Institut für Mathematik, RWTH Aachen

Vorlesung im
Wintersemester 2004/2005, 2006/2007, 2008/2009, 2010/2011, 2012/2013, 2014/2015, 2016/2017
Sommersemester 2005, 2007, 2009, 2011, 2013, 2015, 2017

Lineare Algebra I, II

Sätze und Definitionen – ein Gerüst zur Vorlesung

Prof. Dr. Heiko von der Mosel

Letzte Änderung: 20. Januar 2025

© Copyright 2025 Prof. Dr. H. von der Mosel

Dieses Skript ist unter `http://www.instmath.rwth-aachen.de/teaching/la` verfügbar.

Anregungen und Korrekturvorschläge bitte an `heiko@instmath.rwth-aachen.de`.

Inhaltsverzeichnis

1 Der euklidische Raum $\mathbb{R}^n$ **1**

2 Geometrie im $\mathbb{R}^n$ **7**

3 Vektorräume **13**

4 Lineare Gleichungssysteme **23**

5 Determinanten **29**

6 Eigenwerte und Eigenvektoren **35**

7 Komplexe Zahlen, Fundamentalsatz der Algebra und Jordannormalformen **47**

8 Lineare Differentialgleichungssysteme **63**

9 Lineare Optimierung **67**

Kapitel 1

Der euklidische Raum $\mathbb{R}^n$

Anders als in der Differentialrechnung starten wir im ersten Kapitel zunächst mit grundlegenden Beispielen von linearen Räumen, insbesondere dem *euklidischen Raum* $\mathbb{R}^n$, dessen Elemente Raumpunkte oder *Vektoren* sind, und erläutern die zugehörigen Rechenoperationen. Dies erlaubt auch die präzise Formulierung von geometrischen Konzepten wie *Geraden*, *Strecken* oder *Ebenen* im euklidischen Raum. Damit lässt sich z.B. die Zugehörigkeit von Punkten zu gegebenen Geraden einfach berechnen. Das damit verbundene Lösen von *linearen Gleichungssystemen* wird später in Kapitel 4 systematisch behandelt.

Definition 1.3 [Der euklidische Raum]
Der *euklidische Raum* $\mathbb{R}^n$ ist definiert durch

$$\mathbb{R}^n := \{x = (x_1, x_2, \ldots, x_n) : x_1, \ldots, x_n \in \mathbb{R}\}.$$

Die Zahl $n \in \mathbb{N}$ wird *Dimension* genannt.
Meist schreiben wir die *Vektoren* $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ in Spaltenform

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Es gibt zwei Operationen in $\mathbb{R}^n$, die *Addition* "+" und die *Multiplikation mit einem Skalar* "$\cdot$", die in der folgenden Weise definiert sind: Seien

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n \text{ und } \alpha \in \mathbb{R},$$

dann setzen wir

$$x + y := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} \in \mathbb{R}^n$$

und

$$\alpha \cdot x = \alpha x := \begin{pmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{pmatrix} \in \mathbb{R}^n.$$

Für die Vektoren $x, y \in \mathbb{R}^n$ gilt $x = y :\Leftrightarrow x_i = y_i$ für alle $i = 1, \ldots n$.
Der *Nullvektor* im $\mathbb{R}^n$ ist gegeben durch

$$0 := \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^n,$$

und die Vektoren

$$e_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{pmatrix}, \quad e_j := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

heißen der 1. bzw. 2. bzw. j-te *Einheitsvektor*, wobei bei e_j die 1 an der j-ten Stelle steht.

Lemma 1.4 [Rechenregeln im $\mathbb{R}^n$]
$\forall\ x, y, z \in \mathbb{R}^n$, $\alpha, \beta \in \mathbb{R}$ gilt:

(A1) $x + y = y + x$

(A2) $x + (y + z) = (x + y) + z$

(A3) $x + 0 = x$

(A4) $\forall\ x \in \mathbb{R}^n \quad \exists!\ y \in \mathbb{R}^n : x + y = 0 \quad (y =: -x)$

(S1) $(\alpha + \beta)x = \alpha x + \beta x$

(S2) $\alpha(x + y) = \alpha x + \alpha y$

(S3) $1 \cdot x = x$

(S4) $\alpha(\beta x) = (\alpha\beta)x$.

Wir schreiben $x - y := x + (-y)$.

Definition 1.5
(i) Für $p, v \in \mathbb{R}^n$, $v \neq 0$, heißt die Menge

$$G := \{x \in \mathbb{R}^n : \ \exists\, t \in \mathbb{R} : x = p + tv\} = \{x \in \mathbb{R}^n : x = p + tv,\ t \in \mathbb{R}\}$$

Gerade durch p mit Richtungsvektor $v \neq 0$. Man nennt $x = p + tv$ auch eine *Parameterdarstellung* der Geraden G und schreibt manchmal auch $G = p + \mathbb{R}v$ oder $G = p + \operatorname{span}\{v\}$.

(ii) Für $p, q \in \mathbb{R}^n$ mit $p \neq q$ heißt die Menge

$$G := \{x \in \mathbb{R}^n : x = p + t(q - p), t \in \mathbb{R}\}$$

Gerade durch p und q.

(iii) Die Menge

$$S_1 := \{x \in \mathbb{R}^n : x = p + tv, t \in [a, b]\}$$

ist eine *Strecke*. Speziell ist für $q \in \mathbb{R}^n$, $q \neq p$, die Menge

$$S_2 := \{x \in \mathbb{R}^n : x = p + t(q - p), t \in [0, 1]\}$$

die *Verbindungsstrecke zwischen p und q.*

Lemma 1.6
Seien die Geraden

$$\begin{aligned} G_1 &= \{x \in \mathbb{R}^n : x = p + tu,\ t \in \mathbb{R}\}, \quad u \neq 0, \ \textit{und} \\ G_2 &= \{y \in \mathbb{R}^n : y = q + sv,\ s \in \mathbb{R}\}, \quad v \neq 0, \end{aligned}$$

gegeben. Dann gilt:

$$G_1 = G_2 \Leftrightarrow \ \exists\, \lambda, \mu \in \mathbb{R} : q = p + \lambda u \ \ \textit{und} \ \ v = \mu u.$$

Definition 1.7
(i) $u, v \in \mathbb{R}^n$ heißen *linear abhängig (l.a.)* $:\Leftrightarrow \exists\, \lambda \in \mathbb{R} : u = \lambda v$ oder $\lambda u = v$. Sonst heißen u und v *linear unabhängig (l.u.)*.

(ii) Allgemein heißen k Vektoren $v_1, v_2, \ldots, v_k$ *linear unabhängig*, wenn gilt:
Aus $\lambda_1 v_1 + \cdots + \lambda_k v_k = 0$ folgt $\lambda_i = 0\ \forall\ i = 1, \ldots, k$.

Definition 1.8
(i) Für $p, u, v \in \mathbb{R}^n$, u, v l.u., heißt die Menge

$$E := \{x \in \mathbb{R}^n : x = p + su + tv,\ s, t \in \mathbb{R}\}$$

die *von u und v aufgespannte Ebene durch p.* Man schreibt auch $E = p + \operatorname{span}\{u, v\}$ oder $E = p + \mathbb{R}u + \mathbb{R}v$.

(ii) Für $p, q, r \in \mathbb{R}^n$ mit $q - p, r - p$ l.u. heißt die Menge

$$E = \{x \in \mathbb{R}^n : x = p + t(q - p) + s(r - p),\ s, t \in \mathbb{R}\}$$

die *Ebene durch p, q, r*, oder *die von den Punkten p, q, r aufgespannte Ebene.* Man schreibt dann auch $E = p + \operatorname{span}\{q - p, r - p\}$ oder $E = p + \mathbb{R}(q - p) + \mathbb{R}(r - p)$.

Wir betrachten nun das *lineare Gleichungssystem*

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 \\ a_{21}x_1 + a_{22}x_2 &= b_2 \end{aligned} \tag{1.1}$$

mit gegebenen Zahlen $a_{ij}, b_k \in \mathbb{R}$, $i, j, k = 1, 2$. Wir analysieren nun die Lösungsmenge

$$L := \{x = (x_1, x_2) \in \mathbb{R}^2 : x \text{ löst } (1.1)\}$$

in Abhängigkeit von den *Koeffizienten* a_{ij} und von der rechten Seite $b = (b_1, b_2)$:

Satz 1.10
Sei $D := a_{11}a_{22} - a_{12}a_{21}$.
(i) *Falls* $D \neq 0$, *dann gibt es genau eine Lösung* $x = (x_1, x_2) \in \mathbb{R}^2$, *so dass* (1.1) *wahr ist. Tatsächlich ist dann*

$$\begin{aligned} x_1 &= \frac{1}{D}(b_1a_{22} - b_2a_{12}), \\ x_2 &= \frac{1}{D}(b_2a_{11} - b_1a_{21}). \end{aligned}$$

(ii) *Falls* $D = 0$, *dann ist entweder a)* (1.1) *nicht lösbar, oder b)* (1.1) *ist nicht eindeutig lösbar. Im Fall b) können für die Lösungsmenge L die folgenden Situationen eintreten:*

α) $L = \{x \in \mathbb{R}^2 : x_1 = -\frac{a_{12}}{a_{11}}x_2 + \frac{b_1}{a_{11}},\ x_2 \in \mathbb{R}\}$

β) $L = \{x \in \mathbb{R}^2 : x_2 = -\frac{a_{21}}{a_{22}}x_1 + \frac{b_2}{a_{22}},\ x_1 \in \mathbb{R}\}$

γ) $L = \{x \in \mathbb{R}^2 : x_1 = \frac{b_2}{a_{21}},\ x_2 \in \mathbb{R}\}$

δ) $L = \{x \in \mathbb{R}^2 : x_2 = \frac{b_1}{a_{12}},\ x_1 \in \mathbb{R}\}$

ϵ) $L = \mathbb{R}^2$.

Kapitel 2

Geometrie im $\mathbb{R}^n$

Die Einführung des *Skalarproduktes* zweier Vektoren im euklidischen Raum $\mathbb{R}^n$ erlaubt es, Winkel und Abstände zu bestimmen. Damit können auch kompliziertere geometrische Aufgaben im $\mathbb{R}^n$ rein rechnerisch gelöst werden. Normalenvektoren von Ebenen lassen sich im $\mathbb{R}^3$ mit Hilfe des *Dachproduktes* oder *Kreuzproduktes* angeben, und dieses Produkt zweier Vektoren, dass wiederum einen Vektor liefert, kann auch zur Bestimmung von Flächen- und Rauminhalten benutzt werden. Ganz allgemein werden wir im $\mathbb{R}^n$ Unterräume definieren und den Begriff des Senkrechtstehens, der sogenannten *Orthogonalität* erklären. Die *orthogonale Projektion* ist die mathematisch stringente Definition des "Lotfällens", und dabei wird auch die wichtige Rolle von speziellen Vektorsystemen deutlich, den *Orthonormalsystemen*, die wir mit der *Gram-Schmidtschen* Methode einfach erzeugen können.

Definition 2.1
(i) Das *Skalarprodukt* zweier Vektoren $x, y \in \mathbb{R}^n$ ist definiert durch

$$\begin{aligned} x \cdot y &:= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \\ &= \sum_{i=1}^{n} x_i y_i. \end{aligned}$$

(ii) $x \cdot y = 0 \Leftrightarrow: x \perp y$ heißt, dass x senkrecht (*orthogonal*) zu y ist.
(iii) Die *Norm* (=Länge) von $x \in \mathbb{R}^n$ ist gegeben durch

$$\|x\| := \sqrt{x \cdot x} = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

(iv) Die *Distanz* $d(x, y)$, also der Abstand von $x, y \in \mathbb{R}^n$, ist definiert durch

$$d(x, y) := \|x - y\|.$$

(v) Der kleinere, nichtorientierte, also stets positiv zu nehmende Winkel $\phi = \sphericalangle(x, y) \in [0, \pi]$ zwischen $x, y \in \mathbb{R}^n \setminus \{0\}$ ist gegeben durch die Beziehung

$$\cos \phi = \frac{x \cdot y}{\|x\| \cdot \|y\|}.$$

Satz 2.3 [Cauchy-Schwarzsche Ungleichung]
$\forall\ x, y \in \mathbb{R}^n : |x \cdot y| \leq \|x\| \cdot \|y\|$ *mit Gleichheit genau dann, wenn* x, y *linear abhängig sind.*

Satz 2.4 [Parallelogramm-Identitäten]
Für alle $x, y \in \mathbb{R}^n$ *gelten die Identitäten*

$$\begin{aligned} \|x+y\|^2 + \|x-y\|^2 &= 2\|x\|^2 + 2\|y\|^2, \\ \|x+y\|^2 - \|x-y\|^2 &= 4x \cdot y. \end{aligned}$$

Weiterhin gilt für alle $x, y \in \mathbb{R}^n$ die *Dreiecksungleichung*

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

Im Spezialfall des $\mathbb{R}^3$ definieren wir ein weiteres Produkt zwischen zwei Vektoren, dass wieder auf einen Vektor im $\mathbb{R}^3$ führt:

Definition 2.5
Das *Vektorprodukt* (oder *Kreuzprodukt* oder *äußeres Produkt*) $x \wedge y \in \mathbb{R}^3$ zweier Vektoren $x, y \in \mathbb{R}^3$ ist definiert durch

$$x \wedge y = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \wedge \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} := \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ x_3 y_1 - x_1 y_3 \\ x_1 y_2 - x_2 y_1 \end{pmatrix} \in \mathbb{R}^3.$$

Es gelten die folgenden Rechenregeln für das Vektorprodukt:

Satz 2.6
Seien $x, y \in \mathbb{R}^3$. Dann gilt

(i) x, y *linear abhängig* $\Leftrightarrow x \wedge y = 0$

(ii) $x \wedge y = -y \wedge x$

(iii) $(x \wedge y) \cdot x = 0$ *und* $(x \wedge y) \cdot y = 0$

(iv) $\|x \wedge y\| = \|x\| \cdot \|y\| \sin\phi$ *und* $\|x \wedge y\|^2 = \|x\|^2 \|y\|^2 - (x \cdot y)^2$, *wobei* $\phi \in [0, \pi]$ *der von x und y eingeschlossene Winkel ist, vgl. Definition 2.1 (v).*

Bemerkung. Elementargeometrisch läßt sich verifizieren, dass der Ausdruck $\|x \wedge y\|$ dem Flächeninhalt des von x und y aufgespannten Parallelogramms gleich ist.

Weitere Rechenregeln für das Vektorprodukt sind zusammengefasst in dem

Satz 2.7
Für $x, y, z \in \mathbb{R}^3$, $\alpha, \beta, \gamma \in \mathbb{R}$ gilt:

(i) $(\alpha x + \beta y) \wedge z = \alpha(x \wedge z) + \beta(y \wedge z)$ *und* $x \wedge (\beta y + \gamma z) = \beta(x \wedge y) + \gamma(x \wedge z)$

(ii) $x \wedge (y \wedge z) = y(x \cdot z) - z(x \cdot y)$ *("bac-cab-Regel" oder Grassmannscher Entwicklungssatz)*

(iii) $x \wedge (y \wedge z) + y \wedge (z \wedge x) + z \wedge (x \wedge y) = 0$ *(Jacobi-Identität).*

Definition 2.8
Seien $x, y, z \in \mathbb{R}^3$, dann heißt die Zahl $z \cdot (x \wedge y) \in \mathbb{R}$ das *Spatprodukt* der Vektoren x, y und z.

Bemerkung. Das Spatprodukt $z \cdot (x \wedge y)$ gibt bis auf Vorzeichen das Volumen des von x, y, z aufgespannten Parallelepipeds (oder Spats) an. Es gilt $z \cdot (x \wedge y) = x \cdot (y \wedge z) = y \cdot (z \wedge x)$.

Definition 2.9
(i) Für $v_1, v_2, \ldots, v_k \in \mathbb{R}^n$ heißt die Menge

$$U_k := \{x \in \mathbb{R}^n : x = t_1 v_1 + \cdots + t_k v_k,\ t_i \in \mathbb{R}, i = 1, \ldots, k\}$$

der von $v_1, \ldots, v_k$ aufgespannte Unterraum des $\mathbb{R}^n$. Man schreibt auch
$U_k = \operatorname{span}\{v_1, \ldots, v_k\}$ (oder $U_k = \mathbb{R}v_1 + \ldots + \mathbb{R}v_k$).

(ii) Die Menge

$$V_k := \{x \in \mathbb{R}^n : x = p + t_1 v_1 + \cdots + t_k v_k,\ t_i \in \mathbb{R}, i = 1, \ldots, k\} = p + \operatorname{span}\{v_1, \ldots, v_k\}$$

heißt *der von $v_1, \ldots, v_k$ aufgespannte affine Unterraum durch $p \in \mathbb{R}^n$*.

