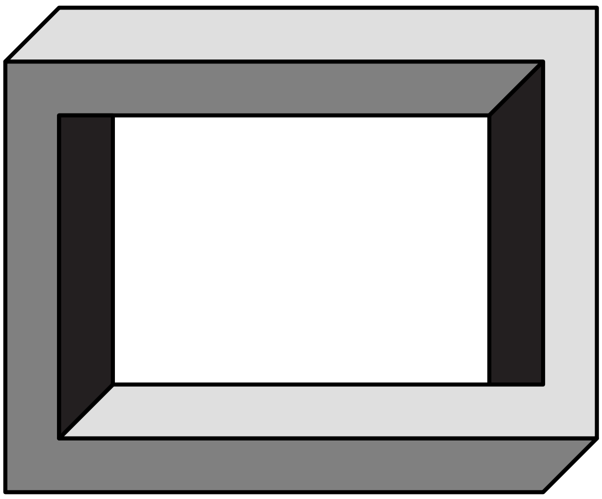
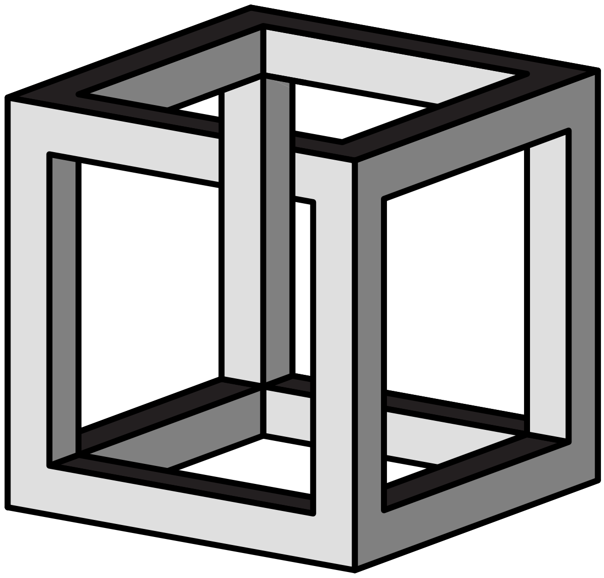
Lineare Algebra

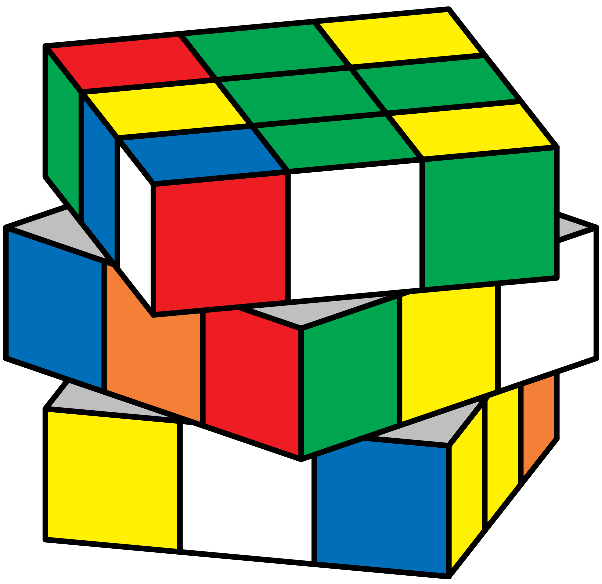
AOR Dr. Thoralf Räsch

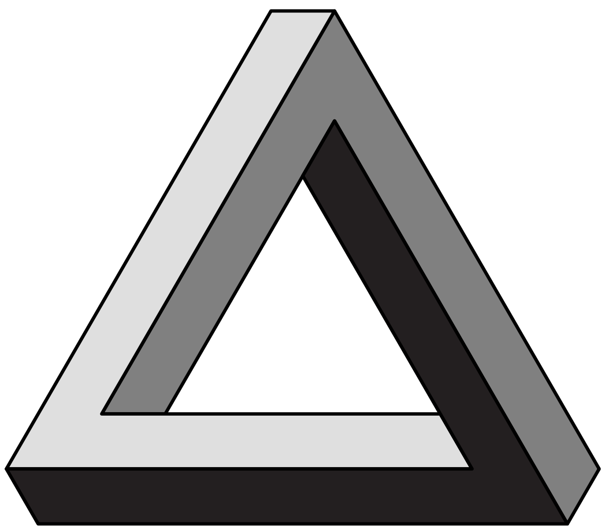
Mathematisches Institut
der Universität Bonn

Version 4.1.0
6. September 2022









Danksagung

Dieses Skript entstand während meiner Vorlesung im Sommersemester 2009. Es basiert auf Skript-Vorlagen von Herrn Professor Koepke aus vergangenen Semestern und wurde wegen der aktualisierten Inhalte der neuen Studiengänge erweitert bzw. angepasst.

⋆ ⋆ ⋆

Für die Erstbearbeitung im Sommersemester 2009 danke ich meiner Tutorin Anika Markgraf für die tatkräftige Unterstützung bei der Erstellung einer LaTeX-Vorlage der Tafelbilder, so dass ich überhaupt in diesem turbulenten Semester zeitlich in der Lage war, ein solches Skript zu verfassen. Darüber hinaus danke ich den Studenten der Informatik, die mir die Finanzierung einer Tutorin für diese Zwecke durch ihre Studienbeiträge überhaupt ermöglichten.

⋆ ⋆ ⋆

Ich danke Herrn Angel Koutev für die intensive und engagierte Unterstützung bei der Überarbeitung des Skriptes im Sommersemester 2011. Außerdem bedanke ich mich bei den Studierenden aus dem Kurs im Wintersemester 2011/12 für die hilfreichen Kommentare; insbesondere bei Lea Krüger, Cina Razzaghi, Roman Schmitz, Ruben Sparka und Martin Üding.

⋆ ⋆ ⋆

Und schließlich danke ich den Studierenden aus dem Wintersemester 2017/18 für die hilfreichen Kommentare. Insbesondere gilt hierbei mein Dank Herrn Peter Rosinsky für das unermüdliche Aufspüren von Typos im gesamten Skript. Des weiteren danke ich meiner Tutorin Christiane Langel für die anschließende geduldige Um- und Einarbeitung.

Inhaltsverzeichnis

1 Lineare Gleichungssysteme 9

Gleichungen . . . 9

Lineare Gleichungen in einer Unbekannten . . . 9

Quadratische Gleichungen in einer Unbekannten . . . 9

Lineare Gleichungssysteme . . . 10

Lineare Gleichungssyteme spezieller Gestalt . . . 11

Der Gauß-Algorithmus . . . 13

2 Vektorräume 16

Der Raum $\mathbb{K}^n$ . . . 16

Untervektorräume . . . 20

3 Gruppen, Ringe, Körper 24

Gruppen . . . 25

Ringe . . . 26

Teilbarkeit und modulare Arithmetik . . . 29

4 Komplexe Zahlen 32

Definition der komplexen Zahlen . . . 32

Komplexe quadratische Gleichungen . . . 34

Komplexe Zahlen als reelle Ebene . . . 34

Komplexe Zahlen als Polarkoordinaten . . . 35

Wurzelziehen aus komplexen Zahlen . . . 37

Anwendungen . . . 40

Jenseits der komplexen Zahlen: Quaternionen und Oktonionen . . . 40

5 Linearkombinationen und Basen 42

Linearkombinationen und lineare Unabhängigkeit . . . 42

Existenzsätze für Basen . . . 45

6 Lineare Abbildungen 56

Grundlagen Linearer Abbildungen . . . 56

Endliche Beschreibung von Homomorphismen . . . 62

Klassifikation endlich-dimensionaler Vektorräume . . . 63

Der Dimensionssatz . . . 64

7 Die Welt der Matrizen 68

Darstellende Matrizen . . . 68

Matrizenrechnung . . . 70

Lineare Abbildungen und Matrizen . . . 77

8 Anwendungen von Matrizen 80

Matrizen als Kodierwerkzeug: Lineare Codes . . . 80

Matrizen als Drehungen in der reellen Ebene ... 81
Matrizen als Spiegelungen in der reellen Ebene ... 85
Matrizen als elementare Umformung: Vertauschen von zwei Zeilen ... 87
Matrizen als elementare Umformung: Skalarmultiplikation einer Zeile ... 88
Matrizen als elementare Umformung: Addition zweier Zeilen ... 89

9 Matrizen und lineare Gleichungssysteme **91**
Geometrie der Lösungsmengen ... 91
gaußscher Algorithmus ... 96
Algorithmus zum Invertieren von Matrizen ... 96

10 Koordinatentransformation **99**
Basiswechsel ... 99
Darstellende Matrizen von linearen Abbildungen bezüglich beliebiger Basen ... 102
Darstellende Matrizen von Endomorphismen ... 106

11 Eigenwerte und Eigenvektoren **111**
Grundlegende Begriffe ... 111
Eigenwerte und Eigenvektoren an bekannten Beispielen ... 112
Eigenräume zu Eigenwerten ... 114
Lineare Unabhängigkeit von Eigenvektoren ... 114
Vorläufige Strategie des Diagonalisierens ... 117

12 Determinanten von Matrizen **120**
Determinanten von Matrizen ... 121
Determinaten und Gaußscher Algorithmus ... 123
Praktisch Determinaten berechnen ... 127
Die Cramersche Regel ... 130
Determinanten und Volumina ... 131

13 Charakteristische Polynome & Diagonalisierbarkeit **134**
MEGA-Beispiel zur Diagonalisierbarkeit ... 138
Ausblick Hauptachsentransformation ... 144

14 Euklidische Vektorräume **147**
Geometrische Begriffe in der reellen Ebene ... 147
Skalarprodukte ... 149
Normen ... 150
Orthogonalität ... 153
Orthonormalsysteme ... 154
Orthonormieren über nicht-triviale Skalarprodukte ... 158
Orthogonale Zerlegungen und Projektionen ... 160
Orthogonale Abbildungen ... 162

15 Über selbstadjungierte Endomorphismen und reell symmetrische Matrizen 166
Selbstadjungierte Endomorphismen . . . 166
Hauptachsentransformation mittels des Spektralsatzes . . . 168

16 Definitheit von Matrizen 169
Definition und Charakterisierungen der Definitheit . . . 169
Anwendung der Definitheit . . . 170

17 Die Jordansche Normalform – die Königsklasse der Darstellungsformen 172
Erste Gedanken zur Jordanschen Normalform . . . 172
Wie die jordansche Normalform aufgebaut ist und funktioniert . . . 174
Mit Jordanketten zum Ziel . . . 175
Anwendung der Jordanschen Normalform bei Differenzialgleichungen . . . 176

18 Lösungshinweise 180
Aufgaben zu Kapitel 1 . . . 180
Aufgaben zu Kapitel 2 . . . 181
Aufgaben zu Kapitel 6 . . . 181
Aufgaben zu Kapitel 7 . . . 182
Aufgaben zu Kapitel 8 . . . 183
Aufgaben zu Kapitel 9 . . . 183
Aufgaben zu Kapitel 10 . . . 185
Aufgaben zu Kapitel 11 . . . 186
Aufgaben zu Kapitel 12 . . . 188
Aufgaben zu Kapitel 13 . . . 188
Aufgaben zu Kapitel 14 . . . 189

Abbildungen 190

Definitionsverzeichnis 191

Stichwortverzeichnis 193

1. Lineare Gleichungssysteme

Gleichungssysteme finden wir überall in der Mathematik. In der Linearen Algebra werden wir uns vor allem mit den so genannten linearen Gleichungen beschäftigen.

Gleichungen. Wir arbeiten über einem Ring oder Körper. Formal lernen wir diese Begriffe erst im dritten Kapitel dieses Skripts kennen, ich werde die Sätze hier allerdings jetzt schon so allgemein halten. Bis zum dritten Kapitel können Sie sich jedes Mal bei Ring oder Körper die bekannten reellen Zahlen vorstellen – das reicht für den Anfang und das Verständnis.

Eine Gleichung ist eine Formel $t(x_1, \ldots, x_n) = b$, wobei t ein arithmetischer Term mit Unbekannten $x_1, \ldots, x_n$ ist.

Beispiel 1.1: Es gibt verschiedene Typen von Gleichungen:

(a) $a \cdot x = b$ lineare Gleichung mit einer Unbekannten
(b) $a \cdot x^2 + b \cdot x = 0$ quadratische Gleichung mit einer Unbekannten
(c) $a_n \cdot x^n + \ldots + a_1 \cdot x + a_0 = 0$ polynomielle Gleichung n-ten Grades mit einer Unbekannten
(d) $a \cdot x + b \cdot y = c$ lineare Gleichung mit zwei Unbekannten
(e) $a \cdot x^2 + b \cdot y^2 = r^2$ quadratische Gleichung mit zwei Unbekannten

Eine Gleichung ist eine Bedingung an die Unbekannten der Gleichung. Gesucht ist jeweils die Lösungsmenge $\{(x_1, \ldots, x_n) \mid t(x_1, \ldots, x_n) = b\}$. Die Bestimmung der Lösungsmenge verläuft so, dass die Gleichung $t(\bar{x}) = b$ zu einer übersichtlichen oder expliziten Formel $\varphi(\bar{x})$ umgeformt wird, sodass

$$\{(x_1, \ldots, x_n) \mid t(x_1, \ldots, x_n) = b\} = \{(x_1, \ldots, x_n) \mid \varphi(x_1, \ldots, x_n)\}$$

ist.

Lineare Gleichungen in einer Unbekannten. Wir betrachten die lineare Gleichung $a \cdot x = b$ über einem Körper $\mathbb{K}$.

Fall 1 $a \neq 0$

Dann ist $a \cdot x = b$ äquivalent zu $x = a^{-1} \cdot a \cdot x = a^{-1} \cdot b$. In diesem Fall ist die Lösungsmenge einelementig:

$$\left\{x \mid x = a^{-1} \cdot b\right\} = \left\{a^{-1} \cdot b\right\}$$

Fall 2 $a = 0$ und $b \neq 0$

Dann ist die Gleichung $a \cdot x = b$ immer unlösbar. Die Lösungsmenge ist dann die leere Menge: $\{x \mid a \cdot x = b\} = \emptyset$

Fall 3 $a = 0$ und $b = 0$

Dann ist die Gleichung $a \cdot x = b$ immer wahr. Die Lösungsmenge ist dann der gesamte Körper: $\{x \in \mathbb{K} \mid a \cdot x = b\} = \mathbb{K}$

Insgesamt erhalten wir folgende Gesamtlösungsmenge :

$$\left\{x \in \mathbb{K} \mid (a \neq 0 \text{ und } x = a^{-1} \cdot b) \text{ oder } (a = 0 \text{ und } b = 0)\right\}$$

Quadratische Gleichungen in einer Unbekannten. Wir betrachten die quadratischen Gleichungen $a \cdot x^2 + b \cdot x + c = 0$ über dem Körper $\mathbb{R}$.

Fall 1 $a = 0$

Dann ist die Gleichung linear.

Fall 2 $a \neq 0$

Dann ist die Gleichung äquivalent zu $x^2 + \frac{b}{a} \cdot x + \frac{c}{a} = 0$. Wir setzen abkürzend $p = \frac{b}{a}$ und $q = \frac{c}{a}$ an und betrachten somit die Gleichung $x^2 + p \cdot x + q = 0$. Nach quadratischer Ergänzung gilt: $x^2 + p \cdot x + \left(\frac{p}{2}\right)^2 = -q + \left(\frac{p}{2}\right)^2$ und nach binomischer Formel gilt:

$$\left(x + \frac{p}{2}\right)^2 = -q + \left(\frac{p}{2}\right)^2$$

Fall 2.1 $-q + \left(\frac{p}{2}\right)^2 \geq 0$

Dann kann man die Quadratwurzel ziehen und erhält

$$x + \frac{p}{2} = \pm\sqrt{-q + \left(\frac{p}{2}\right)^2}.$$

Fall 2.2 $-q + \left(\frac{p}{2}\right)^2 < 0$

Dann können wir über $\mathbb{R}$ keine Wurzel ziehen.

Allgemein erhalten wir als Lösungsmenge für die Gleichung $x^2 + p \cdot x + q = 0$:

$$\left\{ x \in \mathbb{R} \;\middle|\; -q + \left(\frac{p}{2}\right)^2 \geq 0 \quad \text{und} \quad \left(x = -\frac{p}{2} + \sqrt{-q + \left(\frac{p}{2}\right)^2} \quad \text{oder} \quad x = -\frac{p}{2} - \sqrt{-q + \left(\frac{p}{2}\right)^2} \right) \right\}$$

und die Lösungsmenge der Ausgangsgleichung $a \cdot x^2 + b \cdot x + c = 0$ wird deutlich komplizierter (logische Verknüpfung des linearen Falls mit dem gerade behandelten Fall $a \neq 0$).

Bemerkung 1.2: Über dem Grundkörper $\mathbb{C}$ gibt es auch für den letzten Fall Lösungen. Dies untersuchen wir in Kapitel 4.

⋆ ⋆ ⋆

Lineare Gleichungssysteme. Gerade wenn man Gleichungen in mehreren Unbekannten betrachtet, spielen Systeme von Gleichungen eine wichtige Rolle. Mittels einer Gleichung in mehreren Unbekannten beschreibt man einen Zusammenhang dieser Variablen. Je mehr Gleichungen man in Konjunktion zusammenfasst, desto einschränkender ist die beschriebene Bedingung und somit desto kleiner die Lösungsmenge.

Definition 1.3 – lineares Gleichungssystem:

Sei $\mathbb{K}$ ein Körper. Ein **lineares Gleichungssystem** (kurz **LGS**) für die Unbekannten $x_1, \dots, x_m$ ist ein System von linearen Gleichungen der Gestalt:

$$\begin{array}{ccc} a_{11}x_1+ & \dots & +a_{1m}x_m = b_1 \\ \vdots & & \vdots \\ a_{n1}x_1+ & \dots & +a_{nm}x_m = b_n \end{array}$$

Dabei sind die a_{ij} und b_i Körperelemente. Die a_{ij} werden als **Koeffizienten** des Systems bezeichnet. Wir bezeichnen ein solches System gelegentlich auch mit $Ax = b$.

Das System heißt **homogen**, falls $b_1 = b_2 = \ldots = b_n = 0$ gilt. In diesem Fall schreiben wir auch kurz: $Ax = 0$. Ansonsten heißt das System **inhomogen**.

Die **Lösungsmenge** des linearen Gleichungssystems $Ax = b$ ist

$$\text{Lös}(A,b) = \Big\{(x_1,\dots,x_m) \,\Big|\, x_1,\dots,x_m \in \mathbb{K} \text{ und } \\ a_{11}x_1 + \ldots + a_{1m}x_m = b_1, \dots,\; a_{n1}x_1 + \ldots + a_{nm}x_m = b_n\}$$

Das lineare Gleichungssystem heißt **lösbar**, wenn Lös(A,b) $\neq \emptyset$. Das lineare Gleichungssystem heißt **eindeutig lösbar**, wenn die Lösungsmenge genau ein Element enthält.

⋆ ⋆ ⋆

Lineare Gleichungssyteme spezieller Gestalt. Wir betrachten das lineare Gleichungssystem $Ax = b$ über dem Körper $\mathbb{K}$ in Diagonalgestalt:

$$\begin{array}{ccccc} a_{11}x_1 & & & & = b_1 \\ & a_{22}x_2 & & & = b_2 \\ & & \ddots & & \vdots \\ & & & a_{nn}x_n & = b_n \end{array}$$

Dieses lineare Gleichungssystem können wir schnell lösen:

Fall 3 Für alle $i \leq n$ gilt $a_{ii} \neq 0$.
Dann ist das lineare Gleichungssystem eindeutig lösbar mit der Lösungsmenge:

$$\text{Lös}(A,b) = \left\{ \left(\frac{b_1}{a_{11}}, \dots, \frac{b_n}{a_{nn}} \right) \right\}$$

Fall 4 Es gibt ein $i \leq n$ mit $a_{ii} = 0$ und $b_i \neq 0$.
Dann ist das gesamte lineare Gleichungssystem nicht lösbar, da die i-te Gleichung nicht lösbar ist. Es gilt also:

$$\text{Lös}(A,b) = \emptyset$$

Fall 5 Für alle $i \leq n$ mit $a_{ii} = 0$ gilt $b_i = 0$.
Dann ist das lineare Gleichungssystem lösbar mit :

$$\text{Lös}(A, b) = \left\{ (x_1, \ldots, x_n) \in \mathbb{K}^n \;\middle|\; \text{für } i \leq n \quad \text{mit} \quad a_{ii} \neq 0 \text{ gilt } x_i = \frac{b_i}{a_{ii}} \right\}$$

Bemerkung 1.4: Ein homogenes LGS ist immer lösbar!
Betrachten wir nun ein lineares Gleichungssystem $Ax = b$ über einem Körper $\mathbb{K}$ in Zeilenstufenform (ZSF):

$$\begin{array}{cccccccc}
a_{1j(1)}x_{j(1)} & + & & \cdots & & +\ldots+ & a_{1m}x_m & = & b_1 \\
& & a_{2j(2)}x_{j(2)} & + & \cdots & +\ldots+ & a_{2m}x_m & = & b_2 \\
& & & & \ddots & \vdots & & & \vdots \\
& & & & a_{ij(i)}x_{j(i)} & +\ldots+ & a_{im}x_m & = & b_i
\end{array}$$

mit $i \leq n$ und Pivot-Koeffizienten, oder Pivot-Elementen,

$$a_{1j(1)}, a_{2j(2)}, \ldots, a_{ij(i)} \neq 0.$$

Menge der Lösungen $(x_1, \ldots, x_m)$:

Die Menge der Lösungen ergibt sich durch Auflösen des Systems von unten nach oben und von hinten nach vorne. Die letzte Zeile im linearen Gleichungssystem sieht wie folgt aus:

$$a_{ij(i)}x_{j(i)} + a_{i(j(i)+1)}x_{j(i)+1} + \ldots + a_{im}x_m = b_i$$

Das heißt $x_{j(i)+1}, \ldots, x_m \in \mathbb{K}$ können beliebige Werte annehmen, da sie durch keine weitere Gleichung eingeschränkt werden. Daraus ergibt sich $x_{j(i)}$ wie folgt:

$$x_{j(i)} = \frac{1}{a_{ij(i)}} \cdot (b_i - a_{ij(i)+1}x_{j(i)+1} - \ldots - a_{im}x_m)$$

Außerdem verfahren Sie mit den restlichen Gleichungen analog. Die Lösungsmenge ist dann:

$$\text{Lös}(A, b) = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in \mathbb{K}^m \;\middle|\; \forall\, k = 1, \ldots, i \text{ gilt } x_{j(k)} = \frac{1}{a_{kj(k)}} \cdot \left(b_k - \sum_{l=j(k)+1}^{m} a_{kl}x_l \right) \right\}$$

Beispiel 1.5: Betrachte das lineare Gleichungssystem über $\mathbb{R}$:

$$\begin{aligned} 1 \cdot x_1 + 2 \cdot x_2 + 3 \cdot x_3 + 4 \cdot x_4 &= 2 \\ 2 \cdot x_3 + 3 \cdot x_4 &= 1 \end{aligned}$$

Dann sind die Unbekannten x_4 und x_2 frei wählbar, während sich x_3 und x_1 in Abhängigkeit von x_4 und x_2 ergeben. Die Lösungsmenge ist:

$$\begin{aligned} \mathbb{L} &= \left\{ (x_1, \ldots, x_4) \in \mathbb{R}^4 \;\middle|\; x_3 = \frac{1 - 3x_4}{2} \text{ und } x_1 = 2 - 2x_2 - 3x_3 - 4x_4 \right\} \\ &= \left\{ (x_1, \ldots, x_4) \in \mathbb{R}^4 \;\middle|\; x_3 = \frac{1 - 3x_4}{2} \text{ und } x_1 = \frac{1}{2} - 2x_2 + \frac{1}{2}x_4 \right\} \end{aligned}$$

Man kann dieses lineare Gleichungssystem zum Beispiel auch als System über dem Körper $\mathbb{Z}_5$ auffassen:
In $\mathbb{Z}_5$ ist $\frac{1}{2} = 2^{-1} = 3$ (denn in $\mathbb{Z}_5$ gilt: $2 \cdot 3 = 1$), $\frac{3}{2} = 3 \cdot \frac{1}{2} = 3 \cdot 3 = 4$ und $-2 = 3$. Damit ist die Lösungsmenge zu $\mathbb{Z}_5$:

$$\mathbb{L} = \left\{ (x_1, \ldots, x_4) \;\middle|\; x_3 = 3 + x_4 \text{ und } x_1 = 3 + 3x_2 + 3x_4 \right\}$$

⋆ ⋆ ⋆

Der Gauß-Algorithmus. Ziel ist es ein beliebiges lineares Gleichungssystem logisch äquivalent in Zeilenstufenform zu überführen, da wir ein lineares Gleichungssystem in Zeilenstufenform einfacher lösen können. Grundlage der Transformation ist die folgende Beobachtung:

Satz 1.6:

Sei $\mathbb{K}$ ein Körper und seien $a_1, \ldots, a_m, b, a'_1, \ldots, a'_m, b' \in \mathbb{K}$. Weiterhin sei $\lambda \in \mathbb{K}, \lambda \neq 0$ und $\mu \in \mathbb{K}$. Dann gilt:

(a) für $x_1, \ldots, x_m \in \mathbb{K}$ ist die Gleichung

$$a_1x_1 + \ldots + a_mx_m = b \tag{1}$$

äquivalent zu der Gleichung

$$\lambda a_1x_1 + \ldots + \lambda a_mx_m = \lambda b \tag{2}$$

(b) für $x_1, \ldots, x_m \in \mathbb{K}$ ist:

$$a_1x_1 + \ldots + a_mx_m = b \text{ \underline{und} } a'_1x_1 + \ldots + a'_mx_m = b'$$

äquivalent zu:

$$a_1x_1 + \ldots + a_mx_m = b \text{ \underline{und} } (a'_1 + \mu a_1)x_1 + \ldots + (a'_m + \mu a_m)x_m = b' + \mu b$$

Beweis:

(a) Seien $x_1, \ldots, x_m \in \mathbb{K}$. Wir zeigen die Äquivalenz (1) $\Leftrightarrow$ (2).
Zunächst zeigen wir „$\Rightarrow$":
Es gelte also (1). Dann gilt insbesondere:

$$\lambda \cdot (a_1x_1 + \ldots + a_mx_m) = \lambda \cdot b$$

Durch wiederholtes Anwenden des Distributivgesetzes in $\mathbb{K}$ ergibt sich die Gleichung (2).
Es bleibt „$\Leftarrow$" zu zeigen:
Es gelte nun (2). Nach Anwendung des Distributivgesetzes gilt:

$$\lambda \cdot (a_1x_1 + \ldots + a_mx_m) = \lambda b$$

da $\lambda \neq 0$, existiert das multiplikative Inverse. Es folgt

$$\lambda^{-1} \cdot \lambda(a_1x_1 + \ldots + a_mx_m) = \lambda^{-1} \cdot \lambda \cdot b$$

und daraus folgt (1).

Die Aussage aus (b) kann analog gezeigt werden. ☒

Die Äquivalenz in (b) erlaubt es, eine Unbekannte x_i zu eliminieren, indem μ so gewählt wird, dass gilt:

$$a_i' + \mu a_i = 0$$

Beispiel 1.7:

$$\left|\begin{array}{rcrcr} x & + & 2y & = & 5 \\ 2x & + & 3y & = & 8 \end{array}\right| \quad \rightsquigarrow \quad \left|\begin{array}{rcrcr} -2x & - & 4y & = & -10 \\ 2x & + & 3y & = & 8 \end{array}\right| \quad \rightsquigarrow \quad \left|\begin{array}{rcrcr} -2x & - & 4y & = & -10 \\ & - & y & = & -2 \end{array}\right|$$

Also $y = 2$ und $x = 1$.
Wir können das System auch kürzer fassen:

$$\begin{array}{cc|c} 1 & 2 & 5 \\ 2 & 3 & 8 \end{array} \quad \rightsquigarrow \quad \begin{array}{cc|c} -2 & -4 & -10 \\ 2 & 3 & 8 \end{array} \quad \rightsquigarrow \quad \begin{array}{cc|c} -2 & -4 & -10 \\ 0 & -1 & -2 \end{array}$$

Beispiel 1.8: Schauen wir uns noch ein Beispiel an:

$$\begin{aligned} 2x_1 + 3x_2 - x_3 &= 12 \\ x_1 - 3x_2 + 4x_3 &= -12 \\ 5x_1 - 6x_2 + 11x_3 &= -24 \end{aligned}$$

$$\begin{array}{ccc|c} 2 & 3 & -1 & 12 \\ 1 & -3 & 4 & -12 \\ 5 & -6 & 11 & -24 \end{array} \rightsquigarrow \begin{array}{ccc|c} 1 & -3 & 4 & -12 \\ 2 & 3 & -1 & 12 \\ 5 & -6 & 11 & -24 \end{array} \rightsquigarrow \begin{array}{ccc|c} 1 & -3 & 4 & -12 \\ 0 & 9 & -9 & 36 \\ 0 & 9 & -9 & 36 \end{array} \rightsquigarrow \begin{array}{ccc|c} 1 & -3 & 4 & -12 \\ 0 & 1 & -1 & 4 \\ 0 & 0 & 0 & 0 \end{array}$$

Dieses System entspricht den Gleichungen:

$$\begin{aligned} x_1 - 3x_2 + 4x_3 &= -12 \\ x_2 - x_3 &= 4 \end{aligned}$$

Hierbei ist x_3 beliebig wählbar: $x_2 = 4 + x_3$ und $x_1 = -x_3$, also ergibt sich als Lösungsmenge:

$$\mathbb{L} = \big\{(x_1, x_2, x_3) \;\big|\; x_2 = 4 + x_3 \;,\; x_1 = -x_3\big\}$$

Beispiel 1.9: Betrachte nachfolgendes Gleichungssystem:

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

oder etwas kompakter geschrieben erhalten wir:

$$\left(\begin{array}{cccc|c} 1 & 2 & 4 & 3 & 3 \\ 1 & 2 & 5 & 6 & 3 \\ 2 & 4 & 8 & 7 & 4 \\ 2 & 4 & 8 & 6 & 6 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 2 & 4 & 3 & 3 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 2 & 4 & 3 & 3 \\ 0 & 0 & 1 & 0 & 6 \\ 0 & 0 & 0 & 1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Wir können jetzt einfach die zweite mit der dritten und dann die dritte mit der vierten Spalte vertauschen um unsere gewünschte Zeilenstufenform zu erhalten. Aber Vorsicht! Das Vertauschen von Spalten entspricht einer Umbenennung der Variablen und muss deshalb immer sorgfältig notiert und beim Aufschreiben der Lösungsmenge wieder rückgängig gemacht werden. In unserem Fall notieren wir beispielsweise (2,3) und (3,4) und erhalten:

$$\left(\begin{array}{cccc|c} 1 & 4 & 3 & 2 & 3 \\ 0 & 1 & 0 & 0 & 6 \\ 0 & 0 & 1 & 0 & -2 \end{array}\right)$$

Für dieses Gleichungssystem bekommen wir zunächst die Lösungsmenge:

$$\mathbb{L} = \left\{(x_1, x_2, x_3, x_4) \;\middle|\; x_1 = -15 - 2x_4 \;,\; x_2 = 6 \;,\; x_3 = -2\right\}$$

Jetzt müssen wir noch die Spaltenvertauschungen rückgängig machen. Dazu tauschen wir zuerst den Wert von x_3 mit x_4 und anschließend x_2 mit x_3, wodurch wir die richtige Lösungsmenge für unser Ausgangsproblem erhalten:

$$\mathbb{L} = \left\{(x_1, x_2, x_3, x_4) \;\middle|\; x_1 = -15 - 2x_2 \;,\; x_3 = 6 \;,\; x_4 = -2\right\}$$

Beispiel 1.10: Schauen wir uns einmal folgendes LGS an:

$$\begin{aligned} x_1 + x_2 + 0x_3 + x_4 &= 1 \\ 2x_1 + x_2 + x_3 + 0x_4 &= 2 \\ 3x_1 + 0x_2 + 0x_3 + 2x_4 &= 2 \\ 4x_1 + x_2 + 0x_3 + 3x_4 &= 3 \end{aligned}$$

Dann errechnen wir wieder die Lösungsmenge:

$$\left(\begin{array}{cccc|c} 1 & 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 0 & 2 \\ 3 & 0 & 0 & 2 & 2 \\ 4 & 1 & 0 & 3 & 3 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 1 & 0 & 1 & 1 \\ 0 & -1 & 1 & -2 & 0 \\ 0 & -3 & 0 & -1 & -1 \\ 0 & -3 & 0 & -1 & -1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cccc|c} 1 & 1 & 0 & 1 & 1 \\ 0 & 3 & 0 & 1 & 1 \\ 0 & 0 & 3 & -5 & 1 \end{array}\right)$$

$$\rightsquigarrow \left(\begin{array}{cccc|c} 3 & 0 & 0 & 2 & 2 \\ 0 & 3 & 0 & 1 & 1 \\ 0 & 0 & 3 & -5 & 1 \end{array}\right)$$

Die daraus resultierende Lösungsmenge ist:

$$\mathbb{L} = \left\{(x_1, x_2, x_3, x_4) \;\middle|\; x_1 = 2\frac{1 - x_4}{3} \;,\; x_2 = \frac{1 - x_4}{3} \;,\; x_3 = \frac{1 + 5x_4}{3}\right\}$$

⋆ ⋆ ⋆

Aufgabe 1.11: Berechnen Sie die Lösungsmenge des zugehörigen homogenen LGS aus dem vorherigen Beispiel. Welchen Zusammenhang gibt es zwischen diesen beiden Lösungsmengen?

Aufgabe 1.12: Konstruieren Sie Lineare Gleichungssysteme mit den folgenden Lösungsmengen:

$$\begin{aligned} \mathbb{L}_1 &= \left\{(x_1, x_2, x_3, x_4) \;\middle|\; x_1 = 3x_4 \;,\; x_2 = 6x_4 \;,\; x_3 = -2x_4\right\} \\ \mathbb{L}_2 &= \left\{(x_1, x_2, x_3, x_4) \;\middle|\; x_1 = 5\lambda + \rho \;,\; x_2 = 6\lambda - 2\rho \;,\; x_3 = -2\lambda \;,\; x_4 = -2\rho, \text{ für } \lambda, \rho \in \mathbb{R}\right\} \end{aligned}$$

2. Vektorräume

Lineare Gleichungssysteme in den Unbekannten $x_1, \ldots, x_n$ über einem Körper $\mathbb{K}$ haben Folgen von Elementen von $\mathbb{K}$ als Lösungen. Die Struktur dieser Folgen werden wir im Folgenden näher betrachten.

Der Raum $\mathbb{K}^n$.

Definition 2.1 – mathematischer Körper:

Sei $\mathbb{K} = (\mathbb{K}, +, \cdot, 0, 1)$ ein Körper und $n \in \mathbb{N}$. Definiere die Struktur $\mathbb{K}^n := (\mathbb{K}^n, +, \cdot, 0)$ durch:

(a) $\mathbb{K}^n = \left\{(x_1, \ldots, x_n) \;\middle|\; x_1 \in \mathbb{K}, \ldots, x_n \in \mathbb{K}\right\}$. Die Elemente von $\mathbb{K}^n$ heißen Vektoren.
(b) Für $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathbb{K}^n$ definiere die Vektoraddition „+" durch

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

(c) Für $(x_1, \ldots, x_n) \in \mathbb{K}^n$ und $\lambda \in \mathbb{K}$ definiere die Skalarmultiplikation „·" durch

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

(d) $0 = (0, \ldots, 0) \in \mathbb{K}^n$ ist der Nullvektor.
(e) Für $x = (x_1, \ldots, x_n)$ definiere das additive Inverse durch

$$-x = (-x_1, \ldots, -x_n)$$

(f) Für $(x_1, \ldots, x_n)$ und $(y_1, \ldots, y_n)$ definiere die Vektorsubtraktion „−" durch

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

Beispiel 2.2: Die Struktur $\mathbb{K}^n$ wird vielfältig angewendet:

(a) Die Menge $(\mathbb{Z}_2)^n$ ist die Menge der 0-1-Folgen der Länge n, das heißt aller Bitfolgen der Länge n. Ein Element von $(\mathbb{Z}_2)^8$ bezeichnet man als Byte.
(b) Die Menge $(\mathbb{Z}_m)^n$ kann man als die Menge der n-elementigen Wörter über dem Alphabet $\mathbb{Z}_m = \{0, 1, \ldots, m-1\} \mathrel{\hat{=}} \{A, B, C, \ldots\}$ auffassen. So kann man dies etwa für die CAESAR-Verschlüsselung benutzen:

$$(x_1, \ldots, x_n) \mapsto (x_1, \ldots, x_n) + (c, \ldots, c)$$

Etwa `INFORMATIK` + `EEEEEEEEEE` = `NSKTWRFYNP`
(c) Analytische Geometrie im $\mathbb{R}^2$ (vgl. Abbildung 1):

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

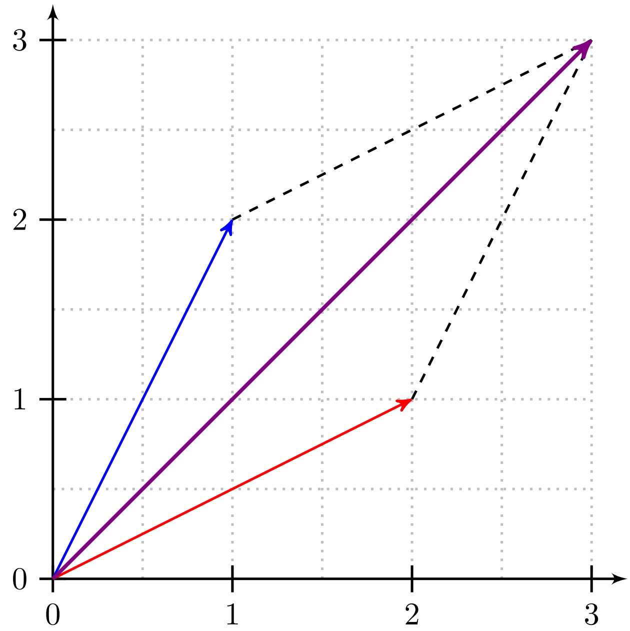


Abbildung 1: Analytische Geometrie im $\mathbb{R}^2$

Motiviert durch die Beispiele studieren wir die Gesetze des vektoriellen Rechnens:

Satz 2.3:

Sei $\mathbb{K} = (\mathbb{K}, +, \cdot, 0, 1)$ ein Körper und $n \in \mathbb{N}$.

(a) Für $x, y, z \in \mathbb{K}^n$ gilt
$$(x + y) + z = x + (y + z) \qquad \text{(Assoziativgesetz)}$$
(b) Für $x, y \in \mathbb{K}^n$ gilt
$$x + y = y + x \qquad \text{(Kommutativgesetz)}$$
(c) Für $x \in \mathbb{K}^n$ gilt
$$x + 0 = x \qquad \text{(Neutralität der Null)}$$
(d) Für $x \in \mathbb{K}^n$ gilt
$$x + (-x) = 0 \qquad \text{(Existenz des Inversen)}$$
(e) Für $\lambda, \mu \in \mathbb{K}$ und $x \in \mathbb{K}^n$ gilt
$$(\lambda \cdot \mu) \cdot x = \lambda \cdot (\mu \cdot x) \qquad \text{(Assoziativgesetz)}$$
(f) Für $x \in \mathbb{K}^n$ gilt
$$1 \cdot x = x \qquad \text{(Neutralität der 1)}$$
(g) Für $\lambda \in \mathbb{K}$ und $x, y \in \mathbb{K}^n$ gilt
$$\lambda \cdot (x + y) = \lambda x + \lambda y \qquad \text{(1. Distributivgesetz)}$$
(h) Für $\lambda, \mu \in \mathbb{K}$ und $x \in \mathbb{K}^n$ gilt
$$(\lambda + \mu) \cdot x = \lambda x + \mu x \qquad \text{(2. Distributivgesetz)}$$

Beweis: Klar mit den Rechengesetzen im Körper. Beispielsweise für das Kommutativgesetz:
Seien $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ und $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$. Dann gilt:

$$x + y = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} = \begin{pmatrix} y_1 + x_1 \\ \vdots \\ y_n + x_n \end{pmatrix} = y + x$$

$\boxtimes$

Beachten Sie, dass das Symbol 0 im Satz 2.3 formell für den Nullvektor im Vektorraum steht, also für 0_V, während die 1 dem Einselement im Körper entspricht.

Beispiel 2.4: Sei $M = \{f \mid f : [-1,1] \to \mathbb{R}\}$ die Menge der reellwertigen Funktionen auf $[-1,1]$. Für $f, g \in M$ setze $(f+g)(x) := f(x) + g(x)$.
Für $\lambda \in \mathbb{R}$ und $f \in M$ setze $(\lambda \cdot f)(x) = \lambda \cdot f(x)$. Dann gilt: $(f+g) \in M$ und $(\lambda \cdot f) \in M$. Definiere $0 \in M$ als $0(x) = 0$. Dann erfüllt die Struktur $(M, +, \cdot, 0)$ die im Satz beschriebenen Gesetze:

Satz 2.5:

Sei die Menge $M = \{f \mid f : [-1,1] \to \mathbb{R}\}$ mit den o.g. Operationen gegeben.

(a) Für $f, g, h \in M$ gilt
$(f+g)+h = f+(g+h)$ (Assoziativgesetz)

(b) Für $f, g \in M$ gilt
$f+g = g+f$ (Kommutativgesetz)

(c) Für $f \in M$ gilt
$f+0 = f$ (Neutralität der Null)

(d) Für $f \in M$ gilt
$f+(-f) = 0$ (Existenz des Inversen)

(e) Für $\lambda, \mu \in \mathbb{R}$ und $f \in M$ gilt
$(\lambda \cdot \mu) \cdot f = \lambda \cdot (\mu \cdot f)$ (Assoziativgesetz)

(f) Für $f \in M$ gilt
$1 \cdot f = f$ (Neutralität der 1)

(g) Für $\lambda \in \mathbb{R}$ und $f, g \in M$ gilt
$\lambda \cdot (f+g) = \lambda f + \lambda g$ (1. Distributivgesetz)

(h) Für $\lambda, \mu \in \mathbb{R}$ und $f \in M$ gilt
$(\lambda + \mu) f = \lambda f + \mu f$ (2. Distributivgesetz)

Beweis: Zur Kommutativität:

$$(f+g) = (g+f) \qquad \Leftrightarrow \qquad \forall x \left((f+g)(x) = (g+f)(x) \right)$$

Sei $x \in [-1,1]$ gegeben. Dann gilt:

$$\begin{aligned} (f+g)(x) &= f(x) + g(x) && \text{(Nach Definition von } +) \\ &= g(x) + f(x) && \text{(Kommutativität in } \mathbb{R}) \\ &= (g+f)(x) \end{aligned}$$

Rest klar, mit den Rechengesetzen im Körper $\mathbb{R}$. ⊠

Definition 2.6 – Vektorraum:

Es sei $\mathbb{K}$ ein Körper und sei $V = (V, +, \cdot)$ eine Struktur mit einer Abbildung „+" (genannt Addition)

$$+ : V \times V \to V \qquad (x, y) \mapsto x + y$$

und einer Abbildung „·" (genannt „skalare Multiplikation" oder „Skalarmultiplikation")

$$\cdot\ : \mathbb{K} \times V \to V \qquad (\lambda, x) \mapsto \lambda \cdot x$$

Dann ist $V = (V, +, \cdot)$ ein $\mathbb{K}$- **Vektorraum** oder **Vektorraum über** $\mathbb{K}$, wenn die folgenden Axiome gelten:

(a) Für $x, y, z \in V$ gilt
$(x + y) + z = x + (y + z)$ (Assoziativgesetz)

(b) Für $x, y \in V$ gilt
$x + y = y + x$ (Kommutativgesetz)

(c) Es gibt ein „0" $\in V$, sodass für $x \in V$ gilt
$x + 0 = x$ (Neutralität der Null)

(d) Für $x \in V$ gibt es ein $y \in V$, sodass
$x + y = 0$ (Existenz des Inversen)

(e) Für $\lambda, \mu \in \mathbb{K}$ und $x \in V$ gilt
$(\lambda \cdot \mu) \cdot x = \lambda \cdot (\mu \cdot x)$ (Assoziativgesetz)

(f) Für $x \in V$ gilt
$1 \cdot x = x$ (Neutralität der 1)

(g) Für $\lambda \in \mathbb{K}$ und $x, y \in V$ gilt
$\lambda \cdot (x + y) = \lambda x + \lambda y$ (1. Distributivgesetz)

(h) Für $\lambda, \mu \in \mathbb{K}$ und $x \in V$ gilt
$(\lambda + \mu)x = \lambda x + \mu x$ (2. Distributivgesetz)

Satz 2.7:

Für alle Vektorräume $V = (V, +, \cdot)$ gibt es genau ein Element, das der Nullvektor ist.

Beweis: Angenommen es gäbe die Nullvektoren 0 und $0'$ in V. Dann gilt nach den Axiomen:

$$\begin{aligned} 0 &= 0 + 0' && \text{((c) für 0')} \\ &= 0' + 0 && \text{(nach (b))} \\ &= 0' && \text{((c) für 0)} \end{aligned}$$

⊠

Nach diesem Satz können wir ohne Einschränkung den Nullvektor mit in die Signatur aufnehmen und schreiben daher im Folgenden stets:

Sei $V = (V, +, \cdot, 0)$ ein Vektorraum.

Satz 2.8:

Für alle Vektorräume $V = (V, +, \cdot, 0)$ gilt: Für x aus V existiert genau ein Element y, welches das additive Inverse in V ist, d.h. $x + y = 0$ erfüllt, das wir mit $-x$ bezeichnen.

Beweis: Betrachte ein beliebiges $x \in V$. Angenommen a und b sind additive Inverse von x in V, d.h. $x + a = 0$ und $x + b = 0$.

$$\begin{aligned}
a &= a + 0 && \text{(Neutrales Element)} \\
&= a + (x + b) && \text{(b als inverses Element)} \\
&= (a + x) + b && \text{(Assoziativität)} \\
&= (x + a) + b && \text{(Kommutativität)} \\
&= 0 + b && \text{(a als inverses Element)} \\
&= b + 0 && \text{(Kommutativität)} \\
&= b
\end{aligned}$$

⊠

Definition 2.9 – Schreibweise des inversen Elements der Addition:

Schreibe $x - y$ statt $x + (-y)$.

⋆ ⋆ ⋆

Untervektorräume. Sie wissen bereits, dass folgende Inklusionskette gilt: $\mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$. Alle drei Strukturen sind Körper und somit selbst für sich Vektorräume. Man kann sich beispielsweise jetzt fragen, ob etwa die reellen Zahlen als $\mathbb{R}$-Vektorraum als Teilmenge des $\mathbb{R}$-Vektorraums $\mathbb{C}$ die einzige solche Teilmenge ist, die unter den Vektorraumoperationen abgeschlossen ist. Solche Teilmengen (Teilräume) nennen wir Untervektorräume.

Definition 2.10 – Unter(vektor)raum:

Sei $V = (V, +, \cdot, 0)$ ein $\mathbb{K}$-Vektorraum und sei $U \subseteq V$ eine Teilmenge von V. Dann ist U ein Unter(vektor)raum (kurz: UVR) von V, wenn $U \neq \emptyset$ und für alle $x, y \in U$ und alle $\lambda \in \mathbb{K}$ gilt: $x + y \in U$ und $\lambda \cdot x \in U$. (Abgeschlossenheit bzgl. $+$ und $\cdot$)

Satz 2.11:

Sei U ein Untervektorraum eines fixierten $\mathbb{K}$-Vektorraums $V = (V, +, \cdot, 0)$. Dann ist $U = (U, +, \cdot, 0)$ auch ein $\mathbb{K}$-Vektorraum, wobei die Operationen $+$ und $\cdot$ auf $U \times U$ bzw. $\mathbb{K} \times U$ eingeschränkt werden.

Beweis: Wir beweisen die geforderten Bedingungen:

(a) $0 \in U$. (eigentlich $0_V \in U$)
Wähle $x \in U$ als Element von $U \neq \emptyset$. Dann gilt:

$0 \cdot x = (0+0) \cdot x = 0 \cdot x + 0 \cdot x$.
Weiterhin gilt:

$$\begin{aligned} 0 &= 0 \cdot x - 0 \cdot x \\ &= (0 \cdot x + 0 \cdot x) - 0 \cdot x && \text{(Nach obiger Gleichung)} \\ &= 0 \cdot x + (0 \cdot x - 0 \cdot x) && \text{(Assoziativität)} \\ &= 0 \cdot x + 0 \\ &= 0 \cdot x \end{aligned}$$

Da $0 \in \mathbb{K}$ und $x \in U$ und U abgeschlossen unter $\cdot$ ist, gilt: $0 = 0 \cdot x \in U$

(b) Für $x \in U$ ist $-x \in U$:

$$\begin{aligned} 0 &= 0 \cdot x \\ &= (1 + (-1)) \cdot x \\ &= 1 \cdot x + (-1) \cdot x \\ &= x + (-1) \cdot x \end{aligned}$$

Nach dem obigen Satz ist das additive Inverse eindeutig, so dass gelten muss: $-x = (-1) \cdot x \in U$.

Aus diesen Existenzaussagen folgt, dass U ein $\mathbb{K}$-Vektorraum ist. ⊠

Satz 2.12:

Es seien U_1, U_2 Untervektorräume des $\mathbb{K}$-Vektorraums V. Dann ist $U_1 \cap U_2$ ein Untervektorraum von V.

Beweis: Wir hatten gezeigt, dass $0 \in U_1$ und $0 \in U_2$. Damit ist $0 \in U_1 \cap U_2$ und somit $U_1 \cap U_2 \neq \emptyset$. Betrachte beliebige $x, y \in U_1 \cap U_2$ und $\lambda \in \mathbb{K}$. Da U_1 und U_2 Untervektorräume sind, ist $x + y \in U_1, \lambda \cdot x \in U_1$ und $x + y \in U_2, \lambda \cdot x \in U_2$. Damit ist $x + y \in U_1 \cap U_2$ und auch $\lambda \cdot x \in U_1 \cap U_2$. ⊠

Wir können nun Varianten der Eingangsfrage des Abschnittes noch einmal aufgreifen:

Beispiel 2.13: Es gibt drei grundsätzliche Arten von Unterräumen der reellen Ebene. Zunächst finden wir sofort den trivialen Unterraum, der nur aus dem Nullvektor besteht, bzw. dem gesamten Raum. Echt dazwischen liegen nur noch die Ursprungsgeraden.

Beispiel 2.14: Die reelle Achse als reeller Vektorraum hat nur die beiden trivialen Untervektorräume (Nullraum und $\mathbb{R}$ selbst) als Unterräume. Unterräume des reellen Vektorraums $\mathbb{C}$ sind daher zunächst $\{0\}$, $\mathbb{R}$ und $\mathbb{C}$, aber auch alle Teilmengen, die Ursprungsgeraden entsprechen, wenn man die komplexen Zahlen als reelle Ebene auffasst.

Wir rechnen noch ein weiteres konkretes Beispiel durch und definieren:

Definition 2.15 – Abbildung:

Sei $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$ die Menge aller Funktionen von $\mathbb{R}$ nach $\mathbb{R}$, das heißt:

$$\mathrm{Abb}(\mathbb{R}, \mathbb{R}) = \left\{ f \;\middle|\; f : \mathbb{R} \to \mathbb{R} \right\}$$

Wir wissen bereits, dass wir $\mathrm{Abb}(\mathbb{R}, \mathbb{R})$ als Vektorraum ansehen können, das heißt es gilt:

$$(f + g)(x) = f(x) + g(x)$$

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

Beispiel 2.16: Setze $U = \{f \in \text{Abb}(\mathbb{R}, \mathbb{R}) \mid f(x) = f(-x) \text{ für } x \in \mathbb{R}\}$. Dann ist $U \subseteq \text{Abb}(\mathbb{R}, \mathbb{R})$ ein Untervektorraum.
Wir zeigen die Unterraumeigenschaften:

(a) $U \neq \emptyset$:
Für die Nullabbildung $0(x) = 0$ gilt insbesondere $0(x) = 0(-x)$, also auch $0 \in U$.

(b) Abgeschlossenheit bzgl. „+“:
Zu zeigen gilt: $f, g \in U \Rightarrow f + g \in U$. Seien also $f, g \in U$. Es gilt also $f(x) = f(-x)$ und $g(x) = g(-x)$.

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

(c) Abgeschlossenheit bzgl. „·“:
Sei $f \in U$, also $f(x) = f(-x)$ und $\lambda \in \mathbb{R}$.

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

Beispiel 2.17: Wir versuchen jetzt an einem konkreten Beispiel den Schnitt zweier Untervektorräume des $\mathbb{R}^3$ zu berechnen:

$$U_1 := \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \;\middle|\; \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \lambda_1, \lambda_2 \in \mathbb{R} \right\}$$

$$U_2 := \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \;\middle|\; \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \lambda_3 \begin{pmatrix} 6 \\ 3 \\ 6 \end{pmatrix}, \quad \lambda_3 \in \mathbb{R} \right\}$$

Dann können wir den Schnitt $U_1 \cap U_2$ mit dem Ansatz:

$$\lambda_1 \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \lambda_3 \begin{pmatrix} 6 \\ 3 \\ 6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

berechnen und gelangen zu dem folgenden homogenen LGS:

$$\left(\begin{array}{ccc|c} 1 & 1 & -6 & 0 \\ 0 & 1 & -3 & 0 \\ 2 & 0 & -6 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 1 & -6 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & -2 & 6 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 1 & -6 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Jetzt können wir die Lösungsmenge dieses LGS angeben:

$$\mathbb{L} = \{(\lambda_1, \lambda_2, \lambda_3) \mid \lambda_1 = \lambda_2 = 3\lambda_3\}$$

und erhalten unseren gewünschten Schnitt $U_1 \cap U_2 = U_2$, indem wir uns anschauen, was man alles mit den aus $\mathbb{L}$ erhaltenen Koeffizienten an Vektoren erzeugen kann. Da unter diesen Bedingungen das x_3 frei wählbar ist, muss der komplette Untervektorraum U_2 den Schnitt bilden.

⋆ ⋆ ⋆

Aufgabe 2.18: Welche der folgenden Teilmengen sind Untervektorräume des $\mathbb{R}^3$?

$$U_1 := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \;\middle|\; 3x + 7y - 8z = 0, \quad x, y, z \in \mathbb{R} \right\}$$

$$U_2 := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \;\middle|\; x + y + z = 1, \quad x, y, z \in \mathbb{R} \right\}$$

$$U_3 := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \;\middle|\; x \cdot y \cdot z = 0, \quad x, y, z \in \mathbb{R} \right\}$$

$$U_4 := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \;\middle|\; x^2 + y^2 = z^2, \quad x, y, z \in \mathbb{R} \right\}$$

3. Gruppen, Ringe, Körper

Wenn wir uns die bekannten Zahlbereiche mit den dazugehörigen Operationen anschauen, so können wir Eigenschaften erkennen, die immer wieder auftauchen – etwa die Kommutativität der Addition. Um hier Ordnung in die Begrifflichkeiten zu bringen, führen wir der Reihe nach Strukturbegriffe ein, die immer wieder gemeinsam auftretende Eigenschaften zusammenfassen.

Definition 3.1 – Verknüpfung:

Für eine Menge A heißt eine Abbildung $\circ\colon A \times A \to A,\ (a, a') \mapsto a \circ a'$ Verknüpfung auf A. Eine Verknüpfung heißt

- assoziativ, wenn für alle $a, b, c \in A$ gilt $(a \circ b) \circ c = a \circ (b \circ c)$;
- kommutativ, wenn für alle $a, b \in A$ gilt $a \circ b = b \circ a$.

Beispiel 3.2: Die Addition und Multiplikation sind assoziative und kommutative Verknüpfungen auf $\mathbb{R}$. Die Subtraktion $a \circ b = a - b$ ist weder assoziativ noch kommutativ, denn es gilt

$$\begin{aligned} 0 \circ (0 \circ 1) \;=\; 0 - (0 - 1) \;=\; 1 \;\neq\; -1 \;=\; (0 - 0) - 1 \;=\; (0 \circ 0) \circ 1 \\ 1 \circ 0 \;=\; 1 - 0 \;=\; 1 \;\neq\; -1 \;=\; 0 - 1 \;=\; 0 \circ 1 \end{aligned}$$

Man kann mit vollständiger Induktion zeigen, dass bei einer assoziativen Verknüpfung von n Elementen Klammern weggelassen werden können.

Definition 3.3 – neutrales Element:

Ist A eine Menge mit einer Verknüpfung „$\circ$“, so heißt $e \in A$ neutrales Element oder Einselement, wenn für alle $a \in A$

$$e \circ a = a \circ e = a.$$

Gibt es so ein e, so ist es eindeutig bestimmt, denn ist e' auch ein neutrales Element, so gilt $e' = e' \circ e = e$.

Beispiel 3.4: Sei M eine beliebige (nicht-leere) Menge und $A = \{f \mid f : M \to M\}$, so ist $\mathrm{id}_M : M \to M,\ m \mapsto m$ das neutrale Element bezüglich der Komposition.
Dies sieht man wie folgt: Für alle $f \in A$ gilt $f \circ \mathrm{id}_M = f$ und $\mathrm{id}_M \circ f = f$, da für alle $m \in M$ gilt:

$$(f \circ \mathrm{id}_M)(m) = f\left(\mathrm{id}_M(m)\right) = f(m)$$

und

$$(\mathrm{id}_M \circ f)(m) = \mathrm{id}_M\left(f(m)\right) = f(m).$$

Definition 3.5 – inverses Element:

Es sei A eine Menge mit einer Verknüpfung „$\circ$“ und einem neutralen Element e. Ist $a \in A$, so heißt $b \in A$ inverses Element oder Inverses von a, wenn $a \circ b = b \circ a = e$ gilt. Existiert für a ein Inverses, so heißt a invertierbar oder Einheit.

Bemerkung 3.6: Bezüglich „+" besitzt $\mathbb{Z}$ als neutrales Element die Zahl $0 = 0$ und es ist sogar jedes $a \in \mathbb{Z}$ invertierbar. Bezüglich „·" ist in $\mathbb{Z}$ die Zahl $1 = 1$ das neutrale Element, aber nur die Zahlen 1 und -1 sind invertierbar.

Satz 3.7:

Ist A eine Menge mit einer assoziativen Verknüpfung „$\circ$" und einem neutralen Element e, so gibt es zu jedem invertierbaren Element $a \in A$ genau ein $b \in A$ mit $a \circ b = e$ und $b \circ a = e$.

Beweis: Sind b, b' Inverse von a, so gilt:

$$b = e \circ b = (b' \circ a) \circ b = b' \circ (a \circ b) = b' \circ e = b'.$$

$\boxtimes$

Das unter den Voraussetzungen von Satz 3.7 eindeutig bestimmte Inverse von a wird mit a^{-1} bezeichnet.

Gruppen.

Definition 3.8 – abelsche Gruppe:

Ist G eine Menge und „$\circ$" eine Verknüpfung auf G, so heißt G **Gruppe** (bzgl. „$\circ$"), wenn gilt

- **(a)** „$\circ$" ist assoziativ.
- **(b)** Es existiert ein neutrales Element $e \in G$.
- **(c)** Jedes $g \in G$ ist invertierbar.

Ist weiterhin „$\circ$" kommutativ, so heißt G **abelsche Gruppe**.

Beispiel 3.9: Offensichtlich bildet $\mathbb{Z}$ bezüglich „+" eine abelsche Gruppe. Darüber hinaus bildet $\mathbb{R}$ zusammen mit „·" keine Gruppe; dagegen ist $\mathbb{R}^\times := \mathbb{R} \setminus \{0\}$ bezüglich „·" eine abelsche Gruppe.

Achtung!

Eine Verknüpfung „$\circ$" auf G bedeutet, dass $\circ : G \times G \to G$, insbesondere ist eine Verknüpfung immer abgeschlossen innerhalb der Menge und kann also nicht aus der Grundmenge G herausführen.

Satz 3.10:

Ist M eine Menge mit einer assoziativen Verknüpfung „$\circ$" und einem neutralen Element e, so ist die Menge $G = \left\{g \;\middle|\; g^{-1} \in M\right\}$ eine Gruppe.

Ist M selbst eine Gruppe, so gilt offenbar $M = G$.

Beweis von Satz 3.10: Wir zeigen zunächst, dass die auf M gegebene Verknüpfung „$\circ$" auch eine Verknüpfung auf G ist – eigentlich zeigen wir dies für die Einschränkung von „$\circ$" auf $G \times G$. Wir weisen also nach, dass für alle $a, b \in G$ stets gilt $a \circ b \in G$: Seien a, b aus G, so sind diese invertierbar, also folgt $a^{-1}, b^{-1} \in M$. Wegen

$$(a \circ b) \circ \left(b^{-1} \circ a^{-1}\right) \quad = \quad \left(a \circ \left(b \circ b^{-1}\right)\right) \circ a^{-1} \quad = \quad (a \circ e) \circ a^{-1} \quad = \quad a \circ a^{-1} \quad = \quad e$$

und analog

$$\left(b^{-1} \circ a^{-1}\right) \circ (a \circ b) \quad = \quad \left(b^{-1} \circ \left(a^{-1} \circ a\right)\right) \circ b \quad = \quad \left(b^{-1} \circ e\right) \circ b \quad = \quad b^{-1} \circ b \quad = \quad e$$

ist $a \circ b$ invertierbar, also $a \circ b \in G$.

Die restlichen drei Gruppeneigenschaften folgen unmittelbar: Die Aussage $(a \circ b) \circ c = a \circ (b \circ c)$ gilt sogar für alle $a, b, c \in M$, insbesondere für $G \subseteq M$. Wegen $e \circ e = e$ gilt $e \in G$ und e ist damit insbesondere das neutrale Element von G. Schließlich gilt für alle $a \in G$ jeweils $a \circ a^{-1} = e$ und $a^{-1} \circ a = e$. Also ist auch a^{-1} in M invertierbar, das heißt $a^{-1} \in G$ und a ist das Inverse von a^{-1}.
☒(Satz 3.10)

Wir haben sogar etwas mehr im letzten Beweis gezeigt: Für invertierbare Gruppenelemente a, b gilt $\left(a^{-1}\right)^{-1} = a$, sowie $(a \circ b)^{-1} = b^{-1} \circ a^{-1}$.

Ringe. Wir führen nun eine etwas mächtigere Struktur ein – den Ring.

Definition 3.11 – Ring:

Ist R eine Menge mit den Verknüpfungen „+“ und „·“, so heißt $(R, +, \cdot)$ Ring, wenn gilt

(a) R ist bzgl. „+“ eine abelsche Gruppe,
(b) „·“ ist assoziativ,
(c) Es gelten die Distributivgesetze, das heißt, für alle $a, b, c \in R$ gilt:

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c) \text{ und } (a + b) \cdot c = (a \cdot c) + (b \cdot c)\,.$$

Ist weiterhin „·“ kommutativ, so heißt R kommutativer Ring.
Hat ein Ring R ein neutrales Element bezüglich „·“, so heißt R Ring mit Eins und dann heißt weiterhin $a \in R$ invertierbar oder Einheit, wenn a bezüglich „·“ invertierbar ist.

Wir führen folgende Vereinbarungen ein:

- Punktrechnung geht vor Strichrechnung, das heißt, um Produkte werden keine Klammern gesetzt.
- Das neutrale Element eines Ringes bezüglich „+“ wird mit dem Symbol „0“ bezeichnet.
- Ist R ein Ring mit Eins, so wird das neutrale Element bezüglich „·“ mit dem Symbol „1“ bezeichnet. In diesem Falle ist

  $$E(R) := \left\{a \in R \;\middle|\; a \text{ Einheit in } R\right\}$$

  eine Gruppe bezüglich der Multiplikation (siehe Satz 3.7) und wird als *Einheitengruppe* bezeichnet.
- Für alle $a \in R$ bezeichnet $-a$ das (additive) Inverse von a bezüglich „+“. Für alle $a, b \in R$ definiert man weiterhin $a - b := a + (-b)$.

Wir wissen, dass die ganzen Zahlen $\mathbb{Z}$ bezüglich der gewöhnlichen Addition und Multiplikation einen kommutativen Ring mit Eins bilden und es gilt $E(\mathbb{Z}) = \{1, -1\}$.

Beachten Sie, dass die Angabe der neutralen Elemente (sofern existent) mit den Symbolen 0 und 1 lediglich ein symbolischer Akt ist. Welche Ring-Elemente dann letztendlich hinter diesen Symbolen stehen, geht dabei aus dem Kontext hervor.

Beispiel 3.12: Wir können etwa eine einelementige Menge $R := \{0\}$ betrachten und auf eindeutige Art und Weise eine Addition und Multiplikation einführen, nämlich:

$$0 + \underbrace{0}_{=0} = 0 \qquad \text{und} \qquad 0 \cdot \underbrace{0}_{=1} = 0$$

Wie Sie aber sehen können, ist in diesem Fall das einzige Element sowohl das neutrale Element der Addition als auch der Multiplikation, so dass in diesem Fall (in diesem Ring R) $0 = 1$ gilt. Solche pathologischen Fälle möchten wir insbesondere mit der nächsten Definition ausschließen:

Definition 3.13 – Körper:

Ist $\mathbb{K}$ ein kommutativer Ring mit Eins, so heißt K Körper, wenn $0 \neq 1$ und jedes $a \in \mathbb{K}$ mit $a \neq 0$ eine Einheit ist.

Bemerkung 3.14: Wir können feststellen, dass bezüglich der gewöhnlichen Addition und Multiplikation die bekannten Zahlbereiche $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$ (und $\mathbb{C}$) kommutative Ringe mit Eins sind. Darüber hinaus sind $\mathbb{Q}, \mathbb{R}$ (und $\mathbb{C}$) sogar Körper, $\mathbb{Z}$ aber nicht.

Wir werden jetzt exemplarisch interessante Eigenschaften in Ringen untersuchen. Beachten Sie, dass die Eigenschaften, die wir jetzt nachweisen, in beliebigen (kommutativen) Ringen (mit Eins) gelten.

Satz 3.15:

Sei $(R, +, \cdot, 0, 1)$ ein kommutativer Ring mit Eins. Dann erfüllt R für beliebige $\lambda \in R$ die folgenden Bedingungen:

(a) $0 \cdot \lambda = 0$
(b) $(-1) \cdot \lambda = -\lambda$
(c) $-(-\lambda) = \lambda$
(d) $(-1) \cdot (-1) = 1$

Beweis: Wir beweisen die erste Behauptung mit der folgenden Gleichungskette:

$$\begin{aligned}
0 \cdot \lambda &= \lambda \cdot 0 && \text{(Kommutativität der Mult.)} \\
&= \lambda \cdot 0 + 0 && \text{(Neutrales Element der Add.)} \\
&= \lambda \cdot 0 + \Big(\lambda \cdot 0 + \big(-(\lambda \cdot 0)\big)\Big) && \text{(Inverses Element der Add.)} \\
&= (\lambda \cdot 0 + \lambda \cdot 0) + \big(-(\lambda \cdot 0)\big) && \text{(Assoziativität der Add.)} \\
&= \lambda \cdot (0 + 0) + \big(-(\lambda \cdot 0)\big) && \text{(Distributivität)} \\
&= \lambda \cdot 0 + \big(-(\lambda \cdot 0)\big) && \text{(Neutrales Element der Add.)} \\
&= 0 && \text{(Inverses Element der Add.)}
\end{aligned}$$

$$\begin{aligned}
0 \cdot \lambda &= \lambda \cdot 0 && \text{(Kommutativität der Mult.)} \\
&= \lambda \cdot 0 + 0 && \text{(Neutrales Element der Add.)} \\
&= \lambda \cdot 0 + \Big(\lambda \cdot 0 + \big(-(\lambda \cdot 0)\big)\Big) && \text{(Inverses Element der Add.)} \\
&= (\lambda \cdot 0 + \lambda \cdot 0) + \big(-(\lambda \cdot 0)\big) && \text{(Assoziativität der Add.)}
\end{aligned}$$

$$\begin{aligned}
&= \lambda \cdot (0+0) + \big(-(\lambda \cdot 0)\big) && \text{(Distributivität)}\\
&= \lambda \cdot 0 + \big(-(\lambda \cdot 0)\big) && \text{(Neutrales Element der Add.)}\\
&= 0 && \text{(Inverses Element der Add.)}
\end{aligned}$$

Für die zweite Behauptung lässt sich ähnlich argumentieren:

$$\begin{aligned}
(-1)\cdot\lambda &= (-1)\cdot\lambda + 0 && \text{(Neutrales Element der Add.)}\\
&= (-1)\cdot\lambda + \big(\lambda + (-\lambda)\big) && \text{(Inverses Element der Add.)}\\
&= \big((-1)\cdot\lambda + \lambda\big) + (-\lambda) && \text{(Assoziativität der Add.)}\\
&= \big((-1)\cdot\lambda + \lambda\cdot 1\big) + (-\lambda) && \text{(Neutrales Element der Mult.)}\\
&= \big(\lambda\cdot(-1) + \lambda\cdot 1\big) + (-\lambda) && \text{(Kommutativität der Mult.)}\\
&= \Big(\lambda\cdot\big((-1)+1\big)\Big) + (-\lambda) && \text{(Distributivität)}\\
&= \lambda\cdot 0 + (-\lambda) && \text{(Inverses Element der Add.)}\\
&= 0\cdot\lambda + (-\lambda) && \text{(Kommutativität der Mult.)}\\
&= 0 + (-\lambda) && \text{(nach (a))}\\
&= -\lambda && \text{(Neutrales Element der Add.)}
\end{aligned}$$

Die restlichen beiden Teile bleiben eine Übungsaufgabe. ⊠

Abschließend stellen wir fest, dass wir auch die bekannten binomischen Formeln allgemein wie folgt in beliebige Ringe übertragen können:

Satz 3.16:

Sei $(R, +, \cdot, 0, 1)$ ein kommutativer Ring mit Eins. Dann gilt für alle λ und μ in R

(a) $(\lambda + \mu)\cdot(\lambda+\mu) = \lambda\cdot\lambda + (1+1)\lambda\cdot\mu + \mu\cdot\mu$
(b) $(\lambda + (-\mu))\cdot(\lambda + (-\mu)) = \lambda\cdot\lambda + (-(1+1)\cdot\lambda\cdot\mu) + \mu\cdot\mu$
(c) $(\lambda+\mu)\cdot(\lambda+(-\mu)) = \lambda\cdot\lambda + (-\mu\cdot\mu)$

Beweis: Wir beschränken uns darauf, die erste Aussage zu beweisen. Die restlichen beiden bleiben eine (leichte) Übungsaufgabe.

Es gilt

$$\begin{aligned}
(\lambda+\mu)\cdot(\lambda+\mu) &= (\lambda+\mu)\cdot\lambda + (\lambda+\mu)\cdot\mu\\
&= \lambda\cdot(\lambda+\mu) + \mu\cdot(\lambda+\mu)\\
&= \lambda\cdot\lambda + (\lambda\cdot\mu + \mu\cdot\lambda) + \mu\cdot\mu\\
&= \lambda\cdot\lambda + (\mu\cdot\lambda + \mu\cdot\lambda) + \mu\cdot\mu\\
&= \lambda\cdot\lambda + (\mu\cdot\lambda\cdot 1 + \mu\cdot\lambda\cdot 1) + \mu\cdot\mu\\
&= \lambda\cdot\lambda + \big(\mu\cdot\lambda\cdot(1+1)\big) + \mu\cdot\mu\\
&= \lambda\cdot\lambda + (1+1)\cdot\mu\cdot\lambda + \mu\cdot\mu
\end{aligned}$$

⊠

Teilbarkeit und modulare Arithmetik. Bisher haben wir die bekannten Zahlbereiche betrachtet, die alle unendlich sind. Wir werden uns jetzt endlichen Ringen und Körpern widmen, die insbesondere in der Informatik zu Kodierungszwecken eine Anwendung finden. Aber zunächst werden wir grundlegende Begriffe klären und arbeiten dabei innerhalb der ganzen Zahlen.

Definition 3.17 – Teiler:

Seien m und n ganze Zahlen. Dann definieren wir „m teilt n“, wenn es ein $l \in \mathbb{Z}$ gibt mit $m \cdot l = n$. Wir schreiben in diesem Fall „m *ist Teiler von* n“ bzw. $m \mid n$.

Satz 3.18:

Es gelten folgende einfache Eigenschaften:

(a) Für jede ganze Zahl m gilt $1 \mid m$ und $m \mid m$.
(b) Für alle ganzen Zahlen a, b und c mit $a \mid b$ und $b \mid c$ gilt $a \mid c$.

Beweis:
Die Aussagen in (a) sind offensichtlich wahr, da $1 \cdot m = m$ und $m \cdot 1 = m$. Wir zeigen die zweite Behauptung: Wegen $a|b$ existiert ein $l_1 \in \mathbb{Z}$ mit $a \cdot l_1 = b$. Wegen $b|c$ existiert $l_2 \in \mathbb{Z}$ mit $b \cdot l_2 = c$. Also gilt $c = b \cdot l_2 = (a \cdot l_1) \cdot l_2 = a \cdot (l_1 \cdot l_2) = a \cdot l$, wobei $l = l_1 \cdot l_2$. Damit ist $a \mid c$ bewiesen. ⊠

Definition 3.19 – Primzahl:

Sei p eine natürliche Zahl, dann heißt p eine Primzahl, wenn $p \neq 0$, $p \neq 1$ und für alle t mit $t|p$ gilt: $t = 1$ oder $t = p$.

