

Inhaltsverzeichnis

Inhaltsverzeichnis	i
Einführung	ii
1 Der Vektorraum $\mathbb{R}^n$	1
1.1 Konstruktion des Vektorraumes	1
1.1.1 Unterräume	5
1.2 Basis und Dimension	7
1.2.1 Spalten vs. Zeilen	14
1.3 Der Vektorraum $\mathbb{K}^n$	15
1.4 Anschauliche Vektorrechnung	16
1.4.1 Addition von Vektoren	17
1.4.2 Multiplikation mit Skalar	20
1.5 Verständnisfragen	22
1.6 Rechenaufgaben	23
2 Matrizen	24
2.1 Definition und Beispiele	25
2.2 Matrizen als Vektorraum	29
2.3 Multiplikation von Matrizen	32
2.3.1 Falk-Schema	36
2.3.2 Warnungen	38
2.4 Inverse von Matrizen	39
2.5 Lineare Abbildungen	41
2.5.1 Matrixdarstellungen lin. Abbildungen	49
2.6 Verständnisfragen	53
2.7 Rechenaufgaben	54
3 Lineare Gleichungssysteme	56
3.1 Berechnung von Lösungen	60
3.1.1 Beschreibung des Gauß-Algorithmus	62

3.1.2	Der Gauß-Jordan Algorithmus	65
3.2	Inverse von Matrizen	68
3.3	Was man mit Gauß/Gauß-Jordan alles berechnen kann	69
3.4	Verständnisfragen	76
3.5	Rechenaufgaben	79
4	Determinanten	83
4.1	Geometrische Interpretation	89
4.2	Inverse von Matrizen	90
4.3	Verständnisfragen	91
4.4	Rechenaufgaben	92
5	Längen-, Abstands- und Winkelmessung im $\mathbb{R}^n$	93
5.1	Längen- und Abstandsmessung im $\mathbb{R}^n$	94
5.2	Das (Standard)Skalarprodukt	100
5.3	Gram-Schmidtsches Orthogonalisierungsverfahren	106
5.4	Das Vektorprodukt in $\mathbb{R}^3$	107
5.5	Das Spatprodukt in $\mathbb{R}^3$	114
5.6	Elemente der analytischen Geometrie	116
5.7	Matrix-Multiplikation	123
5.8	Verständnisfragen	124
5.9	Rechenaufgaben	126
6	Orthogonale Abbildungen	128
6.1	Drehungen im $\mathbb{R}^2$ und $\mathbb{R}^3$	128
6.2	Spiegelungen im $\mathbb{R}^2$ und $\mathbb{R}^3$	134
6.3	Allgemeine Orthogonale Abbildungen	136
6.4	Verständnisfragen	138
7	Eigenwerte und Eigenvektoren	139
7.1	Definition und erste Ergebnisse	143
7.2	Ähnliche Matrizen	147
7.3	Basistransformationen	149
7.4	Diagonalisierbare Matrizen	151

7.5	Jordansche Normalform	156
7.6	Spektralsatz	158
7.7	Eigenwerte spezieller Matrizen	159
7.8	Anwendung: Hauptachsentransformation	161
7.9	Verständnisfragen	164
7.10	Rechenaufgaben	165
8	Allgemeine Vektorräume	167
8.1	Definition und Beispiele	167
8.2	Lineare Abhängigkeit	173
8.3	Linearkombinationen	176
8.4	Basen	178
8.5	Dimension	180
8.6	Lineare Abbildungen	182
8.7	Übungsaufgaben	183
9	Fourier-Reihen	184
9.1	Definition	185
9.2	Konvergenz	187
9.3	Welche Funktionen sind entwickelbar?	190
9.4	Das Gibbs-Phänomen	193
9.5	Bessel und Parseval	196
9.6	Konvergenz im quadratischen Mittel	198
9.7	Optimalitätseigenschaft trig. Polynome	199
9.7.1	Gliedweise Integrierbarkeit und Differenzierbarkeit	200
9.8	Komplexe Darstellung reeller Fourier-Reihen	203
	Literaturverzeichnis	207

Einführung

Der Vektorraum $\mathbb{R}^n$

1.1 Konstruktion des Vektorraumes

Wir starten unsere Diskussion mit der Konstruktion des Vektorraumes $\mathbb{R}^n$ für $n \in \mathbb{N}$. Der Fall $n = 2, 3$ sollte aus der Analytischen Geometrie im Abitur den meisten vertraut sein. Es empfiehlt sich trotzdem die nachfolgenden Konstruktionen genau zu studieren damit Sie die Verallgemeinerungen in Kapitel 8 ohne Probleme verstehen.

Mit $\mathbb{R}^n$ bezeichnen wir die Menge aller **Tupel**

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad x_i \in \mathbb{R}, \quad 1 \leq i \leq n$$

bezeichnet, die wir im folgenden **Vektoren** nennen werden. Wir definieren auf diesen Vektoren eine **Vektoraddition**

$$\begin{cases} \mathbb{R}^n \times \mathbb{R}^n & \rightarrow \mathbb{R}^n \\ (\mathbf{x}, \mathbf{y}) & \mapsto \mathbf{x} + \mathbf{y} \end{cases}$$

durch

$$\mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} := \begin{bmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{bmatrix}$$

und wir setzen

$$\mathbf{0} := \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \quad -\mathbf{x} := \begin{bmatrix} -x_1 \\ \vdots \\ -x_n \end{bmatrix}.$$

Dabei sind die $x_i + y_i$, $i = 1, \dots, n$ durch die gewöhnliche Addition reeller Zahlen erklärt. Zu den Eigenschaften siehe Kapitel 2 in Mathematik 1. Zur graphischen Interpretation siehe auch Abschnitt 1.4.

Weiterhin definieren wir eine **Skalarmultiplikation**¹

$$\begin{cases} \mathbb{R} \times \mathbb{R}^n & \rightarrow \mathbb{R}^n \\ (\lambda, \mathbf{x}) & \mapsto \lambda \mathbf{x} \end{cases}$$

durch

$$\lambda \mathbf{x} = \lambda \begin{bmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{bmatrix} := \begin{bmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{bmatrix}.$$

Dabei sind die λx_i , $i = 1, \dots, n$ durch die gewöhnliche Multiplikation in den reellen Zahlen erklärt. Siehe zu den Eigenschaften Kapitel 2 in Mathematik 1.

Definition 1.1 (Der Vektorraum $\mathbb{R}^n$).

*Wir nennen die Menge $\mathbb{R}^n$ ausgestattet mit den oben eingeführten Operationen $+$ und $\cdot$ den **Vektorraum** $\mathbb{R}^n$.*

*Die Rechenoperationen nennen wir die **Vektorraumoperationen**.*

Wir sollten eigentlich $(\mathbb{R}^n, +, \cdot)$ schreiben, da die Operationen zum Begriff Vektorraum gehören, werden dies aber wie in der Literatur üblich, nicht tun.

Wir sollten hier unterstreichen, dass wir das Objekt $(\mathbb{R}^n, +, \cdot)$ einfach als Vektorraum bezeichnen. Wir werden später weitere Objekte einführen die wir Vektorräume nennen werden. Es ist der folgende Satz, der diese Bezeichnung rechtfertigt. Wir werden sehen, daß die genau gleichen Rechenregeln für andere Objekte gelten. Sie Sätze 2.2 und 2.6. Die strukturelle Ähnlichkeit dieser Objekte werden uns später motivieren die allgemeine Definition von Vektorräumen einzuführen.

¹Nicht mit dem Ihnen vielleicht bekannten Skalarprodukt verwechseln!

Aus unserer Definition der Vektorraumoperationen können wir die folgenden Eigenschaften ableiten:

Satz 1.1.

Die Vektorraumoperationen

$$\left\{ \begin{array}{ll} +: \mathbb{R}^n \times \mathbb{R}^n & \rightarrow \mathbb{R}^n, \\ (\mathbf{x}, \mathbf{y}) & \mapsto \mathbf{x} + \mathbf{y} \end{array} \right. \quad \text{und} \quad \left\{ \begin{array}{ll} \cdot: \mathbb{R} \times \mathbb{R}^n & \rightarrow \mathbb{R}^n, \\ (\lambda, \mathbf{x}) & \mapsto \lambda \mathbf{x}. \end{array} \right.$$

des Vektorraumes $\mathbb{R}^n$ erfüllen die folgenden Eigenschaften:

- (i) Für alle $\mathbf{x} \in \mathbb{R}^n$ gilt $\mathbf{x} + \mathbf{0} = \mathbf{x}$.
- (ii) Für alle $\mathbf{x} \in \mathbb{R}^n$ gilt $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$.
- (iii) Für alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ gilt $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$.
- (iv) Für alle $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ gilt $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$.
- (v) Für alle $\mathbf{x} \in \mathbb{R}^n$ gilt $1\mathbf{x} = \mathbf{x}$.
- (vi) Für alle $\lambda, \mu \in \mathbb{R}$ und alle $\mathbf{x} \in \mathbb{R}^n$ gilt $\lambda(\mu\mathbf{x}) = (\lambda\mu)\mathbf{x}$.
- (vii) Für alle $\lambda \in \mathbb{R}$ und alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ gilt $\lambda(\mathbf{x} + \mathbf{y}) = \lambda\mathbf{x} + \lambda\mathbf{y}$.
- (viii) Für alle $\lambda, \mu \in \mathbb{R}$ und alle $\mathbf{x} \in \mathbb{R}^n$ gilt $(\lambda + \mu)\mathbf{x} = \lambda\mathbf{x} + \mu\mathbf{x}$.

Satz 1.2.

In $\mathbb{R}^n$ gelten weiterhin die folgenden Regeln:

- $0\mathbf{x} = \lambda\mathbf{0} = \mathbf{0}$ für alle $\lambda \in \mathbb{R}$ und alle $\mathbf{x} \in \mathbb{R}^n$.
- $(-\lambda)\mathbf{x} = \lambda(-\mathbf{x}) = -(\lambda\mathbf{x})$ für alle $\lambda \in \mathbb{R}$ und alle $\mathbf{x} \in \mathbb{R}^n$.
- $(-\lambda)(-\mathbf{x}) = \lambda\mathbf{x}$ für alle $\lambda \in \mathbb{R}$ und alle $\mathbf{x} \in \mathbb{R}^n$.

Übungsaufgabe 1.1.

Beweisen Sie Satz 1.2 mit Hilfe der Aussagen in Satz 1.1.

1.1.1 Unterräume

Definition 1.2 (Unterraum).

Sei $U \subseteq \mathbb{R}^n$, $U \neq \emptyset$. Wenn

- $\mathbf{x} + \mathbf{y} \in U$ für alle $\mathbf{x}, \mathbf{y} \in U$ und
- $\lambda \mathbf{x} \in U$ für alle $\mathbf{x} \in U$ und alle $\lambda \in \mathbb{R}$,

dann nennt man U einen **Unterraum** (Untervektorraum) von $\mathbb{R}^n$.

Bemerkung 1.1.

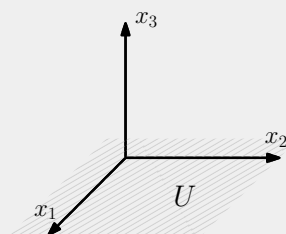
Für die in Definition 1.2 genannten Punkte verwendet man auch zusammenfassend die Sprechweise: U ist **abgeschlossen** (in V) unter Addition und Skalarmultiplikation, also **abgeschlossen unter den Vektorraumoperationen**.

Beispiel 1.1.

Natürlich sind die Vektorräume $\mathbb{R}^n$ und $\{\mathbf{0}\}$ (**Nullraum**) jeweils Unterräume des $\mathbb{R}^n$. Diese werden **triviale Unterräume** genannt. Die Menge

$$U = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ 0 \end{bmatrix} : x_1, x_2 \in \mathbb{R} \right\}$$

ist ein Unterraum von $\mathbb{R}^3$.



Beispiel 1.2.

Die Menge

$$U = \left\{ \begin{bmatrix} x_1 \\ 0 \\ x_3 \end{bmatrix} : x_1, x_3 \in \mathbb{R} \right\} \subseteq \mathbb{R}^3$$

ist ein Unterraum von $\mathbb{R}^3$. Die Details zu prüfen bleibt der Leserin als Übungsaufgabe überlassen.

1.2 Basis und Dimension

Definition 1.3 (Lineare (Un)Abhängigkeit).

Eine Menge von Vektoren $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ heißt genau dann

- (i) **linear abhängig**, falls eine korrespondierende Menge $\{\alpha_1, \dots, \alpha_k\} \subseteq \mathbb{R}$ existiert mit $\alpha_i \neq 0$ für ein $1 \leq i \leq k$ und

$$\mathbf{0} = \alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k = \sum_{i=1}^k \alpha_i \mathbf{x}_i.$$

- (ii) **linear unabhängig**, falls aus

$$\mathbf{0} = \alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k = \sum_{i=1}^k \alpha_i \mathbf{x}_i.$$

stets $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$ folgt.

Wir werden später effiziente Möglichkeiten kennenlernen (Gauß-Algorithmus), Informationen über lineare Unabhängigkeit zu erhalten. Versuchen Sie sich trotzdem bereits hier an folgender Aufgabe.

Übungsaufgabe 1.2.

Für welche der folgenden Mengen $X_j \subseteq \mathbb{R}^2$ sind die Vektoren aus X_j linear unabhängig? Geben Sie jeweils eine schlüssige Begründung.

$$X_1 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}, \quad X_2 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

$$X_3 = \left\{ \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\}, \quad X_4 = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 4 \\ 0 \end{bmatrix} \right\}$$

Definition 1.4 (Linearkombination).

Ein Vektor $\mathbf{x} \in \mathbb{R}^n$ heißt genau dann eine **Linearkombination** von $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$, wenn

$$\mathbf{x} = \alpha_1 \mathbf{x}_1 + \dots + \alpha_k \mathbf{x}_k = \sum_{i=1}^k \alpha_i \mathbf{x}_i$$

für $\{\alpha_1, \dots, \alpha_k\} \subseteq \mathbb{R}$.

Man sagt auch, $\mathbf{x}$ läßt sich aus $\mathbf{x}_1, \dots, \mathbf{x}_k$ linear kombinieren.

Die $\alpha_1, \dots, \alpha_k$ heißen **Koeffizienten** der Linearkombination.

Definition 1.5 (Lineare Hülle).

Es sei $\{\mathbf{x}_1, \dots, \mathbf{x}_k\} \subseteq \mathbb{R}^n$. Die Menge

$$\text{span}(\mathbf{x}_1, \dots, \mathbf{x}_k) = \left\{ \sum_{i=1}^k \alpha_i \mathbf{x}_i : \alpha_i \in \mathbb{R}, 1 \leq i \leq k \right\}$$

heißt dann der **Spann** oder die **lineare Hülle** von $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$.

Allgemeiner ist für $L \subseteq \mathbb{R}^n$ die lineare Hülle $\text{span}(L)$ von L die Menge aller **endlichen Linearkombinationen** von Elementen in L .

Es gilt

Satz 1.3.

Es sei $\{\mathbf{x}_1, \dots, \mathbf{x}_k\} \subseteq \mathbb{R}^n$. Dann ist

$$U = \text{span}(\mathbf{x}_1, \dots, \mathbf{x}_k)$$

ein Unterraum von $\mathbb{R}^n$.

Übungsaufgabe 1.3.

Beweisen Sie Satz 1.3.

Definition 1.6 (Erzeugendensystem).

Eine Menge $\{\mathbf{x}_1, \dots, \mathbf{x}_k\} \subseteq \mathbb{R}^n$ heißt genau dann ein **Erzeugendensystem** von $\mathbb{R}^n$, wenn

$$\text{span}(\mathbf{x}_1, \dots, \mathbf{x}_k) = \mathbb{R}^n.$$

Man sagt dann auch, daß die $\mathbf{x}_1, \dots, \mathbf{x}_k$ den $\mathbb{R}^n$ aufspannen.

Beispiel 1.3.

Natürlich ist $\mathbb{R}^n$ ein Erzeugendensystem für sich selbst, d.h. $\text{span}(\mathbb{R}^n) = \text{span}(\mathbb{R}^n)$.

Interessanter ist das Beispiel

$$\text{span} \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\} = \mathbb{R}^3.$$

Definition 1.7 (Basis).

Eine Menge $B = \{\mathbf{x}_1, \dots, \mathbf{x}_k\} \subseteq \mathbb{R}^n$ heißt genau dann eine **Basis**, wenn sie ein **linear unabhängiges Erzeugendensystem** von $\mathbb{R}^n$ ist.

Satz 1.4 (Standardbasis des $\mathbb{R}^n$).

Die Vektoren

$$\left\{ \mathbf{e}_i = \begin{bmatrix} \delta_{1i} \\ \delta_{2i} \\ \vdots \\ \delta_{ni} \end{bmatrix} : i = 1, \dots, n \right\}, \quad \delta_{ij} = \begin{cases} 1 & : i = j \\ 0 & : i \neq j \end{cases}$$

bilden eine Basis von $\mathbb{R}^n$ die sogenannte **Standardbasis** oder **kanonische Basis**.

Den folgenden Satz nehmen wir hin. Er ist nötig um sicherzustellen, daß das Konzept der Dimension ein sinnvolles ist.

Satz 1.5.

Alle Basen des $\mathbb{R}^n$ haben die gleiche Anzahl an Elementen.

Insbesondere: Da die Standardbasis n Elemente hat, hat jede Basis von $\mathbb{R}^n$ genau n Elemente.

Für interessierte:

Beweis. Angenommen wir haben die Standardbasis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ und eine weitere $\{\mathbf{b}_1, \dots, \mathbf{b}_m\}$. Nach dem Steinitzschen Austauschlemma durch, wenn nötig Umnummerierung, eine Basis

$$\{\mathbf{b}_1, \dots, \mathbf{b}_m, \mathbf{e}_{m+1}, \dots, \mathbf{e}_n\}$$

herstellen also muß $m \leq n$. Ebenso lässt sich eine Basis

$$\{\mathbf{e}_1, \dots, \mathbf{e}_n, \mathbf{b}_{n+1}, \dots, \mathbf{b}_m\}$$

herstellen also muß $n \leq m$ gelten. Damit gilt $m = n$. □

Mit dem letzten Satz können wir den Begriff der Dimension einführen und damit Vektorräume wie $\mathbb{R}^2$ und $\mathbb{R}^3$ in ihrer 'Größe' unterscheiden. Größe hat hier nichts mit der Anzahl der Elemente zu tun (beide haben gleich viel Elemente) sondern mit 'Freiheitsgraden' wie man Dimensionen in der Technik und Physik nennen würde.

Definition 1.8 (Dimension).

Die **Dimension** des Vektorraumes $\mathbb{R}^n$ ist die Anzahl n der Elemente einer Basis von $\mathbb{R}^n$.

Satz 1.6.

Eine Menge von $n+1$ paarweise verschieden Vektoren $\{\mathbf{x}_1, \dots, \mathbf{x}_{n+1}\} \subseteq \mathbb{R}^n$ ist stets linear abhängig.

Bemerkung 1.2.

Viele Studierende machen den Fehler zu glauben, daß Sie nur 2 und 3 dimensionale Vektorräume brauchen.

Dieser Glaube ist so falsch, wie es nur geht. Nicht nur sind schon physikalische Anwendungen hochdimensional, beispielsweise haben Erde, Mond und Sonne einen 3 · 6-dimensionalen Phasenraum ($\text{Ort} \in \mathbb{R}^3$, $\text{Impuls} \in \mathbb{R}^3$ jeweils drei mal) sondern auch Bild- und Tonverarbeitung, Tomographie, Data Analysis, kommen alle hochdimensional aus der Praxis.

Die Diskretisierung kontinuierlicher Systeme um Sie der praktischen Berechnung zugänglich zu machen ist Quelle weiterer hochdimensionaler Vektorräume. Dimensionen von Millionen oder Milliarden sind im 21. Jahrhundert keine Seltenheit.

Außerdem haben Sie schon in der Schule in unendlichdimensionalen Vektorräumen gerechnet ohne das es Ihnen aufgefallen wäre. Man braucht also keine Angst zu haben.

Wir zitieren noch

Satz 1.7 (Basisergänzungssatz).

Sei U ein Unterraum von $\mathbb{R}^n$. Dann gilt $0 \leq \dim(U) \leq n$. Weiterhin läßt sich jede Basis von U zu einer Basis von $\mathbb{R}^n$ ergänzen.

Beweis. Es gibt eine Menge von n linear unabhängigen Vektoren, beispielsweise $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$, die eine Basis von $\mathbb{R}^n$ sind. Es sei $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ eine Basis von U . Entweder ist dies auch eine Basis von $\mathbb{R}^n$, also $r = n$, oder nicht. Dann ist

$$\{\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{e}_1, \dots, \mathbf{e}_n\}$$

jedenfalls ein Erzeugendensystem von $\mathbb{R}^n$. Es gibt einen Vektor $\mathbf{e}_i$ der sich als Linearkombination der vorangehenden Vektoren darstellen läßt; sonst wäre die Menge ja unabhängig. Dann können wir die Menge

$$\{\mathbf{v}_1, \dots, \mathbf{v}_r, \mathbf{e}_1, \dots, \mathbf{e}_{i-1}, \mathbf{e}_{i+1}, \dots, \mathbf{e}_n\}$$

betrachten. Dies ist entweder eine Basis, dann sind wir fertig, oder wir wiederholen die gleiche Konstruktion. Der Prozeß bricht offensichtlich nach endlich vielen Schritten mit einer Basis ab. $\square$

Definition 1.9 (Koordinaten).

Sei $(\mathbf{b}_1, \dots, \mathbf{b}_n)$ eine (geordnete) Basis des $\mathbb{R}^n$. Dann gibt es für jedes $\mathbf{x} \in \mathbb{R}^n$ eine Menge $\{\alpha_1, \dots, \alpha_n\} \subseteq \mathbb{R}$ mit

$$\mathbf{x} = \sum_{i=1}^n \alpha_i \mathbf{b}_i.$$

Die Zahlen $\{\alpha_1, \dots, \alpha_n\}$ heißen dann **Koordinaten** von $\mathbf{x}$ bzgl. der Basis $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ und wir schreiben

$$\mathbf{x}_B = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

Bemerkung 1.3.

Die Leserin bemerkt, daß die Reihenfolge der α_i wird durch die Reihenfolge der $\mathbf{b}_i$ festgelegt wird. Darum muß man die Reihenfolge der $\mathbf{b}_i$ mit spezifizieren. Andernfalls ist nicht klar, wie bspw. der **Koordinatenvektor**

$$\mathbf{x}_B = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

bzgl. der Basis

$$\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\}$$

zu lesen ist. Meinen wir

$$1 \cdot \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + 2 \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 3 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

oder vielleicht

$$1 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 2 \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 3 \cdot \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Wenn wir statt der Mengenschreibweise $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ die Tupelschreibweise $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ verwenden, dann ist klar, daß wir

$$1 \cdot \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + 2 \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 3 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

meinen.

1.2.1 Spalten vs. Zeilen

Natürlich ist es für die obigen Betrachtungen völlig egal ob wir die Vektoren als Spalten, wie ich es oben gemacht habe, oder als Zeilen schreiben so lange wir nicht versuchen die beiden zu mixen. Es gibt gute Gründe für die gewählte Darstellung aber die sind ästhetischer Natur.

Wenn es sich auch Platzgründen anbietet, werden wir von hier an ab und zu $[x_1, \dots, x_n]^T$ schreiben wenn wir einen Spaltenvektor meinen. Das es sich um einen Spaltenvektor handelt sieht man an dem T das für **transponiert** steht: eine Zeile transponiert ist eine Spalte, eine Spalte transponiert ist eine Zeile.

1.3 Der Vektorraum $\mathbb{K}^n$

Alle Konstruktionen und Sätze die wir in diesem Kapitel diskutiert haben, gelten weiterhin, wenn man $\mathbb{R}$ durch $\mathbb{Q}$ oder $\mathbb{C}$ bzw. einen allgemeinen Körper $\mathbb{K}$ ersetzt. In keinem unserer Argumente haben wir Eigenschaften von $\mathbb{R}$ benutzt, die über die Körper-Eigenschaften hinaus gehen. Schreiben Sie am besten ein paar Sachen wie lineare (Un)Abhängigkeit oder lineare Hülle in $\mathbb{C}$ auf um zu trainieren die Resultate zu übersetzen. Zur Wiederholung

Definition 1.10 (Körper).

*Ein **Körper** ist eine Menge $\mathbb{K}$ mit zwei binären Operationen*

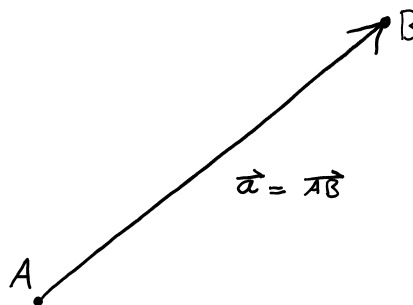
$$\left\{ \begin{array}{l} +: \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K}, \\ (\alpha, \beta) \mapsto \alpha + \beta \end{array} \right. \quad \text{und} \quad \left\{ \begin{array}{l} \cdot: \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K}, \\ (\alpha, \beta) \mapsto \alpha \cdot \beta. \end{array} \right.$$

die die folgenden Axiome erfüllen:

- (i) **Null:** Es existiert ein Element $0 \in \mathbb{K}$ mit $\alpha + 0 = \alpha$ für alle $\alpha \in \mathbb{K}$.
- (ii) **Add. Inverses:** Für alle $\alpha \in \mathbb{K}$ existiert ein Element $-\alpha \in \mathbb{K}$ mit $\alpha + (-\alpha) = 0$.
- (iii) **Kommutativität:** Für alle $\alpha, \beta \in \mathbb{K}$ gilt $\alpha + \beta = \beta + \alpha$.
- (iv) **Assoziativität:** Für alle $\alpha, \beta, \gamma \in \mathbb{K}$ gilt $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$.
- (v) **Eins:** Es existiert ein Element $1 \in \mathbb{K}$ mit $1\alpha = \alpha$ für alle $\alpha \in \mathbb{K}$.
- (vi) **Mult. Inverses:** Für alle $\alpha \in \mathbb{K} \setminus \{0\}$ existiert ein Element $\alpha^{-1} \in \mathbb{K}$ mit $\alpha \cdot \alpha^{-1} = 1$.
- (vii) **Kommutativität:** Für alle $\alpha, \beta \in \mathbb{K}$ gilt $\alpha\beta = \beta\alpha$.
- (viii) **Assoziativität:** Für alle $\alpha, \beta, \gamma \in \mathbb{K}$ gilt $(\alpha\beta)\gamma = \alpha(\beta\gamma)$.
- (ix) **Distributivität:** Für alle $\alpha, \beta, \gamma \in \mathbb{K}$ gilt $\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$.

1.4 Anschauliche Vektorrechnung

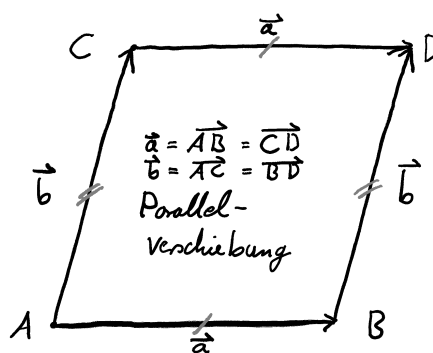
Wir arbeiten im 2 bzw. 3 dimensionalen **Anschauungsraum**, den wir mit $\mathbb{R}^2$ bzw. $\mathbb{R}^3$ identifizieren. Ein geordnetes Punktepaar (A, E) definiert einen Vektor $\overrightarrow{AE}$, den wir als **gerichtete**² **Strecke** darstellen.



Wir nennen A den **Anfangspunkt** und E den **Endpunkt** der gerichteten Strecke von A nach B .

Hinsichtlich der **Gleichheit von Vektoren** treffen wir die folgende Vereinbarung:

Zwei Vektoren $\vec{x}$ und $\vec{y}$ sind gleich, wenn sie durch **Parallelverschiebung**³ auseinander hervorgehen.



²d.h. eine Strecke mit einer angezeichneten Richtung. Hier verwenden wir Richtung im anschaulichen Sinne. Eine mathematisch korrekte Definition geben wir später. Siehe dazu das Kapitel zum inneren Produkt/Skalar-Produkt.

³Wieder berufen wir uns auf die Anschauung. Parallelverschiebung hat bis jetzt keine math. Bedeutung aber Sie erinnern sich aus der Schulgeometrie, was es bedeutet.

Streng genommen bedeutet diese Festlegung natürlich, daß wir nicht mehr von **dem** Anfangspunkt und **dem** Endpunkt eines Vektors $\vec{a}$ sprechen können. Dennoch werden wir diese Sprechweise beibehalten. Wir meinen dann natürlich den Anfangs und Endpunkt der gerichteten Strecke die wir als Repräsentation des Vektors $\vec{a} = \overrightarrow{AE}$ gewählt haben.

Wir setzen:

Ein Vektor ist die Gesamtheit aller Strecken von gleicher Richtung und gleicher Länge.

Behalten Sie im Hinterkopf, daß diese Definition unscharf ist. Es ist weder klar was eine Richtung ist noch was eine Länge sein soll. Wie mißt man die denn? Wir berufen uns hier rein auf die Anschauung! Wir benutzen dies einfach als 'da wo der Pfeil hin zeigt' und auch Länge benutzen wir anschaulich als Abstand der Punkte, was auch immer das heißt. Wenn $A = B$, dann heißt der Vektor $\vec{0} = \overrightarrow{AB}$ der Nullvektor. Der Nullvektor hat natürlich keine definierte Richtung.

Wir bezeichnen die **Länge eines Vektors** $\vec{a} = \overrightarrow{AB}$, d.h den Abstand der beiden Punkte A und B bezeichnen wir mit $|\vec{a}|$ ⁴ und wird auch der **Betrag** von $\vec{a}$ genannt. Der Nullvektor hat natürlich den Betrag 0. Diese Eigenschaft ist charakterisierend:

$$\vec{x} = \vec{0} \quad \Leftrightarrow \quad |\vec{x}| = 0.$$

Dies ist anschaulich klar, da eine Strecke nur dann die Länge 0 hat (gemessen mit Pythagoras), wenn Anfangs und Endpunkt übereinstimmen.

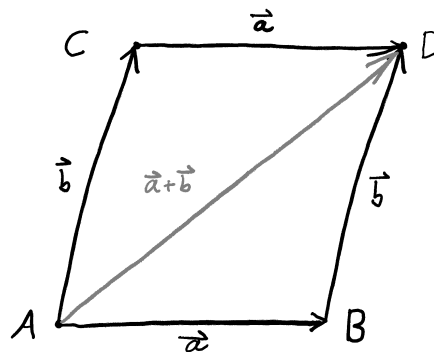
Wir führen jetzt die Addition und die Skalarmultiplikation unserer Vektoren ein. Am ende der kommenden beiden Sektionen sollten Sie (A1)-(A4) und (M1)-(M4) mit den korrespondierenden Eigenschaften in Satz 1.1 vergleichen.

1.4.1 Addition von Vektoren

Wir setzen für die **Summe zweier Vektoren**:

⁴Später werden wir $\|a\|_2$ verwenden. Siehe dazu auch die Sektion zu Normen.

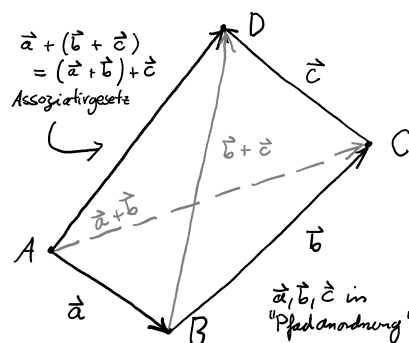
Für gegebene Vektoren $\vec{a}$ und $\vec{b}$ wird die **Summe** $\vec{a} + \vec{b}$ gebildet, indem man den Anfangspunkt von $\vec{a}$ an den Endpunkt von $\vec{b}$ legt und die gerichtete Strecke vom Anfangspunkt von $\vec{b}$ und dem Endpunkt von $\vec{a}$ nimmt. (Parallelogrammregel)



Damit gilt: Wenn $\vec{a} = \overrightarrow{AB}$ und $\vec{b} = \overrightarrow{BC}$, dann gilt $\vec{a} + \vec{b} = \overrightarrow{AC}$. Es gilt aber natürlich auch $\vec{b} = \overrightarrow{BD}$ und $\vec{a} = \overrightarrow{CD}$ und damit $\vec{b} + \vec{a} = \overrightarrow{AC}$. Also gilt für die Addition wie oben definiert das **Kommutativgesetz**:

$$\vec{a} + \vec{b} = \vec{b} + \vec{a}. \quad (\text{A1})$$

Die Summe von drei Vektoren $\vec{a}, \vec{b}, \vec{c}$ ergibt sich durch die folgende Figur:



Damit erhalten wir das **Assoziativgesetz**

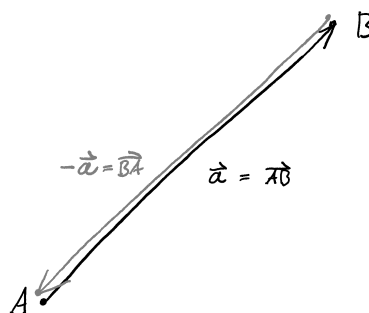
$$\vec{a} + (\vec{b} + \vec{c}) = (\vec{a} + \vec{b}) + \vec{c}. \quad (\text{A2})$$

Ist $\vec{a} = \overrightarrow{AB}$, dann ist $\vec{0} = \overrightarrow{AA} = \overrightarrow{BB}$ und es folgt

$$\vec{a} + \vec{0} = \overrightarrow{AB} + \overrightarrow{BB} = \overrightarrow{AB} = \vec{a} \quad (\text{A3})$$

und in gleicher Weise $\vec{0} + \vec{a} = \vec{a}$. Also ist $\vec{0}$ das **neutrale Element** der von uns oben definierten Addition.

Für $\vec{a} = \overrightarrow{AB}$ definieren wir den Vektor $-\vec{a} = \overrightarrow{BA}$ und nennen ihn den **negativen Vektor** von $\vec{a}$.



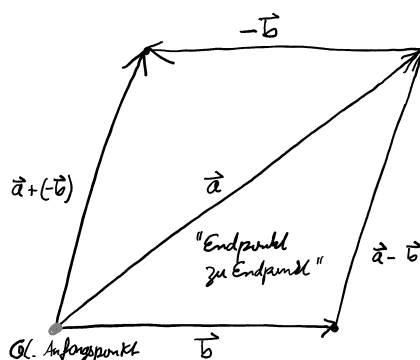
Es gilt

$$\vec{a} + (-\vec{a}) = \overrightarrow{AB} + \overrightarrow{BA} = \overrightarrow{AA} = \vec{0}. \quad (\text{A4})$$

Damit ist $-\vec{a}$ das **additiv Inverse** von $\vec{a}$. (Diese Eigenschaft ist charakterisierend.)

Wir definieren die **Differenz zweier Vektoren** $\vec{a} - \vec{b}$ zweier Vektoren durch

$$\vec{a} - \vec{b} = \vec{a} + (-\vec{b}).$$



Merkregel für die Differenz:

$\vec{a} - \vec{b}$ ist der Vektor, der vom Endpunkt von $\vec{b}$ zum Endpunkt von $\vec{a}$, wenn $\vec{a}$ und $\vec{b}$ so aneinandergelegt sind, daß ihre Anfangspunkte zusammenfallen.

Es gilt

$$(\vec{a} - \vec{b}) + \vec{b} = \vec{0}. \quad (1.4.1)$$

Zum einen kann man dies am Bild erkennen und zum anderen nachrechnen:

$$\begin{aligned} (\vec{a} - \vec{b}) + \vec{b} &= (\vec{a} + (-\vec{b})) + \vec{b} \quad (\text{Definition Differenz}) \\ &= \vec{a} + (\vec{b} + (-\vec{b})) \quad (\text{A2}) \\ &= \vec{a} + \vec{0} \quad (\text{Definition Add. Inverses}) \\ &= \vec{a} \quad (\text{A3}). \end{aligned}$$

Damit ist die Differenz (Subtraktion) von Vektoren wie oben eingeführt die Umkehroperation der Addition.

1.4.2 Multiplikation mit Skalar

Sei $\lambda \in \mathbb{R}$ mit $\lambda > 0$ und ein Vektor $\vec{x}$ vorgelegt. (Wir nennen reelle Zahlen in diesem Zusammenhang **Skalare**). Wir definieren den Vektor $a\vec{x}$ als denjenigen Vektor, der in die gleiche Richtung zeigt wie $\vec{x}$, aber a -mal so lang ist (also kürzer (gestaucht) für $\lambda < 1$ und länger (gestreckt) für $a > 1$), d.h. $|\lambda\vec{x}| = \lambda|\vec{x}|$.

Wir setzen $0\vec{x} = \vec{0}$ und erklären weiter für $\lambda > 0$ den Vektor $(-\lambda)\vec{x}$ als denjenigen Vektor, der die entgegengesetzte Richtung von $\vec{x}$ hat aber seine a -fache Länge, also $|(-\lambda)\vec{x}| = a|\vec{x}|$. Es gilt also allgemein $|\lambda\vec{x}| = |\lambda||\vec{x}|$ wobei $|\lambda|$ der Betrag der reellen Zahl λ und $|\vec{x}|$ der Betrag von $\vec{x}$ ist. Beachten Sie, daß beides völlig verschiedene Sachen sind auch wenn Sie die gleichen Symbole haben. Die gleichen Symbole rühren daher das es sich konzeptionell (Konzept einer Länge, später allgemeiner Norm) um das gleiche handelt.

Wie in der letzten Sektion lassen sich für diese Definition die folgenden Eigenschaften prüfen:

$$\lambda(\vec{x} + \vec{y}) = \lambda\vec{x} + \lambda\vec{y} \quad (\text{M1})$$

$$(\lambda + \mu)\vec{x} = \lambda\vec{x} + \mu\vec{x} \quad (\text{M2})$$

$$\lambda(\mu\vec{x}) = (\lambda\mu)\vec{x} \quad (\text{M3})$$

$$1\vec{x} = \vec{x}. \quad (\text{M4})$$

Auch hier sei nochmals darauf hingewiesen, daß $+$ und $\cdot$ hier jeweils andere Operationen sind auch wenn wir stets das gleiche Symbol verwenden. Sie sollten jetzt langsam ein Gefühl dafür haben.

Übungsaufgabe 1.4.

Zeigen Sie die Eigenschaften (M1)–(M4) mit Hilfe der Definition.

Wir verabreden noch folgende Schreibweise: Wenn $\lambda \neq 0$ ist, dann schreiben wir $\frac{\vec{x}}{\lambda}$ für $\frac{1}{\lambda}\vec{x}$. Mit $\vec{x} \neq 0$ haben wir, daß $\frac{\vec{x}}{|\vec{x}|}$ die Länge 1. Man nennt Vektoren $\vec{x}$ mit $|\vec{x}| = 1$ **Einheitsvektoren**. Den Übergang von $\vec{x}$ zu seinem **Einheitsvektor** $\frac{\vec{x}}{|\vec{x}|}$ nennt man **Normierung**.

1.5 Verständnisfragen

1. Wie begründet man gegebenenfalls, daß die angegebenen Mengen Unterräume von $\mathbb{R}^3$ sind oder nicht.

a) $\left\{ \begin{bmatrix} x_1 \\ 0 \\ x_2 \end{bmatrix} : x_1, x_2 \in \mathbb{R} \right\},$

c) $\left\{ \lambda \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} : \lambda \in \mathbb{Z} \right\},$

b) $\left\{ \lambda \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} : \lambda \in \mathbb{R} \right\},$

d) $\left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} : x_i \in \mathbb{R}, x_1 + x_2 + x_3 = 0 \right\}.$

2. Ist $\{\mathbf{a}, \mathbf{b}\} \subseteq \mathbb{R}^n$ genau dann linear abhängig, wenn $\mathbf{a} = \alpha \mathbf{b}$ für ein $\alpha \in \mathbb{R}$ gilt?
3. Wie kann man die lineare Abhängigkeit bzw. Unabhängigkeit von $\{\mathbf{a}, \mathbf{b}\} \subseteq \mathbb{R}^2$ und $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\} \subseteq \mathbb{R}^3$ geometrisch interpretieren?
4. Läßt sich jeder Vektor $\mathbf{c} \in \mathbb{R}^3$ als Linearkombination der Vektoren $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ darstellen, wenn $\{\mathbf{a}, \mathbf{b}\}$ als linear unabhängig vorausgesetzt wird?
5. Wie begründet man, daß jeder Vektor $\mathbf{x} \in \mathbb{R}^n$ sich als eindeutige Linearkombination der Elemente einer Basis $B \subseteq \mathbb{R}^n$ von $\mathbb{R}^n$ darstellen läßt?
6. Was versteht man unter den Koordinaten eines Vektors $\mathbf{x} \in \mathbb{R}^n$ bezüglich einer Basis $B \subseteq \mathbb{R}^n$ von $\mathbb{R}^n$? Mit welcher Begründung bezeichnet man beispielsweise die Zahlen x_1, x_2, x_3 eines Vektors $\mathbf{x} = [x_1, x_2, x_3]^T$ also dessen Koordinaten?
7. Wie beweist man, daß die Elemente der Standard-Basis von $\mathbb{R}^n$ eine Basis des $\mathbb{R}^n$ bilden?
8. Wie begründet man, daß $\mathbf{x} \in \mathbb{R}^n$ eine Linearkombination von $S = \{\mathbf{x}_1, \dots, \mathbf{x}_k\} \subseteq \mathbb{R}^n$ ist, wenn S linear unabhängig, aber $S' = S \cup \{\mathbf{x}\}$ linear abhängig ist?

1.6 Rechenaufgaben

1. Für welche $\lambda \in \mathbb{R}$ gilt $[1, 0, 3]^T = \lambda[2, 0, 6]^T$.
2. Geben Sie die zu folgenden Pfeilen gehörenden Vektoren an:
 - a) $\overrightarrow{OP}$ mit $O = (0, 0)$, $P = (-1, 3)$,
 - b) $\overrightarrow{OX}$ mit $O = (0, 0, 0)$, $X = (0, 1, 0)$,
 - c) $\overrightarrow{AB}$ mit $A = (-1, 0, 2)$, $B = (1, 1, 3)$.
3. Es seien $A_1 = (2, 2, 0)$, $A_2 = (2, 5, 0)$, $A_3 = (4, 2, 0)$ und $A_4 = (2, 2, 2)$ die Eckpunkte eines Tetraeders und $\overrightarrow{A_i A_{i+1}} =: \mathbf{a}_i$, $i = 1, 2, 3$ drei Kantenpfeile bzw. Vektoren. Wählen Sie für die dabei nicht berücksichtigten Kanten einen Durchlaufsinne, und stellen Sie die sich so ergebenden Kantenpfeile bzw. Vektoren als Linearkombination der $\mathbf{a}_i$, $i = 1, 2, 3$ dar!
4. Gegeben sind die Vektoren $\mathbf{a} = [2, -2, 1]^T$ und $\mathbf{b} = [4, 1, -3]^T$. Berechnen Sie $\mathbf{a} + \mathbf{b}$, $2\mathbf{a} + 5\mathbf{b}$, $5\mathbf{b} - 3\mathbf{a}$.

Matrizen

Zur Wiederholung

Definition 2.1 (Körper).

