

Lineare Algebra

AOR Dr. Thoralf Räsch
Mathematisches Institut
der Universität Bonn

Version 0.99

19. März 2010

* * *

Dieses vorläufige 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.

* * *

Inhaltsverzeichnis

1. Lineare Gleichungssysteme	5
Gleichungen	5
Gleichungssysteme	6
Lineare Gleichungssysteme spezieller Gestalt	7
Der Gauß-Algorithmus	8
2. Vektorräume	10
Der Raum \mathbb{K}^n	10
Untervektorräume	14
3. Komplexe Zahlen	17
Definition des Zahlenbereichs und der Operationen darauf	17
Komplexe quadratische Gleichungen	19
Komplexe Zahlen als reelle Ebene	20
Komplexe Zahlen als Polarkoordinaten	21
Anwendungen	22
4. Linearkombinationen und Basen	22
Linearkombinationen und lineare Unabhängigkeit	23
Existenzsätze für Basen	26
5. Lineare Abbildungen	31
Grundlagen Linearer Abbildungen	31
Endliche Beschreibung von Homomorphismen	36
Klassifikation endlich-dimensionaler Vektorräume	38
Der Dimensionssatz	39
6. Die Welt der Matrizen	40
Darstellende Matrizen	40
Matrizenrechnung	42
Lineare Abbildungen und Matrizen	48
7. Anwendungen von Matrizen	49
Matrizen als Kodierwerkzeug: Lineare Codes	49
Matrizen als Drehungen in der reellen Ebene	51
Matrizen als Spiegelungen in der reellen Ebene	54
Matrizen als Elementare Umformung: Vertauschen von zwei Zeilen	56
Matrizen als Elementare Umformung: Skalarmultiplikation einer Zeile	57
Matrizen als Elementare Umformung: Addition zweier Zeilen	58
8. Matrizen und Lineare Gleichungssysteme	59
Geometrie der Lösungsmengen	60
Gaußscher Algorithmus	63

Algorithmus zum Invertieren von Matrizen	64
9. Koordinatentransformation	65
Basiswechsel	65
Darstellende Matrizen von linearen Abbildungen bezüglich beliebiger Basen	68
Darstellende Matrizen von Endomorphismen	72
10. Eigenwerte und Eigenvektoren	74
Grundlegende Begriffe	75
Eigenwerte und Eigenvektoren an bekannten Beispielen	76
Eigenräume zu Eigenwerten	78
Lineare Unabhängigkeit von Eigenvektoren	79
Vorläufige Strategie des Diagonalisierens	81
11. Determinanten und charakteristische Polynome	82
Determinanten von Matrizen	83
Determinanten und Gaußscher Algorithmus	86
Die Cramersche Regel	92
Charakteristische Polynome und Eigenwerte	93
Diagonalisieren an Beispielen	96
Ausblick Hauptachsentransformation	100
Determinanten und Volumina	101
12. Euklidische Vektorräume	103
Geometrische Begriffe in der reellen Ebene	104
Skalarprodukte	105
Normen	107
Orthogonalität	109
Orthonormalsysteme	111
Orthogonale Zerlegungen und Projektionen	114
Orthogonale Abbildungen	116
Index	119

1. LINEARE GLEICHUNGSSYSTEME

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

Gleichungen. Wir arbeiten über einem Ring oder Körper. Eine Gleichung ist eine Formel $t(x_1, \dots, x_n) = b$, wobei t ein arithmetischer Term mit Unbekannten x_1, \dots, x_n ist.

Beispiel 1.1. Es gibt verschiedene Arten von Gleichungen:

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

Eine Gleichung ist eine Bedingung an die Unbekannten der Gleichung. Gesucht ist jeweils die Lösungsmenge $\{(x_1, \dots, x_n) \mid t(x_1, \dots, x_n) = b\}$. 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

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

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 eindeutig:

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

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

Dann ist die Gleichung $a \cdot x = b$ immer falsch. 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 :

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

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: $x^2 + p \cdot x + (\frac{p}{2})^2 = -q + (\frac{p}{2})^2$ und nach binomischer Formel gilt: $(x + \frac{p}{2})^2 = -q + (\frac{p}{2})^2$

Fall 1: $-q + (\frac{p}{2})^2 \geq 0$

Dann kann man die Quadratwurzel ziehen und erhält

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

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

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

Allgemein erhalten wir als Lösungsmenge:

$$\{x \in \mathbb{R} \mid -q + (\frac{p}{2})^2 \geq 0 \text{ und } (x = -\frac{p}{2} + \sqrt{-q + (\frac{p}{2})^2} \\ \text{oder } x = -\frac{p}{2} - \sqrt{-q + (\frac{p}{2})^2})\}$$

Bemerkung 1.2. Über dem Grundkörper \mathbb{C} gibt es auch für den letzten Fall Lösungen. Dies untersuchen wir in Kapitel 3.

* * *

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. 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:

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

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

Die Lösungsmenge des linearen Gleichungssystems $Ax = b$ ist

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

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

* * *

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

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

Dieses lineare Gleichungssystem können wir schnell lösen:

Fall 1: Für alle $i \leq n$ gilt $a_{ii} \neq 0$.

Dann ist das lineares Gleichungssystem eindeutig lösbar mit der Lösungsmenge: $\text{Lös}(A, b) = \{(\frac{b_1}{a_{11}}, \dots, \frac{b_n}{a_{nn}})\}$

Fall 2: Es gibt ein $i \leq n$ mit $a_{ii} = 0$ und $b_i \neq 0$.

Dann ist das gesamte lineares Gleichungssystem nicht lösbar, da die i -te Gleichung nicht lösbar ist. Es gilt also:

$$\text{Lös}(A, b) = \emptyset$$

Fall 3: Für alle $i \leq n$ mit $a_{ii} = 0$ gilt $b_i = 0$.

Dann ist das lineares Gleichungssystem lösbar mit :

$$\text{Lös}(A, b) = \{(x_1, \dots, x_n) \in \mathbb{K}^n \mid \text{für } i \leq n \text{ mit } a_{ii} \neq 0 \text{ gilt } x_i = \frac{b_i}{a_{ii}}\}$$

Betrachten wir nun ein lineares Gleichungssystem $Ax = b$ über einem Körper \mathbb{K} in Zeilenstufenform (ZSF):

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

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

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

Menge der Lösungen (x_1, \dots, 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 lineares Gleichungssystem sieht wie folgt aus:

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

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

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

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

$$\begin{aligned} \text{Lös}(A, b) &= \{(x_1, \dots, x_m) \in \mathbb{K}^m \mid \text{für } k = 1, \dots, i \text{ gilt} \\ & x_{j(k)} = \frac{1}{a_{kj(k)}} \cdot (b_k - \sum_{l=j(k)+1}^m a_{kl}x_l)\} \end{aligned}$$

Beispiel 1.4. 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:

$$\begin{aligned} \mathbb{L} &= \{(x_1, \dots, x_4) \in \mathbb{R}^4 \mid x_3 = \frac{1-3x_4}{2} \text{ und } x_1 = 2 - 2x_2 - 3x_3 - 4x_4\} \\ &= \{(x_1, \dots, x_4) \in \mathbb{R}^4 \mid x_3 = \frac{1-3x_4}{2} \text{ und } x_1 = \frac{1}{2} - 2x_2 + \frac{1}{2}x_4\} \end{aligned}$$

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 :

$$\mathbb{L} = \{(x_1, \dots, x_4) \mid x_3 = 3 + x_4 \text{ und } x_1 = 3 + 3x_2 + 3x_4\}$$

* * *

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

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

$$a_1x_1 + \dots + a_mx_m = b \quad (*)$$

äquivalent zu der Gleichung

$$\lambda a_1x_1 + \dots + \lambda a_mx_m = \lambda b \quad (**)$$

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

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

äquivalent zu:

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

Beweis:

zu (a) Seien $x_1, \dots, x_m \in \mathbb{K}$. Wir zeigen die Äquivalenz $(*) \Leftrightarrow (**)$.

Zunächst zeigen wir „ \Rightarrow “:

Es gelte also $(*)$. Dann gilt insbesondere:

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

Durch wiederholtes Anwenden des Distributivgesetzes in \mathbb{K} ergibt sich die Gleichung $(**)$.

Es bleibt „ \Leftarrow “ zu zeigen:

Es gelte nun $(**)$. Nach Anwendung des Distributivgesetzes gilt:

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

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

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

und daraus folgt $(*)$.

Die Aussage aus (b) kann analog gezeigt werden. \(\square\)

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 1.6.

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

Also $y = 2$ und $x = 1$.

Wir können das System auch kürzer fassen:

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

Beispiel 1.7. 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:

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

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

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

2. VEKTORRÄUME

Lineare Gleichungssysteme in den Unbekannten x_1, \dots, 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 . Wir starten gleich mit der entscheidenden

Definition 2.1. 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) $\mathbb{K}^n = \{(x_1, \dots, x_n) \mid x_1 \in \mathbb{K}, \dots, x_n \in \mathbb{K}\}$. Die Elemente von \mathbb{K}^n heißen **Vektoren**.

(b) Für $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbb{K}^n$ definiere die **Vektoraddition** “+” durch

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

(c) Für $(x_1, \dots, x_n) \in \mathbb{K}^n$ und $\lambda \in \mathbb{K}$ definiere die **Skalarmultiplikation** “ \cdot ” durch

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

(d) $0 = (0, \dots, 0) \in \mathbb{K}^n$ ist der **Nullvektor**.

(e) Für $x = (x_1, \dots, x_n) \in \mathbb{K}^n$ definiere ein **additiv Inverses** durch

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

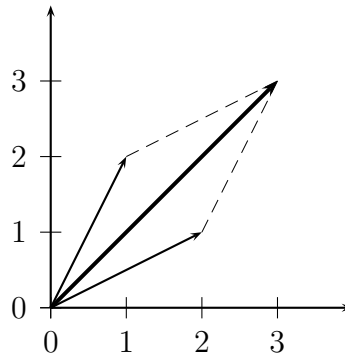
- (f) Für (x_1, \dots, x_n) und (y_1, \dots, y_n) definiere die Vektorsubtraktion “ $-$ ” durch $(x_1, \dots, x_n) - (y_1, \dots, y_n) := (x_1 - y_1, \dots, x_n - y_n)$

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 $\mathbb{Z}_m = \{0, 1, \dots, m-1\} \hat{=} \{A, B, C, \dots\}$ auffassen. So kann man dies benutzen etwa für die CAESAR- Verschlüsselung: $(x_1, \dots, x_n) \mapsto (x_1, \dots, x_n) + (c, \dots, c)$. Etwa

INFORMATIK + EEEEEEEEEE = NSKTWRFYNP

- (c) Analytische Geometrie im \mathbb{R}^2 : $\begin{pmatrix} 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$



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

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

□

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 $M = \{f \mid f: [-1, 1] \rightarrow \mathbb{R}\}$ 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 $M = \{f \mid f: [-1, 1] \rightarrow \mathbb{R}\}$ 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
 $(\lambda + \mu) \cdot f = \lambda f + \mu f$ (2. Distributivgesetz)

Beweis: Zur Kommutativität:

$$(f + g) = (g + f) \Leftrightarrow \forall x((f + g)(x) = (g + f)(x))$$

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

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

Rest klar, mit den Rechengesetzen im Korper \mathbb{R} . □

Definition 2.6. *Es sei \mathbb{K} ein Korper und sei $V = (V, +, \cdot)$ eine Struktur mit einer Abbildung “+” (genannt Addition)*

$$+: V \times V \rightarrow V \quad ; \quad (x, y) \mapsto x + y$$

und einer Abbildung “ \cdot ” (genannt “skalare Multiplikation” oder “Skalarmultiplikation”)

$$\cdot: \mathbb{K} \times V \rightarrow V \quad ; \quad (\lambda, x) \mapsto \lambda \cdot x$$

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

- (a) *Fur $x, y, z \in V$ gilt*
 $(x + y) + z = x + (y + z)$ *(Assoziativgesetz)*
- (b) *Fur $x, y \in V$ gilt*
 $x + y = y + x$ *(Kommutativgesetz)*
- (c) *Es gibt ein “0” $\in V$, sodass fur $x \in V$ gilt*
 $x + 0 = x$ *(Neutralitat der Null)*
- (d) *Fur $x \in V$ gibt es ein $y \in V$, sodass*
 $x + y = 0$ *(Existenz des Inversen)*
- (e) *Fur $\lambda, \mu \in \mathbb{K}$ und $x \in V$ gilt*
 $(\lambda \cdot \mu) \cdot x = \lambda \cdot (\mu \cdot x)$ *(Assoziativgesetz)*
- (f) *Fur $x \in V$ gilt*
 $1 \cdot x = x$ *(Neutralitat der 1)*
- (g) *Fur $\lambda \in \mathbb{K}$ und $x, y \in V$ gilt*
 $\lambda \cdot (x + y) = \lambda x + \lambda y$ *(1. Distributivgesetz)*
- (h) *Fur $\lambda, \mu \in \mathbb{K}$ und $x \in V$ gilt*
 $(\lambda + \mu)x = \lambda x + \mu x$ *(2. Distributivgesetz)*

Satz 2.7. *Fur alle Vektorraume $V = (V, +, \cdot)$ gibt es genau ein Element, das der Nullvektor ist.*

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

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

☒

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 additiv 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 additiv Inverse von x in V , d.h. $x + a = 0$ und $x + b = 0$.

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

☒

Definition 2.9. *Schreibe $x - y$ statt $x + (-y)$.*

* * *

Untervektorräume. Sie wissen bereits, dass etwa 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 beispielweise 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. *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:

(1) $0 \in U$. (eigentlich $0_V \in U$)

Wähle $x \in U$ als Zeuge 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: $0 = 0 \cdot x \in U$

(2) 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 additiv 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 noch 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. den 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. Sei $\text{Abb}(\mathbb{R}, \mathbb{R})$ die Menge aller Funktionen von \mathbb{R} nach \mathbb{R} , das heißt:

$$\text{Abb}(\mathbb{R}, \mathbb{R}) = \{f \mid f: \mathbb{R} \rightarrow \mathbb{R}\}$$

Wir wissen bereits, dass wir $\text{Abb}(\mathbb{R}, \mathbb{R})$ 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 $g(x) = g(-x)$.

$$\begin{aligned} (f + g)(x) &= f(x) + g(x) && \text{per Definition} \\ &= f(-x) + g(-x) && f, g \in U \\ &= (f + g)(-x) \end{aligned}$$

(c) Abgeschlossenheit bzgl. “·”:

Sei $f \in U$, also $f(x) = f(-x)$ und $\lambda \in \mathbb{R}$.

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

3. 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:

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

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

Definition des Zahlenbereichs und der Operationen darauf. 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 $b \cdot i$ oder kurz bi bzw. ib . Weiterhin können wir gemischte Summen bilden: Die Summe aus einer reellen Zahl a und einer rein imaginären Zahl $b \cdot i$ heißt dann *komplexe Zahl*. Die Menge der komplexen Zahlen wird mit \mathbb{C} bezeichnet:

$$\mathbb{C} := \{a + i \cdot b \mid a, b \in \mathbb{R}\}.$$

Wir vereinbaren, dass zwei komplexe Zahlen genau dann gleich sind, wenn sie sowohl im Realteil, als auch im Imaginärteil übereinstimmen. Hierbei bezeichnet für $z := a + i \cdot b \in \mathbb{C}$ das a den *Realteil* von z und b den *Imaginärteil*; kurz $a = \operatorname{Re}(z)$ und $b = \operatorname{Im}(z)$. Insbesondere gilt, dass für $0 = b = \operatorname{Im}(z)$ 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 $z_1 := a + i \cdot b$ und $z_2 := c + i \cdot d$ setze:

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

Damit ist $+: \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$ offensichtlich eine Fortsetzung der reellen Addition, denn für $b = d = 0$ sind $z_1, z_2 \in \mathbb{R}$, $\operatorname{Im}(z_1 + z_2) = 0$ 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 $z_1 := a + i \cdot b$ und $z_2 := c + i \cdot d$ gegeben durch

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

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

Satz 3.1. *Die Struktur $(\mathbb{C}, +, \cdot, 0, \mathbb{1})$ ist ein Körper, wobei $0 := 0 + i \cdot 0$ und $\mathbb{1} := 1 + i \cdot 0 = 1$. 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 $\mathbb{1} = 1$.

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

$$\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2 \quad \text{und} \quad \overline{z_1 \cdot z_2} = \bar{z}_1 \cdot \bar{z}_2.$$

Wir betrachten nun die Division zweier komplexer Zahlen.

Betrachte für $z = a + i \cdot b$ 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:

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

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

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

$$\frac{(3-2i)(4-5i)}{(4+5i)(4-5i)} = \frac{12-15i-8i+10i^2}{16-25i^2} = \frac{12-10-23i}{16+25}$$

$$= \frac{2 - 23i}{41} = \frac{2}{41} - i \cdot \frac{23}{41}$$

Insbesondere gilt für den Real- und Imaginärteil dieser komplexen Zahl: $\operatorname{Re}\left(\frac{3-2i}{4+5i}\right) = \frac{2}{41}$ und $\operatorname{Im}\left(\frac{3-2i}{4+5i}\right) = -\frac{23}{41}$. Beachten Sie, dass der Imaginärteil einer komplexen Zahl immer eine reelle Zahl ist; es gilt für $z = a + ib$ stets: $\operatorname{Re}(z) = a$ und $\operatorname{Im}(z) = b$, 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 $\sqrt{-4} = \sqrt{(-1) \cdot 4} = \sqrt{-1} \cdot \sqrt{4} = \pm 2i$. Dieses Argument zeigt auch, dass für $z = \sqrt{a}$ mit $a < 0$ stets gilt: $z = \pm i \cdot \sqrt{-a}$.

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:

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

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 3.2 (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 $z = a + ib$ mit einem geordneten Paar reeller Zahlen $(a, b) \in \mathbb{R} \times \mathbb{R}$ identifizieren, das heißt, wir können folgende Abbildung angeben:

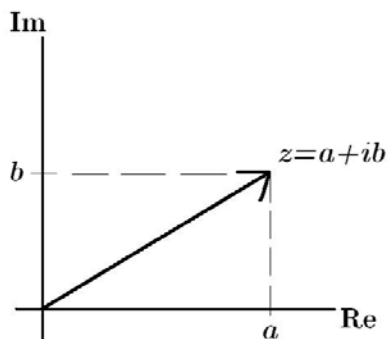
$$\mathbb{C} \rightarrow \mathbb{R} \times \mathbb{R}, \quad a + ib \mapsto (a, b)$$

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

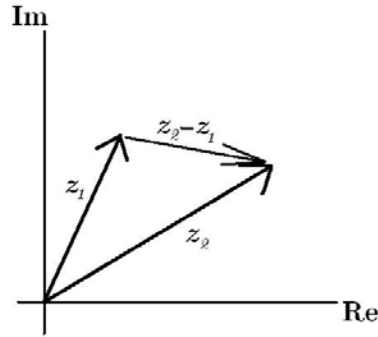
$$\begin{aligned} (a, b) + (c, d) &:= (a + c, b + d) \\ (a, b) \cdot (c, d) &:= (ac - bd, ad + bc) \end{aligned}$$

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

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 einzuzichnen – die so genannte Gaußsche Zahlenebene:



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: $z_2 - z_1 = z_2 + (-z_1)$).

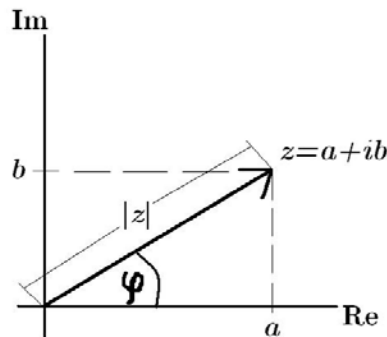


Komplexe Zahlen als Polarkoordinaten. Wenn wir uns die Multiplikation geometrisch überlegen wollen, dann ist eine andere Sichtweise der komplexen Zahlen besser geeignet: die *Darstellung der komplexen Zahlen durch Polarkoordinaten*.

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

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

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 Hypotenuse $|z|$ 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:



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

$$z = a + ib = |z| \cdot \cos \varphi + i \cdot |z| \cdot \sin \varphi = |z| \cdot (\cos \varphi + i \cdot \sin \varphi)$$

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

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

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

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 Aspekte.

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:

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

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

$$(|z_1| \cdot e^{i\varphi_1}) \cdot (|z_2| \cdot e^{i\varphi_2}) = |z_1| \cdot |z_2| \cdot e^{i(\varphi_1 + \varphi_2)}$$

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.

* * *

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.

4. LINEARKOMBINATIONEN UND BASEN

Um Vektorräume im Allgemeinen verstehen zu können, werden wir diese etwas tiefer analysieren und 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 4.1. Sei V ein \mathbb{K} -Vektorraum und $v_1, \dots, v_n \in V$. Ein $v \in V$ ist eine **Linearkombination** (kurz: LK) von v_1, \dots, v_n , wenn es $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ gibt, sodass

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

Die Körperelemente $\lambda_1, \dots, \lambda_n$ heißen **Koeffizienten** dieser Linearkombination.

Die Menge $\mathcal{L}(v_1, \dots, v_n) := \{\sum_{i=1}^n \lambda_i v_i \mid \lambda_1, \dots, \lambda_n \in \mathbb{K}\} \subseteq V$ aller Linearkombinationen von v_1, \dots, v_n heißt **Lineare Hülle** von v_1, \dots, v_n .

Im Fall $n = 0$ setzen wir $\mathcal{L}(\emptyset) := \{0\}$.

Satz 4.2. Sei V ein \mathbb{K} -Vektorraum und $v_1, \dots, v_n \in V$. Dann ist $\mathcal{L}(v_1, \dots, v_n)$ ein Untervektorraum von V .

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

Es bleibt die Abgeschlossenheit bezüglich der Vektorraum-Addition und Skalarmultiplikation 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:

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

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

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

!TO-DO!

[Bild]

Definition 4.3. Sei V ein \mathbb{K} -Vektorraum. Dann heißt V **endlich erzeugt**, wenn es eine endliche Folge (v_1, \dots, v_n) von Vektoren in V gibt, sodass $\mathcal{L}(v_1, \dots, v_n) = V$ 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 4.18 betrachtet.

