz 1.29 (n-dimensionaler Pythagoras) *Es seien* $\mathbf{v}_1, ..., \mathbf{v}_r \in \mathbb{R}^n$ *Vektoren, die paarweise aufeinander senkrecht stehen:*

$$(\mathbf{v}_k.\mathbf{v}_l) = 0 \text{ für alle } k \neq l.$$

Dann gilt

$$\parallel \mathbf{v}_1 + \mathbf{v}_2 + ... + \mathbf{v}_r \parallel^2 = \parallel \mathbf{v}_1 \parallel^2 + \parallel \mathbf{v}_2 \parallel^2 + ... + \parallel \mathbf{v}_r \parallel^2 .$$

Beweis. Aus der Voraussetzung folgt, dass die linke Seite gleich

$$(\mathbf{v}_1 + ... + \mathbf{v}_r \,.\, \mathbf{v}_1 + ... + \mathbf{v}_r) = \sum_{k,l=1}^{r} (\mathbf{v}_k.\mathbf{v}_l) = \sum_{k=1}^{r} (\mathbf{v}_k.\mathbf{v}_k)$$

ist. □

Aufgabe 1.52 *a) Beweisen Sie, dass sich die drei Mittelsenkrechten eines Dreiecks in einem Punkt schneiden.*
b) Beweisen Sie, dass sich die drei Höhen eines Dreiecks in einem Punkt schneiden.

Aufgabe 1.53 *Es seien* $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$. *Zeigen Sie:*
a) $|\parallel \mathbf{x} \parallel - \parallel \mathbf{y} \parallel| \leq \parallel \mathbf{x} - \mathbf{y} \parallel$,
b) $\parallel \mathbf{x} \parallel = \parallel \mathbf{y} \parallel \quad \Leftrightarrow \quad (\mathbf{x} - \mathbf{y}) \perp (\mathbf{x} + \mathbf{y})$,
c) ist $\mathbf{x} \neq \mathbf{0}$ *und* $\mathbf{y} \neq \mathbf{0}$, *so gilt*

$$\parallel \frac{\mathbf{x}}{\parallel \mathbf{x} \parallel^2} - \frac{\mathbf{y}}{\parallel \mathbf{y} \parallel^2} \parallel = \frac{\parallel \mathbf{x} - \mathbf{y} \parallel}{\parallel \mathbf{x} \parallel \cdot \parallel \mathbf{y} \parallel},$$

d) $\parallel \mathbf{x} - \mathbf{y} \parallel \cdot \parallel \mathbf{z} \parallel \leq \parallel \mathbf{y} - \mathbf{z} \parallel \cdot \parallel \mathbf{x} \parallel + \parallel \mathbf{z} - \mathbf{x} \parallel \cdot \parallel \mathbf{y} \parallel$.

Aufgabe 1.54 *Zeigen Sie: In der Cauchy-Schwarz-Ungleichung* $|(\mathbf{x}.\mathbf{y})| \leq \| \mathbf{x} \| \cdot \| \mathbf{y} \|$ *gilt das Gleichheitszeichen genau dann, wenn die Vektoren* $\mathbf{x}$ *und* $\mathbf{y} \in \mathbb{R}^n$ *linear abhängig sind.*

Aufgabe 1.55 *Man bestimme im* $\mathbb{R}^4$ *den Winkel zwischen den Vektoren*

$$\mathbf{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

Aufgabe 1.56 *Zeige, dass zwei Vektoren* $\mathbf{x}$ *und* $\mathbf{y} \in \mathbb{R}^n$ *genau dann zueinander orthogonal sind, wenn für alle* $\lambda \in \mathbb{R}$ *gilt*

$$\| \mathbf{x} + \lambda \mathbf{y} \| \geq \| \mathbf{x} \| .$$

Aufgabe 1.57 *Im* $\mathbb{R}^3$ *seien* E_1 *und* E_2 *zwei 2-dimensionale Untervektorräume, die sich in einer Geraden* A *schneiden. Der Schnittwinkel* $\varphi = \angle(E_1, E_2)$ *von* E_1 *mit* E_2 *sei wie folgt definiert: Man wähle Vektoren* $\mathbf{v}_i \in E_i \, (i = 1, 2)$ *der Länge 1, die senkrecht auf* A *stehen. Man darf annehmen, dass* $(\mathbf{v}_1.\mathbf{v}_2) \geq 0$ *(andernfalls ersetze man* $\mathbf{v}_2$ *durch* $-\mathbf{v}_2$*). Dann ist*

$$cos(\varphi) = (\mathbf{v}_1.\mathbf{v}_2).$$

Sei

$$\begin{aligned} E_1 &= \{(x, y, z) \in \mathbb{R}^3 : x - y + 2z = 0\}, \\ E_2 &= \{(x, y, z) \in \mathbb{R}^3 : x + 2y - 3z = 0\}. \end{aligned}$$

Man berechne den Cosinus des Schnittwinkels von E_1 *und* E_2*.*

1.5.2 Skalarprodukt und Norm in $\mathbb{R}$-Vektorräumen

Wir verallgemeinern hier den Begriff des Skalarprodukts auf beliebige $\mathbb{R}$-Vektorräume. Das SKP aus Abschnitt 1.5.1 auf dem $\mathbb{R}^n$ nennen wir künftig *euklidisches* SKP. Die Norm aus Abschnitt 1.5.1 nennen wir entsprechend künftig *euklidische* Norm.

Definition 1.26 *Es sei* V *ein* $\mathbb{R}$*-Vektorraum. Ein Skalarprodukt (SKP) auf* V *ist eine Abbildung*

$$(\ . \ ) : V \times V \to \mathbb{R}, \qquad (\mathbf{x}, \mathbf{y}) \mapsto (\mathbf{x}.\mathbf{y}),$$

welche bilinear, symmetrisch und definit ist, also die Eigenschaften (i),(ii), (iii) aus Abschnitt 1.5.1 besitzt.

Beispiel 1.38 *Wir bezeichnen mit* $C[a, b]$ *den* $\mathbb{R}$*-Vektorraum der auf einem Intervall* $[a, b]$ *stetigen Funktionen. Mit* $f, g \in C[a, b]$ *ist auch das Produkt* $f \cdot g$ *auf* $[a, b]$ *stetig und insbesondere integrierbar. Deswegen ist*

$$(f.g) := \int_a^b f(x) \cdot g(x) \, dx$$

wohldefiniert. Dies ist ein SKP auf $C[a, b]$*. Für die Details müssen wir hier auf die Vorlesung Analysis I verweisen.*

Beispiel 1.39 *Das SKP aus Beispiel 1.38 ist auch auf dem Vektorraum $T(\Delta)$ der Treppenfunktionen (Beispiel 1.8) definiert. Sind $f, g \in T(\Delta)$ solche Treppenfunktionen mit den konstanten Werten f_i, g_i auf den Intervallen $]x_{i-1}, x_i[$, so ist*

$$(f.g) = \int_a^b f(x) \cdot g(x)\, dx = h \cdot \sum_{i=1}^n f_i \cdot g_i.$$

Bis auf den Faktor $h = x_{i+1} - x_i$ ist es das euklidische SKP für die n-tupel $(f_i), (g_i) \in \mathbb{R}^n$.

Die Eigenschaften (iv),(v) aus Abschnitt 1.5.1 sind fundamental für den Begriff der Länge im $\mathbb{R}^n$. Wir verallgemeinern sie wie folgt:

Definition 1.27 *Es sei V ein $\mathbb{R}$-Vektorraum. Eine Norm auf V ist eine Abbildung*

$$\| \quad \|: V \to \mathbb{R}, \qquad \mathbf{x} \mapsto \| \mathbf{x} \|,$$

welche definit und homogen ist und die Dreiecksungleichung erfüllt. Das sind also genau die Eigenschaften (iv),(v),(vii) aus Abschnitt 1.5.1. Ein $\mathbb{R}$-Vektorraum zusammen mit einer Norm heißt normierter *Vektorraum.*

Die Beweise der Eigenschaften (iv)-(vii) für die euklidische Norm in Abschnitt 1.5.1 haben nur die Eigenschaften (i)-(iii) des euklidischen Skalarprodukts auf $V = \mathbb{R}^n$ benutzt. Deswegen gilt

Satz 1.30 *Es sei $(\quad . \quad)$ ein SKP auf dem $\mathbb{R}$-Vektorraum V. Dann wird durch*

$$\| \mathbf{x} \| := \sqrt{(\mathbf{x}.\mathbf{x})}$$

eine Norm auf V definiert. Insbesondere gilt die Cauchy-Schwarz-Ungleichung. Man sagt auch: Diese Norm wird vom SKP $(\quad . \quad)$ erzeugt.

Jede Norm auf einem $\mathbb{R}$-Vektorraum definiert einen Abstand (Metrik) durch

$$d(\mathbf{x}, \mathbf{y}) = \| \mathbf{x} - \mathbf{y} \| .$$

Ist die Norm von einem SKP erzeugt, so erfüllt sie die *Parallelogrammgleichung*

$$\begin{aligned} \| \mathbf{x} + \mathbf{y} \|^2 + \| \mathbf{x} - \mathbf{y} \|^2 &= \| \mathbf{x} \|^2 + 2(\mathbf{x}.\mathbf{y}) + \| \mathbf{y} \|^2 + \| \mathbf{x} \|^2 - 2(\mathbf{x}.\mathbf{y}) + \| \mathbf{y} \|^2 \\ &= 2\left(\| \mathbf{x} \|^2 + \| \mathbf{y} \|^2 \right) \end{aligned}$$

Beispiel 1.40 *Schon auf $V = \mathbb{R}^n$ gibt es Normen, die nicht durch ein Skalarprodukt erzeugt werden, z.B.*

$$\| \mathbf{x} \|_1 := \sum_{i=1}^n |x_i|, \qquad \text{oder} \qquad \| \mathbf{x} \|_\infty := \max_{i=1}^n |x_i|, \quad \text{die Maximumsnorm.}$$

Analog dazu lassen sich auf $C[a, b]$ Normen definieren durch

$$\| f \|_1 := \int_a^b |f(x)|\, dx \qquad \| f \|_\infty := \max_{x \in [a,b]} |f(x)|.$$

Das SKP aus Beispiel 1.38 definiert die Norm

$$\| f \|_2 = \sqrt{\int_a^b f(x)^2\, dx}.$$

Aufgabe 1.58 *Den $\mathbb{R}$-Vektorraum $\mathbb{R}^\infty$ kann man auffassen als den Raum aller Folgen $(a_\nu)_{\nu\in\mathbb{N}}$ reeller Zahlen. Mit $\ell^2 \subset \mathbb{R}^\infty$ bezeichnet man die Teilmenge aller* quadrat-konvergenten *Folgen, also die Menge der Folgen $(a_\nu) \in \mathbb{R}^n$, für welche eine reelle Konstante A existiert mit*

$$\sum_\nu a_\nu^2 < A.$$

Zeigen Sie: ℓ^2 ist ein Untervektorraum von $\mathbb{R}^\infty$ und durch

$$(\mathbf{a}.\mathbf{b}) := \sum_{\nu\in\mathbb{N}} a_\nu \cdot b_\nu$$

wird auf diesem Untervektorraum ein SKP definiert.

1.5.3 Orthogonalität

Zwei Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ stehen senkrecht aufeinander, wenn sie einen Winkel von 90^0 einschließen. Das ist genau dann der Fall, wenn ihr Zwischenwinkel den Cosinus-Wert $= 0$ hat. Und das wiederum passiert genau dann, wenn ihr euklidisches Skalarprodukt $(\mathbf{x}.\mathbf{y}) = 0$ ist. Das verallgemeinern wir jetzt auf $\mathbb{R}$-Vektorräume mit SKP.

In diesem Abschnitt 1.5.3 sei also V ein $\mathbb{R}$-Vektorraum mit einem SKP, das wir wie immer $(\quad.\quad)$ schreiben.

Definition 1.28 *Zwei Vektoren $\mathbf{x}, \mathbf{y} \in V$ heißen* orthogonal, *in Zeichen $\mathbf{x} \perp \mathbf{y}$, wenn $(\mathbf{x}.\mathbf{y}) = 0$ ist. Hier ist $\mathbf{x} = \mathbf{0}$ oder $\mathbf{y} = \mathbf{0}$ ausdrücklich zugelassen.*

Beispiel 1.41 *Der Nullvektor $\mathbf{0} \in V$ ist orthogonal zu jedem Vektor $\mathbf{x} \in V$. Wenn $\mathbf{x} \in V$ zu sich selbst orthogonal ist, dann folgt aus (iii), dass $\mathbf{x} = \mathbf{0}$ ist. Deswegen ist $\mathbf{0}$ auch der einzige Vektor in V, der zu allen $\mathbf{x} \in V$ orthogonal ist.*

Satz 1.31 (Abstrakter Pythagoras) *Es seien $\mathbf{x}_1, ..., \mathbf{x}_r \in V$ paarweise zueinander orthogonal, also*

$$(\mathbf{x}_k.\mathbf{x}_l) = 0 \text{ für alle } k \neq l.$$

Dann ist

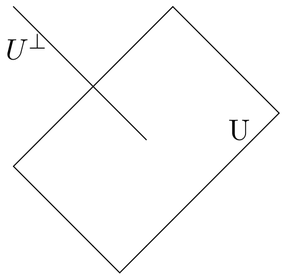
$$\| \mathbf{x}_1 + ... + \mathbf{x}_r \|^2 = \| \mathbf{x}_1 \|^2 + ... + \| \mathbf{x}_k \|^2 .$$

Der Beweis ist der gleiche wie für Satz 1.29. □

Definition 1.29 *Ist $A \subset V$ eine beliebige Menge, so sei*

$$A^\perp := \{\mathbf{x} \in V : (\mathbf{x}.\mathbf{a}) = 0 \text{ für alle } \mathbf{a} \in A\}$$

die Menge der Vektoren $\mathbf{x}$, die zu allen Vektoren aus A orthogonal sind. Ist insbesondere $A = U \subset \mathbb{R}^n$ ein Untervektorraum, so nennen wir $U^\perp$ das orthogonale Komplement *zu U in V. Wenn $\mathbf{a} \in V$ ein einziger Vektor ist, dann schreiben wir kurz $\mathbf{a}^\perp$ für $\{\mathbf{a}\}^\perp$.*



Unmittelbar aus der Definition folgt, dass $A^\perp$ ein Untervektorraum von V ist. Und es ist $A^\perp = (span(A))^\perp$. Für Teilmengen $A_1 \subset A_2 \subset V$ gilt immer $A_2^\perp \subset A_1^\perp$. Für jeden Untervektorraum $U \subset V$ ist

$$U \cap U^\perp = \{\mathbf{0}\}, \qquad U \subset (U^\perp)^\perp.$$

Beispiel 1.42 *Es sei $V = \mathbb{R}^n$ mit dem euklidischen SKP und $A = \{\mathbf{a}_1, ..., \mathbf{a}_m\} \subset \mathbb{R}^n$ eine endliche Menge. Dann ist also*

$$\begin{aligned} \mathbf{x} \in A^\perp \quad &\Leftrightarrow \quad (\mathbf{a}_1.\mathbf{x}) = ... = (\mathbf{a}_m.\mathbf{x}) = 0 \\ &\Leftrightarrow \quad \sum_{\nu=1}^n a_{1,\nu} x_\nu = ... = \sum_{\nu=1}^n a_{m,\nu} x_\nu = 0 \\ &\Leftrightarrow \quad \sum_{\nu=1}^n a_{\mu,\nu} x_\nu = 0 \qquad \text{für} \qquad \mu = 1, ..., m \end{aligned}$$

Die Vektoren $\mathbf{x} \in \{\mathbf{a}_1, ..., \mathbf{a}_m\}^\perp$ sind also genau die Lösungen des linearen Gleichungssystems, dessen Koeffizientenmatrix aus den Zeilenvektoren $\mathbf{a}_1, ..., \mathbf{a}_m$ zusammengesetzt ist. Anders ausgedrückt: Ist U der Lösungsraum des LGS $A \cdot \mathbf{x} = \mathbf{0}$ und $Z(A) = span(\mathbf{a}_1, ..., \mathbf{a}_m)$ der Zeilenraum der Koeffizientenmatrix A, dann ist

$$\{\mathbf{a}_1, ..., \mathbf{a}_m\}^\perp = Z(A)^\perp = U.$$

Satz 1.26 zeigt in dieser Situation

$$dim\{\mathbf{a}_1, ..., \mathbf{a}_m\}^\perp = n - dim\, span(\mathbf{a}_1, ..., \mathbf{a}_m).$$

Im Fall $m = 1$ sehen wir: Für $\mathbf{0} \neq \mathbf{a} \in V$ ist $\mathbf{a}^\perp$ eine Hyperebene.

Beispiel 1.43 *Es sei $V = span(\mathbf{v}_1, ..., \mathbf{v}_n)$ endlich-erzeugt und $U = span(\mathbf{u}_1, ..., \mathbf{u}_m) \subset V$ ein Untervektorraum. Ein Vektor $\mathbf{v} = \sum c_\nu \mathbf{v}_\nu$ steht senkrecht auf U genau dann, wenn*

$$(\mathbf{v}, \mathbf{u}_\mu) = \sum_{\nu=1}^n c_\nu (\mathbf{v}_\nu.\mathbf{u}_\mu) = 0 \quad \text{für} \quad \mu = 1, ..., m.$$

Das ist also genau dann der Fall, wenn der Koeffizientenvektor $\mathbf{c} = (c_1, ..., c_n)^t$ Lösungsvektor des LGS mit der Koeffizientenmatrix

$$A = (a_{\mu,\nu}) = (\mathbf{u}_\mu.\mathbf{v}_\nu), \quad \mu \text{ Zeile, } \nu \text{ Spalte}$$

ist.

Das Skalarprodukt (.) definiert eine Norm $\parallel \quad \parallel$ auf V und damit einen Abstand $\parallel \mathbf{x} - \mathbf{y} \parallel$ für zwei Vektoren $\mathbf{x}, \mathbf{y} \in V$.

Satz 1.32 *Gegeben seien ein Untervektorraum $U \subset V$ und ein Vektor $\mathbf{v} \in V$. Für $\mathbf{u} \in U$ sind äquivalent:*

a) $(\mathbf{v} - \mathbf{u}) \perp U$.

b) $\| \mathbf{v} - \mathbf{u} \| = \min_{\mathbf{x} \in U} \| \mathbf{v} - \mathbf{x} \|$.

Beweis a) $\Rightarrow$ b): Für $\mathbf{x} \in U$ ist nach Pythagoras

$$\| \mathbf{v} - \mathbf{x} \|^2 = \| (\mathbf{v} - \mathbf{u}) + (\mathbf{u} - \mathbf{x}) \|^2 = \| \mathbf{v} - \mathbf{u} \|^2 + \| \mathbf{u} - \mathbf{x} \|^2 \geq \| \mathbf{v} - \mathbf{u} \|^2$$

wegen $\mathbf{u} - \mathbf{x} \in U$ und $(\mathbf{v} - \mathbf{u}) \perp (\mathbf{u} - \mathbf{x})$.

b) $\Rightarrow$ a): Für jedes $\mathbf{x} \in U$ hat die Funktion

$$f(t) := \| \mathbf{v} - (\mathbf{u} + t \cdot \mathbf{x}) \|^2, \quad t \in \mathbb{R},$$

bei $t = 0$ ein Minimum. Nun ist

$$f(t) = \| (\mathbf{v} - \mathbf{u}) - t \cdot \mathbf{x} \|^2 = \| \mathbf{v} - \mathbf{u} \|^2 - 2t \cdot (\mathbf{v} - \mathbf{u}.\mathbf{x}) + t^2 \cdot \| \mathbf{x} \|^2$$

und im Minimum $t = 0$ ist

$$\frac{df}{dt}(0) = 2 \cdot (\mathbf{v} - \mathbf{u}.\mathbf{x}) = 0.$$

Weil diese Orthogonalität für alle $\mathbf{x} \in U$ gilt, haben wir $(\mathbf{v} - \mathbf{u}) \perp U$. □

Satz 1.33 *Falls in Satz 1.32 der Vektor $\mathbf{u} \in U$ existiert, dann ist er durch $\mathbf{v}$ eindeutig bestimmt.*

Beweis. Es sei $(\mathbf{v} - \mathbf{u}_1) \perp U$ und $(\mathbf{v} - \mathbf{u}_2) \perp U$. Dann ist auch

$$\mathbf{u}_1 - \mathbf{u}_2 = (\mathbf{u}_1 - \mathbf{v}) + (\mathbf{v} - \mathbf{u}_2) \quad \perp \quad U.$$

Wegen $\mathbf{u}_1 - \mathbf{u}_2 \in U$ folgt daraus $\mathbf{u}_1 - \mathbf{u}_2 = \mathbf{0}$. □

Definition 1.30 *Der nach Satz 1.33 eindeutig bestimmte Vektor $\mathbf{u} \in U$ heißt - wenn er existiert - die* Orthogonalprojektion $P_U(\mathbf{v})$ *von $\mathbf{v}$ in den Untervektorraum U. Der Abstand*

$$\| \mathbf{v} - P_U(\mathbf{v}) \|,$$

der nach Satz 1.32 der minimale Abstand von $\mathbf{v}$ zu Vektoren $\mathbf{x} \in U$ ist, heißt der Abstand $d(\mathbf{v}, U)$ *des Vektors $\mathbf{v}$ zum Unterraum U.*

Beispiel 1.44 *Es sei $U = \mathbb{R} \cdot \mathbf{w} \subset V$ eine Gerade. Für $\mathbf{u} = t \cdot \mathbf{w}$ ist*

$$(\mathbf{v} - \mathbf{u}.\mathbf{w}) = (\mathbf{v}.\mathbf{w}) - t \cdot \| \mathbf{w} \|^2 = 0$$

genau dann, wenn

$$t = \frac{(\mathbf{v}.\mathbf{w})}{\| \mathbf{w} \|^2},$$

und

$$P_U(\mathbf{v}) = \frac{(\mathbf{v}.\mathbf{w})}{\| \mathbf{w} \|^2} \cdot \mathbf{w}$$

ist die Orthogonalprojektion.

Dieses Beispiel verallgemeinern wir jetzt auf Untervektorräume $U \subset V$ mit einer Basis $\mathbf{u}_1, ..., \mathbf{u}_r$.

Definition 1.31 *Die* Gramsche Matrix *der Basis* $\mathbf{u}_1, ..., \mathbf{u}_r$ *ist die* $r \times r$*-Matrix*

$$G(\mathbf{u}_1, ..., \mathbf{u}_r) = ((\mathbf{u}_i.\mathbf{u}_j)_{i,j=1,...,r})$$

der Skalarprodukte.

Satz 1.34 *Die Gramsche Matrix einer Basis* $\mathbf{u}_1, ..., \mathbf{u}_r$ *von* $U \subset V$ *hat den Vollrang* r*.*

Beweis. Wir testen die Spalten G^j der Gramschen Matrix auf lineare Unabhängigkeit. Sei dazu

$$\sum_{j=1}^{r} G^j c_j = \mathbf{0}.$$

Ausgeschrieben lautet diese Vektorgleichung

$$\begin{array}{rcl} (\mathbf{u}_1.\sum c_j \mathbf{u}_j) & = & 0 \\ & \vdots & \\ (\mathbf{u}_r.\sum c_j \mathbf{u}_j) & = & 0 \end{array}$$

Der Vektor $\mathbf{u} = \sum c_j \mathbf{u}_j \in U$ ist also orthogonal zu allen Basisvektoren $\mathbf{u}_i \in U$ und damit zum ganzen Unterraum U. Dann ist er auch orthogonal zu sich selbst und damit der Nullvektor

$$\sum_{j=1}^{r} c_j \mathbf{u}_j = \mathbf{0}.$$

Weil die Vektoren $\mathbf{u}_1, ..., \mathbf{u}_r$ linear unabhängig sind, folgt daraus $c_1 = ... = c_r = 0$. $\square$

Die Orthogonalprojektion in endlich-dimensionale Untervektorräume existiert:

Satz 1.35 *Es sei* $U \subset V$ *ein Untervektorraum mit einer Basis* $\mathbf{u}_1, ..., \mathbf{u}_r$*. Für* $\mathbf{v} \in V$ *ist die Orthogonalprojektion*

$$P_U(\mathbf{v}) = \sum_{j=1}^{r} c_j \mathbf{u}_j,$$

wobei der Koeffizientenvektor $\mathbf{c} = (c_1, ..., c_r)^t$ *eindeutig bestimmt ist als Lösungsvektor des LGS*

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

Dabei ist A *die Gramsche Matrix der Basis* $\mathbf{u}_1, ..., \mathbf{u}_r$ *und die rechte Seite der Vektor*

$$\mathbf{b} = ((\mathbf{v}.\mathbf{u}_1), ..., (\mathbf{v}.\mathbf{u}_r))^t .$$

Beweis. Wir betrachten einen Vektor $\mathbf{u} = \sum_{j=1}^{r} c_j \mathbf{u}_j \in U$ mit Koeffizienten $c_1, ..., c_r \in \mathbb{R}$. Für diesen Vektor ist

$$\begin{array}{rcl} \mathbf{u} - \mathbf{v} \perp U & \Leftrightarrow & (\mathbf{u} - \mathbf{v}.\mathbf{u}_i) = 0 \text{ für } i = 1, ..., r \\ & \Leftrightarrow & \left((\sum_{j=1}^{r} c_j \mathbf{u}_j - \mathbf{v}).\mathbf{u}_i \right) = 0 \\ & \Leftrightarrow & \sum_{j=1}^{r} (\mathbf{u}_i.\mathbf{u}_j) c_j = (\mathbf{u}_i.\mathbf{v}) \text{ für } i = 1, ..., r \\ & \Leftrightarrow & A \cdot \mathbf{c} = \mathbf{b} \end{array}$$

Die Gramsche Matrix A hat den Maximalrang r (Satz 1.34). Deswegen ist dieses LGS eindeutig lösbar und definiert mit seiner Lösung die Orthogonalprojektion $P_U(\mathbf{v})$. $\square$

Satz 1.36 *Es sei U ein endlich-dimensionaler Untervektorraum von V.*
a) Dann ist $V = U \oplus U^\perp$ eine direkte Summe.
b) Für $\mathbf{v} \in V$ ist $P_{U^\perp}(\mathbf{v}) = \mathbf{v} - P_U(\mathbf{v})$.

Beweis. a) Jeder Vektor $\mathbf{v} \in V$ schreibt sich

$$\mathbf{v} = P_U(\mathbf{v}) + (\mathbf{v} - P_U(\mathbf{v})) \quad \text{mit} \quad P_U(\mathbf{v}) \in U, \, (\mathbf{v} - P_U(\mathbf{v})) \in U^\perp.$$

Dies zeigt $V = U + U^\perp$. Die Eigenschaft $U \cap U^\perp = \{\mathbf{0}\}$ wurde oben schon gezeigt.

b) Es ist klar, dass $\mathbf{v} - P_U(\mathbf{v})$ zu $U^\perp$ gehört. Zu zeigen ist noch

$$\mathbf{v} - (\mathbf{v} - P_U(\mathbf{v})) = P_U(\mathbf{v}) \in (U^\perp)^\perp.$$

Aber dies folgt aus

$$P_U(\mathbf{v}) \in U \subset (U^\perp)^\perp.$$ $\square$

Den Abstand eines Punktes (Vektors) im $\mathbb{R}^3$ von einer Geraden oder Ebene zu bestimmen, das ist ein Problem von praktischer Bedeutung in der Geometrie. Das ist der Fall, wo V kleine Dimension hat. Aber auch der Fall, wo die Dimension von V groß oder gar $= \infty$ ist, hat beachtliche praktische Bedeutung.

Beispiel 1.45 *Es sei $V = C[a, b]$ der Vektorraum der stetigen Funktionen auf dem Intervall $[a, b]$ mit dem SKP aus Beispiel 1.38. Weiter sei $U \subset V$ der Unterraum T der Treppenfunktionen aus Beispiel 1.8. Eine Basis von U sind die Treppenfunktionen t_i, $i = 1, ..., n$ aus Beispiel 1.29. Die Gramsche Matrix dieser Basis hat die Einträge*

$$(t_i.t_j) = \int_a^b t_i(x)t_j(x)\,dx = \begin{cases} h & \textit{falls} \quad i = j \\ 0 & \textit{falls} \quad i \neq j \end{cases}$$

Die Orthogonalprojektion einer Funktion $f \in C[a, b]$ in den Unterraum T ist die Funktion $\sum_{i=1}^n c_i t_i$, wo die Koeffizienten c_i nach Satz 1.35 durch das Gleichungssystem

$$h \cdot c_i = (f.t_i) = \int_{x_{i-1}}^{x_i} f(x)\,dx, \quad i = 1, ..., n,$$

bestimmt sind. Damit wird die Orthogonalprojektion

$$P_U(f) = \frac{1}{h} \cdot \sum_{i=1}^n \left(\int_{x_{i-1}}^{x_i} f(x)\,dx \right) \cdot t_i.$$

Dahinter steht die allgemeine Aufgabe, einen Vektor aus V durch einen Vektor $\mathbf{u} \in U$ optimal zu approximieren.

Satz 1.37 *Für jeden linearen Unterraum U eines* endlich-dimensionalen *Vektorraums V gilt*
a) $dim(U^\perp) = n - dim(U)$,
b) $(U^\perp)^\perp = U$,

Beweis. a) folgt aus Satz 1.36 a).

b): Allgemein gilt $U \subset (U^\perp)^\perp$. Nach a) ist $dim(U^\perp)^\perp = n - (n - dimU) = dimU$. Nun folgt $U = (U^\perp)^\perp$ aus Satz 1.23. □

Definition 1.32 *a) Eine Teilmenge $A \subset V$ heißt* orthogonal, *wenn die Elemente von A paarweise aufeinander senkrecht stehen. D.h., für $\mathbf{u} \neq \mathbf{v} \in A$ gilt*

$$(\mathbf{u}.\mathbf{v}) = 0.$$

Eine Basis B von V heißt Orthogonalbasis, *wenn die Menge B orthogonal ist.*

b) Die Teilmenge A heißt orthonormal, *wenn zusätzlich zu a) gilt*

$$\| \mathbf{u} \| = 1 \quad \textit{für alle } \mathbf{u} \in A \qquad \textit{(Normalität).}$$

Eine Basis B heißt Orthonormalbasis *(ONB), wenn B als Menge orthonormal ist.*

Die Gramsche Matrix einer Orthogonalbasis hat Diagonalform, die Gramsche Matrix einer ONB ist die *Einheitsmatrix*

$$\mathbb{1} := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Das hat sehr erfreuliche Konsequenzen:

Ist nämlich $U \subset V$ ein Untervektorraum mit einer Orthogonalbasis $\mathbf{u}_1, ..., \mathbf{u}_r$. Dann ist für $\mathbf{v} \in V$ nach Satz 1.35

$$P_U(\mathbf{v}) = \sum_{i=1}^r c_i \cdot \mathbf{u}_i \quad \text{mit} \quad c_i = \frac{\mathbf{v}.\mathbf{u}_i}{\| \mathbf{u}_i \|^2}.$$

Für die Länge von $P_U(\mathbf{v})$ folgt daraus

$$\| P_U(\mathbf{v}) \|^2 = \sum_{i,j=1}^r c_i c_j \cdot (\mathbf{u}_i.\mathbf{u}_j) = \sum_{i=1}^r c_i^2 \cdot \| \mathbf{u}_i \|^2 = \sum_{i=1}^r \frac{(\mathbf{v}.\mathbf{u}_i)^2}{\| \mathbf{u}_i \|^2}.$$

Insbesondere für $\mathbf{u} \in U$ ist $\mathbf{u} = P_U(\mathbf{u})$ und

$$\mathbf{u} = \sum_{i=1}^r \frac{(\mathbf{u}.\mathbf{u}_i)}{\| \mathbf{u}_i \|^2} \cdot \mathbf{u}_i.$$

Satz 1.38 *Ist A orthogonal mit $\mathbf{0} \notin A$, dann ist A linear unabhängig.*

Beweis. Sei eine lineare Relation

$$\sum_{i=1}^r c_i \cdot \mathbf{a}_i = \mathbf{0}$$

zwischen Vektoren $\mathbf{a}_i \in A$ gegeben. Dann ist für jedes $j = 1, ..., r$

$$0 = \left(\mathbf{a}_j . \sum_{i=1}^r c_i \cdot \mathbf{a}_i \right) = \sum_{i=1}^r c_i \cdot (\mathbf{a}_j.\mathbf{a}_i) = c_j \| \mathbf{a}_j \|^2$$

und daraus folgt $c_j = 0$. □

Satz 1.39 (Schmidtsche Orthonormalisierung) *Jeder endlich-dimensionale Vektorraum V mit SKP besitzt eine* Orthonormalbasis.

Beweis. Wir gehen von einer beliebigen Basis $\mathbf{v}_1, ..., \mathbf{v}_m$ des Vektorraums V aus. Für $i = 1, ..., m$ sind die Vektoren $\mathbf{v}_1, ..., \mathbf{v}_i$ eine Basis des Unterraums $V_i := span(\mathbf{v}_1, ..., \mathbf{v}_i)$.

Als erstes normalisieren wir $\mathbf{v}_1$:

$$\mathbf{u}_1 := \frac{1}{\parallel \mathbf{v}_1 \parallel} \mathbf{v}_1.$$

Der Vektor $\mathbf{u}_1$ ist dann eine ONB von V_1.

Dann ersetzen wir $\mathbf{v}_2$ durch

$$\mathbf{u}'_2 := \mathbf{v}_2 - (\mathbf{u}_1 . \mathbf{v}_2)\mathbf{u}_1 = P_{U_1{}^\perp}(\mathbf{v}_2)$$

(vgl. Satz 1.36). Als nächstes normieren wir $\mathbf{u}'_2$:

$$\mathbf{u}_2 := \frac{1}{\parallel \mathbf{u}'_2 \parallel} \mathbf{u}'_2.$$

Damit wird $\mathbf{u}_1, \mathbf{u}_2$ eine ONB von V_2.

Dieses Verfahren können wir mit jedem der Vektoren $\mathbf{v}_k$ wiederholen: Haben wir für ein $k \le m$ schon erreicht, dass

$$\left\{ \begin{array}{l} (\mathbf{u}_j . \mathbf{u}_l) = 0 \text{ falls } j \ne l \le k \\ \parallel \mathbf{u}_j \parallel = 1 \text{ für } j = 1, ..., k, \end{array} \right.$$

wobei $\mathbf{u}_1, ..., \mathbf{u}_k \in V_k$ Linearkombinationen der Vektoren $\mathbf{v}_1, ..., \mathbf{v}_k$ sind, so definieren wir

$$\mathbf{u}'_{k+1} := \mathbf{v}_{k+1} - (\mathbf{u}_1 . \mathbf{v}_{k+1})\mathbf{u}_1 - ... - (\mathbf{u}_k . \mathbf{v}_{k+1}) u_k = \mathbf{v}_{k+1} - P_{U_k}(\mathbf{v}_{k+1}) = P_{U_k{}^\perp}(\mathbf{v}_{k+1})$$

und normieren

$$\mathbf{u}_{k+1} := \frac{1}{\parallel \mathbf{u}'_{k+1} \parallel} \mathbf{u}'_{k+1}.$$

Damit haben wir eine ONB $\mathbf{u}_1, ..., \mathbf{u}_{k+1} \in V_{k+1}$ konstruiert. Nach endlich vielen derartigen Schritten haben wir eine Orthonormalbasis für V. $\square$

Dieser Beweis funktionierte ja ganz schön, aber ein Wort der Warnung: Will man dieses Verfahren konkret anwenden, so bringen die Normierungen Zahlenfaktoren ins Spiel, die sich meist äußerst unangenehm aufschaukeln!

Aufgabe 1.59 *Es sei $U \subset \mathbb{R}^5$ der von den Vektoren $(1, 2, 0, 2, 1)$ und $(1, 1, 1, 1, 1)$ aufgespannte Unterraum. Bestimmen Sie eine Orthonormalbasis von U und von $U^\perp$.*

Aufgabe 1.60 *Es seien $L_1 : \mathbf{v}_1 + \mathbb{R}\mathbf{w}_1$ und $L_2 : \mathbf{v}_2 + \mathbb{R}\mathbf{w}_2$ zwei Geraden im $\mathbb{R}^n$ mit linear unabhängigen Richtungsvektoren $\mathbf{w}_1, \mathbf{w}_2$. Zeigen Sie: Auf L_1 existiert genau ein Punkt $\mathbf{x}_1$ und auf L_2 genau ein Punkt $\mathbf{x}_2$ so, dass $\mathbf{x}_1 - \mathbf{x}_2$ auf $\mathbf{w}_1$ und $\mathbf{w}_2$ senkrecht steht. ($\parallel \mathbf{x}_1 - \mathbf{x}_2 \parallel$ heißt Abstand der Geraden L_1 und L_2.)*

Aufgabe 1.61 *a) Ergänzen Sie die beiden Vektoren*

$$\mathbf{v}_1 = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right) \quad \textit{und} \quad \mathbf{v}_2 = \left(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right)$$

zu einer Orthonormalbasis des $\mathbb{R}^4$.
b) Es sei $E \subset \mathbb{R}^4$ die von $\mathbf{v}_1$ und $\mathbf{v}_2$ aufgespannte Ebene. Bestimmen Sie für jeden Vektor $\mathbf{x} = (x_1, x_2, x_3, x_4) \in \mathbb{R}^4$ die Orthogonalprojektion $\mathbf{x}' \in E$, d.h., den Vektor $\mathbf{x}' \in E$ mit $(\mathbf{x} - \mathbf{x}') \perp E$.

Aufgabe 1.62 *Im* $\mathbb{R}^4$ *seien die aufeinander senkrecht stehenden Vektoren* $\mathbf{v}_1 = (1,2,3,4)$ *und* $\mathbf{v}_2 = (4,1,-2,0)$ *gegeben. Man ergänze* $\mathbf{v}_1, \mathbf{v}_2$ *zu einer Orthogonalbasis des* $\mathbb{R}^4$*.*

Aufgabe 1.63 *Im* $\mathbb{R}^4$ *seien die Vektoren*

$$\mathbf{u}_1 = (1,1,0,0)^t, \quad \mathbf{u}_2 = (0,-1,-1,0)^t, \quad \mathbf{u}_3 = (0,0,1,1,)^t$$

gegeben, sowie der Vektor $\mathbf{v} = (1,2,0,0)$*. Bestimmen Sie die Orthogonalprojektionen*

$$\begin{array}{lll} P_L(\mathbf{v}) & \textit{von } \mathbf{v} & \textit{in die Gerade } L = \mathbb{R} \cdot \mathbf{u}_1 \\ P_E(\mathbf{v}) & \textit{von } \mathbf{v} & \textit{in die Ebene } E = \mathbb{R} \cdot \mathbf{u}_1 + \mathbb{R} \cdot \mathbf{u}_2 \\ P_H(\mathbf{v}) & \textit{von } \mathbf{v} & \textit{in die Hyperebene } H = span(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3) \end{array}$$

sowie die Abstände $d(\mathbf{v}, L)$*,* $d(\mathbf{v}, E)$ *und* $d(\mathbf{v}, H)$*.*

Aufgabe 1.64 *a) Die Hyperebene* $H \subset \mathbb{R}^n$ *sei gegeben durch die Gleichung* $(\mathbf{a}.\mathbf{x}) = 0$ *mit* $\mathbf{0} \neq \mathbf{a} \in \mathbb{R}^n$*. Zeigen Sie: Der Vektor* $\mathbf{v} \in \mathbb{R}^n$ *hat von dieser Hyperebene den Abstand*

$$d(\mathbf{v}, H) = \frac{|(\mathbf{a}.\mathbf{v})|}{\| \mathbf{a} \|}.$$

b) Jetzt sei die Hyperebene durch eine Gleichung $(\mathbf{a}.\mathbf{x}) + \alpha = 0$ *mit* $0 \neq \alpha \in \mathbb{R}$ *gegeben. Zeigen Sie:*

$$d(\mathbf{v}, H) = \frac{|(\mathbf{a}.\mathbf{v}) + \alpha|}{\| \mathbf{a} \|}.$$

c) Im $\mathbb{R}^2$ *seien zwei Geraden* L, M *durch ihre Gleichungen*

$$L: (\mathbf{a}.\mathbf{x}) + \alpha = 0, \quad M: (\mathbf{b}.\mathbf{x}) + \beta = 0$$

mit $\alpha, \beta \in \mathbb{R}$ *und linear unabhängigen Vektoren* $\mathbf{a}, \mathbf{b} \in \mathbb{R}^2$ *gegeben. Eine Winkelhalbierende zu diesen Geraden ist eine Gerade, deren Punkte von* L *und* M *den gleichen Abstand haben. Bestimmen Sie die Gleichungen der beiden Winkelhalbierenden zu* L *und* M*.*

Aufgabe 1.65 *Bestimmen Sie eine ONB des Unterraums*

$$U := \{(u_1 + ... + u_5)^t \in \mathbb{R}^5 : \sum_{i=1}^{5} u_i = 0\} \subset \mathbb{R}^5.$$

2 Matrizen

Wir betrachten hier Abbildungen $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ oder allgemeiner $\Phi : V \to W$ zwischen $\mathbb{R}$-Vektorräumen V und W. Eine derartige Abbildung ordnet jedem Vektor $\mathbf{x} \in V$ einen Bildvektor $\Phi(\mathbf{x}) \in W$ zu.

Definition 2.1 *Eine Abbildung Φ heißt* injektiv*, falls gilt: $\Phi(\mathbf{x}) = \Phi(\mathbf{y}) \Rightarrow \mathbf{x} = \mathbf{y}$.*
Φ heißt surjektiv*, falls gilt: Zu jedem $\mathbf{y} \in W$ gibt es ein $\mathbf{x} \in V$ mit $\mathbf{y} = \Phi(\mathbf{x})$. Ein solches $\mathbf{x}$ heißt* Urbildvektor *von $\mathbf{y}$.*
Φ heißt bijektiv*, falls Φ zugleich injektiv und surjektiv ist, d.h., falls es zu jedem $\mathbf{y} \in V$ genau ein $\mathbf{x} \in W$ gibt mit $\mathbf{y} = \Phi(\mathbf{x})$. In diesem Fall ist die* Umkehrabbildung $\mathbf{y} \mapsto \mathbf{x}$ *auch wieder eine Abbildung, sie wird mit $\Phi^{-1} : W \to V$ bezeichnet.*

Für die Umkehrabbildung Φ^{-1} einer bijektiven Abbildung Φ ist nach Definition $\Phi^{-1}(\Phi(\mathbf{x})) = \Phi^{-1}(\mathbf{y}) = \mathbf{x}$. Aber es gilt auch $\Phi(\Phi^{-1}(\mathbf{y})) = \Phi(\mathbf{x}) = \mathbf{y}$.

Die Abbildung $\mathbf{x} \mapsto \mathbf{x}$, die jeden Vektor auf sich selber abbildet, kriegt einen eigenen Namen. Sie heißt die Identität $id = id_V : V \to V$. Zwei Abbildungen $\Phi_1 : U \to V$ und $\Phi_2 : V \to W$ kann man hintereinanderschalten: $\Phi_2 \circ \Phi_1 : U \to W$ ist die Abbildung

$$\mathbf{x} \mapsto \Phi_1(\mathbf{x}) \mapsto \Phi_2(\Phi_1(\mathbf{x})).$$

Wenn $\Phi_2 = \Phi_1^{-1}$, dann gilt also

$$\Phi_2 \circ \Phi_1 = \Phi_1 \circ \Phi_2 = id.$$

Aufgabe 2.1 *Beweisen oder widerlegen Sie: Für alle Mengen A, B, C und Abbildungen $f : A \to B, g : B \to C$ gilt:*
a) Sind f und g injektiv, so auch $g \circ f$.
b) Sind f und g surjektiv, so auch $g \circ f$.
c) Ist f injektiv und g surjektiv, so ist $g \circ f$ bijektiv.
d) Ist $g \circ f$ bijektiv, so ist g surjektiv und f injektiv.
e) Ist $g \circ f$ bijektiv, so ist g injektiv und f surjektiv.

2.1 Bewegungen im $\mathbb{R}^n$ und lineare Abbildungen

Definition 2.2 *Eine* Bewegung *im $\mathbb{R}^n$ ist eine Abbildung*

$$\Phi : \begin{cases} \mathbb{R}^n \to \mathbb{R}^n \\ \mathbf{x} \mapsto \Phi(\mathbf{x}), \end{cases}$$

die den Abstand erhält, d.h. eine Abbildung mit der Eigenschaft

$$\parallel \Phi(\mathbf{x}) - \Phi(\mathbf{y}) \parallel = \parallel \mathbf{x} - \mathbf{y} \parallel$$

für alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Wenn man einen „starren Körper“ bewegt, ändern sich die Abstände von Punkten in seinem Inneren nicht. Bei einer Bewegung des $\mathbb{R}^n$ im eben definierten Sinn stellt man sich vor, den ganzen $\mathbb{R}^n$ so zu bewegen wie einen starren Körper. Wir geben Beispiele:

Beispiel 2.1 *Die* Translation *um einen festen Vektor* $\mathbf{a}$

$$T : \mathbf{x} \mapsto \mathbf{x} + \mathbf{a}$$

ist eine Bewegung wegen

$$\parallel T(\mathbf{x}) - T(\mathbf{y}) \parallel = \parallel \mathbf{x} + \mathbf{a} - (\mathbf{y} + \mathbf{a}) \parallel = \parallel \mathbf{x} - \mathbf{y} \parallel .$$

Beispiel 2.2 *Die* Punktspiegelung *am Ursprung*

$$\Phi : \mathbf{x} \mapsto -\mathbf{x}$$

ist eine Bewegung, weil

$$\parallel \Phi(\mathbf{x}) - \Phi(\mathbf{y}) \parallel = \parallel -\mathbf{x} + \mathbf{y} \parallel = \parallel \mathbf{x} - \mathbf{y} \parallel .$$

Beispiel 2.3 *Wir betrachten eine* Hyperebene

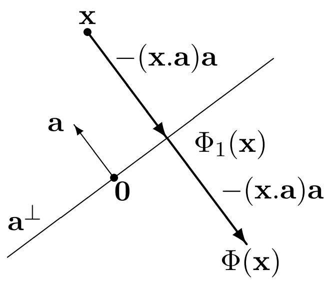
$$a^\perp = \{\mathbf{x} \in \mathbb{R}^n : (\mathbf{a}.\mathbf{x}) = 0\}$$

wobei $\mathbf{a} \neq 0$ *sein soll. Deswegen können wir* $\mathbf{a}$ *als normiert annehmen:* $\parallel \mathbf{a} \parallel = 1$. *In diesem Fall hat die Abbildung*

$$\Phi_1 : \mathbf{x} \mapsto \mathbf{x} - (\mathbf{x}.\mathbf{a})\mathbf{a}$$

die Eigenschaften

$$\Phi_1(\mathbf{x}) \in \mathbf{a}^\perp, \quad (\Phi_1(\mathbf{x}) - \mathbf{x}) \perp \mathbf{a}^\perp,$$



d.h. Φ_1 *ist die Orthogonalprojektion auf* $\mathbf{a}^\perp$. *Wenn wir von* $\mathbf{x}$ *nicht nur einmal* $(\mathbf{x}.\mathbf{a})\mathbf{a}$ *abziehen, sondern zweimal, so ist dies die* Spiegelung an der Hyperebene $\mathbf{a}^\perp$:

$$\Phi : \mathbf{x} \mapsto \mathbf{x} - 2(\mathbf{x}.\mathbf{a})\mathbf{a}.$$

Auch diese Abbildung ist eine Bewegung:

$$\parallel \Phi(\mathbf{x}) - \Phi(\mathbf{y}) \parallel = \parallel \mathbf{x} - 2(\mathbf{x}.\mathbf{a})\mathbf{a} - \mathbf{y} + 2(\mathbf{y}.\mathbf{a})\mathbf{a} \parallel = \parallel \mathbf{x} - \mathbf{y} - 2(\mathbf{x} - \mathbf{y}.\mathbf{a})\mathbf{a} \parallel = \parallel \Phi(\mathbf{x} - \mathbf{y}). \parallel,$$

und es genügt also, zu zeigen $\parallel \Phi(\mathbf{x}) \parallel = \parallel \mathbf{x} \parallel$. *Aber dies folgt aus*

$$\parallel \Phi(\mathbf{x}) \parallel^2 = (\mathbf{x} - 2(\mathbf{x}.\mathbf{a})\mathbf{a} \,.\, \mathbf{x} - 2(\mathbf{x}.\mathbf{a})\mathbf{a}) = \parallel \mathbf{x} \parallel^2 - 4(\mathbf{x}.\mathbf{a})(\mathbf{a}.\mathbf{x}) + 4(\mathbf{x}.\mathbf{a})^2 = \parallel \mathbf{x} \parallel^2 .$$

Beispiel 2.4 *Sind* Φ_1 *und* Φ_2 *Bewegungen, so ist auch* $\Phi_1 \circ \Phi_2$ *eine Bewegung, denn*

$$\parallel \Phi_1(\Phi_2(\mathbf{x})) - \Phi_1(\Phi_2(\mathbf{y})) \parallel = \parallel \Phi_2(\mathbf{x}) - \Phi_2(\mathbf{y}) \parallel = \parallel \mathbf{x} - \mathbf{y} \parallel .$$

Sei Φ eine beliebige Bewegung im $\mathbb{R}^n$ und $\mathbf{a} := \Phi(\mathbf{0}) \in \mathbb{R}^n$. Sei T die Translation $\mathbf{x} \mapsto \mathbf{x} - \mathbf{a}$. Dann ist auch $T \circ \Phi$ eine Bewegung (Beispiele 2.1 und 2.4), und sie hat die Eigenschaft

$$(T \circ \Phi)(\mathbf{0}) = T(\Phi(\mathbf{0})) = T(\mathbf{a}) = \mathbf{a} - \mathbf{a} = \mathbf{0}.$$

Zu jeder Bewegung Φ gibt es also eine Translation T mit $(T \circ \Phi)(\mathbf{0}) = \mathbf{0}$.

Definition 2.3 *Eine Bewegung im* $\mathbb{R}^n$*, die den Nullvektor fest lässt, heißt* orthogonale Transformation.