Ein **Körper** ist eine Menge $\mathbb{K}$ mit zwei binären Operationen

$$\left\{ \begin{array}{l} +: \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K}, \\ (\alpha, \beta) \mapsto \alpha + \beta \end{array} \right. \quad \text{und} \quad \left\{ \begin{array}{l} \cdot: \mathbb{K} \times \mathbb{K} \rightarrow \mathbb{K}, \\ (\alpha, \beta) \mapsto \alpha \cdot \beta. \end{array} \right.$$

die die folgenden Axiome erfüllen:

- (i) **Null:** Es existiert ein Element $0 \in \mathbb{K}$ mit $\alpha + 0 = \alpha$ für alle $\alpha \in \mathbb{K}$.
- (ii) **Add. Inverse:** Für alle $\alpha \in \mathbb{K}$ existiert ein Element $-\alpha \in \mathbb{K}$ mit $\alpha + (-\alpha) = 0$.
- (iii) **Kommutativität:** Für alle $\alpha, \beta \in \mathbb{K}$ gilt $\alpha + \beta = \beta + \alpha$.
- (iv) **Assoziativität:** Für alle $\alpha, \beta, \gamma \in \mathbb{K}$ gilt $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$.
- (v) **Eins:** Es existiert ein Element $1 \in \mathbb{K}$ mit $1\alpha = \alpha$ für alle $\alpha \in \mathbb{K}$.
- (vi) **Mult. Inverse:** Für alle $\alpha \in \mathbb{K} \setminus \{0\}$ existiert ein Element $\alpha^{-1} \in \mathbb{K}$ mit $\alpha \cdot \alpha^{-1} = 1$.
- (vii) **Kommutativität:** Für alle $\alpha, \beta \in \mathbb{K}$ gilt $\alpha \cdot \beta = \beta \cdot \alpha$.
- (viii) **Assoziativität:** Für alle $\alpha, \beta, \gamma \in \mathbb{K}$ gilt $(\alpha \cdot \beta) \cdot \gamma = \alpha \cdot (\beta \cdot \gamma)$.
- (ix) **Distributivität:** Für alle $\alpha, \beta, \gamma \in \mathbb{K}$ gilt $\alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$.

Bemerkung 2.1.

Für uns sind die Körper $\mathbb{Q}$, $\mathbb{R}$ und $\mathbb{C}$ die wichtigsten. Alles funktioniert aber auch mit endlichen Körpern wie $\mathbb{Z}_2$. Siehe Kapitel 2 Mathematik 1.

2.1 Definition und Beispiele

Definition 2.2 (Matrix).

Für Zahlen $a_{i,j} \in \mathbb{K}$, $1 \leq i \leq m$, $1 \leq j \leq n$ heißt das Zahlenschema

$$\mathbf{A} := \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} =: [a_{i,j}]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$$

eine **Matrix** vom Format $m \times n$ mit Einträgen/Koeffizienten in $\mathbb{K}$.

Die Zahlen $a_{i,j}$ heißen der (i, j) -te Eintrag/**Koeffizient** in $\mathbf{A}$.

Die Menge aller solcher Matrizen vom Format $m \times n$ mit $\mathbb{K}$ -Koeffizienten wird mit $\mathbb{K}^{m \times n}$ bezeichnet.

Schreibregelung: Wir setzen in $[a_{i,j}]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ stets **Zeile vor Spalte**. Wir bezeichnen die Spalten von $\mathbf{A}$ mit $\mathbf{A}[\cdot, j]$ für $j = 1, \dots, n$ und die Zeilen von $\mathbf{A}$ mit $\mathbf{A}[i, \cdot]$, $i = 1, \dots, m$. Auch hier geht Zeile vor Spalte um sich die Notation zu merken.

Sprachregelungen: Wir benennen ein paar spezielle Matrizen:

(i) Wenn $m = n$, dann nennen wir die Matrix **quadratisch**. Wir schreiben

$$\text{auch } A = [a_{i,j}]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} =: [a_{i,j}]_{i,j=1}^n$$

(ii) Die Matrix

$$\mathbf{I}_n := [a_{i,j}]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}} = [\delta_{i,j}]_{\substack{1 \leq i \leq n \\ 1 \leq j \leq n}}$$

, d.h.

$$\mathbf{I}_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

heißt $n \times n$ **Einheitsmatrix**.

(iii) Eine Matrix der Form

$$\mathbf{A} := \text{diag}(\lambda_1, \dots, \lambda_n) := [\lambda_i \delta_{i,j}]_{\substack{1 \leq i \leq n, \\ 1 \leq j \leq n}},$$

d.h.

$$\mathbf{A} = \text{diag}(\lambda_1, \dots, \lambda_n) = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}$$

für $\lambda_i \in \mathbb{K}$, $i = 1, \dots, n$ heißt **Diagonalmatrix**.

Es gilt natürlich $I_n = \text{diag}(1, \dots, 1)$.

(iv) Eine Matrix $[a_{i,j}]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ mit $a_{i,j} = 0$ für $j < i$, d.h. von der Form

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ 0 & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{m,n} \end{bmatrix}$$

heißt **obere Dreiecksmatrix**.

Als nächstes führen wir noch ein paar weitere Begriffe ein um Matrizen zu klassifizieren. Wir beginnen mit

Definition 2.3 (Transponierte Matrix).

Sei $\mathbf{A} = [a_{i,j}]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ eine Matrix mit Koeffizienten in $\mathbb{K}$.

Dann heißt die Matrix

$$[a_{j,i}]_{\substack{1 \leq i \leq m, \\ 1 \leq j \leq n}} \quad (2.1.1)$$

in Zeichen $\mathbf{A}^T$, die transponierte Matrix zu $\mathbf{A}$.^a

^aMan schreibt Spalten als Zeilen, bzw. Zeilen als Spalten.

Definition 2.4 ((Anti)Symmetrische Matrizen).

Sei $\mathbf{A} = [a_{i,j}]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ eine Matrix mit Koeffizienten in $\mathbb{K}$.

Dann heißt $\mathbf{A}$ genau dann

(i) **symmetrisch**, wenn $\mathbf{A} = \mathbf{A}^T$.

(ii) **antisymmetrisch**, wenn $\mathbf{A} = -\mathbf{A}^T$.

Antisymmetrische Matrizen nennt man auch **chiefsymmetrisch** (skew symmetric).

Definition 2.5 (Adjungierte Matrix).

Sei $\mathbf{A} = [a_{i,j}]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ eine Matrix mit Koeffizienten in $\mathbb{C}$. Dann heißt die Matrix

$$[\overline{a_{j,i}}]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} = \overline{\mathbf{A}^T} = \overline{\mathbf{A}}^T,$$

in Zeichen $\mathbf{A}^*$ oder $\mathbf{A}^H$, die adjungierte Matrix zu $\mathbf{A}$.^a

^aMan schreibt Spalten als Zeilen, bzw. Zeilen als Spalten und konjugiert die Elemente.

Definition 2.6 ((Anti)Hermitesche Matrix).

Sei $\mathbf{A} = [a_{i,j}]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ eine Matrix mit Koeffizienten in $\mathbb{C}$. Dann heißt A genau dann

(i) **hermitesch**, wenn $\mathbf{A} = \mathbf{A}^H$.

(ii) **antihermitesch**, wenn $\mathbf{A} = -\mathbf{A}^H$.

Man nennt solche Matrizen auch **schiefhermitesch** (skew hermitian).

Es bleibt der Leserin als Übungsaufgabe überlassen die behaupteten Eigenschaften im folgenden Satz nachzurechnen.

Satz 2.1.

Für $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{m \times n}$ und $\lambda \in \mathbb{K}$ gelten:

(i) $(\mathbf{A}^T)^T = \mathbf{A}$,

(ii) $(\lambda \mathbf{A})^T = \lambda \mathbf{A}^T$ und

(iii) $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$.

2.2 Matrizen als Vektorraum

Wir definieren auf der Menge $\mathbb{K}^{m \times n}$ die folgenden **Vektorraumoperationen**:

(i) **Vektoraddition**:

$$\left\{ \begin{array}{ll} +: \mathbb{K}^{m \times n} \times \mathbb{K}^{m \times n} & \rightarrow \mathbb{K}^{m \times n}, \\ (\mathbf{A}, \mathbf{B}) & \mapsto \mathbf{A} + \mathbf{B} \end{array} \right.$$

definiert durch

$$\left[a_{i,j} \right]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} + \left[b_{i,j} \right]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} := \left[a_{i,j} + b_{i,j} \right]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$$

Dabei benutzen wir auf der rechten Seite einfach die Addition im Körper $\mathbb{K}$.

(ii) **Skalarmultiplikation**:

$$\left\{ \begin{array}{ll} \cdot: \mathbb{K} \times \mathbb{K}^{m \times n} & \rightarrow \mathbb{K}^{m \times n}, \\ (\lambda, \mathbf{A}) & \mapsto \lambda \cdot \mathbf{A}. \end{array} \right.$$

definiert durch

$$\lambda \left[a_{i,j} \right]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} := \left[\lambda a_{i,j} \right]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$$

Dabei benutzen wir auf der rechten Seite einfach die Multiplikation im Körper $\mathbb{K}$.

Weiterhin setzen wir

$$\mathbf{0} := \left[0 \right]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \quad \text{und} \quad -\mathbf{A} := \left[-a_{i,j} \right]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}.$$

Definition 2.7 (Der Vektorraum $\mathbb{K}^{m \times n}$).

Die Menge $\mathbb{K}^{m \times n}$ der $m \times n$ Matrizen mit Koeffizienten in $\mathbb{K}$, ausgestattet mit den oben definierten Operationen heißt der Vektorraum $\mathbb{K}^{m \times n}$ der $m \times n$ Matrizen.

Man rechnet leicht nach, wie im Falle $\mathbb{K}^n$ (Sektion 1.1), daß

Satz 2.2.

Die Vektorraumoperationen des Vektorraumes $\mathbb{K}^{m \times n}$ erfüllen die folgenden Eigenschaften:

- (i) Für alle $\mathbf{A} \in \mathbb{K}^{m \times n}$ gilt $\mathbf{A} + \mathbf{O} = \mathbf{A}$.
- (ii) Für alle $\mathbf{A} \in \mathbb{K}^{m \times n}$ gilt $\mathbf{A} + (-\mathbf{A}) = \mathbf{O}$.
- (iii) Für alle $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{m \times n}$ gilt $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$.
- (iv) Für alle $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{K}^{m \times n}$ gilt $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$.
- (v) Für alle $\mathbf{A} \in \mathbb{K}^{m \times n}$ gilt $1\mathbf{A} = \mathbf{A}$.
- (vi) Für alle $\lambda, \mu \in \mathbb{K}$ und alle $\mathbf{A} \in \mathbb{K}^{m \times n}$ gilt $\lambda(\mu\mathbf{A}) = (\lambda\mu)\mathbf{A}$.
- (vii) Für alle $\lambda \in \mathbb{K}$ und alle $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{m \times n}$ gilt $\lambda(\mathbf{A} + \mathbf{B}) = \lambda\mathbf{A} + \lambda\mathbf{B}$.
- (viii) Für alle $\lambda, \mu \in \mathbb{K}$ und alle $\mathbf{A} \in \mathbb{K}^{m \times n}$ gilt $(\lambda + \mu)\mathbf{A} = \lambda\mathbf{A} + \mu\mathbf{A}$.

Bemerkung 2.2.

Wie in unserer Konstruktion von $\mathbb{K}^n$ haben wir auch hier einfach auf einer gegebenen Menge von Objekten Operationen definiert und Eigenschaften gezeigt. Wir haben dann einfach festgelegt, daß wir diese Objekte, zusammen mit den Operationen jeweils Vektorraum nennen. Dies geschieht im Vorgriff auf die abstrakte Definition in Kapitel 8. Unser Konstruktionen in Kapitel 1 und 2 zeigen dann, daß hier tatsächlich Vektorräume nach Definition 8.1 sind. Machen Sie sich die Mühe und vergleichen die Sätze 2.2 und 1.1 und stellen

Sie fest, daß die Struktur auf den beiden Mengen genau die gleiche ist; das ist auch der Grund für die abstrakte Definition. Alles was man nur mit den Vektorraumeigenschaften zeigen kann, gilt dann automatisch für alle Vektorräume. Außerdem schafft es Ordnung in einem sonst kaum überschaubaren Zoo an wichtigen Objekten.

Definition 2.8 (Standardbasis).

Die **Standard-Basis** (siehe auch Definition 1.4) von $\mathbb{R}^{m \times n}$ ist gegeben durch

$$E_{ij} = [\delta_{i,k} \delta_{j,l}]_{\substack{1 \leq k \leq m \\ 1 \leq l \leq n}} \quad 1 \leq i \leq m, 1 \leq j \leq m$$

wobei

$$\delta_{ij} = \begin{cases} 1 & : i = j \\ 0 & : i \neq j \end{cases}$$

Damit gilt

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

Wir können $\mathbb{R}^{m \times n}$ und $\mathbb{R}^{m \cdot n}$ mit der Abbildung

$$[a_{i,j}]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \mapsto [a_{11}, \dots, a_{1,n}, a_{2,1}, \dots, a_{2,n}, \dots, a_{m,1}, \dots, a_{m,n}]^T$$

identifizieren.

2.3 Multiplikation von Matrizen

Die Matrizen tragen mehr Struktur als nur die eines Vektorraumes. Die Matrizen tragen eine natürliche Definition der Multiplikation.¹

Als erstes wollen wir sehen, ob wir eine Matrix mit einem Vektor multiplizieren können. Moment. Hatten wir nicht gerade gesagt das Matrizen Vektoren sind? Das ist richtig, da haben Sie gut aufgepaßt. Es hat sich allerdings eingebürgert, Matrizen in diesem Zusammenhang nicht Vektoren zu nennen, auch wenn Sie Vektoren sind. Man bemerke auch, daß Vektoren natürlich Matrizen sind. Nach unserer Definition ist der Spaltenvektor $\begin{bmatrix} x_1, \dots, x_n \end{bmatrix}^T$ eine $n \times 1$ Matrix und der Zeilenvektor $\begin{bmatrix} x_1, \dots, x_n \end{bmatrix}$ eine $1 \times n$ -Matrix.

Okay, nun zur Definition von $\mathbf{Ax}$ für $\mathbf{A} = [a_{i,j}]_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ und $\mathbf{x} = [x_i]_{1 \leq i \leq n}$. Es ist wichtig, daß die Anzahl der Spalten von A gleich der Anzahl der Zeilen von $\mathbf{x}$ ist. Wir setzen

$$\mathbf{Ax} := \begin{bmatrix} \sum_{j=1}^n a_{1,j}x_j \\ \vdots \\ \sum_{j=1}^n a_{m,j}x_j \end{bmatrix}.$$

¹Das wird uns in der Definition vielleicht nicht so vorkommen aber wir werden das später verstehen. Jetzt wollen wir erst einmal rechnen.

Übungsaufgabe 2.1.*Berechnen Sie*

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & -1 & 1 \\ 3 & 0 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}.$$

Nachdem Sie die Übungsaufgabe gelöst haben, schauen wir uns $\mathbf{Ax}$ etwas genauer an um besser zu verstehen, was da eigentlich passiert:

$$\begin{aligned} \mathbf{Ax} &= \begin{bmatrix} \sum_{j=1}^n a_{1,j}x_j \\ \vdots \\ \sum_{j=1}^n a_{m,j}x_j \end{bmatrix} \\ &= \begin{bmatrix} a_{1,1} \\ \vdots \\ a_{m,1} \end{bmatrix} x_1 + \begin{bmatrix} a_{1,2} \\ \vdots \\ a_{m,2} \end{bmatrix} x_2 + \dots + \begin{bmatrix} a_{1,n} \\ \vdots \\ a_{m,n} \end{bmatrix} x_n. \end{aligned}$$

Diese Gleichung bedeutet, daß das **Matrix-Vektor-Produkt** eine **Linear-kombination** der Spalten der Matrix $\mathbf{A}$ mit den Koeffizienten $\{x_1, \dots, x_n\}$ ist. Mit diesen Vorbereitungen führen wir nun die allgemeine Definition des Matrix-Produktes ein, wo wir für zwei Matrizen

$$\mathbf{A} \cdot \mathbf{B} = \begin{bmatrix} a_{1,1} & \dots & a_{1,n} \\ \vdots & & \vdots \\ a_{m,1} & \dots & a_{m,n} \end{bmatrix} \cdot \begin{bmatrix} b_{1,1} & \dots & b_{1,k} \\ \vdots & & \vdots \\ b_{n,1} & \dots & b_{n,k} \end{bmatrix} = \begin{bmatrix} c_{1,1} & \dots & c_{1,k} \\ \vdots & & \vdots \\ c_{m,1} & \dots & c_{m,k} \end{bmatrix}$$

einfach die obige Multiplikation spaltenweise mit den Spalten von $\mathbf{B}$ durchführen. Dann erhalten wir eine $m \times k$ -Matrix als Resultat.

Definition 2.9 (Matrix-Produkt).

Für $\mathbf{A} = [a_{i,j}] \in \mathbb{K}^{m \times n}$ und $\mathbf{B} = [b_{i,j}] \in \mathbb{K}^{n \times p}$ ist das Produkt $\mathbf{C} = \mathbf{A} \cdot \mathbf{B} = [c_{i,j}] \in \mathbb{K}^{m \times p}$ definiert durch

$$c_{i,j} = \sum_{k=1}^n a_{i,k} b_{k,j} \quad (i = 1, 2, \dots, m \text{ und } j = 1, 2, \dots, p).$$

Bemerkung 2.3.

Wie am Anfang schon erklärt ist das Produkt $\mathbf{C} = \mathbf{AB}$ ist nur dann erklärt, wenn $\mathbf{A}$ so viele Spalten wie $\mathbf{B}$ Zeilen hat. In diesem Fall übernimmt das Ergebnis $\mathbf{C}$ die Zeilenanzahl von $\mathbf{A}$ und die Spaltenanzahl von $\mathbf{B}$, symbolisch:

$$\boxed{A} \cdot \boxed{B} = \boxed{C}$$

Bemerkung 2.4.

Wie schon am Anfang erklärt, ergibt sich sofort die folgende Interpretation der Matrixmultiplikation: Die Spalten der Matrix $\mathbf{AB}$ sind Linearkombinationen der Spalten von $\mathbf{A}$ wobei die Koeffizienten durch die Spalten von $\mathbf{B}$ gegeben sind.

Natürlich kann man die Zeilen des Produktes $\mathbf{AB}$ auch als Linearkombinationen der Zeilen von $\mathbf{B}$ auffassen wobei die Koeffizienten durch die Zeilen von $\mathbf{A}$ gegeben sind.

Satz 2.3 (Rechenregeln Matrixprodukt).

Es gelten die folgenden Rechenregeln für das Matrix-Produkt:

- (i) $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$ für alle $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{B} \in \mathbb{K}^{n \times p}$, $\mathbf{C} \in \mathbb{K}^{p \times q}$,
- (ii) $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$ für alle $\mathbf{A} \in \mathbb{K}^{m \times n}$, $\mathbf{B}, \mathbf{C} \in \mathbb{K}^{n \times p}$,
- (iii) $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$ für alle $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{m \times n}$, $\mathbf{C} \in \mathbb{K}^{n \times p}$,
- (iv) $\lambda(\mathbf{AB}) = (\lambda\mathbf{A})\mathbf{B} = \mathbf{A}(\lambda\mathbf{B})$ für alle $\lambda \in \mathbb{K}$, $\mathbf{A} \in \mathbb{K}^{m \times n}$, $\mathbf{B} \in \mathbb{K}^{n \times p}$,
- (v) $(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$ für alle $\mathbf{A} \in \mathbb{K}^{m \times n}$, $\mathbf{B} \in \mathbb{K}^{n \times p}$,

Satz 2.4.

Für die Multiplikation mit Einheits- und Nullmatrizen gelten:

- (i) Für die m -dimensionale **Einheitsmatrix**

$$\mathbf{I}_m := \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & & 0 & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \in \mathbb{R}^{m \times m},$$

gilt $\mathbf{I}_m \mathbf{A} = \mathbf{A}$ für alle $\mathbf{A} \in \mathbb{R}^{m \times n}$.

- (ii) Für die n -dimensionale Einheitsmatrix $\mathbf{I}_n \in \mathbb{R}^{n \times n}$ gilt $\mathbf{A} \mathbf{I}_n = \mathbf{A}$ für alle $\mathbf{A} \in \mathbb{R}^{m \times n}$,
- (iii) Für die wie gekennzeichnet dimensionierten Nullmatrizen gilt $\mathbf{A} \mathbf{O}_{n \times p} = \mathbf{O}_{m \times p}$ und $\mathbf{O}_{q \times m} \mathbf{A} = \mathbf{O}_{q \times n}$ für alle $\mathbf{A} \in \mathbb{R}^{m \times n}$.

2.3.1 Falk-Schema

Hier ein besprechen wir kurz das sogenannte Falk-Schema zur Berechnung von Matrix-Produkten

$$\begin{bmatrix} a_{i,j} \end{bmatrix}_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \cdot \begin{bmatrix} b_{i,j} \end{bmatrix}_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}} = \begin{bmatrix} c_{i,j} \end{bmatrix}_{\substack{1 \leq i \leq m \\ 1 \leq j \leq p}}.$$

Die Matrizen $\mathbf{A} = \begin{bmatrix} a_{i,j} \end{bmatrix}_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ und $\mathbf{B} = \begin{bmatrix} b_{i,j} \end{bmatrix}_{\substack{1 \leq i \leq n \\ 1 \leq j \leq p}}$ sind wie im unteren Bild anzuordnen. Die Berechnung der c_{ij} wird nach dem Bild erklärt.

						$b_{1,1}$	$\cdots$	$b_{1,j}$	$\cdots$	$b_{1,p}$
						$b_{2,1}$	$\cdots$	$b_{2,j}$	$\cdots$	$b_{2,p}$
						$\vdots$		$\vdots$		$\vdots$
						$\vdots$		$\vdots$		$\vdots$
						$b_{n,1}$	$\cdots$	$b_{n,j}$	$\cdots$	$b_{n,p}$
$a_{1,1}$	$a_{1,2}$	$\cdots$	$\cdots$	$a_{1,n}$		$c_{1,1}$	$\cdots$	$c_{1,j}$	$\cdots$	$c_{1,p}$
$\vdots$				$\vdots$		$\vdots$		$\vdots$		$\vdots$
$a_{i,1}$	$a_{i,2}$	$\cdots$	$\cdots$	$a_{i,n}$		$c_{i,1}$	$\cdots$	$c_{i,j}$	$\cdots$	$c_{i,p}$
$\vdots$				$\vdots$		$\vdots$		$\vdots$		$\vdots$
$a_{m,1}$	$a_{m,2}$	$\cdots$	$\cdots$	$a_{m,n}$		$c_{m,1}$	$\cdots$	$c_{m,j}$	$\cdots$	$c_{m,p}$

Nur die grau markierte Zeile bzw. Spalte von A und B geht in die Berechnung von $c_{i,j}$ ein. Folgendes Vorgehen:

- Bilden Sie 'von außen kommend' Zahlenpaare der Form $(a_{i,k}, b_{k,j})$.
- Multiplizieren Sie jeweils die beiden Zahlen und summieren Sie sämtliche Ergebnisse.

Übungsaufgabe 2.2.

Berechnen Sie für

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 3 & -1 & 2 \end{bmatrix} \quad \text{und} \quad \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 0 & -2 \\ 2 & 2 \end{bmatrix}$$

die Produkte $\mathbf{AB}$ und $\mathbf{BA}$ und vergleichen Sie die Ergebnisse. Was können Sie aus dem Vergleich schließen?

2.3.2 Warnungen

- Auch wenn beide Produkte $\mathbf{AB}$ und $\mathbf{BA}$ definiert sind (was beispielsweise für $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{n \times n}$ der Fall ist), gilt im allgemeinen $\mathbf{AB} \neq \mathbf{BA}$.
- Aus $\mathbf{AB} = \mathbf{O}$ (Nullmatrix) folgt **keineswegs** $\mathbf{A} = \mathbf{O}$ oder $\mathbf{B} = \mathbf{O}$. Man sagt, daß $\mathbb{K}^{n \times n}$ nicht nullteilerfrei ist.
- Selbst aus $\mathbf{A}^2 = \mathbf{A} \cdot \mathbf{A} = \mathbf{O}$ folgt **nicht** $\mathbf{A} = \mathbf{O}$.

Übungsaufgabe 2.3.

Berechnen Sie $\mathbf{A}^2$ für

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

Finden Sie eine Matrix $\mathbf{B} \in \mathbb{R}^{2 \times 3}$, so daß $\mathbf{AB} = \mathbf{O}$.

2.4 Inverse von Matrizen

Definition 2.10 (Invertierbare Matrix).

Eine Matrix $A \in \mathbb{K}^{n \times n}$ heißt genau dann invertierbar, wenn eine Matrix $B \in \mathbb{K}^{n \times n}$ existiert mit

$$AB = BA = I_n.$$

Wenn diese Matrix existiert, dann ist Sie eindeutig bestimmt und wird mit A^{-1} bezeichnet.

Bemerkung 2.5.

Behauptungen wie

Wenn diese Matrix existiert, dann ist Sie eindeutig bestimmt.

haben in Definitionen eigentlich nicht zu suchen. Stimmt die Aussage? Angenommen $\mathbf{A} \in \mathbb{K}^{n \times n}$ ist so, daß zwei Inverse B und B' existieren. Dann gilt

$$\mathbf{B} = \mathbf{I} \cdot \mathbf{B} = (\mathbf{B}' \mathbf{A}) \mathbf{B} = \mathbf{B}' (\mathbf{A} \mathbf{B}) = \mathbf{B}' \cdot \mathbf{I} = \mathbf{B}'.$$

Damit ist die Bezeichnung **die Inverse** gerechtfertigt.

Bemerkung 2.6.

Nicht jede Matrix $A \in \mathbb{K}^{n \times n}$ ist invertierbar. Beispielsweise existiert für

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \end{bmatrix}$$

keine Matrix $\mathbf{B} \in \mathbb{K}^{n \times n}$ mit $\mathbf{A} \mathbf{B} = \mathbf{I}_n$.

Satz 2.5 (Rechenregeln für Inverse).

Seien $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{n \times n}$ invertierbar. Dann gelten:

- $\mathbf{A}^T$ und $\mathbf{A}^H$ sind invertierbar, und es gilt

$$(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T \quad \text{bzw.} \quad (\mathbf{A}^H)^{-1} = (\mathbf{A}^{-1})^H.$$

- $\mathbf{A}^{-1}$ ist invertierbar, und es gilt

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A}.$$

- $\mathbf{AB}$ ist invertierbar, und es gilt

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}.$$

Schreibregelung. Da es nach dem ersten Anstrich egal ist, welche der beiden Operationen zuerst ausgeführt wird, schreibt man auch, etwas weniger umständlich, $\mathbf{A}^{-T}$ bzw. $\mathbf{A}^{-H}$.

Übungsaufgabe 2.4.

Beweisen Sie Satz 2.5.

2.5 Lineare Abbildungen

Jede Matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ induziert nun über die Matrix-Vektor-Multiplikation eine Abbildung (die wir wieder mit $\mathbf{A}$ bezeichnen):

$$\begin{cases} \mathbf{A}: \mathbb{K}^n \rightarrow \mathbb{K}^m \\ \mathbf{x} \mapsto \mathbf{y} = \mathbf{Ax}. \end{cases}$$

Der Definitionsbereich von $\mathbf{A}$ ist also $\mathbb{K}^n$ und der Wertebereich von $\mathbf{A}$ ist in $\mathbb{K}^m$ enthalten.

Die Abbildung $\mathbf{A}$ besitzt die bemerkenswerte Eigenschaft

$$\mathbf{A}(\alpha \mathbf{x} + \beta \mathbf{y}) = \alpha \mathbf{Ax} + \beta \mathbf{Ay}$$

für alle $\mathbf{x}, \mathbf{y} \in \mathbb{K}^n$. Diese Eigenschaft wird mit **Linearität** bezeichnet.

Definition 2.11 (Lineare Abbildungen).

Eine Abbildung $f: \mathbb{K}^n \rightarrow \mathbb{K}^m$ heißt genau dann linear wenn für alle $\mathbf{x}, \mathbf{y} \in \mathbb{K}^n$ und $\alpha, \beta \in \mathbb{K}$ gilt

$$f(\alpha \mathbf{x} + \beta \mathbf{y}) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y}).$$

Übungsaufgabe 2.5.

überprüfen Sie, ob die folgenden Abbildungen $f: \mathbb{K}^2 \rightarrow \mathbb{K}^2$ lineare Abbildungen sind:

$$(i) \quad f(\mathbf{x}) = \begin{bmatrix} x_1 + 1 \\ x_1 - x_2 \end{bmatrix}$$

$$(ii) \quad f(\mathbf{x}) = \begin{bmatrix} \sqrt{|x_1|} \\ x_2 \end{bmatrix}$$

$$(iii) \quad f(\mathbf{x}) = \begin{bmatrix} x_1 - 2x_2 \\ x_1 + 3x_2 \end{bmatrix}.$$

Wir setzen

$$L(\mathbb{K}^n, \mathbb{K}^m) := \{f: \mathbb{K}^n \rightarrow \mathbb{K}^m : f \text{ ist linear}\}.$$

Dann erweist sich auch $L(\mathbb{K}^n, \mathbb{K}^m)$ in natürlicher Weise als Vektorraum. Wir definieren die Vektoraddition

$$\begin{cases} +: L(\mathbb{K}^n, \mathbb{K}^m) \times L(\mathbb{K}^n, \mathbb{K}^m) & \rightarrow & L(\mathbb{K}^n, \mathbb{K}^m), \\ (f, g) & \mapsto & f + g \end{cases}$$

punktweise durch

$$x \in \mathbb{K}^n: \quad (f + g)(x) := f(x) + g(x),$$

wobei wir rechts die Vektoraddition aus [Sektion 1.1](#) verwenden. Ebenso definieren wir die Skalarmultiplikation

$$\begin{cases} \cdot: \mathbb{K} \times L(\mathbb{K}^n, \mathbb{K}^m) & \rightarrow & L(\mathbb{K}^n, \mathbb{K}^m), \\ (\lambda, f) & \mapsto & \lambda f. \end{cases}$$

punktweise durch

$$x \in \mathbb{K}^n: \quad (\lambda f)(x) := \lambda \cdot f(x),$$

wobei wir rechts die Skalarmultiplikation aus [Sektion 1.1](#) benutzen. Wir setzen 0 einfach als die Abbildung, die identisch auf die $\mathbf{0}$ in $\mathbb{K}^m$ abbildet und $-f$ wird wieder punktweise durch

$$x \in \mathbb{K}^n: \quad (-f)(x) = -f(x)$$

definiert mit der Definition in [Sektion 1.1](#).

Definition 2.12 (Vektorraum $L(\mathbb{K}^n, \mathbb{K}^m)$).

Mit den obigen Operationen für Vektoraddition und Skalarmultiplikation nennen wir $L(\mathbb{K}^n, \mathbb{K}^m)$ den **Vektorraum der linearen Abbildungen** zwischen $\mathbb{K}^n$ und $\mathbb{K}^m$.

Satz 2.6 (Rechengesetze in $L(\mathbb{K}^n, \mathbb{K}^m)$).

In $L(\mathbb{K}^n, \mathbb{K}^m)$ gilt:

- (i) Für alle $f \in L(\mathbb{K}^n, \mathbb{K}^m)$ gilt $f + 0 = f$.
- (ii) Für alle $f \in L(\mathbb{K}^n, \mathbb{K}^m)$ gilt $f + (-f) = 0$.
- (iii) Für alle $f, g \in L(\mathbb{K}^n, \mathbb{K}^m)$ gilt $f + g = g + f$.
- (iv) Für alle $f, g, h \in L(\mathbb{K}^n, \mathbb{K}^m)$ gilt $(f + g) + h = f + (g + h)$.
- (v) Für alle $f \in L(\mathbb{K}^n, \mathbb{K}^m)$ gilt $1f = f$.
- (vi) Für alle $\lambda, \mu \in \mathbb{K}$ und alle $f \in L(\mathbb{K}^n, \mathbb{K}^m)$ gilt $\lambda(\mu f) = (\lambda\mu)f$.
- (vii) Für alle $\lambda \in \mathbb{K}$ und alle $f, g \in L(\mathbb{K}^n, \mathbb{K}^m)$ gilt $\lambda(f + g) = \lambda f + \lambda g$.
- (viii) Für alle $\lambda, \mu \in \mathbb{K}$ und alle $f \in \mathbb{K}^n$ gilt $(\lambda + \mu)f = \lambda f + \mu f$.

Übungsaufgabe 2.6.

- (i) Vergleichen Sie die Aussage von Satz 2.6 mit denen der Sätze 1.1 und 2.2.
- (ii) Beweisen Sie alle Aussagen von Satz 2.6. Natürlich nur mit der Definition der Rechenoperationen und dann geeigneten Aussagen die für diese gelten.

Definition 2.13 (Kern (Nullraum)/Bild).

Es sei $f: \mathbb{K}^n \rightarrow \mathbb{K}^m$ eine lineare Abbildung. Dann heißt

(i) Die Menge

$$\ker(f) := \{\mathbf{x} \in \mathbb{K}^n : f(\mathbf{x}) = \mathbf{0}\}$$

der **Kern** oder **Nullraum** von f .

(ii) die Menge

$$\operatorname{im}(f) := \{\mathbf{y} = f(\mathbf{x}) : \mathbf{x} \in \mathbb{K}^n\}$$

das **Bild** von f .

Wenn f durch eine Matrix $\mathbf{A}$ gegeben ist, dann schreiben wir auch $N(\mathbf{A})$ (N für nullspace) für den Kern und $R(\mathbf{A})$ (R für range) für das Bild.

Auf Grund der Linearität der Abbildung f in der Definition von $\ker(f)$ und $\operatorname{im}(f)$, haben diese Mengen eine spezielle Struktur. Sie sind Vektorräume.

Es gilt

Satz 2.7 (Dimensionssatz/Rangsatz).

Sei $f: \mathbb{K}^n \rightarrow \mathbb{K}^m$ eine lineare Abbildung. Dann gilt

(i) $\ker(f)$ ist ein Unterraum von $\mathbb{K}^n$,

(ii) $\operatorname{im}(f)$ ist ein Unterraum von $\mathbb{K}^m$,

(iii) $\dim(\ker(f)) + \dim(\operatorname{im}(f)) = n$.

Wir nennen $\dim(\ker(f))$ den **Defekt** von f und $\dim(\operatorname{im}(f)) =: \operatorname{rank}(f)$ den **Rang** von f .

Bemerkung 2.7.

Nach unserer Diskussion der Matrixmultiplikation ist klar, daß für $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$, $\mathbf{A} \in \mathbb{R}^{m \times n}$ gilt

$$\text{im}(f) = R(A) = \text{span}(\mathbf{A}[\cdot, 1], \dots, \mathbf{A}[\cdot, n]).$$

Damit ist die Anzahl der linear unabhängigen Spalten in der Matrix A gleich dem Rang der Matrix A . Manchmal wird dies auch der **Spaltenrang** genannt. In den Übungen werden Sie sehen, daß $\dim(R(A)) = \dim(R(A^T))$ gilt und damit der Zeilenrang gleich dem Spaltenrang ist. Damit ist der Name Rang gerechtfertigt.

Beispiel 2.1.

Wir geben ein paar Beispiele an. Die Leserin versuche die an dieser Stelle nachzuvollziehen. Nach dem Studium des Abschnittes zu Abbildungsmatrizen sollten die Beispiele nochmals angesehen werden.

$$(i) \ f: \mathbb{R}^2 \rightarrow \mathbb{R}^3, f(x) = \begin{bmatrix} x_1 - x_2 \\ x_2 \\ x_1 \end{bmatrix}. \text{ Dann gilt}$$

$$\ker(f) = \{0\}, \quad \text{im}(f) = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$$

Der Defekt ist 0 (d.h. die Abbildung ist injektiv (Wirklich?)) und der Rang ist 2. Man sagt in diesem Fall auch, daß die Abbildung vollen Rang hat.

(ii) $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$, $f(\mathbf{x}) = \begin{bmatrix} x_1 + x_2 + x_3 \\ 0 \end{bmatrix}$. Dann gilt

$$\begin{aligned} \ker(f) &= \text{span} \left\{ \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} \right\} \\ &= \text{span} \left\{ \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\} \end{aligned}$$

da

$$-\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} - \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$$

und

$$\text{im}(f) = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}.$$

Der Defekt von f ist also 2 und der Rang gleich 1.

Übungsaufgabe 2.7.

Zeigen Sie (i) und (ii) von Satz 2.7.

Bemerkung 2.8.

Sämtliche Sätze von Kapitel 1 gelten auch, wenn man $\mathbb{R}$ durch $\mathbb{K}$ ersetzt. (Siehe Sektion 1.3.) Die Beweise sind identisch und die Leserin sollte sich die klar machen.

Wenn Sie das $\mathbb{K}$ für den Anfang verwirrt, ersetzen Sie es durch $\mathbb{R}$ und denken das Argument damit durch.

Beweis. Wir zeigen noch (iii). Wir setzen

$$\dim(\ker(f)) = k \quad \text{und} \quad \dim(\operatorname{im}(f)) = r$$

und zeigen $k + r = n$. Sei nun $\{\mathbf{x}_1, \dots, \mathbf{x}_r\}$ eine Basis von $\operatorname{im}(f)$. Wir wählen nun Vektoren $\{\mathbf{x}'_1, \dots, \mathbf{x}'_r\} \subseteq \mathbb{K}^n$ mit $f(\mathbf{x}'_i) = \mathbf{x}_i$ sowie $\{\mathbf{x}''_1, \dots, \mathbf{x}''_k\}$ eine Basis von $\ker(f)$. Es bleibt uns nun, nach Satz 8.3 zu zeigen, daß dann

$$\{\mathbf{x}'_1, \dots, \mathbf{x}'_r, \mathbf{x}''_1, \dots, \mathbf{x}''_k\}$$

eine Basis von $\mathbb{K}^n$ ist. Wir zeigen zuerst, daß es ein Erzeugendensystem ist. Sei dazu $\mathbf{x} \in \mathbb{K}^n$ beliebig. Dann gilt $f(\mathbf{x}) \in \operatorname{im}(f)$ und damit

$$f(\mathbf{x}) = \sum_{i=1}^r \lambda'_i \mathbf{x}_i = f\left(\sum_{i=1}^r \lambda'_i \mathbf{x}'_i\right)$$

für eine gewisse Menge $\{\lambda'_1, \dots, \lambda'_r\}$. Für

$$\mathbf{x}' = \mathbf{x} - \sum_{i=1}^r \lambda'_i \mathbf{x}'_i$$

folgt dann $f(\mathbf{x}') = \mathbf{0}$, d.h. $\mathbf{x}' \in \ker(f)$. Damit gilt also

$$\mathbf{x}' = \sum_{i=1}^k \lambda''_i \mathbf{x}''_i$$

für eine gewisse Menge $\{\lambda''_1, \dots, \lambda''_k\}$. Damit gilt dann wiederum

$$\mathbf{x} = \sum_{i=1}^r \lambda'_i \mathbf{x}'_i + \sum_{i=1}^k \lambda''_i \mathbf{x}''_i$$

bzw.

$$x \in \text{span}(\{\mathbf{x}_1, \dots, \mathbf{x}_r, \mathbf{x}_1'', \dots, \mathbf{x}_k''\}).$$

Damit ist $\{\mathbf{x}_1, \dots, \mathbf{x}_r, \mathbf{x}_1'', \dots, \mathbf{x}_k''\}$ ein Erzeugendensystem von $\mathbb{R}^n$. Wir müssen noch die lineare Unabhängigkeit zeigen. Angenommen es gilt

$$\sum_{i=1}^r \lambda'_i \mathbf{x}'_i + \sum_{i=1}^k \lambda''_i \mathbf{x}''_i = \mathbf{0} \quad (2.5.1)$$

für eine geeignete Menge $\{\lambda'_1, \dots, \lambda'_r, \lambda''_1, \dots, \lambda''_k\}$. Dann folgt

$$\begin{aligned} \mathbf{0} &= f\left(\sum_{i=1}^r \lambda'_i \mathbf{x}'_i + \sum_{i=1}^k \lambda''_i \mathbf{x}''_i\right) \\ &= f\left(\sum_{i=1}^r \lambda'_i \mathbf{x}'_i\right) + f\left(\sum_{i=1}^k \lambda''_i \mathbf{x}''_i\right) \\ &= \sum_{i=1}^r \lambda'_i \mathbf{x}_i, \end{aligned}$$

da die $\mathbf{x}''_i$ alle in $\ker(f)$ sind. Da $\{\mathbf{x}_1, \dots, \mathbf{x}_r\}$ eine Basis von $\text{im}(f)$ ist, muß

$$\lambda'_1 = \lambda'_2 = \dots = \lambda'_r = 0$$

gelten. Damit haben wir nun, wegen (2.5.1) daß

$$\sum_{i=1}^k \lambda''_i \mathbf{x}''_i = \mathbf{0}$$

gelten muß. Da die $\{\mathbf{x}''_1, \dots, \mathbf{x}''_k\}$ eine Basis von $\ker(f)$ bilden, folgt also auch hier

$$\lambda''_1 = \lambda''_2 = \dots = \lambda''_k = 0.$$

Damit ist die lineare Unabhängigkeit gezeigt. □

2.5.1 Matrixdarstellungen lin. Abbildungen

Wir haben schon gesehen, daß Matrizen lineare Abbildungen induzieren. Wir werden nun sehen, daß die Umkehrung auch gilt und lineare Abbildungen mit natürlichen Matrizen daher kommen.

Wir erinnern uns zuerst, daß $\mathbf{x}_B$ der Koordinatenvektor $[\alpha_1, \dots, \alpha_n]^T$ ist wobei die α_i durch

$$\mathbf{x} = \alpha_1 \mathbf{b}_1 + \dots + \alpha_n \mathbf{b}_n$$

gegeben sind. In Matrix-Schreibweise gilt

$$\mathbf{x} = \begin{bmatrix} | & & | \\ \mathbf{b}_1 & \cdots & \mathbf{b}_n \\ | & & | \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

Damit ist die Koordinatentransformation eine lineare Abbildung und es gilt $(\mathbf{x} + \mathbf{y})_B = \mathbf{x}_B + \mathbf{y}_B$. Wenn wir einen Vektor in Koordinaten bzgl. B gegeben haben, erhalten wir die Standard-Koordinaten durch $\mathbf{x} = \mathbf{B}^{-1} \mathbf{x}_B$. (wieso ist $\mathbf{B}$ invertierbar?)