Definition 4.4. Sei V ein \mathbb{K} -Vektorraum und $v_1, \dots, v_n \in V$. Das n -Tupel (v_1, \dots, v_n) heißt *linear unabhängig* (kurz: l.u.), wenn für alle Darstellungen

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

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

Ist das n -Tupel (v_1, \dots, 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, \dots, v_n betrachten Sie das zugehörige homogene LGS

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

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

Beispiel 4.5. Die Vektoren $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \begin{pmatrix} 7 \\ 8 \\ 9 \end{pmatrix}$ im \mathbb{R}^3 sind linear abhängig, da gilt:

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

Satz 4.6. Sei V ein \mathbb{K} -Vektorraum und seien $v_1, \dots, v_n \in V$. Dann ist (v_1, \dots, v_n) linear unabhängig genau dann, wenn für alle $i \leq n$ gilt: v_i ist nicht Linearkombination der Vektoren $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$.

Beweis: Wir beweisen zuerst “ \Rightarrow ”:

Sei (v_1, \dots, v_n) linear unabhängig. Betrachte $i \leq n$ beliebig. Angenommen v_i wäre eine Linearkombination von $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$. Wähle $\lambda_1, \dots, \lambda_{i-1}, \lambda_{i+1}, \dots, \lambda_n \in \mathbb{K}$, sodass

$$\begin{aligned} v_i &= \lambda_1 v_1 + \dots + \lambda_{i-1} v_{i-1} + \lambda_{i+1} v_{i+1} + \dots + \lambda_n v_n \\ &= \sum_{j=1, j \neq i}^n \lambda_j v_j \end{aligned}$$

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 ($0 \neq -1$) und dies ist ein Widerspruch zur linearen Unabhängigkeit der v_1, \dots, v_n .

Es bleibt “ \Leftarrow ” zu zeigen:

Sei also für alle $i \leq n$ v_i keine Linearkombination der Vektoren $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$. Angenommen (v_1, \dots, v_n) wären linear abhängig. Wähle Koeffizienten $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, sodass $0 = \lambda_1 v_1 + \dots + \lambda_n v_n$ 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 $v_i = -\frac{\lambda_1}{\lambda_i} v_1 - \dots - \frac{\lambda_{i-1}}{\lambda_i} v_{i-1} - \frac{\lambda_{i+1}}{\lambda_i} v_{i+1} - \dots - \frac{\lambda_n}{\lambda_i} v_n$. Das ist eine Linearkombination von $v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n$ und somit ein Widerspruch zur Voraussetzung. \square

Definition 4.7. Sei V ein \mathbb{K} -Vektorraum und v_1, \dots, v_n Vektoren, dann ist v_1, \dots, v_n eine Basis von V , falls v_1, \dots, v_n linear unabhängig sind und $\mathcal{L}(v_1, \dots, v_n) = V$.

Satz 4.8. Sei V ein \mathbb{K} -Vektorraum mit Basis (v_1, \dots, v_n) . Für alle Vektoren $v \in V$ existiert genau ein n -Tupel $(\lambda_1, \dots, \lambda_n)$, sodass $v = \sum_{i=1}^n \lambda_i v_i$. Wir nennen $(\lambda_1, \dots, \lambda_n)$ die Darstellung von v in der Basis v_1, \dots, 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 nicht-triviale 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. \square

Definition 4.9. Sei $n \in \mathbb{N}$ eine natürliche Zahl und sei \mathbb{K} ein Körper, dann ist

$$(e_1, \dots, e_n) \text{ mit } e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \text{ die kanonische Basis des}$$

Vektorraumes \mathbb{K}^n .

Lemma 4.10. (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 $\mathbb{K}^n = \{(\lambda_1, \dots, \lambda_n) \mid \lambda_1, \dots, \lambda_n \in \mathbb{K}\}$ 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

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

\square

* * *

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

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

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

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

Induktionsanfang: Sei $s \neq 0$. Dann ist (v_1, \dots, v_r) linear unabhängig in V und $\mathcal{L}(v_1, \dots, v_r) = V$. Also ist (v_1, \dots, 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, \dots, v_r) linear unabhängig in V , $w_1, \dots, w_{s+1} \in V$ mit

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

Fall 1: Für alle $i \in \{1, \dots, s + 1\}$ gilt: $w_i \in \mathcal{L}(v_1, \dots, v_r)$.

Dann ist jede Linearkombination der Vektoren

$v_1, \dots, v_r, w_1, \dots, w_s, w_{s+1}$ eine Linearkombination der Vektoren v_1, \dots, v_r und es gilt:

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

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

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

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

Beweis der Behauptung: Sei $0 = \lambda_1 v_1 + \dots + \lambda_r v_r + \lambda w_i$. Wenn $\lambda \neq 0$ wäre, dann gilt: $w_i = -\frac{1}{\lambda}(\sum_{j=1}^r \lambda_j v_j) \in \mathcal{L}(v_1, \dots, v_r)$. Dies ist ein Widerspruch.

Also ist $\lambda = 0$ und somit gilt $\sum_{j=1}^r \lambda_j v_j = 0$. Da (v_1, \dots, v_r) linear unabhängig sind, gilt: $\lambda_1 = \dots = \lambda_r = 0$. □(Beh.)

Weiter gilt (nach Umsortierung)

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

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

Damit ist die Aussage per vollständiger Induktion bewiesen. \square

Satz 4.12 (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 . \square

Satz 4.13 (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 $i \leq m$.

- (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 $w_j \in \mathcal{L}(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_m)$. 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:

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

Weiter ist

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

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 4.14. *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 4.15. *Sei V ein endlich erzeugter \mathbb{K} -Vektorraum. Dann ist die Dimension von V die Zahl n der Vektoren in einer Basis (v_1, \dots, v_n) von V . V heißt in diesem Fall auch n -dimensional. Die Dimension von V wird mit $\dim(V)$ (oder auch einfach $\dim V$) bezeichnet.*

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

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

Satz 4.16. *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 $\lambda_1 \cdot v_1 + \dots + \lambda_r \cdot v_r + \lambda \cdot u = 0$. Dann ist $\lambda = 0$, da $u \notin \mathcal{L}(v_1, \dots, v_r)$. Da (v_1, \dots, v_r) linear unabhängig ist, gilt:

$$\lambda_1 = \dots = \lambda_r = 0$$

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

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

Satz 4.17. *Sei V ein endlich-dimensionaler \mathbb{K} -Vektorraum und $U \subseteq V$ ein Unterraum. Dann ist*

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

Beweis: Wir zeigen zunächst " \Rightarrow ":

Angenommen $U \neq V$. Wähle eine Basis (v_1, \dots, v_r) von U . Die Folge (v_1, \dots, v_r) ist linear unabhängig in V . Nach dem Basisergänzungssatz kann man diese Folge zu einer Basis $(v_1, \dots, v_r, w_1, \dots, w_s)$ von V ergänzen. Da $U \neq V$ ist, gilt $s \neq 0$. Damit gilt

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

" \Leftarrow " ist klar. \square

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

Der Satz 4.17 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 4.18. *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 $(\lambda_1, \dots, \lambda_n)$ aus \mathbb{K} . (Basisdarstellung zur fixierten Basis.) Es gibt k^n -viele derartiger n -Tupel. \square

Beispiel 4.19. 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 jeden 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 verschiedene Grundkörper sich die Länge 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 $(1, i)$ als Basis. Die

komplexen Zahlen als \mathbb{Q} -Vektorraum sind aus Gründen der Mächtigkeit nicht endlich dimensional. Eine Basis bildet etwa die Folge der irrationalen Zahlen ergänzt um die Elemente der Menge der Zahlen der Gestalt ir für irrationale Zahlen r .

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

Beispiel 4.20. 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:

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

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

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

allerdings ist $\sqrt{6} \notin \mathbb{Q}$. Dieses 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 4.21. Die Vektoren $1, \sqrt{2}, \sqrt{3}$ im \mathbb{R} -Vektorraum 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 $\sqrt{3}$.

5. 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 von Homomorphismus zwischen allgemeinen Strukturen.

Definition 5.1. Seien V und W \mathbb{K} - Vektorräume und $f: V \rightarrow W$. 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 5.2. Seien V, W, X drei \mathbb{K} -Vektorräume und $f: V \rightarrow W$ und $g: W \rightarrow X$ zwei lineare Abbildungen. Dann sind die beiden Abbildungen $g \circ f: V \rightarrow X$ und $\text{id}_V: V \rightarrow V$ 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 5.3. 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

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

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

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

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

- (c) *Drehung des \mathbb{R}^2* : Wir hatten gesehen, dass die Multiplikation mit der imaginären Einheit i einer Drehung um 90 entspricht. Damit ist eine Drehung des \mathbb{R}^2 um 90 linear. Auch Drehungen um andere Winkel sind lineare Abbildungen.

Beispiel 5.4. Betrachten wir die Abbildung $f: \mathbb{Q} \times \mathbb{Q} \rightarrow \mathbb{R}$, 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):

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

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

Satz 5.5. Seien V und W \mathbb{K} -Vektorräume und $f: V \rightarrow W$ linear. Dann gilt:

- (a) Das Bild $f[V]$ ist ein Untervektorraum von W . Man bezeichnet diesen auch als $\text{Bild}(f)$.
- (b) Das Urbild $f^{-1}[\{0\}]$ ist ein Untervektorraum von V . Man nennt $f^{-1}[\{0\}]$ auch den Kern der Abbildung f und bezeichnet ihn auch als $f^{-1}(0)$ oder $\text{Kern}(f)$.

Beweis: Wir beweisen jeweils die Eigenschaften eines Untervektorraums:

- (a) 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

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

- (b) Es gilt $f(0) = 0$, also $0 \in f^{-1}[\{0\}]$, also $\text{Kern}(f) \neq \emptyset$. Betrachte $x, y \in \text{Kern}(f)$ 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 $x + y \in \text{Kern}(f)$. Ebenso gilt:

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

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

□

Definition 5.6. Die Dimension des Bildes einer linearen Abbildung bezeichnet man als Rang dieser Abbildung, d.h. für $f: V \rightarrow W$ linear schreiben wir kurz für den Rang

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

Die Dimension des Kerns wird als Defekt bezeichnet, kurz:

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

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

Satz 5.7. Sei $F: V \rightarrow V$ eine lineare Abbildung und $V = \mathcal{L}(v_1, \dots, v_n)$. Dann gilt:

$$\text{Bild}(F) = \mathcal{L}(F(v_1), \dots, F(v_n)).$$

Beweis:

Wir zeigen beide Inklusionen:

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

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

“ \subseteq ” Sei $v \in \text{Bild}(F)$, das heißt es existiert ein $\bar{v} \in V$ mit $v = F(\bar{v})$. Da $V = \mathcal{L}(v_1, \dots, v_n)$ 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 $v \in \mathcal{L}(F(v_1), \dots, F(v_n))$.

□

Beispiel 5.8. Sei $F: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ 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 $\text{Kern}(f)$ und $\text{Bild}(f)$.

Bild(f): Offenbar gilt $\mathbb{R}^3 = \mathcal{L}(e_1, e_2, e_3)$. Nach Satz 5.7 gilt:

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

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 5.17 beweisen, der besagt, dass die Summe des Rangs und des Defekts eines Homomorphismus gleich der Dimension des Urbild-Vektorraums ist:

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

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: $x_2 = -x_3$ 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:

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

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

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

Definiere die Addition “+” durch:

$$+ : \text{Hom}(V, W) \times \text{Hom}(V, W) \rightarrow \text{Hom}(V, W)$$

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

Definiere die skalare Multiplikation “ \cdot ” durch:

$$\cdot : \mathbb{K} \times \text{Hom}(V, W) \rightarrow \text{Hom}(V, W)$$

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

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

$$0(x) := 0$$

Beachten Sie, dass wir eigentlich schreiben müssten:

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

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

Satz 5.10. Seien V, W zwei \mathbb{K} -Vektorräume, dann ist $\text{Hom}(V, W)$ ein \mathbb{K} -Vektorraum.

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

Diese ist allerdings klar, denn es gilt:

$$\begin{aligned} (f + g)(x) &= f(x) + g(x) && \text{per Definition} \\ &= g(x) + f(x) && W \text{ ist Vektorraum} \\ &= (g + f)(x) && \text{per Definition} \end{aligned}$$

⊠

Definition 5.11. Sei $f: V \rightarrow W$ ein Homomorphismus. Dann definieren wir:

- (a) f ist ein **Monomorphismus**, wenn f injektiv ist.
- (b) f ist ein **Epimorphismus**, wenn f surjektiv ist.
- (c) f ist ein **Isomorphismus**, wenn f bijektiv ist.

- (d) f ist ein *Endomorphismus*, wenn $V = W$ ist.
 (e) f ist ein *Automorphismus*, wenn f bijektiv und $V = W$ ist.

Satz 5.12. Sei $f: V \rightarrow W$ ein Isomorphismus von \mathbb{K} -Vektorräumen. Dann ist $f^{-1}: W \rightarrow V$ ebenfalls ein Isomorphismus.

Beweis: Da f bijektiv ist, dürfen wir auch von der Umkehrabbildung f^{-1} sprechen. Die Bijektivität von $f^{-1}: W \rightarrow V$ 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:

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

und

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

□

Beispiel 5.13. Setze $\text{Aut}(V) := \{f: V \rightarrow V \mid f \text{ Automorphismus}\}$. Mittels der Hintereinanderausführung von Funktionen \circ bildet

$$(\text{Aut}(V), \circ)$$

eine Gruppe, die so genannte *Automorphismengruppe*: Die notwendigen Eigenschaften Assoziativität, Existenz des neutralen Elements und von inverses 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 5.14. 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 $f: V \rightarrow W$, sodass 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:

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

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

(Eindeutigkeit) Angenommen $f: V \rightarrow W$ und $f': V \rightarrow W$ 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:

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

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

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

Bemerkung 5.16. 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 eine Bijektion haben können wir also anstatt auf V zu rechnen, mit Vektoren rechnen.

Beweis von Satz 5.15: Wähle eine Basis (v_1, \dots, v_n) von V . Dann gibt es eine lineare Abbildung $f: V \rightarrow \mathbb{K}^n$ (nach Satz 5.14) mit:

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

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

Wir zeigen, dass $f: V \rightarrow \mathbb{K}^n$ bijektiv ist:

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

$$\begin{aligned} (\lambda_1, \dots, \lambda_n) &= \sum_{i=1}^n \lambda_i \cdot e_i = \sum_{i=1}^n \lambda_i \cdot f(v_i) \\ &= f\left(\sum_{i=1}^n \lambda_i v_i\right) = f(v) \in \text{Bild}(f) \end{aligned}$$

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

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

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

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

☐(Satz 5.15)

* * *

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

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

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

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

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

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

$$\begin{aligned} \text{Bild}(f) &= \{f(v) \mid v \in V\} \\ &= \{f(\lambda_1 v_1 + \dots + \lambda_r v_r + \lambda_{r+1} v_{r+1} + \dots + \lambda_n v_n) \mid \lambda_1, \dots, \lambda_n \in \mathbb{K}\} \\ &= \{\underbrace{\lambda_1 \cdot f(v_1)}_{=0} + \dots + \underbrace{\lambda_r \cdot f(v_r)}_{=0} + \lambda_{r+1} \cdot f(v_{r+1}) + \dots + \lambda_n \cdot f(v_n) \mid \\ &\hspace{20em} \lambda_1, \dots, \lambda_n \in \mathbb{K}\} \\ &= \{\lambda_{r+1} f(v_{r+1}) + \dots + \lambda_n f(v_n) \mid \lambda_1, \dots, \lambda_n \in \mathbb{K}\} \\ &= \mathcal{L}(f(v_{r+1}), \dots, f(v_n)) \end{aligned}$$

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

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

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

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

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

Dies ist offenbar äquivalent zu der Gleichung:

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

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

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

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

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

☒

6. 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 Gleichungssystemen 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 6.1. Sei \mathbb{K} ein Körper und sei $f: \mathbb{K}^m \rightarrow \mathbb{K}^n$ 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 $f(e_i) = \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix}$. Sei \mathbb{K} ein Körper und sei $f: \mathbb{K}^m \rightarrow \mathbb{K}^n$ 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

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

gegeben.

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

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

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:

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

□

Der Satz 6.1 rechtfertigt die folgende

Definition 6.2. Sei \mathbb{K} ein Körper und $n, m \in \mathbb{N}$. Eine $n \times m$ -Matrix über \mathbb{K} ist eine Folge $M = (a_{ij} \mid i = 1, \dots, n; j = 1, \dots, m)$ von Elementen $a_{ij} \in \mathbb{K}$. Wir schreiben die Matrix als zweidimensionales Schema wie folgt:

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

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

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

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

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

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

Ist die Matrix $M = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix}$ durch $f(e_i) = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix}$ wie in Satz 6.1 definiert, so nennen wir M die **darstellende Matrix** von f . Wir schreiben kurz

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

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

Bemerkung 6.3. (Merkregel) Die Spalten der darstellenden Matrix sind die Bilder der Basisvektoren.

Satz 6.4. Für jede $n \times m$ -Matrix M über \mathbb{K} gibt es eine eindeutig bestimmte lineare Abbildung $f: \mathbb{K}^m \rightarrow \mathbb{K}^n$, sodass M die darstellende Matrix von f ist. Damit ist die Abbildung

$$\begin{aligned} \text{Hom}(\mathbb{K}^m, \mathbb{K}^n) &\rightarrow \text{Mat}(n \times m, \mathbb{K}) \\ f &\mapsto \text{DM}(f) \end{aligned}$$

eine Bijektion.

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

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

* * *

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

(a) Definiere die **Matrizenaddition** durch:

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

(b) Definiere die *skalare Multiplikation* durch:

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

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

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

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

Beweis: Offenbar ist die Zuordnung

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

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

An dieser Stelle können wir den Satz 6.4 verstärken und den bereits als Bijektion erkannten Zusammenhang zum Isomorphismus erweitern:

Satz 6.7 (Isomorphissatz). Sei \mathbb{K} ein Körper und $n, m \in \mathbb{N}$. Dann ist die Abbildung

$$\text{Hom}(\mathbb{K}^m, \mathbb{K}^n) \rightarrow \text{Mat}(n \times m, \mathbb{K});$$

$$f \mapsto \text{DM}(f)$$

(sogar) ein Isomorphismus von Vektorräumen.

Beweis: Nach Satz 6.4 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 $\text{Hom}(\mathbb{K}^n, \mathbb{K}^m)$ und $\text{Mat}(n \times m, \mathbb{K})$, x und y sind Elemente aus dem Startraum, in unsrem Fall also Abbildungen aus $\text{Hom}(\mathbb{K}^n, \mathbb{K}^m)$. Nennen wir die Abbildung aus dem Satz ϕ und Elemente aus $\text{Hom}(\mathbb{K}^n, \mathbb{K}^m)$ mit f und g , so erhalten wir die Schreibweise:

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

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

$$(a) \text{DM}(f+g) = \text{DM}(f) + \text{DM}(g)$$

$$(b) \text{ DM}(\lambda \cdot f) = \lambda \cdot \text{DM}(f)$$

Zu (a): Seien $f, g \in \text{Hom}(\mathbb{K}^n, \mathbb{K}^m)$ mit darstellenden Matrizen $\text{DM}(f)$ und $\text{DM}(g)$. Die Spalten von $\text{DM}(f)$ sind gerade die Vektoren $f(e_1), \dots, f(e_m)$ und die Spalten von $\text{DM}(g)$ sind die Vektoren $g(e_1), \dots, g(e_m)$. Die Spalten von $\text{DM}(f+g)$ sind die Vektoren $(f+g)(e_1) \stackrel{(*)}{=} f(e_1)+g(e_1), \dots, (f+g)(e_m) \stackrel{(*)}{=} f(e_m)+g(e_m)$. $(*)$ gilt wegen Linearität.

Nach Definition der Matrizenaddition ist damit

$$\text{DM}(f+g) = \text{DM}(f) + \text{DM}(g)$$

Zu (b): Sei f ein Homomorphismus aus $\text{Hom}(\mathbb{K}^m, \mathbb{K}^n)$ mit darstellender Matrix $\text{DM}(f)$ und sein $\lambda \in \mathbb{K}$. Die Spalten von $\text{DM}(\lambda \cdot f)$ sind die Vektoren $(\lambda \cdot f)(e_1) = \lambda \cdot f(e_1), \dots, (\lambda \cdot f)(e_m) = \lambda \cdot f(e_m)$. Nach Definition der Skalarmultiplikation von Matrizen gilt damit wie gewünscht $\text{DM}(\lambda \cdot f) = \lambda \cdot \text{DM}(f)$. \square

* * *

Wir werden jetzt eine Multiplikation von Matrizen einführen, sodass wir die schreibweise des Beispiels 5.8 rechtfertigen:

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

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

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

$$\begin{aligned}
&= \begin{pmatrix} b_{1j}a_{11} + \dots + b_{nj}a_{1n} \\ \vdots \\ b_{1j}a_{r1} + \dots + b_{nj}a_{rn} \end{pmatrix} \\
&= \begin{pmatrix} a_{11}b_{1j} + \dots + a_{1n}b_{nj} \\ \vdots \\ a_{r1}b_{1j} + \dots + a_{rn}b_{nj} \end{pmatrix} \quad \text{Körperaxiome}
\end{aligned}$$

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

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

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:

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

Damit ist die folgende Definition motiviert:

Definition 6.8. *Definiere das Matrizenprodukt für zwei Matrizen $A = (a_{ik}) \in \text{Mat}(r \times n, \mathbb{K})$ und $B = (b_{kj}) \in \text{Mat}(n \times m, \mathbb{K})$ als*

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

Beispiel 6.9. Allgemein gilt für Matrizen $A, B \in \text{Mat}(2 \times 2, \mathbb{K})$:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} a \cdot e + b \cdot g & a \cdot f + b \cdot h \\ c \cdot e + d \cdot g & c \cdot f + d \cdot h \end{pmatrix} \quad (\star)$$

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 6.10. Im Allgemeinen ist das Matrizenprodukt nicht kommutativ.

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

Beispiel 6.11. 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:

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

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{matrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{matrix}$$

wobei

$$c_{11} := 1 \cdot 1 + 2 \cdot 2 + 3 \cdot 3 = 14, \quad c_{12} := 1 \cdot 2 + 2 \cdot 3 + 3 \cdot 4 = 20,$$

$$c_{21} := 3 \cdot 1 + 2 \cdot 2 + 1 \cdot 3 = 10, \quad c_{22} := 3 \cdot 2 + 2 \cdot 3 + 1 \cdot 4 = 16$$

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 6.12. Sei $f: \mathbb{K}^m \rightarrow \mathbb{K}^n$ eine lineare Abbildung und $A = \text{DM}(f)$, 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 (★) aus Beispiel 6.9 und Satz 6.1 ist die Behauptung klar. \square

Satz 6.13. 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 6.14. 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 6.7 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 :

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

Damit gilt wie gewünscht: $f \circ (g + h) = f \circ g + f \circ h$

- (c) Folgt analog zu (b). ☒

Bemerkung 6.15. Die Gesetze aus Satz 6.13 können auch durch (mühseliges) Nachrechnen mit der formalen Definition der Matrixmultiplikation bewiesen werden.