Satz 2.1 *Jede Bewegung* Φ *im* $\mathbb{R}^n$ *ist ein Produkt* $\Phi = T \circ \Psi$ *einer Translation* T *mit einer orthogonalen Transformation* Ψ*.*

Beweis. Sei die Bewegung Φ gegeben. Ist T irgend eine Translation, so ist $\Psi := T^{-1}\Phi$ orthogonal genau dann, wenn $\Psi(\mathbf{0}) = \mathbf{0}$, d.h. $T(\mathbf{0}) = \Phi(\mathbf{0})$. Wir definieren also ganz einfach $T : \mathbf{x} \mapsto \mathbf{x} + \Phi(\mathbf{0})$. Dann ist $\Psi := T^{-1} \circ \Phi$ eine orthogonale Transformation mit $\Phi = T \circ \Psi$. □

Orthogonale Transformationen Φ haben folgende Eigenschaften:

- $\Phi(\mathbf{0}) = \mathbf{0}$ (nach Definition),
- $\| \Phi(\mathbf{x}) - \Phi(\mathbf{y}) \| = \| \mathbf{x} - \mathbf{y} \|$ (nach Definition einer Bewegung),
- $\| \Phi(\mathbf{x}) \| = \| \mathbf{x} \|$ (vorige Eigenschaft mit $\mathbf{y} = \mathbf{0}$).

Satz 2.2 *Eine orthogonale Transformation erhält das Skalarprodukt zweier Vektoren, d.h. für alle* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *gilt*

$$(\Phi(\mathbf{x}).\Phi(\mathbf{y})) = (\mathbf{x}.\mathbf{y}).$$

Beweis. Es ist

$$\| \Phi(\mathbf{x}) - \Phi(\mathbf{y}) \|^2 = (\Phi(\mathbf{x}) - \Phi(\mathbf{y}).\Phi(\mathbf{x}) - \Phi(\mathbf{y})) = \| \Phi(\mathbf{x}) \|^2 + \| \Phi(\mathbf{y}) \|^2 - 2(\Phi(\mathbf{x}).\Phi(\mathbf{y})).$$

Mit $\| \Phi(\mathbf{x}) \| = \| \mathbf{x} \|$, $\| \Phi(\mathbf{y}) \| = \| \mathbf{y} \|$ und $\| \Phi(\mathbf{x}) - \Phi(\mathbf{y}) \| = \| \mathbf{x} - \mathbf{y} \|$ folgt

$$\begin{aligned} (\Phi(\mathbf{x}).\Phi(\mathbf{y})) &= -\frac{1}{2}(\| \Phi(\mathbf{x}) - \Phi(\mathbf{y}) \|^2 - \| \Phi(\mathbf{x}) \|^2 - \| \Phi(\mathbf{y}) \|^2) \\ &= -\frac{1}{2}(\| \mathbf{x} - \mathbf{y} \|^2 - \| \mathbf{x} \|^2 - \| \mathbf{y} \|^2) \\ &= (\mathbf{x}.\mathbf{y}) \end{aligned}$$

□

Wir haben Bewegungen und damit orthogonale Abbildungen durch die Eigenschaft der *Längentreue* definiert. Satz 2.2 sagt, aus der Längentreue folgt die *Winkeltreue.*

Die Bilder der Vektoren $\mathbf{u}_\nu$ einer ONB $\mathbf{v}_1 := \Phi(\mathbf{u}_1), ..., \mathbf{v}_n := \Phi(\mathbf{u}_n)$ unter einer orthogonalen Transformation Φ haben wegen Satz 2.2 dieselben Skalarprodukte

$$(\mathbf{v}_k.\mathbf{v}_l) = (\mathbf{u}_k.\mathbf{u}_l) = \begin{cases} 1 & \text{falls} \quad k = l, \\ 0 & \text{falls} \quad k \neq l. \end{cases}$$

Daraus folgt mit Satz 1.38 dass die Vektoren $\mathbf{v}_1, ..., \mathbf{v}_n$ linear unabhängig sind. *Also ist das Bild der ONB* $\mathbf{u}_1, ..., \mathbf{u}_k$ *wieder eine ONB.*

Das Bild $\Phi(\mathbf{x})$ eines Vektors $\mathbf{x} = \sum_1^n c_\nu \mathbf{u}_\nu$ ist $\Phi(\mathbf{x}) = \sum_1^n d_\nu \mathbf{v}_\nu$ mit

$$d_\nu = (\Phi(\mathbf{x}).\mathbf{v}_\nu) = (\Phi(\mathbf{x}).\Phi(\mathbf{u}_\nu)) = (\mathbf{x}.\mathbf{u}_\nu) = c_\nu.$$

Für eine orthogonale Abbildung Φ gilt also:

$$\boxed{\Phi(\sum_1^n c_\nu \mathbf{u}_\nu) = \sum_1^n c_\nu \Phi(\mathbf{u}_\nu)}$$

Hieraus folgt insbesondere

$$\boxed{\begin{array}{lll} \Phi(c \cdot \mathbf{x}) = c \cdot \Phi(\mathbf{x}) & \text{für } \mathbf{x} \in \mathbb{R}^n, c \in \mathbb{R} & \text{(Homogenität)} \\ \Phi(\mathbf{x}+\mathbf{y}) = \Phi(\mathbf{x}) + \Phi(\mathbf{y}) & \text{für } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n & \text{(Additivität)} \end{array}} \quad \text{(LINEARITÄT).}$$

Diese Eigenschaft der *Linearität einer Abbildung* hat der linearen Algebra ihren Namen gegeben. Die fundamentalen Beziehungen in der Linearen Algebra werden durch lineare Abbildungen vermittelt. Deswegen halten wir diese Eigenschaft für - nicht nur orthogonale - Abbildungen als Definition fest.

Definition 2.4 *Eine Abbildung $F : V \to W$ von einem $\mathbb{R}$-Vektorraum V in einen $\mathbb{R}$-Vektorraum W heißt* linear, *wenn*

$$F(c_1\mathbf{x}_1 + c_2\mathbf{x}_2) = c_1F(\mathbf{x}_1) + c_2F(\mathbf{x}_2) \text{ für alle } c_1, c_2 \in \mathbb{R}, \mathbf{x}_1, \mathbf{x}_2 \in V.$$

Satz 2.3 *Eine Abbildung $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ ist orthogonal genau dann, wenn sie folgende beiden Eigenschaften hat:*
(i) Φ ist linear.
(ii) Es gibt eine ONB $\mathbf{u}_1, ..., \mathbf{u}_n \in \mathbb{R}^n$, welche unter Φ wieder auf eine ONB $\Phi(\mathbf{u}_1), ..., \Phi(\mathbf{u}_n)$ abgebildet wird.

Beweis. „$\Rightarrow$“ Nach Satz 2.2 bildet eine orthogonale Abbildung jede (nicht nur eine einzige) ONB auf eine ONB ab. Und dass die Linearität eine Konsequenz der Orthogonalität ist, haben wir soeben gesehen.

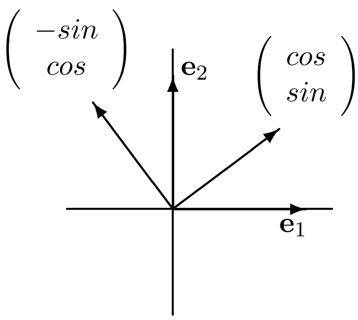
„$\Leftarrow$“ Aus der Linearität folgt $\| \Phi(\mathbf{x}) - \Phi(\mathbf{y}) \| = \| \Phi(\mathbf{x}-\mathbf{y}) \|$ für alle Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Es genügt deswegen $\| \Phi(\mathbf{x}) \| = \| \mathbf{x} \|$ für jeden Vektor $\mathbf{x} \in \mathbb{R}^n$ zu zeigen. Wir schreiben den Vektor $\mathbf{x}$ in unserer ONB als $\mathbf{x} = \sum_1^n c_\nu \mathbf{u}_\nu$. Aus der Linearität folgt $\Phi(\mathbf{x}) = \sum_1^n c_\nu \Phi(\mathbf{u}_\nu)$. Und da sowohl die $\mathbf{u}_\nu$ als auch ihre Bilder $\Phi(\mathbf{u}_\nu)$ eine ONB bilden, ist nach Pythagoras

$$\| \Phi(\mathbf{x}) \|^2 = \sum_1^n c_\nu^2 = \| \mathbf{x} \|^2 .$$

$\square$

Beispiel 2.5 *Rotation im $\mathbb{R}^2$ um einen Winkel φ.*
Rotiert man die beiden Vektoren $\mathbf{e}_1 = (1,0)$ und $\mathbf{e}_2 = (0,1)$ der Standardbasis des $\mathbb{R}^2$ um einen Winkel φ, so erhält man die ONB

$$\Phi(\mathbf{e}_1) = \begin{pmatrix} cos(\varphi) \\ sin(\varphi) \end{pmatrix}, \quad \Phi(\mathbf{e}_2) = \begin{pmatrix} -sin(\varphi) \\ cos(\varphi) \end{pmatrix}$$



des $\mathbb{R}^2$. Es gibt deswegen eine einzige lineare (und dann auch orthogonale) Abbildung $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$, welche diese Drehung der Basisvektoren bewirkt, nämlich

$$\Phi : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} x_1 cos(\varphi) - x_2 sin(\varphi) \\ x_1 sin(\varphi) + x_2 cos(\varphi) \end{pmatrix}.$$

Die Orthogonalität dieser linearen Abbildung ist leicht direkt nachzurechen:

$$\begin{aligned}(x_1cos(\varphi)-x_2sin(\varphi))^2+(x_1sin(\varphi)+x_2cos(\varphi))^2 &= x_1^2cos(\varphi)^2+x_2^2sin(\varphi)^2+x_1^2sin(\varphi)^2+x_2^2cos(\varphi)^2\\ &= x_1^2+x_2^2.\end{aligned}$$

Wir betrachten jetzt allgemein eine lineare Abbidung $\Phi : V \to W$ eines $\mathbb{R}$-Vektorraums V in einen $\mathbb{R}$-Vektorraum W.

Definition 2.5 *Der* Kern *von Φ ist die Teilmenge*

$$Kern(\Phi) := \{\mathbf{v} \in V : \Phi(\mathbf{v}) = \mathbf{0}\}$$

von V und das Bild von Φ ist die Teilmenge

$$Bild(\Phi) := \{\mathbf{w} : \text{ es gibt ein } \mathbf{v} \in V \text{ mit } \mathbf{w} = \Phi(\mathbf{v})\}$$

von W.

Satz 2.4 *a) $Kern(\Phi) \subset V$ und $Bild(\Phi) \subset W$ sind Untervektorräume.*
b) Φ ist genau dann injektiv, wenn $Kern(\Phi) = \{\mathbf{0}\}$.
c) Φ ist genau dann surjektiv, wenn $Bild(\Phi) = W$.

Beweis a) Für $\mathbf{0} \in V$ ist wegen der Linearität von Φ

$$\Phi(\mathbf{0}) = \Phi(0 \cdot \mathbf{0}) = 0 \cdot \Phi(\mathbf{0}) = \mathbf{0},$$

also gehört $\mathbf{0} \in V$ zu $Kern(\Phi)$, und dieser Kern ist nicht leer. Weiter seien $\mathbf{v}_1, \mathbf{v}_2 \in Kern(V)$, also $\Phi(\mathbf{v}_1) = \Phi(\mathbf{v}_2) = \mathbf{0}$, und $c_1, c_2 \in \mathbb{R}$. Aus der Linearität von Φ folgt

$$\Phi(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1\Phi(\mathbf{v}_1) + c_2\Phi(\mathbf{v}_2) = \mathbf{0}$$

und damit $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \in Kern(\Phi)$. Also erfüllt $Kern(\Phi)$ die Bedingung LIN aus Definition 1.10.

Wegen $\Phi(\mathbf{0}) = \mathbf{0} \in W$ gehört $\mathbf{0} \in W$ zu $Bild(\Phi)$, und dieses Bild ist nicht leer. Seien weiter $\mathbf{w}_1 = \Phi(\mathbf{v}_1)$ und $\mathbf{w}_2 = \Phi(\mathbf{v}_2)$ Vektoren in $Bild(\Phi)$, sowie $c_1, c_2 \in R$. Aus der Linearität von Φ folgt

$$c_1\mathbf{w}_1 + c_2\mathbf{w}_2 = c_1\Phi(\mathbf{v}_1) + c_2\Phi(\mathbf{v}_2) = \Phi(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = \Phi(\mathbf{v})$$

mit $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 \in V$. Also hat auch $Bild(\Phi)$ die Eigenschaft LIN aus Definition 1.10.

b) Für $\mathbf{v}_1, \mathbf{v}_2 \in V$ ist

$$\Phi(\mathbf{v}_1) = \Phi(\mathbf{v}_2) \quad \Leftrightarrow \quad \Phi(\mathbf{v}_1 - \mathbf{v}_2) = \mathbf{0} \quad \Leftrightarrow \quad \mathbf{v}_1 - \mathbf{v}_2 \in Kern(\Phi).$$

Ist $Kern(\Phi) = \{\mathbf{0}\}$ so folgt hieraus $\mathbf{v}_1 = \mathbf{v}_2$. Somit ist Φ injektiv. Und wenn Φ injektiv ist, so folgt aus $\Phi(\mathbf{v}) = \mathbf{0} = \Phi(\mathbf{0})$, dass $\mathbf{v} = \mathbf{0}$ ist. Der Kern von Φ ist der Nullraum.

c) Es ist genau dann $Bild(\Phi) = W$, wenn jeder Vektor $\mathbf{w} \in W$ ein Urbild $\mathbf{v} \in V$ mit $\Phi(\mathbf{v}) = \mathbf{w}$ hat. $\square$

Satz 2.5 (Dimensionsformeln) *a) Es ist stets $dim(Bild(\Phi)) \leq dim(V)$.*
b) Es gilt

$$dim(V) = dim(Kern(\Phi)) + dim(Bild(\Phi)).$$

Beweis a). Im Fall $dim(V) = \infty$ ist nichts zu zeigen. Sei also $dim(V) = n$ und $\mathbf{v}_1, ..., \mathbf{v}_n$ eine Basis von V. Jeder Vektor $\mathbf{v} \in V$ ist dann also eine Linearkombination $c_1\mathbf{v}_1 + ... + c_n\mathbf{v}_n$ und jeder Bildvektor $\mathbf{w} \in Bild(\Phi)$ ist eine Linearkombination $\mathbf{w} = c_1\Phi(\mathbf{v}_1) + ... + c_n\Phi(\mathbf{v}_n)$. Dies zeigt, dass die n Vektoren $\Phi(\mathbf{v}_1), ..., \Phi(\mathbf{v}_n)$ den Bildraum $Bild(\Phi)$ aufspannen. Die Behauptung folgt aus dem Basisauswahlsatz Satz 1.21.

b) Wenn einer der beteiligten Vektorräume unendliche Dimension hat, kann man über den Sinn der Dimensionsgleichung streiten. Aber, wenn $Kern(\Phi)$ unendliche Dimension hat, dann auch V wegen $Kern(\Phi) \subset V$. Und wenn $Bild(\Phi)$ unendliche Dimension hat, dann ist V wegen a) unendlich-dimensional. Es bleibt die Aufgabe, die Gleichung für den Fall zu zeigen, wo $dim(Kern(\Phi)) = r$ und $dim(Bild(\Phi)) = s$ endlich sind.

Seien also $\mathbf{v}_1, ..., \mathbf{v}_r \in Kern(\Phi)$ und $\mathbf{w}_1, ..., \mathbf{w}_r \in Bild(\Phi)$ Basen. Nach Definition von $Bild(\Phi)$ gibt es Urbilder $\mathbf{v}_{r+1}, ..., \mathbf{v}_{r+s} \in V$ mit $\Phi(\mathbf{v}_{r+1}) = \mathbf{w}_1, ..., \Phi(\mathbf{v}_{r+s}) = \mathbf{w}_s$. Wir beweisen die Aussage, indem wir zeigen, dass die Vektoren $\mathbf{v}_1, ..., \mathbf{v}_{r+s}$ eine Basis von V bilden.

Aufspannen: Sei $\mathbf{v} \in V$ mit

$$\Phi(\mathbf{v}) = \mathbf{w} = c_1\mathbf{w}_1 + ... + c_s\mathbf{w}_s = c_1\Phi(\mathbf{v}_{r+1}) + ... + c_s\Phi(\mathbf{v}_{r+s}).$$

Dann ist

$$\Phi(\mathbf{v} - c_1\mathbf{v}_{r+1} - ... - c_{r+s}\mathbf{v}_{r+s}) = \mathbf{0}, \quad \text{also} \quad \mathbf{v} - c_1\mathbf{v}_{r+1} - ... - c_s\mathbf{v}_{r+s} \in Kern(\Phi)$$

eine Linearkombination

$$\mathbf{v} - c_1\mathbf{v}_{r+1} - ... - c_s\mathbf{v}_{r+s} = b_1\mathbf{v}_1 + ... + b_r\mathbf{v}_r.$$

Damit folgt

$$\mathbf{v} = b_1\mathbf{v}_1 + ... + b_r\mathbf{v}_r + c_1\mathbf{v}_{r+1} + ... + c_s\mathbf{v}_{r+s}.$$

Lineare Unabhängigkeit: Wir setzen unseren bewährten Test an, also

$$c_1\mathbf{v}_1 + ... + c_r\mathbf{v}_r + c_{r+1}\mathbf{v}_{r+1} + ... + c_{r+s}\mathbf{v}_{r+s} = \mathbf{0}.$$

Auf diese Gleichung wenden wir die lineare Abbildung Φ an und erhalten

$$c_{r+1}\mathbf{w}_1 + ... + c_{r+s}\mathbf{w}_s = \mathbf{0}.$$

Weil die Vektoren $\mathbf{w}_1, ..., \mathbf{w}_s$ linear unabhängig sind, folgt daraus $c_{r+1} = ... = c_{r+s} = 0$. Unsere ursprüngliche lineare Relation schrumpft deswegen auf

$$c_1\mathbf{v}_1 + ... + c_r\mathbf{v}_r + = \mathbf{0}$$

zusammen. Weil auch die Vektoren $\mathbf{v}_1, ..., \mathbf{v}_r$ linear unabhängig sind, erhalten wir daraus $c_1 = ... = c_r = 0$. □

Definition 2.6 (Sprachregelungen) *Eine*

	heißt auch
lineare Abbildung $V \to W$	*Homomorphismus*
injektive lineare Abbildung $V \to W$	*Monomorphismus*
surjektive lineare Abbildung $V \to W$	*Epimorphismus*
bijektive lineare Abbildung $V \to W$	*Isomorphismus*
lineare Abbildung $V \to V$	*Endomorphismus*
bijektive lineare Abbildung $V \to V$	*Automorphismus*

Gibt es einen Isomorphismus $V \to W$, so nennt man die Vektorräume V und W isomorph, in Zeichen $V \simeq W$.

Insbesondere ist jeder Vektorraum zu sich selbst isomorph. Nicht völlig überraschend. Aber, zumindest wenn er $\neq \{\mathbf{0}\}$ ist und eine Basis besitzt, dann hat er unendlich viele Automorphismen. Er ist also auf unendlich viele Weisen zu sich selbst isomorph (höchstgradig schizophren).

Weil die Hintereinanderschaltung von zwei Isomorphismen wieder ein Isomorphismus ist, deswegen ist die Isomorphie zwischen Vektorräumen transitiv im folgenden Sinn: Sind $U \simeq V$ und $V \simeq W$, dann ist auch $U \simeq W$. Isomorphie zwischen Vektorräumen ist also so etwas, wie eine Äquivalenzrelation. Genauer: Nicht ganz, denn die Menge aller $\mathbb{R}$-Vektorräume ist keine Menge.

Satz 2.6 *Es sei $\Phi : V \simeq W$ linear und bijektiv. Dann ist die Umkehrabbildung $\Phi^{-1} : W \to V$ auch wieder linear.*

Beweis. Es ist $\Phi^{-1}(\mathbf{w}) = \mathbf{v}$, wenn $\Phi(\mathbf{v}) = \mathbf{w}$ ist. Seien $\mathbf{w}_1 = \Phi(\mathbf{v}_1)$ und $\mathbf{w}_2 = \Phi(\mathbf{v}_2) \in W$. Aus der Linearität von Φ folgt $\Phi(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = c_1\mathbf{w}_1 + c_2\mathbf{w}_2$ und deswegen

$$\Phi^{-1}(c_1\mathbf{w}_1 + c_2\mathbf{w}_2) = c_1\Phi^{-1}(\mathbf{w}_1) + c_2\Phi^{-1}(\mathbf{w}_2). \qquad \square$$

Aufgabe 2.2 *Zeigen Sie, dass die folgende Abbildung $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ orthogonal ist:*

$$\Phi : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} x_1 cos(\varphi) + x_2 sin(\varphi) \\ x_1 sin(\varphi) - x_2 cos(\varphi) \end{pmatrix}$$

Aufgabe 2.3 *$\Phi : \mathbb{R}^n \to \mathbb{R}^p$ sei eine lineare Abbildung. Zeigen Sie:*
a) Φ ist genau dann injektiv, wenn gilt: Sind Vektoren $\mathbf{v}_1, ..., \mathbf{v}_r \in \mathbb{R}^n$ linear unabhängig, so sind auch die Bildvektoren $\Phi(\mathbf{v}_1), ..., \Phi(\mathbf{v}_r) \in \mathbb{R}^p$ linear unabhängig.
b) Φ ist genau dann surjektiv, wenn gilt: Spannen die Vektoren $\mathbf{v}_1, ..., \mathbf{v}_r$ den Raum $\mathbb{R}^n$ auf, so spannen ihre Bilder $\Phi(\mathbf{v}_1), ..., \Phi(\mathbf{v}_r)$ den Raum $\mathbb{R}^p$ auf.

Aufgabe 2.4 *Es seien $U, W \subset \mathbb{R}^n$ Untervektorräume und $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ eine orthogonale Abbildung mit $\Phi(U) = W$. Beweisen Sie, dass Φ das orthogonale Komplement von U auf das orthogonale Komplement von W abbildet.*

Aufgabe 2.5 *Es seien $\mathbf{a}$ und $\mathbf{b} \in \mathbb{R}^2$ zwei Einheitsvektoren und $S_\mathbf{a}$, bzw. $S_\mathbf{b}$ die Spiegelung an der Geraden senkrecht zu $\mathbf{a}$, bzw. $\mathbf{b}$.*
a) Leiten Sie Formeln her für $S_\mathbf{a} \circ S_\mathbf{b}$ und $S_\mathbf{b} \circ S_\mathbf{a}$.
b) Zeigen Sie: Es ist $S_\mathbf{a} \circ S_\mathbf{b} = S_\mathbf{b} \circ S_\mathbf{a}$ genau dann, wenn

$$\mathbf{a} = \pm\mathbf{b} \quad \textit{oder} \quad (\mathbf{a}.\mathbf{b}) = 0.$$

Aufgabe 2.6 *Es sei a, b, c ein pythagoräisches Zahlentripel. D.h., es seien $a, b, c \in \mathbb{N}$ mit*

$$a^2 + b^2 = c^2 > 0.$$

Zeigen Sie: Die Matrix

$$\frac{1}{c}\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

ist orthogonal.

2.2 Lineare Abbildungen und Matrizen

Wenn nichts anderes festgelegt wird, seien im folgenden V und W immer $\mathbb{R}$-Vektorräume.

Bei der Beschreibung der Rotation im letzten Abschnitt haben wir von folgendem Prinzip Gebrauch gemacht: Ist V endlich-dimensional mit einer Basis $\mathbf{v}_1, ..., \mathbf{v}_n \in V$, $\Phi : V \to W$ eine lineare Abbildung, welche die Basisvektoren auf die Vektoren $\mathbf{w}_1 = \Phi(\mathbf{v}_1), ..., \mathbf{w}_n = \Phi(\mathbf{v}_n) \in W$ abbildet, so ist das Bild eines jeden Vektors $\mathbf{x} = (x_1, ..., x_n) = \sum_1^n x_\nu \mathbf{v}_\nu \in V$ bereits festgelegt durch

$$\Phi(\mathbf{x}) = \Phi\left(\sum_1^n x_\nu \mathbf{v}_\nu\right) = \sum_1^n x_\nu \Phi(\mathbf{v}_\nu) = \sum_1^n x_\nu \mathbf{w}_\nu.$$

Umgekehrt kann man Vektoren $\mathbf{w}_1, ..., \mathbf{w}_n \in W$ beliebig vorgeben, durch diese Formel wird dann eine lineare Abbildung $\Phi : V \to W$ definiert mit $\Phi(\mathbf{v}_1) = \mathbf{w}_1, ..., \Phi(\mathbf{v}_n) = \mathbf{w}_n$. Daraus folgt

Satz 2.7 (Prinzip der linearen Ausdehnung) *Es sei $\{\mathbf{v}_i\}_{i\in I}$ eine Basis von V mit der Indexmenge I. Weiter seien $\mathbf{w}_i$, $i \in I$, beliebige Vektoren in W. Dann gibt es eine lineare Abbildung $\Phi : V \to W$ mit $\Phi(\mathbf{v}_i) = \mathbf{w}_i$ für alle $i \in I$. Diese lineare Abbildung ist durch die Vektoren $\mathbf{w}_i$, $i \in I$, eindeutig festgelegt.*

Beweis. Wir müssen uns nur noch über den Fall Gedanken machen, wo I unendlich ist. Aber weil die Vektoren $\mathbf{v}_i \in V$ eine Basis bilden, ist jeder Vektor $\mathbf{v} \in V$ eine *endliche* Linearkombination

$$\mathbf{v} = c_{i_1}\mathbf{v}_{i_1} + ... + c_{i_k}\mathbf{v}_{i_k}, \quad i_1, ..., i_k \in I.$$

Dabei sind die Koeffizienten $c_{i_1}, ..., c_{i_k}$ durch $\mathbf{v}$ eindeutig festgelegt. Dann ist auch die *endliche* Linearkombination

$$\Phi(\mathbf{v}) := c_{i_1}\Phi(\mathbf{v}i_1) + ... + c_{i_k}\Phi(\mathbf{v}_{i_k}) \in W$$

durch $\mathbf{v}$ eindeutig festgelegt. Die Zuordnung $\mathbf{v} \mapsto \Phi(\mathbf{v})$ definiert also eine Abbildung $\Phi : V \to W$ mit $\Phi(\mathbf{v}_i) = \mathbf{w}_i$ für alle $i \in I$.

Wegen

$$c \cdot \mathbf{v} = cc_{i_1}\mathbf{v}_{i_1} + ... + cc_{i_k}\mathbf{v}_{i_k}$$

und

$$(c_{i_1}\mathbf{v}_{i_1} + ... + c_{i_k}\mathbf{v}_{i_k}) + \left(c'_{i_1}\mathbf{v}_{i_1} + ... + c'_{i_k}\mathbf{v}_{i_k}\right) = (c_{i_1} + c'_{i_1})\mathbf{v}_{i_1} + ... + (c_{i_k} + c'_{i_k})\mathbf{v}_{i_k}$$

ist die Linearität von Φ offensichtlich. Umgekehrt stimmt jede lineare Abbildung $\Psi : V \to W$ mit $\Psi(\mathbf{v}_i) = \mathbf{w}_i$ für alle $i \in I$ mit der Abbildung Φ überein. □

Eine erste Anwendung des Prinzips der linearen Ausdehnung auf endlich-dimensionale Vektorräume liefert

Satz 2.8 *a) Es sei V endlich-dimensional mit einer Basis $\mathbf{v}_1, ..., \mathbf{v}_n$. Dann wird durch*

$$\mathbf{e}_i \mapsto \mathbf{v}_i, \quad i = 1, ..., n$$

ein Isomorphismus $\mathbb{R}^n \to V$ definiert. (Dabei ist $\mathbf{e}_1, ..., \mathbf{e}_n \in \mathbb{R}^n$ die kanonische Basis des $\mathbb{R}^n$.)

b) Zwei endlich-dimensionale Vektorräume sind genau dann isomorph, wenn sie die gleiche Dimension haben.

Beweis. a) ist eine offensichtliche Konsequenz von Satz 2.7.

b) Wenn zwei Vektorräume die gleiche Dimension n haben, dann sind beide zu $\mathbb{R}^n$ isomorph und damit auch zueinander. Umgekehrt führt jeder Isomorphismus $V \to W$ eine Basis von V in eine Basis von W über. Isomorphe Vektorräume haben also die gleich Dimension. □

Isomorphe Vektorräume sind mit den Mitteln der linearen Algebra nicht zu unterscheiden. Deswegen ist der Vektorraum V in Satz 2.8 a) trotzdem nicht dasselbe, wie der Vektorraum $\mathbb{R}^n$. Man kann die beiden Vektorräume erst dann identifizieren, wenn die Basis für V gewählt ist. Ein ziemlich subtiler Punkt!

Im Rest dieses Abschnitts betrachten wir nur noch die Vektorräume $V = \mathbb{R}^n$ und $W = \mathbb{R}^m$. Aber nach Wahl von Basen kann man alle Aussagen auf beliebige endlich-dimensionale Vektorräume übertragen.

Satz 2.9 *Für eine lineare Abbildung $\Phi : V \to W$ zwischen $\mathbb{R}$-Vektorräumen gleicher endlicher Dimension $dim(V) = dim(W) = n$ sind äquivalent: Die Abbildung Φ ist*

a) ein Isomorphismus, *b) injektiv,* *c) surjektiv.*

Beweis. Die Aussagen a) $\Rightarrow$ b) und a) $\Rightarrow$ c) sind nach Definition 2.1 klar. Für die Umkehrrichtungen wird auf Aufgabe 2.3 verwiesen. □

Definition 2.7 *Sei $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ linear. Wir schreiben die Vektoren $\Phi(\mathbf{e}_\nu) = \mathbf{v}_\nu$ als Spaltenvektoren*

$$\mathbf{v}_\nu = \begin{pmatrix} v_{1,\nu} \\ \cdot \\ \cdot \\ \cdot \\ v_{m,\nu} \end{pmatrix}$$

und fügen sie zu einer Matrix

$$(v_{\mu,\nu})_{\binom{\mu=1,\ldots,m}{\nu=1,\ldots,n}} = m \left\{ \underbrace{\begin{pmatrix} v_{1,1} & \cdot & \cdot & \cdot & v_{1,n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ v_{m,1} & \cdot & \cdot & \cdot & v_{m,n} \end{pmatrix}}_{n} \right.$$

mit m Zeilen und n Spalten zusammen. Diese Matrix heißt die darstellende Matrix *der linearen Abbildung Φ.*

Aus dieser darstellenden Matrix $(v_{\mu,\nu})$ erhält man das Bild $\Phi(\mathbf{x})$ eines Vektors $\mathbf{x} = (x_\nu)$ durch

$$\Phi(\mathbf{x}) = (\sum_{\nu=1}^{n} v_{\mu,\nu} x_\nu)_{\mu=1,\ldots,m}.$$

Dies ist genau dieselbe Formel wie für ein lineares Gleichungssystem. Natürlich ist das kein Zufall, denn auch ein lineares Gleichungssystem kann man als lineare Abbildung auffassen: Ist $(a_{\mu,\nu})$ die Koeffizientenmatrix des Systems, so ist die Abbildung

$$\Phi : \mathbb{R}^n \to \mathbb{R}^m, \qquad (x_\nu) \mapsto \sum_1^n a_{\mu,\nu} x_\nu,$$

linear und das Lösen der linearen Gleichung

$$\sum_{1}^{n} a_{\mu,\nu} x_\nu = b_\mu$$

besteht darin, für den Vektor $\mathbf{b} \in \mathbb{R}^m$ alle Vektoren $\mathbf{x} \in \mathbb{R}^n$ zu finden, welche unter Φ auf $\mathbf{b}$ abgebildet werden.

Definition 2.8 *Der Kern oder Nullraum der $m \times n$-Matrix A ist die Menge*

$$\{\mathbf{x} \in \mathbb{R}^n : A \cdot \mathbf{x} = \mathbf{0}\}.$$

Beispiel 2.6 *Die Identität*

$$id : \mathbb{R}^n \to \mathbb{R}^n \qquad \mathbf{x} \mapsto \mathbf{x}$$

bildet jeden Vektor auf sich selbst ab, also auch die Standardbasis auf die Standardbasis. Ihre Matrix ist die Einheitsmatrix

$$\mathbb{1}_n := \begin{pmatrix} 1 & 0 & . & . & . & 0 \\ 0 & 1 & 0 & & & . \\ . & 0 & . & . & & . \\ . & & . & . & . & . \\ . & & & . & . & 0 \\ 0 & . & . & . & 0 & 1 \end{pmatrix} = (\delta_{\mu,\nu})_{\mu,\nu=1,\ldots,n}.$$

Bei dieser Gelegenheit führen wir folgende Konvention ein: Die Zahl

$$\delta_{\mu,\nu} := \begin{cases} 0 \text{ wenn } \mu \neq \nu \\ 1 \text{ wenn } \mu = \nu \end{cases}$$

heißt Kronecker-Delta.

Beispiel 2.7 *Es sei $c \in \mathbb{R}$. Die Streckung*

$$\Phi : \begin{cases} \mathbb{R}^n \to \mathbb{R}^n \\ \mathbf{x} \mapsto c \cdot \mathbf{x} \end{cases}$$

bildet jeden Vektor $\mathbf{e}_\nu$ auf $c \cdot \mathbf{e}_\nu$ ab. Ihre Matrix ist deswegen

$$\begin{pmatrix} c & 0 & . & . & . & 0 \\ 0 & c & 0 & & & . \\ . & 0 & c & . & & . \\ . & & . & \ddots & . & . \\ . & & & . & c & 0 \\ 0 & . & . & . & 0 & c \end{pmatrix} = (c \cdot \delta_{\mu,\nu})_{\mu,\nu=1,\ldots,n}.$$

Spezialfälle sind die Identität *($c = 1$), die* Punktspiegelung *am Nullpunkt ($c = -1$) und die* Nullabbildung *($c = 0$).*

Beispiel 2.8 *Für jeden Einheitsvektor $\mathbf{a}, \parallel \mathbf{a} \parallel = 1$, haben wir gesehen, dass die Spiegelung an der Hyperebene $\mathbf{a}^\perp$ durch $\mathbf{x} \mapsto \mathbf{x} - 2(\mathbf{x}.\mathbf{a})\mathbf{a}$ gegeben wird. Dabei wird der Vektor $\mathbf{e}_\nu$ auf*

$$\mathbf{e}_\nu - 2(\mathbf{e}_\nu.\mathbf{a})\mathbf{a} = \mathbf{e}_\nu - 2a_\nu \mathbf{a} = (\delta_{\mu,\nu} - 2a_\nu a_\mu)_{\mu=1,\ldots,n}$$

abgebildet. Die zugehörige Matrix ist also $(\delta_{\mu,\nu} - 2a_\mu a_\nu)$.

Beispiel 2.9 *Die Matrix zu einer Rotation in der Ebene um den Winkel φ ist eine* Drehmatrix

$$\begin{pmatrix} cos(\varphi) & -sin(\varphi) \\ sin(\varphi) & cos(\varphi) \end{pmatrix}.$$

Dies erhält man aus der Form für die Bilder der Einheitsvektoren $\mathbf{e}_1$ und $\mathbf{e}_2$, die wir in Beispiel 2.5 angegeben haben.

Beispiel 2.10 *Auch eine reine Vertauschung von Basisvektoren definiert eine lineare Abbildung. So gehört z.B. zu der Vertauschung $\mathbf{e}_1 \leftrightarrow \mathbf{e}_2$ die Matrix*

$$\begin{pmatrix} 0 & 1 & 0 & . & . & 0 \\ 1 & 0 & 0 & & & 0 \\ 0 & 0 & 1 & . & & . \\ . & & . & . & . & . \\ . & & & . & 1 & 0 \\ 0 & . & . & . & 0 & 1 \end{pmatrix}.$$

Beispiel 2.11 *Es sei $U \subset \mathbb{R}^n$ ein m-dimensionaler Unterraum, der von einer ONB $\mathbf{v}_1, ..., \mathbf{v}_m$ aufgespannt wird. Die Orthogonalprojektion P_U auf diesen Unterraum wird nach Satz 1.35 gegeben durch*

$$P_U(\mathbf{x}) = \sum_{\mu=1}^{m} (\mathbf{x}.\mathbf{v}_\mu) \cdot \mathbf{v}_\mu.$$

Sie bildet $\mathbf{e}_\nu$ auf $\sum_{\mu=1}^{m} v_{\mu,\nu}\mathbf{v}_\mu$ ab und ihre Matrix ist

$$\left(\sum_{\mu=1}^{m} v_{\mu,k} v_{\mu,l} \right)_{k,l=1,...,n}.$$

Beispiel 2.12 *Sind $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ und $\Psi : \mathbb{R}^m \to \mathbb{R}^l$ linear, so ist auch $\Psi \circ \Phi : \mathbb{R}^n \to \mathbb{R}^l$ wieder linear, denn*

$$(\Psi \circ \Phi)(c_1\mathbf{v}_1 + c_2\mathbf{v}_2) = \Psi(\Phi(c_1\mathbf{v}_1 + c_2\mathbf{v}_2)) = \Psi(c_1\Phi(\mathbf{v}_1) + c_2\Phi(\mathbf{v}_2)) = c_1\Psi\Phi(\mathbf{v}_1) + c_2\Psi\Phi(\mathbf{v}_2).$$

Was ist die zu $\Psi \circ \Phi$ gehörige Matrix? Dazu sei

$\mathbf{e}_1, ..., \mathbf{e}_n$ Standardbasis des $\mathbb{R}^n$, $\Phi(\mathbf{e}_\nu) = \sum_{\mu=1}^m a_{\mu,\nu}\mathbf{f}_\mu$, $(a_{\mu,\nu})$ die Matrix für Φ,
$\mathbf{f}_1, ..., \mathbf{f}_m$ Standardbasis des $\mathbb{R}^m$, $\Psi(\mathbf{f}_\mu) = \sum_{\lambda=1}^l b_{\lambda,\mu}\mathbf{g}_\lambda$, $(b_{\lambda,\mu})$ die Matrix für Ψ,
$\mathbf{g}_1, ..., \mathbf{g}_l$ Standardbasis des $\mathbb{R}^l$.

Wir berechnen die Bilder der Basisvektoren $\mathbf{e}_\nu$:

$$(\Psi \circ \Phi)(\mathbf{e}_\nu) = \Psi(\sum_{\mu=1}^{m} a_{\mu,\nu}\mathbf{f}_\mu) = \sum_{\lambda=1}^{l} b_{\lambda,\mu}a_{\mu,\nu}\mathbf{g}_\lambda = \sum_{\lambda=1}^{l} c_{\lambda,\nu}\mathbf{g}_\lambda$$

mit der Matrix

$$\boxed{(c_{\lambda,\nu})_{\lambda,\nu} = (\sum_{\mu=1}^{m} b_{\lambda,\mu} a_{\mu,\nu})_{\lambda,\nu}.}$$

Schließlich formulieren wir noch

Satz 2.10 *Es sei $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ linear. Dann ist der Rang der darstellenden Matrix von Φ die Dimension des Bildraums $Bild(\Phi)$.*

Beweis. Die Spalten der darstellenden Matrix spannen im $\mathbb{R}^m$ den Bildraum $Bild(\Phi)$ auf. $\square$

Aufgabe 2.7 *Geben Sie die $n \times n$-Matrizen an, die zu den folgenden linearen Abbildungen $\Phi_i : \mathbb{R}^n \to \mathbb{R}^n$ $(i = 1, ..., 4)$ gehören:*
a) Φ_1 ist die Spiegelung an der Hyperebene $\mathbf{a}^\perp$, wo $\mathbf{a} = (a_1, ..., a_n) \in \mathbb{R}^n$ mit $\parallel \mathbf{a} \parallel = 1$ gegeben ist.
b) Φ_2 ist die Orthogonalprojektion auf einen p-dimensionalen linearen Unterraum U, der durch eine Orthonormalbasis $\mathbf{u}_1 = (u_{1,1}, ..., u_{1,n}), ..., \mathbf{u}_p = (u_{p,1}, ..., u_{p,n})$ aufgespannt wird.
c) $\Phi_3 = \Phi_1 \circ \Phi_1$.
d) $\Phi_4 = \Phi_2 \circ \Phi_2$.

Aufgabe 2.8 *Die lineare Abbildung $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ werde definiert durch*

$$\Phi(\mathbf{e}_\nu) = \mathbf{e}_\nu - \mathbf{e}_{\nu+1},\ \nu = 1, ..., n-1, \qquad \textit{und} \qquad \Phi(\mathbf{e}_n) = \mathbf{e}_n - \mathbf{e}_1.$$

a) Geben Sie die Matrix für Φ an.
b) Bestimmen Sie eine Basis für den linearen Unterraum $\Phi(\mathbb{R}^n) \subset \mathbb{R}^n$.

Aufgabe 2.9 *Es sei $\Phi : \mathbb{R}^3 \to \mathbb{R}^3$ die lineare Abbildung mit der Matrix*

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Bestimmen Sie die Matrizen für die linearen Abbildungen Φ^2 und Φ^3.

Aufgabe 2.10 *Bestimmen Sie die Matrizen für die Orthogonalprojektionen des $\mathbb{R}^3$ auf die*
a) vom Vektor $\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3$ aufgespannte Gerade;
b) von den Vektoren $\mathbf{e}_1 + \mathbf{e}_2$ und $\mathbf{e}_2 + \mathbf{e}_3$ aufgespannte Ebene.

Aufgabe 2.11 *Es sei $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ die lineare Abbildung, die durch die Matrix*

$$\begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$$

beschrieben wird. Zeigen Sie:
a) $\Phi \circ \Phi = id$.
b) Jeder Vektor $\mathbf{x} \in \mathbb{R}^2$ lässt sich eindeutig darstellen als $\mathbf{x} = \mathbf{u} + \mathbf{v}$ mit $\Phi(\mathbf{u}) = \mathbf{u}$ und $\Phi(\mathbf{v}) = -\mathbf{v}$.

Aufgabe 2.12 *Im* $\mathbb{R}^2$ *seien die vier Punkte*

$$\mathbf{v}_1 = (1,1), \quad \mathbf{v}_2 = (1,-1), \quad \mathbf{v}_3 = (-1,-1), \quad \mathbf{v}_4 = (-1,1)$$

gegeben. Man bestimme zwei lineare Abbildungen $\varphi, \psi : \mathbb{R}^2 \to \mathbb{R}^2$, *welche die* $\mathbf{v}_i$ *permutieren und außerdem die Bedingungen*

$$\varphi^2, \psi \neq id, \quad \varphi\psi = \psi\varphi^{-1}$$

erfüllen.

Aufgabe 2.13 (Katzenauge) *Die Spiegelung*

$$S_z : (x,y,z) \mapsto (x,y,-z)$$

an der x,y*-Ebene kann man interpretieren als die Änderung des Richtungsvektors eines Lichtstrahls bei Reflexion an dieser Ebene. Wie ändert sich der Richtungsvektor eines Lichtstrahls, der nacheinander an der* x,y*-Ebene, der* y,z*-Ebene und der* x,z*-Ebene gespiegelt wird?*

2.3 Matrizenrechnung

2.3.1 Matrizenprodukt

In 1.2.2, Beispiel 1.9 haben wir den $\mathbb{R}$-Vektorraum $M(m \times n)$ der reellen $m \times n$-Matrizen eingeführt. In Abschnitt 2.2 haben wir jeder linearen Abbildung $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ eine darstellende $m \times n$-Matrix zugeordnet.

Definition 2.9 *Die Menge aller linearen Abbildungen* $\Phi : V \to W$ *zwischen zwei* $\mathbb{R}$*-Vektorräumen* V *und* W *bezeichnen wir mit* $Hom(V,W)$. *Die Menge* $Hom(V,W)$ *versehen wir mittels der folgenden Rechenoperationen mit der Struktur eines* $\mathbb{R}$*-Vektorraums:*

Für $\Phi \in Hom(V,W)$ *und* $c \in \mathbb{R}$ *definieren wir* $c \cdot \Phi \in Hom(V,W)$ *durch*

$$(c \cdot \Phi)(\mathbf{v}) := c \cdot \Phi(\mathbf{v}).$$

Für $\Phi, \Psi \in Hom(V,W)$ *definieren wir* $\Phi + \Psi \in Hom(V,W)$ *durch*

$$(\Phi + \Psi)(\mathbf{v}) := \Phi(\mathbf{v}) + \Psi(\mathbf{v}).$$

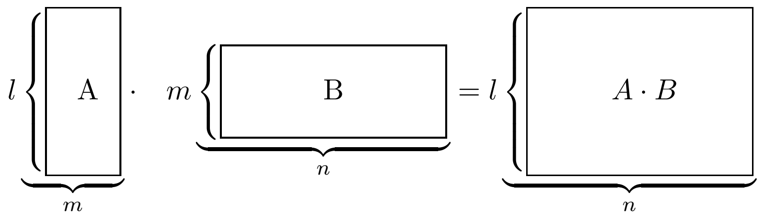
Es ist klar, dass die so definierten Abbildungen linear sind, also zu $Hom(V,W)$ gehören. Ebenso klar ist für $V = \mathbb{R}^n$ und $W = \mathbb{R}^m$: Ist A die darstellende Matrix von Φ, so ist $c \cdot A$ die darstellende Matrix von $c \cdot \Phi$. Ist weiter B die darstellende Matrix von Ψ, so ist $A + B$ die darstellende Matrix von $\Phi + \Psi$. Etwas akademischer kann man das so ausdrücken:

Satz 2.11 *Ordnet man jeder Abbildung* $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ *ihre darstellende Matrix zu, so erhält man einen Vektorraum-Isomorphismus*

$$Hom(\mathbb{R}^n, \mathbb{R}^m) \to M(m \times n).$$

Das ist nicht viel Neues. Neu ist aber die nach Beispiel 2.12 definierte Matrizenmultiplikation:

$$\begin{array}{rcl} A & = & (a_{\lambda,\mu}) \in M(l \times m) \\ B & = & (b_{\mu,\nu}) \in M(m \times n) \\ \hline A \cdot B & = & (\sum_{\mu=1}^m a_{\lambda,\mu} b_{\mu,\nu}) \in M(l \times n) \end{array}$$



Das Matrizenprodukt $A \cdot B$ ist nur definiert, wenn die Matrix A genau so viele Spalten wie B Zeilen hat! Und Dann ist dieses Produkt

$$A \cdot B := C = (c_{\lambda,\nu}) \quad \text{mit} \quad c_{\lambda,\nu} = \sum_{\mu=1}^m a_{\lambda,\mu} \cdot b_{\mu,\nu}.$$

Bevor wir Beispiele behandeln, wollen wir diese fundamentale Rechenoperation noch etwas analysieren.

Wir bezeichnen die Zeilenvektoren von A mit $\mathbf{a}_\lambda^t$ und die Spaltenvektoren von B mit $\mathbf{b}^\nu$, haben also

$$A = \begin{pmatrix} \mathbf{a}_1^t \\ \vdots \\ \mathbf{a}_l^t \end{pmatrix}, \quad B = (\mathbf{b}^1, ..., \mathbf{b}^n).$$

Dann ist $c_{\lambda,\nu}$ gerade das Skalarprodukt

$$(\mathbf{a}_\lambda . \mathbf{b}^\nu) = \sum_{\mu=1}^m a_{\lambda,\mu} b_{\mu,\nu}.$$

Das ist die (in Bezug auf A) zeilenweise Sicht

$$A \cdot B = ((\mathbf{a}_\lambda . \mathbf{b}^\nu))$$

der Matrizenmultiplikation. Sie gehört zur 'Zeile·Spalte'-Berechnungsregel: Man erhält den Eintrag $c_{\lambda,\nu}$ der Produktmatrix, indem man den Zeilenvektor $\mathbf{a}_\lambda$ von A und den Spaltenvektor $\mathbf{b}^\nu$ von B 'aufeinanderlegt, komponentenweise multipliziert und addiert.'

Die (in Bezug auf A) spaltenweise Sicht der Matrizenmultiplikation erhalten wir, wenn wir die Spaltenvektoren von A mit $\mathbf{a}^1, ..., \mathbf{a}^m$ bezeichnen, also

$$A = (\mathbf{a}^1, ..., \mathbf{a}^m), \quad B = (b_{\mu,\nu})$$

haben. Dann ist

$$A \cdot B = \left(\sum_{\mu=1}^m \mathbf{a}^\mu \cdot b_{\mu,1}, ..., \sum_{\mu=1}^m \mathbf{a}^\mu \cdot b_{\mu,n} \right)$$

eine Matrix, deren Spalten gerade Linearkombinationen der m Spalten von A sind.