Der nächste Satz wird uns die so genannte Division mit Rest garantieren, die wir im Folgenden ausnutzen werden.

Satz 3.20:

Sei n eine natürliche Zahl, $n \neq 0$. Dann gibt es für jede ganze Zahl a eindeutig bestimmte ganze Zahlen q und r, so dass $a = q \cdot n + r$ gilt, mit $0 \leq r < n$.

Beweis: Sei n eine natürliche Zahl. Wie Sie sich leicht klarmachen können, ist

$$\mathbb{Z} = \bigcup_{q \in \mathbb{Z}} \left\{ m \in \mathbb{Z} \mid q \cdot n \leq m \leq (q+1) \cdot n \right\}.$$

Sei nun $a \in \mathbb{Z}$, so gibt es also ein $q \in \mathbb{Z}$ mit $q \cdot n \leq a \leq (q+1) \cdot n$. Definieren wir für dieses q nun $r := a - q \cdot n$, so gilt $a = q \cdot n + r$ und $0 \leq r < n$. Letzteres ergibt sich sofort, wenn Sie von $q \cdot n \leq a \leq (q+1) \cdot n$ die Zahl $q \cdot n$ subtrahieren. ⊠

Das heißt, dass wir jede natürliche Zahl durch n mit Rest r teilen können; dabei bezeichnet r den *Rest von a modulo n* und dieser wird mit $[a]_n$ bezeichnet:

$$[a]_n := \left\{ r \mid a = q \cdot n + r \quad q \in \mathbb{Z}, n \in \mathbb{N}, 0 \leq r < n \right\}$$

Wenn n im Kontext fixiert ist, schreiben wir manchmal auch nur $[a]$.

Definition 3.21 – modulo:

Seien a, b ganze Zahlen und n eine natürliche Zahl. Dann sei

$$a \equiv b \bmod n,$$

wenn n ein Teiler der Differenz $a - b$ ist.

Wenn „$a \equiv b \bmod n$" gilt, spricht man dies als „a ist kongruent zu b modulo n". Insbesondere bedeutet es, dass a und b denselben Rest modulo n haben.

Mithilfe der Division mit Rest, können wir nun Operationen auf einer endlichen Menge (der Reste) einführen.

Definition 3.22 – Addition modulo und Multiplikation modulo:

Sei n eine natürliche Zahl, $n \geq 2$. Wir definieren die Menge $\mathbb{Z}_n := \{0, 1, \ldots, n-1\}$ und auf ihr folgende zwei Verknüpfungen:

- Die Addition modulo n ist die folgende Operation

$$\oplus_n : \mathbb{Z}_n \times \mathbb{Z}_n \to \mathbb{Z}_n, \quad a \oplus_n b = [a+b]_n .$$

- Die Multiplikation modulo n ist die folgende Operation

$$\otimes_n : \mathbb{Z}_n \times \mathbb{Z}_n \to \mathbb{Z}_n, \quad a \otimes_n b = [a \cdot b]_n .$$

Beispiel 3.23: Diese modulare Arithmetik kennen Sie aus Ihrem täglichen Leben. Betrachten Sie etwa die Tageszeit. Dann ist dies formal die Addition modulo 24; so ist etwa $20 \oplus_{24} 8 = 4$ (Uhr). Oder betrachten Sie einfache 8-Bit Mikroprozessoren; diese benutzen die Arithmetik modulo $2^8 (= 256)$. Dabei gilt $100 \oplus_{256} 156 = 0$, das heißt, dass das additive Inverse von 100 bezüglich dieser Addition 156 ist und somit die Gleichung „$-100 = 156$" gilt.

Satz 3.24:

Für beliebige n ist $(\mathbb{Z}_n, \oplus, \otimes, 0, 1)$ ein kommutativer Ring mit Eins. Dabei ist $\mathbb{Z}_n$ genau dann ein Körper, wenn n eine Primzahl ist.

Beweis: Der Nachweis der Ringeigenschaften ist einfaches Nachrechnen. Für die zusätzliche Körpereigenschaft betreffs existierender Inversen zeigen wir nun beide geforderten Richtungen der Äquivalenz: Sei zunächst $n = p$ eine Primzahl, sowie $a \in \mathbb{Z}_p, a \neq 0$. Wegen $a \neq 0$ ist p kein Teiler von a. Da p eine Primzahl ist, haben a und p keine gemeinsamen Teiler außer der Eins und es existieren $x, y \in \mathbb{Z}$ mit $ax + py = 1$. (Die Existenz von x, y garantiert uns ein weiteres Theorem, auf welches wir hier nicht weiter eingehen werden) Insbesondere heißt das aber, dass gilt $ax = -py + 1$ und somit auch

$$a \cdot [x] = [a] \cdot [x] = [ax] = \big[-py + 1 \big] = \big[p \cdot (-y) + 1 \big] = 1.$$

Somit ist $[x]_p$ das gewünschte multiplikative Inverse zu a in $\mathbb{Z}_p$. Wegen $1 \neq 0$ ist $\mathbb{Z}_p$ damit ein Körper.

Sei nun andererseits n keine Primzahl. Gilt $n = 1$, so folgt $0 = 1$, das heißt, das neutrale Element der Addition ist gleich dem neutralen Element der Multiplikation. Somit ist $\mathbb{Z}_1$ kein Körper.

Sei daher im Folgenden $n > 1$. Dann existieren $a, b \in \mathbb{N}$ mit $n = a \cdot b$, $1 < a < n$ und $1 < b < n$, da n keine Primzahl ist. Also gilt $a \neq 0$, $b \neq 0$ und $a \otimes b = [a] \otimes [b] = [a \cdot b] = [n] = 0 \notin E(\mathbb{Z}_n)$. Folglich ist a oder b keine Einheit in $\mathbb{Z}_n$, da wir schon aus dem Beweis von Satz 3.10 wissen, dass sonst $(ab)^{-1} = b^{-1}a^{-1}$ gelten würde und damit 0 invertierbar wäre. Damit ist $\mathbb{Z}_n$ insbesondere kein Körper. $\boxtimes$

4. Komplexe Zahlen

Wie Sie wissen, besitzen quadratische Gleichungen nicht immer eine (reelle) Lösung, wie etwa die Gleichung

$$x^2 + 1 = 0 \text{ oder äquivalent } x^2 = -1.$$

Um u.a. trotzdem mit Lösungen von solchen Gleichungen rechnen zu können, führte Euler 1777 eine neue Zahl i ein. Für dieses i gilt dann per Definition:

$$\mathrm{i}^2 = -1 \quad \text{bzw.} \quad \mathrm{i} = \sqrt{-1}.$$

Man bezeichnet diese neue Zahl i als *imaginäre Einheit*. Offensichtlich ist i keine reelle Zahl.

Definition der komplexen Zahlen. Wir führen ganz naiv, ausgehend von dieser neuen Zahl, die so genannten komplexen Zahlen ein, indem wir zunächst mit i rechnen, als würden die Gesetze gelten, die wir von den reellen Zahlen her kennen. So können wir beispielsweise Vielfache dieser imaginären Einheit bilden, indem wir eine reelle Zahl b an i heran multiplizieren, etwa $b \cdot \mathrm{i}$ oder kurz $b\mathrm{i}$ bzw. $\mathrm{i}b$. Weiterhin können wir gemischte Summen bilden: Die Summe aus einer reellen Zahl a und einer rein imaginären Zahl $b \cdot \mathrm{i}$ heißt dann *komplexe Zahl*. Die Menge der komplexen Zahlen wird mit $\mathbb{C}$ bezeichnet und wir definieren noch den Realteil und Imaginärteil einer komplexen Zahl:

$$\mathbb{C} := \{z \mid z = a + \mathrm{i} \cdot b;\ a, b \in \mathbb{R}\} \qquad \text{wobei} \quad \mathfrak{Re}(z) := a \quad \text{und} \quad \mathfrak{Im}(z) := b$$

Wir vereinbaren, dass zwei komplexe Zahlen genau dann gleich sind, wenn sie sowohl im Realteil, als auch im Imaginärteil übereinstimmen. Insbesondere gilt, dass für $0 = b = \mathfrak{Im}(z)$ die komplexe Zahl z reell ist; auf diese Weise haben wir unsere bekannten Zahlen in den neuen Zahlbereich eingebettet: $\mathbb{R} \subsetneq \mathbb{C}$.

Eigentlich haben wir bisher nur die Zahlenmengen ineinander eingebettet; es wäre sehr schön, wenn sich die Operationen auch übertragen lassen würden, das heißt, dass die komplexe Addition und Multiplikation so definiert wird, dass sie eine Fortführung der reellen ist – mit anderen Worten: Wenn wir die komplexen Operationen auf die reellen Zahlen einschränken, sollten wir wieder unsere bekannten reellen Verknüpfungen erhalten. Außerdem wäre es wünschenswert, dass die Fortsetzung der uns bekannten Operationen auf den neuen Zahlbereich dennoch eine schöne Struktur hervorbringt: Unser Ziel ist es, die komplexen Zahlen als Körper zu definieren.

Diese Ziele vor Augen definieren wir die gewünschten Verknüpfungen wie folgt – zunächst die *komplexe Addition.*

Für $z_1 := a + \mathrm{i} \cdot b$ und $z_2 := c + \mathrm{i} \cdot d$ setze:

$$z_1 + z_2 := (a + c) + \mathrm{i} \cdot (b + d) \in \mathbb{C}.$$

Damit ist $+ : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ offensichtlich eine Fortsetzung der reellen Addition, denn für $b = d = 0$ sind $z_1, z_2 \in \mathbb{R}$, $\mathfrak{Im}(z_1 + z_2) = 0$ und $z_1 +_{\mathbb{C}} z_2 = z_1 +_{\mathbb{R}} z_2$. In diesem Sinne verzichten wir auf die Indizierung beim Operationszeichen.

Die *komplexe Multiplikation* ist für $z_1 := a + \mathrm{i} \cdot b$ und $z_2 := c + \mathrm{i} \cdot d$ gegeben durch

$$z_1 \cdot z_2 := (ac - bd) + \mathrm{i} \cdot (ad + bc) \in \mathbb{C}.$$

Wie man leicht sieht, ist auch $\cdot : \mathbb{C} \times \mathbb{C} \to \mathbb{C}$ eine Fortsetzung der reellen Multiplikation. Es ist eine leichte Übungsaufgabe nachzurechnen, dass folgender Satz gilt:

Satz 4.1:

Die Struktur $(\mathbb{C}, +, \cdot, 0, 1)$ der komplexen Zahlen ist ein Körper, wobei $0 := 0 + \mathrm{i} \cdot 0$ und $1 := 1 + \mathrm{i} \cdot 0 = 1$. Insbesondere ist der Körper $\mathbb{C}$ eine Erweiterung des Körpers $\mathbb{R}$ (inklusive der Verknüpfungen).

Im Folgenden verzichten wir –aufgrund der erfolgreichen Einbettung der reellen Zahlen in die komplexen– auf die formale Unterscheidung der neutralen Elemente und stellen fest: $0 = 0$ und $1 = 1$.

Bevor wir nun die Division komplexer Zahlen behandeln, führen wir den dabei nützlichen Begriff der Konjugierten einer komplexen Zahl ein: Für $z = a + \mathrm{i} \cdot b$ nennen wir $\overline{z} := a - \mathrm{i} \cdot b$ die *Konjugierte* zu z. Diese Operation hat beispielsweise folgende Eigenschaften, die man leicht als Übungsaufgabe nachrechnet:

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2} \quad \text{und} \quad \overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}.$$

Wir betrachten nun die Division zweier komplexer Zahlen.

Betrachte für $z = a + \mathrm{i} \cdot b$ mit $z \neq 0$ die komplexe Zahl $z' := \frac{1}{a^2+b^2} \cdot \bar{z}$. Beachten Sie, dass insbesondere $a^2 + b^2 \neq 0$ gilt und weiterhin:

$$z \cdot z' = z' \cdot z = \frac{1}{a^2 + b^2} \cdot \bar{z} \cdot z = \frac{1}{a^2 + b^2} \cdot (a^2 + b^2) = 1$$

Damit ist z' das multiplikative Inverse von z und wir bezeichnen im Folgenden z' mit z^{-1} oder $\frac{1}{z}$.

Die *Division komplexer Zahlen* können wir jetzt wie folgt einführen:

$$\frac{z_1}{z_2} := z_1 : z_2 := z_1 \cdot \frac{1}{z_2} \, (= z_1 \cdot z_2^{-1})$$

Insbesondere gilt:

$$z^{-1} = \frac{1}{z} = \frac{\overline{z}}{\overline{z} \cdot z} = \frac{1}{\overline{z} \cdot z} \cdot \overline{z},$$

wobei $\overline{z} \cdot z = a^2 + b^2$ für $z = a + ib$ eine reelle Zahl ist, so dass man diese Formel bequem für die Berechnung komplexer Inversen ausnutzen kann.

Schauen wir uns ein Beispiel an und berechnen: $\frac{3-2\mathrm{i}}{4+5\mathrm{i}}$. Dann gilt:

$$\begin{aligned} \frac{(3-2\mathrm{i})(4-5\mathrm{i})}{(4+5\mathrm{i})(4-5\mathrm{i})} &= \frac{12 - 15\mathrm{i} - 8\mathrm{i} + 10\mathrm{i}^2}{16 - 25\mathrm{i}^2} &= \frac{12 - 10 - 23\mathrm{i}}{16 + 25} \\ &= \frac{2 - 23\mathrm{i}}{41} &= \frac{2}{41} - \mathrm{i} \cdot \frac{23}{41} \end{aligned}$$

Insbesondere gilt für den Real- und Imaginärteil dieser komplexen Zahl: $\mathfrak{Re}(\frac{3-2\mathrm{i}}{4+5\mathrm{i}}) = \frac{2}{41}$ und $\mathfrak{Im}(\frac{3-2\mathrm{i}}{4+5\mathrm{i}}) = -\frac{23}{41}$. Beachten Sie, dass der Imaginärteil einer komplexen Zahl immer eine reelle Zahl ist; es gilt für $z = a + ib$ stets: $\mathfrak{Re}(z) = a$ und $\mathfrak{Im}(z) = b$, wobei a und b reelle Zahlen sind.

⋆ ⋆ ⋆

Komplexe quadratische Gleichungen. Als eine erste Anwendung komplexer Zahlen betrachten wir *quadratische Gleichungen* und suchen nach Lösungen. Zur Erinnerung: Eine quadratische Gleichung über den reellen Zahlen hat allgemein die Form

$$ax^2 + bx + c = 0,$$

wobei $a, b, c \in \mathbb{R}$ und $a \neq 0$. Aus der Theorie der reellen Zahlen kennen wir die Lösungsformel:

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

sofern $D := b^2 - 4ac \geq 0$. Hierbei wird D als *Diskriminante* bezeichnet.

Mithilfe der komplexen Zahlen können wir Wurzeln aus negativen Zahlen ziehen, beispielsweise ist $\sqrt{-4} = \sqrt{(-1) \cdot 4} = \sqrt{-1} \cdot \sqrt{4} = \pm 2\mathrm{i}$. Dieses Argument zeigt auch, dass für $z = \sqrt{a}$ mit $a < 0$ stets gilt: $z = \pm \mathrm{i} \cdot \sqrt{-a}$.

Man kann leicht zeigen, dass sich dies auch für den Fall einer negativen Diskriminante bei quadratischen Gleichungen ausnutzen lässt; in diesem Fall (wenn $D < 0$) finden wir auch die beiden komplexen Lösungen:

$$z_{1,2} = \frac{-b \pm \mathrm{i} \cdot \sqrt{4ac - b^2}}{2a} = -\frac{b}{2a} \pm \mathrm{i} \cdot \frac{\sqrt{4ac - b^2}}{2a}.$$

Man kann sogar noch mehr zeigen: Diese Lösungsformel gilt auch für komplexe Koeffizienten a, b, c. Allerdings muss man dann ggf. die Quadratwurzel einer komplexen Zahl berechnen und dies kann aufwendig sein.

In diesem Zusammenhang möchte ich den folgenden Satz erwähnen, der grundlegend ist:

Satz 4.2 – Fundamentalsatz der Algebra:

Jede polynomielle Gleichung n-ten Grades hat genau n komplexe Lösungen (Vielfachheiten mitgezählt).

⋆ ⋆ ⋆

Komplexe Zahlen als reelle Ebene. Kommen wir zu einem anderen Thema und befassen uns mit der *Darstellung komplexer Zahlen als Paare reeller Zahlen*: Wir können eine komplexe Zahl $z = a + \mathrm{i}b$ mit einem geordneten Paar reeller Zahlen $(a, b) \in \mathbb{R} \times \mathbb{R}$ identifizieren, das heißt, wir können folgende Abbildung angeben:

$$\mathbb{C} \to \mathbb{R} \times \mathbb{R}, \quad a + \mathrm{i}b \mapsto (a, b)$$

Diese Abbildung bildet eine komplexe Zahl z auf das Paar $(\mathfrak{Re}(z), \mathfrak{Im}(z))$ ab und ist somit insbesondere bijektiv. Die Operationen „+“ und „·“ sehen in dieser Darstellung wie folgt aus:

$$\begin{aligned}(a, b) + (c, d) &:= (a + c, b + d)\\ (a, b) \cdot (c, d) &:= (ac - bd, ad + bc)\end{aligned}$$

Insbesondere sieht man leicht, dass gilt: $1 \mapsto (1, 0)$ und $\mathrm{i} \mapsto (0, 1)$.

Mithilfe dieser Überlegung können wir uns eine geometrische Darstellung komplexer Zahlen überlegen. Da eine komplexe Zahl genau einem Paar von reellen Zahlen entspricht, können wir versuchen, komplexe Zahlen in eine Ebene einzuzeichnen – die so genannte gaußsche Zahlenebene (Abbildung 2).

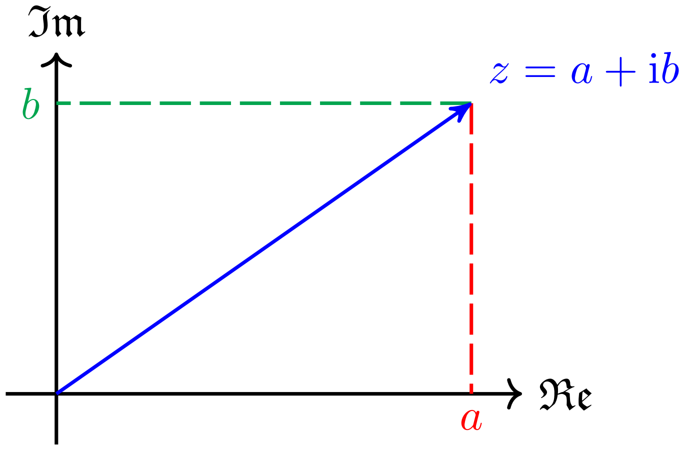


Abbildung 2: grafische Definition der komplexen Zahlen

Dabei interpretieren wir eine komplexe Zahl entweder als den Punkt (a, b) in der Ebene oder als den dazugehörigen so genannten *Ortsvektor*. Im Folgenden werden wir beides parallel verwenden, vorzugsweise aber mit den Ortsvektoren arbeiten. Diese Art der Sichtweise können wir insbesondere ausnutzen, wenn wir die *Addition* geometrisch interpretieren wollen, wie etwa im folgenden Bild dargestellt (hier sogar der spezielle Fall der *Subtraktion*, denn es gilt: $z_2 - z_1 = z_2 + (-z_1)$).

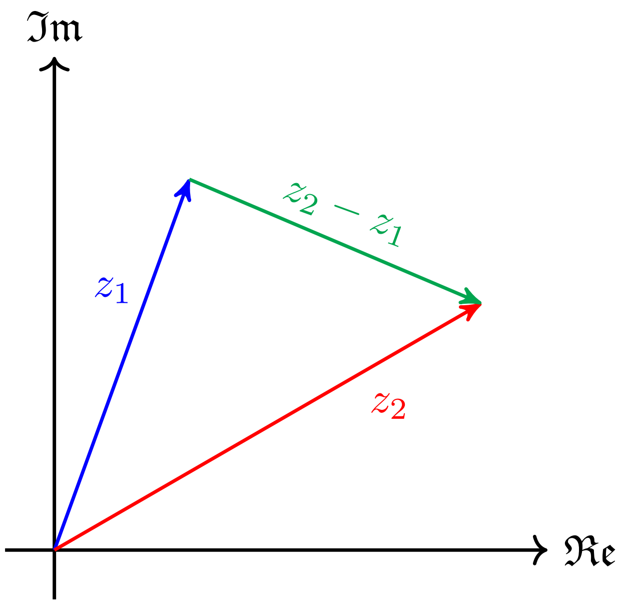


Abbildung 3: Spezieller Fall der Subtraktion

Komplexe Zahlen als Polarkoordinaten. Wenn wir uns die Multiplikation geometrisch überlegen wollen, dann ist eine andere Sichtweise der komplexen Zahlen besser geeignet: die *Darstellung der komplexen Zahlen durch Polarkoordinaten*.

Zunächst definieren wir den *Betrag* einer komplexen Zahl $z = a + ib$. Dieser ist gegeben durch:

$$|z| := \sqrt{a^2 + b^2} = \sqrt{\overline{z} \cdot z}$$

Wenn Sie sich überlegen, dass in der Darstellung mittels Ortsvektoren immer ein rechtwinkliges Dreieck entsteht, welches aus den beiden Katheten a und b und der Hypothenuse $|z|$ besteht, dann wird Ihnen auch klar, dass der Betrag einer komplexen Zahl gerade der Länge des Ortsvektors entspricht. Schauen wir uns dafür folgende Abbildung an:

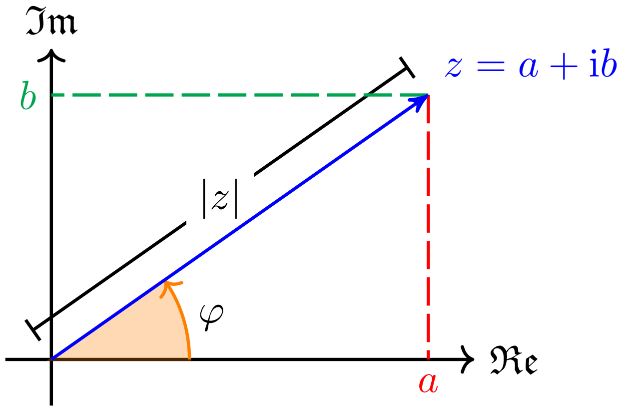


Abbildung 4: Polarkoordinatendarstellung komplexer Zahlen

Der Ortsvektor eines Punktes (a, b) kann auch durch den Winkel φ und die Länge des Ortsvektors charakterisiert werden – was wiederum einfach eine andere Sichtweise der komplexen Zahlen ist.

Dabei gilt –aufgrund des Kosinussatzes im rechtwinkligen Dreieck– die Gleichung $a = |z| \cdot \cos\varphi$ und entsprechend $b = |z| \cdot \sin\varphi$ wegen des Sinussatzes. Somit gilt insbesondere

$$z = a + \mathrm{i}b = |z| \cdot \cos\varphi + \mathrm{i} \cdot |z| \cdot \sin\varphi = |z| \cdot (\cos\varphi + \mathrm{i} \cdot \sin\varphi)$$

In der Vorlesung über Analysis werden Sie die komplexe Exponentialfunktion kennenlernen und beweisen, dass folgendes gilt:

$$\mathrm{e}^{\mathrm{i}\varphi} = \cos\varphi + \mathrm{i} \cdot \sin\varphi$$

Dann gilt offenbar auch: $z = |z| \cdot \mathrm{e}^{\mathrm{i}\varphi}$.

Damit haben Sie eine weitere Darstellung komplexer Zahlen gefunden. Wir können einer komplexen Zahl $z = a + ib$ mit $z \neq 0$ auf eindeutige Art und Weise das Paar $(|z|, \varphi)$ zuordnen, wobei φ der zum Ortsvektor gehörende Winkel entsprechend der obigen Abbildung ist. Dieser Winkel wird dabei derart gewählt, dass $-\pi < \varphi \leq \pi$ gilt; damit werden die beiden Halbkreise beschrieben. Dass wir hier nicht einen Vollkreis (also $0 \leq \varphi < 2\pi$) nutzen, hat technische Gründe.

Schauen wir uns die nicht ganz einfache *Umwandlung in Polarkoordinaten* an: Es sei dafür eine komplexe Zahl $z = a + ib$ gegeben. Dann gilt: $|z| = \sqrt{a^2 + b^2}$ und den gewünschten Winkel erhalten wir durch:

$$\varphi = \begin{cases} \arccos\left(\frac{a}{|z|}\right) & \text{für } b \geq 0 \\ -\arccos\left(\frac{a}{|z|}\right) & \text{für } b < 0 \end{cases}$$

Mit dieser Darstellung wird auch die geometrische Deutung der Multiplikation komplexer Zahlen einfacher, denn es gilt:

$$\left(|z_1| \cdot \mathrm{e}^{\mathrm{i}\alpha}\right) \cdot \left(|z_2| \cdot \mathrm{e}^{\mathrm{i}\beta}\right) = |z_1| \cdot |z_2| \cdot \mathrm{e}^{\mathrm{i}(\alpha+\beta)}$$

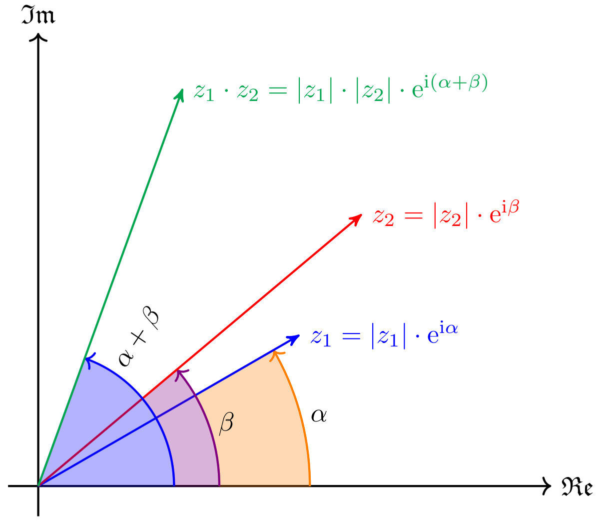


Abbildung 5: Multiplikation komplexer Zahlen in Polarkoordianten

Wir erkennen, dass sich die Winkel bei der Multiplikation addieren, das heißt, der eine Ortsvektor wird um den Winkel des anderen Ortsvektors gedreht, wobei sich die Längen der Ortsvektoren multiplizieren.

Wurzelziehen aus komplexen Zahlen. Jetzt sind wir auch dazu in der Lage, die n-te Wurzel einer komplexen Zahl zu berechnen. Beachten Sie, dass dieses Thema durchaus „komplexer“ ist. Auf den reellen Zahlen bezeichnen wir mit dem Ausdruck $\sqrt{x}$ den eindeutig bestimmten Wert y auf den nicht-negativen reellen Zahlen, so dass $y^2 = x$ gilt. Dabei ist diese reelle Wurzelfunktion auch nur auf den nicht-negativen reellen Zahlen definiert. Wenn wir nun Wurzeln auf *allen* komplexen Zahlen ziehen wollen, also etwa $\sqrt{z}$, dann ist dies streng genommen keine Funktion mehr, da es hier mehrere Lösungen gibt, an denen wir aber auch alle interessiert sind. Wir suchen also in diesem Fall nach den Lösungen z_1 und z_2, so dass $z_1^2 = z_2^2 = z$ gilt. Und es gibt entsprechend mehr Lösungen, wenn der Wurzelexponent größer wird. Exakter formuliert suchen wir also nicht nach den Werten $\sqrt[n]{z}$, sondern eigentlich nach den Lösungen z_i, so dass $z_i^n = z$ gilt.

Wir wollen nun zeigen, wie diese „Wurzeln“, also die Lösungen der entsprechenden polynomiellen Gleichung, praktisch gefunden werden können. Sei dazu $z \in \mathbb{C}$ gegeben. Dann können wir die Darstellung in Polarkoordinaten nutzen, also $z = |z| \cdot \mathrm{e}^{\mathrm{i}\varphi}$, und in salopper Art und Weise die gesuchten Wurzeln finden:

$$\sqrt[n]{z} = \sqrt[n]{|z| \cdot \mathrm{e}^{\mathrm{i}\varphi}} = \sqrt[n]{|z|} \cdot \mathrm{e}^{\mathrm{i}\frac{\varphi}{n}}$$

Wenn wir nun noch $\mathrm{e}^{\mathrm{i}(2\pi+\varphi)} = \mathrm{e}^{\mathrm{i}\varphi}$ beachten, kommen wir auf die Werte $z_k = \sqrt[n]{|z|} \cdot \mathrm{e}^{\mathrm{i}\frac{\varphi+2k\pi}{n}}$ für $k = 0, \ldots, n-1$. Man überprüft das durch einfaches Nachrechnen:

$${z_k}^n = \sqrt[n]{|z|}^n \cdot \left(\mathrm{e}^{\mathrm{i}\frac{\varphi+2k\pi}{n}}\right)^n = |z| \cdot \mathrm{e}^{\mathrm{i}(\varphi+2k\pi)} = |z| \cdot \mathrm{e}^{\mathrm{i}\varphi} = z$$

Wir haben also alle n-ten Wurzeln von z gefunden und wissen wegen des Fundamentalsatzes der Algebra, dass es keine weiteren Lösungen der Gleichung $x^n - z = 0$ geben kann.

Beispiel 4.3:

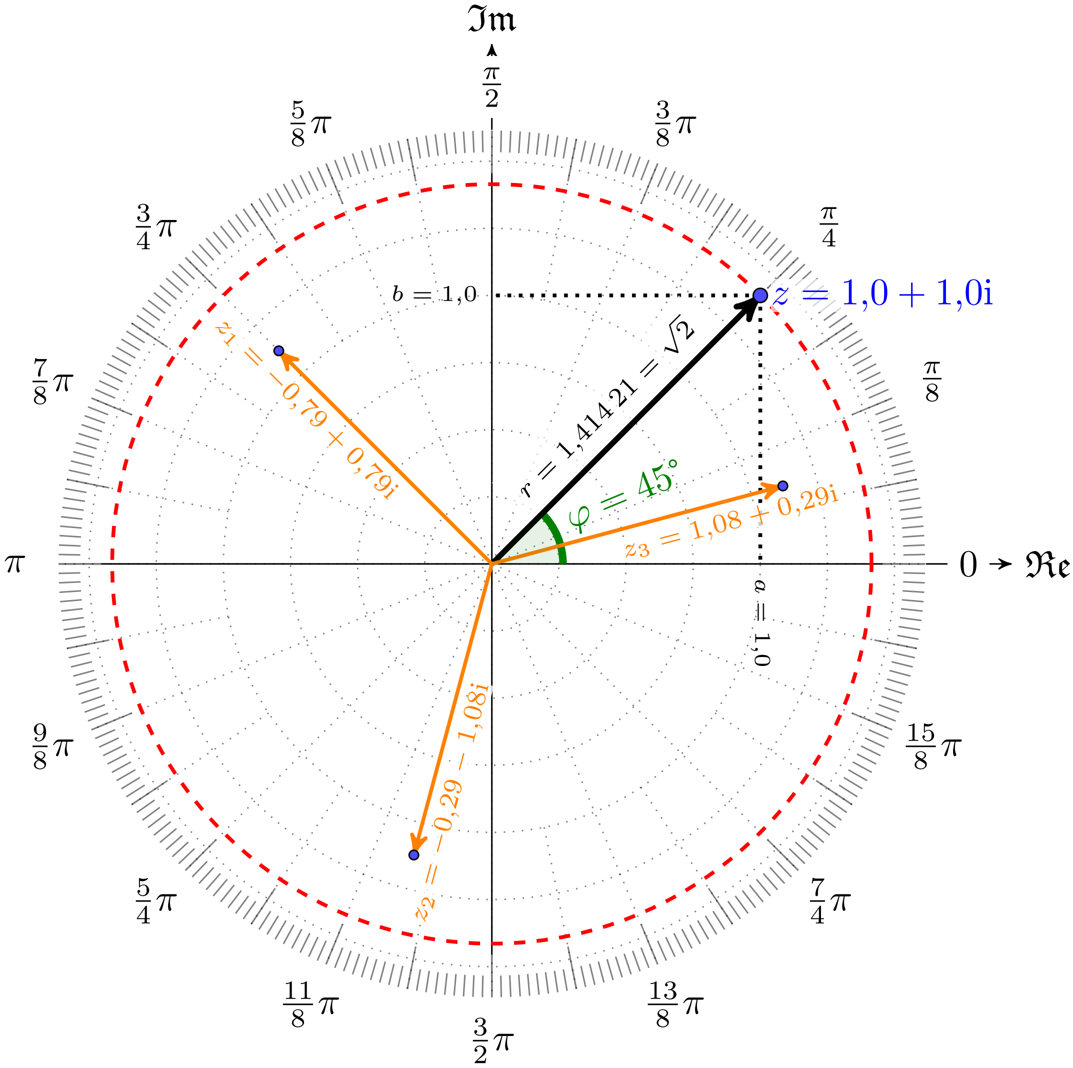


Abbildung 6: $1 + \mathrm{i}$ und die drei 3-ten Wuzeln in Polarkoordinaten

Sei $z = 1 + \mathrm{i}$. Wir wollen die 3-te Wurzel berechnen. Dazu müssen wir zunächst den Betrag und den Winkel φ bestimmen:
Wir wissen, dass

$$
\begin{aligned}
|z| &= \sqrt{1+1} = \sqrt{2} \\
\varphi &= \begin{cases} \arccos\left(\frac{a}{|z|}\right) & \text{für } b \geq 0 \\ -\arccos\left(\frac{a}{|z|}\right) & \text{für } b < 0 \end{cases}
\end{aligned}
$$

Da $b = 1$ ($z = a + \mathrm{i}b = 1 + \mathrm{i}$), also $b \geq 0$ berechnen wir den Winkel $\varphi = \arccos\left(\frac{a}{|z|}\right) = \arccos\left(\frac{1}{\sqrt{2}}\right) = \frac{\pi}{4} \hat{=} 45°$. Damit gilt:

$$
\begin{aligned}
z &= \sqrt{2} \cdot \mathrm{e}^{\mathrm{i}\frac{\pi}{4}} \\
\sqrt[3]{z} &= \sqrt[3]{\sqrt{2} \cdot \mathrm{e}^{\mathrm{i}\left(\frac{\pi}{4} + 2k\pi\right)}} = \sqrt[6]{2} \cdot \mathrm{e}^{\mathrm{i}\frac{(1+8k)\pi}{12}} \\
\to z'_k &= \sqrt[6]{2} \cdot \left(\cos\left(\frac{(1+8k)\pi}{12}\right) + \mathrm{i}\sin\left(\frac{(1+8k)\pi}{12}\right)\right) \\
z'_1 &= \sqrt[6]{2} \cdot \left(\cos\left(\frac{3\pi}{4}\right) + \mathrm{i}\sin\left(\frac{3\pi}{4}\right)\right) \approx -0{,}794 + 0{,}794\mathrm{i}
\end{aligned}
$$

$$z_2' = \sqrt[6]{2}\cdot\left(\cos\left(\frac{17\pi}{12}\right)+\mathrm{i}\sin\left(\frac{17\pi}{12}\right)\right) \approx -0{,}291-1{,}084\mathrm{i}$$
$$z_3' = \sqrt[6]{2}\cdot\left(\cos\left(\frac{\pi}{12}\right)+\mathrm{i}\sin\left(\frac{\pi}{12}\right)\right) \approx 1{,}084+0{,}291\mathrm{i}$$

Den Winkel hätten wir auch durch scharf nachdenken ermitteln können. Klar ist, dass $1+\mathrm{i}$ Diagonal nach oben gehen muss, also mit 45°. Vergleichen Sie dazu Abbildung 6.

Beispiel 4.4: Hier einmal kurz die geometrische Darstellung der 16 16-ten Wurzeln von $1{,}2\cdot\mathrm{e}^{\frac{29}{16}\mathrm{i}\pi}\approx 0{,}9977-0{,}666\,64\mathrm{i}$. Wenn wir uns nun die Koordinaten ansehen und daraus den Winkel berechnen wollen, so müssen wir berücksichtigen, dass $b=-0{,}666\,64<0$. Damit berechnen wir φ durch:

$$\varphi = -\arccos\left(\frac{a}{r}\right) \quad \xrightarrow[r=1{,}2]{a=0{,}9977} = -\frac{3}{16}\pi = \frac{29}{16}\pi = 1{,}8125\pi$$

Der Radius der Wurzeln ist (für alle Ordnungen k) dann gegeben durch:

$$r_k = \sqrt[16]{1{,}2}\approx 1{,}011\,44$$

Die Winkel werden berechnet durch:

$$\varphi_k = \frac{\varphi+2k\pi}{n} = \frac{\frac{29}{16}\pi+2k\pi}{16} = \left(\frac{29}{256}+\frac{2k}{16}\right)\pi$$

Also erhalten wir die 16 Wurzeln durch:

$$z_k = \sqrt[16]{1{,}2}\cdot\mathrm{e}^{\mathrm{i}\left(\frac{29}{256}+\frac{2k}{16}\right)\pi} = \sqrt[16]{1{,}2}\cdot\left(\cos\left(\left(\frac{29}{256}+\frac{2k}{16}\right)\pi\right)+\mathrm{i}\sin\left(\left(\frac{29}{256}+\frac{2k}{16}\right)\pi\right)\right)$$

Die 16 Wurzeln sind geometrisch in Abbildung 7 dargestellt.

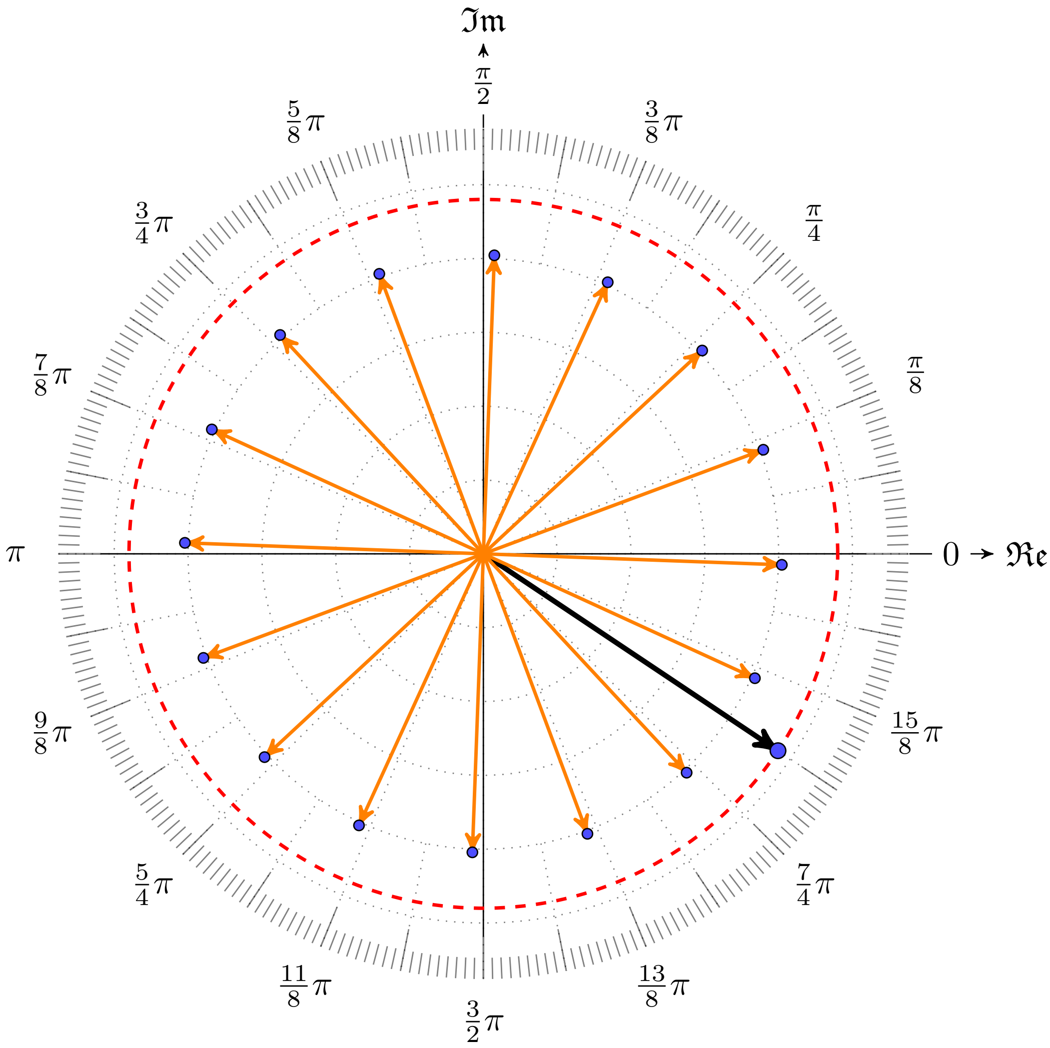


Abbildung 7: $\sqrt[16]{1{,}2 \cdot \mathrm{e}^{\frac{29}{16}\mathrm{i}\pi}} = \sqrt[16]{0{,}9977 - 0{,}6666\mathrm{i}}$

⋆ ⋆ ⋆

Anwendungen. Abschließend noch eine kurze Bemerkung zu den *Anwendungen* komplexer Zahlen: Komplexe Zahlen werden beispielsweise in der Physik als sehr nützlich angesehen und verdienen daher den Namen „imaginäre Zahlen“ eigentlich nicht (allerdings ist dieser historisch gewachsen und so bleibt man natürlich dabei). Die komplexen Zahlen werden in der Physik u.a. in der Quantentheorie und Relativitätstheorie angewendet, um Schwingungsvorgänge oder Phasenverschiebungen zu untersuchen.

⋆ ⋆ ⋆

Jenseits der komplexen Zahlen: Quaternionen und Oktonionen. Wir haben gesehen, dass wir die komplexen Zahlen als zweidimensionalen Vektorraum $\mathbb{R}^2$ mit zusätzlicher Multiplikation (neben der Skalarmultiplikation) auffassen können. Damit wird der Vektorraum $\mathbb{R}^2$ zum Körper, wie schon der $\mathbb{R}^1 = \mathbb{R}$. Daher entsteht ganz natürlich die Frage, welche anderen Dimensionen n noch möglich sind, um $\mathbb{R}^n$ zu einem Körper mittels Einführung einer geeigneten Multiplikation zu machen.

Hamilton versuchte jahrzehntelang ein solche Erweiterung in drei Dimensionen zu finden - erfolglos. Wie wir heute wissen, gab es dafür einen guten Grund: Dies ist auch nicht möglich! In vier Dimensionen ist dies in der Tat (immerhin fast!) möglich und auch für die Physik wichtig: Quaternionen. Als Erweiterung der komplexen Zahlen gibt es nun nicht eine sondern gleich drei imaginäre Einheiten: i, j und k, wobei jeweils gilt: $\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = -1$. Allerdings ist die Multiplikation *nicht* kommutativ: $\mathrm{i} \cdot \mathrm{j} = -\mathrm{j} \cdot \mathrm{i} = \mathrm{k}$, $\mathrm{j} \cdot \mathrm{k} = -\mathrm{k} \cdot \mathrm{j} = \mathrm{i}$, $\mathrm{k} \cdot \mathrm{i} = -\mathrm{i} \cdot \mathrm{k} = \mathrm{j}$. Eine Quaternion ist als vierdimensionaler Vektor definiert, nämlich als Linearkombination der Basisvektoren 1, i, j und k. Die Körperaxiome können alle nachgerechnet werden - außer der Kommutativität der Multiplikation. Solche Strukturen nennt man Schiefkörper.

Auf der Suche nach weiteren Strukturen dieser Art wird man erst wieder in der achten Dimension fündig. Die so genannten Oktonionen bilden als Erweiterung des $\mathbb{R}^8$ um eine geeignete Multiplikation nicht einmal mehr einen Schiefkörper, da neben der Kommutativität nun auch noch die Assoziativität auf der Strecke bleibt. (Man spricht von einer 8-dimensionalen reellen und normierten Divisonalgebra.) Sie werden sich nicht wundern, dass Mathematiker und auch Physiker diese Zahlenerweiterung in einigen Bereichen anwendungsfreundlich finden. Eine solche assoziative kann es auch nicht geben, denn nach dem Satz von Frobenius gilt:

Satz 4.5:

Jede reelle, endlichdimensionale und assoziative Divisionsalgebra ist zu den reellen oder komplexen Zahlen oder zu den Quaternionen isomorph.

Aber selbst wenn Sie auf die Assoziativität der Multiplikation scheinbar großzügig verzichten und nur nach endlich-dimensionalen Divisionsalgebren über den reellen Zahlen fragen, werden Sie nicht viel weiter fündig.

Satz 4.6:

Endlich-dimensionale Divisionsalgebren über den reellen Zahlen sind nur in den Dimensionen 1, 2, 4 und 8 möglich. Die normierten Versionen sind bis auf Isomorphie gerade die reellen und komplexen Zahlen, die Quaternionen (Hamilton-Zahlen) und die Oktonionen (Caylay-Zahlen).

5. Linearkombinationen und Basen

Um Vektorräume im Allgemeinen verstehen zu können, werden wir diese etwas tiefer analysieren, um mithilfe neuer Begrifflichkeiten Zusammenhänge besser verstehen zu können.

Linearkombinationen und lineare Unabhängigkeit. Wir haben bei Vektoren im bekannten Raum $\mathbb{R}^2$ bereits gesehen, dass wir jeden beliebigen Punkt in der Ebene durch Addition von gestauchten bzw. gestreckten Versionen von zwei gegebenen Vektoren erzeugen können, sofern die Ausgangsvektoren nicht auf einer Geraden liegen. Dieses Phänomen untersuchen wir als Erstes:

Definition 5.1 – Linearkombination:

Sei V ein $\mathbb{K}$-Vektorraum und $v_1, \ldots, v_n \in V$. Ein $v \in V$ ist eine Linearkombination (kurz: LK) von $v_1, \ldots, v_n$, wenn es $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ gibt, sodass

$$v = \lambda_1 v_1 + \ldots + \lambda_n v_n$$

Die Körperelemente $\lambda_1, \ldots, \lambda_n$ heißen Koeffizienten dieser Linearkombination.
Die Menge $\mathcal{L}(v_1, \ldots, v_n) := \left\{ \sum_{i=1}^{n} \lambda_i v_i \;\middle|\; \lambda_1, \ldots, \lambda_n \in \mathbb{K} \right\} \subseteq V$ aller Linearkombinationen von $v_1, \ldots, v_n$ heißt Lineare Hülle von $v_1, \ldots, v_n$.
Im Fall $n = 0$ setzen wir $\mathcal{L}(\emptyset) := \{0\}$.

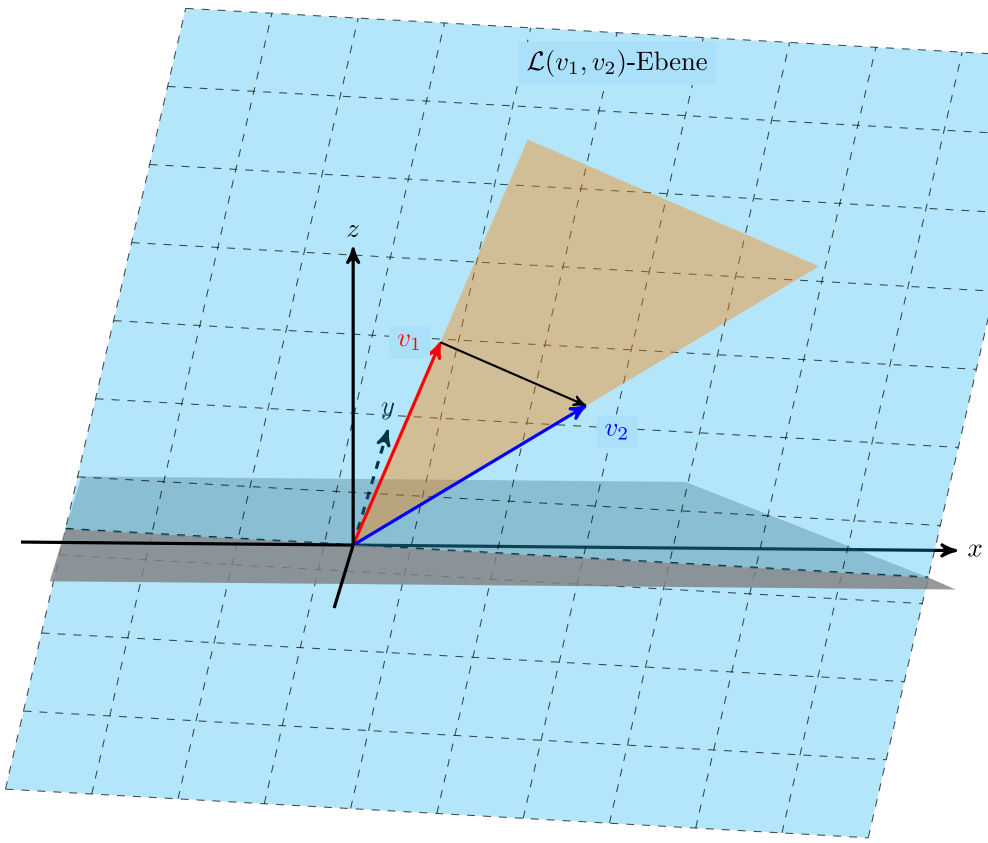


Abbildung 8: $\mathcal{L}(v_1, v_2)$-Ebene

Satz 5.2:

Sei V ein $\mathbb{K}$-Vektorraum und $v_1, \ldots, v_n \in V$. Dann ist $\mathcal{L}(v_1, \ldots, v_n)$ ein Untervektorraum von V.

Beweis: Zunächst einmal gilt, dass $0 \in \mathcal{L}(v_1, \ldots, v_n)$ ist und daher $\mathcal{L}(v_1, \ldots, v_n) \neq \emptyset$. (Klar per Definition für $n = 0$, für $n \neq 0$ setze in der Linearkombination $\lambda_1 = \ldots = \lambda_n = 0$.)

Es bleibt die Abgeschlossenheit bezüglich der Vektorraum-Addition und Skalar-Multiplikation zu zeigen:

Bzgl. „+“: Seien $v = \sum_{i=1}^{n} \lambda_i v_i$ und $w = \sum_{i=1}^{n} \mu_i v_i$. Dann gilt:

$$v + w = \sum_{i=1}^{n} \lambda_i v_i + \sum_{i=1}^{n} \mu_i v_i = \sum_{i=1}^{n} (\lambda_i + \mu_i) v_i \in \mathcal{L}(v_1, \ldots, v_n)$$

Analog gilt $\lambda v \in \mathcal{L}(v_1, \ldots, v_n)$ für $\lambda \in \mathbb{K}$ und $v \in \mathcal{L}(v_1, \ldots, v_n)$. ⊠

Geometrisch ist die lineare Hülle $\mathcal{L}(v_1, \ldots, v_n)$ der kleinste Unterraum von V, der die erzeugenden Vektoren $v_1, \ldots, v_n$ enthält.

Abbildung 9: [Bild] – Bild folgt

Definition 5.3 – endlich erzeugend:

Sei V ein $\mathbb{K}$-Vektorraum. Dann heißt V endlich erzeugt, wenn es eine endliche Folge $(v_1, \ldots, v_n)$ von Vektoren in V gibt, sodass $\mathcal{L}(v_1, \ldots, v_n) = V$ gilt.

Beachten Sie, dass ein endlich erzeugter Vektorraum nicht notwendig endlich ist. Dies hängt gravierend vom Grundkörper ab. So ist beispielsweise der Vektorraum $\mathbb{R}^2$ offenbar endlich erzeugt, aber nicht endlich. Endliche Vektorräume werden etwa in Satz 5.18 betrachtet.

Definition 5.4 – lineare Abhängigkeit:

Sei V ein $\mathbb{K}$-Vektorraum und $v_1, \ldots, v_n \in V$. Das n-Tupel $(v_1, \ldots, v_n)$ heißt linear unabhängig (kurz: l.u.), wenn für alle Darstellungen

$$\lambda_1 v_1 + \ldots + \lambda_n v_n = 0$$

gilt: $\lambda_1 = 0, \ldots, \lambda_n = 0$.

Ist das n-Tupel $(v_1, \ldots, v_n)$ nicht linear unabhängig, so sagen wir, dass es linear abhängig ist. Wir vereinbaren: für $n = 0$ sei das 0-Tupel $() = \emptyset$ immer linear unabhängig.

Diese Definition wird im ersten Moment häufig unterschätzt. Beachten Sie folgenden praktischen Algorithmus zur Überprüfung: Bei gegebenen Vektoren $v_1, \ldots, v_n$ betrachten Sie das zugehörige homogene LGS

$$\lambda_1 v_1 + \ldots + \lambda_n v_n = 0$$

in den Unbekannten $\lambda_1, \ldots, \lambda_n$ und überprüfen, ob es nur genau eine Lösung, nämlich die triviale Lösung, gibt.

Beispiel 5.5: Die Vektoren $v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, v_2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, v_3 = \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix}$ im $\mathbb{R}^3$ sind linear abhängig, da gilt:

$$2 \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} - \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix} \qquad \Rightarrow \qquad (-1) \cdot v_1 + 2 \cdot v_2 + (-1) \cdot v_3 = 0$$

Lineare Unabhängigkeit kann wie folgt charakterisiert werden:

Satz 5.6:

Sei V ein $\mathbb{K}$-Vektorraum und seien $v_1, \ldots, v_n \in V$. Dann ist $(v_1, \ldots, v_n)$ linear unabhängig genau dann, wenn für alle $i \leq n$ gilt: v_i ist nicht Linearkombination der Vektoren $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$.

Beweis: Wir beweisen zuerst „$\Rightarrow$“:
Sei $(v_1, \ldots, v_n)$ linear unabhängig. Betrachte $i \leq n$ beliebig. Angenommen v_i wäre eine Linearkombination von $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$, Wähle $\lambda_1, \ldots, \lambda_{i-1}, \lambda_{i+1}, \ldots, \lambda_n \in \mathbb{K}$, sodass

$$v_i = \lambda_1 v_1 + \ldots + \lambda_{i-1} v_{i-1} + \lambda_{i+1} v_{i+1} + \ldots + \lambda_n v_n = \sum_{\substack{j=1 \\ j \neq i}}^{n} \lambda_j v_j$$

Dann gilt: $0 = \sum_{j=1, j \neq i}^{n} \lambda_j v_j + (-1) v_i$. Damit haben wir einen Koeffizienten, der nicht 0 ist (es gilt $0 \neq -1$, warum eigentlich?!) und dies ist ein Widerspruch zur linearen Unabhängigkeit der $v_1, \ldots, v_n$.

Es bleibt „$\Leftarrow$“ zu zeigen:
Sei also für alle $i \leq n$, v_i keine Linearkombination der Vektoren $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$. Angenommen $(v_1, \ldots, v_n)$ wären linear abhängig. Wähle Koeffizienten $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$, sodass $0 = \lambda_1 v_1 + \ldots + \lambda_n v_n$ und es existiert ein $i \leq n$ mit $\lambda_i \neq 0$.

Dann ist $(-\lambda_i) v_i = \sum_{j=1, j \neq i}^{n} \lambda_j v_j$ und damit auch

$$v_i = -\frac{\lambda_1}{\lambda_i} v_1 - \ldots - \frac{\lambda_{i-1}}{\lambda_i} v_{i-1} - \frac{\lambda_{i+1}}{\lambda_i} v_{i+1} - \ldots - \frac{\lambda_n}{\lambda_i} v_n = \sum_{\substack{k=1 \\ k \neq i}}^{n} \frac{\lambda_k}{\lambda_i} v_k.$$

Das ist eine Linearkombination von $v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n$ und somit ein Widerspruch zur Voraussetzung. ⊠

Kommen wir zum zentralen Begriff einer Basis eines Vektorraums:

Definition 5.7 – Basis:

Sei V ein $\mathbb{K}$-Vektorraum und $v_1, \ldots, v_n$ Vektoren, dann ist $v_1, \ldots, v_n$ eine **Basis** von V, falls $v_1, \ldots, v_n$ linear unabhängig sind und $\mathcal{L}(v_1, \ldots, v_n) = V$.

Satz 5.8:

Sei V ein $\mathbb{K}$-Vektorraum mit Basis $(v_1, \ldots, v_n)$. Für alle Vektoren $v \in V$ existiert genau ein n-Tupel $(\lambda_1, \ldots, \lambda_n)$, sodass $v = \sum_{i=1}^{n} \lambda_i v_i$. Wir nennen $(\lambda_1, \ldots, \lambda_n)$ die **Darstellung** von v in der Basis $v_1, \ldots, v_n$.

Beweis:
Betrachte ein $v \in V$. Da $\mathcal{L}(v_1, \dots, v_n) = V$, existiert ein Tupel $(\lambda_1, \dots, \lambda_n)$, sodass $v = \sum_{i=1}^n \lambda_i v_i$. Angenommen es gäbe außerdem die Darstellung $v = \sum_{i=1}^n \lambda'_i v_i$. Dann gilt:

$$0 = v - v = \sum_{i=1}^n \lambda_i v_i - \sum_{i=1}^n \lambda'_i v_i = \sum_{i=1}^n (\lambda_i - \lambda'_i) v_i$$

(In dieser Gleichungskette werden massiv die Vektorraumaxiome angewendet.) Dies ist eine nichttriviale Darstellung des Nullvektors, denn bei den zwei verschiedenen Darstellungen von v existiert (mindestens) ein Index $i \leq n$ mit $\lambda_i \neq \lambda'_i$, also insbesondere $\lambda_i - \lambda'_i \neq 0$ und dies ist ein Widerspruch zur Annahme. ⊠

Definition 5.9 – kanonische Basis:

Sei $n \in \mathbb{N}$ und sei $\mathbb{K}$ ein Körper, dann ist $(e_1, \dots, e_n)$ mit $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$ die **kanonische Basis** des Vektorraumes $\mathbb{K}^n$.

Lemma 5.10:

$(e_1, \dots e_n)$ ist eine Basis des $\mathbb{K}^n$.

Beweis:
Klar, denn zu zeigen ist $\mathcal{L}(v_1, \dots, v_n) = \mathbb{K}^n$ und die Vektoren $e_1, \dots, e_n$ sind linear unabhängig. Betrachten wir dazu die Definition des $\mathbb{K}^n = \left\{(\lambda_1, \dots, \lambda_n) \;\middle|\; \lambda_1, \dots, \lambda_n \in \mathbb{K}\right\}$ und dieses Tupel entspricht gerade der Linearkombination $\sum_{i=1}^n \lambda_i e_i$.

Für die lineare Unabhängigkeit müssen wir überprüfen, ob es für die Darstellung des Nullvektors als einzige Lösung $\lambda_1 = 0, \dots, \lambda_n = 0$ gibt. Dies ist offensichtlich für die Linearkombination

$$0 = \sum_{i=1}^n \lambda_i e_i = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.$$

⊠

⋆ ⋆ ⋆

Existenzsätze für Basen. Bisher haben wir gesehen, welche Vorteile Basen haben, können uns aber noch nicht sicher sein, dass wir überhaupt in beliebigen Vektorräumen eine solche schöne Folge von Vektoren finden können. Zumindest für endlich-dimensionale Vektorräume wird diese Existenz durch die folgenden drei Sätze garantiert. In unendlich-dimensionalen Vektorräumen ist dies nicht so einfach zu erhalten, da dies gravierend von der zugrunde liegenden Mengenlehre abhängt (Stichwort: Zornsches Lemma bzw. Auswahlaxiom).

Satz 5.11 – Basisergänzungssatz:

Sei V ein beliebiger Vektorraum. Sei die Folge $(v_1, \ldots, v_r)$ linear unabhängig in V, $w_1, \ldots, w_s \in V$ und

$$\mathcal{L}(v_1, \ldots, v_r, w_1, \ldots, w_s) = V.$$

Dann gibt es Indizes $i_1, \ldots, i_k \leq s$, sodass $(v_1, \ldots, v_r, w_{i_1}, \ldots, w_{i_k})$ eine Basis von V ist.

Beachten Sie: Jede linear unabhängige Teilmenge eines Vektorraums kann zu einer Basis ergänzt werden!

Beweis: Durch vollständige Induktion über $s \in \mathbb{N}$.

Induktionsanfang

Sei $s = 0$. Dann ist $(v_1, \ldots, v_r)$ linear unabhängig in V und $\mathcal{L}(v_1, \ldots, v_r) = V$. Also ist $(v_1, \ldots, v_r)$ eine Basis von V.

Induktionsschritt

Voraussetzung: Sei $s \in \mathbb{N}$ und der Satz gelte für $s \in \mathbb{N}$.
Behauptung: Der Satz gilt dann auch für $s + 1$.
Betrachte $(v_1, \ldots, v_r)$ linear unabhängig in V, $w_1, \ldots, w_{s+1} \in V$ mit

$$\mathcal{L}(v_1, \ldots, v_r, w_1, \ldots, w_s, w_{s+1}) = V.$$

Fall 1 Für alle $i \in \{1, \ldots, s + 1\}$ gilt: $w_i \in \mathcal{L}(v_1, \ldots, v_r)$.
Dann ist jede Linearkombination der Vektoren
$v_1, \ldots, v_r, w_1, \ldots, w_s, w_{s+1}$ eine Linearkombination der Vektoren $v_1, \ldots, v_r$ und es gilt:

$$\mathcal{L}(v_1, \ldots, v_r) = \mathcal{L}(v_1, \ldots, v_r, w_1, \ldots, w_s, w_{s+1}) = V.$$

Also ist $(v_1, \ldots, v_r)$ eine Basis von V.

Fall 2 Es gibt ein $i \in \{1, \ldots, s + 1\}$ mit $w_i \notin \mathcal{L}(v_1, \ldots, v_r)$.

Behauptung: $(v_1, \ldots, v_r, w_i)$ ist linear unabhängig.

Beweis der Behauptung: Sei $0 = \lambda_1 v_1 + \ldots + \lambda_r v_r + \lambda w_i$. Wenn $\lambda \neq 0$ wäre, dann gilt: $w_i = -\frac{1}{\lambda}(\sum_{j=1}^{r} \lambda_j v_j) \in \mathcal{L}(v_1, \ldots, v_r)$. Dies ist ein Widerspruch.
Also ist $\lambda = 0$ und somit gilt $\sum_{j=1}^{r} \lambda_j v_j = 0$. Da $(v_1, \ldots, v_r)$ linear unabhängig sind, gilt: $\lambda_1 = \ldots = \lambda_r = 0$. ⊠(Beh.)

Weiter gilt (nach Umsortierung)

$$\mathcal{L}(v_1, \ldots, v_r, w_i, \underbrace{w_1, \ldots, w_{i-1}, w_{i+1}, \ldots, w_{s+1}}_{\text{s-viele Elemente}}) = V.$$

Nach Anwendung der Induktionsvoraussetzung auf die letzten s Elemente existieren Indizes $i_2, \ldots, i_k \in \{1, \ldots, i - 1, i + 1, \ldots, s + 1\}$, sodass insgesamt die Folge $(v_1, \ldots, v_r, w_i, w_{i_2}, \ldots, w_{i_k})$ eine Basis von V ist.

Damit ist die Aussage per vollständiger Induktion bewiesen. ⊠

Satz 5.12 – Basisexistenzsatz:

Sei V ein endlich erzeugter $\mathbb{K}$-Vektorraum. Dann existiert eine Basis von V.

Beweis: Wähle eine Folge von Vektoren $(w_1, \dots, w_s)$ aus V, die V aufspannen:

$$V = \mathcal{L}(w_1, \dots, w_s).$$

Die leere Folge $() = \emptyset$ der Länge $r = 0$ ist linear unabhängig in V. Nach dem Basisergänzungssatz mit $r = 0$ existieren Indizes $i_1, \dots, i_k \leq s$, sodass $(w_{i_1}, \dots, w_{i_k})$ eine Basis von V ist. Somit existiert eine Basis von V. ⊠

Satz 5.13 – Basisaustauschsatz:

Sei V ein $\mathbb{K}$-Vektorraum und seien $(v_1, \dots, v_m)$ und $(w_1, \dots, w_n)$ zwei Basen von V. Dann existiert für alle $i \leq m$ ein $j \leq n$, sodass $(v_1, \dots, v_{i-1}, w_j, v_{i+1}, \dots, v_m)$ eine Basis von V ist.

Beweis: Betrachte ein beliebiges $i \leq m$.

Behauptung 1 ***Es gibt ein $j \leq n$, sodass $w_j \notin \mathcal{L}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_m)$.***

Nehmen wir an, dass für alle $j \leq n$ dies nicht der Fall wäre. Also gilt für beliebige $j \leq n$, dass $w_j \in \mathcal{L}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n)$. Dann ist

$$V = \mathcal{L}(w_1, \dots, w_n) \subseteq \mathcal{L}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_m).$$

Also ist $v_i \in V$ und somit $v_i \in \mathcal{L}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_m)$. Dies ist ein Widerspruch.

Wähle ein $j \leq n$, sodass $w_j \notin \mathcal{L}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_m)$. Dann zeigt man ähnlich wie im Beweis des Basisergänzungssatzes:

Behauptung 2 $(v_1, \dots, v_{i-1}, w_j, v_{i+1}, \dots, v_m)$ ***ist linear unabhängig.***

Es gilt

$$\mathcal{L}(v_1, \dots, v_{i-1}, w_j, v_{i+1}, \dots, v_m, v_i) \supseteq \mathcal{L}(v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_m) = V.$$

Nach dem Basisergänzungssatz ist nun (entweder) die Folge

$$(v_1, \dots, v_{i-1}, w_j, v_{i+1}, \dots, v_m, v_i)$$

oder die Folge $(v_1, \dots, v_{i-1}, w_j, v_{i+1}, \dots, v_m)$ eine Basis von V. Da aber $w_j \in V = \mathcal{L}(v_1, \dots, v_m)$ gilt, ist die Folge $(v_1, \dots, v_{i-1}, w_j, v_{i+1}, \dots, v_m, v_i)$ linear abhängig und somit $(v_1, \dots, v_{i-1}, w_j, v_{i+1}, \dots, v_m)$ die gewünschte Basis.

⊠

⋆ ⋆ ⋆

Wir nähern uns dem Dimensionsbegriff eines Vektorraums. Um diesen zu rechtfertigen, benötigen wir den folgenden

Satz 5.14:

Sei V ein $\mathbb{K}$-Vektorraum und seien zwei beliebige Basen $(v_1, \dots, v_m)$ und $(w_1, \dots, w_n)$ von V gegeben. Dann ist $m = n$.

Beweis: Angenommen, es gelte $m \neq n$. O.B.d.A. sei $m > n$. Dann definiere für jedes $i \in \{1, \dots, m\}$ rekursiv einen Index $j(i) \in \{1, \dots, n\}$ als das kleinste $j \leq n$, sodass

$(w_{j(1)},\dots,w_{j(i-1)},w_j,v_{i+1},\dots,v_m)$ eine Basis von V ist. Die Funktion $i\mapsto j(i)$ ist nach dem Basisaustauschsatz wohldefiniert. Somit ist also $(w_{j(1)},\dots,w_{j(m)})$ eine Basis von V. Da $m>n$ gilt, gibt es $i,i'<m$, sodass $j(i)=j(i')$. Somit enthält die Folge $(w_{j(1)},\dots,w_{j(m)})$ (mindestens) zwei gleiche Vektoren und ist damit linear abhängig. Dies ist ein Widerspruch zur Basiseigenschaft. ⊠

Schließlich können wir den Begriff der Dimension einführen.

Definition 5.15 – Dimension:

Sei V ein endlich erzeugter $\mathbb{K}$-Vektorraum. Dann ist die **Dimension** von V die Anzahl n der Vektoren in einer Basis $(v_1,\dots,v_n)$ von V. In diesem Fall heißt V auch **n-dimensional**. Die Dimension von V wird mit $\dim(V)$ (oder auch einfach $\dim V$) bezeichnet[a].

Ein endlich erzeugter $\mathbb{K}$-Vektorraum wird auch als **endlich-dimensionaler** $\mathbb{K}$-Vektorraum bezeichnet. Wenn V nicht endlich erzeugt ist, so ist V **unendlich-dimensional**. Wir schreiben hierfür auch kurz: $\dim(V)=\infty$.

[a] Sie werden noch sehen, dass der Dimensionsbegriff zwar nicht von der gewählten Basis, wohl aber vom zugrunde liegenden Körper abhängt. Ändern Sie den Grundkörper, so hat der neue Vektorraum in der Regel neue Eigenschaften — insbesondere kann sich auch die Dimension ändern. Daher wird die Dimension des $\mathbb{K}$-Vektorraums V zur Betonung des Körpers gelegentlich auch mit $\dim_{\mathbb{K}} V$ bezeichnet. Siehe dazu auch die Überlegungen rund um das Beispiel 5.19.

Die Dimension von Untervektorräumen ist nicht größer als die des Grundraums:

Satz 5.16:

Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $U\subseteq V$ ein Untervektorraum. Dann ist U ebenfalls endlich-dimensional und es gilt:

$$\dim(U)\leq\dim(V)$$

Beweis: Jede linear unabhängige Folge $(v_1,\dots,v_r)$ in U kann nach dem Basisergänzungssatz zu einer Basis von V ergänzt werden. Also ist $r\leq\dim(V)$. Wähle nun eine linear unabhängige Folge $(v_1,\dots,v_r)$ von Elementen in U, deren Länge maximal ist.

Behauptung: Dann gilt $(v_1,\dots,v_r)$ ist eine Basis von U.
Angenommen nicht, dann gilt: $\mathcal{L}(v_1,\dots,v_r)\neq U$. Wähle ein $u\in U$, sodass $u\notin\mathcal{L}(v_1,\dots,v_r)$. Betrachte eine Linearkombination $\lambda_1\cdot v_1+\ldots+\lambda_r\cdot v_r+\lambda\cdot u=0$. Dann ist $\lambda=0$, da $u\notin\mathcal{L}(v_1,\dots,v_r)$. Da $(v_1,\dots,v_r)$ linear unabhängig ist, gilt:

$$\lambda_1=\ldots=\lambda_r=0$$

Damit ist $(v_1,\dots,v_r,u)$ linear unabhängig. Dies ist ein Widerspruch zur maximalen Länge der Folge $(v_1,\dots,v_r)$. ⊠

Der folgende Satz wird uns das Rechnen mit Unterräumen erheblich erleichtern:

Satz 5.17:

Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $U\subseteq V$ ein Unterraum. Dann ist

$$U\neq V \text{ genau dann, wenn } \dim(U)<\dim(V).$$

Beweis: Wir zeigen zunächst „$\Rightarrow$“:
Angenommen $U \neq V$. Wähle eine Basis $(v_1, \ldots, v_r)$ von U. Die Folge $(v_1, \ldots, v_r)$ ist linear unabhängig in V. Nach dem Basisergänzungssatz kann man diese Folge zu einer Basis $(v_1, \ldots, v_r, w_1, \ldots, w_s)$ von V ergänzen. Da $U \neq V$ ist, gilt $s \neq 0$. Damit gilt

$$\dim(U) = r < r + s = \dim(V).$$

„$\Leftarrow$“ ist klar. ⊠

Beachten Sie: Da jeder Untervektorraum auch ein Vektorraum ist, können wir über seine Dimension sprechen.

Der Satz 5.17 lässt sich effektiv einsetzen, wenn man beispielsweise zeigen möchte, dass zwei Unterräume gleich sind. Seien also konkret U_1, U_2 Untervektorräume von V. Wir möchten die Gleichheit $U_1 = U_2$ zeigen. Dann reicht es zu zeigen, dass $U_1 \subseteq U_2$ und $\dim(U_1) = \dim(U_2)$. Damit sind offenbar beide Unterräume gleich.

Der folgende Satz gibt die Anzahl der Elemente eines endlich-dimensionalen Vektorraums zu einem endlichen Körper an:

Satz 5.18:

Sei V ein n-dimensionaler $\mathbb{K}$-Vektorraum und $\mathbb{K}$ habe k Elemente. Dann hat V genau k^n-viele Elemente.

Beweis: Die Elemente des Vektorraums V entsprechen eindeutig den n-Tupeln $(\lambda_1, \ldots, \lambda_n)$ aus $\mathbb{K}$. (Basisdarstellung zur fixierten Basis.) Es gibt k^n-viele derartiger n-Tupel. ⊠

Beispiel 5.19: Bei den Vektorräumen ist entscheidend, welcher Körper zugrunde liegt. Es ist also für einen Körper $\mathbb{K}$ die Struktur $\mathbb{K}$ insbesondere ein $\mathbb{K}$-Vektorraum, aber auch ein $\mathbb{K}'$-Vektorraum für geeignete[1] Teilkörper $\mathbb{K}' \subseteq \mathbb{K}$.
So ist $\mathbb{C}$ ein $\mathbb{R}$-Vektorraum, ein $\mathbb{Q}$-Vektorraum, sowie $\mathbb{R}$ auch ein $\mathbb{Q}$-Vektorraum ist. Beachten Sie, dass bei verschiedenen Grundkörpern sich die Längen der Basen, also die Dimensionen, ändern können.

Die Dimension des $\mathbb{C}$-Vektorraums $\mathbb{C}$ ist offenbar 1, denn sowohl die Folge (1) aber auch die Folge (i) ist jeweils eine Basis. Dagegen ist die Dimension der komplexen Zahlen als $\mathbb{R}$-Vektorraum 2, denn wir haben etwa die Folge $(1, \mathrm{i})$ als Basis. Die komplexen Zahlen als $\mathbb{Q}$-Vektorraum sind aus Gründen der Mächtigkeit nicht endlich dimensional.

Achtung!