Definition 6.16. Sei \mathbb{K} ein Körper.

- (a) Eine Matrix A über \mathbb{K} heißt **quadratisch**, wenn es eine natürliche Zahl n gibt, so dass $A \in \text{Mat}(n \times n, \mathbb{K})$.
- (b) Die n -dimensionale Einheitsmatrix ist die darstellende Matrix der identischen Abbildung $\text{id}_{\mathbb{K}^n}$:

$$E_n = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Satz 6.17. Sei \mathbb{K} ein Körper und $A \in \text{Mat}(n \times m, \mathbb{K})$. 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 6.18. Sei \mathbb{K} ein Körper und $n \in \mathbb{N}$. Eine $n \times n$ -Matrix A über \mathbb{K} heißt **invertierbar**, wenn es ein $B \in \text{Mat}(n \times n, \mathbb{K})$ 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 6.19. Sei \mathbb{K} ein Körper und $n \in \mathbb{N}$, dann gilt:

- (a) Sei $A \in \text{Mat}(n \times n, \mathbb{K})$ invertierbar. Dann gibt es genau ein B mit $A \cdot B = E_n$.
- (b) Sei $A \in \text{Mat}(n \times n, \mathbb{K})$ invertierbar. Dann gilt $A \cdot A^{-1} = A^{-1} \cdot A = E_n$.
- (c) Sei $A \in \text{Mat}(n \times n, \mathbb{K})$ invertierbar. Dann ist A^{-1} ebenfalls invertierbar und es gilt $(A^{-1})^{-1} = A$.
- (d) Seien $A, B \in \text{Mat}(n \times n, \mathbb{K})$ 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 6.7 mit ihren darstellenden Matrizen. Angenommen, A ist invertierbar und es gilt $A \cdot B = E_n$ (Existenz von B ist nach Definition 6.18 gesichert). Dann gilt:

$$\begin{aligned} \text{Rg}(A) &= \dim(\text{Bild}(A)) \\ &\geq \dim(\text{Bild}(A \cdot B)) = \dim(\text{Bild}(E_n)) \\ &= \dim(\mathbb{K}^n) = n \end{aligned}$$

Wegen $\text{Rg}(A) \leq n$ gilt daher $\text{Rg}(A) = n$. Damit ist die Abbildung A surjektiv (da nun $\text{Bild}(A) \subseteq \mathbb{K}^n$, n -dimensional, also $\text{Bild}(A) = \mathbb{K}^n$). Nach der Dimensionsformel gilt: $\dim(\text{Kern}(A)) + \dim(\text{Bild}(A)) = n$, also $\dim(\text{Kern}(A)) = 0$ und somit gilt $\text{Kern}(A) = \{0\}$. Wenn der $\text{Kern}(A) = \{0\}$ ist, dann ist A auch injektiv (Übungsaufgabe!).

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 $C = \text{DM}(A^{-1})$. 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 6.7. \square

* * *

Lineare Abbildungen und Matrizen. Wir hatten Matrizen als darstellende Matrizen linearer Abbildungen motiviert und eingeführt, das heißt für eine lineare Abbildung $f: V \rightarrow V$ ist $\text{DM}(f)$ eine Matrix. Eine lineare Abbildung $f: \mathbb{K}^m \rightarrow \mathbb{K}^n$ besitzt eine darstellende Matrix $\text{DM}(f) \in \text{Mat}(n \times m, \mathbb{K})$ und es gilt:

$$f(x) = \text{DM}(f) \cdot x,$$

wobei x als Spaltenvektor $x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \in \mathbb{K}^m$ aufgefasst wird.

Umgekehrt definiert jede Matrix $A \in \text{Mat}(n \times m, \mathbb{K})$ die Abbildung $\mathbb{K}^m \rightarrow \mathbb{K}^n$; $x \mapsto A \cdot x$. 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 $A = \text{DM}(A)$, daher wird oftmals die so definierte Abbildung mit demselben Buchstaben A bezeichnet:

$$A: \mathbb{K}^m \rightarrow \mathbb{K}^n$$

Damit lassen sich von Abbildungen bekannte Begriffe auf A übertragen:

- $A(x) := A \cdot x$
- $A[X] := \{A \cdot x \mid x \in X\}$
- $A^{-1}[Y] := \{x \in \mathbb{K}^m \mid A \cdot x \in Y\}$
- $\text{Kern}(A) := A^{-1}[\{0\}]$

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.

7. 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 Anwendungen von Matrizen in Form der Drehungen und Spiegelungen in der reellen Ebene und schliesslich im letzten Teil wenden wir 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 eine k -dimensionaler 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 $\{(0, 0, 0), (1, 1, 1)\}$ eine 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 $A \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$ 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, in dem wir nun in die korrekte Lösung

$$c := \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} \text{ zwei Fehler } e_1 := \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \text{ bzw. } e_2 := \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \text{ einarbeiten}$$

$$\text{und so } c + e_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \text{ bzw. } c + e_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \text{ erhalten.}$$

Mittels der Prüfmatrix erhalten wir schliesslich die beiden Vektoren

$$H \cdot (c + e_1) = \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix} \text{ bzw. } H \cdot (c + e_2) = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$$

Da die Einträge modulo 2 gesehen werden müssen, entsprechen die beiden Vektoren gerade den folgenden beiden:

$$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ bzw. } \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Jetzt können wir ablesen, dass der erste Fehler an der Stelle

$$0 \cdot 2^0 + 1 \cdot 2^1 + 0 \cdot 2^2 = 2$$

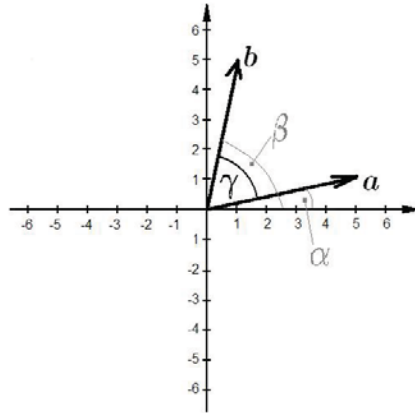
bzw. der zweite Fehler an der Position $0 \cdot 2^0 + 0 \cdot 2^1 + 1 \cdot 2^2 = 4$ liegt. Und dieses entspricht genau der Position der 1 in den Fehlervektoren e_1 bzw. e_2 .

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 $\alpha + \gamma = \beta$.



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):

$$\begin{aligned}
 x_2 &= \cos(\beta) \cdot |b| \\
 &= \cos(\alpha + \gamma) \cdot |b| \\
 &= \cos(\alpha) \cdot \cos(\gamma) \cdot |b| - \sin(\alpha) \cdot \sin(\gamma) \cdot |b| \\
 &= \cos(\gamma) \cdot (\cos(\alpha) \cdot |a|) - \sin(\gamma) \cdot (\sin(\alpha) \cdot |a|) \quad (\text{es gilt } |a| = |b|) \\
 &= \cos(\gamma) \cdot x_1 - \sin(\gamma) \cdot y_1
 \end{aligned}$$

Darüber hinaus gilt:

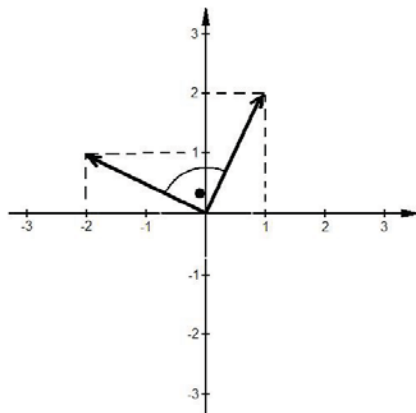
$$\begin{aligned}
 y_2 &= \sin(\beta) \cdot |b| \\
 &= \sin(\alpha + \gamma) \cdot |b| \\
 &= \sin(\alpha) \cdot \cos(\gamma) \cdot |b| + \cos(\alpha) \cdot \sin(\gamma) \cdot |b| \\
 &= \cos(\gamma) \cdot (\sin(\alpha) \cdot |a|) + \sin(\gamma) \cdot (\cos(\alpha) \cdot |a|) \quad (\text{es gilt } |a| = |b|) \\
 &= \cos(\gamma) \cdot y_1 + \sin(\gamma) \cdot x_1
 \end{aligned}$$

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 7.1. Für $\gamma = \frac{\pi}{2}$ ergibt sich $D_{90^\circ} = D_{\frac{\pi}{2}} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.



Für die Drehmatrizen können wir die von Drehungen bekannten Eigenschaften nachrechnen:

Satz 7.2. Für Winkel $\phi, \psi \in \mathbb{R}$ gilt:

$$D_{\phi+\psi} = D_{\phi} \cdot D_{\psi}$$

Der Beweis ist klar. Allerdings ergeben sich hieraus einfache trigonometrische Zusammenhänge:

$$\begin{aligned} D_{\phi+\psi} &= \begin{pmatrix} \cos(\phi + \psi) & -\sin(\phi + \psi) \\ \sin(\phi + \psi) & \cos(\phi + \psi) \end{pmatrix} \\ &= D_{\phi} \cdot D_{\psi} \\ &= \begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix} \cdot \begin{pmatrix} \cos(\psi) & -\sin(\psi) \\ \sin(\psi) & \cos(\psi) \end{pmatrix} \\ &= \begin{pmatrix} \cos(\phi) \cos(\psi) - \sin(\phi) \sin(\psi) & -\sin(\psi) \cos(\phi) - \sin(\phi) \cos(\psi) \\ \sin(\phi) \cos(\psi) + \sin(\psi) \cos(\phi) & -\sin(\phi) \sin(\psi) + \cos(\phi) \cos(\psi) \end{pmatrix} \\ &= \begin{pmatrix} \cos(\phi) \cos(\psi) - \sin(\phi) \sin(\psi) & -(\cos(\phi) \sin(\psi) + \sin(\phi) \cos(\psi)) \\ \sin(\phi) \cos(\psi) + \cos(\phi) \sin(\psi) & \cos(\phi) \cos(\psi) - \sin(\phi) \sin(\psi) \end{pmatrix} \end{aligned}$$

Ausserdem gilt:

$$\begin{aligned} &\begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix} \cdot \begin{pmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix} \\ &= \begin{pmatrix} \cos^2(\phi) + \sin^2(\phi) & \sin(\phi) \cos(\phi) - \sin(\phi) \cos(\phi) \\ \sin(\phi) \cos(\phi) - \sin(\phi) \cos(\phi) & \sin^2(\phi) + \cos^2(\phi) \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2 \end{aligned}$$

Damit ist $\begin{pmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix}$ die Inverse von D_{ϕ} . Aus diesen algebraischen Überlegungen erhalten wir folgende (analytischen) Aussagen:

Satz 7.3. *Es gilt:*

- (a) $\cos(\phi + \psi) = \cos(\phi)\cos(\psi) - \sin(\phi)\sin(\psi)$
- (b) $\sin(\phi + \psi) = \cos(\phi)\sin(\psi) + \sin(\phi)\cos(\psi)$
- (c) $\cos(-\phi) = \cos(\phi)$
- (d) $\sin(-\phi) = -\sin(\phi)$

Beweis: Die ersten beiden Punkte haben wir bereits gesehen.

Für die Aussagen (c) und (d) beachten Sie, dass $\begin{pmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix}$ invers zur Matrix $\begin{pmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix} = D_\phi$ ist.

Auf der anderen Seite ist aber die Drehung um den Winkel $-\phi$ ebenfalls invers zu D_ϕ , sodass gilt:

$$D_{-\phi} = \begin{pmatrix} \cos(-\phi) & -\sin(-\phi) \\ \sin(-\phi) & \cos(-\phi) \end{pmatrix} = \begin{pmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix}.$$

□

* * *

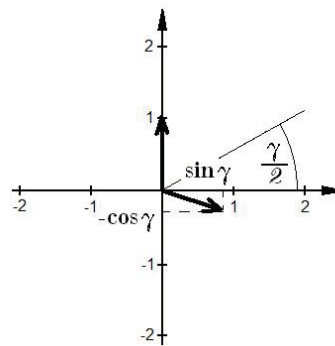
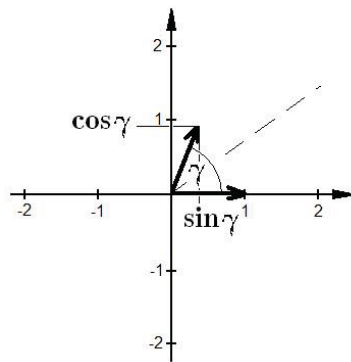
Matrizen als Spiegelungen in der reellen Ebene. Betrachte die Matrix

$$S_{\frac{\gamma}{2}} := \begin{pmatrix} \cos(\gamma) & \sin(\gamma) \\ \sin(\gamma) & -\cos(\gamma) \end{pmatrix}.$$

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:

$$S_{\frac{\gamma}{2}}(e_1) = \begin{pmatrix} \cos(\gamma) & \sin(\gamma) \\ \sin(\gamma) & -\cos(\gamma) \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos(\gamma) \\ \sin(\gamma) \end{pmatrix} \text{ und}$$

$$S_{\frac{\gamma}{2}}(e_2) = \begin{pmatrix} \cos(\gamma) & \sin(\gamma) \\ \sin(\gamma) & -\cos(\gamma) \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \sin(\gamma) \\ -\cos(\gamma) \end{pmatrix}$$



Die Bildkoordinaten für den ersten Basisvektor 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: $\frac{\pi}{2} - \frac{\gamma}{2} - \frac{\gamma}{2} = \frac{\pi}{2} - \gamma$. Damit erhalten wir schliesslich wieder mittels der Kosinus- bzw. Sinussätze wie oben.

Die x -Koordinate ist dann gegeben durch:

$$\begin{aligned}\cos\left(\frac{\pi}{2} - \gamma\right) &= \cos\left(\frac{\pi}{2}\right)\cos(\gamma) + \sin\left(\frac{\pi}{2}\right)\sin(\gamma) \\ &= 0 \cdot \cos(\gamma) + 1 \cdot \sin(\gamma) = \sin(\gamma).\end{aligned}$$

Die y -Koordinate ist analog:

$$\begin{aligned}\sin\left(\frac{\pi}{2} - \gamma\right) &= \cos\left(\frac{\pi}{2}\right)\sin(\gamma) - \sin\left(\frac{\pi}{2}\right)\cos(\gamma) \\ &= 0 \cdot \sin(\gamma) - 1 \cdot \cos(\gamma) = -\cos(\gamma)\end{aligned}$$

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 7.4. *Es gilt $S_{\frac{\gamma}{2}} \cdot S_{\frac{\gamma}{2}} = E_2$.*

Beweis:

$$\begin{aligned}&\begin{pmatrix} \cos(\gamma) & \sin(\gamma) \\ \sin(\gamma) & -\cos(\gamma) \end{pmatrix} \begin{pmatrix} \cos(\gamma) & \sin(\gamma) \\ \sin(\gamma) & -\cos(\gamma) \end{pmatrix} \\ &= \begin{pmatrix} \cos^2(\gamma) + \sin^2(\gamma) & \cos(\gamma)\sin(\gamma) - \sin(\gamma)\cos(\gamma) \\ \sin(\gamma)\cos(\gamma) - \cos(\gamma)\sin(\gamma) & \sin^2(\gamma) + \cos^2(\gamma) \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2\end{aligned}$$

□

Schliesslich erhält man durch einfaches Nachrechnen den folgenden

Satz 7.5. *Die Verknüpfung zweier Spiegelungen ist eine Drehung:*

$$S_{\frac{\phi}{2}} \cdot S_{\frac{\psi}{2}} = D_{\frac{\phi}{2} - \frac{\psi}{2}}$$

Damit schliessen wir unseren kurzen zwei-dimensionalen Exkurs. Wir kommen später im Kapitel über Eigenwerte noch einmal auf das Thema Drehungen und Spiegelungen zurück.

!TO-DO!

* * *

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 $r < s < m$ natürliche Zahlen. Definiere nun die Matrix $C(r, s)$, gegeben als $(c_{ij})_{i=1, \dots, m; j=1, \dots, m}$, 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:

$$C(r, s) = \begin{pmatrix} 1 & 0 & \dots & & & & \dots & 0 \\ 0 & \ddots & & & & & & \vdots \\ \vdots & & 1 & & & & & \\ & & & 0 & \dots & \dots & 0 & 1 \\ & & & \vdots & 1 & & & 0 \\ & & & \vdots & & \ddots & & \vdots \\ & & & 0 & & & 1 & \vdots \\ & & & 1 & 0 & \dots & & 0 \\ & & & & & & & 1 \\ \vdots & & & & & & & \ddots & 0 \\ 0 & \dots & & & & & \dots & 0 & 1 \end{pmatrix}$$

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 7.6. Für eine beliebige Matrix $A \in \text{Mat}(m \times n, \mathbb{K})$ 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 $A \in \text{Mat}(m \times n, \mathbb{K})$ gilt:

$$\begin{pmatrix} 1 & 0 & \dots & & \dots & 0 \\ 0 & \ddots & & & & \vdots \\ \vdots & \ddots & 1 & \ddots & & \vdots \\ & & & \lambda & & \\ & & & \ddots & 1 & \ddots \\ \vdots & & & & \ddots & 0 \\ 0 & \dots & & \dots & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{r1} & \dots & a_{rn} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ \lambda a_{r1} & \dots & \lambda a_{rn} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

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 7.7. Für $A \in \text{Mat}(m \times n, \mathbb{K})$ 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 $E(r, s, \lambda) = (e_{ij})_{i=1, \dots, m; j=1, \dots, m}$ 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 der selben Gestalt wie die Einheitsmatrix E_m , außer dass $e_{rs} = \lambda$ ist:

$$E(r, s, \lambda) = \begin{pmatrix} 1 & 0 & \dots & & \dots & 0 \\ 0 & \ddots & & & & \vdots \\ \vdots & & 1 & \dots & \lambda & \\ & & \ddots & \ddots & \vdots & \\ & & & & 1 & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \dots & & \dots & 0 & 1 \end{pmatrix}$$

Sei $A = (a_{ij}) \in \text{Mat}(m \times n, \mathbb{K})$. Dann ist:

$$E(r, s, \lambda) \cdot \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{r1} & \dots & a_{rn} \\ \vdots & & \vdots \\ a_{s1} & \dots & a_{sn} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{r1} + \lambda a_{s1} & \dots & a_{rn} + \lambda a_{sn} \\ \vdots & & \vdots \\ a_{s1} & \dots & a_{sn} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

Die Koeffizienten $\sum_{i=1}^m e_{ij} a_{jk}$ 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 7.8. Für $A \in \text{Mat}(m \times n, \mathbb{K})$ 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 7.9. Eine elementare Zeilenumformung ist eine 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**.

8. MATRIZEN UND LINEARE GLEICHUNGSSYSTEME

Wir betrachten lineare Gleichungssysteme aus Sicht der Matrizenrechnung. Sei \mathbb{K} ein Körper. Für $x_1, \dots, x_m \in \mathbb{K}$ ist das lineare Gleichungssystem über \mathbb{K} von der Gestalt

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

Dies ist äquivalent zu der Matrixgleichung $A \cdot x = b$:

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

Man nennt die Matrix A die **Koeffizientenmatrix** und

$$(A, b) = \begin{pmatrix} a_{11} & \dots & a_{1m} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \dots & a_{nm} & b_n \end{pmatrix}$$

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

$$\text{Lös}(A, b) = \{x \in \mathbb{K}^m \mid Ax = b\} = A^{-1}[\{b\}]$$

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 8.1. *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*

$$\text{Lös}(A, b) = z + \text{Lös}(A, 0) := \{z + x \mid x \in \text{Lös}(A, 0)\}.$$

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 $\{\emptyset\}$, eine Ursprungsgerade, also einen 1-dimensionalen Untervektorraum, oder der gesamte Raum ist, da es nur einen 2-dimensionalen Untervektorraum gibt.

Lösungsmengen von inhomogenen linearen Gleichungssystemen sind um einen Vektor verschobene Untervektorräume.

[Bild]

Wir zeigen umgekehrt, dass sich die geometrischen Grundelemente wie Punkte, Geraden, Ebenen usw. in allgemeiner Lage als Lösungsmengen darstellen lassen. damit lassen dich geometrische Aufgaben wie dir Berechnung von Schnittmengen von Geraden, Ebenen usw. auf die Lösung linearer Gleichungssysteme zurückführen.

Satz 8.2. 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, \dots, v_l) von U . Nach dem Basisergänzungssatz wähle v_{l+1}, \dots, v_m , so dass $(v_1, \dots, v_l, v_{l+1}, \dots, v_m)$ eine Basis von \mathbb{K}^m ist. Wir definieren die folgende lineare Abbildung $A: \mathbb{K}^m \rightarrow \mathbb{K}^{m-l}$ durch die Angabe der Bilder der obigen Basisvektoren:

$$A(v_1) = \dots = A(v_l) = 0,$$

$$A(v_{l+1}) = e_1, A(v_{l+2}) = e_2, \dots, A(v_m) = e_{m-l},$$

wobei e_1, \dots, e_{m-l} die kanonische Basis von \mathbb{K}^{m-l} ist.

Behauptung. Dann ist $U = \text{Kern}(A)$.

Beweis der Behauptung: Wir zeigen beide Inklusionen:

“ \subseteq ” Betrachte ein $x \in U$. Wähle $\lambda_1, \dots, \lambda_l \in \mathbb{K}$, 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 $x \in \text{Kern}(A)$.

“ \supseteq ” Sei $x \in \text{Kern}(A)$. Wähle $\lambda_1, \dots, \lambda_l, \lambda_{l+1}, \dots, \lambda_m \in \mathbb{K}$ mit $x = \sum_{i=1}^m \lambda_i v_i$.

Dann gilt:

$$\begin{aligned} 0 &= A \cdot x = A \left(\sum_{i=1}^m \lambda_i v_i \right) = \sum_{i=1}^m \lambda_i A(v_i) \\ &= \sum_{i=1}^l \lambda_i \underbrace{A(v_i)}_{=0} + \sum_{i=l+1}^m \lambda_i A(v_i) \\ &= \lambda_{l+1} e_1 + \dots + \lambda_m e_{m-l} \end{aligned}$$

Da die kanonische Basis e_1, \dots, e_{m-l} linear unabhängig ist, gilt:

$$\lambda_{l+1} = \dots = \lambda_m = 0.$$

Damit ist $x = \sum_{i=1}^l \lambda_i v_i \in \mathcal{L}(v_1, \dots, v_l) = U$. Indem wir A als Matrix in $\text{Mat}(m-l \times m, \mathbb{K})$ auffassen, ist $(A, 0)$ ein homogenes Lösungssystem mit

$$U = \text{Kern}(A) = \text{Lös}(A, 0). \quad \boxtimes(\text{Beh.})$$

(b) Setze $b := A(z)$. Dann ist $z \in \text{Lös}(A, b)$. Nach Satz 6.1 gilt

$$z + U = z + \text{Lös}(A, 0) = \text{Lös}(A, b).$$

\(\boxtimes\)

Auf dem Weg zum Gaußschen Algorithmus betrachten wir folgende Definition:

Definition 8.3. Sei \mathbb{K} ein Körper und $A = (a_{ij}) \in \text{Mat}(n \times m, \mathbb{K})$.

