Brückenkurs Lineare Algebra

29. Juli 2021

Inhaltsverzeichnis

Vorwort

Dieser Brückenkurs soll eine Einführung in die Lineare Algebra sein, die ausreicht, um mehrdimensionale Analysis zu verstehen. Es werden also alle dafür nötigen Konzepte, Zusammenhänge und Rechenmethoden eingeführt. Die Vorlesungen Lineare Algebra I und II kann dieser Kurs nicht ersetzen.
Im Wesentlichen soll dieses Skript eine Einführung zum Selbststudium sein, die auch im Laufe der Analysis II immer wieder konsultiert werden kann.

1. Matrizen

Matrizen sind wichtige Rechenobjekte in der linearen Algebra. Eigentlich immer, wenn man irgendetwas konkret ausrechnen möchte, benutzt man dazu Matrizen.

1.1. Definition und Rechenregeln

Erst einmal schauen wir uns an, was Matrizen eigentlich sind und was man damit machen kann.

Definition 1.1. Eine $(m \times n)$ - *Matrix* ist ein rechteckiges Schema

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \vdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

wobei die *Einträge* a_{ij} aus einer beliebigen Menge X stammen. Für uns in Analysis sind aber vor allem reelle Matrizen interessant (also $a_{ij} \in \mathbb{R}$) [1].

Notation $A = (a_{ij})_{1 \leq i \leq m,\ 1 \leq j \leq n}$ oder $A = (a_{ij})_{ij}$ oder noch kürzer $A = (a_{ij})$.

$\begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix}$ heißt i-te *Zeile* von A, $\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$ heißt j-te *Spalte* von A. Die Zeilen und Spalten einer Matrix A bilden Vektoren, welche als i-ter Zeilen- bzw. j-ter Spaltenvektor bezeichnet werden.
Die Menge aller $(m \times n)$-Matrizen mit Einträgen in einer Menge X bezeichnen wir mit $\mathrm{Mat}_X(m \times n)$, die Mengen $\mathrm{Mat}_{\mathbb{R}}(m \times n)$ nennen wir auch *reelle Matrizen*.

Im Folgenden beschränken wir uns auf reelle Matrizen (für komplexe Matrizen geht alles analog und für Matrizen, deren Einträge reell- oder komplexwertige Funktionen sind auch).

Definition 1.2. Gegeben seien zwei reelle $(m \times n)$ - Matrizen $A = (a_{ij})$ und $B = (b_{ij})$. Dann definieren wir ihre *Summe* als eintragsweise Summe, also

$$A + B := (a_{ij} + b_{ij}).$$

Für $\lambda \in \mathbb{R}$ definieren wir eine *skalare Multiplikation* durch

$$\lambda A := (\lambda a_{ij}).$$

[1] Ebenso treten Matrizen auf, deren Einträge reellwertigen Funktionen sind (also $a_{ij} : \mathbb{R}^k \to \mathbb{R}$ für ein $k \in \mathbb{N}_{>0}$). Aber dafür reicht es sich immer vorzustellen, dass man die Funktionen an einem $x \in \mathbb{R}^k$ auswertet und dann wie mit reellen Matrizen rechnet.

Man sieht leicht, dass sich verschiedene Eigenschaften der Addition und skalaren Multiplikation auf Matrizen übertragen, z.B. Kommutativität und Assoziativität der Addition und Distributivgesetze. Auch haben wir ein neutrales Elements bzgl. der Addition – die *Nullmatrix*, d.h. alle Einträge sind Null. Damit wird $\mathrm{Mat}_{\mathbb{R}}(m \times n)$ mit den so definierten Operationen der Addition und der skalaren Multiplikation ein $\mathbb{R}$-Vektorraum der Dimension mn sein - genaue Definition kommt später.

Vektoren im $\mathbb{R}^n$ schreiben wir in einer Spalte und betrachten diese als $(n \times 1)$ - Matrix. Dann ist die Addition und skalare Multiplikation von diesen Matrizen einfach die Vektoraddition und Vektorstreckung.

Man kann Matrizen auch multiplizieren, solange der erste Faktor so viele Spalten hat wie der zweite Zeilen hat:

Definition 1.3. Sei $A = (a_{ij})$ eine $(\ell \times m)$ - Matrix und $B = (b_{ij})$ eine $(m \times n)$ - Matrix. Dann definieren wir das *Matrixprodukt* $AB := C = (c_{ij})$ als die $(\ell \times n)$ - Matrix, deren Einträge berechnet werden durch

$$c_{ij} := \sum_{k=1}^{m} a_{ik} b_{kj}.$$

Die Formel sagt, dass wir den ij-ten Eintrag als Skalarprodukt [1] des i-ten Zeilenvektors mit dem j-ten Spaltenvektor erhalten:

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ \boxed{a_{i1}} & \boxed{\cdots} & \boxed{a_{in}} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & \cdots & \boxed{b_{1j}} & \cdots & b_{1n} \\ \vdots & & \boxed{\vdots} & & \vdots \\ b_{n1} & \cdots & \boxed{b_{nj}} & \cdots & b_{nn} \end{pmatrix} = \begin{pmatrix} c_{11} & \cdots & \cdots & \cdots & c_{1n} \\ \vdots & & \vdots & & \vdots \\ \vdots & \cdots & \boxed{c_{ij}} & \cdots & \vdots \\ \vdots & & \vdots & & \vdots \\ c_{n1} & \cdots & \cdots & \cdots & c_{nn} \end{pmatrix}$$

Beispiel 1.4. (a) Durch

$$\mathrm{Id}_n = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & & & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix} \tag{1.1}$$

wird die *Einheitsmatrix* in $\mathrm{Mat}_{n \times n}$ definiert.

[1] Zur Definition des Skalarproduktes in $\mathbb{R}^n$ siehe Beispiel 2.7

(b) Durch

$$\operatorname{diag}(a_{11}, a_{22}, \ldots, a_{nn}) := \begin{pmatrix} a_{11} & 0 & 0 & \ldots & 0 & 0 \\ 0 & a_{22} & 0 & \ldots & 0 & 0 \\ 0 & 0 & a_{33} & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & & & \vdots \\ 0 & 0 & 0 & \ldots & a_{(n-1)(n-1)} & 0 \\ 0 & 0 & 0 & \ldots & 0 & a_{nn} \end{pmatrix},$$

wird die Matrix mit den Einträgen $a_{11}, a_{22}, \ldots, a_{nn}$ auf der Diagonalen und Nullen auf allen anderen Plätzen definiert. Matrizen von diesem Typ heißen *Diagonalmatrizen.*

(c)

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 \cdot 0 - 1 \cdot 1 & 0 \cdot 1 - 1 \cdot 0 \\ 1 \cdot 0 + 0 \cdot 1 & 1 \cdot 1 + 0 \cdot 0 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

(d) Die Matrizen

$$A = \begin{pmatrix} 1 & 2 & 1 \\ -1 & 0 & 3 \\ 4 & -3 & 2 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 1 & 2 & -2 \\ 2 & 1 & 5 \\ 3 & 1 & 0 \end{pmatrix}$$

sollen multipliziert werden. Zur einfacheren Orientierung verwenden wir das sogenannte *Falk-Schema*, in welchem stets das Skalarprodukt der links des zu berechnenden Elements (der Lösungsmatrix) stehenden Zeile mit der darüberstehenden Spalte berechnet wird:

$$\begin{array}{rrr|rrr} & & & 1 & 2 & -2 \\ & & B & 2 & 1 & 5 \\ & A & & 3 & 1 & 0 \\ \hline 1 & 2 & 1 & 8 & 5 & 8 \\ -1 & 0 & 3 & 8 & 1 & 2 \\ 4 & -3 & 2 & 4 & 7 & -23 \end{array}$$

Es ist also

$$A\,B = \begin{pmatrix} 8 & 5 & 8 \\ 8 & 1 & 2 \\ 4 & 7 & -23 \end{pmatrix}.$$

Wir werden später in Gleichung (1.3) sehen, warum die Multiplikation so definiert wird und nicht anders.

Zunächst listen wir auf, welche Rechenregeln von dieser Multiplikation erfüllt sind.

Lemma 1.5. *Die Matrixmultiplikation ist assoziativ und mit der Addition und skalaren Multiplikation verträglich, d.h. für reelle Matrizen A, B und C passender Größe und $\lambda \in \mathbb{R}$ gelten die Distributivgesetze:*

$$\lambda(A+B) = \lambda A + \lambda B \quad \textit{und} \quad C(A+B) = CA + CB.$$

und es gilt

$$(\lambda A)B = \lambda(AB) = A(\lambda B).$$

Mit den Einheitsmatrizen $\mathrm{Id}_n, \mathrm{Id}_m$ *gilt für jede* $(m \times n)$ *- Matrix* A*:*

$$A\,\mathrm{Id}_n = A = \mathrm{Id}_m\, A.$$

!! Was gilt nicht? Matrixmultiplikation ist *nicht kommutativ*, d.h. es kommt auf die Reihenfolge an bei der Multiplikation[1]. Nimmt man zum Beispiel die beiden Matrizen aus Beispiel 1.4 und multipliziert diese in der anderen Reihenfolge ist das neue Produkt (-1) mal dem aus dem Beispiel.
Außerdem folgt aus $AB = 0$ (0 ist hier die Nullmatrix) i.A. nicht, dass A oder B Null sein müssen (bei reellen Zahlen schon) – z.B. ist

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

Es gibt noch eine weitere Operation für Matrizen, die gelegentlich auftaucht.

Definition 1.6. Sei $A = (a_{ij})$ eine $(n \times m)$ - Matrix. Dann ist die *Transponierte* von A eine $(m \times n)$ - Matrix, die wir mit A^T bezeichnen. Ihre Zeilen sind gerade die Spalten von A und umgekehrt, es gilt also $(A^T)_{ij} = a_{ji}$.

Es gilt $(A^T)^T = A$.

Besonders nützlich ist diese Operation, wenn wir Vektoren aus Platzgründen lieber als Zeile schreiben, auch wenn wir Vektoren als Spaltenvektoren auffassen wollen. Wir schreiben dann einen Spaltenvektor als $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$:

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (x_1, \ldots, x_n)^T$$

Bemerkung 1.7. (i) Für $A \in \mathrm{Mat}_{\mathbb{R}}(n \times m)$ und $B \in \mathrm{Mat}_{\mathbb{R}}(m \times k)$ gilt $(AB)^T = B^T A^T$. (Für ein Beispiel hierzu siehe Aufgabe 2.)

(ii) Mit Hilfe der Transponierten können wir das euklidische Skalarprodukt[2] zweier Vektoren $x, y \in \mathbb{R}^n$ schreiben als

$$\langle x, y \rangle = x^T y,$$

[1] Es müssen quadratische Matrizen sein (d.h. gleich viele Zeilen wie Spalten) damit wir überhaupt beide Produkte bilden können. Aber selbst dann ist das Produkt i.A. verschieden.

[2] Zur Definition des Skalarproduktes in $\mathbb{R}^n$ siehe Beispiel 2.7

wobei das Produkt auf der rechten Seite aufgefasst wird als das Produkt von einer $(1 \times n)$ - Matrix mit einer $(n \times 1)$ - Matrix.

Konkret bedeutet das für $\mathbb{R}^3$ mit $x = (x_1, x_2, x_3)^T$ und $y = (y_1, y_2, y_3)^T$

$$\langle x, y \rangle = (x_1, x_2, x_3) \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = x_1 y_1 + x_2 y_2 + x_3 y_3.$$

(iii) Die i-ten Zeile einer Matrix $A = (a_{ij})$ heißt auch *i - ter Zeilenvektor von A*, notiert durch

$$v_i = (a_{i1}, \dots, a_{in}).$$

Die j- te Spalte heißt entsprechend *j - ter Spaltenvektor von A*, notiert durch

$$w_j = (a_{1j}, \dots, a_{mj})^T.$$

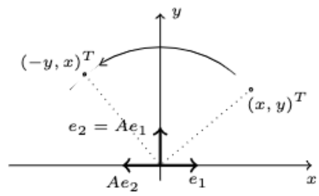
1.2. Matrizen als Abbildungen

Wir können eine reelle $m \times n$-Matrix A wie folgt als Abbildung auffassen:

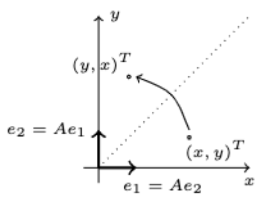
$$F_A \colon \mathbb{R}^n \to \mathbb{R}^m, \quad x \mapsto Ax$$

Beispiel 1.8. (i)

Die Matrix $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ ergibt die Abbildung $(x, y)^T \mapsto (-y, x)^T$. Geometrisch ist das die Rotation um $90°$ in mathematisch positive Richtung. In der Abbildung ist die Wirkung auf die Vektoren $e_1 = (1, 0)^T$ und $e_2 = (0, 1)^T$ dargestellt.



(ii)



Die Matrix $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ ergibt die Abbildung $(x, y)^T \mapsto (y, x)^T$. Geometrisch ist das die Spiegelung an der Winkelhalbierenden $y = x$.

Lemma 1.9. *Sei $A \in \mathrm{Mat}_{\mathbb{R}}(n \times m)$ eine Matrix. Seien $x, y \in \mathbb{R}^n$ Vektoren und sei $\lambda \in \mathbb{R}$ ein Skalar. Dann ist die Abbildung $F_A : \mathbb{R}^n \to \mathbb{R}^m$ linear: Es gilt*

$$F_A(x + \lambda y) = A(x + \lambda y) = Ax + \lambda Ay = F_A(x) + \lambda F_A(y). \tag{1.2}$$

Solche Abbildungen werden *lineare Abbildungen* genannt und wir werden später sehen, dass alle linearen Abbildungen von $\mathbb{R}^n$ nach $\mathbb{R}^m$ durch eine reelle $(m \times n)$ - Matrix beschrieben werden können.