$\mathbb{R}$ ist kein $\mathbb{C}$-Vektorraum, da in einem solchen Fall die Skalarmultiplikation nicht wohldefiniert ist (da z.B. $\mathrm{i} \cdot 1 = \mathrm{i} \notin \mathbb{R}$ gilt).

Beispiel 5.20: Betrachten Sie die Vektoren $1, \sqrt{2}, \sqrt{3}$ im $\mathbb{Q}$-Vektorraum $\mathbb{R}$. Dann sind diese Vektoren linear unabhängig.
Dies gilt, weil: Sei $a \cdot 1 + b \cdot \sqrt{2} + c \cdot \sqrt{3} = 0$ eine beliebige, fixierte Linearkombination des Nullvektors, wobei $a, b, c \in \mathbb{Q}$. Wir wollen zeigen, dass daraus folgt: $a = b = c = 0$. Zunächst formen wir um und erhalten:

$$b \cdot \sqrt{2} + c \cdot \sqrt{3} = -a \in \mathbb{Q}, \qquad (-a)^2 = 2b^2 + 2bc \cdot \sqrt{6} + 3c^2 \in \mathbb{Q}$$

[1] Wir werden an dieser Stelle nicht weiter spezifizieren, welche Teilkörper hier in Frage kommen. Wenn Sie mehr darüber wissen möchten, dann spielt der Begriff *Charakteristik eines Körpers* eine große Rolle.

Fall 1 Seien $b, c \neq 0$. Dann gilt offenbar:

$$\sqrt{6} = \frac{1}{2bc}\left((-a)^2 - 2b^2 - 3c^2\right) \in \mathbb{Q},$$

allerdings ist $\sqrt{6} \notin \mathbb{Q}$. Dies führt zum Widerspruch.

Fall 2 Sei $b \neq 0$ und $c = 0$. Somit erhalten wir:

$$\sqrt{2} = -\frac{a}{b} \in \mathbb{Q},$$

allerdings ist $\sqrt{2} \notin \mathbb{Q}$. Dieses führt zum Widerspruch.

Fall 3 Sei $b = 0$ und $c \neq 0$. In diesem Fall folgt:

$$\sqrt{3} = -\frac{a}{c} \in \mathbb{Q},$$

allerdings ist $\sqrt{3} \notin \mathbb{Q}$, Widerspruch.

Damit müssen $b, c = 0$ sein, also insgesamt: $a = b = c = 0$.

Bemerkung 5.21: Die Vektoren $1, \sqrt{2}, \sqrt{3}$ im $\mathbb{R}$-Vektorraum $\mathbb{R}$ sind linear abhängig. Es gilt offenbar:

$$\sqrt{2} = \sqrt{2} \cdot 1,$$

wobei die linke Seite der Gleichung als Element in V aufgefasst wird, die rechte Seite als skalare Multiplikation mit $\sqrt{2} \in \mathbb{K} = \mathbb{R}$ und dem Element $1 \in V$. Analoges gilt für $\sqrt{3}$.

Beispiel 5.22: Wir betrachten nun einen sehr wichtigen endlich-dimensionalen Vektorraum, der nicht unterschätzt werden sollte, nämlich der Vektorraum der Polynome n-ten Grades über einen unendlichen Körper $\mathbb{K}$, den wir mit $\mathbb{K}[x]$ bezeichnen. Dabei handelt es sich um alle Funktionen[2] der Form

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x^1 + a_0 x^0$$

mit skalaren Vielfachen $a_n, a_{n-1}, \ldots, a_0 \in \mathbb{K}$. Die einfachsten Vertreter dieser Klasse sind lineare Funktionen, etwa $f(x) = ax + b$, quadratische Funktionen, etwa $f(x) = x^2$ oder auch kubische Funktionen, etwa $f(x) = x^3$, die Ihnen bestimmt schon aus der Schule vertraut sind.

Bitte machen Sie sich an dieser Stelle klar, dass es sich hier wirklich um einen Vektorraum handelt (die Summe und skalaren Vielfache von Polynomen sind wieder Polynome etc. ...). Wie bestimmt man nun eine Basis des Polynom(vektor)raums?

Als Erstes liegt es nahe die Menge der Monome bis zum Grad n, also

$$b_0(x) = x^0 = 1, \quad b_1(x) = x^1, \quad b_2 = x^2, \quad \ldots \quad , \quad b_n(x) = x^n$$

als Basis zu vermuten. In der Tat gilt für jedes $p(x) \in \mathbb{K}[x]$:

$$\begin{aligned} p(x) &= a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x^1 + a_0 x^0 \\ &= a_n \cdot b_n + a_{n-1} \cdot b_{n-1} + \ldots + a_1 \cdot b_1 + a_0 \cdot b_0 \end{aligned}$$

[2]Sie mögen bemerkt haben, dass wir hier die Einschränkung haben, nur unendliche Grundkörper zu betrachten. Grundsätzlich können wir auch Polynome über endliche Körper ins Spiel bringen. Dann müssen wir allerdings zwischen formalen Termausdrücken und den dahinter stehenden Funktionen wohl unterscheiden. So ist etwa der Ausdruck $x^k - x$ zum Ausdruck 0 zu unterscheiden, die dahinter zu verstehenden Funktionen allerdings sind gleich, sowie der Körper genau k Elemente enthält. Dieser Sachverhalt sollte Sie an das Rechnen mit Restklassen erinnern.

also besitzt jedes $p(x)$ eine Darstellung als Linearkombination der b_i. Die b_i sind auch wirklich linear unabhängig, denn:

$$\lambda_n \cdot b_n + \lambda_{n-1} \cdot b_{n-1} + \ldots + \lambda_1 \cdot b_1 + \lambda_0 \cdot b_0 = 0 \quad \rightsquigarrow \quad \lambda_n = \lambda_{n-1} = \ldots = \lambda_0 = 0$$

Damit haben wir gezeigt, dass die b_i eine Basis von $\mathbb{K}[x]$ bilden.

Schauen wir uns jetzt die Polynome $c_i(x) := (x+1)^i,\ i = 0, \ldots, n$, an.
Bilden die c_i eine Basis des $\mathbb{K}[x]$?

Dazu nehmen wir uns wieder ein beliebiges $p(x) \in \mathbb{K}[x]$, wobei

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x^1 + a_0,$$

und multiplizieren die c_i aus:

$$\begin{aligned}
c_0(x) &= (x+1)^0 \ = \ 1 \\
c_1(x) &= (x+1) \\
c_2(x) &= x^2 + 2x + 1 \\
c_3(x) &= x^3 + 3x^2 + 3x + 1 \\
c_i(x) &= \mu_{i,i} x^i + \mu_{i,i-1} x^{i-1} + \ldots + \mu_{i,0} x^0
\end{aligned}$$

Wenn wir jetzt eine Darstellung von $p(x)$ als Linearkombination der c_i erhalten möchten, dann müssen wir $\lambda_0, \ldots, \lambda_n \in \mathbb{K}$ finden, so dass:

$$\begin{aligned}
p(x) = \quad & a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x^1 + a_0 \\
= \quad & \lambda_n \cdot c_n + \lambda_{n-1} \cdot c_{n-1} + \ldots + \lambda_1 \cdot c_1 + \lambda_0 \cdot c_0 \\
= \quad & \lambda_n \left(\mu_{n,n} x^n + \mu_{n,n-1} x^{n-1} + \ldots + \mu_{n,0} \right) + \\
& \qquad + \lambda_{n-1} \left(\mu_{n-1,n-1} x^{n-1} + \mu_{n-1,n-2} x^{n-2} + \ldots + \mu_{n-1,0} \right) + \ldots + \lambda_0 \cdot 1
\end{aligned}$$

Dadurch gelangen wir zu den folgenden Gleichungen:

$$\begin{aligned}
a_0 = \ & \lambda_0 + \lambda_1 \cdot \mu_{1,0} + \lambda_2 \cdot \mu_{2,0} + \ldots + \lambda_n \cdot \mu_{n,0} \\
a_1 = \ & \lambda_1 \cdot \mu_{1,1} + \lambda_2 \cdot \mu_{2,1} + \ldots + \lambda_n \cdot \mu_{n,1} \\
\vdots \ & \\
a_n = \ & \lambda_n \cdot \mu_{n,n}
\end{aligned}$$

Wir schreiben dies jetzt als LGS mit $\lambda_0, \ldots, \lambda_n$ als Unbekannte:

$$\begin{array}{ccccc|c}
1 & \mu_{1,0} & \mu_{2,0} & \cdots & \mu_{n,0} & a_0 \\
0 & \mu_{1,1} & \mu_{2,1} & \cdots & \mu_{n,1} & a_1 \\
0 & 0 & \mu_{2,2} & \cdots & \mu_{n,2} & a_2 \\
\vdots & & \ddots & \ddots & \vdots & \vdots \\
0 & \cdots & \cdots & 0 & \mu_{n,n} & a_n
\end{array}$$

Diesen Typ Gleichungssystem(Zeilenstufenform) haben wir uns aber bereits im Kapitel 1 angesehen. Man kann hier durch Auflösen des Systems von unten nach oben und von hinten nach vorne die λ_i bestimmen, da alle Pivotelemente nicht Null sind.

Darüberhinaus stellen wir mit diesem LGS schnell fest, dass die c_i linear unabhängig sind: Der Ansatz $\lambda_0 \cdot c_0 + \lambda_1 \cdot c_1 + \ldots + \lambda_n \cdot c_n = 0$ bringt uns mit der obigen Rechnung zu folgendem LGS:

$$\begin{array}{ccccc|c} 1 & \mu_{1,0} & \mu_{2,0} & \cdots & \mu_{n,0} & 0 \\ 0 & \mu_{1,1} & \mu_{2,1} & \cdots & \mu_{n,1} & 0 \\ 0 & 0 & \mu_{2,2} & \cdots & \mu_{n,2} & 0 \\ \vdots & & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & \cdots & 0 & \mu_{n,n} & 0 \end{array}$$

Daraus folgt aber wieder $\lambda_n = \lambda_{n-1} = \ldots = \lambda_0 = 0$

Beispiel 5.23: Wir wollen das gerade Gelernte nun an einem praktischen Beispiel anwenden. Betrachten Sie die folgenden Polynome über $\mathbb{R}$ und überprüfen Sie diese auf lineare Unabhängigkeit:

$$p_1(x) = 3x^2 + 2x \qquad p_2(x) = 2x^2 + x + 2 \qquad p_3(x) = x^2 + 2x + 4$$

Das Überprüfen der linearen Unabhängigkeit sollte für Sie bereits Routine sein:

$$\begin{aligned} 0 &= \lambda_1 p_1(x) + \lambda_2 p_2(x) + \lambda_3 p_3(x) \\ &= (3\lambda_1 + 2\lambda_2 + \lambda_3)\, x^2 + (2\lambda_1 + \lambda_2 + 2\lambda_3)\, x + (2\lambda_2 + 4\lambda_3) \end{aligned}$$

Denken Sie daran, dass die Monome x^2, x^1, x^0 eine Basis unseres Polynomraumes sind und wir deshalb das Nullpolynom bzgl. dieser Basis nur durch $0 = 0 \cdot x^2 + 0 \cdot x^1 + 0 \cdot x^0$ darstellen können. Dadurch gelangen wir zu folgenden Gleichungen:

$$3\lambda_1 + 2\lambda_2 + \lambda_3 = 0 \qquad 2\lambda_1 + \lambda_2 + 2\lambda_3 = 0 \qquad 2\lambda_2 + 4\lambda_3 = 0$$

oder als LGS:

$$\left(\begin{array}{ccc|c} 3 & 2 & 1 & 0 \\ 2 & 1 & 2 & 0 \\ 0 & 2 & 4 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|c} 3 & 2 & 1 & 0 \\ 0 & 1 & -4 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|c} 3 & 2 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 6 & 0 \end{array}\right)$$

Wir erhalten $\lambda_1 = \lambda_2 = \lambda_3 = 0$, also sind unsere Polynome linear unabhängig.

Beispiel 5.24: Wir betrachten nun die Menge der $(n \times n)$ quadratischen Zahlenschemata mit Einträgen in $\mathbb{R}$, wobei alle Zeilen- und Spaltensummen Null ergeben sollen. Nun überlegen Sie sich bitte, dass diese so definierten „Sudokus" einen Vektorraum bilden. Welche Dimension hat dieser Vektorraum?

Dies ist ganz einfach: Um unsere Sudokus zu charakterisieren brauchen wir für jede Position in unserem Zahlenschema eine Variable und führen für alle Zeilen- und Spaltensummen eine Gleichung ein. Dadurch erhalten wir ein lineares homogenes Gleichungssystem mit n^2 Unbekannten und $2n$ Gleichungen. In unserem Gleichungssystem ist jede Zeile genau n^2 lang und es stehen genau n Einsen an jeweils unterschiedlichen Stellen. Daraus folgt, dass wir dieses LGS in Zeilenstufenform bringen müssen und beim Berechnen der Lösung die Anzahl der frei wählbaren Unbekannten als die Dimension des Vektorraums erhalten.

Im Fall $n = 3$ sieht das beispielsweise wie folgt aus: Wir starten mit unserem Quadrat:

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

und kommen auf folgende Gleichungen:

$$\begin{aligned} x_1+x_2+x_3&=0\\ x_4+x_5+x_6&=0\\ x_7+x_8+x_9&=0\\ x_1+x_4+x_7&=0\\ x_2+x_5+x_8&=0\\ x_3+x_6+x_9&=0 \end{aligned}$$

nach einigen Umformungsschritten erhalten wir:

$$\begin{array}{ccccccccc|c} 1&1&1&0&0&0&0&0&0&0\\ 0&0&0&1&1&1&0&0&0&0\\ 0&0&0&0&0&0&1&1&1&0\\ 1&0&0&1&0&0&1&0&0&0\\ 0&1&0&0&1&0&0&1&0&0\\ 0&0&1&0&0&1&0&0&1&0 \end{array} \rightsquigarrow \begin{array}{ccccccccc|c} 1&0&0&1&0&0&1&0&0&0\\ 0&1&0&0&1&0&0&1&0&0\\ 0&0&1&0&0&1&0&0&1&0\\ 0&0&0&-1&-1&-1&-1&-1&-1&0\\ 0&0&0&1&1&1&0&0&0&0\\ 0&0&0&0&0&0&1&1&1&0 \end{array} \rightsquigarrow$$

$$\begin{array}{ccccccccc|c} 1&0&0&0&-1&-1&0&-1&-1&0\\ 0&1&0&0&1&0&0&1&0&0\\ 0&0&1&0&0&1&0&0&1&0\\ 0&0&0&1&1&1&0&0&0&0\\ 0&0&0&0&0&0&1&1&1&0 \end{array}$$

Die Lösungsmenge dieses LGS ist somit:

$$\mathbb{L}=\left\{ \lambda_1\begin{pmatrix}1\\-1\\0\\-1\\1\\0\\0\\0\\0\end{pmatrix}+\lambda_2\begin{pmatrix}1\\0\\-1\\-1\\0\\1\\0\\0\\0\end{pmatrix}+\lambda_3\begin{pmatrix}1\\-1\\0\\0\\0\\0\\-1\\1\\0\end{pmatrix}+\lambda_4\begin{pmatrix}1\\0\\-1\\0\\0\\0\\-1\\0\\1\end{pmatrix} \;\middle|\; \lambda_1,\lambda_2,\lambda_3,\lambda_4\in\mathbb{R} \right\}$$

und wie Sie sicher schon festgestellt haben, sind die zu den λ_i gehörigen Vektoren eine Basis des Vektorraums, oder wenn wir diese in Form unserer Zahlenschemata schreiben wollen, sieht die Basis wie folgt aus:

$$b_1=\begin{pmatrix}1&-1&0\\-1&1&0\\0&0&0\end{pmatrix},\quad b_3=\begin{pmatrix}1&0&-1\\-1&0&1\\0&0&0\end{pmatrix},\quad b_2=\begin{pmatrix}1&-1&0\\0&0&0\\-1&1&0\end{pmatrix},\quad b_4=\begin{pmatrix}1&0&-1\\0&0&0\\-1&0&1\end{pmatrix}$$

Beispiel 5.25: Jetzt noch eine kleine Übung in Sachen Basisergänzungssatz: Wir wollen beispielsweise nachfolgende Menge von Vektoren zu einer Basis des $\mathbb{R}^5$ ergänzen:

$$u_1 = \begin{pmatrix} 2 \\ 3 \\ 0 \\ 8 \\ 0 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 4 \\ 0 \end{pmatrix}, \quad u_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

Wie fängt man an? – Ganz einfach: Als Erstes müssen wir überprüfen, ob die Vektoren linear unabhängig sind, sonst ist unser Unterfangen unmöglich: Man kann keine linear abhängige Menge von Vektoren zu einer Basis ergänzen, weil eine Basis immer aus linear unabhängigen Vektoren bestehen muss. Die ganze Idee besteht nun darin, solche Linearkombinationen unserer Ausgangsvektoren zu betrachten, die wir besonders einfach (durch kanonische Einheitsvektoren) zu einer Basis des $\mathbb{R}^5$ ergänzen können.

Das Ganze wird Sie an die Zeilenstufenform erinnern, die wir aus dem Kapitel über Lineare Gleichungssysteme kennen. Das Gute an dieser maximalen Zeilenstufenform ist, dass wir nur Einsen und Nullen in den Zeilenvektoren stehen haben und bei diesen ist es besonders einfach eine Ergänzung zur Basis zu finden. Wenn wir beispielsweise

$$a_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad a_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad a_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

angeben würden, dann wäre die naheliegendste Ergänzung zu einer Basis durch folgende Vektoren gegeben:

$$e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad e_5 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

Wir schreiben nun unsere Ausgangsvektoren u_1, u_2, u_3 als Zeilenvektoren in ein Gleichungssystem ohne rechte Seite:

$$\begin{pmatrix} 2 & 3 & 0 & 8 & 0 \\ 1 & 1 & 0 & 4 & 0 \\ 1 & 1 & 0 & 1 & 0 \end{pmatrix}$$

Dieses bringen wir mit den uns erlaubten Umformungsschritten in Zeilenstufenform.

$$\begin{pmatrix} 2 & 3 & 0 & 8 & 0 \\ 1 & 1 & 0 & 4 & 0 \\ 1 & 1 & 0 & 1 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 2 & 3 & 0 & 8 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 6 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Wir haben nun aus unseren Ausgangsvektoren eine Linearkombination (mit derselben linearen Hülle) erzeugt, welche wir auf recht einfache Art und Weise (nämlich durch kanonische Einheitsvektoren) ergänzen können. Insgesamt ergänzen wir also u_1, u_2, u_3 durch

$$e_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad e_5 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

zu einer Basis des $\mathbb{R}^5$.

6. Lineare Abbildungen

In diesem Zusammenhang betrachten wir Abbildungen zwischen Vektorräumen, die die Vektorraumstruktur respektieren. Diese Abbildungen sind in der Linearen Algebra so grundlegend wie Lineare Gleichungssysteme.

Grundlagen Linearer Abbildungen. Wir definieren, was wir als Homomorphismus zwischen den Vektorraumstrukturen verstehen. Diese Definition entspricht der Definition des bekannten Begriffs vom Homomorphismus zwischen allgemeinen Strukturen.

Definition 6.1 – lineare Abbildung/Homomorphismus:

Seien V und W zwei $\mathbb{K}$- Vektorräume und $f : V \to W$. Dann ist die Abbildung f linear oder ein Homomorphismus von V nach W, wenn

(a) für alle $x, y \in V$ gilt:
$$f(x) + f(y) = f(x + y)$$

(b) für alle $x \in V$ und $\lambda \in \mathbb{K}$ gilt:
$$f(\lambda \cdot x) = \lambda \cdot f(x)$$

Satz 6.2:

Seien V, W, X drei $\mathbb{K}$-Vektorräume und $f : V \to W$ und $g : W \to X$ zwei lineare Abbildungen. Dann sind die beiden Abbildungen $g \circ f : V \to X$ und $\mathrm{id}_V : V \to V$ linear.

Beweis: Betrachte $x, y \in V$ und $\lambda \in \mathbb{K}$. Dann ist

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

und

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

Es ist klar, dass id_V linear ist. ⊠

Beispiel 6.3: Die anschaulichsten Beispiele für lineare Abbildungen finden sich in der 2- und 3-dimensionalen analytischen Geometrie.

(a) *Streckung* („Zoomen"): Sei $\lambda \in \mathbb{R}$ und definiere die Abbildung

$$f : \mathbb{R}^2 \to \mathbb{R}^2 \; ; \; f(x) = \lambda \cdot x$$

durch Skalar-Multiplikation mit dem Faktor λ. Wir hatten bereits gesehen, dass die Skalar-Multiplikation mit λ eine Vergrößerung um den Faktor λ darstellt. Die Abbildung f ist linear, denn für alle $x, y \in \mathbb{R}^2$ und $\lambda \in \mathbb{R}$ gilt unter Benutzung der Vektorraum-Axiome:

$$f(x + y) = \lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y = f(x) + f(y)$$

und

$$f(\mu \cdot x) = \lambda \cdot (\mu \cdot x) = (\lambda \cdot \mu) \cdot x = (\mu \cdot \lambda) \cdot x = \mu \cdot (\lambda \cdot x) = \mu \cdot f(x).$$

(b) *Komplexe Multiplikation:* Die komplexe Multiplikation war auf $\mathbb{C} = \mathbb{R} \times \mathbb{R}$ definiert durch

$$(x, y) \cdot (a, b) = (xa - yb, xb + ya).$$

Sei $z = (x, y) = x + \mathrm{i}y \in \mathbb{C}$ und definiere die Abbildung

$$f : \mathbb{R}^2 \to \mathbb{R}^2 \ ; \ f(a, b) = (x, y) \cdot (a, b) = z \cdot (a, b).$$

Die Linearität der Abbildung f folgt ähnlich wie in (a) aus den Körper-Axiomen für $\mathbb{C}$.

(c) *Drehung des* $\mathbb{R}^2$: Wir hatten gesehen, dass die Multiplikation mit der imaginären Einheit i einer Drehung um $\frac{\pi}{2}$ entspricht. Damit ist eine Drehung des $\mathbb{R}^2$ um $\frac{\pi}{2}$ linear. Auch Drehungen um andere Winkel sind lineare Abbildungen.

Beispiel 6.4: Betrachten wir die Abbildung $f : \mathbb{Q} \times \mathbb{Q} \to \mathbb{R}$, gegeben durch $f(x, y) = x + \sqrt{2}y$. Hierbei fassen wir den Urbildraum $\mathbb{Q} \times \mathbb{Q} = \mathbb{Q}^2$ als $\mathbb{Q}$-Vektorraum auf. Dann ist f linear.

Es ist zu zeigen:

(i) $f\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) = f\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}\right) + f\left(\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right)$ und

(ii) $f\left(\lambda \cdot \begin{pmatrix} x \\ y \end{pmatrix}\right) = \lambda \cdot f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right)$

Zu (i):

$$\begin{aligned} f\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) &= f\left(\begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix}\right) &&= (x_1 + x_2) + \sqrt{2}(y_1 + y_2) \\ &= x_1 + \sqrt{2}y_1 + x_2 + \sqrt{2}y_2 &&= f\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}\right) + f\left(\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right) \end{aligned}$$

Zu (ii):

$$\begin{aligned} f\left(\lambda \cdot \begin{pmatrix} x \\ y \end{pmatrix}\right) &= f\left(\begin{pmatrix} \lambda \cdot x \\ \lambda \cdot y \end{pmatrix}\right) &&= \lambda \cdot x + \sqrt{2}\lambda \cdot y \\ &= \lambda \cdot (x + \sqrt{2}y) &&= \lambda \cdot f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) \end{aligned}$$

Beispiel 6.5: Sei P_4 der Vektorraum der Polynome vom Grade kleiner gleich 3 mit reellen Koeffizienten. Es sei $F : P_4 \to P_4$ die Abbildung $F\big(p(x)\big) := \frac{\mathrm{d}}{\mathrm{d}x}\big(x \cdot p(x)\big)$. Diese Abbildung ist ebenfalls linear, denn es gilt für beliebige Vektoren $p(x)$, $p_1(x)$, $p_2(x)$ (Polynome) des Raums P_4 und einem Skalar $\lambda \in \mathbb{R}$:

$$\begin{aligned} F\big(p_1(x) + p_2(x)\big) &= \frac{\mathrm{d}}{\mathrm{d}x}\Big[x \cdot \big(p_1(x) + p_2(x)\big)\Big] &&= \frac{\mathrm{d}}{\mathrm{d}x}\big(x \cdot p_1(x) + x \cdot p_2(x)\big) \\ &= \frac{\mathrm{d}}{\mathrm{d}x}\big(x \cdot p_1(x)\big) + \frac{\mathrm{d}}{\mathrm{d}x}\big(x \cdot p_2(x)\big) &&= F\big(p_1(x)\big) + F\big(p_2(x)\big) \end{aligned}$$

$$F\left(\lambda \cdot p(x)\right) = \frac{\mathrm{d}}{\mathrm{d}x}\left(x \cdot \left(\lambda \cdot p(x)\right)\right) = \lambda \cdot \frac{\mathrm{d}}{\mathrm{d}x}\left(x \cdot p(x)\right)$$
$$= \lambda \cdot F\left(p(x)\right)$$

Sie sehen, wenn Sie die Definition von Linearität verstanden haben, ist das Nachprüfen nicht mehr schwer!

⋆ ⋆ ⋆

Wir definieren nun zwei interessante Mengen für eine gegebene lineare Abbildung, die sich als sehr nützlich herausstellen werden.

Definition 6.6 – Kern und Bild:

Seien V und W zwei $\mathbb{K}$-Vektorräume und $f : V \to W$ linear.

(a) Das Urbild $f^{-1}\left[\{0\}\right]$ ist die Menge $\left\{v \in V \mid f(v) = 0\right\}$ und wird als **Kern** der Abbildung f, als $f^{-1}(0)$ oder auch als $\mathrm{Kern}(f)$ bezeichnet.

(b) Das Bild von f ist die Menge $\left\{w \in W \mid \exists v \in V : f(v) = w\right\}$ und wird mit $f[V]$ oder auch $\mathrm{Bild}(f)$ bezeichnet.

Schließlich entpuppen sich diese Mengen als Untervektoräume:

Satz 6.7:

Seien V und W zwei $\mathbb{K}$-Vektorräume und $f : V \to W$ linear. Dann gilt:

(a) Die Menge $\mathrm{Kern}(f)$ ist ein Untervektorraum von V.

(b) Die Menge $\mathrm{Bild}(f)$ ist ein Untervektorraum von W.

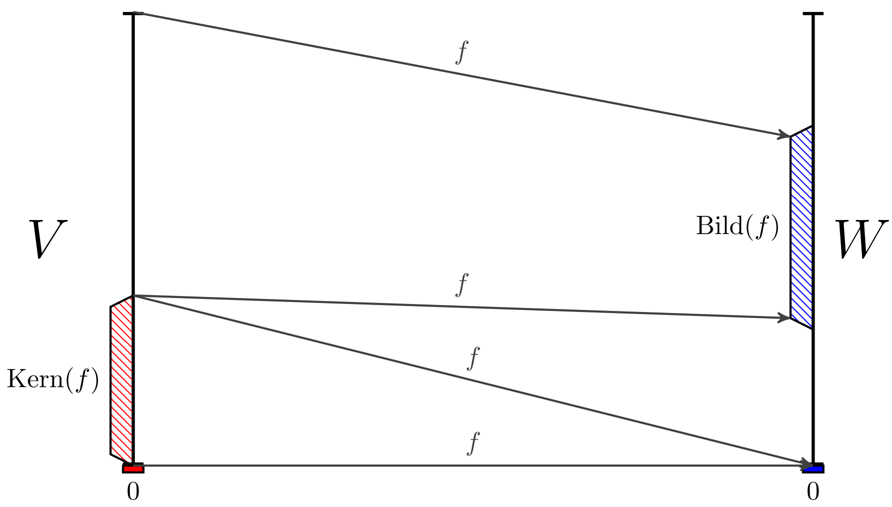


Abbildung 10: Bild und Kern einer Funktion f

Beweis: Wir beweisen jeweils die Eigenschaften eines Untervektorraums:

Kern(f) Es gilt $f(0) = 0$, also $0 \in f^{-1}[\{0\}]$, also Kern(f) $\neq \varnothing$. Betrachte $x, y \in$ Kern(f) und $\lambda \in \mathbb{K}$. Dann gilt $0 = f(x) = f(y)$ und somit

$$f(x + y) = f(x) + f(y) = 0 + 0 = 0$$

also ist $x + y \in$ Kern(f). Ebenso gilt:

$$f(\lambda \cdot x) = \lambda \cdot f(x) = \lambda \cdot 0 = 0$$

also ist $\lambda \cdot x \in$ Kern(f).

Bild(f) Es gilt: $0 = f(0) \in f[V]$, also ist das Bild von f nicht leer. Betrachte beliebige $x, y \in f[V]$ und $\lambda \in \mathbb{K}$. Wähle $\bar{x}, \bar{y} \in V$ mit $x = f(\bar{x})$ und $y = f(\bar{y})$. Dann gilt

$$x + y = f(\bar{x}) + f(\bar{y}) = f(\bar{x} + \bar{y}) \in f[V]$$

und

$$\lambda \cdot x = \lambda \cdot f(\bar{x}) = f(\lambda \cdot \bar{x}) \in f[V].$$

⊠

Definition 6.8 – Rang und Defekt:

Die Dimension des Bildes einer linearen Abbildung bezeichnet man als **Rang** dieser Abbildung, d.h. für $f : V \to W$ linear schreiben wir kurz für den Rang

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

Die Dimension des Kerns wird als **Defekt** bezeichnet, kurz:

$$\mathrm{Df}(f) := \dim(\mathrm{Kern}(f)).$$

Der folgende Satz hilft uns, wenn wir eine Basis des Bildes finden wollen:

Satz 6.9:

Sei $F : V \to W$ eine lineare Abbildung und $V = \mathcal{L}(v_1, \ldots, v_n)$. Dann gilt:

$$\mathrm{Bild}(F) = \mathcal{L}(F(v_1), \ldots, F(v_n)).$$

Beweis:
Wir zeigen beide Inklusionen:

„$\supseteq$"

Sei $v \in \mathcal{L}(F(v_1), \ldots, F(v_n))$. Dann existieren $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$, so dass

$$v = \sum_{i=1}^{n} \lambda_i F(v_i) = F\left(\sum_{i=1}^{n} \lambda_i v_i\right) \in \mathrm{Bild}(F)$$

„$\subseteq$"

Sei $v \in \mathrm{Bild}(F)$, das heißt es existiert ein $\bar{v} \in V$ mit $v = F(\bar{v})$. Da $V = \mathcal{L}(v_1, \ldots, v_n)$ existieren $\lambda_i \in \mathbb{K}$ mit $\bar{v} = \sum_{i=1}^{n} \lambda_i v_i$. Damit gilt

$$v = F(\bar{v}) = F\left(\sum_{i=1}^{n} \lambda_i v_i\right) = \sum_{i=1}^{n} \lambda_i F(v_i)$$

somit ist $v \in \mathcal{L}(F(v_1), \ldots, F(v_n))$.

$\boxtimes$

Beispiel 6.10: Sei $F : \mathbb{R}^3 \to \mathbb{R}^2$ mit

$$F\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} x_1 + 2x_2 + 3x_3 \\ 4x_1 + 5x_2 + 6x_3 \end{pmatrix}.$$

Gesucht sind Basen von Kern(f) und Bild(f).

Bild(f): Offenbar gilt $\mathbb{R}^3 = \mathcal{L}(e_1, e_2, e_3)$. Nach Satz 6.9 gilt:

$$\text{Bild}(F) = \mathcal{L}\left(F(e_1), F(e_2), F(e_3)\right) = \mathcal{L}\left(\begin{pmatrix} 1 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 5 \end{pmatrix}, \begin{pmatrix} 3 \\ 6 \end{pmatrix}\right)$$

Eine Basis bilden etwa $\begin{pmatrix} 1 \\ 4 \end{pmatrix}$ und $\begin{pmatrix} 2 \\ 5 \end{pmatrix}$.

Kern(f): Später in diesem Kapitel werden wir den Dimensionssatz 6.19 beweisen, der besagt, dass die Summe des Rangs und des Defekts eines Homomorphismus gleich der Dimension des Urbild-Vektorraums ist:

$$\text{Df}(F) = \dim\left(\text{Kern}(F)\right) = \dim(V) - \dim\left(\text{Bild}(F)\right) = 3 - 2 = 1$$

Daher wüssten wir schon, welche Dimension die gesuchte Lösungsmenge hat, die wir jetzt durch das Lösen bestimmen:

$$\left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 4 & 5 & 6 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & -3 & -6 & 0 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|c} 1 & 2 & 3 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right)$$

Sei etwa x_3 beliebig gewählt. Dann gilt: $x_2 = -2x_3$ und

$$x_1 = 0 - 3x_3 - 2x_2 = -3x_3 + 4x_3 = x_3,$$

so dass wir schließlich als Lösungsmenge des homogenen Gleichungssystems und somit auch als Kern der Abbildung F erhalten:

$$\begin{aligned} \text{Lös}(A, b) &= \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 \;\middle|\; x_1 = x_3, x_2 = -2x_3\right\} \\ &= \left\{\begin{pmatrix} t \\ -2t \\ t \end{pmatrix} \;\middle|\; t \in \mathbb{R}\right\} \end{aligned}$$

Eine Basis des Kerns von F ist mit dem Vektor $\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ gegeben.

⋆ ⋆ ⋆

Im Folgenden betrachten wir die Homomorphismen noch einen Schritt abstrakter: Wir führen Operationen auf der Menge der linearen Abbildungen ein und betrachten damit die Menge der Homomorphismen selbst als Vektorraum.

Definition 6.11 – $\mathrm{Hom}(V, W)$:

Seien V, W zwei $\mathbb{K}$-Vektorräume. Definiere die Struktur
$\mathrm{Hom}(V, W) := (\mathrm{Hom}(V, W), +, \cdot, 0)$ aller Homomorphismen von V nach W

$$\mathrm{Hom}(V, W) := \left\{ f \;\middle|\; f : V \to W \text{ linear} \right\}$$

Definiere die Addition „+“ durch:

$$+ \;:\; \mathrm{Hom}(V, W) \times \mathrm{Hom}(V, W) \to \mathrm{Hom}(V, W)$$

$$(f + g)(x) := f(x) + g(x)$$

Definiere die skalare Multiplikation „·“ durch:

$$\cdot \;:\; \mathbb{K} \times \mathrm{Hom}(V, W) \to \mathrm{Hom}(V, W)$$

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

Definiere die Nullabbildung $0 \in \mathrm{Hom}(V, W)$ durch:

$$0(x) := 0$$

Beachten Sie, dass wir eigentlich schreiben müssten:

$$\left(f +_{\mathrm{Hom}(V,W)} g \right)(x) = f(x) +_W g(x),$$

ebenso bei $(\lambda \cdot_{\mathrm{Hom}(V,W)} f)(x) = \lambda \cdot_W f(x)$ und $0_{\mathrm{Hom}(V,W)}(x) = 0_W$. Aber wir unterdrücken diese Art der Signatur.

Satz 6.12:

Seien V, W zwei $\mathbb{K}$-Vektorräume, dann ist $\mathrm{Hom}(V, W)$ ein $\mathbb{K}$-Vektorraum.

Beweis: Wir müss(t)en die Vektorraum-Axiome der Reihe nach prüfen. Exemplarisch zeigen wir die Kommutativität: $f + g = g + f$:

Diese ist allerdings klar, denn es gilt:

$(f + g)(x)$	$=$	$f(x) + g(x)$	per Definition
	$=$	$g(x) + f(x)$	W ist Vektorraum
	$=$	$(g + f)(x)$	per Definition

$\boxtimes$

Definition 6.13 – Morphismen:

Sei $f : V \to W$ ein Homomorphismus. Dann definieren wir:

(1) f ist ein Monomorphismus, wenn f injektiv ist.
(2) f ist ein Epimorphismus, wenn f surjektiv ist.
(3) f ist ein Isomorphismus, wenn f bijektiv ist.
(4) f ist ein Endomorphismus, wenn $V = W$ ist.
(5) f ist ein Automorphismus, wenn f bijektiv und $V = W$ ist.

Satz 6.14:

Sei $f : V \to W$ ein Isomorphismus von $\mathbb{K}$-Vektorräumen. Dann ist ebenfalls $f^{-1} : W \to V$ ein Isomorphismus.

Beweis: Da f bijektiv ist, dürfen wir auch von der Umkehrabbildung f^{-1} sprechen. Die Bijektivität von $f^{-1} : W \to V$ ist klar. Wir zeigen, dass f^{-1} linear ist: Betrachte $\lambda \in \mathbb{K}$ und $w_0, w_1 \in W$. Wegen der Surjektivität von f existieren $v_0, v_1 \in V$ mit $f(v_0) = w_0$ und $f(v_1) = w_1$, also gilt auch: $v_i = f^{-1}(w_i)$ für $i = 0, 1$. Dann gilt:

$$\begin{aligned} f^{-1}(w_0 + w_1) &= f^{-1}\big(f(v_0) + f(v_1)\big) && \\ &= f^{-1}\big(f(v_0 + v_1)\big) && \text{wegen der Linearität von } f \\ &= v_0 + v_1 &=& f^{-1}(w_0) + f^{-1}(w_1) \end{aligned}$$

und

$$\begin{aligned} f^{-1}(\lambda \cdot w_0) &= f^{-1}\big(\lambda \cdot f(v_0)\big) && \\ &= f^{-1}\big(f(\lambda \cdot v_0)\big) && \text{Linearität von } f \\ &= \lambda \cdot v_0 &=& \lambda \cdot f^{-1}(w_0) \end{aligned}$$

$\boxtimes$

Beispiel 6.15: Setze $\mathrm{Aut}(V) := \big\{ f : V \to V \;\big|\; f \text{ Automorphismus} \big\}$. Mittels der Hintereinanderausführung von Funktionen $\circ$ bildet

$$\big(\mathrm{Aut}(V), \circ\big)$$

eine Gruppe, die so genannte *Automorphismengruppe*: Die notwendigen Eigenschaften Assoziativität, Existenz des neutralen Elements und von inversen Elementen sind klar.

⋆ ⋆ ⋆

Endliche Beschreibung von Homomorphismen. Der große Vorteil von linearen Abbildungen zwischen endlich dimensionalen Vektorräumen ist der, dass wir diese bereits endlich erfassen können, nämlich bereits durch Angabe der Bilder einer Basis:

Satz 6.16:

Seien V und W $\mathbb{K}$-Vektorräume und sei $(v_1, \dots, v_n)$ eine Basis von V. Seien $w_1, \dots, w_n \in W$. Dann existiert genau eine lineare Abbildung $f : V \to W$, so dass für $i = 1, \dots, n$ gilt: $f(v_i) = w_i$.

Beweis: (Existenz) Für ein $v \in V$ definieren wir $f(v)$ wie folgt: Wähle $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ mit $v = \sum_{i=1}^n \lambda_i \cdot v_i$. Dann setze

$$f(v) := \sum_{i=1}^n \lambda_i \cdot f(v_i).$$

Wegen der Eindeutigkeit der Basisdarstellung von v bezüglich $(v_1, \dots, v_n)$ ist f wohldefiniert. Wir zeigen, dass f linear ist:
Betrachte $\lambda \in \mathbb{K}$, $v, v' \in V$ und wähle $\lambda'_1, \dots, \lambda'_n \in \mathbb{K}$ mit $v = \sum_{i=1}^n \lambda_i \cdot v_i$ und sei $v' = \sum_{i=1}^n \lambda'_i \cdot v_i$. Dann gilt, dass

$$v + v' = \sum_{i=1}^n \lambda_i \cdot v_i + \sum_{i=1}^n \lambda'_i \cdot v_i = \sum_{i=1}^n (\lambda_i + \lambda'_i) \cdot v_i$$

und

$$\lambda \cdot v = \lambda \cdot \sum_{i=1}^n \lambda_i \cdot v_i = \sum_{i=1}^n (\lambda \cdot \lambda_i) \cdot v_i$$

die eindeutigen Basisdarstellungen von $v + v'$ und $\lambda \cdot v$ sind. Damit gilt:

$$f(v + v') = \sum_{i=1}^n (\lambda_i + \lambda'_i) \cdot f(v_i) = \sum_{i=1}^n \lambda_i \cdot f(v_i) + \sum_{i=1}^n \lambda'_i \cdot f(v_i) = f(v) + f(v')$$

Analog folgt : $f(\lambda \cdot v) = \lambda \cdot f(v)$

(Eindeutigkeit) Angenommen $f : V \to W$ und $f' : V \to W$ sind lineare Abbildungen, für die gilt: $f(v_i) = w_i = f'(v_i)$. Wir müssen zeigen, dass $f(v) = f'(v)$ für beliebiges $v \in V$ ist.

Sei $v = \sum_{i=1}^n \lambda_i \cdot v_i$ die Basisdarstellung von v. Dann gilt:

$$f(v) = \sum_{i=1}^n \lambda_i \cdot f(v_i) = \sum_{i=1}^n \lambda_i \cdot w_i = \sum_{i=1}^n \lambda_i \cdot f'(v_i) = f'(v)$$

Damit gilt $f = f'$. ⊠

⋆ ⋆ ⋆

Klassifikation endlich-dimensionaler Vektorräume. Wir werden jetzt einen Satz beweisen, der uns im Wesentlichen zeigt, dass es für festes $n \in \mathbb{N}$ bis auf Isomorphie nur einen Vektorraum gibt, nämlich den $\mathbb{K}^n$ selbst:

Satz 6.17:

Sei $n \in \mathbb{N}$ und V ein n-dimensionaler $\mathbb{K}$-Vektorraum. Dann gibt es einen Isomorphismus

$$f : V \to \mathbb{K}^n$$

Bemerkung 6.18: Das ist hilfreich, weil V ein ganz beliebiger Vektorraum sein kann, der eine beliebig komplizierte Struktur haben kann. Dagegen kennen wir die Struktur $\mathbb{K}^n$ als die Menge der n-dimensionalen Vektoren sehr gut, denn mit denen können wir konkret rechnen. Da wir hier einen Isomorphismus haben, können wir also anstatt auf V zu rechnen, mit Vektoren aus $\mathbb{K}^n$ rechnen.

Beweis von Satz 6.17: Wähle eine Basis $(v_1, \ldots, v_n)$ von V. Dann gibt es eine lineare Abbildung $f : V \to \mathbb{K}^n$ (nach Satz 6.16) mit:

$$\text{Für } i = 1, \ldots, n \text{ gilt } f(v_i) = e_i,$$

wobei $e_1, \ldots, e_n$ die kanonische Basis des $\mathbb{K}^n$ ist.

Wir zeigen, dass $f : V \to \mathbb{K}$ bijektiv ist:

(Surjektivität) Betrachte $(\lambda_1, \ldots, \lambda_n) \in \mathbb{K}^n$. Dann gilt:

$$(\lambda_1, \ldots, \lambda_n) = \sum_{i=1}^{n} \lambda_i \cdot e_i = \sum_{i=1}^{n} \lambda_i \cdot f(v_i) = f\left(\sum_{i=1}^{n} \lambda_i v_i\right) = f(v) \in \mathrm{Bild}(f)$$

(Injektivität) Betrachte $v, v' \in V$ mit $f(v) = f(v')$. Wähle $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ und $\lambda'_1, \ldots, \lambda'_n \in \mathbb{K}$ mit $v = \sum_{i=1}^{n} \lambda_i \cdot v_i$ und $v' = \sum_{i=1}^{n} \lambda'_i \cdot v_i$. Dann gilt:

$$\begin{aligned}
(\lambda_1, \ldots, \lambda_n) &= f\left(\sum_{i=1}^{n} \lambda_i v_i\right) && \text{siehe oben} \\
&= f(v) = f(v') = f\left(\sum_{i=1}^{n} \lambda'_i v_i\right) && \text{nach Voraussetzung} \\
&= (\lambda'_1, \ldots, \lambda'_n)
\end{aligned}$$

Also gilt $\lambda_i = \lambda'_i$ für alle $i = 1, \ldots, n$ und somit gilt:

$$v = \sum_{i=1}^{n} \lambda_i v_i = \sum_{i=1}^{n} \lambda'_i v_i = v'$$

⊠(Satz 6.17)

⋆ ⋆ ⋆

Der Dimensionssatz. Der folgende Satz wird uns noch oft hilfreich zur Seite stehen und erklärt uns, wie lineare Abbildungen mit den Dimensionen des Urbildraums umgehen: Entweder diese werden vom Kern geschluckt oder wir finden diese im Bild wieder.

Satz 6.19 – Dimensionssatz:

Seien V, W zwei $\mathbb{K}$-Vektorräume, wobei V endlich dimensional ist mit $\dim(V) = n \in \mathbb{N}$, und $f : V \to W$ linear. Dann gilt:

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

das heißt, es gilt: $\quad \mathrm{Df}(f) + \mathrm{Rg}(f) = n.$

Beispiel 6.20: Betrachten Sie die folgende Projektion $f : \mathbb{R}^3 \to \mathbb{R}^3$ auf die x-y-Ebene, das heißt, es gilt $f(x, y, z) = (x, y, 0)$. Dann ist dies eine lineare Abbildung. (Überlegen Sie sich dies!) Weiterhin sehen Sie, dass die z-Achse gerade den Kern dieser Abbildung ausmacht. Im Bild dagegen sind alle Vektoren der x-y-Ebene. Somit wird der dreidimensionale Urbildraum zum einen in den eindimensionalen Kern und zum anderen in das zweidimensionale Bild aufgespalten – dies entspricht dem Dimensionssatz.

Beweis von Satz 6.19: Es ist $\text{Kern}(f)$ ein Untervektorraum von V mit

$$\dim\big(\text{Kern}(f)\big) \leq \dim(V).$$

Setze $r := \dim\big(\text{Kern}(f)\big)$. Wähle eine Basis $(v_1, \ldots, v_r)$ von $\text{Kern}(f)$. Nach dem Basisergänzungssatz wähle $v_{r+1}, \ldots, v_n \in V$, so dass die Folge $(v_1, \ldots, v_r, v_{r+1}, \ldots, v_n)$ eine Basis von V ist. Dann gilt:

$$\begin{aligned}
\text{Bild}(f) &= \big\{f(v) \bigm| v \in V\big\} \\
&= \big\{f(\lambda_1 v_1 + \ldots + \lambda_r v_r + \lambda_{r+1} v_{r+1} + \ldots + \lambda_n v_n) \bigm| \lambda_1, \ldots, \lambda_n \in \mathbb{K}\big\} \\
&= \left\{\lambda_1 \underbrace{f(v_1)}_{=0} + \ldots + \lambda_r \underbrace{f(v_r)}_{=0} + \lambda_{r+1} f(v_{r+1}) + \ldots + \lambda_n f(v_n) \Bigm| \lambda_1, \ldots, \lambda_n \in \mathbb{K}\right\} \\
&= \big\{\lambda_{r+1} f(v_{r+1}) + \ldots + \lambda_n f(v_n) \bigm| \lambda_{r+1}, \ldots, \lambda_n \in \mathbb{K}\big\} \\
&= \mathcal{L}\big(f(v_{r+1}), \ldots, f(v_n)\big)
\end{aligned}$$

Offenbar erzeugen $f(v_{r+1}), \ldots, f(v_n)$ das Bild von f. Es bleibt zu zeigen, dass $f(v_{r+1}), \ldots, f(v_n)$ linear unabhängig sind.

Setze dazu $w_{r+1} := f(v_{r+1}), \ldots, w_n := f(v_n)$ und betrachte Skalare $\alpha_{r+1}, \ldots, \alpha_n \in \mathbb{K}$, so dass $\alpha_{r+1} w_{r+1} + \ldots + \alpha_n w_n = 0$. Ziel ist es also zu zeigen, dass alle $\alpha_{r+1} = \cdots = \alpha_n = 0$ sind.

$$\begin{aligned}
0 &= \alpha_{r+1} w_{r+1} + \ldots + \alpha_n w_n &&= \alpha_{r+1} f(v_{r+1}) + \ldots + \alpha_n f(v_n) \\
&= f(\alpha_{r+1} v_{r+1}) + \ldots + f(\alpha_n v_n) &&= f\Bigg(\underbrace{\alpha_{r+1} v_{r+1} + \ldots + \alpha_n v_n}_{=:v}\Bigg)
\end{aligned}$$

Damit liegt $v = \sum_{i=r+1}^{n} \alpha_i v_i$ im Kern von f, so dass wir eine Basisdarstellung durch $(v_1, \ldots, v_r)$ erhalten:

$$v = \alpha_1 v_1 + \ldots + \alpha_r v_r = \sum_{i=r+1}^{n} \alpha_i v_i = \alpha_{r+1} v_{r+1} + \ldots + \alpha_n v_n$$

Dies ist offenbar äquivalent zu der Gleichung:

$$0 = \alpha_1 v_1 + \ldots + \alpha_r v_r - \alpha_{r+1} v_{r+1} - \ldots - \alpha_n v_n = \sum_{i=1}^{r} \alpha_i v_i - \sum_{i=r+1}^{n} \alpha_i v_i$$

Da $(v_1, \ldots, v_r, v_{r+1}, \ldots, v_n)$ eine Basis von V ist, gilt:

$$\alpha_1 = \ldots = \alpha_r = \alpha_{r+1} = \ldots = \alpha_n = 0$$

Damit ist $(w_{r+1}, \ldots, w_n)$ eine Basis von $\text{Bild}(f)$ und es gilt schließlich wie gewünscht:

$$\text{Df}(f) + \text{Rg}(f) = r + (n - r) = n$$

$\boxtimes$(Satz 6.19)

Beispiel 6.21: Wir betrachten eine Abbildung f mit den Bedingungen $f(v_1) = b_1$ und $f(v_2) = b_2$ für $v_1, v_2, b_1, b_2 \in V$. Kann man f dann zu einem Endomorphismus $f : V \to V$ ergänzen?

Dies ist nur dann möglich wenn wir noch zusätzliche Bedingungen an die Vektoren v_1, v_2, b_1, b_2 stellen: Es kann im WorstCase sein, dass $v_1 = \lambda v_2$ gilt. Dann muss auch $b_1 = f(v_1) = f(\lambda v_2) = \lambda f(v_2) = \lambda b_2$ gelten. Andernfalls sind die Vektoren v_1, v_2 linear unabhängig (der Leser überlege sich den trivialen Fall $v_i = 0$ oder $b_i = 0$ selbst) und wegen der gewünschten Linearität von f müssen auch b_1, b_2 linear unabhängig sein.

Nun können wir aber mit dem Basisergänzungssatz zuerst einmal v_1, v_2 und dann b_1, b_2 zu einer Basis von V ergänzen. Damit sind $(v_1, \ldots, v_n)$ und $(b_1, \ldots, b_n)$ Basen von V und durch die Festsetzung $f(v_i) := b_i$ ist bereits eindeutig eine lineare Abbildung auf ganz V definiert.

⋆ ⋆ ⋆

Wir nutzen die bisherigen Zusammenhänge zwischen diesen einzelnen neuen Begriffen, um zwei wesentliche Aussagen zu formulieren:

Satz 6.22:

Für eine lineare Abbildung $f : V \to W$ gilt: f ist genau dann injektiv, wenn $\mathrm{Kern}(f) = \{0\}$.

Beweis: Sei $f : V \to W$ ein Monomorphismus. Angenommen, der Kern wäre nicht trivial – dann gäbe es einen Vektoren $v \in \mathrm{Kern}(f)$, so dass $v \neq 0$ gilt. Somit gilt aber auch $f(0) = 0 = f(v)$, ein Widerspruch zur Injektivität von f.

Sei nun umgekehrt $f : V \to W$ linear und ihr Kern sei trivial. Angenommen, f wäre nicht injektiv. Dann gäbe es zwei Vektoren $v_1, v_2 \in V$, $v_1 \neq v_2$ und $f(v_1) = f(v_2)$. Somit gilt in diesem Fall: $0 = f(v_1) - f(v_2) = f(v_1 - v_2)$, also liegt der Vektor $v_1 - v_2$ im Kern der Abbildung und ist selbst vom Nullvektor verschieden. Dies ist ein Widerspruch zur Voraussetzung. $\boxtimes$

Und schließlich beweisen wir mithilfe des Dimensionssatzes 6.19 den folgenden –aus meiner Sicht überraschenden– Zusammenhang:

Satz 6.23:

Für einen endlich-dimensionalen Vektorraum V und einen Endomorphismus $f : V \to V$ gilt: f ist genau dann injektiv, wenn f surjektiv ist.

Beweis: Sei der Endomorphismus $f : V \to V$ für einen endlich-dimensionalen Vektorraum V gegeben. Nun gilt nach Satz 6.22 und der Dimensionsformel 6.19:

$$\begin{array}{lllll} f \text{ ist injektiv} & \text{gdw.} & \mathrm{Kern}(f) = \{0\} & \text{gdw.} & \mathrm{Df}(f) = 0 \\ & \text{gdw.} & \mathrm{Rg}(f) = \dim(V) & \text{gdw.} & f \text{ ist surjektiv} \end{array}$$

Beachten Sie, die letzte Äquivalenz gilt, da $\mathrm{Bild}(f)$ einen Unterraum von V bildet. $\boxtimes$

Der Beweis zeigt, dass die Aussage auch für Abbildungen $f : V \to W$ gilt, sofern V und W zwei Vektorräume der gleichen endlichen Dimension sind. Da solche Räume V und W dann aber sofort isomorph sind, ist diese Aussage nicht wirklich allgemeiner und wir können uns auf Endomorphismen $f : V \to V$ beschränken.

Beispiel 6.24: Um den Umgang mit den neuen Begrifflichkeiten injektiv, surjektiv und bijektiv zu üben, schauen wir uns jetzt zwei interessante Endomorphismen auf dem Vektorraum aller Folgen $(x_1, x_2, x_3, \ldots)$ aus Elementen des Körpers $\mathbb{K}$ an:

$$f_1\left((x_1, x_2, x_3, \ldots)\right) := (0, x_1, x_2, x_3, \ldots)$$
$$f_2\left((x_1, x_2, x_3, \ldots)\right) := (x_2, x_3, \ldots)$$

Wir prüfen nun diese Abbildungen auf Injektivität und Surjektivität. Beachten Sie, dass der hier betrachtete Vektorraum nicht endlich-dimensional ist und daher der Satz 6.23 nicht greift.

Das Ergebnis ist dennoch nicht überraschend:

Die Abbildung f_1 ist injektiv, aber nicht surjektiv, da mit

$$f_1\left((x_1, x_2, x_3, \ldots)\right) = f_1\left((y_1, y_2, y_3, \ldots)\right)$$

auch sofort für die Bilder gilt:

$$(0, x_1, x_2, x_3, \ldots) = (0, y_1, y_2, y_3, \ldots)\,.$$

Und somit offenbar auch $(x_1, x_2, x_3, \ldots) = (y_1, y_2, y_3, \ldots)$. Surjektivität liegt allerdings nicht vor, denn es existiert keine Folge $(x_1, x_2, x_3, \ldots)$ mit

$$f_1\left((x_1, x_2, x_3, \ldots)\right) = (1, \star, \star, \ldots)\,.$$

Bei f_2 ist es genau andersherum. Man sieht leicht, dass f_2 surjektiv und nicht injektiv ist, da für jedes $(x_1, x_2, x_3, \ldots)$ gilt $f_2\left((\star, x_1, x_2, x_3, \ldots)\right) = (x_1, x_2, x_3, \ldots)$ aber aus $f_2\left((x_1, x_2, x_3, \ldots)\right) = f_2\left((y_1, y_2, y_3, \ldots)\right)$ folgt nur $(x_2, x_3, \ldots) = (y_2, y_3, \ldots)$ und nicht unbedingt $x_1 = y_1$.

⋆ ⋆ ⋆

Aufgabe 6.25: Kann es auf einem endlich dimensionalen Vektorraum V solche Endomorphismen wie f_1 und f_2 aus dem vorherigen Beispiel geben? Das bedeutete also, Sie suchten eine lineare Abbildungen $f : V \to V$ die injektiv und nicht surjektiv oder nicht injektiv und surjektiv sind. Beweisen oder widerlegen Sie ihre Vermutung!

Hinweis: Beachten Sie Satz 6.22.

Aufgabe 6.26: Sei $f : V \to W$ für V, W zwei $\mathbb{Q}-$Vektorräume. Man zeige:
Aus $f(x + y) = f(x) + f(y)$ für alle $x, y \in V$ folgt, dass f linear ist.

7. Die Welt der Matrizen

Matrizen, anfangs als endliche Folgen definiert, anschließend sofort als zweidimensionales Schema interpretiert, werden uns das (algebraische) Leben im Umgang mit linearen Abbildungen aber auch bei den Grundlagen in Form von Linearen Gleichungssystem deutlich erleichtern.

Darstellende Matrizen. Zunächst wollen wir Homomorphismen zwischen $\mathbb{K}^m$ und $\mathbb{K}^n$ schematisch als Matrix darstellen und werden zeigen, dass wir dadurch keinem Informationsverlust unterliegen.

Satz 7.1:

Sei $\mathbb{K}$ ein Körper und sei $f : \mathbb{K}^m \to \mathbb{K}^n$ eine lineare Abbildung mit $n, m \in \mathbb{N}$. Sei weiterhin $(e_1, \dots, e_m)$ die kanonische Basis des $\mathbb{K}^m$ und für $i = 1, \dots, m$ sei

$$f(e_i) = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix}$$

dabei schreiben wir die n-Tupel in $\mathbb{K}^n$ aus rechnerischen Gründen als Spaltenvektoren. Dann ist für Argumente $v = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in \mathbb{K}^m$ der Funktionswert

$$f(v) = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nm}x_m \end{pmatrix} \in \mathbb{K}^n$$

gegeben.

Beweis: Sei $v = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in \mathbb{K}^m$. Dann gilt offenbar:

$$v = x_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \ldots + x_m \cdot \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = \sum_{i=1}^{m} x_i e_i.$$

Wegen der Linearität von f gilt:

$$f(v) = f\left(\sum_{i=1}^{m} x_i e_i\right) = \sum_{i=1}^{m} x_i f(e_i)$$

Daraus folgt:

$$f(v) = x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} + \ldots + x_m \begin{pmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1m}x_m \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2m}x_m \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nm}x_m \end{pmatrix}$$

⊠

Der Satz 7.1 rechtfertigt die folgende

Definition 7.2 – Matrix:

Sei $\mathbb{K}$ ein Körper und $n, m \in \mathbb{N}$. Eine $(\mathsf{n} \times \mathsf{m})$-Matrix über $\mathbb{K}$ ist eine Folge $M = \big(a_{ij} \mid i = 1, \ldots, n\,;\ j = 1, \ldots, m\big)$ von Elementen $a_{ij} \in \mathbb{K}$. Wir schreiben die Matrix als zweidimensionales Schema wie folgt:

$$M = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{pmatrix}$$

Es sei $\mathrm{Mat}(n \times m, \mathbb{K})$ die Menge aller $(n \times m)$-Matrizen über $\mathbb{K}$.

Die Elemente von $\mathrm{Mat}(n \times m, \mathbb{K})$ werden abkürzend bezeichnet als

$$(a_{ij})_{i=1,\ldots,n\,;\,j=1,\ldots,m}$$

oder einfach als $(a_{ij})_{ij}$ oder auch als (a_{ij}). Die Klammer () bei (a_{ij}) steht also für die Folge und nicht etwa für einen einzelnen Eintrag. Hierbei nennen wir i und j die Laufindizes.

Für $i = 1, \ldots, n$ ist der **Zeilenvektor** (waagerechter Vektor) $\quad (a_{i1}, \ldots, a_{im})$ die i-te Zeile von M.

Für $j = 1, \ldots, m$ ist der **Spaltenvektor** (senkrecht notierter Vektor) $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}$ die j-te Spalte von M.

Ist die Matrix $M = \begin{pmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nm} \end{pmatrix}$ durch $f(e_i) = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix}$ wie in Satz 7.1 definiert, so nennen wir M die **darstellende Matrix** von f. Wir schreiben kurz

$$M = \mathrm{DM}(f).$$

Die darstellende Matrix von f besteht also aus den Spaltenvektoren $f(e_1), \ldots, f(e_m)$.

Bemerkung 7.3: (**Merkregel**) Die Spalten der darstellenden Matrix sind die Bilder der Basisvektoren.

Beispiel 7.4: Betrachte die Abbildung $f : \mathbb{R}^3 \to \mathbb{R}^2$, gegeben durch die Vorschrift

$$f(\begin{pmatrix} x \\ y \\ z \end{pmatrix}) = \begin{pmatrix} x + 2z \\ x - y + z \end{pmatrix}.$$

Für die darstellende Matrix berechnen Sie die Werte $f(e_1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $f(e_2) = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ sowie $f(e_3) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ und schreiben diese spaltenweise in eine Matrix:

$$\mathrm{DM}(f) = \begin{pmatrix} 1 & 0 & 2 \\ 1 & -1 & 1 \end{pmatrix}$$

Mithilfe der darstellenden Matrizen können wir schließlich die Menge der linearen Abbildungen und die Menge der Matrizen wunderbar miteinander identifizieren:

Satz 7.5:

Für jede $(n \times m)$-Matrix M über $\mathbb{K}$ gibt es eine eindeutig bestimmte lineare Abbildung $f : \mathbb{K}^m \to \mathbb{K}^n$, sodass M die darstellende Matrix von f ist. Damit ist die Abbildung

$$\mathrm{Hom}(\mathbb{K}^m, \mathbb{K}^n) \to \mathrm{Mat}(n \times m, \mathbb{K})$$
$$f \to \mathrm{DM}(f)$$

eine Bijektion.

Beweis: Bisher haben wir gesehen, dass wir einer Abbildung f die darstellende Matrix $\mathrm{DM}(f)$ eineindeutig zuordnen konnten. Damit ist die Abbildung
$\mathrm{Hom}(\mathbb{K}^m, \mathbb{K}^n) \to \mathrm{Mat}(n \times m, \mathbb{K})$ injektiv.

Ebenso kann man einer Matrix M eine Abbildung f zuordnen, sodass schließlich gilt: $M = \mathrm{DM}(f)$. Damit ist die Abbildung surjektiv. Dies werden wir später im letzten Abschnitt des Kapitels noch näher beleuchten. ⊠

⋆ ⋆ ⋆

Matrizenrechnung. Wir möchten mit Matrizen rechnen und führen daher zunächst einfache Operationen ein, die die Menge der Matrizen zum Vektorraum erheben. Ziel ist es, die Operationen so zu definieren, dass wir einen starken Zusammenhang zu dem Vektorraum der Homomorphismen zwischen dem $\mathbb{K}^m$ und $\mathbb{K}^n$ haben.

Definition 7.6 – Matrizenaddition, Matrizenmultiplikation, Nullmatrix:

Sei $\mathbb{K}$ ein Körper und sei $\mathrm{Mat}(n \times m, \mathbb{K})$ die Menge aller $(n \times m)$-Matrizen über $\mathbb{K}$.

(a) Definiere die Matrizenaddition durch:

$$\begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} + \begin{pmatrix} b_{11} & \dots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{n1} & \dots & b_{nm} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & \dots & a_{1m}+b_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1}+b_{n1} & \dots & a_{nm}+b_{nm} \end{pmatrix}$$

(b) Definiere die skalare Multiplikation durch:

$$\lambda \cdot \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \dots & \lambda a_{1m} \\ \vdots & \ddots & \vdots \\ \lambda a_{n1} & \dots & \lambda a_{nm} \end{pmatrix}$$

(c) Definiere die Nullmatrix $0 \in \mathrm{Mat}(n \times m, \mathbb{K})$ als:

$$0 = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{pmatrix}$$

Damit ist die Nullmatrix das neutrale Element der Matrizenaddition.

Damit haben wir eine skalare Multiplikation, eine Addition und ein neutrales Element definiert, so dass wir folgenden Satz beweisen können:

Satz 7.7:

Die Struktur $(\mathrm{Mat}(n \times m, \mathbb{K}), +, \cdot, 0)$ ist ein $\mathbb{K}$-Vektorraum.

Beweis: Offenbar ist die Zuordnung

$$\begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} \mapsto (a_{11}, \dots, a_{1m}, a_{21}, \dots, a_{2m}, \dots, a_{n1}, \dots, a_{nm})$$

ein Isomorphismus zwischen $(\mathrm{Mat}(n \times m, \mathbb{K}), +, \cdot, 0)$ und $\mathbb{K}^{n \cdot m}$: Wir wissen bereits, dass $\mathbb{K}^{nm}$ ein $\mathbb{K}$-Vektorraum ist. ⊠

An dieser Stelle können wir den Satz 7.5 verstärken und den bereits als Bijektion erkannten Zusammenhang zum Isomorphismus erweitern:

Satz 7.8 – Isomorphiesatz:

Sei $\mathbb{K}$ ein Körper und $n, m \in \mathbb{N}$. Dann ist die Abbildung

$$\begin{aligned} \mathrm{Hom}(\mathbb{K}^m, \mathbb{K}^n) &\to \mathrm{Mat}(n \times m, \mathbb{K}); \\ f &\to \mathrm{DM}(f) \end{aligned}$$

(sogar) ein Isomorphismus von Vektorräumen.

Beweis: Nach Satz 7.5 ist die Abbildung eine Bijektion. Es bleibt zu zeigen, dass diese Abbildung ein Homomorphismus ist. Wie wir wissen muss allgemein für die Linearität gelten: $h(x+y) = h(x)+h(y)$ und $\lambda \cdot h(x) = h(\lambda \cdot x)$. In unserem Fall entspricht das h einer Abbildung zwischen $\mathrm{Hom}(\mathbb{K}^n, \mathbb{K}^m)$ und $\mathrm{Mat}(n \times m, \mathbb{K})$, x und y sind Elemente aus dem Startraum, in unserem Fall also Abbildungen aus $\mathrm{Hom}(\mathbb{K}^n, \mathbb{K}^m)$. Nennen wir die Abbildung aus dem Satz Φ und Elemente aus $\mathrm{Hom}(\mathbb{K}^n, \mathbb{K}^m)$ mit f und g, so erhalten wir die Schreibweise:

$$\Phi : \mathrm{Hom}(\mathbb{K}^n \mathbb{K}^m) \to \mathrm{Mat}(n \times m, \mathbb{K}) \ ; \ \Phi(f) = \mathrm{DM}(f)$$

Die erste Eigenschaft der Linearität, die wir zeigen müssen heißt also:
$\Phi(f + g) = \Phi(f) + \Phi(g)$. Wir wissen, dass $\Phi(f) = \mathrm{DM}(f)$ gilt und somit müssen wir zeigen:

(a) $\mathrm{DM}(f + g) = \mathrm{DM}(f) + \mathrm{DM}(g)$
(b) $\mathrm{DM}(\lambda \cdot f) = \lambda \cdot \mathrm{DM}(f)$

Zu (a): Seien $f, g \in \mathrm{Hom}(\mathbb{K}^n, \mathbb{K}^m)$ mit darstellenden Matrizen $\mathrm{DM}(f)$ und $\mathrm{DM}(g)$. Die Spalten von $\mathrm{DM}(f)$ sind gerade die Vektoren $f(e_1), \ldots, f(e_m)$ und die Spalten von $\mathrm{DM}(g)$ sind die Vektoren $g(e_1), \ldots, g(e_m)$. Die Spalten von $\mathrm{DM}(f + g)$ sind die Vektoren

$$(f + g)(e_1) = f(e_1) + g(e_1), \ldots, (f + g)(e_m) = f(e_m) + g(e_m).$$

Nach Definition der Matrizenaddition ist damit

$$\mathrm{DM}(f + g) = \mathrm{DM}(f) + \mathrm{DM}(g).$$

Zu (b): Sei f ein Homomorphismus aus $\mathrm{Hom}(\mathbb{K}^m, \mathbb{K}^n)$ mit darstellender Matrix $\mathrm{DM}(f)$ und sein $\lambda \in \mathbb{K}$. Die Spalten von $\mathrm{DM}(\lambda \cdot f)$ sind die Vektoren

$$(\lambda \cdot f)(e_1) = \lambda \cdot f(e_1), \ldots, (\lambda \cdot f)(e_m) = \lambda \cdot f(e_m).$$

Nach Definition der Skalarmultiplikation von Matrizen gilt damit wie gewünscht $\mathrm{DM}(\lambda \cdot f) = \lambda \cdot \mathrm{DM}(f)$.

⊠

⋆ ⋆ ⋆

Wir werden jetzt eine Multiplikation von Matrizen einführen, sodass wir die Schreibweise des Beispiels 6.10 rechtfertigen:

$$F(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}) = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 + 3x_3 \\ 4x_1 + 5x_2 + 6x_3 \end{pmatrix}$$

Diese Matrizenmultiplikation entspricht in der Sprache von Abbildungen der Hintereinanderausführung von Abbildungen. Betrachten wir zwei Homomorphismen f, g wie folgt:

$$\mathbb{K}^m \xrightarrow{f} \mathbb{K}^n \xrightarrow{g} \mathbb{K}^r$$

mit darstellenden Matrizen $\mathrm{DM}(g) = A = (a_{ik}) \in \mathrm{Mat}(r \times n, \mathbb{K})$ und
$\mathrm{DM}(f) = B = (b_{kj}) \in \mathrm{Mat}(n \times m, \mathbb{K})$. Die darstellende Matrix von $g \circ f : \mathbb{K}^m \to \mathbb{K}^r$ besteht dann aus den Spaltenvektoren $(g \circ f)(e_j)$:

$$\begin{aligned}
(g \circ f)(e_j) &= g(f(e_j)) = g(\begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix}) &&= g(b_{1j}e_1 + \ldots + b_{nj}e_n) \\
&= b_{1j}g(e_1) + \ldots + b_{nj}g(e_n) && \text{wegen Linearität} \\
&= b_{1j} \begin{pmatrix} a_{11} \\ \vdots \\ a_{r1} \end{pmatrix} + \ldots + b_{nj} \begin{pmatrix} a_{1n} \\ \vdots \\ a_{rn} \end{pmatrix} &&= \begin{pmatrix} b_{1j}a_{11} + \ldots + b_{nj}a_{1n} \\ \vdots \\ b_{1j}a_{r1} + \ldots + b_{nj}a_{rn} \end{pmatrix} \\
&= \begin{pmatrix} a_{11}b_{1j} + \ldots + a_{1n}b_{nj} \\ \vdots \\ a_{r1}b_{1j} + \ldots + a_{rn}b_{nj} \end{pmatrix} && \text{Körperaxiome}
\end{aligned}$$

Der i-te Koeffizient dieses Spaltenvektors ist von der Form:

$$a_{i1}b_{1j} + \ldots + a_{in}b_{nj} = \sum_{k=1}^{n} a_{ik}b_{kj}$$

Beachten Sie, dass diese Summe wirklich nur ein einzelnes Element aus der gesuchten Matrix ist, nämlich zu finden in der i-ten Zeile und j-ten Spalte.

Insgesamt ist die darstellende Matrix von der Form:

$$\mathrm{DM}(g \circ f) = \begin{pmatrix}
\sum_{k=1}^{n} a_{1k}b_{k1} & \ldots & \ldots & \ldots & \sum_{k=1}^{n} a_{1k}b_{km} \\
\vdots & \ddots & & & \vdots \\
\vdots & & \sum_{k=1}^{n} a_{ik}b_{kj} & & \vdots \\
\vdots & & & \ddots & \vdots \\
\sum_{k=1}^{n} a_{rk}b_{k1} & \ldots & \ldots & \ldots & \sum_{k=1}^{n} a_{rk}b_{km}
\end{pmatrix}$$

Damit ist die folgende Definition motiviert:

Definition 7.9 – Matrizenprodukt:

Definiere das **Matrizenprodukt** für zwei Matrizen

$$A = (a_{ik}) \in \mathrm{Mat}(r \times n, \mathbb{K}) \text{ und } B = (b_{kj}) \in \mathrm{Mat}(n \times m, \mathbb{K}) \text{ als}$$

$$A \cdot B = AB = \left(\sum_{k=1}^{n} a_{ik}b_{kj} \right)_{\substack{i=1,\ldots,r \\ j=1,\ldots,m}} \in \mathrm{Mat}(r \times m, \mathbb{K})$$

Beispiel 7.10: Allgemein gilt für Matrizen $A, B \in \mathrm{Mat}(2 \times 2, \mathbb{K})$:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} a \cdot e + b \cdot g & a \cdot f + b \cdot h \\ c \cdot e + d \cdot g & c \cdot f + d \cdot h \end{pmatrix} \tag{3}$$

Betrachte also zwei Matrizen A, B wie folgt:

$$\begin{aligned} A \cdot B = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\ \neq \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} = B \cdot A \end{aligned}$$

Bemerkung 7.11: Im Allgemeinen ist das Matrizenprodukt nicht kommutativ.

Folgendes Schema, welches wir am Beispiel vorführen, kann helfen, die Matrizenmultiplikation übersichtlicher und verständlicher darzustellen:

Beispiel 7.12: Sei $A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ und $B = \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{pmatrix}$. Berechne nun das Matrizenprodukt $A \cdot B$ wie folgt:

$$\begin{array}{cc} & \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{pmatrix} , \\ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} & \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \end{array}$$

wobei

$$\begin{aligned} c_{11} &:= 1 \cdot 1 + 2 \cdot 2 + 3 \cdot 3 = 14, & c_{12} &:= 1 \cdot 2 + 2 \cdot 3 + 3 \cdot 4 = 20, \\ c_{21} &:= 3 \cdot 1 + 2 \cdot 2 + 1 \cdot 3 = 10, & c_{22} &:= 3 \cdot 2 + 2 \cdot 3 + 1 \cdot 4 = 16 \end{aligned}$$

sind.