Sei $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ eine Basis von $\mathbb{K}^n$, $\mathbf{x} \in \mathbb{K}^n$ und $\tilde{B} = \{\tilde{\mathbf{b}}_1, \dots, \tilde{\mathbf{b}}_m\}$ eine Basis von $\mathbb{K}^m$. Dann gilt

$$\begin{aligned} f(\mathbf{x})_{\tilde{B}} &= f\left(\sum_{i=1}^n \alpha_i \mathbf{b}_i\right)_{\tilde{B}} = \sum_{i=1}^n \alpha_i f(\mathbf{b}_i)_{\tilde{B}} \\ &= \begin{bmatrix} | & & | \\ f(\mathbf{b}_1)_{\tilde{B}} & \cdots & f(\mathbf{b}_n)_{\tilde{B}} \\ | & & | \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} \\ &= \begin{bmatrix} | & & | \\ f(\mathbf{b}_1)_{\tilde{B}} & \cdots & f(\mathbf{b}_n)_{\tilde{B}} \\ | & & | \end{bmatrix} \mathbf{x}_B \end{aligned}$$

Es reicht also, die Abbildung auf einer bel. Basis B Basis zu kennen um das

Bild $f(\mathbf{x})_{\tilde{B}}$ für ein beliebiges $\mathbf{x}_B \in \mathbb{K}^n$ auszurechnen. Die Matrix

$$A_{f,B,\tilde{B}} = \begin{bmatrix} | & & | \\ f(\mathbf{b}_1)_{\tilde{B}} & \cdots & f(\mathbf{b}_n)_{\tilde{B}} \\ | & & | \end{bmatrix}$$

wird als Abbildungsmatrix von f bzgl. der Basen B und $\tilde{B}$ bezeichnet und es gilt

$$A_{f,B,\tilde{B}} \mathbf{x}_B = f(\mathbf{x})_{\tilde{B}}.$$

Im folgenden beschränken wir uns darauf, in beiden Räumen die Standard-Basis zu wählen. Wir werden sehen, daß dies keine Einschränkung darstellt.

Wenn wir die Standard-Basen nehmen, dann schreiben wir einfach A_f anstelle von $A_{f,B,\tilde{B}}$. Es gilt

$$\mathbf{A}_f = [f(\mathbf{e}_1) | \cdots | f(\mathbf{e}_n)]$$

und $f(\mathbf{x}) = \mathbf{A}_f \mathbf{x}$ für alle $\mathbf{x} \in \mathbb{R}^n$.

Wie hängen A_f und $A_{f,B,\tilde{B}}$ nun zusammen?

Übungsaufgabe 2.8.

Zeigen Sie, daß die Zuordnung

$$L(\mathbb{R}^n, \mathbb{R}^m) \ni f \mapsto \mathbf{A}_f \in \mathbb{R}^{m \times n} \quad (2.5.2)$$

bzgl. der kanonischen Basen linear (d.h. $f + g \mapsto \mathbf{A}_f + \mathbf{A}_g$, $\lambda f \mapsto \lambda \mathbf{A}_f$) und bijektiv ist. Wir fassen zusammen.

Wir fassen zusammen:

Satz 2.8 (Abbildungsmatrix).

Sei $f: \mathbb{K}^n \rightarrow \mathbb{K}^m$ eine lineare Abbildung.

Dann existiert eine eindeutig bestimmte Matrix $\mathbf{A}_f \in \mathbb{K}^{m \times n}$, so daß

$$f(\mathbf{x}) = \mathbf{A}_f \mathbf{x} \quad \text{für alle } \mathbf{x} \in \mathbb{K}^n.$$

In den Spalten $\mathbf{A}[\cdot, j]$, $1 \leq j \leq n$ von $\mathbf{A}$ stehen dabei gerade die Bilder der Einheitsvektoren $\mathbf{e}_j$, d. h.

$$\mathbf{A}_f[\cdot, j] = f(\mathbf{e}_j) \quad j = 1, \dots, n.$$

Beispiel 2.2.

Wir betrachten die lineare Abbildung (Nachweis?)

$$\left\{ \begin{array}{l} f : \mathbb{R}^2 \rightarrow \mathbb{R}^3 \\ \mathbf{x} \mapsto f(\mathbf{x}) = \begin{bmatrix} x_1 + x_2 \\ x_1 \\ x_2 \end{bmatrix} \end{array} \right. .$$

und wollen die Abbildungsmatrix bestimmen. Da wir auf beiden Seiten der Abbildung, also in $\mathbb{R}^2$ und $\mathbb{R}^3$ die standard-Basis verwenden, brauchen wir nur die Bilder $f(\mathbf{e}_1)$ und $f(\mathbf{e}_2)$ bestimmen.

Da die Abbildung von $\mathbb{R}^2$ nach $\mathbb{R}^3$ ist, wissen wir, daß unsere Matrix 2 Spalten und 3 Zeilen hat:

$$f(\mathbf{e}_1) = \begin{bmatrix} 1+0 \\ 1 \\ 0 \end{bmatrix} \quad \text{und} \quad f(\mathbf{e}_2) = \begin{bmatrix} 0+1 \\ 0 \\ 1 \end{bmatrix}$$

und damit

$$A_f = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Mit einem scharfen Auge hätte man das auch direkt ablesen können aber Sie müssen natürlich wissen, wie man die Abbildungsmatrix systematisch ausrechnet.

Übungsaufgabe 2.9 (Verkettung linearer Abbildungen).

Zeigen Sie, daß für $g \in L(\mathbb{K}^n, \mathbb{K}^k)$, $f \in L(\mathbb{K}^k, \mathbb{K}^m)$ gilt $f \circ g \in L(\mathbb{K}^n, \mathbb{K}^m)$.

Satz 2.9 (Verkettungen lin. Abbildungen).

Seien $g \in L(\mathbb{K}^n, \mathbb{K}^k)$, $f \in L(\mathbb{K}^k, \mathbb{K}^m)$. Dann gilt für die Matrixdarstellung von $f \circ g$ bzgl. der Standardbasen

$$A_{f \circ g} = A_f A_g.$$

2.6 Verständnisfragen

1. Für welche Matrizen gilt $\mathbf{A}^T = \mathbf{A}$?
2. Wenn $\mathbf{A} \in \mathbb{R}^{m \times n}$, dann $\mathbf{A}^T \in ??$
3. Unter welchen Voraussetzungen existiert für eine Matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$
 - a) $\mathbf{A} + \mathbf{A}^T$
 - b) $\mathbf{A} \cdot \mathbf{A}^T$ und $\mathbf{A}^T \cdot \mathbf{A}$ und in welchem Raum liegen die Ergebnisse?
4. Wie lautet der i -te Zeilen und j -te Spaltenvektor der Einheitsmatrix $\mathbf{I}_n$.
5. Welche Dimension hat die Matrix $\mathbf{A}$, wenn $\mathbf{A}^T \in \mathbb{R}^n$ gilt?
6. Welche Matrizen sind zugleich obere und untere Dreiecksmatrizen?
7. Welche Matrizen können mit sich selbst multipliziert werden?
8. Unter welchen Voraussetzungen an $\mathbf{A}$ und $\mathbf{B}$ existiert sowohl das Produkt $\mathbf{A} \cdot \mathbf{B}$ als auch $\mathbf{B} \cdot \mathbf{A}$. Sind dann auch beide Produkte vom gleichen Typ (haben gleiche Dimensionen) oder gilt sogar Kommutativität?
9. Welche Gestalt hat die Matrix $\mathbf{A}^k = \mathbf{A} \cdot \dots \cdot \mathbf{A}$ (k Faktoren), wenn $\mathbf{A}$ eine Diagonalmatrix ist?
10. Ist das Produkt zweier Diagonalmatrizen wieder eine Diagonalmatrix? Sind diese Produkte kommutativ?
11. Welche spezielle Matrizen $\mathbf{A} = [a_{ij}]_{\substack{1 \leq i,j \\ 1 \leq n}}$ ergeben sich, wenn
 - a) $a_{i,j} = \begin{cases} 1 & : i \neq j \\ 0 & : i = j \end{cases}$
 - b) $a_{ij} = i + j$.
12. Unter welchen Voraussetzungen sind 2×2 obere Dreiecksmatrizen zu sich selbst invers?
13. Wieviele 3×3 -Diagonalmatrizen sind zu sich selbst invers?

2.7 Rechenaufgaben

Aufgabe 1

Gegeben ist $\mathbf{A} = \begin{bmatrix} \sin(\varphi) & \cos(\varphi) \\ \cos(\varphi) & -\sin(\varphi) \end{bmatrix}$. Berechnen Sie $\mathbf{A}^2$ und $\mathbf{A}^{-1}$ (siehe Sektion 3.2).

Welche Eigenschaft hat $\mathbf{A}$?

Aufgabe 2

Für die Matrix $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix}$ ($a \in \mathbb{R}$ Parameter) berechne man $\mathbf{A}^3$ und per vollständiger Induktion $\mathbf{A}^n$.

Aufgabe 3

Berechnen Sie für die Matrizen

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 \\ 3 & 2 & 1 \\ 0 & -1 & 1 \end{bmatrix} \quad \text{und} \quad \mathbf{B} = \begin{bmatrix} 1 & 1 & -2 \\ -2 & 0 & 0 \\ 0 & 1 & 3 \end{bmatrix}$$

die Produkte $\mathbf{AB}$, $\mathbf{BA}$, $\mathbf{AB}^T$, $\mathbf{BA}^T$, $\mathbf{A}^T\mathbf{B}^T$.

Aufgabe 4

Es sei M die Menge aller Matrizen der Form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, wobei $a, b \in \mathbb{R}$ ist.

Zeigen Sie, daß die Summe und das Produkt zwei beliebiger Matrizen aus M wieder zu M gehören.

Aufgabe 5

Lösen Sie die Matrixgleichung $\mathbf{A}\mathbf{X}^T\mathbf{B} = \mathbf{C}$ nach $\mathbf{X}$ auf wobei wir annehmen, daß die Voraussetzungen für die Auflösbarkeit, über die Sie nachdenken sollten, erfüllt sein mögen.

Berechnen Sie $\mathbf{X}$ für

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 0 & 1 & 1 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad \text{und} \quad \mathbf{C} = \begin{bmatrix} 1 & 2 & -3 \\ 4 & 0 & 1 \end{bmatrix}.$$

Aufgabe 6

Berechnen Sie $\mathbf{X}$ aus der Matrixgleichung $\mathbf{A}\mathbf{X} + \mathbf{C} = \mathbf{B}$ für

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 3 & 1 & 1 & 1 \\ 1 & -1 & -1 & 3 \\ -1 & 6 & -5 & 0 \end{bmatrix}$$

$$\text{und} \quad \mathbf{C} = \begin{bmatrix} 1 & 2 & 3 & -3 \\ 0 & -1 & 2 & 2 \\ -1 & 2 & -4 & -2 \end{bmatrix}.$$

Aufgabe 7

Für die Matrizen $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{X}$ gelte die Gleichung $\mathbf{A}^T\mathbf{X}\mathbf{B} = \mathbf{C}$, wobei $\mathbf{A}$ und $\mathbf{B}$ invertierbar sind.

- Stellen Sie die Gleichung nach $\mathbf{X}$ um.
- Von welchem Typ müssen $\mathbf{A}$ und $\mathbf{B}$ gewählt werden, damit die Gleichung für $\mathbf{C} \in \mathbb{R}^{5 \times 3}$ sinnvoll wird?
- Berechnen Sie $\mathbf{X}$ für

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ -1 & 2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \\ 6 & 3 & 4 \end{bmatrix} \quad \text{und} \quad \mathbf{C} = \begin{bmatrix} 0 & 1 & -1 \\ 2 & 0 & 2 \end{bmatrix}.$$

Lineare Gleichungssysteme

Ein System aus Gleichungen der Form

$$\begin{array}{ccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \cdots & + & a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \cdots & + & a_{2,n}x_n & = & b_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \cdots & + & a_{m,n}x_n & = & b_m \end{array} \quad (3.0.1)$$

heißt **lineares Gleichungssystem** (LGS), genauer: ein System von m linearen algebraischen Gleichungen in n Unbekannten.

Dabei sind die **Koeffizienten** $a_{i,j} \in \mathbb{K}$ sowie die Zahlen $b_i \in \mathbb{K}$ vorgegeben, während man die Zahlen $x_i \in \mathbb{K}$ zu bestimmen sucht.

Mit der **Koeffizientenmatrix**

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix} \in \mathbb{K}^{m \times n},$$

dem Vektor der **Unbekannten**

$$\mathbf{x} = [x_1, x_2, \dots, x_n]^T \in \mathbb{K}^n$$

sowie der **rechten Seite**

$$\mathbf{b} = [b_1, b_2, \dots, b_m]^T \in \mathbb{K}^m$$

schreibt man (3.0.1) kürzer als:

$$\mathbf{Ax} = \mathbf{b}. \quad (3.0.2)$$

Ein lineares Gleichungssystem $\mathbf{Ax} = \mathbf{b}$ heißt **homogen**, wenn $\mathbf{b} = \mathbf{0}$ und **inhomogen**, wenn $\mathbf{b} \neq \mathbf{0}$.

Ein **homogenes** System besitzt immer (mindestens) eine Lösung, nämlich $\mathbf{x} = \mathbf{0}$. Die Lösungsmenge von $\mathbf{Ax} = \mathbf{0}$ ist gerade der Kern $N(\mathbf{A})$ von $\mathbf{A}$. Bei inhomogenen Systemen ist die Sache komplizierter.

Wir beginnen mit folgendem Ergebnis:

Satz 3.1.

Ist $A \in \mathbb{K}^{m \times n}$ und ist $\mathbf{b} \in R(\mathbf{A})$, d. h. es gibt (mindestens) ein $\mathbf{x}_0 \in \mathbb{K}^n$ mit $A\mathbf{x}_0 = \mathbf{b}$, dann gilt

$$\begin{aligned} \{\mathbf{x} \in \mathbb{K}^n : \mathbf{Ax} = \mathbf{b}\} &= \{\mathbf{x}_0\} + N(\mathbf{A}) \\ &= \{\mathbf{x}_0 + \mathbf{y} : \mathbf{y} \in N(\mathbf{A})\}. \end{aligned} \quad (3.0.3)$$

Bemerkung 3.1.

In anderen Worten: die allgemeine Lösung eines Gleichungssystems $\mathbf{Ax} = \mathbf{b}$ ergibt sich aus einer speziellen/partikulären Lösung $\mathbf{x}_0 \in \mathbb{K}^n$ mit $A\mathbf{x}_0 = \mathbf{b}$ und beliebigen Lösungen des homogenen Systems $\mathbf{Ax} = \mathbf{0}$:

$$\{\mathbf{x} \in \mathbb{K}^n : \mathbf{Ax} = \mathbf{b}\} = \mathbf{x}_0 + \{\mathbf{x} \in \mathbb{K}^n : \mathbf{Ax} = \mathbf{0}\}. \quad (3.0.4)$$

Übungsaufgabe 3.1.

Beweisen Sie Satz 3.1.

Im Falle der Existenz einer Lösung erhält man also die Lösungsmenge des inhomogenen Systems durch Verschieben der Lösungsmenge des homogenen Systems zur gleichen Koeffizientenmatrix.

- die Leserin mache sich dies anhand des relativ trivialen Gleichungssystems

$$x_1 + x_2 = 7$$

graphisch klar. Verwenden Sie im Ansatz (3.0.3) aus Satz 3.1 auch verschiedene $\mathbf{x}_0$.

- Die Leserin verifiziere die Beziehung (3.0.3).

Zur weiteren Analyse der Lösbarkeit von LGS beleuchten wir die Lösungssuche noch unter einem weiteren Aspekt. Sind $\mathbf{A}[\cdot, 1], \mathbf{a}_2, \dots, \mathbf{A}[\cdot, n]$ die Spalten der Matrix $A \in \mathbb{K}^{m \times n}$, dann läßt sich $A\mathbf{x} = \mathbf{b}$ auch schreiben als

$$x_1 \mathbf{A}[\cdot, 1] + x_2 \mathbf{A}[\cdot, 2] + \dots + x_n \mathbf{A}[\cdot, n] = \mathbf{b}$$

Die Suche nach einer Lösung von $A\mathbf{x} = \mathbf{b}$ ist also gleichbedeutend mit der Suche nach Linearkombinationen der Spalten $\mathbf{A}[\cdot, j]$, die die rechte Seite $\mathbf{b}$ ergeben.

Wir halten somit fest:

Satz 3.2 (Lösbarkeit linearer Gleichungssysteme).

Es seien $A \in \mathbb{K}^{m \times n}$ mit den Spalten $\mathbf{A}[\cdot, 1], \mathbf{A}[\cdot, 2], \dots, \mathbf{A}[\cdot, n]$ und $\mathbf{b} \in \mathbb{K}^m$.

Das lineare Gleichungssystem $A\mathbf{x} = \mathbf{b}$ ist genau dann lösbar, wenn

$$\mathbf{b} \in R(\mathbf{A}) = \text{span}\{\mathbf{A}[\cdot, 1], \mathbf{A}[\cdot, 2], \dots, \mathbf{A}[\cdot, n]\}.$$

Das lineare Gleichungssystem $A\mathbf{x} = \mathbf{b}$ ist genau dann **eindeutig** lösbar, wenn $\mathbf{b} \in R(\mathbf{A})$ gilt, und die Spalten $\mathbf{A}[\cdot, 1], \mathbf{A}[\cdot, 2], \dots, \mathbf{A}[\cdot, n]$ linear unabhängig sind:

$$A\mathbf{x} = \mathbf{b} \text{ ist eindeutig lösbar} \Leftrightarrow \mathbf{b} \in R(\mathbf{A}) \text{ und } \text{rank}(\mathbf{A}) = n.$$

Zusammenfassung: Es können somit genau drei Fälle eintreten:

- $A\mathbf{x} = \mathbf{b}$ besitzt **keine** Lösung, d.h. $\mathbf{b} \notin R(\mathbf{A})$ (kann nur bei inhomogenen LGS passieren).
- $A\mathbf{x} = \mathbf{b}$ besitzt **genau eine** Lösung, d.h. $\mathbf{b} \in R(\mathbf{A})$ und $\text{rank}(\mathbf{A}) = n$.
Dieser Fall kann nur für $m \geq n$ eintreten (mindestens so viele Gleichungen wie Unbekannte).
- $A\mathbf{x} = \mathbf{b}$ besitzt **unendlich viele** Lösungen, d.h. $\mathbf{b} \in R(\mathbf{A})$ und $\text{rank}(\mathbf{A}) < n$. Mit einer beliebigen Lösung $\mathbf{x}_0$ erhält man als Lösungsmenge

$$\mathbf{x}_0 + N(\mathbf{A}) = \mathbf{x}_0 + \{\text{alle Lösungen von } A\mathbf{x} = \mathbf{0}\}.$$

Lineare Gleichungssysteme mit mehr als einer, aber nur endlich vielen Lösungen existieren nicht.

3.1 Berechnung von Lösungen

Bislang wissen wir zwar über die Lösbarkeit eines LGS $\mathbf{Ax} = \mathbf{b}$ Bescheid, können aber die Lösung selbst noch nicht konkret *berechnen*. Jedenfalls nicht sehr effektiv. Für diese Aufgabe stehen verschiedenste Algorithmen zur Verfügung, so zum Beispiel die in dieser Vorlesung behandelten:

- Gauß-Algorithmus,
- Gauß-Jordan-Algorithmus – eine Erweiterung des Gauß-Algorithmus.

Des weiteren gibt es eine ganze Reihe numerischer Algorithmen, die Sie mittels Literatur oder in einer Numerik-Vorlesung erlernen können. Die hier vorgestellten exakten Verfahren beruhen auf Umformungen in ein äquivalentes Gleichungssystem $\tilde{\mathbf{A}}\mathbf{x} = \tilde{\mathbf{b}}$, für das man die Lösung relativ mühelos angeben kann.

Zunächst erzeugen wir aus der Koeffizientenmatrix $A \in \mathbb{K}^{m \times n}$ und der rechten Seite $b \in \mathbb{K}^m$ die **erweiterte Koeffizientenmatrix** $[A | \mathbf{b}]$. Diese enthält die vollständige Information über das betrachtete LGS. Auf dieser werden wir dann die sogenannten **elementaren Umformungen**

- Multiplikation einer Gleichung (Zeile) mit $\lambda \in \mathbb{K}$, $\lambda \neq 0$,
- Addition des Vielfachen einer Gleichung (Zeile) zu einer anderen Gleichung (Zeile),
- Vertauschung von zwei Gleichungen (Zeilen),
- Umnumerierung von zwei Unbekannten (Vertauschung von zwei Spalten),

durchführen. Diese Operationen ändern die Lösungsmenge von $\mathbf{Ax} = \mathbf{b}$ nicht wie im nächsten Satz festhalten.

Satz 3.3.

Die folgenden **elementaren Umformungen** der erweiterten Koeffizientenmatrix verändern die Lösungsmenge eines linearen Gleichungssystems nicht:

- Multiplikation einer Gleichung (Zeile) mit $\lambda \in \mathbb{K}$, $\lambda \neq 0$,
- Addition des Vielfachen einer Gleichung (Zeile) zu einer anderen Gleichung (Zeile),
- Vertauschung von zwei Gleichungen (Zeilen),
- Umnummerierung von zwei Unbekannten (Vertauschung von zwei Spalten). Hier muß man sich allerdings merken, welche Unbekannten man umnummeriert hat.

Beim Durchführen des **Gauß-Algorithmus** wird das LGS $\mathbf{Ax} = \mathbf{b}$ in ein äquivalentes System $\tilde{\mathbf{A}}\mathbf{x} = \tilde{\mathbf{b}}$ mit einer Matrix $[\tilde{\mathbf{A}} | \tilde{\mathbf{b}}]$ in Trapezform/Zeilenstufenform überführt:

$$\left[\begin{array}{ccccccc|c} \tilde{a}_{1,1} & \tilde{a}_{1,2} & \tilde{a}_{1,3} & \dots & \tilde{a}_{1,r} & \tilde{a}_{1,r+1} & \dots & \tilde{a}_{1,n} & \tilde{b}_1 \\ 0 & \tilde{a}_{2,2} & \tilde{a}_{2,3} & \dots & \tilde{a}_{2,r} & \tilde{a}_{2,r+1} & \dots & \tilde{a}_{2,n} & \tilde{b}_2 \\ 0 & 0 & \tilde{a}_{3,3} & \dots & \tilde{a}_{3,r} & \tilde{a}_{3,r+1} & \dots & \tilde{a}_{3,n} & \tilde{b}_3 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \tilde{a}_{r,r} & \tilde{a}_{r,r+1} & \dots & \tilde{a}_{r,n} & \tilde{b}_r \end{array} \right]$$

mit $\tilde{a}_{j,j} \neq 0$ ($j = 1, \dots, r$). Die Lösung solcher 'gestaffelter' Gleichungssysteme ist durch Rückwärtsauflösen leicht zu bestimmen: Die Variablen $x_{r+1}, \dots, x_n$ sind unbestimmt und werden als freie Parameter gewählt und die sogenannten Leitvariablen/Kopfvariablen $x_1, \dots, x_r$ sind dann durch sukzessives einsetzen zu bestimmen. Zu beachten ist, siehe unten, daß sich die Namen der Variablen geändert haben können.

Bevor wir den Algorithmus im Detail beschreiben, noch zwei kleine Vorüberlegungen:

- Vor jedem Einzelschritt werden zunächst alle Zeilen der Form

$$0 \ 0 \ \dots \ 0 \mid 0$$

ersatzlos gestrichen.

- Gibt es Zeilen der Form

$$0 \ 0 \ \dots \ 0 \mid \tilde{b}_j,$$

(mit $\tilde{b}_j \neq 0$) so wird der Algorithmus abgebrochen; das LGS hat dann keine Lösung.

3.1.1 Beschreibung des Gauß-Algorithmus

Wir beschreiben den Gauß-Algorithmus:

- Schritt 1
 - Ist $a_{1,1} = 0$, so suche in der ersten Spalte nach einem Eintrag $a_{j,1} \neq 0$. Existiert ein solcher, tausche Zeilen 1 und j .
 - Ist $a_{1,1} = 0$ und der erste Punkt war nicht erfolgreich, so existiert in der ersten Zeile ein Eintrag $a_{1,k} \neq 0$. Tausche in diesem Fall Spalten 1 und k ; merke wie die Unbekannten umnummeriert wurden,
 - Ist jetzt $\tilde{a}_{1,1} \neq 0$, so erzeuge unterhalb von $\tilde{a}_{1,1}$ lauter Nullen durch Addition des jeweils $(-\frac{\tilde{a}_{i,1}}{\tilde{a}_{1,1}})$ -fachen der ersten Zeile zur Zeile i .

Die erste Zeile und Spalte von $[\tilde{\mathbf{A}} \mid \tilde{\mathbf{b}}]$ ist damit ermittelt und wird nicht mehr angefaßt.

- Schritt 2 etc. Nach Ausführung der Vorarbeiten wird nun mit der um die erste Zeile und Spalte reduzierten Matrix analog verfahren usw.

Im Falle der Existenz einer Lösung steht am Ende des Gauß-Algorithmus ein äquivalentes System der Form

$$\begin{bmatrix} \tilde{a}_{1,1} & \tilde{a}_{1,2} & \tilde{a}_{1,3} & \dots & \tilde{a}_{1,r} & \tilde{a}_{1,r+1} & \dots & \tilde{a}_{1,n} \\ 0 & \tilde{a}_{2,2} & \tilde{a}_{2,3} & \dots & \tilde{a}_{2,r} & \tilde{a}_{2,r+1} & \dots & \tilde{a}_{2,n} \\ 0 & 0 & \tilde{a}_{3,3} & \dots & \tilde{a}_{3,r} & \tilde{a}_{3,r+1} & \dots & \tilde{a}_{3,n} \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & \tilde{a}_{r,r} & \tilde{a}_{r,r+1} & \dots & \tilde{a}_{r,n} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ \tilde{b}_3 \\ \vdots \\ \tilde{b}_r \end{bmatrix}.$$

Die Variablen $x_{r+1}, \dots, x_n$ können als Parameter frei gewählt werden ($x_{r+1} = \lambda_{r+1}, \dots, x_n = \lambda_n$).

Die anderen Variablen $x_r, x_{r-1}, \dots, x_1$ lassen sich dann bestimmen, indem man die Gleichungen von unten her sukzessive auflöst.

Achtung: Umnummerierungen, d.h. Vertauschungen von Spalten, muß man natürlich berücksichtigen!

Die **elementaren Umformungen** lassen nicht nur die Lösung eines LGS unverändert.

Sie erhalten des weiteren natürlich auch den Rang der Koeffizientenmatrix. Es gilt also mit den bisherigen Bezeichnungen und den Erkenntnissen:

$$\text{rank}(\mathbf{A}) = \text{rank}(\tilde{\mathbf{A}}) = r.$$

Somit steht uns mit dem Gauß-Verfahren auch ein effizientes Verfahren zur Rangbestimmung von Matrizen zur Verfügung.

Übungsaufgabe 3.2.

Trainieren Sie den Gauß-Algorithmus:

1. Bestimmen Sie die Lösungen von $\mathbf{Ax} = \mathbf{b}$ und $\mathbf{Ax} = \mathbf{0}$ für

$$A = \begin{bmatrix} 2 & -1 & -1 & 2 \\ 6 & -2 & 3 & -1 \\ -4 & 2 & 3 & -2 \\ 2 & 0 & 4 & -3 \end{bmatrix} \quad \text{und} \quad \mathbf{b} = \begin{bmatrix} 3 \\ -3 \\ -2 \\ -1 \end{bmatrix}$$

2. Bestimmen Sie die Lösungen von

$$6x_1 + 4x_2 + 8x_3 + 17x_4 = -20$$

$$3x_1 + 2x_2 + 5x_3 + 8x_4 = -8$$

$$3x_1 + 2x_2 + 7x_3 + 7x_4 = -4$$

$$2x_3 - x_4 = 4$$

3. Wieviele Lösungen besitzt das LGS $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$?

Übungsaufgabe 3.3.

Geben Sie die Ränge der Koeffizientenmatrizen aus den Beispielen in Aufgabe 3.2 an.

3.1.2 Der Gauß-Jordan Algorithmus

Beim **Gauß-Jordan-Algorithmus** wird das LGS $\mathbf{Ax} = \mathbf{b}$, $\mathbf{A} \in \mathbb{K}^{m \times n}$, ebenfalls in ein äquivalentes System mit Trapezform überführt.

Allerdings führt man die Umformungen noch weiter und erzeugt schließlich eine Matrix $[\tilde{\mathbf{A}} \mid \tilde{\mathbf{b}}]$, in der $\tilde{\mathbf{A}}$ folgende Blockstruktur besitzt:

$$\tilde{\mathbf{A}} = [\mathbf{I}_r \mid \mathbf{R}] = \begin{bmatrix} 1 & 0 & \dots & 0 & * & \dots & * \\ 0 & 1 & \dots & 0 & * & \dots & * \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & * & \dots & * \end{bmatrix}. \quad (3.1.1)$$

Dabei ist $\mathbf{I}_r$ die r -dimensionale Einheitsmatrix und $\mathbf{R} \in \mathbb{K}^{r \times (n-r)}$ eine beliebige Restmatrix, die auch entfallen kann

Wir wollen hier nur den Fall betrachten, daß $\mathbf{Ax} = \mathbf{b}$ eindeutig lösbar ist ($r = n$). Das Gleichungssystem wird dann in ein äquivalentes System $\mathbf{I}_n \mathbf{x} = \tilde{\mathbf{b}}$ überführt, dessen Lösung $\mathbf{x} = \tilde{\mathbf{b}}$ direkt abgelesen werden kann.

Vorgehen im k -ten Schritt von Gauß-Jordan

- Ist $a_{k,k} = 0$, dann tausche die Zeilen (oder Spalten) wie beim Gauß-Algorithmus, so dass $\tilde{a}_{k,k} \neq 0$.
- Teile nun die k -te Zeile durch $\tilde{a}_{k,k}$, d. h. erzeuge eine 1 an Position (k, k) .
- Erzeuge unter- und oberhalb der Position (k, k) in **allen** Zeilen Nullen durch Addition des jeweils $(-\tilde{a}_{i,k})$ -fachen der k -ten Zeile zur Zeile i .

Achtung: Wie beim Gauß-Algorithmus muß dabei über unterwegs vorgenommene Umnumerierungen von Variablen (Spaltenvertauschungen) sorgfältig buchgeführt werden.

Bestimmung der Lösung im Gauß-Jordan-Algorithmus

Fehlt in (3.1.1) der Block R , so ist das Gleichungssystem $A\mathbf{x} = \mathbf{b}$ äquivalent zu $I_n\mathbf{x} = \tilde{\mathbf{b}}$. Es gibt eine eindeutige Lösung, die man direkt ablesen kann:

$$\mathbf{x} = \tilde{\mathbf{b}}.$$

Ist das Gleichungssystem äquivalent zu $[I_r \ R]\mathbf{x} = \tilde{\mathbf{b}}$ mit $r < n$, so teilt man $\mathbf{x}$ in zwei Teile $\mathbf{x} = \begin{bmatrix} \mathbf{x}^1 \\ \mathbf{x}^2 \end{bmatrix}$ mit $\mathbf{x}^1 = [x_1, \dots, x_r]^T$ und $\mathbf{x}^2 = [x_{r+1}, \dots, x_n]^T$.

Die Variablen $x_{r+1}, \dots, x_n$ können als Parameter frei gewählt werden. Wegen

$$\tilde{\mathbf{b}} = \begin{bmatrix} I_r & R \end{bmatrix} \begin{bmatrix} \mathbf{x}^1 \\ \mathbf{x}^2 \end{bmatrix} = \mathbf{x}^1 + R\mathbf{x}^2$$

ist die Lösung von $A\mathbf{x} = \mathbf{b}$ gegeben durch

$$\mathcal{L} = \left\{ \mathbf{x} = \begin{bmatrix} \mathbf{x}^1 \\ \mathbf{x}^2 \end{bmatrix} : \mathbf{x}^2 \in \mathbb{R}^{n-r} \text{ beliebig und } \mathbf{x}^1 = \tilde{\mathbf{b}} - R\mathbf{x}^2 \right\}. \quad (3.1.2)$$

Ob man ein LGS mittels Gauß- oder Gauß-Jordan-Verfahren löst, ist letztlich Geschmackssache.

Beim Gauß-Jordan-Verfahren spart man das Rückwärtseinsetzen, muß aber eine höhere Anzahl elementarer Umformungen in Kauf nehmen.

Besonders vorteilhaft ist Gauß-Jordan jedoch dann, wenn es um die Lösung mehrerer LGS mit gleicher Koeffizientenmatrix $A \in \mathbb{R}^{n \times n}$, $\text{rank}(\mathbf{A}) = n$, und verschiedenen rechten Seiten geht:

$$A\mathbf{x} = \mathbf{b}_1, \quad A\mathbf{x} = \mathbf{b}_2, \quad \dots, \quad A\mathbf{x} = \mathbf{b}_\ell$$

Diese Aufgabe ist gleichbedeutend mit der Lösung der Matrixgleichung

$$AX = B,$$

wobei $X \in \mathbb{R}^{n \times \ell}$ gesucht ist, und $B \in \mathbb{R}^{n \times \ell}$ als Spalten gerade die rechten Seiten $\mathbf{b}_i$ enthält.

Übungsaufgabe 3.4.

Mit Hilfe des Gauß-Jordan-Algorithmus bestimme man die Lösung der Matrixgleichung $AX = B$, wobei

$$A = \begin{bmatrix} 3 & 5 & 1 \\ 2 & 4 & 5 \\ 1 & 2 & 2 \end{bmatrix} \quad \text{und} \quad B = \begin{bmatrix} 6 & 13 \\ 17 & 10 \\ 7 & 5 \end{bmatrix}.$$

3.2 Inverse von Matrizen

Um für eine gegebene Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ die Inverse $\mathbf{A}^{-1}$ zu berechnen, d.h. eine Matrix $\mathbf{B} \in \mathbb{R}^{n \times n}$ mit

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n$$

müssen wir sämtliche LGSe

$$\mathbf{Ax} = \mathbf{e}_i, \quad i = 1, \dots, n$$

lösen um die Spalten von $\mathbf{B}$ zu erhalten. Für diese Aufgabe bietet sich der Gauß-Jordan-Algorithmus an. Schema:

$$\begin{array}{c|c} A & I_n \\ \hline \vdots & \vdots \\ \hline I_n & A^{-1} \end{array}$$

Dabei werden alle n Gleichungssysteme simultan gelöst.

Übungsaufgabe 3.5.

Berechnen Sie die Inversen von

$$A = \begin{bmatrix} 1 & 23 \\ 0 & 42 \end{bmatrix} \quad \text{und} \quad B = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 3 & 0 \\ 3 & 4 & 1 \end{bmatrix}.$$

3.3 Was man mit Gauß/Gauß-Jordan alles berechnen kann

Berechnung allgemeine Lösung eines LGS.

Input: $\mathbf{A} \in \mathbb{K}^{m \times n}$, $\mathbf{b} \in \mathbb{K}^n$.

Output: Lösungsmenge $\{\mathbf{x} \in \mathbb{K}^n : \mathbf{Ax} = \mathbf{b}\}$

(S1) Bringe die erweiterte Koeffizientenmatrix $[A|b]$ in Parallelogrammform/Zeilenstufenform.

(S2) Löse die Gleichungen auf.

Wenn in S1 Gauß-Jordan verwendet wurde und das System in normierte Zeilenstufenform vorliegt, dann lies die Lösung ab bzw. bringe die Nichtleitvariablen auf die Seite der Inhomogenität und drücke die Leitvariablen in den Nichtleitvariablen aus. Siehe (3.1.2).

Beispiel 3.1.

Wir betrachten die erweiterte Koeffizientenmatrix

$$\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 2 \\ 1 & 2 & 1 & 3 & 1 \\ 0 & 1 & 0 & 1 & 1 \end{array}$$

Wir ziehen die erste Zeile von der zweiten ab und dann die zweite von der dritten und erhalten

$$\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 2 \\ 0 & 1 & 0 & 2 & -1 & II - I \\ 0 & 1 & 0 & 1 & 1 \end{array}$$

und

$$\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 2 \\ 0 & 1 & 0 & 2 & -1 \\ 0 & 0 & 0 & -1 & 2 & III - II \end{array}$$

Damit erhalten wir

$$x_4 = -2$$

$$x_2 = -1 - 2x_4 = 3$$

$$x_1 = 2 - x_4 - x_3 - x_2 = 1 - x_3.$$

Dabei ist x_3 ein freier Parameter und wir erhalten als Lösungsmenge

$$L = \left\{ \mathbf{x} = \begin{bmatrix} 1 \\ 3 \\ 0 \\ -2 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \in \mathbb{R}^4: t \in \mathbb{R} \right\}.$$

Beispiel 3.2.

Wir betrachten die erweiterte Koeffizientenmatrix

$$\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 2 \\ 1 & 2 & 1 & 3 & 1 \\ 0 & 1 & 0 & 1 & 1 \end{array}$$

Wir wenden diesmal Gauß-Jordan an. Die Tableaus einzeln aufzuschreiben überlassen wir der Leserin. Wir erhalten als vorletztes Tableau

$$\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 0 & 1 & -2 \end{array}$$

und wir Tauschen Spalten 3 und 4 und erhalten

$$\begin{array}{cccc|c} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 3 \\ 0 & 0 & 1 & 0 & -2 \end{array}.$$

Dies gibt die Gleichungen

$$x_1 = 1 - x_3$$

$$x_2 = 3$$

$$x_4 = -2.$$

Dabei ist x_3 der frei Parameter und wir erhalten die Lösungsmenge wie oben. Siehe auch (3.1.2).

Ausführlicher korrespondierend zu (3.1.2): Wir setzen die neue Variable x_4 als Parameter, und splitten $\mathbf{x} \in \mathbb{R}^4$ in die zwei Teile $\mathbf{x}^1 = [x_1, x_2, x_3]^T$ und $\mathbf{x}^2 = [x_4]$. Das R in unserer Situation ist gegeben durch $R = [1, 0, 0]^T$ und wir haben $R\mathbf{x}^2 = [x_4, 0, 0]^T$. Damit dann

$$\begin{aligned} \tilde{\mathcal{L}} &= \left\{ \begin{bmatrix} \mathbf{x}^1 \\ \mathbf{x}^2 \end{bmatrix} : \mathbf{x}^1 = \begin{bmatrix} 1 \\ 3 \\ -2 \end{bmatrix} - \begin{bmatrix} x_4 \\ 0 \\ 0 \end{bmatrix}, \mathbf{x}^2 = [x_4] \in \mathbb{R} \text{ bel.} \right\} \\ &= \left\{ \mathbf{x} = \begin{bmatrix} 1 \\ 3 \\ -2 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} : t \in \mathbb{R} \right\}. \end{aligned}$$

Um die Lösungsmenge des Ausgangssystems zu erhalten müssen wir noch die dritte und 4. Zeile tauschen.

$$\mathcal{L} = \left\{ \mathbf{x} = \begin{bmatrix} 1 \\ 3 \\ 0 \\ -2 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} : t \in \mathbb{R} \right\}$$

Berechnung einer Basis des Kerns.

Input: $\mathbf{A} \in \mathbb{K}^{m \times n}$, $r = \text{rank}(\mathbf{A}) > 0$.

Output: Basis $\{\mathbf{b}_1, \dots, \mathbf{b}_{n-r}\}$ des Kerns $N(\mathbf{A})$.

(S1) Bringe die Matrix A mit elementaren Umformungen auf normierte reduzierte Zeilenstufenform $\tilde{A}$, d.h. über den Leitvariablen stehen stets 0en und die Leitvariablen haben Koeffizienten 1.

Betrachte das homogene System $\tilde{A}\mathbf{x} = \mathbf{0}$.

(S2) Zerlege die Variable $\mathbf{x} = [\mathbf{x}^1, \mathbf{x}^2]^T$ in Kopf- $\mathbf{x}^1 = [x_1, \dots, x_r]^T$ und Nichtkopfvariablen $\mathbf{x}^2 = [x_{r+1}, \dots, x_n]$. Zu jeder der $n - r$ Nichtkopfvariablen in $\mathbf{x}^2$ konstruieren wir dann einen Basisvektor $\mathbf{b}_i$:

- (i) Setze die $r + i - 1$ -te Koordinate von $\mathbf{b}_i$, $1 \leq i \leq n - r$ zu 1.
- (ii) Für alle anderen Nichtkopfvariablen x_k setze die k -te Komponente von $\mathbf{b}_i$ zu 0.
- (iii) Die restlichen Komponenten ergeben sich aus den i ten Spalten der Matrix R .

Hintergrund: Es gilt $[\mathbf{I}_r \ \mathbf{R}][\mathbf{x}^1, \mathbf{x}^2]^T = \mathbf{0}$ also

$$R\mathbf{x}^2 = -\mathbf{x}^1.$$

Wir haben die Vektoren $\mathbf{x}^2$ als Einheitsvektoren $\mathbf{e}_i$, $1 \leq i \leq n - r$ der korrekten Dimension festgesetzt und damit ergeben sich entsprechend die $\mathbf{x}^1$ also $\mathbf{x}^1 = -\mathbf{R}\mathbf{e}_i$, also die i -te Spalte von $\mathbf{R}$.

- (iv) Beachte evtl. Vorgenommene Umnummerierung der Variablen durch Spaltentausch!

Beispiel 3.3.

Wir wollen eine Basis des Kerns von

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 2 & 0 & 1 & 0 & 1 \end{bmatrix}$$

berechnen. Dazu bringen wir die Matrix zuerst in die Form $\tilde{A} = [I_r R]$:

$$\tilde{\mathbf{A}} = \begin{bmatrix} 1 & 0 & 0 & \frac{1}{3} & \frac{2}{3} \\ 0 & 1 & 0 & \frac{4}{3} & \frac{2}{3} \\ 0 & 0 & 1 & -\frac{2}{3} & -\frac{1}{3} \end{bmatrix}.$$

Die Berechnung dieser Form ist der Leserin zur Übung überlassen. Die Basisvektoren ergeben sich nach der obigen Beschreibung wie folgt

$$\mathbf{b}_1 = \begin{bmatrix} \\ \\ \\ 1 \\ 0 \end{bmatrix} \quad \text{und} \quad \mathbf{b}_2 = \begin{bmatrix} \\ \\ \\ 0 \\ 1 \end{bmatrix}$$

Nur x_4 und x_5 sind Nichtkopfvariablen. Der erste Vektor korrespondiert zu x_4 , also in der 4. Komponente eine 1 und in der 5. Komponente eine 0. Für $\mathbf{b}_2$ entsprechend. Weiter erhalten wir

$$\mathbf{b}_1 = \begin{bmatrix} -\frac{1}{3} \\ \frac{2}{3} \\ -\frac{4}{3} \\ 1 \\ 0 \end{bmatrix} \quad \text{und} \quad \mathbf{b}_2 = \begin{bmatrix} -\frac{2}{3} \\ \frac{1}{3} \\ -\frac{2}{3} \\ 0 \\ 1 \end{bmatrix}.$$

Natürlich können wir noch mit 3 multiplizieren um eine schönere Basis ohne Brüche zu bekommen.

Bemerkung 3.2.

Wenn sie die allgemeine Lösung eines Gleichungssystems $\mathbf{Ax} = \mathbf{b}$ bestimmt haben, dann haben Sie nach Satz 3.1 natürlich auch eine Basis des Kerns $N(\mathbf{A})$ und in einer entsprechenden Aufgabe müßte diese Information nur noch extrahiert werden. Extrahieren Sie eine Basis von $N(\mathbf{A})$ aus Beispiel 3.1.

Berechnung einer Basis des Bildes von $\mathbf{A}$

Input: $\mathbf{A} \in \mathbb{K}^{m \times n}$.

Output: Basis $\{\mathbf{b}_1, \dots, \mathbf{b}_r\}$ des Bildraumes $R(\mathbf{A})$.