Bemerkung. Jeder affine Raum, der die $0 \in \mathbb{R}^n$ enthält, ist automatisch ein Unterraum.

Definition 2.10
Die Menge $\{v_1, \ldots, v_k\} \subset \mathbb{R}^n$ heißt *Orthonormalsystem* $:\Leftrightarrow$

$$v_i \cdot v_j = \delta_{ij} := \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j \end{cases} \qquad \forall\, i, j \in \{1, \ldots, k\}.$$

Die Größe δ_{ij} bezeichnet man auch als das *Kronecker-Symbol*.

Definition 2.12
Sei $x \in \mathbb{R}^n$ und $\{v_1, \ldots, v_k\} \subset \mathbb{R}^n$ ein Orthonormalsystem und $E := \operatorname{span}\{v_1, \ldots, v_k\}$.
Dann heißt der Vektor

$$P_E(x) := \sum_{i=1}^{k} (x \cdot v_i) v_i$$

die *orthogonale Projektion* von x auf E.

Bemerkung. (i) Es gilt $P_E(x) \in E$ und $(x - P_E(x)) \perp y\ \forall\, y \in E$, wofür man auch
$(x - P_E(x)) \perp E$ schreibt.

(ii) Erfüllen $v_1, \ldots, v_k \in \mathbb{R}^n \setminus \{0\}$ lediglich $v_i \cdot v_j = 0$ für alle $i \neq j$, so ist $\{\frac{v_1}{\|v_1\|}, \ldots, \frac{v_k}{\|v_k\|}\}$ ein Orthonormalsystem, und es gilt $P_E(x) := \sum_{i=1}^{k} \frac{x \cdot v_i}{\|v_i\|^2} v_i$.

Definition 2.13
Seien $v_1, v_2 \in \mathbb{R}^3$ linear unabhängig, $E := \operatorname{span}\{v_1, v_2\}$. Dann heißt die Menge

$$E^\perp := \{x \in \mathbb{R}^3 : x \perp E\}$$

die *Normale* zu der Ebene E.
Allgemeiner definiert man für eine Teilmenge $A \subset \mathbb{R}^n$ das *orthogonale Komplement* $A^\perp$ als

$$A^\perp := \{x \in \mathbb{R}^n : x \cdot a = 0 \ \forall\ a \in A\}.$$

Lemma 2.14
Für linear unabhängige Vektoren $u, v \in \mathbb{R}^3$ und $E := \operatorname{span}\{u, v\}$ gilt

$$E^\perp = \{y \in \mathbb{R}^3 : y = tu \wedge v,\ t \in \mathbb{R}\}.$$

Definition 2.15
Für linear unabhängige Vektoren $u, v \in \mathbb{R}^3$ und $E := \operatorname{span}\{u, v\}$ heißt

$$(n_E =) \quad n := \frac{u \wedge v}{\|u \wedge v\|}$$

der *Einheits-Normalenvektor* von E.

Lemma 2.16
$E^{\perp\perp} := (E^\perp)^\perp = E$.

Definition 2.17
Sei $n \in \mathbb{R}^3 \setminus \{0\}$. Dann ist die Menge $E := \{x \in \mathbb{R}^3 : x \cdot n = d\}$ eine Ebene gegeben in *Normalform*.
Falls $\|n\| = 1$, dann nennt man diese Darstellung genauer die *Hessesche Normalform* der Ebene E.

Lemma 2.18
Falls $\{v_1, v_2, v_3\} \subset \mathbb{R}^3$ ein Orthonormalsystem ist, dann gilt $v_3 = v_1 \wedge v_2$ oder $v_3 = -v_1 \wedge v_2$.

Bemerkung. Wir vereinbaren: falls $v_3 = v_1 \wedge v_2$, dann nennen wir das Orthonormalsystem $\{v_1, v_2, v_3\}$ *positiv orientiert*, falls $v_3 = -v_1 \wedge v_2$, dann nennen wir das Orthonormalsystem $\{v_1, v_2, v_3\}$ *negativ orientiert*.

Satz 2.19
Falls $\{v_1, v_2, v_3\} \subset \mathbb{R}^3$ ein Orthonormalsystem ist, dann gibt es für jedes $x \in \mathbb{R}^3$ eine eindeutige Darstellung der Form

$$x = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3$$

mit $\alpha_i := x \cdot v_i \in \mathbb{R}$ für $i = 1, 2, 3$.

Bemerkung. Wir nennen deshalb das Orthonormalsystem $\{v_1, v_2, v_3\}$ auch eine *Orthonormalbasis* des $\mathbb{R}^3$.

Satz 2.20 [Gram-Schmidtsches Orthonormalisierungsverfahren]
Seien $v_1, \ldots, v_k \in \mathbb{R}^n$ linear unabhängige Vektoren. Dann kann ein Orthonormalsystem $\{w_1, \ldots, w_k\}$ des $\mathbb{R}^n$ in der folgenden Weise (rekursiv) konstruiert werden:

$$\begin{aligned} w_1 &:= \frac{v_1}{\|v_1\|} \\ w_l &:= \frac{v_l - \sum_{i=1}^{l-1}(v_l \cdot w_i)w_i}{\|v_l - \sum_{i=1}^{l-1}(v_l \cdot w_i)w_i\|}, \quad l = 2, \ldots, k. \end{aligned}$$

Es gilt $\operatorname{span}\{v_1, \ldots, v_k\} = \operatorname{span}\{w_1, \ldots, w_k\}$.

Kapitel 3

Vektorräume

Die in den ersten beiden Kapiteln im Wesentlichen nur im $\mathbb{R}^n$ behandelten linearen Strukturen werden wir nun abstrakt als *Vektorraumeigenschaften* fordern und die Konzepte von *Unterräumen* und *Linearkombinationen* definieren, um möglichst viele mathematische Eigenschaften nur aus diesen Definitionen zu ziehen. Das hat den Vorteil, dass jede Situation, in der man eine solche Vektorraumstruktur hat (wie zum Beispiel bei der Menge aller Polynome vom Grade n), all diese Eigenschaften zur Verfügung hat. Dieser auf den ersten Blick vielleicht etwas ungewohnte Abstraktionsgrad vereinfacht also in Wirklichkeit die Lage, weil man nur auf ein begrenztes Instrumentarium zurückgreifen muss. Wir werden den zentralen Begriff einer *Basis* eines Vektorraums kennenlernen, der es erlaubt, jeden Vektor als Linearkombination der Basisvektoren auszudrücken. Die Zusammenfassung der zugehörigen reellen Koeffizienten als *Spalten-* oder *Zeilenvektor* erlaubt dann wieder das Rechnen mit den Regeln im $\mathbb{R}^n$, wie wir sie im ersten Kapitel kennengelernt haben. Dasselbe kann man für *lineare Abbildungen* zwischen zwei Vektorräumen tun; entsprechende Entwicklungen von Urbildvektor und Bildvektor in zugehörigen Basen von Urbild- und Bildvektorraum erlauben die Darstellung von linearen Abbildungen als *Matrizen*. Verschiedene Operationen wie die Addition oder die Verkettung von linearen Abbildungen lassen sich dann durch zugehörigen *Matrixoperationen* darstellen und berechnen. Dies lässt sich auch später z.B. bei der systematischen Berechnung von Lösungen linearer Gleichungssysteme in Kapitel 4 gewinnbringend anwenden.

Definition 3.1
Eine Menge $V \neq \emptyset$ heißt *Vektorraum (über $\mathbb{R}$)*, wenn es eine Addition "+" und eine Multiplikation "·" in V gibt, so dass für alle $u, v, w \in V$ und für alle $\alpha, \beta \in \mathbb{R}$ gilt:

(A1) $u + v = v + u$

(A2) $u + (v + w) = (u + v) + w$

(A3) $\exists\, 0 \in V : u + 0 = u$

(A4) $\forall\, u \in V\ \exists\, v \in V : u + v = 0 \quad (v =: -u)$

(S1) $(\alpha + \beta)u = \alpha u + \beta u$

(S2) $\alpha(u + v) = \alpha u + \alpha v$

(S3) $1 \cdot u = u$

(S4) $\alpha(\beta u) = (\alpha\beta)u.$

Die Elemente $u, v, w \in V$ heißen *Vektoren* und $\alpha, \beta \in \mathbb{R}$ *Skalare*.

Definition 3.2
Sei V ein Vektorraum (über $\mathbb{R}$). Dann heißt $\emptyset \neq U \subset V$ *Unterraum (Untervektorraum, Teilraum)* von $V :\Leftrightarrow$

(i) $0 \in U$

(ii) $u, v \in U \Rightarrow u + v \in U$

(iii) $u \in U \Rightarrow \alpha u \in U \ \forall\ \alpha \in \mathbb{R}$.

Definition 3.3
(i) Für $v_1, v_2, \ldots, v_n \in V$, V ein Vektorraum, und $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ heißt

$$v := \alpha_1 v_1 + \cdots + \alpha_n v_n = \sum_{i=1}^{n} \alpha_i v_i$$

eine *Linearkombination* der v_i, $i = 1, \ldots, n$, und wir schreiben

$$\operatorname{span}\{v_1, \ldots, v_n\} := \{v \in V : v = \sum_{i=1}^{n} \alpha_i v_i,\ \alpha_i \in \mathbb{R}\}$$

für die Menge aller Linearkombinationen der Vektoren $v_1, \ldots, v_n$.

(ii) Die Vektoren $v_1, \ldots, v_n \in V$ heißen *linear abhängig* $:\Leftrightarrow$

$$\exists\, \alpha_1, \ldots, \alpha_n \in \mathbb{R} \quad \text{mit } \alpha_1^2 + \cdots + \alpha_n^2 \neq 0, \quad \text{so dass} \quad \sum_{i=1}^{n} \alpha_i v_i = 0.$$

Sind $v_1, \ldots, v_n \in V$ nicht linear abhängig, so heißen sie *linear unabhängig*.

Bemerkung. Gilt z.B.

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

und $\alpha_j \neq 0$, für ein $j \in \{1, \ldots, n\}$, dann kann man v_j als Linearkombination der übrigen Vektoren v_i, $i \neq j$, darstellen:

$$v_j = \frac{-\alpha_1 v_1 - \cdots - \alpha_{j-1} v_{j-1} - \alpha_{j+1} v_{j+1} - \cdots - \alpha_n v_n}{\alpha_j} = -\frac{1}{\alpha_j} \sum_{i \neq j} \alpha_i v_i.$$

Lemma 3.4
Es gilt

$$v_1, \ldots, v_n \quad \text{sind linear unabhängig} \Leftrightarrow \left(\sum_{i=1}^{n} \alpha_i v_i = 0 \Rightarrow \alpha_i = 0\ \forall\ i = 1, \ldots, n \right).$$

Definition 3.6
Die Menge $\mathscr{B} := \{v_1, \ldots, v_n\} \subset V$ heißt *Basis* des Vektorraums V, wenn es für alle $w \in V$ eindeutig bestimmte Zahlen $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ gibt, so dass

$$w = \sum_{i=1}^{n} \alpha_i v_i.$$

Die Zahlen $\alpha_i, i = 1, \ldots, n$, heißen die *Koordinaten* von w in der Basis $\mathscr{B}$, und wir schreiben auch

$$w_{\mathscr{B}} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}.$$

Lemma 3.7
Die Menge $\mathscr{B} := \{v_1, \ldots, v_n\} \subset V$ ist Basis des Vektorraums $V \Leftrightarrow$

(i) $V = \operatorname{span}\{v_1, \ldots, v_n\}$ *und*

(ii) $v_1, \ldots, v_n$ *sind linear unabhängig.*

Lemma 3.9
Es sei $V = \operatorname{span}\{v_1, \ldots, v_n\}$. *Dann sind Vektoren* $w_1, w_2, \ldots, w_m \in V$ *linear abhängig, wenn* $m > n$.

Satz 3.10 [Basisergänzung]
Sei $V = \operatorname{span}\{v_1, \ldots, v_n\}$. *Eine Menge linear unabhängiger Vektoren* $\{w_1, \ldots, w_m\} \subset V$, *die keine Basis darstellt, kann man durch geeignete Vektoren* $v_i, i \in \{1, \ldots, n\}$ *zu einer Basis ergänzen.*

Satz 3.11 [Basisauswahl]
Es sei $V = \operatorname{span}\{v_1, \ldots, v_n\}$, *wobei* $V \neq \{0\}$. *Dann gibt es eine Basis* $\mathscr{B}$ *von* V *mit* $\mathscr{B} \subset \{v_1, \ldots, v_n\}$.

Bemerkung. Sei $\mathscr{B}$ eine Basis eines Vektorraums V mit $\sharp\mathscr{B} = n$, d.h. $\mathscr{B}$ besitzt n Elemente. Dann gilt $\sharp\mathscr{B}' = n$ für alle Basen $\mathscr{B}'$ von V.

Definition 3.12 [Dimension]
Sei $\mathscr{B}$ eine Basis eines Vektorraums V mit $\sharp\mathscr{B} = n$, dann heißt die Zahl n die *Dimension* von V ("$\dim V = n$").

Bemerkung. Falls die Vektoren $w_1, \ldots, w_n$ linear unabhängig in einem Vektorraum V mit der Dimension n sind, dann ist $\{w_1, \ldots, w_n\}$ eine Basis von V.

Definition 3.14
Seien V und W Vektorräume (über $\mathbb{R}$). Dann heißt die Abbildung $L : V \to W$ *linear* $:\Leftrightarrow$

$$L(\alpha u + \beta v) = \alpha L(u) + \beta L(v) \quad \forall\ \alpha, \beta \in \mathbb{R} \text{ und } \forall\ u, v \in V.$$

Satz 3.15
Seien $\mathscr{B} = \{v_1, \ldots, v_n\} \subset V$ und $\mathscr{C} = \{w_1, \ldots, w_m\} \subset W$ Basen von V bzw. W. Dann gilt:

(i) *Eine lineare Abbildung $L : V \to W$ ist auf V allein durch die Werte $L(v_1), \ldots, L(v_n)$ festgelegt.*

(ii) *Zu n Vektoren $y_1, \ldots, y_n \in W$ gibt es genau eine lineare Abbildung $L : V \to W$, so dass*

$$L(v_i) = y_i \quad \forall\ i = 1, \ldots, n.$$

(iii) *Eine lineare Abbildung $L : V \to W$ ist eindeutig durch ein Zahlenschema*

$$\mathscr{L} := \begin{pmatrix} l_{11} & l_{12} & \cdots & l_{1n} \\ l_{21} & l_{22} & \cdots & l_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ l_{m1} & l_{m2} & \cdots & l_{mn} \end{pmatrix}, \quad l_{ij} \in \mathbb{R},\ i = 1, \ldots, m,\ j = 1, \ldots, n,$$

charakterisiert.

Definition 3.16
Das Zahlenschema

$$\mathscr{L} = (l_{ij})_{\substack{i=1,\ldots,m \\ j=1,\ldots,n}} = \begin{pmatrix} l_{11} & l_{12} & \cdots & l_{1n} \\ l_{21} & l_{22} & \cdots & l_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ l_{m1} & l_{m2} & \cdots & l_{mn} \end{pmatrix}$$

heißt $(m \times n)$-*Matrix* mit *Koeffizienten* $l_{ij} \in \mathbb{R}$. Das n-Tupel $(l_{i1}, \ldots, l_{in})$ ist die i-te *Zeile*, der Spaltenvektor

$$\begin{pmatrix} l_{1j} \\ l_{2j} \\ \vdots \\ l_{mj} \end{pmatrix}$$

die j-te *Spalte* der Matrix $\mathscr{L}$.
Wenn $\mathscr{L}$ eine lineare Abbildung $L : V \to W$ bzgl. der Basen $\mathscr{B} \subset V$ und $\mathscr{C} \subset W$ beschreibt, dann nennt man $\mathscr{L}$ auch die *darstellende Matrix* von L bzgl. $\mathscr{B}$ und $\mathscr{C}$.
Den Raum der $(m \times n)$-Matrizen mit reellen Koeffizienten nennen wir $\mathbb{R}^{m \times n}$.

Lemma 3.18 [Matrix-Vektor-Multiplikation]
Seien $L : V \to W$ linear, $\mathscr{B} = \{v_1, \ldots, v_n\} \subset V$ und $\mathscr{C} = \{w_1, \ldots, w_m\} \subset W$ Basen von V bzw. W und

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

die darstellende Matrix von L bzgl. $\mathscr{B}$ und $\mathscr{C}$. Wenn

$$x_{\mathscr{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

die Koordinatendarstellung von $x \in V$ in der Basis $\mathscr{B}$ ist (vgl. Definition 3.6), dann ist

$$(L(x))_{\mathscr{C}} = Ax_{\mathscr{B}} := \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \sum_{j=1}^n a_{2j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix}$$

die Koordinatendarstellung von $L(x) \in W$ in der Basis $\mathscr{C}$.