Also ist:

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 14 & 20 \\ 10 & 16 \end{pmatrix}.$$

Übrigens, analog lässt sich $B \cdot A$ berechnen:

$$\begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 7 & 6 & 5 \\ 11 & 10 & 9 \\ 15 & 14 & 13 \end{pmatrix}$$

und wir erkennen, dass das Produkt eine (3×3)-Matrix ist, wobei $A \cdot B$ dagegen eine (2×2)-Matrix war.

Satz 7.13:

Sei $f : \mathbb{K}^m \to \mathbb{K}^n$ eine lineare Abbildung und $A = \mathrm{DM}(f)$, dann gilt für alle Spaltenvektoren $x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}$ stets: $f(x) = A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = A \cdot x.$

Beweis: Wegen (3) aus Beispiel 7.10 und Satz 7.1 ist die Behauptung klar. ⊠

Satz 7.14:

Seien $\mathbb{K}$ ein Körper und A, B, C Matrizen über $\mathbb{K}$ mit geeigneten Dimensionen. Dann gelten folgende Rechengesetze:

(a) Assoziativität: $A \cdot (B \cdot C) = (A \cdot B) \cdot C$
(b) Distributivgesetz I : $A \cdot (B + C) = A \cdot B + A \cdot C$
(c) Distributivgesetz II : $(A + B) \cdot C = A \cdot C + B \cdot C$

Bemerkung 7.15: Wir brauchen beide Distributivgesetze, da die Matrizenmultiplikation im Allgemeinen nicht kommutativ ist. Außerdem benutzen wir auch hier die allgemeinen mathematischen Vereinbarungen wie „Punkt- vor Strichrechnung".

Beweis: Wir nehmen nach Satz 7.8 jeweils an, dass A, B, C die darstellenden Matrizen von linearen Abbildungen f, g, h sind.

(a) Dann ist $B \cdot C$ die darstellende Matrix von $g \circ h$ und $A \cdot (B \cdot C)$ die darstellende Matrix von $f \circ (g \circ h)$. Analog ist $(A \cdot B) \cdot C$ die darstellende Matrix von $(f \circ g) \circ h$ und wegen der Gleichheit $f \circ (g \circ h) = (f \circ g) \circ h$ gilt wie gewünscht $A \cdot (B \cdot C) = (A \cdot B) \cdot C$.

(b) Es ist $B + C$ die darstellende Matrix von $g + h$ und $A \cdot (B + C)$ die darstellende Matrix von $f \circ (g + h)$. Analog ist $A \cdot B + A \cdot C$ die darstellende Matrix von $f \circ g + f \circ h$. Nun gilt für alle x:

$$\begin{aligned} \big(f \circ (g + h)\big)(x) &= f\big((g + h)(x)\big) &&= f\big(g(x) + h(x)\big) \\ &= f\big(g(x)\big) + f\big(h(x)\big) &&= (f \circ g)(x) + (f \circ h)(x) \\ &= (f \circ g + f \circ h)(x) \end{aligned}$$

Damit gilt wie gewünscht: $f \circ (g + h) = f \circ g + f \circ h$

(c) Folgt analog zu (b).

⊠

Bemerkung 7.16: Die Gesetze aus Satz 7.14 können auch durch (mühseliges) Nachrechnen der formalen Definition der Matrixmultiplikation bewiesen werden.

Definition 7.17 – quadratische Matrix, Einheitsmatrix:

Sei $\mathbb{K}$ ein Körper.

(a) Eine Matrix A über $\mathbb{K}$ heißt **quadratisch**, wenn es eine natürliche Zahl n gibt, so dass $A \in \mathrm{Mat}(n \times n, \mathbb{K})$.

(b) Die n-dimensionale **Einheitsmatrix** ist die darstellende Matrix der identischen Abbildung $\mathrm{id}_{\mathbb{K}^n}$:

$$E_n = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Satz 7.18:

Sei $\mathbb{K}$ ein Körper und $A \in \operatorname{Mat}(n \times m, \mathbb{K})$. Dann gilt:

$$A \cdot E_m = A \quad \text{und} \quad E_n \cdot A = A.$$

Beweis: Klar, sowohl über die darstellenden Matrizen, als auch mittels der formalen Definition. ☒

Definition 7.19 – invertierbar (Matrix):

Sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$. Eine $(n \times n)$-Matrix A über $\mathbb{K}$ heißt **invertierbar**, wenn es ein $B \in \operatorname{Mat}(n \times n, \mathbb{K})$ gibt mit $A \cdot B = E_n$.
Wenn es genau ein solches B gibt, so heißt B die **Inverse** zu A und wir schreiben $B = A^{-1}$.

Satz 7.20:

Sei $\mathbb{K}$ ein Körper und $n \in \mathbb{N}$, dann gilt:

(a) Sei $A \in \operatorname{Mat}(n \times n, \mathbb{K})$ invertierbar. Dann gibt es genau ein B mit $A \cdot B = E_n$.
(b) Sei $A \in \operatorname{Mat}(n \times n, \mathbb{K})$ invertierbar. Dann gilt $A \cdot A^{-1} = A^{-1} \cdot A = E_n$.
(c) Sei $A \in \operatorname{Mat}(n \times n, \mathbb{K})$ invertierbar. Dann ist A^{-1} ebenfalls invertierbar und es gilt $(A^{-1})^{-1} = A$.
(d) Seien $A, B \in \operatorname{Mat}(n \times n, \mathbb{K})$ invertierbar. Dann ist $A \cdot B$ ebenfalls invertierbar und es gilt $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$.

Beweis: Wir identifizieren Homomorphismen von $\mathbb{K}^n$ nach $\mathbb{K}^n$ wieder nach Satz 7.8 mit ihren darstellenden Matrizen. Angenommen, A ist invertierbar und es gilt
$A \cdot B = E_n$ (Existenz von B ist nach Definition 7.19 gesichert). Dann gilt:

$$\begin{aligned}\operatorname{Rg}(A) &= \dim(\operatorname{Bild}(A)) \\ &\geq \dim(\operatorname{Bild}(A \cdot B)) = \dim(\operatorname{Bild}(E_n)) \\ &= \dim(\mathbb{K}^n) = n\end{aligned}$$

Wegen $\operatorname{Rg}(A) \leq n$ gilt daher $\operatorname{Rg}(A) = n$. Damit ist die Abbildung A surjektiv (da nun $\operatorname{Bild}(A) \subseteq \mathbb{K}^n$, n-dimensional, also $\operatorname{Bild}(A) = \mathbb{K}^n$). Nach der Dimensionsformel gilt:

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

also

$$\dim(\operatorname{Kern}(A)) = 0$$

und somit gilt

$$\operatorname{Kern}(A) = \{0\}.$$

Nach Satz 6.22 ist A damit injektiv.

Insbesondere ist A also bijektiv, also ein Isomorphismus. Dieser besitzt eine Inverse Abbildung A^{-1} (im Sinne der Abbildungsinversen). Es sei C die darstellende Matrix von A^{-1}, das heißt $C = \operatorname{DM}(A^{-1})$. Es gilt $A \cdot C = C \cdot A = E_n$. Ziel ist es zu zeigen: $C = A^{-1}$.

Für alle Vektoren $x \in \mathbb{K}^n$ ist $(A \cdot C)(x) = E_n(x) = x = (A \cdot B)(x)$. Wegen der Injektivität von A gilt schließlich $C(x) = B(x)$, also $C = B$, und somit ist B die darstellende Matrix der zu A inversen Abbildung. Damit folgt Teil (a).

Die restlichen Eigenschaften folgen leicht, etwa wieder durch den Übergang von Matrizen zu Abbildungen mittels des Isomorphismus, gegeben durch Satz 7.8. ⊠

⋆ ⋆ ⋆

Lineare Abbildungen und Matrizen. Wir hatten Matrizen als darstellende Matrizen linearer Abbildungen motiviert und eingeführt, das heißt für eine lineare Abbildung $f : V \to V$ ist $\mathrm{DM}(f)$ eine Matrix. Eine lineare Abbildung $f : \mathbb{K}^m \to \mathbb{K}^n$ besitzt eine darstellende Matrix $\mathrm{DM}(f) \in \mathrm{Mat}(n \times m, \mathbb{K})$ und es gilt:

$$f(x) = \mathrm{DM}(f) \cdot x,$$

wobei x als Spaltenvektor $x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in \mathbb{K}^m$ aufgefasst wird.

Umgekehrt definiert jede Matrix $A \in \mathrm{Mat}(n \times m, \mathbb{K})$ offenbar die Abbildung $\mathbb{K}^m \to \mathbb{K}^n;\quad x \mapsto A \cdot x$. Diese Abbildung ist linear, denn es gilt

$$A \cdot (\lambda \cdot x) = \lambda(A \cdot x)$$

(andere Schreibweise dafür $A(\lambda x) = \lambda A(x)$) und

$$A(x + y) = A(x) + A(y).$$

Ihre darstellende Matrix ist gerade die Matrix $A = \mathrm{DM}(A)$, daher wird oftmals die so definierte Abbildung mit demselben Buchstaben A bezeichnet:

$$A : \mathbb{K}^m \to \mathbb{K}^n$$

Damit lassen sich von Abbildungen bekannte Begriffe auf A übertragen:

- $A(x) := A \cdot x$
- $A[X] := \{A \cdot x \mid x \in X\}$
- $A^{-1}[Y] := \{x \in \mathbb{K}^m \mid A \cdot x \in Y\}$
- $\mathrm{Kern}(A) := A^{-1}[\{0\}]$

Damit können wir Matrizen als lineare Abbildungen und lineare Abbildungen mittels ihrer darstellenden Matrizen auch als Matrizen auffassen, so dass sich die jeweils bekannten Begriffe auch leicht übertragen.

Beispiel 7.21: Wir werden jetzt mit Endomorphismen und ihren darstellenden Matrizen experimentieren. Wir haben schon festgestellt, dass die darstellende Matrix einer linearen Abbildung f von der Basis von V abhängt. Das bedeutet, dass je nach Wahl der Basis von V, sich im Allgemeinen auch die darstellende Matrix von f unterscheiden wird. Uns interessiert deswegen eine möglichst gutartige Basis, so dass die darstellende Matrix von f eine eher einfache (zum Rechnen handliche) Form hat.

Sei also f ein Endomorphismus auf dem n dimensionalen Vektorraum V über einem Körper $\mathbb{K}$ mit $1 + 1 \neq 0$ und es soll $f^2 = id$ gelten.

Wir werden eine Basis $v_1, \dots, v_n$ angeben, so dass f bezüglich dieser Basis die darstellende Matrix

$$\mathrm{DM}(f) = \begin{pmatrix} E_r & 0 \\ 0 & -E_k \end{pmatrix} \qquad \text{mit } k + r = n$$

besitzt.

Zuerst betrachten wir eine beliebige Basis $v_1, \dots, v_n$ von V und stellen fest, dass für jedes v_i folgendes gilt:

$$v_i = \frac{2 \cdot v_i + f(v_i) - f(v_i)}{2} = \frac{v_i + f(v_i)}{2} + \frac{v_i - f(v_i)}{2} =: a_i + b_i$$

Interessanterweise ist nun :

$$f(a_i) = f\left(\frac{v_i + f(v_i)}{2}\right) = \frac{f(v_i) + f^2(v_i)}{2} = \frac{v_i + f(v_i)}{2} = a_i \quad \text{ und}$$
$$f(b_i) = f\left(\frac{v_i - f(v_i)}{2}\right) = \frac{f(v_i) - f^2(v_i)}{2} = -\frac{v_i - f(v_i)}{2} = -b_i$$

Sei o.B.d.A. die Folge der Vektoren $a_1, \dots, a_r$ eine Basis des Untervektorraums:

$$U_a := \left\{ x \;\middle|\; x = \lambda_1 \cdot a_1 + \ldots + \lambda_n \cdot a_n, \quad \lambda_1, \dots, \lambda_n \in \mathbb{K} \right\}$$

und analog die Folge der Vektoren $b_1, \dots, b_k$ eine Basis des Untervektorraums:

$$U_b := \left\{ x \;\middle|\; x = \lambda_1 \cdot b_1 + \ldots + \lambda_n \cdot b_n, \quad \lambda_1, \dots, \lambda_n \in \mathbb{K} \right\}.$$

An dieser Stelle ist es wichtig sich klar zu machen, dass U_a und U_b nur den Nullvektor als Schnittmenge haben. Angenommen $v \in U_a \cap U_b$, dann wäre $f(v) = v$ und $f(v) = -v$ also $v = 0$.

Jetzt bleibt nur noch zu zeigen, dass $a_1, \dots, a_r, b_1, \dots, b_k$ eine Basis von V ist:

Jeder Vektor $v \in V$ besitzt eine Darstellung bezüglich der Basis $v_1, \dots, v_n$ und da wir diese Basis für die Konstruktion der a_i und b_i einfach nur zerlegt haben, besitzt er auch eine Darstellung bezüglich unserer $a_1, \dots, a_r, b_1, \dots, b_k$.

Es bleibt also noch die lineare Unabhängigkeit unserer Vektoren zu zeigen:
Sei

$$\lambda_1 \cdot a_1 + \ldots + \lambda_r \cdot a_r + \lambda_{r+1} \cdot b_1 + \ldots + \lambda_{r+k} \cdot b_k = 0$$

Angenommen, $\lambda_1 \neq 0$ (die anderen Fälle folgen genauso), dann können wir umformen zu:

$$\underbrace{-\lambda_1 \cdot a_1 - \ldots - \lambda_r \cdot a_r}_{=:a \in U_a} = \underbrace{\lambda_{r+1} \cdot b_1 + \ldots + \lambda_{r+k} \cdot b_k}_{=:b \in U_b}$$

Daraus folgt, dass das so erhaltene Element im Schnitt liegt und für die rechte Seite der Gleichung muss $\lambda_{r+1} = \cdots = \lambda_{r+k} = 0$ gelten (da nur der Nullvektor im Schnitt liegt). Hierdurch erhalten wir auf der linken Seite eine Linearkombination der a_i's, die Null ergibt, ohne dass alle Koeffizienten Null sind! Das ist aber ein Widerspruch zur linearen Unabhängigkeit der a_i's.

Schließlich müssen Sie sich nur noch klar werden, wie die darstellende Matrix von f bezüglich unserer neuen Basisvektoren aussieht:

$$\mathrm{DM}(f) = \begin{pmatrix} E_r & 0 \\ 0 & -E_k \end{pmatrix}$$

Dies liegt daran, dass in den ersten r Spalten das Bild der a_i steht, also $f(a_i) = 0 + \ldots + 0 + 1 \cdot a_i + 0 + \ldots + 0$ (deswegen E_r) und für die restlichen Spalten das Bild der b_i mit $f(b_i) = 0 + \ldots + 0 - 1 \cdot b_i + 0 + \ldots + 0$ (deswegen $-E_k$)

⋆ ⋆ ⋆

Aufgabe 7.22: Sei $f : V \to V$ ein Endomorphismus auf dem n dimensionalen Vektorraum V über einem Körper $\mathbb{K}$ mit $1 + 1 \neq 0$ und $f^2 = f$. Man zeige die Existenz einer Basis von V, so dass

$$\mathrm{DM}(f) = \begin{pmatrix} E_k & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{mit } k \in \mathbb{N}, \text{ bzgl. dieser Basis}$$

Tipp: Überlegen Sie sich, wie die darstellende Matrix bezüglich den Vektoren $f(b_1), \ldots, f(b_l)$, wobei die b_i's hier eine Basis von Bild(f) sind, aussieht.

8. Anwendungen von Matrizen

In diesem Kapitel beschäftigen wir uns zunächst kurz mit einer Anwendung von Matrizen in der Informatik, den linearen Codes. Anschließend zeige ich Ihnen eine weitere klassische Anwendung von Matrizen in Form von Drehungen und Spiegelungen in der reellen Ebene und schließlich wenden wir im letzten Teil die Theorie der Matrizen auf die einzelnen elementaren Zeilenumformungen an, um diese als Matrizenmultiplikation darstellen zu können. Dies werden wir im nächsten Kapitel benutzen, um dann auf den Gaußschen Algorithmus zurückzukommen.

Matrizen als Kodierwerkzeug: Lineare Codes. Man kann Unterräume von endlich-dimensionalen Vektorräumen zur Kodierung binärer Informationen verwenden.

Wir sagen: Ein linearer (n,k)-Code C ist ein k-dimesionaler Unterraum von $(\mathbb{Z}_2)^n$. Jedes Element von C ist eine binäre Folge der Länge n. Die 2^k Elemente von C sind in den 2^n Elementen von $(\mathbb{Z}_2)^n$ enthalten. Diese Unterräume entsprechen isomorphen Kopien von $(\mathbb{Z}_2)^k$.

Die Art der Lage des Unterraumes in $(\mathbb{Z}_2)^n$ kann zur Fehlererkennung und teilweise auch zur Korrektur benutzt werden. So ist etwa $\big\{(0,0,0),(1,1,1)\big\}$ ein linearer $(3,1)$-Code; $(0,0,0)$ kodiert die 0 und $(1,1,1)$ die 1. Wenn bei der Übertragung von Tripeln höchstens 1-Bit-Fehler erwartet werden, so ist eine Fehlerkorrektur möglich, indem die Tripel $(0,0,0),(1,0,0),(0,1,0),(0,0,1)$ als 0 und schließlich $(1,1,1),(0,1,1),(1,0,1),(1,1,0)$ als 1 dekodiert werden.

Betrachten wir nun ein anderes Beispiel, den so genannten $(7,4)$-Code. Eine mögliche Darstellung der Kodierung ist etwa durch die folgende Matrix gegeben:

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

Das Wort $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \in (\mathbb{Z}_2)^4$ wird durch den Vektor $A \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ kodiert.

Wir können auch eine Prüfmatrix erstellen, die so gewählt ist, dass sie auf Codewörter angewendet gleich 0 ist. Insbesondere gilt, dass bei fehlerhaften Übertragungen das Produkt mit der Prüfmatrix ungleich 0 ist und sogar bei 1-Bit-Fehlern die Stelle des Fehlers angibt:

So erhalten wir beispielsweise für den $(7,4)$-Code die folgende Prüfmatrix:

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

Diese Matrix können wir wie folgt anwenden: Wenn c ein Codewort ist, welches durch einen Einfachfehler (1-Bit-Fehler) e in das Wort $x := c + e$ verfälscht wurde, dann gilt:

$$H \cdot x = H \cdot (c+e) = \underbrace{H \cdot c}_{=0} + H \cdot e = H \cdot e.$$

Betrachten wir etwa das obige Beispiel, indem wir nun in die korrekte Lösung $c := \begin{pmatrix}1\\0\\1\\1\\0\\1\\0\end{pmatrix}$ zwei Fehler $e_1 := \begin{pmatrix}0\\1\\0\\0\\0\\0\\0\end{pmatrix}$ bzw. $e_2 := \begin{pmatrix}0\\0\\0\\1\\0\\0\\0\end{pmatrix}$ einarbeiten und so $c + e_1 = \begin{pmatrix}1\\1\\1\\1\\0\\1\\0\end{pmatrix}$ bzw. $c + e_2 = \begin{pmatrix}1\\0\\1\\0\\0\\1\\0\end{pmatrix}$ erhalten.

Mittels der Prüfmatrix erhalten wir schließlich die beiden Vektoren

$$H \cdot (c + e_1) = \begin{pmatrix}2\\3\\2\end{pmatrix} \quad \text{bzw.} \quad H \cdot (c + e_2) = \begin{pmatrix}2\\2\\1\end{pmatrix}$$

Da die Einträge modulo 2 gesehen werden müssen, entsprechen die beiden Vektoren gerade den folgenden beiden:

$$\begin{pmatrix}0\\1\\0\end{pmatrix} \quad \text{bzw.} \quad \begin{pmatrix}0\\0\\1\end{pmatrix}$$

Jetzt können wir ablesen, dass der erste Fehler an der Stelle

$$0 \cdot 2^0 + 1 \cdot 2^1 + 0 \cdot 2^2 = 2$$

bzw. der zweite Fehler an der Position $0 \cdot 2^0 + 0 \cdot 2^1 + 1 \cdot 2^2 = 4$ liegt. Und dies entspricht genau der Position der 1 in den Fehlervektoren e_1 bzw. e_2.

Mehr zu den linearen Codes lernen Sie in Ihren Informatik-Vorlesungen.

⋆ ⋆ ⋆

Wir kommen nun zu zwei speziellen Matrizen vom Typ (2×2), den Drehungen und Spiegelungen.

Matrizen als Drehungen in der reellen Ebene. Betrachten Sie einen beliebigen Vektor $a = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ und drehen diesen um den Winkel γ. Das Ergebnis sei der Vektor $b = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$. Offenbar gilt $\alpha + \gamma = \beta$.

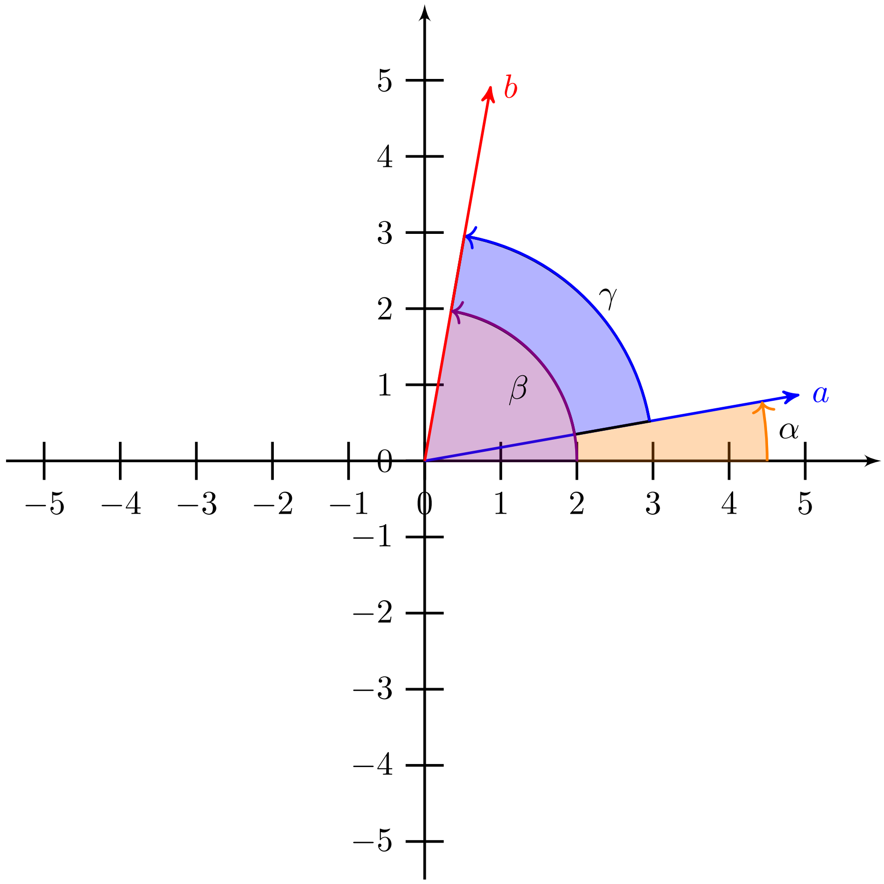


Abbildung 11: Matrizen als Drehungen in der reellen Ebene

Entsprechend der Sinus- und Kosinussätze am rechtwinkligen Dreieck gilt:
$\sin(\alpha) = \frac{y_1}{|a|}$, $\cos(\alpha) = \frac{x_1}{|a|}$, $\sin(\beta) = \frac{y_2}{|b|}$, $\cos(\beta) = \frac{x_2}{|b|}$, wobei mit $|a|$ und $|b|$ die Länge der Vektoren a, b gemeint ist. Damit erhalten wir (mittels trigonometrischer Zusammenhänge):

$$\begin{aligned}
x_2 &= \cos(\beta)\cdot|b| = \cos(\alpha+\gamma)\cdot|b| \\
&= \cos(\alpha)\cdot\cos(\gamma)\cdot|b| - \sin(\alpha)\cdot\sin(\gamma)\cdot|b| \\
&= \cos(\gamma)\cdot(\cos(\alpha)\cdot|a|) - \sin(\gamma)\cdot(\sin(\alpha)\cdot|a|) && (\text{es gilt } |a| = |b|) \\
&= \cos(\gamma)\cdot x_1 - \sin(\gamma)\cdot y_1
\end{aligned}$$

Darüber hinaus gilt:

$$\begin{aligned}
y_2 &= \sin(\beta)\cdot|b| = \sin(\alpha+\gamma)\cdot|b| \\
&= \sin(\alpha)\cdot\cos(\gamma)\cdot|b| + \cos(\alpha)\cdot\sin(\gamma)\cdot|b| \\
&= \cos(\gamma)\cdot\big(\sin(\alpha)\cdot|a|\big) + \sin(\gamma)\cdot\big(\cos(\alpha)\cdot|a|\big) && \big(\text{es gilt } |a| = |b|\big) \\
&= \cos(\gamma)\cdot y_1 + \sin(\gamma)\cdot x_1
\end{aligned}$$

Mittels Matrixschreibweise ergibt sich:

$$\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} \cos(\gamma) & -\sin(\gamma) \\ \sin(\gamma) & \cos(\gamma) \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$$

das heißt der Vektor a geht bei Anwendung der oben betrachteten linearen Abbildung

$$D_\gamma := \begin{pmatrix} \cos(\gamma) & -\sin(\gamma) \\ \sin(\gamma) & \cos(\gamma) \end{pmatrix}$$

auf den Vektor b über. Diese Matrix D_γ nennt man auch die **Drehmatrix** zum Winkel γ.

Beispiel 8.1: Für $\gamma = \frac{\pi}{2}$ ergibt sich $D_{90^\circ} = D_{\frac{\pi}{2}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

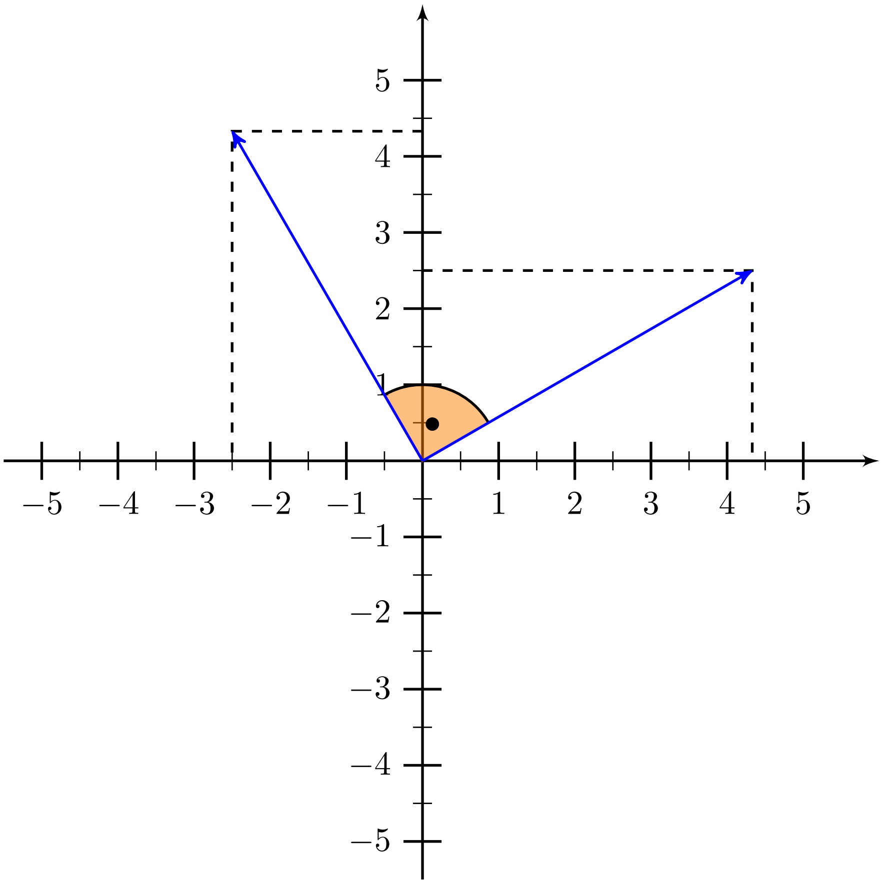


Abbildung 12

Für die Drehmatrizen können wir die von Drehungen bekannten Eigenschaften nachrechnen:

Satz 8.2:

Für Winkel $\varphi, \psi \in \mathbb{R}$ gilt:

$$D_{\varphi+\psi} = D_\varphi \cdot D_\psi$$

Beweis: Der Beweis ist klar. ⊠

Aus dem letzten Satz ergeben sich sofort einfache trigonometrische Zusammenhänge:

$$\begin{aligned}
D_{\varphi+\psi} &= \begin{pmatrix} \cos(\varphi+\psi) & -\sin(\varphi+\psi) \\ \sin(\varphi+\psi) & \cos(\varphi+\psi) \end{pmatrix} \\
&= D_\varphi \cdot D_\psi = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \cdot \begin{pmatrix} \cos(\psi) & -\sin(\psi) \\ \sin(\psi) & \cos(\psi) \end{pmatrix} \\
&= \begin{pmatrix} \cos(\varphi)\cos(\psi) - \sin(\varphi)\sin(\psi) & -\sin(\psi)\cos(\varphi) - \sin(\varphi)\cos(\psi) \\ \sin(\varphi)\cos(\psi) + \sin(\psi)\cos(\varphi) & -\sin(\varphi)\sin(\psi) + \cos(\varphi)\cos(\psi) \end{pmatrix} \\
&= \begin{pmatrix} \cos(\varphi)\cos(\psi) - \sin(\varphi)\sin(\psi) & -(\cos(\varphi)\sin(\psi) + \sin(\varphi)\cos(\psi)) \\ \sin(\varphi)\cos(\psi) + \cos(\varphi)\sin(\psi) & \cos(\varphi)\cos(\psi) - \sin(\varphi)\sin(\psi) \end{pmatrix}
\end{aligned}$$

Außerdem gilt:

$$\begin{aligned}
&\begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} \cdot \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix} \\
&= \begin{pmatrix} \cos^2(\varphi) + \sin^2(\varphi) & \sin(\varphi)\cos(\varphi) - \sin(\varphi)\cos(\varphi) \\ \sin(\varphi)\cos(\varphi) - \sin(\varphi)\cos(\varphi) & \sin^2(\varphi) + \cos^2(\varphi) \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2
\end{aligned}$$

Damit ist $\begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix}$ die Inverse von D_φ.

Aus diesen algebraischen Überlegungen erhalten wir die folgenden (analytischen) Aussagen:

Satz 8.3:

Es gilt:

(a) $\cos(\varphi \pm \psi) = \cos(\varphi)\cos(\psi) \mp \sin(\varphi)\sin(\psi)$
(b) $\sin(\varphi \pm \psi) = \sin(\varphi)\cos(\psi) \pm \cos(\varphi)\sin(\psi)$
(c) $\cos(-\varphi) = \cos(\varphi)$
(d) $\sin(-\varphi) = -\sin(\varphi)$

Beweis: Die ersten beiden Punkte haben wir bereits gesehen.

Für die Aussagen (c) und (d) beachten Sie, dass

$$\begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix} = D_\varphi^{-1}$$

invers zur Matrix

$$\begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} = D_\varphi$$

ist.

Auf der anderen Seite ist aber die Drehung um den Winkel $-\varphi$ ebenfalls invers zu D_φ, sodass gilt:

$$D_{-\varphi} = \begin{pmatrix} \cos(-\varphi) & -\sin(-\varphi) \\ \sin(-\varphi) & \cos(-\varphi) \end{pmatrix} = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix}.$$

⊠

Beachten Sie, dass wir hier in diesem Unterkapitel scheinbar einen Kreisschluss geschaffen haben: Wir haben mittels der trigonometrischen Additionstheoreme die Drehmatrix hergeleitet, um dann mit ihr die Additionstheoreme zu beweisen. Das ist natürlich keine logisch korrekte Schlussfolgerung. Das Dilemma löst sich allerdings rasch auf, indem wir die Drehmatrix anders herleiten: Wir hätten genauso gut die Drehmatrix angeben können und dann aufgrund der Eindeutigkeit einer linearen Abbildung auf einer Basis leicht überpüfen können, dass es sich wirklich um eine Drehung handelt. Ich habe dennoch diesen Weg hier gewählt, damit Sie schrittweise mit mir gemeinsam die Drehung herleiten – den anderen Weg werden wir gleich im Unterkapitel über Spiegelungen gehen.

⋆ ⋆ ⋆

Matrizen als Spiegelungen in der reellen Ebene. Betrachte die Matrix

$$S_{\frac{\gamma}{2}} := \begin{pmatrix} \cos(\gamma) & \sin(\gamma) \\ \sin(\gamma) & -\cos(\gamma) \end{pmatrix}.$$

Dies ist offenbar eine lineare Abbildung. Um zu verstehen, was $S_{\frac{\gamma}{2}}$ macht, nutzen wir unser algebraisches Verständnis über lineare Abbildungen und schauen uns die Bilder der kanonischen Basis e_1 und e_2 an:

$$S_{\frac{\gamma}{2}}(e_1) = \begin{pmatrix} \cos(\gamma) & \sin(\gamma) \\ \sin(\gamma) & -\cos(\gamma) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos(\gamma) \\ \sin(\gamma) \end{pmatrix} \text{ und}$$
$$S_{\frac{\gamma}{2}}(e_2) = \begin{pmatrix} \cos(\gamma) & \sin(\gamma) \\ \sin(\gamma) & -\cos(\gamma) \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \sin(\gamma) \\ -\cos(\gamma) \end{pmatrix}$$

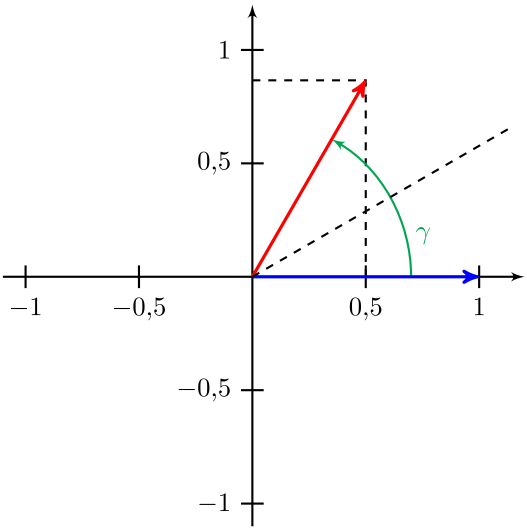


Abbildung 13

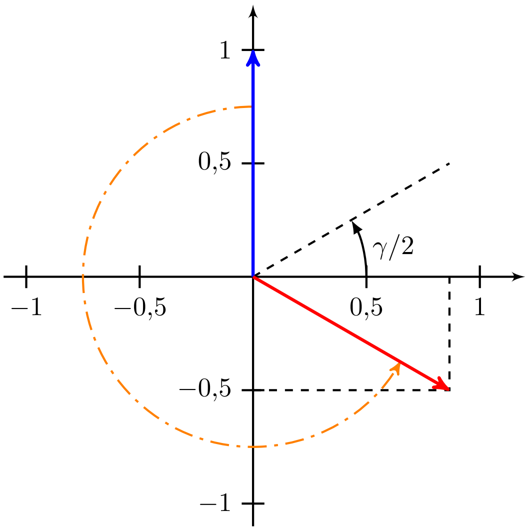


Abbildung 14

Die Bildkoordinaten des ersten Basisvektors e_1 lassen sich rasch ablesen – anders sieht es für das Bild von e_2 aus.

Diese Koordinaten können wir aber auch schrittweise nachvollziehen: Betrachten Sie das rechtwinklige Dreieck, bestehend aus dem Bildvektor des zweiten Einheitsvektors und der x-Achse. Der Winkel zwischen Bildvektor und x-Achse beträgt $\frac{\pi}{2} - \gamma$, denn der Winkel zwischen Spiegelachse und Ausgangsvektor beträgt gerade $\frac{\pi}{2} - \frac{\gamma}{2}$ und daher entspricht dies auch dem Winkel zwischen Spiegelachse und Bildvektor. Die Spiegelachse schließt aber mit der x-Achse einen Winkel von $\frac{\gamma}{2}$ ein, so dass wir den gesuchten Winkel im Dreieck erhalten durch: $\frac{\pi}{2} - \frac{\gamma}{2} - \frac{\gamma}{2} = \frac{\pi}{2} - \gamma$.

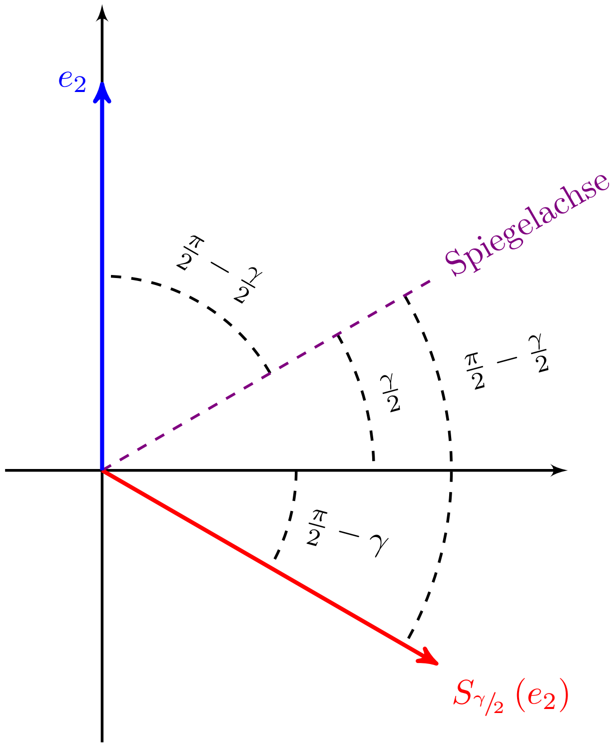


Abbildung 15

Beachten Sie, dass wir den Winkel unterhalb der x-Achse nun noch korrekt im mathematisch positiven Sinne umrechnen müssen: Somit wird aus dem Winkel $\frac{\pi}{2} - \gamma$ durch Negation $-\left(\frac{\pi}{2} - \gamma\right) = \gamma - \frac{\pi}{2}$ oder für die 2π-periodischen trigonometrischen Funktionen äquivalent $2\pi - (\frac{\pi}{2} - \gamma) = \frac{3\pi}{2} + \gamma$. Mit diesen Überlegungen erhalten wir schließlich die gesuchten Koordinationabschnitte. Die x-Koordinate ist gegeben durch:

$$\begin{aligned}\cos\left(-\left(\frac{\pi}{2} - \gamma\right)\right) &= \cos\left(\frac{\pi}{2} - \gamma\right) = \cos\left(\frac{\pi}{2}\right)\cos(\gamma) + \sin\left(\frac{\pi}{2}\right)\sin(\gamma) \\ &= 0 \cdot \cos(\gamma) + 1 \cdot \sin(\gamma) = \sin(\gamma).\end{aligned}$$

Für die y-Koordinate rechnen wir analog:

$$\begin{aligned}\sin\left(-\left(\frac{\pi}{2} - \gamma\right)\right) &= \sin\left(\gamma - \frac{\pi}{2}\right) = \sin(\gamma)\cos\left(\frac{\pi}{2}\right) - \cos(\gamma)\sin\left(\frac{\pi}{2}\right) \\ &= \sin(\gamma) \cdot 0 - \cos(\gamma) \cdot 1 = -\cos(\gamma).\end{aligned}$$

Damit ist $S_{\frac{\gamma}{2}}$ in der Tat die Spiegelung an der Geraden zum Winkel $\frac{\gamma}{2}$, so dass wir die folgende Aussage beweisen können:

Satz 8.4:

Es gilt $S_{\frac{\gamma}{2}} \cdot S_{\frac{\gamma}{2}} = E_2$.

Beweis:

$$\begin{pmatrix} \cos(\gamma) & \sin(\gamma) \\ \sin(\gamma) & -\cos(\gamma) \end{pmatrix} \cdot \begin{pmatrix} \cos(\gamma) & \sin(\gamma) \\ \sin(\gamma) & -\cos(\gamma) \end{pmatrix}$$

$$= \begin{pmatrix} \cos^2(\gamma)+\sin^2(\gamma) & \cos(\gamma)\sin(\gamma)-\sin(\gamma)\cos(\gamma) \\ \sin(\gamma)\cos(\gamma)-\cos(\gamma)\sin(\gamma) & \sin^2(\gamma)+\cos^2(\gamma) \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2$$

⊠

Schließlich erhält man durch einfaches Nachrechnen den folgenden

Satz 8.5:

Die Verknüpfung zweier Spiegelungen ist eine Drehung:

$$S_{\frac{\varphi}{2}} \cdot S_{\frac{\psi}{2}} = D_{\varphi-\psi}$$

Damit schließen wir unseren kurzen zwei-dimensionalen Exkurs. Wir kommen später im Kapitel über Eigenwerte noch einmal auf das Thema Drehungen und Spiegelungen zurück.

⋆ ⋆ ⋆

Im ersten Kapitel haben wir bereits den Algorithmus zum Lösen von linearen Gleichungssystemen kennengelernt. Im nächsten Kapitel werden wir unsere Kenntnisse über Matrizen benutzen, um diesen wichtigen Algorithmus auch als geeignetes Produkt von Matrizen nachvollziehen zu können.

Wir zeigen daher jetzt, dass die bekannten **elementaren Zeilenumformungen** von Gleichungssystemen, konkret von den entsprechenden Koeffizientenmatrizen, bereits durch Matrizenmultiplikation realisiert werden können.

Matrizen als elementare Umformung: Vertauschen von zwei Zeilen. Sei $\mathbb{K}$ ein Körper und seien $r < s \leq m$ natürliche Zahlen. Definiere nun die Matrix $C(r,s)$, gegeben als $(c_{ij})_{i=1,\ldots,m\,;\,j=1,\ldots,m}$, durch:

$$c_{ij} = \begin{cases} 1, & \text{falls } i=j \text{ und } i \neq r, i \neq s \\ 1, & \text{falls } i=r \text{ und } j=s \\ 1, & \text{falls } i=s \text{ und } j=r \\ 0, & \text{sonst} \end{cases}$$

Somit hat $C(r,s)$ die folgende Form:

$$C(r,s) = \begin{pmatrix} 1 & 0 & \ldots & & & & & & & \ldots & 0 \\ 0 & \ddots & \ddots & & & & & & & & \vdots \\ \vdots & \ddots & 1 & & & & & & & & \\ & & & 0 & \ldots & \ldots & 0 & 1 & & & \\ & & & \vdots & 1 & & & 0 & & & \\ & & & \vdots & & \ddots & & \vdots & & & \\ & & & 0 & & & 1 & \vdots & & & \\ & & & 1 & 0 & \ldots & & 0 & & & \\ & & & & & & & & 1 & \ddots & \vdots \\ \vdots & & & & & & & & \ddots & \ddots & 0 \\ 0 & \ldots & & & & & & & \ldots & 0 & 1 \end{pmatrix}$$

Die Anwendung von $C(r,s)$ ist wie gewünscht:

$$C(r,s)\cdot\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{r1} & \dots & a_{rn} \\ \vdots & & \vdots \\ a_{s1} & \dots & a_{sn} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{s1} & \dots & a_{sn} \\ \vdots & & \vdots \\ a_{r1} & \dots & a_{rn} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

Die Koeffizienten $\sum_{j=1}^{m} c_{ij}a_{jk}$ des Matrizenprodukts lassen sich folgendermaßen bestimmen: Für $i \leq m,\ i \neq r,s$ besteht die i-te Zeile aus den Koeffizienten

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

Die r-te Zeile besteht somit aus den Koeffizienten

$$\sum_{j=1}^{m} c_{rj}a_{jk} = c_{rs}a_{sk} = a_{sk}$$

umgekehrt gilt dies für die s-te Zeile. Damit ist gezeigt:

Satz 8.6:

Für eine beliebige Matrix $A \in \mathrm{Mat}(m \times n, \mathbb{K})$ ergibt sich das Matrizenprodukt $C(r,s)\cdot A$ aus A durch Vertauschen der r-ten Zeile mit der s-ten Zeile. Weiterhin gilt

$$C(r,s)\cdot C(r,s) = E_m.$$

Damit ist $C(r,s)$ invertierbar und es gilt $C(r,s)^{-1} = C(r,s)$.

Matrizen als elementare Umformung: Skalarmultiplikation einer Zeile. Sei $\mathbb{K}$ ein Körper. Seien n und $r \leq m$ natürliche Zahlen sowie $\lambda \in \mathbb{K}$. Definiere die Matrix $D(r,\lambda) = (d_{ij})_{i=1,\dots,m\,;\,j=1,\dots,m}$ durch:

$$d_{ij} = \begin{cases} 1 & \text{falls } i = j \text{ und } i \neq r \\ \lambda & \text{falls } i = j = r \\ 0 & \text{sonst} \end{cases}$$

Diese Matrix $D(r,\lambda)$ arbeitet wie gewünscht, denn für beliebig gegebene Matrizen $A \in \mathrm{Mat}(m\times n, \mathbb{K})$ gilt:

$$\begin{pmatrix} 1 & 0 & \dots & & & \dots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & 1 & & & & \\ & & & \lambda & & & \\ & & & & 1 & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \dots & & & \dots & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{r1} & \dots & a_{rn} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ \lambda a_{r1} & \dots & \lambda a_{rn} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

Denn die Koeffizienten des Produkts $\sum_{j=1}^{m} d_{ij}a_{jk}$ ergeben sich wie folgt:
Für $i \leq m, i \neq r$ besteht die i-te Zeile aus den Koeffizienten

$$\sum_{j=1}^{m} d_{ij}a_{jk} = d_{ii}a_{ik} = a_{ik}$$

und an der r-ten Zeile

$$\sum_{j=1}^{m} d_{rj}a_{jk} = d_{rr}a_{rk} = \lambda a_{rk}.$$

Damit ist gezeigt:

Satz 8.7:

Für $A \in \mathrm{Mat}(m \times n, \mathbb{K})$ ergibt sich das Matrizenprodukt $D(r, \lambda) \cdot A$ aus A durch Multiplikation der r-ten Zeile mit dem Faktor λ. Für $\lambda, \mu \in \mathbb{K}$ gilt

$$D(r, \lambda) \cdot D(r, \mu) = D(r, \lambda \cdot \mu).$$

Für $\lambda \neq 0$ ist $D(r, \lambda)$ invertierbar und es gilt $D(r, \lambda)^{-1} = D(r, \lambda^{-1})$
(da $D(r, \lambda \cdot \lambda^{-1}) = D(r, 1) = E_m$).

Matrizen als elementare Umformung: Addition zweier Zeilen. Sei $\mathbb{K}$ ein Körper. Sei $m \in \mathbb{N}$, $r, s \leq m$, $r \neq s$ und sei $\lambda \in \mathbb{K}$. Definiere die Matrix
$E(r, s, \lambda) = (e_{ij})_{i=1,\ldots,m\,;\,j=1,\ldots,m}$ durch

$$e_{ij} = \begin{cases} 1 & \text{falls } i = j \\ \lambda & \text{falls } i = r \text{ und } j = s \\ 0 & \text{sonst} \end{cases}$$

Diese Matrix ist von derselben Gestalt wie die Einheitsmatrix E_m, außer dass $e_{rs} = \lambda$ ist:

$$E(r, s, \lambda) = \begin{pmatrix} 1 & 0 & \ldots & & & \ldots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & 1 & \ldots & \lambda & & \\ & & & \ddots & \vdots & & \\ & & & & 1 & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \ldots & & & \ldots & 0 & 1 \end{pmatrix}$$

Sei $A = (a_{ij}) \in \mathrm{Mat}(m \times n, \mathbb{K})$. Dann ist:

$$E(r, s, \lambda) \cdot \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{r1} & \ldots & a_{rn} \\ \vdots & & \vdots \\ a_{s1} & \ldots & a_{sn} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{r1} + \lambda a_{s1} & \ldots & a_{rn} + \lambda a_{sn} \\ \vdots & & \vdots \\ a_{s1} & \ldots & a_{sn} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}$$

Die Koeffizienten $\sum_{j=1}^m e_{ij}a_{jk}$ des Matrizenprodukts lassen sich wie folgt berechnen: Für $i \leq m$, $i \neq r$ besteht die i-te Zeile aus den Koeffizienten

$$\sum_{j=1}^{m} e_{ij}a_{jk} = e_{ii}a_{ik} = a_{ik}$$

und die r-te Zeile besteht aus den Koeffizienten

$$\sum_{j=1}^{m} e_{rj}a_{jk} = e_{rr}a_{rk} + e_{rs}a_{sk} = a_{rk} + \lambda a_{sk}.$$

Satz 8.8:

Für $A \in \mathrm{Mat}(m \times n, \mathbb{K})$ ergibt sich das Matrizenprodukt $E(r, s, \lambda) \cdot A$ aus A durch Addition des λ-fachen der s-ten Zeile zur r-ten Zeile. Weiterhin gilt

$$E(r, s, \lambda) \cdot E(r, s, \mu) = E(r, s, \lambda + \mu).$$

Dann ist $E(r, s, \lambda)$ invertierbar und es gilt $E(r, s, \lambda)^{-1} = E(r, s, -\lambda)$
(da $E(r, s, \lambda + (-\lambda)) = E(r, s, 0) = E_m$).

Zusammenfassend halten wir fest:

Definition 8.9 – elementare Zeilenumformung:

Eine **elementare Zeilenumformung** ist eine der in den obigen Abschnitten beschriebenen Matrizenumformungen:

- Das Vertauschen zweier Zeilen (durch Anwendung von $C(r, s)$).
- Die Multiplikation einer Zeile mit einem skalaren Faktor ungleich 0 (durch Anwendung von $D(r, \lambda)$).
- Die Addition des Vielfachen einer Zeile zu einer anderen (durch Anwendung von $E(r, s, \lambda)$).

Die Hintereinanderausführung von elementaren Zeilenumformungen entspricht der sukzessiven Linksmultiplikation mit invertierbaren Matrizen der obigen Typen C, D und E.

Matrizen der Gestalt C, D und E heißen **Elementarmatrizen**.

⋆ ⋆ ⋆

Aufgabe 8.10: Warum gibt die Prüfmatrix H (aus dem obigen Beispiel) die Stelle eines 1 Bit Fehlers an?

Tipp 8.11:

Multiplizieren Sie H mit allen kanonischen Einheitsvektoren und betrachten Sie hinterher die Struktur des Resultats.

9. Matrizen und lineare Gleichungssysteme

Wir betrachten lineare Gleichungssysteme aus Sicht der Matrizenrechnung. Sei $\mathbb{K}$ ein Körper. Für $x_1, \ldots, x_m \in \mathbb{K}$ ist das linere Gleichungssystem über $\mathbb{K}$ von der Gestalt

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1m}x_m &= b_1 \\ \vdots \quad & \quad \vdots \\ a_{n1}x_1 + \ldots + a_{nm}x_m &= b_n \end{aligned}$$

Dies ist äquivalent zu der Matrizengleichung $A \cdot x = b$:

$$\begin{pmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nm} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

Man nennt die Matrix A die Koeffizientenmatrix und

$$(A, b) = \begin{pmatrix} a_{11} & \ldots & a_{1m} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \ldots & a_{nm} & b_n \end{pmatrix}$$

die erweiterte Koeffizientenmatrix des Gleichungssystems $A \cdot x = b$. Diese Schreibweise hatten wir bereits im ersten Kapitel gesehen. Wir können nun Eigenschaften von linearen Gleichungssystemen mit Hilfe der Matrizenrechnung ausdrücken.

- Das lineare Gleichungssystem $Ax = b$ heißt homogen, wenn $b = 0$ ist
- und inhomogen, wenn $b \neq 0$ ist.
- Für das lineare Gleichungssystem $Ax = b$ heißt

$$\text{Lös}(A, b) = \left\{ x \in \mathbb{K}^m \mid Ax = b \right\} = A^{-1}[\{b\}]$$

  Lösungsmenge.
- Das lineare Gleichungssystem heißt lösbar, wenn $b \in \text{Bild}(A)$.

Geometrie der Lösungsmengen. Sehen wir uns den Zusammenhang zwischen Lösungsmengen von linearen Gleichungssystemen und Untervektorräumen an:

Satz 9.1:

Betrachte das Gleichungssystem $A \cdot x = b$ mit $A \in \text{Mat}(n \times m, \mathbb{K})$. Dann gilt:

(a) Die Lösungsmenge des homogenen Systems $A \cdot x = 0$ ist

$$\text{Lös}(A, 0) = A^{-1}[\{0\}] = \text{Kern}(A).$$

Die Lösungsmenge ist ein Untervektorraum des $\mathbb{K}^m$.

(b) Das homogene Gleichungssystem $(A, 0)$ besitzt immer eine Lösung, nämlich den Nullvektor.

(c) Wenn (A, b) lösbar ist und $z \in \text{Lös}(A, b)$, dann ist

$$\text{Lös}(A, b) = z + \text{Lös}(A, 0) := \left\{ z + x \mid x \in \text{Lös}(A, 0) \right\}.$$

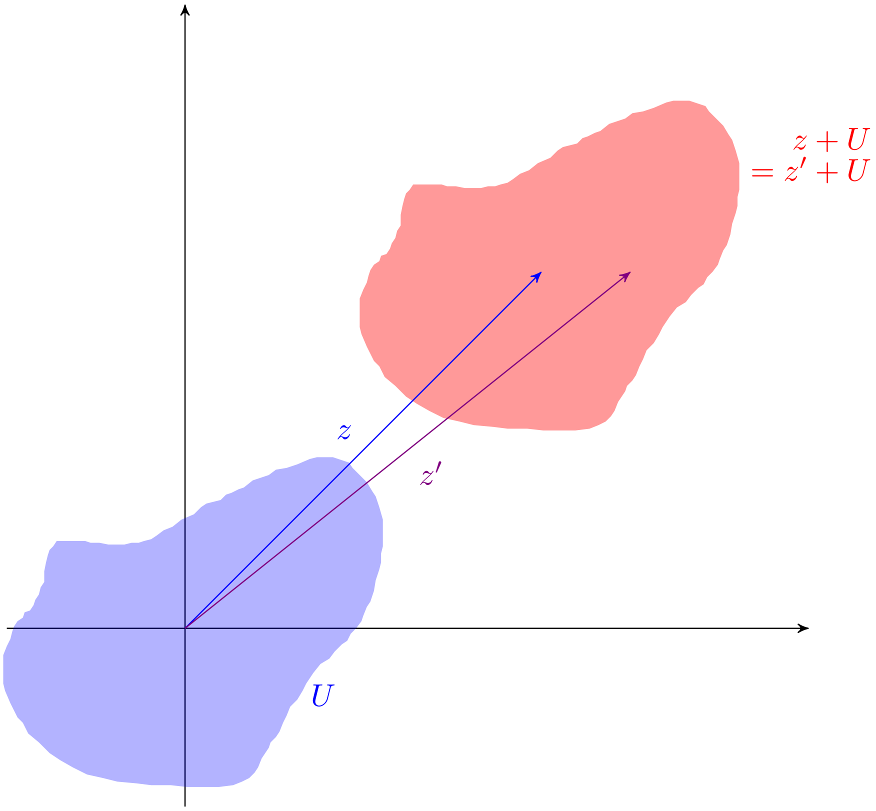


Abbildung 16: Unterraum U und affiner (verschobener) Unterraum $z+U$ – beachte: $z'+U = z+U$

Beweis:
Die ersten zwei Punkte sind klar. Wir zeigen (c):
Zu zeigen ist $\text{Lös}(A, b) = z + \text{Lös}(A, 0)$. Wir zeigen die Gleichheit der beiden Mengen durch beide Inklusionen:

„$\subseteq$" Betrachte $y \in \text{Lös}(A, b)$. Dann ist $A \cdot y = b$, $A \cdot z = b$. Insbesondere gilt: $A \cdot (-z) = -A \cdot z = -b$ und daher:

$$0 = b - b = b + (-b) = A \cdot y + A \cdot (-z) = A \cdot (y + (-z)) = A \cdot (y - z)$$

Somit ist $y - z \in \text{Kern}(A)$ und da $\text{Kern}(A) = \text{Lös}(A, 0)$ ist, ist $y \in z + \text{Lös}(A, 0)$ (da $y \in z + \text{Kern}(A)$).

„$\supseteq$" Betrachte $y \in z + \text{Lös}(A, 0)$. Wähle $x \in \text{Lös}(A, 0)$ mit $y = z + x$. Dann gilt:

$$A \cdot y = A \cdot (z + x) = A \cdot z + A \cdot x = b + 0 = b$$

Somit ist $y \in \text{Lös}(A, b)$.

⊠

Damit sind Lösungsmengen von homogenen linearen Gleichungssystemen stets Untervektorräume des Grundraums. Über $\mathbb{R}^2$ bedeutet dies, dass wir die triviale Lösung, also den Nullraum $\{0\}$, eine Ursprungsgerade, also einen 1-dimensionalen Untervektorraum, oder den gesamten Raum erhalten, da es nur einen 2-dimensionalen Untervektorraum gibt.

Lösungsmengen von inhomogenen linearen Gleichungssystemen sind um einen Vektor verschobene Untervektorräume.

Diese Verschiebung des Unterraumes kann man sich wie folgt vorstellen:
Der Himmel besitzt Sternenbilder. Je nach Standort hat man eine Verschiebung, auch wenn sich die

Sternenbilder selbst nicht ändern. Das Gesamtbild bleibt das gleiche, auch wenn man jeweils nur einen anderen Ausschnitt erkennt.
Zudem fällt auf, dass am Horizont keine weiße Fläche auftaucht, weil kein schwarzer Himmel mehr vorhanden ist. Hier rückt ein zuvor verborgener Teil in den Fokus.
Der Himmel an sich (Betrachtet als die gesamte Kugelinnenschale, unabhängig vom Beobachtungspunkt und -ausschnitt), ändert sich während all dem nicht.

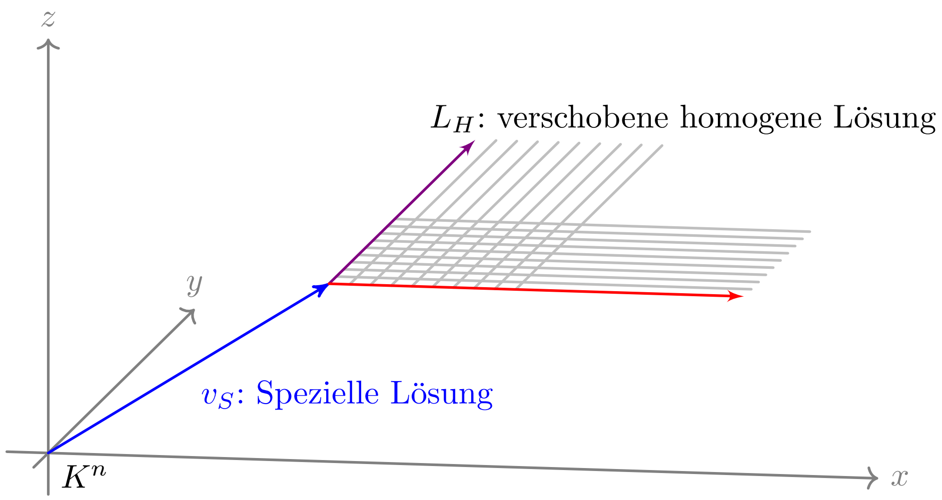


Abbildung 17: die spezielle Lösung verschiebt die homogene Lösung

Wir zeigen umgekehrt, dass sich die geometrischen Grundelemente wie Punkte, Geraden, Ebenen usw. in allgemeiner Lage als Lösungsmengen darstellen lassen. Damit lassen sich geometrische Aufgaben wie die Berechnung von Schnittmengen von Geraden, Ebenen usw. auf die Lösung linearer Gleichungssysteme zurückführen.

Satz 9.2:

Sei $\mathbb{K}$ ein Körper, $m \in \mathbb{N}$ und U ein Untervektorraum von $\mathbb{K}^m$.

(a) Es gibt ein homogenes lineares Gleichungssystem $(A, 0)$, sodass $U = \text{Lös}(A, 0)$.

(b) Sei $z \in \mathbb{K}^m$. Dann gibt es ein lineares Gleichungssystem (A, b), sodass

$$z + U = \{z + x \mid x \in U\} = \text{Lös}(A, b).$$

Beweis:

(a) Nach dem Basisexistenzsatz wähle eine Basis $(v_1, \ldots, v_l)$ von U. Nach dem Basisergänzungssatz wähle $v_{l+1}, \ldots, v_m$, so dass $(v_1, \ldots, v_l, v_{l+1}, \ldots, v_m)$ eine Basis von $\mathbb{K}^m$ ist. Wir definieren die folgende lineare Abbildung $A : \mathbb{K}^m \to \mathbb{K}^{m-l}$ durch die Angabe der Bilder der obigen Basisvektoren:

$$A(v_1) = \ldots = A(v_l) = 0,$$

$$A(v_{l+1}) = e_1,\ A(v_{l+2}) = e_2,\ \ldots,\ A(v_m) = e_{m-l},$$

wobei $e_1, \ldots, e_{m-l}$ die kanonische Basis von $\mathbb{K}^{m-l}$ ist.

Behauptung 9.3: Dann ist $U = \text{Kern}(A)$.

Beweis der Behauptung: Wir zeigen beide Inklusionen:

„$\subseteq$“ Betrachte ein $x \in U$. Wähle $\lambda_1, \ldots, \lambda_l \in \mathbb{K}$, sodass gilt $x = \sum_{i=1}^{l} \lambda_i v_i$. Dann haben wir:

$$A \cdot x = A \cdot \left(\sum_{i=1}^{l} \lambda_i v_i \right) = \sum_{i=1}^{l} \lambda_i \cdot A(v_i) = 0,$$

also ist $x \in \mathrm{Kern}(A)$.

„$\supseteq$“ Sei $x \in \mathrm{Kern}(A)$. Wähle $\lambda_1, \ldots, \lambda_l, \lambda_{l+1}, \ldots, \lambda_m \in \mathbb{K}$ mit $x = \sum_{i=1}^{m} \lambda_i v_i$. Dann gilt:

$$\begin{aligned} 0 &= A \cdot x = A \left(\sum_{i=1}^{m} \lambda_i v_i \right) = \sum_{i=1}^{m} \lambda_i A(v_i) \\ &= \sum_{i=1}^{l} \lambda_i \underbrace{A(v_i)}_{=0} + \sum_{i=l+1}^{m} \lambda_i A(v_i) \\ &= \lambda_{l+1} e_1 + \ldots + \lambda_m e_{m-l} \end{aligned}$$

Da die kanonische Basis $e_1, \ldots, e_{m-l}$ linear unabhängig ist, gilt:

$$\lambda_{l+1} = \ldots = \lambda_m = 0.$$

Damit ist $x = \sum_{i=1}^{l} \lambda_i v_i \in \mathcal{L}(v_1, \ldots, v_l) = U$. Indem wir A als Matrix in $\mathrm{Mat}(m-l \times m, \mathbb{K})$ auffassen, ist $(A, 0)$ ein homogenes Lösungssystem mit

$$U = \mathrm{Kern}(A) = \mathrm{Lös}(A, 0).$$

$\boxtimes$(Beh.)

(b) Setze $b := A(z)$. Dann ist $z \in \mathrm{Lös}(A, b)$. Nach Satz 9.1 gilt

$$z + U = z + \mathrm{Lös}(A, 0) = \mathrm{Lös}(A, b).$$

⊠

Auf dem Weg zum Gaußschen Algorithmus betrachten wir folgende Definition:

Definition 9.4 – obere Dreiecksmatrix, Diagonalmatrix, Zeilenstufenmatrix:

Sei $\mathbb{K}$ ein Körper und $A = (a_{ij}) \in \mathrm{Mat}(n \times m, \mathbb{K})$.

(a) Die Matrix A heißt obere Dreiecksmatrix, wenn A quadratisch ist und für $i, j \leq n$ mit $i > j$ gilt: $a_{ij} = 0$:

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

(b) Die Matrix A heißt Diagonalmatrix, wenn A quadratisch ist und für alle $i, j \leq n$ mit $i \neq j$ gilt: $a_{ij} = 0$

$$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}$$

(c) Die Matrix A heißt Zeilenstufenmatrix (kurz: ZSM), wenn für $i < i' \leq n$ gilt: wenn $j(i)$ das kleinste j mit $a_{ij} \neq 0$ ist und wenn $j(i')$ das kleinste j mit $a_{i'j} \neq 0$ ist, dann ist $j(i) < j(i')$, d.h. das Pivotelement der oberen Zeile i steht weiter links als das Pivotelement der unteren Zeile i':

$$A = \begin{pmatrix} 0 & a_{1j(1)} & \dots & a_{1j(2)} & a_{1j(3)} & \dots & \dots \\ 0 & \dots & 0 & a_{2j(2)} & a_{2j(3)} & \dots & \\ 0 & \dots & \dots & \dots & 0 & a_{3j(3)} & \dots \\ 0 & \dots & \dots & \dots & \dots & 0 & \ddots \end{pmatrix}$$

oder schematisch, wenn wir $*$ als Zeichen für beliebige Körperelemente und # als Zeichen für invertierbare Körperelemente benutzen:

$$A = \begin{pmatrix} 0 & \# & \dots & * & * & \dots & \dots \\ 0 & \dots & 0 & \# & * & \dots & \\ & & & & 0 & \# & \dots \\ & & & & & & \ddots \end{pmatrix}$$

Das lineare Gleichungssystem $Ax = b$ ist von quadratischer, oberer Dreiecksform, Diagonal- bzw. Zeilenstufenform, wenn die Koeffizientenmatrix A von der entsprechenden Form ist.

⋆ ⋆ ⋆

gaußscher Algorithmus. Um ein beliebiges lineares Gleichungssystem zu lösen, wird das System in ein System in Zeilenstufenform mit gleicher Lösungsmenge transformiert. Dieses wird dann wie bisher gelöst. Grundlage dieses Verfahrens sind die folgenden Sätze:

Satz 9.5:

Sei $\mathbb{K}$ ein Körper und seien $(A,b),(A',b')$ lineare Gleichungssysteme über $\mathbb{K}$, so dass die erweiterte Koeffizientenmatrix (A',b') durch eine Folge elementarer Zeilenumformungen aus (A,b) hervorgeht. Dann ist

$$\text{Lös}(A,b)=\text{Lös}(A',b').$$

Beweis: Wähle ein Produkt $P=F_1\cdot F_2\cdot\ldots\cdot F_r$ von Elementarmatrizen, sodass $A'=P\cdot A$ und $b'=P\cdot b$. Wir wissen, dass P als elementare Zeilenumformung invertierbar ist, sodass gilt:

$$P^{-1}=(F_1\cdot F_2\cdot\ldots\cdot F_r)^{-1}=F_r^{-1}\cdot\ldots\cdot F_2^{-1}\cdot F_1^{-1}$$

Insgesamt definiert P eine injektive Abbildung. Damit gilt also:

$$\begin{aligned}\text{Lös}(A,b)&=\left\{x\mid Ax=b\right\}\\&=\left\{x\mid P\cdot A\cdot x=P\cdot b\right\}\\&=\text{Lös}\left(P\cdot A,P\cdot b\right)\end{aligned}$$

⊠

Satz 9.6:

Sei $\mathbb{K}$ ein Körper und $A\in\text{Mat}(n\times m,\mathbb{K})$. Dann lässt sich A durch eine Folge von elementaren Zeilenumformungen in eine Matrix A' in Zeilenstufenform umformen.

Beweis: Klar, nach dem früheren Kapitel über lineare Gleichungssysteme. ⊠

⋆ ⋆ ⋆

Algorithmus zum Invertieren von Matrizen. Das Invertieren einer Matrix $A=(a_{ij})\in\text{Mat}(m\times m,\mathbb{K})$ entspricht der Lösung des Systems von Gleichungen

$$\begin{pmatrix}a_{11}&\ldots&a_{1m}\\\vdots&\ddots&\vdots\\a_{m1}&\ldots&a_{mm}\end{pmatrix}\cdot\begin{pmatrix}x_{11}&\ldots&x_{1m}\\\vdots&\ddots&\vdots\\x_{m1}&\ldots&x_{mm}\end{pmatrix}=\begin{pmatrix}1&0&\ldots&0\\0&\ddots&\ddots&\vdots\\\vdots&\ddots&\ddots&0\\0&\ldots&0&1\end{pmatrix}$$

Dies ist ein lineares Gleichungssystem der Gestalt

$$\sum_{j=1}^{m}a_{ij}x_{ji}=1$$

beziehungsweise

$$\sum_{j=1}^{m}a_{ij}x_{jk}=0\quad\text{für}\quad i\neq k$$

Das Gleichungssystem hat m^2 viele Unbekannte x_{jk} und m^2 viele Gleichungen.

Wir stellen ein Verfahren zum Invertieren von A vor, das die von der Gaußelimination bekannten Zeilenumformungen auf effiziente Weise ausnutzt. Wenn die Matrix durch elementare Zeilenumformungen in die Einheitsmatrix überführt werden kann, dann existieren elementare Zeilenumformungen $F_1, \ldots, F_l$ mit

$$F_1 \cdot \ldots \cdot F_l \cdot A = E_m.$$

(Klar: Wenn A invertierbar ist, dann lässt sich A mittels elementarer Zeilenumformungen in die Einheitsmatrix überführen.)

Wir wissen bereits, dass dann das Produkt $F_1 \cdot \ldots \cdot F_l$ invers zu A ist, das heißt es gilt:

$$A^{-1} = F_1 \cdot \ldots \cdot F_l = F_1 \cdot \ldots \cdot F_l \cdot E_m$$

Also: Man erhält daher dieses Produkt, indem man die elementaren Zeilenumformungen wie bei A parallel auf die Einheitsmatrix E_m anwendet.

Beispiel 9.7: Betrachte die Matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$. Gesucht ist die Inverse A^{-1}:

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -2 & -2 & -1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \rightsquigarrow \cdots \rightsquigarrow$$
$$\rightsquigarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{2} & \frac{3}{2} & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & \frac{1}{2} & -\frac{1}{2} & -1 \end{array}\right)$$

Damit ist die Matrix $\begin{pmatrix} -\frac{1}{2} & \frac{3}{2} & 1 \\ 0 & 0 & 1 \\ \frac{1}{2} & -\frac{1}{2} & -1 \end{pmatrix}$ die Inverse der Matrix $\begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.

Der <u>Algorithmus</u> Schritt für Schritt:

- links steht A, rechts steht E_m
- solange umformen, bis E_m links steht
- dann steht rechts A^{-1}

⋆ ⋆ ⋆

Aufgabe 9.8: Invertieren Sie die nachfolgenden Matrizen:

$$A_1 = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 2 & 2 & 2 \\ 3 & 1 & 1 \\ 2 & 1 & 4 \end{pmatrix} \qquad A_3 = \begin{pmatrix} 1 & 1 & 3 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

Aufgabe 9.9: Zeigen Sie, dass die (2×2)-Matrix:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad a, b, c, d \in \mathbb{R}$$

genau dann invertierbar ist, wenn $ad - bc \neq 0$

Aufgabe 9.10: Wann ist die Matrix:

$$B = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & e \end{pmatrix} \qquad a, b, c, d, e \in \mathbb{R}$$

invertierbar und wie sieht ihre Inverse aus?

Tipp 9.11:

Konstruieren Sie die Inverse der (2×2)-Blockmatrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, dann lässt sich die Inverse von B aus 2 Blockmatrizen „zusammenbauen“.

10. Koordinatentransformation

Oft kann man durch Wahl geeigneter Koordinaten Berechnungen und Argumente vereinfachen. Im Kontext von Vektorräumen werden durch die Darstellung von Vektoren in einer Basis Koordinaten definiert.

Basiswechsel. Zunächst überlegen wir uns, wie man die Koordinaten von ein und demselben Vektor bezüglich verschiedener Basen übertragen kann.

Definition 10.1 – Transformationsmatrix:

Sei V ein $\mathbb{K}$-Vektorraum und seien $B = (v_1, \ldots, v_m)$ und $C = (w_1, \ldots, w_m)$ Basen von V. Die Transformationsmatrix des Basiswechsels von B nach C ist die Matrix $T_C^B = (t_{ij}) \in \operatorname{Mat}(m \times m, \mathbb{K})$, so dass für alle Spaltenindizes $j \leq m$ die Spalte $\begin{pmatrix} t_{1j} \\ \vdots \\ t_{mj} \end{pmatrix}$ die Darstellung von v_j in der Basis $C = (w_1, \ldots, w_m)$ ist:

$$v_j = \sum_{i=1}^{m} t_{ij} w_i = t_{1j} w_1 + \ldots + t_{mj} w_m,$$

das heißt $\begin{pmatrix} t_{1j} \\ \vdots \\ t_{mj} \end{pmatrix}$ sind gerade die Koordinaten der Basisdarstellung des Vektors v_j bezüglich C.