Beispiel 2.13 *Das Produkt einer $m \times n$-Matrix A mit einer $n \times 1$-Matrix $\mathbf{a} \in \mathbb{R}^n$ (d.h., mit einem Vektor $\mathbf{a} \in \mathbb{R}^n$) ist das Matrix-Vektorprodukt aus Definition 1.12. Insbesondere kann man das Skalarprodukt $(\mathbf{a}.\mathbf{b})$ im $\mathbb{R}^n$ auffassen als das Matrizen-Produkt der $(1 \times n)$-Matrix $\mathbf{a}^t$ mit der $n \times 1$-Matrix $\mathbf{b}$. Das Resultat ist eine 1×1-Matrix, also ein Skalar.*

Beispiel 2.14 *Ist* $\mathbb{1}_m$ *die* $m \times m$*-Einheitsmatrix und* $A \in M(m \times n)$*, so ist*

$$\mathbb{1}_m \cdot A = (\delta_{\lambda,\mu}) \cdot (a_{\mu,\nu}) = (\sum_{\mu=1}^{m} \delta_{\lambda,\mu} \cdot a_{\mu,\nu}) = (a_{\lambda,\nu}) = A.$$

Ebenso ist

$$A \cdot \mathbb{1}_n = A.$$

Beispiel 2.15 *Sind* $A(\alpha)$ *und* $A(\beta)$ *die Drehmatrizen*

$$\begin{pmatrix} cos(\alpha) & -sin(\alpha) \\ sin(\alpha) & cos(\alpha) \end{pmatrix}, \qquad \begin{pmatrix} cos(\beta) & -sin(\beta) \\ sin(\beta) & cos(\beta) \end{pmatrix},$$

so ist das Produkt

$$\begin{aligned} A(\alpha) \cdot A(\beta) &= \begin{pmatrix} cos(\alpha)cos(\beta) - sin(\alpha)sin(\beta) & -cos(\alpha)sin(\beta) - sin(\alpha)cos(\beta) \\ sin(\alpha)cos(\beta) + cos(\alpha)sin(\beta) & -sin(\alpha)sin(\beta) + cos(\alpha)cos(\beta) \end{pmatrix} \\ &= \begin{pmatrix} cos(\alpha+\beta) & -sin(\alpha+\beta) \\ sin(\alpha+\beta) & cos(\alpha+\beta) \end{pmatrix} \end{aligned}$$

die Drehmatrix zum Winkel $\alpha+\beta$*. Dieses Ergebnis ist eine direkte Konsequenz der Additionstheoreme für die Winkelfunktionen.*

Beispiel 2.16 *Sei* E *die Vertauschungsmatrix zum ersten und zweiten Basisvektor:*

$$E = \begin{pmatrix} 0 & 1 & 0 & . & . & . & 0 \\ 1 & 0 & 0 & . & & & . \\ 0 & 0 & 1 & . & & & . \\ . & . & . & 1 & . & & . \\ . & & & . & . & . & . \\ . & & & & . & . & 0 \\ 0 & . & . & . & . & 0 & 1 \end{pmatrix} \in M(m \times m).$$

Dann ist für jede $m \times n$*-Matrix* A

$$E \cdot A = \begin{pmatrix} 0 & 1 & 0 & . & . & . & 0 \\ 1 & 0 & 0 & . & & & . \\ 0 & 0 & 1 & . & & & . \\ . & . & . & 1 & . & & . \\ . & & & . & . & . & . \\ . & & & & . & . & 0 \\ 0 & . & . & . & . & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_{1,1} & a_{1,2} & . & . & . & a_{1,n} \\ a_{2,1} & a_{2,2} & . & . & . & a_{2,n} \\ a_{3,1} & a_{3,2} & . & . & . & a_{3,n} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ a_{m,1} & a_{m,2} & . & . & . & a_{m,n} \end{pmatrix} = \begin{pmatrix} a_{2,1} & a_{2,2} & . & . & . & a_{2,n} \\ a_{1,1} & a_{1,2} & . & . & . & a_{1,n} \\ a_{3,1} & a_{3,2} & . & . & . & a_{3,n} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ a_{m,1} & a_{m,2} & . & . & . & a_{m,n} \end{pmatrix}.$$

Das heißt: Die Linksmultiplikation von E *an* A *vertauscht in* A *die erste mit der zweiten Zeile. Allgemeiner können alle drei Arten elementarer Zeilenumformungen durch Linksmultiplikation mit den folgenden Matrizen (*Elementarmatrizen*) erreicht werden:*

Vertauschen zweier Zeilen (Elementarmatrix vom Typ I):

$$\begin{pmatrix}
1 &&&&&&&&&& \\
& \ddots &&&&&&&&& \\
&& 1 &&&&&&&& \\
&&& 0 & \cdots & \cdots & \cdots & 1 &&& \\
&&& \vdots & 1 &&& \vdots &&& \\
&&& \vdots && \ddots && \vdots &&& \\
&&& \vdots &&& 1 & \vdots &&& \\
&&& 1 & \cdots & \cdots & \cdots & 0 &&& \\
&&&&&&&& 1 && \\
&&&&&&&&& \ddots & \\
&&&&&&&&&& 1
\end{pmatrix}$$

Multiplikation einer Zeile mit $c \in \mathbb{R}$ (Elementarmatrix vom Typ II):

$$\begin{pmatrix}
1 &&&&&& \\
& \ddots &&&&& \\
&& 1 &&&& \\
&&& c &&& \\
&&&& 1 && \\
&&&&& \ddots & \\
&&&&&& 1
\end{pmatrix}$$

Addieren des c-fachen einer Zeile zu einer anderen (Elementarmatrix vom Typ III):

$$\begin{pmatrix}
1 &&&&&& \\
& \ddots &&&&& \\
&& 1 & \cdots & c && \\
&&& \ddots & \vdots && \\
&&&& 1 && \\
&&&&& \ddots & \\
&&&&&& 1
\end{pmatrix}$$

Satz 2.12 *Für die Matrizenmultiplikation gelten folgende Rechenregeln:*
a) Distributivgesetze:

$$\begin{aligned}
&(A_1 + A_2) \cdot B = A_1 \cdot B + A_2 \cdot B \\
&A \cdot (B_1 + B_2) = A \cdot B_1 + A \cdot B_2 \\
&(c \cdot A) \cdot B = c \cdot (A \cdot B) = A \cdot (c \cdot B) \quad \textit{für } c \in \mathbb{R}
\end{aligned}$$

b) Assoziativität:

$$(A \cdot B) \cdot C = A \cdot (B \cdot C).$$

Beweis. Man kann die Rechenregeln nachrechnen, indem man alle vorkommenden Matrizenprodukte elementweise ausrechnet. Wegen der Korrespondenz zwischen Hintereinanderschaltung von Abbildungen und Matrizenmultiplikation genügt es aber auf die entsprechenden Regeln für Abbildungen zu verweisen. □

Ein ganz entscheidender Unterschied zur Multiplikation reeller Zahlen besteht darin, dass die Matrizenmultiplikation *nicht kommutativ* ist.

$$\boxed{\text{Im Allgemeinen ist} \quad A \cdot B \neq B \cdot A.}$$

Wir berechnen dafür als Beispiel

$$\begin{aligned} \begin{pmatrix} a_1 & a \\ 0 & a_2 \end{pmatrix} \cdot \begin{pmatrix} b_1 & b \\ 0 & b_2 \end{pmatrix} &= \begin{pmatrix} a_1 b_1 & a_1 b + a b_2 \\ 0 & a_2 b_2 \end{pmatrix}, \\ \begin{pmatrix} b_1 & b \\ 0 & b_2 \end{pmatrix} \cdot \begin{pmatrix} a_1 & a \\ 0 & a_2 \end{pmatrix} &= \begin{pmatrix} b_1 a_1 & b_1 a + b a_2 \\ 0 & b_2 a_2 \end{pmatrix}. \end{aligned}$$

Im allgemeinen (z.B. wenn $a = b = 1$ und $a_1 + b_2 \neq a_2 + b_1$) unterscheiden sich die beiden Dreiecksmatrizen durch ihren Eintrag rechts oben.

Aufgabe 2.14 *Es sei*

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 0 \\ 3 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Berechnen Sie die Matrizen $A \cdot A$, $A \cdot B$, $B \cdot A$ und $(B - 1\!\!1_3)(B - 1\!\!1_3)(B - 1\!\!1_3)$.

Aufgabe 2.15 *a) Bestimmen Sie alle Matrizen*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M(2 \times 2)$$

mit $A \cdot A = 1\!\!1_2$ und alle Matrizen A mit $A \cdot A = 0$.
b) Bestimmen Sie alle Matrizen

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \in M(2 \times 2)$$

mit $A \cdot A \cdot A \cdot A \cdot A = 1\!\!1_2$.

Aufgabe 2.16 *Es sei $A \in M(n \times n)$. Zeigen Sie, dass die folgenden Aussagen äquivalent sind:*
i) Für jede Matrix $X \in M(n \times n)$ ist $A \cdot X = X \cdot A$.
ii) Es gibt ein $c \in \mathbb{R}$ mit $A = c \cdot 1\!\!1_n$.

Aufgabe 2.17 *Für $r \in \mathbb{N}$ sei die Matrix $A_r = (a_{i,j}) \in M(n \times n)$ definiert durch $a_{i,j} = \delta_{i,j-r}$, $i, j = 1, ..., n$. Beweisen Sie:*
a) $A_r \cdot A_s = A_{r+s}$.
b) $A_1^n = 0$.

Aufgabe 2.18 *Sind $A, B \in M(n \times n)$, so heißen*

$$\begin{array}{ll} [A, B] := A \cdot B - B \cdot A & \textit{Kommutator} \\ \{A, B\} := A \cdot B + B \cdot A & \textit{Anti-Kommutator} \end{array}$$

der Matrizen A und B. Die Matrizen $\sigma_i \in M(2 \times 2)$ seien definiert durch

$$\sigma_1 := \frac{1}{2}\sqrt{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 := \frac{1}{2}\sqrt{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_3 := \frac{1}{2}\sqrt{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Berechnen Sie die Kommutatoren $[\sigma_i, \sigma_j]$ und die Antikommutatoren $\{\sigma_i, \sigma_j\}$.

Aufgabe 2.19 *Zur Matrix*

$$A = \begin{pmatrix} 2 & 1 & -1 \\ 8 & 3 & -3 \end{pmatrix}$$

bestimme man alle 3×2 Matrizen B mit $A \cdot B = \mathbb{1}_2$.

Aufgabe 2.20 *Gegeben sei die Matrix*

$$B = \begin{pmatrix} -\dfrac{1}{2} & -\frac{1}{2}\sqrt{3} \\ \dfrac{1}{2}\sqrt{3} & -\dfrac{1}{2} \end{pmatrix}.$$

Berechnen Sie B^{2000} (d.h., die 2000-te Potenz von B).

Aufgabe 2.21 *Es sei $C \in M(m \times n)$ eine Matrix von Rang k. Man beweise: Es gibt Matrizen $A \in M(m \times k)$ und $B \in M(k \times n)$ mit $C = A \cdot B$.*

Aufgabe 2.22 *Sei E die zweireihige Einheitsmatrix und für reelle Zahlen p und q sei*

$$A = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}, \quad B = \begin{pmatrix} -p & p \\ q & -q \end{pmatrix}.$$

a) Bestimmen Sie für $k = 2, 3, ...$ reelle Zahlen c_k mit $B^k = c_k B$.

b) Zeigen Sie die binomische Formel $(E+B)^k = \sum_{i=0}^{k} \binom{k}{i} B^i$.

c) Ist $p + q \neq 0$, so zeige man für $k \in \mathbb{N}$

$$A^k = (E+B)^k = E + \frac{1}{p+q}B - \frac{(1-p-q)^k}{p+q}B.$$

Was gilt im Falle $p + q = 0$?

Aufgabe 2.23 *Das Vektorprodukt $[\mathbf{a}, \mathbf{b}] = (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1)$ für Vektoren $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ definiert bei festem $\mathbf{a}$ einen Endomorphismus (= lineare Abbildung) $\varphi_{\mathbf{a}} : \mathbb{R}^3 \to \mathbb{R}^3$, $\mathbf{b} \mapsto [\mathbf{a}, \mathbf{b}]$.*
a) Wie sieht die darstellende (3×3)-Matrix von $\varphi_{\mathbf{a}}$ aus?
b) Man bestimme alle Werte, die Rang $\varphi_{\mathbf{a}}$ bei passender Wahl von $\mathbf{a}$ annehmen kann.

Aufgabe 2.24 *Auf dem mit dem Vektorprodukt versehenen $\mathbb{R}^3$ betrachte man die linearen Abbildungen*

$$f(\mathbf{x}) = \mathbf{a} \times \mathbf{x} \quad \text{und} \quad g(\mathbf{x}) = \mathbf{x} \times \mathbf{a} \quad \text{mit } \mathbf{a} = \frac{1}{3}\begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}.$$

a) Man stelle fest, für welche $\mathbf{u} \in \mathbb{R}^3$ die Gleichung $f(\mathbf{x}) = \mathbf{u}$ lösbar ist.

b) Für die Vektoren

$$\mathbf{b} = \frac{1}{3}\begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} \qquad \textit{bzw.} \qquad \mathbf{b}' = \frac{1}{3}\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$$

bestimme man alle Lösungen der Gleichungen $g(\mathbf{x}) = \mathbf{b}$ *bzw* $g(\mathbf{x}) = \mathbf{b}'$.
c) Man zeige $f \cdot g = g \cdot f$ *und* $(f \cdot g)^2 = f \cdot g$.

2.3.2 Tensorprodukt von Vektoren

In Beispiel 2.8 haben wir für die Spiegelung an der Hyperbene $\mathbf{a}^\perp \subset \mathbb{R}^n$ wo $\parallel \mathbf{a} \parallel = 1$, gezeigt, dass ihre Matrix die Form

$$\mathbb{1}_n - 2\mathbf{a} \cdot \mathbf{a}^t$$

hat. Das ist eigenartig! Da wird der Vektor $\mathbf{a}$ mit sich selbst multipliziert, aber ziemlich anders, als beim Skalarprodukt. Links wird er als $n \times 1$-Matrix aufgefasst, rechts als $1 \times n$-Matrix, und das Resultat ist eine $n \times n$-Matrix.

Definition 2.10 *Es seien* $\mathbf{a} = (a_\mu)_{\mu=1,..,m} \in \mathbb{R}^m$ *und* $\mathbf{b} = (b_\nu)_{\nu=1,..,n} \in \mathbb{R}^n$ *Vektoren. Ihr* dyadisches *oder* Tensorprodukt *ist die* $m \times n$*-Matrix*

$$\mathbf{a} \otimes \mathbf{b} = \mathbf{a} \cdot \mathbf{b}^t = (a_\mu \cdot b_\nu)_{\mu,\nu}.$$

Die Matrix $\mathbf{a} \otimes \mathbf{b}$ hat also den Zeilenraum $\mathbb{R} \cdot \mathbf{b}$ und den Spaltenraum $\mathbb{R} \cdot \mathbf{a}$. Insbesondere hat sie den Rang $= 1$ (wenn man von den trivialen Fällen $\mathbf{a} = \mathbf{0}$ oder $\mathbf{b} = \mathbf{0}$ absieht). Davon gilt auch die Umkehrung:

Satz 2.13 (Rang-1-Matrizen) *Hat die* $m \times n$*-Matrix* C *den Rang* 1*, so gibt es Vektoren* $\mathbf{a} \in \mathbb{R}^m$ *und* $\mathbf{b} \in \mathbb{R}^n$ *mit*

$$A = \mathbf{a} \otimes \mathbf{b}.$$

Beweis. Weil der Spaltenrang von A den Wert $= 1$ hat, gibt es eine Spalte $\mathbf{a}$, derart, dass alle Spalten Vielfache

$$\mathbf{c}^\nu = b^\nu \cdot \mathbf{a}$$

dieser einen Spalte sind. Das bedeutet

$$A = (b^1 \cdot \mathbf{a}, ..., b^n \cdot \mathbf{a}) = \mathbf{a} \otimes \mathbf{b} \quad \text{mit} \quad \mathbf{b} = (b^1, ..., b^n). \qquad \square$$

Wenn $\mathbf{a} \neq \mathbf{0}$ ist, dann ist

$$(\mathbf{a} \otimes \mathbf{b}) \cdot \mathbf{x} = \mathbf{a} \cdot (\mathbf{b}^t \cdot \mathbf{x}) = \mathbf{0}, \quad \mathbf{x} \in \mathbb{R}^n,$$

genau dann, wenn $\mathbf{x} \in \mathbf{b}^\perp$. Wir haben ausgerechnet:

$$Kern(\mathbf{a} \otimes \mathbf{b}) = \mathbf{b}^\perp \subset \mathbb{R}^n.$$

Wir wenden dieses Tensorprodukt an auf Orthogonalprojektionen. In Beispiel 2.11 haben wir für die Orthogonalprojektion P_U auf einen Unterraum $U \subset \mathbb{R}^n$ mit ONB $\mathbf{v}_1, ..., \mathbf{v}_m$ die Formel

$$P_U(\mathbf{x}) = \sum_{\mu=1}^{m} (\mathbf{x}.\mathbf{v}_\mu) \cdot \mathbf{v}_\mu$$

angegeben. Die rechte Seite können wir umschreiben

$$\sum_{\mu=1}^{m} \mathbf{v}_\mu \cdot (\mathbf{v}_\mu.\mathbf{x}) = \sum_{\mu=1}^{m} \mathbf{v}_\mu \cdot \mathbf{v}_\mu^t \cdot \mathbf{x} = \sum_{\mu=1}^{m} (\mathbf{v}_\mu \otimes \mathbf{v}_\mu) \cdot \mathbf{x}.$$

Damit haben wir die Darstellungsmatrix von P_U identifiziert als

$$\sum_{\mu=1}^{m} \mathbf{v}_\mu \otimes \mathbf{v}_\mu.$$

Insbesondere im Fall $m = 1$ wird U eine Gerade $\mathbb{R} \cdot \mathbf{v}$ mit $\parallel \mathbf{v} \parallel = 1$. Die Orthogonalprojektion auf eine derartige Gerade hat die Matrix

$$\mathbf{v} \otimes \mathbf{v}.$$

Und die Orthogonalprojektion auf einen m-dimensionalen Unterraum U kann man als Summe von m derartigen Orthogonalprojektionen auf Geraden auffassen.

In Satz 1.36 haben wir für die Orthogonalprojektionen auf U und$U^\perp$ gesehen:

$$P_{U^\perp}(\mathbf{x}) = \mathbf{x} - P_U(\mathbf{x}), \quad \text{bzw.} \quad \mathbf{x} = P_U(\mathbf{x}) + P_{U^\perp}(\mathbf{x}),$$

$$P_U + P_{U^\perp} = id.$$

Die Orthogonalprojektion auf $U^\perp \subset \mathbb{R}^n$ hat also die darstellende Matrix

$$\mathbb{1}_n - \sum_{\mu=1}^{m} \mathbf{v}_\mu \otimes \mathbf{v}_\mu.$$

Aufgabe 2.25 *Es seien*

$$\mathbf{a} \in \mathbb{R}^m, \quad \mathbf{b}, \mathbf{c} \in \mathbb{R}^n, \quad \mathbf{d} \in \mathbb{R}^d.$$

Zeigen Sie

$$(\mathbf{a} \otimes \mathbf{b}) \cdot (\mathbf{c} \otimes \mathbf{d}) = (\mathbf{b}.\mathbf{c})\, \mathbf{a} \otimes \mathbf{d}.$$

Aufgabe 2.26 *a) Es sei*

$$E_1(i,j) = (\mathbf{e}_1, ..., \mathbf{e}_{i-1}, \mathbf{e}_j, \mathbf{e}_{i+1}, ..., \mathbf{e}_{j-1}, \mathbf{e}_i, \mathbf{e}_{j+1}, ..., \mathbf{e}_n)$$

die $n \times n$-Elementarmatrix zur Vertauschung von i-ter und j-ter Zeile. Zeigen Sie:

$$E_1(i,j) = \mathbb{1}_n - (\mathbf{e}_i \otimes \mathbf{e}_i + \mathbf{e}_j \otimes \mathbf{e}_j) + \mathbf{e}_i \otimes \mathbf{e}_j + \mathbf{e}_j \otimes \mathbf{e}_i.$$

b) Es sei

$$E_2(i,c) = (\mathbf{e}_1, ..., \mathbf{e}_{i-1}, c \cdot \mathbf{e}_i, \mathbf{e}_{i+1}, ..., \mathbf{e}_n$$

die $n \times n$-Elementarmatrix, welche die Multiplikation der i-ten Zeile mit dem Faktor c beschreibt. Zeigen Sie:

$$E_2(i,c) = \mathbb{1}_n + (c-1)\mathbf{e}_i \otimes \mathbf{e}_i.$$

c) Es sei

$$E_3(i,j,c) = (\mathbf{e}_1, ..., \mathbf{e}_{i-1}, \mathbf{e}_i + c \cdot \mathbf{e}_j, \mathbf{e}_{i+1}, ..., \mathbf{e}_n)$$

die $n \times n$-Elementarmatrix, welche die Addition des c-fachen der i-ten Zeile zur j-ten Zeile beschreibt. Zeigen Sie

$$E_3(i,j,c) = \mathbb{1}_n + c \cdot \mathbf{e}_j \otimes \mathbf{e}_i.$$

Aufgabe 2.27 *Zeigen Sie mit der Notation der vorhergehenden Aufgabe und mit den Rechenregeln des Tensorprodukts*

$$E_1(i,j)\cdot E_1(i,j) = \mathbb{1}_n, \quad E_2(i,c)\cdot E_2(i,\frac{1}{c}) = \mathbb{1}_n, \quad E_3(i,j,c)\cdot E_3(i,j,-c) = \mathbb{1}_n.$$

Aufgabe 2.28 *Zeigen Sie: Die $m\times n$-Matrix A hat einen Rang $\leq r$ genau dann, wenn*

$$A = \mathbf{a}_1\otimes\mathbf{b}_1 + ... + \mathbf{a}_r\otimes\mathbf{b}_r \quad \textit{mit} \quad \mathbf{a}_1,...,\mathbf{a}_r \in \mathbb{R}^m,\ \mathbf{b}_1,...,\mathbf{b}_r\in\mathbb{R}^n.$$

2.3.3 Projektionen

Definition 2.11 *Es seien $U\subset V$ reelle Vektorräume und $P: V\to V$ linear. Die Abbildung P heißt* Projektion *auf U, wenn*

i) $Bild(P) = U$, *ii) $P(\mathbf{u}) = \mathbf{u}$ für alle $\mathbf{u}\in U$.*

Satz 2.14 *a) Eine lineare Abbildung $P: V\to V$ ist genau dann eine Projektion, wenn $P\circ P = P$.*
b) Mit P ist auch $id_V - P$ eine Projektion.
c) Die Abbildung $id_V - P$ ist die Projektion auf $Kern(P)$.

Beweis. a) Die Eigenschaft $P\circ P = P$ einer Projektion folgt unmittelbar aus der Definition. Sei also diese Eigenschaft vorausgesetzt. Wir definieren den Unterraum $U\subset V$ als $Bild(P)$. Damit ist i) erfüllt. Jedes $\mathbf{u}\in U$ schreibt sich als $\mathbf{u} = P(\mathbf{v})$ mit $\mathbf{v}\in V$ Daraus folgt

$$P(\mathbf{u}) = P(P(\mathbf{v})) = P(\mathbf{v}) = \mathbf{u}.$$

Also ist auch ii) erfüllt.
b) Es ist

$$(id_V - P)\circ(id_V - P) = id_V - P - P + P\circ P = id_V - P.$$

c) Es sei $\mathbf{v} = \mathbf{x} - P(\mathbf{x})$ ein Vektor im Bild der Abbildung $id_V - P$. Dann folgt

$$P(\mathbf{v}) = P(\mathbf{x}) - P(P(\mathbf{x})) = \mathbf{0}.$$

Also gilt $Bild(id_V - P)\subset Kern(P)$.
Sei umgekehrt $\mathbf{v}\in Kern(P)$. Dann ist also $P(\mathbf{v}) = \mathbf{0}$ und

$$\mathbf{v} = \mathbf{v} - P(\mathbf{v}) = (id_V - P)(\mathbf{v}) \in Bild(id_V - P).$$

□

Beispiel 2.17 *Der Vektorraum V sei mit einem SKP versehen. Eine Projektion $P: V\to U\subset V$ ist genau dann die Orthogonalprojektion, wenn $Kern(P)\subset U^\perp$. Denn dann ist für alle $\mathbf{v}\in V$*

$$P(\mathbf{v}) - \mathbf{v} \in Kern(P)\subset U^\perp.$$

In diesem Fall gilt sogar $Kern(P) = U^\perp$. Denn für $\mathbf{v}\in U^\perp$ ist $(\mathbf{v}-\mathbf{0})\perp U$ und deswegen $P(\mathbf{v}) = \mathbf{0}$.

Beispiel 2.18 *Wir betrachten* $\mathbb{R}^n$ *mit dem euklidischen SKP* $(\quad . \quad)$. *Weiter sei* $\mathbf{0} \neq \mathbf{v} \in \mathbb{R}^n$. *Für jeden Vektor* $\mathbf{w} \in \mathbb{R}^n$ *mit* $(\mathbf{v}.\mathbf{w}) = 1$ *gilt*

$$(\mathbf{v} \otimes \mathbf{w}) \cdot (\mathbf{v} \otimes \mathbf{w}) = \mathbf{v} \cdot \mathbf{w}^t \cdot \mathbf{v} \cdot \mathbf{w}^t = \mathbf{v} \cdot (\mathbf{w}.\mathbf{v}) \cdot \mathbf{w}^t = \mathbf{v} \otimes \mathbf{w}.$$

Deswegen beschreibt die Matrix $\mathbf{v} \otimes \mathbf{w}$ *eine Projektion* P. *Wegen* $(\mathbf{v} \otimes \mathbf{w}) \cdot \mathbf{x} = (\mathbf{w}.\mathbf{x}) \cdot \mathbf{v}$ *ist dies eine Projektion auf die Gerade* $L = \mathbb{R} \cdot \mathbf{v}$.

Dabei ist $Kern(P) = \mathbf{w}^\perp$. *Ist speziell* $\| \mathbf{v} \| = 1$ *und* $\mathbf{w} = \mathbf{v}$ *so erhält man die darstellende Matrix* $\mathbf{v} \otimes \mathbf{v}^t$ *der Orthogonalprojektion auf* L.

Ist allgemeiner $\mathbf{w} \in \mathbb{R}^n$ *ein beliebiger Vektor mit* $(\mathbf{v}.\mathbf{w}) \neq 0$, *so ist*

$$\frac{1}{(\mathbf{v}.\mathbf{w})} \cdot \mathbf{v} \otimes \mathbf{w}$$

die darstellende Matrix einer Projektion auf L. *Daher ist*

$$\mathbb{1}_n - \frac{1}{(\mathbf{v}.\mathbf{w})} \cdot \mathbf{v} \otimes \mathbf{w}$$

darstellende Matrix einer Projektion $P_{\mathbf{v}}$ *auf* $Kern(P) = \mathbf{w}^\perp$. *Wegen*

$$P_{\mathbf{v}}(\mathbf{x}) - \mathbf{x} = P(\mathbf{x}) \in L = \mathbb{R} \cdot \mathbf{v}$$

ist dies die Projektion auf die Hyperebene $\mathbf{w}^\perp$ *in Richtung von* $\mathbf{v}$.

Satz 2.15 (Projektion und direkte Summe) *Es sei* V *ein* $\mathbb{R}$-*Vektorraum.*

a) Ist $P : V \to V$ *eine Projektion, dan gibt es eine direkte Summenzerlegung*

$$V = Bild(P) \oplus Kern(P).$$

b) Ist $V = U \oplus W$ *eine direkte Summen-Zerlegung, dann gibt es genau eine Projektion* $P : V \to V$ *mit*

$$Bild(P) = U, \quad Kern(P) = W.$$

c) Es sei $dim(V) = n$ *endlich,* $P : V \to V$ *eine Projektion, sowie* $\mathbf{u}_1, ..., \mathbf{u}_r$ *eine Basis von* U *und* $\mathbf{w}_1, ..., \mathbf{w}_s$ *eine Basis von* W. *Dann ist* $\mathbf{u}_1, ..., \mathbf{u}_r, \mathbf{w}_1, ..., \mathbf{w}_s$ *eine Basis von* V, *und in dieser Basis hat* P *die darstellende Matrix*

$$\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \begin{array}{l} \left.\vphantom{\begin{matrix}1\\ \ddots\\ 1\end{matrix}}\right\} r \\ \left.\vphantom{\begin{matrix}0\\ \ddots\\ 0\end{matrix}}\right\} s \end{array}$$

Beweis. a) Für jeden Vektor $\mathbf{v} \in V$ ist

$$\mathbf{v} = id_V(\mathbf{v}) = (id_V - P)(\mathbf{v}) + P(\mathbf{v}) \quad \text{mit} \quad P(\mathbf{v}) \in Bild(P), \quad (id_V - P)(\mathbf{v}) \in Kern(P),$$

vgl. Satz 2.14 c). Das zeigt $V = Bild(P) + Kern(P)$. Diese Summe ist direkt, denn für $\mathbf{w} \in Bild(P) \cap Kern(P)$ gilt

$$\mathbf{w} = P(\mathbf{w}) = \mathbf{0}.$$

b) Nach Voraussetzung hat jedes $\mathbf{v} \in V$ eine eindeutige Darstellung $\mathbf{v} = \mathbf{u} + \mathbf{w}$ mit $\mathbf{u} \in U$ und $\mathbf{w} \in W$. Aus der Eindeutigkeit dieser Darstellung folgt, dass die Abbildung

$$P : V \to V, \quad \mathbf{v} \mapsto \mathbf{u},$$

linear und sogar eine Projektion auf U ist. Weiter besteht ihr Kern genau aus den Vektoren $\mathbf{v} = \mathbf{0} + \mathbf{w} = \mathbf{w} \in W$.

c) Wegen $V = Bild(P) \oplus Kern(P)$ ist $\mathbf{u}_1, ..., \mathbf{u}_r, \mathbf{w}_1, ..., \mathbf{w}_s$ tatsächlich eine Basis von V. Für $\rho = 1, ..., r$ ist $P(\mathbf{u}_\rho) = \mathbf{u}_\rho$ und für $\sigma = 1, ..., s$ ist $P(\mathbf{w}_\sigma) = \mathbf{0}$. $\square$

Der Begriff der direkten Summe von $\mathbb{R}$-Vektorräumen verallgemeinert sich in naheliegender Weise auf mehr als zwei Summanden:

Definition 2.12 *Es seien V ein $\mathbb{R}$-Vektorraum und $V_1, ..., V_m \subset V$ lineare Unterräume.*

a) In Verallgemeinerung von Beispiel 1.14 schreiben wir

$$V = V_1 + ... + V_m,$$

wenn $V = span(V_1 \cup ... \cup V_m)$.

b) Die Summe aus a) heißt direkt, *in Zeichen*

$$V = V_1 \oplus ... \oplus V_m, \qquad bzw. \quad V = \bigoplus_{\mu=1}^{m} V_\mu,$$

wenn zusätzlich für $i = 1, ..., m$ gilt

$$V_i \cap \sum_{\mu \neq i} V_\mu = \{\mathbf{0}\}.$$

Offensichtlich genügt es in b), die Bedingung

$$V_i \cap \sum_{\mu < i} V_\mu = \{\mathbf{0}\}$$

nachzuweisen. Ist $V = V_1 + ... + V_m$, so hat jeder Vektor $\mathbf{v}$ eine Darstellung $\mathbf{v} = \mathbf{v}_1 + ... + \mathbf{v}_m$ mit $\mathbf{v}_\mu \in V_\mu$. Genau dann ist $V = V_1 \oplus ... \oplus V_m$, wenn diese Darstellung eindeutig ist. D.h., die Vektoren $\mathbf{v}_\mu \in V_\mu$ sind durch $\mathbf{v}$ eindeutig bestimmt. Die Zuordnungen

$$P_\mu : V \to V_\mu, \quad \mathbf{v} \mapsto \mathbf{v}_\mu$$

sind lineare Abbildungen. Offenbar ist $P_\mu \cdot P_\mu = P_\mu$. Also sind die P_μ Projektionen auf die Unterräume V_μ. Weiter gelten

$$P_\mu \circ P_\nu = 0 \text{ für } \mu \neq \nu, \quad P_1 + ... + P_m = id_V.$$

Satz 2.16 *a) Für $V = V_1 + ... + V_m$ sind äquivalent:*

i) $V = V_1 \oplus ... \oplus V_m$, *ii) je m beliebige Vektoren $\mathbf{0} \neq \mathbf{v}_\mu \in V_\mu$ sind linear unabhängig.*

b) Für $V = V_1 \oplus ... \oplus V_m$ gilt

$$dim(V) = dim(V_1) + ... + dim(V_m).$$

Beweis. a) i) $\Rightarrow$ ii): Seien $\mathbf{0} \neq \mathbf{v}_\mu \in V_\mu$ und

$$c_1\mathbf{v}_1 + ... + c_m\mathbf{v}_m = \mathbf{0}, \quad c_1, ..., c_r \in \mathbb{R}.$$

Für $i = 1, ..., m$ folgt daraus

$$c_i\mathbf{v}_i = -\sum_{\mu \neq i} c_\mu \mathbf{v}_\mu \in V_i \cap \sum_{\mu \neq i} V_\mu = \{\mathbf{0}\}.$$

Wegen $\mathbf{v}_i \neq \mathbf{0}$ folgt daraus $c_i = 0$.

ii) $\Rightarrow$ i): Für $i = 1, ..., m$ ist Eigenschaft b) aus Definition 2.12 zu zeigen. Sei also

$$\mathbf{v}_i \in V_i \cap \sum_{\mu \neq i} V_\mu.$$

Daraus folgt

$$\mathbf{v}_i = \sum_{\mu \neq i} \mathbf{v}_\mu \text{ mit } \mathbf{v}_\mu \in V_\mu.$$

Wäre hier $\mathbf{v}_i \neq \mathbf{0}$, so hätten wir einen Widerspruch zu ii).

b) Sei $d_\mu = dim(V_\mu)$. Wir wählen Basen

$$\mathbf{v}_{1,1}, ..., \mathbf{v}_{1,d_1} \in V_1, ..., \mathbf{v}_{m,1}, ..., \mathbf{v}_{m,d_m} \in V_m.$$

Deren Vereinigungsmenge

$$\mathbf{v}_{1,1}, ..., \mathbf{v}_{m,d_m}$$

spannt $V = V_1 + ... + V_m$ auf. Aus a) folgt, dass diese Menge linear unabhängig ist. Sie ist also eine Basis von V der Länge $d_1 + ... + d_m$. □

Aufgabe 2.29 *Es seien $U, V \subset \mathbb{R}^n$ Unterverktorräume und P_U, P_V die Orthogonalprojektionen auf diese Unterräume. Zeigen Sie:*
a) $U \subset V \quad \Rightarrow \quad P_U \circ P_V = P_V \circ P_U = P_U$.
b) $P_U \circ P_V = P_V \circ P_U \quad \Rightarrow \quad P_U \circ P_V = P_{U \cap V}$.

2.3.4 Invertierbare Matrizen

Wir wollen nun die Matrix zur Umkehrabbildung Φ^{-1} bestimmen, wenn diese existiert. Dazu sei $\Phi : \mathbb{R}^m \to \mathbb{R}^n$ linear und bijektiv. Die Umkehrabbildung

$$\Phi^{-1} : \begin{cases} \mathbb{R}^n \to \mathbb{R}^m \\ \mathbf{y} \mapsto \mathbf{x} \text{ falls } \Phi(\mathbf{x}) = \mathbf{y} \end{cases}$$

kann wegen Satz 2.6 und Satz 2.5a) nur dann existieren, wenn $m = n$.

Sei nun $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ linear und invertierbar mit zugehöriger Matrix A. Die zu Φ^{-1} gehörige Matrix sei B. Da $\Phi^{-1} \circ \Phi = \Phi \circ \Phi^{-1} = id$, und da dem Hintereinanderausführen linearer Abbildungen die Matrizenmultiplikation entspricht, folgern wir

$$A \cdot B = B \cdot A = \mathbb{1}_n.$$

Definition 2.13 *Eine Matrix $A \in M(n \times n)$ heißt* invertierbar, *wenn es eine Matrix $B \in M(n \times n)$ gibt mit $AB = BA = 1\!\!1_n$. Weil Φ^{-1} durch Φ eindeutig bestimmt ist, ist auch B durch A eindeutig bestimmt. Wie schreiben*

$$B =: A^{-1}$$

und nennen es die inverse Matrix *zu A.*

Natürlich braucht für die Invertierbarkeit nur $A \cdot B = 1\!\!1_n$ oder $B \cdot A = 1\!\!1_n$ zu gelten, denn in beiden Fällen folgt, dass die lineare Abbildung mit Matrix B die Umkehrabbildung zur linearen Abbildung mit Matrix A ist.

Wir betrachten ein LGS

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

mit *invertierbarer* Koeffizientenmatrix $A \in M(n \times n)$. Dann ist die lineare Abbildung $\mathbb{R}^n \ni \mathbf{x} \mapsto A \cdot \mathbf{x} \in \mathbb{R}^n$ surjektiv. Das LGS ist damit für jede Wahl der rechten Seite $\mathbf{b}$ lösbar. In die Äquivalenzliste von Satz 1.28 kann also noch aufgenommen werden:

v) Die Matrix A ist invertierber.

Beispiel 2.19 *Die Inversen der Elementarmatrizen sind:*

$$\begin{pmatrix}
1 &&&&&&&&&& \\
& . &&&&&&&&& \\
&& 1 &&&&&&&& \\
&&& 0 & . & . & . & 1 &&& \\
&&& . & 1 &&& . &&& \\
&&& . && . && . &&& \\
&&& . &&& 1 & . &&& \\
&&& 1 & . & . & . & 0 &&& \\
&&&&&&&& 1 && \\
&&&&&&&&& . & \\
&&&&&&&&&& 1
\end{pmatrix}^{-1} = \begin{pmatrix}
1 &&&&&&&&&& \\
& . &&&&&&&&& \\
&& 1 &&&&&&&& \\
&&& 0 & . & . & . & 1 &&& \\
&&& . & 1 &&& . &&& \\
&&& . && . && . &&& \\
&&& . & . && 1 & . &&& \\
&&& 1 & . & . & . & 0 &&& \\
&&&&&&&& 1 && \\
&&&&&&&&& . & \\
&&&&&&&&&& 1
\end{pmatrix},$$

$$\begin{pmatrix}
1 &&&&&& \\
& . &&&&& \\
&& 1 &&&& \\
&&& c &&& \\
&&&& 1 && \\
&&&&& . & \\
&&&&&& 1
\end{pmatrix}^{-1} = \begin{pmatrix}
1 &&&&&& \\
& . &&&&& \\
&& 1 &&&& \\
&&& 1/c &&& \\
&&&& 1 && \\
&&&&& . & \\
&&&&&& 1
\end{pmatrix},$$

$$\begin{pmatrix}
1 &&&&&& \\
& . &&&&& \\
&& 1 & . & c && \\
&&& . & . && \\
&&&& 1 && \\
&&&&& . & \\
&&&&&& 1
\end{pmatrix}^{-1} = \begin{pmatrix}
1 &&&&&& \\
& . &&&&& \\
&& 1 & . & -c && \\
&&& . & . && \\
&&&& 1 && \\
&&&&& . & \\
&&&&&& 1
\end{pmatrix}.$$

Satz 2.17 *Sind A und B invertierbare $n \times n$-Matrizen, so ist auch ihr Produkt $A \cdot B$ invertierbar mit Inverser*

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

Beweis. Es seien $\Phi, \Psi : \mathbb{R}^n \to \mathbb{R}^n$ die linearen Abbildungen mit darstellender Matrix A, bzw. B. Beide Abbildungen sind surjektiv. Dann ist auch $\Phi \circ \Psi$ surjektiv. Weil diese Abbildung die darstellende Matrix $A \cdot B$ hat, ist diese Produktmatrix wieder invertierbar.

Die Umkehrabbildung zu $\Phi \circ \Psi$ ist

$$(\Phi \circ \Psi)^{-1} = \Psi^{-1} \circ \Phi^{-1}.$$

(Umgekehrte Reihenfolge!) Dazu gehört die darstellende Matrix $B^{-1} \cdot A^{-1}$. □

Beispiel 2.20 *Wann ist eine 2×2-Matrix*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

invertierbar? Nach Satz 2.10 ist dies genau dann der Fall, wenn wir A auf eine Stufenform

$$\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$$

bringen können. Falls $a \neq 0$ ist, dividieren wir erst die erste Zeile durch a und subtrahieren dann c-mal die neue erste Zeile von der zweiten. Wir erhalten die Stufenform

$$\begin{pmatrix} 1 & \frac{b}{a} \\ 0 & d - \frac{bc}{a} \end{pmatrix}.$$

In diesem Fall ist

$$\boxed{a \cdot d - b \cdot c \neq 0}$$

die Bedingung dafür, dass A invertierbar ist. Falls $a = 0$ und $c \neq 0$ ist, vertauschen wir erste und zweite Zeile und kommen zur selben Bedingung. Wenn aber $a = c = 0$ ist, ist die Dreiecksform nie zu erreichen. Es folgt: Unsere Bedingung $ad - bc \neq 0$ ist notwendig und hinreichend dafür, dass A invertierbar ist.

Wenn A invertierbar ist, so wollen wir A^{-1} auch ermitteln. Wieder wenden wir elementare Zeilenumformungen an, und zwar in einer Form, die man auch für größere Matrizen durchaus zur praktischen Bestimmung von A^{-1} heranzieht. (Wir diskutieren nur den Fall $a \neq 0$.):

Elementarmatrix	umgeformtes A	umgeformte Einheitsmatrix
	$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
$\begin{pmatrix} 1/a & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & b/a \\ c & d \end{pmatrix}$	$\begin{pmatrix} 1/a & 0 \\ 0 & 1 \end{pmatrix}$
$\begin{pmatrix} 1 & 0 \\ -c & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & b/a \\ 0 & d-bc/a \end{pmatrix}$	$\begin{pmatrix} 1/a & 0 \\ -c/a & 1 \end{pmatrix}$
$\begin{pmatrix} 1 & 0 \\ 0 & a/(ad-bc) \end{pmatrix}$	$\begin{pmatrix} 1 & b/a \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1/a & 0 \\ -c/(ad-bc) & a/(ad-bc) \end{pmatrix}$
$\begin{pmatrix} 1 & -b/a \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} d/(ad-bc) & -b/(ad-bc) \\ -c/(ad-bc) & a/(ad-bc) \end{pmatrix}$

Hier haben wir in der dritten Spalte dieselben elementaren Zeilenumformungen auf die Einheitsmatrix angewendet, wie auf die Matrix A. Am Schluss der dritten Spalte ergibt sich das Produkt der Elementarmatrizen, welche man von links an die Matrix A heranmultiplizieren muss, um A in die Einheitsmatrix zu überführen. Dieses Produkt ist deswegen die inverse Matrix

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

Dieses Verfahren ist nichts Neues, es handelt sich nur um eine Variante des Gauß-Verfahrens. Sind nämlich $\mathbf{c}^1, ..., \mathbf{c}^n$ die Spalten der (gesuchten) Matrix A^{-1}, so ist

$$A \cdot A^{-1} = \mathbb{1}_n \quad \text{bzw. äquivalent dazu} \quad A \cdot \mathbf{c}^\nu = \mathbf{e}_\nu \text{ für } \nu = 1, ..., n.$$

Die ν-te Spalte $\mathbf{c}^\nu$ ist also Lösungsvektor eines LGS mit der rechten Seite $\mathbf{e}_\nu$. Aber alle diese n Gleichungssysteme haben dieselbe Koeffizientenmatrix A. Wir lösen sie simultan, indem wir A nicht um eine rechte Seite erweitern, sondern um n rechte Seiten $\mathbf{e}_\nu$, d.h., um die Matrix $\mathbb{1}_n$. Ausgangspunkt der Umformungen ist also die $n \times 2n$-Matrix

$$(A|\mathbb{1}_n).$$

Hierauf wenden wir wie immer elementare Zeilenumformungen an und überführen A in Zeilenstufenform Z und die Matrix $(A|\mathbb{1}_n)$ in

$$(Z|B).$$

Hier könnten wir Schluss machen, und mit jeder Spalte $\mathbf{b}^\nu$ der Matrix B den Lösungsvektor $\mathbf{c}^\nu$ berechnen.

Aber: Mit A hat auch Z maximalen Rang $= n$. Mit dem Verfahren im Anschluss an Satz 1.2 können wir weitere Zeilenumformungen vornehmen und Z in die Einheitsmatrix $\mathbb{1}_n$ überführen. Dabei verändert sich $(Z|B)$ in

$$(\mathbb{1}_n|C).$$

Hier kann man die Lösungsvektoren sofort ablesen. Sie sind identisch mit den Spalten der rechten Seite C, d.h.

$$A^{-1} = C.$$

Es sollte aber festgehalten werden: die inverse Matrix A^{-1} ist vor allem theoretisch von Interesse. Praktisch ist ihre Bestimmung meist sehr aufwendig.

Aufgabe 2.30 *Bestimmen Sie die Inversen der Matrizen*

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Aufgabe 2.31 *Es bezeichne $\Delta \subset M(n \times n)$ die Menge der oberen Dreiecksmatrizen*

$$A = \begin{pmatrix} a_1 & * & * \\ 0 & \ddots & * \\ 0 & 0 & a_n \end{pmatrix}.$$

Zeigen Sie:
a) Für alle $A, B \in \Delta$ gilt $A \cdot B \in \Delta$.
b) Zu $A \in \Delta$ gibt es ein $B \in \Delta$ mit $A \cdot B = \mathbb{1}_n$ genau dann, wenn $a_1 \cdot ... \cdot a_n \neq 0$.

Aufgabe 2.32 *Es sei A eine reelle $n \times n$-Matrix, E die Einheitsmatrix, es sei $(A - E)$ invertierbar, und es sei $B := (A + E)(A - E)^{-1}$. Man beweise*
a) $(A + E)(A - E)^{-1} = (A - E)^{-1}(A + E)$ durch Betrachtung von

$$(A - E + 2E)(A - E)^{-1} - (A - E)^{-1}(A - E + 2E).$$

b) $(B - E)$ ist invertierbar, indem man $B - (A - E)(A - E)^{-1} = 2(A - E)^{-1}$ zeigt.
c) $(B + E)(B - E)^{-1} = A$.

Aufgabe 2.33 *Man berechne die Inverse der Matrix*

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

Aufgabe 2.34 *In der Matrix*

$$A = \begin{bmatrix} P & Q \\ R & S \end{bmatrix} \in M((m + n) \times (m + n))$$

sei die Untermatrix $P \in M(m \times m)$ invertierbar. Zeigen Sie: A ist genau dann invertierbar, wenn dies für $S - RP^{-1}Q$ gilt.

2.3.5 Die transponierte Matrix

Wir verallgemeinern das Transponieren von Vektoren $\mathbf{x} \in \mathbb{R}^n$ (Abschnitt 1.2.1) durch

Definition 2.14 *Es sei*

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

eine reelle $m \times n$-Matrix. Dann heißt die $n \times m$-Matrix

$$A^t = \begin{pmatrix} a_{1,1} & a_{2,1} & \cdots & a_{m,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{m,2} \\ \vdots & \vdots & & \vdots \\ a_{1,n} & a_{2,n} & \cdots & a_{m,n} \end{pmatrix}$$

die transponierte *Matrix zu A. Sie entsteht aus A durch Vertauschen der Zeilen mit den Spalten. Die Abbildung*

$$A \mapsto A^t, \quad M(m \times n) \to M(n \times m)$$

heißt Transposition*.*

Beispiel 2.21 *Es ist*

$$\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}^t = \begin{pmatrix} 1 & 0 \\ 2 & 3 \end{pmatrix}.$$

Offensichtliche Eigenschaften der Transposition sind

$$A^{tt} = A, \quad (A+B)^t = A^t + B^t, \quad (c \cdot A)^t = c \cdot A^t \text{ für } c \in \mathbb{R}.$$

Die Tranposition ist also eine lineare Abbildung $M(m \times n) \to M(n \times m)$, die ihre eigene Umkehrabbildung ist.

Bei der Transposition werden Zeilen und Spalten vertauscht, also auch Zeilenrang und Spaltenrang. Aber mit Satz 1.25 folgt:

Satz 2.18 *Der Rang einer Matrix stimmt mit dem Rang ihrer transponierten Matrix überein:*

$$Rang(A) = Rang(A^t).$$

Insbesondere ist die $n \times n$-Matrix A invertierbar genau dann, wenn A^t invertierbar ist.

Satz 2.19 *a) Für $A \in M(l \times m)$ und $B \in M(m \times n)$ gilt*

$$(A \cdot B)^t = B^t \cdot A^t.$$

b) Für eine invertierbare $n \times n$-Matrix A ist

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

Beweis. a) Wegen der Distributivität des Matrizenprodukts genügt es, die Aussage für Matrizen A und B zu zeigen, die nur eine einzige Eins und sonst lauter Nullen als Einträge enthalten. Dazu kürzen wir ab $E_{i,j} := \mathbf{e}_i \otimes \mathbf{e}_j$. Die Transponierte dieser Matrix ist $E_{j,i}$. Wir betrachten eine Produktmatrix $E_{i,j} \cdot E_{k,l}$.

Für $j \neq k$ ist $E_{i,j} \cdot E_{k,l} = 0$ die Nullmatrix, ebenso wie $E^t_{k,l} \cdot E^t_{i,j} = E_{l,k} \cdot E_{j,i}$.

Für $j = k$ ist $E_{i,j} \cdot E_{j,k} = E_{i,k}$ und

$$E^t_{j,k} \cdot E^t_{i,j} = E_{k,j} \cdot E_{j,i} = E_{k,i} = E^t_{i,k} = (E_{i,j} \cdot E_{j,k})^t.$$

b) Mit a) folgt

$$(A^{-1})^t \cdot A^t = (A \cdot A^{-1})^t = \mathbb{1}^t_n = \mathbb{1}_n.$$

□

Transponierte Matrizen treten in Zusammenhang mit dem euklidischen Skalarprodukt im $\mathbb{R}^n$ auf.

Satz 2.20 *Es sei A eine $n \times n$-Matrix. Für je zwei Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ gilt*

$$((A \cdot \mathbf{x}).\mathbf{y}) = (\mathbf{x}.(A^t \cdot \mathbf{y})).$$

Beweis. Wegen $(\mathbf{x}.\mathbf{y}) = \mathbf{x}^t \cdot \mathbf{y}$ ist

$$((A \cdot \mathbf{x}).\mathbf{y}) = (A \cdot \mathbf{x})^t \cdot \mathbf{y} = \mathbf{x}^t \cdot A^t \cdot \mathbf{y} = \mathbf{x}^t \cdot (A^t \cdot \mathbf{y}) = (\mathbf{x}.(A^t \cdot \mathbf{y})).$$

□

Satz 2.21 *Für eine $n \times n$-Matrix A sind äquivalent:*

a) Die Abbildung

$$\mathbb{R}^n \to \mathbb{R}^n, \quad \mathbf{x} \mapsto A \cdot \mathbf{x},$$

ist orthogonal.

b) Es gilt $A^t \cdot A = \mathbb{1}_n$. Insbesondere ist A invertierbar und $A \cdot A^t = \mathbb{1}_n$.

c) Die Spaltenvektoren von A bilden eine ONB des $\mathbb{R}^n$.

d) Die Zeilenvektoren von A bilden eine ONB des $\mathbb{R}^n$.

Beweis. a) $\Leftrightarrow$ b): Die Abbildung $\mathbf{x} \mapsto A \cdot \mathbf{x}$ ist genau dann orthogonal, wenn für alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ gilt

$$((A \cdot \mathbf{x}).(A \cdot \mathbf{y})) = (\mathbf{x}.\mathbf{y}).$$

Dies heißt

$$(\mathbf{x}.\mathbf{y}) = (A \cdot \mathbf{x})^t \cdot (A \cdot \mathbf{y}) = \mathbf{x}^t \cdot A^t \cdot A \cdot \mathbf{y}.$$

Setzen wir hier insbesondere $\mathbf{x} = \mathbf{e}_i, \mathbf{y} = \mathbf{e}_j$ mit $(\mathbf{e}_i.\mathbf{e}_j) = \delta_{i,j}$, so folgt daraus

$$\delta_{i,j} = (A^t \cdot A)_{i,j}, \quad A^t \cdot A = \mathbb{1}_n.$$

Ist umgekehrt $A^t \cdot A = \mathbb{1}^n$, so folgt

$$((A \cdot \mathbf{x}).(A \cdot \mathbf{y})) = (\mathbf{x}.(A^t \cdot A) \cdot \mathbf{y}) = (\mathbf{x}.\mathbf{y}).$$

b) $\Leftrightarrow$ c): Die Matrixgleichung $A^t \cdot A = \mathbb{1}_n$ bedeutet für die Spaltenvektoren $\mathbf{a}^1, ..., \mathbf{a}^n$ von A, dass $(\mathbf{a}^i.\mathbf{a}^j) = \delta_{i,j}$.

b) $\Leftrightarrow$ d): Die Matrixgleichung $A \cdot A^t = \mathbb{1}_n$ bedeutet für die Zeilenvektoren $\mathbf{a}_1, ..., \mathbf{a}_n$ von A, dass $(\mathbf{a}_i.\mathbf{a}_j) = \delta_{i,j}$. □

Definition 2.15 *eine $n \times n$-Matrix A heißt* orthogonal*, wenn sie die Eigenschaften aus Satz 2.21 hat. Hier liegt ein gewisser Sprach-Missbrauch vor: Die Matrix heißt ortho***gonal***, aber ihre Spalten (Zeilen) sind ortho***normal***.*

Beispiel 2.22 *Jede* 2×2*-Drehmatrix*

$$\begin{pmatrix} cos(\alpha) & -sin(\alpha) \\ sin(\alpha) & cos(\alpha) \end{pmatrix}$$

ist orthogonal.