Lemma 3.20 [Linearkombination von Matrizen]
Seien $L, M : V \to W$ lineare Abbildungen, dargestellt in den Basen $\mathscr{B} = \{v_1, \dots, v_n\} \subset V$ und $\mathscr{C} = \{w_1, \dots, w_m\} \subset W$ durch die Matrizen $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ und $B = (b_{ij}) \in \mathbb{R}^{m\times n}$. Dann ist die die lineare Abbildung $\alpha L + \beta M$ darstellende Matrix C gegeben durch

$$(c_{ij}) = C = \alpha A + \beta B = \alpha(a_{ij}) + \beta(b_{ij}) := (\alpha a_{ij} + \beta b_{ij}), \quad i \in \{1, \dots, m\},\ j \in \{1, \dots, n\}.$$

Korollar 3.21
$\mathbb{R}^{m\times n}$ ist ein Vektorraum der Dimension $m \cdot n$.

Bemerkung. Falls $m = n$, dann nennt man die Matrix

$$\mathrm{Id}_{\mathbb{R}^n} := \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \in \mathbb{R}^{n\times n}$$

die *Einheitsmatrix* des $\mathbb{R}^n$.

Lemma 3.23 [Matrizenmultiplikation]
Seien $\mathscr{B} = \{v_1, \dots, v_n\} \subset V$, $\mathscr{C} = \{w_1, \dots, w_m\} \subset W$, und $\mathscr{D} = \{z_1, \dots, z_l\} \subset Z$ Basen, $L : V \to W$ eine lineare Abbildung mit darstellender Matrix A (bzgl. $\mathscr{B}, \mathscr{C}$) und $M : W \to Z$ eine lineare Abbildung mit darstellender Matrix B (bzgl. $\mathscr{C}, \mathscr{D}$). Dann hat die lineare Abbildung $M \circ L : V \to Z$ die darstellende Matrix

$$(c_{kj}) = C = BA = B \cdot A := \left(\sum_{i=1}^{m} b_{ki} a_{ij} \right) \quad \forall\ k = 1, \dots, l,\ j = 1, \dots, n.$$

Korollar 3.25 (i) $C(DE) = (CD)E \quad \forall\ E \in \mathbb{R}^{m\times n}, D \in \mathbb{R}^{l\times m}, C \in \mathbb{R}^{k\times l}$

(ii) $C(D + E) = CD + CE \quad \forall\ E, D \in \mathbb{R}^{m\times n}, C \in \mathbb{R}^{l\times m}$

(iii) $(C + D)E = CE + DE \quad \forall\ C, D \in \mathbb{R}^{l\times m}, E \in \mathbb{R}^{m\times n}$.

Definition 3.27
Einer Matrix $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \in \mathbb{R}^{m\times n}$ kann man ihre *transponierte Matrix*

$$A^T := (a_{ji})_{\substack{j=1,\dots,n \\ i=1,\dots,m}} \in \mathbb{R}^{n\times m}$$

zuordnen. Gilt speziell für $n = m$ und für eine Matrix $A \in \mathbb{R}^{n\times n}$ die Identität

$$A^T = A,$$

so heißt A *symmetrisch*.

Kapitel 4

Lineare Gleichungssysteme

Nicht nur geometrische Fragestellungen, sondern auch diskretisierte Formen von komplizierten Differentialgleichungen oder hochdimensionale Optimierungsaufgaben aus Ökonomie und Technik führen auf zum Teil sehr umfangreiche lineare Gleichungssysteme. In diesem Kapitel werden wir mit Hilfe der Matrizentheorie des dritten Kapitels Methoden herleiten, um solche Gleichungen systematisch zu lösen. Ziel ist das Verständnis des *Gaußschen Eliminationsverfahrens*, welches es erlaubt, ein beliebiges lineares Gleichungssystem durch eine Kette von *Zeilenoperationen* ohne Änderung der Lösungsmenge auf eine Form zu bringen, aus der man die Lösungsmenge einfach ablesen kann. Von zentralem theoretischen Interesse ist hierbei auch die *Dimensionsformel*, die die Dimensionen des *Kerns* und des *Bildes* einer linearen Abbildung zueinander in Beziehung setzt.

Definition 4.2
Seien $a_{ij} \in \mathbb{R}$, $i = 1, \dots, m,\ j = 1, \dots, n$ und $b_1, \dots, b_m \in \mathbb{R}$ gegeben. Dann heißt

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

ein *lineares Gleichungssystem* mit m Gleichungen und n Unbekannten $x_1, \dots, x_n$.
Wir schreiben dafür auch

$$\sum_{j=1}^{n} a_{ij}x_j = b_i, \qquad i = 1, \dots, m,$$

oder $Ax = b$ mit einer *Koeffizientenmatrix* $A = (a_{ij})_{\substack{i=1,\dots,m \\ j=1,\dots,n}} \in \mathbb{R}^{m\times n}$ und dem Vektor $b = (b_1, \dots, b_m) \in \mathbb{R}^m$, d.h.

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

Falls $b = (b_1, b_2, \dots, b_m) = 0$, dann heißt das System (LGS) *homogen*, falls $b \neq 0$, dann heißt (LGS) *inhomogen*.
Alle $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, für die (LGS) wahr ist, heißen *Lösungen* des Gleichungssystems (LGS).

Bemerkung. Interpretiert man den Vektor $b = (b_1, \ldots, b_m)$ als Koordinatenvektor bezüglich einer Basis $\mathscr{C} \subset \mathbb{R}^m$, und die Koeffizientenmatrix $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ als darstellende Matrix einer linearen Abbildung $L : \mathbb{R}^n \to \mathbb{R}^m$ bezüglich der Basen $\mathscr{B}$ und $\mathscr{C}$, dann ist das Lösen des Gleichungssystems (LGS) gleichbedeutend mit dem Lösen der Gleichung

$$L(x) = b,$$

wobei x Koordinatenvektor bezüglich der Basis $\mathscr{B}$ ist.

Lemma 4.3
Seien $L : V \to W$ eine lineare Abbildung und V und W Vektorräume. Dann gilt:

(i) *Die Menge $L(V) := \mathscr{W}(L) = \{w \in W : w = L(v) \text{ für ein } v \in V\}$ ist ein Unterraum von W.*

(ii) *Die Menge $\{v \in V : L(v) = 0\}$ ist ein Unterraum von V.*

(iii) *L ist injektiv $\Leftrightarrow \{v \in V : L(v) = 0\} = \{0\}$.*

Definition 4.4
Sei $L : V \to W$ linear, und V und W seien Vektorräume. Dann bezeichnen wir mit

$$\mathrm{Bild}(L) := L(V) = \mathscr{W}(L)$$

das *Bild* von L. Die Dimension von $\mathrm{Bild}(L)$ heißt *Rang* von L, und wir schreiben $\dim \mathrm{Bild}(L) = \mathrm{rang} L$.
Die Menge $\{v \in V : L(v) = 0\}$ bezeichnen wir mit *Kern* von L, und wir schreiben dafür $\mathrm{Kern}(L)$.

Bemerkung. Wir ziehen die folgenden Konsequenzen aus Lemma 4.3 für das Studium des Gleichungssystems (LGS):

(i) Ist $\dim W > \dim \mathrm{Bild}(L) = \mathrm{rang} L$, dann ist (LGS) nicht immer lösbar.

(ii) Wenn (LGS) mehrere Lösungen besitzt, erhalten wir als Lösungsmenge

$$\{x \in \mathbb{R}^n : x = x_0 + z,\ z \in \mathrm{Kern}(L)\}$$

für eine (beliebige) spezielle Lösung $x_0 \in \mathbb{R}^n$.

(iii) Um zu entscheiden, ob (LGS) *eindeutig* lösbar ist, genügt es, das homogene System $Ax = 0$ zu studieren.

Satz 4.5 [Dimensionsformel]
Sei $L : V \to W$ linear und $\dim V = n$*, dann gilt*

$$\dim \mathrm{Kern}(L) + \dim \mathrm{Bild}(L) = n.$$

Satz 4.6
Für $A \in \mathbb{R}^{m\times n}$ sind die folgenden Zahlen gleich:

(i) die maximale Anzahl linear unabhängiger Zeilenvektoren, d.h. der Zeilenrang *von A,*

(ii) die maximale Anzahl linear unabhängiger Spaltenvektoren, d.h. der Spaltenrang *von A,*

(iii) der Rang der zu A gehörenden linearen Abbildung L (wofür wir nun auch Rang *von A sagen, und wir setzen* $\mathrm{rang}A := \mathrm{rang}L$*).*

Bemerkung. Sind die Vektoren $v_1, v_2, \ldots, v_k$ eines Vektorraums linear unabhängig, so sind es auch die Vektoren $v_1 + \alpha v_j, v_2, \ldots, v_k$ für alle $j \in \{1, \ldots, k\}$, $j \neq 1$ und für alle $\alpha \in \mathbb{R}$.

Satz 4.8
Sei $Ax = b$ ein lineares Gleichungssystem, wobei $A \in \mathbb{R}^{m\times n}$, $x \in \mathbb{R}^n$ und $b \in \mathbb{R}^m$. Dann gilt:

(i) *Falls* $\mathrm{rang}A = m$*, dann gibt es für alle $b \in \mathbb{R}^m$ mindestens eine Lösung $x \in \mathbb{R}^n$.*

(ii) *Falls* $\mathrm{rang}A = n$*, dann gibt es für alle $b \in \mathbb{R}^m$ höchstens eine Lösung $x \in \mathbb{R}^n$.*

Zusammenfassung zum GAUSS***schen Eliminationsverfahren.***
Das lineare Gleichungssystem (LGS) mit m Gleichungen und n Unbekannten

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

kann man durch das GAUSS*sche Eliminationsverfahren* ohne Abänderung der Lösungsmenge auf eine der beiden folgenden Gestaltvarianten bringen:

(i)

$$\begin{pmatrix} 1 & \star & \star & \star & \cdots & \star \\ 0 & 1 & \star & \star & \cdots & \star \\ 0 & 0 & 1 & \star & \cdots & \star \\ \vdots & & & \ddots & \star & \star \\ 0 & 0 & 0 & 0 & 1 & \star \\ \vdots & & & & & 1 \\ 0 & \cdots & & & & 0 \\ \vdots & & & & & \vdots \\ 0 & \cdots & & & & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ \\ \vdots \\ \\ x_n \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ \\ \vdots \\ \\ c_n \\ c_{n+1} \\ \vdots \\ c_m \end{pmatrix},$$

wobei $\star$ für einen beliebigen reellen Eintrag steht. Man kann nun direkt ablesen, dass $\operatorname{rang} A = n$, da es n linear unabhängige Spalten gibt. Deshalb ist das Gleichungssystem eindeutig lösbar, wenn $c_{n+1} = \cdots c_m = 0$ ist. Falls nur ein $c_i \neq 0$, $i \in \{n+1, \ldots, m\}$, dann gibt es keine Lösung.

(ii)

$$\begin{pmatrix} 1 & \star & \star & \star & \cdots & & & & & & & & \star \\ 0 & 1 & \star & \star & \cdots & & & & & & & & \star \\ 0 & 0 & 0 & 1 & \star & \cdots & & & & & & & \star \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & \star & \cdots & & & & \star \\ \vdots & & & & & & \ddots & & & & & & \vdots \\ & & & & & & & & & 0 & 1 & \star & \cdots & \star \\ 0 & \cdots & & & & \cdots & 0 & \cdots & & & 0 & \cdots & 0 & 0 \\ \vdots & & & & & & \vdots & & & & & & & \vdots \\ 0 & \cdots & & & & \cdots & 0 & \cdots & & & & & \cdots & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ \\ \vdots \\ \\ x_n \end{pmatrix} = \begin{pmatrix} c_1 \\ \\ \vdots \\ \\ c_k \\ c_{k+1} \\ \vdots \\ c_m \end{pmatrix},$$

wobei hier $\operatorname{rang} A = k < n$ abzulesen ist, womit die Lösungsmenge $n - k$ freie Parameter enthält, falls $c_{k+1} = \cdots = c_m = 0$ ist. Falls aber nur ein $c_i \neq 0$, $i \in \{k+1, \ldots, m\}$, dann gibt es keine Lösung.

Bemerkung. Für die Operationen im Gaußschen Eliminationsverfahren stellen wir fest, dass sich der (Zeilen-)Rang der Matrix und die Lösungsmenge des Gleichungssystems nicht ändert, wenn wir

(i) zwei Zeilen vertauschen,

(ii) eine Zeile mit $\alpha \neq 0$ multiplizieren, oder

(iii) die i-te Zeile durch die Summe aus i-ter und j-ter Zeile ersetzen.

Da die Lösungsmenge unverändert bleibt, wird auch die Dimension des Kerns der durch A dargestellten linearen Abbildung L nicht verändert. Aus der Dimensionsformel, Satz 4.5, ergibt sich damit

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

und damit einen nicht geänderten Spaltenrang von A, was Satz 4.6 beweist.

Definition 4.12
Die zu $A \in \mathbb{R}^{n\times n}$ *inverse Matrix* $A^{-1} \in \mathbb{R}^{n\times n}$ ist durch die folgende Identität definiert:

$$A^{-1}A = AA^{-1} = \operatorname{Id}_{\mathbb{R}^n} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \\ 0 & \cdots & & & 1 \end{pmatrix} \in \mathbb{R}^{n\times n}.$$

Bemerkung. Nicht jede Matrix besitzt eine Inverse!

Kapitel 5

Determinanten

Induktiv führen wir hier *Determinanten* von quadratischen Matrizen ein, um dann ihre Rechenregeln herzuleiten. Insbesondere der *Laplacesche Entwicklungssatz* ist hier von zentraler Bedeutung. Als Anwendung stellen wir dar, wie die Determinante genutzt werden kann, um schnell Aussagen über die Lösbarkeit von linearen Gleichungssystemen zu gewinnen. Außerdem ergibt sich mit Hilfe der *Cramerschen Regel* eine Alternative zur Berechnung der Lösungen linearer Gleichungssysteme, aber auch zum Bestimmen der *Inversen* einer Matrix. Determinanten spielen aber auch im folgenden sechsten Kapitel eine entscheidende Rolle für die Berechnung von Eigenwerten.

Definition 5.1
(i) Für $A \in \mathbb{R}^{1\times 1}$, also für $A = (a)$, $a \in \mathbb{R}$, ist die *Determinante* von A ($= \det A$) definiert durch

$$\det A := a \qquad =: |A|.$$

(ii) Für

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbb{R}^{2\times 2}$$

setzen wir

$$\det A := a_{11}a_{22} - a_{12}a_{21} \qquad =: \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}.$$

(iii) Für

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in \mathbb{R}^{3\times 3}$$

setzen wir

$$\begin{aligned} \det A \quad := \quad & a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23} \\ & -a_{31}a_{22}a_{13} - a_{23}a_{32}a_{11} - a_{33}a_{12}a_{21} \qquad =: \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}. \end{aligned}$$

(iv) Für

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \in \mathbb{R}^{n\times n}$$

setzen wir

$$\det A := \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \det \hat{A}_{i1},$$

wobei $\hat{A}_{ij} \in \mathbb{R}^{(n-1)\times(n-1)}$ die Matrix ist, die aus der ursprünglichen Matrix A entsteht, wenn man dort die i-te Zeile und j-te Spalte streicht. Man nennt $\det \hat{A}_{ij}$ auch den (i, j)-ten *Minor* von A.

Satz 5.3 [Laplacescher Entwicklungssatz]
Für $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ gilt

$$\begin{aligned} \det A &= \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det \hat{A}_{ij} \quad \forall\ j = 1, \ldots, n \\ &= \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det \hat{A}_{ij} \quad \forall\ i = 1, \ldots, n, \end{aligned}$$

wobei der erste Term auf der rechten Seite die Entwicklung nach der j-ten Spalte, der zweite die Entwicklung nach der i-ten Zeile darstellt.

