

LINEARE ALGEBRA

HERBSTSEMESTER 2007

STUDIENGANG INFORMATIK

MARTIN H. GUTKNECHT

ETH ZÜRICH

Inhaltsverzeichnis

0	Vorkenntnisse —	
	Was man hätte lernen sollen	0-1
0.1	Lineare Gleichungssysteme	0-1
0.2	Vektorrechnung in der Ebene	0-4
0.3	Komplexe Zahlen	0-5
0.4	Polynome	0-7
0.5	Das griechische Alphabet	0-8
0.6	Notation	0-8
1	Lineare Gleichungssysteme —	
	Gauss-Elimination	1-1
1.1	Gauss-Elimination: der reguläre Fall	1-1
1.2	Gauss-Elimination: der allgemeine Fall	1-9
1.3	Die Lösungsmenge eines linearen Gleichungssystems	1-17
2	Matrizen und Vektoren im \mathbb{R}^n und \mathbb{C}^n	2-1
2.1	Matrizen, Zeilen- und Spaltenvektoren	2-1
2.2	Das Rechnen mit Matrizen und Vektoren	2-4
2.3	Die Transponierte einer Matrix; symmetrische und Hermitesche Matrizen	2-12
2.4	Das Skalarprodukt und die Norm von Vektoren; Längen und Winkel	2-16
2.5	Das äussere Produkt und die orthogonale Projektion auf eine Gerade	2-23
2.6	Matrizen als lineare Abbildungen	2-25
2.7	Die Inverse einer Matrix	2-26
2.8	Orthogonale und unitäre Matrizen	2-29
2.9	Strukturierte Matrizen	2-31

3	Die LR–Zerlegung	3-1
3.1	Die Gauss-Elimination als LR–Zerlegung	3-1
3.2	Die Gauss-Elimination als LR–Zerlegung: der allgemeine Fall	3-9
3.3	Block–LR–Zerlegung und LR–Updating	3-12
3.4	Die Cholesky-Zerlegung	3-16
4	Vektorräume	4-1
4.1	Definition und Beispiele	4-1
4.2	Unterräume, Erzeugendensysteme	4-6
4.3	Lineare Abhängigkeit, Basen, Dimension	4-8
4.4	Basiswechsel, Koordinatentransformation	4-18
5	Lineare Abbildungen	5-1
5.1	Definition, Beispiele, Matrixdarstellung	5-1
5.2	Kern, Bild und Rang	5-7
5.3	Matrizen als lineare Abbildungen	5-11
5.4	Affine Räume und die allgemeine Lösung eines inhomogenen Gleichungssystems	5-17
5.5	Die Abbildungsmatrix bei Koordinaten- transformation	5-18
6	Vektorräume mit Skalarprodukt	6-1
6.1	Normierte Vektorräume	6-1
6.2	Vektorräume mit Skalarprodukt	6-2
6.3	Orthonormalbasen	6-5
6.4	Orthogonale Komplemente	6-13
6.5	Orthogonale und unitäre Basiswechsel und Koordinatentransformationen	6-15
6.6	Orthogonale und unitäre Abbildungen	6-18
6.7	Normen von linearen Abbildungen (Operatoren) und Matrizen	6-19
7	Die Methode der kleinsten Quadrate und die QR–Zerlegung einer Matrix	7-1
7.1	Orthogonalprojektionen	7-1
7.2	Die Methode der kleinsten Quadrate	7-5
7.3	Die QR–Zerlegung einer Matrix	7-8
7.4	Die QR–Zerlegung mit Pivotieren	7-12

8	Determinanten	8-1
8.1	Permutationen	8-1
8.2	Determinante: Definition, Eigenschaften	8-3
8.3	Entwicklung nach Zeilen und Kolonnen	8-9
8.4	Determinanten von Blockdreiecksmatrizen	8-12
9	Eigenwerte und Eigenvektoren	9-1
9.1	Eigenwerte und Eigenvektoren von Matrizen und linearen Abbildungen	9-1
9.2	Ähnlichkeitstransformationen; die Eigenwertzerlegung	9-9
9.3	Eigenwerte und Eigenvektoren symmetri- scher und Hermitescher Matrizen	9-16
9.4	Die Jordansche Normalform	9-19
10	Anwendungen der Eigenwertzerlegung	10-1
10.1	Homogene lineare Differentialgleichungen mit konstanten Koeffizienten	10-1
10.2	Funktionen von Matrizen	10-7
10.3	Reelle lineare Differentialgleichungen mit komplexen Eigenwerten	10-8
10.4	Quadratische Formen und ihre Hauptachsentransformation	10-10
10.5	Quadratische Funktionen	10-15
10.6	Die Spektralnorm	10-16
11	Die Singulärwertzerlegung	11-1
11.1	Die Singulärwertzerlegung: Herleitung	11-1
11.2	Die Singulärwertzerlegung: Folgerungen	11-5

Kapitel 0

Vorkenntnisse — Was man hätte lernen sollen

Wir wollen hier kurz einige Themen aufgreifen, die die meisten Studierenden in der Mittelschule durchgenommen haben sollten.

Keine Angst: alles wird in dieser oder einer anderen Vorlesung explizit oder implizit (in verallgemeinerter Form) noch einmal behandelt. Lücken sind also keine Katastrophe, aber ein gewisser, zeitlich begrenzter Nachteil.

Bemerkung: Die Figuren sind zu ergänzen.

0.1 Lineare Gleichungssysteme

Eine einzige lineare Gleichung in einer Unbekannten x ,

$$ax = b,$$

hat genau eine Lösung, nämlich

$$x = \frac{b}{a},$$

außer wenn $a = 0$. In dieser Ausnahmesituation gibt es zwei Fälle:

- $b = 0$: jedes x ist Lösung \Rightarrow es gibt ∞ Lösungen,
- $b \neq 0$: kein x ist Lösung \Rightarrow es gibt keine Lösungen.

Betrachten wir als nächstes zwei Gleichungen in zwei Unbekannten.

BEISPIEL 0.1:

$$\begin{aligned} 2x_1 - 4x_2 &= 8 \\ 5x_1 + 3x_2 &= 7. \end{aligned} \tag{1}$$

Frage: Gibt es eine Lösung und wenn ja, gibt es nur eine? Diese Frage nach Existenz und Eindeutigkeit tritt in der Mathematik bekanntlich immer wieder auf. Wir können sie hier auf konstruktive Art beantworten: wir können alle Lösungen ausrechnen.

Wir lösen die eine Gleichung (z.B. die erste) nach der einen Unbekannten (z.B. der ersten) auf:

$$x_1 = 4 + 2x_2. \tag{2}$$

Dann setzen wir dies in die andere Gleichung ein:

$$5(4 + 2x_2) + 3x_2 = 7 \quad \text{oder} \quad 13x_2 = -13,$$

d.h. $x_2 = -1$. Einsetzen in (2) liefert dann sofort $x_1 = 2$, also:

$$x_1 = 2, \quad x_2 = -1.$$

Durch Einsetzen ins ursprüngliche System (1) verifiziert man, dass dieses Zahlenpaar effektiv eine Lösung ist. Wir werden sehen, dass diese Kontrolle gar nicht nötig ist. Es ist auch die einzige Lösung, denn wir haben sie ohne Wahlmöglichkeit abgeleitet. \blacklozenge

Gibt es zu zwei Gleichungen in zwei Unbekannten immer genau eine Lösung?

BEISPIEL 0.2:

$$\begin{aligned} 2x_1 - 4x_2 &= 8 \\ x_1 - 2x_2 &= 3. \end{aligned}$$

Dieses Gleichungssystem hat offensichtlich keine Lösung. Multipliziert man nämlich die zweite Gleichung mit 2, erhält man $2x_1 - 4x_2 = 6$. Dies widerspricht der ersten Gleichung. \blacklozenge

BEISPIEL 0.3:

$$\begin{aligned} 2x_1 - 4x_2 &= 8 \\ x_1 - 2x_2 &= 4. \end{aligned}$$

Nun ist die erste Gleichung einfach das Doppelte der zweiten, und offensichtlich besteht nun kein Widerspruch mehr. Man kann sogar die eine Variable, z.B. x_2 , frei wählen und erhält für jede Wahl eine andere Lösung. Das heisst, es gibt eine ganze (einparametrische) Schar von Lösungen. \blacklozenge

Die Beispiele 2 und 3 zeichnen sich offensichtlich dadurch aus, dass auf der linken Seite die eine Gleichung ein Mehrfaches der anderen ist. Bei Beispiel 3 besteht diese Abhängigkeit auch auf der rechten Seite. Beides sind offensichtlich Ausnahmefälle.

Also: In der Regel gibt es zu einer linearen Gleichung in einer Unbekannten und zu zwei linearen Gleichungen in zwei Unbekannten genau eine Lösung; aber in Ausnahmefällen kann es keine oder unendlich viele geben.

Wir werden sehen, dass das Analoge auch bei n Gleichungen in n Unbekannten richtig ist.

Im Falle von n linearen Gleichungen in n Unbekannten gilt:

- In der Regel gibt es genau eine Lösung.
- Es gibt eine Schar von Lösungen, wenn eine Gleichung als eine Summe von Mehrfachen (“eine Linearkombination”) der anderen dargestellt werden kann.
- Es gibt keine Lösung, wenn eine solche Abhängigkeit zwar auf der linken Seite besteht, nicht aber gleichzeitig für die rechte Seite gilt.

Was gilt bei m Gleichungen in n Unbekannten, wenn $m \neq n$?

BEISPIEL 0.4:

$$\begin{aligned}2x_1 - 4x_2 &= 8 \\5x_1 + 3x_2 &= 7 \\x_1 + x_2 &= 3.\end{aligned}$$

Dieses System mit $m = 3$, $n = 2$ hat keine Lösung. Denn aus Beispiel 1 wissen wir, dass die ersten zwei Gleichungen nur die Lösung $x_1 = 2, x_2 = -1$ zulassen, und die steht im Widerspruch zur dritten Gleichung. ♦

Ersetzt man die 3 auf der rechten Seite der letzten Gleichung durch eine 1, so gibt es offenbar genau eine Lösung. Ergänzt man Beispiel 0.3 durch die Gleichung $-3x_1 + 6x_2 = -12$, so findet man sogar ein (ziemlich triviales) Beispiel von drei Gleichungen in zwei Unbekannten mit unendlich vielen Lösungen.

Wenn $m > n$ (“mehr Gleichungen als Unbekannte”), so gibt es in der Regel keine Lösung. Es kann aber eine oder sogar unendlich viele geben, wenn gewisse Abhängigkeiten zwischen den Gleichungen bestehen.

BEISPIEL 0.5:

$$\begin{aligned}x_1 - x_2 + 2x_3 &= 1 \\2x_1 + x_2 + x_3 &= 5.\end{aligned}\tag{3}$$

Hier ist $m = 2$, $n = 3$. Wenn man von der zweiten Gleichung das Doppelte der ersten subtrahiert, folgt $3x_2 - 3x_3 = 3$ oder $x_2 = x_3 + 1$. Einsetzen in die erste oder die zweite Gleichung ergibt dann noch $x_1 = 2 - x_3$. Man kann also x_3 frei wählen. Nennen wir den freien Parameter t , so lässt sich jede Lösung als Zahlentripel der Form

$$x_1 = 2 - t, \quad x_2 = 1 + t, \quad x_3 = t$$

darstellen. Mit anderen Worten, die zwei Gleichungen (3) haben die Lösungsmenge

$$\{(2 - t, 1 + t, t); t \in \mathbb{R}\},$$

wenn wir uns auf reelle Lösungen beschränken. ♦

Man könnte offensichtlich auch ein Beispiel mit $m = 2$, $n = 3$ konstruieren, das keine Lösung hat. Wir werden sehen, dass gilt:

Wenn $m < n$ (“weniger Gleichungen als Unbekannte”), so gibt es in der Regel eine Schar von Lösungen. Es kann aber in Ausnahmefällen auch keine geben, aber nie nur endlich viele.

Wir werden im folgenden Kapitel 1 ganz allgemein ein lineares Gleichungssystem von m Gleichungen in n Unbekannten lösen und seine Lösungsmenge charakterisieren.

0.2 Vektorrechnung in der Ebene

Die Punkte in der Ebene kann man einerseits durch Ortsvektoren, andererseits durch kartesische Koordinaten charakterisieren:

$$\overrightarrow{OW} = \mathbf{w} = \begin{pmatrix} u \\ v \end{pmatrix}, \quad \overrightarrow{OZ} = \mathbf{z} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad \dots$$

Es gibt noch viele andere Bezeichnungsweisen, aber wir verwenden hier diese, weil sie mit der später verwendeten konsistent ist.

Solche Vektoren kann man addieren und mit einer Zahl (“Skalar”) multiplizieren:

$$\mathbf{w} + \mathbf{z} = \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} u + x \\ v + y \end{pmatrix},$$

$$\alpha \mathbf{w} = \alpha \begin{pmatrix} u \\ v \end{pmatrix} := \begin{pmatrix} \alpha u \\ \alpha v \end{pmatrix}.$$

Diese Operationen haben einfache geometrische Interpretationen.

Insbesondere kann der Vektor von W nach Z als Differenz dargestellt werden:

$$\overrightarrow{WZ} = \mathbf{z} - \mathbf{w}.$$

Ergibt sich $\overrightarrow{OW} = \mathbf{w}$ durch Drehung von $\overrightarrow{OZ} = \mathbf{z}$ um den Winkel φ , so kann man dies ausdrücken durch

$$u = x \cos \varphi - y \sin \varphi, \quad v = x \sin \varphi + y \cos \varphi. \quad (4)$$

Der Spezialfall $(x, y) = (1, 0)$ illustriert die Definition von Sinus und Cosinus.

Die Beziehungen (4) kann man zusammenfassen in

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad (5)$$

oder kürzer als

$$\mathbf{w} = \mathbf{Q} \mathbf{z} \quad (6)$$

mit der 2×2 Matrix

$$\mathbf{Q} = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}. \quad (7)$$

$\mathbf{Q} \mathbf{z}$ ist ein Matrix-Vektor-Produkt.

Man könnte hier noch vieles anfügen, was vielleicht schon bekannt ist: die Länge eines Vektors, das Skalarprodukt von zwei Vektoren, die Determinante und die Eigenwerte einer quadratischen Matrix.

Die Verallgemeinerung des Vektor- und des Matrixbegriffes und die Behandlung der obgenannten Operationen und Begriffe (nicht aber des sogenannten Vektorproduktes!) ist das Thema dieser Vorlesung: **lineare Algebra**.

0.3 Komplexe Zahlen

Bemerkung: Komplexe Zahlen werden ausführlich in der Analysis behandelt. Wir beschränken uns auf wenige Grundeigenschaften, die wir gelegentlich brauchen werden und kümmern uns nicht um vollständige Herleitungen.

Man könnte die Punkte in der Ebene auch in der Form

$$w = u + i v, \quad z = x + i y, \quad \dots$$

darstellen, wobei man zunächst i als eine die zweiten Komponenten charakterisierende Flagge auffassen könnte.

Dann ist klar, dass das Folgende vernünftige Definitionen für die **Addition** zweier solcher "Punkte" und für die **Multiplikation mit einer reellen Zahl** α ist:

$$w + z = (u + i v) + (x + i y) = (u + x) + i (v + y), \quad (8)$$

$$\alpha w = \alpha(u + i v) = (\alpha u + i \alpha v). \quad (9)$$

Von entscheidender Bedeutung ist jetzt aber, dass man für diese Zahlen eine die üblichen Regeln erfüllende **Multiplikation** definieren kann, wenn man annimmt, dass formal die binomischen Formeln und die Relation $i^2 = -1$ gelten:

$$w z = (u + i v)(x + i y) = u x + i (v x + u y) + i^2 (v y) \quad (10)$$

$$= (u x - v y) + i (v x + u y). \quad (11)$$

Mit dieser Interpretation von i nennt man $z = x + i y$ eine **komplexe Zahl** [*complex number*]; x ist der **Realteil** [*real part*], y der **Imaginärteil** [*imaginary part*]:

$$z = x + i y \iff x = \operatorname{Re} z, \quad y = \operatorname{Im} z.$$

Die Zahl $\bar{z} = x - i y$ heisst **konjugiert-komplex** [*conjugate complex*] zu z .

Die Menge der komplexen Zahlen bezeichnet man mit \mathbb{C} . In Anbetracht der geometrischen Interpretation als Punkte in der Ebene nennt man \mathbb{C} auch die **komplexe Ebene** [*complex plane*]. Die Menge der reellen Zahlen \mathbb{R} ist diejenige Teilmenge von \mathbb{C} , deren Elemente Imaginärteil 0 haben:

$$\mathbb{R} = \{z \in \mathbb{C}; \operatorname{Im} z = 0\}.$$

In der komplexen Ebene ist das die **reelle Achse** [*real axis*], während die Zahlen mit Realteil null die **imaginäre Achse** [*imaginary axis*] bilden.

Man beachte, dass

$$z \bar{z} = x^2 + y^2 \geq 0, \quad (12)$$

wobei das Gleichheitszeichen nur gilt, wenn $z = 0$ ist.

Somit kann man stets die normale Quadratwurzel von $z\bar{z}$ bilden: die nicht-negative reelle Zahl

$$|z| := \sqrt{z\bar{z}} \quad (13)$$

heisst **Betrag** [*modulus*] oder **Absolutbetrag** [*absolute value*] der komplexen Zahl z . Nach Pythagoras ist $|z|$ gerade die Länge des Ortsvektors zum Punkt $z = x + iy$ der komplexen Ebene.

Die Beziehung (12) zeigt, wie man die **Division** definieren muss:

$$\frac{z}{w} = \frac{z\bar{w}}{w\bar{w}} = \tau z\bar{w}, \quad \text{wo } \tau := \frac{1}{|w|^2}. \quad (14)$$

In Real- und Imaginärteil von Zähler und Nenner ausgedrückt ergibt sich

$$\frac{z}{w} = \frac{xu + yv}{u^2 + v^2} + i \frac{uy - xv}{u^2 + v^2}. \quad (15)$$

Mit diesen Definitionen gelten für die komplexen Zahlen die gleichen Rechenregeln wie für die reellen Zahlen.

Für den Absolutbetrag ist insbesondere folgendes richtig:

$$|z + w| \leq |z| + |w|, \quad (16)$$

$$|zw| = |z| |w|, \quad (17)$$

$$\left| \frac{z}{w} \right| = \frac{|z|}{|w|}. \quad (18)$$

Die komplexen Zahlen als Punkte der komplexen Ebene lassen sich auch in Polarkoordinaten darstellen:

$$z = x + iy \iff \begin{cases} x = r \cos \varphi \\ y = r \sin \varphi \end{cases} \text{ falls } \begin{cases} r = |z|, \\ \varphi = \arccos(x/r) \\ \quad = \arcsin(y/r). \end{cases}$$

Es gilt also

$$z = r(\cos \varphi + i \sin \varphi). \quad (19)$$

Mit der berühmten **Eulerschen Formel** [*Euler formula*]

$$e^{i\varphi} = \cos \varphi + i \sin \varphi \quad (20)$$

erhalten wir schliesslich

$$z = r e^{i\varphi}. \quad (21)$$

Diese Darstellungen heisst **Polarform** [*polar representation*] von z .

Für die Exponentialfunktion mit komplexem Argument gelten die gleichen Regeln wie bei reellem Argument. Zum Beispiel ist

$$e^{i(\varphi+\psi)} = e^{i\varphi} e^{i\psi}, \quad e^{i(\varphi-\psi)} = \frac{e^{i\varphi}}{e^{i\psi}}. \quad (22)$$

Wenn $z = r e^{i\varphi}$ und $w = s e^{i\psi}$, so hat man also

$$zw = rs e^{i(\varphi+\psi)}, \quad \frac{z}{w} = \frac{r}{s} e^{i(\varphi-\psi)}. \quad (23)$$

0.4 Polynome

Sind a_0, a_1, \dots, a_n reelle oder komplexe Zahlen, so kann man das **Polynom** [polynomial]

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0$$

vom **Höchstgrade** [maximum degree] n definieren. Ist $a_n \neq 0$, so sagen wir, n sei der **exakte Grad**. Die Zahlen a_0, a_1, \dots, a_n bezeichnet man als die **Koeffizienten** [coefficients] des Polynoms.

Für uns ist vor allem der Fall interessant, wo auch z reell oder komplex ist. Dann kann man $p(z)$ einen (reellen oder komplexen) Wert zuweisen:

$$p(z) \in \mathbb{R} \quad \text{oder} \quad p(z) \in \mathbb{C}.$$

Die (reelle oder komplexe) Zahl z_1 heisst **Nullstelle** [zero] von p , wenn $p(z_1) = 0$ ist. Ist dies der Fall, so ist

$$q(z) := \frac{p(z)}{z - z_1} \quad (24)$$

nach dem Kürzen des **Linearfaktors** [linear factor] $z - z_1$ ein Polynom vom Grade $n - 1$. Der Übergang von $p(z)$ zu $q(z)$ heisst **Deflation** [deflation] der Nullstelle z_1 .

Sind die Koeffizienten a_0, a_1, \dots, a_n reell und beschränkt man sich auf reelle z , so braucht ein Polynom bekanntlich keine Nullstelle zu haben (Bsp.: $z^2 + 1$). Anders ist es, wenn wir komplexe Nullstellen zulassen; dann sind sogar komplexe Koeffizienten erlaubt:

Satz 0.1 *Ein (reelles oder komplexes) Polynom p vom exakten Grade $n \geq 1$ hat in \mathbb{C} (mindestens) eine Nullstelle:*

$$\exists z_1 \in \mathbb{C} : \quad p(z_1) = 0.$$

Mit Induktion folgert man daraus mittels Deflation:

Korollar 0.2 *Ein (reelles oder komplexes) Polynom p vom exakten Grade $n \geq 1$ hat in \mathbb{C} genau n Nullstellen, die aber nicht alle verschieden sein müssen; genauer, p lässt sich als Produkt von n Linearfaktoren darstellen:*

$$p(z) = a_n (z - z_1)(z - z_2) \cdots (z - z_n) \equiv a_n \prod_{k=1}^n (z - z_k).$$

0.5 Das griechische Alphabet

Das griechische Alphabet hat 24 Buchstaben, aber zu einzelnen gibt es zwei Schreibweisen:

α	alpha	ι	iota	σ	sigma
β	beta	κ	kappa	ς	sigma
γ	gamma	λ	lambda	τ	tau
δ	delta	μ	mu	υ	ypsilon
ϵ	epsilon	ν	nu	ϕ	phi
ε	epsilon	ξ	xi	φ	phi
ζ	zeta	\omicron	omikron	χ	chi
η	eta	π	pi	ψ	psi
θ	theta	ρ	rho	ω	omega
ϑ	theta	ϱ	rho		
Γ	Gamma	Ξ	Xi	Φ	Phi
Δ	Delta	Π	Pi	Ψ	Psi
Θ	Theta	Σ	Sigma	Ω	Omega
Λ	Lambda	Υ	Ypsilon		

Die restlichen Grossbuchstaben sind gleich wie die entsprechenden des lateinischen Alphabets und sind deshalb für uns nutzlos.

0.6 Notation

Im Prinzip ist jedermann frei, die Mathematik in eigenen Bezeichnungen aufzuschreiben. Aber eine gute und systematische Notation erleichtert das Verständnis.

Wir werden für Vektoren und Matrizen kleine bzw. grosse fette Buchstaben verwenden, damit sie in jeder Formel leicht erkennbar sind:

$$\mathbf{x}, \mathbf{y}, \dots, \quad \mathbf{A}, \mathbf{B}, \dots$$

Der Nachteil ist, dass man das so nicht gut von Hand schreiben kann. Aber man kann in Handschrift auf den Unterschied zwischen gewöhnlichen und fetten Buchstaben verzichten, oder die Vektoren und Matrizen durch Unterstreichen markieren, d.h. zum Beispiel A schreiben statt **A**.

Das Gleichheitszeichen hat in der Mathematik und Informatik mindestens drei verschiedene Bedeutungen, und wir werden versuchen, diese auseinander zu halten (was allerdings nicht immer eindeutig möglich ist). Wir schreiben

- = für eine Gleichheit von mathematischen Objekten,
- := für eine Zuweisung, wie sie in Algorithmen und Programmen vorkommt,
- ≐ für eine Definition.

Kapitel 1

Lineare Gleichungssysteme — Gauss-Elimination

Die **Gauss-Elimination**¹ [*Gauss elimination*], oft einfach **Gauss-Algorithmus** [*Gauss algorithm*] genannt, ist die grundlegendste und weitverbreitetste Methode um lineare Gleichungssysteme zu lösen. Es ist im Prinzip jene Methode, die man schon in der Grundschule lernt. Wir betrachten sie hier in systematischerer Weise. Zunächst einige Fakten:

- Die Gauss-Elimination ist von grundlegender Bedeutung für das **wissenschaftliche Rechnen** [*scientific computing*] und allgemeiner, für die **rechnergestützten Wissenschaften** [*computational science and engineering*]. Fast bei jeder Computersimulation sind irgendwo lineare Gleichungssysteme zu lösen, und oft sind diese sehr gross.
- Im wissenschaftlichen Rechnen wird die Methode aber fast nur im Falle $m = n$ eingesetzt.
- Es kommen dabei allerdings z.T. komplizierte Varianten zum Einsatz, die in bezug auf Stabilität, Rechengeschwindigkeit (auch Speicherzugriff), Speicherplatz und Parallelisierbarkeit optimiert sind.
- Wir wollen den allgemeinen Fall von m linearen Gleichungen in n Unbekannten analysieren und die entsprechende Lösungsmenge beschreiben.
- Die daraus ableitbare Theorie ist fundamental für die lineare Algebra. Wir werden uns später immer wieder auf sie beziehen.

1.1 Gauss-Elimination: der reguläre Fall

Wir führen in diesem Abschnitt zunächst die Notation für allgemeine Systeme von m linearen Gleichungen in n Unbekannten ein und diskutieren dann die Gauss-Elimination im Normalfall $m = n$, beginnend mit Beispielen mit $m = n = 3$. Im nächsten Abschnitt werden wir dann den allgemeinen Fall behandeln.

Ein **Gleichungssystem** [*system of equations*] von m linearen Gleichungen in n Unbekannten hat die Form

¹CARL FRIEDRICH GAUSS (30.4.1777 – 23.2.1855), ab 1807 Professor für Astronomie in Göttingen; einer der bedeutendsten Mathematiker aller Zeiten.

Wir können uns damit nicht nur Schreibarbeit sparen, wichtiger ist, dass wir die in diesem Schema auszuführenden Operationen besser veranschaulichen und damit auch leichter auf den Computer übertragen — sprich programmieren — können als bei Verwendung des ausgeschriebenen Systems (1.1). Mit anderen Worten, das Schema wird einerseits zum Rechnen von Hand angewandt (was wir nur zur Illustration ausführen) und erlaubt uns andererseits besser zu verstehen, was in der Rechnung abläuft.

Das Schema (1.3) hat offenbar einerseits **Zeilen** [*rows*] und andererseits **Spalten** oder **Kolonnen** [*columns*].

Wir wollen nun zuerst ein Zahlenbeispiel mit $m = n = 3$ auf konventionelle Weise lösen und dann die Rechnung auf das Schema übertragen.

BEISPIEL 1.1:

$$\begin{aligned} 2x_1 - 2x_2 + 4x_3 &= 6 \\ -5x_1 + 6x_2 - 7x_3 &= -7 \\ 3x_1 + 2x_2 + x_3 &= 9. \end{aligned} \quad (1.4)$$

Wir könnten wie üblich x_1 eliminieren, indem wir die erste Gleichung nach x_1 auflösen und den so erhaltenen Ausdruck für x_1 in die anderen beiden Gleichungen einsetzen. Wir können das Gleiche erreichen, indem wir geeignete Vielfache der ersten Zeile von den beiden anderen subtrahieren.

Hier subtrahieren wir das $-5/2$ -fache der ersten Gleichung von der zweiten und das $3/2$ -fache der ersten Gleichung von der dritten.

$$\begin{aligned} 2x_1 - 2x_2 + 4x_3 &= 6 \\ x_2 + 3x_3 &= 8 \\ 5x_2 - 5x_3 &= 0. \end{aligned}$$

Um auch noch x_2 zu eliminieren, können wir das 5-fache der zweiten Gleichung von der dritten subtrahieren:

$$\begin{aligned} 2x_1 - 2x_2 + 4x_3 &= 6 \\ x_2 + 3x_3 &= 8 \\ -20x_3 &= -40. \end{aligned} \quad (1.5)$$

Dieses System hat **obere Dreiecksgestalt** [*upper triangular form*] und lässt sich auflösen, indem man aus der letzten Gleichung die letzte Variable bestimmt, dann aus der zweitletzten Gleichung die zweitletzte Variable, und so fort.

$$\begin{aligned} -20x_3 = -40 &\implies x_3 = (-40)/(-20) = 2 \\ x_2 + 3x_3 = 8 &\implies x_2 = 8 - 3 \cdot 2 = 2 \\ 2x_1 - 2x_2 + 4x_3 = 6 &\implies x_1 = (6 + 2 \cdot 2 - 4 \cdot 2)/2 = 1. \end{aligned}$$

Man nennt dieses rekursive Lösen **Rückwärtseinsetzen** [*back substitution*], weil man bei x_n mit Auflösen anfängt. ♦

Wir können die Reduktion des gegebenen Systems (1.4) auf das Dreieckssystem (1.5) kompakter darstellen, indem wir die relevanten Zahlen in ein Eliminationsschema eintragen. Wir zeigen links

das Ausgangsschema für ein allgemeines 3×3 -System, rechts jenes für unser konkretes Beispiel:

$$\begin{array}{l}
 \begin{array}{c} \\ l_{21} \\ l_{31} \end{array} \begin{array}{c|ccc|c}
 x_1 & x_2 & x_3 & 1 \\
 a_{11} & a_{12} & a_{13} & b_1 \\
 a_{21} & a_{22} & a_{23} & b_2 \\
 a_{31} & a_{32} & a_{33} & b_3
 \end{array} \\
 \end{array} \quad \begin{array}{c} \\ -\frac{5}{2} \\ \frac{3}{2} \end{array} \begin{array}{c|ccc|c}
 x_1 & x_2 & x_3 & 1 \\
 \textcircled{2} & -2 & 4 & 6 \\
 -5 & 6 & -7 & -7 \\
 3 & 2 & 1 & 9 \\
 1 & -1 & 2 & 3
 \end{array} \quad (1.6)$$

Wir haben links vom Eliminationsschema die Faktoren $l_{21} = -\frac{5}{2}$ und $l_{31} = \frac{3}{2}$ notiert, mit denen wir die erste Zeile (oder Gleichung) multiplizieren, bevor wir sie von der entsprechenden Zeile subtrahieren. Als Rechenhilfe haben wir schliesslich im Zahlenbeispiel unter dem Schema die durch zwei dividierte erste Zeile angegeben. Sie muss nur noch mit dem Zähler (und dem Vorzeichen) der Faktoren multipliziert werden, falls diese als ungekürzter Bruch geschrieben sind; so kann man die im Kopf auszuführenden Operationen vereinfachen. Auf dem Computer wird man diese Zeile nicht berechnen; für ihn spielt es (meist) keine Rolle, ob mit einer ganzen Zahl oder einem Bruch multipliziert wird.

Nach einem Eliminationsschritt hat man

$$\begin{array}{l} \\ l_{32} \end{array} \begin{array}{c|ccc|c}
 x_1 & x_2 & x_3 & 1 \\
 a_{11} & a_{12} & a_{13} & b_1 \\
 0 & a_{22}^{(1)} & a_{23}^{(1)} & b_2^{(1)} \\
 0 & a_{32}^{(1)} & a_{33}^{(1)} & b_3^{(1)}
 \end{array} \quad \begin{array}{c} \\ 5 \end{array} \begin{array}{c|ccc|c}
 x_1 & x_2 & x_3 & 1 \\
 2 & -2 & 4 & 6 \\
 0 & \textcircled{1} & 3 & 8 \\
 0 & 5 & -5 & 0 \\
 0 & 1 & 3 & 8
 \end{array} \quad (1.7)$$

und nach einem weiteren Eliminationsschritt kommt man zum folgenden Endschema:

$$\begin{array}{c|ccc|c}
 x_1 & x_2 & x_3 & 1 \\
 a_{11} & a_{12} & a_{13} & b_1 \\
 0 & a_{22}^{(1)} & a_{23}^{(1)} & b_2^{(1)} \\
 0 & 0 & a_{33}^{(2)} & b_3^{(2)}
 \end{array} \quad \begin{array}{c|ccc|c}
 x_1 & x_2 & x_3 & 1 \\
 2 & -2 & 4 & 6 \\
 0 & 1 & 3 & 8 \\
 0 & 0 & \textcircled{-20} & -40 \\
 & & 1 & 2
 \end{array} \quad (1.8)$$

Dabei haben wir das Element, das jeweils zum Auflösen nach einer Variablen dient, eingekreist. Es heisst **Pivotelement** [*pivot element*] oder kurz **Pivot**. Die Zeile bzw. Kolonne, in der das Pivotelement liegt, heisst **Pivotzeile** [*pivot row*] bzw. **Pivotkolonne** [*pivot column*]. Die (fakultative) Hilfszeile unter dem Eliminationsschema, auch **Kellerzeile** genannt, ist also die durch das Pivotelement dividierte Pivotzeile. Da eine Zeile einer Gleichung entspricht, kann man die Pivotzeile auch als **Pivotgleichung** [*pivot equation*] bezeichnen. Das nach jedem Schritt weiter verkleinerte System heisst **Restgleichungssystem** [*reduced system*]. Es ist in obigem Beispiel durch eine zusätzliche Box hervorgehoben.

Konkret ergeben sich im Beispiel als erstes und zweites Restgleichungssystem (mit eingekreisten Pivotelementen)

$$\begin{array}{ccc|c} x_2 & x_3 & & 1 \\ \hline \textcircled{1} & 3 & & 8 \\ 5 & -5 & & 0 \end{array} \quad \begin{array}{c|c} x_3 & 1 \\ \hline \textcircled{-20} & -40 \end{array}$$

Ausgehend von den Daten im Endschema (1.8) müsste man jetzt noch rückwärts einsetzen, um die Lösung zu berechnen. Dies macht man genau gleich wie oben in Beispiel 1.1.

Im obigen Beispiel ist für $k = 1, 2, 3$ im k -ten Eliminationsschritt das Element $a_{kk}^{(k-1)}$ als Pivotelement gewählt worden, und entsprechend ist die k -te Zeile Pivotzeile und die k -te Kolonne Pivotkolonne. Wir werden aber gleich sehen, dass das nicht immer so geht, ohne dass man Zeilen vertauscht, was problemlos ist, da Zeilen ja Gleichungen sind.

Zunächst sei aber noch darauf hingewiesen, dass beim ersten Schritt ja die erste Gleichung unverändert bleibt, beim zweiten Schritt sogar die zwei ersten Gleichungen. Im ersten Schritt wird eben das 3×3 -System auf ein 2×2 -System reduziert oder allgemeiner, ein $m \times n$ -System auf ein $(m-1) \times (n-1)$ -System.

Da nach dem k -ten Schritt die ersten k Gleichungen unverändert bleiben und die ersten k Koeffizienten der restlichen Zeilen null sind, kann man sich Schreibearbeit sparen, indem man jeweils nur das neue Restgleichungssystem aufschreibt. Für die Programmierung bedeutet dies, dass der Zugriff auf jenen Teil der Daten beschränkt bleibt, die zum Restgleichungssystem gehören.

Wir illustrieren diese neue, kompaktere Darstellung, die besser zeigt, was wirklich zu berechnen ist, an einem zweiten Beispiel.

BEISPIEL 1.2: Die Reduktion des Systems

$$\begin{aligned} x_2 + 4x_3 &= 1 \\ 2x_1 + 4x_2 - 4x_3 &= 1 \\ 4x_1 + 8x_2 - 3x_3 &= 7 \end{aligned} \tag{1.9}$$

auf obere Dreiecksgestalt wird nachfolgend links in der bisherigen Darstellung gezeigt, rechts in der kompakteren. Dabei tritt noch die Schwierigkeit auf, dass $a_{11} = 0$ ist, dieses Element also nicht als Pivotelement in Frage kommt. Wir wählen stattdessen a_{21} als erstes Pivotelement. In unserer linken Darstellung vertauschen wir in diesem Falle die erste und die zweite Zeile. Auf der rechten Seite entfällt dieser triviale Zwischenschritt. Es genügt, das Pivotelement zu markieren. Die Kellerzeile lassen wir diesmal weg.

$$\begin{array}{c}
 \begin{array}{c|ccc|c}
 x_1 & x_2 & x_3 & 1 \\
 \hline
 0 & 1 & 4 & 1 \\
 \textcircled{2} & 4 & -4 & 1 \\
 4 & 8 & -3 & 7
 \end{array} & & \begin{array}{c|ccc|c}
 x_1 & x_2 & x_3 & 1 \\
 \hline
 0 & 1 & 4 & 1 \\
 \textcircled{2} & 4 & -4 & 1 \\
 2 & 4 & 8 & 7
 \end{array} \\
 \\
 \begin{array}{c|ccc|c}
 x_1 & x_2 & x_3 & 1 \\
 \hline
 \textcircled{2} & 4 & -4 & 1 \\
 0 & 1 & 4 & 1 \\
 2 & 4 & 8 & 7
 \end{array} & & \begin{array}{c|cc|c}
 x_2 & x_3 & 1 \\
 \hline
 \textcircled{1} & 4 & 1 \\
 0 & 5 & 5
 \end{array} \\
 \\
 \begin{array}{c|ccc|c}
 x_1 & x_2 & x_3 & 1 \\
 \hline
 2 & 4 & -4 & 1 \\
 0 & \textcircled{1} & 4 & 1 \\
 0 & 0 & 5 & 5
 \end{array} & & \begin{array}{c|c|c}
 x_3 & 1 \\
 \hline
 \textcircled{5} & 5
 \end{array} \\
 \\
 \begin{array}{c|ccc|c}
 x_1 & x_2 & x_3 & 1 \\
 \hline
 \textcircled{2} & 4 & -4 & 1 \\
 0 & \textcircled{1} & 4 & 1 \\
 0 & 0 & \textcircled{5} & 5
 \end{array} & &
 \end{array}$$

Hier passiert im zweiten Eliminationsschritt (mit Pivot 1) nichts mehr, denn nach dem Zeilenvertauschen und dem ersten Schritt hat das Schema bereits obere Dreiecksgestalt. Mit anderen Worten, hier hat das erste Restgleichungssystem (der Größe 2×2) per Zufall Dreiecksgestalt.

Im letzten Schema links haben wir nun alle Pivotelemente markiert; das wird später im allgemeinen Fall wichtig sein.

Es bleibt noch das Rückwärtseinsetzen:

$$\begin{aligned}
 5x_3 &= 5 \implies x_3 = 5/5 = 1 \\
 x_2 + 4x_3 &= 1 \implies x_2 = 1 - 4 \cdot 1 = -3 \\
 2x_1 + 4x_2 - 4x_3 &= 1 \implies x_1 = [1 - 4 \cdot (-3) + 4 \cdot 1]/2 = 17/2.
 \end{aligned}$$



Dieses Eliminationsverfahren kann auch so interpretiert werden, dass man schrittweise das gegebene System durch elementare Operationen in andere Systeme umformt, die die gleiche Lösungsmenge haben. Z.B. hat das gegebene System (1.4) die gleiche Lösung wie das System (1.4), das Dreiecksgestalt hat. Da dieses offensichtlich nur eine Lösung hat (denn beim Rückwärtseinsetzen haben wir hier keine Wahlmöglichkeit), ist auch das gegebene System eindeutig lösbar.

In diesem Beispiel gibt es nur eine Lösung, aber unser Vorgehen würde die Lösungsmenge auch nicht verändern, wenn es mehrere Lösungen gäbe. Dies folgt daraus, dass alle Eliminationsschritte umkehrbar sind.

Als **elementare Zeilen-Operation** [*elementary row operation*] zählen:

- i) das Vertauschen von Gleichungen (Zeilen),
- ii) die Addition/Subtraktion eines Vielfachen der Pivotgleichung (Pivotzeile) zu/von einer anderen Gleichung (Zeile).

Man kann als weitere elementare Zeilen-Operation die Multiplikation einer Gleichung mit einer von null verschiedenen Zahl zulassen. Es zeigt sich aber, dass man ohne diese Operation auskommt. Beim Rechnen von Hand wird sie aber manchmal mit Vorteil angewandt.

Zwei Gleichungssysteme, welche die gleiche Lösungsmenge besitzen, nennen wir **äquivalent** [*equivalent*].

Wir können das Vorgehen, das wir im Prinzip auch im allgemeinen Falle $m \neq n$ beibehalten werden, damit wie folgt zusammenfassen:

Grundidee der Gauss-Elimination:

Ein lineares System aus m Gleichungen in n Unbekannten wird durch die elementaren Zeilen-Operationen i) und ii) schrittweise in äquivalente, gleich grosse Systeme verwandelt, wobei es nach j Schritten ein Teilsystem von $m - j$ Gleichungen gibt, das nur noch höchstens $n - j$ Unbekannte enthält.

Normalerweise wird man zunächst x_1 eliminieren, dann x_2 , und so weiter bis und mit x_{n-1} . Dann kann man x_n berechnen, daraus x_{n-1} , und so weiter, bis man schliesslich x_1 erhält. Ist $m = n$ und ist x_n eindeutig bestimmt, hat man auf diese Weise die eindeutige Lösung des Systems bestimmt. Im allgemeinen geht das nicht immer so, aber im Normalfall funktioniert das mindestens theoretisch, d.h. wenn man von allfälligen Problemen mit Rundungsfehlern absieht.

In der Box 1.1 auf Seite 1-8 ist dieser Algorithmus zunächst nochmals formuliert für diesen Normalfall. Kann man in der angedeuteten Weise eine eindeutige Lösung eines quadratischen linearen Gleichungssystems berechnen, so nennen wir dieses **regulär** [*non-singular*], andernfalls heisst es **singulär** [*singular*].

BEMERKUNGEN:

- 1) Für ein reguläres $n \times n$ -System braucht es $n - 1$ Eliminations-schritte; dann hat das resultierende äquivalente System obere Dreiecksgestalt. Das letzte, n te Pivotelement $a_{nn}^{(n-1)}$ braucht man nur beim Rückwärtseinsetzen.
- 2) Wir nummerieren die Zeilen (Gleichungen) und Kolonnen (Unbekannte) im Restgleichungssystem immer entsprechend der Position im ganzen System. Die Pivotzeile im j -ten Eliminationsschritt hat anfänglich den Index p , nach dem Zeilenvertauschen den Index j .
- 3) Im Prinzip kann jedes von 0 verschiedene Element in der vordersten Kolonne eines Restgleichungssystems als Pivot gewählt werden. Für die Handrechnung ist natürlich eine Eins besonders praktisch. Beim Rechnen mit runden Zahlen beschränkter Genauigkeit (insbesondere in Gleitkomma-Arithmetik) ist es aber wichtig, Elemente zu wählen, die einen grossen Absolutbetrag haben relativ

Box 1.1 Gauss-Elimination im Falle eines regulären Systems**Algorithmus 1.1 (Gauss-Algorithmus, regulärer Fall)**

Zum Lösen des $n \times n$ -Gleichungssystems $\mathbf{Ax} = \mathbf{b}$ führe wie folgt für $j = 1, \dots, n - 1$ je einen Eliminationsschritt durch. Das gegebene System bezeichnen wir hier als 0-tes Restgleichungssystem mit Koeffizientenmatrix $\mathbf{A}^{(0)} := \mathbf{A}$ und rechter Seite $\mathbf{b}^{(0)} := \mathbf{b}$.

- (a) Wähle in der vordersten Kolonne des Restgleichungssystems das j -te Pivotelement $a_{pj}^{(j-1)} \neq 0$ aus. Für den Zeilenindex p der Pivotzeile gilt dabei $p \in \{j, \dots, n\}$. Falls $p \neq j$, vertausche die Zeile p mit der Zeile j und nummeriere die Koeffizienten und rechten Seiten entsprechend um; das Pivotelement heisst dann $a_{jj}^{(j-1)}$.

Lässt sich kein j -tes Pivotelement finden, weil die vorderste Kolonne des Restgleichungssystems lauter Nullen enthält, so bricht der Algorithmus ab: das System ist singulär.

- (b) Berechne für $k = j + 1, \dots, n$ die Faktoren

$$l_{kj} := a_{kj}^{(j-1)} / a_{jj}^{(j-1)} \quad (1.10)$$

und subtrahiere das l_{kj} -fache der Zeile mit Index j von der Zeile mit Index k :

$$a_{ki}^{(j)} := a_{ki}^{(j-1)} - l_{kj} a_{ji}^{(j-1)}, \quad i = j + 1, \dots, n, \quad (1.11a)$$

$$b_k^{(j)} := b_k^{(j-1)} - l_{kj} b_j^{(j-1)}. \quad (1.11b)$$

Ist nach $n - 1$ Eliminationsschritten $a_{nn}^{(n-1)} \neq 0$, so gilt dieses Element als n -tes Pivotelement; andernfalls bricht der Algorithmus ab: das System ist singulär.

Berechne durch Rückwärtseinsetzen die eindeutig bestimmte Lösung des Systems: berechne dazu für $k = n, n - 1, \dots, 1$

$$x_k := \left(b_k^{(k-1)} - \sum_{i=k+1}^n a_{ki}^{(k-1)} x_i \right) \frac{1}{a_{kk}^{(k-1)}}. \quad (1.12)$$

zu den anderen Elementen derselben Zeile. Die Umsetzung dieser Tatsache führt auf Regeln zur Wahl der Pivotelemente (**Pivotstrategien**). Wir kommen auf Seite 3-8 in Abschnitt 3.1 darauf zurück.

4) Die Faktoren l_{kj} ergeben sich als Quotienten aus den Koeffizienten der vordersten Kolonne des j -ten Restgleichungssystems und dem Pivot.

5) Die Lösung ist eindeutig bestimmt (wenn der Algorithmus nicht abbricht), weil das Rückwärtseinsetzen zu einer eindeutigen Lösung führt und das nach den $n - 1$ Eliminationsschritten erhaltene System von Dreiecksgestalt äquivalent ist mit dem gegebenen System. ▼

1.2 Gauss-Elimination: der allgemeine Fall

Im Gauss-Algorithmus 1.1 tritt eine Ausnahmesituation ein, wenn die erste Kolonne des gegebenen Systems oder die erste Kolonne eines Restgleichungssystems lauter Nullen enthält. Dann kann man nicht nach der entsprechenden Variablen auflösen. Aber man braucht diese Variable auch nicht mehr zu eliminieren; sie kommt ja dann scheinbar gar nicht mehr vor. Genauer ausgedrückt, hat der Wert einer solchen Variablen keinen Einfluss, das heisst man kann diese Variable frei wählen. Man nennt sie deshalb **freie Variable** [*free variable*] oder **freier Parameter** [*free parameter*].

Wir wollen diese Ausnahmesituation eines singulären quadratischen linearen Systems ($m = n$) gerade im Rahmen des allgemeinen Falles eines "rechteckigen" Systems, wo $m \neq n$ erlaubt ist, behandeln. Zuerst betrachten wir aber ein kleines 3×3 -Beispiel, in dem freie Variablen auftreten.

BEISPIEL 1.3: Bei der Reduktion des Systems

$$\begin{aligned} 2x_1 - x_2 + 2x_3 &= 2 \\ 4x_1 - 2x_2 + 2x_3 &= 6 \\ 8x_1 - 4x_2 + 6x_3 &= 10 \end{aligned} \quad (1.13)$$

auf obere Dreiecksgestalt werden im ersten Eliminationsschritt nicht nur in der ersten Kolonne, sondern auch in der zweiten zwei Nullen unterhalb der Pivotzeile erzeugt. Wir zeigen dies hier wieder links mit den vollen Eliminationsschemata und rechts in der kompakten Darstellung.

$$\begin{array}{c} \begin{array}{c|ccc|c} & x_1 & x_2 & x_3 & 1 \\ \hline 2 & \textcircled{2} & -1 & 2 & 2 \\ & 4 & -2 & 2 & 6 \\ & 8 & -4 & 6 & 10 \end{array} & \begin{array}{c|ccc|c} & x_1 & x_2 & x_3 & 1 \\ \hline 2 & \textcircled{2} & -1 & 2 & 2 \\ & 4 & -2 & 2 & 6 \\ & 8 & -4 & 6 & 10 \end{array} \\ \\ \begin{array}{c|ccc|c} & x_1 & x_2 & x_3 & 1 \\ \hline & 2 & -1 & 2 & 2 \\ & 0 & 0 & \textcircled{-2} & 2 \\ 1 & 0 & 0 & -2 & 2 \end{array} & \begin{array}{c|cc|c} & x_2 & x_3 & 1 \\ \hline & 0 & \textcircled{-2} & 2 \\ 1 & 0 & -2 & 2 \end{array} \end{array}$$

Die vorderste Kolonne des 2×2 -Restgleichungssystems besteht aus lauter Nullen, was offensichtlich daher kommt, dass im gegebenen System die zweite Kolonne gerade ein Vielfaches der ersten ist. Die Wahl von x_2 hat keinen Einfluss auf den Wert der linken Seite des Restgleichungssystems: x_2 darf also irgend einen Wert annehmen, das heisst, es ist frei wählbar. Wir können nun in der vordersten Kolonne des Restgleichungssystems aber kein Pivotelement festlegen und gehen stattdessen direkt zur nächsten Kolonnen über. Das eingekreiste Pivotelement -2 in der oberen Zeile liefert im nächsten Schritt die folgenden Endschemata:

$$\begin{array}{ccc|c}
 x_1 & x_2 & x_3 & 1 \\
 \hline
 \textcircled{2} & -1 & 2 & 2 \\
 0 & 0 & \textcircled{-2} & 2 \\
 0 & 0 & 0 & 0
 \end{array}
 \quad
 \begin{array}{c}
 1 \\
 \hline
 0
 \end{array}
 \quad (1.14)$$

Hier ist die letzte Gleichung ($0x_1 + 0x_2 + 0x_3 = 0$ bzw. $0 = 0$ — auf der linken Seite kommen keine Variablen mehr vor) für beliebige (x_1, x_2, x_3) erfüllt. Rückwärtseinsetzen in der zweiten und der ersten Gleichung liefert $x_3 = -1$ und

$$x_1 = (2 - 2x_3 + x_2)/2 = (4 + x_2)/2,$$

also die allgemeine Lösung

$$(x_1, x_2, x_3) = (2 + \frac{1}{2}x_2, x_2, -1) \quad (1.15)$$

mit frei wählbarem x_2 .

Wir hatten grosses Glück: Wäre die Konstante in der letzten gegebenen Gleichung in (1.13) nicht gerade 10 sondern zum Beispiel 13, so hätten wir stattdessen die folgenden Endschemata bekommen:

$$\begin{array}{ccc|c}
 x_1 & x_2 & x_3 & 1 \\
 \hline
 \textcircled{2} & -1 & 2 & 2 \\
 0 & 0 & \textcircled{-2} & 2 \\
 0 & 0 & 0 & 3
 \end{array}
 \quad
 \begin{array}{c}
 1 \\
 \hline
 3
 \end{array}
 \quad (1.16)$$

Nun hat die letzte Gleichung und damit das ganze System offensichtlich keine Lösung. \blacklozenge

Wir schliessen aus diesem Beispiel, dass es bei einem singulären System (d.h. einem $n \times n$ System, bei dessen Reduktion freie Variablen auftreten) rechte Seiten gibt, für die das System keine Lösung hat. Offenbar ist dies sogar der Normalfall bei einem singulären System. Denn es kann nur Lösungen geben, wenn im letzten Restgleichungssystem die rechte Seite null ist.

Zu betonen ist nochmals, dass das Auftreten freier Variablen bei quadratischen linearen Systemen eine Ausnahme darstellt. Wir werden aber später, beim Lösen von Eigenwertproblemen in Kapitel 9 bewusst solche Ausnahmefälle herbeiführen.

Nun betrachten wir ein grösseres, "rechteckiges" Gleichungssystem.

BEISPIEL 1.4:

$$\begin{array}{cccccccccc|c}
 & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 & 1 \\
 0 & \textcircled{1} & 0 & 5 & 0 & 4 & 0 & 0 & 1 & 0 & 1 \\
 1 & 0 & 0 & 0 & 5 & 0 & 24 & 16 & 8 & 6 & 12 \\
 1 & 1 & 5 & 9 & 3 & 6 & 1 & 0 & 1 & 0 & 3 \\
 1 & 1 & 10 & 13 & 11 & 8 & 6 & 3 & 1 & 2 & 8 \\
 0 & 0 & 5 & 4 & 18 & 2 & 18 & 12 & 2 & 7 & 13 \\
 1 & 1 & 10 & 13 & 21 & 8 & 24 & 16 & 5 & 8 & 19 \\
 0 & 0 & 5 & 4 & 13 & 2 & 24 & 17 & 6 & 7 & 16
 \end{array} \quad (1.17)$$

Das ist ein System von sieben Gleichungen in neun Unbekannten. Damit es überhaupt auf die Seite passt, haben wir es direkt als Eliminationsschema hingeschrieben. Da es darin mehr Unbekannte als Gleichungen gibt, dürfen wir erwarten, dass das System eine ganze Schar von Lösungen hat. Es wäre allerdings auch in dieser Situation möglich, dass es keine Lösung gibt, was allerdings ein ganz besonderer Ausnahmefall wäre, wie wir in Kürze sehen werden.

Wir haben durch den Kreis bereits angedeutet, dass wir $a_{11} = 1$ als erstes Pivot wählen wollen, und wir haben die entsprechenden Faktoren für den ersten Eliminationsschritt links angefügt. Er liefert das folgende zweite Schema:

$$\begin{array}{cccccccccc|c}
 & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 & 1 \\
 1 & 1 & 0 & 5 & 0 & 4 & 0 & 0 & 1 & 0 & 1 \\
 0 & 0 & 0 & 0 & 5 & 0 & 24 & 16 & 8 & 6 & 12 \\
 0 & 0 & 5 & 4 & 3 & 2 & 1 & 0 & 0 & 0 & 2 \\
 0 & 0 & 10 & 8 & 11 & 4 & 6 & 3 & 0 & 2 & 7 \\
 0 & 0 & 5 & 4 & 18 & 2 & 18 & 12 & 2 & 7 & 13 \\
 0 & 0 & 10 & 8 & 21 & 4 & 24 & 16 & 4 & 8 & 18 \\
 0 & 0 & 5 & 4 & 13 & 2 & 24 & 17 & 6 & 7 & 16
 \end{array}$$

Nun ist im Restgleichungssystem allerdings das erste Element der vorersten Kolonne null, also nicht als Pivot brauchbar. Die Vertauschung der zweiten und dritten Zeile (d.h. der ersten und zweiten Gleichung des Restgleichungssystems) liefert ein modifiziertes zweites Schema mit Pivotelement 5 in der linken oberen Ecke. Die für den zweiten Eliminationsschritt nötigen Faktoren schreiben wir wieder links hin:

$$\begin{array}{cccccccccc|c}
 & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 & 1 \\
 0 & 1 & 0 & 5 & 0 & 4 & 0 & 0 & 1 & 0 & 1 \\
 2 & 0 & \textcircled{5} & 4 & 3 & 2 & 1 & 0 & 0 & 0 & 2 \\
 0 & 0 & 0 & 0 & 5 & 0 & 24 & 16 & 8 & 6 & 12 \\
 1 & 0 & 10 & 8 & 11 & 4 & 6 & 3 & 0 & 2 & 7 \\
 1 & 0 & 5 & 4 & 18 & 2 & 18 & 12 & 2 & 7 & 13 \\
 2 & 0 & 10 & 8 & 21 & 4 & 24 & 16 & 4 & 8 & 18 \\
 1 & 0 & 5 & 4 & 13 & 2 & 24 & 17 & 6 & 7 & 16
 \end{array}$$

Das resultierende dritte Schema sieht so aus:

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	1
1	0	5	0	4	0	0	1	0	1
0	5	4	3	2	1	0	0	0	2
0	0	0	5	0	24	16	8	6	12
0	0	0	5	0	4	3	0	2	3
0	0	0	15	0	17	12	2	7	11
0	0	0	15	0	22	16	4	8	14
0	0	0	10	0	23	17	6	7	14

Nun enthält die zu x_3 gehörende vorderste Kolonne des Restgleichungssystems lauter Nullen. Dies gilt auch für die Kolonne zu x_5 , was aber im Moment noch nicht wichtig ist. Die Variable x_3 braucht also nicht eliminiert zu werden. Wir können im nächsten Schritt direkt x_4 eliminieren, und wir könnten dazu das Pivot 5 in der obersten Zeile wählen. Es zeigt sich allerdings, dass man so auf grosse Zahlen stösst. Wir wollen deshalb, obwohl nicht unbedingt nötig, die dritte und die vierte Zeile vertauschen:

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	1
1	0	5	0	4	0	0	1	0	1
0	5	4	3	2	1	0	0	0	2
0	0	0	5	0	4	3	0	2	3
1	0	0	5	0	24	16	8	6	12
3	0	0	15	0	17	12	2	7	11
3	0	0	15	0	22	16	4	8	14
2	0	0	10	0	23	17	6	7	14

Nach einem weiteren Schritt ist die x_5 -Kolonne, die wieder nur Nullen enthält, im Restgleichungssystem zuvorderst, so dass man direkt zur Elimination von x_6 übergehen kann. Wir wollen dabei aber nochmals, um die Zahlen klein zu halten, die vierte und die fünfte Gleichung freiwillig vertauschen. Dann ergibt sich der Reihe nach:

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	1
1	0	5	0	4	0	0	1	0	1
0	5	4	3	2	1	0	0	0	2
0	0	0	5	0	4	3	0	2	3
0	0	0	0	0	0	20	13	8	9
0	0	0	0	0	0	5	3	2	2
0	0	0	0	0	0	10	7	4	5
0	0	0	0	0	0	15	11	6	8

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	1
1	0	5	0	4	0	0	1	0	1
0	5	4	3	2	1	0	0	0	2
0	0	0	5	0	4	3	0	2	3
0	0	0	0	0	0	5	3	2	2
4	0	0	0	0	0	20	13	8	9
2	0	0	0	0	0	10	7	4	5
3	0	0	0	0	0	15	11	6	8

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	1
	1	0	5	0	4	0	0	1	0	1
	0	5	4	3	2	1	0	0	0	2
	0	0	0	5	0	4	3	0	2	3
	0	0	0	0	0	5	3	2	1	2
	0	0	0	0	0	0	①	0	0	1
1	0	0	0	0	0	0	1	0	0	1
2	0	0	0	0	0	0	2	0	0	2

Jetzt kommt im Restgleichungssystem nur noch x_7 vor. Wir wollen dieses noch aus den zwei letzten Gleichungen eliminieren. Dann erhalten wir das folgende Endschema, in dem wir wieder alle fünf Pivotelemente markiert haben:

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	1
	①	0	5	0	4	0	0	1	0	1
	0	⑤	4	3	2	1	0	0	0	2
	0	0	0	⑤	0	4	3	0	2	3
	0	0	0	0	0	⑤	3	2	1	2
	0	0	0	0	0	0	①	0	0	1
	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0	0

(1.18)

Das Endschema hat nun **Zeilenstufenform** [row echelon form] statt der spezielleren Dreiecksgestalt.

Aus der Zeilenstufenform bekommt man wiederum durch Rückwärtseinsetzen die Lösungen des letzten und damit auch des gegebenen Systems (denn diese sind ja äquivalent). In unserem Beispiel sind die letzten zwei Gleichungen für beliebige Werte der x_k erfüllt. Die drittletzte Zeile kann man nach x_7 auflösen, die viertletzte nach x_6 , die fünftletzte (d.h. die dritte) nach x_4 , die zweite nach x_2 und die erste nach x_1 . Dabei darf man die übrigen Variablen x_9 , x_8 , x_5 und x_3 frei wählen: es sind freie Variable. Dagegen sind die Werte jener Variablen, nach denen man auflöst, eindeutig bestimmt, sobald man die Werte der freien Variablen vorgegeben hat. Jene Variablen stehen alle am Kopf einer Pivotzeile und können deshalb als **Pivotvariable** [pivot variable] bezeichnet werden.

Eine "einfache" Lösung bekommt man, wenn man alle freien Variablen null wählt. In unserem Beispiel ergibt sich dann

$$\begin{aligned}
 x_7 &= 1, \\
 x_6 &= [2 - 3 \cdot 1]/5 = -1/5, \\
 x_4 &= [3 - 3 \cdot 1 - 4 \cdot (-1/5)]/5 = 4/25, \\
 x_2 &= [2 - 0 \cdot 1 - 1 \cdot (-1/5) - 3 \cdot (4/25)]/5 = 43/125, \\
 x_1 &= 1 - 0 \cdot 1 - 0 \cdot (-1/5) - 0 \cdot (4/25) - 0 \cdot (43/125) = 1.
 \end{aligned}$$

Also ist das 9-Tupel

$$(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9) = \left(1, \frac{43}{125}, 0, \frac{4}{25}, 0, -\frac{1}{5}, 1, 0, 0 \right) \quad (1.19)$$

eine Lösung unseres Systems (1.17). Man sieht hier übrigens, dass "einfach" in diesem Zusammenhang nicht unbedingt bedeutet, dass die Lösung eines ganzzahligen Systems nur ganze Zahlen und Brüche mit möglichst kleinen Nennern hat.

Die **allgemeine Lösung** [*general solution*] des Systems bekommt man, indem man die freien Variablen als Parameter auffasst, was man dadurch betonen kann, dass man sie entsprechend umbenennt:

$$x_3 = \alpha, \quad x_5 = \beta, \quad x_8 = \gamma, \quad x_9 = \delta. \quad (1.20)$$

Dann ergibt das Rückwärtseinsetzen:

$$\begin{aligned} x_7 &= 1, \\ x_6 &= -[1 + \delta + 2\gamma]/5, \\ x_4 &= [4 - 6\delta + 8\gamma]/25, \\ x_2 &= [43 + 23\delta - 14\gamma - 50\beta - 100\alpha]/125, \\ x_1 &= 1 - \gamma - 4\beta - 5\alpha. \end{aligned} \quad (1.21)$$

Die allgemeine Lösung besteht aus (1.20) und (1.21). Es ist eine vierparametrische Schar von Lösungen.

Man kann sich fragen, ob es nicht auch eine ganzzahlige Lösung gibt. Das ist alles andere als klar und im allgemeinen nur mit grossem Aufwand entscheidbar. Man könnte natürlich alle 9-Tupel mit z.B. $x_i \in \{-20, -19, \dots, -1, 0, 1, \dots, 19, 20\}$ durchtesten, das sind $41^9 \approx 3.274 \cdot 10^{14}$ Tupel. (Es gibt auch effizientere Vorgehensweisen!) So würde man unter anderem die folgenden drei ganzzahligen Lösungen finden:

$$(1, 1, -1, 0, 1, -1, 1, 1, 2), \quad (1.22a)$$

$$(-15, -1, 1, 4, -1, -7, 1, 15, 4), \quad (1.22b)$$

$$(-17, -2, 1, 2, 2, -2, 1, 5, -1). \quad (1.22c)$$



Wie erkennt man wohl allgemein die Fälle, wo es keine Lösung gibt? Das einfachste Beispiel ist die Gleichung $0x_1 = 1$. Im Schlusschema (1.16) des modifizierten Beispiels 1.3 hatten wir analog die unlösbare Gleichung $0x_1 + 0x_2 + 0x_3 = 3$.

BEISPIEL 1.5: Im Schlusschema (1.18) wären wir hier ebenso auf ein unlösbares Teilsystem gestossen, wenn in den zwei letzten Gleichungen die rechte Seite nicht null gewesen wäre. Man verifiziert leicht, dass dies eintritt, wenn man im Ausgangssystem (1.17) die zwei letzten Konstanten 19 und 16 durch ein anderes Zahlenpaar ersetzt. Genauer: das Gleichungssystem

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	1
1	0	5	0	4	0	0	1	0	1
0	0	0	5	0	24	16	8	6	12
1	5	9	3	6	1	0	1	0	3
1	10	13	11	8	6	3	1	2	8
0	5	4	18	2	18	12	2	7	13
1	10	13	21	8	24	16	5	8	$c_6 + 19$
0	5	4	13	2	24	17	6	7	$c_7 + 16$

(1.23)

mit noch wählbaren Parametern c_6 und c_7 hat die Zeilenstufenform

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	1
①	0	5	0	4	0	0	1	0	1
0	⑤	4	3	2	1	0	0	0	2
0	0	0	⑤	0	4	3	0	2	3
0	0	0	0	0	⑤	3	2	1	2
0	0	0	0	0	0	①	0	0	1
0	0	0	0	0	0	0	0	0	c_6
0	0	0	0	0	0	0	0	0	c_7

(1.24)

Dieses Endschema hat offensichtlich nur dann eine Lösung, wenn $c_6 = c_7 = 0$ gilt. Man beachte, dass der Zusammenhang der Konstanten auf den rechten Seiten von (1.23) und (1.24) nur in den letzten zwei Zeilen so einfach ist. ◆

Der allgemeine Fall der Gauss-Elimination ist als Algorithmus 1.2 in der Box 1.2 auf Seite 1-16 formuliert. Wir lassen dabei sogar zu, dass das gegebene $m \times n$ -System Kolonnen hat, die aus lauter Nullen bestehen. In den Restgleichungssystemen können weitere Kolonnen mit lauter Nullen entstehen. Der Eliminationsprozess ist zu Ende, falls es keine potentielle neue Pivotvariable mehr gibt (wenn $n_j + \ell > n$ in (a) oder $n_j = n$ in (b)) oder keine Gleichungen übrig bleiben (wenn $j = m$ in (b)). Den ersten Fall kann man noch aufteilen in zwei: entweder sind alle Koeffizienten im Restgleichungssystem null oder es bleiben gar keine zu eliminierende Unbekannte mehr übrig.

Nach dem Eliminationsprozess hat das Gleichungssystem die folgende **Zeilenstufenform** bestehend aus r Pivotzeilen und, falls $m > r$, weiteren $m - r$ Zeilen mit lauter Nullen auf der linken Seite:

x_{n_1}	\dots	x_{n_2}	\dots	\dots	x_{n_r}	$x_{n_{r+1}}$	\dots	x_n	1
a_{1,n_1}	\dots	*	\dots	\dots	*	*	\dots	*	c_1
0	\dots	$a_{2,n_2}^{(1)}$	\dots	\dots	*	*	\dots	*	c_2
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
0	\dots	0	\dots	\dots	$a_{r,n_r}^{(r-1)}$	*	\dots	*	c_r
0	\dots	0	\dots	\dots	0	\dots	\dots	0	c_{r+1}
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
0	\dots	0	\dots	\dots	0	\dots	\dots	0	c_m

(1.28)

In diesem Schema haben wir allerdings angenommen, dass die erste Kolonne von \mathbf{A} nicht null ist (d.h. $n_1 = 1$) und dass $m > r$ ist. Zudem haben wir die rechten Seiten umbenannt:

$$c_k := b_k^{(k-1)} \quad (k = 1, \dots, r), \quad c_l := b_l^{(r)} \quad (l = r + 1, \dots, m). \tag{1.29}$$

DEFINITION: In der Zeilenstufenform (1.28) des Systems (1.1)–(1.2) heisst die Anzahl r der Pivotelemente in Algorithmus 1.2 der **Rang [rank]** der Matrix \mathbf{A} und des Gleichungssystems. Die im Falle $m > r$ für die Existenz einer Lösung notwendigen Bedingungen

$$c_{r+1} = c_{r+2} = \dots = c_m = 0 \tag{1.30}$$

sind die **Verträglichkeitsbedingung [consistency conditions]** des Systems. ▲

Box 1.2 Reduktion eines $m \times n$ Systems auf Zeilenstufenform mittels Gauss-Elimination.

Algorithmus 1.2 (Gauss-Algorithmus, allgemeiner Fall)

Zum Lösen des $m \times n$ -Gleichungssystems $\mathbf{Ax} = \mathbf{b}$ setze $j := 1$, $n_1 := 1$, $\mathbf{A}^{(0)} := \mathbf{A}$, $\mathbf{b}^{(0)} := \mathbf{b}$, und betrachte das gegebene System als 0-tes Restgleichungssystem.

j ter Eliminationsschritt:

- (a) Sind alle Koeffizienten in den ℓ vordersten Kolonnen des $(j-1)$ sten Restgleichungssystems null, so sind $x_{n_j}, \dots, x_{n_j+\ell-1}$ freie Variablen; streiche diese ℓ Kolonnen und setze $n_j := n_j + \ell$; falls dann $n_j > n$, setze $r := j - 1$ und gehe zum Verträglichkeitstest.

Wähle in der vordersten Kolonne des verbleibenden Restgleichungssystems das j -te Pivotelement $a_{p,n_j}^{(j-1)} \neq 0$ aus. Für den Zeilenindex p der Pivotzeile gilt dabei $p \in \{j, \dots, m\}$. Falls $p \neq j$, vertausche die Zeile p mit der Zeile j und nummeriere die Koeffizienten und rechten Seiten entsprechend um; das Pivotelement heisst dann $a_{j,n_j}^{(j-1)}$.

- (b) Falls $j = m$, setze $r = m$ und gehe zum Rückwärtseinsetzen. Andernfalls berechne für $k = j + 1, \dots, m$ die Faktoren

$$l_{kj} := a_{k,n_j}^{(j-1)} / a_{j,n_j}^{(j-1)} \quad (1.25)$$

und subtrahiere das l_{kj} -fache der Pivotzeile (mit Index j) von der Zeile mit Index k :

$$a_{ki}^{(j)} := a_{ki}^{(j-1)} - l_{kj} a_{ji}^{(j-1)}, \quad i = n_j + 1, \dots, n, \quad (1.26a)$$

$$b_k^{(j)} := b_k^{(j-1)} - l_{kj} b_j^{(j-1)}. \quad (1.26b)$$

Falls $n_j = n$, so dass (1.26a) leer ist, setze $r = j$ und gehe zum Verträglichkeitstest. Andernfalls setze $n_{j+1} := n_j + 1$ und $j := j + 1$ und beginne den nächsten Eliminationsschritt.

Verträglichkeitstest: Falls $m > r$ und $b_k^{(r)} \neq 0$ für ein $k > r$, so hat das System keine Lösung; Abbruch.

Rückwärtseinsetzen: Bestimme eine Lösung des resultierenden Systems in Zeilenstufenform: berechne dazu für $k = r, r - 1, \dots, 1$

$$x_{n_k} := \left(b_k^{(k-1)} - \sum_{i=n_k+1}^n a_{ki}^{(k-1)} x_i \right) \frac{1}{a_{k,n_k}^{(k-1)}}, \quad (1.27)$$

wobei die freien Variablen (d.h. jene x_i mit $i \notin \{n_1, \dots, n_r\}$) frei wählbar sind.

1.3 Die Lösungsmenge eines linearen Gleichungssystems

In diesem Abschnitt wollen wir aus der Zeilenstufenform (1.28) eines allgemeinen linearen Gleichungssystems (1.1) aus m Gleichungen in n Unbekannten eine Reihe von einfachen theoretischen Aussagen ableiten.

Der Rang r des Systems spielt dabei eine wichtige Rolle. Er ist nach Definition gleich der Anzahl der Pivotelemente, die bei der Reduktion auf Zeilenstufenform mittels Gauss-Elimination (Algorithmus 1.2) auftreten oder auch die Zahl der Zeilen in der Zeilenstufenform der Koeffizientenmatrix, die nicht lauter Nullen enthalten.

Erinnern wir uns daran, dass aufgrund unserer Konstruktion das gegebene System (1.1) und das daraus abgeleitete System (1.28) in Zeilenstufenform äquivalent sind, d.h. die gleichen Lösungen haben. Beim letzteren können wir aber sofort sehen, ob es Lösungen hat, und wenn ja, können wir die allgemeine Lösung leicht berechnen.

Die Anzahl der freien Variablen ist $n - r$, und die Anzahl der Verträglichkeitsbedingungen ist $m - r$. Aufgrund der Definition von r ist klar, dass gilt: $r \leq \min\{m, n\}$. Daraus ergibt sich bereits der folgende wichtige Satz über Existenz und Eindeutigkeit der Lösung.

Satz 1.1 *Das System (1.28) und damit das System (1.1) hat genau dann (mindestens) eine Lösung, wenn*

- *entweder $r = m$ ist (d.h. keine Verträglichkeitsbedingungen auftreten),*
- *oder $r < m$ und $c_{r+1} = c_{r+2} = \dots = c_m = 0$ ist (d.h. die Verträglichkeitsbedingungen erfüllt sind).*

Gibt es Lösungen, so gilt

- *falls $r = n$: die Lösung ist eindeutig,*
- *falls $r < n$: es gibt eine $(n - r)$ -parametrische Lösungsschar (denn die $n - r$ freien Variablen sind frei wählbar).*

Wir können daraus leicht weitere Schlüsse ziehen. Zur Formulierung verwenden wir wieder die Bezeichnung $\mathbf{Ax} = \mathbf{b}$ aus (1.2b).

Korollar 1.2 *Der Rang r hängt nur von der Koeffizientenmatrix \mathbf{A} ab, aber nicht von den gewählten Pivotelementen oder von der rechten Seite \mathbf{b} .*

BEWEIS: Der Rang bestimmt die Zahl $n - r$ freier Parameter in der Lösungsmenge. Letztere bleibt aber während der Gauss-Elimination unverändert. Zudem ist r unabhängig von der rechten Seite des Systems. Also muss r eindeutig bestimmt sein durch die Koeffizientenmatrix \mathbf{A} . (Die Stichhaltigkeit dieses Beweises wird später untermauert werden.)

■

Aufgrund von Korollar 1.2 schreiben wir $r = \text{Rang } \mathbf{A}$ [$r = \text{rank } \mathbf{A}$].

Korollar 1.3 *Es gelten die folgenden Äquivalenzen:*

$$\begin{aligned}
 (i) \quad & \left. \begin{array}{l} \mathbf{Ax} = \mathbf{b} \\ \text{hat genau} \\ \text{eine Lösung} \end{array} \right\} \iff \left\{ \begin{array}{l} r = n = m \quad \text{oder} \\ \left\{ \begin{array}{l} r = n < m \quad \text{und} \\ c_{r+1} = \dots = c_m = 0 \end{array} \right\} \end{array} \right\}, \\
 (ii) \quad & \left. \begin{array}{l} \mathbf{Ax} = \mathbf{b} \\ \text{hat für jedes } \mathbf{b} \\ \text{(mindestens) eine Lösung} \end{array} \right\} \iff r = m, \\
 (iii) \quad & \left. \begin{array}{l} \mathbf{Ax} = \mathbf{b} \\ \text{hat für jedes } \mathbf{b} \\ \text{genau eine Lösung} \end{array} \right\} \iff r = m = n.
 \end{aligned}$$

BEWEIS: (i) Dies folgt sofort aus Satz 1.1.

(ii) Man kann sich überlegen — wir werden das später leicht einsehen — dass das gegebene System (1.1) genau dann für alle b_1, \dots, b_m lösbar ist, wenn das Zeilenstufensystem (1.28) für alle c_1, \dots, c_m lösbar ist, was nach Satz 1.1 genau dann der Fall ist, wenn $r = m$, d.h. wenn keine Verträglichkeitsbedingungen existieren.

(iii) Zusätzlich zur Bedingung aus (ii) verlangen wir hier noch die Eindeutigkeit, die gemäss Satz 1.1 $r = n$ erfordert. ■

Man beachte, dass der Fall (iii) im Korollar 1.3 gerade bedeutet, dass das System und \mathbf{A} regulär sind. Dies setzt immer voraus, dass das System quadratisch ist, d.h. $m = n$ gilt. In diesem Falle ergibt sich aus (ii) und (iii) noch die folgende interessante Aussage.

Korollar 1.4 *Die Lösung eines quadratischen linearen Gleichungssystems (mit $m = n$) ist genau dann eindeutig, wenn das System für beliebige rechte Seiten lösbar ist.*

Reguläre Systeme sind in den Anwendungen mit Abstand am wichtigsten. Fast alle Computerprogramme für die Lösung linearer Gleichungssysteme setzen ein reguläres System voraus.

In der Theorie von spezieller Bedeutung sind Systeme mit einer rechten Seite aus lauter Nullen. Wir schreiben dann $\mathbf{b} = \mathbf{o}$, das heisst $\mathbf{Ax} = \mathbf{o}$. Ein solches System hat immer $x_1 = x_2 = \dots = x_n = 0$ als Lösung.

DEFINITION: Ein lineares Gleichungssystem heisst **homogen** [*homogeneous*], falls die rechte Seite aus Nullen besteht. Andernfalls heisst es **inhomogen** [*inhomogeneous*]. Die Lösung $x_1 = x_2 = \dots = x_n = 0$ eines homogenen Systems heisst **triviale Lösung** [*trivial solution*]. Jede andere Lösung nennt man **nichttrivial** [*non-trivial*]. ▲

Für homogene Systeme kann man die vorangehenden Aussagen über die Lösungsmenge präzisieren. Aus $b_1 = \dots = b_m = 0$ im gegebenen System (1.1), folgt in der Tat, dass $c_1 = \dots = c_m = 0$ im Zeilenstufensystem (1.28). Damit sind allfällige Verträglichkeits-

bedingungen immer erfüllt.

Der erste Teil von Satz 1.1 ist im Falle $\mathbf{b} = \mathbf{o}$ trivial, und der zweite Teil kann so umformuliert werden:

Korollar 1.5 *Ein homogenes Gleichungssystem hat genau dann nichttriviale Lösungen, wenn $r < n$ ist. Falls letzteres zutrifft, gibt es eine $(n - r)$ -parametrische Lösungsschar. Insbesondere hat ein homogenes System mit $m < n$, d.h. mit mehr Unbekannten als Gleichungen, stets eine mindestens $(n - m)$ -parametrische Schar nichttrivialer Lösungen.*

Aus den Teilen (ii) und (iii) von Korollar 1.3 können wir weiter folgern, dass ein Zusammenhang besteht zwischen der Lösungsmenge eines inhomogenen linearen $n \times n$ Gleichungssystems und jener des zugehörigen homogenen Systems, das dieselbe Koeffizientenmatrix besitzt:

Korollar 1.6 *Ein quadratisches lineares Gleichungssystem (mit $m = n$) ist genau dann für beliebige rechte Seiten lösbar, wenn das zugehörige homogene System nur die triviale Lösung besitzt.*

BEWEIS: Beide Bedingungen sind äquivalent zu $r = n = m$. ■

Fassen wir zum Schluss noch einige Fakten über quadratische lineare Systeme zusammen:

Korollar 1.7 *Für ein quadratisches lineares Gleichungssystem $\mathbf{Ax} = \mathbf{b}$ von n Gleichungen in n Unbekannten gelten entweder die vier äquivalenten Aussagen*

- (i) $r := \text{Rang } \mathbf{A} = n$ (d.h. \mathbf{A} ist regulär),
- (ii) für jedes \mathbf{b} gibt es (mindestens) eine Lösung,
- (iii) für jedes \mathbf{b} gibt es genau eine Lösung,
- (iv) das entsprechende homogene System hat nur die triviale Lösung,

oder es gelten die fünf ebenfalls äquivalenten Aussagen

- (v) $r := \text{Rang } \mathbf{A} < n$ (d.h. \mathbf{A} ist singulär),
- (vi) für gewisse \mathbf{b} gibt es keine Lösung,
- (vii) für kein \mathbf{b} gibt es eine eindeutige Lösung,
- (viii) für gewisse \mathbf{b} gibt es unendlich viele Lösungen,
- (ix) das entsprechende homogene System hat nichttriviale Lösungen.

Wir werden in Kapitel 3 auf die Gauss-Elimination zurückkommen und sie neu interpretieren als Matrixzerlegung.

Kapitel 2

Matrizen und Vektoren im \mathbb{R}^n und \mathbb{C}^n

Matrizen, Zeilen- und Spaltenvektoren sind die Basiselemente der Anwendungs-orientierten linearen Algebra, also insbesondere der auf Computern implementierten numerischen linearen Algebra. Es lassen sich aber auch typische Fragen der analytischen Geometrie des Raumes als Matrixaufgaben formulieren. Wir werden später sogar sehen, dass sich viele Aufgabenstellungen aus abstrakteren Räumen, jedenfalls solange diese endlichdimensional sind, auf Matrixaufgaben zurückführen lassen, nämlich indem man im Raum ein Koordinatensystem festlegt. Matrizen und die entsprechenden Vektoren sind also für das Weitere von zentraler Bedeutung.

2.1 Matrizen, Zeilen- und Spaltenvektoren

DEFINITIONEN: Eine $m \times n$ -**Matrix** [$m \times n$ matrix; m -by- n matrix] (pl. **Matrizen** [matrices]) ist ein rechteckiges Schema von mn (reellen oder komplexen) Zahlen, genannt **Elemente** [elements] und angeordnet in m **Zeilen** [rows] und n **Spalten** oder **Kolonnen** [columns]. Das Paar (m, n) definiert die **Grösse** [size] der Matrix.

Das (i, j) -Element der Matrix¹ \mathbf{A} , welches in der i -ten Zeile und in der j -ten Kolonne steht, bezeichnen wir mit a_{ij} oder $(\mathbf{A})_{ij}$. Also:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}. \quad (2.1)$$

Manchmal schreibt man auch etwas ungenau $\mathbf{A} = (a_{ij})$.

Die Elemente a_{jj} ($j = 1, 2, \dots, \min\{m, n\}$) heissen **Diagonalelemente** [diagonal elements]. Die Gesamtheit der Diagonalelemente bildet die **(Haupt-)Diagonale** [(main) diagonal] der Matrix \mathbf{A} .

Eine $m \times 1$ -Matrix heisst **Spaltenvektor** oder **Kolonnenvektor** [column vector] oder auch m -**Vektor** [m -vector].

Eine $1 \times n$ -Matrix nennen wir **Zeilenvektor** [row vector] oder n -**Tupel** [n -tuple].

¹Wir wählen für Matrizen Grossbuchstaben. Zur besseren Kennzeichnung sind sie in diesem Skript halbfett gesetzt, was in Handschrift durch Unterstreichen oder Unterwellen markiert wird, aber auch weggelassen werden kann, wenn keine Verwechslungsgefahr besteht.

Die k -ten Elemente des Kolonnenvektors \mathbf{x} und des Zeilenvektors² $\underline{\mathbf{w}}$ nennen wir k -te **Komponente** [*component*] des Vektors und bezeichnen sie mit x_k bzw. w_k :

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}, \quad \underline{\mathbf{w}} = (w_1 \quad w_2 \quad \dots \quad w_n).$$



BEISPIELE 2.1:

$$\mathbf{A} = \begin{pmatrix} 5 & 3 & 1 \\ 4 & -1 & 4 \end{pmatrix}$$

ist eine reelle, ganzzahlige 2×3 -Matrix. Das zweite Element in ihrer ersten Zeile ist $(\mathbf{A})_{12} = a_{12} = 3$. Dagegen ist

$$\mathbf{B} = \begin{pmatrix} 1.234567 + 4.567890i & 9.876543 + 6.543210i \\ 2.345678 + 5.432109i & 8.765432 + 5.432109i \\ 3.456789 + 3.210987i & 7.654321 + 4.321098i \end{pmatrix}$$

eine komplexe 3×2 -Matrix, wobei z.B. $(\mathbf{B})_{31} = b_{31} = 3.456789 + 3.210987i$.

Das Folgende sind zwei Kolonnenvektoren und ein Zeilenvektor:

$$\mathbf{x} = \begin{pmatrix} 1.05 \\ 2.16 \\ 3.27 \\ 4.38 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} \gamma \\ 2\gamma \end{pmatrix}, \quad \underline{\mathbf{w}} = (0 \quad 1 \quad \pi \quad 2\pi).$$



BEMERKUNGEN:

- 1) Wir arbeiten vorzugsweise mit Kolonnenvektoren und nur selten mit Zeilenvektoren.
- 2) Einige Autoren verwenden für Matrizen und Vektoren statt runde eckige Klammern.
- 3) Oft trennt man die Komponenten eines Zeilenvektors durch Kommas, schreibt also zum Beispiel

$$(w_1, w_2, \dots, w_5) = (4, -7, 12, 1, -9).$$



Einige spezielle Matrizentypen:

Eine $n \times n$ -Matrix (d.h. eine Matrix mit n Kolonnen und gleich vielen Zeilen) ist eine **quadratische Matrix** [*square matrix*] der **Ordnung** [*order*] n .

²Wir bezeichnen Kolonnen- und Zeilenvektoren mit kleinen lateinischen Buchstaben, die ebenfalls halbfett gesetzt sind. Wir kennzeichnen Zeilenvektoren vorderhand durch unterstreichen, obwohl das nicht üblich ist. Später werden wir statt $\underline{\mathbf{w}}$ meist \mathbf{w}^T schreiben, wobei dann \mathbf{w} der entsprechende Kolonnenvektor ist.

Eine $m \times n$ -Matrix deren Elemente alle null sind heisst **Nullmatrix** [*zero matrix*]. Sie wird mit \mathbf{O} bezeichnet³. Analog ist der **Nullvektor** [*zero vector*] ein Kolonnenvektor mit lauter Nullen als Komponenten; er wird hier mit \mathbf{o} bezeichnet⁴.

BEISPIEL 2.2: Die 2×3 -Nullmatrix und der Nullvektor mit 3 Komponenten sind gegeben durch

$$\mathbf{O} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{o} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \quad \blacklozenge$$

Eine $n \times n$ -Matrix \mathbf{D} heisst **diagonal** [*diagonal*], d.h. ist eine **Diagonalmatrix** [*diagonal matrix*], falls $(\mathbf{D})_{ij} = 0$ für $i \neq j$. Für die Diagonalmatrix mit gegebenen Diagonalelementen $d_{11}, d_{22}, \dots, d_{nn}$ schreiben wir

$$\mathbf{D} = \text{diag} (d_{11}, d_{22}, \dots, d_{nn}).$$

BEISPIEL 2.3:

$$\mathbf{D} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \text{diag} (1, 5, 3, -1). \quad \blacklozenge$$

Die $n \times n$ -Matrix $\mathbf{I}_n = \text{diag} (1, 1, \dots, 1)$ ist die **Einheitsmatrix** [*unit matrix*] oder **Identität** [*identity*] der Ordnung n . Oft schreiben wir einfach \mathbf{I} und entnehmen die Grösse dem Kontext.

BEISPIEL 2.4:

$$\mathbf{I} = \mathbf{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad \blacklozenge$$

Eine Matrix \mathbf{R} ist eine **obere Dreiecksmatrix** [*upper triangular matrix*] oder **Rechtsdreiecksmatrix**, falls $(\mathbf{R})_{ij} = 0$ für $i > j$.

BEISPIEL 2.5:

$$\mathbf{R} = \begin{pmatrix} 1 & 3 & 5 & 7 \\ 0 & 2 & 4 & 6 \\ 0 & 0 & 3 & 6 \\ 0 & 0 & 0 & 4 \end{pmatrix}. \quad \blacklozenge$$

Eine Matrix \mathbf{L} ist eine **untere Dreiecksmatrix** [*lower triangular matrix*] oder **Linksdreiecksmatrix**, falls $(\mathbf{L})_{ij} = 0$ für $i < j$.

BEISPIEL 2.6:

$$\mathbf{L} = \begin{pmatrix} 9.8765 & 0 & 0 \\ 7.6543 & 8.7654 & 0 \\ 4.3210 & 5.4321 & 6.5432 \end{pmatrix}. \quad \blacklozenge$$

³Viele Autoren schreiben allerdings 0 (Null) statt \mathbf{O} (gross "Oh").

⁴Üblich ist auch hier die nicht gerade konsequente Bezeichnung durch eine gewöhnliche Null: 0.

Die Mengen der reellen bzw. komplexen $m \times n$ -Matrizen bezeichnen wir mit

$$\mathbb{R}^{m \times n} \quad \text{und} \quad \mathbb{C}^{m \times n} \quad (2.2)$$

und jener der reellen bzw. komplexen n -Vektoren mit

$$\mathbb{R}^n \quad \text{und} \quad \mathbb{C}^n. \quad (2.3)$$

Es wird Aussagen geben die sowohl im Reellen als auch im Komplexen gelten, wo aber im ersten Falle *alle* Grössen reell sein müssen. Diese können wir für beide Fälle formulieren, indem wir die Mengen der $m \times n$ -Matrizen und n -Vektoren bezeichnen mit

$$\mathbb{E}^{m \times n} \quad \text{und} \quad \mathbb{E}^n, \quad \text{wobei} \quad \mathbb{E} := \mathbb{R} \text{ oder } \mathbb{C}. \quad (2.4)$$

2.2 Das Rechnen mit Matrizen und Vektoren

In diesem Abschnitt definieren wir die *Addition und Multiplikation zweier Matrizen* sowie deren *Multiplikation mit einem Skalar*. Zeilen- und Kolonnenvektoren sind abgedeckt als Spezialfälle von Matrizen mit nur einer Zeile oder Kolonne.

Es ist naheliegend, was man unter dem Vielfachen einer Matrix versteht, d.h. wie man sie mit einer Zahl (oder, wie man oft sagt, mit einem Skalar) multipliziert. Für Matrizen sind aber auch Addition und Multiplikation definiert, vorausgesetzt, dass die Operanden passende Grösse haben: während Summanden die gleiche Grösse haben müssen, kommt es beim Produkt darauf an, dass die "Breite" des linken Faktors mit der "Höhe" des rechten Faktors übereinstimmt. Die Definitionen sind unter diesen Voraussetzungen auch auf Zeilen- und Kolonnenvektoren anwendbar.

DEFINITION *der Multiplikation einer Matrix mit einer Zahl:* Eine $m \times n$ -Matrix \mathbf{A} wird **mit einer Zahl (einem Skalar) α multipliziert** [*multiplied by a scalar*], indem man jedes Element von \mathbf{A} mit α multipliziert. Die resultierende $m \times n$ -Matrix $\alpha\mathbf{A}$ mit

$$(\alpha\mathbf{A})_{ij} := \alpha(\mathbf{A})_{ij} \quad (i = 1, \dots, m; j = 1, \dots, n)$$

heisst **Vielfaches** [*multiple*] (genauer: α -faches) der Matrix \mathbf{A} . ▲

BEISPIELE 2.7:

$$5 \begin{pmatrix} 1 & -3 & 5 \\ -2 & 4 & -6 \end{pmatrix} = \begin{pmatrix} 5 & -15 & 25 \\ -10 & 20 & -30 \end{pmatrix}, \quad \frac{1}{4} \begin{pmatrix} 4 \\ 8 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}. \quad \blacklozenge$$

DEFINITION *der Addition zweier Matrizen:* Zwei $m \times n$ -Matrizen \mathbf{A} und \mathbf{B} werden **addiert** [*added*], indem man entsprechende Elemente addiert. Die resultierende $m \times n$ -Matrix $\mathbf{A} + \mathbf{B}$ mit

$$(\mathbf{A} + \mathbf{B})_{ij} := (\mathbf{A})_{ij} + (\mathbf{B})_{ij} \quad (i = 1, \dots, m; j = 1, \dots, n)$$

heisst **Summe** [*sum*] der Matrizen \mathbf{A} und \mathbf{B} . ▲

BEISPIELE 2.8:

$$\begin{pmatrix} 5 & 2 \\ -1 & -5 \end{pmatrix} + \begin{pmatrix} -1 & 1 \\ 6 & 5 \end{pmatrix} = \begin{pmatrix} 4 & 3 \\ 5 & 0 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + \begin{pmatrix} 0.9 & 0.8 & 0.7 \\ 0.6 & 0.5 & 0.4 \end{pmatrix} = \begin{pmatrix} 1.9 & 2.8 & 3.7 \\ 4.6 & 5.5 & 6.4 \end{pmatrix},$$

$$\begin{pmatrix} 1 \\ -1 \\ -3 \end{pmatrix} + \begin{pmatrix} 2 \\ 5 \\ 9 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 6 \end{pmatrix}.$$

◆

Im Gegensatz zur Summe ist das Produkt zweier Matrizen nicht etwa durch elementweise Multiplikation definiert; eine solche Operation würde in Anwendungen nur äusserst selten gebraucht.