(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' :

$$A = \begin{pmatrix} 0 & a_{1j(1)} & \dots & a_{1j(2)} & a_{1j(3)} & \dots & \dots \\ & & & a_{2j(2)} & a_{2j(3)} & \dots & \\ & & & & & a_{3j(3)} & \dots \\ & & & & & & \ddots \end{pmatrix}$$

O-DO!

oder schematisch, wenn wir $*$ als Zeichen für beliebige Körperelemente und $\#$ als Zeichen für invertierbare Körperelemente benutzen:

$$A = \begin{pmatrix} 0 & \# & \dots & * & * & \dots & \dots \\ 0 & \dots & 0 & \# & * & \dots & \\ & & & & 0 & \# & \dots \\ & & & & & & \ddots \end{pmatrix}$$

!TO-DO!

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 8.4. 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 $P = F_1 \cdot F_2 \cdot \dots \cdot F_r$ von Elementarmatrizen, sodass $A' = P \cdot A$ und $b' = P \cdot b$. Wir wissen, dass P als elementare Zeilenumformung invertierbar ist, sodass gilt:

$$P^{-1} = (F_1 \cdot F_2 \cdot \dots \cdot F_r)^{-1} = F_r^{-1} \cdot \dots \cdot F_2^{-1} \cdot F_1^{-1}$$

Insgesamt definiert P eine injektive Abbildung. Damit gilt also:

$$\begin{aligned} \text{Lös}(A, b) &= \{x \mid Ax = b\} \\ &= \{x \mid P \cdot A \cdot x = P \cdot b\} \\ &= \text{Lös}(P \cdot A, P \cdot b) \end{aligned}$$

☒

Satz 8.5. 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

$$\begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mm} \end{pmatrix} \cdot \begin{pmatrix} x_{11} & \dots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{m1} & \dots & x_{mm} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

Dies ist ein lineares Gleichungssystem der Gestalt

$$\sum_{j=1}^m a_{ij} x_{jk} = 1$$

beziehungsweise

$$\sum_{j=1}^m a_{ij} x_{jk} = 0$$

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, \dots, F_l mit

$$F_1 \cdot \dots \cdot F_l \cdot A = E_m.$$

(Klar: Wenn A invertierbar ist, dann lässt sich A mittels elementarer Zeilenumformungen in die Einheitsmatrix überführen.)

Wir wissen bereits, dass dann das Produkt $F_1 \dots F_l$ invers zu A ist, das heißt es gilt:

$$A^{-1} = F_1 \cdot \dots \cdot F_l = F_1 \cdot \dots \cdot F_l \cdot E_m$$

Also: Man erhält daher dieses Produkt, indem man die elementaren Zeilenumformungen wie bei A parallel auf die Einheitsmatrix E_m anwendet.

Beispiel 8.6. Betrachte die Matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$. Gesucht ist die Inverse A^{-1} :

$$\begin{array}{ccccccc} 1 & 2 & 3 & 1 & 0 & 0 & 1 & 2 & 3 & 1 & 0 & 0 & 1 & 0 & 0 & \frac{1}{2} & \frac{3}{2} & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & \rightsquigarrow & 0 & -2 & -2 & -1 & 1 & 0 & \rightsquigarrow & \dots & \rightsquigarrow & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 & & 0 & 1 & 0 & 0 & 0 & 1 & & & & 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} & 1 \end{array}$$

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 Algorithmus Schritt für Schritt:

- links steht A , rechts steht E_m
- solange umformen, bis E_m links steht
- dann steht rechts A^{-1}

9. 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.

!TO-DO!

Basiswechsel. Zunächst überlegen wir uns, wie man die Koordinaten ein und desselben Vektors aber bezüglich verschiedener Basen leicht übertragen kann.

Definition 9.1. Sei V ein \mathbb{K} -Vektorraum und seien $B = (v_1, \dots, v_m)$ und $C = (w_1, \dots, w_m)$ Basen von V . Die Transformationsmatrix des Basiswechsels von B nach C ist die Matrix $T_C^B = (t_{ij}) \in \text{Mat}(m \times m, \mathbb{K})$, so dass für alle Spaltenindizes $j < m$

die Spalte $\begin{pmatrix} t_{1j} \\ \vdots \\ t_{mj} \end{pmatrix}$ die Darstellung von v_j in der Basis $C = (w_1, \dots, w_m)$ ist:

$$v_j = \sum_{i=1}^m t_{ij} w_i = t_{1j} w_1 + \dots + t_{mj} w_m,$$

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 9.2. Sei V ein \mathbb{K} -Vektorraum mit Basen $B = (v_1, \dots, v_m)$ und $C = (w_1, \dots, w_m)$ und der Transformationsmatrix $T_C^B = (t_{ij})$. Sei $z \in V$ und seien $(\lambda_1, \dots, \lambda_m)$ und (μ_1, \dots, μ_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:

$$\begin{aligned} \sum_{i=1}^m \mu_i w_i &= \sum_{i=1}^m \lambda_i v_i = \lambda_1 \cdot \sum_{i=1}^m t_{i1} \cdot w_i + \dots + \lambda_m \cdot \sum_{i=1}^m t_{im} \cdot w_i \\ &= \sum_{j=1}^m (\lambda_1 \cdot t_{j1} + \dots + \lambda_m \cdot t_{jm}) \cdot w_j \\ &= \sum_{j=1}^m (t_{j1} \lambda_1 + \dots + t_{jm} \lambda_m) w_j \end{aligned}$$

Da C eine Basis ist und die Darstellung bzgl. einer Basis eindeutig ist, gilt:

$$\mu_j = t_{j1} \lambda_1 + \dots + t_{jm} \lambda_m$$

Damit folgt die Behauptung. □

Die Transformationsmatrizen erfüllen einfache, zu erwartene Eigenschaften:

Satz 9.3. 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 9.1 und Satz 9.2 □

Beispiel 9.4. Sei $V = \mathbb{K}^m$ und B die kanonische Basis (e_1, \dots, e_m) . Weiterhin sei $C = (w_1, \dots, w_m)$ 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 B als

$$T_B^C = (w_1 \dots w_m) = \begin{pmatrix} w_{11} & \dots & w_{1m} \\ \vdots & & \vdots \\ w_{m1} & \dots & w_{mm} \end{pmatrix}$$

das heißt die Basis C ist im wesentlichen die Transformationsmatrix von der Basis C zur kanonischen Basis B , also T_C^B ; diese erhalten wir nach dem Satz 9.3 durch Invertieren von T_B^C .

Beispiel 9.5. 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:

$$\begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{bzgl. } K$$

$$= 1 \cdot \begin{pmatrix} 2 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{bzgl. } C$$

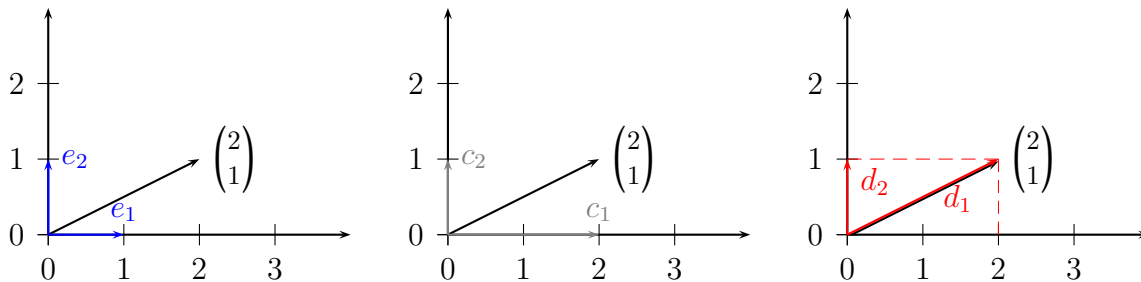
$$= 1 \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{bzgl. } D$$

oder kurz: $\begin{pmatrix} 2 \\ 1 \end{pmatrix}_K = \begin{pmatrix} 1 \\ 1 \end{pmatrix}_C = \begin{pmatrix} 1 \\ 0 \end{pmatrix}_D$

Bemerkung 9.6. 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 $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Somit können wir eine Koordinatenabbildung $K_D^{\mathbb{R}^2} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definieren, wobei einem Vektor v des Vektorraums seine Koordinaten in \mathbb{R}^2 zugeordnet werden:

$$v \mapsto \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad K_D^{\mathbb{R}^2}(v) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

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 wird: $K_B^V : V \rightarrow \mathbb{K}^n$.



• Ziel: $T_C^K \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

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:

$$\begin{pmatrix} 2 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

also gilt: $T_C^K = (T_K^C)^{-1} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix}$ und damit:

$$T_C^K \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (\checkmark)$$

- Ziel: $T_D^K \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

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:

$$\begin{array}{cccc} 2 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array} \rightsquigarrow \begin{array}{cccc} 1 & 0 & \frac{1}{2} & 0 \\ 1 & 1 & 0 & 1 \end{array} \rightsquigarrow \begin{array}{cccc} 1 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & -\frac{1}{2} & 1 \end{array}$$

also gilt: $T_D^K = (T_K^D)^{-1} = \begin{pmatrix} \frac{1}{2} & 0 \\ -\frac{1}{2} & 1 \end{pmatrix}$ und damit:

$$T_D^K \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ -\frac{1}{2} & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (\checkmark)$$

- Ziel: $T_D^C \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

Wir benutzen die Eigenschaft (b) aus Satz 9.3:

$$T_D^C = T_D^K \cdot T_K^C,$$

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:

$$T_D^C \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (\checkmark)$$

Bemerkung 9.7. 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 9.8. Sei $f: V \rightarrow W$ eine lineare Abbildung zwischen endlich dimensionalen \mathbb{K} -Vektorräumen. Sei $B = (v_1, \dots, v_m)$ eine Basis für V und $C = (w_1, \dots, w_n)$

eine Basis für W . Dann definiere $DM_{B,C}(f) = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} \in \text{Mat}(n \times m, \mathbb{K})$ die darstellende Matrix von f bezüglich der Basen B und C durch: für $j \leq m$ ist

$$f(v_j) = a_{1j}w_1 + \dots + a_{nj}w_n.$$

Das bedeutet, dass jeweils die Spalte $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix}$ die Darstellung von $f(v_j)$ in der Basis C ist.

Bemerkung 9.9. 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 6.3 zur Aufstellung der darstellenden Matrix allgemein erhalten.

Das Verhalten der darstellenden Matrizen können wir sehr einfach mittels Transformationsmatrizen zeigen:

Satz 9.10. Sei $f: V \rightarrow W$ eine lineare Abbildung zwischen endlich-dimensionalen \mathbb{K} -Vektorräumen. Seien $B = (v_1, \dots, v_m)$ und $B' = (v'_1, \dots, v'_m)$ Basen von V und $C = (w_1, \dots, w_n)$ und $C' = (w'_1, \dots, w'_n)$ Basen für W . Dann gilt:

$$DM_{B',C'}(f) = T_{C'}^C \cdot DM_{B,C}(f) \cdot T_B^{B'}$$

Beweis: Betrachte $j \leq m$. Es sei e_j der j -te kanonische Basisvektor von \mathbb{K}^m . Dann ist $DM_{B',C'}(f) \cdot e_j$ 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

$$DM_{B,C}(f) \cdot (T_B^{B'} \cdot e_j)$$

die Darstellung von $f(v'_j)$ in der Basis C .

$$T_{C'}^C \cdot (DM_{B,C}(f) \cdot (T_B^{B'} \cdot e_j))$$

ist dann die Darstellung von $f(v'_j)$ in der Basis C' . Damit gilt:

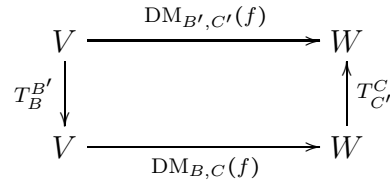
$$DM_{B',C'}(f) \cdot e_j = (T_{C'}^C \cdot DM_{B,C}(f) \cdot T_B^{B'}) \cdot e_j$$

Also sind die Matrizen spaltenweise identisch. □

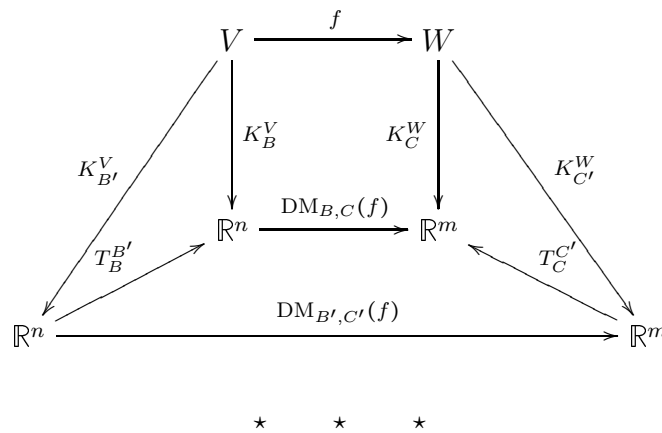
Bemerkung 9.11. Mittels des obigen Satzes erhalten wir durch den Übergang zur kanonischen Basis schnell allgemeine darstellende Matrizen, denn für $V = \mathbb{K}^m, W = \mathbb{K}^n$ und B, C die kanonischen Basen (hier in beiden Fällen übersichtlichkeithalber mit K dargestellt) gilt:

$$DM_{B',C'}(f) = T_{C'}^K \cdot DM(f) \cdot T_K^{B'}$$

Bemerkung 9.12. Das im Satz 9.10 dargestellte Transformationsverhalten der darstellenden Matrizen können wir durch ein kommutatives Diagramm darstellen:



Zusammenfassend erhalten wir für einen Endomorphismus $f : V \rightarrow W$ 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:



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.

O-DO!

Eine Teillösung liefert bereits der folgende Satz:

Satz 9.13. Sei $f: V \rightarrow W$ eine lineare Abbildung vom m -dimensionalen \mathbb{K} -Vektorraum V in den n -dimensionalen \mathbb{K} -Vektorraum W vom Rang: $\text{Rg}(f) = \dim(\text{Bild}(f)) = r$. Dann existieren Basen $B = (v_1, \dots, v_m)$ von V und $C = (w_1, \dots, w_n)$ von W , so dass

$$DM_{B,C}(f) = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & 1 & & \\ & & & 0 & \ddots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 0 \end{pmatrix}$$

wobei r viele Einsen auf der Hauptdiagonalen der Matrix stehen und sonst nur Nullen.

Warum ist dies nur eine Teillösung? Schliesslich brauchen wir doch genau r Einsen, da $\text{Rg}(f) = r$, 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 9.13: Nach der Dimensionsformel für lineare Abbildungen gilt:

$$\dim(\text{Kern}(f)) + \dim(\text{Bild}(f)) = m$$

Setze $s = \dim(\text{Kern}(f)) = m - r$. Wir benutzen die Idee aus dem Beweis des Dimensionssatzes 5.17. Wähle eine Basis (v_{r+1}, \dots, v_m) von $\text{Kern}(f)$. Nach dem Basisergänzungssatz wähle $v_1, \dots, v_r \in V$, sodass $B = (v_1, \dots, v_r, v_{r+1}, \dots, v_m)$ eine Basis von V ist. Somit gilt:

$$\begin{aligned} \text{Bild}(f) &= \{f(v) \mid v \in V = \mathcal{L}(v_1, \dots, v_r, v_{r+1}, \dots, v_m)\} \\ &= \{f(\lambda_1 v_1 + \dots + \lambda_r v_r + \lambda_{r+1} v_{r+1} + \dots + \lambda_m v_m) \mid \lambda_i \in \mathbb{K}\} \\ &= \{\lambda_1 f(v_1) + \dots + \lambda_r f(v_r) + \lambda_{r+1} f(v_{r+1}) + \dots + \lambda_m f(v_m) \mid \lambda_i \in \mathbb{K}\} \\ &= \{\lambda_1 f(v_1) + \dots + \lambda_r f(v_r) + 0 + \dots + 0 \mid \lambda_i \in \mathbb{K}\} \\ &= \mathcal{L}(f(v_1), \dots, f(v_r)) \end{aligned}$$

Setze $w_1 := f(v_1), \dots, w_r := f(v_r)$. Wir wissen bereits aus dem Beweis des Dimensionssatzes, dass (w_1, \dots, w_r) linear unabhängig sind. Nach dem Basisergänzungssatz wähle $w_{r+1}, \dots, w_n \in W$, sodass $C = (w_1, \dots, w_r, w_{r+1}, \dots, w_n)$ eine Basis für W ist.

Für $j = 1, \dots, r$ ist $f(v_j) = w_j$ und damit ist $(0, \dots, 0, 1, 0, \dots, 0)$ mit einer 1 an der j -ten Stelle die Darstellung von $f(v_j)$ in der Basis C .

Für $j = r + 1, \dots, m$ ist $f(v_j) = 0$, da diese Vektoren im Kern der Abbildung f liegen, sodass $(0, \dots, 0, 0, 0, \dots, 0)$ die Darstellung von $f(v_j)$ in der Basis C ist. Nach Definition gilt dann wie erwünscht

$$\text{DM}_{B,C}(f) = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & 1 & & \\ & & & 0 & \ddots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 0 \end{pmatrix}$$

□(Satz 9.13)

Bemerkung 9.14. Wie wir im Beweis gesehen haben, können wir nicht von $B = C$ ausgehen, da B durch den Basisergänzungssatz gewählt wurde und C durch die Bilder 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 $f: V \rightarrow V$ (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 9.15. Sei $f: V \rightarrow V$ 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

$$\text{DM}_B(f) := \text{DM}_{B,B}(f) \in \text{Mat}(m \times m, \mathbb{K}).$$

Die folgende Definition führt hilfreiche Begriffe ein.

Definition 9.16. Sei $A = (a_{ij}) \in \text{Mat}(n \times m, \mathbb{K})$ eine Matrix über dem Körper \mathbb{K} .

(a) Die Matrix $A^t = (a_{ij}^t) \in \text{Mat}(m \times n, \mathbb{K})$, sodass für alle $i \leq m$ und $j \leq n$ gilt

$$a_{ij}^t = a_{ji}$$

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*.

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.

O-DO!

Wir zeigen nun, dass wir bereits die bekannten Spiegelungen durch geeignete Wahl der Basen durch Basistransformation auf eine symmetrische Form bringen können.

Beispiel 9.17. 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 $B = (e_1, e_2)$ die kanonische Basis von \mathbb{R}^2 bezeichnet, so sind die Transformationsmatrizen

$$T_B^C = (u_0 \ u_1) = \begin{pmatrix} a & a \cdot \cos(\gamma) + b \cdot \sin(\gamma) \\ b & a \cdot \sin(\gamma) - b \cdot \cos(\gamma) \end{pmatrix} \quad \text{und} \quad T_C^B = (T_B^C)^{-1}.$$

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:

$$T_B^C = \begin{pmatrix} 1 & \cos(\gamma) \\ 0 & \sin(\gamma) \end{pmatrix} \quad \text{und} \quad T_C^B = (T_B^C)^{-1} = \frac{1}{\sin(\gamma)} \cdot \begin{pmatrix} \sin(\gamma) & -\cos(\gamma) \\ 0 & 1 \end{pmatrix}.$$

Wir berechnen die darstellende Matrix $DM_C(f)$. Nach dem Transformationssatz gilt:

$$\begin{aligned} DM_C(f) &= T_C^B \cdot DM_B(f) \cdot T_B^C \\ &= \begin{pmatrix} 1 & -\frac{\cos(\gamma)}{\sin(\gamma)} \\ 0 & \frac{1}{\sin(\gamma)} \end{pmatrix} \cdot \begin{pmatrix} \cos(\gamma) & \sin(\gamma) \\ \sin(\gamma) & -\cos(\gamma) \end{pmatrix} \cdot \begin{pmatrix} 1 & \cos(\gamma) \\ 0 & \sin(\gamma) \end{pmatrix} \\ &= \begin{pmatrix} 1 & -\frac{\cos(\gamma)}{\sin(\gamma)} \\ 0 & \frac{1}{\sin(\gamma)} \end{pmatrix} \cdot \begin{pmatrix} \cos(\gamma) & 1 \\ \sin(\gamma) & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{aligned}$$

Wir hätten $DM_C(f)$ auch einfacher bestimmen können, indem wir verstehen, dass die Spiegelung die beiden Vektoren u_0 und u_1 vertauscht, das heißt es gilt:

$$f(u_0) = u_1 \quad \text{und} \quad f(u_1) = u_0.$$

Inbesondere muss gelten:

$$DM_C(f) \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{und} \quad DM_C(f) \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Damit ist wegen der Eindeutigkeit einer linearen Abbildung auf einer Basis klar, dass $DM_C(f) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ ist. Diese darstellende Matrix ist symmetrisch.

Beispiel 9.18. 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 diesbzüglich auch die Bemerkung nach der Definition 12.2.) 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 $v_1 = \begin{pmatrix} \cos(\frac{\gamma}{2}) \\ \sin(\frac{\gamma}{2}) \end{pmatrix}$ und $v_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} \cos(\frac{\gamma}{2}) \\ \sin(\frac{\gamma}{2}) \end{pmatrix} = \begin{pmatrix} -\sin(\frac{\gamma}{2}) \\ \cos(\frac{\gamma}{2}) \end{pmatrix}$.