Satz 2.22 *Für eine* $n \times n$*-Matrix* A *sind äquivalent:*
a) Es ist $A^t = A$.
b) Für alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ *gilt*

$$((A \cdot \mathbf{x}).\mathbf{y}) = (\mathbf{a}.(A \cdot \mathbf{y})).$$

Beweis. Es ist

$$((A \cdot \mathbf{x}).\mathbf{y}) = (A \cdot \mathbf{x})^t \cdot \mathbf{y} = \mathbf{x}^t \cdot (A^t \cdot \mathbf{y}) = (\mathbf{x}.(A^t \cdot \mathbf{y})).$$ □

Definition 2.16 *Eine* $n \times n$ *Matrix* S *mit* $S^t = S$ *heißt* symmetrisch. *(Wie sonst ?)*

Beispiel 2.23 *Das Tensorprodukt* $\mathbf{x} \otimes \mathbf{x}$ *eines Vektors* $\mathbf{x} \in \mathbb{R}^n$ *mit sich selbst ist symmetrisch. Daraus folgt, dass die Darstellungsmatrizen für die folgenden linearen Abbildungen symmetrisch sind:*

- *Spiegelungen (Beispiel 2.8),*
- *orthogonale Projektionen auf Geraden oder Hyperebenen (2.3.2),*
- *Orthogonalprojektion auf einen Untervektorraum (Beispiel 2.11 und 2.3.2).*

Den Begriff der transponierten Matrix verallgemeinert man wie folgt auf lineare Abbildungen zwischen Vektorräumen mit SKP:

Definition 2.17 *Es seien* V, W *reelle Vektorräume mit SKP (das wir in beiden Fällen als* $(\ . \ )$ *schreiben) und* $\Phi : V \to W$ *linear. Eine Abbildung* $\Phi^t : W \to V$ *heißt transponiert zu* Φ *(oder adjungiert zu* Φ*), wenn*

$$(\Phi(\mathbf{v}).\mathbf{w}) = (\mathbf{v}.\Phi^t(\mathbf{w})) \quad \textit{für alle } \mathbf{v} \in V, \mathbf{w} \in W.$$

In dieser allgemeinen Situation bleibt offen, ob Φ^t existiert. Aber die Eindeutigkeit ist klar:

Satz 2.23 *Wenn* Φ^t *existiert, dann ist es durch* Φ *eindeutig bestimmt.*

Beweis. Es seien $\Psi_1, \Psi_2 : W \to V$ zwei adjungierte Abbildungen zu Φ. Dann ist also für alle $\mathbf{v} \in V, \mathbf{w} \in W$

$$(\mathbf{v}.\Psi_1(\mathbf{w})) = (\Phi(\mathbf{v}).\mathbf{w}) = (\mathbf{v}.\Psi_2(\mathbf{w})).$$

Anders hingeschrieben heißt dies

$$(\mathbf{v}.(\Psi_1 - \Psi_2)(\mathbf{w})) = 0 \quad \text{für alle } \mathbf{v} \in V.$$

Mit Beispiel 1.41 folgt daraus $(\Psi_1 - \Psi_2)(\mathbf{w}) = \mathbf{0}$, und das für alle $\mathbf{w} \in W$. Das bedeutet $\Psi_1 = \Psi_2$. □

Im Fall endlicher Dimension führt auch diese allgemeinere Definition genau auf die transponierte Matrix.

Satz 2.24 *Es sei $\mathbf{v}_1, ..., \mathbf{v}_n \in V$ eine ONB und $\mathbf{w}_1, ..., \mathbf{w}_m \in W$ eine Basis. Dann gibt es zu jeder linearen Abbildung $\Phi : V \to W$ eine transponierte Abbildung $\Phi^t : W \to V$ definiert durch*

$$\Phi^t(\mathbf{w}_\mu) := \sum_{\nu=1}^{n} (\Phi(\mathbf{v}_\nu).\mathbf{w}_\mu)\mathbf{v}_\nu.$$

Beweis. Durch die Bilder $\Phi^t(\mathbf{w}_\mu)$ wird nach dem Prinzip der linearen Ausdehnung (Satz 2.7) eine lineare Abbildung $\Phi^t : W \to V$ eindeutig festgelegt. Für diese gilt

$$(\mathbf{v}_i.\Phi^t(\mathbf{w}_j)) = \left(\mathbf{v}_i. \sum_{\nu=1}^{n} (\Phi(\mathbf{v}_\nu).\mathbf{w}_j)\mathbf{v}_\nu \right) = (\Phi(\mathbf{v}_i).\mathbf{w}_j).$$

Aus dieser Formel für die Basisvektoren folgt

$$(\Phi(\mathbf{v}).\mathbf{w}) = (\mathbf{v}.\Phi^t(\mathbf{w}))$$

für alle $\mathbf{v} \in V$ und $\mathbf{w} \in W$. □

Insbesondere für $V = \mathbb{R}^n$ mit der Basis $\mathbf{e}_1, ..., \mathbf{e}_n$ und $W = \mathbb{R}^m$ mit der Basis $\mathbf{e}_1, ..., \mathbf{e}_m$ erhalten wir: Hat $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ die darstellende Matrix

$$A = (a_{\mu,\nu}), \quad a_{\mu,\nu} = (\Phi(\mathbf{e}_\nu).\mathbf{e}_\mu),$$

so ist

$$\Phi^t(\mathbf{e}_\mu) = \sum_{\nu=1}^{n} a_{\mu,\nu}\mathbf{e}_\nu.$$

Somit ist A^t die darstellende Matrix von Φ^t. Ganz wie es sein soll.

Weil der Rang einer Matrix mit dem Rang ihrer Transponierten übereinstimmt, sehen wir insbesondere:

Satz 2.25 *Es seien V und W endlich-dimensionale $\mathbb{R}$-Vektorräume mit SKP. Ist $\Phi : V \to W$ linear, so gilt*

$$dim(Bild(\Phi)) = dim(Bild(\Phi^t)).$$

Satz 2.26 *a) Es sei $\Phi : V \to W$ eine lineare Abbildung zwischen (beliebig-dimensionalen) Vektorräumen V und W mit SKP. Dann gilt (falls Φ^t existiert)*

$$Bild(\Phi)^\perp = Kern(\Phi^t).$$

b) Ist V endlich-dimensional, so haben wir auch

$$(Kern(\Phi))^\perp = Bild(\Phi^t).$$

Beweis. a) Für $\mathbf{w} \in W$ gilt

$$\begin{aligned} \mathbf{w} \in Bild(\Phi)^\perp \quad &\Leftrightarrow \quad (\mathbf{w}.\Phi(\mathbf{v})) = 0 \text{ für alle } \mathbf{v} \in V \\ &\Leftrightarrow \quad (\Phi^t(\mathbf{w}).\mathbf{v}) = 0 \text{ für alle } \mathbf{v} \in V \\ &\Leftrightarrow \quad \Phi^t(\mathbf{w}) = \mathbf{0}. \end{aligned}$$

b) Wegen $\Phi^{tt} = \Phi$ folgt aus a), dass

$$Bild(\Phi^t)^\perp = Kern(\Phi).$$

Wenn V endlich-dimensional ist, dann haben wir für $U \subset V$ $(U^\perp)^\perp = U$ und deswegen

$$Bild(\Phi^t) = (Bild(\Phi^t)^\perp)^\perp = Kern(\Phi)^\perp. \qquad \square$$

Wir verallgemeinern die Bezeichnungen 'orthogonal' und 'symmetrisch' für Matrizen auf lineare Abbildungen.

Definition 2.18 *Es seien V ein $\mathbb{R}$-Vektorraum mit SKP und $\Phi : V \to V$ linear. Die Abbildung Φ heißt*

a) orthogonal*, wenn Φ bijektiv und $\Phi^{-1} = \Phi^t$ ist;*

b) symmetrisch *(oder* selbstadjungiert*), wenn $\Phi = \Phi^t$.*

Es seien $\mathbf{v}, \mathbf{w} \in V$. Für jede orthogonale Abbildung $\Phi : V \to V$ gilt

$$(\Phi(\mathbf{v}).\Phi(\mathbf{w})) = (\mathbf{v}.\Phi^t\Phi(\mathbf{w})) = (\mathbf{v}.\mathbf{w}),$$

und für jede symmetrische Abbildung

$$(\Phi(\mathbf{v}).\mathbf{w}) = (\mathbf{v}.\Phi^t(\mathbf{w})) = (\mathbf{v}.\Phi(\mathbf{w})).$$

Satz 2.27 *a) Für jede symmetrische Abbildung $\Phi : V \to V$ ist*

$$Kern(\Phi) = Bild(\Phi)^\perp.$$

Insbesondere ist für endlich-dimensionales V

$$V = Bild(\Phi) \oplus Kern(\Phi)$$

eine orthogonale direkte Summe.

b) Eine Projektion $P : V \to V$ ist genau dann eine Orthogonalprojektion, wenn P symmetrisch ist.

Beweis. a) Wegen $\Phi = \Phi^t$ folgt die Aussage aus Satz 2.26 a).

b) $\Rightarrow$: Es sei P die Orthogonalprojektion auf den Unterraum $U \subset V$. Wir haben

$$(P(\mathbf{v}).\mathbf{w}) = (\mathbf{v}.P(\mathbf{w})) \quad \text{für alle } \mathbf{v}, \mathbf{w} \in V$$

zu zeigen. Für alle $\mathbf{v} \in V$ ist aber $\mathbf{v} - P(\mathbf{v}) \in U^\perp$. Deswegen ist für alle $\mathbf{w} \in V$

$$(P(\mathbf{v}) - \mathbf{v}.P(\mathbf{w})) = 0 \text{ und ebenso } (P(\mathbf{w}) - \mathbf{w}.P(\mathbf{v})) = 0.$$

Daraus folgt

$$(P(\mathbf{v}).\mathbf{w}) = (\mathbf{w}.P(\mathbf{v})) = (P(\mathbf{w}).P(\mathbf{v})) = (P(\mathbf{v}).P(\mathbf{w})) = (\mathbf{v}.P(\mathbf{w})).$$

$\Leftarrow$: Jetzt ist $P^t = P$ symmetrisch vorausgesetzt und $Kern(P) \perp Bild(P)$ zu zeigen. Das ist aber in a) geschehen. $\square$

Für jede $m \times n$-Matrix A ist die $m \times m$-Matrix $A^t \cdot A$ symmetrisch. Denn nach Satz 2.19 a) gilt

$$(A^t \cdot A)^t = A^t \cdot (A^t)^t = A^t \cdot A.$$

Satz 2.28 *Für jede $m \times n$-Matrix A hat die symmetrische $n \times n$-Matrix $A^t \cdot A$ den gleichen Rang wie A*

Beweis. Es sei $\mathbf{v} \in \mathbb{R}^n$. Aus

$$(A^t \cdot A) \cdot \mathbf{v} = \mathbf{0}$$

folgt

$$\mathbf{v}^t \cdot (A^t \cdot A) \cdot \mathbf{v} = (A \cdot \mathbf{v})^t \cdot (A \cdot \mathbf{v}) = \parallel A \cdot \mathbf{v} \parallel^2 = 0$$

und damit $A \cdot \mathbf{v} = \mathbf{0}$. Deswegen definieren die beiden Matrizen A und $A^t \cdot A$ lineare Abbildungen mit demselben Kern $U \subset \mathbb{R}^n$. Nach der Dimensionsformel (Satz 2.5 b) haben beide Abbildungen den gleichen Rang.

Aufgabe 2.35 *Es sei*

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

Bestimmen Sie Kern und Bild der linearen Abbildungen, welche durch die symmetrischen Matrizen $A \cdot A^t$, *bzw.* $A^t \cdot A$ *gegeben werden.*

Aufgabe 2.36 *Für die Matrizen* $A \in M(s \times m)$, $B \in M(s \times n)$ *und* $C \in M(n \times m)$ *soll*

$$A = BC \quad \text{und} \quad rang(A) = rang(C)$$

gelten. $N(A)$ *sei der Nullraum von* A, *d.h.* $N(A) = \{\mathbf{x} \in \mathbb{R}^m : A\mathbf{x} = \mathbf{0}\}$. *Zeigen Sie:*

$$N(A) = N(C).$$

2.4 Ausgleichsrechnung und Pseudo-Inverse

Es sei A eine $m \times n$-Matrix. In 1.4.2 haben wir ihren Spaltenraum $S(A)$ und ihren Zeilenraum $Z(A)$ definiert. Ist $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ die lineare Abbildung $\mathbf{x} \mapsto A \cdot \mathbf{x}$, so sagen wir - etwas unkorrekt - auch

$$Kern(A) := Kern(\Phi) \subset \mathbb{R}^n, \quad Bild(A) = Bild(\Phi) \subset \mathbb{R}^m.$$

Damit haben wir

$$Bild(A) = S(A) \subset \mathbb{R}^m, \quad Bild(A^t) = Z(A) \subset \mathbb{R}^n$$

und

- $Kern(A) \subset \mathbb{R}^n$ ist der Lösungsraum des LGS $A \cdot \mathbf{x} = \mathbf{0}$,
- $Kern(A^t) \subset \mathbb{R}^m$ ist der Lösungsraum des LGS $A^t \cdot \mathbf{x} = \mathbf{0}$ bzw. $\mathbf{x}^t \cdot A = \mathbf{0}$.

Satz 2.26 in seiner Version für Matrizen lautet:

Satz 2.29 *Es ist*

$$Bild(A)^\perp = Kern(A^t), \quad \textit{bzw. dazu äquivalent} \quad Bild(A) = Kern(A^t)^\perp \quad \subset \mathbb{R}^m$$

und

$$Kern(A)^\perp = Bild(A^t), \quad \textit{bzw. dazu äquivalent} \quad Kern(A) = Bild(A^t)^\perp \quad \subset \mathbb{R}^n.$$

Die erste Gleichung liefert folgendes Lösungskriterium für ein LGS:

Satz 2.30 *Es sei A eine $m \times n$-Matrix und $\mathbf{b} \in \mathbb{R}^m$. Das LGS $A \cdot \mathbf{x} = \mathbf{b}$ ist genau dann lösbar, wenn*

$$\mathbf{b} \in Kern(A^t)^\perp,$$

d.h.

$$(\mathbf{b}.\mathbf{x}) = 0 \quad \textit{für alle } \mathbf{x} \in \mathbb{R}^m \textit{ mit } A^t \cdot \mathbf{x} = \mathbf{0}.$$

Weil Zeilenrang und Spaltenrang von A übereinstimmen, haben wir

$$dim(Bild(A)) = dim(Bild(A^t)).$$

Aus der Dimensionsformel Satz 2.5 b) folgern wir

$$dim(Kern(A)) + dim(Bild(A)) = n \quad \Rightarrow \quad dim(Kern(A)) + dim(Bild(A^t)) = n,$$

$$dim(Kern(A^t)) + dim(Bild(A^t)) = m, \quad \Rightarrow \quad dim(Kern(A^t)) + dim(Bild(A)) = m.$$

Wenn ein Mathematiker in der Wirtschaft zu seinem Chef kommt und ihm mitteilt:
„Chef, das Problem, was Sie mir vor vier Wochen gegeben haben, ist ganz toll, es führt auf ein unlösbares LGS!“
dann wird er wohl kaum zur Beförderung vorgemerkt. Besser ist es, wenn er sagt:
„Chef, ich bin auf ein LGS

$$A \cdot \mathbf{x} = \mathbf{b}, \quad A \in M(m \times n), \mathbf{x} \in \mathbb{R}^n, \mathbf{b} \in \mathbb{R}^m,$$

gekommen, das ich zwar nicht ganz exakt lösen kann, aber so genau kommt es ja auch nicht darauf an. Ich kann Ihnen nämlich genau sagen, für welches $\mathbf{x} \in \mathbb{R}^n$ der Fehler

$$A \cdot \mathbf{x} - \mathbf{b}$$

minimal ist. “

Das ist das *lineare Ausgleichsproblem:* Gegeben ist ein (unlösbares) LGS

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

Gesucht ist ein Vektor $\mathbf{x} \in \mathbb{R}^n$, für den

$$\| A \cdot \mathbf{x} - \mathbf{b} \| \leq \| A \cdot \mathbf{y} - \mathbf{b} \| \quad \text{für alle } \mathbf{y} \in \mathbb{R}^n$$

ist. Hier ist $\| \quad \|$ die euklidische Norm im $\mathbb{R}^m$. Die Vektoren $\mathbf{u} = A \cdot \mathbf{y}$ durchlaufen den Untervektorraum $Bild(A) \subset \mathbb{R}^m$. Und nach Satz 1.32 ist $A \cdot \mathbf{x}$ die Orthogonalprojektion von $\mathbf{b}$ in diesen Unterraum. Nach Satz 1.33 ist diese eindeutig bestimmt und nach Satz 1.35 existiert sie. Sie ist dadurch charakterisiert, dass

$$A \cdot \mathbf{x} - \mathbf{b} \in (Bild(A))^\perp = Kern(A^t),$$

also $\mathbf{x}$ Lösung des LGS

$$A^t \cdot A \cdot \mathbf{x} = A^t \cdot \mathbf{b}$$

ist. Dieses LGS heißt die *Normalengleichung.*

Satz 2.31 *a) Für $A \in M(m \times n)$ und $\mathbf{b} \in \mathbb{R}^m$ ist das lineare Ausgleichsproblem immer lösbar. Alle Lösungen $\mathbf{x}$ erfüllen die Normalengleichung.*

b) Die Lösung $\mathbf{x} \in \mathbb{R}^n$ ist genau dann eindeutig bestimmt, wenn Rang$(A) = n$.

Beweis. Mit der Identität $(Bild(A))^\perp = Kern(A^t)$ ist die Existenz einer Lösung für das Ausgleichsproblem nur eine Umformulierung der Existenz der Orthogonalpropjektion. Diese Orthogonalprojektion $A \cdot \mathbf{x}$ ist durch $\mathbf{b}$ eindeutig bestimmt. Aber die Lösung $\mathbf{x} \in \mathbb{R}^n$ ist durch $A \cdot \mathbf{x} \in \mathbb{R}^m$ genau dann eindeutig bestimmt, wenn die Abbildung

$$\mathbb{R}^n \to \mathbb{R}^m, \quad \mathbf{x} \mapsto A \cdot \mathbf{x}$$

injektiv ist. Das ist genau dann der Fall, wenn A den Rang n hat. $\square$

Beispiel 2.24 (Lineare Regression) *Gesucht sei eine lineare Funktion $p(x) = a + b \cdot x$, $a, b, \in \mathbb{R}$, welche die (Mess-) Punkte*

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \quad \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} x_3 \\ y_3 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad \begin{pmatrix} x_4 \\ y_4 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

optimal interpoliert. Dabei soll der Gesamtfehler als

$$\sqrt{\sum_{i=1}^{4} (p(x_i) - y_i)^2}$$

angesetzt werden.

Es geht also um die lineare Abbildung

$$\mathbb{R}^2 \to \mathbb{R}^4, \quad \begin{pmatrix} a \\ b \end{pmatrix} \mapsto \begin{pmatrix} a + b \cdot x_1 \\ a + b \cdot x_2 \\ a + b \cdot x_3 \\ a + b \cdot x_4 \end{pmatrix} = A \cdot \begin{pmatrix} a \\ b \end{pmatrix}$$

mit der 4×2-Matrix

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

Der Wertevektor

$$\mathbf{b} = (2, 1, 3, 1)^t \in \mathbb{R}^4$$

liegt nicht im Bild dieser Abbildung. Aber es soll ein $\mathbf{x} = (a, b)^t \in \mathbb{R}^2$ gefunden werden, derart, dass der Abstand $\parallel A \cdot \mathbf{x} - \mathbf{b} \parallel$ minimal wird. Das ist genau das soeben behandelte lineare Ausgleichsproblem. Wir berechnen

$$A^t \cdot A = \begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix}, \quad A^t \cdot \mathbf{b} = \begin{pmatrix} 7 \\ 10 \end{pmatrix}$$

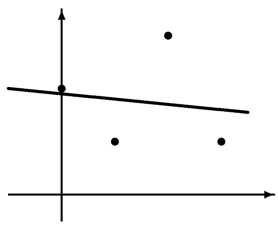
und kommen auf die Normalengleichung

$$\begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 7 \\ 10 \end{pmatrix}.$$

Wie üblich findet man die Lösung

$$a = 1.9, \quad b = -0.1,$$

welche zum Polynom $p(x) = 1.9 - 0.1 \cdot x$ gehört.



2.4.1 Spezialfall: Voller Spaltenrang

Im diesem Unterabschnitt nehmen wir an, die $m \times n$-Matrix habe vollen Spaltenrang, also den Rang n. Dann ist die Zeilenzahl $m \geq$ Spaltenzahl n. Nach Satz 2.28 hat die $n \times n$-Matrix $A^t \cdot A$ den gleichen Rang n wie A. Diese Matrix $A^t \cdot A$ ist also invertierbar. Die Lösung $\mathbf{x}$ der Normalengleichung ist eindeutig bestimmt, und wir können

$$\mathbf{x} = (A^t \cdot A)^{-1} \cdot A^t \cdot \mathbf{b}$$

schreiben.

Definition 2.19 *Die $m \times n$-Matrix A habe den Rang n. Dann heißt die $n \times m$-Matrix*

$$A^+ := (A^t \cdot A)^{-1} \cdot A^t$$

die Pseudo-Inverse *zu A.*

Beispiel 2.25 *In Beispiel 2.24 haben wir für die Matrix*

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

berechnet

$$A^t \cdot A = \begin{pmatrix} 4 & 6 \\ 6 & 14 \end{pmatrix}.$$

Damit erhalten wir

$$(A^t \cdot A)^{-1} = \frac{1}{10} \begin{pmatrix} 7 & -3 \\ -3 & 2 \end{pmatrix}$$

und

$$A^+ = (A^t \cdot A)^{-1} \cdot A^t = \frac{1}{10} \begin{pmatrix} 7 & 4 & 1 & -2 \\ -3 & -1 & 1 & 3 \end{pmatrix}.$$

Satz 2.32 (Pseudo-Inverse, $Rang(A) = n$**)** *a) Für $m = n$ ist $A^+ = A^{-1}$.*

b) Es ist stets $A^+ \cdot A = \mathbb{1}_n$.

c) Die $m \times m$-Matrix $A \cdot A^+$ beschreibt die Orthogonalprojektion $P_{Bild(A)}\colon \mathbb{R}^m \to \mathbb{R}^m$ auf den Unterraum $Bild(A)$.

Beweis. a) Für $m = n$ ist $(A^t \cdot A)^{-1} = A^{-1} \cdot (A^t)^{-1}$ und $A^+ = A^{-1} \cdot (A^t)^{-1} \cdot A^t = A^{-1}$.
b) Wir berechnen
$$A^+ \cdot A = (A^t \cdot A)^{-1} \cdot A^t \cdot A = \mathbb{1}_n.$$
c) Für $\mathbf{b} = A \cdot \mathbf{x} \in \mathbb{R}^m$ im Bild von A gilt
$$(A \cdot A^+) \cdot A \cdot \mathbf{x} = A \cdot (A^+ A) \cdot \mathbf{x} = A \cdot \mathbf{x}$$
wegen b). Deswegen ist $A \cdot A^+$ die Matrix zu einer Projektion auf $Bild(A)$. Sei nun $\mathbf{b} \in \mathbb{R}^m$ orthogonal zu $Bild(A)$. Das heißt $A^t \cdot \mathbf{b} = \mathbf{0}$. Daraus folgt
$$A \cdot A^+ \cdot \mathbf{b} = A \cdot (A^t \cdot A)^{-1} \cdot A^t \cdot \mathbf{b} = \mathbf{0}.$$
Also ist diese Projektion die Orthogonalprojektion $P_{Bild(A)}$. □

Satz 2.32 a) bedeutet insbesondere, dass für eine invertierbare $n \times n$-Matrix A gilt
$$A^+ = A^{-1}.$$

2.4.2 Allgemeinfall

Im letzten Unterabschnitt haben wir vorausgesetzt, dass die durch A beschriebene lineare Abbildung $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ injektiv ist. Dann definiert Φ einen Isomorphismus von $\mathbb{R}^n$ auf $Bild(\Phi) \subset \mathbb{R}^m$. Dies verallgemeinern wir jetzt auf den Fall, wo Φ nicht (notwendig) injektiv ist.

Wenn Φ nicht injektiv ist, dann ist $Kern(\Phi)$ nicht nur der Nullraum und $Kern(\Phi)^\perp \neq \mathbb{R}^n$. Durch Einschränkung definieren wir die Abbildung $\varphi := \Phi | Kern(\Phi)^\perp$ vermöge
$$\varphi(\mathbf{v}) := \Phi(\mathbf{v}) \text{ für } \mathbf{v} \perp Kern(\Phi).$$

Satz 2.33 (Homomorphiesatz) *Die durch $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ definierte lineare Abbildung*
$$\varphi : \; Kern(\Phi)^\perp \to Bild(\Phi)$$
ist ein Isomorphismus.

Beweis. Wir gehen aus von der orthogonalen direkten Summen-Zerlegung
$$\mathbb{R}^n = Kern(\Phi) \oplus Kern(\Phi)^\perp.$$
Für jeden Vektor $\mathbf{v} \in \mathbb{R}^n$ mit $\mathbf{v} = \mathbf{v}' + \mathbf{v}'', \mathbf{v}' \in Kern(\Phi), \mathbf{v}'' \in Kern(\Phi)^\perp$ ist
$$\Phi(\mathbf{v}) = \Phi(\mathbf{v}') + \Phi(\mathbf{v}'') = \Phi(\mathbf{v}'') = \varphi(\mathbf{v}'').$$
Daraus folgt, dass die eingeschränkte Abbildung φ surjektiv ist. Nach der Dimensionsformel (Satz 2.5 b)) sind die Dimensionen
$$dim(Kern(\Phi)^\perp) = n - dim(Kern(\Phi)) = dim(Bild(\Phi))$$
gleich. Dann muss φ ein Isomorphismus sein. □

Nach Satz 2.6 ist die Umkehrabbildung
$$\psi := \phi^{-1} : Bild(\Phi) \to Kern(\Phi)^\perp$$
linear, also auch ein Isomorphismus.

Definition 2.20 *Es sei* $P_{Bild(\Phi)} : \mathbb{R}^n \to Bild(\Phi)$ *die Orthogonalprojektion. Die lineare Abbildung*

$$P^+ := \psi \circ P_{Bild(\Phi)} : \mathbb{R}^m \to Kern(\Phi)^\perp \subset \mathbb{R}^n$$

heißt die Pseudo-Inverse *der Abbildung* Φ*.*

$$\boxed{\begin{array}{cccccccccc} \Phi: & \mathbb{R}^n & \stackrel{P}{\to} & Kern(\Phi)^\perp & \stackrel{\varphi}{\to} & Bild(\Phi) & \subset & \mathbb{R}^m & \\ & \mathbb{R}^n & \supset & Kern(\Phi)^\perp & \stackrel{\psi}{\leftarrow} & Bild(\Phi) & \stackrel{P}{\leftarrow} & \mathbb{R}^m & :\Phi^+ \end{array}}$$

Wir überlegen uns, was die Pseudo-Inverse für das lineare Ausgleichsproblem

$$A \cdot \mathbf{x} = \mathbf{b}, \quad A \in M(m \times n), \mathbf{x} \in \mathbb{R}^n, \mathbf{b} \in \mathbb{R}^m$$

bedeutet: Das LGS $A \cdot \mathbf{x} = \mathbf{b}$ ist unlösbar und wird ersetzt durch das lösbare System $A \cdot \mathbf{x} = P_{Bild(A)}\mathbf{b}$. Wenn A den Rang n hat, dann ist die Abbildung $\Phi : \mathbf{x} \mapsto A \cdot \mathbf{x}$ injektiv und $\mathbf{x}$ ist eindeutig bestimmt. I.A. ist Φ aber nicht injektiv und alle Lösungen $\mathbf{x}$ bilden einen affinen Unterraum von $\mathbb{R}^n$. Das zugehörige homogene System hat als Lösungsraum $Kern(A)$. Ist $\mathbf{x}_0$ eine spezielle Lösung, so sind alle Lösungen des inhomogenen Systems von der Form $\mathbf{x}_0 + \mathbf{x}'$ mit $\mathbf{x}' \in Kern(A) = Kern(\Phi)$. Die Pseudo-Inverse wählt darunter eine Lösung $\mathbf{x}_0 = \Phi^+(\mathbf{b}) \in Kern(A)^\perp$ aus. Für alle anderen Lösungen $\mathbf{x}_0 + \mathbf{x}', \mathbf{x}' \in Kern(A)$ ist dann nach Pythagoras

$$\| \mathbf{x}_0 + \mathbf{x}' \| = \| \mathbf{x}_0 \| + \| \mathbf{x}' \| .$$

D.h., mit der Pseudo-Inversen bekommt man die Lösung $\Phi^+\mathbf{b}$ *kleinster Länge.*

Unmittelbar aus der Definition der Pseudo-Inversen folgt

Satz 2.34 *a) Es ist* $(\Phi^+)^+ = \Phi$*.*
b) Die Abbildung $\Phi^+ : \mathbb{R}^m \to \mathbb{R}^n$ *hat die Eigenschaften*

$$\Phi^+ \circ \Phi = P_{Kern(\Phi)^\perp} : \mathbb{R}^n \to \mathbb{R}^n, \quad \Phi \circ \Phi^+ = P_{Bild(\Phi)} : \mathbb{R}^m \to \mathbb{R}^m.$$

Beispiel 2.26 **(**$Rang(\Phi) = n$**)** *Wenn* $\Phi : \mathbb{R}^n \to \mathbb{R}^m$ *den Rang* n *besitzt, dann hat nach Satz 2.28 auch* $\Phi^t \circ \Phi : \mathbb{R}^n \to \mathbb{R}^n$ *diesen Rang* n *und ist ein Automorphismus des* $\mathbb{R}^n$*. Für seine Umkehrabbildung* $(\Phi^t \circ \Phi)^{-1}$ *berechnen wir*

$$((\Phi^t \circ \Phi)^{-1} \circ \Phi^t) \circ \Phi = (\Phi^t \circ \Phi)^{-1} \circ (\Phi^t \circ \Phi) = id_{\mathbb{R}^n}.$$

Auf $Bild(\Phi)$ *ist deswegen* $(\Phi^t \circ \Phi)^{-1} \circ \Phi^t$ *die Umkehrabbildung* ψ *zu* φ*.*
Jedes $\mathbf{y} \in \mathbb{R}^m$ *schreibt sich*

$$\mathbf{y} = \mathbf{y}' + \mathbf{y}'' \textit{ mit } \mathbf{y}' \in Bild(\Phi), \mathbf{y}'' \in Bild(\Phi)^\perp = Kern(\phi^t).$$

Damit folgt für alle $\mathbf{y} \in \mathbb{R}^m$

$$\Phi^t(\mathbf{y}) = \Phi^t(\mathbf{y}') = \Phi^t(P_{Bild(\Phi)}(\mathbf{y})),$$

also $\Phi^t = \Phi^t \circ P_{Bild(\Phi)}$*.*
Damit haben wir identifiziert

$$\Phi^+ = \psi \circ P_{Bild(\Phi)} = (\Phi^t \circ \Phi)^{-1} \circ \Phi^t \circ P_{Bild(\Phi)} = (\Phi^t \circ \Phi)^{-1} \circ \Phi^t.$$

Hat Φ *die darstellende Matrix* A*, so hat* Φ^+ *die darstellende Matrix aus Definition 2.19.*

Satz 2.35 *a) Für die Pseudo-Inverse* Φ^+ *gelten*
i) $(\Phi^+\Phi)^t = \Phi^+\Phi, \quad (\Phi\Phi^+)^t = \Phi\Phi^+,$
ii) $\Phi^+\Phi\Phi^+ = \Phi^+, \quad \Phi\Phi^+\Phi = \Phi.$
b) Durch die Eigenschaften i) und ii) ist Φ^+ *eindeutig bestimmt.*

Beweis. a) i) Nach Satz 2.34 b) sind $\Phi^+\Phi$ und $\Phi\Phi^+$ Orthogonalprojektionen. Damit sind sie symmetrisch.

ii) Die Abbildung $\Phi^+\Phi$ ist die Orthogonalprojektion auf $Kern(\Phi)^\perp = Bild(\Phi^+)$. Daraus folgt

$$(\Phi^+\Phi)\Phi^+ = id_{Kern(\Phi)^\perp} \circ \Phi^+ = \Phi^+.$$

Die Abbildung $\Phi\Phi^+$ ist die Orthogonalprojektion auf $Bild(\Phi)$. Daraus folgt

$$(\Phi\Phi^+)\Phi = id_{Bild(\Phi)} \circ \Phi = \Phi.$$

b) Sei nun $\Psi : \mathbb{R}^m \to \mathbb{R}^n$ eine weitere Abbildung, welche genauso wie Φ^+ die Eigenschaften i) und ii) hat. Dann sind also

$$P := \Psi\Phi : \mathbb{R}^n \to \mathbb{R}^n \quad \text{und} \quad Q := \Phi\Psi : \mathbb{R}^m \to \mathbb{R}^m$$

symmetrisch. Außerdem folgt aus ii)

$$P^2 = (\Psi\Phi\Psi)\Phi = \Psi\Phi = P, \quad Q^2 = (\Phi\Psi\Phi)\Psi = \Phi\Psi = Q.$$

Deswegen sind

$$P : \mathbb{R}^n \to \mathbb{R}^n \quad \text{und} \quad Q : \mathbb{R}^m \to \mathbb{R}^m$$

Orthogonalprojektionen.

Offensichtlich gilt $Kern(\Phi) \subset Kern(P)$. Und umgekehrt folgt für $\mathbf{x} \in Kern(P)$ mit ii)

$$\Phi(\mathbf{x}) = \Phi\Psi\Phi(\mathbf{x}) = \Phi(P(\mathbf{x})) = \mathbf{0}.$$

Es ist also

$$Kern(P) = Kern(\Phi) \quad \text{und} \quad Bild(P) = Kern(P)^\perp = Kern(\Phi^\perp).$$

Damit ist $P = P_{Kern(\Phi)^\perp}$ als Orthogonalprojektion auf $Kern(\Phi)^\perp$ identifiziert. Es folgt

$$\Psi\Phi = \Phi^+\Phi.$$

Offensichtlich gilt $Bild(Q) \subset Bild(\Phi)$. Und umgekehrt folgt für $\mathbf{y} \in Bild(\Phi)$ mit ii)

$$\mathbf{y} = \Phi(\mathbf{x}) = \Phi\Psi\Phi(\mathbf{x}) = Q(\Phi(\mathbf{x})).$$

Es ist also $Bild(Q) = Bild(\Phi)$, und $Q = \Phi\Psi$ ist die Orthogonalprojektion $\Phi\Phi^+$ auf diesen Bildraum.

Mit ii) folgt schließlich

$$\Psi = (\Psi\Phi)\Psi = \Phi^+(\Phi\Psi) = \Phi^+\Phi\Phi^+ = \Phi^+.$$ □

Aufgabe 2.37 *Bestimmen Sie eine lineare Funktion* $p(x) = a \cdot x + b$, *welche die Punkte*

$$(0,1)^t,\ (1,1)^t,\ (2,1)^t,\ (3,2)^t,\ (4,2)^t$$

im Sinn von Beispiel 2.24 optimal interpoliert.

Aufgabe 2.38 *Bestimmen Sie eine quadratische Funktion* $q(x) = a + b \cdot x + c \cdot x^2$, $a, b, c \in \mathbb{R}$, *welche die Messpunkte aus Beispiel 2.24 optimal interpoliert. Wie in diesem Beispiel soll der Gesamtfehler angesetzt werden als*

$$\sqrt{\sum_{i=1}^{4} (q(x_i) - y_i)^2}.$$

Aufgabe 2.39 *Es sei* A *eine* $m \times n$*-Matrix,* R *eine orthogonale* $m \times m$*-Matrix und* S *eine orthogonale* $n \times n$*-Matrix. Zeigen Sie:*

$$(RAS)^+ = S^{-1} A^+ R^{-1}.$$

Aufgabe 2.40 *Die* $m \times n$*-Matrix habe den Rang* m*. Zeigen Sie:*

$$A^+ = A^t \cdot (A \cdot A^t)^{-1}.$$

Aufgabe 2.41 *Gegeben seien die Matrizen*

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}.$$

Bestimmen Sie

$$A^+,\ B^+,\ B^+ \cdot A^+ \ \text{ und } \ (A \cdot B)^+.$$

Aufgabe 2.42 *Bestimmen Sie die Pseudo-Inverse der Matrix*

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

Aufgabe 2.43 *Bestimmen Sie die Pseudo-Inversen der Matrizen*

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

und der Matrix $A \cdot B$*. Gilt für das Produkt der Pseudo-Inversen die gleiche Formel*

$$(A \cdot B)^+ = B^+ \cdot A^+$$

wie für das Produkt der Inversen?

2.5 Permutationen und Permutationsmatrizen

Definition 2.21 *Eine* Permutation von n Elementen, *z.B. der Zahlen* $1, 2, ..., n$, *ist eine bijektive Abbildung*

$$\sigma : \{1, ..., n\} \to \{1, ..., n\}.$$

Eine solche Permutation schreiben wir auch

$$\sigma = \begin{pmatrix} 1 & ... & n \\ \sigma(1) & ... & \sigma(n) \end{pmatrix}.$$

Die Menge aller Permutationen von n Elementen bezeichnen wir mit Σ_n. *Jedes* $\sigma \in \Sigma_n$ *ist eine bijektive Abbildung und besitzt daher eine Umkehrabbildung* $\sigma^{-1} \in \Sigma_n$.

Beispiel 2.27 *n=1:* $\Sigma_1 = \{id\}$,

n=2: $\Sigma_2 = \left\{ id, \quad \sigma_{1,2} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \right\}$,

n=3: $\Sigma_3 = \left\{ \begin{array}{c} id,\ \sigma_{1,2} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},\ \sigma_{1,3} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix},\ \sigma_{2,3} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \\ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \end{array} \right\}$.

Hier haben wir die Bezeichnung $\sigma_{k,l}$ *für die* Vertauschung

$$\begin{pmatrix} 1 & ... & k & ... & l & ... & n \\ 1 & ... & l & ... & k & ... & n \end{pmatrix}$$

benutzt.

Mit je zwei Permutationen $\sigma, \tau \in \Sigma_n$ gehört auch die Hintereinanderschaltung (oder das Produkt)

$$\sigma \circ \tau : \nu \mapsto \sigma(\tau(\nu))$$

wieder zu Σ_n. Es ist zu beachten, dass (wie immer)

$$(\sigma \circ \tau)^{-1} = \tau^{-1} \circ \sigma^{-1}.$$

Satz 2.36 *Die Menge* Σ_n *der Permutationen von n Zahlen enthält*

$$n! = 1 \cdot 2 \cdot 3 \cdot ... \cdot n$$

Elemente. Für fest gewähltes $\sigma \in \Sigma_n$ *ist die Abbildung*

$$\Sigma_n \ni \tau \mapsto \tau \circ \sigma \in \Sigma_n$$

bijektiv.

Beweis. Die Anzahlformel wird durch vollständige Induktion gezeigt: Die Anzahl der Elemente in Σ_1 ist $1 = 1!$ (Induktionsanfang). Nehmen wir nun $n \geq 2$ an und dass Σ_{n-1} aus $(n-1)!$ Elementen bestünde. Daraus schließen wir die Behauptung für Σ_n: Jede Permutation $\sigma \in \Sigma_n$ ist aber bestimmt durch ihren Wert $s := \sigma(n)$ (dafür gibt es n Möglichkeiten) und eine bijektive Abbildung $\{1, ..., n-1\} \to \{1, ..., n\} \setminus \{s\}$. Solche Abbildungen gibt es genauso viele, wie Σ_{n-1} Elemente enthält, nach Induktionsannahme also $(n-1)!$. Deswegen enthält die Menge Σ_n insgesamt

$$n \cdot (n-1)! = n!$$

Elemente.

Die angegebene Abbildung $\tau \mapsto \tau \circ \sigma$ ist bijektiv, weil $\tau \mapsto \tau \circ \sigma^{-1}$ deren Umkehrabbildung ist. □

Definition 2.22 *Jede Permutation $\sigma \in \Sigma_n$ definiert eine Permutationsmatrix*

$$E_\sigma = (\mathbf{e}_{\sigma(1)}, ..., \mathbf{e}_{\sigma(n)}).$$

Beispiel 2.28

Permutation σ	E_σ
$\sigma_{1,2} : 1 \leftrightarrow 2$	$\begin{pmatrix} 0 & 1 & 0 & ... & 0 \\ 1 & 0 & 0 & ... & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & ... & 1 \end{pmatrix}$
$\begin{pmatrix} 1 & 2 & 3 & ... & n \\ n & 1 & 2 & ... & n-1 \end{pmatrix}$	$\begin{pmatrix} 0 & 1 & 0 & ... & 0 \\ 0 & 0 & 1 & ... & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & ... & 1 \\ 1 & 0 & 0 & ... & 0 \end{pmatrix}$

Allgemein geht die Matrix $E_{\sigma_{k,l}}$ aus der Einheitsmatrix durch Vertauschen der k-ten mit der l-ten Spalte hervor. Das Resultat ist aber dasselbe, wie wenn man die k-te mit der l-ten Zeile vertauscht. Somit ist die Permutationsmatrix $E_{\sigma_{k,l}}$ die Elementarmatrix $E_1(k,l)$ aus Aufgabe 2.26 a).

Eine Permutationsmatrix E entsteht aus der Einheitsmatrix $\mathbb{1}$ durch Vertauschen von Spalten. Deswegen steht in jeder Spalte und in jeder Zeile von E genau eine 1. Es ist

$$E_\sigma \cdot \mathbf{e}_k = (\mathbf{e}_{\sigma(1)}, ..., \mathbf{e}_{\sigma(k)}, ..., \mathbf{e}_{\sigma(n)}) \cdot \mathbf{e}_k = \mathbf{e}_{\sigma(k)}.$$

Die zugehörige lineare Abbildung ist

$$E_\sigma : \mathbf{e}_k \mapsto \mathbf{e}_{\sigma(k)}, \quad k = 1, ..., n.$$

Zur Permutationsmatrix $E_{\sigma\circ\tau}$ gehört deswegen die lineare Abbildung

$$\mathbf{e}_k \mapsto \mathbf{e}_{\sigma\circ\tau(k)} = \mathbf{e}_{\sigma(\tau(k))} = E_\sigma \cdot E_\tau \cdot \mathbf{e}_k,$$

d.h., die Abildung zur Produktmatrix $E_\sigma \cdot E_\tau$.

Satz 2.37 *i) Für die Zuordnung*

$$\Sigma_n \to M(n \times n), \quad \sigma \mapsto E_\sigma$$

gilt

$$E_{\sigma\circ\tau} = E_\sigma \cdot E_\tau, \quad E_{id} = \mathbb{1}_n, \quad E_{\sigma^{-1}} = E_\sigma^{-1}.$$

ii) Jede Permutationsmatrix ist orthogonal.

iii) Jede Permutationsmatrix E_σ ist ein Produkt von Vertauschungsmatrizen $E_{k,l}$. Jede Permutation ist ein Produkt von Vertauschungen $\sigma_{k,l}$. Hierbei ist auch das leere Produkt erlaubt, das zur Identität gehört.

Beweis. i) Die Produktformel haben wir soeben bewiesen. Die Identität $E_{id} = \mathbb{1}_n$ ist klar. Wegen der Produktformel ist

$$E_{\sigma^{-1}} \cdot E_\sigma = E_{\sigma^{-1}\circ\sigma} = E_{id} = \mathbb{1}_n,$$

und $E_{\sigma^{-1}}$ ist die inverse Matrix zu E_σ.

ii) Die zu einer Permutationsmatrix gehörende lineare Abbildung bildet die ONB der Einheitsvektoren $\mathbf{e}_i$ wieder auf diese ONB ab. Die Behauptung folgt aus Satz 2.3.

iii) Jede Permutationsmatrix kann durch elementare Zeilenumformungen vom Typ I, d.h., Vertauschung zweier Zeilen, auf ihre Zeilenstufenform gebracht werden. Diese Zeilenstufenform ist die Einheitsmatrix. Das Resultat ist also

$$\mathbb{1}_n = E_1(k_r,l_r)\cdot\ldots\cdot E_1(k_1,l_1)\cdot E, \quad \text{bzw.} \quad E = E_1(k_r,l_r)^{-1}\cdot\ldots\cdot E_1(k_1,l_1)^{-1}\cdot\mathbb{1}_n = E_1(k_r,l_r)\cdot\ldots\cdot E_1(k_1,l_1).$$

Das beweist iii). □

Nach Satz 2.37,ii) ist die Transponierte der Permutationsmatrix E_σ die Matrix $E_\sigma^t = E_\sigma^{-1}$. Die Zeilen von E_σ sind die Spalten von $E_{\sigma^{-1}}$, also die Vektoren $\mathbf{e}^t_{\sigma^{-1}(k)}$, $k = 1, ..., n$. Wir können also auch schreiben

$$E_\sigma = \begin{pmatrix} \mathbf{e}^t_{\sigma^{-1}(1)} \\ \mathbf{e}^t_{\sigma^{-1}(2)} \\ \vdots \\ \mathbf{e}^t_{\sigma^{-1}(n)} \end{pmatrix}.$$

Definition 2.23 *Unter der* zyklischen Permutation $(i_1, i_2, ..., i_k)$, *bzw. unter dem* Zyklus $(i_1, i_2, ..., i_k)$, *versteht man diejenige Permutation, welche*

$$i_1 \mapsto i_2 \mapsto ... \mapsto i_{k-1} \mapsto i_k \mapsto i_1$$

abbildet und alle anderen $i \neq i_1, ..., i_k$ fest lässt. Hierbei müssen die k Zahlen $i_1, ..., i_k$ alle voneinander verschieden sein.

Zwei Zyklen

$$\sigma = (i_1, i_2, ..., i_k) \quad \text{und} \quad \tau = (j_1, j_2, ..., j_l)$$

heißen elementfremd, *wenn die zugehörigen Mengen*

$$\{i_1, i_2, ..., i_k\} \cap \{j_1, j_2, ..., j_l\} = \emptyset$$

disjunkt sind.

Beispiel 2.29 *Dieser Begriff des Zyklus für Permutationen ist viel eleganter, als unsere bisherige Schreibweise. Hierzu Beispiele:*

Zyklus	*bisherige Schreibweise*
(k, l)	$\sigma_{k,l}$
$(1, 2, 3)$	$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$
$(1, 3, 2)$	$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$
$(1, 2, 3, ..., n)$	$\begin{pmatrix} 1 & 2 & 3 & ... & n \\ 2 & 3 & 4 & ... & 1 \end{pmatrix}$

Allerdings ist die Schreibweise einer Permutation als Zyklus nicht eindeutig: Es ist ja

$$(i_1, i_2, i_3, ..., i_k) = (i_2, i_3, ..., i_k, i_1).$$

Aber die Verwendung dieser Notation ist sehr vorteilhaft, vor allem wegen des folgenden Satzes:

Satz 2.38 *Es seien*

$$\sigma = (i_1, i_2, ..., i_k) \quad \textit{und} \quad \tau = (j_1, j_2, ..., j_l)$$

zwei elementfremde Zyklen. Dann ist

$$\sigma \circ \tau = \tau \circ \sigma.$$

Beweis. Weil die Zyklen elementfremd sind, lässt σ alle j_λ fest und τ alle i_κ. Und ob wir nun zuerst die i_κ zyklisch vertauschen, und danach die j_λ, oder umgekehrt, das läuft deswegen auf die gleiche Permutation hinaus. □

Sehr praktisch ist auch folgende Tatsache:

Satz 2.39 *Jede Permutation $\sigma \in \Sigma_n$ ist ein Zyklus, oder ein Produkt von zwei oder mehr paarweise elementfremden Zyklen.*

Beweis. Es sei $i_1 \in \{1, ..., n\}$ die kleinste Zahl, welche von σ nicht festgelassen wird, also etwa $\sigma(i_1) = i_2$ mit $i_1 \neq i_2$. Dann kann nicht $\sigma(i_2) = i_2$ sein, denn sonst wäre σ wegen $\sigma(i_1) = \sigma(i_2)$ und $i_1 \neq i_2$ nicht injektiv. Wenn $\sigma(i_2) = i_1$ ist, haben wir unseren ersten Zyklus (i_i, i_2). Andernfalls ist $i_3 := \sigma(i_2) \neq i_1, i_2$. Wieder kann $\sigma(i_3) = i_1$ sein, und wir haben (i_1, i_2, i_3) als ersten Zyklus. Andernfalls ist $i_4 := \sigma(i_3) \neq i_1, i_2, i_3$. So machen wir weiter. Weil wir insgesamt nur n Zahlen zur Verfügung haben, muss es irgend wann eine erste Zahl i_k geben, wofür $\sigma(i_k)$ nicht von allen Zahlen $i_1, ..., i_{k-1}$ verschieden ist. Wegen der Injektivität von σ kann nicht $\sigma(i_k) = i_2, i_3, ...$ oder i_k sein. Es bleibt nur $\sigma(i_k) = i_1$ übrig, und wir haben einen Zyklus $(i_1, i_2, ..., i_k)$.

Entweder lässt σ alle anderen Zahlen i, $1 \leq i \leq n$, $i \neq i_1, ..., i_k$ fest, dann ist $\sigma = (i_1, ..., i_k)$ selbst ein Zyklus. Oder es gibt eine kleinste Zahl j_1, $j_1 \neq i_1, ..., i_k$, mit $j_2 := \sigma(j_1) \neq j_1$. Wie eben finden wir, ausgehend von j_1 einen Zyklus $j_1, ..., j_l$ mit $\sigma(j_\lambda) = j_{\lambda+1}, \lambda < l$, und $\sigma(j_l) = j_1$. Aus der Injektivität von σ folgt, dass die beiden Zyklen $(i_1, ..., i_k)$ und $(j_1, ..., j_l)$ elementfremd sind.