DEFINITION *der Multiplikation zweier Matrizen:* Eine $m \times n$ -Matrix \mathbf{A} kann mit einer $n \times p$ -Matrix \mathbf{B} **multipliziert** [*multiplied*] werden, indem man die Elemente des resultierenden **Produktes** [*product*] \mathbf{AB} wie folgt definiert:

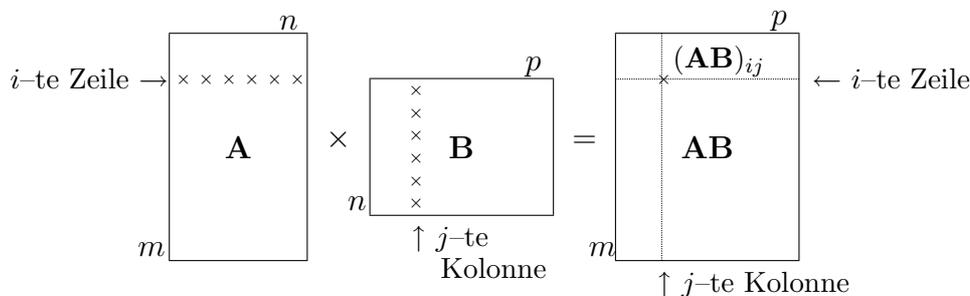
$$\begin{aligned} (\mathbf{AB})_{ij} &::= \sum_{k=1}^n (\mathbf{A})_{ik} (\mathbf{B})_{kj} \\ &= (\mathbf{A})_{i1} (\mathbf{B})_{1j} + (\mathbf{A})_{i2} (\mathbf{B})_{2j} + \cdots + (\mathbf{A})_{in} (\mathbf{B})_{nj} \\ &\quad (i = 1, \dots, m; j = 1, \dots, p). \end{aligned}$$

Das Produkt $\mathbf{C} ::= \mathbf{AB}$ ist also eine $m \times p$ -Matrix, die — in einfacherer Notation — berechnet wird gemäss

$$\begin{aligned} c_{ij} &::= \sum_{k=1}^n a_{ik} b_{kj} \\ &= a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{in} b_{nj} \\ &\quad (i = 1, \dots, m; j = 1, \dots, p). \end{aligned} \tag{2.5}$$

▲

Das Produkt \mathbf{AB} kann also nur gebildet werden, wenn die Anzahl Kolonnen von \mathbf{A} mit der Anzahl Zeilen von \mathbf{B} übereinstimmt. Es lässt sich wie folgt veranschaulichen:



Das Element $(\mathbf{AB})_{ij}$ in der i -ten Zeile und der j -ten Spalte von \mathbf{AB} bekommt man also, indem man die i -te Zeile der Matrix \mathbf{A}

mit der j -ten Kolonne der Matrix \mathbf{B} multipliziert, wobei man die Zeile als Zeilenvektor (d.h. $1 \times n$ -Matrix) auffasst und die Kolonne als Spaltenvektor (d.h. $m \times 1$ -Matrix).

BEISPIELE 2.9: Der Zeilenvektor $(1 \ 2 \ 3)$ kann mit einem Spaltenvektor mit drei Komponenten multipliziert werden; der erste ist eine 1×3 -Matrix, der zweite eine 3×1 -Matrix. Das Produkt ist also eine 1×1 -Matrix:

$$(1 \ 2 \ 3) \begin{pmatrix} 7 \\ -8 \\ 9 \end{pmatrix} = (18),$$

wobei das Resultat aus $1 \cdot 7 + 2 \cdot (-8) + 3 \cdot 9 = 7 - 16 + 27 = 18$ folgt. Wir kommen später auf solche Produkte eines Zeilenvektors mit einem Spaltenvektor zurück. Wir werden in Zukunft eine 1×1 -Matrix als Zahl (Skalar) auffassen und das obige Resultat einfach als 18 schreiben statt (18) .

Die 2×3 -Matrix \mathbf{A} und die 3×4 -Matrix \mathbf{B} seien gegeben als

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 7 & -1 & 1 & 0 \\ -8 & -2 & 0 & -1 \\ 9 & -3 & 0 & 0 \end{pmatrix}.$$

Ihr Produkt ist dann die 2×4 -Matrix

$$\mathbf{AB} = \begin{pmatrix} 18 & -14 & 1 & -2 \\ 42 & -32 & 4 & -5 \end{pmatrix}.$$

Dabei ergibt sich zum Beispiel das $(2, 1)$ -Element 42 gemäss

$$\begin{aligned} a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} &= 4 \cdot 7 + 5 \cdot (-8) + 6 \cdot 9 \\ &= 28 - 40 + 54 \\ &= 42. \end{aligned}$$



BEMERKUNG: Der tiefere Grund für die etwas komplizierte Art das Matrixprodukt zu definieren, wird später ersichtlich werden, wenn wir lineare Abbildungen mit Hilfe von Matrizen beschreiben. Die Matrizen-Multiplikation entspricht dann dem Zusammensetzen von linearen Abbildungen, und das Matrix-Vektor-Produkt erlaubt, Bildpunkte auszurechnen. ▼

Aus den Definitionen der drei eingeführten Matrizenoperationen lassen sich leicht eine ganze Reihe von Eigenschaften dieser Operationen herleiten, von denen die wichtigsten nachfolgend aufgeführt sind. Wir können sie auch als “Regeln” bezeichnen, betonen aber, dass wir diese herleiten können und nicht voraussetzen müssen.

Satz 2.1 Die Addition, die Multiplikation und die skalare Multiplikation von Matrizen (und Vektoren) haben die folgenden Eigenschaften, vorausgesetzt dass die Operationen definiert sind:

$$(\alpha\beta)\mathbf{A} = \alpha(\beta\mathbf{A}), \quad (2.6)$$

$$(\alpha\mathbf{A})\mathbf{B} = \alpha(\mathbf{A}\mathbf{B}) = \mathbf{A}(\alpha\mathbf{B}), \quad (2.7)$$

$$(\alpha + \beta)\mathbf{A} = (\alpha\mathbf{A}) + (\beta\mathbf{A}), \quad (2.8)$$

$$\alpha(\mathbf{A} + \mathbf{B}) = (\alpha\mathbf{A}) + (\alpha\mathbf{B}), \quad (2.9)$$

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} \quad (\text{Add. kommutativ}), \quad (2.10)$$

$$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}) \quad (\text{Add. assoziativ}), \quad (2.11)$$

$$(\mathbf{A}\mathbf{B})\mathbf{C} = \mathbf{A}(\mathbf{B}\mathbf{C}) \quad (\text{Mult. assoziativ}), \quad (2.12)$$

$$(\mathbf{A} + \mathbf{B})\mathbf{C} = (\mathbf{A}\mathbf{C}) + (\mathbf{B}\mathbf{C}) \quad (\text{Add./Mult. distributiv}), \quad (2.13)$$

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = (\mathbf{A}\mathbf{B}) + (\mathbf{A}\mathbf{C}) \quad (\text{Add./Mult. distributiv}), \quad (2.14)$$

BEWEIS: Die Eigenschaften (2.6) und (2.8)–(2.11) ergeben sich sofort aus entsprechenden Regeln für reelle und komplexe Zahlen und der Tatsache, dass die skalare Multiplikation und die Addition von Matrizen elementweise definiert sind, wobei natürlich vorausgesetzt werden muss, dass die in einer Regel auftretenden Matrizen die gleiche Grösse haben.

Etwas zu zeigen bleibt also nur, wenn Matrizen-Multiplikationen auftreten wie in (2.7) und (2.12)–(2.14). Es ist zunächst zu verifizieren, dass jeweils die linke und die rechte Seite unter den gleichen Bedingungen an die Grösse der Matrizen definiert sind. Dann zeigt man, dass das (i, j) -Element links und rechts gleich ist. Für (2.7) ist dies einfach. Für die Assoziativität der Multiplikation (2.12) müssen wir annehmen, dass

$$\mathbf{A} \in \mathbb{E}^{m \times n}, \quad \mathbf{B} \in \mathbb{E}^{n \times p}, \quad \mathbf{C} \in \mathbb{E}^{p \times q}.$$

Dann ist

$$\begin{aligned} ((\mathbf{A}\mathbf{B})\mathbf{C})_{ik} &= \sum_{l=1}^p (\mathbf{A}\mathbf{B})_{il} c_{lk} = \sum_{l=1}^p \sum_{j=1}^n a_{ij} b_{jl} c_{lk}, \\ (\mathbf{A}(\mathbf{B}\mathbf{C}))_{ik} &= \sum_{j=1}^n a_{ij} (\mathbf{B}\mathbf{C})_{jk} = \sum_{j=1}^n \sum_{l=1}^p a_{ij} b_{jl} c_{lk}. \end{aligned}$$

Die zwei Summenzeichen darf man vertauschen, also sind die beiden Ausdrücke auf den rechten Seiten identisch.

Für das erste Distributivgesetz (2.13) brauchen wir stattdessen

$$\mathbf{A} \in \mathbb{E}^{m \times n}, \quad \mathbf{B} \in \mathbb{E}^{m \times n}, \quad \mathbf{C} \in \mathbb{E}^{n \times p},$$

und bekommen dann

$$\begin{aligned} ((\mathbf{A} + \mathbf{B})\mathbf{C})_{ij} &= \sum_{k=1}^n (\mathbf{A} + \mathbf{B})_{ik} c_{kj} = \sum_{k=1}^n (a_{ik} + b_{ik}) c_{kj} \\ &= \sum_{k=1}^n (a_{ik} c_{kj} + b_{ik} c_{kj}) = \sum_{k=1}^n a_{ik} c_{kj} + \sum_{k=1}^n b_{ik} c_{kj} \\ &= (\mathbf{A}\mathbf{C})_{ij} + (\mathbf{B}\mathbf{C})_{ij} = (\mathbf{A}\mathbf{C} + \mathbf{B}\mathbf{C})_{ij}. \end{aligned}$$

Der Beweis für das zweite Distributivgesetz (2.14) verläuft natürlich genau analog. ■

BEMERKUNGEN:

- 1) Mehrere der Aussagen von Satz 2.1 bedeuten, dass man im entsprechenden Ausdruck auf die Klammern verzichten kann, z.B. in $\mathbf{A}(\mathbf{BC}) = \mathbf{ABC}$. Weitere fallen weg, weil wir vereinbaren, dass (wie in \mathbb{R} und \mathbb{C}) die skalare und die Matrizen-Multiplikation stärker binden als die Addition. Es ist also etwa $(\mathbf{AB}) + (\mathbf{CD}) = \mathbf{AB} + \mathbf{CD}$.
- 2) Die Matrizen-Multiplikation ist (selbst für quadratische Matrizen gleicher Ordnung) *nicht kommutativ*, d.h. im allgemeinen gilt

$$\boxed{\mathbf{AB} \neq \mathbf{BA}.} \quad (2.15)$$

Gilt für zwei Matrizen \mathbf{A} und \mathbf{B} , dass $\mathbf{AB} = \mathbf{BA}$, so sagt man, dass diese Matrizen **kommutieren** [*commute*].

BEISPIEL 2.10: Für die drei Matrizen

$$\mathbf{A} = \begin{pmatrix} 2 & 6 \\ 1 & 7 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} -3 & -1 \\ 2 & 1 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 15 & 6 \\ 1 & 20 \end{pmatrix}$$

gilt

$$\begin{aligned} \mathbf{AB} &= \begin{pmatrix} 6 & 4 \\ 11 & 6 \end{pmatrix}, & \mathbf{BA} &= \begin{pmatrix} -7 & -25 \\ 5 & 19 \end{pmatrix}, \\ \mathbf{AC} &= \begin{pmatrix} 36 & 132 \\ 22 & 146 \end{pmatrix}, & \mathbf{CA} &= \begin{pmatrix} 36 & 132 \\ 22 & 146 \end{pmatrix}. \end{aligned}$$

Es ist also

$$\mathbf{AB} \neq \mathbf{BA}, \quad \mathbf{AC} = \mathbf{CA}.$$



- 3) Für jede $m \times n$ -Matrix \mathbf{A} gilt dagegen $\mathbf{I}_m \mathbf{A} = \mathbf{A} = \mathbf{A} \mathbf{I}_n$. Die Einheitsmatrix kommutiert also mit jeder quadratischen Matrix der gleichen Ordnung. ▼

Drei weitere wichtige, einfache Eigenschaften der Matrixaddition sind die folgenden. Sie ergeben sich wiederum direkt aus den entsprechenden Eigenschaften der reellen und komplexen Zahlen.

Satz 2.2 (i) *Es gibt eine $m \times n$ -Nullmatrix \mathbf{O} definiert durch $(\mathbf{O})_{ij} := 0$ ($\forall i, j$), so dass für jede $m \times n$ -Matrix \mathbf{A} gilt*

$$\mathbf{A} + \mathbf{O} = \mathbf{O} + \mathbf{A} = \mathbf{A}. \quad (2.16)$$

(ii) Zu jeder $m \times n$ -Matrix \mathbf{A} gibt es eine $m \times n$ -Matrix $-\mathbf{A}$ definiert durch $(-\mathbf{A})_{ij} := -(\mathbf{A})_{ij}$ ($\forall i, j$), so dass

$$\mathbf{A} + (-\mathbf{A}) = (-\mathbf{A}) + \mathbf{A} = \mathbf{O}. \quad (2.17)$$

(iii) Zu zwei $m \times n$ -Matrizen \mathbf{A} und \mathbf{B} gibt es eine $m \times n$ -Matrix \mathbf{X} definiert durch $(\mathbf{X})_{ij} := (\mathbf{B})_{ij} - (\mathbf{A})_{ij}$ ($\forall i, j$), so dass

$$\mathbf{A} + \mathbf{X} = \mathbf{B}. \quad (2.18)$$

Die Nullmatrix in Teil (i) des Satzes ist natürlich jene, die wir bereits in Abschnitt 2.1 erwähnt haben.

Teil (iii) des Satzes könnte man im Prinzip auch aus den Teilen (i) und (ii) und der Assoziativität der Addition (2.11) herleiten. Die Bedeutung von Teil (iii) liegt darin, dass er auf die genaue Definition der Matrixsubtraktion führt:

$$\mathbf{B} - \mathbf{A} := \mathbf{X}, \quad \text{wo } \mathbf{X} \text{ Lösung von (2.18)}. \quad (2.19)$$

BEMERKUNGEN:

1) Beschränken wir uns auf die Matrixaddition, so gelten die Regeln (2.11), (2.16) und (2.17) für beliebige Elemente der Menge der $m \times n$ -Matrizen. Das bedeutet dass diese Menge bezüglich der Addition eine **Gruppe** [group] bildet, die wegen der Kommutativität (2.10) der Addition sogar **kommutativ** [commutative] oder sogenannt **Abelsch**⁵ [Abelian] ist.

2) Beschränken wir uns auf die Menge der quadratischen Matrizen der Ordnung n , so gelten die Eigenschaften aus den Sätzen 2.1 und 2.2 für beliebige Elemente dieser Menge. Die Eigenschaften (2.10)–(2.14) und (2.16)–(2.17) bedeuten dabei gerade, dass diese Menge (bezüglich Addition und Multiplikation) einen sogenannten **Ring** [ring] Matrizen bildet, der wegen (2.15) **nicht-kommutativ** [non-commutative] ist, falls $n > 1$. Weil $\mathbf{I}_n \mathbf{A} = \mathbf{A} \mathbf{I}_n = \mathbf{A}$ für jede $n \times n$ -Matrix \mathbf{A} , sagt man genauer auch, dass die Menge einen **Ring mit Eins** [ring with identity] bildet.

3) Für reelle und komplexe Zahlen folgt aus $\alpha\beta = 0$, dass $\alpha = 0$ oder $\beta = 0$ sein muss. Beim Matrizenprodukt (2.5) ist das nicht der Fall, wenn $n > 0$ ist. Auch dann nicht, wenn man sich auf quadratische Matrizen beschränkt. Zwei $n \times n$ Matrizen \mathbf{A} und \mathbf{B} mit $\mathbf{AB} = \mathbf{O}$ heißen **Nullteiler** [divisor of zero].

BEISPIEL 2.11: Wegen

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

ist klar, dass

$$\underbrace{\begin{pmatrix} 1 & -1 \\ 3 & -3 \end{pmatrix}}_{\mathbf{A}} \underbrace{\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}}_{\mathbf{B}} = \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}}_{\mathbf{O}}. \quad \blacklozenge$$

4) Nach den Bemerkungen 2) und 3) bilden die reellen $n \times n$ Matrizen ($n > 1$) also einen nicht-kommutativen Ring mit Eins und Nullteilern. ▼

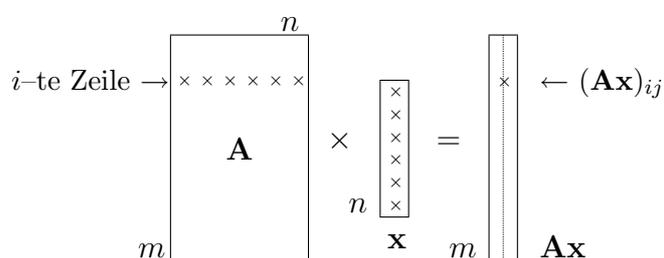
⁵NIELS HENRIK ABEL (5.8.1802 – 6.4.1829), Norwegischer Mathematiker, bewies z.B. 1824, dass im allgemeinen eine Gleichung 5. Grades nicht durch Wurzelziehen lösbar ist.

Von besonderer Bedeutung ist das Produkt einer Matrix mit einem Kolonnenvektor:

DEFINITION *des Matrix-Vektor-Produktes:* Das **Matrix-Vektor-Produkt** [*matrix vector product*] (oder kurz: **Produkt**) einer $m \times n$ -Matrix \mathbf{A} mit einem n -Vektor \mathbf{x} ist definiert als Spezialfall des Produktes zweier Matrizen: $\mathbf{b} \equiv \mathbf{Ax}$ ist gegeben durch

$$\begin{aligned} b_i & \equiv \sum_{k=1}^n a_{ik} x_k \\ & = a_{i1} x_1 + a_{i2} x_2 + \cdots + a_{in} x_n \\ & \quad (i = 1, \dots, m). \end{aligned} \tag{2.20}$$

Als Figur:



Ein spezielles, aber wichtiges Matrixprodukt ist das *Produkt eines Kolonnenvektors \mathbf{x} mit einer 1×1 -Matrix (α), die man als Zahl (Skalar) auffassen kann:* es gilt

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} (\alpha) = \begin{pmatrix} x_1 \alpha \\ \vdots \\ x_n \alpha \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{pmatrix} = \alpha \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Man beachte, dass hier links ein spezielles Matrixprodukt, rechts aber ein äquivalentes Produkt eines Skalars mit einer Matrix steht. Wir definieren deshalb

$$\mathbf{x} \alpha \equiv \mathbf{x} (\alpha) = \alpha \mathbf{x}. \tag{2.21}$$

Wir werden sehen, dass es oft Sinn macht, einen Skalar auf der rechten Seite eines Kolonnenvektors zu schreiben statt wie üblich auf der linken.

Die Summe solcher Produkte des Typs Skalar mal Vektor, das heisst die Summe von Vielfachen von Vektoren, führt zu einem weiteren grundlegenden Begriff der linearen Algebra:

DEFINITION: Eine **Linearkombination** [*linear combination*] der Vektoren $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ ist ein Ausdruck der Form

$$\alpha_1 \mathbf{a}_1 + \alpha_2 \mathbf{a}_2 + \cdots + \alpha_n \mathbf{a}_n, \tag{2.22}$$

worin $\alpha_1, \alpha_2, \dots, \alpha_n$ Zahlen (Skalare) sind. ▲

Man beachte, dass man dank der Priorität der Multiplikation vor der Addition und dank der Assoziativität der Matrixaddition (siehe Satz 2.1) keine Klammern setzen muss. Derartige Vereinfachungen werden laufend benutzt, ohne dass wir darauf hinweisen.

Oft ist es von Vorteil, die Vektoren in einer Linearkombination (2.22) als Kolonnen einer Matrix aufzufassen:

$$\mathbf{A} := \left(\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n \right). \quad (2.23)$$

Um die Struktur hervorzuheben, kann man das auch so schreiben:

$$\mathbf{A} := \left(\begin{array}{c|c|c|c} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{array} \right). \quad (2.24)$$

Mit dieser Notation und der Regel (2.21) kommen wir rasch zu einer Neuinterpretation des Matrix-Vektor-Produktes:

Satz 2.3 Sind $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ gemäss (2.23) die Kolonnenvektoren der $m \times n$ -Matrix \mathbf{A} , und ist \mathbf{x} ein n -Vektor, so gilt:

$$\mathbf{Ax} = \mathbf{a}_1x_1 + \mathbf{a}_2x_2 + \cdots + \mathbf{a}_nx_n = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n. \quad (2.25)$$

Ist insbesondere \mathbf{e}_j der j -te Kolonnenvektor der Einheitsmatrix \mathbf{I}_n , so gilt:

$$\mathbf{Ae}_j = \mathbf{a}_j. \quad (2.26)$$

BEWEIS: Um (2.25) zu beweisen, betrachten wir die i -te Komponente. Wir bezeichnen jene von \mathbf{a}_k mit $(\mathbf{a}_k)_i$, so dass also $(\mathbf{a}_k)_i = a_{ik}$. Dann ist

$$(\mathbf{Ax})_i = \sum_{k=1}^n a_{ik}x_k = \sum_{k=1}^n (\mathbf{a}_k)_i x_k = \sum_{k=1}^n (\mathbf{a}_k x_k)_i = \sum_{k=1}^n (x_k \mathbf{a}_k)_i,$$

wobei wir zuletzt noch (2.21) benutzt haben. Weiter ist Formel (2.26) bloss ein Spezialfall von (2.25). ■

Auf ähnliche Weise kann man das Produkt zweier Matrizen neu interpretieren:

Satz 2.4 Ist \mathbf{A} eine $m \times n$ -Matrix und $\mathbf{B} = \left(\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_p \right)$ eine $n \times p$ -Matrix, so gilt:

$$\mathbf{AB} = \left(\mathbf{Ab}_1 \quad \mathbf{Ab}_2 \quad \cdots \quad \mathbf{Ab}_p \right), \quad (2.27)$$

oder, in anschaulicherer Schreibweise

$$\mathbf{AB} = \left(\begin{array}{c|c|c|c} \mathbf{Ab}_1 & \mathbf{Ab}_2 & \cdots & \mathbf{Ab}_p \end{array} \right). \quad (2.28)$$

BEWEIS: Nach Satz 2.3 ist $\mathbf{b}_j = \mathbf{Be}_j$, $j = 1, \dots, p$. Die j -te Kolonne der Produktmatrix \mathbf{AB} lässt sich deshalb schreiben als $(\mathbf{AB})\mathbf{e}_j$. Dank der Assoziativität der Matrizen-Multiplikation folgt damit aus Satz 2.3 für $j = 1, \dots, p$:

$$(\mathbf{AB})\mathbf{e}_j = \mathbf{A}(\mathbf{Be}_j) = \mathbf{Ab}_j. \quad \blacksquare$$

BEISPIEL 2.12: *Matrixnotation für lineare Gleichungssysteme.*

Mit Hilfe des Matrix-Vektor-Produktes lässt sich offensichtlich ein lineares Gleichungssystem aus m Gleichungen in n Unbekannten,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m, \end{aligned} \quad (2.29)$$

wie angekündigt in kompakter Form schreiben:

$$\boxed{\mathbf{Ax} = \mathbf{b}.} \quad (2.30)$$

Die Matrix \mathbf{A} aus (2.1) ist die **Koeffizientenmatrix** [*coefficient matrix, system matrix*], \mathbf{x} ist der **Lösungsvektor** [*solution vector*] und \mathbf{b} ist die **rechte Seite** [*right-hand side*].

Hat man ℓ verschiedene rechte Seiten $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_\ell$ und entsprechend ℓ Lösungsvektoren $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_\ell$, so kann man diese als Kolonnen zweier Matrizen \mathbf{B} und \mathbf{X} wählen:

$$\mathbf{B} := (\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_\ell), \quad \mathbf{X} := (\mathbf{x}_1 \quad \mathbf{x}_2 \quad \cdots \quad \mathbf{x}_\ell).$$

Auf diese Art lassen sich ℓ Gleichungssysteme zusammenfassen in

$$\boxed{\mathbf{AX} = \mathbf{B}.} \quad (2.31)$$

◆

Ferner sehen wir aus dem Vorangehenden unmittelbar, dass die folgende Aussage gilt:

Satz 2.5 *Das Gleichungssystem $\mathbf{Ax} = \mathbf{b}$ hat genau dann eine Lösung, wenn \mathbf{b} eine Linearkombination der Kolonnen von \mathbf{A} ist.*

2.3 Die Transponierte einer Matrix; symmetrische und Hermitesche Matrizen

DEFINITION: Ist \mathbf{A} eine $m \times n$ -Matrix, so heisst die $n \times m$ -Matrix \mathbf{A}^\top mit $(\mathbf{A}^\top)_{ij} := (\mathbf{A})_{ji}$ die zu \mathbf{A} **transponierte** [*transposed*] Matrix oder die **Transponierte** [*transpose*] von \mathbf{A} .

Ist \mathbf{A} eine komplexe Matrix, so ist $\overline{\mathbf{A}}$ mit $(\overline{\mathbf{A}})_{ij} := \overline{(\mathbf{A})_{ij}}$ die zu \mathbf{A} **konjugiert-komplexe** [*complex conjugate*] Matrix, und

$$\boxed{\mathbf{A}^H := (\overline{\mathbf{A}})^\top = \overline{\mathbf{A}^\top}}$$

ist die zu \mathbf{A} **konjugiert-transponierte** [*conjugate transpose*] oder **Hermitesch-transponierte**⁶ [*Hermitian transpose*] Matrix⁷. ▲

⁶CHARLES HERMITE (24.12.1822 – 14.1.1901), französischer Mathematiker, ab 1870 Professor an der Sorbonne, bewies z.B. die Transzendenz von e .

⁷Oft wird \mathbf{A}^H auch die zu \mathbf{A} **adjungierte** [*adjoint*] Matrix oder kurz die **Adjungierte** [*adjoint*] genannt und mit \mathbf{A}^* bezeichnet.

BEISPIELE 2.13: Für die Matrizen

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} 1.2 + 3.4i & 5.6 - 7.8i \\ 8.7 + 6.5i & 4.3 - 2.1i \end{pmatrix}$$

ist

$$\mathbf{A}^T = \begin{pmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{pmatrix}, \quad \mathbf{C}^H = \begin{pmatrix} 1.2 - 3.4i & 8.7 - 6.5i \\ 5.6 + 7.8i & 4.3 + 2.1i \end{pmatrix}.$$

Weiter ist zum Beispiel allgemein die Transponierte einer unteren Dreiecksmatrix eine obere Dreiecksmatrix. Und es gilt natürlich: *Die Transponierte eines Kolonnenvektors ist ein Zeilenvektor — und umgekehrt.* Dies wird oft ausgenützt, um Kolonnenvektoren Platz sparend aufzuschreiben:

$$\mathbf{x} = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix}^T = \begin{pmatrix} 1 & 3 & 5 & 3 & 5 & 3 & 1 \end{pmatrix}^T. \quad \blacklozenge$$

DEFINITION: Eine Matrix \mathbf{A} heisst **symmetrisch** [*symmetric*], falls

$$\mathbf{A}^T = \mathbf{A}, \quad \text{d.h.} \quad (\mathbf{A})_{ij} = (\mathbf{A})_{ji} \quad (\forall i, j).$$

Eine Matrix \mathbf{A} heisst **Hermitesch** [*Hermitian*], falls

$$\mathbf{A}^H = \mathbf{A}, \quad \text{d.h.} \quad (\mathbf{A})_{ij} = \overline{(\mathbf{A})_{ji}} \quad (\forall i, j). \quad \blacktriangle$$

BEISPIELE 2.14: Die Matrizen

$$\mathbf{B} = \begin{pmatrix} 2 & 3 & -5 \\ 3 & -1 & 2 \\ -5 & 2 & 7 \end{pmatrix}, \quad \mathbf{E} = \begin{pmatrix} 1 & 2 + 3i & 4 + 5i \\ 2 - 3i & 6 & 7 + 8i \\ 4 - 5i & 7 - 8i & 9 \end{pmatrix}$$

sind (reell) symmetrisch bzw. (komplex) Hermitesch. \blacklozenge

Man beachte, dass die Diagonalelemente einer Hermiteschen Matrix reell sind. Ferner ist eine reelle Hermitesche Matrix natürlich symmetrisch. Die Menge der Hermiteschen Matrizen enthält also als Teilmenge die reellen symmetrischen Matrizen. Meist denkt man aber an komplexe Matrizen, wenn man von Hermiteschen Matrizen spricht.

Es gibt auch komplexe symmetrische Matrizen, das heisst komplexe Matrizen mit $\mathbf{A}^T = \mathbf{A}$. Solche kommen z.B. in der Elektrotechnik und in der Teilchenphysik vor.

BEISPIEL 2.15: Die Matrix

$$\begin{pmatrix} 8 + 5i & 1 + 2i & 6i \\ 1 + 2i & 7 + 6i & 2 + 3i \\ 6i & 2 + 3i & 4 \end{pmatrix}$$

ist komplex symmetrisch. Die Diagonalelemente brauchen nicht reell zu sein. \blacklozenge

Eine weitere, oft auftretende Symmetrieeigenschaft ist die folgende:

DEFINITION: Eine quadratische Matrix \mathbf{A} ist **schiefsymmetrisch** [*skew-symmetric*], falls $\mathbf{A}^\top = -\mathbf{A}$. ▲

BEISPIEL 2.16: Die Matrix

$$\begin{pmatrix} 0 & -3 & 5 \\ 3 & 0 & -4 \\ -5 & 4 & 0 \end{pmatrix}$$

ist schiefsymmetrisch und hat wie jede andere solche Matrix lauter Nullen in der Diagonale. ◆

Für das Transponieren gelten folgende einfache Rechenregeln:

Satz 2.6 (i) Für jede Matrix \mathbf{A} gilt

$$\boxed{(\mathbf{A}^\top)^\top = \mathbf{A}}, \quad \boxed{(\mathbf{A}^H)^H = \mathbf{A}}, \quad (2.32)$$

(ii) Für jede Matrix \mathbf{A} und jeden Skalar α gilt

$$\boxed{(\alpha \mathbf{A})^\top = \alpha \mathbf{A}^\top}, \quad \boxed{(\alpha \mathbf{A})^H = \bar{\alpha} \mathbf{A}^H}. \quad (2.33)$$

(iii) Für beliebige $m \times n$ -Matrizen \mathbf{A} und \mathbf{B} gilt

$$\boxed{(\mathbf{A} + \mathbf{B})^\top = \mathbf{A}^\top + \mathbf{B}^\top}, \quad \boxed{(\mathbf{A} + \mathbf{B})^H = \mathbf{A}^H + \mathbf{B}^H}. \quad (2.34)$$

(iv) Für jede $m \times n$ -Matrix \mathbf{A} und jede $n \times p$ -Matrix \mathbf{B} gilt

$$\boxed{(\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top}, \quad \boxed{(\mathbf{AB})^H = \mathbf{B}^H \mathbf{A}^H}. \quad (2.35)$$

BEWEIS: Die Aussagen (2.32)–(2.34) sollten klar sein.

Um (2.35) zu beweisen bemerken wir, dass \mathbf{AB} eine $m \times p$ -Matrix und somit $(\mathbf{AB})^\top$ eine $p \times m$ -Matrix ist. Das Produkt $\mathbf{B}^\top \mathbf{A}^\top$ ist ebenfalls eine $p \times m$ -Matrix. Für die entsprechenden Elemente gilt

$$\begin{aligned} ((\mathbf{AB})^\top)_{ij} &= (\mathbf{AB})_{ji} = \sum_{k=1}^n a_{jk} b_{ki} = \sum_{k=1}^n b_{ki} a_{jk} \\ &= \sum_{k=1}^n (\mathbf{B}^\top)_{ik} (\mathbf{A}^\top)_{kj} = (\mathbf{B}^\top \mathbf{A}^\top)_{ij}. \end{aligned}$$

Der Beweis von $(\mathbf{AB})^H = \mathbf{B}^H \mathbf{A}^H$ verläuft völlig analog. ■

Im allgemeinen ist das Produkt zweier symmetrischer Matrizen nicht symmetrisch. Zum Beispiel ist

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

Aber man hat die folgenden Aussagen:

Satz 2.7 (i) Für symmetrische (quadratische) Matrizen \mathbf{A} , \mathbf{B} gilt:

$$\mathbf{AB} = \mathbf{BA} \iff \mathbf{AB} \text{ symmetrisch.} \quad (2.36)$$

(ii) Für beliebige Matrizen gilt:

$$\mathbf{A}^\top \mathbf{A} \text{ und } \mathbf{AA}^\top \text{ sind symmetrisch,} \quad (2.38)$$

Analoge Aussagen gelten im Hermiteschen Fall.

BEWEIS: (i) Für symmetrische Matrizen mit $\mathbf{AB} = \mathbf{BA}$ folgt nach (2.35)

$$(\mathbf{AB})^\top = (\mathbf{BA})^\top = \mathbf{A}^\top \mathbf{B}^\top = \mathbf{AB},$$

also ist \mathbf{AB} symmetrisch. Ist umgekehrt \mathbf{AB} symmetrisch, hat man

$$\mathbf{AB} = (\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top = \mathbf{BA}.$$

(ii) Zum Beispiel ist nach (2.32) und (2.35)

$$(\mathbf{A}^\top \mathbf{A})^\top = \mathbf{A}^\top (\mathbf{A}^\top)^\top = \mathbf{A}^\top \mathbf{A},$$

also ist $\mathbf{A}^\top \mathbf{A}$ symmetrisch. ■

Mit Satz 2.6 kann man leicht die Aussagen der Sätze 2.3 und 2.4 betreffend die Interpretation der Kolonnenstruktur einer Matrix übertragen auf eine analoge Interpretation der Zeilenstruktur. Um die Zeilenstruktur einer Matrix explizit zu machen, schreiben wir auf unsere unkonventionelle Weise

$$\mathbf{A} \equiv: \begin{pmatrix} \underline{\mathbf{a}}_1 \\ \underline{\mathbf{a}}_2 \\ \vdots \\ \underline{\mathbf{a}}_m \end{pmatrix} \quad \text{oder} \quad \mathbf{A} \equiv: \left(\begin{array}{c} \underline{\mathbf{a}}_1 \\ \underline{\mathbf{a}}_2 \\ \vdots \\ \underline{\mathbf{a}}_m \end{array} \right), \quad (2.40)$$

worin $\underline{\mathbf{a}}_k$ der k -te Zeilenvektoren von \mathbf{A} bezeichnet.

Nun können wir die Sätze 2.3 und 2.4 “transponieren”:

Korollar 2.8 Werden die Zeilenvektoren der $m \times n$ -Matrix \mathbf{A} und der $n \times p$ -Matrix \mathbf{B} gemäss

$$\mathbf{A} = \begin{pmatrix} \underline{\mathbf{a}}_1 \\ \vdots \\ \underline{\mathbf{a}}_m \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} \underline{\mathbf{b}}_1 \\ \vdots \\ \underline{\mathbf{b}}_n \end{pmatrix}$$

bezeichnet, und ist $\underline{\mathbf{y}} = (y_1 \ \dots \ y_n)$ ein Zeilenvektor, so gilt:

$$\boxed{\underline{\mathbf{y}}\mathbf{B} = y_1 \underline{\mathbf{b}}_1 + y_2 \underline{\mathbf{b}}_2 + \dots + y_n \underline{\mathbf{b}}_n}, \quad \boxed{\mathbf{e}_i^\top \mathbf{B} = \underline{\mathbf{b}}_i} \quad (2.41)$$

und

$$\mathbf{AB} = \begin{pmatrix} \underline{\mathbf{a}}_1 \mathbf{B} \\ \underline{\mathbf{a}}_2 \mathbf{B} \\ \vdots \\ \underline{\mathbf{a}}_m \mathbf{B} \end{pmatrix} = \left(\begin{array}{c} \underline{\mathbf{a}}_1 \mathbf{B} \\ \underline{\mathbf{a}}_2 \mathbf{B} \\ \vdots \\ \underline{\mathbf{a}}_m \mathbf{B} \end{array} \right). \quad (2.42)$$

In (2.41) ist \mathbf{e}_i^\top wie üblich der i -te Zeilenvektor von \mathbf{I}_n .

2.4 Das Skalarprodukt und die Norm von Vektoren; Längen und Winkel

Ein n -Vektor \mathbf{x} lässt sich als Ortsvektor im reellen (oder komplexen) n -dimensionalen Euklidischen⁸ Raum \mathbb{R}^n (bzw. \mathbb{C}^n) auffassen, das heisst als Pfeil, der den Ursprung O mit einem anderen Punkt $X = (x_1, x_2, \dots, x_n)$ verbindet. Obwohl wir gewohnt sind Punkte durch ein n -Tupel, d.h. einen Zeilenvektor zu bezeichnen, wollen wir den Vektor von nun an als Spaltenvektor

$$\mathbf{x} = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix}^T$$

wählen.

Die Vektoraddition entspricht offensichtlich dem Zusammensetzen von Pfeilen und die Multiplikation mit einer Zahl einer Verlängerung oder Verkürzung des Pfeiles. Das gilt im n -dimensionalen Raum genau gleich wie im 2- und 3-dimensionalen Raum.

Geometrisch ist klar, was man im \mathbb{R}^n unter der (Euklidischen) Länge $\|\mathbf{x}\|$ eines solchen Pfeiles versteht, und was der Winkel φ zwischen zwei Pfeilen \mathbf{x} und \mathbf{y} ist.

Im \mathbb{R}^2 gilt bekanntlich der Satz von Pythagoras⁹

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2} \quad (2.43)$$

und der daraus folgende Cosinussatz

$$\|\mathbf{y} - \mathbf{x}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - 2\|\mathbf{x}\|\|\mathbf{y}\|\cos\varphi. \quad (2.44)$$

Zum Beweis lesen wir aus einer Figur ab, dass mit

$$a := \|\mathbf{x}\|, \quad b := \|\mathbf{y}\|, \quad c := \|\mathbf{y} - \mathbf{x}\|$$

gilt

$$\begin{aligned} c^2 &= (a - b \cos\varphi)^2 + b^2(\sin\varphi)^2 \\ &= a^2 - 2ab \cos\varphi + b^2 \cos^2\varphi + b^2 \sin^2\varphi \\ &= a^2 + b^2 - 2ab \cos\varphi, \end{aligned}$$

was mit (2.44) identisch ist. Falls $\|\mathbf{x}\| \neq 0$, $\|\mathbf{y}\| \neq 0$ ergibt sich daraus

$$\cos\varphi = \frac{\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - \|\mathbf{y} - \mathbf{x}\|^2}{2\|\mathbf{x}\|\|\mathbf{y}\|}. \quad (2.45)$$

Wenn wir das übliche Skalarprodukt im \mathbb{R}^2 definieren,

$$\langle \mathbf{x}, \mathbf{y} \rangle := x_1 y_1 + x_2 y_2,$$

⁸EUKLID, genauer: EUKLEIDES (um 300 v. Chr.), griechischer Mathematiker in Alexandrien, oft fälschlich als Autor der "Elemente" betrachtet, des Geometrie-Lehrbuchs, das über 2000 Jahre als Standard galt.

⁹PYTHAGORAS (ca. 580 – 500 v.Chr.), griechischer Mathematiker, Astronom, Musikwissenschaftler und Philosoph; ab ca. 530 in Kroton (Oberitalien) Begründer seiner Schule.

erhalten wir wegen

$$\begin{aligned}\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - \|\mathbf{y} - \mathbf{x}\|^2 &= x_1^2 + x_2^2 + y_1^2 + y_2^2 - (y_1 - x_1)^2 - (y_2 - x_2)^2 \\ &= 2(x_1 y_1 + x_2 y_2) \\ &= 2\langle \mathbf{x}, \mathbf{y} \rangle\end{aligned}$$

schliesslich

$$\cos \varphi = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|}. \quad (2.46)$$

Wir wollen die Ergebnisse (2.43), (2.44) und (2.46) auf den \mathbb{R}^n und den \mathbb{C}^n verallgemeinern. Dazu schreiben wir wieder \mathbb{E}^n statt “ \mathbb{R}^n oder \mathbb{C}^n ”.

DEFINITION: Das **(Euklidische) Skalarprodukt** oder **innere Produkt** [*inner product*] zweier Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{E}^n$ ist die Zahl $\langle \mathbf{x}, \mathbf{y} \rangle$ definiert durch

$$\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{x}^H \mathbf{y} = \sum_{k=1}^n \overline{x_k} y_k, \quad (2.47)$$

was sich im Falle reeller Vektoren reduziert auf

$$\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{x}^T \mathbf{y} = \sum_{k=1}^n x_k y_k. \quad (2.48)$$



Es gelten die folgenden Eigenschaften:

Satz 2.9 Für das Skalarprodukt (2.48) im \mathbb{R}^n gilt:

(S1) Es ist **linear** im zweiten Faktor:

$$\begin{aligned}\langle \mathbf{x}, \mathbf{y} + \mathbf{z} \rangle &= \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{z} \rangle && \text{für alle } \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n, \\ \langle \mathbf{x}, \alpha \mathbf{y} \rangle &= \alpha \langle \mathbf{x}, \mathbf{y} \rangle && \text{für alle } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \alpha \in \mathbb{R}.\end{aligned}$$

(S2) Es ist **symmetrisch**:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle \quad \text{für alle } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n.$$

(S3) Es ist **positiv definit**:

$$\begin{aligned}\langle \mathbf{x}, \mathbf{x} \rangle &\geq 0 && \text{für alle } \mathbf{x} \in \mathbb{R}^n, \\ \langle \mathbf{x}, \mathbf{x} \rangle &= 0 && \implies \mathbf{x} = \mathbf{o}.\end{aligned}$$

Für das Skalarprodukt (2.47) im \mathbb{C}^n gelten (S1) und (S3) analog für alle $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{C}^n$ und $\alpha \in \mathbb{C}$, aber (S2) wird ersetzt durch:

(S2') Es ist **Hermiteisch**:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle} \quad \text{für alle } \mathbf{x}, \mathbf{y} \in \mathbb{C}^n.$$

BEWEIS: Wenn man gemäss der Definition $\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{x}^H \mathbf{y}$ das Skalarprodukt als Matrizenprodukt schreibt, so folgt (S1) für \mathbb{R}^n und \mathbb{C}^n sofort aus den Matrix-Eigenschaften (2.14) und (2.7) in Satz 2.1. Das gilt aber nicht für (S2) und (S2'). Dort beachten wir, dass $\mathbf{x}^T \mathbf{y}$ bzw. $\mathbf{x}^H \mathbf{y}$ eine Zahl ist, also ohne Wirkung transponiert werden kann; im übrigen wenden wir (2.35) an: $\mathbf{x}^T \mathbf{y} = (\mathbf{x}^T \mathbf{y})^T = \mathbf{y}^T \mathbf{x}$ und, im komplexen Fall,

$$\mathbf{x}^H \mathbf{y} = (\mathbf{x}^H \mathbf{y})^T = \mathbf{y}^T (\mathbf{x}^H)^T = \overline{\mathbf{y}^H \mathbf{x}}.$$

(Denn es ist ja $\mathbf{y}^T = \overline{\mathbf{y}^H}$ und $\mathbf{x}^H = (\overline{\mathbf{x}})^T$.) Schliesslich: Eigenschaft (S3) ist klar, denn

$$\langle \mathbf{x}, \mathbf{x} \rangle = \mathbf{x}^H \mathbf{x} = \sum_{k=1}^n \overline{x_k} x_k = \sum_{k=1}^n |x_k|^2 \geq 0.$$

Das Gleichheitszeichen gilt nur, wenn alle x_k null sind. ■

Aus den Eigenschaften (S1) und (S2) bzw. (S1) und (S2') folgt sofort:

Korollar 2.10 *Das Skalarprodukt (2.48) im \mathbb{R}^n ist **bilinear** [bilinear], d.h. zusätzlich zu (S1) gilt:*

(S4) *Es ist auch **linear** im ersten Faktor:*

$$\begin{aligned} \langle \mathbf{w} + \mathbf{x}, \mathbf{y} \rangle &= \langle \mathbf{w}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle && \text{für alle } \mathbf{w}, \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \\ \langle \alpha \mathbf{x}, \mathbf{y} \rangle &= \alpha \langle \mathbf{x}, \mathbf{y} \rangle && \text{für alle } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \alpha \in \mathbb{R}. \end{aligned}$$

*Das Skalarprodukt (2.47) im \mathbb{C}^n ist **sesquilinear** [sesquilinear], d.h. zusätzlich zu (S1) gilt:*

(S4') *Es ist **konjugiert-linear** im ersten Faktor:*

$$\begin{aligned} \langle \mathbf{w} + \mathbf{x}, \mathbf{y} \rangle &= \langle \mathbf{w}, \mathbf{y} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle && \text{für alle } \mathbf{w}, \mathbf{x}, \mathbf{y} \in \mathbb{C}^n, \\ \langle \alpha \mathbf{x}, \mathbf{y} \rangle &= \overline{\alpha} \langle \mathbf{x}, \mathbf{y} \rangle && \text{für alle } \mathbf{x}, \mathbf{y} \in \mathbb{C}^n, \alpha \in \mathbb{C}. \end{aligned}$$

Es ist auch klar, dass man im reellen Fall (S1) und (S2) bzw. (S1) und (S4) kombinieren kann zu

$$\begin{aligned} \langle \mathbf{x}, \alpha \mathbf{y} + \beta \mathbf{z} \rangle &= \alpha \langle \mathbf{x}, \mathbf{y} \rangle + \beta \langle \mathbf{x}, \mathbf{z} \rangle, \\ \langle \alpha \mathbf{w} + \beta \mathbf{x}, \mathbf{y} \rangle &= \alpha \langle \mathbf{w}, \mathbf{y} \rangle + \beta \langle \mathbf{x}, \mathbf{y} \rangle. \end{aligned}$$

Im komplexen Fall gilt nach (S1) und (S2') bzw. (S1) und (S4')¹⁰:

$$\begin{aligned} \langle \mathbf{x}, \alpha \mathbf{y} + \beta \mathbf{z} \rangle &= \alpha \langle \mathbf{x}, \mathbf{y} \rangle + \beta \langle \mathbf{x}, \mathbf{z} \rangle, \\ \langle \alpha \mathbf{w} + \beta \mathbf{x}, \mathbf{y} \rangle &= \overline{\alpha} \langle \mathbf{w}, \mathbf{y} \rangle + \overline{\beta} \langle \mathbf{x}, \mathbf{y} \rangle. \end{aligned}$$

Das Skalarprodukt kann nun auch als Basis für die Definition der Länge eines Vektors dienen:

¹⁰Es gibt auch viele Autoren, die das komplexe Skalarprodukt linear im ersten Argument und konjugiert-linear im zweiten Argument wählen. Unsere Wahl ist hier von Vorteil und in der Physik die Übliche.

DEFINITION: Die **Länge** [*length*] oder **2-Norm** [*2-norm*] oder **Euklidische Norm** [*Euclidean norm*] eines Vektors $\mathbf{x} \in \mathbb{E}^n$ ist die nichtnegative reelle Zahl $\|\mathbf{x}\|$ definiert durch

$$\|\mathbf{x}\| := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}. \quad (2.49)$$

Man beachte, dass nach Eigenschaft (S2) gilt $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$, weshalb die Wurzel wohldefiniert ist. ▲

Nach (2.48) gilt im reellen Fall

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}^T \mathbf{x}} = \sqrt{\sum_{k=1}^n x_k^2} \quad (2.50)$$

und nach (2.47) im komplexen Fall

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}^H \mathbf{x}} = \sqrt{\sum_{k=1}^n |x_k|^2}. \quad (2.51)$$

Dass diese Norm (2.49) die übliche Länge eines Vektors liefert, folgt durch rekursives Anwenden des Satzes von Pythagoras auf Paare von Vektoren der Form

$$\begin{pmatrix} x_1 & \dots & x_k & 0 & \dots & 0 \end{pmatrix}^T, \quad \begin{pmatrix} 0 & \dots & 0 & x_{k+1} & 0 & \dots & 0 \end{pmatrix}^T.$$

Um die Länge eines komplexen Vektors $\mathbf{z} = \mathbf{x} + i \mathbf{y} \in \mathbb{C}_n$ geometrisch zu interpretieren, kann man diesen als Vektor im \mathbb{R}^{2n} auffassen, denn es ist ja $|z_k|^2 = x_k^2 + y_k^2$.

BEISPIELE 2.17: Es ist für $\mathbf{x} := \begin{pmatrix} 1 & 2 & 2 \end{pmatrix}^T$

$$\|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle = 1^2 + 2^2 + 2^2 = 9,$$

also $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \sqrt{9} = 3$. Für den komplexen Vektoren

$$\mathbf{z} = \begin{pmatrix} i \\ 3 + 4i \\ 5 + 12i \\ 8 + 15i \end{pmatrix}$$

erhalten wir

$$\|\mathbf{z}\|^2 = 1^2 + (3^2 + 4^2) + (5^2 + 12^2) + (8^2 + 15^2) = 1 + 25 + 169 + 289 = 484,$$

also $\|\mathbf{z}\| = \sqrt{484} = 22$. ◆

Aus der Definition der 2-Norm und den Eigenschaften des Skalarproduktes folgt eine aus dem \mathbb{R}^2 wohlbekannte Ungleichung:

Satz 2.11 Für alle Paare $\mathbf{x}, \mathbf{y} \in \mathbb{E}^n$ gilt die **Schwarzsche Ungleichung** [Schwarz inequality]

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|. \quad (2.52)$$

Das Gleichheitszeichen gilt genau dann, wenn \mathbf{y} ein Vielfaches ist von \mathbf{x} oder umgekehrt.

Statt Schwarzsche Ungleichung sagt man oft **Cauchy-Schwarz-Ungleichung**¹¹ [Cauchy-Schwarz inequality] oder sogar **Cauchy-Bunjakovski-Schwarz-Ungleichung**¹² [CBS inequality].

Die Schwarzsche Ungleichung ist eine Eigenschaft des Skalarproduktes, nicht der Norm, denn quadriert lautet sie

$$|\langle \mathbf{x}, \mathbf{y} \rangle|^2 \leq \langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle. \quad (2.53)$$

Der folgende Beweis, der nur auf den Eigenschaften des Skalarproduktes beruht, liefert deshalb zunächst diese Form.

BEWEIS: Für beliebiges $\alpha \in \mathbb{E}$ ist

$$0 \leq \langle \alpha \mathbf{x} + \mathbf{y}, \alpha \mathbf{x} + \mathbf{y} \rangle = \alpha \bar{\alpha} \langle \mathbf{x}, \mathbf{x} \rangle + \bar{\alpha} \langle \mathbf{x}, \mathbf{y} \rangle + \alpha \overline{\langle \mathbf{x}, \mathbf{y} \rangle} + \langle \mathbf{y}, \mathbf{y} \rangle.$$

Für $\mathbf{x} = \mathbf{o}$ gilt (2.52) offensichtlich, und zwar mit dem Gleichheitszeichen. Wir dürfen also $\mathbf{x} \neq \mathbf{o}$ annehmen und $\alpha = -\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle}$ wählen, womit nach Multiplikation mit $\langle \mathbf{x}, \mathbf{x} \rangle$ folgt:

$$0 \leq |\langle \mathbf{x}, \mathbf{y} \rangle|^2 - |\langle \mathbf{x}, \mathbf{y} \rangle|^2 - |\langle \mathbf{x}, \mathbf{y} \rangle|^2 + \langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle,$$

also (2.53), was äquivalent ist mit (2.52).

Das Gleichheitszeichen gilt genau, wenn $\mathbf{x} = \mathbf{o}$ oder $\alpha \mathbf{x} + \mathbf{y} = \mathbf{o}$ ist, was gerade heisst, dass \mathbf{y} ein Vielfaches von \mathbf{x} ist oder umgekehrt. (Ausser wenn $\mathbf{x} = \mathbf{o}$ oder $\mathbf{y} = \mathbf{o}$ gilt, ist dann \mathbf{y} ein Vielfaches von \mathbf{x} ist und umgekehrt.) ■

BEISPIEL 2.18: Für beliebige reelle Zahlen $x_1, \dots, x_n, y_1, \dots, y_n$ gilt nach (2.53)

$$\left(\sum_{k=1}^n x_k y_k \right)^2 \leq \left(\sum_{k=1}^n x_k^2 \right) \left(\sum_{j=1}^n y_j^2 \right). \quad (2.54)$$

Zum Beispiel erhält man für die zwei Vektoren

$$\mathbf{x} := \begin{pmatrix} 2 & 2 & 1 \end{pmatrix}^T, \quad \mathbf{y} := \begin{pmatrix} 0 & 15 & 8 \end{pmatrix}^T \quad (2.55)$$

¹¹AUGUSTIN LOUIS CAUCHY (21.8.1789 – 23.5.1857), französischer Mathematiker, ab 1916 Professor an der Ecole Polytechnique, später an der Sorbonne; königstreu und streng katholisch; fast 800 Publikationen, darunter wichtige Beiträge zur Gruppentheorie, Analysis und komplexen Funktionentheorie.

HERMANN AMANDUS SCHWARZ (25.1.1843 – 30.11.1921), deutscher Mathematiker, Professor in Zürich, Göttingen und Berlin.

¹²VICTOR JAKOWLEWITSCH BUNJAKOVSKI [englische Transkription: Bunyakovskii], (16.12.1804 – 12.12.1889), russischer Mathematiker in Peterburg.

$\|\mathbf{x}\|^2 = 9$, $\|\mathbf{y}\|^2 = 289$, $\langle \mathbf{x}, \mathbf{y} \rangle = 38$. Also lautet die quadrierte Schwarzsche Ungleichung hier $1444 = 38^2 \leq 9 \cdot 289 = 2601$ und die Schwarzsche Ungleichung selbst ist $38 \leq 3 \cdot 17 = 51$. ♦

Grundlegende Eigenschaften der Norm sind die folgenden:

Satz 2.12 Für die 2-Norm (2.49) im \mathbb{E}^n gilt:

(N1) Sie ist positiv definit:

$$\begin{aligned} \|\mathbf{x}\| &\geq 0 && \text{für alle } \mathbf{x} \in \mathbb{E}^n, \\ \|\mathbf{x}\| &= 0 && \implies \mathbf{x} = \mathbf{o}. \end{aligned}$$

(N2) Sie ist dem Betrage nach homogen:

$$\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\| \quad \text{für alle } \mathbf{x} \in \mathbb{E}^n, \alpha \in \mathbb{E}.$$

(N3) Die **Dreiecksungleichung [triangle inequality]** gilt:

$$\|\mathbf{x} \pm \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \quad \text{für alle } \mathbf{x}, \mathbf{y} \in \mathbb{E}^n.$$

BEWEIS: (N1) und (N2) folgen leicht aus den entsprechenden Eigenschaften des Skalarproduktes. Um (N3) zu beweisen, schätzen wir in

$$\|\mathbf{x} \pm \mathbf{y}\|^2 = \langle \mathbf{x} \pm \mathbf{y}, \mathbf{x} \pm \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle \pm 2 \langle \mathbf{x}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle$$

den gemischten Term $\langle \mathbf{x}, \mathbf{y} \rangle$ mittels der Schwarzschen Ungleichung (2.52) ab durch $2 \|\mathbf{x}\| \|\mathbf{y}\|$

$$\begin{aligned} \|\mathbf{x} \pm \mathbf{y}\|^2 &\leq \|\mathbf{x}\|^2 + 2 \|\mathbf{x}\| \|\mathbf{y}\| + \|\mathbf{y}\|^2 \\ &= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2. \end{aligned}$$

■

Schreibt man im Zähler der Formel (2.45) für $\cos \varphi$ die Quadrate der Normen als Skalarprodukte und wendet man die Rechenregeln aus Satz 2.9 an, so kann man nun allgemein im \mathbb{R}^n und \mathbb{C}^n diesen Zähler durch ein einziges Skalarprodukt ausdrücken (wie wir das in (2.46) im Spezialfall des \mathbb{R}^2 gemacht haben):

$$\begin{aligned} \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - \|\mathbf{y} - \mathbf{x}\|^2 &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - \langle \mathbf{y} - \mathbf{x}, \mathbf{y} - \mathbf{x} \rangle \\ &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 - \langle \mathbf{y}, \mathbf{y} \rangle + \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle - \langle \mathbf{x}, \mathbf{x} \rangle \\ &= 2 \operatorname{Re} \langle \mathbf{x}, \mathbf{y} \rangle. \end{aligned}$$

Also ist

$$\varphi = \arccos \frac{\operatorname{Re} \langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|}. \quad (2.56)$$

Im reellen Fall entfällt natürlich der Realteil:

$$\varphi = \arccos \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|}. \quad (2.57)$$

Diese Formeln liefern den **Winkel [angle]** zwischen zwei Vektoren, wobei $0 \leq \varphi \leq \pi$ gilt. Dass sie nicht nur im \mathbb{R}^2 sondern auch im \mathbb{R}^n und im \mathbb{C}^n gelten, lässt sich daraus schliessen, dass die zwei Vektoren in einer (zweidimensionalen) Ebene liegen, dass der Winkel gemäss (2.45) durch Längen ausgedrückt werden kann und dass, wie wir später sehen werden, Längen und Skalarprodukt nicht vom Koordinatensystem abhängen, solange dieses orthonormiert bleibt.

(2.57) liefert auch eine neue Interpretation der Schwarzschen Ungleichung (2.52) im \mathbb{R}^n : der Faktor, um den die linke Seite von (2.52) kleiner ist als die rechte, ist gerade $|\cos \varphi|$. Das Gleichheitszeichen gilt in (2.52) also genau dann, wenn $|\cos \varphi| = 1$ ist, das heisst wenn der Zwischenwinkel φ null oder π ist.

DEFINITION: Zwei n -Vektoren \mathbf{x} und \mathbf{y} sind zueinander **orthogonal [orthogonal]** (oder: **stehen senkrecht aufeinander [are perpendicular]**) falls $\langle \mathbf{x}, \mathbf{y} \rangle = 0$. Wir schreiben: $\mathbf{x} \perp \mathbf{y}$. ▲

BEISPIEL 2.19: In Beispiel 2.18 erhielten wir für die zwei Vektoren (2.55): $\|\mathbf{x}\|^2 = 9$, $\|\mathbf{y}\|^2 = 289$, $\langle \mathbf{x}, \mathbf{y} \rangle = 38$. Also ist

$$\cos \varphi = \frac{38}{3 \cdot 17} = \frac{38}{51} = 0.745098\dots,$$

was $\varphi = 0.7301145\dots$ oder, in Grad, $\varphi = 41.83248\dots^\circ$ ergibt. ◆

BEISPIEL 2.20: Die Vektoren

$$\mathbf{u} := (1 \ 2 \ 3 \ 4 \ 5 \ 6)^\top, \quad \mathbf{v} := (-6 \ 5 \ -4 \ 3 \ -2 \ 1)^\top$$

stehen wegen $\langle \mathbf{u}, \mathbf{v} \rangle = -6+10-12+12-10+6 = 0$ senkrecht aufeinander. ◆

Zwei senkrechte Vektoren bilden zusammen mit dem an die Spitze von \mathbf{y} verschobenen Vektor $\mathbf{x} - \mathbf{y}$ ein rechtwinkliges Dreieck. Der Satz von Pythagoras nimmt damit folgende Form an:

Satz 2.13 (Satz von Pythagoras) Für $\mathbf{x}, \mathbf{y} \in \mathbb{E}^n$ mit $\mathbf{x} \perp \mathbf{y}$ gilt

$$\|\mathbf{x} \pm \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 \quad (2.58)$$

BEWEIS: Aus der Definition (2.49), den Eigenschaften (S1), (S4) und (S2) bzw. (S2'), sowie der Voraussetzung $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ folgt:

$$\begin{aligned} \|\mathbf{x} \pm \mathbf{y}\|^2 &= \langle \mathbf{x} \pm \mathbf{y}, \mathbf{x} \pm \mathbf{y} \rangle \\ &= \langle \mathbf{x}, \mathbf{x} \rangle \pm \langle \mathbf{x}, \mathbf{y} \rangle \pm \langle \mathbf{y}, \mathbf{x} \rangle + \langle \mathbf{y}, \mathbf{y} \rangle \\ &= \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2, \end{aligned}$$

denn $\langle \mathbf{y}, \mathbf{x} \rangle = \overline{\langle \mathbf{x}, \mathbf{y} \rangle} = 0$. ■

Neben dem Euklidischen Skalarprodukt (2.47)–(2.48) und der davon abgeleiteten (Euklidischen) 2-Norm werden in der Praxis ab und zu noch andere Skalarprodukte und Normen gebraucht. Dabei werden auch Normen eingeführt, die nicht gemäss (2.49) auf einem Skalarprodukt beruhen.

DEFINITION: Ein **Skalarprodukt** oder **inneres Produkt** [*inner product*] im \mathbb{R}^n (bzw. \mathbb{C}^n) ist eine Funktion, die jedem Paar von n -Vektoren eine reelle (bzw. komplexe) Zahl zuordnet, wobei die Eigenschaften (S1)–(S3) [bzw. (S1), (S2'), (S3)] aus Satz 2.9 gelten.

Eine **Norm** [*norm*] im \mathbb{E}^n (d.h. \mathbb{R}^n oder \mathbb{C}^n) ist eine Funktion, die jedem n -Vektor eine nichtnegative reelle Zahl zuordnet, wobei die Eigenschaften (N1)–(N3) aus Satz 2.12 gelten. ▲

Die gebräuchlichsten Normen sind Spezialfälle der **p -Norm**:

$$\|\mathbf{x}\|_p := (|x_1|^p + \cdots + |x_n|^p)^{\frac{1}{p}} \quad (1 \leq p \leq \infty), \quad (2.59)$$

wobei die Fälle 1, 2 und ∞ von grösstem Interesse sind. Dabei ist

$$\|\mathbf{x}\|_1 := (|x_1| + \cdots + |x_n|), \quad \|\mathbf{x}\|_\infty := \max_{k=1, \dots, n} |x_k|. \quad (2.60)$$

BEISPIEL 2.21: Betrachten wir in der normalen Euklidischen Ebene die Vektoren \mathbf{x} , für die $\|\mathbf{x}\|_1 = 1$ ist, so bilden diese den Rand eines Quadrates mit Seitenlänge $\sqrt{2}$, dessen Diagonalen auf den Achsen liegen.

Betrachten wir stattdessen die Vektoren mit $\|\mathbf{x}\|_\infty = 1$, so bilden diese den Rand eines Quadrates mit Seitenlänge 2, dessen Seitenmittelsenkrechten auf den Achsen liegen.

Im \mathbb{R}^3 liefert $\|\mathbf{x}\|_1 = 1$ ein reguläres Oktaeder, $\|\mathbf{x}\|_\infty = 1$ dagegen einen Würfel. ◆

2.5 Das äussere Produkt und die orthogonale Projektion auf eine Gerade

Sind ein m -Vektor \mathbf{x} und ein n -Vektor \mathbf{y} gegeben, so ist das Skalarprodukt (innere Produkt) $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^H \mathbf{y}$ nur definiert wenn $m = n$, in welchem Falle es eine 1×1 -Matrix, d.h. ein Skalar ist.

Dagegen können wir das Produkt $\mathbf{x} \mathbf{y}^H$ immer bilden; es ist eine $m \times n$ -Matrix. Im reellen Fall hat man

$$\begin{aligned} \mathbf{x} \mathbf{y}^T &= \begin{pmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_m \end{pmatrix} \begin{pmatrix} y_1 & \cdots & y_j & \cdots & y_n \end{pmatrix} \\ &= \begin{pmatrix} x_1 y_1 & \cdots & x_1 y_j & \cdots & x_1 y_n \\ \vdots & & \vdots & & \vdots \\ x_i y_1 & \cdots & x_i y_j & \cdots & x_i y_n \\ \vdots & & \vdots & & \vdots \\ x_m y_1 & \cdots & x_m y_j & \cdots & x_m y_n \end{pmatrix}. \end{aligned}$$

DEFINITION: Das **dyadische Produkt** oder **äussere Produkt** [*outer product*] eines m -Vektors \mathbf{x} und eines n -Vektors \mathbf{y} ist die $m \times n$ -Matrix $\mathbf{x} \mathbf{y}^H$ bzw. im reellen Fall die Matrix $\mathbf{x} \mathbf{y}^T$.

Satz 2.14 Eine $m \times n$ Matrix hat genau dann den Rang 1, wenn sie das äussere Produkt eines m -Vektors $\mathbf{x} \neq \mathbf{o}$ und eines n -Vektors $\mathbf{y} \neq \mathbf{o}$ ist.

BEWEIS: Hat eine Matrix \mathbf{A} Rang 1, so gibt es eine Zeile \mathbf{y}^\top , die nicht null ist, und alle andere Zeilen sind Vielfache davon (denn im Gauss'schen Algorithmus kann man Vielfache dieser Zeile von den anderen subtrahieren und erhält dabei lauter Nullzeilen). Im komplexen Fall darf man diese Zeile auch mit \mathbf{y}^H bezeichnen. Ihr Index sei ℓ . Nennen wir den Multiplikator für die i -te Zeile x_i , und setzen wir $x_\ell := 1$ und $\mathbf{x} := (x_1 \ \dots \ x_n)^\top$, so ist gerade $\mathbf{A} = \mathbf{x} \mathbf{y}^H$, wobei $\mathbf{x} \neq \mathbf{o}$ und $\mathbf{y} \neq \mathbf{o}$.

Ist umgekehrt $\mathbf{A} = \mathbf{x} \mathbf{y}^H$ mit $\mathbf{x} \neq \mathbf{o}$, $\mathbf{y} \neq \mathbf{o}$, so gibt es ein $x_\ell \neq 0$, womit die ℓ -te Zeile von \mathbf{A} nicht null ist. Wählt man diese Zeile als Pivotzeile, so kann man, für alle $i \neq \ell$, durch Subtraktion des x_i/x_ℓ -fachen der ℓ -ten Zeile die i -te Zeile zu null machen. Es folgt, dass \mathbf{A} (nach unserer, auf der Anwendung des Gauss-Algorithmus beruhenden Definition des Ranges) den Rang 1 hat. ■

Das äussere Produkt eines Vektors \mathbf{y} mit sich selbst wird gebraucht für die orthogonale Projektion auf die Gerade (durch O) mit der durch \mathbf{y} gegebenen Richtung:

Satz 2.15 Die orthogonale Projektion $\mathbf{P}_y \mathbf{x}$ des n -Vektors \mathbf{x} auf die durch die Vielfachen von \mathbf{y} ($\neq \mathbf{o}$) erzeugte Gerade durch den Ursprung ist gegeben durch

$$\mathbf{P}_y \mathbf{x} := \frac{1}{\|\mathbf{y}\|^2} \mathbf{y} \mathbf{y}^H \mathbf{x} = \mathbf{u} \mathbf{u}^H \mathbf{x}, \quad (2.61)$$

worin $\mathbf{u} := \mathbf{y}/\|\mathbf{y}\|$.

Im Reellen kann man \mathbf{y}^H durch \mathbf{y}^\top ersetzen.

BEWEIS: Die Projektion $\mathbf{P}_y \mathbf{x}$ zeichnet sich geometrisch dadurch aus, dass für alle \mathbf{x} gilt:

- (i) $\mathbf{P}_y \mathbf{x} = \alpha \mathbf{y}$ für einen von \mathbf{x} abhängigen Skalar α ,
- (ii) $\mathbf{x} - \mathbf{P}_y \mathbf{x} \perp \mathbf{y}$.

Diese zwei Eigenschaften sind für den in (2.61) definierten Ausdruck für $\mathbf{P}_y \mathbf{x}$ zu verifizieren, wobei wir annehmen dürfen, dass $\|\mathbf{y}\| = 1$ und damit der Bruch wegfällt.

Zu (i): Es ist $\mathbf{P}_y \mathbf{x} = \mathbf{y} (\mathbf{y}^H \mathbf{x})$, wobei in der Klammer ein Skalar steht. Also gilt (i) mit $\alpha := \mathbf{y}^H \mathbf{x}$. Beachte, dass hier (2.21) angewendet wird.

Zu (ii): Unter Verwendung von $\mathbf{y}^H \mathbf{y} = \|\mathbf{y}\|^2 = 1$ folgt

$$\langle \mathbf{y}, \mathbf{x} - \mathbf{P}_y \mathbf{x} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle - \langle \mathbf{y}, \mathbf{y} \mathbf{y}^H \mathbf{x} \rangle = \mathbf{y}^H \mathbf{x} - \mathbf{y}^H \mathbf{y} \mathbf{y}^H \mathbf{x} = 0. \quad \blacksquare$$

$\mathbf{P}_y \mathbf{x}$ kann man so interpretieren: Auf den zu projizierenden Vektor \mathbf{x} wird die Rang-1-Matrix

$$\mathbf{P}_y := \frac{1}{\|\mathbf{y}\|^2} \mathbf{y} \mathbf{y}^H = \mathbf{u} \mathbf{u}^H \equiv: \mathbf{P}_u \quad (2.62)$$

(von links) angewendet, wobei wieder $\mathbf{u} := \mathbf{y}/\|\mathbf{y}\|$ ist

BEISPIEL 2.22: Gegeben seien die zwei Punkte $X = (4, 8, -1)$ und $Y = (1, 2, 2)$ im \mathbb{R}^3 , also die zwei Vektoren $\mathbf{x} = \begin{pmatrix} 4 & 8 & -1 \end{pmatrix}^\top$ und $\mathbf{y} = \begin{pmatrix} 1 & 2 & 2 \end{pmatrix}^\top$. Man projiziere die Strecke \overline{OX} orthogonal auf die durch die Punkte O und Y laufende Gerade.

Dazu ist bloss die reelle Version der Formel (2.61) auszuwerten, wobei es am effizientesten ist, zuerst $\mathbf{y}^\top \mathbf{x} = 18$ und $\|\mathbf{y}\|^2 = 9$ zu berechnen (und nicht etwa die Matrix $\mathbf{y}\mathbf{y}^\top$). Dann wird

$$\mathbf{P}_y \mathbf{x} = \frac{1}{\|\mathbf{y}\|^2} \mathbf{y} (\mathbf{y}^\top \mathbf{x}) = 2 \mathbf{y} = \begin{pmatrix} 2 & 4 & 4 \end{pmatrix}^\top.$$

Es wird also X auf $X' = (2, 4, 4)$ projiziert. \blacklozenge

Die Projektionsmatrix \mathbf{P}_y aus (2.62) hat zwei interessante Eigenschaften:

$$\boxed{\mathbf{P}_y^H = \mathbf{P}_y, \quad \mathbf{P}_y^2 = \mathbf{P}_y.} \quad (2.63)$$

Die erste bedeutet, dass \mathbf{P}_y Hermitesch oder, falls reell, symmetrisch ist. Aufgrund der zweiten bezeichnet man \mathbf{P}_y als **idempotent** [*idempotent*]. Diese zwei Eigenschaften sind charakteristisch für orthogonale Projektionen und können zu deren Definition verwendet werden. \mathbf{P}_y ist aber eine spezielle orthogonale Projektion, weil ja auf eine Gerade projiziert wird. Im \mathbb{R}^3 (und allgemeiner im \mathbb{E}^n) kann man aber auch auf eine Ebene projizieren. Wir werden darauf zurückkommen.

2.6 Matrizen als lineare Abbildungen

Ist $\mathbf{A} \in \mathbb{E}^{m \times n}$ irgend eine $m \times n$ Matrix, so kann man die ebenfalls mit \mathbf{A} bezeichnete Abbildung

$$\boxed{\mathbf{A} : \mathbb{E}^n \rightarrow \mathbb{E}^m, \quad \mathbf{x} \mapsto \mathbf{A}\mathbf{x}} \quad (2.64)$$

betrachten. Sie hat zwei grundlegende Eigenschaften: für alle $\mathbf{x}, \tilde{\mathbf{x}} \in \mathbb{E}^n$ und alle $\gamma \in \mathbb{E}$ gilt

$$\boxed{\mathbf{A}(\mathbf{x} + \tilde{\mathbf{x}}) = \mathbf{A}\mathbf{x} + \mathbf{A}\tilde{\mathbf{x}}, \quad \mathbf{A}(\gamma\mathbf{x}) = \gamma\mathbf{A}\mathbf{x}.} \quad (2.65)$$

Man kann sie auch zusammenfassen in

$$\boxed{\mathbf{A}(\gamma\mathbf{x} + \tilde{\mathbf{x}}) = \gamma(\mathbf{A}\mathbf{x}) + (\mathbf{A}\tilde{\mathbf{x}}).} \quad (2.66)$$

Wir werden in Kapitel 5 sehen, dass diese Eigenschaften per Definition charakteristisch sind für eine **lineare Abbildung** [*linear transformation*].

Während der Definitionsbereich der Abbildung \mathbf{A} aus (2.64) immer der ganze Raum \mathbb{E}^n ist, ist das **Bild** [*image*] von \mathbf{A} ,

$$\mathbf{im} \mathbf{A} := \{\mathbf{A}\mathbf{x} \in \mathbb{E}^m; \mathbf{x} \in \mathbb{E}^n\} \quad (2.67)$$

im allgemeinen eine Teilmenge von \mathbb{E}^m . Es ist aber immer $\mathbf{o} \in \text{im } \mathbf{A}$, denn $\mathbf{A}\mathbf{o} = \mathbf{o}$.

BEISPIEL 2.23: Die Projektionsmatrix \mathbf{P}_y aus (2.62) ist eine $n \times n$ Matrix, bildet also \mathbb{E}^n in sich ab. Wie wir gesehen haben, stellt sie eine orthogonale Projektion auf die durch den Ursprung verlaufende Gerade mit Richtung \mathbf{y} dar. Das bedeutet insbesondere, dass das Bild im \mathbf{P}_y gerade diese Gerade ist:

$$\text{im } \mathbf{P}_y = \{\alpha \mathbf{y}; \alpha \in \mathbb{E}\}. \quad (2.68)$$



2.7 Die Inverse einer Matrix

Zwei quadratische Matrizen gleicher Ordnung kann man immer zueinander addieren, voneinander subtrahieren und miteinander multiplizieren. Wir werden in diesem Abschnitt sehen, dass man sie oft, aber nicht immer, auch in gewissem Sinne durcheinander dividieren kann.

Vergleicht man $n \times n$ -Matrizen mit reellen oder komplexen Zahlen, so nimmt die Nullmatrix \mathbf{O}_n bei der Matrizen-Addition die Rolle der Null ein, und die Einheitsmatrix \mathbf{I}_n übernimmt bei der Matrizen-Multiplikation die Rolle der Eins. Bei den Zahlen gibt es zu jedem $\alpha \neq 0$ ein ξ , so dass $\alpha \xi = 1 = \xi \alpha$ ist. Gilt dies wohl analog für quadratische Matrizen? Das heisst, gibt es zu $\mathbf{A} \neq \mathbf{O}_n$ eine $n \times n$ -Matrix \mathbf{X} mit $\mathbf{A}\mathbf{X} = \mathbf{I}_n = \mathbf{X}\mathbf{A}$? Ein einfaches Beispiel zeigt, dass das nicht so ist: für

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

folgt bei beliebiger Wahl von \mathbf{X}

$$\mathbf{A}\mathbf{X} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} = \begin{pmatrix} x_{11} & x_{12} \\ 0 & 0 \end{pmatrix} \neq \mathbf{I}_2$$

DEFINITION: Eine $n \times n$ -Matrix \mathbf{A} heisst **invertierbar** [*invertible*], falls es eine $n \times n$ -Matrix \mathbf{X} gibt, so dass $\mathbf{A}\mathbf{X} = \mathbf{I}_n = \mathbf{X}\mathbf{A}$ ist. Die Matrix \mathbf{X} heisst **Inverse** [*inverse*] von \mathbf{A} und wird mit \mathbf{A}^{-1} bezeichnet:

$$\boxed{\mathbf{A} \mathbf{A}^{-1} = \mathbf{I}_n = \mathbf{A}^{-1} \mathbf{A}.} \quad (2.69)$$



BEISPIEL 2.24: Es ist

$$\begin{pmatrix} 2 & 2 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -\frac{1}{2} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} 2 & 2 \\ 1 & 2 \end{pmatrix},$$

es ist also die eine Matrix die Inverse der anderen — und umgekehrt.



BEISPIEL 2.25: Die Einheitsmatrix \mathbf{I}_n ist natürlich invertierbar und ist ihre eigene Inverse, denn $\mathbf{I}_n \mathbf{I}_n = \mathbf{I}_n = \mathbf{I}_n \mathbf{I}_n$. \blacklozenge

Satz 2.16 Ist \mathbf{A} invertierbar, so ist die Inverse eindeutig bestimmt.

BEWEIS: Angenommen \mathbf{X} und \mathbf{Y} sind Inverse von \mathbf{A} , dann ist

$$\mathbf{X} = \mathbf{X} \mathbf{I} = \mathbf{X}(\mathbf{A}\mathbf{Y}) = (\mathbf{X}\mathbf{A})\mathbf{Y} = \mathbf{I}\mathbf{Y} = \mathbf{Y}.$$

■

Man kann die Bedingung für die Inverse effektiv abschwächen: Es genügt, dass $\mathbf{A}\mathbf{X} = \mathbf{I}$ oder $\mathbf{X}\mathbf{A} = \mathbf{I}$ gilt, dann folgt die andere Beziehung automatisch. Dies ist enthalten in

Satz 2.17 Die folgenden Aussagen über eine $n \times n$ -Matrix \mathbf{A} sind äquivalent:

- i) \mathbf{A} ist invertierbar.
- ii) Es gibt eine $n \times n$ -Matrix \mathbf{X} mit $\mathbf{A}\mathbf{X} = \mathbf{I}_n$.
- iii) Es gibt genau eine $n \times n$ -Matrix \mathbf{X} mit $\mathbf{A}\mathbf{X} = \mathbf{I}_n$.
- iv) \mathbf{A} ist regulär, d.h. $\text{Rang } \mathbf{A} = n$.

BEWEIS: Wir zeigen, dass $(i) \implies (ii) \implies (iv) \implies (iii) \implies (i)$.

Wir nehmen also an, \mathbf{A} sei invertierbar, d.h. (i) gelte. Dann ist (ii) mit $\mathbf{X} = \mathbf{A}^{-1}$ erfüllt. Hieraus folgt andererseits, dass das Gleichungssystem $\mathbf{A}\mathbf{x} = \mathbf{b}$ für jedes \mathbf{b} die Lösung $\mathbf{x} := \mathbf{X}\mathbf{b}$ hat, denn nach dem Assoziativgesetz der Matrizen-Multiplikation gilt

$$\mathbf{A}\mathbf{x} = \mathbf{A}(\mathbf{X}\mathbf{b}) = (\mathbf{A}\mathbf{X})\mathbf{b} = \mathbf{I}_n \mathbf{b} = \mathbf{b}.$$

Nach Korollar 1.7 folgt, dass es für jedes \mathbf{b} genau eine Lösung gibt und dass $\text{Rang } \mathbf{A} = n$ ist, also (iv) gilt.

Bezeichnet \mathbf{e}_j wieder die j -te Kolonne von \mathbf{I}_n , so haben dann umgekehrt die Systeme $\mathbf{A}\mathbf{x} = \mathbf{e}_j$ ($j = 1, \dots, n$) alle genau eine Lösung, was bedeutet, dass $\mathbf{A}\mathbf{X} = \mathbf{I}_n$ eine eindeutige Lösung \mathbf{X} hat. Aus $\mathbf{A}\mathbf{X} = \mathbf{I}_n$ ergibt sich weiter

$$\begin{aligned} \mathbf{A}(\mathbf{X} + \mathbf{X}\mathbf{A} - \mathbf{I}_n) &= \mathbf{A}\mathbf{X} + \mathbf{A}(\mathbf{X}\mathbf{A}) - \mathbf{A}\mathbf{I}_n \\ &= \mathbf{I}_n + (\mathbf{A}\mathbf{X})\mathbf{A} - \mathbf{A} \\ &= \mathbf{I}_n + \mathbf{A} - \mathbf{A} \\ &= \mathbf{I}_n. \end{aligned}$$

Da \mathbf{X} die einzige Lösung von $\mathbf{A}\mathbf{X} = \mathbf{I}_n$ ist, folgt, dass $\mathbf{X} + \mathbf{X}\mathbf{A} - \mathbf{I}_n = \mathbf{X}$ ist, also auch $\mathbf{X}\mathbf{A} = \mathbf{I}_n$ gilt, d.h. \mathbf{X} ist die Inverse von \mathbf{A} . Es gilt somit (i). \blacksquare

Die Eigenschaften “regulär” und “invertierbar” sind also äquivalent, und man käme deshalb mit einer der zwei Bezeichnungen aus. Meist bevorzugt man die Bezeichnung regulär.

Als nächstes stellen wir Eigenschaften der Inversen zusammen.

Satz 2.18 Sind \mathbf{A} und \mathbf{B} reguläre $n \times n$ -Matrizen, so gilt:

i) \mathbf{A}^{-1} ist regulär und

$$\boxed{(\mathbf{A}^{-1})^{-1} = \mathbf{A}.} \quad (2.70)$$

ii) \mathbf{AB} ist regulär und

$$\boxed{(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}.} \quad (2.71)$$

iii) \mathbf{A}^T und \mathbf{A}^H sind regulär und

$$\boxed{(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T,} \quad \boxed{(\mathbf{A}^H)^{-1} = (\mathbf{A}^{-1})^H.} \quad (2.72)$$

BEWEIS: i) ergibt sich sofort aus der Definition (2.69).

Zu ii): Aus $(\mathbf{AB})(\mathbf{B}^{-1}\mathbf{A}^{-1}) = \mathbf{A}(\mathbf{BB}^{-1})\mathbf{A}^{-1} = \mathbf{AA}^{-1} = \mathbf{I}_n$ und Satz 2.17 folgt, dass $\mathbf{B}^{-1}\mathbf{A}^{-1}$ die Inverse von \mathbf{AB} ist.

Zu iii): Gemäss (2.35) ist $\mathbf{A}^T(\mathbf{A}^{-1})^T = (\mathbf{A}^{-1}\mathbf{A})^T = \mathbf{I}_n^T = \mathbf{I}_n$. Also ist wiederum nach Satz 2.17 $(\mathbf{A}^{-1})^T$ die Inverse von \mathbf{A}^T . Ganz analog ergibt sich $(\mathbf{A}^H)^{-1} = (\mathbf{A}^{-1})^H$. ■

Die Berechnung von \mathbf{A}^{-1} :

Wie wir im Beweis von Satz 2.17 gesehen haben, lassen sich die n Kolonnen von \mathbf{A}^{-1} berechnen als Lösungen der n Gleichungssysteme $\mathbf{Ax}_k = \mathbf{e}_k$, $k = 1, \dots, n$. Wir können deshalb die Inverse mit Gauss-Elimination bestimmen durch gleichzeitiges Lösen dieser n Systeme.

Mit Hilfe der Inversen der Koeffizientenmatrix, falls sie existiert, lässt sich ein lineares Gleichungssystem offensichtlich sofort formal lösen, indem man es von links mit \mathbf{A}^{-1} multipliziert. Aus Korollar 1.7 wissen wir ja schon, dass es genau eine Lösung gibt, wenn \mathbf{A} regulär ist. Genauer: Im Beweis von Satz 2.17 haben wir die eindeutige Lösbarkeit des Systems direkt mit der Invertierbarkeit und Regularität der Matrix verknüpft. Zusammen gibt das

Satz 2.19 Ist \mathbf{A} regulär, so hat das Gleichungssystem $\mathbf{Ax} = \mathbf{b}$ für jede rechte Seite \mathbf{b} eine eindeutige Lösung, und zwar ist

$$\boxed{\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}.} \quad (2.73)$$

BEMERKUNG: Da die Berechnung der Inversen \mathbf{A}^{-1} das Lösen von n linearen Gleichungssystemen mit der Matrix \mathbf{A} erfordert, macht es keinen Sinn, ein System $\mathbf{Ax} = \mathbf{b}$ auf diesem Weg zu lösen, denn der Aufwand ist deutlich grösser. Das gilt auch dann noch, wenn das System für mehrere rechte Seiten zu lösen ist. Wir werden in Abschnitt 3.1 den Rechenaufwand diverser Varianten diskutieren. ▼

2.8 Orthogonale und unitäre Matrizen

Orthogonale reelle und unitäre komplexe Matrizen sind spezielle reguläre quadratische Matrizen, die in vielen Anwendungen und Algorithmen vorkommen. Ihre Beliebtheit in der Numerik beruht unter anderem auf sehr guten Eigenschaften betreffend Rundung.

DEFINITION: Eine $n \times n$ -Matrix \mathbf{A} heisst **unitär** [*unitary*], falls $\mathbf{A}^H \mathbf{A} = \mathbf{I}_n$. Eine reelle unitäre Matrix heisst auch **orthogonal** [*orthogonal*]; für sie gilt $\mathbf{A}^T \mathbf{A} = \mathbf{I}_n$. ▲

Satz 2.20 Sind \mathbf{A} und \mathbf{B} unitäre (bzw. orthogonale) $n \times n$ -Matrizen, so gilt:

- i) \mathbf{A} ist regulär und $\mathbf{A}^{-1} = \mathbf{A}^H$ (bzw. $\mathbf{A}^{-1} = \mathbf{A}^T$).
- ii) $\mathbf{A} \mathbf{A}^H = \mathbf{I}_n$ (bzw. $\mathbf{A} \mathbf{A}^T = \mathbf{I}_n$).
- iii) \mathbf{A}^{-1} ist unitär (orthogonal).
- iv) $\mathbf{A} \mathbf{B}$ ist unitär (orthogonal).

BEWEIS: Zu i)–iii): Aus $\mathbf{A}^H \mathbf{A} = \mathbf{I}_n$ folgt gemäss Satz 2.17 (mit $\mathbf{A} := \mathbf{A}^H, \mathbf{X} := \mathbf{A}$), dass \mathbf{A}^H invertierbar ist und $(\mathbf{A}^H)^{-1} = \mathbf{A}$ gilt. Wegen der Symmetrie in der Definition (2.69) der Inversen ist damit auch \mathbf{A} invertierbar und $\mathbf{A}^{-1} = \mathbf{A}^H$. Daraus folgt erstens nach Satz 2.17, dass \mathbf{A} regulär ist. Zweitens ist $\mathbf{A} \mathbf{A}^H = \mathbf{A} \mathbf{A}^{-1} = \mathbf{I}_n$. Drittens ergibt sich mit Hilfe von (2.32), dass $(\mathbf{A}^{-1})^H \mathbf{A}^{-1} = (\mathbf{A}^H)^H \mathbf{A}^{-1} = \mathbf{A} \mathbf{A}^{-1} = \mathbf{I}_n$; also ist \mathbf{A}^{-1} unitär.

Zu iv): Dies folgt aus $(\mathbf{A} \mathbf{B})^H \mathbf{A} \mathbf{B} = \mathbf{B}^H (\mathbf{A}^H \mathbf{A}) \mathbf{B} = \mathbf{B}^H \mathbf{B} = \mathbf{I}_n$.

Die Aussagen über orthogonale Matrizen folgen als Spezialfall. ■

BEISPIEL 2.26: **Givens–Rotationen**¹³ [*Givens rotations, plane rotations*] sind einfache orthogonale Matrizen, welche eine Drehung um einen Winkel $-\phi$ in einer durch zwei Koordinatenachsen aufgespannten Ebene beschreiben; die anderen Koordinatenachsen bleiben fest. Im \mathbb{R}^3 bedeutet dies, dass genau eine Koordinatenachse fest bleibt. Das ist dann die Drehachse.

Als Beispiel betrachten wir eine Drehung im \mathbb{R}^5 in der durch die erste und die dritte Achse aufgespannten Ebene:

$$\mathbf{U}_{13}(\phi) := \begin{pmatrix} \cos \phi & 0 & \sin \phi & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ -\sin \phi & 0 & \cos \phi & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

Es ist $\mathbf{U}_{13}^T(\phi) = \mathbf{U}_{13}(-\phi)$, und für $\mathbf{U}_{13}^T(\phi) \mathbf{U}_{13}(\phi)$ ergibt sich

¹³J. WALLACE GIVENS (1910 – 1993), amerikanischer Numeriker, Direktor der Applied Mathematics Division am Argonne National Laboratory.

$$\begin{pmatrix} \cos \phi & 0 & -\sin \phi & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \sin \phi & 0 & \cos \phi & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \phi & 0 & \sin \phi & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ -\sin \phi & 0 & \cos \phi & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \mathbf{I}_5.$$

Givens–Rotationen werden in diversen Algorithmen eingesetzt als Bausteine für kompliziertere orthogonale Transformationen. Manchmal werden sie auch **Jacobi–Rotationen**¹⁴ [*Jacobi rotations*] genannt. \blacklozenge

BEISPIEL 2.27: Householder-Matrizen¹⁵ [*Householder matrices*] (oder: **Householder-Spiegelungen** [*Householder reflections*]) bilden eine weitere spezielle Klasse orthogonaler Matrizen. Sie beschreiben die Spiegelung an einer Hyperebene (Definition später), d.h. an einer Geraden, falls $n = 2$, und an einer Ebene, wenn $n = 3$.

Es sei \mathbf{u} ein reeller Kolonnenvektor der Länge 1, d.h. $\mathbf{u}^\top \mathbf{u} = 1$. Dann ist das äussere Produkt $\mathbf{u}\mathbf{u}^\top$ ja die $n \times n$ -Matrix mit $(\mathbf{u}\mathbf{u}^\top)_{ij} = u_i u_j$. Die zu \mathbf{u} gehörende Householder-Matrix ist

$$\mathbf{Q}_{\mathbf{u}} := \mathbf{I} - 2\mathbf{u}\mathbf{u}^\top. \quad (2.74)$$

Sie lässt sich offensichtlich durch die Projektion (2.62) auf die Richtung \mathbf{u} ausdrücken (wobei jetzt \mathbf{u} reell und $\|\mathbf{u}\| = 1$ ist):

$$\mathbf{Q}_{\mathbf{u}} = \mathbf{I} - 2\mathbf{P}_{\mathbf{u}}, \quad \text{wo } \mathbf{P}_{\mathbf{u}} := \mathbf{u}\mathbf{u}^\top. \quad (2.75)$$

$\mathbf{P}_{\mathbf{u}}$ und damit auch $\mathbf{Q}_{\mathbf{u}}$ ist symmetrisch. Wir zeigen nun, dass $\mathbf{Q}_{\mathbf{u}}$ auch orthogonal ist. Wegen $\mathbf{P}_{\mathbf{u}}^2 = \mathbf{u}(\mathbf{u}^\top \mathbf{u})\mathbf{u}^\top = \mathbf{u}\mathbf{u}^\top = \mathbf{P}_{\mathbf{u}}$, ist

$$\mathbf{Q}_{\mathbf{u}}^\top \mathbf{Q}_{\mathbf{u}} = \mathbf{Q}_{\mathbf{u}}^2 = (\mathbf{I} - 2\mathbf{P}_{\mathbf{u}})^2 = \mathbf{I} - 4\mathbf{P}_{\mathbf{u}} + 4\mathbf{P}_{\mathbf{u}}^2 = \mathbf{I}.$$

Hier noch ein Beispiel einer 4×4 -Householder-Matrix:

Mit $\mathbf{u} = \begin{pmatrix} 4/5 & -2/5 & 1/5 & -2/5 \end{pmatrix}^\top$ erhalten wir $\mathbf{u}^\top \mathbf{u} = 1$ und

$$\mathbf{P}_{\mathbf{u}} = \mathbf{u}\mathbf{u}^\top = \frac{1}{25} \begin{pmatrix} 16 & -8 & 4 & -8 \\ -8 & 4 & -2 & 4 \\ 4 & -2 & 1 & -2 \\ -8 & 4 & -2 & 4 \end{pmatrix},$$

$$\mathbf{Q}_{\mathbf{u}} = \mathbf{I} - 2\mathbf{P}_{\mathbf{u}} = \frac{1}{25} \begin{pmatrix} -7 & 16 & -8 & 16 \\ 16 & 17 & 4 & -8 \\ -8 & 4 & 23 & 4 \\ 16 & -8 & 4 & 17 \end{pmatrix}. \quad \blacklozenge$$

BEISPIEL 2.28: Permutationsmatrizen [*permutation matrices*] sind quadratische Matrizen, die in jeder Kolonne und in jeder Zeile genau eine Eins und sonst lauter Nullen haben. Zum Beispiel ist

¹⁴CARL GUSTAV JACOBI (10.12.1804 – 18.2.1851), deutscher Mathematiker, Professor in Königsberg und Berlin.

