Handout
Lineare Algebra 1

Dirk Lorenz

Inhaltsverzeichnis

Vorbemerkung 3

1 Mengen 5

2 Äquivalenzrelationen 7

3 Abbildungen 8

4 Körper 11

5 Vektorräume 15

6 Lineare Unabhängigkeit 19

7 Basen 23

8 Die Dimension eines Vektorraums 26

9 Matrizen und lineare Gleichungssysteme 30

10 Der Lösungsraum, Kern, Bild und Rang 34

11 Der Gaußsche Eliminationsalgorithmus 38

12 Lineare Abbildungen 42

13 Darstellung von linearen Abbildungen durch Matrizen 46

14 Matrix Multiplikation und Inversion 50

15 Elementarmatrizen und der Gauß-Algorithmus 54

16 Faktorräume 60

17 Der Homomorphiesatz 63

18 Der Dualraum 64

19 Determinaten 67

20 Permutationen und die Leibniz-Formel **72**

21 Eigenwerte, Eigenvektoren **76**

22 Berechnung von Eigenwerten und das charakteristische Polynom **81**

23 Bilinearformen und Skalarprodukte **87**

24 Orthogonalität und Normen **91**

25 Orthogonale Komplemente **95**

26 Selbstadjungierte Abbildungen **100**

27 Hauptachsentransformation **105**

28 Matrixgruppen **107**

Vorbemerkung

Dies ist das Handput zur Vorlesung „Lineare Algebra 1“, gehalten an der TU Braunschweig im Wintersemester 2018/19. Es keinerlei Anspruch, auch ohne Besuch der Vorlesung verständlich zu sein. Diese Version enthält mit an Sicherheit grenzender Wahrscheinlichkeit zahlreiche Fehler; seien Sie also aufmerksam beim Lesen.

Die Inhalte in diesem Handout sind in zahlreichen Büchern zu finden und ich habe unter anderem folgende Bücher benutzt:

- *Lineare Algebra*, 8., aktualisierte Auflage, Albrecht Beutelspacher, Springer Spektrum, 2014, https://link.springer.com/book/10.1007/978-3-658-02413-0
- *Lineare Algebra*, 18., aktualisierte Auflage, Gerd Fischer, Springer Spektrum, 2014, https://link.springer.com/book/10.1007/978-3-658-03945-5,
- *Lineare Algebra*, 12., überarbeitete Auflage, Hans-Joachim Kowalsky Gerhard O. Michler, de Gruyter, 2003
- *Lineare Algebra 1*, 4. Auflage, Falko Lorenz, Spektrum Akademischer Verlag, 2004

Im Text sind ab und zu Absätze in leichtem grau leicht abgesetzt. Dabei handelt es sich um kurze Argumente, die das darüber Stehende begründen oder weiter erläutern. Ist Ihnen etwas vor einem solchen Abschnitt klar, dann können Sie den grauen Absatz getrost überspringen.

> Es kann sich natürlich trotzdem lohnen, diese Absätze zu lesen, da vielleicht an anderes Argument als Ihres gegeben wird.

Zahlreiche Randbemerkungen erläutern oder illustrieren die Konzepte oder erinnern an etwas, was evtl. schon bekannt war.

Neben Kapiteln enthält dieses Handout auch noch die Gliederung nach den einzelnen Vorlesungen. Diese sind in der Kopfzeile durchnummeriert und mit Datum versehen, so dass Sie leicht mit Ihrer Vorlesungsmitschrift abgleichen können. Selbstverständlich sind Ihre Mitschrift und dies Skript fast nie wortgleich; man bringt eben andere Sätze an die Tafel, als in ein Skript.

Braunschweig, den 27. März 2019 Dirk Lorenz
d.lorenz@tu-braunschweig.de

1 Mengen

Naive Definition von *Menge*: Mengen bestehen aus *Elementen* und wenn x ein Element einer Menge M ist, schreiben wir

$$x \in M,$$

Lies: „x ist ein Element von M“ oder „x ist in M“

(ist x nicht in M, schreiben wir $x \notin M$).

Sind alle Elemente einer Menge N auch Elemente einer anderen Menge M, dann nennen wir N *Teilmenge* von M und schreiben

$$N \subset M.$$

Lies „N ist Teil- (oder Unter-)-menge von M“. Manchmal wird $\subseteq$ benutzt.

Wir können Mengen beschreiben, in dem wir *die Elemente aufzählen*, oder *die Eigenschaften ihrer Elemente angeben*:

$$M = \{a, b, c, \dots\}$$
$$M = \{x \mid E(x)\}.$$

Der Ausdruck $\{x \mid \dots\}$ wird gelesen als „Die Menge aller x mit der Eigenschaft ...“.

Das Symbol $:=$ bedeutet, dass der Ausdruck links durch den Ausdruck rechts definiert wird.

Beispiel.

Die leere Menge.

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

Natürliche Zahlen.

$$\mathbb{N} := \{0, 1, 2, 3, \dots, \}.$$

Ganze Zahlen.

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

Rationale Zahlen.

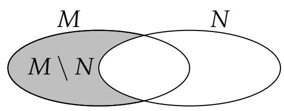
$$\mathbb{Q} := \{\tfrac{p}{q} \mid p \in \mathbb{Z},\ q \in \mathbb{N},\ q \neq 0\}$$

Die **reellen Zahlen** werden in der Analysis beschrieben. △

Definition 1.1. Für zwei Mengen M und N definieren wir:

Mengendifferenz:

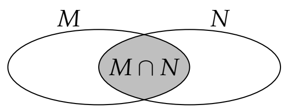
$$M \setminus N := \{x \in M \mid x \notin N\}$$



$M \setminus N$ liest man als „M ohne N“.

Schnitt:

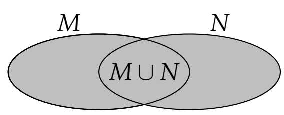
$$M \cap N := \{x \in M \mid x \in N\} = \{x \mid x \in M \text{ und } x \in N\}$$



$M \cap N$ liest man als „M geschnitten mit N“.

Vereinigung:

$$M \cup N := \{x \mid x \in M \text{ oder } x \in N\}$$



$M \cup N$ liest man als „M vereinigt mit N“.

Definition 1.2. Für Mengen $M_1, \dots, M_n$ die nicht leer sind, definieren wir das *kartesische Produkt* als Menge der Folgen der Länge n (oder den n-Tupeln) mit Folgengliedern aus dem Mengen M_i, also

$$M_1 \times \cdots \times M_n := \{(x_1, \dots, x_n) \mid x_i \in M_i,\ i = 1, \dots, n\}.$$

$M \times N$ liest man „Produkt von M und N“ oder „Kreuzprodukt von M und N“.

Beispiel.
Für $M = \{a, b, c\}$ und $N = \{x, y\}$ ist

$$M \times N = \{(a, x), (b, x), (c, x), (a, y), (b, y), (c, y)\}.$$

△

Definition 1.3. Eine Menge heißt *endlich*, wenn sie nur endlich viele Elemente hat. Für eine endliche Menge M bezeichnen wir die Anzahl der Elemente mit $|M|$.

$|M|$ liest man „Anzahl der Elemente in M“.

Satz 1.4. *Für zwei Mengen M und N gilt*

$$|M \cup N| = |M| + |N| - |M \cap N|$$

und

$$|M \times N| = |M|\,|N|.$$

Ist N eine Teilmenge von M, so gilt

$$|M \setminus N| = |M| - |N|.$$

Beweis.
In der Summe $|M| + |N|$ werden die Elemente von $M \cap N$ genau zweimal gezählt, alle anderen genau einmal, d.h. $|M| + |N| = |M \cup N| + |M \cap N|$.

Um die Elemente in $M \times N$ zu zählen, bemerken wir, dass $M \times N$ für jedes Element $x \in M$ die Elemente (x, y) mit $y \in N$ enthält, und dies sind genau $|N|$. Zusammen ergeben sich $|M|\,|N|$ Elemente von $M \times N$.

Die Elemente von $M \setminus N$ sind gerade die aus M ohne die aus N, und da wir annehmen, dass N ganz in M enthalten sind, folgt die letzte Formel. □

2 Äquivalenzrelationen

Definition 2.1. Eine *Relation* auf einer Menge X ist eine Teilmenge ρ von $X \times X$ und wir sagen „x steht in Relation mit y" falls $(x, y) \in \rho$. Wir schreiben auch $x \, \rho \, y$.

Wir könnten auch Relationen zwischen Elementen verschiedener Mengen X und Y definieren; dies wären Teilmengen $\rho \subset X \times Y$.

Oft werden Relationen mit $\sim$ bezeichnet, also $x \sim y$ statt $x \, \rho \, y$.

Beispiel.

- „Gleichheit" ist eine Relation (definiert durch $(x, y) \in \rho$ genau dann, wenn $x = y$).
- Andere Relationen sind zum Beispiel
  - $x \sim y$, genau dann, wenn x mit y verwandt ist, oder
  - $x \sim y$, wenn x schon mal einen Brief an y geschrieben hat.

$\triangle$

Definition 2.2. Eine Relation $\sim$ auf X heißt

- *reflexiv*, wenn
$$\forall x \in X: \; x \sim x,$$
- *symmetrisch*, wenn
$$\forall x, y \in X: \; x \sim y \iff y \sim x,$$
- *transitiv*, wenn
$$\forall x, y, z \in X: x \sim y, y \sim z \implies x \sim z.$$

„$\forall$" liest sich „für alle"
„:" liest sich wie „|" meist als „so dass" oder auch „gilt"

„ $\iff$ " liest sich „genau dann, wenn"

„ $\implies$ "liest sich „impliziert".
$A \implies B$ liest sich auch als „aus A folgt B"oder „B folgt aus A"

Eine Relation die reflexiv, symmetrisch und transitiv ist, heißt *Äquivalenzrelation.*

Beispiel.

- Die Gleichheitsrelation ist eine Äquivalenzrelation.
- Auf der Menge der Geraden in der Ebene ist
$$x \parallel y \iff x \text{ parallel zu } y$$
eine Äquivalenzrelation, die Relation
$$x \perp y \iff x \text{ senkrecht auf } y$$
hingegen nicht.

$\triangle$

Definition 2.3. Sei X eine Menge und $\sim$ eine Äquivalenzrelation auf X. Dann ist

$$A(x) := \{y \in X \mid y \sim x\}$$

die *Äquivalenzklasse* von x.

Satz 2.4. *Sind x und y äquivalent, so sind die Äquivalenzklassen von x und y gleich.*

Beweis.
Wir müssen zeigen, dass aus $x \sim y$ folgt, dass $A(x) = A(y)$. Um die Gleichheit der Mengen zu zeigen, zeigen wir $A(x) \subset A(y)$ und $A(y) \subset A(x)$.

„$A(x) \subset A(y)$“: Sei $z \in A(x)$, d.h. es gilt $z \sim x$. Da $x \sim y$, folgt wegen der Transitivität $z \sim y$, also $z \in A(y)$. Wir haben gezeigt: Jedes Element von $A(x)$ ist auch Element von $A(y)$.

Die Inklusion „$A(y) \subset A(x)$“zeigt man genau so, indem man die Rolle von x und y vertauscht. □

Die Aussage des Satzes können wir auch anders formulieren:

- Zwei Äquivalenzklassen sind entweder gleich oder disjunkt, oder
- jedes Element von X liegt in genau einer Äquivalenzklasse, oder
- die Äquivalenzklassen von X bilden eine Partition von X.

X und Y *disjunkt*, wenn $X \cap Y = \emptyset$

Eine *Partition* von X ist eine Zerlegung von X in disjunkte Teilmengen.

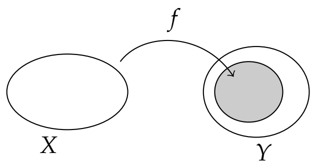
Beispiel.

- Auf $\mathbb{N}$ definieren wir $m \sim n$ falls $n - m$ durch 2 teilbar ist. Dies zerlegt $\mathbb{N}$ in zwei Äquivalenzklassen, nämlich die geraden und die ungeraden Zahlen.
- Auf der Menge der Geraden in der Ebene definieren wir zwei Geraden als äquivalent, wenn sie parallel sind. Die zugehörigen Äquivalenzklassen sind die „Richtungen“.

△

3 Abbildungen

Definition 3.1. Seien X und Y zwei Mengen. Ein *Abbildung* f von X nach Y ordnet jedem Element $x \in X$ genau ein Element $y \in Y$ zu. Wir schreiben $f : X \to Y$ und $y = f(x)$. Die Menge X heißt *Definitionsbereich*, Y heißt *Bildbereich*.



Etwas präziser: Eine Abbildung f ist eine Relation $f \subset X \times Y$ mit der Eigenschaft, dass es zu jedem $x \in X$ genau ein $y \in Y$ mit $(x, y) \in f$ gibt. Man sagt auch, Abbildungen sind „links-total und rechts-eindeutig“.

Beispiel.

- $f : \mathbb{Z} \to \mathbb{N}$, $f(z) = z^2$ ist eine Abbildung, ebenso wie $f : \mathbb{Z} \to \mathbb{Z}$, $f(z) = z^2$.

- $f : \{1,2,3\} \to \{1,2,3\}$ definiert durch $f(1) = 2$, $f(2) = 3$ und $f(3) = 1$ ist eine Abbildung.

  Würden wir nur $f(1) = 2$ und $f(2) = 3$ definieren, so wäre f keine Abbildung (aber immer noch eine Relation).

- Die Multiplikation von reellen Zahlen ist eine Abbildung, nämlich
$$\mathbb{R} \times \mathbb{R} \to \mathbb{R}, \quad (r,s) \mapsto r \cdot s.$$

$\triangle$

Auf jeder Menge X gibt es die *identische Abbildung*

$$\mathrm{id}_X : X \to X, \quad \mathrm{id}_X(x) = x,$$

oft schreiben wir id statt id_X, wenn die Menge X aus dem Kontext klar ist.

Definition 3.2. Sei f eine Abbildung von X nach Y.

- f heißt *injektiv*, wenn aus $x_1 \neq x_2$ folgt, dass $f(x_1) \neq f(x_2)$ gilt.
- f heißt *surjektiv*, wenn es für jedes $y \in Y$ ein $x \in X$ gibt mit $f(x) = y$.
- f heißt *bijektiv*, wenn f injektiv und surjektiv ist.

Eine andere Formulierung: f ist $\left\{\begin{array}{c}\text{surjektiv}\\ \text{injektiv}\\ \text{bijektiv}\end{array}\right\}$, wenn die Gleichung $f(x) = y$ für jedes y $\left\{\begin{array}{c}\text{mindestens}\\ \text{höchstens}\\ \text{genau eine}\end{array}\right\}$ eine Lösung hat.

Beispiel.

- $f : \mathbb{Z} \to \mathbb{N}$, $f(z) = z^2$ ist nicht injektiv (denn $1 \neq -1$, aber $f(1) = 1 = (-1)^2 = f(-1)$) und nicht surjektiv, denn es gibt kein $z \in Z$ mit $z^2 = 3$.

  Anders gesagt: Die Gleichung $z^2 = 1$ hat *zwei* Lösungen in $\mathbb{Z}$, und $z^2 = 3$ hat *keine* Lösung in $\mathbb{Z}$.

- Ändern wir den Definitionsbereich zu $f : \mathbb{N} \to \mathbb{N}$, $f(z) = z^2$, so ist f injektiv, aber immer noch nicht surjektiv.

- Die Abbildung $f : \mathbb{R} \to \mathbb{R}_+$ $(:= \{x \in \mathbb{R} \mid x \geq 0\})$, $f(x) = |x|$ ist nicht injektiv, aber surjektiv.

$\triangle$

Im Allgemeinen kann man Abbildungen z.B. durch Verkleinern ihres Defintionsbereiches injektiv machen und z.B. durch Verkleinern des Bildbereiches surjektiv machen.

Definition 3.3. Zu einer Abbildung $f : X \to Y$ definieren wir das *Bild von f* als

$$\mathrm{Bild}(f) := \{y \in Y \mid \exists x \in X : f(x) = y\}.$$

Für jedes $f : X \to Y$ ist $f : X \to \mathrm{Bild}(f)$ surjektiv.

Das Zeichen $\exists$ liest man „es existiert“ oder „existiert“.

Definition 3.4. Ist $g : X \to Y$ und $f : Y \to Z$, dann ist $f \circ g$ die Abbildung von X nach Z die durch

$$(f \circ g)(x) = f(g(x))$$

definiert wird. Wir nennen $f \circ g$ die *Verknüpfung/Komposition oder Hintereinanderausführung* von f und g.

Man liest $f \circ g$ auch „f nach g" um zu verdeutlichen, dass erst g und dann f angewendet wird.

Satz 3.5 (und Definition)**.** *Ist $f : X \to Y$ bijektiv, so ist f* invertierbar*, d.h. es gibt genau eine Abbildung $f^{-1} : Y \to X$ (genannt* Umkehrabbildung*) mit der Eigenschaft, dass*

$$f \circ f^{-1} = \mathrm{id}_Y, \quad f^{-1} \circ f = \mathrm{id}_X .$$

Anders geschrieben: Für alle $x \in X$ und $y \in Y$ gilt $f(f^{-1}(y)) = y$ und $f^{-1}(f(x)) = x$.

Beweis.
Ein Blick auf die Definition des Begriffes „bijektiv" verrät schon den Beweis der Existenz einer Umkehrabbildung: f ist bijektiv, wenn die Gleichung $f(x) = y$ für alle y genau eine Lösung hat. Also können wir die gesuchte Abbildung f^{-1} so definieren:

$$f^{-1}(y) = \text{„die eindeutige Lösung } x \text{ von } f(x) = y\text{"}.$$

Damit gilt für x und $y = f(x)$:

$$f^{-1}(f(x)) = f^{-1}(y) = \text{„die eindeutige Lösung von } f(x) = y\text{"} = x$$
$$f(f^{-1}(y)) = f(x) = y.$$

Auch die Eindeutigkeit der Umkehrabbildung haben wir eigentlich schon gezeigt, hier noch einmal ausführlich: Wir nehmen an, g und h seien beide Umkehrabbildungen von f und zeigen, dass $g = h$ gilt (was bedeutet, dass für alle y gilt $g(y) = h(y)$).

Sei also $y \in Y$. Dann gibt es genau ein $x \in X$ mit $f(x) = y$ und es gilt $g(y) = x$ und $h(y) = x$, also $g = h$. □

4 Körper

Definition 4.1. Ein *Körper* ist eine Menge K mit zwei Verknüpfungen $+$ und $\cdot$ die $x, y \in K$ das Element $x + y \in K$ bzw. $x \cdot y \in K$ zuordenen, und welche folgende Eigenschaften haben:

Eine Verknüfung auf einer Menge X ist eine Abbildung $X \times X \to X$.

$$
\begin{aligned}
&(x+y)+z = x+(y+z) && \text{(Assoziativität der Addition)}\\
&x+y = y+x && \text{(Kommutativität der Addition)}\\
&\exists 0 \in K \,\forall x : x+0 = x && \text{(Neutrales Element für } +)\\
&\forall x \in K \,\exists -x \in K : x+(-x) = 0 && \text{(Inverses Element für } +)
\end{aligned}
$$

Wir schreiben auch $x - y$ für $x + (-y)$

$$
\begin{aligned}
&(x \cdot y)\cdot z = x \cdot (y \cdot z) && \text{(Assoziativität der Multiplikation)}\\
&x \cdot y = y \cdot x && \text{(Kommutativität der Multiplikation)}\\
&\exists 1 \in K \,\forall x : 1 \cdot x = x && \text{(Neutrales Element für } \cdot)\\
&\forall x \in K \; x \neq 0 \,\exists x^{-1} \in K : x \cdot x^{-1} = 1 && \text{(Inverses Element für } \cdot)
\end{aligned}
$$

Wir schreiben auch $\frac{1}{x}$ für x^{-1} und $\frac{y}{x}$ für yx^{-1}

$$
x \cdot (y+z) = x \cdot y + x \cdot z \qquad 1 \neq 0 \qquad \text{(Distributivität)}
$$

Die Rechenregeln in einem Körper bilden genau die Rechenregeln für die Grundrechenarten für Zahlen ab. In gewissen Sinne sind sie minimal; alle weiteren Regeln folgen aus ihnen, z.B.:

Satz 4.2. *Sei $(K, +, \cdot)$ ein Körper. Dann gilt:*

i) Die neutralen Elemente 0 und 1 sind eindeutig bestimmt, ebenso wie die inversen Elemente bzgl. $+$ und $\cdot$.

ii) Für jedes $x \in K$ gilt $x \cdot 0 = 0$.

iii) Ist $x \cdot y = 0$ so ist $x = 0$ oder $y = 0$.

Beweis.

i) Wir zeigen Eindeutigkeit der 0: Angenommen, 0 und $\tilde{0}$ sind beide neutral bzgl. der Addition. Dann gilt

$$
\begin{aligned}
\tilde{0} &= \tilde{0} + 0 && (0 \text{ ist neutral bzgl. } +)\\
&= 0 + \tilde{0} && \text{(Kommutativität der Addition)}\\
&= 0 && (\tilde{0} \text{ ist neutral bzgl. } +)
\end{aligned}
$$

Die Argumente für die Eindeutigkeit der 1 und der inversen Elemente gehen nach gleichem Muster.

ii) Wir formen um

$$
\begin{aligned}
x \cdot 0 + 0 &= x \cdot 0 && (0 \text{ ist neutral bzgl. } +)\\
&= x \cdot (0+0) && (0 \text{ ist neutral bzgl. } +)\\
&= x \cdot 0 + x \cdot 0. && \text{(Distributivität)}
\end{aligned}
$$

Die Behauptung folgt nun, indem wir zu obiger Gleichung auf beiden Seiten $-x \cdot 0$ addieren.

iii) Gelte $x \cdot y = 0$. Ist nun $x = 0$ ist die Behauptung gezeigt. Ist $x \neq 0$, so existiert x^{-1} und es gilt

$$\begin{aligned} y &= (x^{-1} \cdot x) \cdot y && (x^{-1}x = 1 \text{ ist neutral bzgl. } \cdot) \\ &= x^{-1} \cdot (x \cdot y) && (\text{Assoziativität von } \cdot) \\ &= x^{-1} \cdot 0 && (\text{Annahme}) \\ &= 0 && (\text{nach ii)}) \end{aligned}$$

□

Da die geforderten Eigenschaften von $+$ und $\cdot$ genau die Rechenregeln für Zahlen beschreiben, sind die rationalen Zahlen $\mathbb{Q}$ und auch die reellen Zahlen $\mathbb{R}$ Körper. Die Strukturen $\mathbb{Q}$ und $\mathbb{R}$ haben noch weitere Eigenschaften wie z.B. eine Anordnung (d.h. es gibt eine Relation $\geq$), die aber hier keine Rolle spielen.

Wie gewohnt schreiben wir für eine Multiplikation statt $x \cdot y$ oft einfach xy. Außerdem benutzen wir die Schreibweise $xx = x^2$, usw.

Ein weiterer wichtiger Körper ist der Körper der komplexen Zahlen.

Definition 4.3. Der Körper der *komplexen Zahlen* $\mathbb{C}$ besteht aus Paaren $z = (a, b)$ reeller Zahlen. Die Verknüpfungen sind wie folgt definiert: Für $z = (a, b)$ und $w = (u, v)$ ist

$$\begin{aligned} z + w &= (a, b) + (u, v) &&:= (a + u, b + v) \\ zw &= (a, b)(u, v) &&:= (au - bv, av + bu) \end{aligned}$$

In Wirklichkeit ist die obige Definition gleichzeitig ein Satz, da nicht ohne weiteres klar ist, dass die Verknüpfungen wirklich die geforderten Eigenschaften haben.

Beweis.
Existenz neutraler Elemente: Für die Addition ist $(0, 0)$ das neutrale Element, denn $(a, b) + (0, 0) = (a, b)$. Für die Multiplikation gilt

$$(a, b)(1, 0) = (a - 0, 0 + b) = (a, b).$$

Kommutativität und Assoziativität der Addition folgt direkt aus den entsprechenden Regeln für reelle Zahlen, ebenso wie die Kommutativität der Multiplikation. Assoziativität der Multiplikation ergibt sich aus den Rechnungen

$$\begin{aligned} &\Big((a, b)(u, v)\Big)(x, y) \\ &\qquad = (au - bv, av + bu)(x, y) \\ &= \Big(aux - bvx - (avy + buy), auy - bvy + avx + bux\Big) \end{aligned}$$

und

$$\begin{aligned} &(a, b)\Big((u, v)(x, y)\Big) \\ &\qquad = (a, b)(ux - vy, uy + vx) \\ &= \Big(aux - avy - (buy + bvx), auy + avx + bux - bvy\Big) \end{aligned}$$

und dem Vergleich der rechten Seiten.

Dass das Inverse bzgl. $+$ zu (a, b) einfach $(-a, -b)$ ist, ist klar. Um das Inverse bzgl. der Multiplikation zu finden, nehmen wir an dass $(a, b) \neq 0$ (d.h. a oder b ist nicht Null) und betrachten die Gleichung $(a, b)(u, v) = (1, 0)$, welche den beiden Gleichungen

$$au - bv = 1, \quad \text{und} \quad av + bu = 0$$

entspricht. Um diese Gleichungen nach u und v aufzulösen, multiplizieren wir die erste Gleichung mit a und die zweite mit b und erhalten

$$a^2u - abv = a \quad \text{und} \quad abv + b^2u = 0.$$

Wir lösen die zweite Gleichung nach abv auf, setzen in die erste Gleichung ein und bekommen

$$a^2u + b^2u = a.$$

Da nun $a^2 + b^2 \neq 0$ gilt (wir wissen sogar > 0, was aber keine Rolle spielt), können wir auflösen

$$u = \frac{a}{a^2 + b^2}.$$

Ist $a \neq 0$, können wir aus $av + bu = 0$ folgern, dass

$$av = -\frac{ab}{a^2 + b^2}$$

und damit $v = -\frac{b}{a^2+b^2}$. Ist $a = 0$, so muss $b \neq 0$ gelten und aus $au - bv = 1$ folgt ebenfalls diese Formel. Wir halten fest

$$(a, b)^{-1} = \left(\frac{a}{a^2 + b^2}, -\frac{b}{a^2 + b^2}\right).$$

$\square$

Man kann die komplexen Zahlen auch anders interpretieren: Schreiben wir für $(a, 0)$ einfach a, dann können wir die reellen Zahlen als Teil von $\mathbb{C}$ auffassen. Insbesondere ist

$$r(a, b) = (ra, rb)$$

Kürzen wir ab

$$\mathrm{i} := (0, 1)$$

so bemerken wir

$$\mathrm{i}^2 = (0, 1)(0, 1) = (-1, 0) = -1.$$

Wir können damit eine komplexe Zahl $z = (a, b)$ auch schreiben als

$$z = (a, b) = a(1, 0) + b(0, 1) = a + b\mathrm{i}.$$

Man rechnet einfach nach, dass mit dieser Schreibweise folgt, dass

$$\begin{aligned}(a+b\mathrm{i})(u+v\mathrm{i}) &= au + av\mathrm{i} + bu\mathrm{i} + bv\mathrm{i}^2 = au + (av+bu)\mathrm{i} - bv\\ &= au - bv + (av+bu)\mathrm{i},\end{aligned}$$

und

$$(a+b\mathrm{i}) + (u+v\mathrm{i}) = a+u+(b+v)\mathrm{i}.$$

Man kann also mit komplexen Zahlen $z = a + b\mathrm{i}$ genau so rechnen, wie mit reellen Zahlen, muss nur $\mathrm{i}^2 = -1$ berücksichtigen.

Es gibt auch Körper mit endlich vielen Elementen:

Beispiel.

Der Körper $\mathbb{F}_2$ ist die Menge $\{0,1\}$ versehen mit folgenden Verkünpfungen:

$$\begin{aligned}0+0&=0\\ 0+1&=1+0=1\\ 1+1&=0\\ 0\cdot 0&=0\\ 0\cdot 1&=1\cdot 0=0\\ 1\cdot 1&=1.\end{aligned}$$

Man kann die Verknüpfungen auch in einer Tabelle notieren:

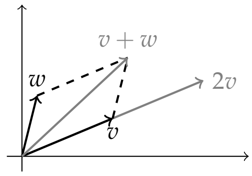
+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

Alle Eigenschaften eines Körpers lassen sich einfach verifizieren, z.B. ist das additiv Inverse zur 1 die 1 selbst, denn $1+1=0$, es gilt also $-1 = 1$ in $\mathbb{F}_2$. △

5 Vektorräume

Während Körper in gewissen Weise die Abstraktion von „Zahlen“ sind, sind Vektorräume die Abstraktion des Begriffes der Ebene und des drei- (oder mehr-)dimensionalen Raumes. Hier können wir uns Punkte als Ortsvektoren vorstellen, diese addieren und skalieren.



Wir legen zum Skalieren statt der reellen Zahlen einen beliebigen Körper zugrunde und definieren:

Definition 5.1. Es sei K ein Körper. Ein K-*Vektorraum* ist eine Menge V mit einer *Addition* $+ : V \times V \to V$ und einer *skalaren Multiplikation* (oder *Skalierung*) $\cdot : K \times V \to V$ welche folgende Regeln erfüllen:

Konvention: u, v, w sind Elemente von V, k, h sind Elemente in K.

$$(u+v)+w = u+(v+w) \qquad \text{(Assoziativität der Addition)}$$
$$v+w = w+v \qquad \text{(Kommutativität der Addition)}$$
$$\exists 0 \in V \forall v \in V:\ v+0 = v \qquad \text{(Nullvektor)}$$
$$\forall v \in V \exists -v : v+(-v) = 0 \qquad \text{(Inverses Element)}$$

$$(k+h)\cdot v = k\cdot v + h\cdot v \qquad \text{(Distributivität)}$$
$$k\cdot(v+w) = k\cdot v + k\cdot w \qquad \text{(Distributivität)}$$
$$(k\cdot h)\cdot v = k\cdot(h\cdot v) \qquad \text{(Distributivität)}$$
$$1v = v \qquad \text{(Neutralität von 1)}$$

Beachte, dass wir in diesen Formeln die Symbole $+$ und $\cdot$ *überladen* habe, d.h. sie haben verschiedene Bedeutungen, je nachdem, was sie verknüpfen.

Wir haben keine Multiplikation von Vektoren definiert!

Wie in Satz 4.2 kann man zeigen, dass der Nullvektor und die additiv Inversen eindeutig sind.

Sind v, w additiv invers zu u (also $u+v=0$ und $u+w=0$), dann gilt $v = v+0 = v+(u+w) = (v+u)+w = (u+v)+w = 0+w = w+0 = w$.

Zuerst die trivialen Beispiele:

Beispiel.

- Der einfachste Vektorraum ist $V = \{0\}$ mit den Rechenregeln $0+0=0$, $a0=0$ (über einem beliebigen Körper).
- Jeder Körper K ist auch ein Vektorraum (mit sich selbst als Grundkörper).

△

Jetzt das einfachste nicht-triviale Beispiel.

Definition 5.2. Ist K eine Körper und $n \in \mathbb{N}$, $n \geq 1$ so ist K^n die Menge der n-Tupel. Wir schreiben

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in V$$

und definieren die Addition und die skalare Multiplikation *komponentenweise*, d.h.

$$
\begin{aligned}
v + w &= \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} := \begin{pmatrix} v_1 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix} \\
kv &= k \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} := \begin{pmatrix} kv_1 \\ \vdots \\ kv_n \end{pmatrix}.
\end{aligned}
$$

Satz 5.3. *Für einen Körper K ist K^n aus Definition 5.2 ein Vektorraum.*

Beweis.
Der Nullvektor und der zu v inverse Vektor sind

$$
\begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \quad \text{und} \quad -\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} -v_1 \\ \vdots \\ -v_n \end{pmatrix}.
$$

Alle Rechenregeln folgen direkt aus den entsprechenden Regeln für das Rechnen im Körper K, z.B. die Kommutativität der Addition von Vektoren:

$$
v + w = \begin{pmatrix} v_1 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix} = \begin{pmatrix} w_1 + v_2 \\ \vdots \\ w_n + v_n \end{pmatrix} = w + v.
$$

□

Der Vektorraum $V = \mathbb{R}^2$ ist die Ebene und $\mathbb{R}^3$ ist der dreidimensionale (reelle) Raum und die Vektoraddition und -skalierung entspricht in beiden Fällen genau der Anschauung.

Definition 5.4. Eine $m \times n$ *Matrix* mit Elementen in einem Körper K ist

$$
A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = (a_{ij})_{i=1,\dots,m, j=1,\dots,m}, \quad (a_{ij} \in K).
$$

$$
\begin{array}{cc}
\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & a_{ij} & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} & i\text{-te Zeile} \\
j\text{-te Spalte} &
\end{array}
$$

Addition und Skalierung von Matrizen ist komponentenweise definiert, also für $A = (a_{ij})_{i=1,\dots,m, j=1,\dots,m}$, $B = (b_{ij})_{i=1,\dots,m, j=1,\dots,m}$

$$
kA = (ka_{ij})_{i=1,\dots,m, j=1,\dots,m}, \quad A + B = (a_{ij} + b_{ij})_{i=1,\dots,m, j=1,\dots,m}.
$$

Satz 5.5. *Die Menge $K^{m \times n}$ aller $m \times n$ Matrizen mit Elementen in K ist ein Vektorraum.*

Der Beweis ist völlig analog zu dem Beweis, dass K^n ein Vektorraum ist. Als Vektorraum verhält sich $K^{m \times n}$ genau wie K^{mn}, nur dass die Einträge in einer Matrix und nicht in einem langen Vektor angeordnet sind.