Aus der Linearitätsbedingung(1.2) folgt insbesondere, dass, wenn wir den Wert der Abbildung F_A auf den Standardeinheitsvektoren $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ (die 1 steht an der i-ten Stelle) des $\mathbb{R}^n$ kennen, wir bereits die ganze Abbildung kennen, da dann für

$$x = (x_1, \ldots, x_n)^T = \sum_{i=1}^{n} x_i e_i$$

gilt:

$$F_A(x) = F_A\left(\sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{n} x_i F_A(e_i).$$

Es bleibt noch zu sehen, warum die Matrixmultiplikation so definiert wurde, wie in Definition 1.3 geschehen. Die Antwort ist das folgende Lemma:

Lemma 1.10. *Für die Hintereinanderausführung der von einer $(m \times n)$ - Matrix B und einer $(\ell \times m)$ - Matrix A definierten Abbildungen gilt:*

$$(F_A \circ F_B)(x) = F_A(F_B(x)) = F_A(Bx) = A(Bx) = (AB)x = F_{AB}(x). \tag{1.3}$$

D.h. die Hintereinanderausführung selbst ist durch die zugehörige Abbildung des Matrixprodukts gegeben.

Aufgabe 1. Welche der folgenden Matrizen sind multiplizierbar (ggf. auf mehr als eine Weise)? Berechnen Sie die zugehörigen Produkte:

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

Aufgabe 2. Gegeben sind die Matrizen

$$A = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 3 & 7 \end{pmatrix} \quad \text{und} \quad B = \begin{pmatrix} 3 & 2 \\ 1 & -3 \\ 0 & 4 \end{pmatrix}$$

Berechnen Sie AB und $B^T A^T$ und vergleichen Sie die Ergebnisse!

Aufgabe 3. Schreiben Sie die Spiegelung an der x-Achse als Abbildung von $f\colon \mathbb{R}^2 \to \mathbb{R}^2$. Finden Sie die zugehörige Matrix.

Aufgabe 4. Überlegen Sie sich an Hand einer Rechnung und einer Skizze, dass die Matrix $D_\varphi \in \mathrm{Mat}_{\mathbb{R}}(2 \times 2)$

$$D_\varphi = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \tag{1.4}$$

einen Vektor $x \in \mathbb{R}^2$ um den Winkel φ gegen den Uhrzeigersinn dreht.

Aufgabe 5. Wählen Sie eine beliebige Matrix $B \in \mathrm{Mat}_{\mathbb{R}}(3 \times 3)$ (sie sollte nicht überwiegend Nullen als Einträge haben) und berechnen Sie die Produkte $M_2^3 B$, $S_{2,3}^2 B$ und $V_{1,2} B$.

$$M_2^3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad S_{2,3}^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, \qquad V_{1,2} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Welche Wirkung haben diese Multiplikationen auf Ihre Matrix?
(Hinweis: Auf diese Matrizen werden wir in Kapitel 3 häufiger zurückgreifen.)

2. Vektorräume und lineare Abbildungen

2.1. Vektorräume

Betrachten wir das homogene reelle lineare Gleichungssystem

$$\begin{aligned} 2x_1 + 3x_2 - x_3 &= 0 \\ x_1 - 2x_2 + x_3 &= 0, \end{aligned}$$

so können wir, bevor wir die Lösungen berechnet haben, bereits zwei Dinge feststellen:

(a) Wenn (x_1, x_2, x_3) eine Lösung des linearen Gleichungssystems ist und (x_1', x_2', x_3') eine zweite, dann ist auch die Summe der Lösungen, zu verstehen als das Tupel $(x_1 + x_1', x_2 + x_2', x_3 + x_3')$, eine Lösung, denn

$$\begin{aligned} 2(x_1 + x_1') + 3(x_2 + x_2') - (x_3 + x_3') &= 2x_1 + 3x_2 - x_3 + 2x_1' + 3x_2' - x_3' \\ &= 0 + 0 = 0 \end{aligned}$$

und entsprechend $(x_1 + x_1') - 2(x_2 + x_2') + (x_3 + x_3') = 0$.

(b) Mit (x_1, x_2, x_3) ist auch das skalierte Tupel $(\lambda x_1, \lambda x_2, \lambda x_3)$ für jedes $\lambda \in \mathbb{R}$ eine Lösung, denn

$$2(\lambda x_1) + 3(\lambda x_2) - (\lambda x_3) = \lambda(2x_1 + 3x_2 - x_3) = \lambda \cdot 0 = 0$$

und entsprechend $(\lambda x_1) - 2(\lambda x_2) + (\lambda x_3) = 0$.

An keiner Stelle haben wir dabei die genauen Koeffizienten des linearen Gleichungssystems, die Anzahl der Variablen oder der Gleichungen benutzt. Das Einzige, was wichtig war, war die Homogenität, also dass auf der rechten Seite Nullen standen.
Ganz ähnlich sieht es aus, wenn wir uns die folgende reelle Differentialgleichung ansehen:

$$f(x) + 2f'(x) - 2f''(x) = 0.$$

Auch hier können wir, ohne die Lösungen zu kennen, bereits feststellen:

(a) Wenn f und g zwei Lösungen der Differentialgleichung sind, dann auch ihre Summe, denn

$$\begin{aligned} &(f+g)(x) + 2(f+g)'(x) - 2(f+g)''(x) \\ &\qquad = f(x) + g(x) + 2(f'(x) + g'(x)) - 2(f''(x) + g''(x)) \\ &\qquad = f(x) + 2f'(x) - 2f''(x) + g(x) + 2g'(x) - 2g''(x) \\ &\qquad = 0 + 0 = 0. \end{aligned}$$

(b) Mit f ist auch die mit einem konstanten Faktor multiplizierte Funktion λf für jedes $\lambda \in \mathbb{R}$ eine Lösung, denn

$$\begin{aligned} (\lambda f)(x) + 2(\lambda f)'(x) - 2(\lambda f)''(x) &= \lambda f(x) + 2(\lambda f'(x)) - 2(\lambda f''(x)) \\ &= \lambda(f(x) + 2f'(x) - 2f''(x)) \\ &= \lambda \cdot 0 = 0. \end{aligned}$$

Wir beobachten, dass in beiden Fällen das Produkt einer Lösung mit einem konstanten Faktor sowie die Summe zweier Lösungen wieder eine Lösung ergibt. (Hinweis: Für das Produkt zweier Lösungen gilt so eine Aussage im allgemeinen nicht.)
Diese Ähnlichkeit in der Struktur der Lösungsmengen ist die Motivation für die folgende Definition, die das Zentrum der linearen Algebra bildet:

Definition 2.1. Sei $(K, +, \cdot)$ ein Körper[1]. Ein K-*Vektorraum* ist ein Tupel $(V, +, \cdot)$ bestehend aus einer kommutativen Gruppe [2]. $(V, +)$ und einer Skalarmultiplikation $K \times V \to V$, sodass für alle $\lambda, \mu \in K$ und $v, w \in V$ gilt:

(a) $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$

(b) $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$

(c) $\lambda \cdot (\mu \cdot v) = (\lambda \cdot \mu) \cdot v$

(d) $1 \cdot v = v$

Die Gruppenverknüpfung $+$ heißt dabei *(Vektor-)Addition*. Die Elemente von V heißen *Vektoren*, die des Körpers K *Skalare*. Das neutrale Element 0 von V wird als *Nullvektor* bezeichnet.
Ist $K = \mathbb{R}$ der Körper der reellen Zahlen, so heißen $\mathbb{R}$-Vektorräume auch *reelle Vektorräume*. Ist $K = \mathbb{C}$ der Körper der komplexen Zahlen, so heißen $\mathbb{C}$-Vektorräume auch *komplexe Vektorräume*.

Hier müssen wir mit der Notation aufpassen: Die Multiplikation in K und die Multiplikation eines Vektors mit einem Skalar werden beide mit $\cdot$ bezeichnet. Man lässt auch oft in beiden Fällen das Symbol ganz weg und schreibt nur λv für $\lambda \cdot v$. Ebenso kann von nun an 0 sowohl die Null im Körper K als auch den Nullvektor im Vektorraum V bezeichnen. Es ist jeweils aus dem Kontext klar, welche Null gemeint ist, wenn es nicht dabei steht.

Bemerkung 2.2. Achtung! Wir können zwei Vektoren addieren und einen Vektor mit einem Skalar multiplizieren. Zwei Vektoren zu multiplizieren geht nicht auf die gleiche intuitive Weise.

Beispiel 2.3. (a) Sei $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R} = \{(x_1, \ldots, x_n) \mid x_1, \ldots, x_n \in \mathbb{R}\}$ das kartesische Produkt von n Kopien von $\mathbb{R}$. Wir definieren für zwei Elemente $(x_1, \ldots, x_n)$ und $(y_1, \ldots, y_n)$ aus $\mathbb{R}^n$ ihre Summe

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) := (x_1 + y_1, \ldots, x_n + y_n)$$

und für $\lambda \in \mathbb{R}$ setzen wir

$$\lambda(x_1, \ldots, x_n) := (\lambda \cdot x_1, \ldots, \lambda \cdot x_n).$$

[1] Die Definition und wichtigsten Eigenschaften des hier benötigten Begriffes *Körper* finden Sie in Anhang B

[2] Die Definition und wichtigsten Eigenschaften des hier benötigten Begriffes *Gruppe* finden Sie in Anhang A

Mit diesen beiden Verknüpfungen und dem Nullvektor $(0,\dots,0)$ ist $\mathbb{R}^n$ ein $\mathbb{R}$-Vektorraum:

Seien $(x_1,\dots,x_n),(y_1,\dots,y_n)\in\mathbb{R}^n$ und $\lambda\in\mathbb{R}$ beliebig. Dann gilt:

$$\begin{aligned}
\lambda((x_1,\dots,x_n)+(y_1,\dots,y_n)) &= \lambda(x_1+y_1,\dots,x_n+y_n)\\
&= \big(\lambda\cdot(x_1+y_1),\dots,\lambda\cdot(x_n+y_n)\big)\\
&= (\lambda\cdot x_1+\lambda\cdot y_1,\dots,\lambda\cdot x_n+\lambda\cdot y_n)\\
&= (\lambda\cdot x_1,\dots,\lambda\cdot x_n)+(\lambda\cdot y_1,\dots,\lambda\cdot y_n)\\
&= \lambda(x_1,\dots,x_n)+\lambda(y_1,\dots,y_n)
\end{aligned}$$

Die anderen Vektorraumaxiome zeigt man auf ähnliche Weise.

(b) Vektorräume K^n lassen sich analog zu $\mathbb{R}^n$ mit anderen Körpern K definieren. Ein Beispiel hierfür ist der $\mathbb{C}^n$.

(c) Die Menge der reellen $m\times n$-Matrizen $\mathrm{Mat}_{\mathbb{R}}(m\times n)$ bilden mit der in Abschnitt 1.1 eingeführten Addition und skalaren Multiplikation einen reellen Vektorraum.

(d) Die Menge $\mathcal{P}_n$ der Polynome vom Grad n auf $\mathbb{R}$ bilden einen reellen Vektorraum.

(e) **Bezug zur Analysis.** Die Menge $C(\mathbb{R})$ der stetigen Funktionen auf $\mathbb{R}$ oder $C([a,b])$ auf einem Intervall $[a,b]\subset\mathbb{R}$ sind reelle Vektorräume. Hier ist das neutrale Element der Addition die Funktion, die alles auf $0\in\mathbb{R}$ abbildet – die Nullfunktion.

(f) Allgemeiner als in Beispiel (c) gilt: Ist V ein K-Vektorraum und X eine beliebige Menge, dann ist die Menge

$$\mathrm{Abb}(X,V):=\{f\colon X\to V\}$$

aller Abbildungen von X nach V mit der Vektoraddition $(f+g)(x)=f(x)+g(x)$, der Skalarmultiplikation $(\lambda f)(x)=\lambda\cdot f(x)$ ein K-Vektorraum:

Seien f und g Abbildungen von X nach V und $\lambda\in K$. Wir müssen zeigen, dass die Abbildungen $\lambda(f+g)$ und $\lambda f+\lambda g$ gleich sind, also auf jedem Element von X dasselbe bewirken. Sei also $x\in X$ beliebig. Dann haben wir

$$\begin{aligned}
(\lambda(f+g))(x) &= \lambda\cdot((f+g)(x))\\
&= \lambda\cdot(f(x)+g(x))\\
&= \lambda\cdot f(x)+\lambda\cdot g(x)\\
&= (\lambda f)(x)+(\lambda g)(x)\\
&= (\lambda f+\lambda g)(x)
\end{aligned}$$

Auch hier zeigt man die übrigen Vektorraumaxiome auf ähnliche Weise.

Definition 2.4. Eine *Norm* ist eine Abbildung $\|\cdot\|$ von einem Vektorraum V über $\mathbb{R}$ oder $\mathbb{C}$ nach $\mathbb{R}$

$$\|\cdot\|\colon V \to \mathbb{R}$$

mit den Eigenschaften

(a) Definitheit	$\|x\| \geq 0 \quad \forall x \in V$ und aus	$\|x\| = 0$ folgt $x = 0$	(2.1)
(b) Homogenität	$\|\alpha x\| = \lvert\alpha\rvert\, \|x\|$	$\forall x \in V,\ \alpha \in \mathbb{R}$	(2.2)
(c) Dreiecksungleichung	$\|x + y\| \leq \|x\| + \|y\|$	$\forall x, y \in V$	(2.3)

Ein Vektorraum, auf welchem eine Norm definiert ist, bezeichnet man als *normierten Vektorraum.*

Beispiel 2.5. (a) Durch $\|x\| = \sqrt{x_1^2 + \ldots + x_n^2}$ wird eine Norm auf $\mathbb{R}^n$ definiert. Sie heißt auch *euklidische Norm.*

(b) Durch $\|x\|_{\max} = \max\{|x_i| \mid i = 1, \ldots, n\}$ wird ebenfalls eine Norm auf $\mathbb{R}^n$ definiert. Sie heißt *Maximumsnorm.*

(c) **Bezug zur Analysis.** Durch

$$\|f\|_{\sup} = \max_{x \in [a,b]} |f(x)|$$

wird eine Norm auf dem Vektorraum $C([a, b])$ der auf einem Intervall stetigen Funktionen definiert. Sie heißt ebenfalls Maximumsnorm oder auch *Supremumsnorm.*

Wir definieren ein Produkt für 2 Vektoren, dessen Ergebnis nicht wieder ein Vektor sondern ein Skalar ist:

Definition 2.6. (a) Eine *symmetrische Bilinearform* auf einem reellen Vektorraum V ist eine in beiden Argumenten lineare Abbildung

$$(\,\cdot\,,\,\cdot\,) : V \times V \to \mathbb{R},$$

für welche $(x, y) = (y, x)$ für alle $x, y \in V$ gilt.