¹⁵ALSTON SCOTT HOUSEHOLDER (5.5.1904 – 4.7.1993), amerikanischer Mathematiker, 1946 – 1969 am Oak Ridge National Laboratory, ab 1948 Direktor der Mathematics Division, Vorreiter der mathematischen Biologie und der numerischen linearen Algebra.

$$\mathbf{P} := \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

eine Permutationsmatrix der Ordnung $n = 5$. Man verifiziert leicht, dass jede Permutationsmatrix orthogonal ist.

Wendet man \mathbf{P} von links auf eine Matrix \mathbf{A} an, so enthält das Produkt \mathbf{PA} permutierte Zeilen von \mathbf{A} . Wendet man \mathbf{P} von rechts auf \mathbf{A} an, so besteht das Produkt \mathbf{AP} aus permutierten Kolonnen von \mathbf{A} . \blacklozenge

Die durch orthogonale (reelle) und unitäre (komplexe) Matrizen definierten linearen Abbildungen haben besondere geometrische Eigenschaften:

Satz 2.21 Die durch eine orthogonale oder unitäre $n \times n$ Matrix \mathbf{A} definierte Abbildung ist **längentreu** [length preserving] (oder: **isometrisch** [isometric]) und **winkeltreu** [angle preserving], d.h. es ist für alle $\mathbf{x}, \mathbf{y} \in \mathbb{E}^n$

$$\|\mathbf{Ax}\| = \|\mathbf{x}\|, \quad \langle \mathbf{Ax}, \mathbf{Ay} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle. \quad (2.76)$$

Da gemäss (2.56)–(2.57) der Winkel zwischen zwei Vektoren mittels Skalarprodukt und Norm definiert ist, folgt aus (2.76) in der Tat, dass bei der Abbildung \mathbf{A} der Winkel zwischen zwei Vektoren invariant (unverändert) bleibt, d.h. gleich jenem zwischen den Bildvektoren ist.

BEWEIS von Satz 2.21: Nach Voraussetzung ist $\mathbf{A}^H \mathbf{A} = \mathbf{I}$, also

$$\langle \mathbf{Ax}, \mathbf{Ay} \rangle = (\mathbf{Ax})^H (\mathbf{Ay}) = \mathbf{x}^H \mathbf{A}^H \mathbf{Ay} = \mathbf{x}^H \mathbf{y} = \langle \mathbf{x}, \mathbf{y} \rangle.$$

Wählt man $\mathbf{y} := \mathbf{x}$, so folgt

$$\|\mathbf{Ax}\|^2 = \langle \mathbf{Ax}, \mathbf{Ax} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|^2,$$

also $\|\mathbf{Ax}\| = \|\mathbf{x}\|$. \blacksquare

2.9 Strukturierte Matrizen

Wir haben bereits mehrere Typen von Matrizen mit spezieller Struktur angetroffen, etwa die symmetrischen Matrizen, die Dreiecksmatrizen und die Rang-1-Matrizen, welche äussere Produkte von zwei Vektoren sind. In der Matrixtheorie und den Anwendungen kommen noch viele weitere Matrizen mit speziellen Strukturen vor, von denen wir hier ein paar vorstellen wollen.

Eine Matrix \mathbf{B} ist eine **untere Bidiagonalmatrix** [lower bidiagonal matrix] bzw. eine **obere Bidiagonalmatrix** [upper bidiagonal matrix], falls $(\mathbf{B})_{ij} = 0$ für $i < j$ und $i > j + 1$ bzw. für $j < i$ und $j > i + 1$.

BEISPIEL 2.29:

$$\mathbf{B} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 3 & 4 & 0 & 0 \\ 0 & 2 & 6 & 0 \\ 0 & 0 & 5 & 3 \end{pmatrix}$$

ist eine untere Bidiagonalmatrix der Ordnung 4. \blacklozenge

Eine quadratische Matrix \mathbf{T} heisst **tridiagonal** [*tridiagonal*], d.h. ist eine **Tridiagonalmatrix** [*tridiagonal matrix*], falls $(\mathbf{T})_{ij} = 0$ für $i > j + 1$ und $j > i + 1$.

BEISPIEL 2.30:

$$\mathbf{T} = \begin{pmatrix} 4 & 2 & 0 & 0 \\ 3 & 5 & 4 & 0 \\ 0 & 8 & 6 & 7 \\ 0 & 0 & 9 & 0 \end{pmatrix}$$

ist tridiagonal. Natürlich dürfen einzelne Elemente innerhalb des Bandes null sein. \blacklozenge

Eine $m \times n$ Matrix \mathbf{B} ist eine **Bandmatrix** [*banded matrix*] mit **unterer Bandbreite** [*lower bandwidth*] p und **oberer Bandbreite** [*upper bandwidth*] q , wenn

$$(\mathbf{B})_{ij} = 0 \quad \text{falls} \quad i > j + p \quad \text{oder} \quad j > i + q.$$

Die **(gesamte)Bandbreite** [*(total) bandwidth*] ist $p + q + 1$.

BEISPIEL 2.31:

$$\mathbf{B} = \begin{pmatrix} 9 & 6 & 3 & 0 & 0 \\ 7 & 8 & 5 & 2 & 0 \\ 0 & 5 & 7 & 4 & 1 \\ 0 & 0 & 3 & 6 & 3 \\ 0 & 0 & 0 & 1 & 5 \end{pmatrix}$$

hat untere Bandbreite 1, obere Bandbreite 2 und gesamte Bandbreite 4. (Die Diagonale wird also nur bei der gesamten Bandbreite mitgezählt.) \blacklozenge

Tridiagonalmatrizen sind natürlich spezielle Bandmatrizen mit oberer und unterer Bandbreite eins. Bei Bidiagonalmatrizen ist eine der Bandbreiten eins, die andere null.

Eine wichtige Klasse einseitiger Bandmatrizen sind die **Hessenberg-Matrizen**¹⁶ [*Hessenberg matrices*], die untere Bandbreite 1 haben.

BEISPIEL 2.32:

$$\mathbf{H} = \begin{pmatrix} 9 & 6 & 3 & 2 & 1 \\ 7 & 8 & 5 & 2 & 1 \\ 0 & 5 & 7 & 4 & 1 \\ 0 & 0 & 3 & 6 & 3 \\ 0 & 0 & 0 & 1 & 5 \end{pmatrix}$$

ist eine Hessenberg-Matrix. \blacklozenge

¹⁶GERHARD HESSENBERG (16.8.1874 – 16.11.1925), deutscher Mathematiker, Professor in Breslau und Tübingen.

Wir werden später sehen, dass Matrizen lineare Abbildungen darstellen bezüglich einem oder zwei fest gewählten Koordinatensystemen. Durch geeignete Wahl der Koordinaten kann man eine Matrix oft durch eine “einfachere” ersetzen. Hessenberg-, Tridiagonal- und Bidiagonalmatrizen spielen dabei eine wichtige Rolle.

In vielen Anwendungen, vor allem beim numerischen Lösen von partiellen Differentialgleichungen, treten Matrizen auf, die sehr gross sind, aber nur sehr wenige von Null verschiedene Elemente haben, welche nicht auf ein schmales Band um die Diagonale beschränkt sind. Man nennt diese Matrizen **dünn besetzt** [*sparse*].

In anderen Anwendungen kommen ganz speziell strukturierte Matrizen vor. Die folgenden zwei Typen sind weitverbreitet.

Die Matrix $\mathbf{A} = (a_{ij})$ ist eine **Toeplitz-Matrix**¹⁷ [*Toeplitz matrix*], falls a_{ij} nur von der Differenz $i - j$ abhängt, und es ist eine **Hankel-Matrix**¹⁸ [*Hankel matrix*], falls a_{ij} nur von der Summe $i + j$ abhängt.

BEISPIELE 2.33: Hier sind Beispiele einer Toeplitz-Matrix \mathbf{T} und einer Hankel-Matrix \mathbf{H} :

$$\mathbf{T} = \begin{pmatrix} 9 & 7 & 5 & 3 & 1 \\ 8 & 9 & 7 & 5 & 3 \\ 7 & 8 & 9 & 7 & 5 \\ 6 & 7 & 8 & 9 & 7 \\ 5 & 6 & 7 & 8 & 9 \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} 9 & 8 & 7 & 6 & 5 \\ 8 & 7 & 6 & 5 & 4 \\ 7 & 6 & 5 & 4 & 3 \\ 6 & 5 & 4 & 3 & 2 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix}. \quad \blacklozenge$$

Jede Matrix lässt sich durch horizontale und/oder vertikale Trennlinien in Blöcke aufteilen, die selbst Matrizen sind. Wir bezeichnen die Matrix dann als **Blockmatrix** [*block matrix*]. Wir nennen die Blöcke von \mathbf{A} zum Beispiel $\mathbf{A}_{k\ell}$. Sie brauchen nicht von derselben Grösse zu sein, aber für festes k sind die (nebeneinanderliegenden) Blöcke $\mathbf{A}_{k\ell}$ alle gleich hoch, und für festes ℓ sind die (übereinanderliegenden) Blöcke $\mathbf{A}_{k\ell}$ alle gleich breit.

BEISPIEL 2.34:

$$\mathbf{A} = \left(\begin{array}{cc|cc} \mathbf{A}_{11} & \mathbf{A}_{12} & & \\ \mathbf{A}_{21} & \mathbf{A}_{22} & & \\ \mathbf{A}_{31} & \mathbf{A}_{32} & & \end{array} \right) = \left(\begin{array}{cc|cc} 9 & 8 & 7 & 6 & 5 \\ 7 & 9 & 5 & 7 & 6 \\ \hline 7 & 6 & 5 & 4 & 3 \\ 5 & 7 & 3 & 5 & 4 \\ 3 & 5 & 1 & 3 & 5 \\ \hline 1 & 2 & 3 & 4 & 5 \end{array} \right)$$

ist eine Blockmatrix mit der speziellen Eigenschaft, dass alle Blöcke Toeplitz-Struktur haben. \blacklozenge

¹⁷OTTO TOEPLITZ (1.8.1881 – 15.2.1940), deutscher Mathematiker, Professor in Kiel (1913–1927) und Bonn (1927–1933); 1938 Emigration nach Palästina.

¹⁸HERMANN HANKEL (14.2.1839 – 29.8.1873), deutscher Mathematiker, Professor in Erlangen und Tübingen.

Die Blöcke \mathbf{C}_{kj} des Produktes \mathbf{C} von zwei Blockmatrizen $\mathbf{A} = (\mathbf{A}_{k\ell})$ und $\mathbf{B} = (\mathbf{B}_{\ell j})$ lassen sich nach der Formel

$$\mathbf{C}_{kj} = \sum_{\ell} \mathbf{A}_{k\ell} \mathbf{B}_{\ell j} \quad (2.77)$$

berechnen, vorausgesetzt dass die ℓ -te Blockkolonne von \mathbf{A} so breit ist wie die ℓ -te Blockzeile von \mathbf{B} hoch ist.

BEISPIEL 2.35: Das Produkt \mathbf{C} der Matrix \mathbf{A} aus Beispiel 2.34 mit

$$\mathbf{B} = \left(\begin{array}{c|c|c} \mathbf{B}_{11} & \mathbf{B}_{12} & \mathbf{B}_{13} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{B}_{23} \end{array} \right) = \left(\begin{array}{cc|ccc|c} 1 & -1 & 2 & -2 & 2 & -3 \\ -1 & 1 & -2 & 2 & -2 & 3 \\ \hline 4 & -4 & 5 & -5 & 5 & -6 \\ -4 & 4 & -5 & 5 & -5 & 6 \\ \hline 4 & -4 & 5 & -5 & 5 & -6 \end{array} \right)$$

kann man nach Formel (2.77) bestimmen. Zum Beispiel wird

$$\begin{aligned} \mathbf{C}_{11} &= \begin{pmatrix} 9 & 8 \\ 7 & 9 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \begin{pmatrix} 7 & 6 & 5 \\ 5 & 7 & 6 \end{pmatrix} \begin{pmatrix} 4 & -4 \\ -4 & 4 \\ 4 & -4 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -1 \\ -2 & 2 \end{pmatrix} + \begin{pmatrix} 24 & -24 \\ 16 & -16 \end{pmatrix} \\ &= \begin{pmatrix} 25 & -25 \\ 14 & -14 \end{pmatrix}. \quad \blacklozenge \end{aligned}$$

Blockmatrizen spielen bei der Implementation der Matrix-Operationen auf Vektor- und Parallelrechnern (und mittlerweile sogar auf PCs) eine wichtige Rolle. Die Matrizen werden derart in Blöcke aufgeteilt, dass die Blockgrösse mit der bestehenden Hardware, insbesondere dem Cache und den Vektorregistern optimal abgestimmt ist. Dadurch lässt sich das zeitaufwendige Laden von Daten aus dem Speicher minimieren.

Für FORTRAN und C Programme gibt es spezielle, von den Prozessor-Herstellern unterstützte Subroutine-Libraries für die grundlegenden Vektor- und Matrix-Operationen, die **BLAS** (oder, ausgeschrieben, Basic Linear Algebra Subroutines). BLAS1 enthält die Vektor-Operationen, BLAS2 die Matrix-Vektor-Operationen und BLAS3 die Matrix-Matrix-Operationen. Sie sind ihrerseits in **LA-PACK** integriert, einer optimierten Library von Subroutinen für die numerische lineare Algebra (ohne Berücksichtigung dünn besetzter Matrizen).

Kapitel 3

Die LR-Zerlegung

3.1 Die Gauss-Elimination als LR-Zerlegung

Wir kommen zurück auf die Gauss-Elimination und werden hier die Reduktion einer Matrix auf Dreiecks- oder Zeilenstufenform als Matrixzerlegung interpretieren. Dazu modifizieren wir unsere Bezeichnungen aus Kapitel 1: Die resultierende Zeilenstufenmatrix wird neu mit $\mathbf{R} = (r_{kj})$ bezeichnet und die dazugehörige rechte Seite als $\mathbf{c} = (c_1 \ \dots \ c_n)^\top$. Diese neuen Bezeichnungen werden in den Formeln berücksichtigt, sobald die entsprechenden Elemente definiert worden sind, also sobald die Pivotzeile gewählt worden ist.

Wir betrachten zuerst in diesem Abschnitt den Fall eines Systems $\mathbf{Ax} = \mathbf{b}$ mit regulärer quadratischer Matrix, in dem \mathbf{R} ja eine reguläre $n \times n$ -Rechtsdreiecksmatrix ist:

$$\mathbf{R} = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{nn} \end{pmatrix}.$$

Wir nehmen anfänglich an, dass keine Zeilenvertauschungen nötig sind. Für den j -ten Schritt lauten die Umbenennungen und die relevanten Formeln (1.10)–(1.11b) in den neuen Bezeichnungen:

$$r_{ji} := a_{ji}^{(j-1)} \quad (i = j, \dots, n), \quad (3.1)$$

$$c_j := b_j^{(j-1)}, \quad (3.2)$$

$$l_{kj} := a_{kj}^{(j-1)} / r_{jj} \quad (k = j + 1, \dots, n), \quad (3.3)$$

$$a_{ki}^{(j)} := a_{ki}^{(j-1)} - l_{kj} r_{ji}, \quad (i = j + 1, \dots, n, \\ k = j + 1, \dots, n), \quad (3.4)$$

$$b_k^{(j)} := b_k^{(j-1)} - l_{kj} c_j \quad (k = j + 1, \dots, n). \quad (3.5)$$

Dabei sind die Koeffizienten l_{kj} ($k > j$) ja die Multiplikatoren, mit denen die Pivotzeilen multipliziert werden. Wir setzen zusätzlich

$$l_{kj} := 0 \quad (k < j), \quad l_{jj} := 1, \quad (3.6)$$

so dass $\mathbf{L} := (l_{kj})$ eine $n \times n$ -Linksdreiecksmatrix ist:

$$\mathbf{L} = \begin{pmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{pmatrix}.$$

Falls $k \leq i$ ist, folgt aus (3.4) durch rekursive Anwendung mit $j = 1, 2, \dots, k-1$

$$\begin{aligned} a_{ki} &= a_{ki}^{(0)} = l_{k1} r_{1i} + a_{ki}^{(1)} = \dots \\ &= l_{k1} r_{1i} + l_{k2} r_{2i} + \dots + l_{k,k-1} r_{k-1,i} + a_{ki}^{(k-1)}. \end{aligned}$$

Ersetzen wir das letzte Glied gemäss (3.1) und (3.6) durch $l_{kk} r_{ki}$, so erhalten wir

$$a_{ki} = l_{k1} r_{1i} + \dots + l_{k,k-1} r_{k-1,i} + l_{kk} r_{ki} = \sum_{j=1}^k l_{kj} r_{ji}. \quad (3.7)$$

Analog ergibt sich im Falle $k > i$ aus (3.4) mit $j = 1, 2, \dots, i-1$

$$\begin{aligned} a_{ki} &= a_{ki}^{(0)} = l_{k1} r_{1i} + a_{ki}^{(1)} = \dots \\ &= l_{k1} r_{1i} + l_{k2} r_{2i} + \dots + l_{k,i-1} r_{i-1,i} + a_{ki}^{(i-1)}. \end{aligned}$$

Hier ersetzen wir das letzte Glied gemäss (3.3), so dass

$$a_{ki} = l_{k1} r_{1i} + \dots + l_{k,i-1} r_{i-1,i} + l_{ki} r_{ii} = \sum_{j=1}^i l_{kj} r_{ji}. \quad (3.8)$$

Wegen der Dreiecksgestalt der Matrizen \mathbf{L} und \mathbf{R} lassen sich die zwei Formeln (3.7) und (3.8) zusammenfassen zu

$$a_{ki} = \sum_{j=1}^n l_{kj} r_{ji},$$

das heisst es gilt

$$\boxed{\mathbf{A} = \mathbf{LR}.} \quad (3.9)$$

Durch die Gauss-Elimination wird also die Koeffizientenmatrix \mathbf{A} eines linearen Gleichungssystems implizit in ein Produkt einer Linksdreiecksmatrix \mathbf{L} und einer Rechtsdreiecksmatrix \mathbf{R} zerlegt. Man nennt deshalb diesen Hauptteil der Gauss-Elimination auch **LR-Zerlegung** oder, in Anlehnung an die englische Terminologie, **LU-Zerlegung** [*LU decomposition*], weil in der englischen Literatur unsere Matrix \mathbf{R} oft mit \mathbf{U} bezeichnet wird.

Die Wirkung der Gauss-Elimination auf die Konstantenkolonne lässt sich analog deuten: Statt (3.7) gilt

$$b_k = l_{k1} c_1 + \dots + l_{k,k-1} c_{k-1} + l_{kk} c_k = \sum_{j=1}^k l_{kj} c_j, \quad (3.10)$$

also

$$\boxed{\mathbf{Lc} = \mathbf{b}.} \quad (3.11)$$

Schliesslich gilt nach (1.12) für das **Rückwärtseinsetzen** [*back substitution*]:

$$\boxed{x_k := \left(c_k - \sum_{i=k+1}^n r_{ki} x_i \right) \frac{1}{r_{kk}}} \quad (k = n, n-1, \dots, 1). \quad (3.12)$$

Wenn wir dies nach c_k auflösen, erhalten wir

$$c_k = \sum_{i=k}^n r_{ki} x_i \quad (k = n, n-1, \dots, 1), \quad (3.13)$$

also, weil \mathbf{R} eine Rechtsdreieckmatrix ist,

$$\boxed{\mathbf{R}\mathbf{x} = \mathbf{c}.} \quad (3.14)$$

So gut man dieses System mit Rückwärtseinsetzen rekursiv lösen kann, so gut kann man auch das System (3.11) rekursiv lösen durch **Vorwärtseinsetzen** [*forward substitution*]:

$$\boxed{c_k := \left(b_k - \sum_{j=1}^{k-1} l_{kj} c_j \right) \frac{1}{l_{kk}}} \quad (k = 1, 2, \dots, n), \quad (3.15)$$

wobei hier ja der Bruch wegfällt, weil nach Definition $l_{kk} = 1$ ist.

Mit anderen Worten, wir brauchen die Konstantenkolonne gar nicht im Eliminationsprozess mitzuführen, wir können die rechte Seite \mathbf{c} des reduzierten Systems gemäss (3.15) bestimmen. In Bezug auf die Zahl der Rechenoperationen und die Grösse der Rundungsfehler ändert das allerdings nichts, doch kann die separate Behandlung von Vorteil sein, sowohl aus Gründen der Programmstruktur als auch, weil die rechte Seite \mathbf{b} manchmal während der LR-Zerlegung von \mathbf{A} noch gar nicht bekannt ist.

BEISPIEL 3.1: In unserem Beispiel 1.1 auf Seite 1-3 ist

$$\mathbf{A} = \begin{pmatrix} 2 & -2 & 4 \\ -5 & 6 & -7 \\ 3 & 2 & 1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 6 \\ -7 \\ 9 \end{pmatrix}. \quad (3.16)$$

Die Gauss-Elimination liefert, wie man (1.6)–(1.8) entnehmen kann,

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{5}{2} & 1 & 0 \\ \frac{3}{2} & 5 & 1 \end{pmatrix}, \quad \mathbf{R} = \begin{pmatrix} 2 & -2 & 4 \\ 0 & 1 & 3 \\ 0 & 0 & -20 \end{pmatrix}. \quad (3.17)$$

$$\mathbf{c} = \begin{pmatrix} 6 \\ 8 \\ -40 \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}. \quad (3.18)$$

Man überprüft leicht, dass effektiv $\mathbf{A} = \mathbf{LR}$, $\mathbf{Lc} = \mathbf{b}$ und $\mathbf{Rx} = \mathbf{c}$ gilt. \blacklozenge

Allerdings lässt sich ja auch bei regulärer quadratischer Matrix \mathbf{A} ein Gleichungssystem im allgemeinen nicht ohne Zeilenvertauschungen lösen. Bekanntlich hängen diese vom Verlauf der Rechnung ab: Nullen oder (in gewissem Sinne) relativ kleine Zahlen dürfen nicht als Pivot gewählt werden. Wenn wir die Vertauschungen alle am Anfang der Elimination kennen würden, könnten wir sie aber vorgängig ausführen und dann im so präparierten Eliminationschema alle Pivot auf der Diagonalen wählen. Formelmässig kann man solche vorgängige Vertauschungen ausdrücken durch die Multiplikation von links mit einer Permutationsmatrix \mathbf{P} .

BEISPIEL 3.2: Man hätte in Beispiel 3.1 die Zeilen von \mathbf{A} etwa wie folgt vertauschen können:

$$\underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}}_{\mathbf{P}} \underbrace{\begin{pmatrix} 2 & -2 & 4 \\ -5 & 6 & -7 \\ 3 & 2 & 1 \end{pmatrix}}_{\mathbf{A}} = \underbrace{\begin{pmatrix} -5 & 6 & -7 \\ 3 & 2 & 1 \\ 2 & -2 & 4 \end{pmatrix}}_{\mathbf{PA}}. \quad (3.19)$$

Natürlich müsste man dann auch die Komponenten der rechten Seite entsprechend vertauschen:

$$\underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}}_{\mathbf{P}} \underbrace{\begin{pmatrix} 6 \\ -7 \\ 9 \end{pmatrix}}_{\mathbf{b}} = \underbrace{\begin{pmatrix} -7 \\ 9 \\ 6 \end{pmatrix}}_{\mathbf{Pb}}. \quad (3.20)$$

Die LR-Zerlegung $\mathbf{PA} = \tilde{\mathbf{L}}\tilde{\mathbf{R}}$ mit diagonalen Pivotwahl erweist sich als möglich und liefert

$$\tilde{\mathbf{L}} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{3}{5} & 1 & 0 \\ -\frac{2}{5} & \frac{1}{14} & 1 \end{pmatrix}, \quad \tilde{\mathbf{R}} = \begin{pmatrix} -5 & 6 & -7 \\ 0 & \frac{28}{5} & -\frac{16}{5} \\ 0 & 0 & \frac{10}{7} \end{pmatrix}, \quad (3.21)$$

wobei nun also $\mathbf{PA} = \tilde{\mathbf{L}}\tilde{\mathbf{R}}$ gilt. Vor- und Rückwärtseinsetzen ergibt

$$\begin{aligned} \tilde{\mathbf{c}} &= \tilde{\mathbf{L}}^{-1}(\mathbf{Pb}) = \left(-7 \quad \frac{24}{5} \quad \frac{20}{7} \right)^{\top}, \\ \mathbf{x} &= \tilde{\mathbf{R}}^{-1}\tilde{\mathbf{c}} = \left(1 \quad 2 \quad 2 \right)^{\top}. \end{aligned} \quad (3.22)$$

Die in der LR-Zerlegung erzeugten Matrizen und auch die reduzierte rechte Seite hängen natürlich von der Wahl der Pivotelemente (d.h. von den ausgeführten Zeilenvertauschungen) ab. Erst das Schlussresultat für \mathbf{x} ist wieder identisch. Wir sehen hier, dass wir durch die neue Pivotwahl das grosse Element -20 in \mathbf{R} (vgl. (3.17)) vermeiden konnten. \blacklozenge

Zusammenfassend erhalten wir aus den vorhergehenden Überlegungen die folgende Aussage:

Satz 3.1 *Angewandt auf ein quadratisches Gleichungssystem $\mathbf{Ax} = \mathbf{b}$ mit regulärer Matrix \mathbf{A} liefert die Gauss-Elimination eine die Zeilenvertauschungen beschreibende Permutationsmatrix \mathbf{P} , eine Rechtsdreiecksmatrix \mathbf{R} des reduzierten Systems und eine entsprechende rechte Seite \mathbf{c} sowie, durch Zusammenziehen der Zeilenmultiplikatoren l_{kj} , eine Linksdreiecksmatrix \mathbf{L} , wobei alles quadratische Matrizen gleicher Ordnung sind und die Beziehungen*

$$\boxed{\mathbf{PA} = \mathbf{LR}, \quad \mathbf{Lc} = \mathbf{Pb}, \quad \mathbf{Rx} = \mathbf{c}.} \quad (3.23)$$

gelten.

Ist die LR-Zerlegung $\mathbf{PA} = \mathbf{LR}$ einmal berechnet, so lässt sich irgend ein System mit Koeffizientenmatrix \mathbf{A} lösen durch Auflösen von $\mathbf{Lc} = \mathbf{Pb}$ nach \mathbf{c} (Vorwärtseinsetzen) und Auflösen von $\mathbf{Rx} = \mathbf{c}$ nach \mathbf{x} (Rückwärtseinsetzen).

BEMERKUNG: Dass die Beziehungen (3.23) die richtige Lösung \mathbf{x} liefern, sieht man auch daraus, dass

$$\mathbf{Ax} = \mathbf{P}^{-1}\mathbf{PAx} = \mathbf{P}^{-1}\mathbf{LRx} = \mathbf{P}^{-1}\mathbf{Lc} = \mathbf{P}^{-1}\mathbf{Pb} = \mathbf{b}. \quad \blacktriangledown$$

Es gibt viele Varianten der LR-Zerlegung. Wir haben in jedem Eliminationsschritt eine neue Kolonne von \mathbf{L} , eine neue Zeile von \mathbf{R} und eine neue Komponente von \mathbf{c} bestimmt und zudem das verkleinerte Restgleichungssystem nachgeführt. Man kann aber auch \mathbf{L} und \mathbf{R} direkt aus \mathbf{A} berechnen, ohne dass man Restgleichungssysteme aufstellt. Zum Beispiel lassen sich beide Matrizen \mathbf{L} und \mathbf{R} direkt kolonnenweise (oder zeilenweise) berechnen. In der Tat bekommen wir, wenn wir wieder von Zeilenvertauschungen absehen, aus (3.7)–(3.8) sofort die nachfolgenden Rekursionen:

Algorithmus 3.1 (Zeilenweise, direkte LR-Zerlegung, “regulärer Fall”) *Zur LR-Zerlegung einer regulären Matrix \mathbf{A} berechne man für $i = 1, \dots, n$, mit $l_{ii} := 1$,*

$$l_{ik} := \left(a_{ik} - \sum_{j=1}^{k-1} l_{ij} r_{jk} \right) \frac{1}{r_{kk}} \quad (k = 1, \dots, i-1), \quad (3.24a)$$

$$r_{ik} := \left(a_{ik} - \sum_{j=1}^{i-1} l_{ij} r_{jk} \right) \frac{1}{l_{ii}} \quad (k = i, \dots, n). \quad (3.24b)$$

Zeilenvertauschungen können dabei wie üblich durch Umnummerierung der Elemente von \mathbf{A} berücksichtigt werden.

Will man statt den Diagonalelementen von \mathbf{L} jene von \mathbf{R} auf 1 normieren, so wendet man (3.24a) auch für $k = i$ an, dafür (3.24b) erst ab $k = i + 1$.

BEISPIEL 3.3: Um etwa die Matrix \mathbf{A} aus (3.16) in Beispiel 3.1 zu zerlegen in die Faktoren \mathbf{L} und \mathbf{R} in (3.17) kann man nach (3.24a)–(3.24b) zeilenweise wie folgt vorgehen (wir wählen dabei $l_{ii} := 1$):

$$\begin{aligned}
r_{11} &:= a_{11} = 2, & r_{12} &:= a_{12} = -2, & r_{13} &:= a_{13} = 4, \\
l_{21} &:= a_{21} \frac{1}{r_{11}} = -5 \cdot \frac{1}{2} = -\frac{5}{2}, \\
r_{22} &:= a_{22} - l_{21} r_{12} = 6 - \left(-\frac{5}{2}\right) \cdot (-2) = 1, \\
r_{23} &:= a_{23} - l_{21} r_{13} = -7 - \left(-\frac{5}{2}\right) \cdot 4 = 3, \\
l_{31} &:= a_{31} \frac{1}{r_{11}} = 3 \cdot \frac{1}{2} = \frac{3}{2}, \\
l_{32} &:= \left(a_{32} - l_{31} r_{12}\right) \frac{1}{r_{22}} = \left(2 - \frac{3}{2} \cdot (-2)\right) \frac{1}{1} = 5, \\
r_{33} &:= a_{33} - l_{31} r_{13} - l_{32} r_{23} = 1 - \frac{3}{2} \cdot 4 - 5 \cdot 3 = -20.
\end{aligned}$$

Genau so gut kann man die Matrizen \mathbf{L} und \mathbf{R} kolonnenweise bestimmen: man führt dieselben Rechnungen aus, aber in der Reihenfolge $r_{11}, l_{21}, l_{31}, r_{12}, r_{22}, l_{32}, r_{13}, r_{23}, r_{33}$. \blacklozenge

Eine weitere Option ist, die Pivotelemente in eine Diagonalmatrix \mathbf{D} einzubringen und dafür sowohl \mathbf{L} als auch \mathbf{R} durch Diagonalelemente 1 zu normieren. Das erfordert nur eine kleine Modifikation der Formeln (3.24a)–(3.24b). Formal bedeutet es, dass

$$\mathbf{D} \equiv \text{diag}(r_{11}, r_{22}, \dots, r_{nn}), \quad \mathbf{R}_1 \equiv \mathbf{D}^{-1} \mathbf{R} \quad (3.25)$$

und

$$\boxed{\mathbf{PA} = \mathbf{LDR}_1}. \quad (3.26)$$

Dies ist die **LDR-Zerlegung** oder **LDU-Zerlegung** [*LDU decomposition*] von \mathbf{A} .

Wir werden in Abschnitt 3.4 auf sie zurückkommen.

In grossen Problemen, die auf der Diskretisation von partiellen Differentialgleichungen beruhen, spielen Versionen der LR-Zerlegung eine wichtige Rolle, die angepasst sind an die Struktur dünn besetzter Matrizen, die pro Zeile nur relativ wenige Elemente haben, die nicht null sind.

Das grundsätzliche Vorgehen beim Lösen von $\mathbf{Ax} = \mathbf{b}$ mittels LR-Zerlegung sei noch einmal zusammengefasst:

Algorithmus 3.2 (Gauss-Elimination durch LR-Zerlegung, Vor- und Rückwärtseinsetzen)

Zum Lösen eines regulären Gleichungssystems $\mathbf{Ax} = \mathbf{b}$ kann man, als Alternative zur Gauss-Elimination von Algorithmus 1.1, wie folgt vorgehen:

1. LR-Zerlegung von \mathbf{A} zum Beispiel gemäss Algorithmus 1.1 (aber ohne rechte Seite \mathbf{b}) oder gemäss Algorithmus 3.1; sie erzeugt die Matrizen \mathbf{L} und \mathbf{R} sowie die durch \mathbf{P} darstellbare Permutation der Zeilen.
2. Vorwärtseinsetzen: $\mathbf{Lc} = \mathbf{b}$ auflösen nach \mathbf{c} .
3. Rückwärtseinsetzen: $\mathbf{Rx} = \mathbf{c}$ auflösen nach \mathbf{x} .

Speicherbedarf und Rechenaufwand

Wie man unseren Beispielen aus Kapitel 1 entnimmt, siehe etwa Beispiel 1.9 auf Seite 1-5, lassen sich die im Laufe einer LR-Zerlegung erzeugten und noch gebrauchten Daten jederzeit in einem Datenfeld der Grösse $n \times n$ unterbringen. Die nicht-null Elemente von \mathbf{L} (ausser die auf 1 normierten Diagonalelemente) kann man nämlich genau dort abspeichern, wo Nullen erzwungen werden.

Bei Zeilenvertauschungen braucht man eine zusätzliche Kolonne, um zum Beispiel die ursprünglichen Zeilennummern abzuspeichern.

Der Rechenaufwand für die LR-Zerlegung und das Vor- und Rückwärtseinsetzen für ℓ verschiedene rechte Seiten ist genau gleich wie für die Gauss-Elimination gemäss Algorithmus 1.1 mit ℓ Konstantenkolonnen. Es kommt nicht darauf an, ob man Restgleichungssysteme nachführt oder die kompakten Formeln aus Algorithmus 3.1 benutzt. Es ist allerdings von Vorteil, wenn man die reziproken Pivotelemente r_{kk}^{-1} berechnet und abspeichert. Es sind dann total nur n Divisionen nötig. (Divisionen brauchen auf vielen Computern mehr Zeit als Multiplikationen.)

Im j -ten Schritt von Algorithmus 1.1 braucht man dann

1	Division für	$(a_{jj}^{(j-1)})^{-1}$	[Formel (1.10)]
$n - j$	Multiplikationen für die	l_{kj}	[Formel (1.10)]
$(n - j)^2$	Add. & Mult. für die	$a_{ki}^{(j)}$	[Formel (1.11a)]
$\ell(n - j)$	Add. & Mult. für die	$b_k^{(j)}$	[Formel (1.11b)]

und für den k -ten Teilschritt des ℓ -maligen Rückwärtseinsetzens

$\ell(n - k)$	Additionen für	x_k	[Formel (1.12)]
$\ell(n - k + 1)$	Multiplikationen für	x_k	[Formel (1.12)]

Ersetzen wir in der letzten Formel den Index k durch j , so können wir alle Beiträge für festes j addieren:

$$\begin{aligned}
 & 1 \text{ Division} \\
 & (n - j)^2 + (2\ell + 1)(n - j) + \ell \text{ Multiplikationen} \\
 & (n - j)^2 + 2\ell(n - j) \text{ Additionen}
 \end{aligned}$$

Diese Zahlen sind zu summieren von $j = 1$ bis $j = n$ (wobei es im letzten Schritt bloss eine Division im Hinblick auf das Rückwärtseinsetzen gibt). Wenn wir $k := n - j$ setzen, so erhalten wir zum Beispiel für die Multiplikationen

$$\begin{aligned}
 \sum_{k=0}^{n-1} (k^2 + k(2\ell + 1) + \ell) &= \frac{n(n-1)(2n-1)}{6} + \frac{n(n-1)(2\ell+1)}{2} + \ell n \\
 &= \frac{1}{3}n^3 + \ell n^2 + \mathcal{O}(n),
 \end{aligned}$$

wobei im Term $\mathcal{O}(n)$ alle Beiträge untergebracht sind, die linear in n oder konstant sind. Der ℓn -Term fällt hier weg; bei der Addition ist er negativ. Für $\ell = 0$ erhalten wir den Rechenaufwand für die reine LR-Zerlegung. Wir können das so zusammenfassen:

Satz 3.2 *Der Rechenaufwand für die LR-Zerlegung einer voll besetzten, regulären $n \times n$ -Matrix beträgt n Divisionen (inklusive der letzten, die man erst beim Rückwärtseinsetzen braucht), $\frac{1}{3}n^3 + \mathcal{O}(n)$ Multiplikationen und $\frac{1}{3}n^3 - \frac{1}{2}n^2 + \mathcal{O}(n)$ Additionen. Für Vor- und Rückwärtseinsetzen braucht man zusätzlich pro rechte Seite je $n^2 - n$ Additionen und n^2 Multiplikationen.*

Das Lösen eines Systems $\mathbf{A}\mathbf{X} = \mathbf{B}$ mit n rechten Seiten erfordert dementsprechend n Divisionen, $\frac{4}{3}n^3 + \mathcal{O}(n)$ Multiplikationen und $\frac{4}{3}n^3 - \frac{3}{2}n^2 + \mathcal{O}(n)$ Additionen.

Der zweite Teil des Satzes gibt im wesentlichen auch den Aufwand für die Berechnung der Inversen an, wobei man allerdings in diesem Falle (wo $\mathbf{B} = \mathbf{I}$ ist), noch beim Vorwärtseinsetzen Operationen sparen kann, weil die rechten Seiten viele Nullen haben. Dabei ist der führende Term von $\frac{4}{3}n^3$ um ein Viertel auf je n^3 Additionen und Multiplikationen reduzierbar.

Pivotstrategien

Es ist wichtig, daran zu erinnern, dass selbst bei quadratischen Matrizen der Gauss-Algorithmus und die dazu äquivalente LR-Zerlegung in einer Arithmetik mit endlicher Genauigkeit (wo Rundungsfehler auftreten) nur dann zuverlässige Resultate liefert, wenn man die Pivots geeignet wählt und wenn die Matrix nicht beinahe singular ist (was immerhin gleichzeitig überprüft werden kann).

Die meistens benutzte Regel (**Pivotstrategie** [*pivot strategy*]) besagt, dass man im j ten Eliminationsschritt als Pivot $a_{pj}^{(j-1)}$ jenes Element wählen soll, das unter allen $m - j + 1$ Elementen der ersten Kolonne der Restgleichungsmatrix das betragsmässig grösste ist:

$$|a_{pj}^{(j-1)}| = \max_{j \leq k \leq m} |a_{kj}^{(j-1)}|. \quad (3.27)$$

Diese Regel nennt man **Kolonnenmaximumstrategie** oder **partielles Pivotieren** [*partial pivoting*]. Damit lässt sich fast immer vermeiden, dass in \mathbf{L} und \mathbf{R} Elemente entstehen, die viel grösser sind als jene von \mathbf{A} und zu grossen Rundungsfehlern führen können. Allerdings setzt dies voraus, dass die Zeilen von \mathbf{A} von der gleichen Grössenordnung sind.

Noch sicherer wäre es, auch Kolonnenvertauschungen zuzulassen und in der ganzen Restgleichungsmatrix nach dem grössten oder sogar dem relativ grössten Element (im Vergleich zu den anderen derselben Zeile) zu suchen. Aber dieses **vollständige Pivotieren** [*complete pivoting*] ist zu aufwendig und wird kaum je angewandt.

3.2 Die Gauss-Elimination als LR-Zerlegung: der allgemeine Fall

Im allgemeinen Falle der Reduktion irgend einer $m \times n$ -Matrix auf Zeilenstufenform ist \mathbf{R} von der Form in (1.28), das heisst

$$\mathbf{R} = \left(\begin{array}{cccc|cccc} 0 & \cdots & r_{1,n_1} & \cdots & * & \cdots & \cdots & * & * & \cdots & * \\ 0 & \cdots & 0 & \cdots & r_{2,n_2} & \cdots & \cdots & * & * & \cdots & * \\ \vdots & & \vdots & & \vdots & & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & \cdots & r_{r,n_r} & * & \cdots & * \\ \hline 0 & \cdots & 0 & \cdots & 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \end{array} \right), \quad (3.28)$$

wobei die eingezeichneten Pivotelemente $r_{1,n_1}, r_{2,n_2}, \dots, r_{r,n_r}$ nicht null sind und wir jetzt auch noch die möglichen Null-Kolonnen am linken Rand eingetragen haben.

Die Formeln (3.1)–(3.5) für den j -ten Eliminationsschritt sind nun zu ersetzen durch

$$r_{ji} := a_{ji}^{(j-1)} \quad (i = n_j + 1, \dots, n), \quad (3.29)$$

$$c_j := b_j^{(j-1)}, \quad (3.30)$$

$$l_{kj} := a_{k,n_j}^{(j-1)} / r_{j,n_j} \quad (k = j + 1, \dots, n), \quad (3.31)$$

$$a_{ki}^{(j)} := a_{ki}^{(j-1)} - l_{kj} r_{ji}, \quad (i = n_j + 1, \dots, n, \\ k = j + 1, \dots, n), \quad (3.32)$$

$$b_k^{(j)} := b_k^{(j-1)} - l_{kj} c_j \quad (k = j + 1, \dots, n). \quad (3.33)$$

In (3.29), (3.30) und (3.33) gibt es keine Änderung, in den anderen Formeln ist zum Teil der Kolonnenindex j durch n_j ersetzt worden.

Die Multiplikatoren l_{kj} passen in die $m \times m$ -Linksdreiecksmatrix

$$\mathbf{L} = \left(\begin{array}{cccc|cccc} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ l_{21} & 1 & \ddots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \ddots & 0 & \vdots & & \vdots & \vdots \\ l_{r1} & l_{r2} & \cdots & 1 & 0 & \cdots & 0 & 0 \\ \hline l_{r+1,1} & l_{r+1,2} & \cdots & l_{r+1,r} & 1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & & \vdots & 0 & \ddots & 0 & \vdots \\ \vdots & \vdots & & \vdots & \vdots & \ddots & 1 & 0 \\ l_{m1} & l_{m2} & \cdots & l_{mr} & 0 & \cdots & 0 & 1 \end{array} \right) \quad (3.34)$$

mit

$$l_{kj} := 0 \quad (1 \leq k < j \leq m), \quad (3.35)$$

$$l_{kj} := 0 \quad (r < j < k \leq m), \quad (3.36)$$

$$l_{jj} := 1 \quad (1 \leq k < j \leq m). \quad (3.37)$$

Schliesslich wird die Formel (3.12) für das Rückwärtseinsetzen ersetzt durch (1.27), das heisst

$$x_{n_k} := \left(b_k^{(k-1)} - \sum_{i=n_k+1}^n a_{ki}^{(k-1)} x_i \right) \frac{1}{a_{k,n_k}^{(k-1)}}, \quad (3.38)$$

Nun ist es nicht mehr schwierig nachzuprüfen, dass mit diesen Definitionen Satz 3.1 mit geringen Änderungen weiterhin gilt:

Satz 3.3 *Angewandt auf ein beliebiges $m \times n$ -System $\mathbf{Ax} = \mathbf{b}$ liefert die Gauss-Elimination eine die Zeilenvertauschungen beschreibende $m \times m$ -Permutationsmatrix \mathbf{P} , eine $m \times n$ Zeilenstufenmatrix \mathbf{R} des reduzierten Systems und eine entsprechende rechte Seite $\mathbf{c} \in \mathbb{E}^m$ sowie, durch Zusammenziehen der Zeilenmultiplikatoren l_{kj} , eine reguläre $m \times m$ Linksdreiecksmatrix \mathbf{L} , wobei gilt*

$$\boxed{\mathbf{PA} = \mathbf{LR}, \quad \mathbf{Lc} = \mathbf{Pb}, \quad \mathbf{Rx} = \mathbf{c}.} \quad (3.39)$$

Ist die LR-Zerlegung $\mathbf{PA} = \mathbf{LR}$ einmal berechnet, so lässt sich irgend ein System mit Koeffizientenmatrix \mathbf{A} lösen durch Auflösen von $\mathbf{Lc} = \mathbf{Pb}$ nach \mathbf{c} (Vorwärtseinsetzen) und Auflösen von $\mathbf{Rx} = \mathbf{c}$ nach \mathbf{x} (Rückwärtseinsetzen). Beim letzteren sind allfällige freie Variablen (x_k mit Indizes $k \neq n_j(\forall j)$) frei wählbar.

Im Fall $\text{Rang } A = r < m$ lässt sich die LR-Zerlegung von Satz 3.3 vereinfachen: da alle Elemente der letzten $m - r > 0$ Zeilen von \mathbf{R} null sind und im Produkt \mathbf{LR} auf die Elemente in den letzten $m - r$ Kolonnen von \mathbf{L} treffen, kann man sowohl diese Kolonnen als auch die Nullzeilen von \mathbf{R} wegstreichen. Dabei wird \mathbf{R} zu einer $r \times n$ -Matrix $\tilde{\mathbf{R}}$ und \mathbf{L} zu einer $m \times r$ -Matrix $\tilde{\mathbf{L}}$, wobei gilt

$$\tilde{\mathbf{L}}\tilde{\mathbf{R}} = \mathbf{LR} = \mathbf{PA}.$$

Das liefert das folgende Korollar von Satz 3.3:

Korollar 3.4 *Zu jeder $m \times n$ -Matrix \mathbf{A} mit Rang r gibt es eine $m \times r$ -Matrix $\tilde{\mathbf{L}}$ mit Einsen in der Diagonalen, eine $r \times n$ -Matrix $\tilde{\mathbf{R}}$ in Zeilenstufenform und eine $m \times m$ -Permutationsmatrix \mathbf{P} , so dass*

$$\tilde{\mathbf{L}}\tilde{\mathbf{R}} = \mathbf{PA}. \quad (3.40)$$

Auch mit dieser Zerlegung ist $\mathbf{Ax} = \mathbf{b}$ natürlich äquivalent zu $\tilde{\mathbf{L}}\tilde{\mathbf{R}}\mathbf{x} = \mathbf{Pb}$ und kann deshalb durch Vor- und Rückwärtseinsetzen gelöst werden. Nun ist zuerst das $m \times r$ -System $\tilde{\mathbf{L}}\tilde{\mathbf{c}} = \mathbf{Pb}$ zu lösen. Weil $r < m$ ist, ist das aber nicht immer möglich. Die ersten r Gleichungen ergeben $\tilde{\mathbf{c}}$ eindeutig durch Vorwärtseinsetzen. Die restlichen $m - r$ Gleichungen sind die Verträglichkeitsbedingungen. Nun treten die Verträglichkeitsbedingungen also bereits beim Vorwärtseinsetzen auf. Falls sie erfüllt sind, können alle Lösungen aus dem $r \times n$ -System $\tilde{\mathbf{R}}\mathbf{x} = \tilde{\mathbf{c}}$ durch Rückwärtseinsetzen bestimmt werden, wobei hier wieder die freie Variablen wählbar sind, wenn $n > r$ ist.

BEISPIEL 3.4: In unserem Beispiel 1.4 auf Seite 1-11 ist $m = 7$ und $n = 9$. Gegeben sind

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 5 & 0 & 4 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 5 & 0 & 24 & 16 & 8 & 6 \\ 1 & 5 & 9 & 3 & 6 & 1 & 0 & 1 & 0 \\ 1 & 10 & 13 & 11 & 8 & 6 & 3 & 1 & 2 \\ 0 & 5 & 4 & 18 & 2 & 18 & 12 & 2 & 7 \\ 1 & 10 & 13 & 21 & 8 & 24 & 16 & 5 & 8 \\ 0 & 5 & 4 & 13 & 2 & 24 & 17 & 6 & 7 \end{pmatrix} \quad (3.41a)$$

und

$$\mathbf{b} = (1 \ 12 \ 3 \ 8 \ 13 \ 19 \ 16)^T. \quad (3.41b)$$

Wir haben bei der Lösung im zweiten Schritt die Zeilen 2 und 3, im dritten Schritt die Zeilen 3 und 4, und im vierten Schritt die Zeilen 4 und 5 vertauscht, wodurch die Permutation $2 \rightarrow 5, 3 \rightarrow 2, 4 \rightarrow 3, 5 \rightarrow 4$ definiert wird, die durch die folgende Permutationsmatrix \mathbf{P} beschrieben wird:

$$\mathbf{P} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \quad (3.42)$$

Unserer Lösung entnimmt man weiter, dass $r = 5$ ist, und, wenn man die Auswirkung der Zeilenvertauschungen auf bereits berechnete Kolonnen von \mathbf{L} berücksichtigt, dass

$$\mathbf{L} = \left(\begin{array}{cccccc|cc} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 3 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 4 & 1 & 0 & 0 & 0 \\ \hline 1 & 2 & 3 & 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 3 & 2 & 0 & 1 & 0 \end{array} \right), \quad (3.43a)$$

$$\mathbf{R} = \begin{pmatrix} 1 & 0 & 5 & 0 & 4 & 0 & 0 & 1 & 0 \\ \boxed{0} & 5 & 4 & 3 & 2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \boxed{5} & 0 & 4 & 3 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & \boxed{5} & 3 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \boxed{1} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad (3.43b)$$

$$\mathbf{c} = (1 \ 2 \ 3 \ 2 \ 1 \ 0 \ 0)^T. \quad (3.43c)$$



BEMERKUNG: Der hier behandelte allgemeine Gauss-Algorithmus, mit dem man jede $m \times n$ -Matrix auf Zeilenstufenform bringen kann, ist primär ein theoretischer Algorithmus für exakte Arithmetik. Man muss ja entscheiden, ob eine potentielle Pivotkolonne null ist und übersprungen werden soll. Mit anderen Worten, der Rang einer Matrix ist numerisch nicht immer exakt definiert. ▼

3.3 Block-LR-Zerlegung und LR-Updating

Die Gauss-Elimination und ihre Interpretation als LR-Zerlegung ist auch mit Blockmatrizen machbar, deren Blöcke als unteilbare Einheiten aufgefasst werden. Allerdings muss man die den Pivotelementen entsprechenden Blöcke invertieren können, d.h. sie müssen regulär sein, also insbesondere quadratisch. Wir wollen nur den Fall einer regulären 2×2 Blockmatrix betrachten,

$$\tilde{\mathbf{A}} := \left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array} \right), \quad (3.44)$$

wobei \mathbf{A} auch regulär sein muss (und damit \mathbf{D} sicher quadratisch).

In Anlehnung an die Gauss-Elimination subtrahieren wir das \mathbf{CA}^{-1} -fache der ersten Blockzeile von der zweiten, was

$$\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{O} & \mathbf{S} \end{array} \right) \quad \text{mit} \quad \mathbf{S} := \mathbf{D} - \mathbf{CA}^{-1}\mathbf{B} \quad (3.45)$$

ergibt. Der Block \mathbf{S} heisst **Schur-Komplement**¹ [*Schur complement*]. Die Operation, die (3.44) in (3.45) überführt, kann man als Matrixmultiplikation darstellen:

$$\left(\begin{array}{c|c} \mathbf{I} & \mathbf{O} \\ \hline -\mathbf{CA}^{-1} & \mathbf{I} \end{array} \right) \left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array} \right) = \left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{O} & \mathbf{S} \end{array} \right). \quad (3.46)$$

Hier ist die erste 2×2 Blockmatrix ganz einfach zu invertieren; man muss bloss das Vorzeichen des $(2, 1)$ -Blockes ändern:

$$\left(\begin{array}{c|c} \mathbf{I} & \mathbf{O} \\ \hline -\mathbf{CA}^{-1} & \mathbf{I} \end{array} \right)^{-1} = \left(\begin{array}{c|c} \mathbf{I} & \mathbf{O} \\ \hline \mathbf{CA}^{-1} & \mathbf{I} \end{array} \right). \quad (3.47)$$

Damit ergibt sich aus (3.46) die **Block-LR-Zerlegung** [*Block LR decomposition*]

$$\underbrace{\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array} \right)}_{\equiv: \tilde{\mathbf{A}}} = \underbrace{\left(\begin{array}{c|c} \mathbf{I} & \mathbf{O} \\ \hline \mathbf{CA}^{-1} & \mathbf{I} \end{array} \right)}_{\equiv: \tilde{\mathbf{L}}} \underbrace{\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{O} & \mathbf{S} \end{array} \right)}_{\equiv: \tilde{\mathbf{R}}}. \quad (3.48)$$

Wir können diese Zerlegung noch insofern modifizieren, als wir den $(1, 1)$ -Block \mathbf{A} auf die zwei Faktoren aufteilen können. Ist zum Beispiel eine LR-Zerlegung $\mathbf{A} = \mathbf{LR}$ dieses Blockes möglich und bekannt, so können wir wegen $\mathbf{A}^{-1} = \mathbf{R}^{-1}\mathbf{L}^{-1}$ statt (3.48) schreiben

$$\underbrace{\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array} \right)}_{\equiv: \tilde{\mathbf{A}}} = \underbrace{\left(\begin{array}{c|c} \mathbf{L} & \mathbf{O} \\ \hline \mathbf{CR}^{-1} & \mathbf{I} \end{array} \right)}_{\equiv: \tilde{\mathbf{L}}} \underbrace{\left(\begin{array}{c|c} \mathbf{R} & \mathbf{L}^{-1}\mathbf{B} \\ \hline \mathbf{O} & \mathbf{S} \end{array} \right)}_{\equiv: \tilde{\mathbf{R}}}, \quad (3.49)$$

¹ISSAI SCHUR (10.1.1875 – 10.1.1941), ab 1916 Professor in Berlin, 1938 Emigration nach Palästina. Seine theoretischen Arbeiten enthalten auch wichtige Algorithmen.

mit nun neuen Matrizen $\tilde{\mathbf{L}}$ und $\tilde{\mathbf{R}}$.

Die Formeln (3.48) und (3.49) haben interessante Spezialfälle und viele Anwendungen. Insbesondere kann man auch theoretische Ergebnisse aus ihnen ableiten. Wir fangen mit einem solchen an.

Lemma 3.5 *In der 2×2 Blockmatrix $\tilde{\mathbf{A}} \in \mathbb{E}^{n \times n}$ aus (3.44) sei der $(1,1)$ -Block \mathbf{A} regulär. Dann ist $\tilde{\mathbf{A}}$ genau dann regulär, wenn das Schur-Komplement \mathbf{S} aus (3.45) regulär ist.*

BEWEIS: Weil \mathbf{A} regulär ist, lässt sich der Schritt von (3.44) zu (3.45) durchführen, und es gilt (3.46), wobei dort die erste Matrix regulär ist, denn (3.47) ist ja eine Formel für die Inverse.

Weil diese Matrix regulär ist, lässt sich für jede rechte Seite das System

$$\left(\begin{array}{c|c} \mathbf{I} & \mathbf{O} \\ \hline -\mathbf{CA}^{-1} & \mathbf{I} \end{array} \right) \tilde{\mathbf{y}} = \left(\begin{array}{c} \mathbf{b} \\ \mathbf{c} \end{array} \right) \quad (3.50)$$

lösen, wobei wir hier annehmen, dass diese entsprechend den Blockgrößen in der Matrix aufgeteilt sei. Nun ist auch $\tilde{\mathbf{A}}$ regulär genau dann, wenn das Gleichungssystem $\tilde{\mathbf{A}}\tilde{\mathbf{x}} = \tilde{\mathbf{y}}$ für jede rechte Seite $\tilde{\mathbf{y}} \in \mathbb{E}^n$ eine Lösung $\tilde{\mathbf{x}}$ hat, also insbesondere für jede Lösung $\tilde{\mathbf{y}}$ von (3.50). Diese ist dann eine Lösung von

$$\underbrace{\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{O} & \mathbf{S} \end{array} \right)}_{=\tilde{\mathbf{R}}} \tilde{\mathbf{x}} = \underbrace{\left(\begin{array}{c|c} \mathbf{I} & \mathbf{O} \\ \hline -\mathbf{CA}^{-1} & \mathbf{I} \end{array} \right)}_{=\tilde{\mathbf{L}}^{-1}} \underbrace{\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array} \right)}_{=\tilde{\mathbf{y}}} \tilde{\mathbf{x}} = \left(\begin{array}{c} \mathbf{b} \\ \mathbf{c} \end{array} \right).$$

Teilen wir nun $\tilde{\mathbf{x}} \equiv: \left(\begin{array}{cc} \mathbf{x}^\top & \mathbf{u}^\top \end{array} \right)^\top$ passend in zwei Blöcke auf, so folgt aus

$$\left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{O} & \mathbf{S} \end{array} \right) \left(\begin{array}{c} \mathbf{x} \\ \mathbf{u} \end{array} \right) = \left(\begin{array}{c} \mathbf{b} \\ \mathbf{c} \end{array} \right), \quad (3.51)$$

dass

$$\mathbf{S}\mathbf{u} = \mathbf{c} \quad (3.52)$$

ist. Wir können auf diese Weise also zu jedem \mathbf{c} eine Lösung dieser Gleichung konstruieren, wobei wir in (3.50) sogar $\mathbf{b} = \mathbf{o}$ wählen dürften. Es folgt, dass \mathbf{S} regulär ist.

Umgekehrt können wir, falls \mathbf{S} regulär ist, zu jedem \mathbf{c} eine Lösung \mathbf{u} von (3.52) gewinnen, womit sich mittels der ersten Blockzeile von (3.51),

$$\mathbf{A}\mathbf{x} = \mathbf{b} - \mathbf{B}\mathbf{u},$$

für beliebiges \mathbf{b} ein reguläres System für \mathbf{x} ergibt. Mit anderen Worten, wir können zu jedem $\tilde{\mathbf{b}} \equiv: \left(\begin{array}{cc} \mathbf{b}^\top & \mathbf{c}^\top \end{array} \right)^\top$ eine Lösung $\tilde{\mathbf{x}} \equiv: \left(\begin{array}{cc} \mathbf{x}^\top & \mathbf{u}^\top \end{array} \right)^\top$ von $\tilde{\mathbf{A}}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}$ finden, was bedeutet, dass $\tilde{\mathbf{A}}$ regulär ist. (Effektiv lösen wir hier $\tilde{\mathbf{A}}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}$ durch Block-Rückwärtseinsetzen.) ■

Wir werden später in Kapitel 8 ein Lemma beweisen, aus dem man die obige Aussage von Lemma 3.5 sofort ablesen kann.

Zwei Spezialfälle obiger Block-LR-Zerlegung sind besonders interessant: wenn entweder der (1, 1)-Block oder der (2, 2)-Block ein Skalar ist. Im ersten Fall hat man aus (3.48) oder (3.49)

$$\underbrace{\left(\begin{array}{c|c} \alpha & \mathbf{b}^\top \\ \mathbf{c} & \mathbf{D} \end{array}\right)}_{=\tilde{\mathbf{A}}} = \underbrace{\left(\begin{array}{c|c} 1 & \mathbf{o}^\top \\ \mathbf{c}\alpha^{-1} & \mathbf{I} \end{array}\right)}_{=\tilde{\mathbf{L}}} \underbrace{\left(\begin{array}{c|c} \alpha & \mathbf{b}^\top \\ \mathbf{o} & \mathbf{S} \end{array}\right)}_{=\tilde{\mathbf{R}}} \quad (3.53)$$

mit dem Schur-Komplement

$$\mathbf{S} = \mathbf{D} - \mathbf{c}\alpha^{-1}\mathbf{b}^\top, \quad (3.54)$$

das eine Rang-1-Modifikation von \mathbf{D} ist.

Diese zwei Formeln beschreiben offenbar gerade den ersten Schritt der (gewöhnlichen) LR-Zerlegung, und man sieht, dass das Schur-Komplement gerade die Koeffizientenmatrix des ersten Restgleichungssystems ist.

Analog gilt im zweiten Fall, in dem wir uns auf (3.49) stützen wollen, also annehmen, dass der (1, 1)-Block \mathbf{A} bereits in $\mathbf{A} = \mathbf{L}\mathbf{R}$ zerlegt sei:

$$\underbrace{\left(\begin{array}{c|c} \mathbf{A} & \mathbf{b} \\ \mathbf{c}^\top & \delta \end{array}\right)}_{=\tilde{\mathbf{A}}} = \underbrace{\left(\begin{array}{c|c} \mathbf{L} & \mathbf{o} \\ \mathbf{c}^\top\mathbf{R}^{-1} & 1 \end{array}\right)}_{=\tilde{\mathbf{L}}} \underbrace{\left(\begin{array}{c|c} \mathbf{R} & \mathbf{L}^{-1}\mathbf{b} \\ \mathbf{o}^\top & \sigma \end{array}\right)}_{=\tilde{\mathbf{R}}} \quad (3.55)$$

mit dem skalaren Schur-Komplement

$$\sigma = \delta - \mathbf{c}^\top\mathbf{A}^{-1}\mathbf{b} = \delta - \mathbf{c}^\top\mathbf{R}^{-1}\mathbf{L}^{-1}\mathbf{b}. \quad (3.56)$$

Wenn wir diese zwei Formeln rekursiv anwenden, eröffnen sie uns eine neue, rekursive Variante der LR-Zerlegung, die allerdings keine Zeilenvertauschungen zulässt. Wir führen dazu noch geeignete Bezeichnungen ein.

DEFINITION: Die einer $m \times n$ Matrix $\mathbf{A} = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}$ zugeordneten $\min\{m, n\}$ quadratischen Teilmatrizen

$$\mathbf{A}_k = (a_{ij})_{1 \leq i \leq k, 1 \leq j \leq k} \quad (k = 1, \dots, \min\{m, n\}) \quad (3.57)$$

heissen **führende Hauptuntermatrizen** [*leading principal submatrices*] von \mathbf{A} . ▲

Unter Verwendung der Bezeichnungen

$$\mathbf{A}_k = \mathbf{L}_k\mathbf{R}_k, \quad \mathbf{A}_{k+1} = \left(\begin{array}{c|c} \mathbf{A}_k & \mathbf{b}_k \\ \mathbf{c}_k^\top & a_{k+1,k+1} \end{array}\right) \quad (3.58)$$

können wir Formeln (3.55)–(3.56) umschreiben in

$$\underbrace{\left(\begin{array}{c|c} \mathbf{A}_k & \mathbf{b}_k \\ \mathbf{c}_k^\top & a_{k+1,k+1} \end{array}\right)}_{=:\mathbf{A}_{k+1}} = \underbrace{\left(\begin{array}{c|c} \mathbf{L}_k & \mathbf{o} \\ \mathbf{c}_k^\top\mathbf{R}_k^{-1} & 1 \end{array}\right)}_{=:\mathbf{L}_{k+1}} \underbrace{\left(\begin{array}{c|c} \mathbf{R}_k & \mathbf{L}_k^{-1}\mathbf{b}_k \\ \mathbf{o}^\top & \sigma_{k+1} \end{array}\right)}_{=:\mathbf{R}_{k+1}}, \quad (3.59)$$

$$\sigma_{k+1} := a_{k+1,k+1} - \mathbf{c}_k^\top\mathbf{R}_k^{-1}\mathbf{L}_k^{-1}\mathbf{b}_k. \quad (3.60)$$

Dabei braucht man für die Berechnung von \mathbf{R}_{k+1} , σ_{k+1} und \mathbf{L}_{k+1} gemäss (3.59)–(3.60) die Inversen \mathbf{R}_k^{-1} und \mathbf{L}_k^{-1} nicht, denn man kann ja $\mathbf{L}_k^{-1}\mathbf{b}_k$ und $(\mathbf{R}_k^\top)^{-1}\mathbf{c}_k$ durch Vorwärtseinsetzen berechnen.

Die Rekursionen für \mathbf{R}_{k+1} , σ_{k+1} und \mathbf{L}_{k+1} brechen zusammen, sobald σ_{k+1} null wird, denn dann existiert \mathbf{R}_{k+1}^{-1} nicht. In der Tat ist σ_{k+1} das Pivotelement im $(k+1)$ ten Eliminationsschritt. Nachfolgend ein Kriterium dafür, dass es nicht zum Abbruch kommt.

Satz 3.6 *Eine $m \times n$ Matrix \mathbf{A} vom Rang r lässt sich genau dann ohne Zeilenvertauschungen LR-zerlegen, wenn die r führenden Hauptuntermatrizen \mathbf{A}_k ($k = 1, \dots, r$) regulär sind.*

BEWEIS: Für Matrizen vom Rang 1 ist der Satz richtig, denn eine LR-Zerlegung existiert genau dann, wenn a_{11} nicht null ist. Nach einem Schritt ist die Zerlegung fertig.

Allgemein: Lässt sich \mathbf{A} ohne Zeilenvertauschungen LR-zerlegen, so gilt $\mathbf{A} = \mathbf{LR}$, wobei $r_{11}, \dots, r_{rr} \neq 0$. Dabei ist $\mathbf{A}_k = \mathbf{L}_k\mathbf{R}_k$ ($k = 1, \dots, r$), wobei die Faktoren und damit auch $\mathbf{A}_1, \dots, \mathbf{A}_r$ regulär sind.

Für die Umkehrung verwenden wir Induktion nach r , wobei die Verankerung ($r = 1$) schon erledigt ist. Wir nehmen also an, dass wenn $\text{Rang}(\mathbf{A}) \equiv r = \ell$ ist und $\mathbf{A}_1, \dots, \mathbf{A}_r$ regulär sind, die LR-Zerlegung $\mathbf{A} = \mathbf{LR}$ existiert und wollen zeigen, dass das auch für $r = \ell + 1$ gilt.

Es gelte also $\text{Rang}(\mathbf{A}) \equiv r = \ell + 1$, und es seien $\mathbf{A}_1, \dots, \mathbf{A}_r$ regulär. Weil $\mathbf{A}_1, \dots, \mathbf{A}_{r-1}$ regulär sind, lässt sich \mathbf{A}_{r-1} ohne Zeilenvertauschungen LR-zerlegen, d.h. $\mathbf{A}_{r-1} = \mathbf{L}_{r-1}\mathbf{R}_{r-1}$, wobei \mathbf{L}_{r-1} und \mathbf{R}_{r-1} auch die führenden Hauptuntermatrizen der Faktoren \mathbf{L} und \mathbf{R} sind, die zu einer allfälligen LR-Zerlegung von \mathbf{A} gehören. Da auch \mathbf{A}_r regulär ist, folgt aus Lemma 3.5, dass das Schurkomplement von \mathbf{A}_{r-1} bezüglich \mathbf{A}_r nicht null ist: $\sigma = r_{rr} \neq 0$. Damit sind r nichtverschwindende Pivots gefunden; die allenfalls verbleibenden $m - r$ Zeilen von \mathbf{R} (in der Zeilenstufenform von \mathbf{A}) enthalten lauter Nullen, und es gilt $\mathbf{A} = \mathbf{LR}$. ■

Zusammenfassend erhalten wir:

Algorithmus 3.3 (LR-Zerlegung durch Updating) *Zur LR-Zerlegung einer regulären Matrix \mathbf{A} , deren $n - 1$ führende Hauptuntermatrizen \mathbf{A}_k ($k = 1, \dots, n - 1$) regulär sind, setze man zunächst $\mathbf{L}_1 := \begin{pmatrix} 1 \end{pmatrix}$, $\mathbf{R}_1 := \begin{pmatrix} a_{11} \end{pmatrix}$. Für $k = 1, \dots, n - 1$ teile man dann \mathbf{A}_{k+1} gemäss (3.58) in vier Blöcke auf und berechne*

$$\sigma_{k+1} := a_{k+1,k+1} - \mathbf{c}_k^\top \mathbf{R}_k^{-1} \mathbf{L}_k^{-1} \mathbf{b}_k, \quad (3.61a)$$

$$\mathbf{L}_{k+1} := \left(\begin{array}{c|c} \mathbf{L}_k & \mathbf{o} \\ \mathbf{c}_k^\top \mathbf{R}_k^{-1} & 1 \end{array} \right), \quad (3.61b)$$

$$\mathbf{R}_{k+1} := \left(\begin{array}{c|c} \mathbf{R}_k & \mathbf{L}_k^{-1} \mathbf{b}_k \\ \mathbf{o}^\top & \sigma_{k+1} \end{array} \right). \quad (3.61c)$$

Dabei erhält man $\mathbf{L}_k^{-1}\mathbf{b}_k$ und $(\mathbf{R}_k^\top)^{-1}\mathbf{c}_k$ durch Vorwärtseinsetzen. Der Algorithmus bricht nur ab, wenn entgegen der Voraussetzung eine der Hauptuntermatrizen singulär ist, und zwar ist $\sigma_k = 0$ (und damit \mathbf{R}_k singulär) genau dann, wenn \mathbf{A}_k singulär ist.

3.4 Die Cholesky-Zerlegung

In Abschnitt 3.1 haben wir darauf hingewiesen, siehe (3.25)–(3.26), dass man die LR-Zerlegung $\mathbf{PA} = \mathbf{LR}$ (mit Zeilenvertauschungen) jederzeit durch eine äquivalente LDR-Zerlegung $\mathbf{PA} = \mathbf{LDR}_1$ ersetzen kann, wo \mathbf{D} die Diagonale von \mathbf{R} und $\mathbf{R}_1 := \mathbf{D}^{-1}\mathbf{R}$ ist.

Ist \mathbf{A} symmetrisch oder Hermitesch und die Zerlegung ohne Zeilenvertauschen möglich, so ist dabei $\mathbf{P} = \mathbf{I}$ und automatisch $\mathbf{R}_1 = \mathbf{L}^\top$ bzw. $\mathbf{R}_1 = \mathbf{L}^H$, so dass gilt:

$$\boxed{\mathbf{A} = \mathbf{LDL}^\top} \quad \text{bzw.} \quad \boxed{\mathbf{A} = \mathbf{LDL}^H} \quad (3.62)$$

Sind zudem die Diagonalelemente $d_{ii} := r_{ii}$ von \mathbf{D} positiv, so kann man ihre Quadratwurzeln berechnen und jene von \mathbf{D} definieren als

$$\mathbf{D}^{\frac{1}{2}} := \text{diag}(\sqrt{r_{11}}, \sqrt{r_{22}}, \dots, \sqrt{r_{nn}}). \quad (3.63)$$

Setzt man nun

$$\tilde{\mathbf{R}} := \mathbf{D}^{\frac{1}{2}} \mathbf{L}^\top \quad \text{bzw.} \quad \tilde{\mathbf{R}} := \mathbf{D}^{\frac{1}{2}} \mathbf{L}^H, \quad (3.64)$$

erhält man die **Cholesky-Zerlegung**² [*Cholesky decomposition*]

$$\boxed{\mathbf{A} = \tilde{\mathbf{R}}^\top \tilde{\mathbf{R}}} \quad \text{bzw.} \quad \boxed{\mathbf{A} = \tilde{\mathbf{R}}^H \tilde{\mathbf{R}}}. \quad (3.65)$$

Bei der linken Version setzt man dabei voraus, dass die Matrizen reell sind, aus einem Grund, auf den wir gleich kommen werden.

Nun haben wir in Abschnitt 3.3 gesehen, dass man die LR-Zerlegung (theoretisch) ohne Zeilenvertauschungen ausführen kann genau dann, wenn die führenden Hauptuntermatrizen \mathbf{A}_k ($k = 1, \dots, n-1$) regulär sind. Aber wann können wir das garantieren? Und wann sind dann zudem die Diagonalelemente von \mathbf{D} positiv?

Es gibt eine wichtige Klasse reell symmetrischer oder Hermitescher Matrizen, wo das gewährleistet ist.

DEFINITION: Eine reell symmetrische $n \times n$ Matrix \mathbf{A} heisst **positiv definit** [*positive definite*] oder, kurz, eine **spd** Matrix, falls

$$\mathbf{x}^\top \mathbf{A} \mathbf{x} > 0 \quad \text{für alle } \mathbf{x} \in \mathbb{R}^n, \mathbf{x} \neq \mathbf{o}. \quad (3.66)$$

Eine Hermitesche $n \times n$ Matrix \mathbf{A} heisst **positiv definit** [*positive definite*] oder, kurz, eine **Hpd** Matrix, falls

$$\mathbf{x}^H \mathbf{A} \mathbf{x} > 0 \quad \text{für alle } \mathbf{x} \in \mathbb{C}^n, \mathbf{x} \neq \mathbf{o}. \quad (3.67)$$

Gilt (3.66) bzw. (3.67) nur mit dem \geq -Zeichen, so heisst die Matrix **positiv semidefinit** [*positive semidefinite*]. ▲

²ANDRÉ LOUIS CHOLESKY (15.10.1875 – 31.10.1918), französischer Offizier, der unter anderem Vermessungen in Algerien und Tunesien durchführte und im ersten Weltkrieg als Regimentskommandant fiel; Ritter der Ehrenlegion.

Lemma 3.7 *Eine reell symmetrische oder Hermitesche Matrix, die positiv definit ist, ist regulär.*

BEWEIS: Wäre eine spd oder Hpd Matrix \mathbf{A} singular, gäbe es nach Korollar 1.7 nichttriviale Lösungen \mathbf{x} des homogenen Systems $\mathbf{Ax} = \mathbf{o}$. Für diese \mathbf{x} wäre aber $\mathbf{x}^\top \mathbf{Ax} = 0$ bzw. $\mathbf{x}^\mathbf{H} \mathbf{Ax} = 0$ im Widerspruch zu (3.66) und (3.67). ■

Im folgenden betrachten wir den Hermiteschen Fall; den reell symmetrischen erhält man daraus als Spezialfall: im wesentlichen ist $^\mathbf{H}$ durch $^\top$ zu ersetzen.

Es sei also \mathbf{A} eine Hpd Matrix der Ordnung n . Wählt man $\mathbf{x} \in \mathbb{C}^n$ speziell von der Form

$$\mathbf{x} = \begin{pmatrix} \mathbf{u} \\ \mathbf{o} \end{pmatrix} \quad \text{mit } \mathbf{u} \in \mathbb{C}^k, \mathbf{u} \neq \mathbf{o},$$

so sieht man, dass auch jede Hauptuntermatrix \mathbf{A}_k positiv definit ist:

$$0 < \mathbf{x}^\mathbf{H} \mathbf{Ax} = \mathbf{u}^\mathbf{H} \mathbf{A}_k \mathbf{u} \quad \text{für alle } \mathbf{u} \in \mathbb{C}^k, \mathbf{u} \neq \mathbf{o}.$$

Insbesondere ist nach Lemma 3.7 \mathbf{A}_k für $k = 1, \dots, n$ regulär. Dies bedeutet nach Satz 3.6, dass man \mathbf{A} ohne Zeilen zu vertauschen LR-zerlegen kann: $\mathbf{A} = \mathbf{LR} = \mathbf{LDL}^\mathbf{H}$, wobei \mathbf{L} und \mathbf{D} regulär sind. Damit gilt für beliebiges $\mathbf{x} \in \mathbb{C}^n$, $\mathbf{x} \neq \mathbf{o}$, dass

$$0 < \mathbf{x}^\mathbf{H} \mathbf{Ax} = \mathbf{x}^\mathbf{H} \mathbf{LDL}^\mathbf{H} \mathbf{x} = \mathbf{y}^\mathbf{H} \mathbf{Dy}. \quad (3.68)$$

falls $\mathbf{y} := \mathbf{L}^\mathbf{H} \mathbf{x}$. Weil \mathbf{L}^\top regulär ist, gibt es zu jedem $\mathbf{y} \neq \mathbf{o}$ ein $\mathbf{x} \neq \mathbf{o}$, das heisst (3.68) gilt auch für alle $\mathbf{y} \neq \mathbf{o}$. Da \mathbf{D} diagonal ist, bedeutet das, dass $d_{ii} > 0$ ($i = 1, \dots, n$), womit man $\mathbf{D}^{\frac{1}{2}}$ und $\tilde{\mathbf{R}}$ wie in (3.63) und (3.64) definieren kann. Mit anderen Worten, die Cholesky-Zerlegung von \mathbf{A} existiert.

Umgekehrt, folgt aus der Existenz der Cholesky-Zerlegung $\mathbf{A} = \tilde{\mathbf{R}}^\mathbf{H} \mathbf{R}$ sofort, dass \mathbf{A} Hermitesch ist. Weiter gilt mit der Definition $\mathbf{y} := \tilde{\mathbf{R}} \mathbf{x}$, dass

$$\mathbf{x}^\mathbf{H} \mathbf{Ax} = \mathbf{x}^\mathbf{H} \tilde{\mathbf{R}}^\mathbf{H} \tilde{\mathbf{R}} \mathbf{x} = \mathbf{y}^\mathbf{H} \mathbf{y} > 0,$$

falls $\mathbf{y} \neq \mathbf{o}$. Dabei bedingt $\mathbf{y} = \mathbf{o}$ wieder, dass $\mathbf{x} = \mathbf{o}$ ist. Also ist \mathbf{A} positiv definit.

Damit ist folgender Satz bewiesen:

Satz 3.8 *Eine reell symmetrische oder Hermitesche Matrix \mathbf{A} ist genau dann positiv definit, wenn sie eine Cholesky-Zerlegung (3.65) hat, worin $\tilde{\mathbf{R}}$ eine (reelle bzw. komplexe) Rechtsdreiecksmatrix mit positiven Diagonalelementen ist.*

Es gibt wiederum verschiedene Varianten der Berechnung von $\tilde{\mathbf{R}}$: zeilenweise mit Berechnung des Schurkomplementes (analog zum Gauss-Algorithmus 1.1), oder zeilenweise direkt (analog zur LR-Zerlegung in Algorithmus 3.1, oder auf ähnliche Art kolonnenweise, oder durch Updating wie in Algorithmus 3.3. Die zeilenweise

Berechnung entspricht auch der Reihenfolge der führenden Hauptuntermatrizen.

Algorithmus 3.4 (Zeilenweise, direkte Cholesky-Zerlegung) Zur Cholesky-Zerlegung einer positiv definiten reell symmetrischen oder Hermiteschen $n \times n$ Matrix \mathbf{A} berechne man für $i = 1, \dots, n$:

$$\tilde{r}_{ii} := \sqrt{a_{ii} - \sum_{j=1}^{i-1} |\tilde{r}_{ji}|^2}, \quad (3.69a)$$

$$\tilde{r}_{ik} := \left(a_{ik} - \sum_{j=1}^{i-1} \tilde{r}_{ji} \tilde{r}_{jk} \right) \frac{1}{\tilde{r}_{ii}} \quad (k = i + 1, \dots, n). \quad (3.69b)$$

Dabei sind die Summen leer, wenn $i = 1$ ist. Im letzten Schritt mit $i = n$ entfällt die zweite Formel (3.69b).

Anwendung auf Gleichungssysteme; Rechenaufwand

Natürlich kann man die Cholesky-Zerlegung genau wie die LR-Zerlegung zum Lösen eines Gleichungssystems einsetzen.

Algorithmus 3.5 (Gauss-Elimination durch Cholesky-Zerlegung, Vor- und Rückwärtseinsetzen)

Zum Lösen eines Gleichungssystems $\mathbf{Ax} = \mathbf{b}$ mit symmetrisch oder Hermitesche positiv definiten Matrix \mathbf{A} kann man wie folgt vorgehen:

1. Cholesky-Zerlegung $\mathbf{A} = \tilde{\mathbf{R}}^H \tilde{\mathbf{R}}$ von \mathbf{A} , z.B. mittels Algorithmus 3.4.
2. Vorwärtseinsetzen: $\tilde{\mathbf{R}}^H \mathbf{c} = \mathbf{b}$ auflösen nach \mathbf{c} .
3. Rückwärtseinsetzen: $\tilde{\mathbf{R}} \mathbf{x} = \mathbf{c}$ auflösen nach \mathbf{x} .

Weil in der Cholesky-Zerlegung und in der symmetrischen LDR-Zerlegung (3.62) nur eine Dreiecksmatrix zu bestimmen ist, reduziert sich der Rechenaufwand im Vergleich zur LR-Zerlegung und zur Gauss-Elimination um fast die Hälfte. Im Detail gilt analog zu Satz 3.2 folgende Aussage:

Satz 3.9 Der Rechenaufwand für die Cholesky-Zerlegung einer voll besetzten, symmetrisch oder Hermitesch positiv definiten $n \times n$ -Matrix beträgt n Divisionen (inklusive der letzten, die man erst im Rückwärtseinsetzen braucht), $\frac{1}{6}n^3 + \frac{1}{2}n^2 + \mathcal{O}(n)$ Multiplikationen und $\frac{1}{6}n^3 + \mathcal{O}(n)$ Additionen. Für Vor- und Rückwärtseinsetzen braucht man zusätzlich pro rechte Seite je $n^2 - n$ Additionen und $n^2 + n$ Multiplikationen.

Kapitel 4

Vektorräume

In der Mathematik trifft man vielerorts folgende Struktur an: Die Elemente einer bestimmten Menge lassen sich in natürlicher Weise addieren und mit einer reellen (oder komplexen) Zahl multiplizieren (“strecken”), wobei das Resultat beider Operationen je wieder ein Element der Menge ist und gewisse einfache Regeln gelten. Diese Struktur, genannt Vektorraum oder linearer Raum, kann als Verallgemeinerung des n -dimensionalen Euklidischen Raumes, den wir in Kapitel 2 behandelt haben, aufgefasst werden.

4.1 Definition und Beispiele

DEFINITION: Ein **Vektorraum** [*vector space*] V über \mathbb{E} ($:\equiv \mathbb{R}$ oder \mathbb{C}) ist eine nichtleere Menge auf der eine **Addition** [*Addition*]

$$x, y \in V \quad \longmapsto \quad x + y \in V$$

und eine **skalare Multiplikation** [*scalar multiplication*]

$$\alpha \in \mathbb{E}, x \in V \quad \longmapsto \quad \alpha x \in V$$

definiert sind, wobei folgende Grundregeln gelten:

- (V1) $x + y = y + x$ ($\forall x, y \in V$),
- (V2) $(x + y) + z = x + (y + z)$ ($\forall x, y, z \in V$),
- (V3) es gibt ein ausgezeichnetes Element $o \in V$ mit
 $x + o = x$ ($\forall x \in V$),
- (V4) zu jedem $x \in V$ gibt es ein eindeutig bestimmtes $-x \in V$ mit
 $x + (-x) = o$,
- (V5) $\alpha(x + y) = \alpha x + \alpha y$ ($\forall \alpha \in \mathbb{E}, \forall x, y \in V$),
- (V6) $(\alpha + \beta)x = \alpha x + \beta x$ ($\forall \alpha, \beta \in \mathbb{E}, \forall x \in V$),
- (V7) $(\alpha \beta)x = \alpha(\beta x)$ ($\forall \alpha, \beta \in \mathbb{E}, \forall x \in V$),
- (V8) $1x = x$ ($\forall x \in V$).

Die Elemente von V heissen **Vektoren** [*vector*]. Das Element $o \in V$ ist der **Nullvektor** [*zero vector*]. Die Elemente von \mathbb{E} bezeichnet man als **Skalare** [*scalars*]. Statt Vektorraum sagt man auch **linearer Raum** [*linear space*]. Ist $\mathbb{E} = \mathbb{R}$, heisst der Vektorraum **reell**, falls $\mathbb{E} = \mathbb{C}$ **komplex**. ▲

Die Grundregeln (V1)–(V8) nennt man wieder **Axiome** [*axioms*], weil sie postulierte Grundeigenschaften der mathematischen Struktur “Vektorraum” sind. Allein auf ihnen aufbauend wird die ganze Theorie abstrakter Vektorräume aufgestellt. Will man andererseits nachweisen, dass eine bestimmte Menge zusammen mit geeignet gewählten Operationen “Addition” und “skalare Multiplikation” effektiv einen Vektorraum bildet, muss man verifizieren, dass wirklich in diesem Falle (V1) bis (V8) gelten. Solche Beispiele werden wir gleich betrachten. Häufig ist dabei der Nachweis der meisten Axiome so einfach, dass es sich nicht lohnt, die Details aufzuschreiben.

Die Axiome (V1)–(V4) besagen gerade, dass ein Vektorraum bezüglich der Addition eine kommutative Gruppe bildet; vgl. Seite 2-9.

Wir betrachten nun eine Reihe von Beispielen von Vektorräumen:

BEISPIEL 4.1: Der Raum \mathbb{R}^n der reellen n -Vektoren

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad \dots$$

mit der Addition

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \equiv \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix} \quad (4.1a)$$

und der skalaren Multiplikation

$$\alpha \mathbf{x} = \alpha \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \equiv \begin{pmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{pmatrix}. \quad (4.1b)$$

◆

BEISPIEL 4.2: Der Raum $C[a, b]$ der auf dem Intervall $[a, b]$ definierten und dort **stetigen reellen Funktionen** [*continuous real functions*] mit der (punktweisen) Addition

$$f + g : \quad x \in [a, b] \quad \longmapsto \quad f(x) + g(x) \in \mathbb{R} \quad (4.2a)$$

und der (punktweisen) skalaren Multiplikation

$$\alpha f : \quad x \in [a, b] \quad \longmapsto \quad \alpha f(x) \in \mathbb{R}. \quad (4.2b)$$

◆

BEISPIEL 4.3: Der Raum $C^m[a, b]$ der auf dem Intervall $[a, b]$ **m -mal stetig differenzierbaren Funktionen** [*m -times continuously differentiable functions*] mit (4.2a) und (4.2b). Hier ist $(0 < m \leq \infty)$. Zudem setzt man $C^0[a, b] \equiv C[a, b]$. ◆

BEISPIEL 4.4: Der Raum \mathcal{P}_m aller **Polynome vom Grad $\leq m$** [*polynomials of degree $\leq m$*] mit reellen Koeffizienten,

$$p(t) := a_0 + a_1 t + \cdots + a_m t^m, \quad q(t) := b_0 + b_1 t + \cdots + b_m t^m, \dots$$

mit der Addition

$$(p + q)(t) := (a_0 + b_0) + (a_1 + b_1)t + \cdots + (a_m + b_m)t^m \quad (4.3a)$$

und der skalaren Multiplikation:

$$(\alpha p)(t) := (\alpha a_0) + (\alpha a_1)t + \cdots + (\alpha a_m)t^m. \quad (4.3b)$$

Dies ist eine “algebraische” Definition der Addition und skalaren Multiplikation für Polynome. Die Potenzen $1, t, \dots, t^m$ sind nur “Platzhalter”; es wird nicht benützt, dass man für x einen reellen Wert einsetzen und so einen Funktionswert $p(x)$ berechnen kann. Wir könnten (4.3a), (4.3b) durch (4.2a), (4.2b) (mit p statt f , q statt g) ersetzen, was eine “analytische” Definition ergäbe.

Mit (4.3a), (4.3b) ist \mathcal{P}_m “äquivalent” zu \mathbb{R}^{m+1} ; mit (4.2a), (4.2b) ist \mathcal{P}_m dagegen ein “Teilraum” von $C(\mathbb{R})$. Die genauere Bedeutung von “äquivalent” und “Teilraum” folgt später.

\mathcal{P}_0 ist der Raum der Polynome vom Grade 0, das heisst der reellen Konstanten. Er ist offensichtlich “äquivalent” zu \mathbb{R}^1 , das heisst zu den reellen Zahlen aufgefasst als Vektorraum. \blacklozenge

BEISPIEL 4.5: Auch die Menge *aller* Polynome,

$$\mathcal{P} := \bigcup_{m=0}^{\infty} \mathcal{P}_m \quad (4.4)$$

(die Vereinigung aller Mengen \mathcal{P}_m) ist zusammen mit der durch (4.3a) und (4.3b) definierten Addition und skalaren Multiplikation ein Vektorraum. (Bei der Addition muss man m gleich dem grösseren der zwei Grade von p und q wählen.) Zu \mathcal{P} gehören Polynome beliebig hohen Grades, jedoch keine nichtabbrechenden Potenzreihen.

Polynome könnte man natürlich auch miteinander multiplizieren, aber diese Operation lassen wir hier ausser acht. Lässt man sie anstelle der skalaren Multiplikation zu, so bilden die Polynome einen Ring mit Eins (die Konstante 1), vgl. Seite 2-9. Anders als der Ring der $n \times n$ Matrizen ist dieser jetzt aber kommutativ (weil die Polynom-Multiplikation kommutativ ist: $p(t)q(t) = q(t)p(t)$) und hat keine Nullteiler. Er hat damit die gleichen Grundeigenschaften wie der Ring der ganzen Zahlen. \blacklozenge

BEISPIEL 4.6: Der Raum ℓ der **reellen Zahlfolgen** $\{x_k\}_{k=0}^{\infty}$ mit gliedweiser Addition

$$\{x_k\} + \{y_k\} := \{x_k + y_k\} \quad (4.5a)$$

und gliedweiser skalarer Multiplikation

$$\alpha \{x_k\} := \{\alpha x_k\}. \quad (4.5b)$$

\blacklozenge

Zu jedem dieser Beispiele reeller Vektorräume gibt es einen analog definierten komplexen Vektorraum. Zum Beispiel analog zu $C[a, b]$ den Raum der auf $[a, b]$ stetigen komplexwertigen Funktionen.

Es sei nochmals betont, dass man für jedes Beispiel verifizieren müsste, dass die Vektorraum-Axiome erfüllt sind.

Aus den Vektorraum-Axiomen lassen sich eine Reihe weiterer einfacher Regeln ableiten:

Satz 4.1 *Es sei V ein Vektorraum über \mathbb{E} . Für alle $x, y \in V$ und alle $\alpha \in \mathbb{E}$ gilt:*

$$0x = o, \quad (4.6)$$

$$\alpha o = o, \quad (4.7)$$

$$\alpha x = o \implies \alpha = 0 \text{ oder } x = o, \quad (4.8)$$

$$(-\alpha)x = \alpha(-x) = -(\alpha x). \quad (4.9)$$

Obwohl diese Aussagen im Hinblick auf den Spezialfall $V = \mathbb{R}^n$ trivial aussehen, sind ihre Beweise zum Teil etwa mühselig. Als Beispiel eines solchen Beweises geben wir jenen für (4.6) an.