Im Folgenden ist K immer ein Körper und V bezeichnet stets einen K-Vektorraum, es sei denn, es wird etwas anderes gesagt.

Definition 5.6. Eine Teilmenge $U \subset V$ einer Vektorraumes heißt *Untervektorraum* oder nur *Unterraum*, falls gilt

$$\begin{aligned} &U \neq \emptyset && \text{(nichtleer)}\\ &u, v \in U \implies u + v \in U && \text{(abgeschlossen bzgl. Addition)}\\ &u \in U, k \in K \implies ku \in U. && \text{(abgeschlossen bzgl. Skalierung)} \end{aligned}$$

Beispiel.
Jeder Vektorraum V hat die Unterräume $\{0\}$ und V. △

Auf U sollen die gleiche Addition und Skalierung wir in V gelten. Aus der Definition folgt, dass diese auch wirklich Verknüpfungen auf U sind (und nicht evtl. „aus U heraus abbilden").

Satz 5.7. *Ein Untervektorraum U eines Vektorraumes V ist selbst ein Vektorraum.*

Beweis.
Da die Addition und die Skalierung die gleichen sind wie in V, gelten auch die gleichen Rechenregeln. Es ist also nur noch zu zeigen, dass Nullvektor und inverse Elemente existieren. Da U nicht leer ist, gibt es ein $u \in U$ und daher ist $0 = 0 \cdot u \in U$. Außerdem ist $(-1)u \in U$ und wegen $u + (-1)u = 0$ ist dies das additiv inverse Element. □

Beispiel.

- In der Analysis betrachtet man Funktionen $f : \mathbb{R} \to \mathbb{R}$. Die Menge aller solcher Funktionen ist ein $\mathbb{R}$-Vektorraum, wenn wir Addition und Skalierung wie folgt definieren: Für $f, g : \mathbb{R} \to \mathbb{R}$ und $a \in \mathbb{R}$

$$(f + g)(x) := f(x) + g(x), \quad (af)(x) := af(x).$$

- Die Menge

$$\{f : \mathbb{R} \to \mathbb{R} \mid \forall x : f(x) \geq 0\}$$

  der positiven Funktionen ist kein Vektorraum (es gibt keine additiv Inversen).

- Etwas allgemeiner: Ist X eine Menge und V ein K-Vektorraum, so ist die Menge

$$V^X := \{f : X \to V\}$$

  versehen mit den Operationen

$$(f + g)(x) = f(x) + g(x), \quad (af)(x) = af(x)$$

  wieder ein K-Vektorraum.

  Die Schreibweise V^X für die Menge der Abbildungen von X nach V erklärt sich dadurch, dass für endliche Mengen X und V gilt $|V^X| = |V|^{|X|}$.

- Ist K ein Körper, so ist

$$K^\infty := \{(x_n)_{n\in\mathbb{N}} \mid x_n \in K\}$$

die Menge aller Folgen mit Folgengliedern in K. Auch dies ist ein Vektorraum (mit komponentenweiser Addition und Skalierung). Die Vektorräume K^n sind in naheliegender Weise Untervektorräume. Genauer: Der K^n kann mit dem Unterraum

$$U = \{v \in K^\infty \mid v_k = 0 \text{ für } k > n\}$$

identifiziert werden.

- Auch folgende Menge ist ein Untervektorraum von K^∞:

$$K^{(\infty)} = \{v \in K^\infty \mid \exists n : v_k = 0 \quad \text{für } k > n\}$$

(die Menge aller Folgen, die schließlich Null sind). Der Raum K^n ist für jedes $n \in \mathbb{N}$ ein Unterraum von $K^{(\infty)}$ und $K^{(\infty)}$ ist wiederum ein (echter!) Unterraum von K^∞.

$\triangle$

6 Lineare Unabhängigkeit

Definition 6.1. Eine *Linearkombination* der Vektoren $v_1, \ldots, v_n \in V$ ist eine Vektor der Form

$$k_1 v_1 + \cdots + k_n v_n \quad \text{mit Skalaren} \quad k_1, \ldots, k_n \in K.$$

Achtung, hier bezeichnet v_i einen Vektor und nicht den i-ten Eintrag eines Vektors v!

Sind alle $k_i = 0$, so ergibt die Linearkombinaton der Nullvektor und wir sprechen von der *trivialen Linearkombination.*

Für Linearkombinationen (und auch sonst) ist die Summenschreibweise mit dem Zeichen $\sum$ praktisch:

Der Ausdruck $\sum_{k=m}^{n} a_k$ liest sich zum Beispiel „Die Summe über a_k für k von m bis n“.

$$\sum_{i=1}^{n} a_i := a_1 + \cdots + a_n.$$

Schreibweisen wie

$$\sum_{3 \le k \le 10} a_k \text{ oder } \sum_{\substack{0 \le n \le N \\ n \text{ Primzahl}}} a_n$$

erklären sich von selbst.

Im $\mathbb{R}^3$ beschreiben Vektoren Richtungen. Zwei Vektoren können in die gleiche oder in verschiedene Richtungen zeigen. Bei drei Vektoren kann man sich ebenfalls fragen, ob „alle drei in verschiedene Richtungen zeigen oder nicht.“ In einem allgemeinen Vektorraum gibt es hierfür den folgenden Begriff:

Definition 6.2. Die Vektoren $v_1, \ldots v_n$ heißen *linear unabhängig*, wenn die triviale Linearkombination die einzige ist, die den Nullvektor liefert; in Formeln: $v_1, \ldots, v_n$ sind linear unabhängig, falls

$$k_1 v_1 + \cdots + k_n v_n = 0 \implies k_1 = \cdots = k_n = 0.$$

Sind Vektoren nicht linear unabhängig, so nennen wir sie *linear abhängig.*

Anders gesagt: $v_1, \ldots, v_n$ sind linear abhängig, wenn es Skalare $k_1, \ldots, k_n$ gibt, welche nicht alle Null sind und trotzdem gilt $k_1 v_1 + \cdots + k_n v_n = 0$.

Beispiel.

Die Vektoren

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{R}^3$$

sind linear unabhängig, denn die Gleichung

$$k_1 v_1 + k_2 v_2 + k_3 v_3 = 0$$

entspricht den drei Gleichungen

$$\begin{aligned} k_1 + k_2 + k_3 &= 0 \\ k_1 + k_2 + 0 \;\; &= 0 \\ k_1 + 0 \;\; + 0 \;\; &= 0 \end{aligned}$$

und wir können von oben nach unten ablesen, dass $k_1 = 0, k_2 = 0$ und $k_3 = 0$ gelten muss.

Die Vektoren

$$v_1 = \begin{pmatrix}1\\0\\1\end{pmatrix}, \quad v_2 = \begin{pmatrix}0\\1\\1\end{pmatrix}, \quad v_3 = \begin{pmatrix}2\\-1\\1\end{pmatrix} \in \mathbb{R}^3$$

hingegen sind linear abhängig, denn es gilt

$$2v_1 - v_2 - v_3 = \begin{pmatrix}2-0-2\\0-1-(-1)\\2-1-1\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}.$$

△

Satz 6.3. *Die Vektoren $v_1, \dots, v_n$ sind genau dann linear abhängig, wenn es einen Vektor v_i gibt, welcher sich als Linearkombination der anderen schreiben lässt.*

Beweis.
„$\implies$“: Sind $v_1, \dots, v_n$ linear abhängig, so gibt es eine nicht-triviale Linearkombination der Null, d.h. es gilt

$$\sum_{i=1}^{n} k_i v_i = 0$$

und mindestens ein Skalar ist nicht Null. Das heißt, es gibt $j \in \{1, \dots, n\}$, so dass $k_j \neq 0$ und wir können nach v_j auflösen, nämlich

$$v_j = -\tfrac{k_1}{k_j} v_1 \cdots - \tfrac{k_{j-1}}{k_j} v_{j-1} - -\tfrac{k_{j+1}}{k_j} v_{j+1} - \cdots - \tfrac{k_n}{k_j} v_n.$$

„$\Longleftarrow$“: Lasse sich einer der Vektoren v_i als Linearkombination der anderen schreiben. Der Einfachheitheit halber nehmen wir $i = 1$ an (dies könnten wir z.B. durch umnummerieren der Vektoren erzeugen).

Es gilt also

$$v_1 = \sum_{i=2}^{n} k_i v_i$$

und wir bekommen eine nicht-triviale Linearkombination der Null durch Umstellen

$$v_1 - \sum_{i=2}^{n} k_i v_i = 0.$$

□

Satz 6.4 (Koeffizientenvergleich). *Sind $v_1, \dots, v_n$ linear unabhängig und gilt*

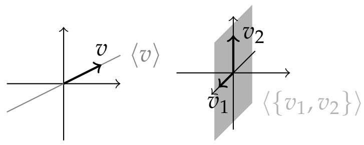
$$\sum_{i=1}^{n} k_i v_i = \sum_{i=1}^{n} h_i v_i,$$

so folgt $k_i = h_i$ für alle $i = 1, \dots, n$.

Beweis.
Es folgt nämlich
$$\sum_{i=1}^{n} (k_i - h_i) v_i = 0$$
und da dies eine Linearkombination ist, die Null ergibt und die Vektoren linear unabhängig sind, folgt, dass für alle i gilt $k_i - h_i = 0$. □

Definition 6.5. Für eine Menge $M \subset V$ (endlich, oder unendlich) ist die Menge aller Linearkombinationen von Vektoren aus M definiert durch
$$\langle M \rangle = \{ \sum_{i=1}^{n} k_i v_i \mid n \in \mathbb{N},\ k_i \in K, v_i \in M \}$$
und wird *Erzeugnis von M* genannt. Andere Namen sind die *lineare Hülle von M* oder der *Spann von M* und eine andere Schreibweise ist $\operatorname{Span}(M)$.

Achtung: Linearkombinationen sind immer endliche Summen. Im Erzeugnis sind beliebig lange, aber nur endliche Summen zugelassen!



Beachte: Oft werden die Mengenklammern weggelassen, d.h. man schreibt $\langle v_1, \dots, v_n \rangle$ statt $\langle \{ v_1, \dots, v_n \} \rangle$.

Beispiel.

- Für einen Vektor v ist $\langle \{v\} \rangle$ die Menge aller skalaren Vielfache von v.
- Im $\mathbb{R}^3$ ist
$$\langle \{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \} \rangle = \{ \begin{pmatrix} s \\ t \\ 0 \end{pmatrix} \mid s, t \in \mathbb{R}$$
oder
$$\begin{aligned} \langle \{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \} \rangle &= \{ s \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \mid s, t \in \mathbb{R} \} \\ &= \{ \begin{pmatrix} s+t \\ 0 \\ t \end{pmatrix} \mid s, t \in \mathbb{R} \}. \\ &= \{ \begin{pmatrix} s \\ 0 \\ t \end{pmatrix} \mid s, t \in \mathbb{R} \}. \end{aligned}$$

△

Satz 6.6. *Ist M eine nicht leere Teilmenge eines Vektorraums V, so ist die Menge $\langle M \rangle$ ist ein Vektorraum.*

Die Bedingung, dass M nicht leer ist, kann man weglassen, wenn man die Konvention benutzt, dass leere Summen Null ergeben, in diesem Fall also $\sum_{i \in \emptyset} = 0$ (Nullelement des Vektorraumes).

Beweis.
Die Menge $U = \langle M \rangle$ ist offensichtlich nichtleer. Sind u und v in M, dann lassen sich die Vektoren schreiben als

$$u = \sum_{i=1}^{n} k_i v_i, \quad v = \sum_{j=1}^{m} h_j w_j$$

mit Vektoren v_i und w_j aus M. Dann ist

$$u + v = \sum_{i=1}^{n} k_i v_i + \sum_{j=1}^{m} h_j w_j, \quad ku = \sum_{i=1}^{n} k k_i v_i$$

für jeden Skalar k. Das bedeutet, dass $u + v$ und ku ebenfalls Linearkombinationen von Elementen aus M sind, also gilt $u + v \in \langle M \rangle$ und $ku \in \langle M \rangle$. Also ist U ein Unterraum und nach Satz 5.7 ein Vektorraum. □

Beispiel.

1. Sind $U, W \subset V$ zwei Unterräume, so ist $\langle U \cup W \rangle$ nach dem vorigen Satz ein Unterraum und man erkennt leicht, dass $\langle U \cup W \rangle = U + W$ (da $U + W$ alle Summen von Elementen aus U und W enthält; vgl. große Übung).

2. Im K^∞ bezeichnen wir mit e_i die Folge, die genau im i-ten Eintrag eine 1 und sonst Nullen enthält. Sei nun

$$M = \{e_i \mid i \in \mathbb{N}\}.$$

   Was ist $U = \langle M \rangle$?

   Sicherlich ist $U \neq K^\infty$, denn $e = (1, 1, \dots)$ liegt in K^∞ aber nicht in U. Letzteres sieht man, wenn man versucht, e als Summe von *endlich vielen* e_i zu schreiben. Das geht nicht, denn jede endliche Summe der e_i hat nur an endlich vielen Einträgen einen Eintrag, der von Null verschieden ist. Anders gesagt: Es würden unendlich viele benötigt, aber „Summen"der Form $\sum_{i=1}^{\infty}$ haben wir noch nicht erklärt und sie sind in allgemeinen Vektorräumen auch nicht sinnvoll. Sie werden in der Analysis eingeführt und heißen Reihen. Zu ihrer Untersuchung benötigt man den Grenzwertbegriff.

   Es gilt allerdings

$$U = K^{(\infty)} = \{a \in K^\infty \mid a_i \neq 0 \text{ höchstens endlich oft}\}.$$

△

7 Basen

Definition 7.1. Ist V ein Vektorraum, $M \subset V$ und $V = \langle M \rangle$, so heißt M *Erzeugendensystem* von V.

Gibt es für einen Vektorraum V ein Erzeugendensystem mit endlich vielen Elementen, so nennen wir V *endlich erzeugt.*

Ein Erzeugendensystem von V heißt *Basis* von V, wenn es linear unabhängig ist.

Beispiel.

- Im reellen Vektorraum $\mathbb{R}^n$ betrachten wir

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

  Diese Vektoren sind linear unabhängig (denn aus $\sum_{i=1}^n k_i e_i = 0$ folgt, dass alle k_i Null sind) und die bilden ein Erzeugendensystem, denn für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ gilt

$$x = \sum_{i=1}^n x_i e_i.$$

  Also ist $\{e_1, \dots, e_n\}$ eine Basis von $\mathbb{R}^n$; diese heißt auch *kanonische Basis* oder *Standardbasis* und also ist $\mathbb{R}^n$ endlich erzeugt.

- Die Vektoren

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

  bilden keine Basis des $\mathbb{R}^2$, da sie nicht linear unabhängig sind. (Spannen sie $\mathbb{R}^2$ auf? Ja, denn $(x_1,\ x_2) = \frac{1}{2}(x_2 - x_1)(-1,\ 1) + \frac{1}{2}(x_2 + x_1)(1,\ 1)$.)

- Im Vektorraum K^∞ der Folgen in K bildet die (unendliche) Menge $M = \{e_1, e_2, \dots\}$ *keine* Basis. Die Vektoren sind zwar linear unabhängig, bilden aber kein Erzeugendensystem (der Vektor $(1,\ 1, \dots)$ liegt nicht in $\langle M \rangle$; insbesondere ist K^∞ nicht endlich erzeugt.)

$\triangle$

Wir schreiben statt $\{v_1, v_2, \dots\}$ in Zukunft auch $\{v_i\}_{i \in I}$ wobei I eine beliebige Menge darstellt, die wir als Indexmenge benutzen. Dabei kann I eine endliche Menge (wie $I = \{1, 2, 3\}$) sein, aber auch unendlich wie $I = \mathbb{N}$ oder $I = \mathbb{R}$.

Satz 7.2. *Ist V ein Vektorraum und $\{v_i\}_{i \in I}$ eine Basis, so lässt sich jedes $v \in V$ eindeutig als Linearkombination von Vektoren v_i darstellen.*

Beweis.
Da eine Basis ein Erzeugendensystem ist, lässt sich jedes $v \in V$ als eine Linearkombination der v_i schreiben. Bleibt die Eindeutigkeit der Darstellung zu zeigen. Nehmen wir an, es gäbe zwei Darstellungen

$$v = \sum_{i=1}^{n} k_i v_i, \quad v = \sum_{i=1}^{n} h_i v_i.$$

Beachte, dass wir ohne Einschränkung das gleiche n und die gleichen $v_1, \ldots, v_n$ genommen haben (wir können nämlich beliebige endlich viele Vektoren aus $\{v_i\}$ hinzunehmen wenn wir die zugehörigen Skalare auf 0 setzen).

Dann gilt aber

$$\begin{aligned} 0 &= v - v \\ &= \sum_{i=1}^{n} k_i v_i - \sum_{i=1}^{n} h_i v_i \\ &= \sum_{i=1}^{n} (k_i - h_i) v_i \end{aligned}$$

und da die v_i $(i = 1, \ldots, n)$ linear unabhängig sind, folgt $k_i - h_i = 0$ für alle i, also $k_i = h_i$. $\square$

Satz 7.3 (Basisergänzungssatz). *Eine Teilmenge B eines Vektorraumes V ist eine Basis genau dann, wenn sie eine maximale linear unabhängige Menge ist. Dabei heißt „maximal linear unabhängig", dass $B \cup \{v\}$ für jedes $v \in V \setminus B$ nicht mehr linear unabhängig ist.*

Eine äquivalente Formulierung dazu ist: Ist B eine linear unabhängige Teilmenge eines Vektorraumen V, so lässt sich B durch Hinzunahme von Elementen aus V zu einer Basis ergänzen. Daher der Name „Basisergänzungssatz".

Beweis.
Sei B eine Basis von V. Nach Definition des Begriffs Basis ist B linear unabhängig. Bleibt die Maximalität zu zeigen. Sei $v \in V \setminus B$. Da B eine Basis ist, gibt es Skalare k_i und $v_i \in B$ $(i = 1, \ldots, n)$, so dass

$$v = \sum_{i=1}^{n} k_i v_i.$$

Dann ist aber $0 = v - \sum_{i=1}^{n} k_i v_i$, was zeigt, dass es eine nichttriviale Linearkombination der Null mit Elementen aus $B \cup \{v\}$ gibt. Also ist $B \cup \{v\}$ nicht linear unabhängig.

Sei nun B eine maximale linear unabhängige Menge. Wir müssen nur zeigen, dass B ein Erzeugendensystem für V ist. Wäre dies nicht so, so gäbe es $v \in V$, welches keine Linearkombination von Elementen in V wäre. Dann wäre aber $B \cup \{v\}$ linear unabhängig, was eine Widerspruch zur Maximalität ist. $\square$

Für endlich erzeugte Vektorräume können wir den vorigen Satz auch so schreiben:

> Ist $\{v_1, \ldots, v_n\}$ eine linear unabhängige Teilmenge von V, so gibt es $v_{n+1}, \ldots, v_m$, so dass $\{v_1, \ldots, v_m\}$ eine Basis von V ist.

Satz 7.4 (Basisauswahlsatz)**.** *Eine Teilmenge B eines Vektorraumes V ist eine Basis genau dann, wenn B ein minimales Erzeugendensystem ist, d.h. wenn für jedes $v \in B$ gilt, dass $B \setminus \{v\}$ kein Erzeugendensystem ist.*

Eine äquivalente Formulierung ist, dass jedes Erzeugendensystem B eine Teilmenge $B' \subset B$ hat, welches eine Basis von V ist. Daher der Name „Basisauswahlsatz".

Beweis.
Sei B eine Basis und sei $v \in B$. Nach Definition ist B ein Erzeugendensystem. Da B linear unabhängig ist, lässt sich v nicht als Linearkombination von Elementen in $B \setminus \{v\}$ darstellen (denn sonst hätten wir eine nicht-triviale Linearkombination der Null mit Elementen aus B). Es folgt $v \notin \langle B \setminus \{v\} \rangle$, und daher ist $B \setminus \{v\}$ kein Erzeugendensystem.

Sei nun B eine minimales Erzeugendensystem von V. Es bleibt zu zeigen, dass B linear unabhängig ist. Wäre B nicht linear unabhängig, so könnten wir einen der Vektoren in B als Linearkombination der anderen schreiben. Dann aber wäre $\langle B \rangle = \langle B \setminus \{v\} \rangle$ und auch $B \setminus \{v\}$ wäre ein Erzeugendensystem. □

Insbesondere können wir folgern:

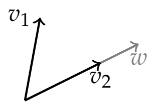
Satz 7.5. *Jeder endlich erzeugte Vektorraum V hat hat eine Basis mit endlich vielen Elementen.*

Beweis.
Nach Vorraussetzung gibt es ein endliches Erzeugendensystem E von V. Ist E ein minimales Erzeugendensystem, so ist E eine (endliche) Basis. Ist E nicht minimal, so gibt es $v_0 \in E$, so dass $E_1 = E \setminus \{v_0\}$ immer noch ein Erzeugendensystem ist. So verkleinern wir E weiter, bis ein minimales Erzeugendensystem ensteht. Der Prozess muss enden, da E endlich war. □

8 Die Dimension eines Vektorraums

Ein Lemma ist ein Hilfssatz, also ein Satz, den man hauptsächlich zum Beweis anderer Sätze benutzt.

Lemma 8.1 (Austauschlemma). *Es sei B eine Basis eines Vektorraums V. Dann gibt es zu jedem $w \in V$, $w \neq 0$ ein $v \in B$, so dass $B' = (B \setminus \{v\}) \cup \{w\}$ wieder eine Basis von V ist.*



Basis $B = \{v_1, v_2\}$

$B' = \{v_1, w\}$ Basis, aber $\{w, v_2\}$ keine Basis

Beweis.
Da B eine Basis ist, gibt es Vektoren $v_1, \dots, v_n \in B$ und Koeffizienten $k_1, \dots, k_n$, so dass

$$w = \sum_{i=1}^{n} k_i v_i. \qquad (*)$$

Da w nicht Null ist, muss ein Koeffizient ungleich Null sein. Ohne Einschränkung sei dies $k_1 \neq 0$.

Wir zeigen, dass $B' = (B \setminus \{v_1\}) \cup \{w\}$ wieder eine Basis von V ist.

Zuerst bemerken wir, dass $v_1 = \frac{1}{k_1}\left(w - \sum_{i=2}^{n} k_i v_i\right)$, und daher ist $v_1 \in \langle B' \rangle$. Es folgt, $V = \langle B \rangle \subset \langle B' \rangle$, also ist B' ein Erzeugendensystem von V.

Wäre B' linear abhängig, dann gäbe es Koeffizienten $h_1, \dots, h_n$, so dass

$$0 = h_1 w + \sum_{i=2}^{n} h_i v_i$$

und mindestens ein h_i wäre nicht Null. Außerdem muss $h_1 \neq 0$ gelten (sonst wäre B nicht linear unabhängig gewesen). Ersetzen wir w durch seine Darstellung (*), so bekommen wir nach Division durch h_1

$$0 = \sum_{i=1}^{n} k_i v_i + \sum_{i=2}^{n} \frac{h_i}{h_1} v_i = k_1 v_1 + \sum_{i=2}^{n} \left(k_i - \frac{h_i}{h_1}\right) v_i.$$

Da B eine Basis ist, müsste $k_1 = 0$ gelten, was aber ein Widerspruch ist. $\square$

Mit diesem Lemma zeigen wir:

Satz 8.2 (Austauschsatz von Steinitz). *Ist $B = \{v_1, \dots, v_n\}$ eine Basis von V, und ist $C = \{w_1, \dots, w_m\}$ eine Menge linear unabhängiger Vektoren aus V, dann gibt es $n - m$ viele Vektoren aus B, ohne Einschränkung seien dies $v_{m+1}, \dots, v_n$, so dass $\{w_1, \dots, w_m, v_{m+1}, \dots, v_n\}$ eine Basis von V ist.*

Beweis.
Wir führen Induktion über m durch.

Ist $m = 1$, so ist die Aussage des Satzes genau die des Lemmas 8.1.

Sei nun $m > 1$, aber $m \leq n$ und wir nehmen an, dass die Aussage für $m - 1$ gezeigt ist. Sei $C = \{w_1, \dots, w_m\}$ linear unabhängig und $C' = \{w_1, \dots, w_{m-1}\}$. Da C' nur $m - 1$ Elemente hat, können wir die Aussage für $m - 1$ benutzen und sehen, dass es eine Menge $\{v_m, \dots, v_n\}$ von Vektoren aus B gibt, so dass

$\{w_1, \dots, w_{m-1}, v_m, \dots, v_n\}$ eine Basis von V ist. Das heißt, wir können schreiben

$$w_m = \sum_{i=1}^{m-1} k_i w_i + \sum_{i=m}^{n} k_i v_i.$$

Es kann nicht $k_m = \cdots = k_n = 0$ gelten, denn sonst wäre w_m eine Linearkombination der $w_1, \dots, w_{m-1}$ und C damit nicht linear unabhängig. Ohne Einschränkung sei $k_m \neq 0$. Wieder nach dem Lemma 8.1 folgt, dass

$$\{w_1, \dots, w_m, v_{m+1}, \dots, v_n\}$$

eine Basis von V ist.

□

Ein Korollar ist eine Folgerung aus einem Satz. Man benutzt „Korollar"meist dann, wenn die Aussage „ohne weitere Argumente" aus dem Satz folgt.

Korollar 8.3. *Je zwei Basen eines endlich erzeugten Vektorraumes haben gleich viele Elemente.*

Wären $B = \{v_1, \dots, v_n\}$ und $C = \{w_1, \dots, w_m\}$ zwei Basen mit $m < n$, so könnte man nach dem Austauschsatz von Steinitz C vergrößern um eine Basis zu erhalten. Das widerspricht dem Basisergänzungssatz, da C ja schon eine Basis ist.

Diese Definition ist nur sinnvoll, da Korollar 8.3 zeigt, dass die Anzahl der Elemente eines Basis (im endlich erzeugten Fall) unabhängig von der Basis ist.

Definition 8.4. Für einen endlich erzeugten Vektorraum V ist die *Dimension* die Anzahl der Elemente einer beliebigen Basis. Ist also $\{v_1, \dots, v_n\}$ eine Basis von V, so ist $\dim(V) = n$.

Korollar 8.5. *Ist* $\dim(V) = n$*, so hat jede linear unabhängige Menge höchstens n Elemente; hat sie genau n Elemente, ist sie eine Basis.*

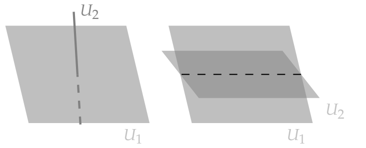
Sei B Basis mit n Elementen und C linear unabhängig. Dann ist $\langle C \rangle \subset V$. Ist $\langle C \rangle \neq V$, dann kann C zu Basis ergänzt werden, also $|C| < n$ und ist $\langle C \rangle = V = \langle B \rangle$, dann ist C Basis und es folgt nach Korollar 8.3 $|C| = |B| = n$.

Korollar 8.6. *Ist $U \subset V$ ein Unterraum, so ist* $\dim(U) \leq \dim(V)$.

Eine Basis von U ist eine linear unabhängige Teilmenge in V und kann zu Basis von V ergänzt werden.

Korollar 8.7. *Sind U_1, U_2 zwei Unterräume von V mit $U_1 \subset U_2$, so gilt* $\dim(U_1) \leq \dim(U_2)$.

U_1 ist auch ein Unterraum von U_2.



U_1 Komplement von U_2 U_1 kein Komplement von U_2

Definition 8.8. Sind U_1 und U_2 Unterräume von V, so heißt U_1 ein *Komplement* von U_2 (oder *komplementär* zu U_1), wenn

1. $U_1 \cap U_2 = \{0\}$ und
2. $\langle U_1 \cup U_2 \rangle = V$.

Beispiel.

Ist $U_1 = \langle\{e_1, e_2\}\rangle \subset V = \mathbb{R}^3$ so ist $U_2 = \langle e_3\rangle$ komplementär zu U_1 ebenso wie $U_3 = \langle e_2 + e_3\rangle$. Nicht komplementär zu U_1 sind zum Beispiel $U_4 = \langle e_2\rangle$, $U_5 = \langle\{e_2, e_3\}\rangle$ oder $U_6 = \{0\}$. $\triangle$

Satz 8.9. *Jeder Unterraum U eines endlich erzeugten Vektorraumes V hat einen komplementären Unterraum.*

Beweis.

Wir wählen eine Basis $\{u_1, \dots, u_r\}$ von U, ergänzen diese so zu $\{u_1, \dots, u_r, v_{r+1}, \dots, v_n\}$, dass Basis von V entsteht und sehen, dass $W = \langle\{v_{r+1}, \dots, v_n\}\rangle$ komplementär zu U ist. $\square$

Satz 8.10. *Sind U_1 und U_2 komplementäre Unterräume von V, so gibt es zu jedem $v \in V$ eindeutige $u_1 \in U$, $u_2 \in U_2$ mit $u_1 + u_2 = v$.*

Beweis.

Die Existenz von u_1 und u_2 folgt aus $\langle U_1 \cup U_2\rangle = V$ (also, daraus, dass $U_1 \cup U_2$ den ganzen Raum V erzeugt).

Für die Eindeutigkeit nehmen wir an, dass $v = u_1 + u_2 = w_1 + w_2$ mit $u_1, w_1 \in U_1$ und $u_2, w_2 \in U_2$. Es folgt

$$u_1 - w_1 = w_2 - u_2.$$

Die linke Seite liegt in U_1, die rechte Seite liegt in U_2, da beide gleich sind, liegen beide in $U_1 \cap U_2 = \{0\}$. Also gilt $u_1 = w_1$ und $u_2 = w_2$. $\square$

Definition 8.11. Ein Unterraum U eines n-dimensionalen Vektorraums V heißt *Hyperebene*, falls $\dim(U) = n - 1$.

Beispiel.

Für $V = \mathbb{R}^3$ ist $\langle\{e_1, e_2\}\rangle$ eine Hyperebene (die Ebene, die durch die ersten beiden Koordinatenachsen aufgespannt wird).

Für $V = \mathbb{R}^2$ sind die Geraden $\langle v\rangle$ für jedes $v \neq 0$ Hyperebenen. $\triangle$

Satz 8.12. *Sind U_1, U_2 Unterräume von V, so gilt die* Dimensionsformel

$$\dim(\langle U_1 \cup U_2\rangle) = \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2).$$

Beweis.

Wir wählen eine Basis $B = \{v_1, \dots, v_m\}$ von $U_1 \cap U_2$. Diese Ergänzen wir auf zwei Arten zu Basen

$$B_1 = \{v_1, \dots, v_m, v_{m+1}, \dots, v_r\}$$

von U_1 und

$$B_2 = \{v_1, \dots, v_m, w_{m+1}, \dots, w_s\}$$

von U_2. Dann bilden wir

$$C = \{v_1, \dots, v_m, v_{m+1}, \dots, v_r, w_{m+1}, \dots, w_s\} = B_1 \cup B_2$$

und bemerken, dass die Anzahl der Elemente in C genau $m + (r - m) + (s - m) = r + s - m = \dim(U_1) + \dim(U_2) - \dim(U_1 \cap U_2)$ ist. Es bleibt also nur noch zu zeigen, dass C eine Basis von $\langle U_1 \cup U_2 \rangle$ ist.

Klar ist, dass C den Raum $\langle U_1 \cup U_2 \rangle$ erzeugt. Zeigen wir, dass C linear unabhängig ist. Dazu gelte

$$0 = \sum_{i=1}^{m} k_i v_i + \sum_{i=m+1}^{r} k_i v_i + \sum_{j=m+1}^{s} h_j w_j. \qquad (*)$$

Es folgt

$$-\sum_{i=1}^{r} k_i v_i = \sum_{j=m+1}^{s} h_j w_j.$$

Da die linke Seite in U_1 und die rechte Seite in U_2 liegt, liegen beide Seiten in $U_1 \cap U_2$ und insbesondere ist die rechte Seite eine Linearkombination aus den Elementen in B, d.h. es gibt Koeffizienten $a_i, \dots, a_m$ mit

$$\sum_{i=1}^{m} a_i v_i = \sum_{j=m+1}^{s} h_j w_j.$$

Da aber die Menge $\{v_1, \dots, v_m, w_{m+1}, \dots, w_s\}$ linear unabhängig ist, folgt $h_{m+1} = \cdots, h_s = 0$. Gehen wir mit dieser Erkenntnis in (*), und benutzen, dass auch $\{v_1, \dots, v_m, v_{m+1}, \dots, v_r\}$ linear unabhängig ist, folgt, dass auch die Koeffizienten k_i alle Null sind. Also ist C linear unabhängig. $\square$

Korollar 8.13. *Ist H eine Hyperebene und U eine Unterraum von V, so gilt $U \subset H$ oder* $\dim(U \cap H) = \dim(U) - 1$.