So machen wir weiter. Weil wir nur endlich viele Zahlen zur Verfügung haben, muss nach endlich vielen Schritten Schluss sein. Dann haben wir endlich viele Zyklen $\sigma_1, ..., \sigma_m$ konstruiert mit $\sigma = \sigma_1 \circ ... \circ \sigma_m$. □

Das Rechnen mit Permutationen in Zyklenschreibweise ist auch deswegen vorteilhaft, weil Zyklen sehr schön einfach zu multiplizieren sind. Statt der allgemeinen Aussage, hierzu

Beispiel 2.30 *Wir berechnen das Produkt* $\sigma := \sigma_1 \circ \sigma_2$ *wo*

$$\sigma_1 = (1,2,3) \quad \textit{und} \quad \sigma_2 = (2,3,4)$$

ist. Wir berechnen das Bild von 1*: Wegen* $\sigma_2(1) = 1$ *ist*

$$\sigma(1) = \sigma_1(1) = 2.$$

Wir berechnen das Bild von 2*:*

$$\sigma(2) = \sigma_1(\sigma_2(2)) = \sigma_1(3) = 1,$$

also enthält σ *den Zyklus* $(1,2)$*. Wir berechnen das Bild von* 3*:*

$$\sigma(3) = \sigma_1(\sigma_2(3)) = \sigma_1(4) = 4,$$

und das Bild von 4*:*

$$\sigma(4) = \sigma_1(\sigma_2(4)) = \sigma_1(2) = 3.$$

Das Ergebnis ist:

$$(1,2,3) \circ (2,3,4) = (1,2) \circ (3,4).$$

Unser nächstes Ziel ist die Konstruktion der sogenannten *Signum-Funktion.*

Satz 2.40 (Existenz des Signums) *Es gibt eine Abbildung* $sign : \Sigma_n \to \{\pm 1\}$ *mit den Eigenschaften*

a) $sign((k,l)) = -1$ *für jede Vertauschung* (k,l)*.*

b) $sign(\sigma \circ \tau) = sign(\sigma) \cdot sign(\tau)$ *für alle* $\sigma, \tau \in \Sigma_n$*.*

Beweis. Nur für diesen Beweis führen wir folgende Bezeichnung ein: Ein *Fehlstand* in der Permutation $\sigma \in \Sigma_n$ ist ein Paar i, j, $1 \leq i < j \leq n$, mit $\sigma(i) > \sigma(j)$. Eine Vertauschung (k,l) zum Beispiel hat die Bilder

$$(\sigma(1), ..., \sigma(n)) = (1, ..., k-1, l, \underbrace{k+1, ..., l-1}_{l-k-1}, k, l+1, ..., n).$$

Sie hat damit $2(l-k-1)+1 = 2(l-k)-1$ Fehlstände, da k,l ein Fehlstand ist, und zwei weitere durch l, bzw. k mit jedem $j, k < j < l$, entstehen.

Wir definieren die Signum-Funktion durch

$$sign(\sigma) := (-1)^f, \quad f = \text{Anzahl der Fehlstände in } \sigma.$$

Beweis von a). Die Anzahl der Fehlstände in (k,l) ist, wie wir soeben bemerkt haben, ungerade.

Beweis von b). Wir wissen, dass jede Permutation σ ein Produkt von Vertauschungen $\prod(k_\mu, l_\mu)$ ist. Wenn wir b) für den Fall beweisen können, dass $\sigma = (k,l)$ eine Vertauschung ist, folgt deshalb

$$sign(\sigma \circ \tau) = sign((k_1,l_1) \circ ... \circ (k_m,l_m) \circ \tau) = sign((k_1,l_1)) \cdot ... \cdot sign((k_m,l_m)) \cdot sign(\tau) = sign(\sigma) \cdot sign(\tau)$$

ganz allgemein. Somit genügt es, die Behauptung nur für $\sigma = (k,l)$ zu beweisen. Genauso sieht man, dass es genügt, den Fall zu behandeln, wo auch $\tau = (r,s)$ eine Vertauschung ist.

Wir nehmen also $k < l$ und $r < s$ an und müssen zeigen: Die Anzahl der Fehlstände in der Permutation $(k,l) \circ (r,s)$ ist gerade.

Wenn $l > k+1$, dann ist

$$(k,l) = \underbrace{(k,k+1)(k+1,k+2)...(l-2,l-1)(l-1,l)(l-1,l-2)...(k+1,k+2)(k,k+1)}_{\text{ungerade}}.$$

Deswegen genügt es, die Behauptung für Vertauschungen $(k,k+1)$ zu beweisen. Ebenso genügt es, die Vertauschung $(r,r+1)$ zu betrachten. Wir zählen die Fehlstände von $(k,k+1) \circ (r,r+1)$:

- Wenn $\{k, k+1\} \cap \{r, r+1\} = \emptyset$ ist, dann haben wir genau die beiden Fehlstände $k, k+1$ und $r, r+1$.
- Wenn $r = k$ ist, dann hat $(k, k+1)(r, r+1) = id$ keinen Fehlstand.
- Wenn $r = k+1$ ist, dann haben wir

  $$(k, k+1)(r, r+1) = (k, k+1)(k+1, k+2) = (k, k+1, k+2)$$

  mit den beiden Fehlständen $k, k+2$ und $k+1, k+2$.
- Wenn $r = k-1$ ist, dann haben wir

  $$(k, k+1)(k-1, k) = (k, k-1, k+1)$$

  mit den beiden Fehlständen $k-1, k$ und $k, k+1$.

Damit ist die Behauptung bewiesen. □

In Σ_3 beispielsweise gibt es die drei Vertauschungen $(1,2), (1,3)$ und $(2,3)$ mit $signum = -1$ und die drei Permutationen

σ	Anzahl der Vertauschungen	$signum$
id	0	$+1$
$(1,2,3) = (1,3)(1,2)$	2	$+1$
$(1,3,2) = (1,2)(1,3)$	2	$+1$

mit $signum = +1$.

Aufgabe 2.44 *Schreiben Sie die Permutationen*

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 5 & 4 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 6 & 4 & 1 & 2 & 3 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 6 & 5 & 8 & 7 & 2 & 1 & 4 \end{pmatrix}$$

als Produkt von Vertauschungen

Aufgabe 2.45 *Stellen Sie alle Permutationen $\sigma \in \Sigma_4$ als Zyklus oder als Produkt zyklischer Permutationen dar.*

Aufgabe 2.46 *Zeigen Sie für die zyklische Permutation $\sigma = (i_1, i_2, ..., i_k)$*

$$sign(\sigma) = (-1)^{k+1}.$$

Aufgabe 2.47 *Zeigen Sie: Für $n \geq 2$ gibt es in Σ_n genauso viele Permutationen σ mit $sign(\sigma) = 1$ wie mit $sign(\sigma) = -1$.*

2.6 Die Determinante

2.6.1 Definition

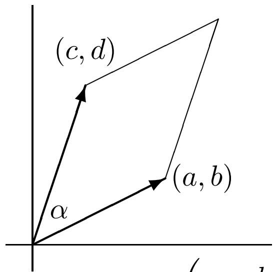
Wir betrachten eine $n \times n$-Matrix

$$A = \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{pmatrix} = \begin{pmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,n} \end{pmatrix}$$

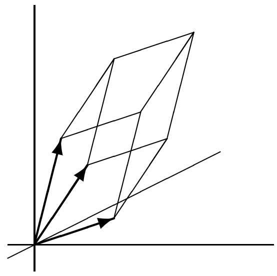
mit den Zeilenvektoren $\mathbf{a}_1, ..., \mathbf{a}_n \in \mathbb{R}^n$. Diese Zeilenvektoren spannen einen *Spat* , bzw. ein *Parallelotop*

$$P(\mathbf{a}_1, ..., \mathbf{a}_n) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \sum_1^n c_k \mathbf{a}_k, \quad c_1, ..., c_n \in \mathbb{R}, 0 \leq c_k \leq 1\}$$

auf.



$n = 2, \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$



$n = 3$

Wir möchten das *Volumen* $|A|$ dieses Spats berechnen. Der Fall $n = 2$ ist aus der Elementargeometrie bekannt: Die Fläche des Parallelogramms ist das Produkt der Seitenlängen mal $sin(\alpha)$:

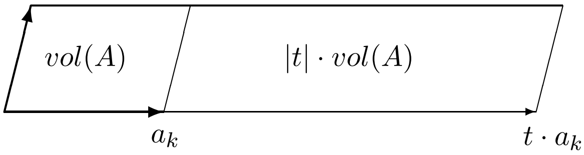
$$\begin{aligned}
\left| \begin{pmatrix} a & b \\ c & d \end{pmatrix} \right| &= \| \,(a,b)\, \| \cdot \| \,(c,d)\, \| \cdot sin(\alpha) \\
&= \| \,(a,b)\, \| \cdot \| \,(c,d)\, \| \cdot \sqrt{1 - cos^2(\alpha)} \\
&= \| \,(a,b)\, \| \cdot \| \,(c,d)\, \| \cdot \sqrt{1 - \frac{(\,(a,b)\,.\,(c,d)\,)^2}{\| \,(a,b)\, \|^2 \cdot \| \,(c,d)\, \|^2}} \\
&= \sqrt{\| \,(a,b)\, \|^2 \cdot \| \,(c,d)\, \|^2 - (\,(a,b)\,.\,(c,d)\,)^2} \\
&= \sqrt{(a^2 + b^2)(c^2 + d^2) - (ac + bd)^2} \\
&= \sqrt{a^2 d^2 + b^2 c^2 - 2 \cdot abcd} \\
&= \sqrt{(ad - bc)^2} \\
&= |ad - bc|
\end{aligned}$$

Es ist ziemlich einsichtig, dass das Volumen $|A|$ des Spats $P(\mathbf{a}_1, ..., \mathbf{a}_n)$ folgende Eigenschaften haben sollte:

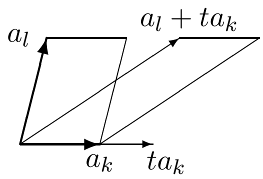
(I) Beim Vertauschen zweier Zeilen in der Matrix A ändert sich das Volumen $|A|$ nicht.

(II) Streckt man einen Zeilenvektor mit einem Faktor $t \in \mathbb{R}$, so ändert sich $|A|$ mit dem Faktor $|t|$, d.h. in Formeln

$$|\mathbf{a}_1, ..., \mathbf{a}_{k-1}, t \cdot \mathbf{a}_k, \mathbf{a}_{k+1}, ..., \mathbf{a}_n| = |t| \cdot |\mathbf{a}_1, ..., \mathbf{a}_n| \text{ für } t \in \mathbb{R}.$$



(III) $|\mathbf{a}_1, ..., \mathbf{a}_k, ..., \mathbf{a}_l + t\mathbf{a}_k, ..., \mathbf{a}_n| = |\mathbf{a}_1, ..., \mathbf{a}_k, ..., \mathbf{a}_l, ..., \mathbf{a}_n|$ für alle $1 \leq k \neq l \leq n$ und $t \in \mathbb{R}$.



(0) Für die Einheitsmatrix $\mathbb{1}$ ist

$$|\mathbb{1}| = 1.$$

Die Eigenschaften (I)-(III) beschreiben die Änderung des Volumens von $P(\mathbf{a}_1, ..., \mathbf{a}_n)$, wenn man die Vektoren elementaren Zeilentransformationen vom Typ I-III unterwirft. Wir wollen hier nicht weiter über die formulierten Eigenschaften des Volumens spekulieren, sondern eine Funktion

$$det : M(n \times n) \to \mathbb{R}$$

die *Determinante* der Matrix A, konstruieren, deren Absolutbetrag das Volumen $|A|$ ist: $|A| = |det(A)|$. Von der Funktion det verlangen wir die folgenden Eigenschaften, die natürlich bis auf das Vorzeichen mit den Eigenschaften von $|\ |$ übereinstimmen:

(I) Vertauscht man in der Matrix $A \in M(n \times n)$ zwei Zeilen, so ändert sich das Vorzeichen von $det(A)$.

(II) $det(\mathbf{a}_1, ..., \mathbf{a}_{k-1}, t \cdot \mathbf{a}_k, \mathbf{a}_{k+1}, ..., \mathbf{a}_n) = t \cdot det(\mathbf{a}_1, ..., \mathbf{a}_n)$ für alle $t \in \mathbb{R}$.

(III) $det(\mathbf{a}_1, ..., \mathbf{a}_k, ..., \mathbf{a}_l + t\mathbf{a}_k, ..., \mathbf{a}_n) = det(\mathbf{a}_1, ..., \mathbf{a}_k, ..., \mathbf{a}_l, ..., \mathbf{a}_n)$ für alle $1 \leq k \neq l \leq n$ und $t \in \mathbb{R}$.

(0) (Normierung) Für die Einheitsmatrix $\mathbb{1}$ gilt

$$det(\mathbb{1}) = 1.$$

Äquivalent können wir also det auffassen als eine Abbildung

$$det : \underbrace{\mathbb{R}^n \times ... \times \mathbb{R}^n}_{n-mal} \to \mathbb{R},$$

wobei sich $A \in M(n \times n)$ und $\mathbf{a}_1, ..., \mathbf{a}_n$ dadurch entsprechen, dass die Vektoren $\mathbf{a}_i$ die Zeilen von A sind.

Beispiel 2.31 *Die Funktion*

$$det \begin{pmatrix} a & b \\ c & d \end{pmatrix} := ad - bc$$

hat die Eigenschaften (0),(I),(II),(III). Hiervon sind (0), (I), und (II) unmittelbar einsichtig. Zum Beweis von (III) betrachten wir nur den Fall $k = 1$ und $l = 2$. Dann ist

$$det \begin{pmatrix} a & b \\ c + ta & d + tb \end{pmatrix} = a(d + tb) - b(c + ta) = ad - bc + t(ab - ba) = det \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Satz 2.41 (Eindeutigkeit der Determinante) *Wenn eine Funktion $det : M(n \times n) \to \mathbb{R}$ mit den Eigenschaften (0) bis (III) existiert, dann ist sie durch diese Eigenschaften eindeutig festgelegt. Genau dann, wenn die Matrix A einen Rang $< n$ hat, dann ist ihre Determinante*

$$det(A) = 0.$$

Beweis. Wir wissen, dass man A durch elementare Zeilenumformungen auf Zeilenstufenform bringen kann, bzw. umgekehrt, dass A durch elementare Zeilenumformungen aus einer Matrix Z in Zeilenstufenform hervorgeht. Da die Eigenschaften (I),(II),(III) genau festlegen, wie sich die Determinante bei einer elementaren Zeilenumformung ändert, genügt es, die Eindeutigkeit für Matrizen Z in Zeilenstufenform zu beweisen. Dazu unterscheiden wir die Fälle:

Rang$(A) < n$. In diesem Fall ist der letzte Zeilenvektor $\mathbf{z}_n$ in Z ein Nullvektor. Dann ist $0 \cdot \mathbf{z}_n = \mathbf{z}_n$, und aus (II) folgt

$$det(Z) = det(\mathbf{z}_1, ..., \mathbf{z}_n) = det(\mathbf{z}_1, ..., \mathbf{z}_{n-1}, 0 \cdot \mathbf{z}_n) = 0 \cdot det(\mathbf{z}_1, ..., \mathbf{z}_n) = 0.$$

Rang$(A) = n$. Nun ist Z eine Dreiecksmatrix und der letzte Zeilenvektor ist $\mathbf{z}_n = \mathbf{e}_n$. Durch Addition geeigneter Vielfacher dieses Vektors zu den vorhergehenden Zeilen (Umformung vom Typ (III)) können wir erreichen, dass der letzte Eintrag in den ersten $n - 1$ Zeilen 0 ist. Jetzt ist der vorletzte Zeilenvektor $\mathbf{z}_{n-1} = \mathbf{e}_{n-1}$, und durch elementare Zeilenumformungen vom Typ III können wir erreichen, dass auch der vorletzte Eintrag in den ersten $n - 2$ Zeilen 0 ist. Mit endlich vielen elementaren Zeilenumformungen vom Typ III, können wir also Z in die Einheitsmatrix $\mathbb{1}$ überführen (Gauß-Jordan-Verfahren). Aus Eigenschaft (III) und (0) folgt

$$det(Z) = det(\mathbb{1}) = 1.$$

□

Ein ganz anderes Problem ist es, nachzuweisen, dass eine Funktion *det* mit den Eigenschaften (0),...,(III) tatsächlich existiert. Im Wesentlichen läuft dies auf die Existenz des Signums (Satz 2.40) hinaus, denn wenn eine Determinantenfunktion $det(A)$ mit den Eigenschaften (0) und (I) existiert, dann gilt für jede Permutationsmatrix E_σ

$$det(E_\sigma) = sign(\sigma).$$

Dies ist ein Zusammenhang zwischen Determinante und *signum*-Funktion. Wir benützen die *signum*-Funktion nun für unsere Definition der Determinante:

Definition 2.24 *Es sei $A = (a_{k,l})_{k,l=1,...,n} \in M(n \times n)$ eine $n \times n$-Matrix. Die Zahl*

$$\boxed{det(A) := \sum_{\sigma \in \Sigma_n} sign(\sigma) \cdot a_{1,\sigma(1)} \cdot ... \cdot a_{n,\sigma(n)}}$$

heißt Determinante der Matrix A. *(Diese Formel für die Determinante stammt von Leibniz und ist nach ihm benannt.)*

Dass diese Determinante tatsächlich die Eigenschaften (0),...,(III) besitzt, weisen wir im nächsten Abschnitt nach. Zuerst einige einfache Beispiele, die zeigen sollen, was diese Formel bedeutet.

n=1: Im Fall $n = 1$ ist $det(a) = a$.

n=2: Für $n = 2$ ist

$$det \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} = \underbrace{sign(id) \cdot a_{1,1}a_{2,2}}_{\sigma=id} + \underbrace{sign(\sigma_{1,2})a_{1,2}a_{2,1}}_{\sigma=\sigma_{1,2}} = a_{1,1}a_{2,2} - a_{1,2}a_{2,1}.$$

Wenn wir die Matrix $\begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ schreiben, dann wird

$$\boxed{det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.}$$

n=3: Für $n = 3$ haben wir

$$\begin{array}{rclll} det \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} & = & a_{1,1}a_{2,2}a_{3,3} & \sigma = id & \\ & + & a_{1,2}a_{2,3}a_{3,1} & \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} & = \sigma_{1,3} \circ \sigma_{1,2} \\ & + & a_{1,3}a_{2,1}a_{3,2} & \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} & = \sigma_{1,2} \circ \sigma_{1,3} \\ & - & a_{1,1}a_{2,3}a_{3,2} & \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} & = \sigma_{2,3} \\ & - & a_{1,2}a_{2,1}a_{3,3} & \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} & = \sigma_{1,2} \\ & - & a_{1,3}a_{2,2}a_{3,1} & \sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} & = \sigma_{1,3} \end{array}.$$

Dies ist die klassische „Regel von Sarrus“:

$$\begin{array}{ccc|cc} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,1} & a_{3,2} \end{array} \quad - \quad \begin{array}{ccc|cc} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,1} & a_{3,2} \end{array}$$

Aufgabe 2.48 *Berechnen Sie die Determinanten der Matrizen*

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix} \quad und \quad \begin{pmatrix} 11 & 12 & 13 \\ 14 & 15 & 14 \\ 13 & 12 & 11 \end{pmatrix}.$$

2.6.2 Eigenschaften

Wir wollen jetzt einige wichtige Eigenschaften der Determinante angeben. Insbesondere suchen wir nach praktischen Möglichkeiten, die Determinante einer gegebenen Matrix zu berechnen, da die Leibnizsche Formel hierfür bei großen n ungeeignet ist.

Satz 2.42 (Fundamentaleigenschaften der Determinante) *Die Funktion* $det : M(n \times n) \to \mathbb{R}, \quad A \mapsto det(A)$, *hat folgende Eigenschaften:*

1) Linearität in Bezug auf jede Zeile:

$$det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{k-1} \\ s\mathbf{a}_k + t\mathbf{a}'_k \\ \mathbf{a}_{k+1} \\ \vdots \\ a_n \end{pmatrix} = s \cdot det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{k-1} \\ \mathbf{a}_k \\ \mathbf{a}_{k+1} \\ \vdots \\ a_n \end{pmatrix} + t \cdot det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{k-1} \\ \mathbf{a}'_k \\ \mathbf{a}_{k+1} \\ \vdots \\ \mathbf{a}_n \end{pmatrix}.$$

2) Schiefsymmetrie in Bezug auf je zwei Zeilen:

$$det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{k-1} \\ \mathbf{a}_k \\ \mathbf{a}_{k+1} \\ \vdots \\ \mathbf{a}_{l-1} \\ \mathbf{a}_l \\ \mathbf{a}_{l+1} \\ \vdots \\ \mathbf{a}_n \end{pmatrix} = -det \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_{k-1} \\ \mathbf{a}_l \\ \mathbf{a}_{k+1} \\ \vdots \\ \mathbf{a}_{l-1} \\ \mathbf{a}_k \\ \mathbf{a}_{l+1} \\ \vdots \\ \mathbf{a}_n \end{pmatrix}$$

3) Normierung: $det(\mathbb{1}_n) = 1$.

Beweis. 1) Wir werten die Determinante auf der linken Seite der Gleichung mit der Leibnizformel aus:

$$\begin{aligned} &\textstyle\sum_{\sigma \in \Sigma_n} sign(\sigma) \cdot a_{1,\sigma(1)} \cdot \ldots \cdot (s \cdot a_{k,\sigma(k)} + t \cdot a'_{k,\sigma(k)}) \cdot \ldots \cdot a_{n,\sigma(n)} = \\ &s \cdot \textstyle\sum_{\sigma \in \Sigma_n} sign(\sigma) \cdot a_{1,\sigma(1)} \cdot \ldots \cdot a_{k,\sigma(k)} \cdot \ldots \cdot a_{n,\sigma(n)} + \\ &t \cdot \textstyle\sum_{\sigma \in \Sigma_n} sign(\sigma) \cdot a_{1,\sigma(1)} \cdot \ldots \cdot a'_{k,\sigma(k)} \cdot \ldots \cdot a_{n,\sigma(n)}. \end{aligned}$$

2) Wieder mit der Leibnizformel und mit Satz 2.40 ist die Determinante auf der rechten Seite der Gleichung

$$\begin{aligned}
&\textstyle\sum_{\sigma\in\Sigma_n} sign(\sigma)\cdot a_{1,\sigma(1)}\cdot\ldots\cdot a_{l,\sigma(k)}\cdot\ldots\cdot a_{k,\sigma(l)}\cdot\ldots\cdot a_{n,\sigma(n)} =\\
&\textstyle\sum_{\sigma\in\Sigma_n} sign(\sigma)\cdot a_{1,\sigma\sigma_{k,l}(1)}\cdot\ldots\cdot a_{l,\sigma\sigma_{k,l}(l)}\cdot\ldots\cdot a_{k,\sigma\sigma_{k,l}(k)}\cdot\ldots\cdot a_{n,\sigma\sigma_{k,l}(n)} =\\
-&\textstyle\sum_{\sigma\in\Sigma_n} sign(\sigma\sigma_{k,l})\cdot a_{1,\sigma\sigma_{k,l}(1)}\cdot\ldots\cdot a_{l,\sigma\sigma_{k,l}(l)}\cdot\ldots\cdot a_{k,\sigma\sigma_{k,l}(k)}\cdot\ldots\cdot a_{n,\sigma\sigma_{k,l}(n)} =\\
-&\textstyle\sum_{\tau=\sigma\sigma_{k,l}\in\Sigma_n} sign(\tau)\cdot a_{1,\tau(1)}\cdot\ldots\cdot a_{l,\tau(l)}\cdot\ldots\cdot a_{k,\tau(k)}\cdot\ldots\cdot a_{n,\tau(n)}.
\end{aligned}$$

3) Es ist $det(\mathbb{1}_n) = \sum_\sigma sign(\sigma)\cdot\delta_{1,\sigma(1)}\cdot\ldots\cdot\delta_{n,\sigma(n)}$, und der Summand ist nur dann $\neq 0$, wenn alle Kronecker-Deltas $= 1$ sind, d.h. wenn $k = \sigma(k)$ für alle $k = 1, ..., n$. Also bleibt nur der Summand für $\sigma = id$ übrig, und die Determinante wird $= 1$. □

Die Abbildung $det : \mathbb{R}^n \times ... \times \mathbb{R}^n \to \mathbb{R}$ ist also *multilinear* in dem Sinn, dass für alle i und fest gewählte $\mathbf{a}_1, ..., \mathbf{a}_{i-1}, \mathbf{a}_{i+1}, ..., \mathbf{a}_n \in \mathbb{R}^n$ die Funktion

$$\mathbb{R}^n \to \mathbb{R}, \quad \mathbf{x} \mapsto det(\mathbf{a}_1, ..., \mathbf{a}_{i-1}, \mathbf{x}, \mathbf{a}_{i+1}, ..., \mathbf{a}_n)$$

linear ist. Deswegen ist $det : M(n\times n) \to \mathbb{R}$ i.a. nicht linear. Vielmehr folgt aus der Multilinearität

$$det(\lambda A) = \lambda^n \cdot det(A) \quad \text{für} \quad \lambda \in \mathbb{R}, A \in M(n\times n).$$

Und für $det(A + B)$ gibt es keine einfache Beziehung zu $det(A)$ und $det(B)$.

Satz 2.43 (Korollar 1 zu Satz 2.42) *Hat die $n\times n$-Matrix A zwei gleiche Zeilen, so ist $det(A) = 0$.*

Beweis. Sind die Zeilenvektoren $\mathbf{a}_k$ und $\mathbf{a}_l$ gleich, so ändert sich A und damit $det(A)$ nicht, wenn wir beide Zeilen vertauschen. Andererseits ändert sich dabei wegen der Schiefsymmetrie das Vorzeichen von $det(A)$. Es folgt:

$$det(A) = -det(A), \qquad 2\cdot det(A) = 0, \qquad det(A) = \frac{1}{2}(2\cdot det(A)) = 0.$$

(Für Mathematiker: Hier brauchen wir zum ersten Mal wirklich eine andere reelle Zahl als 0 und 1, nämlich $\frac{1}{2}$. Gäbe es diese Zahl nicht, wäre das Argument unrichtig.) □

Satz 2.44 (Korollar 2 zu Satz 2.42) *Die mit der Leibnizformel definierte Determinante hat die Eigenschaften (0),(I),(II),(III) aus Abschnitt 2.6.1.*

Beweis. Normierung (0) und Schiefsymmetrie beim Vertauschen von Zeilen (I) sind die Eigenschaften 3) und 2) von Satz 2.42. Eigenschaft (II) ist Teil der Linearität der Determinante und Eigenschaft (III) folgt aus der Linearität mit Hilfe von Korollar 1. □

Satz 2.45 (Weitere Eigenschaften der Determinante)
a) $A, B \in M(n\times n) \Rightarrow det(A\cdot B) = det(A)\cdot det(B)$ (Determinanten-Multiplikations-Satz).
b) $det(A^t) = det(A)$.

Beweis. a) Wir beweisen die Aussage zunächst für den Fall, dass $A = E$ eine Elementarmatrix ist.

Eine Elementarmatrix E vom Typ (I) entsteht aus der Einheitsmatrix durch Vertauschen zweier Zeilen. Also ist $det(E) = -det(\mathbb{1}) = -1$. Die Matrix $E\cdot B$ entsteht aus B ebenfalls durch Vertauschen zweier Zeilen. Und deswegen ist $det(E\cdot B) = -det(B) = det(E)\cdot det(B)$.

Eine Elementarmatrix E vom Typ (II) multipliziert in B eine Zeile mit einem Faktor $c \in \mathbb{R}$. Für E gilt $det(E) = c$ und wegen der Linearität der Determinante ist $det(E \cdot B) = c \cdot det(B)$.

Eine Elementarmatrix E vom Typ (III) entsteht aus der Einheitsmatrix indem man eine Vielfaches einer Zeile zu einer anderen addiert. Wegen Eigenschaft (III) der Determinante ist also $det(E) = 1$. Da weiter $det(E \cdot B) = det(B)$ ist, folgt die Behauptung auch in diesem Fall.

Wenn $Rang(A) < n$ ist, sind die Zeilen von A linear abhängig und es gibt einen Zeilenvektor $\mathbf{x} = (x_1, ..., x_n)$ mit $\mathbf{x} \cdot A = 0$. Dann ist auch $\mathbf{x} \cdot (A \cdot B) = 0$ und $Rang(A \cdot B) < n$. Mit Satz 2.41 folgt $det(A) \cdot det(B) = det(A) = 0$ und $det(A \cdot B) = 0$.

Wenn $Rang(A) = n$ ist, gibt es Elementarmatrizen $E_1, ..., E_k$ so, dass $A = E_1 \cdot ... \cdot E_k$. Es folgt

$$det(A \cdot B) = det(E_1 \cdot ... \cdot E_k \cdot B) = det(E_1) \cdot ... \cdot det(E_k) \cdot det(B) = det(A) \cdot det(B).$$

b) Mit der Leibnizformel und $sign(\sigma) = sign(\sigma^{-1})$ ist

$$\begin{aligned} det(A^t) &= \sum_{\sigma \in \Sigma_n} sign(\sigma) \cdot a_{\sigma(1),1} \cdot ... \cdot a_{\sigma(n),n} \\ &= \sum_{\sigma \in \Sigma_n} sign(\sigma) \cdot a_{1,\sigma^{-1}(1)} \cdot ... \cdot a_{n,\sigma^{-1}(n)} \\ &= \sum_{\tau = \sigma^{-1} \in \Sigma_n} sign(\tau) \cdot a_{1,\tau(1)} \cdot ... \cdot a_{n,\tau(n)} \\ &= det(A) \end{aligned}$$

Eigenschaft b) bedeutet, dass alles, was für die Zeilen einer Determinante gilt, auch für Spalten stimmt. Insbesondere ist also $det(A)$ auch linear in Bezug auf jede Spalte und ändert beim Vertauschen zweier Spalten das Vorzeichen. Die Funktion $det : \mathbb{R}^n \times ... \times \mathbb{R}^n \to \mathbb{R}$ kann also auch als Abbildung der Spalten $\mathbf{a}^1, ..., \mathbf{a}^n$ in der Matrix A aufgefasst werden, welche auch hier multilinear und schiefsymmetrisch ist.

Für orthogonale Matrizen ($A^t = A^{-1}$) folgt

$$det(A)^2 = det(A) \cdot det(A^t) = det(A \cdot A^t) = det(\mathbb{1}) = 1,$$

also $det(A) = \pm 1$.

Wir benötigen noch zwei häufig anwendbare Methoden zur Berechnung von Determinanten, die wir aber nicht extra als Satz formulieren möchten.

Kästchenregel. Die $n \times n$-Matrix A habe *Kästchenform*

$$A = \left(\begin{array}{c|c} A_1 & * \\ \hline 0 & A_2 \end{array} \right),$$

wo A_1 eine $r \times r$-Matrix und A_2 eine $(n-r) \times (n-r)$-Matrix ist. Das heißt also, das linke untere $(n-r) \times r$-Nullkästchen (oder das rechte obere $r \times (n-r)$-Kästchen) reicht bis zur Diagonale der Matrix. Wegen $det(A^t) = det(A)$ genügt es, den ersten Fall zu betrachten. In der Leibnizformel

$$det(A) = \sum_{\sigma \in \Sigma_n} sign(\sigma) \cdot a_{1,\sigma(1)} \cdot ... \cdot a_{r,\sigma(r)} \cdot a_{r+1,\sigma(r+1)} \cdot ... \cdot a_{n,\sigma(n)}$$

sind dann alle Produkte $a_{1,\sigma(1)} \cdot ... \cdot a_{r,\sigma(r)}$ =0, wo die Permutation σ eine Zahl $k, r+1 \leq k \leq n$ auf eine Zahl $\sigma(k) \leq r$ abbildet. Die Summe ist also nur über solche Permutationen zu erstrecken, welche die Teilmengen

$$\{1, ..., r\} \text{ und } \{r+1, ..., n\}$$

in sich abbilden. Diese Permutationen bestehen also aus zwei Permutationen

$$\sigma_1 : \{1, ..., r\} \to \{1, ..., r\} \in \Sigma_r, \quad \sigma_2 : \{r+1, ..., n\} \to \{r+1, ..., n\} \in \Sigma_{n-r}.$$

Schreiben wir dies in die Leibnizformel, dann wird

$$\begin{aligned} det(A) &= \sum_{\sigma_1 \in \Sigma_r, \sigma_2 \in \Sigma_{n-r}} sign(\sigma_1\sigma_2) \cdot (a_{1,\sigma_1(1)} \cdot ... \cdot a_{r,\sigma_1(r)}) \cdot (a_{r+1,\sigma_2(r+1)} \cdot ... \cdot a_{n,\sigma(n)}) \\ &= \left(\sum_{\sigma_1 \in \Sigma_r} sign(\sigma_1) \cdot a_{1,\sigma_1(1)} \cdot ... \cdot a_{r,\sigma_1(r)} \right) \cdot \left(\sum_{\sigma_2 \in \Sigma_{n-r}} sign(\sigma_2) \cdot a_{r+1,\sigma_2(r+1)} \cdot ... \cdot a_{n,\sigma_2(n)} \right) \\ &= det(A_1) \cdot det(A_2). \end{aligned}$$

Diese Kästchenregel kann man iterieren und findet für eine obere (oder untere) Dreiecksmatrix A, dass

$$det(A) = a_{1,1} \cdot a_{2,2} \cdot ... \cdot a_{n,n}$$

das Produkt der Diagonalelemente ist. Dies führt auch auf die effektivste Methode, eine Determinante zu berechnen: Man überführt die Matrix durch elementare Zeilenumformungen in obere Dreiecksform, merkt sich die dabei auftretenden Änderungen der Determinante, berechnet die Determinante der Dreiecksmatrix, und macht die Veränderungen der Determinante wieder rückgängig.

Entwicklung nach Spalten oder Zeilen. Wir schreiben den ersten Zeilenvektor $\mathbf{a}_1$ unserer Matrix A als

$$(a_{1,1}, ..., a_{1,k}, ..., a_{1,n}) = (a_{1,1}, 0, ..., 0) + ... + (0, ..., 0, a_{1,k}, 0, ..., 0) + ... + (0, ..., 0, a_{1,n})$$

und wenden die Linearität der Determinante auf die erste Zeile an:

$$\begin{aligned} det(A) &= det \left(\begin{array}{c|ccc} a_{1,1} & 0 & ... & 0 \\ \hline \vdots & & A_{1,1} & \end{array} \right) \\ &\vdots \\ &+ det \left(\begin{array}{ccc|c|ccc} 0 & ... & 0 & a_{1,k} & 0 & ... & 0 \\ \hline & A'_{1,k} & & \vdots & & A''_{1,k} & \end{array} \right) \\ &\vdots \\ &+ det \left(\begin{array}{ccc|c} 0 & ... & 0 & a_{1,n} \\ \hline & A_{1,n} & & \vdots \end{array} \right). \end{aligned}$$

Hier bezeichnen wir mit $A_{k,l}$ die *Streichungsmatrix* von A zur Stelle (k, l), d.h. die $(n-1) \times (n-1)$-Matrix, welche aus der $n \times n$-Matrix A entsteht, indem man die k-te Zeile und die l-te Spalte streicht. Die erste Determinante auf der rechten Seite hat Kästchenform und berechnet sich zu

$$det \left(\begin{array}{c|c} a_{1,1} & 0 \\ \hline * & A_{1,1} \end{array} \right) = a_{1,1} \cdot det(A_{1,1}).$$

Die anderen Matrizen können auch auf diese Kästchenform gebracht werden. Und zwar müssen wir dazu die k-te Spalte mit der $(k-1)$-ten Spalte vertauschen, dann mit der $(k-2)$-ten usw. Insgesamt ergeben sich dabei $k-1$ Änderungen des Vorzeichens.

$$det \left(\begin{array}{c|c|c} 0 & a_{1,k} & 0 \\ \hline A^l_{1,k} & \cdot & A^r_{1,k} \end{array} \right) = (-1)^{1+k} det \left(\begin{array}{c|c} a_{1,k} & 0 \\ \cdot & A_{1,k} \end{array} \right) = (-1)^{1+k} a_{1,k} \cdot det(A_{1,k}).$$

Damit haben wir die *Entwicklung von $det(A)$ nach der ersten Zeile:*

$$det(A) = \sum_{k=1}^{n} (-1)^{1+k} \cdot a_{1,k} \cdot det(A_{1,k}).$$

Ebenso kann man nach einer anderen (etwa der l-ten) Zeile entwickeln, wenn man diese erst durch $l-1$ Vertauschungen nach oben bringt. Und genauso, wie man nach einer Zeile entwickeln kann, kann man die Determinante nach einer Spalte entwickeln.

$$\begin{array}{ll} \text{Entwicklung nach der } l\text{-ten Zeile:} & det(A) = \sum_{k=1}^{n} (-1)^{k+l} \cdot a_{l,k} \cdot det(A_{l,k}) \\ \text{Entwicklung nach der } k\text{-ten Spalte:} & det(A) = \sum_{l=1}^{n} (-1)^{k+l} \cdot a_{l,k} \cdot det(A_{l,k}) \end{array}$$

Diese Formeln können in speziellen Fällen, wo die Entwicklungszeile/Spalte viele Nullen enthält, sehr nützlich sein. I.a. ist der Rechenaufwand aber nicht wesentlich geringer als bei der Leibniz-Formel. Meist ist die Berechnung mittels Überführung in Dreiecksform vorzuziehen.

Adjunkten und die Inverse Matrix. Mit Hilfe der Determinante lassen sich 'explizite' Formeln für A^{-1} und $A^{-1} \cdot \mathbf{b}$ angeben, welche für theoretische Zwecke, nicht aber zur Berechnung günstig sind.

Die Streichungsdeterminanten $det(A_{l,k})$ kann man zu einer $n \times n$-Matrix zusammenfassen. Transponiert und mit Vorzeichen versehen heißen diese Determinanten die *Adjunkten* von A, und die Matrix

$$A^{adj} = ((-1)^{l+k} det(A_{l,k}))^t$$

heißt die Matrix der Adjunkten. Diese Matrix wurde transponiert, damit das Produkt

$$A \cdot A^{adj} = (a_{\mu,\nu})_{\substack{\mu:\text{Zeile} \\ \nu:\text{Spalte}}} \cdot ((-1)^{k+l} det(A_{k,l}))_{\substack{l:\text{Zeile} \\ k:\text{Spalte}}} = (\sum_{\nu=1}^{n} a_{\mu,\nu} (-1)^{k+\nu} det(A_{k,\nu}))_{\mu,k}$$

leicht auszurechnen ist: Die Entwicklung nach Zeilen hat zur Folge, dass alle Diagonaleinträge

$$(A \cdot A^{adj})_{l,l} = \sum_{k=1}^{n} (-1)^{k+l} \cdot a_{l,k} \cdot det(A_{l,k}) = det(A)$$

sind. Und die Nicht-Diagonaleinträge $(l_1 \neq l_2)$

$$\sum_{k=1}^{n} (-1)^{k+l_2} a_{l_1,k} det(A_{l_2,k})$$

kann man interpretieren als Entwicklung nach der l_2-ten Zeile für die Determinante derjenigen Matrix, welche aus A entsteht, indem die l_2-te Zeile durch die l_1-te Zeile ersetzt worden ist. Diese Matrix hat zwei gleiche Zeilen, ihre Determinante ist $= 0$, und

$$(A \cdot A^{adj})_{l_1,l_2} = det(A) \cdot \delta_{l_1,l_2}.$$

Damit haben wir das Matrixprodukt

$$\boxed{A \cdot A^{adj} = det(A) \cdot \mathbb{1}_n}$$

berechnet. Wenn $det(A) \neq 0$ ist, erhalten wir daraus die Formel

$$\boxed{A^{-1} = \frac{1}{det(A)} A^{adj}}$$

für die inverse Matrix A^{-1}.

Cramersche Regel. Ist die Matrix A quadratisch (d.h. eine $n \times n$-Matrix), und ist A invertierbar, so kann man das Gleichungssystem $A \cdot \mathbf{x} = \mathbf{b}$ ganz einfach so

$$\mathbf{x} = A^{-1} \cdot \mathbf{b}$$

lösen. Leider ist dies aber keine Zauberformel, sondern zeigt nur, dass es genauso schwer ist, die inverse Matrix zu berechnen, wie das Gleichungssystem zu lösen. Trotzdem wollen wir unsere Formel für A^{-1} hierauf anwenden.

Die Lösung wird also

$$\mathbf{x} = \frac{1}{det(A)} \cdot A^{adj} \cdot \mathbf{b}.$$

Die k-te Komponente des Lösungsektors $\mathbf{x}$ ist dann

$$x_k \;=\; \frac{1}{det(A)} \cdot \sum_{l=1}^{n} (A^{adj})_{k,l} \cdot b_l \;=\; \frac{1}{det(A)} \cdot \sum_{l=1}^{n} (-1)^{k+l} \cdot det(A_{l,k}) \cdot b_l.$$

Die Summe kann interpretiert werden als die Entwicklung der modifizierten Koeffizientenmatrix

$$A^{(k)} := \begin{pmatrix} a_{1,1} & . & . & . & a_{1,k-1} & b_1 & a_{1,k+1} & . & . & . & a_{1,n} \\ \vdots & & & & \vdots & \vdots & \vdots & & & & \\ a_{n,1} & . & . & . & a_{n,k-1} & b_n & a_{n,k+1} & . & . & . & a_{n,n} \end{pmatrix},$$

nach der k-ten Spalte, wobei diese in A durch die rechte Seite $\mathbf{b}$ ersetzt worden ist. Mit dieser Matrix $A^{(k)}$ erhält man also die Lösung $\mathbf{x} = (x_1, ..., x_n)$ in der Form

$$\boxed{x_k = \frac{det(A^{(k)})}{det(A)}.}$$

Dies ist die *Cramersche Regel* zum Lösen linearer Gleichungssysteme mit *quadratischer und invertierbarer* Koeffizientenmatrix.

Aufgabe 2.49 *Berechnen Sie die Determinante der Matrix*

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

Aufgabe 2.50 *Für $A \in M(n \times n)$ zeige man*

$$det(A) = 0 \quad \Longleftrightarrow \quad \exists B \in M(n \times n) \setminus \{0\} \textit{ mit } AB = 0.$$

Aufgabe 2.51 *Berechnen Sie*

$$det\begin{pmatrix} 2 & 3 & 4 & 2 \\ 1 & 3 & 5 & -2 \\ 3 & 7 & 9 & 2 \\ 2 & 5 & 12 & -9 \end{pmatrix}.$$

Aufgabe 2.52 *Es seien $a_1, ..., a_n \in \mathbb{R}$. Beweisen Sie*

$$det\begin{pmatrix} 1 & 1 & ... & 1 \\ a_1 & a_2 & ... & a_n \\ a_1^2 & a_2^2 & ... & a_n^2 \\ \vdots & \vdots & & \vdots \\ a_1^{n-1} & a_2^{n-1} & ... & a_n^{n-1} \end{pmatrix} = \prod_{1 \leq i < j \leq n} (a_j - a_i).$$

(„Vandermondesche Determinante")

Aufgabe 2.53 *Es seien $a_1, ..., a_n, b_1, ..., b_n \in \mathbb{R}$. Berechnen Sie*

$$det\begin{pmatrix} 1 & 1 & 1 & ... & 1 \\ b_1 & a_1 & a_1 & ... & a_1 \\ b_1 & b_2 & a_2 & ... & a_2 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ b_1 & b_2 & ... & b_n & a_n \end{pmatrix}.$$

Aufgabe 2.54 *Es seien $a, b, c \in \mathbb{R}$. Berechnen Sie die Determinanten der Matrizen*

$$\begin{pmatrix} bc & a^2 & a^2 \\ b^2 & ac & b^2 \\ c^2 & c^2 & ab \end{pmatrix}, \qquad \begin{pmatrix} bc & ab & ac \\ ab & ac & bc \\ ac & bc & ab \end{pmatrix},$$

$$\begin{pmatrix} a^2 & ab & ab & b^2 \\ ab & ac & b^2 & bc \\ ab & b^2 & ac & bc \\ b^2 & bc & bc & c^2 \end{pmatrix}, \qquad \begin{pmatrix} a & b & b & b \\ b & a & b & b \\ b & b & a & b \\ b & b & b & a \end{pmatrix}.$$

Aufgabe 2.55 *Es seien $x_1, x_2, x_3, y_1, y_2, y_3 \in \mathbb{R}$. Berechnen Sie*

$$det\begin{pmatrix} 0 & x_1 & x_2 & x_3 \\ -x_1 & 0 & y_3 & -y_2 \\ -x_2 & -y_3 & 0 & y_1 \\ -x_3 & y_2 & -y_1 & 0 \end{pmatrix}.$$

Aufgabe 2.56 *Sei A eine relle $n \times n$-Matrix, deren Koeffizienten außerhalb der Diagonale alle 1 sind und deren Diagonalkoeffizienten alle Null sind. Man zeige $det(A) = (-1)^{n-1} \cdot (n-1)$.*

Aufgabe 2.57 *Für reelle Zahlen $\lambda_1, \lambda_2, ..., \lambda_n$ berechne man die Determinante*

$$\begin{vmatrix} 1-\lambda_1 & -\lambda_2 & -\lambda_3 & ... & -\lambda_n \\ -\lambda_1 & 1-\lambda_2 & -\lambda_3 & ... & -\lambda_n \\ -\lambda_1 & -\lambda_2 & 1-\lambda_3 & ... & -\lambda_n \\ \vdots & \vdots & & \ddots & \\ -\lambda_1 & -\lambda_2 & -\lambda_3 & ... & 1-\lambda_n \end{vmatrix}.$$

Aufgabe 2.58 *Es sei* $A = (a_{ij})_{1 \le i,j \le n} \in M(n \times n)$ *mit* $a_{ij} = (-1)^i \cdot i$ *für* $i + j > n$ *und* $a_{ij} = 0$ *sonst, also z.B.*

$$A_1 = -1, \quad A_2 = \begin{pmatrix} 0 & -1 \\ 2 & 2 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 2 & 2 \\ -3 & -3 & -3 \end{pmatrix}, \quad A_4 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 2 & 2 \\ 0 & -3 & -3 & -3 \\ 4 & 4 & 4 & 4 \end{pmatrix}.$$

Man berechne $det(A_n)$ *für beliebiges* n.

Aufgabe 2.59 *Seien* A, B, C, D *reelle* $n \times n$*-Matrizen und* $X = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ *die durch sie in Blockschreibweise gegebene* $2n \times 2n$*-Matrix. Gelte* $AC = CA$*. Man zeige:* $det(X) = det(AD - CB)$*, wenn* $det(A) \neq 0$.

Aufgabe 2.60 *Es seien* $a, a_1, ..., a_n \in \mathbb{R}$*. Beweisen Sie*

$$det \begin{pmatrix} a & 0 & ... & 0 & a_1 \\ 0 & a & & 0 & a_2 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & a & a_n \\ a_1 & a_2 & ... & a_n & a \end{pmatrix} = a^{n+1} - a^{n-1}(a_1^2 + ... + a_n^2).$$

Aufgabe 2.61 *Zeigen Sie:*
a) Der Rang der Matrix

$$A = \begin{pmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & a & 0 \\ 0 & c & 0 & a \end{pmatrix}$$

ist stets eine gerade Zahl.
b) A ist genau dann invertierbar, wenn $a^2 \neq bc$ *ist.*
c) Geben Sie in diesem Fall A^{-1} *an.*

Aufgabe 2.62 *a) Für* $n \in \mathbb{N}$ *wird die* $n \times n$*-Matrix* A_n *definiert durch*

$$A_n := \begin{pmatrix} -2 & 1 & 0 & ... & 0 \\ 1 & -2 & 1 & \ddots & \vdots \\ 0 & 1 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & -2 & 1 \\ 0 & ... & 0 & 1 & -2 \end{pmatrix}.$$

Berechnen Sie $det(A_n)$ *für alle* $n \in \mathbb{N}$.

b) Für $n \geq 4$ wird die $n \times n$-Matrix D_n definiert durch

$$D_n := \begin{pmatrix} -2 & 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & -2 & 1 & 0 & & & \vdots \\ 1 & 1 & -2 & 1 & \ddots & & \vdots \\ 0 & 0 & 1 & -2 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \ddots & -2 & 1 \\ 0 & \dots & \dots & \dots & 0 & 1 & -2 \end{pmatrix}.$$

Berechnen Sie $det(D_n)$ für alle $n \geq 4$.
c) Berechnen Sie die Determinanten der Matrizen

$$E_6 := \begin{pmatrix} -2 & 1 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & 0 & 0 & -2 \end{pmatrix}, \quad E_7 := \begin{pmatrix} -2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -2 \end{pmatrix}$$

und

$$E_8 := \begin{pmatrix} -2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & -2 \end{pmatrix}.$$

2.6.3 Orientierung und Determinante

Der uns umgebende Raum hat eine Orientierung. Wie jeder weiß wird die im Spiegel geändert, (das ist richtig) weil der Spiegel die rechte und die linke Hand vertauscht (das weiß jeder, es ist aber falsch). Trotzdem: Es gibt zwei Orientierungen, welch beim Spiegeln vertauscht werden aber bei Drehungen nicht. Nur, was ist das: Eine Orientierung?

Erinnern wir uns an Drehungen und Spiegelungen in der Ebene $\mathbb{R}^2$:

	Drehung um		Spieglung an	
	0^0	180^0	$x-$ Achse	$y-$ Achse
	$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$
Determinante	1	1	-1	-1

Die zugehörigen Matrizen unterscheiden sich um das Vorzeichen ihrer Determinante. Natürlich haben nur invertierbare Matrizen eine Determinante $\neq 0$ und damit eine Determinante mit Vorzeichen. In

Verallgemeinerung der Spiegelungen in der Ebene definieren wir also: Eine lineare Abbildung $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ ändert die Orientierung des Raums $\mathbb{R}^n$, wenn *ihre Determinante negativ ist.* Damit wissen wir, wann sich die Orientierung ändert. Aber wir wissen immer noch nicht, was das ist: Eine Orientierung?

Definition 2.25 *Eine Orientierung des Vektorraums $\mathbb{R}^n$ wird definiert durch eine Basis $\mathbf{a}_1, ..., \mathbf{a}_n \in \mathbb{R}^n$. Zwei Basen $\mathbf{a}_1, ..., \mathbf{a}_n$ und $\mathbf{b}_1, ..., \mathbf{b}_n$ definieren die gleiche Orientierung, wenn beide $n \times n$-Matrizen*

$$(\mathbf{a}_1, ..., \mathbf{a}_n) \quad \textit{und} \quad (\mathbf{b}_1, ..., \mathbf{b}_n)$$

Determinanten mit dem gleichen Vorzeichen haben.