(b) Ist eine Bilinearform *positiv definit*, das heißt, wenn

$$(x, x) \geq 0 \quad \text{und} \qquad (x, x) = 0 \Longleftrightarrow x = 0$$

gilt, so heißt sie *Skalarprodukt* auf V. Ein Skalarprodukt wird häufig mit spitzen Klammern $\langle\,\cdot\,,\,\cdot\,\rangle$ notiert. Ein Vektorraum mit einem Skalarprodukt heißt *euklidischer Vektorraum.*

Ist auf einem Vektorraum ein Skalarprodukt definiert, so erhält man durch

$$\|x\| = \sqrt{\langle x, x\rangle}$$

eine Norm. Sie heißt auch die durch das Skalarprodukt *induzierte* Norm.

Beispiel 2.7. (a) Die euklidische Norm des $\mathbb{R}^n$, siehe Beispiel 2.5, wird induziert durch das mittels

$$\langle x, y \rangle = x_1 y_1 + \ldots + x_n y_n$$

definierte Skalarprodukt auf $\mathbb{R}^n$.

(b) Auf der Menge $C([a, b])$ der stetigen Funktionen auf einem abgeschlossenen Intervall wird durch

$$\langle f, g \rangle = \int_a^b f(x) g(x)\, dx$$

ein Skalarprodukt definiert. Die durch dieses Skalarprodukt induzierte Norm heißt L^2 - Norm.

(c) Nicht jede Norm wird durch ein Skalarprodukt induziert. So gibt es zum Beispiel kein Skalarprodukt, welches die Maximumsnorm auf dem Vektorraum $C([a, b])$ induziert.

Bemerkung 2.8. Zusammenfassung der Begriffe:

(a) Ein Vektorraum über $\mathbb{R}$ heißt *reeller Vektorraum* oder $\mathbb{R}$ - Vektorraum.

(b) Ein Vektorraum mit einer Norm heißt *normierter Vektorraum*.

(c) Ein Vektorraum mit einem Skalarprodukt heißt *euklidischer Vektorraum*.

Bemerkung 2.9. Jede symmetrische Bilinearform b lässt sich durch eine symmetrische $(n \times n)$ - Matrix

$$B = (b_{ij} := b(e_i, e_j))_{ij}$$

darstellen, und umgekehrt definiert jede symmetrische $(n \times n)$ - Matrix B durch

$$b(x, y) := \langle x, By \rangle$$

eine symmetrische Bilinearform.

Bezug zur Analysis. Für eine zweimal stetig differenzierbare Funktion $f \colon \mathbb{R}^n \to \mathbb{R}$ ist die Hessesche

$$\operatorname{Hess} f(x) := \left(\frac{\partial^2}{\partial x_i \partial x_j}(x) \right)_{ij}$$

für jedes $x \in \mathbb{R}^n$ eine symmetrische $(n \times n)$ - Matrix und definiert somit eine symmetrische Bilinearform. Diese entspricht der Summe aller Terme zweiten Grades in der Taylorentwicklung von f.

Aufgabe 6. Überprüfen Sie die Vektorraumeigenschaften von $\mathcal{P}_n$, siehe Beispiel 2.3 (d).

Aufgabe 7. Zeigen Sie, dass die Maximumsnorm aus Beispiel 2.5 (b) die Eigenschaften einer Norm erfüllt.

Aufgabe 8. Überlegen Sie sich, warum durch $N(x) = x_1$ keine Norm auf $\mathbb{R}^n$ definiert wird.

2.2. Lineare Abbildungen

Jetzt, da wir Vektorräume kennengelernt haben, können wir lineare Abbildungen zwischen Vektorräumen (auch Vektorraumhomomorphismen genannt) definieren. Dieses sind Abbildungen zwischen 2 Vektorräumen, welche die Vektorraumstruktur respektieren. Da wir es in der Analysis 2 in der Regel mit $\mathbb{R}$-Vektorräumen zu tun haben, formulieren wir hier alle Aussagen für diese Situation. Sie gelten aber entsprechend allgemein für Vektorräume über einem Körper K.

Definition 2.10. Seien V und W zwei reelle Vektorräume. Ein *reeller Vektorraumhomomorphismus* oder auch eine $\mathbb{R}$*-lineare Abbildung* von V nach W ist eine Abbildung $f\colon V \to W$, für die gilt:

(a) Für alle $v_1, v_2 \in V$ gilt $f(v_1 + v_2) = f(v_1) + f(v_2)$ (Additivität).

(b) Für alle $\lambda \in \mathbb{R}$ und $v \in V$ gilt $f(\lambda \cdot v) = \lambda \cdot f(v)$ (Homogenität).

Beispiel 2.11. (a) Lineare Abbildungen zwischen den Vektorräumen $V = \mathbb{R}^n$ und $W = \mathbb{R}^m$ haben wir bereits im Abschnitt 1.2 Matrizen als Abbildungen kennengelernt, siehe hierzu Gleichung (1.2)

(b) Betrachten wir die Vektorräume $V = C^1(\mathbb{R})$ der einmal stetig differenzierbaren Funktionen und $W = C(\mathbb{R})$ der stetigen Funktionen auf $\mathbb{R}$, dann wird durch die Differentiation eine lineare Abbildung $'\colon V \to W$, $f \mapsto f'$, definiert. Dieses folgt aus den Differentiationsregeln

$$(f+g)'(x) = f'(x) + g\prime(x) \qquad \text{und} \qquad f'(\lambda x) = \lambda f'(x).$$

Nun interessieren wir uns für die Umkehrungen und Verkettungen linearer Abbildungen:

Lemma 2.12. *(a) Seien V und W zwei reelle Vektorräume. Wenn $f\colon V \to W$ eine bijektive lineare Abbildung ist, so ist die Umkehrabbildung f^{-1} auch linear.*

(b) Seien U, V und W reelle Vektorräume. Wenn die Abbildungen $f\colon U \to V$ und $g\colon V \to W$ beide linear sind, dann ist auch die Komposition $g \circ f\colon U \to W$ eine lineare Abbildung.

Beweis. (a) Wir zeigen nur die Homogenität:

$$\begin{aligned} f^{-1}(\lambda \cdot w) &= f^{-1}(\lambda \cdot f(f^{-1}(w))) \\ &= f^{-1}(f(\lambda \cdot f^{-1}(w))) \\ &= \lambda \cdot f^{-1}(w). \end{aligned}$$

(b) Hier zeigen wir die Additivität: Seien $x, y \in U$ beliebig. Dann gilt:

$$\begin{aligned} (g \circ f)(x+y) &= g(f(x+y)) \\ &\overset{f \text{ linear}}{=} g(f(x) + f(y)) \\ &\overset{g \text{ linear}}{=} g(f(x)) + g(f(y)) \\ &= (g \circ f)(x) + (g \circ f)(y) \end{aligned}$$

$\square$

Bezug zur Analysis. Ist $f\colon \mathbb{R}^n \to \mathbb{R}^m$ differenzierbar, dann ist die Ableitung $D_x f\colon \mathbb{R}^n \to \mathbb{R}^m$ von f in $x \in \mathbb{R}^n$ eine lineare Abbildung. Geschrieben in lokalen Koordinaten ist diese durch Multiplikation mit der zugehörigen Jacobimatrix gegeben. Ist auch $g\colon \mathbb{R}^m \to \mathbb{R}^\ell$ differenzierbar, dann auch $g \circ f$. In der zugehörigen Kettenregel kommt dann die Hintereinanderausführung von linearen Abbildungen vor.

Aufgabe 9. Zeigen Sie, dass durch die Projektion auf die erste Koordinate, $P_1(x) = x_1$ für $x \in \mathbb{R}^n$, eine lineare Abbildung $P_1 : \mathbb{R}^n \to \mathbb{R}$ definiert wird.

2.3. Untervektorräume

Nun betrachten wir Teilmengen von Vektorräumen, die unter der Vektoraddition und der Multiplikation mit einem Skalar abgeschlossen sind:

Definition 2.13. Eine Teilmenge $W \subseteq V$ heißt *Untervektorraum*, wenn die folgenden drei Eigenschaften erfüllt sind:

(a) Es gilt $W \neq \emptyset$

(b) Für alle $x, y \in W$ gilt auch $x + y \in W$, also ist W abgeschlossen unter der Addition.

(c) Für alle $x \in W$ und $\lambda \in K$ ist auch $\lambda x \in W$, also ist W auch abgeschlossen unter der Multiplikation mit einem Skalar.

Häufig verkürzt man das Wort Untervektorraum und spricht nur von Unterraum, wenn klar ist, dass es um Vektorräume geht.

Beispiel 2.14. (a) Einfache Untervektorräume des $\mathbb{R}^2$ sind

$$M_1 = \{v \in \mathbb{R}^2 \mid v = \begin{pmatrix} x \\ 0 \end{pmatrix}, x \in \mathbb{R}\},$$
$$M_2 = \{w \in \mathbb{R}^2 \mid v = \begin{pmatrix} 0 \\ y \end{pmatrix}, y \in \mathbb{R}\}.$$

(b) Im Vektorraum $\mathbb{R}^3$ ist die Menge der Lösungen $(x_1, x_2, x_3)^T$ des linearen Gleichungssystems

$$\begin{aligned} 2x_1 + 3x_2 - x_3 &= 0 \\ x_1 - 2x_2 + x_3 &= 0 \end{aligned}$$

ein Untervektorraum.

Die erste Eigenschaft eines Unterraums ist erfüllt, da $(0, 0, 0)$ offensichtlich eine Lösung ist. Die anderen beiden Eigenschaften haben wir bereits in der Einleitung des Kapitels auf Seite 10 nachgerechnet.

(c) **Bezug zur Analysis.** Im Vektorraum $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$ aller Funktionen $f\colon \mathbb{R} \to \mathbb{R}$ ist die Menge der Lösungen der Differentialgleichung

$$f(x) + 2f'(x) - 2f''(x) = 0$$

(insbesondere muss f dann zweimal differenzierbar sein) ein Untervektorraum aus denselben Gründen.

(d) In jedem Vektorraum V sind die Menge $\{0\}$ und der ganze Vektorraum V Untervektorräume. Man nennt sie die *trivialen Unterräume*.

(e) Die Menge $\mathcal{P}_n$ der Polynome vom Grad n sind ein Untervektorraum des Vektorraums $C(\mathbb{R})$.

Eine ganze Klasse von Beispielen für Untervektorräume findet man mit Hilfe linearer Abbildungen:

Beispiel 2.15. Seien V und W zwei $\mathbb{R}$-Vektorräume und $f\colon V \to W$ eine lineare Abbildung. Der *Kern* von f ist definiert durch

$$\ker(f) = f^{-1}(0) = \{v \in V \mid f(v) = 0\}$$

und dies ist ein Untervektorraum von V. Um dies zu zeigen, müssen die drei Eigenschaften aus Definition 2.13 überprüft werden: Zunächst gilt für lineare Abbildungen $f(0) = 0$, so dass $\ker(f) \neq \emptyset$ gegeben ist. Die Abgeschlossenheit unter der Addition folgt daraus, dass

$$f(v_1 + v_2) = f(v_1) + f(v_2) = 0 \qquad \forall v_1, v_2 \in \ker(f)$$

gilt, also auch $v_1 + v_2 \in \ker(f)$. Bleibt also nur noch die Abgeschlossenheit unter der Skalarmultiplikation zu zeigen. Sei dazu $v \in \ker(f)$ und $\lambda \in \mathbb{R}$. Dann gilt $f(\lambda v) = \lambda f(v) = \lambda \cdot 0 = 0$ und also auch $\lambda v \in \ker(f)$.

2.4. Lineare Unabhängigkeit

Definition 2.16. Sei V ein $\mathbb{R}$-Vektorraum.

(a) Eine *Linearkombination* von endlich vielen Vektoren $v_1, \ldots, v_r$ ist ein Ausdruck der Form

$$\lambda_1 v_1 + \ldots + \lambda_r v_r$$

für $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$. Wir sagen ein Vektor $v \in V$ ist als Linearkombination der Vektoren $v_1, \ldots, v_r$ darstellbar, wenn es $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$t mit $v = \lambda_1 v_1 + ?. + \lambda_r v_r$ gibt.

(b) Ist $E \subseteq V$ eine Teilmenge, so nennen wir

$$\mathrm{span}_{\mathbb{R}}(E) := \{v \in V \mid \exists e_1, \ldots, e_r \in E, \lambda_1, \ldots, \lambda_r \in K\colon v = \lambda_1 e_1 + \ldots + \lambda_r v_r\}$$

das *Erzeugnis* von E über $\mathbb{R}$.

Bemerkung 2.17. Eine andere Schreibweise für das Erzeugnis ist $\langle E\rangle$. Für endliche Mengen $E = \{v_1, \ldots, v_k\}$ schreiben wir auch $\langle v_1, \ldots, v_k\rangle$.

Ohne Beweis notieren wir das folgende Lemma:

Lemma 2.18. *Sei V ein $\mathbb{R}$-Vektorraum und $E \subseteq V$ beliebig. Dann ist* $\operatorname{span}_{\mathbb{R}}(E)$ *der kleinste Untervektorraum, der E enthält.*

Als nächstes beschäftigt uns die Frage, *wie viele* und *welche* Vektoren notwendig sind, um einen Vektorraum zu erzeugen.