Es gilt $H \subset \langle U \cup H \rangle \subset V$, also $n - 1 = \dim(H) \leq \dim\langle U \cup H \rangle \leq n$, also entweder $\dim\langle U \cup H \rangle = n$ oder $= n - 1$. Im zweiten Fall gilt $U \subset H$ und im ersten Fall folgt die Behauptung aus der Dimensionsformel $\dim\langle U \cup H \rangle = \dim U + \dim H - \dim(U \cap H)$.

Korollar 8.14. *Sind U_1 und U_2 komplementäre Unterräume von V, so folgt*

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

Folgt mit $\dim(U_1 \cap U_2) = 0$ aus der Dimensionsformel.

9 Matrizen und lineare Gleichungssysteme

Mit Matrizen werden wir lineare Gleichungssysteme beschreiben.

Definition 9.1. Ist K eine Körper und sind $a_{ij} \in K$ und $b_i \in K$, $i = 1, \dots m$, $j = 1, \dots, n$ dann ist

Standard-Nummerierung: a_{ij} ist der Koeffzient der j-ten Variable in der i-ten Gleichung.

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

ein *lineares Gleichungssystem* mit m Gleichungen und n Unbekannten.

Wichtige Fragen zu linearen Gleichungssytemen sind:

- Gibt es eine Lösung?
- Wieviele Lösungen gibt es?
- Wie kann man eine (oder alle) Lösungen ausrechnen?
- Wie aufwändig ist das berechnen einer (aller) Lösungen?

Matrizen haben wir schon in Definition 5.4 eingeführt und werden sie nun benutzen, um lineare Gleichungssysteme in kompakter Form zu schreiben.

Definition 9.2. Zu einer Matrix $A = (a_{ij}) \in K^{m \times n}$ und einem Vektor $x \in K^n$ definieren wir das *Matrix-Vektor-Produkt* Ax als Element in K^m durch

$$(Ax)_i = \sum_{j=1}^{n} a_{ij}x_j.$$

Mit $(Ax)_i$ ist der i-te Eintrag des Vektors Ax gemeint.

Damit das Matrix-Vektor-Produkt Ax erklärt ist, muss die Anzahl der Spalten von A gleich der Länge von x sein.

$$\underset{K^{m\times n}}{\overset{A}{\cap}} \quad \underset{K^{n}}{\overset{x}{\cap}}$$

Damit ist das Gleichungssystem aus Definition 9.1 nichts weiter als

$$Ax = b$$

mit $A \in K^{m \times n}$, $b \in K^m$ und gesucht ist $x \in K^n$.

Beispiel.

Für

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \end{pmatrix} \in \mathbb{R}^{3\times 4} \quad \text{und} \quad x = \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} \in \mathbb{R}^4$$

ist

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

$\triangle$

Beispiel.
Wir können eine Matrix $A \in K^{m \times n}$ zum Beispiel durch ihre *Spaltenvektoren* beschreiben: Ist

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & & & \vdots \\ \vdots & & & \vdots \\ a_{m1} & \cdots & \cdots & a_{mn} \end{pmatrix}$$

so sind die Spaltenvektoren

$$a_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix}, \dots, a_n = \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

und man schreibt auch

$$A = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} \quad \text{mit} \quad a_i \in K^n,\ i = 1, \dots, n$$

Es gilt

$$Ae_k = a_k,$$

d.h. die Multiplikation von A mit dem k-ten Standard-Basisvektor gibt die k-te Spalte. Etwas allgemeiner: Für $x \in K^n$ ist

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

d.h. Ax ist eine Linearkombination der Spalten von A und die Koeffizienten sind die Einträge in x. $\triangle$

Satz 9.3. *Die Matrix-Vektor-Multiplikation ist additiv und homogen, d.h. für $A \in K^{m \times n}$, $x, y \in K^n$ und $k \in K$ gilt*

$$A(x + y) = Ax + Ay \textit{ und } A(kx) = kAx.$$

Beweis.
Einfacher anwenden der Definition: Für die Additivität berechnen wir den i-ten Eintrag der linken Seite:

$$\begin{aligned} (A(x+y))_i &= \sum_{j=1}^{n} a_{ij}(x_j + y_j) = \sum_{j=1}^{n} a_{ij}x_j + \sum_{j=1}^{n} a_{ij}y_j \\ &= (Ax)_i + (Ay)_i \end{aligned}$$

und die Homogenität geht ebenso einfach. $\square$

Definition 9.4. Wir nennen das Gleichungssystem $Ax = b$ *homogen*, wenn $b = 0$, sonst inhomogen.

Die Menge aller Lösungen des Gleichungssystems $Ax = b$ bezeichnen wir mit $L(A, b)$.

Beispiel.
Wir betrachten das reelle inhomogene Gleichungssystem zu

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

und suchen die Lösungen in $\mathbb{R}^3$, d.h. wir betrachten die Gleichungen

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

Die zweite Gleichung ist äquivalent zu $x_2 = x_3 - 2$, was wir in die erste Gleichung einsetzen und $x_1 + x_3 - 2 + x_3 = 2$ bekommen. Wir folgern $x_1 = 4 - 2x_3$. Damit haben wir alle Lösungen gefunden:

$$L(A, b) = \{ \begin{pmatrix} 4 - 2s \\ -2 + s \\ s \end{pmatrix} \mid s \in \mathbb{R} \}$$

Anders geschrieben, die Lösungen haben die Form

$$\begin{pmatrix} 4 \\ -2 \\ 0 \end{pmatrix} + s \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, \quad s \in \mathbb{R}.$$

Wir bemerken, dass die Vektoren

$$x = s \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix}, \quad s \in \mathbb{R}$$

die Lösungsmenge des homogenen Gleichungssystems $Ax = 0$ bilden; es gilt

$$Ax = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} -2s \\ s \\ s \end{pmatrix} = \begin{pmatrix} -2s + s + s \\ s - s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

△

Beispiel.
Es sei

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Die ersten beiden Gleichungen von $Ax = b$ lauten

$$\begin{aligned} x_1 + 2x_2 &= 1 \\ 3x_1 + 4x_2 &= 0 \end{aligned}$$

Die zweite Gleichung gibt $x_1 = -\frac{4}{3}x_2$ und Einsetzen in die erste Gleichung gibt $x_2 = \frac{3}{2}$. Es folgt, dass diese Gleichungen genau eine

Lösung haben, nämlich $x_1 = -2$, $x_2 = \frac{3}{2}$. Diese erfüllen jedoch nicht die dritte Gleichung, denn für diese x_1, x_2 gilt $5x_1 + 6x_2 = -10 + \frac{18}{2} = -1 \neq 0$. In diesem Fall ist also

$$L(A, b) = \emptyset.$$

Für die gleiche Matrix A und rechte Seite

$$c = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

gilt hingegen

$$L(A, c) = \{\begin{pmatrix} -2 \\ \frac{3}{2} \end{pmatrix}\}.$$

△

10 Der Lösungsraum, Kern, Bild und Rang

Satz 10.1. *Ist $A \in K^{m \times n}$, so ist die Menge $L(A, 0)$ der Lösungen des homogenen Gleichungssystems ein Unterraum von K^n.*

Beweis.
Wir prüfen die Unterraumkriterien:

- $0 \in L(A, 0)$, denn $A0 = 0$
- Sind x und y Lösungen, dann gilt $Ax = 0$ und $Ay = 0$ und damit
$$A(x + y) = Ax + Ay = 0 + 0 = 0;$$
also ist $x + y \in L(A, 0)$.
- Ist x Lösung und $k \in K$, dann
$$A(kx) = kAx = k0 = 0;$$
also ist $kx \in L(A, 0)$.

□

Definition 10.2. Die Lösungsmenge $L(A, 0)$ heißt *Kern* der Matrix $A \in K^{m \times n}$, geschrieben

$$\operatorname{Kern}(A) := \{x \in K^n \mid Ax = 0\}.$$

Die Menge

$$\operatorname{Bild}(A) := \{y \in K^m \mid \exists x \in K^n : Ax = y\}$$

heißt *Bild* von A.

Per Definition ist klar, dass $Ax = b$ genau dann eine Lösung hat, wenn $b \in \operatorname{Bild}(A)$ gilt.

Satz 10.3. *Für eine Matrix $A \in K^{m \times n}$ ist* $\operatorname{Kern}(A)$ *ein Unterraum von K^n und* $\operatorname{Bild}(A)$ *ein Unterraum von K^m.*

Beweis.
Die Aussage zum Kern haben wir in Satz 10.1 gezeigt (der Kern ist genau $L(A, 0)$).

Das Bild von A ist das Erzeugnis der Spalten von A, vgl. Abschitt 9. Anders geschrieben: Hat A die Spalten $a_1, \dots, a_n \in K^m$, so ist $\operatorname{Bild}(A) = \langle\{a_1, \dots, a_n\}\rangle$ und nach Satz 6.6 ein Unterraum.
□

Satz 10.4. *Hat das Gleichungssystem $Ax = b$ eine Lösung x_0, dann ist der Lösungsraum*

$$L(A, b) = x_0 + \operatorname{Kern}(A) = \{x_0 + x \mid Ax = 0\}.$$

Beweis.
Sei $x \in L(A,b)$. Dann ist $x = x_0 + (x - x_0)$ und wegen $Ax = b$ und $Ax_0 = b$ folgt $A(x - x_0) = b - b = 0$, d.h. $x - x_0 \in \mathrm{Kern}(A)$.

Sei nun $x \in x_0 + \mathrm{Kern}(A)$, d.h. es gibt ein $z \in \mathrm{Kern}(A)$ mit $x = x_0 + z$. Dann gilt $Ax = A(x_0 + z) = Ax_0 + Az = b + 0 = b$, also $x \in L(A,b)$. $\square$

Definition 10.5. Der *Rang* einer Matrix $A \in K^{m\times n}$ ist die Dimension des Bildes von A, also

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

Da das Bild von A der Spann der Spalten von A ist, ist der Rang gleich der Dimension des Spanns der Spalten oder der maximalen Anzahl von linear unabhängigen Spalten von A.

Entsprechend können wir den Raum betrachten, der von den Zeilen von A aufgespannt wird. Seine Dimension wird *Zeilenrang* genannt (und der Rang entsprechend auch *Spaltenrang*).

Beispiel.
Die Matrix

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

hat Rang 3 (der Zeilenrang ist ebenfalls 3), die Matrix

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

hingegen hat Rang 2, denn es gilt

$$-\begin{pmatrix} 1 \\ 4 \\ 7 \end{pmatrix} + 2\begin{pmatrix} 2 \\ 5 \\ 8 \end{pmatrix} = \begin{pmatrix} -1+4 \\ -4+10 \\ -7+16 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \\ 9 \end{pmatrix},$$

die ersten beiden Spalten sind jedoch linear unabhängig. Wie ist der Zeilenrang? $\triangle$

Satz 10.6. *Für jede Matrix ist Zeilenrang gleich dem (Spalten-)Rang.*

Beweis.
Sei $A \in K^{m\times n}$ mit den Spalten $a_1, \dots, a_n \in K^m$. Der i-te Eintrag der j-ten Spalte ist also $(a_j)_i = a_{ij}$, $i = 1, \dots, m$, $j = 1, \dots, n$. Es sei r der Spaltenrang von A, d.h. $\mathrm{Bild}(A)$ hat Dimension r. Ist

$$b_1 = \begin{pmatrix} b_{11} \\ \vdots \\ b_{1m} \end{pmatrix}, \dots, b_r = \begin{pmatrix} b_{r1} \\ \vdots \\ b_{rm} \end{pmatrix}$$

eine Basis von $\mathrm{Bild}(A)$, so können wir jede Spalte von A in dieser

Basis darstellen, d.h. es gilt

$$\begin{aligned} a_1 &= c_{11}b_1 + \cdots + c_{1r}b_r, \\ \vdots \quad & \quad \vdots \\ a_n &= c_{n1}b_1 + \cdots + c_{nr}b_r \end{aligned}$$

mit Koeffizienten c_{jk}, $j = 1, \ldots, n$, $k = 1, \ldots, r$.

Für die einzelnen Matrixeinträge von A heißt das

$$a_{ij} = c_{j1}b_{1i} + \cdots + c_{jr}b_{ri}. \tag{*}$$

Bezeichnen wir die Zeilen von A mit $a_1^*, \ldots, a_m^* \in K^n$, d.h. es gilt

$$a_i^* = \begin{pmatrix} a_{i1} & \cdots & a_{in} \end{pmatrix},$$

(also $(a_i^*)_j = a_{ij}$) so können wir (*) auch „zeilenweise" lesen. Für den Zeilenvektor a_1^* ergibt sich für den j-ten Eintrag nämlich

$$a_{1j}^* = c_{j1}b_{11} + \cdots + c_{jr}b_{r1},$$

also gilt für diese Zeile

$$a_1^* = b_{11}c_1^* + \cdots + b_{r1}c_r^*$$

mit $c_k^* = \begin{pmatrix} c_{1k} & \cdots & c_{nk} \end{pmatrix}$, $k = 1, \ldots, r$. Ebenso gilt für die Zeilenvektoren a_i^*

$$a_i^* = b_{1i}c_1 + \cdots b_{ri}c_r.$$

Das aber bedeutet, $a_i^* \in \langle \{c_1^*, \ldots, c_r^*\} \rangle$, d.h. die Dimension des Raumes, der von den Zeilen von A aufgespannt wird, ist höchstens r. Wir haben also „Zeilenrang $\leq$ Spaltenrang" gezeigt.

Genau analog zeigt man „Spaltenrang $\leq$ Zeilenrang". □

Satz 10.7. *Das Gleichungssystem $Ax = b$ hat genau dann eine Lösung, wenn die Matrix A den gleichen Rang wie die erweiterte Matrix $(A \mid b)$ (welche dadurch entsteht, dass b als weitere Spalte zu A hinzugefügt wird).*

Ist $A \in K^{m \times n}$ und $b \in K^m$, so ist

$$(A \mid b) = \begin{pmatrix} A & b \end{pmatrix} \in K^{m \times n+1}$$

Beweis.
Die Tatsache $\operatorname{Rang}(A) = \operatorname{Rang}(A \mid b)$ bedeutet genau, dass die Spalten von A und $(A \mid b)$ den gleichen Vektorraum aufspannen. Das aber heißt, b liegt im Spann der Spalten von A, also im Bild von A. □

In Abschnitt 13 werden wir dann folgenden Satz zeigen:

Satz 10.8. *Ist $A \in K^{m \times n}$, so hat gilt*

$$\dim(\operatorname{Kern}(A)) = n - \dim(\operatorname{Bild}(A))$$

oder anders geschrieben

$$\dim(L(A, 0)) = n - \operatorname{Rang}(A),$$

d.h. die Dimension des Lösungsraumes des homogenen Gleichungssystems ist die Anzahl der Spalten minus den Rang.

Der Lösungsraum eines linearen Gleichungssystems ist invariant bei einigen wichtigen Umformungen.

Satz 10.9. *Der Lösungsmenge $L(A, b)$ bleibt gleich, wenn eine der folgenden elementaren Zeilenumformungen mit der erweiterten Matrix $(A \mid b)$ durchgeführt wird:*

Typ 1 Addition eines Vielfachen einer Zeile zu einer anderen Zeile.

Typ 2 Vertauschen zweier Zeilen.

Typ 3 Skalieren einer Zeile mit einem Skalar $\neq 0$.

Beweis.
Eine Umformung vom Typ 2 ändert nur die Nummerierung der Gleichungen. Eine Umformung vom Typ 3 multipliziert alle Summanden und die rechte Seite der entsprechenden Gleichung mit einem Skalar. Ist dieser Skalar nicht Null, ändert das die Lösungsmenge nicht (wäre der Skalar Null, so würde diese Gleichung nach Multiplikation immer erfüllt sein, d.h. die Lösungsmenge wäre eventuell größer geworden).

Für Umformungen vom Typ 1 betrachten wir zwei Zeilen

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2. \end{aligned}$$

Die Lösungen dieser Zeilen sind für jedes $k \in K$, genau die Lösungen der Zeilen

$$\begin{aligned} a_{11}x_1 + \quad\quad a_{12}x_2 + \cdots + \quad\quad a_{1n}x_n &= b_1 \\ (ka_{11} + a_{21})x_1 + (ka_{12} + a_{22})x_2 + \cdots + (ka_{1n} + a_{2n})x_n &= kb_1 + b_2. \end{aligned}$$

$\square$

11 Der Gaußsche Eliminationsalgorithmus

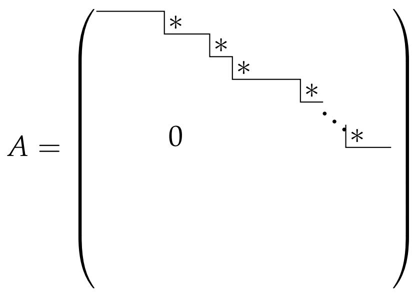
Beschäftigen wir uns nun mit der Frage, wie man tatsächlich Lösungen eines Gleichungssystems ausrechnen kann.

Eine Methode besteht darin, eine beliebige Zeile nach einer beliebigen Variable aufzulösen, das Ergebnis in eine beliebige anderen Zeile einzusetzen und so weiter zu Verfahren, bis am Ende nach einer Variable aufgelöst ist mit Hilfe derer auf ähnliche Weise eine weitere Variable bestimmt werden kann, usw.

Nach Satz 10.9 ändert sich die Lösungsmenge nicht bei elementaren Zeilenumformungen der erweiterten Matrix. Die Idee des Gaußschen Eliminationsalgorithmus ist, dies Auszunutzen um das Gleichungssystem auf eine einfacher Form zu bringen.

Der Gauß-Algorithmus ist das systematische Anwenden von elementaren Zeilenumformungen, um das Gleichungssystem auf eine Form zu bringen, in der man die Lösung einfacher ausrechnen kann. Dies gilt zum Beispiel, wenn die Matrix folgende Form hat:

Definition 11.1. Ein Matrix $A \in K^{m\times n}$ ist in *Zeilenstufenform*, wenn Sie die folgende Form hat



wobei die Einträge $*$ nicht Null sind und unterhalb der Linie nur Nullen stehen.

Formal: Bezeichnet j_i den kleinsten Index der nicht-Null Einträge in der i-ten Zeile, d.h.

$$j_i = \min\{j \mid a_{ij} \neq 0\}$$

dann muss gelten, dass $j_1 < j_2 < \cdots < j_r$.

Beispiel.
Die Matrix

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

ist in Zeilenstufenform. Für Gleichungssysteme mit dieser Matrix können wir den Lösungsraum fast ablesen: Ist die erweiterte Matrix

$$\left(\begin{array}{ccccc|c} 0 & 3 & 1 & -1 & 4 & 2 \\ 0 & 0 & 1 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right)$$

so gibt es keine Lösung, denn die letzte Gleichung lautet $0 = 1$.

Im Fall

$$\left(\begin{array}{ccccc|c} 0 & 3 & 1 & -1 & 4 & 2 \\ 0 & 0 & 1 & -1 & 0 & -1 \\ 0 & 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

so gibt es Lösungen und wir können uns von unten nach oben arbeiten: Die vierte Gleichung ist $0 = 0$ und spielt keine Rolle. Die dritte Gleichung ist

$$2x_5 = 0, \text{ also } x_5 = 0.$$

Die zweite Gleichung ist

$$x_3 - x_4 = -1.$$

Wir lösen nach der Variable mit dem kleineren Index auf, also nach x_3, und bekommen $x_3 = x_4 - 1$. Die erste Gleichung lautet

$$3x_2 + 1x_3 - x_4 + 4x_5 = 2.$$

Wir setzten die Ergebnisse für x_3 und x_5 ein (x_4 ist ja ein freier Parameter) und bekommen

$$3x_2 + (x_4 - 1) - x_4 + 0 = 2, \text{ also } x_2 = 1.$$

Für x_1 gibt es keine Bedingung und wir erhalten den Lösungsraum

$$\begin{aligned} L(A,b) &= \{ \begin{pmatrix} x_1 \\ 1 \\ x_4 - 1 \\ x_4 \\ 0 \end{pmatrix} \mid x_1, x_4 \in \mathbb{R} \} \\ &= \{ \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \mid s, t \in \mathbb{R} \} \end{aligned}$$

△

Allgemein gilt: Hat die Zeilenstufenform von A genau r Stufen, dann ist der Spalten- und der Zeilenrang r (wir wussten schon aus Satz (*) dass beide gleich sind).

Kommen wir nun zum Gaußschen Eliminationsalgorithmus, der im Wesentlichen aus nur einer Idee besteht:

Das Arbeiten mit der Zeilenstufenform wird noch einfacher, wenn wir auch noch Spaltenvertauschungen erlauben. Ein Vertauschung der Spalten j_1 und j_2 bedeutet allerdings, dass die Variablen x_{j_1} und x_{j_2} ihre Rolle tauschen. Wir müssen also über Spaltenvertauschungen Buch führen, um am Ende den Lösungsraum eines Gleichungssystems angeben zu können. Durch Spaltenvertauschungen können wir folgende Form erreichen:

$$\begin{pmatrix} * & & & & \\ & * & & & \\ & & \ddots & & \\ & & & * & \\ & & & & \\ & & & & \end{pmatrix}$$

Sei unsere Matrix (oder erweiterte Matrix) in der Form

$$\left(\begin{array}{cccc|c} a_{11} & & & & \\ & a_{22} & * & & * \\ & & \ddots & & \\ & 0 & & a_{kk} & \\ \hline & & 0 & & B \end{array}\right)$$

wobei alle $a_{11}, \dots, a_{kk}$ nicht Null sind und unten rechts eine Matrix B steht.

Ist B die Nullmatrix, sind wir fertig, wenn nicht, hat B einen Eintrag, der nicht Null ist, den wir durch vertauschen von Zeilen und Spalten an die Stelle $(k+1, k+1)$ bringen können. Durch Addition von geeigneten Vielfachen der $k+1$-ten Zeile zu den Zeilen darunter bekommen wir die Form

$$\left(\begin{array}{ccccc|c} a_{11} & & & & & \\ & a_{22} & * & & & * \\ & & \ddots & & & \\ & & & a_{kk} & & \\ & 0 & & & a_{k+1,k+1} & \\ \hline & & 0 & & & B' \end{array}\right)$$

mit einer neuen, kleineren, Matrix B'. Wir verfahren so weiter, bis die Methode abbricht.

Mit dem Gaußsche Eliminationsverfahren können wir nicht nur den Lösungsraum von Gleichungssystemen bestimmen, sondern auch den Rang einer Matrix und ihren Kern.

Beispiel.

Es sei

$$A = \begin{pmatrix} 2 & 3 & 0 & -1 \\ 4 & 3 & 2 & 0 \\ -1 & 0 & -2 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

und gesucht sind der Lösungsraum, Rang von A und der Kern von A.

Wir bilden die erweiterte Matrix und führen das Gaußsche

Eliminationsverfahren durch:

$$
\begin{aligned}
&\begin{pmatrix} 2 & 3 & 0 & -1 & 1 \\ 4 & 3 & 2 & 0 & 1 \\ -1 & 0 & -2 & 1 & 1 \end{pmatrix} \begin{array}{l} \cdot(-2) \quad \cdot\tfrac{1}{2} \\ \leftarrow + \\ \leftarrow \quad + \end{array} \\
\rightsquigarrow &\begin{pmatrix} 2 & 3 & 0 & -1 & 1 \\ 0 & -3 & 2 & 2 & -1 \\ 0 & \frac{3}{2} & -2 & \frac{1}{2} & \frac{3}{2} \end{pmatrix} \begin{array}{l} \\ \\ \mid \cdot 2 \end{array} \\
\rightsquigarrow &\begin{pmatrix} 2 & 3 & 0 & -1 & 1 \\ 0 & -3 & 2 & 2 & -1 \\ 0 & 3 & -4 & 1 & 3 \end{pmatrix} \begin{array}{l} \\ \\ \leftarrow + \end{array} \\
\rightsquigarrow &\begin{pmatrix} 2 & 3 & 0 & -1 & 1 \\ 0 & -3 & 2 & 2 & -1 \\ 0 & 0 & -2 & 3 & 2 \end{pmatrix}.
\end{aligned}
$$

Die Matrix ist ohne Spaltentausch in Zeilenstufenform überführt und wir können den Lösungsraum durch Rückwärtseinsetzen erhalten: Die letzte Gleichung liefert

$$-2x_3 + 3x_4 = 2 \rightsquigarrow x_3 = \tfrac{3}{2}x_4 - 1.$$

Damit bekommen wir in der zweiten Gleichung

$$-3x_2 + 2(\tfrac{3}{2}x_4 - 1) + 2x_4 = -1 \rightsquigarrow x_2 = \tfrac{5}{3}x_4 - \tfrac{1}{3}$$

was in der ersten Zeile

$$2x_1 + 3(\tfrac{5}{3}x_4 - \tfrac{1}{3}) + 0x_3 - x_4 = 1 \rightsquigarrow x_1 = -2x_4 + 1.$$

Die Lösungen haben also die Form

$$\begin{pmatrix} -2x_4 + 1 \\ \frac{5}{3}x_4 - \frac{1}{3} \\ \frac{3}{2}x_4 - 1 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ -\frac{1}{3} \\ -1 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -2 \\ \frac{5}{3} \\ \frac{3}{2} \\ 1 \end{pmatrix}$$

Das beschreibt den Lösungsraum sowie den Kern:

$$L(A,b) = \{ \begin{pmatrix} 1 \\ -\frac{1}{3} \\ -1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -2 \\ \frac{5}{3} \\ \frac{3}{2} \\ 1 \end{pmatrix} \mid t \in \mathbb{R} \},\ \mathrm{Kern}(A) = \langle \begin{pmatrix} -2 \\ \frac{5}{3} \\ \frac{3}{2} \\ 1 \end{pmatrix} \rangle$$

Es gilt also $\mathrm{Rang}(A) = 3$, $\dim(\mathrm{Kern}(A)) = 1$. △

12 Lineare Abbildungen

Die Multiplikation einer Matrix $A \in K^{m\times n}$ können wir als Abbildung von K^n nach K^m auffassen welche $f : v \mapsto Av$ abbildet. Diese Abbildung hat die folgenden Eigenschaften:

Siehe Satz 9.3.

Additivität: $f(v+w) = A(v+w) = Av + Aw = f(v) + f(w)$

Homogenität: $f(kv) = A(kv) = kAv = kf(v)$.

Der Begriff *lineare Abbildung* abstrahiert diesen Begriff für allgemeine Vektorräume:

Definition 12.1. Es seien V und W zwei K-Vektorräume. Eine Abbildung $f : V \to W$ heißt *linear*, wenn für alle $v_1, v_2 \in V$ und alle $k \in K$ gilt

$$f(v_1 + kv_2) = f(v_1) + kf(v_2).$$

Aus dieser Definition folgen Additivität und Homogenität: Zuerst bemerken wir, dass $f(0) = f(0+0) = f(0) + f(0)$ und daher $f(0) = 0$ gilt. Mit $v_1 = 0$ folgt also $f(kv_2) = kf(v_2)$ und mit $k = 1$ folgt die Additivität.

Eine lineare Abbildung von V nach W heißt auch *Homomorphismus* von V nach W.

Eine lineare und bijektive Abbildung von V nach W heißt *Isomorphismus*. Gibt es zwischen V und W einen Isomorphismus, heißen V und W *isomorph*.

Sind V und W isomorph, so sind sie in gewissem Sinne „gleich“, da sich strukturverträglich und eins-zu-eins aufeinander abbilden lassen.

Beispiel.

Lineare Abbildungen, die nichts mit Matrizen zu tun haben.

- Für jeden Vektorraum V ist die identische Abbildung $\mathrm{id}_V : V \to V$ linear.
- Für alle Vektorräume V und W ist die Nullabbildung linear ($\forall v \in V : f(v) = 0$).
- Sei V der Vektorraum aller Funktionen des Intervalls $[0,1]$ nach $\mathbb{R}$. Dann ist $F : V \to \mathbb{R}$, $F(f) = f(0)$ eine lineare Abbildung.
- Ist V der Vektorraum aller konvergenten Folgen, dann ist $f : V \to \mathbb{R}$ definiert durch $f(a) = \lim_{n\to\infty} a_n$ eine lineare Abbildung.
- Ist V der Vektorraum aller stetig differenzierbaren Funktionen und W der Vektorraum der stetigen Funktionen, so ist $F : V \to W$ definiert durch $F(f) = f'$ eine lineare Abbildung.

$\triangle$

Vollkommen analog zu Matrizen definieren wir

$$\begin{aligned} \mathrm{Bild}(f) &= \{w \in W \mid \exists v \in V :\ w = f(v)\}, \\ \mathrm{Kern}(f) &= \{v \in V \mid f(v) = 0\} \end{aligned}$$

und vollkommen analog zum Fall von Matrizen folgt:

Damit können wir schreiben:
f surjektiv $\iff$ $\mathrm{Bild}(f) = W$
f injektiv $\iff$ $\mathrm{Kern}(f) = \{0\}$.

Satz 12.2. *Ist* $f : V \to W$ *linear, so ist* $\mathrm{Bild}(f)$ *und Unterraum von* W *und* $\mathrm{Kern}(f)$ *ein Unterraum von* V.

Hier eine wichtige Konstruktion von linearen Abbildungen zwischen endlichdimensionalen Vektorräumen:

Satz 12.3. *Es sei* $B = \{v_1, \dots, v_n\}$ *eine Basis von* V *und es seien* $w_1, \dots, w_n$ *beliebige Vektoren in einem Vektorraum* W. *Dann gibt es genau eine lineare Abbildung* $f : V \to W$ *mit* $f(v_i) = w_i, i = 1, \dots, n$.

Beweis.
Wir definieren f für einen beliebigen Vektor $v \in V$ wie folgt: Wir stellen v in der Basis B dar, d.h. wir schreiben

$$v = \sum_{i=1}^{n} k_i v_i$$

und setzen

$$f(v) = \sum_{i=1}^{n} k_i w_i.$$

Dies ist eine lineare Abbildung, denn für $u \in V$ mit $u = \sum_{i=1}^{n} h_i v_i$ und $k \in K$ gilt

$$\begin{aligned} f(u + kv) &= f(\sum_{i=1}^{n} h_i v_i + k \sum_{i=1}^{n} k_i v_i) \\ &= f(\sum_{i=1}^{n} (h_i + k k_i) v_i) \\ &= \sum_{i=1}^{n} (h_i + k k_i) w_i \\ &= \sum_{i=1}^{n} h_i w_i + k \sum_{i=1}^{n} k_i v_i = f(u) + k f(v). \end{aligned}$$

Für die Eindeutigkeit nehmen wir an, dass g eine weitere solche Abbildung ist, d.h., es gilt $g(v_i) = w_i$. Dann gilt aber für jedes $v = \sum_{i=1}^{n} k_i v_i$:

$$\begin{aligned} f(v) &= f(\sum_{i=1}^{n} k_i v_i) = \sum_{i=1}^{n} k_i f(v_i) \\ &= \sum_{i=1}^{n} k_i g(v_i) = g(\sum_{i=1}^{n} k_i v_i) = g(v). \end{aligned}$$

$\square$

Lineare Abbildungen werden also durch die Bilder der Vektoren einer Basis vollständig beschrieben. Man kann auch einige Eigenschaften der Abbildung an den Bildern einer Basis ablesen:

Man sagt auch, dass lineare Abbildungen durch die Bilder der Vektoren einer Basis *charakterisiert* werden.

Satz 12.4. *Es sei* $B = \{v_1, \dots, v_n\}$ *eine Basis von* V *und* $f : V \to W$ *die lineare Abbildung mit* $f(v_i) = w_i$. *Dann gilt: Die Vektoren* $\{w_1, \dots, w_n\}$ *sind* *i) linear unabhängig genau dann, wenn f injektiv ist,*

ii) eine Erzeugendensystem genau dann, wenn f surjektiv ist,

iii) eine Basis, genau dann, wenn f bijektiv ist.

Beweis.

i) Seien die $w_1, \dots, w_n$ linear unabhängig und gelte $f(v) = f(v')$. Zu zeigen ist $v = v'$.

Wir stellen v und v' in der Basis B dar

$$v = \sum_{i=1}^{n} k_i v_i, \quad v' = \sum_{i=1}^{n} k_i' v_i.$$

Da $f(v) = f(v')$, folgt

$$\sum_{i=1}^{n} k_i w_i = \sum_{k=1}^{n} k_i' w_i$$

Ist f injektiv, so betrachten wir eine Linearkombination

$$0 = \sum_{i=1}^{n} k_i w_i = f(\sum_{i=1}^{n} k_i v_i).$$

Da f injektiv ist, folgt, dass $\sum_{i=1}^{n} k_i v_i = 0$ gilt, und daher sind alle k_i Null. Dies zeigt, dass die w_i linear unabhängig sind.

ii) Sei $\{w_1, \dots, w_n\}$ ein Erzeugendensystem und sei $w \in W$. Wir zeigen, dass ein $v \in V$ mit $f(v) = w$ existiert. Da w ein Erzeugendensystem ist, können wir w durch die w_i darstellen, also

$$w = \sum_{i=1}^{n} k_i w_i.$$

Mit $v = \sum_{k=1}^{n} k_i v_i$ folgt $f(v) = w$, wie gefordert.

Sei nun f surjektiv und sei $w \in W$. Dann existiert $v \in V$ mit $f(v) = w$. Darstellen von v in der Basis B liefert

$$w = f(v) = f(\sum_{i=1}^{n} k_i v_i) = \sum_{i=1}^{n} k_i w_i$$

was zeigt, dass die w_i ganz W erzeugen.

iii) Ist eine direkte Folgerung aus i) und ii).