Satz 5.5
Für $A \in \mathbb{R}^{n\times n}$ mit der Transponierten $A^T \in \mathbb{R}^{n\times n}$ gilt

$$\det A = \det A^T.$$

Satz 5.7 [Rechenregeln für Determinanten]
Die Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \in \mathbb{R}^{n\times n}$$

habe die Zeilen $z_1, \ldots, z_n$ *und die Spalten* $s_1, \ldots, s_n$*, d.h.*

$$\begin{aligned} A &= \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} \textit{ für } z_i := (a_{i1}, \ldots, a_{in}),\ i = 1, \ldots, n \\ &= (s_1, \ldots, s_n) \textit{ für } s_i := \begin{pmatrix} a_{1i} \\ \vdots \\ a_{ni} \end{pmatrix}. \end{aligned}$$

Dann gelten die folgenden Regeln:
(i)

$$\det \begin{pmatrix} z_1 \\ \vdots \\ z_{k-1} \\ \lambda z_k \\ z_{k+1} \\ \vdots \\ z_n \end{pmatrix} = \lambda \det A \qquad \forall\ \lambda \in \mathbb{R}$$

und

$$\det(s_1, \ldots, s_{k-1}, \lambda s_k, s_{k+1}, \ldots, s_n) = \lambda \det A \qquad \forall\ \lambda \in \mathbb{R}.$$

(ii) $\det(\lambda A) = \lambda^n \det A$
(iii) Für einen (Zeilen-)Vektor $b = (b_1, \ldots, b_n) \in \mathbb{R}^n$ *gilt*

$$\det \begin{pmatrix} z_1 \\ \vdots \\ z_{k-1} \\ z_k + b \\ z_{k+1} \\ \vdots \\ z_n \end{pmatrix} = \det A + \det \begin{pmatrix} z_1 \\ \vdots \\ z_{k-1} \\ b \\ z_{k+1} \\ \vdots \\ z_n \end{pmatrix},$$

und für einen (Spalten-)Vektor $c = (c_1, \ldots, c_n)^T \in \mathbb{R}^n$ *gilt*

$$\det(s_1, \ldots, s_{k-1}, s_k + c, s_{k+1}, \ldots, s_n) = \det A + \det(s_1, \ldots, s_{k-1}, c, s_{k+1}, \ldots, s_n).$$

(iv) Sind zwei Spalten oder zwei Zeilen von A gleich, dann ist $\det A = 0$.
(v) Es gilt $\det \mathrm{Id}_{\mathbb{R}^n} = 1$.
(vi) Falls eine Zeile oder eine Spalte von A verschwindet, dann ist $\det A = 0$.

(vii) Gilt $\operatorname{rang} A < n$*, dann ist* $\det A = 0$.
(viii) Die Addition eines Vielfachen einer Zeile zu einer anderen Zeile oder die Addition eines Vielfachen einer Spalte zu einer anderen Spalte ändert den Wert $\det A$ *nicht.*
(ix) Die Vertauschung zweier Zeilen oder zweier Spalten ändert das Vorzeichen von $\det A$.

Bemerkung. Die Operationen im Gaußschen Eliminationsverfahren, die die Lösungsmenge eines Gleichungssystems unverändert lassen, haben Einfluss auf den Wert der zugehörigen Determinante:

(i) Der Tausch zweier Zeilen führt zu einem Vorzeichenwechsel bei der Determinante.

(ii) Die Multiplikation einer Zeile mit einem nichtverschwindenden reellen Faktor bewirkt eine Multiplikation der Determinante mit diesem Faktor.

(iii) Die Addition eines Vielfachen einer Zeile zu einer anderen Zeile ändert den Wert der Determinante nicht.

Satz 5.9
$A \in \mathbb{R}^{n\times n}$ *ist invertierbar* $\Leftrightarrow \det A \neq 0$.

Bemerkung. Mit Hilfe von Satz 5.9 lässt sich die Lösbarkeit von Gleichungssystemen mit n Gleichungen und n Unbekannten sofort ablesen:
Sei $A \in \mathbb{R}^{n\times n}$ und $Ax = b \in \mathbb{R}^n$ das zu lösende Gleichungssystem. Falls A invertierbar ist, d.h. die zu A inverse Matrix $A^{-1} \in \mathbb{R}^{n\times n}$ existiert, dann ist $\operatorname{rang} A = n$ und das Gleichungssystem ist eindeutig in der folgenden Weise lösbar:

$$\begin{aligned} Ax &= b \\ \Leftrightarrow A^{-1}Ax &= A^{-1}b \\ \Leftrightarrow x &= A^{-1}b. \end{aligned}$$

Falls A nicht invertierbar ist, dann ist $k := \operatorname{rang} A < n$, und es existiert entweder eine Lösungsmenge mit $n - k$ freien Parametern, oder aber es gibt gar keine Lösung, je nachdem, ob $b \in \operatorname{Bild}(A)$ oder $b \notin \operatorname{Bild}(A)$ (vgl. die Zusammenfassung zum Gaußschen Eliminationsverfahren am Ende von Kapitel 4).

Satz 5.11 [Determinantenmultiplikationssatz]
Für $A, B \in \mathbb{R}^{n\times n}$ gilt

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

Speziell gilt für eine invertierbare Matrix $A \in \mathbb{R}^{n\times n}$

$$\det A^{-1} = \frac{1}{\det A}.$$

Satz 5.13 [Inversenberechnung]
Falls $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ invertierbar ist, dann gilt für die Inverse $A^{-1} = (c_{ij}) \in \mathbb{R}^{n\times n}$:

$$c_{ij} = \frac{(-1)^{i+j}}{\det A} \det \hat{A}_{ji},$$

wobei die Matrix $\hat{A}_{ji} \in \mathbb{R}^{(n-1)\times(n-1)}$ aus A hervorgeht, indem man die j-te Zeile und i-te Spalte streicht.

Satz 5.15 [Cramersche Regel]
Für eine invertierbare Matrix $A = (s_1, s_2, \ldots, s_n) \in \mathbb{R}^{n\times n}$ (mit Spalten $s_i \in \mathbb{R}^n$ für $i = 1, \ldots, n$) hat das lineare Gleichungssystem $Ax = b \in \mathbb{R}^n$ die eindeutig bestimmte Lösung

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

mit den Komponenten

$$x_i = \frac{\det A_i}{\det A} \text{ für } i = 1, \ldots, n,$$

wobei die Matrix $A_i \in \mathbb{R}^{n\times n}$ gegeben ist durch

$$A_i = (s_1, s_2, \ldots, s_{i-1}, b, s_{i+1}, \ldots, s_n) \text{ für } i = 1, \ldots, n.$$

Kapitel 6

Eigenwerte und Eigenvektoren

Eigenwerte und die zugehörigen *Eigenvektoren* sind z.B. zentral für das Verständnis von Resonanzphänomenen in der Physik. Aber auch bei der Linearisierung von komplizierten nichtlinearen Differentialgleichungen ist das Verständnis der zugehörigen Eigenwerte z.B. für Stabilitätsfragen und asymptotischen Untersuchungen wichtig. Hier werden diese Begriffe für lineare Abbildungen eingeführt und untersucht. Es stellt sich heraus, dass die Eigenwerte als Nullstellen des *charakteristischen Polynoms* bestimmbar sind, und diese Polynomgleichung ergibt sich aus dem Verschwinden einer Determinante. Wir ergänzen die Theorie durch hinreichende Kriterien dafür, dass Eigenvektoren linear unabhängig sind und (im besten Fall) eine Basis des $\mathbb{R}^n$ bilden. Das führt auf den Begriff der *Diagonalisierbarkeit* von Matrizen oder linearen Abbildungen. Der für die Theorie bedeutende *Satz von Cayley-Hamilton* zeigt, dass die resultierende Matrix, die sich ergibt, wenn man eine Matrix in ihr eigenes charakteristisches Polynom einsetzt, die Nullmatrix ist. In vielen Anwendungen aus der Physik und Mechanik spielen *symmetrische Matrizen* eine wichtige Rolle, und diese Matrizen lassen spezifischere Aussagen zu. So haben sie immer reelle Eigenwerte, was längst nicht für alle Matrizen stimmt, und die zugehörigen Eigenvektoren zu verschiedenen Eigenwerten stehen immer senkrecht aufeinander. Aus symmetrischen Matrizen kann man *quadratische Formen* generieren, für die dann analoge Aussagen gelten. Durch spezielle *Hauptachsentransformationen* lassen sich symmetrische Matrizen immer auf *Diagonalgestalt* bringen. Allgemeiner kann man quadratische Formen durch lineare Zusatzterme ergänzen, um sogenannte *Quadriken* zu erzeugen, deren Nullstellenmengen man im $\mathbb{R}^2$ und im $\mathbb{R}^3$ nicht nur allgemein durch bestimmte geometrische Formen klassifizieren kann, sondern die man auch durch konkrete Transformationen auf eine solche Normalform bringen kann. Zum Abschluss dieses Kapitels erklären und untersuchen wir noch die verschiedenen *Definitheitsbegriffe* für Matrizen und erläutern ein einfaches *Definitheitskriterium von Jacobi* für symmetrische Matrizen. Das ist insbesondere in der mehrdimensionalen Analysis für die Untersuchung von Extremwerten von reellwertigen Funktionen nützlich, die beispielsweise von zwei oder drei Variablen abhängen.

Definition 6.1
(i) Sei V ein Vektorraum (über $\mathbb{R}$) und $L : V \to V$ eine lineare Abbildung. Dann heißt $\lambda \in \mathbb{R}$ ein reeller *Eigenwert* von $L :\Leftrightarrow$

$$\exists\, v \in V \setminus \{0\} : \qquad L(v) = \lambda v.$$

Die Menge $E_L(\lambda) := \{w \in V : L(w) = \lambda w\}$ heißt der zu λ gehörige *Eigenraum*, die Elemente $w \in E_L(\lambda) \setminus \{0\}$ heißen die zu λ gehörigen *Eigenvektoren* von L.

(ii) Sei $A \in \mathbb{R}^{n\times n}$. Dann heißt $\lambda \in \mathbb{R}$ ein reeller *Eigenwert* von A :$\Leftrightarrow$

$$\exists\, x \in \mathbb{R}^n \setminus \{0\} : \qquad Ax = \lambda x.$$

Die Menge $E_A(\lambda) := \{y \in \mathbb{R}^n : Ay = \lambda y\}$ heißt der zu λ gehörige *Eigenraum*, die Elemente $y \in E_A(\lambda) \setminus \{0\}$ heißen die zu λ gehörigen *Eigenvektoren* von A.

Bemerkung. Einen systematischen Zugang zur Bestimmung der Eigenwerte einer Matrix $A \in \mathbb{R}^{n\times n}$ liefert das lineare Gleichungssystem

$$Ax = \lambda x, \tag{EWG}$$

wobei ein $\lambda \in \mathbb{R}$ gesucht ist, so dass für das Gleichungssystem (EWG) (mindestens) eine Lösung $x \in \mathbb{R}^n \setminus \{0\}$ existiert. Das Gleichungssystem (EWG) ist äquivalent zu

$$\begin{aligned} \lambda x - Ax &= 0 \\ \Leftrightarrow \lambda \operatorname{Id}_{\mathbb{R}^n} x - Ax &= 0 \\ \Leftrightarrow (\lambda \operatorname{Id}_{\mathbb{R}^n} - A)x &= 0. \end{aligned}$$

Wäre die Matrix $(\lambda \operatorname{Id}_{\mathbb{R}^n} - A) \in \mathbb{R}^{n\times n}$ invertierbar, dann wäre $x = 0 \in \mathbb{R}^n$ die eindeutige, aber von uns nicht gewünschte Lösung. Also sind nur solche $\lambda \in \mathbb{R}$ gesucht, so dass die Matrix $(\lambda \operatorname{Id}_{\mathbb{R}^n} - A)$ *nicht* invertierbar ist. Dies ist nach Satz 5.9 genau dann der Fall, wenn die zugehörige Determinante verschwindet, d.h. wir suchen $\lambda \in \mathbb{R}$, so dass

$$\det(\lambda \operatorname{Id}_{\mathbb{R}^n} - A) = 0. \tag{char}$$

Definition 6.3
Für eine Matrix

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \in \mathbb{R}^{n\times n}$$

heißt der Ausdruck

$$p_A(\lambda) := \det(\lambda \operatorname{Id}_{\mathbb{R}^n} - A) = \begin{vmatrix} (\lambda - a_{11}) & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & (\lambda - a_{22}) & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & \cdots & & (\lambda - a_{nn}) \end{vmatrix}$$

das *charakteristische Polynom* von A.

Bemerkung. Ist man (z.B. zum Bestimmen der Eigenwerte) nur an der Bestimmung der Nullstellen von $p_A(\lambda)$ interessiert, so kann man stattdessen auch $\tilde{p}_A(\lambda) := \det(A - \lambda \operatorname{Id}_{\mathbb{R}^n}) = (-1)^n \cdot p_A(\lambda)$ betrachten. Die hier verwendete (und auch in der Literatur stärker verbreitete) Definition von $p_A(\lambda)$ sorgt jedoch stets für ein normiertes Polynom (d.h. der Koeffizient vor x^n ist gleich 1), was für die Formulierung einiger der noch folgenden Sätze von Bedeutung ist.

Definition 6.5
Sei $A \in \mathbb{R}^{n\times n}$, $\lambda \in \mathbb{R}$ sei k-fache Nullstelle des charakteristischen Polynoms $p_A(\lambda)$. Dann heißt die Zahl $a_A(\lambda) := k$ die *algebraische Vielfachheit* des Eigenwerts λ, während die Dimension des zugehörigen Eigenraums $E_A(\lambda)$, also $g_A(\lambda) := \dim E_A(\lambda)$ die *geometrische Vielfachheit* von λ genannt wird.

Bemerkung. Man kann zeigen, dass die geometrische Vielfachheit eines Eigenwertes stets kleiner oder gleich der algebraischen Vielfachheit ist.

Satz 6.7
Falls $\lambda_1, \dots, \lambda_k$ paarweise verschiedene Eigenwerte von $A \in \mathbb{R}^{n\times n}$ mit zugehörigen Eigenvektoren $x_1, x_2, \dots, x_k \in \mathbb{R}^n$ sind, dann sind die Eigenvektoren $x_1, x_2, \dots, x_k$ linear unabhängig.

Korollar 6.8
Falls $A \in \mathbb{R}^{n\times n}$ n paarweise verschiedene Eigenwerte $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}$ mit zugehörigen Eigenvektoren $x_1, x_2, \dots, x_n \in \mathbb{R}^n$ hat, dann gilt

$$\mathbb{R}^n = \operatorname{span}\{x_1, x_2, \dots, x_n\},$$

d.h. das System $\{x_1, x_2, \dots, x_n\}$ bildet eine Basis des $\mathbb{R}^n$.

Satz 6.10 [Diagonalisierbarkeit]
Hat eine Matrix $A \in \mathbb{R}^{n\times n}$ n (nicht notwendig verschiedene) Eigenwerte $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ und sind die zugehörigen Eigenvektoren $x_1, x_2, \dots, x_n \in \mathbb{R}^n$ linear unabhängig, dann gilt für die Matrix $B := (x_1, x_2, \dots, x_n) \in \mathbb{R}^{n\times n}$

$$B^{-1}AB = \begin{pmatrix} \lambda_1 & 0 & \cdots & & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & & \\ & & & & \\ 0 & \cdots & & 0 & \lambda_n \end{pmatrix} =: \operatorname{Diag}(\lambda_1, \dots, \lambda_n).$$

Bemerkung.

(i) Gibt es im $\mathbb{R}^n$ eine Basis aus Eigenvektoren einer Matrix $A \in \mathbb{R}^{n\times n}$, dann heißt diese Matrix *diagonalisierbar* auf $\mathbb{R}$.

(ii) Besitzt eine Matrix $A \in \mathbb{R}^{n\times n}$ mindestens einen komplexen Eigenwert, so ist diese Matrix nicht diagonalisierbar auf $\mathbb{R}$ (vgl. Kapitel 7, Satz 7.30).

Korollar 6.11
Falls $A \in \mathbb{R}^{n\times n}$ n paarweise verschiedene Eigenwerte $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ hat, dann gilt für jede Matrix $B = (x_1, \dots, x_n) \in \mathbb{R}^{n\times n}$ von zugehörigen Eigenvektoren $x_1, \dots, x_n \in \mathbb{R}^n$

$$B^{-1}AB = \operatorname{Diag}(\lambda_1, \dots, \lambda_n).$$

Satz 6.13 [Cayley–Hamilton]
Für $A \in \mathbb{R}^{n\times n}$ gilt $p_A(A) = 0 \in \mathbb{R}^{n\times n}$, d.h. eine Matrix ist eine Nullstelle ihres charakteristischen Polynoms.