Definition 2.19. Sei V ein reeller Vektorraum und W ein Untervektorraum von V. Ein *Erzeugendensystem* von W ist eine Teilmenge $E \subseteq W$ mit

$$\operatorname{span}_{\mathbb{R}}(E) = W.$$

Hat W ein Erzeugendensystem, das nur endlich viele Elemente hat, so nennen wir W *endlich erzeugt.*

Es ist leicht, für einen beliebigen Vektorraum ein Erzeugendensystem anzugeben, nämlich einfach den gesamten Vektorraum. Es ist aber klar, dass das ineffizient ist in dem Sinne, als wir Vektoren hätten weglassen können, sodass die restlichen immer noch ein Erzeugendensystem bilden.
Wir hätten gerne eine Möglichkeit herauszufinden, ob ein gegebenes Erzeugendensystem minimal ist.

Definition 2.20. Sei V ein reeller Vektorraum. Eine endliche Teilmenge $U = \{v_i \in V \mid i = 1, \ldots, r\}$ von Vektoren aus V heißt *linear unabhängig*, wenn sich der Nullvektor nur dadurch als Linearkombination der Vektoren aus U darstellen lässt, dass alle Koeffizienten gleich null gewählt werden:

$$\lambda_1 v_1 + \ldots + \lambda_r v_r = 0, \quad \Rightarrow \quad \lambda_1 = \ldots = \lambda_r = 0.$$

Eine Menge von Vektoren aus V, die nicht linear unabhängig ist, heißt *linear abhängig.*

Bemerkung 2.21. (a) Achtung! Wir gehen im folgenden immer davon aus, dass unterschiedlich indizierte Vektoren nicht gleich sind, d.h.

$$v_i \neq v_j \qquad \text{für } i \neq j.$$

(b) Betrachtet man in Definition 2.20 den Fall $r = 2$, so sieht man, dass 2 Vektoren genau dann linear unabhängig sind, wenn sie nicht Vielfache voneinander sind.

(c) Wenn 3 oder mehr Vektoren gegeben sind, die jeweils paarweise linear unabhängig sind, so folgt daraus *nicht*, dass sie zusammen auch linear unabhängig sind. Dieses macht man sich an einem Beipiel klar:

$$v_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad v_3 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

sind paarweise linear unabhängig, aber es ist $v_2 - v_1 - v_3 = 0$.

Zwei einfache Folgerungen aus Definition 2.20 enthält das folgende Lemma:

Lemma 2.22. *(a) Eine Teilmenge linear unabhängiger Vektoren aus einem Vektorraum V ist wieder linear unabhängig.*

(b) Seien die Vektoren $\{v_1, \dots v_n\}$ linear abhängig. Dann gibt es mindestens einen Index i_0, so dass sich der Vektor v_{i_0} als Linearkombination der übrigen Vektoren darstellen lässt:

$$v_{i_0} = \sum_{i=0, i \neq i_0}^{n} a_i v_i.$$

2.5. Basen von Vektorräumen und Dimension

Lemma 2.23. *Sei V ein reeller Vektorraum. Für Vektoren $v_1, \dots, v_r \in V$ gilt:*

(a) Die Vektoren $v_1, \dots, v_r$ sind genau dann ein Erzeugendensystem von V, wenn es für jeden Vektor $v \in V$ mindestens eine Darstellung $v = \lambda_1 v_1 + \dots + \lambda_r v_r$ gibt.

(b) Die Vektoren $v_1, \dots, v_r$ sind genau dann linear unabhängig, wenn es für jeden Vektor $v \in V$ höchstens eine Darstellung $v = \lambda_1 v_1 + \dots + \lambda_r v_r$ gibt.

Es bietet sich also an, beide Eigenschaften zu kombinieren, um für jeden Vektor eine eindeutige Darstellung zu bekommen.

Definition 2.24. Sei V ein reeller Vektorraum. Ein linear unabhängiges Erzeugendensystem $\mathcal{B} = (v_1, \dots, v_n)$ von V heißt *Basis* des Vektorraums V.

Bemerkung 2.25. Wichtige Eigenschaften einer Basis sind:

(a) Eine Basis $\mathcal{B}$ eines Vektorraumes V ist *unverkürzbar*. Das heißt, sobald man einen Vektor aus der Basis $\mathcal{B}$ entfernt, erzeugen die verbleibenden Vektoren nicht mehr den ganzen Raum V.

(b) Eine Basis $\mathcal{B} = (v_1, \dots, v_n)$ eines Vektorraumes V ist *unverlängerbar*. Fügt man zu einer Basis einen weiteren Vektor $w \in V$ hinzu, so sind die Vektoren $(v_1, \dots, v_n, w)$ nicht mehr linear unabhängig.

Wofür sind Basen gut? Eine erste Antwort liefert uns der folgende wichtige Satz:

Theorem 2.26. *Sei V eine Vektorraum und $\mathcal{B}$ eine Basis von V. Dann kann jeder Vektor aus V auf eindeutige Weise als Linearkombination der Basisvektoren dargestellt werden.*

Dieses gilt, da eine Basis ein Erzeugendensystem ist, es also für jeden Vektor aus V eine Darstellung als Linearkombination aus Basisvektoren geben muss. Da die Basisvektoren linear unabhängig sind, ist diese Darstellung aber auch eindeutig. Gäbe es nämlich zwei verschiedene Darstellungen, so wäre die Differenz der beiden eine Darstellung des 0 - Vektors, was der Definition von linear unabhängig widerspricht, siehe Definition 2.20.

Definition 2.27. Gegeben sei der $\mathbb{R}^n$. Die Familie $\mathcal{E} := (e_1, \ldots, e_n)$ der Vektoren

$$e_j := (0, \ldots, 0, \underbrace{1}_{j\text{-te Stelle}}, 0, \ldots, 0)$$

heißt *Standardbasis* des Vektorraums $\mathbb{R}^n$.

Bemerkung 2.28. Jede lineare Abbildung $f\colon V \to W$ ist eindeutig festgelegt durch die Bilder der Vektoren einer beliebigen Basis von V. Diese wichtige und nützliche Aussage überlegt man sich leicht auf die folgende Weise:

(a) Angenommen, die Werte einer linearen Abbildung f seien auf einer beliebigen Basis $(v_1, \ldots v_n)$ bekannt. Ist der Wert für einen Vektor $v \in V$ gesucht, so stellt man v als Linearkombination der Basisvektoren dar, $v = \sum_{i=1}^n a_i v_i$ und verwendet die Linearität von f:

$$f(v) = f\left(\sum_{i=1}^n a_i v_i\right) = \sum_{i=1}^n a_i f(v_i).$$

(b) Mit der gleichen Argumentation folgt, dass es genügt, die Werte einer linearen Abbildung auf einer beliebigen Basis zu kennen.

Beispiel 2.29. Von einer linearen Abbildung $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ sei bekannt, dass

$$f\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}\cos(\varphi)\\\sin(\varphi)\end{pmatrix} \quad \text{und} \quad f\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}-\sin(\varphi)\\\cos(\varphi)\end{pmatrix}$$

für ein festes $\varphi \in [0, 2\pi)$. Dann berechnet man

$$\begin{aligned} f\begin{pmatrix}2\\1\end{pmatrix} &= f\left(2\begin{pmatrix}1\\0\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}\right) = 2f\begin{pmatrix}1\\0\end{pmatrix} + f\begin{pmatrix}0\\1\end{pmatrix} \\ &= 2\begin{pmatrix}\cos(\varphi)\\\sin(\varphi)\end{pmatrix} + \begin{pmatrix}-\sin(\varphi)\\\cos(\varphi)\end{pmatrix} = \begin{pmatrix}2\cos(\varphi) - \sin(\varphi)\\2\sin(\varphi) + \cos(\varphi)\end{pmatrix} \end{aligned}$$

Nachdem wir gesehen haben, wofür Basen nützlich sind, interessieren wir uns dafür, wie sie gefunden werden können und ob jeder Vektorraum eine Basis hat. Der folgende Satz beantwortet das zumindest schon einmal für alle endlich erzeugten Vektorräume.

Theorem 2.30 (Basisauswahlsatz)**.** *Sei V ein reeller Vektorraum. Seien die Vektoren $(v_1, \ldots, v_r)$ ein Erzeugendensystem von V. Dann gibt es Indizes $i_1, \ldots, i_n \in \{1, \ldots, r\}$, sodass die Vektoren $(v_{i_1}, \ldots, v_{i_n})$ eine Basis bilden.*

Beweis. Wir verkleinern das Erzeugendensystem Schritt für Schritt. Wenn das Erzeugendensystem linear unabhängig ist, so haben wir bereits eine Basis und sind fertig.
Wenn das Erzeugendensystem linear abhängig ist, gibt es nach Lemma 2.22 einen Index $i_0 \in \{1, \ldots, r\}$, sodass v_{i_0} sich als Linearkombination der übrigen Vektoren darstellen lässt. Also ist bereits $\{v_1, \ldots, v_r\} \setminus \{v_{i_0}\}$ ein Erzeugendensystem, welches aber ein Element weniger hat als das ursprüngliche Erzeugendensystem.

Wir fahren so fort, bis die verbleibenden Vektoren linear unabhängig sind. Da wir mit endlich vielen Vektoren im Erzeugendensystem angefangen haben, erreichen wir auch nach endlich vielen Schritten ein linear unabhängiges Erzeugendensystem als die gesuchte Basis. □

Bemerkung 2.31. Interessant ist, dass auch jeder unendlichdimensionale Vektorraum eine Basis besitzt. Der Beweis ist aber recht aufwändig und führt für diesen Brückenkurs zu weit. Dieser Beweis ist dann Teil der Vorlesungen Lineare Algebra I und II.

Man sieht am Beweis oben direkt, dass die Basis nicht eindeutig sein muss. Denn für den Index i_0 kann man jeden Index nehmen, so dass der Koeffizient von v_i in einer Darstellung des Nullvektors gemäß Definition 2.20 nicht verschwindet. Also hat man die Wahl, welchen Vektor man weglassen möchte und erhält dann auch verschiedene Basen. Genauer gilt sogar:

Lemma 2.32 (Austauschlemma). *Sei V ein reeller Vektorraum. Sei $(v_1, \ldots, v_n)$ eine Basis von V und $w = \lambda_1 v_1 + \ldots + \lambda_n v_n \in V$. Falls $w \neq 0$, d.h., falls ein $j \in \{1, \ldots, n\}$ existiert mit $\lambda_j \neq 0$, dann ist auch $(v_1, \ldots, v_{j-1}, w, v_{j+1}, \ldots, v_n\}$ eine Basis.*

Diese Aussage verdeutlichen wir uns an einem Beispiel:

Beispiel 2.33. Es sei $\mathcal{E} = (e_1, e_2)$ die Standardbasis des $\mathbb{R}^2$. Wir setzen $w = (3, 2)^T = 3e_1 + 2e_2$. Dann bilden sowohl $\mathcal{E}_1 = (e_1, w)$ wie auch $\mathcal{E}_2 = (e_2, w)$ Basen des $\mathbb{R}^2$. Dieses macht man sich leicht an einer Skizze deutlich.

Theorem 2.34 (Basisaustauschsatz von Steinitz). *Sei V ein reeller Vektorraum. Sei $(v_1, \ldots, v_n)$ eine Basis von V und $(w_1, \ldots, w_m)$ seien linear unabhängige Vektoren in V.*

(a) Es gilt $m \leq n$.

(b) Die Vektoren $(w_1, \ldots w_m)$ können durch Vektoren aus der Basis $(v_1, \ldots, v_n)$ zu einer Basis ergänzt werden.

Auch diese Aussage veranschaulichen wir uns an einem Beispiel:

Beispiel 2.35. Es sei $V = \mathbb{R}^3$ mit der Standardbasis $\mathcal{E} = (e_1, e_2, e_3)$ und (w_1, w_2) seien zwei linear unabhängige Vektoren, $w_1 = (2, 0, -7)^T$ und $w_2 = (7, 0, 8)^T$. Dann lassen sich diese Vektoren durch (w_1, w_2, e_2) zu einer Basis des $\mathbb{R}^3$ ergänzen.

Hieraus folgt eine wichtige Aussage:

Korollar 2.36. *Sei V ein reeller Vektorraum. Wenn V eine Basis mit n Elementen hat, dann hat jede Basis von V genau n Elemente.*

Beweis. Seien $(v_1, \ldots, v_n)$ und $(w_1, \ldots, w_m)$ zwei Basen von V. Da beide Basen linear unabhängige Familien sind, gilt sowohl $m \leq n$ als auch $n \leq m$, also $m = n$. □

Diese Erkenntnis motiviert dann auch die folgende Definition.

Definition 2.37. Sei V ein reeller Vektorraum. Wenn V endlich erzeugt ist und eine Basis mit n Elementen hat (und damit jede Basis von V aus n Elementen besteht), nennen wir n die *Dimension* von V und schreiben

$$\dim(V) := n.$$

In diesem Fall sagen wir, V sei *endlichdimensional*. Wenn V keine Basis mit endlich vielen Elementen hat, dann nennen wir V *unendlichdimensional* und schreiben $\dim_K(V) := \infty$.

2.6. Darstellende Matrix

In Abschnitt 1.1 hatten wir gesehen, dass jede reelle $(m \times n)$ - Matrix A durch

$$\begin{aligned} f\colon \mathbb{R}^n &\to \mathbb{R}^m \\ x &\mapsto Ax \end{aligned}$$

eine lineare Abbildung definiert. Andersherum kann man jede lineare Abbildung zwischen *endlichdimensionalen* Vektorräumen durch eine Multiplikation *Matrix* $\cdot$ *Vektor* darstellen. Allerdings muss man dazu Basen auf den Vektorräumen wählen.