Setze die Basis D an als $D = (v_1, v_2)$. Dann gilt für die Transformationsmatrizen:

$$T_B^D = (v_1 v_2) = \begin{pmatrix} \cos(\frac{\gamma}{2}) & -\sin(\frac{\gamma}{2}) \\ \sin(\frac{\gamma}{2}) & \cos(\frac{\gamma}{2}) \end{pmatrix}$$

Die Transformationsmatrix T_D^B ist dann die Drehmatrix für den Winkel $-\frac{\gamma}{2}$:

$$T_D^B = (T_B^D)^{-1} = \begin{pmatrix} \cos(-\frac{\gamma}{2}) & -\sin(-\frac{\gamma}{2}) \\ \sin(-\frac{\gamma}{2}) & \cos(-\frac{\gamma}{2}) \end{pmatrix} = \begin{pmatrix} \cos(\frac{\gamma}{2}) & \sin(\frac{\gamma}{2}) \\ -\sin(\frac{\gamma}{2}) & \cos(\frac{\gamma}{2}) \end{pmatrix}$$

Schliesslich gilt:

$$\text{DM}_D(f) = T_D^B \cdot \text{DM}_B(f) \cdot T_B^D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Dies hätte man auch wieder schnell durch Überlegung entscheiden können, denn die Basisvektoren werden wie folgt abgebildet:

$$f(v_1) = v_1 \text{ und } f(v_2) = -v_2.$$

Im nächsten Kapitel werden diese Art der Beziehung schätzen lernen, nämlich dass $f(v_1) = 1 \cdot v_1$ und $f(v_2) = (-1) \cdot v_2$ gilt.

10. EIGENWERTE UND EIGENVEKTOREN

Es ist offenbar vorteilhaft die darstellende Matrix eines Endomorphismus f in Diagonalform zu bringen:

$$\text{DM}_C(f) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix} \in \text{Mat}(m \times m, \mathbb{K}),$$

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

$$f(v_i) = \lambda_i \cdot v_i. \quad (*)$$

Grundlegende Begriffe. Das Ziel des Diagonalisierens können wir folgendermaßen mithilfe der Wahl einer geeigneten Basis formalisieren:

Definition 10.1. Sei V ein \mathbb{K} -Vektorraum und $f: V \rightarrow V$ ein Endomorphismus. Dann ist f *diagonalisierbar*, wenn es eine Basis C von V gibt, so dass $DM_C(f)$ eine Diagonalmatrix ist.

Bereits durch die obige Einleitung ist offenbar die Gleichung (*) von besonderer Bedeutung.

Definition 10.2. Sei V ein \mathbb{K} -Vektorraum und $f: V \rightarrow V$ 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 $f(v) = \lambda \cdot v$.
- (c) Ein Vektor $v \in V$ heißt *Eigenvektor* von f , wenn es ein $\lambda \in \mathbb{K}$ gibt, so dass v ein Eigenvektor von f zum Eigenwert λ ist.

Damit erhalten wir sofort den folgenden

Satz 10.3. Sei V ein endlich-dimensionaler \mathbb{K} -Vektorraum und $f: V \rightarrow V$ 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

$$DM_C(f) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{pmatrix}$$

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 $C = (v_1, \dots, v_m)$ eine Basis von V , die aus Eigenvektoren besteht. Wähle $\lambda_1, \dots, \lambda_m$ 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

$$DM_C(f) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_m \end{pmatrix}$$

□

Eigenwerte und Eigenvektoren an bekannten Beispielen. Wir betrachten bekannte Endomorphismen des \mathbb{R}^2 und deren Bezüge zu Eigenwerten und -vektoren.

Beispiel 10.4. Wie bereits im Beispiel 9.17 gesehen, lässt sich eine Spiegelung des \mathbb{R}^2 in eine Matrix der Form $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ diagonalisieren.

Beispiel 10.5. 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 10.6. 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 zugehörigen 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:

$$(i) \quad a + b = \lambda \cdot a \text{ und}$$

$$(ii) \quad b = \lambda b$$

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 10.3 gilt also: Die Scherung h ist nicht diagonalisierbar.

Beispiel 10.7. 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.

$$\begin{aligned} A - \lambda_1 \cdot E_2 &= \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} - 2 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 2 & -1 \end{pmatrix} \end{aligned}$$

Anwendung des Gaußschen Algorithmus liefert:

$$\begin{array}{cc|cc} 2 & -1 & & 1 & -\frac{1}{2} \\ 2 & -1 & \rightsquigarrow & 0 & 0 \end{array}$$

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 $\lambda_2 = 3$:

$$\begin{aligned} 0 &= A \cdot v - \lambda_2 \cdot v = (A - \lambda_2 E_2)v : \\ A - \lambda_2 \cdot E_2 &= \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} - 3 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 4 & -1 \\ 2 & 1 \end{pmatrix} - \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -1 \\ 2 & -2 \end{pmatrix} \end{aligned}$$

Anwendung des Gaußschen Algorithmus liefert:

$$\begin{array}{cc|cc} 1 & -1 & & 1 & -1 \\ 2 & -2 & \rightsquigarrow & 0 & 0 \end{array}$$

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 10.8. *Sei V ein \mathbb{K} -Vektorraum und $f: V \rightarrow V$ ein Endomorphismus. Dann gilt für einen Vektor $v \in V \setminus \{0\}$ und $\lambda \in \mathbb{K}$:*

$$\begin{aligned} v \text{ ist genau dann Eigenvektor von } f \text{ zum Eigenwert } \lambda, \\ \text{wenn } v \in \text{Kern}(f - \lambda \cdot \text{Id}_V). \end{aligned}$$

Beweis: v ist Eigenvektor von f zum Eigenwert λ

$$\Leftrightarrow f(v) = \lambda \cdot v$$

$$\Leftrightarrow 0 = f(v) - \lambda \cdot v = f(v) - (\lambda \cdot \text{Id}_V)(v) = (f - \lambda \cdot \text{Id}_V)(v)$$

$$\Leftrightarrow v \in \text{Kern}(f - \lambda \cdot \text{Id}_V) \quad \square$$

Beachte die Analogie zu der Umformung im letzten Beispiel 10.7.

Offenbar sind die obigen Kerne des verschobenen Endomorphismus A von Interesse, so dass wir diese gesondert bezeichnen:

Definition 10.9. *Sei V ein \mathbb{K} -Vektorraum und $f: V \rightarrow V$ ein Endomorphismus mit Eigenwert $\lambda \in \mathbb{K}$. Dann heißt*

$$E(\lambda) := \text{Kern}(f - \lambda \cdot \text{Id}_V)$$

der Eigenraum von f zum Eigenwert λ . Die geometrische Vielfachheit des Eigenwerts λ ist die Dimension von $E(\lambda)$.

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 10.3 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 10.10. *Sei V ein \mathbb{K} -Vektorraum und $f: V \rightarrow V$ ein Endomorphismus. Seien $\lambda_1, \dots, \lambda_r$ Eigenvektoren zu paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_r$. Dann ist (v_1, \dots, 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, \dots, v_r, v_{r+1} sind Eigenvektoren zu paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_r, \lambda_{r+1}$.

Ziel: v_1, \dots, v_r, v_{r+1} sind linear unabhängig.

Seien dazu (nach Definition) $\alpha_1, \dots, \alpha_r, \alpha_{r+1} \in \mathbb{K}$ mit

$$\alpha_1 v_1 + \dots + \alpha_r v_r + \alpha_{r+1} v_{r+1} = 0 \quad (*)$$

gegeben.

Noch zu zeigen: $\alpha_1 = \dots = \alpha_r = \alpha_{r+1} = 0$

Wende dazu f auf diese Gleichung an. Da v_1, \dots, v_r, v_{r+1} Eigenvektoren sind, ergibt sich

$$\begin{aligned} f(\alpha_1 v_1) + \dots + f(\alpha_r v_r) + f(\alpha_{r+1} v_{r+1}) &= f(0) \\ \Leftrightarrow \alpha_1 f(v_1) + \dots + \alpha_r f(v_r) + \alpha_{r+1} f(v_{r+1}) &= 0 \\ \Leftrightarrow \alpha_1 \lambda_1 v_1 + \dots + \alpha_r \lambda_r v_r + \alpha_{r+1} \lambda_{r+1} v_{r+1} &= 0 \quad (**) \end{aligned}$$

Multiplikation der ersten Gleichung (*) mit λ_{r+1} ergibt:

$$\alpha_1 \lambda_{r+1} v_1 + \dots + \alpha_r \lambda_{r+1} v_r + \alpha_{r+1} \lambda_{r+1} v_{r+1} = 0 \quad (***)$$

Subtraktion von (***) von (**) ergibt:

$$\alpha_1 (\lambda_1 - \lambda_{r+1}) v_1 + \dots + \alpha_r (\lambda_r - \lambda_{r+1}) v_r = 0$$

Nach Induktionsvoraussetzung sind die Vektoren v_1, \dots, v_r linear unabhängig. Also gilt:

$$\alpha_1 (\lambda_1 - \lambda_{r+1}) = \dots = \alpha_r (\lambda_r - \lambda_{r+1}) = 0$$

Da die Eigenwerte $\lambda_1, \dots, \lambda_r, \lambda_{r+1}$ paarweise verschieden sind, gilt

$$\lambda_1 - \lambda_{r+1} \neq 0, \dots, \lambda_r - \lambda_{r+1} \neq 0.$$

Somit folgt: $\alpha_1 = \dots = \alpha_r = 0$. Da v_{r+1} ein Eigenvektor ist, gilt $v_{r+1} \neq 0$ und somit wie gewünscht nach (*) auch $\alpha_{r+1} = 0$. Damit ist alles per Induktion bewiesen. \square

Aus dem letzten Satz folgt sofort:

Satz 10.11. *Sei V ein m -dimensionaler \mathbb{K} -Vektorraum und $f: V \rightarrow V$ ein Endomorphismus. Dann besitzt f höchstens m verschiedene Eigenwerte.*

Beweis: Bei $m + 1$ vielen Eigenwerten $\lambda_1, \dots, \lambda_{m+1}$ wähle jeweils dazugehörige Eigenvektoren v_1, \dots, v_{m+1} . Nach dem letzten Satz 10.10 ist v_1, \dots, v_{m+1} eine linear unabhängige Menge von Vektoren im m -dimensionalen Vektorraum. Dies ist ein Widerspruch. \square

Wir können sogar die Aussage über die lineare Unabhängigkeit wie folgt verstärken:

Satz 10.12. *Sei V ein m -dimensionaler \mathbb{K} -Vektorraum und $f: V \rightarrow V$ ein Endomorphismus. Seien $\lambda_1, \dots, \lambda_r$ paarweise verschiedene Eigenwerte von f mit geometrischen Vielfachheiten m_1, \dots, m_r . Für $i \leq r$ sei $(v_1^i, \dots, v_{m_i}^i)$ eine Basis des Eigenraums $E(\lambda_i)$. Dann gilt:*

(a) *Das System*

$$\left(\underbrace{(v_1^1, \dots, v_{m_1}^1)}_{\text{Basis } E(\lambda_1)}, \underbrace{(v_1^2, \dots, v_{m_2}^2)}_{\text{Basis } E(\lambda_2)}, \dots, \underbrace{(v_1^r, \dots, v_{m_r}^r)}_{\text{Basis } E(\lambda_r)} \right)$$

ist linear unabhängig.

(b) *Für die Summe der geometrischen Vielfachheiten gilt:*

$$m_1 + \dots + m_r \leq m$$

(c) *Der Endomorphismus f ist genau dann diagonalisierbar, wenn*

$$m_1 + \dots + m_r = m.$$

Beweis: zu (a). Wir betrachten eine Darstellung des Nullvektor und hierfür konkret Skalare $\alpha_j^i \in \mathbb{K}$, so dass

$$\alpha_1^1 v_1^1 + \dots + \alpha_{m_1}^1 v_{m_1}^1 + \dots + \alpha_1^r v_1^r + \dots + \alpha_{m_r}^r v_{m_r}^r = 0$$

Wir sehen, dass die Teilsummen

$$\alpha_1^i v_1^i + \dots + \alpha_{m_i}^i v_{m_i}^i$$

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 10.10. Also gilt für alle $i \leq r$, dass

$$\alpha_1^i = \dots = \alpha_{m_i}^i = 0.$$

Damit sind aber wie gewünscht alle Skalare $\alpha_j^i = 0$ und somit ist das System

$$(v_1^1, \dots, v_{m_1}^1, v_1^2, \dots, v_{m_2}^2, \dots, v_1^r, \dots, v_{m_r}^r)$$

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:

$$m_1 + \dots + m_r \leq m.$$

zu (c). Wir zeigen beide Richtungen der geforderten Äquivalenz. Angenommen, dass $m_1 + \dots + m_r = m$ gilt, dann ist

$$(v_1^1, \dots, v_{m_1}^1, v_1^2, \dots, v_{m_2}^2, \dots, v_1^r, \dots, v_{m_r}^r)$$

sogar eine Basis von V . Diese Basis besteht aus Eigenvektoren und nach Satz 10.3 ist f diagonalisierbar.

Umgekehrt sei nun f diagonalisierbar. Nach Satz 10.3 wähle eine Basis $C = (v_1, \dots, v_m)$ aus Eigenvektoren. Für $i \leq r$ sei n_i die Anzahl der Vektoren aus C , deren Eigenwerte λ_i ist. Dann ist $n_i \leq m_i$ und

$$m = n_1 + \dots + n_r \leq m_1 + \dots + m_r \leq m$$

und damit ist

$$m_1 + \dots + m_r = m.$$

□

Vorläufige Strategie des Diagonalisierens. Damit zeichnet sich eine Strategie für das Diagonalisieren eines Endomorphismus $f: V \rightarrow V$ 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:

$$\text{Kern}(f - \lambda_i \cdot \text{Id}_V) \neq \{0\}$$

Dann bestimme man eine Basis des Eigenraums $E(\lambda_i)$ durch Lösen des linearen Gleichungssystems

$$(f - \lambda_i \cdot \text{Id}_V)(x) = 0.$$

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.

11. DETERMINANTEN UND CHARAKTERISTISCHE POLYNOME

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 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

$$ad - bc \neq 0. \quad (*)$$

Wir benutzen nun dieses Kriterium in der folgenden einfachen Umformung: Sei also $A = \begin{pmatrix} e & f \\ g & h \end{pmatrix} \in \text{Mat}(2 \times 2, \mathbb{K})$ darstellende Matrix eines Endomorphismus f . Bei der Suche nach Eigenwerten sind die Abbildungen $A - \lambda \cdot \text{Id}_{\mathbb{K}^2}$ für $\lambda \in \mathbb{K}$ zu studieren. Nach der Dimensionsformel gilt:

$$\dim(\text{Kern}(A - \lambda \cdot \text{Id}_{\mathbb{K}^2})) + \text{Rg}(A - \lambda \cdot \text{Id}_{\mathbb{K}^2}) = 2$$

Damit gilt:

$$\begin{aligned} & \dim(\text{Kern}(A - \lambda \cdot \text{Id}_{\mathbb{K}^2})) > 0 \\ \iff & \text{Rg}(A - \lambda \cdot \text{Id}_{\mathbb{K}^2}) < 2, \\ \iff & A - \lambda \cdot \text{Id}_{\mathbb{K}^2} \text{ nicht invertierbar ist,} \\ \iff & \begin{pmatrix} e & f \\ g & h \end{pmatrix} - \lambda \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e - \lambda & f \\ g & h - \lambda \end{pmatrix} \text{ nicht invertierbar ist} \end{aligned}$$

$$\iff p_A(\lambda) := (e - \lambda)(h - \lambda) - gf = 0.$$

Das heißt, die Eigenwerte des durch die Matrix A dargestellten Endomorphismus sind die Nullstellen dieses quadratischen Polynoms $p_A(\lambda)$, welches als **charakteristische 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 11.1. Im Fall der Spiegelung

$$A = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ \sin(\varphi) & -\cos(\varphi) \end{pmatrix}$$

lautet das charakteristische Polynom

$$\begin{aligned} p_A(\lambda) &= (\cos(\varphi) - \lambda)(-\cos(\varphi) - \lambda) - \sin^2(\varphi) \\ &= -\cos^2(\varphi) + \lambda^2 - \sin^2(\varphi) \\ &= \lambda^2 - (\sin^2(\varphi) + \cos^2(\varphi)) \\ &= \lambda^2 - 1 = (\lambda + 1)(\lambda - 1) \end{aligned}$$

Damit hat das Polynom $p(\lambda)$ die Nullstellen 1 und -1 . Da die Dimensionen der Eigenräume $E(1)$ und $E(-1)$ jeweils mindestens 1 sind, ist die geometrische Vielfachheit der Eigenräume jeweils gleich (klar!). Da $1 + 1 = 2$ ist, ist die Matrix A nach Satz 10.12 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 mit Hilfe des Invertierkriteriums $ad - bc \neq 0$ definiert. Der Term $ad - bc$ wird sich als Determinante der Matrix herausstellen. Ziel ist es die Determinante für beliebige Matrizen zu definieren.

Definition 11.2. Sei $A = (a_{ij}) \in \text{Mat}((m+1) \times (m+1), \mathbb{K})$ und $k, l \leq m$. Dann definiere $A_{kl} = (\bar{a}_{ij}) \in \text{Mat}(m \times m, \mathbb{K})$ aus A durch Weglassen der k -ten Zeile und l -ten Spalte:

$$\bar{a}_{ij} := \begin{cases} a_{ij} & \text{falls } i < k \text{ und } j < l \\ a_{i+1j} & \text{falls } k \leq i < m \text{ und } j < l \\ a_{ij+1} & \text{falls } i < k \text{ und } l \leq j < m \\ a_{i+1j+1} & \text{falls } k \leq i < m \text{ und } l \leq j < m \end{cases}$$

Skizzenhaft heißt dies, dass wir die entsprechende Zeile und Spalte löschen:

$$A_{kl} = \begin{pmatrix} a_{11} & \dots & a_{1l} & \dots & a_{1m+1} \\ \vdots & & \vdots & & \vdots \\ a_{k1} & \dots & a_{kl} & \dots & a_{km+1} \\ \vdots & & \vdots & & \vdots \\ a_{m+11} & \dots & a_{m+1l} & \dots & a_{m+1m+1} \end{pmatrix}$$

O-DO!

[Boxen setzen]

Nach diesem technischen Hilfsmittel folgt nun die eigentliche

Definition 11.3 (Laplacescher Entwicklungssatz). *Sei \mathbb{K} ein Körper. Definiere die Determinantenfunktion*

$$\det: \text{Mat}(m \times m, \mathbb{K}) \rightarrow \mathbb{K}$$

für $m \in \mathbb{N} \setminus \{0\}$ durch Rekursion über m :

(a) Für $A = (a) \in \text{Mat}(1 \times 1, \mathbb{K})$ sei $\det(A) = \det A := a$

(b) Angenommen $\det: \text{Mat}(m \times m, \mathbb{K}) \rightarrow \mathbb{K}$ sei bereits definiert. Dann setze die gesuchte Abbildung $\det: \text{Mat}((m+1) \times (m+1), \mathbb{K}) \rightarrow \mathbb{K}$ an wie folgt:

Für $A = (a_{ij}) \in \text{Mat}(m+1 \times m+1, \mathbb{K})$ sei $\det(A) := \det A$ die folgende alternierende Summe:

$$\begin{aligned} \det(A) &= a_{11} \det A_{11} - a_{21} \det A_{21} + a_{31} \det A_{31} - \dots \pm a_{m+1,1} \det A_{m+1,1} \\ &= \sum_{k=1}^{m+1} (-1)^{k+1} \cdot a_{k1} \det A_{k1} \end{aligned}$$

Man schreibt oft abkürzend auch $\begin{vmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mm} \end{vmatrix}$ statt $\det \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mm} \end{pmatrix}$.

Beispiel 11.4. 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 Merkmregel wird auch als die *Regel von Sarrus* bezeichnet.

Das gleiche Ergebnis erhalten wir, wenn wir die ersten beiden Spalten rechts hinzufügt $\begin{pmatrix} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{pmatrix}$. Diese Regel gilt allerdings nur für die Dimensionen $m = 2$

und $m = 3$. Für $m = 2$ gilt erwartungsgemäss: $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$.

Strategie für $m = 4$: 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 11.5. 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:

$$\begin{array}{cccc} + & - & + & - & \dots \\ - & + & - & + & \dots \\ + & - & + & - & \dots \\ - & + & - & + & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{array}$$

Betrachten Sie dazu das folgende

Beispiel 11.6. Betrachte $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 0 & 1 \\ 2 & 0 & 3 \end{pmatrix}$ und das dazugehörige Schachbrettmuster

$$\begin{array}{ccc} + & - & + \\ - & + & - \\ + & - & + \end{array}$$

. 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.

Determinanten 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 Determinanten und der Invertierbarkeit einer Matrix (siehe Satz 11.10) nachweisen.

Wir starten mit einfachen Umformungen von benachbarten Zeilen.

Satz 11.7. *Sei $A \in \text{Mat}(m \times m, \mathbb{K})$. 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 darauf folgenden Zeilen 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:

$$\det A = \det(a' + a'') = \det A' + \det A''.$$

Zu (d): Klar!

Induktionsschritt: Wir nehmen an, dass die Behauptungen (a)–(d) für m gelten. Betrachte $A = (a_{ij}) \in \text{Mat}(m+1 \times m+1, \mathbb{K})$.

Zu (a): Betrachte $k \leq m$. Die Matrix $A' = (a'_{ij}) \in \text{Mat}(m+1 \times m+1, \mathbb{K})$ entstehe aus A durch Vertauschen der Zeilen k und $k+1$. Dann gilt:

$$\begin{aligned}
 \det A' &= \sum_{i=1}^{m+1} (-1)^{i+1} a'_{i,1} \det A'_{i,1} \\
 &= (-1)^{k+1} a'_{k,1} \det A'_{k,1} + (-1)^{k+2} a'_{k+1,1} \det A'_{k+1,1} + \sum_{i=1, i \neq k, i \neq k+1}^{m+1} (-1)^{i+1} a'_{i,1} \det A'_{i,1} \\
 &= (-1)^{k+1} a_{k+1,1} \det A'_{k+1,1} + (-1)^{k+2} a_{k,1} \det A'_{k,1} + \sum_{i=1, i \neq k, i \neq k+1}^{m+1} (-1)^{i+1} a_{i,1} \det A'_{i,1} \\
 &= -(-1)^{k+2} a_{k+1,1} \det A_{k+1,1} + (-1)^{k+2} a_{k,1} \det A_{k,1} - \sum_{i=1, i \neq k, i \neq k+1}^{m+1} (-1)^{i+1} a_{i,1} \det A_{i,1} \\
 &= - \sum_{i=1}^{m+1} (-1)^{i+1} a_{i,1} \det A_{i,1} = -\det A
 \end{aligned}$$

Zur Begründung der Umformung beachte man, dass

- $a'_{k,1} = a_{k+1,1}$ und $a'_{k+1,1} = a_{k,1}$
- für $i \neq k, k+1$ ist $a'_{i,1} = a_{i,1}$
- $A'_{k,1} = A_{k+1,1}$ und $A'_{k+1,1} = A_{k,1}$
- für $i \neq k, k+1$ entsteht $A'_{i,1}$ aus $A_{i,1}$ durch Vertauschen benachbarter Zeilen, nach Induktionsannahme ist dann $\det A'_{i,1} = -\det A_{i,1}$

Zu (b): Betrachte $k \leq m$. Die Matrix $A' = (a'_{i,j}) \in \text{Mat}(m+1 \times m+1, \mathbb{K})$ entstehe aus A durch Multiplikation der k -ten Zeile mit dem Faktor $\lambda \in \mathbb{K}$. Dann gilt:

$$\begin{aligned}
 \det A' &= \sum_{i=1}^{m+1} (-1)^{i+1} a'_{i,1} \det A'_{i,1} \\
 &= (-1)^{k+1} a'_{k,1} \det A'_{k,1} + \sum_{i=1, i \neq k}^{m+1} (-1)^{i+1} a'_{i,1} \det A'_{i,1} \\
 &= (-1)^{k+1} \lambda a_{k,1} \det A_{k,1} + \sum_{i=1, i \neq k}^{m+1} (-1)^{i+1} a_{i,1} \lambda \det A_{i,1} \\
 &= \lambda \cdot \sum_{i=1}^{m+1} (-1)^{i+1} a_{i,1} \det A_{i,1} \\
 &= \lambda \cdot \det A
 \end{aligned}$$

Zur Begründung der Umformung beachte man, dass

- $a'_{k,1} = \lambda a_{k,1}$
- für $i \neq k$ ist $a'_{i,1} = a_{i,1}$
- $A'_{k,1} = A_{k,1}$
- für $i \neq k$ entsteht $A'_{i,1}$ aus $A_{i,1}$ durch skalare Multiplikation einer Zeile mit dem Faktor λ , nach Induktionsannahme ist dann $\det A'_{i,1} = \lambda \cdot \det A_{i,1}$

Zu (c): es seien $A = (a_{ij}), A' = (a'_{ij}), A'' = (a''_{ij}) \in \text{Mat}(m+1 \times m+1, \mathbb{K})$ wie angegeben.