BEWEIS von (4.6): Es ist nach (V6) zunächst

$$0x = (0 + 0)x = 0x + 0x.$$

Nach (V4) gibt es zu $y := 0x$ einen Vektor $-y$ mit $0x + (-y) = o$. Unter Verwendung von (V2) und (V3) folgt weiter, dass

$$\begin{aligned} o &= 0x + (-y) = (0x + 0x) + (-y) \\ &\stackrel{(V2)}{=} 0x + (0x + (-y)) = 0x + o \stackrel{(V3)}{=} 0x. \end{aligned}$$

■

Ähnlich beweist man

Satz 4.2 *Zu $x, y \in V$ existiert $z \in V$ mit $x + z = y$, wobei z eindeutig bestimmt ist und gilt: $z = y + (-x)$.*

Die erste Aussage (Existenz) kann an die Stelle der Axiome (V3) und (V4) treten. Der Satz liefert auch eine Definition der Subtraktion in V :

DEFINITION: In einem Vektorraum V ist die **Subtraktion** [*subtraction*] zweier Vektoren $x, y \in V$ definiert durch

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

▲

AUSBLICK: Allgemeiner kann man in der Definition des Vektorraumes anstelle von \mathbb{R} oder \mathbb{C} einen beliebigen **Skalarenkörper** [*field of scalars*] \mathbb{K} zulassen und Vektorräume über \mathbb{K} definieren. Ein Körper ist eine weitere durch Axiome definierte mathematische Struktur. Es ist ein kommutativer Ring mit Eins und ohne Nullteiler, indem auch die Division definiert ist, solange der Divisor (Nenner) nicht das neutrale Element der Addition (die “Null”) ist.

Die reellen und komplexen Zahlen, \mathbb{R} und \mathbb{C} , mit der üblichen Addition und Multiplikation sind die bekanntesten Beispiele. Obwohl wir im folgenden solche allgemeinere Vektorräume nicht betrachten, fügen wir hier die Definition eines Körpers an:

DEFINITION: Ein **Körper** [*field*] ist eine nichtleere Menge \mathbb{K} auf der eine Addition

$$\alpha, \beta \in \mathbb{K} \quad \mapsto \quad \alpha + \beta \in \mathbb{K}$$

und eine Multiplikation

$$\alpha, \beta \in \mathbb{K} \quad \mapsto \quad \alpha \cdot \beta \in \mathbb{K}$$

definiert sind, wobei folgende Grundregeln gelten:

$$(K1) \quad \alpha + \beta = \beta + \alpha \quad (\forall \alpha, \beta \in \mathbb{K}),$$

$$(K2) \quad (\alpha + \beta) + \gamma = \alpha + (\beta + \gamma) \quad (\forall \alpha, \beta, \gamma \in \mathbb{K}),$$

$$(K3) \quad \text{es gibt ein ausgezeichnetes Element } 0 \in \mathbb{K} \text{ mit} \\ \alpha + 0 = \alpha \quad (\forall \alpha \in \mathbb{K}),$$

$$(K4) \quad \text{zu jedem } \alpha \in \mathbb{K} \text{ gibt es ein eindeutig bestimmtes } -\alpha \in \mathbb{K} \text{ mit} \\ \alpha + (-\alpha) = 0,$$

$$(K5) \quad \alpha \cdot \beta = \beta \cdot \alpha \quad (\forall \alpha, \beta \in \mathbb{K}),$$

$$(K6) \quad (\alpha \cdot \beta) \cdot \gamma = \alpha \cdot (\beta \cdot \gamma) \quad (\forall \alpha, \beta, \gamma \in \mathbb{K}),$$

$$(K7) \quad \text{es gibt ein ausgezeichnetes Element } 1 \in \mathbb{K}, 1 \neq 0, \text{ mit} \\ \alpha \cdot 1 = \alpha \quad (\forall \alpha \in \mathbb{K}),$$

$$(K8) \quad \text{zu jedem } \alpha \in \mathbb{K}, \alpha \neq 0 \text{ gibt es ein eindeutig bestimmtes} \\ \alpha^{-1} \in \mathbb{K} \text{ mit } \alpha \cdot (\alpha^{-1}) = 1,$$

$$(K9) \quad \alpha \cdot (\beta + \gamma) = \alpha \cdot \beta + \alpha \cdot \gamma \quad (\forall \alpha, \beta, \gamma \in \mathbb{K}),$$

$$(K10) \quad (\alpha + \beta) \cdot \gamma = \alpha \cdot \gamma + \beta \cdot \gamma \quad (\forall \alpha, \beta, \gamma \in \mathbb{K}).$$



Die Axiome (K1)–(K4) bedeuten, dass \mathbb{K} bezüglich der Addition eine kommutative Gruppe ist, und wegen den Axiomen (K5)–(K8) ist $\mathbb{K} \setminus \{0\}$ (d.h. \mathbb{K} ohne die Null) auch eine kommutative Gruppe bezüglich der Multiplikation. Schliesslich sind die Axiome (K9) und (K10) die üblichen Distributivgesetze. Den Multiplikationspunkt lässt man oft weg.

Axiom (K8) hat zur Folge, dass man in einem Körper eine Division definieren kann:

$$\alpha/\beta \quad :\equiv \quad \alpha \cdot \beta^{-1}. \quad (4.10)$$

Lässt man dieses Axiom weg, so hat man anstelle eines Körpers einen kommutativen Ring mit Eins.

4.2 Unterräume, Erzeugendensysteme

DEFINITION: Eine nichtleere Teilmenge U eines Vektorraums V heisst **Unterraum** [*subspace*], falls sie bezüglich Addition und skalarer Multiplikation abgeschlossen ist, d.h. falls

$$\boxed{x + y \in U, \quad \alpha x \in U \quad (\forall x, y \in U, \forall \alpha \in \mathbb{E}).} \quad (4.11)$$

In einem Unterraum $U \subseteq V$ sind Addition und Multiplikation gleich definiert wie in V . Zudem gelten die Axiome (V1), (V2), (V5)–(V8), weil sie in V gelten. Schliesslich ist für jedes $x \in U$ auch $0x \in U$ und $(-1)x \in U$, wobei wir nach Satz 4.1 und (V8) wissen, dass

$$0x = o, \quad (-1)x = -(1x) = -x.$$

Also gelten (V3) und (V4) auch, und wir erhalten die folgende Aussage.

Satz 4.3 *Ein Unterraum ist selbst ein Vektorraum.*

BEISPIEL 4.7: Fasst man die Polynome als auf \mathbb{R} definierte Funktionen auf, und ist $0 < m < n$, so gilt

$$\mathcal{P}_0 \subset \mathcal{P}_m \subset \mathcal{P}_n \subset \mathcal{P} \subset C^\infty(\mathbb{R}) \subset C^n(\mathbb{R}) \subset C^m(\mathbb{R}) \subset C(\mathbb{R}), \quad (4.12)$$

wobei nun jeder Raum Unterraum der rechts davon stehenden Vektorräume ist. \blacklozenge

BEISPIEL 4.8: Jeder Vektorraum V ist ein Unterraum von sich selbst. Zudem hat jeder Vektorraum als Unterraum den Raum $\{o\}$, der nur den Nullvektor enthält. Dieser Nullraum ist selbst auch ein Vektorraum, der alle Axiome erfüllt. \blacklozenge

BEISPIEL 4.9: Im geometrisch interpretierten Vektorraum \mathbb{R}^3 der Ortsvektoren ist jede Ebene durch den Ursprung O und jede Gerade durch O ein Unterraum. \blacklozenge

Ein weiteres, besonders wichtiges Beispiel formulieren wir als Satz:

Satz 4.4 *Ist $\mathbf{A} \in \mathbb{R}^{m \times n}$ eine reelle $m \times n$ -Matrix und \mathcal{L}_0 die Menge der Lösungsvektoren $\mathbf{x} \in \mathbb{R}^n$ des homogenen Gleichungssystems*

$$\mathbf{Ax} = \mathbf{o},$$

so ist \mathcal{L}_0 ein Unterraum von \mathbb{R}^n .

BEWEIS: Sind $\mathbf{x}, \mathbf{y} \in \mathcal{L}_0$, d.h. gilt $\mathbf{Ax} = \mathbf{o}$ und $\mathbf{Ay} = \mathbf{o}$, so folgt

$$\mathbf{A}(\mathbf{x} + \mathbf{y}) = \mathbf{Ax} + \mathbf{Ay} = \mathbf{o},$$

und für alle $\alpha \in \mathbb{R}$

$$\mathbf{A}(\alpha \mathbf{x}) = \alpha(\mathbf{Ax}) = \mathbf{o},$$

d.h. es gilt $\mathbf{x} + \mathbf{y} \in \mathcal{L}_0$, $\alpha \mathbf{x} \in \mathcal{L}_0$. \blacksquare

Dieser Satz ist eine Verallgemeinerung von Beispiel 4.9: Der Lösungsraum von $\mathbf{Ax} = \mathbf{o}$ ist eine ‐Ebene‐ (ein Unterraum) im \mathbb{R}^n durch den Ursprung.

Als nächstes übertragen wir den Begriff der Linearkombination von \mathbb{E}^n (siehe Kapitel 2, Seite 2-10) auf beliebige Vektorräume.

DEFINITION: Es seien V ein Vektorraum über \mathbb{E} , und $a_1, \dots, a_\ell \in V$ ausgewählte Vektoren. Ein Vektor der Form

$$x := \gamma_1 a_1 + \dots + \gamma_\ell a_\ell = \sum_{k=1}^{\ell} \gamma_k a_k \quad (4.13)$$

mit $\gamma_1, \dots, \gamma_\ell \in \mathbb{E}$ heisst eine **Linearkombination** [*linear combination*] von a_1, \dots, a_ℓ . ▲

Die Gesamtheit aller Linearkombinationen von a_1, \dots, a_ℓ ist offensichtlich ein Unterraum, denn (4.11) ist für die Vektoren der Form (4.13) mit festen Vektoren a_1, \dots, a_ℓ und freien Skalaren $\gamma_1, \dots, \gamma_\ell$ erfüllt.

DEFINITION: Die Menge aller Linearkombinationen von a_1, \dots, a_ℓ heisst der **von a_1, \dots, a_ℓ aufgespannte** (oder: **erzeugte**) **Unterraum** [*subspace spanned by a_1, \dots, a_ℓ ; span*] oder die **lineare Hülle von a_1, \dots, a_ℓ** [*linear hull*]. Er wird bezeichnet mit:¹

$$\text{span} \{a_1, \dots, a_\ell\} := \left\{ \sum_{k=1}^{\ell} \gamma_k a_k ; \gamma_1, \dots, \gamma_\ell \in \mathbb{E} \right\}. \quad (4.14)$$

Der von einer unendlichen Folge oder Menge $S \subset V$ erzeugte Unterraum ist gleich der Gesamtheit aller Linearkombinationen endlich² vieler Vektoren aus S :

$$\text{span } S := \left\{ \sum_{k=1}^m \gamma_k a_k ; m \in \mathbb{N}; a_1, \dots, a_m \in S; \gamma_1, \dots, \gamma_m \in \mathbb{E} \right\}. \quad (4.15)$$

Die Vektoren a_1, \dots, a_ℓ in (4.14) bzw. die Menge S in (4.15) heissen **Erzeugendensystem** [*spanning set*] von $\text{span} \{a_1, \dots, a_\ell\}$ bzw. $\text{span } S$. ▲

BEISPIEL 4.10: Die vier Vektoren

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{a}_3 = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{a}_4 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$$

spannen offenbar im 3-dimensionalen Raum die Koordinatenebene E_{12} auf, die alle Vektoren enthält, deren dritte Koordinate null ist. Sie sind

¹Oft schreibt man auch $\langle a_1, \dots, a_\ell \rangle$ für diesen Unterraum, was wir aber wegen des Konfliktes mit unserer Notation für das Skalarprodukt vermeiden.

²Man betrachtet anderswo in der Mathematik allerdings auch Vektorräume und Unterräume, die von Linearkombinationen unendlich vieler Vektoren erzeugt werden; siehe den Ausblick am Schluss von Abschnitt 6.3.

also eine Erzeugendensystem für E_{12} . Offensichtlich würden dazu schon die Vektoren $\mathbf{a}_2 = \mathbf{e}_2$ und $\mathbf{a}_3 = 2\mathbf{e}_1$ reichen; diese sind auch ein Erzeugendensystem für E_{12} . In der Tat spannen irgend zwei der vier Vektoren E_{12} auf. Definiert man die Matrix $\mathbf{A} = (\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3 \ \mathbf{a}_4)$, so ist

$$\text{span} \{a_1, a_2, a_3, a_4\} = \{\mathbf{y} = \mathbf{A} \mathbf{x}; \mathbf{x} \in \mathbb{R}^4\} \subset \mathbb{R}^3. \quad \blacklozenge$$

BEISPIEL 4.11: Die $m + 1$ **Monome** [*monomials*] $1, t, t^2, \dots, t^m$ sind ein Erzeugendensystem des Raumes \mathcal{P}_m aller Polynome vom Grade $\leq m$. Offensichtlich reichen weniger als $m + 1$ Polynome nicht. \blacklozenge

BEISPIEL 4.12: Die (unendlich vielen) Monome $1, t, t^2, \dots$ bilden ein Erzeugendensystem des Raumes \mathcal{P} aller Polynome aus Beispiel 4.5. Jedes Polynom hat ja einen endlichen Grad m und lässt sich deshalb als Linearkombination von $1, t, t^2, \dots, t^m$ schreiben. \blacklozenge

4.3 Lineare Abhängigkeit, Basen, Dimension

Die vorangehenden Beispiele lassen vermuten, dass sich gewisse Erzeugendensysteme dadurch auszeichnen, dass sie aus einer minimalen Anzahl von Vektoren bestehen. In grösseren Erzeugendensystemen lassen sich ein Teil der gegebenen Vektoren als Linearkombinationen der anderen Vektoren darstellen.

DEFINITION: Die Vektoren $a_1, \dots, a_\ell \in V$ heissen **linear abhängig** [*linearly dependent*], falls es Skalare $\gamma_1, \dots, \gamma_\ell$ gibt, die nicht alle 0 sind und für die gilt

$$\gamma_1 a_1 + \dots + \gamma_\ell a_\ell = \mathbf{o}. \quad (4.16)$$

Andernfalls, d.h. wenn (4.16) impliziert dass $\gamma_1 = \dots = \gamma_\ell = 0$, sind die Vektoren **linear unabhängig** [*linearly independent*]. \blacktriangle

Lemma 4.5 Die $\ell \geq 2$ Vektoren a_1, \dots, a_ℓ sind linear abhängig genau dann, wenn sich einer dieser Vektoren als Linearkombination der andern schreiben lässt.

BEWEIS: Sind die Vektoren a_1, \dots, a_ℓ linear abhängig, gilt (4.16) mit $\gamma_k \neq 0$ für ein $k, 1 \leq k \leq \ell$; folglich ist

$$a_k = - \sum_{\substack{j=1 \\ j \neq k}}^{\ell} \frac{\gamma_j}{\gamma_k} a_j.$$

Die Umkehrung ist trivial. \blacksquare

BEISPIEL 4.13: Die vier Vektoren von Beispiel 4.10 sind linear abhängig, denn es ist

$$2\mathbf{a}_1 - 4\mathbf{a}_2 - \mathbf{a}_3 + 0\mathbf{a}_4 = \mathbf{o} \quad \text{bzw.} \quad \mathbf{a}_3 = 2\mathbf{a}_1 - 4\mathbf{a}_2. \quad \blacklozenge$$

BEISPIEL 4.14: Die drei Funktionen $t \mapsto 1, \cos^2 t, \cos 2t \in C^\infty(\mathbb{R})$ sind linear abhängig, denn

$$\cos 2t = 2 \cos^2 t - 1. \quad \blacklozenge$$

BEISPIEL 4.15: Sowohl als formale Polynome als auch als reelle Funktionen sind die $m + 1$ Monome $t \mapsto 1, t, t^2, \dots, t^m \in \mathcal{P}_m \subset C^\infty(\mathbb{R})$ linear unabhängig, denn

$$\sum_{k=0}^m \gamma_k t^k = 0$$

impliziert $\gamma_0 = \gamma_1 = \dots = \gamma_m = 0$. Die Summe ist ja selbst ein Polynom vom Grad $\leq m$, muss also das Nullpolynom sein. Dieses kann man aber nicht anders als Linearkombination der Monome $1, t, t^2, \dots, t^m$ darstellen, als wenn alle Koeffizienten null sind. \blacklozenge

BEISPIEL 4.16: Sind die drei Polynome

$$p_1(t) := 1 + t^2, \quad p_2(t) := 1 - t^2, \quad p_3(t) := 1 + t^2 + t^4 \quad (4.17)$$

wohl linear unabhängig? Aus der Bedingung, dass für alle t gilt

$$\begin{aligned} 0 &= \gamma_1 p_1(t) + \gamma_2 p_2(t) + \gamma_3 p_3(t) \\ &= \gamma_1(1 + t^2) + \gamma_2(1 - t^2) + \gamma_3(1 + t^2 + t^4) \\ &= (\gamma_1 + \gamma_2 + \gamma_3) + (\gamma_1 - \gamma_2 + \gamma_3)t^2 + \gamma_3 t^4, \end{aligned}$$

folgt durch Koeffizientenvergleich (wir wissen aus Beispiel 4.15 ja, dass die Monome $1, t^2, t^4$ linear unabhängig sind):

$$\begin{aligned} \gamma_1 + \gamma_2 + \gamma_3 &= 0 \\ \gamma_1 - \gamma_2 + \gamma_3 &= 0 \\ \gamma_3 &= 0. \end{aligned} \quad (4.18)$$

Dies ist ein lineares Gleichungssystem. Ein Eliminationsschritt führt bereits auf die LR-Zerlegung

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (4.19)$$

aus der man schliesst, dass die Matrix des 3×3 -Systems (4.18) regulär ist. Also hat das homogene System nur die triviale Lösung; das heisst $\gamma_1 = \gamma_2 = \gamma_3 = 0$, was bedeutet, dass die Polynome p_1, p_2, p_3 linear unabhängig sind. \blacklozenge

BEISPIEL 4.17: Sind $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{R}$ paarweise voneinander verschieden, so sind die Funktionen

$$t \mapsto e^{\alpha_1 t}, e^{\alpha_2 t}, \dots, e^{\alpha_n t} \in C^\infty(\mathbb{R}) \quad (4.20)$$

linear unabhängig.

BEWEIS: Wir dürfen annehmen, die α_k seien der Grösse nach geordnet: $\alpha_1 < \alpha_2 < \dots < \alpha_n$. Wären die Funktionen linear abhängig, so gäbe es einen grössten Index k mit der Eigenschaft, dass $e^{\alpha_k t}$ eine Linearkombination der übrigen Funktionen ist:

$$e^{\alpha_k t} = \sum_{j=1}^{k-1} \gamma_j e^{\alpha_j t}.$$

Es folgt, dass

$$f(t) := \sum_{j=1}^{k-1} \gamma_j e^{(\alpha_j - \alpha_k)t} \equiv 1.$$

Wegen $\alpha_j - \alpha_k < 0$ ($j = 1, \dots, k-1$) ist aber $\lim_{t \rightarrow \infty} f(t) = 0$ im Widerspruch zu $f(t) \equiv 1$. ■ ◆

Die Verallgemeinerung der linearen Unabhängigkeit auf unendliche Mengen von Vektoren ist heikel. In der linearen Algebra geht man üblicherweise zunächst anders vor als in späteren Anwendungen. Wir kommen in Abschnitt 6.3 auf diesen Punkt zurück.

DEFINITION: Eine unendliche Menge von Vektoren heisst **linear unabhängig** [*linearly independent*], wenn jede endliche Teilmenge linear unabhängig ist. Andernfalls ist die Menge **linear abhängig** [*linearly dependent*]. ▲

BEISPIEL 4.18: Ist $\{\alpha_k\}_{k=0}^{\infty}$ eine Folge verschiedener reeller Zahlen, so sind die unendlich vielen Funktionen

$$t \mapsto e^{\alpha_k t} \quad (k = 1, 2, \dots)$$

gemäss Beispiel 4.17 linear unabhängig.

Wir könnten hier sogar von einer beliebigen Menge verschiedener reeller Zahlen ausgehen, zum Beispiel sogar von ganz \mathbb{R} . Wir könnten auch beliebige komplexe Exponenten zulassen, müssten dann aber den Beweis anpassen. Es gibt hier offenbar eine enorm grosse Zahl linear unabhängiger Funktionen. ◆

Es ist nun naheliegend, unter den Erzeugendensystemen für einen Vektorraum solche auszuwählen, die möglichst klein sind. Umgekehrt möchte man Mengen von linear unabhängigen Vektoren so gross wählen oder so ergänzen, dass sie den ganzen Raum erzeugen. Beides führt auf den folgenden Begriff.

DEFINITION: Ein linear unabhängiges Erzeugendensystem eines Vektorraums V heisst **Basis** [*basis*] von V . ▲

BEISPIEL 4.19: Die n Einheitsvektoren im \mathbb{E}^n ,

$$\mathbf{e}_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \mathbf{e}_n := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad (4.21)$$

(das sind die n Kolonnen der Einheitsmatrix \mathbf{I}_n) spannen den Raum \mathbb{E}^n auf, sind also ein Erzeugendensystem von \mathbb{E}^n . Offensichtlich reichen weniger als n Vektoren dazu nicht. Diese n Vektoren bilden also eine Basis des \mathbb{E}^n (d.h. des \mathbb{R}^n und des \mathbb{C}^n). Wir nennen sie die **Standardbasis** [*standard basis*]. \blacklozenge

BEISPIEL 4.20: Was gibt es für andere Basen in \mathbb{E}^n ?

Die n Kolonnen $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n \in \mathbb{E}^n$ einer $n \times n$ Matrix \mathbf{B} bilden genau dann eine Basis von \mathbb{E}^n , wenn \mathbf{B} regulär ist.

In der Tat gilt nach Korollar 1.7: Ist \mathbf{B} regulär, hat $\mathbf{B}\mathbf{x} = \mathbf{c}$ für jedes $\mathbf{c} \in \mathbb{E}^n$ eine Lösung und für $\mathbf{c} = \mathbf{o}$ nur die Nulllösung. Das heisst gerade, dass die Kolonnen von \mathbf{B} den Raum \mathbb{E}^n erzeugen und linear unabhängig sind.

Ist \mathbf{B} dagegen singulär, so sind seine Kolonnen weder erzeugend noch linear unabhängig.

Sind zum Beispiel $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n \in \mathbb{E}^n$ die n Kolonnen einer oberen Dreiecksmatrix,

$$\mathbf{b}_1 := \begin{pmatrix} b_{11} \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \mathbf{b}_2 := \begin{pmatrix} b_{12} \\ b_{22} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad \mathbf{b}_n := \begin{pmatrix} b_{1n} \\ b_{2n} \\ b_{3n} \\ \vdots \\ b_{nn} \end{pmatrix}$$

und sind die Diagonalelemente $b_{11}, b_{22}, \dots, b_{nn}$ alle nicht 0, so bilden diese Kolonnen von \mathbf{B} eine Basis von \mathbb{E}^n . \blacklozenge

BEISPIEL 4.21: Wir wissen bereits, dass die drei Polynome p_1, p_2, p_3 aus (4.17) linear unabhängig sind. Sie liegen alle in \mathcal{P}_4 und sind alles gerade Funktionen (d.h. $p_k(-t) = p_k(t)$). Sicher bilden sie nicht eine Basis von \mathcal{P}_4 , denn die Polynome t und t^3 gehören nicht zum von p_1, p_2, p_3 erzeugten Unterraum. Man könnte aber vermuten, dass dieser Unterraum gleich dem Unterraum \mathcal{G}_4 aller gerader Polynome vom Höchstgrade 4 ist, das heisst, dass gilt³

$$\mathcal{G}_4 := \text{span} \{1, t^2, t^4\} = \text{span} \{p_1, p_2, p_3\}. \quad (4.22)$$

Dazu müssten wir zeigen, dass ein beliebiges Polynom der Form

$$g(t) = a + bt^2 + ct^4$$

als Linearkombination von p_1, p_2, p_3 dargestellt werden kann. Analog zur Rechnung in Beispiel 4.16 müsste nun die Bedingung

$$\begin{aligned} a + bt^2 + ct^4 &= \gamma_1 p_1(t) + \gamma_2 p_2(t) + \gamma_3 p_3(t) \\ &= (\gamma_1 + \gamma_2 + \gamma_3) + (\gamma_1 - \gamma_2 + \gamma_3)t^2 + \gamma_3 t^4 \end{aligned}$$

gelten, die durch Koeffizientenvergleich das inhomogene System

³Wir sind hier und anderswo etwas schludrig in der Notation und verwenden Bezeichnungen wie $p(t), t^n$ auch für die Funktionen, nicht nur für die Funktionswerte.

$$\begin{aligned}
 \gamma_1 + \gamma_2 + \gamma_3 &= a \\
 \gamma_1 - \gamma_2 + \gamma_3 &= b \\
 \gamma_3 &= c
 \end{aligned}
 \tag{4.23}$$

liefert. Da das homogene System (4.18) nur die triviale Lösung hat, wissen wir aus Korollar 1.6, dass obiges System für beliebige rechte Seiten $\begin{pmatrix} a & b & c \end{pmatrix}^\top$ lösbar ist. Also ist $\{p_1, p_2, p_3\}$ ein Erzeugendensystem für \mathcal{G}_4 , d.h. (4.22) ist richtig.

Wir können dies hier dank Korollar 1.6 schliessen, ohne die Koeffizienten γ_k explizit als Funktion von a, b, c zu bestimmen. Allerdings wäre auch das einfach, denn mittels der LR-Zerlegung (4.19) kann man das System (4.23) mit a, b, c als Parameter lösen. \blacklozenge

BEISPIEL 4.22: Die abzählbar unendlich vielen Monome

$$t \mapsto 1, t, t^2, \dots$$

bilden eine Basis (die Standardbasis) des Raumes \mathcal{P} aller Polynome, vgl. (4.4). Wir sahen ja bereits, dass endlich viele Monome linear unabhängig sind. Nach unserer Definition ist also auch die unendliche Menge aller Monome linear unabhängig. Zudem ist jedes Polynom eine endliche Linearkombination von Monomen. \blacklozenge

BEISPIEL 4.23: Die unendlich vielen Zahlfolgen

$$\begin{aligned}
 e_1 &:= (1 \ 0 \ 0 \ \dots), \\
 e_2 &:= (0 \ 1 \ 0 \ \dots), \\
 &\dots
 \end{aligned}$$

bilden die Standardbasis des Vektorraumes ℓ_0 der abbrechenden (reellen oder komplexen) Zahlfolgen. Denn offensichtlich kann man jedes Element von ℓ_0 als Linearkombination endlich vieler dieser Vektoren darstellen, und andererseits sind endlich viele dieser Vektoren auch immer linear unabhängig. \blacklozenge

VORAUSSETZUNG: Im folgenden beschränken wir uns auf Vektorräume, die ein endliches Erzeugendensystem besitzen.

Für diese Räume sind einige grundlegende Fakten leicht herzuleiten:

Lemma 4.6 *Gibt es zu einem (vom Nullraum verschiedenen) Vektorraum ein endliches Erzeugendensystem, so besitzt er eine Basis, die eine Teilmenge des Erzeugendensystems ist.*

BEWEIS: Ist das Erzeugendensystem nicht ohnehin linear unabhängig und enthält es mehr als einen Vektor, so lässt sich gemäss Lemma 4.5 einer der Vektoren durch die anderen ausdrücken. Man kann also diesen streichen; die restlichen bilden immer noch ein Erzeugendensystem. Durch Wiederholung dieses Vorgehens kann man das Erzeugendensystem schrittweise auf eines reduzieren, das nur noch aus $n \geq 1$ linear unabhängigen Vektoren besteht. Dieses ist per Definition eine Basis. \blacksquare

Wir zeigen als nächstes, dass die Zahl n der Basisvektoren eine charakteristische Grösse des Vektorraumes ist.

Satz 4.7 *Besitzt ein Vektorraum V ein endliches Erzeugendensystem, so besteht jede Basis von V aus derselben Zahl von Vektoren.*

DEFINITION: Die Zahl der Basisvektoren (in jeder Basis) eines Vektorraumes V mit endlichem Erzeugendensystem heisst **Dimension** [*dimension*] von V und wird mit **dim** V bezeichnet. Ein solcher Raum ist also **endlich-dimensional** [*finite dimensional*]. Falls $\dim V = n$ gilt, so sagt man, V sei **n -dimensional** [*n -dimensional*]. ▲

Bevor wir Satz 4.7 beweisen, betrachten wir einfache Beispiele:

BEISPIEL 4.24: \mathcal{P}_m hat die Standardbasis bestehend aus den $m + 1$ Monomen $t \mapsto t^k$ ($0 \leq k \leq m$). Es ist also $\dim \mathcal{P}_m = m + 1$. ◆

BEISPIEL 4.25: Der Raum \mathbb{E}^n mit der in Beispiel 4.19 gegebenen Standardbasis $\mathbf{e}_1, \dots, \mathbf{e}_n$ hat natürlich die Dimension n , und zwar ist \mathbb{R}^n ein n -dimensionaler reeller Vektorraum, \mathbb{C}^n ein n -dimensionaler komplexer Vektorraum.

Fasst man \mathbb{C}^n als reellen Vektorraum auf (d.h. lässt man nur reelle Skalare zu), so hat \mathbb{C}^n die Dimension $2n$. Eine einfache Basis ist

$$\underbrace{\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}}_n, \quad \underbrace{\begin{pmatrix} i \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ i \end{pmatrix}}_n.$$

Offensichtlich lässt sich jeder komplexe n -Vektor als Linearkombination dieser $2n$ Vektoren schreiben, auch wenn nur reelle Koeffizienten zugelassen sind; die Vektoren erzeugen also \mathbb{C}^n . Zudem sind sie offensichtlich linear unabhängig. ◆

Nun kommen wir zum Beweis von Satz 4.7. Er folgt sofort aus dem Beweis von Lemma 4.6 und aus dem folgenden Lemma.

Lemma 4.8 *Ist $\{b_1, \dots, b_m\}$ ein Erzeugendensystem von V , so ist jede Menge $\{a_1, \dots, a_\ell\} \subset V$ von $\ell > m$ Vektoren linear abhängig.*

BEWEIS: Die a_k lassen sich als Linearkombinationen von b_1, \dots, b_m darstellen:

$$a_k = \sum_{i=1}^m \tau_{ik} b_i \quad (1 \leq k \leq \ell). \quad (4.24)$$

Wir betrachten nun das homogene lineare Gleichungssystem

$$\sum_{k=1}^{\ell} \tau_{ik} \xi_k = 0 \quad (1 \leq i \leq m) \quad (4.25)$$

bestehend aus m Gleichungen in ℓ ($> m$) Unbekannten ξ_k . Nach Korollar 1.5 besitzt es eine nichttriviale Lösung $(\xi_1 \ \dots \ \xi_\ell)^T \neq \mathbf{o} \in \mathbb{E}^\ell$.

Mit diesen ξ_k bilden wir die Linearkombination

$$a := \sum_{k=1}^{\ell} \xi_k a_k$$

und erhalten aufgrund von (4.24) und (4.25)

$$a = \sum_{k=1}^{\ell} \xi_k \left(\sum_{i=1}^m \tau_{ik} b_i \right) = \sum_{i=1}^m \left(\underbrace{\sum_{k=1}^{\ell} \tau_{ik} \xi_k}_{=0 \ (\forall i)} \right) b_i = 0,$$

was zeigt, dass a_1, \dots, a_{ℓ} linear abhängig sind. ■

Betrachten wir noch den Fall $V = \mathbb{E}^n$, wo das Erzeugendensystem $\{\mathbf{b}_1, \dots, \mathbf{b}_m\}$ und die Menge $\{\mathbf{a}_1, \dots, \mathbf{a}_{\ell}\}$ (mit $\ell > m$) aus Spaltenvektoren bestehen und als Spalten von Matrizen aufgefasst werden können:

$$\mathbf{B} = \left(\mathbf{b}_1 \ \dots \ \mathbf{b}_m \right), \quad \mathbf{A} = \left(\mathbf{a}_1 \ \dots \ \mathbf{a}_{\ell} \right). \quad (4.26)$$

Mit der $m \times \ell$ Matrix $\mathbf{T} = \left(\tau_{ik} \right)$ kann man dann (4.24) als

$$\mathbf{A} = \mathbf{B}\mathbf{T} \quad (4.27)$$

schreiben. Da $m < \ell$ ist, hat das homogene System $\mathbf{T}\mathbf{x} = \mathbf{o}$ eine nichttriviale Lösung, und für diese gilt

$$\mathbf{A}\mathbf{x} = \mathbf{B}\mathbf{T}\mathbf{x} = \mathbf{B}\mathbf{o} = \mathbf{o}. \quad (4.28)$$

Dies bedeutet gerade, dass die Spalten $\{\mathbf{a}_1, \dots, \mathbf{a}_{\ell}\}$ von \mathbf{A} linear abhängig sind.

Bei unserer Konstruktion einer Basis in Lemma 4.6 sind wir von einem Erzeugendensystem ausgegangen und haben dieses schrittweise verkleinert. Man kann auch eine linear unabhängige Menge schrittweise zu einer Basis vergrößern:

Satz 4.9 *Jede Menge linear unabhängiger Vektoren aus einem Vektorraum V mit endlichem Erzeugendensystem lässt sich ergänzen zu einer Basis von V .*

BEWEIS: Gegeben ist eine linear unabhängige Menge M von Vektoren und ein Erzeugendensystem E . Ist $\text{span } M \neq V$, so gibt es in E einen Vektor, der nicht als Linearkombination von M dargestellt werden kann. Diesen schlagen wir zusätzlich zu M , wodurch $\dim \text{span } M$ um eins zunimmt, aber M linear unabhängig bleibt. Wiederholt man dieses Vorgehen, so erreicht man schrittweise, dass $\dim \text{span } M = \dim \text{span } E = \dim V$ wird. Dann ist M eine Basis. ■

Da gemäss Satz 4.7 jede Basis eines endlichdimensionalen Vektorraumes V aus gleich vielen, nämlich $n = \dim V$ Vektoren besteht, können wir aus Satz 4.9 sofort folgendes schliessen:

Korollar 4.10 *In einem Vektorraum V endlicher Dimension ist jede Menge von $n = \dim V$ linear unabhängigen Vektoren eine Basis von V .*

BEWEIS: Wäre eine solche Menge keine Basis, so liesse sie sich nach Satz 4.9 zu einer ergänzen. Diese würde aber mehr als $\dim V$ Vektoren enthalten, was Satz 4.7 widersprechen würde. ■

BEISPIEL 4.26: Dass die Polynome p_1, p_2, p_3 aus den Beispielen 4.16 und 4.21 eine Basis des Raumes \mathcal{G}_4 der geraden Polynome vom Grade ≤ 4 bilden, können wir nun einfacher einsehen als in Beispiel 4.21. Da die drei Monome $1, t^2, t^4$ eine Basis von \mathcal{G}_4 bilden, ist $\dim \mathcal{G}_4 = 3$. In Beispiel 4.16 haben wir bewiesen, dass p_1, p_2, p_3 linear unabhängig sind; also bilden sie auch eine Basis.

Man kann nun sowohl $1, t^2, t^4$ als auch p_1, p_2, p_3 zu einer Basis von \mathcal{P}_4 ergänzen. Beispielsweise sind die zwei Polynome t und t^3 eine geeignete Ergänzung der beiden Basen, aber wir könnten auch etwa das Paar $t + p_1(t)$ und $t + t^3 + p_3(t)$ verwenden. Die Wahl der Monome scheint attraktiver, weil diese den Unterraum \mathcal{U}_3 der ungeraden Polynome vom Höchstgrade 3 aufspannen, aber als Basen von \mathcal{P}_4 sind $1, t, t^2, t^3, t^4$ und $p_1(t), p_2(t), p_3(t), t + p_1(t), t + t^3 + p_3(t)$ theoretisch gleichwertig. (Beim numerischen Rechnen wird man aber gewisse Basen bevorzugen. Man muss vor allem vermeiden, dass nahezu linear abhängige Basen gewählt werden.) ◆

Man kann zeigen (mit Hilfe des Zornschen Lemmas, einer tiefschürfenden mathematischen Aussage), dass jeder Vektorraum eine Basis hat, also auch jene, die kein endliches Erzeugendensystem haben. Es gilt sogar die folgende Verallgemeinerung von Satz 4.9:

Satz 4.11 *Jede Menge linear unabhängiger Vektoren aus einem Vektorraum V lässt sich ergänzen zu einer Basis von V .*

Für die Räume ℓ_0 und \mathcal{P} aus den Beispielen 4.23 und 4.22 ist das kaum überraschend. Gemäss unseren Definitionen bilden die Vektoren $\mathbf{e}_1, \mathbf{e}_2, \dots$ aber nicht eine Basis des Vektorraumes ℓ aller Zahlfolgen. Das würde erfordern, dass man unendliche Linearkombinationen zulassen würde. Der Raum ℓ hat keine “einfache” Basis im Sinne unserer Definition. Eine Vektorraumbasis von ℓ ist notwendigerweise “sehr gross”. Wir können keine solche Basis angeben, obwohl wir nach obigem Satz wissen, dass welche existieren. Das Gleiche gilt zum Beispiel für $C[0, 1]$. Wir werden in Kapitel 6 auf diese Problematik zurück kommen und einen Ausblick auf die Antworten geben, die die Mathematik bereit hält.

Die Nützlichkeit der Basen beruht auf der nächsten Aussage:

Satz 4.12 *$\{b_1, \dots, b_n\} \subset V$ ist genau dann eine Basis von V , wenn sich jeder Vektor $x \in V$ in eindeutiger Weise als Linearkombination von b_1, \dots, b_n darstellen lässt:*

$$x = \sum_{k=1}^n \xi_k b_k. \quad (4.29)$$

DEFINITION: Die Koeffizienten ξ_k in (4.29) heissen **Koordinaten** [coordinates] von x bezüglich der Basis $\{b_1, \dots, b_n\}$. Der Vektor $\boldsymbol{\xi} = (\xi_1 \ \dots \ \xi_n)^T \in \mathbb{E}^n$ ist der **Koordinatenvektor** [coordina-

te vector]; (4.29) heisst **Koordinatendarstellung** [representation in coordinates] von x . ▲

BEWEIS von Satz 4.12:

(i) Es sei $\{b_1, \dots, b_n\}$ eine Basis von V , und es sei $x \in V$ beliebig. Da $\{b_1, \dots, b_n\}$ ein Erzeugendensystem ist, lässt sich x in der Form (4.29) darstellen. Gäbe es eine zweite Darstellung

$$x = \sum_{k=1}^n \xi'_k b_k$$

von x , so bekäme man durch Subtraktion

$$\sum_{k=1}^n (\xi_k - \xi'_k) b_k = x - x = o,$$

woraus wegen der linearen Unabhängigkeit von $\{b_1, \dots, b_n\}$ folgen würde, dass $\xi_k - \xi'_k = 0$, d.h. $\xi_k = \xi'_k$ ($1 \leq k \leq n$).

(ii) Es sei jedes $x \in V$ auf eindeutige Weise in der Form (4.29) darstellbar. Dann ist $\{b_1, \dots, b_n\}$ ein Erzeugendensystem und, weil $o \in V$ die eindeutige Darstellung

$$o = 0b_1 + \dots + 0b_n$$

hat, sind b_1, \dots, b_n linear unabhängig. ■

BEISPIELE 4.27: Offensichtlich gilt für einen beliebigen n -Vektor $\mathbf{x} = (x_1 \ \dots \ x_n)^\top \in \mathbb{E}^n$

$$\mathbf{x} = \sum_{k=1}^n x_k \mathbf{e}_k. \quad (4.30)$$

Ebenso lässt sich jedes Polynom p vom Grade m per Definition darstellen in der Form

$$p(t) = a_0 + a_1 t + \dots + a_m t^m = \sum_{k=0}^m a_k t^k, \quad (4.31)$$

das heisst als Linearkombination der Standardbasis $1, t, \dots, t^m$.

Um in den Räumen \mathbb{E}^n und \mathcal{P}_m zur Darstellung eines Vektors in einer anderen Basis zu kommen, wird man in der Regel diesen Vektor und die andere Basis in der Standardbasis darstellen und dann einen Koeffizientenvergleich durchführen.

So gelang es uns etwa in Beispiel 4.21 das Polynom g , das allgemeine gerade Polynom vom Grad 4, in der Basis p_1, p_2, p_3 darzustellen; in einem konkreten Beispiel wäre noch das Gleichungssystem (4.23) zu lösen. Im Raum \mathbb{E}^n kommen wir zu diesem Koeffizientenvergleich automatisch, sobald wir die Komponenten der Vektoren betrachten. ◆

Die Sätze 4.9 und 4.12 geben Anlass zu einer Aufspaltung eines Vektorraumes in zwei sich ergänzende Unterräume. Es sei $M = \{b_1, \dots, b_l\}$ eine Menge linear unabhängiger Vektoren, die den Raum V nicht erzeugen. Sie bilden eine Basis des echten Unterraumes

$$U := \text{span } M = \text{span } \{b_1, \dots, b_l\}. \quad (4.32)$$

Nach Satz 4.9 kann man M zu einer Basis von V ergänzen. Es sei $N = \{b_{l+1}, \dots, b_n\}$ die Menge zusätzlicher Basisvektoren und

$$U' := \text{span } N = \text{span } \{b_{l+1}, \dots, b_n\} \quad (4.33)$$

der von ihnen aufgespannte Unterraum. Damit lässt sich die Koordinatendarstellung eines beliebigen $x \in V$ eindeutig aufteilen in

$$x = \underbrace{\sum_{k=1}^l \xi_k b_k}_{u \in U} + \underbrace{\sum_{k=l+1}^n \xi_k b_k}_{u' \in U'}. \quad (4.34)$$

DEFINITION: Zwei Unterräume U und U' eines Vektorraumes V mit der Eigenschaft, dass jedes $x \in V$ eine eindeutige Darstellung

$$x = u + u' \quad \text{mit } u \in U, \quad u' \in U' \quad (4.35)$$

hat, heißen **komplementär** [*complementary*]. Wir nennen dann V die **direkte Summe** [*direct sum*] von U und U' und schreiben

$$V = U \oplus U' \quad (4.36)$$



BEISPIEL 4.28: In der Situation von Beispiel 4.26 ist klar, dass man \mathcal{P}_4 als direkte Summe der geraden Polynome vom Grad 4 und der ungeraden Polynome vom Grad 3 schreiben kann:

$$\mathcal{P}_4 = \mathcal{G}_4 \oplus \mathcal{U}_3. \quad (4.37)$$

Aber es ist auch

$$\mathcal{P}_4 = \text{span } \{p_1, p_2, p_3\} \oplus \mathcal{U}_3$$

oder

$$\mathcal{P}_4 = \text{span } \{p_1, p_2, p_3\} \oplus \text{span } \{t + p_1(t), t + t^3 + p_3(t)\}.$$



Allgemeiner kann man einen Vektorraum als direkte Summe mehrerer Unterräume darstellen.

BEISPIEL 4.29: Wir können die Beispiele 4.16, 4.21, 4.26 und 4.28 fortsetzen: Da p_1, p_2, p_3 linear unabhängig sind, ist zum Beispiel

$$\mathcal{G}_4 = \text{span } \{p_1\} \oplus \text{span } \{p_2\} \oplus \text{span } \{p_3\}$$

eine direkte Summe und ebenso

$$\mathcal{P}_4 = \text{span } \{p_1\} \oplus \text{span } \{p_2\} \oplus \text{span } \{p_3\} \oplus \mathcal{U}_3.$$

Hier könnte man \mathcal{U}_3 auch noch als eine direkte Summe von zwei eindimensionalen Unterräumen darstellen. ◆

4.4 Basiswechsel, Koordinatentransformation

Es sei $\{b_1, \dots, b_n\}$ eine vorgegebene, ‐alte‐ Basis des Vektorraumes V , und es sei $\{b'_1, \dots, b'_n\}$ eine zweite, ‐neue‐ Basis von V . Die neuen Basisvektoren kann man in der alten Basis darstellen:

$$b'_k = \sum_{i=1}^n \tau_{ik} b_i, \quad k = 1, \dots, n. \quad (4.38)$$

Die $n \times n$ -Matrix $\mathbf{T} = (\tau_{ik})$ heisst **Transformationsmatrix** [*transformation matrix*] des **Basiswechsels** [*change of basis*].

Beachte: In der k -ten Kolonne von \mathbf{T} stehen gerade die Koordinaten des k -ten neuen Basisvektors (bezüglich der alten Basis).

Solch ein Basiswechsel hat eine **Koordinatentransformation** [*coordinate transformation*] zur Folge, die sich so beschreiben lässt:

Satz 4.13 Es sei $\xi = (\xi_1 \ \dots \ \xi_n)^\top$ der Koordinatenvektor eines beliebigen Vektors $x \in V$ bezüglich der alten Basis, und es sei $\xi' = (\xi'_1 \ \dots \ \xi'_n)^\top$ der Koordinatenvektor dieses Vektors x bezüglich einer neuen, durch (4.38) gegebenen Basis, so dass

$$\sum_{i=1}^n \xi_i b_i = x = \sum_{k=1}^n \xi'_k b'_k. \quad (4.39)$$

Dann gilt für diese Koordinatentransformation:

$$\xi_i = \sum_{k=1}^n \tau_{ik} \xi'_k, \quad i = 1, \dots, n, \quad (4.40)$$

d.h. es ist

$$\xi = \mathbf{T} \xi', \quad (4.41)$$

wobei die Transformationsmatrix \mathbf{T} regulär ist, so dass also

$$\xi' = \mathbf{T}^{-1} \xi. \quad (4.42)$$

Beachte: In (4.38) wird die neue Basis in der alten dargestellt, aber in (4.40) werden die alten Koordinaten von x als Funktionen der neuen beschrieben.

BEWEIS von Satz 4.13: Nach (4.38) und (4.40) ist

$$x = \sum_{k=1}^n \xi'_k b'_k = \sum_{k=1}^n \xi'_k \left(\sum_{i=1}^n \tau_{ik} b_i \right) = \sum_{i=1}^n \underbrace{\left(\sum_{k=1}^n \tau_{ik} \xi'_k \right)}_{= \xi_i} b_i = \sum_{i=1}^n \xi_i b_i.$$

Weil x laut Satz 4.12 nur auf eine Art als Linearkombination der Basisvektoren b_i darstellbar ist, folgt (4.40).

Der Nullvektor o wird in den beiden Basen eindeutig dargestellt durch $\xi = \mathbf{o}$ bzw. $\xi' = \mathbf{o}$. Fasst man (4.41) als Gleichungssystem für ξ' auf,

so folgt aus $\xi = \mathbf{o}$ also $\xi' = \mathbf{o}$, d.h. das homogene System hat nur die triviale Lösung. Also ist \mathbf{T} regulär, und es gilt (4.42). ■

BEISPIEL 4.30: Der in Beispiel 4.21 betrachtete Raum \mathcal{G}_4 der geraden Polynome vom Grad ≤ 4 hat laut der dort bewiesenen Beziehung (4.22) die zwei Basen $\{1, t^2, t^4\}$ und $\{p_1, p_2, p_3\}$, wobei nach (4.17) gilt

$$p_1(t) := 1 + t^2, \quad p_2(t) := 1 - t^2, \quad p_3(t) := 1 + t^2 + t^4. \quad (4.43)$$

Wir können ein beliebiges Polynom $g \in \mathcal{G}_4$ je auf eindeutige Weise in den beiden Basen darstellen:

$$g(t) = a + bt^2 + ct^4 = a'p_1(t) + b'p_2(t) + c'p_3(t). \quad (4.44)$$

Die entsprechenden Koordinatenvektoren sind also

$$\xi = (a \quad b \quad c)^T, \quad \xi' = (a' \quad b' \quad c')^T.$$

Die neue Basis $\{p_1, p_2, p_3\}$ ist in (4.43) gerade durch die alte Basis $\{1, t^2, t^4\}$ dargestellt; die Transformationsmatrix ist

$$\mathbf{T} = (\tau_{ik}) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}. \quad (4.45)$$

Diese Matrix ist schon in (4.18) und (4.23) als Koeffizientenmatrix aufgetreten. Um nun zum Beispiel das in der alten Basis ausgedrückte Polynom

$$g(t) = 4 + 6t^2 + 8t^4 \quad (4.46)$$

in der neuen Basis darzustellen, müssen wir (4.42) anwenden mit $\xi = (4 \quad 6 \quad 8)^T$:

$$\xi' = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 4 \\ 6 \\ 8 \end{pmatrix}.$$

Wir haben in (4.18) bereits die LR-Zerlegung von \mathbf{T} angegeben. Was zu tun bleibt, ist vor- und rückwärts einzusetzen:

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 6 \\ 8 \end{pmatrix} \implies \omega = \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 8 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \xi'_1 \\ \xi'_2 \\ \xi'_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 8 \end{pmatrix} \implies \xi' = \begin{pmatrix} \xi'_1 \\ \xi'_2 \\ \xi'_3 \end{pmatrix} = \begin{pmatrix} -3 \\ -1 \\ 8 \end{pmatrix}. \quad \blacklozenge$$

Betrachten wir Basiswechsel und Koordinatentransformation nun noch im Spezialfall $V = \mathbb{E}^n$. Dann lassen sich die alte und die neue Basis ja als Kolonnen von $n \times n$ -Matrizen \mathbf{B} und \mathbf{B}' darstellen:

$$\mathbf{B} = (\mathbf{b}_1 \quad \cdots \quad \mathbf{b}_n), \quad \mathbf{B}' = (\mathbf{b}'_1 \quad \cdots \quad \mathbf{b}'_n). \quad (4.47)$$

Damit kann der Basiswechsel (4.38) einfach geschrieben werden als

$$\boxed{\mathbf{B}' = \mathbf{B} \mathbf{T}}. \quad (4.48)$$

Für die alte und die neue Koordinatendarstellung gelten dann nach (4.39)

$$\boxed{\mathbf{B} \boldsymbol{\xi} = \mathbf{x} = \mathbf{B}' \boldsymbol{\xi}' .} \quad (4.49)$$

Die Matrizen \mathbf{B} und \mathbf{B}' sind regulär, denn zum Nullvektor \mathbf{o} gehört in beiden Basen der eindeutige Koordinatenvektor $\boldsymbol{\xi} = \boldsymbol{\xi}' = \mathbf{o}$.

Aus (4.49) folgt durch Einsetzen von (4.48) sofort

$$\mathbf{B} \boldsymbol{\xi} = \mathbf{B} \mathbf{T} \boldsymbol{\xi}' ,$$

also, weil \mathbf{B} regulär ist,

$$\boxed{\boldsymbol{\xi} = \mathbf{T} \boldsymbol{\xi}'}$$

wie in (4.41).

BEISPIEL 4.31: Wir wollen im \mathbb{R}^3 neben der Standardbasis $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ noch die “schiefe” Basis $\{\mathbf{b}'_1, \mathbf{b}'_2, \mathbf{b}'_3\}$ mit

$$\mathbf{b}'_1 = \mathbf{e}_1, \quad \mathbf{b}'_2 = \mathbf{e}_1 + \mathbf{e}_2, \quad \mathbf{b}'_3 = \mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 \quad (4.50)$$

einführen. Es ist in diesem Falle wegen $\mathbf{B} = \mathbf{I}_3$

$$\mathbf{T} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{B}' . \quad (4.51)$$

Die Koordinaten transformieren sich also gemäss

$$\begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \xi'_1 \\ \xi'_2 \\ \xi'_3 \end{pmatrix} ,$$

und, wie man leicht verifiziert, ist

$$\begin{pmatrix} \xi'_1 \\ \xi'_2 \\ \xi'_3 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} .$$



Kapitel 5

Lineare Abbildungen

Zum Strukturbegriff “Vektorraum” gehört der Abbildungsbegriff “lineare Abbildung”. Eine Abbildung zwischen zwei Vektorräumen heisst linear, wenn sie “die linearen Beziehungen invariant (unverändert) lässt”. Wählt man in den Vektorräumen Basen, so induziert eine solche Abbildung eine lineare Abbildung der Koordinaten. Falls die Räume endlich-dimensional sind, wird diese durch eine Matrix repräsentiert. Dies ist, in Kürze, der grundlegende Zusammenhang zwischen linearen Abbildungen und Matrizen.

5.1 Definition, Beispiele, Matrixdarstellung

Zur Erinnerung ein paar Begriffe zum Thema “Abbildung” (oder “Funktion”), die nicht auf lineare Abbildungen beschränkt sind:

DEFINITION: Es sei $F : X \rightarrow Y$ irgend eine Abbildung. Die Menge $F(X)$ der Bildpunkte von F heisst **Wertebereich** [*range*] von F oder **Bild** [*image*] von F . Sie wird auch mit **im** F bezeichnet:

$$F(X) = \text{im } F := \{F(x) \in Y; x \in X\} \subseteq Y.$$

Ist $F(X) = Y$, so ist F eine Abbildung **auf** [*on*] Y und heisst auch **surjektiv** [*surjektive*]. Gilt

$$F(x) = F(x') \implies x = x',$$

so heisst F **eindeutig** [*one-to-one*] oder **injektiv** [*injektive*].

Eine Abbildungen, die surjektiv und injektiv ist, heisst “**eindeutig auf**” [*one-to-one onto*] oder **bijektiv** [*bijektive*].

Ist F bijektiv, so ist die **inverse Abbildung** [*inverse map, inverse mapping, inverse transformation*] F^{-1} definiert. ▲

Im allgemeinen ist jedoch F^{-1} nicht als Abbildung definiert. Der Ausdruck $F^{-1}(Z)$ bezeichnet dann die Menge der Urbilder von Elementen $z \in Z \subseteq Y$.

Von nun an wollen wir “lineare” Abbildungen betrachten. Bei solchen lässt man die Klammer um das Argument meist weg. Das heisst, man schreibt einfach FX statt $F(X)$ und Fx statt $F(x)$. Das geht aber nicht immer: bei $F(x+y)$ braucht es eine Klammer.

DEFINITION: Eine Abbildung $F : X \rightarrow Y$, $x \mapsto Fx$ zwischen zwei Vektorräumen X und Y (über \mathbb{E}) heisst **linear** [*linear transformation*], wenn für alle $x, \tilde{x} \in X$ und alle $\gamma \in \mathbb{E}$ gilt

$$F(x + \tilde{x}) = Fx + F\tilde{x}, \quad F(\gamma x) = \gamma(Fx). \quad (5.1)$$

X ist der **Definitionsraum** [*domain*], Y der **Bildraum** [*image space*] der Abbildung.

Ist der Bildraum Y der Skalarenkörper \mathbb{E} , so bezeichnet man eine lineare Abbildung als **lineares Funktional** [*linear functional*].

Sind Definitions- und Bildraum Funktionenräume, so spricht man statt von einer linearen Abbildung von einem **linearen Operator** [*linear operator*].

Ist der Definitionsraum gleich dem Bildraum, das heisst ist $X = Y$, so hat man eine **Selbstabbildung**. ▲

Die Eigenschaften (5.1) sind gleichwertig mit

$$F(\beta x + \gamma \tilde{x}) = \beta Fx + \gamma F\tilde{x} \quad (\forall x, \tilde{x} \in X, \forall \beta, \gamma \in \mathbb{E}). \quad (5.2)$$

Man beachte auch, dass die Bezeichnung “Bildraum” nicht unbedingt bedeutet, dass $Y = F(X)$ ist, das heisst, dass es zu jedem $y \in Y$ ein $x \in X$ mit $y = f(x)$ gibt.

BEISPIEL 5.1: Ist $\mathbf{A} \in \mathbb{E}^{m \times n}$ eine vorgegebene $m \times n$ -Matrix, so definiert die Zuordnung $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ eine lineare Abbildung des \mathbb{E}^n in den \mathbb{E}^m , die wir ebenfalls mit \mathbf{A} bezeichnen:

$$\mathbf{A} : \mathbb{E}^n \rightarrow \mathbb{E}^m, \quad \mathbf{x} \mapsto \mathbf{A}\mathbf{x}. \quad (5.3)$$

$\mathbf{y} \in \mathbb{E}^m$ ist genau dann Bild eines $\mathbf{x} \in \mathbb{E}^n$, wenn das Gleichungssystem $\mathbf{A}\mathbf{x} = \mathbf{y}$ eine Lösung \mathbf{x} hat. ◆

BEISPIEL 5.2: Der Übergang von einer auf $[a, b]$ stetig differenzierbaren Funktion f zu ihrer Ableitung f' definiert den **Ableitungsoperator**, eine lineare Abbildung von $C^1[a, b]$ in $C[a, b]$:

$$D : C^1[a, b] \rightarrow C[a, b], \quad f \mapsto f'. \quad (5.4)$$

Ihre Linearität folgt aus $(\alpha f + \beta g)' = \alpha f' + \beta g'$.

Allgemeiner ist, bei festen reellen Koeffizienten (oder festen stetigen reellen Funktionen) c_0, c_1, \dots, c_m , der **Differentialoperator**

$$L : C^m[a, b] \rightarrow C[a, b], \quad (5.5)$$

$$f \mapsto c_m f^{(m)} + c_{m-1} f^{(m-1)} + \dots + c_1 f' + c_0 f$$

ein linearer Operator, also eine lineare Abbildung. ◆

BEISPIEL 5.3: Bei festem $t \in [a, b]$ ist die **Evaluationsabbildung** definiert durch

$$E_t : C[a, b] \rightarrow \mathbb{R}, \quad f \mapsto E_t f := f(t) \quad (5.6)$$

wegen $(\alpha f + \beta g)(t) = \alpha f(t) + \beta g(t)$ ein lineares Funktional. Ist zum Beispiel

$$f(t) = \alpha + \beta t + \gamma t^2,$$

so gilt

$$E_0 f = f(0) = \alpha, \quad E_2 f = f(2) = \alpha + 2\beta + 4\gamma. \quad ◆$$

BEISPIEL 5.4: Auch der **Multiplikationsoperator**

$$M : C[a, b] \rightarrow C[a, b], \quad f \mapsto g, \quad \text{wo } g(t) := t f(t) \quad (5.7)$$

ist ein linearer Operator, denn

$$\begin{aligned} (M(\beta f_1 + \gamma f_2))(\tau) &= \beta t f_1(t) + \gamma t f_2(t) \Big|_{t=\tau} \\ &= \beta \tau f_1(\tau) + \gamma \tau f_2(\tau) \\ &= (\beta (M f_1) + \gamma (M f_2))(\tau). \end{aligned}$$

Die lineare Abbildung

$$F : C^1[a, b] \Rightarrow C[a, b], \quad f \mapsto g, \quad g(t) := t f'(t) + f(2) \quad (5.8)$$

kombiniert den Ableitungsoperator D , den Multiplikationsoperator M und die Evaluationsabbildung E_2 . Unter Verwendung des Verknüpfungssymbols \circ für Abbildungen könnte man schreiben

$$F = M \circ D + E_2. \quad (5.9)$$

Man kann diese Abbildungen natürlich analog für Funktionen definieren, die auf ganz \mathbb{R} gegeben sind. Man beachte, dass $C^1[0, 1]$ nicht etwa ein Unterraum von $C^1(\mathbb{R})$ ist.

Beschränkt man die Abbildungen auf Unterräume, kann man sich Gedanken über den Bildraum machen. Zum Beispiel gilt

$$E_2 : \mathcal{P}_m \rightarrow \mathbb{R} = \mathcal{P}_0, \quad (5.10)$$

$$M : \mathcal{P}_m \rightarrow \mathcal{P}_{m+1}, \quad (5.11)$$

$$D : \mathcal{P}_m \rightarrow \mathcal{P}_{m-1}, \quad (5.12)$$

$$F : \mathcal{P}_m \rightarrow \mathcal{P}_m, \quad (5.13)$$

wenn F wie in (5.9) definiert ist. ◆

Wir betrachten nun eine *beliebige* lineare Abbildung F eines n -dimensionalen Vektorraums X in einen m -dimensionalen Vektorraum Y . Es sei $\{b_1, \dots, b_n\}$ eine Basis von X , $\{c_1, \dots, c_m\}$ eine Basis von Y :

$$\boxed{X = \text{span} \{b_1, \dots, b_n\} \xrightarrow{F} Y = \text{span} \{c_1, \dots, c_m\}} \\ \text{(dim } X = n) \qquad \qquad \qquad \text{(dim } Y = m)$$

Die Bilder $F b_l \in Y$ der Basis von X lassen sich als Linearkombinationen der c_k darstellen:

$$\boxed{F b_l = \sum_{k=1}^m a_{kl} c_k} \quad (l = 1, \dots, n). \quad (5.14)$$

Die $m \times n$ -Matrix $\mathbf{A} = (a_{kl})$ heisst **Abbildungsmatrix von F bezüglich den gegebenen Basen in X und Y** [*matrix for F relative to the given bases in X and Y*].

Beachte: Die Abbildungsmatrix enthält in der l -ten Kolonne die Koordinaten (bezüglich der Basis von Y) des Bildes des l -ten Basisvektors (von X).

Umgekehrt gehört zu jeder Matrix $\mathbf{A} = (a_{kl}) \in \mathbb{E}^{m \times n}$ bei vorgegebenen Basen in X und Y eine durch (5.14) eindeutig bestimmte lineare Abbildung $F : X \rightarrow Y$. Durch (5.14) sind die Bilder der Basis von X bestimmt. Das Bild eines beliebigen Vektors $x \in X$ ergibt sich auf Grund der geforderten Linearität: Es wird ja

$$\boxed{x = \sum_{l=1}^n \xi_l b_l} \quad (5.15)$$

mit dem Koordinatenvektor $\boldsymbol{\xi} := (\xi_1 \ \dots \ \xi_n)^\top$ abgebildet auf

$$\begin{aligned} y := Fx &= F\left(\sum_{l=1}^n \xi_l b_l\right) = \sum_{l=1}^n \xi_l Fb_l \\ &= \sum_{l=1}^n \xi_l \sum_{k=1}^m a_{kl} c_k = \sum_{k=1}^m \underbrace{\left(\sum_{l=1}^n a_{kl} \xi_l\right)}_{\equiv: \eta_k} c_k, \end{aligned}$$

mit dem Koordinatenvektor $\boldsymbol{\eta} := (\eta_1 \ \dots \ \eta_m)^\top$:

$$\boxed{y = \sum_{k=1}^m \eta_k c_k.} \quad (5.16)$$

Das Bild $y = Fx$ hat also den Koordinatenvektor

$$\boxed{\begin{pmatrix} \eta_1 \\ \vdots \\ \eta_m \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix},} \quad (5.17)$$

das heisst

$$\boxed{\boldsymbol{\eta} = \mathbf{A}\boldsymbol{\xi}.} \quad (5.18)$$

In Worten: Man erhält den Koordinatenvektor $\boldsymbol{\eta}$ des Bildes $y = Fx$, indem man die Abbildungsmatrix \mathbf{A} auf den Koordinatenvektor $\boldsymbol{\xi}$ von x anwendet.

Im Falle einer Selbstabbildung, wo $X = Y$ und damit auch $m = n$ und $\{b_1, \dots, b_n\} = \{c_1, \dots, c_n\}$ gilt, ist die Abbildungsmatrix \mathbf{A} quadratisch und die Koordinatenvektoren $\boldsymbol{\xi}$ und $\boldsymbol{\eta}$ in (5.18) beziehen sich auf dieselbe Basis.

BEISPIEL 5.5: Die vier linearen Abbildungen (5.10)–(5.13), beschränkt auf den Raum \mathcal{P}_4 (d.h. $m = 4$), haben bezüglich der monomialen Basen $1, t, t^2, \dots, t^n$ in \mathcal{P}_n ($n = 0, 3, 4, 5$) folgende Abbildungsmatrizen:

$$\mathbf{E}_2 = \begin{pmatrix} 1 & 2 & 4 & 8 & 16 \end{pmatrix}, \quad \mathbf{M} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\mathbf{D} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} 1 & 2 & 4 & 8 & 16 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}.$$

◆

Auch hier lassen sich im Falle $X = \mathbb{E}^n$, $Y = \mathbb{E}^m$ die Zusammenhänge etwas eleganter darstellen. Die Basen schreiben wir wieder als Kolonnen von Matrizen \mathbf{B} und \mathbf{C} ,

$$\mathbf{B} = \begin{pmatrix} \mathbf{b}_1 & \dots & \mathbf{b}_n \end{pmatrix}, \quad \mathbf{C} = \begin{pmatrix} \mathbf{c}_1 & \dots & \mathbf{c}_m \end{pmatrix}, \quad (5.19)$$

und wir definieren das Bild $F\mathbf{B}$ von \mathbf{B} durch

$$F\mathbf{B} := \begin{pmatrix} F\mathbf{b}_1 & \dots & F\mathbf{b}_n \end{pmatrix}. \quad (5.20)$$

Dann schreibt sich (5.14)–(5.16) neu als

$$F\mathbf{B} = \mathbf{C}\mathbf{A}, \quad \mathbf{x} = \mathbf{B}\boldsymbol{\xi}, \quad \mathbf{y} = \mathbf{C}\boldsymbol{\eta}, \quad (5.21)$$

und es folgt

$$\mathbf{C}\boldsymbol{\eta} = \mathbf{y} = F\mathbf{x} = F\mathbf{B}\boldsymbol{\xi} = \mathbf{C}\mathbf{A}\boldsymbol{\xi}. \quad (5.22)$$

Da \mathbf{C} linear unabhängige Kolonnen hat, ist also notwendigerweise $\boldsymbol{\eta} = \mathbf{A}\boldsymbol{\xi}$, wie in (5.18).

Nun zurück zur allgemeinen Theorie.

DEFINITION: Eine eindeutige lineare Abbildung von X auf Y heisst **Isomorphismus** [*isomorphism*]. Ist $X = Y$, so heisst sie **Automorphismus** [*automorphism*]. ▲

Lemma 5.1 *Ist $F : X \rightarrow Y$ ein Isomorphismus, so ist die inverse Abbildung $F^{-1} : Y \rightarrow X$ linear und auch ein Isomorphismus.*

BEWEIS: Es seien $y, w \in Y$, $x := F^{-1}y$, $z := F^{-1}w$ und $\beta, \gamma \in \mathbb{E}$. Dann schliessen wir aus der Linearität von f , dass

$$\begin{aligned} \beta F^{-1}y + \gamma F^{-1}w &= F^{-1} F(\beta F^{-1}y + \gamma F^{-1}w) \\ &= F^{-1}(\beta F F^{-1}y + \gamma F F^{-1}w) \\ &\stackrel{\substack{\uparrow \\ F \text{ linear}}}{=} F^{-1}(\beta y + \gamma w), \end{aligned}$$

woraus folgt, dass F^{-1} linear ist. Dass mit F auch F^{-1} bijektiv ist, ist allgemein gültig und klar. ■

Ist $\{b_1, \dots, b_n\}$ eine feste Basis von X , so wird gemäss Satz 4.12 jedem $x \in X$ ein eindeutiger Koordinatenvektor $\xi \in \mathbb{E}^n$ zugeordnet und umgekehrt. Die Abbildung

$$\kappa_X : X \rightarrow \mathbb{E}^n, \quad x \mapsto \xi \quad (5.23)$$

ist also bijektiv. Sie ist zudem linear, wie man sofort sieht, also ein Isomorphismus. Wir nennen sie **Koordinatenabbildung** [*coordinate mapping*].

Betrachten wir nochmals die der Beziehung (5.18) zu Grunde liegende Situation einer linearen Abbildung $F : X \Rightarrow Y$ und deren Darstellung in festen Basen von X und Y durch die Abbildungsmatrix \mathbf{A} . Es seien κ_X und κ_Y die Koordinatenabbildungen in X bzw. Y . Dann gilt das folgende **kommutative Diagramm** [*commutative diagram*]

$$\begin{array}{ccc} X & \xrightarrow{F} & Y \\ \kappa_X \downarrow \uparrow \kappa_X^{-1} & & \kappa_Y \downarrow \uparrow \kappa_Y^{-1} \\ \mathbb{E}^n & \xrightarrow{\mathbf{A}} & \mathbb{E}^m \end{array} \quad (5.24)$$

Es bedeutet, dass

$$\boxed{F = \kappa_Y^{-1} \mathbf{A} \kappa_X, \quad \mathbf{A} = \kappa_Y F \kappa_X^{-1}.} \quad (5.25)$$

Dieses Diagramm illustriert auch den Tatbestand, dass F und \mathbf{A} äquivalente Darstellungen der linearen Abbildung sind, wobei allerdings die Matrix \mathbf{A} abhängig von den gewählten Basen in X and Y ist. Die Eigenschaften von F kann man meist auch in \mathbf{A} finden.

Wir können das Diagramm ergänzen durch die Angabe typischer Elemente der auftretenden Räume:

$$\begin{array}{ccc} x \in X & \xrightarrow{F} & y \in Y \\ \kappa_X \downarrow \uparrow \kappa_X^{-1} & & \kappa_Y \downarrow \uparrow \kappa_Y^{-1} \\ \xi \in \mathbb{E}^n & \xrightarrow{\mathbf{A}} & \eta \in \mathbb{E}^m \end{array} \quad (5.26)$$

Im Falle einer bijektiven Abbildung, also eines Isomorphismus, können wir aus Lemma 5.1 und dem Diagramm leicht schliessen:

Korollar 5.2 *Wird der Isomorphismus F bezüglich festen Basen durch die Matrix \mathbf{A} dargestellt, so ist \mathbf{A} regulär und die Umkehrabbildung F^{-1} wird durch \mathbf{A}^{-1} dargestellt.*

BEWEIS: Die gemäss Lemma 5.1 existierende Umkehrabbildung F^{-1} werde durch die Matrix \mathbf{B} dargestellt. Dann ist nach dem Diagramm (5.24) analog zu (5.25) $\mathbf{B} = \kappa_X F^{-1} \kappa_Y^{-1}$, also

$$\mathbf{A}\mathbf{B} = \kappa_Y F \kappa_X^{-1} \kappa_X F^{-1} \kappa_Y^{-1} = \mathbf{I}.$$

Also ist nach Satz 2.17 \mathbf{A} invertierbar und $\mathbf{A}^{-1} = \mathbf{B}$. ■

Liegt die Bildmenge einer Abbildung F im Definitionsbereich einer zweiten Abbildung G , so kann man diese bekanntlich zusammensetzen zur Komposition $G \circ F$. Wir schreiben oft einfach GF . Zur Illustration dieser Situation kann man das Diagramm 5.26 ergänzen:

$$\begin{array}{ccccc}
 & & G \circ F & & \\
 x \in X & \xrightarrow{F} & y \in Y & \xrightarrow{G} & z \in Z \\
 & & & & \\
 \kappa_X \updownarrow & & \kappa_Y \updownarrow & & \kappa_Z \updownarrow \\
 & & & & \\
 \xi \in \mathbb{E}^n & \xrightarrow{\mathbf{A}} & \eta \in \mathbb{E}^m & \xrightarrow{\mathbf{B}} & \zeta \in \mathbb{E}^p \\
 & & \mathbf{B}\mathbf{A} & &
 \end{array} \tag{5.27}$$

Dabei gilt:

Lemma 5.3 Sind X, Y, Z Vektorräume über \mathbb{E} und $F : X \rightarrow Y$, $G : Y \rightarrow Z$ lineare Abbildungen, so ist auch $G \circ F : X \rightarrow Z$ linear. Sind \mathbf{A} und \mathbf{B} die Abbildungsmatrizen zu F und G bezüglich festen Basen in X, Y, Z , so hat $G \circ F$ die Abbildungsmatrix $\mathbf{B}\mathbf{A}$.

BEWEIS: Die Verifikation des ersten Teils ist einfach. Beim Beweis des zweiten Teils kann man sich wieder auf (5.25) stützen. ■

5.2 Kern, Bild und Rang

In diesem Abschnitt sei stets $F : X \rightarrow Y$ eine lineare Abbildung. Dabei sei $\dim X = n$, $\dim Y = m$.

DEFINITION: Der **Kern** [*kernel*] $\ker F$ von F ist das Urbild von $o \in Y$.

$$\ker F := \{x \in X; Fx = o\} \subseteq X, \quad \blacktriangle$$

Lemma 5.4 $\ker F$ ist ein Unterraum von X , und im F ist ein Unterraum von Y .

BEWEIS: (i) Es seien $x, y \in \ker F$ und $\beta, \gamma \in \mathbb{E}$. Dann ist

$$F(\beta x + \gamma y) = \beta Fx + \gamma Fy = \beta o + \gamma o = o,$$

also $\beta x + \gamma y \in \ker F$.

(ii) Es seien $z = Fx \in \text{im } F$ und $w = Fy \in \text{im } F$ sowie $\beta, \gamma \in \mathbb{E}$. Dann gilt

$$\beta z + \gamma w = \beta Fx + \gamma Fy = F(\beta x + \gamma y) \in \text{im } Y.$$

■

Auf ähnliche Weise zeigt man, dass allgemeiner folgendes gilt:

Lemma 5.5 *Ist U ein Unterraum von X , so ist dessen Bild FU ein Unterraum von Y . Ist W ein Unterraum von $\text{im } F$, so ist dessen Urbild $F^{-1}W$ ein Unterraum von X .*

BEMERKUNG: Wie auf Seite 5-1 erwähnt, ist das Urbild $F^{-1}W$ definiert durch

$$F^{-1}W := \{x \in X ; Fx \in W\}.$$

F^{-1} ist dabei nicht notwendigerweise eine (eindeutige) Abbildung; für festes x kann die Menge $F^{-1}x$ aus mehr als einem Element bestehen, insbesondere hier auch aus unendlich vielen. ▼

BEISPIEL 5.6: Für die durch eine Matrix \mathbf{A} definierte lineare Abbildung, für den Ableitungsoperator D und den Differentialoperator L aus Beispiel 5.2, für die Evaluationsabbildung E_t aus Beispiel 5.3, für den Multiplikationsoperator M aus Beispiel 5.4, sowie für die Koordinatenabbildung κ_X ergeben sich für Kern und Bild die folgenden Unterräume:

$\ker \mathbf{A}$ = Lösungsmenge des homogenen linearen Gleichungssystems $\mathbf{Ax} = \mathbf{o}$;

$\text{im } \mathbf{A}$ = Menge der rechten Seiten \mathbf{b} , für die $\mathbf{Ax} = \mathbf{b}$ eine Lösung hat;

$\ker D$ = Menge der Konstanten $c \in C^1[a, b] \cong \mathbb{R}$;

$\text{im } D = C[a, b]$;

$\ker L$ = Menge der Lösungen $f \in C^m[a, b]$ der homogenen linearen Differentialgleichung $Lf = o$, d.h.

$$c_m f^{(m)} + \dots + c_1 f' + c_0 f = o;$$

$\text{im } L$ = Menge der Funktionen $g \in C[a, b]$, für die die inhomogene Differentialgleichung $Lf = g$, d.h.

$$c_m f^{(m)} + \dots + c_1 f' + c_0 f = g,$$

eine Lösung $f \in C^m[a, b]$ besitzt;

$\ker E_t = \{f \in C[a, b]; f(t) = 0\}$;

$\text{im } E_t = \mathbb{R}$;

$\ker M = \{o \in C[a, b]\} = \text{Nullfunktion}$;

$\text{im } M = \{g \in C[a, b]; f : \tau \mapsto g(\tau)/\tau \in C[a, b]\}$;

$\ker \kappa_X = \{o\} \subset X$;

$\text{im } \kappa_X = \mathbb{E}^n$. ◆

Lineare Abbildungen sind auch insofern speziell als man am Kern erkennen kann, ob sie eineindeutig sind:

Satz 5.6 F ist genau dann injektiv, wenn $\ker F = \{o\}$ ist.

BEWEIS: Ist F injektiv, so hat $o \in Y$ ein einziges Urbild, nämlich $o \in X$. Ist umgekehrt $\ker F = \{o\}$ und $x \neq y$, d.h. $x - y \neq o$, so folgt $F(x - y) \neq o$, also, weil F linear ist, $Fx \neq Fy$. ■

Wir wollen uns nun den Dimensionen der Unterräume $\ker F \subseteq X$ und $\operatorname{im} F \subseteq Y$ zuwenden und setzen dabei $\dim X < \infty$ voraus. Es besteht der folgende grundlegende Zusammenhang:

Satz 5.7 Es gilt die Dimensionsformel

$$\dim X - \dim \ker F = \dim \operatorname{im} F. \quad (5.28)$$

BEWEIS: Es sei $\{c_1, \dots, c_r\}$ eine Basis von $\operatorname{im} F$. Wir bezeichnen mit b_1, \dots, b_r einen Satz von Urbildern von c_1, \dots, c_r . Weiter sei $\{b_{r+1}, \dots, b_{r+k}\}$ eine Basis von $\ker F$. Wir zeigen, dass $\{b_1, \dots, b_r, b_{r+1}, \dots, b_{r+k}\}$ eine Basis von X ist, dass also $r + k = \dim X$ gilt.

(i) Wir zeigen zuerst, dass $\{b_1, \dots, b_{r+k}\}$ ganz X erzeugt. Zu $x \in X$ gibt es $\gamma_1, \dots, \gamma_r$ mit

$$Fx = \gamma_1 c_1 + \dots + \gamma_r c_r = F(\gamma_1 b_1 + \dots + \gamma_r b_r).$$

Also ist $x - \gamma_1 b_1 - \dots - \gamma_r b_r \in \ker F$, d.h. es gibt $\gamma_{r+1}, \dots, \gamma_{r+k}$ mit

$$x - \gamma_1 b_1 - \dots - \gamma_r b_r = \gamma_{r+1} b_{r+1} + \dots + \gamma_{r+k} b_{r+k},$$

d.h. $x \in \operatorname{span} \{b_1, \dots, b_{r+k}\}$.

(ii) Weiter zeigen wir, dass b_1, \dots, b_{r+k} linear unabhängig sind. Aus

$$\sum_{j=1}^{r+k} \gamma_j b_j = o$$

folgt, dass

$$\begin{aligned} \gamma_1 c_1 + \dots + \gamma_r c_r &= F(\gamma_1 b_1 + \dots + \gamma_r b_r) \\ &= F(\underbrace{-\gamma_{r+1} b_{r+1} - \dots - \gamma_{r+k} b_{r+k}}_{\in \ker F}) = o. \end{aligned}$$

Also ist $\gamma_1 = \dots = \gamma_r = 0$, denn c_1, \dots, c_r sind linear unabhängig. Unter Verwendung der Voraussetzung folgt weiter, dass

$$\sum_{j=r+1}^{r+k} \gamma_j b_j = \sum_{j=1}^{r+k} \gamma_j b_j = o,$$

woraus sich wegen der linearen Unabhängigkeit von b_{r+1}, \dots, b_{r+k} ergibt, dass $\gamma_{r+1} = \dots = \gamma_{r+k} = 0$. ■

Die Dimensionsformel gibt eine reiche Ernte. Wir bleiben zunächst bei den abstrakten linearen Abbildungen. Eine Reihe von Folgerungen in der Matrixtheorie werden wir im nächsten Abschnitt besprechen. Zunächst definieren wir auch für lineare Abbildungen einen Rang.

DEFINITION: Der **Rang** [*rank*] einer linearen Abbildung F ist gleich der Dimension des Bildes von F :

$$\boxed{\text{Rang } F \equiv \dim \text{im } F.}$$



Damit kann man die Dimensionsformel auch so schreiben:

$$\boxed{\dim X - \dim \ker F = \text{Rang } F.} \tag{5.29}$$

Korollar 5.8 *Es gelten die folgenden drei Äquivalenzen:*

- (i) $F : X \rightarrow Y$ *injektiv* \iff $\text{Rang } F = \dim X$
- (ii) $F : X \rightarrow Y$ *bijektiv, d.h. Isomorphismus* \iff $\text{Rang } F = \dim X = \dim Y$
- (iii) $F : X \rightarrow X$ *bijektiv, d.h. Automorphismus* \iff $\text{Rang } F = \dim X$
 \iff $\ker F = o$

BEWEIS: Aussage (i) folgt aus (5.29) und Satz 5.6.

Um (ii) zu beweisen, benützt man zusätzlich Lemma 5.4 und die Tatsache, dass jeder echte Unterraum von Y (“echt” heisst $\neq Y$) geringere Dimension als Y hat. Hieraus folgt dann: $\text{im } F = Y \iff \text{Rang } F = \dim Y$.

(iii) ergibt sich aus (ii) als Spezialfall und aus (5.29). ■

DEFINITION: Zwei Vektorräume X and Y heissen **isomorph** [*isomorphic*], falls es einen Isomorphismus $F : X \rightarrow Y$ gibt. ▲

Satz 5.9 *Zwei Vektorräume endlicher Dimension sind genau dann isomorph, wenn sie die gleiche Dimension haben.*

BEWEIS: (i) Sind X und Y isomorph, so gibt es einen Isomorphismus $F : X \rightarrow Y$, und nach Korollar 5.8 (ii) folgt $\dim X = \dim Y$.

(ii) Ist $\dim X = \dim Y = n$, und sind $\{b_1, \dots, b_n\}$ und $\{c_1, \dots, c_n\}$ Basen von X und Y , so ist durch die Forderung

$$Fb_k = c_k \quad (k = 1, \dots, n)$$

eine lineare Abbildung F eindeutig definiert (vgl. (5.14), wo in diesem Falle $\mathbf{A} = \mathbf{I}$ ist). Dabei ist natürlich

$$\text{im } F = \text{span} \{c_1, \dots, c_n\} = Y.$$

Zudem gilt $\ker F = \{o\}$, denn ist $x = \sum \xi_k b_k$ und $Fx = o$, also

$$o = Fx = \sum_{k=1}^n \xi_k Fb_k = \sum_{k=1}^n \xi_k c_k,$$

so folgt wegen der linearen Unabhängigkeit der Basisvektoren c_k , dass $\xi_1 = \dots = \xi_n = 0$. ■

Korollar 5.10 Sind $F : X \rightarrow Y$, $G : Y \rightarrow Z$ lineare Abbildungen (wobei $\dim X, \dim Y < \infty$), so gilt:

- i) $\text{Rang } FG \leq \min\{\text{Rang } F, \text{Rang } G\}$,
- ii) G injektiv $\implies \text{Rang } GF = \text{Rang } F$,
- iii) F surjektiv $\implies \text{Rang } GF = \text{Rang } G$.

BEWEIS: (i) Die Dimensionsformel (5.29) zeigt, dass $\text{Rang } G = \dim \text{im } G \leq \dim Y$ ist, und dass analog, bei Anwendung auf die Restriktion $G|_{\text{im } F}$ von G auf $\text{im } F$, das heisst für

$$\begin{aligned} G|_{\text{im } F} : \underbrace{\text{im } F}_{\equiv: \tilde{Y}} \subseteq Y &\implies \underbrace{\text{im } GF}_{= G(\tilde{Y})} \subseteq Z \end{aligned}$$

folgt:

$$\text{Rang } GF = \dim \underbrace{\text{im } GF}_{= G(\tilde{Y})} \leq \dim \underbrace{\text{im } F}_{= \tilde{Y}} = \text{Rang } F. \quad (5.30)$$

Ferner ist natürlich

$$\text{Rang } GF = \dim \text{im } GF \leq \dim \text{im } G = \text{Rang } G. \quad (5.31)$$

(ii) Ist G injektiv, so ist $G|_{\text{im } F}$ auch injektiv, und nach (5.28) gilt (5.30) mit dem Gleichheitszeichen.

(iii) Ist F surjektiv, so ist $\text{im } GF = \text{im } G$; folglich gilt (5.31) mit dem Gleichheitszeichen. ■

5.3 Matrizen als lineare Abbildungen

Es sei $\mathbf{A} = (a_{kl})$ eine $m \times n$ -Matrix. Wir bezeichnen ihre n Kolonnen wieder mit $\mathbf{a}_1, \dots, \mathbf{a}_n$, so dass

$$\mathbf{A} = (\mathbf{a}_1 \quad \mathbf{a}_2 \quad \dots \quad \mathbf{a}_n). \quad (5.32)$$

DEFINITION: Der von den Kolonnen von \mathbf{A} aufgespannte Unterraum $\mathcal{R}(\mathbf{A}) := \text{span} \{ \mathbf{a}_1, \dots, \mathbf{a}_n \}$ heisst **Kolonnenraum** [column space] oder **Wertebereich** [range] von \mathbf{A} . Der Lösungsraum \mathcal{L}_0 des homogenen Systems $\mathbf{A}\mathbf{x} = \mathbf{o}$ heisst **Nullraum** [null space] $\mathcal{N}(\mathbf{A})$. ▲

Fasst man \mathbf{A} als lineare Abbildung $\mathbf{A} : \mathbb{E}^n \rightarrow \mathbb{E}^m$, $\mathbf{x} \mapsto \mathbf{A}\mathbf{x}$ auf, so ist $\ker \mathbf{A} = \mathcal{N}(\mathbf{A})$, und es gilt mit $\mathbf{x} \equiv: (x_1 \quad \dots \quad x_n)^\top$

$$\begin{aligned} \text{im } \mathbf{A} &= \{ \mathbf{A}\mathbf{x}; \mathbf{x} \in \mathbb{E}^n \} = \left\{ \sum_{k=1}^n \mathbf{a}_k x_k; x_1, \dots, x_n \in \mathbb{E} \right\} \\ &= \text{span} \{ \mathbf{a}_1, \dots, \mathbf{a}_n \} = \mathcal{R}(\mathbf{A}). \end{aligned}$$

Wie wir bereits in Beispiel 5.6 betont haben, ist auch im \mathbf{A} mit dem Gleichungssystem $\mathbf{Ax} = \mathbf{b}$ verknüpft. Zusammengefasst gilt:

Satz 5.11 *Fasst man die Matrix \mathbf{A} als lineare Abbildung auf, so ist das Bild von \mathbf{A} gleich dem Kolonnenraum oder Wertebereich von \mathbf{A} , und der Kern von \mathbf{A} ist gleich dem Nullraum von \mathbf{A} :*

$$\text{im } \mathbf{A} = \mathcal{R}(\mathbf{A}), \quad \ker \mathbf{A} = \mathcal{N}(\mathbf{A}). \quad (5.33)$$

Das Gleichungssystem $\mathbf{Ax} = \mathbf{b}$ ist genau dann lösbar, wenn \mathbf{b} im Kolonnenraum von \mathbf{A} liegt. Eine allfällige Lösung ist genau dann eindeutig, wenn $\mathcal{N}(\mathbf{A}) = \{\mathbf{o}\}$ ist, das heisst das homogene System nur die triviale Lösung hat.

Die Lösungsmenge \mathcal{L}_0 von $\mathbf{Ax} = \mathbf{o}$ haben wir bereits in Kapitel 1 studiert, siehe insbesondere Korollar 1.5. Im homogenen Fall gilt im Endschema des Gauss-Algorithmus $c_1 = \dots = c_m = 0$, das heisst die Verträglichkeitsbedingungen sind immer erfüllt, und die allgemeine Lösung hat $n - r$ frei wählbare Parameter. Nehmen wir der Einfachheit halber an, dass die ersten r Variablen x_1, \dots, x_r gerade die Pivotvariablen sind, d.h. dass \mathbf{A} auf eine Matrix \mathbf{R} mit der speziellen Zeilenstufenform

$$\mathbf{R} := \begin{pmatrix} r_{11} & * & * & \cdots & \cdots & \cdots & * \\ 0 & r_{22} & * & \cdots & \cdots & \cdots & * \\ \vdots & \ddots & \ddots & & & & \vdots \\ \vdots & & \ddots & r_{rr} & * & \cdots & * \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & \cdots & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix} \leftarrow r \quad (5.34)$$

$\begin{matrix} \uparrow & \uparrow \\ r+1 & n \end{matrix}$

reduziert worden ist. (Dies ist stets durch nachträgliches Ummummern der Variablen erreichbar.) Dann sind also x_{r+1}, \dots, x_n frei wählbar. Wählen wir für irgend ein $\ell \in \{1, \dots, n - r\}$

$$x_k^{(\ell)} := \begin{cases} 1 & \text{falls } k = r + \ell, \\ 0 & \text{falls } r < k < r + \ell \text{ or } r + \ell < k \leq n, \end{cases}$$

und bestimmen wir weiter die Komponenten $x_r^{(\ell)}, x_{r-1}^{(\ell)}, \dots, x_1^{(\ell)}$ durch Rückwärtseinsetzen, so erhalten wir einen Lösungsvektor \mathbf{x}_ℓ . Die so bestimmten $n - r$ Lösungen $\mathbf{x}_1, \dots, \mathbf{x}_{n-r}$ haben die Form

$$\mathbf{x}_1 = \begin{pmatrix} * \\ \vdots \\ * \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow r+1, \quad \mathbf{x}_2 = \begin{pmatrix} * \\ \vdots \\ * \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \quad \mathbf{x}_{n-r} = \begin{pmatrix} * \\ \vdots \\ * \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \quad (5.35)$$

und sind damit offensichtlich linear unabhängig. Weil nur $n - r$ Parameter wählbar sind, kann es nicht mehr linear unabhängige Lösungen geben. Also bilden diese $n - r$ Vektoren eine Basis des Lösungsraumes \mathcal{L}_0 , der folglich die Dimension $n - r$ hat:

Satz 5.12 *Bezeichnet r den Rang der Matrix \mathbf{A} und \mathcal{L}_0 den Lösungsraum von $\mathbf{A}\mathbf{x} = \mathbf{o}$, so ist*

$$\dim \mathcal{L}_0 \equiv \dim \mathcal{N}(\mathbf{A}) \equiv \dim \ker \mathbf{A} = n - r. \quad (5.36)$$

Mit dem Identitätssymbol \equiv wollen wir hier betonen, dass im Prinzip drei äquivalente Formulierungen vorliegen: es gilt nicht nur $\dim \mathcal{L}_0 = \dim \mathcal{N}(\mathbf{A}) = \dim \ker \mathbf{A}$, sondern es ist aufgrund der Definitionen $\mathcal{L}_0 = \mathcal{N}(\mathbf{A}) \equiv \ker \mathbf{A}$.

BEISPIEL 5.7: In unserem Beispiel 1.4 auf Seite 1-11 in Kapitel 1 das wir auch als Beispiel 3.4 in Kapitel 3, Seite 3-11, betrachtet haben, ist $r = 5$ und $n - r = 9 - 5 = 4$. Das zugehörige homogene System wird durch die Gauss-Elimination reduziert auf die (1.18) entsprechende Zeilenstufenform

$$\begin{array}{cccccccccc|c} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 & 1 \\ \hline \textcircled{1} & 0 & 5 & 0 & 4 & 0 & 0 & 1 & 0 & 0 \\ 0 & \textcircled{5} & 4 & 3 & 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \textcircled{5} & 0 & 4 & 3 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \textcircled{5} & 3 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \textcircled{1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \quad (5.37)$$

aus der man leicht eine Basis von \mathcal{L}_0 konstruiert. Dazu setzt man nacheinander je eine der freien Variablen x_3 , x_5 , x_8 und x_9 eins, die anderen null und rechnet die Werte der restlichen Variablen aus. Man beachte, dass die Zeilenstufenform nicht die vereinfachte Gestalt (5.34) hat, und deshalb auch die Basis nicht genau die gleiche Gestalt wie in (5.35) hat, wobei hier allerdings die vielen Nullen in \mathbf{R} zu vielen zusätzlichen Nullen in der Basis führen:

$$\mathbf{x}_1 = \begin{pmatrix} -5 \\ -\frac{4}{5} \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{x}_2 = \begin{pmatrix} -4 \\ -\frac{2}{5} \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{x}_3 = \begin{pmatrix} -1 \\ -\frac{14}{125} \\ 0 \\ \frac{8}{25} \\ 0 \\ -\frac{2}{5} \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{x}_4 = \begin{pmatrix} 0 \\ \frac{23}{125} \\ 0 \\ -\frac{6}{25} \\ 0 \\ -\frac{1}{5} \\ 0 \\ 0 \\ 1 \end{pmatrix}. \quad (5.38)$$

◆

Durch Vergleich von (5.36) mit der Dimensionsformel (5.28) sehen wir, dass r gleich dem Rang der durch \mathbf{A} gegebenen linearen Abbildung ist.