Seien V und W endlichdimensionale reelle Vektorräume mit einer Basis $\mathcal{B} = (v_1, \ldots, v_n)$ von V und $\mathcal{C} = (w_1, \ldots, w_m)$ von W. Dann kann jede lineare Abbildung $f\colon V \to W$ mit Hilfe einer Matrix A beschrieben werden, die wir wie folgt erhalten: Das Bild von $v_i \in V$ ist ein Element in W, welches wir in der Basis $\mathcal{C}$ schreiben:

$$f(v_i) = \sum_{j=1}^{m} a_{ij} w_j$$

Dann ist $A = (a_{ij})$. Schreiben wir einen Vektor bzgl. der Basis $\mathcal{B}$ auch wieder als Spaltenvektor

$$v = \sum_{i=1}^{n} b_i v_i =: (b_1, \ldots, b_n)^T$$

und analog

$$f(v) = \sum_{i=1}^{m} c_i w_i =: (c_1, \ldots, c_m)^T$$

bzgl. der Basis $\mathcal{C}$, dann ist bzgl. dieser Koordinaten

$$(c_1, \ldots, c_m)^T = A(b_1, \ldots, b_n)^T$$

Wir nennen A die *darstellende Matrix* bezüglich der gegebenen Basen $\mathcal{B}$ und $\mathcal{C}$ und bezeichnen sie mit

$$M(\mathcal{C}, f, \mathcal{B}).$$

Beispiel 2.38. Nehmen wir die lineare Abbildung f, die die Spiegelung an der Winkelhalbierenden $y = x$ im $\mathbb{R}^2$ darstellt, vgl. Beispiel 1.8. Die Matrix dort ist die darstellende Matrix bzgl. der Standardbasis des $\mathbb{R}^2$ (wenn nichts extra dazu gesagt, nehmen wir im $\mathbb{R}^n$ immer die Standardbasis). Als Basis im $\mathbb{R}^2$ nehmen wir dieses Mal nicht die Standardbasis sondern für $\mathcal{B} = \mathcal{C}$: $v_1 = e_1 + e_2$ und $v_2 = e_1 - e_2$. Dann gilt

$$\begin{aligned} f(v_1) &= f(e_1) + f(e_2) = e_2 + e_1 = v_1 = 1 \cdot v_1 + 0 \cdot v_2 \\ f(v_2) &= f(e_1) - f(e_2) = e_2 - e_1 = -v_2 = 0 \cdot v_1 - 1 \cdot v_2. \end{aligned}$$

Somit ist

$$M(\mathcal{B}, f, \mathcal{B}) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Vergleichen wir das nochmal mit der Matrix aus Beispiel 1.8 (diese war für $\mathcal{B} = \mathcal{C} =$ Standardbasis im $\mathbb{R}^2$), sehen wir, dass die darstellende Matrix von den gewählten Basen abhängt.

Bemerkung 2.39. Bei gegebenen Basen ist die Zuordnung

$$\text{lineare Abbildung} \quad \to \quad \text{darstellende Matrix}$$

bijektiv und mit den Vektorraumoperationen kompatibel, d.h. für zwei lineare Abbildungen $f, g \colon V \to W$ mit darstellender Matrix A bzw. B und einem $\lambda \in \mathbb{R}$ gilt

$$f + \lambda g \quad \to \quad A + \lambda B.$$

Aufgabe 10. Stellen Sie die darstellende Matrix der linearen Abbildung f aus Beispiel 2.38 bzgl. der Basis $(v_1 = (2,1)^T, v_2 = (1,3)^T)$ auf.

3. Invertierbarkeit und Determinante

3.1. Berechnung der Inversen einer Matrix

Hier gehen wir in einer ungewöhnlichen Reihenfolge vor. Wir geben zunächst ein Verfahren dazu an, wie die Inverse einer Matrix berechnet werden kann. Erst danach befassen wir uns mit der Frage, wann diese Inverse existiert.

Definition 3.1. Eine Matrix A heißt *invertierbar*, wenn es eine zweite Matrix gibt, die wir dann A^{-1} nennen, sodass

$$AA^{-1} = \mathrm{Id_n} = A^{-1}A. \tag{3.1}$$

Hieraus ergibt sich bereits, dass die zu einer Matrix A inverse Matrix A^{-1} nur dann berechnet werden kann, wenn die Matrix quadratisch ist, also vom Format $(n \times n)$. Sonst sind die Multiplikationen in (3.1) nicht definiert. Wir demonstrieren das Verfahren an einem Beispiel und gehen wie folgt vor: Wir notieren die gegebene Matrix A in einem um die Einheitsmatrix erweiterten Zahlenschema, d. h. es werden

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix} \quad \text{und} \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

zusammengefasst zu

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}\right). \tag{3.2}$$

Nun wird die Matrix A mit den sogenannten *elementaren Zeilenumformungen* zur Einheitsmatrix Id_3 umgeformt und gleichzeitig werden alle Schritte ebenfalls an der rechten Seite in (3.2) vorgenommen, wo zu Beginn die Einheitsmatrix Id_n steht. Erlaubt sind die folgenden *elementaren Zeilenumformungen*, wobei darauf zu achten ist, dass diese Umformung auch auf der rechten Seite des Systems ausgeführt werden müssen:

(I) Vertauschen zweier Zeilen.

(II) Addition eines Vielfachen einer Zeile zu einer anderen Zeile.

(III) Multiplikation einer Zeile mit einem Skalar $s \in \mathbb{R} \setminus \{0\}$.

Die Rechnung sieht dann aus wie folgt:

$$
\begin{array}{rl}
& \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} \cdot(-2) \\ + \\ \end{array} \\
\rightsquigarrow & \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -2 & -2 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \begin{array}{l} \\ \cdot(-2) \\ + \end{array} \\
\rightsquigarrow & \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -2 & -2 & 1 & 0 \\ 0 & 0 & 5 & 4 & -2 & 1 \end{array}\right) \begin{array}{l} \\ \\ \mid \cdot \frac{1}{5} \end{array} \\
\rightsquigarrow & \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -2 & -2 & 1 & 0 \\ 0 & 0 & 1 & \frac{4}{5} & -\frac{2}{5} & \frac{1}{5} \end{array}\right) \begin{array}{ll} & + \\ + & \\ \cdot(2) & \cdot(-1) \end{array} \\
\rightsquigarrow & \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{5} & \frac{2}{5} & -\frac{1}{5} \\ 0 & 1 & 0 & -\frac{2}{5} & \frac{1}{5} & \frac{2}{5} \\ 0 & 0 & 1 & \frac{4}{5} & -\frac{2}{5} & \frac{1}{5} \end{array}\right)
\end{array}
$$

Auf der rechten Seite steht nun die inverse Matrix A^{-1}:

$$
A^{-1} = \begin{pmatrix} \frac{1}{5} & \frac{2}{5} & -\frac{1}{5} \\ -\frac{2}{5} & \frac{1}{5} & \frac{2}{5} \\ \frac{4}{5} & -\frac{2}{5} & \frac{1}{5} \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 1 & 2 & -1 \\ -2 & 1 & 2 \\ 4 & -2 & 1 \end{pmatrix}.
$$

Bemerkung 3.2. (a) Die hier erlaubten Umformungen sind Ihnen eventuell aus dem *Gaußschen Eliminationsverfahren* zur Lösung linearer Gleichungssysteme aus der Schule bekannt.

(b) Dass man auf diese Weise die Inverse einer Matrix A berechnet, kann man sich auf folgende Weise überlegen: Diese elementaren Zeilenumformungen werden bewirkt durch Mulitplikation von links mit Matrizen wie sie in Aufgabe 5 angegeben wurden. Es bewirkt

(I) $V_{i,j}$ das Vertauschen der i-ten mit der j-ten Zeile,

(II) $S^s_{i,j}$ die Addition des s-fachen der j-ten Zeile zur i-ten Zeile,

(III) M^s_i die Multiplikation der i-ten Spalte mit dem Faktor $s \in \mathbb{R}$,

Man beginnt mit der Gleichung

$$A\,A^{-1} = \mathrm{Id}_n.$$

Diese multipliziert sie so lange von links mit geeigneten Matrizen vom Typ $V_{i,j}$, M^s_i und $S^s_{i,j}$ bis auf der linken Seite A zur Einheitsmatrix Id_n umgeformt geworden ist, so dass dort nur noch A^{-1} steht. Das Ergebnis der Umformungen auf der rechten Seite der Gleichung ist damit A^{-1}. Die oben gezeigte Rechnung ist eine vereinfachte Notation dieses Vorgehens.

Wendet man dieses Verfahren auf eine beliebige (2×2) - Matrix $A := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ an, so erhält man die folgende Formel:

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}. \tag{3.3}$$

Die Richtigkeit von (3.3) prüft man durch Nachrechnen von $AA^{-1} = \mathrm{Id}_2$ und $A^{-1}A = \mathrm{Id}_2$ nach.

Beispiel 3.3. (a) Wir wählen $A = \begin{pmatrix} 3 & 4 \\ -2 & 1 \end{pmatrix}$. Dann ist

$$A^{-1} = \frac{1}{(3 \cdot 1 - 4 \cdot (-2))} \begin{pmatrix} 1 & -4 \\ 2 & 3 \end{pmatrix} = \frac{1}{11} \begin{pmatrix} 1 & -4 \\ 2 & 3 \end{pmatrix}$$

(b) Die Matrix $S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ bewirkt die Spiegelung an der Winkelhalbierenden. Ihre Inverse ist wieder sie selbst

$$S^{-1} = \frac{1}{0 \cdot 0 - 1 \cdot 1} \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = S.$$

Zweifache Spiegelung eines Vektors führt auf den Vektor zurück.

(c) Ist $f\colon \mathbb{R}^n \to \mathbb{R}^n$ eine invertierbare lineare Abbildung und $M(\mathcal{B}, f, \mathcal{B})$ ihre darstellende Matrix. So ist die darstellende Matrix der Umkehrabbildung gegeben durch

$$M(\mathcal{B}, f^{-1}, \mathcal{B}) = M^{-1}(\mathcal{B}, f, \mathcal{B}).$$

Aufgabe 11. Bestimmen Sie die Inverse D_φ^{-1} der in (1.4) definierten Drehmatrix. Beachten Sie dabei die Symmetrieeigenschften der sin - und cos - Funktionen und interpretieren Sie Ihr Ergebnis anschaulich!

Bezug zur Analysis. Sei $f\colon \mathbb{R}^n \to \mathbb{R}^n$ eine bijektive differenzierbare Abbildung, so dass die Umkehrabbildung $g = f^{-1}$ auch wieder differenzierbar ist. Dann gilt für die Ableitungen von f in $x \in \mathbb{R}^n$ und g in $f(x)$:

$$D_{f(x)} g = (D_x f)^{-1}$$

(folgt mit der Kettenregel für differenzierbare Abbildungen).

3.2. Determinante

Die Frage, die uns nun beschäftigt, ist: Wie prüft man, ob eine gegebene $(n \times n)$ - Matrix invertierbar ist? Wie man in Gleichung (3.3) sehen kann, ist die Inverse zu einer (2×2) - Matrix genau dann definiert, wenn $ad - bc \neq 0$ gilt. Dieses veranlasst folgende Definition:

Definition 3.4. Die *Determinante* einer (2×2) - Matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ ist definiert durch

$$\det(A) = ad - bc. \tag{3.4}$$

Nun haben wir also einen Term definiert, mit dessen Hilfe wir die Frage der Invertierbarkeit einer (2×2) - Matrix beantworten können durch

$$A \text{ ist invertierbar} \qquad \Longleftrightarrow \qquad \det(A) \neq 0.$$

Eine solche Aussage würden wir gerne allgemein für $(n \times n)$ - Matrizen formulieren können. Die Frage ist dann, wie $\det(A)$ im allgemeinen Fall *definiert* werden muss. Eine vollständige Herleitung der Antwort auf diese Frage ist für diesen Kurs zu umfangreich. Sie wird in der Vorlesung Lineare Algebra gegeben. Wir beschränken uns darauf zu erklären, wie die Determinante einer größeren Matrix *berechnet* werden kann. Im nächsten Schritt berechnen wir die Determinante einer $(n \times n)$ - Matrix von einer bestimmten Form. Danach geben wir Regeln an, mit denen die Determinante einer allgemeinen $(n \times n)$ - Matrix auf die Determinante einer Matrix dieser speziellen Form zurückgeführt werden kann.

Definition 3.5. (a) Eine Matrix $A \in \mathrm{Mat}_{\mathbb{R}}(n \times n)$ heißt *obere Dreiecksmatrix*, wenn unterhalb der Diagonalelemente nur Nullen stehen:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & \cdots & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & \cdots & a_{2n} \\ & 0 & a_{33} & \cdots & \cdots & a_{3n} \\ \vdots & & 0 & \ddots & & \vdots \\ & \vdots & & \ddots & \ddots & \vdots \\ 0 & & \cdots & & 0 & a_{nn} \end{pmatrix}$$

(b) Die Determinante einer oberen Dreiecksmatrix wird definiert als Produkt ihrer Diagonalelemente:

$$\det(A) = a_{11} \cdot a_{22} \cdot \ldots \cdot a_{nn}. \tag{3.5}$$

(c) Analog kann man eine *untere Dreiecksmatrix* und ihre Determinante definieren. Aus Definition 1.6 und Gleichung (3.5) ergibt sich, dass dann für Dreiecksmatrizen $\det(A) = \det(A^T)$ gilt.

Bemerkung 3.6. (a) Man sieht direkt, dass die durch (3.5) definierte Determinante einer oberen Dreiecksmatrix genau dann von 0 verschieden ist, wenn alle Diagonalelemente $a_{11}, \dots, a_{nn}$ ungleich 0 sind.

(b) Man kann sich ebenfalls überlegen, dass eine obere Dreiecksmatrix, deren Diagonalelemente $a_{11}, \dots, a_{nn}$ ungleich 0 sind, invertierbar ist mit dem in Abschnitt 3 beschriebenen Verfahren.

Nun gilt es, diese Methode zur Berechnung einer Determinante für alle $(n \times n)$ - Matrizen zu verallgemeinern.

Man kann eine Abbildung $\det\colon \mathrm{Mat}_{\mathbb{R}}(n \times n) \to \mathbb{R}$ definieren, die für Dreiecksmatrizen mit der obigen Definition übereinstimmt, und für die die folgenden Regeln gilt:

Seien $A, B \in \mathrm{Mat}_{\mathbb{R}}(n \times n)$ zwei reelle Matrizen. Dann gilt:

(I) Geht B aus A hervor durch Vertauschen zweier Zeilen, so ist

$$\det(B) = -\det(A). \tag{3.6}$$

(II) Geht B aus A hervor durch Addition eines Vielfachen einer Zeile von A zu einer anderen Zeile von A, so ist

$$\det(B) = \det(A). \tag{3.7}$$

Mit diesen Regeln ist die Determinante einer Matrix eindeutig bestimmt.