□

Korollar 12.5. *Ist V ein Vektorraum mit Basis $B = (v_1, \dots, v_n)$, so gibt es genau einen Isomorphismus*

$$\Phi_B : K^n \to V \quad \textit{mit} \quad \Phi_B(e_j) = v_j,$$

insbesondere ist V isomorph zu K^n.

Korollar 12.6. *Zu $f : K^n \to K^m$ linear gibt es genau eine Matrix $A \in K^{m\times n}$, so dass für alle $x \in K^n$ gilt*

$$f(x) = Ax.$$

$A = \begin{pmatrix} f(e_1) & \cdots & f(e_n) \end{pmatrix}$ tut es.

Satz 12.7. *Es sei $f : V \to W$ eine lineare Abbildung, V endlich-dimensional,* $\dim(\mathrm{Kern}(f)) = k$, *$(v_1, \dots, v_k)$ eine Basis von* $\mathrm{Kern}(f)$, $\dim(\mathrm{Bild}(f)) = r$ *und $(w_1, \dots, w_r)$ eine Basis von* $\mathrm{Bild}(f)$. *Weiter seien $u_1, \dots, u_r \in V$, mit $f(u_1) = w_1, \dots f(u_r) = w_r$. Dann ist $B = (v_1, \dots, v_k, u_1, \dots, u_r)$ eine Basis von V.*

Beweis.
Es sei $v \in V$. Wir zeigen, dass v von B erzeugt wird. Da $(w_1, \dots, w_r)$ Basis von $\mathrm{Bild}(f)$ ist, gibt es Koeffizienten $\alpha_1, \dots, \alpha_r$ mit $f(v) = \alpha_1 w_1 + \cdots + \alpha_r w_r$. Setzen wir $v' = \alpha_1 u_1 + \cdots + \alpha_r u_r$, so gilt $f(v) = f(v')$, also $f(v - v') = 0$, d.h. $v - v' \in \mathrm{Kern}(f)$. Also können wir $v - v'$ in $(v_1, \dots, v_k)$ entwickeln, d.h. es gibt $\beta_1, \dots, \beta_k$ mit $v - v' = \beta_1 v_1 + \cdots + \beta_k v_k$. Es folgt

$$v = \beta_1 v_1 + \cdots + \beta_k v_k + \alpha_1 u_1 + \cdots + \alpha_r u_r,$$

d.h. v liegt im Spann von B.

Zeigen wir, dass B linear unabhängig ist. Sei also

$$\beta_1 v_1 + \cdots + \beta_k v_k + \alpha_1 u_1 + \cdots + \alpha_r u_r = 0.$$

Anwendung von f zeigt (wegen $v_i \in \mathrm{Kern}(f)$)

$$\alpha_1 w_1 + \cdots \alpha_r w_k = 0,$$

also $\alpha_1, \dots, \alpha_r = 0$. Jetzt folgt

$$\beta_1 v_1 + \cdots + \beta_k v_k = 0$$

und daher $\beta_1, \dots, \beta_k = 0$. □

Korollar 12.8. *Es gilt die Dimensionsformel*

$$\dim(V) = \dim(\mathrm{Bild}(f)) + \dim(\mathrm{Kern}(f)).$$

Im obigen Beweis ist $\dim(\mathrm{Kern}(f)) = k$, $\dim(\mathrm{Bild}(f)) = r$ und wir zeigen, dass $\dim(V) = k + r$.

13 Darstellung von linearen Abbildungen durch Matrizen

Wie wir schon gesehen haben, lassen sich lineare Abbildungen $f : V \to W$ durch die Bilder von Vektoren einer Basis B in V beschreiben. Wählen wir auch noch eine Basis C in W, so können wir f über eine Matrix beschreiben.

Definition 13.1. Ist $f : V \to W$ linear, $B = \{v_1, \ldots, v_n\}$ eine Basis in V und $\{w_1, \ldots, w_m\}$ eine Basis in W, so bestimmen wir Skalare a_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$ durch die Entwicklung der $f(v_j)$ in die Basis C, nämlich

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i.$$

Die zugehörige Matrix

Merke: Die Koeffizienten des Bildes von v_j bilden die j-te Spalte von A.

$$A = M_C^B(f) = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \in K^{m\times n}$$

nennen wir *Darstellungsmatrix von f bezüglich B und C*.

Ist $v = \sum_{j=1}^n b_j v_j$, so ist $f(v) = f(\sum_j b_j v_j) = \sum_j b_j f(v_j) = \sum_{i=1}^m \sum_{j=1}^n a_{ij} b_j w_j$, also bildet A die Entwicklungskoeffizienten von v auf die Entwicklungskoeffizienten von $w = f(v)$ ab.

Umgekehrt gibt es nach Satz 12.3 zu jeder Matrix $A \in K^{m\times n}$ und Basen B in V und C in W genau eine lineare Abbildung f, die

$$f(v_j) = \sum_{i=1}^{m} a_{ij} w_i$$

erfüllt.

Definition 13.2. Für zwei Vektorräume V und W ist

$$\mathrm{Hom}(V, W) = \{f : V \to W \mid f \text{ linear}\}$$

die Menge der Homomorphismen von V nach W.

Satz 13.3. $\mathrm{Hom}(V, W)$ *ist wieder ein Vektorraum.*

Beweis.
Die Menge *aller* Abbildungen W^V ist ein Vektorraum. Es reicht also zu zeigen, dass $\mathrm{Hom}(V, W)$ ein Unterraum davon ist. Sind also $f, g \in \mathrm{Hom}(V, W)$, so muss $f + kg \in \mathrm{Hom}(V, W)$ gelten. Wir müssen also prüfen, ob $f + kg$ eine lineare Abbildung ist.

Sind $u, v \in V$ und $h \in K$, dann gilt

$$\begin{aligned}(f + kg)(u + hv) &= f(u + hv) + kg(u + hv)\\ &= f(u) + hf(v) + kg(u) + khg(v)\\ &= (f + kg)(u) + h(f + kg)(v),\end{aligned}$$

was die Behauptung zeigt. $\square$

Satz 13.4. *Sind V und W K-Vektorräume mit* $\dim(V) = n, \dim(W) = m$ *und sind B und C Basen von V bzw. W, so ist*

$$M_C^B : \mathrm{Hom}(V, W) \to K^{m \times n}$$

ein Isomorphismus.

Beweis.
Seien $f, g \in \mathrm{Hom}(V, W)$, $R = M_C^B(f)$ und $S = M_C^B(g)$ die zugehörigen Matrizen. Dann gilt für jedes $k \in K$ und jedes j

$$\begin{aligned}(f + kg)(v_j) &= f(v_j) + kg(v_j) \\ &= \sum_{i=1}^m r_{ij} w_i + k \sum_{i=1}^m s_{ij} w_i \\ &= \sum_{i=1}^m (r_{ij} + k s_{ij}) w_i\end{aligned}$$

und das heißt

$$M_C^B(f + kg) = M_C^B(f) + k M_C^B(g).$$

Dies zeigt, dass M_C^B eine lineare Abbildung ist.

Die Surjektivität von M_C^B haben wir direkt nach Definition 13.1 gezeigt. Für die Injektivität folgt aus Satz 12.3 welcher sagt, dass eine linear Abbildung durch die Bilder einer Basis eindeutig bestimmt ist. □

Beispiel.
Es sei $V = W = \mathbb{R}^2$ und $f(x, y) = \begin{pmatrix} 3x - 2y \\ x + y \end{pmatrix}$.

- Wir wählen (willkürlich) die Basen

  $$B = \{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}\}, \quad C = \{\begin{pmatrix} 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 2 \end{pmatrix}\}$$

  Die Einträge a_{11} und a_{21} können wir berechnen, indem wir

  $$f(v_1) = f(1, 1) = \begin{pmatrix} 3 - 2 \\ 1 + 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} = a_{11} w_1 + a_{21} w_2$$

  lösen. Diese Gleichung entspricht dem Gleichungssystem

  $$\begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{12} \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

  Wir erhalten $a_{11} = -3$ und $a_{21} = \frac{5}{2}$. Für die Einträge a_{12} und a_{22} müssen wir

  $$f(v_2) = f(-1, 1) = \begin{pmatrix} -3 - 2 \\ -1 + 1 \end{pmatrix} = \begin{pmatrix} -5 \\ 0 \end{pmatrix} = a_{12} w_1 + a_{22} w_2$$

  lösen. Wir bekommen $a_{12} = 5$ und $a_{22} = -\frac{5}{2}$. Insgesamt:

  $$M_C^B(f) = \begin{pmatrix} -3 & 5 \\ \frac{5}{2} & -\frac{5}{2} \end{pmatrix}.$$

- Bezüglich der Standardbasen, also $B = \{e_1, e_2\} = C$, müssen wir weniger rechnen: Es gilt

$$f(v_1) = f(1,0) = \begin{pmatrix} 3 \\ 1 \end{pmatrix} = 3e_1 + 1e_2,$$
$$f(v_2) = f(0,1) = \begin{pmatrix} -2 \\ 1 \end{pmatrix} = -2e_1 + e_2.$$

und es folgt

$$M_C^B(f) = \begin{pmatrix} 3 & -2 \\ 1 & 1 \end{pmatrix}$$

△

Korollar 13.5. *Für die Identität $\mathrm{id}_V : V \to V$ auf einem n-dimensionalen Vektorraum V und jede Basis B von V gilt*

$$M_B^B(\mathrm{id}_V) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} \in K^{n \times n}.$$

Diese Matrix heißt *Einheitsmatrix* und wird mit E_n oder I_n bezeichnet.

Folgt direkt aus der Tatsache, dass $\mathrm{id}_V(v_j) = v_j$ für jeden Basisvektor v_j.

Das ursprüngliche Korollar war exakt Korollar 12.5 und hätte gelöscht werden sollen...

Satz 13.7. *Es seien V und W endlichdimensionale Vektorräume mit Basen $B = \{v_1, \dots, v_n\}$ und $C = \{w_1, \dots, w_m\}$ von V bzw. W, $f \in \mathrm{Hom}(V, W)$ und sei $A = M_C^B(f)$. Dann gilt*

$$\dim(\mathrm{Bild}(A)) = \dim(\mathrm{Bild}(f)) \text{ und } \dim(\mathrm{Kern}(A)) = \dim(\mathrm{Kern}(f)).$$

Beweis.
Wegen $f(v_j) = \sum_{i=1}^n a_{ij} w_i$ gilt für jede Linearkombination

$$\begin{aligned} \sum_{j=1}^m k_j a_j = 0 &\iff \sum_{j=1}^m k_j a_{ij} = 0 \; \forall i = 1, \dots, n \\ &\iff \sum_{i=1}^n \sum_{j=1}^m k_j a_{ij} w_i = 0 \qquad (\{w_i\} \text{ Basis}) \\ &\iff \sum_{j=1}^m k_j \sum_{i=1}^n a_{ij} w_i = 0 \iff \sum_{j=1}^m k_j f(v_j) = 0 \end{aligned}$$

Wir sehen also, dass eine Linearkombination der a_j genau dann Null ergibt, wenn die entsprechende Linearkombination der $f(v_j)$ Null ergibt. Daraus folgt die Behauptung. Die Behauptung für den Kern folgt analog. □

Insbesondere folgt, dass $\dim(\mathrm{Bild}(A))$ für jede Darstellungsmatrix von f gleich ist. Da wir die Größe $\dim(\mathrm{Bild}(A))$ auch

Rang(A) genannt haben, bietet es sich an auch für lineare Abbildungen f zu definieren

$$\mathrm{Rang}(f) := \dim(\mathrm{Bild}(f)).$$

Als Korollar aus diesem Satz und Korollar 12.8 erhalten wir den in Abschnitt 10 in Satz 10.8 angekündigte Dimensionsformel für Matrizen $A \in K^{m\times n}$:

$$\begin{aligned}\dim(\mathrm{Kern}(A)) &= \dim(\mathrm{Kern}(f))\\ &= \dim(K^n) - \dim(\mathrm{Bild}(f)) = n - \dim(\mathrm{Bild}(A)).\end{aligned}$$

Linear Abbildungen kann man verknüpfen. Was passiert mit den zugehörigen Matrizen? Dies beantwortet dieser Satz:

Satz 13.8. *Es seien V_1, V_2, V_3 Vektorräume mit den Dimensionen n_1, n_2, n_3 und $f : V_1 \to V_2$ und $g : V_2 \to V_3$ linear. Sei weiter B_i eine Basis von V_i und $A = M^{B_1}_{B_2}(f) \in K^{n_2\times n_1}$ und $B = M^{B_2}_{B_3}(g) \in K^{n_3\times n_1}$ dann hat*

$$C = M^{B_1}_{B_3}(g\circ f) \in K^{n_3\times n_1} \textit{ die Einträge } c_{ki} = \sum_{j=1}^{n_2} b_{kj}a_{ji}.$$

Beweis.
Wir bezeichnen die Basen mit $B_1 = \{u_i\}_{i=1,\dots,n_1}$, $B_2 = \{v_j\}_{j=1,\dots,n_2}$, $B_3 = \{w_k\}_{k=1,\dots,n_3}$. Wir finden die Matrix-Einträge c_{ki} in der Darstellung von $g(f(u_i))$ in der Basis $\{w_k\}_{k=1,\dots,n_3}$, nämlich in

$$g(f(u_i)) = \sum_{k=1}^{n_3} c_{ki}w_k.$$

Wir benutzen

$$f(u_i) = \sum_{j=1}^{n_2} a_{ji}v_j \text{ und } g(v_j) = \sum_{k=1}^{n_3} b_{kj}w_k$$

und bekommen

$$\begin{aligned}g(f(u_i)) = g\left(\sum_{j=1}^{n_2} a_{ij}v_j\right) &= \sum_{j=1}^{n_2} a_{ij}g(v_j)\\ = \sum_{j=1}^{n_2} a_{ji}\sum_{k=1}^{n_3} b_{kj}w_k &= \sum_{k=1}^{n_3}\sum_{j=1}^{n_2} b_{kj}a_{ji}w_k.\end{aligned}$$

□

14 Matrix Multiplikation und Inversion

Dieser Satz motiviert die Definition der Multiplikation von Matrizen analog zu deren Verknüpfung (wenn die als lineare Abbildungen gesehen werden).

Definition 14.1. Für zwei Matrizen $B \in K^{\ell\times m}$ und $A \in K^{m\times n}$ definieren wir die *Matrix-Multiplikation* $C = BA \in K^{\ell\times n}$ durch

$$c_{ik} = \sum_{j=1}^{m} b_{ij}a_{jk}.$$

Damit das Matrix-Produkt BA erklärt ist, muss die Anzahl der Spalten von B gleich der Anzahl der Zeilen von A sein.

$$\begin{array}{cc} B & A \\ \cap & \cap \\ K^{\ell\times m} & K^{m\times n} \end{array}$$

Diese Definition und der Satz davor liefern zusammen:

$$M_{B_3}^{B_1}(g\circ f) = M_{B_3}^{B_2}(g)M_{B_2}^{B_1}(f).$$

Anders gesagt: Matrix-Multiplikation, Darstellung von linearen Abbildung und Komposition von linearen Abbildungen sind verträglich.

Beispiel.

- Die Matrix-Multiplikation verallgemeinert die Matrix-Vektor Multiplikation. Ein Vektor $x \in K^n$ ist ebenfalls eine Matrix $x \in K^{n\times 1}$ (und damit auch eine lineare Abbildung $K^1 \to K^n$ definiert durch $t \mapsto tx$). Für eine Matrix $A \in K^{m\times n}$ gilt nämlich $Ax \in K^{m\times 1}$ mit

$$(Ax)_{ik} = \sum_{j=1}^{n} a_{ij}x_{jk},\ i = 1,\dots,m,\ k = 1.$$

- Ein kleines Zahlenbeispiel: Das Produkt von

$$A = \begin{pmatrix} 2 & -1 & 2 \\ 0 & -3 & 1 \\ 4 & 0 & -4 \end{pmatrix},\ \text{und } B = \begin{pmatrix} -1 & 1 \\ 3 & 0 \\ 2 & -1 \end{pmatrix}$$

ist

$$\begin{aligned} AB &= \begin{pmatrix} 2 & -1 & 2 \\ 0 & -3 & 1 \\ 4 & 0 & -4 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 3 & 0 \\ 2 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 2\cdot(-1)+(-1)\cdot 3+2\cdot 2 & 2\cdot 1+(-1)\cdot 0+2\cdot(-1) \\ 0\cdot(-1)+(-3)\cdot 3+1\cdot 2 & 0\cdot 1+(-3)\cdot 0+1\cdot(-1) \\ 4\cdot(-1)+0\cdot 3+(-4)\cdot 2 & 4\cdot 1+0\cdot 0+(-4)\cdot(-1) \end{pmatrix} \\ &= \begin{pmatrix} -1 & 0 \\ -7 & -1 \\ -12 & 8 \end{pmatrix}. \end{aligned}$$

- Die Multiplikation von Matrizen lässt sich auch wie folgt verstehen: Ist $A \in K^{\ell\times m}$ und $B \in K^{m\times n}$ und schreiben wir B spaltenweise auf als

$$B = \begin{pmatrix} b_1 & \dots & b_n \end{pmatrix} \text{ mit } b_k \in K^m,\ k = 1,\dots n,$$

so sind die Spalten von AB genau die $Ab_k \in K^\ell, k = 1, \dots, n$, also

$$AB = \begin{pmatrix} Ab_1 & \dots & Ab_n \end{pmatrix}.$$

- Falls AB definiert ist, muss BA nicht definiert sein: Für $A \in K^{m\times n}$ und $B \in K^{n\times k}$ mit $m \neq k$ ist AB aber nicht BA definiert.

- Falls AB und BA beide definiert sind, müssen sie nicht gleich groß sein: Für einen Spaltenvektor $x \in K^m$ (aufgefasst als Matrix in $K^{m\times 1}$) und einen Zeilenvektor $y \in K^{1\times n}$ ist das Produkt xy definiert:

$$xy = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \begin{pmatrix} y_1 & \dots y_n \end{pmatrix} = \begin{pmatrix} x_1y_1 & x_1y_2 & \dots & x_1y_n \\ \vdots & \vdots & & \vdots \\ x_my_1 & x_my_2 & \dots & x_my_n \end{pmatrix} \in K^{m\times n}.$$

Ist $m = n$, also $y \in K^{1\times m}$ so ist auch yx definiert, ist aber ein völlig anderes Objekt

$$yx = \begin{pmatrix} y_1 & \dots & y_n \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} x_1y_1 + \cdots x_my_m \end{pmatrix} \in K^{1\times 1} = K$$

Analog: Sind $A \in K^{m\times n}$, $B \in K^{n\times m}$, dann $AB \in K^{m\times m}$ aber $BA \in K^{n\times n}$.

- Falls AB und BA beide definiert und gleich groß sind, so müssen sie nicht gleich sein. Ganz im Gegenteil, typischerweise sind AB und BA *nicht* gleich. Für $A, B \in \mathbb{R}^{2\times 2}$ mit

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

gilt zum Beispiel

$$AB = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 4 \\ 0 & 2 \end{pmatrix},$$

und

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

$\triangle$

Definition 14.2. Eine Matrix $M \in K^{n\times n}$ heißt *invertierbar* (auch *regulär*), wenn es eine andere Matrix M' gibt, für die gilt $M'M = E_n = MM'$. Wir schreiben M^{-1} für die inverse Matrix.

Die Inverse ist übrigens eindeutig, denn ist M'' ebenfalls eine Inverse, dann gilt $M'' = M''E_n = M''MM' = E_nM' = M'$.

Der Exponent $^{-1}$ an der inversen Matrix ist kein Kehrwert und sollte niemals als „eins durch…" gelesen werden.

Fassen wir Matrizen als lineare Abbildungen auf, bedeutet Invertierbarkeit nichts anderes, als dass es eine Umkehrabbildung gibt.

Satz 14.3. *Sei $A \in K^{n\times n}$. Folgende Aussagen sind äquivalent:*

i) Die Abbildung $f(x) = Ax$ ist umkehrbar.

ii) A ist invertierbar

iii) A hat Rang n

Man sagt auch „A hat *Vollrang*".

Beweis.

i) $\implies$ ii) Ist die Abbildung $f(x) = Ax$ umkehrbar, so gibt es $g : K^n \to K^n$ linear, so dass $g(f(x)) = x = f(g(x))$. Wir bemerken $A = M_E^E(f)$ (wobei E die Standardbasis ist) und setzen $A' = M_E^E(g)$ und sehen, dass $AA' = A'A = M_E^E(\mathrm{id}_{K^n}) = E_n$.

ii) $\implies$ iii) Wir zeigen die Kontraposition: Sei $\mathrm{Rang}(A) < n$ (also $\dim(\mathrm{Bild}(A)) < n$). D.h. es gibt $y \notin \mathrm{Bild}(A)$, und daher gibt es kein A' mit $AA'y = y$. Also ist A nicht invertierbar.

iii) $\implies$ i) A habe Rang n, d.h. die Spalten $a_1, \ldots, a_n$ bilden eine Basis von K^n. Also bildet $f(x) = Ax$ die Standardbasis $e_1, \ldots, e_n$ auf eine Basis ab und nach Satz 12.4 iii) ist f bijektiv, also umkehrbar.

□

Beispiel.

Die Matrix
$$A = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$$
(mit belibigem $a \in K$) ist invertierbar mit inverser Matrix
$$A^{-1} = \begin{pmatrix} 1 & -a \\ 0 & 1 \end{pmatrix}$$
Denn es gilt
$$AA^{-1} = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & -a \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -a+a \\ 0 & 1 \end{pmatrix} = E_2$$
und
$$A^{-1}A = \begin{pmatrix} 1 & -a \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a-a \\ 0 & 1 \end{pmatrix} = E_2$$
Die Matrix
$$B = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$$
ist nicht invertierbar, da $\mathrm{Rang}(B) = 1$. △

Es hätte gereicht, wenn wir nur $A^{-1}A = E_n$ in der Definition der inversen Matrix gefordert hätten: Ist A invertierbar, so ist es auch A^{-1} (wegen Satz 14.3 i)) und wir folgern

Die Einheitsmatrix E_n erfüllt übrigens $E_n x = x$ und daher auch $AE_n = A$. Ebenso gilt $E_n A = A$.

$$AA^{-1} = E_n AA^{-1} = (A^{-1})^{-1} \underbrace{A^{-1}A}_{=E_n} A^{-1} = (A^{-1})^{-1}A^{-1} = E_n.$$

Beispiel.
Mit Hilfe der inversen Matrix lassen sich quadratische Gleichungssysteme lösen: Multiplizieren wir beide Seiten von $Ax = b$ von links mit A^{-1} bekommen wir $A^{-1}Ax = A^{-1}b$ und da $A^{-1}A = E_n$ und $E_n x = x$ gilt, folgt $x = A^{-1}b$.

Konkret ist z.B. die inverse Matrix zu

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

gegeben durch

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

(es gilt

$$\begin{aligned} A^{-1}A &= \begin{pmatrix} 2 & 0 & -1 \\ 0 & 1/2 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 2+0-1 & 0+0+0 & 2+0-2 \\ 0+0+0 & 0+1+0 & 0+0+0 \\ -1+0+1 & 0+0+0 & -1+0+2 \end{pmatrix} = E_3.) \end{aligned}$$

Also ist die Lösung von

$$Ax = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

gegeben durch

$$x = A^{-1}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 1/2 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2+0-1 \\ 0+1/2+0 \\ -1+0+1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1/2 \\ 0 \end{pmatrix}$$

△

Satz 14.4. *Sind $A, B \in K^{n\times n}$ invertierbar, so ist auch AB invertierbar, und es gilt $(AB)^{-1} = B^{-1}A^{-1}$.*

Das ist die „Schuhe-Socken-Eigenschaft“: Ist $A =$ „Schuhe anziehen“ und $B =$ „Socken anziehen“ so ist die Umkehrung von $AB =$ „erst Socken, dann Schuhe anziehen“ die Operation „erst Schuhe, dann Socken ausziehen“ $= B^{-1}A^{-1}$

Beweis.
Man rechnet nach, dass

$$B^{-1}A^{-1}AB = B^{-1}B = E$$
$$ABB^{-1}A^{-1} = AA^{-1} = E$$

was die Behauptung zeigt. □

15 Elementarmatrizen und der Gauß-Algorithmus

Wir beschreiben den Gauß-Algorithmus noch einmal mit Hilfe der Matrix-Multiplikation.

Definition 15.1. Zu $i, j \in \{1, \ldots, n\}$ bezeichnen wir mit $\mathbb{1}_{ij} \in K^{n\times n}$ die Matrix, welche genau an der Stelle (i, j) eine 1 und sonst Nullen enthält.

Eine Matrix $M \in K^{n\times n}$ der Form

$$M = E_n + m\mathbb{1}_{ij} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & m & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \tag{I}$$

für eine $m \in K$ heißt *Elementarmatrix vom Typ I.*

Matrizen der Form

$$\begin{aligned} P &= E_n - \mathbb{1}_{ii} - \mathbb{1}_{jj} + \mathbb{1}_{ij} + \mathbb{1}_{ji} \\ &= \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \end{aligned} \tag{II}$$

heißen *Elementarmatrizen von Typ II* und

$$\begin{aligned} D &= E_n - \mathbb{1}_{ii} + a\mathbb{1}_{ii} \\ &= \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & a & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \end{aligned} \tag{III}$$

nennen wir *Elementarmatrizen vom Typ III.*

Die Multiplikation einer Elementarmatrix vom Typ I von links hat folgende Wirkung: Ist M wie in der Definition, $i > j$ (d.h. der nicht-null Eintrag liegt unterhalb der Diagonale) und $A \in K^{n\times k}$, so ist

$$MA = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & m & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{j1} & & a_{jk} \\ \vdots & & \vdots \\ a_{i1} & & a_{ik} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nk} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{j1} & & a_{jk} \\ \vdots & & \vdots \\ ma_{j1} + a_{i1} & & ma_{jk} + a_{ik} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nk} \end{pmatrix}, \qquad (*)$$

das heißt, die Multiplikation von links addiert das m-fache der j-ten Zeile zur i-ten Zeile hinzu.

Ebenso überzeugt man sich davon , dass die Multiplikation von links mit einer Elementarmatrix vom Typ II zwei Zeilen vertauscht und die Multiplikation von links mit Typ III eine Zeile skaliert. Die Multiplikation einer Matrix mit einer Elementarmatrix vom Typ 1 entspricht also genau dem entsprechenden Typ der elementaren Zeilenumformungen aus Satz 10.9.

Im Folgenden kodieren wir auch die weiteren Daten der Elementarmatizen in der Notation, d.h. wir setzen

$$\begin{aligned} M(i,j,m) &= E_n + m\mathbb{1}_{ij} \\ P(i,j) &= E_n - \mathbb{1}_{ii} - \mathbb{1}_{jj} + \mathbb{1}_{ij} + \mathbb{1}_{ji} \\ D(i,a) &= E_n - \mathbb{1}_{ii} + a\mathbb{1}_{ii}. \end{aligned}$$

Lemma 15.2. *Die Elementarmatrizen $M(i,j,m)$ und $P(i,j)$ sind für alle $i,j \in \{1,\dots,n\}$, $m \in K$ invertierbar und es gilt*

$$\begin{aligned} M(i,j,m)^{-1} &= M(i,j,-m) \\ P^{-1} &= P \end{aligned}$$

Ist $a \neq 0$ so ist auch $D(i,a)$ invertierbar mit

$$D(i,a)^{-1} = D(i,a^{-1}).$$

Beweis.
Dass $P(i,j)^{-1} = P(i,j)$ ist, sieht man daran, dass die Multiplikation mit $P(i,j)$ von links die $i$te und die $j$te Spalte vertauscht: Also ist $P(i,j)^{-1}P(i,j) = E_n$.

Die Rechnung (*) zeigt, dass die Multiplikation mit $M(i,j,-m)$

von links dass $-m$-fache der j-ten Zeile zur i-ten Zeile addiert. Also gilt
$$M(i,j,-m)M(i,j,m) = E_n.$$
Ebenso einfach sieht man die behauptete Form der Inversen von $D(i,a)$ für $a \neq 0$ ein. $\square$

Ebenso direkt aus der Rechnung (*) folgt, dass für $M(i,j,m_1)$ und $M(k,j,m_2)$ gilt

$$M(i,j,m_1)M(k,j,m_2) = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & m_1 & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & m_2 & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & m_1 & & 1 & & \\ & & m_2 & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

gilt.

Nach diesen Vorbetrachtungen schauen wir uns den Gauß-Algorithmus noch einmal an: Gegeben sei die quadratische Matrix

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

mit $a_{11} \neq 0$ und wir wollen die Einträge unterhalb von a_{11} eliminieren Die können wir erreichen, indem wir von Links mit der Matrix

$$M^1 = \begin{pmatrix} 1 & & & \\ -\frac{a_{21}}{a_{11}} & 1 & & \\ \vdots & & \ddots & \\ -\frac{a_{n1}}{a_{11}} & & & 1 \end{pmatrix}$$

multiplizieren. Das Ergebnis ist

Nach der Rechnung oben gilt $M^1 = M(n,1,-\frac{a_{n1}}{a_{11}}) \cdots M(2,1,-\frac{a_{21}}{a_{11}})$, d.h. M^1 ist ein Produkt von Elementarmatrizen vom Typ I.

$$\begin{aligned} A^1 &:= M^1 A \\ &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & \tilde{a}_{22} & \ldots & \tilde{a}_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & \tilde{a}_{n2} & \cdots & \tilde{a}_{nn} \end{pmatrix} \end{aligned}$$

Ist nun $\tilde a_{22}$ können wir analog mit

$$M^2 = \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & -\frac{\tilde a_{32}}{\tilde a_{22}} & 1 & & \\ & \vdots & & \ddots & \\ & -\frac{\tilde a_{n2}}{\tilde a_{22}} & & & 1 \end{pmatrix}$$

von links multiplizieren, und haben damit alle Einträge unterhalb von $\tilde a_{22}$ zusätzlich eliminiert. So fortgesetzt bekommen wir (immer unter den Annahme, dass die neuen Diagonalelemente $\tilde a_{kk}$ nicht Null sind)

$$M^{n-1}M^{n-2}\cdots M^2M^1A = \begin{pmatrix} a_{11} & \cdots & & a_{1n} \\ 0 & \tilde a_{22} & \cdots & \tilde a_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & & 0 & \tilde a_{nn} \end{pmatrix} = R.$$

Die Matrix R ist eine *obere Dreiecksmatrix*.

Analog zu unseren Überlegungen oben sieht man, dass die Matrizen M^k invertierbar sind mit inversen

$$(M^k)^{-1} = \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & \frac{\tilde a_{k+1,k}}{\tilde a_{kk}} & 1 & & \\ & \vdots & & \ddots & \\ & \frac{\tilde a_{nk}}{\tilde a_{kk}} & & & 1 \end{pmatrix}.$$

Wir können also A schreiben als

$$A = \underbrace{(M^1)^{-1}(M^2)^{-1}\cdots(M^{n-1})^{-1}}_{=L} R = LR.$$

Die Matrix L ist, wie man sich überzeugen kann, eine *untere Dreiecksmatrix*. Das heißt, wir haben A faktorisiert als Produkt von unterer und oberer Dreiecksmatrix:

$$A = LR = (\llcorner)(\urcorner)$$

Das eben skizzierte Verfahren führt aus die sogenannt *LR-Zerlegung* einer Matrix (welche man auch für rechteckige Matrizen A anwenden kann, und welche im Allgemeinen noch Vertauschungen von Zeilen benötigt, um wirklich $\tilde a_{kk} \neq 0$ zu garantieren).

Liegt eine LR-Zerlegung einer Matrix vor, so lassen sich lineare Gleichungssysteme einfach Lösen: Ist $A = LR$ mit einer unteren Dreiecksmatrix L und einer oberen Dreiecksmatrix R, so

$$Ax = b \iff LRx = b \iff \begin{cases} Ly = b \\ Rx = y. \end{cases}$$

Da L und R Dreiecksmatrizen sind, lass sich die beiden Systeme rechts einfach durch *Vorwärts- bzw. Rückwärtseinsetzen* lösen.

Mit einem ähnlichen Verfahren lassen sich auch inverse Matrizen berechnen. Es gilt nämlich:

Satz 15.3. *Lässt sich $A \in K^{n\times n}$ durch Anwendung von Elementarmatrizen von Typ I, II und III von links in eine Einheitsmatrix umformen, so ist A invertierbar. Wendet man die gleichen Elementarmatrizen von links auf die Einheitsmatrix an, so erhält man die Inverse A^{-1}.*

Beweis.
Da das Produkt von Matrizen genau dann invertierbar ist, wenn alle Faktoren invertierbar sind, folgt aus

$$M^k \cdots M^1 A = E_n,$$

(wobei M^j Elementarmatrizen vom Typ I, II und II sind), dass alle M^j invertierbar sind. Dann ist

$$M^k \cdots M^1 E_n = M^k \cdots M^1 A A^{-1} = E_n A^{-1} = A^{-1},$$

was zu zeigen war. □

Die Methode aus diesem Satz nennt sich auch Gauß-Jordan-Methode zur Berechnung von inverse Matrizen.