Ziemlich einleuchtend! Hat die $n \times n$-Matrix A eine Determinante $det(A) > 0$, so definiert die Basis $\mathbf{a}_1, ..., \mathbf{a}_n$ die gleiche Orientierung des $\mathbb{R}^n$ wie die Basis $A \cdot \mathbf{a}_1, ..., A \cdot \mathbf{a}_n$. Wenn $det(a) < 0$ ist, so ändert definiert sie die andere Orientierung. Aber was ist das eigentlich: eine Orientierung des $\mathbb{R}^n$?

Definition 2.26 *Eine Orientierung des $\mathbb{R}^n$ ist eine Menge von Basen $\mathbf{a}_1, ..., \mathbf{a}_n$ des $\mathbb{R}^n$, und zwar die Menge aller Basen mit demselben Vorzeichen von $det(\mathbf{a}_1, ..., \mathbf{a}_n)$.*

Es gibt also genau zwei Orientierungen des $\mathbb{R}^n$, weil Determinanten invertierbarer Matrizen zwei Vorzeichen haben können. Aber es ist doch angsteinflößend, sich so etwas Grundlegendes wie eine Orientierung vorstellen zu müssen als eine furchtbar große Menge von Basen. Deswegen sagt man auch nicht 'eine Menge von Basen', sondern 'eine Äquivalenzklasse von Basen'. Das beruhigt etwas. Dann hat man nicht mehr unendlich viele Objekte, sondern nur noch zwei: Die Äquivalenzklasse der Basen $\mathbf{a}_1, ..., \mathbf{a}_n$ mit $det(\mathbf{a}_1, ..., \mathbf{a}_n) > 0$ und die der Basen mit $det(\mathbf{a}_1, ..., \mathbf{a}_n) < 0$.

Beispiel 2.32 ($n = 1$) *Die zwei Orientierungen der Geraden $\mathbb{R}^1$ sind genau die beiden Richtungen, in der man sie durchlaufen kann.*

Beispiel 2.33 ($n = 2$) *Im $\mathbb{R}^2$ gibt es die mathematisch positive Orientierung, definiert durch die Basis $\mathbf{e}_1, \mathbf{e}_2$ und die mathematisch negative Orientierung, definiert durch die Basis $\mathbf{e}_1, -\mathbf{e}_2$. Die unterscheiden sich nur dadurch, ob man von oben oder von unten auf das Papier schaut. (Letzteres ist schwieriger.) Wieso Peter Henlein seine Taschenuhr in die mathematisch negative Richtung laufen ließ, das weiß ich nicht. Wieso man den Vektor $\mathbf{e}_2$ in der Zeichenebene nach oben anträgt und nicht nach unten, das weiß ich übrigens genauso wenig.*

Beispiel 2.34 ($n = 3$) *Die beiden Orientierungen des $\mathbb{R}^3$ kann man an den Fingern ablesen. Zeigt der Daumen der rechten Hand nach rechts, der Zeigefinger nach vorne, so zeigt der Mittelfinger nach oben. Das ist näherungsweise die Position der Vektoren $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3 \in \mathbb{R}^3$ (wenn man sie sich konventionell vorstellt). Dies definiert die positive Orientierung des $\mathbb{R}^3$ und wird unter* Rechte-Hand-Regel *verstanden. Zeigt der Daumen der linken Hand nach rechts (geht mit etwas Übung), deren Zeigefinger nach vorne, so zeigt ihr Mittelfinger nach unten. Das definiert die andere Orientierung.*

Eine Orientierung eines endlich-dimensionalen $\mathbb{R}$-Vektorraums kann man genauso definieren als eine Äquivalenzklasse von Basen. Und zwar gehören zwei Basen $\mathbf{a}_1, ..., \mathbf{a}_n$ und $\mathbf{b}_1, ..., \mathbf{b}_n$ zur gleichen Äquivalenzklasse, wenn die Koordinatenvektoren der $\mathbf{b}_\nu$ bezüglich der Basis $\{ veca_\nu\}$ eine $n \times n$-Matrix mit positiver Determinante bilden. Oder Äquivalent dazu:

$$det\left((\mathbf{b}_1, ..., \mathbf{b}_n) \cdot (\mathbf{a}_1, ..., \mathbf{a}_n)\right) > 0.$$

Eine Orientierung des $\mathbb{R}^n$ hat keinerlei Einfluss auf Orientierung eines Untervektorraums. Ist eine Orientierung der Ebene $\mathbb{R}^2$ gewählt, so kann man eine Gerade in dieser Ebene ruhig in jeder ihrer beiden Richtungen durchlaufen. Psychologisch schwierig ist das nur bei den Koordinatenachsen, wo die Einheitsvektoren drauf liegen. Aber da muss man sich eben daran erinnern, dass die gleiche Orientierung des $\mathbb{R}^2$ auch durch jede Basis definiert ist, welche nicht aus den Einheitsvektoren besteht.

Anders ist dies bei Hyperebenen. Eine Hyperebene $H \subset \mathbb{R}^n$ ist ein Untervektorraum der Dimension $n-1$. Eine Orientierung von H wird definiert durch eine Basis $\mathbf{a}_1, ..., \mathbf{a}_{n-1}$ von H. Durch jeden Vektor $\mathbf{a}_n \in \mathbb{R}^n, \mathbf{a}_n \notin H$ kann man sie zu einer Basis des $\mathbb{R}^n$ ergänzen. Ist eine Orientierung des $\mathbb{R}^n$ vorgegeben, so kann die Basis $\mathbf{a}_1, ..., \mathbf{a}_{n-1}, \mathbf{a}_n$ diese Orientierung repräsetieren. Oder auch nicht. In diesem Fall ist $\mathbf{a}_1, ..., \mathbf{a}_{n-1}, -\mathbf{a}_n$ eine Basis mit

$$det(\mathbf{a}_1, ..., \mathbf{a}_{n-1}, -\mathbf{a}_n) = -det(\mathbf{a}_1, ..., \mathbf{a}_{n-1}, \mathbf{a}_n),$$

welche die vorgegebene Orientierung des $\mathbb{R}^n$ definiert. Wir sehen:

Es sei V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum und $H \subset V$ eine Hyperebene. Ist eine Orientierung von V und ein Vektoer $\mathbf{v} \in V, \mathbf{v} \notin H$ gegeben, so wird dadurch eine Orientierung von H gegeben. Und zwar ist diese Orientierung von H definiert durch jede Basis $\mathbf{a}_1, ..., \mathbf{a}_{n-1}$ von H derart, dass die Basis $\mathbf{a}_1, ..., \mathbf{a}_{n-1}, \mathbf{v}$ die vorgegebene Orientierung von V repräsentiert.

2.7 Die LR-Zerlegung

Eine $n \times n$-Matrix $A = (a_{i,j})$ heißt *obere Dreiecksmatrix*, wenn $a_{i,j} = 0$ für $i > j$:

$$\begin{pmatrix} a_{1,1} & * & * & * \\ 0 & \ddots & \ddots & * \\ \vdots & \ddots & \ddots & * \\ 0 & ... & 0 & a_{n,n} \end{pmatrix}$$

Das Produkt $C = A \cdot B$ zweier oberer Dreiecksmatrizen ist wieder eine obere Dreiecksmatrix mit den Diagonalelementen $c_{i,i} = a_{i,i} \cdot b_{i,i}$.

Beweis. Nach Voraussetzung ist $a_{i,j} = 0$ für $i > j$ und $b_{k,l} = 0$ für $k > l$. Für $i > l$ ist deswegen

$$c_{i,l} = \sum_{j=1}^{n} a_{i,j} b_{j,l} = \sum_{j=1}^{i-1} \underbrace{a_{i,j}}_{=0} b_{j,l} + \sum_{j=i}^{n} a_{i,j} \underbrace{b_{j,l}}_{=0} = 0.$$

Und für $i = l$ haben wir

$$c_{i,i} = \sum_{j=1}^{i-1} \underbrace{a_{i,j}}_{=0} b_{j,i} + a_{i,i} b_{i,i} + \sum_{j=i+1}^{n} a_{i,j} \underbrace{b_{j,i}}_{=0} = a_{i,i} b_{i,i}.$$

□

Wegen $A^t \cdot B^t = (B \cdot A)^t$ gilt die analoge Aussage natürlich auch für untere Dreiecksmatrizen.

Definition 2.27 *Eine (untere oder obere) Dreiecksmatrix A heißt* normiert, *wenn ihre Diagonalelemente $a_{i,i} = 1$ sind für alle i.*

Eine normierte untere Dreicksmatrix $M = (m_{i.j})$ heißt Frobenius-Matrix, *wenn es ein k gibt mit $m_{i,j} = 0$ für $j \neq k$.*

Eine Frobenius-Matrix sieht also so aus:

$$\begin{pmatrix} 1 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & & & & \vdots \\ \vdots & \ddots & 1 & 0 & & & & \vdots \\ \vdots & & 0 & 1 & 0 & & & \vdots \\ \vdots & & \vdots & m_{k+1,k} & 1 & \ddots & & \vdots \\ \vdots & & \vdots & m_{k+2,k} & 0 & \ddots & \ddots & \vdots \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & m_{n,k} & 0 & \dots & 0 & 1 \end{pmatrix}.$$

Man kann sie kompakt schreiben als

$$M = \mathbb{1}_n + \mathbf{m} \otimes \mathbf{e}_i \quad \text{mit} \quad \mathbf{m} = (0, ..., 0, m_{k+1,k}, ..., m_{n,k})^t.$$

Satz 2.46 *Das Produkt zweier Frobenius-Matrizen*

$$M_1 = \mathbb{1}_n + \mathbf{m}_1 \otimes \mathbf{e}_i, \quad M_2 = \mathbb{1}_n + \mathbf{m}_2 \otimes \mathbf{e}_j$$

mit $i \leq j$ ist

$$M_1 \cdot M_2 = \mathbb{1}_n + \mathbf{m}_1 \otimes \mathbf{e}_i + \mathbf{m}_2 \otimes \mathbf{e}_j.$$

Beweis. Wir berechnen

$$M_1 \cdot M_2 = \mathbb{1}_n + \mathbf{m}_1 \otimes \mathbf{e}_i + \mathbf{m}_2 \otimes \mathbf{e}_j + \mathbf{m}_1 \otimes \mathbf{e}_i \cdot \mathbf{m}_2 \otimes \mathbf{e}_j.$$

Hier ist

$$\mathbf{m}_1 \otimes \mathbf{e}_i \cdot \mathbf{m}_2 \otimes \mathbf{e}_j = (\mathbf{e}_i.\mathbf{m}_2) \cdot \mathbf{m}_1 \otimes \mathbf{e}_j.$$

Wegen $\mathbf{m}_{k,2} = 0$ für $k \leq j$ ist dabei $(\mathbf{e}_i.\mathbf{m}_2) = 0$. □

Satz 2.47 (Korollar) *Jede normierte untere Dreiecksmatrix ist ein Produkt*

$$\mathbb{1}_n + \sum_{i=1}^{n-1} \mathbf{m}_i \otimes \mathbf{e}_i = (\mathbb{1}_n + \mathbf{m}_1 \otimes \mathbf{e}_1) \cdot (\mathbb{1}_n + \mathbf{m}_2 \otimes \mathbf{e}_2) \cdot ... \cdot (\mathbb{1}_n + \mathbf{m}_{n-1} \otimes \mathbf{e}_{n-1}).$$

Beweis (Induktion). Als normierte untere Dreiecksmatrix hat L eine Form

$$L = \mathbb{1}_n + \mathbf{m}_1 \otimes \mathbf{e}_1 + \mathbf{m}_2 \otimes \mathbf{e}_2 + ... + \mathbf{m}_{n-1} \otimes \mathbf{e}_{n-1}.$$

Für $k < n$ beweisen wir die Formel

$$\prod_{\nu=1}^{k} (\mathbb{1}_n + \mathbf{m}_\nu \otimes \mathbf{e}_\nu) = \mathbb{1}_n + \sum_{\nu=1}^{k} \mathbf{m}_\nu \otimes \mathbf{e}_\nu$$

durch Induktion nach k. Um von k auf $k+1$ zu schließen, multiplizieren wir

$$\left(\mathbb{1}_n + \sum_{\nu=1}^{k} \mathbf{m}_\nu \otimes \mathbf{e}_\nu\right) \cdot (\mathbb{1}_n + \mathbf{m}_{k+1} \otimes \mathbf{e}_{k+1}) = \mathbb{1}_n + \sum_{\nu=1}^{k} \mathbf{m}_\nu \otimes \mathbf{e}_\nu + \mathbf{m}_{k+1} \otimes \mathbf{e}_k + \sum_{\nu=1}^{k} \mathbf{m}_\nu \otimes \underbrace{\mathbf{e}_\nu \cdot \mathbf{m}_{k+1}}_{=0} \otimes \mathbf{e}_{k+1}.$$

Satz 2.48 *a) Es sei* $\mathbf{x} \in \mathbb{R}^n$ *mit* $x_i = 0$*. Dann ist*

$$(\mathbb{1}_n - \mathbf{x} \otimes \mathbf{e}_i)^{-1} = \mathbb{1}_n + \mathbf{x} \otimes \mathbf{e}_i.$$

Insbesondere ist die Inverse der Frobeniusmatrix $L_i = \mathbb{1}_n - \mathbf{m}_i \otimes \mathbf{e}_i$ *die Matrix*

$$L_i^{-1} = (\mathbb{1}_n - \mathbf{m}_i \otimes \mathbf{e}_i)^{-1} = \mathbb{1}_n + \mathbf{m}_i \otimes \mathbf{e}_i.$$

b) Die Inverse einer normierten unteren Dreicksmatrix ist wieder eine solche normierte untere Dreiecksmatrix.

Beweis. a) Wir multiplizieren

$$(\mathbb{1}_n + \mathbf{x} \otimes \mathbf{e}_i) \cdot (\mathbb{1}_n - \mathbf{x} \otimes \mathbf{e}_i) = \mathbb{1}_n - \mathbf{x}\mathbf{e}_i^t + \mathbf{x}\mathbf{e}_i^t - \mathbf{x}\,\underbrace{\mathbf{e}_i^t\mathbf{x}}_{=0}\,\mathbf{e}_i^t = \mathbb{1}_n.$$

b) Für Frobeniusmatrizen folgt die Behauptung aus a). Im Allgemeinen ist eine normierte untere Dreiecksmatrix L ein Produkt von Frobeniusmatrizen (Satz 2.47)

$$L = L_1 \cdot ... \cdot L_{n-1}.$$

Als Produkt normierter unterer Dreiecksmatrizen ist

$$L^{-1} = L_{n-1}1^{-1} \cdot ... \cdot L_1^{-1}$$

wieder eine solche normierte untere Dreiecksmatrix. □

Dreiecksmatrizen sind uns schon mehrmals begegnet. Zuerst, als wir die Lösung eines LGS mit dem Gauß-Algorithmus behandelten. Angenommen, die Koeffizientenmatrix A ist quadratisch und invertierbar, dann ist ihre Zeilenstufenform eine normierte obere Dreiecksmatrix.

Jede elementare Zeilenumformung kann man erreichen als Multiplikation mit einer Elementarmatrix von links. Wir wollen das Produkt dieser Elementarmatrizen etwas genauer analysieren, und zwar in dem Fall, wo

i) A invertierbar ist und

ii) bei der Umformung in Zeilenstufenform keine Zeilenvertauschungen (Umformung vom Typ I) notwendig sind.

Auf Umformungen vom Typ II, welche die Diagonalelemente mit ihrem Inversen multiplizieren, um als Zeilenstufenform eine normierte obere Dreiecksmatrix zu erhalten, wollen wir verzichten. Es bleiben also ausschließlich Umformungen vom Typ III.

Wir beginnen also mit der $n \times n$-Matrix

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

Vereinbarungsgemäß ist hier $a_{1,1} \neq 0$, weil wir sonst die erste Zeile mit einer anderen Zeile vertauschen müssten (Umformung vom Typ I). Weil wir $a_{1,1}$ nicht verändern wollen, müssen wir folgendermaßen Zeilen umformen, um die erste Spalte unterhalb $a_{1,1}$ zu bereinigen:

$$II - \frac{a_{2,1}}{a_{1,1}} \cdot I, \quad III - \frac{a_{3,1}}{a_{1,1}} \cdot I, ...$$

Dies $n-1$ Umformungen zusammen, erhält man, indem man A von links mit der Frobenius-Matrix

$$L_1 := \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ -\frac{a_{2,1}}{a_{1,1}} & 1 & 0 & & \vdots \\ -\frac{a_{3,1}}{a_{1,1}} & 0 & 1 & & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ -\frac{a_{n,1}}{a_{1,1}} & 0 & \dots & 0 & 1 \end{pmatrix} = \mathbb{1}_n - \mathbf{m}_1 \otimes \mathbf{e}_1, \quad \text{wobei} \quad \mathbf{m}_1 := \left(0, -\frac{a_{2,1}}{a_{1,1}}, ..., -\frac{a_{n,1}}{a_{1,1}}\right)^t,$$

multipliziert.

Und hat man A durch $k-1$ derartige Umformungen auf eine Form

$$A'_k = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & \dots & \dots & a_{1,n} \\ 0 & a'_{2,2} & \ddots & & & \vdots \\ \vdots & \ddots & \ddots & a'_{k-1,k} & & \vdots \\ \vdots & & 0 & a'_{k,k} & & \vdots \\ \vdots & & \vdots & a'_{k+1,k} & & \vdots \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & a'_{n,k} & \dots & a'_{n,n} \end{pmatrix}$$

gebracht, so ist hier $a'_{k,k} \neq 0$, und die k-te Spalte wird unterhalb der Diagonale bereinigt durch Multiplikation mit der Frobenius-Matrix

$$L_k := \mathbb{1}_n - \mathbf{m}_k \otimes \mathbf{e}_k \quad \text{wobei} \quad \mathbf{m}_k := \left(0, ..., 0, -\frac{a_{k+1,k}}{a_{k,k}}, ..., -\frac{a_{n,k}}{a_{k,k}}\right)^t.$$

Fassen wir zusammen:

Satz 2.49 *Die $n \times n$-Matrix A sei invertierbar, und es sei möglich, sie in Zeilenstufenform (eine obere Dreiecksmatrix) zu bringen, ohne Zeilen zu vertauschen. Dann gibt es Frobenius-Matrizen $L_i = \mathbb{1}_n - \mathbf{m}_i \otimes \mathbf{e}_i, i = 1, ..., n-1$ und eine obere Dreiecksmatrix R so, dass*

$$L_{n-1} \cdot ... \cdot L_1 \cdot A = R.$$

Dummerweise sind hier die Frobeniusmatrizen so angeordnet, dass Satz 2.47 gerade nicht anwendbar ist. Aber mit Satz 2.47 können wir alle Frobenius-Matrizen L_i auf die rechte Seite bringen, um

$$A = L_1^{-1} \cdot ... \cdot L_{n-1} \cdot R = (E + \mathbf{m}_1 \otimes \mathbf{e}_1) \cdot ... \cdot (E + \mathbf{m}_{n-1}\mathbf{e}_{n-1}) \cdot R$$

zu erhalten. Mit dem Beweis von Satz 2.48 a) sehen wir, das Produkt der Frobenius-Matrizen auf der rechten Seite ist die normierte untere Dreiecksmatrix

$$L := (E + \mathbf{m}_1 \otimes \mathbf{e}_1) \cdot ... \cdot (E + \mathbf{m}_{n-1} \otimes \mathbf{e}_{n-1}) = \mathbb{1} + \sum_{i=1}^{n-1} \mathbf{m}_i \otimes \mathbf{e}_i.$$

Wir haben die Existenzaussage des folgenden Satzes bewiesen:

Satz 2.50 (L-R-Zerlegung) *Die invertierbare $n \times n$-Matrix A habe die Eigenschaft, dass sie ohne Zeilenvertauschungen auf Zeilenstufenform zu bringen ist.*

Existenz: Dann gibt es eine obere Dreiecksmatrix R und eine normierte untere Dreiecksmatrix L mit

$$A = L \cdot R.$$

Eindeutigkeit: die Matrizen L und R sind durch A eindeutig bestimmt.

Beweis. Wir brauchen nur noch die Eindeutigkeit zu zeigen. Seien also R_1, R_2 zwei obere Dreiecksmatrizen und L_1, L_2 zwei normierte untere Dreiecksmatrizen mit

$$A = L_1 \cdot R_1 = L_2 \cdot R_2.$$

Daraus folgt

$$L_2^{-1} \cdot L_1 = R_2 \cdot R_1^{-1}.$$

Wegen Satz 2.48 b) ist hier $L_2^{-1} \cdot L_1$ eine normierte untere Dreiecksmatrix und $R_2 \cdot R_1^{-1}$ eine obere Dreiecksmatrix. Deswegen müssen alle Nicht-Diagonal-Elemente in $L_2^{-1} \cdot L_1$ verschwinden. Damit wird

$$L_2^{-1} \cdot L_1 = \mathbb{1}_n, \quad L_2 = L_1.$$

Aus

$$L_1 \cdot R_1 = L_2 \cdot R_2$$

folgt $R_1 = R_2$ nach Multiplikation mit $L_1^{-1} = L_2^{-1}$. □

Natürlich stört in Satz 2.50 die Voraussetzung, dass A ohne Zeilenvertauschungen auf Zeilenstufenform zu bringen ist. Man kann auf diese Voraussetzung verzichten, muss aber dann die auftretenden Permutationen der Zeilen sorgfältig zusammenfassen.

Satz 2.51 (L-R-Zerlegung) *Zu jeder invertierbaren $n \times n$-Matrix A gibt es*

- *eine Permutationsmatrix P,*
- *eine normierte untere Dreiecksmatrix L*
- *und eine obere Dreiecksmatrix R*

mit

$$P \cdot A = L \cdot R.$$

Beweis. Analog zum Beweis von Satz 2.49 bereinigen wir die k-te Spalte, $k = 1, ..., n-1$ mit einer Frobeniusmatrix $L_k = \mathbb{1}_n - \mathbf{m}_k \otimes \mathbf{e}_k$. Vorher müssen wir aber jetzt die k-te Zeile eventuell mit einer Zeile weiter unten vertauschen. Das geschieht mit einer Elementarmatrix

$$E_k = E_1(k, j) \quad \text{wobei} \quad j > k.$$

Wenn wir nicht vertauschen müssen, dann sei E_k die Einheitsmatrix $\mathbb{1}_n$. Analog zum Beweis von Satz 2.49 erhalten wir jetzt

$$L_{n-1} \cdot E_{n-1} \cdot L_{n-2} \cdot E_{n-2} \cdot ... \cdot L_2 \cdot E_2 \cdot L_1 \cdot E_1 \cdot A = R.$$

Jetzt kommt der Trick: Die Elementarmatrizen sind ihre eigenen Inversen. Die Permutationsmatrix

$$P_k := E_{n-1} \cdot E_{n-2} \cdot ... \cdot E_{k+1}, \quad k = 0, ..., n-2,$$

hat also die Inverse

$$P_k^{-1} = E_{k+1} \cdot ... \cdot E_{n-2} \cdot E_{n-1}.$$

Diese Permutationsmatrizen P_k kommen folgendermaßen zum Einsatz: Das oben erhaltene Produkt

$$L_{n-1} \cdot E_{n-1} \cdot L_{n-2} \cdot E_{n-2} \cdot ... \cdot L_2 \cdot E_2 \cdot L_1 \cdot E_1$$

ändern wir ab durch Einfügen von Faktoren $P_k^{-1} \cdot P_k$ in

$$L_{n-1} \cdot (E_{n-1} \cdot L_{n-2} \cdot E_{n-1}) \cdot (E_{n-1} \cdot E_{n-2} \cdot L_{n-3} \cdot E_{n-2} \cdot E_{n-1}) \cdot .... =$$

$$L_{n-1} \cdot (P_{n-2} \cdot L_{n-2} \cdot P_{n-2}^{-1}) \cdot ... \cdot (P_k \cdot L_k \cdot P_k^{-1}) \cdot ... \cdot (P_1 \cdot L_1 \cdot P_1^{-1}) \cdot P_0.$$

Die Permutationsmatrix P_k permutiert nur Spalten P_j mit $j > k$. Deswegen ist

$$\mathbf{e}_k^t \cdot P_k^{-1} = \mathbf{e}_k$$

und

$$P_k \cdot L_k \cdot P_k^{-1} = P_k \cdot (\mathbb{1}_n - \mathbf{m}_k \otimes \mathbf{e}_k) \cdot P_k^{-1} = \mathbb{1}_n - (P_k \cdot \mathbf{m}_k) \otimes \mathbf{e}_k =: L'_k$$

eine Frobeniusmatrix, die sich wie L_k von der Einheitsmatrix nur in der k-ten Spalte unterscheidet. Mit Satz 2.48 folgt:

$$L := (L'_1)^{-1} \cdot (L'_2)^{-1} \cdot ... \cdot (L'_{n-2})^{-1} \cdot (L_{n-1})^{-1}$$

ist eine normierte untere Dreiecksmatrix. Damit haben wir bewiesen

$$P_0 \cdot A = L \cdot R.$$

□

Anzumerken ist, dass die Permutationsmatrix P_0 gerade das Produkt aller verwendeten Zeilenvertauschungen bewirkt.

Auch beim Schmidtschen Orthonormalisierungsverfahren kommt eine obere Dreiecksmatrix vor. Sei etwa $\mathbf{a}^1, ..., \mathbf{a}^n \in \mathbb{R}^n$ eine Basis. Mit diesen Vektoren als Spaltenvektoren bilden wir die invertierbare $n \times n$-Matrix

$$A = (\mathbf{a}^1, ..., \mathbf{a}^n).$$

Beim sukzessiven Orthonormalisieren werden die Vektoren $\mathbf{a}^i$ ersetzt durch Linearkombinationen

$$\mathbf{q}^i = \sum_{j=1}^{i} r'_{i,j} \mathbf{a}^j$$

mit $r'_{i,i} \neq 0$ (weil $span(\mathbf{a}^1, ..., \mathbf{a}^i) = span(\mathbf{q}^1, ..., \mathbf{q}^i)$). Die Vektoren $\mathbf{q}^1, ..., \mathbf{q}^n$ sind orthonormal, also ist die Matrix

$$Q = (\mathbf{q}^1, ..., \mathbf{q}^n) \in M(n \times n)$$

mit diesen Spaltenvektoren orthogonal. Die Beziehung zwischen den Vektoren $\mathbf{a}^i$ und $\mathbf{q}^i$ kann als Matrixgleichung

$$Q = A \cdot R'$$

mit der invertierbaren oberen Dreiecksmatrix

$$R' = \begin{pmatrix} r'_{1,1} & r'_{2,1} & \cdots & r'_{n,1} \\ 0 & r'_{2,2} & & r'_{n,2} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & r'_{n,n} \end{pmatrix}$$

Satz 2.52 (Q-R-Zerlegung) *Die $n \times n$-Matrix sei invertierbar.*

Existenz: Dann gibt es eine orthogonale Matrix Q und eine obere Dreicksmatrix R mit

$$A = Q \cdot R.$$

Eindeutigkeit: In der Matrix R können alle Diagonalelemente $r_{i,i} > 0$ gewählt werden. Die dann entstehenden Matrizen Q und R sind durch A eindeutig bestimmt.

Beweis. Existenz: Wir haben $Q = A \cdot R'$ bewiesen. Mit $R := (R')^{-1}$ folgt daraus

$$A = Q \cdot R.$$

Eindeutigkeit: Es sei $A = Q_1 \cdot R_1 = Q_2 \cdot R_2$. Wie im Beweis von Satz 2.50 sehen wir

$$Q_2^{-1} \cdot Q_1 = R_2 \cdot R_1^{-1}.$$

Die obere Dreiecksmatrix $R_2 \cdot R_1^{-1}$ ist also orthogonal. Auch ihre Inverse = Transponierte ist eine obere Dreiecksmatrix. Dann kann diese Matrix keine Einträge außerhalb der Diagonale enthalten. Sie ist eine orthogonale Diagonalmatrix, d.h., von der Form

$$I = \begin{pmatrix} \pm 1 & & \\ & \ddots & \\ & & \pm 1 \end{pmatrix}.$$

Mit dieser Matrix haben wir

$$R_2 = I \cdot R_1.$$

D.h., die Diagonaleinträge in R sind bis auf das Vorzeichen bestimmt. Wegen

$$Q \cdot R = (Q \cdot I) \cdot (I \cdot R)$$

mit der orthogonalen Matrix $Q \cdot I$ und der oberen Dreiecksmatrix $I \cdot R$ können wir die Diagonaleinträge von R so abändern, dass sie alle positiv sind.

Aber wenn die Diagonal-Einträge von R_1 und R_2 positiv sind, dann gilt das auch für die Einträge von $R_2 \cdot R_1^{-1}$. Diese Matrix ist deswegen die Einheitsmatrix. Es folgt $R_1 = R_2$ und $Q_1 = Q_2$. □

Aufgabe 2.63 *Gegeben seien die Frobeniusmatrizen*

$$F_1 := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 \\ 4 & 0 & 0 & 1 \end{pmatrix}, \quad F_2 := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 3 & 0 & 1 \end{pmatrix}, \quad F_3 := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 3 & 0 & 1 \end{pmatrix}.$$

a) Berechnen Sie

$$L_1 := F_1 \cdot F_2 \cdot F_3 \quad \textit{und} \quad L_2 := F_3 \cdot F_2 \cdot F_1.$$

b) Berechnen Sie L_1^{-1} und L_2^{-1}.

Aufgabe 2.64 *Berechnen Sie die L-R-Zerlegung für die Matrizen A_2, A_3, A_4 aus Aufgabe 2.59.*

Aufgabe 2.65 *Berechnen Sie die L-R-Zerlegung für die Matrix D_4 aus Aufgabe 2.59.*

Aufgabe 2.66 *Berechnen Sie eine Q-R-Zerlegung für die Matrizen A_2 und A_3 aus Aufgabe 2.59*

3 Rechenstrukturen

In diesem Kapitel werden wir die Theorie, welche wir bisher entwickelt haben (lineare Gleichungssysteme, lineare Unterräume, Matrizen, Determinanten), auf eine möglichst allgemeine, begriffliche Grundlage stellen. Der Bezug zur geometrischen Vorstellung (=Realisierung im $\mathbb{R}^n$), der für das erste Verstehen wahrscheinlich unerlässlich ist, wird dabei allerdings verloren gehen.

Bisher haben wir ausschließlich mit reellen Zahlen gerechnet. Wir werden dies jetzt so verallgemeinern, dass auch komplexe Zahlen vorkommen. Hier sind manche Fragestellungen einfacher zu diskutieren. Aber auch die in der Codierungstheorie auftretenden endlichen Körper werden hier einbezogen.

3.1 Gruppen

Definition 3.1 *Eine* Gruppe *ist eine (nichtleere) Menge G zusammen mit einer Verknüpfungsoperation*

$$G \times G \to G \qquad g, h \mapsto g \cdot h,$$

welche folgende Eigenschaften hat:

Assoziativität: *Für alle $g, h, k \in G$ ist*

$$(g \cdot h) \cdot k = g \cdot (h \cdot k).$$

Existenz der Eins: *Es gibt ein $e \in G$ mit*

$$e \cdot g = g \quad \text{für alle } g \in G.$$

Existenz des Inversen: *Für alle $g \in G$ gibt es ein $g^{-1} \in G$ mit $g^{-1} \cdot g = e$.*

Die Gruppe heißt kommutativ *oder* abelsch*, wenn zusätzlich gilt*

$$g \cdot h = h \cdot g \quad \text{für alle } g, h \in G.$$

Bevor wir aus diesen Eigenschaften Konsequenzen ziehen, beschreiben wir erst Beispiele von Gruppen, die wir schon kennen.

Beispiel 3.1 *Die Menge $\mathbb{R}$ mit der Addition „+“ als Verknüpfung ist eine abelsche Gruppe. Es ist $e = 0$ und $g^{-1} = -g$. Diese Gruppe enthält die* Untergruppen $(\mathbb{Q}, +)$ *der rationalen und* $(\mathbb{Z}, +)$ *der ganzen Zahlen.*

Dabei haben wir einen neuen Begriff benutzt:

Definition 3.2 (Untergruppe) *Es sei G eine Gruppe. Eine Teilmenge $U \subset G$ heißt* Untergruppe*, wenn sie mit der Verknüpfungsoperation aus G selbst eine Gruppe ist. D.h. also:*

- $1 \in U$,
- $g, h \in U \quad \Rightarrow \quad g \cdot h \in U$,

- $g \in U \quad \Rightarrow \quad g^{-1} \in U$.

Beispiel 3.2 *Der Zahlenraum* $\mathbb{R}^n$ *mit der Addition „+" als Verknüpfung ist ebenfalls eine abelsche Gruppe. Es ist* $e = \mathbf{0}$ *der Nullvektor und* $\mathbf{x}^{-1} = -\mathbf{x}$.

Beispiel 3.3 *Die Menge* $\mathbb{R}^* := \mathbb{R} \setminus \{0\}$ *der rellen Zahlen* $\neq 0$ *ist eine abelsche Gruppe mit der Multiplikation „·" als Verknüpfung. Dabei ist* $e = 1$ *und* $g^{-1} = \frac{1}{g}$. *Sie enthält die Untergruppe* $\mathbb{Q}^* := \mathbb{R}^* \cap \mathbb{Q}$. *Auch die zwei-elementige Menge* $\{\pm 1\}$ *ist eine Untergruppe der Gruppe* $\mathbb{R}^*$.

Beispiel 3.4 *Mit* $\mathbb{Z}_n$ *bezeichnen wir die endliche Menge* $\{0, 1, ..., n-1\}$. *Die Addition modulo* n

$$g + h := \left\{ \begin{array}{c} g+h \\ g+h-n \end{array} \right\} \text{ wenn } \left\{ \begin{array}{l} g+h \leq n-1 \\ g+h \geq n \end{array} \right.$$

definiert auf dieser Menge eine Verknüpfung, welche sie zu einer abelschen Gruppe macht. Es ist $e = 0$ *und*

$$g^{-1} = \left\{ \begin{array}{cl} 0 & \text{wenn } g = 0 \\ n-g & \text{wenn } g > 0 \end{array} \right. .$$

Beispiel 3.5 *Die* symmetrische Gruppe Σ_n *ist die Menge aller Permutationen der Zahlen* $1, ..., n$ *mit der Hintereinanderschaltung* $\sigma \cdot \tau = \sigma \circ \tau$ *als Verknüpfung. Es ist* $e = id$ *und* σ^{-1} *die Umkehrabbildung. Diese Gruppe ist für* $n \geq 3$ *nicht abelsch, da z.B.*

$$(1,2)(2,3) = (1,2,3) \neq (1,3,2) = (2,3)(1,2).$$

Wir stellen einige Konsequenzen aus den Gruppeneigenschaften zusammen:

(1) Die Eins $e \in G$ mit der Eigenschaft $e \cdot g = g$ („Linkseins") ist auch eine „Rechtseins", d.h. es gilt $g \cdot e = g$ für alle $g \in G$.

Beweis. Zu beliebigem $g \in G$ gibt es das Inverse g^{-1} mit $g^{-1} \cdot g = e$ und dazu wieder ein Inverses $g' \in G$ mit $g' \cdot g^{-1} = e$. Daraus folgt

$$g = e \cdot g = (g' \cdot g^{-1}) \cdot g = g' \cdot e = g' \cdot (e \cdot e) = (g' \cdot e) \cdot e = g' \cdot (g^{-1} \cdot g) \cdot e = (g' \cdot g^{-1}) \cdot g \cdot e = g \cdot e.$$

(2) Das „Linksinverse" g^{-1} zu g mit der Eigenschaft $g^{-1} \cdot g = e$ ist auch ein „Rechtsinverses", d.h. es gilt $g \cdot g^{-1} = e$.

Beweis. Mit der Notation des vorhergehenden Beweises ist $g = g' \cdot e$ und wegen der Eigenschaft (1) ist dies $= g'$. Also ist auch $g \cdot g^{-1} = g' \cdot g^{-1} = e$.

(3) Das Einselement e ist eindeutig bestimmt.

Beweis. Sei auch $e' \in G$ mit $e' \cdot g = g$ für alle $g \in G$. Setzen wir $g = e$, so folgt daraus $e' \cdot e = e$. Da e aber auch eine Rechtseins ist, gilt $e' \cdot e = e'$.

(4) Das Inverse g^{-1} ist durch g eindeutig bestimmt.

Beweis. Es sei $g^{-1} \cdot g = g' \cdot g = e$. Wegen (2) ist dann

$$g^{-1} = (g^{-1} \cdot g) \cdot g^{-1} = e \cdot g^{-1} = (g' \cdot g) \cdot g^{-1} = g' \cdot (g \cdot g^{-1}) = g'.$$

(5) Kürzungsregel: Seien $a, b, g \in G$. Wenn $g \cdot a = g \cdot b$ gilt, dann auch (Linksmutiplikation mit g^{-1}) die Gleichung $a = b$. Aus $a \cdot g = b \cdot g$ folgt (nach Rechtsmultiplikation mit g^{-1}) die Gleichung $a = b$.

(6) Lösbarkeit von Gleichungen: Zu beliebigen $g, h \in G$ gibt es genau ein $x \in G$ und ein $y \in G$ mit

$$\begin{array}{ll} g \cdot x = h & (\text{ nämlich } x := g^{-1} \cdot h,) \\ y \cdot g = h & (\text{ nämlich } y := h \cdot g^{-1}.) \end{array}$$

Definition 3.3 *Es seien G, H Gruppen. Eine Abbildung $\varphi : G \to H$ heißt (Gruppen-) Homomorphismus, wenn für alle $g_1, g_2 \in G$ gilt*

$$\varphi(g_1 \cdot g_2) = \varphi(g_1) \cdot \varphi(g_2).$$

Satz 3.1 *Für jeden Gruppenhomomorphismus $\varphi : G \to H$ gilt:*
a) Die Eins $1_G \in G$ wird auf die Eins $1_H \in H$ abgebildet.
b) Das Inverse von $g \in G$ wird auf $\varphi(g)^{-1}$ abgebildet.
c) Die Menge

$$Kern(\varphi) = \{g \in G : \varphi(g) = 1_H\} \subset G$$

ist eine Untergruppe von G.

Beweis. a) Wir berechnen

$$\varphi(1_G) = \varphi(1_G \cdot 1_G) = \varphi(1_G) \cdot \varphi(1_G)$$

und multiplizieren diese Gleichung in H (etwa von rechts) mit $\varphi(1_G)^{-1}$ um $1_H = \varphi(1_G)$ zu erhalten.

b) Wegen

$$\varphi(g^{-1}) \cdot \varphi(g) = \varphi(g^{-1} \cdot g) = \varphi(1_G) = 1_H$$

ist $\varphi(g^{-1})$ das Inverse von $\varphi(g)$ in H.

c) Mit $g_1, g_2 \in Kern(\varphi)$ gehört auch $g_1 \cdot g_2$ zu $Kern(\varphi)$ wegen

$$\varphi(g_1 \cdot g_2) = \varphi(g_1) \cdot \varphi(g_2) = 1_H \cdot 1_H = 1_H.$$

Die Multiplikation in G definiert (durch Einschränkung) also eine Multiplikation in $Kern(\varphi)$. Wegen a) gehört 1_G zu $Kern(\varphi)$ und ist das Eins-Element in $Kern(\varphi)$. Mit $g \in Kern(\varphi)$ gehört auch g^{-1} zu $Kern(\varphi)$ wegen

$$\varphi(g^{-1}) = \varphi(g)^{-1} = 1_H^{-1} = 1_H.$$

□

Beispiel 3.6 *Wegen Satz 2.40 b) ist*

$$sign : \Sigma_n \to \{\pm 1\}$$

ein Gruppen-Homomorphismus. Sein Kern ist die alternierende Gruppe A_n *der Permutationen σ mit $sign(\sigma) = 1$.*

Beispiel 3.7 *Die* allgemeine lineare Gruppe *ist die Menge $GL(n, \mathbb{R})$ aller* invertierbaren $n \times n$-*Matrizen mit der Matrizenmultiplikation als Verknüpfung. Das Einselement ist $e = \mathbb{1}_n$, das Inverse ist die inverse Matrix. Für $n = 1$ ist dies die abelsche Gruppe $\mathbb{R}^*$, für $n \geq 2$ ist $GL(n, \mathbb{R})$ nicht abelsch. Aus dem Determinanten-Multiplikationssatz (Satz 2.45) folgt, dass die Abbildung*

$$det : GL(n, \mathbb{R}) \to \mathbb{R}^*$$

ein Gruppenhomomorphismus ist. Sein Kern ist die spezielle lineare Gruppe

$$SL(n, \mathbb{R}) = \{A \in M(n \times n) : \ det(A) = 1\} \subset GL(n, \mathbb{R}).$$

Beispiel 3.8 *Die* reelle orthogonale Gruppe *ist die Menge*

$$\begin{aligned} O(n,\mathbb{R}) &= \{A = \begin{pmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{pmatrix} \in GL(n,\mathbb{R}) \,:\, \mathbf{a}_1,...,\mathbf{a}_n \text{ bilden eine O-N-Basis des } \mathbb{R}^n\} \\ &= \{A \in GL(n,\mathbb{R}) \,:\, A \cdot A^t = \mathbb{1}_n\}. \end{aligned}$$

Sie ist eine Untergruppe der $GL(n,\mathbb{R})$. Ihrerseits enthält sie als Untergruppe die spezielle orthogonale Gruppe

$$SO(n,\mathbb{R}) = O(n,\mathbb{R}) \cap SL(n,\mathbb{R}) = \{A \in O(n,\mathbb{R}) : \, det(A) = 1\}.$$

Wir betrachten die zwei-dimensionale orthogonale Gruppe $O(2,\mathbb{R})$ etwas genauer. Dies ist die Menge aller 2×2-Matrizen $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ mit

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2+b^2 & ac+bd \\ ac+bd & c^2+d^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Die Gleichungen

$$a^2+b^2 = c^2+d^2 = 1, \qquad ac+bd = 0$$

haben die Lösungen

$$(a,b) = (cos(\varphi), -sin(\varphi)), \qquad (c,d) = \pm(sin(\varphi), cos(\varphi)), \varphi \in \mathbb{R}.$$

Also besteht $O(2,\mathbb{R})$ aus den Matrizen

$$\begin{pmatrix} cos(\varphi) & -sin(\varphi) \\ sin(\varphi) & cos(\varphi) \end{pmatrix} \quad \text{(Drehmatrix)},$$

und

$$\begin{pmatrix} cos(\varphi) & -sin(\varphi) \\ -sin(\varphi) & -cos(\varphi) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} cos(\varphi) & -sin(\varphi) \\ sin(\varphi) & cos(\varphi) \end{pmatrix}$$

(Spiegelung an der x_2-Achse $\circ$ Drehung).

Die Drehmatrizen haben die Determinante $= 1$ und bilden die Untergruppe $SO(2,\mathbb{R})$.

Beispiel 3.9 *Die* konforme Gruppe $\mathbb{C}^*$ *ist die Menge*

$$\left\{\begin{pmatrix} a & -b \\ b & a \end{pmatrix} = \sqrt{a^2+b^2}\begin{pmatrix} cos(\varphi) & -sin(\varphi) \\ sin(\varphi) & cos(\varphi) \end{pmatrix} \in GL(2,\mathbb{R})\right\} = \{r \cdot A \,:\, 0 < r \in \mathbb{R}, A \text{ Drehmatrix}\}.$$

Diese Matrizen beschreiben Drehstreckungen. *Die Gruppe ist abelsch.*

Aufgabe 3.1 *Welche der folgenden Teilmengen von $M(n \times n)$ bilden eine Gruppe bezüglich der Matrizenmultiplikation:*
a) die Menge aller oberen Dreiecksmatrizen,
b) die Menge aller oberen Dreiecksmatrizen mit Determinante $\neq 0$
c) die Menge aller normierten oberen Dreiecksmatrizen,
d) für festes $B \in GL(n,\mathbb{R})$ die Menge $\{A \in GL(n,\mathbb{R}) : \, ABA^t = B\}$.

Aufgabe 3.2 *Eine Matrix $C \in GL(n, \mathbb{R})$ werde festgehalten.*
a) Zeigen Sie, dass die Matrizen $A \in GL(n, \mathbb{R})$ mit

$$A \cdot C = C \cdot A$$

eine Gruppe bezüglich der Matrizenmultiplikation bilden.
b) Bestimmen Sie diese Gruppe für $n = 2$ und

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

Aufgabe 3.3 *Zeigen Sie, dass die folgende Menge unter der Matrizenmultiplikation eine Gruppe ist:*

$$Sp(2n) := \left\{ \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in M(2n \times 2n) : \begin{array}{c} A, B, C, D \in M(n \times n), \\ AB^t = BA^t,\ CD^t = DC^t, \\ AD^t - BC^t = \mathbb{1}_n \end{array} \right\}.$$

Aufgabe 3.4 *a) Es seien $G_1, G_2 \subset G$ Untergruppen der Gruppe G. Zeigen Sie, dass auch $G_1 \cap G_2$ eine Untergruppe von G ist.*
b) Es seien G_1, G_2 Gruppen. Zeigen Sie, dass auch ihr kartesisches Produkt $G_1 \times G_2$ eine Gruppe ist mit der Gruppenoperation

$$(g_1, g_2) \cdot (h_1, h_2) = (g_1 \cdot h_1, g_2 \cdot h_2) \quad \text{für} \quad g_1, h_1 \in G_1,\ g_2, h_2 \in G_2.$$

Aufgabe 3.5 *Es sei $\varphi : G \to H$ ein Gruppenhomomorphismus mit Kern U. Zeigen Sie für jedes $h = \varphi(g) \in H$ im Bild von φ: Die Anzahl der Elemente $g \in G$, welche auf h abgebildet werden ist gleich der Anzahl der Elemente, welche im Kern U enthalten sind.*

Aufgabe 3.6 *Zeigen Sie, dass die multiplikative Gruppe $\{\pm 1\}$ eine Untergruppe von $\mathbb{C}^*$ ist, nicht aber die Gruppe $\{\pm 1\} \times \{\pm 1\}$.*

3.2 Körper

Definition 3.4 *Ein* Körper *ist eine (nichtleere) Menge K mit zwei kommutativen Rechenoperationen „+“ und „·“ (D.h. also $x + y = y + x$ und $x \cdot y = y \cdot x$). Für diese Operationen muss gelten:*

(a) K mit „+“ ist eine abelsche Gruppe. (Das neutrale Element wird mit $0 \in K$ bezeichnet, und das Inverse zu $a \in K$ mit $-a$.)

(b) $K^* := K \setminus 0$ mit „·“ ist eine abelsche Gruppe. (Das neutrale Element wird mit $1 \in K$ bezeichnet, und das Inverse zu $0 \neq a \in K$ ist $\frac{1}{a}$.)

(c) Für alle $a, b, c \in K$ gilt das Distributivgesetz

$$(a + b) \cdot c = a \cdot c + b \cdot c.$$

Aus dem Distributivgesetz folgt sofort für alle $x \in K$

$$0 \cdot x = (x - x) \cdot x = x \cdot x - x \cdot x = 0.$$

Also kann $0 \in K$ kein Inverses bezüglich der Multiplikation in K besitzen.

Beispiel 3.10 *Die reellen Zahlen* $\mathbb{R}$ *und die rationalen Zahlen* $\mathbb{Q}$ *mit den üblichen Rechenoperationen bilden einen Körper. Auch in der Menge* $\mathbb{Z}$ *der ganzen Zahlen gibt es die Addition und die Multiplikation. Ganze Zahlen* $\neq \pm 1$ *haben aber kein Inverses bezüglich Multiplikation in* $\mathbb{Z}$*. Also ist* $\mathbb{Z}$ *kein Körper.*

Beispiel 3.11 *Der Körper* $\mathbb{C}$ *der* komplexen Zahlen: *Als Menge ist* $\mathbb{C} = \mathbb{R}^2 = \{(a,b) \,:\, a,b \in \mathbb{R}\}$. *Statt* (a,b) *schreibt man auch* $a + b \cdot i$. *Die Addition ist die übliche Vektoraddition des* $\mathbb{R}^2$. *Damit ist (a) erfüllt. Die Multiplikation wird definiert wie in der konformen Gruppe*

$$\mathbb{C}^* = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} : (0,0) \neq (a,b) \in \mathbb{R}^2 \right\}.$$

Wegen der Formel

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \cdot \begin{pmatrix} a' & -b' \\ b' & a' \end{pmatrix} = \begin{pmatrix} aa' - bb' & -(ab' + a'b) \\ ab' + a'b & aa' - bb' \end{pmatrix}$$

definiert man die Multiplikation also durch

$$(a + b \cdot i) \cdot (a' + b' \cdot i) := aa' - bb' + (ab' + a'b) \cdot i.$$

Da man natürlich $i = 0 + 1 \cdot i$ *setzt, ist insbesondere*

$$\boxed{i^2 = -1.}$$

Beweis von (b). Die so definierte Multiplikation in $\mathbb{C}$ ist assoziativ, weil Multiplikation von Matrizen assoziativ ist (Satz 2.12 b). Man sieht unmittelbar, dass sie kommutativ ist. Das Einselement ist $1 = 1 + 0 \cdot i$, weil dieses Element zur Einheitsmatrix gehört ($a = 1, b = 0$). Die Inverse Matrix ist

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}^{-1} = \frac{1}{a^2 + b^2} \cdot \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

Also ist für $0 \neq a + b \cdot i \in \mathbb{C}$ das Inverse

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

(c) folgt aus der Distributivität der Matrizen-Multiplikation (Satz 2.12 a). □

In $\mathbb{C}$ gibt es die *Konjugation*

$$c = a + b \cdot i \quad \mapsto \quad \bar{c} = a - b \cdot i.$$

Man benutzt sie, um den *Betrag* der komplexen Zahl c (die Länge des Vektors (a,b))

$$|c| = \sqrt{a^2 + b^2} = \sqrt{c \cdot \bar{c}}$$

und ihr Inverses

$$\frac{1}{c} = \frac{1}{|c|^2}\bar{c}$$

kürzer zu beschreiben.

Eine Zahl $c \in \mathbb{C}$ ist genau dann reell, wenn $c = \bar{c}$. Konjugation verträgt sich nicht nur mit der Addition komplexer Zahlen

$$\overline{c_1 + c_2} = \overline{(a_1 + ib_1) + (a_2 + ib_2)} = \overline{a_1 + a_2 + i(b_1 + b_2)} = a_1 + a_2 - i(b_1 + b_2) = \bar{c}_1 + \bar{c}_2$$

sondern auch mit der Multiplikation:

$$\begin{aligned}
\overline{c_1 \cdot c_2} &= \overline{(a_1 + ib_1) \cdot (a_2 + ib_2)} \\
&= \overline{a_1 a_2 - b_1 b_2 + i(a_1 b_2 + a_2 b_1)} \\
&= a_1 a_2 - b_1 b_2 - i(a_1 b_2 + a_2 b_1) \\
&= (a_1 - ib_1) \cdot (a_2 - ib_2) \\
&= \bar{c}_1 \cdot \bar{c}_2.
\end{aligned}$$

Für die Theorie sind die komplexen Zahlen vor allem wegen des folgenden Satzes wichtig.