Bemerkung 3.7. (a) Die Schritte (I) und (II) sind auch in Abschnitt 3.1 erlaubt, um die Inverse A^{-1} einer Matrix zu bestimmen.

(b) Im Gegensatz zu Abschnitt 3.1 ist es hier nicht erlaubt, einfach eine Zeile zu skalieren. Dann ändert sich die Determinante, wie unmittelbar aus der Definition der Determinante für eine obere Dreicksmatrix in (3.5) hervorgeht (siehe hierzu auch Theorem 3.9).

Anders als beim Berechnen der Inversen einer Matrix ist es zum Berechnen der Determinante nicht notwendig, die Matrix in die Einheitsmatrix Id_{n} umzuformen. Jetzt reicht es uns, die Matrix mit Hilfe der erlaubten Schritte (I) und (II) in eine obere Dreiecksmatrix umzuformen. Das ist immer möglich, sofern wir akzeptieren, dass einige Diagonalelemente möglicherweise gleich 0 sind. In diesen Fällen ist die Matrix nicht invertierbar und es gilt $\det(A) = 0$.

Beispiel 3.8. (a) Wir wenden Formel (3.5) auf eine obere Dreiecksmatrix an:

$$\det \begin{pmatrix} 3 & 5 & 2 \\ 0 & 2 & -1 \\ 0 & 0 & -2 \end{pmatrix} = -12$$

(b) Wir multiplizieren die 2. Zeile mit (-2), addieren sie zur 3. Zeile und wenden Regel (3.7) an:

$$\det\begin{pmatrix}3 & 5 & 2\\ 0 & 2 & -1\\ 0 & 4 & -4\end{pmatrix} = \det\begin{pmatrix}3 & 5 & 2\\ 0 & 2 & -1\\ 0 & 0 & -2\end{pmatrix} = -12$$

(c) Die 3. Zeile ist ein Vielfaches der 2. Zeile:

$$\det\begin{pmatrix}3 & 5 & 2\\ 0 & 2 & -1\\ 0 & 4 & -2\end{pmatrix} = \det\begin{pmatrix}3 & 5 & 2\\ 0 & 2 & -1\\ 0 & 0 & 0\end{pmatrix} = 0$$

(d) Wir zeigen, dass Formel (3.5) für (2×2) - Matrizen mit Formel (3.4) übereinstimmt. Hierzu verwenden wir wieder Regel (3.7) und ziehe das $\frac{c}{a}$ - fache der ersten Zeile von der zweiten Zeile ab:

$$\det\begin{pmatrix}a & b\\ c & d\end{pmatrix} = \det\begin{pmatrix}a & b\\ 0 & (d - b\,\frac{c}{a})\end{pmatrix} = ad - bc.$$

Die so für alle $(n\times n)$ - Matrizen definierte Determinante hat die folgenden Eigenschaften:

Theorem 3.9. *Sei eine $(n \times n)$ - Matrix gegeben durch ihre Zeilenvektoren*

$$A = (v_1^T, \dots, v_n^T),$$

sei $t \in \mathbb{R}$ und $b \in \mathbb{R}^n$ ein weiterer Vektor. Dann gilt folgende Linearität:

(a) $\det(v_1^T, \dots, tv_i^T, \dots, v_n^T) = t\,\det(v_1^T, \dots, v_i^T, \dots, v_n^T)$.

(b) $\det(V_1^T, \dots, (v_i + b)^T, \dots, v_n^T) = \det(v_1^T, \dots, v_i^T, \dots, v_n^T) + \det(v_1^T, \dots, b^T, \dots, v_n^T)$

Bemerkung 3.10. Anschauliche Interpretation:

(a) Betrachtet man die Zeilenvektoren einer Matrix A als die Koeffizienten eines linearen Gleichungssystems (LGS), so entsprechen die Umformungen (I) und (II) genau den Umformungen, welche man beim Lösen eines LGS mit dem Gaußschen Eliminationsverfahren vornimmt. Ist eine der Gleichungen des LGS eine Linearkombination der anderen Gleichungen, so kann diese mit Hilfe dieser Umformungen eliminiert werden. Das LGS ist in diesem Fall nicht eindeutig lösbar, was äquivalent dazu ist, dass die Matrix A nicht invertierbar ist. Ist keine der Gleichungen als Linearkombination der anderen darstellbar, so ist A invertierbar und $\det(A) \neq 0$.

(b) Mittels der Umformung $A \mapsto A^T$ werden die Zeilen und die Spalten einer Matrix vertauscht. Die Umformungen (I) und (II) können also statt auf die Zeilen auch auf die Spalten einer Matrix angewendet werden, ohne dass sich dadurch die Determinante ändert.

(c) Nun fügen wir Vektoren $v_1, \ldots, v_n \in \mathbb{R}^n$ als Spaltenvektoren zu einer Matrix A zusammen. Sind die Vektoren linear unabhängig, bilden sie also eine Basis des $\mathbb{R}^n$, so kann mittels der Umformungen (I) und (II) keine der Spalten eliminiert werden und es gilt $\det(A) \neq 0$. Daran sehen wir, dass für eine Matrix genau dann $\det(A) \neq 0$ gilt, wenn die Spaltenvektoren von A eine Basis des $\mathbb{R}^n$ bilden.

Zusammengefasst liefern diese Resultate:

Korollar 3.11. *Für eine Matrix $A \in \mathrm{Mat}_{\mathbb{R}}(n \times n)$ sind die folgenden Bedingungen äquivalent:*

(a) A ist invertierbar.

(b) Die durch A definierte Abbildung $F_A \colon \mathbb{R}^n \to \mathbb{R}^n$ ist invertierbar, hat also eine Umkehrabbildung.

(c) Die Bilder der Standardbasisvektoren des $\mathbb{R}^n$ unter der Abbildung F_A, also die Familie $(F_A(e_1), \ldots, F_A(e_n))$, bilden eine Basis des $\mathbb{R}^n$.

(d) Die Spaltenvektoren von A bilden eine Basis des $\mathbb{R}^n$.

Lemma 3.12. *Sei K ein Körper, $n \in \mathbb{N} \setminus \{0\}$ und V ein n-dimensionaler K-Vektorraum. Bilden die Vektoren $v_1, \ldots, v_n \in V$ keine Basis von V, so gilt $\det(v_1, \ldots, v_n) = 0$.*

Beweis. Da $\dim V = n$ ist, kann die Familie $(v_1, \ldots, v_n)$ nur dann keine Basis sein, wenn sie linear abhängig ist. Dann gibt es aber einen Vektor v_i, der sich als Linearkombination $v_i = \sum_{j \neq i} \lambda_j v_j$ der anderen schreiben lässt. Dann gilt aber

$$\det(v_1, \ldots, v_i, \ldots, v_n) = \det(v_1, \ldots, \sum_{j \neq i} \lambda_j v_j, \ldots, v_n) = \sum_{j \neq i} \lambda_j \det(v_1, \ldots, v_j, \ldots, v_n).$$

Weil $\det(\cdot)$ alternierend ist und in jeder der Determinanten am Ende der Zeile ein Eintrag doppelt vorkommt, ist jede davon 0. □

Aufgabe 12. Sei A eine reelle $n \times n$-Matrix. Wie ändert sich die Determinante, wenn wir λA betrachten?

3.3. Multiplikationsformel für Determinanten

Theorem 3.13 (Multiplikationsregel für Determinanten)**.** *Für alle $(n \times n)$ - Matrizen A, B gilt:*

$$\det(BA) = \det(B)\det(A).$$

Bemerkung 3.14. Interessant ist vielleicht noch, dass man auch die Determinante einer linearen Abbildung $f \colon \mathbb{R}^n \to \mathbb{R}^n$ definieren kann. Zwar kann man eine jede solche Abbildung als Matrix darstellen und davon die Determinante berechnen, jedoch hängt die Matrixdarstellung von der gewählten Basis ab. Man kann sich aber überlegen, dass für zwei Basen $\mathcal{B}$ und $\mathcal{C}$ von $\mathbb{R}^n$ gilt (in der Notation von Abschnitt 2.6)

$$\begin{aligned} M(\mathcal{C}, f, \mathcal{C}) &= M(\mathcal{C}, \mathrm{id}, \mathcal{B}) M(\mathcal{B}, f, \mathcal{B}) M(\mathcal{B}, \mathrm{id}, \mathcal{C}) \\ &= M(\mathcal{C}, \mathrm{id}, \mathcal{B}) M(\mathcal{B}, f, \mathcal{B}) M(\mathcal{C}, \mathrm{id}, \mathcal{B})^{-1} \end{aligned}$$

gilt. Zusammen mit der Multiplikationsformel und der Aufgabe 13 folgt

$$\det M(\mathcal{C}, f, \mathcal{C}) = \det M(\mathcal{B}, f, \mathcal{B}),$$

d.h. diese Determinante hängt nicht von der gewählten Basis ab.

Deshalb definiert man

$$\det f := \det M(\mathcal{B}, f, \mathcal{B}).$$

Bemerkung 3.15. Für die Matrix $S^s_{i,j}$ aus Aufgabe 5, welche durch Multiplikation von links das s - fache der i - ten Zeile zur j -ten Zeile einer Matrix A addiert, gilt $\det(S^s_{i,j}) = 1$. Daher sehen wir hier zusammen mit der Multiplikationsregel aus Definition 3.13, dass durch diese Umformung die Determinante einer Matrix nicht verändert wird:

$$\det(S^s_{i,j} A) = \det(S^s_{i,j}) \det(A) = \det(A)$$

Aufgabe 13. Sei A eine invertierbare Matrix. Bestimmen Sie $\det(A^{-1})$ in Abhängigkeit von $\det(A)$. Warum sieht man auch an dieser Beziehung, dass für eine invertierbare Matrix $\det(A) \neq 0$ sein muss?

Aufgabe 14. Wir betrachten die Matrix

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

(a) Überlegen Sie sich, dass die Multiplikation einer beliebegen Matrix $A \in \mathrm{Mat}_{\mathbb{R}}(n \times n)$ von links mit V zwei Zeilen vertauscht und eine Zeile mit -1 multipliziert.

(b) Überlegen Sie außerdem, dass die Multiplikation mit dieser Matrix auch als Produkt dreier Matrizen vom Typ $S^s_{i,j}$ dargestellt werden kann.

(c) Schlussfolgern Sie daraus, dass bei Vertauschen von zwei Zeilen die Determinante das Vorzeichen wechseln muss.

(d) Was passiert, wenn Sie eine Matrix von *rechts* mit V multiplizieren?

3.4. Rang einer Matrix

Sei A eine reelle $(n \times m)$-Matrix. Der *Rang* von A ist die maximale Anzahl linear unabängiger Spalten von A. Es gilt:

Lemma 3.16. *Sei A eine reelle $(n \times m)$-Matrix mit Rang ℓ. Dann gilt*

(i) $\ell \leq \min\{m,n\}$ *und* ℓ *ist die Dimension des Untervektorraums Bild*$(A) \subset \mathbb{R}^n$.

(ii) Ist $\ell = m = \min\{m,n\}$, *dann ist* $A\colon \mathbb{R}^m \to \mathbb{R}^n$ *injektiv.*

(iii) Ist $\ell = n = \min\{m,n\}$, *dann ist* $A\colon \mathbb{R}^m \to \mathbb{R}^n$ *surjektiv.*

(iv) Ist $\ell = n = m$, *dann ist* $A\colon \mathbb{R}^m \to \mathbb{R}^n$ *bijektiv und damit ist* $\det(A) \neq 0$.

Bezug zur Analysis. Ist $f\colon \mathbb{R}^n \to \mathbb{R}^k$ differenzierbar mit $k \leq n$ und der Rang von $D_u f$ sei k. Dann gibt es k Spalten von $D_u f$, die eine invertierbare Matrix formen. Das kommt im Satz über implizite Funktionen vor (dort wird direkt angenommen, dass es die letzten k Spalten sind und die Annahme über den Rang wird über die Determinante der Matrix aus diesen Spalten formuliert).

In der Definition einer Untermannigfaltigkeit und im Kriterium zum regulären Wert kommt der Rang der Ableitung vor.

3.5. Eigenwerte und Diagonalisierbarkeit

Hier gehen wir nochmal zurück zu Matrizen linearer Abbildungen und erinnern uns, dass die Matrix einer linearen Abbildung des $\mathbb{R}^n$ auf sich selbst von der Wahl einer Basis abhängt. Ein zentraler Teil der linearen Algebra befasst sich damit, wie man Basen findet, bezüglich der diese Matrix möglichst einfach aussieht.

Definition 3.17. Sei $A \in \mathrm{Mat}_{\mathbb{R}}(n \times n)$. Ein Vektor $v \neq 0$ heißt *Eigenvektor* von A, wenn gilt

$$Av = \lambda v \quad \text{für ein } \lambda \in \mathbb{R}. \tag{3.8}$$

Der Skalar λ heißt dann *Eigenwert* von A.

Geometrisch gesehen, wird ein Eigenvektor durch die Anwendung der Matrix also höchstens gestreckt mit Faktor λ aber nicht gedreht.

Beispiel 3.18. Die für Matrix $A = \begin{pmatrix} 2 & 7 \\ 0 & 1 \end{pmatrix}$ gilt

$$\begin{pmatrix} 2 & 7 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} = 2 \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Daher ist $v = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ Eigenvektor von A zum Eigenwert $\lambda = 2$. Aus der Linearität folgt, dass auch jedes Vielfache von v ein Eigenvektor zum gleichen Eigenwert ist:

Theorem 3.19. *Sei* $A \in \mathrm{Mat}_{\mathbb{R}}(n \times n)$ *und es seien* $\lambda, \mu \in \mathbb{R}$ *Eigenwerte von* A *mit Eigenvektoren* $v_\lambda, v_\mu \in \mathbb{R}^n$. *Dann gilt:*

(a) Für jedes $s \in \mathbb{R} \setminus \{0\}$ ist sv_λ Eigenvektor von A zum Eigenwert λ.