Beispiel.
Wir betrachten die reelle Matrix

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

Das Schema zur Inversion der Matrix A schreiben wir die Einheitsmatrix rechts neben A und wenden auf die erweiterte Matrix elementare Zeilenumformungen an:

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

Abziehen der ersten Zeile von der zweiten ergibt

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

Addition der zweite auf die erste Zeile liefert

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

Multipizieren wir nun die zweite Zeile mit -1 liefert

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

Also ist die inverse Matrix

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

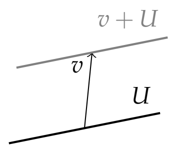
(und man kann leicht nachrechnen, dass $A^{-1}A = E_2 = AA^{-1}$).
$\triangle$

16 Faktorräume

Definition 16.1. Ist $U \subset V$ ein Unterraum und $v \in V$, dann ist

$$v + U := \{v + u \mid u \in U\}$$

die *Nebenklasse von U durch v* und v ein *Repräsentant* der Nebenklasse.



Satz 16.2. *Es gilt, dass $v + U$ genau dann ein Unterraum ist, wenn $v \in U$ gilt.*

Beweis.
Ist $v \in U$, so gilt

$$v + U = \{v + u \mid u \in U\} = U$$

dann heißt, $v + U$ selbst ist gleich dem Unterraum U. Ist andererseits $v + U$ ein Unterraum, so muss $0 \in v + U$ gelten und das heißt, es gibt ein $u \in U$ mit $v + u = 0$. Es folgt $v = -u \in U$. □

Korollar 16.3. *Für zwei Vektoren v, v' gilt $v + U = v' + U$ genau dann, wenn $v - v' \in U$ gilt.*

Es gilt $v + U = v' + U$ genau dann, wenn $v - v' + U = U$. Die Behauptung folgt also aus vorigem Satz.

Lemma 16.4. *Für einen Unterraum $U \subset V$ ist*

$$x \sim y \iff x - y \in U$$

eine Äquivalenzrelation, die zugehörigen Äquivalenzklassen sind genau die Nebenklassen von U.

Beweis.
Wegen $0 \in U$ gilt $x \sim x$. Und da aus $x - y \in U$ auch $y - x = -(x - y) \in U$ folgt, ist $\sim$ auch symmetrisch. Die Transitivität folgt schließlich daraus das mit $x \sim y$ und $y \sim z$ (also $x - y \in U$ und $y - z \in U$) folgt, dass $x - z = x - y + y - z \in U$, also $x \sim z$. □

Korollar 16.5. *Der Vektorraum V ist die disjunkte Vereinigung der Nebenklassen von U.*

Wir können die Menge de Nebenklassen selbst zu einem Vektorraum machen. Dazu führen wir zuerst folgende Schreibweise ein: Ist $U \subset V$ ein Unterraum und im Folgenden fixiert, dann schreiben wir für die Nebenklasse eines $x \in V$

$$[x] = x + U.$$

Jetzt statten wir Nebenklassen mit folgenden Rechenregeln aus:

Es gilt also $[x] = [y]$ falls x und y die gleiche Nebenklassen erzeugen, also genau dann, wenn $x - y \in U$.

Für $x, y \in V$ und $k \in K$ definieren wir

$$[x] + [y] = [x + y], \quad \text{und} \quad k[x] = [kx].$$

Bevor wir zeigen, dass die Menge der Nebenklassen damit zum Vektorraum wird, müssen wir erst zeigen, dass Addition und Skalierung von Nebenklassen *wohldefiniert* sind.

„Wohldefiniert" bedeutet in diesem Fall, dass die Definition unabhängig von den gewählten Repräsentanten ist.

Beweis (Wohldefiniertheit der Addition und Skalierung für Nebenklassen). Es gelte $x \sim x'$ und $y \sim y'$. Es ist zu zeigen, dass $[x + y] = [x' + y']$ gilt. Es gilt $x - x' \in U$ und $y - y' \in U$, und daher ist auch $(x + y) - (x' + y') = x - x' + y - y' \in U$. Dies zeigt $x + y \sim x' + y'$, also $[x + y] = [x' + y']$ wie gefordert.

Für die Skalierung sei wieder $x \sim x'$, also $x - x' \in U$. Dann gilt auch $k(x - x') \in U$, also $kx - kx' \in U$, und das heißt $kx \sim kx'$, also $[kx] = [kx']$. □

Satz 16.6. *(und Definition) Ist V ein Vektorraum und U ein Unterraum, dann heißt die Menge der Nebenklassen*

$$V/U := \{[v] \mid v \in V\}$$

Faktorraum oder Quotientenraum und ist, versehen mit oben definierter Addition und Skalierung, ein Vektorraum.

Beweis.
Die Axiome lassen sich auf dem naheliegenden Weg nachweisen. Zum Beispiel ist die 0 in V/U gleich $[0] = 0 + U$, denn $[0] + [v] = [0 + v] = [v]$.

Die weiteren Axiome folgen ähnlich einfach, zum Beispiel gilt

$$\begin{aligned} [v] + [w] &= [v + w] = [w + v] = [w] + [v] \\ ([u] + [v]) + [w] &= [u + v] + [w] = [(u + v) + w] \\ &= [u + (v + w)] = [u] + [v + w] \\ &= [u] + ([v] + [w]) \end{aligned}$$

□

Satz 16.7. *Es gilt*

$$\dim(V/U) = \dim(V) - \dim(U).$$

Intuitiv: In V/U „fehlen die Richtungen, die in U liegen".

Beweis.
Es sei $s = \dim(U)$ und $B_0 = \{v_1, \ldots, v_s\}$ eine Basis von U. Wir ergänzen diese zu einer Basis $B = \{v_1, \ldots v_s, v_{s+1}, \ldots, v_n\}$ von V.

Wir zeigen jetzt, dass $B^* = \{[v_{s+1}], \ldots, [v_n]\}$ eine Basis von V/U ist, denn damit folgt

$$\dim(V/U) = n - s = \dim(V) - \dim(U).$$

Beweis dieser Behauptung: Einerseits ist B^* linear unabhängig, denn gilt

$$k_{s+1}[v_{s+1}] + \cdots + k_n[v_n] = 0$$

so folgt

$$[k_{s+1}v_{s+1} + \cdots + k_n v_n] = 0$$

und das bedeutet $w = k_{s+1}v_{s+1} + \cdots + k_n v_n \in U$. Also ist w eine Linearkombination der $v_1, \ldots, v_s$ und es folgt $k_{s+1} = \cdots k_n = 0$.

Schließlich erzeugt B^* ganz V/U. Dazu sei $[v] \in V/U$. Wir können v in der Basis B darstellen

$$v = \sum_{i=1}^{n} k_i v_i$$

und da $\sum_{i=1}^{s} k_i v_i \in U$ gilt, folgt

$$[v] = [k_{s+1}v_{s+1} + \cdots k_n v_n]$$

wie gewünscht. $\square$

17 Der Homomorphiesatz

Der Homomorphiesatz stellt eine Verbindung zwischen Kern und Bild einer linearen Abbildung her, die über den Dimensionssatz (Satz 10.8) hinausgeht.

Ziel ist es, mit Hilfe des Faktorraumes lineare Abbildungen „bijektiv machen“, indem wir das Bild einschränken, und „den Kern heraus faktorisieren“.

Lemma 17.1. *Eine lineare Abbildung $f : V \to W$ ist injektiv, genau dann, wenn* $\mathrm{Kern}(f) = \{0\}$.

Beweis.
Sei $\mathrm{Kern}(f) = \{0\}$ und gelte $f(v) = f(v')$. Dann folgt

$$0 = f(v) - f(v') = f(v - v'),$$

also $v - v' \in \mathrm{Kern}(f)$. Es folgt also $v = v'$ und damit die Injektivität von f.

Sei nun f injektiv. Wegen der Linearität von f gilt $f(0) = 0$ und wegen der Injektivität gilt für alle $v \neq 0$ auch $f(v) \neq 0$, d.h. aber gerade, dass der Kern von f nur die Null enthält. □

Satz 17.2 (Homomorphiesatz)**.** *Zu einer linearen Abbildung $f : V \to W$ definieren wir*

$$\tilde{f} : V / \mathrm{Kern}(f) \to \mathrm{Bild}(f)$$

durch

$$\tilde{f}([v]) = f(v).$$

Dann ist $\tilde{f}$ bijektiv; mit anderen Worten: $\tilde{f}$ ist ein Isomorphismus von $V / \mathrm{Kern}(f)$ nach $\mathrm{Bild}(f)$.

Wir nennen zwei Vektorräume *isomorph*, wenn es einen Isomorphismus zwischen ihnen gibt. Wir können hier also folgern, dass $V / \mathrm{Kern}(f)$ isomorph zu $\mathrm{Bild}(f)$ ist.

Beweis.
Zuerst zeigen wir, dass durch $\tilde{f}$ überhaupt eine Abbildung definiert wird, also, dass für zwei $v' \in [v]$ mit $v' \neq v$ gilt $\tilde{f}([v']) = f(v)$. Ist $v' \in [v]$, dann gilt $v' = v + w$ mit $w \in \mathrm{Kern}(f)$, also folgt $f([v']) = f(v') = f(v + w) = f(v) + f(w) = f(v)$, wie gefordert.

Für die Linearität von $\tilde{f}$ bemerken wir

$$\tilde{f}([v + \alpha v']) = f(v + \alpha v') = f(v) + \alpha f(v') = f([v]) + \alpha f([v']).$$

Dass $\tilde{f}$ surjektiv ist, ist klar, da wir den Bildbereich gerade auf $\mathrm{Bild}(f)$ eingeschränkt haben und nach Definition $\mathrm{Bild}(\tilde{f}) = \mathrm{Bild}(f)$ gilt.

Für die Injektivität nehmen wir $v_1, v_2 \in V$ mit $[v_1] \neq [v_2]$. Dass die Äquivalenzklassen nicht gleich sind, bedeutet $v_1 - v_2 \notin \mathrm{Kern}(f)$ und das heißt $f(v_1 - v_2) \neq 0$. Es folgt $f(v_1) \neq f(v_2)$ und das bedeutet wiederum $\tilde{f}([v_1]) \neq \tilde{f}([v_2])$. □

Wir erhalten nun die Dimensionsformel aus Korollar 12.8 noch einmal als Korollar:

Korollar 17.3. *Ist $f : V \to W$ linear, so gilt*

$$\dim(\mathrm{Kern}(f)) + \dim(\mathrm{Bild}(f)) = \dim(V).$$

Es gilt nach dem vorigen Satz und Satz 16.7 $\dim(\mathrm{Bild}(f)) = \dim(V/\,\mathrm{Kern}(f)) = \dim(V) - \dim(\mathrm{Kern}(f))$.

Korollar 17.4. *Ist $f : V \to W$ linear und gilt* $\dim(V) = \dim(W) = n$ *, so ist äquivalent:*

Ist die Dimension von V und W nicht endlich, so gilt der Satz nicht, d.h. es gibt Vektorräume, die nicht endlich erzeugt sind und zwischen denen es injektive Abbildungen gibt, die nicht surjektiv sind (und umgekehrt).

i) f injektiv

ii) f surjektiv

iii) f bijektiv.

Sei $n = \dim(V) = \dim(W)$. Es gilt

$$\begin{aligned} f \text{ injektiv} &\iff \mathrm{Kern}(f) = \{0\} \\ &\iff \dim(\mathrm{Kern}(f)) = 0 \\ &\iff \dim(\mathrm{Bild}(f)) = n \\ &\iff f \text{ surjektiv.} \end{aligned}$$

Die Äquivalenz zur Bijektivität ist dann klar.

18 Der Dualraum

Für zwei K-Vektorräume V und W bezeichnete $\mathrm{Hom}(V, W)$ die Menge der linearen Abbildungen von V nach W. Da auch K selbst ein K-Vektorraum ist, können wir auch $\mathrm{Hom}(V, K)$ betrachten, also die Menge aller lineare Abbildung von V in den Grundkörper. Wie auch im allgemeinen Fall ist dies wieder ein Vektorraum, der jedoch einen speziellen Namen hat:

Definition 18.1. Ist V ein K-Vektorraum, so ist

$$V^* = \mathrm{Hom}(V, K)$$

der *Dualraum* von V. Die Elemente aus V^* heißen *Linearformen* auf V.

Lemma 18.2. *Ist* $\dim(V) = n$*, $f \in V^*$ und ist $f \neq 0$, so gilt* $\dim(\mathrm{Kern}(f)) = n - 1$.

Beweis.
Es gilt $\dim(\operatorname{Bild}(f)) = 1$. Nach der Dimensionsformel ist $\dim(\operatorname{Kern}(f)) = \dim(V) - \dim(\operatorname{Bild}(f)) = n - 1$. $\square$

Haben wir eine Basis von V, so lässt sich dazu eine duale Basis von V^* konstruieren:

Satz 18.3. *(und Definition) Es sei V ein Vektorraum mit Basis $B = \{v_1, \dots, v_n\}$, dann gibt es eine Basis $B^* = \{v_1^*, \dots, v_n^*\}$ des Dualraumes mit*

$$v_i^*(v_j) = \delta_{ij} := \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}.$$

Das hier definierte Symbol δ_{ij} heißt *Kronecker-Delta*.

Beweis.
Fixieren wir ein $i \in \{1, \dots, n\}$. Durch die Forderungen $v_i^*(v_j) = \delta_{ij}$ für $j = 1, \dots, n$ ist v_i^* als lineare Abbildung eindeutig bestimmt (erinnere, dass lineare Abbildungen durch die Bilder auf einer Basis bestimmt sind).

Es bleibt zu zeigen, dass B^* wirklich eine Basis ist.

Sei $k_1 v_1^* + \cdots + k_n v_n^* = 0$ (soll meinen, dass diese Linearkombination die Nullabbildung ist). Dann können wir jeden Basisvektor v_i einsetzen und bekommen immer Null, d.h. für jedes i gilt

$$0 = \Big(k_1 v_1^* + \cdots + k_n v_n^*\Big)(v_i) = k_1 v_1^*(v_i) + \cdots + k_n v_n^*(v_i) = k_i,$$

also ist $k_i = 0$ für alle i. Dies zeigt, dass B^* linear unabhängig ist.

Sei nun $f \in V^*$. Dann ist f durch die Werte $f(v_i) = a_i$ bestimmt. Die Linearform

$$v^* = a_1 v_1^* + \dots a_n v_n^*$$

hat aber auf den v_i die gleichen Werte, also gilt $v^* = f$. Dies zeigt, dass B^* ein Erzeugendensystem ist. $\square$

Korollar 18.4. *Der Dualraum V^* eines n-dimensionalen Vektorraums V ist ebenfalls n-dimensional.*

Beispiel.
Ist $V = K^n$, so ist $V^* = \operatorname{Hom}(K^n, K)$. Der Raum $\operatorname{Hom}(K^n, K)$ der linearen Abbildungen von K^n nach N ist isomorph zum Raum $K^{1\times n}$ der $1 \times n$ Matrizen, also zum Raum der Spaltenvektoren. $\triangle$

Beispiel.
Es sei $V = \mathbb{R}^2$ und wie betrachte die Basis $B = \{v_1, v_2\}$ mit $v_1 = \left(\begin{smallmatrix} 1 \\ 1, \end{smallmatrix}\right)\}, v_2 = \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$. Um die Duale Basis zu bestimmen setzten wir $v_1^* = \left(\begin{smallmatrix} x_1 \\ x_2, \end{smallmatrix}\right)\}$, $v_2 = \left(\begin{smallmatrix} y_1 \\ y_2 \end{smallmatrix}\right)$ und stellen die vier Gleichungen

$v_i^*(v_j) = \delta_{ij}$ auf:

$$\begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1, \qquad \begin{pmatrix} y_1 & y_2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 0,$$
$$\begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 0, \qquad \begin{pmatrix} y_1 & y_2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = 1.$$

Wir schreiben diese als

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Wir bekommen als duale Basis

$$v_1^* = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \qquad v_2^* = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

Wie sehen folgenden Zusammenhang zwischen der dualen Basis und Matrix-Inversion: Schreiben wir die Basisvektoren aus B spaltenweise in

$$M = \begin{pmatrix} v_1 & v_2 \end{pmatrix}$$

und die Basisvektoren der dualen Basis zeilenweise in

$$M^* = \begin{pmatrix} v_1^* \\ v_2^* \end{pmatrix}$$

so muss gelten

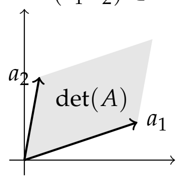
$$M^* M = \begin{pmatrix} v_1^* v_1 & v_1^* v_2 \\ v_2^* v_1 & v_2^* v_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2.$$

Das heißt, die Vektoren der dualen Basis sind die Zeilen der inversen Matrix von M. △

19 Determinaten

Die Determinante ist eine Abbildung welche jeder quadratischen Matrix in $K^{n\times n}$ eine Element aus K zuordnet. Im Fall von $\mathbb{R}^{n\times n}$ hat die Determinante eine geometrische Bedeutung, mit der wir beginnen:

$A = (a_1\ a_2) \in \mathbb{R}^{2\times 2}$,



> Die Determinante $\det(A)$ einer Matrix $A \in \mathbb{R}^{n\times n}$ mit den Spalten $a_1, \dots, a_n$ ist das n-dimensionale orientierte Volumen des n-dimensionalen Parallelepipeds der durch die Vektoren $a_1, \dots, a_n$ bestimmt wird.

Wir klären einige Begriffe (teilweise eher intuitiv, als durch exakte Defintionen):

1. Das Parallelepiped, das durch die Vektoren $a_1, \dots, a_n$ bestimmt wird, ist die Menge $\{\sum_{i=1}^n k_i a_i \mid 0 \leq k_i \leq 1\}$.

2. Das n-dimensionale Volumen entspricht der Verallgemeinerung dessen, was man durch „Grundfläche mal Höhe" erhält, also haben Mengen der Form $[0, k_1] \times \cdots \times [0, k_n]$ das n-dimensionale Volumen $k_1 \cdots k_n$.

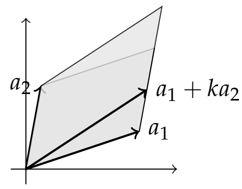
3. Der Einheitswürfel $W = \{x \in \mathbb{R}^n \mid 0 \leq x_i \leq 1\}$ ist das Parallelepiped zur Einheitsmatrix und sein orientiertes Volumen ist 1.

4. Wird ein einzelner Vektor mit einem $k \in \mathbb{R}$ skaliert, so wird das entsprechende n-dimensionale orientierte Volumen ebenfalls mit k multipliziert, insbesondere ändert das orientierte Volumen sein Vorzeichen bei Multiplikation eines Vektors mit -1.

5. Wird die Reihenfolge von zwei Vektoren vertauscht, so ändert das orientierte Volumen nur sein Vorzeichen.

6. Sind die Vektoren $a_1, \dots, a_n$ linear abhängig, so ist das n-dimensionale orientierte Volumen Null.

7. Ein *Scheerung* einens Parallelepipeds zur Matrix A entspricht der Addition des Vielfachen eines Vektors zu einem anderen. Dabei ändert sich das n-dimensionale orientierte Volumen nicht.

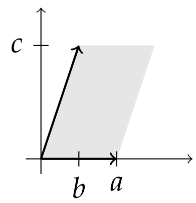


Ein Formel für das n-dimensionale orientierte Volumen geben wir vorerst nicht an. Erstaunlicherweise können wir aus den obigen Forderungen eine Formel ableiten (das Ergebnis ist die sogenannte Leibniz-Formel aus Satz 20.6).

Im Falle von Dreiecksmatrizen $M \in \mathbb{R}^{n\times n}$ ist die Berechnung von Determinaten einfach. Im 2×2-Fall gilt (wegen „Fläche =

Grundseite mal Höhe“).

$$\det\left(\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}\right) = ac.$$



Entsprechend gilt im $n \times n$-Fall (wegen „Volumen = Grundfläche mal Höhe“):

$$\det\left(\begin{pmatrix} a_{11} & \cdots & * \\ 0 & \ddots & \vdots \\ 0 & 0 & a_{nn} \end{pmatrix}\right) = a_{11} \cdots a_{nn}.$$

Wir arbeiten von nun an wieder in einem allgemeinen Körper K und leiten aus den Forderungen für die reelle Determinante folgende Axiome für eine abstrakte Determinentenabbilung ab:

Definition 19.1. Eine Abbildung $\det : K^{n\times n} \to K$ heißt *Determinantenfunktion*, wenn sie die folgenden drei Eigenschaften hat:

(D1) Die Abbildung ist linear in jeder Spalte, d.h. für jedes $i \in \{1,\ldots,n\}$ gilt

$$\begin{aligned}\det(\begin{pmatrix} a_1 & \cdots & a_i + a_i' & \cdots a_n \end{pmatrix}) = \det(\begin{pmatrix} a_1 & \cdots & a_i & \cdots & a_n \end{pmatrix}) \\ + \det(\begin{pmatrix} a_1 & \cdots & a_i' & \cdots & a_n \end{pmatrix})\end{aligned}$$

und für jedes $k \in K$ gilt

$$\det(\begin{pmatrix} a_1 & \cdots & ka_i & \cdots & a_n \end{pmatrix}) = k\det(\begin{pmatrix} a_1 & \cdots & a_i & \cdots & a_n \end{pmatrix})$$

Das ist eine Konsequenz aus den Punkten 4., 5. und 7. der obigen Liste.

(D2) Ist $\operatorname{Rang}(A) < n$, so gilt $\det(A) = 0$.

Das ist Punkt 6 aus obiger Liste.

(D3) $\det(E_n) = 1$

Das ist Punkt 3. aus obiger Liste.

Determinantenfunktionen verhalten sich einfach bezüglich elementarer *Spalten*umformungen (vgl. Satz 10.9) d.h. bei Multiplikation von *rechts* mit Elementarmatrizen:

Satz 19.2. *Es sei $A \in K^{n\times n}$ und $\det : K^{n\times n} \to K$ eine Determinantenfunktion. Dann gilt*

(i) Ist $M(i,j,m)$ eine Elementarmatrix vom Typ I, dann gilt $\det(AM(i,j,m)) = \det(A)$.

Die Typen der Elementarmatrizen wurden in Abschnitt 15 eingeführt.

(ii) Ist $P(i,j)$ eine Elementarmatrix vom Typ II, so gilt $\det(AP(i,j)) = -\det(A)$.

(iii) Ist $D(i,a)$ eine Elementarmatrix vom Typ III, so gilt $\det(AD(i,a)) = a\det(A)$.

Beweis.

(i) Hat A die Spalten $a_1, \dots, a_n$ so gilt

$$\begin{aligned} AM(i,j,m) &= \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & m & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \\ &= \begin{pmatrix} a_1 & \cdots & a_j + m a_i & \cdots & a_n \end{pmatrix}, \end{aligned}$$

d.h. $AM(i,j,m)$ entsteht aus A, indem das m-fach der i-ten Zeile auf die j-te Zeile addiert wird. Bezeichnen wir mit $\tilde{A}$ die Matrix, die aus A entsteht, wenn wir die j-te Zeile von A durch ma_i *ersetzen*, so folgt

$$\det(AM(i,j,m)) = \det(A) + \det(\tilde{A}) = \det(A) + 0 = \det(A),$$

wobei wir in der ersten Gleichung (D1) benutzt haben, und der zweiten Gleichung (D2) (da $\tilde{A}$ zwei gleiche Spalten hat, ist $\operatorname{Rang}(\tilde{A}) < n$).

(ii) Die Multiplikation von A von rechts mit $P(i,j)$ vertauscht die i-te und die j-te Spalte. Es gilt also

$$\begin{aligned} 0 &\overset{(D2)}{=} \det(\begin{pmatrix} a_1 & \cdots & a_i + a_j & \cdots & a_j + a_i & \dots & a_n \end{pmatrix}) \\ &\overset{(D1)}{=} \det(\begin{pmatrix} a_1 & \cdots & a_i & \cdots & a_j + a_i & \dots & a_n \end{pmatrix}) \\ &\quad + \det(\begin{pmatrix} a_1 & \cdots & a_j & \cdots & a_j + a_i & \dots & a_n \end{pmatrix}) \\ &\overset{(i)}{=} \det(\begin{pmatrix} a_1 & \cdots & a_i & \cdots & a_j & \dots & a_n \end{pmatrix}) \\ &\quad + \det(\begin{pmatrix} a_1 & \cdots & a_j & \cdots & a_i & \dots & a_n \end{pmatrix}) \\ &= \det(A) + \det(AP(i,j)), \end{aligned}$$

woraus die Behauptung folgt.

(iii) Da die Multiplikation von rechts mit $D(i,a)$ genau die i-te Spalte skaliert, folgt die Behauptung direkt aus der Linearität von det in jeder Spalte.

□

Satz 19.3. *Für* $A, B \in K^{n \times n}$ *gilt* $\det(AB) = \det(A)\det(B)$.

Beweis.
Gelte zunächst $\operatorname{Rang}(B) < n$. Dann ist $\operatorname{Rang}(AB) = \dim(\operatorname{Bild}(AB)) \leq \dim(\operatorname{Bild}(B)) = \operatorname{Rang}(B) < n$, also ist $\det(B) = 0 = \det(AB)$.

Sei nun $\operatorname{Rang}(B) = n$. Dann können wir A mit alleiniger Hilfe von elementaren Spaltenumformungen vom Typ I in eine Matrix der Form

$$D(n,b) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & 1 & 0 \\ 0 & \cdots & 0 & b \end{pmatrix}$$

(vgl. große Übung). Mit anderen Worten: Es gilt $BM_1 \cdots M_K = D(n,b)$ mit Elementarmatrizen M_i vom Typ I: Es gilt also $\det(B) = \det(BM_1 \cdots M_K) = \det(D(n,b)) = b$. Da aber $D(n,b)$ eine Spaltenskalierung ist, folgt nach (D3)

$$\det(AB) = \det(AD(n,b)) \overset{(D3)}{=} b\det(A) = \det(B)\det(A)$$

wie behauptet. □

Aus den Invarianzeigenschaften folgt, dass es nur eine Determinantenfunktion gibt:

Satz 19.4. *Für jedes $n = 1, 2, \ldots$ gibt es höchstens eine Determinantenfunktion.*

Beweis.
Es seien det und det′ zwei Funktionen $K^{n\times n} \to K$, welche beide (D1)–(D3) erfüllen.

Aus (D2) folgt, dass $\det(M) = \det'(M)$ für alle M mit $\operatorname{Rang}(M) < n$ gilt.

Sei also $\operatorname{Rang}(M) = n$. Dann lässt sich M (analog zum vorigen Beweis) durch elementare Spaltentransformationen vom Typ I auf die Form $D(n,b)$ bringen. Da det und det′ beide die Invarianzen aus Satz 19.2 erfüllen, folgt $\det(M) = \det(D(n,b)) = b$ und $\det'(M) = \det'(D(n,b)) = b$, also folgt $\det(M) = \det'(M)$ auch für alle Matrizen mit vollrang. □

Es folgt:

Korollar 19.5. *Für $A \in K^{n\times n}$ sind äquivalent:*

(i) *A ist invertierbar,*

(ii) $\operatorname{Rang}(A) = n$,

(iii) $\det(A) \neq 0$.

Die Äquivanlenz von (i) und (ii) haben wir in Satz 14.3 gesehen. Die Implikation (iii) $\implies$ (ii) ist gerade (D2). Die Implikation (ii) $\implies$ (iii) haben wir im Beweis von Satz 19.3 mit bewiesen.

Beispiel.

Es sei

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad d \neq 0.$$

Wir bringen A auf obere Dreiecksform:

$$AM = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{c}{d} & 1 \end{pmatrix} = \begin{pmatrix} a - \frac{bc}{d} & b \\ 0 & d \end{pmatrix}.$$

Wegen Satz 19.2 (1) folgt

$$\det(A) = \det(AM) = \det(\begin{pmatrix} a - \frac{bc}{d} & b \\ 0 & d \end{pmatrix}) = \left(a - \frac{bc}{d}\right) d = ad - bc.$$

Im Falle von $d = 0$ aber $a \neq 0$ bekommt man durch Tauschen der Spalten wegen Satz 19.2 (ii)

$$\det(\begin{pmatrix} a & b \\ c & 0 \end{pmatrix}) = -\det(\begin{pmatrix} b & a \\ 0 & c \end{pmatrix}) = -bc.$$

Insgesamt gilt in jedem Fall

$$\det(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) = ad - bc.$$

$\triangle$

20 Permutationen und die Leibniz-Formel

Der vorige Abschnitt lieferte einen Algorithmus zur Berechnung von Determinanten. Für theoretische Zwecke wäre eine Formel für die Determinante hilfreich. Hierfür benötigt man Permutationen als Hilfsmittel:

Der Algorithmus geht wie folgt: Anwendung von Spaltentransformationen vom Typ I zur Umformung auf Dreiecksform, dann Berechnen des Produktes der Diagonalelemente.

Definition 20.1. Eine *Permutation* einer Menge X ist eine bijektive Abbildung $\pi : X \to X$.

Definition 20.2 (und Satz)**.** Die Menge aller Permutationen der Menge $\{1, \ldots, n\}$ bezeichnen wir mit $\mathbb{S}_n$. Die Menge $\mathbb{S}_n$ bildet, mit der Hintereinanderausführung als Verknüpfung eine Gruppe, welche *symmetrische Gruppe* genannt wird.

Das neutrale Element ist die identische Permutation, die wir mir ϵ bezeichnen.

Eine Gruppe ist eine Menge mit eine assoziativen Verknüpfung auf der ein eine neutrales und inverse Elemente gibt.

Die Hintereinanderausführung von zwei Permutationen wird auch *Produkt* der Permutationen genannt.

Ist $n > 1$, so setzen wir

$$\mathbb{S}_n(k) := \{\pi \in \mathbb{S}_n \mid \pi(k) = k\},$$

also die Menge der Permutationen, die k fixiert lassen. Die Menge $\mathbb{S}_n(n)$ sind also die Permutationen, die das letzte Element fixiert lassen, d.h. sie permutieren nur die Elemente $1, \ldots, n-1$. Wir können also $\mathbb{S}_n(n)$ mit $\mathbb{S}_{n-1}$ identifizieren.

Definition 20.3. Eine Permutation heißt *Transposition*, wenn Sie genau zwei Elemente vertauscht.

Für jede Transposition τ gilt $\tau \circ \tau = \epsilon$.

Oder anders gesagt: $\tau^{-1} = \tau$.

Satz 20.4. *Die symmetrische Gruppe $\mathbb{S}_n$ hat $n!$ Elemente und jede Permutation aus $\mathbb{S}_n$ lässt sich als Produkt von Transpositionen darstellen*

Beweis.
Für $n = 1$ und $n = 2$ ist die Aussage klar.

Für $n = 1$ enthält $\mathbb{S}_1$ nur ϵ und $n = 2$ auch noch eine Transposition.

Für den Induktionsschritt sei also $n \geq 3$ und wir bezeichnen mit $\tau^{(i)} \in \mathbb{S}_n$ die Permutation, welche i und n vertauscht (also $\tau^{(n)} = \epsilon$). Damit können wir jedes Element $\pi \in \mathbb{S}_n$ wie folgt faktorisieren: Setze $i = \pi(n)$ und schreibe

$$\pi = \tau^{(i)} \underbrace{\tau^{(i)}\pi}_{=\rho} = \tau^{(i)}\rho, \quad \text{mit} \quad \rho \in \mathbb{S}_n(n).$$

Da wir $\mathbb{S}_n(n)$ mit $\mathbb{S}_{n-1}$ identifizieren können, und $\rho \in \mathbb{S}_{n-1}$ nach Induktionsannahme ein Produkt von Transpositionen ist, ist auch π ein Produkt von Transpositionen. Ebenso folgt die Behauptung zur Anzahl der Elemente, da wir die Elemente von $\mathbb{S}_n$ gerade als $\tau^{(i)}\rho$ mit $i \in \{1, \ldots, n\}$ und $\rho \in \mathbb{S}_{n-1}$ schreiben können. Dies zeigt $|\mathbb{S}_n| = n|\mathbb{S}_{n-1}|$ was rekursiv die Behauptung zeigt. □

Zu jeder Permutation π gibt es eine Permutationsmatrix P_π definiert durch

$$P_\pi = \begin{pmatrix} e_{\pi(1)} & \cdots e_{\pi(n)} \end{pmatrix}.$$

Es gilt

$$P_\pi e_i = e_{\pi(i)},$$

und daher auch für zwei Permutationen π, σ

$$P_\pi P_\sigma = P_{\pi\sigma}.$$

Es gilt nämlich $P_\pi P_\sigma e_i = P_\pi e_{\sigma(i)} = e_{\pi(\sigma(i))}$.

Weiterhin ist für $A = \begin{pmatrix} a_1 & \cdots a_n \end{pmatrix} \in K^{n\times n}$ und $\pi \in S_n$ auch

$$\begin{aligned} AP_\pi &= A\begin{pmatrix} e_{\pi(1)} & \cdots & e_{\pi(n)} \end{pmatrix} = \begin{pmatrix} Ae_{\pi(1)} & \cdots & Ae_{\pi(n)} \end{pmatrix} \\ &= \begin{pmatrix} a_{\pi(1)} & \cdots & a_{\pi(n)} \end{pmatrix}, \end{aligned}$$

also ist AP_π die Matrix, die durch Permutation der Spalten von A gemäß π hervorgeht.