- (S1) Transponiere $\mathbf{A}$: Die Zeilen von $\mathbf{A}^T$ bilden ein Erzeugendensystem des Bildes von $\mathbf{A}$, da die Spalten von $\mathbf{A}$ ein Erzeugendensystem von $R(\mathbf{A})$ bilden.
- (S2) Man bringe also $\mathbf{A}^T$ in Trapezform/Zeilenstufenform. Die Nichtnullzeilen bilden dann eine Basis des Bildes von $\mathbf{A}$.

Beispiel 3.4.

Wir betrachten wieder

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 2 & 0 & 1 & 0 & 1 \end{bmatrix}$$

und wollen diesmal eine Basis des Bildes Berechnen. Da $n = 5$ gilt muß nach Beispiel 3.3 die Dimension des Bildes $R(\mathbf{A})$ gleich 3 sein. Wir berechnen aus

$$\mathbf{A}^T = \begin{bmatrix} 1 & 1 & 2 \\ 2 & 1 & 0 \\ 3 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

durch Anwendung des Gauß-Algorithmus die Basis

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Berechnung der Inversen. Siehe Sektion 3.2

Input: $\mathbf{A} \in \mathbb{K}^{n \times n}$, $r = \text{rank}(\mathbf{A}) = n$.

Output: Inverse A^{-1}

- (S1) Wende den Gauß-Jordan Algorithmus auf $[\mathbf{A} | \mathbf{I}_n]$ an und erzeuge, wenn möglich, die Matrix $[\mathbf{I}_n | \tilde{\mathbf{A}}]$. Die Matrix $\tilde{\mathbf{A}}$ ist die gesuchte Inverse.
- (S2) Wenn $\text{rank}(A) < n$, dann merkt man dies beim berechnen weil auf der Diagonalen der linken $n \times n$ Matrix eine Null entsteht.

Koordinatendarstellung

Input: Sei $\mathbf{x} \in \mathbb{K}^n$, $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ sei eine Basis von $\mathbb{K}^n$

Output: Koordinatenvektor $[\alpha_1, \dots, \alpha_n]^T$ bzgl. B .

- (S1) Löse das Gleichungssystem $\mathbf{A}\alpha = \mathbf{x}$ wobei $\mathbf{A}$ die Matrix

$$\mathbf{A} = [\mathbf{b}_1 | \dots | \mathbf{b}_n]$$

ist.

- (S2) Der Lösungsvektor α ist der Koordinatenvektor.

Es gilt auch $\alpha = A^{-1}\mathbf{x}$, d.h. man kann auch die Inverse von A berechnen um die Koordinaten von $\mathbf{x}$ bzgl. einer Basis zu bestimmen. Dies ist ökonomischer, wenn man die Koordinaten von mehr als einem Vektor bzgl. der gleichen Basis berechnen will.

3.4 Verständnisfragen

Viele der Fragen sind aus [4], eventuell modifiziert. Auf den Einzelnachweis wird verzichtet. Die Studierenden sind angehalten die Aufgaben selbst zu lösen und die Resultate miteinander zu besprechen.

1. Für die Koeffizientenmatrix eines inhomogenen linearen Gleichungssystems $\mathbf{Ax} = \mathbf{b}$ gelte $\mathbf{A} \in \mathbb{R}^{m \times n}$.
 - a) Aus wievielen Gleichungen mit wievielen Variablen besteht das System?
 - b) Von welchem Typ ist die erweiterte Koeffizientenmatrix?
 - c) Wieviele Koordinaten besitzen die Vektoren $\mathbf{x}$ und $\mathbf{b}$?
2. Welche Möglichkeiten gibt es, mit den Matrizen $\mathbf{A}$, $\mathbf{B}$ und $\mathbf{C}$ als Koeffizientenmatrizen und den Vektoren $\mathbf{b}_k \in \mathbb{R}^k$, $k = 1, 2, 3, 4, 5$, lineare Gleichungssysteme zu bilden, wenn $\mathbf{A} \in \mathbb{R}^{5 \times 3}$, $\mathbf{B} \in \mathbb{R}^{2 \times 4}$ und $\mathbf{C} \in \mathbb{R}^{3 \times 3}$?
3. Für die Koeffizientenmatrix von $\mathbf{Ax} = \mathbf{b}$ gelte $\mathbf{A} \in \mathbb{R}^{5 \times 3}$.
 - a) Für welche Beziehung zwischen Rang der Matrix $\mathbf{A}$ und dem Rang der erweiterten Koeffizientenmatrix $[\mathbf{A}|\mathbf{b}]$ ist das System lösbar?
 - b) Unter welchen Bedingungen ist es eindeutig lösbar bzw. existieren unendliche viele Lösungen.
4. Wodurch wird die Dimension d des Lösungsraumes eines homogenen LGS bestimmt, und wie lautet die allgemeine Lösung?
Welche Lösung erhält man für $d = 0$?
5. Warum kann $\text{rank}(\mathbf{A}) < \text{rank}([\mathbf{A}|\mathbf{b}])$ bei einem LGS $\mathbf{Ax} = \mathbf{b}$ im homogenen Fall nicht eintreten?
Welcher Schluß ergibt sich daraus für die Lösbarkeit bzw. Unlösbarkeit des LGS?
6. Kann ein inhomogenes LGS $\mathbf{Ax} = \mathbf{b}$ mit $\mathbf{A} \in \mathbb{R}^{m \times n}$ im Falle $m < n$ eindeutig lösbar sein/?
Kann es im Fall $m > n$ unendlich viele Lösungen besitzen?

7. Wie kann man mit Hilfe einer Rangbetrachtung für die Koeffizientenmatrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ feststellen, ob ein homogenes bzw. inhomogenes LGS eindeutig lösbar ist.
8. Welche Rangabschätzung gilt für ein homogenes LGS mit 4 Variablen und 8 Gleichungen, wenn ein nichttrivialer Lösungsraum existieren soll? Welche Dimension kann der Lösungsraum besitzen?
9. Welche Aussagen gelten für die Anzahl der Variablen bzw. für die Mindestanzahl der Gleichungen sowie den Rang der (erweiterten) Koeffizientenmatrix eines LGS, wenn die allgemeine Lösung die folgenden Gestalt hat?

$$\text{a) } \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = t \begin{bmatrix} -1 \\ 0 \\ 0 \\ 4 \end{bmatrix} + s \begin{bmatrix} 0 \\ 3 \\ 0 \\ -3 \end{bmatrix}, t, s \in \mathbb{R}$$

$$\text{b) } \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 2 \\ 5 \end{bmatrix} + s \begin{bmatrix} 2 \\ -1 \\ 1 \\ 6 \end{bmatrix}, t, s \in \mathbb{R}$$

Charakterisieren die das jeweilige LGS als homogen oder inhomogen.

10. Welcher Zusammenhang besteht im Falle unendlich vieler Lösungen zwischen der allgemeinen Lösung eines inhomogenen LGS $\mathbf{Ax} = \mathbf{b}$ und der allgemeinen Lösung des zugehörigen homogenen LGS $\mathbf{Ax} = \mathbf{0}$?
11. Welche (bestmögliche) obere Abschätzung gilt für den Rang einer Matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$?
12. Welchen Rang hat eine 3×3 Matrix $\mathbf{A}$, bei welcher der erste Spaltenvektor (kein Nullvektor) dem Zweifachen des 2. Spaltenvektors entspricht und der 3. Spaltenvektor der Nullvektor ist?
13. Unter welchen Voraussetzungen sind folgende Matrixgleichungen nach $\mathbf{X}$ auflösbar?

a) $\mathbf{A}^T \mathbf{X} - \mathbf{B} = \mathbf{B}^T \mathbf{A},$

b) $\mathbf{A}^T \mathbf{B} - \mathbf{B}^T (\mathbf{A} - \mathbf{X}) = \mathbf{I}_n.$

14. Welchen Rang hat die Matrix $\mathbf{A} = [a_{ij}]_{1 \leq i, j \leq n}$ mit

$$a_{ij} = \begin{cases} 1 & : i = j \\ 0 & : i \neq j \end{cases} ?$$

Wie begründet man dies?

3.5 Rechenaufgaben

Aufgabe 8

Begründen Sie, wieso die folgenden linearen Gleichungssysteme eindeutig lösbar sind und geben Sie einen Lösungsvektor an.

$$(a) \begin{cases} 3x_1 + 2x_2 + 2x_3 = -1 \\ x_1 + 6x_2 - x_3 = 3 \\ 4x_1 + x_2 + 5x_3 = -6 \end{cases}$$

$$(b) \begin{cases} x_1 + 3x_2 + 2x_3 = 0 \\ 2x_1 - 2x_2 + 5x_3 = 0 \\ -3x_1 + 3x_2 - 2x_3 = 0 \end{cases}$$

Aufgabe 9

Überprüfen Sie die Lösbarkeit folgender linearer Gleichungssysteme, und ermitteln Sie im Falle der Lösbarkeit den Lösungsvektor bzw. die allgemeine Lösung.

$$(a) \begin{cases} x_1 - x_2 + 3x_3 = 0 \\ 2x_1 + 3x_2 - x_3 = 0 \\ 3x_1 + 7x_2 - 5x_3 = 0 \end{cases}$$

$$(b) \begin{cases} 3x_1 - x_2 + 2x_3 = 1 \\ 7x_1 - 4x_2 - x_3 = -2 \\ -x_1 - 3x_2 - 12x_3 = -5 \\ -x_1 + 2x_2 + 5x_3 = 2 \\ 5x_2 + 17x_3 = 7 \end{cases}$$

$$(c) \begin{cases} 6x_1 + 4x_2 + 8x_3 + 17x_4 = -20 \\ 3x_1 + 2x_2 + 5x_3 + 8x_4 = -8 \\ 3x_1 + 2x_2 + 7x_3 + 7x_4 = -4 \\ 2x_3 - x_4 = 4 \end{cases}$$

$$(d) \begin{cases} x_1 + 2x_2 + 2x_3 & = -1 \\ 2x_1 + 4x_2 + 3x_3 - x_4 & = 1 \\ -x_1 - 2x_2 + x_3 + 2x_4 & = -8 \\ -3x_1 - 6x_2 + 2x_3 + 3x_4 & = -21 \end{cases}$$

$$(e) \begin{cases} 3x_1 & + 6x_3 - 3x_4 & = 1 \\ 2x_1 & & + 3x_5 & = 0 \\ & x_2 - 4x_3 + 2x_4 & = -2 \\ 4x_1 + 3x_2 & & - 3x_5 & = -4 \end{cases}$$

$$(f) \begin{cases} 3x_1 + x_2 - x_3 & + 4x_5 = 5 \\ -2x_1 + 2x_2 & + x_4 - 5x_5 = 0 \\ & - x_2 + 2x_3 - 3x_4 - 4x_5 = -5 \\ x_1 + 3x_2 + 3x_3 & - 2x_5 = 3 \\ 2x_1 & + 3x_3 + x_4 + 2x_5 = 18 \end{cases}$$

Aufgabe 10

Welchen Rang haben die folgenden Matrizen.

$$(a) \mathbf{A}_1 = \begin{bmatrix} 2 & 3 \\ 4 & 1 \\ 0 & 2 \end{bmatrix}$$

$$(d) \mathbf{A}_4 = \begin{bmatrix} 1 & 0 & 2 \\ 3 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix}$$

$$(b) \mathbf{A}_2 = \begin{bmatrix} 2 & 3 \\ 6 & 9 \end{bmatrix}$$

$$(e) \mathbf{A}_5 = \begin{bmatrix} 3 & 0 & -2 & 5 \\ 0 & 2 & 7 & -1 \\ 6 & 6 & 17 & 7 \end{bmatrix}$$

$$(c) \mathbf{A}_3 = \begin{bmatrix} 1 & -2 & 7 \\ 2 & 1 & 4 \\ 3 & 4 & 1 \end{bmatrix}$$

$$(f) \mathbf{A}_6 = \begin{bmatrix} 2 & 1 & 2 & -1 & 4 \\ 3 & 7 & 0 & -1 & 5 \\ 1 & 2 & -3 & -1 & -4 \\ 1 & -8 & 12 & 0 & 15 \end{bmatrix}$$

Aufgabe 11

Drei Ebenengleichungen sind gegeben durch

$$\begin{cases} x - y - z = -2 \\ 3x + y - z = b \\ ax + 8y + 2z = 7 \end{cases}$$

mit Parametern $a, b \in \mathbb{R}$. Berechnen Sie a und b so, daß sich die drei Ebenen in einer Geraden schneiden!

Geben Sie für die berechneten Werte eine Parameterdarstellung der Schnittgeraden an!

Aufgabe 12

Gegeben ist das lineare Gleichungssystem

$$\begin{cases} x - 3y + 4z = 4 \\ -5x \quad \quad + z = a \\ -x + 3y + z = -12 \\ x \quad \quad \quad - 3z = -1 \end{cases}.$$

Machen Sie, in Abhängigkeit von a Aussagen über die Lösungsmenge des Gleichungssystems.

Aufgabe 13

Gegeben ist das Gleichungssystem

$$\begin{cases} x - 2y + 3z = -4 \\ 2x + y + z = 2 \\ x + \alpha y + 2z = -\beta \end{cases}$$

mit $\alpha, \beta \in \mathbb{R}$ freie Parameter.

(i) Für welche α und β hat das System genau eine Lösung?

(ii) Für welche α und β ist das System unlösbar?

(iii) Für welche α und β existieren unendlich viele Lösungen?

Aufgabe 14

Gegeben ist das Gleichungssystem

$$\begin{cases} x - 2y + 3z = -4 \\ 2x + y + z = 2 \\ x + \alpha y + 2z = -\beta \end{cases}$$

mit $\alpha, \beta \in \mathbb{R}$.

- (i) Für welche α und β hat das System genau eine Lösung?
- (ii) Für welche α und β ist das System unlösbar?
- (iii) Für welche α und β existieren unendliche viele Lösungen?

Determinanten

Die **Determinante** ist eine Abbildung, die einer quadratischen Matrix $A \in \mathbb{K}^{n \times n}$ eine Zahl aus $\mathbb{K}$ zuordnet:

$$\det : \mathbb{K}^{n \times n} \rightarrow \mathbb{K}$$

$$\mathbf{A} \mapsto \det(\mathbf{A})$$

Sie wird u. A. benutzt zur Beschreibung

- der eindeutigen Lösbarkeit von LGS,
- der linearen Unabhängigkeit von n Vektoren,
- der Veränderung von Flächen/Volumina bei Anwendung der Abbildung $\mathbf{A}$.

Bemerkung 4.1 (Historisches zum Determinantenbegriff).

Determinanten mit $n = 2$ wurden erstmals am Ende des 16. Jh. von Cardano, größere ca. 100 Jahre später von Leibniz behandelt. Ein moderner axiomatischer Ansatz (1864) geht auf Weierstraß zurück.

Wir werden den axiomatischen Aufbau hier umgehen und für die Fälle mit $n \leq 3$ explizite Formeln angeben. Für $n > 3$ definieren wir die Determinante induktiv.

Wir definieren für $n = 1$:

$$\mathbf{A} = \begin{bmatrix} a_{1,1} \end{bmatrix}, \quad \det(\mathbf{A}) := a_{1,1}.$$

Und schließlich für $n = 2$ setzen wir

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}, \quad \det(\mathbf{A}) := a_{1,1}a_{2,2} - a_{1,2}a_{2,1}.$$

und für $n = 3$

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}$$

$$\det(\mathbf{A}) := +a_{1,1}a_{2,2}a_{3,3} + a_{1,2}a_{2,3}a_{3,1} + a_{1,3}a_{2,1}a_{3,2} \\ - a_{1,3}a_{2,2}a_{3,1} - a_{1,1}a_{2,3}a_{3,2} - a_{1,2}a_{2,1}a_{3,3}.$$

Im Fall $n = 3$ verwendet man zur Berechnung einer Determinante gerne folgendes Schema:

$$\begin{array}{ccccc} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,1} & a_{2,2} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,1} & a_{3,2} \\ - & - & -, + & + & + \end{array}$$

Vorgehen im obigen Schema (Regel von Sarrus¹):

- Man schreibt 1. und 2. Spalte nochmals neben die Determinante.
- Entlang der Diagonalen ermittelt man die Produkte der Einträge und versieht die Ergebnisse mit den dargestellten Vorzeichen.
- Man summiert die 6 vorzeichenbehafteten Produkte.

¹Pierre Frédéric Sarrus, 1798-1861, französischer Mathematiker

Für Determinanten mit $n \geq 4$ gehen wir induktiv vor, greifen also auf Determinanten geringerer Größe zurück.

Für $i, j \in \{1, 2, \dots, n\}$ bezeichne $\mathbf{A}_{i,j} \in \mathbb{K}^{(n-1) \times (n-1)}$ diejenige Matrix, die durch Streichen der Zeile i und der Spalte j aus $\mathbf{A}$ entsteht.

Sei j nun ein beliebiger Spaltenindex. Dann definieren wir

$$\det(\mathbf{A}) := \sum_{i=1}^n a_{i,j} (-1)^{i+j} \det(\mathbf{A}_{i,j}) \quad (4.0.1)$$

Diese Darstellung wird **Laplace-Entwicklung** nach der j -ten Spalte genannt.

- Natürlich müßte man zunächst zeigen, daß die Darstellung in (4.0.1) von der gewählten Spalte unabhängig ist.
- Bei Berechnungen mit Hilfe der Laplace-Entwicklung ist es häufig zweckmäßig, Spalten (oder Zeilen, siehe später) mit möglichst vielen Nulleinträgen zu wählen.
- Die Laplace-Entwicklung (4.0.1) gilt auch für Determinanten mit $n = 2$ oder 3 .
- In der Literatur wird die Determinante häufig anders aufgebaut. Formel (4.0.1) ist dann ein Bestandteil des **Laplaceschen Entwicklungssatzes**.

Schreibregelung Man verwendet für Determinanten mit $n \geq 2$ auch folgende verkürzende Schreibweise:

$$\begin{vmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{vmatrix} := \det \left(\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,n} \end{bmatrix} \right).$$

Übungsaufgabe 4.1.

Berechnen Sie folgende Determinanten:

$$\begin{vmatrix} 1 & 2 \\ 2 & 4 \end{vmatrix}, \quad \begin{vmatrix} 1 & 0 & 2 \\ 3 & -1 & 1 \\ 3 & 1 & 5 \end{vmatrix}, \quad \begin{vmatrix} 1 & 42 & 23 \\ 0 & 17 & -110 \\ 0 & 0 & 3 \end{vmatrix}, \quad \begin{vmatrix} 1 & 0 & 1 & 3 \\ -1 & 2 & 0 & 0 \\ 0 & 2 & -1 & 0 \\ 1 & -2 & 0 & -1 \end{vmatrix}.$$

Satz 4.1 (Rechenregeln Determinante).

Für $\mathbf{A} = [a_{i,j}] \in \mathbb{K}^{n \times n}$ und $\lambda \in \mathbb{K}$ gelten:

- (i) Multipliziert man eine Spalte von $\mathbf{A}$ mit λ , so multipliziert sich auch die Determinante mit λ :

$$\begin{vmatrix} a_{1,1} & \cdots & \lambda a_{1,j} & \cdots & a_{1,n} \\ a_{2,1} & \cdots & \lambda a_{2,j} & \cdots & a_{2,n} \\ \vdots & & \vdots & & \vdots \\ a_{n,1} & \cdots & \lambda a_{n,j} & \cdots & a_{n,n} \end{vmatrix} = \lambda \cdot \begin{vmatrix} a_{1,1} & \cdots & a_{1,j} & \cdots & a_{1,n} \\ a_{2,1} & \cdots & a_{2,j} & \cdots & a_{2,n} \\ \vdots & & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j} & \cdots & a_{n,n} \end{vmatrix}$$

- (ii) Die Determinante einer Matrix $\mathbf{A}$ ändert sich nicht, wenn man zu einer Spalte das Vielfache einer **anderen** Spalte addiert.
- (iii) Vertauscht man in einer quadratischen Matrix zwei verschiedene Spalten, so multipliziert sich die Determinante mit -1 .

Satz 4.2.

Seien $\mathbf{A}, \mathbf{B} \in \mathbb{K}^{n \times n}$, $\mathbf{A} = [a_{i,j}]$ und $\lambda \in \mathbb{K}$. Dann gelten:

(i) $\det(\lambda \mathbf{A}) = \lambda^n \det(\mathbf{A})$.

(ii) $\det(\mathbf{A}^T) = \det(\mathbf{A})$.

(iii) $\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$.

(iv) Ist $\mathbf{A}$ eine Dreiecksmatrix (d.h. $a_{i,j} = 0$ für alle $i < j$ oder alle $i > j$), dann gilt

$$\det(\mathbf{A}) = \prod_{i=1}^n a_{i,i},$$

d.h. die Determinante ist das Produkt der (Haupt-) Diagonalelemente.

Bemerkung 4.2.

Ein paar wichtige ergänzende Bemerkungen:

- Wegen (4.0.1) und Satz 4.2, Punkt 2, läßt sich eine Determinante auch nach der i -ten Zeile entwickeln:

$$\det(\mathbf{A}) = \sum_{j=1}^n a_{i,j} (-1)^{i+j} \det(\mathbf{A}_{i,j})$$

- Elementare Zeilen- und Spaltenumformungen wie im Gauß-Algorithmus können die Determinante der Koeffizientenmatrix ändern. Die Eigenschaft, gleich oder ungleich Null zu sein, bleibt dabei aber immer erhalten.
- Die Determinante einer Dreiecksmatrix ist genau dann Null, wenn für eines der Diagonalelemente $a_{i,i} = 0$ gilt.

Bemerkung 4.3.

Kombiniert man die letzten beiden Punkte der Bemerkung mit den Kenntnissen über den Endzustand der Matrix $\tilde{\mathbf{A}}$ beim Gauß-Algorithmus, so ergibt sich weiterhin:

- Ist $\text{rank}(\mathbf{A}) = n$, so ist $\det(\mathbf{A}) \neq 0$.
- Ist $\text{rank}(\mathbf{A}) < n$, so ist $\det(\mathbf{A}) = 0$.

Dabei ist natürlich immer $\mathbf{A} \in \mathbb{K}^{n \times n}$ anzusetzen. Das bedeutet wiederum, daß man die eindeutige Lösbarkeit eines linearen Gleichungssystems leicht mit der Determinante verifizieren kann.

Übungsaufgabe 4.2.

Bestimmen Sie alle $\alpha \in \mathbb{R}$, für die das LGS $\mathbf{A}\mathbf{x} = \mathbf{b}$ mit

$$\mathbf{A} = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & 2\alpha & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

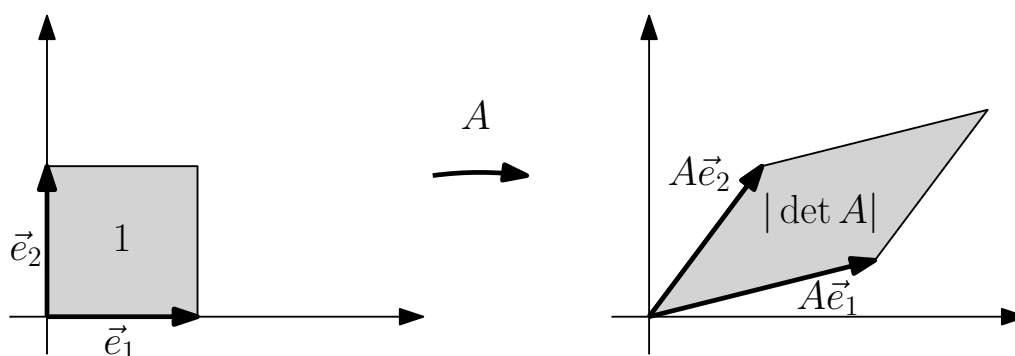
und beliebigem $\mathbf{b} \in \mathbb{R}^3$ eine eindeutige Lösung besitzt

4.1 Geometrische Interpretation

Sei $\mathbf{A} \in \mathbb{R}^{n \times n}$ und $\text{vol}(S)$ das n -dimensionale Volumen² einer geeigneten³ Punktmenge S . Dann ist das Volumen des Bildes $f(S)$ unter $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$ gegeben durch

$$\text{vol}(f(S)) = |\det(\mathbf{A})| \cdot \text{vol}(S).$$

Determinanten können also als Flächen- oder Volumenverzerrungsfaktor interpretiert werden:



²in 2D ist das eine Fläche, in 3D das gewohnte Volumen.

³'Geeignet' lässt sich im Rahmen der Maßtheorie sauber definieren. Gängige geometrische Objekte machen i.d.R. kein Probleme.

4.2 Inverse von Matrizen

Die Inverse einer invertierbaren 2×2 -Matrix

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}$$

kann man explizit angeben:

$$\begin{aligned} \mathbf{A}^{-1} &= \frac{1}{\det(\mathbf{A})} \begin{bmatrix} a_{2,2} & -a_{1,2} \\ -a_{2,1} & a_{1,1} \end{bmatrix} \\ &= \frac{1}{a_{1,1}a_{2,2} - a_{1,2}a_{2,1}} \begin{bmatrix} a_{2,2} & -a_{1,2} \\ -a_{2,1} & a_{1,1} \end{bmatrix}. \end{aligned}$$

Für $n \geq 3$ existieren zwar auch Formeln, diese sind jedoch sperrig und kaum in Gebrauch.

Übungsaufgabe 4.3.

Bestätigen Sie die obige Lösungsformel für $n = 2$ durch nachrechnen!

Satz 4.3.

Sei $A \in \mathbb{K}^{n \times n}$. Dann sind folgende Aussagen äquivalent:

- $\mathbf{A}$ ist invertierbar.
- $\det(\mathbf{A}) \neq 0$.
- Es gilt $\text{rank}(\mathbf{A}) = n$, d.h. die Spalten (Zeilen) von $\mathbf{A}$ bilden eine Basis des $\mathbb{K}^n$.
- Das homogene System $\mathbf{Ax} = \mathbf{0}$ besitzt nur die triviale Lösung $\mathbf{x} = \mathbf{0}$.
- Für jede rechte Seite $\mathbf{b} \in \mathbb{K}^n$ besitzt das System $\mathbf{Ax} = \mathbf{b}$ genau eine Lösung, nämlich

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}.$$

4.3 Verständnisfragen

Die Fragen sind meißt aus [4], eventuell leicht modifiziert. Die Studierenden sind angehalten die Aufgaben selbst zu lösen und die Resultate miteinander zu besprechen.

1. Worauf beruht, daß $\det(\mathbf{A}) = \det(\mathbf{A}^T)$?
2. Es sei $\mathbf{Ax} = \mathbf{0}$ ein quadratischen LGS. Unter welcher Bedingungen für $\det(\mathbf{A})$ ist das System eindeutig bzw. nicht eindeutig lösbar.
3. Welche Forderung ist notwendig, damit eine Diagonalmatrix regulär ist?
4. Welchen Rang hat eine reguläre Matrix?

4.4 Rechenaufgaben

Aufgabe 15

Prüfen Sie ob die folgenden Matrizen regulär sind, und berechnen Sie gegebenenfalls die inverse Matrix.

$$(a) \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}$$

$$(b) \begin{bmatrix} 3 & -6 \\ -4 & 8 \end{bmatrix}$$

$$(c) \begin{bmatrix} 0 & -3 \\ 2 & 5 \end{bmatrix}$$

$$(d) \begin{bmatrix} 2 & 1 & 0 \\ -1 & 3 & 5 \\ 3 & -2 & 2 \end{bmatrix}$$

$$(e) \begin{bmatrix} 2 & 3 & -1 \\ 1 & 2 & 4 \\ 3 & 4 & -6 \end{bmatrix}$$

$$(f) \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -2 & 0 & 3 \end{bmatrix}$$

$$(g) \begin{bmatrix} 1 & n & 0 & 0 \\ 0 & 1 & n & 0 \\ 0 & 0 & 1 & n \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$(h) \begin{bmatrix} 2 & 3 & -1 \\ 1 & 2 & 4 \\ 3 & 4 & -6 \end{bmatrix}$$

Aufgabe 16

Zeigen Sie, daß jede quadratische Matrix $\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ folgenden Gleichungen genügt:

$$1. \mathbf{A}^2 - \operatorname{tr}(\mathbf{A})\mathbf{A} + \det(\mathbf{A})\mathbf{I}_2 = \mathbf{O}.$$

$$2. \frac{1}{c-b}[\mathbf{A}^2 - (\mathbf{A}^T)^2] - (a+d) \left[\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} - \mathbf{I}_2 \right] = \mathbf{O}.$$

5

Längen-, Abstands- und Winkelmessung im $\mathbb{R}^n$

5.1 Längen- und Abstandsmessung im $\mathbb{R}^n$

In der Schule hat man den Satz des Pythagoras gelernt. Mit diesem kann man den Vektoren im geometrischen Modell $\mathbb{R}^2$ bzw. $\mathbb{R}^3$ eine Länge zuordnen. In der Schule haben wir die mit $|\mathbf{x}|$ bezeichnet und

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

gesetzt wobei die Koordinaten bzgl. der Standardbasis zu verstehen sind. Wir wollen ab sofort den etwas gebildeteren Namen **Norm** verwenden und auch die Notation vom Betrag einer reellen/komplexen Zahl abgrenzen. Insbesondere dann, wenn wir eine andere 'Länge' verwenden als die Euklidische (Pythagoras).

Definition 5.1 (Norm).

Eine Funktion $\|\cdot\|: \mathbb{R}^n \rightarrow [0, \infty)$ heißt genau dann **Norm**, falls

- (i) Es gilt $\|\mathbf{x}\| = 0$ genau dann, wenn $\mathbf{x} = \mathbf{0}$.
- (ii) Für alle $\lambda \in \mathbb{R}$ und $\mathbf{x} \in \mathbb{R}^n$ gilt $\|\lambda\mathbf{x}\| = |\lambda|\|\mathbf{x}\|$.
- (iii) Für alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ gilt $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

Wenn man $\mathbb{R}$ überall durch $\mathbb{C}$ ersetzt, ändert sich an dieser Definition nichts.

Beispiel 5.1.

Das wichtigste Beispiel für den Moment ist die Euklidische Norm, oder ℓ^2 -Norm

$$\|x\|_2 := \sqrt{\sum_{k=1}^n |x_k|^2}.$$

Weitere wichtige Normen sind die sogenannte ℓ^1 -Norm

$$\|x\|_1 := \sum_{k=1}^n |x_k|$$

und allgemein für $p \in [1, \infty)$ die ℓ^p -Norm

$$\|\mathbf{x}\|_p = \left(\sum_{k=1}^n |x_k|^p \right)^{\frac{1}{p}}.$$

Für $p = \infty$ modifiziert man

$$\|x\|_\infty := \max_{k=1}^n |x_k|.$$

Diese Norm heißt die ℓ^∞ -Norm bzw. Supremumsnorm oder Maximumsnorm. Für uns sind $\|\cdot\|_i$ für $i = 1, 2, \infty$ besonders relevant und interessant.

Übungsaufgabe 5.1.

Zeichnen Sie die Mengen

(i) $\{\mathbf{x} \in \mathbb{R}^2: \|\mathbf{x}\|_2 = 1\},$

(ii) $\{\mathbf{x} \in \mathbb{R}^2: \|\mathbf{x}\|_1 = 1\}$ und

(iii) $\{\mathbf{x} \in \mathbb{R}^2: \|\mathbf{x}\|_\infty = 1\}$

in $\mathbb{R}^2$.

Der folgende Begriff einer **Metrik** formalisiert den **Abstandsbegriff**. Schon in der Schule hatten wir gelernt, daß ein Längenbegriff einen Abstandsbegriff erzeugt/induziert. Wir haben den Abstand der Punkte A und B durch $\|\overrightarrow{OB} - \overrightarrow{OA}\|_2$ definiert. Das wollen wir jetzt auch tun aber eben erst noch den Begriff des Abstandes etwas klarer fassen.

Definition 5.2 (Metrik).

Eine Funktion $d: \mathbb{R}^n \times \mathbb{R}^n \rightarrow [0, \infty)$ heißt genau dann **Metrik**, wenn für alle $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ gilt:

- (i) $d(\mathbf{x}, \mathbf{y}) = 0$ genau dann, wenn $\mathbf{x} = \mathbf{y}$, (Definitheit)
- (ii) $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ (Symmetrie)
- (iii) $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$ (Dreiecksungleichung)

Wenn Wir eine Norm $\|\cdot\|$ auf $\mathbb{R}^n$ gegeben haben, dann ist mit

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| \tag{5.1.1}$$

auch eine Metrik gegeben. Wenn wir für $\|\cdot\|$ die $\|\cdot\|_2$ -Norm verwenden, dann bekommen wir den aus der Schule bekannten Euklidischen Abstandsbegriff. Wenn die Metrik, wie in (5.1.1) von einer Norm kommt, dann gilt neben den obigen Eigenschaften noch

$$d(\mathbf{x}, \mathbf{y}) = d(\mathbf{x} + \mathbf{z}, \mathbf{y} + \mathbf{z}) \quad (\text{Translationsinvarianz})$$

für alle $\mathbf{z} \in \mathbb{R}^n$.

Die verschiedenen Abstandsbegriffe auf $\mathbb{R}^n$ sind alle in einem gewissen Sinne vergleichbar. Wir führen dazu den Begriff der Äquivalenz ein.

Definition 5.3 (Normäquivalenz).

Es seien $\|\cdot\|$ und $\|\cdot\|'$ zwei Normen auf $\mathbb{R}^n$. Die beiden Normen heißen genau dann **äquivalent**, wenn Konstanten $c, C > 0$ existieren mit

$$\mathbf{x} \in \mathbb{R}^n: \quad c\|\mathbf{x}\| \leq \|\mathbf{x}\|' \leq C\|\mathbf{x}\|.$$

Beispiel 5.2.

Wir zeigen, daß $\|\cdot\|_\infty$ und $\|\cdot\|_2$ äquivalent sind. Als erstes gilt

$$\begin{aligned} \|\mathbf{x}\|_\infty &= |x_j| = \sqrt{|x_j|^2} \\ &\leq \sqrt{|x_1|^2 + \dots + |x_n|^2} = \|\mathbf{x}\|_2. \end{aligned}$$

Dabei haben wir die Monotonie der Wurzel benutzt und das j sei einfach ein Index für den das Maximum in der Definition der Supremumsnorm angenommen wird. Da offensichtlich für alle $i = 1, \dots, n$ gilt das $|x_i| \leq \|\mathbf{x}\|_\infty$ folgt

$$\begin{aligned} \|\mathbf{x}\|_2 &= \sqrt{|x_1|^2 + \dots + |x_n|^2} \\ &\leq \sqrt{\|\mathbf{x}\|_\infty + \dots + \|\mathbf{x}\|_\infty} \\ &= \sqrt{n} \|\mathbf{x}\|_\infty. \end{aligned}$$

Dabei haben wir wieder die Monotonie der Wurzel ausgenutzt. Damit haben wir also

$$\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \sqrt{n} \|\mathbf{x}\|_\infty.$$

Interpretieren Sie diese Ungleichung graphisch!

Es ist leicht zu sehen, daß Normäquivalenz eine Äquivalenzrelation ist, d.h.

- (i) Jede Norm ist zu sich selbst äquivalent. (Reflexivität)
- (ii) Wenn eine Norm zu einer anderen äquivalent ist, dann ist auch die zweite äquivalent zu ersten. (Symmetrie)
- (iii) Wenn eine Norm zu einer zweiten äquivalent ist und diese zu einer dritten, dann sind auch die erste und die dritte äquivalent. (Transitivität)

Folglich ist $\|\cdot\|_\infty$ äquivalent zu $\|\cdot\|_2$ und umgekehrt. Schreiben Sie die fehlenden Ungleichungen auf.

Es gilt

Satz 5.1.

Alle Normen auf $\mathbb{R}^n$ sind äquivalent.

Dieser Satz ist sehr nützlich denn er zieht nach sich, daß wir immer die Norm verwenden können, die wir am bequemsten finden. Verallgemeinern wir die Definition von Konvergenz von Folgen auf $\mathbb{R}^n$. Es ist klar, daß eine Folge in $\mathbb{R}^n$ eine Funktion $\mathbb{N} \rightarrow \mathbb{R}^n$ ist in direkter Verallgemeinerung zu Mathematik 1. Beispielsweise $(\mathbf{a}_n)_{n \in \mathbb{N}}$ mit

$$\mathbf{a}_n := \begin{bmatrix} 1 + \frac{1}{n} \\ 1 - \frac{1}{n} \end{bmatrix}.$$

Wir definieren Konvergenz mit

Definition 5.4 (Konvergenz von Folgen in $\mathbb{R}^n$).

Sei $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge. Dann heißt die Folge genau dann konvergent, wenn ein $\mathbf{a} \in \mathbb{R}^n$ existiert so daß für alle $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert mit

$$n \geq n_0: \quad \|\mathbf{a}_n - \mathbf{a}\|_2 < \varepsilon.$$

Nach Satz 5.1 ist es egal, welche Norm in Definition 5.4 verwendet wird da man, wenn die Folge in der einen Norm konvergent ist, auch Konvergenz in der anderen zeigen kann.

Beispielsweise: Sei $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge in $\mathbb{R}^n$ und sei diese konvergent nach Definition 5.4. Sei $\varepsilon > 0$ beliebig. Dann existiert $n_0 \in \mathbb{N}$ mit

$$n \geq n_0: \quad \|\mathbf{a}_n - \mathbf{a}\|_2 < \frac{\varepsilon}{\sqrt{n}}.$$

Damit haben wir

$$n \geq n_0: \quad \|\mathbf{a}_n - \mathbf{a}\|_\infty \leq \sqrt{n} \|\mathbf{a}_n - \mathbf{a}\|_2 < \varepsilon. \quad (5.1.2)$$

also konvergiert die Folge in der Supremumsnorm. In gleicher Weise kann man dies für jede andere äquivalente Norm tun. Aus (5.1.2) folgt insbesondere, daß die Komponenten-Folgen von $(\mathbf{a}_n)_{n \in \mathbb{N}}$ konvergent sind: Sei $i \in \{1, \dots, n\}$ und $a_{i,n}$ sei die i -te Komponente von $\mathbf{a}_n$ und a_i die i -te Komponente von $\mathbf{a}$. Dann

$$n \geq n_0: \quad |a_{i,n} - a_i| \leq \|\mathbf{a}_n - \mathbf{a}\|_\infty < \varepsilon.$$

Dies aber ist nach Mathematik 1 die Definition für $a_{i,n} \rightarrow a_i$. Damit haben wir im wesentlichen gezeigt, daß eine Folge in $\mathbb{R}^n$ genau dann nach Definition 5.1.2 konvergent ist, wenn die Komponentenfolgen konvergent sind. Vergleiche dazu auch das Hausaufgabenblatt auf dem wir gezeigt haben, daß Folgen komplexer Zahlen genau dann konvergieren, wenn die Real- und Imaginärteilefolgen konvergieren. (Die Umkehrung die ich hier nicht nachgerechnet habe sei der Leserin zur Übung überlassen. Diese braucht nur die andere Ungleichung aus Beispiel 5.2.)

5.2 Das (Standard)Skalarprodukt

Definition 5.5 (Skalarprodukt/Innenprodukt).

Eine Abbildung

$$\begin{cases} \langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n & \rightarrow \mathbb{R} \\ (\mathbf{x}, \mathbf{y}) & \mapsto \langle \mathbf{x}, \mathbf{y} \rangle \end{cases}$$

heißt genau dann **Skalarprodukt** (Innenprodukt) wenn für alle $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ gilt:

(i) $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$, $\langle \mathbf{x}, \mathbf{x} \rangle = 0$ genau dann, wenn $\mathbf{x} = \mathbf{0}$
(positive Definitheit)

(ii) $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$ (Symmetrie)

(iii) $\langle \lambda \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \lambda \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle$. (Linear in 1. Komponente)

Den Vektorraum $\mathbb{R}^n$ zusammen mit dem Skalarprodukt nennt man Skalarproduktraum/Innenproduktraum oder Hilbert-Raum.

Bemerkung 5.1.

Wenn man in Definition 5.5 überall $\mathbb{R}$ durch $\mathbb{C}$ ersetzt, dann muß man (ii) durch $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$ ersetzen. In (iii) ergibt sich dann für die zweite Komponente

$$\langle \mathbf{x}, \lambda \mathbf{y} + \mu \mathbf{z} \rangle = \bar{\lambda} \langle \mathbf{x}, \mathbf{y} \rangle + \bar{\mu} \langle \mathbf{x}, \mathbf{z} \rangle.$$

Diese Eigenschaft heißt anti-linear.

Definition 5.6 (Orthogonalität).

Sei $\mathbb{R}^n$ mit einem Skalarprodukt $\langle \cdot, \cdot \rangle$ ausgestattet. Dann heißen zwei $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ genau dann orthogonal ($\mathbf{x} \perp \mathbf{y}$) (bzgl. $\langle \cdot, \cdot \rangle$), wenn $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

Beispiel 5.3.

Das für den Moment wichtigste Beispiel eines Skalarproduktes auf $\mathbb{R}^n$ ist das sogenannte **Standard-Skalarprodukt**

$$\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{k=1}^n x_k \cdot y_k = \mathbf{x}^T \mathbf{y}.$$

Auf $\mathbb{C}^n$ ist das entsprechende Analogon

$$\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{k=1}^n x_k \cdot \overline{y_k} = \mathbf{y}^H \mathbf{x}.$$

Übungsaufgabe 5.2.

Prüfen Sie folgende Vektoren des $\mathbb{R}^3$ auf paarweise Orthogonalität bzgl. des Standard-Skalarproduktes:

$$\mathbf{a} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Das Standard-Skalarprodukt ist jedoch mitnichten das einzige Skalarprodukt auf $\mathbb{R}^n$.

Welches man wählt hängt von der von ihm zu erfüllenden Aufgabe ab und Sie werden später in weiterführenden Vorlesungen Beispiele für die Nützlichkeit anderer Skalarprodukte sehen. In unendlichdimensionalen Räumen werden wir mit der Untersuchung von Fourier-Reihen mit

$$\int_0^{2\pi} f(t)g(t) \, dt$$

ein Analogon zum Standard-Skalarprodukt auf $\mathbb{R}^n$ auf einem Funktionenraum, sagen wir $C[0, 2\pi]$ betrachten.

Beispiel 5.4.

Sei $\mathbf{A} = \text{diag}(\lambda_1, \dots, \lambda_n)$, $\lambda_i \in (0, \infty)$, $i = 1, \dots, n$. Dann ist

$$\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{x}^T (\mathbf{A} \mathbf{y}), \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$$

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

Hat man einmal ein Skalarprodukt festgelegt, so ist damit immer eine ‘Längenmessung’, also eine Norm verbunden:

Satz 5.2 (und Definition).

Sei der Vektorraum $\mathbb{R}^n$ mit dem Skalarprodukt $\langle \cdot, \cdot \rangle$ ausgestattet.

Dann gilt die sogenannte **Cauchy-Schwarzsche-Ungleichung**

$$|\langle \mathbf{x}, \mathbf{y} \rangle|^2 \leq \langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle, \quad (\mathbf{x}, \mathbf{y} \in \mathbb{R}^n) \quad (\text{CSU})$$

mit Gleichheit genau dann, wenn $\mathbf{x}$ und $\mathbf{y}$ linear abhängig sind.

Die durch

$$\|\mathbf{x}\| := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \quad (\mathbf{x} \in \mathbb{R}^n)$$

definierte Abbildung heißt die von diesem Skalarprodukt **erzeugte Norm**.