Bezeichnen wir noch den durch die Zeilenvektoren von \mathbf{A} aufgespannten Unterraum des \mathbb{E}^n als **Zeilenraum** [row space] von \mathbf{A} , so erhalten wir nun eine ganze Reihe äquivalenter Aussagen:

Satz 5.13 *Der Rang einer $m \times n$ Matrix \mathbf{A} ist gleich*

- (i) *der Anzahl Pivotelemente bei der Reduktion von \mathbf{A} auf Zeilenstufenform;*
- (ii) *dem Rang der linearen Abbildung $\mathbf{A} : \mathbb{E}^n \rightarrow \mathbb{E}^m$, definiert als $\dim \operatorname{im} \mathbf{A}$;*
- (iii) *der Dimension des Spaltenraums (“Spaltenrang”), definiert als Anzahl linear unabhängiger Spalten ($\in \mathbb{E}^m$);*
- (iv) *der Dimension des Zeilenraums (“Zeilenrang”), definiert als Anzahl linear unabhängiger Zeilen ($\in \mathbb{E}^n$).*

BEWEIS:

(i) ist unsere Definition des Ranges r von \mathbf{A} aus Kapitel 1.

(ii) folgt aus der Dimensionsformel (5.28) und Satz 5.12:

$$\dim \operatorname{im} \mathbf{A} = n - \dim \ker \mathbf{A} = n - (n - r) = r.$$

Zu (iii): Nach Satz 5.11 ist $\operatorname{im} \mathbf{A}$ der Spaltenraum von \mathbf{A} , also ist nach (ii) dessen Dimension gleich r .

Zu (iv): Bei der Gauss-Elimination von Kapitel 1 lässt sich leicht folgern, dass der Zeilenraum von \mathbf{A} identisch ist mit dem Zeilenraum der Zeilenstufenmatrix \mathbf{R} . Dieser hat die Dimension r . ■

Ein alternativer Beweis von Teil (iv) folgt aus der LR-Zerlegung: Der Zeilenraum von \mathbf{A} ist der Spaltenraum, d.h. der Bildraum von \mathbf{A}^T . Aus der Beziehung

$$\mathbf{A} = \mathbf{P}^{-1}\mathbf{L}\mathbf{R} = \mathbf{P}^T\mathbf{L}\mathbf{R} \quad (5.39)$$

von Satz 3.3 ergibt sich aber $\mathbf{A}^T = \mathbf{R}^T\mathbf{L}^T\mathbf{P}$, wobei \mathbf{P} und \mathbf{L} regulär sind, also auch \mathbf{L}^T . Hieraus folgt, dass $\mathcal{R}(\mathbf{A}^T) = \mathcal{R}(\mathbf{R}^T)$. Wie erwähnt ist aber natürlich $\dim \mathcal{R}(\mathbf{R}^T) = r$.

Die Aussagen (iii) und (iv) implizieren die “Formel”

$$\boxed{\text{“Zeilenrang”} = \text{“Spaltenrang”} .}$$

Wenn man noch berücksichtigt, dass für eine komplexe Matrix natürlich $\operatorname{Rang} \bar{\mathbf{A}} = \operatorname{Rang} \mathbf{A}$ gilt, folgt sofort:

Korollar 5.14 $\operatorname{Rang} \mathbf{A}^T = \operatorname{Rang} \mathbf{A}^H = \operatorname{Rang} \mathbf{A} .$

Wir haben oben eine Basis von $\ker \mathbf{A} \equiv \mathcal{N}(\mathbf{A})$ konstruiert. Wie findet man wohl eine von $\operatorname{im} \mathbf{A} \equiv \mathcal{R}(\mathbf{A})$?

Das ist einfach, wenn wir an die Zeilenstufenform (1.28) und die entsprechende LR-Zerlegung (3.23) denken. Wie in Beispiel 5.6

vermerkt, ist $\mathcal{R}(\mathbf{A}) \equiv \text{im } \mathbf{A}$ gleich der Menge der $\mathbf{b} \in \mathbb{R}^m$, für die $\mathbf{Ax} = \mathbf{b}$ eine Lösung \mathbf{x} hat. Bei der Reduktion auf Zeilenstufenform wird \mathbf{b} durch $\mathbf{c} = \mathbf{L}^{-1}\mathbf{Pb}$ ersetzt, und anschliessend findet man alle \mathbf{x} zu einem festen \mathbf{b} bzw. \mathbf{c} durch Rückwärtseinsetzen. Dabei gibt es zu jedem \mathbf{b} , für das eine Lösung existiert, auch eine spezielle Lösung \mathbf{x} , in der alle freien Variablen null sind. Mit anderen Worten, wir erhalten alle \mathbf{c} und \mathbf{b} auch, wenn wir in \mathbf{R} bzw. \mathbf{A} die Kolonnen mit den freien Variablen streichen, das heisst, wenn wir nur die r Pivotkolonnen behalten, mit denen wir eine $m \times r$ Matrix $\tilde{\mathbf{A}}$ bilden können. Von diesen lässt sich andererseits keine streichen, ohne dass die Menge zulässiger Vektoren \mathbf{b} reduziert wird. Es gilt also:

Satz 5.15 Für den Kolonnenraum einer $m \times n$ -Matrix \mathbf{A} gilt:

$$\text{im } \mathbf{A} \equiv \mathcal{R}(\mathbf{A}) = \mathcal{R}(\tilde{\mathbf{A}}) = \text{span} \{ \mathbf{a}_{n_1}, \dots, \mathbf{a}_{n_r} \}, \quad (5.40)$$

worin $\mathbf{a}_{n_1}, \dots, \mathbf{a}_{n_r}$ die Pivotkolonnen von \mathbf{A} sind, und $\tilde{\mathbf{A}}$ die daraus gebildete $m \times r$ -Matrix bezeichnet.

BEISPIEL 5.8: Wir setzen das oben behandelte Beispiel 5.7 fort. Gemäss (5.37) ist $r = 5$ und sind x_1, x_2, x_4, x_6, x_7 die Pivotvariablen, also $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_4, \mathbf{a}_6, \mathbf{a}_7$ die Pivotkolonnen von \mathbf{A} . Sie bilden gemäss (5.40) eine Basis des Kolonnenraumes und ergeben zusammen die Matrix $\tilde{\mathbf{A}}$:

$$\tilde{\mathbf{A}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 5 & 3 & 1 & 0 \\ 1 & 10 & 11 & 6 & 3 \\ 0 & 5 & 18 & 18 & 12 \\ 0 & 0 & 5 & 24 & 16 \\ 1 & 10 & 21 & 24 & 16 \\ 0 & 5 & 13 & 24 & 17 \end{pmatrix}. \quad (5.41)$$

Wir werden in Abschnitt 7.4 eine weitere Methode zur Konstruktion einer Basis von $\mathcal{R}(\mathbf{A})$ kennen lernen, die direkt eine orthogonale Basis liefert. Sie ist im praktischen Rechnen vorzuziehen.

Allgemeine Aussagen über den Rang eines Produktes \mathbf{AB} ergeben sich sofort aus Lemma 5.3, Korollar 5.10 und Satz 5.13 (ii):

Satz 5.16 Es seien $\mathbf{A} \in \mathbb{E}^{m \times n}$, $\mathbf{B} \in \mathbb{E}^{p \times m}$. Dann gilt:

- (i) $\text{Rang } \mathbf{BA} \leq \min\{\text{Rang } \mathbf{B}, \text{Rang } \mathbf{A}\}$,
- (ii) $\text{Rang } \mathbf{B} = m \ (\leq p) \implies \text{Rang } \mathbf{BA} = \text{Rang } \mathbf{A}$,
- (iii) $\text{Rang } \mathbf{A} = m \ (\leq n) \implies \text{Rang } \mathbf{BA} = \text{Rang } \mathbf{B}$.

BEWEIS:

(i) ergibt sich aus den genannten Aussagen.

(ii) Bei der Anwendung von Korollar 5.10 benütze man noch, dass nach Satz 5.6 und Satz 5.12 folgt

$$\begin{aligned} \mathbf{B} : \mathbb{E}^m \rightarrow \mathbb{E}^p \text{ injektiv} &\iff \ker \mathbf{B} = \{\mathbf{o}\} \\ &\iff \text{Rang } \mathbf{B} = m. \end{aligned}$$

(iii) Hier wird zusätzlich benützt, dass natürlich gilt:

$$\begin{aligned} \mathbf{A} : \mathbb{E}^n \rightarrow \mathbb{E}^m \quad \text{surjektiv} &\iff \text{im } \mathbf{A} = \mathbb{E}^m \\ &\iff \text{Rang } \mathbf{A} = m. \end{aligned}$$

■

Im Falle von *quadratische* Matrizen ist vieles etwas einfacher, besonders wenn sie regulär sind. Satz 5.16 vereinfacht sich so:

Korollar 5.17 *Es seien $\mathbf{A} \in \mathbb{E}^{m \times m}$, $\mathbf{B} \in \mathbb{E}^{m \times m}$. Dann gilt:*

- (i) $\text{Rang } \mathbf{BA} \leq \min\{\text{Rang } \mathbf{B}, \text{Rang } \mathbf{A}\}$,
- (ii) $\text{Rang } \mathbf{B} = m \implies \text{Rang } \mathbf{BA} = \text{Rang } \mathbf{A}$,
- (iii) $\text{Rang } \mathbf{A} = m \implies \text{Rang } \mathbf{BA} = \text{Rang } \mathbf{B}$.

Im folgenden Satz fassen wir nochmals eine Reihe von gleichwertigen Aussagen zusammen.

Satz 5.18 *Für eine quadratische Matrix $\mathbf{A} \in \mathbb{E}^{n \times n}$ sind folgende Aussagen äquivalent:*

- (i) \mathbf{A} ist invertierbar;
- (ii) \mathbf{A} ist regulär;
- (iii) $\text{Rang } \mathbf{A} = n$;
- (iv) die n Kolonnenvektoren von \mathbf{A} sind linear unabhängig;
- (v) die n Zeilenvektoren von \mathbf{A} sind linear unabhängig;
- (vi) $\text{im } \mathbf{A} \equiv \mathcal{R}(\mathbf{A}) = \mathbb{E}^n$;
- (vii) $\ker \mathbf{A} \equiv \mathcal{N}(\mathbf{A}) = \{\mathbf{o}\}$;
- (viii) die lineare Abbildung $\mathbf{A} : \mathbb{E}^n \rightarrow \mathbb{E}^n$ ist ein Automorphismus;
- (ix) \mathbf{A} ist die Transformationsmatrix einer Koordinatentransformation in \mathbb{E}^n .

BEWEIS: (i) \iff (ii) auf Grund unserer Definition von “regulär”.

(ii) \iff (iii) nach Korollar 1.4 und Satz 2.15.

(iii) \iff (iv) \iff (v) ergibt sich aus Satz 5.13 (iii) und (iv).

(iv) \iff (vi), weil n Vektoren genau dann den \mathbb{E}^n erzeugen, wenn sie linear unabhängig sind.

(iii) \iff (vii) nach der Dimensionsformel (5.36); oder: weil beide Aussagen äquivalent damit sind, dass $\mathbf{Ax} = \mathbf{o}$ nur die triviale Lösung hat.

(iii) \iff (viii) folgt aus Korollar 5.8 (iii) und Satz 5.13 (ii).

(viii) \iff (ix) ist ziemlich klar, soll aber genauer betrachtet werden: $\mathbf{A} = (a_{kl})$ ist Transformationsmatrix genau dann, wenn aus der “alten” Basis $\mathbf{b}_1, \dots, \mathbf{b}_n$ gemäss (4.38),

$$\mathbf{b}'_l := \sum_{k=1}^n a_{kl} \mathbf{b}_k \quad (l = 1, \dots, n) \quad (5.42)$$

eine “neue” Basis $\mathbf{b}'_1, \dots, \mathbf{b}'_n$ hervorgeht, d.h. genau dann, wenn $\mathbf{b}'_1, \dots, \mathbf{b}'_n$ den ganzen Raum \mathbb{E}^n erzeugen (womit sie automatisch linear unabhängig sind). Wir können \mathbf{b}'_l als Bild von \mathbf{b}_l bei einer linearen Abbildung F auffassen: $F\mathbf{b}_l := \mathbf{b}'_l$, vgl. (5.14). Es ist dann $\text{span} \{\mathbf{b}'_1, \dots, \mathbf{b}'_n\} = \mathbb{E}^n$ gleichbedeutend mit $\text{Rang } F = n$, was gemäss Korollar 5.8 heisst, dass F ein Automorphismus ist. Gemäss (5.42) ist auch $\mathbf{A}\mathbf{b}_l = \mathbf{b}'_l$, also, wenn man \mathbf{A} als lineare Abbildung auffasst, $\mathbf{A} = F$. ■

5.4 Affine Räume und die allgemeine Lösung eines inhomogenen Gleichungssystems

DEFINITION: Ist U ein echter Unterraum eines Vektorraumes V und $u_0 \in V$, so heisst die Menge

$$u_0 + U := \{u_0 + u ; u \in U\} \quad (5.43)$$

ein **affiner Teilraum** [*affine (sub)space*].

Ist $F : X \rightarrow Y$ eine lineare Abbildung und $y_0 \in Y$, so ist

$$H : X \rightarrow y_0 + Y, \quad x \mapsto y_0 + Fx \quad (5.44)$$

eine **affine Abbildung** [*affine mapping*]. ▲

BEISPIEL 5.9: Jede Gerade und jede Ebene im \mathbb{R}^3 ist ein affiner Teilraum, auch wenn sie nicht durch den Ursprung verläuft. Eine Translation $\mathbf{x} \mapsto \mathbf{y}_0 + \mathbf{x}$ ist eine spezielle affine Abbildung. ◆

Man beachte, dass ein affiner Teilraum im allgemeinen kein Unterraum ist, nämlich stets wenn $u_0 \notin U$ gilt. Ebenso ist eine affine Abbildung nicht linear, wenn $y_0 \neq o$ ist, denn $H(o) = y_0 \neq o$. Dies gilt selbst, wenn $y_0 \in \text{im } F \subset Y$ und damit das Bild $\text{im } H = \text{im } F$ ein Unterraum ist.

Allgemein ist $\text{im } H = y_0 + \text{im } F$ ein affiner Teilraum.

Affine Teilräume und affine Abbildungen spielen eine grosse Rolle bei geometrischen Aufgaben und Anwendungen. Wir betrachten hier nur den Zusammenhang mit linearen Gleichungssystemen.

Satz 5.19 *Es sei \mathbf{x}_0 irgend eine Lösung des inhomogenen Systems $\mathbf{A}\mathbf{x} = \mathbf{b}$, und es bezeichne \mathcal{L}_0 den Lösungsraum des homogenen Systems $\mathbf{A}\mathbf{x} = \mathbf{o}$. Dann ist die Lösungsmenge \mathcal{L}_b von $\mathbf{A}\mathbf{x} = \mathbf{b}$ gleich dem affinen Teilraum*

$$\mathcal{L}_b = \mathbf{x}_0 + \mathcal{L}_0. \quad (5.45)$$

Mit anderen Worten: Die allgemeine Lösung eines inhomogenen linearen Gleichungssystems lässt sich darstellen als Summe einer speziellen Lösung (Partikulärlösung) des Systems und der allgemeinen Lösung des homogenen Systems.

Analoge Aussagen gibt es für andere Lösungsmengen linearer Beziehungen. Zum Beispiel für lineare Differentialgleichungen, und zwar sowohl für eine Differentialgleichung höherer Ordnung als auch für Systeme erster (oder höherer) Ordnung. In jedem Falle ist die Lösungsmenge ein affiner Teilraum eines geeignet gewählten Funktionenraums.

BEWEIS von Satz 5.19: Aus $\mathbf{A}\mathbf{x}_0 = \mathbf{b}$ und $\mathbf{A}\mathbf{x} = \mathbf{o}$ folgt $\mathbf{A}(\mathbf{x}_0 + \mathbf{x}) = \mathbf{b}$; es gilt also $\mathcal{L}_b \supseteq \mathbf{x}_0 + \mathcal{L}_0$. Ist umgekehrt \mathbf{x}_1 irgend eine Lösung von $\mathbf{A}\mathbf{x} = \mathbf{b}$, so gilt $\mathbf{A}(\mathbf{x}_1 - \mathbf{x}_0) = \mathbf{o}$, es ist also $\mathbf{x}_1 - \mathbf{x}_0 \in \mathcal{L}_0$, d.h., $\mathbf{x}_1 \in \mathbf{x}_0 + \mathcal{L}_0$; also gilt auch $\mathcal{L}_b \subseteq \mathbf{x}_0 + \mathcal{L}_0$. ■

BEISPIEL 5.10: Wir betrachten nochmals das oben behandelte Beispiel 5.7, interessieren uns nun aber für die allgemeine Lösung des inhomogenen Systems (1.17), die wir übrigens ja bereits in Kapitel 1 bestimmt haben, siehe (1.21). Dort haben wir insbesondere die spezielle Lösung

$$\mathbf{x}_0 = \left(1 \quad \frac{43}{125} \quad 0 \quad \frac{4}{25} \quad 0 \quad -\frac{1}{5} \quad 1 \quad 0 \quad 0 \right)^T$$

gefunden, in der alle freien Variablen null gesetzt sind. Anderseit haben wir in (5.38) eine Basis von \mathcal{L}_0 berechnet. In Matrixschreibweise lässt sich damit die allgemeine Lösung ausdrücken als

$$\mathbf{x} = \mathbf{x}_0 + \begin{pmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 & \mathbf{x}_4 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix} \quad (5.46)$$

$$= \begin{pmatrix} 1 \\ \frac{43}{125} \\ 0 \\ \frac{4}{25} \\ 0 \\ -\frac{1}{5} \\ 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} -5 & -4 & -1 & 0 \\ -\frac{4}{5} & -\frac{2}{5} & -\frac{14}{125} & \frac{23}{125} \\ 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{8}{25} & -\frac{6}{25} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\frac{2}{5} & -\frac{1}{5} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix} \quad (5.47)$$

mit frei wählbaren Parametern $\alpha, \beta, \gamma, \delta$. Dies stimmt genau mit der allgemeinen Lösung in (1.21) überein. ◆

5.5 Die Abbildungsmatrix bei Koordinatentransformation

Es seien X und Y Vektorräume der Dimensionen n bzw. m , und weiter seien

$F: X \rightarrow Y, \quad x \mapsto y$ eine lineare Abbildung,

$\mathbf{A}: \mathbb{E}^n \rightarrow \mathbb{E}^m, \quad \boldsymbol{\xi} \mapsto \boldsymbol{\eta}$ die entsprechende Abbildungsmatrix,

$\mathbf{T}: \mathbb{E}^n \rightarrow \mathbb{E}^n, \quad \boldsymbol{\xi}' \mapsto \boldsymbol{\xi}$ eine Transformationsmatrix im \mathbb{E}^n ,

$\mathbf{S}: \mathbb{E}^m \rightarrow \mathbb{E}^m, \quad \boldsymbol{\eta}' \mapsto \boldsymbol{\eta}$ eine Transformationsmatrix im \mathbb{E}^m .

Wir wissen aus Abschnitt 4.4 wie Basiswechsel in X und Y durch Koordinatentransformationen \mathbf{T} und \mathbf{S} ausgedrückt werden. Wir wollen nun die Auswirkungen auf die Abbildungsmatrix \mathbf{A} betrachten. Dazu genügt es, das kommutative Diagramm (5.26) nach unten zu erweitern:

$$\begin{array}{ccc}
 x \in X & \xrightarrow[\text{lin. Abb.}]{F} & y \in Y \\
 \kappa_X \Big\downarrow \Big\uparrow \kappa_X^{-1} & & \kappa_Y \Big\downarrow \Big\uparrow \kappa_Y^{-1} \quad \text{(Koordinaten-} \\
 & & \text{abbildung bzgl.} \\
 \xi \in \mathbb{E}^n & \xrightarrow[\text{Abb.matrix}]{\mathbf{A}} & \eta \in \mathbb{E}^m \quad \text{(Koordinaten} \\
 & & \text{bzgl. "alten" Basen)} \quad (5.48) \\
 \mathbf{T}^{-1} \Big\downarrow \Big\uparrow \mathbf{T} & & \mathbf{S}^{-1} \Big\downarrow \Big\uparrow \mathbf{S} \quad \text{(Koordinaten-} \\
 & & \text{transformation)} \\
 \xi' \in \mathbb{E}^n & \xrightarrow[\text{Abb.matrix}]{\mathbf{B}} & \eta' \in \mathbb{E}^m \quad \text{(Koordinaten} \\
 & & \text{bzgl. "neuen" Basen)}
 \end{array}$$

Es gilt also

$$\boxed{y = Fx, \quad \eta = \mathbf{A}\xi, \quad \xi = \mathbf{T}\xi', \quad \eta = \mathbf{S}\eta', \quad \eta' = \mathbf{B}\xi'}. \quad (5.49)$$

Diesen Formeln oder dem Diagramm entnimmt man, dass für die Abbildungsmatrix \mathbf{B} , die die Abbildung F bezüglich den "neuen" Basen in \mathbb{E}^m und \mathbb{E}^n beschreibt, gilt

$$\mathbf{B}\xi' = \eta' = \mathbf{S}^{-1}\eta = \mathbf{S}^{-1}\mathbf{A}\xi = \mathbf{S}^{-1}\mathbf{A}\mathbf{T}\xi'$$

Da ξ' beliebig ist, ergibt sich

$$\boxed{\mathbf{B} = \mathbf{S}^{-1}\mathbf{A}\mathbf{T}, \quad \mathbf{A} = \mathbf{S}\mathbf{B}\mathbf{T}^{-1}}. \quad (5.50)$$

Aus Satz 5.16 folgt im übrigen wegen $\text{Rang } \mathbf{S}^{-1} = \text{Rang } \mathbf{T} = n$, dass $\text{Rang } \mathbf{B} = \text{Rang } \mathbf{A}$ ist, und in ähnlicher Weise folgt aus Korollar 5.10, dass $\text{Rang } F = \text{Rang } \mathbf{A}$ ist:

$$\boxed{\text{Rang } F = \text{Rang } \mathbf{A} = \text{Rang } \mathbf{B}}. \quad (5.51)$$

Im Falle einer linearen Abbildung von X in sich, ist natürlich $Y = X$, $\kappa_Y = \kappa_X$, $\mathbf{S} = \mathbf{T}$. Aus (5.50) wird damit

$$\boxed{\mathbf{B} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}, \quad \mathbf{A} = \mathbf{T}\mathbf{B}\mathbf{T}^{-1}}. \quad (5.52)$$

Da \mathbf{B} und \mathbf{A} Matrixdarstellungen der selben linearen Selbstabbildung sind, kann man erwarten, dass sie gewisse ähnliche Eigenschaften haben. Das wird sich bestätigen. Es gilt folgende Terminologie:

DEFINITION: Die $n \times n$ -Matrizen \mathbf{A} und \mathbf{B} heissen **ähnlich** [*similar*], falls es eine reguläre Matrix \mathbf{T} gibt, so dass eine der Beziehungen (5.52) gilt (und damit auch die andere).

Der Übergang $\mathbf{A} \mapsto \mathbf{B} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$ heisst **Ähnlichkeitstransformation** [*similarity transformation*].

BEISPIEL 5.11: Wir betrachten den Ableitungsoperator D auf dem in den Beispielen 4.21 und 4.30 von Kapitel 4 betrachteten Unterraum \mathcal{G}_4 der geraden Polynome vom Höchstgrad 4. Er hat die monomiale Basis $1, t^2, t^4$. Für jedes $p \in \mathcal{G}_4$ ist die Ableitung p' ein ungerades Polynom vom Höchstgrad 3, liegt also im entsprechenden Unterraum \mathcal{U}_3 , der von den zwei Monomen t und t^3 aufgespannt wird. Bezüglich diesen Basen wird $D : p \in \mathcal{G}_4 \mapsto p' \in \mathcal{U}_3$ dargestellt durch die 2×3 -Matrix

$$\mathbf{A} = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}, \quad (5.53)$$

die man im übrigen auch aus der Matrix \mathbf{D} in Beispiel 5.5 bekommt, indem man die erste und die dritte Zeile sowie die zweite und die vierte Kolonne streicht.

In \mathcal{G}_4 ersetzen wir nun die monomiale Basis durch jene aus (4.17):

$$p_1(t) := 1 + t^2, \quad p_2(t) := 1 - t^2, \quad p_3(t) := 1 + t^2 + t^4.$$

Diese Basistransformation wird durch die Matrix

$$\mathbf{T} = (\tau_{ik}) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

dargestellt, wie wir ja auch schon in (4.45) gesehen haben. (Zur Erinnerung: In den Kolonnen stehen die Koordinaten der neuen Basis bezüglich der alten Basis.)

In \mathcal{U}_3 wählen wir als neue Basis

$$q_1(t) := t, \quad q_2(t) := 3t + 2t^3. \quad (5.54)$$

Die entsprechende Transformationsmatrix und ihre Inverse sind

$$\mathbf{S} = (\sigma_{ik}) = \begin{pmatrix} 1 & 3 \\ 0 & 2 \end{pmatrix}, \quad \mathbf{S}^{-1} = \begin{pmatrix} 1 & -\frac{3}{2} \\ 0 & \frac{1}{2} \end{pmatrix}. \quad (5.55)$$

Um den Ableitungsoperator bezüglich den Basen $\{p_1, p_2, p_3\}$ von \mathcal{G}_4 und $\{q_1, q_2\}$ von \mathcal{U}_3 darzustellen, braucht man somit gemäss (5.50) die Abbildungsmatrix

$$\begin{aligned} \mathbf{B} &= \mathbf{S}^{-1}\mathbf{A}\mathbf{T} \\ &= \begin{pmatrix} 1 & -\frac{3}{2} \\ 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & -2 & -4 \\ 0 & 0 & 2 \end{pmatrix}. \end{aligned} \quad (5.56)$$



Wie weit lässt sich die Abbildungsmatrix durch geeignete Wahl der Basen vereinfachen?

Im Falle $Y = X$ werden wir uns später im Zusammenhang mit dem Eigenwertproblem ausgiebig mit dieser Frage beschäftigen. Der Fall $Y \neq X$ ist einfacher, denn man kann statt einer zwei Basen geschickt wählen:

Satz 5.20 *Es sei $F : X \rightarrow Y$ eine lineare Abbildung, wobei $\dim X = n$, $\dim Y = m$, $\text{Rang } F = r$ gelte. Dann besitzt F bezüglich geeignet gewählten Basen in X und Y die Abbildungsmatrix*

$$\mathbf{A} = \left(\begin{array}{c|c} \mathbf{I}_r & \mathbf{O} \\ \hline \mathbf{O} & \mathbf{O} \end{array} \right). \quad (5.57)$$

BEWEIS: Es seien $\{b_1, \dots, b_n\}$ und $\{c_1, \dots, c_r\}$ gleich gewählt wie im Beweis von Satz 5.7. Zudem werde $\{c_1, \dots, c_r\}$ gemäss Satz 4.9 zu einer Basis $\{c_1, \dots, c_m\}$ von Y ergänzt. Dann ist also

$$Fb_j = \begin{cases} c_j, & j = 1, \dots, r, \\ o, & j = r + 1, \dots, n. \end{cases}$$

Da in der Kolonne j der Abbildungsmatrix gerade die Koordinaten von Fb_j (bezüglich der Basis des Bildraumes) stehen, bekommen wir genau die Matrix \mathbf{A} von (5.57). ■

BEISPIEL 5.12: Im Falle des Ableitungsoperators $D : p \in \mathcal{G}_4 \mapsto p' \in \mathcal{U}_3$ aus Beispiel 5.11 genügt es, die monomiale Basis von \mathcal{G}_4 umzuordnen in $\{t^4, t^2, 1\}$ und die monomiale Basis von \mathcal{U}_3 zusätzlich zu skalieren: $\{4t^3, 2t\}$. Dann hat man als Abbildungsmatrix statt \mathbf{A} aus (5.53) einfach

$$\tilde{\mathbf{A}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}. \quad (5.58)$$

◆

Kapitel 6

Vektorräume mit Skalarprodukt

Wir verallgemeinern in diesem Kapitel verschiedene in Kapitel 2 für n -Vektoren eingeführten Begriffe wie Norm, Skalarprodukt und Orthogonalität. Wir wissen ja schon, siehe Satz 5.9, dass jeder n -dimensionale Vektorraum V über \mathbb{R} bzw. \mathbb{C} isomorph ist zum \mathbb{R}^n beziehungsweise \mathbb{C}^n im Sinne, dass es eine eindeutige lineare Abbildung von V auf \mathbb{R}^n oder \mathbb{C}^n gibt. Die Existenz eines Skalarproduktes macht einen n -dimensionalen Vektorraum aber “noch ähnlicher” zum \mathbb{R}^n oder \mathbb{C}^n . Es ist zum Beispiel naheliegend, dass man in solchen Räumen orthogonale Basen betrachtet und an linearen Abbildungen interessiert ist, die Längen und Winkel erhalten.

6.1 Normierte Vektorräume

DEFINITION: Eine **Norm** [*norm*] in einem Vektorraum V ist eine Funktion

$$\|\cdot\| : V \rightarrow \mathbb{R}, \quad x \mapsto \|x\| \quad (6.1)$$

mit den folgenden Eigenschaften:

(N1) Sie ist **positiv definit**:

$$\begin{aligned} \|x\| &\geq 0 && \text{für alle } x \in V, \\ \|x\| = 0 &\implies x = o. \end{aligned}$$

(N2) Sie ist **dem Betrage nach homogen**:

$$\|\alpha x\| = |\alpha| \|x\| \quad \text{für alle } x \in V, \alpha \in \mathbb{E}.$$

(N3) Die **Dreiecksungleichung** [*triangle inequality*] gilt:

$$\boxed{\|x + y\| \leq \|x\| + \|y\| \quad \text{für alle } x, y \in V.} \quad (6.2)$$

Ein Vektorraum mit einer Norm heisst **normierter Vektorraum** [*normed vector space*] oder **normierter linearer Raum** [*normed linear space*]. ▲

Man beachte, dass die Norm-Axiome (N1)–(N3) genau jene sind, die wir in Satz 2.12 als Eigenschaften der 2-Norm für \mathbb{R}^n und \mathbb{C}^n aufgezählt haben. Dort wurde die 2-Norm mit Hilfe des Euklidischen Skalarprodukt $\langle \cdot, \cdot \rangle$ definiert: $\|\mathbf{x}\| \equiv \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$. Dies werden wir im folgenden Abschnitt verallgemeinern. Aber hier wird die Norm ohne Bezug auf ein Skalarprodukt definiert.

BEISPIEL 6.1: Auf dem in Beispiel 4.2 betrachteten Raum $C[a, b]$ der auf dem Intervall $[a, b]$ definierten und dort stetigen reell- oder komplexwertigen¹ Funktionen werden durch

$$\|f\|_1 := \int_a^b |f(t)| dt \quad (6.3)$$

und

$$\|f\|_\infty := \max_{t \in [a, b]} |f(t)| \quad (6.4)$$

zwei verschiedene Normen definiert, die **L_1 -Norm** und die **L_∞ -Norm**. Die zweite heisst auch **Maximumnorm** [*maximum norm*]. Ersetzt man in (6.4) das Maximum durch ein Supremum, so können beide Normen auch auf einem “grösseren”, un stetige Funktionen enthaltenden Vektorraum definiert werden. ◆

6.2 Vektorräume mit Skalarprodukt

DEFINITION: Ein **Skalarprodukt** [*inner product*] in einem reellen oder komplexen Vektorraum ist eine Funktion von zwei Variablen

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{E}, \quad x, y \mapsto \langle x, y \rangle \quad (6.5)$$

mit folgenden Eigenschaften:

(S1) Es ist **linear** im zweiten Faktor:

$$\begin{aligned} \langle x, y + z \rangle &= \langle x, y \rangle + \langle x, z \rangle && \text{für alle } x, y, z \in V, \\ \langle x, \alpha y \rangle &= \alpha \langle x, y \rangle && \text{für alle } x, y \in V, \alpha \in \mathbb{E}. \end{aligned}$$

(S2) Falls $\mathbb{E} = \mathbb{R}$, ist es **symmetrisch**:

$$\langle x, y \rangle = \langle y, x \rangle \quad \text{für alle } x, y \in V.$$

Falls $\mathbb{E} = \mathbb{C}$, ist es **Hermiteisch**:

$$\langle x, y \rangle = \overline{\langle y, x \rangle} \quad \text{für alle } x, y \in V.$$

(S3) Es ist **positiv definit**:

$$\begin{aligned} \langle x, x \rangle &\geq 0 && \text{für alle } x \in V, \\ \langle x, x \rangle &= 0 && \implies x = o. \end{aligned}$$

V mit $\langle \cdot, \cdot \rangle$ ist ein **Vektorraum mit Skalarprodukt** [*inner product space*]. Ist V endlichdimensional, nennt man V auch

- falls $\mathbb{E} = \mathbb{R}$: **Euklidischer Vektorraum**² [*Euclidean vector space*] oder **orthogonaler Vektorraum**,
- falls $\mathbb{E} = \mathbb{C}$: **unitärer Vektorraum** [*unitary vector space*]. ▲

¹In diesem Kapitel wollen wir explizit den Fall komplexwertiger Funktionen berücksichtigen, da sie die Modifikation einiger Formeln erfordern.

²Euklidische Vektorräume sind also Verallgemeinerungen des \mathbb{R}^n bzw. \mathbb{C}^n . Letztere haben wir als reellen bzw. komplexen n -dimensionalen Euklidischen Raum bezeichnet.

Aus (S1) und (S3) folgt, dass

$$\begin{aligned}\langle x, \alpha y + \beta z \rangle &= \alpha \langle x, y \rangle + \beta \langle x, z \rangle, \\ \langle \alpha w + \beta x, y \rangle &= \bar{\alpha} \langle w, y \rangle + \bar{\beta} \langle x, y \rangle,\end{aligned}\quad (6.6)$$

wobei im Falle $\mathbb{E} = \mathbb{R}$ natürlich $\bar{\alpha} = \alpha$ und $\bar{\beta} = \beta$ ist, das Skalarprodukt also auch im ersten Argument linear ist.

BEMERKUNG: Gemäss (S1) ist das Skalarprodukt linear im zweiten Argument. Stattdessen wird auch im Falle $\mathbb{E} = \mathbb{C}$ oft Linearität im ersten Argument verlangt. In der komplexen Matrizenrechnung ist aber unsere Version von Vorteil, denn im Falle von n -Vektoren gilt so direkt $\langle \mathbf{y}, \mathbf{x} \rangle = \mathbf{y}^H \mathbf{x}$. ▼

In diesem Kapitel sind V, X, Y, \dots stets Vektorräume mit Skalarprodukt, und zwar ist $\langle \cdot, \cdot \rangle$ oder $\langle \cdot, \cdot \rangle_V$ das Skalarprodukt in V , $\langle \cdot, \cdot \rangle_X$ jenes in X , usw.

DEFINITION: Die **Norm** [*norm*] oder **Länge** [*length*] eines Vektors x in einem Vektorraum V mit Skalarprodukt ist

$$\|x\| := \sqrt{\langle x, x \rangle}. \quad (6.7)$$

Man zeigt wie im Beweis von Satz 2.12, dass die in Abschnitt 6.1 angegebenen Axiome (N1)–(N3) für die Norm (6.7) stets erfüllt sind. ▲

BEISPIEL 6.2: \mathbb{R}^n mit dem Skalarprodukt $\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{x}^T \mathbf{y}$ ist ein Euklidischer Vektorraum. \mathbb{C}^n mit dem Skalarprodukt $\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{x}^H \mathbf{y}$ ist ein unitärer Vektorraum. ◆

BEISPIEL 6.3: Auf dem Raum $C[a, b]$ der auf dem Intervall $[a, b]$ definierten und dort stetigen reell- oder komplexwertigen Funktionen wird durch

$$\langle f, g \rangle := \int_a^b \overline{f(t)} g(t) dt \quad (6.8)$$

ein Skalarprodukt definiert. Die zugehörige Norm ist

$$\|f\|_2 := \sqrt{\int_a^b |f(t)|^2 dt}. \quad (6.9)$$

Ist $w \in C[a, b]$ irgend eine positive reelle **Gewichtsfunktion** [*weight function*], so ist auch durch

$$\langle f, g \rangle_w := \int_a^b \overline{f(t)} g(t) w(t) dt \quad (6.10)$$

ein Skalarprodukt erklärt. Weiter wird auf $C[-1, 1]$ zum Beispiel durch

$$\langle f, g \rangle := \int_{-1}^1 \overline{f(t)} g(t) \frac{dt}{\sqrt{1-t^2}} \quad (6.11)$$

ein Skalarprodukt definiert, denn diese Integrale existieren alle. ◆

Grundlegende Folgerungen aus den Axiomen des Skalarproduktes lassen sich genau gleich herleiten, wie wir das in Kapitel 2 für das spezielle Euklidische Skalarprodukt von \mathbb{R}^n und \mathbb{C}^n taten. Als Beispiele geben wir die allgemeine Form der Schwarzschen Ungleichung, die Definitionen des Winkels und der Orthogonalität, sowie die allgemeine Form des Satzes von Pythagoras an:

Satz 6.1 Für alle Paare $x, y \in V$ gilt die **Schwarzsche Ungleichung** (oder **Cauchy-Bunjakowski-Schwarz-Ungleichung**)

$$|\langle x, y \rangle| \leq \|x\| \|y\|. \quad (6.12)$$

Das Gleichheitszeichen gilt genau dann, wenn x und y linear abhängig sind.

BEWEIS: Der Beweis verläuft völlig analog zu jenem von Satz 2.11. ■

DEFINITION: Der **Winkel [angle]** φ ($0 \leq \varphi \leq \pi$) zwischen zwei Vektoren x und y ist gegeben durch

$$\varphi := \arccos \frac{\langle x, y \rangle}{\|x\| \|y\|}, \quad \text{falls } \mathbb{E} = \mathbb{R}, \quad (6.13)$$

$$\varphi := \arccos \frac{\operatorname{Re} \langle x, y \rangle}{\|x\| \|y\|}, \quad \text{falls } \mathbb{E} = \mathbb{C}. \quad (6.14)$$

Zwei Vektoren $x, y \in V$ sind zueinander **orthogonal [orthogonal]** (oder: **senkrecht [perpendicular]**), falls $\langle x, y \rangle = 0$. Symbolisch: $x \perp y$.

Zwei Teilmengen $M, N \subseteq V$ sind zueinander **orthogonal [orthogonal]**, falls $\langle x, y \rangle = 0$ für alle $x \in M$ und alle $y \in N$. Symbolisch: $M \perp N$. ▲

BEMERKUNG: M und N können insbesondere Unterräume sein. Der Nullvektor ist orthogonal zu allen Vektoren; der Nullraum ist orthogonal zu allen Unterräumen. ▼

Für senkrechte Vektoren kann man nun allgemein den Satz von Pythagoras beweisen, genau wie Satz 2.13:

Satz 6.2 (Satz von Pythagoras) In einem Vektorraum mit Skalarprodukt gilt:

$$x \perp y \implies \|x \pm y\|^2 = \|x\|^2 + \|y\|^2. \quad (6.15)$$

BEISPIEL 6.4: Wir betrachten im Raum $C[0, 1]$, versehen mit dem Skalarprodukt (6.8) die zwei Funktionen $f(t) : t \mapsto t^2$ und $g(t) : t \mapsto t^3$. Hier gilt: $\langle f, g \rangle = 1/6$, $\|f\|^2 = 1/5$, $\|g\|^2 = 1/7$. Der Winkel zwischen diesen Funktionen (Vektoren) beträgt damit

$$\varphi = \arccos \frac{\langle f, g \rangle}{\|f\| \|g\|} = \arccos \frac{\sqrt{5}\sqrt{7}}{6} = \arccos \sqrt{\frac{35}{36}}. \quad \blacklozenge$$

6.3 Orthonormalbasen

Sobald die Orthogonalität von Vektoren definiert ist, kann man wie im \mathbb{R}^3 orthogonale Basen, d.h. orthogonale Koordinatensysteme verwenden. Wir überlegen uns in diesem Abschnitt, dass solche Vorteile haben und immer konstruiert werden können. Insbesondere sind die Koordinaten bezüglich einer orthogonalen Basis besonders einfach zu berechnen.

Satz 6.3 Eine Menge M von paarweise orthogonalen Vektoren ist linear unabhängig, falls $0 \notin M$.

BEWEIS: Annahme: Es seien $x_1, \dots, x_m \in M$, $\gamma_1, \dots, \gamma_m \in \mathbb{E}$, und es gelte

$$\sum_{k=1}^m \gamma_k x_k = 0.$$

Auf Grund der Linearität des Skalarproduktes gilt für $j = 1, \dots, m$:

$$0 = \left\langle x_j, \sum_{k=1}^m \gamma_k x_k \right\rangle = \sum_{k=1}^m \gamma_k \underbrace{\langle x_j, x_k \rangle}_{0, \text{ falls } k \neq j} = \gamma_j \langle x_j, x_j \rangle,$$

also ist $\gamma_j = 0$, weil $\langle x_j, x_j \rangle > 0$. ■

Ist $\dim V = n$, so ergeben n paarweise orthogonale, von 0 verschiedene Vektoren somit eine Basis von V .

DEFINITION: Eine Basis heisst **orthogonal** [*orthogonal*], falls die Basisvektoren paarweise orthogonal sind:

$$\langle b_k, b_l \rangle = 0 \quad \text{falls } k \neq l.$$

Sie heisst **orthonormiert** oder **orthonormal**³ [*orthonormal*], wenn zusätzlich die Basisvektoren die Länge 1 haben, d.h. wenn

$$\langle b_k, b_k \rangle = 1 \quad (\forall k). \quad \blacktriangle$$

Mit dem **Kronecker-Symbol** [*Kronecker symbol*], definiert durch

$$\delta_{kl} := \begin{cases} 0, & \text{falls } k \neq l, \\ 1, & \text{falls } k = l, \end{cases}$$

gilt für eine orthonormierte Basis also: $\langle b_k, b_l \rangle = \delta_{kl}$.

Satz 6.4 Ist V ein n -dimensionaler Vektorraum mit Skalarprodukt und $\{b_1, \dots, b_n\}$ eine Orthonormalbasis, so gilt für alle $x \in V$

$$x = \sum_{k=1}^n \underbrace{\langle b_k, x \rangle}_{\equiv: \xi_k} b_k. \quad (6.16)$$

Das heisst für die Koordinaten bezüglich einer Orthonormalbasis gilt einfach $\xi_k = \langle b_k, x \rangle$.

³Man beachte, dass eine *orthogonale Matrix* orthonormale Kolonnen hat und quadratisch ist. Ist die Matrix nicht quadratisch (d.h. ist $m > n$), so sagt man es sei eine *Matrix mit orthonormalen Kolonnen*; sind diese nicht normiert, so spricht man von einer *Matrix mit orthogonalen Kolonnen*.

BEWEIS: Jedes $x \in V$ lässt sich nach Satz 4.12 in der Form

$$x = \sum_{k=1}^n \xi_k b_k$$

darstellen, mit eindeutig bestimmten Koordinaten ξ_k . Bei orthonormierter Basis gilt wegen der Linearität des Skalarproduktes im zweiten Argument

$$\langle b_j, x \rangle = \left\langle b_j, \sum_{k=1}^n \xi_k b_k \right\rangle = \sum_{k=1}^n \xi_k \langle b_j, b_k \rangle = \xi_j.$$

■

BEISPIEL 6.5: Ein **reelles trigonometrisches Polynom vom Grad $\leq m$** ist eine auf ganz \mathbb{R} definierte 2π -periodische Funktion der Form

$$f(t) = \frac{1}{2} \alpha_0 + \alpha_1 \cos(t) + \beta_1 \sin(t) + \dots + \alpha_m \cos(mt) + \beta_m \sin(mt)$$

oder

$$f(t) = \sum_{k=0}^m (\alpha_k \cos(kt) + \beta_k \sin(kt)), \tag{6.17}$$

wobei der Strich nach dem Summenzeichen bedeutet, dass der Summand $\alpha_0 \cos(0t)$ das Gewicht $\frac{1}{2}$ bekommt und der Summand $\beta_0 \sin(0t)$ nicht auftritt. Wir bezeichnen die Menge dieser Polynome mit \mathcal{T}_m . Mit der üblichen, auf $C(\mathbb{R})$ definierten Addition und skalaren Multiplikation ist \mathcal{T}_m offensichtlich ein Unterraum des Vektorraumes $C(\mathbb{R})$, also selbst ein Vektorraum. Man verifiziert leicht, dass durch

$$\langle f, g \rangle := \int_0^{2\pi} f(t) g(t) dt \tag{6.18}$$

auf \mathcal{T}_m ein Skalarprodukt definiert ist. Wir behaupten, dass die $2m + 1$ Funktionen

$$\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos(t), \frac{1}{\sqrt{\pi}} \sin(t), \dots, \frac{1}{\sqrt{\pi}} \cos(mt), \frac{1}{\sqrt{\pi}} \sin(mt) \tag{6.19}$$

eine orthonormale Basis von \mathcal{T}_m bilden. In der Tat kann man mit Hilfe der Additionstheoreme für Sinus und Cosinus die Integrale

$$\int_0^{2\pi} \cos(kt) \cos(lt) dt, \quad \int_0^{2\pi} \sin(kt) \sin(lt) dt,$$

als Summe bzw. Differenz der zwei Integrale

$$\frac{1}{2} \int_0^{2\pi} \cos((k \pm l)t) dt = \delta_{kl} \pi$$

ausdrücken, während für alle k und l gilt

$$\int_0^{2\pi} \cos(kt) \sin(lt) dt = \frac{1}{2} \int_0^{2\pi} \sin((l-k)t) dt + \frac{1}{2} \int_0^{2\pi} \sin((l+k)t) dt = 0.$$

Schreibt man f statt in der Form (6.17) als Linearkombination der normierten Basisfunktionen (6.19), also in der Form

$$f(t) = \frac{\xi'_0}{\sqrt{2\pi}} + \frac{\xi'_1}{\sqrt{\pi}} \cos(t) + \frac{\xi''_1}{\sqrt{\pi}} \sin(t) + \cdots + \frac{\xi'_m}{\sqrt{\pi}} \cos(mt) + \frac{\xi''_m}{\sqrt{\pi}} \sin(mt), \quad (6.20)$$

so lassen sich die Koeffizienten analog zur Formel $\xi_k = \langle \beta_k, x \rangle$ aus (6.16) bestimmen: zum Beispiel ist für $k = 1, \dots, m$

$$\xi'_k = \frac{1}{\sqrt{\pi}} \langle \cos(kt), f(t) \rangle = \frac{1}{\sqrt{\pi}} \int_0^{2\pi} f(t) \cos(kt) dt.$$

Da $\xi'_k/\sqrt{\pi} = \alpha_k$ ist und sich β_k analog berechnen lässt, folgt schliesslich für die Koeffizienten in (6.17):

$$\alpha_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos(kt) dt, \quad \beta_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin(kt) dt, \quad (6.21)$$

wobei die erste Formel auch für α_0 gilt. ◆

Ist $V = \mathbb{E}^n$ und $\{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ eine orthonormierte Basis, so schreibt man (6.16) vorzugweise wieder so, dass die Skalare auf der rechten Seite der Vektoren stehen. Auf Grund der Regeln der Matrizenrechnung gilt dann für jedes $\mathbf{x} \in \mathbb{E}^n$

$$\mathbf{x} = \sum_{k=1}^n \mathbf{b}_k (\mathbf{b}_k^H \mathbf{x}) = \sum_{k=1}^n \mathbf{b}_k \mathbf{b}_k^H \mathbf{x} = \left(\sum_{k=1}^n \mathbf{b}_k \mathbf{b}_k^H \right) \mathbf{x}. \quad (6.22)$$

Damit ergibt sich folgende additive Zerlegung der Einheitsmatrix in eine Summe von Rang-1-Matrizen

$$\sum_{k=1}^n \mathbf{b}_k \mathbf{b}_k^H = \mathbf{I}_n. \quad (6.23)$$

Sie bedeutet, dass die Identität als Summe von Projektionen auf die Koordinatenachsen dargestellt werden kann.

In Vektorräumen mit Orthonormalbasis sind die Vektoren gleich lang wie die entsprechenden Koordinatenvektoren. Laut der folgenden Parsevalschen Formel⁴ gilt allgemeiner das Analoge für die Skalarprodukte in V und \mathbb{E}^n :

⁴MARC-ANTOINE DE PARSEVAL DES CHÊSNES (27.4.1755 – 16.8.1836), französischer Mathematiker, in Kontakt mit Fourier, Poisson, Bessel; Royalist.

Satz 6.5 (Parsevalsche Formel) *Unter den Voraussetzungen von Satz 6.4 gilt mit $\xi_k := \langle b_k, x \rangle$ und $\eta_k := \langle b_k, y \rangle$ ($k = 1, \dots, n$):*

$$\langle x, y \rangle = \sum_{k=1}^n \overline{\xi_k} \eta_k = \boldsymbol{\xi}^H \boldsymbol{\eta} = \langle \boldsymbol{\xi}, \boldsymbol{\eta} \rangle, \quad (6.24)$$

das heisst das Skalarprodukt zweier Vektoren in V ist gleich dem (Euklidischen) Skalarprodukt ihrer Koordinatenvektoren im \mathbb{E}^n . Insbesondere gilt:

$$\|x\| = \|\boldsymbol{\xi}\|, \quad (6.25)$$

$$\angle(x, y) = \angle(\boldsymbol{\xi}, \boldsymbol{\eta}), \quad (6.26)$$

$$x \perp y \iff \boldsymbol{\xi} \perp \boldsymbol{\eta}. \quad (6.27)$$

BEWEIS: Man stellt x und y mittels Formel (6.16) dar und wendet die Eigenschaften (S1) und (S2) an:

$$\langle x, y \rangle = \left\langle \sum_{k=1}^n \xi_k b_k, \sum_{l=1}^n \eta_l b_l \right\rangle = \sum_{k=1}^n \sum_{l=1}^n \overline{\xi_k} \eta_l \underbrace{\langle b_k, b_l \rangle}_{\delta_{kl}} = \sum_{k=1}^n \overline{\xi_k} \eta_k.$$

Die drei weiteren Formeln folgen unmittelbar aufgrund der Definitionen der auftretenden Grössen. ■

Noch haben wir nicht gezeigt, dass es in einem Vektorraum mit Skalarprodukt immer eine orthonormierte Basis gibt. Dass dies bei endlicher oder abzählbar-unendlicher Basis so ist, zeigt der folgende Algorithmus von Gram-Schmidt⁵. Wir wissen aus Lemma 4.6 und Satz 4.11, dass es in jedem Vektorraum eine Basis gibt. Der Algorithmus erlaubt uns, diese durch eine orthogonale Basis zu ersetzen.

Algorithmus 6.1 (Gram-Schmidt-Orthogonalisierungsverfahren)

Es sei $\{a_1, a_2, \dots\}$ eine endliche oder abzählbare, linear unabhängige Menge von Vektoren. Wir berechnen eine gleich grosse Menge $\{b_1, b_2, \dots\}$ von Vektoren rekursiv gemäss

$$\left. \begin{aligned} b_1 &:= \frac{a_1}{\|a_1\|}, \\ \tilde{b}_k &:= a_k - \sum_{j=1}^{k-1} \langle b_j, a_k \rangle b_j, \\ b_k &:= \frac{\tilde{b}_k}{\|\tilde{b}_k\|} \end{aligned} \right\} (k = 2, 3, \dots). \quad (6.28)$$

⁵JORGEN PEDERSEN GRAM (27.6.1850 – 29.4.1916), dänischer Versicherungsmathematiker.

ERHARD SCHMIDT (14.1.1876 – 6.12.1959), aus dem Baltikum stammender deutscher Mathematiker, Schüler von H.A. Schwarz und D. Hilbert; 1908 Professor in Zürich, 1910 in Erlangen, 1911 in Breslau, 1917 in Berlin.

Satz 6.6 Die im Gram–Schmidt–Orthogonalisierungsverfahren konstruierten Vektoren b_1, b_2, \dots sind normiert und paarweise orthogonal, und es gilt nach k Schritten

$$\text{span} \{a_1, a_2, \dots, a_k\} = \text{span} \{b_1, b_2, \dots, b_k\}.$$

Ist $\{a_1, a_2, \dots\}$ eine Basis von V , so ist $\{b_1, b_2, \dots\}$ eine Orthonormalbasis von V .

Statt Gram–Schmidt–Orthogonalisierungsverfahren werden auch die Bezeichnungen **Schmidtsches Orthogonalisierungsverfahren** (vor allem in der deutschsprachigen Mathematik-Literatur) und **klassisches Gram–Schmidt–Verfahren** [*classical Gram-Schmidt algorithm*] (vor allem im wissenschaftlichen Rechnen) verwendet. Eine in Bezug auf die Rundungsfehler etwas genauere Version heisst **modifiziertes Gram–Schmidt–Verfahren** [*modified Gram-Schmidt algorithm*].

Das Verfahren wird vor allem, aber nicht nur, im Falle $V = \mathbb{E}^n$ verwendet. In diesem Fall gibt es allerdings diverse Alternativen. Wir werden sehen, dass man das Gram–Schmidt–Verfahren nicht nur für die Konstruktion einer Orthonormalbasis brauchen kann.

BEWEIS von Satz 6.6: Es ist klar, dass $\|b_k\| = 1$ ($\forall k$). Wir müssen zeigen, dass $\langle b_l, b_k \rangle = 0$ falls $k \neq l$. Wir dürfen die Induktionsvoraussetzung machen, dass b_1, \dots, b_{k-1} paarweise orthogonal sind. (Falls $k = 1$ ist die Induktionsvoraussetzung leer; es ist deshalb keine eigentliche Verankerung nötig.) Dann folgt für $l = 1, \dots, k - 1$:

$$\begin{aligned} \langle b_l, b_k \rangle &= \|\tilde{b}_k\|^{-1} \langle b_l, \tilde{b}_k \rangle = \|\tilde{b}_k\|^{-1} \left[\langle b_l, a_k \rangle - \underbrace{\sum_{j=1}^{k-1} \langle b_j, a_k \rangle \langle b_l, b_j \rangle}_{= \langle b_l, a_k \rangle} \right] \\ &= \|\tilde{b}_k\|^{-1} [\langle b_l, a_k \rangle - \langle b_l, a_k \rangle] = 0. \end{aligned}$$

Also sind sogar b_1, \dots, b_k paarweise orthogonal, und mit Induktion schließen wir auf die paarweise Orthogonalität aller b_j . Durch Induktion folgt sofort auch, dass $b_k \in \text{span} \{a_1, \dots, a_k\}$, also

$$\text{span} \{b_1, \dots, b_k\} \subseteq \text{span} \{a_1, \dots, a_k\} \quad (\forall k).$$

Nach Satz 6.3 sind b_1, \dots, b_k linear unabhängig, bilden also eine Basis von $\text{span} \{b_1, \dots, b_k\}$. Aus Dimensionsgründen folgt daraus direkt, dass $\text{span} \{b_1, \dots, b_k\} = \text{span} \{a_1, \dots, a_k\}$, vgl. Satz 4.7. ■

Aufgrund der letzten Aussage in Satz 6.6 gilt insbesondere:

Korollar 6.7 Zu einem Vektorraum mit Skalarprodukt von endlicher oder abzählbar unendlicher Dimension gibt es eine orthonormierte Basis.

BEISPIEL 6.6: Die Monome $a_k : t \mapsto t^{k-1}$ ($k = 1, 2, \dots$) bilden ja eine Basis des Vektorraumes \mathcal{P} aller Polynome mit reellen Koeffizienten. (Man könnte stattdessen auch komplexe zulassen.) Diese Basis ist aber nicht orthonormal bezüglich des Skalarproduktes

$$\langle p, q \rangle := \int_{-1}^1 p(t) q(t) dt,$$

denn zum Beispiel ist $\langle t^2, 1 \rangle = \frac{2}{3} \neq 0$. Wir können diese Basis aber mit dem Gram–Schmidt–Verfahren orthonormieren:

$$\|a_1\|^2 = \int_{-1}^1 1 dt = t \Big|_{-1}^1 = 2, \quad b_1(t) = \frac{a_1(t)}{\|a_1\|} = \frac{1}{\sqrt{2}},$$

$$\langle b_1, a_2 \rangle = \left\langle \frac{1}{\sqrt{2}}, t \right\rangle = \int_{-1}^1 \frac{t}{\sqrt{2}} dt = \frac{t^2}{2\sqrt{2}} \Big|_{-1}^1 = 0,$$

$$\tilde{b}_2(t) = a_2(t) - \langle b_1, a_2 \rangle b_1(t) = t - 0 = t,$$

$$\|\tilde{b}_2\|^2 = \int_{-1}^1 t^2 dt = \frac{t^3}{3} \Big|_{-1}^1 = \frac{2}{3}, \quad b_2(t) = \frac{\tilde{b}_2(t)}{\|\tilde{b}_2\|} = \sqrt{\frac{3}{2}} t,$$

$$\langle b_1, a_3 \rangle = \left\langle \frac{1}{\sqrt{2}}, t^2 \right\rangle = \frac{2}{3} \frac{1}{\sqrt{2}}, \quad \langle b_2, a_3 \rangle = \left\langle \sqrt{\frac{3}{2}} t, t^2 \right\rangle = 0,$$

$$\tilde{b}_3(t) = a_3(t) - \langle b_1, a_3 \rangle b_1(t) - \langle b_2, a_3 \rangle b_2(t) = t^2 - \frac{2}{3} \frac{1}{\sqrt{2}} \frac{1}{\sqrt{2}} = t^2 - \frac{1}{3}$$

$$\|\tilde{b}_3\|^2 = \int_{-1}^1 \left(t^2 - \frac{1}{3}\right)^2 dt = \frac{t^5}{5} - \frac{2t^3}{9} \frac{t}{9} \Big|_{-1}^1 = \frac{8}{45},$$

$$b_3(t) = \sqrt{\frac{45}{8}} \left(t^2 - \frac{1}{3}\right), \quad \text{usw.}$$

Auf diese Art entsteht eine Folge $\{b_k\}_{k=1}^\infty$ orthonormaler Polynome mit Grad $b_k = k - 1$. Die umskalierten orthogonalen Polynome $P_k(t) := b_{k+1}(t)/b_{k+1}(1)$ heissen **Legendre–Polynome**⁶ [*Legendre polynomials*].

Auf analoge Art erhält man bei Benützung des Skalarproduktes

$$\langle p, q \rangle := \int_{-1}^1 \frac{p(t) q(t)}{\sqrt{1-t^2}} dt$$

die **Tschebyscheff–Polynome**⁷ [*Chebyshev polynomials*]. Allgemeiner kann man zu jedem endlichen oder unendlichen Intervall und jeder dazu passenden positiven Gewichtsfunktion eine Folge von **Orthogonal–polynomen** [*orthogonal polynomials*] definieren. Zu deren Konstruktion gibt es aber einen effizienteren Algorithmus als das Verfahren von Gram–Schmidt, denn drei aufeinanderfolgende Polynome sind immer durch eine Rekursionsformel verknüpft. \blacklozenge

⁶ADRIEN MARIE LEGENDRE (18.9.1752 – 10.1.1833), französischer Mathematiker, ab 1755 Professor an der École Militaire in Paris, später an der École Normale.

⁷PAFNUTI LWOWITSCH TSCHEBYSCHEFF (14.5.1821 – 26.11.1894), russischer Mathematiker, ab 1857 Professor in St. Petersburg.

AUSBLICK: Wie wir schon in Kapitel 4, Satz 4.11, erwähnt haben, gibt es in jedem Vektorraum (ausser $\{o\}$) eine Basis gemäss unserer Definition, aber diese Definition führt nur in Räumen wie ℓ_0 und \mathcal{P} mit abzählbarer Basis weiter, nicht aber für die „grossen“ Funktionenräume, wie ℓ oder $C[a, b]$.

Man verwendet für die meisten unendlich-dimensionalen Räume in der Mathematik eine andere Definition von Erzeugendensystem und Basis. Die Vektoren werden als *unendliche* Linearkombinationen der Basisvektoren dargestellt, d.h. als *Reihe*. Dabei müssen diese Reihen in einem gewissen Sinne konvergieren. Notwendig für eine Definition der Konvergenz ist, dass ein Mass für die Differenz (ein Abstand) zweier Vektoren (bzw. Funktionen) eingeführt werden kann oder zumindest ein Umgebungsbegriff. Ein entsprechender Vektorraum heisst⁸

- i) **topologischer Vektorraum** [*topological vector space*] (mit Umgebungsbegriff),
- ii) **Banachraum**⁹ [*Banach space*] (mit Vektornorm) oder
- iii) **Hilbertraum**¹⁰ [*Hilbert space*] (mit Skalarprodukt und zugehöriger Vektornorm).

In den wichtigen Beispielen sind die in diesen Räumen verwendeten Basen abzählbar unendlich gross, aber i.a. viel „kleiner“ als eine Vektorraumbasis, weil jeder Vektor ja nur als unendliche Linearkombination von Basisvektoren darstellbar sein muss, nicht als endliche Linearkombination.

BEISPIEL 6.7: Der Hilbertraum ℓ^2 der **quadrat-summierbaren Folgen** [*square summable series*] ist definiert als Vektorraum der Folgen

$$x = \{x_k\}_{k=0}^{\infty} \quad \text{mit} \quad \sum_{k=0}^{\infty} |x_k|^2 < \infty \quad (6.29)$$

mit dem Skalarprodukt

$$\langle x, y \rangle = \sum_{k=0}^{\infty} \overline{x_k} y_k \quad (6.30)$$

und der Norm

$$\|x\| = \sqrt{\sum_{k=0}^{\infty} |x_k|^2}. \quad (6.31)$$

Die Zahlfolgen e_k ($1 \leq k < \infty$) aus Beispiel 4.23 bilden eine Hilbertraum-Basis: jedes $x \in \ell^2$ ist ja eine unendliche Linearkombination dieser Basisvektoren: $x = \sum x_k e_k$. \blacklozenge

⁸Bei Banach- und Hilberträumen wird noch verlangt, dass sie **vollständig** [*complete*] sind, das heisst, dass mit jeder konvergierenden Folge von Vektoren auch deren Grenzwert im Raum liegt.

⁹STEFAN BANACH (30.3.1892 – 31.8.1945), polnischer Mathematiker, ab 1922 Professor in Lwów.

¹⁰DAVID HILBERT (23.1.1862 – 14.2.1943), deutscher Mathematiker, ab 1895 Professor in Göttingen, das dank ihm neben Berlin zum zweiten deutschen mathematischen Zentrum wurde.

BEISPIEL 6.8: Der Hilbertraum L^2 der **quadrat-integrierbaren 2π -periodischen Funktionen** [*square integrable 2π -periodic functions*] ist definiert als Menge der 2π -periodischen Funktionen¹¹

$$f : \mathbb{R} \longrightarrow \mathbb{E} \quad \text{mit} \quad \int_{k=0}^{2\pi} |f(t)|^2 dt < \infty \quad (6.32)$$

mit dem Skalarprodukt und der Norm

$$\langle f, g \rangle = \int_0^{2\pi} \overline{f(t)} g(t) dt, \quad \|f\| = \sqrt{\int_0^{2\pi} |f(t)|^2 dt}. \quad (6.33)$$

Falls $\mathbb{E} = \mathbb{R}$, bilden die Funktionen

$$\{\cos kt\}_{k=0}^{\infty} \quad \text{und} \quad \{\sin kt\}_{k=1}^{\infty} \quad (6.34)$$

zusammen eine Basis. Im komplexen Fall, $\mathbb{E} = \mathbb{C}$, verwendet man stattdessen die Funktionen

$$\{e^{ikt}\}_{k=-\infty}^{\infty}. \quad (6.35)$$

Die entsprechenden Reihendarstellungen einer Funktion $f \in L^2$ heißen **Fourier-Reihen**¹² [*Fourier series*]

$$f(t) = \alpha_0 + \sum_{k=1}^{\infty} (\alpha_k \cos kt + \beta_k \sin kt) \quad (6.36)$$

bzw.

$$f(t) = \sum_{k=-\infty}^{\infty} \gamma_k e^{ikt}. \quad (6.37)$$

Dabei gilt allerdings an Unstetigkeitsstellen von f das Gleichheitszeichen in der Regel nicht.

Die Fourier-Reihen (6.36) und (6.37) sind Darstellungen der Art (6.16), aber mit einer unendlichen statt einer endlichen Reihe. In der reellen Version (6.36) gelten wie in Beispiel 6.5 für die **Fourier-Koeffizienten** [*Fourier coefficients*] a_k und b_k die Formeln (6.19). Jene für die Koeffizienten γ_k der komplexen Version (6.37) sind noch einfacher:

$$\gamma_k = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-ikt} dt. \quad (6.38)$$

Die Bedeutung der Fourier-Reihen liegt darin, dass man fast jede periodische Funktion f durch eine Fourier-Reihe darstellen kann, die “fast überall” (in einem exakt definierbaren Sinn) gegen f konvergiert.

Zu (6.36) und (6.37) analoge Reihenentwicklungen (Darstellungen in einer unendlichen Basis) für die Vektoren gibt es auch in allen anderen Hilberträumen mit abzählbar unendlicher Basis. Man nennt sie dort **verallgemeinerte Fourier-Reihen** [*generalized Fourier series*]. ◆
▼

¹¹Das Integral ist allerdings hier nicht ein Riemann-Integral sondern ein allgemeineres Lebesgue-Integral. Die Funktionen brauchen nicht stetig zu sein, sondern nur im Sinne von Lebesgue integrierbar.

HENRI LEBESGUE (28.6.1875 – 26.7.1941), französischer Mathematiker, der bereits in seiner Dissertation eine neue Theorie des Masses aufstellte, die zum Lebesgue-Integral führte. Professor am Collège de France.

¹²JEAN-BAPTISTE-JOSEPH FOURIER (21.3.1768 – 16.5.1830), französischer Mathematiker, 1794 Professor an der École Normale, später Napoleons Präfekt in Grenoble.

6.4 Orthogonale Komplemente

Wenn wir den Satz 6.6 über das Gram–Schmidt–Verfahren mit dem früheren Satz 4.11 kombinieren, gemäss dem eine linear unabhängige Menge von Vektoren zu einer Basis ergänzt werden kann, folgt sofort:

Korollar 6.8 *In einem Vektorraum endlicher Dimension mit Skalarprodukt kann man jede Menge orthonormaler Vektoren zu einer orthonormalen Basis ergänzen.*

Das gilt auch noch für Vektorräume mit abzählbar unendlicher Basis, wie z.B. den Raum \mathcal{P} aller Polynome, aber wir wollen uns in diesem Abschnitt der Einfachheit halber auf Räume endlicher Dimension beschränken, wo gewisse Schwierigkeiten nicht auftreten.

Bezeichnen wir wie in (4.32) und (4.33) die gegebene Menge orthonormaler Vektoren mit $M = \{b_1, \dots, b_l\}$ und die Ergänzungsmenge mit $N = \{b_{l+1}, \dots, b_n\}$, so gilt für die von diesen Mengen aufgespannten Unterräume

$$\begin{aligned} U &::= \text{span } M = \text{span } \{b_1, \dots, b_l\}, \\ U' &::= \text{span } N = \text{span } \{b_{l+1}, \dots, b_n\}, \end{aligned} \quad (6.39)$$

gemäss (4.36) dass $V = U \oplus U'$ gilt, aber zusätzlich ist $U \perp U'$. Stellen wir irgend ein $x \in V$ in der Basis dar, so ist

$$x = \underbrace{\xi_1 b_1 + \dots + \xi_l b_l}_{\equiv: u \in U} + \underbrace{\xi_{l+1} b_{l+1} + \dots + \xi_n b_n}_{\equiv: u' \in U'} = u + u'$$

mit

$$u \perp U', \quad u' \perp U.$$

Man bezeichnet den Unterraum U' deshalb oft mit U^\perp . Offensichtlich ist er eindeutig bestimmt, denn er enthält genau alle zu U orthogonale Vektoren von V .

DEFINITION: In einem endlich-dimensionalen Vektorraums V mit Skalarprodukt heisst der zu einem echten Unterraum U orthogonale komplementäre Unterraum das **orthogonale Komplement** [*orthogonal complement*] von U und wird mit U^\perp („**U senkrecht**“ [*“U perp”*]) bezeichnet. Es wird implizit charakterisiert durch die Beziehungen

$$V = U \oplus U^\perp, \quad U \perp U^\perp \quad (6.40)$$

oder explizit durch

$$U^\perp ::= \{x \in V; x \perp U\}. \quad (6.41)$$

Wir nennen dann V eine **direkte Summe orthogonaler Komplemente** [*direct sum of orthogonal complements*]. ▲

Offensichtlich ist

$$(U^\perp)^\perp = U \quad (6.42)$$

und

$$\dim U + \dim U^\perp = \dim V. \quad (6.43)$$

Die letzte Beziehung gilt analog auch für nicht-orthogonale komplementäre Unterräume. Wie bei der allgemeinen direkten Summe von Unterräumen lässt sich der Begriff der Summe orthogonaler Unterräume sofort auf mehrere Terme ausdehnen. Zum Beispiel könnte man oben den Unterraum U selbst wieder in zwei orthogonale Komplemente aufteilen, falls $\dim U > 1$ ist.

BEISPIEL 6.9: Ist b_1, \dots, b_n eine orthogonale Basis von V und setzt man $U_k := \text{span} \{b_k\}$, so gilt

$$V = U_1 \oplus U_2 \oplus \dots \oplus U_n, \quad U_j \perp U_k \quad \text{for } j \neq k, \quad (6.44)$$

das heisst V ist die direkte Summe von n zueinander orthogonalen eindimensionalen Unterräumen aufgespannt durch die einzelnen Basisvektoren. \blacklozenge

Eine interessante Anwendung ergibt sich im Zusammenhang mit Kern (Nullraum) und Bild (Kolonnenraum) einer Matrix. Es sei $\mathbf{A} \in \mathbb{E}^{m \times n}$ eine beliebige Matrix und $\mathbf{x} \in \mathbb{E}^n$. Es gilt $\mathbf{A}\mathbf{x} = \mathbf{o}$ genau dann, wenn \mathbf{x} auf allen Zeilen von \mathbf{A} (bzw. von $\overline{\mathbf{A}}$ falls $\mathbb{E} = \mathbb{C}$) senkrecht steht, das heisst wenn \mathbf{x} auf allen Kolonnen von \mathbf{A}^\top (bzw. \mathbf{A}^H) senkrecht steht. Wir haben also, falls $\mathbb{E} = \mathbb{R}$,

$$\mathbf{x} \in \mathcal{N}(\mathbf{A}) \iff \mathbf{A}\mathbf{x} = \mathbf{o} \iff \mathbf{x} \perp \mathcal{R}(\mathbf{A}^\top) \iff \mathbf{x} \in (\mathcal{R}(\mathbf{A}^\top))^\perp$$

und, falls $\mathbb{E} = \mathbb{C}$,

$$\mathbf{x} \in \mathcal{N}(\mathbf{A}) \iff \mathbf{A}\mathbf{x} = \mathbf{o} \iff \mathbf{x} \perp \mathcal{R}(\mathbf{A}^H) \iff \mathbf{x} \in (\mathcal{R}(\mathbf{A}^H))^\perp.$$

Erinnern wir uns noch daran, dass nach Korollar 5.14 gilt $\text{Rang } \mathbf{A} = \text{Rang } \mathbf{A}^H = \text{Rang } \mathbf{A}^\top$, so erhalten wir:

Satz 6.9 Für eine komplexe $m \times n$ -Matrix mit Rang r gilt

$$\boxed{\mathcal{N}(\mathbf{A}) = (\mathcal{R}(\mathbf{A}^H))^\perp \subset \mathbb{E}^n}, \quad \boxed{\mathcal{N}(\mathbf{A}^H) = (\mathcal{R}(\mathbf{A}))^\perp \subset \mathbb{E}^m}. \quad (6.45)$$

Insbesondere ist

$$\boxed{\mathcal{N}(\mathbf{A}) \oplus \mathcal{R}(\mathbf{A}^H) = \mathbb{E}^n}, \quad \boxed{\mathcal{N}(\mathbf{A}^H) \oplus \mathcal{R}(\mathbf{A}) = \mathbb{E}^m} \quad (6.46)$$

und

$$\begin{aligned} \dim \mathcal{R}(\mathbf{A}) &= r, & \dim \mathcal{N}(\mathbf{A}) &= n - r, \\ \dim \mathcal{R}(\mathbf{A}^H) &= r, & \dim \mathcal{N}(\mathbf{A}^H) &= m - r. \end{aligned} \quad (6.47)$$

Bei reellen Matrizen kann überall \mathbf{A}^H durch \mathbf{A}^\top ersetzt werden.

DEFINITION: Die zwei Paare komplementärer Unterräume $\mathcal{N}(\mathbf{A})$, $\mathcal{R}(\mathbf{A}^H)$ und $\mathcal{N}(\mathbf{A}^H)$, $\mathcal{R}(\mathbf{A})$ nennt man die vier **fundamentalen Unterräume** [fundamental subspaces] der Matrix \mathbf{A} . \blacktriangle

6.5 Orthogonale und unitäre Basiswechsel und Koordinatentransformationen

In einem Vektorraum V mit Skalarprodukt interessiert man sich naturgemäß für Basiswechsel zwischen orthonormierten Basen. Wie in §3.4 stellen wir zunächst die “neue” Basis $\{b'_1, \dots, b'_n\}$ in der “alten” Basis $\{b_1, \dots, b_n\}$ dar:

$$\boxed{b'_k = \sum_{i=1}^n \tau_{ik} b_i}, \quad k = 1, \dots, n. \quad (6.48)$$

Bei einer “gewöhnlichen” Basistransformation ist die Transformationsmatrix $\mathbf{T} = (\tau_{ik})$ irgend eine reguläre $n \times n$ -Matrix (vgl. Satz 4.13). Welche spezielle Eigenschaft hat \mathbf{T} wohl, wenn beide Basen orthonormiert sind? Es gilt dann

$$\begin{aligned} \delta_{kl} &= \langle b'_k, b'_l \rangle = \left\langle \sum_{i=1}^n \tau_{ik} b_i, \sum_{j=1}^n \tau_{jl} b_j \right\rangle \\ &= \sum_{i=1}^n \overline{\tau_{ik}} \sum_{j=1}^n \tau_{jl} \underbrace{\langle b_i, b_j \rangle}_{\delta_{ij}} = \sum_{i=1}^n \overline{\tau_{ik}} \tau_{il}, \end{aligned}$$

das heisst

$$\boxed{\mathbf{I} = \mathbf{T}^H \mathbf{T}.} \quad (6.49)$$

Satz 6.10 Die Transformationsmatrix einer Basistransformation zwischen orthonormierten Basen ist unitär (falls $\mathbb{E} = \mathbb{C}$) bzw. orthogonal (falls $\mathbb{E} = \mathbb{R}$).

Unitäre und orthogonale Koordinatentransformationen sind sehr praktisch, weil $\mathbf{T}^{-1} = \mathbf{T}^H$ ist, also die inverse Transformation immer bekannt ist. Damit vereinfacht sich die Formel (4.42) von Satz 4.13.

Satz 6.11 Es sei die durch (6.48) definierte Basistransformation unitär bzw. orthogonal, d.h. es gelte (6.49). Dann sind die alten und die neuen Koordinatenvektoren $\boldsymbol{\xi}$ bzw. $\boldsymbol{\xi}'$ verknüpft gemäss

$$\boxed{\boldsymbol{\xi} = \mathbf{T} \boldsymbol{\xi}', \quad \boldsymbol{\xi}' = \mathbf{T}^H \boldsymbol{\xi}.} \quad (6.50)$$

Falls der zu Grunde liegende Vektorraum V der \mathbb{E}^n ist, kann man die alten und neuen orthonormierten Basisvektoren wie üblich als Kolonnenvektoren unitärer oder orthogonaler Matrizen auffassen:

$$\mathbf{B} := (\mathbf{b}_1 \ \dots \ \mathbf{b}_n), \quad \mathbf{B}' := (\mathbf{b}'_1 \ \dots \ \mathbf{b}'_n). \quad (6.51)$$

Aus den Formeln (6.50) wird dann

$$\boxed{\mathbf{B} = \mathbf{B}' \mathbf{T}^H, \quad \mathbf{B}' = \mathbf{B} \mathbf{T},} \quad (6.52)$$

wobei nun hier alle drei Matrizen unitär bzw. orthogonal sind.

BEWEIS: (6.54) folgt mit Hilfe von (6.50) und (6.49) gemäss

$$\langle \xi, \eta \rangle = \xi^H \eta = (\xi')^H \mathbf{T}^H \mathbf{T} \eta' = (\xi')^H \eta' = \langle \xi', \eta' \rangle .$$

■

Wir kommen nun noch zurück auf die in Abschnitt 5.5 und besonders in Diagramm (5.48) zusammengefasste Situation einer linearen Abbildung $F : X \rightarrow Y$, wobei wir nun voraussetzen wollen, dass X und Y Vektorräume mit Skalarprodukten $(\cdot, \cdot)_X$ bzw. $(\cdot, \cdot)_Y$ über dem gleichen Zahlkörper \mathbb{R} oder \mathbb{C} sind (also beides Euklidische Räume oder beides unitäre Räume). Wir wollen F bezüglich orthonormalen Basen in X und Y darstellen, und dabei von einem ersten Paar gegebener, “alter” Basen auf ein zweites Paar “neuer” orthonormalen Basen übergehen.

$$\begin{array}{ccc}
 x \in X & \xrightarrow[\text{lin. Abb.}]{F} & z \in Y & \text{Vektorräume} \\
 & & & \text{mit Skalarprodukt} \\
 \kappa_X \left\| \begin{array}{c} \uparrow \\ \downarrow \end{array} \right. \kappa_X^{-1} & & \kappa_Y \left\| \begin{array}{c} \uparrow \\ \downarrow \end{array} \right. \kappa_Y^{-1} & \\
 \xi \in \mathbb{E}^n & \xrightarrow[\text{Abb.matrix}]{\mathbf{A}} & \eta \in \mathbb{E}^m & \text{Koordinaten bzgl.} \\
 & & & \text{“alten” Orthonormal-} \\
 & & & \text{basen in } X \text{ und } Y \\
 \mathbf{T}^{-1} = \mathbf{T}^H \left\| \begin{array}{c} \uparrow \\ \downarrow \end{array} \right. \mathbf{T} & & \mathbf{S}^{-1} = \mathbf{S}^H \left\| \begin{array}{c} \uparrow \\ \downarrow \end{array} \right. \mathbf{S} & \\
 \xi' \in \mathbb{E}^n & \xrightarrow[\text{Abb.matrix}]{\mathbf{B}} & \eta' \in \mathbb{E}^m & \text{Koordinaten bzgl.} \\
 & & & \text{“neuen” Orthonormal-} \\
 & & & \text{basen in } X \text{ und } Y \\
 & & & (6.58)
 \end{array}$$

Weil die Matrizen T und S der Koordinatentransformationen orthogonal bzw. unitär sind, lassen sich nun die Formeln (5.50) und (5.52) für die Transformation der Abbildungsmatrix umschreiben in

$$\boxed{\mathbf{B} = \mathbf{S}^H \mathbf{A} \mathbf{T}, \quad \mathbf{A} = \mathbf{S} \mathbf{B} \mathbf{T}^H,} \quad (6.59)$$

bzw. falls $Y = X$, $\kappa_Y = \kappa_X$ und $\mathbf{S} = \mathbf{T}$,

$$\boxed{\mathbf{B} = \mathbf{T}^H \mathbf{A} \mathbf{T}, \quad \mathbf{A} = \mathbf{T} \mathbf{B} \mathbf{T}^H.} \quad (6.60)$$

Im zweiten Falle, wo es um eine Selbstabbildung geht, ist nach (6.60) klar, dass

- wenn \mathbf{A} Hermitesch bzw. reell symmetrisch ist, so auch \mathbf{B} ,
- wenn \mathbf{A} unitär bzw. orthogonal ist, so auch \mathbf{B} .

6.6 Orthogonale und unitäre Abbildungen

Wir wollen nun noch auf spezielle lineare Abbildungen zu sprechen kommen, die nicht nur die linearen Beziehungen, sondern auch das Skalarprodukt zweier Vektoren invariant lassen.

DEFINITION: Es seien X und Y zwei unitäre (bzw. orthogonale) Vektorräume. Eine lineare Abbildung $F : X \rightarrow Y$ heisst **unitär** [*unitary*] (bzw. **orthogonal** [*orthogonal*]), falls für $x, y \in X$ gilt

$$\langle Fx, Fy \rangle_Y = \langle x, y \rangle_X. \quad (6.61)$$

Diese Abbildungen haben eine Reihe interessanter Eigenschaften, die sich leicht herleiten lassen und die im folgenden Satz zusammengefasst sind. Die ersten zwei haben wir schon in Satz 2.21 angetroffen als Eigenschaften einer durch eine orthogonale oder unitäre Matrix definierten Abbildung.

Satz 6.13 *Für eine orthogonale oder unitäre Abbildung $F : X \rightarrow Y$ gilt:*

- (i) $\|Fx\|_Y = \|x\|_X$ ($\forall x \in X$), d.h. F ist **längentreu** [*length preserving*] (oder: **isometrisch** [*isometric*]);
- (ii) $x \perp y \implies Fx \perp Fy$, d.h. F ist **winkeltreu** [*angle preserving*];
- (iii) $\ker F = \{o\}$, d.h. F ist *injektiv*.

Ist $\dim X = \dim Y < \infty$, so gilt zusätzlich:

- (iv) F ist ein *Isomorphismus*.
- (v) Ist $\{b_1, \dots, b_n\}$ eine Orthonormalbasis von X , so ist $\{Fb_1, \dots, Fb_n\}$ eine Orthonormalbasis von Y ;
- (vi) F^{-1} ist unitär (bzw. orthogonal);
- (vii) Die Abbildungsmatrix \mathbf{A} bezüglich orthonormierten Basen in X und Y ist unitär (bzw. orthogonal).

BEWEIS:

(i) und (ii) folgen direkt aus (6.61).

(iii) folgt aus

$$Fx = o \iff \langle Fx, Fx \rangle_Y = 0 \stackrel{(6.61)}{\iff} \langle x, x \rangle_X = 0 \iff x = o.$$

(iv) ergibt sich aus (iii), der Dimensionsformel (5.29) und ihrem Korollar 5.8.

(v) und (vi) folgen aus (iv) und (6.61).

(vii) ist schliesslich eine Konsequenz der folgenden drei Hilfssätze, angewandt auf die aus dem Diagramm (6.58) folgende Relation $\mathbf{A} = \kappa_Y F \kappa_X^{-1}$ (die bereits in (5.25) auftrat). ■

Weitere Aussagen über die Verknüpfung unitärer Abbildungen, die Koordinatenabbildung und die Abbildungsmatrix folgen mühelos:

Lemma 6.14 Sind $F : X \rightarrow Y$ und $G : Y \rightarrow Z$ zwei unitäre (oder orthogonale) Isomorphismen endlich-dimensionaler unitärer (bzw. orthogonaler) Vektorräume, so auch $G \circ F : X \rightarrow Z$.

BEWEIS: Es ist klar, dass $G \circ F$ ein Isomorphismus ist. Zudem gilt (6.61) für F und G , also

$$\langle GFx, GFy \rangle_Z = \langle Fx, Fy \rangle_Y = \langle x, y \rangle_X . \quad \blacksquare$$

Lemma 6.15 Ist V ein n -dimensionaler unitärer (oder orthogonaler) Vektorraum mit Orthonormalbasis, so ist die Koordinatenabbildung $\kappa_V : V \rightarrow \mathbb{C}^n$ (bzw. $V \rightarrow \mathbb{R}^n$) ein unitärer (bzw. orthogonaler) Isomorphismus.

BEWEIS: Wir haben uns bereits bei der Definition (5.23) der Koordinatenabbildung überlegt, dass sie ein Isomorphismus ist. Zudem ist für sie (6.61) identisch mit der Parsevalschen Formel (6.24) aus Satz 6.5, also ist sie ein unitärer (oder orthogonaler) Isomorphismus. \blacksquare

Lemma 6.16 Die Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ (bzw. $\mathbb{R}^{n \times n}$) ist genau dann unitär (bzw. orthogonal), wenn die lineare Abbildung $\mathbf{A} : \mathbb{C}^n \rightarrow \mathbb{C}^n$ (bzw. $\mathbb{R}^n \rightarrow \mathbb{R}^n$) unitär (bzw. orthogonal) ist.

BEWEIS: Die durch \mathbf{A} definierte Abbildung ist unitär, wenn gemäss (6.61) gilt: $\langle \mathbf{A}\mathbf{x}, \mathbf{A}\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$ ($\forall \mathbf{x}, \mathbf{y}$). Hieraus folgt mit $\mathbf{x} = \mathbf{e}_k$ und $\mathbf{y} = \mathbf{e}_l$, dass

$$(\mathbf{A}^H \mathbf{A})_{kl} = \mathbf{e}_k^H \mathbf{A}^H \mathbf{A} \mathbf{e}_l = (\mathbf{A} \mathbf{e}_k)^H (\mathbf{A} \mathbf{e}_l) = \langle \mathbf{A} \mathbf{e}_k, \mathbf{A} \mathbf{e}_l \rangle = \langle \mathbf{e}_k, \mathbf{e}_l \rangle = \delta_{kl} ,$$

was heisst, dass die Matrix \mathbf{A} unitär ist. Umgekehrt gilt diese Formel genau dann, wenn $\langle \mathbf{A} \mathbf{e}_k, \mathbf{A} \mathbf{e}_l \rangle = \langle \mathbf{e}_k, \mathbf{e}_l \rangle$ ($\forall k, l$). Schreibt man $\mathbf{x} = \sum \mathbf{e}_k \xi_k$, $\mathbf{y} = \sum \mathbf{e}_l \eta_l$, so folgt aufgrund der Eigenschaften (6.6), dass auch $\langle \mathbf{A}\mathbf{x}, \mathbf{A}\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$ ($\forall \mathbf{x}, \mathbf{y}$) gilt, also die Abbildung unitär ist. \blacksquare

6.7 Normen von linearen Abbildungen (Operatoren) und Matrizen

DEFINITION: Es seien X und Y zwei normierte Vektorräume mit den Normen $\|\cdot\|_X$ und $\|\cdot\|_Y$. Eine lineare Abbildung (oder: ein linearer Operator) $F : X \rightarrow Y$ heisst **beschränkt** [*bounded*], wenn es ein $\gamma_F \geq 0$ gibt mit

$$\|F(x)\|_Y \leq \gamma_F \|x\|_X \quad (\forall x \in X) .$$

Die Gesamtheit solcher linearer Abbildungen (Operatoren) F zwischen X und Y heisst $\mathcal{L}(X, Y)$. \blacktriangle

BEMERKUNG: Die zwei Normen in X und Y können gemäss (6.7) durch Skalarprodukte $\langle \cdot, \cdot \rangle_X$ und $\langle \cdot, \cdot \rangle_Y$ definiert sein, müssen es aber nicht. \blacktriangledown

In $\mathcal{L}(X, Y)$ können wir auf natürliche Art eine Addition und eine skalare Multiplikation definieren:

$$\begin{aligned} F + G : x \in X &\mapsto (F + G)(x) := Fx + Gx \\ &(\forall F, G \in \mathcal{L}(X, Y)), \\ \alpha F : x \in X &\mapsto (\alpha F)(x) := \alpha Fx \\ &(\forall \alpha \in \mathbb{E}, \forall F \in \mathcal{L}(X, Y)). \end{aligned} \tag{6.62}$$

Damit wird $\mathcal{L}(X, Y)$ selbst zu einem Vektorraum, wie man leicht verifiziert.

BEISPIEL 6.11: Die $m \times n$ Matrizen können als Elemente von $\mathcal{L}(\mathbb{E}^n, \mathbb{E}^m)$ aufgefasst werden. Dabei sind Addition und skalare Multiplikation wie in Kapitel 2 durch $\mathbf{A} + \mathbf{B}$ und $\alpha \mathbf{A}$ definiert. \blacklozenge

Nun werden wir im Vektorraum $\mathcal{L}(X, Y)$ eine Norm definieren.

DEFINITION: Die auf $\mathcal{L}(X, Y)$ durch die Normen $\|\cdot\|_X$ und $\|\cdot\|_Y$ **induzierte Operatornorm** [*operator norm*] ist definiert durch

$$\|\cdot\| : \mathcal{L}(X, Y) \rightarrow \mathbb{R}, \quad F \mapsto \|F\| := \sup_{x \neq \mathbf{o}} \frac{\|Fx\|_Y}{\|x\|_X}. \tag{6.63}$$

Ist $\mathbf{X} = \mathbf{Y} = \mathbb{E}^n$, so dass F durch eine quadratische Matrix \mathbf{A} gegeben ist, heisst $\|\mathbf{A}\|$ die durch die Vektornorm (in \mathbb{E}^n) **induzierte Matrixnorm** von \mathbf{A} . Sie ist also definiert durch

$$\|\cdot\| : \mathbb{E}^{n \times n} \rightarrow \mathbb{R}, \quad \mathbf{A} \mapsto \|\mathbf{A}\| := \sup_{\mathbf{x} \neq \mathbf{o}} \frac{\|\mathbf{A}\mathbf{x}\|}{\|\mathbf{x}\|}. \tag{6.64}$$

Verwendet man in \mathbb{E}^n die Euklidische 2-Norm, so heisst die induzierte Matrixnorm **Spektralnorm** [*spectral norm*] oder **2-Norm** [*2-norm*]; sie wird auch mit $\|\cdot\|_2$ bezeichnet:

$$\|\mathbf{A}\|_2 := \sup_{\mathbf{x} \neq \mathbf{o}} \frac{\|\mathbf{A}\mathbf{x}\|_2}{\|\mathbf{x}\|_2}. \tag{6.65}$$

Statt “die von den Vektornormen induzierte Operator- oder Matrixnorm” sagt man auch “die den Vektornormen **untergeordnete Operatornorm** bzw. **Matrixnorm** [*subordinate matrix norm*]”. \blacktriangle

Die Definitionen (6.63)–(6.65) lassen sich vereinfachen: führt man zum Beispiel im Bruch in (6.63) $\alpha := \|x\|_X$ und $\tilde{x} := x/\alpha$ ein, so sieht man, dass wegen $\|\tilde{x}\| = 1$ und

$$\|Fx\|_Y = \left\| \alpha F \left(\frac{x}{\alpha} \right) \right\|_Y = \alpha \left\| F \left(\frac{x}{\alpha} \right) \right\|_Y = \|x\|_X \|F\tilde{x}\|_Y$$

folgendes gilt (wir ersetzen \tilde{x} wieder durch x):

Lemma 6.17 Die in (6.63)–(6.65) gegebenen Formeln für $\|F\|$, $\|\mathbf{A}\|$, $\|\mathbf{A}\|_2$ lassen sich ersetzen durch

$$\|F\| = \sup_{\|x\|_X=1} \|Fx\|_Y, \quad (6.66)$$

$$\|\mathbf{A}\| = \sup_{\|\mathbf{x}\|=1} \|\mathbf{A}\mathbf{x}\|, \quad (6.67)$$

$$\|\mathbf{A}\|_2 = \sup_{\|\mathbf{x}\|_2=1} \|\mathbf{A}\mathbf{x}\|_2. \quad (6.68)$$

In dieser Form lässt sich die Operatornorm leicht geometrisch interpretieren: Die Restriktion $\|x\|_X = 1$ bzw. $\|\mathbf{x}\|_2 = 1$ bedeutet ja, dass wir nur Vektoren auf der Oberfläche der “Einheitskugel” in X oder \mathbb{E}^n betrachten. Die Norm entspricht dann der maximalen Länge der Bildvektoren dieser Sphäre. Diese Interpretation passt vor allem für die Spektralnorm. Man muss aber daran denken, dass bei Verwendung einer anderen als der 2-Norm in \mathbb{E}^n die “Einheitskugel” nicht eine Kugel ist.