Definition 6.15
Wir identifizieren Spaltenvektoren

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

mit n-zeiligen Matrizen mit einer Spalte:

$$x \in \mathbb{R}^{n\times 1},$$

und fassen den entsprechenden Zeilenvektor als transponierte Matrix auf:

$$x^T := (x_1, x_2, \ldots, x_n) \equiv (x_1 \;\; x_2 \;\; \ldots \;\; x_n) \in \mathbb{R}^{1\times n}.$$

Damit ergibt sich insbesondere eine andere Darstellung für das Skalarprodukt im $\mathbb{R}^n$:

$$\begin{aligned} x \cdot y &= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \\ &= (x_1 \;\; x_2 \;\; \ldots \;\; x_n) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \\ &= x^T y \in \mathbb{R}. \end{aligned}$$

Lemma 6.16
Für $A \in \mathbb{R}^{m\times n}$ und $B \in \mathbb{R}^{n\times k}$ gilt

$$(AB)^T = B^T A^T \quad \in \mathbb{R}^{k\times m}.$$

Korollar 6.17
Für invertierbare $A \in \mathbb{R}^{n\times n}$ gilt

$$(A^{-1})^T = (A^T)^{-1}.$$

Definition 6.18
Sei $(a_{ij}) = A \in \mathbb{R}^{n\times n}$ symmetrisch (vgl. Definition 3.27). Dann heißt die Funktion $q_A : \mathbb{R}^n \to \mathbb{R}$ definiert durch

$$q_A(x) := x \cdot Ax = x^T Ax = \sum_{i,j=1}^{n} a_{ij} x_i x_j, \quad x \in \mathbb{R}^n,$$

die zu A gehörige *quadratische Form*.

Definition 6.20
$B \in \mathbb{R}^{n\times n}$ heißt *orthogonale Matrix* $:\Leftrightarrow$

$$B^T B = \mathrm{Id}_{\mathbb{R}^n}.$$

Bemerkung. Wenn $B \in \mathbb{R}^{n\times n}$ eine orthogonale Matrix ist, dann ergibt sich aus Satz 5.5 und aus dem Determinantenmultiplikationssatz, Satz 5.11:

$$(\det B)^2 = \det B^T \det B = \det(B^T B) = \det \mathrm{Id}_{\mathbb{R}^n} = 1,$$

also $\det B = 1$ oder $\det B = -1$.

Satz 6.22
Für $B \in \mathbb{R}^{n\times n}$ sind äquivalent:

(i) *B ist orthogonal*

(ii) $(Bx) \cdot (By) = x \cdot y \quad \forall\, x, y \in \mathbb{R}^n$

(iii) *$B = (b_1, b_2, \ldots, b_n)$ besteht aus Spalten $b_i, i = 1, \ldots, n$, die eine Orthonormalbasis des $\mathbb{R}^n$ bilden, d.h.*

$$b_i \cdot b_j = \delta_{ij} = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j, \end{cases}, \quad i, j = 1, \ldots, n.$$

Satz 6.23
Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch. Dann gilt:

(i) *A hat n (nicht notwendig verschiedene) reelle Eigenwerte $\lambda_1, \lambda_2, \ldots, \lambda_n$.*

(ii) *Falls $\lambda_i \neq \lambda_j$, dann gilt für die zugehörigen Eigenvektoren x_i und x_j:*

$$x_i \perp x_j.$$

(iii) *Falls $Ax_i = \lambda_i x_i$, $x_i \neq 0$, dann gilt*

$$(Az) \perp x_i \quad \forall\, z \perp x_i.$$

(iv) *Für alle $i = 1, \ldots, n$ gilt: Die Algebraische Vielfachheit von λ_i ist gleich der Geometrischen Vielfachheit von λ_i, und der $\mathbb{R}^n$ besitzt eine Orthonormalbasis bestehend aus (normierten) Eigenvektoren von A.*

Satz 6.24
Sei $A \in \mathbb{R}^{n\times n}$ symmetrisch, dann gilt für die quadratische Form $q_A(x) = x \cdot Ax = x^T Ax$ die Darstellung

$$q_A(x) = \sum_{i=1}^{n} \lambda_i \xi_i^2$$

mit $\xi_i := (B^T x)_i$, wobei $B = (b_1, b_2, \ldots, b_n)$ eine orthogonale Matrix ist, deren Spalten b_i Eigenvektoren von A sind, so dass $\{b_1, b_2, \ldots, b_n\}$ eine Orthonormalbasis des $\mathbb{R}^n$ bildet.

Bemerkung. (i) Zu jeder symmetrischen Matrix $A \in \mathbb{R}^{n\times n}$ gibt es demnach eine orthogonale Matrix $B \in \mathbb{R}^{n\times n}$, so dass $B^{-1}AB$ diagonal ist.

Gibt es umgekehrt zu einer beliebigen Matrix $A \in \mathbb{R}^{n\times n}$ eine orthogonale Matrix $B \in \mathbb{R}^{n\times n}$, so dass $B^{-1}AB$ diagonal ist, dann ist A symmetrisch, d.h. dann gilt $A = A^T$.

(ii) Falls A symmetrisch (oder auch nur diagonalisierbar) ist, dann gilt

$$\det A = \lambda_1 \cdot \lambda_2 \cdot \cdots \cdot \lambda_n, \tag{6.1}$$

wobei $\lambda_i, i = 1, \ldots, n$ die (nicht notwendig verschiedenen) reellen Eigenwerte von A sind. Die Beziehung (6.1) gilt übrigens auch für Matrizen, die nicht symmetrisch sind, wenn man komplexe Eigenwerte zulässt, siehe Kapitel 7, Corollar 7.34.

Definition 6.26
Ein *Hauptachsensystem* einer quadratischen Form $q_A(x) := x^T Ax$, $A = A^T \in \mathbb{R}^{n\times n}$, ist eine Orthonormalbasis $\{b_1, b_2, \ldots, b_n\}$ des $\mathbb{R}^n$, so dass

$$q_A(x) = \sum_{i=1}^{n} \lambda_i \xi_i^2 \qquad \text{für } \xi = B^T x \in \mathbb{R}^n,$$

wobei $B = (b_1, b_2, \ldots, b_n)$ die orthogonale Matrix mit Spaltenvektoren b_i, $i = 1, \ldots, n$.

Satz 6.27 [Hauptachsentransformation]
Für jede quadratische Form $q_A(x) := x^T Ax$, $A = A^T \in \mathbb{R}^{n\times n}$, gibt es (mindestens ein) Hauptachsensystem. Zu dessen Berechnung bestimmt man

1. *alle (reellen) Eigenwerte λ_i, $\lambda_i \neq \lambda_j$ für $i \neq j$, $i, j \in \{1, \ldots, r\}$, $r \leq n$,*
2. *Orthonormalbasen $\{b_1^{(i)}, \ldots, b_{k_i}^{(i)}\}$ der zugehörigen Eigenräume $E_A(\lambda_i)$, $i = 1, \ldots, r$,*
3. *die orthogonale Matrix B durch*

$$B := (b_1^{(1)}, \ldots, b_{k_1}^{(1)}, b_1^{(2)}, \ldots, b_{k_2}^{(2)}, \ldots, b_1^{(r)}, \ldots, b_{k_r}^{(r)}\},$$

so dass

$$B^T AB = \begin{pmatrix}
\lambda_1 & 0 & \cdots & & & & & & & & 0 \\
\vdots & \ddots & & & & & & & & & \vdots \\
0 & & \lambda_1 & 0 & & & & & & & 0 \\
0 & & 0 & \lambda_2 & 0 & & & & & & 0 \\
\vdots & & & & \ddots & & & & & & \vdots \\
0 & & & & & \lambda_2 & 0 & & & & 0 \\
\vdots & & & & & & \ddots & & & & \vdots \\
\vdots & & & & & & & \ddots & & & \vdots \\
0 & & & & & & & 0 & \lambda_r & 0 & \vdots \\
\vdots & & & & & & & & & \ddots & \\
0 & & & & & & & & & 0 & \lambda_r
\end{pmatrix}.$$

Definition 6.29
(i) Eine *Quadrik* (=*Hyperfläche zweiter Ordnung*) ist eine Menge $Q \subset \mathbb{R}^n$ definiert durch

$$Q := Q(A, a, \beta) := \{x \in \mathbb{R}^n : q(x) := x^T A x + a^T x + \beta = 0\},$$

wobei $A = A^T \in \mathbb{R}^{n\times n}$, $a \in \mathbb{R}^n$ und $\beta \in \mathbb{R}$.

(ii) Eine Quadrik liegt in *Normalform* vor, wenn der Ausdruck $x^T A x$ rein quadratisch ist (d. h. A ist Diagonalmatrix) und $a^T x + \beta$ durch keinen Basiswechsel und Translation des Koordinatenursprungs verkürzt werden kann.

Bemerkung. Die **Transformation einer Quadrik auf Normalform** geschieht in zwei Schritten:
1. Eine Hauptachsentransformation des quadratischen Anteils führt nach Satz 6.27 auf

$$q(x) = \tilde{q}(\xi) = \sum_{i=1}^{n} \lambda_i \xi_i^2 + c^T \xi + \beta$$

mit einem Vektor $c^T = (c_1, c_2, \ldots, c_n) := a^T B \in \mathbb{R}^{1\times n} \simeq \mathbb{R}^n$.

2. Die Substitution (quadratische Ergänzung)

$$\eta_i := \begin{cases} \xi_i & \text{falls } \lambda_i = 0 \\ \xi_i + \frac{c_i}{2\lambda_i} & \text{falls } \lambda_i \neq 0 \end{cases}$$

verschiebt das Koordinatensystem in einen anderen Ursprung und führt unter der Annahme $\lambda_1, \ldots, \lambda_s \neq 0$, $\lambda_{s+1} = \cdots = \lambda_n = 0$ auf

$$q(x) \quad (= \tilde{q}(\xi)) \quad = q^*(\eta) := \sum_{i=1}^{s} \lambda_i \eta_i^2 + \sum_{k=s+1}^{n} \mu_k \eta_k + \gamma = 0.$$

Falls $\mu_l \neq 0$ für nur *ein* $l \in \{s+1, \ldots, n\}$, dann führt eine weitere Substitution

$$\zeta_i := \begin{cases} \eta_i & \text{für } i \neq l \\ \eta_l + \frac{\gamma}{\mu_l} & \text{für } i = l \end{cases}$$

auf die Form

$$q(x) \quad (= \tilde{q}(\xi) = q^*(\eta)) \quad = \hat{q}(\zeta) := \sum_{i=1}^{s} \lambda_i \zeta_i^2 + \sum_{k=s+1}^{n} \mu_k \zeta_k = 0.$$

Eine Übersicht mit der Klassifikation sämtlicher Quadriken in $\mathbb{R}^2$ und $\mathbb{R}^3$ finden Sie auf der Internetseite zu dieser Vorlesung.

Definition 6.31
(i) Eine Matrix $A \in \mathbb{R}^{n\times n}$ heißt *positiv definit* — man schreibt dann auch $A > 0$ — (bzw. *negativ definit*, also $A < 0$)

$$:\Longleftrightarrow \qquad x^T Ax = x \cdot Ax > 0 \text{ für alle } x \in \mathbb{R}^n, x \neq 0 \qquad (\text{bzw. } x^T Ax = x \cdot Ax < 0 \text{ für alle } x \neq 0).$$

(ii) Eine Matrix $A \in \mathbb{R}^{n\times n}$ heißt *positiv semidefinit* — man schreibt dann auch $A \geq 0$ — (bzw. *negativ semidefinit*, also $A \leq 0$)

$$:\Longleftrightarrow \qquad x^T Ax = x \cdot Ax \geq 0 \text{ für alle } x \in \mathbb{R}^n \qquad (\text{bzw. } x^T Ax = x \cdot Ax \leq 0 \text{ für alle } x \in \mathbb{R}^n).$$

(iii) $A = A^T \in \mathbb{R}^{n\times n}$ und die zugehörige quadratische Form $q_A(x) = x^T Ax$ heißen *positiv definit* (bzw. *negativ definit*)

$$:\Longleftrightarrow \qquad q_A(x) > 0 \text{ für alle } x \neq 0 \qquad (\text{bzw. } q_A(x) < 0 \text{ für alle } x \neq 0).$$

(iv) $A = A^T \in \mathbb{R}^{n\times n}$ und $q_A(x) = x^T Ax$ heißen *positiv semidefinit* (bzw. *negativ semidefinit*)

$$:\Longleftrightarrow \qquad q_A(x) \geq 0 \text{ für alle } x \in \mathbb{R}^n \qquad (\text{bzw. } q_A(x) \leq 0 \text{ für alle } x \in \mathbb{R}^n).$$

(v) Erfüllt eine Matrix $A \in \mathbb{R}^{n\times n}$ keine der Bedingungen in (i) oder (ii), dann heißt A *indefinit.*

Satz 6.33 [Charakterisierung der positiven Definitheit]
(i) *Eine* Diagonalmatrix

$$A = \mathrm{Diag}(\alpha_1, \alpha_2, \ldots, \alpha_n) = \begin{pmatrix} \alpha_1 & 0 & \cdots & \\ 0 & \alpha_2 & 0 & \\ \vdots & & \ddots & \\ 0 & & & \alpha_n \end{pmatrix} \in \mathbb{R}^{n\times n}$$

ist positiv definit $\Leftrightarrow \alpha_i > 0$ *für alle* $i = 1, \ldots, n$.

(ii) *Für jede invertierbare Matrix* $P \in \mathbb{R}^{n\times n}$ *gilt:*
$A \in \mathbb{R}^{n\times n}$ *ist positiv definit* $\Leftrightarrow$ *die Matrix* $P^T AP \in \mathbb{R}^{n\times n}$ *ist positiv definit.*
(iii) $A = A^T \in \mathbb{R}^{n\times n}$ *ist positiv definit* $\Leftrightarrow$ *alle Eigenwerte* λ_i, $i = 1, \ldots, n$, *von* A *sind positiv:* $\lambda_i > 0$ *für alle* $i = 1. \ldots, n$.

Bemerkung. 1. Satz 6.33 verlangt keine *orthogonale* Matrix P, jede beliebige, invertierbare Matrix P, mit deren Hilfe man die Matrix A durch Bildung von $P^T AP$ auf Diagonalgestalt bringen kann, erlaubt eine einfache Beurteilung der Definitheit von A.

2. Für eine positiv definite Matrix $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ gilt

$$0 < e_i^T A e_i = a_{ii} \quad \forall\, i = 1, \dots, n,$$

d.h. *alle* Diagonalelemente müssen positiv sein.

3. Für den Fall $n = 2$ ist eine symmetrische Matrix

$$A = \begin{pmatrix} a & b \\ b & d \end{pmatrix} \in \mathbb{R}^{2\times 2}$$

genau dann positiv definit, wenn $a > 0$ und $\det A = ad - b^2 > 0$.

4. Für den Fall $n = 3$ ist eine symmetrische Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} \in \mathbb{R}^{3\times 3}$$

genau dann positiv definit, wenn $a_{11} > 0$ und $a_{11}a_{22} - a_{12}^2 > 0$ und $\det A > 0$.

Satz 6.35 [Definitheitskriterium von C.G. Jacobi**]**
Sei $A = (a_{ij}) \in \mathbb{R}^{n\times n}$ symmetrisch, d.h. $A = A^T$. Es gilt:
A ist positiv definit genau dann, wenn

$$a_{11} > 0, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} > 0, \dots, \quad \textit{und} \quad \det A > 0.$$

Korollar 6.37
Sei $A = (a_{ij}) \in \mathbb{R}^{n\times n}$. Die Matrix $B := \frac{1}{2}\left(A + A^T\right) = (b_{ij})$ ist der symmetrische Anteil *von A. Es sind äquivalent:*

(i) A ist positiv definit

(ii) B ist positiv definit

(iii) B erfüllt das Kriterium von Jacobi (Satz 6.35), d.h.

$$b_{11} > 0, \quad \begin{vmatrix} b_{11} & b_{12} \\ b_{12} & b_{22} \end{vmatrix} > 0, \quad \begin{vmatrix} b_{11} & b_{12} & b_{13} \\ b_{12} & b_{22} & b_{23} \\ b_{13} & b_{23} & b_{33} \end{vmatrix} > 0, \dots, \quad \textit{und} \quad \det B > 0.$$

Kapitel 7

Komplexe Zahlen, Fundamentalsatz der Algebra und Jordannormalformen

Schon im sechsten Kapitel wurde an Beispielen deutlich, dass nicht alle Matrizen reelle Eigenwerte haben, so etwa Drehmatrizen im $\mathbb{R}^2$ für einen positiven Drehwinkel echt kleiner als π. Damit wir trotzdem Eigenwerte für beliebige Matrizen finden, vergrößern wir den Raum der reellen Zahlen auf den der *komplexen Zahlen*. Ursprünglich wurde diese Erweiterung im Rahmen der Elektrizitätslehre nötig, um Phasen von Schwingungen mathematisch zu beschreiben, mittlerweile haben komplexe Zahlen vielschichtige Anwendungen in der Mathematik, in den Natur- und Ingenieurwissenschaften. Wir beschränken uns hier auf grundlegende Eigenschaften dieses erweiterten Zahlenraums und nutzen u.a. auch die *Polardarstellung*, um schließlich über die allgemeine *Lösungsformel für quadratische Gleichungen* auch für negative Diskriminanten den *Fundamentalsatz der Algebra* zu behandeln. Der sichert, dass jedes Polynom, also auch das charakteristische Polynom einer beliebigen Matrix, komplexe Nullstellen besitzt. Daraus resultieren neue *Diagonalisierbarkeitskriterien* für Matrizen. Auch für allgemeine quadratische Matrizen, die nicht notwendig diagonalisierbar sein müssen, gibt es eine Normalform, die sogenannte *Jordansche Normalform*, deren Ermittlung das Hauptziel dieses Kapitels ist. Dazu notwendig ist die Untersuchung eines weiteren Polynoms, des *Minimalpolynoms*, welches z.B. das charakteristische Polynom teilt. Mit Hilfe dieser beiden Polynome lassen sich schnell die möglichen Jordanschen Normalformen einer gegebenen Matrix eingrenzen. Die tatsächliche *Transformation auf Jordannormalform* ist dann ein längere Algorithmus, den wir in allen Einzelheiten studieren werden.