Mit dem Standard-Skalarprodukt (auch Euklidisches Skalarprodukt)

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{k=1}^n x_k y_k$$

folgt

$$\|x\|_2 = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{\sum_{k=1}^n (x_k)^2}.$$

Das Standard-Skalarprodukt erzeugt also die uns bekannte Euklid-Norm und wir hatten erkannt, daß die Euklid-Norm unserer anschaulichen Definition einer Länge eines Vektors im geometrischen Modell des $\mathbb{R}^2$ bzw. $\mathbb{R}^3$ mit Hilfe

KAPITEL 5. LÄNGEN-, ABSTANDS- UND WINKELMESSUNG IM $\mathbb{R}^n$ 103

des Satzes von Pythagoras entspricht.

Bitte beachten Sie, daß die ℓ^p -Normen $\|\cdot\|_p$ für kein $p \neq 2$ von einem Skalarprodukt erzeugt werden!

Wir wollen uns noch überlegen, daß die in Definition 5.6 eingeführte Definition von Orthogonalität unserer Anschauung entspricht. Wenn wir zwei Vektoren $\mathbf{x}$ und $\mathbf{y}$ in $\mathbb{R}^2$ gegeben haben, dann sind diese genau dann orthogonal, wenn $\|\mathbf{x} - \mathbf{y}\| = \|\mathbf{x} + \mathbf{y}\|$ gilt. Um Platz zu sparen vereinbaren wir die Schreibweise $\mathbf{x}^2 = \langle \mathbf{x}, \mathbf{x} \rangle$. Mit ein wenig Algebra erhalten wir dann :

$$\begin{aligned}\|\mathbf{x} - \mathbf{y}\| = \|\mathbf{x} + \mathbf{y}\| &\Leftrightarrow \|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x} + \mathbf{y}\|^2 \\ &\Leftrightarrow (\mathbf{x} - \mathbf{y})^2 = (\mathbf{x} + \mathbf{y})^2 \\ &\Leftrightarrow \mathbf{x}^2 - 2\langle \mathbf{x}, \mathbf{y} \rangle + \mathbf{y}^2 = \mathbf{x}^2 + 2\langle \mathbf{x}, \mathbf{y} \rangle + \mathbf{y}^2 \\ &\Leftrightarrow 4\langle \mathbf{x}, \mathbf{y} \rangle = 0 \\ &\Leftrightarrow \langle \mathbf{x}, \mathbf{y} \rangle = 0\end{aligned}$$

Damit haben wir die Einführung $\mathbf{x} \perp \mathbf{y} \Leftrightarrow \langle \mathbf{x}, \mathbf{y} \rangle = 0$ gerechtfertigt.

Nun wollen wir uns eine geeignete Definition von Winkeln zwischen Vektoren verschaffen. Aus der Cauchy-Schwarz'schen Ungleichung folgt für alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, daß

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

für die von $\langle \cdot, \cdot \rangle$ erzeugte Norm $\|\cdot\|$. Damit können wir $\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}$ für ein gegebenes Skalarprodukt als den Kosinus (Wieso nicht Sinus?) des Winkels zwischen den Vektoren $\mathbf{x}$ und $\mathbf{y}$ setzen.

Damit geben wir die folgende Definition

Definition 5.7 (Winkel).

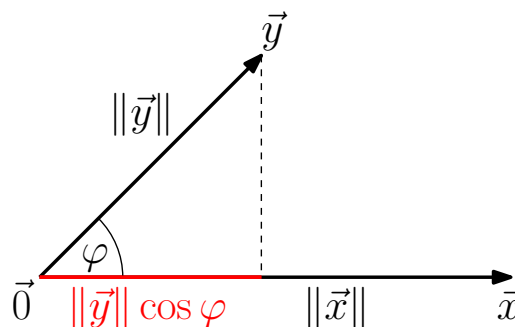
Sei $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf $\mathbb{R}^n$. Dann ist der Winkel $\angle(\mathbf{x}, \mathbf{y})$ zwischen zwei Vektoren $\mathbf{x}, \mathbf{y} \in n$ definiert durch

$$\cos(\angle(\mathbf{x}, \mathbf{y})) = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}.$$

Für das Euklidische Skalarprodukt erhalten wir dann aus der Definition die Formel

$$\mathbf{x}^T \mathbf{y} = \|\mathbf{x}\|_2 \|\mathbf{y}\|_2 \cos(\varphi), \quad \varphi = \angle(\mathbf{x}, \mathbf{y}).$$

Wir interpretieren diese Gleichung an der Skizze



und sehen, daß das Skalarprodukt gleich der (vorzeichenbehafteten) Länge der **orthogonalen Projektion** von $\mathbf{y}$ auf $\mathbf{x}$, multipliziert mit der Länge von $\mathbf{x}$ ist.

Mit dem nun zur Verfügung stehenden Begriff der Orthogonalität verfeinern wir nun den Begriff der Basis. Dazu geben wir

Definition 5.8 (Orthonormalbasis).

Sei $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf $\mathbb{R}^n$. Eine Basis $B = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$ des $\mathbb{R}^n$ heißt **Orthonormalbasis** (ONB), wenn

$$\langle \mathbf{b}_i, \mathbf{b}_j \rangle = \delta_{ij} \quad i, j = 1, \dots, n.$$

Wenn das Skalarprodukt das Standardskalarprodukt ist, dann gilt $\mathbf{b}_j^T \mathbf{b}_i = \delta_{ij}$, $i, j = 1, \dots, n$.

Bei einer Orthonormalbasis stehen die Basisvektoren also paarweise aufeinander senkrecht und haben allesamt die Länge 1.

Beispiel 5.5.

Ein prominentes Beispiel für eine Orthonormalbasis des $\mathbb{R}^n$ ist die Standardbasis $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$.

Definition 5.9 (Orthogonale Matrix).

Sei $\mathbf{A} \in \mathbb{R}^{n \times n}$.

Die Matrix $\mathbf{A}$ heißt genau dann orthogonal, wenn $\mathbf{A}\mathbf{A}^T = \mathbf{I}_n$ gilt.

Die Leserin bemerkt sofort, daß die Inverse einer orthogonalen Matrix $\mathbf{A}$ (per Definition) durch die Transponierte $\mathbf{A}^T$ der Matrix gegeben ist.

Bemerkung 5.2.

An dieser Stelle sollte klar sein, daß eine Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ genau dann orthogonal ist, wenn die Spalten von $\mathbf{A}$ eine ONB von $\mathbb{R}^n$ bilden.

5.3 Gram-Schmidtsches Orthogonalisierungsverfahren

Aus einem Satz Vektoren $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ kann man durch folgendes Verfahren eine Orthonormalbasis des Unterraumes $\text{span}(\{\mathbf{a}_1, \dots, \mathbf{a}_n\})$ konstruieren.

$$\mathbf{b}_1 = \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|},$$

$$\tilde{\mathbf{b}}_2 = \mathbf{a}_2 - \langle \mathbf{a}_2, \mathbf{b}_1 \rangle \mathbf{b}_1,$$

$$\mathbf{b}_2 = \frac{\tilde{\mathbf{b}}_2}{\|\tilde{\mathbf{b}}_2\|},$$

$$\tilde{\mathbf{b}}_3 = \mathbf{a}_3 - \langle \mathbf{a}_3, \mathbf{b}_2 \rangle \mathbf{b}_2 - \langle \mathbf{a}_3, \mathbf{b}_1 \rangle \mathbf{b}_1,$$

$$\mathbf{b}_3 = \frac{\tilde{\mathbf{b}}_3}{\|\tilde{\mathbf{b}}_3\|},$$

$\vdots$

$$\tilde{\mathbf{b}}_n = \mathbf{a}_n - \sum_{k=1}^{n-1} \langle \mathbf{a}_n, \mathbf{b}_k \rangle \mathbf{b}_k,$$

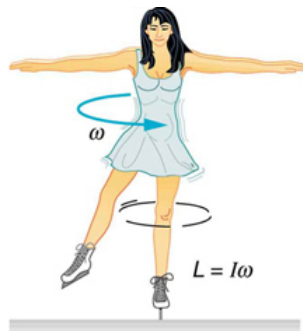
$$\mathbf{b}_n = \frac{\tilde{\mathbf{b}}_n}{\|\tilde{\mathbf{b}}_n\|}.$$

Falls die Vektoren $\mathbf{a}_1, \dots, \mathbf{a}_n$ nur einen Unterraum der Dimension $m < n$ aufspannen, gilt $\mathbf{b}_{m+1} = \dots = \mathbf{b}_n = 0$.

5.4 Das Vektorprodukt in $\mathbb{R}^3$

Bei vielen physikalischen Anwendungen benötigt man eine weitere Operation, das sogenannte **Vektor- oder Kreuzprodukt**. Beispiele sind

- das **Drehmoment** $\vec{M} = \vec{r} \times \vec{F}$ zum Abstandsvektor $\vec{r}$ und zur Kraft $\vec{F}$;



Verwandt: **Drehimpuls** $\vec{L} = \vec{r} \times \vec{p}$ bei Impuls $m\vec{v}$. Bei Drehbewegung mit Winkelgeschwindigkeit ω ist $|\vec{L}| = mrv = mr^2\omega$.

- die **Lorentzkraft** $\vec{F}_L = q(\vec{v} \times \vec{B})$, welche auf ein Teilchen mit Ladung q und Geschwindigkeit $\vec{v}$ im Magnetfeld $\vec{B}$ wirkt;



Abbildung 5.1: Fadenstrahlrohr

- die **Corioliskraft** $\vec{F}_C = 2m(\vec{v} \times \vec{\omega})$, welche auf einen Körper der Masse m wirkt, welcher sich mit Geschwindigkeit $\vec{v}$ relativ zu einem mit Winkelgeschwindigkeit $\vec{\omega}$ rotierenden Bezugssystem bewegt;



Wir gehen bei der Definition des Kreuzprodukts von den gewünschten geometrischen Eigenschaften aus und leiten dann die Rechenregeln sowie die Formel zur Berechnung her.

Sämtliche Vektoren in diesem Kapitel sind als Elemente des $\mathbb{R}^3$ aufzufassen.

Als erstes bringen wir den Begriff eines Rechtssystems.

Definition 5.10 (Rechtssystem).

Drei linear unabhängige Vektoren $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ bilden (in dieser Reihenfolge) ein **Rechtssystem**, wenn sich – von der Spitze von $\mathbf{c}$ aus gesehen – $\mathbf{a}$ durch Drehung um einen Winkel $\phi \in [0, \pi)$ im Gegenuhrzeigersinn in die gleiche Richtung wie $\mathbf{b}$ bringen läßt.

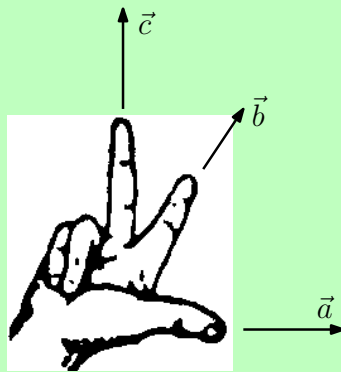


Abbildung 5.2: Quelle: Qniemiec, [Wikimedia Commons](#)

Ob ein Rechtssystem vorliegt, kann mit der Rechten-Hand-Regel (Korkenzieherregel) entschieden werden. Analog kann man ein **Linkssystem** definieren.

Nun führen wir axiomatisch ein, was wir unter dem Kreuzprodukt zweier Vektoren verstehen wollen.

Definition 5.11 (Kreuzprodukt).

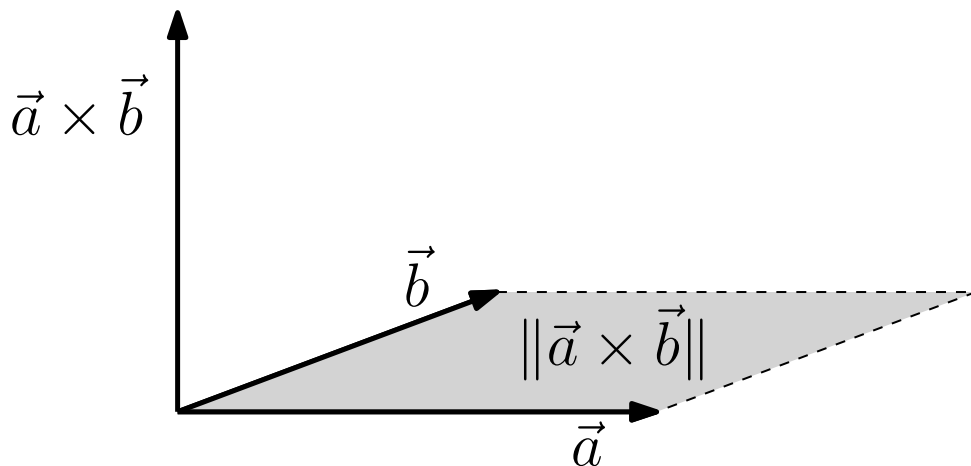
Seien $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ linear unabhängig. Der Vektor

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} \in \mathbb{R}^3$$

heißt **Kreuzprodukt** (auch **Vektorprodukt** oder **äußeres Produkt**) von $\mathbf{a}$ und $\mathbf{b}$, wenn

- (i) $\mathbf{c} \perp \mathbf{a}$ und $\mathbf{c} \perp \mathbf{b}$ (d.h. $\mathbf{c} \perp \text{span}\{\mathbf{a}, \mathbf{b}\}$),
- (ii) $\|\mathbf{c}\|_2 = \|\mathbf{a}\|_2 \|\mathbf{b}\|_2 |\sin(\angle(\mathbf{a}, \mathbf{b}))|$,
- (iii) $\mathbf{a}, \mathbf{b}, \mathbf{c}$ bilden in dieser Reihenfolge ein Rechtssystem.

Sind $\mathbf{a}$ und $\mathbf{b}$ linear abhängig (was die Fälle $\mathbf{a} = \mathbf{0}$ oder $\mathbf{b} = \mathbf{0}$ mit einschließt), so definieren wir $\mathbf{a} \times \mathbf{b} := \mathbf{0}$.



Bemerkung 5.3.

Ein paar Bemerkungen zur Definition des Kreuzproduktes:

- (i) legt die Gerade fest, auf der $\mathbf{a} \times \mathbf{b}$ liegt. Die Richtung von $\mathbf{a} \times \mathbf{b}$ ist damit bis auf das Vorzeichen eindeutig bestimmt.*
- (ii) legt die Länge von $\mathbf{a} \times \mathbf{b}$ fest. Diese stimmt mit der Fläche des von $\mathbf{a}$ und $\mathbf{b}$ aufgespannten Parallelogramms überein.*
- (iii) liefert die Entscheidung, welcher der verbliebenen zwei Vektoren (Richtungen) zu verwenden ist.*

Allein aufgrund der geometrischen Definition ergeben sich folgende Rechenregeln für Kreuzprodukte:

Satz 5.3.

Seien $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ und $\lambda \in \mathbb{R}$.

Dann gelten:

- (i) $\mathbf{b} \times \mathbf{a} = -(\mathbf{a} \times \mathbf{b})$,*
- (ii) $(\mathbf{a} + \mathbf{b}) \times \mathbf{c} = \mathbf{a} \times \mathbf{c} + \mathbf{b} \times \mathbf{c}$,*
- (iii) $\lambda(\mathbf{a} \times \mathbf{b}) = (\lambda\mathbf{a}) \times \mathbf{b} = \mathbf{a} \times (\lambda\mathbf{b})$.*

Übungsaufgabe 5.3.

Illustrieren Sie zumindest den ersten und dritten Punkt in Satz 5.3 anhand geeigneter Skizzen.

Satz 5.4.

Für $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ gilt

$$\mathbf{a} \times \mathbf{b} = \begin{bmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{bmatrix}. \quad (5.4.1)$$

Beweis. Wir geben hier nur die Idee an. Die Leserin ist eingeladen aus den bisher gezeigten Eigenschaften die Rechenvorschrift mit den genannten Ideen abzuleiten.

- Für die Einheitsvektoren des $\mathbb{R}^3$ gilt offenbar $\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3$, $\mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1$ und $\mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2$. (Hat nichts mit der Rechenvorschrift zu tun!)
- Schreibe $\mathbf{a} = a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + a_3 \mathbf{e}_3$ bzw. $\mathbf{b} = b_1 \mathbf{e}_1 + b_2 \mathbf{e}_2 + b_3 \mathbf{e}_3$ und wende Satz 5.3 an.

□

Übungsaufgabe 5.4.

Man führe die Rechnung zum zweitgenannten Punkt im obigen Beweis aus.

Bemerkung 5.4.

Formel (5.4.1) merkt man sich am besten mit Hilfe der formalen 3×3 -Determinanten

$$\mathbf{a} \times \mathbf{b} = \det \left(\begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} \right),$$

die man mit der Regel von Sarrus auswertet.

Übungsaufgabe 5.5.

Berechnen Sie auf diese (Bemerkung 5.4) Weise das Kreuzprodukt

$$\begin{bmatrix} 1 \\ 2 \\ -4 \end{bmatrix} \times \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}.$$

5.5 Das Spatprodukt in $\mathbb{R}^3$

Mitunter ist noch eine Kombination von Kreuz- und Skalarprodukt in Gebrauch. Unter dem **Spatprodukt** (auch **gemischten Produkt**) dreier Vektoren $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ versteht man die reelle Zahl

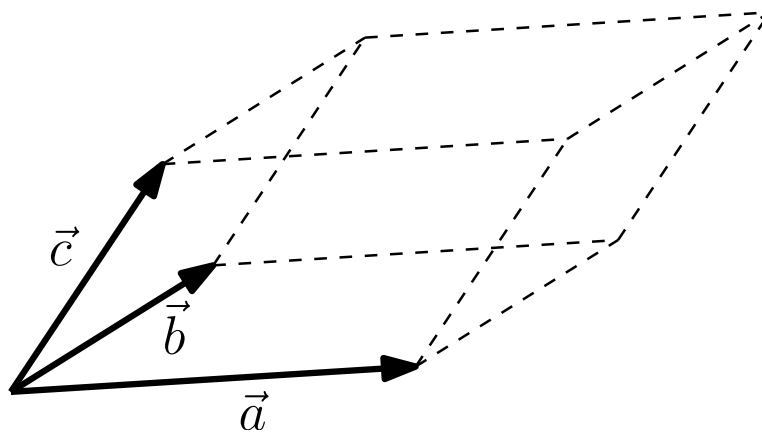
$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] := \mathbf{a}^T(\mathbf{b} \times \mathbf{c}).$$

Eine einfache Rechnung (ausführen!) zeigt, daß sich das Spatprodukt als gewöhnliche Determinante interpretieren läßt:

$$[\mathbf{a}, \mathbf{b}, \mathbf{c}] = \det \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}.$$

Beim Spatprodukt handelt es sich also eigentlich um nichts Neues.

Wie im Kapitel 4 zu Determinanten erörtert, läßt sich damit der Betrag des Spatprodukts als Volumen des von $\mathbf{a}, \mathbf{b}$ und $\mathbf{c}$ aufgespannten Parallelepipeds interpretieren.



Der Begriff **Spat** steht synonym für **Parallelepip**; dies begründet die Namensgebung.

Satz 5.5.

Seien $\mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3$. Dann gelten:

- (i) Sind $\mathbf{a}, \mathbf{b}, \mathbf{c}$ linear abhängig, so ist $[\mathbf{a}, \mathbf{b}, \mathbf{c}] = 0$.
- (ii) Sind $\mathbf{a}, \mathbf{b}, \mathbf{c}$ linear unabhängig, und bilden sie in dieser Reihenfolge ein Rechtssystem, so ist $[\mathbf{a}, \mathbf{b}, \mathbf{c}] > 0$.
- (iii) Sind $\mathbf{a}, \mathbf{b}, \mathbf{c}$ linear unabhängig, und bilden sie in dieser Reihenfolge ein Linkssystem, so ist $[\mathbf{a}, \mathbf{b}, \mathbf{c}] < 0$.
- (iv) $[\mathbf{b}, \mathbf{a}, \mathbf{c}] = [\mathbf{c}, \mathbf{b}, \mathbf{a}] = [\mathbf{a}, \mathbf{c}, \mathbf{b}] = -[\mathbf{a}, \mathbf{b}, \mathbf{c}]$.
- (v) $[\mathbf{a}, \mathbf{b}, \mathbf{c}] = [\mathbf{c}, \mathbf{a}, \mathbf{b}] = [\mathbf{b}, \mathbf{c}, \mathbf{a}]$.

5.6 Elemente der analytischen Geometrie

In diesem Abschnitt betrachten wir nur noch Objekte im $\mathbb{R}^3$. Einige Ergebnisse haben ihre Entsprechungen im $\mathbb{R}^2$, die klar ersichtlich sind.

Konkret wollen wir uns mit den Lagebeziehungen von **Punkt**, **Gerade** und **Ebene** befassen.

Der Stoff gehört zum Standardrepertoire an Gymnasien und wird dort auch in der nötigen Tiefe behandelt. Daher sollen diese Inhalte nur überblicksartig dargestellt werden.

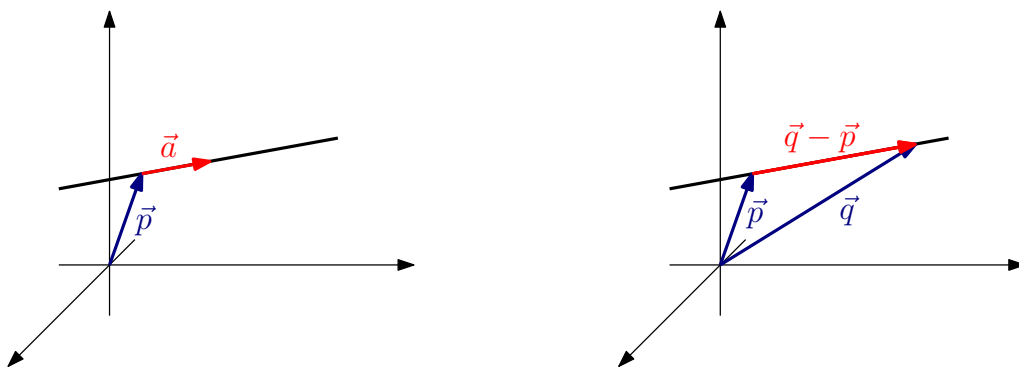
Eine Gerade g , die durch den Punkt $\mathbf{p}$ und parallel zur Richtung $\mathbf{a} \neq \mathbf{0}$ verläuft, besitzt die Parameterdarstellung

$$g: \quad \mathbf{x} = \mathbf{p} + \lambda \mathbf{a}, \quad \lambda \in \mathbb{R}.$$

Diese Form wird auch **Punkt-Richtungs-Form** genannt.

Alternativ ist eine Gerade durch zwei verschiedene Punkte $\mathbf{p}$ und $\mathbf{q}$ eindeutig festgelegt. Man erhält die Parameterdarstellung über die **Zwei-Punkte-Form**

$$g: \quad \mathbf{x} = \mathbf{p} + \lambda (\mathbf{q} - \mathbf{p}), \quad \lambda \in \mathbb{R}.$$



Analog ist eine Ebene E durch einen Punkt $\mathbf{p}$ und zwei linear unabhängige Richtungsvektoren $\mathbf{a}$ und $\mathbf{b}$ eindeutig festgelegt. Sie besitzt also die Parame-

terform

$$E: \quad \mathbf{x} = \mathbf{p} + \lambda \mathbf{a} + \mu \mathbf{b}, \quad \lambda, \mu \in \mathbb{R}.$$

Diese Form wird auch **Punkt-Richtungs-Form der Ebene** genannt.

Alternativ kann man zur eindeutigen Festlegung drei Punkte $\mathbf{p}$, $\mathbf{q}$ und $\mathbf{r}$ verwenden. Man erhält die Parameterdarstellung über die **Drei-Punkte-Form**

$$E: \quad \mathbf{x} = \mathbf{p} + \lambda (\mathbf{q} - \mathbf{p}) + \mu (\mathbf{r} - \mathbf{p}), \quad \lambda, \mu \in \mathbb{R}.$$

Jeder Vektor $\mathbf{n} \neq 0$, der senkrecht auf der Ebene steht, heißt **Normalenvektor** der Ebene.

Auch durch Vorgabe eines Punktes $\mathbf{p}$ und eines Normalenvektors $\mathbf{n}$ ist eine Ebene eindeutig festgelegt. Es entsteht die **Normalenform**

$$E: \mathbf{n}^T(\mathbf{x} - \mathbf{p}) = 0 \quad \text{bzw.} \quad (5.6.1)$$

$$E: n_1 x_1 + n_2 x_2 + n_3 x_3 = c \quad (\text{mit } c = \mathbf{n}^T \mathbf{p}).$$

Ist $\mathbf{n}$ auf Länge Eins normiert, so ist der Abstand des Punktes $\mathbf{x} = (x_1, x_2, x_3)$ von der Ebene gegeben durch (**Hessesche Normalform**¹)

$$d = |n_1 x_1 + n_2 x_2 + n_3 x_3 - c| = \min_{\mathbf{p} \in E} \|\mathbf{x} - \mathbf{p}\|_2.$$

Übungsaufgabe 5.6.

Wie kann man einen Normalenvektor zur Ebene $\mathbf{x} = \mathbf{p} + \lambda \mathbf{a} + \mu \mathbf{b}$ berechnen?

¹Benannt nach Otto Hesse, 1811–1874, Deutscher Mathematiker

Übungsaufgabe 5.7.

Interpretieren Sie Formel (5.6.1) im Kontext orthogonaler Projektionen auf den Normalenvektor.

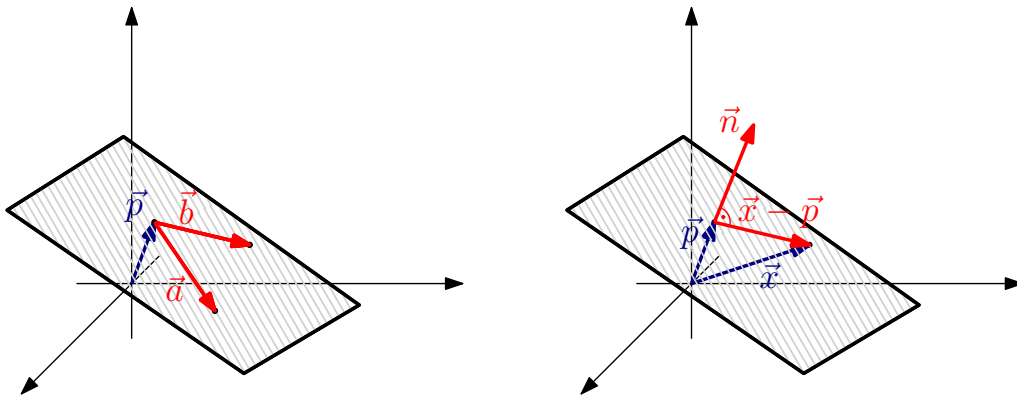


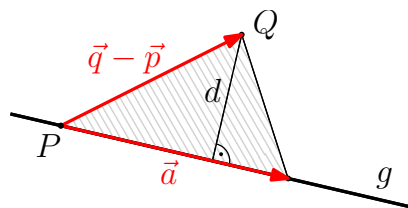
Abbildung 5.3: Illustration von Punkt-Richtungsform (links) und Normalenform (rechts).

Ob ein Punkt $\mathbf{q}$ auf einer Geraden $g : \mathbf{x} = \mathbf{p} + \lambda \mathbf{a}$ liegt oder nicht, kann man direkt mit der Parameterdarstellung prüfen. Alternativ berechnet man einfach den Abstand

$$\begin{aligned} d(\mathbf{q}, g) &= \frac{\|\mathbf{a} \times (\mathbf{q} - \mathbf{p})\|}{\|\mathbf{a}\|} \\ &= \min_{\mathbf{x} \in g} \|\mathbf{q} - \mathbf{x}\|_2. \end{aligned} \quad (5.6.2)$$

Übungsaufgabe 5.8.

Bestätigen Sie die Formel. Berechnen Sie dazu den Flächeninhalt des von $\mathbf{a}$ und $\mathbf{q} - \mathbf{p}$ aufgespannten Dreiecks (siehe Figur unten) auf zwei verschiedene Weisen.



Für zwei Geraden

$$g_1 : \mathbf{x} = \mathbf{p} + \lambda \mathbf{a} \quad \text{und} \quad g_2 : \mathbf{x} = \mathbf{q} + \mu \mathbf{b}$$

im $\mathbb{R}^3$ liegt genau eine der drei folgenden Situationen vor:

- sie sind **parallel** (Spezialfall der Gleichheit einbezogen),
- sie schneiden sich in genau einem Punkt,
- sie sind **windschief**, d.h. sie sind weder parallel noch schneiden sie sich.

Parallelität läßt sich am leichtesten erkennen: In diesem Fall sind die beiden Vektoren $\mathbf{a}$ und $\mathbf{b}$ linear abhängig, unterscheiden sich also nur um einen skalaren Faktor.

Im Falle der Parallelität kann man Formel (5.6.2) zur Abstandsbestimmung nutzen (warum?) und erhält

$$d(g_1, g_2) = \frac{\|\mathbf{a} \times (\mathbf{q} - \mathbf{p})\|}{\|\mathbf{a}\|} = \frac{\|\mathbf{b} \times (\mathbf{p} - \mathbf{q})\|}{\|\mathbf{b}\|}.$$

Für die restlichen Fälle erzeugt man durch Gleichsetzen das lineare Gleichungssystem

$$\mathbf{p} + \lambda \mathbf{a} = \mathbf{q} + \mu \mathbf{b} \tag{5.6.3}$$

(drei Gleichungen für die zwei Unbekannten λ und μ).

Hat (5.6.3) genau eine Lösung (λ^*, μ^*) , so schneiden sich die Geraden im Punkt $\mathbf{p} + \lambda^* \mathbf{a}$ (identisch mit $\mathbf{q} + \mu^* \mathbf{b}$).

Der Schnittwinkel ist gegeben durch

$$\cos(\varphi) = \frac{\mathbf{a}^T \mathbf{b}}{\|\mathbf{a}\| \|\mathbf{b}\|}.$$

Hat (5.6.3) unendlich viele Lösungen, so sind die Geraden g_1 und g_2 identisch. (Das kann man aber im Rahmen des Tests auf Parallelität bereits entscheiden.)

Hat (5.6.3) keine Lösung, so sind die Geraden g_1 und g_2 windschief oder aber parallel und verschieden (s. o.).

Liegt der windschiefe Fall vor, erhält man den Abstand der Geraden mittels

$$d(g_1, g_2) = \frac{|[\mathbf{a}, \mathbf{b}, \mathbf{q} - \mathbf{p}]|}{\|\mathbf{a} \times \mathbf{b}\|}. \tag{5.6.4}$$

Übungsaufgabe 5.9.

Formel (5.6.4) lässt sich in Analogie zu (5.6.2) bestätigen, indem man das Volumen eines geeigneten Körpers auf zwei verschiedene Weisen berechnet. Um welchen Körper handelt es sich?

Ob ein Punkt $\mathbf{q}$ zur Ebene

$$E : \mathbf{n}^T(\mathbf{x} - \mathbf{p}) = 0$$

gehört, läßt sich durch Einsetzen prüfen. (Gilt $\mathbf{n}^T(\mathbf{q} - \mathbf{p}) = 0$?)

Allgemein berechnet sich der Abstand von $\mathbf{q}$ zu E gemäß

$$d(\mathbf{q}, E) = \frac{|\mathbf{n}^T(\mathbf{q} - \mathbf{p})|}{\|\mathbf{n}\|}. \quad (5.6.5)$$

Übungsaufgabe 5.10.

Welcher Ansatz liegt Formel (5.6.5) zugrunde?

KAPITEL 5. LÄNGEN-, ABSTANDS- UND WINKELMESSUNG IM $\mathbb{R}^n$

Eine Gerade und eine Ebene sind im $\mathbb{R}^3$ entweder

- (i) parallel (beinhaltet den Fall, daß die Gerade in der Ebene liegt),
- (ii) oder sie schneiden sich in genau einem Punkt.

Die Gerade $g : \mathbf{x} = \mathbf{q} + \lambda \mathbf{a}$ und die Ebene $E : \mathbf{n}^T(\mathbf{x} - \mathbf{p}) = 0$ sind genau dann parallel, wenn $\mathbf{a}$ und $\mathbf{n}$ orthogonal sind, d. h. wenn

$$\mathbf{n}^T \mathbf{a} = 0.$$

In diesem Fall ist der Abstand Gerade-Ebene nach (5.6.5) gegeben durch

$$\begin{aligned} d &= \frac{|\mathbf{n}^T(\mathbf{q} - \mathbf{p})|}{\|\mathbf{n}\|} \\ &= \min_{\substack{\mathbf{x} \in g \\ \mathbf{y} \in E}} \|\mathbf{x} - \mathbf{y}\|_2. \end{aligned}$$

Sind die Ebene und die Gerade nicht parallel ($\mathbf{n}^T \mathbf{a} \neq 0$), dann ist

$$\mathbf{q} + \frac{\mathbf{n}^T(\mathbf{p} - \mathbf{q})}{\mathbf{n}^T \mathbf{a}} \mathbf{a}$$

ihr Schnittpunkt. Der Schnittwinkel φ erfüllt

$$\sin(\varphi) = \frac{|\mathbf{n}^T \mathbf{a}|}{\|\mathbf{n}\| \cdot \|\mathbf{a}\|}.$$

Übungsaufgabe 5.11.

Leiten Sie beide Formeln her.

KAPITEL 5. LÄNGEN-, ABSTANDS- UND WINKELMESSUNG IM $\mathbb{R}^n$

Zwei Ebenen im $\mathbb{R}^3$ sind

- (i) entweder parallel (Spezialfall: identisch)
- (ii) oder schneiden sich entlang einer Geraden.

Zwei Ebenen

$$E_1 : \mathbf{n}_1^T(\mathbf{x} - \mathbf{p}_1) = 0 \quad \text{und} \quad E_2 : \mathbf{n}_2^T(\mathbf{x} - \mathbf{p}_2) = 0,$$

sind genau dann parallel, wenn $\mathbf{n}_1$ und $\mathbf{n}_2$ linear abhängig (also bis auf Vielfache gleich) sind. In diesem Fall ist der Abstand der Ebenen gegeben durch

$$\begin{aligned} d(E_1, E_2) &= \frac{|\mathbf{n}_1^T(\mathbf{p}_2 - \mathbf{p}_1)|}{\|\mathbf{n}_1\|_2} \\ &= \frac{|\mathbf{n}_2^T(\mathbf{p}_1 - \mathbf{p}_2)|}{\|\mathbf{n}_2\|_2} \\ &= \min_{\substack{\mathbf{x} \in E_1 \\ \mathbf{y} \in E_2}} \|\mathbf{x} - \mathbf{y}\|_2 \end{aligned}$$

Schneiden sich die Ebenen entlang einer Geraden $\mathbf{x} = \mathbf{p} + \lambda \mathbf{a}$ (äquivalent zu $\mathbf{n}_1 \times \mathbf{n}_2 \neq \mathbf{0}$), dann ist

$$\mathbf{a} = \mathbf{n}_1 \times \mathbf{n}_2,$$

und $\mathbf{p}$ ist (jede) Lösung des linearen Gleichungssystems

$$\begin{aligned} \mathbf{n}_1^T(\mathbf{p} - \mathbf{p}_1) &= 0, \\ \mathbf{n}_2^T(\mathbf{p} - \mathbf{p}_2) &= 0 \end{aligned}$$

(zwei Gleichungen für drei Unbekannte, nämlich die drei Komponenten von $\mathbf{p}$). Der Schnittwinkel φ erfüllt

$$\cos(\varphi) = \frac{\mathbf{n}_1^T \mathbf{n}_2}{\|\mathbf{n}_1\| \cdot \|\mathbf{n}_2\|}.$$

5.7 Matrix-Multiplikation

Wenn wir uns an das Standard-**Skalarprodukt** aus [Sektion 5.2](#) erinnern, dann können wir sehen, daß,

$$\langle \mathbf{A}[i, \cdot], \mathbf{x} \rangle = \sum_{j=1}^n a_{i,j} x_j \quad (5.7.1)$$

wenn $\mathbf{A}[i, \cdot]$ die i -te Zeile der Matrix $\mathbf{A}$ bezeichnet. Also

$$\mathbf{Ax} = \begin{bmatrix} \langle \mathbf{A}^T[\cdot, 1], \mathbf{x} \rangle \\ \vdots \\ \langle \mathbf{A}^T[\cdot, m], \mathbf{x} \rangle \end{bmatrix} = \begin{bmatrix} (\mathbf{e}_1^T \mathbf{A}) \mathbf{x} \\ \vdots \\ (\mathbf{e}_m^T \mathbf{A}) \mathbf{x} \end{bmatrix}$$

und damit

$$\mathbf{A} \cdot \mathbf{B} = \begin{bmatrix} (\mathbf{e}_1^T \mathbf{A})(\mathbf{B}\mathbf{e}_1) & \cdots & (\mathbf{e}_1^T \mathbf{A})(\mathbf{B}\mathbf{e}_n) \\ (\mathbf{e}_2^T \mathbf{A})(\mathbf{B}\mathbf{e}_1) & \cdots & (\mathbf{e}_2^T \mathbf{A})(\mathbf{B}\mathbf{e}_n) \\ \vdots & & \vdots \\ (\mathbf{e}_m^T \mathbf{A})(\mathbf{B}\mathbf{e}_1) & \cdots & (\mathbf{e}_m^T \mathbf{A})(\mathbf{B}\mathbf{e}_n) \end{bmatrix}$$

5.8 Verständnisfragen

1. Welche Beziehung besteht zwischen der (definierten) Orthogonalität von Vektoren des $\mathbb{R}^n$ und unserer Anschauung?
2. Was bedeutet $\angle(\mathbf{x}, \mathbf{y})$ im geometrischen Modell des $\mathbb{R}^2$ bzw. $\mathbb{R}^3$.
3. Welche Orthogonalitätsbeziehungen gelten zwischen den Standardbasis-Vektoren $\mathbf{e}_i$, $i = 1, \dots, n$ von $\mathbb{R}^n$?
4. Worauf beruht die Berechnung von Winkeln im $\mathbb{R}^2$ bzw. $\mathbb{R}^3$? Welcher Zusammenhang besteht im $\mathbb{R}^3$ (relevant?) zwischen dem Einheitsvektor $\frac{\mathbf{x}}{\|\mathbf{x}\|_2}$ und den Winkeln zwischen $\mathbf{x}$ und den Koordinatenachsen?
5. Welche Eigenschaft besitzt der Vektor $\mathbf{x} \times \mathbf{y}$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ im geometrischen Modell des $\mathbb{R}^3$.
6. Wie lauten die Ergebnisse der zwischen den Vektoren der Standardbasis des $\mathbb{R}^3$ möglichen Vektorprodukte?
7. Welche geometrische Interpretation folgt für $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3 \setminus \{\mathbf{0}\}$ aus den Gleichungen
 - a) $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ und
 - b) $\mathbf{x} \times \mathbf{y} = \mathbf{0}$.
8. Welche Aussage gilt für alle Produkte $\langle \mathbf{x} \times \mathbf{y}, \mathbf{x} \rangle$ und $\langle \mathbf{x} \times \mathbf{y}, \mathbf{y} \rangle$ mit $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$?
9. Wie kann man das Ergebnis des Spatproduktes geometrisch interpretieren, und was bedeutet es, wenn das Spatprodukt gleich 0 ist?
10. Welche Auswirkung hat die Vertauschung der Reihenfolge der Vektoren auf den Wert des Spatproduktes?
11. Welche besondere Matrix ergibt sich, wenn man eine quadratische Matrix $\mathbf{A}$ mit ihrer Transponierten unter folgenden Bedingungen multipliziert?
 - a) Jeder Zeilenvektor von $\mathbf{A}$ sei ein Einheitsvektor.

b) Alle Zeilenvektoren von $\mathbf{A}$ sind paarweise orthogonal zueinander.

12. Darf man für alle $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ von $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle$ auf $\mathbf{y} = \mathbf{z}$ schließen?

13. Seien E_1 und E_2 zwei Ebenen in $\mathbb{R}^3$ mit Normalenvektoren $\mathbf{n}_i$. Bilde die Matrix $\begin{bmatrix} \mathbf{n}_1^T \\ \mathbf{n}_2^T \end{bmatrix}$. Was sagt der Rang der Matrix über die Lagebeziehung von E_1 und E_2 aus bzw. umgekehrt.

5.9 Rechenaufgaben

1. Berechnen Sie $\lambda \in \mathbb{R}$ derart, daß der Abstand der Punkte $A = (2, 1, -\lambda)$ und $B = (4, -3, 2)$ gleich $\sqrt{29}$ ist.
2. Es seien $A = (2, 0, 5)$, $B = (2, 4, 5)$, $C = (0, 4, 9)$ die Eckpunkte eines Dreiecks im geometrischen Modell des $\mathbb{R}^3$. Berechnen Sie die Länge der Seitenhalbierenden durch B sowie den Winkel $\alpha = \angle(\overrightarrow{AC}, \overrightarrow{AB})$.
3. Sei $\mathbf{a} \in \mathbb{R}^3$. Bestimmen Sie die Länge und den Richtungskosinus von $\mathbf{a}$, wenn
 - a) $\mathbf{a} = \mathbf{e}_1 - 4\mathbf{e}_2 + 8\mathbf{e}_3$,
 - b) $\mathbf{a}$ ist durch die Koordinaten $a_1 = 4$, $a_2 = 7$ und $a_3 = -4$ im $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ -System gegeben.
4. Von einem Vektor $\mathbf{a} \in \mathbb{R}^3$ seien $a_1 = 2$, $a_3 = 3$ und $\|\mathbf{a}\|_2 = 7$ gegeben. Bestimmen Sie alle Vektoren die diese Vorgaben erfüllen.
Welche Winkel schließen die Vektoren mit den Koordinatenachsen ein?
5. Für alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ ist der Vektor $\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle} \mathbf{y}$ als (senkrechte) Projektion von $\mathbf{x}$ auf $\mathbf{y}$ erklärt. Zeigen Sie, daß diese Festlegung für das Geometrische Modell von $\mathbb{R}^3$ mit dem dort anschaulichen Begriff der Projektion übereinstimmt!
6. Berechnen Sie die senkrechte Projektion des Vektors $\mathbf{a}$ auf den Vektor $\mathbf{b}$ für
 - a) $\mathbf{a} = -3\mathbf{e}_1 + 2\mathbf{e}_2$, $\mathbf{b} = 3\mathbf{e}_1 + \mathbf{e}_2$,
 - b) $\mathbf{a} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$,
 - c) $\mathbf{a} = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 5 \\ -1 \\ 1 \end{bmatrix}$.
7. Geben Sie den Einheitsvektor in Richtung der Winkelhalbierenden von
 - a) $\mathbf{a} = [2, 0]^T$ und $\mathbf{b} = [0, 3]^T$,

KAPITEL 5. LÄNGEN-, ABSTANDS- UND WINKELMESSUNG IM $\mathbb{R}^n$ 127

b) $\mathbf{a} = [-1, 2, 2]^T$ und $\mathbf{b} = [2, -6, -3]^T$

an!