Definition 20.5. Für eine Permutation $\pi \in S_n$ definieren wir das *Signum* der Permutation durch

$$\operatorname{sig}(\pi) = \det(P_\pi).$$

Ist τ eine Transposition so entsteht P_τ aus der Einheitsmatrix E_n durch vertauschen von zwei Spalten. Nach Satz 19.2 (ii) gilt also $\det(P_\tau) = -1$.

Es folgt: Lässt sich π als Produkt von r Transpositionen $\tau^1, \ldots, \tau^r$ schreiben, so gilt $\det(P_\pi) = \det(P_{\tau_1} \cdots P_{\tau_r}) = (-1)^r$.

Satz 20.6 (Leibniz-Formel für die Determinante)**.** *Für $A \in K^{n\times n}$ gilt*

$$\det(A) = \sum_{\pi\in S} \operatorname{sig}(\pi) \prod_{i=1}^n a_{\pi(i),i}$$

Das Produktzeichen $\prod_{i=1}^n$ ist analog zum Summenzeichen $\sum_{i=1}^n$ zu lesen, nur eben als Produkt.

Beweis.
Wir leiten die Leibniz-Formel aus den Eigenschaften her. Die j-te Spalte von $A = (a_{ij})$ lässt sich schreiben als

$$a_j = \sum_{i=1}^n a_{ij} e_i.$$

Daher gilt

$$\det(A) = \det(\begin{pmatrix} \sum_{i=1}^n a_{i1} e_i & \cdots & \sum_{i=1}^n a_{in} e_i \end{pmatrix}.$$

Wir nutzen nun (D1) (Linearität in jeder Spalte) für alle diese Summen aus. Da wir in jeder Spalte einen eignen Summationsindex haben, bekommen wir n Summen, je von 1 bis n, nämlich

$$\det(A) = \sum_{i_1=1}^n \cdots \sum_{i_n=1}^n a_{i_1,1} a_{i_2,2} \cdots a_{i_n,n} \det(\begin{pmatrix} e_{i_1} & \cdots & e_{i_n} \end{pmatrix}).$$

Die Determinanten $\det(\begin{pmatrix} e_{i_1} & \cdots & e_{i_n} \end{pmatrix})$ sind immer dann Null, wenn zwei der Summationsindizes $i_1, \dots, i_n$ gleich sind, also genau dann nicht Null, wenn $i_1, \dots, i_n$ eine Permutation von $1, \dots, n$ ist. Daher wird aus den n Summen eine Summe über alle Permutatitionen, nämlich

$$\det(A) = \sum_{\pi \in S_n} a_{\pi(1),1} a_{\pi_2,2} \cdots a_{\pi(n),n} \det(\begin{pmatrix} e_{\pi(1)} & \cdots & e_{\pi(n)} \end{pmatrix}).$$

Unsere Definition des Signums einer Permutation liefert die Behauptung.

□

Wir schauen uns die Determinante nach der Leibniz-Formel für kleine n an. Vorher bemerken wir: Die Leibniz-Formel für die Determinante von $n \times n$-Matrizen hat $n!$ Summanden und jeder Summand ist ein Produkt aus n Faktoren (und dazu das Vorzeichen).

Zählen wir die Anzahl der Additionen und Multiplikationen, die nötig sind, um die Determinante mit der Leibniz-Formel auszurechnen, so kommen wir auf etwas in der Größenordnung $n!n$, was sogar für kleine Matrizen schon sehr groß ist (für $n = 10$ schon über 36 Millionen, für $n = 20$ schon fast $5 \cdot 10^{19}$! Die Methode per Gauß-Eliminition braucht viel weniger Additionen und Multiplikationen.

- $n = 1$: Hier ist $A = (a_{11})$ ein Skalar. Die Leibniz-Formel hat also nur einen Summanden, der aus einem Produkt mit einem Faktor besteht. Das einzige Element der symmetrischen Gruppe S_1 ist die Identität ϵ für welche $\operatorname{sig}(\epsilon) = 1$ gilt. Daher ist

$$\det(A) = \sum_{\pi \in S_1} \operatorname{sig}(\pi) \prod_{i=1}^{1} a_{\pi(i),i} = \operatorname{sig}(\epsilon) a_{\epsilon(1),1} = a_{11}.$$

- $n = 2$: Die symmetrischen Gruppe S_2 hat die zwei Elemente ϵ und τ mit $\tau(1) = 2$ und $\tau(2) = 1$. Daher gilt

$$\begin{aligned}\det(A) &= \sum_{\pi \in S_2} \operatorname{sig}(\pi) \prod_{i=1}^{n} a_{\pi(i),i} \\ &= \operatorname{sig}(\epsilon) a_{\epsilon(1),1} a_{\epsilon(2),2} + \operatorname{sig}(\tau) a_{\tau(1),1} a_{\tau(2),2} \\ &= a_{11}a_{22} - a_{21}a_{12}\end{aligned}$$

- $n = 3$: Hier hat S_3 schon sechs Elemente, nämlich

$$
\begin{array}{llrcl}
 & & 1 & \to & 1 \\
\text{Identität} & \epsilon & 2 & \to & 2 \\
 & & 3 & \to & 3 \\
 & & & & \\
 & & 1 & \to & 2 \\
\text{Zykel 1} & \sigma^1 & 2 & \to & 3 \\
 & & 3 & \to & 1 \\
 & & & & \\
 & & 1 & \to & 3 \\
\text{Zykel 2} & \sigma^2 & 2 & \to & 1 \\
 & & 3 & \to & 2 \\
 & & & & \\
 & & 1 & \to & 2 \\
\text{Transposition 1} & \tau^1 & 2 & \to & 1 \\
 & & 3 & \to & 3 \\
 & & & & \\
 & & 1 & \to & 3 \\
\text{Transposition 2} & \tau^2 & 2 & \to & 2 \\
 & & 3 & \to & 1 \\
 & & & & \\
 & & 1 & \to & 1 \\
\text{Transposition 3} & \tau^3 & 2 & \to & 3 \\
 & & 3 & \to & 2
\end{array}.
$$

Daher ist die Formel für die Determinante

$$
\begin{aligned}
\det(A) &= \sum_{\pi \in S_3} \operatorname{sig}(\pi) \prod_{i=1}^{3} a_{\pi(i),i} \\
&= \operatorname{sig}(\epsilon) a_{\epsilon(1),1} a_{\epsilon(2),2} a_{\epsilon(3),3} \\
&\quad + \operatorname{sig}(\sigma^1) a_{\sigma^1(1),1} a_{\sigma^1(2),2} a_{\sigma^1(3),3} \\
&\quad + \operatorname{sig}(\sigma^2) a_{\sigma^2(1),1} a_{\sigma^2(2),2} a_{\sigma^2(3),3} \\
&\quad + \operatorname{sig}(\tau^1) a_{\tau^1(1),1} a_{\tau^1(2),2} a_{\tau^1(3),3} \\
&\quad + \operatorname{sig}(\tau^2) a_{\tau^2(1),1} a_{\tau^2(2),2} a_{\tau^2(3),3} \\
&\quad + \operatorname{sig}(\tau^3) a_{\tau^3(1),1} a_{\tau^3(2),2} a_{\tau^3(3),3} \\
&= a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} \\
&\quad - a_{21}a_{12}a_{33} - a_{31}a_{22}a_{13} - a_{11}a_{32}a_{23}
\end{aligned}
$$

Die Fälle $n \geq 4$ schreiben wir nicht auf. Allein für $n = 4$ müssten wir 24 Summanden mit je 4 Faktoren aufschreiben…

21 Eigenwerte, Eigenvektoren

Die Multiplikation mit Diagonalmatrizen ist besonders einfach: Ist

$$A = \begin{pmatrix} k_1 & 0 & \cdots 0 & \\ 0 & k_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & k_n \end{pmatrix} =: \operatorname{diag}(k_1, \dots, k_n).$$

so ist

$$Av = \begin{pmatrix} k_1 & 0 & \cdots & 0 \\ 0 & k_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & k_n \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} k_1 v_1 \\ k_2 v_2 \\ \vdots \\ k_n v_n \end{pmatrix},$$

Die Abbildung $\operatorname{diag} : K^n \to K^{n\times n}$ bildet einen Vektor auf die Matrix ab, bei der dieser Vektor auf der Diagonalen steht (und die sonst nur Nullen enthält).

d.h. der i-te Eintrag wird mit k_i skaliert.

Entsprechend ist eine lineare Abbildung $f : V \to V$ eines n-dimensionalen K-Vektorraumes vergleichsweise einfach, wenn es Basen B und B' von V gibt, so dass die Matrixdarstellung von f diagonal ist, also wenn Skalare $k_1, \dots, k_n$ existieren, so dass

$$M_{B'}^{B}(f) = \begin{pmatrix} k_1 & 0 & \cdots & 0 \\ 0 & k_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & k_n \end{pmatrix}.$$

Erlaubt man, dass $B \neq B'$, so ist dies immer möglich: Man wähle eine Basis $\{v_1, \dots, v_k\}$ von $\operatorname{Kern}(f)$ und ergänze diese zur Basis $B = \{v_1, \dots, v_n\}$. Dann setzt man $w_{k+1} = f(v_{k+1})$, ..., $w_n = f(v_n)$ und ergänzt dies zur Basis $B' = \{w_1, \dots, w_n\}$ dann ist $M_{B'}^{B}(f)$ diagonal, sogar von der Form $M_{B'}^{B}(f) = \begin{pmatrix} 0_{k\times k} & 0 \\ 0 & E_{n-k\times n-k} \end{pmatrix}$.

Fordert man hingegen, dass $B = B'$ gelten muss, so lässt sich nicht immer eine Diagonalgestalt erreichen, stattdessen definiert man:

Definition 21.1. Es sei V ein K-Vektorraum, $f : V \to V$ linear. Dann heißt f *diagonalisierbar*, wenn es eine Basis B von V gibt, so dass $M_B^B(f)$ eine Diagonalmatrix ist.

Definition 21.2. Sei $f : V \to V$ linear. Ein Skalar k heißt *Eigenwert* von f, wenn es einen Vektor $v \in V$, $v \neq 0$ gibt, so dass

$$f(v) = kv.$$

Ein Vektor $v \in V$, $v \neq 0$ heißt *Eigenvektor* von f zu Eigenwert k, falls

$$f(v) = kv.$$

Beachte: $v = 0$ ist als Eigenvektor ausgeschlossen! (Wegen $f(0) = 0 = k \cdot 0$ wäre sonst $v = 0$ immer ein Eigenvektor zu jedem Eigenwert k, was im Folgenden nicht sinnvoll wäre.)

Beispiel.

1. Für die Identität id_V gilt $\mathrm{id}_V(v) = 1 \cdot v$, also ist 1 ein Eigenwert und jedes $v \in V$ ist Eigenvektor zu diesem Eigenwert. Es gibt keine anderen Eigenwerte.

2. Die Nullabbildung $f(v) = 0 = 0 \cdot v$ hat 0 als Eigenwert und jedes $v \in V$ ist Eigenvektor zu diesem Eigenwert. Es gibt

keine anderen Eigenwerte.

3. Ist $\mathrm{Kern}(f)$ nicht trivial, d.h. $\mathrm{Kern}(f) \neq \{0\}$, so ist jeder Vektor $v \in \mathrm{Kern}(f), v \neq 0$ Eigenvektor von f zum Eigenwert 0, d.h.

$$\begin{aligned}&\mathrm{Kern}(f)\\&= \{v \in V \mid v \text{ Eigenvektor von } f \text{ zum Eigenwert } 0\} \cup \{0\}.\end{aligned}$$

4. Für $f : \mathbb{R}^2 \to \mathbb{R}^2$ mit $f(x,y) := (x,0)$ gilt: f hat die Eigenwerte 1 und 0. Jeder Vektor der Form $(x,0)$, $x \neq 0$ ist Eigenvektor zum Eigenwert 1 und jeder Vektor der Form $(0,y)$, $y \neq 0$ ist Eigenvektor zum Eigenwert 0.

△

Satz 21.3. *Eine lineare Abbildung $f : V \to V$ ist genau dann diagonalisierbar, wenn V eine Basis aus Eigenvektoren von f hat.*

Beweis.
„$\Longrightarrow$“: Sei f diagonalisierbar, d.h. es gibt eine Basis $B = \{v_1, \dots, v_n\}$ und Skalare $k_1, \dots, k_n$, so dass $M_B^B(f) = \mathrm{diag}(k_1, \dots, k_n)$. Das heißt aber, dass $f(v_j) = k_j v_j$ gilt, also enthält die Basis B nur Eigenvektoren.

„$\Longleftarrow$“: Es sei $B = \{v_1, \dots, v_n\}$ eine Basis aus Eigenvektoren (mit jeweiligen Eigenwerten $k_1, \dots, k_n$, nicht notwendigerweise verschieden). Die j-te Spalte von $M_B^B(f)$ sind die Entwicklungskoeffizienten von $f(v_j)$ bezüglich B. Da v_j Eigenvektor von f zum Eigenwert k_j ist, gilt

$$f(v_j) = k_j v_j$$

und das bedeutet, dass das k_j-fache des j-te Standardbasisvektors e_j genau die j-Spalte von $M_B^B(f)$ bildet. Also ist $M_B^B(f) = \mathrm{diag}(k_1, \dots, k_n)$.

□

Analog zu Eigenwerten und -vektoren von linearen Abbildung betrachten wir auch Eigenwerte und -vektoren von Matrizen $M \in K^{n \times n}$, in dem wir M als lineare Abbildung $K^n \to K^n$ auffassen. D.h. $x \in K^n$ ist Eigenvektor von M zum Eigenwert λ, wenn

$$Mx = \lambda x.$$

Beispiel.
Die Matrix

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

hat die Eigenwerte 1 und 0, denn es gilt $Me_1 = e_1 = 1\,e_1$ und

$Me_2 = 0 = 0\,e_2$. Insbesondere ist jeder Vektor der Form $\left(\begin{smallmatrix} x \\ 0 \end{smallmatrix}\right)$ Eigenvektor zum Eigenwert 1 und jeder Vektor der Form $\left(\begin{smallmatrix} 0 \\ x \end{smallmatrix}\right)$ Eigenvektor zum Eigenwert 0. △

Satz 21.4. *Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig.*

Beweis.
Es seien $v_1, \ldots, v_m$ Eigenvektoren von f zu den Eigenwerten $k_1, \ldots, k_m$, und die $k_1, \ldots, k_m$ seien paarweise verschieden. Wir zeigen per Induktion über m, dass die $v_1, \ldots, v_m$ linear unabhängig sind.

Der Fall $m = 1$ ist klar, $v_1 \neq 0$ per Definition eines Eigenvektors.

Gelte die Behauptung also für $m - 1$ und es gelte

$$h_1 v_1 + h_2 v_2 + \cdots + h_m v_m = 0.$$

Wir wenden f auf diese Gleichung an und bekommen

$$0 = h_1 f(v_1) + \cdots + h_m f(v_m) = h_1 k_1 v_1 + \cdots + h_m k_m v_m.$$

Andererseits gilt ebenso

$$0 = k_m(h_1 v_1 + \cdots + h_m v_m) = k_m h_1 v_1 + \cdots + k_m h_m v_m.$$

Ziehen wir die vorige Gleichung von dieser ab, bekommen wir

$$0 = (k_m - k_1)h_1 v_1 + (k_m - k_2)h_2 v_2 + (k_m - k_{m-1})h_{m-1} v_{m-1}.$$

Nach der Induktionsannahme folgt nun, dass

$$(k_m - k_1)h_1 = 0,\ (k_m - k_2)h_2 = 0, \ldots, (k_m - k_{m-1})h_{m-1} = 0.$$

Da die k_j paarweise verschieden sind, folgt $k_m - k_j \neq 0$ für $j = 1, \ldots, m-1$ und daher auch $h_1 = h_2 = \cdots h_{m-1} = 0$. Damit folgt nun auch $h_m = 0$. □

Korollar 21.5. *Eine lineare Abbildung auf einem n-dimensionalen Vektorraum ist diagonalisierbar, wenn sie n verschiedene Eigenwerte hat.*

Die Umkehrung gilt nicht! Die Identität hat nur den Eigenwert 1 und ist bezüglich jeder Basis diagonal.

Nach Satz 21.4 bilden die Eigenvektoren in diesem Fall eine Basis; die Behauptung folgt also aus Satz 21.3.

Definition 21.6. Für eine lineare Abbildung f und einen Skalar k definieren wir

$$\mathrm{Eig}(f,k) := \{v \in V \mid f(v) = kv\}.$$

Ist k ein Eigenwert von f, so nennen wir $\mathrm{Eig}(f,k)$ den *Eigenraum* von f zu k.

Lemma 21.7. *Für eine lineare Abbildung f und $k \in K$ gilt*

Hier ist k nicht als Eigenwert vorausgesetzt!

i) $\mathrm{Eig}(f,k)$ *ist ein Unterraum von V.*

ii) *k ist Eigenwert von f genau dann, wenn* $\mathrm{Eig}(f,k) \neq \{0\}$.

iii) $\mathrm{Eig}(f,k) = \mathrm{Kern}(f - k\,\mathrm{id})$

iv) *Sind $k_1, \ldots, k_s$ paarweise verschieden, und ist B_r eine Basis von* $\mathrm{Eig}(f,k_r)$, *$r = 1, \ldots, s$, so ist $B = B_1 \cup \cdots \cup B_s$ eine linear unabhängige Menge in V.*

Beweis.

i) Es seien $v, w \in \mathrm{Eig}(f,k)$ und $h \in K$. Dann gilt

$$f(v + hw) = f(v) + hf(w) = kv + khw = k(v + hw)$$

also gilt $v + hw \in \mathrm{Eig}(f,k)$, was zeigt, dass $\mathrm{Eig}(f,k)$ ein Unterraum.

ii) Folgt direkt aus der Definition eines Eigenwertes.

iii) Es gilt

$$\begin{aligned}\mathrm{Eig}(f,k) &= \{v \in V \mid f(v) = kv\}\\ &= \{v \in V \mid f(v) - kv = 0\}\\ &= \{v \in V \mid (f - k\,\mathrm{id}_v)(v) = 0\} = \mathrm{Kern}(f - k\,\mathrm{id}_V)\}.\end{aligned}$$

iv) Wir bezeichnen die Elemente der Basis B_r mit $\{v_1^{(r)}, \ldots, v_{n_r}^{(r)}\}$, $r = 1, \ldots, s$ und betrachten eine Linearkombination der Null durch alle Elemente der Basen $B_1, \ldots, B_s$, also

$$0 = \sum_{i=1}^{n_1} h_i^{(1)} v_i^{(1)} + \cdots + \sum_{i=1}^{n_s} h_i^{(s)} v_i^{(s)}.$$

Die Vektoren $v_r = \sum_{i=1}^{n_r} h_i^{(r)} v_i^{(r)}$ sind dann in $\mathrm{Eig}(f,k_r)$ und es gilt

$$0 = v_1 + \cdots + v_s.$$

Da Eigenvektoren zu verschiedenen Eigenwerten nach Satz 21.4 linear unabhängig sind, müssen die v_i alle $= 0$ sein. Da die B_r alle Basen sind, folgt auch $h_i^{(r)} = 0$ für alle r und i.

□

Korollar 21.8. *Eine lineare Abbildung eines n-dimensionalen Vektorraumes in sich ist genau dann diagonalsierbar, wenn die Summe der Dimensionen der Eigenräume gleich n ist.*

Ist f diagonalisierbar, so gibt es eine Basis aus Eigenvektoren und da jeder Eigenvektor in einem Eigenraum liegt, folgt, dass die Summe der Dimensionen der Eigenräume gleich n ist. Ist umgekehrt die Summe der Dimensionen der Eigenräume $\geq n$, so ist die Vereinigung von Basen der Eigenräume eine Basis des ganzen Raumes.

Beispiel.
Wir untersuchen die Matrix $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) \in K^{2\times 2}$ auf Diagonalisierbarkeit (bzw. die zugehörige lineare Abbildung $f(x) = Ax$).

Ein Vektor x ist Eigenvektor zum Eigenwert λ genau dann, wenn $Ax = \lambda x$, also

$$\begin{aligned} x_2 &= \lambda x_1 \\ 0 &= \lambda x_2. \end{aligned}$$

Ist $\lambda \neq 0$, so folgt aus der zweiten Zeile $x_2 = 0$ und die erste Zeile liefert $x_1 = 0$ (oder $\lambda = 0$, aber das ist ja in diesem Fall ausgeschlossen). Ist $\lambda = 0$, so liefert die erste Zeile $x_2 = 0$ und für x_1 haben wir keine Gleichung.

Das heißt, $\lambda = 0$ ist der einzige Eigenwert von A und der zugehörige Eigenraum ist

$$\operatorname{Eig}(A, 0) = \{x \in \mathbb{R}^2 \mid x_2 = 0\} = \langle e_1 \rangle.$$

Der Eigenraum ist eindimensional und die Summe der Dimensionen der Eigenräume ist 1, daher ist A nicht diagonalisierbar.

$\triangle$

22 Berechnung von Eigenwerten und das charakteristische Polynom

Um zu einer gegebenen Matrix $A \in K^{n\times n}$ einen Eigenwert λ und einen Eigenvektor x zu finden, muss man die Gleichung

$$Ax = \lambda x$$

gleichzeitig nach λ und x lösen. Aufgrund des Produktes λx auf der rechten Seite ist dies ein *nichtlineares* Gleichungssystem (mit $n+1$ Unbekannten). Unsere bisherigen Techniken helfen uns also auf den ersten Blick nicht weiter. Es gilt jedoch folgende einfache Beobachtung (vgl Lemma 21.7):

Lemma 22.1. *Ist $A \in K^{n\times n}$, so ist x Eigenvektor von A zum Eigenwert λ genau dann, wenn $x \in \operatorname{Kern}(A - \lambda E_n)$ und $x \neq 0$.*

Beweis.
Ein Vektor $x \neq 0$ ist Eigenvektor von A zum Eigenwert λ genau dann, wenn $Ax = \lambda x$, was genau dann der Fall ist, wenn $Ax - \lambda x = 0$, was wiederum äquivalent zu $(A - \lambda E_n)x = 0$ ist. Das zeigt die Behauptung. □

Dieses Lemma erlaubt uns, Eigenwerte einer Matrix zu bestimmen, ohne gleichzeitig Eigenvektoren ausrechnen zu müssen. Es gilt nämlich:

Satz 22.2. *Für eine Matrix $A \in K^{n\times n}$ sind äquivalent*

i) λ *ist ein Eigenwert von* A

ii) $\operatorname{Kern}(A - \lambda E_n) \neq \{0\}$

iii) $\operatorname{Rang}(A - \lambda E_n) < n$

iv) $\det(A - \lambda E_n) = 0$.

Beweis.
Die Äquivalenz von i) und ii) ist gerade die Aussage von Lemma 22.1. Die Äquivalenz von ii) und iii) ist die Kern-Rang-Formel aus Satz 10.8. Die Äquivalenz von iii) und iv) ist die Eigenschaft der Determinante, dass sie genau dann Null ist, wenn die Matrix nicht Rang n hat. □

Satz 22.3 (Satz und Defintion). *Die Abbildung $p(\lambda) = \det(A - \lambda E_n)$ ist ein Polynom vom Grad n, genannt charakteristisches Polynom von A. Wir schreiben $\chi_A(\lambda) := \det(A - \lambda E_n)$.*

Beweis.
Wir benutzen die Leibniz-Formel für die Determinante und bemerken dass für $B = A - \lambda E_n$ gilt

$$b_{ij} = \begin{cases} a_{ij} & i \neq j \\ a_{ij} - \lambda & i = j. \end{cases}$$

Daher sind die Faktoren der Summanden auf der rechten Seite von

$$\det(A - \lambda E_n) = \sum_{\pi \in \mathbb{S}_n} \operatorname{sig}(\pi) \prod_{i=1}^{n} b_{\pi(i),i}$$

von der Form a_{ij} oder $(a_{ii} - \lambda)$, also jeweils ein Produkt von Polynomen. Der höchste Polynomgrad, den ein Summand haben kann, ist n und der wird angenommen, für $\pi = \epsilon$, denn dort ist der Summand $(a_{11} - \lambda) \cdots (a_{nn} - \lambda)$. □

Korollar 22.4. *Die Eigenwerte von A sind genau die Nullstellen des charakteristischen Polynoms χ_A.*

Folgt aus Satz 22.2 und Definition 22.3.

Beispiel.

1. Es sei K beliebig und

$$A = \operatorname{diag}(k_1, \ldots, k_n)$$

eine Diagonalmatrix. Dann ist das charakteristische Polynom

$$\begin{aligned}\chi_A(\lambda) &= \det(A - \lambda E_n) = \det(\operatorname{diag}(k_1 - \lambda, \ldots, k_n - \lambda)) \\ &= (k_1 - \lambda) \cdots (k_n - \lambda).\end{aligned}$$

Dies Polynom ist direkt in *Linearfaktoren* angegeben, und daher können wir die Nullstellen und damit die Eigenwerte von A direkt ablesen: Es sind $k_1, \ldots, k_n$. (Dies hätten wir auch direkt sehen können: Der j-te Einheitsvektor e_j is offensichtlich Eigenvektor von A zum Eigenwert k_j.)

2. Für $K = \mathbb{R}$ und $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ ist

$$\chi_A(\lambda) = \det(\begin{pmatrix} 1-\lambda & 0 \\ 0 & 1-\lambda \end{pmatrix}) = (1 - \lambda)^2,$$

also gibt es den Eigenwert $\lambda = 1$, dieser ist jedoch eine „doppelte Nullstelle". Der zugehörige Eigenraum ist

$$\operatorname{Eig}(A, 1) = \operatorname{Kern}(A - E_2) = \operatorname{Kern}(\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}) = \mathbb{R}^2,$$

insbesondere ist er zweidimensional.

3. Für $K = \mathbb{R}$ und $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ ist

$$\chi_A(\lambda) = \det(\begin{pmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix}) = (1 - \lambda)^2,$$

also gibt es wieder den Eigenwert $\lambda = 1$ mit „doppelter Nullstelle". Der zugehörige Eigenraum ist

$$\operatorname{Eig}(A, 1) = \operatorname{Kern}(A - E_2) = \operatorname{Kern}(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}) = \langle e_1 \rangle$$

insbesondere ist er hier nur eindimensional.

4. Für $K = \mathbb{R}$ und $A = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ ist

$$\chi_A(\lambda) = \det(\left(\begin{smallmatrix} -\lambda & 1 \\ 1 & -\lambda \end{smallmatrix}\right)) = \lambda^2 - 1 = (\lambda - 1)(\lambda + 1),$$

also gibt es die beiden Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = -1$. Die zugehörigen Eigenräume sind

$$\operatorname{Eig}(A, 1) = \operatorname{Kern}(A - E_2) = \operatorname{Kern}(\left(\begin{smallmatrix} -1 & 1 \\ 1 & -1 \end{smallmatrix}\right)) = \langle\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)\rangle$$

und

$$\operatorname{Eig}(A, -1) = \operatorname{Kern}(A + E_2) = \operatorname{Kern}(\left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)) = \langle\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)\rangle.$$

5. Es sei $K = \mathbb{R}$ und

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

In diesem Fall ist das charakteristische Polynom

$$\begin{aligned}\chi_A(\lambda) &= \det(A - \lambda E_2) = \det(\begin{pmatrix} -\lambda & 1 \\ 1 & 1-\lambda \end{pmatrix}) \\ &= -\lambda(1-\lambda) - 1 = \lambda^2 - \lambda - 1.\end{aligned}$$

Die Nullstellen erhalten wir aus der bekannten pq-Formel (die Nullstellen von $x^2 + px + q$ sind $-\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}$):

$$\lambda_{1/2} = \tfrac{1}{2} \pm \sqrt{\tfrac{1}{4} + 1} = \tfrac{1}{2} \pm \sqrt{\tfrac{5}{4}} = \tfrac{1 \pm \sqrt{5}}{2}.$$

Um die zugehörigen Eigenvektoren zu berechnen, berechnen wir den Kern von $A - \lambda_{1/2}E_2$. Nehmen wir exemplarisch $\lambda_1 = \frac{1+\sqrt{5}}{2}$ so müssen wir den Kern von

$$\begin{pmatrix} -\frac{1+\sqrt{5}}{2} & 1 \\ 1 & 1 - \frac{1+\sqrt{5}}{2} \end{pmatrix} = \begin{pmatrix} -\frac{1+\sqrt{5}}{2} & 1 \\ 1 & \frac{1-\sqrt{5}}{2} \end{pmatrix}$$

bestimmen. Da wir wissen, dass diese Matrix einen eindimensionalen Kern hat, können wir den Ansatz $v_1 = \left(\begin{smallmatrix} 1 \\ x \end{smallmatrix}\right)$ machen, und sehen dass

$$v_1 = \begin{pmatrix} 1 \\ \frac{1+\sqrt{5}}{2} \end{pmatrix}$$

ein Eigenvektor zum Eigenwert λ_1 ist. Entsprechend ist

$$v_2 = \begin{pmatrix} 1 \\ \frac{1-\sqrt{5}}{2} \end{pmatrix}$$

ein Eigenvektor zum Eigenwert $\lambda_2 = \frac{1-\sqrt{5}}{2}$.

$\triangle$

Um also zu gegebenem A ein Paar aus Eigenwert und Eigenvektor zu finden, können wir wie folgt vorgehen:

1. Finde eine Nullstelle λ des charakteristischen Polynoms χ_A.
2. Berechne ein Element $x \neq 0$ mit $x \in \operatorname{Kern}(A - \lambda E_n)$.

Dies ist ein *nichtlineares* Problem!

Die Problem ist linear und lässt sich mit Hilfe von Gauß-Elimination lösen. Beachte, dass der erste Schritt garantiert, dass ein solches x existiert!

Bei der Berechnung von charakteristischen Polynomen bei $n \times n$ Matrizen mit $n \geq 4$ sind unsere bisherigen Möglichkeiten oft nicht sehr praktisch. Eine Technik, die hier in vielen Fällen gut funktioniert, basiert auf folgendem Satz:

Satz 22.8 (Laplacescher Entwicklungssatz). *Zu $A \in K^{n\times n}$ und $i, j \in \{1, \dots, n\}$ sei $A_{ij} \in K^{(n-1)\times(n-1)}$ die Matrix, welche aus A durch Streichen der i-ten Zeile und j-ten Spalte entsteht.*

Für jedes $i \in \{1, \dots, n\}$ gilt

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) \qquad \text{(Entwicklung nach der } i\text{-ten Zeile)}$$

und für jedes $j \in \{1, \dots, n\}$ gilt

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}). \qquad \text{(Entwicklung nach der } j\text{-ten Spalte)}$$

Beweis.
Wir betrachten den Fall $j = n$ (Entwicklung nach der letzten Spalte) und starten von der Leibniz-Formel aus Satz 20.6:

$$\det(A) = \sum_{\pi \in S_n} \operatorname{sig}(\pi) \prod_{i=1}^{n} a_{\pi(i),i} \tag{*}$$

Da jedes Produkt $\prod_{i=1}^{n} a_{\pi(i),i}$ einen Faktor aus der letzten Zeile enthält, können wir schreiben

$$\det(A) = a_{1n}c_{1n} + a_{2n}c_{2n} + \cdots + a_{nn}c_{nn},$$

wobei c_{in} jeweils die Größe beschreibt, die beim Ausklammern von a_{in} übrig bleibt. Die Behauptung ist gezeigt, wenn wir zeigen, dass

$$c_{in} = (-1)^{i+n} \det(A_{in}).$$

Schauen wir noch einmal auf (*) und sammeln alle Summanden, in denen a_{nn} vorkommt. Wir bekommen

$$a_{nn}c_{nn} = a_{nn} \sum_{\pi \in S_n(n)} \operatorname{sig}(\pi) \prod_{i=1}^{n-1} a_{\pi(i),i}$$

(erinnere: $S_n(n)$ sind die Permutationen, die n fixiert lassen). Die Leibniz-Formel für die Determinante von $A_{n,n} \in K^{(n-1)\times(n-1)}$ liefert

$$\det(A_{n,n}) = \sum_{\pi \in S_{n-1}} \operatorname{sig}(\pi) \prod_{i=1}^{n-1} a_{\pi(i),i}$$

und dies zeigt

$$c_{nn} = \det(A_{n,n}) = (-1)^{n+n} \det(A_{n,n}).$$

Die Fälle der c_{in} führen wir auf diesen Fall zurück: Wir tauschen die i-te Zeile mit der nächsten so lange, bis sie zur letzten Zeile geworden ist (das ist $n - i$ Mal). Dabei ändert sich der Wert $\det(A_{in})$ nicht geändert, jedoch hat sich $\det(A)$ und damit c_{in} um den Faktor $(-1)^{n-i}$ geändert. Wir erhalten also

$$c_{in} = (-1)^{n-i} \det(A_{i,n})$$

was die Behauptung für die Entwicklung nach der n-ten Spalte zeigt.

Die Entwicklung nach anderen Spalten folgt wieder durch entsprechende Permutation von Spalten und die Entwicklung nach den Zeilen folgt aus der Aussage für die Spalten, da $\det(A) = \det(A^T)$ gilt (A^T ist die transponierte Matrix, also $(A^T)_{ij} = A_{ji}$, vgl. Übung). □

Der folgende Rest dieses Kapitels wurde in der Vorlesung übersprungen und steht hier nur der Vollständigkeit halber. Stattdessen wurden der Begriff „Ähnlichkeit" in der großen Übung thematisiert.