Die Spektralnorm einer Matrix oder einer linearen Abbildung ist leider im allgemeinen nicht einfach zu berechnen, aber oft benötigt man effektiv nur eine obere Schranke, und eine solche kann man häufig einfach erhalten. Auf die Berechnung der Spektralnorm werden wir in den Abschnitten 10.6 und 11.2 zurückkommen. Wie die folgende Betrachtung zeigt, ist die Berechnung einfach, wenn \mathbf{A} eine Diagonalmatrix ist.

BEISPIEL 6.12: Für eine Diagonalmatrix $\mathbf{D} = \text{diag}(d_{11}, \dots, d_{nn})$ ist

$$\frac{\|\mathbf{D}\mathbf{x}\|_2^2}{\|\mathbf{x}\|_2^2} = \frac{\sum_{k=1}^n |d_{kk}|^2 |x_k|^2}{\sum_{k=1}^n |x_k|^2}$$

ein gewichtetes Mittel der Betragsquadrate der Diagonalelemente d_{kk} und dieses Mittel wird offensichtlich maximal, wenn man alles Gewicht auf ein betragsmäßig größtes Element legt. Ist dies d_{ll} , so wählen wir also $\mathbf{x}_l = \mathbf{e}_l$ und erhalten

$$\max_{\mathbf{x} \neq \mathbf{o}} \frac{\|\mathbf{D}\mathbf{x}\|_2^2}{\|\mathbf{x}\|_2^2} = \frac{\|\mathbf{D}\mathbf{e}_l\|_2^2}{\|\mathbf{e}_l\|_2^2} = \frac{|d_{ll}|^2}{1} = \max_{1 \leq k \leq n} |d_{kk}|^2.$$

Das heisst, es ist

$$\|\mathbf{D}\|_2 = \max_{\mathbf{x} \neq \mathbf{o}} \frac{\|\mathbf{D}\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \max_{1 \leq k \leq n} |d_{kk}|. \quad (6.69)$$



Ähnlich wie die Vektornormen, insbesondere die bereits in Kapitel 2 diskutierte 2-Norm von \mathbb{E}^n , hat auch die Operatornorm Eigenschaften, die als Axiome eines allgemeineren Normbegriffes für Abbildungen und Matrizen dienen können:

Satz 6.18 Die durch (6.63) definierte induzierte Operatornorm hat die folgenden Eigenschaften:

(OpN1) Sie ist **positiv definit**:

$$\begin{aligned}\|F\| &\geq 0 & (\forall F \in \mathcal{L}(X, Y)), \\ \|F\| = 0 &\implies F = O.\end{aligned}$$

(OpN2) Sie ist **dem Betrage nach homogen**:

$$\|\alpha F\| = |\alpha| \|F\| \quad (\forall F \in \mathcal{L}(X, Y), \forall \alpha \in \mathbb{E}).$$

(OpN3) Die **Dreiecksungleichung** gilt:

$$\boxed{\|F + G\| \leq \|F\| + \|G\| \quad (\forall F, G \in \mathcal{L}(X, Y)).} \quad (6.70)$$

(OpN4) Für **zusammengesetzte Abbildungen** gilt:

$$\boxed{\|G \circ F\| \leq \|G\| \|F\| \quad (\forall F \in \mathcal{L}(X, Y), G \in \mathcal{L}(Y, Z)).} \quad (6.71)$$

(OpN5) Sie ist **kompatibel mit den Vektornormen** in X, Y :

$$\boxed{\|Fx\|_Y \leq \|F\| \|x\|_X \quad (\forall F \in \mathcal{L}(X, Y), \forall x \in X).} \quad (6.72)$$

BEWEIS: (OpN1): Aus der Definition (6.63) ist klar, dass $\|F\| \geq 0$ gilt und dass $\|F\| = 0$ bedeutet, dass $Fx = o$ für alle x , also F die Nullabbildung O ist.

(OpN2): Unter Verwendung der Normeigenschaft (N2) in Y folgt, dass

$$\begin{aligned}\|\alpha F\| &= \sup_{\|x\|_X=1} \|\alpha Fx\|_Y = \sup_{\|x\|_X=1} (|\alpha| \|Fx\|_Y) \\ &= |\alpha| \sup_{\|x\|_X=1} \|Fx\|_Y = |\alpha| \|F\|.\end{aligned}$$

(OpN3): Hier verwendet man u.a. die Dreiecksungleichung (N3) aus Y :

$$\begin{aligned}\|F + G\| &= \sup_{\|x\|_X=1} \|(F + G)(x)\|_Y = \sup_{\|x\|_X=1} \|Fx + Gx\|_Y \\ &\leq \sup_{\|x\|_X=1} (\|Fx\|_Y + \|Gx\|_Y) \\ &\leq \sup_{\|x\|_X=1} \|Fx\|_Y + \sup_{\|x\|_X=1} \|Gx\|_Y \\ &= \|F\| + \|G\|.\end{aligned}$$

(OpN5): Folgt sofort daraus, dass für festes $x = \tilde{x}$ gilt

$$\frac{\|F\tilde{x}\|_Y}{\|\tilde{x}\|_X} \leq \sup_{\|x\|_X=1} \frac{\|Fx\|_Y}{\|x\|_X} = \|F\|.$$

(OpN4): Wir wenden (OpN5) zunächst auf G , dann auf F an, umzu sehen, dass

$$\|(G \circ F)(x)\|_Z = \|G(Fx)\|_Z \leq \|G\| \|Fx\|_Y \leq \|G\| \|F\| \|x\|_X.$$

Dividiert man dies unter der Annahme $x \neq 0$ durch $\|x\|_X$ und nimmt man das Supremum über alle $x \neq 0$, folgt die Behauptung. ■

Für quadratische Matrizen, wo $X = Y = Z := \mathbb{E}^n$, wird der Begriff “Matrixnorm” nun allgemein wie folgt definiert:

DEFINITION: Eine **Matrixnorm** [*matrix norm*] ist eine Funktion

$$\|\cdot\| : \mathbb{E}^{n \times n} \rightarrow \mathbb{R}, \quad \mathbf{A} \mapsto \|\mathbf{A}\| \quad (6.73)$$

mit den (an die Situation angepassten) Eigenschaften (OpN1)–(OpN4). Gilt zusätzlich analog zu (OpN5)

$$\boxed{\|\mathbf{Ax}\| \leq \|\mathbf{A}\| \|\mathbf{x}\| \quad \text{für alle } \mathbf{A} \in \mathbb{E}^{n \times n}, \mathbf{x} \in \mathbb{E}^n,} \quad (6.74)$$

so heisst die Matrixnorm **kompatibel** [*compatible*] mit der in \mathbb{E}^n verwendeten Vektornorm (die nicht die 2–Norm sein muss). ▲

Hier ist ein Beispiel einer Matrixnorm, die einfacher zu berechnen ist als die Spektralnorm:

BEISPIEL 6.13: Die **Frobenius–Norm** ist definiert durch

$$\boxed{\|\mathbf{A}\|_F := \sqrt{\sum_{k=1}^n \sum_{l=1}^n |a_{kl}|^2}.} \quad (6.75)$$

Man müsste natürlich beweisen, dass die Axiome (OpN1)–(OpN4) erfüllt sind. Wir wollen hier nur zeigen, dass diese Matrixnorm kompatibel ist mit der 2–Norm (als Vektornorm im \mathbb{E}^n). Wir wenden dazu die Schwarzsche Ungleichung(2.52) für $k = 1, \dots, n$ an auf den Zeilenvektor $(a_{k1} \ \cdots \ a_{kn})$ und einen n –Vektor \mathbf{x} :

$$\begin{aligned} \|\mathbf{Ax}\|_2^2 &= \sum_{k=1}^n \left(\sum_{l=1}^n a_{kl} x_l \right)^2 \\ &\leq \sum_{k=1}^n \left(\sum_{l=1}^n a_{kl}^2 \cdot \sum_{j=1}^n x_j^2 \right) \\ &= \sum_{j=1}^n x_j^2 \cdot \sum_{k=1}^n \left(\sum_{l=1}^n a_{kl}^2 \right) \\ &= \|\mathbf{x}\|_2^2 \|\mathbf{A}\|_F^2. \end{aligned}$$

◆

Die Genauigkeit, mit der man ein reguläres Gleichungssystem $\mathbf{Ax} = \mathbf{b}$ auf einem Computer mit Gleitkomma–Arithmetik lösen kann, hängt ab vom Produkt $\|\mathbf{A}\|_\infty \|\mathbf{A}^{-1}\|_\infty$, worin $\|\mathbf{A}\|_\infty$ die von der Vektornorm $\|\cdot\|_\infty$ aus (2.60) induzierte Operatornorm bezeichnet. In anderen Situationen spielt das Produkt $\|\mathbf{A}\|_2 \|\mathbf{A}^{-1}\|_2$ eine wichtige Rolle für die erreichbare Genauigkeit einer Rechnung. Man definiert deshalb:

DEFINITION: Die **Konditionszahl** [*condition number*] einer regulären Matrix \mathbf{A} bezüglich einer gewissen Norm¹³ $\|\cdot\|$ oder, kurz, die **Kondition** [*condition*] einer regulären Matrix \mathbf{A} ist die Zahl

$$\kappa(\mathbf{A}) = \|\mathbf{A}\| \|\mathbf{A}^{-1}\|. \quad (6.76)$$

Insbesondere ist die **2-Norm-Konditionszahl** [*2-norm condition number*] definiert als

$$\kappa_2(\mathbf{A}) = \|\mathbf{A}\|_2 \|\mathbf{A}^{-1}\|_2. \quad (6.77)$$



Auf die Berechnung der 2-Norm-Konditionszahl werden wir ebenfalls in den Abschnitten 10.6 und 11.2 zurückkommen.

¹³Gemeint ist eine Matrixnorm oder eine Vektornorm mit der induzierten Operatornorm.

Kapitel 7

Die Methode der kleinsten Quadrate und die QR–Zerlegung einer Matrix

In diesem Abschnitt beschränken wir uns auf den Vektorraum \mathbb{E}^m mit dem Standardskalarprodukt aus Abschnitt 2.4. Alles gilt aber sinngemäss für beliebige m -dimensionale Vektorräume mit Skalarprodukt, wie sie in Kapitel 6 eingeführt worden sind.

In Abschnitt 2.5 haben wir schon die orthogonale Projektion auf eine Gerade durch O , d.h. auf einen eindimensionalen Unterraum betrachtet, nun wollen wir jene auf mehrdimensionale Unterräume bestimmen und verschiedene Methoden zu deren Berechnung und zur Lösung des entsprechenden Minimumproblems kennen lernen. Dieses ist unter anderem als Problem der kleinsten Quadrate bekannt. Dabei gehen wir nun von einer allgemeinen Definition der Projektion aus.

7.1 Orthogonalprojektionen

DEFINITION: Eine lineare Abbildung $\mathbf{P} : \mathbb{E}^m \rightarrow \mathbb{E}^m$ heisst **Projektion** [*projection*] oder **Projektor** [*projector*], falls

$$\boxed{\mathbf{P}^2 = \mathbf{P}.} \quad (7.1)$$

Eine Projektion heisst **Orthogonalprojektion** [*orthogonal projection*] oder **orthogonaler Projektor** [*orthogonal projector*], falls zusätzlich gilt

$$\boxed{\ker \mathbf{P} \perp \operatorname{im} \mathbf{P},} \quad \text{d.h.} \quad \boxed{\mathcal{N}(\mathbf{P}) \perp \mathcal{R}(\mathbf{P}).} \quad (7.2)$$

Andernfalls ist es eine **schiefe** [*oblique*] Projektion. ▲

BEMERKUNG: Die Bedingung (7.1) entspricht der anschaulichen Vorstellung einer Projektion: Übt man auf das Bild $\mathbf{P}\mathbf{y} \in \operatorname{im} \mathbf{P}$ eines beliebigen Punktes \mathbf{y} nochmals dieselbe Projektion aus, so bleibt $\mathbf{P}\mathbf{y}$ “an Ort”: $\mathbf{P}(\mathbf{P}\mathbf{y}) = \mathbf{P}\mathbf{y}$. Die Bedingung (7.2) wird anschaulich, wenn man beachtet, dass $\mathbf{y} - \mathbf{P}\mathbf{y} \in \ker \mathbf{P}$, weil wegen (7.1) gilt $\mathbf{P}(\mathbf{y} - \mathbf{P}\mathbf{y}) = \mathbf{P}\mathbf{y} - \mathbf{P}^2\mathbf{y} = \mathbf{o}$. ▼

Lemma 7.1 *Ist \mathbf{P} ein Projektor, so ist auch $\mathbf{I} - \mathbf{P}$ ein Projektor und es gilt:*

$$\operatorname{im} (\mathbf{I} - \mathbf{P}) = \ker \mathbf{P}, \quad \ker (\mathbf{I} - \mathbf{P}) = \operatorname{im} \mathbf{P}. \quad (7.3)$$

BEWEIS: Wegen (7.1) ist $(\mathbf{I} - \mathbf{P})^2 = \mathbf{I} - 2\mathbf{P} + \mathbf{P}^2 = \mathbf{I} - \mathbf{P}$, d.h. es gilt (7.1) auch für $\mathbf{I} - \mathbf{P}$.

Ist $\mathbf{P}\mathbf{x} = \mathbf{o}$, so ist $\mathbf{x} = \mathbf{x} - \mathbf{P}\mathbf{x} = (\mathbf{I} - \mathbf{P})\mathbf{x} \in \text{im}(\mathbf{I} - \mathbf{P})$. Ist umgekehrt $\mathbf{x} = (\mathbf{I} - \mathbf{P})\mathbf{y} \in \text{im}(\mathbf{I} - \mathbf{P})$, so folgt $\mathbf{P}\mathbf{x} = (\mathbf{P} - \mathbf{P}^2)\mathbf{y} = \mathbf{o}$.

Die zweite Gleichung in (7.3) ergibt sich am einfachsten, wenn man in der ersten den Projektor \mathbf{P} durch den Projektor $\mathbf{I} - \mathbf{P}$ ersetzt. ■

Die Definition (7.2) der Orthogonalität eines Projektors kann man durch äquivalente Bedingungen ersetzen:

Satz 7.2 Für einen Projektor $\mathbf{P} : \mathbb{E}^m \rightarrow \mathbb{E}^m$ sind folgende Aussagen äquivalent:

- (i) \mathbf{P} ist orthogonaler Projektor;
- (ii) $\mathbf{I} - \mathbf{P}$ ist orthogonaler Projektor;
- (iii) $\mathbf{P}^H = \mathbf{P}$.

BEWEIS: Die Äquivalenz von (i) und (ii) folgt sofort aus Lemma 7.1. Nehmen wir also an, \mathbf{P} und $\mathbf{I} - \mathbf{P}$ seien orthogonale Projektoren. Zudem sei $\mathbf{x} \in \ker \mathbf{P}$, so dass wegen (7.2) für alle $\mathbf{y} \in \mathbb{E}^m$ gilt

$$\langle \mathbf{P}^H \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{P}\mathbf{y} \rangle = 0,$$

also $\mathbf{P}^H \mathbf{x} = \mathbf{o}$, d.h. $\mathbf{x} \in \ker \mathbf{P}^H$. Somit ist $\ker \mathbf{P} \subseteq \ker \mathbf{P}^H$, aber weil $\text{Rang } \mathbf{P} = \text{Rang } \mathbf{P}^H$ (Korollar 5.14), muss sogar $\ker \mathbf{P} = \ker \mathbf{P}^H$ sein. Analog ist $\ker(\mathbf{I} - \mathbf{P}) = \ker(\mathbf{I} - \mathbf{P}^H)$, also wegen Lemma 7.1 auch $\text{im } \mathbf{P} = \text{im } \mathbf{P}^H$, denn \mathbf{P}^H ist wegen $(\mathbf{P}^H)^2 = (\mathbf{P}^2)^H = \mathbf{P}^H$ sicher auch ein Projektor. Da ein Projektor durch seinen Kern (der auf \mathbf{o} abgebildet wird) und sein Bild (dessen Punkte fest bleiben), eindeutig bestimmt ist (denn Kern und Bild sind bei einer Projektion komplementäre Unterräume: $\ker \mathbf{P} \oplus \text{im } \mathbf{P} = \mathbb{E}^m$), folgt, dass $\mathbf{P} = \mathbf{P}^H$.

Es sei umgekehrt \mathbf{P} ein Projektor mit $\mathbf{P} = \mathbf{P}^H$, und es seien $\mathbf{x} \in \ker \mathbf{P}$ und $\mathbf{y} \in \text{im } \mathbf{P}$, so dass gilt $\mathbf{P}\mathbf{x} = \mathbf{o}$ und $\mathbf{y} = \mathbf{P}\mathbf{y}$. Dann folgt

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{P}\mathbf{y} \rangle = \langle \mathbf{P}^H \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{P}\mathbf{x}, \mathbf{y} \rangle = 0,$$

was gleichbedeutend ist mit (7.2), also ist \mathbf{P} ein orthogonaler Projektor. ■

Um eine Formel für die Orthogonalprojektion auf einen vorgegebenen Unterraum (“Projektionsebene”) anzugeben, wollen wir annehmen, dieser Unterraum sei aufgespannt durch n linear unabhängige, gegebene Vektoren $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{E}^m$, und wir betrachten diese Vektoren als Spaltenvektoren einer $m \times n$ -Matrix \mathbf{A} . Der vorgegebene Unterraum ist dann gerade der Spaltenraum $\mathcal{R}(\mathbf{A})$. Es gilt:

Lemma 7.3 Sind die Spalten einer $m \times n$ -Matrix \mathbf{A} linear unabhängig, das heißt ist $\text{Rang } \mathbf{A} = n$ ($\leq m$), so ist $\mathbf{A}^H \mathbf{A}$ regulär.

BEWEIS: Aus $\mathbf{A}^H \mathbf{A}\mathbf{x} = \mathbf{o}$ folgt $\mathbf{x}^H \mathbf{A}^H \mathbf{A}\mathbf{x} = 0$, d.h. $\langle \mathbf{A}\mathbf{x}, \mathbf{A}\mathbf{x} \rangle = 0$, also $\mathbf{A}\mathbf{x} = \mathbf{o}$. Weil $\text{Rang } \mathbf{A} = n$ ist, ist aber $\ker \mathbf{A} = \{\mathbf{o}\}$, d.h. $\mathbf{A}^H \mathbf{A}$ ist regulär. ■

Satz 7.4 Die Orthogonalprojektion $\mathbf{P}_A : \mathbb{E}^m \rightarrow \text{im } A \subseteq \mathbb{E}^m$ auf den Kolonnenraum $\mathcal{R}(A) \equiv \text{im } A$ einer $m \times n$ -Matrix A mit Rang n ($\leq m$) ist gegeben durch

$$\mathbf{P}_A := A(A^H A)^{-1} A^H. \quad (7.4)$$

BEWEIS: Zunächst ist

$$\mathbf{P}_A^2 = A(A^H A)^{-1} A^H A (A^H A)^{-1} A^H = \mathbf{P}_A,$$

\mathbf{P}_A ist also ein Projektor. Weiter ist $\text{im } \mathbf{P}_A = \text{im } A$, denn es ist klar, dass $\text{im } \mathbf{P}_A \subseteq \text{im } A$, und andererseits existiert zu $\mathbf{z} = A\mathbf{x} \in \text{im } A$ ein \mathbf{y} mit $\mathbf{P}_A \mathbf{y} = \mathbf{z}$, denn die entsprechende Gleichung für \mathbf{y} ,

$$A(A^H A)^{-1} A^H \mathbf{y} = \mathbf{z} = A\mathbf{x}$$

lässt sich lösen, indem man \mathbf{y} so wählt, dass

$$(A^H A)^{-1} A^H \mathbf{y} = \mathbf{x}, \quad \text{d.h. } A^H \mathbf{y} = A^H A \mathbf{x} = A^H \mathbf{z}.$$

Zum Beispiel ist $\mathbf{y} = \mathbf{z}$ eine Lösung; sie ist aber nicht eindeutig, falls $n < m$. Aus (7.4) sieht man weiter sofort, dass $\mathbf{P}_A^H = \mathbf{P}_A$, also ist \mathbf{P}_A eine Orthogonalprojektion, vgl. Satz 7.2 (iii). Es ist im übrigen auch einfach zu zeigen, dass aus $\mathbf{y} \perp \text{im } A$, d.h. $\mathbf{y}^H A \mathbf{x} = 0$ ($\forall \mathbf{x}$), folgt, dass $\mathbf{y} \in \ker \mathbf{P}_A$. Aus Dimensionsgründen gilt dann $\ker \mathbf{P}_A \perp \text{im } \mathbf{P}_A$, die ursprüngliche Bedingung (7.2) für einen orthogonalen Projektor. ■

Geometrisch ist klar, dass ein Projektor nicht davon abhängt, durch welche Vektoren man die “Projektionsebene” $\mathcal{R}(A) \equiv \text{im } A$ aufspannt. Man kann das auch formal beweisen. Wir dürfen also insbesondere $\mathbf{a}_1, \dots, \mathbf{a}_n$ ersetzen durch n orthonormale Vektoren, d.h. wir ersetzen $A \in \mathbb{E}^{m \times n}$ durch $Q \in \mathbb{E}^{m \times n}$ mit $\mathcal{R}(Q) = \mathcal{R}(A)$ und $Q^H Q = I_n$. Damit vereinfacht sich die Formel (7.4) für den Projektor:

Korollar 7.5 Die Orthogonalprojektion $\mathbf{P}_Q : \mathbb{E}^m \rightarrow \text{im } Q \subseteq \mathbb{E}^m$ auf den Kolonnenraum $\mathcal{R}(Q) \equiv \text{im } Q$ einer $m \times n$ -Matrix $Q = (\mathbf{q}_1 \ \cdots \ \mathbf{q}_n)$ mit orthonormalen Kolonnen ist gegeben durch

$$\mathbf{P}_Q := Q Q^H. \quad (7.5)$$

Es gilt also für jedes $\mathbf{y} \in \mathbb{E}^m$

$$\mathbf{P}_Q \mathbf{y} = Q Q^H \mathbf{y} = \sum_{j=1}^n \mathbf{q}_j \mathbf{q}_j^H \mathbf{y} = \sum_{j=1}^n \mathbf{q}_j \langle \mathbf{q}_j, \mathbf{y} \rangle. \quad (7.6)$$

Soll eine Orthogonalprojektion auf einen Kolonnenraum $\mathcal{R}(A)$ berechnet werden, so kann man also auf zwei Arten vorgehen: Entweder man wendet die allgemeine Formel (7.4) an oder man ersetzt A durch eine Matrix Q mit orthonormalen Kolonnen und wendet dann die einfacheren Formeln (7.5) und (7.6) an. Auf den Übergang von A auf Q werden wir in Abschnitt 7.3 zurückkommen.

BEISPIEL 7.1: Wir wollen im \mathbb{R}^4 orthogonal auf den 2-dimensionalen Unterraum projizieren, der aufgespannt wird durch die Kolonnen der Matrix

$$\mathbf{A} = (\mathbf{a}_1 \quad \mathbf{a}_2) = \begin{pmatrix} 4 & 6 \\ 2 & -2 \\ 2 & 6 \\ 1 & -7 \end{pmatrix}. \quad (7.7)$$

Berechnen wir

$$\mathbf{A}^\top \mathbf{A} = \begin{pmatrix} 25 & 25 \\ 25 & 125 \end{pmatrix}, \quad (\mathbf{A}^\top \mathbf{A})^{-1} = \frac{1}{100} \begin{pmatrix} 5 & -1 \\ -1 & 1 \end{pmatrix}, \quad (7.8)$$

so wird nach (7.4)

$$\mathbf{P}_\mathbf{A} = \frac{1}{100} \begin{pmatrix} 4 & 6 \\ 2 & -2 \\ 2 & 6 \\ 1 & -7 \end{pmatrix} \begin{pmatrix} 5 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 4 & 2 & 2 & 1 \\ 6 & -2 & 6 & -7 \end{pmatrix} \quad (7.9)$$

$$= \frac{1}{100} \begin{pmatrix} 4 & 6 \\ 2 & -2 \\ 2 & 6 \\ 1 & -7 \end{pmatrix} \begin{pmatrix} 14 & 12 & 4 & 12 \\ 2 & -4 & 4 & -8 \end{pmatrix} \quad (7.10)$$

$$= \begin{pmatrix} 0.68 & 0.24 & 0.40 & 0.00 \\ 0.24 & 0.32 & 0.00 & 0.40 \\ 0.40 & 0.00 & 0.32 & -0.24 \\ 0.00 & 0.40 & -0.24 & 0.68 \end{pmatrix}. \quad (7.11)$$

Anwendung des Gram-Schmidt-Verfahrens auf die zwei Kolonnen \mathbf{a}_1 und \mathbf{a}_2 gibt andererseits

$$\begin{aligned} \|\mathbf{a}_1\| &= 5, & \mathbf{q}_1 &= \frac{1}{5} (4 \ 2 \ 2 \ 1)^\top, \\ \langle \mathbf{q}_1, \mathbf{a}_2 \rangle &= 5, & \tilde{\mathbf{q}}_2 &= (2 \ -4 \ 4 \ -8)^\top, \\ \|\tilde{\mathbf{q}}_2\| &= 10, & \mathbf{q}_2 &= \frac{1}{5} (1 \ -2 \ 2 \ -4)^\top, \end{aligned} \quad (7.12)$$

also die folgende Matrix \mathbf{Q} mit orthonormalen Kolonnen:

$$\mathbf{Q} = (\mathbf{q}_1 \quad \mathbf{q}_2) = \frac{1}{5} \begin{pmatrix} 4 & 1 \\ 2 & -2 \\ 2 & 2 \\ 1 & -4 \end{pmatrix}. \quad (7.13)$$

Damit können wir $\mathbf{P}_\mathbf{A} = \mathbf{P}_\mathbf{Q}$ nach (7.5) auch darstellen durch

$$\mathbf{P}_\mathbf{A} = \mathbf{P}_\mathbf{Q} = \frac{1}{25} \begin{pmatrix} 4 & 1 \\ 2 & -2 \\ 2 & 2 \\ 1 & -4 \end{pmatrix} \begin{pmatrix} 4 & 2 & 2 & 1 \\ 1 & -2 & 2 & -4 \end{pmatrix}. \quad (7.14)$$

Man könnte dieses Produkt ausmultiplizieren, um wieder (7.11) zu erhalten, aber je nach dem Verhältnis $n : m$ und der Zahl der zu berechnenden Matrix-Vektor-Produkte $\mathbf{P}_\mathbf{A} \mathbf{x}$ ist es effizienter, den Projektor in einer Produktform (7.9), (7.10) oder (7.14) zu belassen, wobei man im allgemeinen (7.10) aus (7.9) am besten durch Lösen von n Gleichungssystemen berechnen würde. \blacklozenge

Eine Orthogonalprojektion hat die Eigenschaft, dass für Punkte \mathbf{y} , die nicht in der “Projektionsebene” liegen, für die also $\mathbf{y} \notin \text{im } \mathbf{P}$ gilt, das Bild $\mathbf{P}\mathbf{y}$ der am nächsten bei \mathbf{y} liegende Punkt des Unterraumes $\text{im } \mathbf{P}$ ist. Dies ergibt sich sofort aus dem Satz von Pythagoras (Satz 2.13). Die Orthogonalprojektion liefert damit die Lösung eines linearen Approximationsproblems in der 2-Norm:

Satz 7.6 Für eine Orthogonalprojektion \mathbf{P} gilt:

$$\|\mathbf{y} - \mathbf{P}\mathbf{y}\|_2 = \min_{\mathbf{z} \in \text{im } \mathbf{P}} \|\mathbf{y} - \mathbf{z}\|_2. \quad (7.15)$$

Analoges gilt in einem beliebigen Vektorraum mit Skalarprodukt für die durch das Skalarprodukt definierte Norm (6.7) aufgrund des allgemeinen Satzes von Pythagoras (Satz 6.2).

7.2 Die Methode der kleinsten Quadrate

Wir betrachten nun das **überbestimmte lineare Gleichungssystem** [*overdetermined linear system*]

$$\mathbf{A}\mathbf{x} = \mathbf{y} \quad (7.16)$$

mit einer “hohen” $m \times n$ -Matrix ($m > n$). Es hat im allgemeinen ja keine Lösung. Wir wissen, dass genau dann eine existiert, wenn $\mathbf{y} \in \mathcal{R}(\mathbf{A})$, aber das wird hier *nicht* vorausgesetzt. Weil man die Gleichungen (7.16) nur bis auf einen Fehler lösen kann, nennt man sie **Fehlergleichungen**.

Wenn es keine exakte Lösung gibt, ist es naheliegend, $\mathbf{x} \in \mathbb{E}^n$ so zu wählen, dass der **Residuenvektor** [*residual vector*] (kurz: das **Residuum** [*residual*])

$$\mathbf{r} \equiv \mathbf{y} - \mathbf{A}\mathbf{x} \quad (7.17)$$

minimale Euklidische Norm (2-Norm, “Länge”) hat, d.h., dass die Quadratsumme

$$\|\mathbf{r}\|^2 = \sum_{k=1}^m r_k^2 \quad \text{falls } \mathbf{r} \in \mathbb{R}^m \quad (7.18)$$

bzw.

$$\|\mathbf{r}\|^2 = \sum_{k=1}^m |r_k|^2 \quad \text{falls } \mathbf{r} \in \mathbb{C}^m \quad (7.19)$$

minimal wird. Die entsprechende Lösung \mathbf{x} heisst die **Lösung im Sinne der kleinsten Quadrate** [*least squares solution*] des überbestimmten Systems $\mathbf{A}\mathbf{x} = \mathbf{y}$.

Wir nehmen an, dass die Kolonnen von \mathbf{A} linear unabhängig sind, also $\ker \mathbf{A} = \{\mathbf{o}\}$ gilt und $\mathbf{A}^H \mathbf{A}$ nach Lemma 7.3 regulär ist. Auf Grund von Satz 7.6 wissen wir, dass $\|\mathbf{r}\|$ minimal ist, wenn gilt

$$\mathbf{A}\mathbf{x} = \mathbf{P}_{\mathbf{A}}\mathbf{y}, \quad \text{d.h.} \quad \mathbf{A}\mathbf{x} = \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{y}.$$

Da $\ker \mathbf{A} = \{\mathbf{o}\}$, folgt hieraus

$$\boxed{\mathbf{x} = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{y}} \quad (7.20)$$

oder

$$\boxed{\mathbf{A}^H \mathbf{A} \mathbf{x} = \mathbf{A}^H \mathbf{y}}. \quad (7.21)$$

Dies sind die **Normalgleichungen** [*normal equations*]. Man bekommt sie, indem man die Fehlergleichungen von links mit \mathbf{A}^H multipliziert. Sie bedeuten gerade, dass $\mathbf{r} = \mathbf{y} - \mathbf{A} \mathbf{x}$ auf den Kolonnen von \mathbf{A} senkrecht steht:

$$\boxed{\mathbf{A}^H \mathbf{r} = \mathbf{o}} \quad \text{oder} \quad \boxed{\mathbf{r} \perp \mathcal{R}(\mathbf{A})}. \quad (7.22)$$

Die Matrix $(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$ in (7.20) heisst **Pseudoinverse** [*pseudoinverse*] von \mathbf{A} . (Falls $\text{Rang } \mathbf{A} < m$, ist die Definition komplizierter.)

Zusammengefasst gilt:

Satz 7.7 *Es sei $\mathbf{A} \in \mathbb{E}^{m \times n}$, $\text{Rang } \mathbf{A} = n \leq m$, $\mathbf{y} \in \mathbb{E}^m$. Dann hat das überbestimmte Gleichungssystem $\mathbf{A} \mathbf{x} = \mathbf{y}$ eine eindeutig bestimmte Lösung \mathbf{x} im Sinne der kleinsten Quadrate, d.h. \mathbf{x} mit*

$$\|\mathbf{A} \mathbf{x} - \mathbf{y}\|^2 = \min_{\tilde{\mathbf{x}} \in \mathbb{E}^n} \|\mathbf{A} \tilde{\mathbf{x}} - \mathbf{y}\|^2. \quad (7.23)$$

\mathbf{x} kann berechnet werden durch Lösen des regulären Systems (7.21) der Normalgleichungen. Der Residuenvektor $\mathbf{r} := \mathbf{y} - \mathbf{A} \mathbf{x}$ steht senkrecht auf $\mathcal{R}(\mathbf{A})$.

BEISPIEL 7.2: Gegeben seien die 4×2 Matrix \mathbf{A} aus (7.7) und $\mathbf{y} = (3 \ 6 \ -3 \ 9)^\top$. Es soll $\mathbf{x} = (x_1 \ x_2)^\top$ so bestimmt werden, dass

$$\|\mathbf{r}\| = \|\mathbf{y} - \mathbf{A} \mathbf{x}\| = \left\| \begin{pmatrix} 3 \\ 6 \\ -3 \\ 9 \end{pmatrix} - \begin{pmatrix} 4 & 6 \\ 2 & -2 \\ 2 & 6 \\ 1 & -7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\| \quad (7.24)$$

minimal wird. Es ist $\mathbf{A}^\top \mathbf{y} = (27 \ -75)^\top$ und $\mathbf{A}^\top \mathbf{A} = \begin{pmatrix} 25 & 25 \\ 25 & 125 \end{pmatrix}$.

Die Normalgleichungen (7.21) lauten also

$$\begin{pmatrix} 25 & 25 \\ 25 & 125 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 27 \\ -75 \end{pmatrix}. \quad (7.25)$$

Es wird $\mathbf{x} = (2.10 \ -1.02)^\top$, also

$$\mathbf{r} = \mathbf{y} - \mathbf{A} \mathbf{x} = \begin{pmatrix} 3 \\ 6 \\ -3 \\ 9 \end{pmatrix} - \begin{pmatrix} 4 & 6 \\ 2 & -2 \\ 2 & 6 \\ 1 & -7 \end{pmatrix} \begin{pmatrix} 2.10 \\ -1.02 \end{pmatrix} = \begin{pmatrix} 0.72 \\ -0.24 \\ -1.08 \\ -0.24 \end{pmatrix} \quad (7.26)$$

und

$$\|\mathbf{r}\| = 1.3416\dots$$



In der Praxis kann das Lösen der Normalgleichungen für grosse Matrizen ungenau sein, da mit grossen Rundungsfehlern behaftet, falls die Kolonnen von \mathbf{A} nahezu linear abhängig sind. Es gibt jedoch mehrere Alternativen für die Bestimmung von \mathbf{x} . Eine haben wir bereits zur Hand: die Anwendung des Gram-Schmidt-Verfahrens. Geometrisch sollte klar sein, dass folgendes gilt:

Lemma 7.8 *Unter den Voraussetzungen von Satz 7.7 kann man die Kleinste-Quadrate-Lösung \mathbf{x} bestimmen, indem man die Kolonnen $\mathbf{a}_1, \dots, \mathbf{a}_n$ von \mathbf{A} und den Vektor $\mathbf{a}_{n+1} := \mathbf{y}$ dem Schmidtschen Orthogonalisierungsprozess (6.28) unterwirft und so das Lot*

$$\tilde{\mathbf{q}}_{n+1} := \mathbf{y} - \mathbf{A}\mathbf{x} = \mathbf{r} \perp \mathcal{R}(\mathbf{A}), \quad (7.27)$$

von \mathbf{y} auf $\mathcal{R}(\mathbf{A}) = \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ bestimmt. Das Gleichungssystem $\mathbf{A}\mathbf{x} = \mathbf{y} - \tilde{\mathbf{q}}_{n+1}$ ist dann exakt nach \mathbf{x} auflösbar.

Wir werden dieses Vorgehen im Abschnitt 7.3 weiterverfolgen und neu interpretieren. Statt $\mathbf{A}\mathbf{x} = \mathbf{y} - \tilde{\mathbf{q}}_{n+1}$ zu lösen, werden wir dort \mathbf{x} auf elegantere und numerisch genauere Art bestimmen.

BEISPIEL 7.3: Wir wollen die Beispiele 7.1 und 7.2 fortsetzen, aber im Moment nur das Residuum $\mathbf{r} = \mathbf{y} - \mathbf{A}\mathbf{x}$ der Kleinste-Quadrate-Lösung \mathbf{x} bestimmen. In Beispiel 7.1 haben wir bereits die Kolonnen von \mathbf{A} aus (7.7) orthonormalisiert und in (7.13) $\mathbf{Q} = (\mathbf{q}_1 \quad \mathbf{q}_2)$ erhalten. Nun ist das Lot von \mathbf{y} auf $\text{span}\{\mathbf{a}_1, \mathbf{a}_2\} = \text{span}\{\mathbf{q}_1, \mathbf{q}_2\}$ zu fällen: Es wird $\langle \mathbf{q}_1, \mathbf{y} \rangle = 27/5 = 5.4$, $\langle \mathbf{q}_2, \mathbf{y} \rangle = -51/5 = -10.2$, und damit

$$\begin{aligned} \mathbf{r} &= \tilde{\mathbf{q}}_3 = \mathbf{y} - \mathbf{q}_1 \langle \mathbf{q}_1, \mathbf{y} \rangle - \mathbf{q}_2 \langle \mathbf{q}_2, \mathbf{y} \rangle \\ &= \begin{pmatrix} 3 \\ 6 \\ -3 \\ 9 \end{pmatrix} - \frac{27}{5} \begin{pmatrix} 4 \\ 2 \\ 2 \\ 1 \end{pmatrix} + \frac{51}{25} \begin{pmatrix} 1 \\ -2 \\ 2 \\ 44 \end{pmatrix} = \begin{pmatrix} 0.72 \\ -0.24 \\ -1.08 \\ -0.24 \end{pmatrix} \end{aligned}$$

wie in (7.26).



Überbestimmte lineare Gleichungssysteme treten zum Beispiel auf, wenn man Parameter eines linearen Modells (Systems) durch eine *Messreihe* bestimmt. Gewöhnlich wird man viel mehr Messungen machen als es Parameter gibt und dann die Parameter als Lösung im Sinne der kleinsten Quadrate bestimmen (**Ausgleichsrechnung, lineare Regression**).

BEISPIEL 7.4: Für den vertikalen freien Fall eines Massenpunktes im Vakuum gilt

$$y(t) = x_1 t + x_2 t^2,$$

wobei

- t : Zeit,
- $y(t)$: zurückgelegter Weg,
- x_1 : Anfangsgeschwindigkeit,
- x_2 : halbe Erdbeschleunigung.

Es sollen x_1 und x_2 bestimmt werden durch Messung von $y(t)$ zu den Zeiten $t = 1, 2, 3, 4$ sec. Die Messreihe ergibt:

t	1	2	3	4	Sekunden
$y(t)$	13.0	35.5	68.0	110.5	Meter

Also ist

$$\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 2 & 4 \\ 3 & 9 \\ 4 & 16 \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} 13.0 \\ 35.5 \\ 68.0 \\ 110.5 \end{pmatrix}.$$

Für die Normalgleichungen brauchen wir:

$$\mathbf{A}^\top \mathbf{A} = \begin{pmatrix} 30 & 100 \\ 100 & 354 \end{pmatrix}, \quad \mathbf{A}^\top \mathbf{y} = \begin{pmatrix} 730.0 \\ 2535.0 \end{pmatrix}.$$

Zu lösen ist also das Normalgleichungssystem

$$\begin{array}{cc|c} x_1 & x_2 & 1 \\ \hline 30 & 100 & 730.0 \\ 100 & 354 & 2535.0 \end{array}.$$

das $x_1 = 7.9355\dots$, $x_2 = 4.9194\dots$ ergibt und die Approximation $2x_2 = 9.8387\dots$ für die Erdbeschleunigung liefert. \blacklozenge

7.3 Die QR-Zerlegung einer Matrix

Wir wollen in diesem Abschnitt das Gram-Schmidt-Verfahren im Falle $V = \mathbb{E}^m$ neu interpretieren und dann damit das Problem der kleinsten Quadrate lösen. Gegeben sind also zunächst n linear unabhängige Vektoren $\mathbf{a}_1, \dots, \mathbf{a}_n$, die wir als Kolonnen einer $m \times n$ -Matrix \mathbf{A} auffassen können. Statt b_1, \dots, b_n nennen wir die neu konstruierten orthonormalen Vektoren $\mathbf{q}_1, \dots, \mathbf{q}_n$ — wie schon in den Abschnitten 7.1–7.2. Wir werden sie später als Kolonnen einer $m \times n$ -Matrix \mathbf{Q} auffassen. Diese Matrix wird also orthonormierte Kolonnen haben, was heisst, dass

$$\mathbf{Q}^H \mathbf{Q} = \mathbf{I}_n \quad (\text{falls } \mathbb{E} = \mathbb{C}) \quad \text{bzw.} \quad \mathbf{Q}^\top \mathbf{Q} = \mathbf{I}_n \quad (\text{falls } \mathbb{E} = \mathbb{R}). \quad (7.28)$$

Zur Erinnerung: nur wenn $m = n$ ist, ist \mathbf{Q} unitär oder orthogonal, womit dann zusätzlich gilt: $\mathbf{Q}\mathbf{Q}^H = \mathbf{I}_n$ bzw. $\mathbf{Q}\mathbf{Q}^\top = \mathbf{I}_n$.

Wir schreiben zuerst das Gram-Schmidt-Verfahren für diese Situation nochmals auf:

Algorithmus 7.1 (klassisches Gram–Schmidt–Verfahren im \mathbb{E}^m) Es seien $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{E}^m$ gegebene linear unabhängige Vektoren. Man berechne

$$\left. \begin{aligned} \mathbf{q}_1 &:= \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|}, \\ \tilde{\mathbf{q}}_k &:= \mathbf{a}_k - \sum_{j=1}^{k-1} \mathbf{q}_j \langle \mathbf{q}_j, \mathbf{a}_k \rangle, \\ \mathbf{q}_k &:= \frac{\tilde{\mathbf{q}}_k}{\|\tilde{\mathbf{q}}_k\|}, \end{aligned} \right\} (k = 2, \dots, n). \quad (7.29)$$

Durch Vergleich mit (7.6) ist die Summe in der Formel für $\tilde{\mathbf{q}}_k$ nun klar als Projektion von \mathbf{a}_k auf $\text{span}\{\mathbf{q}_1, \dots, \mathbf{q}_{k-1}\}$ erkennbar; es ist also $\tilde{\mathbf{q}}_k$ das Lot von \mathbf{a}_k auf $\text{span}\{\mathbf{q}_1, \dots, \mathbf{q}_{k-1}\}$.

Auflösen nach den gegebenen Vektoren \mathbf{a}_k liefert

$$\mathbf{a}_1 = \mathbf{q}_1 \|\mathbf{a}_1\|, \quad \mathbf{a}_k = \mathbf{q}_k \|\tilde{\mathbf{q}}_k\| + \sum_{j=1}^{k-1} \mathbf{q}_j \langle \mathbf{q}_j, \mathbf{a}_k \rangle \quad (k = 2, \dots, n).$$

Nun definieren wir die Koeffizienten

$$\left. \begin{aligned} r_{11} &:= \|\mathbf{a}_1\|, \\ r_{jk} &:= \langle \mathbf{q}_j, \mathbf{a}_k \rangle, \quad j = 1, \dots, k-1, \\ r_{kk} &:= \|\tilde{\mathbf{q}}_k\|, \end{aligned} \right\} k = 2, \dots, n, \quad (7.30)$$

die wir ergänzen können durch $r_{jk} := 0$ ($j = k+1, \dots, n$). Damit erhalten wir für $k = 1, \dots, n$

$$\mathbf{a}_k = \mathbf{q}_k r_{kk} + \sum_{j=1}^{k-1} \mathbf{q}_j r_{jk} = \sum_{j=1}^k \mathbf{q}_j r_{jk} = \sum_{j=1}^n \mathbf{q}_j r_{jk}. \quad (7.31)$$

Definieren wir die bereits erwähnten $(m \times n)$ -Matrizen

$$\mathbf{A} := (\mathbf{a}_1 \ \dots \ \mathbf{a}_n), \quad \mathbf{Q} := (\mathbf{q}_1 \ \dots \ \mathbf{q}_n)$$

sowie die $(n \times n)$ -Rechtsdreiecksmatrix

$$\mathbf{R} := \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ 0 & r_{22} & \dots & r_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & r_{nn} \end{pmatrix} \quad (7.32)$$

wird aus (7.31) kurz

$$\boxed{\mathbf{A} = \mathbf{QR}.} \quad (7.33)$$

DEFINITION: Die Zerlegung (7.33) einer $m \times n$ Matrix \mathbf{A} mit Maximalrang $n \leq m$ in eine $m \times n$ Matrix \mathbf{Q} mit orthonormalen Kolonnen mal eine $n \times n$ Rechtsdreiecksmatrix mit positiven Diagonalelementen heisst **QR-Faktorisierung** [*QR factorization*] von \mathbf{A} . ▲

Wir können die n orthonormalen Vektoren $\mathbf{q}_1, \dots, \mathbf{q}_n$ zu einer Orthonormalbasis $\{\mathbf{q}_1, \dots, \mathbf{q}_n, \dots, \mathbf{q}_m\}$ von \mathbb{E}^m ergänzen und erhalten so eine unitäre (oder orthogonale) quadratische Matrix

$$\tilde{\mathbf{Q}} := (\mathbf{Q} \mid \mathbf{Q}_\perp) := (\mathbf{q}_1 \ \dots \ \mathbf{q}_n \mid \mathbf{q}_{n+1} \ \dots \ \mathbf{q}_m). \quad (7.34)$$

Um \mathbf{R} kompatibel zu machen mit $\tilde{\mathbf{Q}}$, können wir \mathbf{R} mit $m - n$ Zeilen aus Nullen ergänzen zu einer $m \times n$ -Matrix $\tilde{\mathbf{R}}$:

$$\tilde{\mathbf{R}} := \begin{pmatrix} \mathbf{R} \\ \mathbf{O} \end{pmatrix}. \quad (7.35)$$

Dann gilt

$$\mathbf{A} = \mathbf{QR} = (\mathbf{Q} \mid \mathbf{Q}_\perp) \begin{pmatrix} \mathbf{R} \\ \mathbf{O} \end{pmatrix} = \tilde{\mathbf{Q}}\tilde{\mathbf{R}}. \quad (7.36)$$

Hierbei bilden die Kolonnen $\mathbf{q}_{n+1}, \dots, \mathbf{q}_m$ von \mathbf{Q}_\perp natürlich gerade eine Orthonormalbasis des zu $\mathcal{R}(\mathbf{A})$ orthogonalen komplementären Unterraumes.

DEFINITION: Die Zerlegung $\mathbf{A} = \tilde{\mathbf{Q}}\tilde{\mathbf{R}}$ einer $m \times n$ Matrix \mathbf{A} mit Maximalrang $n \leq m$ in eine unitäre (oder orthogonale) $m \times m$ Matrix $\tilde{\mathbf{Q}}$ mal eine (rechteckige) $m \times n$ Rechtdreiecksmatrix mit positiven Diagonalelementen heisst **QR-Zerlegung**¹ [*QR decomposition*] von \mathbf{A} . ▲

Nun können wir Lemma 7.8 zum gesuchten Satz umformulieren:

Satz 7.9 *Das Gram-Schmidt-Verfahren, angewandt auf die Kolonnen $\mathbf{a}_1, \dots, \mathbf{a}_n$ einer $m \times n$ -Matrix \mathbf{A} liefert die QR-Faktorisierung (7.33) dieser Matrix. Ergänzt man \mathbf{A} durch den m -Vektor \mathbf{y} , so liefert das Verfahren (vor dem Normieren) zusätzlich den zu $\mathcal{R}(\mathbf{A})$ orthogonalen Residuenvektor \mathbf{r} gemäss der Formel*

$$\mathbf{r} = \mathbf{y} - \sum_{j=1}^n \mathbf{q}_j \langle \mathbf{q}_j, \mathbf{y} \rangle = \mathbf{y} - \mathbf{Q}\mathbf{Q}^H\mathbf{y}. \quad (7.37)$$

Die Lösung \mathbf{x} des Kleinste-Quadrate-Problems erfüllt das durch Rückwärtseinsetzen lösbare System

$$\mathbf{R}\mathbf{x} = \mathbf{Q}^H\mathbf{y}. \quad (7.38)$$

BEWEIS: Der erste Teil des Satzes ist eine Zusammenfassung unserer Herleitung von (7.33). Die Formel (7.37) für \mathbf{r} ergibt sich aus Lemma 7.8 und den Formeln (7.29) des Gram-Schmidt-Verfahrens. Es bleibt also nur (7.38) zu zeigen. Aus $\mathbf{A}\mathbf{x} = \mathbf{y} - \mathbf{r}$ folgt durch Einsetzen von $\mathbf{A} = \mathbf{QR}$ und Multiplikation mit \mathbf{Q}^H :

$$\underbrace{\mathbf{Q}^H\mathbf{Q}}_{=\mathbf{I}_n} \mathbf{R}\mathbf{x} = \mathbf{Q}^H\mathbf{y} - \underbrace{\mathbf{Q}^H\mathbf{r}}_{=\mathbf{o}}, \quad (7.39)$$

¹In bezug auf die sprachliche Unterscheidung zwischen “QR-Faktorisierung” und “QR-Zerlegung” folgen wir G.W. Stewart: *Matrix Algorithms, Vol. 1*, SIAM 1998. Sie ist noch nicht vielerorts akzeptiert.

wobei wir benutzt haben, dass \mathbf{r} auf den Kolonnen von \mathbf{Q} senkrecht steht, wie das aus der bekannten, in Satz 7.7 beschriebenen Eigenschaft der Minimallösung folgt. ■

Die Lösung eines Kleinste-Quadrate-Problems durch QR-Zerlegung der Matrix \mathbf{A} ist weniger mit Rundungsfehlern behaftet als die Lösung via die Normalgleichungen, und zwar vor allem, wenn die Kolonnen von \mathbf{A} beinahe linear abhängig sind. Allerdings ist dann auch das klassische Gram-Schmidt-Verfahren nicht sehr exakt. Aber es gibt daneben weitere, numerisch stabilere Methoden zur QR-Zerlegung einer Matrix.

BEISPIEL 7.5: Wir wollen zuerst das Beispiel 7.3 abschliessen. Den Formeln (7.12) aus Beispiel 7.1 entnehmen wir, dass \mathbf{A} aus (7.7) die QR-Faktorisierung

$$\underbrace{\begin{pmatrix} 4 & 6 \\ 2 & -2 \\ 2 & 6 \\ 1 & -7 \end{pmatrix}}_{\mathbf{A}} = \underbrace{\frac{1}{5} \begin{pmatrix} 4 & 1 \\ 2 & -2 \\ 2 & 2 \\ 1 & -4 \end{pmatrix}}_{\mathbf{Q}} \underbrace{\begin{pmatrix} 5 & 5 \\ 0 & 10 \end{pmatrix}}_{\mathbf{R}}$$

hat, und aus Beispiel 7.3 kennen wir schon $\mathbf{Q}^T \mathbf{y} = \begin{pmatrix} 27/5 & -51/5 \end{pmatrix}^T = \begin{pmatrix} 5.4 & -10.2 \end{pmatrix}^T$. Zu lösen bleibt also das System $\mathbf{R}\mathbf{x} = \mathbf{Q}^T \mathbf{y}$, konkret

$$\begin{array}{cc|c} x_1 & x_2 & 1 \\ \hline 5 & 5 & 5.4 \\ 0 & 10 & -10.2 \end{array}$$

das $x_1 = 2.10$, $x_2 = -1.02$ ergibt, was wir in (7.25) schon aus den Normalgleichungen erhalten haben. ◆

BEISPIEL 7.6: Im Beispiel 7.4 erhält man, gerundet auf vier Stellen nach dem Komma, aus dem Gram-Schmidt-Verfahren

$$\underbrace{\begin{pmatrix} 1 & 1 \\ 2 & 4 \\ 3 & 9 \\ 4 & 16 \end{pmatrix}}_{\mathbf{A}} = \underbrace{\begin{pmatrix} 0.1826 & -0.5133 \\ 0.3651 & -0.5866 \\ 0.5477 & -0.2200 \\ 0.7303 & 0.5866 \end{pmatrix}}_{\mathbf{Q}} \underbrace{\begin{pmatrix} 5.4772 & 18.2574 \\ 0 & 4.5461 \end{pmatrix}}_{\mathbf{R}},$$

$$\underbrace{\begin{pmatrix} 0.1826 & 0.3651 & 0.5477 & 0.7303 \\ -0.5133 & -0.5866 & -0.2200 & 0.5866 \end{pmatrix}}_{\mathbf{Q}^T} \underbrace{\begin{pmatrix} 13.0 \\ 35.5 \\ 68.0 \\ 110.5 \end{pmatrix}}_{\mathbf{y}} = \underbrace{\begin{pmatrix} 133.2792 \\ 22.3637 \end{pmatrix}}_{\mathbf{Q}^T \mathbf{y}},$$

also für das System $\mathbf{R}\mathbf{x} = \mathbf{Q}^T \mathbf{y}$

$$\begin{array}{cc|c} x_1 & x_2 & 1 \\ \hline 5.4772 & 18.2574 & 133.2792 \\ 0 & 4.5461 & 22.3637 \end{array}$$

was wie in Beispiel 7.4 wieder $x_1 = 7.9355$, $x_2 = 4.9194$ liefert. ◆

7.4 Die QR-Zerlegung mit Pivotieren

Wir haben bisher bei der in Abschnitt 7.3 mit dem klassischen Gram-Schmidt-Verfahren berechneten QR-Zerlegung angenommen, dass die $m \times n$ Ausgangsmatrix \mathbf{A} Maximalrang n hat, also insbesondere $n \leq m$ ist. In der Tat kann sonst das Verfahren (Algorithmus 7.1) zusammenbrechen, weil $\tilde{\mathbf{q}}_k = \mathbf{0}$ sein kann, womit dann \mathbf{q}_k nicht definiert ist. Dies kann schon für $k \leq \text{Rang } \mathbf{A}$ eintreten, wie die Beispiele

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 1 \\ 3 & 6 & 1 \end{pmatrix}$$

zeigen, wo der Rang zwei ist, aber $\tilde{\mathbf{q}}_2 = \mathbf{0}$ wird. Offensichtlich müsste man hier die zweite und die dritte Kolonne vertauschen um einen zweiten normierten Basisvektor $\mathbf{q}_2 \perp \mathbf{q}_1$ zu finden. Mit anderen Worten, man muss **Kolonnen-Pivotieren** zulassen. Aber wie findet man eine zum Vertauschen geeignete Kolonne?

Dazu modifizieren wir das Gram-Schmidt-Verfahren noch in anderer Hinsicht: durch Vertauschen von Schleife und Summe in (7.29).

Algorithmus 7.2 (modifiziertes Gram-Schmidt-Verfahren mit Kolonnen-Pivotieren) *Es sei $\mathbf{A} = (\mathbf{a}_1 \dots \mathbf{a}_n) \in \mathbb{E}^{m \times n}$. Man berechne:*

$$\left. \begin{aligned} \mathbf{q}_1 &:= \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|}, & \tilde{\mathbf{q}}_i &:= \mathbf{a}_i - \mathbf{q}_1 \langle \mathbf{q}_1, \mathbf{a}_i \rangle \quad (i = 2, \dots, n); \\ & \text{wähle } p \geq k \text{ mit } \|\tilde{\mathbf{q}}_p\| \neq 0 \text{ und} & & \\ & \text{vertausche Kolonnen } p \text{ und } k; \text{ berechne} & & \\ \mathbf{q}_k &:= \frac{\tilde{\mathbf{q}}_k}{\|\tilde{\mathbf{q}}_k\|}, & & \\ \tilde{\mathbf{q}}_i &:= \tilde{\mathbf{q}}_i - \mathbf{q}_k \langle \mathbf{q}_k, \tilde{\mathbf{q}}_i \rangle \quad (i = k + 1, \dots, n); & & \\ \text{ist } \|\tilde{\mathbf{q}}_{k+1}\| &= \dots = \|\tilde{\mathbf{q}}_n\| = 0, & & \\ \text{so gilt } \text{Rang } \mathbf{A} &= k \text{ und man ist fertig} & & \end{aligned} \right\} (k = 2, \dots, n).$$

(7.40)

Um zu beweisen, dass in exakter Arithmetik das modifizierte und das klassische Gram-Schmidt-Verfahren die gleichen Vektoren \mathbf{q}_k erzeugen, müsste man zeigen, dass in (7.40) gilt: $\langle \mathbf{q}_k, \tilde{\mathbf{q}}_i \rangle = \langle \mathbf{q}_k, \mathbf{a}_i \rangle$.

Sind in Algorithmus 7.2 keine Kolonnenvertauschungen nötig, gilt nach wie vor $\mathbf{A} = \mathbf{QR}$, wobei nun $r_{ki} \equiv \langle \mathbf{q}_k, \tilde{\mathbf{q}}_i \rangle$ für $k < i$. Ist $\text{Rang } \mathbf{A} < n$, so werden Vertauschungen ausgeführt, und man hat

$$\boxed{\mathbf{AP} = \mathbf{QR}} \tag{7.41}$$

mit einer $m \times m$ Permutationsmatrix \mathbf{P} , der $m \times r$ Matrix $\mathbf{Q} \equiv (\mathbf{q}_1 \dots \mathbf{q}_r)$ mit orthonormalen Kolonnen, die eine Basis von $\text{im } \mathbf{A} = \mathcal{R}(\mathbf{A})$ bilden, und einer rechteckigen, $r \times n$ Rechtsdreiecksmatrix \mathbf{R} . Dabei ist $r \equiv \text{Rang } \mathbf{A}$.

Kapitel 8

Determinanten

Die Determinante einer quadratischen Matrix ist ein zentraler Begriff und ein wichtiges Hilfsmittel in der klassischen Theorie der Matrizen. Wir werden sie vor allem für die Definition des sogenannten charakteristischen Polynoms brauchen, dessen Nullstellen die Eigenwerte der Matrix sind.

Die Determinante ist dagegen kaum je nützlich, wenn es um numerische Berechnungen geht. Glaubt man sie zu brauchen, hat man in der Regel den falschen Weg gewählt, um das Problem zu lösen! Aber es gibt Ausnahmen, zum Beispiel Anwendungen in der Physik.

Es gibt verschiedene Möglichkeiten, die Determinante einzuführen. Wir wählen hier die klassische Definition durch eine etwas komplizierte, in der Praxis unbrauchbare Formel, aus der aber alles folgt, insbesondere auch eine effiziente Berechnungsart.

8.1 Permutationen

Wir müssen uns zuerst an die Definition und die Eigenschaften der Permutationen der Zahlmenge $\{1, \dots, n\}$ erinnern.

DEFINITION: Eine **Permutation** [*permutation*] von n Elementen ist eine eineindeutige Abbildung der Menge $\{1, \dots, n\}$ auf sich. Die Menge aller dieser Permutationen sei S_n .

Eine **Transposition** [*transposition*] ist eine Permutation, bei der nur zwei Elemente vertauscht werden. ▲

Satz 8.1 *Es gibt $n!$ Permutationen in S_n .*

BEWEIS: Jede Permutation entspricht einer Anordnung der Zahlen $1, \dots, n$. Für $n = 1$ gilt: S_1 enthält nur die Identität, also nur 1 Element. Betrachtet wir nun in S_n jene Abbildungen, die die Reihenfolge der Zahlen $1, \dots, n - 1$ nicht verändern. Es gibt n von ihnen, denn man kann die Zahl n an n Stellen zwischen die restlichen einfügen. Wir dürfen die Induktionsvoraussetzung machen, dass die restlichen $n - 1$ Zahlen auf $(n - 1)!$ Arten angeordnet werden können. Also enthält S_n total $n(n - 1)! = n!$ Elemente. ■

Zwei Permutationen p_1 und p_2 kann man natürlich zusammensetzen zu $p_2 \circ p_1$. Diese Zusammensetzung soll als Produkt von Permutationen aufgefasst werden. S_n bildet mit dieser Operation eine Gruppe mit $n!$ Elementen, die nicht kommutativ ist, wenn $n \geq 3$. In der Algebra nennt man S_n die **symmetrische Gruppe** [*symmetric group*].

Satz 8.2 Für $n > 1$ kann jede Permutation p als Produkt von Transpositionen t_k benachbarter Elemente dargestellt werden:

$$p = t_\nu \circ t_{\nu-1} \circ \cdots \circ t_2 \circ t_1. \quad (8.1)$$

Die Darstellung ist im allgemeinen nicht eindeutig, aber die Anzahl der Transpositionen ist entweder immer gerade oder immer ungerade.

BEISPIEL 8.1: Die Permutation $p : (1, 2, 3, 4) \mapsto (4, 2, 1, 3)$ ist darstellbar als

$$(1, 2, 3, 4) \mapsto (1, 2, 4, 3) \mapsto (1, 4, 2, 3) \mapsto (4, 1, 2, 3) \mapsto (4, 2, 1, 3).$$

Das sind vier Transpositionen benachbarter Elemente, d.h. es ist $p = t_4 \circ t_3 \circ t_2 \circ t_1$. Die inverse Permutation p^{-1} ist gegeben durch $p^{-1} = t_1 \circ t_2 \circ t_3 \circ t_4$, denn es ist ja $t_k^{-1} = t_k$. \blacklozenge

BEWEIS von Satz 8.2: Die Aussagen sind trivial im Falle $n = 2$: es gibt nur zwei Permutationen, die Identität e und die Vertauschung $t : \{1, 2\} \mapsto \{2, 1\}$, die eine Transposition benachbarter Elemente ist. Jede Zusammensetzung einer geraden Anzahl von Exemplaren der Transposition t ergibt e , jedes Produkt einer ungeraden Anzahl wieder t .

Für den Induktionsbeweis nehmen wir an, die Aussagen seien richtig für Permutationen von $1, \dots, n-1$. Wir argumentieren ähnlich wie im Beweis von Satz 8.1: ausgehend von ihrer ursprünglichen Position am Ende lässt sich die Position der Zahl n im Bild einer Permutation

$$p : \{1, \dots, n\} \mapsto \{p(1), \dots, p(n)\} \quad (8.2)$$

offensichtlich durch ein Produkt von Transpositionen benachbarter Elemente erreichen, sagen wir durch $t_\mu \circ t_{\mu-1} \circ \cdots \circ t_1$. Die Permutation der übrigen $n-1$ Zahlen ist nach Induktionsvoraussetzung auch als ein solches Produkt darstellbar, sagen wir als $t_\nu \circ t_{\nu-1} \circ \cdots \circ t_{\mu+1}$. Damit hat p die Darstellung (8.1). Natürlich ist diese Darstellung nicht eindeutig wenn $n > 1$, denn für jede Vertauschung t gilt $t \circ t = e$.

Es seien nun

$$p = t_\nu \circ t_{\nu-1} \circ \cdots \circ t_2 \circ t_1 = t'_{\nu'} \circ t'_{\nu'-1} \circ \cdots \circ t'_2 \circ t'_1 \quad (8.3)$$

zwei verschiedene Darstellungen von p durch Transpositionen benachbarter Elemente. Sie brauchen nicht auf obige Art konstruiert zu sein. Wir betrachten für beliebiges $\mathbf{x} = (x_1 \ \dots \ x_n)^T \in \mathbb{R}^n$ das Produkt

$$\Psi_p = \prod_{i>j} (x_{p(i)} - x_{p(j)}) = \pm \prod_{i>j} (x_i - x_j) = \pm \Psi_e \quad (8.4)$$

Das Vorzeichen hängt nur ab von der Anzahl Transpositionen in der Darstellung (8.1), denn für eine beliebige Transposition t gilt $\Psi_t = -\Psi_e$. Aus (8.3) folgt damit, dass

$$\Psi_p = (-1)^\nu \Psi_e = (-1)^{\nu'} \Psi_e,$$

was impliziert, dass die Differenz $\nu - \nu'$ gerade ist. \blacksquare

DEFINITION: Das **Signum** [*sign*] einer Permutation p ist definiert als

$$\text{sign } p = \begin{cases} +1 & \text{falls } \nu \text{ in (8.1) gerade,} \\ -1 & \text{falls } \nu \text{ in (8.1) ungerade.} \end{cases} \quad (8.5)$$



8.2 Determinante: Definition, Eigenschaften

DEFINITION: Die **Determinante** [*determinant*] einer $n \times n$ -Matrix \mathbf{A} ist definiert als die Summe

$$\det \mathbf{A} = \sum_{p \in S_n} \text{sign } p \cdot a_{1,p(1)} a_{2,p(2)} \cdots a_{n,p(n)}, \quad (8.6)$$

die über die $n!$ Permutationen

$$p : (1, 2, \dots, n) \mapsto (p(1), p(2), \dots, p(n))$$

läuft. Statt $\det(\mathbf{A})$ schreibt man oft auch $|\mathbf{A}|$, wobei man dann die Klammern der Matrix weglässt, wenn sie durch ihre Elemente gegeben ist (und $n > 1$ ist).

BEISPIEL 8.2:

$$\det \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} = \begin{vmatrix} 2 & 1 \\ 3 & 4 \end{vmatrix}. \quad \blacklozenge$$

Man beachte, dass jeder der $n!$ Summanden in (8.6) aus jeder Zeile und jeder Kolonne von \mathbf{A} genau ein Element als Faktor enthält; das heisst, die Faktoren sind verteilt wie die Einsen der zugehörigen Permutationsmatrix.

Die Formel (8.6) ist für praktische Zwecke unbrauchbar, weil der Rechenaufwand mit n wie $n!$ zunimmt; es gibt ja $n!$ Terme in der Summe. Ein Computer, der 10^9 Multiplikationen pro Sekunde (1 GFLOP/s) ausführt, bräuchte für die Auswertung dieser Formel etwa 77 Jahre, wenn $n = 20$ ist, bzw. etwa 10^{140} Jahre, falls $n = 100$.

Brauchbar ist die Formel für $1 \leq n \leq 3$ und in Spezialfällen.

BEISPIEL 8.3: Für Matrizen der Ordnungen 1, 2 und 3 liefert die Formel (8.6) die folgenden Ausdrücke:

$$\det \begin{pmatrix} a_{11} \end{pmatrix} = a_{11}, \quad \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} a_{22} - a_{21} a_{12},$$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} a_{22} a_{33} + a_{21} a_{32} a_{13} + a_{31} a_{12} a_{23} - a_{31} a_{22} a_{13} - a_{21} a_{12} a_{33} - a_{11} a_{32} a_{23}. \quad (8.7)$$

Das letzte ist die **Regel von Sarrus**¹.

¹PIERRE SARRUS (1798 – 1861), französischer Mathematiker, Professor in Strasbourg.

BEISPIEL 8.4: Für eine Dreiecksmatrix reduziert sich die Summe (8.6) auf das Produkt der Diagonalelemente:

$$\begin{vmatrix} r_{11} & \star & \cdots & \star \\ & r_{22} & \cdots & \star \\ & & \ddots & \vdots \\ & & & r_{nn} \end{vmatrix} = r_{11} r_{22} \cdots r_{nn}, \quad (8.8)$$

$$\begin{vmatrix} l_{11} & & & \\ \star & l_{22} & & \\ \vdots & \vdots & \ddots & \\ \star & \star & \cdots & l_{nn} \end{vmatrix} = l_{11} l_{22} \cdots l_{nn}, \quad (8.9)$$

denn von den $n!$ Summanden kann nur einer nicht null sein, jener von $p = e$. In den anderen tritt immer auch ein Paar (k, l) auf mit $p(k) < k$ und $p(l) > l$, also mindesten je ein Faktor von oberhalb und einer von unterhalb der Diagonale, womit mindestens ein Faktor null ist, und damit auch das Produkt $r_{1,p(1)} r_{2,p(2)} \cdots r_{n,p(n)}$. Wir werden in Satz 8.4 noch einen zweiten Beweis dieser Tatsache geben. \blacklozenge

Einige grundlegende Eigenschaften der Determinante sind leicht zu verifizieren:

Satz 8.3 Die durch (8.6) definierte Determinante ist ein Funktional

$$\det : \mathbb{E}^{n \times n} \rightarrow \mathbb{E}, \quad \mathbf{A} \mapsto \det \mathbf{A},$$

mit den folgenden Eigenschaften:

- i) \det ist eine lineare Funktion jeder einzelnen Zeile der Matrix, d.h. für alle $\gamma, \gamma' \in \mathbb{E}$ und alle $l \in \{1, \dots, n\}$ ist

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ \gamma a_{l1} + \gamma' a'_{l1} & \gamma a_{l2} + \gamma' a'_{l2} & \cdots & \gamma a_{ln} + \gamma' a'_{ln} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \\ = \gamma \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{l1} & \cdots & a_{ln} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} + \gamma' \det \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a'_{l1} & \cdots & a'_{ln} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}. \end{aligned}$$

- ii) Werden in \mathbf{A} zwei Zeilen vertauscht, so wechselt $\det(\mathbf{A})$ das Vorzeichen.

- iii) $\det(\mathbf{I}) = 1$.

BEWEIS: Eigenschaft (i) folgt sofort aus

$$\begin{aligned} & \sum_{p \in S_n} (\text{sign } p) a_{1,p(1)} \cdots (\gamma a_{l,p(l)} + \gamma' a'_{l,p(l)}) \cdots a_{n,p(n)} \\ &= \gamma \sum_{p \in S_n} (\text{sign } p) a_{1,p(1)} \cdots a_{l,p(l)} \cdots a_{n,p(n)} \\ & \quad + \gamma' \sum_{p \in S_n} (\text{sign } p) a_{1,p(1)} \cdots a'_{l,p(l)} \cdots a_{n,p(n)}. \end{aligned}$$

Eigenschaft (ii) ergibt sich daraus, dass eine Zeilenvertauschung in \mathbf{A} in der Formel (8.6) dadurch kompensiert werden kann, dass man jede Permutation p mit der Transposition verknüpft, die dieser Zeilenvertauschung entspricht. Dadurch ändert sich aber gerade das Signum jedes Termes, während die Produkte unverändert bleiben.

Eigenschaft (iii) folgt als einfacher Spezialfall von Beispiel 8.4 über die Determinante von Dreiecksmatrizen. ■

Wir ziehen direkt aus den Eigenschaften (i)–(iii) weitere Schlüsse, die zum Teil auch leicht aus der Definition (8.6) gefolgert werden könnten. Es lassen sich in der Tat die meisten weiteren theoretischen Resultate in diesem Kapitel ableiten, wenn man vom Postulat eines eindeutigen Funktionals mit diesen drei Eigenschaften ausgeht.

Satz 8.4 *Das Funktional \det hat folgende weitere Eigenschaften:*

- iv) *Hat \mathbf{A} eine Zeile aus lauter Nullen, so ist $\det \mathbf{A} = 0$.*
- v) *$\det(\gamma \mathbf{A}) = \gamma^n \det \mathbf{A}$.*
- vi) *Hat \mathbf{A} zwei gleiche Zeilen, so ist $\det \mathbf{A} = 0$.*
- vii) *Addiert man zu einer Zeile von \mathbf{A} ein Vielfaches einer anderen Zeile von \mathbf{A} , so ändert sich der Wert von $\det \mathbf{A}$ nicht.*
- viii) *Ist \mathbf{A} eine Diagonalmatrix, so ist $\det \mathbf{A}$ gleich dem Produkt der Diagonalelemente.*
- ix) *Ist \mathbf{A} eine Dreiecksmatrix, so ist $\det \mathbf{A}$ gleich dem Produkt der Diagonalelemente.*

BEWEIS: Die Eigenschaften (iv), (v), (viii) und (ix) lassen sich sofort aus der Definition (8.6) ablesen. Aber (iv) folgt auch sofort aus Eigenschaft (i) in Satz 8.3, wenn wir l als Index dieser Zeile und $\gamma = \gamma' = 0$ wählen. Ebenso folgt (v) aus dieser Eigenschaft, wenn wir sie auf jede Zeile anwenden und $\gamma' = 0$ setzen.

Eigenschaft (vi) ergibt sich direkt aus (ii). Eigenschaft (vii) folgt aus (i) und (vi). Die Regel (viii) für Diagonalmatrizen ergibt sich auch (iii) nach n -maligem Anwenden von (i).

Ist \mathbf{R} eine obere Dreiecksmatrix mit Diagonalelementen $r_{kk} \neq 0$, kann man sie durch die in (vii) genannten elementaren Zeilenoperationen in eine Diagonalmatrix überführen, ohne den Wert der Determinante zu

ändern:

$$\left| \begin{array}{cccc} r_{11} & \star & \cdots & \star \\ & r_{22} & \cdots & \star \\ & & \ddots & \vdots \\ & & & r_{nn} \end{array} \right| = \left| \begin{array}{cccc} r_{11} & & & \\ & r_{22} & & \\ & & \ddots & \\ & & & r_{nn} \end{array} \right|.$$

(Zum Beispiel kann man hier durch Addition Vielfacher der letzten Zeile zu den anderen Zeilen deren letztes Element zu 0 machen.) Damit folgt Eigenschaft (ix) aus (viii). Sind nicht alle Diagonalelemente ungleich 0, so gibt es ein unterstes das 0 ist, sagen wir r_{kk} . Dann kann man durch Addition Vielfacher der Zeilen $k + 1$ bis n erreichen, dass die k -te Zeile nur Nullen enthält. Also ist $\det \mathbf{R} = 0$ nach (vii) und (iv). Analoges gilt für eine untere Dreiecksmatrix. ■

Weil die in Eigenschaft (vii) genannten Zeilenoperationen die Determinante nicht verändern, bleibt die Determinante im Verlaufe des Gauss-Algorithmus invariant, ausser dass jede Zeilenvertauschung zu einem Faktor -1 führt. Das stimmt auch dann, wenn eine Matrix \mathbf{R} in allgemeiner Zeilenstufenform resultiert, d.h. $\text{Rang } \mathbf{A} < n$ gilt. Nach Eigenschaft (ix) ist in diesem Falle aber $\det \mathbf{A} = 0$, während im regulären Falle das Produkt der Diagonalelemente von \mathbf{R} gerade das Produkt der Pivotelemente ist. Es gilt damit folgender Satz:

Satz 8.5 Für jede $n \times n$ -Matrix \mathbf{A} gilt:

$$\det \mathbf{A} \neq 0 \iff \text{Rang } \mathbf{A} = n \iff \mathbf{A} \text{ ist regulär}.$$

Wendet man auf \mathbf{A} den Gauss-Algorithmus an und ist dabei $\text{Rang } \mathbf{A} = n$, so ist $\det \mathbf{A}$ gleich dem Produkt der Pivotelemente multipliziert mit $(-1)^\nu$, wobei ν die Anzahl der ausgeführten Zeilenvertauschungen bezeichnet:

$$\det \mathbf{A} = (-1)^\nu \prod_{k=1}^n r_{kk}. \quad (8.10)$$

Der Faktor $(-1)^\nu$ in (8.10) ist gerade das Signum der Permutation, die den Zeilenvertauschungen entspricht.

Algorithmus 8.1 (Determinante via Gauss-Algorithmus)

Zur Berechnung der Determinante einer $n \times n$ -Matrix \mathbf{A} wende man den Gauss-Algorithmus auf \mathbf{A} an. Falls er den Rang n und die obere Dreiecksmatrix \mathbf{R} liefert, so gilt (8.10). Ist $\text{Rang } \mathbf{A} < n$, so ist $\det \mathbf{A} = 0$.

Der Gauss-Algorithmus 8.1 ist für $n \geq 4$ im allgemeinen die weitaus schnellste Methode zur Berechnung der Determinante. Im wesentlichen ist der Rechenaufwand ja bloss je $\frac{1}{3} n^3$ Additionen und Multiplikationen (vgl. Satz 3.2), was im Vergleich zu den rund $n!$ Additionen und $(n - 1)n!$ Multiplikationen von Definition (8.6) enorm wenig ist. Der Trick bestand in der Anwendung von Eigenschaft (vii) zur Reduktion auf obere Dreiecksform.

BEISPIEL 8.5: In unserem Beispiel 1.2 haben wir das System (1.9),

$$\begin{aligned}x_2 + 4x_3 &= 1 \\2x_1 + 4x_2 - 4x_3 &= 1 \\4x_1 + 8x_2 - 3x_3 &= 7\end{aligned}$$

mit dem Gauss-Algorithmus gelöst, wobei zunächst eine Zeilenvertauschung $1 \leftrightarrow 2$ nötig war und danach die Pivotelemente $r_{11} = 2$, $r_{22} = 1$ und $r_{33} = 5$ resultierten. Nach Formel (8.10) bzw. Algorithmus 8.1 ist also

$$\det \mathbf{A} = (-1)^1 \cdot 2 \cdot 1 \cdot 5 = -10,$$

was man leicht mit der Formel von Sarrus, (8.7), überprüfen kann. \blacklozenge

Es gilt sogar, dass die Eigenschaften (i)–(iii) aus Satz 8.3 charakteristisch sind für die Determinante. Es bleibt dazu zu zeigen, dass aus den Eigenschaften (i)–(iii) die Formel (8.6) folgt.

Satz 8.6 Die durch (8.6) definierte Determinante ist das einzige auf $\mathbb{E}^{n \times n}$ definierte Funktional mit den Eigenschaften (i)–(iii) aus Satz 8.3, das heisst diese Eigenschaften sind charakteristisch für die Determinante.

BEWEIS (Idee): Schreiben wir für $l = 1, \dots, n$ die l te Zeile von \mathbf{A} als

$$a_{l1} \mathbf{e}_1^\top + a_{l2} \mathbf{e}_2^\top + \dots + a_{ln} \mathbf{e}_n^\top$$

und wenden wir Eigenschaft (i) je für $l = 1, \dots, n$, also n -mal auf eine solche Summe von n Termen an, so ergibt sich, wie man leicht sieht,

$$\det \mathbf{A} = \sum_{k_1=1}^n \sum_{k_2=1}^n \dots \sum_{k_n=1}^n a_{1,k_1} a_{2,k_2} \dots a_{n,k_n} \begin{vmatrix} \mathbf{e}_{k_1}^\top \\ \mathbf{e}_{k_2}^\top \\ \vdots \\ \mathbf{e}_{k_n}^\top \end{vmatrix}.$$

In dieser Summe von n^n Termen sind aber aufgrund der Eigenschaft (vi) nur jene Determinanten nicht null, wo (k_1, \dots, k_n) eine Permutation von $(1, \dots, n)$ ist, d.h. es ist

$$\det \mathbf{A} = \sum_{p \in S_n} a_{1,p(1)} a_{2,p(2)} \dots a_{n,p(n)} \begin{vmatrix} \mathbf{e}_{p(1)}^\top \\ \mathbf{e}_{p(2)}^\top \\ \vdots \\ \mathbf{e}_{p(n)}^\top \end{vmatrix},$$

Anwendung der Eigenschaft (ii) liefert schliesslich

$$\det \mathbf{A} = \sum_{p \in S_n} (\text{sign } p) a_{1,p(1)} a_{2,p(2)} \dots a_{n,p(n)} \det \mathbf{I},$$

was wegen Eigenschaft (iii) mit der Formel (8.6) übereinstimmt. \blacksquare

Mit Hilfe von Satz 8.6 können wir nun eine weitere grundlegende Eigenschaft der Determinante beweisen.

Satz 8.7 Für irgend zwei $n \times n$ -Matrizen \mathbf{A} und \mathbf{B} gilt

$$\det(\mathbf{AB}) = \det \mathbf{A} \cdot \det \mathbf{B} . \quad (8.11)$$

BEWEIS (Idee): Ist $\det \mathbf{B} = 0$, so ist nach Satz 8.5 $\text{Rang } \mathbf{B} < n$ und damit nach Satz 5.16 auch $\text{Rang } (\mathbf{AB}) < n$, also $\det(\mathbf{AB}) = 0$. Wir dürfen also im folgenden $\det \mathbf{B} \neq 0$ voraussetzen.

Wir betrachten nun das durch $d(\mathbf{A}) := \det(\mathbf{AB})/\det \mathbf{B}$ definierte Funktional d und zeigen, dass es die Eigenschaften (i)–(iii) hat. Für (iii) ist dies sofort klar, während (i) und (ii) leicht aus der Beziehung (2.42) in Korollar 2.8 folgen. Aufgrund der gemäss Satz 8.6 gültigen Eindeutigkeit dieses Funktionals folgt dann, dass $d(\mathbf{A}) = \det \mathbf{A}$ sein muss. Also ist $\det \mathbf{A} = \det(\mathbf{AB})/\det \mathbf{B}$. ■

Korollar 8.8 Ist \mathbf{A} regulär, so gilt:

$$\det \mathbf{A}^{-1} = (\det \mathbf{A})^{-1} . \quad (8.12)$$

BEWEIS: Anwendung von Satz 8.7 auf $\mathbf{AA}^{-1} = \mathbf{I}$. ■

BEISPIEL 8.6: Wir wissen, dass der Gauss-Algorithmus eine LR-Zerlegung $\mathbf{PA} = \mathbf{LR}$ liefert. Da die Permutationsmatrix \mathbf{P} orthogonal ist, also $\mathbf{P}^{-1} = \mathbf{P}^T$ gilt, haben wir

$$\mathbf{A} = \mathbf{P}^T \mathbf{L} \mathbf{R}$$

und damit

$$\det \mathbf{A} = \det \mathbf{P}^T \cdot \det \mathbf{L} \cdot \det \mathbf{R} . \quad (8.13)$$

Hier ist nach Eigenschaften (ii) und (iii) aus Satz 8.3 gerade $\det \mathbf{P}^T = (-1)^\nu$, wo ν wieder die Zahl der Zeilenvertauschungen im Gauss-Algorithmus bezeichnet (diese führen \mathbf{P} gerade in \mathbf{I} über). Weiter ist nach Eigenschaft (ix) aus Satz 8.4

$$\det \mathbf{L} = 1 , \quad \det \mathbf{R} = \prod_{k=1}^n r_{kk} .$$

Zusammen ergibt sich wieder Formel (8.10). ◆

Nun fehlt uns einzig noch eine Formel für die Determinante der transponierten Matrix.

Satz 8.9 Es gilt:

$$\det \mathbf{A}^T = \det \mathbf{A} , \quad \det \mathbf{A}^H = \overline{\det \mathbf{A}} . \quad (8.14)$$

BEWEIS (1. Version): Nach (8.13) und Satz 8.7 sowie dank $\det \mathbf{P} = \det \mathbf{P}^{-1} = \det \mathbf{P}^T$ und Eigenschaft (ix) aus Satz 8.4 ist

$$\begin{aligned} \det \mathbf{A}^T &= \det \mathbf{R}^T \det \mathbf{L}^T \det \mathbf{P} = \det \mathbf{P} \prod_{k=1}^n r_{kk} \prod_{j=1}^n l_{jj} \\ &= \det \mathbf{P}^T \det \mathbf{L} \det \mathbf{R} = \det \mathbf{A} . \end{aligned}$$

Zudem ist $\det \mathbf{A}^H = \det \overline{\mathbf{A}}^T = \det \overline{\mathbf{A}} = \overline{\det \mathbf{A}}$. ■

BEWEIS (2. Version): Bei Anwendung der Definition (8.6) bekommt man durch Umm Nummerierung der Indizes

$$\begin{aligned} \det \mathbf{A}^T &= \sum_{p \in S_n} (\text{sign } p) a_{p(1),1} a_{p(2),2} \cdots a_{p(n),n} \\ &= \sum_{p^{-1} \in S_n} (\text{sign } p^{-1}) a_{1,p^{-1}(1)} a_{2,p^{-1}(2)} \cdots a_{n,p^{-1}(n)} \\ &= \sum_{\tilde{p} \in S_n} (\text{sign } \tilde{p}) a_{1,\tilde{p}(1)} a_{2,\tilde{p}(2)} \cdots a_{n,\tilde{p}(n)} \\ &= \det \mathbf{A}. \end{aligned}$$

■

Man beachte, dass $\det \mathbf{A}^T = \det \mathbf{A}$ auch für komplexe Matrizen gilt.

BEISPIEL 8.7: Für unitäres \mathbf{U} ist $1 = \det \mathbf{I} = \det \mathbf{U} \det \mathbf{U}^H = |\det \mathbf{U}|^2$, also

$$|\det \mathbf{U}| = 1.$$

Für orthogonale Matrizen \mathbf{Q} hat man speziell

$$\det \mathbf{Q} = \pm 1.$$

◆

Aus Satz 8.9 folgt nun unmittelbar:

Korollar 8.10 Die Eigenschaften (i) und (ii) aus Satz 8.3 sowie die Eigenschaften (iv), (vi) und (vii) aus Satz 8.4 gelten auch, wenn man “Zeile” durch “Kolonne” ersetzt.

8.3 Entwicklung nach Zeilen und Kolonnen

Schliesslich wollen wir eine weitere Möglichkeit der rekursiven Berechnung der Determinante herleiten, die allerdings nur in Spezialfällen effizient ist, aber in solchen Fällen interessante Formeln und Zusammenhänge liefern kann.

DEFINITION: Zu jedem Element a_{kl} einer $n \times n$ -Matrix \mathbf{A} werde die $(n-1) \times (n-1)$ -Untermatrix $\mathbf{A}_{[k,l]}$ definiert durch Streichen der Zeile k und der Kolonne l von \mathbf{A} . Der **Kofaktor** [*cofactor*] κ_{kl} von a_{kl} ist dann die Zahl

$$\kappa_{kl} := (-1)^{k+l} \det \mathbf{A}_{[k,l]}.$$

▲

BEISPIEL 8.8: Es sei

$$\mathbf{A} = \begin{pmatrix} 5 & 2 & 1 \\ 3 & 6 & 2 \\ 1 & 4 & 3 \end{pmatrix}.$$

Dann ist

$$\mathbf{A}_{[2,1]} = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix}, \quad \kappa_{21} = (-1)^3 \begin{vmatrix} 2 & 1 \\ 4 & 3 \end{vmatrix} = -(2 \cdot 3 - 4 \cdot 1) = -2.$$

◆

BEISPIEL 8.9: Für

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & 5 & 1 \\ 2 & 4 & 6 & 3 \\ 3 & 6 & 4 & 2 \\ 1 & 5 & 3 & 1 \end{pmatrix}$$

ist

$$\mathbf{A}_{[2,3]} = \begin{pmatrix} 1 & 3 & 1 \\ 3 & 6 & 2 \\ 1 & 5 & 1 \end{pmatrix}, \quad \kappa_{23} = (-1)^5 \begin{vmatrix} 1 & 3 & 1 \\ 3 & 6 & 2 \\ 1 & 5 & 1 \end{vmatrix} = -2.$$

◆

Lemma 8.11 *Es sei \mathbf{A} eine Matrix, in deren l -ter Kolonne nur das Element $a_{kl} \neq 0$ ist. Dann gilt*

$$\det \mathbf{A} = a_{kl} \kappa_{kl}.$$

BEWEIS: Wir bringen zunächst die l -te Kolonne von \mathbf{A} durch $l - 1$ Kolonnenvertauschungen so ganz nach links, dass die Reihenfolge der anderen Kolonnen erhalten bleibt. Dann bringen wir analog die k -te Zeile durch $k - 1$ Zeilenvertauschungen ganz nach oben. Die Determinante ändert sich dabei um einen Faktor $(-1)^{k+l-2} = (-1)^{k+l}$, und es resultiert die Matrix

$$\left(\begin{array}{c|ccc} a_{kl} & \star & \cdots & \star \\ \hline \mathbf{o} & & \mathbf{A}_{[k,l]} & \end{array} \right).$$

Bei Anwendung des Gauss-Algorithmus auf diese Matrix gibt es im ersten Schritt nichts zu tun: das erste Pivotelement ist a_{kl} und die erste Restgleichungsmatrix ist $\mathbf{A}_{[k,l]}$. Weiterverarbeitung der letzteren wird $\det \mathbf{A}_{[k,l]}$ als Produkt der restlichen Pivotelemente ergeben. Das Produkt aller Pivotelemente ist also $a_{kl} \det \mathbf{A}_{[k,l]}$. Mittels Definition (8.15) ergibt sich damit gerade die Behauptung. ■

Satz 8.12 *Ist \mathbf{A} eine $n \times n$ -Matrix, so gelten für jedes feste $k \in \{1, \dots, n\}$ und jedes feste $l \in \{1, \dots, n\}$ die Formeln*

$$\det \mathbf{A} = \sum_{i=1}^n a_{ki} \kappa_{ki} \quad (8.16)$$

und

$$\det \mathbf{A} = \sum_{i=1}^n a_{il} \kappa_{il}. \quad (8.17)$$

Die Formel (8.16) nennt man **Entwicklung nach der k -ten Zeile** [*expansion along row k*], und die Formel (8.17) heisst **Entwicklung nach der l -ten Kolonne** [*expansion along column l*].

BEWEIS von Satz 8.12: Wir können die l -te Kolonne von \mathbf{A} als eine Summe von n Kolonnenvektoren schreiben, von denen jeder nur eine

einzigste Komponente ungleich 0 hat. Wendet man die (gemäss Korollar 8.10 geltende) Kolonnenversion von Eigenschaft (i) an auf die Matrix \mathbf{A} mit der l -ten Kolonne geschrieben als diese Summe, so folgt unter Berücksichtigung von Lemma 8.11 gerade die Kolonnenentwicklung (8.17).