Diese Transformationsmatrizen verdienen ihren Namen, wie der folgende Satz zeigen wird:

Satz 10.2:

Sei V ein $\mathbb{K}$-Vektorraum mit Basen $B = (v_1, \ldots, v_m)$ und $C = (w_1, \ldots, w_m)$ und der Transformationsmatrix $T_C^B = (t_{ij})$. Sei $z \in V$ und seien $(\lambda_1, \ldots, \lambda_m)$ und $(\mu_1, \ldots, \mu_m)$ die Darstellungen des Vektors z in den Basen B bzw. C

$$z = \sum_{i=1}^{m} \lambda_i v_i = \sum_{i=1}^{m} \mu_i w_i.$$

Dann gilt:

$$\begin{pmatrix} \mu_1 \\ \vdots \\ \mu_m \end{pmatrix} = T_C^B \cdot \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix}$$

Beweis: Mithilfe der Definition der Transformationsmatrix gilt:

$$\begin{aligned} \sum_{i=1}^{m} \mu_i w_i = \sum_{i=1}^{m} \lambda_i v_i &= \lambda_1 \cdot \sum_{i=1}^{m} t_{i1} \cdot w_i + \ldots + \lambda_m \cdot \sum_{i=1}^{m} t_{im} \cdot w_i \\ &= \sum_{j=1}^{m} (\lambda_1 \cdot t_{j1} + \ldots + \lambda_m \cdot t_{jm}) \cdot w_j \end{aligned}$$

$$= \sum_{j=1}^{m} (t_{j1}\lambda_1 + \ldots + t_{jm}\lambda_m) w_j$$

Da C eine Basis ist und die Darstellung bzgl. einer Basis eindeutig ist, gilt:

$$\mu_j = t_{j1}\lambda_1 + \ldots + t_{jm}\lambda_m$$

Damit folgt die Behauptung. ⊠

Die Transformationsmatrizen erfüllen einfache, zu erwartende Eigenschaften:

Satz 10.3:

Sei V ein m-dimensionaler $\mathbb{K}$-Vektorraum mit Basen B, C und D. Dann gilt:

(a) $T_B^B = E_m$
(b) $T_D^B = T_D^C \cdot T_C^B$
(c) $T_B^C = (T_C^B)^{-1}$

Beweis: Klar nach Definition 10.1 und Satz 10.2 ⊠

Beispiel 10.4: Sei $V = \mathbb{K}^m$ und K die kanonische Basis $(e_1, \ldots, e_m)$. Weiterhin sei $C = (w_1, \ldots, w_m)$ eine Basis von V. Für $j \leq m$ ist der Vektor $w_j = \begin{pmatrix} w_{1j} \\ \vdots \\ w_{mj} \end{pmatrix}$ seine eigene Darstellung in der kanonischen Basis. Daher ergibt sich die Transformationsmatrix für den Wechsel von C nach K als

$$T_K^C = (w_1, \ldots, w_m) = \begin{pmatrix} w_{11} & \ldots & w_{1m} \\ \vdots & & \vdots \\ w_{m1} & \ldots & w_{mm} \end{pmatrix}$$

das heißt die Basis C ist im wesentlichen die Transformationsmatrix von der Basis C zur kanonischen Basis K. Die Matrix T_C^K erhalten wir nach dem Satz 10.3 durch Invertieren von T_K^C.

Beispiel 10.5: Seien $C = \left(\begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$, $D = \left(\begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$ zwei beliebige und $K = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$ die kanonische Basis. Wollen wir beispielsweise den Vektor $\begin{pmatrix} 2 \\ 1 \end{pmatrix} \in \mathbb{R}^2$ bezüglich der Basen K, C und D darstellen, so gilt:

$$\begin{aligned} \begin{pmatrix} 2 \\ 1 \end{pmatrix} &= 2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} && \text{bzgl. } K \\ &= 1 \cdot \begin{pmatrix} 2 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} && \text{bzgl. } C \\ &= 1 \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} && \text{bzgl. } D \end{aligned}$$

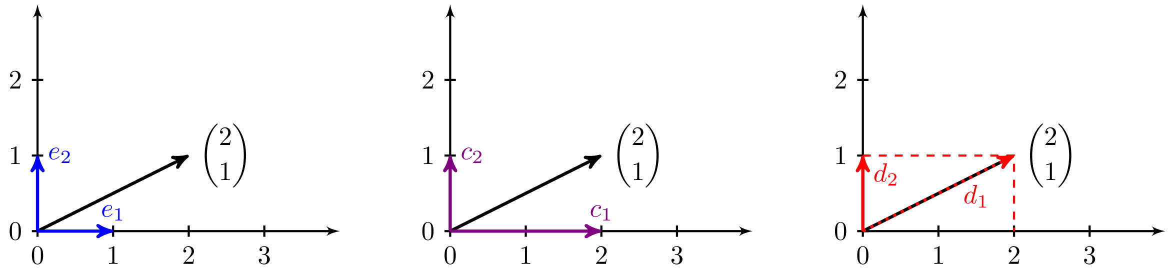
oder kurz: $\begin{pmatrix} 2 \\ 1 \end{pmatrix}_K \equiv \begin{pmatrix} 1 \\ 1 \end{pmatrix}_C \equiv \begin{pmatrix} 1 \\ 0 \end{pmatrix}_D$

Bemerkung 10.6: Wir können diese Koordinatenzuweisung auch als Abbildung darstellen: Sei dafür etwa $v = \begin{pmatrix} 2 \\ 1 \end{pmatrix}_K$ der gegebene Vektor. Dann hat, wie wir gerade gesehen haben, v bezüglich der

Basis D die Koordinaten $\begin{pmatrix} 1 \\ 0 \end{pmatrix}_D$. Somit können wir eine Koordinatenabbildung $K_D^{\mathbb{R}^2} : \mathbb{R}^2 \to \mathbb{R}^2$ definieren, wobei einem Vektor v des Vektorraums seine Koordinaten in $\mathbb{R}^2$ zugeordnet werden:

$$v \mapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix}_D , \qquad K_D^{\mathbb{R}^2}(v) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_D .$$

Allgemein bildet daher eine Koordinatenabbildung K_B^V bezüglich einer Basis B eines n-dimensionalen $\mathbb{K}$-Vektorraums V in den $\mathbb{K}^n$ ab, indem einem Vektor $v \in V$ seine Koordinaten bezüglich der Basis B zugeordnet werden: $K_B^V : V \to \mathbb{K}^2$.



(a) kanonische Basis K **(b)** Basis $C = \left\{ \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$ **(c)** Basis $D = \left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$

Abbildung 18: Darstellung des Vektors $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ durch verschiedene Basen

- <u>Ziel:</u> $T_C^K \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix}_K \overset{!}{=} \begin{pmatrix} 1 \\ 1 \end{pmatrix}_C$

  Einfach zu bestimmen ist T_K^C, denn es gilt: $T_K^C = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$.

  Wir invertieren nun T_K^C um T_C^K zu erhalten:

  $$\begin{array}{cc|cc} 2 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{array} \quad \rightsquigarrow \quad \begin{array}{cc|cc} 1 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & 1 \end{array}$$

  also gilt: $T_C^K = (T_K^C)^{-1} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix}$ und damit:

  $$T_C^K \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad (\checkmark)$$

- <u>Ziel:</u> $T_D^K \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix}_K \overset{!}{=} \begin{pmatrix} 1 \\ 0 \end{pmatrix}_D$

  Einfach zu bestimmen ist T_K^D, denn es gilt: $T_K^D = \begin{pmatrix} 2 & 0 \\ 1 & 1 \end{pmatrix}$

  Wir invertieren nun T_K^D um T_D^K zu erhalten:

  $$\left(\begin{array}{cc|cc} 2 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 0 & \frac{1}{2} & 0 \\ 1 & 1 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & -\frac{1}{2} & 1 \end{array}\right)$$

also gilt: $T_D^K = \left(T_K^D\right)^{-1} = \begin{pmatrix} \frac{1}{2} & 0 \\ -\frac{1}{2} & 1 \end{pmatrix}$ und damit:

$$T_D^K \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ -\frac{1}{2} & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad (\checkmark)$$

- Ziel: $T_D^C \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}_C \overset{!}{=} \begin{pmatrix} 1 \\ 0 \end{pmatrix}_D$
  Wir benutzen die Eigenschaft (b) aus Satz 10.3:
  $$T_D^C = T_D^K \cdot T_K^C,$$
  wobei K die kanonische Basis ist. Damit gilt:
  $$\begin{pmatrix} \frac{1}{2} & 0 \\ -\frac{1}{2} & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$$
  Also gilt:
  $$T_D^C \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad (\checkmark)$$

Bemerkung 10.7: Wir können jetzt alle Basen ineinander überführen. Manchmal ist der Umweg über die kanonische Basis sinnvoll, da dies in der Praxis meist einfacher ist.

⋆ ⋆ ⋆

Darstellende Matrizen von linearen Abbildungen bezüglich beliebiger Basen.

Motivation

Wir möchten nicht nur von Homomorphismen des $\mathbb{K}^n$ in den $\mathbb{K}^m$ darstellende Matrizen haben, sondern von linearen Abbildungen zwischen beliebigen Vektorräumen (und beliebigen Basen).

Definition 10.8 – darstellende Matrix:

Sei $f : V \to W$ eine lineare Abbildung zwischen endlich dimensionalen $\mathbb{K}$-Vektorräumen. Sei $B = (v_1, \ldots, v_m)$ eine Basis für V und $C = (w_1, \ldots, w_n)$ eine Basis für W. Dann definiere $\mathrm{DM}_C^B(f) = \begin{pmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nm} \end{pmatrix} \in \mathrm{Mat}(n \times m, \mathbb{K})$ als die darstellende Matrix von f bezüglich der Basen B und C durch: für $j \leq m$ ist

$$f(v_j) = a_{1j}w_1 + \ldots + a_{nj}w_n.$$

Das bedeutet, dass jeweils die Spalte $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}_C$ die Darstellung von $f(v_j)$ in der Basis C ist.

Bemerkung 10.9: Der bereits definierte Begriff der darstellenden Matrix ist hier ein Spezialfall für $V = \mathbb{K}^n$, $W = \mathbb{K}^m$ und B und C jeweils die kanonische Basis. Insbesondere bleibt auch die Merkregel 7.3 zur Aufstellung der darstellenden Matrix allgemein erhalten, wobei Sie die –wie in der Definition angegeben– Bilder der B-Basisvektoren in den Koordinaten der Basis C darstellen müssen!
Beachten Sie folgendes Beispiel, welches die Theorie der darstellenden Matrizen Ihnen noch einmal näher bringen sollte:

Beispiel 10.10: Sei $\mathtt{Pol}_3$ der Vektorraum der Polynome vom Grade echt kleiner 3 mit reellen Koeffizienten. Es sei $F : \mathtt{Pol}_3 \to \mathtt{Pol}_3$ die Abbildung

$$F\big(p(x)\big) \;:=\; \frac{d}{dx}\big(x \cdot p(x)\big)\,.$$

Diese Abbildung ist nach Beispiel 6.5 linear. Wir berechnen nun die darstellende Matrix $\mathrm{DM}_C^B(f)$ für die beiden Basen $B := \big(1, 1+x, x^2\big)$ und $C := \big(1+x^2, 1-x^2, x\big)$.

Dazu können Sie die Darstellungsmatrix berechnen, indem Sie die Bilder der Basis B berechnen und diese als Linarkombination der Basis C darstellen. So ist $F(1) = \frac{d}{dx}(\,x \cdot 1\,) = 1$ und diesen Funktionswert können wir wie folgt darstellen:

$$F(1) = 1 = \frac{1}{2}\Big((1+x^2) + (1-x^2)\Big)\,.$$

Somit wird der Vektor $\begin{pmatrix}1\\0\\0\end{pmatrix}_B$ auf den Vektor $\begin{pmatrix}\frac{1}{2}\\\frac{1}{2}\\0\end{pmatrix}_C$ abgebildet.

Ähnlich folgen die beiden anderen Darstellungen:

$$\begin{aligned}
F\,(1+x) \quad &= \quad \frac{\mathrm{d}}{\mathrm{d}x}\,(x\cdot(1+x)) \quad &&= \quad 1+2x \quad &&= \quad \frac{1}{2}\left(\left(1+x^2\right)+\left(1-x^2\right)\right)+2x\\
F\left(x^2\right) \quad &= \quad \frac{\mathrm{d}}{\mathrm{d}x}\left(x\cdot x^2\right) \quad &&= \quad 3x^2 \quad &&= \quad \frac{3}{2}\left(\left(1+x^2\right)-\left(1-x^2\right)\right)
\end{aligned}$$

Also wird $\begin{pmatrix}0\\1\\0\end{pmatrix}_B$ auf $\begin{pmatrix}\frac{1}{2}\\\frac{1}{2}\\2\end{pmatrix}_C$ sowie $\begin{pmatrix}0\\0\\1\end{pmatrix}_B$ auf $\begin{pmatrix}\frac{3}{2}\\-\frac{3}{2}\\0\end{pmatrix}_C$ abgebildet. Die gesuchte darstellende Matrix lautet daher zusammenfassend: $\mathrm{DM}_C^B(f) = \begin{pmatrix}\frac{1}{2} & \frac{1}{2} & \frac{3}{2}\\ \frac{1}{2} & \frac{1}{2} & -\frac{3}{2}\\ 0 & 2 & 0\end{pmatrix}$.

⋆ ⋆ ⋆

Das Verhalten der darstellenden Matrizen können wir auch sehr einfach mittels Transformationsmatrizen zeigen:

Satz 10.11:

Sei $f : V \to W$ eine lineare Abbildung zwischen endlich-dimensionalen $\mathbb{K}$-Vektorräumen. Seien $B = (v_1, \ldots, v_m)$ und $B' = (v_1', \ldots, v_m')$ Basen von V und $C = (w_1, \ldots, w_n)$ und $C' = (w_1', \ldots, w_n')$ Basen für W. Dann gilt:

$$\mathrm{DM}_{C'}^{B'}(f) = T_{C'}^{C} \cdot \mathrm{DM}_C^B(f) \cdot T_B^{B'}$$

Beweis: Betrachte $j \leq m$. Es sei e_j der j-te kanonische Basisvektor von $\mathbb{K}^m$. Dann ist $\mathrm{DM}_{C'}^{B'}(f) \cdot e_j$ die Darstellung von $f(v_j')$ bezüglich der Basis C'. Andererseits ist $T_B^{B'} \cdot e_j$ die Darstellung von v_j' in der Basis B. Dann ist

$$\mathrm{DM}_C^B(f) \cdot \left(T_B^{B'} \cdot e_j\right)$$

die Darstellung von $f(v_j')$ in der Basis C.

$$T_{C'}^C \cdot \left(\mathrm{DM}_C^B(f) \cdot \left(T_B^{B'} \cdot e_j\right)\right)$$

ist dann die Darstellung von $f(v_j')$ in der Basis C'. Damit gilt:

$$\mathrm{DM}_{C'}^{B'}(f) \cdot e_j = \left(T_{C'}^C \cdot \mathrm{DM}_C^B(f) \cdot T_B^{B'}\right) \cdot e_j$$

Also sind die Matrizen spaltenweise identisch. ⊠

Bemerkung 10.12: Mittels des obigen Satzes erhalten wir durch den Übergang zur kanonischen Basis schnell allgemeine darstellende Matrizen, denn für $V = \mathbb{K}^m, W = \mathbb{K}^n$ und K_m, K_n die kanonischen Basen gilt:

$$\mathrm{DM}_{C'}^{B'}(f) = T_{C'}^{K_n} \cdot \mathrm{DM}_{K_n}^{K_m}(f) \cdot T_{K_m}^{B'}$$

Bemerkung 10.13: Das im Satz 10.11 dargestellte Transformationsverhalten der darstellenden Matrizen können wir durch ein kommutatives Diagramm darstellen:

$$\begin{tikzcd}
V \arrow[r, "\mathrm{DM}_{C'}^{B'}(f)"] \arrow[d, "T_B^{B'}"'] & W \\
V \arrow[r, "\mathrm{DM}_C^B(f)"'] & W \arrow[u, "T_{C'}^C"']
\end{tikzcd}$$

Zusammenfassend erhalten wir für einen Homomorphismus $f : V \to W$ zu zwei $\mathbb{R}$-Vektorräumen V, W mit Basen B, B' für V bzw. C, C' für W und dazugehörigen Koordinatenabbildungen K_B^V, $K_{B'}^V$ bzw. K_C^W, $K_{C'}^W$ sowie Koordinatentransformationsmatrizen $T_B^{B'}$ bzw. $T_C^{C'}$ (von Koordinatendarstellungen bezüglich der Basis B' bzw. C' in die Darstellung betreffs der Basis B bzw. C) folgendes Diagramm:

$$\begin{tikzcd}
{} & V \arrow[r, "f"] \arrow[d, "K_B^V"] \arrow[ddl, "K_{B'}^V"'] & W \arrow[d, "K_C^W"'] \arrow[ddr, "K_{C'}^W"] & {} \\
 & \mathbb{R}^n \arrow[r, "\mathrm{DM}_C^B(f)"] & \mathbb{R}^m & \\
\mathbb{R}^n \arrow[ur, "T_B^{B'}"] \arrow[rrr, "\mathrm{DM}_{C'}^{B'}(f)"] & & & \mathbb{R}^m \arrow[ul, "T_C^{C'}"']
\end{tikzcd}$$

⋆ ⋆ ⋆

Langfristiges Ziel

Möglichst einfache Darstellung von Matrizen, das heißt durch geschickte Wahl der Basen möchten wir möglichst viele Nullen in der darstellenden Matrix erzeugen können.

Eine Teillösung liefert bereits der folgende Satz:

Satz 10.14:

Sei $f : V \to W$ eine lineare Abbildung vom m-dimensionalen $\mathbb{K}$-Vektorraum V in den n-dimensionalen $\mathbb{K}$-Vektorraum W vom Rang: $\operatorname{Rg}(f) = \dim\left(\operatorname{Bild}(f)\right) = r$. Dann existieren Basen $B = (v_1, \ldots, v_m)$ von V und $C = (w_1, \ldots, w_n)$ von W, so dass

$$\mathrm{DM}_C^B(f) = \left(\begin{array}{ccc|ccc} 1 & 0 & \ldots & & \ldots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & 1 & & & \\ \hline & & & 0 & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \ldots & & \ldots & 0 & 0 \end{array}\right)$$

wobei r viele Einsen auf der Hauptdiagonalen der oben links entstandenen Blockmatrix stehen und sonst nur Nullen.

Warum ist dies nur eine Teillösung? Schließlich brauchen wir doch genau r Einsen, da $\operatorname{Rg}(f) = r$, und ansonsten haben wir bereits nur Nullen! Das Problem ist, dass wir für die Vektorräume V, W Basen B, C wählen mussten, die im Allgemeinen nicht gleich sind, selbst wenn $V = W$ ist. In diesem Fall möchten wir aber eigentlich auch $B = C$ haben. Wir werden innerhalb des Beweises sehen, dass wir dies aber nicht erwarten können.

Beweis von Satz 10.14: Nach der Dimensionsformel für lineare Abbildungen gilt:

$$\dim\left(\operatorname{Kern}(f)\right) + \dim\left(\operatorname{Bild}(f)\right) = m$$

Setze $s = \dim(\operatorname{Kern}(f)) = m - r$. Wir benutzen die Idee aus dem Beweis des Dimensionssatzes 6.19. Wähle eine Basis $(v_{r+1}, \ldots, v_m)$ von $\operatorname{Kern}(f)$. Nach dem Basisergänzungssatz wähle $v_1, \ldots, v_r \in V$, sodass $B = (v_1, \ldots, v_r, v_{r+1}, \ldots, v_m)$ eine Basis von V ist. Somit gilt:

$$\begin{aligned} \operatorname{Bild}(f) &= \left\{ f(v) \;\middle|\; v \in V = \mathcal{L}(v_1, \ldots, v_r, v_{r+1}, \ldots, v_m) \right\} \\ &= \left\{ f(\lambda_1 v_1 + \ldots + \lambda_r v_r + \lambda_{r+1} v_{r+1} + \ldots + \lambda_m v_m) \;\middle|\; \lambda_i \in \mathbb{K} \right\} \\ &= \left\{ \lambda_1 f(v_1) + \ldots + \lambda_r f(v_r) + \lambda_{r+1} f(v_{r+1}) + \ldots + \lambda_m f(v_m) \;\middle|\; \lambda_i \in \mathbb{K} \right\} \\ &= \left\{ \lambda_1 f(v_1) + \ldots + \lambda_r f(v_r) + 0 + \ldots + 0 \;\middle|\; \lambda_i \in \mathbb{K} \right\} \\ &= \mathcal{L}\left(f(v_1), \ldots, f(v_r)\right) \end{aligned}$$

Setze $w_1 := f(v_1), \ldots, w_r := f(v_r)$. Wir wissen bereits aus dem Beweis des Dimensionssatzes, dass $(w_1, \ldots, w_r)$ linear unabhängig sind. Nach dem Basisergänzungssatz wähle $w_{r+1}, \ldots, w_n \in W$, sodass $C = (w_1, \ldots, w_r, w_{r+1}, \ldots, w_n)$ eine Basis für W ist.

Für $j = 1, \ldots, r$ ist $f(v_j) = w_j$ und damit ist $(0, \ldots, 0, 1, 0, \ldots, 0)$ mit einer 1 an der j-ten Stelle die Darstellung von $f(v_j)$ in der Basis C.

Für $j = r+1, \ldots, m$ ist $f(v_j) = 0$, da diese Vektoren im Kern der Abbildung f liegen, sodass $(0, \ldots, 0, 0, 0, \ldots, 0)$ die Darstellung von $f(v_j)$ in der Basis C ist. Nach Definition gilt dann wie erwünscht

$$\mathrm{DM}_C^B(f) = \begin{pmatrix} 1 & 0 & \dots & & \dots & 0 \\ 0 & \ddots & \ddots & & & \vdots \\ \vdots & \ddots & 1 & & & \\ & & & 0 & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \dots & & \dots & 0 & 0 \end{pmatrix}$$

⊠(Satz 10.14)

Bemerkung 10.15: Wie wir im Beweis gesehen haben, können wir nicht von $B = C$ ausgehen, da B durch den Basisergänzungssatz gewählt wurde und C durch Bildwerte gegeben ist, also beide nicht hinreichend frei wählbar sind.
Dies ist nur eine Teillösung des Problems, wenn man den Anspruch hat, dass im Falle eines Endomorphismus $f : V \to V$ (d.h. $V = W$) die Basen B und C auch gleich gewählt werden sollen. Der Beweis zeigt, dass die konkrete Wahl von B notwendig war, um mittels des Kerns die unteren Nullen zu erzeugen, sowie die Wahl von C mittels der Bilder der Basis B (teilweise) notwendig war, um die Einsen im oberen Teil zu erhalten. Wenn man allerdings nicht auf die gleiche Wahl von B und C angewiesen ist, ist die Aussage des Satzes offenbar optimal.

⋆ ⋆ ⋆

Darstellende Matrizen von Endomorphismen. Wir beschäftigen uns nun mit Endomorphismen.

Definition 10.16 – Darstellende Matrix eines Endomorphismus:

Sei $f : V \to V$ ein Endomorphismus eines endlich-dimensionalen $\mathbb{K}$-Vektorraums V. Sei $B = (v_1, \dots, v_m)$ eine Basis von V. Dann definiere die darstellende Matrix von f bezüglich B als

$$\mathrm{DM}_B(f) := \mathrm{DM}_B^B(f) \in \mathrm{Mat}(m \times m, \mathbb{K}).$$

Die folgende Definition führt hilfreiche Begriffe ein.

Definition 10.17 – transponierte Matrix, symmetrische Matrix:

Sei $A = (a_{ij}) \in \mathrm{Mat}(n \times m, \mathbb{K})$ eine Matrix über dem Körper $\mathbb{K}$.

(a) Die Matrix $A^t = \left(a^t_{ij}\right) \in \mathrm{Mat}(m \times n, \mathbb{K})$, sodass für alle $i \le m$ und $j \le n$ gilt

$$a^t_{ij} = a_{ji}$$

ist die transponierte Matrix von A. Das heißt die transponierte Matrix ist eine Spiegelung der Ausgangsmatrix an der Hauptdiagonalen.

(b) Die Matrix A heißt symmetrisch, wenn

$$A^t = A$$

gilt. Offenbar sind symmetrische Matrizen quadratisch.

Beispiel 10.18: Die transponierte Matrix zu finden ist ganz einfach — es gilt:

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

Eine transponierte Matrix lässt sich natürlich auch für nicht quadratische Matrizen finden:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix}^t = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix}$$

Offensichtlich ist jede Diagonalmatrix symmetrisch. Symmetrische Matrizen und Matrizen ähnlicher Gestalt spielen eine wichtige Rolle in den Anwendungen der Linearen Algebra. Wir werden am Ende des Themas Diagonalisierung darauf eingehen.

Wir zeigen nun, dass wir bereits die bekannten Spiegelungen durch geeignete Wahl der Basen durch Basistransformation auf eine symmetrische Form bringen können.

Beispiel 10.19: Die darstellende Matrix einer Spiegelung lässt sich durch geeignete Basistransformationen auf eine einfachere Form bringen: Wähle dazu einen Vektor $u_0 = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$, der nicht auf Spiegelachse und der Achse senkrecht dazu liegt. Sei

$$u_1 \;=\; S_{\frac{\gamma}{2}} \cdot \begin{pmatrix} a \\ b \end{pmatrix} \;=\; \begin{pmatrix} \cos(\gamma) & \sin(\gamma) \\ \sin(\gamma) & -\cos(\gamma) \end{pmatrix} \cdot \begin{pmatrix} a \\ b \end{pmatrix} \;=\; \begin{pmatrix} a \cdot \cos(\gamma) + b \cdot \sin(\gamma) \\ a \cdot \sin(\gamma) - b \cdot \cos(\gamma) \end{pmatrix}$$

das Spiegelbild von u_0. Nach Wahl von u_0 ist $C = (u_0, u_1)$ linear unabhängig in $\mathbb{R}^2$ und somit eine Basis. Wenn $K = (e_1, e_2)$ die kanonische Basis von $\mathbb{R}^2$ bezeichnet, so sind die Transformationsmatrizen

$$T_K^C \quad = \quad (u_0\; u_1) \quad = \quad \begin{pmatrix} a & a \cdot \cos(\gamma) + b \cdot \sin(\gamma) \\ b & a \cdot \sin(\gamma) - b \cdot \cos(\gamma) \end{pmatrix} \qquad \text{und} \qquad T_C^K \quad = \quad (T_K^C)^{-1}.$$

Im Fall, dass $u_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ ist, das heißt wir schließen aus, dass die Spiegelachse gerade die x- oder y-Achse ist, gilt:

$$T_K^C = \begin{pmatrix} 1 & \cos(\gamma) \\ 0 & \sin(\gamma) \end{pmatrix} \qquad \text{und} \qquad T_C^K = (T_K^C)^{-1} = \frac{1}{\sin(\gamma)} \cdot \begin{pmatrix} \sin(\gamma) & -\cos(\gamma) \\ 0 & 1 \end{pmatrix}$$

Wir berechnen die darstellende Matrix $\mathrm{DM}_C(f)$. Nach dem Transformationssatz gilt:

$$\begin{aligned}
\mathrm{DM}_C(f) \quad &= \quad T_C^K \cdot \mathrm{DM}_K(f) \cdot T_K^C \\
&= \quad \begin{pmatrix} 1 & -\frac{\cos(\gamma)}{\sin(\gamma)} \\ 0 & \frac{1}{\sin(\gamma)} \end{pmatrix} \cdot \begin{pmatrix} \cos(\gamma) & \sin(\gamma) \\ \sin(\gamma) & -\cos(\gamma) \end{pmatrix} \cdot \begin{pmatrix} 1 & \cos(\gamma) \\ 0 & \sin(\gamma) \end{pmatrix} \\
&= \quad \begin{pmatrix} 1 & -\frac{\cos(\gamma)}{\sin(\gamma)} \\ 0 & \frac{1}{\sin(\gamma)} \end{pmatrix} \cdot \begin{pmatrix} \cos(\gamma) & 1 \\ \sin(\gamma) & 0 \end{pmatrix} \quad = \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}
\end{aligned}$$

Achtung!

Wir hätten $\mathrm{DM}_C(f)$ auch einfacher bestimmen können, indem wir verstehen, dass die Spiegelung die beiden Vektoren u_0 und u_1 vertauscht, das heißt es gilt:

$$f(u_0) = u_1 \text{ und } f(u_1) = u_0.$$

Insbesondere muss gelten:

$$\mathrm{DM}_C(f)\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}0\\1\end{pmatrix} \text{ und } \mathrm{DM}_C(f)\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix}$$

Damit ist wegen der Eindeutigkeit einer linearen Abbildung auf einer Basis klar, dass $\mathrm{DM}_C(f) = \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}$ ist. Diese darstellende Matrix ist symmetrisch.

Beispiel 10.20: Wir setzen unsere Betrachtungen des letzten Beispiels fort: Eine andere Darstellung einer Spiegelung erhalten wir, indem wir eine Basis aus einem Vektor v_0 der Spiegelachse und einem hierzu senkrechten Vektor v_1 bilden. Nehmen wir weiterhin an, dass beide Vektoren normiert sind, d.h. sie haben die Länge 1. (Siehe diesbezüglich auch die Bemerkung nach der Definition 14.2.) Der senkrechte Vektor lässt sich nun durch Drehung um den Winkel $\frac{\pi}{2}$ mit der Drehmatrix $D_{\frac{\pi}{2}} = \begin{pmatrix}0 & -1\\1 & 0\end{pmatrix}$ berechnen. Sei $v_0 = \begin{pmatrix}\cos\left(\frac{\gamma}{2}\right)\\\sin\left(\frac{\gamma}{2}\right)\end{pmatrix}$ und

$$v_1 = \begin{pmatrix}0 & -1\\1 & 0\end{pmatrix}\cdot\begin{pmatrix}\cos\left(\frac{\gamma}{2}\right)\\\sin\left(\frac{\gamma}{2}\right)\end{pmatrix} = \begin{pmatrix}-\sin\left(\frac{\gamma}{2}\right)\\\cos\left(\frac{\gamma}{2}\right)\end{pmatrix}.$$

Setze die Basis D an als $D = (v_0, v_1)$. Dann gilt für die Transformationsmatrizen:

$$T_K^D = (v_0 v_1) = \begin{pmatrix}\cos\left(\frac{\gamma}{2}\right) & -\sin\left(\frac{\gamma}{2}\right)\\\sin\left(\frac{\gamma}{2}\right) & \cos\left(\frac{\gamma}{2}\right)\end{pmatrix}$$

Die Transformationsmatrix T_D^K ist dann die Drehmatrix für den Winkel $\frac{-\gamma}{2}$:

$$T_D^K = (T_K^D)^{-1} = \begin{pmatrix}\cos\left(\frac{-\gamma}{2}\right) & -\sin\left(\frac{-\gamma}{2}\right)\\\sin\left(\frac{-\gamma}{2}\right) & \cos\left(\frac{-\gamma}{2}\right)\end{pmatrix} = \begin{pmatrix}\cos\left(\frac{\gamma}{2}\right) & \sin\left(\frac{\gamma}{2}\right)\\-\sin\left(\frac{\gamma}{2}\right) & \cos\left(\frac{\gamma}{2}\right)\end{pmatrix}$$

Schließlich gilt:

$$\mathrm{DM}_D(f) = T_D^K\cdot\mathrm{DM}_K(f)\cdot T_K^D = \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}$$

Achtung!

Dies hätte man auch wieder schnell durch Überlegung entscheiden können, denn die Basisvektoren werden wie folgt abgebildet:

$$f(v_0) = v_0 \text{ und } f(v_1) = -v_1.$$

Im nächsten Kapitel werden wir diese Art der Beziehung schätzen lernen, nämlich dass $f(v_0) = 1\cdot v_0$ und $f(v_1) = (-1)\cdot v_1$ gilt.

Beispiel 10.21: Wir wollen nun Eigenschaften eines Endomorphismus $f : V \to V$ untersuchen, die unabhängig von der gewählten Basis von V sind: Betrachten Sie die Situation, dass $f(v) = \lambda v$ für $\lambda \in \mathbb{R}\setminus\{0\}$ gilt, wobei V ein n-dimensionaler $\mathbb{R}$-Vektorraum ist. Wir werden zeigen, dass dann auch $\mathrm{DM}_B(f) \cdot K_B^{\mathbb{R}^n}(v) = \lambda \cdot K_B^{\mathbb{R}^n}(v)$ für eine beliebige Basis B gilt. Hierbei ist $K_B^{\mathbb{R}^n}(v)$ die Darstellung von v bezüglich der Basis B (vgl. Bemerkung 10.6). Somit ist in der Gleichung das Skalar λ unabhängig von der Basiswahl im Vektorraum V.

Nehmen wir daher an, es gelte $\mathrm{DM}_B(f) \cdot K_B^{\mathbb{R}^n}(v) = \lambda \cdot K_B^{\mathbb{R}^n}(v)$. Wir zeigen, dass dann auch für eine beliebige andere Basis C gilt: $\mathrm{DM}_C(f) \cdot K_C^{\mathbb{R}^n}(v) = \lambda \cdot K_C^{\mathbb{R}^n}(v)$.

Dies zeigen wir nur mithilfe unserer Kentnisse über Basiswechsel:

$$\begin{aligned}
\mathrm{DM}_C(f) \cdot K_B^{\mathbb{R}^n}(v) &= \underbrace{T_C^B \cdot \mathrm{DM}_B(f) \cdot T_B^C}_{\mathrm{DM}_C(f)} \cdot \underbrace{T_C^B \cdot K_B^{\mathbb{R}^n}(v)}_{K_C^{\mathbb{R}^n}(v)} \overset{T_B^C \cdot T_C^B = 1}{=} T_C^B \cdot \underbrace{\mathrm{DM}_B(f) \cdot K_B^{\mathbb{R}^n}(v)}_{=\lambda \cdot K_B^{\mathbb{R}^n}(v)} \\
&= \lambda \cdot \underbrace{T_C^B \cdot K_B^{\mathbb{R}^n}(v)}_{=K_C^{\mathbb{R}^n}(v)} = \lambda \cdot K_C^{\mathbb{R}^n}(v)
\end{aligned}$$

Beispiel 10.22: Sei f ein Endomorphismus $f : V \to V$, K die kanonische Basis des $\mathbb{R}^n$ und $\mathrm{DM}_B(f)$ symmetrisch für eine Basis B von V (V sei ein n-dimensionaler $\mathbb{R}$-Vektorraum). Wir betrachten die folgende Frage: Ist dann $\mathrm{DM}_C(f)$ für eine beliebige Basis C von V auch symmetrisch?

Im Allgemeinen ist diese Aussage nicht richtig, aber für Orthonormalbasen (ONB) können wir die Aussage beweisen. Dieser Begriff ist neu, aber wir werden uns bei der Hauptachsentransformation noch ausgiebig mit Orthonormalbasen beschäftigen. An dieser Stelle soll Ihnen folgende Eigenschaft einer Orthonormalbasis B reichen:

$$\left(T_K^B\right)^{-1} = \left(T_K^B\right)^t.$$

Wie Sie wissen, gilt die Gleichung: $T_B^C = T_B^K T_K^C$. Wenn man jetzt noch die Matrixrechenregeln $(AB)^t = B^t A^t$ und $(AB)^{-1} = B^{-1}A^{-1}$ für quadratische Matrizen kennt, dann erhält man:

$$\left(T_B^C\right)^{-1} = \left(T_B^K T_K^C\right)^{-1} = \left(T_K^C\right)^{-1}\left(T_B^K\right)^{-1} = \left(T_K^C\right)^t\left(T_B^K\right)^t = \left(T_B^K T_K^C\right)^t$$

Wir beweisen jetzt die Symmetrie der Matrix $\mathrm{DM}_C(f)$ direkt:

$$\begin{aligned}
\mathrm{DM}_C(f)^t &= \left(\left(T_B^C\right)^{-1} \cdot \mathrm{DM}_B(f) \cdot T_B^C\right)^t &&= \left(T_B^C\right)^t \cdot \left(\left(T_B^C\right)^{-1} \cdot \mathrm{DM}_B(f)\right)^t \\
&= \left(T_B^C\right)^t \cdot \left(\mathrm{DM}_B(f)\right)^t \cdot \left(\left(T_B^C\right)^{-1}\right)^t &&= \left(T_B^K T_K^C\right)^t \cdot \mathrm{DM}_B(f) \cdot \left(T_B^K \cdot T_K^C\right) \\
&= \left(T_B^C\right)^{-1} \cdot \mathrm{DM}_B(f) \cdot \left(T_B^C\right) &&= \mathrm{DM}_C(f)
\end{aligned}$$

⋆ ⋆ ⋆

Aufgabe 10.23: Sei f ein Endomorphismus $f : V \to V$, V ein n-dimensionaler $\mathbb{R}$-Vektorraum und B eine beliebige Basis von V. Dann definieren wir:

$$\mathrm{Spur}(A) := \sum_{i=1}^{n} a_{ii}$$

für eine $(n \times n)$-Matrix $A = \left(a_{ij}\right)_{i;j=1,\ldots,n}$ und

$$\operatorname{Spur}(f) := \operatorname{Spur}\left(\mathrm{DM}_B(f)\right) = \sum_{i=1}^{n} \left(\mathrm{DM}_B(f)\right)_{ii}$$

Zeigen Sie, dass der letzte Ausdruck wohldefiniert (unabhängig von der Basis B) ist!

<u>Anleitung</u>: Sie müssen zeigen, dass

$$\operatorname{Spur}\left(\mathrm{DM}_B(f)\right) = \operatorname{Spur}\left(\mathrm{DM}_C(f)\right)$$

für jede andere Basis C von V gilt. Beweisen Sie dazu die Gleichung

$$\operatorname{Spur}(AB) = \operatorname{Spur}(BA)$$

und dann verwenden Sie

$$\mathrm{DM}_C(f) = \left(T_B^C\right)^{-1} \cdot \mathrm{DM}_B(f) \cdot T_B^C.$$

11. Eigenwerte und Eigenvektoren

Es ist offenbar vorteilhaft die darstellende Matrix eines Endomorphismus f in Diagonalform zu bringen:

$$\mathrm{DM}_C(f) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix} \in \mathrm{Mat}(m \times m, \mathbb{K}),$$

da wir insbesondere mit einer solchen Darstellung viel einfacher rechnen können: Wir sparen bereits bei der Darstellung Speicherplatz und beim Umgang Rechenzeit.

Man sieht leicht, dass für die Vektoren v_i der Basis $C = (v_1, \dots, v_m)$ gilt

$$f(v_i) = \lambda_i \cdot v_i \qquad \lambda_i \in \mathbb{K} \tag{4}$$

Grundlegende Begriffe. Das Ziel des Diagnonalisierens können wir folgendermaßen mit Hilfe der Wahl einer geeigneten Basis formalisieren:

Definition 11.1 – diagonalisierbar:

Sei V ein $\mathbb{K}$-Vektorraum und $f : V \to V$ ein Endomorphismus. Dann ist f diagonalisierbar, wenn es eine Basis C von V gibt, so dass $\mathrm{DM}_C(f)$ eine Diagonalmatrix ist.

Bereits durch die obige Einleitung ist offenbar die Gleichung (4) von besonderer Bedeutung.

Definition 11.2 – Eigenwert, Eigenvektor:

Sei V ein $\mathbb{K}$-Vektorraum und $f : V \to V$ ein Endomorphismus.

(a) Ein Skalar $\lambda \in \mathbb{K}$ heißt Eigenwert von f, wenn es ein $v \in V \setminus \{0\}$ gibt mit $f(v) = \lambda \cdot v$.

(b) Ein Vektor $v \in V$ heißt Eigenvektor von f zum Eigenwert $\lambda \in \mathbb{K}$, wenn $v \neq 0$ und $f(v) = \lambda \cdot v$.

Damit erhalten wir sofort den folgenden

Satz 11.3:

Sei V ein endlich-dimensionaler $\mathbb{K}$-Vektorraum und $f : V \to V$ ein Endomorphismus. Dann ist f genau dann diagonalisierbar, wenn es eine Basis C von V gibt, deren Vektoren Eigenvektoren von f sind.

Beweis:

„$\Rightarrow$“ Sei f diagonalisierbar. Wähle eine Basis $C = (v_1, \dots, v_m)$ von V, so dass die darstellende Matrix von f von der Gestalt

$$\mathrm{DM}_C(f) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{pmatrix}$$

ist. Für alle $i \leq m$ gilt also $f(v_i) = \lambda_i \cdot v_i$. Somit ist jedes v_i ein Eigenvektor von f.

„$\Leftarrow$“ Sei also $C = (v_1, \ldots, v_m)$ eine Basis von V, die aus Eigenvektoren besteht. Wähle $\lambda_1, \ldots, \lambda_m$ aus $\mathbb{K}$, so dass für alle $i \leq m$ gilt:

$$f(v_i) = \lambda_i \cdot v_i$$

Damit haben wir wie gewünscht die Diagonalform

$$\mathrm{DM}_C(f) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{pmatrix}$$

$\boxtimes$

Eigenwerte und Eigenvektoren an bekannten Beispielen. Wir betrachten bekannte Endomorphismen des $\mathbb{R}^2$ und deren Bezüge zu Eigenwerten und -vektoren.

Beispiel 11.4: Wie bereits im Beispiel 10.20 gesehen, lässt sich eine Spiegelung des $\mathbb{R}^2$ in eine Matrix der Form $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ diagonalisieren.

Beispiel 11.5: Eine Drehung des $\mathbb{R}^2$ um den Winkel γ besitzt nur dann einen Eigenvektor v, falls der Drehwinkel ein ganzzahliges Vielfaches von π (also von $180°$) ist. In diesem Fall ist f diagonalisierbar und die entsprechende Diagonalmatrix ist die Einheitsmatrix $E_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ oder ihr Negatives $-E_2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$.

Beispiel 11.6: Betrachte schließlich eine Scherung h des $\mathbb{R}^2$ mit darstellender Matrix

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

Der Endomorphismus h besitzt einen Eigenvektor $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ mit zugehörigem Eigenwert 1, das heißt es gilt:

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

Alle weiteren Eigenvektoren $\begin{pmatrix} a \\ b \end{pmatrix}$ von h sind skalare Vielfache von $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, denn aus

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a+b \\ b \end{pmatrix} \stackrel{!}{=} \lambda \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \lambda a \\ \lambda b \end{pmatrix}$$

folgt:

(i) $a + b = \lambda \cdot a$ und
(ii) $b = \lambda b$

Hierfür betrachten wir folgende Fallunterscheidung: Wir wissen, dass Eigenvektoren ungleich dem Nullvektor sind und unterscheiden daher formal zwei Fälle:

- Fall $b \neq 0$. Nach (ii) muss wegen $b = \lambda b$ also $\lambda = 1$ sein. Nach (i) gilt: $a + b = a$. Dies stimmt nur für $b = 0$ und dies ist nun aber ein Widerspruch zur Voraussetzung des Falls und daher kann dieser nicht eintreten. Es bleibt also der

– Fall $b = 0$: Dann muss $a \neq 0$ gelten. Aus der ersten Formel (i) folgt somit: $a = \lambda a \Rightarrow \lambda = 1$

Insgesamt gilt also: $\lambda = 1$ und $b = 0$.

Damit spannen die Eigenvektoren von h nur den eindimensionalen Untervektorraum

$$\left\{ \begin{pmatrix} a \\ 0 \end{pmatrix} \mid a \in \mathbb{R} \right\}$$

auf und es gibt keine Basis von $\mathbb{R}^2$, die nur aus Eigenvektoren von h besteht. Nach Satz 11.3 gilt also: Die Scherung h ist nicht diagonalisierbar.

Beispiel 11.7: Betrachten wir die Matrix $A := \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix}$. Diese Matrix hat die Eigenwerte $\lambda_1 = 2$ und $\lambda_2 = 3$. Wir berechnen die dazugehörigen Eigenvektoren. Dazu müssen wir das Lineare Gleichungssystem der Gestalt $A \cdot v = \lambda_1 \cdot v$ lösen. Dies ist etwa äquivalent zu

$$0 = Av - \lambda_1 v = A \cdot v - \lambda_1 \cdot E_2 \cdot v = (A - \lambda_1 \cdot E_2) \cdot v.$$

Somit haben wir das Problem auf das Lösen eines homogenen, linearen Gleichungssystems (bzw. des Berechnens des Kerns einer geeigneten linearen Abbildung) reduziert.

$$A - \lambda_1 \cdot E_2 = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} - 2 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 2 & -1 \end{pmatrix}$$

Anwendung des Gaußschen Algorithmus liefert:

$$\begin{pmatrix} 2 & -1 \\ 2 & -1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{pmatrix}$$

Das heißt b ist frei wählbar und $a - \frac{1}{2}b = 0$, also $a = \frac{1}{2}b$. Somit ist die Lösungsmenge die Folgende:

$$\text{Lös}\,(A - \lambda_1 E_2, 0) = \left\{ \begin{pmatrix} \frac{1}{2}b \\ b \end{pmatrix} \mid b \in \mathbb{R} \right\}$$

und eine Basis ist etwa für $b = 2$ gegeben durch den Vektor $v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

Die Probe ergibt tatsächlich:

$$A \cdot v_1 = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \lambda_1 \cdot v_1$$

Analoge Rechnungen ergeben für $\lambda_2 = 3$:

$$0 = A \cdot v - \lambda_2 \cdot v = (A - \lambda_2 E_2)\, v :$$

$$A - \lambda_2 \cdot E_2 = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} - 3 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} - \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 2 & -2 \end{pmatrix}$$

Anwendung des Gaußschen Algorithmus liefert:

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

das heißt b ist frei wählbar und es gilt $a - b = 0$, also $a = b$. Die Lösungsmenge ist dann

$$\text{Lös}\,(A - \lambda_2 E_2, 0) = \left\{ \begin{pmatrix} b \\ b \end{pmatrix} \mid b \in \mathbb{R} \right\}$$

und eine Basis ist etwa für $b = 1$ gegeben durch den Vektor $v_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Eigenräume zu Eigenwerten. Wir haben nun verstanden, was Eigenvektoren sind und können schließlich die Ideen aus dem letzten Beispiel vertiefen. Wir interessieren uns daher als Nächstes, wie man die Eigenvektoren berechnen kann – vorausgesetzt, dass wir bereits Eigenwerte gegeben haben. Dabei hilft die folgende Charakterisierung enorm:

Satz 11.8:

Sei V ein $\mathbb{K}$-Vektorraum und $f: V \to V$ ein Endomorphismus. Dann gilt für einen Vektor $v \in V \setminus \{0\}$ und $\lambda \in \mathbb{K}$:

$$v \text{ ist genau dann Eigenvektor von } f \text{ zum Eigenwert } \lambda,$$
$$\text{wenn } v \in \operatorname{Kern}(f - \lambda \cdot \mathrm{Id}_V).$$

Beweis: v ist Eigenvektor von f zum Eigenwert λ

$$\begin{aligned}
\Leftrightarrow\quad & f(v) = \lambda \cdot v \\
\Leftrightarrow\quad & 0 = f(v) - \lambda \cdot v = f(v) - (\lambda \cdot \mathrm{Id}_V)(v) = (f - \lambda \cdot \mathrm{Id}_V)(v) \\
\Leftrightarrow\quad & v \in \operatorname{Kern}(f - \lambda \cdot \mathrm{Id}_V)
\end{aligned}$$

⊠

Beachte die Analogie im Beweis zu der Umformung im letzten Beispiel 11.7.

Offenbar sind die obigen Kerne des verschobenen Endomorphismus A von Interesse, so dass wir diese gesondert bezeichnen:

Definition 11.9 – Eigenraum, geometrische Vielfachheit:

Sei V ein $\mathbb{K}$-Vektorraum und $f: V \to V$ ein Endomorphismus mit Eigenwert $\lambda \in \mathbb{K}$. Dann heißt

$$\mathrm{E}(\lambda) := \operatorname{Kern}(f - \lambda \cdot \mathrm{Id}_V)$$

der Eigenraum von f zum Eigenwert λ.
Die geometrische Vielfachheit des Eigenwerts λ ist die Dimension von $\mathrm{E}(\lambda)$.

Lineare Unabhängigkeit von Eigenvektoren. Unser Ziel ist es also, die oben genannten Eigenräume zu bestimmen, da sich hier die Eigenvektoren befinden. Wünschenswert ist allerdings nach dem Satz 11.3 letztendlich eine Gesamtbasis des Ausgangsraums V, bestehend nur aus Eigenvektoren. Auf dem Wege zu diesem Ziel hilft uns der folgende Satz, der besagt, dass Eigenvektoren zu paarweise verschiedenen Eigenwerten immer linear unabhängig sind. Dies ist ein erster Schritt in die richtige Richtung.

Satz 11.10:

Sei V ein $\mathbb{K}$-Vektorraum und $f: V \to V$ ein Endomorphismus. Seien $v_1, \ldots, v_r$ Eigenvektoren zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_r$. Dann ist $(v_1, \ldots, v_r)$ linear unabhängig.

Beweis: Durch Induktion über r:

Induktionsanfang

Die Behauptung gilt für $r = 0$, weil das 0-Tupel nach Definition linear unabhängig ist.

Induktionsschritt

Voraussetzung: Die Aussage gilt für r.
Behauptung: Die Aussage gilt auch für $r+1$.

Angenommen $v_1, \ldots, v_r, v_{r+1}$ sind Eigenvektoren zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_r, \lambda_{r+1}$.

Ziel: $v_1, \ldots, v_r, v_{r+1}$ sind linear unabhängig.
Seien dazu (nach Definition) $\alpha_1, \ldots, \alpha_r, \alpha_{r+1} \in \mathbb{K}$ mit

$$\alpha_1 v_1 + \ldots + \alpha_r v_r + \alpha_{r+1} v_{r+1} = 0 \tag{5}$$

gegeben.

Noch zu zeigen: $\alpha_1 = \ldots = \alpha_r = \alpha_{r+1} = 0$
Wende dazu f auf diese Gleichung an. Da $v_1, \ldots, v_r, v_{r+1}$ Eigenvektoren sind, ergibt sich

$$\begin{aligned} f(\alpha_1 v_1) + \ldots + f(\alpha_r v_r) + f(\alpha_{r+1} v_{r+1}) &= f(0) \\ \Leftrightarrow \alpha_1 f(v_1) + \ldots + \alpha_r f(v_r) + \alpha_{r+1} f(v_{r+1}) &= 0 \\ \Leftrightarrow \alpha_1 \lambda_1 v_1 + \ldots + \alpha_r \lambda_r v_r + \alpha_{r+1} \lambda_{r+1} v_{r+1} &= 0 \end{aligned} \tag{6}$$

Multiplikation der ersten Gleichung (5) mit λ_{r+1} ergibt:

$$\alpha_1 \lambda_{r+1} v_1 + \ldots + \alpha_r \lambda_{r+1} v_r + \alpha_{r+1} \lambda_{r+1} v_{r+1} = 0 \tag{7}$$

Subtraktion von (7) von (6) ergibt:

$$\alpha_1 \left(\lambda_1 - \lambda_{r+1}\right) v_1 + \ldots + \alpha_r \left(\lambda_r - \lambda_{r+1}\right) v_r = 0$$

Nach Induktionsvoraussetzung sind die Vektoren $v_1, \ldots, v_r$ linear unabhängig. Also gilt:

$$\alpha_1(\lambda_1 - \lambda_{r+1}) = \ldots = \alpha_r(\lambda_r - \lambda_{r+1}) = 0$$

Da die Eigenwerte $\lambda_1, \ldots, \lambda_r, \lambda_{r+1}$ paarweise verschieden sind, gilt

$$\lambda_1 - \lambda_{r+1} \neq 0, \ldots, \lambda_r - \lambda_{r+1} \neq 0.$$

Somit folgt: $\alpha_1 = \ldots = \alpha_r = 0$. Da v_{r+1} ein Eigenvektor ist, gilt $v_{r+1} \neq 0$ und somit wie gewünscht nach (5) auch $\alpha_{r+1} = 0$. Damit ist alles per Induktion bewiesen.

⊠

Aus dem letzten Satz folgt sofort:

Satz 11.11:

Sei V ein m-dimensionaler $\mathbb{K}$-Vektorraum und $f : V \to V$ ein Endomorphismus. Dann besitzt f höchstens m verschiedene Eigenwerte.

Beweis: Bei $m+1$ vielen Eigenwerten $\lambda_1, \ldots, \lambda_{m+1}$ wähle jeweils dazugehörige Eigenvektoren $v_1, \ldots, v_{m+1}$.

Nach dem letzten Satz 11.10 ist $v_1, \ldots, v_{m+1}$ eine linear unabhängige Menge von Vektoren im m-dimensionalen Vektorraum. Dies ist ein Widerspruch. ☒

Wir können sogar die Aussage über die lineare Unabhängigkeit wie folgt verstärken:

Satz 11.12:

Sei V ein m-dimensionaler $\mathbb{K}$-Vektorraum und $f : V \to V$ ein Endomorphismus. Seien $\lambda_1, \ldots, \lambda_r$ paarweise verschiedene Eigenwerte von f mit geometrischen Vielfachheiten $m_1, \ldots, m_r$. Für $i \leq r$ sei $(v_1^i, \ldots, v_{m_i}^i)$ eine Basis des Eigenraums $\mathrm{E}(\lambda_i)$. Dann gilt:

(a) Das System

$$\left(\underbrace{v_1^1, \ldots, v_{m_1}^1}_{\text{Basis } \mathrm{E}(\lambda_1)}, \underbrace{v_1^2, \ldots, v_{m_2}^2}_{\text{Basis } \mathrm{E}(\lambda_2)}, \ldots, \underbrace{v_1^r, \ldots, v_{m_r}^r}_{\text{Basis } \mathrm{E}(\lambda_r)}\right)$$

ist linear unabhängig.

(b) Für die Summe der geometrischen Vielfachheiten gilt:

$$m_1 + \ldots + m_r \leq m$$

(c) Der Endomorphismus f ist genau dann diagonalisierbar, wenn

$$m_1 + \ldots + m_r = m.$$

Beweis:

zu (a). Wir betrachten eine Darstellung des Nullvektors und hierfür konkret Skalare $\alpha_j^i \in \mathbb{K}$, so dass

$$\alpha_1^1 v_1^1 + \ldots + \alpha_{m_1}^1 v_{m_1}^1 + \ldots + \alpha_1^r v_1^r + \ldots + \alpha_{m_r}^r v_{m_r}^r = 0$$

Wir sehen, dass die Teilsummen

$$\alpha_1^i v_1^i + \ldots + \alpha_{m_i}^i v_{m_i}^i$$

selbst Eigenvektoren zum Eigenwert λ_i oder eben der Nullvektor sind. Wenn derartige Teilsummen aber ungleich Null wären, so wären sie linear abhängige Eigenvektoren zu verschiedenen Eigenwerten, im Widerspruch zu Satz 11.10. Also gilt für alle $i \leq r$, dass

$$\alpha_1^i = \ldots = \alpha_{m_i}^i = 0.$$

Damit sind aber wie gewünscht alle Skalare $\alpha_j^i = 0$ und somit ist das System

$$\left(v_1^1, \ldots, v_{m_1}^1, v_1^2, \ldots, v_{m_2}^2, \ldots, v_1^r, \ldots, v_{m_r}^r\right)$$

linear unabhängig.

zu (b). Da V die Dimension m besitzt, hat dieses System als Folge linear unabhängiger Vektoren in V höchstens m Elemente und wir erhalten:

$$m_1 + \ldots + m_r \leq m.$$

zu (c). Wir zeigen beide Richtungen der geforderten Äquivalenz. Angenommen, dass $m_1 + \ldots + m_r = m$ gilt, dann ist

$$\left(v_1^1, \ldots, v_{m_1}^1, v_1^2, \ldots, v_{m_2}^2, \ldots, v_1^r, \ldots, v_{m_r}^r\right)$$

sogar eine Basis von V. Diese Basis besteht aus Eigenvektoren und nach Satz 11.3 ist f diagonalisierbar.

Umgekehrt sei nun f diagonalisierbar. Nach Satz 11.3 wähle eine Basis $C = (v_1, \ldots, v_m)$ aus Eigenvektoren. Für $i \leq r$ sei n_i die Anzahl der Vektoren aus C, deren Eigenwert λ_i ist. Dann ist $n_i \leq m_i$ und

$$m = n_1 + \ldots + n_r \leq m_1 + \ldots + m_r \leq n$$

und damit ist

$$m_1 + \ldots + m_r = m.$$

⊠

Vorläufige Strategie des Diagonalisierens. Damit zeichnet sich eine Strategie für das Diagonalisieren eines Endomorphismus $f : V \to V$ ab – wir kommen am Ende des folgenden Kapitels noch einmal im Detail darauf zurück:

Man bestimme alle Eigenwerte von f, das heißt alle $\lambda_i \in \mathbb{K}$, für die gilt:

$$\operatorname{Kern}(f - \lambda_i \cdot \operatorname{Id}_V) \neq \{0\}$$

Dann bestimme man eine Basis des Eigenraums $E(\lambda_i)$ durch Lösen des linearen Gleichungssystems

$$(f - \lambda_i \cdot \operatorname{Id}_V)(x) = 0.$$

Wenn die Summe der Dimensionen der Eigenräume gleich der Dimension von V ist, erhalten wir eine Diagonalisierung von f und wir erhalten eine Diagonaldarstellung von f in der Form

$$\begin{pmatrix} \lambda_1 & & & & & & \\ & \ddots & & & & & \\ & & \lambda_1 & & & & \\ & & & \ddots & & & \\ & & & & \lambda_r & & \\ & & & & & \ddots & \\ & & & & & & \lambda_r \end{pmatrix},$$

in der jeder Eigenwert von f entsprechend seiner geometrischen Vielfachheit auftritt.

Sie sehen insbesondere, dass wir fast am Ziel sind: Wir können bis auf eine kleine Blackbox am Anfang des Algorithmus bereits Matrizen diagonalisieren (sofern möglich). Das Einzige, was wir noch lernen müssen, ist Eigenwerte zu bestimmen.

Diese Fähigkeit werden wir im folgenden Kapitel erlernen.

⋆ ⋆ ⋆

Beispiel 11.13: Wir wollen jetzt die Menge der diagonalisierbaren (2×2)-Matrizen über $\mathbb{R}$ untersuchen:

Sei $f : V \to V$ gegeben durch $\operatorname{DM}_B(f) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ für eine Basis B von V. Man kann zeigen, dass die

Menge der diagonalisierbaren Endomorphismen f entweder zwei unterschiedliche Eigenwerte besitzt oder ein Vielfaches der Einheitsmatrix ist.

Um dies zu beweisen nehmen wir an, dass f diagonalisierbar ist und es genau einen Eigenwert hat. Weiterhin starten wir mit unserem gewohnten Ansatz:

$$\begin{aligned} \mathrm{DM}_B(f)\cdot K_B^{\mathbb{R}^2}(v) &= \lambda\cdot K_B^{\mathbb{R}^2}(v) &\Leftrightarrow\quad DM_B(f)\cdot K_B^{\mathbb{R}^2}(v)-\lambda\cdot K_B^{\mathbb{R}^2}(v) &= 0\\ &&\Leftrightarrow\quad \begin{pmatrix} a-\lambda & b\\ c & d-\lambda\end{pmatrix}\cdot K_B^{\mathbb{R}^2}(v) &= 0\end{aligned}$$

Damit diese Gleichung für irgendein $v\neq 0$ erfüllt ist, darf die Matrix $\begin{pmatrix} a-\lambda & b\\ c & d-\lambda\end{pmatrix}$ nicht invertierbar sein (sonst würde man $v=0$ ausrechnen)! Sie wissen bereits, dass dies genau dann erfüllt ist wenn $(a-\lambda)(d-\lambda)-bc=0$.

Dann gibt es genau ein λ, welches die Gleichung $(a-\lambda)(d-\lambda)-bc=0$ erfüllt, da jede weitere Lösung ebenfalls ein Eigenwert wäre und wir nach Voraussetzung nur einen haben.

Somit folgt aus der Gleichung

$$ad+\lambda^2-\lambda(a+d)-bc = 0$$

mittels der Lösungsformel sofort

$$\lambda_{1,2} = \frac{a+d}{2}\pm\sqrt{\frac{(a+d)^2}{4}-ad+bc}$$

und wegen der Eindeutigkeit muss nun $\lambda_1=\lambda_2$, also

$$\sqrt{\frac{(a+d)^2}{4}-ad+bc} = 0$$

gelten.

Als Nächstes können wir sukzessiv alle Koeffizienten berechnen. Es gilt:

$$0 = \frac{(a+d)^2}{4}-ad+bc = \frac{1}{4}(a-d)^2+bc$$

und somit $a=\pm 2\sqrt{-bc}+d$. Schließlich daher $\lambda=a\pm\sqrt{bc}$ und wegen der Eindeutigkeit von λ erhält man $\sqrt{bc}=0$ und folglich $\lambda=d$. Also insgesamt:

$$\mathrm{DM}_B(f) = \begin{pmatrix} d & 0\\ 0 & d\end{pmatrix} = d\begin{pmatrix} 1 & 0\\ 0 & 1\end{pmatrix}$$

Was wir dazu wissen müssen ist zunächst einmal:

Beispiel 11.14: Unter dem Annulator eines Endomorphismus $f:V\to V$ versteht man die Menge aller Polynome $p(x)=a_nx^n+\ldots+a_1x^1+a_0$ für welche gilt:

$$p(f)\,v := a_nf^n(v)+\ldots+a_1f^1(v)+a_0v=0$$

Sei f ein Endomorphismus $f:V\to V$, V ein 2-dimensionaler $\mathbb{R}$-Vektorraum mit $\mathrm{DM}_B(f)=\begin{pmatrix} a & b\\ c & d\end{pmatrix}$ für B eine beliebige Basis von V. Mann kann zeigen, dass

$$p_f(\lambda) := (a-\lambda)(d-\lambda)-bc$$

ein Element des Annulators ist.

Überlegen Sie sich, dass $p_f(\lambda)$ wohldefiniert, also unabhängig von der Basis B ist.

Wir werden dies zuerst für den Fall, dass f diagonalisierbar ist beweisen:
Angenommen f ist diagonalisierbar. Das bedeutet, dass $v = c_1v_1 + c_2v_2,\ \forall v \in V$ und Eigenvektoren v_1, v_2. Wir beweisen jetzt $p_f(f)\,v = 0$ direkt:

$$\begin{aligned}
p_f(f)\,v &= p_f(f)\,(c_1v_1 + c_2v_2)\\
&= c_1p_f(f)\,v_1 + c_2p_f(f)\,v_2\\
&= c_1\left(f^2(v_1) - (a+d)\,f(v_1) + (ad-bc)\,v_1\right)\\
&\qquad + c_2\left(f^2(v_2) - (a+d)\,f(v_2) + (ad-bc)\,v_2\right)\\
&= c_1\underbrace{\left(\lambda_1^2 - (a+d)\,\lambda_1 + ad - bc\right)}_{\star}v_1 + c_2\underbrace{\left(\lambda_2^2 - (a+d)\,\lambda_2 + ad - bc\right)}_{\star\star}v_2\\
&= 0
\end{aligned}$$

weil $\star = \star\star = 0$. Wir haben nämlich im vorherigen Beispiel gesehen, dass aus λ_i ist ein EW $0 = (a-\lambda_i)(d-\lambda_i) - bc = \lambda_i^2 - (a+d)\lambda_i + ad - bc$ folgt.

$\star \quad \star \quad \star$

Aufgabe 11.15: Zeigen Sie mithilfe des obigen Beispiels: Sei f ein Endomorphismus $f : V \to V$ und V ein 2-dimensionaler $\mathbb{R}$-Vektorraum mit $\mathrm{DM}_B(f) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ für B eine beliebige Basis von V. Angenommen f ist nicht diagonalisierbar, dann gilt: $p_f(f)\,v = 0$ für alle $v \in V$.

Anleitung. Man beweise: Es existiert v mit $\{v, f(v)\} := \{b_1, b_2\}$ sind linear unabhängig, also auch eine Basis von V. Man erhält, dass $f(b_1) = b_2$ und $f(b_2) = c_1b_1 + c_2b_2$ und für die darstellende Matrix gilt:

$$\mathrm{DM}_{b_2}^{b_1}(f) = \begin{pmatrix} 0 & c_1 \\ 1 & c_2 \end{pmatrix}.$$

Als Nächstes müssen Sie zeigen, dass $p_f(f)$ unabhängig von der Basiswahl ist. Das bedeutet:

$$A := (a)_{i,j=1,2} = \left(T_B^C\right)^{-1}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\left(T_B^C\right) \quad \text{für } B,C \text{ Basen von V}$$

dann ist:

$$p_f(\lambda) = (a-\lambda)(d-\lambda) - bc \quad = \quad (a_{11}-\lambda)(a_{22}-\lambda) - a_{12}a_{21}.$$

Jetzt zeigt man nur noch $(0-f)(c_2-f) - c_1 = 0$ und erhält wegen der Basisinvarianz von $p_f(\lambda)$, dass $p_f(f) = (0-f)(c_2-f) - c_1 = 0$

Aufgabe 11.16: Sei $f : V \to V$ ein Endomorphismus auf dem $\mathbb{K}$-Vektorraum V. Angenommen für alle $v \in V$ gilt: $f^4(v) + v = 0$. Welche Aussagen bezüglich der Diagonalisierbarkeit und der Eigenwerte von f lassen sich für $\mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{Z}_2$ feststellen?

Aufgabe 11.17: Sei f ein Endomorphismus $f : V \to V$ und V ein zweidimensionaler $\mathbb{C}$-Vektorraum. Es soll weiterhin für $v \in V$ und ein festes $k \in \mathbb{N}$ gelten $f^k(v) = 0$. Man zeige, dass: $\mathrm{Spur}(f) = 0$.

12. Determinanten von Matrizen

Ziel: Bestimmung des noch fehlenden Gliedes in der Algorithmuskette zum Diagonalisieren von Matrizen — das heißt Klärung der Frage, wie Eigenwerte bestimmt werden können.

Wir motivieren die Betrachtungen durch ein einfaches Beispiel in der Matrizen-Dimension (2×2). Wir haben bereits gesehen, dass eine (2×2)-Matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ genau dann invertierbar ist, wenn gilt

$$ad - bc \neq 0.$$

Wir benutzen nun dieses Kriterium in der folgenden einfachen Umformung: Sei also

$$A = \begin{pmatrix} e & f \\ g & h \end{pmatrix} \in \mathrm{Mat}(2 \times 2, \mathbb{K})$$

die darstellende Matrix eines Endomorphismus f. Bei der Suche nach Eigenwerten sind die Abbildungen $A - \lambda \cdot \mathrm{Id}_{\mathbb{K}^2}$ für $\lambda \in \mathbb{K}$ zu studieren. Nach der Dimensionsformel gilt:

$$\dim(\mathrm{Kern}(A - \lambda \cdot \mathrm{Id}_{\mathbb{K}^2})) + \mathrm{Rg}(A - \lambda \cdot \mathrm{Id}_{\mathbb{K}^2}) = 2$$

Damit gilt:

$$\begin{aligned}
&\dim\big(\mathrm{Kern}\,(A - \lambda \cdot \mathrm{Id}_{\mathbb{K}^2})\big) > 0 \\
\iff\quad & \mathrm{Rg}\,(A - \lambda \cdot \mathrm{Id}_{\mathbb{K}^2}) < 2, \\
\iff\quad & A - \lambda \cdot \mathrm{Id}_{\mathbb{K}^2} \text{ ist nicht invertierbar}, \\
\iff\quad & \begin{pmatrix} e & f \\ g & h \end{pmatrix} - \lambda \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e - \lambda & f \\ g & h - \lambda \end{pmatrix} \text{ ist nicht invertierbar} \\
\iff\quad & p_A\,(\lambda) := (e - \lambda)\,(h - \lambda) - gf = 0.
\end{aligned}$$

Das heißt, die Eigenwerte des durch die Matrix A dargestellten Endomorphismus sind die Nullstellen dieses quadratischen Polynoms $p_A\,(\lambda)$, welches als so genanntes charakteristisches Polynom bezeichnet wird. (Beachte, $p_A\,(\lambda)$ ist in der Tat ein Polynom in der Unbekannten λ bei gegebenen (Parametern oder Zahlen) e, f, g, h.)

Beispiel 12.1: Im Fall der Spiegelung

$$A = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ \sin(\varphi) & -\cos(\varphi) \end{pmatrix}$$

lautet das charakteristische Polynom

$$\begin{aligned}
p_A\,(\lambda) &= \big(\cos(\varphi) - \lambda\big)\big(-\cos(\varphi) - \lambda\big) - \sin^2(\varphi) &&= -\cos^2(\varphi) + \lambda^2 - \sin^2(\varphi) \\
&= \lambda^2 - \Big(\sin^2(\varphi) + \cos^2(\varphi)\Big) &&= \lambda^2 - 1 \\
&= (\lambda + 1)\,(\lambda - 1)
\end{aligned}$$

Damit hat das Polynom $p\,(\lambda)$ die Nullstellen 1 und -1. Da die Dimensionen der Eigenräume $\mathrm{E}(1)$ und $\mathrm{E}(-1)$ jeweils mindestens 1 sind, ist die geometrische Vielfachheit der Eigenräume jeweils gleich (klar!). Da $1 + 1 = 2$ ist, ist die Matrix A nach Satz 11.12 diagonalisierbar, mit der Diagonalform

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

Dies hatten wir bereits gesehen.

⋆ ⋆ ⋆

Determinanten von Matrizen. In der obigen Umformung hatten wir das charakteristische Polynom mithilfe des Invertierkriteriums $ad - bc \neq 0$ angesetzt. Der Term $ad - bc$ wird sich als Determinante der Matrix herausstellen. Ziel ist es die Determinante und somit das charakteristische Polynom für beliebige Matrizen zu definieren. Dies werden wir in diesem und dem folgenden Kapitel umsetzen.

Definition 12.2 – Untermatrix zur Bestimmung der Determinante:

Sei $A = \left(a_{ij}\right) \in \operatorname{Mat}\left((m+1) \times (m+1), \mathbb{K}\right)$ und $k, l \leq m$. Dann definiere $A_{kl} = \left(\bar{a}_{ij}\right) \in \operatorname{Mat}(m \times m, \mathbb{K})$ aus A durch Weglassen der k-ten Zeile und l-ten Spalte:

$$\bar{a}_{ij} := \begin{cases} a_{ij} & \text{falls } i < k \text{ und } j < l \\ a_{i+1,j} & \text{falls } k \leq i < m \text{ und } j < l \\ a_{i,j+1} & \text{falls } i < k \text{ und } l \leq j < m \\ a_{i+1,j+1} & \text{falls } k \leq i < m \text{ und } l \leq j < m \end{cases}$$

Skizzenhaft heißt dies, dass wir die entsprechende Zeile und Spalte löschen:

$$A_{kl} = \left(\begin{array}{cc|c|cc} a_{11} & \ldots & a_{1l} & \ldots & a_{1,m+1} \\ \vdots & & \vdots & & \vdots \\ \hline a_{k1} & \ldots & a_{kl} & \ldots & a_{k,m+1} \\ \hline \vdots & & \vdots & & \vdots \\ a_{m+1,1} & \ldots & a_{m+1,l} & \ldots & a_{m+1,m+1} \end{array}\right)$$

Nach diesem technischen Hilfsmittel folgt nun die eigentliche

Definition 12.3 – Laplacescher Entwicklungssatz/Determinante(nfunktion):

Sei $\mathbb{K}$ ein Körper. Definiere die Determinantenfunktion

$$\det : \operatorname{Mat}(m \times m, \mathbb{K}) \to \mathbb{K}$$

für $m \in \mathbb{N} \setminus \{0\}$ durch Rekursion über m:

(a) Für $A = (a) \in \operatorname{Mat}(1 \times 1, \mathbb{K})$ sei $\det(A) = \det A := a$

(b) Angenommen $\det : \operatorname{Mat}(m \times m, \mathbb{K}) \to \mathbb{K}$ sei bereits definiert. Dann setze die gesuchte Abbildung $\det : \operatorname{Mat}\left((m+1) \times (m+1), \mathbb{K}\right) \to \mathbb{K}$ an wie folgt:

Für $A = \left(a_{ij}\right) \in \operatorname{Mat}\left((m+1) \times (m+1), \mathbb{K}\right)$ sei $\det(A) := \det A$ die folgende alternierende Summe:

$$\begin{aligned} \det(A) &= a_{11} \det A_{11} - a_{21} \det A_{21} + a_{31} \det A_{31} - \ldots \pm a_{m+1,1} \det A_{m+1,1} \\ &= \sum_{k=1}^{m+1} (-1)^{k+1} \cdot a_{k1} \det A_{k1} \end{aligned}$$

Man schreibt oft abkürzend auch $\begin{vmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mm} \end{vmatrix}$ statt $\det \begin{pmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mm} \end{pmatrix}$.

Beispiel 12.4: Für $m = 2$ erhalten wir den bekannten Ausdruck

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = a \cdot |d| - c \cdot |b| = a \cdot d - c \cdot b,$$

den wir vorher bereits als Determinante von $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ bezeichnet haben.

Für $m = 3$ erhalten wir:

$$\begin{aligned} \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} &= a \cdot \begin{vmatrix} e & f \\ h & i \end{vmatrix} - d \cdot \begin{vmatrix} b & c \\ h & i \end{vmatrix} + g \cdot \begin{vmatrix} b & c \\ e & f \end{vmatrix} \\ &= aei - ahf - dbi + dhc + gbf - gec \\ &= aei + dhc + gbf - ahf - dbi - gec \end{aligned}$$

Diese Formel kann man sich folgendermaßen merken: Man hänge die erste und zweite Zeile unten an die Matrix heran und erhält:

$$\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \\ a & b & c \\ d & e & f \end{pmatrix}$$

Danach addiere man die Produkte der drei Hauptdiagonalen ($\searrow$) und subtrahiere die Produkte der drei Nebendiagonalen ($\swarrow$). Diese Merkregel wird auch als die *Regel von Sarrus* bezeichnet.