Definition 22.5. Zwei Matrizen $A, B \in K^{n\times n}$ heißen *ähnlich*, wenn es eine invertierbare Matrix $S \in K^{n\times n}$ gibt, so dass

$$B = S^{-1}AS.$$

Man zeigt leicht: Ähnlichkeit von Matrizen ist eine Äquivalenzrelation.

Satz 22.6. *Sind A und B ähnlich, so gilt $\chi_A = \chi_B$.*

Beweis.
Da A und B ähnlich sind, gibt es S mit $B = SAS^{-1}$, also gilt

$$\begin{aligned}
\chi_B(\lambda) &= \det(B - \lambda E_n) \\
&= \det(S^{-1}AS - \lambda S^{-1}S) \\
&= \det(S^{-1}(A - \lambda E_n)S) \\
&= \det(S^{-1}) \det(A - \lambda E_n) \det(S) \\
&= \det(S)^{-1} \chi_A(\lambda) \det(S) \\
&= \chi_A(\lambda).
\end{aligned}$$

□

Korollar 22.7. *Sind A und B ähnlich, so haben sie die gleichen Eigenwerte. Genauer gilt: Ist $B = S^{-1}AS$ und ist x Eigenvektor von A zum Eigenwert λ, so ist $S^{-1}x$ ein Eigenvektor von B zum Eigenwert λ.*

Ist λ Eigenwert von A mit Eigenvektor x, so gilt $Ax = \lambda x$. Schreiben wir nun $x = Sy$, so folgt

$$ASy = \lambda Sy$$

und nach Multiplikation von links mit S^{-1} bekommen wir

$$S^{-1}ASy = \lambda y$$

was nichts anderes bedeutet, als dass $y = S^{-1}x$ ein Eigenvektor von $S^{-1}AS = B$ zum Eigenwert λ ist.

23 Bilinearformen und Skalarprodukte

Definition 23.1. Eine Abbildung

$$B : V \times V \to K$$

auf einem K-Vektorraum V heißt *Bilinearform*, wenn sie in beiden Argumenten linear ist, d.h. für $v, v', w, w' \in V$ und $k \in K$ gilt

$$B(v + kv', w) = B(v, w) + kB(v', w),$$
$$B(v, w + kw') = B(v, w) + kB(v, w').$$

Bilinearität können wir auch so ausdrücken: Für jedes v ist die Abbildung $g_v(w) := B(v, w)$ linear und für jedes w ist die Abbildung $f_w(v) = B(v, w)$ linear.

Ähnlich wie lineare Abbildungen, hängen Bilinearformen (durch die Wahl einer Basis) mit Matrizen zusammen.

Satz 23.2. *Ist V eine K-Vektorraum mit Basis $\{v_1, \ldots, v_n\}$ und $A \in K^{n\times n}$ eine Matrix, so existiert eine Bilinearform B mit*

$$B(v_i, v_j) = a_{ij}$$

Beweis.
Wir entwickeln zwei Vektoren v und w in die Basis

$$v = \sum_{i=1}^{n} k_i v_i, \quad w = \sum_{j=1}^{n} h_j v_j$$

und definieren

$$B(v, w) := \sum_{i=1}^{n} \sum_{j=1}^{n} k_i a_{ij} h_j.$$

Offensichtlich gilt $B(v_i, v_j) = a_{ij}$ und ebenso direkt sieht man Bilinearität. $\square$

Besonders einfach ist die (zur Basis $\{v_1, \ldots, v_n\}$ gehörige) Standardbilinearform

$$\langle v, w\rangle := k_1 h_1 + \cdots + k_n h_n,$$

welche zur Einheitsmatrix gehört und welche im nächsten Kapitel genauer untersucht wird.

Ist eine Bilinearform gegeben, so können wir (bei gewählter Basis) eine Matrix zuordnen:

Definition 23.3. Ist B eine Bilinearform auf V mit der Basis $\mathcal{B} = \{v_1, \ldots, v_n\}$, so ist die *Gramsche Matrix* $A = (a_{ij}) \in K^{n\times n}$ der Bilinearform gegeben durch

Hier benutzen wir ein Skript-B für Basen, da wir B selbst für die Bilinearformen brauchen.

$$a_{ij} = B(v_i, v_j).$$

Die Standardbilinearform $\langle v, w\rangle = \sum_{i=1}^{n} k_i h_i$ hat also die Einheitsmatrix als Gramsche Matrix.

Um die Formel $B(v, w) = B(\sum_{i=1}^{n} k_i v_i, \sum_{j=1}^{n} h_j v_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} k_i a_{ij} h_j$ kompakter zu formulieren, erinnern wir an das Transponieren von Matrizen (und damit von Vektoren).

Definition 23.4. Für eine Matrix $A \in K^{m\times n}$ ist die transponierte Matrix $A^T \in K^{n\times m}$ definiert durch

$$(A^T)_{ij} = A_{ji}.$$

Beispiel.
Das Transponieren bedeutet „spiegeln an der Diagonalen". So ist z.B.

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Außerdem ist das transponierte eines Zeilenvektors $x \in K^n$ ein Spaltenvektor $x^T \in K^{1\times n}$. $\triangle$

Ist V (mit der Standardbasis) und $A \in K^{n\times n}$, so ist die Bilinearform, B die zu A (und zur Basis $\{v_1, \dots, v_n\}$) gehört für $v = \sum_{i=1}^n x_i v_i$, $w = \sum_{i=1}^n y_i v_i$ gegeben durch:

$$B(v,w) = \sum_{i,j=1}^n x_i a_{ij} y_j = \sum_{i=1}^n x_i \underbrace{\sum_{j=1}^n a_{ij} y_j}_{(Ay)_i} = x^T A y.$$

Die Standardbilinearform auf K^n (mit der Standardbasis) lässt sich damit einfach als

$$\langle x, y\rangle = \sum_{i=1}^n x_i y_i = x^T y$$

schreiben.

Definition 23.5. Eine Bilinearform B heißt *nichtausgeartet*, wenn gilt:

$$(\forall w \in V : \; B(v,w) = 0) \implies v = 0.$$

Entsprechend heißt B *ausgeartet*, wenn es ein $v \neq 0$ gibt, so dass für alle w gilt $B(v,w) = 0$.

Man kan zeigen, dass es hier nicht darauf ankommt, ob das linke oder das recht Argument benutzt wird, d.h. B ist ausgeartet, wenn gilt $\exists w \neq 0 \forall v : B(v,w) = 0$.

Satz 23.6. *Ist A die Gramsche Matrix einer Bilinearform B, so gilt: B ist nichtausgeartet genau dann, wenn* $\operatorname{Rang}(A) = n$.

Beweis.
Gelte zuerst $\operatorname{Rang}(A) < n$ (und zu zeigen ist, dass B ausgeartet ist). In diesem Fall sind die Spalten von A linear abhängig, d.h. es gibt Koeffizienten $k_1, \dots, k_n$ (welche nicht alle $= 0$ sind), so dass

$$\begin{aligned} B(v_1, v_1)k_1 + \cdots + B(v_n, v_1)k_n &= 0 \\ \vdots \quad & \quad \vdots \\ B(v_1, v_n)k_1 + \cdots + B(v_n, v_n)k_n &= 0. \end{aligned}$$

Wir definieren $w = \sum_{i=1}^n k_i v_i$ (und bemerken, dass $w \neq 0$) gilt. Es folgt

$$B(w, v_j) = B(\sum_{i=1}^{n} k_i v_i, v_j) = \sum_{i=1}^{n} k_i B(v_i, v_j) = 0$$

was die Behauptung zeigt.

Ist nun B ausgeartet (und wir müssen zeigen, dass $\mathrm{Rang}(A) < n$ gilt), so gibt es w, so dass für alle v_j gilt $B(w, v_j) = 0$. Wir entwickeln w in die Basis: $w = \sum_{i=1}^{n} k_i v_i$ und bemerken, dass wenigstens ein $k_i \neq 0$ sein muss. Es folgt für alle j

$$0 = B(w, v_j) = B(\sum_{i=1}^{n} k_i v_i, v_j) = \sum_{i=1}^{n} k_i \underbrace{B(v_i, v_j)}_{=a_{ij}}$$

was zeigt, dass die Zeilen von A linear abhängig sind. □

Kommen wir zu Skalarprodukten. Hier beschränken wir uns auf reelle Vektorräume V und hier mit einer speziellen Art von Bilinearformen:

Auch für komplexe Vektorräume lassen sich Skalarprodukte definieren - dazu kommen wir in dieser Vorlesung nicht mehr.

Definition 23.7. Eine Bilinearform $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ auf einem $\mathbb{R}$-Vektorraum V heißt *Skalarprodukt*, wenn gilt

$$\begin{aligned} \forall v, w \in V : \ \langle v, w \rangle &= \langle w, v \rangle && \text{(Symmetrie)} \\ v \neq 0 \implies \langle v, v \rangle &> 0. && \text{(positive Definitheit)} \end{aligned}$$

Das $>$-Zeichen bei der positive Definitheit ergibt Sinn, da wir uns auf einem reellen Vektorraum befinden!

Wegen Bilinearität gilt immer $\langle 0, 0 \rangle = 0$, d.h. wir könnten bei der positive Definitheit auch $\iff$ statt $\implies$ schreiben.

Die positive Definitheit impliziert, dass ein Skalarprodukt eine nicht ausgeartete Bilinearform ist: Da immer $\langle v, v \rangle \geq 0$ gilt, kann nicht $\langle w, v \rangle = 0$ für alle w gelten.

Da sich Bilinearformen über die Gramsche Matrix beschreiben lassen, untersuchen wir, welche Eigenschaften die Gramsche Matrix haben muss, damit die zugehörige Bilinearform ein Skalarprodukt ist.

Lemma 23.8. *Ist V ein n-dimensionaler K-Vektorraum mit Basis $\mathcal{B} = \{v_1, \dots, v_n\}$ und B eine Bilinearform darauf mit Gramscher Matrix $A = (a_{ij}) \in K^{n \times n}$. Dann gilt: B ist symmetrisch genau dann, wenn $a_{ij} = a_{ji}$ gilt.*

Die Bedingung $a_{ij} = a_{ji}$ lässt sich auch als $A^T = A$ schreiben. Matrizen mit dieser Eigenschaft heißen *symmetrisch*. Es gilt also: Eine Bilinearform ist symmetrisch, wenn ihre Gramsche Matrix symmetrisch ist.

Beweis.

Ist B symmetrisch, so folgt sofort $a_{ij} = B(v_i, v_j) = B(v_j, v_i) = a_{ji}$.

Für die andere Richtung sei A symmetrisch und $v = \sum_{i=1}^{n} k_i v_i$, $w = \sum_{i=1}^{n} h_i w_i$. Wegen $B(v_i, v_j) = a_{ij} = a_{ji} = B(v_j, v_i)$ gilt dann

$$\begin{aligned} B(v, w) &= \sum_{i=1}^{n} \sum_{j=1}^{n} k_i h_j B(v_i, v_j) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} k_i h_j B(v_j, v_i) = B(w, v). \end{aligned}$$

□

Lemma 23.9. *Betrachte $\mathbb{R}^n$ mit der Standardbasis $\{e_1, \dots, e_n\}$ und $\langle \cdot, \cdot \rangle$ sei ein Skalarprodukt mit Gramscher Matrix A. Dann gilt für $x, y \in \mathbb{R}^n$*

$$\langle x, y \rangle = x^T Ay.$$

Um diese Skalarprodukte vom Standard-Skalarprodukt zu unterscheiden, wird oft eine Schreibweise wie $\langle \cdot, \cdot \rangle_A$ verwendet. Hier machen wir das nicht - es sollte jeweils aus dem Kontext klar sein, ob das Standardskalarprodukt oder ein anderes gemeint ist.

Beweis.
Wegen $x = \sum_{i=1}^n x_i e_i$, $y = \sum_{i=1}^n y_i e_i$ und $\langle e_i, e_j \rangle = a_{ij}$ folgt

$$\begin{aligned}\langle x, y \rangle &= \langle \sum_{i=1}^n x_i e_i, \sum_{i=1}^n e_j \rangle = \sum_{i=1}^n \sum_{j=1}^n x_i y_j \langle e_i, e_j \rangle \\ &= \sum_{i=1}^n \sum_{j=1}^n x_i y_j a_{ij} = x^T Ay.\end{aligned}$$

□

Definition 23.10. Wir nennen eine symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ *positiv definit*, wenn für jedes $x \in \mathbb{R}^n$, $x \neq 0$ gilt, dass

$$x^T Ax > 0.$$

Gilt für jedes $x \in \mathbb{R}^n$

$$x^T Ax \geq 0,$$

so heißt A *positiv semidefinit*.

Entsprechend sind die Begriff *negativ (semi-)definit* definiert.

Korollar 23.11. *Ist $A \in \mathbb{R}^{n \times n}$, so ist die die Bilinearform $B(x, y) = x^T Ay$ ein Skalarprodukt, genau dann, wenn A positiv definit ist.*

Symmetrie von A muss nicht extra gefordert werden, dass sie implizit bei der Definitheit enthalten ist.

Lemma 23.12. *Sind $A, B \in \mathbb{R}^{n \times n}$ positiv (semi-)definit und $\alpha > 0$, so sind $A + B$ und αA positiv (semi-)definit.*

Die Folgerung „positiv semidefinit" bleibt richtig, wenn $\alpha \geq 0$ ist.

Beweis.
Es gilt $(A + B)^T = A^T + B^T$ und $(\alpha A)^T = \alpha A^T$. Also sind $A + B$ und αA symmetrisch, wenn A und B symmetrisch sind.

Weiterhin ist $x^T(A + B)x = x^T(Ax + Bx) = x^T Ax + x^T Bx$ und $x^T(\alpha A)x = \alpha x^T Ax$, woraus die Behauptung folgt. □

Beispiel.
Wir untersuchen die Matrix $A = \begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix}$ auf positive Definitheit:

$$\begin{aligned} x^T Ax &= \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 1 & a \\ a & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} x_1 + ax_2 \\ ax_1 + x_2 \end{pmatrix} = x_1^2 + 2ax_1x_2 + x_2^2 \end{aligned}$$

Für $|a| \leq 1$ folgt

$$\begin{aligned} x^T Ax &\geq x_1^2 - 2|a||x_1||x_2| + x_2^2 \\ &\geq x_1^2 - 2|x_1||x_2| + x_2^2 = (|x_1| - |x_2|)^2 \geq 0. \end{aligned}$$

Also ist A für $|a| \leq 1$ positiv semidefinit. △

24 Orthogonalität und Normen

Definition 24.1. Ist $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf V, so nennen wir zwei Vektoren $v, w \in V$ *orthogonal* (bzgl. $\langle \cdot, \cdot \rangle$), wenn $\langle v, w \rangle = 0$ gilt.

Wieso Skalarprodukte etwas mit Orthogonalität im geometrischen Sinn zu haben sollen, ist bisher noch nicht klar. Um hier Klarheit zu schaffen, führen wir erst noch den Begriff der Norm ein, welcher (abstrakt) Längen von Vektoren beschreiben soll. Noch vorher beweisen wir eine Aussage, die auf den ersten Blick wenig bedeutend aussieht:

Satz 24.2 (Cauchy-Schwarz-Ungleichung)**.** *Für alle v, w aus einem reellen Vektorraum V mit Skalarprodukt $\langle \cdot, \cdot \rangle$ gilt*

$$\langle v, w \rangle^2 \leq \langle v, v \rangle \langle w, w \rangle.$$

Beweis.
Ist $w = 0$, so sind beide Seiten $= 0$ (und damit ist die Ungleichung erfüllt). Sei also $w \neq 0$. Wir benutzen die positive Definitheit, Bilinearität und Symmetrie und sehen, dass für alle Skalare $\lambda \in \mathbb{R}$ gilt

$$\begin{aligned} 0 &\leq \langle v - \lambda w, v - \lambda w \rangle = \langle v, v - \lambda w \rangle - k \langle w, v - \lambda w \rangle \\ &\langle v, v \rangle - \lambda \langle v, w \rangle - \lambda \langle w, v \rangle + \lambda^2 \langle w, w \rangle \\ &= \langle v, v \rangle - 2\lambda \langle v, w \rangle + \lambda^2 \langle w, w \rangle. \end{aligned}$$

Diese Ungleichung gilt für alle λ, d.h. sie ist dann aus „stärksten“ (im Sinne von „am aussagekräftigsten“) wenn die rechte Seite so kleine wie möglich ist. Die rechte Seite ist ein quadratisches Polynom in λ (hier benutzen wir, dass $w \neq 0$ und daher $\langle w, w \rangle > 0$) und kann daher über λ minimiert werden. Betrachten wir die rechte Seite also Funktion von λ, leiten ab und setzen $= 0$, so bekommen wir, dass das minimale Wert für $\lambda = \langle v, w \rangle / \langle w, w \rangle$ erreicht wird. Setzen wir dies ein bekommen wir

$$0 \leq \langle v, v \rangle - 2 \frac{\langle v, w \rangle}{\langle w, w \rangle} \langle v, w \rangle + \left(\frac{\langle v, w \rangle}{\langle w, w \rangle} \right)^2 \langle w, w \rangle.$$

Multiplikation mit $\langle w, w \rangle$ liefert

$$0 \leq \langle w, w \rangle \langle v, v \rangle - \langle v, w \rangle^2$$

woraus die Behauptung folgt. □

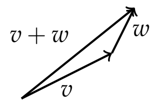
Nur zur Defintion des Begriffes des Norm:

Definition 24.3. Es sei V ein reeller Vektorraum. Eine Abbildung $\|\cdot\| : V \to [0, \infty[$ heißt *Norm*, wenn für alle $v, w \in V$ und $k \in \mathbb{R}$

gilt

$$
\begin{aligned}
\|kv\| &= |k|\|v\| && \text{(positive Homogenität)}\\
\|v+w\| &\leq \|v\| + \|w\| && \text{(Dreiecksungleichung)}\\
\|v\| = 0 &\iff v = 0 && \text{(Definitiheit)}
\end{aligned}
$$

Anschaulich: Positive Homogenität bedeutet, dass die Skalierung die „Länge“ entsprechend skaliert. Die Dreicksungleichung bedeutet, dass ein Umweg „länger“ als der direkte Weg ist



und Definitheit bedeutet, dass nur der Nullvektor „Länge“ Null hat.

Für von Skalarprodukten induzierte Normen gelten Entsprechungen der Binomischen Formeln:

$\|x+y\|^2 = \|x\|^2 + 2\langle x,y\rangle + \|y\|^2$, usw.

Die Cauchy-Schwarz-Ungleichung liest sich mit Hilfe der induzierten Norm als $|\langle x,y\rangle| \leq \|x\|\|y\|$.

Satz 24.4. *Ist $\langle\cdot,\cdot\rangle$ ein Skalarprodukt auf einem reellen Vektorraum V, so ist*

$$\|v\| := \sqrt{\langle v,v\rangle}$$

eine Norm, genannt die von $\langle\cdot,\cdot\rangle$ induzierte Norm.

Beweis.

- Positive Homogenität:

$$
\begin{aligned}
\|kv\| &= \sqrt{\langle kv,kv\rangle} = \sqrt{kk\langle v,v\rangle}\\
&= \sqrt{k^2}\sqrt{\langle v,v\rangle} = |k|\|v\|.
\end{aligned}
$$

- Dreiecksungleichung: Mit der Cauchy-Schwarz-Ungleichung folgt

$$
\begin{aligned}
\|v+w\|^2 &= \langle v+w, v+w\rangle = \langle v,v\rangle + \langle v,w\rangle + \langle w,v\rangle + \langle w,w\rangle\\
&= \|v\|^2 + 2\langle v,w\rangle + \|w\|^2\\
&\leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2\\
&= (\|v\| + \|w\|)^2.
\end{aligned}
$$

- Positive Definitheit: Ist $v=0$, so ist $\|v\|^2 = \langle v,v\rangle = 0$ und ist $\|v\| = 0$, so ist $\langle v,v\rangle = 0$ und es folgt $v = 0$ aus der Definitheit von Skalarprodukten.

$\square$

Beispiel.

1. Auf dem reellen Raum $\mathbb{R}^n$ spielt das Standardskalarprodukt eine besondere Rolle. Dies ist

$$\langle x,y\rangle = \sum_{i=1}^n x_i y_i = x^T y.$$

Dies induziert die sogenannte *Euklidsche Norm*:

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

Im $\mathbb{R}^2$ ist diese

$$\left\|\begin{pmatrix} x_1\\ x_2\end{pmatrix}\right\| = \sqrt{x_1^2 + x_2^2}$$

was nach dem Satz von Pythagoras der geometrischen Länge entspricht. In höheren Dimensionen gilt das Gleiche.

2. Nicht alle Normen werden von Skalarprodukten induziert. So gibt es zum Beispiel noch die 1-Norm:

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$

oder die ∞-Norm (auch Maximums-Norm genannt)

$$\|x\|_\infty = \max_{i=1,\dots,n} |x_i|.$$

In beiden Fällen ist das Nachrechnen der Normaxiome einfach.

3. Auf dem Vektorraum der Riemann-integrierbaren Funktionen auf dem Interval $[a, b]$ ist das folgende ein Skalarprodukt:

$$\langle f, g\rangle = \int_a^b f(x)g(x)dx.$$

Dies induziert die 2-Norm

$$\|f\|_2 = \left(\int_a^b f(x)^2 dx\right)^{1/2}$$

welche z.B. bei partiellen Differentialgleichungen, in der Quantenmechanik und auch sonst an vielen Stellen der Mathematik und Physik eine wichtige Rolle spielt.

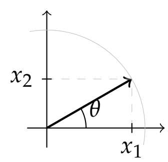
$\triangle$

Kommen wir nun dazu, wieso Skalarprodukte etwas mit geometrischer Orthogonalität und Winkeln zu tun haben: Wir benutzen hier nur das Standardskalarprodukt im $\mathbb{R}^2$ und die zugehörige Euklidsche Norm, also

$$\langle x, y\rangle = x_1y_1 + x_2y_2 \quad \text{und} \quad \|x\|_2 = \sqrt{x_1^2 + x_2^2}.$$

Gilt $\|x\|_2 = 1$ und bezeichnen wir den Winkel zwischen der Rechtsachse und dem Vektor x mit θ, so ist nach elementarer Geometrie (siehe Skizze)

$$\langle x, e_1\rangle = x_1 = \cos(\theta).$$



Für beliebiges $x \in \mathbb{R}^2$ gilt $x = \|x\|_2 \frac{x}{\|x\|_2}$ und $\left\|\frac{x}{\|x\|_2}\right\|_2 = 1$ und daher

$$\langle x, e_1\rangle = \|x\|_2 \cos(\theta).$$

Nun überlegt man sich, dass die Matrix $R_\alpha = \begin{pmatrix}\cos(\alpha) & -\sin(\alpha)\\ \sin(\alpha) & \cos(\alpha)\end{pmatrix}$ eine Drehung des $\mathbb{R}^2$ (um den Winkel α gegen den Uhrzeigersinn)

Die Drehrichtung gegen den Uhrzeigersinn nennt man auch *mathematisch positiv*.

beschreibt. Für diese Matrix gilt

$$\begin{aligned}
\langle R_\alpha x, R_\alpha y\rangle &= \Big\langle \begin{pmatrix}\cos(\alpha)x_1 - \sin(\alpha)x_2\\ \sin(\alpha)x_1 + \cos(\alpha)x_2\end{pmatrix}, \begin{pmatrix}\cos(\alpha)y_1 - \sin(\alpha)y_2\\ \sin(\alpha)y_1 + \cos(\alpha)y_2\end{pmatrix}\Big\rangle\\
&= \cos(\alpha)^2 x_1y_1 - \cos(\alpha)\sin(\alpha)x_1y_2 - \sin(\alpha)\cos(\alpha)x_2y_1 + \sin(\alpha)^2x_2y_2\\
&\quad + \sin(\alpha)^2x_1y_1 + \cos(\alpha)\sin(\alpha)x_1y_2 + \sin(\alpha)\cos(\alpha)x_2y_1 + \cos(\alpha)^2x_2y_2\\
&= x_1y_1 + x_2y_2 = \langle x, y\rangle
\end{aligned}$$

und folglich gilt auch $\|R_\alpha x\|_2 = \sqrt{\langle R_\alpha x, R_\alpha x\rangle)} = \sqrt{\langle x, x\rangle} = \|x\|_2$. Da wir y durch eine Drehung mit einer Matrix R_α auf die Rechtsachse drehen können schließen wir also für $x, y \in \mathbb{R}^2$

$$\langle x, y\rangle = \langle R_\alpha x, R_\alpha y\rangle = \|x\|_2\|y\|_2\cos(\theta),$$

wobei $\cos(\theta)$ den Winkel zwischen den Vektoren $R_\alpha x$ und $R_\alpha y$ und damit auch zwischen den Vektoren x und y beschreibt.

Da sich zwei Vektoren $x, y \in \mathbb{R}^n$ in einer gemeinsamen Ebene befinden, definieren wir den Winkel θ zwischen diesen Vektoren über

$$\cos(\theta) := \frac{\langle x, y\rangle}{\|x\|_2\|y\|_2}.$$

Die Frage, wann eine Norm von einem Skalarprodukt induziert ist, wird im folgenden behandelt:

Lemma 24.5. 1. *Für jedes Skalarprodukt gilt die Polarisationsformel*

$$\langle x, y\rangle = \tfrac{1}{4}\big(\langle x+y, x+y\rangle - \langle x-y, x-y\rangle\big).$$

2. *Ist $\|\cdot\|$ vom Skalarprodukt $\langle\cdot,\cdot\rangle$ induziert, so gilt die Parallelogrammgleichung*

$$\|x+y\|^2 + \|x-y\|^2 = 2\Big(\|x\|^2 + \|y\|^2\Big).$$

Die Umkehrung (erfüllt $\|\cdot\|$ die Parallelogrammgleichung, so ist über $\langle x, y\rangle := \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2)$ ein Skalarprodukt definiert) ist der Satz von Jordan-von Neumann.

Beweis.

1. Wir rechnen nach (und starten mit dem vierfachen rechten Seite):

$$\begin{aligned}
&\langle x+y, x+y\rangle - \langle x-y, x-y\rangle\\
&\quad = \langle x, x+y\rangle + \langle y, x+y\rangle - \langle x, x-y\rangle + \langle y, x-y\rangle\\
&= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle - \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle - \langle y, y\rangle = 4\langle x, y\rangle.
\end{aligned}$$

2. Auch hier rechnen wir einfach nach:

$$\begin{aligned}
\|x+y\|^2 + \|x-y\|^2 &= \|x\|^2 + 2\langle x, y\rangle + \|y\|^2 + \|x\|^2 - 2\langle x, y\rangle + \|y\|^2\\
&= 2\Big(\|x\|^2 + \|y\|^2\Big).
\end{aligned}$$

□

25 Orthogonale Komplemente

Definition 25.1. Ist V reeller Vektorraum mit Skalarprodukt und $X \subset V$, so ist

$$X^{\perp} := \{v \in V \mid \forall w \in X : \langle w, v\rangle = 0\}$$

das *orthogonale Komplement von* X.

Beispiel.
Für $V = \mathbb{R}^2$ und $X = \{e_1\}$ gilt $X^{\perp} = \{\left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right) \mid \langle \left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right), e_1\rangle = 0\} = \{\left(\begin{smallmatrix} 0 \\ x_2 \end{smallmatrix}\right) \mid x_2 \in \mathbb{R}\}$. △

Satz 25.2. *Es ist* $\langle \cdot, \cdot\rangle$ *ein Skalarprodukt auf* V *und* $X, Y \subset V$, *so gilt:*

i) $X^{\perp}$ *ist ein Unterraum von* V.

ii) Ist $X \subset Y$, *so ist* $Y^{\perp} \subset X^{\perp}$.

iii) $X^{\perp} = \langle X\rangle^{\perp}$.

Achtung: Die spitzen Klammern hier bedeuten den Spann von X!

iv) $X \subset (X^{\perp})^{\perp}$.

v) $X^{\perp} = ((X^{\perp})^{\perp})^{\perp}$.

Beweis.

i) Es seien $v, w \in X^{\perp}$ (d.h. für alle $x \in X$ gilt $\langle v, x\rangle = 0$ und $\langle w, x\rangle = 0$) und $k \in K$. Dann gilt für alle $x \in X$

$$\langle v + kw, x\rangle = \langle v, x\rangle + k\langle w, x\rangle = 0.$$

ii) Da $X \subset Y$ gilt, ist die Bedingung für $v \in Y^{\perp}$ restriktiver als die für $v \in X^{\perp}$; etwas genauer

$$\begin{aligned} Y^{\perp} &= \{v \in V \mid \forall y \in Y \ \langle y, v\rangle = 0\} \\ &\subset \{v \in V \mid \forall x \in X \ \langle x, v\rangle = 0\} = X^{\perp}. \end{aligned}$$

iii) Da $X \subset \langle X\rangle$, folgt $\langle X\rangle^{\perp} \subset X^{\perp}$ aus ii). Ist andersherum $v \in X^{\perp}$, so gilt für jedes $x \in \langle X\rangle$ auch $\langle v, x\rangle = 0$ (da x eine Linearkombination aus Elementen in X ist). Dies zeigt $X^{\perp} \subset \langle X\rangle^{\perp}$.

iv) Es sei $x \in X$. Für jedes $v \in X^{\perp}$ gilt dann $\langle x, v\rangle = 0$. Da Skalarprodukte symmetrisch sind, gilt

$$\langle v, x\rangle = 0 \quad \text{für alle } v \in X^{\perp}$$

aber das bedeutet $x \in (X^{\perp})^{\perp}$.

v) Wenden wir iv) auf $X^{\perp}$ an, so sehen wir $X^{\perp} \subset ((X^{\perp})^{\perp})^{\perp}$. Aus $X \subset (X^{\perp})^{\perp}$ folgt aber mit ii) auch $((X^{\perp})^{\perp})^{\perp} \subset X^{\perp}$, was die Gleichheit zeigt.

□

Satz 25.3. *Ist V n-dimensionaler K-Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$, und U ein Unterraum von V, so gilt*

$$\dim(U) + \dim(U^\perp) = n.$$

Beweis.
Es sei $\{v_1, \dots, v_m\}$ eine Basis von U, welche wir zu einer Basis $\{v_1, \dots, v_m, \dots, v_n\}$ von V ergänzen. Ein Vektor $x = \sum_{i=1}^n k_i v_i$ ist in $U^\perp$, genau dann wenn für $j = 1, \dots, m$ gilt

$$0 = \langle x, v_j\rangle = \sum_{i=1}^n k_i \langle v_i, v_j\rangle,$$

d.h. $U^\perp$ ist genau der Lösungsraum eines homogenes Gleichungssystem bestehend aus m Gleichungen mit n Unbekannten. Die Koeffizienten des Gleichungssystems sind genau m Zeilen des Gramschen Matrix $A = (a_{ij}) = (\langle v_i, v_j\rangle)$. Da $\langle\cdot,\cdot\rangle$ eine nichtausgeartete Bilinearform ist, hat A den Rang n, also hat die Matrix des Gleichungssystems Rang m. Daher ist die Dimension des Lösungsraumes nach der Dimensionsformel (Satz 10.8) gerade $n - m$, was die Behauptung zeigt. □

Korollar 25.4. *Ist U ein Unterraum eines n-dimensionalen $\mathbb{R}$-Vektorraums V, so gilt $U = (U^\perp)^\perp$.*

Zweimaliges Anwenden von Satz 25.3 zeigt $\dim((U^\perp)^\perp) = n - \dim(U^\perp) = n - (n - \dim(U)) = \dim(U)$ und mit $U \subset (U^\perp)^\perp$ folgt die Behauptung.

Korollar 25.5. *$U^\perp$ ist ein komplementärer Unterraum zu U.*

Zu zeigen: $U \cap U^\perp = \{0\}$ und $\langle U \cup U^\perp\rangle = V$.
Für ersteres: Sei $v \in U \cap U^\perp$, d.h. v orthogonal zu sich selbst, d.h. $\langle v, v\rangle = 0$. Da $\langle\cdot,\cdot\rangle$ positiv definit ist, folgt $v = 0$.
Zweiteres folgt aus der Dimensionsformel für Unterräume (Satz 8.12) und der Dimensionsformel aus Satz 25.3, denn $\dim(\langle U \cup U^\perp\rangle) = \dim(U) + \dim(U^\perp) - 0 = \dim(V)$.

Beschäftigen wir uns nun mit speziellen Basen, nämlich solchen, in denen die Vektoren paarweise orthogonal sind:

Definition 25.6. Es sei V ein reeller Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$. Eine Menge von Vektoren $\{v_1, \dots, v_m\}$ heißt *orthogonal*, wenn für $i \neq j$ gilt $\langle v_i, v_j\rangle = 0$, sie heißt *orthonormal*, falls zusätzlich für alle i gilt $\|v_i\| = 1$. Eine orthonormale Basis heißt *Orthonormalbasis*.

Satz 25.7. *Ist $\{v_1, \dots, v_m\}$ eine orthogonale Menge von Vektoren aus $V \setminus \{0\}$, so gilt*

i) Die Menge $\{\frac{v_1}{\|v_1\|}, \dots, \frac{v_m}{\|v_m\|}\}$ ist orthonormal.

ii) Die Menge $\{v_1, \dots, v_m\}$ ist linear unabhängig.