Die Entwicklung (8.16) erhält man aus (8.17) wiederum durch Anwendung von $\det \mathbf{A} = \det \mathbf{A}^T$ (Satz 8.9). ■

BEISPIEL 8.10: Für eine 3×3 -Matrix gilt bei Entwicklung nach der ersten Kolonne:

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

Dies stimmt mit der eingangs erwähnten Regel von Sarrus überein. ◆

BEISPIEL 8.11: Entwickeln wir die bereits betrachtete Matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & 5 & 1 \\ 2 & 4 & 6 & 3 \\ 3 & 6 & 4 & 2 \\ 1 & 5 & 3 & 1 \end{pmatrix}$$

nach der zweiten Zeile: Wir kennen bereits $\kappa_{23} = -2$; weiter ist

$$\begin{aligned} \kappa_{21} &= (-1)^3 \begin{vmatrix} 3 & 5 & 1 \\ 5 & 3 & 1 \end{vmatrix} = -12, \\ \kappa_{22} &= (-1)^4 \begin{vmatrix} 1 & 5 & 1 \\ 3 & 4 & 2 \\ 1 & 3 & 1 \end{vmatrix} = -2, \\ \kappa_{24} &= (-1)^6 \begin{vmatrix} 1 & 3 & 5 \\ 3 & 6 & 4 \\ 1 & 5 & 3 \end{vmatrix} = 28. \end{aligned}$$

Zusammen ergibt sich

$$\begin{aligned} \det \mathbf{A} &= a_{21} \kappa_{21} + a_{22} \kappa_{22} + a_{23} \kappa_{23} + a_{24} \kappa_{24} \\ &= 2 \cdot (-12) + 4 \cdot (-2) + 6 \cdot (-2) + 3 \cdot 28 \\ &= 40. \end{aligned}$$
◆

8.4 Determinanten von Blockdreiecksmatrizen

Man könnte hoffen, für eine 2×2 Blockmatrix gelte

$$\begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{vmatrix} = \det \mathbf{A} \det \mathbf{D} - \det \mathbf{B} \det \mathbf{C},$$

aber schon ein ganz simples Beispiel wie

$$\begin{vmatrix} 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \end{vmatrix} = 71 \neq 15 = 4 \cdot 4 - 1 \cdot 1$$

zeigt, dass dies im allgemeinen falsch ist. In der speziellen Situation, wo $\mathbf{B} = \mathbf{O}$ oder $\mathbf{C} = \mathbf{O}$ ist, gilt aber in der Tat folgendes:

Satz 8.13 *Für eine 2×2 Blockdreiecksmatrix ist*

$$\begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{D} \end{vmatrix} = \det \mathbf{A} \det \mathbf{D} \quad \text{bzw.} \quad \begin{vmatrix} \mathbf{A} & \mathbf{O} \\ \mathbf{C} & \mathbf{D} \end{vmatrix} = \det \mathbf{A} \det \mathbf{D}. \tag{8.18}$$

BEWEIS: Wir betrachten den Fall der oberen Blockdreiecksmatrix und nehmen an, \mathbf{A} sei ein $m \times m$ Block und \mathbf{D} ein $(n - m) \times (n - m)$ Block. In unserer Definition (8.6) leisten offenbar nur jene Permutationen p einen Beitrag, für die $p(1), \dots, p(m) \in \{1, \dots, m\}$ gilt; die anderen führen auf mindestens einen Faktor aus $\mathbf{B} = \mathbf{O}$. Die relevanten Permutationen p haben also die Form

$$(1, \dots, m; m+1, \dots, n) \mapsto (p_{\mathbf{A}}(1), \dots, p_{\mathbf{A}}(m); m+p_{\mathbf{D}}(1), \dots, m+p_{\mathbf{D}}(n-m))$$

mit $p_{\mathbf{A}} \in S_m$ und $p_{\mathbf{D}} \in S_{n-m}$. Dabei ist $\text{sign } p = \text{sign } p_{\mathbf{A}} \cdot \text{sign } p_{\mathbf{D}}$. Aus der Summe in (8.6) wird damit

$$\begin{aligned} \begin{vmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{O} & \mathbf{D} \end{vmatrix} &= \sum_{\substack{p_{\mathbf{A}} \in S_m \\ p_{\mathbf{D}} \in S_{n-m}}} \text{sign } p_{\mathbf{A}} \cdot \text{sign } p_{\mathbf{D}} \cdot a_{1,p_{\mathbf{A}}(1)} \cdots a_{m,p_{\mathbf{A}}(m)} \times \\ &\quad \times d_{1,p_{\mathbf{D}}(1)} \cdots d_{n-m,p_{\mathbf{D}}(n-m)} \\ &= \sum_{p_{\mathbf{A}} \in S_m} \text{sign } p_{\mathbf{A}} \cdot a_{1,p_{\mathbf{A}}(1)} \cdots a_{m,p_{\mathbf{A}}(m)} \times \\ &\quad \times \sum_{p_{\mathbf{D}} \in S_{n-m}} \text{sign } p_{\mathbf{D}} \cdot d_{1,p_{\mathbf{D}}(1)} \cdots d_{n-m,p_{\mathbf{D}}(n-m)} \\ &= \det \mathbf{A} \det \mathbf{D}. \end{aligned}$$

■

Durch rekursive Anwendung ergibt sich aus diesem Satz sofort:

Korollar 8.14 *Die Determinante einer Blockdreiecksmatrix ist gleich dem Produkt der Determinanten der diagonalen Blöcke.*

Kapitel 9

Eigenwerte und Eigenvektoren

Eigenwerte und Eigenvektoren spielen sowohl in der Theorie als auch in den Anwendungen eine grosse Rolle. Sie erlauben einfache Abbildungsmatrizen von Selbstabbildungen zu finden und zum Beispiel Normalformen für quadratische Funktionen zu finden. Man kann mit ihnen aber auch lineare Differentialgleichungen mit konstanten Koeffizienten lösen oder, konkret, die Eigenschwingungen eines elastischen Körpers bestimmen, im einfachsten Falle einer schwingenden Saite.

9.1 Eigenwerte und Eigenvektoren von Matrizen und linearen Abbildungen

Wir betrachten lineare Abbildungen $F : V \rightarrow V$, $x \mapsto Fx$ eines endlich-dimensionalen Vektorraumes V in sich und interessieren uns für Vektoren, die bei einer solchen Abbildung bloss gestreckt werden (allenfalls mit einem negativen Faktor).

DEFINITION: Die Zahl $\lambda \in \mathbb{E}$ heisst **Eigenwert** [*eigenvalue*] der linearen Abbildung $F : V \rightarrow V$, falls es einen **Eigenvektor** [*eigenvector*] $v \in V$, $v \neq 0$ gibt, so dass

$$\boxed{Fv = \lambda v}. \quad (9.1)$$

Ist λ ein Eigenwert, so ist der zugehörige **Eigenraum** [*Eigenspace*] E_λ gleich der um den Nullvektor erweiterten Menge der Eigenvektoren zu λ :

$$\boxed{E_\lambda := \{v \in V; Fv = \lambda v\}}. \quad (9.2)$$

Die Menge aller Eigenwerte von F heisst **Spektrum** [*spectrum*] von F und wird mit $\sigma(F)$ bezeichnet.

Als Abkürzungen für die Begriffe ‘Eigenvektor’ und ‘Eigenwert’ schreibt man oft **EV** [*EVec*] und **EW** [*EVal*]. ▲

Diese Definitionen gelten insbesondere für quadratische Matrizen: $\xi \in \mathbb{E}^n$ ist ein Eigenvektor zum Eigenwert λ von $\mathbf{A} \in \mathbb{E}^{n \times n}$ falls

$$\boxed{\mathbf{A}\xi = \xi\lambda}, \quad (9.3)$$

wobei wir hier wieder besser den Skalar rechts schreiben, vgl. (2.21).

Ist \mathbf{A} eine Abbildungsmatrix, die eine lineare Abbildung F bezüglich einer gewissen Basis repräsentiert, so haben F und \mathbf{A} die gleichen Eigenwerte, und die Eigenvektoren entsprechen sich auf natürliche Weise:

Lemma 9.1 *Es sei $F : V \rightarrow V$ eine lineare Abbildung, $\kappa_V : V \rightarrow \mathbb{E}^n$, $x \mapsto \xi$ eine Koordinatenabbildung von V (bezüglich einer gewählten Basis) und $\mathbf{A} = \kappa_V F \kappa_V^{-1}$ die entsprechende Abbildungsmatrix von F . Dann gilt:*

$$\left. \begin{array}{l} \lambda \text{ EW von } F \\ x \text{ EV von } F \end{array} \right\} \iff \left\{ \begin{array}{l} \lambda \text{ EW von } \mathbf{A} \\ \xi \text{ EV von } \mathbf{A} \end{array} \right.$$

BEWEIS: Aus $Fx = \lambda x$ folgt $\kappa_V(Fx) = \kappa_V(\lambda x) = \lambda \kappa_V x = \lambda \xi$. Aufgrund der Definitionen von κ_V und \mathbf{A} , vgl. (5.26), ist aber $\mathbf{A}\xi = \kappa_V(Fx)$, also $\mathbf{A}\xi = \lambda \xi$. Analog schliesst man unter Verwendung von κ_V^{-1} , dass die Umkehrung gilt. ■

Wir können uns also bei der Diskussion von Eigenwertproblemen weitgehend auf Matrizen beschränken.

Der zu einem Eigenwert λ gehörende Eigenvektor v ist natürlich nie eindeutig bestimmt: offensichtlich ist mit v auch jedes (von 0 verschiedene) skalare Vielfache von v ein Eigenvektor, denn

$$Fv = \lambda v \implies F(\alpha v) = \lambda(\alpha v).$$

Das bedeutet, dass jeder Eigenraum E_λ immer mindestens einen eindimensionalen Unterraum enthält.

Kann es wohl mehrere linear unabhängige Eigenvektoren zu einem Eigenwert geben?

Der Eigenraum E_λ ist die Menge der Vektoren v , für die

$$(F - \lambda I)v = 0 \tag{9.4}$$

gilt, wo I die Identität auf V bedeutet und $F - \lambda I$ definiert ist durch $(F - \lambda I)x := Fx - \lambda x$ ($\forall x \in V$). Mit F ist auch $F - \lambda I$ eine lineare Abbildung von V in sich, und die Eigenvektoren v (zusammen mit 0) bilden gerade den Kern von $F - \lambda I$. Dieser Kern besteht eben genau dann nicht nur aus dem Nullvektor, wenn λ ein Eigenwert ist. Unter Berücksichtigung von Lemma 5.4 gilt somit:

Lemma 9.2 *λ ist genau dann ein Eigenwert von $F : V \rightarrow V$, wenn der Kern von $F - \lambda I$ nicht nur aus dem Nullvektor besteht. Der Eigenraum E_λ zu einem Eigenwert λ von F ist ein vom Nullraum verschiedener Unterraum von V , und zwar ist*

$$E_\lambda = \ker(F - \lambda I). \tag{9.5}$$

DEFINITION: Die **geometrische Vielfachheit** [*geometric multiplicity*] eines Eigenwertes λ ist gleich der Dimension von E_λ . ▲

BEISPIEL 9.1: Die Identität I in V hat den Eigenwert 1 und den zugehörigen Eigenraum $E_1 = V$, denn für jeden Vektor gilt ja $Ix = x = 1x$. Also hat der Eigenwert 1 die geometrische Vielfachheit $n := \dim V$.

Für die Nullmatrix $\mathbf{O} \in \mathbb{E}^{n \times n}$ ist analog $E_0 = \mathbb{E}^n$, $\dim E_0 = n$. ◆

BEISPIEL 9.2: Für eine Rotation im \mathbb{R}^3 um den Winkel ϕ um eine durch den Ursprung gehende Achse mit der Richtung \mathbf{a} ist \mathbf{a} ein Eigenvektor zum Eigenwert 1, denn die Punkte auf der Achse bleiben invariant (fest). Für $|\phi| < \pi$ sind die Ortsvektoren dieser Punkte die einzigen, die fest bleiben, das heißt es ist $E_1 = \text{span} \{\mathbf{a}\}$.

Ist $\phi = \pm\pi$, so gibt es auch noch Eigenvektoren zum Eigenwert -1 : Jeder Vektor in der Ebene, die durch O geht und zu \mathbf{a} senkrecht steht, also jedes \mathbf{x} mit $\langle \mathbf{a}, \mathbf{x} \rangle = 0$, ist dann ein Eigenvektor zum Eigenwert -1 . Ergänzt man $\mathbf{a}/\|\mathbf{a}\|$ zu einer orthonormalen Basis von \mathbb{R}^3 , so hat die Rotation um $\phi = \pm\pi$ bezüglich dieser Basis gerade die Abbildungsmatrix

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

◆

BEISPIEL 9.3: Für die Projektion $\mathbf{P} : \mathbb{R}^3 \rightarrow \mathbb{R}^2 \subset \mathbb{R}^3$, die in cartesianischen Koordinaten definiert ist durch

$$\mathbf{P} : \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}^T \mapsto \begin{pmatrix} x_1 & x_2 & 0 \end{pmatrix}^T,$$

gilt

$$\begin{aligned} \text{im } \mathbf{P} &= \left\{ \begin{pmatrix} x_1 & x_2 & 0 \end{pmatrix}^T ; x_1, x_2 \in \mathbb{R} \right\}, \\ \ker \mathbf{P} &= \left\{ \begin{pmatrix} 0 & 0 & x_3 \end{pmatrix}^T ; x_3 \in \mathbb{R} \right\}. \end{aligned}$$

Zudem ist

$$\begin{aligned} \mathbf{P}\mathbf{x} &= \mathbf{x} = 1\mathbf{x} && \text{falls } \mathbf{x} \in \text{im } \mathbf{P}, \\ \mathbf{P}\mathbf{x} &= \mathbf{o} = 0\mathbf{x} && \text{falls } \mathbf{x} \in \ker \mathbf{P}, \end{aligned}$$

Also hat \mathbf{P} die Eigenwerte 1 und 0, wobei 1 die geometrische Vielfachheit 2 hat, während 0 die geometrische Vielfachheit 1 hat:

$$E_1 = \text{im } \mathbf{P}, \quad \dim E_1 = 2, \quad E_0 = \ker \mathbf{P}, \quad \dim E_0 = 1.$$

Wir werden gleich sehen, dass dies alle Eigenwerte (und damit auch alle Eigenräume) sind.

In den oben vorausgesetzten cartesianischen Koordinaten wird die Projektion beschrieben durch die Diagonalmatrix

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

Völlig analog verhält es sich mit jeder anderen Projektion des \mathbb{R}^3 auf eine (zweidimensionale) Ebene: Die Eigenwerte sind wieder 1, 1, 0, aber die Eigenvektoren hängen natürlich sowohl von der gewählten Projektion als auch von der gewählten Basis ab. Wählt man in der Bildebene im P zwei Basisvektoren und in der Projektionsrichtung $\ker P$ den dritten, so wird die Projektion wieder durch obige Matrix beschrieben. ◆

Auf Matrizen übersetzt, lautet Lemma 9.2 wie folgt:

Korollar 9.3 λ ist genau dann ein Eigenwert von $\mathbf{A} \in \mathbb{E}^{n \times n}$, wenn $\mathbf{A} - \lambda \mathbf{I}$ singular ist.

Der Eigenraum E_λ zu einem Eigenwert λ von \mathbf{A} ist ein vom Nullraum verschiedener Unterraum des \mathbb{E}^n , und zwar ist

$$E_\lambda = \ker(\mathbf{A} - \lambda \mathbf{I}), \quad (9.6)$$

das heisst E_λ besteht aus allen Lösungen \mathbf{v} des homogenen Gleichungssystems

$$(\mathbf{A} - \lambda \mathbf{I})\mathbf{v} = \mathbf{0}. \quad (9.7)$$

Die geometrische Vielfachheit von λ beträgt

$$\dim E_\lambda = \dim \ker(\mathbf{A} - \lambda \mathbf{I}) = n - \text{Rang}(\mathbf{A} - \lambda \mathbf{I}). \quad (9.8)$$

BEWEIS: Das zweite Gleichheitszeichen in (9.8) folgt aus der Dimensionsformel (5.29); siehe auch Satz 5.12. Alles andere ergibt sich aus Lemma 9.2 und bereits bekannten Äquivalenzen. ■

Wir können also die Eigenvektoren zu einem bekannten Eigenwert durch Lösen eines singulären homogenen Gleichungssystems berechnen. Dabei wollen wir in der Regel eine Basis von E_λ konstruieren, das heisst $\dim E_\lambda$ linear unabhängige Lösungsvektoren \mathbf{v} berechnen. Wir können diese zudem bei Bedarf “normieren”, das heisst so wählen, dass $\|\mathbf{v}\| = 1$ ist (in der Euklidischen Norm).

Aber wie finden wir die Eigenwerte? Und wieviele gibt es?

Wir müssen λ so wählen, dass $\mathbf{A} - \lambda \mathbf{I}$ singular ist. Nach Satz 8.5 ist $\mathbf{A} - \lambda \mathbf{I}$ genau dann singular, wenn $\det(\mathbf{A} - \lambda \mathbf{I}) = 0$.

BEISPIEL 9.4: Für

$$\mathbf{A} = \begin{pmatrix} -7 & 2 & -6 \\ 12 & -2 & 12 \\ 12 & -3 & 11 \end{pmatrix} \quad (9.9)$$

erhalten wir mit der Formel (8.7) von Sarrus

$$\begin{aligned} \det(\mathbf{A} - \lambda \mathbf{I}) &= \begin{vmatrix} -7 - \lambda & 2 & -6 \\ 12 & -2 - \lambda & 12 \\ 12 & -3 & 11 - \lambda \end{vmatrix} \\ &= (-7 - \lambda)(-2 - \lambda)(11 - \lambda) + 72(-2 - \lambda) \\ &\quad + 36(-7 - \lambda) - 24(11 - \lambda) + 288 + 216 \\ &= (-\lambda)^3 + (-7 - 2 + 11)\lambda^2 \\ &\quad - (-22 - 77 + 14 + 72 + 36 - 24)\lambda \\ &\quad + 154 - 144 - 252 - 264 + 288 + 216 \\ &= -\lambda^3 + 2\lambda^2 + \lambda - 2. \end{aligned}$$

Es ist also

$$\det(\mathbf{A} - \lambda \mathbf{I}) = -\lambda^3 + 2\lambda^2 + \lambda - 2 \equiv: \chi_{\mathbf{A}}(\lambda)$$

ein kubisches Polynom $\chi_{\mathbf{A}}$ in λ , und das gilt offensichtlich analog für jede 3×3 -Matrix. Ein solches Polynom hat im Komplexen genau drei

Nullstellen, wenn man diese mit ihrer Vielfachheit zählt. Es gibt also drei Eigenwerte, wobei aber zwei davon oder alle drei zusammenfallen können. Zudem können zwei Nullstellen konjugiert-komplex sein.

In diesem Beispiel kann man erraten, dass $\lambda_1 = 1$ eine Nullstelle von $\chi_{\mathbf{A}}$ ist. Das **Abspalten** [*deflation*] dieser Nullstelle, das heisst die Division des Polynoms $\chi_{\mathbf{A}}$ durch den Linearfaktor $(1 - \lambda)$, liefert

$$(-\lambda^3 + 2\lambda^2 + \lambda - 2) : (1 - \lambda) = \lambda^2 - \lambda - 2,$$

also ein quadratisches Polynom mit den Nullstellen $\lambda_2 = 2$ und $\lambda_3 = -1$. Somit hat \mathbf{A} das Spektrum

$$\sigma(\mathbf{A}) = \{1, 2, -1\}.$$



BEISPIEL 9.5: Dass es wirklich komplexe Eigenwerte geben kann, selbst wenn \mathbf{A} reell ist, zeigt schon das einfache Beispiel

$$\mathbf{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1.$$

Die Eigenwerte sind hier rein imaginär:

$$\lambda_1 = i, \quad \lambda_2 = -i.$$



Aufgrund der Determinanten-Definition (8.6) sieht man, dass

$$\begin{aligned} \det(\mathbf{A} - \lambda\mathbf{I}) &= \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} - \lambda \end{vmatrix} \\ &= (-\lambda)^n + (a_{11} + \cdots + a_{nn})(-\lambda)^{n-1} + \cdots + \det \mathbf{A} \\ &\equiv: \chi_{\mathbf{A}}(\lambda) \end{aligned} \tag{9.10}$$

ein Polynom $\chi_{\mathbf{A}}$ vom Grade n in λ ist. Dessen Höchstkoeffizient ist $(-1)^n$, der konstante Koeffizient ist gerade $\det \mathbf{A}$.

Der zweithöchste Koeffizient ist das $(-1)^{n-1}$ -fache der Summe der Diagonalelemente von \mathbf{A} .

DEFINITION: Das Polynom

$$\boxed{\chi_{\mathbf{A}}(\lambda) := \det(\mathbf{A} - \lambda\mathbf{I})} \tag{9.11}$$

heisst **charakteristisches Polynom** [*characteristic polynomial*] der Matrix $\mathbf{A} \in \mathbb{E}^{n \times n}$, und die Gleichung

$$\boxed{\chi_{\mathbf{A}}(\lambda) = 0} \tag{9.12}$$

ist die **charakteristische Gleichung** [*characteristic equation*].

Die Summe der Diagonalelemente von \mathbf{A} nennt man **Spur** [*trace*] von \mathbf{A} :

$$\boxed{\text{Spur } \mathbf{A} := a_{11} + a_{22} + \cdots + a_{nn}.} \quad (9.13)$$

Lemma 9.4 *Das charakteristische Polynom $\chi_{\mathbf{A}}$ hat die Form*

$$\begin{aligned} \chi_{\mathbf{A}}(\lambda) &\equiv: \det(\mathbf{A} - \lambda \mathbf{I}) \\ &= (-\lambda)^n + \text{Spur } \mathbf{A} (-\lambda)^{n-1} + \cdots + \det \mathbf{A}. \end{aligned} \quad (9.14)$$

Aus dem Vorgehenden folgt sofort:

Satz 9.5 *$\lambda \in \mathbb{E}$ ist genau dann Eigenwert der Matrix $\mathbf{A} \in \mathbb{E}^{n \times n}$, wenn λ Nullstelle des charakteristischen Polynoms $\chi_{\mathbf{A}}$ ist, das heisst eine Lösung der charakteristischen Gleichung ist.*

Das Polynom $\chi_{\mathbf{A}}$ hat natürlich reelle Koeffizienten, wenn \mathbf{A} reell ist ($\mathbb{E} = \mathbb{R}$); es hat komplexe Koeffizienten, wenn \mathbf{A} komplex ist ($\mathbb{E} = \mathbb{C}$). Nach dem Fundamentalsatz der Algebra hat ein komplexes (oder reelles) Polynom vom exakten Grad n genau n (im allgemeinen komplexe) Nullstellen, falls man mehrfache Nullstellen mehrfach zählt.

Das heisst man kann $\chi_{\mathbf{A}}$, das den Höchstkoeffizienten $(-1)^n$ hat, wie folgt in n Linearfaktoren zerlegen:

$$\chi_{\mathbf{A}}(\lambda) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda),$$

wobei $\lambda_k \in \mathbb{C}$, $k = 1, \dots, n$.

Ein reelles Polynom kann aber auch nicht-reelle Nullstellen haben. Falls \mathbf{A} reell ist, aber λ nicht reell ist, so sind auch die zugehörigen Eigenvektoren nicht reell. (Denn wäre $\mathbf{v} \in \mathbb{R}^n$ und $\lambda \notin \mathbb{R}$, so wäre $\mathbf{A}\mathbf{v} \in \mathbb{R}^n$, aber $\mathbf{v}\lambda \notin \mathbb{R}^n$.)

In der Eigenwerttheorie fasst man deshalb in der Regel reelle Matrizen als komplexe Matrizen auf, das heisst man lässt komplexe Eigenwerte und komplexe Eigenvektoren zu. *Deshalb setzen wir im folgenden $\mathbb{E} = \mathbb{C}$ voraus.*

Es wird sich zeigen, dass die bereits definierte geometrische Vielfachheit eines Eigenwertes λ nicht unbedingt gleich seiner Vielfachheit als Nullstelle von $\chi_{\mathbf{A}}$ ist. Man definiert deshalb:

DEFINITION: Die **algebraische Vielfachheit** [*algebraic multiplicity*] eines Eigenwertes λ ist die Vielfachheit von λ als Nullstelle des charakteristischen Polynoms. ▲

Das Verhältnis zwischen algebraischer und geometrischer Vielfachheit wird in Abschnitt 9.4 diskutiert werden.

Wir können also bei kleinen Matrizen die Eigenwerte und Eigenvektoren nach der folgenden Methode berechnen, die aber für grössere Matrizen ($n \gg 3$) unbrauchbar ist, weil zu aufwendig und zu starken Rundungseffekten unterworfen.

Algorithmus 9.1 (Berechnung von Eigenwerten und Eigenvektoren via charakteristisches Polynom)

Um die Eigenwerte und Eigenvektoren einer Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ zu bestimmen, kann man theoretisch wie folgt vorgehen:

1. Berechnung des charakteristischen Polynoms $\chi_{\mathbf{A}}$,

$$\chi_{\mathbf{A}}(\lambda) := \det(\mathbf{A} - \lambda \mathbf{I}).$$

2. Berechnung der n Nullstellen $\lambda_1, \dots, \lambda_n$ von $\chi_{\mathbf{A}}$. Die Vielfachheit einer Nullstelle ist gleich der algebraischen Vielfachheit dieses Eigenwertes.

3. Für jeden verschiedenen Eigenwert λ_k : Bestimmung einer Basis des Kernes von $\mathbf{A} - \lambda_k \mathbf{I}$, des Eigenraumes zu λ_k . Das heisst Berechnung (maximal vieler) linear unabhängiger Lösungen des singulären homogenen Gleichungssystems

$$(\mathbf{A} - \lambda_k \mathbf{I}) \mathbf{v} = \mathbf{0}.$$

Dazu reduziert man $\mathbf{A} - \lambda_k \mathbf{I}$ mit dem Gauss-Algorithmus auf Zeilenstufenform und wählt von den $n - r$ freien Parametern der Reihe nach immer einen $\neq 0$ und die anderen 0. Die Dimension $n - r$ des Lösungsraumes ist gleich der geometrischen Vielfachheit dieses Eigenwertes.

BEISPIEL 9.6: Für die Matrix

$$\mathbf{A} = \begin{pmatrix} 5 & -1 & 3 \\ 8 & -1 & 6 \\ -4 & 1 & -2 \end{pmatrix} \quad (9.15)$$

wird

$$\begin{aligned} \chi_{\mathbf{A}}(\lambda) &= \begin{vmatrix} 5 - \lambda & -1 & 3 \\ 8 & -1 - \lambda & 6 \\ -4 & 1 & -2 - \lambda \end{vmatrix} \\ &= (5 - \lambda)(-1 - \lambda)(-2 - \lambda) + 12(-1 - \lambda) \\ &\quad - 6(5 - \lambda) + 8(-2 - \lambda) + 24 + 24 \\ &= -\lambda(\lambda^2 - 2\lambda + 1). \end{aligned}$$

Die charakteristische Gleichung lässt sich also schreiben als

$$0 = -\lambda(\lambda^2 - 2\lambda + 1) = -\lambda(\lambda - 1)^2,$$

d.h. die Eigenwerte sind

$$\lambda_1 = 1, \quad \lambda_2 = 1, \quad \lambda_3 = 0. \quad (9.16)$$

Zur Bestimmung der Eigenvektoren sind nacheinander die zwei homogenen Gleichungssysteme

$$\begin{array}{ccc|c} \xi_1 & \xi_2 & \xi_3 & 1 \\ \hline 4 & -1 & 3 & 0 \\ 8 & -2 & 6 & 0 \\ -4 & 1 & -3 & 0 \end{array} \quad \begin{array}{ccc|c} \xi_1 & \xi_2 & \xi_3 & 1 \\ \hline 5 & -1 & 3 & 0 \\ 8 & -1 & 6 & 0 \\ -4 & 1 & -2 & 0 \end{array}$$

durch Reduktion auf Zeilenstufenform zu lösen. Vertauscht man im ersten System keine, im zweiten System die erste mit der dritten Zeile, so ergibt sich:

$$\begin{array}{c|ccc|c}
 \xi_1 & \xi_2 & \xi_3 & 1 \\
 \hline
 \textcircled{4} & -1 & 3 & 0 \\
 \hline
 0 & 0 & 0 & 0 \\
 \hline
 0 & 0 & 0 & 0
 \end{array}
 \quad
 \begin{array}{c|ccc|c}
 \xi_1 & \xi_2 & \xi_3 & 1 \\
 \hline
 \textcircled{-4} & 1 & -2 & 0 \\
 \hline
 0 & \textcircled{1} & 2 & 0 \\
 \hline
 0 & 0 & 0 & 0
 \end{array}$$

Das System links hat Rang 1 und die linear unabhängigen Lösungen $\mathbf{v}_1 = (1 \ 4 \ 0)^\top$ und $\mathbf{v}_2 = (3 \ 0 \ -4)^\top$. Man kann in \mathbf{v}_2 statt die zweite die erste Komponente null setzen und bekommt dann $\mathbf{v}_2 = (0 \ 3 \ 1)^\top$. Das System rechts hat Rang 2 und zum Beispiel die Lösung $\mathbf{v}_3 = (-1 \ -2 \ 1)^\top$. Wir erhalten also zu den Eigenwerten (9.16) die drei Eigenvektoren

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} -1 \\ -2 \\ 1 \end{pmatrix}. \quad (9.17)$$

◆

Wir weisen noch darauf hin, dass die Matrix (9.15) des Beispiels 9.6 singular ist. Es gilt nämlich allgemein:

Lemma 9.6 *Eine (quadratische) Matrix \mathbf{A} ist genau dann singular, wenn sie 0 als Eigenwert hat:*

$$\mathbf{A} \text{ singular} \iff 0 \in \sigma(\mathbf{A}).$$

BEWEIS: Nach Satz 1.7 ist \mathbf{A} genau dann singular, wenn das homogene System $\mathbf{A}\mathbf{x} = \mathbf{o}$ eine nichttriviale Lösung \mathbf{x} hat. Dies ist aber gleichbedeutend damit, dass \mathbf{x} ein Eigenvektor zum Eigenwert 0 ist.

Alternativer Beweis: Nach Satz 8.5 ist \mathbf{A} genau dann singular, wenn $\det \mathbf{A} = 0$. Dies bedeutet aber nach (9.10), dass der konstante Term im charakterischen Polynom verschwindet, dieses also 0 als Nullstelle hat, was nach Satz 9.5 heisst, dass 0 Eigenwert von \mathbf{A} ist. ■

9.2 Ähnlichkeitstransformationen; die Eigenwertzerlegung

Wir betrachten nun wieder eine lineare Abbildung $F : V \rightarrow V$, $x \mapsto Fx$ eines Vektorraumes V der Dimension n in sich und die zugehörigen Abbildungsmatrizen \mathbf{A} und \mathbf{B} bezüglich zwei verschiedenen Basen $\{b_1, \dots, b_n\}$ und $\{b'_1, \dots, b'_n\}$. Analog zu Abschnitt 5.5, wo jedoch Urbild- und Bildraum verschieden waren, gilt das folgende kommutatives Diagramm:

$$\begin{array}{ccc}
 x \in V & \xrightarrow[\text{lin. Abb.}]{F} & y \in V \\
 \kappa_V \Big\| \Big\| \kappa_V^{-1} & & \kappa_V \Big\| \Big\| \kappa_V^{-1} \quad \text{(Koordinaten-} \\
 & & \text{abbildung bzgl.} \\
 & & \text{“alten” Basen)} \\
 \xi \in \mathbb{E}^n & \xrightarrow[\text{Abb.matrix}]{\mathbf{A}} & \eta \in \mathbb{E}^n \quad \text{(Koordinaten} \\
 & & \text{bzgl. “alten”} \\
 & & \text{Basen)} \quad (9.18) \\
 \mathbf{T}^{-1} \Big\| \Big\| \mathbf{T} & & \mathbf{T}^{-1} \Big\| \Big\| \mathbf{T} \quad \text{(Koordinaten-} \\
 & & \text{transformation)} \\
 \xi' \in \mathbb{E}^n & \xrightarrow[\text{Abb.matrix}]{\mathbf{B}} & \eta' \in \mathbb{E}^n \quad \text{(Koordinaten} \\
 & & \text{bzgl. “neuen”} \\
 & & \text{Basen)}
 \end{array}$$

Wir erinnern uns daran, vgl. (5.52), dass nach diesem Diagramm gilt

$$\boxed{\mathbf{B} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}, \quad \mathbf{A} = \mathbf{T}\mathbf{B}\mathbf{T}^{-1}}, \quad (9.19)$$

was bedeutet, dass die $n \times n$ -Matrizen \mathbf{A} und \mathbf{B} **ähnlich** [*similar*] sind. Der Übergang $\mathbf{A} \mapsto \mathbf{B} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$ wird als **Ähnlichkeitstransformation** [*similarity transformation*] bezeichnet.

Wir stellen uns nun wieder die Frage, wieweit sich die Abbildungsmatrix durch geeignete Wahl der Basis vereinfachen lässt, das heißt, wie weit man die Matrix \mathbf{A} durch eine Ähnlichkeitstransformation $\mathbf{A} \mapsto \mathbf{B} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$ vereinfachen kann. Diese Frage ist eng verknüpft mit der Eigenwerttheorie.

Satz 9.7 *Ähnlichen Matrizen haben dasselbe charakteristische Polynom; sie haben also die gleiche Determinante, die gleiche Spur und die gleichen Eigenwerte.*

Sowohl die geometrische als auch die algebraische Vielfachheit eines Eigenwertes ist bei ähnlichen Matrizen gleich.

BEWEIS: Es sei $\mathbf{B} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$. Dann gilt für die charakteristischen Polynome von \mathbf{A} und \mathbf{B}

$$\begin{aligned}\chi_{\mathbf{B}}(\lambda) &= \det(\mathbf{B} - \lambda\mathbf{I}) \\ &= \det(\mathbf{T}^{-1}\mathbf{A}\mathbf{T} - \lambda\mathbf{I}) \\ &= \det(\mathbf{T}^{-1}(\mathbf{A} - \lambda\mathbf{I})\mathbf{T}) \\ &= \det \mathbf{T}^{-1} \det(\mathbf{A} - \lambda\mathbf{I}) \det \mathbf{T} \\ &= \det(\mathbf{A} - \lambda\mathbf{I}) \\ &= \chi_{\mathbf{A}}(\lambda)\end{aligned}$$

Folglich stimmen nach Lemma 9.4 und Satz 9.5 die Determinanten von \mathbf{A} und \mathbf{B} , ihre Spuren und ihre Eigenwerte inklusive der algebraischen Vielfachheit überein.

Dass auch die geometrische Vielfachheit übereinstimmt, folgt schon aus Lemma 9.1. Direkter sehen wir, dass aus $\mathbf{A}\mathbf{v} = \mathbf{v}\lambda$ folgt

$$\underbrace{\mathbf{T}^{-1}\mathbf{A}\mathbf{T}}_{=\mathbf{B}} \underbrace{\mathbf{T}^{-1}\mathbf{v}}_{\equiv:\mathbf{w}} = \underbrace{\mathbf{T}^{-1}\mathbf{v}}_{\equiv:\mathbf{w}} \lambda,$$

also $\mathbf{B}\mathbf{w} = \mathbf{w}\lambda$ falls $\mathbf{w} := \mathbf{T}^{-1}\mathbf{v}$. Umgekehrt schliesst man genauso. Das heisst es ist \mathbf{v} Eigenvektor von \mathbf{A} zum Eigenwert λ genau dann, wenn \mathbf{w} Eigenvektor von \mathbf{B} zu λ ist. Kurz, wenn wir \mathbf{T} und \mathbf{T}^{-1} als Abbildungen auffassen:

$$E_{\lambda}(\mathbf{B}) = \mathbf{T}^{-1}E_{\lambda}(\mathbf{A}), \quad E_{\lambda}(\mathbf{A}) = \mathbf{T}E_{\lambda}(\mathbf{B}) \quad (9.20)$$

Da \mathbf{T} regulär ist, folgt $\dim E_{\lambda}(\mathbf{B}) = \dim E_{\lambda}(\mathbf{A})$. ■

Gibt es in \mathbf{V} eine Basis aus lauter Eigenvektoren (was nicht immer der Fall ist), so ist die entsprechende Abbildungsmatrix diagonal:

Lemma 9.8 *Eine zu $F : V \rightarrow V$ gehörende Abbildungsmatrix ist genau dann diagonal, wenn die gewählte Basis von V aus lauter Eigenvektoren von F besteht.*

BEWEIS: Sind v_1, \dots, v_n Eigenvektoren, die eine Basis bilden, und sind $\lambda_1, \dots, \lambda_n$ die zugehörigen Eigenwerte, so hat $Fv_l = \lambda_l v_l$ den Koordinatenvektor

$$(0 \ \cdots \ 0 \ \lambda_l \ 0 \ \cdots \ 0)^{\top}$$

(mit λ_l als l -ter Komponente), das heisst die Abbildungsmatrix ist $\mathbf{A} = \text{diag}(\lambda_1, \dots, \lambda_n)$, vgl. (5.14). Dieser Schluss ist zudem umkehrbar. ■

DEFINITION: Eine Basis aus Eigenvektoren von F (oder \mathbf{A}) heisst **Eigenbasis** [*eigen basis*] von F (bzw. \mathbf{A}). ▲

Gibt es eine Eigenbasis v_1, \dots, v_n , so gilt also die einfache Abbildungsformel

$$\boxed{x = \sum_{k=1}^n \xi_k v_k \quad \longmapsto \quad Fx = \sum_{k=1}^n \lambda_k \xi_k v_k} \quad (9.21)$$

In diesem Falle gibt es also zu F eine extrem einfache Abbildungsmatrix, ähnlich wie das allgemein gilt für eine lineare Abbildung $F : V \rightarrow W$ zwischen zwei verschiedenen Räumen, wo man in beiden die Basis frei wählen kann, vgl. Satz 5.20, wo die Diagonalelemente sogar alle 0 oder 1 sind. Aber während man dort im wesentlichen nur Basen von Kern und Bild bestimmen muss, sowie die Urbilder der letzteren, ist hier die Bestimmung einer Basis aus Eigenvektoren eine rechnerisch sehr aufwendige Aufgabe, die man zudem im allgemeinen nur approximativ lösen kann, falls $n > 3$.

Bereits in Lemma 9.1 haben wir gesehen, dass man sich für die Bestimmung der Eigenwerte und Eigenvektoren einer linearen Selbstabbildung F auf die Bestimmung der Eigenwerte und Eigenvektoren einer zugehörigen Abbildungsmatrix \mathbf{A} beschränken kann.

Für Matrizen gilt die folgende Neuformulierung von Lemma 9.8:

Satz 9.9 *Zu einer Matrix $\mathbf{A} \in \mathbb{E}^{n \times n}$ gibt es genau dann eine ähnliche Diagonalmatrix $\mathbf{\Lambda}$, wenn es eine Eigenbasis von \mathbf{A} gibt. Für die reguläre Matrix*

$$\mathbf{V} := (\mathbf{v}_1 \ \dots \ \mathbf{v}_n)$$

mit dieser Eigenbasis als Kolonnen gilt dann

$$\boxed{\mathbf{A} \mathbf{V} = \mathbf{V} \mathbf{\Lambda}}, \quad \text{d.h.} \quad \boxed{\mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^{-1}}. \quad (9.22)$$

Gibt es umgekehrt eine reguläre Matrix $\mathbf{V} \in \mathbb{E}^{n \times n}$ und eine Diagonalmatrix $\mathbf{\Lambda} \in \mathbb{E}^{n \times n}$, so dass (9.22) gilt, so sind die Diagonalelemente von $\mathbf{\Lambda}$ Eigenwerte von \mathbf{A} und die Kolonnen von \mathbf{V} entsprechende Eigenvektoren, die eine Eigenbasis bilden.

BEWEIS: Der erste Teil des Satzes ist ein Spezialfall von Lemma 9.8. Setzen wir eine Basis aus Eigenvektoren voraus, so dass

$$\boxed{\mathbf{A} \mathbf{v}_k = \mathbf{v}_k \lambda_k} \quad (k = 1, \dots, n) \quad (9.23)$$

gilt, ist klar, dass man diese n Gleichungen zur Matrix-Gleichung (9.22) zusammenfassen kann, wobei \mathbf{V} linear unabhängige Kolonnen hat, also regulär ist. Dieser Schluss lässt sich auch umkehren. ■

DEFINITION: $\mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^{-1}$ heisst **Spektralzerlegung** [*spectral decomposition*] oder **Eigenwertzerlegung** [*eigenvalue decomposition*] von \mathbf{A} .

Eine Matrix \mathbf{A} , zu der es eine Spektralzerlegung $\mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^{-1}$ mit diagonalem $\mathbf{\Lambda}$ gibt, heisst **diagonalisierbar** [*diagonalizable*]. ▲

Eine diagonalisierbare $n \times n$ -Matrix lässt sich auf Grund der Spektralzerlegung (9.22) als Summe von n Rang-1-Matrizen schreiben. Genauer gilt:

Korollar 9.10 Ist \mathbf{A} diagonalisierbar und zerlegen wir die Matrizen \mathbf{V} und \mathbf{V}^{-1} aus der Spektralzerlegung $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$ in Kolonnen- bzw. Zeilenvektoren,

$$\mathbf{V} = (\mathbf{v}_1 \ \cdots \ \mathbf{v}_n), \quad \mathbf{V}^{-1} = \begin{pmatrix} \mathbf{w}_1^\top \\ \vdots \\ \mathbf{w}_n^\top \end{pmatrix}, \quad (9.24)$$

so lässt sich \mathbf{A} als Summe von n Rang-1-Matrizen darstellen:

$$\mathbf{A} = \sum_{k=1}^n \mathbf{v}_k \lambda_k \mathbf{w}_k^\top. \quad (9.25)$$

Dabei gilt $\mathbf{A}\mathbf{v}_k = \mathbf{v}_k \lambda_k$ und $\mathbf{w}_k^\top \mathbf{A} = \lambda_k \mathbf{w}_k^\top$.

DEFINITION: Ein Vektor \mathbf{w} , für den gilt $\mathbf{w}^\top \mathbf{A} = \lambda_k \mathbf{w}^\top$, heisst **linker Eigenvektor** [left eigenvector] von \mathbf{A} . ▲

BEISPIEL 9.7: Die Matrix-Gleichung

$$\underbrace{\begin{pmatrix} -7 & 2 & -6 \\ 12 & -2 & 12 \\ 12 & -3 & 11 \end{pmatrix}}_{\mathbf{A}} \underbrace{\begin{pmatrix} 1 & 0 & -1 \\ 4 & 3 & 0 \\ 0 & 1 & 1 \end{pmatrix}}_{\mathbf{V}} = \underbrace{\begin{pmatrix} 1 & 0 & -1 \\ 4 & 3 & 0 \\ 0 & 1 & 1 \end{pmatrix}}_{\mathbf{V}} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}}_{\mathbf{\Lambda}}$$

ist gleichwertig mit den Beziehungen $\mathbf{A}\mathbf{v}_k = \mathbf{v}_k \lambda_k$ ($k = 1, \dots, 3$), wenn $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = -1$ und

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

Zudem ist

$$\mathbf{V}^{-1} = \begin{pmatrix} 1 & 0 & -1 \\ 4 & 3 & 0 \\ 0 & 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} -3 & 1 & -3 \\ 4 & -1 & 4 \\ -4 & 1 & -3 \end{pmatrix},$$

also

$$\mathbf{A} = \underbrace{\begin{pmatrix} 1 & 0 & -1 \\ 4 & 3 & 0 \\ 0 & 1 & 1 \end{pmatrix}}_{\mathbf{V}} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}}_{\mathbf{\Lambda}} \underbrace{\begin{pmatrix} -3 & 1 & -3 \\ 4 & -1 & 4 \\ -4 & 1 & -3 \end{pmatrix}}_{\mathbf{V}^{-1}}.$$

Die Zerlegung (9.24) in Rang-1-Matrizen lautet hier

$$\begin{aligned} \mathbf{A} &= \begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix} \begin{pmatrix} -3 & 1 & -3 \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} 2 \begin{pmatrix} 4 & -1 & 4 \end{pmatrix} \\ &\quad - \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \begin{pmatrix} -4 & 1 & -3 \end{pmatrix}, \end{aligned}$$

oder ausgeschrieben

$$\begin{aligned} \mathbf{A} &= \begin{pmatrix} -3 & 1 & -3 \\ 12 & 4 & -12 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 24 & -6 & 24 \\ 8 & -2 & 8 \end{pmatrix} + \begin{pmatrix} -4 & 1 & -3 \\ 0 & 0 & 0 \\ 4 & -1 & 3 \end{pmatrix} \\ &= \begin{pmatrix} -7 & 2 & -6 \\ 12 & -2 & 12 \\ 12 & -3 & 11 \end{pmatrix}. \end{aligned}$$

◆

Wie findet man aber allgemein eine Basis aus Eigenvektoren? Liefert unser Algorithmus 9.1 eine, falls es eine gibt?

In Beispiel 9.7 gab es drei verschiedene Eigenwerte und die zugehörigen Eigenvektoren sind linear unabhängig, denn andernfalls wäre \mathbf{V} nicht invertierbar. In der Tat gilt allgemein:

Satz 9.11 *Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig.*

BEWEIS: Der Satz ist richtig für $m = 1$, da o nicht ein Eigenvektor ist. Wir nehmen an, der Satz sei richtig für m Eigenvektoren v_1, \dots, v_m zu m verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_m$. Es sei nun λ_{m+1} ein weiterer anderer Eigenwert und v_{m+1} ($\neq o$) ein zugehöriger Eigenvektor. Wäre

$$v_{m+1} = \sum_{k=1}^m \gamma_k v_k, \quad (9.26)$$

so würde folgen

$$Fv_{m+1} = \sum_{k=1}^m \gamma_k Fv_k,$$

also

$$\lambda_{m+1} v_{m+1} = \sum_{k=1}^m \gamma_k \lambda_k v_k.$$

Wegen (9.26) hätte man demnach

$$\sum_{k=1}^m \gamma_k \underbrace{(\lambda_{m+1} - \lambda_k)}_{\neq 0} v_k = o.$$

Da v_1, \dots, v_m linear unabhängig sind, folgt $\gamma_1 = \dots = \gamma_m = 0$, im Widerspruch zu (9.26) und $v_{m+1} \neq o$. ■

Korollar 9.12 *Sind die n Eigenwerte von $F : V \rightarrow V$ (wo $n = \dim V$) verschieden, so gibt es eine Basis von Eigenvektoren und die entsprechende Abbildungsmatrix ist diagonal.*

Das folgende Beispiel zeigt, dass die Umkehrung dieses Korollars nicht gilt: es gibt auch Matrizen mit mehrfachen Eigenwerten, die diagonalisierbar sind. Das wissen wir allerdings auch schon von der Einheitsmatrix \mathbf{I} .

BEISPIEL 9.8: Wir setzen Beispiel 9.6 fort: Zur Matrix \mathbf{A} aus (9.15) gehört gemäss (9.16)–(9.17) die Identität

$$\underbrace{\begin{pmatrix} 5 & -1 & 3 \\ 8 & -1 & 6 \\ -4 & 1 & -2 \end{pmatrix}}_{\mathbf{A}} \underbrace{\begin{pmatrix} 1 & 0 & -1 \\ 4 & 3 & -2 \\ 0 & 1 & 1 \end{pmatrix}}_{\mathbf{V}} = \underbrace{\begin{pmatrix} 1 & 0 & -1 \\ 4 & 3 & -2 \\ 0 & 1 & 1 \end{pmatrix}}_{\mathbf{V}} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{\mathbf{\Lambda}}.$$

Die Kolonnen \mathbf{v}_1 und \mathbf{v}_2 von \mathbf{V} sind gemäss Konstruktion linear unabhängige Eigenvektoren zum Eigenwert $\lambda_1 = \lambda_2 = 1$ mit algebraischer und geometrischer Vielfachheit 2. Der Eigenvektor \mathbf{v}_3 gehört zu $\lambda_3 = 0$ und ist gemäss Satz 9.11 linear unabhängig zu jeder Linearkombination von \mathbf{v}_1 und \mathbf{v}_2 . Also sind die drei Vektoren $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ linear unabhängig, das heisst sie bilden eine Eigenbasis, und \mathbf{V} ist regulär. In der Tat ist

$$\mathbf{V}^{-1} = \begin{pmatrix} 5 & -1 & 3 \\ -4 & 1 & -2 \\ 4 & -1 & 3 \end{pmatrix},$$

und somit hat \mathbf{A} die Spektralzerlegung

$$\mathbf{A} = \underbrace{\begin{pmatrix} 1 & 0 & -1 \\ 4 & 3 & -2 \\ 0 & 1 & 1 \end{pmatrix}}_{\mathbf{V}} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}}_{\mathbf{\Lambda}} \underbrace{\begin{pmatrix} 5 & -1 & 3 \\ -4 & 1 & -2 \\ 4 & -1 & 3 \end{pmatrix}}_{\mathbf{V}^{-1}}.$$

◆

Es gibt leider aber auch nicht-diagonalisierbare Matrizen. Hier ist das einfachste Beispiel:

BEISPIEL 9.9: Die Matrix

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

hat das charakteristische Polynom $\chi_{\mathbf{A}}(\lambda) = (1 - \lambda)^2$, also den doppelten Eigenwert 1. Aber das Gleichungssystem $(\mathbf{A} - 1\mathbf{I})\mathbf{v} = 0$ mit der Koeffizientenmatrix

$$\mathbf{A} - 1\mathbf{I} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

hat Rang 1. Es ist also $n - r = 2 - 1 = 1$, das heisst der Eigenraum E_1 ist eindimensional; die geometrische Vielfachheit 1 des Eigenwertes ist damit kleiner als die algebraische Vielfachheit 2. ◆

Dieses Beispiel zeigt, dass die geometrische Vielfachheit kleiner als die algebraische sein kann, was bedeutet, dass es dann keine Eigenbasis geben kann. Wie kann man wohl diese Situation charakterisieren? Und ist auch das Umgekehrte möglich?

Es gelten die folgenden zwei grundlegenden Aussagen:

Satz 9.13 *Für jeden Eigenwert gilt, dass die geometrische Vielfachheit kleiner gleich der algebraischen Vielfachheit ist.*

Weil der Grad von $\chi_{\mathbf{A}}$ genau n ist, sich also die algebraischen Vielfachheiten zu n summieren, folgt sofort die Notwendigkeit der folgenden Bedingung.

Satz 9.14 Eine Matrix ist genau dann diagonalisierbar, wenn für jeden Eigenwert die geometrische gleich der algebraischen Vielfachheit ist.

BEWEIS von Satz 9.13: Es sei λ_0 ein fester Eigenwert mit geometrischer Vielfachheit ν . Wir wählen nun eine Basis $\mathbf{v}_1, \dots, \mathbf{v}_\nu \in \mathbb{E}^n$ von E_{λ_0} aus, die wir nach Satz 4.9 zu einer Basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ von \mathbb{E}^n ergänzen können, welche die Kolonnen einer regulären Matrix \mathbf{V} liefert. Es ist dann

$$\begin{aligned} \mathbf{A}\mathbf{V} &= \mathbf{A} \begin{pmatrix} \mathbf{v}_1 & \dots & \mathbf{v}_n \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{A}\mathbf{v}_1 & \dots & \mathbf{A}\mathbf{v}_\nu & \mathbf{A}\mathbf{v}_{\nu+1} & \dots & \mathbf{A}\mathbf{v}_n \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{v}_1\lambda_0 & \dots & \mathbf{v}_\nu\lambda_0 & \mathbf{A}\mathbf{v}_{\nu+1} & \dots & \mathbf{A}\mathbf{v}_n \end{pmatrix}. \end{aligned}$$

Wegen $\mathbf{V}^{-1}\mathbf{v}_k = \mathbf{e}_k$ folgt, dass $\mathbf{C} := \mathbf{V}^{-1}\mathbf{A}\mathbf{V}$ folgende Form hat:

$$\begin{aligned} \mathbf{C} &:= \mathbf{V}^{-1}\mathbf{A}\mathbf{V} \\ &= \begin{pmatrix} \mathbf{e}_1\lambda_0 & \dots & \mathbf{e}_\nu\lambda_0 & \mathbf{V}^{-1}\mathbf{A}\mathbf{v}_{\nu+1} & \dots & \mathbf{V}^{-1}\mathbf{A}\mathbf{v}_n \end{pmatrix} \\ &= \begin{pmatrix} \lambda_0\mathbf{I}_\nu & \mathbf{E} \\ \mathbf{O} & \mathbf{F} \end{pmatrix}. \end{aligned}$$

Aus der Formel (8.6) für die Determinante kann man sehen, dass gilt

$$\begin{aligned} \det(\mathbf{C} - \lambda\mathbf{I}) &= \det \begin{pmatrix} (\lambda_0 - \lambda)\mathbf{I}_\nu & \mathbf{E} \\ \mathbf{O} & \mathbf{F} - \lambda\mathbf{I}_{n-\nu} \end{pmatrix} \\ &= \det((\lambda_0 - \lambda)\mathbf{I}_\nu) \det(\mathbf{F} - \lambda\mathbf{I}_{n-\nu}), \end{aligned}$$

also nach Satz 9.7 und Determinanten-Eigenschaft v) aus Lemma 8.4

$$\chi_{\mathbf{A}}(\lambda) = \chi_{\mathbf{C}}(\lambda) = \det(\mathbf{C} - \lambda\mathbf{I}) = (\lambda_0 - \lambda)^\nu \det(\mathbf{F} - \lambda\mathbf{I}_{n-\nu}).$$

Wir schliessen daraus, dass λ_0 als Eigenwert von \mathbf{A} mindestens die algebraische Vielfachheit ν hat. ■

BEWEIS von Satz 9.14: Aus Lemma 9.8 wissen wir bereits, dass eine diagonale Abbildungsmatrix genau dann existiert, wenn es eine Eigenbasis gibt. Es gebe m verschiedene Eigenvektoren. Wir wählen nun in jedem Eigenraum E_{λ_k} ($k = 1, \dots, m$) eine Basis $\{\mathbf{v}_{k,j}; j = 1, \dots, \nu_k\}$.

Sind für alle Eigenwerte geometrische und algebraische Vielfachheit gleich, so ist $\nu_1 + \dots + \nu_m = n$, und wir behaupten, dass alle diese n Basisvektoren linear unabhängig sind. Ist nämlich

$$\underbrace{\sum_{k=1}^m \sum_{j=1}^{\nu_k} \mathbf{v}_{k,j} \gamma_{k,j}}_{\equiv: \mathbf{c}_k} = \mathbf{o}, \quad (9.27)$$

so sind die in verschiedenen Eigenräumen liegenden Vektoren \mathbf{c}_k ja entweder Eigenvektoren zu λ_k oder Nullvektoren. Wegen (9.27), das heisst $\mathbf{c}_1 + \dots + \mathbf{c}_m = \mathbf{o}$, folgt aus Satz 9.11 aber, dass sie das letztere sein müssen. Aus der Definition von \mathbf{c}_k sieht man dann, dass alle Koeffizienten $\gamma_{k,j}$ null sein müssen, denn die $\mathbf{v}_{k,j}$ sind ja Basisvektoren des k -ten Eigenraumes. Wir schliessen, dass alle Vektoren $\mathbf{v}_{k,j}$ zusammen eine Eigenbasis bilden.

Gibt es umgekehrt eine Eigenbasis $\{\mathbf{b}_l\}$ von \mathbf{A} , so dass also $\mathbf{A}\mathbf{b}_l = \mathbf{b}_l\lambda_l$ ($l = 1, \dots, n$) gilt, so folgt: Zu einem festen Eigenwert λ_k und Koeffizienten

$$\alpha_l := \begin{cases} \text{beliebig} & \text{falls } \lambda_l = \lambda_k, \\ 0 & \text{falls } \lambda_l \neq \lambda_k, \end{cases} \quad (9.28)$$

gilt für $\mathbf{x} := \sum_{l=1}^n \mathbf{b}_l\alpha_l$, dass

$$\mathbf{A}\mathbf{x} = \mathbf{A} \left(\sum_{l=1}^n \mathbf{b}_l\alpha_l \right) = \sum_{l=1}^n \mathbf{b}_l\lambda_l\alpha_l = \lambda_k \sum_{l=1}^n \mathbf{b}_l\alpha_l = \lambda_k\mathbf{x}.$$

Also ist $\mathbf{x} \in E_{\lambda_k}$, und die Dimension dieses Eigenraums ist wegen (9.28) gleich der algebraischen Vielfachheit des Eigenwertes. ■

Bevor wir den Fall analysieren, wo die geometrische Vielfachheit kleiner als die arithmetische ist, betrachten wir eine wichtige Klasse von diagonalisierbaren Matrizen.

9.3 Eigenwerte und Eigenvektoren symmetrischer und Hermitescher Matrizen

Die Mehrzahl der Eigenwertprobleme, die in der Praxis auftreten sind **selbstadjungiert** [*self-adjoint*], das heisst die Matrizen sind reell symmetrisch oder Hermitesch. In diesem Falle ist die Eigenwertzerlegung einfach: Die Eigenwerte sind reell, und es gibt stets eine orthonormale Eigenbasis. Falls die Matrix selbst reell ist, ist auch die Eigenbasis reell.

Satz 9.15 *Ist $\mathbf{A} \in \mathbb{C}^{n \times n}$ Hermitesch, so gilt:*

- i) *Alle Eigenwerte $\lambda_1, \dots, \lambda_n$ sind reell.*
- ii) *Die Eigenvektoren zu verschiedenen Eigenwerten sind paarweise orthogonal in \mathbb{C}^n .*
- iii) *Es gibt eine orthonormale Basis des \mathbb{C}^n aus Eigenvektoren $\mathbf{u}_1, \dots, \mathbf{u}_n$ von \mathbf{A} .*
- iv) *Für die unitäre Matrix $\mathbf{U} := (\mathbf{u}_1 \ \dots \ \mathbf{u}_n)$ gilt*

$$\mathbf{U}^H \mathbf{A} \mathbf{U} = \mathbf{\Lambda} := \text{diag}(\lambda_1, \dots, \lambda_n). \quad (9.29)$$

Korollar 9.16 Ist $\mathbf{A} \in \mathbb{R}^{n \times n}$ symmetrisch, so gilt:

- i) Alle Eigenwerte $\lambda_1, \dots, \lambda_n$ sind reell.
- ii) Die reellen Eigenvektoren zu verschiedenen Eigenwerten sind paarweise orthogonal in \mathbb{R}^n .
- iii) Es gibt eine orthonormale Basis des \mathbb{R}^n aus Eigenvektoren $\mathbf{u}_1, \dots, \mathbf{u}_n$ von \mathbf{A} .
- iv) Für die orthogonale Matrix $\mathbf{U} := (\mathbf{u}_1 \ \dots \ \mathbf{u}_n)$ gilt

$$\mathbf{U}^T \mathbf{A} \mathbf{U} = \mathbf{\Lambda} := \text{diag}(\lambda_1, \dots, \lambda_n). \quad (9.30)$$

BEWEIS von Satz 9.15:

i) und ii): Es seien \mathbf{u} und \mathbf{v} Eigenvektoren zu den nicht notwendigerweise verschiedenen Eigenwerten λ_u und λ_v . Dann folgt aus $\mathbf{A}^H = \mathbf{A}$, dass

$$\langle \mathbf{u}, \mathbf{A} \mathbf{v} \rangle = \mathbf{u}^H \mathbf{A} \mathbf{v} = \mathbf{u}^H \mathbf{A}^H \mathbf{v} = (\mathbf{A} \mathbf{u})^H \mathbf{v} = \langle \mathbf{A} \mathbf{u}, \mathbf{v} \rangle$$

und damit

$$\lambda_v \langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, \lambda_v \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{A} \mathbf{v} \rangle = \langle \mathbf{A} \mathbf{u}, \mathbf{v} \rangle = \langle \lambda_u \mathbf{u}, \mathbf{v} \rangle = \overline{\lambda_u} \langle \mathbf{u}, \mathbf{v} \rangle. \quad (9.31)$$

Wenden wir dies an auf den Fall wo $\lambda_u = \lambda_v$ und $\mathbf{u} = \mathbf{v}$, so folgt wegen $\langle \mathbf{u}, \mathbf{u} \rangle > 0$, dass $\lambda_v = \overline{\lambda_u}$. Dies bedeutet, dass alle Eigenwerte reell sind.

Im Falle $\lambda_u \neq \lambda_v$ schreiben wir (9.31) als $(\lambda_v - \overline{\lambda_u}) \langle \mathbf{u}, \mathbf{v} \rangle = 0$. Wegen $\lambda_v - \overline{\lambda_u} = \lambda_v - \lambda_u \neq 0$ folgt dann, dass der zweite Faktor null sein muss. Also stehen in diesem Falle \mathbf{u} und \mathbf{v} senkrecht aufeinander.

iii) und iv): Diese zwei Aussagen sind offensichtlich äquivalent. Wir können natürlich in jedem Eigenraum E_λ , weil Unterraum von \mathbb{C}^n , eine orthonormale Basis wählen (vgl. Korollar 6.7). Da nach Teil ii) die Eigenvektoren zu verschiedenen Eigenwerten senkrecht aufeinander stehen, folgt, dass diese Basisvektoren zusammen eine orthogonale Basis des Unterraumes bilden, der von allen Eigenvektoren aufgespannt wird; vgl. den Beweis von Satz 9.14. Was uns allerdings fehlt, ist die Gewissheit, dass dieser Raum der ganze \mathbb{C}^n ist, d.h. dass für keinen Eigenwert die geometrische kleiner als die algebraische Vielfachheit ist. Oder, mit anderen Worten, dass \mathbf{A} diagonalisierbar ist. Dies lässt sich mit Induktion beweisen; siehe z.B. Nipp/Stoffer Seiten 159–161. ■

BEWEIS von Korollar 9.16:

i) Weil jede reelle symmetrische Matrix Hermitesch ist, gelten die Aussagen von Satz 9.15, insbesondere Teil i).

ii)–iv) Weil Matrix und Eigenwerte reell sind, kann man sich auf reelle Eigenvektoren beschränken, für die das Skalarprodukt in \mathbb{R}^n gleich jenem in \mathbb{C}^n ist. Aus der unitären Matrix mit den Kolonnen aus Eigenvektoren wird dabei eine orthogonale Matrix. ■

BEISPIEL 9.10: Die dem Vektor $\mathbf{w} = (1 \ 2 \ 2)^\top / 3$ zugeordnete Householder-Matrix

$$\mathbf{Q}_w := \mathbf{I} - 2\mathbf{w}\mathbf{w}^\top = \frac{1}{9} \begin{pmatrix} 7 & -4 & -4 \\ -4 & 1 & -8 \\ -4 & -8 & 1 \end{pmatrix} \quad (9.32)$$

ist symmetrisch und orthogonal. Da die Matrix einer Spiegelung an einer 2-dimensionalen Ebene entspricht, ist auf Grund von Satz 9.13 klar, dass die Eigenwerte 1, 1, -1 sind; der Eigenraum E_1 ist die Spiegelebene, der Eigenraum E_{-1} eine Gerade senkrecht dazu, erzeugt durch den Vektor $\mathbf{u}_3 := \mathbf{w}$.

Um eine orthonormale Basis der Spiegelebene E_1 zu finden, könnten wir hier einfach \mathbf{w} zu einer orthonormalen Basis von \mathbb{R}^3 ergänzen.

Oder wir können eine orthonormale Basis von E_1 konstruieren, ohne dass wir vom schon bekannten Eigenvektor \mathbf{u}_3 Gebrauch machen. Die Matrix

$$\mathbf{Q}_w - \mathbf{I} = \frac{1}{9} \begin{pmatrix} -2 & -4 & -4 \\ -4 & -8 & -8 \\ -4 & -8 & -8 \end{pmatrix}$$

hat offensichtlich Rang 1, ihre Zeilenstufenform liefert als einzige aus $(\mathbf{Q}_w - \mathbf{I})\mathbf{x} = \mathbf{o}$ resultierende nichttriviale Gleichung

$$\langle 3\mathbf{w}, \mathbf{x} \rangle = x_1 + 2x_2 + 2x_3 = 0,$$

was gerade die Bedingung ist, dass \mathbf{x} orthogonal auf \mathbf{w} steht.

Die Wahl $x_2 = 1, x_3 = 0$ ergibt $x_1 = -2$, also den normierten Eigenvektor $\mathbf{u}_1 = (-2 \ 1 \ 0)^\top / \sqrt{5}$. Die Wahl $x_2 = 0, x_3 = 1$ ergäbe $x_1 = -2$; dieser Vektor wäre zwar ein Eigenvektor, aber nicht senkrecht zu \mathbf{u}_1 . Um eine orthogonale Eigenbasis zu erhalten müssen wir die Bedingung $\langle \mathbf{u}_1, \mathbf{u}_2 \rangle = 0$ als weitere Gleichung beifügen:

$$x_1 + 2x_2 + 2x_3 = 0, \quad -2x_1 + x_2 = 0.$$

Hier sieht man, dass die Wahl $x_1 = 2$ auf $x_2 = 4$ und $x_3 = -5$ führt, also auf $\mathbf{u}_2 = (2 \ 4 \ -5)^\top / \sqrt{45}$.

Man verifiziert leicht, dass in der Tat mit $\mathbf{U} := (\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3)$ die Formel (9.30) bzw. die Spektralzerlegung $\mathbf{Q}_w = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\top$ gilt (auf vier Stellen nach dem Komma gerundet):

$$\frac{1}{9} \begin{pmatrix} 7 & -4 & -4 \\ -4 & 1 & -8 \\ -4 & -8 & 1 \end{pmatrix} = \begin{pmatrix} -0.8944 & 0.2981 & 0.3333 \\ 0.4472 & 0.5963 & 0.6667 \\ 0.0000 & -0.7454 & 0.6667 \end{pmatrix} \times \\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} -0.8944 & 0.4472 & 0.0000 \\ 0.2981 & 0.5963 & -0.7454 \\ 0.3333 & 0.6667 & 0.6667 \end{pmatrix}.$$

◆

9.4 Die Jordansche Normalform

Es stellt sich immer noch die Frage, wie weit sich eine quadratische Matrix im allgemeinen vereinfachen lässt. Ohne Beweis zitieren wir den folgenden Satz, der auch die Analyse der Situation liefert, wo die geometrische Vielfachheit gewisser Eigenwerte kleiner als die arithmetische Vielfachheit ist.

Satz 9.17 *Zu einer quadratischen Matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ gibt es eine ähnliche, blockdiagonale Matrix*

$$\mathbf{J} := \begin{pmatrix} \mathbf{J}_1 & \mathbf{O} & \cdots & \mathbf{O} \\ \mathbf{O} & \mathbf{J}_2 & \cdots & \mathbf{O} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{O} & \mathbf{O} & \cdots & \mathbf{J}_\mu \end{pmatrix} \quad (9.33)$$

mit μ Diagonalblöcken \mathbf{J}_k der Ordnung m_k der bidiagonalen Form

$$\mathbf{J}_k = \begin{pmatrix} \lambda_k & & & & \\ & \lambda_k & & & \\ & & \lambda_k & & \\ & & & \lambda_k & \\ & & & & \lambda_k \end{pmatrix} \quad \text{oder} \quad \mathbf{J}_k = \begin{pmatrix} \lambda_k & 1 & 0 & \cdots & 0 \\ 0 & \lambda_k & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \lambda_k & 1 \\ 0 & \cdots & \cdots & 0 & \lambda_k \end{pmatrix}. \quad (9.34)$$

Dabei sind die λ_k nicht unbedingt verschieden.

Die Matrix \mathbf{J} ist eindeutig bestimmt bis auf die Reihenfolge der Diagonalblöcke \mathbf{J}_k .

Ist \mathbf{A} eine reelle Matrix, die nur reelle Eigenwerte hat, so kann man auch die Ähnlichkeitstransformation reell wählen.

DEFINITION: Eine zu \mathbf{A} ähnliche Matrix der Form (9.33)–(9.34) heisst **Jordansche Normalform** [*Jordan canonical form*] von \mathbf{A} . Die Blöcke \mathbf{J}_k nennen wir **Jordan-Blöcke** [*Jordan blocks*]; jene mit $m_k > 1$ sind die **nicht-trivialen Jordan-Blöcke**. ▲

BEISPIEL 9.11: Ein Beispiel einer nicht-diagonalen Jordanschen Normalform ist

$$\mathbf{J} = \begin{pmatrix} 5 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (9.35)$$

Zur Hervorhebung der Blöcke haben wir die 10×10 -Matrix als Blockmatrix geschrieben. ◆

BEMERKUNGEN: 1) In der Zerlegung

$$\mathbf{A} = \mathbf{V}\mathbf{J}\mathbf{V}^{-1} \tag{9.36}$$

mit $\mathbf{V} \equiv: (\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_n)$ sind jene Vektoren \mathbf{v}_l , deren Index l mit dem der ersten Kolonne eines Blockes \mathbf{J}_k übereinstimmt, die Eigenvektoren, denn für sie gilt $\mathbf{A}\mathbf{v}_l = \mathbf{v}_l\lambda_k$.

Für die anderen Vektoren \mathbf{v}_l im k -ten Block gilt stattdessen

$$\mathbf{A}\mathbf{v}_l = \mathbf{v}_l\lambda_k + \mathbf{v}_{l-1}, \quad \text{d.h.} \quad (\mathbf{A} - \lambda_k\mathbf{I})\mathbf{v}_l = \mathbf{v}_{l-1}. \tag{9.37}$$

Diese Vektoren nennt man **Hauptvektoren** [*principal vectors*] zum Eigenwert λ_k .

Satz 9.17 besagt also auch, dass es zu jeder $n \times n$ -Matrix \mathbf{A} eine Basis von \mathbb{C}^n aus Eigen- und Hauptvektoren von \mathbf{A} gibt.

2) Die Zahl μ der Blöcke ist gleich der Anzahl linear unabhängiger Eigenvektoren. \mathbf{A} ist diagonalisierbar genau dann, wenn $\mu = n$. Für einen bestimmten Eigenwert λ ist die algebraische Vielfachheit gleich

$$\sum_{\lambda_k=\lambda} m_k \tag{9.38}$$

und die geometrische Vielfachheit gleich $\sum_{\lambda_k=\lambda} 1$, das heisst gleich der Anzahl Blöcke mit $\lambda_k = \lambda$.

3) Die Jordansche Normalform ist in der Matrizentheorie ein wichtiges Hilfsmittel. Sie ist aber für praktische numerische Berechnungen unbrauchbar, weil sie nicht stetig von der Matrix (d.h. den Elementen der Matrix) abhängt. Durch eine beliebig kleine Änderung einer Matrix mit Jordan-Blöcken kann man erreichen, dass alle Eigenwerte der Matrix verschieden werden, was bedeutet dass alle diese Jordan-Blöcke verschwinden. Natürlich ändern dabei auch die Matrizen \mathbf{V} aus (9.36) unstetig. ▼

BEISPIEL 9.12: Die Matrix \mathbf{J} aus (9.35) hat die vier verschiedenen Eigenwerte 5, 3, 1 und 0, von denen alle ausser 0 mehrfach sind. Die algebraischen und geometrischen Vielfachheiten der Eigenwerte sind in der nachfolgenden Tabelle zusammengefasst:

EW λ	alg. Vfh.	geom. Vfh.	Blockgrössen m_k
5	5	3	3, 1, 1
3	2	1	2
1	1	1	1
0	2	1	2



Satz 9.17 liefert auch sofort die folgende Verallgemeinerung von Korollar 9.10.

Korollar 9.18 Zerlegen wir die Matrizen \mathbf{V} und \mathbf{V}^{-1} aus (9.36) entsprechend den Blöcken von \mathbf{J} in vertikale bzw. horizontale Streifen,

$$\mathbf{V} = \begin{pmatrix} \mathbf{V}_1 & \cdots & \mathbf{V}_\mu \end{pmatrix}, \quad \mathbf{V}^{-1} = \begin{pmatrix} \mathbf{W}_1^\top \\ \vdots \\ \mathbf{W}_\mu^\top \end{pmatrix} \quad (9.39)$$

(so dass \mathbf{V}_k ein $n \times m_k$ -Block und \mathbf{W}_k ein $n \times m_k$ -Block ist), so lässt sich \mathbf{A} wie folgt als Summe von μ Matrizen des Ranges m_k ($k = 1, \dots, \mu$) darstellen:

$$\mathbf{A} = \sum_{k=1}^{\mu} \mathbf{V}_k \mathbf{J}_k \mathbf{W}_k^\top. \quad (9.40)$$

BEWEIS: Man schreibt \mathbf{J} als Summe von μ block-diagonalen $n \times n$ Matrizen, wobei in der k -ten nur gerade der Block \mathbf{J}_k steht, während die anderen durch Nullblöcke ersetzt sind. Setzt man diese Summen und die Aufteilungen aus (9.39) in $\mathbf{A} = \mathbf{V}\mathbf{J}\mathbf{V}^{-1}$ ein, folgt die Behauptung sofort. ■

Während (9.25) in Korollar 9.10 eine additive Zerlegung einer diagonalisierbaren Matrix \mathbf{A} in eine Summe von n Rang-1-Matrizen war, ist die Eigenwertzerlegung im allgemeinen Fall gemäss (9.40) eine additive Zerlegung in eine Summe von μ Termen, die in der Regel verschiedene Ränge m_k haben.

BEISPIEL 9.13: Die der Matrix \mathbf{J} aus (9.35) entsprechende Zerlegung (9.40) einer Matrix $\mathbf{A} = \mathbf{V}\mathbf{J}\mathbf{V}^{-1}$ lautet in der Notation von Korollar 9.18:

$$\begin{aligned} \mathbf{A} = & \mathbf{V}_1 \begin{pmatrix} 5 & 1 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 5 \end{pmatrix} \mathbf{W}_1^\top + \mathbf{V}_2 \begin{pmatrix} 5 \end{pmatrix} \mathbf{W}_2^\top + \mathbf{V}_3 \begin{pmatrix} 5 \end{pmatrix} \mathbf{W}_3^\top \\ & + \mathbf{V}_4 \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix} \mathbf{W}_4^\top + \mathbf{V}_5 \begin{pmatrix} 1 \end{pmatrix} \mathbf{W}_5^\top + \mathbf{V}_6 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \mathbf{W}_6^\top, \end{aligned}$$

wobei

$$\begin{aligned} \mathbf{V}_1 &= \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{pmatrix}, & \mathbf{V}_2 &= \begin{pmatrix} \mathbf{v}_4 \end{pmatrix}, & \mathbf{V}_3 &= \begin{pmatrix} \mathbf{v}_5 \end{pmatrix}, \\ \mathbf{V}_4 &= \begin{pmatrix} \mathbf{v}_6 & \mathbf{v}_7 \end{pmatrix}, & \mathbf{V}_5 &= \begin{pmatrix} \mathbf{v}_8 \end{pmatrix}, & \mathbf{V}_6 &= \begin{pmatrix} \mathbf{v}_9 & \mathbf{v}_{10} \end{pmatrix}, \\ \mathbf{W}_1^\top &= \begin{pmatrix} \mathbf{w}_1^\top \\ \mathbf{w}_2^\top \\ \mathbf{w}_3^\top \end{pmatrix}, & \mathbf{W}_2^\top &= \begin{pmatrix} \mathbf{w}_4^\top \end{pmatrix}, & \mathbf{W}_3^\top &= \begin{pmatrix} \mathbf{w}_5^\top \end{pmatrix}, \\ \mathbf{W}_4^\top &= \begin{pmatrix} \mathbf{w}_6^\top \\ \mathbf{w}_7^\top \end{pmatrix}, & \mathbf{W}_5^\top &= \begin{pmatrix} \mathbf{w}_8^\top \end{pmatrix}, & \mathbf{W}_6^\top &= \begin{pmatrix} \mathbf{w}_9^\top \\ \mathbf{w}_{10}^\top \end{pmatrix}. \end{aligned}$$

◆

Kapitel 10

Anwendungen der Eigenwertzerlegung

Wir betrachten in diesem Kapitel eine Reihe von Anwendungen der Eigenwertzerlegung. Als erstes behandeln wir verschiedene Typen von Systemen von linearen Differentialgleichungen mit konstanten Koeffizienten. Die gefundene Theorie gibt Anlass zur Definition der Exponentialfunktion und weiterer Funktionen einer Matrix. Dann betrachten wir die Hauptachsentransformation von quadratischen Formen und ihre Anwendung auf die Charakterisierung von Kegelschnitten und Flächen zweiter Ordnung (Ellipsoiden usw.). Schliesslich zeigen wir, dass die Spektralnorm einer Matrix \mathbf{A} durch den grössten Eigenwert von $\mathbf{A}^H\mathbf{A}$ ausgedrückt werden kann.

10.1 Homogene lineare Differentialgleichungen mit konstanten Koeffizienten

Systeme erster Ordnung

Gegeben sei ein **System von homogenen linearen Differentialgleichungen erster Ordnung**

$$\begin{aligned} \dot{y}_1(t) &= a_{11}y_1(t) + a_{12}y_2(t) + \cdots + a_{1n}y_n(t) \\ &\vdots \\ \dot{y}_n(t) &= a_{n1}y_1(t) + a_{n2}y_2(t) + \cdots + a_{nn}y_n(t) \end{aligned} \quad (10.1)$$

mit **konstanten (zeitunabhängigen) Koeffizienten** a_{kl} für die gesuchten Funktionen y_1, \dots, y_n . Der Punkt auf $\dot{y}_k(t)$ bezeichnet die Ableitung nach der Zeit t . Das System ist homogen, weil rechts nur Terme vorkommen, die in einem $y_k(t)$ linear sind und folglich die Nullfunktionen $y_1(t) \equiv \cdots \equiv y_n(t) \equiv 0$ eine triviale Lösung bilden. Eine spezifische Lösung ist erst eindeutig bestimmt durch n zusätzliche Bedingungen, z.B. die Anfangsbedingungen

$$y_1(0) = \eta_1, \quad \dots, \quad y_n(0) = \eta_n, \quad (10.2)$$

worin η_1, \dots, η_n vorgegebene Zahlen sind. Definiert man

$$\mathbf{A} := (a_{kl})_{k,l=1}^n, \quad \mathbf{y}(t) := \begin{pmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{pmatrix}, \quad \mathbf{y}_0 := \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix}, \quad (10.3)$$

so kann man das System (10.1) mit den Anfangsbedingungen (10.2) wie folgt in Matrixform schreiben:

$$\boxed{\dot{\mathbf{y}}(t) = \mathbf{A}\mathbf{y}(t), \quad \mathbf{y}(0) = \mathbf{y}_0.} \quad (10.4)$$

Ist \mathbf{A} diagonalisierbar, $\mathbf{A}\mathbf{V} = \mathbf{V}\mathbf{\Lambda}$, und definieren wir $\mathbf{z}(t) := \mathbf{V}^{-1}\mathbf{y}(t)$, so dass $\dot{\mathbf{z}}(t) = \mathbf{V}^{-1}\dot{\mathbf{y}}(t)$ (weil \mathbf{V} zeitunabhängige Elemente hat), ist (10.4) äquivalent zu

$$\underbrace{\mathbf{V}^{-1}\dot{\mathbf{y}}(t)}_{\dot{\mathbf{z}}(t)} = \underbrace{\mathbf{V}^{-1}\mathbf{A}\mathbf{V}}_{\mathbf{\Lambda}} \underbrace{\mathbf{V}^{-1}\mathbf{y}(t)}_{\mathbf{z}(t)}.$$

Das heisst, es gilt

$$\boxed{\dot{\mathbf{z}}(t) = \mathbf{\Lambda}\mathbf{z}(t)} \quad (10.5)$$

oder

$$\begin{aligned} \dot{z}_1(t) &= \lambda_1 z_1(t) \\ &\vdots \\ \dot{z}_n(t) &= \lambda_n z_n(t). \end{aligned} \quad (10.6)$$

Durch Bestimmen der Eigenwerte und Eigenvektoren von \mathbf{A} ist es uns also gelungen, das System (10.1) gekoppelter linearer Differentialgleichungen zu *entkoppeln*. Man nennt dieses Vorgehen die **Transformationsmethode**.

Das entkoppelte System (10.5)/(10.6) hat die *allgemeine Lösung*

$$\boxed{\begin{aligned} z_1(t) &= \gamma_1 e^{\lambda_1 t} \\ &\vdots \\ z_n(t) &= \gamma_n e^{\lambda_n t} \end{aligned}} \quad (10.7)$$

mit frei wählbaren Konstanten $\gamma_1, \dots, \gamma_n$. Um eine kompakte Darstellung zu erhalten, definieren wir die Diagonalmatrix

$$\boxed{e^{t\mathbf{\Lambda}} := \text{diag}(e^{\lambda_1 t}, \dots, e^{\lambda_n t})} \quad (10.8)$$

und den Konstantenvektor $\mathbf{c} = (\gamma_1 \ \dots \ \gamma_n)^\top$, womit

$$\mathbf{z}(t) = \begin{pmatrix} \gamma_1 e^{\lambda_1 t} \\ \vdots \\ \gamma_n e^{\lambda_n t} \end{pmatrix} = e^{t\mathbf{\Lambda}} \mathbf{c}. \quad (10.9)$$

Die allgemeine Lösung von (10.1) lautet damit

$$\boxed{\mathbf{y}(t) = \mathbf{V}\mathbf{z}(t) = \mathbf{V}e^{t\mathbf{\Lambda}}\mathbf{c}.} \quad (10.10)$$

Um die spezielle Lösung, die (10.2) erfüllt, zu finden, benutzen wir, dass $\mathbf{z}(0) = (\gamma_1 \ \dots \ \gamma_n)^\top = \mathbf{c}$. Aus (10.10) und $\mathbf{y}(0) = \mathbf{y}_0$ folgt, dass das lineare Gleichungssystem

$$\boxed{\mathbf{V}\mathbf{c} = \mathbf{y}_0} \quad \text{oder} \quad \boxed{\mathbf{V} \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix} = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix}} \quad (10.11)$$

nach $\mathbf{c} = (\gamma_1 \ \dots \ \gamma_n)^\top$ aufzulösen ist. Danach ist das berechnete \mathbf{c} in (10.10) einzusetzen.

BEISPIEL 10.1: Gegeben sei das Differentialgleichungssystem 1. Ordnung

$$\begin{aligned} \dot{y}_1(t) &= -2y_1(t) && + 2y_3(t) \\ \dot{y}_2(t) &= -2y_1(t) - 3y_2(t) - 4y_3(t) \\ \dot{y}_3(t) &= && - 3y_3(t) \end{aligned} \quad (10.12)$$

und die Anfangsbedingungen

$$y_1(0) = 0, \quad y_2(0) = 0, \quad y_3(0) = 1. \quad (10.13)$$

Die (10.12) entsprechende Matrix \mathbf{A} ist

$$\mathbf{A} = \begin{pmatrix} -2 & 0 & 2 \\ -2 & -3 & -4 \\ 0 & 0 & -3 \end{pmatrix}. \quad (10.14)$$

Sie hat die Eigenwertzerlegung

$$\mathbf{A} = \underbrace{\begin{pmatrix} 1 & -2 & 0 \\ -2 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}}_{\mathbf{V}} \underbrace{\begin{pmatrix} -2 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{pmatrix}}_{\mathbf{\Lambda}} \underbrace{\begin{pmatrix} 1 & 0 & 2 \\ 0 & 0 & 1 \\ -2 & -1 & -4 \end{pmatrix}}_{\mathbf{V}^{-1}}. \quad (10.15)$$

Es würde allerdings genügen, \mathbf{V} zu kennen, \mathbf{V}^{-1} bräuchte man nicht. Weil hier der Vektor \mathbf{y}_0 der Anfangswerte gerade gleich \mathbf{e}_3 ist, ist die Lösung \mathbf{c} von $\mathbf{V}\mathbf{c} = \mathbf{y}_0$ aber gerade gleich der dritten Kolonne von \mathbf{V}^{-1} :

$$\mathbf{c} = \mathbf{V}^{-1}\mathbf{y}_0 = \mathbf{V}^{-1}\mathbf{e}_3 = \begin{pmatrix} 2 & 1 & -4 \end{pmatrix}^T. \quad (10.16)$$

Nach (10.10) erhält man damit als Lösung des Anfangswertproblems (10.12)–(10.13)

$$\begin{aligned} \mathbf{y}(t) &= \mathbf{V} e^{t\mathbf{\Lambda}} \mathbf{c} \\ &= \begin{pmatrix} 1 & -2 & 0 \\ -2 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} e^{-2t} & 0 & 0 \\ 0 & e^{-3t} & 0 \\ 0 & 0 & e^{-3t} \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ -4 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -2 & 0 \\ -2 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2e^{-2t} \\ e^{-3t} \\ -4e^{-3t} \end{pmatrix} \\ &= \begin{pmatrix} 2e^{-2t} - 2e^{-3t} \\ -4e^{-2t} + 4e^{-3t} \\ e^{-3t} \end{pmatrix}. \end{aligned}$$



Wir betonen, dass die Transformationsmethode in der hier beschriebenen Form voraussetzt, dass \mathbf{A} diagonalisierbar ist. Die Methode ist erweiterbar auf nicht-diagonalisierbare Matrizen, aber dann komplizierter. Wir nehmen auch im Rest dieses Abschnittes an, dass \mathbf{A} diagonalisierbar ist. Dagegen führen mehrfache Eigenwerte mit voller geometrischer Vielfachheit auf keine Schwierigkeiten, wie obiges Beispiel illustriert.

Systeme höherer Ordnung

Neben Systemen von Differentialgleichung erster Ordnung mit konstanten Koeffizienten kann man nach demselben Rezept auch **lineare Differentialgleichungen höherer Ordnung mit konstanten Koeffizienten** lösen oder sogar Systeme von solchen. Bekanntlich lassen sich allgemein Differentialgleichungen höherer Ordnung auf Systeme von Differentialgleichung erster Ordnung reduzieren.

Ist die Differentialgleichung n -ter Ordnung

$$x^{(n)}(t) - \beta_n x^{(n-1)}(t) - \dots - \beta_2 \dot{x}(t) - \beta_1 x(t) = 0 \quad (10.17)$$

mit den Anfangsbedingungen

$$x(0) = \eta_1, \quad \dot{x}(0) = \eta_2, \quad \dots, \quad x^{(n-1)}(0) = \eta_n \quad (10.18)$$

gegeben, so setzen wir

$$y_1(t) := x(t), \quad y_2(t) := \dot{x}(t), \quad \dots, \quad y_n(t) := x^{(n-1)}(t), \quad (10.19)$$

womit (10.17)–(10.18) geschrieben werden kann in der Form (10.4), $\dot{\mathbf{y}}(t) = \mathbf{A}\mathbf{y}(t)$, $\mathbf{y}(0) = \mathbf{y}_0$, mit \mathbf{y} und \mathbf{y}_0 wie in (10.3) und

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 1 \\ \beta_1 & \beta_2 & \dots & \beta_{n-1} & \beta_n \end{pmatrix}. \quad (10.20)$$

BEISPIEL 10.2: Gegeben sei das Anfangswertproblem 2. Ordnung

$$\ddot{x}(t) - \dot{x}(t) - 2x(t) = 0, \quad x(0) = 3, \quad \dot{x}(0) = 0. \quad (10.21)$$

Wir setzen $y_1(t) := x(t)$, $y_2(t) := \dot{x}(t)$ und erhalten so das Anfangswertproblem 1. Ordnung

$$\begin{aligned} \dot{y}_1 &= y_2 & \text{mit} & \quad \eta_1 = y_1(0) = 3, \\ \dot{y}_2 &= 2y_1 + y_2 & & \quad \eta_2 = y_2(0) = 0. \end{aligned}$$

Hier ist

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix}, \quad \mathbf{y}_0 = \begin{pmatrix} 3 \\ 0 \end{pmatrix},$$

$$\chi_{\mathbf{A}}(\lambda) = \begin{vmatrix} -\lambda & 1 \\ 2 & 1 - \lambda \end{vmatrix} = \lambda^2 - \lambda - 2 = (\lambda + 1)(\lambda - 2).$$

Die Eigenwerte sind also $\lambda_1 = -1$, $\lambda_2 = 2$.

Das Auflösen von $(\mathbf{A} - \lambda_k \mathbf{I})\mathbf{v} = \mathbf{0}$ ($k = 1, 2$) gibt nach kurzer Rechnung die Eigenvektoren $\mathbf{v}_1 = (1 \ -1)^\top$ und $\mathbf{v}_2 = (1 \ 2)^\top$, also die Matrix

$$\mathbf{V} = \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix}.$$

Nun muss man die Anfangsbedingungen transformieren durch Lösen von (10.11), was $\gamma_1 = 2, \gamma_2 = 1$, das heisst $\mathbf{c} = \begin{pmatrix} 2 & 1 \end{pmatrix}^\top$ ergibt. Damit wird

$$\begin{aligned} z_1(t) &= 2e^{-t} & \text{und} & & x(t) &= y_1(t) &= 2e^{-t} + e^{2t}, \\ z_2(t) &= e^{2t} & & & \dot{x}(t) &= y_2(t) &= -2e^{-t} + 2e^{2t}. \end{aligned}$$

Die gesuchte Lösung ist also $x(t) = 2e^{-t} + e^{2t}$. \blacklozenge

Systeme zweiter Ordnung ohne Terme erster Ordnung

Nun betrachten wir noch **Systeme von homogenen linearen Differentialgleichungen zweiter Ordnung mit konstanten Koeffizienten und ohne erste Ableitungen**. Sie haben die Form

$$\begin{aligned} \ddot{y}_1(t) &= a_{11}y_1(t) + a_{12}y_2(t) + \cdots + a_{1n}y_n(t) \\ &\vdots \\ \ddot{y}_n(t) &= a_{n1}y_1(t) + a_{n2}y_2(t) + \cdots + a_{nn}y_n(t) \end{aligned} \quad (10.22)$$

Ein solches System kann man direkt mit der Transformationsmethode behandeln, ohne es auf ein System erster Ordnung zu reduzieren. Wir schreiben es zusammen mit den Anfangsbedingungen

$$\begin{aligned} y_1(0) &= \eta_1, & \dot{y}_1(0) &= \eta'_1, \\ &\vdots & &\vdots \\ y_n(0) &= \eta_n, & \dot{y}_n(0) &= \eta'_n, \end{aligned} \quad (10.23)$$

als

$$\boxed{\ddot{\mathbf{y}}(t) = \mathbf{A}\mathbf{y}(t), \quad \mathbf{y}(0) = \mathbf{y}_0, \quad \dot{\mathbf{y}}(0) = \mathbf{y}_1}. \quad (10.24)$$

Wir setzen wieder voraus, dass \mathbf{A} diagonalisierbar ist. Zudem seien die Eigenwerte von \mathbf{A} negativ. Dies ist typischerweise in Anwendungen, wo solche Differentialgleichungen auftreten, der Fall. Mit $\mathbf{A}\mathbf{V} = \mathbf{V}\mathbf{\Lambda}$ und $\mathbf{z}(t) := \mathbf{V}^{-1}\mathbf{y}(t)$ erhalten wir analog zu (10.4) das *entkoppelte System*

$$\boxed{\ddot{\mathbf{z}}(t) = \mathbf{\Lambda}\mathbf{z}(t)}. \quad (10.25)$$

Schreiben wir dies komponentenweise, und führen wir die Größen

$$\boxed{\omega_k := \sqrt{-\lambda_k}} \quad (10.26)$$

ein (hier wird $\lambda_k < 0$ gebraucht), so ergibt sich

$$\boxed{\ddot{z}_k(t) + \omega_k^2 z_k(t) = 0} \quad (k = 1, \dots, n). \quad (10.27)$$

Das ist die Differentialgleichung des **harmonischen Oszillators** [*harmonic oscillator*]. Sie hat bekanntlich die allgemeine Lösung

$$\boxed{z_k(t) = \alpha_k \cos(\omega_k t) + \beta_k \sin(\omega_k t)} \quad (k = 1, \dots, n) \quad (10.28)$$

mit freien Parametern α_k und β_k ($k = 1, \dots, n$). Die allgemeine Lösung von (10.22) ist damit

$$\mathbf{y}(t) = \mathbf{V} \mathbf{z}(t) = \mathbf{V} \begin{pmatrix} \alpha_1 \cos(\omega_1 t) + \beta_1 \sin(\omega_1 t) \\ \vdots \\ \alpha_n \cos(\omega_n t) + \beta_n \sin(\omega_n t) \end{pmatrix}. \quad (10.29)$$

Durch Einführung der Diagonalmatrizen

$$\begin{aligned} \boldsymbol{\Omega} &::= \text{diag}(\omega_1, \dots, \omega_n), \\ \cos(\boldsymbol{\Omega} t) &::= \text{diag}(\cos(\omega_1 t), \dots, \cos(\omega_n t)), \\ \sin(\boldsymbol{\Omega} t) &::= \text{diag}(\sin(\omega_1 t), \dots, \sin(\omega_n t)) \end{aligned} \quad (10.30)$$

und der Parametervektoren

$$\mathbf{a} ::= (\alpha_1 \ \dots \ \alpha_n)^\top, \quad \mathbf{b} ::= (\beta_1 \ \dots \ \beta_n)^\top \quad (10.31)$$

könnte man (10.29) auch schreiben als

$$\mathbf{y}(t) = \mathbf{V} \mathbf{z}(t) = \mathbf{V} (\cos(\boldsymbol{\Omega} t) \mathbf{a} + \sin(\boldsymbol{\Omega} t) \mathbf{b}). \quad (10.32)$$

Um die spezielle Lösung, die (10.23) erfüllt, zu finden, bemerken wir zunächst, dass gilt

$$z_k(0) = \alpha_k, \quad \dot{z}_k(0) = \beta_k \omega_k \quad (k = 1, \dots, n)$$

oder eben

$$\mathbf{z}(0) = \mathbf{a}, \quad \dot{\mathbf{z}}(0) = \boldsymbol{\Omega} \mathbf{b}. \quad (10.33)$$

Wegen $\mathbf{y}(t) = \mathbf{V} \mathbf{z}(t)$ resultieren hier aus den Anfangsbedingungen in (10.24) zwei Gleichungssysteme mit der Matrix \mathbf{V} , die zu lösen sind, um die Parameter α_k und β_k ($k = 1, \dots, n$) festzulegen:

$$\mathbf{V} \mathbf{a} = \mathbf{y}_0, \quad \mathbf{V} \tilde{\mathbf{b}} = \mathbf{y}_1 \quad \text{mit} \quad \tilde{\mathbf{b}} ::= \boldsymbol{\Omega} \mathbf{b}. \quad (10.34)$$

Da $\boldsymbol{\Omega}$ diagonal ist, bedeutet $\mathbf{b} = \boldsymbol{\Omega}^{-1} \tilde{\mathbf{b}}$ bloss eine komponentenweise Division durch die ω_k .