Definition 7.1
Die Menge der *komplexen Zahlen* $\mathbb{C} := (\mathbb{R}^2, +, \cdot)$ ist die Menge der geordneten Paare $(x, y)^T \in \mathbb{R}^2$, auf der Addition "+" und Multiplikation "·" in der folgenden Weise definiert sind:
Für $z_1 = (x_1, y_1)^T$ und $z_2 = (x_2, y_2)^T \in \mathbb{R}^2$ ist die Summe

$$z_1 + z_2 := \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \end{pmatrix} \in \mathbb{R}^2$$

und das Produkt

$$z_1 z_2 := z_1 \cdot z_2 := \begin{pmatrix} x_1 x_2 - y_1 y_2 \\ x_1 y_2 + x_2 y_1 \end{pmatrix} \in \mathbb{R}^2.$$

Die Null ($0 \in \mathbb{C}$) ist $(0, 0)^T \in \mathbb{R}^2$ als Neutrales Element der Addition, die Eins ($1 \in \mathbb{C}$) ist $(1, 0)^T \in \mathbb{R}^2$ als Neutrales Element der Multiplikation.

Bemerkung. (i) $\mathbb{C}$ erfüllt die Additionsaxiome (A1)–(A4) (vgl. Definition 3.1 oder Kapitel 1 aus der Differential- und Integralrechnung I).

(ii) $\mathbb{C}$ erfüllt die Multiplikationsaxiome (M1)–(M4) sowie (AM) (vgl. Kapitel 1 aus der Differential- und Integralrechnung I).

(iii) Insbesondere gilt

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

Definition 7.2
(i) Die komplexe Zahl

$$1 := \begin{pmatrix} 1 \\ 0 \end{pmatrix} \in \mathbb{C} \qquad \text{ist die } \textit{reelle Einheit},$$

die komplexe Zahl

$$i := \begin{pmatrix} 0 \\ 1 \end{pmatrix} \in \mathbb{C} \qquad \text{nennt man die } \textit{imaginäre Einheit}.$$

Mit

$$x := \begin{pmatrix} x \\ 0 \end{pmatrix} \in \mathbb{C}$$

identifiziert man $\mathbb{R}$ als Untermenge von $\mathbb{C}$, und mit

$$iy := \begin{pmatrix} 0 \\ y \end{pmatrix} \in \mathbb{C}$$

bezeichnet man die (rein) *imaginären Zahlen*.
Wir schreiben

$$z = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ y \end{pmatrix} = x + iy \quad \in \mathbb{C}$$

und nennen $\mathrm{Re}(z) := x$ den *Realteil* von z und $\mathrm{Im}(z) := y$ den *Imaginärteil* von z.

(ii) Die komplexe Zahl $\bar{z} := x - iy$ heißt die zu $z = x + iy \in \mathbb{C}$ *konjugiert komplexe Zahl*.

Bemerkung. Es gilt nach Bemerkung (iii)

$$i^2 = i \cdot i = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix} = -1,$$

und daher allgemein für $z_1 := x_1 + iy_1$ und $z_2 := x_2 + iy_2 \in \mathbb{C}$

$$z_1 z_2 = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1).$$

Definition 7.4
(i) $|z| := \sqrt{x^2 + y^2} = \sqrt{z\bar{z}}$ ist der *Betrag* von $z = x + iy \in \mathbb{C}$.

(ii) Für $0 \neq z \in \mathbb{C}$ ist

$$z = |z|(\cos\phi + i\sin\phi)$$

die *Polardarstellung* von $z \in \mathbb{C}$ mit dem orientierten Winkel

$$\phi = \measuredangle\,(1,0)^T, (x,y)^T) =: \arg z$$

als *Argument* von z.

Satz 7.5 [Eigenschaften des komplexen Betrages]
Für alle

$$\begin{aligned} z &= x + iy = |z|(\cos\phi + i\sin\phi) \in \mathbb{C} \ \textit{ und} \\ w &= u + iv = |w|(\cos\psi + i\sin\psi) \in \mathbb{C} \end{aligned}$$

gilt:

(i) $|z| \geq 0$ *und* $|z| = 0 \Leftrightarrow z = 0$

(ii) $|zw| = |z||w|, \qquad |z/w| = |z|/|w|$ *für* $w \neq 0, \qquad |z| = |\bar{z}|$

(iii) $|z + w| \leq |z| + |w|$

(iv) $\max\{|\operatorname{Re}(z)|, |\operatorname{Im}(z)|\} \leq |z| \leq |\operatorname{Re}(z)| + |\operatorname{Im}(z)|$

(v) $1/z = \bar{z}/|z|^2$

(vi) $zw = |z||w|(\cos(\phi + \psi) + i\sin(\phi + \psi))$

Bemerkung. (i) Für festes $w \in \mathbb{C}$ liefert die Abbildung $z \mapsto zw$ eine Drehung des Vektors z um den Winkel ψ und gleichzeitig eine Streckung um den Faktor $|w|$.

(ii) Die Abbildung $z \mapsto 1/z = \bar{z}/|z|^2$ ist eine Spiegelung von z am Einheitskreis gefolgt von einer Spiegelung an der reellen Achse.

Definition 7.6

$$e^z := \sum_{n=0}^{\infty} \frac{1}{n!} z^n$$

ist die *komplexe Exponentialfunktion*.

Satz 7.7 [Eigenschaften von e^z, $z \in \mathbb{C}$] (i) $e^{z+w} = e^z \cdot e^w \quad \forall\, z, w \in \mathbb{C}$

(ii) $e^{iy} = \cos y + i \sin y \quad \forall\, y \in \mathbb{R}$

(iii) $S^1 := \{e^{i\phi} : \phi \in [0, 2\pi)\}$ *ist die* Einheitskreislinie

(iv) $|e^z| = e^{\mathrm{Re}(z)}$

(v) $e^z = e^w \quad \Rightarrow \quad z = w + k \cdot 2\pi i$ *für ein* $k \in \mathbb{Z}$.

Korollar 7.8 [Polarkoordinaten in $\mathbb{C}$]
Für alle $z \in \mathbb{C}$ *gilt* $z = re^{i\phi}$ *mit* $r := |z|$ *und* $\phi := \arg z$.

Bemerkung. Eine quadratische Gleichung der Form

$$ax^2 + bx + c = 0 \tag{7.1}$$

besitzt die Lösungen

$$x_{1/2} := \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad \in \mathbb{R} \text{ für } b^2 - 4ac \geq 0.$$

Falls nun $b^2 - 4ac < 0$ ist, dann liefert das Paar konjugiert komplexer Zahlen

$$z_{1/2} := \frac{-b \pm i\sqrt{-(b^2 - 4ac)}}{2a} \quad \in \mathbb{C}$$

zwei komplexe Lösungen der quadratischen Gleichung (7.1).

Satz 7.11 [Identitätssatz für komplexe Polynome]
(i) *Hat ein komplexes Polynom*

$$P(z) := a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0, \quad z, a_i \in \mathbb{C} \text{ für } i = 0, \ldots, n, \quad n \in \mathbb{N},$$

$n + 1$ *verschiedene Nullstellen, dann ist* $P \equiv 0$.

(ii) *Gilt*

$$a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0 = b_n z^n + b_{n-1} z^{n-1} + \cdots + b_1 z + b_0$$

für $n + 1$ *verschiedene Zahlen* $z = \zeta_1, \zeta_2, \ldots, \zeta_{n+1} \in \mathbb{C}$, *so folgt*

$$a_n = b_n, \quad a_{n-1} = b_{n-1}, \quad \ldots, \quad a_0 = b_0.$$

Lemma 7.12 [Polynomnullstellen in $\mathbb{C}$]
Das Polynom

$$P_n(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0, \quad a_n \neq 0, \quad n \in \mathbb{N},$$

hat höchstens n *verschiedene Nullstellen* $\zeta_1, \zeta_2, \ldots, \zeta_m$, $0 \leq m \leq n$, *und es gilt*

$$P_n(z) = (z - \zeta_1)^{\alpha_1} (z - \zeta_2)^{\alpha_2} \cdots (z - \zeta_m)^{\alpha_m} Q_l(z),$$

mit

$$\alpha_1 + \alpha_2 + \cdots + \alpha_m \leq n$$

und dem Polynom Q_l *vom Grade* $l := n - \alpha_1 - \cdots - \alpha_m$ *mit* $Q_l(z) \neq 0$ *für alle* $z \in \mathbb{C}$.

Satz 7.13 [Fundamentalsatz der Algebra]
Für jedes komplexe Polynom

$$P_n(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0, \quad z, a_i \in \mathbb{C} \text{ für } i = 0, \ldots, n, \quad a_n \in \mathbb{C} \setminus \{0\},\ n \in \mathbb{N},$$

gibt es eine Nullstelle, und es gilt die Faktorisierung

$$P_n(z) = a_n (z - \zeta_1)^{\alpha_1} (z - \zeta_2)^{\alpha_2} \cdots (z - \zeta_m)^{\alpha_m}$$

mit $P_n(\zeta_k) = 0$ *für alle* $k = 1, \ldots, m$ *und mit*

$$\alpha_1 + \alpha_2 + \cdots + \alpha_m = n.$$

Bemerkung. Zählt man die Nullstellen ζ_k gemäß ihrer Vielfachheit α_k, dann hat jedes komplexe Polynom $P_n(z)$ vom Grade n *genau* n Nullstellen in $\mathbb{C}$. In Lemma 7.12 ist insbesondere stets $l = 0$ und $Q_l(z) \equiv a_n \neq 0$.

Satz 7.14 [Faktorisierung reeller Polynome]
Jedes reelle Polynom

$$P_n(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0, \quad x, a_i \in \mathbb{R} \text{ für } i = 0, \ldots, n, \quad a_n \in \mathbb{R} \setminus \{0\},\ n \in \mathbb{N},$$

besitzt die Faktorisierung

$$P_n(x) = a_n (x - c_1)^{k_1} (x - c_2)^{k_2} \cdots (x - c_r)^{k_r} (x^2 + p_1 x + q_1)^{l_1} \cdots (x^2 + p_s x + q_s)^{l_s}$$

mit $r, s \geq 0$, $k_1 + k_2 + \ldots + k_r + 2(l_1 + l_2 + \ldots + l_s) = n$, $P_n(c_k) = 0$, $c_k \in \mathbb{R}$, $k = 1, \ldots, r$, *und*

$$x^2 + p_i x + q_i \neq 0 \text{ für alle } i = 1, \ldots, s \text{ für alle } x \in \mathbb{R}.$$

Korollar 7.16 [Charakteristisches Polynom zerfällt in $\mathbb{C}$]
Für alle $A \in \mathbb{C}^{n\times n}$ gilt

$$p_A(\lambda) = (\lambda - \lambda_1)^{\alpha_1}(\lambda - \lambda_2)^{\alpha_2} \cdots (\lambda - \lambda_m)^{\alpha_m}, \qquad \lambda \in \mathbb{C},$$

mit Eigenwerten $\lambda_k \in \mathbb{C}$, d.h. mit

$$p_A(\lambda_k) = 0 \quad \textit{für alle} \;\; k = 1, \ldots, m$$

und mit

$$\alpha_1 + \alpha_2 + \cdots + \alpha_m = n.$$

Speziell gilt dies auch für $A \in \mathbb{R}^{n\times n}$ mit $\lambda \in \mathbb{R}$, $\lambda_i \in \mathbb{C}$.

Korollar 7.17 [Diagonalisierbarkeit]
Falls für $A \in \mathbb{C}^{n\times n}$

$$p_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n) \qquad \lambda \in \mathbb{C},$$

mit $\lambda_i \neq \lambda_k$ für $i \neq k$, dann gibt es eine invertierbare Matrix $B \in \mathbb{C}^{n\times n}$, so dass

$$B^{-1}AB = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \\ & & \ddots & \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}. = \mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$$

Korollar 7.18 [Existenz von Eigenwerten]
Jede Matrix $A \in \mathbb{C}^{n\times n}$ besitzt mindestens einen Eigenwert in $\mathbb{C}$. Jede reelle Matrix $A \in \mathbb{R}^{(2n-1)\times(2n-1)}$, $n \in \mathbb{N}$, besitzt mindestens einen reellen Eigenwert.

Definition 7.20
Für $A \in \mathbb{C}^{n\times n}$ heißt das Polynom

$$m_A(z) = z^r + a_{r-1}z^{r-1} + \cdots + a_1 z + a_0, \qquad z \in \mathbb{C}$$

Minimalpolynom von $A :\Longleftrightarrow$

(i) $m_A(A) = 0$

(ii) Für alle Polynome q mit $q(A) = 0$ gilt $\mathrm{Grad}(m_A) \leq \mathrm{Grad}(q)$.

Satz 7.21 [Minimalpolynom versus Charakteristisches Polynom]
Das Minimalpolynom $m_A(z)$ einer Matrix $A \in \mathbb{C}^{n\times n}$ teilt jedes Polynom $P(z)$, welches $P(A) = 0$ erfüllt. Speziell teilt $m_A(z)$ das charakteristische Polynom $p_A(z)$. Darüberhinaus haben $m_A(z)$ und $p_A(z)$ dieselben Linearfaktoren (allerdings in möglicherweise unterschiedlichen Potenzen). Insbesondere ist $\lambda \in \mathbb{C}$ genau dann Eigenwert von A, wenn $m_A(\lambda) = 0$.

Lemma 7.23 [Determinanten von Blockmatrizen]
Seien $A \in \mathbb{C}^{k\times k}$, $C \in \mathbb{C}^{m\times m}$ und $B \in \mathbb{C}^{k\times m}$. Dann gilt für die Matrix

$$M := \begin{pmatrix} A & B \\ 0 & C \end{pmatrix} \in \mathbb{C}^{(k+m)\times(k+m)}$$

$$\det M = \det A \cdot \det C.$$

Lemma 7.25 [Minimalpolynome von Blockmatrizen]
Seien $A \in \mathbb{C}^{k\times k}$ und $C \in \mathbb{C}^{m\times m}$. Dann ist das Minimalpolynom $m_M(\lambda)$ der Matrix

$$M := \begin{pmatrix} A & 0 \\ 0 & C \end{pmatrix} \in \mathbb{C}^{(k+m)\times(k+m)}$$

das kleinste gemeinsame Vielfache der Minimalpolynome $m_A(\lambda)$ und $m_C(\lambda)$ von A bzw. C.

Satz 7.27 [Algebraische versus geometrische Vielfachheit]
Eine Matrix $A \in \mathbb{C}^{n\times n}$ ist genau dann diagonalisierbar, wenn für jeden Eigenwert von A die algebraische gleich der geometrischen Vielfachheit ist.

Lemma 7.28 [$m_A = p_A$]
Hat $A \in \mathbb{C}^{n\times n}$ genau n paarweise verschiedene Eigenwerte, so gilt $m_A(\lambda) = p_A(\lambda)$.

Satz 7.30 [Diagonalisierbarkeit]
$A \in \mathbb{C}^{n\times n}$ ist diagonalisierbar $\Leftrightarrow m_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_m)$ für paarweise verschiedene $\lambda_i \in \mathbb{C}$, $i = 1, \ldots, m$.