Dann gilt:

$$\begin{aligned}
 \det A &= \sum_{i=1}^{m+1} (-1)^{i+1} a_{i,1} \det A_{i,1} \\
 &= (-1)^{k+1} a_{k,1} \det A_{k,1} + \sum_{i=1, i \neq k}^{m+1} (-1)^{i+1} a_{i,1} \det A_{i,1} \\
 &= (-1)^{k+1} (a'_{k,1} + a''_{k+1,1}) \det A_{k,1} + \sum_{i=1, i \neq k}^{m+1} (-1)^{i+1} a_{i,1} (\det A'_{i,1} + \det A''_{i,1}) \\
 &= (-1)^{k+1} a'_{k,1} \det A'_{k,1} + \sum_{i=1, i \neq k}^{m+1} (-1)^{i+1} a_{i,1} \det A'_{i,1} \\
 &\quad + (-1)^{k+1} a''_{k+1,1} \det A''_{k,1} + \sum_{i=1, i \neq k}^{m+1} (-1)^{i+1} a''_{i,1} \det A''_{i,1} \\
 &= \sum_{i=1}^{m+1} (-1)^{i+1} a'_{i,1} \det A'_{i,1} + \sum_{i=1}^{m+1} (-1)^{i+1} a''_{i,1} \det A''_{i,1} \\
 &= \det A' + \det A''
 \end{aligned}$$

Zur Begründung der Umformung beachte man, dass

- $a_{k,1} = a'_{k,1} + a''_{k+1,1}$
- für $i \neq k$ entsteht $A_{i,1}$ aus $A'_{i,1}$ und $A''_{i,1}$ durch Addieren einer Zeile und Kopieren der übrigen Zeilen, nach Induktionsannahme ist dann $\det A_{i,1} = \det A'_{i,1} + \det A''_{i,1}$
- Wegen der Übereinstimmung in den Zeilen von A, A' und A'' ist $A_{k,1} = A'_{k,1} = A''_{k,1}$
- für $i \neq k$ ist $a_{i,1} = a'_{i,1} = a''_{i,1}$

Zu (d): Betrachte $k \leq m$ und $\lambda \in \mathbb{K}$. Die Matrix $A' = (a'_{ij}) \in \text{Mat}(m+1 \times m+1, \mathbb{K})$ entstehe aus A durch Addition des λ -fachen der Zeile k zur Zeile $k+1$. Dann gilt:

$$\begin{aligned}
 \det A' &= \sum_{i=1}^{m+1} (-1)^{i+1} a'_{i,1} \det A'_{i,1} \\
 &= (-1)^{k+1} a'_{k,1} \det A'_{k,1} + (-1)^{k+2} a'_{k+1,1} \det A'_{k+1,1} + \sum_{i=1, i \neq k, i \neq k+1}^{m+1} (-1)^{i+1} a'_{i,1} \det A'_{i,1} \\
 &= (-1)^{k+1} a_{k,1} \det A'_{k,1} + (-1)^{k+2} (a_{k+1,1} + \lambda a_{k,1}) \det A_{k+1,1} \\
 &\quad + \sum_{i=1, i \neq k, i \neq k+1}^{m+1} (-1)^{i+1} a_{i,1} \det A'_{i,1} \\
 &= (-1)^{k+1} a_{k,1} (\det A_{k,1} + \lambda \det A_{k+1,1}) \\
 &\quad + (-1)^{k+2} a_{k,1} \det A_{k,1} + (-1)^{k+2} (a_{k+1,1} + \lambda a_{k,1}) \det A_{k+1,1}
 \end{aligned}$$

$$\begin{aligned}
& + \sum_{i=1, i \neq k, i \neq k+1}^{m+1} (-1)^{i+1} a_{i,1} \det A_{i,1} + \sum_{i=1, i \neq k, i \neq k+1}^{m+1} (-1)^{i+1} a_{i,1} \det A_{i,1} \\
& = \sum_{i=1}^{m+1} (-1)^{i+1} a_{i,1} \det A_{i,1} = \det A
\end{aligned}$$

Zur Begründung der Umformung beachte man, dass

- $a'_{k,1} = a_{k+1,1} + \lambda a_{k,1}$
- für $i \neq k$ ist $a'_{i,1} = a_{i,1}$
- $A'_{k+1,1} = A_{k+1,1}$
- für $i \neq k, k+1$ entsteht $A'_{i,1}$ aus $A_{i,1}$ durch addition des λ -fachen einer Zeile, nach Induktionsannahme ist dann $\det A'_{i,1} = \det A_{i,1}$
- die k -te Zeile von $A'_{k,1}$ ist die Summe der k -te Zeile von $A_{k,1}$ und dem λ -fachen der k -ten Zeile von $A_{k+1,1}$, die übrigen Zeilen von $A'_{k,1}, A_{k,1}$ und $A_{k+1,1}$ stimmen überein. Nach der in (b) und (c) formulierten Zeilenlinearität ist $\det A'_{k+1} = \det A_{k,1} + \lambda A_{k+1,1}$

Damit ist der Satz nach dem Prinzip der vollständigen Induktion bewiesen. \square

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:

$$\begin{array}{cccccccc}
\dots & k & k+1 & k+2 & \dots & l-2 & l-1 & l & \dots \\
\dots & k+1 & k & k+2 & \dots & l-2 & l-1 & l & \dots \\
\dots & k+1 & k+2 & k & \dots & l-2 & l-1 & l & \dots \\
& & & & \ddots & & & & \\
\dots & k+1 & k+2 & k+3 & \dots & k & l-1 & l & \dots \\
\dots & k+1 & k+2 & k+3 & \dots & l-1 & k & l & \dots \\
\dots & k+1 & k+2 & k+3 & \dots & l-1 & l & k & \dots \\
\dots & k+1 & k+2 & k+3 & \dots & l & l-1 & k & \dots \\
& & & & \ddots & & & & \\
\dots & k+1 & k+2 & l & \dots & l-2 & l-1 & k & \dots \\
\dots & k+1 & l & k+2 & \dots & l-2 & l-1 & k & \dots \\
\dots & l & k+1 & k+2 & \dots & l-2 & l-1 & k & \dots
\end{array}$$

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 bestet dei 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 11.7 Teil (d) vorliegt. Satz 11.7 impliziert die Invarianz der Determinante. Anschliessend tauscht man die vertauschten Zeilen zurück. Die hierdurch verursachten Vorzeichenwechsel heben sich auf. Damit gilt der folgende

Satz 11.8. *Sei $A \in \text{Mat}(m \times m, \mathbb{K})$. 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 $\det A' = \det A$.*

Schliesslich erhalten wir ebenfalls den folgenden

Satz 11.9. *Sei $A = (a_{ij}) \in \text{Mat}(m \times m, \mathbb{K})$ eine Matrix in Zeilenstufenform. Dann ist $\det A = \prod_{i=1}^m a_{ij}$. Insbesondere gilt für die m -dimensionale Einheitsmatrix $E_m \in \text{Mat}(m \times m, \mathbb{K})$ stets $\det E_m = 1$.*

Beweis: 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

$$\det A = a_{11} \cdot \det A_{11} = a_{11} \cdot \prod_{i=2}^{m+1} a_{ii} = \prod_{i=1}^{m+1} a_{ii}$$

□

Wir erhalten nun die für die Theorie der Eigenwerte und Eigenvektoren wichtigste Eigenschaft von Matrizen:

Satz 11.10. *Sei \mathbb{K} ein Körper, $m \in \mathbb{N} \setminus \{0\}$ und $A \in \text{Mat}(m \times m, \mathbb{K})$. Dann ist A invertierbar genau dann, wenn $\det A \neq 0$.*

Beweis: Überführe $A = (a_{ij})$ durch den Gauß-Algorithmus mittels Folge von elementaren Zeilenumformungen in eine Matrix $A' = (a'_{ij})$ in Zeilenstufenform. Nach den Sätzen 11.7 und 11.8 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:

$$\begin{aligned} A \text{ ist invertierbar genau dann, wenn } \prod_{i=1}^m a'_{ii} &\neq 0 \\ &\text{genau dann, wenn } \det A' \neq 0 \end{aligned}$$

genau dann, wenn $\det A \neq 0$

☒

Bemerkung 11.11. 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 11.12. Wir berechnen die Determinante mittels Gauß-Umformungen:

$$\begin{aligned} \begin{vmatrix} 2 & 0 & -1 & -4 \\ -3 & 1 & 3 & 0 \\ 2 & 0 & -1 & -2 \\ 1 & 0 & -1 & -1 \end{vmatrix} &= - \begin{vmatrix} 1 & 0 & -1 & -3 \\ -3 & 1 & 3 & 0 \\ 2 & 0 & -1 & -2 \\ 2 & 0 & -1 & -4 \end{vmatrix} = - \begin{vmatrix} 1 & 0 & -1 & -3 \\ 0 & 1 & 0 & -9 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 1 & 2 \end{vmatrix} \\ &= - \begin{vmatrix} 1 & 0 & -1 & -3 \\ 0 & 1 & 0 & -9 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & -2 \end{vmatrix} = -(1 \cdot 1 \cdot 1 \cdot (-2)) = 2 \end{aligned}$$

Dies stimmt mit dem Ergebnis überein, wenn wir nach der zweiten Spalte entwickeln:

$$\begin{aligned} \begin{vmatrix} 2 & 0 & -1 & -4 \\ -3 & 1 & 3 & 0 \\ 2 & 0 & -1 & -2 \\ 1 & 0 & -1 & -1 \end{vmatrix} &= 1 \cdot \begin{vmatrix} 2 & -1 & -4 \\ 2 & -1 & -2 \\ 1 & -1 & -1 \end{vmatrix} \\ &= 2 \cdot (-1) \cdot (-1) + (-1) \cdot (-2) \cdot 1 + (-4) \cdot 2 \cdot (-1) \\ &\quad - (-4) \cdot (-1) \cdot 1 - (-2) \cdot (-1) \cdot 2 - (-1) \cdot (-1) \cdot 2 \\ &= 2 + 2 + 8 - 4 - 4 - 2 = 2 \end{aligned}$$

* * *

Folgender Satz faßt noch einmal den übergreifenden Zusammenhang der verschiedenen Begriffe zusammen – staunen Sie selbst:

Satz 11.13. Für eine Matrix $A \in \text{Mat}(n \times n, \mathbb{K})$ 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 eine für eine beliebige rechte Seite b eine eindeutige Lösung.

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.

* * *

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 $x_1 = \frac{D_1}{D}, \dots, x_n = \frac{D_n}{D}$, mit $D := \det A$ und $D_i = \det A_i$ für $i = 1, \dots, n$, 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 11.14. Gegeben sei das lineare Gleichungssystem $Ax = b$ mit

$$A = \begin{pmatrix} 2 & 3 & 1 \\ 0 & 4 & 2 \\ 1 & 0 & -2 \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad \text{und} \quad b = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Berechne $D = \det A = \begin{vmatrix} 2 & 3 & 1 \\ 0 & 4 & 2 \\ 1 & 0 & -2 \end{vmatrix} = -14$,

$$D_1 = \begin{vmatrix} 1 & 3 & 1 \\ 0 & 4 & 2 \\ 1 & 0 & -2 \end{vmatrix} = -6, \quad D_2 = \begin{vmatrix} 2 & 1 & 1 \\ 0 & 0 & 2 \\ 1 & 1 & -2 \end{vmatrix} = -2 \quad \text{und} \quad D_3 = \begin{vmatrix} 2 & 3 & 1 \\ 0 & 4 & 0 \\ 1 & 0 & 1 \end{vmatrix} = 4.$$

Somit gilt:

$$x_1 = \frac{D_1}{D} = \frac{3}{7}, \quad x_2 = \frac{D_2}{D} = \frac{1}{7}, \quad x_3 = \frac{D_3}{D} = -\frac{2}{7}$$

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 Gleichungssystem darstellt.

Charakteristische Polynome und Eigenwerte. Nachdem wir nun wissen, wie wir Determinanten berechnen können, kommen wir zurück zum eigentlichen Thema: die Bestimmung der Eigenwerte einer gegebenen Matrix.

Definition 11.15. Sei \mathbb{K} ein Körper und $m \in \mathbb{N} \setminus \{0\}$. Sei $A = (a_{ij}) \in \text{Mat}(m \times m, \mathbb{K})$. 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 11.16. Sei $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$. Das charakteristische Polynom von A ist:

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda E_3) = \begin{vmatrix} 1 - \lambda & 1 & 0 \\ 0 & -1 - \lambda & 1 \\ 0 & 1 & -1 - \lambda \end{vmatrix} \\ &= (1 - \lambda) \cdot \begin{vmatrix} -1 - \lambda & 1 \\ 1 & -1 - \lambda \end{vmatrix} \quad \text{Entwicklung nach erster Spalte} \\ &= (1 - \lambda)((-1 - \lambda) \cdot (-1 - \lambda) - 1) = (1 - \lambda)(1 + 2\lambda + \lambda^2 - 1) \\ &= (1 - \lambda)(2\lambda + \lambda^2) = (1 - \lambda)\lambda(2 + \lambda) = -\lambda^3 - \lambda^2 + 2\lambda \end{aligned}$$

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 11.17. *Sei \mathbb{K} ein Körper und $m \in \mathbb{N} \setminus \{0\}$. Sei $A = (a_{ij}) \in \text{Mat}(m \times m, \mathbb{K})$. Dann ist das charakteristische Polynom von A von der Gestalt*

$$p_A(\lambda) = (-1)^m \lambda^m + c_{m-1} \lambda^{m-1} + \dots + c_1 \lambda + c_0.$$

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 notwendige verschiedene) Nullstellen.

Mit dem folgenden Satz sind wir schließlich am Ziel unserer Betrachtungen:

Satz 11.18. *Sei $f: V \rightarrow V$ ein Endomorphismus eines m -dimensionalen \mathbb{K} -Vektorraums V , $m \in \mathbb{N} \setminus \{0\}$. Sei B eine Basis von V und $A = \text{DM}_B(f)$. 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:

- λ ist Eigenwert von f
- $\iff \text{Kern}(f - \lambda \text{Id}_V) \neq \{0\}$
- $\iff \dim \text{Kern}(f - \lambda \text{Id}_V) > 0$
- $\iff \text{Rg}(f - \lambda \text{Id}_V) < m$
- $\iff f - \lambda \text{Id}_V$ ist nicht invertierbar
- $\iff \text{DM}_B(f - \lambda \text{Id}_V) = \text{DM}_B(f) - \lambda \text{DM}_B(\text{Id}_V) = A - \lambda E_m$ ist nicht invertierbar
- $\iff p_A(\lambda) = \det(A - \lambda E_m) = 0$ (nach Satz 11.10)

□

Beispiel 11.19. Wir setzen das Beispiel 11.16 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 11.18 sind dies die Eigenwerte von A . Wir bestimmen jetzt die zugehörigen Eigenvektoren durch lösen des homogenen linearen Gleichungs-

systems $(A - \lambda_i E_3) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$:

$\lambda_1 = -2$ 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: $y = -z$

und $3x = -y$, das heißt $x = -\frac{1}{3}y = \frac{1}{3}z$. Wähle $z = 3$ und wir erhalten den

Eigenvektor $u_{-2} = \begin{pmatrix} 1 \\ -3 \\ 3 \end{pmatrix}$.

$\lambda_2 = 0$ Betrachte $A - 0 \cdot E_3 = A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$. Durch den Gauß-Algorithmus er-

halten 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 $u_0 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$.

$\lambda_3 = 1$ 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_B^C , für B die kanonische Basis, ist:

$$T_B^C = \begin{pmatrix} 1 & -1 & 1 \\ -3 & 1 & 0 \\ 3 & 1 & 0 \end{pmatrix}$$

Die Transformationsmatrix T_C^B erhalten wir durch invertieren von T_B^C :

$$\begin{array}{cccccc} 1 & -1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & -\frac{1}{6} & \frac{1}{6} \\ -3 & 1 & 0 & 0 & 1 & 0 & \xrightarrow{\text{Gauß}} & \dots & \xrightarrow{\text{Gauß}} & 0 & 1 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 3 & 1 & 0 & 0 & 0 & 1 & & & & 0 & 0 & 1 & 1 & \frac{2}{3} & \frac{1}{3} \end{array}$$

das heißt: $T_C^B = \begin{pmatrix} 0 & -\frac{1}{6} & \frac{1}{6} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & \frac{2}{3} & \frac{1}{3} \end{pmatrix}$

Die darstellende Matrix von A bezüglich der Basis aus Eigenvektoren hat dann Diagonalform:

$$\begin{aligned} DM_C(A) &= T_C^B \cdot A \cdot T_B^C = \begin{pmatrix} 0 & -\frac{1}{6} & \frac{1}{6} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & \frac{2}{3} & \frac{1}{3} \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 & 1 \\ -3 & 1 & 0 \\ 3 & 1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -\frac{1}{6} & \frac{1}{6} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & \frac{2}{3} & \frac{1}{3} \end{pmatrix} \cdot \begin{pmatrix} -2 & 0 & 1 \\ 6 & 0 & 0 \\ -6 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} -2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

Wir sehen, dass wie erwartet in $DM_C(A)$ die Eigenwerte auf der Diagonalen stehen.

Diagonalisieren an Beispielen. Jetzt sind wir in der Lage ein vollständiges Verfahren ein Einzelschritten zum Diagonalisieren einer Matrix A bzw. eines Endomorphismus $f : V \rightarrow V$ anzugeben.

Algorithmus zur Diagonalisierung

- (0) Wählen einer Basis B von V und Darstellung des Endomorphismus f als $A = DM_B(f)$.
- (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 10.9 eingeführt. Nach den Überlegungen aus Kapitel 10 wissen wir bereits, dass die geometrischen Vielfachheiten höchstens gleich der algebraischen, mindestens jedoch positiv sein müssen, d.h. es gilt:

$$1 \leq \text{geometrische_Vielfachheit}(\lambda) \leq \text{algebraische_Vielfachheit}(\lambda).$$

* * *

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 $\text{Mat}(3 \times 3, \mathbb{R})$:

$$A_1 = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad A_2 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 0 & -1 & -2 \\ 2 & 3 & 2 \\ 1 & 1 & 3 \end{pmatrix} \quad \text{und den Endomorphismus}$$

$$f: \mathbb{R}^3 \rightarrow \mathbb{R}^3 \text{ gegeben durch } \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} x_1 - 4x_2 + 8x_3 \\ -4x_1 + 7x_2 + 4x_3 \\ 8x_1 + 4x_2 + x_3 \end{pmatrix}.$$

Zu (0) Die ersten drei Endomorphismen haben wir bereits in Matrixform A_i gegeben. Für den vierten Endomorphismus bestimmen wir $\text{DM}_K(f)$ bezüglich der kanonischen Basis K des \mathbb{R}^3 :

$$f\left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 1 \\ -4 \\ 8 \end{pmatrix}, \quad f\left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} -4 \\ 7 \\ 4 \end{pmatrix} \quad \text{und} \quad f\left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right) = \begin{pmatrix} 8 \\ 4 \\ 1 \end{pmatrix}. \quad \text{Daraus ergibt sich}$$

$$A_4 := \text{DM}_K(f) = \begin{pmatrix} 1 & -4 & 8 \\ -4 & 7 & 4 \\ 8 & 4 & 1 \end{pmatrix}.$$

Zu (1)

$$\begin{aligned} p_{A_1}(\lambda) &= \det(A_1 - \lambda E_3) = \begin{vmatrix} 3-\lambda & 0 & 0 \\ 0 & -\lambda & -1 \\ 0 & 1 & -\lambda \end{vmatrix} = (3-\lambda) \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} \\ &= (3-\lambda)(\lambda^2 + 1) \end{aligned}$$

$$\begin{aligned}
 p_{A_2}(\lambda) &= \det(A_2 - \lambda E_3) = \begin{vmatrix} -1 - \lambda & 0 & 0 \\ 0 & 1 - \lambda & 1 \\ 0 & & 1 - \lambda \end{vmatrix} \\
 &= (-1 - \lambda) \begin{vmatrix} 1 - \lambda & 1 \\ 0 & 1 - \lambda \end{vmatrix} = (-1 - \lambda)(1 - \lambda)^2
 \end{aligned}$$

$$p_{A_3}(\lambda) = \det(A_3 - \lambda E_3) = \begin{vmatrix} -\lambda & -1 & -2 \\ 2 & 3 - \lambda & 2 \\ 1 & 1 & 3 - \lambda \end{vmatrix} = (1 - \lambda)(2 - \lambda)(3 - \lambda)$$

$$p_{A_4}(\lambda) = \det(A_4 - \lambda E_3) = \begin{vmatrix} 1 - \lambda & -4 & 8 \\ -4 & 7 - \lambda & 4 \\ 8 & 4 & 1 - \lambda \end{vmatrix} = (-9 - \lambda)(9 - \lambda)^2$$

Zu (2) Matrix \mathbf{A}_1 besteht den Test nicht, da $(\lambda^2 + 1)$ 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).