BEISPIEL 10.3: Der Einfachheit halber und um die nötigen Änderungen zu betonen, ersetzen wir einfach in Beispiel 10.1 die erste durch die zweite Ableitung:

$$\begin{aligned} \ddot{y}_1(t) &= -2y_1(t) && + 2y_3(t) \\ \ddot{y}_2(t) &= -2y_1(t) - 3y_2(t) - 4y_3(t) \\ \ddot{y}_3(t) &= && - 3y_3(t) \end{aligned} \quad (10.35)$$

Die Anfangsbedingungen seien nun

$$\begin{aligned} y_1(0) &= 0, & y_2(0) &= 0, & y_3(0) &= 1, \\ \dot{y}_1(0) &= 1, & \dot{y}_2(0) &= 3, & \dot{y}_3(0) &= 0, \end{aligned} \quad (10.36)$$

Die für (10.35) relevante Matrix \mathbf{A} ist also wieder (10.14) mit der Eigenwertzerlegung (10.15). Aus den Eigenwerten $\lambda_1 = -2$, $\lambda_2 = -3$, $\lambda_3 = -3$, bekommen wir

$$\boldsymbol{\Omega} := \text{diag}(\omega_1, \omega_2, \omega_n) = \text{diag}(\sqrt{2}, \sqrt{3}, \sqrt{3}).$$

Wie in (10.16) gilt

$$\mathbf{a} = \mathbf{V}^{-1} \mathbf{y}_0 = \mathbf{V}^{-1} \mathbf{e}_3 = \begin{pmatrix} 2 & 1 & -4 \end{pmatrix}^T$$

und zudem wegen $\mathbf{y}_1 = \mathbf{e}_1 + 3\mathbf{e}_2$

$$\tilde{\mathbf{b}} = \mathbf{V}^{-1} \mathbf{y}_1 = \mathbf{V}^{-1} (\mathbf{e}_1 + 3\mathbf{e}_2) = \begin{pmatrix} 1 & 0 & -5 \end{pmatrix}^T,$$

also

$$\mathbf{b} = \boldsymbol{\Omega}^{-1} \tilde{\mathbf{b}} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{-5}{\sqrt{3}} \end{pmatrix}^T.$$

In der Darstellung (10.29) lautet die gesuchte Lösung damit

$$\mathbf{y}(t) = \underbrace{\begin{pmatrix} 1 & -2 & 0 \\ -2 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}}_{\mathbf{V}} \underbrace{\begin{pmatrix} 2 \cos(\sqrt{2}t) + \frac{1}{\sqrt{2}} \sin(\sqrt{2}t) \\ \cos(\sqrt{3}t) \\ -4 \cos(\sqrt{3}t) - \frac{5}{\sqrt{3}} \sin(\sqrt{3}t) \end{pmatrix}}_{\cos(\boldsymbol{\Omega}t) \mathbf{a} + \sin(\boldsymbol{\Omega}t) \mathbf{b}}.$$

Das Matrix-Vektor-Produkt in dieser Formel könnte man noch ausmultiplizieren, um explizite Ausdrücke für y_1 , y_2 , y_3 zu bekommen. \blacklozenge

10.2 Funktionen von Matrizen

Wir gehen nochmals zurück zu (10.7)–(10.10), wo wir bereits in (10.8) für die Diagonalmatrix $t\mathbf{\Lambda}$ die Exponentialfunktion definiert haben durch

$$e^{t\mathbf{\Lambda}} := \text{diag}(e^{t\lambda_1}, \dots, e^{t\lambda_n}),$$

womit sich (10.7) schreiben liess als

$$\mathbf{z}(t) = e^{t\mathbf{\Lambda}} \mathbf{c}. \quad (10.37)$$

Weil $t\mathbf{\Lambda}$ diagonal ist und weil die Potenzreihe von e^x für alle x konvergiert, ist klar, dass

$$e^{t\mathbf{\Lambda}} = \mathbf{I} + t\mathbf{\Lambda} + \frac{1}{2!} t^2 \mathbf{\Lambda}^2 + \frac{1}{3!} t^3 \mathbf{\Lambda}^3 + \dots.$$

Da $\mathbf{V} (t\mathbf{\Lambda})^k \mathbf{V}^{-1} = t^k (\mathbf{V} \mathbf{\Lambda} \mathbf{V}^{-1})^k = t^k \mathbf{A}^k$ (für alle $k \in \mathbb{N}$), macht es Sinn, für eine diagonalisierbare Matrix \mathbf{A} zu definieren

$$\boxed{e^{t\mathbf{A}} := \mathbf{V} e^{t\mathbf{\Lambda}} \mathbf{V}^{-1}}, \quad (10.38)$$

weil man dann zeigen kann, dass

$$\boxed{e^{t\mathbf{A}} = \mathbf{I} + t\mathbf{A} + \frac{1}{2!} t^2 \mathbf{A}^2 + \frac{1}{3!} t^3 \mathbf{A}^3 + \dots}. \quad (10.39)$$

Damit wird aus (10.10) zunächst

$$\mathbf{y}(t) = \mathbf{V}\mathbf{z}(t) = \mathbf{V}e^{t\mathbf{\Lambda}}\mathbf{c} = \mathbf{V}e^{t\mathbf{\Lambda}}\mathbf{V}^{-1}\mathbf{V}\mathbf{c}$$

und dann nach (10.38) und (10.11)

$$\boxed{\mathbf{y}(t) = e^{t\mathbf{A}}\mathbf{y}_0}. \quad (10.40)$$

Leitet man dies formal ab (d.h. wie wenn \mathbf{A} , \mathbf{y}_0 und \mathbf{y} bloss reelle Zahlen bzw. eine reelle Funktion wären), so ergibt sich gerade

$$\dot{\mathbf{y}}(t) = \mathbf{A}e^{t\mathbf{A}}\mathbf{y}_0 = \mathbf{A}\mathbf{y}(t),$$

also das System (10.4), von dem wir ausgegangen sind.

Analog zu (10.38) kann man auch andere *Funktionen von Matrizen* definieren: sofern die Matrix $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$ diagonalisierbar ist und die Funktion f in den Eigenwerten λ_k der Matrix definiert ist, setzt man

$$\boxed{f(\mathbf{\Lambda}) := \text{diag}(f(\lambda_1), \dots, f(\lambda_n))}. \quad (10.41)$$

und

$$\boxed{f(\mathbf{A}) := \mathbf{V}f(\mathbf{\Lambda})\mathbf{V}^{-1}}. \quad (10.42)$$

10.3 Reelle lineare Differentialgleichungen mit komplexen Eigenwerten

Im Prinzip funktioniert unser Vorgehen für Systeme von linearen Differentialgleichungen erster Ordnung auch, wenn die Matrix \mathbf{A} *komplexe Eigenwerte* hat.

Aber es ist natürlich unschön, wenn \mathbf{A} reell ist und man die Lösung einer reellen Differentialgleichungen mit Hilfe nicht-reeller komplexer Zahlen darstellt. Das ist auch nicht nötig. Wir skizzieren hier das Vorgehen nur, und zwar auf eine neue Art. Detaillierte Beispiele zum “klassischen” Lösungsweg gibt es in Nipp/Stoffer, S. 181ff. Wir fangen an mit einem allgemeinen Resultat über reelle Matrizen.

Die reelle Matrix \mathbf{A} habe den nicht-reellen Eigenwert λ und den zugehörigen (ebenfalls nicht-reellen) Eigenvektor \mathbf{v} . Durch Konjugieren von $\mathbf{A}\mathbf{v} = \mathbf{v}\lambda$ folgt dann $\mathbf{A}\bar{\mathbf{v}} = \bar{\mathbf{v}}\bar{\lambda}$, also gilt zusammengefasst

$$\mathbf{A} \begin{pmatrix} \mathbf{v} & \bar{\mathbf{v}} \end{pmatrix} = \begin{pmatrix} \mathbf{v} & \bar{\mathbf{v}} \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix} \quad (10.43)$$

Nun kann man Real- und Imaginärteil von \mathbf{v} ausdrücken durch

$$\boxed{\begin{pmatrix} \text{Re } \mathbf{v} & \text{Im } \mathbf{v} \end{pmatrix}} = \begin{pmatrix} \mathbf{v} & \bar{\mathbf{v}} \end{pmatrix} \mathbf{K}_2, \quad \text{wo} \quad \boxed{\mathbf{K}_2 := \frac{1}{2} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}}. \quad (10.44)$$

Unter Verwendung der Inversen der 2×2 -Matrix \mathbf{K}_2 ,

$$\mathbf{K}_2^{-1} = \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} = 2\mathbf{K}_2^H, \quad (10.45)$$

können wir schliesslich (10.43) ersetzen durch

$$\begin{aligned} \mathbf{A} \begin{pmatrix} \operatorname{Re} \mathbf{v} & \operatorname{Im} \mathbf{v} \end{pmatrix} &= \mathbf{A} \begin{pmatrix} \mathbf{v} & \bar{\mathbf{v}} \end{pmatrix} \mathbf{K}_2 \\ &= \begin{pmatrix} \mathbf{v} & \bar{\mathbf{v}} \end{pmatrix} \mathbf{K}_2 2\mathbf{K}_2^H \begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix} \mathbf{K}_2 \\ &= \underbrace{\frac{1}{2} \begin{pmatrix} \mathbf{v} & \bar{\mathbf{v}} \end{pmatrix} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}}_{= \begin{pmatrix} \operatorname{Re} \mathbf{v} & \operatorname{Im} \mathbf{v} \end{pmatrix}} \underbrace{\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix} \frac{1}{2} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}}_{= \begin{pmatrix} \operatorname{Re} \lambda & \operatorname{Im} \lambda \\ -\operatorname{Im} \lambda & \operatorname{Re} \lambda \end{pmatrix}}. \end{aligned}$$

Das liefert den folgenden Satz:

Satz 10.1 *Die reelle Matrix \mathbf{A} habe den nicht-reellen Eigenwert λ und den zugehörigen Eigenvektor \mathbf{v} . Dann ist $\bar{\mathbf{v}}$ ein Eigenvektor zu $\bar{\lambda}$ und es gilt (10.44) (wobei $\mathbf{K}_2^{-1} = 2\mathbf{K}_2^H$ ist), und*

$$\mathbf{A} \begin{pmatrix} \operatorname{Re} \mathbf{v} & \operatorname{Im} \mathbf{v} \end{pmatrix} = \begin{pmatrix} \operatorname{Re} \mathbf{v} & \operatorname{Im} \mathbf{v} \end{pmatrix} \begin{pmatrix} \operatorname{Re} \lambda & \operatorname{Im} \lambda \\ -\operatorname{Im} \lambda & \operatorname{Re} \lambda \end{pmatrix}. \quad (10.46)$$

Nehmen wir nun an, dass die Matrix \mathbf{A} in $\dot{\mathbf{y}}(t) = \mathbf{A}\mathbf{y}(t)$ die nicht-reellen Eigenwerte λ_1 und $\lambda_2 = \bar{\lambda}_1$ hat sowie die zugehörigen Eigenvektoren \mathbf{v}_1 und $\mathbf{v}_2 = \bar{\mathbf{v}}_1$. Alle anderen $n - 2$ Eigenwerte und Eigenvektoren seien reell. (Der allgemeine Fall geht analog.) (Der allgemeine Fall lässt sich aber ganz analog behandeln.) Wir sind jetzt nur an Lösungen interessiert, die reell sind. Diese bekommt man, indem man in (10.10) $\gamma_2 = \bar{\gamma}_1$ und $\gamma_3, \dots, \gamma_n \in \mathbb{R}$ wählt, denn dann ist, mit \mathbf{z} aus (10.7),

$$\mathbf{y}(t) = \mathbf{V}\mathbf{z}(t) = \mathbf{v}\gamma_1 e^{\lambda_1 t} + \bar{\mathbf{v}}\bar{\gamma}_1 e^{\bar{\lambda}_1 t} + \mathbf{v}_3\gamma_3 e^{\lambda_3 t} + \dots + \mathbf{v}_n\gamma_n e^{\lambda_n t} \in \mathbb{R}^n.$$

Definieren wir die blockdiagonale Matrix

$$\mathbf{K} := \begin{pmatrix} \mathbf{K}_2 & \mathbf{O} \\ \mathbf{O} & \mathbf{I}_{n-2} \end{pmatrix},$$

so kann man diese allgemeine reelle Lösung auch schreiben als

$$\mathbf{y}(t) = (\mathbf{VK})(\mathbf{K}^{-1}\mathbf{z}(t)), \quad (10.47)$$

wobei jetzt hier beide Klammern nur reelle Größen enthalten.

Zur Bestimmung der Parameter aufgrund der Anfangsbedingungen ersetzen wir schliesslich (10.11), d.h. $\mathbf{V}\mathbf{c} = \mathbf{y}_0$, durch

$$(\mathbf{VK})(\mathbf{K}^{-1}\mathbf{c}) = \mathbf{y}_0, \quad (10.48)$$

wo nun auch wieder die Größen in Klammer reell sind:

$$\tilde{\mathbf{c}} := \mathbf{K}^{-1}\mathbf{c} = \begin{pmatrix} 2\operatorname{Re} c_1 & 2\operatorname{Im} c_1 & c_3 & \dots & c_n \end{pmatrix}^T. \quad (10.49)$$

10.4 Quadratische Formen und ihre Hauptachsentransformation

Wir betrachten hier erstmals Funktionen, die in mehreren Variablen, etwa x_1, \dots, x_n quadratisch sind, und wir werden vorerst annehmen, dass nur quadratische Terme, aber keine linearen oder konstanten Terme vorkommen. Ein Beispiel ist

$$Q(x_1, x_2) := 4x_1^2 - 24x_1x_2 + 11x_2^2. \quad (10.50)$$

Ziel wird es sein, Q als Summe oder Differenz von Quadraten darzustellen, wie das nötig ist, um festzustellen, ob die Gleichung $Q(x_1, x_2) = \text{const}$ eine Ellipse oder eine Hyperbel beschreibt.

Für die drei **Kegelschnitte** [*conic sections*] oder **Kurven zweiter Ordnung** [*quadratic curves*] lauten die Gleichungen in Normalform, d.h. wenn die Achsen des Kegelschnittes mit den Koordinatenachsen zusammenfallen und der Mittel- bzw. Scheitelpunkt in 0 liegt,

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1 \quad \text{Ellipse [ellipse] mit Halbachsen } a \text{ und } b,$$

$$\frac{x_1^2}{a^2} - \frac{x_2^2}{b^2} = 1 \quad \text{nach links und rechts geöffnete Hyperbel [hyperbola] mit reeller Halbachse } a \text{ und Asymptotensteigung } \pm \arctan(b/a),$$

$$x_1^2 - 2p x_2 = 0 \quad \text{nach oben geöffnete Parabel [parabola] mit Halbparameter } p.$$

Dabei unterscheidet sich die Parabelgleichung massgeblich von den beiden anderen: x_2 kommt nur in einem linearen Term vor. Auf diesen Fall werden wir in Abschnitt 10.5 zurückkommen. Ebenso werden wir dort Ellipsen und Hyperbeln betrachten, deren Mittel- bzw. Scheitelpunkt nicht in 0 liegt.

Wir gehen diese Probleme nicht nur in zwei sondern in n Variablen an. Im \mathbb{R}^3 spricht man von **Flächen zweiter Ordnung** [*quadratic surface*]. Wir definieren zuerst drei eng verknüpfte Funktionstypen:

DEFINITION: Eine auf $\mathbb{R}^n \times \mathbb{R}^n$ definierte **symmetrische bilineare Form** [*symmetric bilinear form*] B ist eine Funktion der Form

$$(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^n \quad \longmapsto \quad B(\mathbf{x}, \mathbf{y}) := \mathbf{x}^T \mathbf{A} \mathbf{y} \in \mathbb{R}, \quad (10.51)$$

mit einer reellen, symmetrischen $n \times n$ -Matrix \mathbf{A} . Die zugehörige auf \mathbb{R}^n definierte **quadratische Form** [*quadratic form*] Q ist

$$\mathbf{x} \in \mathbb{R}^n \quad \longmapsto \quad Q(\mathbf{x}) := \mathbf{x}^T \mathbf{A} \mathbf{x} \in \mathbb{R}, \quad (10.52)$$

Eine auf $\mathbb{C}^n \times \mathbb{C}^n$ definierte **Hermiteische Form** [*Hermitian form*] B ist eine Funktion der Form

$$(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^n \quad \longmapsto \quad B(\mathbf{x}, \mathbf{y}) := \mathbf{x}^H \mathbf{A} \mathbf{y} \in \mathbb{R}, \quad (10.53)$$

mit einer Hermiteschen $n \times n$ -Matrix \mathbf{A} . ▲

Offensichtlich gilt gerade, wenn in (10.51) und (10.52) die gleiche Matrix verwendet wird, dass

$$Q(\mathbf{x}) = B(\mathbf{x}, \mathbf{x}). \quad (10.54)$$

Die Voraussetzung einer symmetrischen Matrix \mathbf{A} erfolgt in Q ohne Beschränkung der Allgemeinheit: Definiert man $Q(\mathbf{x}) := \mathbf{x}^T \mathbf{B} \mathbf{x}$ mit einer unsymmetrischen Matrix \mathbf{B} , so kann man diese durch die symmetrische Matrix $\mathbf{A} := \frac{1}{2}(\mathbf{B} + \mathbf{B}^T)$ ersetzen, ohne dass die Werte von Q ändern.

Wenn man $Q(\mathbf{x})$ ausschreibt, so erkennt man, dass (10.50) verallgemeinert wird:

$$Q(\mathbf{x}) := \mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_{k=1}^n \sum_{j=1}^n a_{kj} x_k x_j. \quad (10.55)$$

Wir wollen diese Formel für Q durch Basistransformation so umformen, dass keine gemischten Terme mehr auftreten. Das gelingt uns sofort mit Hilfe der Spektralzerlegung für symmetrische Matrizen (Korollar 9.16). Setzt man

$$\mathbf{A} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T, \quad \tilde{\mathbf{x}} := \mathbf{U}^T \mathbf{x} \quad (10.56)$$

in (10.55) ein, resultiert wegen $\mathbf{x}^T \mathbf{U} = \tilde{\mathbf{x}}^T$

$$Q(\mathbf{x}) = \tilde{Q}(\tilde{\mathbf{x}}) := \tilde{\mathbf{x}}^T \mathbf{\Lambda} \tilde{\mathbf{x}} = \sum_{k=1}^n \lambda_k \tilde{x}_k^2. \quad (10.57)$$

Ellipsen und Hyperbeln ($n = 2$) und Flächen zweiter Ordnung ($n = 3$) werden, wie erwähnt, durch Gleichungen der Form $Q(\mathbf{x}) = \text{const}$ dargestellt. Nach der Koordinatentransformation $\mathbf{x} \mapsto \tilde{\mathbf{x}}$ liegen die Hauptachsen dieser Kurven und Flächen auf den Koordinatenachsen, weshalb man von **Hauptachsen-Transformation** spricht.

Satz 10.2 [Hauptachsen-Transformation]

Jede quadratische Form $Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ lässt sich mittels einer auf der Spektralzerlegung von \mathbf{A} beruhenden orthogonalen Basistransformation auf die Hauptachsen-Darstellung (10.57) bringen.

Ist man bereit, zusätzlich die Länge der neuen Basisvektoren anzupassen, so kann man in (10.57) die λ_k durch ± 1 und 0 ersetzen:

Korollar 10.3 Jede quadratische Form $Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x}$ lässt sich durch Übergang auf eine neue orthogonale (aber im allgemeinen nicht orthonormierte) Basis mit Koordinaten \mathbf{y} auf die folgende Form bringen:

$$Q(\mathbf{x}) = \mathbf{y}^T \mathbf{I}_{\pm} \mathbf{y} = \sum_{k=1}^p y_k^2 - \sum_{l=p+1}^r y_l^2. \quad (10.58)$$

Die Zahl p ist dabei gleich der Anzahl positiver Eigenwerte von \mathbf{A} , und r ist der Rang von \mathbf{A} , d.h. $r - p$ ist die Anzahl negativer Eigenwerte.

BEWEIS: Man permutiert die Komponenten von \mathbf{x} so, dass jene mit positivem λ_k zuerst, jene mit $\lambda_k = 0$ zuletzt stehen. Zudem setzt man

$$y_k := \sqrt{|\lambda_k|} \tilde{x}_k \quad (k = 1, \dots, n). \quad \blacksquare$$

DEFINITION: Das Tripel $(p, r-p, n-r)$ aus Satz 10.3 mit der Zahl der positiven, negativen und null Eigenwerte von \mathbf{A} heisst **Trägheit** [*inertia*] von \mathbf{A} . Die Differenz $p - (r-p)$ der Zahl positiver und negativer Eigenwerte heisst **Signatur** [*signature*] von \mathbf{A} .

BEISPIEL 10.4: Der quadratischen Form \mathbf{Q} von (10.43) entspricht die Matrix

$$\mathbf{A} = \begin{pmatrix} 4 & -12 \\ -12 & 11 \end{pmatrix}$$

mit der Spektralzerlegung

$$\mathbf{A} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T = \begin{pmatrix} 0.6 & 0.8 \\ -0.8 & 0.6 \end{pmatrix} \begin{pmatrix} 20 & 0 \\ 0 & -5 \end{pmatrix} \begin{pmatrix} 0.6 & -0.8 \\ 0.8 & 0.6 \end{pmatrix}.$$

\mathbf{U} ist eine Rotation, \mathbf{A} hat die Trägheit $(1, 1, 0)$ und die Signatur 0. Wir erhalten

$$Q(x_1, x_2) = \tilde{Q}(\tilde{x}_1, \tilde{x}_2) = 20\tilde{x}_1^2 - 5\tilde{x}_2^2, \quad (10.59)$$

wobei

$$\begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{pmatrix} = \begin{pmatrix} 0.6 & -0.8 \\ 0.8 & 0.6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{5}(3x_1 - 4x_2) \\ \frac{1}{5}(4x_1 + 3x_2) \end{pmatrix}. \quad (10.60)$$

Die Gleichung $Q(x_1, x_2) = 20$ hat damit die Normalform

$$\frac{1}{20} \tilde{Q}(\tilde{x}_1, \tilde{x}_2) = 1, \quad \text{d.h.} \quad \tilde{x}_1^2 - \frac{1}{4} \tilde{x}_2^2 = 1.$$

Dies ist eine nach links und rechts geöffnete Hyperbel mit Scheitelpunkten $(\pm 1, 0)$ und Öffnungswinkel $2 \arctan(b/a) = 2 \arctan(2)$. Im ursprünglichen Koordinatensystem haben die Scheitel die Koordinaten $\begin{pmatrix} x_1 & x_2 \end{pmatrix}^T = \mathbf{U} \begin{pmatrix} \pm 1 & 0 \end{pmatrix}^T = \pm \begin{pmatrix} 0.6 & -0.8 \end{pmatrix}^T$. \blacklozenge

BEISPIEL 10.5: Zur quadratischen Form

$$\begin{aligned} Q(\mathbf{x}) &= Q(x_1, x_2, x_3) \\ &= 221x_1^2 + 144x_1x_2 + 480x_1x_3 + 179x_2^2 + 360x_2x_3 + 100x_3^2. \end{aligned} \quad (10.61)$$

gehört die Matrix

$$\mathbf{A} = \begin{pmatrix} 221 & 72 & 240 \\ 72 & 179 & 180 \\ 240 & 180 & 100 \end{pmatrix} \quad (10.62)$$

mit der Spektralzerlegung

$$\begin{aligned} \mathbf{A} &= \begin{pmatrix} 0.60 & 0.48 & 0.64 \\ -0.80 & 0.36 & 0.48 \\ 0.00 & -0.80 & 0.60 \end{pmatrix} \begin{pmatrix} 125 & 0 & 0 \\ 0 & -125 & 0 \\ 0 & 0 & 500 \end{pmatrix} \times \\ &\times \begin{pmatrix} 0.60 & -0.80 & 0.00 \\ 0.48 & 0.36 & -0.80 \\ 0.64 & 0.48 & 0.60 \end{pmatrix}. \end{aligned}$$

Wenn wir die neuen Koordinaten

$$\begin{aligned}\tilde{x}_1 &= 0.60x_1 - 0.80x_2 \\ \tilde{x}_2 &= 0.48x_1 + 0.36x_2 - 0.80x_3 \\ \tilde{x}_3 &= 0.64x_1 + 0.48x_2 + 0.60x_3\end{aligned}$$

eingeführen, so erhalten wir also die Hauptachsendarstellung

$$Q(\mathbf{x}) = \tilde{Q}(\tilde{\mathbf{x}}) = 125\tilde{x}_1^2 - 125\tilde{x}_2^2 + 500\tilde{x}_3^2.$$

Die Gleichung $Q(\mathbf{x}) = 500$ führt damit auf

$$\frac{1}{500} \tilde{Q}(\tilde{\mathbf{x}}) = \frac{1}{4} \tilde{\mathbf{x}}_1^2 - \frac{1}{4} \tilde{\mathbf{x}}_2^2 + \tilde{\mathbf{x}}_3^2 = 1 \quad (10.63)$$

Für die Schnittkurven mit den drei Koordinatenebenen des neuen Koordinatensystems erhalten wir

$$\mathbf{x}_3 = 0 \implies \frac{1}{4} \tilde{\mathbf{x}}_1^2 - \frac{1}{4} \tilde{\mathbf{x}}_2^2 = 1 \quad (\text{Hyperbel}), \quad (10.64)$$

$$\mathbf{x}_2 = 0 \implies \frac{1}{4} \tilde{\mathbf{x}}_1^2 + \tilde{\mathbf{x}}_3^2 = 1 \quad (\text{Ellipse}), \quad (10.65)$$

$$\mathbf{x}_3 = 0 \implies -\frac{1}{4} \tilde{\mathbf{x}}_2^2 + \tilde{\mathbf{x}}_3^2 = 1 \quad (\text{Hyperbel}). \quad (10.66)$$

Eine solche Fläche zweiter Ordnung nennt man **einschaliges Hyperboloid** [*one-sheeted hyperboloid*]. Gäbe es in der Summe in (10.63) zwei Minuszeichen statt nur eines, hätten wir ein **zweischaliges Hyperboloid** [*two-sheeted hyperboloid*]. Wären alle drei Vorzeichen positiv (und damit alle drei Schnittkurven Ellipsen), wäre die Fläche ein **Ellipsoid** [*ellipsoid*]. \blacklozenge

Man kann Korollar 10.3 auch etwas anders formulieren. Wir können schon bei der Spektralzerlegung von \mathbf{A} darauf achten, dass die positiven Eigenwerte zuerst und die Eigenwerte null zuletzt aufgeführt werden. Definieren wir noch die Diagonalmatrizen

$$\mathbf{D} := \text{diag}(\sqrt{|\lambda_1|}, \dots, \sqrt{|\lambda_n|}), \quad (10.67)$$

$$\mathbf{I}_{\pm} := \text{diag}(\underbrace{1, \dots, 1}_p, \underbrace{-1, \dots, -1}_{r-p}, \underbrace{0, \dots, 0}_{n-r}), \quad (10.68)$$

so lässt sich die Transformation von Q ja wie folgt zusammenfassen:

$$Q(\mathbf{x}) = \mathbf{x}^T \mathbf{A} \mathbf{x} = \mathbf{x}^T \mathbf{U} \mathbf{\Lambda} \mathbf{U}^T \mathbf{x} = \mathbf{x}^T \underbrace{\mathbf{U} \mathbf{D}}_{\equiv: \mathbf{S}} \mathbf{I}_{\pm} \underbrace{\mathbf{D} \mathbf{U}^T}_{\equiv: \mathbf{S}^T} \mathbf{x} = \mathbf{y}^T \mathbf{I}_{\pm} \mathbf{y}.$$

$\underbrace{\quad}_{= \mathbf{y}^T} \quad \underbrace{\quad}_{= \mathbf{y}}$

Es wird also \mathbf{A} zerlegt in

$$\mathbf{A} = \mathbf{S} \mathbf{I}_{\pm} \mathbf{S}^T, \quad \text{wobei} \quad \mathbf{S} := \mathbf{U} \mathbf{D} \quad (10.69)$$

orthogonale Kolonnen hat.

DEFINITION: Zwei $n \times n$ -Matrizen \mathbf{A} und \mathbf{B} heißen **kongruent** [*congruent*], wenn es eine reguläre Matrix \mathbf{S} gibt, so dass

$$\mathbf{A} = \mathbf{S} \mathbf{B} \mathbf{S}^T.$$

Der Übergang von $\mathbf{B} \mapsto \mathbf{A} = \mathbf{S} \mathbf{B} \mathbf{S}^T$ (oder umgekehrt) heisst dann **Kongruenztransformation** [*congruence transformation*]. \blacktriangle

Damit ist bereits ein Teil des folgenden Satzes bewiesen.

Satz 10.4 [Trägheitssatz]

Zu jeder reellen symmetrischen Matrix \mathbf{A} gibt es eine kongruente Matrix \mathbf{I}_\pm der Form (10.68). Dabei ist das Tripel $(p, r - p, n - r)$ gleich der Trägheit von \mathbf{A} , die unabhängig ist von der zugrunde liegenden Kongruenztransformation.

BEWEIS: Was zu zeigen bleibt, ist dass p und r unabhängig sind von \mathbf{S} , selbst wenn wir eine beliebige reguläre Matrix \mathbf{S} zulassen, und nicht nur jene von der speziellen Form $\mathbf{S} = \mathbf{U}\mathbf{D}$ wie in (10.69).

Es seien $\mathbf{A} = \mathbf{S}\mathbf{I}_\pm\mathbf{S}^\top = \mathbf{S}'\mathbf{I}'_\pm(\mathbf{S}')^\top$ zwei verschiedene Kongruenztransformationen von \mathbf{A} auf die gewünschte Diagonalform (10.68), und es seien $(p, r - p, n - r)$ und $(p', r' - p', n - r')$ die darin auftauchenden Tripel von Indizes. Ferner seien $\tilde{\mathbf{x}}$ und $\tilde{\mathbf{x}}'$ die entsprechend transformierten Variablen, für die gilt

$$Q(\mathbf{x}) = \mathbf{x}^\top \mathbf{A} \mathbf{x} = \mathbf{y}^\top \mathbf{I}_\pm \mathbf{y} = (\mathbf{y}')^\top \mathbf{I}'_\pm \mathbf{y}'. \quad (10.70)$$

Schliesslich seien $\mathbf{b}_1, \dots, \mathbf{b}_n$ und $\mathbf{b}'_1, \dots, \mathbf{b}'_n$ die diesen Koordinaten entsprechenden Basen des \mathbb{R}^n .

Die Tatsache, dass $r = r' = \text{Rang } \mathbf{A}$, folgt schon daraus, dass nach Voraussetzung \mathbf{S} und \mathbf{S}' regulär sind und bei einer Zusammensetzung von Abbildungen (bzw. einem Produkt von Matrizen) der Rang nicht wachsen kann; siehe Korollar 5.10 und Satz 5.16. Aus (10.70) schliesst man, dass in den Unterräumen

$$U := \text{span} \{\mathbf{b}_1, \dots, \mathbf{b}_p\}, \quad W := \text{span} \{\mathbf{b}'_{p'+1}, \dots, \mathbf{b}'_n\}$$

gilt:

$$Q(\mathbf{x}) > 0 \quad \text{für alle } \mathbf{x} \in U, \quad Q(\mathbf{x}) \leq 0 \quad \text{für alle } \mathbf{x} \in W.$$

Folglich ist $U \cap W = \emptyset$, also $p + (n - p') \leq n$, d.h. $p \leq p'$. Aus Symmetriegründen gilt auch $p' \leq p$, also notwendigerweise $p = p'$. Mit dem gleichen Argument hätte man auch zeigen können, dass $r - p = r' - p'$ sein muss.

Da in der zuvor betrachteten speziellen Kongruenztransformation das Tripel $(p, r - p, n - r)$ gleich der Trägheit von \mathbf{A} ist, muss dies allgemein gelten. ■

Erlaubt man bei der Kongruenztransformation eine beliebige reguläre Matrix, so lässt sich diese auch ohne Lösen des Eigenwertproblems in einem endlichen Algorithmus finden. Im wesentlichen handelt es sich dabei um eine symmetrische LR-Zerlegung mit gleichzeitigen Spalten- und Zeilenpermutationen, also um eine Zerlegung der Form

$$\mathbf{A} = \mathbf{P}\mathbf{L}\mathbf{D}\mathbf{L}^\top\mathbf{P}^\top, \quad (10.71)$$

aber es gibt noch Fälle, wo ein zusätzlicher Trick nötig sind. Einen besonders einfachen Fall, wo man ohne Permutationen auskommt, haben wir bereits in Abschnitt 3.4 betrachtet: die symmetrisch positiv definiten Matrizen. Da die Cholesky-Zerlegung einer solchen

Matrix die Form (10.69) mit $\mathbf{I}_\pm = \mathbf{I}$ hat, folgt aus dem Trägheitssatz 10.4 sofort die folgende Aussage:

Satz 10.5 *Eine symmetrische Matrix ist genau dann positiv definit, wenn alle ihre Eigenwerte positiv sind.*

10.5 Quadratische Funktionen

Ist ein Kegelschnitt in allgemeiner Lage, so enthält seine Gleichung in der Regel auch lineare Terme, wie im Beispiel

$$4x^2 - 24xy + 11y^2 + 8x + 26y = 5. \quad (10.72)$$

Auch hier möchte man die Gleichung auf Normalform bringen, um Lage, Typ und Parameter des Kegelschnittes zu bestimmen. Dazu muss man jetzt das Koordinatensystem nicht nur drehen, sondern auch verschieben.

Wir wollen auch dieses Problem allgemein angehen und quadratische Funktionen in n (statt zwei) Variablen betrachten.

Wir schreiben jetzt eine gegebene, im \mathbb{R}^n definierte, **quadratische Funktion** [quadratic function] F als

$$F(\mathbf{x}) = Q(\mathbf{x}) - 2\mathbf{b}^\top \mathbf{x} - \gamma, \quad (10.73)$$

wobei Q eine quadratische Form und γ eine reelle Zahl ist. Wir wissen bereits, dass man $Q(\mathbf{x})$ mittels Hauptachsentransformation auf die Normalform $\tilde{Q}(\tilde{\mathbf{x}})$ transformieren kann. Man hat dann wegen $\mathbf{x} = \mathbf{U}\tilde{\mathbf{x}}$

$$F(\mathbf{x}) = \tilde{F}(\tilde{\mathbf{x}}) := \tilde{Q}(\tilde{\mathbf{x}}) - 2\tilde{\mathbf{b}}^\top \tilde{\mathbf{x}} - \gamma \quad \text{mit} \quad \tilde{\mathbf{b}}^\top := \mathbf{b}^\top \mathbf{U}. \quad (10.74)$$

Ist zum Beispiel $\lambda_k \neq 0$, so kann man den Term $2\tilde{b}_k \tilde{x}_k$ in $2\tilde{\mathbf{b}}^\top \tilde{\mathbf{x}}$ durch eine Verschiebung in der Koordinatenrichtung \tilde{x}_k wegtransformieren:

$$\lambda_k \tilde{x}_k^2 - 2\tilde{b}_k \tilde{x}_k = \lambda_k \left(\tilde{x}_k - \frac{\tilde{b}_k}{\lambda_k} \right)^2 - \frac{\tilde{b}_k^2}{\lambda_k}.$$

Definieren wir noch

$$\mathbf{d} := \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} \quad \text{mit} \quad d_k := \begin{cases} \tilde{b}_k/\lambda_k & \text{falls } \lambda_k \neq 0, \\ 0 & \text{falls } \lambda_k = 0, \end{cases} \quad (10.75)$$

$$\mathbf{p} := \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix} \quad \text{mit} \quad p_k := \begin{cases} 0 & \text{falls } \lambda_k \neq 0, \\ \tilde{b}_k & \text{falls } \lambda_k = 0, \end{cases} \quad (10.76)$$

$$\tilde{\gamma} := \gamma + \sum_{\lambda_k \neq 0} \frac{\tilde{b}_k^2}{\lambda_k}, \quad (10.77)$$

so erhalten wir schliesslich

$$F(\mathbf{x}) = \tilde{F}(\tilde{\mathbf{x}}) = \tilde{Q}(\tilde{\mathbf{x}} - \mathbf{d}) - 2\mathbf{p}^T \tilde{\mathbf{x}} - \tilde{\gamma} \quad (10.78)$$

oder ausgeschrieben

$$F(\tilde{x}_1, \dots, \tilde{x}_n) = \sum_{\lambda_k \neq 0} \lambda_k (\tilde{x}_k - d_k)^2 - 2 \sum_{\lambda_k = 0} p_k \tilde{x}_k - \tilde{\gamma}. \quad (10.79)$$

Den zweiten Term in (10.78) kann man auch als $-2\mathbf{p}^T(\tilde{\mathbf{x}} - \mathbf{d})$ schreiben, wenn man $-\tilde{\gamma}$ durch $-\tilde{\gamma} - 2\mathbf{p}^T \mathbf{d}$ ersetzt.

Satz 10.6 Eine auf dem \mathbb{R}^n definierte quadratische Funktion (10.73) lässt sich auf die Form (10.78)–(10.79) reduzieren.

Setzt man in einer Gleichung $\tilde{F}(\tilde{\mathbf{x}}) = 0$ mit \tilde{F} aus (10.78) / (10.79) $n - 2$ der Komponenten $\tilde{x}_k - d_k$ von $\tilde{\mathbf{x}} - \mathbf{d}$ gleich Null, so erhält man einen Kegelschnitt in Normalform oder die leere Menge. Solch ein Kegelschnitt entspricht dem Schnitt der durch $F(\mathbf{x}) = 0$ definierten Kegelfläche im \mathbb{R}^n mit einer geeignet verschobenen Koordinatenebene im $\tilde{\mathbf{x}}$ -Koordinatensystem.

Offenbar treten Parabeln als Schnittkurven genau dann auf, wenn es Eigenwerte gibt, die null sind. Reine Ellipsoide (d.h. Ellipsen für alle Schnitte) erfordern, dass alle Eigenwerte das gleiche Vorzeichen wie γ haben müssen. Ist es entgegengesetzt, so ist die Lösungsmenge von $F(\mathbf{x}) = 0$ leer.

10.6 Die Spektralnorm

Nun wollen wir auf die Berechnung der in (6.65) eingeführten Spektralnorm oder 2-Norm einer Matrix,

$$\|\mathbf{A}\|_2 := \sup_{\mathbf{x} \neq 0} \frac{\|\mathbf{A}\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \sup_{\|\mathbf{x}\|_2=1} \|\mathbf{A}\mathbf{x}\|_2,$$

eingehen, die auch die Konditionszahl $\kappa_2(\mathbf{A})$ aus (6.76) liefert.

Satz 10.7 Die Spektralnorm einer Matrix $\mathbf{A} \in \mathbb{E}^{n \times n}$ ist gleich der Wurzel aus dem grössten Eigenwert von $\mathbf{A}^H \mathbf{A}$ oder, äquivalent,

$$\|\mathbf{A}\|_2 = \max\{\sqrt{\omega}; \omega \text{ Eigenwert von } \mathbf{A}^H \mathbf{A}\}. \quad (10.80)$$

Falls \mathbf{A} reell ist, ist natürlich $\mathbf{A}^H \mathbf{A} = \mathbf{A}^T \mathbf{A}$.

BEWEIS: Da $\mathbf{A}^H \mathbf{A}$ Hermitesch ist, gibt es gemäss Satz 9.15 eine Spektralzerlegung $\mathbf{A}^H \mathbf{A} = \mathbf{U} \mathbf{\Omega} \mathbf{U}^H$ mit unitärem \mathbf{U} und diagonalem $\mathbf{\Omega}$. Weil $\mathbf{A}^H \mathbf{A}$ Hermitesch positiv semidefinit ist, sind dabei die Diagonalelemente ω_k von $\mathbf{\Omega}$ nicht negativ. Mit $\mathbf{y} := \mathbf{U}^H \mathbf{x}$ folgt dank $\|\mathbf{y}\|_2 = \|\mathbf{x}\|_2$:

$$\begin{aligned} \|\mathbf{A}\|_2^2 &= \sup_{\|\mathbf{x}\|_2=1} \|\mathbf{A}\mathbf{x}\|_2^2 = \sup_{\|\mathbf{x}\|_2=1} \mathbf{x}^H \mathbf{A}^H \mathbf{A} \mathbf{x} = \sup_{\|\mathbf{x}\|_2=1} \mathbf{x}^H \mathbf{U} \mathbf{\Omega} \mathbf{U}^H \mathbf{x} \\ &= \sup_{\|\mathbf{y}\|_2=1} \mathbf{y}^H \mathbf{\Omega} \mathbf{y} = \sup_{\|\mathbf{y}\|_2=1} \sum_{k=1}^n \omega_k |y_k|^2 = \max_{k=1, \dots, n} \omega_k. \quad \blacksquare \end{aligned}$$

Für eine Hermitesche Matrix vereinfacht sich die Formel: es ist ja $\mathbf{A}^H \mathbf{A} = \mathbf{A}^2$, und allgemein sind die Eigenwerte von \mathbf{A}^2 die Quadrate jener von \mathbf{A} . Also ist in diesem Fall $\omega_k = \lambda_k^2$ und damit $\max \sqrt{\omega_k} = \max |\lambda_k|$. Es gilt somit:

Satz 10.8 Die Spektralnorm einer Hermiteschen (oder reell symmetrischen) Matrix $\mathbf{A} \in \mathbb{E}^{n \times n}$ ist gleich dem Betrag des grössten Eigenwertes dieser Matrix:

$$\|\mathbf{A}\|_2 = \max\{|\omega|; \omega \text{ Eigenwert von } \mathbf{A}\}. \quad (10.81)$$

Noch einfacher ist der Fall einer unitären (oder orthogonalen) Matrix \mathbf{U} : Wegen $\mathbf{U}^H \mathbf{U} = \mathbf{I}$ ist stets $\|\mathbf{U}\|_2 = 1$.

Für die 2-Norm der Inversen \mathbf{A}^{-1} einer Matrix kann man analoge Formeln herleiten wie für die 2-Norm von \mathbf{A} selbst:

Satz 10.9 Die Spektralnorm der Inversen \mathbf{A}^{-1} einer Matrix ist gleich dem Inversen der Wurzel aus dem kleinsten Eigenwert von $\mathbf{A}^H \mathbf{A}$ oder, äquivalent,

$$\|\mathbf{A}^{-1}\|_2 = \max \left\{ \frac{1}{\sqrt{\omega}}; \omega \text{ Eigenwert von } \mathbf{A}^H \mathbf{A} \right\} \quad (10.82)$$

$$= \frac{1}{\min \{ \sqrt{\omega}; \omega \text{ Eigenwert von } \mathbf{A}^H \mathbf{A} \}}. \quad (10.83)$$

Ist \mathbf{A} Hermitesch (oder reell symmetrisch), so ist $\|\mathbf{A}^{-1}\|_2$ gleich dem Inversen des Betrages des kleinsten Eigenwertes von \mathbf{A} :

$$\|\mathbf{A}^{-1}\|_2 = \max \left\{ \frac{1}{|\omega|}; \omega \text{ Eigenwert von } \mathbf{A} \right\} \quad (10.84)$$

$$= \frac{1}{\min \{ |\omega|; \omega \text{ Eigenwert von } \mathbf{A} \}}. \quad (10.85)$$

Die Sätze 10.7, 10.8 und 10.9 erlauben uns insbesondere, die in (6.77) eingeführte 2-Norm-Konditionszahl

$$\kappa_2(\mathbf{A}) = \|\mathbf{A}\|_2 \|\mathbf{A}^{-1}\|_2$$

zu berechnen.

Korollar 10.10 Die 2-Norm-Konditionszahl einer Matrix $\mathbf{A} \in \mathbb{E}^{n \times n}$ ist gegeben durch

$$\kappa_2(\mathbf{A}) = \frac{\max \{ \sqrt{\omega}; \omega \text{ Eigenwert von } \mathbf{A}^H \mathbf{A} \}}{\min \{ \sqrt{\omega}; \omega \text{ Eigenwert von } \mathbf{A}^H \mathbf{A} \}}. \quad (10.86)$$

Ist \mathbf{A} Hermitesch (oder reell symmetrisch), so vereinfacht sich diese Formel zu

$$\kappa_2(\mathbf{A}) = \frac{\max \{ |\omega|; \omega \text{ Eigenwert von } \mathbf{A} \}}{\min \{ |\omega|; \omega \text{ Eigenwert von } \mathbf{A} \}}. \quad (10.87)$$

Insbesondere ist die Konditionszahl immer grösser oder gleich 1.

Ist \mathbf{A} singulär (aber nicht die Nullmatrix), wird $\kappa_2(\mathbf{A}) = \infty$. Ist \mathbf{A} fast singulär, ist $\kappa_2(\mathbf{A})$ gross, das heisst \mathbf{A} ist schlecht konditioniert.

BEISPIEL 10.6: Wir betrachten wieder die bereits mehrmals verwendete Matrix aus (9.9),

$$\mathbf{A} = \begin{pmatrix} -7 & 2 & -6 \\ 12 & -2 & 12 \\ 12 & -3 & 11 \end{pmatrix},$$

die das Spektrum $\sigma(\mathbf{A}) = \{1, 2, -1\}$ hat. Der grösste Betrag eines Eigenwertes ist hier also 2, der kleinste 1, und der Quotient aus grösstem und kleinstem Eigenwertbetrag ist damit bloss 2. Aber da \mathbf{A} nicht symmetrisch ist, sind diese Zahlen nicht relevant für Spektralnorm und Konditionszahl, denn man muss die Eigenwerte von $\mathbf{A}^T \mathbf{A}$ betrachten, beziehungsweise deren Quadratwurzel.

$$\mathbf{A}^T \mathbf{A} = \begin{pmatrix} 337 & -74 & 318 \\ -74 & 17 & -69 \\ 318 & -69 & 301 \end{pmatrix}$$

hat das Spektrum

$$\sigma(\mathbf{A}^T \mathbf{A}) = \{0.0044\dots, 1.4063\dots, 653.5893\dots\}.$$

Es ergibt sich

$$\|\mathbf{A}\|_2 = \sqrt{653.5893\dots} = 25.5654\dots, \quad \|\mathbf{A}^{-1}\|_2 = 1/\sqrt{0.0044\dots} = 15.1587\dots$$

und

$$\kappa_2(\mathbf{A}) = \sqrt{\frac{\lambda_3}{\lambda_1}} = \sqrt{150186.5876\dots} = 387.5391\dots$$

Die Norm einer Matrix kann also viel grösser sein als der grösste Eigenwert und die Kondition kann viel grösser sein als der grösste Quotient der Eigenwerte. \blacklozenge

BEISPIEL 10.7: Sehr bekannte Beispiele liefern auch die sogenannten **Hilbert-Matrizen** [*Hilbert matrices*]

$$\mathbf{H}_n := \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \cdots & \frac{1}{n} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \cdots & \frac{1}{n+1} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \cdots & \frac{1}{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n} & \frac{1}{n+1} & \frac{1}{n+2} & \cdots & \frac{1}{2n-1} \end{pmatrix} \quad (10.88)$$

Es sind Hankel-Matrizen, das heisst h_{ij} hängt nur von der Summe $i + j$ ab (vgl. Abschnitt 2.9). Hankel-Matrizen sind notorisch für ihre schlechte Kondition, und die Hilbert-Matrizen ganz besonders. Es ist zum Beispiel

$$\begin{aligned} \|\mathbf{H}_5\|_2 &= 1.5671, \\ \|\mathbf{H}_5^{-1}\|_2 &= 304142.8416\dots, \\ \kappa_2(\mathbf{H}_5) &= 476607.2502\dots, \end{aligned}$$

\blacklozenge

Kapitel 11

Die Singulärwertzerlegung

In Kapitel 9 haben wir gesehen, dass man eine quadratische Matrix durch geeignete Wahl der Basis in die Jordansche Normalform bringen kann. Diese ist in der Regel diagonal, insbesondere stets wenn alle Eigenwerte verschieden sind oder wenn die Matrix symmetrisch oder Hermitesch ist. Im zweiten Falle kann man sogar die Basis orthogonal bzw. unitär wählen.

Bei einer linearen Abbildung $F : X \rightarrow Y$ kann man dagegen Basen in beiden Räumen wählen. Wie wir bereits in Satz 5.20 gesehen haben, lässt sich so auf einfache Art erreichen, dass die Abbildungsmatrix die Form

$$\mathbf{B} = \left(\begin{array}{c|c} \mathbf{I}_r & \mathbf{O} \\ \hline \mathbf{O} & \mathbf{O} \end{array} \right). \quad (11.1)$$

hat. Insbesondere lässt sich jede $m \times n$ Matrix \mathbf{A} durch zwei Koordinatentransformationen \mathbf{S} und \mathbf{T} in eine Matrix

$$\mathbf{B} = \mathbf{S}^{-1} \mathbf{A} \mathbf{T}$$

dieser Form überführen. Aber die einzige Information, die in der Matrix (11.1) noch enthalten ist, ist der Rang r der Abbildung. Diese Form ist in gewissem Sinne zu einfach, als dass man damit noch eine tiefe Einsicht in die Abbildung gewinnen könnte.

Wir werden in diesem Kapitel sehen, dass das ändert, wenn wir nur orthogonale bzw. unitäre Koordinatentransformationen zulassen. Es resultiert dann eine sowohl für theoretische Fragen als auch in vielen Anwendungen äusserst nützliche Matrixfaktorisierung, die Singulärwertzerlegung.

11.1 Die Singulärwertzerlegung: Herleitung

Gegeben sei irgend eine reelle oder komplexe $m \times n$ Matrix \mathbf{A} . Die Matrix $\mathbf{A}^H \mathbf{A} \in \mathbb{E}^{n \times n}$ ist Hermitesch und positiv semidefinit. Sie hat also reelle, nicht-negative Eigenwerte, sagen wir σ_k^2 ($k = 1, \dots, n$) mit $\sigma_k \geq 0$, und eine orthonormale Eigenbasis $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$. Es gilt also

$$\mathbf{A}^H \mathbf{A} \mathbf{V} = \mathbf{V} \Sigma_n^2 \quad (11.2)$$

mit der unitären $n \times n$ Matrix

$$\mathbf{V} = \left(\begin{array}{cccc} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{array} \right)$$

und der $n \times n$ Diagonalmatrix

$$\Sigma_n^2 := \text{diag} \{ \sigma_1^2, \dots, \sigma_n^2 \}.$$

Dabei können wir Eigenwerte und Eigenvektoren so ordnen, dass

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0, \quad \sigma_{r+1} = \dots = \sigma_n = 0, \quad (11.3)$$

worin $r := \text{Rang } \mathbf{A}^H \mathbf{A}$ ist. Multiplizieren wir (11.2) von links mit \mathbf{V}^H , so gilt wegen $\mathbf{V}^H \mathbf{V} = \mathbf{I}_n$, dass

$$\mathbf{V}^H \mathbf{A}^H \mathbf{A} \mathbf{V} = \mathbf{V}^H \mathbf{V} \boldsymbol{\Sigma}_n^2 = \boldsymbol{\Sigma}_n^2. \quad (11.4)$$

In den Kolonnen \mathbf{v}_k von \mathbf{V} ausgedrückt heisst das, dass

$$\|\mathbf{A} \mathbf{v}_k\|^2 = \langle \mathbf{A} \mathbf{v}_k, \mathbf{A} \mathbf{v}_k \rangle = \mathbf{v}_k^H \mathbf{A}^H \mathbf{A} \mathbf{v}_k = \sigma_k^2 \quad (k = 1, \dots, n),$$

oder

$$\|\mathbf{A} \mathbf{v}_k\| = \sigma_k \quad (k = 1, \dots, n). \quad (11.5)$$

Ist $r < n$, so folgt insbesondere, dass $\mathbf{A} \mathbf{v}_k = \mathbf{o}$ ($k = r+1, \dots, n$), das heisst $\mathbf{v}_{r+1}, \dots, \mathbf{v}_n$ bilden eine Orthonormalbasis eines Unterraumes von $\ker \mathbf{A} \equiv \mathcal{N}(\mathbf{A}) \subset \mathbb{E}^n$. Somit ist $\dim \mathcal{N}(\mathbf{A}) \geq n-r$, also nach der Dimensionsformel (5.36) $\text{Rang } \mathbf{A} = n - \dim \mathcal{N}(\mathbf{A}) \leq r$. Andererseits ist aber nach den Sätzen 5.16 (i) und 5.13

$$r = \text{Rang } \mathbf{A}^H \mathbf{A} \leq \min\{\text{Rang } \mathbf{A}, \text{Rang } \mathbf{A}^H\} = \text{Rang } \mathbf{A}.$$

Folglich gilt $\text{Rang } \mathbf{A}^H \mathbf{A} = \text{Rang } \mathbf{A}$. Also ist $\dim \mathcal{N}(\mathbf{A}) = n-r$, und $\mathbf{v}_{r+1}, \dots, \mathbf{v}_n$ bilden eine Orthonormalbasis von $\mathcal{N}(\mathbf{A})$. Wir unterteilen \mathbf{V} entsprechend in

$$\mathbf{V} = (\mathbf{V}_r \mid \mathbf{V}_\perp) := (\mathbf{v}_1 \ \dots \ \mathbf{v}_r \mid \mathbf{v}_{r+1} \ \dots \ \mathbf{v}_n). \quad (11.6)$$

Die Kolonnen von \mathbf{V}_r spannen dabei den zu $\mathcal{N}(\mathbf{A})$ orthogonalen Unterraum auf. Durch Beschränkung von \mathbf{A} auf diesen Unterraum haben wir statt (11.4)

$$\mathbf{V}_r^H \mathbf{A}^H \mathbf{A} \mathbf{V}_r = \boldsymbol{\Sigma}_r^2 := \text{diag} \{\sigma_1^2, \dots, \sigma_r^2\}. \quad (11.7)$$

Da hier nun $\boldsymbol{\Sigma}_r$ regulär ist, können wir von links und rechts mit $\boldsymbol{\Sigma}_r^{-1} = \boldsymbol{\Sigma}_r^{-T}$ multiplizieren:

$$\underbrace{(\boldsymbol{\Sigma}_r^{-T} \mathbf{V}_r^H \mathbf{A}^H)}_{= \mathbf{U}_r^H} \underbrace{(\mathbf{A} \mathbf{V}_r \boldsymbol{\Sigma}_r^{-1})}_{\equiv: \mathbf{U}_r} = \mathbf{I}_r.$$

Die $m \times r$ Matrix

$$\mathbf{U}_r := \mathbf{A} \mathbf{V}_r \boldsymbol{\Sigma}_r^{-1} \quad (11.8)$$

hat also orthonormale Kolonnen $\mathbf{u}_1, \dots, \mathbf{u}_r \in \mathbb{E}^m$. Wir können diese zu einer Orthonormalbasis von \mathbb{E}^m und damit \mathbf{U}_r zu einer unitären Matrix \mathbf{U} ergänzen:

$$\mathbf{U} = (\mathbf{U}_r \mid \mathbf{U}_\perp) := (\mathbf{u}_1 \ \dots \ \mathbf{u}_r \mid \mathbf{u}_{r+1} \ \dots \ \mathbf{u}_m). \quad (11.9)$$

Nach (11.8) ist $\mathbf{A} \mathbf{V}_r = \mathbf{U}_r \boldsymbol{\Sigma}_r$, zudem haben wir $\mathbf{A} \mathbf{V}_\perp = \mathbf{O}$, was wir zusammenfassen können in

$$\mathbf{A} \underbrace{(\mathbf{V}_r \mid \mathbf{V}_\perp)}_{= \mathbf{V}} = \underbrace{(\mathbf{U}_r \mid \mathbf{U}_\perp)}_{= \mathbf{U}} \underbrace{\left(\begin{array}{c|c} \boldsymbol{\Sigma}_r & \mathbf{O} \\ \hline \mathbf{O} & \mathbf{O} \end{array} \right)}_{\equiv: \boldsymbol{\Sigma}}. \quad (11.10)$$

Beachtet man noch, dass durch Transponieren die Aussage über \mathbf{A} zu einer völlig analogen Aussage über \mathbf{A}^H wird, worin \mathbf{U} und \mathbf{V} vertauscht sind, so ergibt sich der folgende fundamentale Satz.

Satz 11.1 Zu einer komplexen $m \times n$ Matrix \mathbf{A} vom Rang r existieren unitäre Matrizen $\mathbf{U} = (\mathbf{u}_1 \cdots \mathbf{u}_m)$ und $\mathbf{V} = (\mathbf{v}_1 \cdots \mathbf{v}_n)$ sowie eine $m \times n$ Matrix Σ der Form

$$\Sigma := \begin{pmatrix} \Sigma_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix} \quad (11.11a)$$

mit einem diagonalen $r \times r$ Block

$$\Sigma_r := \text{diag} \{ \sigma_1, \dots, \sigma_r \}, \quad (11.11b)$$

dessen Diagonalelemente positiv und gemäss (11.3) geordnet sind, so dass gilt

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^H = \sum_{k=1}^r \mathbf{u}_k \sigma_k \mathbf{v}_k^H. \quad (11.11c)$$

Dabei bilden die Kolonnen von \mathbf{U} und \mathbf{V} orthonormale Eigenbasen von $\mathbf{A}\mathbf{A}^H$ bzw. $\mathbf{A}^H\mathbf{A}$:

$$\mathbf{A}\mathbf{A}^H = \mathbf{U}\Sigma_m^2\mathbf{U}^H, \quad \mathbf{A}^H\mathbf{A} = \mathbf{V}\Sigma_n^2\mathbf{V}^H \quad (11.11d)$$

mit $\Sigma_m := \text{diag} \{ \sigma_1, \dots, \sigma_m \}$, $\Sigma_n := \text{diag} \{ \sigma_1, \dots, \sigma_n \}$, wo $\sigma_k = 0$ falls $k > r$. Äquivalent zu (11.11c) hat man die Beziehungen

$$\mathbf{A}\mathbf{V} = \mathbf{U}\Sigma, \quad \mathbf{A}^H\mathbf{U} = \mathbf{V}\Sigma^T, \quad (11.11e)$$

die zeigen, dass folgendes gilt:

$$\begin{aligned} \{ \mathbf{u}_1, \dots, \mathbf{u}_r \} & \text{ ist Basis von } \text{im } \mathbf{A} \equiv \mathcal{R}(\mathbf{A}), \\ \{ \mathbf{v}_1, \dots, \mathbf{v}_r \} & \text{ ist Basis von } \text{im } \mathbf{A}^H \equiv \mathcal{R}(\mathbf{A}^H), \\ \{ \mathbf{u}_{r+1}, \dots, \mathbf{u}_m \} & \text{ ist Basis von } \text{ker } \mathbf{A}^H \equiv \mathcal{N}(\mathbf{A}^H), \\ \{ \mathbf{v}_{r+1}, \dots, \mathbf{v}_n \} & \text{ ist Basis von } \text{ker } \mathbf{A} \equiv \mathcal{N}(\mathbf{A}), \end{aligned} \quad (11.11f)$$

wobei dies alles Orthonormalbasen sind.

Ist \mathbf{A} reell, so kann man \mathbf{U} und \mathbf{V} als reelle orthogonale Matrizen wählen, und alle Exponenten H können durch T ersetzt werden.

DEFINITION: Die Matrixfaktorisierung (11.11c) heisst **Singulärwertzerlegung** [*singular value decomposition*] oder kurz **SVD** von \mathbf{A} . Die Zahlen

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_{\min\{m,n\}} = 0 \quad (11.12)$$

heissen **Singulärwerte** [*singular values*] von \mathbf{A} . (Dabei kann $r = \min\{m, n\}$ sein, so dass alle $\sigma_k > 0$ sind.)

Die Vektoren $\mathbf{u}_1, \dots, \mathbf{u}_m$ sind die **linken Singulärvektoren** [*left singular vectors*], und die Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_n$ sind die **rechten Singulärvektoren** [*right singular vectors*]. ▲

BEMERKUNG: Man beachte, dass gemäss (11.11c) die Singulärwertzerlegung (ähnlich wie die Eigenwertzerlegung einer diagonalisierbaren quadratischen Matrix) eine additive Zerlegung einer beliebigen Matrix \mathbf{A} in eine Summe von r Rang-1-Matrizen liefert. Ist $r \ll \min\{m, n\}$, so ergibt diese eine effiziente Art der Berechnung von Matrix-Vektor-Produkten $\mathbf{A}\mathbf{x}$. ▼

BEISPIEL 11.1: Es sei

$$\mathbf{A} := \begin{pmatrix} 68 & -74 \\ 14 & -52 \\ 46 & -28 \\ -17 & -44 \end{pmatrix}, \quad (11.13)$$

also

$$\mathbf{A}^T \mathbf{A} = \begin{pmatrix} 7225 & -6300 \\ -6300 & 10900 \end{pmatrix},$$

$$\mathbf{A} \mathbf{A}^T = \begin{pmatrix} 10100 & 4800 & 5200 & 2100 \\ 4800 & 2900 & 2100 & 2050 \\ 5200 & 2100 & 2900 & 450 \\ 2100 & 2050 & 450 & 2225 \end{pmatrix}.$$

Die Eigenwertzerlegung von $\mathbf{A}^T \mathbf{A}$ und $\mathbf{A} \mathbf{A}^T$ liefert die folgenden, in (11.11d) auftretenden Matrizen:

$$\mathbf{\Sigma}_n = \text{diag}(125, 50), \quad \mathbf{\Sigma}_m = \text{diag}(125, 50, 0, 0),$$

$$\mathbf{V} = \begin{pmatrix} 0.60 & 0.80 \\ -0.80 & 0.60 \end{pmatrix},$$

$$\mathbf{U} = \begin{pmatrix} 0.8000 & 0.2000 & -0.5488 & 0.1372 \\ 0.4000 & -0.4000 & 0.5831 & 0.5831 \\ 0.4000 & 0.4000 & 0.5831 & -0.5831 \\ 0.2000 & -0.8000 & -0.1372 & -0.5488 \end{pmatrix},$$

wobei in \mathbf{U} die letzten beiden Kolonnen auf vier Stellen nach dem Komma gerundet sind.

Die 2×2 -Matrix $\mathbf{A}^T \mathbf{A}$ hat also die Eigenwerte $\sigma_1^2 = 125^2 = 15625$ und $\sigma_2^2 = 50^2 = 2500$, wogegen die 4×4 -Matrix $\mathbf{A} \mathbf{A}^T$ die Eigenwerte 15625, 2500, 0, 0 hat. Insbesondere ist $\text{Rang } \mathbf{A}^T \mathbf{A} = \text{Rang } \mathbf{A} \mathbf{A}^T = 2$.

Während $\mathbf{\Sigma}_n$ und $\mathbf{\Sigma}_m$ quadratische Diagonalmatrizen sind, hat die Matrix $\mathbf{\Sigma}$ in der Singulärwertzerlegung (11.11c) die gleiche Form wie \mathbf{A} . Dabei ist der quadratische Block $\mathbf{\Sigma}_r$ hier wegen $r = n = 2$ gleich $\mathbf{\Sigma}_n$:

$$\mathbf{\Sigma} = \begin{pmatrix} 125 & 0 \\ 0 & 50 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Die Singulärwertzerlegung (11.11c) ergibt sich damit hier als

$$\underbrace{\begin{pmatrix} 68 & -74 \\ 14 & -52 \\ 46 & -28 \\ -17 & -44 \end{pmatrix}}_{\mathbf{A}} = \underbrace{\begin{pmatrix} 0.8000 & 0.2000 & -0.5488 & 0.1372 \\ 0.4000 & -0.4000 & 0.5831 & 0.5831 \\ 0.4000 & 0.4000 & 0.5831 & -0.5831 \\ 0.2000 & -0.8000 & -0.1372 & -0.5488 \end{pmatrix}}_{\mathbf{U}} \times \underbrace{\begin{pmatrix} 125 & 0 \\ 0 & 50 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}}_{\mathbf{\Sigma}} \underbrace{\begin{pmatrix} 0.60 & 0.80 \\ -0.80 & 0.60 \end{pmatrix}}_{\mathbf{V}} \quad (11.14)$$

und liefert die additive Rang-1-Zerlegung

$$\mathbf{A} = \underbrace{\begin{pmatrix} 0.8 \\ 0.4 \\ 0.4 \\ 0.2 \end{pmatrix}}_{\mathbf{u}_1} \underbrace{125}_{\sigma_1} \underbrace{\begin{pmatrix} 0.6 & 0.8 \end{pmatrix}}_{\mathbf{v}_1} + \underbrace{\begin{pmatrix} 0.2 \\ -0.4 \\ 0.4 \\ -0.8 \end{pmatrix}}_{\mathbf{u}_2} \underbrace{50}_{\sigma_2} \underbrace{\begin{pmatrix} -0.8 & 0.6 \end{pmatrix}}_{\mathbf{v}_2}. \quad (11.15)$$

◆

Die Schlussformeln (11.14) und (11.15) im vorangehenden Beispiel 11.1 zeigen, dass die letzten zwei Kolonnen von \mathbf{U} effektiv nicht benutzt werden. Das sieht man in der Tat allgemein schon in der Darstellung (11.10) der Singulärwertzerlegung oder in unserer Definition (11.8) von \mathbf{U}_r . Offensichtlich gilt:

Korollar 11.2 *Bezeichnen wir wie in (11.6), (11.9) und (11.10) mit \mathbf{V}_r die Matrix aus den r ersten Kolonnen von \mathbf{V} , mit \mathbf{U}_r die Matrix aus den r ersten Kolonnen von \mathbf{U} und mit $\mathbf{\Sigma}_r$ die führende $r \times r$ Hauptuntermatrix von $\mathbf{\Sigma}$, so lässt sich die Singulärwertzerlegung (11.11c) in der kompakten Form*

$$\mathbf{A} = \mathbf{U}_r \mathbf{\Sigma}_r \mathbf{V}_r^H = \sum_{k=1}^r \mathbf{u}_k \sigma_k \mathbf{v}_k^H \quad (11.16)$$

schreiben, worin die $m \times r$ Matrix \mathbf{U} und die $n \times r$ Matrix \mathbf{V} orthonormale Kolonnen haben und die $r \times r$ Diagonalmatrix $\mathbf{\Sigma}_r$ positive Diagonalelemente $\sigma_1 \geq \dots \geq \sigma_r > 0$ hat.

11.2 Die Singulärwertzerlegung: Folgerungen

Aus der Singulärwertzerlegung lassen sich leicht eine ganze Reihe von interessanten Folgerungen ziehen.

Korollar 11.3 *Die Unterräume $\mathcal{R}(\mathbf{A})$ und $\mathcal{N}(\mathbf{A}^H)$ sind zueinander orthogonal und spannen zusammen den Bildraum \mathbf{E}^m auf. Analog sind die Unterräume $\mathcal{R}(\mathbf{A}^H)$ und $\mathcal{N}(\mathbf{A})$ zueinander orthogonal und spannen zusammen den Definitionsraum \mathbf{E}^n auf.*

BEWEIS: Die Aussage folgt sofort aus (11.11f) und der Tatsache, dass die unitären Matrizen \mathbf{U} und \mathbf{V} orthogonale Kolonnen haben. ■

Die Unterräume $\mathcal{R}(\mathbf{A}) \perp \mathcal{N}(\mathbf{A}^H) \subseteq \mathbf{E}^m$ und $\mathcal{N}(\mathbf{A}) \perp \mathcal{R}(\mathbf{A}^H) \subseteq \mathbf{E}^n$, die in Korollar 11.3 vorkommen und für die die Singulärwertzerlegung gemäss (11.11f) orthonormale Basen liefert, sind gerade die vier fundamentalen Unterräume der Matrix \mathbf{A} , die wir bereits in Abschnitt 6.4 angetroffen haben. Weiter folgt sofort:

Korollar 11.4 *Es sei $\mathbf{A} \in \mathbb{E}^{m \times n}$, $\text{Rang } \mathbf{A} = r$. Dann sind die r positiven Eigenwerte von $\mathbf{A}^H \mathbf{A}$ und $\mathbf{A} \mathbf{A}^H$ gleich, aber die Vielfachheit des Eigenwertes 0 ist $n - r$ bzw. $m - r$.*

BEWEIS: Dass die nicht-null Eigenwerte gleich sind, folgt aus (11.11d). ■

Nun lassen sich auch Satz 10.7 über die Spektralnorm und Korollar 10.10 über die entsprechende Konditionszahl neu formulieren:

Korollar 11.5 *Für die Spektralnorm einer Matrix $\mathbf{A} \in \mathbb{E}^{n \times n}$ gilt*

$$\|\mathbf{A}\|_2 = \sigma_1. \quad (11.17)$$

Ist \mathbf{A} regulär, so gilt für die (der Spektralnorm zugeordnete) Konditionszahl von $\mathbf{A} \in \mathbb{E}^{n \times n}$

$$\kappa_2(\mathbf{A}) = \frac{\sigma_1}{\sigma_n}. \quad (11.18)$$

Die wohl wichtigste Anwendung der Singulärwertzerlegung ist eine beste Approximationseigenschaft der Teilsummen der Summe in (11.11e). Wir formulieren sie nur für quadratische Matrizen, sie gilt aber auch für andere, wenn man für diese die 2-Norm geeignet definiert. Wir verzichten auf den Beweis.

Satz 11.6 *Die Matrix $\mathbf{A} \in \mathbb{E}^{n \times n}$ habe die Singulärwertzerlegung (11.11e) (mit $m = n$), und für $p = 1, \dots, r$ sei*

$$\mathbf{A}_p = \sum_{k=1}^p \mathbf{u}_k \sigma_k \mathbf{v}_k^H. \quad (11.19)$$

Dann ist \mathbf{A}_p im Sinne der Spektralnorm die beste Approximation von \mathbf{A} durch eine Matrix vom Rang p , das heisst es gilt

$$\|\mathbf{A} - \mathbf{A}_p\|_2 = \min \|\mathbf{A} - \mathbf{B}\|_2, \quad (11.20)$$

wenn über alle Matrizen $\mathbf{B} \in \mathbb{E}^{n \times n}$ vom Rang p minimiert wird. Dabei ist

$$\|\mathbf{A} - \mathbf{A}_p\|_2 = \sigma_{p+1}. \quad (11.21)$$

Diese Approximationseigenschaft kann insbesondere auch für die Datenkompression eingesetzt werden.

Index

- ABEL, NIELS HENRIK, 2-9
BANACH, STEFAN, 6-11
BUNJAKOVSKI, VIKTOR, 2-20
CAUCHY, AUGUSTIN LOUIS, 2-20
CHOLESKY, ANDRÉ LOUIS, 3-16
EUKLID, 2-16
FOURIER, JEAN-BAPTISTE-JOSEPH, 6-12
GAUSS, CARL FRIEDRICH, 1-1
GIVENS, J. WALLACE, 2-29
GRAM, JORGEN PEDERSEN, 6-8
HANKEL, HERMANN, 2-33
HERMITE, CHARLES, 2-12
HESSENBERG, GERHARD, 2-32
HILBERT, DAVID, 6-11
HOUSEHOLDER, ALSTON SCOTT, 2-30
JACOBI, CARL GUSTAV, 2-30
LEBESGUE, HENRI, 6-12
LEGENDRE, ADRIEN MARIE, 6-10
PARSEVAL DES CHÊSNES, MARC-ANTOINE
DE, 6-7
PYTHAGORAS, 2-16
SARRUS, PIERRE, 8-3
SCHMIDT, ERHARD, 6-8
SCHUR, ISSAI, 3-12
TOEPLITZ, OTTO, 2-33
TSCHEBYSCHJEFF, PAFNUTI LWOWITSCH,
6-10
- $C[a, b]$, 4-2, 6-2, 6-3
 $C^m[a, b]$, 4-2
 \mathbb{C}^n , 2-4
 $\mathbb{C}^{m \times n}$, 2-4
 \mathbb{E}^n , 2-4
 $\mathbb{E}^{m \times n}$, 2-4
 \mathbb{R}^n , 2-4
 $\mathbb{R}^{m \times n}$, 2-4
 $\mathcal{L}(X, Y)$, 6-19
 \mathcal{P}_m , 4-3
 $\sigma(F)$, 9-1
- 2-Norm, 2-19
einer Matrix, 10-16
- 2-Norm-Konditionszahl, 6-24
 m -Vektor, 2-1
 n -Tupel, 2-1
 n -dimensional, 4-13
- Abbildung
affine, 5-17
auf, 5-1
bijektive, 5-1
eindeutige, 5-1
eindeutige auf, 5-1
injektive, 5-1
inverse, 5-1
isometrische, 2-31, 6-18
längentreue, 2-31, 6-18
lineare, 5-1
surjektive, 5-1
winkeltreue, 2-31, 6-18
- Abbildungsmatrix, 5-3
Abelsche Gruppe, 2-9
Ableitungsoperator, 5-2
Absolutbetrag, 0-6
Abspalten, 9-5
Addition
von Matrizen, 2-4
von Vektoren, 4-1
Adjungierte, 2-12
affine Abbildung, 5-17
affiner Teilraum, 5-17
Ähnlichkeitstransformation, 5-20, 9-9
ähnlich, 5-20, 9-9
algebraische Vielfachheit, 9-6
Alphabet
griechisches, 0-8
äusseres Produkt, 2-23
Axiome, 4-2
- Banachraum, 6-11
Bandbreite
gesamte, 2-32
obere, 2-32
untere, 2-32

- Bandmatrix, 2-32
- Basis, 4-10
 - orthogonale, 6-5
 - orthonormale, 6-5
 - Standard-, 4-11
- Basiswechsel, 4-18
- Berechnung, 8-6
- Betrag (einer komplexen Zahl), 0-6
- Bidiagonalmatrix
 - obere, 2-31
 - untere, 2-31
- bijektiv, 5-1
- Bild, 2-25, 5-1
- Bildraum, 5-2
- bilinear, 2-18
- bilineare Form
 - symmetrische, 10-10
- BLAS, 2-34
- Block-LR-Zerlegung, 3-12
- Blockmatrix, 2-33, 2-34

- Cache, 2-34
- CBS Ungleichung, 6-4
- charakteristische Gleichung, 9-5
- charakteristisches Polynom, 9-5, 9-6
- Cholesky-Zerlegung, 3-16, 3-18
- Cosinussatz, 2-16

- dünn besetzte Matrix, 2-33, 3-6
- Definitionsraum, 5-2
- Deflation, 0-7, 9-5
- Determinante, 8-3, 8-6
 - Entwicklung nach Kolonne, 8-10
 - Entwicklung nach Zeile, 8-10
- Diagonale, 2-1
- diagonalisierbar, 9-11
- Diagonalmatrix, 2-3
- Diagramm
 - kommutatives, 5-6
- Differentialgleichungen
 - entkoppelte, 10-5
 - lineare, erster Ordnung, 10-1
 - lineare, höherer Ordnung, 10-4
 - lineare, zweiter Ordnung, 10-5
 - reelle mit komplexen Eigenwerten, 10-8
- Differentialoperator, 5-2
- Dimension, 4-13
- direkte Summe, 4-17
- direkte Summe orthogonaler Komplemente, 6-13
- Dreiecksgestalt, 1-3
- Dreiecksmatrix
 - obere, 2-3
 - untere, 2-3
- Dreiecksungleichung, 2-21, 6-1, 6-22
- dyadisches Produkt, 2-23

- Eigenbasis, 9-10
- Eigenraum, 9-1, 9-4
- Eigenvektor, 9-1
 - linker, 9-12
- Eigenwert, 9-1
- Eigenwertzerlegung, 9-11
- eindeutig auf, 5-1
- Einheitsmatrix, 2-3
- einschaliges Hyperboloid, 10-13
- elementare Zeilen-Operation, 1-6
- Eliminationsschema, 1-2
- Eliminationsschritt, 1-8, 1-16
- Ellipse, 10-10
- Ellipsoid, 10-13
- endlich-dimensional, 4-13
- Entwicklung
 - nach Kolonne, 8-10
 - nach Zeile, 8-10
- Erzeugendensystem, 4-7
- Euklidische Norm, 2-19
- Euklidischer Vektorraum, 6-2
- Eulerschen Formel, 0-6
- EV, 9-1
- Evaluationsabbildung, 5-2
- EW, 9-1

- Faktorisierung
 - QR-, 7-9
- Fehlergleichungen, 7-5
- Fläche zweiter Ordnung, 10-10
- Folge
 - quadrat-summierbare, 6-11
- Form
 - bilineare, 10-10
 - Hermitesche, 10-10
- Fourier-Koeffizienten, 6-12
- Fourier-Reihe, 6-12
 - verallgemeinerte, 6-12
- freie Variable, 1-9
- freier Parameter, 1-9
- Frobenius-Norm, 6-23

- fundamentale Unterräume, 6-14, 11-6
 Funktion
 differenzierbare, 4-2
 einer Matrix, 10-8
 quadrat-integrierbare, 6-12
 quadratische, 10-15
 stetige, 4-2
 Funktional
 lineares, 5-2
 Funktionen
 stetige, 6-2, 6-3
 Gauss-Algorithmus, 1-1
 allgemeiner Fall, 1-16
 regulärer Fall, 1-8
 Gauss-Elimination, 1-1
 Gauss-Elimination, 3-6, 3-18
 allgemeiner Fall, 1-16
 regulärer Fall, 1-8
 geometrische Vielfachheit, 9-2
 Gewichtsfunktion, 6-3
 Givens-Rotation, 2-29, 6-16
 Gleichung
 charakteristische, 9-5
 Gleichungssystem, 1-1
 äquivalentes, 1-7
 überbestimmtes, 7-5
 homogenes, 1-18
 inhomogenes, 1-18
 lineares, 2-12
 quadratisches, 1-2
 reguläres, 1-7
 singuläres, 1-7
 Grad
 exakter, 0-7
 Gram-Schmidt-Verfahren, 6-8
 klassisches, 6-9, 7-8, 7-12
 modifiziertes, 6-9, 7-12
 griechisches Alphabet, 0-8
 Gruppe, 2-9
 Abelsche, 2-9
 kommutative, 2-9
 symmetrische, 8-1
 Höchstgrad, 0-7
 Hankel-Matrix, 2-33
 harmonischer Oszillator, 10-5
 Hauptachsen-Transformation, 10-11
 Hauptdiagonale, 2-1
 Hauptuntermatrizen
 führende, 3-14
 Hauptvektoren, 9-20
 Hermitesche Form, 10-10
 Hessenberg-Matrix, 2-32
 Hilbert-Matrix, 10-18
 Hilbertraum, 6-11
 homogen, 1-18
 Householder-Matrix, 2-30
 Householder-Spiegelung, 2-30
 Hpd, 3-16
 Hyperbel, 10-10
 Hyperboloid
 einschaliges, 10-13
 zweischaliges, 10-13
 idempotent, 2-25
 Identität, 2-3
 imaginäre Achse, 0-5
 Imaginärteil, 0-5
 induzierte Matrixnorm, 6-20
 induzierte Operatornorm, 6-20
 inhomogen, 1-18
 injektiv, 5-1
 inneres Produkt, 2-17, 2-23
 Inverse, 2-26
 inverse Abbildung, 5-1
 invertierbar, 2-26
 isometrisch, 2-31, 6-18

 Jacobi-Rotation, 2-30, 6-16
 Jordan-Block, 9-19
 nicht-trivialer, 9-19
 Jordansche Normalform, 9-19

 Körper, 4-4, 4-5
 Kegelschnitt, 10-10
 Kellerzeile, 1-4
 Kern, 5-7
 Kleinste-Quadrate-Lösung, 7-5
 Koeffizienten
 eines Gleichungssystems, 1-2
 Koeffizientenmatrix, 1-2, 2-12
 Kofaktor, 8-9
 Kolonne, 1-3, 2-1
 Kolonnen-Pivotieren, 7-12
 Kolonnenmaximumstrategie, 3-8
 Kolonnenraum, 5-11
 Kolonnenvektor, 2-1
 kommutativ, 2-9
 kommutatives Diagramm, 5-6

- kommutieren, 2-8
 Komplement
 orthogonales, 6-13
 komplementäre Unterräume, 4-17
 komplexe Ebene, 0-5
 komplexe Zahl, 0-5
 Komponente, 2-2
 Kondition, 6-24
 Konditionszahl, 6-24, 10-16, 10-17, 11-6
 kongruent, 10-13
 Kongruenztransformation, 10-13
 konjugiert-komplex, 0-5
 Koordinaten, 4-15
 Koordinatenabbildung, 5-6
 Koordinatendarstellung, 4-16
 Koordinatentransformation, 4-18
 Koordinatenvektor, 4-15
 Kronecker-Symbol, 6-5
 Kurve 2. Ordnung, 10-10
- Länge, 2-19
 längentreu, 2-31, 6-18
 Lösung
 allgemeine, 1-14
 im Sinne der kleinsten Quadrate, 7-5
 nichttriviale, 1-18
 triviale, 1-18
 Lösungsvektor, 2-12
 LAPACK, 2-34
 LDR-Zerlegung, 3-6
 LDU-Zerlegung, 3-6
 Legendre-Polynome, 6-10
 linear abhängig, 4-8, 4-10
 linear unabhängig, 4-8
 unendlich viele Vektoren, 4-10
 lineare Abbildung, 2-25, 5-1
 beschränkte, 6-19
 lineare Hülle, 4-7
 linearer Operator, 5-2
 beschränkter, 6-19
 linearer Raum, 4-1
 mit Skalarprodukt, 6-2
 normierter, 6-1
 vollständiger, 6-11
 lineares Funktional, 5-2
 Linearfaktor, 0-7
 Linearkombination, 2-10, 4-7
 linker Eigenvektor, 9-12
 LR-Zerlegung, 3-2, 3-6, 3-10, 3-18
- regulärer Fall, 3-5
 symmetrische, 10-14
 Updating der, 3-15
 LU-Zerlegung, 3-2
- Matrix, 1-2, 2-1
 2-Norm einer, 10-16
 adjungierte, 2-12
 Bandbreite einer, 2-32
 bidiagonale, 2-31
 dünn besetzte, 2-33, 3-6
 diagonale, 2-3
 Diagonale einer, 2-1
 diagonalisierbare, 9-11
 Dreiecks-, 2-3
 Einheits-, 2-3
 Elemente einer, 2-1
 Grösse einer, 2-1
 Hankel-, 2-33
 Hauptdiagonale einer, 2-1
 Hermitesch positiv definite, 3-16
 Hermitesch transponierte, 2-12
 Hermitesche, 2-13
 Hessenberg, 2-32
 Inverse einer, 2-26
 invertierbar, 2-26
 Koeffizienten-, 2-12
 Kolonne einer, 2-1
 konjugiert transponierte, 2-12
 konjugiert-komplexe, 2-12
 Multiplikation mit Skalar, 2-4
 Multiplikation mit Zahl, 2-4
 Null-, 2-3
 orthogonale, 2-29
 quadratische, 2-2
 reguläre, 2-27
 schiefsymmetrische, 2-14
 Spalte einer, 2-1
 Spektralnorm einer, 10-16
 symmetrisch positiv definite, 3-16
 symmetrische, 2-13
 Toeplitz-, 2-33
 transponierte, 2-12
 tridiagonale, 2-32
 unitäre, 2-29
 Vielfaches einer, 2-4
 Zeile einer, 2-1
 Matrix-Vektor-Produkt, 2-10
 Matrixnorm, 6-23

- induzierte, 6-20
- kompatible, 6-23
- untergeordnete, 6-20
- Matrixordnung, 2-2
- Matrizen
 - ähnliche, 5-20, 9-9
 - Addition von, 2-4
 - kommutierende, 2-8
 - kongruente, 10-13
 - Multiplikation von, 2-5
 - Produkt von, 2-5
 - Summe von, 2-4
- Maximumnorm, 6-2
- Monome, 4-8, 4-12
- Multiplikation
 - von Matrizen, 2-5
 - skalare, 4-1
- Multiplikationsoperator, 5-3
- nichttriviale Lösung, 1-18
- Norm, 2-19, 2-23, 6-1, 6-3
 - Euklidische, 2-19
- Normalform
 - Jordansche, 9-19
- Normalgleichungen, 7-6
- normierter linearer Raum, 6-1
- normierter Vektorraum, 6-1
- Nullmatrix, 2-3, 2-9
- Nullraum, 5-11
- Nullstelle, 0-7
- Nullteiler, 2-9
- Nullvektor, 2-3, 4-1
- Operator
 - linearer, 5-2
- Operatornorm
 - untergeordnete, 6-20
- Ordnung
 - einer Matrix, 2-2
- orthogonal, 6-4
- Orthogonalpolynomen, 6-10
- orthogonale Basis, 6-5
- orthogonale Komplement, 6-13
- orthogonale Matrix, 2-29
- orthogonale Vektoren, 2-22
- Orthogonalprojektion, 7-1
- orthonormale Basis, 6-5
- Oszillator
 - harmonischer, 10-5
- Parabel, 10-10
- Parallelrechner, 2-34
- Parameter
 - freier, 1-9
- Parsevalsche Formel, 6-8
- partielles Pivotieren, 3-8
- Permutation, 8-1
- Permutationsmatrix, 2-30
- Pivot, 1-4
- Pivotelement, 1-4, 1-8, 1-16
- Pivotgleichung, 1-4
- Pivotieren
 - Kolonnen-, 7-12
 - partielles, 3-8
 - vollständiges, 3-8
- Pivotkolonne, 1-4
- Pivotstrategie, 1-9, 3-8
- Pivotvariable, 1-13
- Pivotzeile, 1-4, 1-8, 1-16
- Polarform, 0-6
- Polynom, 0-7
 - charakteristisches, 9-5, 9-6
- Polynome
 - Legendre-, 6-10
 - orthogonale, 6-10
 - Tschebyscheff-, 6-10
 - Vektorraum aller, 4-8
 - Vektorraum der, 4-3
 - vom Grad m , 4-3, 4-8
- positiv definit, 3-16
- positiv semidefinit, 3-16
- Produkt
 - äusseres, 2-23
 - dyadisches, 2-23
 - inneres, 2-17
 - Matrix-Vektor-, 2-10
 - Skalar-, 2-17
- Projektion, 7-1, 9-3
 - orthogonale, 7-1
 - schiefe, 7-1
- Projektionen
 - auf Koordinatenachsen, 6-7
- Projektor, 7-1
 - orthogonaler, 7-1
- Pseudoinverse, 7-6
- Pythagoras
 - Satz von, 2-16
- QR-Faktorisierung, 7-9

- QR-Zerlegung, 7-10
- quadratische Funktion, 10-15
- Quadratwurzel
 - einer Diagonalmatrix, 3-16
- Rückwärtseinsetzen, 1-3, 1-8, 1-16, 3-3
- Rang, 1-17
 - einer Matrix, 1-15
 - eines Gleichungssystems, 1-15, 1-17
- Raum
 - affiner, 5-17
 - normierter linearer, 6-1
- Realteil, 0-5
- Rechenaufwand
 - Gauss-Elimination, 3-7
 - LR-Zerlegung, 3-7
- rechte Seite, 1-2, 2-12
- rechte Seiten
 - mehrere, 2-12
- Rechtsdreiecksmatrix, 2-3
- reelle Achse, 0-5
- Regel von Sarrus, 8-3
- Residuenvektor, 7-5
- Residuum, 7-5
- Restgleichungssystem, 1-4
 - 0-tes, 1-8, 1-16
- Ring, 2-9
 - der $n \times n$ Matrizen, 2-9
 - der ganzen Zahlen, 4-3
 - der Polynome, 4-3
 - mit Eins, 2-9
 - nicht-kommutativer, 2-9
- Rotation, 2-29, 6-16, 9-3
- Sarrus
 - Regel von, 8-3
- schiefsymmetrisch, 2-14
- Schmidtsches Orthogonalisierungsverfahren, 6-9
- Schur-Komplement, 3-12
- Schwarz
 - Hermann Amandus, 2-20
- Schwarzsche Ungleichung, 6-4
- Selbstabbildung, 5-2
- senkrecht, 2-22, 6-4
- sesquilinear, 2-18
- Signatur, 10-12
- Signum, 8-3
- Singulärwertzerlegung, 11-3
- Singulärvektor
 - linker, 11-3
 - rechter, 11-3
- Singulärwert, 11-3
- Singulärwertzerlegung, 11-1
- Skalar, 2-4, 4-1
- Skalarenkörper, 4-4
- Skalarprodukt, 2-17, 2-23
- Spalte, 1-3, 2-1
- Spaltenvektor, 2-1
- spd, 3-16
- Spektralnrm, 6-20, 10-16, 11-6
- Spektralzerlegung, 9-11, 9-12
- Spektrum, 9-1
- Spur, 9-6
- Standardbasis, 4-11
- Subtraktion, 4-4
- Summe
 - direkte, 4-17
 - orthogonaler Komplemente, 6-13
- surjektiv, 5-1
- SVD, 11-3
- symmetrische bilineare Form, 10-10
- symmetrische Gruppe, 8-1
- Teilmengen
 - orthogonale, 6-4
- Teilraum
 - affiner, 5-17
- Toeplitz-Matrix, 2-33
- Trägheit, 10-12
- Transformation
 - auf Hauptachsen, 10-11
 - Kongruenz-, 10-13
- Transformationsmatrix, 4-18
- Transponierte, 2-12
- Transposition, 8-1
- Tridiagonalmatrix, 2-32
- triviale Lösung, 1-18
- Tschebyscheff-Polynome, 6-10
- Unbekannte, 1-2
- Ungleichung
 - Cauchy-Schwarz-, 2-20
 - Schwarzsche, 2-20, 6-4
- unitär, 2-29
- unitäre Matrix, 2-29
- unitärer Vektorraum, 6-2
- untergeordnete Operatornorm, 6-20
- Unterräume
 - fundamentale, 6-14, 11-6

- orthogonale, 6-4
- Unterraum, 4-6
 - aufgespannter, 4-7
 - erzeugter, 4-7
- Urbild, 5-1, 5-7, 5-8
- Variable
 - freie, 1-9
- Vektor, 1-2, 4-1
- Vektoren
 - orthogonale, 2-22, 6-4
- Vektorraum, 4-1
 - n -dimensionaler, 4-13
 - Dimension eines, 4-13
 - endlich-dimensionaler, 4-13
 - Euklidischer, 6-2
 - mit Skalarprodukt, 6-2
 - normierter, 6-1
 - orthogonaler, 6-2
 - topologischer, 6-11
 - unitärer, 6-2
 - vollständiger, 6-11
- Vektorrechner, 2-34
- Vektorregister, 2-34
- Verträglichkeitsbedingung, 1-15, 3-10
- Verträglichkeitstest, 1-16
- Vielfachheit
 - algebraische, 9-6, 9-14, 9-15
 - geometrische, 9-2, 9-14, 9-15
- vollständig, 6-11
- vollständige Pivotieren, 3-8
- Vorwärtseinsetzen, 3-3
- Wertebereich, 5-1, 5-11
- Winkel zwischen Vektoren, 2-22, 6-4
- winkeltreu, 2-31, 6-18
- Zeile, 1-3, 2-1
- Zeilen-Operation
 - elementare, 1-6
- Zeilenstufenform, 1-13, 1-15
 - Reduktion auf, 3-9
- Zeilenvektor, 2-1
- Zeilenvertauschung, 1-8, 1-16
- Zerlegung
 - Cholesky-, 3-16, 3-18
 - LDR-, 3-6
 - LDU-, 3-6
 - LR-, 3-2
 - LU-, 3-2
- QR-, 7-10
- SVD, 11-3
- zweischaliges Hyperboloid, 10-13