Satz 7.32 [Jordannormalform]
Für $A \in \mathbb{C}^{n\times n}$ mit charakteristischem Polynom

$$p_A(\lambda) = (\lambda - \lambda_1)^{\alpha_1}(\lambda - \lambda_2)^{\alpha_2} \cdots (\lambda - \lambda_m)^{\alpha_m}$$

und Minimalpolynom

$$m_A(\lambda) = (\lambda - \lambda_1)^{\beta_1}(\lambda - \lambda_2)^{\beta_2} \cdots (\lambda - \lambda_m)^{\beta_m},$$

wobei die λ_i paarweise disjunkt seien, gibt es eine invertierbare Matrix $B \in \mathbb{C}^{n\times n}$, so dass

$$B^{-1}AB = J_A := \begin{pmatrix} J_{11} & 0 & \cdots & & & & & & & & & \\ 0 & J_{12} & 0 & \cdots & & & & & & & & \\ \vdots & & \ddots & & & & & & & & & \\ & & & J_{1g_A(\lambda_1)} & 0 & \cdots & & & & & & \\ & & & 0 & J_{21} & 0 & & & & & & \\ & & & & & \ddots & & & & & & \\ & & & & & & J_{2g_A(\lambda_2)} & 0 & & & & \\ & & & & & & & \ddots & & & & \\ & & & & & & & & J_{m1} & 0 & & \\ & & & & & & & & & \ddots & & \\ & & & & & & & & 0 & & & J_{mg_A(\lambda_m)} \end{pmatrix},$$

wobei ein Jordanblock *J_{ij} für jedes $i = 1, \ldots, m$ von der folgenden Gestalt ist:*

$$J_{ij} = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & \\ 0 & \lambda_i & 1 & 0 & \cdots \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{pmatrix}, \qquad j = 1, \ldots, g_A(\lambda_i),$$

wobei $g_A(\lambda_i) = \dim E_A(\lambda_i)$ die geometrische Vielfachheit des Eigenwerts λ_i bezeichnet.
Dabei existiert für jedes i mindestens ein j mit $J_{ij} \in \mathbb{C}^{\beta_i\times\beta_i}$, und für jedes $j = 1, \ldots, g_A(\lambda_i)$ gibt es $k \in \{1, \ldots, \beta_i\}$ mit $J_{ij} \in \mathbb{C}^{k\times k}$. Weiterhin ist für jedes $i = 1, \ldots, m$

$$\begin{pmatrix} J_{i1} & 0 & \cdots & \\ 0 & J_{i2} & 0 & \\ & & \ddots & \\ & & 0 & J_{ig_A(\lambda_i)} \end{pmatrix} \in \mathbb{C}^{\alpha_i\times\alpha_i}.$$

Die Zahl der Jordanblöcke $J_{ij} \in \mathbb{C}^{k\times k}$ ist für jedes $k \in \{1, \ldots, \beta_i\}$ eindeutig bestimmt, und damit die Jordannormalform *J_A bis auf die Reihenfolge der Jordanblöcke.*

Korollar 7.34
Für $A \in \mathbb{C}^{n\times n}$ gilt

$$\det A = \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_n,$$

wobei $\lambda_i \in \mathbb{C}$, $i = 1, \dots, n$ die (nicht notwendig verschiedenen) Eigenwerte von A sind.

Für die Berechnung der Transformationsmatrix $B \in \mathbb{C}^{n\times n}$, so dass $B^{-1}AB = J_A$ kann man z.B. folgendermaßen vorgehen:

1. Schritt. Bestimme das charakteristische Polynom

$$p_A(\lambda) = (\lambda - \lambda_1)^{\alpha_1}(\lambda - \lambda_2)^{\alpha_2} \cdots (\lambda - \lambda_m)^{\alpha_m}, \qquad \sum_{i=1}^{m} \alpha_i = n,$$

und das Minimalpolynom

$$m_A(\lambda) = (\lambda - \lambda_1)^{\beta_1}(\lambda - \lambda_2)^{\beta_2} \cdots (\lambda - \lambda_m)^{\beta_m}, \qquad \beta_i \le \alpha_i \ \text{ für alle } \ i = 1, \dots, m.$$

2. Schritt. Berechne die unter A invarianten Unterräume

$$W_A(\lambda_i) := \operatorname{Kern}[(A - \lambda_i \operatorname{Id}_{\mathbb{C}^n})^{\beta_i}]$$

durch Lösen von linearen Gleichungssystemen. Beachte, dass $E_A(\lambda_i) \subset W_A(\lambda_i)$.

3. Schritt. Suche einen Vektor $x_1 \in W_A(\lambda_1)$ mit

$$\begin{aligned} (A - \lambda_1 \operatorname{Id}_{\mathbb{C}^n})^{\beta_1} x_1 &= 0 \quad \text{und} \\ (A - \lambda_1 \operatorname{Id}_{\mathbb{C}^n})^{\gamma} x_1 &\neq 0 \quad \text{für alle } \ \gamma = 1, 2, \dots, \beta_1 - 1. \end{aligned}$$

Dann definiere eine absteigende Kette von nichtverschwindenden Vektoren in der folgenden Weise

$$\begin{aligned} x_1^{(\beta_1)} &:= x_1 \\ x_1^{(\beta_1-1)} &:= (A - \lambda_1 \operatorname{Id}_{\mathbb{C}^n}) x_1 \\ x_1^{(\beta_1-2)} &:= (A - \lambda_1 \operatorname{Id}_{\mathbb{C}^n})^2 x_1 \\ &\vdots \\ x_1^{(1)} &:= (A - \lambda_1 \operatorname{Id}_{\mathbb{C}^n})^{\beta_1-1} x_1. \end{aligned} \tag{7.2}$$

Es folgt aufgrund der Konstruktion

$$A x_1^{(j)} = \begin{cases} \lambda_1 x_1^{(j)} & \text{für } \ j = 1 \\ \lambda_1 x_1^{(j)} + x_1^{(j-1)} & \text{für } \ j = 2, 3, \dots, \beta_1, \end{cases} \tag{7.3}$$

Man kann zeigen, dass die Vektoren $x_1^{(j)}$, $j = 1, 2, \dots, \beta_1$, linear unabhängig sind, und wegen (7.3) gilt für

$$U_1 := \operatorname{span}\{x_1^{(1)}, x_1^{(2)}, \dots, x_1^{(\beta_1)}\}$$

die Invarianzeigenschaft

$$A(U_1) \subset U_1.$$

Wegen (7.3) gilt dann

$$A_1 := A|_{U_1} = \left.\begin{pmatrix} \lambda_1 & 1 & 0 & \cdots & 0 \\ 0 & \lambda_1 & 1 & \ddots & \\ & & \ddots & \ddots & 0 \\ & & 0 & \lambda_1 & 1 \\ & & & & \lambda_1 \end{pmatrix}\right\} \beta_1 \text{ Zeilen}$$

$$= \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_1 & 0 & & \\ & & \ddots & & \\ & & 0 & \lambda_1 & 0 \\ & & & 0 & \lambda_1 \end{pmatrix} + \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}$$

in der Basis $\{x_1^{(1)}, x_1^{(2)}, \ldots, x_1^{(\beta_1)}\}$. Wir bemerken, dass $U_1 \cap E_A(\lambda_1) = \operatorname{span}\{x_1^{(1)}\}$.

4. Schritt. Falls $W_A(\lambda_1) \not\subset U_1$, also wenn $g_A(\lambda_1) > 1$, dann wähle unter den nicht in U_1 enthaltenen Basisvektoren von $W_A(\lambda_1)$ einen Vektor x_2 aus, für den ein maximales (!) k_2, $1 \le k_2 \le \beta_1$, existiert, so dass einerseits

$$\begin{aligned} (A - \lambda_1 \operatorname{Id}_{\mathbb{C}^n})^{k_2} x_2 &= 0 \quad \text{und} \\ (A - \lambda_1 \operatorname{Id}_{\mathbb{C}^n})^{\gamma} x_2 &\neq 0 \quad \text{für alle } \gamma = 1, 2, \ldots, k_2 - 1, \text{ falls } k_2 > 1, \end{aligned}$$

und andererseits $(A - \lambda_1 \operatorname{Id}_{\mathbb{C}^n})^{k_2-1} x_2 \notin \operatorname{span}\left\{x_1^{(1)}\right\}$ gilt.
Für dieses x_2 wiederhole man die Konstruktion in (7.2) mit x_2 anstelle von x_1, um einen A-invarianten Unterraum

$$U_2 := \operatorname{span}\{x_2^{(1)}, x_2^{(2)}, \ldots, x_2^{(k_2)}\}$$

mit der Matrixdarstellung

$$A_2 := A|_{U_2} = \left.\begin{pmatrix} \lambda_1 & 1 & 0 & \cdots & 0 \\ 0 & \lambda_1 & 1 & \ddots & \\ & & \ddots & \ddots & 0 \\ & & 0 & \lambda_1 & 1 \\ & & & & \lambda_1 \end{pmatrix}\right\} k_2 \text{ Zeilen}$$

in der Basis $\{x_2^{(1)}, x_2^{(2)}, \ldots, x_2^{(k_2)}\}$ zu erhalten. In gleicher Weise wird die gesamte Basis von $W_A(\lambda_1)$ ausgeschöpft, um schließlich alle $g_A(\lambda_1)$ Jordanblöcke zum Eigenwert λ_1 zu erzeugen.
(Man beachte, dass für x_3 dann $(A - \lambda_1 \operatorname{Id}_{\mathbb{C}^n})^{k_3-1} x_3 \notin \operatorname{span}\left\{x_1^{(1)}, x_2^{(1)}\right\}$ gelten muss, für x_4 entsprechend $(A - \lambda_1 \operatorname{Id}_{\mathbb{C}^n})^{k_4-1} x_4 \notin \operatorname{span}\left\{x_1^{(1)}, x_2^{(1)}, x_3^{(1)}\right\}$ usw.)

5. Schritt. Mit den weiteren Eigenwerten $\lambda_2, \ldots, \lambda_m$ verfährt man genauso und fügt schließlich die so erhaltenen neuen Teilbasen zu einer Basis des $\mathbb{C}^n$ zusammen. Die so erhaltenen Basisvektoren bilden als Spaltenvektoren die gesuchte *Transformationsmatrix B*.

Kapitel 8

Lineare Differentialgleichungssysteme

Viele Naturgesetze werden mathematisch mit Hilfe von Differentialgleichungen formuliert, sie spielen in den Natur-, Ingenieur- und Wirtschaftswissenschaften eine zentrale Rolle. Die bisher erlernten Methoden der Linearen Algebra erlauben uns die Behandlung von linearen Differentialgleichungssystemen, also Relationen, die etwa die zeitliche Entwicklung mehrerer Beobachtungsgrößen beschreiben, z.B. Schwingungszustände von Atomgittern, die zeitliche Entwicklung der Haushalte vieler Unternehmen oder die Population einer Spezies von Raubtieren abhängig von der Menge an Beutetieren.

Definition 8.1
Sei $I \subset \mathbb{R}$ ein Intervall. Die Gleichung

$$y'(t) = A(t)y(t) + b(t), \tag{8.1}$$

mit einer differenzierbaren Vektorfunktion $y : I \to \mathbb{C}^n$, einer Matrixfunktion $A : I \to \mathbb{C}^{n\times n}$ und einer Vektorfunktion $b : I \to \mathbb{C}^n$ heißt *lineares Differentialgleichungssystem*. Falls $b(t) = 0$ für alle $t \in I$, dann heißt das System *homogen*, andernfalls *inhomogen*. Falls in einem homogenen System die Matrixfunktion konstant ist, d.h. $A(t) = A \in \mathbb{C}^{n\times n}$, dann spricht man von einem homogenen linearen Differentialgleichungssystem *mit konstanten Koeffizienten,* also

$$y'(t) = Ay(t), \qquad A \in \mathbb{C}^{n\times n}. \tag{8.2}$$

Lemma 8.2 [Lösungsraum von linearen Differentialgleichungssystemen]
Die Lösungen von (8.1) *für* $A \in C^0(I, \mathbb{C}^{n\times n})$ *und* $b(t) \equiv 0$ *bilden einen n-dimensionalen Unterraum der auf I differenzierbaren Funktionen.*

Satz 8.3 [komplexes Differentialgleichungssystem]
(i) *Die Vektorfunktion*

$$y(t) = e^{\lambda t}c, \qquad t \in I, \quad c \in \mathbb{C}^n \setminus \{0\}, \quad \lambda \in \mathbb{C},$$

löst das System (8.2) *genau dann, wenn λ ein Eigenwert und c ein Eigenvektor von A zum Eigenwert λ ist.*

(ii) *Die Lösungen*

$$y_i(t) = e^{\lambda_i t}c_i, \qquad i = 1, \ldots, m,$$

sind linear unabhängig genau dann, wenn die Eigenvektoren c_i linear unabhängig sind. Insbesondere sind die Lösungen linear unabhängig, wenn die Eigenwerte paarweise verschieden sind, d.h. wenn $\lambda_i \neq \lambda_k$ für $i \neq k$.

Bemerkung. Für das *Anfangswertproblem*

$$\begin{cases} y' & = Ay \\ y(0) & = \xi_0 \quad \in \mathbb{C}^n \end{cases} \tag{8.3}$$

mit n linear unabhängigen Eigenvektoren $c_1, \ldots, c_n$ von A muss man für den Anfangsvektor $\xi_0 \in \mathbb{C}^n$ die (eindeutige) Basisdarstellung

$$\xi_0 = \sum_{i=1}^{n} \alpha_i c_i$$

wählen und erhält als (eindeutig bestimmte) Lösung von (8.3) die Vektorfunktion

$$y(t) := \sum_{i=1}^{n} \alpha_i y_i(t) = \sum_{i=1}^{n} \alpha_i e^{\lambda_i t} c_i.$$

Satz 8.5 [Reelles Differentialgleichungssystem]
Eine Lösung $y(t) = e^{\lambda t}c$ für das Differentialgleichungssystem (8.2) *mit $\lambda = u+iv \in \mathbb{C}$, $c = a+ib \in \mathbb{C}^n$, $A \in \mathbb{R}^{n\times n}$ erzeugt zwei reelle Lösungen*

$$\begin{aligned} z(t) &= \operatorname{Re}(y(t)) = e^{ut}\{a\cos vt - b\sin vt\} \\ z^*(t) &= \operatorname{Im}(y(t)) = e^{ut}\{a\sin vt + b\cos vt\}. \end{aligned}$$

Für $2p$ verschiedene komplexe Eigenwerte

$$\lambda_1, \overline{\lambda_1}, \lambda_2, \overline{\lambda_2}, \ldots, \lambda_p, \overline{\lambda_p} \in \mathbb{C} \setminus \mathbb{R}$$

und q verschiedene reelle Eigenwerte

$$\lambda_{2p+1}, \lambda_{2p+2}, \ldots, \lambda_{2p+q} \in \mathbb{R}$$

gibt es die $2p + q$ linear unabhängigen reellen Lösungen

$$\begin{aligned} z_i &= \operatorname{Re}(c_i e^{\lambda_i t}) = \operatorname{Re}(\overline{c_i} e^{\overline{\lambda_i} t}) \\ z_i^* &= \operatorname{Im}(c_i e^{\lambda_i t}) = -\operatorname{Im}(\overline{c_i} e^{\overline{\lambda_i} t}) \\ y_j &= c_j e^{\lambda_j t} \end{aligned}$$

für $i = 1, \ldots, p$ und $j = 2p + 1, \ldots, 2p + q$.

Lemma 8.7 [Transformation für Differentialgleichungssysteme]
Sei $B \in \mathbb{C}^{n\times n}$ eine invertierbare Matrix. Dann gilt:

$$y' = Ay \qquad \Longleftrightarrow \qquad z' = B^{-1}ABz \quad \text{für } z := B^{-1}y.$$

Lemma 8.8 [Diagonalisierbares Differentialgleichungssystem]
Falls A diagonalisierbar ist, dann bildet die Menge

$$\{b_1 e^{\lambda_1 t}, \dots, b_n e^{\lambda_n t}\}$$

eine Basis (oder Hauptsystem*) des Lösungsraums zur Differentialgleichung*

$$y' = Ay,$$

wobei die Matrix $B = (b_1, b_2, \dots, b_n)$ die Identität

$$B^{-1}AB = \mathrm{Diag}(\lambda_1, \lambda_2, \dots, \lambda_n) = \begin{pmatrix} \lambda_1 & 0 & & \\ 0 & \lambda_2 & 0 & \\ & & \ddots & \\ & & 0 & \lambda_n \end{pmatrix}$$

erfüllt.

Lemma 8.10 [Differentialgleichungssystem mit einem Jordanblock]
Falls für ein $\lambda \in \mathbb{C}$

$$A = \begin{pmatrix} \lambda & 1 & 0 & & \\ 0 & \lambda & 1 & 0 & \\ 0 & 0 & \ddots & \ddots & \\ \vdots & \vdots & & \lambda & 1 \\ 0 & \cdots & & 0 & \lambda \end{pmatrix} \in \mathbb{C}^{n\times n},$$

dann liefern die Spalten der Matrix

$$Y(t) := \begin{pmatrix} e^{\lambda t} & te^{\lambda t} & \frac{1}{2!}t^2 e^{\lambda t} & \cdots & \frac{1}{(n-1)!}t^{n-1}e^{\lambda t} \\ 0 & e^{\lambda t} & te^{\lambda t} & & \frac{1}{(n-2)!}t^{n-2}e^{\lambda t} \\ 0 & 0 & e^{\lambda t} & & \frac{1}{(n-3)!}t^{n-3}e^{\lambda t} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & e^{\lambda t} \end{pmatrix}$$

eine Basis (Hauptsystem) des Lösungsraums zu (8.2) *mit $Y(0) = \mathrm{Id}_{\mathbb{C}^n}$.*

Korollar 8.11 [System mit Jordannormalform]
Sei $A \in \mathbb{C}^{n\times n}$ in Jordannormalform gegeben, dann ergibt sich eine Basis des Lösungsraums von (8.2) *durch Kombination der Hauptsysteme für jeden Jordanblock gemäß Lemma* 8.10.