(b) Aus $\lambda \neq \mu$ folgt dass v_λ und v_μ linear unabhängig sind.

Nun beschäftigen wir uns mit der Frage, wie wir zu einer Matrix A Eigenwerte und Eigenvektoren berechnen können.

Theorem 3.20. *Sei $A \in \mathrm{Mat}_{\mathbb{R}}(n \times n)$. Es ist $\lambda \in \mathbb{R}$ genau dann Eigenwert von A, wenn*

$$\det(A - \lambda \mathrm{Id_n}) = 0. \tag{3.9}$$

Beweis. Ein $\lambda \in \mathbb{R}$ ist Eigenwert von A, wenn $Av = \lambda v$ für ein $v \neq 0$ gilt, was äquivalent ist zu

$$(A - \lambda \mathrm{Id_n})v = 0, \tag{3.10}$$

d. h. λ ist genau dann Eigenwert von A, wenn die Matrix $(A - \lambda \mathrm{Id_n})$ nicht invertierbar ist, d.h. wenn

$$\det(A - \lambda \mathrm{Id_n}) = 0 \tag{3.11}$$

gilt. Denn sonst wäre $v = (A - \lambda \mathrm{Id_n})^{-1}0 = 0$ die einzige Lösung von (3.10). □

Der Term in (3.11) ist ein Polynom vom Grad n in λ. Damit ist folgende Begriffsbildung sinnvoll:

Definition 3.21. Zu einer Matrix $A \in \mathrm{Mat}_{\mathbb{R}}(n \times n)$ heißt

$$p(\lambda) = \det(A - \lambda I_n)$$

das *charakteristische Polynom* von A.

Die Eigenwerte einer Matrix können wir also bestimmen, indem wir die Nullstellen ihres charakteristischen Polynoms berechnen. Die zugehörigen Eigenvektoren bestimmen wir dann, indem wir das lineare Gleichungssystem 3.10 für die berechneten Eigenwerte lösen.

Beispiel 3.22. (a) Es sollen die Eigenwerte und Eigenvektoren der Matrix $A = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$ bestimmt werden. Hierzu berechnen wir

$$\begin{aligned}
& & \det(A - \lambda I_2) &= 0 \\
\iff & & \det \begin{pmatrix} 2-\lambda & 3 \\ 1 & 4-\lambda \end{pmatrix} &= 0 \\
\iff & & (2-\lambda)(4-\lambda) - 3 &= 0 \\
\iff & \quad \lambda_1 = 5 & \lambda_2 &= 1.
\end{aligned}$$

Um die Eigenvektoren zu diesen Eigenwerten zu berechnen, muss das Gleichungssystem (3.10) mit λ_1 und λ_2 gelöst werden. Für $\lambda_1 = 5$ muss gelten

$$\begin{aligned}
&\left(\begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix} - 5\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
\iff \quad &\begin{pmatrix} -3 & 3 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
\iff \quad &\begin{pmatrix} -3x_1 + 3x_2 \\ 1x_1 - 1x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.
\end{aligned}$$

Wir erhalten 2 linear abhängige Gleichungen für x_1 und x_2,

$$-2x_1 + 2x_2 = 0 \quad \text{und} \quad 2x_1 - 2x_2 = 0,$$

welche für $x_1 = x_2$ erfüllt sind. Ein Eigenvektor $\mathbf{v}$ zu $\lambda_1 = 5$ ist also $\mathbf{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. In parametrisierter Form erhält man mit $\mathbf{v} = t\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ mit $t \in \mathbb{R}$ alle Eigenvektoren zu λ_1. Für $\lambda_2 = 1$ muss gelten

$$\begin{aligned}
&\left(\begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix} - 1\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
\iff \quad &\begin{pmatrix} 1 & 3 \\ 1 & 3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
\iff \quad &\begin{pmatrix} x_1 + 3x_2 \\ x_1 + 3x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.
\end{aligned}$$

Wir erhalten zweimal die selbe Gleichung für x_1 und x_2 welche für $x_1 = -3x_2$ erfüllt ist. Ein Eigenvektor $\mathbf{w}$ zu $\lambda_2 = 1$ ist daher $\mathbf{w} = \begin{pmatrix} -3 \\ 1 \end{pmatrix}$. In parametrisierter Form erhält man mit $\mathbf{w} = t\begin{pmatrix} -3 \\ 1 \end{pmatrix}$ für $t \in \mathbb{R}$ wieder alle Eigenvektoren zu λ_2.

Satz 3.23. *Wenn eine lineare Abbildung $f\colon \mathbb{R}^n \to \mathbb{R}^n$ Eigenvektoren $v_1, \dots, v_n$ zu Eigenwerten $\lambda_1, \dots \lambda_n$ hat, so dass $\mathcal{B} = (v_1, \dots, v_n)$eine Basis des $\mathbb{R}^n$ ist, dann ist $M(\mathcal{B}, f, \mathcal{B})$ eine Diagonalmatrix mit den Werten λ_i auf der Diagonalen.*

Beweis. Die k-te Spalte der Matrix ist gerade das Bild von v_k dargestellt in der Basis $\mathcal{B}$. Nach Voraussetzung gilt $f(v_k) = \lambda_k v_k$, also steht in dieser Spalte nur der Koeffizient λ_k in der k-ten Zeile. □

Leider gibt es nicht für jede Endomorphismus eine solche Basis. Wenn es so eine Basis gibt, dann heißt der Endomorphismus auch *diagonalisierbar*[1] in Anlehnung daran, dass man eine Matrix in Diagonalgestalt erhält.

[1] Das charakteristische Polynom kann natürlich auch komplexe Nullstellen haben. Dann sind auch die Eigenvektoren komplex (man betrachtet die Matrix dann als Matrix, die auf $\mathbb{C}^n$ wirkt). Strenggenommen bedeutet *diagonalisierbar*, dass man auch solche komplexen Eigenvektoren zulässt und die Matrix bzgl derer als Diagonalmatrix darstellen kann. Ist hier für uns aber nicht relevant.

Die allgemeineren Normalformen sollen uns hier nicht weiter interessieren, diese sind Teil der Vorlesungsreihe Lineare Algebra.

Aber es gilt:

Satz 3.24. *Für jede symmetrische reelle Matrix A (also $A = A^T$) hat das charakteristische Polynom nur reelle Nullstellen und die Matrix ist diagonalisierbar.*

In Abschnitt 3.3 haben wir uns überlegt, dass die Determinante einer linearen Abbildung $f\colon V \to V$ auf einem endlichdimensionalen Vektorraum einfach als Determinante einer darstellenden Matrix von f definiert werden kann und dies wohldefiniert. Deshalb übersetzt sich der Begriff von Eigenwerten und des charakteristischen Polynoms direkt auf solche f: Ein Vektor $v \in V$, $v \neq 0$, ist ein Eigenwert von f, falls es ein λ mit $f(v) = \lambda v$ gibt. Damit gilt:

Satz 3.25. *Sei $f\colon V \to V$ eine lineare Abbildung. Dann ist*

$$\chi_f(\lambda) := \det(f - \lambda \mathrm{Id}_V)$$

ein Polynom in λ, dessen Nullstellen die Eigenwerte von f sind.

Das Polynom χ_f wird als *charakteristisches Polynom* von f bezeichnet.

Aufgabe 15. Bestimmen Sie die Eigenwerte und Eigenvektoren der Matrizen aus Beispiel 1.8. Hätten Sie auch an der geometrischen Bedeutung dieser Abbildungen ablesen können, was die Eigenvektoren/-werte sind?

A. Gruppen

In der Algebra geht es im Wesentlichen um Mengen mit Verknüpfung. Eine Verknüpfung $*$ ist dabei eine Abbildung, die je zwei Elementen a und b der Menge ein drittes, geschrieben $a * b$, zuordnet. Im Prinzip ist das nichts neues: schon aus der ersten Klasse bekannt ist natürlich die Addition, die aus zwei natürlichen Zahlen die Summe $a + b$ bildet. Später wurde die Addition auch auf anderen Mengen, bis hin zu der Menge der reellen Zahlen, definiert. Das ist auch gleich eines der wichtigsten Beispiele. Eine andere Verknüpfung ist aber zum Beispiel auch die Verkettung oder Komposition von Abbildungen, die zwei Abbildungen f und g mit geeigneten Definitions- und Wertebereichen zu der neuen Abbildung $x \mapsto (f \circ g)(x)$ verknüpft.
Man sieht, dass sich die beiden Verknüpfungen in einer wesentlichen Eigenschaft unterscheiden: Bei der Addition natürlicher oder reeller Zahlen spielt die Reihenfolge der beiden Elemente a und b keine Rolle, bei der Verkettung dagegen ist die Reihenfolge der Funktionen sehr wohl entscheidend, wie man am Beispiel der Funktionen $f\colon \mathbb{R} \to \mathbb{R}, x \mapsto x + 2$ und $g\colon \mathbb{R} \to \mathbb{R}, x \mapsto x^2$ leicht sieht: $f \circ g$ bildet 0 auf 2 ab, $g \circ f$ dagegen auf 4.

A.1. Definitionen und Eigenschaften

Definition A.1. Eine *Verknüpfung* $*$ auf einer Menge X ist eine Abbildung

$$X \times X \to X, \quad (x, y) \mapsto x * y$$

Eine Menge X mit einer Verknüpfung $*$ heißt *kommutative Gruppe* $(X, *)$, wenn

(a) das *Assoziativgesetz* gilt, d. h. wenn

$$(x * y) * z = x * (y * z) \quad \text{für alle } x, y, z \in X,$$

(b) das *Kommutativgesetz* gilt. d. h. wenn

$$x * y = y * x \quad \text{für alle } x, y \in X,$$

(c) ein *neutrales Element* $e \in X$ existert, so dass

$$e * x = x = x * e,$$

(d) es zu jedem Element $a \in X$ ein *inverses Element* $a^{-1} \in X$ gibt, so dass

$$a * a^{-1} = e = a^{-1} * a$$

gilt.

Eine Gruppe $(X, *)$ heißt *nicht kommutativ*, wenn das Kommutativgesetz nicht erfüllt ist.

Bemerkung A.2. (a) Wichtig ist, dass wir für *alle* Paare von Elementen aus der Menge ein Ergebnis der Verknüpfung definiert haben. Die Subtraktion ist zum Beispiel auf den natürlichen Zahlen keine wohldefinierte Verknüpfung, da $2 - 5$ keine natürliche Zahl mehr ist.

Genauso definiert die Division keine Verknüpfung auf den reellen Zahlen $\mathbb{R}$, da wir nicht durch 0 dividieren dürfen. Auf $\mathbb{R} \setminus \{0\}$ ist die Division dagegen eine Verknüpfung. Sogar eine, die nicht assoziativ ist, denn

$$(12/6)/2 = 2/2 = 1, \quad \text{aber} \quad 12/(6/2) = 12/3 = 4$$

(b) Wir werden uns hier eigentlich nur mit assoziativen Verknüpfungen beschäftigen, hauptsächlich mit Addition, Multiplikation und Verkettung von Abbildungen. Subtraktion und Division tauchen dann nur auf, wenn es um Inverse bezüglich der Addition bzw Multiplikation geht. Wie bei Addition und Multiplikation bedeutet das Assoziativgesetz, dass wir Klammern einfach weglassen können. Also schreiben wir einfach $x * y * z$ für den Ausdruck $(x * y) * z = x * (y * z)$, ohne uns um die Klammern zu kümmern.

Beispiel A.3. (a) Die ganzen Zahlen $\mathbb{Z}$, die rationalen Zahlen $\mathbb{Q}$, die reellen Zahlen $\mathbb{R}$ und komplexen Zahlen $\mathbb{C}$ sind mit der üblichen Addition Gruppen. Die inversen Elemente haben einfach das andere Vorzeichen, also ist -1 das Inverse zu 1 und 3 das Inverse zu -3.

(b) Mit der bekannten Multiplikation sind $\mathbb{N}_+, \mathbb{N}_0$ und $\mathbb{Z}$ jeweils keine Gruppen, da es nicht für alle Elemente inverse Elemente gibt. Wegen $0 \cdot x = 0$ für jede Zahl x kann 0 kein multiplikatives Inverses haben, wir bekommen also nie Gruppen. In den natürlichen und ganzen Zahlen können wir nur durch ± 1 dividieren, alle anderen Elemente haben kein multiplikatives Inverses. Für $\mathbb{Q}, \mathbb{R}$ und $\mathbb{C}$ bekommen wir aber Gruppen, wenn wir die 0 weglassen, also jeweils $(Q \setminus \{0\}, \cdot)$, $(\mathbb{R} \setminus \{0\}, \cdot)$ und $(C \setminus \{0\}, \cdot)$ betrachten.

(c) Sei X eine beliebige Menge und G die Menge der bijektiven Abbildungen von X in sich selbst, also

$$G = \{f \colon X \to X \mid f \text{ bijektiv}\}.$$

Dann ist $(G, \circ)$ eine Gruppe, wobei $\circ$ die Verkettung von Abbildungen ist. Das neutrale Element ist die Identität, die jedes Element auf sich abbildet. Die Inversen sind natürlich die Umkehrabbildungen, da jede bijektive Abbildung umkehrbar ist. Wenn X endlich ist, dann nennt man diese Abbildungen auch *Permutationen*. Diese werden später wieder auftauchen.

(d) Mit den Ergebnissen aus Kapitel 3 ergibt sich: Sei K ein Körper. Die invertierbaren $n \times n$-Matrizen bilden mit der Matrizenmultiplikation eine Gruppe. Sie notieren wir mit

$$GL_n(K) := \mathrm{Mat}_K(n \times n)^*$$

und nennen sie die *allgemeine lineare Gruppe der* $n \times n$*-Matrizen.*

Neutrale und inverse Elemente sind eindeutig, wir müssen für eine Gruppe also nur die Menge mit der Verknüpfung angeben. Die Beweise der Aussagen im folgenden Satz sind alle eine gute Übung im Umgang mit abstrakten Verknüpfungen.