Das gleiche Ergebnis erhalten wir, wenn wir die ersten beiden Spalten rechts hinzufügen:

$$\begin{pmatrix} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{pmatrix}.$$

Diese Regel gilt allerdings nur für die Dimensionen $m = 2$ und $m = 3$. Für $m = 2$ gilt erwartungsgemäß:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

<u>Strategie für $m = 4$:</u> Wir entwickeln zunächst nach einer Spalte und nutzen dann diese Regel von Sarrus für die 4 Unterdeterminanten (siehe Übungsaufgabe). Wir lernen allerdings noch wie man mithilfe des Gauß-Algorithmus eine höherdimensionale Matrix zunächst vereinfachen kann, um dann relativ bequem die Determinante zu berechnen.

Bemerkung 12.5: In der Definition der Determinante einer Matrix, dem sogenannten Laplaceschem Entwicklungssatz, wird die Determinante nach der *ersten Spalte* entwickelt. Das gleiche Ergebnis und somit eine vielleicht einfachere Möglichkeit der Berechnung erhält man, wenn man nach einer *beliebigen Spalte* oder einer *beliebigen Zeile* entwickelt. Dabei ändert sich allerdings gegebenenfalls der *Wechsel der Vorzeichen*, entsprechend des folgenden Schachbrettmusters:

$$\begin{matrix} + & - & + & - & \dots \\ - & + & - & + & \dots \\ + & - & + & - & \dots \\ - & + & - & + & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{matrix}$$

Betrachten Sie dazu das folgende

Beispiel 12.6: Betrachte $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 0 & 1 \\ 2 & 0 & 3 \end{pmatrix}$ und das dazugehörige Schachbrettmuster

$$\begin{matrix} + & - & + \\ - & + & - \\ + & - & + \end{matrix}.$$

Aufgrund der Nullen in der zweiten Spalte entwickeln wir nach dieser:

$$\begin{aligned} |A| &= -2 \cdot \begin{vmatrix} 4 & 1 \\ 2 & 3 \end{vmatrix} + 0 \cdot \begin{vmatrix} 1 & 3 \\ 2 & 3 \end{vmatrix} - 0 \cdot \begin{vmatrix} 1 & 3 \\ 4 & 1 \end{vmatrix} &&= -2 \cdot (4 \cdot 3 - 2 \cdot 1) \\ &= -2 \cdot (12 - 2) &&= -20 \end{aligned}$$

Wir könnten auch nach der dritten Zeile entwickeln:

$$\begin{aligned} |A| &= 2 \cdot \begin{vmatrix} 2 & 3 \\ 0 & 1 \end{vmatrix} - 0 \cdot \begin{vmatrix} 1 & 3 \\ 4 & 1 \end{vmatrix} + 3 \cdot \begin{vmatrix} 1 & 2 \\ 4 & 0 \end{vmatrix} &&= 2 \cdot (2 \cdot 1 - 0 \cdot 3) + 3 \cdot (1 \cdot 0 - 4 \cdot 2) \\ &= 2 \cdot (2) + 3 \cdot (-8) &&= -20 \end{aligned}$$

Wir könnten die Determinante auch nach irgendeiner beliebigen anderen Zeile oder Spalte entwickeln oder nach der Regel von Sarrus und würden immer das gleiche Ergebnis erhalten. Schauen Sie daher genau hin, welche Regel am schnellsten zum Ziel führt.

$\star \quad \star \quad \star$

Determinaten und Gaußscher Algorithmus. Wie angekündigt befassen wir uns abschließend mit der Methode der Bestimmung der Determinante mittels des Gaußschen Algorithmus. Mithilfe dessen können wir schließlich auch den gewünschten Zusammenhang zwischen einer nicht verschwindenden Determinante und der Invertierbarkeit einer Matrix (siehe Satz 12.10) nachweisen.

Wir starten mit einfachen Umformungen von benachbarten Zeilen.

Satz 12.7:

Sei $A \in \mathrm{Mat}\,(m \times m, \mathbb{K})$. Dann gilt:

- **(a)** Verwandelt man A durch Vertauschen zweier benachbarter Zeilen in die Matrix A', so ist $\det A' = -\det A$.
- **(b)** Verwandelt man A durch Multiplikation einer Zeile mit $\lambda \in \mathbb{K}$ in die Matrix A', so ist $\det A' = \lambda \det A$.
- **(c)** Sei $k < m$ und seien A' und A'' Matrizen, deren i-te Zeilen für $i \neq k$ mit der i-ten Zeile von A übereinstimmen und so dass die k-te Zeile von A die Summe der k-ten Zeilen von A' und A'' ist. Dann ist $\det A = \det A' + \det A''$.
- **(d)** Verwandelt man A durch Addition eines skalaren Vielfachen einer Zeile zu einer darauf folgenden Zeile in die Matrix A', so ist $\det A' = \det A$.

Beweis: Wir beweisen die Behauptungen durch Induktion über $m \in \mathbb{N} \setminus \{0\}$.
Induktionsanfang: Für $m = 1$ besitzt die Matrix $A = (a)$ nur eine Zeile.

Zu (a): Klar!

Zu (b): Es gilt: $\det A' = \det(\lambda \cdot a) = \lambda a = \lambda \det(a) = \lambda \det A$

Zu (c): Betrachte $A' = (a')$, $A'' = (a'')$ und $A = (a' + a'')$. Dann gilt:

$$\det A = a' + a'' = \det A' + \det A''.$$

Zu (d): Klar!

<u>**Induktionsschritt:**</u> Wir nehmen an, dass die Behauptungen (a)–(d) für m gelten. Betrachte $A = \left(a_{ij}\right) \in \operatorname{Mat}\left((m+1) \times (m+1), \mathbb{K}\right)$.

Zu (a): Betrachte $k \leq m$. Die Matrix $A' = \left(a'_{ij}\right) \in \operatorname{Mat}\left((m+1) \times (m+1), \mathbb{K}\right)$ entstehe aus A durch Vertauschen der Zeilen k und $k+1$. Dann gilt:

$$\begin{aligned}
\det\left(A'\right) &= \sum_{i=1}^{m+1} (-1)^{i+1} a'_{i1} \det\left(A'_{i1}\right) \\
&= (-1)^{k+1} a'_{k1} \det\left(A'_{k1}\right) + (-1)^{k+2} a'_{k+1,1} \det\left(A'_{k+1,1}\right) + \sum_{\substack{i=1,\ i\neq k \\ i\neq k+1}}^{m+1} (-1)^{i+1} a'_{i1} \det\left(A'_{i1}\right) \\
&= (-1)^{k+1} a_{k+1,1} \det A_{k+1,1} + (-1)^{k+2} a_{k1} \det(A_{k1}) + \sum_{\substack{i=1,\ i\neq k \\ i\neq k+1}}^{m+1} (-1)^{i+1} a'_{i1} \det\left(A'_{i1}\right) \\
&= -(-1)^{k+2} a_{k+1,1} \det\left(A_{k+1,1}\right) - (-1)^{k+1} a_{k1} \det(A_{k1}) - \sum_{\substack{i=1,\ i\neq k \\ i\neq k+1}}^{m+1} (-1)^{i+1} a_{i1} \det(A_{i1}) \\
&= -\sum_{i=1}^{m+1} (-1)^{i+1} a_{i1} \det(A_{i1}) \\
&= -\det(A)
\end{aligned}$$

Zur Begründung der Umformung beachte man, dass

- $a'_{k1} = a_{k+1,1}$ und $a'_{k+1,1} = a_{k1}$
- für $i \neq k, i \neq k+1$ ist $a'_{i1} = a_{i1}$
- $A'_{k1} = A_{k+1,1}$ und $A'_{k+1,1} = A_{k1}$
- für $i \neq k, i \neq k+1$ entsteht A'_{i1} aus A_{i1} durch Vertauschen benachbarter Zeilen, nach Induktionsannahme ist dann $\det A'_{i1} = -\det A_{i1}$

Zu (b): Betrachte $k \leq m$. Die Matrix $A' = (a'_{ij}) \in \operatorname{Mat}((m+1) \times (m+1), \mathbb{K})$ entstehe aus A durch Multiplikation der k-ten Zeile mit dem Faktor $\lambda \in \mathbb{K}$. Dann gilt:

$$\begin{aligned}
\det A' &= \sum_{i=1}^{m+1} (-1)^{i+1} a'_{i1} \det A'_{i1} \\
&= (-1)^{k+1} a'_{k1} \det A'_{k1} + \sum_{\substack{i=1 \\ i\neq k}}^{m+1} (-1)^{i+1} a'_{i1} \det A'_{i1} \\
&= (-1)^{k+1} \lambda a_{k1} \det A_{k1} + \sum_{\substack{i=1 \\ i\neq k}}^{m+1} (-1)^{i+1} a_{i1} \lambda \det A_{i1} \\
&= \lambda \cdot \sum_{i=1}^{m+1} (-1)^{i+1} a_{i1} \det A_{i1}
\end{aligned}$$

$$= \quad \lambda \cdot \det A$$

Zur Begründung der Umformung beachte man, dass

- $a'_{k1} = \lambda a_{k1}$
- für $i \neq k$ ist $a'_{i1} = a_{i1}$
- $A'_{k1} = A_{k1}$
- für $i \neq k$ entsteht A'_{i1} aus A_{i1} durch skalare Multiplikation einer Zeile mit dem Faktor λ, nach Induktionsannahme ist dann $\det A'_{i1} = \lambda \cdot \det A_{i1}$

Zu (c): Es seien $A = \left(a_{ij}\right), A' = \left(a'_{ij}\right), A'' = \left(a''_{ij}\right) \in \operatorname{Mat}\left((m+1) \times (m+1), \mathbb{K}\right)$ wie angegeben. Dann gilt:

$$\begin{aligned}
\det A &= \sum_{i=1}^{m+1} (-1)^{i+1} a_{i1} \det A_{i1} \\
&= (-1)^{k+1} a_{k1} \det A_{k1} + \sum_{\substack{i=1 \\ i \neq k}}^{m+1} (-1)^{i+1} a_{i1} \det A_{i1} \\
&= (-1)^{k+1} (a'_{k1} + a''_{k1}) \det A_{k1} + \sum_{\substack{i=1 \\ i \neq k}}^{m+1} (-1)^{i+1} a_{i1} (\det A'_{i1} + \det A''_{i1}) \\
&= (-1)^{k+1} a'_{k1} \det A'_{k1} + \sum_{\substack{i=1 \\ i \neq k}}^{m+1} (-1)^{i+1} a'_{i1} \det A'_{i1} \\
&\qquad + (-1)^{k+1} a''_{k1} \det A''_{k1} + \sum_{\substack{i=1 \\ i \neq k}}^{m+1} (-1)^{i+1} a''_{i1} \det A''_{i1} \\
&= \sum_{i=1}^{m+1} (-1)^{i+1} a'_{i1} \det A'_{i1} + \sum_{i=1}^{m+1} (-1)^{i+1} a''_{i1} \det A''_{i1} \\
&= \det A' + \det A''
\end{aligned}$$

Zur Begründung der Umformung beachte man, dass

- $a_{k1} = a'_{k1} + a''_{k1}$
- für $i \neq k$ entsteht A_{i1} aus A'_{i1} und A''_{i1} durch Addieren einer Zeile und Kopieren der übrigen Zeilen, nach Induktionsannahme ist dann $\det A_{i1} = \det A'_{i1} + \det A''_{i1}$
- Wegen der Übereinstimmung in den Zeilen von A, A' und A'' ist $A_{k1} = A'_{k1} = A''_{k1}$
- für $i \neq k$ ist $a_{i1} = a'_{i1} = a''_{i1}$

Zu (d): Betrachte $k \leq m$ und $\lambda \in \mathbb{K}$. Die Matrix $A' = (a'_{ij}) \in \operatorname{Mat}((m+1) \times (m+1), \mathbb{K})$ entstehe aus A durch Addition des λ-fachen der Zeile k zur Zeile $k+1$. Dann gilt:

$$\begin{aligned}
\det A' &= \sum_{i=1}^{m+1} (-1)^{i+1} a'_{i1} \det A'_{i1} \\
&= (-1)^{k+1} a'_{k1} \det A'_{k1} + (-1)^{k+2} a'_{k+1,1} \det A'_{k+1,1} + \sum_{\substack{i=1,\ i \neq k \\ i \neq k+1}}^{m+1} (-1)^{i+1} a'_{i1} \det A'_{i1} \\
&= (-1)^{k+1} a_{k1} \det A'_{k1} + (-1)^{k+2} (a_{k+1,1} + \lambda a_{k1}) \det A_{k+1,1}
\end{aligned}$$

$$\begin{aligned}
&\qquad + \sum_{\substack{i=1,\ i\neq k\\ i\neq k+1}}^{m+1} (-1)^{i+1} a_{i1} \det A'_{i1}\\
&= (-1)^{k+1} a_{k1} (\det A_{k1} + \lambda \det A_{k+1,1})\\
&\qquad + (-1)^{k+2} a_{k1} \det A_{k1} + (-1)^{k+2} (a_{k+1,1} + \lambda a_{k1}) \det A_{k+1,1}\\
&\qquad + \sum_{\substack{i=1,\ i\neq k\\ i\neq k+1}}^{m+1} (-1)^{i+1} a_{i1} \det A_{i1} + \sum_{\substack{i=1,\ i\neq k\\ i\neq k+1}}^{m+1} (-1)^{i+1} a_{i1} \det A_{i1}\\
&= \sum_{i=1}^{m+1} (-1)^{i+1} a_{i1} \det A_{i1} = \det A
\end{aligned}$$

Zur Begründung der Umformung beachte man, dass

- $a'_{k+1,1} = a_{k+1,1} + \lambda a_{k1}$
- für $i \neq k$ ist $a'_{i1} = a_{i1}$
- $A'_{k+1,1} = A_{k+1,1}$
- für $i \neq k, i \neq k+1$ entsteht A'_{i1} aus A_{i1} durch Addition des λ-fachen einer Zeile, nach Induktionsannahme ist dann $\det A'_{i1} = \det A_{i1}$
- die k-te Zeile von A'_{k1} ist die Summe der k-ten Zeile von A_{k1} und dem λ-fachen der k-ten Zeile von $A_{k+1,1}$, die übrigen Zeilen von A'_{k1}, A_{k1} und $A_{k+1,1}$ stimmen überein. Nach der in (b) und (c) formulierten Zeilenlinearität ist $\det A'_{k1} = \det A_{k1} + \lambda A_{k+1,1}$

Damit ist der Satz nach dem Prinzip der vollständigen Induktion bewiesen. ⊠

Wir wollen jetzt versuchen, diese Aussage für Umformungen benachbarter Zeilen zu verallgemeinern.

Betrachte dazu natürliche Zahlen $1 \leq k < l \leq m$. Die Vertauschung der Zahlen k und l kann durch eine Folge von Vertauschungen benachbarter Zahlen erreicht werden:

$$\begin{array}{ccccccccc}
\dots & k & k+1 & k+2 & \dots & l-2 & l-1 & l & \dots\\
\dots & k+1 & k & k+2 & \dots & l-2 & l-1 & l & \dots\\
\dots & k+1 & k+2 & k & \dots & l-2 & l-1 & l & \dots\\
 & & & & \ddots & & & & \\
\dots & k+1 & k+2 & k+3 & \dots & k & l-1 & l & \dots\\
\dots & k+1 & k+2 & k+3 & \dots & l-1 & k & l & \dots\\
\dots & k+1 & k+2 & k+3 & \dots & l-1 & l & k & \dots\\
\dots & k+1 & k+2 & k+3 & \dots & l & l-1 & k & \dots\\
 & & & & \ddots & & & & \\
\dots & k+1 & k+2 & l & \dots & l-2 & l-1 & k & \dots\\
\dots & k+1 & l & k+2 & \dots & l-2 & l-1 & k & \dots\\
\dots & l & k+1 & k+2 & \dots & l-2 & l-1 & k & \dots
\end{array}$$

Diese Folge enthält $l-k$ Nachbarvertauschungen, um k an die l-te Stelle zu transportieren und $l-k-1$ Nachbarvertauschungen, um anschließend l an die k-te Stelle zu transportieren. Insgesamt besteht die Folge aus $2(l-k)-1$ Vertauschungen, also einer ungeraden Anzahl. Mit dieser Beobachtung können wir den vorangehenden Satz auch auf nicht-benachbarte Vertauschungen und Zeilen-Additionen ausdehnen. Eine beliebige Vertauschung kann man durch eine ungerade Anzahl von Nachbarvertauschungen realisieren. Wenn man das λ-fache der k-ten Zeile zur l-ten Zeile addieren möchte, vertauscht man zunächst Zeilen, so dass die Additionssituation von Satz 12.7 Teil

(d) vorliegt. Satz 12.7 impliziert die Invarianz der Determinante. Anschließend tauscht man die vertauschten Zeilen zurück. Die hierdurch verursachten Vorzeichenwechsel heben sich auf. Damit gilt der folgende

Satz 12.8:

Sei $A \in \operatorname{Mat}(m \times m, \mathbb{K})$. Dann gilt:

(a) Verwandelt man A durch Vertauschen zweier Zeilen in die Matrix A', so ist $\det A' = -\det A$.

(b) Verwandelt man A durch Addition eines skalaren Vielfachen einer Zeile zu einer anderen Zeile in die Matrix A', so ist $\det A' = \det A$.

Praktisch Determinaten berechnen. Fassen wir noch einmal zusammen, wie wir mit Papier und Bleistift Determinanten berechnen, wenn wir keinen Computer benutzen (können):

- Für $(n \times n)$-Matrizen und $n \leq 3$ hilft die Regel von Sarrus.
- Für $n = 4$ könnte es sinnvoll sein, zunächst nach einer Zeile/Spalte mittels des Laplaceschen Entwicklungssatzes zu entwickeln und schließlich für die Unterdeterminanten die Regel von Sarrus zu nehmen.
- Generell ist es sinnvoll, nach einer Zeile und Spalte zu entwickeln, die sehr viele Nullen enthält.
- Bei großen Matrizen ist es auf jeden Fall sinnvoll, zunächst mittels der beschriebenen Gauß-Methode Umformungen anzusetzen, um schließlich über eine Nullzeile oder -spalte zu entwickeln – oder gar die Determinante aus der Zeilenstufenform ganz einfach zu entwickeln, wie auch der folgende Satz zeigt.

Dennoch, Determinanten großer Matrizen seien dem Computer überlassen. Am einfachsten programmiert ist der Laplacesche Entwicklungssatz – dieser ist allerdings aufgrund der Rekursion eher teuer umgesetzt.

⋆ ⋆ ⋆

Wir schauen uns nun weitere Eigenschaften der Determinantenfunktion an.

Satz 12.9:

Sei $A = \left((a_{ij})\right) \in \operatorname{Mat}(m \times m, \mathbb{K})$ eine Matrix in Zeilenstufenform. Dann ist

$$\det A = \prod_{i=1}^{m} a_{ii}.$$

Insbesondere gilt für die Einheitsmatrix $E_m \in \operatorname{Mat}(m \times m, \mathbb{K})$ stets: $\det E_m = 1$.

Beweis von Satz 12.9: Durch Induktion über $m \in \mathbb{N} \setminus \{0\}$:

Induktionsanfang

$\det A = \det (a_{11}) = a_{11} = \prod_{i=1}^{1} a_{ii}$.

Induktionsschritt

Die Behauptung gelte für m. Dann ist nach Induktionsannahme

$$\det A = a_{11} \cdot \det A_{11} = a_{11} \cdot \prod_{i=2}^{m+1} a_{ii} = \prod_{i=1}^{m+1} a_{ii}$$

⊠(Satz 12.9)

Wir erhalten nun die für die Theorie der Eigenwerte und Eigenvektoren wichtigste Eigenschaft von Matrizen:

Satz 12.10:

Sei $\mathbb{K}$ ein Körper, $m \in \mathbb{N} \setminus \{0\}$ und $A \in \mathrm{Mat}\,(m \times m, \mathbb{K})$. Dann ist A invertierbar genau dann, wenn $\det A \neq 0$.

Beweis: Überführe $A = \left((a_{ij})\right)$ durch den Gauß-Algorithmus mittels Folge von elementaren Zeilenumformungen in eine Matrix $A' = \left(a'_{ij}\right)$ in Zeilenstufenform. Nach den Sätzen 12.7 und 12.8 wird die Determinante der Matrix A entlang der Umformungen nur mit Faktoren $\neq 0$ multipliziert. Das bedeutet

$$\det A = 0 \text{ genau dann, wenn } \det A' = 0.$$

Die Matrix A ist genau dann invertierbar, wenn alle Diagonaleinträge a'_{ii} von A' ungleich Null sind. Wir erhalten folgende Äquivalenzen:

$$\begin{aligned} A \text{ ist invertierbar genau dann, wenn } & \prod_{i=1}^{m} a'_{ii} \neq 0 \\ \text{genau dann, wenn } & \det A' \neq 0 \\ \text{genau dann, wenn } & \det A \neq 0 \end{aligned}$$

⊠

Bemerkung 12.11: Der vorangehende Beweis enthält implizit ein Verfahren zur Berechnung der Determinante einer Matrix A mit Hilfe des Gauß-Verfahrens:

Die Matrix A wird durch die elementaren Zeilenumformungen des Gauß-Algorithmus auf Zeilenstufenform gebracht. Die Determinante der Zeilenstufenmatrix ist das Produkt der Diagonaleinträge. Die Determinante der ursprünglichen Matrix A ergibt sich aus der Determinante der Zeilenstufenmatrix durch Vorzeichenwechsel entsprechender Folgen der Zeilenumformungen.

Beispiel 12.12: Wir berechnen die Determinante mittels Gauß-Umformungen:

$$\begin{aligned} \begin{vmatrix} 2 & 0 & -1 & -4 \\ -3 & 1 & 3 & 0 \\ 2 & 0 & -1 & -2 \\ 1 & 0 & -1 & -1 \end{vmatrix} &= -\begin{vmatrix} 1 & 0 & -1 & -1 \\ -3 & 1 & 3 & 0 \\ 2 & 0 & -1 & -2 \\ 2 & 0 & -1 & -4 \end{vmatrix} = -\begin{vmatrix} 1 & 0 & -1 & -3 \\ 0 & 1 & 0 & -9 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & -2 \end{vmatrix} \\ &= -\begin{vmatrix} 1 & 0 & -1 & -3 \\ 0 & 1 & 0 & -9 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -2 \end{vmatrix} = -\left(1 \cdot 1 \cdot 1 \cdot (-2)\right) = 2 \end{aligned}$$

Dies stimmt mit dem Ergebnis überein, wenn wir nach der zweiten Spalte entwickeln:

$$\begin{aligned}
\begin{vmatrix} 2 & 0 & -1 & -4 \\ -3 & 1 & 3 & 0 \\ 2 & 0 & -1 & -2 \\ 1 & 0 & -1 & -1 \end{vmatrix} &= 1 \cdot \begin{vmatrix} 2 & -1 & -4 \\ 2 & -1 & -2 \\ 1 & -1 & -1 \end{vmatrix} \\
&= 2 \cdot (-1) \cdot (-1) + (-1) \cdot (-2) \cdot 1 + (-4) \cdot 2 \cdot (-1) \\
&\qquad - (-4) \cdot (-1) \cdot 1 - (-2) \cdot (-1) \cdot 2 - (-1) \cdot (-1) \cdot 2 \\
&= 2 + 2 + 8 - 4 - 4 - 2 \\
&= 2
\end{aligned}$$

⋆ ⋆ ⋆

Folgender Satz fasst noch einmal den übergreifenden Zusammenhang der verschiedenen Begriffe zusammen – staunen Sie selbst:

Satz 12.13:

Für eine Matrix $A \in \mathrm{Mat}\,(n \times n, \mathbb{K})$ sind die folgenden Aussagen äquivalent:

(a) Die Determinante von A ist ungleich Null.
(b) Der Rang der Matrix ist maximal, nämlich gleich n.
(c) Der Kern von A ist trivial, nämlich gleich der Menge $\{0\}$.
(d) Die Zeilenvektoren von A sind linear unabhängig.
(e) Die Spaltenvektoren von A sind linear unabhängig.
(f) Die Inverse von A existiert.
(g) Das homogene lineare Gleichungssystem $Ax = 0$ besitzt eine eindeutige Lösung, nämlich den Nullvektor.
(h) Das inhomogene lineare Gleichungssystem $Ax = b$ besitzt für eine beliebige rechte Seite b eine eindeutige Lösung.

Achtung!

In dieser Äquivalenz steht nichts über Diagonalisierbarkeit!

Den Beweis werden wir an dieser Stelle nicht bringen, aber es ist eine gute Übungsaufgabe, sich die Zusammenhänge selbst zu überlegen, denn wir haben bereits alle Teile des Beweises in den vorhergehenden Betrachtungen (innerhalb der verschiedenen Kapitel!) gesehen bzw. angerissen.

⋆ ⋆ ⋆

Mit unserem bisherigen Wissen können wir bereits eine wichtige Rechenregel für Determinanten feststellen:

Lemma 12.14:

Für beliebige $(n \times n)$-Matrizen gilt: $\det(AB) = \det(A)\det(B)$.

Beweis: Wir führen eine Fallunterscheidung durch:

<u>**Fall:** $\det(B) = 0$</u>:

Aufgrund des Dimensionssatzes ist die Matrix nicht invertierbar und damit kann B nicht injektiv sein. Also existiert ein Vektor $v \neq 0$ mit $Bv = 0$. Das bedeutet aber wiederum, dass $(AB)v = 0$ und deswegen die Matrix AB nicht injektiv sein kann und somit wegen der Dimensionsformel auch nicht invertierbar.

Zusammengefasst gilt: $\det(AB) = 0 = \det(A) \cdot 0 = \det(A) \cdot \det(B)$.

<u>**Fall:** $\det(B) \neq 0$</u>:

Dann können wir B nur mithilfe der dritten Gaußoperation als $B = D \cdot P_1 \cdot \ldots \cdot P_k$ schreiben. Hier handelt es sich bei den P_i um Elementarmatrizen, die die Determinante nicht verändern und D ist eine Diagonalmatrix. Außerdem wissen wir bereits, dass gilt:

$$\det(D) = \left| \begin{pmatrix} d_{11} & & \\ & \ddots & \\ & & d_{nn} \end{pmatrix} \right| = d_{11} \cdot \ldots \cdot d_{nn}$$

Und es gilt: $\det(AB) = \det(AD \cdot P_1 \cdot \ldots \cdot P_k) = \det(AD)$.

Dabei entsteht die Matrix AD aus der Matrix A wobei die j-te Spalte von A mit dem j-ten Diagonaleintrag von D multipliziert wurde. Die Eigenschaften der Determinante sagen uns jetzt, dass wir aus jeder Spalte einen Faktor d_{jj} herausziehen können:

$$\begin{aligned} \left| \begin{pmatrix} d_{11} \cdot a_{11} & \ldots & d_{jj} \cdot a_{1j} & \ldots & d_{nn} \cdot a_{1n} \\ \vdots & & \vdots & & \vdots \\ d_{11} \cdot a_{n1} & \ldots & d_{jj} \cdot a_{nj} & \ldots & d_{nn} \cdot a_{nn} \end{pmatrix} \right| &= d_{11} \cdot \ldots \cdot d_{nn} \left| \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} \right| \\ &= \det(D) \det(A) \end{aligned}$$

Zusammengefasst gilt wie gewünscht:

$$\det(AB) = \det(AD \cdot P_1 \cdot \ldots \cdot P_k) = \det(AD) = \det(D) \det(A) = \det(B) \det(A)$$

⊠

⋆ ⋆ ⋆

Die Cramersche Regel. In diesem Abschnitt werden wir noch eine weitere wesentliche Anwendung der Determinanten kennenlernen. Wir beschäftigen uns nun mit linearen Gleichungssystemen, die auf einer invertierbaren Matrix beruhen, also sei für ein gegebenes Gleichungssystem $Ax = b$ die Determinante ungleich Null. Damit ist die Lösung eindeutig bestimmt und die Lösungsmenge besteht nur aus einem Vektor, nämlich $x = A^{-1}b$. Die Berechnung der Inversen ist allerdings nicht immer leicht, sodass wir auf eine andere Methode zurückgreifen werden.

Neben dem Gaußschem Algorithmus befassen wir uns nun mit folgender Regel:

Ein lineares Gleichungssystem $Ax = b$, mit $\det A \neq 0$, besitzt die eindeutige Lösung $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$.
Dann ist $x_1 = \frac{D_1}{D}, \ldots, x_n = \frac{D_n}{D}$, mit $D := \det A$ und $D_i = \det A_i$ für $i = 1, \ldots, n$, wobei A_i aus A hervorgeht indem man die i-te Spalte durch die rechte Seite b ersetzt. Diese Art der Lösungsangabe eines linearen Gleichungssystems nennt man die Cramersche Regel.

Betrachten wir folgendes

Beispiel 12.15: Gegeben sei das lineare Gleichungssystem $Ax = b$ mit

$$A = \begin{pmatrix} 2 & 3 & 1 \\ 0 & 4 & 2 \\ 1 & 0 & -2 \end{pmatrix}, \; x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \text{ und } b = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

Berechne $D = \det A = \begin{vmatrix} 2 & 3 & 1 \\ 0 & 4 & 2 \\ 1 & 0 & -2 \end{vmatrix} = -14$,

$D_1 = \begin{vmatrix} 1 & 3 & 1 \\ 0 & 4 & 2 \\ 1 & 0 & -2 \end{vmatrix} = -6$, $D_2 = \begin{vmatrix} 2 & 1 & 1 \\ 0 & 0 & 2 \\ 1 & 1 & -2 \end{vmatrix} = -2$ und $D_3 = \begin{vmatrix} 2 & 3 & 1 \\ 0 & 4 & 0 \\ 1 & 0 & 1 \end{vmatrix} = 4$.

Somit gilt:

$$x_1 = \frac{D_1}{D} = \frac{3}{7}, \qquad x_2 = \frac{D_2}{D} = \frac{1}{7}, \qquad x_3 = \frac{D_3}{D} = -\frac{2}{7}$$

Offenbar kann die Cramersche Regel *nicht* für eine Matrix A mit $\det A = 0$ angewendet werden, aber das ist auch klar, denn eine solche Matrix wäre nicht invertierbar, hätte also einen nicht-trivialen Kern und somit keine eindeutige Lösung.

Für ein homogenes Gleichungssystem $Ax = 0$ mit $\det A \neq 0$ wird kein solches Lösungsverfahren benötigt, da es nur *eine* Lösung geben kann und der Nullvektor selbst eine Lösung eines homogenen Gleichungssystems darstellt.

⋆ ⋆ ⋆

Determinanten und Volumina. Im m-dimensionalen euklidischen Raum $\mathbb{R}^m$ lassen sich mit der Determinantenfunktion Volumina von parallelogrammförmigen Mengen bestimmen. Wir führen diesen Sachverhalt am Beispiel der bekannten reellen Ebene $\mathbb{R}^2$ vor:

Die Vektoren $\begin{pmatrix} a \\ b \end{pmatrix}$ und $\begin{pmatrix} c \\ d \end{pmatrix}$ spannen ein Parallelogramm

$$F = \left\{ \lambda_0 \begin{pmatrix} a \\ b \end{pmatrix} + \lambda_1 \begin{pmatrix} c \\ d \end{pmatrix} \;\middle|\; 0 \leq \lambda_0, \lambda_1 \leq 1 \right\}$$

auf. Der Flächeninhalt $|F|$ von F ergibt sich als Differenz des äußeren Rechtecks und der rechtwinkligen Dreiecke und Rechtecke zwischen äußerem Rechteck und Parallelogramm. Zur Vermeidung von Vorzeichenfragen nehmen wir an, dass $a, b, c, d \geq 0$ sind, so dass wir erhalten:

$$|F| \quad = \quad (a + c)(d + b) - \frac{1}{2}ab - cb - \frac{1}{2}cd - \frac{1}{2}ab - cb - \frac{1}{2}cd$$

$$= \quad ad + ab + cd + cb - ab - 2cb - cd \qquad = ad - cb \qquad = \begin{vmatrix} a & c \\ b & d \end{vmatrix}$$

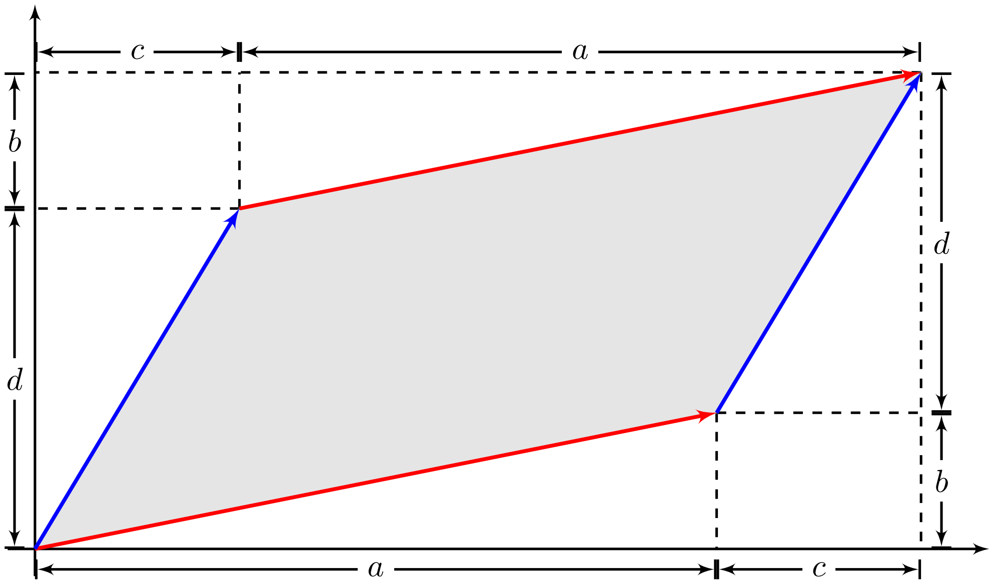


Abbildung 19: Parallelogramm

Damit entspricht der Flächeninhalt des Parallelogramms wie gewünscht der Determinaten (vgl. Abbildung 19). (Insbesondere überlegt man sich leicht[3] mittels Geradengleichung der Diagonalen, dass in diesem Fall der Term $ad - cb$ auch immer positiv ist.)

Diese Situation überträgt sich auf höhere Dimensionen.

Definition 12.16 – aufgespanntes Parallelogramm von Vektoren:

Sei $m \in \mathbb{N} \setminus \{0\}$. Seien $a_1, \ldots, a_m \in \mathbb{R}^m$. Dann ist

$$[a_1, \ldots, a_m] = \left\{ \sum_{i=1}^{m} \lambda_i a_i \;\middle|\; 0 \leq \lambda_i \leq 1 \right\}$$

das von $a_1, \ldots, a_m$ aufgespannte Parallelogramm.

[3] Der Punkt $\begin{pmatrix} c \\ d \end{pmatrix}$ liegt über der Diagonalen des Parallelogramms, d.h. es gilt $d \geq \frac{b+d}{a+c} \cdot c$, wobei $y = \frac{b+d}{a+c} \cdot x$ die Geradengleichung für die Diagonale durch die Punkte $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ und $\begin{pmatrix} a+c \\ b+d \end{pmatrix}$ ist. Somit ist $ad + cd \geq bc + dc$ und damit gilt wie gewünscht $ad - bc \geq 0$.

Der folgende Satz kann mit den Rechengesetzen für Determinanten gezeigt werden:

Satz 12.17:

Sei $m \in \mathbb{N} \setminus \{0\}$ und $a_1, \ldots, a_m \in \mathbb{R}^m$. Für $j \leq m$ sei

$$a_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Dann ist das Volumen des von $a_1, \ldots, a_m$ aufgespannten Parallelogramms der Absolutbetrag von

$$\begin{vmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mm} \end{vmatrix}.$$

Insbesondere gilt: Wenn $a_1, \ldots, a_m$ linear abhängig sind, so ist das Volumen gleich 0. In diesem Fall kollabiert das Parallelogramm zu einem Gebilde von Dimension echt kleiner als m und hat damit auch anschaulich das Volumen 0.

Den anfänglichen Zusammenhang über Flächeninhalt von Parallelogrammen und Determinanten werden wir im übernächsten Kapitel wieder aufgreifen.

⋆ ⋆ ⋆

Aufgabe 12.18: Sei f ein Endomorphismus, $f : V \to V$ für den gilt, dass jeder Vektor $v \neq 0$ ein Eigenvektor ist. Zeigen Sie, dass dann $f = c \cdot \mathrm{id}_v$ gilt.

13. Charakteristische Polynome & Diagonalisierbarkeit

Nachdem wir nun wissen, wie wir Determinanten berechnen können, kommen wir zurück zum eigentlichen Thema: die Bestimmung der Eigenwerte einer gegebenen Matrix.

Definition 13.1 – charakteristische Polynom:

Sei $\mathbb{K}$ ein Körper und $m \in \mathbb{N} \setminus \{0\}$. Sei $A = (a_{ij}) \in \mathrm{Mat}(m \times m, \mathbb{K})$. Dann definiere das charakteristische Polynom von A als

$$p_A(\lambda) = \det(A - \lambda \cdot E_m) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} - \lambda & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{m-1,m} \\ a_{m1} & \dots & a_{m,m-1} & a_{mm} - \lambda \end{vmatrix}$$

Beispiel 13.2: Sei $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$. Das charakteristische Polynom von A ist:

$$\begin{aligned} p_A(\lambda) = \det(A - \lambda E_3) &= \begin{vmatrix} 1-\lambda & 1 & 0 \\ 0 & -1-\lambda & 1 \\ 0 & 1 & -1-\lambda \end{vmatrix} \\ &= (1-\lambda) \cdot \begin{vmatrix} -1-\lambda & 1 \\ 1 & -1-\lambda \end{vmatrix} \qquad \text{Entwicklung nach erster Spalte} \\ &= (1-\lambda)((-1-\lambda) \cdot (-1-\lambda) - 1) = (1-\lambda)(1 + 2\lambda + \lambda^2 - 1) \\ &= (1-\lambda)(2\lambda + \lambda^2) = (1-\lambda)\lambda(2+\lambda) = -\lambda^3 - \lambda^2 + 2\lambda \end{aligned}$$

Das charakteristische Polynom hier im Beispiel ist also ein Polynom dritten Grades in der Variable λ. Aus der vorletzten Zeile in der Gleichungskette kann man die Nullstellen ablesen, nämlich $-2, 0, 1$.

Ganz allgemein sieht man, dass das charakteristische Polynom aus den Summanden der Gestalt $c \cdot \lambda^i$ besteht mit Koeffizienten $c \in \mathbb{K}$ und $i \leq m$. Die höchste Potenz von λ erscheint nur in dem Summanden $(-1)^m \lambda^m$ als Produkt der Hauptdiagonalen. Insbesondere gilt der folgende

Satz 13.3:

Sei $\mathbb{K}$ ein Körper und $m \in \mathbb{N} \setminus \{0\}$. Sei $A = (a_{ij}) \in \mathrm{Mat}(m \times m, \mathbb{K})$. Dann ist das charakteristische Polynom von A von der Gestalt

$$p_A(\lambda) = (-1)^m \lambda^m + c_{m-1}\lambda^{m-1} + \ldots + c_1\lambda + c_0.$$

Die Existenz von Nullstellen derartiger Polynome hängt allerdings vom Körper $\mathbb{K}$ ab. Wenn $\mathbb{K}$ der Körper $\mathbb{C}$ der komplexen Zahlen ist, so hat $p_A(\lambda)$ nach dem Fundamentalsatz der Algebra immer eine Nullstelle und somit hat ein Polynom über $\mathbb{C}$ vom Grade m auch immer m (nicht notwendig verschiedene) Nullstellen (das heißt, die Vielfachen werden mitgezählt).

Mit dem folgenden Satz sind wir schließlich am Ziel unserer Betrachtungen:

Satz 13.4:

Sei $f : V \to V$ ein Endomorphismus eines m-dimensionalen $\mathbb{K}$-Vektorraums V, $m \in \mathbb{N} \setminus \{0\}$. Sei B eine Basis von V und $A = \mathrm{DM}_B(f)$. Dann gilt für alle $\lambda \in \mathbb{K}$: λ ist genau dann ein Eigenwert von f, wenn $p_A(\lambda) = 0$.

Beweis: Betrachte $\lambda \in \mathbb{K}$. Die folgenden Aussagen sind äquivalent:

$$\begin{aligned}
&\lambda && \text{ist Eigenwert von } f\\
\Longleftrightarrow\ &\mathrm{Kern}\,(f - \lambda \mathrm{Id}_V) \neq \{0\} &&\\
\Longleftrightarrow\ &\dim\big(\mathrm{Kern}\,(f - \lambda \mathrm{Id}_V)\big) > 0 &&\\
\Longleftrightarrow\ &\mathrm{Rg}\,(f - \lambda \mathrm{Id}_V) < m &&\\
\Longleftrightarrow\ &f - \lambda \mathrm{Id}_V && \text{ist nicht invertierbar}\\
\Longleftrightarrow\ &\mathrm{DM}_B\,(f - \lambda \mathrm{Id}_V) = \mathrm{DM}_B\,(f) - \lambda \mathrm{DM}_B\,(\mathrm{Id}_V) = A - \lambda E_m && \text{ist nicht invertierbar}\\
\Longleftrightarrow\ &p_A\,(\lambda) = \det\,(A - \lambda E_m) = 0 && \text{(nach Satz 12.10)}
\end{aligned}$$

⊠

Beispiel 13.5: Wir setzen das Beispiel 13.2 fort: Wir haben bereits gesehen, dass die Nullstellen des charakteristischen Polynoms von $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$ die Skalare $-2, 0$ und 1 sind. Nach Satz 13.4 sind dies die Eigenwerte von A. Wir bestimmen jetzt die zugehörigen Eigenvektoren durch Lösen des homogenen linearen Gleichungssystems $(A - \lambda_i E_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$:

$\boxed{\lambda_1 = -2}$ Betrachte $A - (-2)E_3 = A + 2E_3 = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$. Durch den Gauß-Algorithmus erhalten wir: $\begin{pmatrix} 1 & \frac{1}{3} & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, das heißt z ist beliebig wählbar, damit:

$$y = -z \qquad \text{und} \qquad 3x = -y,$$

das heißt

$$x = -\frac{y}{3} = \frac{z}{3} \qquad \text{und} \qquad y = -z.$$

Wähle $z = 3$ und wir erhalten den Eigenvektor $u_{-2} = \begin{pmatrix} 1 \\ -3 \\ 3 \end{pmatrix}$.

$\boxed{\lambda_2 = 0}$ Betrachte $A - 0 \cdot E_3 = A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$. Durch den Gauß-Algorithmus erhalten wir: $\begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, das heißt z ist beliebig wählbar, damit: $y = z$ und $x = -y = -z$. Wir erhalten den Eigenvektor $u_0 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$.

$\boxed{\lambda_3 = 1}$ Betrachte $A - 1 \cdot E_3 = A - E_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & -2 & 1 \\ 0 & 1 & -2 \end{pmatrix}$. Durch den Gauß-Algorithmus erhalten wir: $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -2 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, das heißt x ist beliebig wählbar und es gilt: $y = 0$ und $z = 0$. Wir erhalten den Eigenvektor $u_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

Das System $C = \left(\begin{pmatrix} 1 \\ -3 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right)$ ist eine Basis des $\mathbb{R}^3$, die nur aus Eigenvektoren besteht.

Die Transformationsmatrix T_K^C, für K die kanonische Basis, ist:

$$T_K^C = \begin{pmatrix} 1 & -1 & 1 \\ -3 & 1 & 0 \\ 3 & 1 & 0 \end{pmatrix}$$

Die Transformationsmatrix T_C^K erhalten wir durch Invertieren von T_K^C:

$$\begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ -3 & 1 & 0 & 0 & 1 & 0 \\ 3 & 1 & 0 & 0 & 0 & 1 \end{array} \overset{\text{Gauß}}{\rightsquigarrow} \dots \overset{\text{Gauß}}{\rightsquigarrow} \begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & -\frac{1}{6} & \frac{1}{6} \\ 0 & 1 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & 1 & \frac{2}{3} & \frac{1}{3} \end{array}$$

das heißt: $T_C^K = \begin{pmatrix} 0 & -\frac{1}{6} & \frac{1}{6} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & \frac{2}{3} & \frac{1}{3} \end{pmatrix}$

Die darstellende Matrix von A bezüglich der Basis aus Eigenvektoren hat dann Diagonalform:

$$\begin{aligned} \mathrm{DM}_C(A) = T_C^K \cdot A \cdot T_K^C &= \begin{pmatrix} 0 & -\frac{1}{6} & \frac{1}{6} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & \frac{2}{3} & \frac{1}{3} \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & 1 \\ -3 & 1 & 0 \\ 3 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -\frac{1}{6} & \frac{1}{6} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & \frac{2}{3} & \frac{1}{3} \end{pmatrix} \cdot \begin{pmatrix} -2 & 0 & 1 \\ 6 & 0 & 0 \\ -6 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

Wir sehen, dass wie erwartet in $\mathrm{DM}_C(A)$ die Eigenwerte auf der Diagonalen stehen.

⋆ ⋆ ⋆

Jetzt sind wir in der Lage ein vollständiges Verfahren in Einzelschritten zum Diagonalisieren einer Matrix A bzw. eines Endomorphismus $f : V \to V$ anzugeben.

Algorithmus zur Diagonalisierung

(0) Wählen einer Basis B von V und Darstellung des Endomorphismus f als $A = \mathrm{DM}_B(f)$.

(1) Bestimmen des charakteristischen Polynoms der Matrix A (und somit auch des Endomorphismus f).

(2) **Abbruchkriterium I**:
Zerfällt das charakteristische Polynom nicht in Linearfaktoren, so ist die Matrix A nicht diagonalisierbar.

(3) Berechne für Eigenwerte (also Nullstellen des charakteristischen Polynoms) eine maximale Anzahl von linear unabhängigen Eigenvektoren im dazugehörigen Eigenraum (das heißt bestimme eine Basis des Eigenraums).

(4) **Abbruchkriterium II**:
Wenn diese gerade berechnete maximale Anzahl für einen k-fachen Eigenwert nicht gleich k ist, so ist die Matrix A nicht diagonalisierbar (Im anderen Fall haben wir jeweils eine hinreichend große Basis der Eigenräume gefunden!).

(5) Wenn wir alle in Schritt (3) berechneten Basen für die verschiedenen Eigenräume zusammen in eine Menge schreiben, haben wir die gewünschte Basis aus Eigenvektoren gefunden, bezüglich der die Matrix A Diagonalgestalt hat.

(6) Schreibe nacheinander für jeden Eigenwert die in Schritt (3) bestimmte Basis spaltenweise in eine Matrix: Das Ergebnis ist die gesuchte Transformationsmatrix, die die Ausgangsmatrix A in eine Diagonalgestalt überführt.

Die Stelligkeit eines Eigenwerts als Nullstelle innerhalb des dazugehörigen charakteristischen Polynoms bezeichnet man als seine *algebraische Vielfachheit.* Dagegen bezeichnet man die Dimension des Eigenraums eines Eigenwerts als seine *geometrische Vielfachheit.* Die geometrischen Vielfachheiten haben wir bereits in Definition 11.9 eingeführt. Nach den Überlegungen rund um den Satz 11.12 wissen wir bereits, dass die geometrischen Vielfachheiten höchstens gleich der algebraischen, mindestens jedoch positiv sein müssen, d.h. es gilt:

$$1 \leq \text{geometrische\_Vielfachheit}(\lambda) \leq \text{algebraische\_Vielfachheit}(\lambda).$$

Bemerkung 13.6: Beachten Sie, dass der Gauß-Algorithmus innerhalb der Diagonalisierung an verschiedenen Stellen eine wichtige und wertvolle Rolle spielt.
Beachten Sie aber auch die Details, so ändert sich der Betrag einer Determinante bei den üblichen Gaußschritten nicht, eine mit Gauß manipulierte Matrix wird in der Regel ihre Eigenwerte und -vektoren verändern.

⋆ ⋆ ⋆

Bisher haben wir den Begriff des charakteristischen Polynoms nur für Matrizen (also speziellen linearen Abbildungen) definiert. Um diesen nun auch für beliebige Endomorphismen betrachten zu können, ist es wichtig, dass verschiedene Darstellungen nicht zu verschiedenen Polynomen führen.

Dies ist insbesondere wichtig, wenn wir die Eigenwerte als Nullstellen der charakteristischen Polynome betrachten, denn dies darf dann nicht von der gewählten Basis in der Darstellung abhängen. Dies garantiert uns das nächste

Lemma 13.7:

Sei f ein Endomorphismus über dem Vektorraum V und B, C zwei Basen von V. Dann gilt: $p_{\mathrm{DM}_B(f)}(\lambda) = p_{\mathrm{DM}_C(f)}(\lambda)$.

Beweis: Es gilt:

$$\begin{aligned}
p_{\mathrm{DM}_B(f)}(\lambda) &= \det\big(\mathrm{DM}_B(f) - \lambda E_n\big) \\
&= \det\Big((T_C^B)^{-1} \cdot \mathrm{DM}_C(f) \cdot T_C^B - \lambda \cdot (T_C^B)^{-1} \cdot T_C^B\Big) \\
&= \det\Big((T_C^B)^{-1} \cdot (\mathrm{DM}_C(f) - \lambda \cdot E_n) \cdot T_C^B \Big) \\
&= \det\Big((T_C^B)^{-1}\Big) \cdot \det\big(\mathrm{DM}_C(f) - \lambda \cdot E_n\big) \cdot \det\Big(T_C^B\Big) \\
&= \underbrace{\det\Big((T_C^B)^{-1}\Big) \cdot \det\Big(T_C^B\Big)}_{=1} \cdot \det\big(\mathrm{DM}_C(f) - \lambda \cdot E_n\big) \\
&= \det\big(\mathrm{DM}_C(f) - \lambda \cdot E_n\big) \\
&= p_{\mathrm{DM}_C(f)}(\lambda)
\end{aligned}$$

wobei wir oben noch das Lemma 12.14 in der folgenden Form angewendet haben:

$$1 = \det\left(E_n\right) = \det\left(\left(T_C^B\right)^{-1} \cdot T_C^B\right) = \det\left(\left(T_C^B\right)^{-1}\right) \cdot \det\left(T_C^B\right).$$

$\boxtimes$

⋆ ⋆ ⋆

MEGA-Beispiel zur Diagonalisierbarkeit. Wir gehen nun anhand von vier Beispielen alle möglichen Fälle langsam durch, wobei wir nicht jeden Einzelschritt im Detail für alle Unterfälle durchrechnen werden:

Betrachte die folgenden Matrizen aus $\mathrm{Mat}(3 \times 3, \mathbb{R})$:

$$A_1 = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & -1 & -2 \\ 2 & 3 & 2 \\ 1 & 1 & 3 \end{pmatrix}$$

und den Endomorphismus $f : \mathbb{R}^3 \to \mathbb{R}^3$ gegeben durch

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} x_1 - 4x_2 + 8x_3 \\ -4x_1 + 7x_2 + 4x_3 \\ 8x_1 + 4x_2 + x_3 \end{pmatrix}.$$

Zu (0) Die ersten drei Endomorphismen haben wir bereits in Matrixform A_i gegeben. Für den vierten Endomorphismus bestimmen wir $\mathrm{DM}_K(f)$ bezüglich der kanonischen Basis K des $\mathbb{R}^3$:

$$f(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}) = \begin{pmatrix} 1 \\ -4 \\ 8 \end{pmatrix}, \quad f(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}) = \begin{pmatrix} -4 \\ 7 \\ 4 \end{pmatrix} \quad \text{und} \quad f(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}) = \begin{pmatrix} 8 \\ 4 \\ 1 \end{pmatrix}.$$

Daraus ergibt sich

$$A_4 := \mathrm{DM}_K(f) = \begin{pmatrix} 1 & -4 & 8 \\ -4 & 7 & 4 \\ 8 & 4 & 1 \end{pmatrix}.$$

Zu (1)

$$\begin{aligned} p_{A_1}(\lambda) &= \det(A_1 - \lambda E_3) = \begin{vmatrix} 3-\lambda & 0 & 0 \\ 0 & -\lambda & -1 \\ 0 & 1 & -\lambda \end{vmatrix} = (3-\lambda)\begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} \\ &= (3-\lambda)(\lambda^2+1) \end{aligned}$$

$$\begin{aligned} p_{A_2}(\lambda) &= \det(A_2 - \lambda E_3) = \begin{vmatrix} -1-\lambda & 0 & 0 \\ 0 & 1-\lambda & 1 \\ 0 & 0 & 1-\lambda \end{vmatrix} \\ &= (-1-\lambda)\begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix} = (-1-\lambda)(1-\lambda)^2 \end{aligned}$$

$$p_{A_3}(\lambda) = \det(A_3 - \lambda E_3) = \begin{vmatrix} -\lambda & -1 & -2 \\ 2 & 3-\lambda & 2 \\ 1 & 1 & 3-\lambda \end{vmatrix} = (1-\lambda)(2-\lambda)(3-\lambda)$$

$$p_{A_4}(\lambda) = \det(A_4 - \lambda E_3) = \begin{vmatrix} 1-\lambda & -4 & 8 \\ -4 & 7-\lambda & 4 \\ 8 & 4 & 1-\lambda \end{vmatrix} = (-9-\lambda)(9-\lambda)^2$$

Zu (2) Matrix $\mathbf{A_1}$ besteht den Test nicht, da (λ^2+1) über $\mathbb{R}$ nicht weiter zerfällt. Die anderen drei Matrizen bleiben im Rennen.

Zu (3) Für die verbleibenden Matrizen berechnen wir jeweils die Dimension der jeweiligen Eigenräume:

Die Matrix $\mathbf{A_2}$ hat die Eigenwerte -1 (mit algebraischer Vielfachheit 1) und 1 (mit algebraischer Vielfachheit 2).

$\boxed{\lambda_1 = -1:}$ Löse das homogene lineare Gleichungssystem

$$(A_2 + 1 \cdot E_3)x \quad = \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \cdot x \quad = \quad 0.$$

Der Vektor $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ bildet eine Basis des Eigenraums.

$\boxed{\lambda_2 = 1:}$ Löse das homogene lineare Gleichungssystem

$$(A_2 - 1 \cdot E_3)x \quad = \quad \begin{pmatrix} -2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \cdot x \quad = \quad 0.$$

Der Vektor $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ bildet eine Basis des Eigenraums.

Die Matrix $\mathbf{A_3}$ hat die Eigenwerte 1, 2 und 3 (jeweils mit algebraischer Vielfachheit 1).

$\boxed{\lambda_1 = 1:}$ Löse das homogene lineare Gleichungssystem

$$(A_3 - 1 \cdot E_3)x \quad = \quad \begin{pmatrix} -1 & -1 & -2 \\ 2 & 2 & 2 \\ 1 & 1 & 2 \end{pmatrix} \cdot x \quad = \quad 0.$$

Der Vektor $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$ bildet eine Basis des Eigenraums.

$\boxed{\lambda_2 = 2:}$ Löse das homogene lineare Gleichungssystem

$$(A_3 - 2 \cdot E_3)x \quad = \quad \begin{pmatrix} -2 & -1 & -2 \\ 2 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix} \cdot x \quad = \quad 0.$$

Der Vektor $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ bildet eine Basis des Eigenraums.

$\boxed{\lambda_3 = 3:}$ Löse das homogene lineare Gleichungssystem

$$(A_3 - 3 \cdot E_3)x \quad = \quad \begin{pmatrix} -3 & -1 & -2 \\ 2 & 0 & 2 \\ 1 & 1 & 0 \end{pmatrix} \cdot x \quad = \quad 0.$$

Der Vektor $\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$ bildet eine Basis des Eigenraums.

Die Matrix $\mathbf{A_4}$ hat die Eigenwerte -9 (mit algebraischer Vielfachheit 1) und 9 (mit algebraischer Vielfachheit 2).

$\boxed{\lambda_1 = -9:}$ Löse das homogene lineare Gleichungssystem

$$(A_4 + 9 \cdot E_3)x \quad = \quad \begin{pmatrix} 10 & -4 & 8 \\ -4 & 16 & 4 \\ 8 & 4 & 10 \end{pmatrix} \cdot x \quad = \quad 0.$$

Der Vektor $\begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}$ bildet eine Basis des Eigenraums.

$\boxed{\lambda_2 = 9:}$ Löse das homogene lineare Gleichungssystem

$$(A_4 - 9 \cdot E_3)x \quad = \quad \begin{pmatrix} -8 & -4 & 8 \\ -4 & -2 & 4 \\ 8 & 4 & -8 \end{pmatrix} \cdot x \quad = \quad 0$$

Die Vektoren $\begin{pmatrix}1\\2\\2\end{pmatrix}$ und $\begin{pmatrix}-2\\2\\-1\end{pmatrix}$ bilden eine Basis des Eigenraumes.

Zu (4) Die Matrix A_2 besteht den Test nicht, denn der zweifache Eigenwert $\lambda = 1$ besitzt nur einen eindimensionalen Eigenraum mit Basisvektor $\begin{pmatrix}0 & 1 & 0\end{pmatrix}^t$. Die Matrizen A_3 und A_4 bestehen den Test.

Zu (5) Für A_3 haben wir

$$C_3 \quad = \quad \left(\begin{pmatrix}1\\-1\\0\end{pmatrix}, \begin{pmatrix}1\\0\\-1\end{pmatrix}, \begin{pmatrix}-1\\1\\1\end{pmatrix}\right).$$

Für A_4 haben wir

$$C_4 \quad = \quad \left(\begin{pmatrix}2\\1\\-2\end{pmatrix}, \begin{pmatrix}1\\2\\2\end{pmatrix}, \begin{pmatrix}-2\\2\\-1\end{pmatrix}\right).$$

Zu (6) In C_3 und C_4 sind die Vektoren bezüglich der kanonischen Basis gegeben. Wir definieren

$$T_3 \ := \ T_K^{C_3} \ = \ \begin{pmatrix}1 & 1 & -1\\-1 & 0 & 1\\0 & -1 & 1\end{pmatrix} \qquad \text{und} \qquad T_4 \ := \ T_K^{C_4} \ = \ \begin{pmatrix}2 & 1 & -2\\1 & 2 & 2\\-2 & 2 & -1\end{pmatrix}.$$

Damit erhalten wir wie gewünscht:

$$\begin{aligned} T_3^{-1} \cdot A_3 \cdot T_3 \quad &= \quad \begin{pmatrix}1 & 0 & 1\\1 & 1 & 0\\1 & 1 & 1\end{pmatrix} \cdot \begin{pmatrix}0 & -1 & -2\\2 & 3 & 2\\1 & 1 & 3\end{pmatrix} \cdot \begin{pmatrix}1 & 1 & -1\\-1 & 0 & 1\\0 & -1 & 1\end{pmatrix} \\ &= \quad \begin{pmatrix}1 & 0 & 0\\0 & 2 & 0\\0 & 0 & 3\end{pmatrix} \end{aligned}$$

$$\begin{aligned} T_4^{-1} \cdot A_4 \cdot T_4 \quad &= \quad \frac{1}{9}\begin{pmatrix}2 & 1 & -2\\1 & 2 & 2\\-2 & 2 & -1\end{pmatrix} \cdot \begin{pmatrix}1 & -4 & 8\\-4 & 7 & 4\\8 & 4 & 1\end{pmatrix} \cdot \begin{pmatrix}2 & 1 & -2\\1 & 2 & 2\\-2 & 2 & -1\end{pmatrix} \\ &= \quad \begin{pmatrix}-9 & 0 & 0\\0 & 9 & 0\\0 & 0 & 9\end{pmatrix} \end{aligned}$$

Gut gemeinter Hinweis: Nutzen Sie diese Beispiele und füllen Sie die Lücken, indem Sie Details durchrechnen. Dies übt und bereitet Sie auf die Klausur vor!

Beispiel 13.8: Betrachten Sie die folgende Matrix und lassen sich *nicht* von den darin enthaltenen Parametern a und b abschrecken:

$$A := \begin{pmatrix}-3 & 0 & 0\\2a & b & a\\10 & 0 & 2\end{pmatrix}$$

Wir überlegen uns, unter welchen Bedingungen diese Matrix diagonalisierbar ist.

Dazu untersuchen wir die Matrix A nach der gegebenen Prozedur und berechnen zunächst das charakteristische Polynom:

$$p_A(\lambda) = \det(A - \lambda E_3) = \det\begin{pmatrix} -3-\lambda & 0 & 0 \\ 2a & b-\lambda & a \\ 10 & 0 & 2-\lambda \end{pmatrix} = (-3-\lambda)(b-\lambda)(2-\lambda).$$

Somit sehen wir sofort, dass die Matrix für $b \neq -3$ und $b \neq 2$ schon einmal diagonalisierbar sein muss, denn in diesem Fall hätten wir drei verschiedene Eigenwerte. Aber vielleicht gibt es auch noch zusätzliche Möglichkeiten für diese beiden Ausschlussfälle, die wir uns gesondert überlegen müssen:

$\boxed{\textbf{Fall } b=-3}$ In diesem Fall gilt $p_A(\lambda) = (-3-\lambda)^2 \cdot (2-\lambda)$ und wir müssen den Eigenraum zum (doppelten) Eigenwert $\lambda = -3$ untersuchen:

Betrachte also $E(-3) = \mathrm{Kern}(A + 3E_3)$ für $b = -3$ und bestimme die Dimension des Lösungsraums des homogenen Gleichungssystems:

$$\begin{pmatrix} 0 & 0 & 0 \\ 2a & 0 & a \\ 10 & 0 & 5 \end{pmatrix} \quad \cdots \text{Gauß-Algorithmus} \cdots \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

Somit ist die Dimension des Eigenraums wünschenswert gleich 2 und die Matrix ist für den Fall $b = -3$ diagonalisierbar.

$\boxed{\textbf{Fall } b=2}$ In diesem Fall gilt $p_A(\lambda) = (-3-\lambda)(2-\lambda)^2$ und wir müssen den Eigenraum zum (doppelten) Eigenwert $\lambda = 2$ untersuchen:

Betrachte also $E(2) = \mathrm{Kern}(A - 2E_3)$ für $b = 2$ und bestimme die Dimension des Lösungsraums des homogenen Gleichungssystems:

$$\begin{pmatrix} -5 & 0 & 0 \\ 2a & 0 & a \\ 10 & 0 & 0 \end{pmatrix} \quad \cdots \text{Gauß-Algorithmus} \cdots \quad \begin{pmatrix} -5 & 0 & 0 \\ 0 & 0 & a \\ 0 & 0 & 0 \end{pmatrix}$$

Somit ist die Dimension des Eigenraums genau dann gleich 2, wenn $a = 0$ gilt. Somit ist die Matrix für den Fall $b = 2$ nur für $a = 0$ diagonalisierbar.

Zusammenfassend können wir festhalten: Die Matrix A ist genau dann diagonalisierbar, wenn $b \neq 2$ oder $a = 0$ ist. War also gar nicht so schwer.

Hiermit sind wir dann auch am Ende der Betrachtungen rund um das Diagonalisieren angekommen. Sie sehen, dass wir für diesen Zweck das gesamte Wissen der bisherigen Vorlesung benötigen.

⋆ ⋆ ⋆

Einschiebend bringen wir noch ein kleines (analytisches) Beispiel aus der Welt der Zahlenfolgen. Die Theorie des Diagonalisierens kann hier helfen, explizite Darstellungen zu finden.

Beispiel 13.9: Fibonacci-Zahlen. Bei den Fibonacci-Zahlen handelt es sich um eine sehr wichtige rekursiv definierte Zahlenfolge $(f_i)_{i=1,\ldots,n}$, die Sie vielleicht aus der Analysis kennen. Man definiert diese Folge dadurch, dass man immer die zwei letzten Folgenglieder addiert, um das nächte Folgenglied zu erhalten:

$$f_0 := 0 \quad f_1 := 1 \quad f_2 := 1 \quad f_{n+2} := f_{n+1} + f_n$$

Dieses Wachstumsverhältnis für die f_i fassen wir jetzt in Matrixschreibweise wie folgt zusammen:

$$A := \begin{pmatrix} f_2 & f_1 \\ f_1 & f_0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$$

Diese rekursive Berechnung der Fibonacci-Zahlen kann sehr aufwendig sein und wir sind daher an einer expliziten Formel interessiert, die uns die $(n+1)$-te Fibonacci-Zahl liefert, ohne alle vorherigen Fibonacci-Zahlen ausrechnen zu müssen.

Wir gehen dazu wie folgt vor – es gilt:

$$\begin{aligned}\begin{pmatrix} f_{(n+1)+2} & f_{(n+1)+1} \\ f_{(n+1)+1} & f_{(n+1)} \end{pmatrix} &= \begin{pmatrix} f_{n+2}+f_{n+1} & f_{n+2} \\ f_{n+1}+f_n & f_{n+1} \end{pmatrix} &&= \begin{pmatrix} f_{(n)+2} & f_{(n)+1} \\ f_{(n)+1} & f_{(n)} \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \\ &= \ldots &&= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{n+2}\end{aligned}$$

Der Trick besteht jetzt darin, den letzten Ausdruck mithilfe unserer Diagonalisierungskenntnisse auszurechen – es gilt:

$$p_A(\lambda) = \left| \begin{pmatrix} 1-\lambda & 1 \\ 1 & -\lambda \end{pmatrix} \right| = \lambda^2 - \lambda - 1$$

mit den beiden Nullstellen $\alpha := \frac{1}{2} + \frac{\sqrt{5}}{2}$ und $\beta := \frac{1}{2} - \frac{\sqrt{5}}{2}$.

Jetzt wissen wir auch schon automatisch, dass die Matrix diagonalisierbar ist und rechnen die Eigenvektoren aus.

Das α berechnet man wie folgt:

$$\operatorname{Kern}\left(\begin{pmatrix} 1-\alpha & 1 \\ 1 & -\alpha \end{pmatrix}\right) = \operatorname{Kern}\left(\begin{pmatrix} 0 & 0 \\ 1 & -\alpha \end{pmatrix}\right) = \left\{ v \mid v = \lambda \begin{pmatrix} 1 \\ \frac{1}{\alpha} \end{pmatrix},\ \lambda \in \mathbb{R} \right\}$$

und für β gilt:

$$\operatorname{Kern}\left(\begin{pmatrix} 1-\beta & 1 \\ 1 & -\beta \end{pmatrix}\right) = \operatorname{Kern}\left(\begin{pmatrix} 0 & 0 \\ 1 & -\beta \end{pmatrix}\right) = \left\{ v \mid v = \lambda \begin{pmatrix} 1 \\ \frac{1}{\beta} \end{pmatrix},\ \lambda \in \mathbb{R} \right\}$$

Wir können nun einen Basiswechsel vollziehen und erhalten folgenden Zusammenhang:

$$A = T \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} T^{-1} = \begin{pmatrix} 1 & 1 \\ \frac{1}{\alpha} & \frac{1}{\beta} \end{pmatrix} \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \frac{-1}{\sqrt{5}} \begin{pmatrix} \frac{1}{\beta} & -1 \\ -\frac{1}{\alpha} & 1 \end{pmatrix}$$

daraus folgt speziell:

$$\begin{aligned} A^{n+2} &= \left(T \begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} T^{-1} \right)^{n+2} &&= T \begin{pmatrix} \alpha^{n+2} & 0 \\ 0 & \beta^{n+2} \end{pmatrix} T^{-1} \\ &= \begin{pmatrix} 1 & 1 \\ \frac{1}{\alpha} & \frac{1}{\beta} \end{pmatrix} \begin{pmatrix} \alpha^{n+2} & 0 \\ 0 & \beta^{n+2} \end{pmatrix} \frac{-1}{\sqrt{5}} \begin{pmatrix} \frac{1}{\beta} & -1 \\ -\frac{1}{\alpha} & 1 \end{pmatrix} \end{aligned}$$

Berechnen des Matrixprodukts liefert uns nun insgesamt:

$$A^{n+2} = \frac{-1}{\sqrt{5}} \begin{pmatrix} 1 & 1 \\ \frac{1}{\alpha} & \frac{1}{\beta} \end{pmatrix} \begin{pmatrix} \frac{\alpha^{n+2}}{\beta} & -\alpha^{n+2} \\ -\frac{\beta^{n+2}}{\alpha} & \beta^{n+2} \end{pmatrix} = \frac{1}{\sqrt{5}} \begin{pmatrix} -\frac{\alpha^{n+2}}{\beta} + \frac{\beta^{n+2}}{\alpha} & \alpha^{n+2} - \beta^{n+2} \\ \frac{\beta^{n+1}}{\alpha} - \frac{\alpha^{n+1}}{\beta} & \alpha^{n+1} - \beta^{n+1} \end{pmatrix}$$

Zum Schluss können wir schließlich den Wert von f_{n+1} aus dieser Matrix direkt ablesen und erhalten das gesuchte Ergebnis, nämlich eine explizite Darstellung der Fibonacci-Zahlen:

$$f_{n+1} = \frac{\alpha^{n+1} - \beta^{n+1}}{\sqrt{5}}$$

⋆ ⋆ ⋆

Ausblick Hauptachsentransformation. Abschließend möchte ich Ihnen noch einen kleinen Ausblick geben, was noch alles möglich ist, wir aber an dieser Stelle zeitlich nicht leisten können.

In der Mathematik spricht man bei Basiswechseln auch häufig von *Hauptachsentransformationen*. Stellen Sie sich vor, Sie betrachten Punkte in der Ebene durch Angabe (nichtlinearer) Gleichungen, wie beispielsweise der folgenden:

$$-\frac{7}{2}x^2 - \frac{7}{2}y^2 + 25xy + 30\sqrt{2}x - 66\sqrt{2}y = 256$$

Diese beschriebene Punktmenge der Ebene stellt eine Kurve dar. Genauer ausgedrückt handelt es sich um um eine Hyperbel und im gewohnten x-y-Koordinatensystem hat sie etwa die Gestalt:

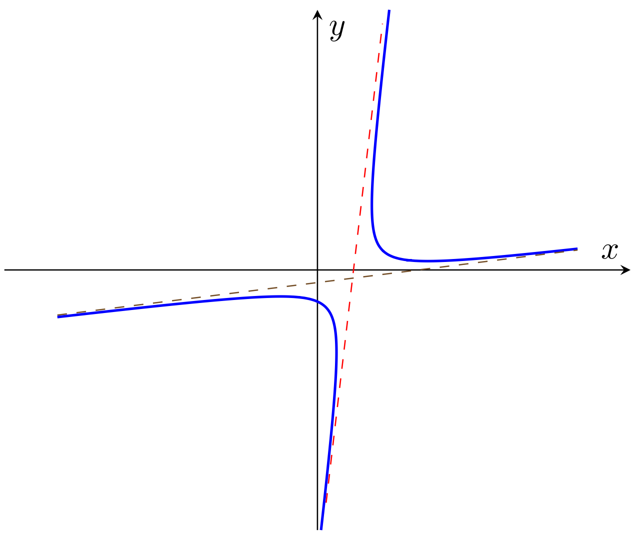


Abbildung 20

Wenn Sie jetzt Ihren Blickwinkel ändern, das heißt, das Koordinatensystem stauchen, strecken, verschieben und drehen, dann können Sie diese Kurve in einem besseren Licht darstellen; anders ausgedrückt wird die Darstellung dieser Kurve mit einfacheren Mitteln möglich.

Betrachten Sie dazu die Abbildung (und drehen Sie Ihren Kopf um 45 Grad nach links):

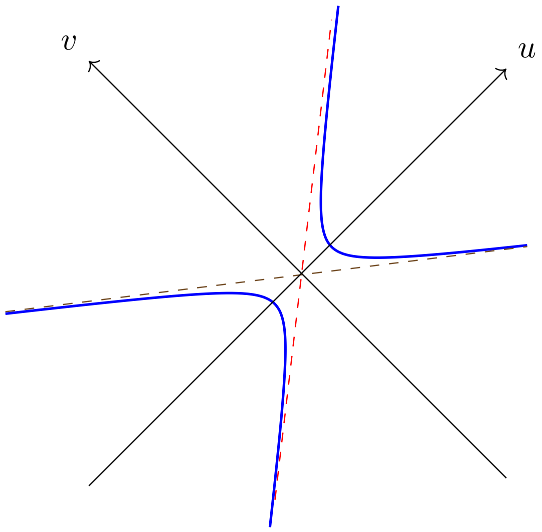


Abbildung 21

Mithilfe der Eigenwerttheorie ist es möglich, in diesem Fall von der kanonischen Basis zur neuen Basis $\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right\}$ zu kommen. Wir führen dies hier exemplarisch vor, ohne die einzelnen Schritte im Detail zu erklären.

Die Ähnlichkeit zum Diagonalisieren ist allerdings unverkennbar – schauen Sie selbst! Zunächst stellen wir die gegebene Punktmenge mittels Matrizenschreibweise dar.