8. Zeigen Sie für das geometrische Modell des $\mathbb{R}^2$ mit Hilfe der Vektorrechnung, daß jeder Peripheriewinkel über dem Durchmesser eines Halbkreises ein rechter Winkel ist. (Satz von Thales)
9. Die Punkte $P_1 = (1, 0, 0)$, $P_2 = (4, 3, 0)$, $P_3 = (0, 5, 2)$ und $P_4 = (0, 0, 5)$ seien die Eckpunkte eines Tetraeders. Berechnen sie Volumen und Oberfläche dieses Körpers.
10. Wie ist die Dritte Koordinate des Punktes P_4 in Aufgabe 8 abzuändern, damit $P_1, \dots, P'_4$ in einer Ebene liegen?
11. Berechnen Sie $\lambda \in \mathbb{R}$ derart, daß die Punkte $P_1 = (2, 2, 1)$, $P_2 = (-1, 4, -1)$ und $P_3 = (6, 5, -\lambda, 4\lambda)$ auf einer gemeinsamen Geraden liegen!
12. Es seien $\mathbf{a} = 2\mathbf{x} - 3\mathbf{y}$ und $\mathbf{b} = \mathbf{x} + \lambda\mathbf{y}$ mit

$$\mathbf{x} = \begin{bmatrix} -1 \\ 0 \\ 2 \\ 2 \end{bmatrix} \quad \text{und} \quad \mathbf{y} = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 2 \end{bmatrix}.$$

Berechnen Sie $\lambda \in \mathbb{R}$ derart, daß $\mathbf{x} \perp \mathbf{y}$.

13. Berechnen Sie $p, q \in \mathbb{R}$ derart, daß der Vektor $\mathbf{x} = -\mathbf{e}_1 + 3\mathbf{e}_2 + 2\mathbf{e}_4 + 6\mathbf{e}_5$ sowohl zu $\mathbf{a} = (p-2)\mathbf{e}_1 - \mathbf{e}_2 + (p+2)\mathbf{e}_3 + (1-q)\mathbf{e}_4 - \mathbf{e}_5$ als auch $\mathbf{b} = 3\mathbf{e}_1 + (p+1)\mathbf{e}_2 + (6-p)\mathbf{e}_3 - (4+p)\mathbf{e}_4 + (q-\frac{1}{2})\mathbf{e}_5$ orthogonal ist.

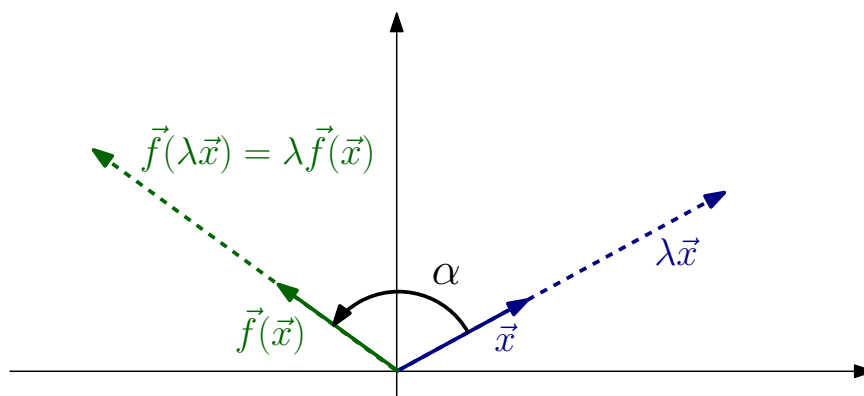
Orthogonale Abbildungen

Eine große Klasse linearer Abbildungen sind die **orthogonalen Abbildungen**. Sie spielen insbesondere bei geometrischen Transformationen eine Rolle und zeichnen sich durch die Eigenschaft aus, **Längen und Winkel unverändert zu lassen**. So lassen sich zum Beispiel Drehungen und Spiegelungen mit orthogonalen Abbildungen mathematisch beschreiben.

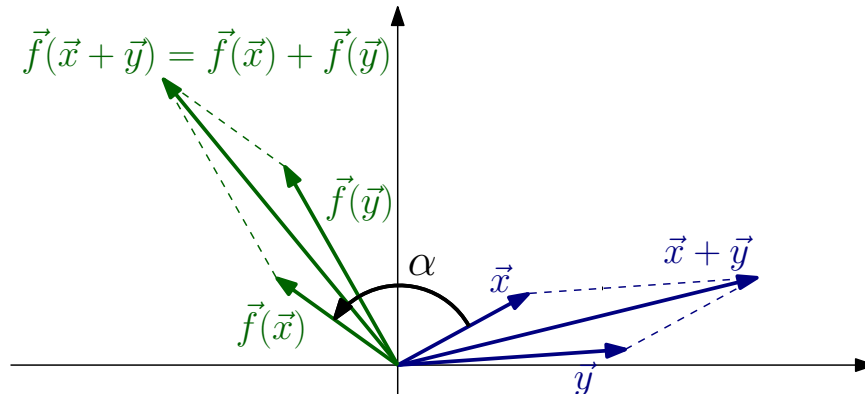
6.1 Drehungen im $\mathbb{R}^2$ und $\mathbb{R}^3$

Drehungen des $\mathbb{R}^2$ um einen Winkel α und Mittelpunkt in $\mathbf{0}$ sind zunächst **lineare** Abbildungen, wie an folgenden Skizzen deutlich wird:

- Homogenität



- Additivität



Übungsaufgabe 6.1.

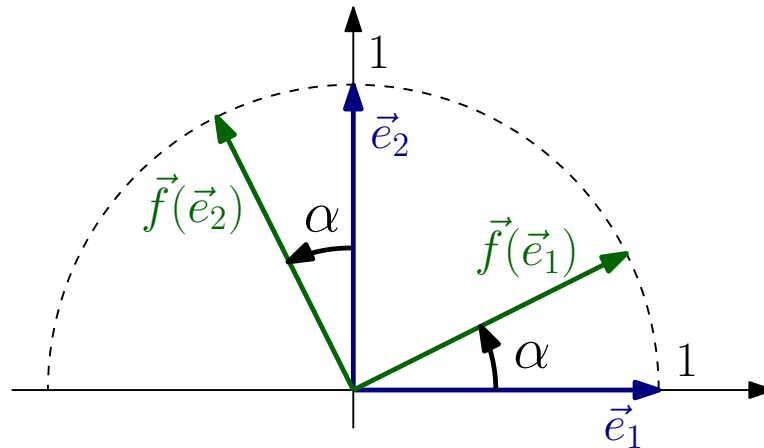
Zeichnen Sie ein analoges Bild für Spiegelungen an einer Achse $\text{span}(\{\mathbf{x}\})$ durch den Ursprung.

Natürlich läßt sich dieses geometrische Argument analog auf Drehungen und auf Spiegelungen an Ebenen im $\mathbb{R}^3$ anwenden.

Sowohl Drehungen um den Ursprung als auch Spiegelungen an einer Geraden (Ebene) durch den Ursprung können also als Matrix-Vektor-Multiplikationen beschrieben werden und wir begeben uns auf die Suche nach den Abbildungsmatrizen zu diesen Abbildungen.

Wie wir in Abschnitt 2.5.1 gelernt haben müssen wir nur die Bilder $\{f(\mathbf{e}_1), f(\mathbf{e}_2)\}$ der Standard-Basis $\{\mathbf{e}_1, \mathbf{e}_2\}$ finden, dann kennen wir die Abbildung.

Skizzieren wir die Situation:



Damit haben wir

$$f(\mathbf{e}_1) = \begin{bmatrix} \cos(\alpha) \\ \sin(\alpha) \end{bmatrix}$$

und

$$f(\mathbf{e}_2) = \begin{bmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{bmatrix}. \quad (6.1.1)$$

folglich ist die Abbildungsmatrix D_α von f , die die Drehung eines Vektors/Punktes um α um den Ursprung beschreibt, gegeben durch

$$D_\alpha = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}.$$

Aus Kapitel 4 wissen wir, daß eine Matrix genau dann invertierbar ist, wenn sie regulär ist, d.h. $\det(D_\alpha) \neq 0$. Da

$$\begin{vmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{vmatrix} = \sin^2(\alpha) + \cos^2(\alpha) = 1$$

nach trigonometrischen Pythagoras, ist D_α also für alle $\alpha \in \mathbb{R}$ invertierbar. Was ist die Inverse? Nun, etwas denken sagt uns, daß die Inverse gerade durch

die Drehung um $-\alpha$ realisiert werden sollte. (Obgleich wir Sie auch ausrechnen könnten aber Sie wissen, Mathematik ist Denkökonomie.) Es sollte also gelten

$$D_{\alpha}^{-1} = D_{-\alpha} = \begin{bmatrix} \cos(-\alpha) & -\sin(-\alpha) \\ \sin(-\alpha) & \cos(-\alpha) \end{bmatrix} = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{bmatrix}.$$

Die Eigenschaft $D_{\alpha}^{-1}D_{\alpha} = \mathbf{I}_2$ zu prüfen überlassen wir der Leserin. Wir bemerken die Eigenschaft

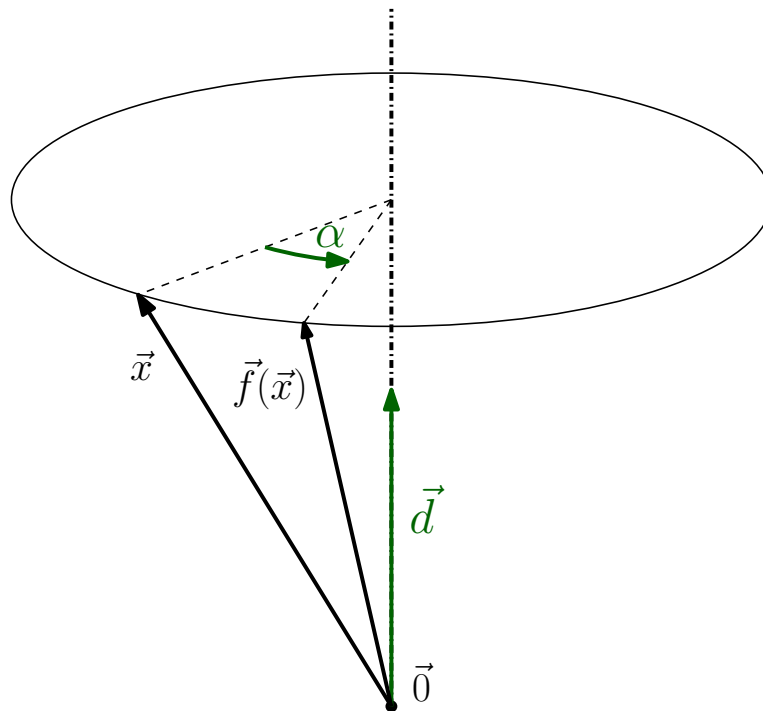
$$D_{\alpha}^{-1} = D_{\alpha}^T \quad (6.1.2)$$

Siehe dazu auch Definition 5.9. Die Matrix ist also orthogonal, was die vermittelte lineare Abbildung zu einer orthogonalen Abbildung macht.

Übungsaufgabe 6.2.

Berechnen Sie die Drehmatrix für eine Drehung des $\mathbb{R}^2$ um den Ursprung mit $\alpha = 30^\circ$. Geben Sie das Bild des Vektors $[2, 3]^T$ an.

Drehungen im $\mathbb{R}^3$ werden durch eine Drehachse (die den Ursprung enthält) und einen Drehwinkel α festgelegt. Die Drehachse wird dabei durch einen Vektor $\vec{d}$ festgelegt, der in die positive Achsenrichtung zeigt.



Besonders einfach wird die Angabe der Drehmatrizen, wenn man die Einheitsvektoren (also die Koordinatenachsen) als Drehachsen verwendet:

- Drehung um die x -Achse

$$D_{x,\alpha} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{bmatrix}$$

- Drehung um die y -Achse

$$D_{y,\alpha} = \begin{bmatrix} \cos(\alpha) & 0 & \sin(\alpha) \\ 0 & 1 & 0 \\ -\sin(\alpha) & 0 & \cos(\alpha) \end{bmatrix}$$

- Drehung um die z -Achse

$$D_{z,\alpha} = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Übungsaufgabe 6.3.

Machen Sie sich klar, daß die angegebenen Matrizen $D_{x,\alpha}$, $D_{y,\alpha}$ und $D_{z,\alpha}$ in der Tat die gewünschten Abbildungen vermitteln.

Bemerkung 6.1.

Jede beliebige Drehung im $\mathbb{R}^3$ läßt sich als Komposition von Drehungen um die Koordinatenachsen schreiben.

Physiker sprechen daher häufig von drei möglichen **Freiheitsgraden** der Rotation.

Wenn die Drehachse durch den Vektor $\mathbf{d} \in \mathbb{R}^3$, $\|\mathbf{d}\|_2 = 1$ gegeben ist und ein Winkel α , dann lautet die Matrix $D_{\mathbf{d},\alpha}$:

$$\begin{bmatrix} \cos \alpha + (1 - \cos \alpha)d_1^2 & (1 - \cos \alpha)d_1d_2 - d_3 \sin \alpha & (1 - \cos \alpha)d_1d_3 + d_2 \sin \alpha \\ (1 - \cos \alpha)d_1d_2 + d_3 \sin \alpha & \cos \alpha + (1 - \cos \alpha)d_2^2 & (1 - \cos \alpha)d_2d_3 - d_1 \sin \alpha \\ (1 - \cos \alpha)d_1d_3 - d_2 \sin \alpha & (1 - \cos \alpha)d_2d_3 + d_1 \sin \alpha & \cos \alpha + (1 - \cos \alpha)d_3^2 \end{bmatrix}$$

Eine explizite Abbildungsvorschrift ist gegeben durch

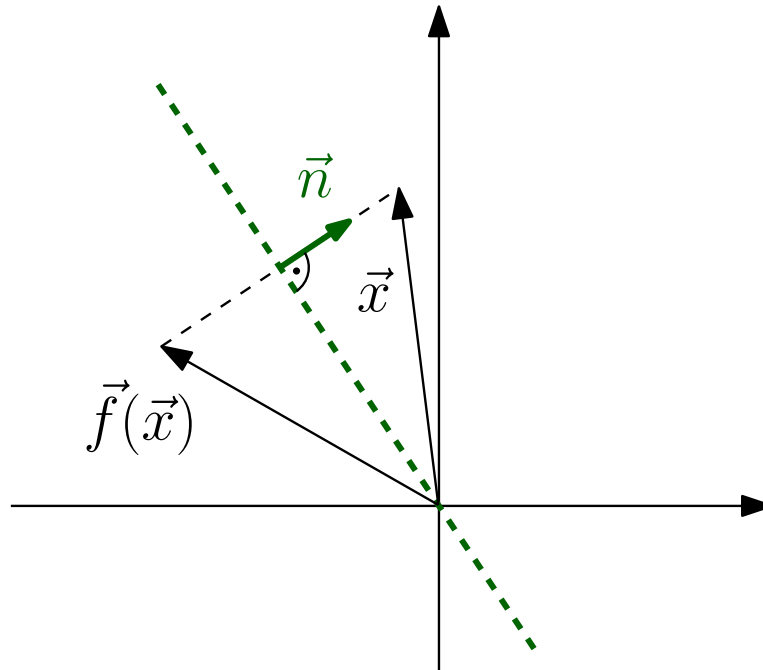
$$f(\mathbf{x}) = \cos(\alpha) \mathbf{x} + (1 - \cos(\alpha))(\mathbf{x}^T \mathbf{d}) \mathbf{d} + \sin(\alpha)(\mathbf{d} \times \mathbf{x}).$$

Dies alles schaut man bei Bedarf aber am besten in der Literatur nach.

6.2 Spiegelungen im $\mathbb{R}^2$ und $\mathbb{R}^3$

Wir betrachten zunächst Spiegelungen im $\mathbb{R}^2$ an einer Geraden durch $\mathbf{0}$ senkrecht zum Vektor $\mathbf{n}$. Dabei sei $\mathbf{n}$ auf Länge 1 normiert ($\|\mathbf{n}\|_2 = 1$).

Wir lesen das Spiegelbild $f(\mathbf{x})$ von $\mathbf{x}$ aus folgender Skizze ab:



Also

$$f(x) = x - 2(\mathbf{n}^T \mathbf{x})\mathbf{n}. \quad (\text{Orth. Projektion})$$

Beachten Sie dabei, daß wegen der Normierung von $\mathbf{n}$ die Länge der Projektion von $\mathbf{x}$ auf $\mathbf{n}$ gerade $\mathbf{n}^T \mathbf{x}$ ist. Es gilt also

$$\begin{aligned} f(x) &= x - 2(\mathbf{n}^T \mathbf{x})\mathbf{n} = x - 2\mathbf{n}(\mathbf{n}^T \mathbf{x}) \\ &= x - 2(\mathbf{n}\mathbf{n}^T)\mathbf{x} = (I - 2\mathbf{n}\mathbf{n}^T)\mathbf{x}, \end{aligned}$$

d. h. die Spiegelung wird durch Multiplikation mit der Matrix

$$S_{\mathbf{n}} = I - 2\mathbf{n}\mathbf{n}^T = \begin{bmatrix} 1 - 2n_1^2 & -2n_1n_2 \\ -2n_1n_2 & 1 - 2n_2^2 \end{bmatrix} = \begin{bmatrix} n_2^2 - n_1^2 & -2n_1n_2 \\ -2n_1n_2 & n_1^2 - n_2^2 \end{bmatrix}$$

beschrieben (beachte $n_1^2 + n_2^2 = 1$).

Übungsaufgabe 6.4.

Geben Sie die Spiegelungsmatrix für die Spiegelung an der Geraden $x_2 = -x_1$ an.

Multiplizieren Sie diese Matrix mit einem Vektor $\mathbf{x}$.

Kann man das Ergebnis auch rein geometrisch begründen?

Bei Spiegelungen im $\mathbb{R}^3$ verwendet man statt der Spiegelachse eine Spiegelebene durch den Ursprung, welche ganz analog durch den Normalenvektor $\mathbf{n}$ festgelegt ist ($\|\mathbf{n}\|_2 = 1$).

Die Spiegelungsmatrix besitzt jetzt drei Zeilen und Spalten, allerdings die gleiche Struktur:

$$S_{\mathbf{n}} = I - 2\mathbf{n}\mathbf{n}^T$$

Übungsaufgabe 6.5.

Zeichnen Sie eine geeignete Skizze, in welcher die Analogie sichtbar wird.

Wie lautet die Matrix $S_{\mathbf{n}}$ in ausgeschriebener Form?

6.3 Allgemeine Orthogonale Abbildungen

Drehungen und Spiegelungen gehören zur Klasse der **orthogonalen linearen Abbildungen**, die sich auch für höhere Dimensionen erklären lassen:

Definition 6.1 (Orthogonale Matrix).

Eine Matrix $\mathbf{U} \in \mathbb{R}^{n \times n}$ heißt **orthogonal**, wenn sie das Skalarprodukt/Innenprodukt nicht verändert^a, d. h. wenn

$$\langle \mathbf{U}\mathbf{x}, \mathbf{U}\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$$

für alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

^a $(\mathbf{U}\mathbf{y})^T(\mathbf{U}\mathbf{x}) = \mathbf{y}^T\mathbf{x}$ mit Standard-Skalarprodukt

Bemerkung 6.2.

Legt man ein anderes Skalarprodukt zu Grunde, dann könnte man andere orthogonale Abbildungen erhalten. Wir interessieren uns hier aber nur für die orthogonalen Abbildungen, die zum Standard-Skalarprodukt assoziiert sind, wie in den letzten beiden Sektionen dieses Kapitels.

Insbesondere ist eine orthogonale Matrix

- **längenerhaltend**, d.h. $\|\mathbf{U}\mathbf{x}\|_2 = \|\mathbf{x}\|_2$ für alle $\mathbf{x} \in \mathbb{R}^n$ und
- **winkeltreu**, d.h. $\angle(\mathbf{U}\mathbf{x}, \mathbf{U}\mathbf{y}) = \angle(\mathbf{x}, \mathbf{y})$ für alle $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Satz 6.1 (Charakterisierung orthogonaler Matrizen).

Sei $\mathbf{U} \in \mathbb{R}^{n \times n}$.

Dann sind die folgenden Aussagen äquivalent:

- (i) $\mathbf{U}$ ist orthogonal. (nach Definition 6.1)
- (ii) Die Spalten (Zeilen) von $\mathbf{U}$ bilden eine Orthonormalbasis des $\mathbb{R}^n$.

(iii) $\mathbf{U}$ ist invertierbar mit $\mathbf{U}^{-1} = \mathbf{U}^T$. (nach Definition 5.9)

Übungsaufgabe 6.6.

Beweisen Sie Satz 6.1. Am besten beweisen Sie

$$(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (i)$$

6.4 Verständnisfragen

1. Welche besondere Eigenschaft hat die Drehmatrix

$$\begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{bmatrix}$$

hinsichtlich der Inversenbildung?

2. Durch welche Drehmatrix werden die Koordinatenachsen eines ebenen kartesischen Koordinatensystems um

- a) 30° im math. positiven Sinn gedreht,
- b) 45° im math. negativen Sinn gedreht?

3. Welcher Zusammenhang besteht zwischen der Drehung aller Punkte einer (Koordinaten-)Ebene um einen Winkel α einerseits und der Drehung des Koordinatensystems durch en gleichen Winkel andererseits, wobei alle Punkte festbleiben.

Eigenwerte und Eigenvektoren

Wir betrachten ein Flugzeug, das auf einer holperigen Piste landet. Das Modell des Flugzeugs, das wir hier zu Grunde legen wollen ist wie folgt:

- drei Massen: m_1 (Rumpf und Motor); m_2, m_3 (Flügel),
- drei Steifigkeiten (k_1, k_2, k_3) für die "federnde" Verbindung der Teile.

Die Piste wird durch eine Sinuskurve

$$r(t) = r_0 \sin(\omega_0 t)$$

modelliert. Der Rumpf ist dann einer externen Kraft

$$f_1(t) = k_1 r_0 \sin(\omega_0 t)$$

ausgesetzt. Die Frequenz ω_0 hängt von der Landungsgeschwindigkeit v ab.

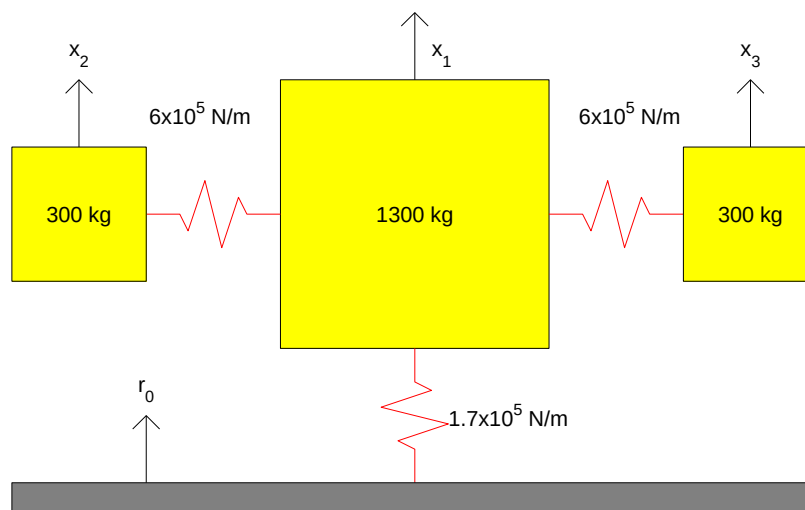
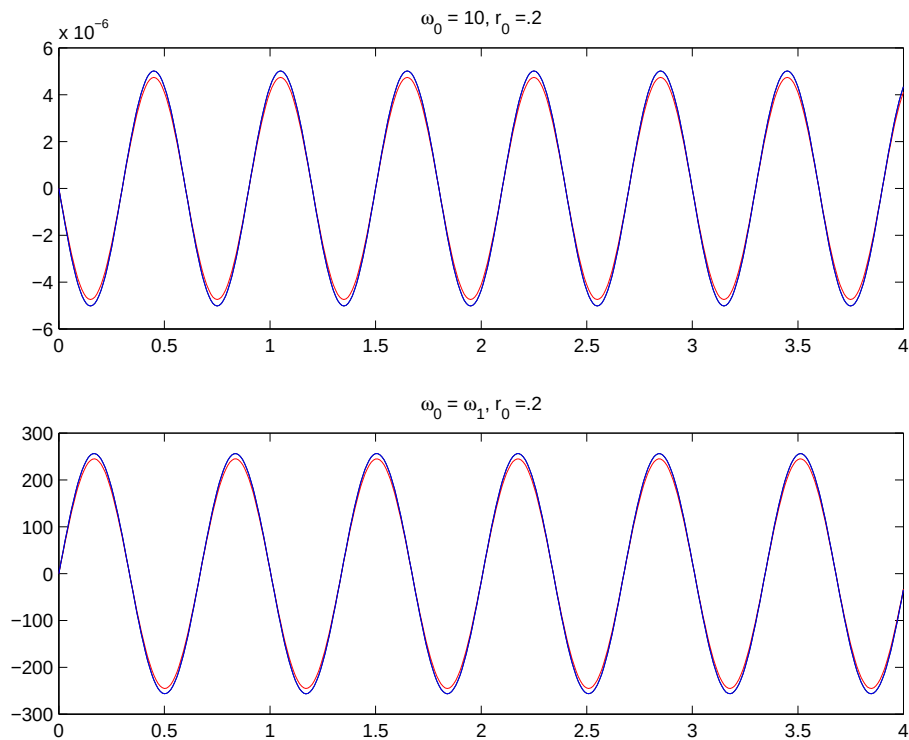


Abbildung 7.1: Skizze der Situation.

Die mathematische Analyse des Beispiels wird erst am Ende des nächsten Semesters gelingen. Wir zeigen hier aber schon die Lösungen $x_{1,2,3}(t)$ [m] über t [s] für $v = 120 \text{ km/h}$ und $v = 108 \text{ km/h}$:



Beachten Sie, daß sich die Amplituden bei den verschiedenen Landungsgeschwindigkeiten um 7 Größenordnungen(!!!) unterscheiden.

Im zweiten Fall tritt ein **Resonanzphänomen** auf: Wenn die Anregungsfrequenz ω_0 (nahezu) mit einer der **Eigenfrequenzen** $\omega_{1,2,3}$ des Flugzeugs übereinstimmt, kommt es zu gefährlich großen Oszillationen.

Die Nichtbeachtung von Eigenfrequenzen und Resonanz kann z. B. bei Brücken oder Hochhäusern katastrophale Auswirkungen haben:



Abbildung 7.2: Tacoma Narrows Bridge (WA, 1940). Bild: Prelinger Archives

Weitere Beispiele:

- Broughton suspension bridge, Manchester 1831
- Millennium footbridge, London 2000

Betrachte System aus n Webseiten (n groß), teilweise untereinander verlinkt. Sei $\mathbf{x} \in \mathbb{R}^n$, $x_i \geq 0$, $i = 1, \dots, n$ sei ein Maß für die Relevanz der i -ten Webseite, etwa für eine Suchanfrage.

Prinzip: x_i hängt von der Wichtigkeit der auf x_i verweisenden Webseiten $\{x_j\}_{j=1}^n$ ab. Sei $A \in \{0, 1\}^{n \times n}$ mit Einträgen

$$a_{i,j} = \begin{cases} 1 & \text{Webseite } j \text{ verweist auf Webseite } i \\ 0 & \text{sonst.} \end{cases}$$

Obigem Prinzip folgend könnte man dann folgendes Modell verwenden:

$$x_i = \sum_{j=1}^n a_{i,j} x_j, \quad \text{oder } \mathbf{Ax} = \mathbf{x}$$

Diese (grundlegende) Ansatz liegt Google's PageRank-Verfahren **Brin & Page** [1] zugrunde.

7.1 Definition und erste Ergebnisse

Um z.B. Resonanzphänomene zu analysieren, benötigt man die folgenden Begriffe:

Definition 7.1 (Eigenwert).

Sei $\mathbf{A} \in \mathbb{R}^{n \times n}$ (oder $\mathbb{C}^{n \times n}$). Eine Zahl $\lambda \in \mathbb{C}$ heißt **Eigenwert** von $\mathbf{A}$, wenn es einen Vektor $\mathbf{v} \in \mathbb{C}^n$, $\mathbf{v} \neq \mathbf{0}$, gibt, so daß

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v}. \quad (7.1.1)$$

Jeder Vektor $\mathbf{v} \in \mathbb{C}^n \setminus \{\mathbf{0}\}$, der (7.1.1) erfüllt, heißt **Eigenvektor** von $\mathbf{A}$ zum Eigenwert λ .

Achtung: Auch wenn wir nur reelle Matrizen betrachten: bei Eigenwerten und Eigenvektoren läßt man immer auch komplexe Zahlen zu!

Die Richtung eines Eigenvektors wird durch die lineare Abbildung $\mathbf{A}$ nicht verändert – der Eigenvektor wird lediglich gestreckt. Das rechte Bild entsteht z. B. aus dem linken durch eine "Scherung" der Leinwand.

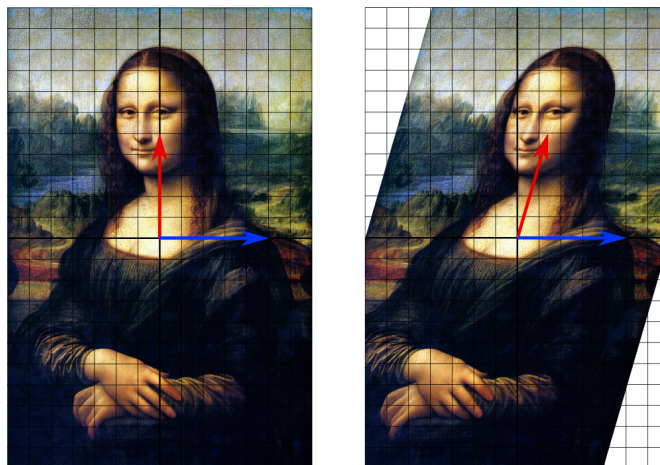


Abbildung 7.3: Bild: [Wikimedia Commons](#)

Der rote Vektor bleibt unverändert und ist damit ein Eigenvektor zum Eigenwert 1. Der blaue Vektor ändert hingegen seine Richtung und ist daher kein Eigenvektor der Scherungsabbildung.

Übungsaufgabe 7.1.

Welche reellen Eigenwerte besitzt eine Drehung im Raum um die x_1 -Achse, und welcher reelle Eigenvektor kommt in Frage?

Wie verhält es sich mit einer Spiegelung in der Ebene an einer Geraden durch 0 und senkrecht zu $\mathbf{n}$?

Argumentieren Sie rein geometrisch!

Übungsaufgabe 7.2.

Bestätigen Sie, dass $\lambda = 1$ Eigenwert von $\mathbf{A} = \begin{bmatrix} 5 & -8 \\ -1 & 3 \end{bmatrix}$ mit zugehörigem Eigenvektor $\mathbf{v} = [2, 1]^T$ ist.

Übungsaufgabe 7.3.

Zeigen Sie, daß jedes komplexe Vielfache des Vektors $[1, i]^T$ ein Eigenvektor von $B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ zum Eigenwert $\lambda = i$ ist.

Formulieren Sie eine allgemeingültige Aussage und bestätigen Sie diese durch Einsetzen in (7.1.1).

Bislang wissen wir weder, wie man EW und EV berechnet, noch ob in den Beispielen alle EW und EV erfaßt wurden. Wir benutzen zur Herleitung der Formel für die Eigenwertberechnung folgende Äquivalenzkette: $\lambda \in \mathbb{C}$ ist, per Definition, Eigenwert von $\mathbf{A} \in \mathbb{R}^{n \times n}$

$\Leftrightarrow$ Es gibt einen Vektor $\mathbf{v} \neq \mathbf{0}$ mit $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$.

$\Leftrightarrow$ Ein Vektor $\mathbf{v} \neq \mathbf{0}$ löst das homogene LGS $(\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{v} = \mathbf{0}$.

$\Leftrightarrow$ Die Matrix $\mathbf{A} - \lambda \mathbf{I}_n$ ist singulär¹, d.h. $\det(\mathbf{A} - \lambda \mathbf{I}_n) = 0$.

Die Eigenwerte sind also gerade die Nullstellen der Funktion

$$\chi_{\mathbf{A}}(\lambda) := \det(\mathbf{A} - \lambda \mathbf{I}_n).$$

Diese Funktion ist ein Polynom (wieso?) vom Grad n und wird das **charakteristische Polynom** von $\mathbf{A}$ genannt.

Satz 7.1.

Für $\mathbf{A} \in \mathbb{R}^{n \times n}$ ist das charakteristische Polynom $\chi_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I}_n)$ ein Polynom vom exakten Grad n mit reellen Koeffizienten^a und Höchstkoeffizienten 1 oder -1 .

Die Eigenwerte von $\mathbf{A}$ sind genau die Nullstellen von $\chi_{\mathbf{A}}$, also die Lösungen der Gleichung

$$\chi_{\mathbf{A}}(\lambda) = \det(\mathbf{A} - \lambda \mathbf{I}_n) = 0. \quad (7.1.2)$$

Gleiches gilt für $\mathbf{A} \in \mathbb{C}^{n \times n}$ nur dass die Koeffizienten nicht mehr reell sind.

^aDie Komplexen Nullstellen kommen in diesem Fall also in Paaren komplex konjugierter Zahlen.

Übungsaufgabe 7.4.

Berechnen Sie sämtliche Eigenwerte der Matrizen

$$\mathbf{A} = \begin{bmatrix} 5 & -8 \\ -1 & 3 \end{bmatrix} \quad \text{und} \quad \mathbf{B} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Sei λ Eigenwert der Matrix $\mathbf{A}$. Die zugehörigen Eigenvektoren von $\mathbf{A}$ sind die (nicht-trivialen) Lösungen des homogenen Gleichungssystems

$$(\mathbf{A} - \lambda \mathbf{I}_n) \mathbf{v} = \mathbf{0}. \quad (7.1.3)$$

¹Eine Matrix ist genau dann singulär, wenn sie nicht regulär ist. Eine Matrix heißt regulär, wenn ihre Determinante von 0 verschieden ist.

Zusammen mit $\mathbf{0}$ bilden sie einen Unterraum des $\mathbb{C}^n$, den sogenannten **Eigenraum** von $\mathbf{A}$ zum Eigenwert λ ,

$$\begin{aligned}\text{Eig}(\mathbf{A}, \lambda) &:= \{\mathbf{x} \in \mathbb{C}^n : \mathbf{A}\mathbf{x} = \lambda\mathbf{x}\} \\ &= \{\mathbf{x} \in \mathbb{C}^n : (\mathbf{A} - \lambda\mathbf{I}_n)\mathbf{x} = \mathbf{0}\} \\ &= N(\mathbf{A} - \lambda\mathbf{I}_n)\end{aligned}$$

Übungsaufgabe 7.5.

Berechnen Sie sämtliche Eigenvektoren zu den Eigenwerten der Matrizen $\mathbf{A}$ und $\mathbf{B}$ von Seite 145 Aufgabe 7.4.

7.2 Ähnliche Matrizen

Definition 7.2 (Ähnlichkeit).

Zwei Matrizen $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n \times n}$ heißen **ähnlich**, wenn es eine invertierbare Matrix $\mathbf{V} \in \mathbb{C}^{n \times n}$ gibt, so daß

$$\mathbf{A} = \mathbf{V}^{-1} \mathbf{B} \mathbf{V}. \quad (7.2.1)$$

Bemerkung 7.1.

Zwei Bemerkungen zur Definition von Ähnlichkeit von Matrizen.

(i) Gleichung (7.2.1) ist äquivalent zu

$$\mathbf{V} \mathbf{A} = \mathbf{B} \mathbf{V} \quad \text{und} \quad \mathbf{B} = \mathbf{V} \mathbf{A} \mathbf{V}^{-1}.$$

(ii) Ähnlichkeit ist eine Äquivalenzrelation, d.h. jede Matrix ist zu sich selbst ähnlich, wenn $\mathbf{A}$ zu $\mathbf{B}$ ähnlich ist, dann ist $\mathbf{B}$ zu $\mathbf{A}$ ähnlich und es gilt, wenn $\mathbf{A}$ zu $\mathbf{B}$ ähnlich ist und $\mathbf{B}$ zu $\mathbf{C}$, dann ist $\mathbf{A}$ ähnlich zu $\mathbf{C}$.

Ähnliche Matrizen haben das gleiche charakteristische Polynom, wie folgende Rechnung zeigt: aus $\mathbf{A} = \mathbf{V}^{-1} \mathbf{B} \mathbf{V}$ erhalten wir

$$\begin{aligned} \chi_{\mathbf{A}}(\lambda) &= \chi_{\mathbf{V}^{-1} \mathbf{B} \mathbf{V}}(\lambda) = \det(\mathbf{V}^{-1} \mathbf{B} \mathbf{V} - \lambda \mathbf{I}_n) \\ &= \det(\mathbf{V}^{-1} (\mathbf{B} - \lambda \mathbf{I}_n) \mathbf{V}) \\ &= \det(\mathbf{V}^{-1}) \det(\mathbf{B} - \lambda \mathbf{I}) \det(\mathbf{V}) \\ &= (\det(\mathbf{V}))^{-1} \det(\mathbf{B} - \lambda \mathbf{I}_n) \det(\mathbf{V}) \\ &= \chi_{\mathbf{B}}(\lambda) \end{aligned}$$

da $\det(\mathbf{V}^{-1}) = \frac{1}{\det(\mathbf{V})}$.

Wir fassen zusammen:

Satz 7.2.

Ähnliche Matrizen haben das gleiche charakteristische Polynom und damit die gleichen Eigenwerte.

Die Eigenvektoren zu einem Eigenwert λ sind dabei allerdings verschieden: $\mathbf{x}$ ist genau dann Eigenvektor von $\mathbf{A} = \mathbf{V}^{-1}\mathbf{B}\mathbf{V}$, wenn $\mathbf{V}\mathbf{x}$ Eigenvektor von $\mathbf{B}$ ist.

7.3 Basistransformationen

Um den Ähnlichkeitsbegriff vollständig zu verstehen, muß man sich mit **Basistransformationen** auseinandersetzen. Siehe auch Abschnitt 2.5.1. Wir gehen dabei von der Darstellung $\mathbf{B} = \mathbf{VAV}^{-1}$ aus. Zu einem gegebenen Vektor $\mathbf{x}$ läßt sich $\mathbf{Vx}$ als Linearkombination der Spalten $\mathbf{V}[\cdot, 1], \dots, \mathbf{V}[\cdot, n]$ von $\mathbf{V}$ schreiben:

$$\mathbf{Vx} = x_1 \mathbf{V}[\cdot, 1] + x_2 \mathbf{V}[\cdot, 2] + \dots + x_n \mathbf{V}[\cdot, n].$$

Faßt man $\mathbf{x}$ als Koordinatenvektor bezüglich der Basis

$$B_{\mathbf{V}} = \{\mathbf{V}[\cdot, 1], \dots, \mathbf{V}[\cdot, n]\}$$

auf, so ist $\mathbf{Vx}$ gerade der zugehörige Koordinatenvektor bezüglich der Standardbasis.

Man sagt daher auch, $\mathbf{V}$ stellt eine **Basistransformation** von der Basis $B_{\mathbf{V}}$ in die Standardbasis dar.

Die Abbildung $\mathbf{V}^{-1}$ macht die Transformation rückgängig und stellt somit eine Basistransformation von der Standardbasis nach $B_{\mathbf{V}}$ dar.

Übungsaufgabe 7.6.

Verifizieren Sie beide Aussagen am Beispiel der Basis

$$B = \left\{ \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$$

und des Vektors

$$\begin{bmatrix} \sqrt{2} \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}_B$$

In $\mathbf{B}\mathbf{x} = \mathbf{V}\mathbf{A}\mathbf{V}^{-1}\mathbf{x}$ kann man die rechte Seite nun von rechts nach links wie folgt lesen:

- Stelle $\mathbf{x}$ als Koordinatenvektor bezüglich der Basis

$$B_{\mathbf{V}} = \{\mathbf{V}[\cdot, 1], \dots, \mathbf{V}[\cdot, n]\}$$

dar (d. h. Multiplikation mit $\mathbf{V}^{-1}$).

- Führe die durch $\mathbf{B}$ beschriebene lineare Abbildung in den Koordinaten bezüglich $B_{\mathbf{V}}$ aus (d. h. Multiplikation mit $\mathbf{A}$).
- Transformiere das Ergebnis wieder zurück in die Standardkoordinaten (Multiplikation mit $\mathbf{V}$).

Wir haben somit erkannt:

Satz 7.3.

Ähnliche Matrizen stellen die gleiche lineare Abbildung bezüglich verschiedener Basen dar.

Folglich besitzen ähnliche Matrizen

- dieselbe Determinante,
- denselben Rang und
- denselben Defekt.

Auch die Aussage von Satz 7.2 wird mit dieser Erkenntnis noch ein Stück verständlicher.

7.4 Diagonalisierbare Matrizen

Definition 7.3 (Diagonalisierbare Matrix).

Eine Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ heißt **diagonalisierbar**, wenn $\mathbf{A}$ ähnlich zu einer Diagonalmatrix $\mathbf{D} \in \mathbb{C}^{n \times n}$ ist, d. h. wenn es eine invertierbare Matrix $\mathbf{V} \in \mathbb{C}^{n \times n}$ gibt mit

$$\mathbf{D} = \mathbf{V}^{-1} \mathbf{A} \mathbf{V}.$$

Aus dem bisher besprochenem ergibt sich unmittelbar

Satz 7.4.

Eine Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ ist genau dann diagonalisierbar, wenn es n linear unabhängige Eigenvektoren von $\mathbf{A}$ – also eine Basis von $\mathbb{C}^n$ aus Eigenvektoren von $\mathbf{A}$ – gibt.

Sei $\mathbf{A} \in \mathbb{C}^{n \times n}$ eine Matrix, für die eine Basis des $\mathbb{C}^n$ aus Eigenvektoren von $\mathbf{A}$ existiert (vgl. Satz 7.4). Dann kann die Abbildung $\mathbf{A}$ bezüglich dieser Basis nur durch Streckungen der Basisvektoren ausgedrückt werden (Multiplikation mit einer Diagonalmatrix).

Beispiel 7.1.

Die Matrix $\mathbf{A} = \begin{bmatrix} 5 & -8 \\ -1 & 3 \end{bmatrix}$ besitzt die Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = 7$ mit den Eigenvektoren $\mathbf{v}_1 = [2, 1]^T$ und $\mathbf{v}_2 = [-4, 1]^T$ (vgl. S. 7.4). Es gilt die Darstellung

$$\mathbf{A} = \mathbf{V} \mathbf{D} \mathbf{V}^{-1} = \begin{bmatrix} 2 & -4 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 7 \end{bmatrix} \begin{bmatrix} 2 & -4 \\ 1 & 1 \end{bmatrix}^{-1}.$$

In der Diagonalmatrix $\mathbf{D}$ stehen die Eigenwerte und in den Spalten von $\mathbf{V}$ die Eigenvektoren.

Wir widmen uns nun der Frage, wie groß die Dimension des von den Eigenvektoren einer Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ aufgespannten Unterraums ist.

Insbesondere wollen wir wissen, **wann** es eine Basis des $\mathbb{C}^n$ aus Eigenvektoren von $\mathbf{A}$ gibt.

Wir beginnen mit folgendem Satz:

Satz 7.5.