$\lambda_1 = -1$: Löse das homogene lineare Gleichungssystem

$$(A_2 + 1 \cdot E_3)x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix} \cdot x = 0. \text{ Der Vektor } \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \text{ bildet eine Basis des}$$

Eigenraums.

$\lambda_2 = 1$: Löse das homogene lineare Gleichungssystem

$$(A_2 - 1 \cdot E_3)x = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \cdot x = 0. \text{ Der Vektor } \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ bildet eine Basis des}$$

Eigenraums.

Die Matrix \mathbf{A}_3 hat die Eigenwerte 1 , 2 und 3 (jeweils mit algebraischer Vielfachheit 1).

$\lambda_1 = 1$: Löse das homogene lineare Gleichungssystem

$$(A_3 - 1 \cdot E_3)x = \begin{pmatrix} -1 & -1 & -2 \\ 2 & 2 & 2 \\ 1 & 1 & 2 \end{pmatrix} \cdot x = 0. \text{ Der Vektor } \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \text{ bildet eine Basis des}$$

Eigenraums.

$\lambda_2 = 2$: Löse das homogene lineare Gleichungssystem

$$(A_3 - 2 \cdot E_3)x = \begin{pmatrix} -2 & -1 & -2 \\ 2 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix} \cdot x = 0. \text{ Der Vektor } \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \text{ bildet eine Basis des}$$

Eigenraums.

$\lambda_3 = 3$: Löse das homogene lineare Gleichungssystem

$$(A_3 - 3 \cdot E_3)x = \begin{pmatrix} -3 & -1 & -2 \\ 2 & 0 & 2 \\ 1 & 1 & 0 \end{pmatrix} \cdot x = 0. \text{ Der Vektor } \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \text{ bildet eine Basis des}$$

Eigenraums.

Die Matrix A_4 hat die Eigenwerte -9 (mit algebraischer Vielfachheit 1) und 9 (mit algebraischer Vielfachheit 2).

$\lambda_1 = -9$: Löse das homogene lineare Gleichungssystem

$$(A_4 + 9 \cdot E_3)x = \begin{pmatrix} 10 & -4 & 8 \\ -4 & 16 & 4 \\ 8 & 4 & 10 \end{pmatrix} \cdot x = 0. \text{ Der Vektor } \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} \text{ bildet eine Basis des}$$

Eigenraums.

$\lambda_2 = 9$: Löse das homogene lineare Gleichungssystem

$$(A_4 - 9 \cdot E_3)x = \begin{pmatrix} -8 & -4 & 8 \\ -4 & -2 & 4 \\ 8 & 4 & -8 \end{pmatrix} \cdot x = 0. \text{ Die Vektoren } \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \text{ und } \begin{pmatrix} -2 \\ 2 \\ -1 \end{pmatrix} \text{ bilden}$$

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 $(0 \ 1 \ 0)$. Die Matrizen A_3 und A_4 bestehen den Test.

Zu (5) Für A_3 haben wir $C_3 = \left(\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right)$.

Für A_4 haben wir $C_4 = \left(\begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 2 \\ -1 \end{pmatrix} \right)$.

Zu (6) In C_3 und C_4 sind die Vektoren bezüglich der kanonischen Basis gegeben.

Wir definieren $T_3 := T_K^{C_3} = \begin{pmatrix} 1 & 1 & -1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix}$ und $T_4 := T_K^{C_4} = \begin{pmatrix} 2 & 1 & -2 \\ 1 & 2 & 2 \\ -2 & 2 & -1 \end{pmatrix}$.

Damit erhalten wir wie gewünscht:

$$\begin{aligned} T_3^{-1} \cdot A_3 \cdot T_3 &= \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & -1 & -2 \\ 2 & 3 & 2 \\ 1 & 1 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 & -1 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} T_4^{-1} \cdot A_4 \cdot T_4 &= \frac{1}{9} \begin{pmatrix} 2 & 1 & -2 \\ 1 & 2 & 2 \\ -2 & 2 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -4 & 8 \\ -4 & 7 & 4 \\ 8 & 4 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 & -2 \\ 1 & 2 & 2 \\ -2 & 2 & -1 \end{pmatrix} \\ &= \begin{pmatrix} -9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix} \end{aligned}$$

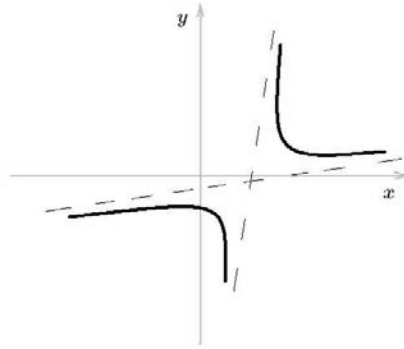
Hiermit sind wir am Ende der Betrachtungen rund um das Diagonalisieren angekommen. Sie sehen, dass wir für diesen Zweck das gesamte Wissen der bisherigen Vorlesung benötigen.

Ausblick Hauptachsentransformation. Abschließend möchte ich Ihnen noch einen kleinen Ausblick geben, was noch alles möglich ist, wir aber an dieser Stelle zietlich 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:

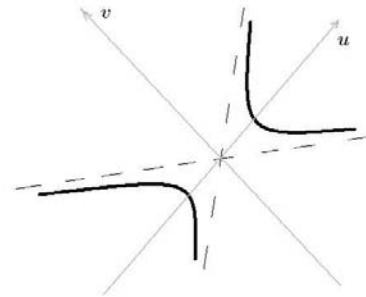
$$-\frac{7}{2}x^2 - \frac{7}{2}y^2 + 25xy + 30\sqrt{2}x - 66\sqrt{2}y = 256$$

Diese beschriebene Punktmenge der Ebene stellt eine Kurve dar. Genauer ausgedrückt handelt es sich um eine Hyperbel und im gewohnten $x-y$ -Koordinatensystem hat sie etwa die Gestalt:



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):



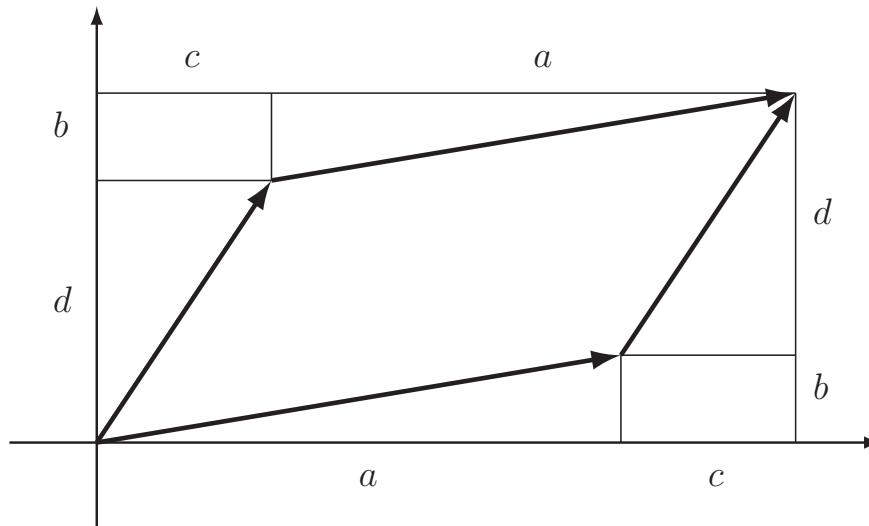
Mithilfe der Eigenwerttheorie ist es möglich, in diesem Fall von der kanonischen Basis zur Basis $\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}\right\}$ zu kommen. Unter dieser Hauptachse, die ich jetzt nicht mehr x und y , sondern u und v nenne, hat die Kurve die Gestalt:

$$9u^2 - 16v^2 = 144$$

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.

Determinanten und Volumina. Im m -dimensionalen euklidischen Raum \mathbb{R}^m lassen sich mit der Determinantenfunktion Volumina von parallelogramm-förmigen

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

$$F = \left\{ \lambda_0 \begin{pmatrix} a \\ b \end{pmatrix} + \lambda_1 \begin{pmatrix} c \\ d \end{pmatrix} \mid 0 \leq \lambda_0, \lambda_1 \leq 1 \right\}$$

auf. der Flächeninhalt $|F|$ von F ergibt sich als Differenz des äußeren Rechtecks und der rechtwinkligen Dreiecke und Rechtecke zwischen äußeren Rechteck und Parallelogramm. Zur Vermeidung von Vorzeichenfragen nehmen wir an, dass $a, b, c, d \geq 0$ sind, so dass wir erhalten:

$$\begin{aligned} |F| &= (a+c)(b+d) - \frac{1}{2}ab - cb - \frac{1}{2}cd - \frac{1}{2}ab - cb - \frac{1}{2}cd \\ &= ad + ab + cd + cb - ab - 2cb - cd = ad - cb = \begin{vmatrix} a & c \\ b & d \end{vmatrix} \end{aligned}$$

Damit entspricht der Flächeninhalt der Parallelogramms wie gewünscht der Determinanten. (Insbesondere überlegt man sich leicht¹ mittels Geradengleichung der Diagonalen, dass in diesem Fall der Term $ad - cb$ auch immer positiv ist.)

¹Der Punkt $\begin{pmatrix} c \\ d \end{pmatrix}$ liegt über der Diagonalen der 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 $ad - bc \geq 0$.

Diese Situation überträgt sich auf höhere Dimensionen.

Definition 11.20. Sei $m \in \mathbb{N} \setminus \{0\}$. Seien $a_1, \dots, a_m \in \mathbb{R}^m$. Dann ist

$$[a_1, \dots, a_m] = \left\{ \sum_{i=1}^m \lambda_i a_i \mid 0 \leq \lambda_i \leq 1 \right\}$$

das von a_1, \dots, a_m aufgespannte Parallelogramm.

Der folgende Satz kann mit den Rechengesetzen für Determinanten gezeigt werden:

Satz 11.21. Sei $m \in \mathbb{N} \setminus \{0\}$ und $a_1, \dots, a_m \in \mathbb{R}^m$. Für $j \leq m$ sei

$$a_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Dann ist das Volumen des von a_1, \dots, a_m aufgespannten Parallelogramms der Absolutbetrag von

$$\begin{vmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mm} \end{vmatrix}.$$

Insbesondere gilt: Wenn a_1, \dots, 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 nächsten Kapitel wieder aufgreifen.

12. 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, d.h. bei Vektorräumen zum Grundkörper \mathbb{R} , sind besonders die so genannten *Skalarprodukte* interessant, mit denen sich dann geometrische Begriffe wie Abstand, Winkel oder Orthogonalität definieren lassen.

Geometrische Begriffe in der reellen Ebene. Wir können dieses Prinzip erneut am Beispiel der wohl-bekannteren reellen Ebene motivieren. Im Folgenden werden wir die bekannten Begriffe wie Abstand, Winkel oder Orthogonalität aus einem (scheinbar) neuem, speziell definiertem Produkt von zwei Vektoren schrittweise herleiten. Wir starten mit der ersten

Definition 12.1. Sei $m \in \mathbb{N} \setminus \{0\}$ gegeben. Definiere eine zweistellige Funktion $\langle \cdot, \cdot \rangle^{st} : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}$ durch

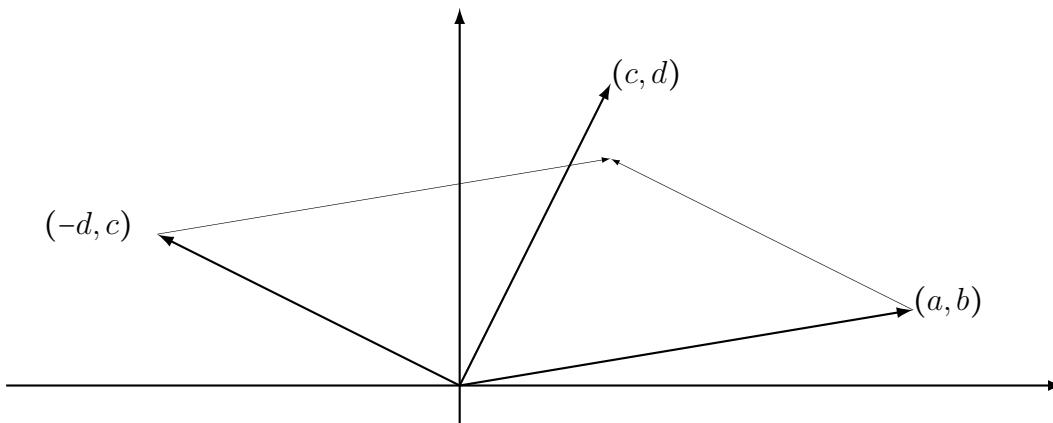
$$\langle x, y \rangle^{st} := \left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} \right\rangle^{st} := x_1 y_1 + \dots + x_m y_m.$$

Die Funktion $\langle \cdot, \cdot \rangle^{st}$ heisst das *Standard-Skalarprodukt* auf dem \mathbb{R}^m .

Es gilt nun in der reellen Ebene:

$$\left\langle \begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix} \right\rangle^{st} = ac + bd = \begin{vmatrix} a & -d \\ b & c \end{vmatrix} = \pm \text{Flächeninhalt} \left(\left[\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} -d \\ c \end{pmatrix} \right] \right)$$

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 11.21 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.



Mithilfe dieser Überlegung können wir schliesslich die gewünschten Begriffe herleiten. Wir wissen, lässt sich der Flächeninhalt eines Parallelogramms bereits aus bekannten trigonometrischen Zusammenhängen berechnen:

$$\text{Flächeninhalt} \left(\left[\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} -d \\ c \end{pmatrix} \right] \right) = \text{Länge} \left(\begin{pmatrix} a \\ b \end{pmatrix} \right) \cdot \text{Länge} \left(\begin{pmatrix} c \\ d \end{pmatrix} \right) \cdot \sin \angle \left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} -d \\ c \end{pmatrix} \right)$$

$$\begin{aligned}
&= \text{Länge}\left(\begin{pmatrix} a \\ b \end{pmatrix}\right) \cdot \text{Länge}\left(\begin{pmatrix} c \\ d \end{pmatrix}\right) \cdot \sin \angle \left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix} + \frac{\pi}{2} \right) \\
&= \text{Länge}\left(\begin{pmatrix} a \\ b \end{pmatrix}\right) \cdot \text{Länge}\left(\begin{pmatrix} c \\ d \end{pmatrix}\right) \cdot \cos \angle \left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix} \right)
\end{aligned}$$

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:

$$\begin{aligned}
\cos \angle \left(\begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix} \right) &= \frac{\left\langle \begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} a \\ b \end{pmatrix} \right\rangle^{\text{st}}}{\text{Länge}\left(\begin{pmatrix} a \\ b \end{pmatrix}\right) \cdot \text{Länge}\left(\begin{pmatrix} c \\ d \end{pmatrix}\right)} \\
&= \frac{\left\langle \begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} a \\ b \end{pmatrix} \right\rangle^{\text{st}}}{\sqrt{\left\langle \begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} a \\ b \end{pmatrix} \right\rangle^{\text{st}}} \cdot \sqrt{\left\langle \begin{pmatrix} c \\ d \end{pmatrix}, \begin{pmatrix} c \\ d \end{pmatrix} \right\rangle^{\text{st}}}}
\end{aligned}$$

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 einen Begriff von Skalarprodukt über die Angabe zu erfüllender Axiome und arbeiten uns langsam vor:

Definition 12.2. Sei V ein reeller Vektorraum und $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{R}$ 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

$$\langle \lambda x + \lambda' x', y \rangle = \lambda \langle x, y \rangle + \lambda' \langle x', y \rangle.$$

(b) $\langle \cdot, \cdot \rangle$ ist linear im zweiten Argument: Für alle $x, y, y' \in V$ und $\lambda, \lambda' \in \mathbb{R}$ ist

$$\langle x, \lambda y + \lambda' y' \rangle = \lambda \langle x, y \rangle + \lambda' \langle x, y' \rangle.$$

(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 $\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{R}$ darf nicht mit der Skalarmultiplikation $\mathbb{R} \times V \rightarrow V$; $(\lambda, v) \mapsto \lambda v$ 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 12.3. Das Standard-Skalarprodukt auf $V = \mathbb{R}^m$ ist ein Skalarprodukt im Sinne der Definition 12.2

Schliesslich können wir den Bogen zur Überschrift spannen:

Definition 12.4. Ein euklidischer Vektorraum ist ein Paar $(V, \langle \cdot, \cdot \rangle)$, das aus einem reellen Vektorraum V und einem Skalarprodukt $\langle \cdot, \cdot \rangle$ auf V besteht.

Wie beweisen einige Eigenschaften von Skalarprodukten:

Satz 12.5. Sei $V = (V, \langle \cdot, \cdot \rangle)$ 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 (d) aus der Definition 12.2 (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 12.6. Sei $V = (V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum. Die Norm auf V ist die Funktion

$$\|\cdot\|: V \rightarrow \mathbb{R}; \|\cdot\| := +\sqrt{\langle x, x \rangle}.$$

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 12.7. Sei $V = (V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum und $x, y \in V$. Dann gilt die Cauchy-Schwarzsche Ungleichung

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|.$$

Beweis: Es gilt:

$$\begin{aligned} 0 &\leq \langle \langle y, y \rangle \cdot x - \langle y, x \rangle \cdot y, \langle y, y \rangle \cdot x - \langle y, x \rangle \cdot y \rangle \quad (\text{pos. Definitheit}) \\ &= \langle y, y \rangle^2 \langle x, x \rangle - \langle y, y \rangle \langle y, x \rangle \langle x, y \rangle - \underbrace{\langle y, x \rangle \langle y, y \rangle \langle y, x \rangle + \langle y, x \rangle^2 \langle y, y \rangle}_{=0} \\ &= \langle y, y \rangle^2 \langle x, x \rangle - \langle y, y \rangle \langle y, x \rangle^2 \end{aligned}$$

Fall 1: $y = 0$ Dann gilt nach Satz 12.5(a) offenbar

$$|\langle x, y \rangle| = |\langle x, 0 \rangle| = 0 \leq \|x\| \cdot \|y\|.$$

Fall 2: $y \neq 0$ 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:

$$\begin{aligned} 0 &\leq \langle y, y \rangle \langle x, x \rangle - \langle y, x \rangle^2 \\ \langle y, x \rangle^2 &\leq \langle y, y \rangle \langle x, x \rangle \\ |\langle y, x \rangle| &\leq \sqrt{\langle y, y \rangle} \sqrt{\langle x, x \rangle} = \|y\| \cdot \|x\|. \end{aligned}$$

□

Die Cauchy-Schwarzsche Ungleichung folgt für das Standard-Skalarprodukt auch aus den anfänglichen geometrischen Überlegungen, denn $|\langle x, y \rangle|$ ist der Flächeninhalt eines Parallelogramms mit den Seitenlängen $\|x\|$ und $\|y\|$, also ist $|\langle x, y \rangle|$ 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 12.7 können wir nun grundlegende Längeneigenschaften der Norm nachweisen.

Satz 12.8. 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 12.5(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 12.7:

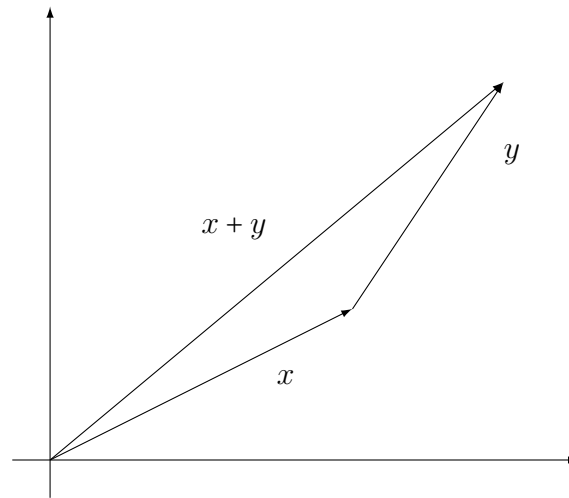
$$\begin{aligned} (\|x\| + \|y\|)^2 &= \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 \\ &\geq \langle x, x \rangle + 2 \cdot \langle x, y \rangle + \langle y, y \rangle \\ &= \langle x + y, x + y \rangle \\ &= \|x + y\|^2 \end{aligned}$$

Ziehen wir auf beide 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:

$$\|x + y\| \leq \|x\| + \|y\|.$$



* * *

Orthogonalität. Zwei Vektoren sind in besonderem Maße linear unabhängig, wenn sie orthogonal zueinander stehen. Geometrisch bedeutet dies, dass sie einen Winkel von $\frac{\pi}{2} = 90^\circ$ bilden. Im Sinne unserer anfänglichen Überlegungen zu euklidischen Vektorräumen können wir allgemein definieren:

Definition 12.9. 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

$$\cos(\alpha(x, y)) = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}; \quad 0 \leq \alpha(x, y) \leq \pi.$$

Nach der Cauchy-Schwarzschen Ungleichung aus Satz 12.7 ist

$$-1 \leq \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \leq 1$$

und somit ist der Kosinus des Öffnungswinkel wohldefiniert.

Anschaulich liegen zwei Vektoren orthogonal, wenn sie einen Öffnungswinkel von $\frac{\pi}{2}$ einschließen. Hierfür ist notwendig, dass $\cos(\alpha(x, y)) = \cos(\frac{\pi}{2})$ und somit $\langle x, y \rangle = 0$.

Definition 12.10. 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 12.11. 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

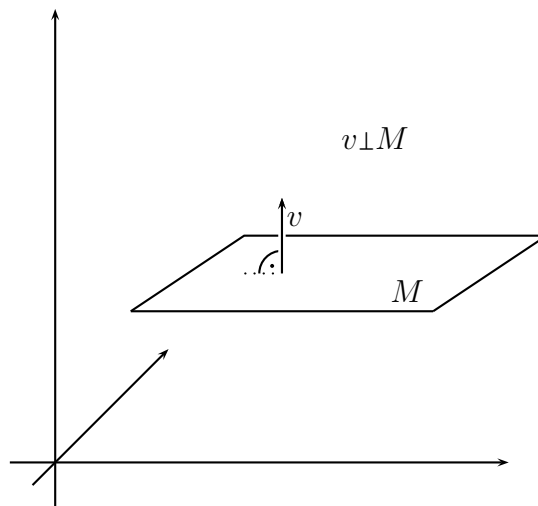
$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -b \\ a \end{pmatrix}.$$

Diese Vektoren sind dann auch orthogonal zueinander, denn

$$\left\langle \begin{pmatrix} a \\ b \end{pmatrix}, \begin{pmatrix} -b \\ a \end{pmatrix} \right\rangle^{\text{st}} = a \cdot (-b) + b \cdot a = 0.$$

Definition 12.12. 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

$$M^\perp := \{v \in V \mid v \perp M\}$$



Satz 12.13. 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 12.5 Teil (b).