Kapitel 9

Lineare Optimierung

Die in der Analysis behandelten Optimierungsaufgaben werden in der Praxis durch eine häufig sehr große Anzahl von Nebenbedingungen in Form von Ungleichungen erschwert – etwa in vielen Bereichen der Ökonomie, wo Ungleichungsnebenbedingungen bestehende Kapazitäten bei Produktionsprozessen modellieren. Im Allgemeinen sind solche Fragestellungen hoch komplex, aber für lineare Optimierungsprobleme, bei denen die zu optimierende Funktion als auch die Nebenbedingungen in linearer Form auftreten, gibt es Algorithmen wie zum Beispiel das Simplexverfahren, das hier in seiner Standardform behandelt werden soll. Solche Verfahren erlauben den Einsatz von leistungsstarken Computern, um komplexe Anwendungsprobleme wie das Entwickeln von effizienten Fahrplänen öffentlicher Verkehrsmittel mit Tausenden von Variablen unter Millionen von Nebenbedingungen zu lösen.

Definition 9.6
Die folgende Maximierungsaufgabe nennt man ein *lineares Optimierungsproblem in Standardform*. Für eine Matrix $A \in \mathbb{R}^{m\times n}$, Vektoren $x, c \in \mathbb{R}^n$, eine Konstante $c_0 \in \mathbb{R}$ und einen Vektor $b = (b_1, \ldots, b_m) \in \mathbb{R}^m$ mit nichtnegativen Komponenten $b_i \geq 0$, $i = 1, \ldots, m$, bestimme man das Maximum der *Zielfunktion*

$$z(x) := c_0 + c \cdot x = c_0 + c^T x$$

unter den *Ungleichungsnebenbedingungen*

$$\begin{cases} x_j \geq 0 & \text{für alle } \; j = 1, \ldots, n, \\ (Ax - b)_i \leq 0 & \text{für alle } \; i = 1, \ldots, m. \end{cases} \tag{9.1}$$

Die Menge

$$\mathscr{C} := \{x \in \mathbb{R}^n : x \text{ erfüllt die Bedingungen (9.1)}\}$$

ist der Schnitt endlich vieler Halbräume, den man auch *Zulässigkeitsbereich* nennt. Jede Lösung $x^* \in \mathscr{C}$ mit

$$z(x^*) = \max_{x\in\mathscr{C}} z(x)$$

heißt *Maximierer* oder *maximale Lösung* des Optimierungsproblems, der zugehörige Wert $z(x^*)$ heißt *Maximum von z auf $\mathscr{C}$*. .

Bemerkung. (i) Es gibt allgemeinere lineare Optimierungsprobleme, die auf Standardform reduziert werden können.

(ii) Der Zulässigkeitsbereich $\mathscr{C}$ ist als Schnitt endlich vieler Halbräume ein *konvexes Polyeder* im $\mathbb{R}^n$. Dabei heißt eine Menge $\Omega \subset \mathbb{R}^n$ konvex, falls mit zwei Punkten $p, q \in \Omega$ auch deren Verbindungsstrecke in Ω liegt.

(iii) Ein Punkt $x_s \in \mathscr{C}$, der die Nebenbedingungen (9.1) mit mindestens einer Gleichung statt Ungleichung erfüllt, liegt auf einer *Sphäre* oder *Seitenfläche* des Polyeders.

(iv) Ein Punkt $x_e \in \mathscr{C}$, der die Nebenbedingungen (9.1) mit mindestens n linear unabhängigen Gleichungen statt Ungleichungen erfüllt, heißt *Ecke* des Polyeders $\mathscr{C}$.

Satz 9.7 [Ecken sind optimal]
Wenn ein lineares Optimierungsproblem in Standardform mit Zulässigkeitsbereich $\mathscr{C}$ eine maximale Lösung besitzt, dann gibt es mindestens eine Ecke $x_e \in \mathscr{C}$, die ein Maximierer ist.

Lemma 9.8 [Nichtmaximale Ecken]
Falls für eine Ecke $x_e \in \mathscr{C}$ der Wert $z(x_e)$ nicht das globale Maximum von z auf dem Zulässigkeitsbereich $\mathscr{C}$ liefert, dann gibt es eine benachbarte Ecke $x'_e \in \mathscr{C}$, so dass $z(x'_e) \geq z(x_e)$. (Hierbei heißt x'_e eine zu x_e benachbarte Ecke*, wenn die Verbindungsstrecke zwischen x_e und x'_e eine Kante des Polyeders ist. Eine* Kante *K des Polyeders zeichnet sich dadurch aus, dass jede in $\mathscr{C}$ enthaltene Strecke, die K schneidet, in K enthalten sein muss.)*

Definition 9.9
Für ein lineares Optimierungsproblem in Standardform mit den Nebenbedingungen (9.1) definiert man die *Schlupfvariablen* $x_{n+i} \geq 0$, $i = 1, \ldots, m$, durch die Gleichungen

$$(Ax - b)_i + x_{n+i} = 0 \qquad \text{für alle} \quad i = 1, \ldots, m.$$

Bemerkung. Falls $x_{n+i} = 0$ für ein $i \in \{1, \ldots, m\}$, dann gilt Gleichheit in der i-ten Ungleichung von (9.1). Damit liegt der zugehörige Punkt $x \in \mathbb{R}^n$ auf derjenigen Sphäre oder Seitenfläche des konvexen Polyeders $\mathscr{C}$, die zu der i-ten Ungleichung gehört.
Allgemein beschreibt eine Lösung

$$x \in \mathbb{R}^{n+m}_{\geq 0} := \{x = (x_1, \ldots, x_{n+m}) \in \mathbb{R}^{n+m} : x_i \geq 0 \ \text{ für alle } \ i = 1, \ldots, n+m\}$$

des linearen Gleichungssystems

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ x_{n+1} \\ \vdots \\ x_{n+m} \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m,$$

oder kurz

$$[\, A | \mathrm{Id}_{\mathbb{R}^m} \,]\, x = b,$$

eine Ecke, wenn $x_j = 0$ für mindestens n verschiedene Indizes $j \in \{1, \ldots, n+m\}$. Die eigentliche geometrische Ecke des Zulässigkeitsbereichs $\mathscr{C}$ ist dann gegeben durch die ersten n Komponenten $(x_1, x_2, \ldots, x_n)$.

Definition 9.12
Für ein lineares Optimierungsproblem in Standardform mit den Nebenbedingungen (9.1) ist das anfängliche *Simplextableau* gegeben durch:

$$\begin{array}{c|c||c} A & \mathrm{Id}_{R^m} & b \\ \hline c^T & 0 & -c_0 \end{array} \quad \sim \quad \left[\begin{array}{c|c} A & \mathrm{Id}_{R^m} \\ \hline c^T & 0 \end{array}\right] x = \left[\begin{array}{c} b \\ \hline z - c_0 \end{array}\right], \qquad x \in \mathbb{R}^{n+m}_{\geq 0}.$$

Die *zu diesem Simplextableau gehörige Ecke* ist

$$x = (0, \ldots, 0, b_1, b_2, \ldots, b_m) \in \mathbb{R}^{n+m}_{\geq 0},$$

und der Wert der Zielfunktion in dieser Ecke ist durch $z(0, \ldots, 0) = c_0$ gegeben.

Allgemein kann man durch Zeilenoperationen nach dem Gauß-Verfahren aus der Matrix $[A|\mathrm{Id}_{\mathbb{R}^m}]$ eine Matrix $\tilde{A} \in \mathbb{R}^{m\times(n+m)}$ erzeugen, die die verschiedenen Standardeinheitsvektoren $e_i \in \mathbb{R}^m$, $i = 1, \ldots, m$ als j_i-te Spalten besitzt ($j_i \neq j_k$ für $k \neq i$ und $j_i \in \{1, \ldots, n+m\}$.) Das zugehörige neue Simplextableau lautet dann

$$\begin{array}{c||c} \tilde{A} & \tilde{b} \\ \hline \tilde{c}^T & -\tilde{c}_0 \end{array},$$

wobei $\tilde{b} \in \mathbb{R}^m$, $\tilde{c} \in \mathbb{R}^{n+m}$ und $\tilde{c}_0 \in \mathbb{R}$ aus $b \in \mathbb{R}^m, (c, 0) \in \mathbb{R}^{n+m}$ und $c_0 \in \mathbb{R}$ durch die Zeilenoperationen hervorgegangen sind. Es gilt $c_{j_1} = c_{j_2} = \cdots = c_{j_m} = 0$, so dass für die zu diesem neuen Simplextableau gehörige Ecke $\tilde{x} \in \mathbb{R}^{n+m}_{\geq 0}$ gilt:

$$\tilde{x}_{j_i} = \tilde{b}_i \quad \text{für } i = 1, \ldots, m, \quad \text{und } \tilde{x}_k = 0 \quad \text{für alle } k \in \{1, \ldots, n+m\} \setminus \{j_1, \ldots, j_m\}.$$

Der Wert der Zielfunktion in dieser neuen Ecke ist $z(\tilde{x}_1, \ldots, \tilde{x}_n) = \tilde{c}_0$.

Definition 9.14 [Engpassbedingung]
Nach Wahl der j-ten Spalte (als *Pivotspalte*), $j \in \{1, \dots, n+m\}$ zur Erzeugung eines Einheitsvektors in der Matrix muss man nach der sogenannten *Engpassbedingung* die i-te Zeile (als *Pivotzeile*) ($i \in \{1, \dots, m\}$) so wählen, dass

$$\frac{b_i}{a_{ij}} := \min_{k \in \{1,\dots,m\}} \left\{ \frac{b_k}{a_{kj}} : a_{kj} > 0 \right\}.$$

Das Element a_{ij} nennt man auch das *Pivotelement*.

Bemerkung. Es kann zu einem Spaltenindex $j \in \{1, \dots, n\}$ mehrere Zeilenindices $i \in \{1, \dots, m\}$ geben, so dass die Engpassbedingung erfüllt ist.

Lemma 9.16 [Optimalitätskriterium]
Für das nach Zeilenoperationen aus dem anfänglichen Simplextableau hervorgegangene Tableau

$$\begin{array}{ccc||c} * & \cdots & * & \tilde{b} \\ \hline \tilde{c}_1 & \cdots & \tilde{c}_{n+m} & -\tilde{c}_0 \end{array}$$

mit zugehöriger Ecke

$$\xi = (\xi_1, \dots, \xi_{n+m}) \in \mathbb{R}^{n+m}_{\geq 0}$$

mit Zielfunktionswert $z(\xi_1, \dots, \xi_n) = \tilde{c}_0$ gilt: Falls $\tilde{c}_i > 0$ für (mindestens) ein $i \in \{1, \dots, n+m\}$, dann gibt es (mindestens) eine andere Ecke

$$\eta = (\eta_1, \dots, \eta_{n+m}) \in \mathbb{R}^{n+m} \geq 0$$

mit Zielfunktionswert $z(\eta_1, \dots, \eta_n) > \tilde{c}_0$. Um eine dieser Ecken zu finden, wählt man die i-te Spalte als Pivotspalte für einen weiteren Simplexschritt.

Andererseits ist die ursprüngliche Ecke $\xi \in \mathbb{R}^{n+m}$ mit $z(\xi_1, \dots, \xi_n) = \tilde{c}_0$ eine (nicht notwendig eindeutige) optimale Lösung, wenn $\tilde{c}_i \leq 0$ für alle $i = 1, \dots, n+m$.

Definition 9.17
Eine Ecke $x_e \in \mathscr{C}$ eines Zulässigkeitsbereichs $\mathscr{C}$ heißt *entartet*, wenn mehr als n Hyperebenen, die $\mathscr{C}$ beranden, sich in x_e schneiden.

Lemma 9.19 [Blandes Regel]
Der Simplexalgorithmus (siehe Definition 9.20) endet garantiert, wenn man immer die erste mögliche Spalte als Pivotspalte und dann die erste mögliche Zeile als Pivotzeile (unter Beachtung der Engpassbedingung Definition 9.14) wählt.

Bemerkung. In der Praxis ist es häufig effizienter, wenn man die Spalte als Pivotspalte wählt, für die der Koeffizient der Zielfunktion maximal ist. Aber für alle bekannten Pivotstrategien kann man "schlechte" Polyeder konstruieren, für die der Simplexalgorithmus jede Ecke durchläuft.

Definition 9.20 [Simplexalgorithmus]
Gegeben sei ein lineares Optimierungsproblem in Standardform, also eine Matrix $A_0 \in \mathbb{R}^{m\times n}$, Vektoren $b_0 \in \mathbb{R}^m_{\geq 0}$, $c \in \mathbb{R}^n$ und eine Konstante $c_0 \in \mathbb{R}$. Gesucht ist eine optimale Lösung $x^* \in \mathbb{R}^n$, so dass die Zielfunktion $z(x) := c^T x + c_0$ bei x^* den maximalen Wert annimmt unter den Nebenbedingungen

$$\begin{cases} x_j \geq 0 & \text{für alle } \; j = 1, \dots, n, \\ (A_0 x - b_0)_i \leq 0 & \text{für alle } \; i = 1, \dots, m. \end{cases} \tag{9.1}$$

Der *Simplexalgorithmus* beschreibt das folgende Vorgehen:

1. Erstelle nach Einführung der Schlupfvariablen $x_{n+1}, \dots, x_{n+m}$ das erste Simplextableau

$$\begin{array}{c|c||c} A_0 & \mathrm{Id}_{R^m} & b_0 \\ \hline c^T & 0 & -c_0 \end{array} \quad \sim \quad \left[\begin{array}{c|c} A_0 & \mathrm{Id}_{R^m} \\ \hline c^T & 0 \end{array}\right] x = \left[\begin{array}{c} b_0 \\ \hline z - c_0 \end{array}\right], \qquad x \in \mathbb{R}^{n+m}_{\geq 0},$$

   mit zugehöriger Ecke $p_0 = (0, \dots, 0, b_1, \dots, b_m)^T \in \mathbb{R}^{n+m}_{\geq 0}$. Setze $\tilde{c}^T := (c^T, 0) \in \mathbb{R}^{n+m}$, $\tilde{c}_0 := c_0 \in \mathbb{R}$ und $\tilde{p} := p_0$, und gehe zu 2.

2. Falls $\tilde{c}_i \leq 0$ für alle $i = 1, \dots, n+m$, dann STOP mit einer optimalen Lösung $x^* := (\tilde{p}_1, \dots, \tilde{p}_n)$ und dem maximalen Zielfunktionswert $z(x^*) = \tilde{c}_0$.

   Falls andererseits $\tilde{c}_i > 0$ für mindestens ein $i \in \{1, \dots, n+m\}$, dann wähle (nach Blandes Regel) den kleinsten Index $s \in \{1, \dots, n+m\}$ mit $\tilde{c}_s > 0$ als Pivotspaltenindex und gehe zu 3.

3. Falls $a_{rs} \leq 0$ für alle $r = 1, \dots, m$, dann ist die Zielfunktion z unbeschränkt auf dem durch (9.1) definierten Zulässigkeitsbereich, STOP.

   Sonst wähle $r \in \{1, \dots, m\}$ mit $a_{rs} > 0$, so dass

$$\frac{b_r}{a_{rs}} = \min\left\{\frac{b_i}{a_{is}} : a_{is} > 0, 1 \leq i \leq m\right\}$$

   als Pivotzeilenindex r (Engpassbedingung), und gehe zu 4.

4. Führe einen Simplexschritt durch, d.h. multipliziere die r-te Zeile mit $1/a_{rs}$ und erzeuge mit der so entstandenen 1 den r-ten Einheitsvektor in der s-ten Spalte durch Gaußsche Zeilenoperationen, um auf ein neues Simplextableau

$$\begin{array}{c||c} A & b \\ \hline d^T & -d_0 \end{array}$$

   mit zugehöriger Ecke $p \in \mathbb{R}^{n+m}_{\geq 0}$ zu kommen. Setze $\tilde{c} := d \in \mathbb{R}^{n+m}$, $\tilde{c}_0 := d_0 \in \mathbb{R}$ und $\tilde{p} := p$, und gehe zu 2.