*Gehören die Eigenvektoren $\mathbf{v}_1, \dots, \mathbf{v}_r$ zu **verschiedenen** Eigenwerten $\lambda_1, \dots, \lambda_r$ der $n \times n$ -Matrix $\mathbf{A}$, dann sind sie linear unabhängig.*

Übungsaufgabe 7.7.

Machen Sie sich dies zumindest für den Fall $r = 2$ klar.

Hinweis: Nehmen Sie an, die Eigenvektoren $\mathbf{v}_1$ und $\mathbf{v}_2$ zu den Eigenwerten λ_1 und λ_2 von $\mathbf{A}$ sind linear abhängig. Benutzen Sie dann die Eigenschaften der Eigenvektoren durch Anwendung von $\mathbf{A}$.

Wenn es n **verschiedene** Eigenwerte zur Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ gibt, existiert also eine Basis des $\mathbb{C}^n$, die nur aus Eigenvektoren von $\mathbf{A}$ besteht. Da Eigenwerte Lösungen von Polynomgleichungen sind, haben wir den Begriff der Vielfachheit. Da es sein kann, daß die Dimension der Eigenräume nicht mit der Vielfachheit des Eigenwertes übereinstimmt brauchen wir zwei Vielfachheitsbegriffe:

Definition 7.4.

Es sei $\mathbf{A} \in \mathbb{C}^{n \times n}$ und $\lambda \in \mathbb{C}$ ein Eigenwert von $\mathbf{A}$. Dann definieren wir

- (i) die **algebraische Vielfachheit** $\nu_{\text{alg}}(\lambda)$ des Eigenwerts λ ist seine Vielfachheit als Nullstelle des charakteristischen Polynoms $\chi_{\mathbf{A}}$ von $\mathbf{A}$.*
- (ii) die **geometrische Vielfachheit** $\nu_{\text{geom}}(\lambda)$ des Eigenwerts λ ist die Dimension des zugehörigen Eigenraums $\dim(\text{Eig}(\mathbf{A}, \lambda))$.*

Bemerkung 7.2.

Bemerkungen zur Berechnung der Vielfachheiten:

- (i) *Um die algebraischen Vielfachheiten zu bestimmen, muß man nur das charakteristische Polynom χ_A kennen.*
- (ii) *Um die geometrische Vielfachheiten zu berechnen, reicht die Kenntnis von χ_A allein nicht aus.*

Nach dem Fundamentalsatz der Algebra summieren sich die algebraischen Vielfachheiten der Eigenwerte einer Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ zu n . Die Summe der geometrischen Vielfachheiten kann dagegen kleiner sein. Es gilt

Satz 7.6.

Sei λ Eigenwert der Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ mit algebraischer Vielfachheit $\nu_{\text{alg}}(\lambda)$ und geometrischer Vielfachheit $\nu_{\text{geom}}(\lambda)$. Dann gilt:

$$1 \leq \nu_{\text{geom}}(\lambda) \leq \nu_{\text{alg}}(\lambda) \leq n.$$

Übungsaufgabe 7.8.

Machen Sie sich am Beispiel der Matrix $\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$ klar, daß die geometrische Vielfachheit eines Eigenwerts tatsächlich kleiner sein kann als die algebraische.

Bei Beispiel 7.8 handelt es sich um einen sogenannten **Jordan-Block**, d. h. eine Matrix der Form

$$\begin{bmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} \in \mathbb{R}^{n \times n}.$$

Dieser besitzt nur den Eigenwert λ mit geometrischer Vielfachheit 1 und algebraischer Vielfachheit n .

Übungsaufgabe 7.9.

Bestätigen Sie die obigen Aussagen über den angegebenen Jordan-Block.

Für die geometrische Vielfachheit nutzen Sie am besten die Beziehung

$$\nu_{\text{geom}}(\lambda) = \dim(\text{Eig}(\mathbf{A}, \lambda)) = n - \text{rank}(\mathbf{A} - \lambda \mathbf{I}_n).$$

Diagonalisierbarkeit läßt sich nun mittels Vielfachheiten auch folgendermaßen charakterisieren:

Satz 7.7.

Eine Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ ist genau dann diagonalisierbar, wenn die algebraische und geometrische Vielfachheit für jeden Eigenwert übereinstimmen. Sind $\lambda_1, \dots, \lambda_n$ die Eigenwerte und $\mathbf{v}_1, \dots, \mathbf{v}_n$ zugeordnete Eigenvektoren, die eine Basis des $\mathbb{C}^n$ bilden, so gilt die Darstellung

$$\mathbf{A} = \mathbf{V} \mathbf{D} \mathbf{V}^{-1}.$$

Dabei gilt $\mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_n)$, und $\mathbf{v}_1, \dots, \mathbf{v}_n$ sind in dieser Reihenfolge die Spalten von $\mathbf{V}$.

Übungsaufgabe 7.10.

Die Matrix $A = \begin{bmatrix} 5 & -8 \\ -1 & 3 \end{bmatrix}$ besitzt die Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = 7$ mit den Eigenvektoren $\mathbf{v}_1 = [2, 1]^T$ und $\mathbf{v}_2 = [-4, 1]^T$. Sie bilden eine Basis des $\mathbb{C}^2$ (und in diesem Falle auch eine Basis des $\mathbb{R}^2$).

Das charakteristische Polynom einer reellen 2×2 -Matrix $\mathbf{A}$ besitzt die Struktur

$$\chi_{\mathbf{A}}(\lambda) = \lambda^2 + p\lambda + q \quad (\text{mit } p, q \in \mathbb{R}).$$

In Kombination mit Satz 7.6 ergeben sich verschiedene Möglichkeiten für die Eigenwerte. Es existieren

- entweder zwei verschiedene *reelle* Eigenwerte λ_1 und λ_2 , die geometrische und algebraische Vielfachheit 1 haben. Zu jedem Eigenwert gibt es reelle Eigenvektoren.
- oder zwei verschiedene *konjugiert-komplexe* Eigenwerte λ_1 und λ_2 , die geometrische und algebraische Vielfachheit 1 haben. Zu jedem Eigenwert gibt es komplexe Eigenvektoren.
- oder nur einen Eigenwert λ , der dann zwangsläufig reell (konjugiert-komplex zu sich selbst) ist.

Er besitzt die algebraische Vielfachheit 2 und

- entweder die geometrische Vielfachheit 2, d. h. jeder Vektor aus $\mathbb{C}^2$ ist Eigenvektor. $\mathbf{A}$ ist dann ähnlich zu $\lambda \mathbf{I}_2$, d. h. es existiert eine invertierbare Matrix $\mathbf{V}$ mit zwei linear unabhängigen Eigenvektoren zu λ als Spalten sodaß

$$\mathbf{V}^{-1}\mathbf{A}\mathbf{V} = \lambda \mathbf{I}_2 = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$$

- oder die geometrische Vielfachheit 1.

Übungsaufgabe 7.11.

Ordnen Sie die bisher besprochenen Beispiele den entsprechenden Fällen zu!

Übungsaufgabe 7.12.

Führen Sie eine ähnliche Analyse für den 3×3 -Fall durch.

7.5 Jordansche Normalform

Gilt $\nu_{\text{geom}}(\lambda) < \nu_{\text{alg}}(\lambda)$ für ein Eigenwert λ einer Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$, so ist $\mathbf{A}$ nicht diagonalisierbar. Folgende "fast diagonale" Form existiert aber im allgemeinen Fall:

Satz 7.8 (Jordansche Normalform).

Jede Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ ist ähnlich zu einer Blockdiagonalmatrix $\mathbf{J} = \text{diag}(\mathbf{J}_1, \dots, \mathbf{J}_s)$ mit Diagonalblöcken der Form

$$\mathbf{J}_k = \begin{bmatrix} \lambda_k & 1 & & \\ & \lambda_k & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_k \end{bmatrix}, \quad \text{d.h.} \quad \mathbf{A} = \mathbf{V}\mathbf{J}\mathbf{V}^{-1} \quad (7.5.1)$$

mit Eigenwerten λ_k von $\mathbf{A}$. Zu jedem Eigenwert von $\mathbf{A}$ existieren ein oder mehrere solcher **Jordan-Blöcke**. Ist n_j die Dimension des größten zum Eigenwert λ_j von $\mathbf{A}$ gehörenden Jordan-Blocks, so heißt $m_A(\lambda) = \prod_j (\lambda - \lambda_j)^{n_j}$ das **Minimalpolynom** von A , welches das charakteristische Polynom von A teilt. Für die Spaltenvektoren $\mathbf{v}$ von $\mathbf{V}$ im Indexbereich eines Jordan-Blocks zum Eigenwert λ_j gilt

$$(\mathbf{A} - \lambda_j \mathbf{I}_n)^\ell \mathbf{v} = \mathbf{0} \quad \text{für ein } \ell \text{ mit } 1 \leq \ell \leq n_j.$$

Bemerkung 7.3.

(i) Die Spalten der Ähnlichkeitstransformation V in (7.5.1) heißen **Hauptvektoren**.

Zu jedem Eigenwert λ von $\mathbf{A}$ existieren $\nu_{\text{alg}}(\lambda)$ linear unabhängige Hauptvektoren.

Da Hauptvektoren zu verschiedenen Eigenwerten linear unabhängig sind existiert somit eine Basis des $\mathbb{C}^n$ aus Hauptvektoren von $\mathbf{A}$.

(ii) Im Fall $\nu_{\text{geom}}(\lambda) = \nu_{\text{alg}}(\lambda)$ sind alle Hauptvektoren von $\mathbf{A}$ zum Eigenwert λ auch Eigenvektoren.

(iii) Es gilt $m_{\mathbf{A}}(\mathbf{A}) = \mathbf{O}$ und $m_{\mathbf{A}}$ ist das Polynom niedrigsten Grades mit dieser Eigenschaft. Somit gilt auch $\chi_{\mathbf{A}}(\mathbf{A}) = \mathbf{O}$, was auch als Satz von Cayley-Hamilton bekannt ist.

7.6 Spektralsatz

Satz 7.9 (Spektralsatz).

Sei $\mathbf{A} \in \mathbb{R}^{n \times n}$.

Dann gelten:

(i) $\mathbf{A}$ und $\mathbf{A}^T$ besitzen dasselbe charakteristische Polynom, also dieselben Eigenwerte (mit i. A. verschiedenen Eigenräumen).

(ii) Besitzt $\mathbf{A}$ den Eigenvektor $\mathbf{x}$ zum Eigenwert λ , dann besitzen

$$\alpha \mathbf{A}, \mathbf{A}^m, \mathbf{A} + \beta \mathbf{I}_n, p(\mathbf{A}) = \alpha_m \mathbf{A}^m + \dots + \alpha_1 \mathbf{A} + \alpha_0 \mathbf{I}_n$$

denselben Eigenvektor $\mathbf{x}$, allerdings zum Eigenwert

$$\alpha \lambda, \lambda^m, \lambda + \beta, p(\lambda) = \alpha_m \lambda^m + \dots + \alpha_1 \lambda + \alpha_0.$$

(iii) $\mathbf{A}$ ist genau dann invertierbar, wenn alle Eigenwerte von $\mathbf{A}$ von 0 verschieden sind. Ist dann λ ein Eigenwert von $\mathbf{A}$ mit Eigenvektor $\mathbf{x}$, so ist λ^{-1} ein Eigenwert von $\mathbf{A}^{-1}$ mit demselben Eigenvektor $\mathbf{x}$.

Übungsaufgabe 7.13.

Beweisen Sie Satz 7.9.

7.7 Eigenwerte spezieller Matrizen

Weiterhin kann man Eigenwerte sehr einfach bestimmen, wenn $\mathbf{A}$ bestimmte strukturelle Eigenschaften besitzt.

Das erste Ergebnis bezieht sich auf Diagonal- und Dreiecksmatrizen.

Satz 7.10.

Ist $\mathbf{A}$ eine (untere oder obere) Dreiecksmatrix, so sind die Hauptdiagonalelemente von $\mathbf{A}$ genau die Eigenwerte von $\mathbf{A}$.

Dies trifft insbesondere dann zu, wenn $\mathbf{A}$ eine Diagonalmatrix ist. In diesem Fall sind die Einheitsvektoren zugehörige Eigenvektoren.

Auch symmetrische Matrizen $\mathbf{A}^T = \mathbf{A}$ haben bemerkenswerte Eigenschaften:

Satz 7.11.

Ist $\mathbf{A} \in \mathbb{R}^{n \times n}$ symmetrisch, so besitzt $\mathbf{A}$ nur reelle Eigenwerte. Es gibt eine Orthonormalbasis des $\mathbb{R}^n$, die aus Eigenvektoren von $\mathbf{A}$ besteht.

Es gibt also eine Diagonalmatrix $\mathbf{D} \in \mathbb{R}^{n \times n}$ und eine orthogonale Matrix $\mathbf{U} \in \mathbb{R}^{n \times n}$ (d. h. $\mathbf{U}^T = \mathbf{U}^{-1}$), sodaß

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^T.$$

Insbesondere ist $\mathbf{A}$ diagonalisierbar.

Übungsaufgabe 7.14.

Es sei $\mathbf{A}$ antisymmetrisch, d. h. $\mathbf{A}^T = -\mathbf{A}$. Zeigen Sie, daß die Eigenwerte von $\mathbf{A}$ alle rein imaginär sind, d. h. $\bar{\lambda} = -\lambda$ für alle Eigenwerte gilt.

Für orthogonale Matrizen gibt es ein ähnliches Ergebnis. Beachten Sie aber, daß die Eigenwerte hier i. A. komplex sind!

Satz 7.12.

Ist $\mathbf{A} \in \mathbb{R}^{n \times n}$ orthogonal, so besitzt $\mathbf{A}$ nur Eigenwerte mit Betrag 1. Es gibt eine Orthonormalbasis des $\mathbb{C}^n$, die aus Eigenvektoren von $\mathbf{A}$ besteht.

Es gibt also eine Diagonalmatrix $\mathbf{D} \in \mathbb{C}^{n \times n}$ und eine unitäre Matrix $\mathbf{U} \in \mathbb{C}^{n \times n}$ (d. h. $\mathbf{U}^H = \mathbf{U}^{-1}$), so dass

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^H.$$

Insbesondere ist $\mathbf{A}$ diagonalisierbar.

Bemerkung 7.4.

Die Klasse der Matrizen, welche eine Basis aus **orthonormalen** Eigenvektoren besitzen, ist größer als die der symmetrischen (Hermiteischen) Matrizen. Genau trifft dies zu für **normale** Matrizen. Diese sind charakterisiert durch die Eigenschaft

$$\mathbf{A}\mathbf{A}^T = \mathbf{A}^T\mathbf{A}, \quad \text{bzw.} \quad \mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A}.$$

Übungsaufgabe 7.15.

Bestimmen Sie die Eigenwerte und Eigenvektoren der Matrizen

$$A = \begin{bmatrix} 8 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & -3 & 1 \end{bmatrix} \quad \text{und} \quad D_\alpha = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}$$

Geben Sie in beiden Fällen eine Darstellung der Form $\mathbf{V}\mathbf{D}\mathbf{V}^{-1}$ an. Erkennen Sie die Matrix $\mathbf{B}$ von S. 7.4 f. wieder?

7.8 Anwendung: Hauptachsentransformation

Die durch die Gleichungen

$$\frac{x^2}{4} + \frac{y^2}{9} = 1 \quad \text{bzw.} \quad y = 4(x+1)^2 + 2$$

festgelegten Punktmengen erkennt man leicht als **Ellipse** symmetrisch zu den Koordinatenachsen mit Halbachsen 2 und 3 bzw. nach oben offene **Parabel** mit Scheitelpunkt $(x_S, y_S) = (-1, 2)$.

Dies ist weniger offensichtlich bei

$$\frac{x^2}{9} + 4xy - \frac{y^2}{4} + x - 2y = 4.$$

Die **Hauptachsentransformation** bietet eine Variablensubstitution, mit Hilfe derer eine Klassifikation quadratischer Funktionen einfach möglich ist.

Jede quadratische Funktion $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ besitzt die allgemeine Form

$$\phi(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c \quad (7.8.1)$$

mit einer symmetrischen Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, einem Vektor $\mathbf{b}$ und einer Zahl $c \in \mathbb{R}$.

Beispiel 7.2.

Im Fall $n = 2$ haben wir beispielsweise die quadratische Funktion

$$\phi(x_1, x_2) = a_{1,1}x_1^2 + a_{1,2}x_1x_2 + a_{2,1}x_2x_1 + a_{2,2}x_2^2 + b_1x_1 + b_2x_2 + c$$

Die Aufgabe der **Hauptachsentransformation** besteht darin, durch einen geeigneten Basiswechsel im $\mathbb{R}^n$ die quadratische Funktion in eine einfache Form zu bringen, an der man leicht ihren Typ ablesen kann.

Dies ist auch hilfreich bei der Klassifikation von **Quadriken** (auch **Kegelschnitte** genannt), also den Punktmengen

$$Q = \{\mathbf{x} \in \mathbb{R}^n : \phi(\mathbf{x}) = 0\}.$$

Da die Matrix $\mathbf{A}$ ohne Beschränkung der Allgemeinheit als symmetrisch angenommen werden kann besitzt diese eine Basis aus orthonormalen Eigenvektoren, d.h. eine orthogonale Matrix $\mathbf{Q}$ mit

$$\mathbf{Q}^T \mathbf{A} \mathbf{Q} = \mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_n)$$

mit reellen Eigenwerten $\lambda_1, \dots, \lambda_n$. Der Hauptteil $x^T \mathbf{A} x$ einer quadratischen Funktion (7.8.1) heißt auch **quadratische Form**.

Die orthogonalen Eigenrichtungen der Matrix $\mathbf{A}$ heißen **Hauptachsen** der quadratischen Funktion.

Für die neue Variable $\mathbf{y} := \mathbf{Q}^T \mathbf{x}$ gilt dann $\mathbf{x} = \mathbf{Q} \mathbf{y}$ und, mit dem neuen Vektor $\mathbf{d} := \mathbf{Q}^T \mathbf{b}$ erhalten wir

$$\begin{aligned} \phi(\mathbf{x}) &= \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c = (\mathbf{Q} \mathbf{y})^T \mathbf{A} (\mathbf{Q} \mathbf{y}) + \mathbf{b}^T (\mathbf{Q} \mathbf{y}) + c \\ &= \mathbf{y}^T \mathbf{Q}^T \mathbf{A} \mathbf{Q} \mathbf{y} + \mathbf{b}^T \mathbf{Q} \mathbf{y} + c = \mathbf{y}^T \mathbf{D} \mathbf{y} + \mathbf{d}^T \mathbf{y} + c \\ &= \sum_{j=1}^n \lambda_j y_j^2 + \sum_{j=1}^n d_j y_j + c. \end{aligned}$$

Als letzter Schritt folgt nun quadratisches Ergänzen in jeder Variablen y_j :

$$\lambda_j y_j^2 + d_j y_j = \lambda_j \left(y_j + \frac{d_j}{2\lambda_j} \right)^2 - \frac{d_j^2}{4\lambda_j}.$$

Mit einer weiteren Substitution $z_j = y_j + \frac{d_j}{2\lambda_j}$, $j = 1, \dots, n$ erhalten wir

$$\phi(\mathbf{x}) = \sum_{j=1}^n \lambda_j z_j^2 + d = \lambda_1 z_1^2 + \dots + \lambda_n z_n^2 + d.$$

mit

$$d := c - \sum_{j=1}^n \frac{d_j^2}{4\lambda_j}.$$

Übungsaufgabe 7.16.

Führen Sie die Hauptachsentransformation durch für die Quadrik

$$2x^2 - y^2 + 4xy - 2x + y - 6 = 0.$$

Alle Eigenwerte von A ungleich Null	
$\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$	Ellipse mit Halbachsen a, b
$\frac{x^2}{a^2} + \frac{y^2}{b^2} + 1 = 0$	leere Menge
$\frac{x^2}{a^2} - \frac{y^2}{b^2} - 1 = 0$	Hyperbel
$x^2 + a^2 y^2 = 0$	Punkt $\{(0, 0)\}$
$x^2 - a^2 y^2 = 0$	Geradenpaar $y = \pm a x$
Ein Eigenwert von A gleich Null	
$x^2 - 2py = 0$	Parabel
$x^2 - a^2 = 0$	paralleles Geradenpaar
$x^2 + a^2 = 0$	leere Menge
$x^2 = 0$	Gerade $x = 0$ (y -Achse)

Tabelle 7.1: Quadrikeneinteilung im $\mathbb{R}^2$.

7.9 Verständnisfragen

1. Wie erfolgt die schrittweise Berechnung der Eigenwerte und der zugehörigen Eigenräume einer Matrix $\mathbf{A} \in \mathbb{R}^n$.
2. Ist jeder Vektor eines zu einem Eigenwert λ einer Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ gehörenden Eigenraumes ein Eigenvektor?
3. Welche (trivialen) Ergebnisse erhält man, wenn man eine Einheitsmatrix auf Eigenwerte untersucht?
4. Gilt für alle Matrizen $\mathbf{A} \in \mathbb{R}^{n \times n}$, daß $\text{rank}(\mathbf{A} - \lambda \mathbf{I}_n) < n$ ist?
5. Warum besitzen eine Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ und ihre Transponierte die selben Eigenwerte.
6. Welche Rolle spielen die Eigenräume einer Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ hinsichtlich der durch $\mathbf{A}$ definierten linearen Abbildung $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$, $f(\mathbf{x}) = \mathbf{A}\mathbf{x}$?
7. Welche für die Eigenwerte einer Matrix notwendigen Bedingungen ergeben Kontrollmöglichkeiten bei der Berechnung der Eigenwerte?
8. Warum ist jede nicht-triviale Linearkombination von k zu einem Eigenwert gehörenden Eigenvektoren $\mathbf{x}_1, \dots, \mathbf{x}_n$ wieder ein zu λ gehörender Eigenvektor?
9. Was folgt für eine Matrix $\mathbf{A} \in \mathbb{R}^n$, wenn Sie (u.a.) den Eigenwert 0 hat.
10. Warum besitzen ähnliche Matrizen gleiche Eigenwerte?
11. Was kann man über $\text{rank}(\mathbf{A} - \lambda_i \mathbf{I}_n)$, $i = 1, \dots, n$ aussagen, wenn $\lambda_1, \dots, \lambda_n$ paarweise verschiedene Eigenwerte von $\mathbf{A} \in \mathbb{R}^{n \times n}$ sind.
Welche Beziehung besteht in diesem Fall zwischen $\mathbb{R}^n$ und den Eigenvektoren von $\mathbf{A}$?
12. Wie kann man mit Hilfe von Eigenwerten zu einer symmetrischen Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ eine ähnliche Diagonalmatrix ermitteln?

7.10 Rechenaufgaben

Aufgabe 17

Berechnen Sie zu folgenden Matrizen die Eigenwerte, geben Sie zu jedem Eigenwert ein Maximalsystem linear unabhängiger Eigenvektoren sowie eine Darstellung des zugehörigen Eigenraumes an. Kontrollieren Sie die berechneten Eigenwerte mit Hilfe der notwendigen Bedingungen für Summen und Produkte der Eigenwerte.

$$(a) \mathbf{A} = \begin{bmatrix} -5 & 2 \\ -3 & 0 \end{bmatrix}$$

$$(c) \mathbf{C} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$(b) \mathbf{B} = \begin{bmatrix} -1 & 6 & -2 \\ -3 & 10 & -3 \\ -1 & 3 & 0 \end{bmatrix}$$

$$(d) \mathbf{D} = \begin{bmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{bmatrix}$$

Aufgabe 18

Welche reellen Eigenwerte besitzen die folgenden Matrizen $\mathbf{A}_i$? Welchen Wert hat jeweils $\text{rank}(\mathbf{A}_i - \lambda I)$ und die Dimension der ggf. existierenden Eigenräume? In welchen Fällen gibt es eine Basis des jeweiligen $\mathbb{R}^n$, die nur aus Eigenvektoren besteht und in welchen Fällen gibt es beliebig viele Vektoren des $\mathbb{R}^n$, die zu keinem Eigenraum gehören?

$$(a) \mathbf{A}_1 = \begin{bmatrix} 6 & 16 \\ -1 & -2 \end{bmatrix}$$

$$(d) \mathbf{A}_4 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$(b) \mathbf{A}_2 = \begin{bmatrix} 2 & -1 \\ 4 & -2 \end{bmatrix}$$

$$(e) \mathbf{A}_5 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$$(c) \mathbf{A}_3 = \begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix}$$

$$(f) \mathbf{A}_6 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

Aufgabe 19

Ermitteln Sie für die folgenden Matrizen die Transformationsmatrix, die diese in eine ähnliche Diagonalmatrix überführt. Führen Sie gegebenenfalls die Transformation durch.

$$(a) \mathbf{A} = \begin{bmatrix} 2 & \sqrt{2} \\ \sqrt{2} & 3 \end{bmatrix}$$

$$(c) \mathbf{C} = \begin{bmatrix} 6 & 16 \\ -1 & -2 \end{bmatrix}$$

$$(b) \mathbf{B} = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 20 & -16 \\ 0 & -16 & 20 \end{bmatrix}$$

$$(d) \mathbf{D} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

$$(e) \mathbf{F} = \begin{bmatrix} -5 & 2 \\ -3 & 0 \end{bmatrix}$$

Aufgabe 20

Zeigen Sie, daß die Matrizen $\mathbf{A}$ und $\mathbf{B}$ ähnlich sind, indem Sie zunächst für beide die Ähnlichkeit zu einer gemeinsamen Diagonalmatrix nachweisen! Ermitteln Sie anschließend eine Matrix $\mathbf{M}$, die die Ähnlichkeit von $\mathbf{A}$ und $\mathbf{B}$ vermittelt $\mathbf{A} = \mathbf{M}^{-1}\mathbf{B}\mathbf{M}$ begründet, wobei

$$\mathbf{A} = \begin{bmatrix} -3 & 1 \\ 0 & -2 \end{bmatrix} \quad \text{und} \quad \mathbf{B} = \begin{bmatrix} -2 & 2 \\ 0 & -3 \end{bmatrix}.$$

Aufgabe 21

Ermitteln Sie eine Matrix $\mathbf{M}$, die die gegebene symmetrische Matrix $\mathbf{S}$ auf Diagonalgestalt transformiert.

$$\mathbf{S} = \begin{bmatrix} 41 & 12 & 0 \\ 12 & 34 & 0 \\ 0 & 0 & 25 \end{bmatrix}$$

Allgemeine Vektorräume

Wir haben uns bis jetzt im wesentlichen mit drei Vektorräumen beschäftigt. Das waren $\mathbb{R}^n$, $\mathbb{R}^{n \times n}$ und $L(\mathbb{R}^n, \mathbb{R}^m)$. Wir haben hier und da $\mathbb{R}$ schon durch $\mathbb{C}$ ersetzt und notiert, daß alles was wir bisher diskutiert haben, weiter funktioniert, wenn man $\mathbb{R}$ durch einen beliebigen Körper $\mathbb{K}$ ersetzt.

In diesem Kapitel werden wir

8.1 Definition und Beispiele

Die Leserin erinnere sich mit einem Blatt Papier und Bleistift an die Definition eines Körpers. Sollte dies Schwierigkeiten bereiten, sei auf Definition 1.10 verwiesen.

Die folgende Definition ist motiviert durch unsere Definitionen von $\mathbb{K}^n$, $\mathbb{K}^{m \times n}$ und $L(\mathbb{K}^n, \mathbb{K}^m)$.

Definition 8.1 (Vektorraum).

Es sei $\mathbb{K}$ ein Körper^a und $V \neq \emptyset$ eine Menge mit zwei binären Operationen

$$\left\{ \begin{array}{l} +: V \times V \rightarrow V, \\ (v_1, v_2) \mapsto v_1 + v_2 \end{array} \right. \quad \text{und} \quad \left\{ \begin{array}{l} \cdot: \mathbb{K} \times V \rightarrow V, \\ (\lambda, v) \mapsto \lambda v \end{array} \right.$$

die die folgenden Axiome erfüllen:

- (i) Es existiert ein eindeutig bestimmtes Element $0 \in V$ (**Nullvektor**/Ursprung) mit $v + 0 = v$ für alle $v \in V$.
- (ii) Für alle $v \in V$ existiert ein $-v \in V$ mit $v + (-v) = 0$.
Dieses Element heißt (additiv) Inverses Element.
- (iii) Für alle $v_1, v_2 \in V$ gilt $v_1 + v_2 = v_2 + v_1$.
- (iv) Für alle $v_1, v_2, v_3 \in V$ gilt $(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$.

(v) Für alle $\lambda, \mu \in \mathbb{K}$ und $v \in V$ gilt $\lambda(\mu v) = (\lambda\mu)v$.

(vi) Für alle $v \in V$ gilt $1v = v$, wobei 1 das neutrale Element von $\mathbb{K}$ ist.

(vii) Für alle $\lambda \in \mathbb{K}$ und $v_1, v_2 \in V$ gilt $\lambda(v_1 + v_2) = \lambda v_1 + \lambda v_2$.

(viii) Für alle $\lambda, \mu \in \mathbb{K}$ und $v \in V$ gilt $(\lambda + \mu)v = \lambda v + \mu v$.

Die Menge V mit den obigen Operationen wird **Vektorraum** über $\mathbb{K}$ genannt und die Elemente **Vektoren**. Die Elemente des Körpers werden **Skalare** genannt.

^atypischerweise $\mathbb{R}$ oder $\mathbb{C}$. Siehe auch Kapitel 2 im Skript Mathematik 1.

Sprachregelung: Die Addition $v_1 + v_2$ zweier Vektoren $v_1, v_2 \in V$ nennen wir **Vektoraddition** um sie von der Addition im Körper zu unterscheiden. Die Multiplikation λv eines Skalars $\lambda \in \mathbb{K}$ mit einem Vektor $v \in V$ wird **Skalarmultiplikation**¹ (skalare Multiplikation) oder **äußeres Produkt** genannt. Wenn $\mathbb{K} = \mathbb{Q}$, dann sagen wir V ist ein **rationaler Vektorraum**, wenn $\mathbb{K} = \mathbb{R}$, dann sagen wir V ist ein **reeller Vektorraum** und wenn $\mathbb{K} = \mathbb{C}$, dann heißt V ein **komplexer Vektorraum**.

Wir diskutieren eine Liste an Beispielen. Die Leserin möge sich durch nachrechnen von der Korrektheit der nachfolgenden Behauptungen überzeugen. Wir sollten klar herausstellen, daß $+$ (und $\cdot$) auf der linken Seite eine andere Bedeutung hat als $+$ (und $\cdot$) auf der rechten Seite. Es wäre sinnvoll bspw. $+$ und $\oplus$ (sowie $\cdot$ und $\odot$) zu verwenden, das führt aber schnell zu überbordender und verwirrender Notation; der Kontext macht stets klar, was die gegebenen Symbole bedeuten sollen. In den Beispielen sollte sich die Leserin dies immer klar machen.

¹Nicht zu verwechseln mit dem Skalarprodukt mit dem Sie vielleicht schon vertraut sind und welches wir später einführen werden.

Beispiel 8.1.

Sei $V = \mathbb{C}$ und $\mathbb{K} = \mathbb{C}$. Die Vektoraddition und die Skalarmultiplikation sind dann durch die Addition komplexer Zahlen und die Multiplikation komplexer Zahlen gegeben. Dann ist V ein komplexer Vektorraum nach den Diskussionen in Kapitel 10 von Mathematik 1.

Allgemeiner kann jeder Körper als Vektorraum über sich selbst aufgefaßt werden.

Beispiel 8.2 (Körper).

Sei $V = \mathbb{C}$ und $\mathbb{K} = \mathbb{R}$. Die Vektoraddition ist durch die übliche Addition komplexer Zahlen definiert und die Skalarmultiplikation über die gewöhnliche Multiplikation komplexer Zahlen eingeschränkt auf die reellen Zahlen ($\mathbb{R} \subseteq \mathbb{C}$). Dann ist $\mathbb{C}$ ein reeller Vektorraum. Die beiden (hier und Beispiel 8.1) Interpretationen von $\mathbb{C}$ sind sehr verschieden. Wir werden das im weiteren noch diskutieren.

Beispiel 8.3 (Polynome).

Es sei $V = \mathcal{P}_2$ die Menge aller Polynome (einer Variable) vom Grade nicht größer als 2 mit komplexen Koeffizienten und inklusive des Nullpolynoms. Die Vektoraddition ist definiert als die normale Addition von Polynomen und die Skalarmultiplikation ist einfach die Multiplikation von Polynomen mit einer komplexen Zahl. Dann ist V ein komplexer Vektorraum.

Allgemeiner sei $V = \mathcal{P}_n$, die Menge aller Polynome (einer Variable) vom Grade nicht größer als n mit rationalen/reellen/komplexen Koeffizienten oder $V = \mathcal{P}$ die Menge aller Polynome (einer Variable) mit rationalen/reellen/komplexen Koeffizienten jeweils inklusive des Nullpolynoms mit den obigen Vereinbarungen und $\mathbb{K}$ sei $\mathbb{Q}/\mathbb{R}/\mathbb{C}$. Dann ist V ein rationaler, reeller oder komplexer Vektorraum.

(Es geht nicht, die Koeffizienten bspw. reell zu wählen und als Körper $\mathbb{C}$. Wieso?)

Beispiel 8.4 (n -Tupel).

Siehe die Diskussion in Kapitel 1. Hier werden wir die Tupel aus Platzgründen in Zeilen schreiben. Sei $n \in \mathbb{N}$. Dann bezeichnet $\mathbb{R}^n$ die Menge aller n -Tupel reeller Zahlen (siehe Kapitel 1 Mathematik 1)

$$\mathbb{R}^n = \{(x_1, \dots, x_n) : x_i \in \mathbb{R}, 1 \leq i \leq n\}.$$

Wenn $x, y \in \mathbb{R}^n = V$ und $\mathbb{K} = \mathbb{R}$, dann definieren wir

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

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

$$0 = (0, \dots, 0),$$

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

Damit ist es nicht schwer die Axiome nachzuprüfen und man erhält einen reellen Vektorraum.

Wenn man $\mathbb{R}$ durch $\mathbb{Q}$ oder $\mathbb{C}$ ersetzt, dann erhält man einen rationalen bzw. komplexen Vektorraum.

Beispiel 8.5 (Funktionen).

Sei $\mathbb{K} \in \{\mathbb{Q}, \mathbb{R}, \mathbb{C}\}$ und $V = \text{Fun}(A, \mathbb{K})$ die Menge der $\mathbb{K}$ -wertigen Funktionen auf einer Menge $A \neq \emptyset$. Dann definieren wir die Vektoraddition punktweise durch

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

für $f, g \in \text{Fun}(A, \mathbb{R})$ wobei auf der rechten Seite die Addition in $\mathbb{K}$ verwendet wird. Auch die Skalarmultiplikation definieren wir punktweise

durch

$$(\lambda, f) \mapsto \lambda f, \quad (\lambda f)(x) = \lambda f(x)$$

für $\lambda \in \mathbb{K}$ und $f \in \text{Fun}(A, \mathbb{K})$ wobei wir auf der rechten Seite die Multiplikation in $\mathbb{K}$ verwenden. Der Nullvektor ist durch die Funktion konstant 0 (also $x \mapsto 0$) gegeben.

Damit ist $\text{Fun}(A, \mathbb{K})$ ein $\mathbb{K}$ -Vektorraum.

Beispiel 8.6 (Stetige Funktionen).

Sei $I \subseteq \mathbb{R}$ ein Intervall und bezeichne $C(I)$ (auch $C^0(I)$) die Menge der stetigen (reellwertigen) Funktionen auf dem Intervall I . Dann ist $V = C(I)$ ein Vektorraum über $\mathbb{R}$. Da es sich bei $f, g \in C(I)$ um Funktionen handelt können wir die Vektoraddition und die Skalarmultiplikation wieder punktweise definieren wie im letzten Beispiel. Nach einem Satz (Rechenregeln stetiger Funktionen) aus Mathematik 1 ist die Summe stetiger Funktionen eine stetige Funktion. Es gilt also $+: C(I) \times C(I) \rightarrow C(I)$. Nach dem gleichen Satz gilt auch für die Skalarmultiplikation $\cdot: \mathbb{R} \times C(I) \rightarrow C(I)$. Der Nullvektor ist die Funktion konstant 0 ($x \mapsto 0$). Die restlichen Axiome lassen sich leicht prüfen und ergeben sich wie in den anderen Beispielen aus den Eigenschaften des Körpers, hier $\mathbb{R}$.

Damit ist $C^0(I)$ ein reeller Vektorraum.

Beispiel 8.7 (Differenzierbare Funktionen).

Es sei $I \subseteq \mathbb{R}$ ein Intervall und $C^k(I)$ bezeichne die Menge aller k mal stetig differenzierbaren Funktionen $f: I \rightarrow \mathbb{R}$ (natürlich könnte man auch komplexwertige Funktionen betrachten).

Nach den entsprechenden Sätzen aus Mathematik 1 gilt für die punktweise definierte Addition $+: C^k(I) \times C^k(I) \rightarrow C^k(I)$ und die punktweise

definierte Skalarmultiplikation $\cdot: \mathbb{R} \times C^k(I) \rightarrow C^k(I)$. Damit ist auch $C^k(I)$ ein reeller Vektorraum.

8.2 Lineare Abhängigkeit

Wir wiederholen Definition 1.3 im allgemeinen Kontext.

Definition 8.2 (Lineare (Un)Abhängigkeit).

Sei $n \in \mathbb{N}$ und V ein $\mathbb{K}$ -Vektorraum. Es sei $\{v_1, \dots, v_n\} \subseteq V$ eine Menge von Vektoren.

- (i) Die Menge $\{v_1, \dots, v_n\}$ heißt genau dann **linear abhängig**, wenn eine korrespondierende Menge $\{\alpha_1, \dots, \alpha_n\} \subseteq \mathbb{K}$ von Skalaren existiert die nicht alle 0 sind, so daß

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

- (ii) Die Menge $\{v_1, \dots, v_n\}$ heißt genau dann **linear unabhängig**, wenn aus

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

stets $\alpha_i = 0$, $i = 1, \dots, n$ folgt.

Offensichtliche Beobachtungen: Wenn eine nichtleere Teilmenge von $\{v_1, \dots, v_n\}$ linear anhängig ist, dann auch $\{v_1, \dots, v_n\}$. Jede Menge $\{v\}$ mit $v \neq 0$ ist linear unabhängig. Die Menge $\{0\}$ ist stets linear abhängig und wenn $0 \in \{v_1, \dots, v_n\}$ dann ist $\{v_1, \dots, v_n\}$ linear abhängig.

Sprachregelung: Lineare Abhängigkeit und Unabhängigkeit sind Eigenschaften von Mengen von Vektoren. Es hat sich jedoch eingebürgert, die Adjektive auch auf die Vektoren anzuwenden und zu sagen $\{v_1, \dots, v_n\}$ sei eine Menge linear (un)abhängiger Vektoren anstatt von $\{v_1, \dots, v_n\}$ sei eine linear unabhängige Menge von Vektoren.

Sprachregelung/Definition: Wenn V ein $\mathbb{K}$ -Vektorraum ist und wir eine unendliche Menge $\{v_i : i \in I\}$ gegeben haben, dann heißt diese linear unabhängig falls jede endliche Teilmenge linear unabhängig ist.

Beispiel 8.8.

Wir fassen $\mathbb{C}$ als Vektorraum über sich selbst auf. siehe Beispiel 8.1. Dann bilden zwei Elemente $x, y \in \mathbb{C}$ stets eine linear abhängige Menge. Wenn $x = 0$ oder $y = 0$, dann ist dies offensichtlich. Wenn beide ungleich Null sind, dann gilt beispielsweise

$$yx + (-x)y = 0.$$

Beispiel 8.9.

Sei $\mathcal{P}_2$ die Menge aller Polynome in einer Variablen vom Grade höchstens 2 inklusive des Nullpolynoms. Nach Beispiel 8.3 ist $\mathcal{P}_2$ ein Vektorraum. Wir betrachten die Polynome

$$p(x) = 1 - x, \quad q(x) = x(1 - x), \quad r(x) = 1 - x^2.$$

Die Menge $\{p, q, r\}$ ist linear abhängig da beispielsweise

$$p + q - r = 0$$

gilt.

Beispiel 8.10.

Sei $\mathcal{P}$ die Menge aller Polynome mit reellen (oder komplexen) Koeffizienten in einer Variablen inklusive des Nullpolynoms. Nach Beispiel 8.3 ist $\mathcal{P}$ ein Vektorraum. Die Menge

$$\{p_n : n \in \mathbb{N}_0\} \subseteq \mathcal{P}, \quad p_n(x) = x^n, n \geq 1 \quad p_0(x) = 1$$

ist linear unabhängig da aus

$$\alpha_0 p_0 + \dots + \alpha_n p_n = 0$$

folgt

$$\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \dots + \alpha_n x^n = 0.$$

(Die α_i sind reell (oder komplex).) Die rechte Seite ist das Nullpolynom und die linke Seite kann nur das Nullpolynom sein, wenn

$$\alpha_0 = \alpha_1 = \dots = \alpha_n = 0$$

gilt.

Definition 8.3 (Unterraum).

Sei V ein Vektorraum über $\mathbb{K}$ und $U \subseteq V$ mit $U \neq \emptyset$. Wenn die Menge U ein Vektorraum über $\mathbb{K}$ ist, dann heißt U ein **Unterraum** von V .

Die Menge aller Unterräume von V bezeichnen wir mit $\text{Sub}(V)$.

Hilfssatz 8.1.

Sei V ein Vektorraum über $\mathbb{K}$ und $U \subseteq V$ mit $U \neq \emptyset$. Dann ist U genau dann ein Unterraum von V , wenn

- (i) Für alle $u_1, u_2 \in U$ gilt $u_1 + u_2 \in U$,
- (ii) Für alle $u \in U$ und $\lambda \in \mathbb{K}$ gilt $\lambda u \in U$.

Wir sagen auch, daß U gegen die Vektorraumoperationen **abgeschlossen** sein muß.

Übungsaufgabe 8.1.

Beweisen Sie Hilfssatz 8.1.

8.3 Linearkombinationen

Definition 8.4 (Linearkombination).

Sei V ein $\mathbb{K}$ -Vektorraum und $\{v_1, \dots, v_n\} \subseteq V$ eine Menge von Vektoren. Dann heißt

$$v = \sum_{i=1}^n \alpha_i v_i, \quad \alpha_i \in \mathbb{K}, i = 1, \dots, n$$

eine **Linearkombination** der Menge $\{v_1, \dots, v_n\}$.

Sprachregelungen: Eine Linearkombination von $\{v_1, \dots, v_n\}$ ist linear abhängig von $\{v_1, \dots, v_n\}$.

Übungsaufgabe 8.2.

In der Situation von Def. 8.4 gilt: Die Menge $\{v_1, \dots, v_n\}$ ist genau dann linear unabhängig, wenn für jede Linearkombination v von $\{v_1, \dots, v_n\}$ die Menge $\{v_1, \dots, v_n, v\}$ linear abhängig ist.

Wir beweisen den ersten wichtigen Satz, ein fundamentales Resultat zur linearen Unabhängigkeit.

Satz 8.1.

Sei V ein Vektorraum und $\{v_1, \dots, v_n\} \subseteq V$ eine Menge von Vektoren. Dann gilt: die Menge $\{v_1, \dots, v_n\}$ ist genau dann linear abhängig, wenn ein v_k , $2 \leq k \leq n$ eine Linearkombination der Vektoren $\{v_1, \dots, v_{k-1}\}$ ist.