Beweis.

i) Für beliebige Skalierungen $w_i = k_i v_i$ gilt $\langle w_i, w_j\rangle = k_i k_j \langle v_i, v_j\rangle$, also sind die skalierten Vektoren $v_i/\|v_i\|$ wieder orthogonal. Außerdem ist $\|v_i/\|v_i\|\| = \|v_i\|/\|v_i\| = 1$, was die Orthonormalität zeigt.

ii) Es seien $k_1, \dots, k_m \in \mathbb{R}$ und $k_1 v_1 + \cdots k_m v_m = 0$. Bilden wir das Skalarprodukt mit einem v_i, so bekommen wir

$$0 = \langle v_i, k_1 v_1\rangle + \cdots + \langle v_i, k_m v_m\rangle = k_i \langle v_i, v_i\rangle.$$

Da $v_i \neq 0$ gilt, gilt auch $\langle v_i, v_i\rangle = \|v_i\|^2 \neq 0$ und es folgt $k_i = 0$.

□

Satz 25.8 (Orthonormalisierungssatz von Schmidt)**.** *Ist V ein n-dimensionaler reeller Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$ und $M = \{v_1, \dots, v_m\}$ eine orthonormale Menge von Vektoren in V, so lässt sich M zu einer Orthonormalbasis $B = \{v_1, \dots, v_n\}$ ergänzen.*

Beweis.
Wir zeigen: Ist $m < n$, so kann M zu einer orthonormalen Menge mit $m+1$ Elementen erweitert werden, denn daraus folgt die Behauptung.

Ist nun $\{v_1, \dots, v_m\}$ bereits orthonormal, aber $m < n$. Dann gibt es einen Vektor w_{m+1}, der nicht im Spann der $v_1, \dots, v_m$ liegt. Dieser ist aber im allgemeinen nicht zu allen v_i orthogonal. Wir ändern nun w_{m+1} mit folgendem Ansatz ab:

$$\tilde{w}_{m+1} = w_{m+1} + k_1 v_1 + \dots k_m v_m$$

und versuchen, die Koeffizienten $k_1, \dots, k_m$ passend zu bestimmen. Aus der Forderung der Orthogonalität $\langle \tilde{w}_{m+1}, v_i\rangle = 0$ und der Orthonormalität der $v_1, \dots, v_m$ bekommen wir

$$\begin{aligned} 0 = \langle \tilde{w}_{m+1}, v_i\rangle &= \langle w_{m+1}, v_i\rangle + k_1\langle v_1, v_i\rangle + \cdots + k_m\langle v_m, v_i\rangle \\ &= \langle w_{m+1}, v_i\rangle + k_i. \end{aligned}$$

Setzen wir also $k_i = -\langle w_{m+1}, v_i\rangle$, ist

$$\tilde{w}_{m+1} = w_{m+1} - \langle w_{m+1}, v_1\rangle v_1 - \cdots - \langle w_{m+1}, v_m\rangle v_m$$

ein Vektor, der zu allen $v_1, \dots, v_m$ orthogonal ist. Durch Normierung

$$v_{m+1} = \tfrac{\tilde{w}_{m+1}}{\|\tilde{w}_{m+1}\|}$$

erhalten wir die geforderte orthonormale Erweiterung. □

Beispiel.
Wir betrachten die Vektoren

$$w_1 = \begin{pmatrix}1\\1\\1\end{pmatrix}, w_2 = \begin{pmatrix}1\\1\\0\end{pmatrix}, w_3 = \begin{pmatrix}1\\0\\0\end{pmatrix}$$

welche bezüglich des Standardskalarproduktes auf $\mathbb{R}^3$ nicht orthonormal sind.

Wir benutzen die Methode aus dem Beweis des Orthonormalisierungssatzes:

1. Normiere w_1 zu

$$v_1 = \frac{w_1}{\|w_1\|} = \frac{1}{\sqrt{3}} w_1 = \begin{pmatrix}\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{3}}\end{pmatrix}$$

2. Nimm w_2 und bilde

$$\begin{aligned}\tilde{w}_2 &= w_2 - \langle w_2, v_1\rangle v_1\\ &= \begin{pmatrix}1\\1\\0\end{pmatrix} - \langle \begin{pmatrix}1\\1\\0\end{pmatrix}, \tfrac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}\rangle \tfrac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}\\ &= \begin{pmatrix}1\\1\\0\end{pmatrix} - \tfrac{2}{\sqrt{3}}\tfrac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}\\ &= \begin{pmatrix}\frac{1}{3}\\\frac{1}{3}\\-\frac{2}{3}\end{pmatrix}.\end{aligned}$$

3. Normiere $\tilde{w}_2$:

$$v_2 = \frac{\tilde{w}_2}{\|\tilde{w}_2\|} = \frac{1}{\sqrt{\frac{1}{9}+\frac{1}{9}+\frac{4}{9}}}\begin{pmatrix}\frac{1}{3}\\\frac{1}{3}\\-\frac{2}{3}\end{pmatrix} = \frac{3}{\sqrt{6}}\begin{pmatrix}\frac{1}{3}\\\frac{1}{3}\\-\frac{2}{3}\end{pmatrix} = \begin{pmatrix}\frac{1}{\sqrt{6}}\\\frac{1}{\sqrt{6}}\\-\frac{2}{\sqrt{6}}\end{pmatrix}.$$

4. Nimm w_3 und bilde

$$\begin{aligned}\tilde{w}_3 &= w_3 - \langle w_3, v_1\rangle v_1 - \langle w_3, v_2\rangle v_2\\ &= \begin{pmatrix}1\\0\\0\end{pmatrix} - \langle \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{3}}\end{pmatrix}\rangle \begin{pmatrix}\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{3}}\end{pmatrix} - \langle \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}\frac{1}{\sqrt{6}}\\\frac{1}{\sqrt{6}}\\-\frac{2}{\sqrt{6}}\end{pmatrix}\rangle \begin{pmatrix}\frac{1}{\sqrt{6}}\\\frac{1}{\sqrt{6}}\\-\frac{2}{\sqrt{6}}\end{pmatrix}\\ &= \begin{pmatrix}1\\0\\0\end{pmatrix} - \tfrac{1}{\sqrt{3}}\begin{pmatrix}\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{3}}\\\frac{1}{\sqrt{3}}\end{pmatrix} - \tfrac{1}{\sqrt{6}}\begin{pmatrix}\frac{1}{\sqrt{6}}\\\frac{1}{\sqrt{6}}\\-\frac{2}{\sqrt{6}}\end{pmatrix}\\ &= \begin{pmatrix}1\\0\\0\end{pmatrix} - \begin{pmatrix}\frac{1}{3}\\\frac{1}{3}\\\frac{1}{3}\end{pmatrix} - \begin{pmatrix}\frac{1}{6}\\\frac{1}{6}\\-\frac{2}{6}\end{pmatrix} = \begin{pmatrix}\frac{1}{2}\\-\frac{1}{2}\\0,\end{pmatrix}\end{aligned}$$

5. Normiere $\tilde{w}_3$:

$$v_3 = \frac{\tilde{w}_3}{\|\tilde{w}_3\|} = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{pmatrix}.$$

Dieses iterative Verfahren zur Produktion einer Orthonormalbasis nennt sich *Gram-Schmidtsches Orthonormalisierungsverfahren.*. △

26 Selbstadjungierte Abbildungen

Definition 26.1. Es sei V ein komplexer Vektorraum. Eine Abbildung $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{C}$ heißt *Skalarprodukt*, wenn folgende Eigenschaften gelten: Für alle $u, v, w \in V$ und $\alpha \in \mathbb{C}$ gilt

$$\begin{aligned} \langle u+v, w\rangle &= \langle u, w\rangle + \langle v, w\rangle \\ \langle \alpha v, w\rangle &= \alpha\langle v, w\rangle \end{aligned} \qquad \text{(Linear in 1. Komponente)}$$

$$\begin{aligned} \langle u, v+w\rangle &= \langle u, v\rangle + \langle u, w\rangle \\ \langle v, \alpha w\rangle &= \bar{\alpha}\langle v, w\rangle \end{aligned} \qquad \text{(Semilinear in 2. Komponente)}$$

$$\langle v, w\rangle = \overline{\langle w, v\rangle} \qquad \text{(Antisymmetrie)}$$

$$v \neq 0 \implies \langle v, v\rangle > 0 \qquad \text{(positive Definitheit)}$$

Erinnerung: Für $\alpha = x + iy$ mit $x, y \in \mathbb{R}$ ist $\bar{\alpha} = x - iy$ das komplex konjugierte.
Beachte dass die Antisymmetrie impliziert, dass für alle v gilt $\langle v, v\rangle = \overline{\langle v, v\rangle}$ und daraus folgt $\langle v, v\rangle \in \mathbb{R}$.
Achtung: Häufig wird bei komplexen Skalarprodukten auch Semilinearität in der ersten und Linearität in der zweiten Komponente gefordert.

Man rechnet schnell nach: Auf $\mathbb{C}^n$ ist durch

$$\langle x, y\rangle = \sum_{k=1}^{n} x_k \overline{y_k}$$

ein Skalarprodukt gegeben, es heißt, das *unitäre Skalarprodukt*. Auch dieses lässt sich durch Matrixmultiplikation beschreiben. Dafür führen wir ein: Ist $A \in \mathbb{C}^{m\times n}$ so ist die *konjuguert-transponierte* oder *adjungierte* Matrix dazu $A^H \in \mathbb{C}^{n\times m}$ definiert durch

$$A^H = \overline{A}^T, \quad \text{d.h.} \quad (A^H)_{ij} = \overline{A_{ji}}.$$

Damit gilt für das unitäre Skalarprodukt

$$\langle x, y\rangle = x^T\overline{y} = y^H x.$$

Völlig analog zum reellen Fall zeigt man die Cauchy-Schwarz-Ungleichung für den komplexen Fall und, dass jedes Skalarprodukt auf einem komplexen Vektorraum via $\|x\| = \sqrt{\langle x, x\rangle}$ eine Norm induziert. Das unitäre Skalarprodukt auf $\mathbb{C}^n$ induziert die *unitäre Norm*

$$\|x\| = \sqrt{\langle x, x\rangle} = \left(\sum_{k=1}^{n} x_k\overline{x_k}\right)^{1/2} = \left(\sum_{k=1}^{n} |x_k|^2\right)^{1/2}.$$

Weiterhin sind die Begriffe *orthogonal*, *orthogonales Komplement* und *Orthonormalbasis* völlig analog zum reellen Fall definiert.

Definition 26.2. Es sei V ein reeller oder komplexer Vektorraum mit Skalarprodukt $\langle\cdot,\cdot\rangle$. Ein linear Abbildung $f : V \to V$ heißt *selbstadjungiert*, wenn für alle $v, w \in V$ gilt

$$\langle f(v), w\rangle = \langle v, f(w)\rangle.$$

Satz 26.3. *Ist $V = \mathbb{R}^n$ bzw . $= \mathbb{C}^n$ mit dem Standard- bzw. unitärem Skalarprodukt, so gilt: Die lineare Abbildung die zur Matrix $A \in \mathbb{R}^{n\times n}$ bzw. $\in \mathbb{C}^{n\times n}$ ist selbstadjungiert genau dann, wenn im reellen Fall $A = A^T$ gilt und im komplexen Fall $A = A^H$ gilt.*

Beweis.
Es gilt im komplexen Fall

$$\langle x, Ay\rangle = \sum_{k=1}^{n} x_k \overline{\sum_{j=1}^{n} a_{kj} y_j} = \sum_{j,k=1}^{n} x_k \overline{a_{kj}} \overline{y_j} = \sum_{j=1}^{n}\sum_{k=1}^{n} \underbrace{\overline{a_{kj}} x_k}_{(A^H x)_j} \overline{y_j} = \langle A^H x, y\rangle,$$

und der reelle Fall ist analog. □

Eine Matrix $A \in \mathbb{C}^{n\times n}$ mit $A^H = A$ heißt *hermitesch*. Für eine hermitesche Matrix A gilt insbesondere $a_{kk} = \overline{a_{kk}}$, d.h. Einträge auf der Diagonalen sind reell.

Lemma 26.4. *Ist f eine selbstadjungierte Abbildung auf einem reellen oder komplexen Vektorraum V, so sind alle Eigenwerte von f reell, insbesondere hat eine hermitesche Matrix nur reelle Eigenwerte.*

Beweis.
Ist λ ein Eigenwert von f, so existiert ein $v \in V$ mit $f(v) = \lambda v$. Es folgt

$$\lambda\langle v, v\rangle = \langle \lambda v, v\rangle = \langle f(v), v\rangle = \langle v, f(v)\rangle = \langle v, \lambda v\rangle = \overline{\lambda}\langle v, v\rangle.$$

Es folgt $\lambda = \overline{\lambda}$, also $\lambda \in \mathbb{R}$. □

Satz 26.5 (Spektralsatz für selbstadjungierte Abbildungen)**.** *Ist f eine selbstadjungierte Abbildung auf einem n-dimensionalen reellen oder komplexen Vektorraum, so gibt es eine Orthonormalbasis aus Eigenvektoren von f. Insbesondere haben hermitesche komplexe Matrizen und symmetrische reelle Matrizen eine Orthonormalbasis aus Eigenvektoren.*

Der Beweis benutzt den *Fundamentalsatz der Algebra*. Dieser besag, dass jedes nicht-konstante Polynom mit komplexen Koeffizienten eine komplexe Nullstelle hat. (Es folgt daraus, dass Polynom $p(z) = \sum_{k=0}^{n} a_k z^k$ vom Grad n in Linearfaktoren zerfällt, d.h. es gibt Nullstellen $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ (nicht notwendigerweise verschieden), so dass $p(z) = a_n \prod_{l=1}^{n}(z - \lambda_l)$.)
Diesen Satz beweisen wir hier nicht - typischerweise wird er im Rahmen der Funktionentheorie bewiesen.

Beweis.
Behandeln wir zuerst den komplexen Fall: Nach dem Fundamentalsatz der Algebra zerfällt das charakteristische Polynom von f in Linearfaktoren, d.h. es gibt $\lambda_1, \ldots, \lambda_n$ mit

$$\chi_f(\lambda) = \pm(\lambda - \lambda_1)\cdots(\lambda - \lambda_n)$$

und nach Lemma 26.4 gilt $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$.

Zeigen wir nun induktiv, dass es eine Orthonormalbasis aus Eigenvektoren von f gibt: Zum Eigenwert λ_1 gibt es einen Eigenvektor v_1, und wir können annehmen, dass $\|v_1\| = 1$ gilt.

Wir setzen

$$W = \{v_1\}^\perp = \{w \in V \mid \langle w, v_1\rangle = 0\}.$$

Es gilt $\dim(W) = n - 1$ und außerdem ist für $w \in W$

$$\langle v_1, f(w)\rangle = \langle f(v_1), w\rangle = \langle \lambda_1 v_1, w\rangle = \lambda_1\langle v_1, w\rangle = 0,$$

d.h. $f(W) \subset W$. Wir können also die Einschränkung $f|_W$ von f auf W betrachten und bekommen mit $f|_W$ eine selbstadjungierte Abbildung auf dem $n-1$-dimensionalen Vektorraum W. Nach

Induktionsannahme hat W eine Orthonormalbasis aus Eigenvektoren von $f|_W$ welche durch v_1 zu einer Orthonormalbasis von V aus Eigenvektoren von f ergänzt wird.

Der reelle Fall geht über den komplexen Fall: Die Darstellungsmatrix von f bezüglich einer Orthonormalbasis ist reell symmetrisch, also auch komplex hermitesch. Also existiert dazu eine Orthormalbasis (des Koeffizientenraumen) welche mit Hilfe der Orthormalbasis, bezüglich derer dargestellt wurde, eine Orthonormalbasis aus Eigenvektoren in V liefert. $\square$

Lemma 26.6. *Ist $v_1, \dots, v_n$ eine Orthonormalbasis in $\mathbb{R}^n$ oder $\mathbb{C}^n$, und ist $A = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix} \in \mathbb{R}^{n\times n}$ bzw. $\in \mathbb{C}^{n\times n}$, so gilt $A^T A = E_n$ bzw $A^H A \in E_n$.*

Insbesondere gilt: Die Spalten von A bilden eine Orthonormalbasis genau dann, wenn $A^{-1} = A^T$ bzw. $A^{-1} = A^H$.

Beweis.

Es gilt $(A^T A)_{ij} = \langle v_i, v_j\rangle$ (im reellen) und $(A^H A)_{ij} = \langle v_i, v_j\rangle$ (im komplexen). In beiden Fällen folgt die Behauptung, da $\langle v_i, v_j\rangle = \delta_{ij}$. $\square$

Definition 26.7. Eine $n \times n$ Matrix deren Spalten paarweise orthonormal sind heißt *orthonormale Matrix.*

Es gilt übrigens: Ist A orthonormal, so ist A^T (im reellen Fall) bzw. A^H (im komplexen Fall) orthonormal. Oder anders gesagt: Sind die Spalten einer Matrix orthogonal, so sind es auch die Zeilen.

Nur der komplexe Fall: Es gilt $A^H = A^{-1}$, d.h. $E_n = A^{-1}A = A^H A$ und ebenso $E_n = AA^{-1} = AA^H$. Aus letzterem folgt, dass A^H eine orthonormale Matrix ist.

Korollar 26.8. *Ist A eine reell-symmetrische bzw. komplex-hermitesche Matrix, so existieren eine orthogonale Matrix S und eine reelle Diagonalmatrix D so dass $S^T AS = D$ (im reellen Fall) bzw. $S^H AS = D$ (im komplexen Fall).*

Man schreibe die Eigenvektoren in die Spalten von S; dann gilt $S^{-1}AS = D$ wobei D die Eigenwerte auf der Diagonalen hat. Bemerke schließlich, dass $S^{-1} = S^T$ (bzw. $= S^H$)

Beispiel.

Wir betrachte $A \in \mathbb{R}^{2\times 2}$ gegeben durch

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

mit charakteristischem Polynom

$$\chi_A(\lambda) = \det(\begin{pmatrix} 2-\lambda & -1 \\ -1 & 2-\lambda \end{pmatrix}) = (2-\lambda)^2 - 1 = \lambda^2 - 4\lambda + 3.$$

Also sind die Eigenwerte

$$\lambda_{1/2} = 2 \pm \sqrt{4-3} = \begin{cases} 3 \\ 1 \end{cases},$$

(und wie erwartet, beide reell). Wir suchen nun normierte Eigenvektoren zu $\lambda_1 = 1$:

$$\mathrm{Kern}(A - E_2) = \mathrm{Kern}(\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}) = \langle \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rangle$$

Ein normierter Eigenvektor ist also z.B.

$$v_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}.$$

Zum Eigenwert $\lambda_2 = 3$ ergibt sich

$$\mathrm{Kern}(A - 3E_2) = \mathrm{Kern}(\begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} = \langle \begin{pmatrix} 1 \\ -1 \end{pmatrix} \rangle.$$

Hier ist ein normierte Eigenvektor zum Beispiel

$$v_2 = \begin{pmatrix} 1/\sqrt{2} \\ -1\sqrt{2}. \end{pmatrix}$$

Also lässt sich A durch die (orthonormale) Matrix $S = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ diagonalisieren. Es gilt $S^{-1} = S^T$ und

$$\begin{aligned} S^T A S &= \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \tfrac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \\ &= \tfrac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 1 & -3 \end{pmatrix} \\ &= \tfrac{1}{2} \begin{pmatrix} 2 & 0 \\ 0 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \end{aligned}$$

und auf der Diagonale stehen, wie erwartet, die Eigenwerte. △

Beispiel.

Für die komplex-hermitesche Matrix

$$A = \begin{pmatrix} -1 & 1+i \\ 1-i & 1 \end{pmatrix}$$

ergibt sich das charakteristische Polynom

$$\begin{aligned} \chi_A(\lambda) &= \det(\begin{pmatrix} -1-\lambda & 1+i \\ 1-i & 1-\lambda \end{pmatrix} \\ &= (-1-\lambda)(1-\lambda) - (1+i)(1-i) = \lambda^2 - 3. \end{aligned}$$

Also sind die beiden Eigenwerte $\lambda_{1/2} = \pm\sqrt{3}$. Als Eigenvektoren ergeben sich z.B.

$$w_1 = \begin{pmatrix} \frac{1+i}{1+\sqrt{3}} \\ 1 \end{pmatrix}, \quad w_2 = \begin{pmatrix} \frac{1+i}{-1+\sqrt{3}} \\ 1 \end{pmatrix}$$

(wir könnten hier noch mit einer beliebigen komplexen Zahl multiplizieren und würden ebenfalls Eigenvektoren erhalten). Das Normieren dieser Vektoren führt auf ziemlich unübersichtliche Ausdrücke. Zum Beispiel ergibt sich als normierter Eigenvektor zum Eigenwert $\lambda_1 = \sqrt{3}$ der Vektor

$$v_1 = \begin{pmatrix} \frac{1+i}{\sqrt{6+2\sqrt{3}}} \\ \sqrt{\frac{2+\sqrt{3}}{3+\sqrt{3}}} \end{pmatrix}.$$

$\triangle$

27 Hauptachsentransformation

Der Spektralsatz für selbstadjungierte Abbildungen hat Konsequenzen für die Darstellung von symmetrischen Bilinearformen:

Satz 27.1. *Es sei $A \in \mathbb{R}^{n\times n}$ symmetrisch und B die Bilinearform definiert durch $B(x,y) = \langle Ax, y\rangle$ (mit dem Standardskalarprodukt). Dann gibt es eine Orthonormalbasis $\mathcal{B} = \{w_1, \dots, w_n\}$ des $\mathbb{R}^n$ bestehend aus Eigenvektoren von A, so dass die Gramsche Matrix von B bezüglich der Basis $\mathcal{B}$ eine Diagonalmatrix* $\operatorname{diag}(\lambda_1, \dots, \lambda_n)$ *ist, d.h. $B(w_i, w_j) = \lambda_i\delta_{ij}$, wobei $\lambda_1, \dots, \lambda_n$ die Eigenwerte von A sind.*

Beweis.
Da A symmetrisch ist, ist A nach Satz 26.5 orthonormal diagonalisierbar, d.h. es gibt eine orthonormale Basis $\mathcal{B} = \{w_1, \dots, w_n\}$, so dass $Aw_i = \lambda_i w_i$. Die Gramsche Matrix von B bezüglich dieser Basis $\mathcal{B}$ ist also

$$(A_{\mathcal{B}})_{ij} = B(w_i, w_j) = \langle Aw_i, w_j\rangle = \langle \lambda_i w_i, w_j\rangle = \lambda_i\delta_{ij}$$

also $A_{\mathcal{B}} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, wie behauptet. □

Definition 27.2. Ist B eine Bilinearform auf $\mathbb{R}^n$ definiert durch $B(x,y) = \langle Ax, y\rangle$ (mit Standardskalarprodukt), dann ist $q : \mathbb{R}^n \to \mathbb{R}$ definiert durch $q(x) = B(x,x)$ die zugehörige *quadratische Form.*

Wir bemerken, dass für jede quadratische Form gilt

$$q(x) = \langle Ax, x\rangle = \langle x, A^T x\rangle = \langle A^T x, x\rangle$$

und es folgt also

$$q(x) = \langle \tfrac{1}{2}(A + A^T)x, x\rangle.$$

Die quadratische Form hängt also nicht wirklich von A, sondern von der symmetrischen Matrix $\frac{1}{2}(A + A^T)$ ab. Wir können bei quadratischen Formen also ohne Einschränkung annehmen, dass die zugehörige Matrix A symmetrisch ist.

Wir halten noch fest:

Lemma 27.3. *Jede Matrix $A \in \mathbb{R}^{n\times n}$ lässt sich als Summe einer symmetrischen Matrix A_S und einer schiefsymmetischen Matrix A_A schreiben.*

Eine Matrix $M \in \mathbb{R}^{n\times n}$ heißt *schiefsymmetrisch*, wenn $M^T = -M$ gilt.

Beweis.
Wir setzen $A_S = \frac{1}{2}(A + A^T)$ und $A_A = \frac{1}{2}(A - A^T)$. Dann ist $A_S^T = A_S$, $A_A^T = \frac{1}{2}(A^T - A) = -A_A$ und $A_S + A_A = A$. □

Beispiel.
Für $A = \begin{pmatrix} \alpha & \beta \\ \beta & \gamma \end{pmatrix}$ ist die zugehörige Bilinearform

$$B(x,y) = \alpha x_1 y_1 + \beta x_1 y_2 + \beta x_2 y_1 + \gamma x_2 y_2.$$

Die quadratische Form ist also

$$q_A(x) = \alpha x_1^2 + 2\beta x_1 x_2 + \gamma x_2^2.$$

Sei umgekehrt $p(x_1, x_2) = ax_1^2 + bx_1x_2 + cx_2^2 + dx_1 + ex_2 + f$ ein quadratisches Polynom in zwei Variablen. Dann gilt

$$p(x_1, x_2) = x^T \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} x + \begin{pmatrix} d & e \end{pmatrix} x + f.$$

Wir untersuchen das Polynom nur bis auf den linearen und konstanten Term, und untersuchen die Eigenwerte der Matrix. Diese hat die beiden Eigenwerte

$$\lambda_{1/2} = \tfrac{1}{2}(a + c \pm \sqrt{(a-c)^2 + b^2}).$$

Die beiden zugehörigen Eigenvektoren w_1 und w_2 rechnen wir nicht aus, doch wir wissen, dass sie orthonormal gewählt werden können. Wir setzen $S = \begin{pmatrix} w_1 & w_2 \end{pmatrix}$ und $z = S^{-1}x = S^T x$, also gilt $x = Sz$. Damit ist

$$q(x) = x^T \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} x = z^T S^T A S z = z^T S^{-1} A S z$$

und wegen $S^{-1}AS = \operatorname{diag}(\lambda_1, \lambda_2)$, folgt

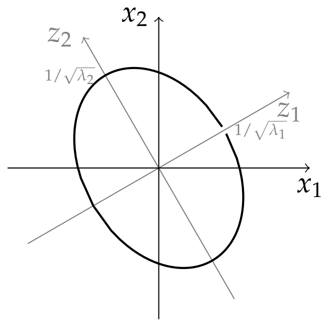
$$q(x) = z^T \operatorname{diag}(\lambda_1, \lambda_2) z = \lambda_1 z_1^2 + \lambda_2 z^2.$$

Wir interpretieren die Situation geometrisch, indem wir die Menge $C = \{x \in \mathbb{R}^2 \mid q(x) = 1\}$ betrachten. Diese lautet $ax_1^2 + bx_1x_2 + cx_2^2 - 1 = 0$ und es ist auf den ersten Blick nicht klar, welche geometrische Form sie hat. In den neuen Koordinaten z wird diese einfach zu $\lambda_1 z_1^2 + \lambda_2 z^2 - 1 = 0$. Formales Auflösen nach z_2 gibt

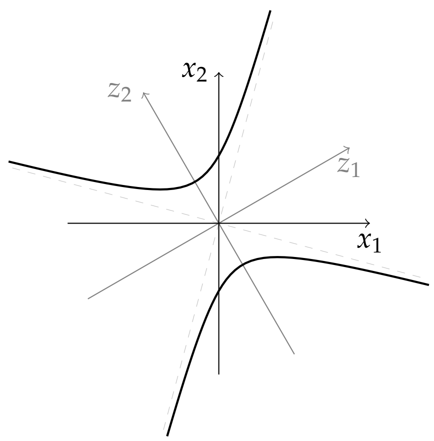
$$z_2 = \pm\sqrt{\tfrac{1}{\lambda_2} - \tfrac{\lambda_1}{\lambda_2} z_1^2}.$$

Die Form der Lösungsmenge hängt von den Vorzeichen von $1/\lambda_2$ und λ_1/λ_2 ab: Es gilt:

1. $\lambda_1, \lambda_2 > 0$: Hier ist der Ausdruck unter der Wurzel nur für $|z_1| \leq 1/\sqrt{\lambda_1}$ nicht negativ und C ist eine Ellipse:



2. $\lambda_1 < 0, \lambda_2 > 0$: Hier ist der Ausdruck unter der Wurzel immer positiv und C ist eine Hyperbel



(Im Fall $\lambda_1 > 0$ und $\lambda_2 < 0$ ist die Hyperbel in die andere Richtung offen.)

3. $\lambda_1 > 0, \lambda_2 = 0$: C besteht aus den beiden Geraden $z_1 = \pm\sqrt{\lambda_1}^{-1}$

4. $\lambda_1, \lambda_2 < 0$: Hier ist C leer.

$\triangle$

28 Matrixgruppen

Satz 28.1 (und Definition). *Es sei K ein Körper und die Menge aller invertierbaren $n \times n$ Matrizen sei*

$$GL_n(K) = \{A \in K^{n\times n} \mid A \text{ invertierbar}\}.$$

Eine Gruppe ist eine Menge G mit einer assoziativen Verknüpfung $\circ$, so dass es ein neutrales Element $e \in G$ gibt und jedes Element $g \in G$ ein inverses Element g^{-1} hat, für das gilt $g \circ g^{-1} = g^{-1} \circ g = e$ hat.

Dann ist $GL_n(K)$ versehen mit der Matrixmultiplikation eine Gruppe, genannt allgemeine lineare Gruppe *der Ordnung n über dem Körper K.*

Beweis.
Sind $A, B \in GL_n(K)$, so ist $AB \in GL_n(K)$ (die Dimension passt und wegen $(AB)^{-1} = B^{-1}A^{-1}$ ist AB invertierbar). Assoziativität der Multiplikation folgt aus der Assozivität der Verknüpfung von Abbildungen (indem wir Matrizen als lineare Abbildungen auffassen). Das neutrale Element ist die Einheitsmatrix E_n und inverse Elemente existieren per Definition. □

Im Fall $K = \mathbb{R}$ bzw. $K = \mathbb{C}$ hat $GL_n(K)$ einige wichtige Untergruppen:

Ist G eine Gruppe, dann ist $H \subset G$ eine Untergruppe, wenn H abgeschlossen bezüglich der Gruppenverknüpfung ist.

Definition 28.2. Wir bezeichnen

$$\begin{aligned} O(n) &:= \{A \in GL_n(\mathbb{R}) \mid A^{-1} = A^T\} && \text{(orthogonale Gruppe)} \\ SO(n) &:= \{A \in O(n) \mid \det(A) = 1\} \\ & && \text{(spezielle orthogonale Gruppe)} \\ U(n) &:= \{A \in GL_n(\mathbb{C}) \mid A^{-1} = A^H\} && \text{(unitäre Gruppe)} \end{aligned}$$

Dass es sich tatsächlich im Untergruppen handelt, ist in allen Fällen schnell nachzurechnen: Sind $A, B \in O(n)$, dann ist $(AB)^{-1} = B^{-1}A^{-1} = B^TA^T = (AB)^T$ (der Nachweis für $U(n)$ geht genau so, nur mit ^{H} statt ^{T}). Sind $A, B \in SO(n)$, dann ist $\det(AB) = \det(A)\det(B) = 1$.

Für $n = 2$ lassen sich $O(2)$ und $SO(2)$ einfach beschreiben:

Lemma 28.3. *Ist $A \in O(2)$, so gibt es ein $\alpha \in [0, 2\pi[$, so dass*

$$A = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \quad \textit{oder} \quad A = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix}.$$

Im ersten Fall ist $A \in SO(2)$ und beschreibt eine Drehung, im zweiten Fall ist $\det(A) = -1$ und beschreibt eine Spiegelung.

Beweis.
Ist $A \in O(2)$, muss $A^TA = E_2$ sein. Ist $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$, so bedeutet das

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & c \\ b & d \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{pmatrix}.$$

Das heißt, es müssen die drei Gleichungen

$$i)\ a^2 + c^2 = 1, \quad ii)\ ab + cd = 0 \quad iii)\ b^2 + d^2 = 1$$

erfüllt sein. Wegen $i)$ und $iii)$ gibt es $\alpha, \alpha' \in [0, 2\pi[$, mit $a = \cos(\alpha)$, $c = \sin(\alpha)$, $b = \sin(\alpha')$, $d = \cos(\alpha')$. Nach $ii)$ und dem Additionstheorem für den Sinus gilt $0 = \cos(\alpha)\sin(\alpha') + \sin(\alpha)\cos(\alpha') = \sin(\alpha + \alpha')$ und daher ist $\alpha - \alpha'$ ein ganzzahliges Vielfaches von π. Ist $\alpha - \alpha'$ ein gradzahliges Vielfaches von π, so gilt $b = \sin(\alpha') = -\sin(\alpha)$ und $d = \cos(\alpha') = \cos(\alpha)$ und ist $\alpha - \alpha'$ und ungradzahliges Vielfaches, so folgt $b = \sin(\alpha') = \sin(\alpha)$ und $d = \cos(\alpha') = -\cos(\alpha)$.

Die Determinanten rechnet man einfach aus, die Geometrie von Drehungen haben wir uns in Abschnitt 24 klar gemacht und dass im zweiten Fall eine Spiegelung an der Geraden mit Winkel $\alpha/2$ beschrieben wird kann man sich z.B. mit Hilfe von Eigenvektoren und Eigenwerten klar machen. $\square$

Für $n > 2$ wird die Beschreibung von $O(n)$ und $SO(n)$ etwas komplizierter.