Satz 3.2 (Fundamentalsatz der Algebra) *Jedes komplexe Polynom*

$$p(x) = a_0 + a_1 \cdot x + ... + a_n \cdot x^n, \quad a_\nu \in \mathbb{C},$$

von einem Grad $n \geq 1$ hat eine Nullstelle in $c \in \mathbb{C}$. Es gibt also eine komplexe Zahl c mit $p(c) = 0$.

Diesen Satz werden wir später anwenden, sein Beweis geht aber weit über die in dieser Vorlesung behandelte Theorie hinaus. Er ist etwa Standardstoff der Vorlesung 'Funktionentheorie'.

Jedes reelle Polynom ist natürlich auch ein komplexes Polynom. Der Fundamentalsdatz der Algebra lehrt also, dass jedes reelle Polynom zumindest komplexe Nullstellen hat.

Beispiel 3.12 *Das reelle Polynom $p(x) = 1+x^2$ hat keine reellen Nullstellen, wohl aber die komplexen Nullstellen $\pm i$.*

Beispiel 3.13 *Die endlichen Körper $\mathbb{F}_p$ (p eine Primzahl). Als Menge ist $\mathbb{F}_p$ die Teilmenge $\{0, 1, ..., p-1\} \subset \mathbb{Z}$. Die Operationen „+“ und „·“ sind die übliche Addition und Multiplikation, aber* modulo p genommen. *Bezeichnen wir die Zahl $m \in \mathbb{Z}, 0 \leq m < p$, aufgefasst als Element in $\mathbb{F}_p$ mit $[m]$, so ist also*

$$[m_1] + [m_2] = [m_1 + m_2 \text{ modulo } p].$$

$\mathbb{F}_p$ mit der Addition ist eine abelsche Gruppe, die wir in Abschnitt 3.1 mit $\mathbb{Z}_p$ bezeichneten. Die Multiplikation ist explizit definiert durch

$$[m] \cdot [n] = [r] \text{ wenn } r + k \cdot p = m \cdot n \text{ für ein } k \in \mathbb{Z} \text{ und } 0 \leq r < p.$$

Diese Multiplikation ist assoziativ und kommutativ, da dies für die Multiplikation in $\mathbb{Z}$ gilt, und das neutrale Element ist $[1] \in \mathbb{F}_p$. Auch die Distributivgesetze übertragen sich aus $\mathbb{Z}$. Schwierigkeiten macht nur die Existenz des Inversen für die Multiplikation mit $0 \neq [m] \in \mathbb{F}_p$. Hierzu ist nachzuweisen, dass die Multiplikation

$$\mathbb{F}_p \to \mathbb{F}_p \quad [n] \mapsto [m] \cdot [n]$$

surjektiv ist. Da $\mathbb{F}_p$ ein endliche Menge ist, genügt es zu zeigen, diese Abbildung ist injektiv, d.h.:

$$[n_1], [n_2] \in \mathbb{F}_p \text{ mit } [m] \cdot [n_1] = [m] \cdot [n_2] \qquad \Rightarrow \qquad [n_1] = [n_2].$$

Wegen des Distributivgesetzes genügt es, dafür zu zeigen

$$[m] \cdot [n] = 0 \qquad \Rightarrow \qquad [n] = 0.$$

Nun bedeutet $[m] \cdot [n] = 0 \in \mathbb{F}_p$ *für die ganzen Zahlen* m *und* n*, dass* mn *durch* p *teilbar ist. Dabei kann* p *nicht* m *teilen, weil* $0 < m < p$*. Also muss der Primfaktor* p *die Zahl* n *teilen. Mit* $0 \leq n < p$ *folgt daraus* $[n] = 0$.

Beispiel 3.14 *Wie über* $\mathbb{R}$ *oder* $\mathbb{C}$ *definiert man auch über einem beliebigen Körper* K *eine rationale Funktion als einen Quotienten*

$$f = \frac{p(X)}{q(X)}, \quad p(X) \in K[x],\, 0 \neq q(x) \in K[X],$$

zweier Polynome. Wegen der Nullstellen des Nenners kann man f *meist nicht als eine richtige Funktion* $K \to K$ *auffassen. Deswegen muss man die Gleichheit zweier rationaler Funktionen folgendermaßen definieren:*

$$\frac{p_1(X)}{q_1(X)} = \frac{p_2(X)}{q_2(X)} \quad \Leftrightarrow \quad p_1(X) \cdot q_2(X) = p_2(X) \cdot q_1(X).$$

Damit bildet diese Menge rationaler Funktionen einen Körper $K(X)$.

Aufgabe 3.7 *Stellen Sie die komplexen Zahlen*

$$(1+i)^{-1}, \quad (2+i)^{-1}, \quad (1+3i)^{-1}$$

dar in der Form $a + bi$ *mit* $a, b \in \mathbb{R}$.

Aufgabe 3.8 *Bestimmen Sie* $\det(A)$, A^2 *und* A^{-1} *für die komplexe* 2×2*-Matrix*

$$A = \begin{pmatrix} 1+i & -i \\ i & 1-i \end{pmatrix}.$$

Aufgabe 3.9 *Berechnen Sie die Determinanten der komplexen Matrizen*

$$A = \begin{pmatrix} 1 & i & 0 \\ i & 0 & 2i \\ 0 & 2i & 2 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & i & -1 \\ i & -1 & i \\ -1 & i & 1 \end{pmatrix}.$$

Aufgabe 3.10 *Es sei* U *die Menge der komplexen* 2×2*-Matrizen*

$$\begin{pmatrix} a & b \\ -\bar{b} & \bar{a} \end{pmatrix}$$

mit $a, b \in \mathbb{C}$ *und* $a\bar{a} + b\bar{b} > 0$*. Zeigen Sie, dass* U *bezüglich der Matrizenmultiplikation eine Gruppe ist.*

Aufgabe 3.11 *Lösen Sie das lineare Gleichungssystem*

$$\begin{array}{ccccccc} x & + & iy & & & = & i \\ & & y & + & iz & = & i \\ ix & + & & + & z & = & i \end{array}$$

Aufgabe 3.12 *a) Bestimmen Sie den Rang der Matrix*

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

über dem Körper $\mathbb{F}_2$ *und über dem Körper* $\mathbb{F}_5$.
b) Lösen Sie das lineare Gleichungssytem

$$\begin{pmatrix} x & + & y & & & = & 1 \\ & & y & + & z & = & 0 \\ x & + & & + & z & = & 1 \end{pmatrix}$$

über $\mathbb{F}_2$ *und über* $\mathbb{F}_5$.

Aufgabe 3.13 (Satz von Vieta) *Jedes komplexe Polynom*

$$p(x) = a_0 + a_1 \cdot x + ... + a_n \cdot x^n$$

ist ein Produkt

$$p(x) = a_n \cdot (x - c_1) \cdot ... \cdot (x - c_n).$$

Aufgabe 3.14 *Jedes reelle Polynom von ungeradem Grad n hat mindestens eine reelle Nullstelle.*

3.3 K-Vektorräume

Mit Elementen aus einem beliebigen Körper kann man genauso, wie mit reellen Zahlen rechnen, wenn man nichts anderes, als die genannten Körpereigenschaften benutzt. *Alles, was wir zu linearen Gleichungssystemen, Matrizenmultiplikation und Determinanten gesehen haben, gilt deswegen über beliebigen Körpern!*

Die einzige Ausnahem ist der Beweis von Satz 2.43, weil er das im Körper $\mathbb{F}_2$ nicht existierende multiplikative Inverse der Zahl 2 benutzt. Das kann man aber leicht reparieren:

Beweis von Satz 2.43 über $\mathbb{F}_2$: Die Matrix A habe zwei gleiche Zeilen $\mathbf{a}_i = \mathbf{a}_j, i \neq j$. Wir bezeichnen mit $\sigma_{i,j}$ die Vertauschung $(i,j) \in \Sigma_n$ und betrachten die Abbildung

$$F : \Sigma_n \to \Sigma_n, \quad \tau \mapsto \tau \circ \sigma_{i,j}.$$

Diese Abbildung ist bijektiv mit Umkehrabbildung $F^{-1} = F$. Offensichtlich ist $F(\tau) \neq \tau$ für alle $\tau \in \Sigma_n$. Die Menge Σ_n zerfällt also in $n!/2$ disjunkte Paare $\{\tau, F(\tau)\}$. (Weil die Matrix A mindestens zwei Zeilen hat, muss $n \geq 2$ sein.) Zu jedem derartigen Paar gehören in der Leibnizformel die beiden Summanden

$$\begin{aligned} s_1 &= sign(\tau) \cdot a_{1,\tau(1)} \cdot ... \cdot a_{i,\tau(i)} \cdot ... a_{j,\tau(j)} \cdot ... \cdot a_{n,\tau(n)} \\ s_2 &= sign(\tau \circ \sigma_{i,j}) \cdot a_{1,\tau(1)} \cdot ... \cdot a_{i,\tau(j)} \cdot ... \cdot a_{j,\tau(i)} \cdot ... \cdot a_{n\tau(n)} \\ &= sign(\tau \circ \sigma_{i,j}) \cdot a_{1,\tau(1)} \cdot ... \cdot a_{j,\tau(j)} \cdot ... \cdot a_{i,\tau(i)} \cdot ... \cdot a_{n\tau(n)} \end{aligned}$$

Wegen $sign(\tau \circ \sigma_{i,j}) = -sign(\tau)$ ist $s_1 = -s_2$. Insgesamt kürzen sich in der Leibnizformel alle derartigen Paare heraus, und es folgt $det(A) = 0$. $\square$

Man kann also in der Definition des $\mathbb{R}$-Vektorraums den Körper $\mathbb{R}$ durch einen beliebigen Körper K ersetzen. Das führt zu

Definition 3.5 *Ein* Vektorraum über dem Körper K *(oder kürzer ausgedrückt: ein* K*-Vektorraum) ist eine abelsche Gruppe* V *(Gruppenoperation „+“ geschrieben, mit neutralem Element* $\mathbf{0} \in V$*) zusammen mit einer Operation*

$$K \times V \to V, \qquad c, \mathbf{v} \mapsto c \cdot \mathbf{v}$$

von K *auf* V*, für die gilt:*

(a) $c_1 \cdot (c_2 \cdot \mathbf{v}) = (c_1 c_2) \cdot \mathbf{v}$ *(Assoziativität)*
für alle $c_1, c_2 \in K$, $\mathbf{v} \in V$,

(b) $(c_1 + c_2) \cdot \mathbf{v} = c_1 \cdot \mathbf{v} + c_2 \cdot \mathbf{v}$ *(Distributivität)*
$c \cdot (\mathbf{v}_1 + \mathbf{v}_2) = c \cdot \mathbf{v}_1 + c \cdot \mathbf{v}_2$ *(Distributivität)*
für alle $c_1, c_2, c \in K, \mathbf{v}, \mathbf{v}_1, \mathbf{v}_2 \in V$,

(c) $1 \cdot \mathbf{v} = \mathbf{v}$ *für alle* $\mathbf{v} \in V$.

Aus den Distributivgesetzen folgt (wie schon für $\mathbb{R}$-Vektorräume gezeigt) für alle $\mathbf{v} \in V$:

$$0 \cdot \mathbf{v} = (0+0) \cdot \mathbf{v} = 0 \cdot \mathbf{v} + 0 \cdot \mathbf{v} \quad \Rightarrow \quad 0 \cdot \mathbf{v} = \mathbf{0} \in V,$$
$$\mathbf{v} + (-1) \cdot \mathbf{v} = (1-1) \cdot \mathbf{v} = 0 \cdot \mathbf{v} = \mathbf{0} \quad \Rightarrow \quad (-1) \cdot \mathbf{v} = -\mathbf{v}.$$

Nur bei der Behandlung des Skalarprodukts haben wir etwas anderes benutzt als die Körperaxiome. Wichtige Aussagen waren hier Ungleichungen. Aber soetwas, wie die $' \leq '$-Relation gibt es in Körpern $\neq \mathbb{R}$ i.a. nicht. Alles, was wir bisher an Begriffen und Aussagen für $\mathbb{R}$-Vektorräume (ohne weitere Voraussetzungen, wie etwa das Skalarprodukt) zusammengestellt haben, gilt auch für K-Vektorräume. In den Definitionen ist überall die Skalarenmenge $\mathbb{R}$ durch den zugrunde gelegten Körper K zu ersetzen, z.B.:

Definition 3.6 *Eine Abbildung* $\Phi : V_1 \to V_2$ *des* K*-Vektorraums* V_1 *in den* K*-Vektorraum* V_2 *heißt* linear *(genauer* K*-linear), wenn*

$$\Phi(s \cdot \mathbf{x} + t \cdot \mathbf{y}) = s \cdot \Phi(\mathbf{x}) + t \cdot \Phi(\mathbf{y})$$

für alle $\mathbf{x}, \mathbf{y} \in V_1$, $s, t \in K$ *gilt. Die Menge aller linearen Abbildungen von* V *nach* W *bezeichnen wir mit* $Hom(V, W)$*, oder wenn die Rolle des Körpers* K *betont werden soll mit* $Hom_K(V, W)$.

Diese Art von Abbildungen ist uns vom Zahlenraum $\mathbb{R}^n$ her wohlbekannt, aber dass es bei der Linearität auf K ankommt, das ist neu. Auch bei der Dimension müssen wir auf den Grundkörper achten. So bezeichnen wir mit $dim_K(V)$ die Dimension von V als K-Vektorraum.

Beispiel 3.15 *Der* $\mathbb{C}$*-Vektorraum* $\mathbb{C}$ *ist auch ein* $\mathbb{R}$*-Vektorraum mit*

$$dim_{\mathbb{C}}(\mathbb{C}) = 1, \quad dim_{\mathbb{R}}(\mathbb{C}) = 2.$$

Die Konjugationsabbildung $\mathbb{C} \to \mathbb{C}$ *ist nicht* $\mathbb{C}$*-linear, wohl aber* $\mathbb{R}$*-linear.*

Also:

Alle Aussagen aus den Kapiteln 1 und 2 für allgemeine $\mathbb{R}$-Vektorräume gelten auch für allgemeine K-Vektorräume.

Bei der Definition der Determinante mit der Leibniz-Formel müssen wir allerdings die Signum-Funktion auffassen als Abbildung

$$sign : \Sigma_n \to \{\pm 1\} \subset K.$$

Beispiel 3.16 *Ebenso wie $\mathbb{R}^n$ ein $\mathbb{R}$-Vektorraum ist, so ist*

$$K^n = \underbrace{K \times ... \times K}_{nmal} = \{(x_1, ..., x_n) : x_1, ..., x_n \in K\}$$

ein Vektorraum über K. Wie früher wollen wir die Elemente in K^n weiterhin als Spaltenvektoren auffassen. Wir hätten also eben $(x_1, ..., x_n)^t \in K^n$ schreiben müssen. Satz 2.8 a) gilt auch über jedem beliebigen Körper K und zeigt: Hat der K-Vektorraum V eine endliche Dimension n, so ist er K-isomorph zu K^n.

Analog bilden die $m \times n$-Matrizen mit Einträgen aus K einen Vektorraum

$$M(m \times n, K) = \{(a_{i,j})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} : a_{i,j} \in K\}$$

einen K-Vektorraum.

Beispiel 3.17 *Der Vektorraum*

$$K_n[X] = \{\sum_0^n a_\nu X^\nu : a_\nu \in K\}$$

der Polynome vom Grad $\leq n$ mit Koeffizienten in K ist isomorph mit K^{n+1} unter der linearen Abbildung

$$\Phi : \sum_0^n a_\nu X^\nu \mapsto (a_0, ..., a_n).$$

Mit $K[X]$ bezeichnet man den K-Vektorraum aller Polynome mit Koeffizienten in K.

Hinsichtlich des Vektorraums K^n der n-tupel und des entsprechenden Matrizenraums $M(m \times n, K)$ ist folgendes zu beachten: Begriffe und Aussagen, bei denen nicht das (euklidische) Skalarprodukt zugrunde gelegt wurde, übertragen sich auf den allgemeinen Fall eines beliebigen Grundkörpers K. Also bleiben insbesondere alle Aussagen zur Umformung einer Matrix in Zeilenstufenform, d.h., das Gauß(-Jordan)-Verfahren, zur LR-Zerlegung, zur Darstellung von linearen Abbildungen zwischen endlichdimensionalen Vektorräumen durch Matrizen, usw. gültig. Alles, was ein Skalarprodukt erfordert, also etwa Orthogonalität, ONB, Schmidtsche Orthonormalisierung, QR-Zerlegung erfordert neue Überlegungen.

Manchmal ist der Grundkörper zu einem Vektorraum nicht eindeutig bestimmt. So kann der Vektorraum $\mathbb{C}^n$ als Vektorraum über $\mathbb{C}$ oder über $\mathbb{R}$ betrachtet werden. Das hat Einfluss auf Aussagen, wie etwa

$$\begin{array}{ll} dim_{\mathbb{C}}(\mathbb{C}^n) = n & \mathbb{C}-\text{Vektorraum} \\ dim_{\mathbb{R}}(\mathbb{C}^n) = 2n & \mathbb{R}-\text{Vektorraum} \end{array}$$

Ist V ein endlich-dimensionaler K-Vektorraum mit einer Basis $\mathbf{v}_1, ..., \mathbf{v}_n$, so ist jede K-lineare Abbildung $\Phi : V \to W$ von V in einen K-Vektorraum V eindeutig festgelegt durch die Bilder $\Phi(\mathbf{v}_\nu) \in W$ der Basisvektoren (Prinzip der linearen Ausdehnung). Ist auch W endlich-dimensional mit einer Basis $\mathbf{w}_1, ..., \mathbf{w}_m$, so ist jeder Bildvektor

$$\Phi(\mathbf{v}_\nu) = \sum_{\mu=1}^{m} a_{\mu,\nu} \cdot \mathbf{w}_\mu$$

eine Linearkombination der Vektoren $\mathbf{w}_\mu$. Die Koeffizienten $a_{\mu,\nu} \in K$ kann man zu einer $m \times n$-Matrix anordnen.

Definition 3.7 *Es sei $\Phi : V \to W$ eine K-lineare Abbildung, es seien $\mathbf{v}_1, ..., \mathbf{v}_n \in V$ und $\mathbf{w}_1, ..., \mathbf{w}_m \in W$ Basen und*

$$\Phi(\mathbf{v}_\nu) = \sum_{\mu=1}^{m} a_{\mu,\nu} \cdot \mathbf{w}_\mu.$$

Dann heißt die Matrix $A = (a_{\mu,\nu}) \in M(m \times n, K)$ (wobei μ der Zeilenindex und ν der Spaltenindex ist) die darstellende Matrix *von Φ bezüglich der fixierten Basen.*

Die Spalten der darstellenden Matrix sind also die Koordinatenvektoren der Bildvektoren $\Phi(\mathbf{v}_\nu)$ in Bezug auf die Basis $\mathbf{w}_1, ..., \mathbf{w}_m$. Ein Vektor $\mathbf{x} = \sum_{\nu=1}^{n} x_\nu \mathbf{v}_\nu$ hat den Bildvektor

$$\Phi(\mathbf{x}) = \sum_{\nu=1}^{n} x_\nu \Phi(\mathbf{v}_\nu) = \sum_{\nu=1}^{n} \sum_{\mu=1}^{m} a_{\mu,\nu} x_\nu \mathbf{w}_\mu.$$

Die Koordinaten des Bildvektors $\Phi(\mathbf{x})$ bilden den Spaltenvektor des Matrix-Vektor-Produkts

$$A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

ganz genau so, wie es bei den Matrizen für eine lineare Abbildung $\mathbb{R}^n \to \mathbb{R}^m$ war. Und genauso sieht man auch

Satz 3.3 *Es seien V, W endlich-dimensionale K-Vektorräume mit Basen wie in Definition 3.7. Die Abbildung*

$$Hom_K(V, W) \to M(m \times n, K),$$

welche jeder linearen Abbildung Φ ihre darstellende Matrix A zuordnet, ist ein K-Isomorphismus. Insbesondere gilt

$$dim_K(Hom_K(V, W)) = dim_K(W) \cdot dim_K(V).$$

Aufgabe 3.15 *Es sei K ein endlicher Körper mit p Elementen und V ein zweidimensionaler Vektorraum über K.*
a) Wieviele Elemente hat V?
b) Wieviele lineare Abbildungen von V in sich gibt es?
c) Wieviele dieser Abbildungen sind bijektiv?
d) Geben Sie für $p = 2$ die Matrizen aller bijektiven linearen Abbildungen von V in sich an.

Aufgabe 3.16 *Sei p eine Primzahl und $\mathbb{F}_p$ der Körper mit p Elementen. Wieviele 2-dimensionale $\mathbb{F}_p$-Unterräume hat $\mathbb{F}_p^3$?*

Aufgabe 3.17 *Es sei K ein Körper mit p Elementen. Zeigen Sie:*
a) Die Anzahl der Elemente in der Gruppe $GL(n, \mathbb{K})$ ist

$$|GL(n, K)| := \prod_{\nu=0}^{n-1} (p^n - p^\nu).$$

b) Die Anzahl der Elemente in der Gruppe $SL(n, K)$ ist

$$\frac{1}{p-1} \cdot |GL(n, K)|.$$

(Hinweis: Aufgabe 3.5)

Aufgabe 3.18 *Bekanntlich trägt $\mathbb{C}^n$ die Struktur eines Vektorraumes über dem Körper $\mathbb{C}$, aber auch über dem Körper $\mathbb{R}$.*
a) Ergänzen Sie die Vektoren $\mathbf{b}_1 = (1, 0, 1)$ und $\mathbf{b}_2 = (1, -1, 0)$ zu einer Basis des $\mathbb{C}$-Vektorraums $\mathbb{C}^3$ und zu einer Basis des $\mathbb{R}$-Vektorraums $\mathbb{C}^3$.
b) Die Abbildung $h : \mathbb{C}^n \to \mathbb{R}^m$ sei eine lineare Abbildung der $\mathbb{R}$-Vektorräume $\mathbb{C}^n$ und $\mathbb{R}^m$. Zeigen Sie, dass

$$f : \mathbb{C}^n \to \mathbb{C}^m,\ f(\mathbf{x}) = h(\mathbf{x}) - ih(i\mathbf{x})$$

eine lineare Abbildung der $\mathbb{C}$-Vektorräume $\mathbb{C}^n$ und $\mathbb{C}^m$ ist.
c) Sei nun $f : \mathbb{C}^n \to \mathbb{C}^m$ eine lineare Abbildung der $\mathbb{C}$-Vektorräume $\mathbb{C}^n$ und $\mathbb{C}^m$. Zeigen Sie, dass es eine lineare Abbildung $h : \mathbb{C}^n \to \mathbb{R}^m$ der $\mathbb{R}$-Vektorräume $\mathbb{C}^n$ und $\mathbb{R}^m$ gibt, so dass $f(\mathbf{x}) = h(\mathbf{x}) - ih(i\mathbf{x})$ für alle $\mathbf{x} \in \mathbb{C}^n$.

Aufgabe 3.19 *Sei K ein Körper und seien $f_1, ..., f_n \in K[X]$ Polynome vom Grad $n-2$ ($2 \le n \in \mathbb{N}$). Zeigen Sie, dass für alle $x_1, ..., x_n \in K$ gilt*

$$det \begin{pmatrix} f_1(x_1) & ... & f_n(x_1) \\ \vdots & & \vdots \\ f_1(x_n) & ... & f_n(x_n) \end{pmatrix} = 0.$$

Aufgabe 3.20 *Es sei K ein Körper, V ein endlich-dimensionaler K-Vektorraum und $f : V \to V$ ein Endomorphismus vom Rang 1. Man beweise:*
a) Es gilt $f^2 = \lambda f$ mit einem eindeutig bestimmten Skalar λ aus K.
b) Es gilt

$$rang(id_V + f) = \begin{cases} dim(V), & falls \quad \lambda \neq -1, \\ dim(V) - 1, & falls \quad \lambda = -1. \end{cases}$$

Aufgabe 3.21 *Als Anwendung der letzten Aufgabe leite man eine notwendige und hinreichende Bedingung für die Invertierbarkeit der* $n \times n$*-Matrix*

$$A = \begin{pmatrix} 1+a_1^2 & a_1a_2 & \dots & a_1a_n \\ a_2a_1 & 1+a_2^2 & \dots & a_2a_n \\ \vdots & & \ddots & \vdots \\ a_na_1 & \dots & & 1+a_n^2 \end{pmatrix} \quad mit\ a_1, ..., a_n \in K$$

her.

Aufgabe 3.22 *a) Eine reelle Zahl* $r \in \mathbb{R}$ *heißt transzendent, wenn es kein Polynom* $p(X) \in \mathbb{Q}(X)$ *gibt mit* $p(r) = 0$*. Es ist eine höchst nicht-triviale Tatsache, dass die Zahl* $\pi = 3.1415...$ *transzendet ist. Beweisen Sie damit*

$$dim_{\mathbb{Q}}(\mathbb{R}) = \infty.$$

b) Beweisen Sie die Formel aus a) mit Hilfe des Begriffs der Abzählbarkeit.

3.4 Der Quotientenvektorraum

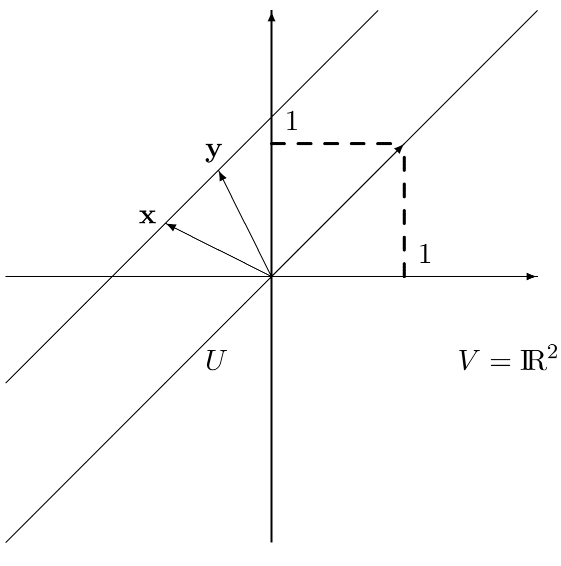
Definition 3.8 *Es sei* $U \subset V$ *ein Untervektorraum des* K*-Vektorraums* V*. Wir definieren eine* Relation *'∼' auf* V *durch*

$$\mathbf{x} \sim \mathbf{y} \quad \Leftrightarrow \mathbf{x} - \mathbf{y} \in U.$$

Relation bedeutet hier weiter nichts, als dass für je zwei Vektoren $\mathbf{x}, \mathbf{y} \in V$ definiert ist, ob sie in Relation stehen oder nicht.

Beispiel 3.18 *Es sei* $U = \mathbb{R} \cdot (1,1)^t \subset \mathbb{R}^2$*. Dann haben wir*

$$\begin{aligned} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \sim \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \quad &\Leftrightarrow \quad \begin{pmatrix} x_1 - y_1 \\ x_2 - y_2 \end{pmatrix} = c \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad c \in \mathbb{R} \\ &\Leftrightarrow \quad x_1 - y_1 = x_2 - y_2 \\ &\Leftrightarrow \quad x_1 - x_2 = y_1 - y_2 \end{aligned}$$



Die oben definierte Relation '$\sim$' ist eine *Äquivalenzrelation*, d.h., sie hat die Eigenschaften

Reflexivität:	$\mathbf{x} \sim \mathbf{x}$ für alle $\mathbf{x} \in V$
Symmetrie:	$\mathbf{x} \sim \mathbf{y} \Rightarrow \mathbf{y} \sim \mathbf{x}$ für alle $\mathbf{x}, \mathbf{y} \in V$
Transitivität:	$\mathbf{x} \sim \mathbf{y}$ und $\mathbf{y} \sim \mathbf{z} \Rightarrow \mathbf{x} \sim \mathbf{z}$ für alle $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$

Beweis dieser drei Eigenschaften: Wegen $\mathbf{x} - \mathbf{x} = \mathbf{0} \in U$ ist die Reflexitivität erfüllt. Wenn $\mathbf{x} \sim \mathbf{y}$, dann ist $\mathbf{x} - \mathbf{y} \in U$ und auch $\mathbf{y} - \mathbf{x} = -(\mathbf{x} - \mathbf{y}) \in U$. Das beweist die Symmetrie. Und aus $\mathbf{x} \sim \mathbf{y}$, $\mathbf{y} \sim \mathbf{z}$ folgt $\mathbf{x} - \mathbf{y} \in U$, sowie $\mathbf{y} - \mathbf{z} \in U$. Daraus folgt $\mathbf{x} - \mathbf{z} = (\mathbf{x} - \mathbf{y}) + (\mathbf{y} - \mathbf{z}) \in U$. Und das war die Transitivität. □

Jeder Vektor $\mathbf{x} \in V$ definiert seine *Äquivalenzklasse*

$$[\mathbf{x}] := \{\mathbf{v} \in V : \mathbf{v} \sim \mathbf{x}\} = \{\mathbf{v} \in V : \mathbf{v} - \mathbf{x} \in U\} = \mathbf{x} + U.$$

Das ist der affine Unterraum $\mathbf{x} + U \subset V$. Diese Äquivalenzklassen sind also Teilmengen von V. Der Vektor $\mathbf{x} \in \mathbf{x} + U$ heißt ein *Repräsentant* seiner Äquivalenzklasse $\mathbf{x} + U$.

Wegen der Reflexivität ist jeder Vektor $\mathbf{x}$ ein Element $\mathbf{x} \in [\mathbf{x}]$ in seiner Äquivalenzklasse. Aus der Symmetrie folgt: Wenn $\mathbf{v}$ zur Äquivalenzklasse $[\mathbf{x}]$ gehört, dann gehört auch $\mathbf{x}$ zur Äquivalenzklasse $[\mathbf{v}]$. Und die Transitivität zeigt: Gehört $\mathbf{v}$ zu $[\mathbf{x}]$ und $\mathbf{w}$ zu $[\mathbf{v}]$, dann gehört auch $\mathbf{w}$ zu $[\mathbf{x}]$.

Die Menge aller Äquivalenzklassen $[\mathbf{x}]$, $\mathbf{x} \in V$, bezeichnen wir mit V/U und nennen sie Quotientenraum (von V nach U).

Satz 3.4 *a) Die Vereinigung aller Äquivalenzklassen ist der gesamte Vektorraum V.*
b) Der Durchschnitt zweier verschiedener Äquivalenklassen ist leer.
c) Auf der Menge V/U aller Äquivalenzklassen kann man die Struktur eines K-Vektorraums definieren durch

$$\begin{array}{llcll} \textit{Addition:} & (\mathbf{x}+U)+(\mathbf{y}+U) & := & (\mathbf{x}+\mathbf{y})+U & \textit{für } \mathbf{x}, \mathbf{y} \in V \\ \textit{Multiplikation:} & c \cdot (\mathbf{x}+U) & := & (c \cdot \mathbf{x})+U & \textit{für } \mathbf{x} \in V, c \in K \end{array}$$

Beweis. a) Wegen $\mathbf{x} \in \mathbf{x} + U$ (Reflexivität) ist

$$V = \bigcup_{\mathbf{x} \in V} \{\mathbf{x}\} \subset \bigcup_{\mathbf{x} \in V} (\mathbf{x} + U).$$

b) Wenn $(\mathbf{x} + U) \cap (\mathbf{y} + U) \neq \emptyset$ ist, Dann gibt es einen Vektor $\mathbf{v} \in (\mathbf{x} + U) \cap (\mathbf{y} + U)$. Für diesen ist $\mathbf{v} = \mathbf{x} + \mathbf{u}_1 = \mathbf{y} + \mathbf{u}_2$ mit $\mathbf{u}_1, \mathbf{u}_2 \in U$. Daraus folgt $\mathbf{x} - \mathbf{y} = \mathbf{u}_2 - \mathbf{u}_1 \in U$. Also gilt $\mathbf{x} \sim \mathbf{y}$ und damit $\mathbf{x} + U = \mathbf{y} + U$.
c) Addition (und Multiplikation) der Restklassen sind repräsentantenweise definiert. Es ist zuerst zu zeigen, dass die Definition von der Wahl des Repräsentanten in der Restklasse unabhängig ist, und damit überhaupt erst sinnvoll. Seien also $\mathbf{x}' \in \mathbf{x} + U$ und $\mathbf{y}' \in \mathbf{y} + U$ weitere Repräsentanten. Dann ist $\mathbf{x}' = \mathbf{x} + \mathbf{u}_1$, $\mathbf{y}' = \mathbf{y} + \mathbf{u}_2$ mit $\mathbf{u}_1, \mathbf{u}_2 \in U$. Daraus folgt

$$(\mathbf{x}' + \mathbf{y}') + U = (\mathbf{x} + \mathbf{u}_1 + \mathbf{y} + \mathbf{u}_2) + U = (\mathbf{x} + \mathbf{y}) + (\mathbf{u}_1 + \mathbf{u}_2 + U) = (\mathbf{x} + \mathbf{y}) + U.$$

Das zeigt, dass die Addition nur von der Restklasse und nicht vom Repräsentanten abhängt. Der Beweis bei der Multiplikation geht analog.

Jetzt müssten eigentlich für die so definierte Addition und Multiplikation auf der Menge V/U die Vektorraum-Eigenschaften (Definition 3.5) nachgewiesen werden. Aber aus diesen Eigenschaften für die Repräsentanten von Restklassen folgen sie für die Restklassen. □

Satz 3.5 *Die Restklassenabbildung*

$$V \to V/U, \quad \mathbf{x} \mapsto \mathbf{x} + U$$

ist K-linear. Ihr Kern ist der Unterraum U.

Beweis. Dass die Abbildung K-linear ist, folgt daraus, dass die Vektorraum-Operationen auf V/U repräsentantenweise definiert sind. Der Nullvektor im Quotientenraum V/U ist die Restklasse $\mathbf{0}+U = U$. Der Kern der Restklassenabbildung ist deswegen die Menge aller $\mathbf{x} \in V$ mit $\mathbf{x}+U = U$, d.h. $\mathbf{x} \in U$. $\square$

Satz 3.6 (Korollar) *Ist V endlich-dimensional, so hat der Quotientenraum die Dimension*

$$dim(V/U) = dim(V) - dim(U).$$

Beweis. Weil die Restklassen-Abbildung offensichtlich surjektiv ist, folgt dies aus der Dimensionsformel Satz 2.5 b). $\square$

Satz 3.7 (Homomorphiesatz) *V und W seien K-Vektorräume und $\Phi : V \to W$ sei K-linear. Dann gibt es einen 'kanonischen' Isomorphismus*

$$V/Kern(\Phi) \to Bild(\Phi), \quad \mathbf{x} + Kern(\Phi) \mapsto \Phi(\mathbf{x}).$$

Beweis. Zunächst ist zu zeigen, dass diese Abbildung wohldefiniert ist. Seien dazu $\mathbf{x}$ und $\mathbf{x}'$ zwei Repräsentanten der gleichen Restklasse $\mathbf{x} + Kern(\Phi) = \mathbf{x}' + Kern(\Phi)$. Daraus folgt $\mathbf{x}' - \mathbf{x} \in Kern(\Phi)$ und $\Phi(\mathbf{x}') = \Phi(\mathbf{x})$. Offensichtlich ist die damit definierte Abbildung $V/Kern(\Phi) \to Bild(\Phi)$ surjektiv. Um zu zeigen, dass sie auch injektiv ist, sei $\mathbf{x} + Kern(\Phi)$ eine Restklasse, die auf $\mathbf{0}_W \in Bild(\Phi)$ abgebildet wird. Das bedeutet aber $\Phi(\mathbf{x}) = \mathbf{0}_W$, $\mathbf{x} \in Kern(\Phi)$ und $\mathbf{x} + Kern(\Phi) = Kern(\Phi) = [\mathbf{0}_V]$. $\square$

Den Isomorphismus aus Satz 3.7 kann man in die lineare Abbildung Φ 'einschieben', man sagt, Φ faktorisiert vermöge

$$\Phi : \quad V \to V/Kern(\Phi) \xrightarrow{\sim} Bild(\Phi) \subset W.$$

Wenn V ein $\mathbb{R}$-Vektorraum mit SKP ist, dann hängt Satz 3.7 folgendermaßen mit dem Homomorphiesatz Satz 2.33 zusammen: Die Abbildung $\Phi : V \to W$ definiert durch Einschränkung einen Isomorphismus

$$\varphi = \Phi | Kern(\Phi)^\perp : Kern(\Phi)^\perp \to Bild(\Phi).$$

Die Restklassenabbildung $V \to V/Kern(\Phi)$ definiert durch Einschränkung eine lineare Abbildung

$$Kern(\Phi)^\perp \to V/Kern(\Phi).$$

Wegen $Kern(\Phi)^\perp \cap Kern(\Phi) = \{\mathbf{0}\}$ ist diese injektiv. Weil beide Räume dieselbe Dimension $n - dim(Kern(\Phi))$ haben, ist sie auch surjektiv. Sie ist also ein Isomorphismus. Und man kann sich den Unterraum $Kern(\Phi)^\perp \subset V$ als eine andere Realisierung des Quotientenraums $V/Kern(\Phi)$ vorstellen.

Wir behandeln hier noch eine Anwendung des Begriffs 'Restklassen' auf die Analysis. Dazu betrachten wir den $\mathbb{R}$-Vektorraum V der auf einem Intervall $[a, b] \subset \mathbb{R}$ Riemann-integrierbaren Funktionen. Für je zwei Funktionen $f, g \in V$ ist auch ihr Produkt $f \cdot g$ auf $[a, b]$ Riemann-integrierbar (Forster I, 1. Aufl. 1976, \$ 18, Satz 7c). Deswegen ist Für $f, g \in V$

$$(f, g) := \int_a^b f(x) \cdot g(x) dx$$

wohldefiniert. Dieses Produkt (f, g) hat alle Eigenschaften eines Skalarprodukts (Definition 1.26) bis auf die Definitheit: Aus

$$(f, f) = \int_a^b f(x)^2 dx = 0$$

folgt nicht $f \equiv 0$. Deswegen ist

$$\| f \| := \sqrt{\int_a^b f(x)^2 dx}$$

auch keine Norm auf V, sondern eine sogenannte *Halbnorm*.

Satz 3.8 (Lemma) *Es sei $f \in V$ mit $\| f \| = 0$. Dann ist $(f, g) = 0$ für alle $g \in V$.*

Beweis. Wir betrachten die reelle Funktion

$$q(c) := (f + c \cdot g, f + c \cdot g) = (f, f) + 2c \cdot (f, g) + c^2(g, g), \quad c \in \mathbb{R}.$$

Es ist $q(c) \geq 0$ und $q(0) = (f, f) = 0$. Also nimmt das quadratische Polynom q in $c = 0$ ein Minimum an und es gilt

$$\frac{d}{dc} q(c) \Big|_{c=0} = 2 \cdot (f, g) = 0.$$

□

Satz 3.9 *Die Menge aller Funktionen $f \in V$ mit $\| f \| = 0$ bildet einen Untervektorraum $U \subset V$.*

Beweis. Wegen $\| c \cdot f \| = |c| \cdot \| f \|$ für $c \in \mathbb{R}$ ist mit $f \in U$ auch $c \cdot f \in U$. Seien nun f_1 und $f_2 \in U$. Dann ist

$$\| f_1 + f_2 \|^2 = (f_1, f_1) + 2 \cdot (f_1, f_2) + (f_2, f_2) = 0$$

wegen Lemma 3.8, also gehört auch $f_1 + f_2$ wieder zu U. □

Wir betrachten den Quotientenvektorraum V/U. Um die Notation zu verschlanken schreiben wir seine Elemente, die Restklassen als

$$[g] := g + U.$$

Wenn g_1, g_2 zwei Funktionen in derselben Restklasse sind, dann ist für alle $h \in V$

$$(g_1, h) - (g_2, h) = (g_1 - g_2, h) = 0 \quad \text{wegen} \quad g_1 - g_2 \in U.$$

Deswegen können wir auf dem Quotientenraum V/U

$$([g_1], [g_2]) := (g_1, g_2)$$

repräsentantenweise definieren, die Zahl $([g_1], [g_2])$ ist unabhängig von der Auswahl der Repräsentanten in $[g_1]$ und $[g_2] \in V/U$. Weil $([g_1], [g_2])$ repräsentantenweise definiert ist, besitzt dieses Produkt alle Eigenschaften eines Skalarprodukts, und hier gilt jetzt auch die Definitheit:

Sei als $[g] \in V/U$ mit $([g], [g]) = 0$. Nach Definition ist dann $g \in U$ und $[g] = 0$. Insbesondere wird durch

$$\| [g] \| := \sqrt{([g], [g])}$$

eine richtige Norm auf dem Quotientenraum V/U definiert.

Aufgabe 3.23 *Es sei V ein K-Vektorraum mit einer Basis $\mathbf{v}_1, ..., \mathbf{v}_n$ und $U \subset V$ der von $\mathbf{v}_1 + ... + \mathbf{v}_n$ erzeugte Unterraum. Bestimmen Sie eine Basis des Quotientenraums V/U.*

3.5 Der Dualraum

Definition 3.9 *Sei V ein K-Vektorraum. Eine lineare Abbildung*

$$f : V \to K$$

von V in den Grundkörper K heißt Linearform. *Der Vektorraum $Hom_K(V, K)$ der Linearformen auf V heißt der* Dualraum V^* *von V.*

Beispiel 3.19 *Sei V der Raum K^n der Spaltenvektoren $\mathbf{x} = (x_1, ..., x_n)^t, x_k \in K$. Die i-te* Koordinatenfunktion

$$f_i : \mathbf{x} \mapsto x_i$$

ist eine Linearform auf V. Man kann f_i auch schreiben als Matrizenprodut

$$f_i(\mathbf{x}) = x_i = (0, ..., 0, 1, 0, ..., 0) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

des Zeilenvektors $\mathbf{e}_i = (0, ..., 0, 1, 0, ..., 0)$ mit $\mathbf{x} \in K^n$. Allgemeiner definiert jeder Zeilenvektor $\mathbf{a} = (a_1, ..., a_n)$ eine Linearform

$$\mathbf{x} \mapsto \mathbf{a} \cdot \mathbf{x} = (a_1, ..., a_n) \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_1^n a_k x_k$$

auf V. Es ist

$$\mathbf{a} \cdot \mathbf{x} = \sum_1^n a_k x_k = \sum_1^n a_k f_k(x).$$

Beispiel 3.19 kann man so deuten: Die Transposition von Spaltenvektoren definiert eine Abbildung

$$\mathbf{a} \mapsto \mathbf{a}^t, \quad K^n \to (K^n)^*.$$

Diese ist offensichtlich linear. Das funktioniert so nur, weil wir speziell den K-Vektorraum K^n betrachtet haben. Ähnlich geht es aber auch bei $\mathbb{R}$-Vektorräumen mit Skalarprodukt.

Beispiel 3.20 *Es sei V ein $\mathbb{R}$-Vektorraum mit Skalarprodukt $(\quad,\quad)$. Für jeden Vektor $\mathbf{a} \in V$ ist die Funktion*

$$(\mathbf{a}, -) : V \to \mathbb{R}, \quad \mathbf{x} \mapsto (\mathbf{a}, \mathbf{x}),$$

eine Linearform. Und die Abbildung $\mathbf{a} \mapsto (\mathbf{a}, -)$ ist ein Homomorphismus $V \to V^*$.

Satz 3.10 (Rieszscher Darstellungssatz, erste Version) *Es sei V ein endlich-dimensionaler $\mathbb{R}$-Vektorraum mit SKP $(-, -)$. Dann gibt es zu jeder Linearform $f \in V^*$ genau einen Vektor $\mathbf{a} \in V$ mit $f = (\mathbf{a}, -)$, d.h., für alle $\mathbf{x} \in V$ gilt*

$$f(\mathbf{x}) = (\mathbf{a}, \mathbf{x}).$$

Mit anderen Worten: Die Abbildung

$$F_V : V \to V^*, \quad \mathbf{a} \mapsto (\mathbf{a}, -),$$

ist ein Isomorphismus.

Beweis. Nach Satz 2.11 ist $dim(V^*) = 1 \cdot n = dim(V)$. Es genügt deswegen zu zeigen, dass F_V injektiv ist. Sei also $\mathbf{a} \in V$ mit $F_V(\mathbf{a}) = 0$. Dann ist $(\mathbf{a}, \mathbf{x}) = 0$ für alle $\mathbf{x} \in V$, also $\mathbf{a} \in V^\perp$, und damit $\mathbf{a} = \mathbf{0}$. $\square$

Satz 3.11 (Dualbasis) *Sei $\mathbf{v}_1, ..., \mathbf{v}_n \in V$ eine Basis. Dann gibt es Linearformen $f_1, ..., f_n \in V^*$, eindeutig bestimmt durch die Eigenschaft*

$$f_i(\mathbf{v}_k) = \delta_{i,k}.$$

Die Linearformen $f_1, ..., f_n$ bilden eine Basis von V^*, die Dualbasis *zur Basis $\mathbf{v}_1, ..., \mathbf{v}_n \in V$.*

Beweis. Nach Satz 2.7 (Prinzip der linearen Ausdehnung) gibt es genau eine lineare Abbildung $f_i : V \to K$ mit $f_i(\mathbf{v}_k) = \delta_{i,k}$. Damit haben wir Linearformen $f_1, ..., f_n \in V^*$ definiert. Bezüglich der gewählten Basis von V und der 'Basis' $1 \in K$ hat f_i die darstellende $1 \times n$-Matrix $\mathbf{e}_i^t$. Diese Zeilenvektoren bilden eine Basis von $M(1 \times n, K)$, und die zugehörigen linearen Abbildungen wegen Satz 2.11 eine Basis von $Hom_K(V, K) = V^*$. $\square$

Jede lineare Abbildung $\Phi : V \to W$ definiert eine *duale Abbildung*

$$\Phi^* : \begin{cases} W^* \to V^* \\ f \mapsto f \circ \Phi \end{cases}$$

In Symbolen:

$$\begin{tikzcd} V \arrow[r, "\Phi"] \arrow[dr, "\Phi^* f = f \circ \Phi"'] & W \arrow[d, "f \in W^*"] \\ & K \end{tikzcd}$$

In Zeichen:
$f \in W^*$, $\mathbf{v} \in V \quad \Rightarrow \quad (\Phi^*(f))(\mathbf{v}) := f(\Phi(\mathbf{v}))$

Satz 3.12 *Es seien Basen*

$$\mathbf{v}_1, ..., \mathbf{v}_n \in V \textit{ und } \mathbf{w}_1, ..., \mathbf{w}_m \in W$$

festgehalten mit den zugehörigen Dualbasen

$$f_1, ..., f_n \in V^* \textit{ und } g_1, ..., g_m \in W^*.$$

Weiter sei $\Phi : V \to W$ eine lineare Abbildung. Ist $A \in M(m \times n, K)$ die beschreibende Matrix für Φ bezüglich der Basen $\mathbf{v}_1, ..., \mathbf{v}_n$ und $\mathbf{w}_1, ..., \mathbf{w}_m$, dann ist die transponierte Matrix $A^t \in M(n \times m, K)$ die beschreibende Matrix für die duale Abbildung $\Phi^* : W^* \to V^*$ bezüglich der Dualbasen.

Beweis. Es seien $A = (a_{\mu,\nu}) \in M(m \times n, K)$ die Matrix für Φ und $B = (b_{\nu,\mu}) \in M(n \times m, K)$ die Matrix für Φ^*. Dann ist

$$\Phi(\mathbf{v}_k) = \sum_{\nu=1}^{n} a_{\mu,k} \mathbf{w}_\mu, \quad \Phi^*(g_l) = \sum_{\nu=1}^{m} b_{\nu,l} f_\nu,$$

und

$$b_{k,l} = \left(\sum_{\nu=1}^{m} b_{\nu,l} f_\nu \right) (\mathbf{v}_k) = (\Phi^*(g_l))(\mathbf{v}_k) = g_l(\Phi(\mathbf{v}_k)) = g_l(\sum_{\mu=1}^{m} a_{\mu,k} \mathbf{w}_\mu) = a_{l,k},$$

also wie behauptet $B = A^t$. $\square$

Aus Satz 3.12 und $Rang(A) = Rang(A^t)$ erhält man eine einfache Folgerung, die aus der Definition von Φ^* zunächst keineswegs einsichtig ist:

Satz 3.13 (Korollar zu Satz 3.12) *Für jede lineare Abbildung Φ zwischen endlich-dimensionalen Vektorräumen gilt*

$$Rang(\Phi) = Rang(\Phi^*).$$

Unmittelbar aus der Definition ergeben sich die folgenden Rechenregeln

$$\begin{aligned}(\Psi \circ \Phi)^* &= \Phi^* \circ \Psi^* \\ (id)^* &= id \\ (\Phi^{-1})^* &= (\Phi^*)^{-1}\end{aligned}$$

Beweis. Seien $\Phi : V \to W$ und $\Psi : W \to U$ linear. Für alle $f \in U^*$ ist dann

$$(\Psi \circ \Phi)^*(f) = f \circ \Psi \circ \Phi = \Phi^*(f \circ \Psi) = \Phi^*(\Psi^*(f)).$$

Natürlich ist $(id)^*(f) = f \circ id = f$ für alle Linearformen f und deswegen $(id)^* = id$. Wenn Φ^{-1} existiert, dann ist $\Phi^{-1} \circ \Phi = id$ und deswegen $\Phi^* \circ (\Phi^{-1})^* = (\Phi^{-1} \circ \Phi)^* = (id)^* = id$. $\square$

Seien V, W endlich-dimensionale $\mathbb{R}$-Vektorräume mit SKP, das wir für beide Räume mit $(-,-)$ bezeichnen. Nach Satz 3.10 haben wir die Darstellungsisomorphismen

$$F_V : V \to V^*,\ \mathbf{v} \mapsto (\mathbf{v}, -) \quad \text{und} \quad F_W : W \to W^*,\ \mathbf{w} \mapsto (\mathbf{w}, -).$$

Sei $\Phi : V \to W$ linear mit dem adjungierten Homomorphismus $\Phi^t : W \to V$ und dem dualen Homomorphismus $\Phi^* : W^* \to V^*$. Zu jedem $\mathbf{w} \in W$ haben wir dann Linearformen in V^*

$$\begin{aligned}F_V\Phi^t(\mathbf{w}) &: \mathbf{v} \mapsto (\Phi^t(\mathbf{w}), \mathbf{v}) = (\mathbf{w}, \Phi(\mathbf{v}), \\ \Phi^* F_W(\mathbf{w}) &: \mathbf{v} \mapsto F_W(\mathbf{w})(\Phi(\mathbf{v}) = (\mathbf{w}, \Phi(\mathbf{v}).\end{aligned}$$

Beide stimmen überein. Das heißt

$$F_V \circ \Phi^t = \Phi^* \circ F_W,$$

bzw., etwas suggestiver ausgedrückt: Das Diagramm

$$\begin{array}{ccc} V & \xleftarrow{\Phi^t} & W \\ F_V \downarrow & & \downarrow F_W \\ V^* & \xleftarrow{\Phi^*} & W^* \end{array}$$

kommutiert. Identifiziert man also V mit V^* unter F_V, sowie W mit W^* unter F_W, so sind die Abbildungen Φ^t und Φ^* identisch.