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 12.5 Teil (b).

□

* * *

Orthonormalsysteme. Die kanonische Basis (e_0, \dots, e_{m-1}) 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 12.14. Sei $V = (V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum mit den Vektoren $v_1, \dots, v_r \in V$. Das r -Tupel (v_1, \dots, 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 12.15. Sei $V = (V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum und (v_1, \dots, v_r) ein Orthonormalsystem in V . Dann ist (v_1, \dots, v_r) linear unabhängig.

Beweis: Betrachte $\lambda_1, \dots, \lambda_r \in \mathbb{R}$ mit $\lambda_1 v_1 + \dots + \lambda_r v_r = 0$. Betrachte $i \leq r$. Die Anwendung des Skalarproduktes mit v_i auf beiden Seiten der Gleichung liefert:

$$\begin{aligned} 0 &= \langle 0, v_i \rangle = \langle \lambda_1 v_1 + \dots + \lambda_r v_r, v_i \rangle \\ &= \lambda_1 \langle v_1, v_i \rangle + \dots + \lambda_r \langle v_r, v_i \rangle = \lambda_i \langle v_i, v_i \rangle = \lambda_i \end{aligned}$$

Damit ist für alle $i \leq r$ der i -te Koeffizient $\lambda_i = 0$. Damit ist (v_1, \dots, v_r) linear unabhängig. □

Besonders wichtig ist, dass sich Koordinatendarstellungen bezüglich orthonormaler Basen durch Anwendung des Skalarproduktes ausrechnen lassen.

Satz 12.16. Sei $V = (V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum und (v_1, \dots, v_m) eine orthonormale Basis von V . Sei weiterhin $v \in V$ gegeben. Dann ist

$$(\langle v, v_1 \rangle, \dots, \langle v, v_m \rangle)$$

die Koordinatendarstellung des Vektors v in der Basis (v_1, \dots, 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, \dots, v_m) .

Beweis: Sei $(\lambda_1, \dots, \lambda_m)$ die eindeutig bestimmte Koordinatendarstellung des Vektors v bezüglich der Basis (v_1, \dots, 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 12.17 (Gram-Schmidtsches Orthonormalisierungsverfahren). Sei $V = (V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum und (v_1, \dots, v_r) linear unabhängig in V . Für $k \leq r$ definiere rekursiv

$$\tilde{v}_k = \frac{v_k - \sum_{i=1}^{k-1} \langle v_k, \tilde{v}_i \rangle \tilde{v}_i}{\|v_k - \sum_{i=1}^{k-1} \langle v_k, \tilde{v}_i \rangle \tilde{v}_i\|}$$

Dann gilt für $s \leq r$, dass $(\tilde{v}_1, \dots, \tilde{v}_s)$ wohldefiniert und orthonormal ist und dass die lineare Hüllen gleich sind:

$$\mathcal{L}(\tilde{v}_1, \dots, \tilde{v}_s) = \mathcal{L}(v_1, \dots, v_s)$$

Beweis: Wir zeigen die Behauptung durch Induktion über $s \leq r$.

Induktionsanfang: Für $s = 0$ gelten die Behauptungen trivialerweise.

Induktionsschritt: Angenommen, die Behauptungen gelten für $s < r$. Da (v_1, \dots, v_r) linear unabhängig ist, ist $v_{s+1} \notin \mathcal{L}(v_1, \dots, v_s) = \mathcal{L}(\tilde{v}_1, \dots, \tilde{v}_s)$. 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

$$\tilde{v}_{s+1} = \frac{v_{s+1} - \sum_{i=1}^s \langle v_{s+1}, \tilde{v}_i \rangle \tilde{v}_i}{\|v_{s+1} - \sum_{i=1}^s \langle v_{s+1}, \tilde{v}_i \rangle \tilde{v}_i\|}$$

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 $\{\tilde{v}_1, \dots, \tilde{v}_s\} \subseteq \mathcal{L}(v_1, \dots, v_s)$ und es gilt nach der obigen Darstellung $\tilde{v}_{s+1} \in \mathcal{L}(\tilde{v}_1, \dots, \tilde{v}_s, v_{s+1}) \subseteq \mathcal{L}(v_1, \dots, v_s, v_{s+1})$, so dass wir schließlich erhalten: $\mathcal{L}(\tilde{v}_1, \dots, \tilde{v}_s, \tilde{v}_{s+1}) \subseteq \mathcal{L}(v_1, \dots, v_s, v_{s+1})$.

Da man die definierende Gleichung für \tilde{v}_{s+1} auch nach v_{s+1} auflösen kann, gilt umgekehrt ebenfalls $\{v_1, \dots, v_s\} \subseteq \mathcal{L}(\tilde{v}_1, \dots, \tilde{v}_s)$, aber auch $v_{s+1} \in \mathcal{L}(\tilde{v}_1, \dots, \tilde{v}_s, \tilde{v}_{s+1})$ und somit wie gewünscht $\mathcal{L}(v_1, \dots, v_s, v_{s+1}) \subseteq \mathcal{L}(\tilde{v}_1, \dots, \tilde{v}_s, \tilde{v}_{s+1})$.

Der Satz gilt damit nach dem Induktionsprinzip. \square

Die Rekursionsvorschrift innerhalb des Satzes 12.17 ist im täglichen (mathematischen) Leben erträglich aufwendig, d.h. wir haben:

$$\begin{aligned} \boxed{\tilde{v}_1} &= \frac{v_1}{\|v_1\|} \\ \boxed{\tilde{v}_2} &= \frac{v_2 - \langle v_2, \tilde{v}_1 \rangle \cdot \tilde{v}_1}{\|v_2 - \langle v_2, \tilde{v}_1 \rangle \cdot \tilde{v}_1\|} \\ \boxed{\tilde{v}_3} &= \frac{v_3 - \langle v_3, \tilde{v}_1 \rangle \cdot \tilde{v}_1 - \langle v_3, \tilde{v}_2 \rangle \cdot \tilde{v}_2}{\|v_3 - \langle v_3, \tilde{v}_1 \rangle \cdot \tilde{v}_1 - \langle v_3, \tilde{v}_2 \rangle \cdot \tilde{v}_2\|} \end{aligned}$$

usw.

Beispiel 12.18. 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:

$$\begin{aligned} \boxed{\tilde{v}_1} &= \frac{v_1}{\|v_1\|} = \frac{\begin{pmatrix} 3 \\ 1 \end{pmatrix}}{\left\| \begin{pmatrix} 3 \\ 1 \end{pmatrix} \right\|} = \frac{1}{\sqrt{3^2 + 1^2}} \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{10}} \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} \\ v'_2 &= v_2 - \langle v_2, \tilde{v}_1 \rangle \cdot \tilde{v}_1 = \begin{pmatrix} 2 \\ 2 \end{pmatrix} - \left\langle \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \frac{1}{\sqrt{10}} \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} \right\rangle \cdot \frac{1}{\sqrt{10}} \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 \\ 2 \end{pmatrix} - \frac{1}{(\sqrt{10})^2} \cdot 8 \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 - \frac{47}{5} \\ 2 - \frac{41}{5} \end{pmatrix} = \frac{1}{5} \cdot \begin{pmatrix} -2 \\ 6 \end{pmatrix} \\ \|v'_2\| &= \sqrt{\left\langle \frac{1}{\sqrt{5}} \cdot \begin{pmatrix} -2 \\ 6 \end{pmatrix}, \frac{1}{\sqrt{5}} \cdot \begin{pmatrix} -2 \\ 6 \end{pmatrix} \right\rangle} = \sqrt{\frac{1}{25} \cdot (4 + 36)} = \frac{\sqrt{4 \cdot 10}}{5} = \frac{2}{5} \sqrt{10} \end{aligned}$$

$$\boxed{\tilde{v}_2} = \frac{v'_2}{\|v'_2\|} = \frac{\frac{1}{5} \cdot \begin{pmatrix} -2 \\ 6 \end{pmatrix}}{\frac{2}{5}\sqrt{10}} = \frac{5}{2 \cdot \sqrt{10} \cdot 5} \cdot \begin{pmatrix} -2 \\ 6 \end{pmatrix} = \frac{1}{\sqrt{10}} \cdot \begin{pmatrix} -1 \\ 3 \end{pmatrix}$$

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:

$$\begin{aligned} \langle \tilde{v}_1, \tilde{v}_1 \rangle &= \left\langle \frac{1}{\sqrt{10}} \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{10}} \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} \right\rangle = \frac{1}{(\sqrt{10})^2} \cdot (3^2 + 1^2) = 1 \\ \langle \tilde{v}_2, \tilde{v}_2 \rangle &= \left\langle \frac{1}{\sqrt{10}} \cdot \begin{pmatrix} -1 \\ 3 \end{pmatrix}, \frac{1}{\sqrt{10}} \cdot \begin{pmatrix} -1 \\ 3 \end{pmatrix} \right\rangle = \frac{1}{(\sqrt{10})^2} \cdot ((-1)^2 + 3^2) = 1 \\ \langle \tilde{v}_1, \tilde{v}_2 \rangle &= \left\langle \frac{1}{\sqrt{10}} \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{10}} \cdot \begin{pmatrix} -1 \\ 3 \end{pmatrix} \right\rangle = \frac{1}{(\sqrt{10})^2} \cdot (3 \cdot (-1) + 1 \cdot 3) = 0 \end{aligned}$$

* * *

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 12.19. *Sei $V = (\langle \cdot, \cdot \rangle)$ 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 12.17 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 :

$$\begin{aligned} \langle v, v_j \rangle &= \langle w - u, v_j \rangle \\ &= \left\langle w - \sum_{i=1}^r \langle w, v_i \rangle v_i, v_j \right\rangle \\ &= \langle w, v_j \rangle - \sum_{i=1}^r \langle w, v_i \rangle \langle v_i, v_j \rangle \end{aligned}$$

$$\begin{aligned}
&= \langle w, v_j \rangle - \langle w, v_j \rangle \langle v_j, v_j \rangle \\
&= \langle w, v_j \rangle - \langle w, v_j \rangle \\
&= 0
\end{aligned}$$

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 12.13 (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 Satz ist damit bewiesen. \square

Nach diesem Satz ist also folgende Zerlegung gegeben:

$$V = M + M^\perp = \{u + v \mid u \in M \text{ und } v \in M^\perp\}.$$

Da weiterhin $M \cap M^\perp = \emptyset$ 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 12.20. Sei $V = (V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum und M ein endlich-dimensionaler Unterraum von V . Definiere die **orthogonale Projektion**

$$P_M: V \rightarrow M$$

auf M durch: Für $w \in V$ ist $f(w)$ das nach Satz 12.19 eindeutig bestimmte $u \in M$, so dass $w - u \in M^\perp$.

Diese orthogonale Projektion hat die gewünschten Eigenschaften, wie wir jetzt sehen werden:

Satz 12.21. Sei $P_M: V \rightarrow M$ die in Definition 12.20 definierte orthogonale Projektion von V auf den Unterraum $M \subseteq V$. Dann gilt:

- (a) $P_M: V \rightarrow M$ ist eine lineare Abbildung.
- (b) $P_M \upharpoonright M = \text{Id}_M$ und $P_M: V \rightarrow M$ ist surjektiv.
- (c) $\text{Kern}(P_M) = M^\perp$.

Beweis:

Zu (a): Betrachte $w, w' \in V$ und $\lambda \in \mathbb{R}$. Dann sind

$$w = P_M(w) + (w - P_M(w)) \text{ und } w' = P_M(w') + (w' - P_M(w'))$$

die Zerlegungen von w und w' nach M und M^\perp . Dann gilt

$$\begin{aligned} w + w' &= (P_M(w) + (w - P_M(w))) + (P_M(w') + (w' - P_M(w'))) \\ &= (P_M(w) + P_M(w')) + ((w - P_M(w)) + (w' - P_M(w'))) \end{aligned}$$

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 $\lambda w = \lambda P_M(w) + \lambda(w - P_M(w))$ die Zerlegung von λw nach M und M^\perp , da $\lambda P_M(w) \in M$ und $\lambda(w - P_M(w)) \in M^\perp$. Damit ist

$$P_M(\lambda w) = \lambda \cdot (P_M(w)).$$

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 $P_M(w) = w = \text{Id}_M(w)$. Damit ist $P_M \upharpoonright M = \text{Id}_M$.

Zu (c): Betrachte $w \in \text{Kern}(P_M)$. Dann ist $P_M(w) = 0$ und

$$w = P_M(w) + (w - P_M(w)) = 0 + w$$

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 $w \in \text{Kern}(P_M)$.

□

* * *

Orthogonale Abbildungen. Wir geben im Folgenden noch abschließend einen kleinen Ausblick. Wir definieren orthogonale Abbildungen und werden sehen, dass wir bereits wohlbekannt Beispiele dafür kennen.

Definition 12.22. Seien $V = (V, \langle \cdot, \cdot \rangle_V)$ und $W = (W, \langle \cdot, \cdot \rangle_W)$ euklidische Vektorräume und sei $f: V \rightarrow W$ eine lineare Abbildung. Dann ist f *isometrisch* oder *orthogonal*, wenn für alle $u, v \in V$ gilt:

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

Beispiel 12.23. 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:

$$\begin{aligned} & \left\langle f\left(\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}\right), f\left(\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}\right) \right\rangle \\ &= \left\langle \begin{pmatrix} u_1 \cos \alpha - u_2 \sin \alpha \\ u_1 \sin \alpha + u_2 \cos \alpha \end{pmatrix}, \begin{pmatrix} v_1 \cos \alpha - v_2 \sin \alpha \\ v_1 \sin \alpha + v_2 \cos \alpha \end{pmatrix} \right\rangle \\ &= u_1 v_1 \cos \alpha \cos \alpha - u_1 v_2 \cos \alpha \sin \alpha - u_2 v_1 \sin \alpha \cos \alpha + u_2 v_2 \sin \alpha \sin \alpha \\ &\quad + u_1 v_1 \sin \alpha \sin \alpha + u_1 v_2 \sin \alpha \cos \alpha + u_2 v_1 \cos \alpha \sin \alpha + u_2 v_2 \cos \alpha \cos \alpha \\ &= u_1 v_1 \cos^2 \alpha + u_1 v_1 \sin^2 \alpha + u_2 v_2 \sin^2 \alpha + u_2 v_2 \cos^2 \alpha \\ &= u_1 v_1 + u_2 v_2 = \left\langle \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \right\rangle. \end{aligned}$$

Damit ist die Drehung f isometrisch.

Orthogonale Abbildungen sind bereits sehr speziell, wie die folgenden beiden Sätze zeigen werden:

Satz 12.24. Seien $V = (V, \langle \cdot, \cdot \rangle_V)$ und $W = (W, \langle \cdot, \cdot \rangle_W)$ euklidische Vektorräume und sei $f: V \rightarrow W$ 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 12.15 ist dann $v - v' = 0$ und somit $v = v'$. \square

Satz 12.25. Sei $V = (V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum und $f: V \rightarrow V$ ein orthogonaler Endomorphismus. Dann ist f ein Automorphismus.

Beweis: Nach Satz 12.24 ist f injektiv. Nach der Dimensionsformel 5.17 ist ein injektiver Homomorphismus surjektiv. \square

Orthogonale Abbildungen haben in sich schöne Eigenschaften, so dass man die Menge solcher Abbildungen zusammenfäkt.

Definition 12.26. Sei $V = (V, \langle \cdot, \cdot \rangle)$ ein endlich-dimensionaler Vektorraum. Dann sei

$$O(V) = \{f \mid f: V \rightarrow V \text{ ist orthogonal}\}$$

die orthogonale Gruppe von V .

Wenn $\mathbb{R}^m = (\mathbb{R}^m, \langle \cdot, \cdot \rangle)$ 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 12.27. 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:
 $f \circ \text{Id}_V = f$ und $\text{Id}_V \circ f = f$ (Neutrales Element)
- (c) Für alle $f \in O(V)$ gilt:
 $f^{-1} \circ f = \text{Id}_V$ und $f \circ f^{-1} = \text{Id}_V$ (Existenz von Inversen)

Ein Element f von $O(m)$ wird in der kanonischen Basis des \mathbb{R}^m eine als $(m \times m)$ -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 12.28. Sei $m \in \mathbb{N} \setminus \{0\}$. Eine Matrix $A \in \text{Mat}(m \times m, \mathbb{R})$ ist orthogonal, wenn das System der Spaltenvektoren von A orthonormal ist.

In der Dimension $m = 2$ sind Drehmatrizen und Spiegelmatrizen

$$\begin{pmatrix} \cos \alpha & \mp \sin \alpha \\ \sin \alpha & \pm \cos \alpha \end{pmatrix}$$

orthonormal, 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 12.29. Jedes $A \in O(2)$ ist von der Form

$$\begin{pmatrix} \cos \alpha & \mp \sin \alpha \\ \sin \alpha & \pm \cos \alpha \end{pmatrix}$$

Damit schliessen wir diesen kleinen Exkurs.

INDEX

- Öffnungswinkel, 109
- Abbruchkriterium, 96
- Abstand, 103
- algebraische Vielfachheit, 97
- Assoziativgesetz, 11, 13, 46
- Automorphismengruppe, 36
- Automorphismus, 36, 117
- Basis, 22, 96
- Basisaustauschsatz, 27
- Basisdarstellung, 25, 65, 111
- Basisergänzungssatz, 26
- Basisexistenzsatz, 27
- Betrag, 21
- Bild, 33, 34, 39, 49, 71, 82, 94
- CAESAR-Verschlüsselung, 11
- Cauchy-Schwarzsche Ungleichung, 107, 109
- characteristisches Polynom, 83, 93, 96
- Cramersche Regel, 92
- darstellende Matrix, 42, 69, 70, 72, 94, 96
- Defekt, 33
- Determinante, 84, 86, 91, 101
- diagonalisierbar, 75
- Diagonalisierung, 81
- Diagonalmatrix, 62
- Dimension, 28, 33
- Dimensionsformel, 71, 82
- Dimensionssatz, 39
- direkte Summe, 115
- Diskriminante, 19
- Distributivgesetz, 11, 13, 46
- Drehung, 51, 73, 76, 110, 118
- Dreiecksmatrix, 62
- Dreiecksungleichung, 108
- Eigenraum, 78, 81, 96
- Eigenvektor, 75, 79, 94, 96
- Eigenwert, 75, 81, 94, 96
- eindeutig lösbar, 7
- Einheitsmatrix, 47, 64
- elementare Zeilenumformung, 56, 59, 63, 90
- Elementarmatrix, 59
- endlich-dimensional, 28, 38, 75
- endlich-erzeugt, 23
- Endomorphismus, 36, 72, 74, 94
- Entwicklungsformel, 112
- Epimorphismus, 35
- erweiterte Koeffizientenmatrix, 59
- euklidischer Vektorraum, 106
- Existenzsätze für Basen, 26
- Flächeninhalt, 102, 104, 107
- Fundamentalsatz der Algebra, 94
- Funktion, 12, 16
- Gaußscher Algorithmus, 8, 63, 86
- geometrische Vielfachheit, 78, 80, 97
- Gleichung, 5, 9
- Gram-Schmidtsche Orthonormalisierungsverfahren, 112
- Gruppe, 118
- Hauptachsentransformation, 100
- homogen, 6, 60
- Homomorphismus, 31, 35, 36
- imaginäre Einheit, 17
- Imaginärteil, 17
- inhomogen, 60
- Inverse, 10, 11, 13, 47, 64, 82, 90, 94
- isometrisch, 116
- Isomorphiesatz, 43
- Isomorphismus, 35
- Körper, 5, 10
- kanonische Basis, 25, 66
- Kern, 33, 34, 39, 49, 60, 71, 81, 82, 94
- Koeffizientenmatrix, 59, 63
- Kommutativgesetz, 11, 13
- komplexe Addition, 17
- komplexe Division, 18
- komplexe Multiplikation, 18
- komplexe Subtraktion, 20
- komplexe Zahlen, 6, 14, 17, 18, 29, 94
- Konjugierte, 18
- Kosinussatz, 21
- Länge, 105, 107, 108
- lösbar, 7, 60
- Lösungsmenge, 6, 60
- Laplacescher Entwicklungssatz, 84, 85

- linear unabhängig, 24, 79, 80, 96, 111
- lineare Abbildung, 31, 39, 46, 48, 68, 115
- lineare Codes, 49
- lineare Hülle, 23
- lineares Gleichungssystem, 5, 6, 31, 60, 63, 77, 94
- Linearfaktor, 96
- Linearkombination, 23

- Matrix, 9, 41, 48
- Matrizenaddition, 42
- Matrizenprodukt, 45
- Monomorphismus, 35

- Neutralität, 11, 13
- Norm, 107
- Nullmatrix, 43
- Nullstelle, 94
- Nullvektor, 10, 60

- orthogonal, 109, 116, 118
- orthogonale Abbildung, 116, 117
- orthogonale Gruppe, 118
- orthogonale Projektion, 115
- orthogonale Zerlegung, 114
- orthogonales Komplement, 110, 114
- Orthogonalität, 103, 105, 109
- Orthonormalsystem, 111
- Ortsvektor, 20, 21

- p-q-Formel, 19
- Parallelogramm, 103
- Pivot-Element, 7
- Polarkoordinaten, 21, 22
- positiv definit, 106
- Projektion, 114

- quadratische Matrix, 47, 62

- Rang, 33, 94
- rationale Zahlen, 14, 29, 30
- Realteil, 17
- reelle Ebene, 20, 51, 54, 104
- reelle Zahlen, 14, 18, 29, 30
- Regel von Sarrus, 85
- Restklasse, 8, 11
- Ring, 5

- Schachbrettmuster, 85
- senkrecht, 109

- skalare Multiplikation, 43
- Skalarmultiplikation, 10
- Skalarprodukt, 103, 105
- Spaltenvektor, 42
- Spiegelung, 54, 72, 73, 76, 83, 118
- Standard-Skalarprodukt, 104
- symmetrisch, 106
- symmetrische Matrix, 72

- Transformationsmatrix, 70, 95, 97
- Transformationsatz, 65
- transponierte Matrix, 72

- Unbekannte, 5
- unendlich-dimensional, 28
- Untervektorraum, 14, 29, 60

- Vektoraddition, 10
- Vektorraum, 10, 13, 14, 35, 38, 43
- Vektorsubtraktion, 11
- Volumen, 101

- Winkel, 103, 105

- Zeilenstufenform, 7, 90
- Zeilenstufenmatrix, 62
- Zeilenvektor, 41
- Zerlegung, 115