Es gilt für $A = \begin{pmatrix} -\frac{7}{2} & \frac{25}{2} \\ \frac{25}{2} & -\frac{7}{2} \end{pmatrix}$ und $a = \begin{pmatrix} 30\sqrt{2} \\ -66\sqrt{2} \end{pmatrix}$:

$$P(x,y) \;=\; (x,y) \cdot A \cdot \begin{pmatrix} x \\ y \end{pmatrix} + a^t \cdot \begin{pmatrix} x \\ y \end{pmatrix} \;=\; 256$$

Beim Diagonalisieren von A erhalten wir dann:

$$\begin{aligned} 0 &= P(\lambda) = \det\left(A - \lambda \cdot \mathrm{Id}\right) \\ &= \left(\frac{7}{2} + \lambda\right)^2 - \left(\frac{25}{2}\right)^2 \end{aligned}$$

Für die Eigenwerte $\lambda_1 = 9$ und $\lambda_2 = -16$. Die zugehörige Basis aus Eigenvektoren berechnen wir jetzt mittels der Kerne:

$$\mathrm{Kern}_{\lambda_1} \begin{pmatrix} -\frac{25}{2} & \frac{25}{2} \\ \frac{25}{2} & -\frac{25}{2} \end{pmatrix} = \mu \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\mathrm{Kern}_{\lambda_2} \begin{pmatrix} \frac{25}{2} & \frac{25}{2} \\ \frac{25}{2} & \frac{25}{2} \end{pmatrix} = \mu \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Nach Normieren der Eigenvektoren ergibt sich für den Basiswechel die Matrix:

$$T = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

und es gilt demnach:

$$T^{-1} \cdot A \cdot T \;=\; T^t \cdot A \cdot T \;=\; \begin{pmatrix} 9 & 0 \\ 0 & -16 \end{pmatrix}$$

Die Matrix T entspricht einer Drehung um 45 Grad des Koordinatensystems:

$$\begin{aligned} P(T(x,y)) &= (T(x,y))^T \cdot A \cdot T(x,y) + a^t \cdot T(x,y) \\ &= (x,y)^T \cdot T^t \cdot A \cdot T \cdot (x,y) + a^t \cdot T(x,y) \\ &= 9x^2 - 16y^2 + (30\sqrt{2}, -66\sqrt{2}) \cdot \frac{1}{\sqrt{2}} \begin{pmatrix} x-y \\ x+y \end{pmatrix} \\ &= 9x^2 - 16y^2 - 36x - 96y \end{aligned}$$

Als Nächstes müssen wir nur noch die linearen Terme, mithilfe der quadratischen Ergänzung, durch eine weitere Substitution beseitigen:

$$\begin{aligned} P\left(T(x,y)\right) &= 9x^2 - 16y^2 - 36x - 96y \\ &= 9\underbrace{\left(x^2 - 4x + 4\right)}_{=(x-2)^2} - 36 - 16\underbrace{\left(y^2 + 6y + 9\right)}_{=(y+3)^2} + 144 \\ &= 9u^2 - 16v^2 + 108 \end{aligned}$$

Für die lineare Transformation $u = x - 2$ und $v = y + 3$. Insgesamt haben wir die Kurve mittels einer Drehung und einer linearen Verschiebung demnach in die folgende Normalform gebracht:

$$P\left(\,T(x+2, y-3)\right) = 9u^2 - 16v^2 \;=\; 256 - 108 = 148$$

Unter den berechneten Hauptachsen, die ich jetzt nicht mehr x und y, sondern u und v nenne, hat die Kurve also die Gestalt:

$$9u^2 - 16v^2 = 148$$

Das ist offenbar eine gravierende Vereinfachung der Darstellung. Dies war durch eine Verschiebung des Koordinatenursprungs bei gleichzeitiger Drehung um 45 Grad des gesamten Koordinatensystems möglich.

⋆ ⋆ ⋆

Aufgabe 13.10: Finden Sie einen Endomorphismus $f : \mathbb{R}^3 \to \mathbb{R}^3$, für den $p_A(\lambda) = -\lambda^3 + 2\lambda^2 + 3\lambda + 2$ gilt.
Kann man für jedes Polynom

$$p(\lambda) = \pm\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_0$$

einen Endomorphismus f finden, so dass $p_A(\lambda) = p(\lambda)$ gilt?

14. Euklidische Vektorräume

Nach dem langen Weg bis einschließlich zum letzten Kapitel mit dem Höhepunkt des Algorithmus zum Diagonalisieren von Endomorphismen, schlagen wir nun eine neue Richtung unserer Betrachtungen ein.

In vielen Anwendungen besitzen die betrachteten Vektorräume zusätzliche Struktur, die wir zu unserem Vorteil ausnutzen können. Bei reellen Vektorräumen, das heißt bei Vektorräumen zum Grundkörper $\mathbb{R}$, sind besonders die so genannten *Skalarprodukte* interessant, mit denen sich dann geometrischen Begriffe wie Abstand, Winkel oder Orthogonalität definieren lassen.

⋆ ⋆ ⋆

Geometrische Begriffe in der reellen Ebene. Wir können dieses Prinzip erneut am Beispiel der wohlbekannten reellen Ebene motivieren. Im Folgenden werden wir die bekannten Begriffe wie Abstand, Winkel oder Orthogonalität aus einem (scheinbar) neuen, speziell definiertem Produkt von zwei Vektoren schrittweise herleiten. Wir starten mit der ersten

Definition 14.1 – Standard-Skalarprodukt:

Sei $m \in \mathbb{N} \setminus \{0\}$ gegeben. Definiere eine zweistellige Funktion $\langle \cdot, \cdot \rangle^{\text{st}} : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ durch

$$\langle x, y \rangle^{\text{st}} := \left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \right\rangle^{\text{st}} := x_1 y_1 + \ldots + x_m y_m.$$

Die Funktion $\langle \cdot, \cdot \rangle^{\text{st}}$ heißt **Standard-Skalarprodukt** auf dem $\mathbb{R}^m$.

Es gilt nun in der reellen Ebene:

$$\left\langle \begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix} \right\rangle^{\text{st}} = ac + bd = \begin{vmatrix} a & -d \\ b & c \end{vmatrix} = \pm \text{Flächeninhalt} \left(\left[\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} -d \\ c \end{pmatrix} \right] \right)$$

Dies ist der Flächeninhalt des Parallelogramms, das von $\begin{pmatrix} a \\ b \end{pmatrix}$ und $\begin{pmatrix} -d \\ c \end{pmatrix}$ aufgespannt wird, wie wir dies bereits im Vorfeld des Satzes 12.17 gesehen haben. Es gilt dabei, dass $\begin{pmatrix} -d \\ c \end{pmatrix}$ der um $\frac{\pi}{2}$ gedrehte Vektor $\begin{pmatrix} c \\ d \end{pmatrix}$ ist (vgl. Abbildung 22).

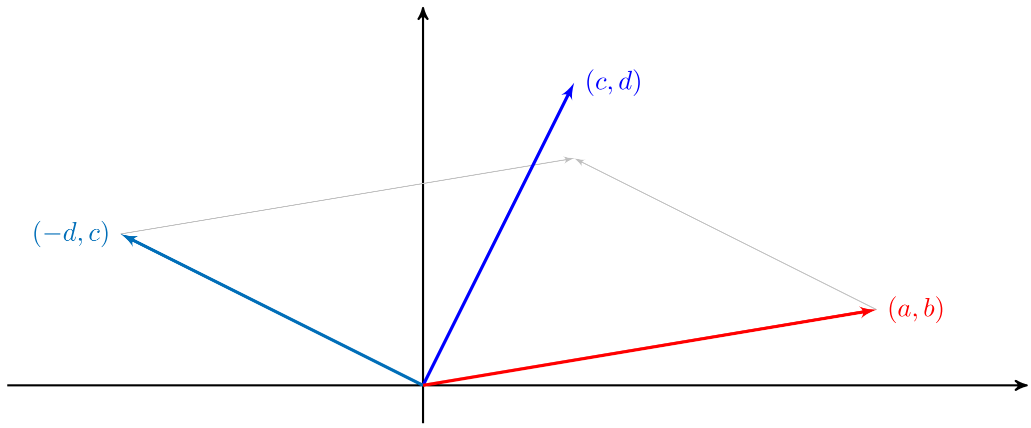


Abbildung 22: gedrehter Vektor

Mithilfe dieser Überlegung können wir schließlich die gewünschten Begriffe herleiten. Wie wir wissen, lässt sich der Flächeninhalt eines Parallelogramms bereits aus bekannten trigonometrischen Zusammenhängen berechnen:

$$\begin{aligned}
&\text{Flächeninhalt}\left(\left[\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} -d \\ c \end{pmatrix}\right]\right) \\
&\qquad = \text{Länge}\left(\begin{pmatrix} a \\ b \end{pmatrix}\right) \cdot \text{Länge}\left(\begin{pmatrix} c \\ d \end{pmatrix}\right) \cdot \sin\angle\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} -d \\ c \end{pmatrix}\right) \\
&\qquad = \text{Länge}\left(\begin{pmatrix} a \\ b \end{pmatrix}\right) \cdot \text{Länge}\left(\begin{pmatrix} c \\ d \end{pmatrix}\right) \cdot \sin\left(\angle\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix}\right) + \frac{\pi}{2}\right) \\
&\qquad = \text{Länge}\left(\begin{pmatrix} a \\ b \end{pmatrix}\right) \cdot \text{Länge}\left(\begin{pmatrix} c \\ d \end{pmatrix}\right) \cdot \cos\angle\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix}\right)
\end{aligned}$$

Im Spezialfall, dass $\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} c \\ d \end{pmatrix}$ ist, erhalten wir dann

$$\left\langle \begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} a \\ b \end{pmatrix} \right\rangle^{\text{st}} = \text{Länge}^2\left(\begin{pmatrix} a \\ b \end{pmatrix}\right)$$

und somit können wir den Begriff *Länge* aus dem Skalarprodukt ableiten:

$$\text{Länge}(a, b) = \sqrt{\left\langle \begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} a \\ b \end{pmatrix} \right\rangle^{\text{st}}}$$

Darüber hinaus kann man auch den Begriff des *Winkels* zwischen den beiden Vektoren $\begin{pmatrix} a \\ b \end{pmatrix}$ und

$\begin{pmatrix} c \\ d \end{pmatrix}$ bestimmen mit:

$$\cos\angle\left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix}\right) = \frac{\left\langle \begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix} \right\rangle^{\text{st}}}{\text{Länge}\left(\begin{pmatrix} a \\ b \end{pmatrix}\right) \cdot \text{Länge}\left(\begin{pmatrix} c \\ d \end{pmatrix}\right)}$$
$$= \frac{\left\langle \begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix} \right\rangle^{\text{st}}}{\sqrt{\left\langle \begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} a \\ b \end{pmatrix} \right\rangle^{\text{st}}} \cdot \sqrt{\left\langle \begin{pmatrix} c \\ d \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix} \right\rangle^{\text{st}}}}$$

Mithilfe des Winkels lässt sich dann natürlich auch der Begriff der Orthogonalität ableiten, in dem man nach einem eingeschlossenen rechten Winkel fragt.

⋆ ⋆ ⋆

Skalarprodukte. Dieses Prinzip der Herleitung der verschiedenen geometrischen Begriffe motiviert uns, dies auch für allgemeine Räume machen zu können. Wir charakterisieren daher allgemein über zu erfüllende Axiome einen Begriff von Skalarprodukt und arbeiten uns langsam vor:

Definition 14.2 – Skalarprodukt:

Sei V ein reeller Vektorraum und $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ eine Funktion. Dann ist $\langle\cdot,\cdot\rangle$ ein Skalarprodukt, wenn die folgenden Axiome gelten:

(a) $\langle\cdot,\cdot\rangle$ ist linear im ersten Argument: Für alle $x, x', y \in V$ und $\lambda, \lambda' \in \mathbb{R}$ ist

$$\left\langle \lambda x + \lambda' x', y \right\rangle = \lambda \langle x, y\rangle + \lambda' \left\langle x', y\right\rangle.$$

(b) $\langle\cdot,\cdot\rangle$ ist linear im zweiten Argument: Für alle $x, y, y' \in V$ und $\lambda, \lambda' \in \mathbb{R}$ ist

$$\left\langle x, \lambda y + \lambda' y' \right\rangle = \lambda \langle x, y\rangle + \lambda' \left\langle x, y'\right\rangle.$$

(c) $\langle\cdot,\cdot\rangle$ ist symmetrisch: Für alle $x, y \in V$ ist

$$\langle x, y\rangle = \langle y, x\rangle.$$

(d) $\langle\cdot,\cdot\rangle$ ist positiv definit: Für alle $x \in V \setminus \{0\}$ ist

$$\langle x, x\rangle > 0.$$

Achtung!

Das Skalarprodukt $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{R}$ darf nicht mit der Skalarmultiplikation $\mathbb{R} \times V \to V\ ;\ (\lambda, v) \mapsto \lambda v$ verwechselt werden. Die Operationen unterscheiden sich bereits in Bezug auf die Definitions- und Wertebereiche.

Wir sehen sofort, dass das zunächst spezielle Standard-Skalarprodukt auch dieser allgemeinen Charakterisierung entspricht:

Bemerkung 14.3: Das Standard-Skalarprodukt auf $V = \mathbb{R}^m$ ist ein Skalarprodukt im Sinne der Definition 14.2

Schließlich können wir den Bogen zur Überschrift spannen:

Definition 14.4 – euklidischer Vektorraum:

Ein **euklidischer Vektorraum** ist ein Paar $\big(V, \langle\cdot,\cdot\rangle\big)$, das aus einem reellen Vektorraum V und einem Skalarprodukt $\langle\cdot,\cdot\rangle$ auf V besteht.

Wir beweisen einige Eigenschaften von Skalarprodukten:

Satz 14.5:

Sei $V = \big(V, \langle\cdot,\cdot\rangle\big)$ ein euklidischer Vektorraum. Dann gilt:

(a) Für alle $x \in V$ ist $\langle x, 0\rangle = \langle 0, x\rangle = 0$.
(b) Für alle $x \in V$ ist $\langle x, x\rangle = 0$ genau dann, wenn $x = 0$.

Beweis:

Zu (a): Aus der Linearität von $\langle\cdot,\cdot\rangle$ folgt: $\langle x, 0\rangle = \langle x, 0 \cdot 0\rangle = 0 \cdot \langle x, 0\rangle = 0$. Aus der Symmetrie von $\langle\cdot,\cdot\rangle$ folgt auch: $\langle x, 0\rangle = \langle 0, x\rangle$.

Achtung!

Eigentlich steht hier: $\langle x, 0_V\rangle = \langle x, 0_{\mathbb{R}} \cdot 0_V\rangle = 0_{\mathbb{R}} \cdot \langle x, 0_V\rangle = 0_{\mathbb{R}}$.

Zu (b): Aus (a) und der Eigenschaft (b) aus der Definition 14.2 (positive Definitheit) ist die Äquivalenz klar.

⊠

⋆ ⋆ ⋆

Normen. Wir führen als Nächstes allgemein den Begriff der Länge ein, den wir an dieser Stelle allerdings anders bezeichnen:

Definition 14.6 – Norm:

Sei $V = \big(V, \langle\cdot,\cdot\rangle\big)$ ein euklidischer Vektorraum. Die **Norm** auf V ist die Funktion

$$\|\cdot\| : V \to \mathbb{R}\ ;\ \|x\| := +\sqrt{\langle x, x\rangle}.$$

Diese Definition entspricht gerade der gesehenen Erfahrung im $\mathbb{R}^2$. Wir werden allerdings auch allgemein zeigen, dass eine aus einem Skalarprodukt definierte Norm die Grundeigenschaft einer

„Längenfunktion“ besitzt. Zunächst dazu eine wichtige und hilfreiche Ungleichung:

Satz 14.7 – Cauchy-Schwarz-Ungleichung:

Sei $V = \big(V, \langle\cdot,\cdot\rangle\big)$ ein euklidischer Vektorraum und $x, y \in V$. Dann gilt die Cauchy-Schwarzsche Ungleichung

$$\big|\langle x, y\rangle\big| \leq \|x\| \cdot \|y\|\,.$$

Beweis: Es gilt:

$$\begin{aligned} 0 &\leq \big\langle \langle y,y\rangle \cdot x - \langle y,x\rangle \cdot y, \langle y,y\rangle \cdot x - \langle y,x\rangle \cdot y \big\rangle \quad \text{(pos. Definitheit)} \\ &= \langle y,y\rangle^2 \langle x,x\rangle - \langle y,y\rangle \langle y,x\rangle \langle x,y\rangle \underbrace{- \langle y,x\rangle \langle y,y\rangle \langle y,x\rangle + \langle y,x\rangle^2 \langle y,y\rangle}_{=0} \\ &= \langle y,y\rangle^2 \langle x,x\rangle - \langle y,y\rangle \langle y,x\rangle^2 \end{aligned}$$

Fall 1 $\boxed{y = 0}$ Dann gilt nach Satz 14.5 (a) offenbar

$$\big|\langle x, y\rangle\big| = \big|\langle x, 0\rangle\big| = 0 \leq \|x\| \cdot \|y\|\,.$$

Fall 2 $\boxed{y \neq 0}$ Dann ist $\langle y,y\rangle > 0$ und wir können die obige Ungleichung durch $\langle y,y\rangle$ dividieren und auf beiden Seiten die positive Quadratwurzel ziehen:

$$\begin{aligned} 0 &\leq \langle y,y\rangle \langle x,x\rangle - \langle y,x\rangle^2 \\ \langle y,x\rangle^2 &\leq \langle y,y\rangle \langle x,x\rangle \\ \big|\langle y,x\rangle\big| &\leq \sqrt{\langle y,y\rangle}\,\sqrt{\langle x,x\rangle} = \|y\| \cdot \|x\|\,. \end{aligned}$$

⊠

Die Cauchy-Schwarzsche Ungleichung folgt für das Standard-Skalarprodukt auch aus den anfänglichen geometrischen Überlegungen, denn $\big|\langle x, y\rangle\big|$ ist der Flächeninhalt eines Parallelogramms mit den Seitenlängen $\|x\|$ und $\|y\|$, also ist $\big|\langle x, y\rangle\big|$ höchstens gleich $\|x\| \cdot \|y\|$, denn

$$\pm\langle x, y\rangle = \text{Flächeninhalt}(a, b) = \|x\| \cdot \|y\| \cdot \cos(x, y).$$

Der Kosinus ist vom Betrag her höchstens 1.

Mittels Satz 14.7 können wir nun grundlegende Längeneigenschaften der Norm nachweisen.

Satz 14.8:

Sei $V = (V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum mit Norm $\|\cdot\|$. Dann gilt:

(a) Für alle $x \in V$ ist $\|x\| \geq 0$.
(b) Für alle $x \in V$ ist $\|x\| = 0$ genau dann, wenn $x = 0$.
(c) Für alle $x \in V$, $\lambda \in \mathbb{R}$ ist $\|\lambda x\| = |\lambda| \cdot \|x\|$.
(d) Für alle $x, y \in V$ gilt die Dreiecksungleichung

$$\|x + y\| \leq \|x\| + \|y\|\,.$$

Beweis:

Zu (a): Betrachte $x \in V$. Dann gilt: $\langle x,x\rangle \geq 0$ und $\|x\| = +\sqrt{\langle x,x\rangle} \geq 0$.

Zu (b): Klar, nach Satz 14.5(b).

Zu (c): Betrachte $x \in V$, $\lambda \in \mathbb{R}$. Dann ist

$$\begin{aligned}\|\lambda \cdot x\| &= \sqrt{\langle \lambda x, \lambda x\rangle} = \sqrt{\lambda \langle x, \lambda x\rangle} = \sqrt{\lambda^2 \langle x, x\rangle} \\ &= \sqrt{\lambda^2}\,\sqrt{\langle x, x\rangle} = |\lambda|\,\|x\|\end{aligned}$$

Zu (d): Betrachte $x, y \in V$. Dann ist nach der Cauchy-Schwarzschen Ungleichung aus Satz 14.7:

$$\begin{aligned}\left(\|x\| + \|y\|\right)^2 &= \|x\|^2 + 2 \cdot \|x\| \cdot \|y\| + \|y\|^2 \\ &\geq \langle x, x\rangle + 2 \cdot \langle x, y\rangle + \langle y, y\rangle \\ &= \langle x + y, x + y\rangle \\ &= \|x + y\|^2\end{aligned}$$

Ziehen wir auf beiden Seiten die positive Quadratwurzel, so erhalten wir wie gewünscht: $\|x\| + \|y\| \geq \|x + y\|$.

⊠

Die Dreiecksungleichung entspricht der Tatsache, dass eine Gerade die kürzeste Verbindung zweier Punkte darstellt. Für das von den Punkten $0, x, x+y$ gebildete Dreieck ist die Seite von 0 nach $x+y$ kürzer als die Summe der Längen der Seiten von 0 nach x und der Seite von x nach $x+y$, diese Längen sind $\|x+y\|$, $\|x\|$ und $\|y\|$, sodass gilt:

$$\|x + y\| \leq \|x\| + \|y\|.$$

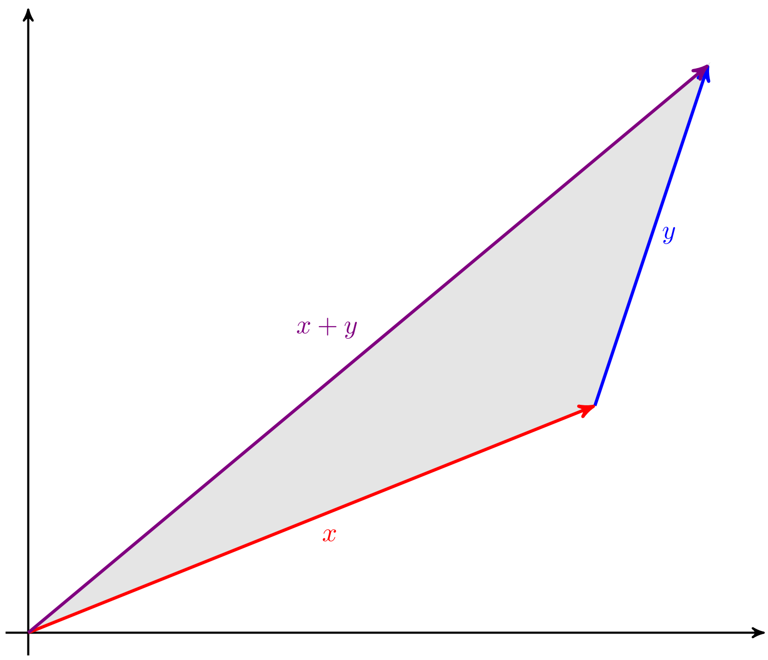


Abbildung 23: Dreiecksungleichung

⋆ ⋆ ⋆

Orthogonalität. Zwei Vektoren sind in besonderem Maße linear unabhängig, wenn sie orthogonal zueinander stehen. Geometrisch bedeutet dies, dass sie einen Winkel von $\frac{\pi}{2} = 90°$ bilden. Im Sinne unserer anfänglichen Überlegungen zu euklidischen Vektorräumen können wir allgemein definieren:

Definition 14.9 – Öffnungswinkel:

Sei $V = (V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum und $x, y \in V \setminus \{0\}$. Der Öffnungswinkel $\alpha(x, y)$ zwischen x und y ist definiert durch

$$\cos\big(\alpha\,(x, y)\big) = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}\ ; \quad 0 \le \alpha(x, y) \le \pi.$$

Nach der Cauchy-Schwarzschen Ungleichung aus Satz 14.7 ist

$$-1 \le \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \le 1$$

und somit ist der Kosinus des Öffnungswinkels wohldefiniert.

Anschaulich liegen zwei Vektoren orthogonal, wenn sie einen Öffnungswinkel von $\frac{\pi}{2}$ einschließen. Hierfür ist notwendig, dass $\cos\big(\alpha(x, y)\big) = \cos\big(\frac{\pi}{2}\big)$ und somit $\langle x, y \rangle = 0$.

Definition 14.10 – Orthogonalität:

Sei $V = (V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum und $x, y \in V$. Die Vektoren x, y heißen orthogonal oder senkrecht zueinander (kurz $x \perp y$), wenn $\langle x, y \rangle = 0$.

Beispiel 14.11: Eine Drehung des $\mathbb{R}^2$ um den Winkel $\frac{\pi}{2}$ wird durch die Drehmatrix $D_{\frac{\pi}{2}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ dargestellt. Eine Drehung des Vektors $\begin{pmatrix} a \\ b \end{pmatrix}$ führt zu dem Vektor

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -b \\ a \end{pmatrix}.$$

Diese Vektoren sind dann auch orthogonal zueinander, denn

$$\left\langle \begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} -b \\ a \end{pmatrix} \right\rangle^{\mathrm{st}} = a \cdot (-b) + b \cdot a = 0.$$

Definition 14.12 – orthogonales Komplement:

Sei $V = (V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum und $M \subseteq V$ eine Teilmenge von V. Ein Vektor $v \in V$ liegt senkrecht zu M, geschrieben $v \perp M$, falls für alle $w \in M$ gilt: $v \perp w$. Das orthogonale Komplement von M ist die Menge

$$M^{\perp} := \{ v \in V \mid v \perp M \}$$

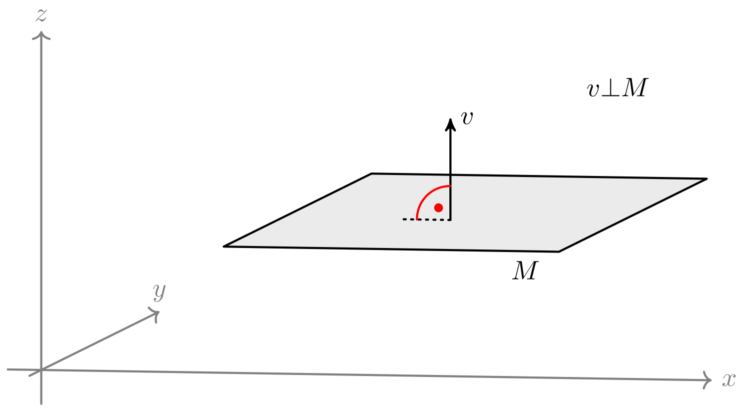


Abbildung 24: orthogonales Komplement

Satz 14.13:

Sei $V = (V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum und $M \subseteq V$. Dann gilt:

(a) $M^\perp$ ist ein Untervektorraum von V.
(b) $M^\perp \cap M \subseteq \{0\}$.

Beweis:

Zu (a): Wir überprüfen die Bedingungen aus der Definition eines Untervektorraums:

(i) $0 \perp M$ und damit ist $0 \in M^\perp$ und somit $M^\perp \neq \emptyset$.
(ii) Betrachte $x, y \in M^\perp$ und $\lambda \in \mathbb{R}$. Betrachte $z \in M$. Dann ist

$$\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle = 0 + 0 = 0$$

und

$$\langle \lambda x, z\rangle = \lambda \langle x, z\rangle = \lambda \cdot 0 = 0.$$

Damit sind $x + y \in M^\perp$ und $\lambda x \in M^\perp$.

Zu (b): Betrachte $x \in M^\perp \cap M$. Jedes Element von $M^\perp$ ist senkrecht zu allen Elementen von M. Daher ist $x \perp x$, sowie $\langle x, x\rangle = 0$ und somit $x = 0$ nach Satz 14.5 Teil (b).

⊠

⋆ ⋆ ⋆

Orthonormalsysteme. Die kanonische Basis $(e_1, \ldots, e_m)$ des euklidischen Vektorraums $\mathbb{R}^m$ mit dem Standard-Skalarprodukt hat die Eigenschaft, dass

$$\langle e_i, e_j\rangle = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases}$$

Verschiedene Vektoren der Basis sind orthogonal und jeder Basisvektor ist auf die Länge $\|e_i\| =$

$\sqrt{\langle e_i, e_i \rangle} = \sqrt{1} = 1$ normiert. Diese Eigenschaft ist sehr nützlich und lässt sich wie folgt verallgemeinern:

Definition 14.14 – Orthonormalsystem:

Sei $V = (V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum mit den Vektoren $v_1, \ldots, v_r \in V$. Das r-Tupel $(v_1, \ldots, v_r)$ ist **orthonormal** oder ein **Orthonormalsystem**, wenn für $i, j \leq r$ gilt:

$$\langle v_i, v_j \rangle = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases}$$

Orthonormale Vektoren sind immer linear unabhängig, wie der folgende Satz zeigt:

Satz 14.15:

Sei $V = (V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum und $(v_1, \ldots, v_r)$ ein Orthonormalsystem in V. Dann ist $(v_1, \ldots, v_r)$ linear unabhängig.

Beweis: Betrachte $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ mit $\lambda_1 v_1 + \ldots + \lambda_r v_r = 0$. Betrachte $i \leq r$. Die Anwendung des Skalarproduktes mit v_i auf beiden Seiten der Gleichung liefert:

$$\begin{aligned} 0 &= \langle 0, v_i \rangle &&= \langle \lambda_1 v_1 + \ldots + \lambda_r v_r, v_i \rangle \\ &= \lambda_1 \langle v_1, v_i \rangle + \ldots + \lambda_r \langle v_r, v_i \rangle &&= \lambda_i \langle v_i, v_i \rangle &&= \lambda_i \end{aligned}$$

Damit ist für alle $i \leq r$ der i-te Koeffizient $\lambda_i = 0$. Damit ist $(v_1, \ldots, v_r)$ linear unabhängig. ⊠

Besonders wichtig ist, dass sich Koordinatendarstellungen bezüglich orthonormaler Basen durch Anwendung des Skalarproduktes ausrechnen lassen.

Satz 14.16:

Sei $V = (V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum und $(v_1, \ldots, v_m)$ eine orthonormale Basis von V. Sei weiterhin $v \in V$ gegeben. Dann ist

$$(\langle v, v_1 \rangle, \ldots, \langle v, v_m \rangle)$$

die Koordinatendarstellung des Vektors v in der Basis $(v_1, \ldots, v_m)$, das heißt

$$v = \sum_{i=1}^{m} \langle v, v_i \rangle v_i.$$

Diese Summenformel in der Variablen v ist die **Entwicklungsformel** nach der Basis $(v_1, \ldots, v_m)$.

Beweis: Sei $(\lambda_1, \ldots, \lambda_m)$ die eindeutig bestimmte Koordinatendarstellung des Vektors v bezüglich der Basis $(v_1, \ldots, v_m)$, d.h. es gilt:

$$v = \sum_{i=1}^{m} \lambda_i v_i.$$

Betrachte ein beliebiges $j \leq m$. Multiplikation der Gleichung mit v_j ergibt dann:

$$\langle v, v_j \rangle = \left\langle \sum_{i=1}^{m} \lambda_i v_i, v_j \right\rangle = \sum_{i=1}^{m} \lambda_i \langle v_i, v_j \rangle = \lambda_j \langle v_j, v_j \rangle = \lambda_j$$

⊠

Diese letzte Eigenschaft können wir ausnutzen, um Orthonormalbasen zu erzeugen. Das so genannte Gram-Schmidtsche Orthonormalisierungsverfahren hilft dabei algorithmisch wie folgt:.

Satz 14.17 – Gram-Schmidtsches Orthonormalisierungsverfahren:

Sei $V = (V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum und $(v_1, \ldots, v_r)$ linear unabhängig in V. Für $k \leq r$ definiere rekursiv

$$\tilde{v}_k = \frac{v_k - \sum_{i=1}^{k-1} \langle v_k, \tilde{v}_i\rangle \tilde{v}_i}{\left\| v_k - \sum_{i=1}^{k-1} \langle v_k, \tilde{v}_i\rangle \tilde{v}_i \right\|}$$

Dann gilt für $s \leq r$, dass $(\tilde{v}_1, \ldots, \tilde{v}_s)$ wohldefiniert und orthonormal ist und dass die linearen Hüllen gleich sind:

$$\mathcal{L}(\tilde{v}_1, \ldots, \tilde{v}_s) = \mathcal{L}(v_1, \ldots, v_s)$$

Beweis: Wir zeigen die Behauptung durch Induktion über $s \leq r$.

Induktionsanfang

Für $s = 0$ (und für $s = 1$) gelten die Behauptungen trivialerweise.

Induktionsschritt

Angenommen, die Behauptungen gelten für $s < r$. Da $(v_1, \ldots, v_r)$ linear unabhängig ist, ist $v_{s+1} \notin \mathcal{L}(v_1, \ldots, v_s) = \mathcal{L}(\tilde{v}_1, \ldots, \tilde{v}_s)$. Damit ist offenbar $v_{s+1} - \sum_{i=1}^{s} \langle v_{s+1}, \tilde{v}_i\rangle \tilde{v}_i \neq 0$ und somit ist

$$\tilde{v}_{s+1} = \frac{v_{s+1} - \sum_{i=1}^{s} \langle v_{s+1}, \tilde{v}_i\rangle \tilde{v}_i}{\left\| v_{s+1} - \sum_{i=1}^{s} \langle v_{s+1}, \tilde{v}_i\rangle \tilde{v}_i \right\|}$$

wohldefiniert und von Norm 1. Zum Nachweis der Orthogonalität betrachte $j \leq s$:

$$\begin{aligned}
\left\langle \tilde{v}_j, v_{s+1} - \sum_{i=1}^{s} \langle v_{s+1}, \tilde{v}_i\rangle \tilde{v}_i \right\rangle &= \langle \tilde{v}_j, v_{s+1}\rangle - \sum_{i=1}^{s} \langle v_{s+1}, \tilde{v}_i\rangle \langle \tilde{v}_j, \tilde{v}_i\rangle \\
&= \langle \tilde{v}_j, v_{s+1}\rangle - \langle v_{s+1}, \tilde{v}_j\rangle \langle \tilde{v}_j, \tilde{v}_j\rangle \\
&= \langle \tilde{v}_j, v_{s+1}\rangle - \langle v_{s+1}, \tilde{v}_j\rangle \\
&= 0
\end{aligned}$$

Nach Induktionsvoraussetzung ist $\{\tilde{v}_1, \ldots, \tilde{v}_s\} \subseteq \mathcal{L}(v_1, \ldots, v_s)$ und es gilt nach der obigen Darstellung $\tilde{v}_{s+1} \in \mathcal{L}(\tilde{v}_1, \ldots, \tilde{v}_s, v_{s+1}) \subseteq \mathcal{L}(v_1, \ldots, v_s, v_{s+1})$, so dass wir schließlich erhalten: $\mathcal{L}(\tilde{v}_1, \ldots, \tilde{v}_s, \tilde{v}_{s+1}) \subseteq \mathcal{L}(v_1, \ldots, v_s, v_{s+1})$.

Da man die definierende Gleichung für $\tilde{v}_{s+1}$ auch nach v_{s+1} auflösen kann, gilt umgekehrt ebenfalls $\{v_1, \ldots, v_s\} \subseteq \mathcal{L}(\tilde{v}_1, \ldots, \tilde{v}_s)$, aber auch $v_{s+1} \in \mathcal{L}(\tilde{v}_1, \ldots, \tilde{v}_s, \tilde{v}_{s+1})$ und somit wie gewünscht $\mathcal{L}(v_1, \ldots, v_s, v_{s+1}) \subseteq \mathcal{L}(\tilde{v}_1, \ldots, \tilde{v}_s, \tilde{v}_{s+1})$.

Der Satz gilt damit nach dem Induktionsprinzip. ⊠

Die Rekursionsvorschrift innerhalb des Satzes 14.17 ist im täglichen (mathematischen) Leben erträglich aufwendig, d.h. wir haben:

$$\boxed{\tilde{v}_1} = \frac{v_1}{\|v_1\|}$$

$$\boxed{\tilde v_2} = \frac{v_2 - \langle v_2, \tilde v_1\rangle \cdot \tilde v_1}{\left\|v_2 - \langle v_2, \tilde v_1\rangle \cdot \tilde v_1\right\|}$$

$$\boxed{\tilde v_3} = \frac{v_3 - \langle v_3, \tilde v_1\rangle \cdot \tilde v_1 - \langle v_3, \tilde v_2\rangle \cdot \tilde v_2}{\left\|v_3 - \langle v_3, \tilde v_1\rangle \cdot \tilde v_1 - \langle v_3, \tilde v_2\rangle \cdot \tilde v_2\right\|}$$

usw.

Beispiel 14.18: Es seien die beiden Vektoren $v_1 = \begin{pmatrix}3\\1\end{pmatrix}$ und $v_2 = \begin{pmatrix}2\\2\end{pmatrix}$ gegeben. Wir suchen die Orthonormalbasis nach dem Gram-Schmidtschen Orthonormalisierungsverfahren:

$$\begin{aligned}
\boxed{\tilde v_1} &= \frac{v_1}{\|v_1\|} &&= \frac{\begin{pmatrix}3\\1\end{pmatrix}}{\left\|\begin{pmatrix}3\\1\end{pmatrix}\right\|}\\
&= \frac{1}{\sqrt{3^2+1^2}} \cdot \begin{pmatrix}3\\1\end{pmatrix} &&= \frac{1}{\sqrt{10}} \cdot \begin{pmatrix}3\\1\end{pmatrix}\\
v_2' &= v_2 - \langle v_2, \tilde v_1\rangle \cdot \tilde v_1 &&= \begin{pmatrix}2\\2\end{pmatrix} - \left\langle \begin{pmatrix}2\\2\end{pmatrix}, \frac{1}{\sqrt{10}} \cdot \begin{pmatrix}3\\1\end{pmatrix}\right\rangle \cdot \frac{1}{\sqrt{10}} \cdot \begin{pmatrix}3\\1\end{pmatrix}\\
&= \begin{pmatrix}2\\2\end{pmatrix} - \frac{1}{(\sqrt{10})^2} \cdot 8 \cdot \begin{pmatrix}3\\1\end{pmatrix} &&= \begin{pmatrix}2 - \frac{4\cdot 3}{5}\\ 2 - \frac{4\cdot 1}{5}\end{pmatrix}\\
&= \frac{1}{5} \cdot \begin{pmatrix}-2\\6\end{pmatrix}\\
\left\|v_2'\right\| &= \sqrt{\left\langle \frac{1}{5} \cdot \begin{pmatrix}-2\\6\end{pmatrix}, \frac{1}{5} \cdot \begin{pmatrix}-2\\6\end{pmatrix}\right\rangle} &&= \sqrt{\frac{1}{25} \cdot (4+36)}\\
&= \frac{\sqrt{4 \cdot 10}}{5} &&= \frac{2}{5}\sqrt{10}\\
\boxed{\tilde v_2} &= \frac{v_2'}{\left\|v_2'\right\|} &&= \frac{\frac{1}{5} \cdot \begin{pmatrix}-2\\6\end{pmatrix}}{\frac{2}{5}\sqrt{10}}\\
&= \frac{5}{2 \cdot \sqrt{10} \cdot 5} \cdot \begin{pmatrix}-2\\6\end{pmatrix} &&= \frac{1}{\sqrt{10}} \cdot \begin{pmatrix}-1\\3\end{pmatrix}
\end{aligned}$$

Damit haben wir unsere beiden Vektoren $\tilde v_1 = \frac{1}{\sqrt{10}} \cdot \begin{pmatrix}3\\1\end{pmatrix}$ und $\tilde v_2 = \frac{1}{\sqrt{10}} \cdot \begin{pmatrix}-1\\3\end{pmatrix}$ im Orthonormalsystem nach Gram-Schmidt gefunden.

Diese sind offenbar orthogonal und jeweils normiert, denn es gilt:

$$\begin{aligned}
\langle \tilde v_1, \tilde v_1\rangle &= \left\langle \frac{1}{\sqrt{10}} \cdot \begin{pmatrix}3\\1\end{pmatrix}, \frac{1}{\sqrt{10}} \cdot \begin{pmatrix}3\\1\end{pmatrix}\right\rangle &&= \frac{1}{(\sqrt{10})^2} \cdot (3^2 + 1^2) &&= 1\\
\langle \tilde v_2, \tilde v_2\rangle &= \left\langle \frac{1}{\sqrt{10}} \cdot \begin{pmatrix}-1\\3\end{pmatrix}, \frac{1}{\sqrt{10}} \cdot \begin{pmatrix}-1\\3\end{pmatrix}\right\rangle &&= \frac{1}{(\sqrt{10})^2} \cdot ((-1)^2 + 3^2) &&= 1
\end{aligned}$$

$$\langle \tilde{v}_1, \tilde{v}_2 \rangle \quad = \quad \left\langle \frac{1}{\sqrt{10}} \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{10}} \cdot \begin{pmatrix} -1 \\ 3 \end{pmatrix} \right\rangle \quad = \quad \frac{1}{(\sqrt{10})^2} \cdot (3 \cdot (-1) + 1 \cdot 3) \quad = \quad 0$$

⋆ ⋆ ⋆

Orthonormieren über nicht-triviale Skalarprodukte. Sie können ganz leicht nicht-triviale Skalarprodukte über Matrizen einführen, wie das das folgende Beispiel macht:

Beispiel 14.19: Betrachten Sie die Abbildung $\langle \cdot, \cdot \rangle : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$, gegebenen durch

$$\langle v_1, v_2 \rangle \; := \; v_1^t \cdot \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \cdot v_2.$$

Dann definiert diese Abbildung ein Skalarprodukt auf $\mathbb{R}^3$: Die Bilinearität folgt sehr leicht. Die Symmetrie folgt sofort aufgrund der Symmetrie der eingehenden Matrix. Es bleibt die positive Definitheit zu überpüfen. Hierbei gilt:

$$\left\langle \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \begin{pmatrix} x \\ y \\ z \end{pmatrix} \right\rangle \; = \; x^2 + 2xy + 2y^2 + 4z^2 \; = \; (x+y)^2 + y^2 + 4z^2 \; \geq \; 0$$

Dabei ist dieser Funktionswert nur gleich null, wenn alle eingehenden Komponenten gleichzeitig gleich null sind.
Allgemein lässt sich Folgendes zeigen:

Satz 14.20:

Für eine positiv definite und symmetrische Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ definiert $\langle x, y \rangle_A := x^t A y$ ein Skalarprodukt auf $\mathbb{R}^n$. In diesem Fall gilt in Bezug auf das Standardskalarproduk: $\langle x, y \rangle_A = \langle x, Ay \rangle^{\mathrm{st}}$. Umgekehrt lässt sich jedes beliebige Skalarprodukt auf $\mathbb{R}^n$ auf diese Art durch eine positiv definite symmetrische Matrix darstellen.

Natürlich können Sie auch zu allgemeinen Skalarprodukten dieser Darstellungsart das Gram-Schmidtsche Verfahren anwenden, wie dies Ihnen das folgende Beispiel zeigen wird:

Beispiel 14.21: Betrachten Sie erneut das im vorhergehenden Beispiel betrachtete Skalarprodukt. Sie können nun das Gram-Schmidtsche Orthonormalisierungsverfahren auf Vektoren bezüglich dieses allgemeinen Skalarprodukts anwenden. Wir orthonormalisieren daher die kanonischen Einheitsvektoren e_1, e_2 und e_3 wie folgt:

Setze nun $v_1 := e_1$. Wegen

$$\|v_1\| = \|e_1\| = \sqrt{\langle e_1, e_1 \rangle} = \sqrt{e_1^t \cdot \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \cdot e_1} = \sqrt{1} = 1$$

ist $\boxed{\tilde{v}_1} = e_1$. Sehen Sie, dass wir hier die Standardbezeichnungen für Skalarprodukt und Norm benutzen? – Wir könnten diese nun gesondert bezeichnen, verzichten aber darauf. Aber aufpassen und nicht verwechseln mit dem jeweiligen Standardskalarprodukt!

Für den zweiten Vektor gilt wegen

$$\langle e_2, e_1\rangle = e_2^t \cdot \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \cdot e_1 = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 1$$

schließlich:

$$\tilde{v}_2 = \frac{e_2 - \langle e_2, \tilde{v}_1\rangle \cdot \tilde{v}_1}{\left\| e_2 - \langle e_2, \tilde{v}_1\rangle \cdot \tilde{v}_1 \right\|} = \frac{e_2 - \langle e_2, e_1\rangle \cdot e_1}{\left\| e_2 - \langle e_2, e_1\rangle \cdot e_1 \right\|} = \frac{e_2 - 1 \cdot e_1}{\| e_2 - 1 \cdot e_1 \|}$$

Für die Normierung folgt:

$$\left\| \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right\| = \sqrt{\begin{pmatrix} -1 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}} = \sqrt{\begin{pmatrix} -1 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}} = \sqrt{1} = 1$$

Also ist $\boxed{\tilde{v}_2} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$.

Die Berechnungen für den ausstehenden dritten Vektor folgen analog:

$$\begin{aligned} \tilde{v}_3 &= \frac{v_3 - \langle v_3, \tilde{v}_1\rangle \cdot \tilde{v}_1 - \langle v_3, \tilde{v}_2\rangle \cdot \tilde{v}_2}{\left\| v_3 - \langle v_3, \tilde{v}_1\rangle \cdot \tilde{v}_1 - \langle v_3, \tilde{v}_2\rangle \cdot \tilde{v}_2 \right\|} \\ &= \frac{e_3 - \langle e_3, e_1\rangle \cdot e_1 - \left\langle e_3, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right\rangle \cdot \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}}{\left\| e_3 - \langle e_3, e_1\rangle \cdot e_1 - \left\langle e_3, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right\rangle \cdot \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right\|} \end{aligned}$$

Es gilt

$$\langle e_3, e_1\rangle = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 0$$

und ähnlich auch

$$\left\langle e_3, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right\rangle = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = 0$$

Es bleibt, den Vektor e_3 zu normieren; es gilt:

$$\langle e_3, e_3\rangle = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 4$$

Damit hat der gesuchte dritte Vektor die Gestalt:

$$\boxed{\tilde{v}_3} = \frac{e_3}{\|e_3\|} = \frac{1}{2} \cdot e_3 = \begin{pmatrix} 0 \\ 0 \\ \frac{1}{2} \end{pmatrix}.$$

⋆ ⋆ ⋆

Orthogonale Zerlegungen und Projektionen. Wir beschäftigen uns jetzt weiter mit dem bisher nur angerissenen Begriff des orthogonalen Komplements einer Menge von Vektoren und zeigen hierfür den

Satz 14.22:

Sei $V = (V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum und $M \subseteq V$ ein endlich-dimensionaler Unterraum von V. Sei $w \in V$. Dann existieren eindeutig bestimmte Vektoren $u \in M$ und $v \in M^\perp$, so dass $w = u + v$. Diese Darstellung von w ist die Zerlegung des Vektors w nach M und seinem orthogonalen Komplement.

Beweis: Wähle eine Basis $(v_1, \dots, v_r)$ von M. Nach dem Orthonormalisierungssatz 14.17 können wir diese Basis zu einer orthonormalen Basis von M modifizieren. Daher können wir annehmen, dass $(v_1, \dots, v_r)$ orthonormal ist. Setze

$$u := \sum_{i=1}^{r} \langle w, v_i\rangle\, v_i$$

und $v := w - u$. Dann ist $u \in M$.

Es bleibt zu zeigen, dass $v \in M^\perp$ gilt. Betrachte dazu Basiselemente v_j von M:

$$\begin{aligned}
\langle v, v_j\rangle &= \langle w - u, v_j\rangle &&= \left\langle w - \sum_{i=1}^{r} \langle w, v_i\rangle\, v_i, v_j \right\rangle \\
&= \langle w, v_j\rangle - \sum_{i=1}^{r} \langle w, v_i\rangle \langle v_i, v_j\rangle &&= \langle w, v_j\rangle - \langle w, v_j\rangle \langle v_j, v_j\rangle \\
&= \langle w, v_j\rangle - \langle w, v_j\rangle &&= 0
\end{aligned}$$

Somit gilt für ein beliebiges Element $\sum_{j=1}^{r} \lambda_j v_j \in M$ also

$$\left\langle v, \sum_{j=1}^{r} \lambda_j v_j \right\rangle = \sum_{j=1}^{r} \lambda_j \langle v, v_j\rangle = 0$$

Also ist wie gewünscht $v \in M^\perp$.

Um die Eindeutigkeit zu zeigen, betrachte zwei Zerlegungen $w = u+v$ und $w = u'+v'$ mit $u, u' \in M$ und $v, v' \in M^\perp$. Dann ist

$$u + v = u' + v' \text{ und } u - u' = v' - v.$$

Es gilt $u - u' \in M$ und $v' - v \in M^\perp$. Nach Satz 14.13 (b) ist $M \cap M^\perp = \{0\}$ und somit gilt

$$u - u' = v' - v = 0.$$

Damit ist $u = u'$ und $v' = v$. Die Aussage des Satzes ist damit bewiesen. ⊠

Nach diesem Satz ist also folgende Zerlegung gegeben:

$$V = M + M^\perp = \left\{ u + v \mid u \in M \text{ und } v \in M^\perp \right\}.$$

Da weiterhin $M \cap M^\perp = \{0\}$ gilt, ist die Zerlegung eines Vektors nach M und $M^\perp$ eindeutig bestimmt. In diesem Fall sagt man, dass V die direkte Summe von M und $M^\perp$ ist.

Mithilfe dieser Begriffe können wir nun leicht orthogonal projizieren:

Definition 14.23 – orthogonale Projektion:

Sei $V = (V, \langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum und M ein endlich-dimensionaler Unterraum von V. Definiere die orthogonale Projektion

$$P_M : V \to M$$

auf M durch: Für $w \in V$ ist $P_M(w) := u$ das nach Satz 14.22 eindeutig bestimmte $u \in M$, so dass $w - u \in M^{\perp}$.

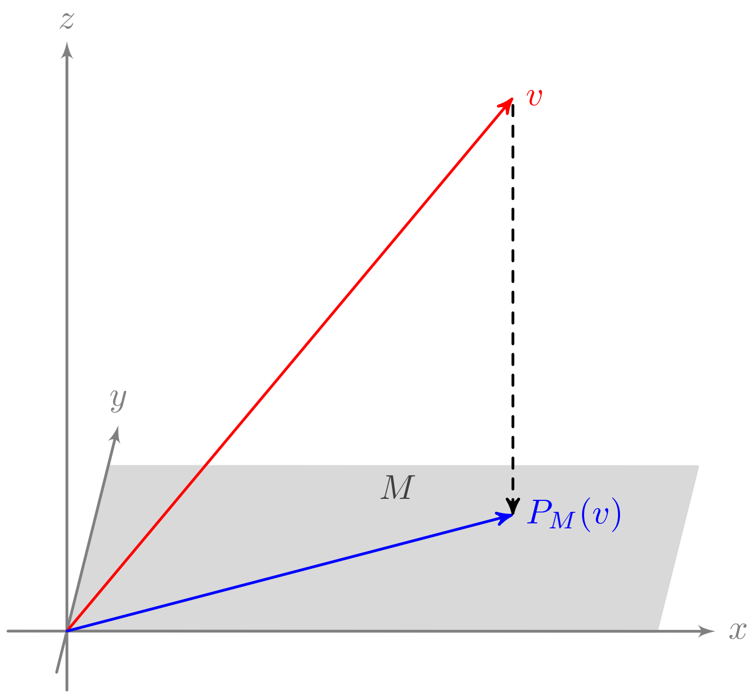


Abbildung 25: orthogonale Projektion

Diese orthogonale Projektion hat die gewünschten Eigenschaften, wie wir jetzt sehen werden:

Satz 14.24:

Sei $P_M : V \to M$ die in Definition 14.23 definierte orthogonale Projektion von V auf den Unterraum $M \subseteq V$. Dann gilt:

(a) $P_M : V \to M$ ist eine lineare Abbildung.
(b) $P_M \restriction M = \mathrm{Id}_M$ und $P_M : V \to M$ ist surjektiv.
(c) $\mathrm{Kern}(P_M) = M^{\perp}$.

Beweis:

Zu (a): Betrachte $w, w' \in V$ und $\lambda \in \mathbb{R}$. Dann sind

$$w = P_M(w) + \left(w - P_M(w)\right) \text{ und } w' = P_M(w') + \left(w' - P_M(w')\right)$$

die Zerlegungen von w und w' nach M und $M^{\perp}$. Dann gilt

$$\begin{aligned} w + w' &= \Big(P_M(w) + (w - P_M(w))\Big) + \Big(P_M(w') + \big(w' - P_M(w')\big)\Big) \\ &= \Big(P_M(w) + P_M(w')\Big) + \Big((w - P_M(w)) + \big(w' - P_M(w')\big)\Big) \end{aligned}$$

ist die Zerlegung von $w + w'$ nach M und $M^{\perp}$. Damit ist

$$P_M(w + w') = P_M(w) + P_M(w').$$

Weiter ist $\lambda w = \lambda P_M(w) + \lambda\big(w - P_M(w)\big)$ die Zerlegung von λw nach M und $M^{\perp}$, da $\lambda P_M(w) \in M$ und $\lambda\big(w - P_M(w)\big) \in M^{\perp}$. Damit ist

$$P_M(\lambda w) = \lambda \cdot \big(P_M(w)\big).$$

und somit die Abbildung linear.

Zu (b): Betrachte $w \in M$. Dann ist $w = w + 0$ die Zerlegung von w nach M und $M^{\perp}$. Also ist $P_M(w) = w = \mathrm{Id}_M(w)$. Damit ist $P_M \restriction M = \mathrm{Id}_M$.

Zu (c): Betrachte $w \in \mathrm{Kern}(P_M)$. Dann ist $P_M(w) = 0$ und

$$w = P_M(w) + \big(w - P_M(w)\big) = 0 + w$$

ist die Zerlegung von w nach M und $M^{\perp}$. Dann ist $w \in M^{\perp}$. Betrachte umgekehrt $w \in M^{\perp}$. Dann ist

$$w = 0 + w$$

die Zerlegung von w nach M und $M^{\perp}$. Somit ist wie gewünscht $P_M(w) = 0$ und $w \in \mathrm{Kern}(P_M)$.

⊠

⋆ ⋆ ⋆

Orthogonale Abbildungen. Wir geben im Folgenden noch abschließend einen kleinen Ausblick. Wir definieren orthogonale Abbildungen und werden sehen, dass wir bereits wohlbekannte Beispiele dafür kennen.

Definition 14.25 – isometrische/orthogonale Abbildung:

Seien $V = (V, \langle\cdot,\cdot\rangle_V)$ und $W = (W, \langle\cdot,\cdot\rangle_W)$ euklidische Vektorräume und sei $f : V \to W$ eine lineare Abbildung. Dann ist f isometrisch oder orthogonal, wenn für alle $u, v \in V$ gilt:

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

Beispiel 14.26: Seien $V = W = \mathbb{R}^2$ mit dem Standard-Skalarprodukt $\langle\cdot,\cdot\rangle$. Sei f die Drehung von $\mathbb{R}^2$ um den Winkel α. Betrachte Vektoren $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{R}^2$. Dann ist

$$f\left(\begin{pmatrix}u_1\\u_2\end{pmatrix}\right)=\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}\begin{pmatrix}u_1\\u_2\end{pmatrix}=\begin{pmatrix}u_1\cos\alpha-u_2\sin\alpha\\ u_1\sin\alpha+u_2\cos\alpha\end{pmatrix},$$
$$f\left(\begin{pmatrix}v_1\\v_2\end{pmatrix}\right)=\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}\begin{pmatrix}v_1\\v_2\end{pmatrix}=\begin{pmatrix}v_1\cos\alpha-v_2\sin\alpha\\ v_1\sin\alpha+v_2\cos\alpha\end{pmatrix}.$$

Für das Skalarprodukt zwischen $f(u)$ und $f(v)$ gilt:

$$\begin{aligned}\left\langle f\left(\begin{pmatrix}u_1\\u_2\end{pmatrix}\right), f\left(\begin{pmatrix}v_1\\v_2\end{pmatrix}\right)\right\rangle &= \left\langle\begin{pmatrix}u_1\cos\alpha-u_2\sin\alpha\\ u_1\sin\alpha+u_2\cos\alpha\end{pmatrix},\begin{pmatrix}v_1\cos\alpha-v_2\sin\alpha\\ v_1\sin\alpha+v_2\cos\alpha\end{pmatrix}\right\rangle\\
&= u_1v_1\cos\alpha\cos\alpha-u_1v_2\cos\alpha\sin\alpha-u_2v_1\sin\alpha\cos\alpha+u_2v_2\sin\alpha\sin\alpha+\\
&\quad+u_1v_1\sin\alpha\sin\alpha+u_1v_2\sin\alpha\cos\alpha+u_2v_1\cos\alpha\sin\alpha+u_2v_2\cos\alpha\cos\alpha\\
&= u_1v_1\cos^2\alpha+u_1v_1\sin^2\alpha+u_2v_2\sin^2\alpha+u_2v_2\cos^2\alpha\\
&= u_1v_1+u_2v_2=\left\langle\begin{pmatrix}u_1\\u_2\end{pmatrix},\begin{pmatrix}v_1\\v_2\end{pmatrix}\right\rangle.\end{aligned}$$

Damit ist die Drehung f isometrisch.

Orthogonale Abbildungen sind bereits sehr speziell, wie die folgenden beiden Sätze zeigen werden:

Satz 14.27:

Seien $V=(V,\langle\cdot,\cdot\rangle_V)$ und $W=(W,\langle\cdot,\cdot\rangle_W)$ euklidische Vektorräume und sei $f: V\to W$ orthogonal. Dann ist f injektiv.

Beweis: Betrachte $v,v'\in V$. mit $f(v)=f(v')$. Dann ist $f(v-v')=f(v)-f(v')=0$ und $\langle v-v',v-v'\rangle_V=\langle f(v-v'),f(v-v')\rangle_W=\langle 0,0\rangle_W=0$. Nach Satz 14.15 ist dann $v-v'=0$ und somit $v=v'$. ⊠

Satz 14.28:

Sei $V=(V,\langle\cdot,\cdot\rangle)$ ein euklidischer Vektorraum endlicher Dimension und $f:V\to V$ ein orthogonaler Endomorphismus. Dann ist f ein Automorphismus.

Beweis: Nach Satz 14.27 ist f injektiv. Nach der Dimensionsformel 6.19 ist ein injektiver Homomorphismus surjektiv. ⊠

Orthogonale Abbildungen haben in sich schöne Eigenschaften, so dass man die Menge solcher Abbildungen zusammenfasst.

Definition 14.29 – orthogonale Gruppe:

Sei $V=(V,\langle\cdot,\cdot\rangle)$ ein endlich-dimensionaler Vektorraum. Dann sei

$$O(V)=\left\{f\;\middle|\; f:V\to V \text{ ist orthogonal}\right\}$$

die orthogonale Gruppe von V.

Wenn $\mathbb{R}^m=\left(\mathbb{R}^m,\langle\cdot,\cdot\rangle\right)$ der m-dimensionale Standard-Raum ist, so schreibt man auch $O(m)$ statt $O(\mathbb{R}^m)$.

Über dieser Menge können wir nun die Gruppengesetze nachweisen:

Satz 14.30:

Sei $V = (V, \langle\cdot,\cdot\rangle)$ ein endlich-dimensionaler euklidischer Vektorraum. Dann ist $O(V)$ mit der Komposition von Abbildungen eine **Gruppe**, d.h. die Struktur $(O(V), \circ)$ erfüllt die **Gruppengesetze**:

(a) Für $f, g, h \in O(V)$ gilt:
$h \circ (g \circ f) = (h \circ g) \circ f$ (Assoziativgesetz)

(b) Für $f \in O(V)$ gilt:
$f \circ \mathrm{id}_V = f$ und $\mathrm{id}_V \circ f = f$ (Neutrales Element)

(c) Für alle $f \in O(V)$ gilt:
$f^{-1} \circ f = \mathrm{id}_V$ und $f \circ f^{-1} = \mathrm{id}_V$ (Existenz von Inversen)

Ein Element f von $O(m)$ wird in der kanonischen Basis des $\mathbb{R}^m$ als eine $(m \times m)$-Matrix A dargestellt. Die kanonische Basis ist selbst ein orthonormales System, das durch die Spalten von A dargestellt wird, so dass wir den Übergang zu Matrizen vollziehen können:

Definition 14.31 – orthogonale Matrix:

Sei $m \in \mathbb{N} \setminus \{0\}$. Eine Matrix $A \in \mathrm{Mat}(m \times m, \mathbb{R})$ ist **orthogonal**, wenn das System der Spaltenvektoren von A orthonormal ist.

Lemma 14.32:

Wenn A orthogonal ist, dann ist $A^t = A^{-1}$, also $A^t \cdot A = E_n$

Beweis: Wenn in A die Spalten ein Orthonormalsystem bilden, dann gilt dies für A^t ebenso. Insbesondere ist dann das gewünschte Matrizenprodukt genau auf der Diagonalen wegen der Normiertheitseigenschaft gleich 1 und außerhalb der Diagonalen wegen der Orthogonalitätseigenschaft der eingehenden Vektoren gleich 0. ⊠

Beispiel 14.33: In der Dimension $m = 2$ sind Drehmatrizen und Spiegelmatrizen

$$\begin{pmatrix} \cos\alpha & \mp\sin\alpha \\ \sin\alpha & \pm\cos\alpha \end{pmatrix}$$

orthogonal, denn es gilt:

$$\begin{aligned} (\cos\alpha)^2 + (\sin\alpha)^2 &= 1, \\ (\mp\sin\alpha)^2 + (\pm\cos\alpha)^2 &= 1, \\ \cos\alpha \cdot (\mp\sin\alpha) + \sin\alpha \cdot (\pm\cos\alpha) &= 0. \end{aligned}$$

Umgekehrt gilt der folgende Satz:

Satz 14.34:

Jedes $A \in O(2)$ ist für geeignetes α von der Form

$$\begin{pmatrix} \cos\alpha & \mp\sin\alpha \\ \sin\alpha & \pm\cos\alpha \end{pmatrix}$$

Damit schließen wir diesen kleinen Exkurs.

⋆ ⋆ ⋆

Aufgabe 14.35: Sei V ein n-dimensionaler euklidischer Vektorraum, $f : V \to V$ ein Endomorpismus, wobei aus $\langle x, y\rangle = 0$ die Gleichung $\big\langle f(x), f(y)\big\rangle = 0$ folgt. Zeigen Sie, dass es ein $\lambda \in \mathbb{R}$ gibt, so dass $\lambda \cdot f$ eine Isometrie ist.

15. Über selbstadjungierte Endomorphismen und reell symmetrische Matrizen

In diesem Kapitel geben wir einen Ausblick, welche Eigenschaften reell symmetrische Matrizen in Bezug auf die Diagonalisierbarkeit haben. Sei für diese Betrachtungen $V = (V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum.

Selbstadjungierte Endomorphismen.

Definition 15.1 – selbstadjungiert:

Ein Endomorphismus $f\colon V \to V$ heißt selbstadjungiert (bezüglich des gegebenen Skalarprodukts $\langle \cdot, \cdot \rangle$), falls gilt:

$$\big\langle f(v), w \big\rangle = \big\langle v, f(w) \big\rangle \qquad \text{für beliebige } v, w \in V.$$

Selbstadjungierte Endomorphismen haben betreffs ihrer Darstellbarkeit bezüglich Orthonomalbasen folgende nützliche Symmetrie-Eigenschaft:

Bemerkung 15.2:

Wenn $B = (b_1, \ldots, b_n)$ eine Orthonormalbasis von V ist, so gilt

$$f \text{ selbstadjungiert} \quad \Leftrightarrow \quad \mathrm{DM}_B(f)^t = \mathrm{DM}_B(f)$$

Beweis: Es gilt:

$$f \text{ selbstadjungiert} \quad \Leftrightarrow \quad \big\langle f(b_j), b_k \big\rangle = \big\langle b_j, f(b_k) \big\rangle \qquad \forall j, k = 1, \ldots, n$$

Da B orthonormal ist, gilt für $(a_{ij})_{1 \leq i,j \leq n} = \mathrm{DM}_B(f)$

$$\begin{array}{rcccl} \big\langle f(b_j), b_k \big\rangle & = & a_{1j} \langle b_1, b_k \rangle + \ldots + a_{nj} \langle b_n, b_k \rangle & = & a_{kj} \\ & & \vdots & & \\ \big\langle b_j, f(b_k) \big\rangle & = & a_{1k} \big\langle b_j, b_1 \big\rangle + \ldots + a_{nk} \big\langle b_j, b_n \big\rangle & = & a_{jk} \end{array}$$

⊠

Beispiel 15.3: Betrachten Sie die Spiegelung an der y-Achse als Endomorphismus $f : \mathbb{R}^2 \to \mathbb{R}^2$ im euklischen Vektorraum $\mathbb{R}^2$ mit dem Standardskalarprodukt. Die Darstellungsmatrix

$$\mathrm{DM}(f) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

bezüglich der kanonischen (orthonormalen) Basis ist offenbar symmetrisch. Außerdem ist eine solche Spiegelung auch selbstadjungiert, wie Ihnen die nachfolgende Abbildung für zwei Ausgangsvektoren $v_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ und $v_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix}$ suggeriert und die folgende Gleichung zeigt:

$$\langle\ f(v_1), v_2\ \rangle = \left\langle f\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}\right), \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} -x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \right\rangle = -x_1 x_2 + y_1 y_2$$

$$= \left\langle \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \begin{pmatrix} -x_2 \\ y_2 \end{pmatrix} \right\rangle = \left\langle \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, f\left(\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \right) \right\rangle = \big\langle v_1, f(v_2) \big\rangle$$

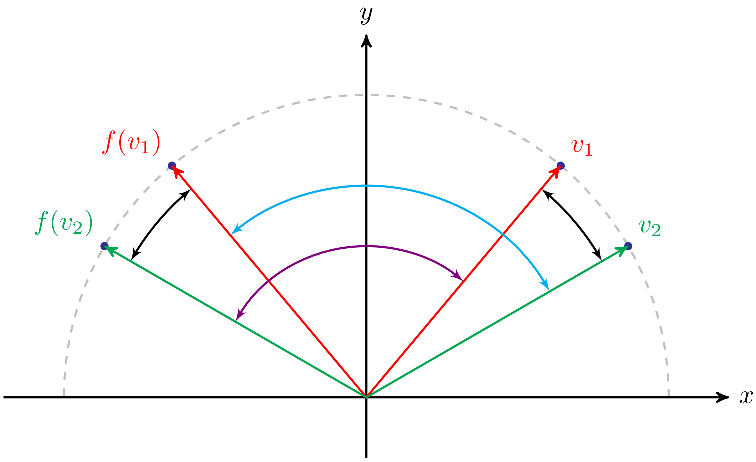


Abbildung 26

Übrigens, eine Spiegelung (etwa wie gerade im Beispiel) ändert natürlich auch nichts am Öffnungswinkel – mathematisch lässt sich dies in folgender Gleichung ausdrücken:

$$\langle v_1, v_2 \rangle = \big\langle f(v_1), f(v_2) \big\rangle$$

Eine Abbildung mit einer solchen Eigenschaft nennt man Isometrie. Sie ist längenerhaltend. Ein weitere selbstadjungierte Isometrie finden Sie im folgenden Beispiel:

Beispiel 15.4: Betrachten Sie den reellen Vektorraum der auf dem abgeschlossenen Intervall $[-1, 1]$ stetigen Funktionen von $\mathbb{R}$ nach $\mathbb{R}$. Dann definiert

$$\langle f, g \rangle := \int_{-1}^{1} f(x)g(x)dx$$

ein Skalarprodukt. Die Abbildung $\Phi(f)(x) := f(-x)$ stellt überdies eine selbstadjungierte Isometrie dar. Mittels Substitution rechnen Sie leicht nach, dass gilt:

$$\begin{aligned} \big\langle \Phi(f), g \big\rangle &= \int_{-1}^{1} \Phi(f)(x)g(x)dx = \int_{-1}^{1} f(-x)g(x)dx = -\int_{1}^{-1} f(z)g(-z)dz \\ &= \int_{-1}^{1} f(z)g(-z)dz = \int_{-1}^{1} f(z)\Phi(g)(z)dz \\ &= \big\langle f, \Phi(g) \big\rangle \end{aligned}$$

$$\begin{aligned} \big\langle \Phi(f), \Phi(g) \big\rangle &= \int_{-1}^{1} \Phi(f)(x)\Phi(g)(x)dx = \int_{-1}^{1} f(-x)g(-x)dx = -\int_{1}^{-1} f(z)g(z)dz \\ &= \int_{-1}^{1} f(z)g(z)dz = \langle f, g \rangle \end{aligned}$$

⋆ ⋆ ⋆

Hauptachsentransformation mittels des Spektralsatzes. Der folgende Satz garantiert uns letztendlich die gewünschte Diagonalisierungseigenschaft für die betrachtete Klasse von Endomorphismen.

Satz 15.5 – Spektralsatz:

Sei $f : V \to V$ ein selbstadjungierter Endomorphismus eines n-dimensionalen euklidischen Vektorraumes V. Dann besitzt V eine Orthonormalbasis, die aus Eigenvektoren von f besteht.

Schließlich lässt sich leicht folgende Folgerung für symmetrische Matrizen ziehen:

Korollar 15.0:

Sei $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ symmetrisch (d.h. $A^t = A$). Dann gibt es eine orthogonale Matrix T (d.h. $T^t \cdot T = E_n$), so dass

$$T^t A T = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}$$

Dabei sind die Skalare $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von A (mit den entsprechenden Vielfachheiten).

Damit haben reell symmetrische Matrizen wunderbare Eigenschaften, die sich wie folgt zusammenfassen lassen:

Satz 15.7:

Es gilt Folgendes:

(a) Jede reelle symmetrische Matrix hat nur reelle Eigenwerte.
(b) Für jede reell symmetrische Matrix stehen die Eigenvektoren zu verschiedenen Eigenwerten senkrecht aufeinander.
(c) Für jede reell symmetrische Matrix existiert stets eine Basis bestehend aus paarweise orthogonalen Eigenvektoren.
(d) Jede reell symmetrische Matrix ist diagonalisierbar, sogar mittels einer orthogonalen Transformationsmatrix.
(e) Der Rang einer reell symmetrischen Matrix ist die Anzahl der von Null verschiedenen Eigenwerten.

16. Definitheit von Matrizen

Dieses Kapitel soll Ihnen ein Hilfsmittel näher bringen, welches Sie an verschiedenen Stellen der Mathematik benötigen werden. Ich werde Ihnen am Ende des kleinen Ausflugs eine analytische Anwendung exemplarisch zeigen.

Definition und Charakterisierungen der Definitheit.

Definition 16.1 – Definitheit von Matrizen:

Sei $A \in \operatorname{Mat}(n \times n, \mathbb{R})$ gegeben. Dann nennt man die Matrix $A \dots$

$$\begin{aligned} \text{positiv definit,} \quad & \text{falls } \langle x, Ax \rangle > 0, \\ \text{positiv semidefinit,} \quad & \text{falls } \langle x, Ax \rangle \geq 0, \\ \text{negativ definit,} \quad & \text{falls } \langle x, Ax \rangle < 0, \\ \text{negativ semidefinit,} \quad & \text{falls } \langle x, Ax \rangle \leq 0, \end{aligned}$$

jeweils für alle Vektoren $x \in \mathbb{R}^n$ mit $x \neq 0$.

Satz 16.2:

Sei $A \in \operatorname{Mat}(n \times n, \mathbb{R})$ symmetrisch mit den Eigenwerten $\lambda_1, \dots, \lambda_n$.

$$\begin{aligned} A \text{ ist positiv definit,} &\Leftrightarrow \lambda_i > 0 \text{ für alle } i = 1, \dots, n \\ A \text{ ist positiv semidefinit,} &\Leftrightarrow \lambda_i \geq 0 \text{ für alle } i = 1, \dots, n \\ A \text{ ist negativ definit,} &\Leftrightarrow \lambda_i < 0 \text{ für alle } i = 1, \dots, n \\ A \text{ ist negativ semidefinit,} &\Leftrightarrow \lambda_i \leq 0 \text{ für alle } i = 1, \dots, n \end{aligned}$$

Beweis: Mit dem Spektralsatz 15.5 erhält man die Existenz von Matrizen T, D

$$T^t A T = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} =: D$$

für eine positiv definite Matrix A bedeutet dies

$$\begin{aligned} & 0 < \langle Tx, ATx \rangle = \left\langle T^t \cdot T \cdot x, T^t \cdot A \cdot Tx \right\rangle = \left\langle x, T^t \cdot A \cdot Tx \right\rangle = \langle x, Dx \rangle \\ \Longleftrightarrow \quad & 0 < \lambda_i \end{aligned}$$

für $x = e_i$ dem i-ten Einheitsvektor.

Hierbei gilt das erste Gleichheitszeichen wegen der Orthogonalität von T und damit auch von T^t. Beachten Sie, dass orthogonale Matrizen als Abbildungen betrachtet natürlich auch orthogonal sind (und damit isometrisch).

Der Rest lässt sich analog zeigen. ⊠

Folgender Begriff lohnt sich in diesem Zusammenhang zu betrachten, wie Sie gleich sehen werden:

Definition 16.3 – Hauptminoren:

Sei $A \in \mathrm{Mat}(n \times n, \mathbb{R})$, $A = (a_{ij})_{1 \leq i,j \leq n}$ gegeben. Dann nennt man

$$\det(a_{11}) = a_{11}, \quad \det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{32} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \quad \dots \quad , \det(A)$$

die Hauptminoren der Matrix A.

Dieser Begriff erweist sich als nützlich, um ein weiteres praktisches Kriterium zur Bestimmung der Definitheit von symmetrischen Matrizen zu bekommen.

Satz 16.4:

Eine symmetrische Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ ist genau dann positiv definit, wenn alle Hauptminoren von A positiv sind. Dementsprechend ist eine symmetrische Matrix $A \in \mathrm{Mat}(n \times n, \mathbb{R})$ genau dann negativ definit, wenn alle Hauptminoren von $-A$ positiv sind.

Bemerkung 16.5:

Achtung! Für *nicht* symmetrische Matrizen gilt das Kriterium nicht.

Betrachten Sie etwa die Matrix

$$C = \begin{pmatrix} 2 & -2 \\ 2 & -1 \end{pmatrix}$$

Ihre Hauptminoren sind 2 und $\det(C) = 2$ positiv, aber C ist offensichtlich nicht positiv definit z.B. $\langle e_2, Ce_2 \rangle = -1 < 0$

⋆ ⋆ ⋆

Anwendung der Definitheit. Der Begriff der Definitheit spielt an verschiedenen Stellen der Mathematik eine Rolle, so auch in der Analysis – beispielsweise wenn innerhalb einer Kurvendiskussion einer mehrdimensionalen Funktion die so genannte Hessematrix aufgestellt wird. Ich zeige Ihnen hier nur ein Beispiel zur Anwendung; die analytischen Grundlagen lesen Sie in der Literatur nach (etwa in „Mathematik der Physik für Dummies“).

Beispiel 16.6: Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \to \mathbb{R}$ eine zweimal stetig differenzierbare Funktion. Unter der Hessematrix von f im Punkt $x \in U$ versteht man die $(n \times n)$-Matrix

$$\mathrm{Hess}(f)(x) = \big(D_i D_j f(x)\big)_{1 \leq i,j \leq n}$$

Beachten Sie, dass die Hessematrix aufgrund des Satz von Schwarz über die doppelten Ableitungen symmetrisch ist.