Aufgabe 3.24 *Es seien $\Phi : \mathbb{R}^3 \to \mathbb{R}^3$ die lineare Abbildung mit der darstellenden Matrix*

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{pmatrix}$$

und $f, g : \mathbb{R}^3 \to \mathbb{R}$ die Linearform

$$f : (x_1, x_2, x_3) \mapsto x_1 + x_2 - x_3, \quad g : (x_1, x_2, x_3) \mapsto 3x_1 - 2x_2 - x_3.$$

Bestimmen Sie die Linearformen $\Phi^*(f) : \mathbb{R}^3 \to \mathbb{R}$ und $\Phi^*(g) : \mathbb{R}^3 \to \mathbb{R}$.

Aufgabe 3.25 *Es seien V, W Vektorräume über einem Körper K und $f : V \to W$ eine lineare Abbildung. Weiter seien V^*, W^* die zu V, W dualen Vektorräume und f^* die zu f duale Abbildung, d.h., $f^* : W^* \to V^*$, $\mu \mapsto \mu \circ f$. Man zeige: f ist genau dann Monomorphismus (=injektiv), wenn f^* Epimorphismus (=surjektiv) ist.*

Aufgabe 3.26 *Für Vektorräume V bezeichne V^* den Dualraum. Seien nun V ein K-Vektorraum, $W \subset V$ ein linearer Unterraum, $W^\perp := \{\varphi \in V^* | \, \varphi(\mathbf{w}) = 0$ für alle $\mathbf{w} \in W\}$ und V/W bzw. $V^*/W^\perp$ seien die entsprechenden Quotientenvektorräume. Definieren Sie*

$$\Phi : V^*/W^\perp \to W^* \text{ durch } \varphi + W^\perp \mapsto \varphi|_W \text{ für } \varphi \in V^*$$

und

$$\Psi : W^\perp \to (V/W)^* \text{ durch } \Psi(\varphi)(\mathbf{v} + W) = \varphi(\mathbf{v}) \text{ für } \varphi \in W^\perp.$$

Zeigen Sie: Beide, Φ und Ψ, sind wohldefiniert und Vektorraumisomorphismen.

Aufgabe 3.27 *Geben Sie zu den Vektoren*

$$\mathbf{x}_1 = (1, 0, -2),\ \mathbf{x}_2 = (-1, 1, 0),\ \mathbf{x}_3 = (0, -1, 1) \text{ aus } \mathbb{R}^3$$

die Linearformen f_i mit $f_i(\mathbf{x}_j) = \delta_{ij}$ an. Hierbei ist δ_{ij} die Kroneckerfunktion.

Aufgabe 3.28 *Sei*

$$V = \{f \in \mathbb{R}[x] | \, Grad(f) \leq 3\}$$

der $\mathbb{R}$-Vektorraum der Polynome vom Grad ≤ 3. durch

$$\begin{array}{ll} \varphi_1(f) = f(1), & \varphi_2(f) = f'(1), \\ \varphi_3(f) = f(-1), & \varphi_4(f) = f'(-1), \end{array}$$

werden Linearformen $\varphi_i : V \to \mathbb{R}$ definiert. (Dabei bezeichne f' die Ableitung von f.)
a) Zeigen Sie, dass $\varphi_1, ..., \varphi_4$ eine Basis des Dualraums V^* von V bilden.
b) Bestimmen Sie die dazu duale Basis von V.

Aufgabe 3.29 *Sei $V = \{p = a_0 + a_1 t + a_2 t^2\} \subset \mathbb{R}[t]$ der Vektorraum aller Polynome vom Grad ≤ 2. Jedes $x \in \mathbb{R}$ liefert vermöge „Auswerten in x“ ein Element $\delta_x : V \to \mathbb{R}$, $\delta_x(p) := p(x)$ des Dualraums V^* von V. Integration über einem Intervall $[a, b]$ gibt $I_a^b : V \to \mathbb{R}$, $I_a^b(p) := \int_a^b p(t)dt$, ein weiteres Element von V^*.*
a) Ist die Abbildung $\mathbb{R} \to V^*$, $x \mapsto \delta_x$ linear?
b) Zeigen Sie: Zu paarweise verschiedenen Elementen $x, y, z \in \mathbb{R}$ bilden $\delta_x, \delta_y, \delta_z$ eine Basis des Dualraums V^*.
c) Stellen Sie $I_0^2 \in V^*$ als Linearkombination von δ_0, δ_1 und δ_2 dar.

3.6 Das Vektorprodukt

Das Vektorprodukt (oder Kreuzprodukt) ist eine sehr nützliche Rechenoperation mit Vektoren im Zahlenraum $\mathbb{R}^3$.

Definition 3.10 *Es seien*

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \quad \textit{und} \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

zwei Vektoren im $\mathbb{R}^3$*. Dann heißt der Vektor*

$$\mathbf{a} \times \mathbf{b} := \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix}$$

das Vektorprodukt *der Vektoren* $\mathbf{a}$ *und* $\mathbf{b}$*.*

Beispiel 3.21 *Es ist*

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 3-2 \\ 1-3 \\ 2-1 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

Die letzten beiden Einträge des Vektors $\mathbf{a} \times \mathbf{b}$ entstehen aus dem ersten Eintrag, indem man die Einträge von $\mathbf{a}$ und $\mathbf{b}$ zyklisch aufsteigend vertauscht. Bei praktischen Rechnungen ist dies keine Erleichterung, theoretisch aber sehr nützlich.

Beispiel 3.22 *Wir berechnen das Vektorprodukt der ersten beiden kanonischen Basisvektoren*

$$\mathbf{e}_1 \times \mathbf{e}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \mathbf{e}_3.$$

Durch zyklisches Vertauschen findet man ohne weitere Rechnung

$$\mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1, \quad \mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2.$$

Die Komponenten des Vektors $\mathbf{a} \times \mathbf{b}$ sind die Unterdeterminanten

$$\begin{array}{rcl} (\mathbf{a} \times \mathbf{b})_1 & = & det^{2,3}(\mathbf{a}, \mathbf{b}) \\ (\mathbf{a} \times \mathbf{b})_2 & = & det^{3,1}(\mathbf{a}, \mathbf{b}) \\ (\mathbf{a} \times \mathbf{b})_3 & = & det^{1,2}(\mathbf{a}, \mathbf{b}) \end{array}$$

der 3×2-Matrix $(\mathbf{a}, \mathbf{b})$. Deswegen erstaunt es nicht, dass das Vektorprodukt eng mit Determinanten zusammenhängt.

Satz 3.14 *Für alle* $\mathbf{x} \in \mathbb{R}^3$ *ist das Skalarprodukt*

$$(\mathbf{a} \times \mathbf{b} \,.\, \mathbf{x}) = det(\mathbf{a}, \mathbf{b}, \mathbf{x})$$

die 3×3*-Determinante der Matrix* $(\mathbf{a}, \mathbf{b}, \mathbf{x})$ *mit den Spaltenvektoren* $\mathbf{a}, \mathbf{b}$ *und* $\mathbf{x}$*.*

Die Aussage des Satzes ist genau die Entwicklung der Determinante nach der dritten Spalte. □

Satz 3.14 identifiziert zwei Linearformen auf $\mathbb{R}^3$. Auf der linken Seite haben wir die Linearform

$$\mathbf{x} \mapsto (\mathbf{a} \times \mathbf{b} \,.\, \mathbf{x}).$$

Durch sie ist der Vektor $\mathbf{a} \times \mathbf{b}$ eindeutig bestimmt (Satz 3.10). Die rechte Seite ist die Linearform

$$\mathbf{x} \mapsto det(\mathbf{a}, \mathbf{b}, \mathbf{c}).$$

Wegen der bekannten Rechenregeln für Determinanten kann man hier meist eleganter rechnen.

Man kann Satz 3.14 auch als Merkregel auffassen, welche formal (Entwicklung nach der dritten Spalte)

$$det \begin{pmatrix} a_1 & b_1 & \mathbf{e}_1 \\ a_2 & b_2 & \mathbf{e}_2 \\ a_3 & b_3 & \mathbf{e}_3 \end{pmatrix} = (a_2b_3 - a_3b_2) \cdot \mathbf{e}_1 + (a_3b_1 - a_1b_3) \cdot \mathbf{e}_2 + (a_1b_2 - a_2b_1) \cdot \mathbf{e}_3$$

liefert. (Natürlich ist das Einsetzen eines Vektors als Matrixelement in eine Determinante streng verboten. Trotzdem ist die Regel nützlich.)

Satz 3.15 (Rechenregeln) *1) Schiefsymmetrie:* $\mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}$,
2) Bilinearität: Der Vektor $\mathbf{a} \times \mathbf{b}$ *hängt linear von* $\mathbf{a}$ *und* $\mathbf{b}$ *ab,*
3) Orthogonalität: $(\mathbf{a}.\mathbf{a} \times \mathbf{b}) = (\mathbf{b}.\mathbf{a} \times \mathbf{b}) = 0$,
4) Es ist $\mathbf{a} \times \mathbf{b} = \mathbf{0}$ *genau dann, wenn die Vektoren* $\mathbf{a}$ *und* $\mathbf{b}$ *linear abhängig (=parallel) sind,*
5) (Grassmann) $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{a}.\mathbf{c}) - \mathbf{c} \cdot (\mathbf{a}.\mathbf{b})$,
6) (Lagrange) $(\mathbf{a} \times \mathbf{b}.\mathbf{c} \times \mathbf{d}) = (\mathbf{a}.\mathbf{c}) \cdot (\mathbf{b}.\mathbf{d}) - (\mathbf{a}.\mathbf{d}) \cdot (\mathbf{b}.\mathbf{c})$,
7) Für $\mathbf{a}, \mathbf{b} \neq \mathbf{0}$ *gilt* $\| \mathbf{a} \times \mathbf{b} \| = \| \mathbf{a} \| \cdot \| \mathbf{b} \| \cdot |sin(\mathbf{a}, \mathbf{b})|$.

Beweis. Die Regeln 1) und 2) folgen aus der Bilinearität und der Schiefsymmetrie der Unterdeterminanten $det^{i,j}(\mathbf{a}, \mathbf{b})$. Regel 3) folgt mit Satz 3.14:

$$(\mathbf{a}.\mathbf{a} \times \mathbf{b}) = det(\mathbf{a}, \mathbf{a}, \mathbf{b}) = 0$$

und

$$(\mathbf{b}.\mathbf{a} \times \mathbf{b}) = det(\mathbf{b}, \mathbf{a}, \mathbf{b}) = 0.$$

4) Es ist $\mathbf{a} \times \mathbf{b} = \mathbf{0}$ genau dann, wenn $(\mathbf{a} \times \mathbf{b}.\mathbf{x}) = det(\mathbf{a}, \mathbf{b}, \mathbf{x}) = 0$ ist für alle $\mathbf{x} \in \mathbb{R}^3$. Sind $\mathbf{a}$ und $\mathbf{b}$ linear unabhängig, so kann man sie zu einer Basis $\mathbf{a}, \mathbf{b}, \mathbf{c}$ des $\mathbb{R}^3$ ergänzen. Die Matrix $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ hat Rang 3 und ihre Determinante ist $\neq 0$. Sind $\mathbf{a}$ und $\mathbf{b}$ linear abhängig, so hat für alle $\mathbf{x} \in \mathbb{R}^3$ die Matrix $(\mathbf{a}, \mathbf{b}, \mathbf{x})$ einen Rang ≤ 2 und ihre Determinante ist $= 0$.

5) Wegen der Bilinearität des Vektorprodukts genügt es, die Formel für Basisvektoren $\mathbf{a} = \mathbf{e}_i, \mathbf{b} = \mathbf{e}_j, \mathbf{c} = \mathbf{e}_k$ zu beweisen. Wegen der zyklischen Symmetrie genügt es, dies für $i = 1$ zu tun. Wenn $j = k$ ist, ist die Formel richtig, weil beide Seiten $= 0$ sind. Wenn $j \neq k$ ist, können wir wegen der Schiefsymmetrie beider Seiten in Bezug auf $\mathbf{b}$ und $\mathbf{c}$ annehmen, dass $j < k$ ist. Dann gibt es die drei Möglichkeiten

$$\begin{array}{ll} j = 1, k = 2 : \mathbf{e}_1 \times (\mathbf{e}_1 \times \mathbf{e}_2) = \mathbf{e}_1 \times \mathbf{e}_3 = -\mathbf{e}_2, & (\mathbf{e}_1.\mathbf{e}_2)\mathbf{e}_1 - (\mathbf{e}_1.\mathbf{e}_1)\mathbf{e}_2 = -\mathbf{e}_2, \\ j = 1, k = 3 : \mathbf{e}_1 \times (\mathbf{e}_1 \times \mathbf{e}_3) = -\mathbf{e}_1 \times \mathbf{e}_2 = -\mathbf{e}_3, & (\mathbf{e}_1.\mathbf{e}_3)\mathbf{e}_1 - (\mathbf{e}_1.\mathbf{e}_1)\mathbf{e}_3 = -\mathbf{e}_3, \\ j = 2, k = 3 : \mathbf{e}_1 \times (\mathbf{e}_2 \times \mathbf{e}_3) = \mathbf{e}_1 \times \mathbf{e}_1 = \mathbf{0}, & (\mathbf{e}_1.\mathbf{e}_3)\mathbf{e}_2 - (\mathbf{e}_1.\mathbf{e}_2)\mathbf{e}_3 = \mathbf{0}. \end{array}$$

6) Mit Satz 3.14 und der Formel 5) finden wir

$$\begin{aligned}
(\mathbf{a}\times\mathbf{b}.\mathbf{c}\times\mathbf{d}) &= det(\mathbf{a},\mathbf{b},\mathbf{c}\times\mathbf{d}) \\
&= det(\mathbf{c}\times\mathbf{d},\mathbf{a},\mathbf{b}) \\
&= ((\mathbf{c}\times\mathbf{d})\times\mathbf{a}\,.\,\mathbf{b}) \\
&= -(\mathbf{a}\times(\mathbf{c}\times\mathbf{d})\,.\,\mathbf{b}) \\
&= -((\mathbf{a}.\mathbf{d}\mathbf{c}-(\mathbf{a}.\mathbf{c})\mathbf{d}\,.\,\mathbf{b}) \\
&= (\mathbf{a}.\mathbf{c})(\mathbf{b}.\mathbf{d})-(\mathbf{a}.\mathbf{d})(\mathbf{b}.\mathbf{c})
\end{aligned}$$

7) Aus der Formel 6) für $\mathbf{a}=\mathbf{c}$ und $\mathbf{b}=\mathbf{d}$ folgt

$$\begin{aligned}
\parallel\mathbf{a}\times\mathbf{b}\parallel^2 &= (\mathbf{a}\times\mathbf{b}\,.\,\mathbf{a}\times\mathbf{b}) \\
&= (\mathbf{a}.\mathbf{a})(\mathbf{b}.\mathbf{b})-(\mathbf{a}.\mathbf{b})^2 \\
&= \parallel\mathbf{a}\parallel^2\cdot\parallel\mathbf{b}\parallel^2\cdot\left(1-\left(\underbrace{\frac{(\mathbf{a}.\mathbf{b})}{\parallel\mathbf{a}\parallel\cdot\parallel\mathbf{b}\parallel}}_{=cos(\mathbf{a},\mathbf{b})}\right)^2\right) \\
&= \parallel\mathbf{a}\parallel^2\cdot\parallel\mathbf{b}\parallel^2\cdot sin^2(\mathbf{a},\mathbf{b}).
\end{aligned}$$

Das ist Behauptung 7) in quadrierter Form. $\square$

Die Bilinearität (Satz 3.15.2) bedeutet insbesondere, dass für festes $\mathbf{a}\in\mathbb{R}^3$ die Abbildung

$$\mathbf{a}\times - \,:\, \mathbf{x}\mapsto\mathbf{a}\times\mathbf{x},\quad \mathbb{R}^3\to\mathbb{R}^3,$$

linear ist. Mit den Vektorprodukten $\mathbf{a}\times\mathbf{e}_i$ berechnet man ihre darstellende Matrix

$$A=\begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}.$$

Man sieht sofort $det(A)=0$ und dass $Rang(A)=2$ ist, wenn $\mathbf{a}\neq\mathbf{0}$.

Satz 3.16 *Für* $\mathbf{a}\neq\mathbf{0}$ *ist die Abbildung*

$$\mathbf{a}\times - \,:\, \mathbb{R}^3\to\mathbf{a}^\perp\subset\mathbb{R}^3$$

surjektiv. Das Urbild eines jeden Vektors $\mathbf{c}\in\mathbf{a}^\perp$ *ist eine affine Gerade mit Richtungsvektor* $\mathbf{a}$.

Beweis. Die Abbildung hat den Rang $Rang(A)=2$ und alle Bildvektoren $\mathbf{a}\times\mathbf{x}$ gehören zum Unterraum $\mathbf{a}^\perp$, welcher damit der Bildraum ist. Das Urbild eines jeden Vektors $\mathbf{c}\in\mathbf{a}^\perp$ ist ein affiner Unterraum der Dimension 1, also eine Gerade $L_{\mathbf{c}}$. Mit $\mathbf{a}\times\mathbf{x}=\mathbf{c}$ ist auch $\mathbf{a}\times(\mathbf{x}+\mathbf{a})=\mathbf{c}$. Also hat jede Gerade L_c den Richtungsbektor $\mathbf{a}$. $\square$

Das Kreuzprodukt $\mathbf{a}\times\mathbf{b}$ besitzt also folgende Eigenschaften:

a) $\mathbf{a}\times\mathbf{b}$ steht senkrecht auf $\mathbf{a}$ und $\mathbf{b}$. Falls $\mathbf{a}\times\mathbf{b}\neq\mathbf{0}$ ist, dann ist hierdurch die *Richtung* von $\mathbf{a}\times\mathbf{b}$ bestimmt.

b) Der Betrag $\parallel\mathbf{a}\times\mathbf{b}\parallel=\parallel\mathbf{a}\parallel\cdot\parallel\mathbf{b}\parallel\cdot|sin(\mathbf{a},\mathbf{b})|$ ist die Fläche des von $\mathbf{a}$ und $\mathbf{b}$ im $\mathbb{R}^3$ aufgespannten Parallelogramms. Durch diesen Betrag und die Richtung aus a) ist aber der Vektor $\mathbf{a}\times\mathbf{b}$ nur bis auf sein Vorzeichen eindeutig bestimmt.

c) Das Vorzeichen ergibt sich aus der Orientierung: Es ist

$$det(\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b}) = (\mathbf{a} \times \mathbf{b} . \mathbf{a} \times \mathbf{b}) = \parallel \mathbf{a} \times \mathbf{b} \parallel > 0,$$

falls nicht $\mathbf{a} \times \mathbf{b} = \mathbf{0}$. In diesem Fall genügen also die drei Vektoren $\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b}$ der Rechte-Hand-Regel.

Problematisch am Vektorprodukt ist, dass man seiner Vektor-Eigenschaft nicht so richtig trauen darf. Er transformiert sich anders als andere Vektoren.

Satz 3.17 *Es sei M eine invertierbare 3×3-Matrix. Für alle $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ gilt dann*

$$(M \cdot \mathbf{a}) \times (M \cdot \mathbf{b}) = det(M) \cdot (M^{-1})^t \cdot (\mathbf{a} \times \mathbf{b}).$$

Beweis. Nach Satz 3.14 ist für alle $\mathbf{x} \in \mathbb{R}^3$

$$((M \cdot \mathbf{a}) \times (M \cdot \mathbf{b}). \mathbf{x}) = det(M \cdot \mathbf{a}, M \cdot \mathbf{b}, \mathbf{x}) = det(M) \cdot det(\mathbf{a}, \mathbf{b}, M^{-1} \cdot \mathbf{x}) =$$

$$= det(M) \cdot (\mathbf{a} \times \mathbf{b} . M^{-1} \cdot \mathbf{x}) = det(M) \cdot ((M^{-1})^t \cdot \mathbf{a} \times \mathbf{b}). \mathbf{x}).$$

Daraus folgt die behauptete Gleichung. □

Im Allgemeinen transformiert sich der Vektor $\mathbf{a} \times \mathbf{b}$ ziemlich anders als seine Faktoren $\mathbf{a}$ und $\mathbf{b}$. Nur wenn M orthogonal ist, haben wir $(M^{-1})^t = M$. Dann ist also

$$\begin{array}{rcll} (M \cdot \mathbf{a}) \times (M \cdot \mathbf{b}) & = & M \cdot (\mathbf{a} \times \mathbf{b}) & \text{falls} \quad M \in SO(3) \\ (M \cdot \mathbf{a}) \times (M \cdot \mathbf{b}) & = & -M \cdot (\mathbf{a} \times \mathbf{b}) & \text{falls} \quad M \notin SO(3) \end{array}$$

Das Vektorprodukt im $\mathbb{R}^3$ hat direkte Anwendungen.

1) In der Linearen Algebra: Betrachtet werde ein einfaches, aber häufig vorkommendes homogenes LGS mit drei Unbekannten und zwei Gleichungen

$$\begin{array}{rcl} a_1 x_1 + a_2 x_2 + a_3 x_3 & = & 0 \\ b_1 x_1 + b_2 x_2 + b_3 x_3 & = & 0, \end{array}$$

wobei die Zeilenvektoren $\mathbf{a} = (a_1, a_2, a_3)^t$ und $\mathbf{b} = (b_1, b_2, b_3)^t$ linear unabhängig sind. Sein Lösungsraum hat die Dimension 1 und besteht aus allen Vektoren, welche gleichzeitig auf $\mathbf{a}$ und $\mathbf{b}$ senkrecht stehen. Er wird erzeugt von $\mathbf{a} \times \mathbf{b}$.

2) In der analytischen Geometrie:
2a) Betrachtet werde eine Ebene $E = \mathbf{v} + \mathbb{R} \cdot \mathbf{a} + \mathbb{R} \cdot \mathbf{b}$ im $\mathbb{R}^3$. Weil $\mathbf{a}$ und $\mathbf{b}$ die Ebene aufspannen, sind sie linear unabhängig, und es ist $\mathbf{a} \times \mathbf{b} \neq \mathbf{0}$ ein Normalenvektor der Ebene. Die Gleichung

$$(\mathbf{a} \times \mathbf{b} \, . \, \mathbf{x}) = 0$$

beschreibt deswegen eine Ebene durch den Nullpunkt, welche von $\mathbf{a}$ und $\mathbf{b}$ aufgespannt wird. Nach dem Struktursatz 1.4 ist die Ebene E Lösungsmenge der inhomogenen Gleichung

$$(\mathbf{a} \times \mathbf{b} \, . \, \mathbf{x}) = (\mathbf{a} \times \mathbf{b} \, . \, \mathbf{v}).$$

2b) Betrachtet werde eine Gerade $\mathbf{a} + \mathbb{R} \cdot \mathbf{v}$ im $\mathbb{R}^3$ mit Aufhängevektor $\mathbf{a}$ und Richtungsvektor $\mathbf{v}$. Der Vektor $\mathbf{w} := \mathbf{a} \times \mathbf{v}$ heißt *Momentenvektor* dieser Gerade. Die sechs Koordinaten des Vektors $(\mathbf{v}, \mathbf{w}) \in \mathbb{R}^6$ heißen *Plücker-Koordinaten* der Gerade L. Der Richtungsvektor $\mathbf{v}$ ist durch die Gerade L

nur bis auf einen konstanten Faktor $\neq 0$ eindeutig bestimmt. Deswegen sind die Plücker-Koordinaten von L auch nur bis auf einen solchen Faktor eindeutig bestimmt. Sind umgekehrt zwei Vektoren

$$\mathbf{v} \perp \mathbf{w} \in \mathbb{R}^3, \quad \mathbf{v} \neq \mathbf{0},$$

gegeben, so gibt es nach Satz 3.16 Vektoren $\mathbf{a} \in \mathbb{R}^3$ mit $\mathbf{a} \times \mathbf{v} = \mathbf{w}$. Die Menge all dieser Vektoren $\mathbf{a}$ ist eine affine Gerade im $\mathbb{R}^3$ mit Richtungsvektor $\mathbf{v}$ und Momentvektor $\mathbf{w}$.

3) In der Mechanik:
3a) Ein Vektorfeld auf $\mathbb{R}^3$ ist eine Abbildung $F : \mathbb{R}^3 \to \mathbb{R}^3$. Das zugehörige Momentenfeld ist

$$G : \mathbb{R}^3 \to \mathbb{R}^3, \quad \mathbf{x} \mapsto \mathbf{x} \times F(\mathbf{x}).$$

Beschreibt etwa F ein Kraftfeld, so heißt G das *Drehmoment*, beschreibt F ein Geschwindigkeitsfeld von Teilchen der Masse m, so heißt mG der Drehimpuls.
3b) Infinitesimale Beschreibung einer Rotation: Wir betrachten die Matrix

$$R_{\mathbf{e}_3}(\omega t) := \begin{pmatrix} cos(\omega t) & -sin(\omega t) & 0 \\ sin(\omega t) & cos(\omega t) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad t \in \mathbb{R}.$$

Sie beschreibt eine gleichförmige Rotation um die $\mathbf{e}_3$-Achse in mathematisch positiver Richtung in Abhängigkeit von der Zeit t. Dabei ist die *Winkelgeschwindigkeit* $\omega = 2\pi/T$, wo T die Dauer einer Rotation um den Winkel 2π ist. Die Geschwindigkeit eines gedrehten Punktes $\mathbf{x} \in \mathbb{R}^3$ zur Zeit $t = 0$ ist

$$\frac{d}{dt} R_{\mathbf{e}_3}(t) \cdot \mathbf{x}|_{t=0} = \begin{pmatrix} 0 & -\omega & 0 \\ \omega & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_2 \\ x_1 \\ 0 \end{pmatrix} = \omega \mathbf{e}_3 \times \mathbf{x}.$$

Wir wollen ähnlich die infinitesimale Drehung $R_{\mathbf{a}}(t)$ um eine beliebige Achse $\mathbb{R} \cdot \mathbf{a}$ beschreiben. Dabei sei $\| \mathbf{a} \| = 1$, und bei Blickreichtung in Richtung von $\mathbf{a}$ soll die Drehung im Urzeigersinn erfolgen. Wir wählen eine Matrix $U \in SO(3)$ mit $U \cdot \mathbf{a} = \mathbf{e}_3$. Dann ist nämlich

$$R_{\mathbf{a}}(t) = U^{-1} \cdot R_{\mathbf{e}_3}(t) \cdot U$$

und die Geschwindigkeit $\mathbf{v}$ in $\mathbf{x}$ zum Zeitpunkt $t = 0$

$$\frac{d}{dt} R_{\mathbf{a}}(t) \cdot \mathbf{x}\Big|_{t=0} = U^{-1} \cdot \frac{d}{dt} R_{\mathbf{e}_3}(t)\Big|_{t=0} \cdot U \cdot \mathbf{x} == U^{-1} \cdot (\omega \mathbf{e}_3 \times U\mathbf{x}) .$$

Mit der Transformationsformel (Satz 3.17) wird daraus

$$(U^{-1} \cdot \omega \mathbf{e}_3) \times (U^{-1} U \cdot \mathbf{x}) = \omega \mathbf{a} \times \mathbf{x}.$$

Hier können wir noch den *Vektor* $\omega = \omega \mathbf{a}$ *der Winkelgeschwindigkeit* einführen und finden

$$\mathbf{v} = \omega \times \mathbf{x}.$$

Aufgabe 3.30 *Bestimmen Sie eine Basis für den Lösungsraum des homogenen LGS*

$$\begin{array}{ccccccc} x & + & 2y & + & 3z & = & 0 \\ 2x & + & 3y & + & z & + & 0 \end{array}$$

Aufgabe 3.31 *Zeigen Sie: Der Punkt* $\mathbf{x} \in \mathbb{R}^3$ *hat von der Ebene:* $\mathbf{w} + \mathbb{R}\mathbf{a} + \mathbb{R}\mathbf{b}$ *den Abstand*

$$\frac{|(\mathbf{w} - \mathbf{x}.\mathbf{a} \times \mathbf{b})|}{\| \mathbf{a} \times \mathbf{b} \|}$$

und deuten Sie diesen Quotienten als

$$\textit{Höhe} = \frac{\textit{Volumen}}{\textit{Grundfläche}}$$

eines Parallelotops.

Aufgabe 3.32 *Es seien*

$$L_1 : \mathbf{v} + \mathbb{R} \cdot \mathbf{a}, \quad L_2 : \mathbf{w} + \mathbb{R} \cdot \mathbf{b}$$

zwei Geraden im $\mathbb{R}^3$ *mit linear unabhängigen Richtungsvektoren* $\mathbf{a}$ *und* $\mathbf{b}$.

a) Zeigen Sie: Die Geraden L_1 *und* L_2 *schneiden sich genau dann, wenn*

$$det(\mathbf{v} - \mathbf{w}, \mathbf{a}, \mathbf{b}) = 0.$$

b) Zeigen Sie, dass der euklidische Abstand der beiden Geraden gegeben wird durch

$$\frac{1}{\| \mathbf{a} \times \mathbf{b} \|} \cdot (\mathbf{v} - \mathbf{w} \,.\, \mathbf{a} \times \mathbf{b}).$$

c) Berechnen Sie mit der Formel aus b) den Abstand der Geraden

$$L_1 : \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \mathbb{R} \cdot \begin{pmatrix} 0 \\ 4 \\ -3 \end{pmatrix}, \quad L_2 : \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + \mathbb{R} \cdot \begin{pmatrix} -1 \\ 0 \\ 3 \end{pmatrix}.$$

Aufgabe 3.33 (Jacobi) *Zeigen Sie für alle* $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) + \mathbf{b} \times (\mathbf{c} \times \mathbf{a}) + \mathbf{c} \times (\mathbf{a} \times \mathbf{b}) = \mathbf{0}.$$

Aufgabe 3.34 *Finden Sie eine Parametrisierung der Gerade*

$$\begin{array}{lll} L_1 & \textit{mit den Plücker-Koordinaten} & (1,0,0,0,1,0), \\ L_2 & \textit{mit den Plücker-Koordinaten} & (1,-1,0,1,1,1) \end{array}$$

3.7 Euklidische und unitäre Vektorräume

Ziel dieses Abschnitts ist es, den Begriff des SKP auf $\mathbb{C}$-Vektorräume zu verallgemeinern. Dafür bezeichnen wir in diesem Abschnitt mit $\mathbb{K}$ den Körper $\mathbb{R}$ oder $\mathbb{C}$ und betrachten $\mathbb{K}$-Vektorräume. Die Konjugation

$$c \mapsto \bar{c}, \quad \mathbb{K} \to \mathbb{K},$$

bezeichnet im komplexen Fall die übliche Konjugation, im Fall $\mathbb{K} = \mathbb{R}$ ist $\bar{c} = c$.

Ausgangspunkt ist die Tatsache, dass jeder Vektorraum $\mathbb{C}^n$, aufgefasst als $\mathbb{R}^{2n}$, eine Norm besitzt, welche folgendermaßen definiert ist: Es sei

$$\mathbf{v} = (v_1, ..., v_n)^t \in \mathbb{C}^n, \quad v_\nu = a_\nu + ib_\nu.$$

Dann ist

$$\| \mathbf{v} \|^2 = \sum_{\nu=1}^{n} a_\nu^2 + b_\nu^2 = \sum_{\nu=1}^{n} v_\nu \bar{v}_\nu.$$

Die letzte Summe ist der Spezialfall $\mathbf{w} = \mathbf{v}$ folgenden Ausdrucks:

$$< \mathbf{v}, \mathbf{w} > := \sum_{\nu=1}^{n} v_\nu \bar{w}_\nu.$$

Dieser Ausdruck $< \mathbf{v}.\mathbf{w} >$ hat folgende Eigenschaften:
(i) $\mathbb{C}$-Linearität im ersten Argument:

$$< c_1\mathbf{v}_1 + c_2\mathbf{v}_2.\mathbf{w} > = c_1 < \mathbf{v}_1.\mathbf{w} > + c_2 < \mathbf{v}_2.\mathbf{w} > \quad \text{für alle } \mathbf{v}_1, \mathbf{v}_2, \mathbf{w} \in \mathbb{C}^n, c_1, c_2 \in \mathbb{C}.$$

(ii) Hermite-Symmetrie:

$$< \mathbf{w}.\mathbf{v} > = \overline{< \mathbf{v}.\mathbf{w} >} \quad \text{für alle } \mathbf{v}, \mathbf{w} \in \mathbb{C}^n.$$

(iii) Definitheit: Für alle $\mathbf{v} \in \mathbb{C}^n$ ist $< \mathbf{v}.\mathbf{v} > \in \mathbb{R}$ und $< \mathbf{v}.\mathbf{v} > \geq 0$. Dabei ist $< \mathbf{v}.\mathbf{v} > = 0$ genau dann, wenn $\mathbf{v} = \mathbf{0}$.

Aus (i) und (ii) folgt für alle $\mathbf{v}, \mathbf{w}_1, \mathbf{w}_2 \in \mathbb{C}^n, c_1, c_2 \in \mathbb{C}$

$$< \mathbf{v}.c_1\mathbf{w}_1 + c_2\mathbf{w}_2 > = \overline{< c_1\mathbf{w}_1 + c_2\mathbf{w}_2.\mathbf{v} >} = \bar{c}_1 \overline{< \mathbf{w}_1.\mathbf{v} >} + \bar{c}_2 \overline{< \mathbf{w}_2.\mathbf{v} >} = \bar{c}_1 < \mathbf{v}.\mathbf{w}_1 > + \bar{c}_2 < \mathbf{v}.\mathbf{w}_2 > .$$

Wir nennen diese Regel die *Antilinearität* im zweiten Argument.

Definition 3.11 *Es sei V ein $\mathbb{K}$-Vektorraum. Ein inneres Produkt auf V ist eine Abbildung*

$$V \times V \to \mathbb{K}, \quad \mathbf{v}, \mathbf{w} \mapsto < \mathbf{v}.\mathbf{w} > \in \mathbb{K},$$

welche die Eigenschaften (i)-(iii) hat und deswegen antilinear im zweiten Argument ist. Für $\mathbb{K} = \mathbb{R}$ ist dies genau dasselbe wie ein SKP. Einen $\mathbb{R}$-Vektorraum versehen mit einem inneren Produkt nennen wir künftig einen euklidischen *Vektorraum. Ein $\mathbb{C}$-Vektorraum versehen mit einem inneren Produkt heißt* unitärer *Vektorraum.*

Es sei V ein unitärer $\mathbb{C}$-Vektorraum mit dem inneren Produkt $< -.- >$. Dann ist $Re(< -.- >)$ eine $\mathbb{R}$-lineare reelle Funktion beider Argumente. Aus der Hermite-Symmetrie folgt die Symmetrie dieser reellen Funktion und die Definitheit ist ohnehin klar. Also ist $(-.-) := Re < -.- >$ ein inneres Produkt auf dem $\mathbb{R}$-Vektorraum V, also ein SKP. Umgekehrt ist $< -.- >$ durch das reelle innere Produkt $(-.-)$ festgelegt vermöge

$$\begin{aligned} < \mathbf{x}.\mathbf{y} > &= Re(< \mathbf{x}.\mathbf{y} >) + iIm(< \mathbf{x}.\mathbf{y} >) \\ &= (\mathbf{x}.\mathbf{y}) + iRe(-i < \mathbf{x}.\mathbf{y} >) \\ &= (\mathbf{x}.\mathbf{y}) + iRe(< \mathbf{x}.i\mathbf{y} >) \\ &= (\mathbf{x}.\mathbf{y}) + i(\mathbf{x}.i\mathbf{y}). \end{aligned}$$

Satz 3.18 *Ein inneres Produkt* $< -.- >$ *auf dem* $\mathbb{K}$*-Vektorraum* V *definiert eine Norm*

$$\| \mathbf{v} \| := \sqrt{< \mathbf{v}.\mathbf{v} >}$$

auf V.

Beweis. Nur für den komplexen Fall ist die Aussage neu. Sei deswegen V mit dem inneren Produkt $< -.- >$ ein unitärer Raum. Wenn wir für den Moment die Norm des komplexen inneren Produkts mit

$$\| \mathbf{x} \|_{\mathbb{C}} = \sqrt{< \mathbf{x}.\mathbf{x} >}$$

abkürzen, und die Norm des zugehörigen reellen inneren Produkts $Re < -.- >$ mit $\| \mathbf{x} \|_{\mathbb{R}}$, so ist

$$\| \mathbf{x} \|_{\mathbb{C}} = \| \mathbf{x} \|_{\mathbb{R}}$$

weil $< \mathbf{x}.\mathbf{x} >$ reell ist. Beide Normen sind also gleich.

Zu zeigen sind die folgenden Eigenschaften.

(iv) Definitheit: Das ist Teil (iii) der Definition eines inneren Produkts.

(v) Homogenität: Für alle $c \in \mathbb{K}$, also eventuell sogar für $c \in \mathbb{C}$, gilt wegen der Antilinearität im zweiten Argument

$$\| c \cdot \mathbf{v} \| = \sqrt{< c \cdot \mathbf{v}.c \cdot \mathbf{v} >} = \sqrt{c\bar{c} \cdot < \mathbf{v}.\mathbf{v} >} = |c| \cdot \| \mathbf{v} \| .$$

(vi) Cauchy-Schwarz-Ungleichung: Für das innere Produkt $Re(< -.- >)$ auf dem $\mathbb{R}$-Vektorraum V gilt die reelle Cauchy-Schwarz-Ungleichung

$$|Re(< \mathbf{x}.\mathbf{y} >)| = |(\mathbf{x}.\mathbf{y})| \leq \| \mathbf{x} \| \cdot \| \mathbf{y} \|, \quad \mathbf{x}, \mathbf{y} \in V.$$

Sei $c := < \mathbf{x}.\mathbf{y} >$. Dann ist

$$< \bar{c}\mathbf{x}.\mathbf{y} > = \bar{c} \cdot c$$

reell. Mit der reellen Cauchy-Schwarz-Ungleichung finden wir deswegen

$$|c| \cdot | < \mathbf{x}.\mathbf{y} > | = | < \bar{c}\mathbf{x}.\mathbf{y} > | \leq \| \bar{c}\mathbf{x} \| \cdot \| \mathbf{y} \| = |c| \cdot \| \mathbf{x} \| \cdot \| \mathbf{y} \| .$$

Für $c = 0$ ist die Ungleichung trivial. Falls $c \neq 0$ ist, können wir $|c|$ kürzen und erhalten die Aussage.

(vii) Dreiecksungleichung: Sie gilt für die Norm des reellen inneren Produkts $Re < -.- >$ und, weil die Norm des reellen und des komplexen inneren Produkts übereinstimmen, auch für das komplexe.

Für das reelle innere Produkt gilt die Polarisationsformel

$$\| \mathbf{x} + \mathbf{y} \| = \| \mathbf{x} \|^2 + 2(\mathbf{x}.\mathbf{y}) + \| \mathbf{y} \|^2 .$$

Sie zeigt, dass die Abstände das innere Produkt bestimmen (aus längentreu folgt winkeltreu). Für das komplexe innere Produkt lautet diese Formel

$$\| \mathbf{x} + \mathbf{y} \| = \| \mathbf{x} \|^2 + 2Re(< \mathbf{x}.\mathbf{y} >) + \| \mathbf{y} \|^2 .$$

Für $\mathbf{x}, \mathbf{y}$ aus einem $\mathbb{K}$-Vektorraum V definieren wir genauso wie früher bei $\mathbb{R}$-Vektorräumen

$$\mathbf{x} \perp \mathbf{y} \quad \text{falls} \quad < \mathbf{x}.\mathbf{y} > = 0.$$

Damit ergibt sich der abstrakte Pythagoras auch in der allgemeinen Situation genau wie früher (Satz 1.29). Für $A \subset V$ wird auch $A^\perp \subset V$ genau wie früher definiert. Allerdings ist des jetzt ein $\mathbb{K}$-Untervektorraum, also im komplexen Fall ein komplexer Untervektorraum. Weil aus $< \mathbf{a}.\mathbf{x} > = 0$

folgt, dass $Re(<\mathbf{a}.\mathbf{x}>) = 0$ ist, ist das komplexe $A^{\perp}_{\mathbb{C}}$ im reellen $A^{\perp}_{\mathbb{R}}$ enthalten. Falls aber $A = U \subset V$ ein komplexer Untervektorraum ist, stimmen beide orthogonalen Komplemente überein.

Beweis. Zu zeigen ist $U^{\perp}_{\mathbb{R}} \subset U^{\perp}_{\mathbb{C}}$. Sei also $\mathbf{x} \in U^{\perp}_{\mathbb{R}}$, d.h., $Re(<\mathbf{x}.\mathbf{u}>) = 0$ für alle $\mathbf{u} \in U$. Weil U ein komplexer Untervektorraum ist, dann ist mit $\mathbf{u} \in U$ auch $i\mathbf{u} \in U$. Damit folgt

$$<\mathbf{x}.\mathbf{u}>= Re<\mathbf{x}.\mathbf{u}> + iRe<\mathbf{x}.i\mathbf{u}>) = 0.$$

□

Damit ist auch die Orthogonalprojektion bezüglich des komplexen und des zugehörigen reellen inneren Produkts identisch. Denn

$$(\mathbf{x}-\mathbf{u}) \perp U$$

bedeutet in beiden Fällen das Gleiche. Man kann hier auch mit dem minimalen Abstand argumentieren: Weil reelle und komplexe Norm identisch sind, sind auch die Abstände $\parallel \mathbf{x}-\mathbf{u} \parallel$ in beiden Fällen dasselbe.

Aus der reellen Theorie folgt:

Eindeutigkeit: Wenn die Orthogonalprojektion $P_U(\mathbf{x})$ existiert, dann ist sie eindeutig bestimmt.

Existenz: Ist $U \subset V$ ein endlich-dimensionaler $\mathbb{C}$-Untervektorraum, so existiert die Orthogonalprojektion $P_U : V \to U$.

Ist $\mathbf{u}_1, ..., \mathbf{u}_m$ eine $\mathbb{C}$-Basis des komplexen Untervektorraums $U \subset V$, so definiert man analog zum reellen Fall ihre Gram-Matrix

$$A = (<\mathbf{u}_i.\mathbf{u}_j>)_{i,j=1,\ldots,m}.$$

Genau wie im Reellen zeigt man auch hier, dass diese Matrix nicht-singulär ist:

Hätte A einen Rang $< m$, so gäbe es einen Vektor $\mathbf{c} \in \mathbb{C}^m, \mathbf{c} \neq \mathbf{0}$, so, dass $\mathbf{c}^t \cdot A = \mathbf{0}$. Dann wäre auch

$$<\sum c_\mu \mathbf{u}_\mu . \sum c_\mu \mathbf{u}_\mu >= \mathbf{c}^t \cdot A \cdot \bar{\mathbf{c}} = 0.$$

Mit der Definitheit folgt $\sum c_\mu \mathbf{u}_\mu = \mathbf{0}$ und daraus $c_1 = ... = c_m = 0$.

Satz 3.19 *Es sei $\mathbf{u}_1, ..., \mathbf{u}_m \in U$ eine $\mathbb{C}$-Basis des komplexen Untervektorraums $U \subset V$. Dann sind äquivalent:*

i) $\mathbf{u} = P_U(\mathbf{x})$*;*

ii) $\mathbf{u} = \sum_{\mu=1}^m c_\mu \mathbf{u}_\mu$ *mit* $\mathbf{c}^t \cdot A = \mathbf{b}^t$*, wobei* $\mathbf{b} = (<\mathbf{x}.\mathbf{u}_\mu>)_{\mu=1,\ldots,m}$.

Beweis. Die Linearkombination $\mathbf{u} = \sum c_\mu \mathbf{u}_\mu$ ist die Orthogonalprojektion $P_U(\mathbf{x})$ genau dann, wenn

$$<\mathbf{x}-\mathbf{u}\,.\mathbf{u}_\nu>= 0 \quad \text{für} \quad \nu = 1, ..., m.$$

Äquivalent dazu ist, dass für alle ν

$$<\mathbf{x}-\sum c_\mu \mathbf{u}_\mu\,.\mathbf{u}_\nu>= 0, \quad <\mathbf{x}.\mathbf{u}_\nu> - \sum c_\mu <\mathbf{u}_\mu.\mathbf{u}_\nu>= 0,$$

bzw.

$$\mathbf{b}^t = \mathbf{c}^t \cdot A.$$

□

Auch das Orthonormalisierungs-Verfahren von Gram-Schmidt lässt sich (mit Beweis) wörtlich übertragen. Allerdings muss man wegen der Hermite-Symmetrie beim Konstruktions-Schritt

$$\mathbf{u}'_{k+1} = \mathbf{v}_{k+1} - <\mathbf{v}_{k+1}.\mathbf{u}_1> \mathbf{u}_1 - ... - <\mathbf{v}_{k+1}.\mathbf{u}_k> \mathbf{u}_k$$

auf die Reihenfolge der Vektoren in den inneren Produkten achten.

Definition 3.12 *Es sei V ein $\mathbb{K}$-Vektorraum mit innerem Produkt und $\Phi : V \to V$ sei $\mathbb{K}$-linear. Genau wie im reellen Fall heißt $\Phi^\dagger$ adjungiert zu Φ, wenn*

$$< \Phi(\mathbf{x}).\mathbf{y} >=< \mathbf{x}.\Phi^\dagger(\mathbf{y}) > .$$

Im komplexen Fall definiert eine endliche ONB in V einen Isomorphismus $V \to \mathbb{C}^n$, unter dem das innere Produkt $< \mathbf{x}.\mathbf{y} >$ in die Standard-Form $\sum_1^n x_\nu \bar{y}_\nu$ übergeht. Es sei A die darstellende Matrix der durch Φ definierten $\mathbb{C}$-linearen Abbildung $\mathbb{C}^n \to \mathbb{C}^n$. Dann ist

$$A^\dagger := \bar{A}^t = \overline{A^t}$$

die Matrix der adjungierten Abbildung. Genau wie früher (mit einer Modifikation bei Multiplikation mit einem Skalar) sehen wir

$$(A^\dagger)^\dagger = A, \quad (A+B)^\dagger = A^\dagger + B^\dagger, \quad (\lambda A)^\dagger = \bar{\lambda} A^\dagger, \quad (AB)^\dagger = B^\dagger A^\dagger.$$

Weiter ist

$$Rang(A) = Rang(\bar{A}) = Rang(A^\perp),$$

wobei nur $Rang(A) = Rang(\bar{A})$ zu zeigen ist. Aber, wenn für Spaltenvektoren $\mathbf{a}_i$ eine lineare Relation

$$\sum_i c_i \cdot \mathbf{a}_i = \mathbf{0}$$

gilt, dann folgt durch Konjugation die Relation

$$\sum_i \bar{c}_i \cdot \bar{\mathbf{a}}_i = \mathbf{0}$$

für die entsprechenden Spaltenvektoren $\bar{\mathbf{a}}_i$ von $\bar{A}$. Die Spaltenvektoren $\mathbf{a}_i$ sind also genau dann linear abhängig, wenn es die entsprechenden Spaltenvektoren $\bar{\mathbf{a}}_i$ sind.

Analog zu orthogonalen, bzw. symmetrischen reellen Matrizen definiert man jetzt:

Definition 3.13 *Die komplexe $n \times n$-Matrix heißt* unitär, *wenn*

$$U^{-1} = U^\dagger.$$

die komplexe $n \times n$-Matrix H heißt hermitesch *(oder selbst-adjungiert), wenn*

$$H = H^\dagger.$$

Für unitäre Matrizen U gilt

$$< U \cdot \mathbf{x}.U \cdot \mathbf{y} >=< \mathbf{x}.U^\dagger U \cdot \mathbf{y} >=< \mathbf{x}.\mathbf{y} >$$

und für hermitesche

$$< H \cdot \mathbf{x}.\mathbf{y} >=< \mathbf{x}.H \cdot \mathbf{y} > .$$

Wie früher folgt für selbst-adjungierte $\mathbb{K}$-lineare Abbildungen $\Phi = \Phi^\dagger$, dass $Kern(\Phi) \perp Bild(\Phi)$. Denn für $\mathbf{x} - Kern(\Phi)$ und $\Phi(\mathbf{y}) \in Bild(\Phi)$ ist

$$< \mathbf{x}.\Phi(\mathbf{y}) >=< \Phi(\mathbf{x}).\mathbf{y} >=< \mathbf{0}.\mathbf{y} >= 0.$$

Damit erhält man auch genau wie bei Satz 2.27 b): Eine Projektion $P : V \to V$ ist genau dann eine Orthogonalprojektion, wenn $P = P^\dagger$ selbst-adjungiert ist.

Aufgabe 3.35 *Beweisen Sie: Die Determinante einer unitären Matrix hat Betrag* 1*, die Determinante einer hermiteschen Matrix ist reell.*

Aufgabe 3.36 *Eine komplexe* $n \times n$*-Matrix* N *heißt normal, wenn*

$$N \cdot N^\dagger = N^\dagger \cdot N$$

gilt.

a) Zeigen Sie: Alle unitären, hermiteschen, reellen orthogonalen und reellen symmetrischen Matrizen sind normal.

b) Welche der folgenden drei Matrizen sind unitär, bzw. hermitesch, bzw. normal:

$$A = \begin{pmatrix} 1 & i \\ 1 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}, \quad C = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}.$$

Aufgabe 3.37 *a) Zeigen Sie, dass die sogenannten Pauli-Matrizen*

$$\sigma_0 := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_4 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

eine Basis des $\mathbb{C}$*-Vektorraums der komplexen* 2×2*-Matrizen bilden.*

b) Sind die Pauli-Matrizen hermitesch? Sind sie unitär?

Aufgabe 3.38 *Es sei* V *ein endlich-dimensionaler unitärer Vektorraum und* $\Phi : V \to V$ *sei* $\mathbb{C}$*-linear. Zeigen Sie:*

$$Rang(\Phi) = Rang(\Phi^\dagger \circ \Phi) = Rang(\Phi \circ \Phi^\dagger).$$

Aufgabe 3.39 *Zeigen Sie: Das Produkt zweier selbst-adjungierter Abbildungen ist genau dann wieder selbst-adjungiert, wenn diese Abbildungen kommutieren.*