Satz A.4. *Sei $(X, *)$ eine Gruppe mit neutralem Element e.*

(a) Dann gibt es höchstens ein neutrales Element bezüglich $*$.

(b) Dann hat jedes Element $x \in X$ höchstens ein Inverses x^{-1}.

(c) Wenn die Verknüpfung der Gruppe nicht kommutativ ist, dann dreht sich beim Invertieren die Reihenfolge um, es gilt also $(x * y)^{-1} = y^{-1} * x^{-1}$.

(d) Es gilt für alle $a \in G$, dass $(a^{-1})^{-1} = a$.

Beweis. (a) Seien e und $\tilde{e}$ zwei neutrale Elemente. Dann gilt $e = e * \tilde{e}$, weil $\tilde{e}$ neutral ist und andererseits $e * \tilde{e} = \tilde{e}$, weil e neutral ist. Insgesamt folgt $e = \tilde{e}$.

(b) Seien a, b zwei Inverse von x. Dann gilt

$$a = a * e = a * (x * b) = (a * x) * b = e * b = b,$$

also ist das Inverse eindeutig.

(c) Sei $z = x * y$ und wir suchen das Inverse z^{-1} von z. Es gilt

$$e = x * x^{-1} = x * e * x^{-1} = x * y * y^{-1} * x^{-1} = z * (y^{-1} * x^{-1}),$$

also $z^{-1} = y^{-1} * x^{-1}$.

(d) Klar nach Definition des Inversen, denn es gilt $a^{-1} * a = e = a * a^{-1}$.

□

Bemerkung A.5. Man nennt c) auch die *Socke-Schuh-Regel*: Beim Ausziehen von Schuhen und Socken muss man die Reihenfolge des Anziehens umkehren. Leicht zu sehen ist das bei bijektiven Abbildungen zwischen verschiedenen Räumen. Wenn $f\colon X \to Y$ und $g\colon Y \to Z$ beide bijektiv ist, dann muss man für die Umkehrabbildung von $g \circ f\colon X \to Z$ zuerst von Z mit g^{-1} nach Y laufen und dann erst mit f^{-1} zurück nach X. Es gilt dann also auch $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.
Vorsicht: Diese Verkettung von Abbildungen ist keine Verknüpfung im Sinne der Definition oben, denn die Verkettung $f \circ g$ ist nicht definiert.

In kommutativen Gruppen wird die Verknüpfung oft mit $+$ bezeichnet, das Inverse eines Elementes $a \in X$ dann mit $-a$. Wie bei Zahlen schreibt man dann auch oft $a - b$ als übersichtliche Kurzform von $a + (-b)$.
Allgemeine Gruppen schreibt man meistens eher multiplikativ, also mit $\cdot$ als Verknüpfungszeichen statt eines abstrakten Symbols wie $*$. Man muss dann nur immer im Kopf haben, um welche Verknüpfung es sich gerade handelt. Oft lässt man das Zeichen auch ganz weg und schreibt nur ab für $a * b$.

Satz A.6 (Kürzungsregel in Gruppen). *Sei $(G,*)$ eine Gruppe. Aus $xa = \tilde{x}a$ folgt bereits $x = \tilde{x}$; ebenso folgt aus $ay = a\tilde{y}$ bereits $y = \tilde{y}$. Insbesondere folgt aus $ax = a$ bereits $x = e$ und aus $ya = a$ bereits $y = e$.*

Beweis. Das folgt direkt aus der Eindeutigkeit der neutralen und inversen Elemente: Sei $xa = \tilde{x}a$. Dann gilt:

$$x = xe = xaa^{-1} = \tilde{x}aa^{-1} = \tilde{x}e = \tilde{x}$$

Und genauso für die andere Aussage. □

A.2. Homomorphismen

Nachdem wir nun Gruppen haben, suchen wir als nächstes Abbildungen zwischen ihnen. Diese sollen natürlich nicht nur Abbildungen zwischen den Mengen sein, sondern auch die Verknüpfungen respektieren.

Definition A.7. Seien $(G,*)$ und $(H,\star)$ Gruppen. Ein *Gruppenhomomorphismus* ist eine Abbildung $f\colon G \to H$, so dass gilt:

$$f(x * y) = f(x) \star f(y) \quad \text{für alle } x, y \in G. \tag{A.1}$$

Ein bijektiver Gruppenhomomorphismus heißt auch *Gruppenisomorphismus*.

Beachte, dass hier natürlich auf der linken Seite der Gleichung (A.1) die Verknüpfung $*$ von G steht und auf der rechten Seite die Verknüpfung $\star$ von H.
Wenn klar ist, dass es um Gruppen geht, lassen wir in allen Begriffen das Gruppen- weg und sagen nur Homomorphismus, Isomorphismus usw.

Lemma A.8. *Ein Gruppenhomomorphismus $f\colon (G,*) \to (H,\star)$, hat folgende Eigenschaften:*

(a) Es gilt $f(e_G) = e_H$.

(b) Für $x \in G$ gilt $f(x^{-1}) = f(x)^{-1}$.

Dabei sind natürlich e_G und e_H die neutralen Elemente in den jeweiligen Gruppen. Meistens schreibt man für beide nur e, denn aus dem Kontext ist oft klar, welches neutrale Element gemeint ist. Genauso schreiben wir beide Inverse mit der gleichen Notation, auch wenn es verschiedene Verknüpfungen sind.

Beweis. (a) Für alle $x \in G$ gilt

$$e_H \star f(x) = f(x) = f(e_G * x) = f(e) \star f(x),$$

also folgt die Behauptung mit der Kürzungsregel Satz A.6

(b) Mit a) gilt

$$e_H = f(e_G) = f(x * x^{-1}) = f(x) \star f(x^{-1})$$

und analog für $x^{-1} * x$, also ist $f(x^{-1})$ das Inverse von $f(x)$.

□

Dass Gruppenhomorphismen die Struktur erhalten, macht es auch einfacher, Eigenschaften der Abbildung zu untersuchen.

Definition A.9. Der *Kern* eines Gruppenhomomorphismus $f\colon (G,*) \to (H,\star)$ ist das Urbild des neutralen Elements. Man schreibt $\ker f = f^{-1}(e)$.

Satz A.10. *(a) Ein Gruppenhomomorphismus $f\colon (G,*) \to (H,\star)$ ist genau dann injektiv, wenn $\ker f = \{e\}$ gilt.*

(b) Die Umkehrabbildung eines Gruppenhomomorphismus $f\colon (G,*) \to (H,\star)$ ist wieder ein Gruppenhomomorphismus.

Wir müssen also nicht mehr das Urbild jedes Elementes anschauen, um Injektivität zu prüfen, sondern nur noch das des neutralen Elements.

Bemerkung A.11. Vorsicht: Die Notationen für Urbild, Umkehrabbildung und Inverses sind gleich oder sehr ähnlich. Die Umkehrabbildung wird mit f^{-1} bezeichnet, wir müssen von nun an also aufpassen und $f^{-1}(x)$ als Urbildmenge unter allgemeinen Abbildungen, $f^{-1}\colon H \to G$ als Abbildung für bijektives $f\colon G \to H$ und außerdem noch $f(x)^{-1} \in H$ als das Inverse von $f(x)$ aueinanderhalten.

Beweis. (a) Wenn f injektiv ist, dann hat jedes Element in H höchstens ein Urbild. Da wir bereits mit dem neutralen Element von G ein Urbild von $e \in H$ kennen, gilt also $\ker f = f^{-1}(e) = e$.

Sei nun $\ker f = \{e\}$ und es gelte $f(x) = f(y)$ für $x, y \in G$. Dann folgt

$$e = f(x) \star f(y)^{-1} = f(x * y^{-1}).$$

Also ist $x * y^{-1} \in f^{-1}(e) = \ker f$ und damit $x * y^{-1} = e$ in G. Da das Inverse eindeutig ist, gilt damit auch $x = y$.

(b) Sei $f\colon (G,*) \to (H,\star)$ ein Gruppenisomorphismus mit Umkehrabbildung $f^{-1}\colon H \to G$ und $x, y \in H$. Dann gilt

$$\begin{aligned} f^{-1}(x \star y) &= f^{-1}\Big(f(f^{-1}(x) \star f(f^{-1}(y)\Big) \\ &= f^{-1}\Big(f(f^{-1}(x) * f^{-1}(y))\Big) \\ &= f^{-1}(x) * f^{-1}(y). \end{aligned}$$

Dabei haben wir zuerst benutzt, dass $f(f^{-1}(a)) = a$ für alle $a \in H$, dann die Homomorphismuseigenschaft von von f und dann $f^{-1}(f(b)) = b$ für alle $b \in G$. □

Beispiel A.12. (a) Die Einbettungen $\mathbb{Z} \to \mathbb{Q} \to \mathbb{R} \to \mathbb{C}$ sind jeweils Gruppenhomomorphismen bezüglich der jeweiligen Addition.

(b) Das erste nicht triviale Beispiel sollte aus der Analysis bekannt sein, nämlich die Exponentialfunktion als Abbildung $\exp\colon (\mathbb{R}, +) \to (R \setminus \{0\}, \cdot)$. Das Exponentialgesetz $e^{x+y} = e^x \cdot e^y$ ist dabei gerade die Homomorphismuseigenschaft.

(c) Wenn man exp als Abbildung $\exp\colon (\mathbb{R}, +) \to (R_+, \cdot)$ in die positiven reellen Zahlen auffasst (mit Multiplikation ebenfalls eine Gruppe), dann ist das sogar ein Isomorphismus mit dem Logarithmus als Umkehrung.

B. Körper

Wenn wir uns die reellen Zahlen vor Augen führen, dann erkennen wir, dass es auf ihnen zwei assoziative Verknüpfungen, $+$ und $\cdot$, gibt, die über das Distributivgesetz miteinander verbunden sind. Das ist die Motivation für die folgende Definition:

Definition B.1. Ein *Körper* ist ein Tupel $(K, +, \cdot)$ bestehend aus einer Menge K mit zwei Verknüpfungen $+$ und $\cdot$, sodass gilt:

(a) $(K, +)$ ist eine kommutative Gruppe, deren neutrales Element wir mit 0 bezeichnen.

(b) $(K \setminus \{0\}, \cdot)$ ist eine kommutative Gruppe, dessen neutrales Element wir mit 1 bezeichnen.

(c) Das neutrale Element der Addition 0 und das neutrale Element der Multiplikation 1 sind verschieden:
$$0 \neq 1$$

(d) Es gelten die Distributivgesetze, also gilt für alle $a, b, c \in K$, dass
$$(a + b) \cdot c = (a \cdot c) + (b \cdot c), \quad \text{und} \quad a \cdot (b + c) = (a \cdot b) + (a \cdot c).$$

Wir bezeichnen die Verknüpfung $+$ als die *Addition*, die Verknüpfung $\cdot$ als die *Multiplikation* des Rings.

Beispiel B.2. Die wichtigsten Körper sind $\mathbb{Q}$, der Körper der rationalen Zahlen, $\mathbb{R}$, der Körper der reellen Zahlen, und $\mathbb{C}$, der Körper der komplexen Zahlen. In der Analysis geht es vor allem um reelle Zahlen, diese sind also unser wichtigstes Beispiel.

Bemerkung B.3. Wir benutzen in Körpern die aus der Schule bekannte Regel

> Punkt vor Strich, sodass wir also statt $(a \cdot b) + c$ auch kürzer $a \cdot b + c$ schreiben können.
> Weiter lassen wir oft den Punkt der zweiten Verknüpfung ganz weg, sodass wir also statt $a \cdot b$ auch einfach ab schreiben können.

Lemma B.4. *Sei $(R, +, \cdot)$ ein Körper.*

(a) Für alle $a \in R$ gilt $0 \cdot a = 0 = a \cdot 0$.

(b) Für alle $a \in R$ gilt $(-1) \cdot a = -a = a \cdot (-1)$.

Aus b) folgen dann sofort die üblichen Regeln für das Vorzeichen, es gilt also
$$a \cdot (-b) = -(a \cdot b) = (-a) \cdot b$$
für alle $a, b \in R$.

Beweis. (a) Mit der Kürzungsregel Satz A.6 in der additiven Gruppe $(R, +)$ folgt

$$0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a.$$

Die andere Aussage beweist man analog.

(b) Es gilt: $(-1) \cdot a + a = (-1) \cdot a + 1 \cdot a = (-1 + 1) \cdot a = 0$, also ist $(-1) \cdot a$ das eindeutige Inverse $-a$ von a.

Auch hier beweist man die zweite Aussage analog.

□

Körper sind also die Strukturen, die uns am vertrautesten sind und auch die schönsten Eigenschaften haben. Sie bilden die Grundlage der linearen Algebra, auf der dann weiter aufgebaut wird. Aber auch Ringe und Gruppen werden immer wieder von selbst auftauchen.
Natürlich gibt es auch Ringhomomorphismen und Körperhomomorphismen. Wie man sich leicht denken kann, sind diese durch die folgenden Eigenschaften charakterisiert:

Definition B.5. Seien $(R, +, \cdot)$ und $(S, +, \cdot)$ Körper. Ein *Körperhomomorphismus* von $(R, +, \cdot)$ nach $(S, +, \cdot)$ ist eine Abbildung $f\colon R \to S$, für die gilt:

(a) Die neutralen Elemente 1_R und 0_R aus R werden auf die neutralen Elemente 1_S und 0_S aus S abgebildet:

$$f(1_R) = 1_S \qquad \text{unr} \qquad f(0_R) = 0_S.$$

(b) Die Addition wird respektiert:

$$f(x + y) = f(x) + f(y) \quad \text{für alle} \quad x, y \in R.$$

(c) Die Multiplikation wird respektiert:

$$f(x \cdot y) = f(x) \cdot f(y) \quad \text{für alle} \quad x, y \in R.$$

Aus praktischen Gründen benutzt man in allen Körpern die Notation $+$ für die Addition und $\cdot$ (oder kein Symbol) für die Multiplikation. Wenn man ganz klar sein möchte, kann man auch jeweils den Körper an das Symbol schreiben, also $a +_R b$ für die Addition in R und $x +_S y$ für die in S und analog für die Multiplikation.