Beweis. Wir beweisen die beiden Richtungen separat.

[$\Rightarrow$] Es sei also $\{v_1, \dots, v_n\}$ linear abhängig und $2 \leq k \leq n$ die kleinste Zahl, sodaß $\{v_1, \dots, v_k\}$ linear abhängig ist. (Nach Voraussetzung gilt dies im schlimmsten Fall für $k = n$, es gibt also immer ein solches k .) Es gibt dann also Zahlen $\{\alpha_1, \dots, \alpha_k\}$ nicht alle 0 mit

$$\alpha v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k = 0.$$

Es gilt nun $\alpha_k \neq 0$, da ansonsten schon $\{v_1, \dots, v_{k-1}\}$ linear abhängig wären. Dann können wir schreiben

$$v_k = -\frac{\alpha_1}{\alpha_k}v_1 - \frac{\alpha_2}{\alpha_k}v_2 - \dots - \frac{\alpha_{k-1}}{\alpha_k}v_{k-1}$$

wie gewünscht.

[$\Leftarrow$] Ausführlich: Angenommen es gibt ein $2 \leq k \leq n$ mit

$$v_k = \alpha_1 v_1 + \dots + \alpha_{k-1} v_{k-1}.$$

Dann gilt natürlich

$$\alpha_1 v_1 + \dots + \alpha_{k-1} v_{k-1} - v_k = 0$$

und damit ist $\{v_1, \dots, v_k\} \subseteq \{v_1, \dots, v_n\}$ linear abhängig. Damit ist auch $\{v_1, \dots, v_n\}$ linear abhängig. Siehe Sprachregelung nach der Definition von Linearkombination.

□

8.4 Basen

Definition 8.5 (Lineare Hülle).

Es sei V ein Vektorraum über $\mathbb{K}$ und $\{v_1, \dots, v_n\} \subseteq V$ eine Menge von Vektoren. Dann bezeichnen wir die Menge $\text{span}(\{v_1, \dots, v_n\})$ aller endlichen Linearkombinationen der Vektoren aus $\{v_1, \dots, v_n\}$ als den **Spann** bzw. die **lineare Hülle** von $\{v_1, \dots, v_n\}$. Genauer:

$$\begin{aligned} \text{span}(\{v_1, \dots, v_n\}) &= \left\{ \sum_{i=1}^n \alpha_i v_i : \alpha_i \in \mathbb{K}, 1 \leq i \leq n \right\} \\ &= \bigcap_{\substack{U \in \text{Sub}(V) \\ \{v_1, \dots, v_n\} \subseteq U}} U \end{aligned}$$

Wenn $\{v_i : i \in I\} \subseteq V$ für eine beliebige Indexmenge I , dann

$$\begin{aligned} \text{span}(\{v_i : i \in I\}) &= \bigcup_{\substack{J \subseteq I \\ J \text{ endlich}}} \text{span}(\{v_i : i \in J\}) \\ &= \bigcap_{\substack{U \in \text{Sub}(V) \\ \{v_i : i \in I\} \subseteq U}} U \end{aligned}$$

Definition 8.6 (Erzeugendensystem).

Ein **Erzeugendensystem** eines Vektorraumes V ist eine Menge $E \subseteq V$ mit der Eigenschaft, daß alle $v \in V$ eine (endliche) Linearkombination von Elementen in E ist.

Definition 8.7 (Basis).

Eine **Basis** (oder Koordinatensystem) in einem Vektorraum V ist eine linear unabhängiges Erzeugendensystem $B \subseteq V$ von V .

In anderen Worten: eine Menge $B \subseteq V$ mit der Eigenschaft, daß alle $v \in V$ eine eindeutige (endliche) Linearkombination von Elementen in B ist.

Wenn B eine endliche Menge ist, dann heißt V **endlichdimensional**.

Das nächste Resultat ist grundlegend für viele Diskussionen in der linearen Algebra. Es sagt, daß jede linear unabhängige Menge in einem (endlichdimensionalen) Vektorraum zu einer Basis erweitert werden kann.

Satz 8.2.

Sei V ein endlichdimensionaler Vektorraum und $\{v_1, \dots, v_m\}$ sei eine beliebige Menge unabhängiger Vektoren in V . Dann ist $\{v_1, \dots, v_m\}$ entweder eine Basis oder es existieren Vektoren $v_{m+1}, \dots, v_{m+k}$, sodaß die Menge $\{v_1, \dots, v_m, v_{m+1}, \dots, v_{m+k}\}$ eine Basis von V ist.

Beweis. Da V endlichdimensional ist, gibt es eine endliche Basis $\{b_1, \dots, b_n\}$. Wir betrachten die Menge

$$M = \{v_1, \dots, v_m, b_1, \dots, b_n\}.$$

Jeder Vektor in V ist eine Linearkombination von Elementen in M . Wir wenden nun Satz 8.1 an um eine Basis daraus zu machen: Die Menge M ist linear abhängig, da die $v_1, \dots, v_m$ jeweils linear abhängig von $\{b_1, \dots, b_n\}$ sind. Damit sind einige der Vektoren in M Linearkombinationen der vorangehenden. Sei z der erste solche Vektor. Da die Menge $\{v_1, \dots, v_m\}$ linear unabhängig ist, ist $z = b_i$ für ein $i = 1, \dots, n$. Damit können wir die Menge

$$M' = \{v_1, \dots, v_m, b_1, \dots, b_{i-1}, b_{i+1}, \dots, b_n\}$$

Betrachten. Jeder Vektor von V ist noch immer eine Linearkombination der Elemente von M' . Wenn M' linear unabhängig ist, dann sind wir fertig, wenn nicht, dann verfahren wir nochmal wie zuvor. Dieses Verfahren bricht nach endlich vielen Schritten mit einer Basis ab. □

8.5 Dimension

Satz 8.3.

Es sei V ein Endlich dimensionaler Vektorraum. Dann ist die Anzahl der Elemente aller Basen gleich.

Beweis. Der Beweis funktioniert ähnlich wie der von Satz 8.1. Es seien

- $\mathcal{V} = \{v_1, \dots, v_n\} \subseteq V$ eine Menge von Vektoren, sodaß jeder Vektor in V eine Linearkombination von $\mathcal{V}$ ist und
- $\mathcal{W} = \{w_1, \dots, w_m\} \subseteq V$ eine Menge von linear unabhängigen Vektoren.

Wir bilden die Menge

$$M = \{w_m, v_1, \dots, v_n\}.$$

Jeder Vektor in V ist eine Linearkombination von Elementen aus M aber die Menge M ist linear abhängig da w_m eine Linearkombination von $\{v_1, \dots, v_n\}$ ist. Damit gibt es, nach Satz 8.1 ein $1 \leq i \leq n$, mit der Eigenschaft, daß v_i eine Linearkombination von $\{w_m, v_1, \dots, v_n\}$ ist. Wir bilden dann

$$M' = \{w_m, v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n\}.$$

Für diese Menge gilt noch immer, daß, jeder Vektor aus V durch eine Linearkombination von Elementen aus M' dargestellt werden kann. Wir bilden nun eine neue Menge

$$M = \{w_{m-1}, w_m, v_{i-1}, v_{i+1}, \dots, v_n\}.$$

und Verfahren wie zuvor. Es kann nicht sein, daß wir die v 's vor den w 's ausschöpfen, also $n \geq m$. Angenommen $n < m$. Dann besteht M' nur aus w 's und jeder Vektor in V kann als Linearkombination in M' dargestellt werden. Folglich sind die verbleibenden w 's linear abhängig von M' . Da $\mathcal{W}$ aber linear unabhängig ist, kann das nicht sein. Wir drücken das noch einmal anders aus: Wir haben nach m Schritten eine Menge M' mit den gleichen Eigenschaft wie $\mathcal{V}$ aber mit m der vs ersetzt durch ws .

Wenn nun $\mathcal{V}$ und $\mathcal{W}$ beide Basen sind, dann gilt damit außerdem $m \leq n$ und damit $m = n$. □

Dieser Satz erlaubt es uns die folgende Definition anzugeben:

Definition 8.8 (Dimension).

Sei V ein endlichdimensionaler Vektorraum. Die Dimension von V , in Zeichen $\dim(V)$, ist die Anzahl der Elemente in einer Basis von V .

8.6 Lineare Abbildungen

8.7 Übungsaufgaben

Aufgabe 22

Es sei V ein Vektorraum. Benutzen Sie nur die Axiome des Vektorraumes und die des Skalar-Körpers um die folgenden Aussagen zu zeigen: $v \in V, \alpha \in \mathbb{K}$

- | | |
|------------------------|---|
| (i) $0 + v = v$, | (v) Wenn $\alpha v = 0$, dann gilt $\alpha = 0$
oder $v = 0$. |
| (ii) $-0 = 0$, | (vi) $-x = (-1)x$, |
| (iii) $\alpha 0 = 0$, | (vii) $w + (v - w) = v, w \in V$ (wir
setzen $v - w = v + (-w)$). |
| (iv) $0v = 0^2$. | |

²bemerke, daß die 0 links ein Vektor ist und rechts ein Skalar

Fourier-Reihen

Fourier-Reihen sind nach dem französischen Mathematiker, Physiker, Ägyptologen und Revolutionär Joseph Fourier (1768–1830) benannt. Die **Fourier-Analyse** ist ein Grundpfeiler der angewandten Mathematik, insbesondere der Signal-, Bild- und allgemein Datenverarbeitung. Neben vielfältigen physikalischen Entdeckungen (Wärmelehre) ist Fourier für seine Entdeckung bekannt, dass – selbst unstetige – Funktionen als Überlagerung (Linearkombination, konvergente Reihe) der Funktionen

$$1, \quad \cos(kx), \quad \sin(kx) \quad (k \in \mathbb{N}).$$



Abbildung 9.1: Joseph Fourier

In vielen Anwendungen sind die Koeffizienten für große k vernachlässigbar klein. Dies gestattet es, Signale in ihre Wellenanteile zu zerlegen, und sie dann zu übermitteln, komprimieren, filtern, entauschen, analysieren, klassifizieren oder verschlüsseln. Die strenge mathematische Begründung der Konvergenz von Fourier-Reihen dauerte bis in die 1960er Jahre (Carleson-Hunt-Theorem).

9.1 Definition

Wir betrachten Reihendarstellungen für periodische Funktionen, zunächst speziell für 2π -periodische Funktionen¹. Als Reihenglieder verwenden wir Funktionen, die selbst 2π -periodisch sind:

Definition 9.1 (Fourier-Reihen).

Eine Reihe der Bauart

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} [a_k \cos(kx) + b_k \sin(kx)] \quad (9.1.1)$$

mit reellen Folgen $(a_k)_{k \in \mathbb{N}_0}$ und $(b_k)_{k \in \mathbb{N}}$ heißt **trigonometrische Reihe**.

Sind alle $a_k = 0$, spricht man von einer **Sinusreihe**.

Sind alle $b_k = 0$, spricht man von einer **Kosinusreihe**.

Eine Teilsumme (bis zum Index $k = n$) heißt **trigonometrisches Polynom** (vom Grade n).

Ist $f: \mathbb{R} \rightarrow \mathbb{R}$ periodisch mit Periode T , bestimmt man die Kreisfrequenz

$$\omega := \frac{2\pi}{T}$$

und setzt

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} [a_k \cos(k\omega x) + b_k \sin(k\omega x)],$$

wobei

$$\begin{aligned} a_0 &= \frac{2}{T} \int_s^{s+T} f(t) \, dt, \\ a_k &= \frac{2}{T} \int_s^{s+T} f(t) \cos(k\omega t) \, dt, \\ b_k &= \frac{2}{T} \int_s^{s+T} f(t) \sin(k\omega t) \, dt \quad (k \in \mathbb{N}). \end{aligned}$$

¹Das bedeutet $f(x + 2\pi) = f(x)$ für alle $x \in \mathbb{R}$.

Man nennt ω die **Kreisfrequenz der Grundschiwingung** und $k\omega$ ($k > 1$) die **Kreisfrequenzen** der harmonischen Oberschwingungen. Die Zahl $s \in \mathbb{R}$ ist beliebig.

9.2 Konvergenz

Wir gehen zunächst von einer gegebenen trigonometrischen Reihe aus und bemerken:

Satz 9.1.

Ist die trigonometrische Reihe (9.1.1) für alle $x \in \mathbb{R}$ konvergent, so ist

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} [a_k \cos(kx) + b_k \sin(kx)] \quad (9.2.1)$$

eine auf $\mathbb{R}$ definierte 2π -periodische Funktion.

Es ist also umgekehrt nur für 2π -periodische Funktionen sinnvoll, nach Darstellungen der Form (9.2.1) zu suchen. Für Funktionen wie $g(x) = x^2$ oder $h(x) = e^x$ ist dies dagegen zwecklos (es sei denn, man betrachtet die periodische Fortsetzung eines endlichen Abschnitts solcher nichtperiodischer Funktionen).

Der folgende Satz liefert u. a. den Schlüssel zur Berechnung der gesuchten Reihendarstellungen:

Satz 9.2.

Sind die Reihen $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=1}^{\infty} b_k$ absolut konvergent, so konvergiert die trigonometrische Reihe (9.1.1) punktweise auf ganz $\mathbb{R}$. Die Summenfunktion

$$f(x) := \frac{a_0}{2} + \sum_{k=1}^{\infty} [a_k \cos(kx) + b_k \sin(kx)]$$

ist dann stetig auf $\mathbb{R}$, und es gelten

$$\begin{aligned} a_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(kt) \quad (k \in \mathbb{N}_0), \\ b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin(kt) \quad (k \in \mathbb{N}). \end{aligned} \tag{9.2.2}$$

Um die Darstellungen in (9.2.2) zu erhalten, versieht man den Raum $L^2(-\pi, \pi)$ aller Funktionen f mit $\int_{-\pi}^{\pi} |f(x)|^2 x < \infty$ mit dem Skalarprodukt

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)g(x) x. \tag{9.2.3}$$

Es lässt sich zeigen, dass die Funktionen $\sin(nx)$ und $\cos(mx)$ ($m \in \mathbb{N}_0, n \in \mathbb{N}$) diesbezüglich ein Orthonormalsystem bilden, d.h.

$$\begin{aligned} \frac{1}{\pi} \int_{-\pi}^{\pi} \sin(nx) \sin(mx) x &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos(nx) \cos(mx) x \\ &= \begin{cases} 0, & \text{falls } n \neq m, \\ 1, & \text{falls } n = m, \end{cases} \end{aligned}$$

und

$$\frac{1}{\pi} \int_{-\pi}^{\pi} \sin(nx) \cos(mx) x = 0 \quad (n \in \mathbb{N}, m \in \mathbb{N}_0).$$

Die Darstellung (9.2.2) der Koeffizienten ergibt sich damit wie in Satz 3.29 aus

$$a_k = \langle f(\cdot), \cos(k\cdot) \rangle \quad \text{und} \quad b_k = \langle f(\cdot), \sin(k\cdot) \rangle.$$

Dieses Orthonormalsystem ist darüber hinaus **vollständig**, d.h. nur die Nullfunktion ist orthogonal zu allen darin enthaltenen Funktionen. Ein solches

vollständiges Orthonormalsystem heißt auch **Orthonormalbasis** der Hilbert-Raumes $L^2(-\pi, \pi)$.

Bemerkung 9.1.

Beachte die Analogie zu Definition 5.8.

Bemerkung 9.2.

Auch wenn es üblich ist ein vollständiges ONS wie oben beschrieben eine Basis zu nennen ist dies jedoch keine Basis im eigentlichen Sinne. Dazu müßte jedes Element mittels einer endlichen Linearkombination darstellbar sein.

9.3 Welche Funktionen sind entwickelbar?

Wir untersuchen nun für eine vorgegebene 2π -periodische Funktion f , ob sie sich in eine trigonometrische Reihe entwickeln lässt.

Definition 9.2.

Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine 2π -periodische, auf $[-\pi, \pi]$ integrierbare Funktion. Dann heißen die Zahlen

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(kt) \quad (k \in \mathbb{N}_0),$$

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin(kt) \quad (k \in \mathbb{N}),$$

Fourier-Koeffizienten von f . Die Reihe

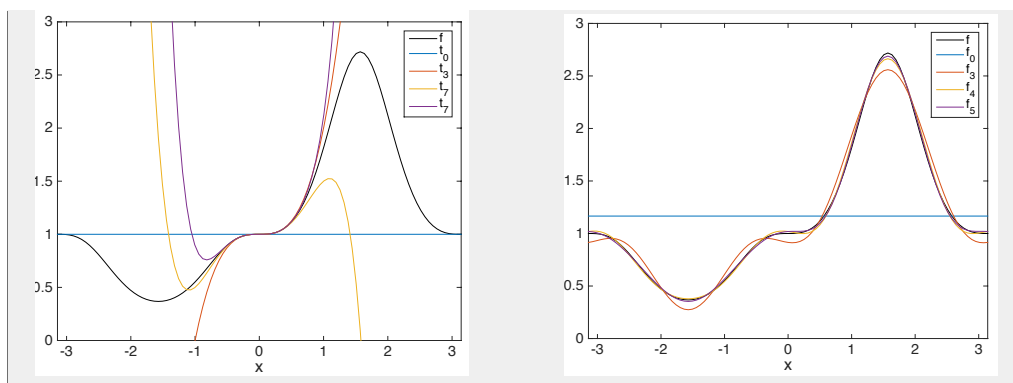
$$R_f(x) := \frac{a_0}{2} + \sum_{k=1}^{\infty} [a_k \cos(kx) + b_k \sin(kx)]$$

heißt **Fourier-Reihe** von f .

Beispiel 9.1.

Wir vergleichen die Taylor-Entwicklung an der Stelle $x_0 = 0$ (links) mit der Fourier-Reihe (rechts) der (2π -periodischen) Funktion

$$f: (-\pi, \pi) \rightarrow \mathbb{R}, \quad f(x) = \exp(\sin^3 x).$$



Da die Integranden die Periode 2π besitzen, kann auch jedes andere Intervall der Länge 2π als Integrationsbereich verwendet werden. Das konkrete Rechnen erleichtert häufig:

Satz 9.3.

Ist $f : \mathbb{R} \rightarrow \mathbb{R}$ eine 2π -periodische, in $[-\pi, \pi]$ integrierbare, und auf $(-\pi, \pi)$ gerade [ungerade] Funktion, dann ist die Fourier-Reihe von f eine Kosinusreihe [eine Sinusreihe].

Bemerkung 9.3.

Man berechne die Fourier-Reihe zum **Rechteckpuls**

$$f(x) = \begin{cases} A, & \text{für } |x| \leq \pi/2, \\ 0, & \text{für } \pi/2 \leq |x| \leq \pi. \end{cases}$$

Wie bei den Taylor-Reihen stellen sich nun folgende Fragen:

- Wann konvergiert die Fourier-Reihe $R_f(x)$?
- Falls sie konvergiert, unter welchen Bedingungen gilt dann auch $R_f(x) = f(x)$?

Zur Beantwortung brauchen wir einen weiteren Begriff:

Definition 9.3.

Eine Funktion $f: [a, b] \rightarrow \mathbb{R}$ heißt auf $[a, b]$ **stückweise glatt**, wenn es eine Partition

$$a = x_0 < x_1 < \dots < x_n = b$$

von $[a, b]$ gibt, so dass f auf jedem der Teilintervalle $[x_{i-1}, x_i]$ stetig differenzierbar ist.

Satz 9.4.

Ist die 2π -periodische Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ stückweise glatt auf $[-\pi, \pi]$, so konvergiert ihre Fourier-Reihe R_f punktweise auf $\mathbb{R}$. Dabei gilt

$$R_f(x_0) = \frac{1}{2} \left[\lim_{x \rightarrow x_0^-} f(x) + \lim_{x \rightarrow x_0^+} f(x) \right] \quad \text{für alle } x_0 \in \mathbb{R}.$$

Ist f stetig in x_0 , so folgt insbesondere $R_f(x_0) = f(x_0)$.

9.4 Das Gibbs-Phänomen

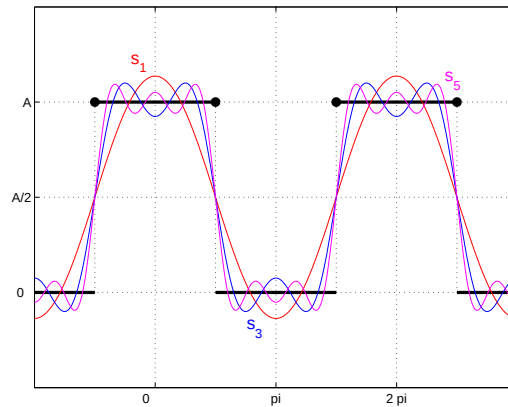


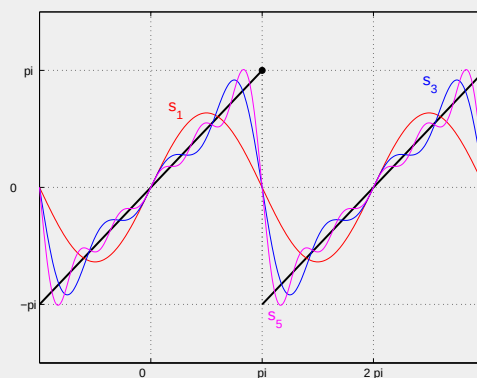
Abbildung 9.2: Teilsummen der Fourierentwicklung zum Rechteckpuls, vgl. Bsp. S. 9.3.

Beispiel 9.2.

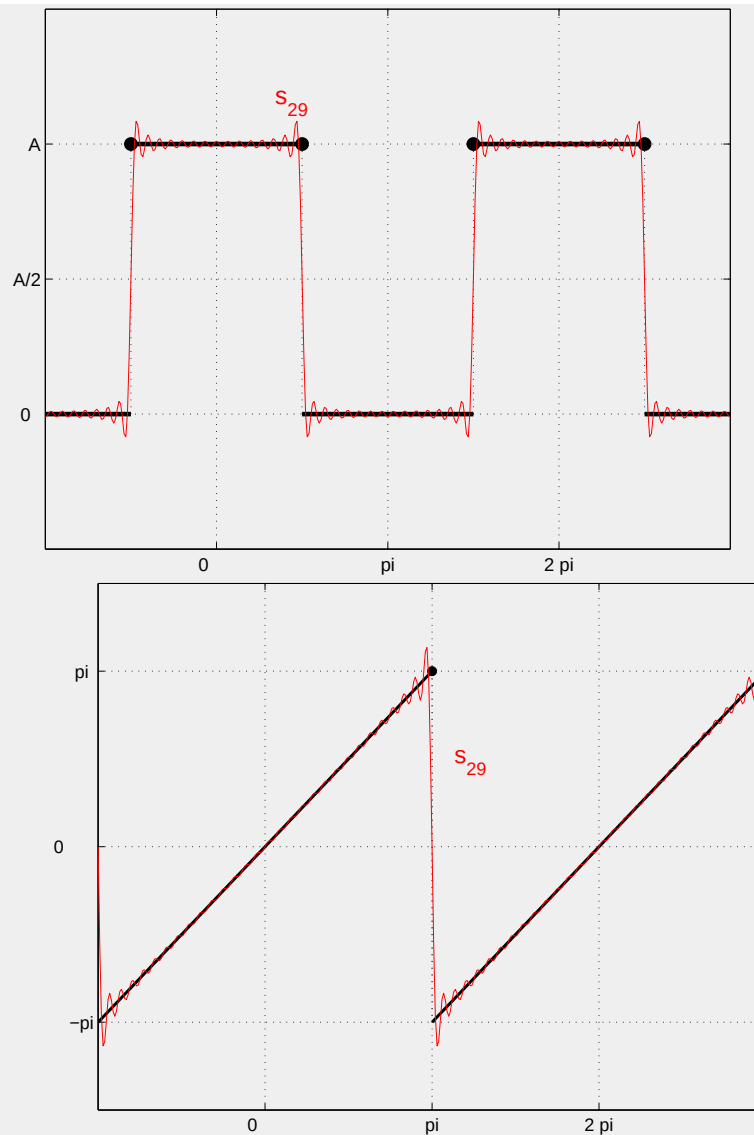
Für die **Sägezahnfunktion** $f(x) = x$ ($|x| \leq \pi$) ergibt sich

$$a_k = 0 \quad (k \in \mathbb{N}_0) \quad \text{sowie} \quad b_k = 2 \frac{(-1)^{k+1}}{k} \quad (k \in \mathbb{N}).$$

Auch hier stellen wir die ersten Teilsummen dar:



Wir betrachten die Teilsummen s_{29} für Rechteckpuls und Sägezahnfunktion:



In einer kleinen Umgebung der Sprungstelle 'überschwingen' die Partialsummen s_n um etwa 9 % der Sprunghöhe ('overshoot'). Dieses **Gibbsche Phänomen**^a verschwindet nicht für $n \rightarrow \infty$, bewegt sich aber näher an die Sprungstelle.

^aJosiah Willard Gibbs, 1839-1903, US-amerikanischer Physiker

Beispiel 9.3.

Ein typisches Problem, welches durch Überspringen verursacht wird, sind Artefakte im JPG-Bildformat in der Nähe scharfer Kanten.



Grund ist u. a. die Verwendung einer Kosinustransformation im Kompressionsalgorithmus, die ganz ähnliche Eigenschaften wie die Fouriertransformation aufweist. Insbesondere für qualitativ hochwertige Balkengrafiken und Diagramme ist JPG daher ein denkbar ungeeignetes Format.

9.5 Bessel und Parsevall

Zur Wiederholung. Der Raum $L^2(-\pi, \pi)$ ist der Vektorraum aller Funktionen $f: (-\pi, \pi)$, für die gilt

$$\int_{-\pi}^{\pi} |f(t)|^2.$$

Satz 9.5.

Für jede Funktion $f \in L^2(-\pi, \pi)$ mit den Fourier-Koeffizienten $\{a_k\}_{k \in \mathbb{N}_0}$ und $\{b_k\}_{k \in \mathbb{N}}$ gilt die **Besselsche Ungleichung**

$$\frac{a_0^2}{2} + \sum_{k=1}^n (a_k^2 + b_k^2) \leq \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)^2 x, \quad n \in \mathbb{N}_0. \quad (9.5.1)$$

Aus der Vollständigkeit unserer Orthonormalbasis $\{\cos(kx), k \in \mathbb{N}_0; \sin(kx), k \in \mathbb{N}\}$ folgt aus der Besselschen Ungleichung für $n \rightarrow \infty$.

Satz 9.6.

Für jede Funktion $f \in L^2(-\pi, \pi)$ mit den Fourier-Koeffizienten $\{a_k\}_{k \in \mathbb{N}_0}$ und $\{b_k\}_{k \in \mathbb{N}}$ gilt die **Parsevalsche Gleichung**

$$\frac{a_0^2}{2} + \sum_{k=1}^{\infty} (a_k^2 + b_k^2) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)^2 x, \quad n \in \mathbb{N}_0. \quad (9.5.2)$$

Bemerkung 9.4.

Parsevall ist im wesentlichen ein unendlichdimensionaler Satz des Pythagoras. Dies wird offensichtlicher in der komplexen Darstellung. Siehe Satz 9.13.

Gleichung (9.5.2) (eigentlich bereits (9.5.1)) zeigt, dass die Fourier-Koeffizienten von L^2 -Funktionen Nullfolgen bilden.

Satz 9.7.

Die Fourier-Reihe einer stetigen, stückweise glatten 2π -periodischen Funktion f konvergiert gleichmäßig und absolut gegen f . Für ihre Fourier-Koeffizienten $\{a_k\}_{k \in \mathbb{N}_0}$ und $\{b_k\}_{k \in \mathbb{N}}$ konvergieren ferner die Reihen

$$\sum_{k=0}^{\infty} |a_k|, \quad \text{und} \quad \sum_{k=1}^{\infty} |b_k|.$$

Mit der Definition der **Supremumsnorm** auf einem reellen Intervall I

$$\|f\|_{\infty} := \sup_{x \in I} |f(x)|$$

und der Bezeichnung $R_f^n(x)$ für die n -te Teilsumme der Fourier-Reihe einer Funktion f lässt sich die gleichmäßige Konvergenz der Fourier-Reihe gegen f ausdrücken als

$$\|f - R_f^n\|_{\infty} \rightarrow 0 \quad \text{mit} \quad n \rightarrow \infty.$$

9.6 Konvergenz im quadratischen Mittel

Eine weitere Norm für Funktionen auf $[-\pi, \pi]$ ist die zum Skalarprodukt (9.2.3) gehörende L^2 -**Norm**

$$\|f\|_2 := \sqrt{\langle f, f \rangle} = \left(\frac{1}{\pi} \int_{-\pi}^{\pi} f(x)^2 dx \right)^{1/2}.$$

Konvergenz in dieser Norm bezeichnet man als **Konvergenz im quadratischen Mittel**.

Satz 9.8.

Die Fourier-Reihe einer Funktion $f \in L^2(-\pi, \pi)$ konvergiert im quadratischen Mittel gegen f .

9.7 Optimalitätseigenschaft trig. Polynome

Die Teilsummen R_f^n der Fourier-Reihe einer Funktion f bilden ein trigonometrisches Polynom vom Grad n . Jede dieser Teilsummen besitzt die Optimalitätseigenschaft, daß sie die unter allen trigonometrischen Polynomen von Grad n

$$T_n(x) = \frac{\tilde{a}_0}{2} + \sum_{k=1}^n [\tilde{a}_k \cos(kx) + \tilde{b}_k \sin(kx)]$$

die Funktion f in der L^2 -Norm am besten approximieren:

Satz 9.9.

Für jedes $n \in \mathbb{N}_0$ wird der Quadratmittelfehler

$$\|f - T_n\|_2^2 = \frac{1}{\pi} \int_{-\pi}^{\pi} [f(x) - T_n(x)]^2 x$$

genau dann minimal, wenn $T_n = R_f^n$. Ferner gilt

$$\|f - R_f^n\|_2^2 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)^2 x - \left[\frac{a_0^2}{2} + \sum_{k=1}^n (a_k^2 + b_k^2) \right].$$

9.7.1 Gliedweise Integrierbarkeit und Differenzierbarkeit

Die Fourier-Reihe einer stetigen, stückweise glatten periodischen Funktion ist gliedweise integrierbar (gleichmäßige Konvergenz). Für die gliedweise Differenzierbarkeit muss auch die Reihe der Ableitungen gleichmäßig konvergieren. Dies ist oft nicht erfüllt, etwa bei der Modellierung unstetiger oder nichtdifferenzierbarer periodischer Vorgänge.

Wann kann eine Fourier-Reihe dennoch gliedweise integriert/differenziert werden? Für die Integration gilt

Satz 9.10.

Eine punktweise konvergente Fourier-Reihe $R(x)$ kann gliedweise integriert werden und es gilt

$$\begin{aligned} F(x) &:= \int_0^x R(t) \\ &= \frac{a_0}{2}x + \sum_{k=1}^{\infty} \left[\frac{a_k}{k} \sin(kx) - \frac{b_k}{k} \cos(kx) \right] + \sum_{k=1}^{\infty} \frac{b_k}{k}, \end{aligned}$$

wobei die Reihe gleichmäßig für alle $x \in \mathbb{R}$ gegen $F(x)$ konvergiert.

Beispiel 9.4.

Ungerade fortgesetzte 2-periodische Funktion

$$f(x) = \begin{cases} x - 1, & 0 < x < 2, \\ 0, & x = 0. \end{cases}$$

Fourier-Reihe

$$R_f(x) = -\frac{2}{\pi} \sum_{k=1}^{\infty} \frac{\sin(k\pi x)}{k}$$

konvergiert an allen Stetigkeitsstellen, also insbesondere in $(0, 2)$, punkt-

weise gegen $f(x) = x - 1$. Die Ableitungsreihe

$$R_{f'}(x) = -2 \sum_{k=1}^{\infty} \cos(k\pi x)$$

divergiert an der Stelle $x = 1$, obwohl f dort stetig, sogar differenzierbar ist.

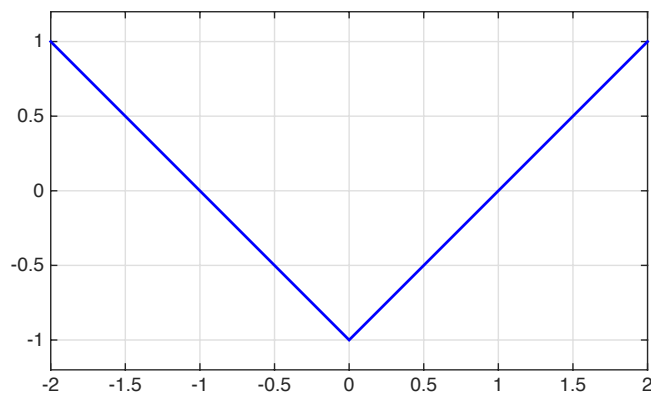
Beispiel 9.5.

Gerade und stetig fortgesetzte 4-periodische Funktion

$$\tilde{f}: [-2, 2] \rightarrow \mathbb{R}, \quad \tilde{f}(x) = |x| - 1$$

mit Fourier-Reihe

$$R_{\tilde{f}}(x) = \sum_{k=1}^{\infty} \frac{(-1)^k - 1}{\left(\frac{k\pi}{2}\right)^2} \cos\left(\frac{\pi}{2}kx\right)$$



Die Ableitungsreihe

$$R_{\tilde{f}'}(x) = \frac{4}{\pi} \sum_{k=1}^{\infty} \frac{\sin\left((2k-1)\frac{\pi}{2}x\right)}{2k-1}$$

konvergiert (Leibniz-Kriterium) an der Stelle $x = 1$ gegen $\tilde{f}'(1) = 1$.

Satz 9.11.

Eine punktweise konvergente Fourier-Reihe, die eine Funktion f darstellt, kann man nur dann gliedweise an einer Stelle x differenzieren, wenn die Ableitungsreihe im Punkt x konvergent ist. Im Fall der Konvergenz stellt die Ableitungsreihe $f'(x)$ dar. Hinreichend für die Konvergenz der Ableitungsreihe ist die Stetigkeit und die stückweise stetige Differenzierbarkeit von f .

9.8 Komplexe Darstellung reeller Fourier-Reihen

Periodische Funktionen mit Werten in den komplexen Zahlen lassen sich ebenfalls in einer Fourier-Reihe entwickeln. Aufgrund der Beziehung zwischen der Sinus-, Kosinus- und Exponentialfunktion besitzt die komplexe Schreibweise sogar eine einfachere Form als jeweils eine reelle Fourier-Reihe für Real- und Imaginärteil.

Besitzt die stückweise glatte 2π -periodische Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ die Fourier-Reihe

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} [a_k \cos(kx) + b_k \sin(kx)],$$

so lassen sich Sinus und Kosinus vermöge der Eulerschen Formel ausdrücken durch

$$\cos(kx) = \frac{e^{ikx} + e^{-ikx}}{2}, \quad \sin(kx) = \frac{e^{ikx} - e^{-ikx}}{2i},$$

und wir erhalten nach Einsetzen

$$\begin{aligned} f(x) &= \frac{a_0}{2} + \sum_{k=1}^{\infty} \left[a_k \frac{e^{ikx} + e^{-ikx}}{2} + b_k \frac{e^{ikx} - e^{-ikx}}{2i} \right] \\ &= \frac{a_0}{2} + \sum_{k=1}^{\infty} \left[\frac{a_k - ib_k}{2} e^{ikx} + \frac{a_k + ib_k}{2} e^{-ikx} \right]. \end{aligned}$$

Wir setzen nun

$$b_0 := 0, \quad a_{-k} := a_k, \quad b_{-k} := -b_k, \quad k \in \mathbb{N}, \quad (9.8.1)$$

sowie $c_k := (a_k - ib_k)/2$, $k \in \mathbb{Z}$, und erhalten

$$f(x) = c_0 + \sum_{k=1}^{\infty} c_k e^{ikx} + c_{-k} e^{-ikx} = \sum_{k=-\infty}^{\infty} c_k e^{ikx}, \quad (9.8.2)$$

wobei der Grenzwert als $\lim_{n \rightarrow \infty} \sum_{k=-n}^n c_k e^{-ikx}$ zu verstehen ist.

Durch Multiplikation von (9.8.2) mit e^{-inx} , $n \in \mathbb{Z}$, Integration über $[-\pi, \pi]$, Vertauschung von Integration und Summation erhält man die komplexe Darstellung der Fourier-Koeffizienten

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx, \quad k \in \mathbb{Z}. \quad (9.8.3)$$

Formel (9.8.3) gilt unter denselben Voraussetzungen wie die entsprechenden Formeln für a_k und b_k . Für die Rückrechnung erhalten wir

$$a_k = 2 \operatorname{Re} c_k, \quad b_k = -2 \operatorname{Im} c_k \quad k \in \mathbb{N}_0.$$

An (9.8.1) erkennt man sofort, dass für reelle Funktionen f gilt $c_k = \overline{c_{-k}}$. Die Konvergenzsätze 9.4 und 9.7 gelten unverändert für die komplexe Darstellung.

Oft ist es praktischer, auch bei der Modellierung reellwertigen periodischer Funktionen $f = f(t)$ direkt die komplexe Darstellung

$$f(t) = \sum_{k=-\infty}^{\infty} c_k e^{ik\omega t}$$

mit einer Kreisfrequenz $\omega > 0$ anzusetzen. So lassen sich etwa die Fourier-Reihe einer phasenverschobene Schwingungen $g(t) = f(t - t_0)$ leicht darstellen als

$$g(t) = f(t - t_0) = \sum_{k=-\infty}^{\infty} c_k e^{ik\omega(t-t_0)} = \sum_{k=-\infty}^{\infty} \underbrace{[c_k e^{-ik\omega t_0}]}_{=: \tilde{c}_k} e^{ik\omega t},$$

was mit der trigonometrischen Variante deutlich umständlicher ginge.

Außer bei der Feststellung, dass bei reellen Funktionen die komplexen Fourier-Koeffizienten $c_k = \overline{c_{-k}}$ erfüllen, wurde bisher an keiner Stelle verwendet, dass die betrachteten Funktionen reellwertig sind. Wir können daher viele der hergeleiteten Ergebnisse auf periodische Funktionen

$$f: \mathbb{R} \rightarrow \mathbb{C}$$

übertragen.

Bei den Integralformeln für die Koeffizienten c_k ist lediglich zu beachten, dass Real- und Imaginärteile für sich integriert werden, d.h.

$$\int f(t) = \int \operatorname{Re} f(t) + i \int \operatorname{Im} f(t)$$

Satz 9.12 (Rechenregeln).

Sind $f, g: \mathbb{R} \rightarrow \mathbb{C}$ zwei T -periodische, stückweise glatte Funktionen mit den Fourier-Reihen $f(t) = \sum_{k=-\infty}^{\infty} f_k e^{ik\omega t}$ und $g(t) = \sum_{k=-\infty}^{\infty} g_k e^{ik\omega t}$ mit $\omega = 2\pi/T$, so gelten

(1) Linearität

$$\alpha f + \beta g = \sum_{k=-\infty}^{\infty} (\alpha f_k + \beta g_k) e^{ik\omega t}, \quad \alpha, \beta \in \mathbb{C}$$

(2) Konjugation:

$$\overline{f(t)} = \sum_{k=-\infty}^{\infty} \overline{f_{-k}} e^{ik\omega t},$$

(3) Zeitumkehr:

$$f(-t) = \sum_{k=-\infty}^{\infty} f_{-k} e^{ik\omega t}$$

(4) Streckung, Ähnlichkeit:

$$f(\alpha t) = \sum_{k=-\infty}^{\infty} f_k e^{ik\alpha\omega t}$$

(5) Translation, Phasenverschiebung:

$$f(t + \tau) = \sum_{k=-\infty}^{\infty} (e^{ik\omega\tau} f_k) e^{ik\omega t}$$

(6) Translation im Frequenzbereich:

$$e^{in\omega t} f(t) = \sum_{k=-\infty}^{\infty} f_{k-n} e^{ik\omega t}, \quad n \in \mathbb{Z}$$

Bemerkung 9.5.

Verbindung zu 2π -periodischen Funktionen: Besitzt f die Periode T , so besitzt $F(t) := f(\frac{t}{\omega})$ die Periode 2π .

Satz 9.13 (Parsevallsche Gleichung).

Sind f und g zwei T -periodische, stückweise stetige Funktionen mit den Fourier-Reihen $f(t) = \sum_{k=-\infty}^{\infty} f_k e^{ik\omega t}$ und $g(t) = \sum_{k=-\infty}^{\infty} g_k e^{ik\omega t}$, so gelten

$$\sum_{k=-\infty}^{\infty} f_k \overline{g_k} = \frac{1}{T} \int_0^T f(t) \overline{g(t)} dt, \quad (9.8.4)$$

$$\sum_{k=-\infty}^{\infty} |f_k|^2 = \frac{1}{T} \int_0^T |f(t)|^2 dt \quad (\text{Parsevallsche Gleichung}). \quad (9.8.5)$$

Aus (9.8.5) folgt für reellwertige Funktionen die schon behandelte reelle Version der Parsevallschen Gleichung (9.5.2). Die Verbindung zwischen (9.8.5) und (9.5.2) ergibt sich durch Einsetzen der Beziehung $c_k = (a_k - ib_k)/2$ und Zusammenfassung der Summanden mit Indices k und $-k$. In technischen Anwendungen liegen Funktionen (Signale) typischerweise nicht in kontinuierlicher Form vor, sondern als diskrete Messwerte oder als digitale Daten. Da die Abtastrate oft gleichabständig ist gehen wir von einer 2π -periodischen Funktion $f = f(x)$ aus, für die die Funktionswerte $y_j = f(x_j)$ an den $N + 1$ Punkten

$$x_j = j \cdot \frac{2\pi}{N}, \quad j = 0, \dots, N$$

gegeben sind. Aufgrund der Periodizität gilt $y_0 = y_N$. Dabei ist es eigentlich egal, ob die Funktionswerte $\{y_j\}_{j=0}^N$ nur als Messwerte oder durch Auswertung einer expliziten Formel für f an den Stützstellen x_j entstanden sind. Ziel ist es nun, die diskreten Werte in analoger Weise durch eine geeignete Orthogonalbasis darzustellen und ggf. durch Abschneiden zu approximieren.

Literaturverzeichnis

- [1] Sergey Brin and Lawrence Page *The anatomy of a large-scale hypertextual Web search engine*, Computer Networks and ISDN Systems, 30 (1998), pp. 107-117. doi:[10.1016/S0169-7552\(98\)00110-X](https://doi.org/10.1016/S0169-7552(98)00110-X)
- [2] Paul Halmos *Finite-Dimensional Vector Spaces*, Springer, Undergraduate Texts in Mathematics, Korrig. 2. Druck der 1993 Auflage. ISBN: 978-0387900933.
- [3] Lothar Papula, *Übungen zur Mathematik für Ingenieure.: Anwendungsorientierte Übungsaufgaben aus Naturwissenschaft und Technik mit ausführlichen Lösungen. 187 Übungsaufgaben mit Lösungen, 310 Bilder und ein Anhang Physikalische Grundlagen*, ISBN 3-528-04355-5, 1990, Vieweg.
- [4] Karl-Heinz Gärtner und Roland Schmieder *Lineare Algebra und Analytische Geometrie in Fragen und Übungsaufgaben*. B.G. Teubner Stuttgart, Leipzig, 1998, ISBN 3-519-00220-5