Nun gilt als hinreichendes Kriterium für ein lokales Extremum von f folgendes: Es gilt $\mathrm{grad}(f)(x) = 0$ sowie

(a) $\text{Hess}(f)(x)$ positiv definit, so besitzt f in x ein striktes lokales Minimum.
(b) $\text{Hess}(f)(x)$ negativ definit, so besitzt f in x ein striktes lokales Maximum.

Betrachten Sie nun die Funktion $f(x,y) = x^2+y^2+xy$. Diese Funktion hat im Nullpunkt ein striktes lokales Minimum, da $\text{grad}\big((f)\,(0,0)\big) = (0,0)$ und die Hessesche Matrix

$$\text{Hess}(f)(x,y) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

positiv definit ist. (Betrachten Sie dazu beispielsweise das Kriterium mittels der Minoren; so gilt: $\text{Hess}(f)_{11} = 2$ und $\det \text{Hess}(f) = 3$.)

Ein anderes Beispiel, nun im $\mathbb{R}^3$, ist $f(x,y,z) = -\frac{1}{2}x^2 - \frac{5}{2}y^2 - 2xy - \frac{3}{2}z^2 - yz$. Diese Funktion hat im Nullpunkt ein striktes lokales Maximum, da $\text{grad}(f)(0,0,0) = (0,0,0)$ und das Negative der Hessematrix

$$-\text{Hess}(f)(x,y,z) = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 5 & 1 \\ 0 & 1 & 3 \end{pmatrix}$$

positiv definit ist.

17. Die Jordansche Normalform – die Königsklasse der Darstellungsformen

In diesem Kapitel...

- Grundlagen der Jordanschen Normalform verstehen
- Jordanblöcke und Jordankästchen in der Normalform verstehen
- Erste Eigenschaften der Jordankästchen kennenlernen
- Jordansche Normalform bei Differenzialgleichungen anwenden

In diesem Kapitel geht es um eine weitere, möglichst einfache (und damit schöne) Darstellungsform von Matrizen. Sie erinnern sich, dass bei der Diagonalisierung zwei Abbruchkriterien eine Rolle spielen, die eine Diagonalform verhindern können: So ist eine Matrix nicht diagonalisierbar, wenn das zugehörige charakteristische Polynom nicht vollständig in Linearfaktoren zerfällt oder die entsprechenden Eigenräume nicht die gewünschte Anzahl von Eigenvektoren liefern. Wir versuchen nun, das zweite Kriterium geschickt zu entschärfen.

Erste Gedanken zur Jordanschen Normalform. Der vollständige Zerfall in Linearfaktoren ist zumindest beim Übergang von den reellen zu den komplexen Zahlen überbrückbar, denn über den komplexen Zahlen hat jedes Polynom vom Grade n auch n Nullstellen (mit den entsprechenden Vielfachheiten). Dies löst natürlich nicht das Problem, dass es nicht immer hinreichend viele Eigenvektoren zum Zusammenstellen einer Basis gibt: Die geometrische Vielfachheit eines Eigenwertes kann auch über den komplexen Zahlen echt kleiner als die algebraische sein.

Die Grundidee in diesem Kapitel ist, dass wir von der sehr einfachen Diagonalgestalt etwas ablassen und auch eine leicht allgemeinere Matrix erlauben – nämlich eine Matrix, bei der die Hauptdiagonale immer noch die Eigenwerte enthält, zusätzlich in der ersten Nebendiagonalen (oberhalb der Hauptdiagonalen) neben Nullen aber auch Einsen erlaubt sind. Die restlichen Einträge verbleiben null. Eine solche Matrix hat immer noch eine sehr einfache Gestalt, und man kann hoffen, eine nicht diagonalisierbare Matrix zumindest in diese Form zu bringen.

Sei nun V in diesem Kapitel ein fixierter, endlich-dimensionaler $\mathbb{R}$- oder $\mathbb{C}$-Vektorraum.

Satz 17.1 – Jordansche Normalform:

Sei $f : V \to V$ ein Endomorphismus. Das charakteristische Polynom p_f zerfalle vollständig in Linearfaktoren. Dann gibt es eine Basis C von V, so dass

$$\mathrm{DM}_C(f) = \begin{pmatrix} \boxed{J_1} & & 0 \\ & \ddots & \\ 0 & & \boxed{J_k} \end{pmatrix}, \text{ wobei } J_i = \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix}.$$

Hierbei sind λ_i die Eigenwerte von f. Die Matrix $\mathrm{DM}_C(f)$ heißt **jordansche Normalform** oder kurz **Jordanform** von f. Die J_i heißen **Jordankästchen**. Alle Jordankästchen zu *einem* Eigenwert bilden einen **Jordanblock**.

Jeder Eigenwert besitzt also einen Jordanblock, der aus seinen Jordankästchen besteht. In jedem Jordankästchen stehen auf der oberen Nebendiagonalen nur Einsen. Da es auch Jordankästchen der

Größe (1×1) geben kann, sind auf der Nebendiagonalen eines Jordanblockes Einsen oder auch Nullen möglich.

Beispiel 17.2: Die folgende Matrix ist bereits eine Jordanform (von sich selbst):

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

Ihr charakteristisches Polynom lautet:

$$p_A(\lambda) = (1-\lambda)^2(2-\lambda)^3(3-\lambda)^5.$$

Schauen Sie sich die Matrix genau an. Sie erkennen die drei Jordanblöcke, für jeden Eigenwert einen. Innerhalb eines solchen Blockes erkennen Sie die Jordankästchen. Der Eigenwert 2 hat das größte Jordankästchen, das gleichzeitig schon der gesamte Jordanblock zu diesem Eigenwert ist.

Die Jordanform ist im Allgemeinen zwar keine Diagonalform, aber immer noch sehr nah dran. Außerdem sehen Sie aufgrund des Fundamentalsatzes der Algebra – ein Polynom n-ten Grades besitzt über dem Körper der komplexen Zahlen n Nullstellen –, dass Sie für jede lineare Abbildung über $\mathbb{C}$ eine solche einfache Normalform erhalten können. Das verrät Ihnen auch die Aussage des nächsten Tipps.

Tipp 17.3:

Für jeden Endomorphismus $f : V \to V$, dessen charakteristisches Polynom vollständig in Linearfaktoren zerfällt, existiert eine Basis C von V, so dass $\mathrm{DM}_C(f)$ eine Jordanform bildet.

Dies ist ein wesentlicher Unterschied und Vorteil gegenüber der Diagonalisierbarkeit. Weiterhin lassen sich folgende erste einfache Zusammenhänge betreffs der Jordankästchen und der (Nicht-)Abhängigkeit von der Wahl der Basis beweisen – schauen Sie mal.

Tipp 17.4:

- ✓ Die Anzahl der Jordankästchen zu einem Eigenwert λ ist gleich der Dimension des Eigenraums von λ.
- ✓ Die Anzahl der Einsen pro Eigenwert entspricht der Differenz seiner algebraischen und geometrischen Vielfachheit.
- ✓ Die Jordanform ist bis auf Permutationen der Jordankästchen eindeutig bestimmt.

Sie erkennen, dass die Einsen im Wesentlichen genau das Fehlen der Eigenvektoren in den Eigenräumen ausgleichen müssen, falls es einmal nicht genügend im Eigenraum geben sollte. Interessant ist dabei, dass dies bei der Jordanform stets (unter den angegebenen Voraussetzungen) möglich ist: Das ist ein fantastischer Fortschritt gegenüber der Diagonalisierbarkeit und für einen kleinen Preis von ein paar Einsen auf der Nebendiagonalen zu erreichen.

Die Trigonalisierung einer Matrix, die Sie im zwölften Kapitel kennengelernt haben, ist unter den gleichen Voraussetzungen ebenfalls möglich, allerdings ist der Preis viel höher: Das Ergebnis könnte eine volle obere Dreiecksmatrix sein.

Das Ziel ist also klar: Wir möchten die Jordanform verstehen lernen. Lassen Sie uns erste einfache Matrizen(-typen) betrachten und uns behutsam weiter vortasten.

Beispiel 17.5:

✓ Jede Diagonalmatrix ist offenbar bereits in jordanscher Normalform.

✓ Ist ein Endomorphismus in eine Matrix D diagonalisierbar, so ist seine Jordanform genau diese Diagonalmatrix D.

✓ Die Matrix $\begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 5 & 1 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$ besteht aus genau zwei Jordanblöcken, für jeden Eigenwert einen.

Der Block zum Eigenwert fünf hat zwei Jordankästchen; der Block zum Eigenwert vier besteht dagegen nur aus dem einen Jordankästchen.

Wie die jordansche Normalform aufgebaut ist und funktioniert. Schauen wir uns ein Jordankästchen im Detail an, um besser zu verstehen, wie eine solche Form zustande kommt. Dadurch wollen wir auch eine Idee bekommen, wie wir unsere gesuchte (Jordan-)Basis, die zum Jordankästchen führt, aufzubauen haben. Sei also ein Jordankästchen zu einem Eigenwert λ gegeben:

$$\begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \in \mathrm{Mat}\,(k \times k,\ \mathbb{K})\,.$$

Um ein solches Kästchen zu bekommen, brauchen wir die Transformationsmatrix T, die eine gegebene Abbildung $A \in \mathrm{Mat}\,(n \times n, \mathbb{K})$ in eine Jordanform J überführt. Hierbei sei einfachheitshalber J das oben angeführte gesamte Jordankästchen. Im Allgemeinen können natürlich mehrere Jordankästchen in einem Jordanblock vorkommen – wie Sie dies im vorangegangenen Beispiel am Eigenwert fünf erkennen können. In einem solchen Fall kann auch $k < n$ sein. Ich werde daher weiterhin allgemein über k und n sprechen. Mit unserer Annahme ist dann allerdings $k = n$ und es gilt für eine Transformationsmatrix T die Gleichung:

$$T^{-1} \cdot A \cdot T = J.$$

Dann erhalten wir durch Umformung: $A \cdot T = T \cdot J$. Lesen Sie dieses Matrizenprodukt spaltenweise als einzelne Multiplikationen und bezeichnen Sie die Spaltenvektoren von T dabei mit $v_1, \ldots, v_k$. Dann kommen Sie zu den folgenden wichtigen Grundgleichungen:

$$\begin{aligned} Av_1 &= \lambda v_1\,, \\ Av_2 &= 1v_1 + \lambda v_2\,, \\ Av_3 &= 1v_2 + \lambda v_3\,, \\ Av_4 &= 1v_3 + \lambda v_4\,, \\ &\text{und so weiter.} \end{aligned}$$

Lassen Sie uns diese Gleichungen so umformen, dass gleiche Variablenvorkommen einfach zusammengeführt werden:

$$\begin{aligned}(A-\lambda E_n)\,v_1 &= 0\,,\\ (A-\lambda E_n)\,v_2 &= v_1\,,\\ (A-\lambda E_n)\,v_3 &= v_2\,,\\ (A-\lambda E_n)\,v_4 &= v_3\,,\\ &\text{und so weiter.}\end{aligned}$$

Schauen Sie genau über diese Gleichungen. Wenn Sie v_4 kennen, können Sie anhand dieser Gleichungen schrittweise die gesamte Kette von Vektoren bestimmen: Mit v_4 zunächst v_3, dann v_2 und schließlich v_1.

Für solche endlich viele gegebene Vektoren $v_1, \ldots, v_k$ bekommen wir daher folgende später die Grundlage bildende Gleichungen, welche später die Grundlage für die zu findenden (so genannten Jordan-)Ketten bilden werden:

$$\begin{aligned}&v_k\\ v_{k-1} &= (A-\lambda E_n)\,v_k\,,\\ v_{k-2} &= (A-\lambda E_n)\,v_{k-1} = (A-\lambda E_n)^2\,v_k\,,\\ v_{k-3} &= (A-\lambda E_n)\,v_{k-2} = (A-\lambda E_n)^3\,v_k\,,\\ &\text{und so weiter.}\end{aligned}$$

Das bedeutet, dass Sie tief nach einem geeigneten Vektor v_k graben müssen, um dann in einem Schwung gleich durch die Methode der iterierten Anwendung von $(A-\lambda E_n)$ eine Kette von insgesamt k gesuchten Vektoren zu bekommen. Das wird unser Ziel sein. Betrachten Sie dazu auch die Abbildung 27.

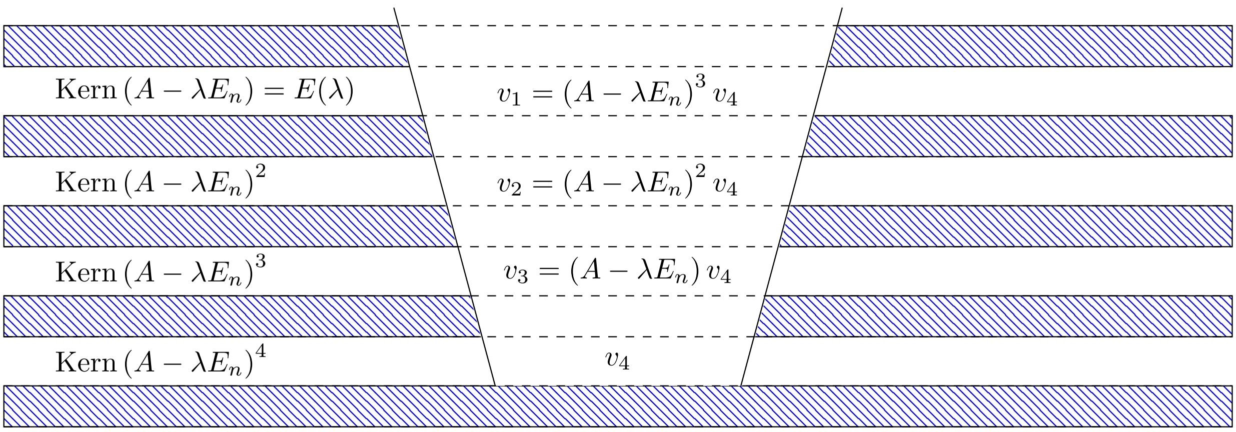


Abbildung 27: Durch Abstieg entlang der iterierten Kerne die wertvollen Jordanketten bergen

Mit Jordanketten zum Ziel. Sei nun also v_1 ein Eigenvektor von A. Dann liegt dieser in der Menge Kern $(A-\lambda E_n)$. Weiterhin sei $(A-\lambda E_n)\,v_2 = v_1$. Dann ist v_1 als Eigenvektor verschieden von null und somit ist v_2 kein Eigenvektor von A, aber es ist $(A-\lambda E_n)^2\,v_2 = (A-\lambda E_n)\,v_1 = 0$. Es gilt also insgesamt:

$$v_2 \,\in\, \text{Kern}\,(A-\lambda E_n)^2 \,\setminus\, \text{Kern}\,(A-\lambda E_n)\,.$$

Ähnlich zeigt man die folgenden wichtigen Gleichungen für alle $j = 1, \ldots, l$:

$$v_j \in \operatorname{Kern}(A - \lambda E_n)^j \setminus \operatorname{Kern}(A - \lambda E_n)^{j-1}.$$

Betrachten Sie unter diesem Gesichtspunkt noch einmal die Abbildung **??**. Solche Ketten von Vektoren spielen eine dermaßen ausgezeichnete Rolle und zwar für sich allein *und* im gegenseitigen Zusammenspiel, dass Sie einen Namen bekommen.

Definition 17.6 – Jordankette, Hauptvektor:

Die Vektoren $v_1, \ldots, v_k$ mit den oben beschriebenen Eigenschaften bezeichnet man als **Jordankette**. Der Vektor v_k heißt **Hauptvektor k-ter Stufe zum Eigenwert λ von A**, falls gilt:

$$v_k \in \operatorname{Kern}(A - \lambda E_n)^k \setminus \operatorname{Kern}(A - \lambda E_n)^{k-1}.$$

Der gerade betrachtete Vektor v_k ist also ein Hauptvektor der k-ten Stufe. Solche Vektoren wollen wir fortan finden. Dabei werden wir pro Jordankästchen *eine* Jordankette suchen, die von einem solchen Hauptvektor bestimmt wird.

Tipp 17.7:

Eine Jordankette der Länge k startet immer mit einem Hauptvektor der Stufe k. Somit hat sie die Gestalt:

$$\Big((A - \lambda E_n)^{k-1} v\ ,\ (A - \lambda E_n)^{k-2} v\ ,\ \ldots\ ,\ (A - \lambda E_n)^2 v\ ,\ (A - \lambda E_n) v\ ,\ v\ \Big)$$

für einen (Haupt-)Vektor v, der zwar in der k-ten Stufe $\operatorname{Kern}(A - \lambda E_n)^k$, nicht aber in der $(k-1)$-ten Stufe $\operatorname{Kern}(A - \lambda E_n)^{k-1}$ liegt.

Im nächsten Kapitel zeige ich Ihnen etwas detaillierter, wie Sie solche Methoden verfeinern und sowohl die einzelnen Größen der Jordankästchen bestimmen als auch letztendlich eine solche Jordanbasis finden, so dass die diesbezüglich darstellende Matrix gerade die Jordanform hervorbringt. Dafür müssen Sie allerdings etwas Zeit und Kraft investieren – das vierzehnte Kapitel wird bei der Aufnahme der Strategien etwas Schweiß kosten. Im Fünfzehnten können Sie gleich weiter schwitzen: Hier rechnen wir praktische Beispiele durch und finden erfolgreich Jordanbasen. Schnallen Sie sich an – es kann holprig werden!

⋆ ⋆ ⋆

Anwendung der Jordanschen Normalform bei Differenzialgleichungen. Bevor Sie den steinigen Weg gehen, sollen Sie wissen, warum es sich lohnt, soviel Kraft für die Jordanform zu investieren. Das folgende Beispiel möge Ihnen zeigen, wie sinnvoll und praktisch in der Mathematik Normalformen anwendbar sind. Konkret wird es um homogene lineare Differenzialgleichungen erster Ordnung mit konstanten Koeffizienten gehen. Ich werde hier nicht weiter auf die analytische Theorie eingehen. Schauen Sie in die Literatur, beispielsweise in das Buch *Vorkurs Mathematik für Ingenieure für Dummies*.

Sie können mir auch gern jetzt schon vertrauen und zum nächsten Kapitel springen, falls Sie noch nichts über Differenzialgleichungen wissen. Solch analytisches Grundwissen brauchen Sie nämlich für diesen Abschnitt.

Beispiel 17.8: Wir wenden die Jordanform im Lösungsalgorithmus für homogene lineare Differenzialgleichungen erster Ordnung mit konstanten Koeffizienten an.

Sei $A \in \mathrm{Mat}\,(n \times n, \mathbb{C})$ samt Anfangswert $f_0 \in \mathbb{C}^n$.
Betrachten Sie dann für Funktionen $f : \mathbb{C} \to \mathbb{C}^n$ die Differenzialgleichung:

$$\frac{\mathrm{d}}{\mathrm{d}t} f(t) = A \cdot f(t) \text{ mit der Anfangswertbedingung } f(0) = f_0.$$

Mithilfe von analytischen Betrachtungen erkennen Sie, dass analog zum eindimensionalen Fall auch der hier folgende Ansatz die gegebene Differenzialgleichung löst:

$$f(t) = \mathrm{e}^{At} \cdot f_0 \quad \text{mit} \quad \mathrm{e}^{At} = \sum_{j=0}^{\infty} \frac{A^j t^j}{j!} \quad \text{und} \quad A^0 = E_n.$$

Es ist allerdings Folgendes zu prüfen:

Konvergiert die Matrizenreihe $\sum_{j=0}^{\infty} \frac{A^j t^j}{j!}$ innerhalb von $\mathrm{Mat}(n \times n, \mathbb{C})$?

Hierbei kommen Techniken des Diagonalisierens und der Jordanform ins Spiel, wie Sie gleich sehen werden.

Die Existenz (also die Konvergenz der dahinter liegenden Reihe) von e^{At} kann man in mehreren Schritten wie folgt zeigen. Ich deute diese hier auch nur an – keine Sorge, dieser Schnellkurs ist und bleibt einer über lineare Algebra:

(a) Für eine Diagonalmatrix

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

konvergiert die Reihe e^{Dt} sogar absolut. (Absolute Konvergenz ist eine spezielle Art der Konvergenz: Hier konvergiert nicht nur die Reihe, sondern sogar die Reihe der Beträge der eingehenden Zahlenfolge.) Für den Grenzwert und damit den Wert der Reihe gilt:

$$\mathrm{e}^{Dt} = \begin{pmatrix} \mathrm{e}^{\lambda_1 t} & & 0 \\ & \ddots & \\ 0 & & \mathrm{e}^{\lambda_n t} \end{pmatrix}.$$

(b) Für eine *nilpotente* Matrix N, also eine Matrix mit der Eigenschaft, dass es eine Zahl $k \in \mathbb{N}$ mit $N^k = 0$ gibt, überzeugt man sich leicht davon, dass die Reihe $\mathrm{e}^{Nt} = \sum_{j=0}^{k-1} \frac{N^j t^j}{j!}$ ebenfalls absolut konvergiert.

(c) Sei nun $C \in \mathrm{Mat}(n \times n, \mathbb{C})$ derart, dass die Reihe e^{Ct} absolut konvergiert und weiterhin $A = B^{-1} \cdot C \cdot B$ gilt. Dann konvergiert die folgende Reihe:

$$\mathrm{e}^{At} = \sum_{j=0}^{\infty} \frac{A^j t^j}{j!} = B^{-1} \cdot \left(\sum_{j=0}^{\infty} \frac{C^j t^j}{j!} \right) \cdot B = B^{-1} \cdot \mathrm{e}^{Ct} \cdot B.$$

(d) Sei $A = B + C$ mit $B \cdot C = C \cdot B$. Weiterhin konvergieren die Reihen e^{Bt} und e^{Ct} jeweils absolut. Dann konvergiert aufgrund der Cauchyschen Summenformel auch die Reihe $\sum_{j=0}^{\infty} \frac{A^j t^j}{j!}$ absolut. Dabei gilt: $\mathrm{e}^{At} = \mathrm{e}^{Bt} \cdot \mathrm{e}^{Ct}$.

Dies waren harte analytische Brocken. Versuchen Sie diese bitte hinzunehmen. Vergessen Sie nicht: Wenn Konvergenz von Reihen bisher kein Thema für Sie war, dann können Sie das Beispiel hier auch ignorieren. Es wird den Lesefluss nicht weiter beeinflussen.

Wir benutzen diese Erkenntnisse wie folgt: Sei also J die Jordanform von A. Damit gilt $A = T^{-1} \cdot J \cdot T$ für eine geeignete Transformationsmatrix T, die durch die Jordanbasis gegeben ist. Wir erhalten:

$$J = \underbrace{\begin{pmatrix} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}}_{=D} + \underbrace{\begin{pmatrix} 0 & \star & & 0 \\ & \ddots & \ddots & \\ & & \ddots & \star \\ 0 & & & 0 \end{pmatrix}}_{=N}, \text{ wobei } \quad \star \in \{0,1\}\,.$$

Sie sehen, dass wir die Jordanform in zwei Summanden zerlegen: Einmal in den Anteil einer Diagonalmatrix, in der die Eigenwerte stehen, zum anderen in den nilpotenten Anteil, in dem außer auf der ersten Nebendiagonalen nur Nullen stehen. Auf dieser sind allerdings die noch verbleibenden Einsen zu finden.

Man kann zeigen, dass dann $D \cdot N = N \cdot D$ gilt.
Damit folgt schließlich mittels der Argumente von oben:

$$\begin{aligned} \mathrm{e}^{At} \quad & \underbrace{=}_{(c)} \quad T^{-1} \cdot \mathrm{e}^{Jt} \cdot T \underbrace{=}_{(d)} T^{-1} \cdot \mathrm{e}^{Dt} \cdot \mathrm{e}^{Nt} \cdot T \\ & \underbrace{=}_{(a),(b)} \quad T^{-1} \cdot \begin{pmatrix} \mathrm{e}^{\lambda_1 t} & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & \mathrm{e}^{\lambda_n t} \end{pmatrix} \cdot \left(\sum_{j=0}^{k-1} \frac{N^j t^j}{j!} \right) \cdot T. \end{aligned}$$

Dies zeigt die Konvergenz und damit die gewünschte Existenz von e^{At}. Geschafft!

Tipp 17.9:

Übrigens, mit der gerade betrachteten Formel kann man den Ausdruck e^{At} auch praktisch berechnen! Das lasse ich aber einfach mal so wirken. Sie soll(t)en nur wahrnehmen, dass diese theoretischen Gedanken in der Analysis auch praktisch anwendbar sind.

Das war gerade zu schnell? – Macht nichts. Bedenken Sie, Ausblicke auf praktische Anwendungen steigern vielleicht Ihre Motivation, unsere algebraische Theorie durchzustehen. Wenn Sie nun also mehr Details über die Jordanform kennenlernen wollen, dann folgen Sie mir ins nächste Kapitel, in dem wir viel tiefer einsteigen werden. Der Lohn wird es sein, dass wir eine Basis finden können, in der eine gegebene Abbildung in ihrer Jordanform erstrahlt. Bereit? – Dann los zum nächsten Kapitel!

⋆ ⋆ ⋆

Aufgabe 17.10: Überlegen Sie sich, ob Sie bereits mit den Kenntnissen aus diesem Kapitel entscheiden können, wie die Jordanform samt Jordanbasis einer Matrix aussehen kann. Entscheiden Sie dies beispielsweise anhand der folgenden Matrix:

$$A = \begin{pmatrix} 5 & 8 \\ -2 & -3 \end{pmatrix}.$$

Aufgabe 17.11: Betrachten Sie die Matrix:

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

(a) Geben Sie die Anzahl der Jordankästchen zu den jeweiligen Eigenwerten an.
(b) Wie viele Einsen befinden sich in der Jordanform auf der Nebendiagonalen?
(c) Geben Sie eine mögliche Jordanform an.

Aufgabe 17.12: Gegeben sei die Matrix $A = \begin{pmatrix} 1 & 1 & -3 \\ 0 & 1 & 1 \\ 1 & 0 & 4 \end{pmatrix}$. Überlegen Sie sich, wie viele Einsen auf der Nebendiagonalen stehen. Berechnen Sie anschließend eine Jordankette maximaler Länge zum Eigenwert Ihrer Wahl. Was fehlt Ihnen noch zum Bestimmen einer Jordanbasis, die zur Jordanform führt?

⋆ ⋆ ⋆

Auf einen Blick ...

- Die Jordansche Normalform ist fast eine Diagonalform, die potentiell auf einer Nebendiagonalen zusätzliche Einsen beherbergt.
- Die Anzahl der Einsen entspricht beim Diagonalisieren gerade der Anzahl der fehlenden Eigenvektoren zu einer Basis.
- Pro Jordankästchen werden die entsprechend benötigten Basisvektoren zur Darstellung in Jordanform aus den Jordanketten gewonnen: Dazu benötigt man einen Hauptvektor aus einem hinreichend oft iterierten Kern und kämpft sich anschließend entlang der Jordankette wieder zum Eigenraum hoch.
- Die Einsen auf der Nebendiagonalen sind nicht beliebig verteilbar, sondern werden durch Jordankästchen bestimmt: Je mehr Eigenvektoren in einer potentiellen Basis fehlen, desto mehr Einsen existieren. Eine solche Eins entspricht einem Vektor aus einem iterierten Kern. Diese Vektoren stammen aus einer Jordankette, die auf einem Hauptvektor basiert. Je tiefer Sie für den Hauptvektor dieser Kette in den Kerniteraten graben müssen, desto größer ist das Jordankästchen, in dem die betrachtete Eins steht.

18. Lösungshinweise

In diesem Kapitel finden Sie Lösungshinweise zu ausgewählten Aufgaben der einzelnen Kapitel. Es handelt sich teilweise auch wirklich nur um Hinweise oder Lösungen und soll Sie beim selbstständigen Lösen der Aufgaben unterstützen.

Aufgaben zu Kapitel 1. **Aufgabe 18.1:** Berechnen Sie die Lösungsmenge des nachfolgenden homogenen LGS:

$$\begin{aligned} x_1 + x_2 + 0x_3 + x_4 &= 0 \\ 2x_1 + x_2 + x_3 + 0x_4 &= 0 \\ 3x_1 + 0x_2 + 0x_3 + 2x_4 &= 0 \\ 4x_1 + x_2 + 0x_3 + 3x_4 &= 0 \end{aligned}$$

Welchen Zusammenhang gibt es zwischen dieser Lösungsmengen und der Lösungsmenge:

$$\mathbb{L}_{\text{inh}} = \left\{ (x_1, x_2, x_3, x_4) \;\middle|\; x_1 = \frac{2}{3}(1 - x_4)\ ,\ x_2 = \frac{1}{3}(1 - x_4)\ ,\ x_3 = \frac{1}{3} + \frac{5}{3}x_4 \right\}$$

des inhomogenen LGS:

$$\begin{aligned} x_1 + x_2 + 0x_3 + x_4 &= 1 \\ 2x_1 + x_2 + x_3 + 0x_4 &= 2 \\ 3x_1 + 0x_2 + 0x_3 + 2x_4 &= 2 \\ 4x_1 + x_2 + 0x_3 + 3x_4 &= 3 \end{aligned}$$

Lösung.
Man berechnet mit dem Gauß-Algorithmus:

$$\begin{array}{cccc|c} 1 & 1 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 0 \\ 3 & 0 & 0 & 2 & 0 \\ 4 & 1 & 0 & 3 & 0 \end{array} \rightsquigarrow \begin{array}{cccc|c} 4 & 1 & 0 & 3 & 0 \\ 0 & 1 & 2 & -3 & 0 \\ 0 & 0 & -6 & 10 & 0 \end{array}$$

und erhält die Lösungsmenge:

$$\mathbb{L}_{\text{Hom}} = \left\{ (x_1, x_2, x_3, x_4) \;\middle|\; x_1 = -\frac{2}{3}x_4\ ,\ x_2 = -\frac{1}{3}x_4\ ,\ x_3 = \frac{5}{3}x_4 \right\}$$

Man erkennt nun, dass:

$$\mathbb{L}_{\text{inh}} = \mathbb{L}_{\text{Hom}} + \left(\frac{2}{3}, \frac{1}{3}, \frac{1}{3}, 0 \right)$$

wobei der letzte Vektor eine spezielle Lösung des inhomogenen LGS ist!

Aufgabe 18.2: Konstruieren Sie lineare Gleichungssysteme mit den folgenden Lösungsmengen:

$$\begin{aligned} &\mathbb{L}_1 = \left\{ (x_1, x_2, x_3, x_4) \;\middle|\; x_1 = 3x_4\ ,\ x_2 = 6x_4\ ,\ x_3 = -2x_4 \right\} \\ &\mathbb{L}_2 = \left\{ (x_1, x_2, x_3, x_4) \;\middle|\; x_1 = 5\lambda + \rho\ ,\ x_2 = 6\lambda - 2\rho\ ,\ x_3 = -2\lambda\ ,\ x_4 = -2\rho,\ \text{für } \lambda, \rho \in \mathbb{R} \right\} \end{aligned}$$

Lösung.

zu $\mathbb{L}_1$: $\left(\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & -3 & 0 \\ 0 & 1 & 0 & 0 & -6 & 0 \\ 0 & 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \end{array}\right)$ zu $\mathbb{L}_2$: $\left(\begin{array}{cccc|c} 1 & 0 & \frac{5}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 3 & -1 & 0 \end{array}\right)$

⋆ ⋆ ⋆

Aufgaben zu Kapitel 2. **Aufgabe 18.3:** Welche der folgenden Teilmengen sind Untervektorräume des $\mathbb{R}^3$?

$$U_1 := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : 3x + 7y - 8z = 0, \quad x, y, z \in \mathbb{R} \right\}$$

$$U_2 := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : x + y + z = 1, \quad x, y, z \in \mathbb{R} \right\}$$

$$U_3 := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : x \cdot y \cdot z = 0, \quad x, y, z \in \mathbb{R} \right\}$$

$$U_4 := \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : x^2 + y^2 = z^2, \quad x, y, z \in \mathbb{R} \right\}$$

Lösung.
Sie müssen nur die Untervektorraumaxiome überprüfen: zu U_1: U_1 ist ein UVR (Lösungsmengen von homogenen LGS sind immer Untervektorräume)

zu U_2: U_2 ist kein UVR: $\underbrace{(0,0,1)}_{\in U_2} + \underbrace{(0,1,0)}_{\in U_2} \notin U_2$

zu U_3: U_3 ist kein UVR: $\underbrace{(0,1,1)}_{\in U_3} + \underbrace{(1,0,0)}_{\in U_3} \notin U_3$

zu U_4: U_4 ist kein UVR: $\underbrace{(0,1,1)}_{\in U_4} + \underbrace{(1,0,1)}_{\in U_4} \notin U_4$

⋆ ⋆ ⋆

Aufgaben zu Kapitel 6. **Aufgabe 18.4:** Kann es auf einem endlich dimensionalen Vektorraum V Endomorphismen $f : V \to V$ geben, die injektiv und nicht surjektiv oder nicht injektiv und surjektiv sind. Beweisen oder widerlegen Sie ihre Vermutung!

Tipp 18.5:

Beachten Sie, dass f injektiv ist genau dann, wenn $\operatorname{Kern}(f) = 0$ und verwenden Sie den Dimensionssatz.

Lösung.
Die Antwort ist, dass es keine solchen Endomorphismen geben kann!
Angenommen f injektiv, dann ist $\text{Kern}(f) = 0$ und mit der Dimensionsformel erhält man:

$$\underbrace{\dim(\text{Kern}(f))}_{=0} + \dim(\text{Bild}(f)) = \dim(V), \text{ also } \text{Bild}(f) = V,$$

das bedeutet, f ist surjektiv.

Analog dazu: Angenommen f surjektiv, dann ist $\dim(\text{Bild}(f)) = \dim(V)$ und mit der Dimensionsformel erhält man:

$$\dim(\text{Kern}(f)) + \dim(\text{Bild}(f)) = \dim(V) \quad \Leftrightarrow \quad \text{Kern}(f) = 0,$$

also muss f injektiv sein.

Aufgabe 18.6: Sei $f : V \to W$ für V, W zwei $\mathbb{Q}-$Vektorräume. Man zeige:
Aus $f(x+y) = f(x) + f(y)$ für alle $x, y \in V$ folgt, dass f linear ist.
Lösung.
Jedes $\lambda \in \mathbb{Q}$ hat die Darstellung $\lambda = \frac{p}{q}$ mit $p \in \mathbb{Z}, q \in \mathbb{N}$. Damit können wir Folgendes nachrechnen:

$$f(qx) = f(\underbrace{x + \ldots + x}_{q\text{-mal}}) = qf(x)$$

und analog:

$$f\left(q\frac{x}{q}\right) = qf\left(\frac{x}{q}\right) \quad \Leftrightarrow \quad \frac{1}{q}f(x) = f\left(\frac{x}{q}\right)$$

Außerdem gilt $f(-x) = -f(x)$, denn $f(x) + f(-x) = f(x - x) = f(0) = 0$.
Somit folgt auch $f(px) = pf(x)$ für $p \in \mathbb{Z}$.

Schließlich erhalten wir:

$$f(\lambda x) = f\left(\frac{p}{q}x\right) = pf\left(\frac{x}{q}\right) = \frac{p}{q}f(x) = \lambda f(x)$$

und somit ist f tatsächlich linear.

⋆ ⋆ ⋆

Aufgaben zu Kapitel 7. **Aufgabe 18.7:** Sei $f : V \to V$ ein Endomorphismus auf dem n dimensionalen Vektorraum V über einem Körper $\mathbb{K}$ mit $1 + 1 \neq 0$ und $f^2 = f$. Man zeige die Existenz einer Basis von V, so dass

$$\text{DM}(f) = \begin{pmatrix} E_k & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{mit } k \in \mathbb{N}, \text{ bzgl. dieser Basis}$$

Hinweis: Überlegen Sie sich, wie die darstellende Matrix bezüglich der Vektoren $f(b_1), \ldots, f(b_l)$, wobei die b_i hier eine Basis von $\text{Bild}(f)$ sind, aussieht. **Lösung.**
Wir wissen, dass der Vektorraum V in $\text{Bild}(f)$ und $\text{Kern}(f)$ zerfällt. Sei $b_1, \ldots, b_l$ eine Basis von $\text{Bild}(f)$ und $c_{l+1}, \ldots, c_n$ eine Basis von $\text{Kern}(f)$. Jetzt ist für jedes b_i: $b_i = f(x)$ für ein $x \in V$, daraus folgt $f(b_i) = f^2(x) = f(x) = b_i$. Trivialerweise ist natürlich auch $f(c_i) = 0$ und damit wissen wie f auf die Basis $\{b_1, \ldots, b_l, c_{l+1}, \ldots, c_n\}$ angewendet aussieht.

⋆ ⋆ ⋆

Aufgaben zu Kapitel 8. **Aufgabe 18.8:** Wir betrachten die Prüfmatrix H:

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

Warum gibt die Prüfmatrix die Stelle eines 1 Bit Fehlers an?

Hinweis: Multiplizieren Sie H mit allen kanonischen Einheitsvektoren und betrachten Sie hinterher die Struktur des Resultats.

Lösung.

Multipliziert man die Prüfmatrix mit einem fehlerfreien Vektor x, so gilt $Hx = 0$. Ist der Vektor mit einem 1-Bit Fehler an der Position i behaftet, so lässt er sich darstellen als $x + e_i$. Jetzt ergibt das Testen mit der Prüfmatrix: $H(x + e_i) = He_i$. Alles was man also noch nachzurechnen hat ist, dass $He_i = i$ in der b-adischen Darstellung für $b = 2$.

⋆ ⋆ ⋆

Aufgaben zu Kapitel 9. **Aufgabe 18.9:** Invertieren Sie die nachfolgenden Matrizen:

$$A_1 = \begin{pmatrix} 1 & 2 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 2 & 2 & 2 \\ 3 & 1 & 1 \\ 2 & 1 & 4 \end{pmatrix} \qquad A_3 = \begin{pmatrix} 1 & 1 & 3 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

Lösung.

Mit dem Invertierungs-Algorithmus kommt man auf:

$$A_1^{-1} = \frac{1}{3}\begin{pmatrix} -1 & 2 & 2 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \qquad A_2^{-1} = \frac{1}{12}\begin{pmatrix} -3 & 6 & 0 \\ 10 & -4 & -4 \\ -1 & -2 & 4 \end{pmatrix}$$

$$A_3^{-1} = \begin{pmatrix} -3 & -1 & 5 \\ 1 & 1 & -2 \\ 1 & 0 & -1 \end{pmatrix}$$

Aufgabe 18.10: Zeigen Sie, dass die (2×2)-Matrix:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad a, b, c, d \in \mathbb{R}$$

genau dann invertierbar ist, wenn $ad - bc \neq 0$

Lösung.

Wir verwenden unseren Invertierungs-Algorithmus und gehen davon aus, dass mindestens einer der Koeffizienten ungleich Null ist. Sei also o.B.d.A $a \neq 0$ (ansonsten vertauscht man die Zeilen und Spalten der Matrix bis der erste Eintrag in der ersten Zeile ungleich Null ist):

$$\left(\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|cc} a & b & 1 & 0 \\ 0 & \frac{-bc}{a} + d & \frac{-c}{a} & 1 \end{array}\right) \rightsquigarrow$$

$$\left(\begin{array}{cc|cc} a & 0 & \frac{bc}{ad-bc}+1 & \frac{-ab}{ad-bc} \\ 0 & -bc+da & -c & a \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 0 & \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ 0 & 1 & \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{array}\right)$$

Dadurch erhalten wir unsere bekannte Matrixinverse einer (2×2)-Matrix:

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

und dafür muss natürlich $ad - bc \neq 0$ gelten.

Aufgabe 18.11: Wann ist die Matrix

$$B = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & e \end{pmatrix} \text{ für } a, b, c, d, e \in \mathbb{R}$$

invertierbar und wie sieht ihre Inverse aus?

Hinweis: Konstruieren Sie die Inverse der (2×2)-Blockmatrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, dann lässt sich die Inverse von B aus 2 Blockmatrizen „zusammenbauen".

Lösung.

Beachten Sie, dass wegen der Struktur der Matrixmultiplikation folgender Zusammenhang gilt:

$$\begin{pmatrix} A_1 & 0 \\ 0 & B_1 \end{pmatrix}\begin{pmatrix} A_2 & 0 \\ 0 & B_2 \end{pmatrix} = \begin{pmatrix} A_1A_2 & 0 \\ 0 & B_1B_2 \end{pmatrix} \qquad \text{für } A_1, A_2, B_1, B_2 \text{ Blockmatrizen}$$

Wir wissen außerdem aus der vorherigen Aufgabe, dass

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

und erhalten insgesamt:

$$BB^{-1} = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & e \end{pmatrix}\begin{pmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} & 0 \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} & 0 \\ 0 & 0 & \mathrm{e}^{-1} \end{pmatrix} = E_3$$

Es muss also $ad - bc \neq 0$ und $e \neq 0$ gelten.

Aufgabe 18.12: Sei $A = (a_{ij})_{i;j=1,\ldots,n}$ eine $(n\times n)$-Matrix über $\mathbb{R}$ mit $|a_{ij}| \leq q < \frac{1}{2n}$ für jedes $1 \leq i, j \leq n$. Man zeige, dass

$$(E_n - A)^{-1} = \sum_{k=0}^{\infty} A^k.$$

Hinweis: Man zeige, dass $\lim_{k\to\infty} A^k = 0$ und die Existenz von $\sum_{k=0}^{\infty} A^k$ mithilfe der geometrischen Summenformel. Anschließend betrachtet man $(E_n - A)\cdot\sum_{k=0}^{N} A^k = E_n - A^{N+1}$ und nutzt aus, dass dieser Ausdruck für $N\to\infty$ gegen die Einheitsmatrix konvergiert.

Lösung.

Schritt 1:

Sei $A^k = \left(a_{ij}^{(k)}\right)_{i,j=1,\ldots,n}$, dann gilt:

$$\left|a_{ij}^{(2)}\right| \leq \sum_{m=1}^{n} |a_{im}|\left|a_{mj}\right| \leq \sum_{m=1}^{n} \frac{1}{2n} \cdot \frac{1}{2n} \leq \frac{1}{2^2 n}$$

und analog:

$$\left|a_{ij}^{(3)}\right| \leq \sum_{m=1}^{n} \left|a_{im}^{(2)}\right|\left|a_{mj}\right| \leq \sum_{m=1}^{n} \frac{1}{2^2 n} \cdot \frac{1}{2n} \leq \frac{1}{2^3 n}$$

also erhält man insgesamt induktiv:

$$\left|a_{ij}^{(k)}\right| \leq \sum_{m=1}^{n} \left|a_{im}^{(k-1)}\right|\left|a_{mj}\right| \leq \sum_{m=1}^{n} \frac{1}{2^{k-1} n} \cdot \frac{1}{2n} \leq \frac{1}{2^k n}$$

und wie gewünscht die Aussage $\lim_{k\to\infty} a_{ij}^{(k)} = 0$.

<u>Schritt 2:</u>
Mit der Ungleichung aus dem ersten Schritt erhalten wir für jeden Koeffizienten:

$$\left|\left(\sum_{k=0}^{N} A^k\right)_{ij}\right| = \left|\sum_{k=0}^{N} a_{ij}^k\right| \leq \sum_{k=0}^{N} \left|a_{ij}^k\right| \leq \sum_{k=0}^{N} \frac{1}{2^k n} \leq \sum_{k=0}^{\infty} \frac{1}{2^k n} = \frac{2}{n}$$

also ist die Summendarstellung jedes Koeffizienten absolut konvergent.

<u>Schritt 3:</u>
Jetzt fehlt nur noch die einfache Feststellung:

$$\begin{aligned}(E_n - A) \cdot \sum_{k=0}^{\infty} A^k &= \lim_{N\to\infty} (E_n - A) \cdot \sum_{k=0}^{N} A^k &&= \lim_{N\to\infty} \sum_{k=0}^{N} A^k - A^{k+1} \\ &= \lim_{N\to\infty} E_n - A^{N+1} &&= E_n\end{aligned}$$

⋆ ⋆ ⋆

Aufgaben zu Kapitel 10. **Aufgabe 18.13:** Sei f ein Endomorphismus $f : V \to V$, V ein n-dimensionaler $\mathbb{R}$-Vektorraum und B eine Basis von V. Dann definieren wir:

$$\mathrm{Spur}(A) := \sum_{i=1}^{n} a_{ii}$$

für eine $(n \times n)$-Matrix $A = \left(a_{ij}\right)_{i,j=1\ldots n}$ und

$$\mathrm{Spur}(f) := \mathrm{Spur}(\mathrm{DM}_B(f)) = \sum_{i=1}^{n} (\mathrm{DM}_B(f))_{ii}$$

Zeigen Sie, dass der letzte Ausdruck wohldefiniert (unabhängig von der Basis B) ist!

<u>Hinweis:</u> Sie müssen zeigen, dass $\mathrm{Spur}(\mathrm{DM}_B(f)) = \mathrm{Spur}(\mathrm{DM}_C(f))$ für jede andere Basis C von V. Beweisen Sie dazu $\mathrm{Spur}(AB) = \mathrm{Spur}(BA)$ und dann verwenden Sie $\mathrm{DM}_C(f) = (T_B^C)^{-1}\mathrm{DM}_B(f)T_B^C$.

Lösung.

<u>Schritt 1:</u>

Wir beweisen $\operatorname{Spur}(AB) = \operatorname{Spur}(BA)$:

$$\operatorname{Spur}(AB) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ji} = \underbrace{\sum_{\hat{j}=1}^{n}\sum_{\hat{\imath}=1}^{n} a_{\hat{j}\hat{\imath}}b_{\hat{\imath}\hat{j}}}_{\text{umbenennen: } \hat{j}:=i,\ \hat{\imath}:=j} = \sum_{\hat{\imath}=1}^{n}(BA)_{\hat{\imath}} = \operatorname{Spur}(BA)$$

<u>Schritt 2:</u>

Wir erhalten jetzt:

$$\begin{aligned}
\operatorname{Spur}\left(\mathrm{DM}_C(f)\right) &= \operatorname{Spur}\left(\left\{(T_B^C)^{-1}\mathrm{DM}_B(f)\right\}T_B^C\right)\\
&= \operatorname{Spur}\left(T_B^C\left\{(T_B^C)^{-1}\mathrm{DM}_B(f)\right\}\right)\\
&= \operatorname{Spur}\left(E_n\mathrm{DM}_B(f)\right)\\
&= \operatorname{Spur}(\mathrm{DM}_B(f))
\end{aligned}$$

⋆ ⋆ ⋆

Aufgaben zu Kapitel 11. **Aufgabe 18.14:** Sei f ein Endomorphismus $f : V \to V$, V ein 2-dimensionaler $\mathbb{R}$-Vektorraum mit $\mathrm{DM}_B(f) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ für B eine beliebige Basis von V. Angenommen f ist nicht diagonalisierbar, dann gilt: $p_f(f)v = 0,\ \forall v \in V$.

<u>Hinweis</u>: Man beweise: Es existiert v, so dass $\{v, f(v)\} =: \{b_1, b_2\}$ linear unabhängig sind, also auch eine Basis von V. Man erhält, dass $f(b_1) = b_2$ und $f(b_2) = c_1b_1 + c_2b_2$ und für die darstellende Matrix gilt:

$$\mathrm{DM}_{b_1,b_2}(f) = \begin{pmatrix} 0 & c_1 \\ 1 & c_2 \end{pmatrix}$$

Als Nächstes müssen Sie zeigen, dass $p_f(f)$ unabhängig von der Basiswahl ist. Das bedeutet:

$$A := (a)_{i,j=1,2} = (T_B^C)^{-1}\begin{pmatrix} a & b \\ c & d \end{pmatrix}(T_B^C) \quad \text{für } B,C \text{ Basen von V}$$

dann ist:

$$p_f(\lambda) = (a-\lambda)(d-\lambda) - bc \quad = \quad (a_{11}-\lambda)(a_{22}-\lambda) - a_{12}a_{21}$$

Jetzt zeigt man nur noch $(0-f)(c_2-f) - c_1 = 0$ und erhält wegen der Basisinvarianz von $p_f(\lambda)$, dass $p_f(f) = (0-f)(c_2-f) - c_1 = 0$

Lösung.

Schritt 1 Angenommen, es existiert kein v, so dass $\{v, f(v)\} =: \{b_1, b_2\}$ linear unabhängig sind, dann wäre $\lambda_v v = f(v)$ für jedes $v \in V$ und damit f diagonalisierbar mit den Eigenwerten: $\lambda_{e_1}, \lambda_{e_2}$. Dies ist aber ein Widerspruch zu den Voraussetzungen, also muss es ein v geben, so dass $\{v, f(v)\} := \{b_1, b_2\}$ eine Basis von V ist. Natürlich ist jetzt auch

$$\mathrm{DM}_{b_1,b_2}(f) = \begin{pmatrix} 0 & c_1 \\ 1 & c_2 \end{pmatrix}$$

Schritt 2 Die Tatsache, dass $p_f(\lambda)$ unabhängig von der Basiswahl ist können Sie jetzt nachrechnen oder wenn Sie das Kapitel zu den Determinanten bereits gelesen haben:

$$(a_{11}-\lambda)(a_{22}-\lambda)-a_{12}a_{21} = \det\left((T_B^C)^{-1}\begin{pmatrix} a & b \\ c & d \end{pmatrix}(T_B^C)-\lambda E_2\right)$$

Schritt 3 wegen der Basisinvarianz von p_f erhalten wir nun:

$$p_f(\lambda) = \det\bigl(\mathrm{DM}_{b_1,b_2}(f)-\lambda E_2\bigr) = \det\left(\begin{pmatrix} -\lambda & c_1 \\ 1 & c_2-\lambda \end{pmatrix}\right) = -\lambda(c_2-\lambda)-c_1$$

und angewendet auf die Basis b_1, b_2:

$$\begin{aligned} p_f(f)b_1 &= (-f(c_2-f)-c_1)b_1 = -c_2f(b_1)+f^2(b_1)-c_1b_1 = -c_2b_2+f(b_2)-c_1b_1 = 0 \\ p_f(f)b_2 &= -c_2f(b_2)+f^2(b_2)-c_1b_2 = -c_2f(b_2)+f(c_1b_1+c_2b_2)-c_1b_2 = 0 \end{aligned}$$

insgesamt gilt also für jeden Vektor v: $p_f(f)v = 0$

Aufgabe 18.15: Sei $f : V \to V$ ein Endomorphismus auf dem $\mathbb{K}$-Vektorraum V. Angenommen $f^4(v)+v = 0 \quad \forall v \in V$. Welche Aussagen bezüglich der Diagonalisierbarkeit und der Eigenwerte von f lassen sich für $\mathbb{K} = \mathbb{R}, \mathbb{C}, \mathbb{Z}_2$ feststellen?

Lösung.

Angenommen $\mathbb{K} = \mathbb{R}$

Für jeden Eigenvektor v zum Eigenwert λ müsste $0 = f^4(v)+v = (\lambda^4+1)v$. Da $v \neq 0$ muss also $(\lambda^4+1) = 0$ gelten, was im Körper $\mathbb{R}$ nicht sein kann!

Angenommen $\mathbb{K} = \mathbb{C}$

Nun bekommen wir mit derselben Rechnung wie oben die möglichen Eigenwerte:

$$\lambda_1 = \mathrm{e}^{\mathrm{i}\frac{\pi}{4}}, \quad \lambda_2 = \mathrm{e}^{\mathrm{i}\frac{3\pi}{4}}, \quad \lambda_3 = \mathrm{e}^{\mathrm{i}\frac{-\pi}{4}}, \quad \lambda_4 = \mathrm{e}^{\mathrm{i}\frac{-3\pi}{4}}$$

als Nst. der Gleichung: $\lambda^4+1 = 0$

Angenommen $\mathbb{K} = \mathbb{Z}_2$

Die Gleichung $\lambda^4+1 = 0$ hat über dem Körper $\mathbb{Z}_2$ nur die Lösung $\lambda = 1$. Das bedeutet, dass $f(v) = v$ sein muss, falls f diagonalisierbar ist.

Aufgabe 18.16: Sei f ein Endomorphismus $f : V \to V$, wobei V ein zweidimensionaler $\mathbb{C}$-Vektorraum ist und es soll $f^k(v) = 0$ für alle $v \in V$ und ein festes $k \in \mathbb{N}$ gelten. Man zeige, dass $\mathrm{Spur}(f) = 0$ gilt.

Lösung.

Angenommen es gibt einen Eigenvektor $v \in V$ mit $f(v) = \lambda v$, dann ist

$$0 = f^k(v) = \lambda^k v$$

und daraus folgt, dass $\lambda = 0$ gelten muss. Wir wissen nun, dass f über $\mathbb{C}$ diagonlisierbar ist (algebraisch abgeschlossener Körper) und es dementsprechend mindestens einen solchen Eigenvektor v zu $\lambda = 0$ geben muss. Dadurch erhalten wir für $\mathrm{DM}_B(f) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, dass die Gleichung:

$$\det\left(\begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix}\right) = \lambda^2 - \lambda \underbrace{(a+d)}_{=\mathrm{Spur}(f)} + ad - bc = 0$$

nur für $\lambda = 0$ erfüllt ist! Das bedeutet aber auch, dass die algebraischen Lösungen (durch Lösen der p-q-Formel für quadratische Gleichungen)

$$\lambda_{1,2} = \frac{\mathrm{Spur}(f)}{2} \pm \sqrt{\frac{\mathrm{Spur}(f)^2}{4} - (ad - bc)}$$

null sein müssen! Daraus folgt $\frac{\mathrm{Spur}(f)}{2} = 0$ und somit die gewünschte Aussage.

⋆ ⋆ ⋆

Aufgaben zu Kapitel 12. **Aufgabe 18.17:** Sei f ein Endomorphismus, $f : V \to V$ für den gilt, dass jeder Vektor $v \neq 0$ ein Eigenvektor ist. Zeigen Sie, dass dann $f = c \cdot id_v$ gilt. **Lösung.**
Sei $\{b_1, \ldots, b_n\}$ eine beliebige Basis von V. Jetzt ist für jedes $i = 1, \ldots, n : f(b_i) = \lambda_i b_i$. Aber es gilt auch $f(\sum_{k=1}^n b_k) = \hat{\lambda} \sum_{k=1}^n b_k$. Wenn wir jetzt die Linearität von f ausnutzen, erhalten wir:

$$\sum_{k=1}^n \lambda_i b_i = \sum_{k=1}^n f(b_i) = f\left(\sum_{k=1}^n b_k\right) = \hat{\lambda} \sum_{k=1}^n b_k = \sum_{k=1}^n \hat{\lambda} b_k$$

Da $\{b_1, \ldots, b_n\}$ einen Basis von V ist, muss aufgrund der linearen Unabhängigkeit der b_i aus dieser Gleichung $\lambda_1 = \cdots = \lambda_n = \hat{\lambda}$ folgen. Damit ist $f = \hat{\lambda} id_v$.

⋆ ⋆ ⋆

Aufgaben zu Kapitel 13. **Aufgabe 18.18:** Finden Sie einen Endomorphismus $f : \mathbb{R}^3 \to \mathbb{R}^3$, für den $p_A(\lambda) = -\lambda^3 + 2\lambda^2 + 3\lambda + 2$ gilt.
Kann man für jedes Polynom $p(\lambda) = \pm\lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_0$ einen Endomorphismus f finden, so dass $p_A(\lambda) = p(\lambda)$ gilt?
Lösung.
Sei:

$$\mathrm{DM}_B(f) := \begin{pmatrix} 0 & 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & 0 & \ldots & 0 & -a_2 \\ \vdots & & \ddots & \vdots & 0 & \vdots \\ 0 & 0 & 0 & \ldots & 1 & -a_{n-1} \end{pmatrix}$$

Dann rechnet man für das charakteristische Polynom (Entwickeln der n-ten Spalte, Laplacescher Entwicklungssatz):

$$p_f(\lambda) = \det \begin{pmatrix} -\lambda & 0 & 0 & \dots & 0 & -a_0 \\ 1 & -\lambda & 0 & \dots & 0 & -a_1 \\ 0 & 1 & -\lambda & \dots & 0 & -a_2 \\ \vdots & & \ddots & \vdots & -\lambda & \vdots \\ 0 & 0 & 0 & \dots & 1 & -a_{n-1} - \lambda \end{pmatrix} = (-1)^n \left(\lambda^n + \sum_{k=0}^{n-1} a_k \lambda^k \right)$$

Man kann also zu jedem Polynom $p(\lambda) = (-1)^n \lambda^n + (-1)^n a_{n-1} \lambda^{n-1} + \ldots + (-1)^n a_0$ einen Endomorphismus f finden, so dass $p_f(\lambda) = p(\lambda)$.

Für $p_A(\lambda) = -\lambda^3 + 2\lambda^2 + 3\lambda + 2$ ist also, der durch

$$\mathrm{DM}_B(f) = \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & 3 \\ 0 & 1 & 2 \end{pmatrix}$$

gegebene Endomorphismus f die gesuchte Antwort.

⋆ ⋆ ⋆

Aufgaben zu Kapitel 14. **Aufgabe 18.19:** Sei V ein n-dimensionaler euklidischer Vektorraum, $f : V \to V$ ein Endomorphismus, wobei aus $\langle x, y \rangle = 0$ die Gleichung $\langle f(x), f(y) \rangle = 0$ folgt. Man zeige, dass es ein $\lambda \in \mathbb{R}$ gibt, so dass λf eine Isometrie ist.

Lösung.

Sei $\{b_1, \ldots, b_n\}$ eine Orthonormalbasis von V.

Jetzt ist $\langle b_i + b_m, b_i - b_m \rangle = \langle b_i, b_i \rangle - \langle b_i, b_m \rangle + \langle b_m, b_i \rangle - \langle b_m, b_m \rangle = 1 + 0 - 0 - 1 = 0$. Daraus folgt:

$$0 = \langle f(b_i + b_m), f(b_i - b_m) \rangle = \langle f(b_i), f(b_i) \rangle - \langle f(b_m), f(b_m) \rangle$$

und wir können die Konstante $C := \sqrt{\langle f(b_1), f(b_1) \rangle} = \ldots = \sqrt{\langle f(b_n), f(b_n) \rangle}$ eindeutig definieren. Damit gilt für die Funktion $\frac{1}{C} f$:

$$\begin{aligned} \left\langle \frac{1}{C} f(x), \frac{1}{C} f(y) \right\rangle &= \left\langle \frac{1}{C} f \left(\sum\nolimits_{k=1}^{n} \lambda_k b_k \right), \frac{1}{C} f \left(\sum\nolimits_{l=1}^{n} \mu_l b_l \right) \right\rangle \\ &= \left\langle \sum_{k=1}^{n} \lambda_k \frac{1}{C} f(b_k), \sum_{l=1}^{n} \mu_l \frac{1}{C} f(b_l) \right\rangle \\ &= \sum_{k=1}^{n} \lambda_k \mu_k \underbrace{\frac{1}{C^2} \langle f(b_k), f(b_k) \rangle}_{=1} \quad = \quad \langle x, y \rangle \end{aligned}$$

Für die Vektoren $x = \sum^n{}_{k=1} \lambda_k b_k$ und $y = \sum^n{}_{l=1} \mu_l b_l$.

Somit ist $\frac{1}{C} f$ die gesuchte Isometrie.

Abbildungen

1 Analytische Geometrie im $\mathbb{R}^2$ . . . 17
2 grafische Definition der komplexen Zahlen . . . 35
3 Spezieller Fall der Subtraktion . . . 35
4 Polarkoordinatendarstellung komplexer Zahlen . . . 36
5 Multiplikation komplexer Zahlen in Polarkoordianten . . . 37
6 $1 + \mathrm{i}$ und die drei 3-ten Wuzeln in Polarkoordinaten . . . 38
7 $\sqrt[16]{1{,}2 \cdot \mathrm{e}^{\frac{29}{16}\mathrm{i}\pi}} = \sqrt[16]{0{,}9977 - 0{,}6666\mathrm{i}}$ . . . 40
8 $\mathcal{L}(v_1, v_2)$-Ebene . . . 42
9 `[Bild]` – Bild folgt . . . 43
10 Bild und Kern einer Funktion f . . . 58
11 Matrizen als Drehungen in der reellen Ebene . . . 82
12 . . . 83
13 . . . 85
14 . . . 85
15 . . . 86
16 Unterraum U und affiner (verschobener) Unterraum $z + U$ – beachte: $z' + U = z + U$ . . . 92
17 die spezielle Lösung verschiebt die homogene Lösung . . . 93
18 Darstellung des Vektors $\begin{pmatrix} 2 & 1 \end{pmatrix}$ durch verschiedene Basen . . . 101
18a kanonische Basis K . . . 101
18b Basis $C = \left\{ \begin{pmatrix} 2 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \end{pmatrix} \right\}$ . . . 101
18c Basis $D = \left\{ \begin{pmatrix} 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \end{pmatrix} \right\}$ . . . 101
19 Parallelogramm . . . 132
20 . . . 144
21 . . . 145
22 gedrehter Vektor . . . 148
23 Dreiecksungleichung . . . 152
24 orthogonales Komplement . . . 154
25 orthogonale Projektion . . . 161
26 . . . 167
27 Durch Abstieg entlang der iterierten Kerne die wertvollen Jordanketten bergen . . . 175

Definitionsverzeichnis

1.3 lineares Gleichungssystem . . . 11
2.1 mathematischer Körper . . . 16
2.6 Vektorraum . . . 19
2.9 Schreibweise des inversen Elements der Addition . . . 20
2.10 Unter(vektor)raum . . . 20
2.15 Abbildung . . . 21
3.1 Verknüpfung . . . 24
3.3 neutrales Element . . . 24
3.5 inverses Element . . . 24
3.8 abelsche Gruppe . . . 25
3.11 Ring . . . 26
3.13 Körper . . . 27
3.17 Teiler . . . 29
3.19 Primzahl . . . 29
3.21 modulo . . . 30
3.22 Addition modulo und Multiplikation modulo . . . 30
5.1 Linearkombination . . . 42
5.3 endlich erzeugend . . . 43
5.4 lineare Abhängigkeit . . . 43
5.7 Basis . . . 44
5.9 kanonische Basis . . . 45
5.15 Dimension . . . 48
6.1 lineare Abbildung/Homomorphismus . . . 56
6.6 Kern und Bild . . . 58
6.8 Rang und Defekt . . . 59
6.11 $\mathrm{Hom}(V, W)$ . . . 61
6.13 Morphismen . . . 62
7.2 Matrix . . . 69
7.6 Matrizenaddition, Matrizenmultiplikation, Nullmatrix . . . 71
7.9 Matrizenprodukt . . . 73
7.17 quadratische Matrix, Einheitsmatrix . . . 75
7.19 invertierbar (Matrix) . . . 76
8.9 elementare Zeilenumformung . . . 90
9.4 obere Dreiecksmatrix, Diagonalmatrix, Zeilenstufenmatrix . . . 95
10.1 Transformationsmatrix . . . 99
10.8 darstellende Matrix . . . 102
10.16 Darstellende Matrix eines Endomorphismus . . . 106
10.17 transponierte Matrix, symmetrische Matrix . . . 106
11.1 diagonalisierbar . . . 111
11.2 Eigenwert, Eigenvektor . . . 111
11.9 Eigenraum, geometrische Vielfachheit . . . 114
12.2 Untermatrix zur Bestimmung der Determinante . . . 121
12.3 Laplacescher Entwicklungssatz/Determinante(nfunktion) . . . 121
12.16 aufgespanntes Parallelogramm von Vektoren . . . 132
13.1 charakteristische Polynom . . . 134
14.1 Standard-Skalarprodukt . . . 147

14.2 Skalarprodukt 149
14.4 euklidischer Vektorraum 150
14.6 Norm 150
14.9 Öffnungswinkel 153
14.10 Orthogonalität 153
14.12 orthogonales Komplement 153
14.14 Orthonormalsystem 155
14.23 orthogonale Projektion 161
14.25 isometrische/orthogonale Abbildung 162
14.29 orthogonale Gruppe 163
14.31 orthogonale Matrix 164
15.1 selbstadjungiert 166
16.1 Definitheit von Matrizen 169
16.3 Hauptminoren 170
17.6 Jordankette, Hauptvektor 176

Stichwortverzeichnis

— **Symbole** —
i (imaginäre Einheit) . . . 32

— **A** —
Abbildung . . . 24
Abbruchkriterium . . . 137
abelsche Gruppe . . . 25
Abstand . . . 147
algebraische Vielfachheit . . . 137
Assoziativgesetz . . . 17, 19, 75
Automorphismengruppe . . . 62
Automorphismus . . . 62, 163

— **B** —
Basis . . . 42, 137
Basisaustauschsatz . . . 47
Basisdarstellung . . . 44, 99, 155
Basisergänzungssatz . . . 46
Basisexistenzsatz . . . 46
Basiswechsel . . . 99
Betrag . . . 35
Bild . . . 60, 64, 77, 105, 120, 135

— **C** —
CAESAR-Verschlüsselung . . . 16
Cauchy-Schwarzsche Ungleichung 151 ff.
Caylay-Zahlen . . . 41
charakteristisches Polynom 120, 134, 137
Cramersche Regel . . . 131

— **D** —
darstellende Matrix . . . 69, 102, 104, 106, 135, 137
Defekt . . . 59
Definitheit . . . 169
Determinante . . . 121, 123, 128, 131
Determinante – Matrix . . . 120
diagonalisierbar . . . 111
Diagonalisierung . . . 117
Diagonalmatrix . . . 95
Dimension . . . 48, 59
Dimensionsformel . . . 105, 120
Dimensionssatz . . . 64
direkte Summe . . . 160
Diskriminante . . . 34
Distributivgesetz . . . 17, 19, 75
Divisionsalgebra . . . 41
Drehmatrix . . . 83
Drehung . . . 81, 108, 112, 153, 164
Dreiecksmatrix . . . 95
Dreiecksungleichung . . . 151 f.

— **E** —
Eigenraum . . . 114, 117, 137
Eigenvektor . . . 111, 114, 135, 137
Eigenwert . . . 111, 117, 135, 137
eindeutig lösbar . . . 11
Einheit . . . 24
Einheitengruppe . . . 26
Einheitsmatrix . . . 75, 97
Einselement . . . 18, 24
elementare Zeilenumformung 87, 90, 96, 127
Elementarmatrix . . . 90
endlich-dimensional . . . 48, 63, 111
endlich-erzeugt . . . 43
Endomorphismus . . . 62, 66, 106, 111, 135
Entwicklungsformel . . . 155
Epimorphismus . . . 62
erweiterte Koeffizientenmatrix . . . 91
euklidische Vektorräume . . . 147
euklidischer Vektorraum . . . 150
Existenzsätze für Basen . . . 45

— **F** —
Flächeninhalt . . . 131, 147, 151
Fundamentalsatz der Algebra . . . 134
Funktion . . . 18, 21

— **G** —
ganze Zahl . . . 29
Gaußscher Algorithmus . . . 13, 123
gaußscher Algorithmus . . . 96
geometrische Vielfachheit . 114, 116, 137
Gleichung . . . 9, 13
Gram-Schmidtsches Orthonormalisierungsverfahren . . . 156
Gruppe . . . 24 f., 164
Gruppe
  abelsche . . . 25
  Definition . . . 25

— **H** —
Hamilton-Zahlen . . . 41
Hauptachsentransformation 144, 168
Hauptminor . . . 170
Hauptvektor . . . 176
Hessematrix . . . 170
homogen . . . 11, 91
Homomorphismus . . . 56, 61 f.

— **I** —
imaginäre Einheit . . . 32
Imaginärteil . . . 32
inhomogen . . . 91
injektiv . . . 66
Inverse 16 f., 19, 24, 76, 96, 120, 128, 135
inverses Element . . . 24
invertierbar . . . 24
Isometrie . . . 167
isometrisch . . . 162
Isomorphiesatz . . . 72
Isomorphismus . . . 62

— **J** —
Jordanblock . . . 172
Jordanform . . . 172
Jordankette . . . 176
Jordankästchen . . . 172
jordansche Normalform . . . 172

— **K** —
kanonische Basis . . . 45, 100
Kern 58, 60, 64, 66, 77, 91, 105, 117, 120, 135
Koeffizientenmatrix . . . 91, 96
Kommutativgesetz . . . 17, 19
komplexe Addition . . . 32
komplexe Division . . . 33
komplexe Multiplikation . . . 32
komplexe Subtraktion . . . 35
komplexe Zahlen . . . 10, 20, 32 f., 49, 134
komplexe Zahlen
  Addition . . . 32
  Definition . . . 32
  Division . . . 33
  Gleichheit . . . 32
  Imaginärteil . . . 32
  Multiplikation . . . 32
  Realteil . . . 32
Konjugierte . . . 33
Koordinatentransformation . . . 99
Kosinussatz . . . 36
Körper . . . 9, 16, 24, 30
Körper
  Definition . . . 27

— **L** —
Laplacescher Entwicklungssatz 121 f.
linear unabhängig 43, 114, 116, 137, 155
lineare Abbildung 56, 64, 66, 74, 77, 102, 161
lineare Codes . . . 80
lineare Hülle . . . 42
lineares Gleichungssystem . 9, 11, 56, 91, 96, 113, 135
Linearfaktor . . . 137
Linearkombination . . . 42
Länge . . . 148, 150, 152
lösbar . . . 11, 91
Lösungsmenge . . . 11, 91

— **M** —
Matrix . . . 14, 69, 77
Matrix
  Definitheit . . . 169
  Definitheit
    negativ definit . . . 169
    negativ semidefinit . . . 169
    positiv definit . . . 169
    positiv semidefinit . . . 169
Matrizen
  Anwendungen . . . 80
  Charakteristisches Polynom 134
  Darstellende Matrix . . . 68
  Darstellungsmatrizen
    joransche Normalform . . 172
  Determinante . . . 120
  Diagonalisierbarkeit . . . 134
  Eigenvektoren . . . 111
  Eigenwerte . . . 111
  linear Gleichungssysteme . . . 91
Matrizen, Matrix . . . 68
Matrizenaddition . . . 71
Matrizenprodukt . . . 73
Minor . . . 170
modulo . . . 29

Addition . . . 30
Multiplikation . . . 30
Monomorphismus . . . 62

— N —
negativ definit . . . 169
neutrales Element . . . 24
Neutralität . . . 17, 19
Norm . . . 150
Normalform
Jordansche . . . 172
Nullmatrix . . . 71
Nullstelle . . . 134
Nullvektor . . . 16, 18, 91

— O —
Oeffnungswinkel . . . 153
Oktonionen . . . 41
orthogonal . . . 153, 162, 164
orthogonale Abbildung . . . 162 f.
orthogonale Gruppe . . . 163
orthogonale Projektion . . . 161
orthogonale Zerlegung . . . 160
orthogonales Komplement . 153, 160
Orthogonalität . . . 147, 149, 153
Orthonormalbasis . . . 166
Orthonormalisierungsverfahren
Gram-Schmidt . . . 156
Orthonormalsystem . . . 155
Ortsvektor . . . 35 f.

— P —
p-q-Formel . . . 34
Parallelogramm . . . 132
Pivot-Element . . . 12
Polarkoordinaten . . . 35 f.
positiv definit . . . 169
Primzahl . . . 29 f.
Primzahl
Definition . . . 29
Projektion . . . 160

— Q —
quadratische Matrix . . . 75, 95
Quaternionen . . . 41

— R —
Rang . . . 59, 135
rationale Zahlen . . . 20, 49
Realteil . . . 32
reelle Ebene . . . 34, 81, 85, 147
reelle Zahlen . . . 20, 33, 49
Regel von Sarrus . . . 122
Rest . . . 29
Restklasse . . . 13, 16
Ring . . . 9, 24, 26, 30
Ring
additive Inverse . . . 26
Definition . . . 26
Einheit . . . 26
invertierbar . . . 26
kommutativ . . . 26
kommutativ (mit Eins)
Bedingungen . . . 28
mit Eins . . . 26
neutrales Element
Addition . . . 26
Multiplikation . . . 26

— S —
Schachbrettmuster . . . 122
Schiefkörper . . . 41
selbstadjungiert . . . 166
senkrecht . . . 153
skalare Multiplikation . . . 71
Skalarmultiplikation . . . 16
Skalarprodukt . . . 147, 149
Skalarprodukt
Linearität . . . 149
Linearität
erstes Argument . . . 149
zweites Argument . . . 149
positive Definitheit . . . 149
Symmetrie . . . 149
Spaltenvektor . . . 69
Spektralsatz . . . 168
Spiegelung . 85, 107 f., 112, 120, 164, 166
Standard-Skalarprodukt . . . 147
surjektiv . . . 66
symmetrisch . . . 166
symmetrische Matrix . . . 106

— T —
Teilbarkeit . . . 29
Transformationsmatrix . . . 104, 136 f.
Transformationssatz . . . 99
transponierte Matrix . . . 106

— U —
Unbekannte . . . 9
unendlich-dimensional . . . 48
Untervektorraum . . . 20, 48, 91

— V —
Vektoraddition . . . 16
Vektorraum . . . 16, 19, 61, 63, 71
Vektorsubtraktion . . . 16
Verknüpfung . . . 24 f.
Volumen . . . 131

— W —
Winkel . . . 147 f.

— Z —
Zahl
ganze . . . 29
Prim- . . . 29 f.
Prim-
Definition . . . 29
Zeilenstufenform . . . 12, 127
Zeilenstufenmatrix . . . 95
Zeilenvektor . . . 69
Zerlegung . . . 160